Conserved differences in protein sequence determine the human pathogenicity of Ebolaviruses

Morena Pappalardo, Miguel Juliá, Mark J. Howard, Jeremy S. Rossman, Martin Michaelis, Mark N. Wass

Centre for Molecular Processing and School of Biosciences, University of Kent, Canterbury, Kent CT2 7NJ, UK.

Email: mp465@kent.ac.uk

Abstract-This work describes the analysis of 196 Ebolavirus genomes and the identification of specificity determining positions (SDPs) in all nine Ebolavirus proteins that distinguish the non human pathogenic Reston viruses from the four human pathogenic Ebolaviruses. Structural analysis was performed to identify those SDPs that are likely to have a functional effect. This analysis revealed novel functional insights, in particular for Ebolavirus proteins VP40 and VP24. The VP40 SDP P85T interferes with VP40 function by altering octamer formation. The VP40 SDP Q245P affects the structure and hydrophobic core of the protein and consequently protein function. Three VP24 SDPs (T131S, M136L, Q139R) are likely to impair VP24 binding to human karyopherin alpha5 (KPNA5) and therefore inhibition of interferon signaling. Since VP24 is critical for Ebolavirus adaptation to novel hosts, and only a few SDPs distinguish Reston virus VP24 from VP24 of other Ebolaviruses, human pathogenic Reston viruses may emerge.

I. INTRODUCTION

Four of the five members of the genus Ebolavirus (Ebola viruses, Sudan viruses, Bundibugyo viruses, Taï Forest viruses) cause hemorrhagic fever in humans associated with fatality rates of up to 90% while Reston viruses are non-pathogenic to humans[1,2] (see Materials and Methods for the Ebolavirus nomenclature). So far there have been three Reston virus outbreaks in nonhuman primates: 1989-1990 in Reston Virginia, USA, 1992-1993 in Sienna, Italy, and 1996 in a licensed commercial quarantine facility in Texas. All cases were traced back to a single monkey breeding facility in the Philippines.

During these outbreaks five human individuals were tested positive for IgG antibodies directed against Reston virus. Moreover, Reston virus was found in 2008 in domestic pigs in the Philippines. Seroconversion was detected in six human individuals. None of the 11 individuals that were seropositive for Reston virus antibodies reported an Ebola-like disease [3]. Our large scale analysis of nearly 200 different Ebolavirus genomes focussed on combining computational methods with detailed structural analysis to identify the genetic causes of the difference in pathogenicity between Reston viruses and the human pathogenic Ebolavirus species. Central to our approach was the identification of Specificity Determining Positions

(SDPs), which are positions in the proteome that are conserved within protein subfamilies but differ between them[11,12] and thus distinguish between the different functional specificities of proteins from the different Ebolavirus species. SDPs have been demonstrated to be typically associated with functional sites, such as protein-protein interface sites and enzyme active sites[12].

The SDPs that we have identified and that distinguish Reston viruses from human pathogenic Ebolaviruses, arguably, contain within them a set of amino acid changes that explain the differences in pathogenicity between Reston viruses and the four human pathogenic species, although a contribution of non-coding RNAs (that may exist but remain to be detected) cannot be excluded[6,13]. The subsequent structural analysis was performed to identify the SDPs that are most likely to affect Ebolavirus pathogenicity, using an approach that is similar to those used to investigate candidate single nucleotide variants in human genome wide association and sequencing studies by us and others[14-17].

II. RESULTS

Specificity Determining Positions (SDP) Analysis. 196 Ebolavirus genomes were obtained from the Virus Pathogen Resource (ViPR[18]), consisting of 156 Ebola viruses, 7 Bundibugyo viruses, 13 Sudan viruses, 3 Taï Forest viruses, and 17 Reston viruses (online Methods). Phylogenetic analysis of the whole genomes and the individual proteins separated the Ebolavirus species from each other (Supplementary Figure 1).

In accordance with previous studies[19-23], we observed high intra-species conservation with greater inter-species variation (Figure 1 and Supplementary Table 1). The surface protein GP exhibited the greatest variation (Figure 1), most likely as a consequence of selective pressure exerted by the host immune response[21].

