An emergent Ebola virus nucleoprotein variant influences virion budding, oligomerization, transcription and replication

Aaron E Lin[1,2,3,11], William E Diehl[4,11], Yingyun Cai[5], Courtney Finch[5], Chidiebere Akusobi[6], Robert N Kirchdoerfer[7], Laura Bollinger[5], Stephen F Schaffner[2,3], Elizabeth A Brown[2,3], Erica Ollmann Saphire[7,8], Kristian G Andersen[7,9], Jens H Kuhn[5,12], Jeremy Luban[4,12], Pardis C Sabeti[1,2,3,10,12]

[1]Harvard Program in Virology, Harvard Medical School, Boston, MA 02115, USA. [2]FAS Center for Systems Biology, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. [3]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA. [4]Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA 01605, USA. [5]Integrated Research Facility at Fort Detrick, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Frederick, MD 21702. [6]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA 02120, USA. [7]Department of Immunology and Microbial Sciences, The Scripps Research Institute, La Jolla, CA 92037, USA. [8]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. [9]Scripps Translational Science Institute, La Jolla, CA 92037, USA. [10]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA. [11]Co-first author. [12]Co-last author.

Abstract

To investigate how Ebola virus phenotypes changed during the 2013–2016 Western African Ebola virus disease epidemic, we examined a key viral mutation that rose to high frequency: an R111C substitution in the viral nucleoprotein (NP). Though NP plays many essential roles during infection, there are a limited number of assays for studying these functions. We developed new reporter assays to measure virion-like particle (VLP) production and NP oligomerization in live cells under biosafety level 2 conditions. We found that NP-R111C significantly enhanced VLP production and slightly increased NP oligomerization without impairing viral transcription and replication. By contrast, a synthetic charge-reversal mutant, NP-R111E, greatly increased oligomerization but dramatically reduced transcription and replication. We detected an interaction of NP with the cellular clathrin adaptor protein-1 (AP-1) complex, which may explain how NP facilitates VLP production. Our study provides enhanced methods to study NP and indicates a complex interplay between NP's roles in virion budding, protein structure, and transcription and replication.

Introduction

Given the past and current public health threats caused by Ebola virus disease (EVD) outbreaks, rapidly evaluating whether Ebola virus (EBOV) genomic mutations change viral phenotypes is critically important. As an RNA virus, EBOV generates many mutations over the course of an outbreak. The vast majority of these mutations likely will not be adaptive and will instead have negligible or negative effects on EBOV viability and replication (Holmes 2009). Yet, changes in the EBOV genome over time can have important implications for clinical patient care, epidemiological modeling, and vaccine development, and thus influence prospective outbreak prediction and outbreak response.

The need to better understand EBOV evolution became clear during the 2013–2016 EVD epidemic in Western Africa caused by the EBOV Makona variant. This epidemic is the largest EVD epidemic on record with over 28,000 infections and more than 11,000 deaths (WHO 2016). EBOV replication generated thousands of mutations over numerous rounds of human-to-human transmission (Baize et al. 2014; Gire et al. 2014; Carroll et al. 2015; Ladner et al. 2015; Tong et al. 2015; Park et al. 2015; Simon-Loriere et al. 2015; Kugelman et al. 2015; T. Hoenen et al. 2015; Thomas Hoenen et al. 2016; Smits et al. 2015; Quick et al. 2016), but only a handful of mutations became common enough to have had a sizeable impact on the epidemic (Baize et al. 2014; Gire et al. 2014; Park et al. 2015; Carroll et al. 2015; Ladner et al. 2015; Tong et al. 2015; Simon-Loriere et al. 2015). These mutations define four distinct genetic EBOV Makona clades, which emerged as the outbreak was accelerating in May 2014. Each of the clades, termed “SL1” through “SL4,” descended from one another sequentially (e.g., SL2 derived from SL1), with a subsequent increase in EVD cases. Because non-synonymous mutations directly change protein sequence, we focused on the only two non-synonymous clade-defining mutations—one each defining the SL1 and SL2 clades.

The SL1 clade-defining mutation, C6283U, results in an A82V substitution in the EBOV glycoprotein (GP-A82V) and has been studied extensively through well-established biosafety level 2 (BSL-2) surrogate model systems and live virus BSL-4 studies. Because EBOV GP mediates EBOV particle entry into cells and is a major target for host antibodies, GP-A82V is an obvious priority for experimental studies. During outbreaks of other viral diseases, mutation in viral surface proteins such as GP has altered cell tropism (e.g., Zika virus (Yuan et al. 2017)) or host tropism (e.g., chikungunya virus (Tsetsarkin et al. 2007)). Based on multiple *in vitro* studies, GP-A82V increases EBOV infectivity for a variety of human and non-human primate cell types. These data suggest that the mutation confers a selective advantage to EBOV (Diehl et al. 2016; Urbanowicz et al. 2016; Dietzel et al. 2017; Ueda et al. 2017; Hoffmann et al. 2017; Wang et al. 2017), which however has not yet been demonstrated clearly *in vivo* (Marzi et al. 2018; Wong et al. 2018).

The SL2 clade-defining mutation, C800U, results in an R111C substitution in the EBOV nucleoprotein (NP-R111C) and has not been thoroughly studied because, though NP has many functions, there are a limited number of assays for studying these functions. While surface glycoproteins like EBOV GP alter viral tropism primarily by affecting cell susceptibility (e.g., increased viral entry), mutations in non-glycoproteins can also affect viral tropism by changing cell permissiveness (e.g., increased viral genome replication) (Cauldwell et al. 2014; Kirmaier et al. 2010; Krupp et al. 2013; Zimmermann et al. 2011). For this study, we focused our attention on EBOV NP.

The N-terminus of EBOV NP, which contains the R111 residue, interacts with viral RNA and connects multiple viral phenotypes such as virion structure, and transcription and replication. During virion assembly and budding, the EBOV matrix protein VP40 dimerizes in the cytoplasm, traffics to the cell membrane, and oligomerizes. VP40 oligomers engage with the EBOV ribonucleoprotein (RNP) complex composed of NP, polymerase cofactor VP35, transcription regulator VP30, and RNA-dependent polymerase L, and shape the cell's plasma

membrane around them into EBOV virions, which exit the cell by budding (Luke D. Jasenosky and Kawaoka 2004). Expression of VP40 in the absence of other viral proteins generates similarly shaped particles, dubbed virion-like particles (VLPs) (Harty et al. 2000; L. D. Jasenosky et al. 2001; Noda et al. 2002). Co-expression of VP40 with NP or other viral proteins significantly increases the number of VLPs in cellular supernatant (Licata et al. 2004), suggesting that NP plays a structural role in assembling and stabilizing VLPs. Mutations in NP that affect its ability to enhance VLP production are not known.

EBOV NP also plays an essential role in viral transcription and replication. By directly interacting with EBOV RNA, VP35, and VP30, NP recruits L to enact both of these essential functions (Groseth et al. 2009). Based on EBOV NP structural data, homology modeling versus other viral nucleoproteins, and site-directed mutagenesis experiments, key EBOV NP residues that interact with EBOV RNA (Dong et al. 2015; Leung et al. 2015; Kirchdoerfer et al. 2015), VP35 (Leung et al. 2015; Kirchdoerfer et al. 2015), and VP30 (Kirchdoerfer et al. 2016) have been identified. However, NP residue 111, the site of the Makona variant SL2 clade-defining mutation, lies outside of any of these annotated interaction surfaces.

Yet another property of EBOV NP is its ability to form long oligomers; these oligomers coat EBOV RNA during multiple viral life cycle events. Many studies have determined the structure of the core of NP and modeled how it interacts with itself and RNA (Dong et al. 2015; Leung et al. 2015; Kirchdoerfer et al. 2015; Wan et al. 2017; Su et al. 2018; Sugita et al. 2018). However, the oligomerization domain (OD) at the very N-terminus of NP has eluded crystallization, presumably due to its disordered structure (Su et al. 2018). Alterations to NP oligomerization could affect virion assembly and transcription and replication, but the interplay of these functions is not obvious. Aside from deletion of the oligomerization domain (ΔOD), no other mutations are known to affect EBOV NP oligomerization, in part because oligomerization is a challenging phenotype to assay in cell culture.

For this study, we have adapted and, when necessary, created new BSL-2 methods to study key functions of EBOV NP in cell culture. Such tools are critical for rapidly characterizing unknown or emerging mutants since studying live EBOV requires scarce maximum containment (BSL-4) facilities. Moreover, existing recombinant live virus systems typically use the same genetic backbone that is different from the Makona C-15 EBOV (Volchkov et al. 2001; Neumann et al. 2002; Towner et al. 2005; Thomas Hoenen et al. 2013), and generating new recombinant systems remains logistically and financially challenging due to restrictions on their use and associated synthesis costs. Using these straightforward, modular reporter assays, we found that NP-R111C increases EBOV VLP production. In addition, NP residue 111 is positioned to control NP oligomerization and viral transcription and replication, highlighting the multi-functionality of NP. Our findings support the possibility that NP-R111C is beneficial for viral replication.

Results

Ebola virus nucleoprotein mutation R111C emerged alongside a GP-A82V mutation during the 2013–2016 Western African epidemic

Among the viral mutations that rose to dominate the EBOV population during the 2013–2016 Western African EVD epidemic, NP-R111C is of great interest because it shares features with the GP-A82V mutation that enhances viral infectivity *in vitro* (Diehl et al. 2016; Urbanowicz et al. 2016; Dietzel et al. 2017; Ueda et al. 2017; Hoffmann et al. 2017; Wang et al. 2017). GP-A82V and NP-R111C are two major clade-defining mutations that rose to high frequency during the epidemic; other mutations do not affect the amino acid sequence of EBOV proteins (Gire et al. 2014; Park et al. 2015). Based on phylogeny of EBOV genomes from clinical samples, the NP-R111C mutation occurred soon after the emergence of the GP-A82V substitution (Figure 1A) and temporally preceded the inflection point of the epidemic (Figure 1B). Indeed, few EBOV Makona variant genomes encode the GP-A82V mutation in the absence of the NP-R111C mutation (23 cases, 1.26% of total), and the overwhelming majority of genomes encode both mutations (1653 cases, 90.67% of total).

Location of the Ebola virus nucleoprotein R111 residue

To investigate the functional importance of the NP 111 residue, we examined existing annotations and potential functions of NP. The R111 residue lies outside of key sites known to be interaction surfaces (i.e., for binding EBOV RNA and VP35) (Figure 1C). Based on NP crystal structures (Kirchdoerfer et al. 2015; Dong et al. 2015), the R111 residue appears on the same face of the protein as the NP oligomerization domain (Figure 1D, left, Figure S1A), opposite to the key VP35 and RNA interaction residues (Figure 1D, right). Electron microscopy (EM) subtomogram averaging indicates that R111 is proximally located to key oligomerization

residues (Figure S1A, blue) (Wan et al. 2017). Interestingly, R111 lies amidst a conserved stretch of 3 basic residues, K109/K110/R111, on the surface of the NP protein. A recent cryo-EM structure identifies K110, adjacent to R111, as a residue forming a key electrostatic interstrand NP-NP interaction (Sugita et al. 2018). Therefore, we focused on whether NP-R111C affects structural phenotypes during the EBOV life cycle, and further queried this residue by generating charge-reversed mutants (NP-R111E and NP-K109E/K110E/R111E).

Ebola virus nucleoprotein mutation R111C increases budding of virion-like particles

To determine whether the NP-R111C mutation plays a structural role in infectivity, we designed and performed a VLP budding assay. Traditionally, researchers assess viral budding efficiency by harvesting cell culture supernatants, purifying VLPs by ultracentrifugation through sucrose, and detecting VLPs by western blot (WB) using antibodies to specific VLP components (Licata et al. 2004; McCarthy, Licata, and Harty 2006; Bornholdt et al. 2013; Liu et al. 2010). However, WBs are often not sensitive to modest changes in VLP numbers and can suffer from high technical variability. By contrast, luminescence can be reproducibly detected over a larger linear dynamic range, so one would prefer to generate luciferase-fused VLPs for cell-based expression, purification, and luminescence detection instead of WBs. However, the size of firefly luciferase (FLuc; 60 kDa) can severely interfere with incorporation into budding VLPs. Indeed, although the EBOV matrix protein VP40 (40 kDa) alone is sufficient to bud VLPs (Harty et al. 2000; L. D. Jasenosky et al. 2001; Noda et al. 2002), fusion of VP40 to FLuc decreased luciferase activity to undetectable levels in a budding assay (McCarthy, Licata, and Harty 2006). Here, we took advantage of the smaller size of NanoLuc (NLuc; 19 kDa) (Hall et al. 2012) and fused it to VP40. We expressed NLuc-VP40 in cell culture, purified VLPs following established protocols, and measured NLuc reporter activity (Figure 2A).

To verify that our assay was truly measuring VLP production, we generated loss-of-function (LOF) mutants, measured thermal stability, and visualized VLPs via electron microscopy (EM). We generated VP40-L117R, an LOF mutant that was defective in dimerization, membrane trafficking, and VLP budding as judged by immunofluorescence microscopy and WB (Bornholdt et al. 2013). As expected, budding of this mutant was >400-fold impaired compared to wild-type VP40-L117 in our NLuc-based VLP assay (adj. $p < 0.002$; Dunnett's test) (Figure 2B). Additionally, since monomeric VP40 can be expelled from cells in exosomes (Pleet et al. 2016), we tested whether ultracentrifugation was purifying VP40-VLPs specifically. As NLuc (Hall et al. 2012) has a higher melting point than VP40 VLPs (Hu et al. 2011), we heated the total supernatant to denature VLPs. Compared to heating supernatant to 60.2 °C alone, heating and subsequent ultracentrifugation reduced NLuc activity 15-fold ($p < 0.007$; paired t-test). This reduction in NLuc activity suggests that NLuc-VP40-bearing VLPs were denatured and thus were not pelleted and detected (Figure S2A). Lastly, using EM, we observed VLPs in culture supernatant of cells expressing VP40 and NP-R111 or NP-R111C (Figure S2B), which appeared similar in size and volume (Figure S2C).

Next, we tested the NP-R111 mutants in our VLP budding assay and found that only NP-R111C improved VLP production. Expression of viral nucleoproteins, including EBOV, is known to significantly increase matrix protein-induced VLP production (Licata et al. 2004). We verified that ancestral NP-R111 expression increased NLuc-VP40 VLP production 1.93-fold compared to enhanced green fluorescent protein (eGFP) control inserted in place of NP-R111 (adj. $p < 0.0002$, Dunnett's test) (Figure 2C). NP-R111C significantly increased VLP production above NP-R111 (1.26-fold; adj. $p < 0.039$, Dunnett's test), whereas the charge-reversed NP-R111E (1.07-fold; adj. $p < 0.847$, Dunnett's test) and NP-K109E/K110E/R111E (1.13-fold; adj. $p < 0.484$, Dunnett's test) did not have a reproducible effects.

Ebola virus nucleoprotein position 111 significantly affects oligomerization of NP

To illuminate why NP-R111C increases VLP production, we developed an assay to measure intracellular NP oligomerization using bioluminescence resonance energy transfer (BRET). Traditional oligomerization assays in cell culture involve tagging a protein separately with two different tags, co-expressing both, and then co-immunoprecipitation (co-IP) targeting one tag and WB targeting the other tag (Watanabe, Noda, and Kawaoka 2006; Ng et al. 2012; Ortiz-Riano et al. 2012). However, WB often has linear dynamic range issues; furthermore, co-IPs can introduce non-specific or spurious protein-protein interactions under different cell lysis and binding buffers. To overcome these deficits of co-IPs and WBs, we used BRET to study NP oligomerization in live cells. We tagged NP with either NLuc or HaloTag (which covalently binds to an acceptor fluorophore), co-expressed the tagged NPs in cells, and activated the NP-NLuc with substrate, resulting in emission of light at 465 nm. Spatial proximity of NP-NLuc to NP-HaloTag due to NP oligomerization results in energy transfer and a second light emission at a longer wavelength, 625 nm (Figure 3A) (Machleidt et al. 2015).

To verify that our assay was truly measuring NP oligomerization, we generated NP LOF mutants and disrupted oligomerization with biologically relevant EBOV VP35. We first generated NP-ΔOD, an LOF mutant that size exclusion chromatography and multiangle light scattering indicate to be defective in oligomerization (Kirchdoerfer et al. 2015). Then, we confirmed that the mutant lost oligomerization capability using the traditional dual-tag co-IP-WB strategy (Figure S3). We then performed our BRET assay in live cells and, as expected, the lack of NP-NLuc or NP-HaloTag, expression of NP-ΔOD, or free NLuc reduced BRET signal appreciably (Figure 3B). To confirm our assay in a biologically relevant context, we additionally expressed the NP-binding peptide (NPBP) of EBOV VP35, which disrupts NP oligomerization (Kirchdoerfer et al. 2015). To quantitatively detect VP35 NPBP expression, we fused eGFP to NPBP via a bridging porcine teschovirus 1 2A 'self-cleaving' peptide (Kim et al. 2011) (eGFP-P2A-

VP35(NPBP)). Titrating increasing amounts of VP35 NPBP led to a quantitative decrease in BRET oligomerization signal (residual ÷ total sum of squares = 0.95) (Figure 3C).

