**Cell Reports, Volume *35***

## Supplemental information

# Cellular mRNA triggers structural transformation of Ebola virus matrix protein VP40 to its essential regulatory form

**Sara Landeras-Bueno, Hal Wasserman, Glenn Oliveira, Zachary L. VanAernum, Florian Busch, Zhe Li Salie, Vicki H. Wysocki, Kristian Andersen, and Erica Ollmann Saphire**

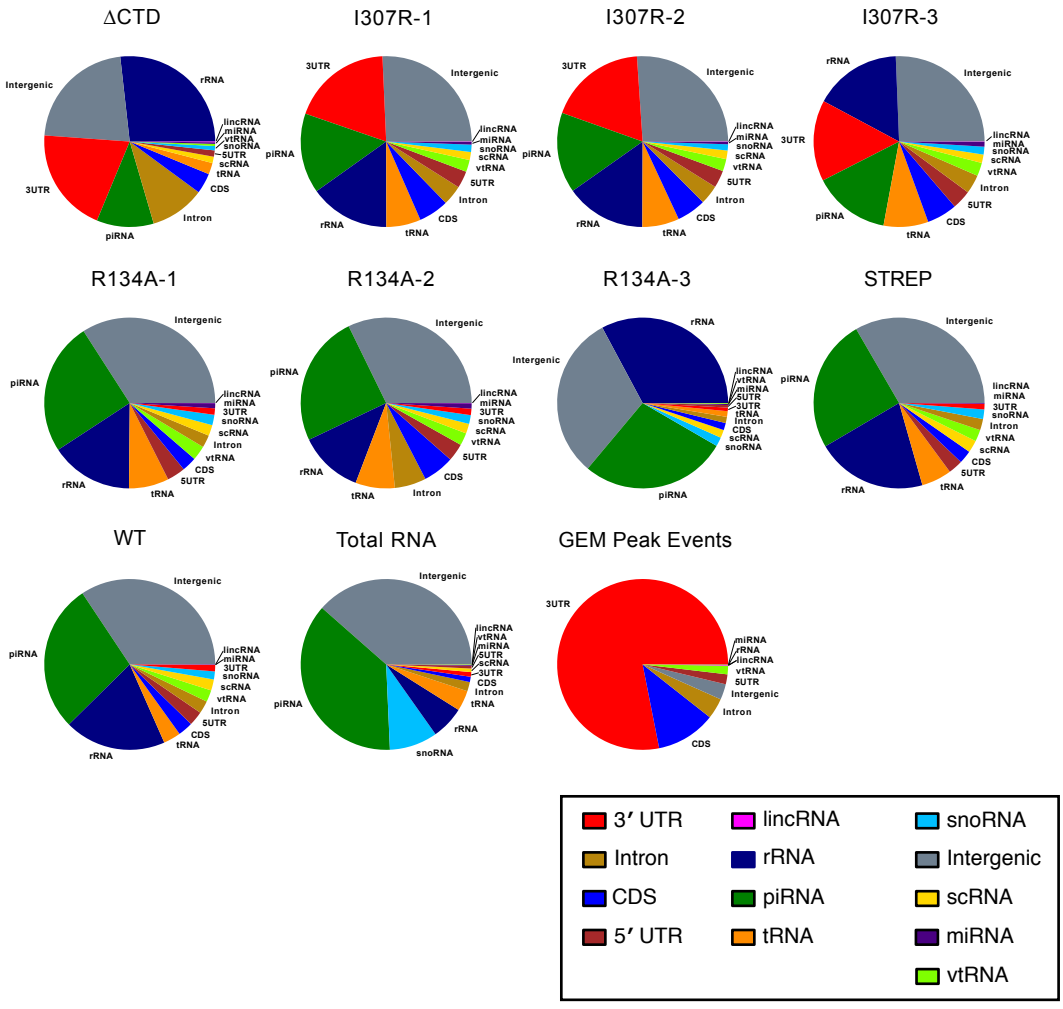


**Figure S1: The cellular 3′ UTRs are enriched in all of our ring-forming VP40 mutant samples. Related to Figures 2 and 3.**