Table 1. SDPs that are likely to alter Reston virus protein structure and function.

Protein	SDP	Interface	Protein Integrity
VP24	T131S	KPNA5 interface	
VP24	M136L	KPNA5 interface	
VP24	Q139R	KPNA5 interface	
VP24	T226A		Loss of Hydrogen bond
VP40	P85T	Octamer interface	
VP40	Q245P		Breaks α helix
VP30	R262A	Dimer interface – loss of Hydrogen bond	
VP35	E269D	Dimer interface	

Using the S3Det algorithm12(Materials and Methods), we identified 189 SDPs that are differentially conserved between Reston viruses and human pathogenic Ebolaviruses (Figure 2, Supplementary Figure 2, Supplementary Tables 2-9). These SDPs represent the most significant changes between the Reston virus and the human pathogenic Ebolaviruses so a subset of these SDPs must explain the difference in pathogenicity. SDPs were present in each of the Ebolavirus proteins representing between 2.4% of residues in sGP to 5.9% of residues in VP30 (Figure 2B). Comparison of the SDPs with previously published mutagenesis studies24 (online Methods) provided no explanation for their functional consequences (Supplementary Table 10).

Structural Analysis. Full-length structures for VP24 and VP40 were available, as well as structures for the globular domains of GP, sGP, NP, VP30, and VP35 (Supplementary Table 11). It was not possible to model the oligeromerization domains of VP30 and VP35 nor the structure of L apart from a short 105 residue segment of the 2239 residue protein, which contained a single SDP. 47 SDPs could be mapped onto Ebolavirus protein structures (or structural models where structures were not available, see online Methods). Most SDPs are located on protein surfaces (Supplementary Figure 3) and are therefore potentially involved in interaction with cellular and viral binding partners and/or immune evasion. Based on our combined computational and structural analysis we find evidence for eight SDPs that are very likely to alter protein structure/function, with six affecting protein-protein interfaces and two that with the potential to influence protein integrity and hence

affect stability, flexibility and conformations of the protein (Table 1). Five additional SDPs may alter protein structure/function but the evidence supporting them is weaker (Supplementary Tables 12-18). Two of these weaker SDPs were present in NP (A705R, R105K - all SDPs are referred to using Ebola virus residue numbering and show the human pathogenic Ebolavirus amino acid first and the Reston virus amino acid second). A705R is likely to introduce a salt bridge with E694 and R105K will alter hydrogen bonding (Supplementary Table 12). The three other SDPs with weaker evidence were present in the glycan cap in GP (see below). The eight confident SDPs were present in V24, VP30, VP35, and VP40. The VP40 and VP24 SDPs revealed the most changes that may relate to differences in human pathogenicity (see below).
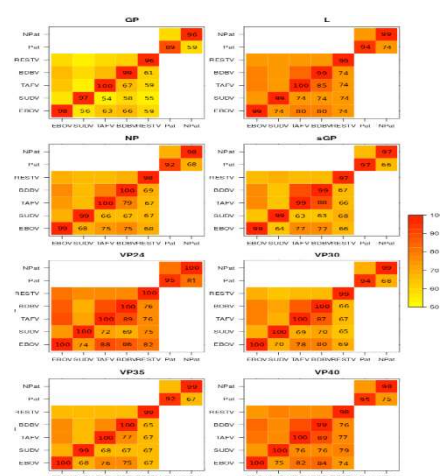

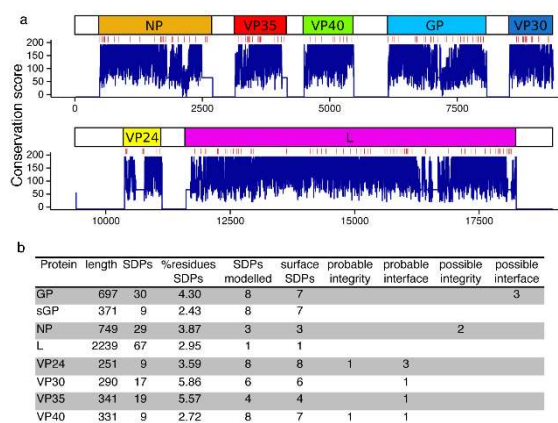Figure 1: Conservation in Ebolavirus proteins

