Viral Evolution

It Is All About Mutations

Adi Stern[1], Raul Andino[2]

[1]*Department of Molecular Microbiology and Biotechnology, Tel-Aviv University, Tel-Aviv, Israel;* [2]*Department of Microbiology and Immunology, University of California, San Francisco, CA, USA*

Chapter Outline

1. INTRODUCTION

The extremely high mutation rates of viruses are not matched by any other organism in the kingdom of life. The high mutation rates of viruses, coupled with short generation times and large population sizes, allow viruses to rapidly evolve and adapt to the host environment. This has important implications for the pathogenesis of viral infections. In the course of the chapter, we will address a number of questions about virus evolution: (1) How are mutation rates defined and what are substitution rates? (2) Why are mutation rates so high? How do they differ for RNA and DNA viruses? (3) How are these rates measured and what are the shortcomings of standard measures? (4) Why is multiplicity of infection so relevant to the accumulation of mutations? (5) What are phylodynamics? What is a molecular clock and how is it estimated? (6) What drives virus evolution? What role does the host response play in virus evolution? How does this impact pathogenesis? Do viruses evolve to a benign relationship with their hosts? This chapter sets forth the basic tenets that govern the evolution of viruses: mutation rates, population size, selection, and the MOI. We will explore how viruses evolve within a host, during transmission to novel susceptible hosts, and establish infections in new host species.

Virus mutations create genetic diversity, which is subject to the opposing actions of selection and random genetic drift, both of which are directly affected by the size of the virus population. When the population size is large, selection will be predominant and random drift less common. This means that deleterious alleles will be efficiently removed from the population, while adaptive alleles will have an opportunity to take over the population. However, when the population size is small, random effects may obscure the effects of selection. Under these conditions, slightly deleterious alleles may rise to an unexpectedly high frequency in the population, and adaptive alleles may be lost by chance.

High mutation rates create many viral variants. During an infection with human immunodeficiency virus (HIV), all genotypes that are one mutation away from the infecting genotype will be created every day. The rich cloud of mutants, often termed a "quasispecies," has the potential to encode viruses with elevated resistance to a drug, or the ability to evade neutralizing antibodies created by the host. As a corollary, this complicates efforts to design effective vaccines, as evolution can greatly increase the number of virus serotypes that circulate in human populations. Furthermore, the unique ability of viruses to change allows them to cross species barriers, resulting in zoonotic infections.

Virus evolution is further characterized by additional layers of complexity. One unique characteristic of viruses is their MOI, which is the ratio between the number of viruses and the infecting cells. MOI has several consequences for evolution that are discussed in a later section, and these are subject to the constantly changing size of the virus population. The typical view of viral evolution is that viruses create huge population sizes within the infected host. However, this huge population size is punctuated by frequent bottlenecks during host-to-host transmission, and population

structure within an infected host, where different organs and tissues may support different independently replicating populations. These differences in population size will affect both the selection-drift balance mentioned above, and the MOI of different virus subpopulations. In the rest of this chapter, we discuss the different factors affecting the virus population, and how these factors intertwine to shape virus evolution.

2. MUTATION AND SUBSTITUTION RATES IN VIRUSES

How are mutation rates defined, and what are substitution rates? Why are mutation rates so high? How do they differ for RNA and DNA viruses?

Mutation rate is typically defined as the average number of errors created in genomes of viral progeny, per base, per replication cycle (mut/nuc/rep). Viruses possess mutation rates that are orders of magnitude higher than any other replicating entity (Table 1). These rates range from approximately 1.5×10^{-3} mut/nuc/rep in the RNA bacteriophage Qβ (Batschelet et al., 1976) to ~10^{-8} mut/nuc/rep in the DNA virus Herpes simplex (Drake and Hwang, 2005). These examples highlight the interesting difference between RNA viruses, which replicate with their own RNA-dependent RNA-polymerase (RdRp), and DNA viruses which replicate with either their own or the cellular DNA-dependent DNA polymerase. RdRps all lack the proofreading capabilities present in DNA polymerases, and thus RNA viruses have much higher mutation rates than DNA viruses. Strikingly, it has been found that both increasing and decreasing the mutation rate of a virus leads to reduced virulence of the virus population (Pfeiffer and Kirkegaard, 2005; Vignuzzi et al., 2006). This suggests there is a close link between the mutation rate of a virus, the diversity created in a virus population, and pathogenesis in an infected host.

