The complexities of viral mutation rates

Kayla M. Peck[a] and Adam S. Lauring[a,b,*]

Division of Infectious Diseases, Department of Internal Medicine[a], Department of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan

*Corresponding Author: Adam S. Lauring (alauring@med.umich.edu)

University of Michigan
1150 W. Medical Center Dr.
MSRB1 5510B
Ann Arbor, MI 48109
734-764-7731

Abstract word count: 61
Text word count: 3549

Abstract

Many viruses evolve rapidly. This is due, in part, to their high mutation rates. Mutation rate estimates for over 25 viruses are currently available. Here, we review the population genetics of virus mutation rates. We specifically cover the topics of mutation rate estimation, the forces that drive the evolution of mutation rates, and how the optimal mutation rate can be context-dependent.

Keywords: virus evolution; viral mutation rates; mutation rate evolution; polymerase fidelity

What are viral mutation rates and why are they important?

Many viruses have high rates of evolution. These high evolutionary rates have been attributed to the large population sizes, short generation times, and high mutation rates of viruses. Mutation rate, specifically, is an important determinant of evolutionary rate across taxa (1-4). In the context of viruses, mutation rate is the rate at which errors are made during replication of the viral genome. This is in contrast to substitution rate, which is the rate at which mutations become fixed, or present within all individuals, in a population. Whereas mutation rates are used to estimate the amount of genetic diversity generated within a population of offspring, substitution rates are used to estimate the rate of evolution for a particular lineage or taxon.

In population genetics, an important parameter is the genomic mutation rate, which is the product of the per-nucleotide site mutation rate and genome size. The genomic mutation rate determines the average number of mutations each offspring will have compared to the parental (or ancestral) genome. On a per-site level, DNA viruses typically have mutation rates on the order of 10^{-8} to 10^{-6} substitutions per nucleotide site per cell infection (s/n/c). RNA viruses, however, have higher mutation rates that range between 10^{-6} and 10^{-4} s/n/c (**Figure 1**). Despite variable per-site rates, species with smaller genomes exhibit a negative correlation between

genomic mutation rate and genome size, such that the per-genome mutation rate is relatively constant (**Figure 1C**) (5, 6).

The higher per-site mutation rates of RNA viruses can be explained, in part, by the RNA-dependent RNA polymerases (RdRp) that replicate their genomes. Unlike many DNA polymerases, RdRp do not have proofreading activity and are thus unable to correct mistakes during replication. Notable exceptions are members of the Nidovirales family, including corona-, toro-, and roniviruses, which have an RdRp-independent proofreading activity and thus lower mutation rates. This proofreading is thought to be a key factor in explaining how these viruses have much larger genomes (>26 kb) compared to other RNA viruses (7). Retroviruses also have high mutation rates, because reverse transcriptase, like most RdRp, lacks proofreading activity. Finally, for unclear reasons, single-stranded viruses tend to mutate more rapidly than double-stranded viruses, causing some ssDNA viruses to have rates comparable to those of dsRNA viruses (**Figure 1A**) (5).

Mutation rates determine the amount of genetic variation generated in a population, which is the material upon which natural selection can act. For this reason, a higher mutation rate correlates with a higher evolutionary rate, but only to a point (**Figure 1A-B**). While the high mutation rates of retroviruses and RNA viruses may explain their higher evolutionary rates relative to DNA viruses, several DNA viruses exhibit evolutionary rates comparable to those of RNA viruses (3, 5). This highlights the importance of additional factors in determining the evolutionary rate, such as within-host dynamics (4) or cell tropism (8). Overall, mutation rates are important, because they determine the probability that a mutation conferring drug resistance, antibody escape, or expanded host range will arise. Additionally, mutation rates can determine whether a virus population will be susceptible to drug-induced lethal mutagenesis (9).

