Did a Single Amino Acid Change Make Ebola Virus More Virulent?

Trevor Bedford[1,*] and Harmit S. Malik[2,3]
[1]Vaccine and Infectious Disease Division
[2]Basic Sciences Division
[3]Howard Hughes Medical Institute
Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
*Correspondence: tbedford@fredhutch.org
http://dx.doi.org/10.1016/j.cell.2016.10.032

A mutation in the Ebola virus glycoprotein arose early during the 2013–2016 epidemic and dominated the viral population. Two studies by Diehl et al. and Urbanowicz et al. now reveal that this mutation is associated with higher infectivity to human cells, representing the clearest example of Ebola's functional adaptation to human hosts.

Since its original detection 40 years ago, the Ebola virus has been on epidemiologists' radar due to its virulence and high mortality index. Nevertheless, the West African Ebola epidemic of 2013–2016 was unprecedented in human impact, eventually resulting in more than 28,000 cases and 11,000 deaths and exhausting the public health system of multiple countries. After initial introduction into Guinea in late 2013, ongoing human-to-human transmission allowed the virus to evolve and accumulate mutations relative to this initial infection (Gire et al., 2014). In early 2014, the viral glycoprotein (GP) mutant A82V appeared and spread, eventually coming to dominate the viral population with more than 90% of sequenced viruses bearing the A82V mutation (Figure 1). In this issue of *Cell,* two complementary studies characterize the infectivity patterns of A82V across different cell types (Diehl et al., 2016; Urbanowicz et al., 2016) and other GP mutants (Urbanowicz et al., 2016). Both studies find a clear pattern of increased infectivity in human cells, consistent with a fitness advantage of 82V over 82A. This represents the clearest example to date of functional adaptation of Ebola virus to the human host after spill-over from its animal reservoir.

Both studies build off of three important molecular clues that emerged from earlier work by several labs. First was the unambiguous identification of the NPC1 protein (Nieman-Pick Disease protein) as the bona fide receptor for Ebola virus (Carette et al., 2011; Côté et al., 2011). Second was the identification of the binding interface between the GP protein of Ebolaviruses and NPC1 via structural studies (Bornholdt et al., 2016; Gong et al., 2016; Wang et al., 2016). Third was the discovery that this battle for binding affinity had led to rapid evolution at the interaction interface, such that NPC1s from different mammals are differentially susceptible to binding and entry by Ebolaviruses (Ng et al., 2015). This suggested the possibility that, in the course of the human epidemic, the Ebolavirus GP protein could change its NPC1 binding preference from reservoir species (most likely fruit bats) to improve binding to human NPC1. This possibility was further suggested by the fact that a single amino acid change that mapped right in the middle of the presumed NPC1 binding interface spread widely in the West African epidemic. Armed with this important insight, both research teams took slightly different tactics to investigate the case for Ebola virus adaptation to their (new) human hosts.

The first study (Diehl et al., 2016) focused primarily on the A82V mutation (the authors also studied and discarded additional mutations T230A and D637G that occurred later in the A82V lineage). Using pseudotyping assays in which a reporter lentivirus' glycoprotein is switched with a test GP—in this case from Ebolavirus—the authors showed that the single A82V mutation is able to confer increased infectivity in a variety of primate cells, including human dendritic cells, which are a target for Ebolavirus during the course of an infection. Furthermore, they show that this increased infectivity by A82V is a direct result of virus entry into host cytoplasm and not due to an intrinsic protein processing or stability determinant. Intriguingly, this increased infectivity is only observed in primate cells, which share a conserved NPC1 domain that interacts with Ebola GP. In contrast, this mutation does not impact infection in rodent or carnivore cells, making the strong case for a primate- (and likely human-) specific adaptation, although Diehl et al. (2016) do not directly examine infectivity in bat cells. To examine the epidemiological consequences of this mutation in the context of the epidemic, the authors then examined whether the A82V mutation conferred increased viremia and mortality. Indeed, they find that there was an increased mortality associated with A82V, although the effect size is modest, reflecting the inherent difficulty of evaluating an ongoing epidemic rather than a controlled trial.