Figure 2 Ebolavirus SDPs

Multiple SDPs are present in the GP glycan cap. GP is highly glycosylated and mediates Ebolavirus host cell entry. Subunit GP1 binds to the host cell receptor(s). Subunit GP2 is responsible for the fusion of viral and host cell membranes. However, their cellular binding partners remain to be defined1,25-27. Reverse genetics experiments have suggested that GP contributes to human pathogenicity but is insufficient for virulence on its own28. We identified SDPs in both GP1 and GP2 (Supplementary Figure 4 and Supplementary Table 12). Three SDPs (I260L, T269S, S307H) are located in the glycan cap that contacts the host cell membrane (Supplementary Figure 4B-C). These changes (particularly S307H at the top of the glycan cap) alter the electrostatic surface of GP (Supplementary Figure 4D) and may therefore alter GP interactions with cellular proteins, however given the glycosylation of GP, it is unlikely that these residues would physically contact the host cell membrane and none of them are near glycosylation sites. So it is not clear what role they may have. GP binding to the endosomal membrane protein NPC1 is necessary for membrane fusion25. However, residues important for NPC1 binding (identified by mutagenesis studies in25) were conserved in all analyzed Ebolaviruses and the SDPs were not located close to them (Supplementary Figure 5). Thus differences in NPC1 binding do not account for differences in Ebolavirus human pathogenicity. This finding is in concert with very recent data indicating that NPC1 is essential for Ebolavirus replication as NPC1-deficient mice were insusceptible to Ebolavirus infection27. It was not possible to predict the consequences of SDPs in sGP and ssGP (Fig. S23), as there is a lack of functional information available for these proteins3,4. A 17 amino acid peptide derived from Ebola virus or Sudan virus GP exerted immunosuppressive effects on human CD4+ T cells and CD8+ T cells while the

respective Reston virus peptide did not29. We identified one SDP in the peptide, which represents the single amino acid change (I604L) previously observed between Reston virus and Ebola virus29, demonstrating that this difference is conserved between Reston viruses and all human pathogenic Ebolaviruses.
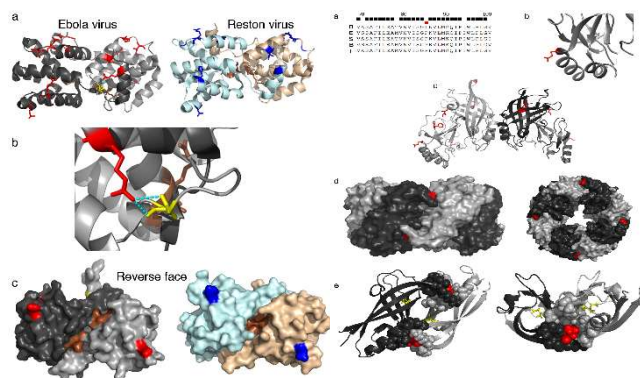

Figure 3-4: SDPs in protein VP30 and in protein VP40

Changes in the VP30 dimer may affect pathogenicity. Analysis of the VP30 SDPs provided novel mechanistic insights into the structural differences previously observed between Reston virus and Ebola virus VP3010 and that may contribute to the differences observed in human pathogenicity between Reston virus and Ebola virus. VP30 is an essential transcriptional co-factor that forms dimers via its C-terminal domain and hexamers via an oligomerization domain (residues 94-112)30.