Next, we measured the propensity of NP-R111 variants to form oligomers and found that NP-R111C, and to an even greater extent NP-R111E and NP-K109E/K110E/R111E, increased NP oligomerization. To quantify changes in oligomerization, we titrated an increasing concentration of acceptor NP-HaloTag to saturate the donor NP-NLuc signal. The resulting binding curves fit well to Michaelis-Menten kinetics, described by the parameters v_{max} (maximum signal) and K_m (concentration of NP-HaloTag needed to reach half v_{max}) (Figure 3D). As expected, control eGFP substituted for NP resulted in no detectable BRET signal, and NP-ΔOD-HaloTag resulted in background signal detected only at high concentrations (Figure 3D). Relative to NP-R111, NP-R111C increased oligomerization slightly (12% lower K_m; adj. $p < 0.003$, Dunnett's test), whereas the charge-reversed NP-R111E (36% lower K_m; adj. $p < 0.0001$, Dunnett's test) and NP-K109E/K110E/R111E (28% lower K_m; adj. $p < 0.0001$, Dunnett's test) mutants oligomerized at even lower concentrations (Figure 3D). These results support our hypothesis that the NP 111 residue affects NP-NP interactions (Figure 1D, S1A). Indeed, deuterium exchange mass spectrometry indicates that this residue is partially buried in wild-type NP compared to an oligomerization-incompetent NP (Su et al. 2018). Moreover, cryo-EM identifies the adjacent residue, NP-K110, as forming a key electrostatic interaction with NP-E349 on a neighboring NP molecule (Sugita et al. 2018).

Ebola virus nucleoprotein interacts with the AP-1 clathrin adaptor complex independent of the nucleoprotein residue 111 allele

As NP-NP interactions appeared to be affected by the NP 111 residue, we searched for host binding proteins that may explain how NP enhances VLP budding. Only a single interactome study has been performed on EBOV NP fused to eGFP (García-Dorival et al. 2016), which

utilized the NP amino acid sequence from the Mayinga isolate of the EBOV Yambuku variant, the first EBOV isolated in 1976.

To build upon these previous results, we performed immunoprecipitation tandem mass spectrometry (IP-MS/MS) using myc-tagged NP from EBOV/Makona bearing either R111 or NP-K109E/K110E/R111E. Our approached yielded multiple members of the adaptor related protein 1 (AP-1) complex as strong candidate interactors (Figure 4A, right), which were identified previously (García-Dorival et al. 2016) but were not further confirmed in that study. Here, we confirm that both NP-R111 and NP-K109E/K110E/R111E strongly interact with AP-1 subunit M1 (AP1M1) and AP1G1 by reciprocal IP-WB (Figure 4B). Yet, NP-R111C, NP-R11E, and NP-K109E/K110E/R111E all bind to the AP-1 complex with similar affinity as NP-R111 **(Figure 4C)**, suggesting that the AP-1 interaction does not explain why the epidemic mutation NP-R111C produced more VLPs than the ancestral NP-R111.

Ebola virus nucleoprotein position 111 influences viral transcription and replication

The mechanism by which changes in NP's structural phenotypes (budding, oligomerization) affect viral transcription and replication is not obvious because NP is highly multi-functional. We quantified viral transcription and replication using a minigenome reporter assay (Luke D. Jasenosky, Neumann, and Kawaoka 2010). In this assay, we express the components of the EBOV RNP complex (NP, VP35, VP30, and L) in the presence of a 'minigenome' consisting of a reporter FLuc-encoding gene flanked by the EBOV promoter-like genomic leader and trailer sequences. Transcription is essential for minimal FLuc activity; replication is further required to achieve maximum signal (T. Hoenen et al. 2010).

Intriguingly, we found that NP-R111C caused similar transcription and replication activity as NP-R111, whereas the charge-reversal NP-R111E abrogated these activities. As expected,

absence of VP30, L, or the minigenome resulted in <5% normalized luminescence compared to cells expressing the minigenome and the entire RNP complex with NP-R111 (Figure 5). Substitution of NP-R111C in place of NP-R111 yielded similar activity (99%). On the other hand, the charge-reversal mutants NP-R111E (23% reporter activity; adj. $p < 0.003$; Dunnett's test) and NP-K109E/K110E/R111E (44% reporter activity; adj. $p < 0.017$; Dunnett's test) greatly attenuate transcription and/or replication. These results indicate that the NP 111 residue is connected to both the structural and transactivation roles of EBOV NP.

Discussion

Here, we developed and modified BSL-2 assays to study in-depth a key EBOV NP mutant, NP-R111C, which arose during the 2013–2016 Western African EVD epidemic. Though the NP-R111 residue has not been previously annotated as functional, NP-R111C increases VLP production and NP oligomerization, and the charge-reversal mutation NP-R111E dramatically increases NP oligomerization while hindering viral transcription and replication.

Many viral proteins are highly multi-functional, making study of individual mutations challenging without high-throughput, robust assays that are sensitive to subtle changes in viral phenotype. Since luciferase-based reporter systems fit the aforementioned requirements, we took advantage of these systems to develop VLP detection assays and BRET assays for NP oligomerization. As new assays always require thorough testing, we have verified that luciferase activity in these assays indeed reflects phenotype using LOF mutations (NLuc-VP40-L117R for VLP production, NP-ΔOD to assess oligomerization) and biologically relevant disruptions (heating to denature VLPs, VP35 expression to reduce NP oligomerization). These BSL-2 assays are simple and flexible for testing new viral mutations as they emerge during epidemics. With more rigorous screening and quantification of key metrics of variability, like Z-factor, these assays could potentially be used for high-throughput screens of hundreds of EBOV NP mutants, interactions with host factors, or antagonism by drug candidates.

Although these reporter assays show that NP-R111C increases VLP production, the mechanism behind this increase remains unclear. In this study, we used co-IP-MS/MS to identify many new putative NP binding partners, since there has only been a single previous interactome study on EBOV NP (García-Dorival et al. 2016). We confirmed an interaction with the clathrin adaptor AP-1 complex (García-Dorival et al. 2016), which is of particular interest. In retroviruses, the Gag protein facilitates budding by hijacking the AP-1 complex (Camus et al. 2007). Because retroviral and EBOV particles bud using the same cellular pathway and

machinery (Liu et al. 2010), we hypothesize that EBOV NP also co-opts the AP-1 complex for virion egress and trafficking. This interaction with the AP-1 complex could explain why expression of NP significantly enhances budding of EBOV VLPs (Licata et al. 2004). However, NP-R111 and NP-R111C bound to AP-1 with similar affinity, suggesting that this interaction was not the reason behind NP-R111C's increased ability to promote VLP production. To identify protein-protein interactions in this study, we used a standard label-free co-IP-MS/MS approach. To gain deeper mechanistic insight into this complex NP surface, researchers of future studies could apply more powerful label-based quantification approaches to identify differential interacting partners between NP-R111, NP-R111C, and NP-R111E.

Charge reversal at the NP residue R111, NP-R111E, further demonstrates the importance of this position to multiple essential viral life cycle functions. We show that NP-R111E is not aberrantly misfolded and degraded since it expresses similarly as NP-R111 and NP-R111C. In fact, we find that NP-R111E oligomerizes at significantly lower protein concentration compared to NP-R111, yet NP-R111E is unable to support normal levels of viral replication and transcription. Basicity at residues K109, K110, and R111 is highly conserved among nearly all immediate relatives of EBOV in the genus *Ebolavirus*, including the newly discovered Bombali virus (Goldstein et al. 2018). This high degree of conservation, despite significant evolutionary divergence between ebolaviruses, emphasizes the importance of this highly basic region to NP functions.

Indeed, a recent cryo-EM structure identified K110 as forming an electrostatic interaction with E349 on a neighboring NP molecule (Sugita et al. 2018), but it it not immediately apparent how changing the NP-R111 residue affects multiple viral phenotypes. The triple charge-reversal mutant NP-K109E/K110E/R111E, which presumably disrupts the K110-E349 electrostatic interaction, increases oligomerization while ablating minigenome transcription and replication. Because NP-R111E phenocopies the triple charge-reversal mutant, it is possible that NP-R111E disrupts the K110-E349 interaction as well. Intriguingly, the epidemic substitution NP-

R111C slightly increases oligomerization compared to NP-R111, but to a lesser extent than NP-R111E. Biochemically, 14% of cysteine residues (pK_a = 8.18) (Nelson, Lehninger, and Cox 2008) will be negatively charged at a typical intracellular pH of 7.4, in between lysine (100% positively charged) and arginine (100% negatively charged). Yet, NP-R111 and NP-R111C produce similar levels of minigenome transcription and replication, while NP-R111E and NP-K109E/K110E/R111E are significantly ablated. This correlation between side chain charge, oligomerization, and transcription and replication hints at the possibility of subtle shifts in the NP structure and/or additional electrostatic interactions that coordinate NP's ability to influence all these different functions simultaneously.

Findings from reporter-based assays are chiefly limited because such assays must still be supported by live virus experiments. Viral proteins like NP have multiple essential and accessory roles during infection, and despite our best efforts, we did not assay every function. There are innumerable molecular phenotypes that could not all be assayed (e.g., alterations to NP protein structure, localization, interaction binding sites, immune epitopes). Even for the functions studied here (VLP production, NP oligomerization, viral transcription/replication), predicting the effect of a mutation in an authentic live virus setting is challenging because active viral replication alters cellular pathways, and antiviral and other host pathways respond to infection. Here, we prioritized testing those phenotypes most likely to be affected based on the location of the R111 residue in the available NP crystal structures.

Testing whether mutations affect fitness with live virus experiments brings additional challenges because different viral stocks and cultured cells versus animal models can cause discrepancies in results. For example, the SL1 clade-defining mutation, GP-A82V, has been shown numerous times to enhance EBOV infectivity in cell culture (Diehl et al. 2016; Urbanowicz et al. 2016; Dietzel et al. 2017; Ueda et al. 2017; Hoffmann et al. 2017; Wang et al. 2017), using multiple EBOV surrogate systems (EBOV VLPs, retroviral particles pseudotyped with EBOV GP, and recombinant live virus (Dietzel et al. 2017)) and multiple cell types (e.g.,

human monocyte-derived dendritic cells (Diehl et al. 2016)). However, results from a recent study using immunocompromised laboratory mice and non-human primates (Marzi et al. 2018) indicated that EBOV Makona viral isolates encoding GP-A82V lead to modestly decreased viral load compared those without the GP-A82V mutation. A second recent study found that GP-A82V may induce slightly more morbidity and mortality in immunocompromised mice but not in ferrets (Wong et al. 2018). Findings could differ between studies because (Marzi et al. 2018) use clinical EBOV isolates, which contain multiple non-GP-A82V mutations in the reference sequence as well as dozens of minor allele mutations. By contrast, other researchers (Diehl et al. 2016; Urbanowicz et al. 2016; Dietzel et al. 2017; Ueda et al. 2017; Hoffmann et al. 2017; Wang et al. 2017) utilize monoclonal plasmids to drive EBOV gene expression, so only the desired mutation is present. Marzi and colleagues (Marzi et al. 2018) did not observe that EBOV viral isolates encoding GP-A82V replicate more efficiently in cell culture than those without the mutation, in direct contrast to another study (Dietzel et al. 2017) which used live EBOV generated from recombinant DNA plasmids. Additional mutations in the EBOV viral isolates or differing protocols may explain the discrepancy between the two studies.

Limited recombinant live EBOV studies have been performed to test the impact of the NP-R111C substitution specifically. (Dietzel et al. 2017) showed that EBOV with GP-A82V and NP-R111C outcompetes an ancestral EBOV Makona in a head-to-head format in cultured cells, but did not measure the impact of NP-R111C alone. (Wong et al. 2018) found that NP-R111C alone increases viral replication in cell culture, decreases morbidity and mortality in immunocompromised mice, and does not differ from ancestral EBOV Makona in ferrets.

The most direct way to test whether a viral mutation has functional consequence would be to generate a recombinant live EBOV bearing the mutation of interest, and to then infect non-human primates in an animal BSL-4 setting (Marzi et al. 2018; Basler 2017). However, the vast majority of recombinant virus backbones are based on the EBOV Yambuku-Mayinga isolate (Volchkov et al. 2001; Neumann et al. 2002; Towner et al. 2005; Thomas Hoenen et al. 2013),

which differs from the desired EBOV Makona-C15 reference sequence by hundreds of mutations. Even an existing recombinant EBOV Makona virus backbone (Albariño et al. 2015) would require 7 or 8 back-mutations to achieve the Makona-C15 reference sequence. Future improvements in large plasmid assembly methodology (Gibson et al. 2009) will greatly facilitate the ability to rapidly generate new recombinant DNA plasmid sequence as novel EBOV variants emerge. Given current limitations and restrictions of BSL-4 settings and the severe impact and potential global threat of EVD epidemics, using BSL-2 model systems allows initial and more rapid and comprehensive exploration of any potentially consequential EBOV mutations.

Our finding of the importance of the R111 residue and experimental systems established represent steps towards characterizing key EBOV mutations that arose during the 2013–2016 Western African EVD epidemic. These findings provide additional insight into the interplay between the many functions of NP in viral assembly and budding, oligomerization, and transcription and replication.

Materials and methods

Key resources table

Reagent type (species) or source	Designation	Source or reference	Identifiers	Additional information
cell line (H. sapiens)	Human embryonic kideny (HEK) 293FT	Thermo Fisher Scientific	R70007	
cell line (H. sapiens)	HEK 293	ATCC	CRL-1573	
antibody	mouse α-myc, clone 9B11	Cell Signaling Technologies	2276	
antibody	mouse α-V5, clone SV5-Pk1	Bio-Rad Laboratories	MCA1360	
antibody	mouse α-AP1G1, clone 100/3	Sigma-Aldrich	A4200-.2ML	
antibody	mouse α-VPS35, clone B-5	Santa Cruz Biotechnology	sc-374372	
antibody	mouse α-tubulin, clone B-7	Santa Cruz Biotechnology	sc-5286	
antibody	normal mouse IgG	Santa Cruz Biotechnology	sc-2025	
antibody	Peroxidase AffiniPure Goat Anti-Mouse IgG, Light Chain Specific	Jackson ImmunoResearch	115-035-174	
antibody	rabbit α-V5, clone D3H8Q	Cell Signaling Technologies	13202	
antibody	rabbit α-AP1M1, polyclonal	Proteintech	12112-1-AP	
antibody	rabbit α-VPS26, clone EPR13456	Abcam	ab181352	

antibody	normal rabbit IgG	Santa Cruz Biotechnology	sc-2027	
antibody	Peroxidase AffiniPure Goat Anti-Rabbit IgG (H+L)	Jackson ImmunoResearch	111-035-144	
recombinant DNA reagent	pGL4.23-CMV/NP-V5	this paper		
recombinant DNA reagent	pGL4.23-CMV/NP-myc	this paper		
recombinant DNA reagent	pGL4.23-CMV/eGFP-V5	this paper		
recombinant DNA reagent	pcDNA3.1(+)/Bla-VP40	BEI Resources, (Manicassamy and Rong 2009)	NR-19813	
recombinant DNA reagent	pcDNA3.1(+)/NLuc-VP40	this paper		
recombinant DNA reagent	pNL1.1/NLuc	Promega	N1001	
recombinant DNA reagent	pcDNA3.3/KLF4	Addgene (Derrick Rossi), (Warren et al. 2010)	26815	
recombinant DNA reagent	pLV-WPRE/mCherry	Addgene (Pantelis Tsoulfas)	36084	
recombinant DNA reagent	pcDNA3.3/eGFP-V5	this paper		
recombinant DNA reagent	pGL4.23-CMV/NP	this paper		
recombinant DNA reagent	pGL4.23-CMV/GP-A82V	(Diehl et al. 2016)		
recombinant DNA reagent	pGL4.23-CMV/NP-NLuc	this paper		
recombinant DNA reagent	pGL4.23-CMV/NP-HaloTag	this paper		
recombinant	pcDNA3.3-	this paper		