Here, we discuss viral mutation rates from a population genetics perspective, including informative work from non-viral systems. Specifically, we address how mutation rates are measured and how evolutionary forces shape viral mutation rates over different time scales. For further discussion on other aspects of viral mutation rates, we refer the reader to articles on polymerase structure and function (10-12), viral mutational fitness effects (13-16), evolutionary rates (2-4, 8, 17), and genome evolution (7, 18).

How are viral mutation rates measured?

The mutation rates of over 25 viruses have been experimentally measured in a laboratory setting. These rates range from 10^{-8} to 10^{-6} s/n/c for DNA viruses and 10^{-6} to 10^{-4} s/n/c for RNA viruses (5, 19). Importantly, even though mutation rates are reported as a single value, each of the 12 mutation classes will have its own rate. In the literature, polymerase error is typically measured and expressed as mutation frequency, mutation rate, or misincorporation rate (**Table 1**). Due to the variety of methods for measuring mutation rates, it is important to consider what is being measured and the strengths and limitations of each approach.

Mutation frequency refers to the proportion of mutants identified in a virus sample or population. As such, mutation frequency is biased toward non-lethal mutations, and typically captures only those mutations that are able to persist in the population. This means that frequency estimates can be confounded by selection and drift. When a *de novo* mutation appears in a population of size *N*, its initial frequency is 1/*N*. Because virus populations are typically very large, this makes the initial frequency of *de novo* mutations extremely low. Under this scenario, most mutations that increase in frequency to a detectable level are either beneficial or selectively neutral. However, we know from studies of viral mutational fitness effects that most mutations are lethal or deleterious, a minority are neutral, and only a few are beneficial (13, 15, 16). Thus, methods

that rely on mutation frequencies are biased against lethal and/or deleterious mutations and may dramatically underestimate the mutation rate.

Common methods for measuring mutation frequencies include the sequencing of clones, populations, or mutation accumulation lines (**Table 1**). The most common approach is to sequence individual clones or perform high depth of coverage sequencing of an entire population after a defined number of passages. Advantages of these methods include the ability to capture mutational spectra (i.e., estimates of the 12 mutation classes) and context-dependent effects. Although sequencing approaches are powerful, disadvantages include an inherent bias toward sampling genomes with higher fitness and the possibility of sequencing errors being mislabeled as mutations. Furthermore, for RNA viruses, genomes must undergo a reverse transcription step prior to sequencing. The impact of reverse transcription error on the processing of RNA virus samples likely leads to high levels of background errors (20) and may skew estimates of mutational bias. Newer sequencing methods, such as circular sequencing (21), may improve accuracy by making it easier to exclude sequencing errors from the final analysis.

Mutation accumulation involves subjecting virus populations to serial bottlenecks, as in plaque-to-plaque transfers, to avoid selection against deleterious mutations. Plaque-to-plaque transfers involve selecting a single viral plaque at random to inoculate the next passage. By continuously selecting a single plaque each passage, deleterious mutations can become fixed in the population and are thus detected upon sequencing. Notably, as a lineage accumulates multiple mutations, selection may play a more prominent role because additional deleterious mutations will reduce the fitness such that a plaque is unable to form. While this method, overall, is less biased against deleterious mutations, it still cannot capture lethal mutations. However, it is possible to adjust the mutation frequency to account for lethal mutations by leveraging data on

viral mutational fitness effects (19). Furthermore, if enough passages are performed, mutational spectra and context-dependent effects can be captured.

Whereas mutation frequency is the proportion of mutants in a population, mutation rate is the rate at which mutations occur in a viral genome. Mutation rates are reported as either substitutions per nucleotide per strand copying (s/n/r) or substitutions per nucleotide per cell infection (s/n/c), where cell infection is an estimate of a viral generation. The distinction between these units is based on whether viruses replicate via a “stamping machine” model, where a single template is used, or if replication is semi-conservative, where replicated strands themselves act as templates (5). Many assays will capture s/n/c, which can be converted to s/n/r if the replication mode (e.g., number of copying cycles) is known (19). Alternatively, fluctuation tests, as described below, directly measure the s/n/r regardless of the mode of replication.

