Deep Dive: Oct 15

Submitted by admin on Wed, 10/15/2014 - 11:39

Ebola Deep Dive Discussion: Outline, Figures, and Notes

Outline:

* Background Information and Definitions
* Impact of Ebola Mutations
* Ebola 2014 Mutation Rate: Comparison to previous Ebola outbreaks & Other Viruses
* Superinfection: Mathematical Properties / Evolutionary Dynamics
* Ebola Virulence vs Infectivity: Confounding Variables
* Recombination: Evidence for Horizontal Gene Transfer in Ebola

NOTE: This post represents notes from a 'deep dive' scientific conference call regarding the Ebola 2014 outbreak.

Definitions and Background Information:

Virulence: A virus's adverse impact on host fitness (ie. host mortality and morbidity)
Infectivity: A virus's inherent ability to spread (ie. reflected in basic reproduction number)

There have long been discussions on virulence vs infectivity. Currently, the consensus is that these two parameters are absolutely related within a virus to some sort of fitness optimum, but the confounding variables present make straightforward analysis impossible. The important point is to understand that Virulence and Infectivity are linked, and form a sort of 'optimum' for a virus under a given set of conditions.

The Ebola Polymerase is an RNA-Dependent RNA Polymerase, which means that it is far more error prone than DNA viruses. This reduces the reproductive fidelity per cycle and introduces genetic instability in the form in single-nucleotide polymorphisms (SNPs), insertions, and deletions. Additionally, recent evidence has indicated the Ebola RdRp is capable of engaging in genetic recombination. This gives Ebola a rich source of genetic change from which to explore it's optimum fitness landscape.

Deep Dive Discussion 1:

· The Ebola 2014 viral outbreak is of a has a common genetic, lineage which goes back to the discovery of the species of the Zaire Ebolavirus in 1976.

· The Ebola 2014 outbreak is considered a 'strain' or 'sub-clade' of Ebola Zaire.

· Ebola Zaire viral species are defined by in reference to a 'consensus strain' -- the Mayinga-76 Ebola Virus (named for Ebola victim Nurse Mayinga N'Seka in 1976).

· The Ebola 2014 outbreak's RNA genome **is only 97% similar** to the 1976 Ebola Zaire consensus strain.

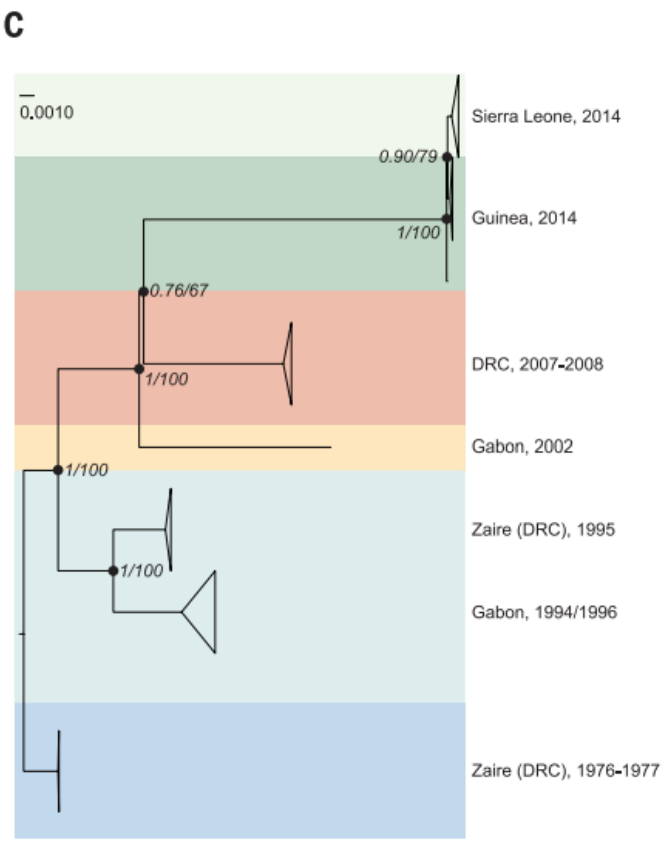
· This means out of the 20kb Ebola 2014 genome, **approx. 600 nucleotides are different** from the 1976 consensus strain.

· The Ebola 2014 outbreak's RNA Genome can be considered to be a separate 'clade' within the species of Ebola Zaire due to these genetic differences.

· The Ebola 2014 outbreak's RNA Genome is most closely related to Ebola strains from 2007-2008 isolated and sequenced in the DRC.

· Computer analysis indicates the Ebola 2014 outbreak and the Ebola DRC 2007-2008 outbreak had a common viral ancestor, perhaps around 2004.

Deep Dive Figure 1:



Ebola 2014 Genetic Lineages

· Notice in the above diagram that Guinea and Sierra Leone both have distinct Ebola 2014 'sub-clades'

· Also notice in the above diagram that the current Ebola 2014 Guinea and Sierra Leone strains are most closely related to the DRC outbreak in 2007-2008.

· The current 2014 Ebola outbreak and the 2007-2008 DRC outbreak have an **unidentified parent lineage**, which ultimately goes back to 1976.

Deep Dive Discussion 2:

· Notice the Ebola 2014 outbreak and the DRC 2007-08 outbreak diverged from common ancestor strain in 2004 (Deep Dive Figure 2A).

· With only 97% sequence homology to the Mayinga-76 strain, the current Ebola outbreak could be substantially changed it's reproductive fitness -- but this is unknown.

· Notice on (Figure 2B) we have what looks like **3 or 4 sub-clades ('strains') present** in the 2014 Ebola outbreak.