Human genome-aligned reads categorized by RNA feature: the 3′ UTR, intron, coding sequence (CDS), and 5′ UTR components of mRNA, as well as several forms of non-coding RNA. Our samples include ring-forming VP40 mutants (VP40-I307R biological replicates, VP40ΔCTD) as well as non-ring-rich cases (VP40-R134A biological replicates, VP40-WT) and VP40-free controls (Strep-tag control, Total RNA). An Ebola minigenome replicon system was co-transfected in all cases except I307R-3 and Total RNA. We observe 3′ UTR enrichment in all of the ring-forming VP40 mutant samples relative to all of the other samples. Also included is a similar categorization of the VP40-RNA binding event loci called by the GEM peak-finding algorithm (Figure 3).

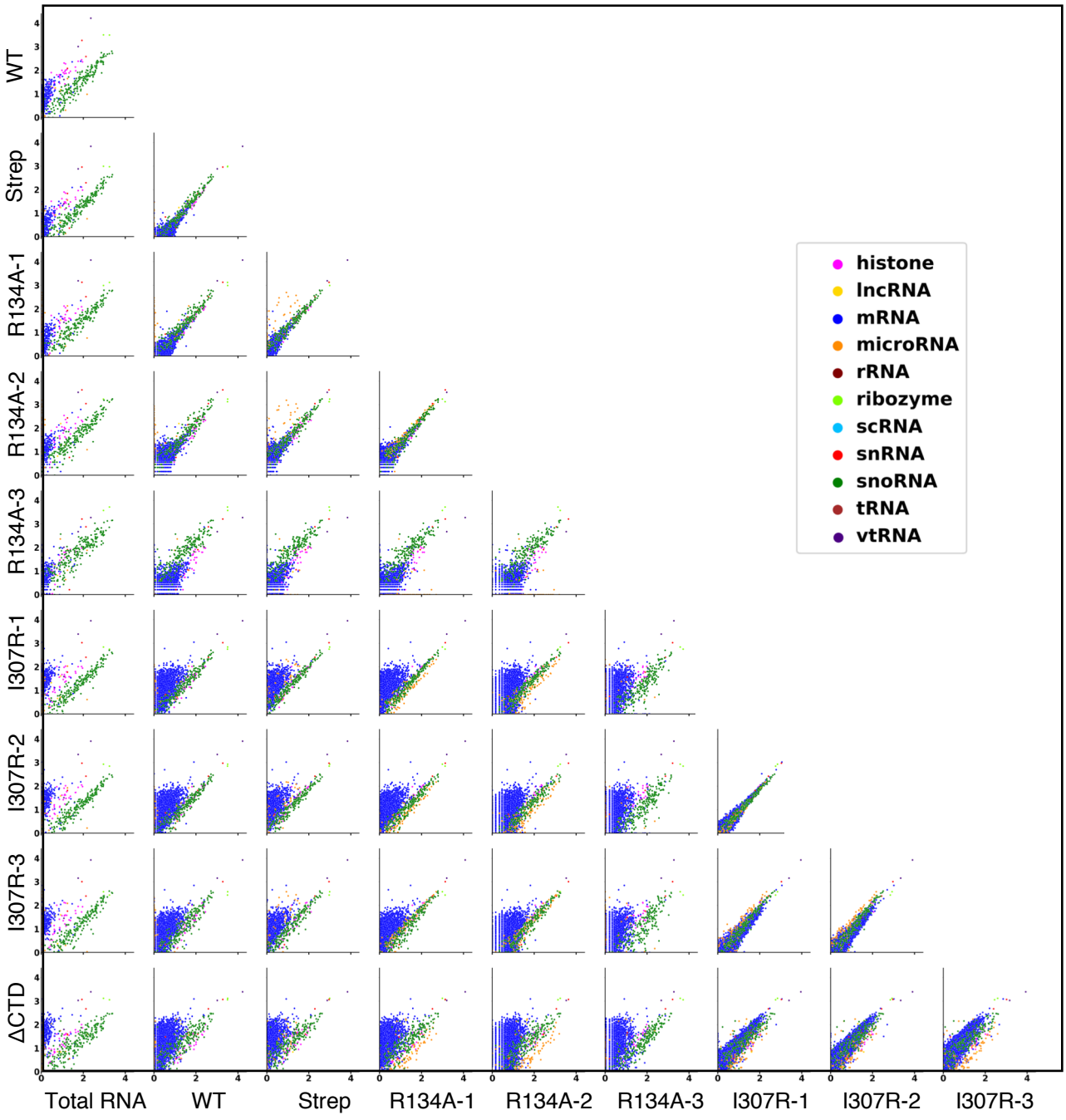


**Figure S2: The cellular mRNA is enriched in all of our ring-forming VP40 mutant samples. Related to Figure 2.