The VP30 hexamers activate transcription while the dimers do not, and the balance of hexamers and dimers has been suggested to control the balance between transcription and replication31. Crystallization studies have shown that Ebola virus and Reston virus dimers are rotated relative to each other10. We observed two SDPs (T150I, R262A) in the dimer interface that can at least partially explain the structural differences between Ebola virus and Reston virus VP30 dimers. Ebola virus R262 is part of the dimer interface and forms a hydrogen bond with the backbone of residue 141 in the other subunit, whereas Reston A262 does not and is not part of the dimer interface (Figure 3). The removal of the two hydrogen bonds (in the symmetrical dimer) is likely to lead to the different Reston and Ebola virus dimer structures. mCSM predicts this change to be destabilizing with a $\Delta\Delta G$ -0.969 Kcal/mol. T

he Reston virus conformation also buries functional residues A179 and K180 potentially affecting protein function10 (Figure 2). Moreover, our findings show that the Ebola virus confirmation is

conserved in all human-pathogenic Ebolaviruses suggesting that it is relevant for human pathogenicity.

VP35 SDP present in dimer interface. VP35 is a multifunctional protein that antagonizes interferon signaling by binding double stranded RNA (dsRNA). Structural data are available for both the Ebola virus and Reston virus VP35 monomer and an asymmetric dsRNA bound dimer 9,32-34. These structures are highly conserved, however functional studies have demonstrated that Reston virus VP35 is more stable, has a reduced affinity for dsRNA, and exerts weaker effects on interferon signaling32. The increased stability is proposed to be due to a linker between the two subdomains having a short alpha helix in the Reston virus structure32. Our analysis shows that the sequence of this linker region is completely conserved in all of the genomes, however an SDP is located close to the linker (A290V). One SDP (E269D) is present in the dimer interface and the shorter aspartate side chain in Reston virus VP35 results in increased distances with the atoms that this aspartate forms hydrogen bonds with: R312, R322, and W324 (Ebola virus numbering; Supplementary Table 13). mCSM predicts this change to be slightly destabilizing to the complex ($\Delta\Delta$G -0.11Kcal/mol). This has the potential to alter the stability of the dimer and thus the ability of VP35 to prevent interferon signaling. It has recently been demonstrated that a VP35 peptide binds NP and modulates NP oligomerization and RNA binding to NP35. There are two SDPs (S26T, E48D) in this region. S26T is located on the periphery of the interface. E48D lies outside the solved structure but is within the region required for binding to NP. Both SDPs represent minor changes that maintain the chemical properties of the side chains. Thus, there is no evidence suggesting substantial differences in the binding of this peptide to NP.

VP40 SDPs may alter oligomeric structure. VP40 exists in three known oligomeric forms36. Dimeric VP40 is responsible for VP40 trafficking to the cellular membrane. Hexameric VP40 is essential for budding and forms a filamentous matrix structure. Octameric VP40 regulates viral transcription by binding RNA. Two SDPs (P85T and Q245P) can affect VP40 structure. P85T occurs at the VP40 octamer interface site (Figure 4) in the middle of a run of 14 residues that are completely conserved in all Ebolaviruses (Figure 4B). In the Ebola virus structure, it is located in an S-G-P-K beta-turn, where the proline at position 85 (P85) confers backbone rigidity. The change to threonine (T) at this residue in Reston viruses introduces backbone flexibility and also provides a side chain with a hydrogen bond donor, potentially

affecting octamer structure and/or formation. mCSM predicted this change to have a destabilizing effect ($\Delta\Delta$G -0.626Kcal/mol). The Q245P SDP introduces a proline residue into an alpha helix (Figure 4B), which most likely breaks and shortens helix five, resulting in the destabilization of helices five and six and a change in the hydrophobic core. Interestingly mCSM predicted this change to have little effect on the stability of the protein (predicted $\Delta\Delta$G 0.059Kcal/mol). Thus, P85T and Q245P may affect VP40 function and human pathogenicity.
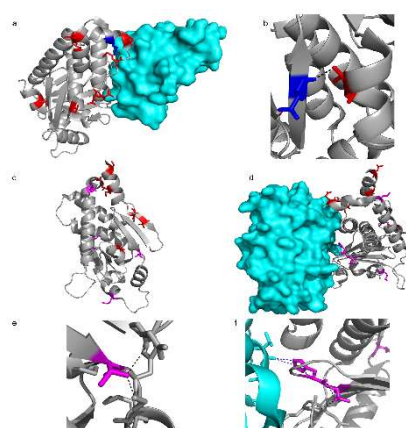