While mutations create raw genetic diversity, it is the coupled action of mutation and selection that will determine which mutations will *persist* in the viral population. The rate at which mutations fix in a population is termed the *substitution rate*, or evolutionary rate, which is measured by comparing the genomes of different isolates of a virus collected at several different time points (Duffy et al., 2008). Once again, RNA viruses possess much higher substitution rates than DNA viruses, ranging from 0.01 substitutions per site per year (sub/site/yr) in the RNA poliovirus type 1 to 7×10^{-7} in the DNA virus monkeypox. As suggested by theory, in most viruses substitution rates correlate well with mutation rates (Table 1). This suggests that the short-term mutation rate is an important determinant of the rate of long-term molecular evolution. However, for the fastest mutators (mostly RNA viruses), there appears to be an upper limit to the rate of evolution. This is due to the exceptionally high load of deleterious variant viruses in these small RNA viruses, which slows down their rate of molecular evolution. This high load dictates a threshold beyond which populations may go extinct. Indeed, it has been shown that by artificially increasing the mutation rate of different RNA viruses, the population will collapse through a process termed lethal mutagenesis. This finding has led to the development of therapeutic drugs that induce lethal mutagenesis, which are used to treat a variety of viral infections such as Hepatitis C and West Nile virus (Beaucourt and Vignuzzi, 2014).

2.1 Measurement of Mutation Rates

How are mutation rates measured and what are the shortcomings of standard measures?

It is important to define mutation rate in a consistent and unbiased manner. Mutation rates refer to the rate of mutation per site per genome replication, or to the rate of mutation per site per round of viral replication (Duffy et al.,

---

TABLE 1 Some Examples of a Range of Mutation Rates and Evolutionary Rates Determined for Different Groups of Viruses

Group	Virus	Mutation Rates (Mutations per Nucleotide per Replication Cycle)	Evolutionary Rate (Substitutions per Nucleotide Site per Year)
Positive-stranded RNA	Poliovirus 1	2.2×10^{-5}–3×10^{-4}	1.17×10^{-2}
Negative-stranded RNA	Influenza A virus	7.1×10^{-6}–3.9×10^{-5}	9×10^{-4}–7.84×10^{-3}
Retrovirus	Human immunodeficiency virus 1	7.3×10^{-7}–1.0×10^{-4}	1.13×10^{-3}–1.08×10^{-2}
Single-stranded DNA	Bacteriophage phiX174	1×10^{-6}–1.3×10^{-6}	Unknown
Double-stranded DNA	Herpes simplex 1	5.9×10^{-8}	8.21×10^{-5}

Data taken from the VirMut website (http://www.uv.es/rasanve2/virmut.htm) (Sanjuan et al., 2010; Sanjuan, 2012).

2008; Sanjuan et al., 2010). Critically, these two measures can differ when viruses use *stamping machine* replication versus *geometric genome* replication (Figure 1). During stamping machine replication one single virus genome is used as the template for replication, leading to linear accumulation of mutations. With geometric genome replication, progeny strands can become templates for replication themselves, and thus there is an exponential (or geometric) increase in progeny genomes. This will lead to a completely different distribution of mutations in the progeny genomes.

Classically, the mutation rate of an organism is determined in one of the two ways: the Lurie–Delbruck fluctuation test or measurement of mutation accumulation. We describe both methods, their caveats, and present a novel sequencing technique, which has the potential to alleviate some of these problems.

Luria–Delbruck fluctuation assay. In this method, a number of parallel populations are grown in a nonselective environment. Next, a selective environment is used to measure a given phenotype caused by a single mutation (e.g., resistance to phage). In each clonal population, the frequency of mutants with the phenotype is measured. Since mutations arose spontaneously in the parallel clonal populations, it is possible to use the mutant frequency to backtrack the mutation rate. If the number of rounds of replication can be estimated and the number of initial input genomes is known, it is possible to obtain a mutation rate per site per round of replication. There are several caveats in this assay leading to a biased estimation of the mutation rate: Mutations at multiple sites may alter the mutant phenotype; or the number of rounds of replication may be incorrectly estimated when the mode (stamping machine versus geometric) of genome replication is unknown.

Mutation accumulation studies. Multiple lineages of one progenitor strain are propagated over many generations, often with severe bottlenecks between propagations. These bottlenecks should reduce the effectiveness of selection, and thus mutations are expected to accumulate at the unbiased basic mutation rate. Sequencing the input and output viruses identifies the number of accumulated mutations; with the size of the genome known, and the number of rounds of replication estimated, the mutations per site per round of replication can be calculated. The major caveat here is the underlying assumption: If selection does operate during this propagation scheme, even at a minor level, this will skew the calculated mutation rate, often in an unknown way.

