Adaptive Molecular Evolution for 13,000 Phage Generations: A Possible Arms Race

Holly A. Wichman,*,[1] Jack Millstein* and J. J. Bull†

*Department of Biological Sciences and Initiative for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, Idaho 83844-3051 and †Department of Integrative Biology and Institute of Cellular and Molecular Biology, University of Texas, Austin, Texas 78712-1023

ABSTRACT

Bacteriophage φX174 was evolved on a continuous supply of sensitive hosts for 180 days (~13,000 phage generations). The average rate of nucleotide substitution was nearly 0.2% (11 substitutions)/20 days, and, surprisingly, substitutions accumulated in a clock-like manner throughout the study, except for a low rate during the first 20 days. Rates of silent and missense substitutions varied over time and among genes. Approximately 40% of the 71 missense changes and 25% of the 58 silent changes have been observed in previous adaptations; the rate of parallel substitution was highest in the early phase of the evolution, but 7% of the later changes had evolved in previous studies of much shorter duration. Several lines of evidence suggest that most of the changes were adaptive, even many of the silent substitutions. The sustained, high rate of adaptive evolution for 180 days defies a model of adaptation to a constant environment. We instead suggest that continuing molecular evolution reflects a potentially indefinite arms race, stemming from high levels of co-infection and the resulting conflict among genomes competing within the same cell.

MOLECULAR evolution is an established fact, but the relative contributions of various factors to that evolution remain controversial (KIMURA 1983; NEI 1987; GILLESPIE 1991; LI 1997). Advantageous, neutral, and even deleterious mutations can become common in a population, and their respective rates of evolution depend on mutation, the strength of selection, and population size, plus the influence of other loci through linkage and epistasis. While much of the work in the field has focused on neutral evolution, a current goal is to identify adaptive evolution by finding its signatures in molecular data. Yet despite considerable evidence of adaptive molecular evolution, much uncertainty remains about the nature of the selective agents.

Toward the goal of revealing the diverse mechanisms of molecular evolution, we have conducted several experimental studies of bacteriophage evolution (BULL *et al.* 1997, 2003; WICHMAN *et al.* 1999, 2000; CRILL *et al.* 2000; HOLDER and BULL 2001). In some of those studies, phages were grown in continuous culture of a two-stage chemostat, whereby actively growing, sensitive bacteria were continuously pumped into a vessel containing phage, allowing the phage to infect and reproduce as if their population were expanding exponentially (BULL *et al.* 1997; WICHMAN *et al.* 1999, 2000). Populations of those phages accumulated nucleotide changes over the duration of the continuous culture, but those previous studies tended to be short (mostly 10 or 11 days, with approximately three phage generations per hour). The nucleotide changes in those systems exhibited characteristics of adaptive evolution: a high rate of missense/silent substitution and high rates of parallel changes. Where the phenotypic effects of changes were identified, they could be interpreted as facilitating phage growth under the culture conditions (BULL *et al.* 2000; CRILL *et al.* 2000).

To further explore the nature of molecular evolution in this system, we now report on a much longer duration of phage growth in continuous culture. In addition, the culture conditions are more benign and are meant to approximate "control conditions," which should impose less selection for change. By conducting a longer adaptation with frequent sampling of the population, the temporal characteristics of molecular evolution may now be studied. If adaptation in this system is primarily to the constant ("control") culture environment, we expect that the initial rate of molecular evolution will be low and then wane as fitness approaches its maximum (FISHER 1958; GILLESPIE 1983, 1984; ORR 1998, 2002).

In two previous adaptations to inhibitory conditions, an initial high rate of molecular evolution decreased by 66 and 80% during extended adaptation in the same environment (BULL *et al.* 1997), supporting the theoretical expectations of a plateau. While the initial rate of molecular evolution in this study was low, consistent with the benign growth conditions, the rate accelerated during extended evolution and showed no sign of a

[1]*Corresponding author:* Department of Biological Sciences, P.O. Box 443051, University of Idaho, Moscow, ID 83844-3051.
E-mail: hwichman@uidaho.edu

plateau even after 13,000 generations. We provide evidence that molecular evolution in this system is driven by high-density interactions, as observed in some other chemostat systems (Paquin and Adams 1983; DePolo *et al.* 1987). This system allows us to examine the temporal characteristics of the molecular evolution driven by these high-density interactions. The sequences here cover the entire phage genome, allowing a comprehensive level of analysis of tempo and mode of this evolution.

The experimental evolution described below used the icosahedral bacteriophage ϕX174. This phage has a small, single-stranded DNA genome of 5386 nucleotides and encodes 11 genes (Sanger *et al.* 1977). The biology of icosahedral phages has been intensively studied and most gene functions are known (Hayashi *et al.* 1988; Bangham and Kirkwood 1993), but more recently a variety of structural studies has been coupled with mutational analyses to infer gene function and interactions at a detailed level (Dokland *et al.* 1997, 1999; Burch *et al.* 1999; Burch and Fane 2003; Bernal *et al.* 2004; Novak and Fane 2004). ϕX174 does not encode its own polymerase, but rather relies on host enzymes for replication. The virion binds to the lipopolysaccaride (LPS) in the outer membrane of its host, which must lack the O antigen. The natural host is thought to be a rough strain of *Escherichia coli*, but the phage can be isolated on or evolved to infect various LPS variants of *E. coli* and Salmonella as well as some closely related bacteria.

MATERIALS AND METHODS

Strains: A wild-type isolate of ϕX174 was used as the ancestor for these studies (GenBank no. AF176034). The bacteria used for phage propagation was *E. coli* C.

Two-stage chemostat: A chemostat consisting of two connected 100×15-mm glass test tubes was used to select phage populations (Bull *et al.* 1997). LB broth with 0.005% antifoam B (Sigma, St. Louis) and 2 mm $CaCl_2$ was pumped continuously into the first tube, which maintained host cells at a density suitable for phage growth; phage were absent from this "cell" tube. Overflow from the cell tube was drawn continuously into the second tube, which contained phage (the "phage" tube). The volume of liquid in the phage tube was maintained at 1–2 ml, with a flow-through rate of 7–10 ml/hr; thus the volume of the phage tube was replaced on the order of 100 times/day. Because of the continuous nature of this replacement, however, this exchange rate is equivalent to a growth of 144 phage population doublings/day. Both tubes were aerated by bubbling with filtered air. Both tubes were immersed in a water bath whose temperature was maintained at 37.5°, 0.5° above the designated chemostat temperature to compensate for the observed cooling effect of the aeration. Chemostat tubes and cells were changed at least every 2 days, thus discarding any phage-resistant cells in the phage tube. Samples from the phage tube were run through a 0.2-µm filter to remove the host (MillexGP, Millipore, Bedford, MA) and 0.2 ml was used to reinoculate the phage tube of the chemostat.

The duration of a chemostat line is measured in days or phage generations. The time from infection to burst is not as synchronous in ϕX174 as with the T-phages and is reported to range from 15 to 30 min with a mean of 21 min at 37° (Hutchison and Sinsheimer 1963; Denhardt and Sinsheimer 1965). This asynchrony stems from a passive lysis mechanism that prevents synthesis of the cell septum and causes rupture when the cell attempts to divide; thus asynchrony in cell division leads to asynchrony of lysis. Use of the mean lysis time gives an approximate estimate of generation time, but when the population is expanding (as here), early lysis has a disproportionate effect and causes the population to behave as if its generation time is shorter than the mean (Abedon *et al.* 2001). We thus consider the generation time in our chemostat to be 20 min (three phage generations/hour). This number is necessarily approximate (indeed, generation time will evolve over the course of an experiment), but the exact value is unimportant to the conclusions here.

Fitness assays: Phage growth rate, represented as doublings per hour, was measured as the rate at which a phage increased its numbers when placed in a culture where cells outnumbered phages. Fitness was estimated as the $\log_2$ of the increase in phage concentration per hour. The assay was conducted in 125-ml flasks with 10 ml LB. Hosts were grown at 37° with aeration to a density of $\sim 10^8$ cells/ml. At this point (t_0), phage were added to a concentration of 10^3–10^4/ml and the mixture was grown for 40 min with shaking before growth was terminated with chloroform; the final concentration of phage (after chloroform) was not allowed to exceed the cell density. Phage titers were then determined immediately after harvest by plating. Fitness was calculated as 1.5 [$\log_2$(phage concentration at 40 min) $-$ $\log_2$(phage concentration at t_0)]. Other considerations that apply to this assay have been discussed elsewhere (Bull *et al.* 2000).

Sequencing: Genomes of phage isolates were amplified by PCR in two overlapping pieces with r*Tth* DNA polymerase, XL (Applied Biosystems, Foster City, CA). Gel-purified PCR products were sequenced on an ABI capillary sequencer. Ambiguities were resequenced, and if necessary a different primer or different DNA extraction was used.

Statistics: Where feasible, standard statistical tests were used, as described in the results. However, in many cases the significance of associations in the data was evaluated with empirical permutation tests; the null distributions for these tests were generated by randomly associating the observed data without replacement for at least tens of thousands of trials.