Figure 5: SDPs in protein VP24

VP24 SDPs affect KPNA5 binding. VP24 is involved in the formation of the viral nucleocapsid and the regulation of virus replication1,19,37-39. VP24 also interferes with interferon signaling through binding of the karyopherins α1 (KPNA1), α5, (KPNA5), and α6 (KPNA6) and subsequent inhibition of nuclear accumulation of phosphorylated STAT1 and through direct interaction with STAT124,40-42. Eight VP24 SDPs are in regions with available structural information (Supplementary Tables 17-18). Seven of these are present on the same face of VP24 (Figure 5A) suggesting that they affect VP24 interaction with viral and/or host cell binding partners. The SDPs T131S, M136L, and Q139R are present in the KPNA5 binding site (Figure 5). M136 and Q139 are part of multi-residue mutations in Ebola virus VP24 that removed KPNA5 interactions (Supplementary Table 17)24 and are adjacent to K142 (Figure 5A), mutants of which have shown reduced interferon antagonism43. Therefore, M136L and Q139R can exert significant effects on VP24-KPNA5 binding. Additionally, T226A results in the loss of a hydrogen bond between T226 and D48 in Reston virus VP24 (Figure 5B), with the potential to alter structural integrity and influence protein function. Analysis

using mCSM predicts the T226A change to be destabilizing with a $\Delta\Delta G$ -0.935 Kcal/mol. mCSM predicted seven of the eight analysed SDPs to be destabilizing (Supplementary Table 2). VP24-mediated inhibition of interferon signaling may be critical for species-specific pathogenicity[24,38,40-42]. In this context, VP24 was a critical determinant of pathogenicity in studies in which Ebola viruses were adapted to mice and guinea pigs that are normally insusceptible to Ebola virus disease[5,38,44-46]. The adaptation-associated VP24 mutations in rodents are located in the KPNA5 binding site with some of them being very close to the VP24 SDPs T131S, M136L, and Q139R that we determined to be in the KPNA5 binding site (Figure 5C-D, Supplementary Table 19). Additionally some of the mutations are similar to the SDPs in that they would remove hydrogen bonds within VP24 (e.g. T187I, T50I, Figure 5E-F, & Supplementary Table 19) or alter hydrogen bonding with KPNA5 (H186Y, Figure 5F & Supplementary Table 19). Thus there is strong evidence suggesting that the VP24 SDPs have a role in rendering the Reston virus non-pathogenic in humans.

III.DISCUSSION

In this study, we have combined the computational identification of residues that distinguish Reston viruses from human pathogenic Ebolavirus species with protein structural analysis to identify determinants of Ebolavirus pathogenicity. The results from this first comprehensive comparison of all available genomic information on Reston viruses and human pathogenic Ebolaviruses detected SDPs in all proteins but only few of them may be responsible for the lack of Reston virus human pathogenicity. Our analysis mapped 47 of the 189 SDPs onto protein structure, so additional SDPs may be relevant but the structural data needed to reliably identify them is missing. Although it is difficult to conclude the extent to which each individual SDP contributes to the differences in human pathogenicity between Reston viruses and the other Ebolaviruses, we can identify certain SDPs that have a particularly high likelihood to be involved. SDPs present in the oligomer interfaces of VP30, VP35, and VP40 may affect viral protein function. VP24 SDPs may interfere with VP24-KPNA5 binding and affect viral inhibition of the host cell interferon response. These findings suggest that changes in protein-protein interactions represent a central cause for the variations in human pathogenicity observed in Ebolaviruses. VP24 and VP40 in particular contain multiple SDPs that are likely to contribute to differences in human pathogenicity. Where possible the SDPs have been considered collectively, such as for VP24, where most of the SDPs are present on a single face of the protein (Figure 5A) and three of them are present in the interface with KPNA5. Beyond this it is difficult to interpret how any combination of SDPs might be responsible for the differences in human pathogenicity. Our data also demonstrate that relevant changes explaining differences in virulence between closely related viruses can be identified by computational analysis of protein sequence and structure. Such computational studies are particularly important for the investigation of Risk Group 4 pathogens like Ebolaviruses whose investigation is limited by the availability of appropriate containment laboratories. The role of VP24 appears to be central given the large number of SDPs we identify as likely to affect function, particularly KPNA5 binding. This is also highlighted by the similarity between these SDPs and the mutations that occur in adaptation experiments in mice and guinea pigs[6,33,39-41]. Consequently, the mutation of a few VP24 SDPs could result in a human pathogenic Reston virus. Given that Reston viruses circulate in domestic pigs, can be spread by asymptomatically infected pigs, and can be transmitted from pigs to humans (possibly by air)[2,47,48], there is a concern that (a potentially airborne) human pathogenic Reston viruses may emerge and pose a significant health risk to humans. Notably, asymptomatic Ebolavirus infections have also been described in dogs[2] and Ebola virus shedding was found in an asymptomatic woman[49]. Thus, there may be further unanticipated routes by which Reston viruses may spread in domestic animals and/or humans enabling them to adapt and cause disease in humans. In summary our combined computational and structural analysis of a large set of Ebolavirus genomes has identified amino acid changes that are likely to have a crucial role in altering Ebolavirus pathogenicity. In particular the differences in VP24 together with the observation that Ebolavirus adaptation to originally non-susceptible rodents results in rodent pathogenic viruses[6,33,39-41] suggest that a few mutations could lead to a human pathogenic Reston virus.

