

REVIEW

Predicting evolution from genomics: experimental evolution of bacteriophage T7

JJ Bull¹ and IJ Molineux²

¹*Integrative Biology, Institute for Cellular and Molecular Biology, University of Texas, Austin, TX, USA* and ²*Molecular Genetics and Microbiology, Institute for Cellular and Molecular Biology, University of Texas, Austin, TX, USA*

A wealth of molecular biology has been exploited in designing and interpreting experimental evolution studies with bacteriophage T7. The modest size of its genome (40 kb dsDNA) and the ease of making genetic constructs, combined with the many genetic resources for its host (*Escherichia coli*), have enabled comprehensive and detailed studies of experimental adaptations. In several studies, the genome was specifically altered (gene knockouts, gene replacements, reordering of genetic elements) such that *a priori* knowledge of genetics and biochemistry of the phage could be used to predict the pathways of compensatory evolution when the modified phage is adapted to recover fitness. In other work, the phage has been adapted to specific environmental conditions chosen to select phenotypic outcomes with a quantitative basis, and the

molecular bases of that evolution have been explored. Predicting the outcomes of these adaptations has been challenging. In hindsight, one-third to one-half of the compensatory nucleotide changes observed during the adaptation can be rationalized based on T7 biology. This rationalization usually only applies at the genetic level—a gene product may be known to be involved in the affected pathway, but it usually remains unknown how the observed change affects activity. The progress is encouraging, but the prediction of experimental evolution pathways remains far from complete, and is still sometimes confounded by observation when an adaptation yields a completely unexpected outcome.

Heredity (2008) **100**, 453–463; doi:10.1038/sj.hdy.6801087; published online 23 January 2008

Keywords: experimental evolution; bacteriophage; genomics; sequence; adaptation

The field of genomics promises to provide the long sought genotype–phenotype map for complex traits in many organisms. Armed with a thorough understanding of how phenotypes are genetically determined, the evolutionary biology of adaptation could become a truly predictive science, in which we anticipate which genes will evolve to change function in specific environments. Programs of this sort have already been initiated for simple phenotypes (for example, those involving single genes) and in studies of bacterial metabolism, where a knowledge of molecular pathways paired with metabolic flux theory has yielded unique insights into evolution in response to the nutrients available to the cell (Dean *et al.*, 1986; Dykhuizen *et al.*, 1987; Ibarra *et al.*, 2002; Lunzer *et al.*, 2002; Fong *et al.*, 2003; Zhong *et al.*, 2004; Dekel and Alon, 2005; Fong *et al.*, 2005a). For most phenotypes and most organisms, however, this approach has been inaccessible, but we appear to be at the dawn of an era in which these studies may become commonplace. It is thus instructive to review what has been accomplished so far, to foreshadow what difficulties lie ahead and to appreciate what will be needed to render the approach viable.

This review summarizes a set of experimental adaptations with the bacteriophage T7. The experiments were

designed to exploit the well-known genotype–phenotype map of the phage to understand adaptation. In many cases, the challenge put before the phage was a targeted disruption of the genome whose effects led to clear predictions and a hoped-for easy interpretation of the compensatory evolution at the level of genes and gene networks. In this respect, the T7 work is similar to experimental evolution of bacterial metabolic pathways (see below) and broader than experiments limited to phenotypic or fitness assays (Chao and Tran, 1997; Lenski *et al.*, 1998; Papadopoulos *et al.*, 1999; Turner and Chao, 1999; Burch and Chao, 2000; Riley *et al.*, 2001). In different studies, phenotypic optimality models were tested by adapting the wild-type virus to different environmental conditions that had been predicted to select specific phenotypic changes, and the phenotypic and genetic bases of the evolution were studied.

Bacteriophage T7

T7 is a strictly lytic phage; hence its entry into a bacterial (host) cell is quickly followed by the death of that host to release phage progeny. Details of T7 biology are known to varying degrees for different genes (Dunn and Studier, 1983; Molineux, 2006). Aside from enzymes providing biosynthetic capacity, the host provides few proteins that are essential for T7 development. T7 also has the advantage in this work of being a largely modular phage, with minor overlap of genes and other genetic elements. Most importantly, the phage encodes its own RNA

Correspondence: Dr JJ Bull, Integrative Biology C0930, Institute for Cellular and Molecular Biology, University of Texas, 1 University Station, Austin, TX 78712, USA.

E-mail: bull@mail.utexas.edu

Received 11 June 2007; accepted 26 September 2007; published online 23 January 2008

polymerase (RNAP) and phage-specific promoters, which are responsible for transcription of all middle genes (primarily DNA metabolism) and late genes (primarily for virion assembly). T7 RNAP is also required for packaging progeny DNA into virions and for initiating replication. Two promoters, one at each end of the genome are thought to be used in cutting unit-length genomes or as replication origins and are thus not primarily involved in gene expression. This expression system, combined with many *in vitro* studies of transcription from the different promoters, has led to a strong sense that the T7 regulatory network is well understood, culminating in an empirically parameterized virtual model that qualitatively captures many aspects of the infection cycle (v2.5; Endy *et al.*, 2000). The high fitness and the relatively small 40 kb dsDNA genome of T7 further suits it to a combination of genomic manipulations, easy propagation and full-genome sequencing to determine all adaptive changes.

