ΦX174 Attenuation by Whole Genome Codon Deoptimization

James T. Van Leuven1,2,‡, Martina M. Ederer1,‡, Katelyn Burleigh1,3, LuAnn Scott1, Randall A. Hughes4,5, Vlad Codrea6, Andrew D. Ellington4,6, Holly Wichman1,2, Craig Miller1,2,*

1Department of Biological Science, University of Idaho, Moscow, ID 83844
2Institute for Modeling Collaboration and Innovation, University of Idaho, Moscow, ID 83844
3Current address: Seattle Children’s Research Institute, Seattle, WA
4Applied Research Laboratories, University of Texas, Austin, TX 78758
5Current address: Biotechnology Branch, CCDC US Army Research Laboratory, 2800 Powder Mill Rd, Adelphi, MD 20783
6Institute for Cellular and Molecular Biology, University of Texas, Austin, TX 78712
‡JTV and MME contributed equally to the work

*Corresponding author:
Craig Miller
Dept. of Biological Sciences
University of Idaho
Moscow, ID 83844-3051
crmiller@uidaho.edu

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License
(http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the
original work is properly cited.
Abstract
Natural selection acting on synonymous mutations in protein-coding genes influences genome composition and evolution. In viruses, introducing synonymous mutations in genes encoding structural proteins can drastically reduce viral growth, providing a means to generate potent, live attenuated vaccine candidates. However, an improved understanding of what compositional features are under selection and how combinations of synonymous mutations affect viral growth is needed to predictably attenuate viruses and make them resistant to reversion. We systematically recoded all non-overlapping genes of the bacteriophage ΦX174 with codons rarely used in its E. coli host. The fitness of recombinant viruses decreases as additional deoptimizing mutations are made to the genome, although not always linearly, and not consistently across genes. Combining deoptimizing mutations may reduce viral fitness more or less than expected from the effect size of the constituent mutations and we point out difficulties in untangling correlated compositional features. We test our model by optimizing the same genes and find that the relationship between codon usage and fitness does not hold for optimization, suggesting that wild-type ΦX174 is at a fitness optimum. This work highlights the need to better understand how selection acts on patterns of synonymous codon usage across the genome and provides a convenient system to investigate the genetic determinants of virulence.

Key words
Bacteriophage, epistasis, fitness landscape, synthetic biology, live attenuated vaccine, codon bias

Significance statement
Attenuating viruses by inserting many synonymous, deleterious mutations offers a means to make potent and reversion-resistant vaccines. We investigate where in a virus genome attenuating mutations should be made and how they should be combined by generating a combinatorial network of codon deoptimized bacteriophage strains. By analyzing the effects of genome editing using mathematical models of epistasis, we find that fitness effects differ between genes in how these deleterious mutations combine. These results show how synonymous mutations can have large effects and will help researchers design synonymously recoded, live attenuated vaccines.
Introduction

**Synonymous compositional features of viral genomes**

The unequal use of synonymous codons is known as codon usage bias. Codon biases are the result of an interaction between mutational and selective pressures (Bulmer 1991; Long et al. 2018) and are observed in organisms across the tree of life (Grantham et al. 1980; Ikemura 1985; Hershberg and Petrov 2008). A commonly accepted explanation for why organisms do not evenly use synonymous codons is that the rate of translation (i.e., the amount of proteins being made) is affected by the abundance of tRNAs that pair with codons on a strand of mRNA (Ikemura 1985). Codons that result in the optimal amount of protein then confer a selective advantage. However, there are many other compositional features within the DNA sequences of protein-coding genes upon which selection acts and understanding what features are most influential on selection is a complicated endeavor. These features include: the genic GC content (Raghavan et al. 2012; Kelkar et al. 2015; Newman et al. 2016), CpG or TpA dinucleotides (Burns et al. 2009; Atkinson et al. 2014; Gaunt et al. 2016; Fros et al. 2017; Giallonardo et al. 2017), codon pairs (Gutman and Hatfield 1989; Irwin et al. 1995; Tats et al. 2008; Gamble et al. 2016), endonuclease recognition sites (Karlin et al. 1992; Levin 1993; Rocha et al. 2001; Pleška and Guet 2017), intron splicing motifs (Cáceres and Hurst 2013), mRNA folding stability (Kudla et al. 2009; Presnyak et al. 2015; Boël et al. 2016; Kelsic et al. 2016; Burkhardt et al. 2017; Jack et al. 2019), ribosomal pausing sites (Ponnala 2010; Li et al. 2012), concentration of unpreferred codons at the 5' ends of transcripts (Chen and Inouye 1990; Tuller et al. 2010; Goodman et al. 2013), autocorrelation of codons on transcripts (Cannarozzi et al. 2010), and capacity of codon order to influence cotranslational folding of proteins (Zhang et al. 2009; Yu et al. 2015). Natural selection acting on one or more of these features can favor the use of certain codons over other synonymous ones. The strength of selection acting on these synonymous codons can be quite strong (Agashe et al. 2016; Lawrie et al. 2013; Bailey et al. 2014; Knöppel et al. 2016; Machado et al. 2017; Kristofich et al. 2018) and some studies compared the relative impact of altering different features (e.g., codon bias vs. mRNA folding). Kudla et al. (2009) generated 154 versions of green fluorescent protein (GFP) that varied only at synonymous sites and found that mRNA folding around the ribosomal binding site to be most predictive of exogenous GFP expression. In an analysis of 6,348 cloned and expressed genes, Boël et al. (2016) found mRNA folding around the translation initiation site to be the second most important predictor of expression level, behind overall codon usage.
As viruses must utilize their hosts’ cellular machinery, there is an expectation that virus genomes are enriched for host-preferred codons to maximize production of viral proteins. This appears to be only partially true. Many viral genomes do contain more host-preferred codons than expected by chance, especially for genes encoding viral structural proteins of dsDNA phages (Carbone 2008; Lucks et al. 2008; Chithambaram et al. 2014). However, many viral genes are not enriched in host-preferred codons. Sometimes unpreferred codons are used to temporally regulate viral gene expression, potentially to avoid host immune responses (Shin et al. 2015). Other virus genomes appear to have little preference for codons abundant in the host genome. For instance, Lucks et al. (2008) found that the majority of 74 bacteriophage genomes show no significant preference for host-preferred codons. Similar discordance between host and viral codon usage patterns are observed in other studies (Kunisawa 1992; Kunisawa et al. 1998; Sau et al. 2005). This discordance could be caused by insufficient selection on codon usage, host-phage relationships that are too short-lived for selection to fine-tune codon usage in the phage, or an inadequate understanding of what features are being selected for. Synonymously editing viral genomes provide an opportunity to learn about viral adaptation to hosts.

Vaccine development by synonymous recoding

Empirically developed (e.g., serial passage viral adaptation) vaccines have saved millions of lives over the last century, yet methodological improvements make rationally designed, recombinant vaccines attractive because they can be rapidly produced and specifically engineered for safety and effectiveness (Lim et al. 2006; Rueckert and Guzmán 2012; Nabel 2013; Minor 2015; Ramezanpour et al. 2016). One proposed method of generating recombinant vaccines involves making many synonymous, attenuating changes to viral genomes, i.e., “deoptimizing” the viral genes (Mueller et al. 2010). The recombinant vaccine can be made by either editing the genome of the wild-type virus or by generating a viral genome entirely from synthesized nucleic acids. Synonymous deoptimization offers a potentially efficient and effective way of making vaccines: the protein sequences of recoded vaccines are identical to their target viruses, they replicate in their host to provide prolonged exposure to the antigen, and the introduction of many synonymous changes presumably assures evolutionary robustness, preventing the evolution of virulence by reversion.

Poliovirus serves as a very good example for the synonymous recoding strategy. Development of a robust, live-attenuated poliovirus vaccine is desired because in some areas on
Earth wild poliovirus and the emergent vaccine-derived polioviruses (cVDPVs) continue to cause concern over the resurgence of poliomyelitis (Cann et al. 1984; Kew 2012; Famulare et al. 2015; Jorba et al. 2016). A synthetic poliovirus was assembled in 2002 (Cello et al. 2002), codon deoptimized in 2006 (Burns et al. 2006; Mueller et al. 2006), codon pair deoptimized in 2008 (Coleman et al. 2008), and dinucleotide deoptimized in 2009 (Burns et al. 2009). In all cases, attenuated viruses were produced by recoding the P1/capsid region of the genome. \textit{In vitro}, these viruses replicate slower and produce lower viral titers than wild-type virus. \textit{In vivo}, the codon pair deoptimized strain protected mice against challenge by wild poliovirus (Coleman et al. 2008). While the mechanism of attenuation is not yet fully elucidated, reduced protein expression of the deoptimized genes is observed (Burns et al. 2006; Mueller et al. 2006; Coleman et al. 2008; Burns et al. 2009). These deoptimized poliovirus constructs are genetically stable and remain non-virulent for up to 25 passages in cell culture (Burns et al. 2006; Coleman et al. 2008).

The apparent success of building poliovirus vaccine candidates using synonymous recoding led to similar attempts to develop vaccines for influenza, adeno-associated, human immunodeficiency, papilloma, chikungunya, respiratory syncytial, simian immunodeficiency, porcine reproductive and respiratory syndrome, echovirus 7, tick-borne encephalitis, vesicular stomatitis, dengue, T7, Lassa, adeno, and swine fever viruses (reviewed in Martínez et al. 2016). The most common method for synonymously deoptimizing viruses is recoding wildtype genes with increased proportions of unpreferred codons (Mueller et al. 2006; Luan et al. 2009; Bull et al. 2012; Cladel et al. 2013; Meng et al. 2014; Nogales et al. 2014; Cheng et al. 2015; Rostad et al. 2016; Velazquez-Salinas et al. 2016; Cheng et al. 2017) although other methods of recoding have been successful as well. For example, viral fitness was decreased when synonymous substitutions were randomly introduced (Nougairede et al. 2013; Fabritus et al. 2015; Fabritus et al. 2016), when codons were replaced by those infrequently used in viral (not host) genes (Burns et al. 2006; Meng et al. 2014), when the proportion of optimal codons was \textit{increased} (Cladel et al. 2013; Vabret et al. 2014; Liang et al. 2017; Villanueva et al. 2016), or when codons were exchanged for codons one substitution away from a translational termination codon (Moratorio et al. 2017).

\textit{No predictive understanding of synonymous recoding}

While it is clear that synonymous recoding causes attenuation and the strategy holds promise for vaccine development, we lack a predictive understanding of the process. Part of this
results from the biological complexity and variation in the systems involved. In many cases, the fitness impact of recoding is cell-line dependent (Martrus et al. 2013; Nougairede et al. 2013; Le Nouën et al. 2014; Meng et al. 2014; Cheng et al. 2015; Shen et al. 2015; Rostad et al. 2016; Velazquez-Salinas et al. 2016), is inconsistent between in vivo and in vitro experiments (Shen et al. 2015; Velazquez-Salinas et al. 2016; Cheng et al. 2017), or is temporally variable (Villanueva et al. 2016). Another obstacle is the nature of the genetic code itself. It is generally challenging to manipulate one synonymous feature of the genome and hold all the others fixed. For example, when codons are shuffled to change codon pair frequency, mRNA stability may be affected, or when codons are deoptimized, codon pair frequencies also change. This makes it difficult to attribute the cause of fitness decreases to one factor (e.g., codon usage adaptation), especially when the features are correlated. As we do here, most studies have focused on manipulating a single compositional feature of the genome and measuring its impact on fitness. Standardizing recoding methodologies and features measured across studies would greatly improve our understanding of the factors that drive fitness decreases and other phenotypic effects caused by synonymous deoptimization.