REFERENCES

1. Feldmann, H. & Geisbert, T. W. Ebola haemorrhagic fever. *Lancet* **377,** 849–862 (2011).

2. Weingartl, H. M., Nfon, C. & Kobinger, G. Review of Ebola virus infections in domestic animals. *Dev Biol (Basel)* **135,** 211–218(2013).

3. Miranda, M. E. G. & Miranda, N. L. J. Reston ebolavirus in humans and animals in the Philippines: a review. *J. Infect. Dis.* **204** Suppl3, S757–60 (2011).

4. Mehedi, M. *et al.* A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *J. Virol.* **85,** 5406–5414 (2011).

5. La Vega, de M.-A., Wong, G., Kobinger, G. P. & Qiu, X. The multiple roles of sGP in Ebola pathogenesis. *Viral Immunol.* **28,** 3–9 (2015).

6. Basler, C. F. Portrait of a killer: genome of the 2014 EBOV outbreak strain. *Cell Host Microbe* **16,** 419–421 (2014).

7. Hoenen, T. *et al.* Soluble Glycoprotein Is Not Required for Ebola Virus Virulence in Guinea Pigs. *J. Infect. Dis.* jiv111, doi: 10.1093/infdis/jiv111 (2015).

8. Zhang, A. P. P. *et al.* The ebola virus interferon antagonist VP24 directly binds STAT1 and has a novel, pyramidal fold. *PLoS Pathog.* **8,** e1002550 (2012).

9. Bale, S. *et al.* Ebolavirus VP35 coats the backbone of double-stranded RNA for interferon antagonism. *J. Virol.* **87,** 10385–10388 (2013).

10. Clifton, M. C. *et al.* Structure of the Reston ebolavirus VP30 C-terminal domain. *Acta Crystallogr F Struct Biol Commun* **70,** 457–460 (2014).

11. Casari, G., Sander, C. & Valencia, A. A method to predict functional residues in proteins. *Nat Struct Biol* **2,** 171–178 (1995).

12. Rausell, A., Juan, D., Pazos, F. & Valencia, A. Protein interactions and ligand binding: From protein subfamilies to functional specificity. *Proc. Natl. Acad. Sci. USA* **107,** 1995–2000 (2010).

13. Teng, Y. *et al.* Systematic Genome-wide Screening and Prediction of microRNAs in EBOV During the 2014 Ebolavirus Outbreak.*Sci Rep* **5,** 9912 (2015).

14. Chambers, J. C. *et al.* Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *NatGenet* **43,** 1131–1138 (2011).

15. Chambers, J. C. *et al.* Genetic loci influencing kidney function and chronic kidney disease. *Nat Genet* **42,** 373–375 (2010).

