Title

Codon Usage Determines the Mutational Robustness, Evolutionary Capacity, and Virulence of an RNA Virus

Journal

Cell Host & Microbe, 12(5)

ISSN

1931-3128

Authors

Lauring, Adam S
Acevedo, Ashley
Cooper, Samantha B
et al.

Publication Date

2012-11-01

DOI

10.1016/j.chom.2012.10.008

Peer reviewed

Published in final edited form as:
Cell Host Microbe. 2012 November 15; 12(5): 623–632. doi:10.1016/j.chom.2012.10.008.

Codon usage determines the mutational robustness, evolutionary capacity and virulence of an RNA virus

Adam S. Lauring[1,3,†], **Ashley Acevedo**[2,3], **Samantha B. Cooper**[4], and **Raul Andino**[3,*]

[1]Department of Medicine, University of California, San Francisco, CA, San Francisco, CA 94143

[2]Tetrad Graduate Program, University of California, San Francisco, CA, San Francisco, CA 94143

[3]Department of Microbiology and Immunology, University of California, San Francisco, CA, San Francisco, CA 94143

[4]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, San Francisco, CA 94143

Summary

RNA viruses exist as dynamic and diverse populations shaped by constant mutation and selection. Yet little is known about how the mutant spectrum contributes to virus evolvability and pathogenesis. Because several codon choices are available for a given amino acid, a central question concerns whether viral sequences have evolved to optimize not only the protein coding consensus, but also the DNA/RNA sequences accessible through mutation. Here we directly test this hypothesis by comparing wild type poliovirus to synthetic viruses carrying reengineered capsid sequences with hundreds of synonymous mutations. Strikingly, such rewiring of the population's mutant network reduced its robustness and attenuated the virus in an animal model of infection. We conclude that the position of a virus in sequence space defines its mutant spectrum, evolutionary trajectory, and pathogenicity. This organizing principle for RNA virus populations confers tolerance to mutations and facilitates replication and spread within the dynamic host environment.

Introduction

The genetic architecture of RNA virus populations is often depicted as a network of variants organized in sequence space around a single master sequence (Domingo et al., 1978; Eigen, 1993). The structure and composition of these mutational neighborhoods is determined by the selective landscape and may have a significant impact on viral fitness, evolvability, and virulence (Burch and Chao, 2000). Although elevated mutation rates can be adaptive, most mutations have deleterious effects on viral fitness, and RNA virus populations exist perilously close to a theoretical error threshold of viability (Holmes, 2003; Sanjuán et al.,

*Corresponding Author 600 16th Street, GH-S572 UCSF Box 2280 San Francisco, CA 94143-2280 (415) 502-6358 phone (415) 514-4112 fax raul.andino@ucsf.edu.
†Present Address Department of Medicine, Division of Infectious Diseases Department of Microbiology & Immunology University of Michigan Medical School 5510B MSRB I, SPC 5680 1150 W. Medical Center Dr. Ann Arbor, MI 48109-5680 (734) 764-7731 alauring@med.umich.edu

2004). Population genetic models suggest that error-prone replication will favor the evolution of mutational robustness, whereby populations buffer the negative effects of mutation by migrating to regions of sequence space corresponding to more neutral fitness landscapes (de Visser et al., 2003; Montville et al., 2005; Codoñer et al., 2006; Sanjuán et al., 2007). In turn, robustness may facilitate evolvability by increasing the number of adaptive pathways available within a given landscape (Draghi et al., 2010). Therefore, the position of a genome within sequence space may be optimized to determine the network of neighboring sequences rapidly accessible through mutation (Figure 1).

In experimental virology, the concept of the master sequence is best approximated by the consensus, which represents the average sequence of the population. Studies of mutational robustness and evolvability in viral systems have generally relied on demonstrating differences in replicative speed and robustness in populations derived from biological clones with divergent consensus sequences (Montville et al., 2005; Codoñer et al., 2006; Sanjuán et al., 2007). While each population occupies a distinct region of genotypic sequence space, uncontrolled changes in the amino acid consensus sequence confound analysis of the impact of fitness landscapes on the mutant spectrum and its association with viral phenotype. Therefore, the importance of robustness and local sequence space to wild type populations replicating at their natural mutation rates remains unclear.

In evolutionary models, robustness is closely linked to the degenerate relationship between traits and genes (Huynen et al., 1996; Meyers et al., 2005). That is, multiple genotypes can give rise to the same phenotype. This phenomenon may be linked to the genetic code, where nearly all amino acids are encoded by more than one synonymous codon. While synonymous mutation is often thought to be selectively neutral, the observed variation in codon usage across both viral and organismal taxa suggests the presence of mutational bias and/or selective pressure (Jenkins and Holmes, 2003; Plotkin and Kudla, 2011). Translational efficiency has been invoked as a selective pressure to explain the evolution of codon bias, codon pair bias, and codon order (Coleman et al., 2008; Cannarozzi et al., 2010; Tuller et al., 2010). In RNA viruses, constraints on RNA structures necessary for replication and packaging have been recognized as another important reason for codon bias (Simmonds and Smith, 1999; Goodfellow et al., 2000).

However, an unexplored aspect of the genetic architecture of RNA viruses is how codon choice influences population diversity. Given the structure of the genetic code, different synonymous codons will yield distinct amino acid changes upon mutation. This may be particularly important for RNA viruses, which replicate with extremely high mutation rates. Thus, the structure of the virus consensus sequence may significantly influence the expressed diversity and connectivity of mutant networks, and ultimately evolutionary trajectory. Theoretical work suggests that the organization of mutant networks in genetic sequence space can profoundly influence the speed of adaptation (Draghi et al., 2010). This model has been validated by experimental studies of small RNA, where robust populations are more evolvable by virtue of accumulated cryptic variation (Hayden et al., 2011). While both RNA virus and ribozyme populations can exhibit significant genetic diversity, the importance of cryptic variation and mutant networks to RNA virus evolution and pathogenesis is unclear.

Here we used large-scale synonymous mutation to relocate viral populations within genetic sequence space. This allowed us to dissociate the properties of the mutant spectrum from those of the coding consensus. We find that these synonymous populations exhibit unique mutant spectra with significantly different distributions of polymorphic amino acid substitutions in the capsid. On a phenotypic level, they have similar replicative capacity, but differ in clonal and population fitness. These three synonymous populations reside in

distinct fitness landscapes, and one of the engineered populations is less mutationally robust than the others. Strikingly, this population is also attenuated in an animal model of infection, demonstrating the importance of mutant neighborhoods in determining viral pathogenesis. Our results suggest that mutational robustness is inherent to a virus' natural position in sequence space and facilitates evolvability in the dynamic host environment.

Results

We compared wild type (WT) poliovirus to two variants with large-scale synonymous substitution in the viral capsid. These previously described variants, Max and SD, were derived by *de novo* gene synthesis and contain 566 and 934 substitutions, respectively, within the 2643 nucleotide sequence coding for the viral capsid protein (Mueller et al., 2006; Coleman et al., 2008). Because they were designed as negative controls for studies of codon bias and virus translation, the GC content and codon usage frequency across all 64 codons are nearly identical to that of the wild type virus. While synonymous mutation could disrupt RNA structural elements within the capsid sequence, Max and SD were specifically designed to have similar RNA folding free energy across the capsid sequence relative to the wild type. Furthermore, no such structures have been identified in poliovirus, and deletion or replacement of the capsid sequence has no effect on genome replication, translation, or genome encapsidation (Kaplan and Racaniello, 1988; Johansen, 2000). Most importantly, Max and SD have a consensus amino acid sequence identical to wild type but occupy very distinct positions within sequence space (Figure 1).

