Complementation and Epistasis in Viral Coinfection Dynamics

Hong Gao* and Marcus W. Feldman^{†,1}

*Department of Genetics and [†]Department of Biology, Stanford University School of Medicine, Stanford, California 94305

Manuscript received December 15, 2008

Accepted for publication March 2, 2009

ABSTRACT

Coinfection in RNA virus populations results in two important phenomena, complementation and recombination. Of the two, complementation has a strong effect on selection against deleterious mutations, as has been confirmed in earlier studies. As complementation delays the purging of less-fit mutations, coinfection may be detrimental to the evolution of a virus population. Here we employ both deterministic modeling and stochastic simulation to explore the mechanisms underlying the interactions between complementation and other evolutionary factors, namely, mutation, selection, and epistasis. We find that strong complementation reduces slightly the overall fitness of a virus population but substantially enhances its diversity and robustness, especially when interacting with selection and epistasis.

RNA viruses infect many organisms. They have high mutation rates, compact segmented genomes, very short generation times, huge population sizes, and substantial diversity within populations (CHAO 1988; WILKE and NOVELLA 2003; FROISSART *et al.* 2004; ELENA *et al.* 2006). These properties enable RNA viruses to change continuously and adapt easily to adverse environments, posing continual challenges for antiviral therapies and vaccinations. Current research on RNA viruses has turned to theoretical population genetics and evolutionary theory for novel and practical approaches for management of viral diseases (*e.g.*, TURNER and CHAO 1999, 2003; BURCH *et al.* 2003; BURCH and CHAO 2004; DENNEHY and TURNER 2004; MOYA *et al.* 2004; BRETSCHER *et al.* 2004; SANJUÁN *et al.* 2005; BONI *et al.* 2006).

Coinfection, where multiple viruses infect a cell simultaneously, is important in the life cycle of RNA viruses. During coinfection, haploid viral genomes act in a manner similar to those of diploid organisms, exchanging genetic material randomly or preferentially. Reassortment of genome segments may create severely deleterious mutational combinations, thus speeding up the elimination of mutational load. Coinfection may, however, result in complementation, where viruses carrying different deleterious mutations may benefit from the normal products that each can produce, so that both types of viruses can be represented in the offspring. In contrast to recombination, complementation weakens the selection against deleterious mutations. In this way, it contributes to

the stability of the whole virus population as it maintains a high level of diversity without sacrificing the overall fitness of the population.

FROISSART *et al.* (2004) investigated whether complementation or random reassortment would have the stronger effect on selection against deleterious mutations in viral life cycles. They proposed a simple model combining the processes of viral replication and free recombination and, via both experiments and simulation, confirmed the major role of complementation in reducing the selective pressure against deleterious mutations. However, their model assumes only two independent loci and does not accommodate such key evolutionary factors as mutation or higher-order epistasis. They simplified their model by ignoring mutation. In reality, as RNA viruses are highly mutable, a viral population does not completely eliminate deleterious mutations after a period of time, contrary to what appears to be the case in FROISSART *et al.* (2004), but reaches a dynamic balance between selection and mutation. When studying the long-term evolution of a viral population, mutation becomes a key factor that should be taken into account appropriately. Epistasis is another important evolutionary factor that affects many evolutionary processes (DESAI *et al.* 2007). Epistasis may be classified as synergistic or antagonistic, which may act on the mutation–selection balance in different ways and also interact with different levels of complementation. Epistasis can also be classified as magnitude and sign epistasis, where magnitude epistasis at a locus indicates that the magnitude of fitness change (loss or gain) of a mutation depends on the genotypic background; sign epistasis exists if the sign, as well as the magnitude, of the fitness cost of a mutation is affected by the genotypic background (WEINREICH *et al.* 2005). To date, the roles of these types of epistasis interacting

Supporting information is available online at <http://www.genetics.org/cgi/content/full/genetics.108.099796/DC1>.

¹Corresponding author: Department of Biology, Stanford University, 478 Herrin Hall, Stanford, CA 94305. E-mail: mfeldman@stanford.edu

with coinfection have not been studied systematically. In modeling viral evolution, it may be important to incorporate epistasis and, in particular, to study how the effects exerted by complementation interact with different types of epistasis.

Here we extend Froissart's model to allow for multiple loci as well as mutation and complex interactions among mutations. We investigate how complementation interacts with epistasis and how this interaction affects the evolutionary dynamics of viruses. We utilize both deterministic modeling and stochastic simulation to investigate these interactions between complementation and mutation, selection, and epistasis. To introduce mutation, we divide the process of viral replication into three stages, prereplication preparation, replication, and assembly. For each segment of the viral genome, complementation and epistasis determine the strength of selection, and these occur before replication. Mutation occurs during replication, and random reassortment take place after replication during assembly. Mutation and random reassortment can be considered as a time-reversible process as the process of replication of one segment multiple times and then randomly choosing one copy for each assembly is equivalent to randomly choosing one of the two parental viral segments and replicating that segment, and repeating these two steps multiple times. Therefore, it is reasonable to consider mutation as a separate step. For epistasis, we employ a model of interactions similar to that used by DESAI *et al.* (2007), including higher-order epistasis. Our results indicate that although it has only moderate effects on the overall mean fitness, complementation interacting with epistasis has a strong effect on the composition and diversity of the virus population.

MODEL

We first consider an RNA virus population evolving under selection, complementation, random reassortment, and epistasis. Suppose the virus is haploid with L segments. Without complementation, we assume each segment carrying one or more deleterious mutation incurs a fitness cost s and $s \ll 1$. In this setting, following the multiplicative selection model proposed by DESAI *et al.* (2007), a virus with k mutated segments has fitness e^{-ks} .

We start with Froissart's model combining complementation, selection, and reassortment and also assume that the frequency of coinfection with more than two virions per cell is negligible (OLKKONEN and BAMFORD 1989; TURNER *et al.* 1999; FROISSART *et al.* 2004). The effect of interactions among mutations is similar to that in DESAI *et al.* (2007), but with additional higher-order interactions. We assume that double mutants have fitness $e^{-2s(1+\epsilon)}$ and triple mutants have fitness $e^{-3s(1+\epsilon+\eta)}$, where ϵ is the first-order epistasis parameter and η is the

second-order epistasis coefficient. For each viral life cycle, we define F_0 to be the frequency of the wild-type viruses in the population at the beginning of this cycle and $F_{(i,j,\dots,k)}$ to be the initial frequency of viruses with mutations on segments $i, j, \dots, k$, where $1 \leq i, j, \dots, k \leq L$. As in our model mutations are deleterious, we assume four or more deleterious mutations are lethal; *i.e.*, the frequencies of haploids carrying more than three mutations are negligible; $F_{i,j,k,m} \approx 0$. We denote by F_0^* and $F_{(i,j,\dots,k)}^*$ the frequencies of wild type and mutants, respectively, after complementation, epistasis, selection, and reassortment. Denote p_c as the fraction of coinfections from the total number of infections occurring per generation. Then the wild-type genotype frequency F_0^* after the action of these four forces is