16. Chambers, J. C. *et al.* The South Asian genome. *PLos One* **9,** e102645 (2014).

17. Palles, C. *et al.* Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas andcarcinomas. *Nat Genet* **45,** 136–144 (2013). Scientific **Reports** | 6:23743 | DOI: 10.1038/srep23743 10

18. Pickett, B. E. *et al.* ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res.* **40,** D593–8 (2012).

19. Morikawa, S., Saijo, M. & Kurane, I. Current knowledge on lower virulence of Reston Ebola virus (in French: Connaissancesactuelles sur la moindre virulence du virus Ebola Reston). *Comparative Immunology, Microbiology and Infectious Diseases* **30,** 391–398 (2007).

20. Gire, S. K. *et al.* Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science* **345,** 1369–1372 (2014).

21. Liu, S.-Q., Deng, C.-L., Yuan, Z.-M., Rayner, S. & Zhang, B. Identifying the pattern of molecular evolution for Zaire ebolavirus in the 2014 outbreak in West Africa. *Infect Genet Evol* **32,** 51–59 (2015).

22. Vogel, G. Infectious Diseases. A reassuring snapshot of Ebola. *Science* **347,** 1407–1407 (2015).

23. Hoenen, T. *et al.* Virology. Mutation rate and genotype variation of Ebola virus from Mali case sequences. *Science* **348,** 117–119 (2015).

24. Xu, W., Edwards, M. R., Borek, D. M., Feagins, A. R. & Mittal, A. Ebola Virus VP24 Targets a Unique NLS Binding Site on Karyopherin Alpha 5 to Selectively Compete with Nuclear Import of Phosphorylated STAT1. *Cell Host Microbe* **13,** 187–200 (2014).

25. Miller, E. H. *et al.* Ebola virus entry requires the host-programmed recognition of an intracellular receptor. *EMBO J.* **31,** 1947–1960 (2012).

26. Dahlmann, F. *et al.* Analysis of Ebola Virus Entry Into Macrophages. *J. Infect. Dis.* jiv140, doi: 10.1093/infdis/jiv140 (2015).

27. Herbert, A. S. *et al.* Niemann-pick c1 is essential for ebolavirus replication and pathogenesis *in vivo*. *MBio* **6,** e00565–15 (2015).

28. Groseth, A. *et al.* The Ebola virus glycoprotein contributes to but is not sufficient for virulence *in vivo*. *PLos Pathog.* **8,** e1002847 (2012).

29. Yaddanapudi, K. *et al.* Implication of a retrovirus-like glycoprotein peptide in the immunopathogenesis of Ebola and Marburg viruses. *FASEB J.* **20,** 2519–2530 (2006).

30. Hartlieb, B., Modrof, J., Muhlberger, E., Klenk, H.-D. & Becker, S. Oligomerization of Ebola virus VP30 is essential for viral transcription and can be inhibited by a synthetic peptide. *J. Biol. Chem.* **278,** 41830–41836 (2003).

31. Hartlieb, B., Muziol, T., Weissenhorn, W. & Becker, S. Crystal structure of the C-terminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association. *Proc. Natl. Acad. Sci. USA* **104,** 624–629 (2007).

32. Leung, D. W. *et al.* Structural and functional characterization of Reston Ebola virus VP35 interferon inhibitory domain. *J. Mol. Biol.* **399,** 347–357 (2010).

33. Leung, D. W. *et al.* Structure of the Ebola VP35 interferon inhibitory domain. *Proc. Natl. Acad. Sci. USA* **106,** 411–416 (2009).

34. Kimberlin, C. R. *et al.* Ebolavirus VP35 uses a bimodal strategy to bind dsRNA for innate immune suppression. *Proc. Natl. Acad. Sci. USA* **107,** 314–319 (2010).

35. Leung, D. W. *et al.* An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis by Modulating Nucleoprotein-RNA Interactions. *Cell Rep* doi: 10.1016/j.celrep.2015.03.034 (2015).