**

The cellular RNA transcript abundances compared between all pairs of our samples. The highest-expressed genes are displayed in each case. We observe mRNA enrichment in all of the ring-forming VP40 mutant samples (VP40-I307R biological replicates, VP40ΔCTD) relative to all of the other samples.

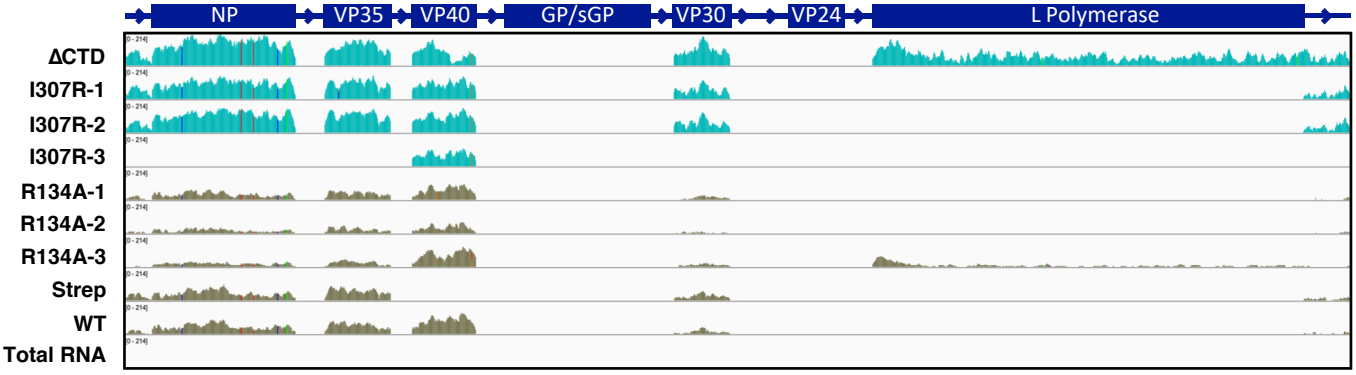


**Figure S3: VP40 binding events are not observed for the *Zaire ebolavirus* genome alignment. Related to Figure 3.**

Reads aligned to the *Zaire ebolavirus* genome (AF086833.2). Uniform enrichment is observed in the ring-forming VP40 mutant samples (VP40-I307R biological replicates, VP40ΔCTD), but we do not observe region-specific enrichment or well-defined read peaks. A plasmid encoding VP40 (without the natural Ebola virus UTRs) was transfected for all of the samples except the Strep-tag control and Total RNA. An Ebola minigenome replicon system (which includes the NP, VP35, VP30, and L genes without natural Ebola virus UTRs, as well as a reporter gene flanked by the Ebola virus genomic leader and trailer) was co-transfected for all of the samples except I307R-3 and Total RNA.