The second study (Urbanowicz et al., 2016) explicitly tested for infectivity differences in Ebolavirus infectivity of cells from the presumed reservoir hosts, fruit bats and humans. This experimental design and use of slightly different human cell lines allowed them to uncover more subtle adaptations in the transition from bat to human hosts. As a result, Urbanowicz et al. (2016) find that, in addition to a large effect of A82V, there were additional increases in primate-specific infectivity associated with subsequent changes in GP: R29V, T206M, and T230A. Studying

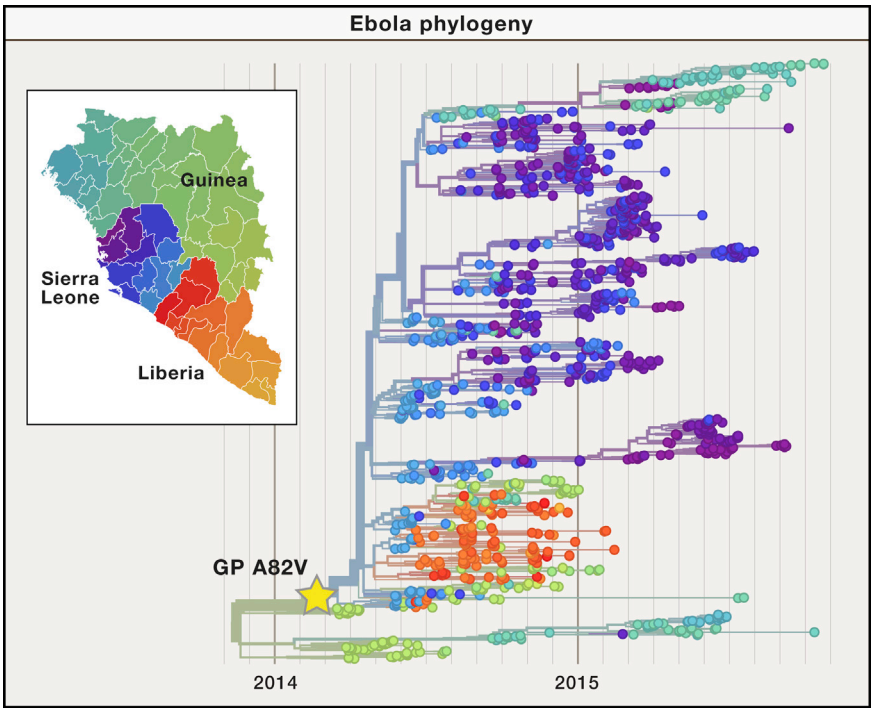

Figure 1. Evolutionary Course of the 2013–2016 Ebola Epidemic

A time-resolved phylogenetic tree showing the evolutionary relationships among 1,261 Ebola viruses sampled during the 2013–2016 West African epidemic is depicted. Viruses (circles) are colored according to geographic division within Guinea (green hues), Sierra Leone (blue hues), and Liberia (orange hues). Internal branches are colored based on phylogeographic reconstruction given sampled locations and evolutionary relationships. This reconstruction estimates that the glycoprotein (GP) 82 mutation occurred around February 2014 in the Guéckédou prefecture of Guinea but immediately spread to the Kailahun district of Sierra Leone and from there drove the majority of overall epidemic. This phylogenetic reconstruction is taken from the website http://www.nextstrain.org/ebola/ by Richard Neher and Trevor Bedford that has maintained a real-time view of Ebola evolution (subject to data availability) since June 2015.