Despite the optimistic results achieved in studies on synonymous recoding to date, basic questions underlying the method itself remain unanswered (Martínez et al. 2016). What is the best strategy to perform synonymous recoding to achieve attenuation? Can generalities be made about the extent of recoding and the degree of attenuation—or will the biological details and idiosyncrasies of each system preclude this? What is the mechanistic cause of attenuation from synonymous recoding? Are viruses recoded this way robust against fitness recovery? Is it more effective to maximally recode less of the genome (say one gene), or make the recoding less severe, but distribute it across the genome? As an increasing proportion of the genome is recoded, or equivalently, as multiple recoded parts are combined, does attenuation respond in an additive or non-additive manner? A deeper understanding of genome evolution and synonymous sequence choice is required to answer these questions.

In this paper, we focus on two issues related to the codon deoptimization of viruses. First, we seek to compare the fitness effects of mutating different genes in the same virus. Second, we seek to understand how, when attenuating mutations are combined, they interact to affect fitness (i.e., epistasis of deleterious mutations). The nature of epistasis among fragments is crucial for modeling fitness effects: if mutational effects combine synergistically (i.e., the combined fitness being even lower than predicted from the observed individual effects), the range in the number of mutations needed to achieve the targeted attenuation level would tend to be reduced (Figure 1).
Conversely, if they combine antagonistically (i.e., the combination of mutations are less attenuated than predicted from individual effects), it may be easier to achieve a target attenuation level, but there may be a limit to how much attenuation is possible. If mutations, in combination, display sign epistasis, irregular magnitude epistasis, or even vary between synergistic and antagonistic epistasis, it will suggest the underlying process is complex and difficult to predict and generalize. To evaluate these issues, herein we have recoded all the non-overlapping genes of the bacteriophage ΦX174 in fragments, combined recoded fragments in all possible within-gene permutations, and measured the fitness of the resulting recoded bacteriophage.

Fig. 1 - Invented data illustrates that epistasis affects how a desired level of attenuation is achieved. When a substantial amount of attenuation is desired (the “targeted attenuation range” is at a low fitness level), the amount of attenuation (e.g., number of deleterious mutations) will be harder to achieve if mutational effects combine synergistically (negative epistasis) because fitness declines at an increasing rate. In this case, the targeted amount of attenuation will be easier to achieve if mutations combine antagonistically (positive epistasis) because a larger range in the number of deleterious mutations results in the same level of fitness effects. Notice that this pattern is reversed if a slight level of attenuation is desired (near y=0). If sign epistasis, or even irregular magnitude epistasis is observed, then the underlying nature of interactions is more difficult to predict and generalize.

Results

**Synonymously deoptimizing ΦX174 genes**
ΦX174 is a bacteriophage with a 5.4kb single-stranded DNA genome containing 11 genes (Figure 2a, table 1). We measured codon usage bias of ΦX174 genes using the codon adaptation index (CAI). CAI is a gene-level statistic running from zero to one that summarizes the extent that codons in a gene are used, rarely (CAI nearer 0) or commonly (CAI nearer 1) among highly expressed host genes (Sharp and Li 1987). We found that most ΦX174 genes are not particularly enriched for preferred E. coli codons (Figure 2b). Only gene J has a CAI value in the upper quartile of E. coli genes. Gene K uses the most unpreferred codons and has a CAI value near the lowest value observed for genes in the E. coli genome. Genes J and K were the only wildtype ΦX174 genes found to be significantly different from simulated genes of similar length using the host’s codon usage preferences (p<0.01, supplementary materials). All other ΦX174 genes have CAI values within the range of most E. coli protein-coding genes. ΦX174 structural proteins (B, D, F, G, H, J) have higher CAI values than non-structural genes (A, C, E, K), suggesting that high expression of these proteins is important for viral fitness. When we computationally deoptimized entire ΦX174 genes (i.e., recoded them to use the least-preferred codons throughout), the resulting CAI values were in the lower tail or even below the tail for all E. coli genes (Figure 2, supplementary materials). These reductions in CAI were the result of changing between 42% (20/48 codons for gene C) and 75% (24/32 codons for gene J) of the codons of a gene, corresponding to 15-32% of its base pairs (Supplementary table S1). All the other ΦX174 genes fall within this range of recoding (42-75% of codons changed). We calculated additional metrics of codon adaptation including an alternative version of CAI (Xia 2007), tRNA adaptation index: tAI (dos Reis et al. 2004), index of translation elongation: I_{TE} (Xia 2015), relative codon adaptation: RCA (Fox and Erill 2010), the number of effective codons: Nc (Wright 1990), COdon Usage Similarity INdex (Bourret et al. 2019), and the starvation codon adaptation index: sCAI (Elf et al. 2003). Nc is a simple index of codon bias that measures the deviation from uniform codon usage. RCA is similar to Nc in that it can be calculated for genes without additional genetic information, but it provides a measure of codon preference that is corrected for gene length and nucleotide content. Like Nc and RCA, tAI does not rely on a list of preferred codons, but it does require the tRNA gene copy number of a genome. A list of highly expressed genes (or the frequency of preferred codons) is needed to calculate both CAI and I_{TE}, but I_{TE} differs from CAI in its handling of R- and Y-ending codon subfamilies. To calculate sCAI, empirical measures of tRNA concentrations are needed. Like CAI and I_{TE}, COUSIN requires a codon usage table from a user-defined gene set (e.g., highly expressed genes), but is uniquely powerful in comparing codon usage to a random codon null and a reference gene set. Calculation details are in the Materials
and Methods. All analyses produced qualitatively similar rank-orders for ΦX174 genes (Supplementary materials).

Table 1 - ΦX174 gene function and protein copy number required for the assembly of one virion.

| Gene | Protein Function                 | Copies | Gene length (bp) |
|------|----------------------------------|--------|------------------|
| A    | DNA replication                  |        | 1,541            |
| A*   | DNA packaging, regulation of DNA replication | 60     | 1,025            |
| B    | internal procapsid scaffolding   | 60     | 362              |
| K    | unknown, not essential           | 240    | 260              |
| C    | regulation of DNA replication    | 240    | 170              |
| D    | external procapsid scaffolding   |        | 362              |
| E    | cell lysis                       |        | 458              |
| J    | DNA packaging                     | 60     | 275              |
| F    | major capsid protein             | 60     | 1,283            |
| G    | major spike protein              | 60     | 527              |
| H    | minor spike, pilot protein       | 12     | 986              |

* Genes encoding structural proteins are bolded.
Fig. 2 - ΦX174 genome organization and capacity for deoptimization relative to host genes. (a) Genes on the ΦX174 genome are labeled at the top and shown as white boxes. Recoded regions are shown as filled blue boxes. These fragments are named consecutively (e.g., A1, A2, A3, A4, F1, F2, F3,...). Transcript expression levels are shown as filled grey bands. The band heights are proportional to the relative number of transcripts by RT-qPCR (Zhao et al. 2012). (b) Codon adaptation index (CAI) of E. coli, wild-type ΦX174, and recoded ΦX174 genes. Genes highly expressed in E. coli are enumerated on the secondary y-axis. The white space between recoded fragments (the BsmBI sites) were enlarged for visualization. Lines connect genes that were fully recoded. Genes containing only some recoded fragments (e.g., deoptimized A1) do not have black boarders.

Codon deoptimization of ΦX174 genes reduces viral fitness

We codon deoptimized whole ΦX174 genes by exchanging wildtype ΦX174 codons for synonymous codons less commonly used by its E. coli host. Because amino acid usage is not random in ΦX174, our recoding resulted in some amino acids being changed synonymously more often than others (Supplementary figure S1). We did not recode regions of genes that overlapped with other genes nor the first six codons of each gene since these codons are known to have
strong effects on gene expression (Bentele et al. 2013). This leaves six genes (A, C, J, F, G, H) that could be deoptimized. The construct containing the fully deoptimized G gene could not be recovered, even after growing the strain overnight in an attempt to obtain a recovery mutation. Of the remaining five constructs, four were less fit than wild type \(\Phi X174\) (Figure 3a). Recoding highly expressed genes (J, F, and G) resulted in larger fitness decreases than recoding lowly expressed genes (A and H). Although the number of variants built was small, the fitness effects of deoptimization were correlated to the amount of recoding performed and the change in CAI of the recoded genes (Figure 3b and supplementary table S1). However, the change in CAI is highly correlated with the number of changes made to the gene \(R^2=0.92\) and although a stepwise regression analysis does not suggest removing either from a multivariate model, the variance inflation factor (VIF) values for the % of codons changed, the # of codons changed, the change in CAI, and the relative expression of wild-type genes are 32.8, 3.8, 31.4, and 3.8, respectively.

Fig. 3 - Fitness effects of deoptimizing \(\Phi X174\) genes. (a) The fitness of wild type and deoptimized \(\Phi X174\) strains containing recoded genes are shown in replicate (colored dots). Means and standard error bars are shown in black. Fitness values that are significantly different from wild type are indicated with asterisks (ANOVA, \(p<0.01\)). (b) Fitness is plotted against...
measures that potentially explain fitness decreases. Fitness is the number of doublings per hour (log2 of the ratio of the phage concentration at 60 min divided by the phage concentration at time zero). Gene expression levels are from Logel and Jaschke (2020) and are normalized to gene A. Structural proteins are shown as empty circles. Deoptimizing gene G yielded no viable phage. The total number of independent fitness measurements is provided in supplementary materials. At least 3 replicates were performed for every strain. Y-axes are the same in panels (a) and (b). The percent of codons changed and change in CAI are highly correlated (R^2=0.92).