Codon usage modulates mutational robustness of the virus population

Low fidelity replication ensures that the majority of newly synthesized genomes will differ from the consensus sequence. Because WT, Max, and SD occupy very different positions in sequence space, we hypothesized that each population would contain a unique set of variants with distinct phenotypes. Plaque morphology provides a useful read-out of the phenotype of individual members of the population, and we thus measured the area of individual plaques to quantify the fitness distribution of clones derived from each population (Burch and Chao, 2000; Sanjuán et al., 2007). The mean plaque size of the Max population was similar to wild type (Figure 2a, mean log plaque area: Max =2.836, WT=2.835), although the size distribution was slightly different (KS test for distribution, $D=0.065$, $p=0.021$). In contrast, clones from the SD population produced on average smaller plaques than wild type (mean log plaque area: SD 2.515, WT 2.835) and had a wider size distribution (KS test for distribution, $D=0.371$, $p<0.001$). We obtained nearly identical results with samples derived from independent viral stocks (data not shown). These data indicate that populations derived from synonymous genomic sequences can give rise to variants with different fitness values.

The wider fitness variance exhibited by SD suggests that it inhabits a region of sequence space with a lower density of neutral mutations. We predicted that SD's position in sequence space would reduce its robustness because the effects of mutations in the SD background are more deleterious for each individual virus in the population, resulting in an overall reduction in population fitness at lower multiplicities of infection. Notably, we observed that SD typically produced populations of lower infectious titer, suggesting that its mutant spectrum harbors a greater proportion of replication-defective variants (data not shown). We thus examined the link between sequence space and robustness by exploring the contours of the fitness landscape as a population is driven further out into sequence space (de Visser et al., 2003). To this end, we quantified the sensitivity of wild type, Max and SD to ribavirin, a nucleoside analogue and RNA virus mutagen (Crotty et al., 2000; 2001). As previously reported, we found a 2–3 log reduction in wild type titer at higher ribavirin concentrations compared to the drug-resistant mutant, $3D^{G64S}$ (Vignuzzi et al., 2006; 2008). While Max displayed a similar sensitivity to mutagen relative to wild type, SD was hypersensitive

(Figure 2b, p= 0.016 and 0.006 for 400μM and 1000μM ribavirin, respectively). We also examined the impact of ribavirin mutagenesis on clonal fitness by comparing the plaque size of viable progeny following drug treatment (Figure 2c). Consistent with the mutagenic effect of ribavirin, we found a significant reduction in plaque size for WT, Max, and SD ($p<0.001$ by either ks-test or t-test for mock vs. ribavirin treated population). In addition, for SD we found that ribavirin also reduced the total number of plaques, indicating that many individual viruses are no longer viable when the mutational load is artificially increased by ribavirin. Importantly, this experiment assessed fitness in the absence of ribavirin; thus the effect of the drug on plaque size cannot be attributed to potential pleiotropic effects of the drug. Rather, these data indicate that SD is less mutationally robust than either Max or wild type, consistent with a model in which the position of a virus population within sequence space determines its tolerance of mutation.

Mutational robustness reduces population fitness without affecting overall replication or translational efficiency of the consensus sequence

We next examined whether fitness differences among the spectrum of mutants generated by each poliovirus synonym result in fitness differences among the three populations. In direct competition assays (Figure 3a), the Max population was somewhat fitter than the wild type (w_{rel} 1.74 ± 0.17, n=5 replicates) and SD was markedly less fit (w_{rel} 0.18 ± 0.02, n=5 replicates). We also observed these fitness differences at a ten-fold higher multiplicity of infection (moi ~ 10), where frequent co infection results in complementation of defective genomes and proteins. Here, Max exhibited a fitness closer to the wild type reference (1.054 ± 0.004, n=5 replicates), while SD remained considerably less fit (0.712 ± 0.019, n=5 replicates). These results are consistent with the idea that the less robust SD population is less fit because of the impact of mutation on the variants generated during each replication cycle. An alternative explanation for the observed fitness differences, however, would be a direct effect of specific synonymous mutations on viral replication or translation. We therefore measured the kinetics of genome replication in a single replication cycle, using a sensitive and precise quantitative polymerase chain reaction assay (Figure 3b). We found no statistically significant difference in the exponential growth rate between WT and either SD or Max (p=0.1837 and p=0.5178, respectively). Despite their ten-fold difference in relative fitness, Max and SD had nearly identical growth curves, with only small differences evident at the plateau phase of replication (Figure S1). In contrast, we were able to measure a replication defect for a poliovirus mutant with a much more modest fitness difference. Under identical assay conditions, this guanidine resistant poliovirus ($2C^{N179G}$) exhibits a fitness of 0.56 and a significant lag in replication relative to wild type (Figure S1, Pincus et al., 1986; Lauring and Andino, 2011). To exclude a subtle translation defect in SD, we directly assayed viral capsid translation in infected and metabolically labeled cells (Figure 3c). Again we found no difference to suggest a replication or translation defect of the viral master sequence that could account for the fitness differences among WT, Max, and SD (Figure 3c and 3d). Rather our data demonstrate that SD is located in a region of sequence space with a lower density of neutral mutations, where the mutant neighborhood surrounding the consensus exhibits reduced fitness. That is, SD produces genomes and proteins with the same kinetics as wild type, but its progeny are less fit.

We further explored the relationship between robustness and fitness by performing competition assays in the presence of ribavirin. Rather than comparing synonymous populations to each other, we competed each virus (WT, Max, or SD) against an isogenic, drug resistant reference containing a single point mutation, $3D^{G64S}$ that increases the fidelity of the viral polymerase, and therefore reduces quasispecies diversity and prevents error-catastrophe. In this assay, the test and reference strain differ only in their replicative fidelity and sensitivity to ribavirin. This experimental set-up allowed us to determine how codon

reorganization influences fitness and robustness, independent of its potential pleiotropic effects on genome replication or translation. We found that each population behaved differently over six, single-cycle, low-multiplicity passages. As above, the SD populations were hypersensitive to ribavirin relative to Max and WT. While we could easily detect WT and Max after five passages in 200μM, the SD populations were extinguished after just 3 passages. As a result, we could only obtain accurate fitness measurements for the 0μM and 100μM concentrations (see Figure 3e). Applying ANOVA to this data set, we found that fitness measurements varied significantly as a function of ribavirin ($p<0.001$), virus ($p=0.0016$), and both factors together ($p<0.001$). Fitting the data to a linear model, we found the fitness of SD varied significantly with ribavirin relative to WT and Max (SD vs. Max $p<0.001$; SD vs. WT $p=0.0051$; Max vs. WT $p=0.0177$). These data support our conclusions that SD occupies a less favorable region of sequence space with consequent impacts on viral fitness.

Synonymous populations exhibit distinct mutant spectra

Next, we used ultra-deep sequencing to define the mutant spectrum of each virus and explore the genotypic basis for the observed differences in fitness and robustness. Sequence libraries were derived from purified virions; thus, the vast majority of our reads mapped to the reference genome (Table S1). We applied a neighborhood quality standard (NQS) to reduce aberrant base calls, and obtained an average coverage across the viral genome of 20,000–56,000 reads per position. As expected, single nucleotide variants (SNV) were extremely rare (frequency 10^{-5} to 10^{-3}, Figure S2). Importantly, the frequency and location of SNV in the three virus populations indicated that synonymous populations acquire distinct mutation distributions.