Fitness is commonly measured in studies of experimental evolution, but the measure used for fitness often varies. The measure used in most of the T7 work has been the intrinsic rate of increase (the ecologist's *r*), usually expressed as doublings per hour of the phage population. This measure is highly dependent on the environment, most notably on the host cell type, cell physiology and cell density. The measure makes little sense unless the phage is grown while surrounded by an excess of hosts; however, when the same environment is used across different phage strains, the intrinsic rate of increase is a fitness measure that is directly comparable among phages, regardless of differences in phage generation time, burst size or other fitness component.

Compensatory evolution

The design of the following studies has been to adapt a phage in response to some genomic or environmental manipulation and interpret the results. In such work, it is important to discriminate mutations (substitutions) that are beneficial specifically in response to the manipulation (and thus 'compensatory') from those that are beneficial in more general circumstances. For example, if a phage is 'pre-adapted' to the culture conditions to be used in an experimental treatment, many beneficial substitutions are typically observed. If the treatment is instead started before the pre-adaptation, then many of those same mutations will evolve, but have nothing to do with the experimental treatment being investigated, and it is important to distinguish them. The methodology is simple. Suppose that two substitutions, A and B, evolve when an organism is adapted to a drug. Did both evolve in response to the drug? By recombining the beginning and end points of the adaptation, all four possible genetic combinations are created: AB, Ab, aB and ab. If this genetically mixed population is then propagated in the absence of the drug, any mutation whose benefit is general will ascend; any mutation that is beneficial only in response to the drug will not. Recombination ensures that each mutation is selected independently of the others. T7 is extremely well suited to this test, rates of recombination are much higher than its bacterial host, and the procedure has been routinely used in various studies.

Experiments in regulatory evolution

Four experiments have challenged T7 in ways that favor changes in genome regulation. Two of these were designed *a priori* as regulatory evolution experiments, two were reinterpreted as such *a posteriori*.

RNAP exchange (Bull *et al.*, 2007)

The phage RNAP is essential for phage reproduction. Early genes are those expressed by the host RNAP and lie in the first ~20% of the slowly entering phage genome, but expression of most phage genes—the middle and late genes—is due to the phage RNAP. This study took advantage of prior knowledge that T7 and T3 are close relatives, with similar genome organization (except for some non-essential genes) and infection strategy, but the two phages have diverged so that the RNAP of one phage transcribes only at low levels from the promoters of the other phage. The experimental design forced T7 to use T3 RNAP. This RNAP exchange was accomplished by deleting the RNAP gene from T7 and supplying T3 RNAP *in trans* as part of the host. Although the T3 RNAP gene could have been inserted into the T7 backbone instead of the host genome, doing so would have enabled the phage to evolve changes in T3 RNAP rather than force a re-evolution of regulation throughout the genome.

Discrimination between T7 and T3 promoters by the T3 RNAP is attributed primarily to a 3-base region in the promoter, but a G→C change at position -11 is sufficient to confer moderate expression by T3 RNAP (Klement *et al.*, 1990). The expected outcome in this study was thus that T7 would adapt to T3 RNAP by a series of G→C changes at the -11 position of many promoters.

There were only three predicted promoter changes with a solid biological basis: activation of any of the promoters upstream of gene 2 (the first essential gene normally transcribed by the phage RNAP) would be required to express all DNA metabolism genes. Activation of the late promoters for the genes for the scaffolding and major capsid proteins (genes 9 and 10) would also be needed to ensure high expression of the two proteins known to be needed at highest levels for progeny phage production. It was also expected that one or both of the terminally located replication promoters would evolve.

The phage RNAP is also known to interact with two phage proteins, lysozyme (effecting a change in preference by the RNAP for late over middle promoters), and the large subunit of terminase (Zhang and Studier, 1997, 2004; Molineux, 2006). It was thus anticipated that changes in both these genes would also evolve during the adaptation.

Predictions at the level of fitness were less clear. Fitness of T7 was expected to be low at the outset. Yet given that a single mutation in the promoter can apparently confer some expression by T3 RNAP, it seemed likely that the phage would evolve to recover most of its lost fitness. This expectation had to be mitigated somewhat by a lack of understanding of the molecular details of phage RNAP–protein interactions. First, it was not known how effectively T3 RNAP would interact with T7 lysozyme and terminase. Furthermore, and in contrast to the understanding of promoter function, there was no basis for predicting how

Table 1 T7 compensatory evolution in response to T3 RNAP

Expected evolution	Observed evolution
(1) G→C at position -11 of potentially all 16 promoters, but at least one upstream of 2 and promoters for genes 9 and 10	(1) Changes in nine promoters, mostly at positions -11 and -2; G→A most common at -11
(2) Missense changes in lysozyme and terminase genes	(2) Missense changes in lysozyme and terminase genes (3) Missense changes in four essential, two non-essential genes with no connection to RNAP

Abbreviation: RNAP, RNA polymerase.

completely compensatory evolution could correct any misfits in these interactions.