RESULTS

Sustained rates of change

For the 180 days of evolution, crudely estimated as 12,960 phage generations and a net expansion of 25,920 population doublings, complete genomes were sequenced from four phage isolates taken every 20 days (Figure 1). In the complete data set of 36 isolate sequences, 71 missense, 58 silent, and eight intergenic substitutions as well as four intergenic insertions/deletions were detected. Overall, once a substitution first appeared in an isolate, it tended to remain common for the duration of the experiment, although some changes remained demonstrably polymorphic, and $\sim 20\%$ made only a transitory appearance before dropping to low frequency or disappearing.

In previous articles (Bull *et al.* 1997; Wichman *et al.* 1999, 2000) we typically observed 12–15 changes in the

first 10 days of evolution in a strongly selective environment. As expected for a virus evolving under benign conditions, the rate of molecular evolution was initially quite low here, with only a single change during the first 10 days of evolution (base 1301; data not shown). Unexpectedly, over the extended evolution we observed a striking, near-linear increase in substitutions averaging 11/20 days. Figure 2 shows not only the total number of differences from the starting genome for each isolate but also the pairwise distance between isolates from the same time point, supporting the impression from Figure 1 that evolution is dominated by a series of selective sweeps and that most time points also had considerable variation. Figure 2 (bottom) shows that this clock-like accumulation applied to both missense and silent substitutions, with similar slopes over most of the experiment, but with a lag before the rise of silent substitutions.

Signatures of adaptive evolution

When comparing sequences evolved in nature, the null model is that divergence represents neutral evolution. The experimental system is not necessarily expected to mirror natural evolution, both because the time span of experiments is too short and because the experimental environment is likely novel and thus should be selective. It is thus plausible that the experimental evolution is more adaptive than natural evolution (in terms of the proportion of changes that are adaptive), and there are several measures of molecular evolution that can be observed to assess the degree of adaptive molecular evolution here.

Missense/silent rates: An excess of missense substitutions (nonsynonymous) over silent substitutions (synonymous) compared to the expected mutational ratio (*i.e.*, $d_N/d_S > 1$) is considered a signature of adaptive evolution, while an excess of silent over missense substitutions ($d_N/d_S < 1$) is generally assumed to indicate purifying selection (Li 1997). For ϕX174, the expected ratio of missense-to-silent mutations in nonoverlapping coding regions, given the mutational bias observed in this system, is 2.7 (analysis not shown), so a missense-to-silent substitution ratio of 2.7 would indicate a d_N/d_S ratio of 1. Previous work with ϕX174 in chemostats reported an excess of missense substitutions, using 10- to 33-day adaptations with temperatures in excess of 42° (Bull *et al.* 1997; Wichman *et al.* 2000). That pattern held only for early passages in this study (Figure 2, bottom). The cumulative d_N/d_S ratio was 3.70 (a missense-to-silent ratio of 10) by day 40, but fell below 1 at day 60 and was 0.27 by day 180. The ratio of new silent *vs.* new missense changes observed at each of the 20-day time points is significantly heterogeneous [$\chi^2(8) = 19.2$, $P < 0.02$], and the correlation of the number (proportion) of new silent changes with day is 0.76 (0.75) and is

statistically significant ($P < 0.02$ for both, by a permutation test). The slight decline in the number of missense changes with time is not significant. Taken alone, this pattern might indicate that early evolution in this study was driven by adaptive evolution while later evolution was driven by purifying selection and neutral evolution. However, we present arguments below that do not support this notion.

Variation in silent/missense changes across the genome: Silent changes were significantly clustered in the genome (Figure 3, top). The numbers of silent changes in the single reading frame regions of genes A, F, G, and H deviated significantly from the expected numbers on the basis of sequence length, with a large excess in A and deficits in F, G, and H [$\chi^2(3) = 34.8$, $P < 2 \times 10^{-7}$]; the equivalent test for missense substitutions is not statistically significant, but the ratio of silent-to-missense changes is significantly heterogeneous across these genes [$\chi^2(3) = 12.7$, $P < 0.006$]. A clustering of silent substitutions is consistent either with a selective advantage of silent changes that were observed or with a deleterious effect of silent changes that were not observed. To help discriminate among these alternatives, we counted the numbers of silent and missense differences between the same regions of the ancestral version ϕX174 and a closely related phage, S13. The distribution of silent changes in this comparison differs profoundly from the distribution of silent changes in the experimental line, whether the test treats each of the four gene regions separately or combines F, G, and H ($P < 5 \times 10^{-7}$). The comparison of missense changes between the experimental line and the ϕX174-S13 contrast is not statistically significant, however. This evidence supports the model that the clustered silent changes in gene A of the experimental line reflect positive selection, unless it is argued that negative selection on silent changes operated differently in the chemostat than in the wild.

Codon bias: Although silent substitutions are often considered neutral, it is widely acknowledged that selection in large populations can operate on silent changes through codon bias, nucleic acid secondary structure, or binding of the DNA by proteins. Of the 137 substitutions identified during the 180 days, 58 produced silent changes. The distribution of base changes differs significantly between the silent substitutions and the others [$\chi^2(11) = 26$, $P < 0.007$]. This difference could stem from some mutational classes being more likely than others to create missense changes (*e.g.*, transversions *vs.* transitions). However, half of the chi-square value comes from a strong excess of T $\rightarrow$ C changes among silent substitutions, which may be the reversal of the large excess of T in third positions found in the wild type.