Reconstructing a combinatorial fitness landscape for deoptimized genes

We segmented the ΦX174 genome into 14 fragments (Figure 2) and measured the fitness of all of the possible within-gene combinations of deoptimized gene fragments (Figure 4 and supplementary materials). Since genes C and J are short and encoded entirely on one fragment each, we analyzed combinations of the remaining 12 fragments. Of the 12 deoptimized strains with only one deoptimized fragment, only 6 have fitness values below wild type. The moderate fitness effects of these partially recoded genes allowed us to observe how deleterious effects combine. As additional deoptimized fragments are joined, the fitness of the resulting viruses decreases (Figure 4b). In most cases combining deoptimized fragments results in less fit viruses. The exception is gene A where instances of sign epistasis are observed. Specifically, the average fitness of A1+A3, A2+A3, A1+A3+A4, A1+A2+A4, A2+A3+A4, and A1+A2+A3+A4 are all higher than at least one of their constituent fitness values (Figure 4b). To further investigate how deleterious effects combine, we employed a statistical framework for calculating the best-fitting model of epistasis.
Fitting models of epistasis to combinatorial fitness data

The combinatorial network of genotypes that we generated in this work can be analyzed by applying simple models of epistasis (Miller et al. 2017) to determine how the effects of mutations combine. We fit the data for genes A, F and H to three basic models—additive, multiplicative, and stickbreaking—which gave rise to no, antagonistic, and synergistic epistasis, respectively (see figure 1 and (Nagel et al. 2012)). In fitting the three models, we conducted two analyses for each gene: one of absolute fit where we assess if the data is consistent with each model individually, and one of relative fit wherein one of the three models is assumed to be correct. The results from this analysis were not highly conclusive, but suggest the nature of epistasis is heterogeneous across different genes. For genes F and H, none of the three models could be rejected based on absolute goodness of fit (Table 2). For gene F, the additive model provides the best fit to the data. For gene H, stickbreaking gives the best fit ($R^2=0.885$), consistent with synergistic epistasis. This is visually clear in figure 4b, where the fully recoded gene H (3
recoded fragments) has far lower fitness than one would expect based on the individually recoded fragments—all of which were basically neutral.

Table 2 - Models of epistasis fit to combinatorial fitness data.

| Gene | Add  | Mult | Stick | Add  | Mult | Stick | Add  | Mult | Stick |
|------|------|------|-------|------|------|-------|------|------|-------|
| A    | 0.011| 0.066| <0.001| 0.145| 0.845| 0.009 | 0.761| 0.860| 0.388 |
| F    | 0.388| 0.071| 0.235 | 0.609| 0.090| 0.301 | 0.965| 0.843| 0.950 |
| H    | 0.513| 0.489| 0.733 | 0.056| 0.054| 0.891 | 0.709| 0.562| 0.885 |

Note - A regression of each recoded fragment's fitness effect (against background) under each model was performed. The p-values of each regression were combined by taking the sum of their logs. Using parametric bootstrap, the distribution of this sum was simulated. The overall p-value is estimated by the proportion of simulations where the sum of logs is ≤ the observed value.

a Absolute goodness of fit. Small p-values indicate that the data is inconsistent with the model (gray-filled). When a model is correct, a recoded block's effect is uncorrelated to background fitness. The p-value indicates how often, under parametric bootstrapping, the correlation of effect to background fitness is as strong as or stronger than that observed in the real data.

b The posterior probability assumes that one of the three models—additivity, multiplicative, stickbreaking—is correct.

Visually, a pattern of antagonistic epistasis was observed for gene A, as several of the variants with two recoded fragments had fitness values as low as or even slightly lower than the three- and four-recoded fragment variants (Figure 4b). Indeed, the additive and stickbreaking models were rejected for gene A based on absolute goodness of fit (Table 2). The multiplicative model, with its antagonistic pattern of epistasis, was not rejected, but the p-value was marginal (p=0.066). Strong antagonistic epistasis is occurring for gene A—even stronger than that predicted under the multiplicative model. This was revealed by regressing background fitness against fitness effect (Figure 5a and supplementary figures S2-S4). When effects were measured as differences (the additive model), negative/positive slopes corresponded to antagonistic/synergistic epistasis. Under the correct model, no correlation exists and slopes are expected to be random deviations around zero. Under the additive model (Figure 5a), a clear pattern across all four recoded fragments where the effect of the fragment becomes more strongly
deleterious on higher fitness backgrounds (negative regression slopes) was observed. When the p-values of the individual fragments were combined, their result is significant (Supplementary figure S2). The analogous regression under the multiplicative model was less extreme, but even here the slopes were consistently negative, indicating a level of antagonistic epistasis beyond multiplicative (Supplementary figure S3).

Fig. 5 - Effect of recoded fragment on all possible backgrounds. (a) Regressions of each recoded fragment’s fitness effect under the additive model against background fitness for gene A. Fitness effect is the difference between the background and the background plus the recoded fragment (e.g., A2 -> A1+A2, A2+A3 -> A1+A2+A3). Horizontal lines indicate a perfect fit to the additive model with no residual effects of background. Sloped regression lines indicate antagonistic/synergistic epistasis. Solid regression lines indicate that the additive model can be rejected (linear model, p<0.05). The overall fit of epistatic models for each gene are shown in Table 2. In (b) fragment fitness effects are shown against the change in CAI. Slight point jitter was used for visualization. Linear regressions are shown with p and r-squared values. Change in CAI (Xia 2007 method) is proportional (CAI of recoded gene over CAI of background). AIC values of alternative models are shown in supplementary table S2.

Correlating codon deoptimization to combinatorial fitness data

Ultimately, our goal was to correlate changes in genomic properties (e.g., codon preference) to changes in viral fitness. The most straightforward method of analysis would be to regress the two measurements, and indeed the fitness of deoptimized variants was linearly correlated to CAI even when all genes are considered together ($R^2=0.36$, $p=2E-16$, Figure 6a). However, it is worth noting that the data points used in this regression were not independent.
because the deoptimized fragments were combined to achieve higher levels of deoptimization. Our combinometric method of making variants also allowed us to correct for the cumulative fitness effect of combined fragments by calculating the effect of adding any particular fragment to different backgrounds (Figure 5b). For example, the effect of deoptimizing the F1 fragment was measured by comparing the fitness values of WT to deoptimized F1 (20-21 = -1), or F2 to F1+F2 (18-20 = -2), or F3 to F1+F3 (14-16 = -2), or F2+F3 to F1+F2+F3 (8-10 = -2). Thus, deoptimizing F1 resulted in an average fitness effect of about -2. This background subtraction approach corrected for the non-independence of data points in regressions between change in fitness from WT and change in CAI from WT.

When we applied this correction, we observed a wide variance in fitness effects (Figure 5b). For example, in some backgrounds, adding deoptimized H1 reduced fitness by only ~1 doublings/hr. In other backgrounds, H1 reduced fitness by ~8 doublings per hour. Despite this variation, there is a good correlation between change in CAI and change in fitness (R²=0.29, p=0.01, Figure 5 and supplementary table S2). Applying this background correction indicates that only a portion of the fitness changes can be explained by changes in codon usage bias. This is particularly true for genes A and H. Fragments in gene F seem to have more consistent effects (Figure 5b).

**How different synonymous features correlate with fitness**

We replaced ΦX174 codons with less preferred codons without consideration for how alterations might affect other features in the genome. As mentioned in the introduction, many such features may be under selection. To investigate unintended consequences of codon deoptimization, we calculated numerous genome characteristics to see if any correlate with the fitness decreases observed in deoptimized fragments (supplementary materials). We included many different measures of codon usage bias (CAI, tAI, I_TE, etc.), codon pair bias (CPB), frequency of Shine-Dalgarno motifs, mRNA folding stability, as well as simply the number of changes made. The best predictor of fitness is the folding stability of the codon deoptimized mRNA (R²=0.33, p.adj=5.1E-3), which performed better than the best measure of codon usage bias which was CAI using the Xia2007 method (ΔAIC=1.6). This correlation is easily observed when mRNA stability values are plotted against change in fitness (Figure 6b) and when the change in folding stability from background is regressed against the change in fitness (Supplementary figure S5). However, many of the measures are highly correlated (Supplementary
materials), thus we performed a stepwise regression analysis which indicated that the change in $I_{TE}$, FOP, tAI, CPB, mfold, and Nc as well as the fraction of the gene edited should all be included as predictor variables in a multivariate model of fitness. This multivariate model has an adjusted $R^2$ of 0.91 and an AIC value of 134, which is better than the best univariate predictor of fitness change (fraction of the gene edited), which has $R^2$ and AIC values of 0.34 and 177, respectively (ANOVA, $p=2.1\times10^{-6}$). Note that CAI is not included, likely because CAI is so well correlated ($R^2=0.88$) with mfold.

We were interested to see if the correlation between genomic features like CAI and fitness held up even when features were optimized, so we replaced ΦX174 codons with codons frequently used in *E. coli* (Supplementary table S1). In all cases, fitness was either unaffected (8/11 viruses) or reduced (3/11 viruses) (ANOVA, $p<0.05$, figure 6a and supplementary materials). Because of this, if these optimized constructs are included in the regression models, the number of sites changed and fraction of gene edited become the metrics that best predict fitness from genomic measures. None of the other indices are significantly correlated with change in fitness when the optimized constructs were considered independently (linear model, $p<0.01$), likely because most of the optimized viruses have fitness values very near wildtype. We observed a peak-shaped fitness landscape when combining the optimized and deoptimized data set; this is discussed below.
Fig. 6 - Fitness of recoded viruses correlates with codon usage bias and mRNA folding stability. (a) Codon usage bias (CAI) compared to fitness for viruses optimized and deoptimized in genes A, F, G, and H. Fitness and CAI of wild type are indicated with grey horizontal and vertical lines. Points to the right of these lines are optimized. Points to the left are deoptimized. (b) Viral fitness compared to mRNA folding stability (mfold). Wild type values are indicated with grey horizontal and vertical lines. Less stable transcripts (mostly deoptimized genes) have less negative values and are right of wild type. More stable transcripts (most optimized genes) have more negative values. $R^2$ and p values shown are from individual (for each gene) linear regressions. A model with all deoptimized genes is shown in supplementary table S2. Codon optimized viruses are shown with empty circles. Those significantly different from wildtype are labelled (ANOVA, p<0.05).

Discussion

Patterns of synonymous codon usage biases

Synonymous codon usage biases are present in genomes across the tree of life (Granath et al. 1980). We often think of these biases as having little consequence during the natural evolution of organisms because the strength of selection acting on a single synonymous mutation is generally weak. Nevertheless, the presence of biases shows that selection acts with sufficient strength to maintain them in the face of genetic drift. The prevailing theories on the preservation of codon biases suggest that codon choice is primarily driven by selection on translational speed and mRNA stability (Granath et al. 1980; Robinson et al. 1984; Plotkin and Kudla 2010; Gamble et al. 2016). The enrichment of codons that use abundant tRNAs in highly expressed genes points towards a model where translation speed is correlated to tRNA abundance. We find that the most highly expressed ΦX174 protein ranks second best in its use of host-preferred codons, but only marginally better than the average E. coli gene (Figure 2). That ΦX174 genes are average in preferred codon usage bias according to E. coli usage patterns is not surprising—many viruses do not favor the most preferred host codons. Carbone et al. (2008) also found that most ΦX174 genes do not use host-preferred codons and in the 116 DNA phages they studied, capsid proteins were the most codon adapted to their hosts. ssDNA phages are particularly poorly matched to their hosts’ codon biases, which is likely due to mutational pressures (Chithambaram et al. 2014). Selection could also be acting to keep viral genes from evolving to their full codon usage potential. Codon usage could control the stoichiometric expression ratio between viral genes (Cherwa et al. 2011; Quax et al. 2013), temporally regulate gene expression (Aragonès et al. 2010; Shin et al. 2015; Villanueva et al. 2016; Mioduser et al. 2017), facilitate co-translational folding (Yu et al. 2015), dampen protein expression to avoid host immune responses (Zhao and Chen 2011; Cladel et al. 2013), be linked to global transcription patterns (Andersson and Kurland 1990; Frumkin et al. 2018), or be limited by other compositional

http://mc.manuscriptcentral.com/gbe
features. Regardless of cause, codon biases among related viruses are conserved, even when they infect different hosts that have variable codon preferences (Cardinale et al. 2013; Kula et al. 2018).