$$\begin{aligned} \bar{w}F_0^* = & (1 - p_c)F_0 \\ & + p_c \left\{ F_0^2 + \sum_i \frac{2}{2} F_0 F_i + \sum_{i,j;i < j} \frac{2}{2^2} F_0 F_{i,j} \right. \\ & + \sum_{i,j;i < j} \frac{2}{2^2} F_i F_j + \sum_{i,j,k;i < j < k} \frac{2}{2^3} F_0 F_{i,j,k} \\ & + \sum_{i,j,k;i \neq j, k < j < k} \frac{2}{2^3} F_i F_{j,k} + \sum_{i,j,k,m;i \neq j, k, m < k < m} \frac{2}{2^4} F_i F_{j,k,m} I_{\{L \geq 4\}} \\ & + \sum_{i,j,k,m;i \neq j \neq k \neq m; i < j, k < m} \frac{2}{2^4} F_i F_j F_{k,m} I_{\{L \geq 4\}} \\ & + \sum_{i,j,k,m,n;i \neq j \neq k \neq m \neq n; i < j, k < m < n} \frac{2}{2^5} F_i F_j F_{k,m,n} I_{\{L \geq 5\}} \\ & \left. + \sum_{i,j,k,l,m,n;i \neq j \neq k \neq l \neq m \neq n; i < j < k < l < m < n} \frac{2}{2^6} F_i F_j F_{k,l,m,n} I_{\{L \geq 6\}} \right\}, \end{aligned} \quad (1)$$

where $I_{\{\Phi\}}$ is an indicator function and is equal to one if Φ is true; otherwise, it is zero. The complicated formulas for frequencies F_i^* , $F_{i,j}^*$, and $F_{i,j,k}^*$ are given in the APPENDIX. Summing these frequency formulas, we obtain the mean fitness $\bar{w}$.

For viruses, mutations mostly occur during the replication processes. Selection, as a result of complementation and epistasis, occurs mainly during protein synthesis before replication. Random reassortment occurs during assembly of replicates. In our simulation, we can consider mutation and random reassortment as reversible. We can reverse the process of replication of one segment and then randomly choose one copy by randomly choosing one of the two parental viral segments and then replicating the chosen segment, since the two processes are, indeed, equivalent. Thus it is reasonable to separate mutation from the other events. (We also show in the APPENDIX that this separation step does not alter the resulting frequencies, but speeds up the calculation). Denote by F'_0 , F'_i , $F'_{i,j}$, and $F'_{i,j,k}$ the frequencies of haploids carrying zero, one, two, and three mutations, respectively, after mutation. We now assume that in each segment the wild-type allele can mutate to the deleterious allele at rate μ per individual

per generation. We further assume no back mutation occurs and ignore the terms with higher orders of μ as $\mu \ll 1$. This produces the following recursions:

$$\begin{aligned} F'_0 &= (1 - L\mu)F_0^*, \\ F'_i &= \mu F_0^* + (1 - (L - 1)\mu)F_i^*, \\ F'_{i,j} &= \mu(F_i^* + F_j^*) + (1 - (L - 2)\mu)F_{i,j}^*, \\ F'_{i,j,k} &= \mu(F_{i,j}^* + F_{j,k}^* + F_{i,k}^*) + F_{i,j,k}^*. \end{aligned} \quad (2)$$

SIMULATION

As these recursions cannot be solved analytically, we conducted both deterministic and stochastic simulations to study the stationary distribution of allele frequencies. For deterministic simulations, we iterated the recursions (1), (2), and (A1)–(A3) 5000 times (corresponding to 5000 generations) with different values of the parameters. For stochastic simulations, we simulated 100 lineages of haploid viruses, with three segments in the genome, independently over 5000 generations. For generation t in lineage l ($t = 1, 2, \dots, 5000$; $l = 1, 2, \dots, 100$), we used the following sampling procedure:

According to FROISSART *et al.* (2004), the proportions of bacteria infected by zero, one, and two viruses follow a Poisson distribution with mean equal to the multiplicity-of-infection (MOI) (ratio of phages to bacterial cells). We denote by $\text{Pr}(0)$, $\text{Pr}(1)$, and $\text{Pr}(2)$ the Poisson probabilities that a cell is infected by zero, one, or two viruses, respectively. Assuming no more than two viruses affect one cell simultaneously, the expected proportion of coinfection with two viruses from the total number of infections is $\text{Pr}(\text{Coinfection}) = \text{Pr}(2)/(1 - \text{Pr}(0)) = p_c$. Thus each cell is coinfecting with probability p_c or has a single infection with probability $\text{Pr}(\text{Infection}) = \text{Pr}(1)/(1 - \text{Pr}(0)) = 1 - p_c$. The frequency of viruses that coinfect is $f_{co} = 2 \text{Pr}(\text{Coinfection})/(2 \text{Pr}(\text{Coinfection}) + \text{Pr}(\text{Infection})) = 2p_c/(1 + p_c)$. Suppose we have N_0 viruses at the beginning of generation i and whether a virus infects or coinfects occurs as a series of Bernoulli trials with the probability of success equal to f_{co} ; *i.e.*, we sample the number of coinfecting viruses k_{co} from a binomial distribution $\text{Binom}(f_{co}, N_0)$ and need to ensure it is an even and positive number. (If an odd number is sampled, we do an extra sampling step from the uniform distribution on $[0, 1]$. If the sampled value is < 0.5 , we deduct one from the odd number; otherwise, we add one to this odd number to make it an even number. We can also discard this odd number and continue to sample until we find an even number.) Within the group of the k_{co} chosen coinfecting viruses, we randomly group them into pairs by iterating the following steps:

1. Sample an individual virus with probability $1/k_{co}$.
2. Sample the other virus of this pair with probability $1/(k_{co} - 1)$.
3. Let $k_{co} = k_{co} - 2$.
4. Repeat steps 1 and 2 until $k_{co} = 2$.

Each cell, either infected by a wild-type virus or coinfecting by two viruses that can complement each other's defects, produces approximately the same number of offspring, which can be treated as a normal random variable $m \sim N(\delta, \sigma^2)$, where δ is the expected number of offspring produced by a wild-type virus and σ^2 is its variance. If a virus carrying one, two, or three mutant segments infects a cell, the number of offspring produced follows $m \sim N(\delta e^{-s}, \sigma^2 \times e^{-2s})$, $N(\delta e^{-2s(1+\epsilon)}, \sigma^2 \times e^{-4s(1+\epsilon)})$, and $N(\delta e^{-3s(1+\epsilon+\eta)}, \sigma^2 \times e^{-6s(1+\epsilon+\eta)})$, respectively, where s is the selection coefficient, ϵ is the pairwise epistasis coefficient, and η is the second-order epistasis coefficient. Similarly, if two mutant viruses coinfect a cell with one, two, or three mutant segments in common, they will produce a number of progeny similar to the case of a single infection above.

When viruses replicate genetic material in coinfecting cells, assuming random reassortment, each virion randomly inherits a segment from one of its parents; *i.e.*, each segment is selected from one of its parents with chance 0.5.

During the replication, we assume that deleterious mutations take place at $\mu = 0.01$ per segment per virus generation. According to DRAKE and HOLLAND (1999) and DRAKE (2006), the median spontaneous mutation rate of RNA viruses is ~ 0.76 per genome per replication cycle. Thus it is reasonable to assume a deleterious mutation rate of 0.01 per segment per generation after accounting for advantageous mutations. We can sample mutations according to this probability and keep track of them. Each mutation occurs on a wild-type segment and may be inherited by offspring. All mutations suffer the same fitness loss s . If a mutation takes place on a mutant segment, we assume this will not have any additional effect on the overall fitness of a virus.