· We can see an Ebola strain in Guinea ("GN") appeared earliest (Feb - March), but then died out by May 2014. (Fig 2B)

· After the Ebola 2014 (GN) strain disappeared, new Ebola (SLx) strains took it's place. (Fig 2B)

· The Ebola strains from Sierra Leone ("SL1, SL2...") appeared after the GN strain, and these continued to spread in May and June 2014. (Fig 2B)

· Within the Ebola 2014 outbreak, **we are dealing with multiple genetic sub-clades of Ebola** ('sub-strains') **which circulate and compete**. (Fig 2B)

· The resurgence of the Ebola in May 2014 coincided with the appearance of genetically distinct Ebola viral sub-clades SL2 and SL3. (Fig 2B)

· Deep Dive Figure 2B does not tell us about reproductive fitness, but this is a mystery that must be resolved (do these Ebola genetic changes play any role?).

Deep Dive Figure 2:

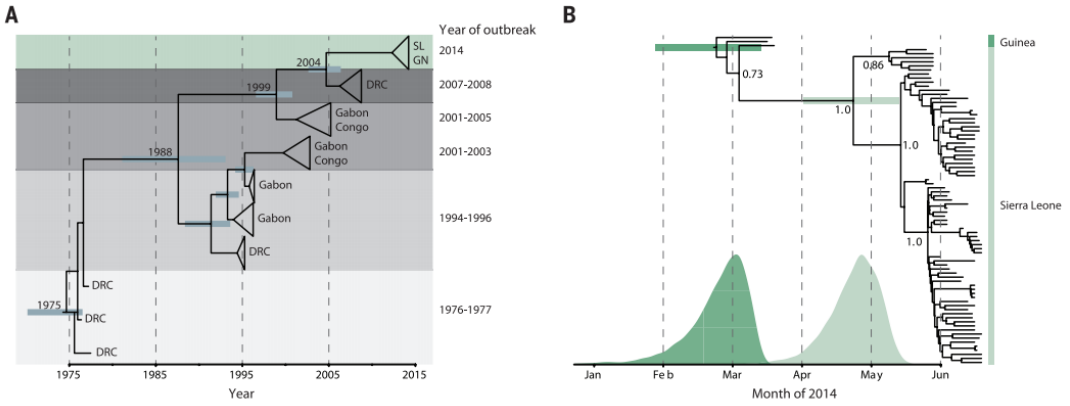


Fig. 3. Molecular dating of the 2014 outbreak. (**A**) BEAST dating of the separation of the 2014 lineage from central African lineages [SL, Sierra Leone; GN, Guinea; DRC, Democratic Republic of Congo; time of most recent common ancestor (tMRCA), September 2004; 95% highest posterior density (HPD), October 2002 to May 2006]. (**B**) BEAST dating of the tMRCA of the 2014 West African outbreak (23 February; 95% HPD, 27 January to 14 March) and the tMRCA of the Sierra Leone lineages (23 April; 95% HPD, 2 April to 13 May). Probability distributions for both 2014 divergence events are overlaid below. Posterior support for major nodes is shown.

Deep Dive Discussion 3:

· Above Deep Dive Figure 3A, we can view the Ebola 2014 Virus Genome, and it's accumulated mutations as of August 28th 2014.

· We can see that circulating Ebola viruses have **substantial genetic changes**, including non-synonymous mutations (protein changes) in:

NP gene (nucleoprotein)
VP35 gene (L cofactor/immune suppression)
VP40 gene (Ebola Matrix Protein)
GP gene (Ebola Spike Glycoprotein)
VP24 gene (Minor Matrix Protein)
L gene (Ebola RdRp)

· This means that the Ebola 2014 Virus has protein-changes (red color) to EVERY gene except highly-conserved VP30 -- probably since VP30 is required for transcription activation.

· We can clearly see the Ebola 2014 Virus Genome has accumulated a substantial number of changes, including non-synonymous mutations.

· What is especially curious (h/t to IBM) is the amount of mutations that accumulated in the **intragenic region** -- the grey lines between VP30 and VP24.

· The implication of these non-synonymous mutations in the 2014 Ebola Virus Genome is **unknown at this time**. They could substantially impact viral replication, tissue tropisms, virulence, etc.... Or these mutations could have absolutely no effect.

Deep Dive Figure 3:

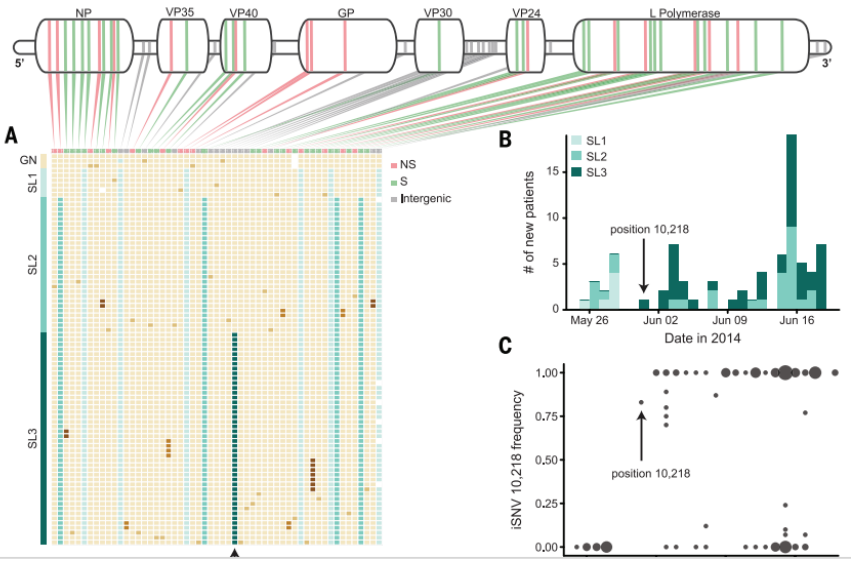


Fig. 4. Viral dynamics during the 2014 outbreak. (**A**) Mutations, one patient sample per row; beige blocks indicate identity with the Kissidougou Guinean sequence (GenBank accession KJ660346). The top row shows the type of mutation (green, synonymous; pink, nonsynonymous; gray, intergenic), with genomic locations indicated above. Cluster assignments are shown at the left. (**B**) Number of EVD-confirmed patients per day, colored by cluster. Arrow indicates the first appearance of the derived allele at position 10,218, distinguishing clusters 2 and 3. (**C**) Intrahost frequency of SNP 10,218 in all 78 patients (absent in 28 patients, polymorphic in 12, fixed in 38). (**D** and **E**) Twelve patients carrying iSNV 10,218 cluster geographically and

· Notice in Deep Dive Figure 3A we see the Ebola virus genome organized from 5' to 3' end divided into a grid of boxes. The rows are grouped by 2014 viral sub-clade (or 'sub-strain'). . . "GN", "SL1", "SL2", and "SL3". The columns represent genetic changes across the Ebola -ssRNA genome. The Ebola 2014 "SL3" strain can be distinguished by a unique SNP at position 10,218 in the genome.