The fluctuation test is a method for mutation rate estimation that is derived from the classic Luria-Delbruck experiment (22). This method requires the presence of a marker that can acquire mutations linked to a scorable phenotype, such as resistance to a drug. While the ideal marker is selectively neutral, fluctuation tests are often robust to ones that are not. One important advantage of this method is that it is less biased against lethal mutations, allowing for an estimate that may more accurately capture the true mutation rate. Additionally, fluctuation tests avoid sequencing errors and the error-prone reverse transcription step in the processing of RNA virus samples. Disadvantages of this method include bias toward the subset of sites and mutational classes that yield the scorable phenotype and sampling of only one location in the genome, which limits the detection of context-dependent effects (**Table 1**). To overcome the disadvantage of a limited mutational spectrum, Pauly et al. (2017) expanded on the traditional fluctuation test to individually probe all 12 mutational classes. Briefly, a GFP gene was

introduced into the influenza A virus genome. Three codons crucial for the fluorescence of the GFP protein were mutated, such that only reversion of a specific nucleotide could result in restoration of the fluorescent phenotype. Advantages of this method include the ability to measure an independent rate for all 12 mutation classes and to explore context-dependent effects by changing the location of the GFP gene.

While the above approaches estimate mutation rates in cells, several cell-free systems allow for measurement of viral misincorporation rates. Here, the nucleotide misincorporation (or discrimination) rate of purified polymerases can be directly quantified in a biochemical reaction. A variety of techniques have been developed and used to investigate RdRp fidelity for many RNA viruses (11, 23). Because these assays define misincorporation dynamics independent of the mutation's effect on the virus, they are less biased against lethal and deleterious mutations. Polymerase kinetics, such as elongation speed, can also be easily measured. However, measurements of misincorporation rates from cell-free assays are often quite different from mutation rate estimates in cell-based assays, although still well-correlated (11) (**Table 1**).

How do viral mutation rates evolve?

Mutation rates are clearly subject to natural selection and can evolve over time. But what forces drive their evolution? Fidelity variants are useful tools for understanding how mutation rates evolve. A small number of fidelity variants have been identified for DNA viruses (24), but many more have been identified in RNA viruses (25-27). Fidelity variants for DNA organisms can have as high as a 100- to 400-fold change in mutation rate while RNA virus fidelity variants have between 2 and 20-fold changes in mutation rate (25, 26). Fidelity variants with an increase in mutation rate are known as mutators, whereas those with a decrease in mutation rate are known as antimutators. Theoretical studies have shown that the strength of the mutator, or magnitude of change in the mutation rate, can play an important role in both its behavior in the

population and the population's rate of adaptation (28, 29), highlighting an important difference between mutators in DNA versus RNA virus populations.

Mutation rates can evolve through the fixation of mutator and antimutator alleles in a population. In theory, when a population is well-adapted to its environment, mutators are not favored (but see 30). Because most mutations are deleterious, mutators will suffer from an increased mutation load and lose out to competitors with a lower mutation rate. Thus, selection against mutators should result in mutation rates being pushed ever lower. Given the detrimental effect of mutational load, a longstanding question in evolutionary biology is why mutation rates have not evolved to be zero. Three major hypotheses have been presented to address this question (**Figure 2**). One explanation, known as the drift-barrier hypothesis, suggests that genetic drift, or the stochastic loss of alleles, prevents selection from being able to reduce the mutation rate to zero (6). The impact of drift relies upon the effective population size, which captures the timescale of population turnover (31) and can roughly be thought of as the number of genomes that contribute genetic material to the next generation. When the effective population size is low, as is thought to be the case for many viral populations (31), selection is weak, and drift could play a large role in the random fixation of mutator alleles. In contrast, large effective populations are subject to less drift, allowing selection to more effectively reduce the mutation rate (**Figure 2**). The drift-barrier hypothesis has been well-studied in other species (6), but has not yet been thoroughly investigated in the context of viruses.