Adaptation of a single T7 line to T3 RNAP (Bull *et al.*, 2007) supported many of the expectations (Table 1). Fitness recovery was profound (Figure 1). For comparison, a T3 deleted of its RNAP gene was adapted to the same host. The fitness of this T3 serves as an empirical upper limit for the T7 phage because T3 genetic elements are already evolved for interaction with T3 RNAP. By this criterion, T7 did not reach the maximum attainable fitness, but the fitness trajectory suggests that it may have done so (Figure 1). Nonetheless, at this point there was sufficient fitness gain to anticipate the accumulation of many changes for testing our predictions.

Over 30 changes were observed in the evolved genome, most of which were compensatory for the novel RNAP (Bull *et al.*, 2007). As expected, many promoters acquired changes, but the results were surprising in two aspects. First, only just over half of the promoters acquired changes, including several upstream of gene 2 and those for genes 9 and 10. Second, the most common positions to evolve were -11 and -2, but the most common change at -11 was G→A, not the G→C that was expected from previous biochemical work (Klement *et al.*, 1990). Third, the phage evolved changes in the two genes whose proteins are known to interact with the RNAP (lysozyme, terminase large subunit; Zhang and Studier, 1997, 2004), but six other compensatory genic changes were also observed with no obvious role in regulation or interaction with RNAP.

Overall, this adaptation supports the thesis that T7 is capable of substantial regulatory evolution, and that the current understanding of T7 biology contributes substantially but incompletely to the understanding of adaptive evolution. In particular, it should be noted that the mere presence of a regulatory element does not imply it has a major impact on fitness. Second, *in vitro* characterizations do not necessarily translate exactly to the *in vivo* situation.

Genome rearrangement (Springman *et al.*, 2005)

As T7 is obligately lytic and the growth conditions used in these experiments favor a short generation time, there is a fitness premium on rapid genome entry into the host and rapid gene expression. Entry and expression are coupled, as most of the T7 genome are internalized via transcription (Molineux, 2006). In concert with this

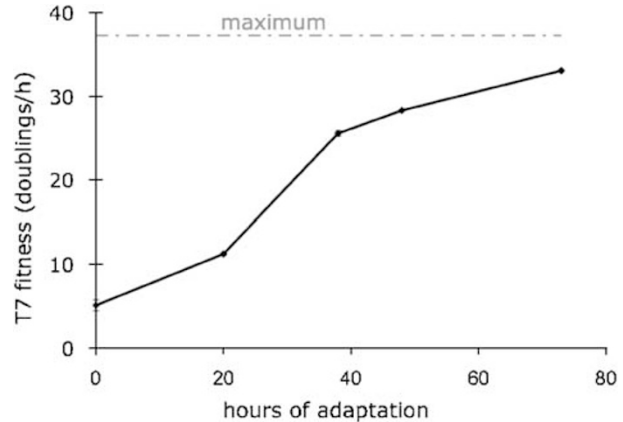


Figure 1 Adaptation of T7 to T3 RNA polymerase (RNAP). T7 was deleted of its RNAP gene and grown on a host supplying the RNAP from T3. Phage fitness was initially quite low, but after many hours of passage, it recovered to nearly maximal levels. The black line indicates the fitness of T7 during the adaptation; the horizontal gray line gives the presumed maximum (a T3 deleted of its RNAP gene adapted to the host providing T3 RNAP *in trans*). Although the T7 line did not reach the maximum, its trajectory suggests that it may have done so upon further transfer. Fitness is measured as doublings per hour of the phage concentration.

premium on rapid genome entry and expression, the rate-limiting step in the phage life cycle following attachment to the host is expression of the phage RNAP gene. The position of the RNAP gene in the early region of the genome is therefore critical to its early expression. *Escherichia coli* RNAP is necessary to transcribe the phage RNAP gene, and shifting the latter's location away from the entering end of the genome (referred to as the genetic 'left' end) delays the life cycle and correspondingly reduces fitness (Endy *et al.*, 2000).

The RNAP gene is located 8–14% of the genome length from the left end in the wild-type phage. As a test of the T7 virtual model, several constructs were made in which the RNAP gene was moved further from the left end of the genome. The most extreme shift was to 66–72% from the left end, and the predicted delays in life cycle were qualitatively, albeit not quantitatively, matched by the observed delays (Endy *et al.*, 2000).

Two questions arising from that study were whether a genome with an ectopically placed RNAP gene would evolve to recover fitness, and whether the observed evolution would match expectations. Two phages with an ectopic RNAP gene placed 66–72% from the left end were adapted for high fitness. One phage lacked the complete gene and flanking sequences at the original location, the second retained a 5' segment of the original RNAP gene at its normal location but lacked most of the coding region and sequences immediately 3' to the gene. Obviously, the simplest solution for the latter phage was to recombine the ectopic RNAP gene back in the wild-type location. T7 recombination is facilitated by a minimum of approximately 40 bp of identical sequence at each end of the donor and recipient regions. Although there was extensive sequence identity at the 5' end of the gene for recombination to restore the complete RNAP gene to its wild-type location, there was little at the 3' end. It was therefore considered unlikely that the RNAP gene would recombine back to its wild-type location in

either phage, and that evolution would be primarily to accommodate the RNAP gene at its ectopic location. For this latter scenario, only three evolutionary changes were expected during the adaptation (Table 2):