As the virus population size increases exponentially, a sampling step is necessary to keep a constant sample size that is computationally tractable, namely, 10,000. That is, if the population size is $> 10,000$, we apply a Bernoulli sampling process without replacement to generate a sample of 10,000.

RESULTS

To evaluate the effects of interactions between complementation and other factors, we ran both deterministic and stochastic simulation at five levels of coinfection (*i.e.*, $p_c \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$) under nine

TABLE 1

Parameter combinations for nine simulation scenarios

Scenario	L	μ	s	ε	η	Initial frequencies
I	3	0.0	0.1	0.0	0.0	$F_1 = F_2 = F_3 = \frac{1}{3}$
II	3	0.01	0.1	0.0	0.0	$F_1 = F_2 = F_3 = \frac{1}{3}$
III	3	0.01	0.2	0.0	0.0	$F_1 = F_2 = F_3 = \frac{1}{3}$
IV	3	0.01	0.1	0.3	0.0	$F_1 = F_2 = F_3 = \frac{1}{3}$
V	3	0.01	0.1	-0.3	0.0	$F_1 = F_2 = F_3 = \frac{1}{3}$
VI	3	0.01	0.1	0.8	0.0	$F_1 = F_2 = F_3 = \frac{1}{3}$
VII	3	0.01	0.1	-0.8	0.0	$F_1 = F_2 = F_3 = \frac{1}{3}$
VIII	3	0.01	0.1	0.1	0.8	$F_1 = F_2 = F_3 = \frac{1}{3}$
IX	3	0.01	0.1	0.1	-0.8	$F_1 = F_2 = F_3 = \frac{1}{3}$

L , the number of segments in a viral genome; μ , mutation rates per generation per segment; s , selection coefficients; ε , epistasis coefficients; η , second-order epistasis coefficients. F_1 , F_2 , and F_3 : the initial frequencies of single mutants.

genetic scenarios (see Table 1). For stochastic modeling, we averaged results across 100 lineages and assessed the variation between lineages. The results from deterministic iterations are mostly similar to those from the stochastic simulation and both are shown in the supporting information. Here we focus mainly on the stationary distribution using the stochastic simulation (as the viral system might get stuck in an unstable equilibrium in the deterministic model; see the DISCUSSION) and also show the three-way interactions among complementation, selection, and epistasis based on deterministic modeling.

Interaction of complementation with mutation: We modified the complementation model proposed by FROISSART *et al.* (2004) to take mutation dynamics into account. In reality, RNA viruses mutate at such high rates that mutations cannot be neglected even over a short time period and, in the long run, a virus population reaches a mutation–selection balance instead of completely eliminating deleterious mutations from the population. The life cycle of an RNA virus is typically composed of five steps, absorption and penetration, synthesis of viral proteins, RNA replication, assembly, and release. Complementation and selection take place during the period of viral protein synthesis, and reassortment happens just before the replication of each segment, whereas nucleotides mutate during the replication process. Therefore, it is reasonable to separate mutation from the other processes in our model.

In the absence of mutation, the virus population clears all the initial polymorphisms efficiently at all levels of coinfection, as illustrated in Figure 1, A1–A5, and in the supporting information, Figure S1 and Figure S6. With a mutation rate of 0.01 per generation per segment, mutations always exist in the population and the mean fitness is ~ 0.03 lower when the population reaches the mutation–selection equilibrium than in the no mutation case. This is reasonable; according to

classical population genetic theory the expected mean fitness of a viral population at the mutation–selection balance is $1 - L\mu = 0.97$. We observed drastic changes in the patterns of frequencies of wild type, single mutants, double mutants, and triple mutants. With no mutation, the wild type frequency is fixed at one whereas mutants' frequencies quickly drop to zero. Generally the frequencies of mutants decrease faster with low coinfection rates. When mutation occurs, the wild-type frequency increases until it reaches a limit and this stationary frequency decreases as p_c increases. The stationary wild-type frequency drops more at high levels of coinfection than at low coinfection levels. The frequencies of single mutants decrease from one to a lower level under low coinfection rates than with high coinfection, and the difference between two consecutive coinfection levels remains approximately the same irrespective of the coinfection rate. Both double and triple mutants show a similar frequency pattern with their equilibrium frequencies substantially higher with frequent coinfections than with rare coinfections.

Interaction of complementation with selection: We then investigated how complementation affects the evolutionary dynamics of virus populations when the selection pressure varies. From the mean fitness plot (see Figure 1, B1–B5, Figure S2, and Figure S7), different selection strengths drive the virus population to a similar equilibrium mean fitness, but the composition of this population varies markedly when the selection coefficient increases from 0.1 to 0.2. When selection becomes stronger, more mutants are eliminated from the population; thus the balanced frequencies of wild-type viruses generally increase for all coinfection levels and this increase is larger and more obvious at a high coinfection level. The opposite pattern is observed for single-mutant frequencies, which drop more with rare coinfection than with frequent coinfection. Both double and triple mutants show a drastic frequency drop at high coinfection levels and their frequencies remain close to zero with low coinfection rates when the selection parameter increases to 0.2. We observed that the virus system reaches equilibrium faster under stronger selection, which implies that increasing selection pressure will weaken the effect of complementation on maintaining population diversity and drive the population to purge mutations faster.

Interaction of complementation with pairwise epistasis: Synergistic epistasis entails that mutations show a stronger effect in combination than the sum of their individual effects. Antagonistic epistasis ($\varepsilon < 0$) makes this effect weaker. When the epistatic effect is moderate, e.g., $\varepsilon = \pm 0.3$ as in Figure 1, C1–C5, Figure S3, and Figure S8 [known as magnitude epistasis (WEINREICH *et al.* 2005)], the overall mean fitness is similar at different levels of coinfection, with slight differences among the cases of positive, negative, and no epistasis. Compared with the case of no epistasis, synergistic

FIGURE 1.—Interaction patterns at equilibrium under different levels of coinfection ratios, 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. (A1–A5) The interaction between complementation and mutation (μ) under scenarios I and II, with $s = 0.1$, $\varepsilon = 0.0$, $\eta = 0.0$, red for $\mu = 0.0$, and light blue for $\mu = 0.01$. (B1–B5) The interaction between complementation and selection (s) under scenarios II and III, with $\mu = 0.01$, $\varepsilon = 0.0$, $\eta = 0.0$, red for $s = 0.1$, and light blue for $s = 0.2$. (C1–C5) The interaction of complementation with magnitude epistasis (ε) under scenarios II, IV, and V, with $\mu = 0.01$, $s = 0.1$, $\eta = 0.0$, red for $\varepsilon = 0.3$, yellow for $\varepsilon = 0.0$, and light blue for $\varepsilon = -0.3$. (D1–D5) The interaction between complementation and sign epistasis (ε) under scenarios II, VI, and VII, with $\mu = 0.01$, $s = 0.1$, $\eta = 0.0$, red for $\varepsilon = 0.8$, yellow for $\varepsilon = 0.0$, and light blue for $\varepsilon = -0.8$. (E1–E5) The interaction of comple-

FIGURE 2.—(a) Trajectories of mean fitness changes under scenarios II, VI, and VII using stochastic simulation. (b, c, d, and e) Under scenarios II, VI, and VII the frequency trajectories of wild type, single mutants, double mutants, and triple mutants, respectively. Dashed lines represent the results assuming epistasis coefficient $\epsilon = 0.8$. Dotted lines represent the results assuming no epistasis. Solid lines represent the results assuming epistasis coefficient $\epsilon = -0.8$. Each color represents a coinfection fraction as labeled in the inset (a).