mutations that occurred independently in multiple sub-lineages, the authors are also able to identify convergent evolution in a pair of mutations—P330S and G480D—that together further increase infectivity in the ancestral 82A lineage. The authors speculate that still other changes (I371V and P375S) may be required to epistatically compensate for any folding or conformational defects that may be associated with A82V. These findings are interesting because they argue that adaptation was not a single event. Although the original A82V was important for increased human infectivity, Ebola virus lineages continued to adapt and evolve to their human hosts, in some cases employing the same adaptive path for increased virulence in humans. Finally, they show a clear trade-off in terms of use of host NPC1; many of the mutations associated with primate-specific adaptation are associated with a loss of infectivity in bat cells. This further makes the case for primate-specific adaptation because these mutations would be less tolerated during infections of reservoir species. Intriguingly, the new findings also allow the authors to bolster the case that fruit bats, and not insectivorous bat species, are indeed the reservoir species for Ebola viruses; the latter are non-permissive for all of the variants tested.

It would appear that both studies have the "smoking gun" evidence to make the case for molecular adaptation in Ebola leading to increased human virulence. In many respects, the dense sampling and sequencing of Ebola genomes due to the herculean efforts of many researchers (some of whom were fatally infected during their studies) makes this perhaps one of the best examples of "catching the virus in the act." However, as both research teams point out, although it seems unarguably clear that the A82V mutation affects viral binding and infectivity in human cells, it is necessarily difficult to connect this single molecular change to observed transmission patterns in the epidemic. This is because there is an unavoidable confounding of the genotype of the virus (82A versus 82V) and its environmental circumstances. In fact, in a remarkable coincidence, the A82V mutation appears to have occurred alongside the initial movement of virus from Guinea into Sierra Leone around April 2014 (Gire et al., 2014). This lineage arriving into Sierra Leone was remarkably successful, spreading throughout Sierra Leone into Liberia and subsequently back into Guinea (Figure 1). However, without the possibility of historical replicates, it's impossible to say whether it was the epidemiological circumstances surrounding the arrival into Sierra Leone or mutations to the virus itself that propelled this lineage to dominance.

Equally difficult to pin down is the question of why A82V may have been apparently more successful in its new human hosts. Adaptation to the human host may occur at different scales. Viruses within a host will compete with one another to spread from cell to cell most rapidly. If there's variation of the virus population within an individual host, we expect the variant that's better able to replicate to increase in frequency. This, in turn, can promote the transmission of the adaptive variant, even if selection is acting entirely at the within-host level. Conversely, selection may also act at the between-host level, wherein hosts that are infected with a more transmissible variant will more often spread virus to secondary infections, and the transmissible variant will increase in frequency in the overall epidemic. There is an important difference in the strength of natural selection between these two scenarios. In the within-host scenario, assuming an adaptive variant has a 1% replication advantage (and assuming a cellular doubling time of 8 hr), then we expect the adaptive variant to increase by 42% in frequency in the course of 15 days. In the between-host scenario, assuming an adaptive variant confers a 1% transmission advantage (and assuming a 23-day epidemic doubling rate), then we expect the adaptive variant to increase by 0.7% in

frequency in the course of 15 days. This is ignoring stochastic effects, which will have a larger impact on the between-host transmission process. Thus, a simple back-of-the-envelope calculation suggests that selection for within-host cell-to-cell spread is likely to have a strong knock-on effect on what variants end up transmitting between hosts.

This leaves us with the hypothesis that A82V may well have promoted spread of Ebola during the West African epidemic, but that A82V likely did not arise due to selection to promote spread. A similar observation exists in the influenza literature, where transient spill-over events from the animal reservoir often show a characteristic pattern in which signature adaptations occur very early on in the human-to-human transmission chain (Chen et al., 2006). However, in previous influenza pandemics, the acquisition of genetic data lagged behind epidemic spread, and so the initial key events in the transmission chain were not obvious. For the West African Ebola epidemic, case detection and sequencing was rapid enough to observe adaptation in real time. This creates the opportunity for tools to track and identify the spread of adaptive variants to improve public health response. In that respect, the two studies appearing in this issue of *Cell* set the stage and the standard to take field observations and forensically tie them to molecular adaptations via detailed laboratory reconstructions

REFERENCES

Bornholdt, Z.A., Ndungo, E., Fusco, M.L., Bale, S., Flyak, A.I., Crowe, J.E., Jr., Chandran, K., and Saphire, E.O. (2016). MBio *7*, e02154–e02115.