DNA reagent	WPRE/eGFP-P2A-VP35(NPBP)-V5			
recombinant DNA reagent	mTagBFP2-pBAD	Addgene (Michael Davidson), (Subach et al. 2011)	54572	
recombinant DNA reagent	pcDNA3.3-WPRE/mTagBFP2-V5	this paper		
recombinant DNA reagent	pCAGGS/L	(Luke D. Jasenosky, Neumann, and Kawaoka 2010)		
recombinant DNA reagent	pCAGGS/VP30	(Luke D. Jasenosky, Neumann, and Kawaoka 2010)		
recombinant DNA reagent	pCAGGS/T7pol	(Luke D. Jasenosky, Neumann, and Kawaoka 2010)		
recombinant DNA reagent	pCAGGS/FLuc	(Luke D. Jasenosky, Neumann, and Kawaoka 2010)		
recombinant DNA reagent	pCAGGS/RLuc	(Luke D. Jasenosky, Neumann, and Kawaoka 2010)		
recombinant DNA reagent	pCAGGS/NP-2A-VP35	(Luke D. Jasenosky, Neumann, and Kawaoka 2010)		
sequence-based reagent	NP-R111-NTerm	Integrated DNA Technologies		GBlock with N-terminus of EBOV/Mak-C15 NP
sequence-based reagent	NP-R111-CTerm	Integrated DNA Technologies		GBlock with C-terminus of EBOV/Mak-C15

				NP
sequence-based reagent	NP-R111C-fwd	Integrated DNA Technologies	CGTTTTGAAGTCAAGAAGTGTGATGGAGTGAAGCGCC	SDM primer to generate NP-R111C
sequence-based reagent	NP-R111C-rev	Integrated DNA Technologies	GGCGCTTCACTCCATCACACTTCTTGACTTCAAAACG	SDM primer to generate NP-R111C
sequence-based reagent	NP-R111E-fwd	Integrated DNA Technologies	GGTTCCGTTTTGAAGTCAAGAAGGAGGATGGAGTGAAGCGCCTTG	SDM primer to generate NP-R111E
sequence-based reagent	NP-R111E-rev	Integrated DNA Technologies	CAAGGCGCTTCACTCCATCCTCCTTCTTGACTTCAAAACGGAACC	SDM primer to generate NP-R111E
sequence-based reagent	NP-K109E-K110E-fwd	Integrated DNA Technologies	GGGTTCCGTTTTGAAGTCGAGGAGGAGGATGGAGTGAAGC	SDM primer to generate NP-K109E & -K110E
sequence-based reagent	NP-K109E-K110E-fwd	Integrated DNA Technologies	GCTTCACTCCATCCTCCTCCTCGACTTCAAAACGGAACCC	SDM primer to generate NP-K109E & -K110E
sequence-based reagent	NP-dOD-fwd	Integrated DNA Technologies	CGAGTCTCACTGAATCTGACATGAGAGTCATCCCAGTGTATC	SDM primer to generate NP-ΔOD
sequence-based reagent	NP-dOD-rev	Integrated DNA Technologies	GATACACTGGGATGACTCTCATGTCAGATTCAGTGAGACTCG	SDM primer to generate NP-ΔOD
sequence-based reagent	NP-dC50-fwd	Integrated DNA Technologies	GGTAAAGAGTACACGTATCCGTCTAGAGGTAAGCCTATCCC	SDM primer to generate NP-ΔC50
sequence-based	NP-dC50-rev	Integrated DNA	GGGATAGGCT	SDM primer to

reagent		Technologies	TACCTCTAGAC GGATACGTGT ACTCTTTACC	generate NP-ΔC50
sequence-based reagent	BspEI_eGFP-fwd	Integrated DNA Technologies	GATGCTTCCG GACCACCATG GTGAGCAAGG GCGA	PCR primer to generate eGFP with restriction enzyme sites as insert
sequence-based reagent	XbaI_eGFP-rev	Integrated DNA Technologies	TGTCCTTCTAG ACTTGTACAGC TCGTCCATGC	PCR primer to generate eGFP with restriction enzyme sites as insert
sequence-based reagent	EBOV-Mak-C15-VP40	Integrated DNA Technologies		GBlock with EBOV/Mak-C15 to replace EBOV/Yam-May in original vector
sequence-based reagent	VP40-L117R-fwd	Integrated DNA Technologies	CTACGGCCGC CATCATGCGT GCTTCATATAC TATCACCC	SDM primer to generate VP40-L117R
sequence-based reagent	VP40-L117R-rev	Integrated DNA Technologies	GGGTGATAGT ATATGAAGCAC GCATGATGGC GGCCGTAG	SDM primer to generate VP40-L117R
sequence-based reagent	NP-GSGGGSGGGT-rev	Integrated DNA Technologies	TCCTCCAGATC CTCCTCCGGA TCCCTGATGAT GTTGCAGGAT TGC	Inversion PCR primer to generate linear pGL4.23-CMV-NP as vector for Gibson assembly
sequence-based reagent	pGL423-CTerm-fwd	Integrated DNA Technologies	TAATCTAGAGT CGGGGCGG	Inversion PCR primer to generate linear pGL4.23-CMV-NP as vector for Gibson assembly
sequence-based reagent	GSGGGSGGGT-HaloTag-fwd	Integrated DNA Technologies	GGAGGAGGAT CTGGAGGAGG TACCGAAATC GGTACTGGCT	PCR primer to generate linear HaloTag with Gly-rich spacer

			TTCC	as insert for Gibson assembly
sequence-based reagent	pGL423-HaloTag-rev	Integrated DNA Technologies	CCGCCCCGAC TCTAGATTAAC CGGAAATCTC CAGAGT	PCR primer to generate linear HaloTag with Gly-rich spacer as insert for Gibson assembly
sequence-based reagent	GSGGGSGGGT -nLuc-fwd	Integrated DNA Technologies	GGAGGAGGAT CTGGAGGAGG TACCATGGTCT TCACACTCGAA GA	PCR primer to generate linear NLuc with Gly-rich spacer as insert for Gibson assembly
sequence-based reagent	pGL423-nLuc-rev	Integrated DNA Technologies	CCGCCCCGAC TCTAGATTACG CCAGAATGCG TTCG	PCR primer to generate linear NLuc with Gly-rich spacer as insert for Gibson assembly
sequence-based reagent	pcDNA-EGFP-fwd	Integrated DNA Technologies	GACCGATCCA GCCTCCACCA TGGTGAGCAA GG	PCR primer to generate linear EGFP with P2A peptide as insert for Gibson assembly
sequence-based reagent	P2A-EGFP-rev	Integrated DNA Technologies	AGGTCCAGGG TTCTCCTCCAC GTCTCCAGCC TGCTTCAGCA GGCTGAAGTT AGTAGCTCCG CTTCCCTTGTA CAGCTCGTCC ATGC	PCR primer to generate linear EGFP with P2A peptide as insert for Gibson assembly
sequence-based reagent	pGL423-N-P2A-rev	Integrated DNA Technologies	GCCCTTTGTTC TAGTTGTCATA GGTCCAGGGT TCTCCTC	PCR primer to generate linear EGFP with P2A peptide as insert for Gibson assembly
sequence-based	EBOV-Mak-C15-	Integrated DNA	GCCCTTTGTTC	GBlock with full-

reagent	VP35	Technologies	TAGTTGTCATAGGTCCAGGGTTCTCCTC	length EBOV/Mak-C15 VP35
sequence-based reagent	VP35-fwd	Integrated DNA Technologies	ATGACAACTAGAACAAAGGGCA	PCR primer to generate linear VP35(NPBP) as insert for Gibson assembly
sequence-based reagent	V5-VP35-80aa-rev	Integrated DNA Technologies	GTTAGGGATAGGCTTACCTTATTAATTGCAAATTGGGTCCGTTT	PCR primer to generate linear VP35(NPBP) as insert for Gibson assembly
sequence-based reagent	BspEI_mTagBFP2-fwd	Integrated DNA Technologies	GATGTCTCCGGACCACCATGGTGTCTAAGGGCGAAGA	PCR primer to generate mTagBFP2 with restriction enzyme sites as insert
sequence-based reagent	XbaI_mTagBFP2-rev	Integrated DNA Technologies	TGTCCATCTAGAATTAAGCTTGTGCCCCAGTT	PCR primer to generate mTagBFP2 with restriction enzyme sites as insert
peptide, recombinant protein	SureBeads Protein A Magnetic Beads, 3 ml	Bio-Rad Laboratories	1614013	
peptide, recombinant protein	SureBeads Protein G Magnetic Beads, 3 ml	Bio-Rad Laboratories	1614023	
peptide, recombinant protein	Protein A/G PLUS-Agarose	Santa Cruz Biotechnology	sc-2003	
peptide, recombinant protein	Phusion High-Fidelity PCR Master Mix with GC Buffer	New England Biolabs	M0532L	
commercial assay or kit	Nano-Glo Luciferase Assay System	Promega	N1120	

commercial assay or kit	NanoBRET Nano-Glo Detection System	Promega	N1662	
commercial assay or kit	Dual-Luciferase Reporter Assay System	Promega	E1980	
chemical compound, drug	DMEM, high glucose, GlutaMAX Supplement, HEPES	Thermo Fisher Scientific	10564029	
chemical compound, drug	Fetal Bovine Serum, certified, One Shot format, US origin	Thermo Fisher Scientific	A3160402	
chemical compound, drug	Penicillin-Streptomycin	Thermo Fisher Scientific	15140122	
chemical compound, drug	MEM Non-Essential Amino Acids Solution (100X)	Thermo Fisher Scientific	11140050	
chemical compound, drug	Sodium Pyruvate (100mM)	Thermo Fisher Scientific	11360070	
chemical compound, drug	Opti-MEM I Reduced Serum Media	Thermo Fisher Scientific	31985062	
chemical compound, drug	Opti-MEM I Reduced Serum Medium, no phenol red	Thermo Fisher Scientific	11058021	
chemical compound, drug	Lipofectamine 2000 Transfection Reagent	Thermo Fisher Scientific	11668019	
chemical compound, drug	TransIT-LT1	Mirus Bio	MIR 2300	
chemical compound, drug	PBS, pH 7.4	Thermo Fisher Scientific	10010049	

chemical compound, drug	Sucrose	Sigma-Aldrich	84097-250G	
chemical compound, drug	Polyvinylpyrrolidone	Sigma-Aldrich	234257-5G	
chemical compound, drug	Dimethyl sulfoxide, ReagentPlus, ≥99.5%	Sigma-Aldrich	D5879-100ML	
chemical compound, drug	PageBlue Protein Staining Solution	Thermo Fisher Scientific	24620	
chemical compound, drug	2x Laemmli Sample Buffer	Bio-Rad Laboratories	1610737	
chemical compound, drug	SuperSignal West Pico Chemiluminescent Substrate	Thermo Fisher Scientific	34078	
software, algorithm	Virus Pathogen Database and Analysis Resource (ViPR)	(Pickett et al. 2012)		Accessed October 2017
software, algorithm	MAFFT	(Katoh and Standley 2013)	v6.902b	
software, algorithm	trimAl	(Capella-Gutiérrez, Silla-Martínez, and Gabaldón 2009)	v1.4	
software, algorithm	RAxML	(Stamatakis, Ludwig, and Meier 2005)	v7.3.0	
software, algorithm	PyMOL	Schrödinger, (Schrödinger, LLC 2015)	v2.0.3	
software, algorithm	ImageJ	(Schneider, Rasband, and Eliceiri 2012)		
software, algorithm	SEQEST	Thermo Fisher Scientific, (Eng, McCormack, and Yates 1994)		

software, algorithm	Contaminant Repository for Affinity Purification (CRAPome)	(Mellacheruvu et al. 2013)	v1.1	Accessed November 2015
software, algorithm	Search Tool for the Retrieval of Interacting Genes/proteins (STRING)	(Snel 2000; Szklarczyk et al. 2015)	v10	Accessed November 2015
software, algorithm	Cytoscape	(Shannon et al. 2003)	v3.4.0	
software, algorithm	Prism	GraphPad Software	v7.0c	
software, algorithm	R	(R Core Team 2016)	v3.3.1	
software, algorithm	nlstools' package in R	(Baty et al. 2015)	v1.0-2	
software, algorithm	ggplot2' package in R	(Wickham 2016)	v2.2.1	
other	0.45 µm, Acrodisc Syringe Filter with HT Tuffryn Membrane	Pall Laboratory	4184	
other	BioCoat Poly-D-Lysine 6-well Clear Flat Bottom TC-treated Multiwell Plate	Corning	356413	
other	96-well Black Flat Bottom Polystyrene NBS Microplate	Corning	3991	
other	BioCoat Poly-D-Lysine 96-well Black/Clear Flat Bottom TC-treated Microplate	Corning	356640	

other	4–15% Mini-PROTEAN TGX Precast Protein Gels, 10-well, 30 µL	Bio-Rad Laboratories	4561083	

Ebola virus genome sequences and phylogenetic analysis

We obtained Ebola virus (EBOV) genomes from the US National Institute of Allergy and Infectious Diseases (NIAID) Virus Pathogen Database and Analysis Resource (ViPR) through the web site at http://www.viprbrc.org/ (Pickett et al. 2012) on October 2017. We removed short sequences, sequences from tissue-cultured EBOV isolates, duplicate sequences from the same clinical EVD case, and sequences with >0.2% ambiguous or missing nucleotide calls. The final dataset consisted of 1,823 EBOV complete or near-complete genomes.

We aligned these genomes with MAFFT v6.902b (Katoh and Standley 2013) using the parameters (L-INS-i):–localpair–maxiterate 1000–reorder–ep 0.123. We trimmed the alignment using trimAl v1.4 (Capella-Gutiérrez, Silla-Martínez, and Gabaldón 2009) with -automated1. Lastly, we generated a maximum likelihood tree with RAxML v7.3.0 (Stamatakis, Ludwig, and Meier 2005) under a generalized time-reversible (GTRɣ) nucleotide substitution model with 100 bootstrap pseudoreplicates.

For functional characterization, we used the genome sequence of Ebola virus/H.sapiens-wt/GIN/2014/Makona-C15 (EBOV/Mak-C15; GenBank #KJ660346.2; *Filoviridae: Zaire ebolavirus*) as the EBOV Makona variant reference sequence for NP, VP40, VP35, and GP analyses and cloning, unless otherwise noted. The structural analysis of EBOV NP was based on the Ebola virus/H.sapiens-tc/COD/1976/Yambuku-Mayinga NP (EBOV/Yam-May) crystal structure under Protein Data Bank (PDB) #4YPI (Leung et al. 2015), with manual annotation of key residues based on results of other studies (Leung et al. 2015; Kirchdoerfer et al. 2015). The structure of EBOV/Mak-C15 NP has not yet been elucidated, but the amino acid sequence is 98% identical (14 mutations / 739 residues) to EBOV/Yam-May NP, and the N-terminal 450

amino acids of the two variants are 99.3% identical (3 mutations / 450 residues). The subtomogram averaged electron microscopy (EM) NP structure (also derived from EBOV/Yam-May) was accessed from PDB #6EHL (Wan et al. 2017). We visualized all structures using PyMOL (Schrödinger, New York City, NY) (Schrödinger, LLC 2015).

Constructs and cloning

We performed all assays with the same mammalian expression vector for EBOV NP and its mutants. We synthesized EBOV NP-R111 in 2 dsDNA gBlocks (Integrated DNA Technologies [IDT], Coralville, IA) and cloned these into pGL4.23-CMV (described in (Diehl et al. 2016)) modified with a C-terminal V5 peptide tag. To generate all NP mutants, we performed a modified site-directed mutagenesis (SDM) protocol on this plasmid, as described in (Diehl et al. 2016). For many assays, we expressed enhanced green fluorescent protein (eGFP) in place of NP as a negative control; we generated the corresponding vector by cloning eGFP into pcDNA3.3-CMV (Thermo Fisher Scientific, Waltham, MA) modified by a sequence encoding an in-frame C-terminal V5 peptide tag.

For the VLP budding assay, we additionally constructed a plasmid to express a NanoLuc-EBOV Makona-C15 VP40 (NLuc-VP40) fusion protein. To create this plasmid, we obtained a pcDNA3.1(+)-based vector expressing β-lactamase (Bla) fused to EBOV/Yam-May (Bla-VP40) through the US National Institutes of Health (NIH)/NIAID Biodefense and Emerging Infections Research Resources Repository (BEI Resources, Manassas, VA; #NR-19813) (Manicassamy and Rong 2009). We replaced the *Bla* gene with the gene encoding NanoLuc (NLuc) from pNL1.1 (Promega, Madison, WI), and replaced the EBOV/Yam-May VP40 sequence with that of EBOV/Mak-C15 from a gBlock (IDT). pNL1.1, which expresses NLuc alone without VP40, was used as a negative control. As an additional negative control, we performed SDM to introduce a VP40 L117R mutation into the NLuc-VP40-encoding vector based on loss-of-function (LOF) induced by this mutation as reported in (Bornholdt et al. 2013).

For EM experiments, we additionally co-expressed EBOV glycoprotein (GP) from a pGL4.23-CMV vector (Diehl et al. 2016).