A second hypothesis is that a physicochemical limit prevents polymerases from attaining perfect fidelity (32). Lower mutation rates impose a higher biochemical cost, for example, by requiring more resources to be spent on producing proteins or protein systems that prevent, or fix, mutations. When the mutation rate is low enough, this cost is so high that a system with perfect fidelity will never be favored in the population (**Figure 2**). This hypothesis, however, is not well-

supported, because we would expect it to act similarly in species with comparable polymerases, and yet we see high levels of variability in the mutation rates of these organisms.

A third hypothesis is that there is countervailing selection *for* higher mutation rates. Two main selective forces that might favor higher mutation rates are increased adaptability and faster replicative speed (**Figure 2**). In the former, a higher mutation rate could be favored because it enhances the adaptability of a population (33). More mutations result in more genetic variation upon which selection can act. Therefore, a genetic background harboring a mutator allele is more likely to produce a beneficial mutation, such as one that mediates escape from host defenses, than those backgrounds with a lower mutation rate. This beneficial mutation is important because it allows the mutator allele to increase in frequency by a process known as hitchhiking. Here, a mutator allele can increase in frequency to fixation by hitchhiking along with a beneficial mutation that it produced, thereby increasing the mutation rate of the population as a whole (34). Importantly, and in contrast to quasispecies models, the mutator allele is not selected for directly. Instead, it gets carried along by selection for a beneficial mutation to which it is linked. Because a mutator allele must remain linked to a beneficial mutation to hitchhike to fixation, this is more likely to occur in virus populations with low recombination rates. When recombination is present, a mutator allele is likely to be separated from its beneficial mutation and no longer benefit from indirect selection.

The adaptability model for high mutation rates is tempting because viruses are constantly facing strong selective pressures from the immune system. When an organism is well-adapted to its environment and no selective pressures are present, all mutations are neutral or deleterious. However, in the face of a selective pressure, beneficial mutations become available. Higher mutation rates would increase the probability of producing a beneficial allele. This is supported by numerous experiments performed in yeast and bacteria showing that a mutator allele will

increase in frequency in a population when a selective pressure, such as an antibiotic, is present. As above, the mutator allele reaches fixation by hitchhiking along with a beneficial mutation (i.e. one mediating antibiotic resistance) and not because of direct selection for adaptability. However, once that selective pressure is no longer present, the supply of beneficial mutations is dramatically reduced, and the population favors a lower mutation rate. Thus, higher mutation rates are thought to be favored when environments are fluctuating, such that beneficial mutations are always available (29, 35). However, theoretical work has shown that the highest levels of adaptability occur at intermediate mutation rates, regardless of how rapidly the environment fluctuates (1).

As an alternative hypothesis, a higher mutation rate might be favored because it correlates with increased replication speed (18, 36, 37). According to this model, a virus that is replicating its genome faster will produce more copies of itself and have a higher fitness than a slower-replicating competitor. However, that speedy replication will, as a byproduct, result in more errors being produced during the replication process. This model may be particularly applicable to RNA viruses (18, 36, 38). With short genomes and large burst sizes, a faster replicator would have a substantial fitness advantage over a slower one. Here, changes in mutation rate are subject to a speed-fidelity tradeoff and are explained by the kinetic proofreading hypothesis. This model posits that the inclusion of a time delay between the formation of the RdRp enzymatic complex and the incorporation of new nucleotides can act as a crude proofreading mechanism and result in increased fidelity (39). Indeed, biochemical assays of purified RdRp have shown that faster polymerases have lower fidelity (11). Additionally, studies in RNA viruses have shown a cost to replication fidelity, with more faithful replication resulting in a direct fitness cost relative to wildtype (37).