· Notice in Deep Dive Figure 3B we see that over time, the Ebola 2014 SL1 strain became less and less dominant in the population, and burned itself out by June 2014, meaning that by June 2014, both the Ebola SL1 and GN strains were not actively circulating in humans in West Africa.

· During June 2014, the Ebola SL2 and SL3 strains began to become dominant and co-circulate. Eventually, both became widespread.

· Deep Dive Figure 3C shows that the iSNP at position 10,218 (associated with SL3) became increasingly frequent over the month of June 2014, indicating reproductive success (for whatever reason) of the Ebola 2014 SL3 strain.

· Current 'deep sequencing' data from the ongoing outbreak as of Oct 16 2014 is not available. The diagram referenced here stops analysis at August 2014.

Deep Dive Discussion 4:

· The phylogenetic tree below contains isolates from patients infected with the Ebola Virus in 2014.

· Their isolates were 'deep-sequenced' and the Ebola RNA sequences were deposited in Genbank.

· Phylogenetic analysis of these Ebola 2014 RNA sequences show (as of August 2014) that there are Four distinct Ebola sub-clades ('strains')

· The earliest strains, named GN and SL1, correlate with Guinea and Cluster1 respectively.

· The circulating strains, named SL2 and SL3, correlate with Cluster2 and Cluster3 respectively.

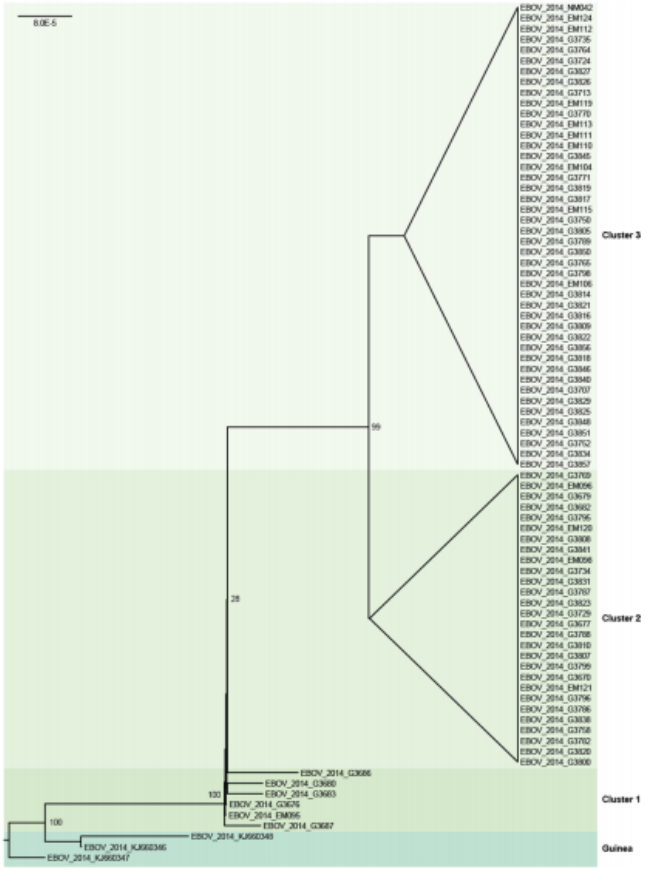
· It is undetermined if there is any molecular biological or sociological factor which would favor SL2 and SL3 over GN and SL1.

· Observed data regarding changes in Ebola sub-clades 'strains' may simply represent sampling bias of a small number of isolates.

Deep Dive Figure 4:

Fig. S8.

Phylogenetic tree showing the individual lineages in the 2014 EVD outbreak. A maximum likelihood tree was created using RAxML and the four main clusters (Guinea, as well as three Sierra Leone clusters) are displayed. Bootstrap values (500 pseudoreplicates) are shown for each node. Scale bar = nucleotide substitutions/site.



Deep Dive Discussion 5:

· This diagram shows the Ebola 2014 mutation rate compared to various parameters

· Mutation Rate and Substitution Rate are not technically the same measurement. For simplicity, by mutation rate we mean substitution rate.

· Notice in Deep Dive Figure 5F we have a brown and a blue probability distribution.

· The brown distribution (Figure 5F) shows Ebola mutation rates using 'all prior known human Ebola outbreaks'. This results in a previous 'all-outbreak' Ebola substitution (mutation) rate average of about **0.9 x 10^{-3}**substitutions / base pair / year.