An independent sequence analysis was carried out by cloning 400–500 individual viruses by endpoint dilution and sequencing pooled libraries (Figure 4a). This approach circumvented the current limitations of deep-sequencing methods given the error rate of the process (Figures S2, S3). Since cloning fixes 1–2 mutations per genome on average (~600 per pool), we were able to use a frequency-based cut-off 0.001 to identify SNV with a higher probability of representing authentic non-lethal variants (Figure S3). The WT, Max, and SD populations had a similar distribution of synonymous and nonsynonymous SNV across the capsid and genome as a whole, which is expected given their identical codon usage (Figure 4b). We observed a significant overlap in variant positions among WT, Max, and SD within the non-capsid portion of the genome (Figure 4c), consistent with their identical nucleotide consensus in the nonstructural proteins and noncoding regions. Importantly, we found much less overlap in the capsid region, where the three synonymous viruses have widely divergent nucleotide sequence (Chi squared goodness of fit, $\chi^2=183$, $df=3$, $p<0.001$). We used a conservation-based scoring system to characterize the types of amino acid substitutions present in each pool of clones (Henikoff and Henikoff, 1992). Again, we found striking differences in the location of capsid mutations identified in the synonymous populations, particularly within neutralizing antibody epitopes and other highly variable capsid domains (Figure 5, Hogle and Filman, 1989). Our data strongly suggest that the distinct region of sequence space inhabited by wild type, Max, and SD populations gives rise to a distinct spectrum of mutants, accounting for their observed phenotypic differences.

Location in sequence space, mutational robustness and virulence

Given that location within nucleotide sequence space determines the mutation distribution of the viral population (Figures 4 and 5), we examined if it also affects the course of viral infection in susceptible animals (Ren et al., 1990; Crotty et al., 2002). Mice infected with SD by the intramuscular route exhibited delayed and decreased mortality compared to mice inoculated with equivalent doses of either wild type or Max ($p<0.001$ by log rank sum test,

Figure 6a). We also quantified virus from the brains of mice infected intravenously. This route of infection presents additional bottlenecks and selective pressures to the virus population (Pfeiffer and Kirkegaard, 2006; Vignuzzi et al., 2006). Similar to the intramuscular route, we found that SD was attenuated *in vivo* compared to both Max and wild type (Figure 6b). The striking differences in neurovirulence phenotype between Max and SD strongly suggest that robustness is a determinant of virus pathogenicity. To determine whether the observed differences were due to specific variants selected *in vivo*, we sequenced ~500 clones from the brains of mice infected with each virus (~50 clones from each of 10 mice). If the same pathogenic mutants were repeatedly selected in independent mice, one would expect to find them in a majority of brain-derived clones. For example, a mutation present in all clones from a single mouse would be present at ~10% in the pool, while a mutation found in five mice would have a frequency of ~50%. Our data show that this is not the case, as the consensus sequence of the infecting population and brain-derived pool were identical; furthermore, no dominant neurotropic clones were found in mice infected with wild type, Max, or SD (Figure 6c). These data confirm the importance of population diversity to virulence (Pfeiffer and Kirkegaard, 2005; Vignuzzi et al., 2006), and suggest that mutational robustness plays a key role in determining quasispecies diversity, and therefore pathogenicity.

Discussion

Our work links mutational robustness, mutant spectrum, and phenotype in RNA viruses. While the consensus sequence forms the genetic backbone of a population, error-prone replication ensures that RNA virus quasispecies will be inherently diverse. The use of synonymous mutations to place the poliovirus master sequence in different locations within sequence space, without affecting the amino acid sequence, produced populations with distinct mutant spectra while avoiding the confounding effect of divergent consensus sequences. In a single replication cycle, at high multiplicity of infection, all three synonymous viruses efficiently replicate with wild type kinetics and express near identical amounts of capsid protein. The observed differences in fitness and virulence among these three synonymous viruses underlie the importance of the mutant distribution to the behavior of RNA virus populations. Our sequencing data indicate that many of the single nucleotide variants in poliovirus populations are extremely rare yet profoundly affect virus phenotype. Thus, an astonishing amount of low-level genotypic and phenotypic variation lies beneath the consensus in viral populations and drives adaptability and pathogenesis.

Is it possible that small pleiotropic effects of synonymous mutation on local or global RNA structures are responsible for the phenotypes describe here? Several lines of evidence make this possibility an unlikely confounder. First, the poliovirus capsid sequence is devoid of functional RNA elements and is not required for either replication or packaging (Kaplan and Racaniello, 1988; Johansen, 2000). Second, the poliovirus capsid tolerates a high frequency of heterologous recombination without evident disruption of relevant structures (Liu et al., 2003; Yang et al., 2005). Third, Max and SD exhibit similar folding free energy across the capsid sequence relative to wild type (see Supplemental Data in Coleman et al., 2008). Fourth, we find that Max and SD replicate with wild type kinetics and express similar amounts of protein. Fifth, it is difficult to define an RNA structural basis for the fitness differences between Max and SD as both have hundreds of potentially disruptive mutations. Finally, our mutagen sensitivity assay and sequencing data offer alterations in mutation distribution and robustness as a clear alternative explanation for the observed phenotypes.

Synonymous codons differ in their propensity to mutate both nonsynomyously and nonconservatively, and directed evolution experiments in bacteria suggest that this differential access to protein sequence space influences adaptive pathways (Cambray and

Mazel, 2008). It is interesting that SD gave rise to a less robust population than either wild type or Max. These three viruses have near identical codon usage across both the capsid sequence and the entire genome. Therefore, on a genomic level, they will access the same nonsynonymous substitutions. While the synonymous substitutions in Max were based on altering local codon pair bias, the codons in the SD capsid are randomly shuffled compared to wild type (Mueller et al., 2006; Coleman et al., 2008). Recent work has highlighted the potential importance of codon volatility in rapidly evolving protein domains (Meyers et al., 2005), and studies of several RNA viruses indicate that the highly mutable surface epitopes of viral proteins may be enriched for codons that mutate nonsynonymously (Stephens and Waelbroeck, 1999; Plotkin and Dushoff, 2003). In this model, the genomic reorganization of SD would disrupt patterns of local evolvability essential to inter- and intra-host survival by poliovirus. Our data suggest that selection for mutational robustness at high mutation rates may play an important role in viral genome evolution. This effect may work either independently or in concert with other non-coding constraints such as codon bias, codon pair bias, or RNA structure (Wimmer and Paul, 2011).

The relationship between mutational robustness and evolvability has been a subject of intense study among evolutionary theorists (de Visser et al., 2003; Wagner, 2005; Draghi et al., 2010). In principle, robustness could slow the adaptive process by limiting the phenotypic expression of genetic variation. However, recent theoretical work suggests that cryptic genetic variation can profoundly impact the phenotypic neighborhood through epistatic effects with subsequent mutations (Draghi et al., 2010). Experimental data supporting this model derives largely from studies of small RNA where the impact of mutation on structural phenotype can be simulated (Fontana and Schuster, 1998; Wagner and Stadler, 1999; Wagner, 2008). Because these RNA populations exist on selectively neutral networks, they accumulate significant genotypic diversity and access greater phenotypic diversity in relatively few mutational steps. A recent study of a mutationally robust ribozyme further indicates that a population harboring cryptic genetic variation can be “pre-adapted” to a new environment, and therefore, more evolvable (Hayden et al., 2011).