One notable point for either hypothesis (adaptability or speed) is the importance of context for the evolution of high mutation rates. Selection varies through time, likely between hosts or even within a host, such that the selective pressures acting on mutation rates also vary over time. When we analyze a sample of a natural population, that sample is just a snapshot in evolutionary time. This analysis occurs outside of the context of selection, and it is not clear what forces are acting on the mutation rate or what mutation rate is favored. For example, computational studies have found that a high mutation rate can enhance adaptation in the short-term, but ultimately causes extinction in the long-term due to the accumulation of deleterious mutations (40). Overall, this subject can be quite complex, with the evolution of higher mutation rates potentially dependent on population size, mutator strength, the supply of beneficial mutations, the deleterious mutation rate, life-history properties, the frequency of environmental fluctuations, and the time frame being considered (28, 29, 35, 38, 41).

Separating the causes and consequences of mutation rate evolution has been widely discussed in the fields of virology and population genetics (34, 36). Further studies that can disentangle specific mechanisms will help to elucidate what selective forces shape the observed mutation rates and which of the above hypotheses might be the most applicable to viruses.

Future Directions

Within the last decade, numerous studies have expanded our knowledge of virus mutation rates and highlighted some of the complexities associated with their measurement and evolution. Mutation rates can be estimated using a variety of methods (**Table 1**) and newer approaches have the requisite accuracy and power for comparative studies (20, 21, 42). Although most mutation frequency and rate estimates do not capture context-dependent effects or mutational spectra, recent studies have addressed these factors explicitly (20, 42, 43). Both of these factors may be important for considering the ability of a virus to evade the immune system,

evolve resistance to antivirals, or emerge into a new host species. Improvements in the measurement of mutation rates will enhance our ability to understand mechanisms of fidelity and build on a strong foundation of structural studies (10-12).

It is clear that mutation rates are subject to natural selection and that deleterious mutation load drives mutation rates lower. What is less clear is why the mutation rates of many viruses are so high, and why there is so much variation in mutation rates among species. It is often assumed that wildtype mutation rates represent an “optimum,” simply because selection has been acting on a virus’ mutation rate for a long period of time. However, the concept of an optimum mutation rate is context-specific. In the literature, the optimum mutation rate has been variably defined as the rate that maximizes the fitness (38, 44), growth rate (33), or adaptability (1, 36, 45) of a population. For any of these definitions, what may be considered optimal in the short-term may not be optimal in the long-term (40). In fact, computational studies have shown that selection can result in a mutation rate that is suboptimal, specifically in terms of long-term adaptability (45). Other studies suggest that some viruses replicate at an optimal mutation rate, but specifically within the context of immune escape (33). Again, the definition of what is optimal is important – the rate that maximizes immune escape may not be optimal in the context of transmission, within-host persistence, or over longer evolutionary time scales. The wildtype rates that we observe today are mere snapshots from a complex evolutionary history, with mutation rates likely changing over time within a virus species.

The long-term stability of viral mutation rates is important for understanding the tractability of using fidelity variants as vaccine candidates (26) and how virus populations will respond to the application of a mutagen, such as those used in lethal mutagenesis (9). Accurate measurements of viral mutation rates, identification of naturally arising fidelity variants, and evaluation of mutation rates over time will improve our understanding of whether mutation rates

are likely to change in the future. These approaches, combined with experimental evolution and deeper mechanistic studies of polymerase fidelity will better reveal the complexities of viral mutation rates.

Acknowledgments

We apologize to colleagues whose work was not directly cited in this article due to space limitations. We thank Ben Sprung and Mike Imperiale for their critical review of the manuscript.

Funding Acknowledgement

The work of the authors is supported by R01 AI118886, a Burroughs Wellcome Fund Investigators in the Pathogenesis of Infectious Diseases Award (to ASL), and NIH T32 AI007528 (to KMP).