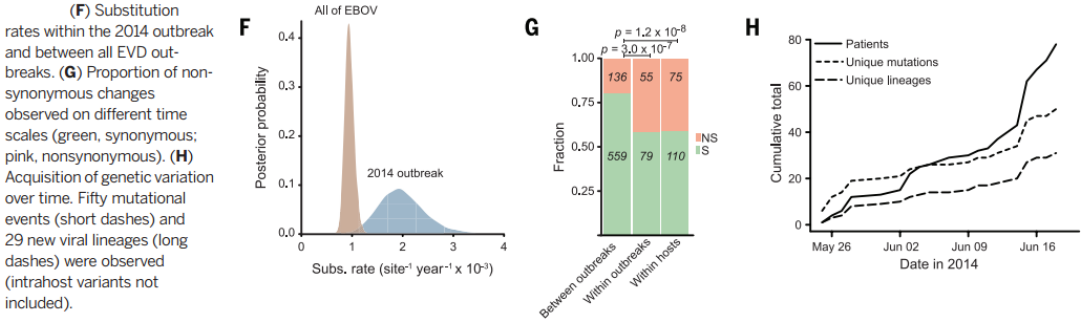
· The blue distribution shows (Figure 5F) Ebola mutation rates using sequences 'only from the 2014 Ebola outbreak'. This results in a previous 'all-outbreak' Ebola substitution (mutation) rate average of about **2.0 x 10^{-3}** substitutions / base pair / year.

· The wide probability distribution of Ebola 2014 mutation rate ranges from as low as **1.0 x 10^{-3}** subs/bp/year to as high as **3.1 x 10^{-3}** subs/bp/year.

· This data indicates that the Ebola 2014 outbreak is **undergoing genetic mutation at a rate 220% to 330% faster** than previous Ebola outbreaks.

· Part of this may reflect an acceleration of genetic change in order for the virus to be adaptable to human hosts, as it explores the fitness landscape.

Deep Dive Figure 5:



· Notice in Deep Dive Figure 5G that a significant fraction of Ebola 2014 viral mutations were detected and sequenced WITHIN patients (Fig 5G, Within Hosts).

· What this means if someone is infected with a single copy of Ebola, by the time they are sick, they may actually posses MULTIPLE Ebola viral-substrains within themselves. (Fig 5G)

· In other words, an original Ebola genome might be in a host liver cell, but a mutated Ebola genetic copy might be in the same host's spleen cell. This is how fast the virus is changing. (Fig 5G)

· Another important point from Figure 5G is that a substantial fraction of intra-host Ebola genetic mutation involves non-synonymous mutations, which can result in changes to amino acid residues which comprise the Ebola proteins. Amino acid substitutions can have no effect, they can be beneficial for the virus, or they can be detrimental to the virus.

· This is how the virus explores the 'fitness landscape'.

· Lastly, in Figure 5H, notice that from May to June 2014, the virus acquired significant genetic variation which seemed to correlate with the number of hosts it infected. As of June 16th, the Ebola virus had acquired 29 new viral lineages, which seemed to scale very closely with the number of new Ebola patients. (Fig 5H)

· Bottom Line: The larger the pool of individuals sick with Ebola 2014, the more opportunities the virus will have to adapt for better genetic fitness (better transmissibility, etc)

Deep Dive Figure 6:

SUMMARY

- The basic reproductive ratio of an infectious agent (parasite) is the number of secondary infections caused by one infected individual that has been introduced into a population of uninfected individuals.
- Parasite evolution tends to maximize the basic reproductive ratio.
- If there is a functional relationship between infectivity and virulence, then well-adapted parasites need not be harmless. Parasite evolution can lead to intermediate levels of virulence.
- Superinfection means that an already infected host can be infected by another parasite strain.

EVOLUTIONARY DYNAMICS

- Superinfection triggers intrahost competition for increased levels of virulence and reduced transmission rates.
- Superinfection increases the average level of virulence above what would be optimum for the parasite population.
- Superinfection does not maximize the basic reproductive ratio. Even the strain with the highest R_0 can become extinct.
- Superinfection leads to a coexistence of parasite strains with many different levels of virulence within a well-defined range.
- Superinfection can maintain strains with very high levels of virulence, including strains that are so virulent that they themselves could not persist alone in an otherwise uninfected host population.
- Superinfection can lead to very complicated dynamics, such as heteroclinic cycles, with sudden and dramatic changes in the average level of virulence.
- The higher the rate of superinfection, the smaller the number of infected hosts. Hence superinfection is not advantageous for the parasite population as a whole.

Deep Dive Discussion 7:

· This diagram shows how mutation rate and substitution rates relate to Viral classifications.

· Notice that the viruses with the highest substitution rates include are ssRNA viruses (which includes Ebola).

Deep Dive Figure 7:

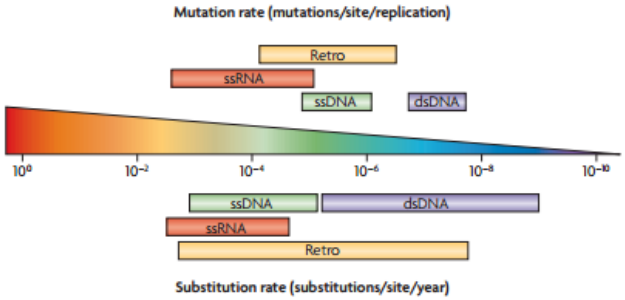


Figure 2 | **Comparison between viral mutation and substitution rates.** The ranges of mutation rates, given as mutations per site per round of replication, for viruses with different genomic architectures are summarized in the upper part of the figure (values from FIG. 1). The ranges of average substitution rates, given as substitutions per site per year, are shown in the lower part of the figure. The ranges are bound by the extremes of average substitution rates: for the retrotranscribing viruses (Retro), human immunodeficiency virus type-1 (REF. 104) and simian foamy virus[41]; for single-stranded (ss) RNA viruses, swine vesicular stomatitis virus[32] and European bat lyssaviruses[105]; for ssDNA viruses, the non-coding region of the Tomato yellow leaf curl virus[70] and canine parvovirus[68]; for double-stranded (ds) DNA viruses, the BK polyomavirus[95] and herpesvirus[37].

Deep Dive Discussion 8:

· Deep Dive Figure 8 shows the average mutation rate of the Influenza A virus NS genes.

· The average mutation rate for Influenza A is 2.6 x 10^{-3} subs / bp / year. (Influenza A / Seasonal Flu)

· The average mutation rate for Influenza B is 0.5 x 10^{-3} subs / bp / year.

