

Deep Sequencing Identifies Noncanonical Editing of Ebola and Marburg Virus RNAs in Infected Cells

Reed S. Shabman,^{a,b} Omar J. Jabado,^c Chad E. Mire,^d Timothy B. Stockwell,^b Megan Edwards,^a Milind Mahajan,^c Thomas W. Geisbert,^d Christopher F. Basler^a

Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, New York, USA^a; Virology Group, J. Craig Venter Institute, Rockville, Maryland, USA^b; Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, New York, USA^c; Department of Microbiology and Immunology, Galveston National Laboratory, University of Texas Medical Branch, Galveston, Texas, USA^d

R.S.S. and O.J.J. contributed equally to this work.

ABSTRACT Deep sequencing of RNAs produced by Zaire ebolavirus (EBOV) or the Angola strain of Marburgvirus (MARV-Ang) identified novel viral and cellular mechanisms that diversify the coding and noncoding sequences of viral mRNAs and genomic RNAs. We identified previously undescribed sites within the EBOV and MARV-Ang mRNAs where apparent cotranscriptional editing has resulted in the addition of non-template-encoded residues within the EBOV glycoprotein (GP) mRNA, the MARV-Ang nucleoprotein (NP) mRNA, and the MARV-Ang polymerase (L) mRNA, such that novel viral translation products could be produced. Further, we found that the well-characterized EBOV GP mRNA editing site is modified at a high frequency during viral genome RNA replication. Additionally, editing hot spots representing sites of apparent adenosine deaminase activity were found in the MARV-Ang NP 3'-untranslated region. These studies identify novel filovirus-host interactions and reveal production of a greater diversity of filoviral gene products than was previously appreciated.

IMPORTANCE This study identifies novel mechanisms that alter the protein coding capacities of Ebola and Marburg virus mRNAs. Therefore, filovirus gene expression is more complex and diverse than previously recognized. These observations suggest new directions in understanding the regulation of filovirus gene expression.

Received 25 September 2014 Accepted 2 October 2014 Published 4 November 2014

Citation 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(6):e02011-14. doi:10.1128/mBio.02011-14.

Editor Anne Moscona, Weill Cornell Medical College

Copyright © 2014 Shabman et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution-Noncommercial-ShareAlike 3.0 Unported license](https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Address correspondence to Christopher Basler, chris.basler@mssm.edu.

This article is a direct contribution from a Fellow of the American Academy of Microbiology.

The recent outbreak in West Africa of Zaire ebolavirus (EBOV) is of unprecedented scope in terms of the number of cases and geographic breadth (1). This event, coupled with the increased frequency of outbreaks caused by the *Filoviridae* family, which includes the ebolaviruses and marburgviruses (MARV), highlights the need for a complete understanding of filovirus replication (2–4).

Previous studies have provided a detailed description of EBOV and MARV genome transcription and replication strategies (5–7). The filoviruses possess nonsegmented negative-sense (NNS) viral RNA (vRNA) genomes of approximately 19 kb in length. Filoviral genome replication results in the production of full-length, positive-sense antigenomic RNA and negative-sense genomic RNA (7, 8). In contrast, filovirus transcription likely proceeds along the viral genomic RNA template via a stop-start mechanism that results in the production of individual 5'-7-methylguanosine (m⁷G)-capped, 3'-polyadenylated mRNAs (9–11). This is analogous to the transcription of several other NNS RNA viruses, such as vesicular stomatitis virus (VSV) (12). Transcripts are produced from each of the nucleoprotein (NP), VP35, VP40, viral glycoprotein (GP), VP30, VP24, and large (L) protein genes. For EBOV, there are seven transcription start signals of 12 nucleotides (nt) in

length that differ at only 1 nt; transcriptional stop signals are present for each gene and are also conserved, and each gene possesses, on the genomic vRNA, a poly(U) stretch that serves as a signal for polyadenylation (6). Notably, the VP24/L junction contains two stop signals (13), but most transcription stops at the first stop signal (14). For MARV, the transcriptional start signals are also 12 nt in length with minor variations depending on the specific gene, and the consensus stop signal is also conserved. Between NNS RNA virus start and stop sequences are intergenic regions that differ in length.

Whereas most filoviral genes have been reported to yield a single mRNA species encoding a single major translation product, the EBOV GP gene is an exception. The major mRNA species produced by the GP gene encodes soluble glycoprotein (sGP). However, two additional proteins, the full-length GP (a structural protein that serves as the viral attachment and fusion protein) and ssGP (a protein of unknown function) are produced via an RNA editing process. This RNA editing is carried out by the viral polymerase, such that one or more non-template-encoded A nucleotides are inserted within the EBOV GP open reading frame (ORF) at a stretch of seven U nucleotides spanning genome positions 6918 to 6925 (15, 16). Therefore, RNA editing by the viral poly-

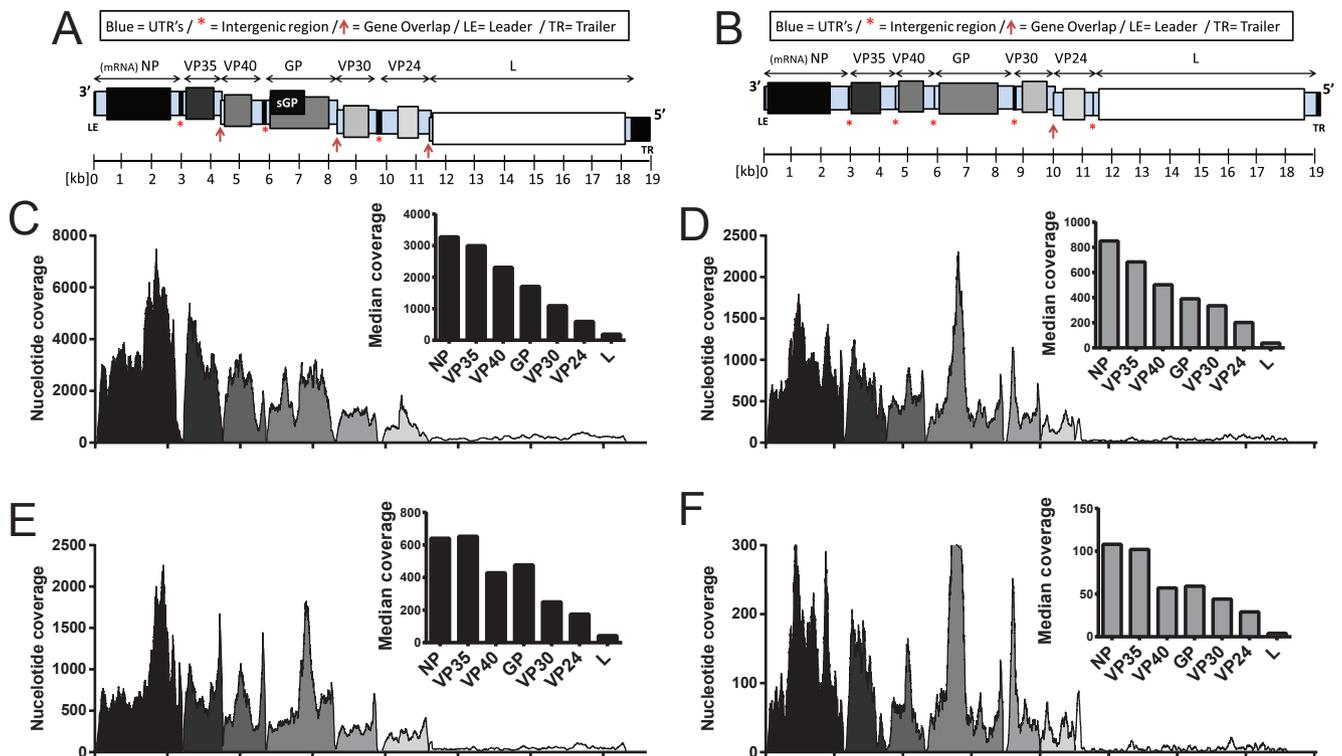


FIG 1 Transcriptional profiles of Zaire EBOV and MARV-Ang in multiple cell lines. Two representative filoviruses, EBOV and MARV-Ang, were used to infect both Vero cells and Thp1 cells at a multiplicity of infection of 3. At 12 hpi, mRNA was isolated from each cell line and Illumina HiSeq libraries were generated; nucleotide coverage at each genomic position in EBOV and MARV-Ang was determined. (A) EBOV genomic organization. (B) MARV-Ang genomic organization. (C) EBOV-infected Vero cells at 12 hpi. (D) MARV-Ang-infected Vero cells at 12 hpi. (E) EBOV-infected Thp1 cells at 12 hpi. (F) MARV-Ang-infected Thp1 cells at 12 hpi. For panels C to F, the x axis denotes the genomic position and the y axis denotes nucleotide coverage. Inset graphs for panels represent median nucleotide coverage for each of the seven filovirus transcriptional units.

merase increases the number of proteins produced by a limited number of viral genes (17, 18). In contrast to EBOV, there are no reports of RNA editing by the MARV polymerase, and MARV is only known to produce a membrane-bound GP from an unedited mRNA.