Gene A (and A*) has the largest concentration of silent changes (36 in the region of nonoverlapping reading frames). We evaluated the impact of those changes on the codon adaptation index (CAI), which measures
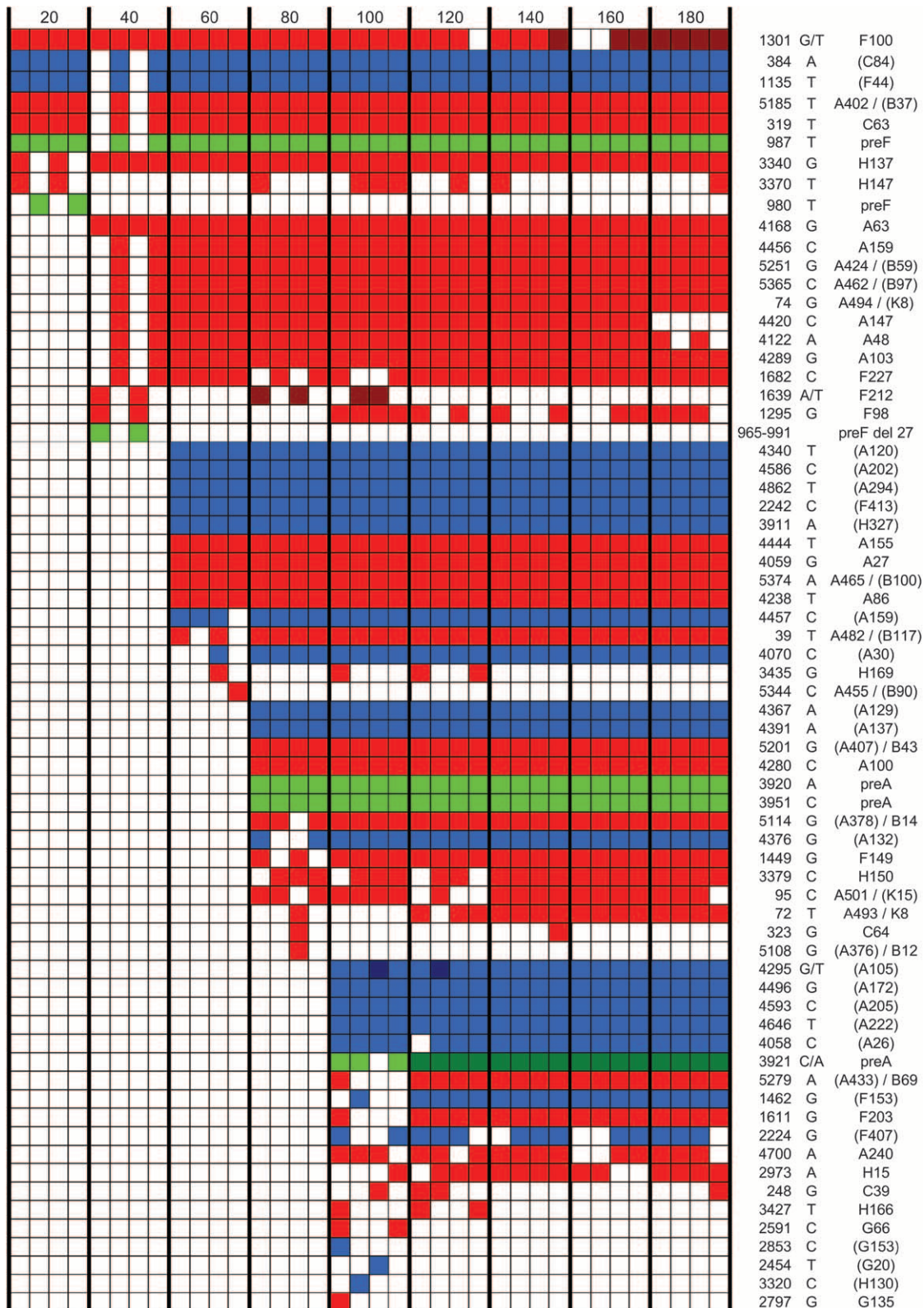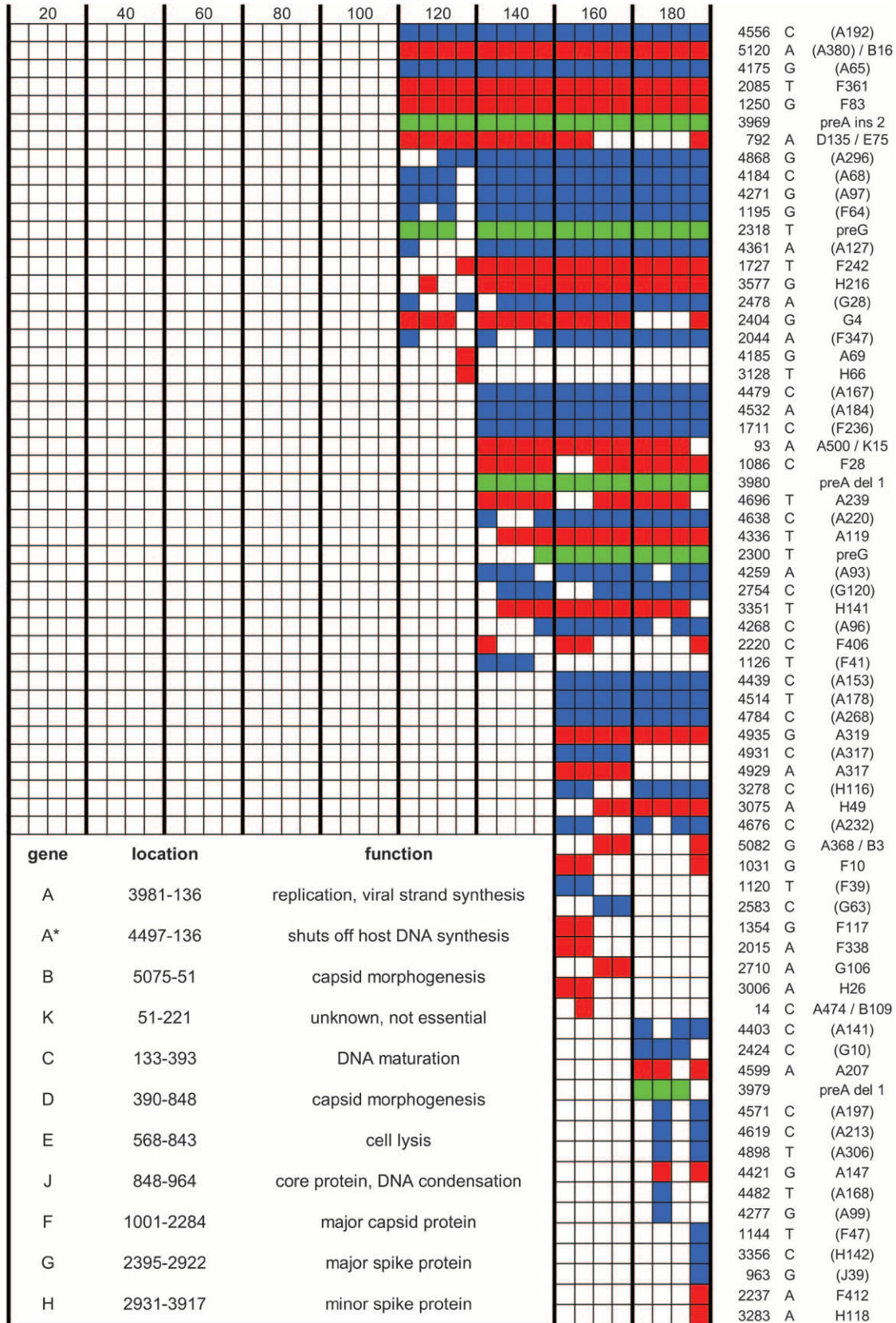
Figure 1.—Matrix of all changes seen in φX174 chemostat evolution (four isolates per time point). The top row indicates the day on which isolates were sampled, and isolates from different time points are separated by vertical black bars. For each time point, each of the four columns indicates the complete set of changes found in an isolate. The presence/absence of these changes in isolates from different time points is indicated by color across that row. From top to bottom, changes are presented in order of first occurrence. Changes resulting in amino acid replacements are indicated in red, silent changes in blue, and intergenic changes in green. Second changes at the same nucleotide position are indicated by a darker shade of the same color. Counts of all changes given in the text included these same-site changes. The columns following the isolate changes indicate genomic position of the change, the base change relative to GenBank accession no. AF176034, and the amino acid number in

20 40 60 80 100 120 140 160 180

gene	location	function
A	3981-136	replication, viral strand synthesis
A*	4497-136	shuts off host DNA synthesis
B	5075-51	capsid morphogenesis
K	51-221	unknown, not essential
C	133-393	DNA maturation
D	390-848	capsid morphogenesis
E	568-843	cell lysis
J	848-964	core protein, DNA condensation
F	1001-2284	major capsid protein
G	2395-2922	major spike protein
H	2931-3917	minor spike protein

4556 C (A192)
5120 A (A380) / B16
4175 G (A65)
2085 T F361
1250 G F83
3969 preA ins 2
792 A D135 / E75
4868 G (A296)
4184 C (A68)
4271 G (A97)
1195 G (F64)
2318 T preG
4361 A (A127)
1727 T F242
3577 G H216
2478 A (G28)
2404 G G4
2044 A (F347)
4185 G A69
3128 T H66
4479 C (A167)
4532 A (A184)
1711 C (F236)
93 A A500 / K15
1086 C F28
3980 preA del 1
4696 T A239
4638 C (A220)
4336 T A119
2300 T preG
4259 A (A93)
2754 C (G120)
3351 T H141
4268 C (A96)
2220 C F406
1126 T (F41)
4439 C (A153)
4514 T (A178)
4784 C (A268)
4935 G A319
4931 C (A317)
4929 A A317
3278 C (H116)
3075 A H49
4676 C (A232)
5082 G A368 / B3
1031 G F10
1120 T (F39)
2583 C (G63)
1354 G F117
2015 A F338
2710 A G106
3006 A H26
14 C A474 / B109
4403 C (A141)
2424 C (G10)
4599 A A207
3979 preA del 1
4571 C (A197)
4619 C (A213)
4898 T (A306)
4421 G A147
4482 T (A168)
4277 G (A99)
1144 T (F47)
3356 C (H142)
963 G (J39)
2237 A F412
3283 A H118

the affected protein. (Tabular inset) φX174 gene location and function. The phage has a single-stranded DNA genome of 5386 nucleotides and encodes 11 genes. Genes A, B, K, C, D, and E are at least partially overlapping in different reading frames; gene A* overlaps A but is in the same reading frame. A text version of this table and an alignment of sequenced isolates is provided as supplementary material at http://www.genetics.org/supplemental/.
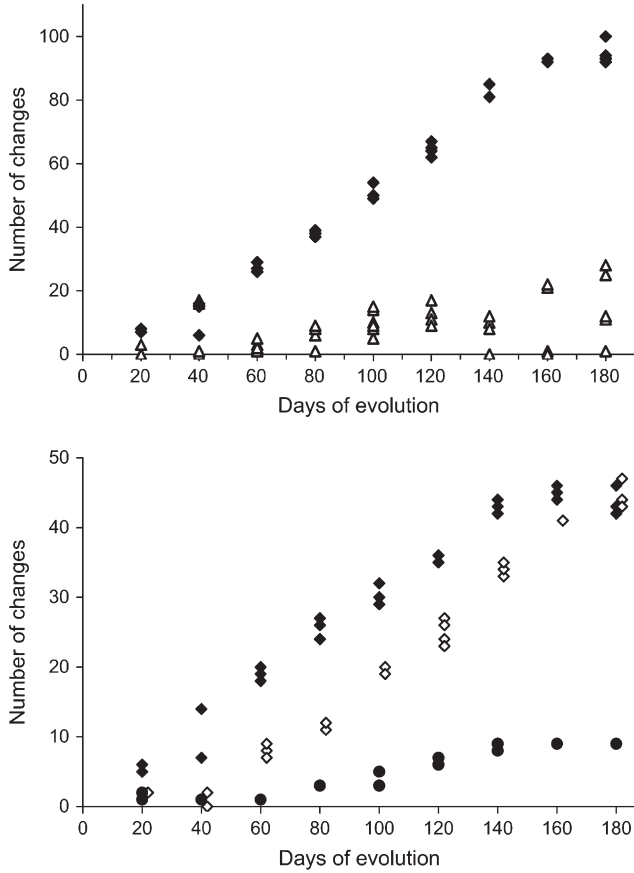
FIGURE 2.—Cumulative changes in the φX174 chemostat line through day 180, ∼13,000 generations. (Top) Solid diamonds show the number of differences between each sequenced isolate and the ancestral φX174 sequence. Open triangles show the distance between pairs of isolates from the same time point (six comparisons per time). (Bottom) The number of differences between each sequenced isolate and the ancestral φX174 sequence is partitioned into missense (solid diamonds), silent (open diamonds), and intergenic (solid circles) changes.

the similarity of codon usage in a sequence to the codon usage known for highly expressed genes of *E. coli* (SHARP and LI 1987). For each amino acid, the most-used codon in highly expressed genes is given a score of 1, and less-used codons are given lower, positive scores. The CAI of a sequence is then the geometric mean of the scores for its codons. A geometric mean is highly sensitive to extremely low values of any of its elements, so improvements can occur in some codons with little overall change in the CAI.