epistasis increases the frequency of wild type while antagonistic epistasis reduces its frequency, and their effects are generally enhanced as p_c increases. For single mutants, both types of epistasis have subtle effects in the presence of strong or weak complementation (*e.g.*, 0.9 or 0.1, respectively), but their effects become stronger at intermediate levels of coinfection (*e.g.*, ~ 0.5). Strong complementation has a much larger impact on the stationary frequencies of double and triple mutants

when interacting with both types of epistasis as illustrated in Figure 1, C4 and C5. At high p_c , the frequencies of double and triple mutants increase more with antagonistic epistasis than they are reduced with synergistic epistasis. When the epistatic effect increases dramatically ($|\epsilon| > 0.5$), say ± 0.8 , we have sign epistasis (WEINREICH *et al.* 2005); *i.e.*, double and triple mutants can have much higher or lower fitness than single mutants, resulting in remarkably variable frequency

mentation with second-order epistasis (η) under scenarios VIII and IX, with $\mu = 0.01$, $s = 0.1$, $\epsilon = 0.1$, red for $\eta = 0.8$, and light blue for $\eta = -0.8$. The first column of bar plots represents the pattern of the overall mean fitness under various scenarios, the second column shows that for wild-type frequencies, the third column shows that for single-mutant frequencies, the fourth column shows that for double-mutant frequencies, and the fifth column shows that for triple-mutant frequencies. Note that the vertical scales of the last three columns of bar plots are different and vary from 0.02 to 1.0.

trajectories before stabilizing (see Figures 1, D1–D5, and 2 and Figure S4). When mutations interact synergistically, the mean fitness equilibria shift to just above those without epistasis, and the higher the coinfection rate, the greater the mean fitness. Synergistic interaction affects the patterns of frequencies of wild type and single mutants in a manner similar to strong selection. But it has a much stronger effect in eliminating double mutants and triple mutants, as illustrated in Figure 2 where their frequencies are seen to decrease dramatically to close to zero. Antagonistic epistasis, in contrast, drives wild type and single mutants to elimination while double and triple mutants thrive in the population. Strong complementation slows down elimination of the wild type and single mutants. It is reasonable that double and triple mutants constitute the ultimate population as our simulation starts with pure single mutants, and double and triple mutants are superior to single mutants; random reassortment shuffles these segments and generates a lot more combinations of mutations that accumulate over time, while wild types and single mutants are removed. The frequencies of double mutants and triple mutants fluctuate dramatically and it takes a much longer time to reach the balanced state. Frequent coinfection tends to elevate the frequencies of triple mutants over double mutants as triple mutants are easily generated through recombination. As double and triple mutants are frequent and they have a slightly lower fitness than the wild type, the mean fitness with antagonistic epistasis is ~ 0.02 lower than that with synergistic interaction.

Interaction of complementation with higher-order epistasis: We also compared the results from scenarios VIII and IX (described in Table 1) to assess how higher-order epistasis interacts with complementation and how this interaction affects the virus population. Second-order epistasis tends to have a much milder effect on mean fitness and frequencies of wild type and single mutants (see Figure 1, E1–E5, Figure S5, and Figure S9). The difference in wild-type proportions and double-mutant proportions between positive and negative second-order epistasis becomes slightly apparent as the coinfection fraction increases to 0.9. Second-order epistasis has a fairly strong effect on the fraction of triple mutants when the fraction of coinfections is high. This is reasonable as the initial virus population is composed of single mutants, and double mutants are less fit than single mutants or wild types. Few double mutants exist, leading to a low number of triple mutants even though they are strongly favored by selection.

Three-way interaction among complementation, selection, and epistasis: To observe evolutionary trajectories under the simultaneous influence of complementation, selection, and epistasis, we iterated deterministically for 5000 generations, assuming mutation rate 0.01 and initial frequencies as in Table 1, using a series of parameters, $p_c \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, $s \in \{0.0, 0.01, \dots, 0.49, 0.50\}$, and $\varepsilon \in \{-0.50, -0.48, \dots, 0.48, 0.50\}$ (only magnitude epistasis is

considered here; *i.e.*, fitness decreases monotonically with increasing k , the number of mutant segments in a virus). From Figure S10, we observe that the overall mean fitness at equilibrium decreases as selection intensity increases and epistasis decreases, and strong complementation tends to weaken the effect of selection and synergistic epistasis on mean fitness. The equilibrium frequency of wild type increases gradually with the strength of selection, while frequent coinfection and antagonistic epistasis substantially reduce the speed at which the wild-type frequency increases, as shown in Figure 3. Figure 4 shows an intriguing frequency pattern of single mutants that starts at zero in the absence of selection against deleterious mutations, then sharply increases and attains its maximum when the selection coefficient increases to 0.05– ~ 0.1 , but gradually decreases as selection becomes stronger. Strong complementation and negative epistasis generally slow down the reduction of the single-mutant frequencies. Double mutants have an equilibrium frequency pattern similar to that of single mutants, “start low–increase–decrease,” although they disappear more quickly as the selection coefficient increases, as in Figure S11. Triple mutants take over the virus population when the selection coefficient is zero, as mutations without fitness cost continually accumulate. Then the frequencies of triple mutants decrease dramatically as stronger selection is applied (see Figure S12). Even strong complementation and antagonistic epistasis have moderate effects on selection against triple mutants. Through these three-way interaction patterns, we find that both selection and synergistic epistasis tend to eliminate mutations and combinations of mutations, whereas strong complementation and antagonistic epistasis drive the virus population in the opposite direction.

DISCUSSION

Coinfection generates two important phenomena in the viral life cycle, reassortment and complementation. It is commonly acknowledged that reassortment benefits the virus population by generating combinations of deleterious mutations, thus accelerating the removal of these mutations, whereas complementation is detrimental to the overall fitness of the group as it allows the population to harbor some deleterious mutations for long periods of time. FROISSART *et al.* (2004) showed that complementation, rather than reassortment, is a major force in the evolution of viruses. What is the real role of complementation in viral evolution that allows viruses to sacrifice higher mean fitness, especially when complementation interacts with mutation, selection, or epistasis?

Here we extend the complementation model of FROISSART *et al.* (2004) to multiple loci and incorporate the key evolutionary factors, mutation and epistasis, into the model. This more realistic model enables the assessment of the evolutionary impact of the interac-

FIGURE 3.—Wild-type frequency pattern under the three-way interaction among complementation, selection, and epistasis. (a, b, c, d, and e) The interaction patterns between selection and epistasis with the co-infection ratios 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. Colors represent the values of wild-type frequency: red = 0.00, yellow = $\frac{1}{3}$, green = $\frac{2}{3}$, and blue = 1.00.

tions between complementation and mutation, selection, and epistasis as well as higher-order epistasis. Our simulation has confirmed Froissart's result that in the absence of mutation, a high coinfection ratio delays the clearance of deleterious mutants in a virus population. We also find that with mutation, the virus population achieves almost the same overall mean fitness at equilibrium irrespective of the levels of coinfection. Scrutinizing the population composition, we observe that strong complementation lowers the wild-type frequencies and increases the frequencies of mutants. Similar results are found in the case of selection and epistasis; complementation has a moderate effect on the overall ability of the virus population to reproduce, but plays a major role in the population composition. In this way, complementation increases the diversity and stability of a virus population substantially while maintaining similar or slightly lower fitness, which in turn improves the chance of survival of a virus population when coping with adverse environments.