For co-immunoprecipitation (co-IP) and oligomerization studies, we generated numerous versions of EBOV/Mak-C15 NP in the pGL4.23-CMV backbone. For a traditional dual-tag co-IP-western blot (WB) strategy, we generated pGL4.23-CMV/NP-myc and pGL4.23-CMV/NP-V5, with both tags at the C-terminus of the fusion protein. For the bioluminescence resonance energy transfer (BRET) oligomerization assay, we replaced the C-terminal V5 tag with either a HaloTag or a NLuc tag from the NanoBRET Nano-Glo Detection System (Promega). As a negative control, we generated NP-ΔOD by SDM to remove NP amino acids 20–38, thereby abrogating NP oligomerization (Kirchdoerfer et al. 2015).

For BRET experiments aimed at studying the NP-VP35 interaction, we modified the pcDNA3.3 backbone (pcDNA3.3/KLF4 was a gift from Derrick Rossi; Addgene, Cambridge, MA; plasmid #26815) (Warren et al. 2010) with a V5 peptide tag and a woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) to increase insert expression, using pLV-WPRE/mCherry, a gift from Pantelis Tsoulfas (Addgene; plasmid #36084) as source material. We then cloned in eGFP, porcine teschovirus 1 2A peptide (P2A) (Kim et al. 2011), and EBOV/Mak-C15 VP35 from a gBlock (IDT) into a single open reading frame (ORF) upstream of the WPRE. The NP-binding peptide (NPBP) of EBOV VP35 was cloned in a similar manner, but only included amino acids 1–80 (Kirchdoerfer et al. 2015). As a negative control, we also cloned mTagBFP2 into the pcDNA3.3-WPRE-V5 backbone, using mTagBFP2-pBAD, a gift from Michael Davidson (Addgene; plasmid #54572) (Subach et al. 2011) as source material .

For minigenome experiments, plasmids are described in (Luke D. Jasenosky, Neumann, and Kawaoka 2010). In this system, EBOV RNA-dependent RNA polymerase (L), viral cofactor proteins (VP30 and VP35), and NP were derived from EBOV/Yam-May and expressed from a

pCAGGS vector. We replaced EBOV/Yam-May NP with EBOV/Mak-C15 NP and its variants before measuring minigenome activity.

Cell culture and plasmid transfections

Unless otherwise specified, we grew human embryonic kidney (HEK) 293FT cells (Thermo Fisher Scientific; #R70007) in Dulbecco's modified Eagle medium (DMEM) containing 10% fetal bovine serum, 100 U/mL penicillin/streptomycin, non-essential amino acids, and sodium pyruvate (Thermo Fisher Scientific), at 37 °C with 5% CO_2.

For most assays, we performed lipid-based reverse transfection using Lipofectamine 2000 (Thermo Fisher Scientific). For a 6-well plate, we incubated 2 µg of plasmid DNA with 125 µL of Opti-MEM (Thermo Fisher Scientific) at room temperature for 5 min. We incubated this mixture with 10 µL of Lipofectamine 2000 in 115 µL of Opti-MEM at room temperature for 45 min. We added all 250 µL of the DNA:lipid mixture to a well of a 6-well plate, and then added trypsin-harvested cells. For smaller or larger plates, amounts were scaled accordingly. For BRET experiments, we used Opti-MEM without phenol red (Thermo Fisher Scientific) to minimize background fluorescence from culture media.

For the minigenome assay and electron microscopy, we performed forward transfection by incubating DNA with TransIT-LT1 (Mirus Bio, Madison, WI) in a 1:3 DNA:reagent ratio in Opti-MEM for 20 min at room temperature, and then added the mixture dropwise onto cells in 6- or 12-well plates.

Virion-like particle (VLP) budding assay

We grew cells to near confluency, harvested following trypsinization, reverse-transfected, and plated them in 6-well poly-D-lysine plates (Corning, Corning, NY). We reverse-transfected each

well with 50 ng of pcDNA3.1/NLuc-VP40 or pNL1.1/NLuc negative control and 2000 ng of pGL4.23-CMV/NP-V5, NP mutants, or pcDNA3.3/eGFP negative control.

At 16 h post-transfection, we removed supernatant, washed the cells with DMEM, and added 1.5 mL of fresh DMEM. 24 h later (40 h post-transfection), we harvested culture supernatant and filtered it through an Acrodisc 0.45 µm low protein-binding filter (Pall Laboratory, Port Washington, NY). We underlaid 1 mL of filtered supernatant with 1 mL of 20% (w/v) sucrose (Sigma-Aldrich) in phosphate-buffered saline (PBS) and ultracentrifuged at 222,000 x *g* at 4 °C for 2 h. We aspirated the supernatant and sucrose, and resuspended the pellet in 170 µL of PBS and rocked at room temperature for 1 h. We aliquoted resuspended VLPs into 3 x 50 µL as technical triplicates, added 50 µL of Nano-Glo assay reagent (Promega) to each replicate, and incubated in 96-well non-binding-surface plates (Corning) in the dark at room temperature for 10 min. We measured total luminescence on a SpectraMax L (Molecular Devices, Sunnyvale, CA) over a 1 s integration. Technical triplicates were averaged and considered as a single biological replicate. For each NP mutant and control, we performed 6 biological replicates.

For the thermal stability assay, the same protocol was used with reverse transfection of 50 ng of pcDNA3.1/NLuc-VP40 per well, except that we heated 1.2 mL of filtered supernatant on a Mastercycler pro S thermocycler (Eppendorf, Hamburg, Germany) at 4, 22, 37.1, 43.8, 60.2, or 95 °C, for 30 min. We saved 50 µL of sample at each temperature point for direct NLuc measurement. Subsequently, we carried out the remainder of the standard protocol described above by ultracentrifugation of 1 mL of heated supernatant through sucrose to purify VLPs, and subsequent measurement of NLuc activity as described above. For each temperature, we performed 3 biological replicates.

Electron Microscopy (EM)

We seeded 6 x 10^5 HEK 293 cells per well in 6-well plates. The following day, we transfected each well with 1250 ng of pcDNA3.1(+)-VP40 (untagged), 930 ng of pGL4.23-CMV/NP or pGL4.23-CMV/NP-R111C (both untagged), and 310 ng of pGL4.23-CMV/GP-A82V (Diehl et al. 2016) using 6.25µl of TransIT-LT1 (Mirus Bio). We changed media the next morning. 48 h later, we harvested culture supernatant, filtered it through a 0.45 µm filter, and overlaid it on a 20% (w/v) sucrose in TNE (10 mM Tris-Cl, 100 mM NaCl, 1 mM EDTA pH 7.5) cushion. VLPs were pelleted by ultracentrifugation at 222,000 x *g* for 2 h at 4 °C. We aspirated the supernatant and sucrose, washed the pellet gently with 1 mL of ice-cold PBS, and resuspended VLPs in 100 µL of 2% FBS in PBS and stored VLPs at 4 °C prior to EM.

We prepared samples for EM based on our previously described protocol (Gao and Hendricks 2012). Briefly, we performed all spreads on freshly prepared Carbon stabilized Formvar Support films on 200 mesh copper grids. We adsorbed VLPs onto a carbon-coated Formvar support films for 30 s. We removed excess liquid with filter paper and negatively stained the samples immediately by running 6 drops of 1% uranyl acetate over the grid to contrast the VLPs. We removed excess stain and air-dried the samples in a controlled humidity chamber. We then examined the samples using a FEI Tecnai 12 Spirit BioTwin transmission electron microscope (Thermo Fisher Scientific) using an accelerating voltage of 120 Kv. We captured micrographs at various magnifications to record the fine structure of VLPs and exported micrographs into ImageJ (Schneider, Rasband, and Eliceiri 2012) to measure the length and volume of individual particles.

Co-immunoprecipitation (co-IP)

We washed cells in 6-well plates with PBS, harvested by scraping, pelleted, and resuspended cells in 30 uL of 1.2% (w/v) polyvinylpyrrolidine (PVP) in 20 mM 4-(2-hydroxyethyl)-1-

piperazineethanesulfonic acid (HEPES) buffer pH 7.4, and snap-froze with liquid nitrogen. We lysed cells with 250–500 µL of pre-chilled lysis buffer with end-over-end rotation at 4 °C for 30 min, pelleted membranous debris at 8000 x *g* at 4 °C for 10 min, and saved an aliquot as input. To capture the target protein, we prepared a mixture of 25 µL each of Protein A and Protein G SureBeads Magnetic Beads (Bio-Rad), and bound 1–2 µg of primary antibody by rotation at room temperature for 20 min. We washed the bead-antibody complexes thrice with lysis buffer, and then incubated with cleared cell lysate while rotating the mixture at 4 °C for 2 h. After capture, we washed beads six times with wash buffer followed by a final wash with PBS, and eluted proteins by boiling in 50 µL of Laemmli sample buffer (Bio-Rad) at 95 °C for 10 min. We separated proteins by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and detected proteins of interest by chemiluminescent WB.

For reciprocal co-IP experiment in Figure 4B, we instead captured protein complexes with Protein A/G PLUS-Agarose beads (Santa Cruz Biotechnology, Dallas, TX). In this setup, we first incubated cleared cell lysate with 1–2 µg of primary antibody, rotated at 4 °C for 2–4 h, and then added 40 µL of protein A/G agarose beads, and rocked at 4 °C overnight. We washed bead-antibody complexes four times with wash buffer, twice with PBS, and eluted proteins as described above.

For the dual-tag co-IP-WB for NP oligomerization, we used RIPA buffer (50 mM Tris pH 6.8, 150 mM NaCl, 0.5% (w/v) sodium deoxycholate, 1% (w/v) Triton X-100 (Sigma-Aldrich)) as both the lysis and wash buffer because the NP-NP interaction is very strong (Watanabe, Noda, and Kawaoka 2006).

Western blot (WB)

We loaded the specified amount of input into 10% acrylamide SDS gels, and ran at 180 V until complete. We transferred protein to Immun-Blot PVDF membranes (Bio-Rad) in a wet tank either at 200 mA for 1.5 h at 4 °C, or at 40 V overnight at 4 °C. We blocked membranes by

rocking in blocking buffer consisting of 5% non-fat dry milk (Santa Cruz Biotechnology) dissolved in tris-buffered saline with 0.1% Tween 20 (TBS-T) buffer for 1 hr at room temperature. We incubated membranes with primary antibody in blocking buffer for 45 mins, washed the membrane three times in TBS-T, incubated with HRP-conjugated secondary antibody in blocking buffer for 1 hr, and washed the membrane three times. We detected chemiluminesence with SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific), and imaged with an AlphaInnotech ChemiImager (ProteinSimple, San Jose, CA) or FluorChem E (ProteinSimple) CCD camera.

BRET NP oligomerization assay

We grew cells to near confluency, harvested by trypsinization, reverse-transfected, and plated cells in poly-D-lysine 96-well black/clear flat bottom plates (Corning). We reverse-transfected each well with 10 ng of pGL4.23-CMV/NP-NLuc or pNL1.1/NLuc negative control and 100 ng of pGL4.23-CMV/NP-HaloTag or pcDNA3.3/eGFP negative control. At the start of transfection, we also added dimethylsulfoxide (DMSO) control or HaloTag NanoBRET 618 ligand in DMSO (NanoBRET Nano-Glo Detection System, Promega) to cell culture media at a final concentration of 100 nM.

At 24 h post-transfection, we added 1:100 NanoBRET Nano-Glo Substrate, incubated cells in the dark at room temperature for 45 min, and measured luminescence on a DTX880 Multimode Detector (Beckman Coulter, Brea, CA) with emission filters of 625/35 nm (HaloTag ligand acceptor signal), and then 465/35 nm (NLuc donor signal), both over 1 s integrations. We calculated BRET signal as the 625 nm / 465 nm ratio for a sample subtracted by the same ratio for its corresponding DMSO (no HaloTag ligand) control, per manufacturer's protocol.

For the VP35 inhibition experiment, we reverse-transfected each well with 2 ng of pGL4.23-CMV/NP-NLuc and 10 ng of pGL4.23-CMV/NP-HaloTag. To test a range of

VP35(NPBP) expression, we co-transfected decreasing amounts of pcDNA3.3-WPRE/eGFP-P2A-VP35(NPBP) plasmid. To ensure that each well received the same total amount of DNA, we also co-transfected increasing amounts of control pcDNA3.3-WPRE/mTagBFP2 plasmid. We serially diluted pcDNA3.3-WPRE/eGFP-P2A-VP35(NPBP) plasmid in control pcDNA3.3-WPRE/mTagBFP2 plasmid, as described in the manufacturer's protocol. We performed the remainder of the standard BRET protocol as described above and collected 3 biological replicates.

For the donor saturation setup, we reverse-transfected all wells with 2 ng of pGL4.23-CMV/NP-NLuc. To test a range of NP-HaloTag expression, we co-transfected decreasing amounts (80, 20, 5, 0 ng) of pGL4.23-CMV/NP-HaloTag or pcDNA3.3/eGFP negative control. To ensure that each well received the same total amount of DNA, we also co-transfected increasing amounts of control pcDNA3.3/eGFP plasmid. We serially diluted pGL4.23-CMV/NP-HaloTag or pcDNA3.3/eGFP in control pcDNA3.3/eGFP plasmid, as described in the manufacturer's protocol. We collected 6 biological replicates for each NP mutant and controls.

Minigenome assay

Screening of NP mutants was done as described in (Luke D. Jasenosky, Neumann, and Kawaoka 2010). We seeded HEK 293T cells into 12-well plates, grew to 70% confluence, and transfected with 2 µg of pCAGGS/L, 0.25 µg of pCAGGS/VP30, 0.5 µg of pCAGGS/T7pol, 0.5 µg of 3E5E/T7-FLuc, 0.1 µg of pCAGGS/RLuc, and 0.75 µg of pCAGGS/NP-2A-VP35 for each NP mutant. After 2 days, we washed and lysed cells with 100 µL of 1X Passive Lysis Buffer (Dual Luciferase Assay Kit, Promega), freeze-thawed lysates, and cleared by centrifugation. We incubated 10 µL of lysate with 50 µL of Luciferase Assay Reagent II, allowed the mixture to settle for 2 s, and integrated luminescence for 10 s on a Spark 10M microplate reader (Tecan, Zürich, Switzerland) to measure FLuc activity. We then added 50 µL of Stop & Glow reagent and integrated luminescence for 10 s to measure RLuc activity. FLuc activity from the EBOV

minigenome was normalized to the co-transfected RLuc control. Data are presented as the percent of the ancestral EBOV/Mak-C15 NP normalized FLuc activity and represent the result of three replicates.

Co-IP and Tandem Mass Spectrometry (MS/MS)

For mass spectrometry and reciprocal co-IPs of AP1 protein complexes, we used the mild lysis and wash buffers, slightly modified from another study (Lin et al. 2013). Mild buffer consisted of 20 mM K-HEPES buffer pH 7.4, 100 mM NaOAc, 2 mM $MgCl_2$, 0.1% (v/v) Tween 20, 250 mM NaCl, 0.5% (v/v) Triton X-100, 4 µg/mL DNase I (QIAgen, Hilden, Germany), 2 µg/mL RNase A (QIAgen), 1/200 (v/v) each phosphatase inhibitor cocktails 2 and 3 (Sigma-Aldrich), and 1/100 (v/v) protease inhibitor mixture (Sigma-Aldrich). Wash buffer consisted of 20 mM K-HEPES pH 7.4, 100 mM NaOAc, 2 mM $MgCl_2$, 0.1% (v/v) Tween 20, 500 mM NaCl, and 0.5% (v/v) Triton X-100.

To assess protein-protein interactions of NP, we scaled up our co-IP protocol. We grew two 15-cm^2 plates of cells to 40–60% confluence and transfected with 32 µg of pGL4.23-CMV/NP-myc encoding either the NP-R111 or the NP-K109E/K110E/R111E mutants. After 48 h, we harvested cells by scraping and lysed in 2.5 mL of mild lysis buffer.

We performed co-IP of myc-tagged NP complexes using 25 µg of mouse α-myc IgG or irrelevant normal mouse IgG at 4° C overnight, and bound complexes to 250 µL of protein A/G agarose beads at 4° C for 2 h. We washed beads as described above and eluted proteins in 120 µL of Laemmli sample buffer at 95 °C for 10 min. We separated proteins by SDS-PAGE, visualized with PageBlue Protein Staining Solution (Thermo Fisher), and excised lanes excluding IgG chains.