- (i) Abolish the pause generated by the entry protein. Upon virion attachment to the cell, the gene 16 protein, together with other virion proteins, creates a pore across the bacterial membranes and drives the genome part way into the cell (Kemp *et al.*, 2005). This process normally pauses when ~850 bp of the genome have entered (Garcia and Molineux, 1996). Those 850 bp include the three strong *E. coli* promoters, to which the host RNAP binds and begins transcribing. The second phase, host-mediated transcription pulls in the next ~15% of the genome, expressing the phage RNAP gene, but transcription by host RNAP is halted at the phage early terminator, TE. When the phage RNAP gene is moved far from the left end, *E. coli* RNAP must read through TE and transcribe ~50% of the phage genome before expressing the ectopic RNAP gene. This is a slow process. Mutations in gene 16 were already known that abolish this pause by the entry protein (Garcia and Molineux, 1996; Struthers-Schlinke *et al.*, 2000) and allow the entire T7 genome to enter the cell at almost twice the rate than by transcription with *E. coli* RNAP. Thus, those gene 16 mutations were expected as a way to ensure fast expression of the ectopic RNAP gene.
- (ii) Loss of the early terminator function. The phage is entirely dependent on *E. coli* RNAP to express the phage RNAP gene, no matter where it is located. With the phage RNAP gene moved downstream, *E. coli* RNAP must transcribe further into the phage genome than it normally does, but TE terminates most transcription from the three early promoters and thus limits expression of middle and late phage genes. As a potential benefit during adaptation (perhaps beneficial only until the gene 16 mutation in (i) occurred), it was expected that this early terminator activity would be lost, so that *E. coli* RNAP would transcribe more efficiently across the entire phage genome, and thus express the ectopic RNAP gene.
- (iii) Evolution of new *E. coli* promoter activity. Even if rapid genome entry is achieved by a gene 16 change, the phage still faces the problem that the only strong *E. coli* promoters lie near the left end of the genome and transcription from them must read through terminators to express the ectopic RNAP gene. Several weak *E. coli* promoters had been identified in the wild-type genome (Dunn and Studier, 1983), and it was expected that one or more of those also might evolve greater strength.

In three comparable adaptations, two retained the RNAP gene at its ectopic location, but in the one where recombination was formally possible (though considered unlikely) the gene was restored back to its wild-type location (Table 2; Springman *et al.*, 2005). However, the recombination did more than merely restore the RNAP gene to its wild-type location, it also exchanged the location of four contiguous late genes with nine early genes. Curiously, the recombination process used a

Table 2 T7 compensatory evolution in response to RNAP gene translocation

Expected evolution	Observed evolution: replicate 1
(1) Recombine RNAP gene back to wt location	(1a) Recombination of RNAP gene back to wt location, accompanied by exchange of four late genes with nine early genes (1b) 17 mutations in 11 genes & 1 promoter
OR	Observed evolution: replicates 2,3
(2a) Mutation in entry protein gp16 to abolish pause during genome entry	(2a) Mutation in gp16 to abolish pause
(2b) Mutation in early terminator (TE) to allow host RNAP to read through	(2b) TE mutation
(2c) New <i>E. coli</i> promoter activity upstream of ectopic RNAP gene	(3) 11 and 12 compensatory mutations in the two replicates, respectively, with no biological justification; both lines evolved the same A2 <i>E. coli</i> promoter change, different changes in the ectopic RNAP gene, and different changes in the tail fiber gene. No other changes were in common

Abbreviation: RNAP, RNA polymerase.

shorter stretch of nucleotide identity near the 3' end of the gene than the process that would have restored the RNAP gene to its original location without affecting other genes. Fitness showed a large jump coincident with the recombination but, at the end of the adaptation, remained somewhat shy of the maximum (evolved wild-type T7, Figure 2). Presumably, the fitness gain from the rare recombination event must have allowed the rearranged genome to dominate the population before the ectopic RNAP parent phage recombined to yield wild type. In addition, two compensatory mutations evolved before, and 15 (affecting 10 genes) and 1 promoter change evolved subsequent to the recombination, but there is no basis for understanding the relevance of any of these mutations to the reordered genome. Reflecting our lack of complete understanding of T7, the virtual model v2.5 (Endy *et al.*, 2000) predicted fitness loss due to the observed final gene order to be much smaller than actually observed.

In the two adaptations in which the RNAP gene remained at its ectopic location (replicates 2 and 3), fitness increased only modestly. Two anticipated changes did evolve in both lines (an early terminator mutation and a gene 16 mutation), but there was no change that could be interpreted as enhancing *E. coli* promoter activity upstream of the ectopic RNAP gene. Surprisingly, the terminator mutation extended the base pairing of the stem-loop hairpin, which naively might be predicted to increase terminator efficiency. Although the gene 16 mutation was known to allow genome entry at a faster rate than by host RNAP-catalyzed transcription (Struthers-Schlinke *et al.*, 2000), fitness assays