Mutation composition in a virus population. With the advent of next-generation sequencing (NGS), it is possible to capture accurate information on rare mutations present in a population. Low MOI will tend to select for viruses that are the "fittest" under the specific growth conditions. Using a sequencing technique that reduces the high error rate of prevalent NGS techniques, Acevedo et al. (2014) could accurately record the frequency of lethal mutations, which are expected to be present in a population at a frequency equal to the basic unbiased mutation rate (Figure 2). Applying this method to Poliovirus 1 populations confirmed a mean mutation rate of 3.97×10^{-4}, consistent with previous measurements. However, these results yield a level of detail previously less appreciated: different pairs of bases are replaced at different rates. A further intriguing study has shown that measurements of viral mutation rates vary substantially when measured across different cell types. Thus, there are previously unappreciated layers of complexity in the ascertainment of viral mutation rates.

3. MULTIPLICITY OF INFECTION

Why is multiplicity of infection so relevant to the accumulation of mutations?

The MOI is the ratio between the number of viruses and the number of cells. When MOI is high, cells are coinfected with multiple viruses, and when MOI is one or lower, each cell is most likely infected by one virus only. High MOI leads to a myriad of complex and contrasting effects (Table 2). First, in recombining viruses (or those that undergo reassortment), high MOI will lead to increased levels of recombination.
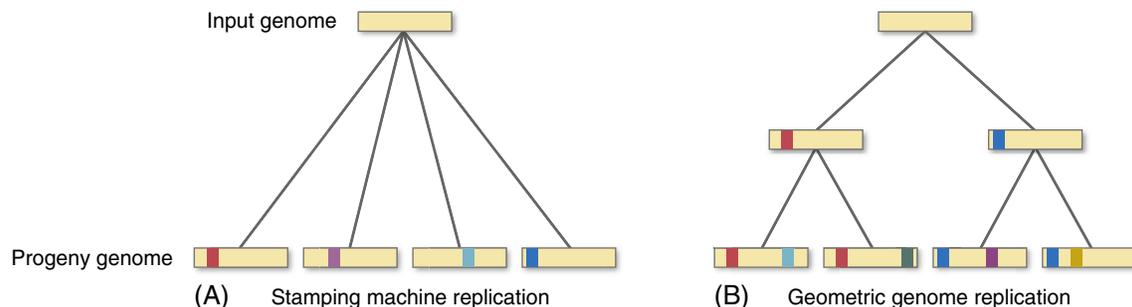
FIGURE 1   An illustration of (A) stamping machine replication, and (B) geometric genome replication. Yellow bars represent genomes, colored boxes represent mutations. In stamping machine replication, a single virus genome is used as the template for replication, leading to linear accumulation of mutations. With geometric genome replication, progeny strands can become templates for replication themselves, and thus there is an exponential increase in the number of mutations in the genomes of newly synthesized virions.
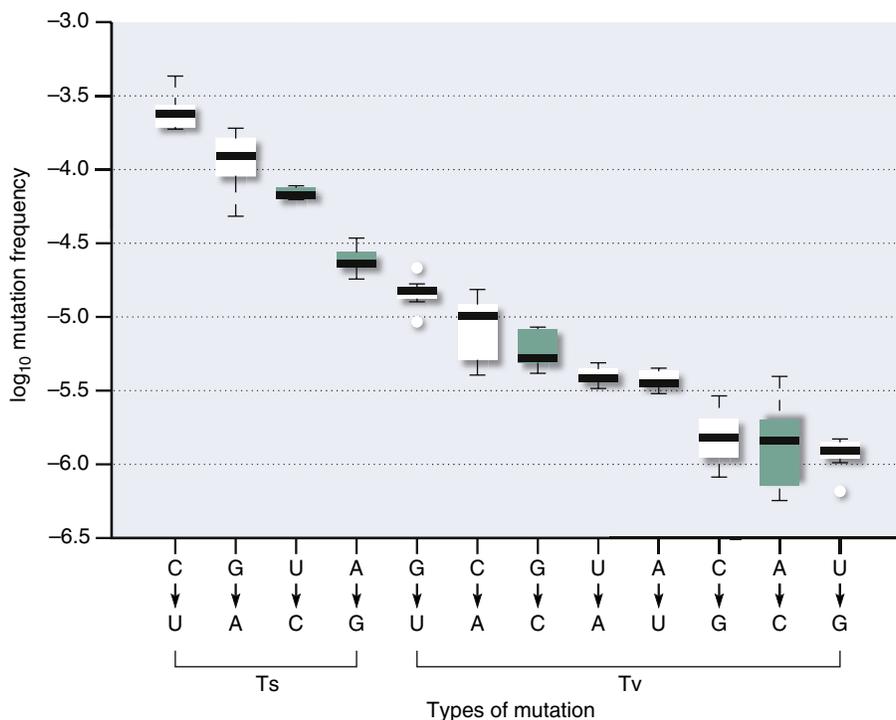
FIGURE 2 An improved method for the determination of virus mutation rates, as applied to poliovirus. This method is based on the premise that sequence data can be used to identify potential lethal mutations: in this case either stop codons within proteins or amino acid changes within catalytic sites of viral enzymes were enumerated. Based on population-genetics theory, the rate of lethal mutations was assumed to be equal to the basic unbiased mutation rate. Open boxes used both kinds of mutations and shaded boxes used only catalytic site mutations. Ts refers to a transition mutation; Tv refers to a transversion mutation (purine to pyrimidine or the reverse). *After Acevedo et al. (2014).*