References

1. Orr HA. 2000. The rate of adaptation in asexuals. Genetics 155:961-8.
2. Duffy S, Shackelton LA, Holmes EC. 2008. Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet 9:267-76.
3. Sanjuan R. 2012. From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. PLoS Pathog 8:e1002685.
4. Peck KM, Chan CH, Tanaka MM. 2015. Connecting within-host dynamics to the rate of viral molecular evolution. Virus Evol 1:vev013.
5. Sanjuan R, Domingo-Calap P. 2016. Mechanisms of viral mutation. Cell Mol Life Sci 73:4433-4448.
6. Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, Foster PL. 2016. Genetic drift, selection and the evolution of the mutation rate. Nat Rev Genet 17:704-714.
7. Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. 2006. Nidovirales: evolving the largest RNA virus genome. Virus Res 117:17-37.
8. Hicks AL, Duffy S. 2014. Cell tropism predicts long-term nucleotide substitution rates of mammalian RNA viruses. PLoS Pathog 10:e1003838.
9. Perales C, Martin V, Domingo E. 2011. Lethal mutagenesis of viruses. Curr Opin Virol 1:419-22.
10. Ferrer-Orta C, Arias A, Escarmis C, Verdaguer N. 2006. A comparison of viral RNA-dependent RNA polymerases. Curr Opin Struct Biol 16:27-34.
11. Campagnola G, McDonald S, Beaucourt S, Vignuzzi M, Peersen OB. 2015. Structure-function relationships underlying the replication fidelity of viral RNA-dependent RNA polymerases. J Virol 89:275-86.

12. Venkataraman S, Prasad B, Selvarajan R. 2018. RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. Viruses 10.
13. Domingo-Calap P, Cuevas JM, Sanjuan R. 2009. The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. PLoS Genet 5:e1000742.
14. Sanjuan R. 2010. Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. Philos Trans R Soc Lond B Biol Sci 365:1975-82.
15. Cuevas JM, Domingo-Calap P, Sanjuan R. 2012. The fitness effects of synonymous mutations in DNA and RNA viruses. Mol Biol Evol 29:17-20.
16. Visher E, Whitefield SE, McCrone JT, Fitzsimmons W, Lauring AS. 2016. The Mutational Robustness of Influenza A Virus. PLoS Pathog 12:e1005856.
17. Jenkins GM, Rambaut A, Pybus OG, Holmes EC. 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J Mol Evol 54:156-65.
18. Belshaw R, Gardner A, Rambaut A, Pybus OG. 2008. Pacing a small cage: mutation and RNA viruses. Trends Ecol Evol 23:188-93.
19. Sanjuan R, Nebot MR, Chirico N, Mansky LM, Belshaw R. 2010. Viral mutation rates. J Virol 84:9733-48.
20. Pauly MD, Procario MC, Lauring AS. 2017. A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses. Elife 6.
21. Acevedo A, Andino R. 2014. Library preparation for highly accurate population sequencing of RNA viruses. Nat Protoc 9:1760-9.
22. Luria SE, Delbruck M. 1943. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. Genetics 28:491-511.
23. Dulin D, Arnold JJ, van Laar T, Oh H, Lee C, Perkins AL, Harki DA, Depken M, Cameron CE, Dekker NH. 2017. Signatures of nucleotide analog incorporation by an RNA-depenent RNA polymerase revealed using high-throughput magnetic tweezers. Cell Rep 21:1063-76.
24. Mansky LM, Cunningham KS. 2000. Virus mutators and antimutators: roles in evolution, pathogenesis and emergence. Trends Genet 16:512-7.
25. Pfeiffer JK, Kirkegaard K. 2005. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. PLoS Pathog 1:e11.
26. Graham RL, Becker MM, Eckerle LD, Bolles M, Denison MR, Baric RS. 2012. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. Nat Med 18:1820-6.
27. Li C, Wang H, Yuan T, Woodman A, Yang D, Zhou G, Cameron CE, Yu L. 2018. Foot-and-mouth disease virus type O specific mutations determine RNA-dependent RNA polymerase fidelity and virus attenuation. Virology 518:87-94.
28. Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, Godelle B. 1997. Role of mutator alleles in adaptive evolution. Nature 387:700-2.
29. Travis JM, Travis ER. 2002. Mutator dynamics in fluctuating environments. Proc Biol Sci 269:591-7.
30. Sniegowski PD, Gerrish PJ, Lenski RE. 1997. Evolution of high mutation rates in experimental populations of E. coli. Nature 387:703-5.
31. Bedford T, Cobey S, Pascual M. 2011. Strength and tempo of selection revealed in viral gene genealogies. BMC Evol Biol 11:220.
32. Kimura M. 1967. On the evolutionary adjustment of spontaneous mutation rates. Genet Res 9:23-34.
33. Kamp C, Wilke CO, Adami C, Bornholdt S. 2002. Viral evolution under the pressure of an adaptive immune system: optimal mutation rates for viral escape. Complexity 8:28-33.

