

Patterns of epistasis in RNA viruses: a review of the evidence from vaccine design

C. L. BURCH,* P. E. TURNER† & K. A. HANLEY‡

*Department of Biology, University of North Carolina, Chapel Hill, NC, USA

†Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

‡Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

Keywords:

attenuation;
epistasis;
evolution;
mutation;
RNA virus;
vaccine.

Abstract

Epistasis results when the fitness effects of a mutation change depending on the presence or absence of other mutations in the genome. The predictions of many influential evolutionary hypotheses are determined by the existence and form of epistasis. One rich source of data on the interactions among deleterious mutations that has gone untapped by evolutionary biologists is the literature on the design of live, attenuated vaccine viruses. Rational vaccine design depends upon the measurement of individual and combined effects of deleterious mutations. In the current study, we have reviewed data from 29 vaccine-oriented studies using 14 different RNA viruses. Our analyses indicate that (1) no consistent tendency towards a particular form of epistasis exists across RNA viruses and (2) significant interactions among groups of mutations within individual viruses occur but are not common. RNA viruses are significant pathogens of human disease, and are tractable model systems for evolutionary studies – we discuss the relevance of our findings in both contexts.

Introduction

Resolving the way in which deleterious mutations interact in their effects on fitness is crucial to many problems in evolutionary biology and population genetics (Lynch & Gabriel, 1990; Kondrashov & Crow, 1991; Kondrashov, 1993; Partridge & Barton, 1993; Barton & Charlesworth, 1998; Wolf *et al.*, 2000; Kondrashov & Kondrashov, 2001). Under the multiplicative model, mutations act independently and the total fitness of a genotype equals the product of the fitness effects of its component mutations. Independent effects of deleterious mutations can be expressed graphically as log fitness that decreases linearly with increasing mutation number (Fig. 1). When mutations do not act independently, defined as epistasis, the fitness effect of a mutation can change depending on the total abundance of deleterious mutations in the genome. Epistasis changes the

relationship between log fitness and mutation number, making it curvilinear (Fig. 1). Synergistic (negative) epistasis results if deleterious mutations are more harmful together than would be expected from their separate effects, and produces accelerated fitness loss with increasing mutation number. In contrast, antagonistic (positive) epistasis produces decelerated loss of fitness as deleterious mutations accumulate.

The predictions of several prominent theories in evolutionary biology depend critically on the nature of epistasis. Synergistic epistasis allows sexual reproduction to be favoured over asexual reproduction because sex creates mutational combinations of very low fitness that are more quickly eliminated by selection (the mutational deterministic hypothesis; Kondrashov, 1988; but see Agrawal, 2001; Siller, 2001). Thus, synergistic epistasis provides a possible explanation for the widespread occurrence of sex despite its associated costs (Maynard Smith, 1978). Synergistic epistasis also can explain how outcrossing allows species with relatively small local or inbred populations (for instance, mammals and trees) to survive the accumulation of deleterious mutations through genetic drift (Peck *et al.*, 1997).

Correspondence: Christina L. Burch, Department of Biology, CB No. 3280, University of North Carolina, Chapel Hill, NC 27599, USA.
Tel.: 919-843-2691; fax: 919-962-1625;
e-mail: cburch@bio.unc.edu

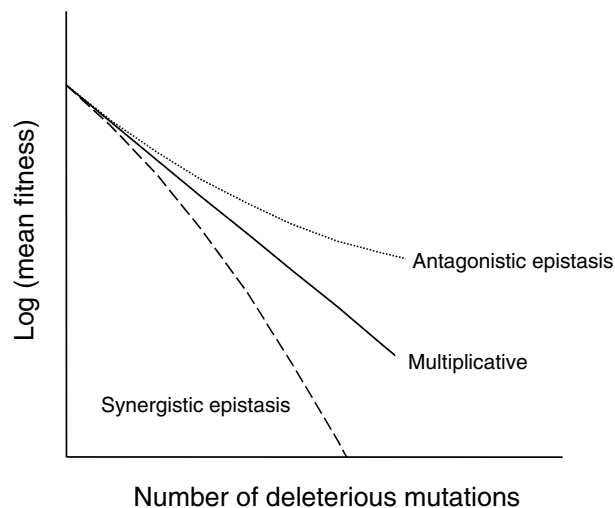


Fig. 1 Hypothetical fitness effects of increasing deleterious mutation number. Each curve is based on the general model described in the text. The solid line shows multiplicative effects, where each additional mutation causes a comparable reduction in fitness. The dashed line shows synergistic epistasis; each additional mutation causes a disproportionate fitness reduction. The dotted line represents antagonistic epistasis; each additional mutation causes a diminished reduction in fitness.

Compensatory adaptation (restoration of fitness) may occur faster under synergistic epistasis because it allows the marginal effect of each compensatory mutation to be greater (Moore *et al.*, 2000). In contrast, if epistasis is nonexistent or antagonistic, a number of the above predictions can be reversed. For example, antagonistic epistasis favours the evolution of asexual over sexual reproduction (Charlesworth, 1990; Otto, 1997). Although the existence of epistasis has been firmly established through population genetic studies (Fenster *et al.*, 1997), much less is known about the extent and form of epistatic interactions (Otto, 1997; Barton & Charlesworth, 1998; West *et al.*, 1999).

Theory has suggested mechanisms for generating both forms of epistasis, but empirical studies have yet to reveal any consistent pattern (Szathmary, 1993; Fenster *et al.*, 1997; Peck & Waxman, 2000). It is not known whether the lack of pattern is real or simply the result of experimental and statistical limitations. Limitations arise both from difficulties inherent in conducting and interpreting empirical studies of epistasis (West *et al.*, 1999), and from an inability to measure fitness in the natural environment of study organisms (Waxman & Peck, 1999). As reviewed elsewhere (West *et al.*, 1999; Peters & Keightley, 2000), studies of epistasis in most organisms have taken two general forms: (i) examining the expected distribution of fitnesses among sexually produced offspring of a pair of parents, or (ii) directly determining whether or not log fitness is a linear function of mutation

number. The first approach can determine the average form of epistasis, but does not estimate the average magnitude of epistasis between pairs of mutations, or whether the form and magnitude of epistasis varies among mutation pairs. The second approach, when pursued by examining fitness effects of known mutations alone and in combination, provides a more rigorous test of epistasis because the average magnitude and variance of epistasis can be determined.

Empirical studies of epistasis

The earliest examination of log fitness as a function of mutation number comes from a mutation accumulation experiment in *Drosophila melanogaster* (Mukai, 1969). In that experiment, competitive viability of a wild type chromosome (relative to a nonrecombining balancer chromosome) was shown to decrease faster with increasing generations of mutation accumulation, suggesting a nonlinear decline in log fitness. However, generation number provides only an indirect estimate of mutation number and allows alternative explanations for the apparent decline in viability, such as selective improvement of balancer chromosomes or nonlinear increases in mutation rates caused by instability of transposon copy number (Keightley, 1996; Nuzhdin *et al.*, 1996). More recent studies allowed mutation number to be determined directly. Elena & Lenski (1997) used transposon mutagenesis to generate 225 genotypes of *Escherichia coli* with varying numbers of mutations, and then measured genotype fitness. No overall pattern in epistasis was found; rather, some mutation combinations showed synergistic epistasis whereas others displayed antagonistic epistasis. Similarly, de Visser *et al.* (1997) created lines of *Aspergillus niger* with different combinations of known deleterious mutations and found no epistasis on average. However, the latter data cannot exclude synergistic epistasis because all the possible mutational combinations were not generated, implying that the worst combinations could have been removed by selection prior to testing.

Whitlock & Bourguet (2000) tested the action of epistasis using a set of deleterious visible mutations in *D. melanogaster*. Of two fitness components, male mating success and productivity (combination of female fecundity and offspring survivorship), only the latter showed a pattern of strong average synergistic epistasis. However, both fitness measures featured some antagonistic and some synergistic interactions between specific sets of mutations. These experiments provide strong evidence that epistasis between individual pairs of mutations is common and that the form and magnitude of epistasis varies between mutation pairs. They provide little, if any, evidence that the average form of epistasis is synergistic, and do not suggest that the form of epistasis can be generalized across organisms.