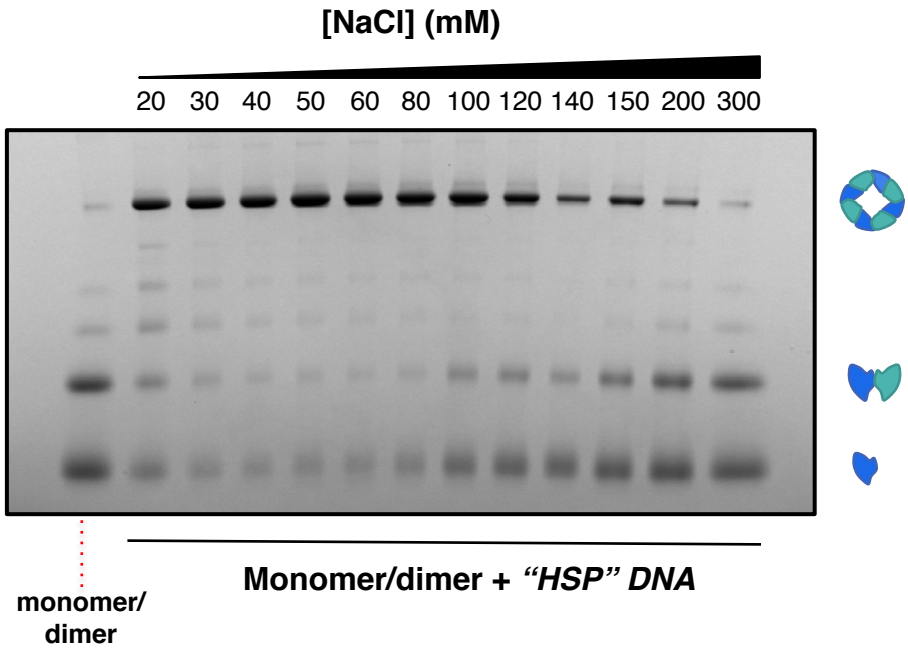


**Figure S4: A low salt concentration is necessary for the nucleic acid-induced transformation to the VP40 ring. Related to Figure 4.**
Lanes 2-13: A gradient of salt concentrations for the overnight incubation of the VP40 monomer/dimer with the "HSP" DNA oligo (Table S2). Lane 1 is a DNA-free control.