Journal of Virology

34. Sniegowski PD, Gerrish PJ, Johnson T, Shaver A. 2000. The evolution of mutation rates: separating causes from consequences. Bioessays 22:1057-66.
35. Tanaka MM, Bergstrom CT, Levin BR. 2003. The evolution of mutator genes in bacterial populations: the roles of environmental change and timing. Genetics 164:843-54.
36. Elena SF, Sanjuan R. 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. J Virol 79:11555-8.
37. Furio V, Moya A, Sanjuan R. 2005. The cost of replication fidelity in an RNA virus. Proc Natl Acad Sci U S A 102:10233-7.
38. Regoes RR, Hamblin S, Tanaka MM. 2013. Viral mutation rates: modelling the roles of within-host viral dynamics and the trade-off between replication fidelity and speed. Proc Biol Sci 280:20122047.
39. Hopfield JJ. 1974. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proc Natl Acad Sci U S A 71:4135-9.
40. Gerrish PJ, Colato A, Perelson AS, Sniegowski PD. 2007. Complete genetic linkage can subvert natural selection. Proceedings of the National Academy of Sciences of the United States of America 104:6266-6271.
41. Tenaillon O, Toupance B, Le Nagard H, Taddei F, Godelle B. 1999. Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. Genetics 152:485-93.
42. Risso-Ballester J, Cuevas JM, Sanjuan R. 2016. Genome-Wide Estimation of the Spontaneous Mutation Rate of Human Adenovirus 5 by High-Fidelity Deep Sequencing. PLoS Pathog 12:e1006013.
43. Geller R, Domingo-Calap P, Cuevas JM, Rossolillo P, Negroni M, Sanjuan R. 2015. The external domains of the HIV-1 envelope are a mutational cold spot. Nat Commun 6:8571.
44. Loh E, Salk JJ, Loeb LA. 2010. Optimization of DNA polymerase mutation rates during bacterial evolution. Proc Natl Acad Sci U S A 107:1154-9.
45. Clune J, Misevic D, Ofria C, Lenski RE, Elena SF, Sanjuan R. 2008. Natural selection fails to optimize mutation rates for long-term adaptation on rugged fitness landscapes. PLoS Comput Biol 4:e1000187.

Figure Legends

Figure 1. Viral mutation rates. A) Viral evolutionary rates (substitutions per nucleotide site per year, s/n/y) increase with mutation rate (substitutions per nucleotide site per cell infection, s/n/c), up to a point (adapted from 3). B) Evolutionary rates against mutation rate for individual viruses (adapted from 4). C) Mutation rate (substitutions per nucleotide site per generation, s/n/g) against genome size (base pairs, bp) for viruses and other organisms (data from 5, 6). For complete data used to generate figures, visit github.com/lauringlab/JVI_Gem_2018.