36. Bornholdt, Z. A. *et al.* Structural rearrangement of ebola virus VP40 begets multiple functions in the virus life cycle. *Cell* **154,** 63–774 (2013).

37. Mateo, M. *et al.* Knockdown of Ebola virus VP24 impairs viral nucleocapsid assembly and prevents virus replication. *J. Infect. Dis.* **204** Suppl 3, S892–6 (2011).

38. Mateo, M. *et al.* VP24 is a molecular determinant of Ebola virus virulence in guinea pigs. *J. Infect. Dis.* **204** Suppl 3, S1011–20 (2011).

39. Watt, A. *et al.* A novel life cycle modeling system for Ebola virus shows a genome length-dependent role of VP24 in virus infectivity. *J. Virol.* **88,** 10511–10524 (2014).

40. Reid, S. P. *et al.* Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation. *J. Virol.* **80,** 5156–5167 (2006).

41. Reid, S. P., Valmas, C., Martinez, O., Sanchez, F. M. & Basler, C. F. Ebola virus VP24 proteins inhibit the interaction of NPI-1 subfamily karyopherin alpha proteins with activated STAT1. *J. Virol.* **81,** 13469–13477 (2007).

42. Zhang, A. P. P. *et al.* The ebolavirus VP24 interferon antagonist: know your enemy. *Virulence* **3,** 440–445 (2012).

43. Ilinykh, P. A. *et al.* Different temporal effects of Ebola virus VP35 and VP24 proteins on the global gene expression in human dendritic cells. *J. Virol.* JVI. 00924–15, doi: 10.1128/JVI.00924-15 (2015).

44. Volchkov, V. E., Chepurnov, A. A., Volchkova, V. A., Ternovoj, V. A. & Klenk, H. D. Molecular characterization of guinea pigadapted variants of Ebola virus. *Virology* **277,** 147–155 (2000).

45. Ebihara, H. *et al.* Molecular determinants of Ebola virus virulence in mice. *PLos Pathog.* **2,** e73 (2006).

46. Dowall, S. D. *et al.* Elucidating variations in the nucleotide sequence of Ebola virus associated with increasing pathogenicity. *Genome Biol.* **15,** 540 (2014).

47. Barrette, R. W. *et al.* Discovery of swine as a host for the Reston ebolavirus. *Science* **325,** 204–206 (2009).

48. Marsh, G. A. *et al.* Ebola Reston virus infection of pigs: clinical significance and transmission potential. *J. Infect. Dis.* **204** Suppl 3,S804–9 (2011).

49. Akerlund, E., Prescott, J. & Tampellini, L. Shedding of Ebola Virus in an Asymptomatic Pregnant Woman. *N. Engl. J. Med.* **372,**2467–2469 (2015).

50. Kuhn, J. H. *et al.* Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virusabbreviations. *Archives of Virology* **155,** 2083–2103 (2010).51. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* **16,** 276–277 (2000).

52. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.***7,** 539 (2011).
53. Mistry, J. *et al.* Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions *Nucleic Acids Res.* **41,**e121 (2013).

54. UniProt Consortium. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **42,** D191–8 (2014).
55. Bouckaert, R. *et al.* BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **10,** e1003537 (2014).

56. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30,**1312–1313 (2014).

57. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20,** 289–290 (2004).

58. Rose, P. W. *et al.* The RCSB Protein Data Bank: views of structural biology for basic and applied research and education. *NucleicAcids Res.* **43,** D345–56 (2015).

59. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction andanalysis. *Nat Protoc* **10,** 845–858 (2015).

60. Joosten, R. P. *et al.* A series of PDB related databases for everyday needs. *Nucleic Acids Res.* **39,** D411–9 (2011).

61. Smith, N. *et al.* DelPhi web server v2: incorporating atomic-style geometrical figures into the computational protocol. *Bioinformatics***28,** 1655–1657 (2012).

62. Pettersen, E. F. *et al.* UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **25,** 1605–1612(2004).

63. Pires, D. E. V., Ascher, D. B. & Blundell, T. L. mCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics 30, 335–342 (2014)