Prior characterizations of the products of filovirus RNA synthesis have largely relied upon sequencing of PCR products or Northern blotting (e.g., references 5, 11, and 19). Here, the single-base resolution of next-generation sequencing was applied to the products of Zaire EBOV and MARV-Ang, a strain that caused an outbreak with a nearly 90% case fatality rate. This approach revealed individual mRNA expression levels, confirmed transcriptional stop sites and known editing sites, and identified novel editing events of various frequencies at multiple homopolymer regions (regions of 6 or 7 identical nucleotides) found in mRNAs produced in both EBOV- and MARV-Ang-infected cells. Furthermore, variants within substantial portions of the MARV NP mRNA population considered indicative of cellular adenosine deaminase activity were detected in MARV-Ang-infected cells. The comprehensive investigation of filoviral mRNA synthesis at a single-base resolution therefore has identified novel means by which viral and host factors increase the diversity of gene products produced during filovirus infection.

RESULTS

Filoviral genes accumulate in a 3' → 5' gradient. We employed next-generation sequencing technology to map mRNA transcript

levels in both EBOV- and MARV-Ang-infected cells. Both Vero and Thp1 cell lines were infected at a high multiplicity of infection (MOI) of 3. At 6, 12, and 24 h postinfection (hpi), sequencing libraries derived from mRNA were constructed and sequenced on the Illumina HiSeq 2500 apparatus. The 100-bp reads were mapped to either EBOV or MARV reference genomes by using the TopHat RNA-Seq (RNA sequencing) suite. The total and filovirus-specific read counts are detailed in Table S1 in the supplemental material, and these data indicate that both MARV-Ang and EBOV grew in both cell lines and represented between ~1.5 and 7% of the total sequencing reads by 24 hpi.

Figure 1A and B depict EBOV and MARV-Ang genome organization, respectively. Key genomic regions are indicated, including the 3' leader and 5' trailer region, each of the seven mRNA transcripts, the conserved transcriptional start and stop sites, intergenic regions, and overlapping transcriptional start and stop sequences (6). The read depth at each genomic nucleotide position was determined following sequencing at 12 hpi of mRNAs from EBOV- and MARV-Ang-infected Vero cells (Fig. 1C and D) and Thp1 cells (Fig. 1E and F). Sequencing coverage is a proxy for the mRNA transcript abundance, and the insets in Fig. 1C to F display the median coverages for each transcriptional unit. At 12 hpi, the average nucleotide coverage decreased in a 3'-to-5' direction across the genomic RNA template, with the highest coverage corresponding to the NP mRNA and the lowest corresponding to the L mRNA (Fig. 1C and D). For both viruses, the total

number of sequencing reads and nucleotide coverage was higher in Vero cells than in Thp1 cells, suggesting more efficient transcription in Vero cells, perhaps due to the lack of the beta-1 interferon (IFN- β 1) and IFN- α genes in this cell line (20, 21). Median coverage at each time point is displayed in Fig. S1 in the supplemental material and illustrates a 3'-to-5' gradient at 6 and 12 h that diminished by 24 hpi. A complete list of nucleotide coverage at each time point is provided in Table S2 in the supplemental material.

Quasispecies analysis suggests RNA editing by cellular enzymes. An analysis was performed on the sequencing data to identify sites in viral mRNAs which may be prone to errors during viral transcription. For EBOV mRNAs, 8 sites were identified where substantial percentages of reads contained substitutions (see Table S3 in the supplemental material). Interestingly, C-to-U changes (indicated as T in the table) and G-to-A changes predominated, suggesting possible cytosine deaminase activity. Four sites were within the GP mRNA, one occurring in the mucin domain and three in the GP2 region (22). The same substitutions were present and found at similar frequencies at 12 and 24 hpi and in both Vero and Thp1 cells. Nonsynonymous substitutions at single sites in NP and VP35 mRNAs were observed at an approximately 20% frequency.

Interestingly, a different pattern of substitutions was observed for the MARV mRNAs (see Table S4 in the supplemental material). Most notably, in the 12-hpi MARV-Ang-infected Vero cells, 6 U-to-C substitutions and one A-to-G substitution were present in a substantial fraction of the reads corresponding to the 3'-untranslated region (UTR) of the NP mRNA. The same 7 substitutions were present in the 24-h samples from Vero and Thp1 cells. At these later time points, additional U-to-C substitutions were also present. The clustered U-to-C substitutions suggest adenosine deaminase (ADAR) activity on As within the viral negative-sense genomic RNA. The A-to-G change could represent ADAR editing on either the positive-sense mRNA or antigenomic viral RNA. The editing occurs in a selective manner, since it was restricted to a specific region of the genome and because the frequencies of edited mRNAs at 24 hpi ranged from 8 to 24%. Consistent with ADAR editing, the edited adenosines in the NP 3'-UTR have 5' neighbors that are either U or A (data not shown).

EBOV GP editing is detected during both transcription and replication. The single-base resolution and available depth of the sequencing data were next used to determine insertion frequencies (RNA editing) at the canonical GP site that lies at positions 6918 to 6924 of the EBOV genome (Fig. 2A). At 24 hpi, the average nucleotide coverage at each A residue in this region (positions 6918 to 6924) was 13,870 in Vero cells and 11,273 in Thp1 cells (see coverage statistics in Table S2 in the supplemental material). To calculate editing frequencies, we identified reads that perfectly matched the reference genome at 10 nt to the left and 10 nt to the right of the homopolymer region. Of these reads, we determined the percentage that had no additional A residues and those that contained extra (i.e., non-template-encoded) A residues. At 24 hpi in EBOV-infected Vero cells, 71.65% of sequencing reads had 7 As, 25.90% of reads had 8 As, and 1.26% of reads had 9 As (Fig. 2B; see also Table S5 in the supplemental material). An analysis of this region in Thp1 mRNAs indicated 67.92% of reads had the wild-type 7-A sequence, 28.04% had 8 As, 1.91% had 9 As, and 1.73% had 10 As (Fig. 2C; see also Table S5). These data confirmed a similar GP editing frequency in the human macrophages and

Vero cells and were comparable to those previously quantified by other methods (17, 18, 23–25). The above insertion frequencies were derived from libraries generated by chemically shearing mRNA. This frequency was confirmed by a second method, where an ~200-bp region flanking the GP editing site was amplified by reverse transcription-PCR (RT-PCR) from the Thp1 cell RNA, and the subsequent “amplicon” was analyzed by deep sequencing (Fig. 2D).

Cell culture-passaged EBOV can acquire an additional U within the negative-sense viral genomic RNA at the GP editing site (positions 6918 to 6924). For such viruses, GP becomes the main translation product, as it is translated from the nonedited mRNA, and sGP is translated from edited mRNAs (18, 26). These observations indicate that some editing occurs at the level of EBOV genome replication. To quantify the frequency of editing at the homopolymer region between residues 6918 and 6924 in vRNA, RT-PCR was performed by using a method that specifically amplified negative-sense EBOV genomic RNA (vRNA) (7, 19). This vRNA-specific amplicon was generated and analyzed by deep sequencing. Insertion frequencies within the vRNA were assessed at 1 hpi (the input viral genome) and 24 hpi (measuring both the input and replicated virus genome). The data indicated that the majority of the EBOV input genome is 7 Us (97.55%), but by 24 hpi 82.87% of total vRNA contained 7 Us and 15.50% contained 8 Us (Fig. 2E and F). The 1-hpi vRNA frequency agrees with the frequency obtained from deep sequencing of the stock virus (95.3% of sequencing reads had a 7-U sequence [data not shown]) and suggests that the majority of vRNA editing measured at 24 hpi is a product of virus replication.