TABLE 2 The Impact of Multiplicity of Infection (MOI) on Virus Genomic Evolution

Effect on Genome	High MOI	Low MOI
Recombination or reassortment	Frequent	Rare
Rescue of inferior genomes	Frequent	Rare
Rescue and maintenance of defective interfering genomes	Frequent	Rare
Selection for fittest genomes	Mixed	High
Bacteriophage infection	Lysogenic	Lytic

Frequent recombination may lead to more efficient selection, allowing the efficient removal of deleterious alleles and the incorporation of adaptive alleles. High rates of recombination/reassortment may also lead to the emergence of strains with a more virulent phenotype, as is thought to occur in cross-species transmission events of influenza virus. However, high MOI may also have contrary effects, whereby inferior genotypes are rescued, and maintained in the population, by products of superior genotypes. Complementation at high MOI also leads to propagation of defective particles. Thus under high MOI conditions, the beneficial effect of adaptive alleles may be masked (Stern et al., 2014).

High MOI also produces a higher gene copy number, that is multiple genomic copies of the same gene in one infected cell. In phages, copy number variation is highly influential: when copy number is one, phages will be lytic and kill the host cell, whereas if it exceeds one, phages become lysogenic and the bacterial host cell remains alive. Finally, competition for resources at high MOI may also have complex effects on viral replication. In fact, following complementation, viruses enter a "Prisoner's dilemma" regime where selfish genotypes evolve reducing the mean fitness of the viral population (had complementation been absent). Of these contrasting effects, the negative effects of complementation are more dominant that the positive effects of recombination, at least for bacteriophage Φ6 (Froissart et al., 2004) and for polioviruses (Stern et al., 2014). To summarize, it is evident that high MOI leads to complex and conflicting effects on genome selection.

There are a number of other implications of MOI. (1) The distribution of viral particles at different sites of an infection is unknown, and will also directly affect MOI and the efficiency of selection. (2) The MOI and ensuing diversity of the transmitting population affect the probability of establishing an infection in a novel host. Interestingly, the population with highest fitness in the original host does not necessarily fare well in new hosts, while

low-frequency genotypes from the original host may be the ones that prevail in the new host. (3) It is likely that different types of viruses will be affected differently by MOI. For example, persistent viruses will likely replicate at an MOI that is lower than viruses that cause acute infections. Since MOI has such important effects on viral replication, it will require additional research to precisely determine how it affects selection in vivo, viral virulence, and the course of infection.

4. PHYLODYNAMICS: EVOLUTION IN A HOST POPULATION

What are phylodynamics? What is a molecular clock and how is it estimated?

There is an increasing interest, combined with new tools and methodology, to investigate transmission networks caused by viral epidemics. Such studies represent a collaboration between epidemiology and evolutionary biology, and the term phylodynamics has been coined to describe them. Phylodynamic methods are rooted in the powerful methodologies of phylogenetics, which emphasize the phylogenetic tree as key to investigating evolutionary processes. The ever-increasing availability of viral sequences has fueled this field and has made it possible to address a range of questions such as: "when did a virus emerge?," "what is the progenitor strain of a circulating epidemic?," and "what is the timing of the spread of a virus across countries and continents?" The phylodynamic approach has yielded remarkable insights into viral evolution.

Molecular clocks are based on previous observations that the number of nucleotide substitutions accumulates roughly linearly over time. This will be true when most nucleotide substitutions are neutral, and are driven directly by the mutation rate. Phylogeny provides a practical method to calibrate the molecular clock. When a viral phylogeny is reconstructed, it furnishes the distance between an ancestral sequence and an extant sequence in units of nucleotide substitutions. When the nucleotide difference (from the ancestral sequence) of each extant sequence is plotted against time, it is possible to infer the rate of nucleotide substitution and thus track back the date at which the ancestral sequence emerged. As an example, Kew et al. (1995) collected a set of isolates of poliovirus from a 10-year sequential infection chain in South America, and used this to determine the number of mutations per year (9×10^{-3} nucleotide substitutions per site per year for a 150 nucleotide window within the ~7500 nucleotide genome) (Figure 3). Most of these changes were synonymous mutations that would have little, if any, influence on the biological properties of the virus.