Recoding of φX174

Synonymous mutations can have substantial phenotypic effects, but it is often difficult to explain why and in what parts of a genome/gene these effects are most substantial. Our work on φX174 confirms that synonymous mutations can have massive (even lethal) fitness effects and that these effects combine in a predictable manner that is gene dependent. Although we introduced many synonymous mutations in each recoded φX174 strain, the largest observed fitness impact of any single deoptimized fragment contained only 29 synonymous codon changes. These 29 synonymous changes resulted in a 50% decrease in fitness, which is a decrease of about 10 doublings per hour or about a 1000-fold change in the number of offspring. Kula et al. (2018) also observed phenotypic effects of deoptimizing the J gene. They observed a 25% reduction in burst size caused by 12 synonymous changes made to a 23 codon region of the J gene. Similar amounts of decrease were observed when 11 synonymous changes were made to a 22 codon region of gene F. Interestingly, over the course of 35-50 serial transfers, Kula et al. (2018) observed substitutions replacing deoptimized codons with more optimal codons in these small recoded regions. Domingo-Calap et al. (2009) also used site-directed mutagenesis to make single synonymous substitutions to φX174. None of these were lethal, but several in gene F did reduce viral fitness. Both studies suggest that codon usage is important for φX174, but neither compares effects across different regions of the genome nor do they combine multiple deoptimizing mutations. Bull et al. (2012) built increasingly deoptimized versions of the phage T7 and observed near-linear decreases in viral fitness. They deoptimized the gene encoding the T7 capsid protein, which is interesting because we also observed linear decreases for the φX174 capsid gene F. This relationship does not hold up for other φX174 genes. Janschke et al. (2019) provides the only other combinatorial data set on par with the results presented in this paper. In their study, φX174 was broken into five fragments of similar lengths. Within these fragments, synonymous changes (120 in total) were made to disrupt cryptic open reading frames (ORFs) and plaque size was measured for 30 combinations of the five mutated fragments. Plaque sizes were reduced about one third, but the effects were not additive. It is difficult to directly compare this study to ours because the goal was not codon deoptimization and changes were made to multiple genes within each fragment (except fragment 1, which only contained gene A). Janschke
et al. (2019) did observe some interesting mRNA stability effects that are corroborated by our results. This is discussed in detail below.

Speculating on the mechanisms causing reduced fitness

In recoding ΦX174 with better or worse codons, we observed correlative changes in other compositional features (Figure 6, supplementary table S2). The best predictor of ΦX174 fitness was mRNA secondary structure, which is a determinant of translation rate. Tightly folded mRNA around translational start sites reduces initiation (Kudla et al. 2009; Bentele et al. 2013; Goodman et al. 2013). Folding in other parts of mRNAs slows elongation (Kelsic et al. 2016; Peeri and Tuller 2020). Translational speed is also influenced by the availability of charged tRNAs. During elongation, ribosomes must wait for cognate tRNAs. The wait-time for codons corresponding to rare tRNAs is longer than common tRNAs. The extreme result of slowed translation is ribosomal stalling and drop-off. However, faster translation is not necessarily beneficial. mRNA structure can slow translation where pausing is needed, most notably before protein structures that require co-translational folding (Zhang et al. 2009; Yu et al. 2015; Faure et al. 2016) and at the 5' end of a transcript, where proper loading of mRNAs onto the ribosome foreshadows correct and efficient translation (Tuller et al. 2010). Mutations that affect the folding stability and codon usage biases of bacteriophage genes tend to cause fitness effects. In the Jaschke et al. (2019) study on cryptic ORFs in ΦX174, a single point mutation was identified in gene H that repeatedly evolved and ablates the detrimental effects caused by the genome-editing that they performed. An investigation of this mutation revealed that it brings the mRNA folding stability of the mutated H mRNA closer to wild-type and increases H protein expression in the mutated strain of ΦX174. Jack et al. (2017) found that deoptimization of T7’s capsid gene (gene 10) reduced protein (but not mRNA) expression. The deoptimization of gene 10 was expected to slow translational elongation because of tRNA limitations (Jack et al. 2017; Jack et al. 2019). Measuring mRNA and protein expression in our deoptimized and optimized strains would provide valuable insight on the role of codon bias and mRNA stability on protein expression. Our data support the importance of mRNA folding stability for organismal fitness, as the stability of recoded ΦX174 genes is correlated with ΦX174 fitness (Figure 6b). Interestingly, the optimized ΦX174 strains almost always have increased mRNA folding stabilities and increased GC content, while folding stability and GC content are uncoupled in the deoptimized strains (folding stability decreases but GC does not). Since we did not change the 5’ end of recoded genes (confirmed by checking the folding stability of a -4 to +37 window of recoded genes), we suggest that changes to initiation rates are
minimal and that the mechanisms driving decreased fitness are potentially different between optimized and deoptimized viruses. Codon deoptimization (and the resulting decrease in mRNA stability) should result in faster elongation, perhaps altering protein expression for other genes or the ratio of protein expression across the genome (see Frumkin et al. 2018 for a model of global translation). Codon optimization (and the resulting increase in mRNA stability) should result in slower elongation and decreased protein production of the targeted gene. However, the changes resulting from recoding are clearly multifaceted and require investigation aimed at understanding the mechanisms causing fitness declines. Since none of the indices that we calculated fully explain the observed variance in fitness, we expect there to be other important features (e.g., cotranslational folding, ssDNA packaging, etc.) in the genome that we have not considered here.

Combining mutations and epistasis

Genomes accumulate deleterious mutations over time. The detrimental effect of accumulating deleterious mutations is prevented by sex, recombination, and purifying selection which purge them from populations. In contrast to beneficial mutations which generally combine with diminishing returns (Chou et al. 2011; Couce and Tenaillon 2015), the way that the individual effects of deleterious mutations combine is less well understood. Among many issues preventing these predictions is a paucity of empirical phenotype data for networks of deleterious mutations (West et al. 1998; Kouyos et al. 2007; de Visser et al. 2011). This is especially true for combinations beyond two. A number of studies have investigated epistasis among pairs or triple sets of deleterious mutations, but the findings are mixed (Elena and Lenski 1997; Sanjuán et al. 2004; Segrè et al. 2005). Sometimes the combined effect is the sum of the individual effects (additive/no epistasis), sometimes it is less than predicted from the individual effects (antagonistic/positive epistasis) (Jasnos and Korona 2007; Guerrero et al. 2017), and sometimes it is more than predicted (synergistic/negative epistasis) (Parera et al. 2009). Among these three scenarios, antagonistic epistasis seems to be most common (Wang et al. 2002; Kouyos et al. 2007). If one considers sign epistasis to be an extreme form of antagonistic epistasis, then more support is garnished for this model as a number of studies on deleterious mutations uncover some degree of deleterious mutations becoming beneficial in combination (Lalić and Elena 2012). Our data is novel in that it builds several complete combinatorial networks of deleterious mutations, but it does lack large sample sizes. Of the networks we built, only the one for gene A had a sufficient number of data points to reject poorly fitting models. For gene A, strong antagonistic epistasis was observed. Johnson et al. (2019) recently found that this type of epistasis is common
among loss-of-function mutations in yeast. They called it “increasing cost epistasis” because a given deleterious mutation tends to have a greater cost on more fit backgrounds (Johnson et al. 2019). For genes F and H, no models can be rejected, but the data suggest that mutations in gene F are additive while mutations in gene H combine synergistically. For the purposes of building synonymously recoded viruses for vaccines, it is promising to see gene A displaying antagonistic epistasis. With this type of epistasis where fitness flattens out, less trial and error should be required to build attenuated, but still viable, viruses.

**Synonymous virus genome recoding for vaccines**

Synonymously recoding viral genomes has a potentially useful application in making live-attenuated vaccines. The antigenicity of synonymously attenuated viruses is maintained because the viral protein sequences remain unchanged. However, the process of choosing how many codons to change and what type of synonymous changes to implement is currently done without guiding principles. In fact, which synonymous features most strongly affect recoded viruses is debated (Futcher et al. 2015; Shen et al. 2015; Kunec and Osterrieder 2016). Of the dozens of viruses that have been deoptimized, a minority of them measure compositional features different from the one being directly targeted for deoptimization. At the very least, we suggest that researchers must measure a variety of compositional features when designing deoptimized constructs. A better approach would be to develop construct design software that supports researchers to engineer deoptimized viral genes (see Jorge et al. 2015) for an example using codon shuffling. This software exists for optimizing genes for expression in host cells (Grote et al. 2005; Chin et al. 2014) and may be co-opted for deoptimization purposes. In our experiments we made no effort to isolate changes to one type of compositional feature. In exploring this possibility, we found it difficult to generate sufficient deoptimization of one feature (CAI) while keeping other features (mfold, CpB, Shine-Dalgarno frequency) unchanged. Recently, Paff et al. demonstrated that promoter ablation attenuated T7 bacteriophage in a predictable manner (Paff et al. 2018). Combining these edits with previous codon deoptimized strains showed increased attenuation. Targeting intragenic attenuating mutations is a promising way to test how deleterious effects combine without the added complication of trying to isolate correlative compositional features.

Like many other studies, our data showed virus codon deoptimization is an effective way to generate attenuated viruses. In cell culture and animal studies, deoptimized viruses were shown to protect from viral challenge and were stable over small numbers of passages (Martínez et al. 2016). However, much concern remains about the potential for attenuated viruses to recover
virulence. It is therefore important to understand how many and what types of synonymous mutations can be made to viral genomes without completely ablating their ability to replicate in host cells. What viral genes should be attenuated? How many attenuating mutations should be made to the genome? What synonymous features should be targeted for deoptimization? In most studies to date, a limited number of deoptimized constructs (usually structural proteins) were tested. We showed that fitness decreases can be obtained by deoptimizing many of the ΦX174 genes, indicating that nonstructural genes may also be good targets for attenuation. One approach to avoid evolutionary reversion might be recoding multiple genes or entire viral genomes, balancing optimization and deoptimization to maintain sufficient virulence while increasing the genetic distance to wild type. This strategy could prevent recovery by mutation or by recombination with wild-type viruses. However, our work suggests that the effects of recoding will not be uniform across a genome. We found that the attenuating effects of recoding and the nature of epistatic interactions from combining fragments differ dramatically between genes.

Acknowledgments
This study was supported by the National Institutes of Health grants P20 GM104420 and GM076040. We thank the Institute for Bioinformatic and Evolutionary Studies (IBEST) Genomics Resources Core for assistance with DNA sequencing. We appreciated incredibly helpful suggestions from four anonymous reviewers and Ignacio Bravo.

Methods and Methods

*Bacterial cultures and phage stocks*

A laboratory strain of bacteriophage ΦX174 (GenBank accession number AF176034) was used in this study. All experiments were carried out using *E. coli* C (strain WG5, accession number CP024090) as a host in modified Luria-Bertani media (10 g/L tryptone, 5 g/L Bacto yeast extract, 10 g/L NaCl, 2 mM CaCl₂).