To control for possible insertions derived from reverse transcription, PCR, or the sequencing method, we generated an amplicon to the same region from plasmid DNA. Deep sequencing of this product yielded an insertion rate of 0.24%. We performed a similar analysis with mRNA derived from an expression plasmid (transcribed by the host RNA polymerase II [Pol II]), which was expected to have a higher fidelity than the EBOV polymerase due to proofreading and exonuclease activities (27). Deep sequencing of an RT-PCR amplicon from the plasmid-derived mRNA yielded an insertion frequency of 0.80%. The insertion rates of these two controls indicated that the additional A residues from the EBOV RNAs are present in the viral RNA and are not artifacts of the sequencing methods (Fig. 2G and H; see also Table S5 in the supplemental material).

Novel editing events in EBOV mRNAs. Within the EBOV genome, there are 25 homopolymer sites of 6 or 7 identical nucleotides. We examined insertion frequencies at each of these sites for our mRNA-Seq data (data not shown). Insertions at select homopolymer regions would lead to a frameshift during protein translation, resulting in truncated protein products with novel C termini of various lengths (Fig. 3A; see also Fig. S2 in the supplemental material). Of these, a site with a high insertion frequency was located in a region corresponding to a stretch of 6 Us (nt 6378 to 6383) on the EBOV vRNA within the EBOV GP gene. This sequence is conserved among all 20 Zaire EBOV full-length annotated genomes outlined in a recent study (28) and in isolates from the 2014 West Africa outbreak from Guinea and Sierra Leone (29) (data not shown), but it appears to be specific for the Zaire species, since there is no corresponding 6-U region in the Sudan, Tai Forest, Bundibugyo, or Reston ebolaviruses (28). The number of A residues in GP mRNAs was enumerated within this homopolymer

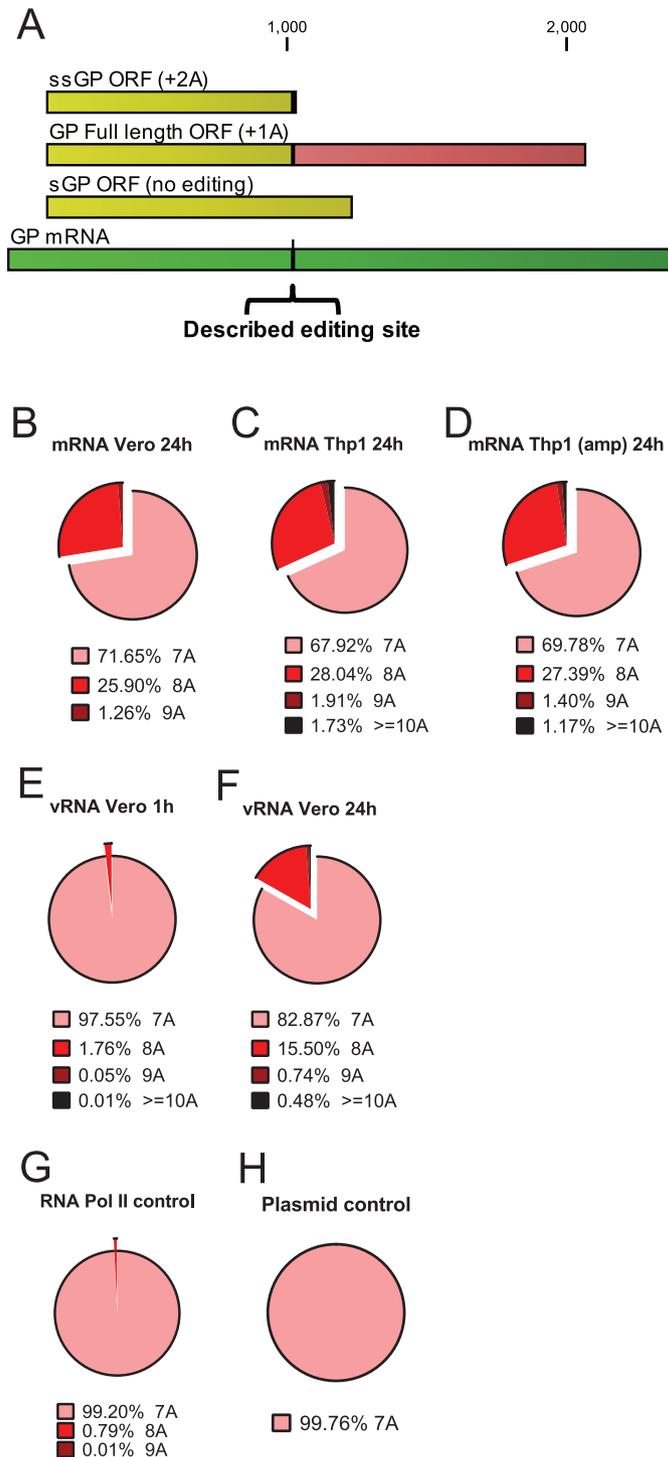


FIG 2 Insertion frequencies at the described EBOV GP homopolymer region. (A) Depiction of the GP mRNA (dark green) and predicted translation products, depending on which editing events occur. The previously described editing site is identified by the vertical black line and inverted bracket. (B to H) The pie charts depict the numbers of adenosine residues at the canonical EBOV GP editing sites from either EBOV RNA, control RNA, or control DNA. Illumina deep sequencing reads with 10 nt of identity (underlined in the sequence) to each side of the poly(A) stretch at positions 6907 to 6934 (CTGGGAAACTAAAAAA CCTCACTAGA) were identified. The number of reads with 7, 8, or 9 A residues were then enumerated. (B) Insertion frequency in mRNA from Vero cells at 24 hpi, determined from libraries constructed from chemically sheared RNA. (C) mRNA from Thp1 cells at 24 hpi, determined from libraries constructed from chemically sheared RNA. (D) mRNA from Thp1 cells at 24 hpi determined from an RT-PCR amplicon encompassing the region of interest. (E) Homopolymer insertion frequency of genomic RNA (vRNA) at 1 hpi, determined from an amplicon encompassing the region of interest. (F) The same experiment as shown in the middle left pie chart, but vRNA was assessed at 24 hpi. (G) Insertion frequency from mRNA derived from an RNA polymerase II-driven plasmid expressing the sequence of the 7-A EBOV sGP mRNA. (H) Insertion frequency from a DNA plasmid harboring the EBOV sGP ORF.