· The average mutation rate for Ebola 2014 is 2.0 x 10^{-3} subs / bp / year. (Ebola 2014 Outbreak)

· Thus, the mutation rate for the Ebola 2014 outbreak is comparable to that of seasonal flu (Influenza A is one of the fastest changing viruses known).

· The future mutations in Ebola will be impossible to predict. This why it is critical to get the outbreak contained.

Deep Dive Figure 8:

TABLE 1. Analysis of mutation rate for NS genes of human influenza viruses

Virus type	Expt no.	Virus[a]	PFU of parent plaque	No. of descendant plaques analyzed	Total no. of nucleotides analyzed[b]	No. of nucleotide changes[c]	Mutation frequency/site[d]	Mutation rate	
								Per site per cycle[e]	Per site per year[f]
B	1	B/Aichi/29/99	6.4×10^6	143	147,576[g]	1	0.7×10^{-5}	1.0×10^{-6}	0.8×10^{-3}
	2	B/Aichi/44/01	2.0×10^6	146	150,672[g]	0	0	0	0
	3	B/Aichi/44/01	4.8×10^6	158	163,056[g]	1	0.6×10^{-5}	0.9×10^{-6}	0.8×10^{-3}
	Avg						$0.4 \pm 0.3 \times 10^{-5}$	$0.6 \pm 0.4 \times 10^{-6}$	$0.5 \pm 0.4 \times 10^{-3}$
A	4	A/Aichi/12/92	3.0×10^6	149	124,117[h]	1	0.8×10^{-5}	1.2×10^{-6}	1.5×10^{-3}
	5	A/Aichi/1/87	2.4×10^6	136	113,288[h]	2	1.8×10^{-5}	2.6×10^{-6}	3.2×10^{-3}
	6	A/Aichi/1/87[i]	1.2×10^6	138	114,954[h]	2	1.7×10^{-5}	2.5×10^{-6}	3.2×10^{-3}
	Avg						$1.4 \pm 0.6 \times 10^{-5}$	$2.0 \pm 0.9 \times 10^{-6}$	$2.6 \pm 1.2 \times 10^{-3}$

Deep Dive Discussion 9:

· The average mutation rate for Ebola 2014 is 2.0 x 10-3 subs / bp / year. (Ebola 2014 Outbreak)

· The average mutation rate for all previous Ebola Outbreaks is under 0.9 x 10-3 subs / bp / year. (Previous Ebola outbreaks)

· The average mutation rate for Influenza A NS genes is 2.6 x 10-3 subs / bp / year. (Seasonal Influenza)

· In Deep Dive Figure 9, we can compare the Ebola 2014 viral mutation rate to 50 other common RNA Viruses.

· In Deep Dive Figure 9, we see that Ebola 2014's viral mutation rate is among the highest mutation rates in the literature for RNA Viruses.

· In Deep Dive Figure 9, any virus with a mutation rate comparable to Ebola 2014 Virus (2.0 x 10-3 subs / bp / year) is marked in red.

· The Ebola 2014 Virus is currently mutating at a very fast pace by any metric, so much so that it is changing as fast as Seasonal Flu.

Deep Dive Figure 9:

Table 1. Overall and synonymous substitution rates for a range of 50 RNA viruses

Virus	Abbreviation	Gene[a]	Substitutions/site/year[b] $\times 10^{-3}$ All sites	Synonymous sites
Picornaviridae				
Foot and mouth disease A	FMDV-A	VP1	1.4 (0.88, 1.9)	2.5 (0.36, 4.6)
Foot and mouth disease C	FMDV-C	P1	0.73 (0.40, 1.2)	3.2 (1.8, 4.7)
Foot and mouth disease O	FMDV-O	VP1	1.3 (0.66, 1.9)	0 (0, 2.4)
Human enterovirus 71	EV71	VP1	3.4 (2.7, 4.0)	12 (7.9, 18)
Swine vesicular disease	SVDV	3BC	3.4 (2.8, 4.0)	7.1 (3.6, 15)
Calciviridae				
E. brown hare syndrome	EBHSV	VP60 (p)	0 {0, 0.73)[c]	0 (0, 3.3)[c]
Rabbit hemorrhagic disease	RHDV	VP60 (p)	1.3 (0.59, 2.1)	0.19 (0, 3.6)[c]
Flaviviridae				
Dengue 1	DENV-1	E	0.49 (0.19, 0.79)	1.7 (0.16, 3.0)
Dengue 2	DENV-2	E	0.62 (0.49, 0.74)	1.5 (0.94, 2.1)
Dengue 3	DENV-3	E	0.72 (0.53, 0.91)	2.1 (1.3, 3.0)
Dengue 4	DENV-4	E	0.77 (0.59, 1.0)	2.2 (1.5, 3.3)
Japanese encephalitis	JPEV	E	0.35 (0.089, 0.53)	0.80 (0.24, 1.3)
Louping ill	LIV	E	0.098 (0, 0.26)[c]	0 (0, 0.50)[c]
St. Louis encephalitis	SLEV	E (p)	0.16 (0, 0.38)[c]	0.37 (0, 1.5)[c]
Tick-borne encephalitis	TBEV	E	0 (0, 0.12)[c]	0.29 (0, 0.41)[c]
Yellow fever	YFV	E	0.28 (0.10, 0.44)	1.26 (0, 1.31)[c]
Classical swine fever	CSFV	NS5B (p)	2.0 (1.5, 2.7)	4.9 (2.5, 7.7)
Hepatitis C	HCV	E1 (p)	0.79 (0.61, 1.0)	0.49 (0.082, 1.5)
Togaviridae				
Barmah Forest	BFV	E2 (p)	0.010 (0, 0.16)[c]	0.096 (0, 1.0)[c]
Eastern equine encephalitis	EEEV	26S	0.20 (0.16, 0.26)	0.27 (0.21, 0.68)
Highlands J	HJV	E1	0.14 (0.086, 0.26)	0.36 (0.091, 0.90)
Ross River	RRV	E2	0.24 (0, 0.51)[c]	0.78 (0, 1.6)[c]
Sinbis	SINV	E2 (p)	0.19 (0, 0.50)[c]	0 (0, 0.83)[c]
Venezuelan equine encephalitis	VEEV	E3-E2 (p)	0.13 (0, 0.32)[c]	0 (0, 1.0)[c]
Western equine encephalitis	WEEV	E1 (p)	0.055 (0.015, 0.16)	0.010 (0, 0.53)[c]
Rubella	RUBV	E1	0.61 (0.45, 0.76)	0.51 (0, 1.2)[c]
Coronaviridae				
Equine arteritis	EAV	GL (p)	0.67 (0.39, 0.96)	0 (0, 1.1)[c]
Rhabdoviridae				
Rabies	RABV	G (p)	0 (0, 1.8)[c]	0 (0, 0.76)[c]
Vesicular stomatitis	VSV	P (p)	0.015 (0, 0.29)[c]	0 (0, 0.75)[c]
Paramyxoviridae				
Bovine respiratory syncytial	BRSV	G	0.79 (0.43, 1.6)	2.8 (0.56, 5.4)
Human respiratory syncytial A	HRSV-A	G	1.6 (1.3, 2.2)	3.3 (1.8, 5.2)
Human respiratory syncytial B	HRSV-B	G	0.27 (0, 0.88)[c]	4.8 (3.0, 7.2)
Measles	MV	HE	0.40 (0.31, 0.49)	1.0 (0.71, 1.4)
Mumps	MUV	F-SH-HN	0.25 (0, 0.43)[c]	0.65 (0, 1.4)[c]
Newcastle disease	NDV	M-F (p)	0 (0, 0.075)[c]	0 (0, 0.32)[c]
Human parainfluenza 1	HPIV-1	HN	0.22 (0.037, 0.41)	0.70 (0.35, 1.2)
Filoviridae				
Ebola	EBOV	G (p)	0 (0, 0.20)[c]	0.36 (0, 1.6)
Orthomyxoviridae				
Avian influenza A	AFLUV-A	NP	1.1 (0.94, 1.3)	3.3 (2.4, 4.2)
Classical swine influenza A	CSFLUV-A	NP	1.5 (1.3, 1.7)	3.6 (3.0, 5.5)
Equine influenza A	EQFLUV-A	NP	1.1 (0.46, 1.8)	4.2 (1.3, 7.7)
European swine influenza A	ESFLUV-A	NP	0 (0, 0.72)[c]	0.38 (0, 6.3)[c]
Human influenza A	HFLUV-A	NP	1.8 (1.6, 2.1)	3.4 (2.3, 4.5)
Influenza B	FLUV-B	HA (p)	1.6 (1.3, 1.9)	2.3 (1.7, 3.7)
Human influenza C	HFLUV-C	HE	0.24 (0.13, 0.34)	0.80 (0.42, 1.3)
Bunyaviridae				
Rift valley fever	RVFV	NSs (p)	0 (0, 0.12)[c]	0.79 (0.010, 1.7)
Reoviridae				
Bluetongue	BTV	S3	0.36 (0.28, 0.48)	0.60 (0.36, 1.1)
Human rotavirus	HROTAV	VP4	0.58 (0.23, 0.95)	1.4 (0, 3.4)[c]
Reovirus	REOV	S3	0.30 (0, 0.64)[c]	1.0 (0, 2.4)[c]
Birnaviridae				
Infectious bursal disease	IBDV	VP2 (p)	0.44 (0, 0.90)[c]	0.71 (0, 2.6)[c]
Retroviridae				
Human immunodeficiency 1	HIV-1	Gag-Env (p)	2.5 (1.1, 4.0)	0 (0, 0.48)[c]

[a] (p) indicates partial gene sequences.
[b] Numbers in parentheses are 0.95 confidence limits.
[c] In this case the SRDT model was not significantly better than the SR model.

Deep Dive Discussion 10:

· The pages below are extracts from the book Evolutionary Dynamics by Martin Nowak.

· These diagrams show how it is mathematically possible for a pathogen NOT maximize it's Ro under conditions of Superinfection.

· These diagrams also illustrate how it is possible for Nature to permit the evolution of Virulence in pathogens, when it seems such virulence is not in the pathogen's self-interest.

Deep Dive Figure 10:

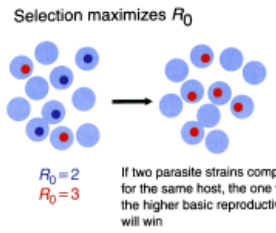


Figure 11.3 In simple models of infection dynamics, selection acts to maximize the parasite's basic reproductive ratio. If two parasites compete for the same host, then the parasite with higher R_0 will outcompete the other parasite. Therefore well-adapted parasites have a high R_0, but not necessarily low virulence.

evolutionary dynamics will increase β and reduce v. This represents the conventional wisdom that infectious diseases will evolve to become less virulent.

In general, however, we expect an association between virulence v and infectivity β; usually the harm done to hosts (v) is associated with the production of transmission stages (β). For certain functional relations between v and β there is an evolutionarily stable degree of virulence, corresponding to the maximum value of R_0. Other situations allow evolution toward the extreme values of very high or low virulences. The detailed dynamics depend on the shape of β as a function of v. It is interesting to note that along some trajectories where virulence increases, parasite evolution can lead to lower and lower parasite population sizes (in terms of total number of infected hosts).

If the infectivity is proportional to virulence, $\beta = av$, where a is some constant, then the basic reproductive ratio, R_0, is an increasing function of virulence, v. In this case selection will always favor more virulent (and therefore more infectious) strains.

If the infectivity is a saturating function of virulence, $\beta = av/(c + v)$, then the basic reproductive ratio, R_0, is a one-humped function of virulence. The maximum R_0 is achieved at an intermediate optimum level of virulence given by $v_{opt} = \sqrt{cu}$. If the virulence of a parasite population is greater than v_{opt}, then selection will reduce virulence. If it is less than v_{opt}, then selection will increase virulence.