Vaccine design and interactions among deleterious mutations

One rich source of data on the interactions among deleterious mutations that has gone untapped by evolutionary biologists is the literature on the design of live, attenuated vaccine viruses. Rational vaccine design depends upon the measurement of individual and combined effects of deleterious (attenuating) mutations. Appropriate mutations can then be sequentially added to a vaccine virus to minimize reactivity and the likelihood of reversion to virulence, while maintaining satisfactory immunogenicity (Murphy & Chanock, 2001). These manipulations generally utilize modern molecular biology techniques such as reverse genetics and site-directed mutagenesis, and most have been conducted in the last decade. In addition, the rigorous requirements of vaccine development necessitate *in vivo* (naturalistic) measures of viral fitness. Thus in creating vaccines for viruses such as dengue, influenza and human immunodeficiency virus (HIV), virologists often inadvertently conduct an ideal test of epistasis. Indeed, reports of such studies have specifically noted that the effect of combined mutations could not always be predicted by the effect of single mutations (e.g. Skiadopoulos *et al.*, 1998) implicitly acknowledging the action of epistasis. Here we compile and analyse data from vaccine-related studies involving RNA viruses, to determine the occurrence and nature of epistatic mutations in these pathogens.

Methods

Review of the literature

In order to analyse the evidence for epistasis in RNA viruses, we first reviewed the biomedical literature using Medline (National Library of Medicine, National Institutes of Health, USA) to identify studies directly or indirectly related to development of live-attenuated RNA virus vaccines. Additionally, we traced relevant references included in the studies identified on Medline. We included only those studies which focused on deleterious mutations and that measured: (i) fitness of a 'wild type' virus (although in some cases this virus contained other mutations which were not manipulated during the study), (ii) fitness effects of individual mutations, groups of mutations, or segments bearing mutations in the otherwise 'wild type' background, and (iii) fitness effects of some combination of paired mutations in the same 'wild type' background. In some cases, the measurement of the effect of combined mutations preceded the measurement of the effect of individual mutations. We included studies that featured a variety of types of mutation, including substitutions and deletions in coding and noncoding regions of the genome, but we excluded studies of whole-gene deletions. Although we do not

claim to have included every relevant study on RNA viruses, we have collected a representative and unbiased sample of the literature.

Although our literature search was unbiased, the possibility exists that the mutations targeted in studies of vaccine design might comprise a biased set. The mutations in our literature sample can be grouped into three categories: (i) site-directed mutations in genes of known function, (ii) site-directed mutations in conserved regions of the viral genome, and (iii) mutations acquired during selection experiments, such as adaptation of viruses to cold temperature or to nonhuman tissue culture. Studies of mutations in categories (i) and (ii) rarely had an *a priori* expectation as to whether the mutations would be of large effect, and in fact many were of minor effect (Brown *et al.*, 1999). In category (iii), the studies often sought to identify mutations responsible for attenuation in humans and, therefore, screened mutations of both major and minor effect. Even so, it seems likely that our literature sample will contain an excess of mutations with obvious effects on viral fitness, and consequently, a deficiency of mutations with barely detectable effects. However, mutations with catastrophic effects on viral fitness will also be under-represented in our data sample because viruses that were not viable or that replicated to very low titres could not be used in the phenotypic assays presented in the compiled studies. Thus, bias in our data set is expected to be against both extremes of fitness effects, making our sample strongly representative only of mutations with intermediate effects on fitness.

We collected data from 29 studies using 14 different viruses (Appendix 1). From each study we recorded the type of each individual mutation, the fitness measures used and the fitnesses of the 'wild type' virus, viruses containing individual mutations, and viruses containing combinations of two and (in some studies) three individual mutations. A few studies reported performance measures for viruses containing four or more mutations. These data are described in their entirety in Fig. 2; however, for brevity only the first three mutations are included in Appendix 1.

Compiling fitness data

As studies of human viruses are limited to measuring performance *in vitro* or in an animal model, there is no standard method for measuring fitness. From each study, we recorded all the reported measures of viral performance. These measures included, but were not limited to, growth rate, peak titre, lethal dose, infectious dose, percent mortality in animal model, difference in titre between low and high sodium bicarbonate concentrations, difference in titre between low and high temperatures, plaque size and polymerase activity. In some studies, viral titres were measured at a permissive temperature and over a large range of higher (restrictive)

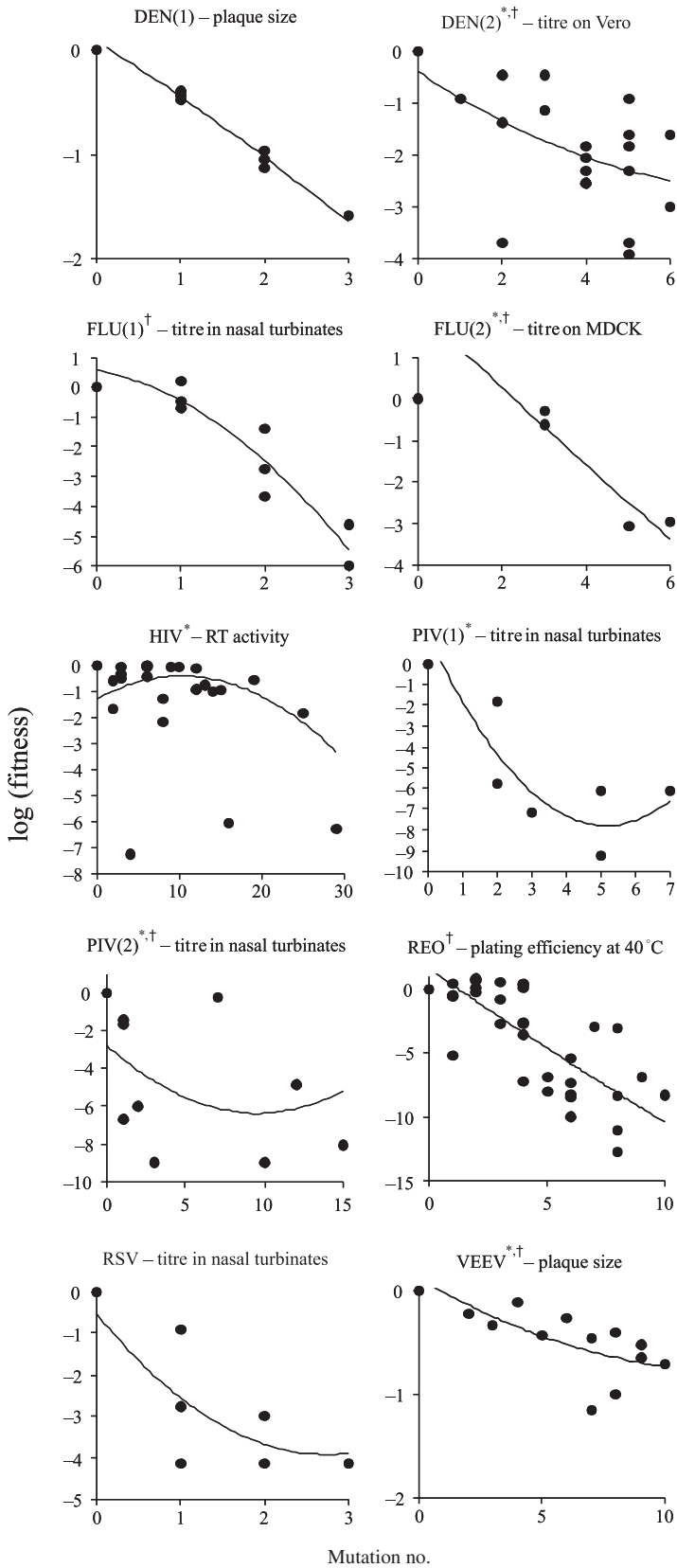


Fig. 2 Observed effects of deleterious mutation number on fitness in eight pathogenic RNA viruses. Each point represents the fitness, relative to the wild type virus, of an individual mutant carrying the specified number of mutations. The curves represent the best fitting model of the form $\ln W = \alpha k + \beta k^2$ where estimates of α and β were obtained from a bootstrap analysis. *For the data shown in this figure, it was possible to count the number of nucleotide (or amino acid) substitutions that comprised the mutations listed in Appendix 1. †There are mutations included in this figure that do not appear in Appendix 1.

temperatures. These cases provide the only exception to our reporting of complete data; rather, we report viral titre at the permissive (baseline) temperature, and the reduction in titre at one higher temperature: 39 °C if mutants showed growth at this temperature, or 38 °C otherwise.