Our data extend this model beyond RNA structures to a more complex, viral system and demonstrate its importance in the context of a host-pathogen interaction. Low replicative fidelity and population diversity are closely linked to fitness and virulence in RNA viruses (Pfeiffer and Kirkegaard, 2005; Vignuzzi et al., 2006). A population with extensive genetic diversity will contain, or be able to quickly acquire, the necessary adaptive mutations to survive within the dynamic host environment. In this model, a delicate balance between replication fidelity and genomic flexibility creates a reservoir of individual variants within the population that facilitates adaptation to changing selective conditions. Here, we have demonstrated a relationship between rare, cryptic variation and viral pathogenesis. We speculate that a successful population will harbor individual variants that are pre-adapted to a broad range of selective pressures, ensuring viral survival and spread. This paradigm applies to both mammalian hosts, with their the diverse spatial and immune environments, and cellular monocultures, where changes in gene expression, membrane structures, and innate immune responses generate shifting fitness landscapes (for theoretical discussion of robustness and spatial environments, see Sardanyés et al., 2008). Given that we have sampled just two points in the vast expanse of sequence space, it is interesting that SD was attenuated, and Max was phenotypically closer to the wild type. Our assays in HeLa and transgenic mice only measure a subset of the selective pressures that poliovirus will encounter in nature, and perhaps the three poliovirus populations differ in other important ways. By analogy, many mutations will alter the structure of a protein, but only a subset will affect its enzymatic activity in a particular set of assay conditions. We have shown that

cryptic genetic variation is important to pathogenesis, and future work should clarify the diverse phenotypic consequences of these shifts in sequence space.

The evolution of RNA virus populations is often discussed in the context of quasispecies theory, which explores the consequences of low replicative fidelity for molecular evolution (Eigen, 1971). In quasispecies models, high mutation rates will favor the evolution of mutational robustness. Selection will therefore maximize the mean fitness of the interconnected mutant network rather than the replicative fitness of any individual variant. This phenomenon, “survival of the flattest,” has been demonstrated in digital organisms, plant viroids, and a mammalian virus (Wilke et al., 2001; Codoñer et al., 2006; Sanjuán et al., 2007). However, the distinction between the fittest and flattest may be an artificial one since experimental fitness assays measure replicative and translational rate as well as robustness to mutation (Wilke et al., 2001). Indeed, we found that robustness was a significant determinant of fitness relative to replicative and translational efficiency. Furthermore, the less robust SD population was attenuated relative to the more robust wild type and Max ones. The role of robustness in balancing the positive and negative effects of mutation is also underscored by our observation of this effect at basal poliovirus mutation rates. This finding addresses a key shortcoming of prior studies, in which the superiority of “flat” populations could only be shown at supranormal mutation rates (see Holmes, 2010)). Importantly, our definition of a less robust alternative consensus (SD) strongly suggests that mutational robustness is adaptive and may play a key role in shaping the neutral networks and phenotypic neighborhoods explored by RNA virus populations.

The theoretical and experimental framework described in this study could be used to design strategies to restrict the evolutionary capacity of live-attenuated viral vaccine strain candidates. The traditional process of attenuation through serial passage operates by forcing a virus to adapt to a foreign host environment. The accumulated mutations move the population to a new region of sequence space that is unable to support efficient replication and spread within the vaccinee. Because this process is highly empiric and somewhat unpredictable, there is a clear need for new approaches (reviewed in Lauring et al., 2010). We have previously shown that relatively small changes in the error rate of an RNA virus can significantly reduce population diversity, resulting in profound attenuation (Vignuzzi et al., 2006; 2008). Here we find that manipulation of the mutant spectrum through synonymous mutation can achieve the same goal. In our model, the SD population is trapped in a region of sequence space with a low density of neutral mutations. The ruggedness of the associated fitness landscape reduces its capacity to generate fit progeny and adapt to host selective pressures. As others have shown, synonymous mutation has the additional advantage of preserving the antigenic composition of the wild type and reducing the risk of reversion through recombination with circulating strains (Burns et al., 2006; Mueller et al., 2006; Coleman et al., 2008; Burns et al., 2009). We speculate that further definition of a virus' mutant spectrum and native fitness landscape may allow finer control of a the populations' evolutionary capacity, limiting its ability to escape immune surveillance and spread within a vaccinated individual and through the human population.

Experimental Procedures

Plasmids, viruses and cells

HeLa S3 cells were maintained in 50% Dulbecco's modified Eagle media (DMEM) / 50% F-12 media supplemented with 10% newborn calf serum and 2 mM glutamine. For measurements of tissue culture infectious dose ($TCID_{50}$), cells were maintained in media supplemented with 2% serum. Plasmids containing poliovirus cDNA for type 1 Mahoney, G64S, SD, and Max have been described previously. Viral RNA was generated by *in vitro* transcription of the corresponding plasmid, and 20 μg of each RNA was electroporated into

4×10^6 HeLa S3 cells. The resulting passage 0 (P0) stock was passed three times on HeLa S3 in serum free media at an MOI of ~1 to generate a P3 stock used for subsequent experiments. Cell free viral supernatants were recovered at each passage after three successive freeze-thaw cycles and low speed centrifugation at 1000 rpm in a benchtop clinical centrifuge. All viruses were titered by $TCID_{50}$.

Real time PCR analysis of viral replication kinetics

Subconfluent monolayers (~ 5×10^5 cells) in 12 well dishes were infected with virus in 150 μl serum-free media at a multiplicity of 10. Thirty-five replicate wells were used for each virus (7 time points, 5 replicate samples per time point). After one hour, the virus was removed, all wells were washed twice with 1 ml PBS and given fresh media. We used this as our 0 hour time point. RNA was harvested for each virus at each time point (0–6 hours), and random hexamers were used to prime cDNA synthesis on 1/6 of each RNA preparation (5 μl) in a 10 μl volume. Complementary DNA were diluted 50-fold and 5 μl of this dilution was analyzed by real time PCR. Each 50 μl PCR reaction contained 1.5 mM $MgCl_2$, 150 μM each dNTP, 0.025 U/μl Taq polymerase, 0.2× SYBR green, 0.4 μl Rox reference dye, and 0.6 μM of primers COM2F (5'-CATGGCAGCCCCGGAACAGG-3') and COM 2R (5'-TGTGATGGATCCGGGGGTAGCG-3'). Reactions were subjected to 40 cycles of 95°C / 30 seconds, 58°C / 30 seconds, 72°C / 30 seconds. A standard curve, correlating cycle threshold and approximate RNA copy number, was generated using cDNA synthesized from 10-fold serial dilutions of *in vitro* transcribed poliovirus RNA.

Analysis of radiolabeled viral proteins in infected HeLa cells

Subconfluent monolayers (~ 2×10^6 cells) in 6 well dishes were infected with virus in 350 μl serum-free media at a multiplicity of 40. Residual virus was removed at 5 hours. The cells were washed twice with 1 ml PBS and incubated for two hours in L-methionine- and L-cysteine-deficient DMEM containing 100 μCi/ml ^{35}S L-methionine / L-cysteine. Cells were lysed 7 hours post infection in 1% NP40, 10 mM Tris HCl pH 7.5, 10 mM NaCl, 1.5 mM $MgCl_2$ and spun at 2000 × g for 5 minutes at 4°C. One quarter of each lysate was resolved by SDS-PAGE. Immunoprecipation of capsid proteins was performed using a polyclonal capsid antisera and Protein G magnetic beads. The 12% gel was fixed in 30% methanol / 10% acetic acid. Autoradiographs were obtained using a Typhoon Phosphorimager, and quantified using ImageQuant software.