Based on the branching of the viral phylogeny, Holmes (2008) has suggested that a number of different patterns can
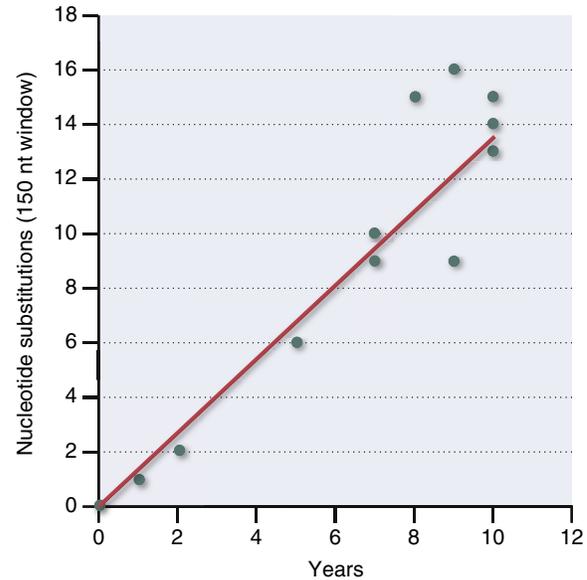
FIGURE 3 Deriving the molecular clock, the substitutions per nucleotide per year, in a host population. In this instance, a wild strain of poliovirus was introduced into a population and isolates were available over a 10-year time period. For each isolate, the number of nucleotide substitutions (from the original introduced strain) was plotted against the date of the isolate, to determine the substitution rate (substitutions per 150 nucleotides per year). *After Kew et al. (1995).*

explain viral dynamics across the globe. Examples include the following:

- *Random mixing*: frequent viral traffic among different regions of the world cause lack of spatial structure in the tree, with no correlation between the geographic location of a sample and its location on the phylogeny
- *Population subdivision*: strong spatial subpopulations of viruses localized to certain regions of the world, suggesting that these populations do not mingle
- *Source–sink transmission*: one viral population acts as the source for all other viral populations in the world.

Different viruses follow distinctly different patterns. For influenza A, a source–sink model of viral population structure was found to best describe the global evolution of the virus. Accordingly, a persistent reservoir in south-east Asia continually seeds epidemics worldwide and drives viral diversity around the globe. On the other hand, hepatitis C virus displays a pattern of population subdivision, consisting of well-defined subtypes with distinct geographical locations.

The phylogenetic analysis yields some of the most striking examples of cross-species emergences. For HIV, as well as for other viruses where sufficient sequence data are available, it is now possible to trace back the evolution of a virus and how it acquired the ability to replicate efficiently in human cells (Sharp and Hahn, 2010). To date, four groups of HIV-1 have been identified: M, N, O, and P, with group M being responsible for the vast majority of infections

worldwide. Phylodynamics showed that the progenitor of HIV-1 was an SIV strain from the chimpanzee strain *Pan t. troglodytes*, denoted SIVcpz-Ptt, as illustrated in Figure 4. Using this phylogeny, it was further possible to identify the consensus ancestor of HIV-1. Using sequences of dated isolates of HIV-1, a molecular clock was calibrated, and was used to calculate that the common ancestor of HIV-1 group M arose between 1910 and 1930.

Phylogenetic studies have provided insights about known pathogens such as SARS Coronavirus, the H5N1 avian influenza, monkeypox, and others, which have emerged in the human population following a cross-species transfer (see Chapter 16, Emerging viral diseases).

In a recent epidemic of Ebola virus, NGS compared the sequence of viruses obtained at different times during an outbreak in Mali (Hoenen et al., 2015). It appeared that the mutation rate was similar to that observed in prior outbreaks of Ebola virus. Importantly, there was no evidence that the 2013 epidemic strain of Ebola virus was evolving toward a strain with increasing transmissibility or virulence.

Multiple examples, together with several other lines of evidence (Parrish et al., 2008) highlight some important concepts in cross-species transfers: (1) The host range of a virus is usually well defined, and viruses only rarely gain the ability to spread efficiently to a new host. (2) The ability to spread to a new host species is compounded of many different factors, ranging from the probability of contact, demographic factors, the ability to bind a receptor, and the ability to overcome intracellular host restriction factors (Strebel, 2013). (3) Cross-species transmission events often involve crucial genetic adaptations on the part of the virus.

4.1 Drivers of Virus Evolution

What drives virus evolution? What role does the host response play in virus evolution? How does this impact pathogenesis?