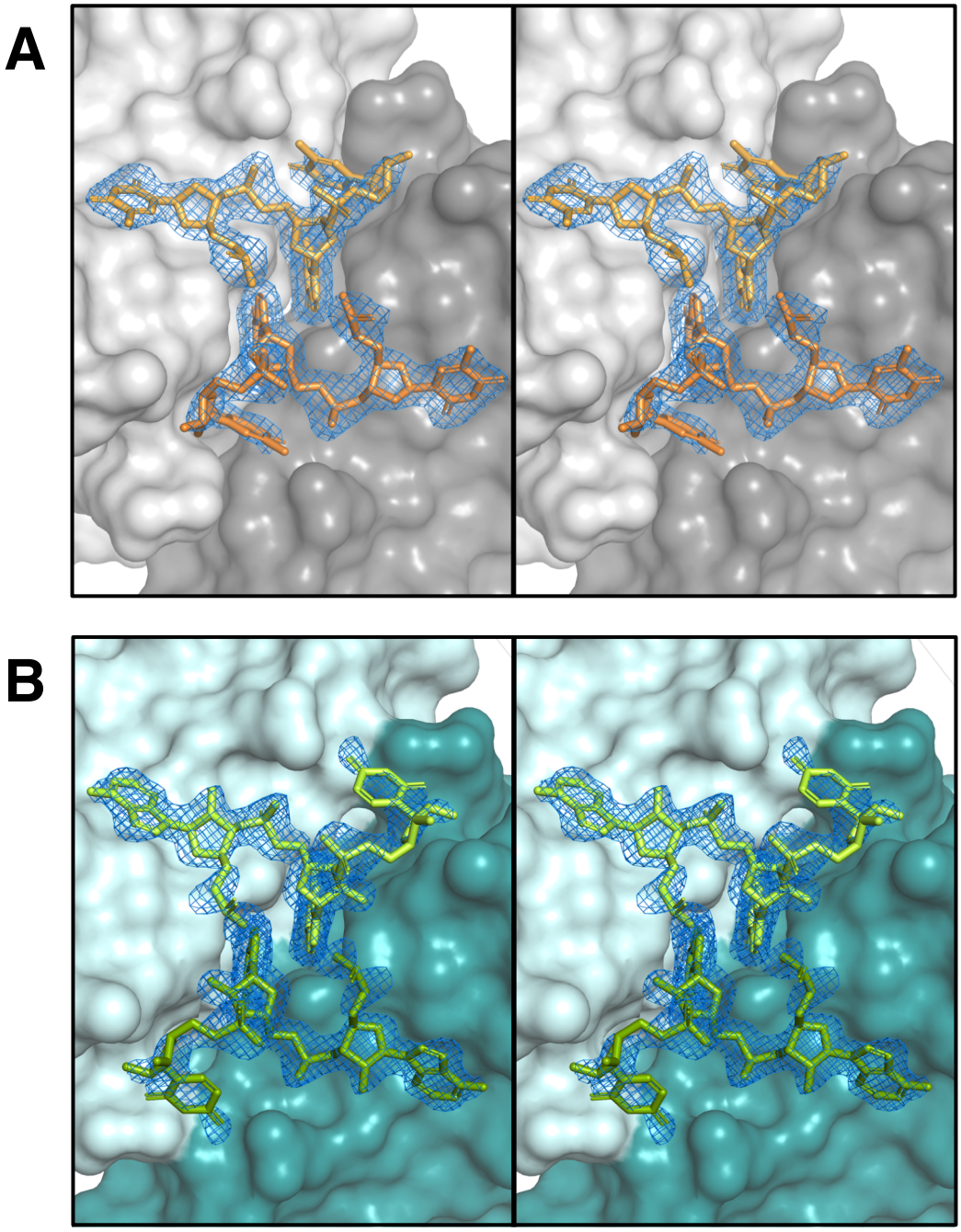


**Figure S5: Nucleic acid binds to the protein surface in our VP40 ring structures. Related to Figure 5.**
**A:** A cross-eye stereo view of the VP40 octameric ring generated using the "HSP" DNA oligo (PDB:7K5D). Two symmetrical DNA fragments are seen bound to the protein surface (as in Figure 5E), along with their electron density maps (2Fo-Fc at 1 sigma). Each DNA fragment has been built as 5′-TGT-3′, but incomplete or indeterminate density for the first and third bases prevents the conclusive identification of these bases.
**B:** A similar cross-eye stereo view of the VP40 octameric ring generated using the "HSP" RNA oligo (PDB:7K5L). Two symmetrical RNA fragments are seen bound to the protein surface (as in Figure 5F). Each RNA fragment has been built as 5′-AGU-3′, but incomplete or indeterminate density for the first and third bases prevents the conclusive identification of these bases.