We cut gel bands into approximately 1-mm^3 pieces and performed a modified in-gel trypsin digestion procedure (Shevchenko et al. 1996). We dehydrated pieces with acetonitrile for 10 min, dried them completely in a speed-vac pump, and rehydrated with 50 mM ammonium

bicarbonate solution containing 12.5 ng/µl of modified sequencing-grade trypsin (Promega) at 4 °C for 45 min. To extract peptides, we replaced the solution with 50 mM trypsin-free ammonium bicarbonate solution and incubated at 37 °C overnight. We washed once with 50% acetonitrile and 1% formic acid, dried in a speed-vac pump for ~1 h and then stored at 4 °C. On the day of analysis, we reconstituted peptides in 5–10 µl of high-performance liquid chromatography (HPLC) solvent A (2.5% acetonitrile, 0.1% formic acid). We packed nano-scale reverse-phase HPLC capillary columns with 2.6 µm C18 spherical silica beads into fused silica capillary tubes (100 µm inner diameter x ~25 cm length) using flame-drawn tips (Peng and Gygi 2001). After equilibrating the columns, we loaded each sample via a Famos autosampler (LC Packings, San Francisco, CA). Peptides were eluted with increasing concentrations of solvent B (97.5% acetonitrile, 0.1% formic acid).

Peptides were detected by MS/MS on an LTQ Orbitrap Velos Pro ion-trap mass spectrometer (Thermo Fisher). We matched MS/MS fragmentation spectra to human forward protein databases and against reverse databases to a 1–2% false discovery rate using the SEQUEST database search program (Thermo Fisher) (Eng, McCormack, and Yates 1994). We computed unique and total peptide spectra matches (PSMs) for each identified protein.

To generate a list of putative NP interacting partners, we filtered proteins with at least 2 unique PSMs in co-IPs of both NP-R111 and NP-K109E/K110E/R111E, and at least 2-fold greater-than-average PSM enrichment of α-myc co-IP over both IgG controls combined. To eliminate abundant and ‘sticky’ proteins, we normalized average PSM enrichment against PSMs identified in all 411 Contaminant Repository for Affinity Purification (CRAPome) version 1.1 experiments (Mellacheruvu et al. 2013), a collection of proteins identified in negative control isolations. From each replicate, we used the top 10% proteins enriched versus CRAPome experiments for Search Tool for the Retrieval of Interacting Genes/proteins (STRING) version 10 analysis (Snel 2000; Szklarczyk et al. 2015) and visualized interactions with Cytoscape

(Shannon et al. 2003). See Supplementary file S1 for raw and filtered peptide/protein PSM counts.

Statistical analysis

We performed all hypothesis testing using Prism 7 (GraphPad Software, La Jolla, CA) and all non-linear curve fitting using R (R Core Team 2016) and the 'nlstools' package (Baty et al. 2015). We generated most plots using the 'ggplot2' package in R (Wickham 2016).

We quantified raw NLuc intensities from VLPs produced from expression of NLuc-VP40, NLuc-VP40-L117R, or NLuc alone with n = 6 biological replicates each. To assess statistical significance, we performed a repeated measures ANOVA (rANOVA) with Dunnett's post-test in which each condition was compared to NLuc-VP40 to generate an adjusted p-value.

To measure the impact of NP genotype on VLP production, we co-expressed NLuc-VP40 and NP-R111 or NP mutants (R111C, R111E, K109E/K110E/R111E, ΔC50 - a 50 amino acid truncation of the NP C-terminus (Licata et al. 2004)) or eGFP control. rANOVA revealed significant day-to-day (replicate-to-replicate) variability; therefore, we normalized NLuc intensities for all NP mutants to NP-R111 for each replicate. We performed Dunnett's post-test with the normalization group NP-R111 removed (since variance and degrees of freedom of NP-R111 are both 0 after normalization) and compared each NP mutant or eGFP versus 1 to generate an adjusted p-value.

To determine whether heating disrupted VLPs, we expressed NLuc-VP40 and heated cell culture supernatant to 4, 22, 37.1, 43.8, 60.2, or 95 °C either before or after purifying VLPs via ultracentrifugation with n = 3 biological replicates in a repeated measures design. We normalized NLuc values for all temperatures to the 4 °C value for each replicate, log-transformed the normalized values, and fit the data to sigmoidal curves using the 'nls' function in R:

$$\log_{10}(\text{NLuc.Norm}) \sim \text{min} + \text{max}/(1 + e^{(\text{midpt} - \text{temp})/\text{scale}}) \quad (1)$$

, where temperature is the independent variable, NLuc.Norm is the dependent variable, and min, max, midpt, and scale are all constants to be fitted. Additionally, we tested whether NLuc intensity differed with heating to 60.2 °C either before or after purifying VLPs with a paired t-test.

For the VLP35 inhibition using BRET, we expressed varying amounts of VP35 NP binding peptide in the presence of NP-NLuc and NP-HaloTag with n = 3 biological replicates for each VP35 expression level. We fit the inverse function to data using the 'nls' function in R:

$$\text{BRET} \sim \text{scale}/(\text{VP35} + \text{max}) + \text{min} \quad (2)$$

, where VP35 expression is the independent variable, BRET is the dependent variable, and scale, max, and min are constants to be fitted. In the non-linear regression, scale = 1.9×10^5, max = 3.4×10^4, and min = 0.49, leading to the model:

$$\text{BRET} \sim 1.88 \times 10^5/(\text{VP35} + 3.53 \times 10^4) + 0.49 \quad (3)$$

This model suggests that, in the absence VP35 expression (VP35 = 0), the maximum BRET signal would be 5.80; with very high VP35 expression ($\text{VP35} \rightarrow \infty$), the minimum BRET signal would be 0.49. To determine the appropriateness of the non-linear regression, we calculated the residual/total sum of squares to be 0.95, very close to the perfect regression value of 1. To calculate confidence intervals, we used the 'nlstools' package to generate 999 bootstrap pseudoreplicates, inferred parameters for each pseudoreplicate, and plotted the central 95% of values as a shaded region.

For donor saturation assay using BRET, we performed the BRET protocol with NLuc- and HaloTag-tagged version of NP-R111 or NP mutants or eGFP control with n = 6 biological replicates for each HaloTag expression level. We fit the data to Michaelis-Menten curves using the 'nls' function in R:

$$\text{BRET} \sim v_{max} * \text{NP-HaloTag} / (K_m + \text{NP-HaloTag}) \quad (4)$$

, where the concentration of NP-HaloTag is the independent variable, BRET is the dependent variable, and v_{max} and K_m are constants to be fitted. For NP-R111 or each NP mutant or eGFP

control, we inferred v_{max} and K_m and generated 95% confidence intervals using 'nlstools' as described above. Data points from eGFP and NP-ΔOD controls failed to generate appropriate curve fits. To determine whether the remaining curve fits were significantly different from each other, we performed ANOVA with Dunnett's post-test in which NP-R111C, NP-R111E, and NP-K109E/K110E/R111E were compared to NP-R111 to generate an adjusted p-value using Prism 7.

Acknowledgements

We thank Ross Tomaino of the Harvard Taplin Mass Spectrometry Facility for sample and instrument handling following immunoprecipitation of NP-containing protein complexes. We thank Gregory Hendricks of the University of Massachusetts Medical School Core Electron Microscopy Facility for sample and instrument handling following production of Ebola virion-like particles. Electron Microscopy (EM) was supported by Award Number S10RR027897 from the National Center for Research Resources to the University of Massachusetts Medical School Core EM Facility. Ebola virus phylogenetics was made possible by the Virus Pathogen Database and Analysis Resource (ViPR) (Pickett et al. 2012), which has been wholly funded with federal funds from NIAID, the US National Institutes of Health, and the US Department of Health and Human Services (DHSS), under Contract No. HHSN272201400028C. The authors are solely the responsibility for the content of this paper and do not necessarily represent the official views of the the National Center for Research Resources or of US DHHS or of the institutions and companies affiliated with the authors.

Additional information

Funding

Funder	Grant reference number	Author
National Science Foundation	DGE 1144152	Aaron E Lin
National Institute of Allergy and Infectious Diseases (NIAID)	HHSN272200700016I	Yingyun Cai
NIAID	HHSN272200700016I	Courtney Finch
National Institute of General Medical Sciences	T32GM007753	Chidiebere Akusobi
NIAID	R01AI118016	Erica Ollmann Saphire
Pew Charitable Trusts	Pew Biomedical Scholar	Kristian G Andersen
National Center for Advancing Translational Studies Clinical and Translational Science	UL1TR001114	Kristian G Andersen
NIAID	HHSN272201400048C	Kristian G Andersen
NIAID	HHSN272200700016I	Jens H Kuhn
National Institute on Drug Abuse	DP1DA034990	Jeremy Luban
NIAID	R01AI111809	Jeremy Luban
NIAID	U19AI110818	Pardis C Sabeti
Howard Hughes Medical Institute	Investigator Award	Pardis C Sabeti

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Aaron E Lin, Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Visualization, Writing—original draft, Writing—review & editing; William E Diehl, Conceptualization, Methodology, Validation, Formal analysis, Investigation, Visualization, Writing—original draft, Writing—review & editing; Yingyun Cai, Methodology, Validation,

Investigation, Writing—review & editing; Courtney Finch, Methodology, Validation, Investigation, Writing—review & editing; Chidiebere Akusobi, Methodology, Investigation, Writing—review & editing; Robert N Kirchdoerfer, Methodology, Validation, Formal analysis, Investigation, Writing—review & editing; Laura Bollinger, Visualization, Writing—review & editing; Stephen F Schaffner, Visualization, Writing—original draft, Writing—review & editing; Elizabeth A Brown, Writing—original draft, Writing—review & editing; Erica Ollmann Saphire, Writing—review & editing, Supervision; Kristian G Andersen, Software, Formal analysis, Writing—original draft, Writing—review & editing, Supervision; Jens H Kuhn, Methodology, Writing—review & editing, Supervision; Jeremy Luban, Conceptualization, Methodology, Writing—original draft, Writing—review & editing, Supervision, Project administration; Pardis C Sabeti, Conceptualization, Writing—original draft, Writing—review & editing, Supervision, Project administration

Author ORCIDs

Aaron E Lin, https://orcid.org/0000-0001-7400-4125

William E Diehl, https://orcid.org/0000-0001-7943-1214

Yingyun Cai, https://orcid.org/0000-0002-1101-0920

Courtney Finch, https://orcid.org/0000-0002-3674-3282

Chidiebere Akusobi, https://orcid.org/0000-0002-1611-0015

Laura Bollinger, https://orcid.org/0000-0002-6676-8894

Stephen F Schaffner, https://orcid.org/0000-0001-6699-3568

Erica Ollmann Saphire, https://orcid.org/0000-0002-1206-7451

Kristian G Andersen, https://orcid.org/0000-0001-6431-5982

Jens H Kuhn, https://orcid.org/0000-0002-7800-6045

Jeremy Luban, https://orcid.org/0000-0001-5650-4054

Pardis C Sabeti, https://orcid.org/0000-0002-9843-1890

Additional files

Supplementary files

- Supplementary file S1. List of all proteins identified by co-immunoprecipitation tandem mass spectrometry (co-IP-MS/MS), related to Figure 4A. We expressed V5-tagged Ebola virus NP-R111 or NP-K109E/K110E/R111E in HEK 293FT cells, and performed co-IP-MS/MS with either normal mouse IgG control or α-V5 mouse IgG. We matched MS/MS fragmentation spectra to human forward protein databases and against reverse databases to a 1–2% false discovery rate using the SEQUEST database search program (Thermo Fisher) (Eng, McCormack, and Yates 1994). We computed unique and total peptide spectra matches (PSMs) for each identified protein. To generate a list of putative NP interacting partners, we filtered proteins with at least 2 unique PSMs in co-IPs of both NP-R111 and NP-K109E/K110E/R111E, and at least 2-fold greater-than-average PSM enrichment of α-myc co-IP over both IgG controls combined. To eliminate abundant and 'sticky' proteins, we normalized average PSM enrichment against PSMs identified in all 411 Contaminant Repository for Affinity Purification (CRAPome) version 1.1 experiments (Mellacheruvu et al. 2013), a collection of proteins identified in negative control isolations . From each replicate, we used the top 10% proteins enriched versus CRAPome experiments for Search Tool for the Retrieval of Interacting Genes/proteins (STRING) version 10 analysis (Snel 2000; Szklarczyk et al. 2015) and visualized interactions with Cytoscape (Shannon et al. 2003).
- Transparent reporting form

Data availability

All data generated or analyzed during this study are included in the manuscript and supporting files.

References

Albariño, César G., Lisa Wiggleton Guerrero, Michael K. Lo, Stuart T. Nichol, and Jonathan S. Towner. 2015. “Development of a Reverse Genetics System to Generate a Recombinant Ebola Virus Makona Expressing a Green Fluorescent Protein.” *Virology* 484: 259–64.

Baize, Sylvain, Delphine Pannetier, Lisa Oestereich, Toni Rieger, Lamine Koivogui, N ’faly Magassouba, Barrè Soropogui, et al. 2014. “Emergence of Zaire Ebola Virus Disease in Guinea.” *The New England Journal of Medicine* 371 (15): 1418–25.

Basler, Christopher F. 2017. “West African Ebola Virus Strains: Unstable and Ready to Invade?” *Cell Host & Microbe* 21 (3): 316–18.

Baty, Florent, Christian Ritz, Sandrine Charles, Martin Brutsche, Jean-Pierre Flandrois, and Marie-Laure Delignette-Muller. 2015. “A Toolbox for Nonlinear Regression in R: The Package NIstools.” *Journal of Statistical Software* 66 (5). https://doi.org/10.18637/jss.v066.i05.

Bornholdt, Zachary A., Takeshi Noda, Dafna M. Abelson, Peter Halfmann, Malcolm R. Wood, Yoshihiro Kawaoka, and Erica Ollmann Saphire. 2013. “Structural Rearrangement of Ebola Virus VP40 Begets Multiple Functions in the Virus Life Cycle.” *Cell* 154 (4): 763–74.

Camus, Grégory, Carolina Segura-Morales, Dorothee Molle, Sandra Lopez-Vergès, Christina Begon-Pescia, Chantal Cazevieille, Peter Schu, Edouard Bertrand, Clarisse Berlioz-Torrent, and Eugenia Basyuk. 2007. “The Clathrin Adaptor Complex AP-1 Binds HIV-1 and MLV Gag and Facilitates Their Budding.” *Molecular Biology of the Cell* 18 (8): 3193–3203.

Capella-Gutiérrez, Salvador, José M. Silla-Martínez, and Toni Gabaldón. 2009. “trimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses.” *Bioinformatics* 25 (15): 1972–73.

Carroll, Miles W., David A. Matthews, Julian A. Hiscox, Michael J. Elmore, Georgios Pollakis, Andrew Rambaut, Roger Hewson, et al. 2015. “Temporal and Spatial Analysis of the 2014-2015 Ebola Virus Outbreak in West Africa.” *Nature* 524 (7563): 97–101.

Cauldwell, Anna V., Jason S. Long, Olivier Moncorgé, and Wendy S. Barclay. 2014. “Viral Determinants of Influenza A Virus Host Range.” *The Journal of General Virology* 95 (Pt 6): 1193–1210.

Diehl, William E., Aaron E. Lin, Nathan D. Grubaugh, Luiz Max Carvalho, Kyusik Kim, Pyae Phyo Kyawe, Sean M. McCauley, et al. 2016. “Ebola Virus Glycoprotein with Increased Infectivity Dominated the 2013-2016 Epidemic.” *Cell* 167 (4): 1088–98.e6.

Dietzel, Erik, Gordian Schudt, Verena Krähling, Mikhail Matrosovich, and Stephan Becker. 2017. “Functional Characterization of Adaptive Mutations during the West African Ebola Virus Outbreak.” *Journal of Virology* 91 (2). https://doi.org/10.1128/JVI.01913-16.

Dong, Shishang, Peng Yang, Guobang Li, Baocheng Liu, Wenming Wang, Xiang Liu, Boran Xia, et al. 2015. “Insight into the Ebola Virus Nucleocapsid Assembly Mechanism: Crystal Structure of Ebola Virus Nucleoprotein Core Domain at 1.8 Å Resolution.” *Protein & Cell* 6 (5): 351–62.

Eng, J. K., A. L. McCormack, and J. R. Yates. 1994. “An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database.” *Journal of the American Society for Mass Spectrometry* 5 (11): 976–89.

Gao, Guangping, and Gregory M. Hendricks. 2012. “Introducing Genes into MAmmalian Cells: Viral Vectors. Protocol 14: Analysis of rAAv Sample Morphology Using Negative Staining and High Resolution Electron Microscopy.” In *Molecular Cloning (A Laboratory Manual)*, edited by Michael R. Green and Joseph Sambrook, 2:1301–3. Cold Spring Harbor Press.

García-Dorival, Isabel, Weining Wu, Stuart D. Armstrong, John N. Barr, Miles W. Carroll, Roger Hewson, and Julian A. Hiscox. 2016. “Elucidation of the Cellular Interactome of Ebola Virus

Nucleoprotein and Identification of Therapeutic Targets." *Journal of Proteome Research* 15 (12): 4290–4303.