As already noted, there are potentiators on viral evolution that have important implications for the control of viral
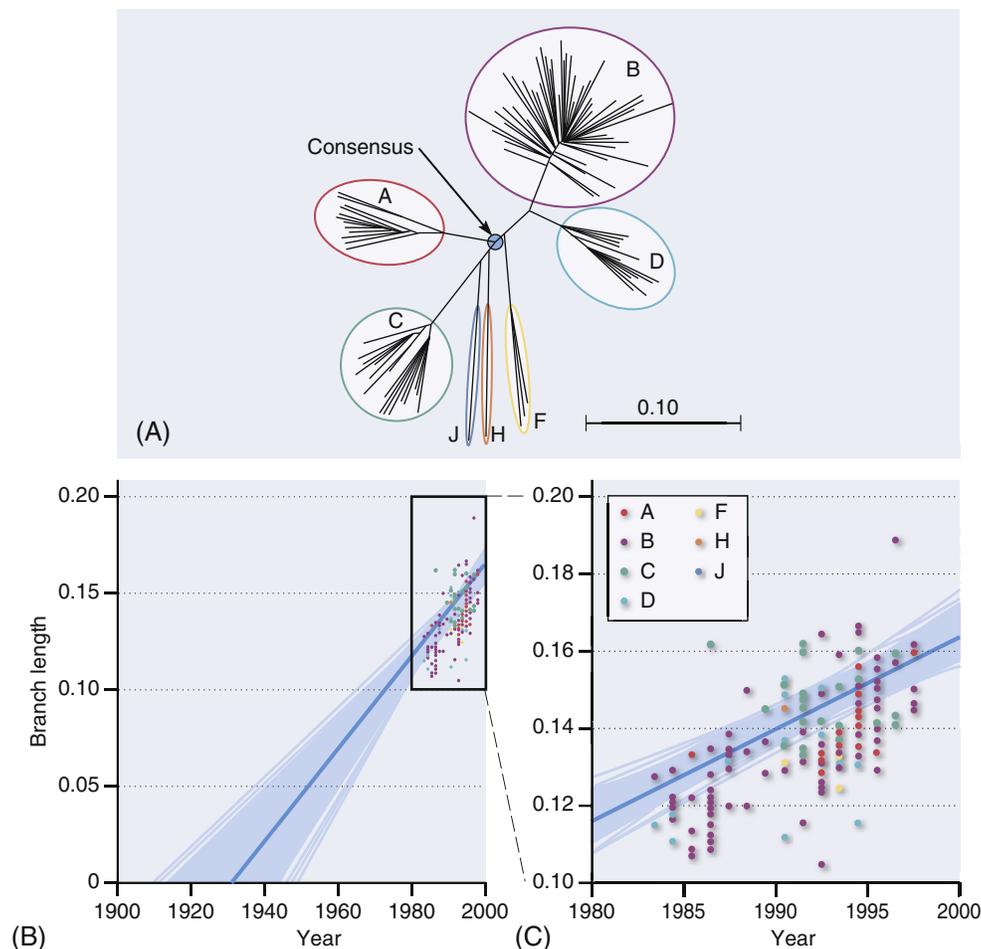
FIGURE 4 A phylogenetic tree of gp160 (the HIV surface glycoprotein) sequences from HIV-1 group M isolates. These data were used to identify an ancestral consensus sequence. Panel B: The nucleotide distances of individual virus isolates from the consensus sequence are plotted against their date of isolation. These data are used to project back to the date of the consensus sequence, giving a range from about 1915 to 1940. Panel C: A magnified view of the box in panel B. *Redrawn from Korber et al. (2000).*

diseases of humans. Two notable examples are escape from the host immune response and from antiviral drugs.

Immune escape mutants. A large armamentarium of effective viral vaccines have been developed over the last 75 years (see Chapter 19, Viral vaccines). These vaccines are based on "ancient" isolates (often at least 50 years old) of viruses which induce homologous neutralizing antibodies. Since these vaccines continue to provide protection against currently circulating viral strains, it must be concluded that the cognate viruses cannot evolve antibody-escape mutants that are sufficiently "fit" to survive and replace the original circulating strains. There are two notable exceptions to this general rule, influenza and HIV viruses. In both instances, antibody-escape mutants can survive to circulate in the population. For this reason, influenza virus continues to "drift" and new virus isolates must be used each year to produce protective vaccines. HIV has resisted the development of an effective vaccine mainly due to the difficulty of inducing "broadly neutralizing" antibodies that are effective against the population of escape mutants that have evolved over the last 30 years. In both instances, there is a structural explanation. Neutralizing antibodies to HIV and influenza viruses are not necessarily directed at the receptor-attachment site on the viral envelope, so neutralization escape mutants can still attach to cellular receptors, replicate, and persist (Figure 5).