Carette, J.E., Raaben, M., Wong, A.C., Herbert, A.S., Obernosterer, G., Mulherkar, N., Kuehne, A.I., Kranzusch, P.J., Griffin, A.M., Ruthel, G., et al. (2011). Nature *477*, 340–343.

Chen, G.W., Chang, S.C., Mok, C.K., Lo, Y.L., Kung, Y.N., Huang, J.H., Shih, Y.H., Wang, J.Y., Chiang, C., Chen, C.J., and Shih, S.R. (2006). Emerg. Infect. Dis. *12*, 1353–1360.

Côté, M., Misasi, J., Ren, T., Bruchez, A., Lee, K., Filone, C.M., Hensley, L., Li, Q., Ory, D., Chandran, K., and Cunningham, J. (2011). Nature *477*, 344–348.

Diehl, W.E., Lin, A.E., Grubaugh, N.D., Carvalho, L.M., Kim, K., Kyawe, P.P., McCauley, S.M., Donnard, E., Kucukural, A., McDonel, P., et al. (2016). Cell *167*, this issue, 1088–1098.

Gire, S.K., Goba, A., Andersen, K.G., Sealfon, R.S., Park, D.J., Kanneh, L., Jalloh, S., Momoh, M., Fullah, M., Dudas, G., et al. (2014). Science *345*, 1369–1372.

Gong, X., Qian, H., Zhou, X., Wu, J., Wan, T., Cao, P., Huang, W., Zhao, X., Wang, X., Wang, P., et al. (2016). Cell *165*, 1467–1478.

Ng, M., Ndungo, E., Kaczmarek, M.E., Herbert, A.S., Binger, T., Kuehne, A.I., Jangra, R.K., Hawkins, J.A., Gifford, R.J., Biswas, R., et al. (2015). eLife *4*, e11785.

Urbanowicz, R.A., McClure, C.P., Sakuntabhai, A., Sall, A.A., Kobinger, G., Müller, M.A., Holmes, E.C., Rey, F.A., Simon-Loriere, E., and Ball, J.K. (2016). Cell *167*, this issue, 1079–1087.

Wang, H., Shi, Y., Song, J., Qi, J., Lu, G., Yan, J., and Gao, G.F. (2016). Cell *164*, 258–268.

The Human Functional Genomics Project: Understanding Generation of Diversity

Jenna L. Pappalardo[1] and David A. Hafler[1,*]

[1]Departments of Immunobiology and Neurology, Yale School of Medicine, New Haven, CT 06520, USA

*Correspondence: david.hafler@yale.edu

http://dx.doi.org/10.1016/j.cell.2016.10.040

Generation of biologic diversity is a cornerstone of immunity, yet the tools to investigate the causal influence of genetic and environmental factors have been greatly limited. Studies from the Human Functional Genomics Project, presented in *Cell* and other Cell Press journals, integrate environmental and genetic factors with the direction and magnitude of immune responses to decipher inflammatory disease pathogenesis.

The first step in defining the inflammatory nature of human infectious and autoimmune diseases came from breakthroughs in light microscopy, beginning over two centuries ago, through the examination of organs in diseased patients. Following waves of subsequent technological advance, many distinct clinical syndromes have since been identified, and it has become clear that both genetic inheritability and environmental factors play important roles. Yet, until recently, the tools to investigate causality by these factors were greatly limited, and why human immune responses are so variable between individuals has largely remained unknown. A critical turning point toward addressing these questions has been the sequencing of the human genome and the subsequent adaptation of technologies to facilitate whole-genome investigation of germlines, transcriptomes, and epigenomes. Genome-wide association scans (GWAS) have, for example, allowed the unbiased clustering of genetic variation defining human autoimmune diseases (Fahr et al., 2015), and we now know that