In our modeling of recombination, we assume implicitly that reassortment is always random during replication and segments of two viruses are randomly incorporated into offspring genomes with probability

0.5 when coinfecting a cell. These assumptions are not necessarily met in reality. Preferential associations among RNA segments have been observed in several types of RNA viruses (GRAHAM *et al.* 1987; URQUIDI and BISHOP 1992). And the abundance of offspring segments may also vary due to the lengths of virus segments, the concentration of RNA polymerase, etc. Although the effect of recombination is critical when interacting with epistasis, as shown in BRETSCHER *et al.* (2004), we focus here on the interaction between complementation and other factors, thus simplifying our model but without loss of generality. We also assume a fixed selection coefficient for all mutations in each simulation scenario, which, although not necessarily true in reality, helps to reveal the mechanism underlying the interaction between complementation and different types of epistasis.

To model sign epistasis, which has been detected in RNA viruses [*e.g.*, HIV-1 (MAMMANO *et al.* 2000)] and is mostly induced by antiviral drugs, we modify the epistasis model of DESAI *et al.* (2007) by taking the product of the number of mutations and the epistatic effect ($e^{-sk(1+\epsilon)}$) instead of raising the number of mutations to the power of the epistatic effect ($e^{-sk^{1+\epsilon}}$).

FIGURE 4.—Single-mutant frequency pattern under the three-way interaction among complementation, selection, and epistasis. (a, b, c, d, and e) The interaction patterns between selection and epistasis with the coinfection ratios, 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. Colors are as in Figure 3.

This enables us to contrast magnitude epistasis with sign epistasis.

In the deterministic simulation with sign epistasis, we observe that the frequencies of the wild type and the single mutants reach an equilibrium above zero when coinfection rates are >0.7 , as illustrated in Figure S4, while in the stochastic simulation, these frequencies decrease to zero. We conducted further simulations to investigate this phenomenon under different initial conditions and observe that initial frequencies affect the equilibrium state of the deterministic system (see Figure S13), which clearly exhibits more than one equilibrium state when $p_c > 0.7$. In the stochastic simulation, the stationary distribution with wild type absent appears to be strongly attracting, which produces the results in Figure 2. Experiments with other initial conditions also appear to result in distributions very similar to Figure 2.

We also observe that it takes a much longer time for a virus system to reach an equilibrium with antagonistic sign epistasis in the stochastic simulation than in the deterministic model. This may result from the strong fluctuations produced in the stochastic case that add variation to the frequencies of double and triple mutants.

The variance of frequencies among the 100 simulated lineages is generally quite low (data not shown), probably owing to the congruent initial frequencies for all simulation replications.

Complementation with synergistic epistasis functions differently from that with antagonistic epistasis. Synergistic epistasis tends to strengthen the selection against the combinations of less-fit mutations, while frequent coinfection generally weakens the selection against all mutants, with the result that wild types and single mutants constitute a major proportion of the virus population, and the stronger the complementation is, the lower the frequency of wild types. In the presence of antagonistic epistasis, the pattern of frequency variation becomes complex. With magnitude antagonistic epistasis, [*i.e.*, the magnitude of their effect may depend on the genetic background although the sign of the fitness effect of mutations is unconditional (WEINREICH *et al.* 2005)], wild types and single mutants are slightly advantageous over double mutants. At equilibrium, wild type, single mutants, and double mutants all constitute a major fraction of the population when coinfection is frequent. With sign epistasis, *i.e.*, both wild type and combinations of mutations are more advantageous than single mu-

tants in terms of selection (WEINREICH *et al.* 2005), at the beginning we observe a sharp increase in the frequencies of both wild type and double mutants and then a quick drop in single mutants through random reassortment and selection. As mutations accumulate over time in the population, the wild-type frequency decreases gradually until it is close to zero. The frequency of single mutants continues to drop as these are the least fit genotypes. The number of double mutants increases to a peak and decreases sharply as coinfection begins to play a major role, which generates a sharp rise in the frequency of triple mutants, although double mutants are slightly superior to triple mutants. The fluctuations in the frequencies of double and triple mutants before equilibrium are mainly due to stochasticity, which does not appear in the deterministic simulation. Higher coinfection rates lead to higher frequencies of triple mutants at stationarity, because with high p_c , triple mutants have a greater chance of sharing normal gene products and transmitting their mutant genes to offspring; thus the viral system tends to produce more triple-mutant offspring. Therefore, the overall mean fitness of the population reaches almost the same equilibrium value at all levels of coinfection with antagonistic epistasis.

When complementation, selection, and epistasis interact simultaneously, the overall mean fitness decreases as selection, *i.e.*, fitness cost, increases. Increase in the coinfection rate and decrease of the epistasis coefficient also reduce the mean fitness since both coinfection and antagonistic epistasis ($\varepsilon < 0$) result in accumulation of mutations, and double and triple mutants have lower fitness than wild type or single mutants although the effects are greater than additive. The “low–high–low” trajectories of single-mutant frequencies along the selection axis occur because when selection against deleterious mutations is very weak, *e.g.*, 0.01, triple mutants become a major part of the population. As the selection strength increases to ~ 0.1 , the fitness of single mutants is slightly less than that of wild type but the fitness cost increases exponentially with the number of mutations; thus double and triple mutants are inferior and the frequency of single mutants reaches a peak. High levels of coinfection tend to expand this stage until the selection coefficient increases to 0.2, as complementation and random reassortment can mitigate the increased fitness cost for single mutants. Strong negative epistasis retards the increase in frequency of single mutants as it reduces the fitness cost of double and triple mutants. As selection strength continues to increase, the wild type becomes much more superior to single mutants, resulting in a gradual reduction in single-mutant frequencies. The frequency pattern of double mutants is similar to that of single mutants although double mutants are more sensitive to the increase of either selection strength or synergistic epistasis. Triple mutants start at the highest frequency

with weak selection and then their frequency drops sharply as selection increases. Coinfection and strong negative epistasis tend to exert mild effects on the decrease of triple-mutant frequencies.

We anticipate that our results will have broad implications in viral disease management. Strong complementation results in weaker selection against deleterious mutations generally and may even favor mutants over wild type when complementation interacts with antagonistic epistasis. A large number of mutations persist in the population, and through random reassortment, multiple combinations of mutant segments can be generated. Both magnitude and sign antagonistic epistasis can exist (MAMMANO *et al.* 2000; WEINREICH *et al.* 2005), which substantially reduces the efficacy of antiviral drugs. Therefore, reducing the coinfection ratio is critical in controlling the spread and variation of viruses. Isolation of virus-infected patients or organisms and alternating drug usage over a period of time are commonly used to reduce the coinfection level. We also suggest using multiple antiviral drugs simultaneously instead of one at a time. As antiviral drugs are designed mainly according to the structure of viral proteins, there is a greater chance that not all viruses are eliminated in a population with high p_c when these kinds of therapies are applied, as some mutated viral molecules will not be affected by these specifically designed drugs. In addition, as viruses mutate with high frequency and have short generation times, these drug-resistant viruses will, with high probability, either mutate into fitter individuals or recombine with those carrying strongly favored genes, thus spreading their drug-resistant genes rapidly. This implies that using a single type of antiviral therapy will not work efficiently. It is more rational to design antiviral drugs for both wild-type proteins and potential mutant proteins and apply them simultaneously. In this way, the number of viruses will be reduced dramatically even if it is a stable system with strong complementation. When the density of viruses is very low, the chance of coinfection becomes minimal and the effect of complementation in reducing the selection against deleterious mutations is ameliorated.