The CAI over the 1092-base portion of gene A with a single reading frame improved from 0.32 to 0.33 when 35 of the silent changes were added. One substitution appeared simultaneously with a missense change in the same codon and hence was excluded from this calculation. The null distribution was simulated by introducing 35 random silent substitutions into the ancestral, single-reading frame portion of gene A using the substitution matrix of observed point mutations and calculating the
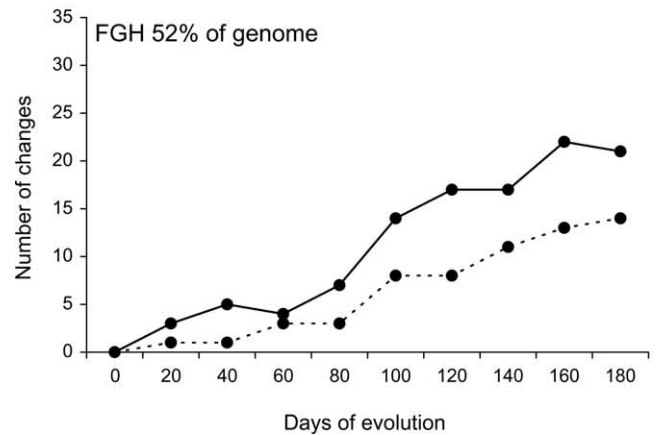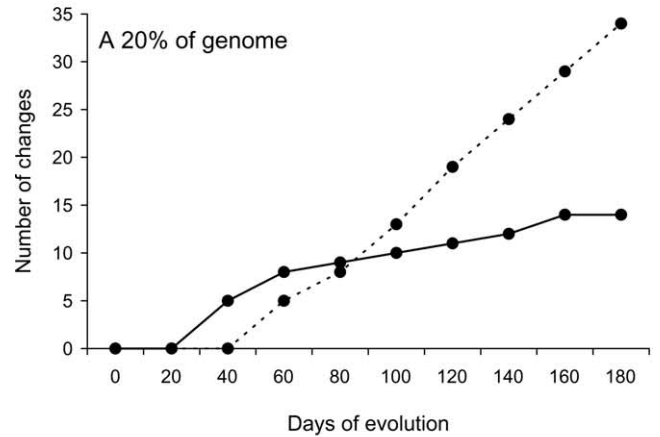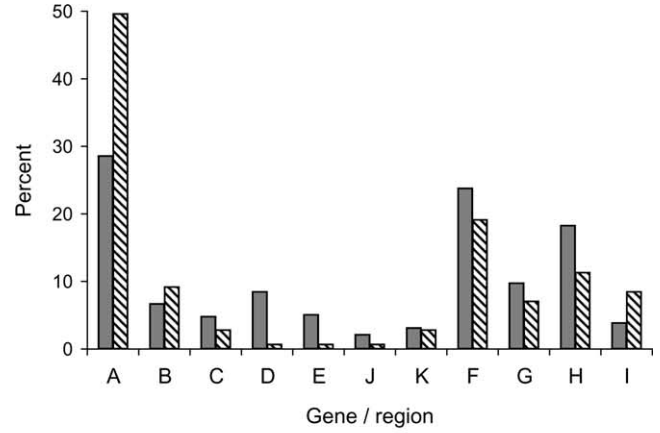

FIGURE 3.—Point mutations appearing in different parts of the genome during 180 days of evolution. (Top) The percentage of the total genome represented by each gene is shown by a solid bar, and the percentage of the total changes observed in the gene is shown by a hatched bar. (Intergenic regions from throughout the genome are combined in category I.) Overlapping reading frames contribute to multiple genes, so the totals exceed 100%. (Middle) Missense changes in the single-reading-frame codons of gene A (bases 3981–5072) are shown by the solid line, silent changes with the dashed line. (Bottom) Missense changes in genes F, G, and H (solid line), silent changes (dashed). Shown here are the total numbers of different mutations/substitutions seen in any of the isolate sequences from the respective time point. Thus, a mutation seen in one or more isolates at day 60 but not at day 80 contributes to the total at day 60 but not at day 80.
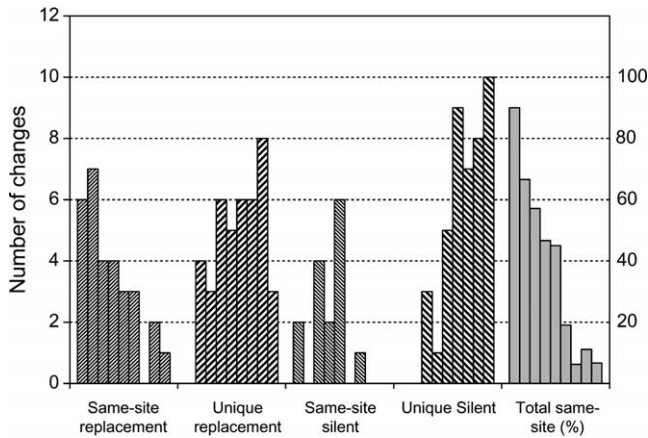
FIGURE 4.—Comparison of all changes seen during 180 days of evolution in the φX174 chemostat line to changes seen in other chemostat evolutions of φX174. Same-site amino acid replacements, unique amino acid replacements, same-site silent changes within genes, and unique silent changes within genes are plotted in the first four blocks, with one bar for each time point from day 20 through day 180. Changes are plotted at the time of first occurrence. A same-site parallel change is a mutation at the same genome position as seen before, but not necessarily to the same base; a unique change is one that has been observed only in this study. The rightmost block gives the percentage of new changes seen at each time point that were also seen in previous experiments. This block also includes intergenic changes and indels. The data set used for comparison included changes seen in BULL *et al.* (1997), WICHMAN *et al.* (1999, 2000), CRILL *et al.* (2000), and unpublished chemostats.

new CAI, repeated 100,000 times both with and without replacement. The mean CAI of the null distribution was 0.006 lower than the CAI of the wild-type sequence, indicating that random additions of 35 silent changes would have lowered the value. The slight increase observed in the actual changes is nearly statistically significant ($P \sim 0.06$). This test merely indicates whether codon bias was a factor in the evolution of these silent substitutions. We clearly cannot rule out codon bias as a factor impacting the evolution of silent substitutions here, but the fact that the CAI barely increased suggests that codon bias was not a major force favoring these silent substitutions (to the extent that the CAI truly reflects the impact of codon bias).

Parallel evolution: Parallel evolution is an especially useful signature of adaptive evolution because independent substitutions of the same base in different lineages (during brief adaptations) are plausibly attributed only to selection. High levels of parallel evolution have been observed in our previous studies, but because none of those phage populations were evolved as long as 180 days and few have been evolved under these conditions, we do not have the ideal database against which to make comparisons. Despite this limitation, a large fraction of substitutions during the 180-day adaptation have been observed in previous lines (Figure 4): 30 of 71 missense

changes, 15 of 58 silent changes, and 4 of 12 intergenic changes were the same as or at the same nucleotide site as seen in other experiments. The proportion of previously seen changes declines with time, although the trend for intergenic changes is not statistically significant (on the basis of permutation tests and using the day each substitution was first detected, $\rho = -0.87$, $P < 0.002$ for missense changes; $\rho = -0.93$, $P < 0.002$ for silent changes; and $\rho = -0.77$, $P \sim 0.06$ for intergenic changes).

Repeatability of evolution—an extension from day 100 isolates: Beyond day 60, the samples exhibited considerable polymorphism and sustained a high rate of evolution, with much of the change from silent variants. To study the repeatability of this evolution, we initiated a 20-day population with the four sequenced isolates from day 100; four new isolates were sequenced from the population after 20 days (referred to as 120R). The 100- to 120-day time frame was chosen specifically to ask whether the high rate of silent substitutions would be repeatable.

This replicate differs in some obvious ways from the original day 100 to day 120 population. The culture conditions maintained a large population size throughout the experiment, so the original day 100 population undoubtedly had more genetic and phenotypic variants than the four isolates used to evolve 120R. Thus we expect to see fewer new changes in day 120R than in the original day 120. On the other hand, using sequenced isolates allowed us to know the exact state of all fixed and polymorphic sites in the founding population and to more definitively identify any new parallel changes as adaptive. Figure 5 shows all sites that were polymorphic at day 100, 120R, or 120.