Gibson, Daniel G., Lei Young, Ray-Yuan Chuang, J. Craig Venter, Clyde A. Hutchison 3rd, and Hamilton O. Smith. 2009. "Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases." *Nature Methods* 6 (5): 343–45.

Gire, Stephen K., Augustine Goba, Kristian G. Andersen, Rachel S. G. Sealfon, Daniel J. Park, Lansana Kanneh, Simbirie Jalloh, et al. 2014. "Genomic Surveillance Elucidates Ebola Virus Origin and Transmission during the 2014 Outbreak." *Science* 345 (6202): 1369–72.

Goldstein, Tracey, Simon J. Anthony, Aiah Gbakima, Brian H. Bird, James Bangura, Alexandre Tremeau-Bravard, Manjunatha N. Belaganahalli, et al. 2018. "The Discovery of Bombali Virus Adds Further Support for Bats as Hosts of Ebolaviruses." *Nature Microbiology*. https://doi.org/10.1038/s41564-018-0227-2.

Groseth, A., J. E. Charton, M. Sauerborn, F. Feldmann, S. M. Jones, T. Hoenen, and H. Feldmann. 2009. "The Ebola Virus Ribonucleoprotein Complex: A Novel VP30–L Interaction Identified." *Virus Research* 140 (1-2): 8–14.

Hall, Mary P., James Unch, Brock F. Binkowski, Michael P. Valley, Braeden L. Butler, Monika G. Wood, Paul Otto, et al. 2012. "Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate." *ACS Chemical Biology* 7 (11): 1848–57.

Harty, R. N., M. E. Brown, G. Wang, J. Huibregtse, and F. P. Hayes. 2000. "A PPxY Motif within the VP40 Protein of Ebola Virus Interacts Physically and Functionally with a Ubiquitin Ligase: Implications for Filovirus Budding." *Proceedings of the National Academy of Sciences of the United States of America* 97 (25): 13871–76.

Hoenen, Thomas, Allison Groseth, Julie Callison, Ayato Takada, and Heinz Feldmann. 2013. "A Novel Ebola Virus Expressing Luciferase Allows for Rapid and Quantitative Testing of Antivirals." *Antiviral Research* 99 (3): 207–13.

Hoenen, Thomas, Allison Groseth, Kyle Rosenke, Robert J. Fischer, Andreas Hoenen, Seth D. Judson, Cynthia Martellaro, et al. 2016. "Nanopore Sequencing as a Rapidly Deployable Ebola Outbreak Tool." *Emerging Infectious Diseases* 22 (2): 331–34.

Hoenen, T., S. Jung, A. Herwig, A. Groseth, and S. Becker. 2010. "Both Matrix Proteins of Ebola Virus Contribute to the Regulation of Viral Genome Replication and Transcription." *Virology* 403 (1): 56–66.

Hoenen, T., D. Safronetz, A. Groseth, K. R. Wollenberg, O. A. Koita, B. Diarra, I. S. Fall, et al. 2015. "Virology. Mutation Rate and Genotype Variation of Ebola Virus from Mali Case Sequences." *Science* 348 (6230): 117–19.

Hoffmann, Markus, Lisa Crone, Erik Dietzel, Jennifer Paijo, Mariana González-Hernández, Inga Nehlmeier, Ulrich Kalinke, Stephan Becker, and Stefan Pöhlmann. 2017. "A Polymorphism within the Internal Fusion Loop of the Ebola Virus Glycoprotein Modulates Host Cell Entry." *Journal of Virology* 91 (9). https://doi.org/10.1128/JVI.00177-17.

Holmes, Edward C. 2009. *The Evolution and Emergence of RNA Viruses*. Oxford University Press.

Hu, Lei, Jared M. Trefethen, Yuhong Zeng, Luisa Yee, Satoshi Ohtake, David Lechuga-Ballesteros, Kelly L. Warfield, et al. 2011. "Biophysical Characterization and Conformational Stability of Ebola and Marburg Virus-like Particles." *Journal of Pharmaceutical Sciences* 100 (12): 5156–73.

Jasenosky, L. D., G. Neumann, I. Lukashevich, and Y. Kawaoka. 2001. "Ebola Virus VP40-Induced Particle Formation and Association with the Lipid Bilayer." *Journal of Virology* 75 (11): 5205–14.

Jasenosky, Luke D., and Yoshihiro Kawaoka. 2004. "Filovirus Budding." *Virus Research* 106 (2): 181–88.

Jasenosky, Luke D., Gabriele Neumann, and Yoshihiro Kawaoka. 2010. "Minigenome-Based Reporter System Suitable for High-Throughput Screening of Compounds Able to Inhibit

Ebolavirus Replication And/or Transcription." *Antimicrobial Agents and Chemotherapy* 54 (7): 3007–10.

Katoh, Kazutaka, and Daron M. Standley. 2013. "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability." *Molecular Biology and Evolution* 30 (4): 772–80.

Kim, Jin Hee, Sang-Rok Lee, Li-Hua Li, Hye-Jeong Park, Jeong-Hoh Park, Kwang Youl Lee, Myeong-Kyu Kim, Boo Ahn Shin, and Seok-Yong Choi. 2011. "High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines, Zebrafish and Mice." *PloS One* 6 (4): e18556.

Kirchdoerfer, Robert N., Dafna M. Abelson, Sheng Li, Malcolm R. Wood, and Erica Ollmann Saphire. 2015. "Assembly of the Ebola Virus Nucleoprotein from a Chaperoned VP35 Complex." *Cell Reports* 12 (1): 140–49.

Kirchdoerfer, Robert N., Crystal L. Moyer, Dafna M. Abelson, and Erica Ollmann Saphire. 2016. "The Ebola Virus VP30-NP Interaction Is a Regulator of Viral RNA Synthesis." *PLoS Pathogens* 12 (10): e1005937.

Kirmaier, Andrea, Fan Wu, Ruchi M. Newman, Laura R. Hall, Jennifer S. Morgan, Shelby O'Connor, Preston A. Marx, et al. 2010. "TRIM5 Suppresses Cross-Species Transmission of a Primate Immunodeficiency Virus and Selects for Emergence of Resistant Variants in the New Species." *PLoS Biology* 8 (8). https://doi.org/10.1371/journal.pbio.1000462.

Krupp, Annabel, Kevin R. McCarthy, Marcel Ooms, Michael Letko, Jennifer S. Morgan, Viviana Simon, and Welkin E. Johnson. 2013. "APOBEC3G Polymorphism as a Selective Barrier to Cross-Species Transmission and Emergence of Pathogenic SIV and AIDS in a Primate Host." *PLoS Pathogens* 9 (10): e1003641.

Kugelman, Jeffrey R., Mariano Sanchez-Lockhart, Kristian G. Andersen, Stephen Gire, Daniel J. Park, Rachel Sealfon, Aaron E. Lin, et al. 2015. "Evaluation of the Potential Impact of Ebola Virus Genomic Drift on the Efficacy of Sequence-Based Candidate Therapeutics." *mBio* 6 (1). https://doi.org/10.1128/mBio.02227-14.

Ladner, Jason T., Michael R. Wiley, Suzanne Mate, Gytis Dudas, Karla Prieto, Sean Lovett, Elyse R. Nagle, et al. 2015. "Evolution and Spread of Ebola Virus in Liberia, 2014-2015." *Cell Host & Microbe* 18 (6): 659–69.

Leung, Daisy W., Dominika Borek, Priya Luthra, Jennifer M. Binning, Manu Anantpadma, Gai Liu, Ian B. Harvey, et al. 2015. "An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis by Modulating Nucleoprotein-RNA Interactions." *Cell Reports* 11 (3): 376–89.

Licata, Jillian M., Reed F. Johnson, Ziying Han, and Ronald N. Harty. 2004. "Contribution of Ebola Virus Glycoprotein, Nucleoprotein, and VP24 to Budding of VP40 Virus-like Particles." *Journal of Virology* 78 (14): 7344–51.

Lin, Aaron E., Todd M. Greco, Katinka Döhner, Beate Sodeik, and Ileana M. Cristea. 2013. "A Proteomic Perspective of Inbuilt Viral Protein Regulation: pUL46 Tegument Protein Is Targeted for Degradation by ICP0 during Herpes Simplex Virus Type 1 Infection." *Molecular & Cellular Proteomics: MCP* 12 (11): 3237–52.

Liu, Yuliang, Luis Cocka, Atsushi Okumura, Yong-An Zhang, J. Oriol Sunyer, and Ronald N. Harty. 2010. "Conserved Motifs within Ebola and Marburg Virus VP40 Proteins Are Important for Stability, Localization, and Subsequent Budding of Virus-like Particles." *Journal of Virology* 84 (5): 2294–2303.

Machleidt, Thomas, Carolyn C. Woodroofe, Marie K. Schwinn, Jacqui Méndez, Matthew B. Robers, Kris Zimmerman, Paul Otto, Danette L. Daniels, Thomas A. Kirkland, and Keith V. Wood. 2015. "NanoBRET--A Novel BRET Platform for the Analysis of Protein-Protein Interactions." *ACS Chemical Biology* 10 (8): 1797–1804.

Manicassamy, Balaji, and Lijun Rong. 2009. "Expression of Ebolavirus Glycoprotein on the Target Cells Enhances Viral Entry." *Virology Journal* 6 (June): 75.

Marzi, Andrea, Spencer Chadinah, Elaine Haddock, Friederike Feldmann, Nicolette Arndt, Cynthia Martellaro, Dana P. Scott, et al. 2018. “Recently Identified Mutations in the Ebola Virus-Makona Genome Do Not Alter Pathogenicity in Animal Models.” *Cell Reports* 23 (6): 1806–16.

McCarthy, Sarah E., Jillian M. Licata, and Ronald N. Harty. 2006. “A Luciferase-Based Budding Assay for Ebola Virus.” *Journal of Virological Methods* 137 (1): 115–19.

Mellacheruvu, Dattatreya, Zachary Wright, Amber L. Couzens, Jean-Philippe Lambert, Nicole A. St-Denis, Tuo Li, Yana V. Miteva, et al. 2013. “The CRAPome: A Contaminant Repository for Affinity Purification-Mass Spectrometry Data.” *Nature Methods* 10 (8): 730–36.

Nelson, David L., Albert L. Lehninger, and Michael M. Cox. 2008. *Lehninger Principles of Biochemistry*. Macmillan.

Neumann, G., H. Feldmann, S. Watanabe, I. Lukashevich, and Y. Kawaoka. 2002. “Reverse Genetics Demonstrates That Proteolytic Processing of the Ebola Virus Glycoprotein Is Not Essential for Replication in Cell Culture.” *Journal of Virology* 76 (1): 406–10.

Ng, Andy Ka-Leung, Mandy Ka-Han Lam, Hongmin Zhang, Jinhuan Liu, Shannon Wing-Ngor Au, Paul Kay-Sheung Chan, Jiahuai Wang, and Pang-Chui Shaw. 2012. “Structural Basis for RNA Binding and Homo-Oligomer Formation by Influenza B Virus Nucleoprotein.” *Journal of Virology* 86 (12): 6758–67.

Noda, Takeshi, Hiroshi Sagara, Emiko Suzuki, Ayato Takada, Hiroshi Kida, and Yoshihiro Kawaoka. 2002. “Ebola Virus VP40 Drives the Formation of Virus-like Filamentous Particles along with GP.” *Journal of Virology* 76 (10): 4855–65.

Ortiz-Riano, E., B. Y. H. Cheng, J. C. de la Torre, and L. Martinez-Sobrido. 2012. “Self-Association of Lymphocytic Choriomeningitis Virus Nucleoprotein Is Mediated by Its N-Terminal Region and Is Not Required for Its Anti-Interferon Function.” *Journal of Virology* 86 (6): 3307–17.

Park, Daniel J., Gytis Dudas, Shirlee Wohl, Augustine Goba, Shannon L. M. Whitmer, Kristian G. Andersen, Rachel S. Sealfon, et al. 2015. “Ebola Virus Epidemiology, Transmission, and Evolution during Seven Months in Sierra Leone.” *Cell* 161 (7): 1516–26.

Peng, J., and S. P. Gygi. 2001. “Proteomics: The Move to Mixtures.” *Journal of Mass Spectrometry: JMS* 36 (10): 1083–91.

Pickett, Brett E., Eva L. Sadat, Yun Zhang, Jyothi M. Noronha, R. Burke Squires, Victoria Hunt, Mengya Liu, et al. 2012. “ViPR: An Open Bioinformatics Database and Analysis Resource for Virology Research.” *Nucleic Acids Research* 40 (Database issue): D593–98.

Pleet, Michelle L., Allison Mathiesen, Catherine DeMarino, Yao A. Akpamagbo, Robert A. Barclay, Angela Schwab, Sergey Iordanskiy, et al. 2016. “Ebola VP40 in Exosomes Can Cause Immune Cell Dysfunction.” *Frontiers in Microbiology* 7 (November): 1765.

Quick, Joshua, Nicholas J. Loman, Sophie Duraffour, Jared T. Simpson, Ettore Severi, Lauren Cowley, Joseph Akoi Bore, et al. 2016. “Real-Time, Portable Genome Sequencing for Ebola Surveillance.” *Nature* 530 (7589): 228–32.

R Core Team. 2016. “R: A Language and Environment for Statistical Computing.” *R Foundation for Statistical Computing, Vienna, Austria*. https://www.R-project.org/.

Schneider, Caroline A., Wayne S. Rasband, and Kevin W. Eliceiri. 2012. “NIH Image to ImageJ: 25 Years of Image Analysis.” *Nature Methods* 9 (7): 671–75.

Schrödinger, LLC. 2015. “The PyMOL Molecular Graphics System, Version 1.8.”

Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. 2003. “Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks.” *Genome Research* 13 (11): 2498–2504.

Shevchenko, A., M. Wilm, O. Vorm, and M. Mann. 1996. “Mass Spectrometric Sequencing of Proteins Silver-Stained Polyacrylamide Gels.” *Analytical Chemistry* 68 (5): 850–58.

Simon-Loriere, Etienne, Ousmane Faye, Oumar Faye, Lamine Koivogui, Nfaly Magassouba,

Sakoba Keita, Jean-Michel Thiberge, et al. 2015. “Distinct Lineages of Ebola Virus in Guinea during the 2014 West African Epidemic.” *Nature* 524 (7563): 102–4.

Smits, Saskia L., Suzan D. Pas, Chantal B. Reusken, Bart L. Haagmans, Peirro Pertile, Corrado Cancedda, Kerry Dierberg, et al. 2015. “Genotypic Anomaly in Ebola Virus Strains Circulating in Magazine Wharf Area, Freetown, Sierra Leone, 2015.” *Euro Surveillance: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin* 20 (40). https://doi.org/10.2807/1560-7917.ES.2015.20.40.30035.

Snel, B. 2000. “STRING: A Web-Server to Retrieve and Display the Repeatedly Occurring Neighbourhood of a Gene.” *Nucleic Acids Research* 28 (18): 3442–44.

Stamatakis, A., T. Ludwig, and H. Meier. 2005. “RAxML-III: A Fast Program for Maximum Likelihood-Based Inference of Large Phylogenetic Trees.” *Bioinformatics* 21 (4): 456–63.

Subach, Oksana M., Paula J. Cranfill, Michael W. Davidson, and Vladislav V. Verkhusha. 2011. “An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore.” *PloS One* 6 (12): e28674.

Sugita, Yukihiko, Hideyuki Matsunami, Yoshihiro Kawaoka, Takeshi Noda, and Matthias Wolf. 2018. “Cryo-EM Structure of the Ebola Virus nucleoprotein–RNA Complex at 3.6 Å Resolution.” *Nature*, October. https://doi.org/10.1038/s41586-018-0630-0.

Su, Zhaoming, Chao Wu, Liuqing Shi, Priya Luthra, Grigore D. Pintilie, Britney Johnson, Justin R. Porter, et al. 2018. “Electron Cryo-Microscopy Structure of Ebola Virus Nucleoprotein Reveals a Mechanism for Nucleocapsid-like Assembly.” *Cell* 172 (5): 966–78.e12.

Szklarczyk, Damian, Andrea Franceschini, Stefan Wyder, Kristoffer Forslund, Davide Heller, Jaime Huerta-Cepas, Milan Simonovic, et al. 2015. “STRING v10: Protein-Protein Interaction Networks, Integrated over the Tree of Life.” *Nucleic Acids Research* 43 (Database issue): D447–52.

Tong, Yi-Gang, Wei-Feng Shi, Di Liu, Jun Qian, Long Liang, Xiao-Chen Bo, Jun Liu, et al. 2015. “Genetic Diversity and Evolutionary Dynamics of Ebola Virus in Sierra Leone.” *Nature* 524 (7563): 93–96.