Competition assays

Cells were infected at a total MOI of 1 with an equal $TCID_{50}$ of wild type and either Max or SD. Five replicate wells were infected with each pair of viruses. Passage 1 virus was harvested after complete cytopathic effect (~24 hours), diluted 1000-fold and applied to fresh cells. For high multiplicity and $3D^{G64S}$-ribavirin competition experiments, each supernatant was diluted 100-fold or 400-fold respectively. This process was repeated for a total of 7 passages. RNA was harvested from each passage using PureLink 96 well RNA mini kits. Random hexamers were used to prime cDNA synthesis with 1/10 of the RNA. Complementary DNA were diluted 100-fold and analyzed by real time PCR as above. Each cDNA was analyzed using three different primer sets with duplicate PCR reactions for each sample/primer set. The first set, COM2F and COM2R, was used to quantify the total amount of viral genomic RNA. The second set, M2F (5'-TACGGGACAGCTCCAGCCCG-3') and M2R (5'-GCCGTAGTACGCCACTGCCC-3'), was used to quantify the total amount of wild type RNA. The third set, either Max4F (5'-TGTGCGGGAGACTGTGGGGG-3') and Max4R (5'-TGTTGGTGGTGGAGGCTGGG-3') or SD4F (5'-CAGCGGCCCTGGGAGATTCG-3') and SD4R (5'-CGCGTGGAGGTCTTGGACACC-3'), was used to quantify the amount of Max or SD RNA, respectively. We used custom TaqMan probes (Applied Biosystems) to distinguish the two variants in $3D^{G64S}$-ribavirin competition

experiments (e.g. WT and WT-3D^{G64S}). Duplicate wells were averaged and relative amounts of WT, Max, and SD RNA were determined by normalizing the cycle thresholds for each to those of the COM primer set ($\Delta Ct=Ct_{Virus}-Ct_{COM}$). The normalized values for each virus passages 2–7 were then compared to passage 1 to obtain a ratio relative to P1 ($\Delta\Delta Ct=Ct_{PX}-Ct_{P1}$). This relative Ct value was converted to reflect the fold change in the ratio ($\Delta ratio = 2^{-\log\Delta\Delta Ct}$). The change in ratio of the mutant relative to the change in ratio of the wild type as a function of passage is the fitness ($[\Delta ratio_{Mut}-\Delta ratio_{WT}]$/time).

Determination of plaque size variability

Plaque assays were performed on subconfluent monolayers (~ 10^7) in 10 cm dishes. The amount of virus was determined empirically to ensure well spaced out plaques (~30–50 per 10 cm dish). Each plate was scanned individually at 300 dpi using a flat-bed scanner. Sixteen bit image files were analyzed using ImageJ. Brightness, contrast, and circularity thresholds for plaque identification were set using uninfected plates. We controlled for the validity of the Kolmogorov-Smirnov test in detecting true differences in plaque size distributions by splitting the data for each virus and verifying that both subsets were not statistically different. In each case, we found that we were unable to reject the null hypothesis ($p>0.05$), suggesting that each came from the same distribution. We also obtained similar data, test statistics, and p values in replicate assays with different virus stocks.

Ribavirin resistance assay

Subconfluent monolayers were pretreated with 0–1000 μM ribavirin in serum free media for four hours then infected with virus at an MOI of 0.1 for 45 minutes. The cells were washed twice in phosphate buffered saline and incubated in ribavirin for an additional 24 hours. Viral supernatants were recovered as above and titered by tissue culture infectious dose.

Next generation sequencing of poliovirus populations

Sequencing was performed on viral genomic RNA purified from P3 stocks. Each population was generated by infecting $2.5 – 3.5 \times 10^7$ cells at an MOI of 0.5–1. As above, virus was harvested after complete cytopathic effect by three freeze-thaw cycles and low speed centrifugation. Pools of clones were derived by end point dilution of the corresponding P3 stock in 96 well plates. Plates were scanned at days 4–6 after infection and 150 μl from each positive well were pooled for a total of 400 wells. Pools of clones and P3 stocks were purified by ultracentrifugation, and sequence libraries were prepared following the Illumina mRNA sequencing sample preparation protocol. Details can be found in Supplementary Experimental Procedures. Brain derived viruses were sequenced using an Illumina HiSeq instrument, and all other libraries were sequenced using a Genome Analyzer II. Fifty (HiSeq) or thirty-six (GA II) base pair reads were aligned to the corresponding reference sequence (i.e. WT, Max, or SD), using Bowtie v0.12.3. All sequences were then filtered using a neighborhood quality standard (NQS) algorithm. This NQS algorithm required each base to be followed by 5 bases of phred ≥ 25. We identified single nucleotide variants and their frequencies using two different analytic pipelines (dep and alternate, see Supplementary Experimental Procedures). Downstream analysis of this SNV data is summarized in the Results section and described in detail in Supplementary Experimental Procedures.

Infection of susceptible mice

The UCSF Institutional Animal Care and Use Committee approved the protocols for the mouse studies described here. In these experiments, we used 6- to 10-week-old cPVR mice, which express the poliovirus receptor under control of the beta-actin promoter. For tissue

distribution studies, mice were infected under anesthesia with 100 μl of viral supernatant via the tail vein. Mice were monitored daily for the onset of paralysis and euthanized when death was imminent. Whole organs were isolated from all mice and homogenized in 2 ml PBS. Viral supernatants were recovered from the tissue homogenates following three freeze-thaw cycles and centrifugation at 3000 rpm for 5 minutes in a benchtop clinical centrifuge. For survival analyses, we infected 25 mice per group intramuscularly (50 μl of inoculum administered in each hindleg) with serial dilutions of virus.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Eckard Wimmer and members of his laboratory for providing plasmid constructs containing the synthetic viruses, Max and SD. We acknowledge members of the Andino laboratory for their comments and suggestions and thank Ron Geller, Oscar Westesson, Bryan Greenhouse, and Ben Schiller for assistance with statistics and sequence data analysis. We also thank Henry Bigelow, Michael Zody, and Matthew Henn, who developed the DEP software and provided it to us ahead of publication. This work was supported by grants from the NIAID to RA (R01 AI36178 and AI40085) and ASL (K08 AI081754). ASL and RA designed the project. ASL performed all of the experiments. AA assisted with the mouse infections. SC assisted with sequence analysis. ASL and RA analyzed the results and wrote the manuscript.

References

Arnold JJ, Vignuzzi M, Stone JK, Andino R, Cameron CE. Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. J. Biol. Chem. 2005; 280:25706–25716. [PubMed: 15878882]

Burch CL, Chao L. Evolvability of an RNA virus is determined by its mutational neighborhood. Nature. 2000; 406:625–628. [PubMed: 10949302]

Burns CC, Campagnoli R, Shaw J, Vincent A, Jorba J, Kew O. Genetic inactivation of poliovirus infectivity by increasing the frequencies of CpG and UpA dinucleotides within and across synonymous capsid region codons. J Virol. 2009; 83:9957–9969. [PubMed: 19605476]

Burns CC, Shaw J, Campagnoli R, Jorba J, Vincent A, Quay J, Kew O. Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region. J Virol. 2006; 80:3259–3272. [PubMed: 16537593]

Cambray G, Mazel D. Synonymous genes explore different evolutionary landscapes. PLoS Genet. 2008; 4:e1000256. [PubMed: 19008944]

Cannarozzi G, Cannarrozzi G, Schraudolph NN, Faty M, Rohr, von P, Friberg MT, Roth AC, Gonnet P, Gonnet G, Barral Y. A role for codon order in translation dynamics. Cell. 2010; 141:355–367. [PubMed: 20403329]

Codoñer FM, Darós J-A, Solé RV, Elena SF. The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. PLoS Pathog. 2006; 2:e136. [PubMed: 17196038]

Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S. Virus attenuation by genome-scale changes in codon pair bias. Science. 2008; 320:1784–1787. [PubMed: 18583614]

Crotty S, Arnold JJ, Zhong W, Lau JY, Hong Z, Andino R, Cameron CE. The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. Nat Med. 2000; 6:1375–1379. null. [PubMed: 11100123]

Crotty S, Cameron CE, Andino R. RNA virus error catastrophe: direct molecular test by using ribavirin. Proc Natl Acad Sci USA. 2001; 98:6895–6900. [PubMed: 11371613]