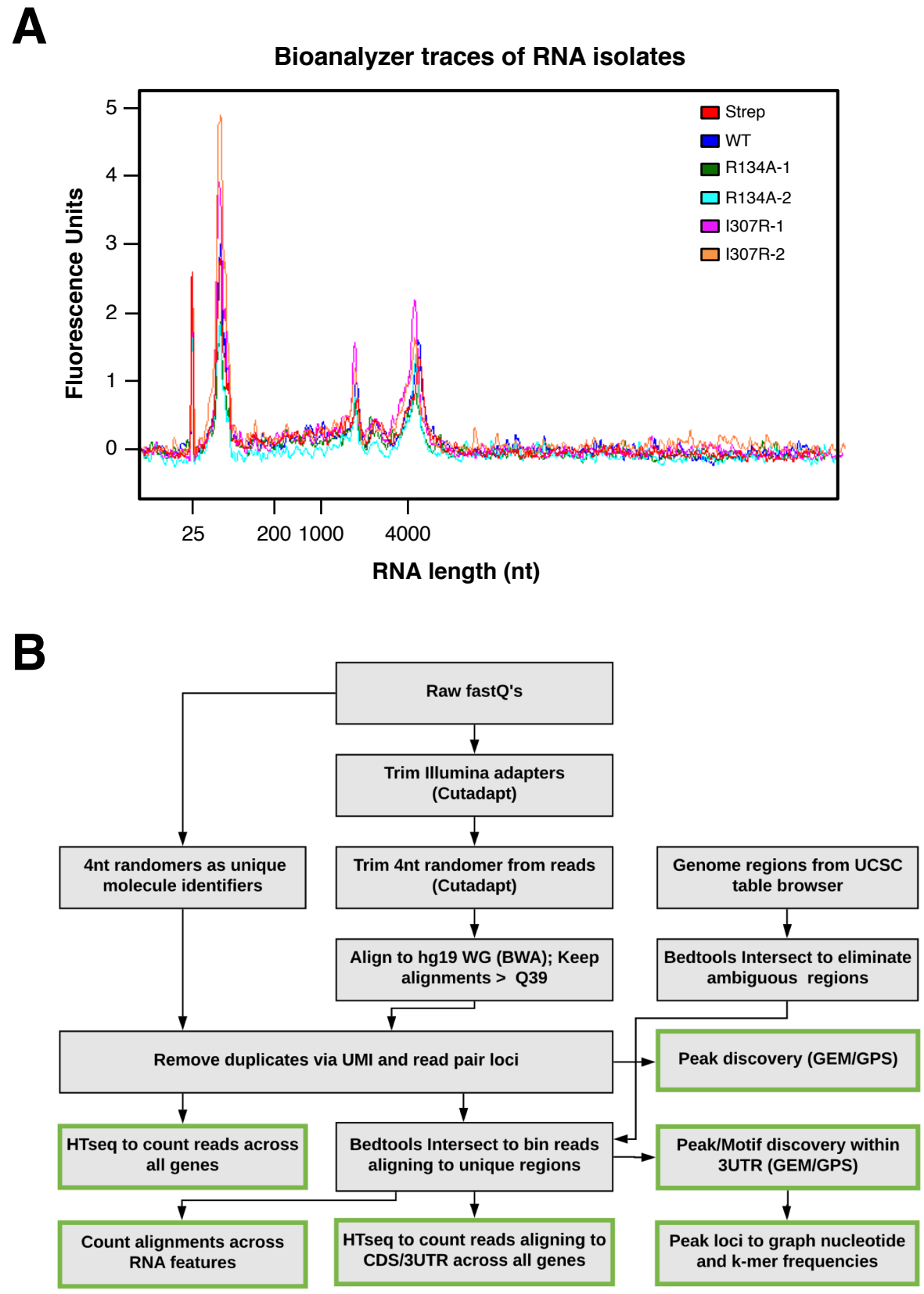


**Figure S6: Our RNA deep sequencing workflow. Related to STAR Methods.**
**A:** The Bioanalyzer 2100 (Agilent) RNA size distributions for the RNA samples sequenced and reported here.
**B:** Our RNA deep sequencing data analysis workflow.

| Sample | Raw read count | Filtered human reads | Filtered Ebola virus reads |
|---|---|---|---|
| VP40ΔCTD | 6,698,953 | 3,042,154 (45.4%) | 10,060 (0.15%) |
| VP40-I307R-1 | 7,840,083 | 1,767,111 (22.5%) | 6,715 (0.09%) |
| VP40-I307R-2 | 9,020,549 | 1,781,536 (19.7%) | 6,422 (0.07%) |
| VP40-I307R-3 | 7,438,083 | 1,307,299 (17.6%) | 960 (0.01%) |
| VP40-R134A-1 | 7,120,235 | 1,011,149 (14.2%) | 2,069 (0.03%) |
| VP40-R134A-2 | 5,332,156 | 710,604 (13.3%) | 1,148 (0.02%) |
| VP40-R134A-3 | 5,235,137 | 1,063,976 (20.3%) | 1,481 (0.03%) |
| Strep control | 9,461,806 | 1,312,057 (13.9%) | 2,110 (0.02%) |
| VP40-WT | 7,207,790 | 1,031,859 (14.3%) | 2,681 (0.04%) |
| Total RNA control | 7,391,743 | 856,051 (11.6%) | 0 (0.00%) |