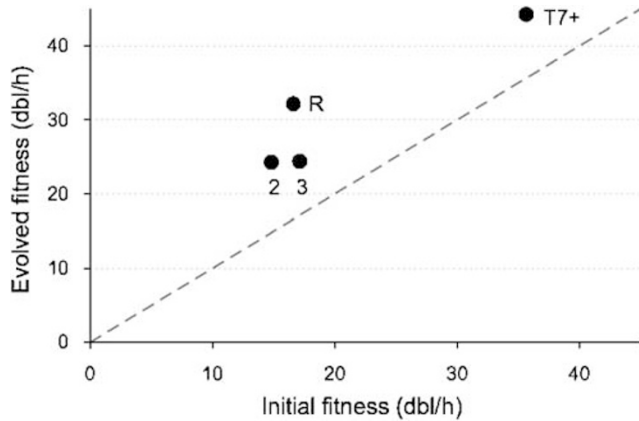


Figure 2 Fitness evolution of T7 wild type and phages with the RNA polymerase (RNAP) gene ectopically placed far from the end of the genome that enters the cell. Two different constructs were adapted, as indicated by the different initial fitnesses (see text). The replicate labeled R evolved a recombination that restored the RNAP gene to its wild-type location, but 13 other genes were misplaced by the recombination. Final fitness fell short of the maximum (evolved wild type, T7⁺). The RNAP gene remained at its ectopic location in replicates 2 and 3, and final fitnesses were similar in both, well below the maximum and also below that of phage R. Fitness is measured as doublings per hour (dbl/h) of the phage concentration.

revealed that nearly all the increase in fitness of the evolved line resulted from the TE change. Presumably, levels of transcription over the ectopic RNAP gene were more important than the rate of genome entry.

The results of this study challenge the conclusions from the study in which T7 was forced to grow on T3 RNAP. In particular, the failure of both lines that maintained the RNAP gene in its ectopic location to approach fitness levels of the evolved wild-type phage indicates that T7 is limited in its ability to evolve new regulation. One additional point to note is that a theoretically less-favored recombination pathway led to the dominance (over its parent) of a phage that had lower fitness than that achievable by a simple recombinatorial restoration of the ectopic RNAP to its original location. However, the fitness gain associated with the recombinant that was found precluded reversal of the process.

Loss of a DNA metabolism gene (Rokyta *et al.*, 2002)

Aside from RNAP, T7 encodes 6–7 genes that have been shown to be intimately associated with DNA metabolism: DNA polymerase, helicase/primase, endo- and exo-nuclease, ssDNA-binding protein and DNA ligase. A few other genes have a peripheral role in DNA metabolism. The replication of T7 is highly recombinogenic, and the resulting Holliday junctions are resolved by a combination of endonuclease and ligase activities. Ligase is not thought to have important physical interactions with any other phage proteins, and its presence in the phage genome is dispensable if the cellular ligase is active. However, if cellular ligase is rendered largely defective (it cannot be deleted, because some ligase activity is required for cell growth), deletion of the ligase gene from T7 is lethal to the phage.

T7 was deleted of its ligase gene and adapted for rapid growth on a ligase-defective host. Since the host ligase

Table 3 Compensatory evolution in response to deletion of DNA ligase gene

Expected changes	Observed changes
(1) Phage capture and evolution of <i>E. coli</i> DNA ligase gene	
OR	
(2) Evolution of remaining phage DNA metabolism genes	(2) Evolution of three phage DNA metabolism genes
	(3) A few changes with no known role in DNA metabolism

could not evolve, one anticipated outcome of this adaptation was recombination of the host ligase gene into T7, followed by compensatory evolution to improve it. Aside from a general requirement for DNA homology to allow recombination, this pathway is not trivial as T7 efficiently degrades all host DNA, obviating much opportunity for crossovers. An alternative pathway was suggested by an old study, which found that a conditionally lethal amber mutation in T7 ligase, fatal on a ligase-defective host, was rescued by an otherwise obligate lethal amber mutation in the phage endonuclease gene (Sadowski, 1974).

One deletion line was adapted for the long term, two briefly (Rokyta *et al.*, 2002). Starting from a negative fitness (a net loss of phage after an infection cycle), considerable improvement was observed in the long-term line. An apparent fitness plateau was reached well below that of wild type adapted to the same host: compensatory evolution resulted in recovery from a fitness low of –2 doublings/h to a maximum of 20 doublings/h, relative to the ligase⁺ phage of ~30 doublings/h. The genetic basis of compensatory evolution was found to lie largely in other DNA metabolism genes: endonuclease, DNA polymerase and helicase/primase. Compensatory changes in this pathway likely represent a type of regulatory evolution (Table 3). As originally suggested by Sadowski (1974), levels and activities of different proteins in a DNA metabolism network need to be balanced. For example, endonuclease cuts and ligase seals nicks in DNA. If the activity of one protein is drastically altered, altering the activity of the other should compensate to maintain the balance.

The endonuclease mutations observed by Rokyta *et al.* (2002) were clearly regulatory: one was an amber mutation in the endonuclease gene, another was a frameshift mutation in the gene immediately upstream that suppressed translation of the endonuclease gene. The mutations that arose in other DNA metabolism genes were missense changes whose effects on expression and activity were not assayed. The expectation, from extrapolation of the endonuclease results, is that those missense mutations would reduce activity of the protein, having a similar effect as if expression levels were reduced.