Crotty S, Hix L, Sigal LJ, Andino R. Poliovirus pathogenesis in a new poliovirus receptor transgenic mouse model: age-dependent paralysis and a mucosal route of infection. J Gen Virol. 2002; 83:1707–1720. [PubMed: 12075090]

de Visser JAGM, Hermisson J, Wagner GP, Ancel Meyers L, Bagheri-Chaichian H, Blanchard JL, Chao L, Cheverud JM, Elena SF, Fontana W, et al. Perspective: Evolution and detection of genetic robustness. Evolution. 2003; 57:1959–1972. [PubMed: 14575319]

Domingo E, Sabo D, Taniguchi T, Weissmann C. Nucleotide sequence heterogeneity of an RNA phage population. Cell. 1978; 13:735–744. [PubMed: 657273]

Draghi JA, Parsons TL, Wagner GP, Plotkin JB. Mutational robustness can facilitate adaptation. Nature. 2010; 463:353–355. [PubMed: 20090752]

Eigen M. Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften. 1971; 58:465–523. [PubMed: 4942363]

Eigen M. Viral quasispecies. Sci Am. 1993; 269:42–49. [PubMed: 8337597]

Fontana W, Schuster P. Continuity in evolution: on the nature of transitions. Science. 1998; 280:1451–1455. [PubMed: 9603737]

Goodfellow I, Chaudhry Y, Richardson A, Meredith J, Almond JW, Barclay W, Evans DJ. Identification of a cis-acting replication element within the poliovirus coding region. J Virol. 2000; 74:4590–4600. [PubMed: 10775595]

Hayden EJ, Ferrada E, Wagner A. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. Nature. 2011; 474:92–95. [PubMed: 21637259]

Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA. 1992; 89:10915–10919. [PubMed: 1438297]

Hogle JM, Filman DJ. The antigenic structure of poliovirus. Philos Trans R Soc Lond, B, Biol Sci. 1989; 323:467–478. [PubMed: 2569204]

Holmes EC. Error thresholds and the constraints to RNA virus evolution. Trends Microbiol. 2003; 11:543–546. [PubMed: 14659685]

Holmes EC. The RNA virus quasispecies: fact or fiction? J Mol Biol. 2010; 400:271–273. [PubMed: 20493194]

Huynen MA, Stadler PF, Fontana W. Smoothness within ruggedness: the role of neutrality in adaptation. Proc Natl Acad Sci USA. 1996; 93:397–401. [PubMed: 8552647]

Jenkins GM, Holmes EC. The extent of codon usage bias in human RNA viruses and its evolutionary origin. Virus Res. 2003; 92:1–7. [PubMed: 12606071]

Johansen L. The RNA Encompassing the Internal Ribosome Entry Site in the Poliovirus 5′ Nontranslated Region Enhances the Encapsidation of Genomic RNA. Virology. 2000; 273:391–399. [PubMed: 10915610]

Kaplan G, Racaniello VR. Construction and characterization of poliovirus subgenomic replicons. J Virol. 1988; 62:1687–1696. [PubMed: 2833619]

Lauring AS, Andino R. Exploring the fitness landscape of an RNA virus by using a universal barcode microarray. J Virol. 2011; 85:3780–3791. [PubMed: 21289109]

Lauring AS, Jones JO, Andino R. Rationalizing the development of live attenuated virus vaccines. 2010; 28:573–579.

Liu H-M, Zheng D-P, Zhang L-B, Oberste MS, Kew OM, Pallansch MA. Serial recombination during circulation of type 1 wild-vaccine recombinant polioviruses in China. J Virol. 2003; 77:10994–11005. [PubMed: 14512548]

Meyers LA, Ancel FD, Lachmann M. Evolution of genetic potential. PLoS Comput Biol. 2005; 1:236–243. [PubMed: 16158095]

Montville R, Froissart R, Remold SK, Tenaillon O, Turner PE. Evolution of mutational robustness in an RNA virus. PLoS Biol. 2005; 3:e381. [PubMed: 16248678]

Mueller S, Papamichail D, Coleman JR, Skiena S, Wimmer E. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. J Virol. 2006; 80:9687–9696. [PubMed: 16973573]

Pfeiffer JK, Kirkegaard K. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. PLoS Pathog. 2005; 1:e11. [PubMed: 16220146]

Pfeiffer JK, Kirkegaard K. Bottleneck-mediated quasispecies restriction during spread of an RNA virus from inoculation site to brain. Proc Natl Acad Sci USA. 2006; 103:5520–5525. [PubMed: 16567621]

Pincus SE, Diamond DC, Emini EA, Wimmer E. Guanidine-selected mutants of poliovirus: mapping of point mutations to polypeptide 2C. J Virol. 1986; 57:638–646. [PubMed: 3003395]

Plotkin JB, Dushoff J. Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus. Proc Natl Acad Sci USA. 2003; 100:7152–7157. [PubMed: 12748378]

Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet. 2011; 12:32–42. [PubMed: 21102527]

Ren RB, Costantini F, Gorgacz EJ, Lee JJ, Racaniello VR. Transgenic mice expressing a human poliovirus receptor: a new model for poliomyelitis. Cell. 1990; 63:353–362. [PubMed: 2170026]

Sanjuán R, Cuevas JM, Furió V, Holmes EC, Moya A. Selection for robustness in mutagenized RNA viruses. PLoS Genet. 2007; 3:e93. [PubMed: 17571922]

Sanjuán R, Moya A, Elena SF. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc Natl Acad Sci USA. 2004; 101:8396–8401. [PubMed: 15159545]

Sardanyés J, Elena SF, Solé RV. Simple quasispecies models for the survival-of-the-flattest effect: The role of space. J. Theor. Biol. 2008; 250:560–568. [PubMed: 18054366]

Simmonds P, Smith DB. Structural constraints on RNA virus evolution. J Virol. 1999; 73:5787–5794. [PubMed: 10364330]

Stephens CR, Waelbroeck H. Codon bias and mutability in HIV sequences. J Mol Evol. 1999; 48:390–397. [PubMed: 10079277]

Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, Pan T, Dahan O, Furman I, Pilpel Y. An evolutionarily conserved mechanism for controlling the efficiency of protein translation. Cell. 2010; 141:344–354. [PubMed: 20403328]

Valdar WSJ. Scoring residue conservation. Proteins. 2002; 48:227–241. [PubMed: 12112692]

van Nimwegen E. Epidemiology. Influenza escapes immunity along neutral networks. Science. 2006; 314:1884–1886. [PubMed: 17185589]

Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature. 2006; 439:344–348. [PubMed: 16327776]

Vignuzzi M, Wendt E, Andino R. Engineering attenuated virus vaccines by controlling replication fidelity. Nat Med. 2008; 14:154–161. [PubMed: 18246077]

Wagner A. Robustness, evolvability, and neutrality. FEBS Lett. 2005; 579:1772–1778. [PubMed: 15763550]

Wagner A. Robustness and evolvability: a paradox resolved. Proc Biol Sci. 2008; 275:91–100. [PubMed: 17971325]

Wagner A, Stadler PF. Viral RNA and evolved mutational robustness. J. Exp. Zool. 1999; 285:119–127. [PubMed: 10440723]

Wilke CO, Wang JL, Ofria C, Lenski RE, Adami C. Evolution of digital organisms at high mutation rates leads to survival of the flattest. Nature. 2001; 412:331–333. [PubMed: 11460163]

Wimmer E, Paul AV. Synthetic poliovirus and other designer viruses: what have we learned from them? Annu Rev Microbiol. 2011; 65:583–609. [PubMed: 21756105]

Yang C-F, Chen H-Y, Jorba J, Sun H-C, Yang S-J, Lee H-C, Huang Y-C, Lin T-Y, Chen P-J, Shimizu H, et al. Intratypic recombination among lineages of type 1 vaccine-derived poliovirus emerging during chronic infection of an immunodeficient patient. J Virol. 2005; 79:12623–12634. [PubMed: 16188964]