*Synthetic ΦX174 genomes*

The phage assembly platform for ΦX174 was used following (Faber et al. 2020). The ΦX174 chromosome was divided into 14 genomic fragments. Each segment is flanked by unique four nucleotide overlaps of WT ΦX174 sequence so that they can be amplified from the ancestral
ΦX174 using PCR primers designed to incorporate terminal BsmB1 restriction sites. Amplicons were cloned into pCR2.1 using the Invitrogen TOPO TA cloning system (Life Technologies, Grand Island, NY). We pooled plasmid DNA containing all 14 of the phage DNA fragments in equimolar amounts and digested them with BsmB1 (Fermentas Fast Digest, Life Technologies, Grand Island, NY) for 15-30 minutes at 37°C. The digested plasmids were subjected to agarose gel electrophoresis for 15 minutes using a 1.2% agarose gel to separate the vector from the inserts. The inserts were excised from the gel, purified using the GeneJET gel extraction kit (Fermentas), ligated overnight at 14°C with T4 DNA ligase (Promega Corporation, Madison, WI), and transformed by electroporation into 100 μL competent E. coli C cells. The transformation mix was resuspended with 0.5 mL of ΦLB and plated immediately. The ΦLB was added to 3 mL of ΦLB top agar and plated onto a ΦLB agar plate. After four to five hours of incubation at 37°C, recombinant phage plaques were visible, and plates were removed from the incubator. Three plaques for each genotype were cored from the agar, suspended in 750 μL of ΦLB, and extracted off 50 μL of chloroform to kill the host cells. These stocks were used for sequencing and fitness assays. To verify that the recombinant phage contained the intended sequence, the resulting phage genome was sequenced in its entirety as previously described (Wichman et al. 2005). There was no difference in fitness between wildtype assembled phage and freezer stock phage (p=0.8, t-test, supplementary materials).

**Codon deoptimization of ΦX174**

Codon deoptimized and optimized fragments were synthesized in-house at the University of Texas at Austin, Applied Research Laboratories’ Gene Synthesis Facility or purchased from Biomatik USA, LLC (Wilmington, DE) according to the codon usage of five representative E. coli genomes (E. coli 536, E. coli UT 189, E. coli O157:H7 str. Sakai, E. coli O157:H7EDL933, and E. coli CFT073). Codon usage was calculated by averaging each codon’s usage frequency in CDS of these E. coli genomes. These hosts were chosen before the lab strain was sequenced and identified as WG5. WG5 has the same most and least-commonly used codons and would have resulted in the same recoding. Wild-type ΦX174 codons that could be changed to a more or less frequently used codon were exchanged for the most or least commonly used codon according to the host. Regions that we were unable to modify included overlapping genes on different reading frames, promoter regions that occurred within other reading frames, two codons at each fragment junction that contain the BsmBI sticky ends, and the region from 4299-4328 which encodes the ΦX174 origin of replication. In addition, 6 bases in front of the initiation codon (AUG) and the first
21 bases of each gene were left unmodified to assure efficient translation initiation. See supplementary materials for exact coordinates of recoded regions. Unsuccessful attempts to create live virus using synthesized fragments were repeated at least three times, then passaged in liquid culture for 24 hours to allow for recovery mutants to arise.

Measuring viral fitness

Fitness assays and fitness calculations were performed as previously described (Wichman et al. 2005). The assay is a determination of growth rate at low MOI in 10 mL ΦLB and is carried out at 37°C. Host cells were prepared by growing to ~10^8 cells/mL and aliquoted into 8.5 mL of warm LB just prior to adding phage. Phage from chloroform stocks were added at a concentration of 103-104 per mL. Phage fitness is expressed as the log2-fold increase in the total number of phage per hour (doublings per hour). All measurements were done in triplicate. At 40 minutes, virus titers were determined on ΦLB-agar plates with 0.7% top agar. Assembled wildtype phage were used as controls.

Calculating genome statistics

Calculations of genome statistics were done as follows: Codon Adaptation Index (CAI) was calculated using the seqinr cai() function in R which uses the Sharp and Li (1987) method and the E. coli codon usage table. An alternative form of CAI was also calculated using an updated method outlined in Xia (2007). The number of effective codons (Nc) was calculated following Wright (1990). The Index of translational elongation (I_{TE}) was calculated according to Xia (2015). The Starvation Codon Adaptation Index (sCAI) was calculated according to Elf et al. (2003). This measure scores genes by how susceptible their codons are to a scarcity of amino-acylated tRNAs. The Xia (2007) CAI and Nc indices were calculated using Puigbò et al. (2008) with the codon table from E. coli WG5. tAI was calculated according to dos Reis et al. (2004) using the tAI R package (github.com/mariodosreis/tai). COUSIN18 values were calculated according to Bourret et al. (2019) using the online server at http://cousin.ird.fr/. The strength of Shine-Dalgarno sequences were included in the model by calculating the per-codon average binding strength of all Shine-Dalgarno motifs in a gene (sum of binding strengths over gene length). An empirically-derived measure of translational speed was calculated according to Chevance et al. (2014) and normalized by codon family. Folding stabilities were calculated for entire gene transcripts using
mfold v3.6 Mathews et al. (1999) with default parameters. Similarly, we calculated the folding stability of 42 bases (-4 to +37) around the initiation site according to Kudla et al. (2009) of each recoded gene to ensure that stability effects that might affect the initiation of translation were negligible. For figure 2, all protein coding sequences were parsed from the *E. coli* WG5 genome (CP024090) and CAI was calculated as described above. The list of the most highly expressed genes is from Karlin et al. (2001). We compared individual (one independent variable) linear models (glm(fitness~deltaMetric, link="identity", family=gaussian)) using the model.sel() function in the MuMIn R package (reported in Supplementary table S2) and global models using the stepAIC(direction='both') function in the MASS R package (reported in the Results section). The p-values were adjusted using p.adjust(method="fdr", n=number of indices).

**Analysis of Epistasis**

We analyzed the network fitness data from genes A, F, and H using the Stickbreaker R package (Miller et al. 2017) and functions therein. This package fits such data to the additive, multiplicative, and stickbreaking models. While the additive and multiplicative models assume a mutation (a recoded block in this context) changes background fitness by a difference or a factor, respectively, the stickbreaking model assumes a mutation’s effect is scaled by the distance between the background and a fitness boundary. For fitting the stickbreaking model, we could not obtain reasonable estimates for the fitness boundary from the data (beneficial mutations are much more useful for estimating the boundary than deleterious ones). Instead, we assumed a fitness boundary of 24.5 dbl/hr doublings per hour (wildtype has fitness of 20.5); using a larger fitness boundary simply makes the stickbreaking model more like the additive model. Relative fit (posterior probabilities) was calculated following the methods in Miller et al. (2017). To estimate the absolute goodness of fit, we used parametric bootstrap. Specifically, for each gene and each model, we extracted the observed effect of each block on each background it appeared on. For each recoded fragment, we then regressed the background’s fitness against the fitness effect by fitting a simple linear model and obtained a p-value associated with a slope of zero (illustrated for gene A in Figure 5a). When a model is correct, the slope of this line is expected to be zero. For a given gene and model, we take the sum of the logs of the p-values, $P_{obs}$, as a summary statistic. We noticed that the data points involved in these regressions were not independent and, as such, the p-values were not valid. We accounted for this by simulating 10,000 datasets (using the estimated coefficient of each block and the estimated Gaussian noise parameter, that captures both experimental noise and variation from model expectations). For each simulated dataset, we
repeated the regression for each block and combined across blocks to obtain a summary $P_{\text{sim}}$. Across 10,000 simulations, this generated the approximate distribution of $P$ when the model is correct. We then located $P_{\text{obs}}$ in this distribution and calculated the p-value as the proportion of simulations where $P_{\text{sim}} < P_{\text{obs}}$. (Zhao et al. 2012; Jack et al. 2017)

**Data Availability**

Sequences for all recoded $\Phi X174$ genes are available on Genbank under accession numbers MN045299-MN045346.

**References**

Agashe D, Sane M, Phalnikar K, Diwan GD, Habibullah A, Martinez-Gomez NC, Sahasrabuddhe V, Polacheck W, Wang J, Chubiz LM, et al. 2016. Large-Effect Beneficial Synonymous Mutations Mediate Rapid and Parallel Adaptation in a Bacterium. *Mol. Biol. Evol.* 33:1542–1553.

Andersson SG, Kurland CG. 1990. Codon preferences in free-living microorganisms. *Microbiol. Rev.* 54:198–210.

Aragonès L, Guix S, Ribes E, Bosch A, Pintó RM. 2010. Fine-Tuning Translation Kinetics Selection as the Driving Force of Codon Usage Bias in the Hepatitis A Virus Capsid. *PLOS Pathog.* 6:e1000797.

Atkinson NJ, Witteveldt J, Evans DJ, Simmonds P. 2014. The influence of CpG and UpA dinucleotide frequencies on RNA virus replication and characterization of the innate cellular pathways underlying virus attenuation and enhanced replication. *Nucleic Acids Res.* 42:4527–4545.

Bailey SF, Hinz A, Kassen R. 2014. Adaptive synonymous mutations in an experimentally evolved Pseudomonas fluorescens population. *Nat. Commun.* 5:4076.

Bentele K, Saffert P, Rauscher R, Ignatova Z, Blüthgen N. 2013. Efficient translation initiation dictates codon usage at gene start. *Mol. Syst. Biol.* 9:675.

Boël G, Letso R, Neely H, Price WN, Wong K-H, Su M, Luff JD, Valecha M, Everett JK, Acton TB, et al. 2016. Codon influence on protein expression in E. coli correlates with mRNA levels. *Nature* 529:358–363.

Bourret J, Alizon S, Bravo IG. 2019. COUSIN (COdon Usage Similarity INdex): A Normalized Measure of Codon Usage Preferences. *Genome Biol. Evol.* 11:3523–3528.

Bull JJ, Molineux IJ, Wilke CO. 2012. Slow Fitness Recovery in a Codon-Modified Viral Genome. *Mol. Biol. Evol.* 29:2997–3004.

Bulmer M. 1991. The selection-mutation-drift theory of synonymous codon usage. *Genetics* 129:897–907.

Burrkhart DH, Rouskin S, Zhang Y, Li G-W, Weissman JS, Gross CA. 2017. Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. *eLife* 6:e22037.

Burns CC, Campagnoli R, Shaw J, Vincent A, Jorba J, Kew O. 2009. Genetic Inactivation of Poliovirus Infectivity by Increasing the Frequencies of CpG and UpA Dinucleotides within and across Synonymous Capsid Region Codons. *J. Virol.* 83:9957–9969.

Burns CC, Shaw J, Campagnoli R, Jorba J, Vincent A, Quay J, Kew O. 2006. Modulation of Poliovirus Replicative Fitness in HeLa Cells by Deoptimization of Synonymous Codon
Usage in the Capsid Region. *J. Virol.* 80:3259–3272.