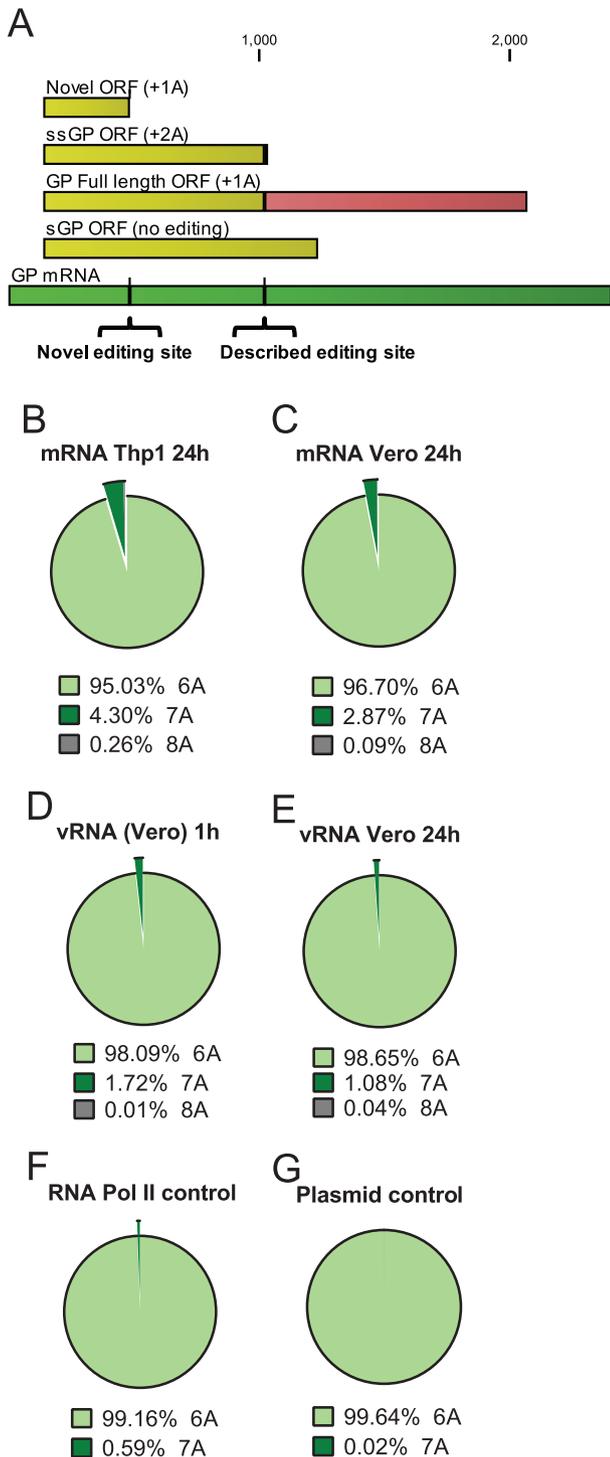


FIG 3 Insertion frequencies in the EBOV GP homopolymer region at nt 6378 to 6383. (A) Diagram similar to that in Fig. 2, except the novel GP editing site and the predicted translation product arising from the novel edited mRNA are depicted. (B to G) Pie charts depicting the number of adenosine residues at a novel location with the EBOV GP ORF from either EBOV RNA, control RNA, or control DNA. The numbers of A residues were enumerated within the homopolymer region from Illumina sequencing reads, with 10 nt of matching sequence (underlined in the sequence) directly flanking each side from position 6367 to 6393 (TCTTGAAATCAAAAAACCTGACGGGA). (B) Homopolymer insertion frequency in mRNA of Thp1 cells at 24 hpi, determined from libraries constructed from chemically sheared RNA. (C) mRNA from Vero cells at 24 hpi, determined from libraries constructed from chemically sheared RNA. (Continued)

region. The frequency of a single A insertion ranged between 4.30% and 2.87% in infected Thp1 and Vero cells at 24 hpi (Fig. 3B and C). Nontemplated insertions were present, but at a lower frequency in vRNA-specific amplicons flanking the homopolymer region at 1 and 24 hpi (Fig. 3D and E). Control reactions on plasmid DNA and RNA Pol II-derived mRNA indicated that the observation was specific to the viral polymerase, such that 99.64% and 99.16% of all reads were wild-type sequence, respectively (Fig. 3F and G; see also Table S5 in the supplemental material). These data therefore identify a new site within the EBOV GP gene that undergoes substantial editing, albeit at a lower frequency than the described GP site.

Two homopolymer regions corresponding to stretches of 6 Us in the negative-sense EBOV vRNA within the NP gene (nt 850 to 855 and 1288 to 1293), a 6-U region in the EBOV VP30 gene (nt 8767 to 8772), and two 6-A regions in the L gene (nt 12146 to 12151 and 17521 to 17526) were also examined, since these products could translate truncated proteins in infected cells (see Table S5 and the theoretical translation products displayed in Fig. S2 in the supplemental material). For NP, the 11,588 reads in Thp1 cells and 10,728 reads from Vero cells corresponding to positions 850 to 855 had insertion frequencies of only 1.27% and 0.48%, respectively. For the second NP site (nt 1288 to 1293), the 2,598 reads in Thp1 cells and 6,670 reads from Vero cells showed only 1.30% and 0.88% editing, respectively (see Table S5). The site at nt 8767 to 8772 in the VP30 gene also exhibited low levels of editing of 2.53% in Thp1 cells and 0.81% in Vero cells (see Table S5). Finally, the L regions had more substantial insertion frequencies of between 2.30% and 7.30% (see Table S5). Taken together, these data demonstrate that while one site in L displays an insertion frequency of 7.30%, not all homopolymer runs outside the GP gene are sites of transcriptional editing.

Detection of nucleotide insertions at homopolymer sites within MARV-Ang mRNAs. There are no previous reports describing MARV mRNA editing. Therefore, we evaluated the MARV-Ang samples for regions of the genome prone to RNA editing. Of 20 homopolymer regions (6 or more identical nucleotides), our analysis identified two regions where insertions of additional As occurred at a high frequency. In each region, the number of sequencing reads with extra A residues within the homopolymer region were enumerated as described above for EBOV. One site in the NP mRNA (a 6-U repeat on the negative-sense vRNA between positions 816 and 821) had an extra-A insertion frequency of 11.73% (Vero cells) and 7.9% (Thp1 cells) at 24 hpi (Fig. 4A to C). A second region in the L mRNA (a 6-U repeat on the negative-sense vRNA at positions 17810 to 17815) was also identified with extra-A insertion frequencies of 16.54% in Vero cells and 19.54% in Thp1 cells at 24 hpi (Fig. 5A to C; see also Table S6 in the supplemental material). In both regions, editing events were specific to the mRNA, as insertion frequencies derived from amplicons specific for vRNA or plasmid DNA were at or below 1.00% (Fig. 4D to F and 5D to F; see also Table S6).

Figure Legend Continued

RNA. (D) Homopolymer insertion frequency in genomic RNA (vRNA) at 1 hpi, determined from an amplicon encompassing the region of interest. (E) The same experiment as depicted in panel D, except that vRNA was assessed at 24 hpi. (F) Insertion frequency from mRNA derived from an RNA polymerase II-driven plasmid expressing EBOV GP mRNA. (G) Insertion frequency from a DNA plasmid harboring the EBOV GP ORF.

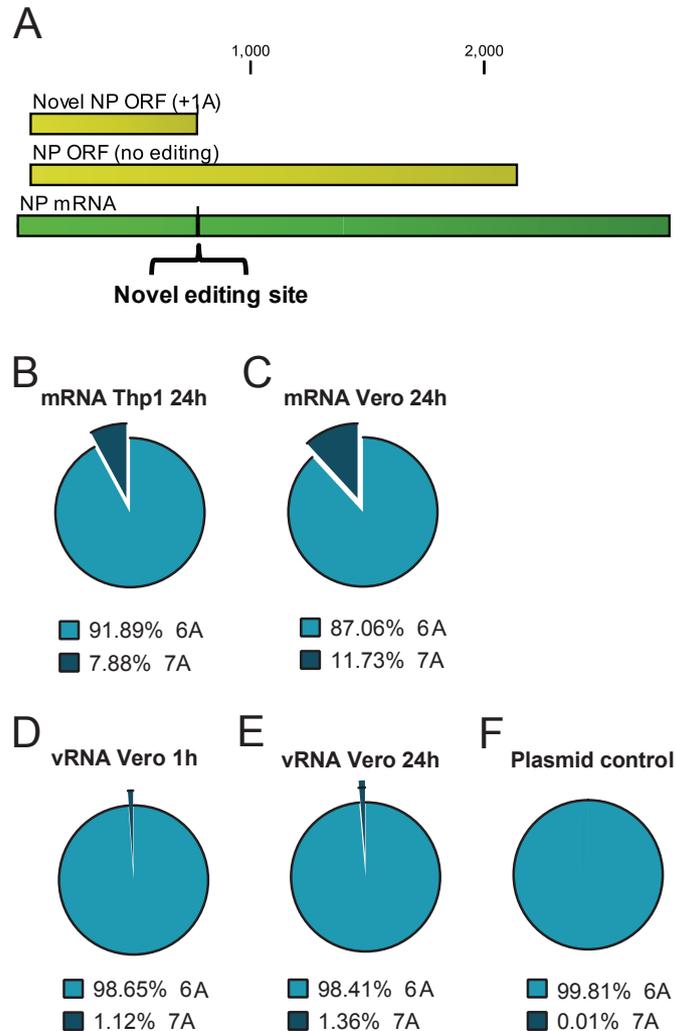


FIG 4 Insertion frequencies in the MARV-Ang NP homopolymer region at nt 816 to 821. (A) Depiction of the NP mRNA (dark green) and different translation products produced, depending on editing. (B to F) Pie charts depicting the numbers of adenosine residues inserted at the novel site within the MARV-Ang NP ORF. The numbers of A residues were enumerated within the homopolymer region from sequencing reads with 10 nt of matching sequence (underlined in the sequence) directly flanking each side from positions 805 to 831 (GTTCATCTTGCAAAAACTGATTCAGG). (B) Homopolymer insertion frequency in mRNA of Thp1 cells at 24 hpi, determined from libraries constructed from chemically sheared RNA. (C) mRNA from Vero cells at 24 hpi, determined from libraries constructed from chemically sheared RNA. (D) Homopolymer insertion frequency in genomic RNA (vRNA) at 1 hpi, determined from an amplicon encompassing the region of interest. (E) The same experiment as in panel D, except that vRNA was assessed at 24 hpi. (F) Insertion frequency from a DNA plasmid containing the MARV-Ang NP ORF.