Antiviral drugs. A vast experience with antiretroviral drugs for the treatment of HIV has shown that the virus swarm contains about 10^{-5} variants that will resist any single drug. Many such mutants will be produced each day in an HIV-infected patient (see Chapter 20, Antiviral therapy and Chapter 15, Mathematical modeling). For this reason, it has been found that effective control of HIV infections in humans usually requires the simultaneous use of at least three drugs, each of which requires a different set of mutations to escape drug control. The frequency of three escape mutations in a single genome is so low (~10^{-15}) that the drug combination will suppress HIV replication.

Do viruses evolve to a benign relationship with their hosts?

How does virulence influence viral evolution? It has been proposed that when a virus and host undergo co-evolution over a long period of time, a benign relationship will develop such that the virus will not cause disease. SIV is cited as an example. SIV strains are relatively benign in their natural simian hosts, where they have presumably been long established. However, when the SIVsmm strain crossed from its natural host—sooty mangabeys—into macaques, a novel host, it caused an AIDS syndrome.

However, a survey of mammalian viruses and their cognate hosts suggest that there is no necessary correlation between virulence and coevolution. Long-established virus infections range from inapparent to fatal. For instance, in the prevaccine era, poliovirus paralyzed only 1 person in every 150 infected (149 infections were inapparent), while rabies is 100% fatal in most of its animal and human hosts. Prior to smallpox eradication, there were two strains of variola virus. Variola major caused 30% mortality but variola minor only 1% mortality; yet each strain was maintained in the human population. It appears that viruses have used many strategies to perpetuate themselves in their host populations. Some strategies are benign, while others cause serious disease in their hosts.

5. SUMMARY AND CHALLENGES AHEAD

This chapter addresses some of the key features in viral evolution. (1) Viruses have mutation rates that are higher than any other member of the kingdom of life. This gives them the ability to evolve, even within the course of a single infection, and to evade multiple host defenses, thereby impacting pathogenesis. (2) There are several methods to estimate the mutation rates of viruses (mutations per nucleotide per genome replication), but each has its limitations. New improved methods utilize next-generation sequencing to take into account the large "swarm" of viral quasispecies. (3) The MOI has a critical effect on viral evolution. Low MOI will favor selection of the fittest viruses, while high MOI can have several opposing effects, and the net result is difficult to predict. (4) Phylodynamics, the collaboration between evolutionary biology and epidemiology, has generated data on the molecular clock that captures the rate of viral genome evolution (nucleotide substitutions per base per year) in host populations. (5) The drivers of viral evolution include
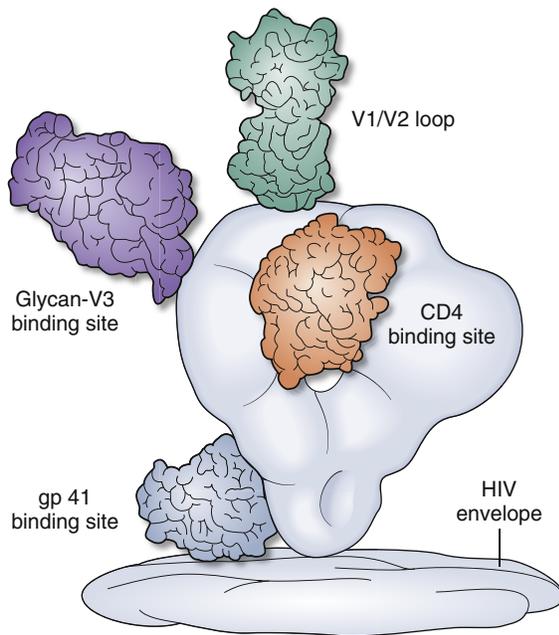
FIGURE 5 Neutralizing epitopes can be located distant from the cellular receptor-binding site on the virion, in which case antibody-escape variants can replicate and persist. This example is a cartoon of the principal neutralizing epitopes on the HIV envelope protein. Orange: CD4 binding site; purple: Glycan-V3 binding site; green: V1/V2 loop; gray: gp41 binding site. *Redrawn from Klein et al. (2013).*

evasion of host defenses, escape from antiviral drugs, and circumvention of vaccine-induced immunity.

What does the future hold for the study of virus evolution? For the first time, virus evolution can be informed by computational modeling based on experimental data. Evolutionary studies have now begun to compare related viruses that infect phylogenetically similar species, yielding a wealth of insights in viral evolution and host responses (Daugherty and Malik, 2012; Sawyer and Elde, 2012). A recent study has shown that evolution of viruses is dictated by strict protein constraints (Wylie and Shakhnovich, 2011), underlining the impact of protein structure on viral evolution. The ability of viruses to create antigenic variation is key to understanding epidemiologic dynamics, and the striking differences between viruses in their ability to escape immune responses reflects underlying structural and genomic determinants yet to be explained.

