Physica A 427 (2015) 277–281

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Similarity is not enough: Tipping points of Ebola Zaire mortalities

J.C. Phillips

Department of Physics and Astronomy, Rutgers University, Piscataway, NJ, 08854, United States

HIGHLIGHTS

- Ebola virus is the most deadly virus known.
- Bioinformatic scaling explains the subtype dependence of its virulence.
- The 2014 West Africa strain is only slightly mutated, but its very different properties can be explained by critical fluctuations.
- Structural considerations identify and explain an inverse correlation between virulence and transmissibility.

ARTICLE INFO

Article history:
Received 17 November 2014
Available online 11 February 2015

Keywords:
Scaling
Bioinformatic
Criticality
Self-organized
Thermodynamic
Strand

ABSTRACT

In early 2014 an outbreak of a slightly mutated Zaire Ebola subtype appeared in West Africa which is less virulent than 1976 and 1994 strains. The numbers of cases per year appear to be ∼1000 times larger than the earlier strains, suggesting a greatly enhanced transmissibility. Although the fraction of the 2014 spike glycoprotein mutations is very small (∼3%), the mortality is significantly reduced, while the transmission appears to have increased strongly. Bioinformatic scaling had previously shown similar inversely correlated trends in virulence and transmission in N1 (H1N1) and N2 (H3N2) influenza spike glycoprotein mutations. These trends appear to be related to various external factors (migration, availability of pure water, and vaccination programs). The molecular mechanisms for Ebola's mutational response involve mainly changes in the disordered mucin-like domain (MLD) of its spike glycoprotein amino acids. The MLD has been observed to form the tip of an oligomeric amphiphilic wedge that selectively pries apart cell–cell interfaces via an oxidative mechanism.

1. Introduction

Protein amino acid sequences (aas) are rich in information, especially when combined with structural data. There are many Web-based tools for analyzing aas, but by far the most utilized is BLAST (**B**asic **L**ocal **A**lignment **S**earch **T**ool), which compares two given sequences, or searches for sequences similar to a given sequence. The original BLAST paper [1] was the most highly cited paper published in the 1990s. Ebola filoviruses exhibit ∼30% spike glycoprotein (GP) subtype aas differences, with wide variations in virulence (from nearly 90% to nearly 0% mortality) [2]. Although the Marburg species has only 30% GP similarity to Ebola, its length and viral morphology are similar, and it also has high mortality levels ∼50% [3,4]. The static GP domain structure of the most virulent Ebola subtype, Zaire or ZEBOV, whose structure bound to an antibody is shown in Fig. 1 of Ref. [5], is the basis of many studies [6–11]. The molecular basis for explaining the widely differential pathogenicity of the Ebola and Marburg filoviruses, which depends on multiple mechanisms involving many steps from molecular membrane penetration to selective disruption of cell–cell binding, remains elusive [2,7].

E-mail address: jcphillips8@comcast.net.

http://dx.doi.org/10.1016/j.physa.2015.02.056
0378-4371/

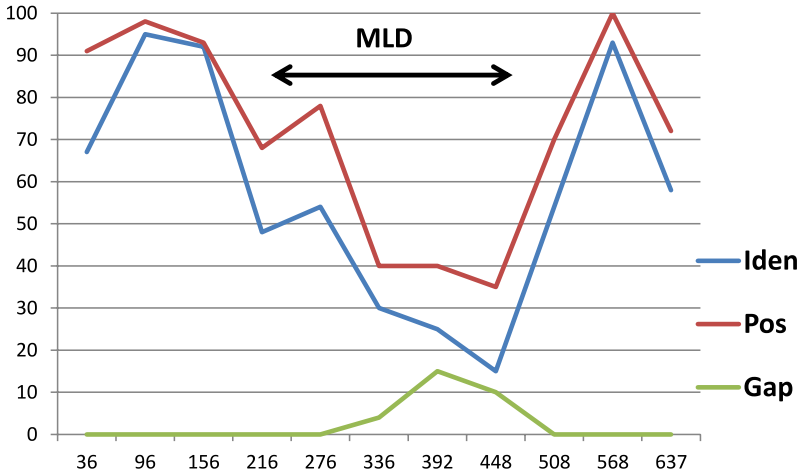

Fig. 1. Running BLAST differences (Identity, Positives and Gaps) between GP ZEBOV and REBOV (Uniprot O11457 and Q66799). Similarity differences below 40% suggest structural differences in the MLD. The values shown refer to (roughly) $W = 60$ windows centered on the indicated sites, for instance, 216 is calculated for 187–246.