We next confirmed that the NP and L regions in the MARV-Ang mRNA were prone to insertions by using Sanger sequencing for RT-PCR amplicons that encompassed 200 nt flanking each homopolymer region. An amplicon flanking the described EBOV GP site was used as a positive control for this analysis. Each PCR product was cloned, and the number of A residues present in individual clones was determined. The EBOV GP region yielded an insertion frequency of 15.7% (3/19 clones had a single A insertion). For both MARV-Ang NP and L regions, an insertion frequency of 2 to 3% was observed (2/86 and 2/64, respectively). While these percentages differ from the Illumina analysis, they clearly indicate that MARV-Ang transcription generates nucleotide insertions at homopolymer regions within the NP and L ORFs.

Editing in MARV-Ang-infected primate tissue. Finally, we asked if noncanonical transcriptional editing can be detected in

MARV-Ang-infected animals. Macaques were infected with MARV-Ang, and at day 9 postinfection (9 dpi), mRNA was isolated from the adrenal gland, the liver, and the axillary and mesenteric lymph nodes. Amplicons that spanned the NP and L homopolymer regions were sequenced to determine the percentage of insertions at each region. Both the NP and L insertion frequencies from the adrenal gland are displayed in Fig. 6, and NP insertion frequencies from the other tissues are listed in Table S6 in the supplemental material. Only 0.79% and 2.71% of all adrenal gland reads had nontemplated insertions for the NP and L mRNA, respectively (Fig. 6A and B). Several factors could contribute to the lower editing rates observed *in vivo* versus cell culture. These factors include an asynchronous infection *in vivo* compared to a high-MOI synchronized *in vitro* infection, differing editing frequencies between cell culture and animal infection, or another undefined reason. Nonetheless, these data do demonstrate editing

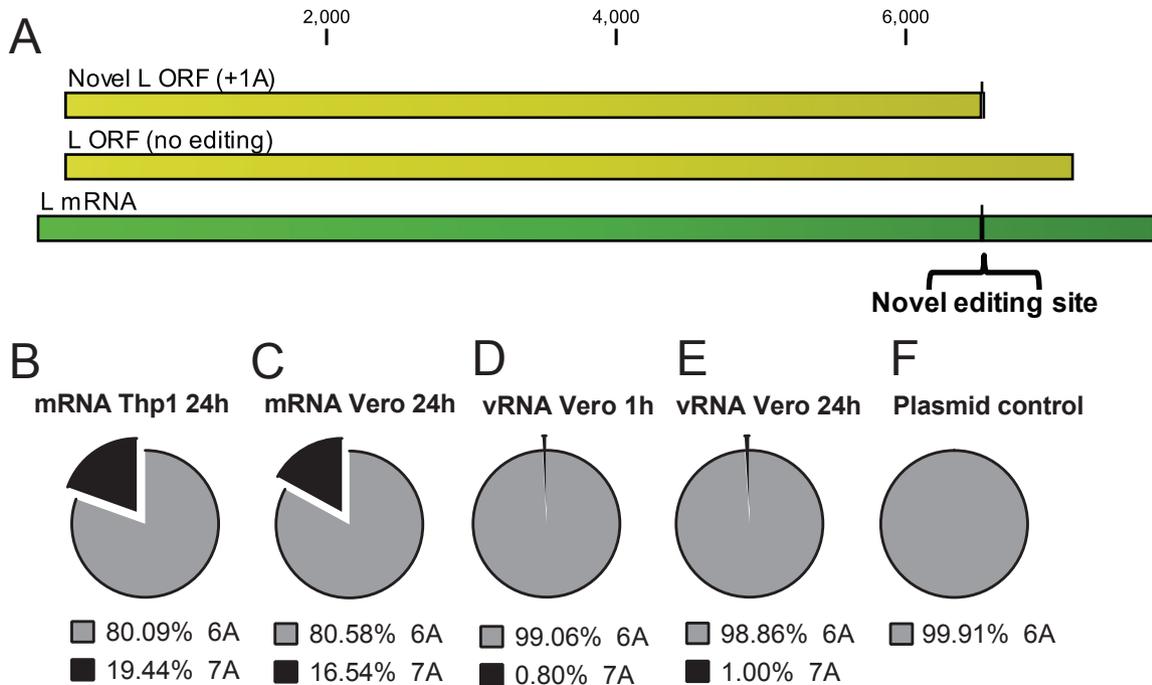


FIG 5 Insertion frequencies in the MARV-Ang L homopolymer region, nt 17810 to 17815. (A) Depiction of the L mRNA (dark green) and different translation products produced, depending on editing. (B to F) Pie charts depicting the numbers of adenosine residues at a novel location in the MARV-Ang L ORF. The numbers of A residues were enumerated within the homopolymer region from sequencing reads with 10 nt of matching sequence (underlined in the sequence) directly flanking each side from positions 17799 to 17825 (GCTCAAATGCAAAAAACTCAGAATGG). (B) Homopolymer insertion frequency in mRNA of Thp1 cells at 24 hpi, determined from libraries constructed from chemically sheared RNA. (C) mRNA from Vero cells at 24 hpi, determined from libraries constructed from chemically sheared RNA. (D) Homopolymer insertion frequency of genomic RNA (vRNA) at 1 hpi, determined from an amplicon encompassing the region of interest. (E) The same experiment as in panel D, except that vRNA was assessed at 24 hpi. (F) Insertion frequency from a DNA plasmid harboring the MARV-Ang L ORF.

of L in MARV-Ang *in vivo*. This observation, coupled with data indicating editing frequencies in EBOV GP that differ between *in vitro* and *in vivo* infections, suggest that the functional significance of MARV editing deserves further investigation (17).

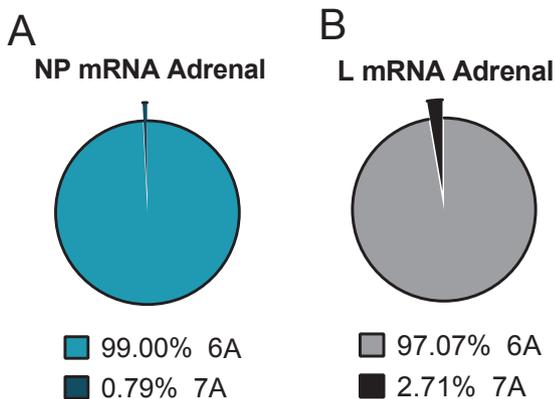


FIG 6 Insertion frequencies within homopolymer regions of MARV-Ang NP and L from the adrenal glands of infected macaques. Pie charts depict the numbers of adenosine residues at the novel locations within the MARV-Ang L ORF detected in RNA extracted from infected macaque mRNA. The numbers of A residues were enumerated within the homopolymer region from sequencing reads with 10 nt of matching sequence (underline in the sequence) directly flanking each side from positions 17799 to 17825 (GCTCAAATGCAAAAAACTCAGAATGG). (A) Homopolymer insertion frequency in MARV-Ang NP mRNA at 9 dpi within adrenal gland, determined from an amplicon encompassing the region of interest. (B) Homopolymer insertion frequency in MARV-Ang L mRNA at 9 dpi within adrenal gland, determined from an amplicon encompassing the region of interest.