It remains to be seen if the results here generalize to other types of network disruptions. There is little genetic redundancy in the DNA metabolism network of T7, in contrast to the large gene networks reported for eukaryotes (Lee *et al.*, 2004). Redundancy will necessarily

reduce the impact of gene deletions and no doubt facilitate compensatory evolution. Notably, the effects of many deletions in eukaryotes are minor, reflecting overlapping or partially overlapping metabolic pathways (Thatcher *et al.*, 1998). Perhaps the most striking aspect of the Rokyta *et al.* (2002) results is that, since the ligase gene was removed, it may well have been predicted that the mutant phage could not evolve at all—no known phage gene whose function could compensate for the loss of ligase, and the host cell counterpart was destroyed every time T7 infected. This system might have been billed as ‘immune’ to viral evolution. Yet, the virus underwent considerable evolution by adjusting the levels and activities of other genes in the network of the lost gene. This result bodes poorly for attempting to experimentally limit the evolution of viral resistance in general.

Suicide plasmid (Djordjevic and Klaenhammer, 1997; Bull *et al.*, 2001)

T7 promoters have been placed on plasmids to express cloned genes, either upon infection by T7 or when placed in a host that expresses T7 RNAP. If the cloned gene downstream of a T7 promoter is toxic and kills the cell quickly enough, a T7 infection will be aborted before any progeny is formed, and the plasmid will serve as a T7-specific, anti-phage construct (altruistic, killing the cell but protecting other bacteria). Furthermore, since the phage cannot simultaneously evolve changes in its many promoters, such a system would seem to be aloof to evolution of resistance by the phage. This approach was used in two studies (Djordjevic and Klaenhammer, 1997; Bull *et al.*, 2001), the former using a *Lactococcus* phage that encoded a transcriptional activator rather than an RNAP. The anticipated resistance mechanism, at least when the T7 study was designed, was evolution by the phage to block the action of the toxic gene. In principle, therefore, when such resistance evolved, one could again block the phage by inserting a new toxic gene into the plasmid.

From the foregoing perspective, this viral challenge does not appear to address the evolution of regulation. Yet from another angle, the challenge is broadly similar to that of the first study (section RNAP exchange (Bull *et al.*, 2007)), in which the viral RNAP was changed. Here, the challenge was to evolve a regulatory system that remained functional within the phage genome but no longer expressed the plasmid system. Regulatory evolution of this nature seems difficult because the viral genome must evolve coordinated changes, so that the RNAP recognizes phage promoters but not the plasmid promoter. It thus seemed that the phage would need to simultaneously evolve changes in the RNAP gene and in several promoters, a difficult prospect.

Contrary to this wisdom, the phages in both studies evolved resistance by reducing expression from the plasmid but without evolving changes in their own promoters. Thus, the phages responded by specifically altering regulation of the plasmid-based promoter. Sequence changes in T7 associated with resistance affected the RNAP gene and either of two non-essential genes of the early region, but the possible function of the latter genes in the resistance mechanism is unknown.

The biochemical bases for this evolution were not explored at a mechanistic level in either study, but it was argued that the supercoiling of a plasmid-borne promoter differs from that in phage, thereby providing a basis by which the phage could evolve discrimination between otherwise identical promoter sequences. Alternatively, perhaps that the cellular locations of phage and plasmid promoters differ. Involvement of the T7 RNAP in resistance evolution is thus not surprising. In both studies, the magnitude of resistance was assayed only semiquantitatively, and it is not known how much of the maximum attainable resistance was realized. Nonetheless, the phage evolved a change to its gene regulatory system transparently, altering expression of extra-genomic elements without requiring changes to its own promoters.

Suicide plasmids provide one means of inhibiting phages through genetic modification of the host. The design of anti-phage vectors is greatly enhanced by an understanding of phage biology, and any resulting viral escape (whether evolutionary or just due to poor design) can give insight into phage genomics. Most research targeting the destruction and inhibition of phages comes from the dairy industry because phages pose major problems to the use of bacteria in milk processing. The milk cannot be sterilized, so phage cannot be excluded. Several types of phage suppression systems have been developed, and many seem robust to single-mutation escape by phages, but the long-term evolution of escape has rarely been studied (Sturino and Klaenhammer, 2006).

Overview: engineered genomes and the limits of adaptation

Much of evolutionary biology addresses fitness in a relative sense. Yet it is the *absolute* fitness or function of a microbe or molecule that is relevant to industry and medicine. Engineered genomes cannot be designed with such perfect knowledge that they are without faults, but it is commonly understood that subsequent adaptation of an engineered genome might be employed to correct any defects and restore fitness to its maximum.

The lessons from experimental evolution with T7 offer some guidance on the power of experimental evolution to correct design imperfections. Clearly, additional studies of this nature are warranted, but we can at least offer suggestions from what is so far evident.

- (1) Regulation is critical to fitness/function but is not always readily evolved. The suicide plasmid and RNAP exchange studies suggest that regulation can be easily evolved, whereas the genome rearrangement study suggests the opposite. It is sobering to consider that, even if one assembled all the correct genes and other elements, something as seemingly trivial as gene order can profoundly impact fitness in a way that is largely resistant to compensatory evolution.
- (2) Irreversible suppression of gene activity is partly correctable. Removal of a gene from a network can be compensated by changes in other components of the network, without restoring the suppressed function. This result points to the power of compensatory evolution to coordinate activities among components

of a network but that overall network function remains constrained by elements in the network that cannot undergo compensatory evolution.