In 2014 a new GP ZEBOV* strain appeared, with apparently much increased transmissibility, and a recently estimated WHO mortality reduced from 90% to ∼30%. Heroic efforts have sequenced GP ZEBOV*, which differs (BLAST) from the 1976 GP ZEBOV strain by only 3% [12]. Sequential similarity differences up to 60% normally do not alter structural homologies (folds), so conventional structure-function methods are uninformative here. Indeed, the surviving authors of Ref. [12] note that their data alone "do not address whether these (aas) differences are related to the severity of the outbreak".

Bioinformatic scaling had previously shown inverse correlated trends in virulence and transmission in N1 (H1N1) and N2 (H3N2) influenza spike glycoprotein strains [13]. These trends appeared to be related to various external factors (migration, cold and crowded conditions, and vaccination programs). As these conditions worsen, a "tipping point" (technically a thermodynamic critical point) can be reached, beyond which a pandemic occurs, its most famous example being the 1918 H1N1 world-wide influenza pandemic (3% mortality, 50 million deaths). There are collective social tipping points [14], which have become more prevalent as the Internet has facilitated "viral" information transmission.

For self-organized networks very near thermodynamic equilibrium there is a general theory of critical tipping points which appears to be relevant to the ZEBOV* outbreak [15]. Studies of toy models had shown that thermodynamic functions near critical points have self-similar (power law) behavior describable by fractals [16,17]. Proteins are compacted into globules, and a landmark bioinformatic paper identified the twenty amino-acid specific fractals MZ that describe hydropathic compaction forces [18]. These twenty fractals are associated with the differential geometry of solvent-accessible Voronoi surfaces centered on each amino acid. While the differences between the MZ hydropathic parameters discovered bioinformatically and the standard KD water–air enthalpy differences [19] are small (correlation (MZ, KD) $= 0.85$), it is just such small differences that become important near a thermodynamic critical point [20]. A striking example is β amyloid aggregation, responsible for Alzheimer's disease, where the MZ scale is nearly twice as effective as the KD scale in identifying molecular β sandwich structure, given only the β amyloid aas as input [21].

Two GP mechanisms have been suggested as possible origins of large Zebola subtype virulence variations, different folds (Ebola Zaire is the only structure known) [5], and different flexibility [7]. Ebola subtypes share 70% sequence similarity, and subtypes with more than 40% similarity generally have similar folds, so the fold explanation is unpromising [22]. A characteristic feature of filoviruses is a large, disordered, flexible and highly glycosylated mucin-like domain (MLD, 305–485), which is a natural candidate for disruption of cell–cell binding [7,11]. There are large differences between Ebola subtypes in the MLD, and strong similarities outside the MLD. This is shown in Fig. 1, which presents the results of a standard BLAST similarity analysis of the full (GP1, GP2) aas differences of ZEBOV and the least virulent subtype REBOV, using running windows. From this plot one suspects that the virulence differences between the Ebola subtypes arise from differences in their disordered MLD, but the elastic mechanism for these differences is unclear.

2. Results and discussion

Given a bioinformatic scale ψ (aa), from Fig. 1 it is clear that the subtype functional differences associated with the 305–485 MLD occur on a large scale, which we have taken to be defined by a sliding window of width $W = 115$. The resulting ψ (aa, 115) profiles of the three subtypes with variable virulences (ZEBOV, 80%–90% mortality, SEBOV, 40%–60% mortality, and REBOV, ∼0% mortality) are shown in Fig. 2. With the MZ scale, hydroneutral is near 155, so the ZEBOV profile consists of two hydroneutral peaks separated by an elastic hinge associated with the deep V-like hydrophilic minimum of the disordered MLD. The disordered structure of the MLD is highly variable (Fig. 1) and may be responsible for most of the