Towner, Jonathan S., Jason Paragas, Jason E. Dover, Manisha Gupta, Cynthia S. Goldsmith, John W. Huggins, and Stuart T. Nichol. 2005. “Generation of eGFP Expressing Recombinant Zaire Ebolavirus for Analysis of Early Pathogenesis Events and High-Throughput Antiviral Drug Screening.” *Virology* 332 (1): 20–27.

Tsetsarkin, Konstantin A., Dana L. Vanlandingham, Charles E. McGee, and Stephen Higgs. 2007. “A Single Mutation in Chikungunya Virus Affects Vector Specificity and Epidemic Potential.” *PLoS Pathogens* 3 (12): e201.

Ueda, Mahoko Takahashi, Yohei Kurosaki, Taisuke Izumi, Yusuke Nakano, Olamide K. Oloniniyi, Jiro Yasuda, Yoshio Koyanagi, Kei Sato, and So Nakagawa. 2017. “Functional Mutations in Spike Glycoprotein of Zaire Ebolavirus Associated with an Increase in Infection Efficiency.” *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* 22 (2): 148–59.

Urbanowicz, Richard A., C. Patrick McClure, Anavaj Sakuntabhai, Amadou A. Sall, Gary Kobinger, Marcel A. Müller, Edward C. Holmes, Félix A. Rey, Etienne Simon-Loriere, and Jonathan K. Ball. 2016. “Human Adaptation of Ebola Virus during the West African Outbreak.” *Cell* 167 (4): 1079–87.e5.

Volchkov, V. E., V. A. Volchkova, E. Muhlberger, L. V. Kolesnikova, M. Weik, O. Dolnik, and H. D. Klenk. 2001. “Recovery of Infectious Ebola Virus from Complementary DNA: RNA Editing of the GP Gene and Viral Cytotoxicity.” *Science* 291 (5510): 1965–69.

Wang, May K., Sun-Young Lim, Soo Mi Lee, and James M. Cunningham. 2017. “Biochemical Basis for Increased Activity of Ebola Glycoprotein in the 2013-16 Epidemic.” *Cell Host & Microbe* 21 (3): 367–75.

Wan, William, Larissa Kolesnikova, Mairi Clarke, Alexander Koehler, Takeshi Noda, Stephan Becker, and John A. G. Briggs. 2017. “Structure and Assembly of the Ebola Virus Nucleocapsid.” *Nature* 551 (7680): 394–97.

Warren, Luigi, Philip D. Manos, Tim Ahfeldt, Yuin-Han Loh, Hu Li, Frank Lau, Wataru Ebina, et al. 2010. “Highly Efficient Reprogramming to Pluripotency and Directed Differentiation of Human Cells with Synthetic Modified mRNA.” *Cell Stem Cell* 7 (5): 618–30.

Watanabe, Shinji, Takeshi Noda, and Yoshihiro Kawaoka. 2006. “Functional Mapping of the Nucleoprotein of Ebola Virus.” *Journal of Virology* 80 (8): 3743–51.

WHO. 2016. “Ebola Situation Report - 30 March 2016.” http://apps.who.int/ebola/current-situation/ebola-situation-report-30-march-2016.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer.

Wong, Gary, Shihua He, Anders Leung, Wenguang Cao, Yuhai Bi, Zirui Zhang, Wenjun Zhu, et al. 2018. “Naturally-Occurring Single Mutations in Ebola Observably Impact Infectivity.” *Journal of Virology*, October, JVI.01098–18.

Yuan, Ling, Xing-Yao Huang, Zhong-Yu Liu, Feng Zhang, Xing-Liang Zhu, Jiu-Yang Yu, Xue Ji, et al. 2017. “A Single Mutation in the prM Protein of Zika Virus Contributes to Fetal Microcephaly.” *Science* 358 (6365): 933–36.

Zimmermann, Petra, Benjamin Mänz, Otto Haller, Martin Schwemmle, and Georg Kochs. 2011. “The Viral Nucleoprotein Determines Mx Sensitivity of Influenza A Viruses.” *Journal of Virology* 85 (16): 8133–40.

Figure Legends

Figure 1. The Ebola virus nucleoprotein mutation R111C emerged alongside a GP-A82V mutation and lies outside of established NP known functional domains.

(**A**) Phylogenetic analysis of the 2013–2016 Ebola virus disease (EVD) epidemic. We constructed a maximum likelihood tree based on 1,823 EBOV genome sequences, and colored branches based on GP-82 and NP-111 alleles. No GP-A82/NP-R111C sequences were detected. Arrowheads point to the emergence of the GP-A82V (green) and NP-R111C (blue) mutations compared to genomes encoding the ancestral GP-A82/NP-R111 alleles (tan). The scale bar denotes substitutions/nucleotide.

(**B**) Number of EVD cases over time, stratified by genotype. Coloring is identical to Figure 1A.

(**C**) Schematic of NP. R111 (yellow) lies in an un-annotated region within the N-terminal lobe. Key residues for known NP interactions are highlighted.

(**D**) Crystal structure (PDB #4YPI) of NP. Though the precise location of the oligomerization domain has yet to be determined by crystallography (orange dashed line), the R111 residue (yellow) is located on the same face as residues proximal to the oligomerization domain (orange), but opposite to the VP35 (magenta) and RNA (red) interaction interfaces.

Figure 2. Ebola virus nucleoprotein mutation R111C increases budding of virion-like particles.

(**A**) Schematic of the virion-like particle (VLP) budding assay. We transfect plasmid encoding NLuc-VP40 to form luminescent VLPs, and co-transfect NP-expressing plasmids to measure the impact of NP genotype on VLP budding.

(**B**) VLP budding assay control. VP40 loss-of-function mutant L117R fails to form VLPs. n = 6

biological replicates.

(**C**) VLP budding with NP mutants. NP-R111C (red) significantly increases budding compared to wild-type NP-R111 (tan). The charge-reversal mutants NP-R111E (light blue) and NP-K109E/K110E/R111E (dark blue) do not affect VLP budding compared to NP-R111 as indicated by lack of statistical significance. p-values were calculated using Dunnett's test. n = 6 biological replicates.

Figure 3. Ebola virus nucleoprotein residue 111 significantly affects oligomerization of NP.

(**A**) Schematic of the NP oligomerization assay. We co-express NP fused to NLuc (donor) and HaloTag (acceptor). Binding and oligomerization brings the tags into close spatial proximity, allowing bioluminescence resonance energy transfer (BRET) and emission at 625 nm.

(**B**) Oligomerization assay controls. Absence of either tag, free NLuc, or deletion of the NP oligomerization domain (ΔOD) reduces BRET signal.

(**C**) EBOV VP35 NP-binding peptide (NPBP) disrupts NP oligomerization. In addition to expressing NP-NLuc and NP-HaloTag, we co-expressed varying amounts of eGFP-P2A-VP35(NPBP). Data (n = 3 biological replicates) are fit to an inverse function. Shading indicates 95% confidence intervals based on 999 bootstrap pseudoreplicates.

(**D**) Donor saturation assay with NP mutants. We expressed a constant amount of NP-NLuc donor and titrated increasing amounts of NP-HaloTag acceptor to generate saturation curves (n = 6 biological replicates). We fitted data to Michaelis-Menten kinetics and calculated maximum oligomerization (Max) and K_m for each NP mutant. NP-R111C (red), NP-R111E (light blue), and NP-K109E/K110E/R111E (dark blue) mutants significantly increased oligomerization compared to NP-R111 (tan). NP-ΔOD (gray) yielded much weaker oligomerization, and eGFP (black dots

near x-axis) did not produce data suitable for curve fitting. Shading indicates 95% confidence intervals based on 999 bootstrap pseudoreplicates. N = 6 biological replicates.

Figure 4. Ebola virus nucleoprotein interacts with the AP-1 clathrin adaptor complex independent of the nucleoprotein residue 111 allele.

(**A**) IP-MS/MS and STRING analysis of proteins co-immunoprecipitating with NP-myc, with both ancestral R111 and triple charge-reversal (EEE) NP-myc mutants.

(**B**) Reciprocal co-IP of NP with AP-1 and vacuolar protein sorting (VPS) antibodies. Adaptor related protein complex 1 subunit gamma 1 (AP1G1), and mu 1 (AP1M1) are strong interactors of NP, whereas VPS35 is one of several weak interactors.

(**C**) Reciprocal co-IP of all NP mutants and eGFP negative control with α-AP1G1 antibody. No apparent difference was observed between any NP mutant or NP-R111 ancestor.

Figure 5. Ebola virus nucleoprotein position 111 influences viral transcription and replication.

We expressed NP mutants or ancestral NP-R111 in the presence of the EBOV replication complex (L, VP30, VP24), a minigenome (MG) encoding a firefly luciferase reporter gene, and a *Renilla* luciferase loading control. Absence of L, VP30, or minigenome abolished firefly luciferase signal. Both NP-R111E and NP-K109E/K110E/R111E charge-reversal mutants significantly decrease MG activity. p-values are calculated using Dunnett's test. n = 3 biological replicates.

Figure S1. Additional EBOV NP structural data.

(**A**) Electron microscopy subtomogram average (PDB #6EHL) of NP. The R111 residue (yellow) lies in β strand 2 (β-2), anti-parallel to strand β-1 (blue) and the oligomerization domain (orange).

(**B**) Ebolavirus sequence alignment of residues surrounding R111. We compared sequences from the National Center for Biotechnology Information RefSeq database for 6 ebolaviruses and overlaid secondary structure as in another study (Wan et al. 2017). The shaded yellow region from residues 109–111 indicates a patch of basic residues that is relatively well conserved. The oligomerization domain (orange) is nearly completely conserved. A key electrostatic NP-NP interaction between K110 on one NP monomer and E349 on a different NP monomer (Sugita et al. 2018), is also highly conserved.

Figure S2. VLP budding assay.

(**A**) Heating supernatant prior to ultracentrifugation results in loss of VLP luminescence signal. We expressed NLuc-VP40 in cells, collected total supernatant, and heated at a gradient of temperatures. We then either measured luminescence directly to assess NLuc thermal stability (blue), or pelleted VLPs and then measured luminescence to assess VLP stability (tan). Data (n = 3 biological replicates) were normalized to 4 °C, log-transformed, and fit to sigmoidal curves.

(**B**) EM of VP40 VLPs created by with co-expression of GP-A82V and NP mutants or ancestral NP-R111.

(**C**) Quantification of VLP size and volume.

Figure S3. BRET NP oligomerization assay.

Co-immunoprecipitation (co-IP) western blot (WB) NP oligomerization assay. Deletion of NP

residues 20–38 (ΔOD) eliminates oligomerization. We co-expressed myc- and V5-tagged NP or NPΔOD, lysed cells, and performed co-IP targeting either the myc (blue) or the V5 (orange) tags, or IgG isotype controls, followed by detection of co-eluting proteins by western blot (WB). Oligomerization is indicated by heterologous detection (IP myc and WB band for V5, and vice versa) and did not occur with NPΔOD. NP and NPΔOD run at the same apparent molecular weight. The H and L chains of the co-IP antibody were detected by WB using a secondary antibody.