We made an attempt to consider the relevance of the various performance measures to fitness under natural conditions. Performance measures such as replication rate and infectious dose have a very clear relationship to viral fitness. Even measures such as shut off temperature and titre reduction from low to high temperature are clearly related to fitness because they serve as an indicator for the extent to which the virus can spread to warmer areas of the body (e.g. into the warmer lower respiratory tract from the cooler upper respiratory tract) or survive elevated temperatures during host fever. However, the relevance of some performance measures, such as plaque size, to fitness under natural conditions is questionable. Therefore, measures of *in vivo* replication in an animal model were analysed separately from other performance data, as *in vivo* replication is the most widely accepted estimator of fitness under natural conditions.

For each performance measure, the natural log of fitness is expressed relative to that of the wild type virus. Log fitness of the single mutants is determined using equation 1:

$$\ln W_i = \ln(r_i/r_{wt}), \quad (1)$$

where r_{wt} and r_i are the performance measures (e.g. growth rates), respectively, of the wild type virus and of the virus containing mutation i , and W_i is the relative fitness of the virus containing mutation i . Observed log fitness values for double mutants were calculated in the same manner:

$$\ln W_{ij} = \ln(r_{ij}/r_{wt}), \quad (2)$$

where r_{ij} and W_{ij} are, respectively, the performance measure and relative fitness of the virus containing both mutations i and j .

In cases where fitness was too low to be accurately measured, i.e. not distinguishable from zero, the natural log of fitness is undefined. Therefore, we do not report log fitnesses in these cases. If zero values most often occurred in double or triple mutants, our exclusion of these data would bias against a finding of synergistic epistasis by creating a threshold below which fitness cannot fall. However, in our data set zero values occur in single mutants nearly as often as they occur in double mutants (seven vs. nine instances, respectively). In addition, the exclusion of zero values might affect our findings if it produced a tendency to exclude a particular type of performance measure. We examined whether *in vivo* performance measures were more often excluded because of zero fitness measures than were *in vitro* performance measures. Thirty percent of the *in vivo* measures (17 of 56) were excluded, compared with 19%

of the *in vitro* measures (12 of 63). Although these numbers suggest that proportionally more *in vivo* measures are lost because of zero fitness values, the difference is not statistically significant ($\chi^2 = 2.346$, d.f. = 1, $P = 0.1256$). Thus, we have no *a priori* expectation of how excluding these data should bias our findings.

Results

Statistical tests for epistasis: the sign test

To test for epistasis using the multiplicative model, we first compared the observed log fitness of viruses containing two mutations with that expected based on the effect of each mutation in isolation. Expected log fitness values for viruses bearing two mutations were generated under the assumption of no epistasis, and were calculated using equation 3 (Elena & Lenski, 1997; Wade *et al.*, 2001):

$$\ln \hat{W}_{ij} = \ln W_i + \ln W_j \quad (3)$$

Each combination of mutations yielded one comparison of observed and expected log fitnesses. For example, consider the first entry in Appendix 1 of mutations in dengue virus (Men *et al.*, 1996). The observed, or measured, log fitness of a virus containing both mutations 1 and 2 was -1.3 . The expected log fitness, determined by summing the log fitness of each of the single mutants, was -1.4 because $\ln W_1 + \ln W_2 = -0.1 - 1.3 = -1.4$. In this case the observed log fitness of -1.3 is greater than the expected log fitness of -1.4 , so the interaction between mutations 1 and 2 is antagonistic.

For studies containing combinations of more than two mutations, we analysed only nonintersecting pairs of those mutations, so that no mutation was contained in more than one analysed pair. In this manner, we ensured that the mutation pairs analysed were independent and avoided pseudo-replication. Each comparison was assigned a value of 1, 0 or -1 if the virus carrying two mutations was, respectively, more fit, equally fit (exactly the same value as expected), or less fit than expected. We had no means to distinguish performance measures that accurately reflected viral fitness from those that did not, therefore, our first pass analysis examined all the reported performance measures for each mutation pair. As the different performance measures cannot be considered independently, we used the median comparison value across all fitness measures for each virus. Thus, for each mutation pair examined, synergistic epistasis was concluded only if a majority of performance measures indicated that the observed $\ln W_{ij}$ was less than the expected $\ln \hat{W}_{ij}$, and antagonistic epistasis was concluded only if a majority of performance measures indicated the opposite (i.e. $\ln W_{ij} > \ln \hat{W}_{ij}$).

We then conducted a sign test to determine whether observed and predicted fitness differed in a consistent

direction across studies. The sign test was used instead of a more powerful parametric test because the nature of the data, collected by different researchers and using different fitness measures, ensured that measurement errors were not normally distributed. A comparison of the number of mutation pairs that yielded a higher fitness than expected to the number that yielded a lower fitness than expected revealed no significant difference [sign test, all fitness measures: $N(1) = 17$, $N(0) = 3$, $N(-1) = 11$; $P = 0.1725$]. To determine whether the results of the first analysis were influenced by the inclusion of performance measures that had little relevance to the natural environment, we also conducted a sign test using only the *in vivo* performance measures. This test also failed to reveal a significant predominance of one type of epistasis [$N(1) = 9$, $N(0) = 1$, $N(-1) = 4$; $P = 0.1334$].

Statistical tests for epistasis: regression analysis

From studies in which three or more mutations were combined into a single virus, we additionally attempted to detect epistasis by investigating the linearity of the relationship between log fitness and mutation number (Elena & Lenski, 1997; Whitlock & Bourguet, 2000). Linear regressions were performed using log fitness as the dependent variable and mutation number (1, 2, 3 mutations, etc.) as the independent variable. In this regression, log fitness is described by an equation of the form

$$\ln W_k = \alpha k + \beta k^2, \quad (4)$$

where W_k is the average fitness of genotypes with k mutations, $\alpha < 0$ for deleterious mutations, and β defines the interaction between mutations: $\beta < 0$ for synergistic interaction, $\beta > 0$ for antagonistic interaction (Elena & Lenski, 1997).

We conducted regression analyses using data from 10 studies in which viruses contained at least three deleterious mutations [DEN (two studies), FLU (two studies), HIV, PIV (two studies), REO, RSV, VEEV; Appendix 1). In most cases, the data used for regression analyses were derived directly from Appendix 1. However, in some cases the mutations described in the appendix actually consist of blocks of nucleotide or amino acid substitutions. In these cases, we used the number of nucleotide substitutions as our measure of mutation number where they were reported, and used the number of amino acid substitutions otherwise. It turned out that all the studies containing at least three deleterious mutations investigated the effects of nucleotide substitutions, and none investigated deletions, so we did not have to decide whether large deletions should be counted as single or multiple mutations. The only other deviation from the data reported in Appendix 1 resulted because some studies reported data for many more than the three mutations. Although there was not space to

include these additional mutations in Appendix 1, we included data for the additional mutations in the regression analyses.

As the mutants within these studies do not contain independent sets of mutations, use of a least squares linear regression is inappropriate. Therefore, we conducted a bootstrap analysis (Sokal & Rohlf, 1995) in which the data were resampled 1000 times, and estimates of α and β were obtained using least squares regression for each resampled data set. The bootstrap approach produces unbiased estimates of the regression parameters α and β , and their standard errors. This approach maximizes statistical power because it allows us to use all the fitness measures provided by each study, and it is robust to variance heterogeneities which exist in these data. We report here the results of regression analyses performed on only one fitness measure for each virus per study. We did conduct regression analyses using the other reported fitness measures, however, the results did not differ qualitatively from those reported here.

The α and β estimated from each of the 10 regressions are reported in Table 1. Although seven of the 10 regressions yielded positive estimates of the quadratic (interaction) term β , none of the 10 regressions yielded a β that differed significantly from zero. A power analysis (included in Table 1) demonstrates that few data sets had the power to detect interaction effects (β) smaller than 20% of the independent effects (α) of mutations. Thus, the regression analysis was not sufficiently powerful to detect weak epistasis. In sum, the results of the regression analyses parallel the results of the sign test in that they provide no conclusive evidence for directional epistasis.

Discussion

The negative results of the sign tests and the regression analyses can be interpreted in a number of ways. As is the intention, both statistical tests will yield negative results if interactions are not sufficiently common or sufficiently strong to be detected. However, interactions could be common, but equally divided between antagonism and synergism, and sign tests and regression analyses would still be unable to detect epistasis.