Fate of the day 100 polymorphisms: The day 100 isolates exhibited 20 variable sites, 19 of which could be used to test for differences between day 120 and 120R. One site (3921) could not be included in this test because all day 120 isolates contained a mutation that had not been seen at day 100. The differences between day 120 and 120R are not statistically significant for these 19 variable sites ($P \cong 0.935$; at each site, the exact probability of an outcome as extreme or more extreme than the results was calculated and conditioned on the marginal frequencies, and Fisher's combined probability test was used to measure the combined significance across all 19 sites). At 14 of the sites, the day 120 and 120R frequencies were as similar to one another as possible, given the total numbers of mutations observed in the 120 and 120R samples. It is noteworthy that, of the six polymorphic silent substitutions in the day 100 isolates, the three that went to high frequency by day 120 were found in three or four of the 120R isolates, and the three that were lost by day 120 were found in at most one isolate of 120R. For one of the silent changes that went to high frequency, the associations of closely linked missense polymorphisms were not preserved at

120R; hitchhiking thus does not seem to be a likely explanation for the increase of that silent change, but hitchhiking is consistent with the increases of the other two silent mutations.

New substitutions in day 120R: Eight new substitutions that were absent from the four founding genomes ap-
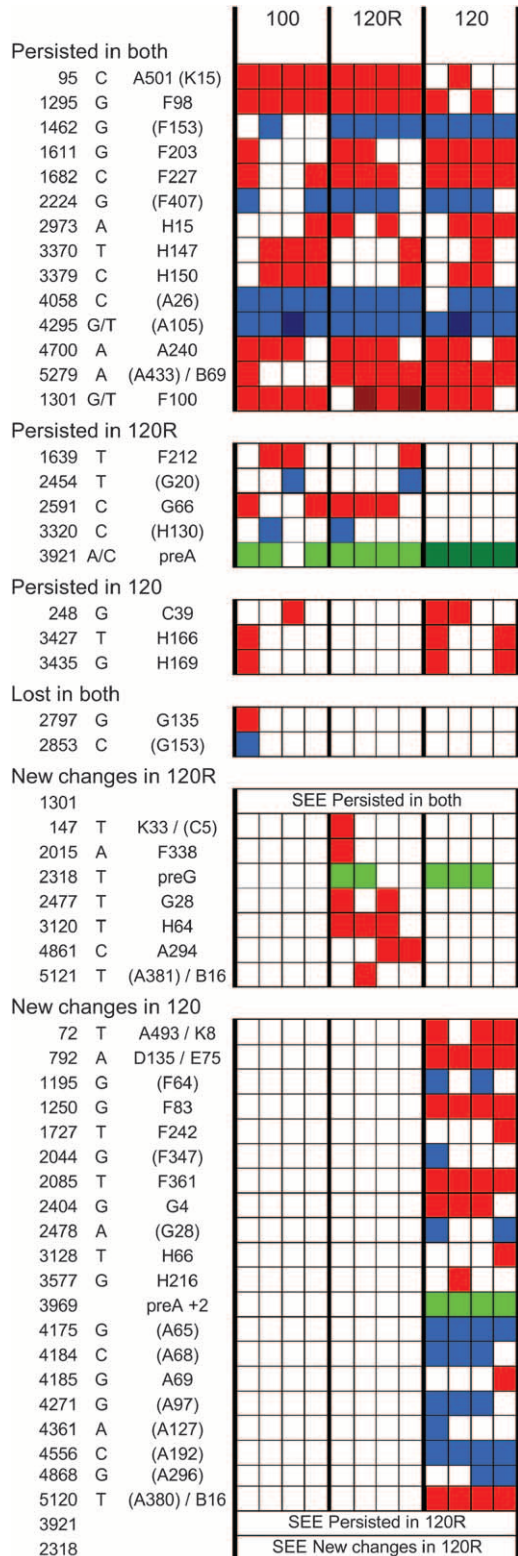

FIGURE 5.—Changes seen in the four isolates of day 100, day 120, and day 120R of the replicate evolution initiated with the four day 100 isolates. Color coding follows the scheme in Figure 1. Changes are grouped according to the description at the top of each block. Note that changes at bases 1301, 2318, and 3921 fall into two categories and are colored only the first time each is listed.

peared in the four isolates of 120R. At least seven of these are likely to have been adaptive. At three sites, new mutations that were absent from day 100 isolates but were identical to substitutions that appeared in the original 180-day chemostat appeared in 120R: 2318T was an intergenic change that also appeared at day 120, 1301T was a missense change that appeared at day 140, and 2015A was a missense change that appeared at day 160. In addition, two of the eight changes (3120T and 4861C) have been seen multiple times in previous experiments, and two others (2477T and 5121T) were in the same codons but at different sites as changes observed at day 120. This further supports the growing picture that most of the changes in this long-term chemostat are adaptive.

New substitutions in day 120: The most apparent difference between day 120 and 120R was the appearance of new mutations: 21 new substitutions and a 2-base insertion were seen at day 120, but only 8 new mutations appeared in 120R [$\chi^2(1) = 6.533$, $P \cong 0.011$]. The difference appears to be chiefly in new silent changes: 0:7:1 silent:missense:intergenic new mutations for 120R *vs.* 10:9:3 at day 120. The null model of no difference is marginally rejected [$\chi^2(2) = 6.02$, $P < 0.05$], and a one-tailed Fisher's exact test of the silent:missense differences is significant at $P \cong 0.018$. Thus we did not observe the same rate of silent substitution in 120R as in the original day 100 to day 120 time period. It seems likely that the polymorphism in the day 100 population contributed to this difference. Mutations present at modest frequency in the day 100 population but not present in the four isolates could have ascended to detectable levels by day 120. In that sense the difference between day 120 and 120R is not surprising, merely reflecting the time required for different types of substitutions to come to high frequency in the population once they arise. One possible interpretation of this difference is that most silent substitutions have a smaller selective coefficient than the missense or intergenic substitutions. This might help to explain why silent substitutions did not begin to show up in appreciable numbers until day 60 of the 180-day chemostat.

DISCUSSION

The major pattern of interest in these data is ongoing molecular evolution during 13,000 generations in a seemingly constant external environment. A result of

continual molecular evolution in a constant environment is puzzling if most changes are adaptive, but is not puzzling if most of the changes are neutral or deleterious. The large population sizes maintained in the chemostat (see below) should have been inimical to fixation of deleterious changes, but the (near) neutrality of substitutions is an option. In fact, several superficial observations are seemingly compatible with neutrality: the linear increase of changes over time, the excess of silent substitutions after day 40, and (as shown below) the absence of a "fitness" increase over time. Were it not for the consideration of other data, we might well be unable to support adaptive evolution as the explanation for most of these changes. Below we present arguments that (i) the rate of substitution was in excess of the expected rate of neutral substitution and (ii) much of the molecular evolution seen here has signatures of adaptation and is driven by high levels of within-host competition. We then consider how our perspective might have been different if we had not examined genome-wide evolution.

Low expected rate of neutral evolution: In considering the causes of molecular evolution in this system, a useful baseline quantity is the expected rate of neutral substitution. Two issues need to be assessed with respect to the fixation of neutral mutations: the time it takes such mutations to become fixed in the population and the expected ongoing rate of neutral substitution once the population reaches the equilibrium rate of neutral substitution. In the absence of selection, the expected time from origin of a successful neutral mutation to its fixation is $2N_e$ generations (half that of diploids; CROW and KIMURA 1970), a time vastly exceeding the duration of this experiment. The ascent of highly beneficial mutations can accelerate the fixation of neutral mutations through hitchhiking, but in the long-term, the per-generation fixation rate of neutral changes is expected to equal the genomic rate of neutral mutations, regardless of linkage and the possibility of hitchhiking (BIRKY and WALSH 1988). Thus hitchhiking cannot increase the equilibrium neutral fixation rate beyond what is expected in the absence of selection; it can only accelerate the time to equilibrium. But the effects of hitchhiking are countered by recombination, which breaks up linkage groups and pushes the time to equilibrium toward $2N_e$ generations.

If recombination were high enough in this system, therefore, few if any of the substitutions in this population would be expected to be neutral. This gives us a lower limit of zero for neutral substitutions. What can we say about the upper limit? The maximum assumes that the neutral substitution rate reached equilibrium immediately. We can calculate the expected equilibrium rate of neutral substitutions—the genomic neutral mutation rate—if we know the genomic mutation rate and the proportion of mutations that are neutral. ϕX174 mutation rates have been estimated for a few sites on

the basis of reversions of amber mutations at $\sim 2 \times 10^{-7}$ (FERSHT and KNILL-JONES 1981), which is 1/5 that of Drake's generic estimate of one mutation per genome per 200 replications for DNA genomes ($\sim 10^{-6}$; DRAKE 1999). Using these estimates and assuming that ϕX174 replicates from the infecting genome as template, this would maximally yield one to five substitutions per 1000 phage generations if *all* changes were neutral. If we guess that 1/3 of mutations are nearly neutral, this would set an upper limit of 0.33–1.67 neutral substitutions per 1000 phage generations, or 4–21 over the course of this experiment. This gives us a rough upper limit if the population instantly reached the equilibrium rate of neutral mutation due to hitchhiking and hence with low rates of recombination and without clonal interference. In view of the evidence for recombination presented below, this upper limit is an overestimate.