**Table S1: RNA deep sequencing read counts. Related to Figure 1.**
For each of our RNA samples, the raw read count is given, followed by the read count obtained after aligning to the human genome (hg19) and filtering to exclude the duplicate reads and the low-quality alignments (< Q40). Also given is the corresponding filtered read count for the alignment to the *Zaire ebolavirus* genome (AF086833.2). The RNA samples were extracted from Strep-Tactin pulldowns of ring-forming VP40 mutants (VP40ΔCTD; three VP40-I307R biological replicates) as well as other forms of VP40 (VP40-WT; three VP40-R134A biological replicates). The Strep control was pulled down from cells in which a Strep-tag without any attached VP40 had been expressed. The Total RNA control is RNA extracted from HEK-293T cellular lysate. An Ebola minigenome replicon system was co-transfected in all cases except I307R-3 and Total RNA.

| Oligo name | Sequence (5′ – 3′) |
|---|---|
| “HSP” RNA | **UACAUUCCCAGCCUUUGUAGUGUUUUCGCCAAGCA** |
| “HSP” DNA | **TACATTCCCAGCCTTTGTAGTGTTTTCGCCAAGCA** |
| “HSP18” DNA | **CAGCCTTTGTAGTGTTTT** |
| “A36” DNA | **AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA** |
| “AGGGG” DNA | **AGGGG** |
| “S2R” DNA | **TAAAATAATTTTATTGAAT** |

**Table S2: The 3′ UTR-inspired RNA/DNA single-stranded oligonucleotides employed in Figures 4, 5, 6, S4, S5.**
The "HSP" sequence is based on a highly ring-enriched subsequence of the human *HSPA1B* 3′ UTR. "HSP18" is a length-18 subsequence of "HSP". "A36" was tested to investigate the role of the mRNA polyadenylate tail, and "AGGGG" due to an observed binding preference for G/A-rich RNA (Figure 3).

| | EBOV VP40 RNA-generated ring (PDB: 7K5L) | EBOV VP40 DNA-generated ring (PDB: 7K5D) |
|---|---|---|
| **Crystal parameters** | | |
| Resolution (Å) | 80.63 – 1.38 (1.41 – 1.38) | 40.76 – 1.78 (1.82 – 1.78) |
| Space group | P 4 2 2 | P 4 2 2 |
| Unit cell dimensions (Å) | a = b = 80.63 c = 47.08 α = β = γ = 90° | a = b = 80.44 c = 47.28 α = β = γ = 90° |
| Wavelength (Å) | 0.97931 | 0.97946 |
| **Data Collection** | | |
| No. unique reflections | 30,977 (1,580) | 15,385 (844) |
| $R_{merge}$ | 0.059 (1.519) | 0.098 (1.411) |
| I/σ(I) | 27.8 (2.4) | 15.5 (2.4) |
| $CC_{1/2}$ | 0.999 (0.853) | 0.998 (0.874) |
| Completeness (%) | 95.7 (100.0) | 99.9 (98.8) |
| Multiplicity | 24.8 (25.7) | 12.3 (11.8) |
| Wilson B-factor (Å$^2$) | 18.9 | 27.0 |
| **Refinement** | | |
| $R_{work}/R_{free}$ (%) | 18.1/19.8 | 18.2/19.8 |
| RMSD bond lengths (Å) | 0.005 | 0.007 |
| RMSD bond angles (°) | 0.784 | 0.927 |
| Average B, all atoms (Å$^2$) | 27.0 | 33.0 |
| No. protein residues | 124 | 123 |
| **Ramachandran analysis** | | |
| Favored (%) | 99.2 | 99.2 |
| Allowed (%) | 0.8 | 0.8 |
| Outlier (%) | 0.0 | 0.0 |

**Table S3: Crystallographic data collection and refinement statistics. Related to Figure 5.**
Statistics for our structures of the Ebola virus VP40 ring generated using the "HSP" RNA and the "HSP" DNA, respectively (Figures 5 and S5, Table S2). Numbers in parentheses represent the value for the highest-resolution shell.