DISCUSSION

The application of deep sequencing technologies allows for a rapid and comprehensive analysis of viral transcription and replication that could not be readily achieved with earlier technologies. There are previous reports that described deep RNA sequencing from virus-infected cells (30). However, the present study represents the first to profile viral mRNAs from either Ebola or Marburg viruses. Utilizing the Illumina platform, our deep sequencing analysis provides detailed characterizations of the mRNAs produced by these viruses and identifies previously unrecognized mechanisms that diversify substantial percentages of viral mRNAs. We provide evidence of a filovirus transcriptional gradient in a 3'-to-5' direction for both EBOV and MARV-Ang. This model is consistent with other NNS viruses, such as VSV (12). Previous work from our group suggested a gradient was indeed present in EBOV-infected cells (19); however, these data provide evidence for a transcriptional gradient at a single-nucleotide resolution, while the previous experiments relied solely on quantitative RT-PCR analysis. Furthermore, these are the first published data profiling the ratios of MARV transcription products. Moreover, all of these observations were confirmed in both Vero and Thp1 cells

and across three time points (see Fig. S1 in the supplemental material), indicating the reproducibility of this analysis.

Previous studies indicated that transcription between the gene borders is present, but at much-reduced levels relative to the primary transcript (14). Our analysis agrees with these results and indicates that nucleotide coverage is drastically reduced within intergenic regions (Fig. 1; see also Table S2 in the supplemental material). This indicates that the predicted start and stop sequences flanking the seven mRNAs prevent a large number of “readthrough” transcripts and that the majority of filovirus mRNAs contain only a single ORF. This is noteworthy, since the EBOV genome has three overlapping transcriptional stop and start sequences (VP35 and VP40, GP and VP30, and VP24 and L), and MARV contains one such overlapping region (VP30 and VP24). This suggests that transcribing viral polymerases encountering an overlapping start-stop signal still stop efficiently. How the polymerase manages to restart and begin transcription of the downstream gene and whether this arrangement has consequences for regulation of gene expression were not revealed by the present experiments.

Notably, our study also identified previously unappreciated mechanisms that diversify filoviral gene products. These include hot spots in the EBOV GP mRNA for modification by cytosine deaminases, although this remains to be proven, and the MARV NP 3'-UTR, where frequent U-C substitutions are highly suggestive of editing by ADARs. ADARs are host-encoded enzymes that deaminate adenosines to inosines in RNAs with a double-stranded nature (31). Two ADARs with adenosine deaminase activity, ADAR1 and ADAR2, have been described. ADAR1 exhibits two isoforms, p150 and p100, with the p150 isoform being IFN inducible and exhibiting both cytoplasmic and nuclear localization (31). Given that the majority of changes detected in our positive-sense viral mRNAs were U-C changes, this suggests that the editing detected occurs predominately on the negative-sense viral genomic RNA. Because MARV replication is cytoplasmic, the editing likely reflects activity of the p150 ADAR1 isoform. ADAR editing sites display a bias based on the 5' neighbor where $A = U > C > G$ (32). The changes in the negative-sense genomic RNA corresponding to the NP 3'-UTR follow this rule and have either an A or U preceding the edited A. For the 24-hpi Vero cell NP mRNAs, apparent ADAR editing was present in 8 to 24% of reads counted in our analysis. Similar rates of editing were seen in the Thp1 cell-derived samples. This suggests that a significant minority of templates from which the reads were derived were edited. Coupled with the localized editing, where the MARV-Ang NP 3'-UTR is preferentially targeted, this reflects a selective editing mechanism (33).

ADAR editing has been detected in a variety of viruses, including negative-sense RNA viruses (31, 34–36). It is notable that during the course of mouse adaptation, the MARV variant Ci67 and the Angola strain and the Ravn strain of MARV each underwent apparent ADAR editing, with clusters of A-G changes accumulating in the genomic RNA (37, 38). The functional consequences of such editing can vary but can result in both proviral and antiviral effects (31). In the case of measles virus, the proviral effects include enhanced replication, reduced apoptosis, decreased activation of PKR, and decreased phosphorylation of interferon regulatory factor 3 (39, 40). Therefore, the editing detected in the present study deserves attention in future experiments.

It is well documented that multiple paramyxovirus poly-

merases stutter to either polyadenylate the 3' end of their mRNA or to add nontemplated nucleotides within open reading frames to produce alternative protein products (41). Our analyses highlight homopolymer regions in both EBOV and MARV-Ang where the filovirus polymerase inserts nontemplated nucleotides. These results confirmed the previously described mRNA editing frequencies at the EBOV nt 6918 to 6924 site within GP (approximately 25 to 30% of all sequencing reads had at least one extra nucleotide). Our data also demonstrate the frequent introduction of nontemplated insertions in the EBOV genomic vRNA at 24 hpi at this site. This indicates that (i) RNA editing at the canonical GP site occurs during both transcription and replication, (ii) editing occurs with high frequency *in vitro*, and (iii) viruses with edited genomes are selected against *in vivo* since the viruses isolated *in vivo* (including human samples from the current outbreak) predominately have 7-U genomes.

Additional 6-A and 6-U regions are present within the EBOV genome. Several of these also displayed increased insertion frequencies relative to the corresponding controls, demonstrating that the viral polymerase is prone to insertions within homopolymer regions. However, of those where editing was detected, the insertion frequencies were significantly lower than for the described GP editing site. Editing frequencies also varied, suggesting that other factors regulate the insertion frequency. One of the highest editing rates was at the homopolymer region at nt 6377 to 6383 within GP, where insertions were present in 4.30% of reads in Thp1 cells. A single-nucleotide insertion here would result in a truncated GP product of 117 amino acids (identical sequence to wild-type GP, followed by a single amino acid and then a stop codon [Fig. 3]). This protein would retain a signal peptide and about 60 amino acids of the receptor binding domain (22). Finally, it may be relevant to the regulation of editing that the flanking sequences surrounding the canonical EBOV GP editing site are 5'-GAAACUAAAAAACCUCAC-3', while the flanking sequences surrounding the newly identified editing site are 5'-GAAAUCAAAAAACCUGAC-3'. For the canonical GP editing site, sequences immediately surrounding the editing site and predicted secondary structures upstream of the editing site have been implicated as critical regulatory features (24). These similar motifs immediately surrounding the 6377 to 6383 site may be recognized in a similar way by the viral polymerase, although this requires further investigation.

We also report the first editing events described for MARVs, with single-nucleotide insertions in both the NP and L mRNAs. These insertions, which according to the deep sequencing results occur at substantial frequencies, would result in the translation of truncated versions of the NP and L proteins. These proteins may have novel functions within an infected cell or perhaps modulate the previously described function of their corresponding full-length proteins; further study is required to address this possibility.

The *in vivo* analysis of MARV-Ang mRNA in infected animals also demonstrated insertions in the NP and L mRNAs in infected animals. However, the frequency of insertions was substantially lower *in vivo* than in the cell culture samples (a reduction of at least 10-fold at each site). Differences between *in vivo* and *in vitro* infection conditions could account for these results, such as a high (*in vitro*) versus low (*in vivo*) MOI, asynchronous infection *in vivo* (at the end stage of disease) versus synchronous infection *in vitro*,

or the absence of host pressures that might select against editing *in vivo*.

It is notable that the sites of noncanonical editing are highly conserved. For example, the 2014 West Africa EBOVs belong to a separate clade than the 1976 strain evaluated in the present study (42). At the nucleotide level, isolates of the 2014 outbreak displayed 97% identity to the 1976 Mayinga strain used here (data not shown). Despite the divergence, all but one of the polymeric sites listed in Table S5 in the supplemental material are completely conserved among 18 strains from the 2014 outbreak, including 3 early isolates from Guinea and 15 later isolates from Sierra Leone. Only the site in NP beginning at position 850 is not conserved; the 1976 Mayinga strain has 6 As, but the outbreak sequences have the sequence 5'-GAGAAA-3' (data not shown). In another example, the homopolymer stretch in MARV NP is completely conserved among all MARV full-length genomic sequences recently analyzed, even between the two different MARV clades (data not shown). Conservation of these sites indicates that they are not detrimental and suggests that they may provide some advantage. These observations highlight the need for future studies to determine the mechanistic basis for the novel editing events identified, their functional consequences, including characterization of the predicted novel viral gene products, and their frequencies in different viral growth contexts.