We thank Marcel Salathé, Daniel Weissman, and Hua Tang for constructive comments on the manuscript. We are grateful to Raul Andino and Adam Luring for discussions that stimulated us to work on this problem. We also thank Yidong Lei for optimizing the C code for stochastic simulation. He is supported by the National Natural Science Foundation of China (grant 40771084). This research project is supported in part by National Institutes of Health (NIH) grant GM28016 to M.W.F. and NIH grant RO1 GM073059 to Hua Tang.

LITERATURE CITED

- BONI, M., J. GOG, V. ANDREASEN and M. M. FELDMAN, 2006 Epidemic dynamics and antigenic evolution in a single season of influenza A. *Proc. Biol. Sci.* **273**: 1307–1316.

- BRETSCHER, M., C. ALTHAUS, V. MÜLLER and S. BONHOEFFER, 2004 Recombination in HIV and the evolution of drug resistance: For better or for worse? *BioEssays* **26**: 180–188.
- BURCH, C., and L. CHAO, 2004 Epistasis and its relationship to canalization in the RNA virus *phi6*. *Genetics* **167**: 559–567.
- BURCH, C., P. TURNER and K. HANLEY, 2003 Patterns of epistasis in RNA viruses: a review of the evidence from vaccine design. *J. Evol. Biol.* **16**: 1223–1235.
- CHAO, L., 1988 Evolution of sex in RNA viruses. *J. Theor. Biol.* **133**: 99–112.
- DENNEHY, J., and P. P. TURNER, 2004 Reduced fecundity is the cost of cheating in RNA virus *phi6*. *Proc. Biol. Sci.* **271**: 2275–2282.
- DESAI, M., D. WEISSMAN and M. FELDMAN, 2007 Evolution can favor antagonistic epistasis. *Genetics* **177**: 1001–1010.
- DRAKE, J., 2006 Chaos and order in spontaneous mutation. *Genetics* **173**: 1–6.
- DRAKE, J., and J. HOLLAND, 1999 Mutation rates among RNA viruses. *Proc. Nat. Acad. Sci. USA* **96**: 13910–13913.
- ELENA, S., P. CARRASCO, J. DARÒS and R. SANJUÁN, 2006 Mechanisms of genetic robustness in RNA viruses. *EMBO Rep.* **7**: 168–173.
- FROISSART, R., C. WILKE, R. MONTVILLE, S. REMOLD, L. CHAO *et al.*, 2004 Co-infection weakens selection against epistatic mutations in RNA viruses. *Genetics* **168**: 9–19.
- GRAHAM, A., G. KUDESIA, A. M. ALLEN and U. DESSELBERGER, 1987 Reassortment of human rotavirus possessing genome rearrangements with bovine rotavirus: evidence for host cell selection. *J. Gen. Virol.* **68**: 115–122.
- MAMMANO, F., V. TROUPLIN, V. ZENNOU and F. CLAVEL, 2000 Retracing the evolutionary pathways of human immunodeficiency virus type 1 resistance to protease inhibitors: virus fitness in the absence and in the presence of drugs. *J. Virol.* **74**: 8524–8531.
- MOYA, A., E. HOLMES and F. GONZALEZ-CANDELAS, 2004 The population genetics and evolutionary epidemiology of RNA viruses. *Nat. Rev. Microbiol.* **2**: 279–288.
- OLKKONEN, V., and D. BAMFORD, 1989 Quantitation of the absorption and penetration stages of bacteriophage *phi6* infection. *Virology* **171**: 229–238.
- SANJUÁN, R., J. CUEVAS, A. MOYA and S. ELENA, 2005 Epistasis and the adaptability of an RNA virus. *Genetics* **170**: 1001–1008.
- TURNER, P., and L. CHAO, 1999 Prisoner's dilemma in an RNA virus. *Nature* **398**: 441–443.
- TURNER, P., and L. CHAO, 2003 Escape from prisoner's dilemma in RNA phage *phi6*. *Am. Nat.* **161**: 497–505.
- TURNER, P., C. BURCH, K. HANLEY and L. CHAO, 1999 Hybrid frequencies confirm limit to coinfection in the RNA bacteriophage *phi6*. *J. Virol.* **73**: 2420–2424.
- URQUIDI, V., and D. H. L. BISHOP, 1992 Non-random reassortment between the tripartite RNA genomes of La Crosse and snowshoe hare viruses. *J. Gen. Virol.* **73**: 2255–2265.
- WEINREICH, D., R. WATSON and L. CHAO, 2005 Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* **59**: 1165–1174.
- WILKE, C., and I. NOVELLA, 2003 Phenotypic mixing and hiding may contribute to memory in viral quasispecies. *BMC Microbiol.* **3**: 11.

Communicating editor: N. TAKAHATA

APPENDIX: DETAILS OF DETERMINISTIC MODEL

Frequencies of mutants: The single-mutant frequency F_i^* after the action of complementation, epistasis, selection, and reassortment is

$$\begin{aligned}
 \bar{w}F_i^* = & (1 - p_c)F_i e^{-s} + p_c \left\{ F_i^2 e^{-s} + F_0 F_i + \sum_{j:j \neq i} F_i F_{i,j} e^{-s} + \sum_{j:j \neq i} \frac{1}{2} (F_i F_j + F_0 F_{i,j}) \right. \\
 & + \sum_{j,k;j,k \neq i;j < k} \frac{1}{2^2} (F_i F_{j,k} + F_0 F_{i,j,k}) + \sum_{j,k;j,k \neq i;j < k} \frac{1}{2} (F_{i,j} F_{i,k} + F_i F_{i,j,k}) e^{-s} \\
 & + \sum_{j,k;j \neq k \neq i} \frac{1}{2^2} F_{i,j} F_k + \sum_{j,k,m;j \neq k \neq m \neq i;j < k} \frac{1}{2^2} F_{i,j,k} F_{i,m} e^{-s} I_{\{L \geq 4\}} \\
 & + \sum_{j,k,m;j \neq k \neq m \neq i;k < m} \frac{1}{2^3} F_{i,j} F_{k,m} I_{\{L \geq 4\}} + \sum_{j,k,m;j \neq k \neq m \neq i;j < k < m} \frac{1}{2^3} F_i F_{j,k,m} I_{\{L \geq 4\}} \\
 & + \sum_{j,k,m;j \neq k \neq m \neq i;j < k} \frac{1}{2^3} F_{i,j,k} F_m I_{\{L \geq 4\}} + \sum_{j,k,l,m;j \neq k \neq l \neq m \neq i;j < k,l < m} \frac{1}{2^3} F_{i,j,k} F_{i,l,m} e^{-s} I_{\{L \geq 5\}} \\
 & + \sum_{j,k,l,m;j \neq k \neq l \neq m \neq i;j < k,l < m} \frac{1}{2^4} F_{i,j,k} F_{l,m} I_{\{L \geq 5\}} + \sum_{j,k,l,m;j \neq k \neq l \neq m \neq i;k < l < m} \frac{1}{2^4} F_{i,j} F_{k,l,m} I_{\{L \geq 5\}} \\
 & \left. + \sum_{j,k,m,n;j \neq k \neq l \neq m \neq n \neq i;j < k,l < m < n} \frac{1}{2^5} F_{i,j,k} F_{l,m,n} I_{\{L \geq 6\}} \right\}. \tag{A1}
 \end{aligned}$$