Cáceres EF, Hurst LD. 2013. The evolution, impact and properties of exonic splice enhancers. *Genome Biol.* 14:R143.

Cann AJ, Stanway G, Hughes PJ, Minor PD, Evans DM, Schild GC, Almond JW. 1984. Reversion to neurovirulence of the live-attenuated Sabin type 3 oral poliovirus vaccine. *Nucleic Acids Res.* 12:7787–7792.

Cannarozzi G, Schraudolph NN, Faty M, von Rohr P, Friberg MT, Roth AC, Gonnet G, Barral Y. 2010. A Role for Codon Order in Translation Dynamics. *Cell* 141:355–367.

Carbone A. 2008. Codon Bias is a Major Factor Explaining Phage Evolution in Translationally Biased Hosts. *J. Mol. Evol.* 66:210–223.

Cardinale DJ, DeRosa K, Duffy S. 2013. Base Composition and Translational Selection are Insufficient to Explain Codon Usage Bias in Plant Viruses. *Viruses* 5:162–181.

Cello J, Paul AV, Wimmer E. 2002. Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template. *Science* [Internet]. Available from: http://science.sciencemag.org/content/early/2002/07/11/science.1072266

Chen GF, Inouye M. 1990. Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. *Nucleic Acids Res.* 18:1465–1473.

Cheng BYH, Nogales A, de la Torre JC, Martínez-Sobrido L. 2017. Development of live-attenuated arenavirus vaccines based on codon deoptimization of the viral glycoprotein. *Virology* 501:35–46.

Cheng BYH, Ortiz-Riaño E, Nogales A, Torre JC de la, Martínez-Sobrido L. 2015. Development of Live-Attenuated Arenavirus Vaccines Based on Codon Deoptimization. *J. Virol.* 89:3523–3533.

Cherwa JE, Young LN, Fane BA. 2011. Uncoupling the functions of a multifunctional protein: the isolation of a DNA pilot protein mutant that affects particle morphogenesis. *Virology* 411:9–14.

Chevance FFV, Guyon SL, Hughes KT. 2014. The Effects of Codon Context on In Vivo Translation Speed. *PLOS Genet.* 10:e1004392.

Chin JX, Chung BK-S, Lee D-Y. 2014. Codon Optimization OnLine (COOL): a web-based multi-objective optimization platform for synthetic gene design. *Bioinformatics* 30:2210–2212.

Chithambaram S, Prabhakaran R, Xia X. 2014. Differential Codon Adaptation between dsDNA and ssDNA Phages in *Escherichia coli*. *Mol. Biol. Evol.* 31:1606–1617.

Chou H-H, Chiu H-C, Delaney NF, Segrè D, Marx CJ. 2011. Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. *Science* 332:1190–1192.

Cladel NM, Budgeon LR, Hu J, Balogh KK, Christensen ND. 2013. Synonymous codon changes in the oncogenes of the cottontail rabbit papillomavirus lead to increased oncogenicity and immunogenicity of the virus. *Virology* 438:70–83.

Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S. 2008. Virus Attenuation by Genome-Scale Changes in Codon Pair Bias. *Science* 320:1784–1787.

Couce A, Tenaillon OA. 2015. The role of declining adaptability in microbial evolution experiments. *Front. Genet.* [Internet] 6. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2015.00099/full

Elena SF, Lenski RE. 1997. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* 390:395–398.

Elf J, Nilsson D, Tenson T, Ehrenberg M. 2003. Selective Charging of tRNA Isoacceptors Explains Patterns of Codon Usage. *Science* 300:1718–1722.

Faber MS, Van Leuven JT, Ederer MM, Sapožnikov Y, Wilson ZL, Wichman HA, Whitehead TA, Miller CR. 2020. Saturation Mutagenesis Genome Engineering of Infective ΦX174 Bacteriophage via Unamplified Oligo Pools and Golden Gate Assembly. *ACS Synth. Biol.* 9:125–131.
Fabritus L de, Nougairède A, Aubry F, Gould EA, Lamballerie X de. 2015. Attenuation of Tick-Borne Encephalitis Virus Using Large-Scale Random Codon Re-encoding. PLOS Pathog 11:e1004738.

Fabritus L de, Nougairède A, Aubry F, Gould EA, Lamballerie X de. 2016. Utilisation of ISA Reverse Genetics and Large-Scale Random Codon Re-Encoding to Produce Attenuated Strains of Tick-Borne Encephalitis Virus within Days. PLOS ONE 11:e0159564.

Famulare M, Chang S, Iber J, Zhao K, Adeniji JA, Bukbuk D, Baba M, Behrend M, Burns CC, Oberste MS. 2015. Sabin Vaccine Reversion in the Field: a Comprehensive Analysis of Sabin-Like Poliovirus isolates in Nigeria. J. Virol. 90:317–331.

Faure G, Ogurtsov AY, Shabalina SA, Koonin EV. 2016. Role of mRNA structure in the control of protein folding. Nucleic Acids Res. 44:10898–10911.

Fox JM, Erill I. 2010. Relative Codon Adaptation: A Generic Codon Bias Index for Prediction of Gene Expression. DNA Res. 17:185–196.

Fros JJ, Dietrich I, Alshaikhahmed K, Passchier TC, Evans DJ, Simmonds P. 2017. CpG and UpA dinucleotides in both coding and non-coding regions of echovirus 7 inhibit replication initiation post-entry. eLife 6:e29112.

Frumkin I, Lajoie MJ, Gregg CJ, Hornung G, Church GM, Pilpel Y. 2018. Codon usage of highly expressed genes affects proteome-wide translation efficiency. Proc. Natl. Acad. Sci. 115:E4940–E4949.

Futcher B, Gorbatsevych O, Shen SH, Stauff CB, Song Y, Wang B, Leatherwood J, Gardin J, Yurovsky A, Mueller S, et al. 2015. Reply to Simmonds et al.: Codon pair and dinucleotide bias have not been functionally distinguished. Proc. Natl. Acad. Sci. 112:E3635–E3636.

Gamble CE, Brule CE, Dean KM, Fields S, Grayhack EJ. 2016. Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast. Cell 166:679–690.

Gaunt E, Wise HM, Zhang H, Lee LN, Atkinson NJ, Nicol MQ, Highton AJ, Klenerman P, Beard PM, Dutia BM, et al. 2016. Elevation of CpG frequencies in influenza A genome attenuates pathogenicity but enhances host response to infection. eLife 5:e12735.

Giallonardo FD, Schlub TE, Shi M, Holmes EC. 2017. Dinucleotide composition in animal RNA viruses is shaped more by virus family than host species. J. Virol.:JVI.02381-16.

Goodman DB, Church GM, Kosuri S. 2013. Causes and Effects of N-Terminal Codon Bias in Bacterial Genes. Science 342:475–479.

Granatham R, Gautier C, Gouy M, Mercier R, Pavé A. 1980. Codon catalog usage and the genome hypothesis. Nucleic Acids Res. 8:r49–r62.

Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, Jahn D. 2005. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res. 33:W526–W531.

Guerrero RF, Muir CD, Josway S, Moyle LC. 2017. Pervasive antagonistic interactions among hybrid incompatibility loci. PLOS Genet. 13:e1006817.

Gutman GA, Hatfield GW. 1989. Nonrandom utilization of codon pairs in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 86:3699–3703.

Hershberg R, Petrov DA. 2008. Selection on codon bias. Annu. Rev. Genet. 42:287–299.

Ikemura T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. 2:13–34.

Irwin B, Heck JD, Hatfield GW. 1995. Codon Pair Utilization Biases Influence Translational Elongation Step Times. J. Biol. Chem. 270:22801–22806.

Jack BR, Boutz DR, Paff ML, Smith BL, Bull JJ, Wilke CO. 2017. Reduced Protein Expression in a Virus Attenuated by Codon Deoptimization. G3 Genes Genomes Genet. 7:2957–2968.

Jack BR, Boutz DR, Paff ML, Smith BL, Wilke CO. 2019. Transcript degradation and codon usage regulate gene expression in a lytic phage. Virus Evol. [Internet] 5. Available from: https://academic.oup.com/ve/article/5/2/vez055/5692930
Jaschke PR, Dotson GA, Hung KS, Liu D, Endy D. 2019. Definitive demonstration by synthesis of genome annotation completeness. *PNAS*, 116, 24206–24213.

Jasnos L, Korona R. 2007. Epistatic buffering of fitness loss in yeast double deletion strains. *Nat. Genet.* 39:550–554.

Johnson MS, Martsuol A, Kryazhimskiy S, Desai MM. 2019. Higher-fitness yeast genotypes are less robust to deleterious mutations. *Science* 366:490–493.

Jorba J, Diop OM, Iber J, Sutter RW, Wassilak SG, Burns CC. 2016. Update on Vaccine-Derived Polioviruses — Worldwide, January 2015–May 2016. *MMWR Morb. Mortal. Wkly. Rep.* 65:763–769.

Jorge DM de M, Mills RE, Lauring AS. 2015. CodonShuffle: a tool for generating and analyzing synonymously mutated sequences. *Virus Evol.* [Internet] 1. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5014483/

Karlin S, Burge C, Campbell AM. 1992. Statistical analyses of counts and distributions of restriction sites in DNA sequences. *Nucleic Acids Res.* 20:1363–1370.

Knöppel A, Näsvall J, Andersson DI. 2016. Compensating the Fitness Costs of Synonymous Mutations. *Mol. Biol. Evol.* 33:1461–1477.

Kouyos RD, Silander OK, Bonhoeffer S. 2007. Epistasis between deleterious mutations and the evolution of recombination. *Trends Ecol. Evol.* 22:308–315.

Kristofich J, Morgenthaler AB, Kinney WR, Ebmeier CC, Snyder DJ, Old WM, Cooper VS, Copley SD. 2018. Synonymous mutations make dramatic contributions to fitness when growth is limited by a weak-link enzyme. *PLOS Genet.* 14:e1007615.

Kudla G, Murray AW, Tollervey D, Plotkin JB. 2009. Coding-Sequence Determinants of Gene Expression in Escherichia coli. *Science* 324:255–258.

Kula A, Saelens J, Cox J, Schubert AM, Travisano M, Putonti C. 2018. The Evolution of Molecular Compatibility between Bacteriophage ΦX174 and its Host. *Sci. Rep.* 8:8350.

Kunec D, Osterrieder N. 2016. Codon Pair Bias Is a Direct Consequence of Dinucleotide Bias. *Cell Rep.* 14:55–67.

Kunisawa T. 1992. Synonymous codon preferences in bacteriophage T4: A distinctive use of transfer RNAs from T4 and from its host Escherichia coli. *J. Theor. Biol.* 159:287–298.

Kunisawa T, Kanaya S, Kutter E. 1998. Comparison of synonymous codon distribution patterns of bacteriophage and host genomes. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* 5:319–326.

Lalić J, Elena SF. 2012. Magnitude and sign epistasis among deleterious mutations in a positive-sense plant RNA virus. *Heredity* 109:71–77.

Lawrie DS, Messer PW, Hershberg R, Petrov DA. 2013. Strong Purifying Selection at Synonymous Sites in D. melanogaster. *PLOS Genet.* 9:e1003527.