MATERIALS AND METHODS

Ebolavirus and Marburgvirus infections and RNA-Seq of infected cell lines. Total RNA was isolated by using Trizol (Invitrogen) from Vero cells (African Green monkey kidney epithelial cells) or differentiated Thp1 cells (human monocytic leukemia cells) infected with EBOV Mayinga (CDC isolate number 808011) or MARV-Ang (CDC isolate number 200501379) at an MOI of 3. Alternatively, total RNA was isolated from tissues of MARV-Ang-infected macaques (described below). Poly(A) mRNA was purified from 2 μ g of RNA by using oligo(dT) magnetic beads (Invitrogen), and cDNA libraries were established for sequencing on the Illumina HiSeq 2500 platform by using the NEB Next mRNA library preparation kit (New England Biolabs).

RNA-Seq reference genome mapping and identification of putative editing sites. Illumina HiSeq 100-nt reads were mapped to the EBOV (GenBank accession number NC_002549.1) or MARV-Ang genome (GenBank accession number DQ447653.1) by using the TopHat/Cufflinks software suite v2 with the default settings. No more than two mismatches between the reference genome and sequencing read were allowed. Insertion or deletion sites identified by the program were the basis for further investigation. Genome coverage was computed by using the BAM Tools package. The TopHat aligner program performs best when reads have high identity to the genome. To capture reads with multiple inserted bases, we employed a more sensitive aligner, NCBI Blast, to search for all EBOV/MARV-Ang reads and extracted them for further analysis. Sequences flanking putative editing sites were extracted from raw read data by using custom programs written in Perl. The program identified reads with 10 nt of exactly matching sequence flanking the putative editing site and then counted the number and composition of bases. In addition, the Illumina quality scores of the regions were extracted.

Identification of minor variants within infected cells. Sequencing reads were first quality trimmed to remove sections of reads where greater than 10% of the bases had quality values less than 20. The quality-trimmed reads were then mapped to the specified GenBank reference sequence for each sample by using the `clc_ref_assemble_long` software. Consensus sequence differences between the read data and the specified GenBank reference sequence were identified by the CLCbio `find_variations` software. Single nucleotide polymorphism variations were identified

by using JCVI custom software (see the description of methods in the supplemental material for further details).

Validation of putative editing regions from RNA. Regions that differed from viral isolate sequences were identified by bioinformatic analysis. To control for polymerase insertions that could occur during reverse transcription, RNA-directed RNA transcription and Illumina sequencing were performed on products amplified by RT-PCR or PCR from RNA and plasmid DNA, respectively. Primers flanking the regions were designed and synthesized with Illumina-specific adapters. Amplicons were designed to position the putative editing site of interest ~50 nt into the sequencing read, because Illumina base calling is most reliable between bases 21 and 75. Four nucleic acid contexts were tested: mRNA from genes transcribed by viral polymerase during infection, viral genomic RNA synthesized during infection by virus L polymerase, mRNA from ectopically expressed genes transcribed by host RNA polymerase II, and plasmid DNA of cloned viral genes. To minimize chances of cross-contamination, for each PCR experiment we employed different DNA bar codes and the experiments were performed separately.

For the *in vivo* analysis of putative MARV-Ang RNA editing, MARV-Ang and total RNAs were isolated from the indicated tissues of MARV-Ang-infected rhesus macaques (*Macaca mulatta*) at day 9 postinfection. RT-PCR was employed as described above to generate amplicons that spanned the NP and L homopolymer regions, and the percentage of insertions at each region was evaluated.

Evaluation of viral RNA from rhesus macaque tissues. Select tissues from an adult rhesus macaque lethally challenged with MARV-Ang were used to isolate total RNA for RNA-Seq analysis. This animal was a non-treated, nonvaccinated control of a previous study. Use of these tissues within this study is in support of the “reuse” portion of the 3 Rs principles of animal ethics. The animal from the previous study was handled in animal biosafety level 2 (BSL-2) and BSL-4 containment spaces in the Galveston National Laboratory (GNL) at the University of Texas Medical Branch (UTMB), Galveston, TX. Research was conducted in compliance with the Animal Welfare Act and other federal statutes and regulations relating to animals and experiments involving animals, and it adhered to principles stated in the 8th edition (2011) of the *Guide for the Care and Use of Laboratory Animals* of the National Research Council (43) (see also the methods described in the supplemental material).

Sequence data accession numbers. Sequencing data are publicly available in NCBI BioProjects under accession numbers PRJNA258131 (EBOV) and PRJNA264121 (MARV).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.02011-14/-/DCSupplemental>.

Table S1, PDF file, 0.01 MB.
Table S2, XLSX file, 1.6 MB.
Table S3, PDF file, 0.02 MB.
Table S4, PDF file, 0.02 MB.
Table S5, PDF file, 0.04 MB.
Table S6, PDF file, 0.02 MB.
Figure S1, EPS file, 1 MB.
Figure S2, EPS file, 1.3 MB.
Text S1, DOCX file, 0.02 MB.

ACKNOWLEDGMENTS

We thank Christine Schwall and Kari Dillely for helpful review and editing of the manuscript.

This work was supported by NIH grants R01AI059536 and U19AI109664 (Basler-PI) to C.F.B., by R01AI8945403 to T.W.G., and U19AI109945 (Basler-PI) to C.F.B. and T.W.G. R.S.S. was supported by the J. Craig Venter Institute (Start-up 9260) and by the Department of Homeland Security (contract HSHQDC-13-C-B0016).