If this neutral calculation could be trusted, it would suggest that most of the substitutions in this system were adaptive. However, the uncertainty at every step in the calculation warrants other evidence on the issue. We thus address indicators of the nature of selection in this system.

Chemostat design and the enigma of selection for faster growth: Our intent in developing this culture system was to create an environment with a continual surplus of cells, whereby selection was primarily toward faster phage growth rate. At face value, the conclusion that most changes were adaptive is consistent with the goal of the chemostat design, and previous work identified the phenotypic bases of several substitutions with growth-enhancing effects (BULL *et al.* 2000; CRILL *et al.* 2000). However, two contradictions challenge an interpretation of the chemostat as primarily selection for faster growth in a constant environment.

The first contradiction concerns the tempo and absolute magnitude of change. The expectation for an adaptive walk in a constant environment is that fitness should rise most rapidly early in the process and then plateau (KAUFFMAN and LEVIN 1987; GERRISH and LENSKI 1998; ORR 1998; DE VISSER and LENSKI 2002). Likewise, there is an expectation that the number of adaptive substitutions should plateau (neutral changes would continue), but no basis for predicting the total number of beneficial changes without understanding the shape of the fitness landscape and where the starting genotype lies in this landscape (VOIGT *et al.* 2000). As noted above, we have previously observed 12–15 changes in the first 10 days of evolution in a strongly selective environment (BULL *et al.* 1997; WICHMAN *et al.* 1999, 2000). In this experiment we observed only a single change during the first 10 days of evolution, followed by a near-linear increase in substitutions averaging 11/20 days throughout the entire 180 days. Thus the tempo of evolution was not as expected for adaptation to a constant environment. The magnitude of change that could be accounted for solely by faster growth would then depend
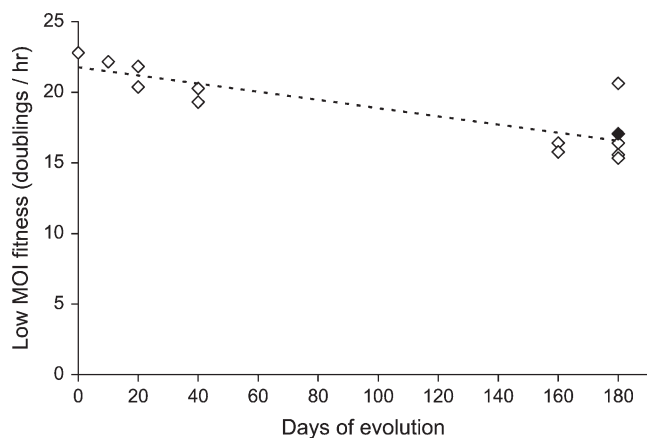
FIGURE 6.—Phage growth rate for selected isolates in the φX174 chemostat line through 180 days of evolution. In this assay, the rate of phage population increase is measured in a culture where hosts outnumber phage, thus representing phage ability to increase without interference from other phage. The decline with time is statistically significant. (The correlation between growth rate and day is negative and highly significant: $\rho = -0.84$, $P < 0.001$; the correlation used the average growth rate of each phage isolate and the day of its isolation for the isolates shown.) The diamond at day 0 indicates the mean fitness of the φX174 ancestor. Open diamonds indicate the mean growth rates of isolates, and the solid diamond indicates the growth rate from a large sample of the day 180 population. The means of the four isolates at day 180 are significantly heterogeneous, with one mean being substantially above the others.

on how far the ancestral φX174 started from the local fitness peak. In previous work, φX174 was adapted to an inhibitory high temperature (44°) with a serial transfer protocol. Under these strongly selective conditions, fitness reached a plateau after a mere three substitutions per genome (HOLDER and BULL 2001). Although differences between the two protocols could account for some differences in the outcomes, the 92–100 substitutions obtained here thus seem excessive for adaptation to a constant environment, especially given that the conditions here were less inhibitory to the virus than previous conditions.

The second contradiction to selection for fast growth rate comes from assays of phage growth rate throughout the experiment. Phage growth rate actually declined during the 180-day adaptation (Figure 6; growth rate is measured with an excess of hosts).

High levels of within-host competition: If a chemostat does not primarily select faster phage growth, what might be the nature of selection? Various lines of evidence support the existence of high levels of co-infection. Co-infection selects the ability to compete with different genomes in the same cell, and this selection can cause the virus to evolve a reduced ability to grow when hosts are abundant (TURNER and CHAO 1998). Furthermore, there is not necessarily an optimal genotype with co-infection, so evolution can potentially con-

tinue indefinitely—as a new genotype comes to prominence, its prevalence provides the foundation on which yet other mutants may gain an advantage. We thus cautiously interpret our results as supporting a model in which sustained molecular evolution is driven by phage-phage competition within cells. This model of selection in our system is supported by four lines of evidence, as described below.

High titers: Titers taken at 5- to 10-day intervals usually exceeded 10^9/ml (data not shown). Since the chemostat was maintained in a relatively steady state, the rate of co-infection can be determined from this value alone: the per-minute rate at which phages adsorb to individual cells is given by δP, where P is the phage concentration and δ is the adsorption rate constant. δ typically lies in the range of 10^{-8}–10^{-9} for sensitive cells across most phages (ADAMS 1959). With P maintained in excess of 10^9/ml, the multiple infection of single cells within minutes is ensured. Superinfection exclusion (whereby the expression and replication of an infecting genome is suppressed by phage genomes that entered the cell earlier) does not begin until 5–10 min after the first infection (HUTCHISON and SINSHEIMER 1971), so this rate of phage infection is more than sufficient to ensure competition among coinfecting genomes.

Decline in growth rate when hosts are in excess: It was noted above that phage growth rate declined over the course of the evolution. A decline in φX174 growth rate was also noted previously in a 10-day, high-temperature, Salmonella chemostat line (WICHMAN et al. 1999). These declines provide direct evidence that selection in the chemostat was not simply to maximize phage growth rate. A similar decline in growth rate was observed in φ6 when the virus was grown at high multiplicity but not at low multiplicity (TURNER and CHAO 1998, 1999).

Heterogeneity in growth rate: The four isolates at day 180 were significantly heterogeneous in growth rate, with one isolate substantially higher than the others (Figure 6 shows the means of four isolates; $F(3,8) = 33.7$, $P < 0.001$). As with the decline in growth rate when hosts are in excess, this heterogeneity suggests that other components in addition to growth rate were important to fitness. It is noteworthy that the growth rates of three isolates were low, indicating that the low growth rate was not likely the result of a spontaneous mutation.