422 **Figure 2.** Three hypotheses for why mutation rates have not evolved to be zero. Left panel: The
423 drift-barrier hypothesis posits that drift, which weakens as the effective population size grows
424 larger, prevents mutation rates from reaching zero (6). Center panel: The physicochemical limit
425 hypothesis posits that the cost of perfect polymerase function pushes the mutation rate away
426 from zero. Right panel: The selection hypothesis posits that selection *for* adaptability and/or
427 replicative speed drives mutation rates higher. Figures are approximate trends and are not
428 meant to indicate exact relationships (e.g. linear).
429
430 **Tables**

Method	Advantages	Disadvantages
Sanger sequencing (mutation frequency)	• Mutational spectra • Context-dependent effects	• Biased against lethal and deleterious mutations • Sequencing errors • Low power for comparison • RT-PCR errors (for RNA viruses)
Deep sequencing (mutation frequency)	• Mutational spectra • Context-dependent effects • Easier to exclude sequencing errors	• Biased against lethal and deleterious mutations • RT-PCR errors (for RNA viruses)
Mutation accumulation (mutation frequency)	• Less biased against deleterious mutations • Mutational spectra • Context-dependent effects	• Biased against lethal mutations • Requires extensive passaging to accumulate large sample size • Decline in population fitness (for RNA viruses)
Fluctuation tests	• No RT-PCR errors	• Requires marker with a scorable

(mutation rate)	• No sequencing errors • Less biased against lethal and deleterious mutations	phenotype • Limited context-dependent effects • Limited mutational spectra
Cell-free assays (misincorporation rate)	• Less biased against lethal and deleterious mutations • Can capture polymerase kinetics	• Requires purification of polymerase • Often differ from estimates of mutation rate

Table 1. Methods for estimating viral mutation rates, including mutation frequency, mutation rate, and misincorporation rate, along with the advantages and disadvantages for each approach. RT = reverse transcription.

Figure 1

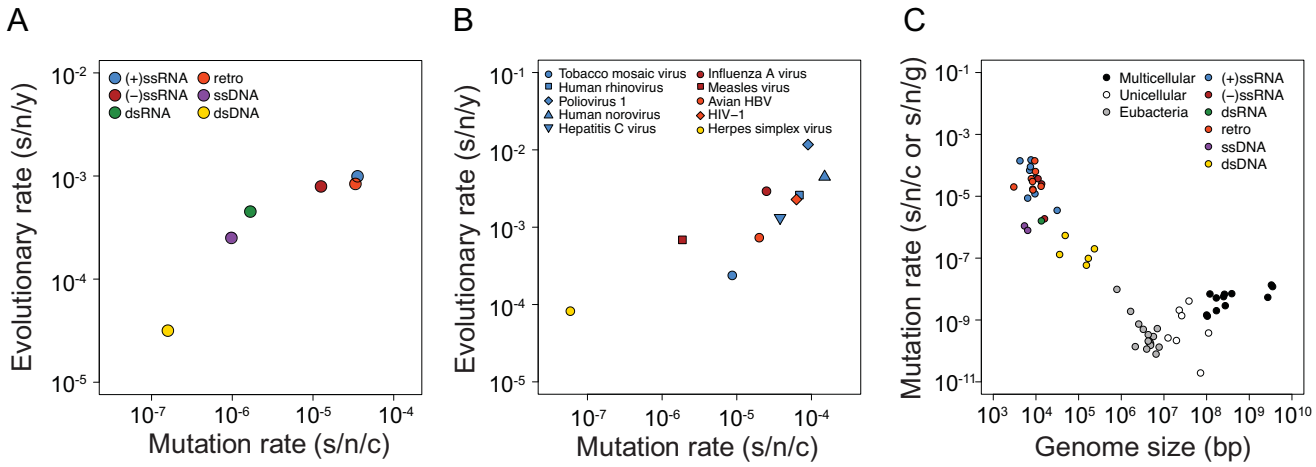

Figure 2