REFERENCES

- WHO Ebola Response Team. 23 September 2014. Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections. *N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1411100>.
- Feldmann H, Geisbert TW. 2011. Ebola haemorrhagic fever. *Lancet* 377:849–862. [http://dx.doi.org/10.1016/S0140-6736\(10\)60667-8](http://dx.doi.org/10.1016/S0140-6736(10)60667-8).
- Hartman AL, Towner JS, Nichol ST. 2010. Ebola and marburg hemorrhagic fever. *Clin. Lab. Med.* 30:161–177. <http://dx.doi.org/10.1016/j.cll.2009.12.001>.
- Bagcchi S. 2014. Ebola haemorrhagic fever in West Africa. *Lancet Infect. Dis.* 14:375. [http://dx.doi.org/10.1016/S1473-3099\(14\)70034-9](http://dx.doi.org/10.1016/S1473-3099(14)70034-9).
- Feldmann H, Mühlberger E, Randolph A, Will C, Kiley MP, Sanchez A, Klenk HD. 1992. Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. *Virus Res.* 24:1–19. [http://dx.doi.org/10.1016/0168-1702\(92\)90027-7](http://dx.doi.org/10.1016/0168-1702(92)90027-7).
- 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. [http://dx.doi.org/10.1016/0168-1702\(93\)90063-S](http://dx.doi.org/10.1016/0168-1702(93)90063-S).
- 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. <http://dx.doi.org/10.1128/JVI.01525-12>.
- 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. <http://dx.doi.org/10.1038/srep01206>.
- Weik M, Modrof J, Klenk HD, Becker S, Mühlberger E. 2002. Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J. Virol.* 76:8532–8539. <http://dx.doi.org/10.1128/JVI.76.17.8532-8539.2002>.
- Ferron F, Longhi S, Henrissat B, Canard B. 2002. Viral RNA-polymerases—a predicted 2'-O-ribose methyltransferase domain shared by all mononegavirales. *Trends Biochem. Sci.* 27:222–224. [http://dx.doi.org/10.1016/S0968-0004\(02\)02091-1](http://dx.doi.org/10.1016/S0968-0004(02)02091-1).
- Sanchez A, Kiley MP. 1987. Identification and analysis of Ebola virus messenger RNA. *Virology* 157:414–420. [http://dx.doi.org/10.1016/0042-6822\(87\)90283-2](http://dx.doi.org/10.1016/0042-6822(87)90283-2).
- Iverson LE, Rose JK. 1981. Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. *Cell* 23:477–484. [http://dx.doi.org/10.1016/0092-8674\(81\)90143-4](http://dx.doi.org/10.1016/0092-8674(81)90143-4).
- Neumann G, Watanabe S, Kawaoka Y. 2009. Characterization of Ebolavirus regulatory genomic regions. *Virus Res.* 144:1–7. <http://dx.doi.org/10.1016/j.virusres.2009.02.005>.
- Brauburger K, Boehmann Y, Tsuda Y, Hoenen T, Olejnik J, Schumann M, Ebihara H, Mühlberger E. 2014. Analysis of the highly diverse gene borders in Ebola virus reveals a distinct mechanism of transcriptional regulation. *J. Virol.* 88:12558–12571.
- 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. <http://dx.doi.org/10.1073/pnas.93.8.3602>.
- Volchkov VE, Becker S, Volchkova VA, Ternovoj VA, Kotov AN, Netesov SV, Klenk HD. 1995. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. *Virology* 214:421–430. <http://dx.doi.org/10.1006/viro.1995.0052>.
- Kugelman JR, Lee MS, Rossi CA, McCarthy SE, Radoshitzky SR, Dye JM, Hensley LE, Honko A, Kuhn JH, Jahrling PB, Warren TK, Whitehouse CA, Bavari S, Palacios G. 2012. Ebola virus genome plasticity as a marker of its passaging history: a comparison of in vitro passaging to non-human primate infection. *PLoS One* 7:e50316. <http://dx.doi.org/10.1371/journal.pone.0050316>.
- Volchkova VA, Dolnik O, Martinez MJ, Reynard O, Volchkov VE. 2011. Genomic RNA editing and its impact on Ebola virus adaptation during serial passages in cell culture and infection of guinea pigs. *J. Infect. Dis.* 204:S941–S946. <http://dx.doi.org/10.1093/infdis/jir321>.
- 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.
- Diaz MO, Ziemins S, Le Beau MM, Pitha P, Smith SD, Chilcote RR, Rowley JD. 1988. Homozygous deletion of the alpha- and beta 1-interferon genes in human leukemia and derived cell lines. *Proc. Natl. Acad. Sci. U. S. A.* 85:5259–5263. <http://dx.doi.org/10.1073/pnas.85.14.5259>.
- Mosca JD, Pitha PM. 1986. Transcriptional and posttranscriptional regulation of exogenous human beta interferon gene in simian cells defective in interferon synthesis. *Mol. Cell. Biol.* 6:2279–2283.
- Lennemann NJ, Rhein BA, Ndungo E, Chandran K, Qiu X, Maury W. 2014. Comprehensive functional analysis of N-linked glycans on Ebola virus GP1. *mBio* 5(1):e00862-13. <http://dx.doi.org/10.1128/mBio.00862-13>.
- 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. <http://dx.doi.org/10.1128/JVI.02190-10>.
- 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. <http://dx.doi.org/10.1371/journal.ppat.1003677>.
- Volchkov VE, Volchkova VA, Mühlberger 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. <http://dx.doi.org/10.1126/science.1057269>.
- Volchkov VE, Chepurinov AA, Volchkova VA, Ternovoj VA, Klenk HD. 2000. Molecular characterization of guinea pig-adapted variants of Ebola virus. *Virology* 277:147–155. <http://dx.doi.org/10.1006/viro.2000.0572>.
- Thomas MJ, Platas AA, Hawley DK. 1998. Transcriptional fidelity and proofreading by RNA polymerase II. *Cell* 93:627–637. [http://dx.doi.org/10.1016/S0092-8674\(00\)81191-5](http://dx.doi.org/10.1016/S0092-8674(00)81191-5).
- 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. <http://dx.doi.org/10.1128/JVI.03118-12>.
- Gire SK, Goba A, Andersen KG, Sealfon RS, Park DJ, Kanneh L, Jalloh S, Momoh M, Fullah M, Dudas G, Wohl S, Moses LM, Yozwiak NL, Winnicki S, Matranga CB, Malboeuf CM, Qu J, Gladden AD, Schaffner SF, Yang X, Jiang PP, Nekoui M, Colubri A, Coomber MR, Fonnies M, Moigboi A, Gbalkie M, Kamara FK, Tucker V, Koneuwa E, Saffa S, Sellu J, Jalloh AA, Kovoma A, Koninga J, Mustapha I, Kargbo K, Foday M, Yillah M, Kanneh F, Robert W, Massally JL, Chapman SB, Bochicchio J, Murphy C, Nusbaum C, Young S, Birren BW, Grant DS, Scheiffelin JS, Lander ES, Hapoi C, Gevaou SM, Gnirke A, Rambaut A, Garry RF, Khan SH, Sabeti PC. 2014. Genomic surveillance elucidates Ebola virus origin and transmission during the outbreak. *Science* 345:1369–1372. <http://dx.doi.org/10.1126/science.1259657>.
- Acevedo A, Brodsky L, Andino R. 2014. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* 505:686–690.
- Pfaller CK, Li Z, George CX, Samuel CE. 2011. Protein kinase PKR and RNA adenosine deaminase ADAR1: new roles for old players as modulators of the interferon response. *Curr. Opin. Immunol.* 23:573–582. <http://dx.doi.org/10.1016/j.coi.2011.08.009>.
- Polson AG, Bass BL. 1994. Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J.* 13:5701–5711.
- Lehmann KA, Bass BL. 1999. The importance of internal loops within RNA substrates of ADAR1. *J. Mol. Biol.* 291:1–13. <http://dx.doi.org/10.1006/jmbi.1999.2914>.
- 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. [http://dx.doi.org/10.1016/0092-8674\(88\)90048-7](http://dx.doi.org/10.1016/0092-8674(88)90048-7).
- Murphy DG, Dimock K, Kang CY. 1991. Numerous transitions in human parainfluenza virus 3 RNA recovered from persistently infected cells. *Virology* 181:760–763. [http://dx.doi.org/10.1016/0042-6822\(91\)90913-V](http://dx.doi.org/10.1016/0042-6822(91)90913-V).
- O'Hara PJ, Nichol ST, Horodyski FM, Holland JJ. 1984. Vesicular stomatitis virus defective interfering particles can contain extensive genomic sequence rearrangements and base substitutions. *Cell* 36:915–924. [http://dx.doi.org/10.1016/0092-8674\(84\)90041-2](http://dx.doi.org/10.1016/0092-8674(84)90041-2).
- Lofts LL, Wells JB, Bavari S, Warfield KL. 2011. Key genomic changes necessary for an in vivo lethal mouse Marburgvirus variant selection process. *J. Virol.* 85:3905–3917. <http://dx.doi.org/10.1128/JVI.02372-10>.

38. Qiu X, Wong G, Audet J, Cutts T, Niu Y, Booth S, Kobinger GP. 2014. Establishment and characterization of a lethal mouse model for the Angola strain of Marburg virus. *J. Virol.* **88**:12703–12714.
39. Ward SV, George CX, Welch MJ, Liou LY, Hahm B, Lewicki H, de la Torre JC, Samuel CE, Oldstone MB. 2011. RNA editing enzyme adenosine deaminase is a restriction factor for controlling measles virus replication that also is required for embryogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **108**:331–336. <http://dx.doi.org/10.1073/pnas.1017241108>.
40. Toth AM, Li Z, Cattaneo R, Samuel CE. 2009. RNA-specific adenosine deaminase ADAR1 suppresses measles virus-induced apoptosis and activation of protein kinase PKR. *J. Biol. Chem.* **284**:29350–29356. <http://dx.doi.org/10.1074/jbc.M109.045146>.
41. Kolakofsky D, Roux L, Garcin D, Ruigrok RW. 2005. Paramyxovirus mRNA editing, the “rule of six” and error catastrophe: a hypothesis. *J. Gen. Virol.* **86**:1869–1877. <http://dx.doi.org/10.1099/vir.0.80986-0>.
42. 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**:1417–1425. <http://dx.doi.org/10.1056/NEJMoa1404505>.
43. National Research Council. 2011. Guide for the care and use of laboratory animals, 8th ed. National Academies Press, Washington, DC.