We are unable to differentiate between these two possibilities using the existing data. Further, we have only limited power even to suggest that the net effect of epistasis is small. The sign test is especially weak. A power analysis shows that, given 28 observations that deviate from the null expectation (as in the 'all fitness measures' data set; see Results), a sign test could reject the null hypothesis only if 68% or more of the 28 interactions were of one type (1 or -1). Thus, epistatic effects could be fairly strong and mostly of one type and still be undetectable by a sign test. Although the regression analyses also showed an inability to detect epistasis, they are more descriptive than the sign test because they allow us to estimate the net effect of epistasis. When we

Table 1 Regression analysis.

Virus	Model parameters	Best fit model	<i>P</i> -value*	Smallest detectable interaction effect (% of α)†
DEN (1)	α	$\ln W = 0.08 - 0.553k$	<0.0001	
	α & β	$\ln W = 0.083 - 0.512k - 0.022k^2$	0.3368	23.1
DEN (2)	α	$\ln W = -0.615 - 0.340k$	0.0100	
	α & β	$\ln W = -0.394 - 0.541k + 0.032k^2$	0.3522	34.9
FLU (1)	α	$\ln W = 1.352 - 2.064k$	<0.0001	
	α & β	$\ln W = 0.591 - 0.548k - 0.494k^2$	0.1513	195.8
FLU (2)	α	$\ln W = 0.963 - 0.647k$	0.0108	
	α & β	$\ln W = 2.21 - 0.987k + 0.010k^2$	0.4836	52.7
HIV	α	$\ln W = -0.361 - 0.101k$	0.0631	
	α & β	$\ln W = -1.296 + 0.117k - 0.008k^2$	0.1591	15.9
PIV (1)	α	$\ln W = -0.361 - 0.101k$	0.0631	
	α & β	$\ln W = 1.359 - 3.553k + 0.345k^2$	0.1238	18.8
PIV (2)	α	$\ln W = -3.193 - 0.305k$	0.1721	
	α & β	$\ln W = -2.768 - 0.755k + 0.040k^2$	0.3845	40.0
REO	α	$\ln W = 1.401 - 1.184k$	<0.0001	
	α & β	$\ln W = 1.612 - 1.282k + 0.008k^2$	0.4530	12.0
RSV	α	$\ln W = -1.036 - 1.256k$	0.0098	
	α & β	$\ln W = -0.531 - 2.469k + 0.449k^2$	0.2179	52.3
VEEV	α	$\ln W = -0.031 - 0.076k$	<0.0001	
	α & β	$\ln W = 0.117 - 0.138k + 0.005k^2$	0.3353	20.2

**P*-values are associated with addition of the highest order term to the model. Thus, for models containing only α , $P < 0.05$ denote a significant fit of the linear parameter α . For models containing both α and β , *P*-values denote the fit of the quadratic parameter β .

†To determine the power of our analyses, we calculated the smallest detectable interaction effect by dividing the 95% confidence interval for β by α . Thus, the smallest detectable interaction effect is reported as a percentage of the independent (linear) effect of mutations. Any interaction effect weaker than that reported here will not achieve significance in this analysis.

consider the magnitude of β (the interaction effect) relative to the magnitude of α (the independent effect), the median value of β/α is 0.059. Thus, the net effect of epistasis is small, only 5.9% of the independent effect of mutations. This number is in agreement with findings of other studies of RNA viruses, in which measures of epistatic effects ranged between 5 and 9% (Elena, 1999; Crotty *et al.*, 2001). Detection of epistatic effects this small in magnitude would require both the use of numerous mutation combinations, and an extremely accurate measure of fitness. Few, if any, of the studies in our data set meet both these criteria.

Although we could detect no overall trend, it is clear that there are interactions between particular pairs of mutations within our data set. Some studies reported both mean values and standard errors for each of their performance measures. In these cases it is possible to determine the existence of a significant interaction effect on a case by case basis, for example, the first combination of mutations examined in one of the DEN data sets (Hanley *et al.*, 2002). This mutation combination yielded an observed log fitness of -0.46 ± 0.21 and an expected log fitness of -4.14 ± 0.48 (titre in mouse brain). The large difference between these values is statistically

significant ($t_{20} = 7.01$, $P < 0.0001$) and indicates antagonistic epistasis. Likewise, the first mutation combination examined in RSV (Whitehead *et al.*, 1998) yielded an observed log fitness of -4.14 ± 0.55 and an expected log fitness of -6.91 ± 0.74 (titre in mouse nasal turbinates). These values also differ significantly ($t_{16} = 3.0$, $P = 0.0071$), indicating antagonistic epistasis. These mutation pairs are just two of the many pairs that exhibit significant interactions.

Overall, our findings are consistent with other empirical studies in which no evidence, or mixed evidence for epistasis was detected (de Visser *et al.*, 1997; Elena & Lenski, 1997; Elena, 1999; Whitlock & Bourguet, 2000). However, the limitations of the data should be considered. First, and most importantly, the mutations analysed were not chosen randomly. Many studies chose mutations in particular genes of known function, so that mutations were not distributed randomly across the genome. This nonrandom sampling of mutations would bias the results, for example, if many mutation combinations were within individual genes, and if such combinations were more likely to interact than mutations between genes. As the mutation combinations included in our study only rarely consisted of mutations

within a single gene, we have no expectation that these mutations are more likely to show interactions than a random sample. Secondly, many studies used multiple measures of fitness, and for a given pair of mutations these different measures sometimes yielded inconsistencies in terms of expected fitness. This pattern is not uncommon. Whitlock & Bourguet (2000) detected epistasis in a set of mutations affecting productivity in *Drosophila*, but the same mutations were additive in their effects on male mating success. These observations suggest that studies of epistasis should incorporate multiple measures of fitness. Thirdly, the studies of human pathogens did not measure fitness in the natural host or, where relevant, in the natural vector used for transmission. Rather, the studies used a combination of proxy measures of fitness (e.g. replication *in vitro*) and replication in animal models. However, these measures represent the best available approximation of attenuation in humans and are used to guide vaccine design. Moreover, measures of replication are measures of Darwinian fitness, albeit in an experimentally circumscribed environment. Nonetheless, the data analysed here suggest that no consistent tendency towards a particular form of epistasis between deleterious mutations exists across RNA viruses.

RNA viruses are important to biomedical scientists as pathogens, and to evolutionary biologists as tractable model systems (Morse, 1994; Turner, 2003). Our inability to find a consistent pattern of epistasis between deleterious mutations, despite the existence of interactions between particular mutation pairs, has implications for both medicine and evolutionary biology. From a biomedical perspective, evolutionary models can be used to inform vaccine design regardless of the average form of epistasis, as long as interactions of both forms exist. Vaccine design differs from other applications because researchers can choose the particular mutations used to create the vaccine virus. Thus, evolutionary theory can be used to predict the consequences of choosing to incorporate a greater number of antagonistically or synergistically interacting mutations into the vaccine virus. Specifically, evolutionary models can inform vaccine design by predicting the effects of epistasis on the persistence of vaccine-derived viruses and their reversion to virulence. That such reversion is a significant threat was highlighted by the recent cases of paralytic polio in Haiti and the Dominican Republic. These infections were caused by vaccine-derived virulent poliovirus that evolved after vaccination strategies failed to confine the attenuated virus (CDC, 2000).

Incorporation of mutations with antagonistic interactions into candidate vaccines could reduce persistence and reversion to virulence in a number of ways. Recombination or co-infection with other components of a multiple-component vaccine, with endogenous virus, or with a circulating wild type virus could contribute to persistence of the vaccine virus genome

and to reversion to virulence. However, the use of antagonistically interacting mutations would reduce these risks because the fitness improvement resulting from recombination or co-infection is smaller when interactions between attenuating mutations are antagonistic (Kondrashov, 1988). Reversion to virulence can also occur through mutations that compensate for the effects of attenuating mutations. The use of antagonistically interacting mutations will reduce this risk as well, because compensatory mutations have been shown to yield smaller fitness improvements when interactions between attenuating mutations are antagonistic (Moore *et al.*, 2000). Finally, the use of antagonistically interacting mutations in vaccine design makes intuitive sense, because fitness effects diminish with mutation number, allowing more mutations to be incorporated into the viral genome, thereby safeguarding attenuation (Murphy & Chanock, 2001). Thus, further studies of epistasis in RNA viruses will not only clarify evolutionary patterns but also generate relevant information for the creation of vaccines. Unlike vaccine design, applications in evolutionary biology (such as species conservation) require knowledge of the general form of epistasis rather than just the form of epistasis between selected mutation pairs. In conservation biology, epistasis plays a role in determining the rate of fitness loss via genetic drift in small populations (Kondrashov, 1994; Schultz & Lynch, 1997). It is difficult to accurately determine the form of epistasis in most endangered populations because they are typically composed of organisms featuring long generation times; rather, several attempts have been made to determine the form of epistasis in model systems. Unfortunately, the absence of a predictable pattern for epistasis in tractable models (e.g. this study; Elena & Lenski, 1997, de Visser *et al.*, 1997) makes it difficult for these studies to inform conservation policy and other applications. More generally, evolutionary biologists should reconsider hypotheses that assume any overall form of epistasis; at least among RNA viruses there is little support for such assumptions.