- Codon usage influences the mutational robustness and variant distribution in a viral population
- A population's mutant spectrum determines its phenotype, independent of the protein coding consensus
- Synonymous populations exhibit distinct mutant spectra
- Mutational robustness and virulence are highly correlated in an animal model of infection

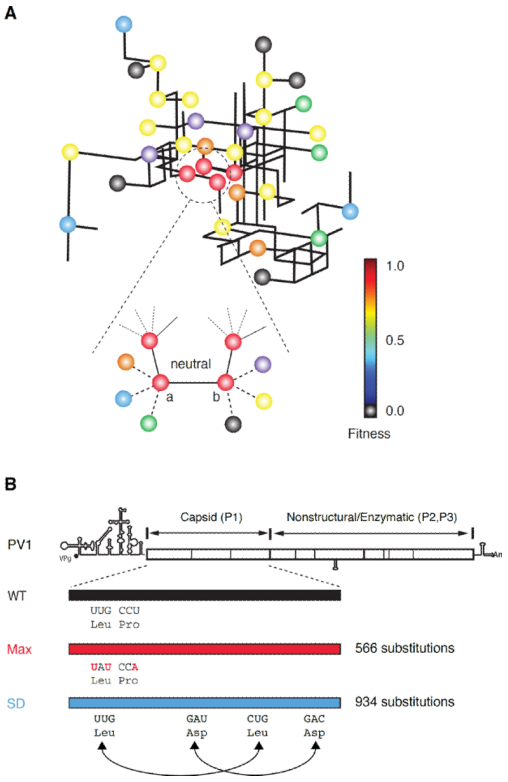

Figure 1. Using synonymous mutation to place viral populations in unique regions of sequence space with distinct fitness landscapes

(A) Schematic of codon networks and their hypothesized role in viral evolution. In this "pipe" network (top), connections represent evolutionary pathways. Coloring is according to hypothetical fitness values for the corresponding viral variant (red, highest; blue, lowest; gray, zero fitness). The bottom panel shows two variants with distinct mutant neighborhoods. While these variants are connected by a neutral, synonymous mutation, they produce progeny with distinct fitness values and adaptive capacity. Because large-scale synonymous mutation places populations in distinct regions of sequence space, the networks are hypothesized to have different genetic structures, which will affect the evolutionary capacity of the population (adapted from van Nimwegen, 2006; Draghi et al., 2010).

(B) Schematic of synthetic viruses used in this study. The genomic organization of poliovirus type 1 (PV1, wild type) is shown. All viral proteins are derived by proteolytic cleavage of a single polyprotein. The synthetic viruses Max and SD were designed as negative controls for studies of codon usage and polyprotein translation. Each has hundreds of synonymous substitutions across the capsid sequence relative to the wild type (WT), while preserving overall codon bias and GC content. Max has codon pair bias matched closely to that of the human genome. In SD, the codons are shuffled randomly throughout the capsid coding region.

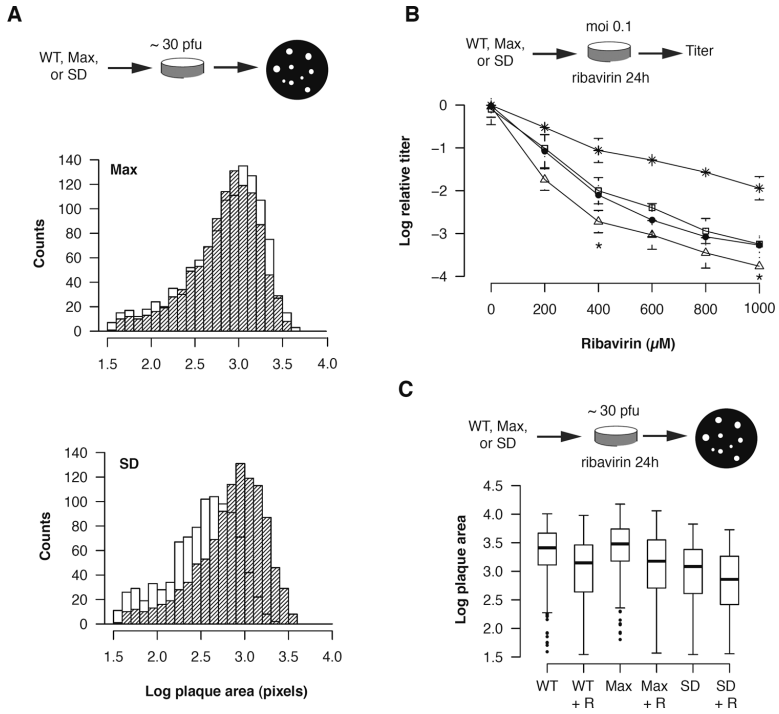

Figure 2. Synonymous populations differ in their robustness to mutation

(A) Fitness distribution of variants in virus populations. Histograms of plaque size for WT (hatched bars, both panels, n=1053), Max (white bars, top panel, n=1146), and SD (white bars, bottom panel, n=937).

(B) Synonymous populations exhibit differential sensitivity to a mutagenic nucleoside. The titers of ribavirin-treated populations were normalized to mock-treated controls. Data for WT (filled circles), Max (open squares), SD (open triangles), and G64S (ribavirin resistant variant, asterisks) populations are shown. Data points and error bars represent mean ± 95% confidence interval for 5 replicates, and significant differences between WT and SD are indicated (t-test, * $p < 0.05$).

(C) Reduction in plaque size in ribavirin treated clones. Each population was passed in 400μM ribavirin (+R) or mock-treated media for 24 hours. Plaques from each population were quantified as in (b). Boxplots indicate median titer, 25% and 75% quartiles, and 1.5 * interquartile range.

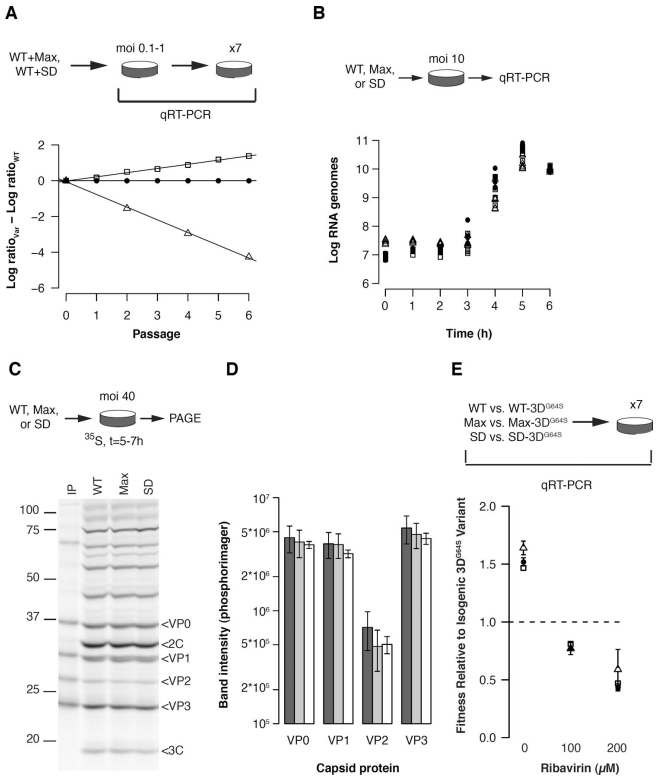

Figure 3. Synonymous mutation affects the fitness of virus populations

(A) Population fitness as measured by direct competition. The amount of each virus at each passage was compared to the input. Data are expressed as the $\log_{10}$ change in ratio for Max (open squares) and SD (open triangles) over time relative to WT (filled circles, shown here as reference). The slope reflects the relative fitness of each population. Representative data are shown and values for mean ± standard deviation, n=5, are given in the text.