The increasing numbers of viral sequences has led to unprecedented observation of viruses as they evolve, from laboratory experiments, from evolution within a host, and from epidemiological sequences of patients around the globe. The integration of rich NGS data on evolving virus populations is opening the door to a better understanding of factors that facilitate adaptation and lead to disease. By establishing the rules that govern viral evolution, research is empowering the design of new strategies that control, treat, and possibly eradicate viral threats.

FURTHER READING

Acevedo, A., Brodsky, L., and Andino, R. (2014). Mutational and fitness landscapes of an RNA virus revealed through population sequencing. Nature 505, 686–690.

Batschelet, E., Domingo, E., and Weissmann, C. (1976). The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. Gene 1, 27–32.

Beaucourt, S., and Vignuzzi, M. (2014). Ribavirin: a drug active against many viruses with multiple effects on virus replication and propagation. Molecular basis of ribavirin resistance. Current opinion in virology 8C, 10–15.

Coffin, J.M. (1995). HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science 267, 483–489.

Daugherty, M.D., and Malik, H.S. (2012). Rules of engagement: molecular insights from host-virus arms races. Annu Rev Genet 46, 677–700.

Drake, J.W., and Hwang, C.B. (2005). On the mutation rate of herpes simplex virus type 1. Genetics 170, 969–970.

Duffy, S., Shackelton, L.A., and Holmes, E.C. (2008). Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet 9, 267–276.

Froissart, R., Wilke, C.O., Montville, R., Remold, S.K., Chao, L., and Turner, P.E. (2004). Co-infection weakens selection against epistatic mutations in RNA viruses. Genetics 168, 9–19.

Hoenen T.1, Safronetz D.1, Groseth A.1, Wollenberg K.R.2, Koita O.A.3, Diarra B.3, Fall I.S.4, Haidara F.C.5, Diallo F.5, Sanogo M.3, Sarro Y.S.3, Kone A.3, Togo A.C.3, Traore A.5, Kodio M.5, Dosseh A.6, Rosenke K.1, de Wit E.1, Feldmann F.7, Ebihara H.1, Munster V.J.1, Zoon K.C.8, Feldmann H.9, Sow S.10. Virology. Mutation rate and genotype variation of Ebola virus from Mali case sequences. Science. April 3, 2015;348(6230):117–119. doi: 10.1126/science.aaa5646. Epub 2015 Mar 26.

Holmes, E.C. (2008). Evolutionary history and phylogeography of human viruses. Annu Rev Microbiol 62, 307–328.

Kew O.M., Mulders M.N., Lipskaya G.Y., da Silva E.E., Pallansch M.A. Molecular epidemiology of poliovirus. Virology 1995, 6: 401–414.

Kirchhoff, F. (2010). Immune evasion and counteraction of restriction factors by HIV-1 and other primate lentiviruses. Cell host & microbe 8, 55–67.

Klein F., Mouquet H., Dosenovic P., et al. Antibodies in HIV-1 vaccine development and therapy. Science 2013 341: 1199–1204.

Korber B., Muldoon M., Theiler J., et al. Timing the ancestor of the HIV-1 pandemic strains. Science 2000, 288: 1789–1796.

Parrish, C.R., Holmes, E.C., Morens, D.M., Park, E.C., Burke, D.S., Calisher, C.H., Laughlin, C.A., Saif, L.J., and Daszak, P. (2008). Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol Mol Biol Rev 72, 457–470.

Sanjuan, R. (2012). From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. PLoS pathogens 8, e1002685.

Sanjuan, R., Nebot, M.R., Chirico, N., Mansky, L.M., and Belshaw, R. (2010). Viral mutation rates. Journal of virology 84, 9733–9748.

Sawyer, S.L., and Elde, N.C. (2012). A cross-species view on viruses. Current opinion in virology 2, 561–568.

Sharp, P.M., and Hahn, B.H. (2010). The evolution of HIV-1 and the origin of AIDS. Philosophical transactions of the Royal Society of London Series B, Biological sciences 365, 2487–2494.

Sharp, P.M. and Hahn, B.H. Origins of HIV and the AIDS pandemic. Cold Spring Harb. Perspect. Biol., 1: a006841, 2011, 1–22.

Stern, A., Bianco, S., Yeh, M.T., Wright, C., Butcher, K., Tang, C., Nielsen, R., and Andino, R. (2014). Costs and benefits of mutational robustness in RNA viruses. Cell reports 8, 1026–1036.

Strebel, K. (2013). HIV accessory proteins versus host restriction factors. Current opinion in virology 3, 692–699.

Wylie, C.S., and Shakhnovich, E.I. (2011). A biophysical protein folding model accounts for most mutational fitness effects in viruses. Proc Natl Acad Sci USA 108, 9916–9921.