Evidence of recombination: In φX174, recombination can occur between the double-stranded replication intermediates and requires co-infection by different phage types (TESSMAN and SHLESER 1963; BAKER et al. 1971). Absence of recombination would be indicated by strong linkage disequilibrium of polymorphic sites. In this experiment, some isolates taken from a single time point differed by >20 changes, but at many time points these polymorphisms did not exhibit complete linkage disequilibrium (Figure 7). Polymorphic sites frequently exhibited linkage disequilibrium over large contiguous tracts, but different associations in adjacent tracts. Multi-
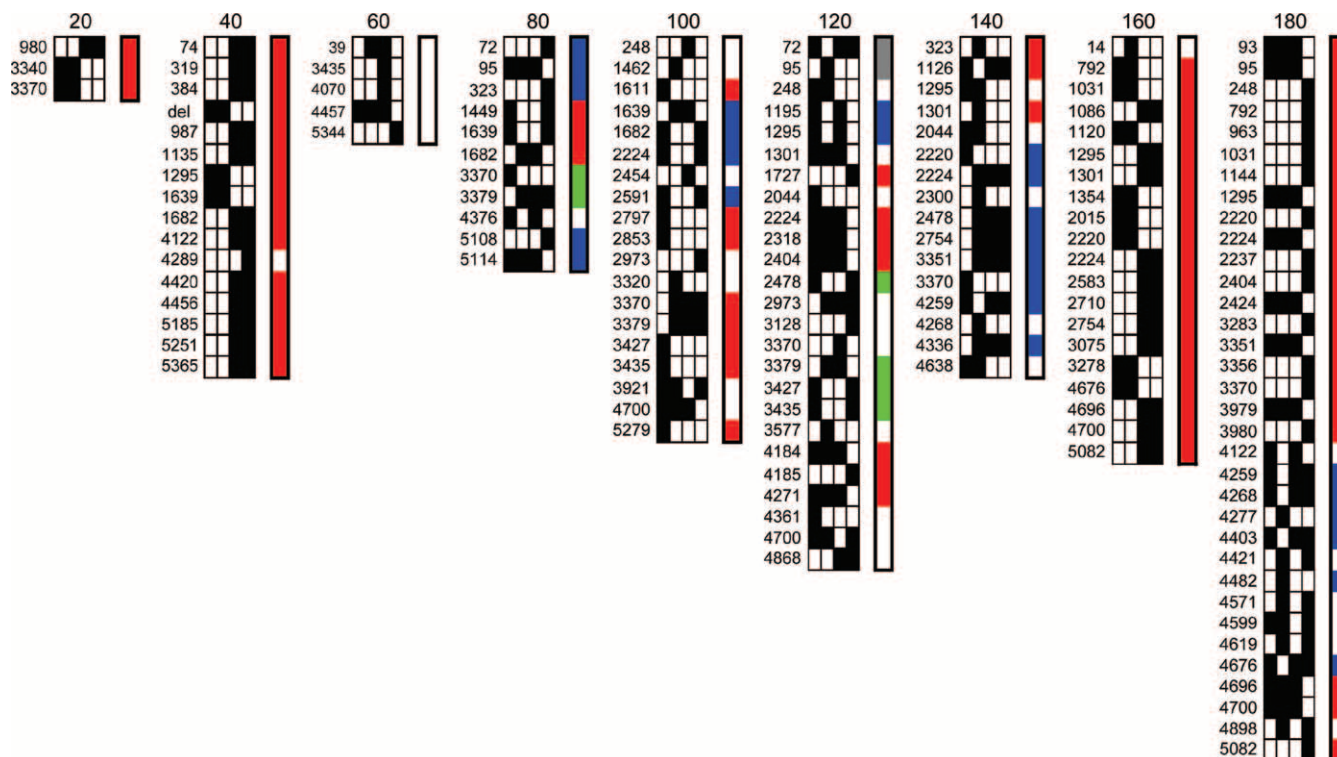
Figure 7.—Schematic of variation between isolates in the φX174 chemostat line through 180 days. Each time point is represented by four isolates and a bar to the right. Each isolate is represented by a column of squares showing sites variable at that time point; the order of isolates has been arranged for ease of viewing and differs from Figure 1. Differences from the ancestor are in black, and ancestral states are in white. At least two adjacent squares exhibiting the same pattern of association among isolates are indicated by a colored block in the column to the right of the isolates (white otherwise). Other sites showing that pattern on that day are indicated in the same color if they are not interrupted by another tract, or in white if they are singletons within another tract. This is a conservative measure of the number of tracts. (There is no relationship among bars of the same color across days because most polymorphic sites become fixed or are lost between time points and recombination can break apart the linkage of those that remain.) Patterns are considered to be consistent between variable sites if they are the same as or are the mirror image of each other. For example, □□■■ is considered the same pattern as ■■□□. A continuous bar of one color at the right thus indicates that two distinct genotypes are present in the population at that time point (days 20, 40, and 160). Changing patterns of color are consistent with recombination between tracts. Recombination is most dramatic at day 120, but also is evident at days 80, 100, 140, and 180.

ple tracts in the same genome are inconsistent with a single haplotype. This pattern of mixed association among variable sites is compatible with a buildup of alternate types over some or all of the genome (as seen at day 40 and at day 160), followed by recombination during co-infection. Note also that the linkage association of sites that remained polymorphic for >20 days could be different across time points. While we cannot formally rule out numerous same-site mutation events without knowing mutation rates and selective coefficients of the variants involved, this pattern is highly suggestive of intermediate levels of recombination (but not so much that linkage disequilibrium is destroyed) and possibly compatible with the rates of 10^{-4}–10^{-5} per co-infection observed by TESSMAN and SHLESER (1963). Additional evidence for recombination was detected in the 120R experiment: by starting the replicate with known isolates and sequencing multiple isolates after 20 days of evolution, we could detect three instances in which the linked association of adjacent markers in the founding genomes had been altered. Other instances of changes in marker association in 120R were observed as well.

The difference between what we had hoped to achieve in the chemostat and what we now think operates has many similarities to the difference between r- and K-selection or density-dependent *vs.* density-independent regulation in ecology (see FLEGR 1997 for history and extensions of the meanings of r- and K-selection in flow-through cultures). In its original formulation, r-selection represented selection for rapid growth, as would occur in populations whose numbers were maintained well below carrying capacity by extrinsic factors that were largely nonselective; K-selection represented selection at high density, dominated by competition for resources. Our environments approximated a type of K-selection. We suggest, therefore, that the high phage density promoted continuing adaptive evolution, where the adaptation was specifically to the presence of competing genomes in the cell.

In some respects our study mirrors results from earlier studies of chemostats and other protocols of high-density passage with microbes (reviewed in ADAMS 2004). Many studies have shown a failure to exhibit progressive increases in fitness measured as growth rate (HELLING *et al.* 1987; ROSENZWEIG *et al.* 1994; TURNER and CHAO 1998; ROZEN and LENSKI 2000; DE VISSER and LENSKI 2002), and even in fact a decline in fitness over time (PAQUIN and ADAMS 1983; TURNER and CHAO 1998, 1999) as we observed. Another common feature is the maintenance of phenotypic variants (HELLING *et al.* 1987; KURLANDZKA *et al.* 1991; ROSENZWEIG *et al.* 1994; ROZEN and LENSKI 2000). In our study, competing variants at two time points could be inferred from sequence analysis (Figure 7, days 40 and 160), but the phenotypic differences are as yet unknown. High-density passaging can give rise to defective viral genomes (HUANG 1973; BARRETT and DIMMOCK 1986; DEPOLO *et al.* 1987; ROUX *et al.* 1991; BANGHAM and KIRKWOOD 1993), which in turn can drive continuing molecular evolution (DEPOLO *et al.* 1987). Our observation of a high substitution rate is compatible with the sequence analysis of vesicular stomatitis virus (VSV) and its defective particle (DEPOLO *et al.* 1987); these viral studies contrast with the low substitution rate seen in bacteria passaged 20,000 generations (LENSKI *et al.* 2003).

Whole-genome perspective: For most studies of experimental evolution it is not feasible to sequence the entire genome even once, much less at multiple intervals throughout the study. Here, we ask how the snapshot of evolution that we see in this study might have been different if we had been forced to examine a single gene or region of the genome. We illustrate this with two regions that would have been logical choices for detailed examination if only a subset of the genome were sequenced. Protein A is involved in replication and thus would be a likely target of the selection for higher growth rate that we expected in this system. Likewise, proteins F, G, and H are known to be involved in attachment, which is another likely target in selection for higher growth rate.

Gene A makes up 29% of the genome but accounted for nearly half the changes (Figure 3, top). Evolution in the nonoverlapping region of A (Figure 3, middle) began slowly, and by the end silent changes were more than twice as numerous as missense changes. Taken alone, analysis of gene A would have supported the view that most substitutions during the latter half of this experiment were neutral and that the excess of silent substitutions was driven by strong purifying selection. In contrast, missense changes in genes F, G, and H increased in a near-linear fashion over time and were never less than silent changes in those genes (Figure 3, bottom), but d_N/d_S was <1 after day 40—again calling into question how much of the observed evolution is adaptive. The most striking aspect of this study—the continuous clock-like rate of substitutions shown in Fig-

ure 2—is less evident when a subset of the genome is examined. It is only when viewed from a total genome perspective that the clock-like evolution, the similar slopes of missense and silent rates, and the pronounced clustering of silent substitutions become evident.

Conclusions: Even if much of the continuing molecular evolution in this system can be attributed to the high-density environment, it will be a far greater challenge to partition the evolution into various underlying causes such as the types of selection and the specific functional effects of different mutations. Despite these difficulties, however, systems like the present one can continue to yield new insights about molecular evolution in nature. This study raises the interesting question of how much molecular evolution is driven by intraspecific competition compared to other aspects of the biotic and the abiotic environment. A second noteworthy question is whether the clock-like rate of substitutions on a genomic scale is a general feature of adaptive molecular evolution, and if so what factors regulate the tempo of the clock. This system is especially amenable to further exploration of these questions not only because of the rapid generation time and ease of sequencing ϕX174 on a genomic scale, but also because of the ability to apply reverse genetics, molecular biology, and structural analysis. We thus envision many opportunities to exploit this system for further evolutionary insight.

LITERATURE CITED

ABEDON, S. T., T. D. HERSCHLER and D. STOPAR, 2001 Bacteriophage latent-period evolution as a response to resource availability. Appl. Environ. Microbiol. **67:** 4233–4241.

ADAMS, J., 2004 Microbial evolution in laboratory environments. Res. Microbiol. **155:** 311–318.

ADAMS, M. H., 1959 *Bacteriophages*. Interscience, New York.

BAKER, R., J. DONIGER and I. TESSMAN, 1971 Roles of parental and progeny DNA in 2 mechanisms of phage S13 recombination. Nat. New Biol. **230:** 23–25.

BANGHAM, C. R., and T. B. KIRKWOOD, 1993 Defective interfering particles and virus evolution. Trends Microbiol. **1:** 260–264.