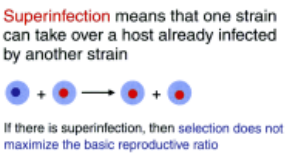


Figure 11.4 Superinfection means that an already-infected host can be infected by another parasite strain. There is competition between the two parasite strains in the superinfected individual; one parasite strain may win this competition and outcompete the other. A consequence of superinfection is that natural selection no longer maximizes the basic reproductive ratio. Instead there can be coexistence of different parasite strains with different levels of virulence. In general, superinfection leads to increased virulence beyond what would be optimum for the parasite. Superinfection introduces competition on two levels: within an infected host and in the population of hosts.

11.3 SUPERINFECTION

The analysis of the previous section did not include the possibility of superinfection. An infected host is not susceptible to another infection. We will now remove this limitation and allow for an infected host to be superinfected by another parasite strain (Figure 11.4).

We will consider a heterogeneous parasite population with a range of different virulences, and assume that more virulent strains outcompete less virulent strains within an infected individual. Thus increased virulence provides a competitive advantage over other parasites in the same host.

For simplicity, we assume that the infection of a single host is always dominated by one parasite strain. Therefore superinfection means that a more virulent strain takes over a host infected by a less virulent strain. This can be described by the following system of ordinary differential equations:

$$\begin{aligned} \dot{x} &= k - ux - x\sum_{i=1}^{n}\beta_i y_i \\ \dot{y}_i &= y_i(\beta_i x - u - v_i + s\beta_i\sum_{j=1}^{i-1} y_j - s\sum_{j=i+1}^{n}\beta_j y_j) \qquad i = 1,\ldots,n \end{aligned} \tag{11.10}$$

Here v_i denotes the virulence of strain i. We order the strains such that $v_1 < v_2 < \ldots < v_n$. A more virulent strain can superinfect a host already infected with a less virulent strain. The parameter s describes the rate at which superinfection occurs relative to infection of uninfected hosts. If either the host or the parasite has evolved mechanisms to make superinfection more difficult, then s is smaller than one. If already-infected hosts are more susceptible to acquiring a second infection, then s is greater than one, which means superinfection occurs at increased rates.

For the numerical simulations shown in Figure 11.5, we assume a functional relation between virulence and infectivity given by

$$\beta_i = \frac{av_i}{c + v_i}. \tag{11.11}$$

For low virulence, infectivity increases linearly with virulence. For high virulence, there is a saturation of infectivity at a maximum level. The basic reproductive ratio is given by

$$R_{0,i} = \frac{akv_i}{u(c + v_i)(u + v_i)}. \tag{11.12}$$

The optimal virulence, which maximizes R_0, is given by

$$v_{opt} = \sqrt{cu}. \tag{11.13}$$

Figure 11.5 shows the equilibrium population structure of the parasite for various values of s between 0 and 2. We have assumed $k = 1$, $u = 1$, and $\beta_i =$

Figure 11.5 The equilibrium distribution of parasite strains with different levels of virulence. The simulation is performed according to equation (11.10) with $k = 1$, $u = 1$, $n = 50$, $\beta_i = 8v_i/(1 + v_i)$, and $s = 0$, 0.2, 1, 2 as indicated. The individual v_i are randomly distributed between 0 and 5. In the absence of superinfection, $s = 0$, the strain with the maximum basic reproductive rate, R_0, is selected. With superinfection, $s > 0$, we find the coexistence of many different strains with different virulences, v_i, within a range v_{min} and v_{max}, but the strain with the largest R_0 is not selected. Superinfection does not optimize parasite reproduction. For increasing s, the values of v_{min} and v_{max} increase, as well. The x-axis denotes virulence, the y-axes indicate equilibrium frequencies (always scaled to the same largest value).

$8v_i/(1+v_i)$. We simulated $n = 100$ strains of parasites with virulences randomly distributed between 0 and 5. For this choice of parameters, the strain with a virulence closest to 1 has the largest R_0. Indeed we find that this strain is selected in the absence of superinfection, $s = 0$. If superinfection is possible ($s > 0$), then there is selection of an ensemble of strains with a range of virulences between two boundaries, v_{min} and v_{max}, with $v_{min} > v_{opt}$. Thus superinfection has two important effects: (i) it shifts parasite virulence to higher levels, beyond the level that would maximize the parasite's reproductive rate; and (ii) it leads to a coexistence between a number of different parasite strains with a range of virulences. There are amusing ups and downs in the equilibrium densities of strains. A strain has a high equilibrium frequency if the strain with a slightly larger virulence has low frequency, and vice versa. Only a subset of strains survive at equilibrium. What determines this complicated and unexpected equilibrium structure?

11.4 AN ANALYTICAL MODEL OF SUPERINFECTION

Let us now derive an analytical understanding of the complexities introduced by superinfection. Instead of using a constant immigration rate k for uninfected hosts, we choose a variable immigration rate that balances exactly the death of uninfected and infected hosts. This can be done by setting

$$k = ux + uy + \sum v_i y_i \tag{11.14}$$

in equation (11.10). The total number of infected hosts is given by $y = \sum_{i=1}^{n} y_i$. The sum $x + y$ remains constant and without loss of generality we choose $x + y = 1$. We obtain the following system of n equations