Acknowledgments

We thank Christian Mandl, Thomas Kinney and Alan Engelman for kindly providing numerical data sets that served as the basis for published graphical figures. We thank Timothy Wright for careful review of the manuscript. P.E.T. acknowledges financial support from the US National Science Foundation.

References

- Agrawal, A.F. 2001. Sexual selection and the maintenance of sexual reproduction. *Nature* **411**: 692–695.
- Barton, N.H. & Charlesworth, B. 1998. Why sex and recombination? *Science* **281**: 1986–1990.

- Basler, C.F., Garcia-Sastre, A. & Palese, P. 1999. Mutation of neuraminidase cysteine residues yields temperature-sensitive influenza viruses. *J. Virol.* **73**: 8095–8103.
- Blaney, J.E. Jr, Johnson, D.H., Firestone, C.Y., Hanson, C.T., Murphy, B.R. & Whitehead, S.S. 2001. Chemical mutagenesis of dengue virus type 4 yields mutant viruses which are temperature sensitive in vero cells or human liver cells and attenuated in mice. *J. Virol.* **75**: 9731–9740.
- Boot, H.J., ter Huurne, A.A., Vastenhouw, S.A., Kant, A., Peeters, B.P. & Gielkens, A.L. 2001. Rescue of infectious bursal disease virus from mosaic full-length clones composed of serotype I and II cDNA. *Arch. Virol.* **146**: 1991–2007.
- Brown, H.E., Chen, H. & Engelman, A. 1999. Structure-based mutagenesis of the human immunodeficiency virus type 1 DNA attachment site: effects on integration and cDNA synthesis. *J. Virol.* **73**: 9011–9020.
- Butrapet, S., Huang, C.Y., Pierro, D.J., Bhamarapravati, N., Gubler, D.J. & Kinney, R.M. 2000. Attenuation markers of a candidate dengue type 2 vaccine virus, strain 16681 (PDK-53), are defined by mutations in the 5' noncoding region and nonstructural proteins 1 and 3. *J. Virol.* **74**: 3011–3019.
- CDC 2000. Outbreak of poliomyelitis – Dominican Republic and Haiti, 2000. *MMWR Morb. Mortal. Wkly. Rep.* **49**: 1094–1103.
- Charlesworth, B. 1990. Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**: 199–221.
- Cheng, X., Zhou, H., Tang, R.S., Munoz, M.G. & Jin, H. 2001. Chimeric subgroup A respiratory syncytial virus with the glycoproteins substituted by those of subgroup B and RSV without the M2–2 gene are attenuated in African green monkeys. *Virology* **283**: 59–68.
- Crotty, S., Cameron, C.E. & Andino, R. 2001. RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc. Natl Acad. Sci. USA* **98**: 6895–6900.
- de Visser, J.A.G.M., Hoekstra, R.F. & van den Ende, H. 1997. Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. *Evolution* **51**: 1499–1505.
- Ebihara, H., Yoshimatsu, K., Oginio, M., Araki, K., Ami, Y., Kariwa, H., Takashima, I., Li, D. & Arikawa, J. 2000. Pathogenicity of Hantaan virus in newborn mice: genetic reassortant study demonstrating that a single amino acid change in glycoprotein G1 is related to virulence. *J. Virol.* **74**: 9245–9255.
- Elena, S.F. 1999. Little evidence for synergism among deleterious mutations in a nonsegmented RNA virus. *J. Mol. Evol.* **49**: 703–707.
- Elena, S.F. & Lenski, R.E. 1997. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**: 395–398.
- Fenster, C.B., Galloway, L.F. & Chao, L. 1997. Epistasis and its consequences for the evolution of natural populations. *Trends Ecol. Evol.* **12**: 282–287.
- Fodor, E., Palese, P., Brownlee, G.G. & Garcia-Sastre, A. 1998. Attenuation of influenza A virus mRNA levels by promoter mutations. *J. Virol.* **72**: 6283–6290.
- Gritsun, T.S., Desai, A. & Gould, E.A. 2001. The degree of attenuation of tick-borne encephalitis virus depends on the cumulative effects of point mutations. *J. Gen. Virol.* **82**: 1667–1675.
- Hanley, K.A., Lee, J.J., Blaney, J.E.J., Murphy, B.R. & Whitehead, S.S. 2002. Paired charge-to-alanine mutagenesis of dengue virus type 4 NS5 generates mutants with temperature-sensitive, host-range and mouse-attenuation phenotypes. *J. Virol.* **76**: 525–531.
- Hazelton, P.R. & Coombs, K.M. 1995. The reovirus mutant tsA279 has temperature-sensitive lesions in the M2 and L2 genes: the M2 gene is associated with decreased viral protein production and blockade in transmembrane transport. *Virology* **207**: 46–58.
- Kanno, T., Mackay, D., Wilsden, G. & Kitching, P. 2001. Virulence of swine vesicular disease virus is determined at two amino acids in capsid protein VP1 and 2A protease. *Virus Res.* **80**: 101–107.
- Keightley, P.D. 1996. Nature of deleterious mutation load in *Drosophila*. *Genetics* **144**: 1993–1999.
- Kinney, R.M., Chang, G.-J., Tsuchiya, K.R., Sneider, J.M., Roehrig, J.T., Woodward, T.M. & Trent, D.W. 1993. Attenuation of Venezuelan equine encephalitis virus strain TC-83 is encoded by the 5'-noncoding region and the E2 envelope glycoprotein. *J. Virol.* **67**: 1269–1277.
- Kondrashov, A.S. 1988. Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**: 435–440.
- Kondrashov, A.S. 1993. Classification of hypotheses on the advantage of amphimixis. *J. Hered.* **84**: 372–387.
- Kondrashov, A.S. 1994. Muller's ratchet under epistatic selection. *Genetics* **136**: 1469–1473.
- Kondrashov, A.S. & Crow, J.F. 1991. Haploidy or diploidy: which is better? *Nature* **351**: 314–315.
- Kondrashov, F.A. & Kondrashov, A.S. 2001. Multidimensional epistasis and the disadvantage of sex. *Proc. Natl Acad. Sci. USA* **98**: 12089–12092.
- Latorre, P., Cadd, T., Itoh, M., Curran, J. & Kolakofsky, D. 1998. The various sendai virus C proteins are not functionally equivalent and exert both positive and negative effects on viral RNA accumulation during the course of infection. *J. Virol.* **72**: 5984–5993.
- Lynch, M. & Gabriel, W. 1990. Mutation load and the survival of small populations. *Evolution* **44**: 1725–1737.
- Mandl, C.M., Allison, S.L., Holzmann, H., Meixner, T. & Heinz, F.X. 2000. Attenuation of tick-borne encephalitis virus by structure-based site-specific mutagenesis of a putative flavivirus receptor site. *J. Virol.* **74**: 9601–9609.
- Maynard Smith, J. 1978. *The Evolution of Sex*. Cambridge University Press, Cambridge.
- McGoldrick, A., Macadam, A.J., Dunn, G., Rowe, A., Burlison, J., Minor, P.D., Meredith, J., Evans, D.J. & Almond, J.W. 1995. Role of mutations G-480 and C-6203 in the attenuation phenotype of Sabin type 1 poliovirus. *J. Virol.* **69**: 7601–7605.
- Men, R., Bray, M., Clark, D., Chanock, R.M. & Lai, C.-J. 1996. Dengue type 4 virus mutants containing deletions in the 3' noncoding region of the RNA genome: analysis of growth restriction in cell culture and altered viremia pattern and immunogenicity in rhesus monkeys. *J. Virol.* **70**: 3930–3937.
- Moore, F.B., Rozen, D.E. & Lenski, R.E. 2000. Pervasive compensatory adaptation in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **267**: 515–522.
- Morse, S.S., ed. 1994. *The Evolutionary Biology of Viruses*. Raven Press Ltd, New York, NY.
- Mukai, T. 1969. The genetic structure of natural populations of *D. melanogaster*. VIII. Synergistic interactions of spontaneous mutant polygenes controlling viability. *Genetics* **61**: 749–761.