(B) Kinetics of genome replication for wild type (WT, filled circles), Max (open squares), and SD (open triangles) with 5 replicates for each virus and time point. Mean ± bootstrap derived standard deviation for key parameters include: exponential growth rate (WT 1.8972 ± 0.1293, Max 2.0076 ± 0.1114, SD 1.6818 ± 0.0977), maximum genome number (WT 10.7099 ± 0.0522, Max 10.4984 ± 0.0672, SD 10.13889 ± 0.08492), and area under the curve (WT 50.9156± 0.2178, Max 49.6591± 0.2497, SD 50.1618± 1.6402). See also Figure S1.

(C) Protein expression in radiolabeled, infected cells. Capsid proteins (VP0–VP3) were identified by comparison to an anti-capsid immunoprecipitate, left.

(D) Quantitation of capsid gel bands for WT (dark grey), Max (light grey), and SD (white). Mean ± standard deviation, n=5.

(E) Changes in fitness at elevated mutation rates. WT (filled circle), Max (open square), and SD (open triangle) populations were competed against isogenic, ribavirin resistant references containing a single point mutation, 3D^{G64S}. Competition assays were performed over seven, single cycle, low moi (< 0.5) passages. Shown are fitness values (mean ± standard deviation, n=4) at varying ribavirin concentration.

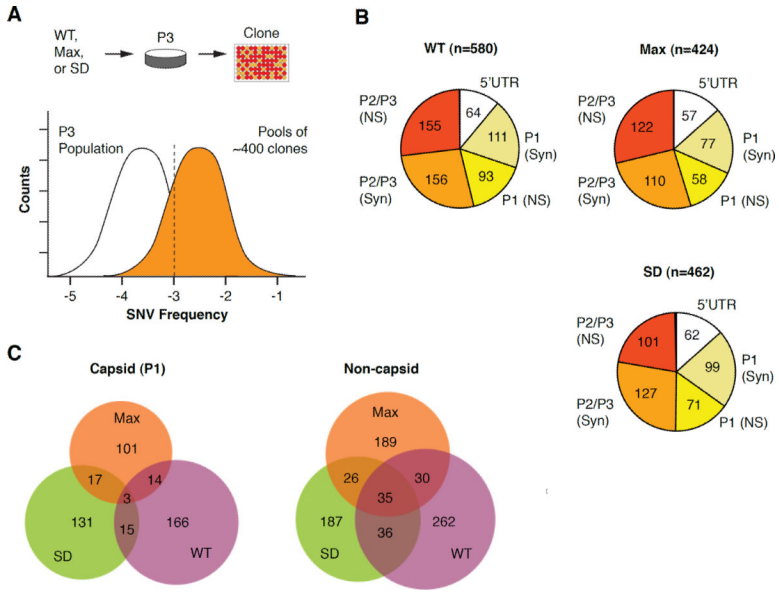

Figure 4. Deep sequencing of mutant spectra

(A) Variants in WT, Max, and SD were cloned by end-point dilution, and 400–500 clones were pooled and sequenced. The cloning process shifted the frequency of observed single nucleotide variants, reducing the likelihood of false-positive base calls. Variants with a frequency >0.001 were used in downstream analyses. Two replicate pools of clones were sequenced for a total of ~850 clones for each virus. See also Figures S2, S4., Table S1.

(B) Pie charts show the number of unique variants identified in the 5' noncoding region (5'UTR), capsid protein (P1), non-structural proteins (P2/P3), and 3' noncoding region (3'UTR). The numbers of nonsynonymous (NS) and synonymous (Syn) substitutions in coding regions are also indicated.

(C) Venn diagrams show overlap in mutant nucleotide position in clones derived from each population.

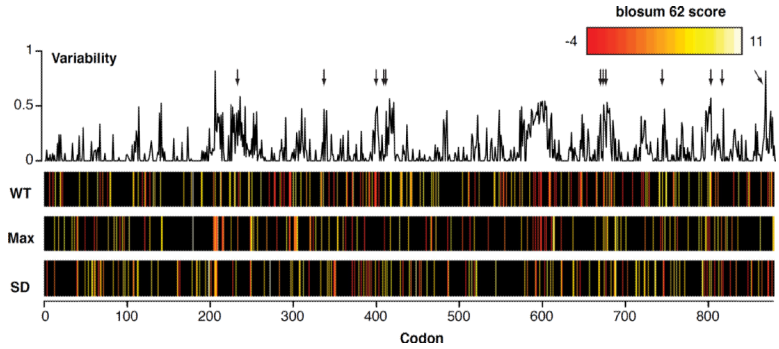

Figure 5. Synonymous populations exhibit distinct sets of polymorphic amino acid substitutions

An amino acid alignment of enterovirus capsids was used to derive a variability score for each codon (top). Higher values indicate more observed variation (Valdar, 2002). Arrows highlight antigenic regions of poliovirus capsid. Heatmap (bottom) of location and blosum62 conservation score for each codon (Henikoff and Henikoff, 1992). Both nonsynonymous and synonymous substitutions are included, and invariant amino acids are shown in black. For codons with ≥2 distinct variants, the less conserved substitution (lower blosum62 score) is shown. See also Figures S3, S4.

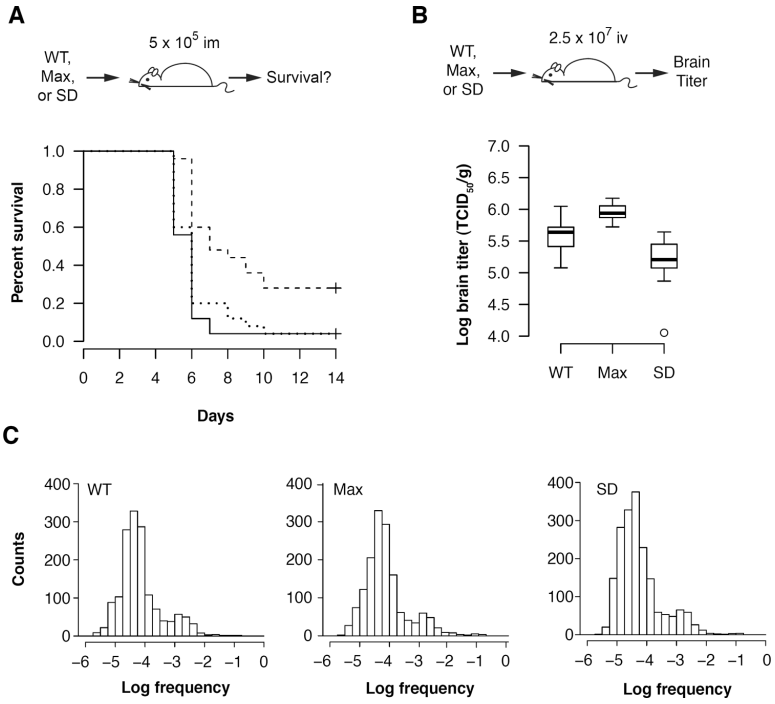

Figure 6. Mutational robustness of poliovirus populations is strongly correlated with virulence in a mouse model of infection

(A) Survival analysis of poliovirus receptor transgenic mice infected intramuscularly with WT (solid line), Max (dotted line), or SD (dashed line) populations, n=25 mice per group.

(B) Titer of virus recovered from brains of mice 5 days after intravenous inoculation. Boxplots indicate median titer, 25% and 75% quartiles, and 1.5 * interquartile range for each virus population, n=10 mice per group.

(C) Sequence analysis of brain-derived mutant spectra. For each virus, approximately 540 clones were recovered from the brains of 10 infected mice and pooled for deep sequencing. Histograms show the frequency of single nucleotide variants identified in each pool of clones. See also Table S1.