The double-mutant frequency F_{ij}^* after the action of complementation, epistasis, selection, and reassortment is

$$\begin{aligned}
\bar{w}F_{ij}^* = & (1 - p_c)F_{ij}e^{-2s(1+\epsilon)} + p_c \left\{ F_{ij}^2 e^{-2s(1+\epsilon)} + \frac{1}{2}(F_0F_{ij} + F_iF_j) + (F_i + F_j)F_{ij}e^{-s} \right. \\
& + \sum_{k \neq i,j} \frac{1}{2}F_{ij}(F_{i,k} + F_{j,k})e^{-s} + \sum_{k \neq i,j} \frac{1}{2^2}(F_{ij}F_k + F_iF_{j,k} + F_jF_{i,k} + F_0F_{i,j,k}) \\
& + \sum_{k \neq i,j} \frac{1}{2}(F_i + F_j)F_{ij,k}e^{-s} + \sum_{k,m; k \neq m \neq i,j} \frac{1}{2^2}(F_{i,m} + F_{j,m})F_{ij,k}e^{-s}I_{\{L \geq 4\}} \\
& + \sum_{k \neq i,j} F_{ij}F_{ij,k}e^{-2s(1+\epsilon)} + \sum_{k,m; k \neq m \neq i,j; k < m} \frac{1}{2}F_{ij,k}F_{ij,m}e^{-2s(1+\epsilon)}I_{\{L \geq 4\}} \\
& + \sum_{k,m;j \neq k \neq m \neq i; k < m} \frac{1}{2^3}(F_{i,j}F_{k,m} + F_{i,k}F_{j,m})I_{\{L \geq 4\}} + \sum_{k,m;j \neq k \neq m \neq i; k < m} \frac{1}{2^3}(F_iF_{j,k,m} + F_jF_{i,k,m})I_{\{L \geq 4\}} \\
& + \sum_{k,l,m;j \neq k \neq l \neq m \neq i; l < m} \frac{1}{2^3}(F_{i,j,k}F_{i,l,m} + F_{i,j,k}F_{j,l,m})e^{-s}I_{\{L \geq 4\}} + \sum_{k,l,m;j \neq k \neq l \neq m \neq i; l < m} \frac{1}{2^4}F_{i,j,k}F_{l,m}I_{\{L \geq 5\}} \\
& + \sum_{k,l,m;j \neq k \neq l \neq m \neq i; l < m} \frac{1}{2^4}(F_{i,k}F_{j,l,m} + F_{j,k}F_{i,l,m})I_{\{L \geq 5\}} + \sum_{k,l,m,n;j \neq k \neq l \neq m \neq n \neq i; l < m < n} \frac{1}{2^5}F_{i,j,k}F_{l,m,n}I_{\{L \geq 6\}} \left. \right\}. \tag{A2}
\end{aligned}$$

The triple-mutant frequency $F_{i,j,k}^*$ after the action of complementation, epistasis, selection, and reassortment is

$$\begin{aligned}
\bar{w}F_{i,j,k}^* = & (1 - p_c)F_{i,j,k}e^{-3s(1+\epsilon+\eta)} + p_c \left\{ F_{i,j,k}^2 e^{-3s(1+\epsilon+\eta)} + \frac{1}{2}(F_i + F_j + F_k)F_{i,j,k}e^{-s} \right. \\
& + \frac{1}{2^2}(F_iF_{j,k} + F_jF_{i,k} + F_kF_{i,j} + F_0F_{i,j,k}) + \frac{1}{2}(F_{i,j}F_{i,k} + F_{i,j}F_{j,k} + F_{i,k}F_{j,k})e^{-s} \\
& + (F_{i,j} + F_{i,k} + F_{j,k})F_{i,j,k}e^{-2s(1+\epsilon)} + \sum_{m; m \neq i \neq j \neq k} \frac{1}{2^2}(F_{i,m} + F_{j,m} + F_{k,m})F_{i,j,k}e^{-s}I_{\{L \geq 4\}} \\
& + \sum_{m; m \neq i \neq j \neq k} \frac{1}{2^2}(F_{i,j,m}F_{i,k} + F_{i,j,m}F_{j,k} + F_{i,k,m}F_{j,k})e^{-s}I_{\{L \geq 4\}} \\
& + \sum_{m; m \neq i \neq j \neq k} \frac{1}{2^2}(F_{i,j}F_{i,k,m} + F_{i,j}F_{j,k,m} + F_{i,k}F_{j,k,m})e^{-s}I_{\{L \geq 4\}} \\
& + \sum_{m; m \neq i \neq j \neq k} \frac{1}{2}(F_{i,j,m} + F_{i,k,m} + F_{j,k,m})F_{i,j,k}e^{-2s(1+\epsilon)}I_{\{L \geq 4\}} \\
& + \sum_{l,m;j \neq k \neq l \neq m \neq i} \frac{1}{2^3}(F_{i,j}F_{k,m} + F_{i,k}F_{j,m} + F_{j,k}F_{i,m} + F_{i,j,k}F_m)I_{\{L \geq 4\}} \\
& + \sum_{l,m;j \neq k \neq l \neq m \neq i} \frac{1}{2^3}(F_iF_{j,k,m} + F_jF_{i,k,m} + F_kF_{i,j,m})I_{\{L \geq 4\}} \\
& + \sum_{l,m;j \neq k \neq l \neq m \neq i} \frac{1}{2^4}(F_{i,j}F_{k,l,m} + F_{i,k}F_{j,l,m} + F_{j,k}F_{i,l,m})I_{\{L \geq 5\}} \\
& + \sum_{l,m;j \neq k \neq l \neq m \neq i} \frac{1}{2^4}(F_{i,l}F_{j,k,m} + F_{j,l}F_{i,k,m} + F_{k,l}F_{i,j,m})I_{\{L \geq 5\}} \\
& + \sum_{l,m,n;j \neq k \neq l \neq m \neq n \neq i; l < m < n} \frac{1}{2^5}F_{i,j,k}F_{l,m,n}I_{\{L \geq 6\}} \\
& + \sum_{l,m,n;j \neq k \neq l \neq m \neq n \neq i; m < n} \frac{1}{2^5}(F_{i,j,l}F_{k,m,n} + F_{i,k,l}F_{j,m,n} + F_{j,k,l}F_{i,m,n})I_{\{L \geq 6\}} \left. \right\}. \tag{A3}
\end{aligned}$$