- Murphy, B.R. & Chanock, R.M. (2001). Immunization against viral diseases. In: *Fields Virology* (P. M. Howley, ed.), pp. 435–467. Lippincott Williams and Wilkins, Philadelphia, PA.
- Nuzhdin, S.V., Pasyukova, E.G. & Mackay, T.F. 1996. Positive association between copia transposition rate and copy number in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **263**: 823–831.
- Omata, T., Kohara, M., Kuge, S., Komatsu, T., Abe, S., Semler, B., Kameda, A., Itoh, H., Arita, M., Wimmer, E. & Nomoto, A. 1986. Genetic analysis of the attenuation of phenotype of poliovirus type 1. *J. Virol.* **58**: 348–358.
- Otto, S.P. 1997. Unravelling gene interactions. *Nature* **390**: 343.
- Parkin, N.T., Chiu, P. & Coelingh, K. 1997. Genetically engineered live attenuated influenza A virus vaccine candidates. *J. Virol.* **71**: 2772–2778.
- Partridge, L. & Barton, N.H. 1993. Optimality, mutation and the evolution of ageing. *Nature* **362**: 305–311.
- Peck, J.R., Barreau, G. & Heath, S.C. 1997. Imperfect genes, Fisherian mutation and the evolution of sex. *Genetics* **145**: 1171–1199.
- Peck, J.R. & Waxman, D. 2000. Mutation and sex in a competitive world. *Nature* **406**: 399–404.
- Peters, A.D. & Keightley, P.D. 2000. A test for epistasis among induced mutations in *Caenorhabditis elegans*. *Genetics* **156**: 1635–1647.
- Roner, M.R., Neplouev, I., Sherry, B. & Joklik, W.K. 1997. Construction and characterization of a reovirus double temperature-sensitive mutant. *Proc. Natl Acad. Sci. USA* **94**: 6826–6830.
- Schmidt, A.C., Wenzke, D.R., McAuliffe, J.M., St Claire, M., Elkins, W.R., Murphy, B.R. & Collins, P.L. 2002. Mucosal immunization of rhesus monkeys against respiratory syncytial virus subgroups A and B and human parainfluenza virus type 3 by using a live cDNA-derived vaccine based on a host range-attenuated bovine parainfluenza virus type 3 vector backbone. *J. Virol.* **76**: 1089–1099.
- Schultz, S.T. & Lynch, M. 1997. Mutation and extinction: the role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. *Evolution* **51**: 1363–1371.
- Siller, S. 2001. Sexual selection and the maintenance of sex. *Nature* **411**: 689–692.
- Skiadopoulos, M.H., Durbin, A.P., Tatem, J.M., Wu, S.-L., Paschalis, M., Tao, T., Collins, P.R. & Murphy, B.R. 1998. Three amino acid substitutions in the L protein of the human parainfluenza virus type 3 cp45 live attenuated vaccine candidate contribute to its temperature-sensitive and attenuation phenotypes. *J. Virol.* **72**: 1762–1768.
- Skiadopoulos, M.H., Surman, S., Tatem, J.M., Paschalis, M., Wu, S.-L., Udem, S.A., Durbin, A.P., Collins, P.L. & Murphy, B.R. 1999. Identification of mutations contributing to the temperature-sensitive, cold-adapted, and attenuation phenotypes of the live-attenuated, cold-passage 45 (cp 45) human parainfluenza virus 3 candidate vaccine. *J. Virol.* **73**: 1374–1381.
- Sokal, R.R. & Rohlf, F.J. 1995. *Biometry*. W. H. Freeman and Company, New York, NY.
- Subbarao, E.K., Park, E.J., Lawson, C.M., Chen, A.Y. & Murphy, B.R. 1995. Sequential addition of temperature-sensitive missense mutations into the PB2 gene of influenza A transfectant viruses can effect an increase in temperature sensitivity and attenuation and permits the rational design of a genetically engineered live influenza A virus vaccine. *J. Virol.* **69**: 5969–5977.
- Szathmary, E. 1993. Do deleterious mutations act synergistically? Metabolic control theory provides a partial answer. *Genetics* **133**: 127–132.
- Tang, R.S., Nguyen, N., Zhou, H. & Jin, H. 2002. Clustered charge-to-alanine mutagenesis of human respiratory syncytial virus L polymerase generates temperature-sensitive viruses. *Virology* **302**: 207–216.
- Tao, T., Skiadopoulos, M.H., Davoodi, F., Riggs, J.M., Collins, P.L. & Murphy, B.R. 2000. Replacement of the ectodomains of the hemagglutinin-neuraminidase and fusion glycoproteins of recombinant parainfluenza virus type 3 (PIV3) with their counterparts from PIV2 yields attenuated PIV2 vaccine candidates. *J. Virol.* **74**: 6448–6458.
- Turner, P.E. 2003. Searching for the advantages of virus sex. *Orig. Life Evol. Biosph* **33**: 95–108.
- Wade, M.J., Winther, R.G., Agrawal, A.F. & Goodnight, C.J. 2001. Alternative definitions of epistasis: dependence and interaction. *Trends Ecol. Evol.* **16**: 498–504.
- Waxman, D. & Peck, J.R. 1999. Sex and adaptation in a changing environment. *Genetics* **153**: 1041–1053.
- West, S.A., Lively, C.M. & Read, A.F. 1999. A pluralistic approach to sex and recombination. *J. Evol. Biol.* **12**: 1003–1012.
- Whitehead, S.S., Firestone, C.-Y., Collins, P.L. & Murphy, B.R. 1998. A single nucleotide substitution in the transcription start signal of the M2 gene of respiratory syncytial virus vaccine candidate cpts 248/404 is the major determinant of the temperature-sensitive and attenuation phenotypes. *Virology* **247**: 232–239.
- Whitehead, S.S., Firestone, C.-Y., Karron, R.A., Crowe, J.E., Elkins, W.R., Collins, P.L. & Murphy, B.R. 1999. Addition of a missense mutation present in the L gene of respiratory syncytial virus (RSV) cpts 530/1030 to RSV vaccine candidate cpts 248/404 increases its attenuation and temperature sensitivity. *J. Virol.* **73**: 871–877.
- Whitlock, M.C. & Bourguet, D. 2000. Factors affecting the genetic load in *Drosophila*: synergistic epistasis and correlations among fitness components. *Evolution* **54**: 1654–1660.
- Wolf, J.D., Brodie, E.D. & Wade, M.J., eds. 2000. *Epistasis and the Evolutionary Process*. Oxford University Press, Oxford.
- Zhang, L., Marriott, K.A., Harnish, D.G. & Aronson, J.F. 2001. Reassortant analysis of guinea-pig virulence of Pichinde virus variants. *Virology* **290**: 30–38.