Le Nouën CL, Brock LG, Luongo C, McCarty T, Yang L, Mehedi M, Wimmer E, Mueller S, Collins PL, Buchholz UJ, et al. 2014. Attenuation of human respiratory syncytial virus by genome-scale codon-pair deoptimization. *Proc. Natl. Acad. Sci.* 111:13169–13174.

Levin BR. 1993. The accessory genetic elements of bacteria: existence conditions and (co)evolution. *Curr. Opin. Genet. Dev.* 3:849–854.

Li G-W, Oh E, Weissman JS. 2012. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* 484:538–541.
Liang B, Ngwuta JO, Surman S, Kabatova B, Liu Xiang, Lingemann M, Liu Xueqiao, Yang L, Herbert R, Swerczek J, et al. 2017. Improved Prefusion Stability, Optimized Codon Usage, and Augmented Virion Packaging Enhance the Immunogenicity of Respiratory Syncytial Virus Fusion Protein in a Vectored-Vaccine Candidate. *J. Virol.* 91:e00189-17.

Lim K, Lang T, Lam V, Yin J. 2006. Model-Based Design of Growth-Attenuated Viruses. *PLOS Comput. Biol.* 2:e116.

Logel DY and Jaschke PR. 2020. A high-resolution map of bacteriophage ϕX174 transcription. *Virology*, 547, 47–56.

Long H, Sung W, Kucukyildirim S, Williams E, Miller SF, Guo W, Patterson C, Gregory C, Strauss C, Stone C, et al. 2018. Evolutionary determinants of genome-wide nucleotide composition. *Nat. Ecol. Evol.*:1.

Luan S, Pan W, Li T, Yang H, Zhang B, Li F, Chen L. 2009. Rescued influenza A virus with codon deoptimized NS1 gene is attenuated both in vitro and in vivo. *Chin. J. Biotechnol.* 25:720–726.

Lucks JB, Nelson DR, Kudla GR, Plotkin JB. 2008. Genome Landscapes and Bacteriophage Codon Usage. *PLoS Comput. Biol.* [Internet] 4. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2266997/

Machado HE, Lawrie DS, Petrov DA. 2017. Strong purifying selection on codon usage bias. *bioRxiv*:106476.

Martínez MA, Jordan-Paiz A, Franco S, Nevot M. 2016. Synonymous Virus Genome Recoding as a Tool to Impact Viral Fitness. *Trends Microbiol.* 24:134–147.

Martrus G, Nevot M, Andres C, Clotet B, Martínez MA. 2013. Changes in codon-pair bias of human immunodeficiency virus type 1 have profound effects on virus replication in cell culture. *Retrovirology* 10:78.

Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure11Edited by I. Tinoco. *J. Mol. Biol.* 288:911–940.

Meng J, Lee S, Hotard AL, Moore ML. 2014. Refining the Balance of Attenuation and Immunogenicity of Respiratory Syncytial Virus by Targeted Codon Deoptimization of Virulence Genes. *mBio* 5:e01704-14.

Miller CR, Van Leuven JT, Wichman HA, Joyce P. 2017. Selecting among three basic fitness landscape models: Additive, multiplicative and stickbreaking. *Theor. Popul. Biol.* [Internet]. Available from: http://www.sciencedirect.com/science/article/pii/S0040580917300424

Minor PD. 2015. Live attenuated vaccines: Historical successes and current challenges. *Virology* 479–480:379–392.

Mioduser O, Goz E, Tuller T. 2017. Significant differences in terms of codon usage bias between bacteriophage early and late genes: a comparative genomics analysis. *BMC Genomics* [Internet] 18. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5683454/

Moratorio G, Henningsson R, Barbezange C, Carrau L, Borderia AV, Blanc H, Beaucourt S, Poirier EZ, Vallet T, Boussier J, et al. 2017. Attenuation of RNA viruses by redirecting their evolution in sequence space. *Nat. Microbiol.* 2:17088.

Mueller S, Coleman JR, Papamichail D, Ward CB, Nimnual A, Futcher B, Skiena S, Wimmer E. 2010. Live attenuated influenza virus vaccines by computer-aided rational design. *Nat. Biotechnol.* 28:723–726.

Mueller S, Papamichail D, Coleman JR, Skiena S, Wimmer E. 2006. Reduction of the Rate of Poliovirus Protein Synthesis through Large-Scale Codon Deoptimization Causes Attenuation of Viral Virulence by Lowering Specific Infectivity. *J. Virol.* 80:9687–9696.

Nabel GJ. 2013. Designing Tomorrow’s Vaccines. *N. Engl. J. Med.* 368:551–560.

Nagel AC, Joyce P, Wichman HA, Miller CR. 2012. Stickbreaking: A Novel Fitness Landscape
Model That Harbors Epistasis and Is Consistent with Commonly Observed Patterns of Adaptive Evolution. *Genetics* 190:655–667.

Newman ZR, Young JM, Ingolia NT, Barton GM. 2016. Differences in codon bias and GC content contribute to the balanced expression of TLR7 and TLR9. *Proc. Natl. Acad. Sci.* 113:E1362–E1371.

Nogales A, Baker SF, Ortiz-Riaño E, Dewhurst S, Topham DJ, Martínez-Sobrido L. 2014. Influenza A Virus Attenuation by Codon Deoptimization of the NS Gene for Vaccine Development. *J. Virol.* 88:10525–10540.

Nougairede A, Fabritus LD, Aubry F, Gould EA, Holmes EC, Lamballerie X de. 2013. Random Codon Re-encoding Induces Stable Reduction of Replicative Fitness of Chikungunya Virus in Primate and Mosquito Cells. *PLOS Pathog* 9:e1003172.

Paff ML, Jack BR, Smith BL, Bull JJ, Wilke CO. 2018. Combinatorial Approaches to Viral Attenuation. *mSystems* 3:e00046-18.

Parera M, Perez-Alvarez N, Clotet B, Martínez MA. 2009. Epistasis among Deleterious Mutations in the HIV-1 Protease. *J. Mol. Biol.* 392:243–250.

Peeri M, Tuller T. 2020. High-resolution modeling of the selection on local mRNA folding strength in coding sequences across the tree of life. *Genome Biol.* 21:63.

Pleška M, Guet CC. 2017. Effects of mutations in phage restriction sites during escape from restriction–modification. *Biol. Lett.* 13:20170646.

Plotkin JB, Kudla G. 2010. Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* 12:32–42.

Ponnala L. 2010. A Plausible Role for the Presence of Internal Shine-Dalgarno Sites. *Bioinforma. Biol. Insights* 4:55–60.

Presnyak V, Alhusaini N, Chen Y-H, Martin S, Morris N, Kline N, Olson S, Weinberg D, Baker KE, Gravelle BR, et al. 2015. Codon Optimality Is a Major Determinant of mRNA Stability. *Cell* 160:1111–1124.

Puigbò P, Bravo IG, Garcia-Vallve S. 2008. CAIcal: A combined set of tools to assess codon usage adaptation. *Biol. Direct* 3:38.

Quax TEF, Wolf YI, Koehorst JJ, Wurtzel O, van der Oost R, Ran W, Blombach F, Makarova KS, Brouns SJJ, Forster AC, et al. 2013. Differential Translation Tunes Uneven Production of Operon-Encoded Proteins. *Cell Rep.* 4:938–944.

Raghavan R, Kelkar YD, Ochman H. 2012. A selective force favoring increased G+C content in bacterial genes. *Proc. Natl. Acad. Sci.* 109:14504–14507.

Ramezanpour B, Haan I, Osterhaus A, Claassen E. 2016. Vector-based genetically modified vaccines: Exploiting Jenner’s legacy. *Vaccine* 34:6436–6448.

dos Reis M, Savva R, Wernisch L. 2004. Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res.* 32:5036–5044.

Robinson M, Lilley R, Little S, Erntage JS, Yarranton G, Stephens P, Millican A, Eaton M, Humphreys G. 1984. Codon usage can affect efficiency of translation of genes in Escherichia coli. *Nucleic Acids Res.* 12:6663–6671.

Rocha EPC, Danchin A, Viari A. 2001. Evolutionary Role of Restriction/Modification Systems as Revealed by Comparative Genome Analysis. *Genome Res.* 11:946–958.

Rostad CA, Stobart CC, Gilbert BE, Pickles RJ, Hotard AL, Meng J, Blanco JCG, Moin SM, Graham BS, Piedra PA, et al. 2016. A Recombinant Respiratory Syncytial Virus Vaccine Candidate Attenuated by a Low-Fusion F Protein Is Immunogenic and Protective against Challenge in Cotton Rats. *J. Virol.* 90:7508–7518.

Rueckert C, Guzmán CA. 2012. Vaccines: From Empirical Development to Rational Design. *PLOS Pathog.* 8:e1003001.

Sanjuán R, Moya A, Elena SF. 2004. The contribution of epistasis to the architecture of fitness in an RNA virus. *Proc. Natl. Acad. Sci. U. S. A.* 101:15376–15379.

Sau K, Gupta SK, Sau S, Ghosh TC. 2005. Synonymous codon usage bias in 16
Staphylococcus aureus phages: Implication in phage therapy. *Virus Res.* 113:123–131.
Segrè D, DeLuna A, Church GM, Kishony R. 2005. Modular epistasis in yeast metabolism. *Nat. Genet.* 37:77–83.
Sharp PM, Li WH. 1987. The codon Adaptation Index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15:1281–1295.
Shen SH, Stauf CB, Gorbatevych O, Song Y, Ward CB, Yurovsky A, Mueller S, Futcher B, Wimmer E. 2015. Large-scale recoding of an arbovirus genome to rebalance its insect versus mammalian preference. *Proc. Natl. Acad. Sci.* 112:4749–4754.
Shin YC, Bischof GF, Lauer WA, Desrosiers RC. 2015. Importance of codon usage for the temporal regulation of viral gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 112:14030–14035.
Tats A, Tenson T, Remm M. 2008. Preferred and avoided codon pairs in three domains of life. *BMC Genomics* 9:463.
Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, Pan T, Dahan O, Furman I, Pilpel Y. 2010. An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation. *Cell* 141:344–354.
Vabret N, Bailly-Bechet M, Lepelley A, Najburg V, Schwartz O, Verrier B, Tangy F. 2014. Large-Scale Nucleotide Optimization of Simian Immunodeficiency Virus Reduces Its Capacity To Stimulate Type I Interferon In Vitro. *J. Virol.* 88:4161–4172.
Velazquez-Salinas L, Risatti GR, Holinka LG, O'Donnell V, Carlson J, Alfano M, Rodriguez-Luis L, Carrillo C, Gladue DP, Borca MV. 2016. Recoding structural glycoprotein E2 in classical swine fever virus (CSFV) produces complete virus attenuation in swine and protects infected animals against disease. *Virology* 494:178–189.
Villanueva E, Marti-Solano M, Fillat C. 2016. Codon optimization of the adenoviral fiber negatively impacts structural protein expression and viral fitness. *Sci. Rep.* 6:srep27546.
de Visser JAGM, Cooper TF, Elena SF. 2011. The causes of epistasis. *Proc. R. Soc. B Biol. Sci.* 278:3617–3624.
Wang X, Minasov G, Shoichet BK. 2002. Evolution of an Antibiotic Resistance Enzyme Constrained by Stability and Activity Trade-offs. *J. Mol. Biol.* 320:85–95.
West SA, Peters AD, Barton NH. 1998. Testing for Epistasis Between Deleterious Mutations. *Genetics* 149:435–444.
Wichman HA, Millstein J, Bull JJ. 2005. Adaptive Molecular Evolution for 13,000 Phage Generations. *Genetics* 170:19–31.
Wright F. 1990. The ‘effective number of codons’ used in a gene. *Gene* 87:23–29.
Xia X. 2007. An Improved Implementation of Codon Adaptation Index. *Evol. Bioinforma. Online* 3:53–58.
Xia X. 2015. A Major Controversy in Codon-Anticodon Adaptation Resolved by a New Codon Usage Index. *Genetics* 199:573–579.
Yu C-H, Dang Y, Zhou Z, Wu C, Zhao F, Sachs MS, Liu Y. 2015. Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Mol. Cell* 59:744–754.
Zhang G, Hubalewska M, Ignatova Z. 2009. Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.* 16:274–280.
Zhao K-N, Chen J. 2011. Codon usage roles in human papillomavirus. *Rev. Med. Virol.* 21:397–411.
Zhao L, Stancik AD, Brown CJ. 2012. Differential Transcription of Bacteriophage φX174 Genes at 37°C and 42°C. *PLoS ONE* 7:e35909.