Figure 1

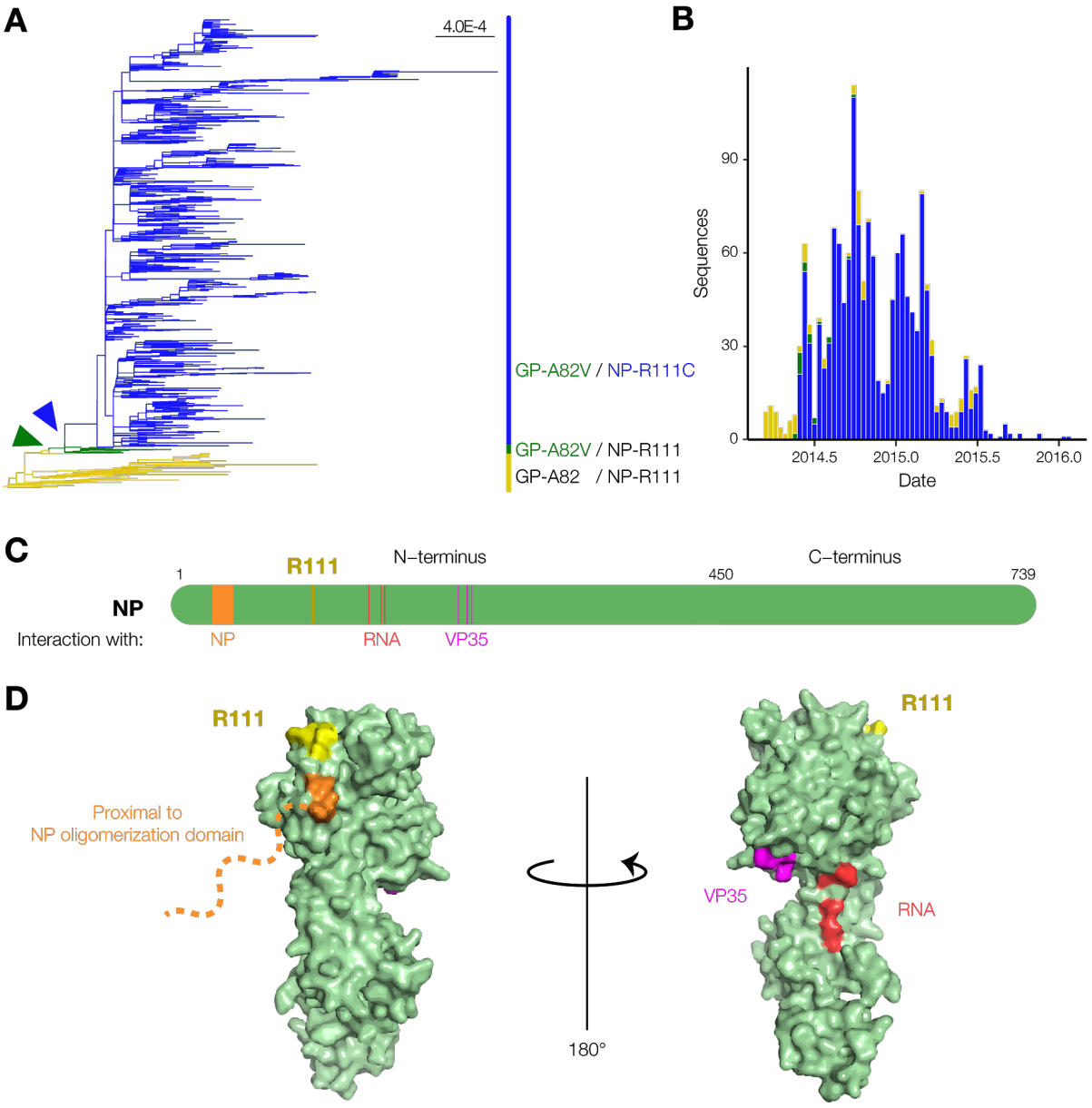


Figure 2

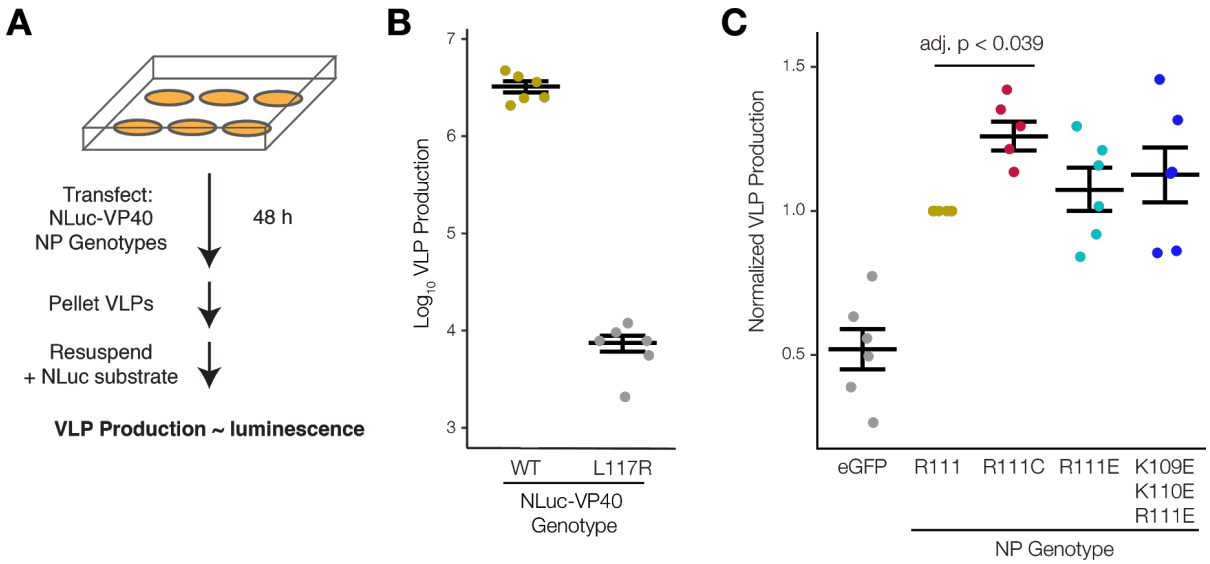
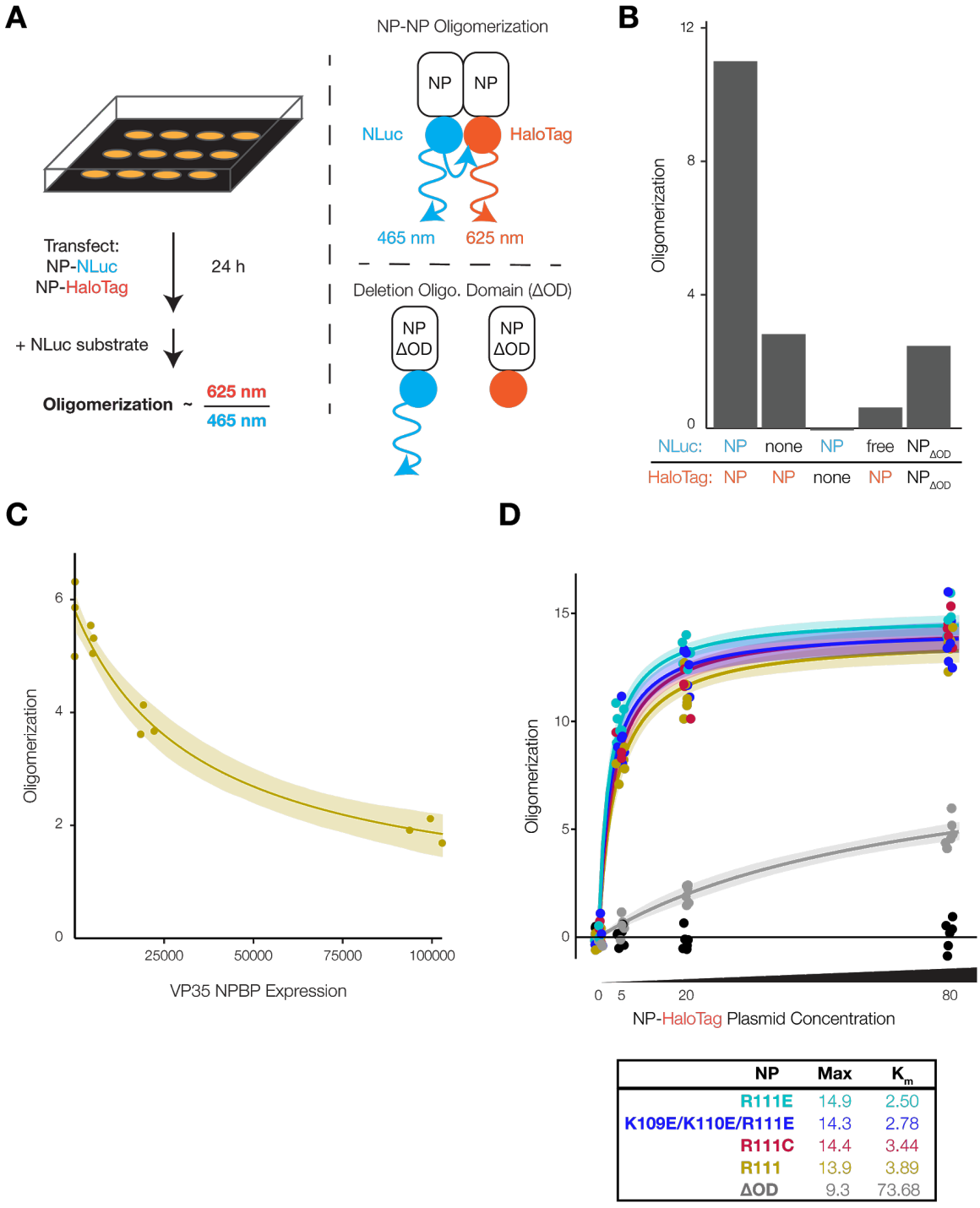


Figure 3



NP	Max	K_m
R111E	14.9	2.50
K109E/K110E/R111E	14.3	2.78
R111C	14.4	3.44
R111	13.9	3.89
ΔOD	9.3	73.68

Figure 4

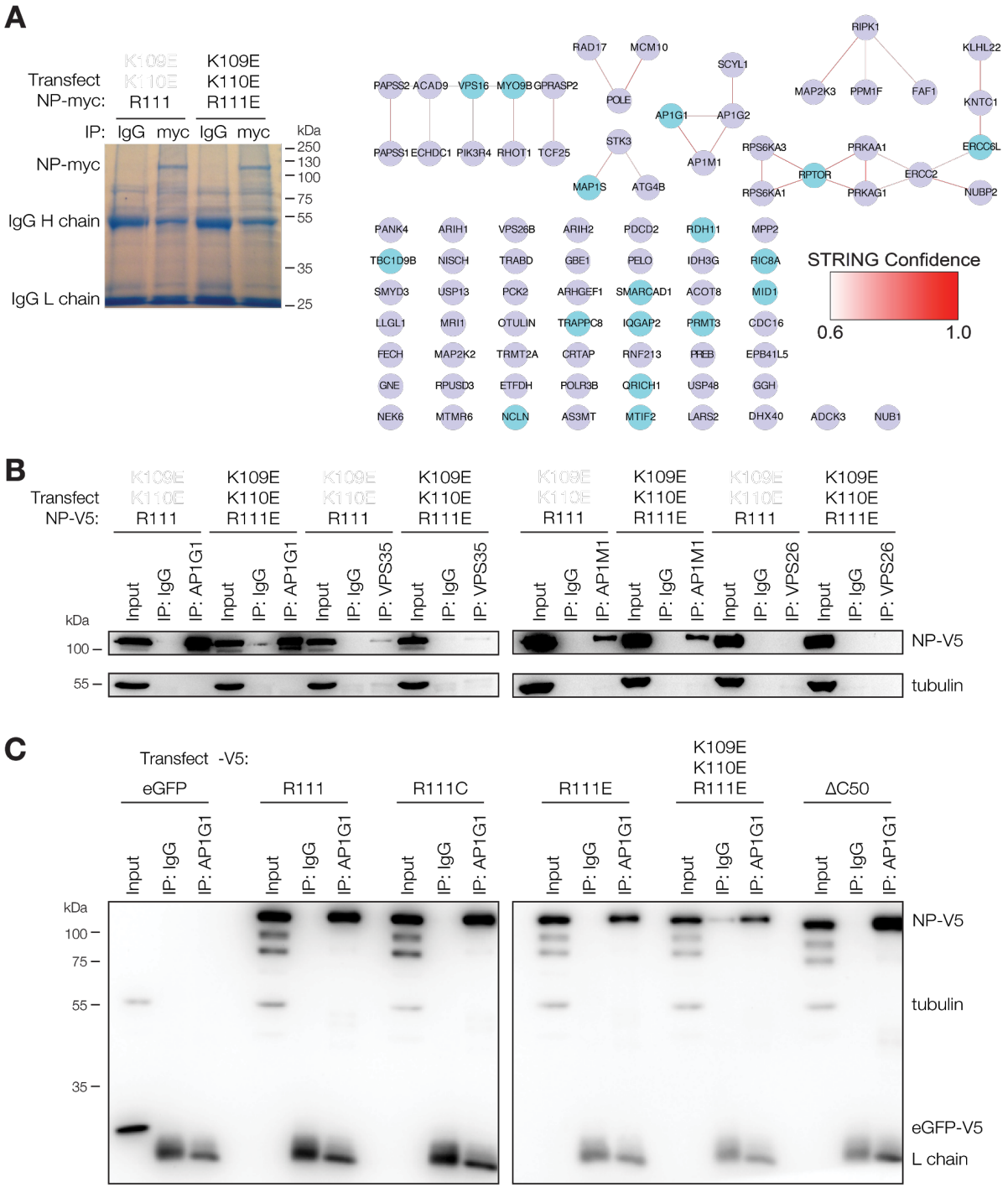


Figure 5

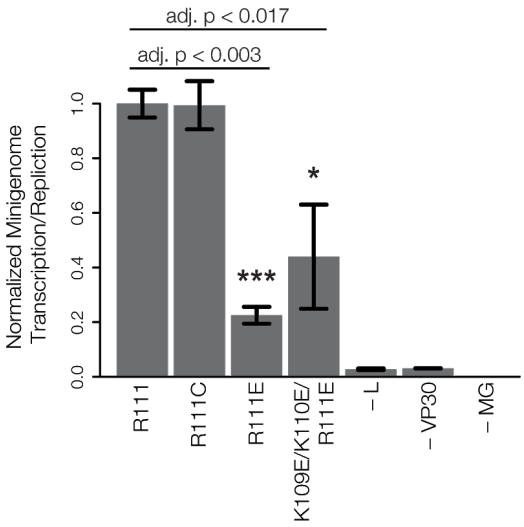
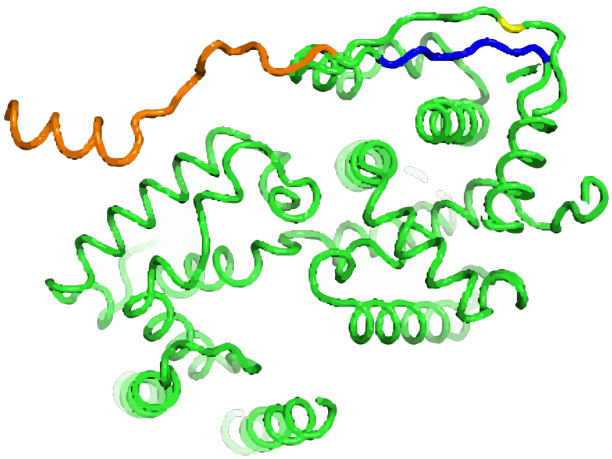


Figure S1

A



B

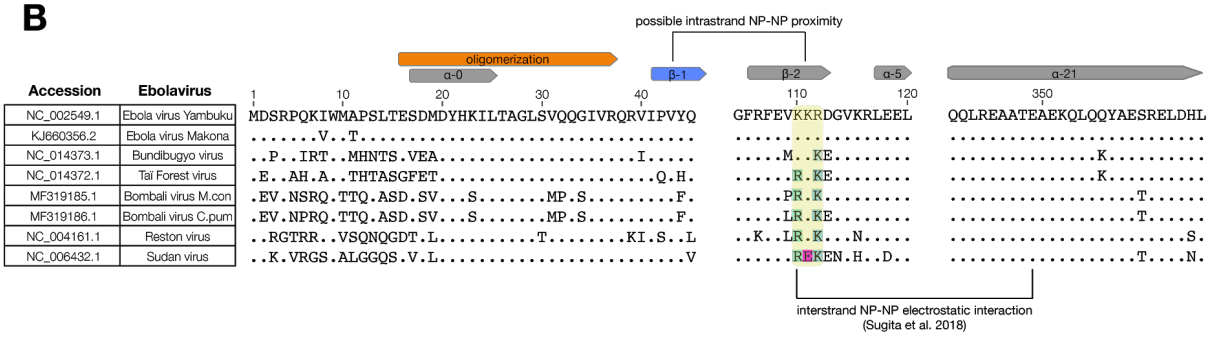


Figure S2

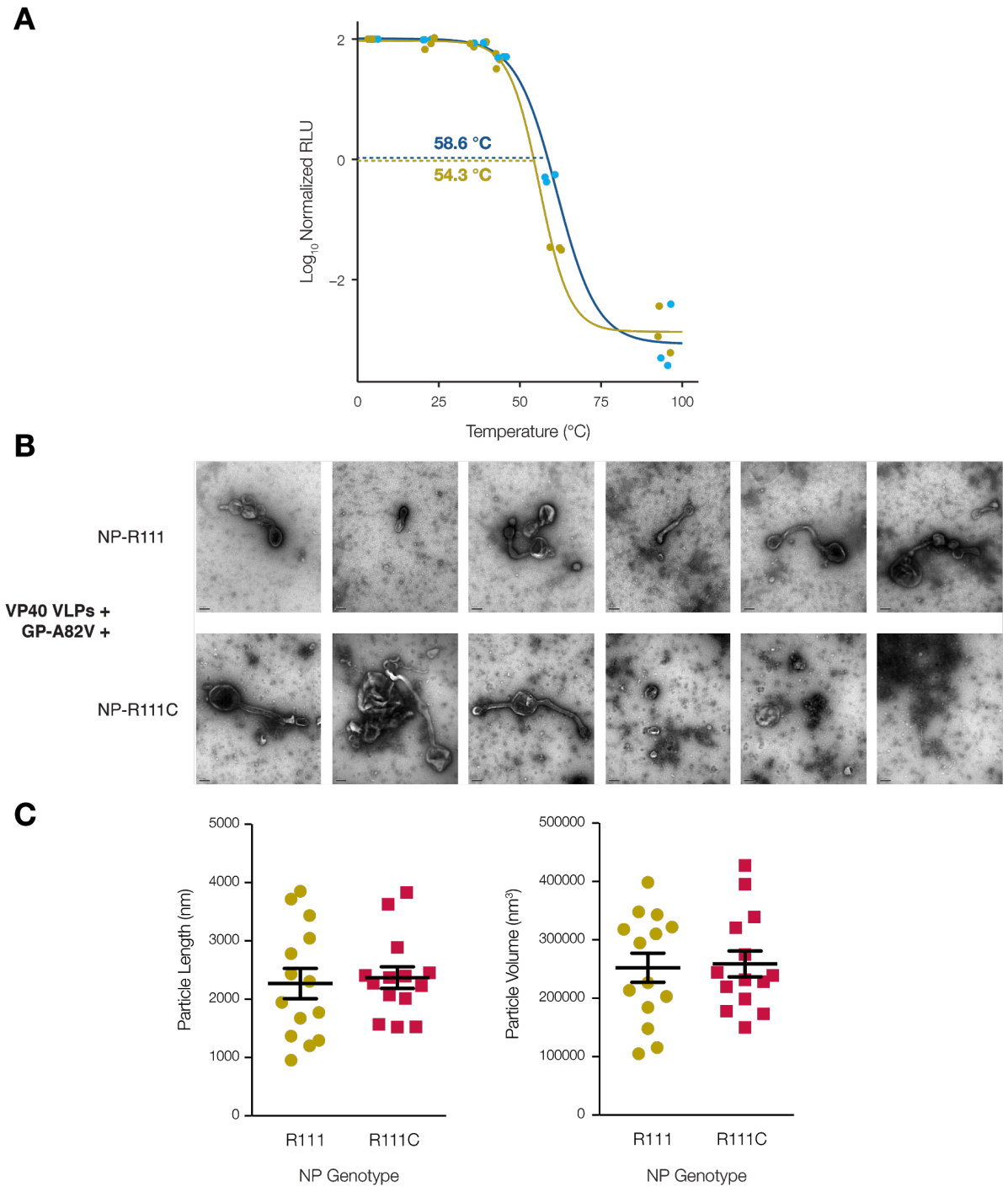
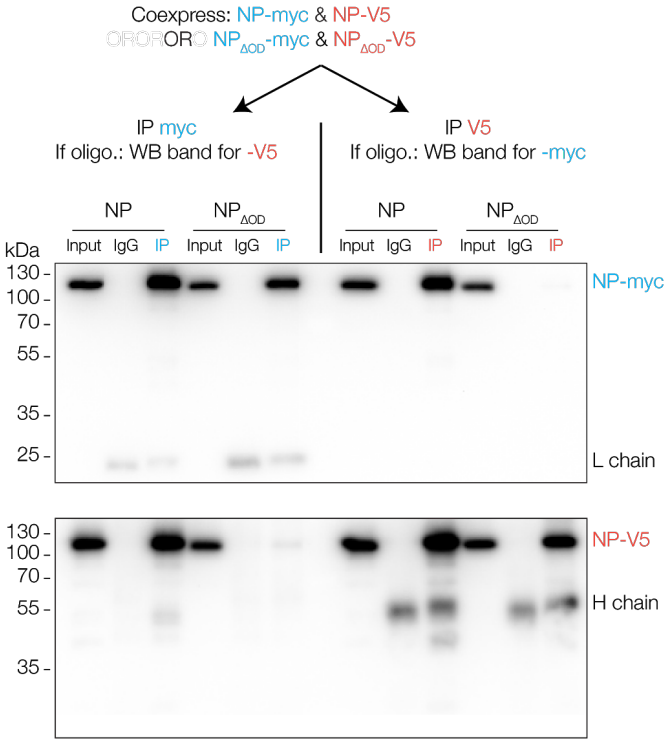


Figure S3

1



2

3