Received 11 March 2003; revised 12 June 2003; accepted 10 July 2003

Appendix 1

Virus* [Wild type]†	Type of mutation‡			Fitness measure	In relative fitness of virus with mutation§								
	1	2	3		1	2	3	1 + 2	2 + 3	1 + 3	1 + 2 + 3		
DEN [WtDEN4] Men <i>et al.</i> (1996)	LD	LD		Plaque size [C6/36 cells]	-0.10	-1.30	na¶	-1.30	na	na	na		
DEN [16681] Butrapet <i>et al.</i> (2000)	SS	SS	SS	% Mouse mortality	-6.14	-2.64	-0.15	undef**	-2.64	-1.54	undef		
				Peak titre [LLC-MK2 cells]	-0.36	-0.26	-0.47	0.07	-0.68	0.01	-0.81		
				Peak titre [C6/36 cells]	-6.40	-5.93	-0.54	-10.60	-4.87	-3.03	-9.81		
				Plaque size [LLC-MK2 cells]	-0.39	-0.48	-0.44	-1.04	-0.96	-1.13	-1.58		
DEN4 [rDEN4] Blaney <i>et al.</i> (2001)	SS	LD		Titre at 35 °C [Vero cells]	0.00	-2.99	na	-2.53	na	na	na		
				Titre at 38 °C/titre at 35 °C [Vero cells]	-0.92	-3.91	na	-3.68	na	na	na		
				Titre at 35 °C [Huh-7 cells]	0.00	-2.53	na	-2.53	na	na	na		
				Titre at 38 °C/titre at 35 °C [Huh-7 cells]	-0.46	-3.68	na	-2.76	na	na	na		
				Titre in mouse brain	-0.46	-6.68	na	-7.14	na	na	na		
DEN4 [rDEN4] Hanley <i>et al.</i> (2002)	SS	LS	LS	Titre at 35 °C [Vero cells]	-3.22	-6.45	-2.53	-2.30	na	-4.14	na		
				Titre at 38 °C/titre at 35 °C [Vero cells]	-1.15	0.46	-0.69	-0.69	na	undef	na		
				Titre at 35 °C [HuH-7 cells]	-2.76	-6.22	-1.84	-1.15	na	-2.76	na		
				Titre at 38 °C/titre at 35 °C [HuH-7 cells]	-1.61	-2.53	1.84	undef	na	-1.61	na		
				Titre in mouse brain	-2.99	-1.15	-1.38	-0.46	na	-5.30	na		
Flu [A/WSN/33] Basler <i>et al.</i> (1999)	SS	SS		Neuraminidase (NA) activity at 33 °C [COS-1 cells]	0.26	-1.29	na	-1.53	na	na	na		
				NA activity at 39.5 °C/NA activity at 33 °C	-0.05	-0.46	na	-0.75	na	na	na		
				NA activity at 33 °C [COS-1 cells]	-3.77	-3.73	na	undef	na	na	na		
				NA activity at 39.5 °C/NA activity at 33 °C	-2.04	-2.48	na	undef	na	na	na		
Flu [A/WSN/33] Fodor <i>et al.</i> (1998)	SS	SS		Titre at 40 h p.i. [MDCK cells]	-0.23	-5.76	na	-8.06	na	na	na		
FLU [A/LA/2/87 AA wtPB2] Subbarao <i>et al.</i> (1995)	SS	SS	SS	Shut off temperature [MDCK cells]	-0.03	-0.03	-0.03	-0.05	-0.03	nm††	-0.08		
				Titre at 32 °C [MDCK cells]	-0.46	-0.92	-1.15	-0.92	-1.38	nm	-2.76		
				Titre at 39 °C/titre at 32 °C [MDCK cells]	-2.53	-2.99	-5.76	-8.98	undef	nm	undef		
				Peak titre in hamster nasal turbinates	-0.46	-0.69	0.23	-3.68	-1.38	nm	-4.61		
				% Hamsters with detectable virus in turbinates	0.02	0.02	0.02	-0.05	0.02	nm	-0.17		
				Peak titre in hamster lungs	-1.38	-1.84	-0.92	undef	undef	nm	undef		
				% Hamsters with detectable virus in lungs	-0.63	-0.34	-0.15	undef	undef	nm	undef		
FLU [LA wt] Parkin <i>et al.</i> (1997)	LS	LS		Titre at 34 °C [MDCK cells]	-0.28	-0.62	na	-2.95	na	na	na		
				Titre at 39 °C/titre at 34 °C [MDCK cells]	-2.00	-1.63	na	-10.48	na	na	na		
				Titre in mouse nasal turbinates	-0.46	-1.13	na	-5.66	na	na	na		
				Titre in mouse lungs	-4.79	undef	na	undef	na	na	na		
HN+A101 [clone 1] Ebihara <i>et al.</i> (2000)	SS	SS		% Mortality in newborn mice (s.c. inoculation)	-0.04	-1.62	na	undef	na	na	na		
				1/mean survival time (i.c. inoculation)	-0.13	-0.16	na	-0.20	na	na	na		
				Virus titre at day 2 p.i. [MBMEC cells]	-0.69	-1.38	na	-0.46	na	na	na		
				Virus titre at day 8 p.i. [MBMEC cells]	0.23	-1.15	na	-1.61	na	na	na		
HIV [pNL43/Xma1] Brown <i>et al.</i> (1999)	SS	SS		Peak RT activity [Jurkat cells]	-0.58	-1.68	na	-7.27	na	na	na		
				Peak RT activity [Jurkat cells]	-0.34	-0.50	na	-0.01	na	na	na		
			SS††	Peak RT activity [Jurkat cells]	-0.05	0.07	na	-0.43	na	na	na		
			LS††	Peak RT activity [Jurkat cells]	-0.03	-0.05	na	-0.52	na	na	na		
			LS††	Peak RT activity [Jurkat cells]	-0.02	-0.02	na	-0.94	na	na	na		
			LS††	Peak RT activity [Jurkat cells]	-1.31	-2.16	na	-6.02	na	na	na		
			LS	Peak RT activity [Jurkat cells]	-0.76	-0.11	na	-1.82	na	na	na		
			LS††	Peak RT activity [Jurkat cells]	-0.96	-1.02	na	-6.28	na	na	na		
IBDV [rCEF94] Boot <i>et al.</i> (2001)	LS	LS		Mean titre after transfection [QM5 cells]	-4.61	-0.69	na	-9.21	na	na	na		
				Peak titre after 3 passages [QM5 cells]	-2.07	-0.23	na	-2.30	na	na	na		
PIC [S18L18] Zhang <i>et al.</i> (2001)	SS	SS		Days of fever [guinea-pig]	-0.32	-1.30	na	-1.70	na	na	na		
				% Original body weight at day 13 [guinea-pig]	-0.30	-0.44	na	-0.36	na	na	na		
				Titre in guinea-pig serum day 12 p.i.	-4.61	-9.44	na	-9.21	na	na	na		
				Titre in guinea-pig spleen day 12 p.i.	-4.84	-11.97	na	-11.51	na	na	na		
PIV [rPIV3] Skiadopoulos <i>et al.</i> (1998)	SS	SS	SS	Titre at 32 °C [LL-MK2 cells]	-3.45	-2.99	-3.68	-3.45	-3.91	-6.91	-4.14		
				Titre at 39 °C/titre at 32 °C [LL-MK2 cells]	-0.92	-2.07	-6.45	undef	-3.22	undef	undef		
				Titre in hamster nasal turbinates 4 days p.i.	-1.84	-5.76	-7.14	undef	-6.10	-9.21	-6.10		
				Titre in hamster lungs 4 days p.i.	-3.68	-5.41	-2.65	undef	-4.03	-8.40	undef		

Appendix 1 (continued)