Integrated deterministic model: We propose a model to integrate the effects of complementation, epistasis, selection, mutation, and random recombination and show it is equivalent to the previous deterministic model with the

separation step. We denote the mean fitness of viral populations using this integrated model $\bar{w}^{**}$ and frequencies of mutants with F^{**} . The frequency of wild-type genotype at the end of the viral replicating cycle is

$$\begin{aligned}
 \bar{w}^{**} F_0^{**} &= (1 - L\mu)(1 - p_c)F_0 + (1 - L\mu)p_c \\
 &\times \left\{ F_0^2 + \sum_i F_0 F_i + \sum_{i,j;i < j} \frac{1}{2} F_0 F_{i,j} \right. \\
 &\quad + \sum_{i,j;i < j} \frac{1}{2} F_i F_j + \sum_{i,j,k;i < j < k} \frac{1}{2^2} F_0 F_{i,j,k} \\
 &\quad + \sum_{i,j,k;i \neq j, k; j < k} \frac{1}{2^2} F_i F_{j,k} + \sum_{i,j,k,m;i \neq j, k, m; j < k < m} \frac{1}{2^3} F_i F_{j,k,m} I_{\{L \geq 4\}} \\
 &\quad + \sum_{i,j,k,m;i \neq j \neq k \neq m; i < j; k < m} \frac{1}{2^3} F_{i,j} F_{k,m} I_{\{L \geq 4\}} \\
 &\quad + \sum_{i,j,k,m,n;i \neq j \neq k \neq m \neq n; i < j; k < m < n} \frac{1}{2^4} F_{i,j} F_{k,m,n} I_{\{L \geq 5\}} \\
 &\quad \left. + \sum_{i,j,k,l,m,n;i \neq j \neq k \neq l \neq m \neq n; i < j < k; l < m < n} \frac{1}{2^5} F_{i,j,k} F_{l,m,n} I_{\{L \geq 6\}} \right\} \\
 &= (1 - L\mu) \bar{w} F_0^*.
 \end{aligned} \tag{A4}$$

Similarly, the frequencies of mutants are

$$\begin{aligned}
 \bar{w}^{**} F_i^{**} &= (1 - (L - 1)\mu) \bar{w} F_i^* + \mu \bar{w} F_0^* \\
 \bar{w}^{**} F_{i,j}^{**} &= (1 - (L - 2)\mu) \bar{w} F_{i,j}^* + \mu (\bar{w} F_i^* + \bar{w} F_j^*) \\
 \bar{w}^{**} F_{i,j,k}^{**} &= \bar{w} F_{i,j,k}^* + \mu \bar{w} (F_{i,j}^* + F_{i,k}^* + F_{j,k}^*).
 \end{aligned} \tag{A5}$$

After simple arithmetic operations, it is clear that F^{**} 's are equal to F^* 's.

GENETICS

Supporting Information

<http://www.genetics.org/cgi/content/full/genetics.108.099796/DC1>

Complementation and Epistasis in Viral Coinfection Dynamics

Hong Gao and Marcus W. Feldman

Copyright © 2009 by the Genetics Society of America

DOI: 10.1534/genetics.108.099796

FIGURE S1.—Trajectories of mean fitness changes under scenarios I and II using deterministic simulation. (b), (c), (d), and (e) show under scenarios I and II the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively. Dashed lines represent the results assuming no mutation. Solid lines represent the results assuming mutation rate $\mu = 0.01$. Each color represents a co-infection fraction as labeled in the legend.

FIGURE S2.—(a) Trajectories of mean fitness changes under scenarios II and III using deterministic simulation. (b), (c), (d), and (e) show under scenarios II and III the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively. Solid lines represent the results assuming selection coefficient $s = 0.1$. Dashed lines represent the results assuming selection coefficient $s = 0.2$. Each color represents a co-infection fraction as labeled in the legend.

FIGURE S3.—Trajectories of mean fitness changes under scenarios II, IV and V using deterministic simulation. (b), (c), (d), and (e) show under scenarios II, IV and V the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively. Dashed lines represent the results assuming epistasis coefficient $\epsilon = 0.3$. Dotted lines represent the results assuming no epistasis. Solid lines represent the results assuming epistasis coefficient $\epsilon = -0.3$. Each color represents a co-infection fraction as labeled in the legend.

FIGURE S4.—(a) Trajectories of mean fitness changes under scenarios II, VI and VII using deterministic simulation. (b), (c), (d), and (e) show under scenarios II, VI and VII the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively. Dashed lines represent the results assuming epistasis coefficient $\epsilon = 0.8$. Dotted lines represent the results assuming no epistasis. Solid lines represent the results assuming epistasis coefficient $\epsilon = -0.8$. Each color represents a co-infection fraction as labeled in the legend.

FIGURE S5.—(a) Trajectories of mean fitness changes under scenarios VIII and IX using deterministic simulation. (b), (c), (d), and (e) show under scenarios VIII and IX the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively. Dashed lines represent the results assuming second order epistasis coefficient $\eta = 0.8$. Solid lines represent the results assuming second order epistasis coefficient $\eta = -0.8$. Each color represents a co-infection fraction as labeled in the legend.

Figure S6.—Trajectories of mean fitness changes under scenarios I and II using stochastic simulation. (b), (c), (d), and (e) show under scenarios I and II the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively. Dashed lines represent the results assuming no mutation. Solid lines represent the results assuming mutation rate $\mu = 0.01$. Each color represents a co-infection fraction as labeled in the legend.

FIGURE S7.—(a) Trajectories of mean fitness changes under scenarios II and III using stochastic simulation. (b), (c), (d), and (e) show under scenarios II and III the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively. Solid lines represent the results assuming selection coefficient $s = 0.1$. Dashed lines represent the results assuming selection coefficient $s = 0.2$. Each color represents a co-infection fraction as labeled in the legend.

Figure S8.—Trajectories of mean fitness changes under scenarios II, IV and V using stochastic simulation. (b), (c), (d), and (e) show under scenarios II, IV and V the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively. Dashed lines represent the results assuming epistasis coefficient $\epsilon = 0.3$. Dotted lines represent the results assuming no epistasis. Solid lines represent the results assuming epistasis coefficient $\epsilon = -0.3$. Each color represents a co-infection fraction as labeled in the legend.

FIGURE S9.—(a) Trajectories of mean fitness changes under scenarios VIII and IX using stochastic simulation. (b), (c), (d), and (e) show under scenarios VIII and IX the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively. Dashed lines represent the results assuming second order epistasis coefficient $\eta = 0.8$. Solid lines represent the results assuming second order epistasis coefficient $\eta = -0.8$. Each color represents a co-infection fraction as labeled in the legend.

FIGURE S10.—Mean fitness pattern under the three-way interaction among complementation, selection and epistasis. (a), (b), (c), (d), and (e) show the interaction patterns between selection and epistasis with the co-infection ratios, 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. Colors represent values of the overall mean fitness under each combination of parameters, red for value 0.95, yellow for 0.967, green for 0.983, and blue for 1.00.

FIGURE S11.—Double mutant frequency pattern under the three-way interaction among complementation, selection and epistasis. (a), (b), (c), (d), and (e) show the interaction patterns between selection and epistasis with the co-infection ratios, 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. Colors represent the values of wildtype frequency, red= 0.00, yellow= 1/3, green= 2/3, and blue= 1.00.

Figure S12.—Triple mutant frequency pattern under the three-way interaction among complementation, selection and epistasis. (a), (b), (c), (d), and (e) show the interaction patterns between selection and epistasis with the co-infection ratios, 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. Colors represent the values of wildtype frequency, red= 0.00, yellow= $1/3$, green= $2/3$, and blue= 1.00.

FIGURE S13.—(a), (b), (c), and (d) show under various starting frequencies the frequency trajectories of wildtype, single mutants, double mutants, and triple mutants, respectively, using deterministic simulation and assuming parameters $p_e = 0.7$, $\mu = 0.01$, $s = 0.1$, $\epsilon = -0.8$, and $\eta = 0.0$. Colors represent different initial frequencies as labeled in the legend. F_0 represents the initial frequency of wildtype, F_1 , F_2 , and F_3 represents the initial frequencies of single mutants and F_{123} represents the initial frequencies of triple mutants.