Figure legends

http://mc.manuscriptcentral.com/gbe
Fig. 1 - Invented data illustrates that epistasis affects how a desired level of attenuation is achieved. When a substantial amount of attenuation is desired (the “targeted attenuation range” is at a low fitness level), the amount of attenuation (e.g., number of deleterious mutations) will be harder to achieve if mutational effects combine synergistically (negative epistasis) because fitness declines at an increasing rate. In this case, the targeted amount of attenuation will be easier to achieve if mutations combine antagonistically (positive epistasis) because a larger range in the number of deleterious mutations results in the same level of fitness effects. Notice that this pattern is reversed if a slight level of attenuation is desired (near y=0). If sign epistasis, or even irregular magnitude epistasis is observed, then the underlying nature of interactions is more difficult to predict and generalize.

Fig. 2 - ΦX174 genome organization and capacity for deoptimization relative to host genes. (a) Genes on the ΦX174 genome are labeled at the top and shown as white boxes. Recoded regions are shown as filled blue boxes. These fragments are named consecutively (e.g., A1, A2, A3, A4, F1, F2, F3,…). Transcript expression levels are shown as filled grey bands. The band heights are proportional to the relative number of transcripts by RT-qPCR (Zhao et al. 2012). (b) Codon adaptation index (CAI) of E. coli, wild-type ΦX174, and recoded ΦX174 genes. Genes highly expressed in E. coli are enumerated on the secondary y-axis. The white space between recoded fragments (the BsmBI sites) were enlarged for visualization. Lines connect genes that were fully recoded. Genes containing only some recoded fragments (e.g., deoptimized A1) do not have black boarders.

Fig. 3 - Fitness effects of deoptimizing ΦX174 genes. (a) The fitness of wild type and deoptimized ΦX174 strains containing recoded genes are shown in replicate (colored dots). Means and standard error bars are shown in black. Fitness values that are significantly different from wild type are indicated with asterisks (ANOVA, p<0.01). (b) Fitness is plotted against measures that potentially explain fitness decreases. Fitness is the number of doublings per hour (log2 of the ratio of the phage concentration at 60 min divided by the phage concentration at time zero). Gene expression levels are from (Logel and Jaschke 2020) and are normalized to gene A. Structural proteins are shown as empty circles. Deoptimizing gene G yielded no viable phage. The total number of independent fitness measurements is provided in supplementary materials. At least 3 replicates were performed for every strain. Y-axis are the same in panels (a) and (b). The percent of codons changed and change in CAI are highly correlated (R²=0.92).
**Fig. 4 - Fitness of ΦX174 when deoptimized gene fragments are combinatorially joined.**
The fitness of variants containing one deoptimized fragment (a) and all possible within-gene variants (b) was measured and compared to wild type (grey horizontal line). Significant differences (ANOVA, p<0.01) are indicated with asterisks. In (b), fragment lengths are drawn to scale. Filled colors indicate the deoptimized fragments while unfilled blocks indicate wild-type fragments. In both (a) and (b), fitness is shown as log2-fold increase in the number of phage per hour. At least 3 replicates were performed for every strain (see supplementary materials). Y-axis are the same in panels (a) and (b).

**Fig. 5 - Effect of recoded fragment on all possible backgrounds.** (a) Regressions of each recoded fragment’s fitness effect under the additive model against background fitness for gene A. Fitness effect is the difference between the background and the background plus the recoded fragment (e.g., A2 -> A1+A2, A2+A3 -> A1+A2+A3). Horizontal lines indicate a perfect fit to the additive model with no residual effects of background. Sloped regression lines indicate antagonistic/synergistic epistasis. Solid regression lines indicate that the additive model can be rejected (linear model, p<0.05). The overall fit of epistatic models for each gene are shown in Table 2. In (b) fragment fitness effects are shown against the change in CAI. Slight point jitter was used for visualization. Linear regressions are shown with p and r-squared values. Change in CAI (Xia 2007 method) is proportional (CAI of recoded gene over CAI of background). AIC values of alternative models are shown in supplementary table S2.

**Fig. 6 - Fitness of recoded viruses correlates with codon usage bias and mRNA folding stability.** (a) Codon usage bias (CAI) compared to fitness for viruses optimized and deoptimized in genes A, F, G, and H. Fitness and CAI of wild type are indicated with grey horizontal and vertical lines. Points to the right of these lines are optimized. Points to the left are deoptimized. (b) Viral fitness compared to mRNA folding stability (mfold). Wild type values are indicated with grey horizontal and vertical lines. Less stable transcripts (mostly deoptimized genes) have less negative values and are right of wild type. More stable transcripts (most optimized genes) have more negative values. R-squared and p values shown are from individual (for each gene) linear regressions. A complete model comparing all indices with adjusted p values for multiple comparisons is shown in supplementary table S2. Codon optimized viruses are shown with empty circles. Those significantly different from wildtype are labelled (ANOVA, p<0.05).
Invented data illustrates that epistasis affects how a desired level of attenuation is achieved. When a substantial amount of attenuation is desired (the "targeted attenuation range" is at a low fitness level), the amount of attenuation (e.g., number of deleterious mutations) will be harder to achieve if mutational effects combine synergistically (negative epistasis) because fitness declines at an increasing rate. In this case, the targeted amount of attenuation will be easier to achieve if mutations combine antagonistically (positive epistasis) because a larger range in the number of deleterious mutations results in the same level of fitness effects. Notice that this pattern is reversed if a slight level of attenuation is desired (near y=0). If sign epistasis, or even irregular magnitude epistasis is observed, then the underlying nature of interactions is more difficult to predict and generalize.
ΦX174 genome organization and capacity for deoptimization relative to host genes. (a) Genes on the ΦX174 genome are labeled at the top and shown as white boxes. Recoded regions are shown as filled blue boxes. These fragments are named consecutively (e.g., A1, A2, A3, A4, F1, F2, F3,...). Transcript expression levels are shown as filled grey bands. The band heights are proportional to the relative number of transcripts by RT-qPCR (Zhao et al. 2012). (b) Codon adaptation index (CAI) of E. coli, wild-type ΦX174, and recoded ΦX174 genes. Genes highly expressed in E. coli are enumerated on the secondary y-axis. The white space between recoded fragments (the BsmBI sites) were enlarged for visualization. Lines connect genes that were fully recoded. Genes containing only some recoded fragments (e.g., deoptimized A1) do not have black boarders.

188x140mm (300 x 300 DPI)
Fitness effects of deoptimizing ΦX174 genes. (a) The fitness of wild type and deoptimized ΦX174 strains containing recoded genes are shown in replicate (colored dots). Means and standard error bars are shown in black. Fitness values that are significantly different from wild type are indicated with asterisks (ANOVA, p<0.01). (b) Fitness is plotted against measures that potentially explain fitness decreases. Fitness is the number of doublings per hour (log2 of the ratio of the phage concentration at 60 min divided by the phage concentration at time zero). Gene expression levels are from (Logel and Jaschke 2020) and are normalized to gene A. Structural proteins are shown as empty circles. Deoptimizing gene G yielded no viable phage. The total number of independent fitness measurements is provided in supplementary materials. At least 3 replicates were performed for every strain. Y-axis are the same in panels (a) and (b). The percent of codons changed and change in CAI are highly correlated (R2=0.92).
Fitness of ΦX174 when deoptimized gene fragments are combinatorially joined. The fitness of variants containing one deoptimized fragment (a) and all possible within-gene variants (b) was measured and compared to wild type (grey horizontal line). Significant differences (ANOVA, p<0.01) are indicated with asterisks. In (b), fragment lengths are drawn to scale. Filled colors indicate the deoptimized fragments while unfilled blocks indicate wild-type fragments. In both (a) and (b), fitness is shown as log2-fold increase in the number of phage per hour. At least 3 replicates were performed for every strain (see supplementary materials). Y-axis are the same in panels (a) and (b).
Effect of recoded fragment on all possible backgrounds. (a) Regressions of each recoded fragment’s fitness effect under the additive model against background fitness for gene A. Fitness effect is the difference between the background and the background plus the recoded fragment (e.g., A2 -> A1+A2, A2+A3 -> A1+A2+A3). Horizontal lines indicate a perfect fit to the additive model with no residual effects of background. Sloped regression lines indicate antagonistic/synergistic epistasis. Solid regression lines indicate that the additive model can be rejected (linear model, p<0.05). The overall fit of epistatic models for each gene are shown in Table 2. In (b) fragment fitness effects are shown against the change in CAI. Slight point jitter was used for visualization. Linear regressions are shown with p and r-squared values. Change in CAI (Xia 2007 method) is proportional (CAI of recoded gene over CAI of background). AIC values of alternative models are shown in supplementary table S2.

307x146mm (300 x 300 DPI)
Fitness of recoded viruses correlates with codon usage bias and mRNA folding stability. (a) Codon usage bias (CAI) compared to fitness for viruses optimized and deoptimized in genes A, F, G, and H. Fitness and CAI of wild type are indicated with grey horizontal and vertical lines. Points to the right of these lines are optimized. Points to the left are deoptimized. (b) Viral fitness compared to mRNA folding stability (mfold). Wild type values are indicated with grey horizontal and vertical lines. Less stable transcripts (mostly deoptimized genes) have less negative values and are right of wild type. More stable transcripts (mostly optimized genes) have more negative values. R-squared and p values shown are from individual (for each gene) linear regressions. A complete model comparing all indices with adjusted p values for multiple comparisons is shown in supplementary table S2. Codon optimized viruses are shown with empty circles. Those significantly different from wildtype are labelled (ANOVA, p<0.05).