Virus* [Wild type]†	Type of mutation‡			Fitness measure	ln relative fitness of virus with mutation§								
	1	2	3		1	2	3	1 + 2	2 + 3	1 + 3	1 + 2 + 3		
PIV [rPIV3] Skiadopoulos <i>et al.</i> (1999)	LS	SS		Titre at 32 °C [LL-MK2 cells]	-0.92	-0.69	na	-0.46	na	na	na		
				Titre at 39 °C/titre at 32 °C [LL-MK2 cells]	-3.45	-9.44	na	undef	na	na	na		
			Titre in hamster nasal turbinates	-0.23	-5.99	na	-4.84	na	na	na			
			Titre in hamster lungs	-0.23	-8.98	na	-8.98	na	na	na			
PIV [rPIV3] Tao <i>et al.</i> (2000)	LS	LS		Titre at 32 °C [LLC-MK2 cells]	-0.23	-2.30	na	0.23	na	na	na		
				Titre at 38 °C/titre at 32 °C [LLC-MK2 cells]	-8.52	-0.46	na	-10.59	na	na	na		
				Titre in hamster nasal turbinate day 4 p.i.	-3.22	-3.22	na	-8.06	na	na	na		
				Titre in hamster lung day 4 p.i.	-10.82	-7.83	na	-11.51	na	na	na		
PV [PV1(M)pDS306] Omata <i>et al.</i> (1986)	LS	LS		Lesion score [monkey]	-3.90	-0.20	na	-3.57	na	na	na		
				Spread value [monkey]	-3.21	0.01	na	-2.92	na	na	na		
				% Monkey paralysis	undef	undef	na	undef	na	na	na		
				Titre at high [Na(CO ₃) ₂]/titre at low [Na(CO ₃) ₂]	-9.65	-1.36	na	-8.96	na	na	na		
				Titre at 40 °C/titre at 36 °C	-8.13	-5.09	na	undef	na	na	na		
				Plaque size [monkey kidney cells]	-0.63	-0.10	na	-0.83	na	na	na		
				LS‡‡	LS‡‡	Lesion score [monkey]	-1.78	-3.12	na	-3.57	na	na	na
				LS‡‡	LS‡‡	Spread value [monkey]	-1.20	-2.15	na	-2.92	na	na	na
				% Monkey paralysis	-1.79	undef	na	undef	na	na	na		
				Titre at high [Na(CO ₃) ₂]/titre at low [Na(CO ₃) ₂]	-0.46	-10.50	na	-8.96	na	na	na		
				Titre at 40 °C/titre at 36 °C	-4.14	-9.60	na	undef	na	na	na		
				Plaque size [monkey kidney cells]	-0.17	-0.90	na	-0.83	na	na	na		
		LS‡‡	LS‡‡	Lesion score [monkey]	-1.24	-1.13	na	-3.57	na	na	na		
		LS‡‡	LS‡‡	Spread value [monkey]	-0.70	-0.61	na	-2.92	na	na	na		
				% monkey paralysis	-1.79	undef	na	undef	na	na	na		
				Titre at high [Na(CO ₃) ₂]/titre at low [Na(CO ₃) ₂]	-2.28	-7.16	na	-8.96	na	na	na		
			Titre at 40 °C/titre at 36 °C	-6.88	-10.43	na	undef	na	na	na			
			Plaque size [monkey kidney cells]	-0.13	-1.27	na	-0.83	na	na	na			
PV [S1F/480A/6203U] McGoldrick <i>et al.</i> (1995)	SS	SS		Fraction macaque paralysis	0.29	0.00	na	undef	na	na	na		
				Lesion score [macaque]	-0.09	-0.02	na	-0.40	na	na	na		
				Titre at 38 °C/titre at 35 °C [BGM cells]	-0.46	-0.23	na	-0.46	na	na	na		
REO [wt REO st3] Roner <i>et al.</i> (1997)	SS	U		Plaquing efficiency at 39 °C	-4.99	-7.17	na	-12.90	na	na	na		
				Plaquing efficiency at 30 °C [L929 cells]									
REO [T1L/tsA279 hybrids] Hazelton & Coombs (1995)	U	U		Plaquing efficiency at 39 °C [L929 cells]	-0.35	-0.33	-1.65	nm	nm	-0.21	-0.15		
				Plaquing efficiency at 40 °C [L929 cells]	-0.60	1.78	0.43	nm	nm	0.79	0.60		
RSV [rA2] Cheng <i>et al.</i> (2001)	LD	LS		Titre in cotton rat nasal turbinate	undef	undef	na	undef	na	na	na		
				Titre in cotton rat lung	-3.57	-3.75	na	undef	na	na	na		
				Virus shedding from infected monkey	-0.75	-0.75	na	-1.04	na	na	na		
				Peak titer in monkey nasopharyngeal swab	-8.01	-5.32	na	undef	na	na	na		
				Peak titre in monkey tracheal lavage fluid	-7.91	-5.35	na	-8.85	na	na	na		
				Peak titre in monkey nasopharyngeal swab	-2.76	-1.84	na	-3.91	na	na	na		
RSV [rB/HPIV3] Schmidt <i>et al.</i> (2002)	LI	LI		Peak titre in monkey nasopharyngeal swab	0.92	0.23	na	-0.23	na	na	na		
				Peak titre in monkey tracheal lavage fluid	-2.99	-3.91	na	-3.45	na	na	na		
				Peak titre in monkey nasopharyngeal swab	1.38	0.46	na	0.23	na	na	na		
				Peak titre in monkey tracheal lavage fluid	1.38	0.46	na	0.23	na	na	na		
RSV [rA2] Tang <i>et al.</i> (2002)	SS	LS	SS	Titre at 33 °C [HEp-2 cells]	0.60	0.00	-2.12	-3.50	na	na	-0.23		
				Titre at 39 °C/titre at 33 °C [HEp-2 cells]	-0.18	-0.92	undef	0.76	na	na	undef		
				Titre in cotton rat lung	nd	0.23	-7.53	nd	na	na	nd		
	LS	SS		Titre at 33 °C [HEp-2 cells]	-2.30	-0.51	na	-2.76	na	na	na		
				Titre at 39 °C/titre at 33 °C [HEp-2 cells]	0.53	undef	na	undef	na	na	na		
				Titre in cotton rat lung	nd	0.23	na	-5.78	na	na	na		
RSV [rA2CP] Whitehead <i>et al.</i> (1998)	SS	SS	SS	Titre in mouse nasal turbinates	-2.76	-4.14	-0.92	-4.14	nm	-2.99	-4.14		
				Titre in mouse lungs	-1.38	-3.45	1.38	-4.61	nm	-0.69	-4.61		
RSV [rA2CP] Whitehead <i>et al.</i> (1999)				Titre in mouse nasal turbinates	-1.40	-1.40	-1.80	undef	na	undef	na		
				Titre in mouse lungs	-0.20	-1.20	-2.30	undef	na	undef	na		
SeV [rSeV] Latorre <i>et al.</i> (1998)	SC	SC		Titre in chicken eggs (day 9)	-2.30	-2.30	na	-15.61	na	na	na		
SVDV [svVLS201M00] Kanno <i>et al.</i> (2001)	SS	SS		Fraction of pigs showing vesicles	0.00	-0.92	na	-1.61	na	na	na		
				Lesion score of pigs at 10 days p.i.	-0.24	-2.30	na	-2.48	na	na	na		

Appendix 1 (continued)

Virus* [Wild type]†	Type of mutation‡			Fitness measure	ln relative fitness of virus with mutation§							
	1	2	3		1	2	3	1 + 2	2 + 3	1 + 3	1 + 2 + 3	
				Serum titre of pigs at 14 days p.i.	-0.74	-0.83	na	0.16	na	na	na	
TBEV [wtNeudoerfl]	SS	SS		Titre at 12 h p.i. [CE cells]	-0.69	0.62	3.26	-0.36	0.30	nm	nm	
Mandl <i>et al.</i> (2000)				Titre at 21 h p.i. [CE cells]	0.65	0.25	1.08	-1.95	0.63	nm	nm	
				1/Lethal dose 50 [5-week-old mice]	-2.86	0.80	-10.36	-7.30	-4.10	nm	nm	
				1/Infectious dose 50 [5-week-old mice]	-3.32	0.00	-4.80	-3.07	-2.59	nm	nm	
TBEV [Vs-c]	SS	SS		Plaque size [PS cells]	-0.30	0.05	na	-0.23	na	na	na	
Gritsun <i>et al.</i> (2001)				CPE (% cell destruction after 48 h; PS cells)	-0.41	0.00	na	-1.10	na	na	na	
				Kill rate = 1/average mouse survival time	-0.06	0.01	na	-0.19	na	na	na	
				% Mouse mortality	-0.56	-0.15	na	-0.75	na	na	na	
VEEV [VE/IC-115]	SS	SS	LS	% Mouse mortality following i.c. challenge	nm	-1.39	0.00	-0.47	-0.69	0.00	-0.13	
Kinney <i>et al.</i> (1993)				% Mouse mortality following i.p. challenge	0.00	undef	0.00	undef	undef	-0.47	-0.98	
				Plaque size [Vero cells]	-0.32	-1.04	-0.15	-0.89	-0.53	-0.35	-0.41	

*DEN, dengue virus; FLU, influenza virus type A; HN, Hanta virus; HIV, human immunodeficiency virus; IBVD, infectious bursal disease virus; PIC, Pichinde virus; PIV, parainfluenza virus; PV, poliovirus; REO, Reovirus; RSV, respiratory syncytial virus; SeV, Sendai virus; SVDV, swine vesicular disease virus; TBEV, tick-borne encephalitis virus; VEEV, Venezuelan equine encephalitis virus.

†'Wild type virus': designation of virus strain to which mutations were added; not necessarily a true wild type.

‡Individual mutations are reported exactly as they are reported in the primary literature, thus large deletions and substitutions are often counted as a single mutation. SS, small (≤ 3 nt) substitution; SD, small (≤ 3 nt) deletion; LS: large (> 3 nt) substitution; LD, large (> 3 nt) deletion; LI, large (> 3 nt) insertion; SC, addition of stop codon; U, unknown number of changes.

§Natural log of the ratio of the fitness of the mutated virus to the fitness of the 'wild type' virus.

¶na, not applicable.

**undef, undefined because fitness of the mutated virus was not distinguishable from zero.

††nm, not measured.

‡‡These mutations are not independent of other mutation pairs in the study. Nonindependent mutation pairs were not included in the sign test, but were included in regression analyses.