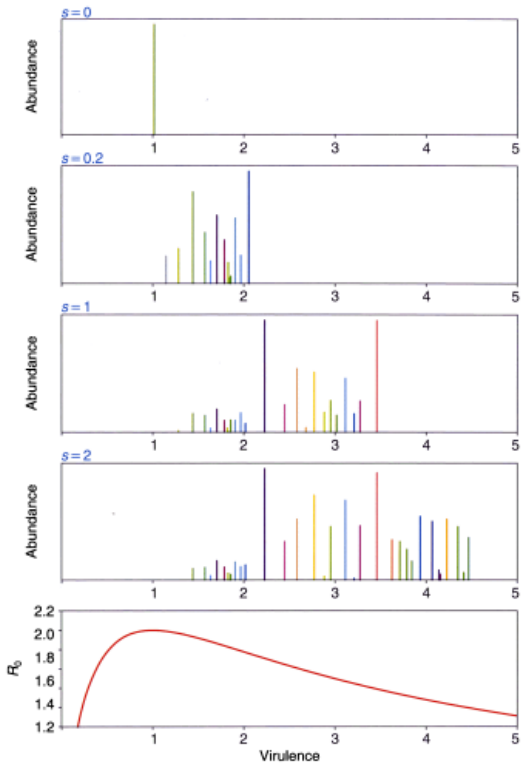
$$\dot{y}_i = y_i \left[\beta_i (1-y) - u - v_i + s \left(\beta_i \sum_{j=1}^{i-1} y_j - \sum_{j=i+1}^{n} \beta_j y_j \right) \right] \tag{11.15}$$

$$i = 1, \ldots, n$$

Note that y remains in the closed interval $[0, 1]$.

System (11.15) is a Lotka-Volterra equation. It can be written in the form

Deep Dive Figure 11:



Deep Dive Discussion 12:

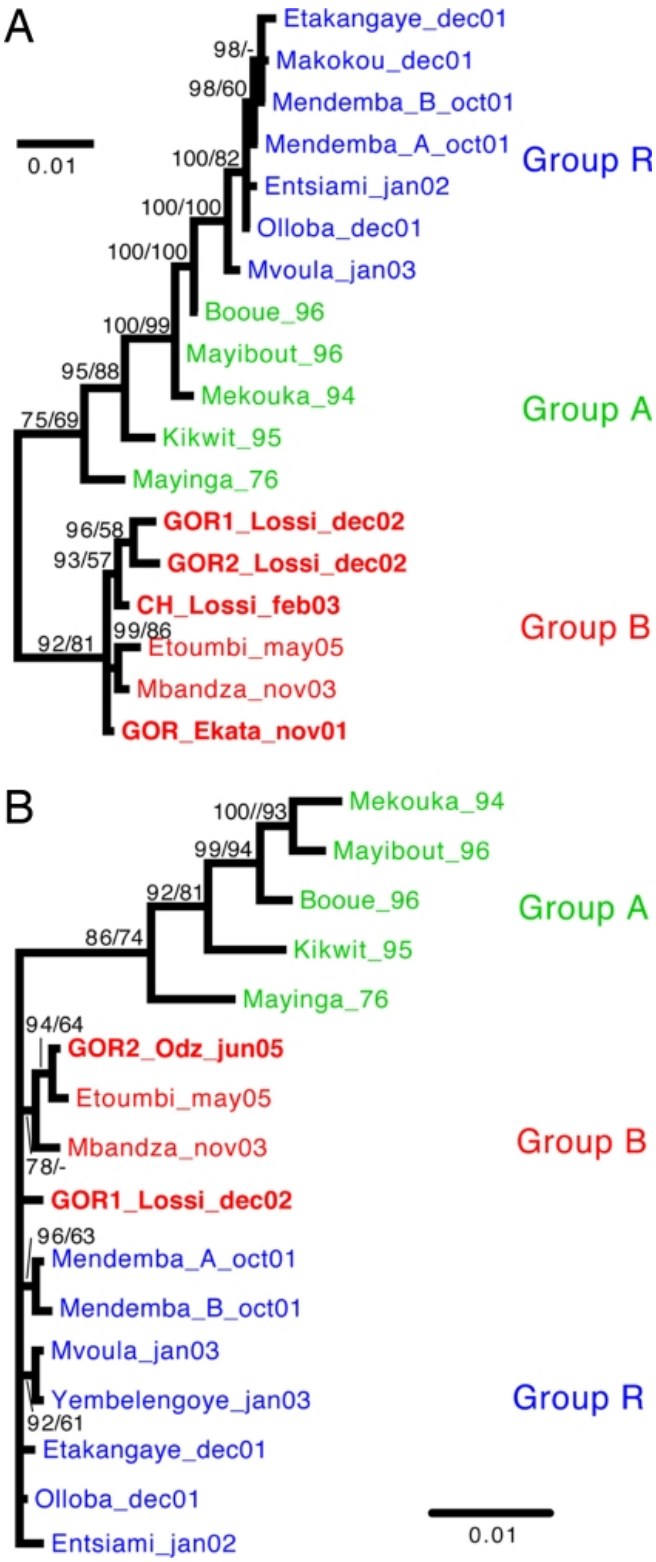
· Contrary to previous thinking, negative-sense ssRNA viruses can undergo genetic recombination.

· Recombination is where the viral RdRp essentially 'swaps' reading from one RNA strand to another, creating a chimeric viral RNA.

· Evidence has emerged to show Ebola Virus has undergone recombination events, event within the last 15 to 20 years.

· This evidence should seriously question our perspective and approach to how viruses evolve, especially if they can 'swap' genetic material so easily in a non-segmented genome.

Deep Dive Figure 12:



source: Isolates of Zaire ebolavirus from wild apes reveal genetic lineage and recombinants

---

Deep Dive Bibliography:

[1] Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. (Gire et al, 2014).

[2] Comparison of the Mutation Rates of Human Influenza A and B Viruses. (Nobusawa et al, 2006).

[3] Rates of Molecular Evolution in RNA Viruses: A Quantitative Phylogenetic Analysis. (Jenkins et al, 2002).

[4] Rates of evolutionary change in viruses: patterns and determinants. (Duffy, 2008).

[5] Viral Mutation Rates. (Sanjuan et al, 2010).

[6] Evolutionary Dynamics: Exploring the Equations of Life. (book), (Nowak, 2006).

[7] Isolates of Zaire ebolavirus from wild apes reveal genetic lineage and recombinants. (Wittman et al, 2007).