Testing optimality models

Specific genome manipulations with T7 enable the use of known gene interactions in making rational predictions of compensatory evolution. They also provide insight into genome engineering. For the most part, however, these manipulations do not match natural evolution or natural challenges faced by a genome. An alternative approach to exploiting genotype–phenotype maps to understand evolution is to adapt organisms to particular environments for which there is a predictive basis of phenotype evolution, to see how the latter is affected by the underlying genetics. This alternative perspective is easily addressed using optimality models.

Optimality models in evolutionary biology attempt to predict stable end points of evolution: the phenotypes expected to evolve in the long term. Their focus is on natural selection as the mechanism driving evolutionary change, so optimality models are usually devoid of genetic details except for constraint functions in the form of trade-offs. Indeed, their intent is to reveal generalities that transcend genetic details and thus are broadly applicable. Yet this aloofness to details is considered their weakness and is the basis of controversy about their utility.

Until recently, the omission of genetic details from optimality models has been almost a necessity, due to the difficulty of uncovering the genetic bases of phenotypes. Now the field of genomics promises to change that, enabling a detailed understanding of the genetics of phenotypes or the mechanistic bases by which phenotypes change. With this new understanding, optimality models can be tested at both phenotypic and genetic levels to directly assess how genetic mechanisms impact evolution at the phenotypic level, and thus, when and how optimality fails because of genetics. Although phages are limited in their phenotypic diversity, there are some problems that can be addressed. T7 optimality has been studied from two perspectives.

Lysis time evolution (Heineman and Bull, 2007)

Lytic phages exhibit what is known in life history theory as semelparous reproduction—they reproduce once and then die. The time of lysis is the ‘age’ of reproduction for a phage. In an environment with unlimited hosts (which may be achieved experimentally by serial passage), early reproduction has the advantage of a short generation time. Yet in the four phages that have been studied, early reproduction has a downside—it reduces burst size (fecundity). Following ‘eclipse’ (the time from infection until the first mature progeny virion is assembled), the number of phage inside a cell increases linearly with time (Hausmann and Harle, 1971; Wang *et al.*, 1996; Wang, 2006). Late lysis therefore extends generation time but yields more offspring. This ‘trade-off’ between generation time and offspring number leads to a lysis time optimum that is suitable for experimental testing (Wang *et al.*, 1996; Bull, 2006), satisfying the first criterion for developing this problem as a model for studying the interplay between optimality and genetic details.

Lysis affords the additional advantage for studies of experimental evolution in having a well-studied genetic basis in many phages (Young, 1992, 2002; Wang *et al.*, 2000). Lysis of dsDNA phages is often controlled by a two-component system: the phage encodes an endolysin, capable of degrading the cell wall, and a holin, which is a membrane protein that creates pores in the inner membrane to allow access of the endolysin to the cell wall. When the latter is degraded, the internal osmotic pressure of the cell causes lysis. By itself, the endolysin cannot achieve lysis, because it cannot get through the inner membrane to access the cell wall. Holin thus acts as a timekeeper to control the release of endolysin through the inner membrane.

The holin in phage λ is known to be highly evolvable—many missense mutations allow function but alter lysis time (Grundling *et al.*, 2000). T7 encodes an endolysin and a holin, so by extension from other phages, it should be able to evolve changes in lysis time with a variety of point mutations. As demonstrated in several phages, including T3, progeny accumulates linearly following eclipse (Hausmann and Harle, 1971; Wang *et al.*, 1996; Bull, 2006; Wang, 2006).

Evolution of T7 lysis time was first studied with the compensatory evolution approach, using knockouts. It was suspected from prior work that the T7 holin is not critical for lysis (amber mutations in the gene are not lethal to phage growth), so a phage deleted of its endolysin gene was used. T7 endolysin also regulates the phage RNAP activity, so a second mutant was also adapted, one in which lysis activity was removed by a small deletion but the protein’s regulatory role remained intact. Although the initial phages were profoundly delayed in lysis time and exhibited great fitness losses, compensatory changes in a virion protein normally involved in genome entry into the cell almost fully restored lysis time and fitness, as compared to an adapted T7⁺. The virion protein contains a transglycosylase domain (the same activity as many phage endolysins) and the compensatory mutations of largest effect were found in this domain. In hindsight, the change made perfect sense, a protein normally involved in cell wall hydrolysis during genome entry at the initiation of infection gained the function of promoting lysis of the bacterium at the end of the infection. This study thus suggested that lysis time in T7 is capable of rapid and extreme evolution, setting the stage for the direct test of the optimality model.

The nature of the optimality test follows from the formula for the optimum. A simple model gives the optimum lysis time as fulfilling

$$\hat{L} = E + \frac{1}{\hat{r}}$$

where E is the eclipse time and \hat{r} is the intrinsic rate of increase of the phage population when lysis time is at its optimum ($\hat{}$ indicates the optimum value). The intrinsic rate of increase depends on many phage parameters, but it also depends strongly on host density—low host density ensures low \hat{r} , because phages spend most of their lives in the media waiting to encounter a host. Thus, this model predicts that high host density selects early lysis, low host density selects late lysis, an environmental property that is easily manipulated.