BARRETT, A. D., and N. J. DIMMOCK, 1986 Defective interfering viruses and infections of animals. Curr. Top. Microbiol. Immunol. **128:** 55–84.

BERNAL, R. A., S. HAFENSTEIN, R. ESMERALDA, B. A. FANE and M. G. ROSSMANN, 2004 The ϕX174 protein J mediates DNA packaging and viral attachment to host cells. J. Mol. Biol. **337:** 1109–1122.

BIRKY, C. W., JR., and J. B. WALSH, 1988 Effects of linkage on rates of molecular evolution. Proc. Natl. Acad. Sci. USA **85:** 6414–6418.

BULL, J. J., M. R. BADGETT, H. A. WICHMAN, J. P. HUELSENBECK, D. M. HILLIS *et al.*, 1997 Exceptional convergent evolution in a virus. Genetics **147:** 1497–1507.

BULL, J. J., M. R. BADGETT and H. A. WICHMAN, 2000 Big-benefit mutations in a bacteriophage inhibited with heat. Mol. Biol. Evol. **17:** 942–950.

BULL, J. J., M. R. BADGETT, D. ROKYTA and I. J. MOLINEUX, 2003 Experimental evolution yields hundreds of mutations in a functional viral genome. J. Mol. Evol. **57:** 241–248.

BURCH, A. D., and B. A. FANE, 2003 Genetic analyses of putative conformation switching and cross-species inhibitory domains in Microviridae external scaffolding proteins. Virology **310:** 64–71.

BURCH, A. D., J. TA and B. A. FANE, 1999 Cross-functional analysis of the Microviridae internal scaffolding protein. J. Mol. Biol. **286:** 95–104.

CRILL, W. D., H. A. WICHMAN and J. J. BULL, 2000 Evolutionary reversals during viral adaptation to alternating hosts. Genetics **154:** 27–37.

CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory.* Harper & Row, New York.

DENHARDT, D. T., and R. L. SINSHEIMER, 1965 The process of infection with bacteriophage φX174. 3. Phage maturation and lysis after synchronized infection. J. Mol. Biol. **12:** 641–646.

DEPOLO, N. J., C. GIACHETTI and J. J. HOLLAND, 1987 Continuing coevolution of virus and defective interfering particles and of viral genome sequences during undiluted passages: virus mutants exhibiting nearly complete resistance to formerly dominant defective interfering particles. J. Virol. **61:** 454–464.

DE VISSER, J. A., and R. E. LENSKI, 2002 Long-term experimental evolution in *Escherichia coli.* XI. Rejection of non-transitive interactions as cause of declining rate of adaptation. BMC Evol. Biol. **2:** 19.

DOKLAND, T., R. MCKENNA, L. L. ILAG, B. R. BOWMAN, N. L. INCARDONA *et al.*, 1997 Structure of a viral procapsid with molecular scaffolding. Nature **389:** 308–313.

DOKLAND, T., R. A. BERNAL, A. BURCH, S. PLETNEV, B. A. FANE *et al.*, 1999 The role of scaffolding proteins in the assembly of the small, single-stranded DNA virus φX174. J. Mol. Biol. **288:** 595–608.

DRAKE, J. W., 1999 The distribution of rates of spontaneous mutation over viruses, prokaryotes, and eukaryotes. Ann. NY Acad. Sci. **870:** 100–107.

FERSHT, A. R., and J. W. KNILL-JONES, 1981 DNA polymerase accuracy and spontaneous mutation rates: frequencies of purine.purine, purine.pyrimidine, and pyrimidine.pyrimidine mismatches during DNA replication. Proc. Natl. Acad. Sci. USA **78:** 4251–4255.

FISHER, R. A., 1958 *The Genetical Theory of Natural Selection.* Dover, New York.

FLEGR, J., 1997 Two distinct types of natural selection in turbidostat-like and chemostat-like ecosystems. J. Theor. Biol. **188:** 121–126.

GERRISH, P. J., and R. E. LENSKI, 1998 The fate of competing beneficial mutations in an asexual population. Genetica **102/103:** 127–144.

GILLESPIE, J. H., 1983 A simple stochastic gene substitution model. Theor. Popul. Biol. **23:** 202–215.

GILLESPIE, J. H., 1984 Molecular evolution over the mutational landscape. Evolution **38:** 1116–1129.

GILLESPIE, J. H., 1991 *The Causes of Molecular Evolution.* Oxford University Press, New York.

HAYASHI, M., A. AOYAMA, D. L. J. RICHARDSON and M. N. HAYASHI, 1988 Biology of the bacteriophage φX174, pp. 1–17 in *The Bacteriophages,* edited by R. CALENDAR. Plenum Press, New York.

HELLING, R. B., C. N. VARGAS and J. ADAMS, 1987 Evolution of *Escherichia coli* during growth in a constant environment. Genetics **116:** 349–358.

HOLDER, K. K., and J. J. BULL, 2001 Profiles of adaptation in two similar viruses. Genetics **159:** 1393–1404.

HUANG, A. S., 1973 Defective interfering viruses. Annu. Rev. Microbiol. **27:** 101–117.

HUTCHISON, C. A., III, and R. L. SINSHEIMER, 1963 Kinetics of bacteriophage release by single cells of φX174-infected *E. coli.* J. Mol. Biol. **41:** 206–208.

HUTCHISON, C. A., III, and R. L. SINSHEIMER, 1971 Requirement of protein synthesis for bacteriophage φX174 superinfection exclusion. J. Virol. **8:** 121–124.

KAUFFMAN, S., and S. LEVIN, 1987 Towards a general theory of adaptive walks on rugged landscapes. J. Theor. Biol. **128:** 11–45.

KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge, UK.

KURLANDZKA, A., R. F. ROSENZWEIG and J. ADAMS, 1991 Identification of adaptive changes in an evolving population of *Escherichia coli:* the role of changes with regulatory and highly pleiotropic effects. Mol. Biol. Evol. **8:** 261–281.

LENSKI, R. E., C. L. WINKWORTH and M. A. RILEY, 2003 Rates of DNA sequence evolution in experimental populations of *Escherichia coli* during 20,000 generations. J. Mol. Evol. **56:** 498–508.

LI, W.-H., 1997 *Molecular Evolution.* Sinauer Associates, Sunderland, MA.

NEI, M., 1987 *Molecular Evolutionary Genetics.* Columbia University Press, New York.

NOVAK, C. R., and B. A. FANE, 2004 The functions of the N terminus of the φX174 internal scaffolding protein, a protein encoded in an overlapping reading frame in a two scaffolding protein system. J. Mol. Biol. **335:** 383–390.

ORR, H. A., 1998 The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution **52:** 935–949.

ORR, H. A., 2002 The population genetics of adaptation: the adaptation of DNA sequences. Evolution **56:** 1317–1330.

PAQUIN, C. E., and J. ADAMS, 1983 Relative fitness can decrease in evolving asexual populations of *S. cerevisiae.* Nature **306:** 368–371.

ROSENZWEIG, R. F., R. R. SHARP, D. S. TREVES and J. ADAMS, 1994 Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli.* Genetics **137:** 903–917.

ROUX, L., A. E. SIMON and J. J. HOLLAND, 1991 Effects of defective interfering viruses on virus replication and pathogenesis in vitro and in vivo. Adv. Virus Res. **40:** 181–211.

ROZEN, D. E., and R. E. LENSKI, 2000 Long-term experimental evolution in *Escherichia coli.* VIII. Dynamics of a balanced polymorphism. Am. Nat. **155:** 24–35.

SANGER, F., G. M. AIR, B. G. BARRELL, N. L. BROWN, A. R. COULSON *et al.*, 1977 Nucleotide sequence of bacteriophage φX174 DNA. Nature **265:** 687–695.

SHARP, P. M., and W. H. LI, 1987 The Codon Adaptation Index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. **15:** 1281–1295.

TESSMAN, E. S., and R. SHLESER, 1963 Genetic recombination between phages S13 and φX174. Virology **19:** 239–240.

TURNER, P. E., and L. CHAO, 1998 Sex and the evolution of intrahost competition in RNA virus φ6. Genetics **150:** 523–532.

TURNER, P. E., and L. CHAO, 1999 Prisoner's dilemma in an RNA virus. Nature **398:** 441–443.

VOIGT, C. A., S. KAUFFMAN and Z. G. WANG, 2000 Rational evolutionary design: the theory of in vitro protein evolution. Adv. Protein Chem. **55:** 79–160.

WICHMAN, H. A., M. R. BADGETT, L. A. SCOTT, C. M. BOULIANNE and J. J. BULL, 1999 Different trajectories of parallel evolution during viral adaptation. Science **285:** 422–424.

WICHMAN, H. A., L. A. SCOTT, C. D. YARBER and J. J. BULL, 2000 Experimental evolution recapitulates natural evolution. Philos. Trans. R. Soc. Lond. B Biol. Sci. **355:** 1677–1684.

Communicating editor: H. OCHMAN