Adaptation of T7 wild type to a high density of hosts (10^8 cells ml⁻¹) resulted in several changes, notably in both eclipse and lysis time, and phenotypic evolution led to a closer match to predicted lysis time than at the start (Table 4). As the wild-type phage had not been knowingly adapted to the experimental conditions, the evolution of many phage phenotypes is expected. Once this phage was adapted to high density, however, the expected evolution in response to low density (10^6) was purely in lysis time, and presumably thus in the timekeeper (holin). Although one might expect cell physiology to differ between high and low density, the low-density environment in fact used a mix of permissive and non-permissive cells, achieving the same overall cell density as in the high-density adaptation. The low-density experiment thus used a low-density of permissive cells, not of total cells.

The phenotypic evolution at low density fell far short of the optimum, the phage failed to evolve much change in lysis time (Table 4). This was not due to a failure to evolve. Rather, seven substitutions arose in the low-density experiment, one with a presumed regulatory effect on the holin gene. Unexpectedly, eclipse time also evolved a slight increase. Most of the seven substitutions were shown to be specifically beneficial in the low-density environment, rather than being beneficial at high density as well. Thus, the failure to match the optimum at low-density adaptation was not due to a lack of beneficial mutations. If the T7 holin is as evolvable as the λ holin appears to be, adaptation of T7 to low-density growth should have at least approached the optimum, and evolution should have been confined to the holin gene. The main importance of this study is that it challenges the utility of simple optimality models, even for simple genomes.

Optimal foraging (Heineman *et al.*, 2008)

Animals make many choices in feeding, from avoiding or neglecting certain types of food items in favor of others, to how and when to move in search of new resources (MacArthur and Pianka, 1966; Pyke *et al.*, 1977). There is a 40-year history of modeling these decisions from the perspective of maximizing food/energy intake. The optimal foraging theory has virtually always been applied to plastic behaviors, in which the same individual can behave one way in some circumstances but differently in others. However, with slight modification, some of the theory can be applied to phages, where the 'decision' is genetically hard-wired. For example, the optimal lysis time considered above has been argued to be a form of optimal foraging in a 'patchy' environment, with each host being a 'patch' (Wang *et al.*, 1996).

In a more traditional extension of the animal-based theories, the host range of a phage is akin to choosing some food items and not others. The evolution of phage host range in fact obeys some of the most basic optimal foraging principles (Bull, 2006). While it might seem that a phage should never avoid a host from which it can produce offspring, this intuition fails when there is a high density of alternative hosts. Phage mutants that avoid a poor but suitable host can be favored because of the offsetting chance of encountering a better one.

This theory was tested with T7 (Heineman *et al.*, 2008). In contrast to the optimal lysis time study, the optimal

Table 4 Lysis time evolution

Host density	Eclipse time	Lysis time (obs)	Lysis time (exp)
<i>Initial (wild type)</i>			
10^8	9	13.3	11.4
<i>Evolved</i>			
10^8	7.1	10.4	9.1
10^6	7.8	12.4	18.6

Abbreviations: exp, expected from the optimality model; obs, observed.

Host density units are cells per milliliter; all time units are minutes.

foraging study was broken down into two steps that facilitated separating the genetic details from the nature of selection. The first step was to determine whether T7 could evolve to discriminate among hosts that it normally infects. If T7 could not pass this 'genetic details test' then it would not pass the second test, either. Stringent selection to avoid a host was achieved by a knockout of an essential (for T7) host protein (thioredoxin) such that the phage could infect but not produce progeny. Absence of thioredoxin is considered insurmountable, so evolution of resistance by T7 was not an issue.

Using this scheme, T7 was selected to (i) infect *E. coli* C and avoid *E. coli* K-12 and (ii) infect *E. coli* C and avoid *E. coli* B. Both selections were successful. It was expected that the genetic bases of these discriminations would affect the tail fiber (gene 17), and different amino-acid changes were indeed found in this gene from the two adaptations. The sufficiency of a tail fiber change was demonstrated by site-directed mutagenesis to create the parental phage with just the tail fiber change (this was attempted only for the adaptation to avoid *E. coli* K-12), and this phage was then used for the tests of optimal foraging.

Optimal foraging theory was tested by growing a mix of discriminating and non-discriminating T7 in a population of hosts and observing changes in the frequency of each type. This test was therefore evolutionary, expedited by starting with high levels of the two types of phages. The tests used two *E. coli* hosts (C and K-12), in which the quality of one host relative to the other was manipulated using the antibiotic tetracycline at sublethal levels. Host C carried resistance to tetracycline, so increasing concentrations of tetracycline rendered C a progressively better host than K-12. Three optimal foraging predictions were tested

- (i) Holding host densities constant, avoidance of the poor host will be favored when the difference in quality of the two hosts is high, but not when the difference is slight.
- (ii) Holding host quality constant, avoidance of the poor host will be favored at high densities of the good host but not at low densities
- (iii) For a density and quality of the good host at which avoidance of the poor host is favored, the evolution of discrimination will be insensitive to the density of the poor host.

Predictions (i) and (ii) were supported qualitatively, but (iii) was not. The failure of (iii) appeared to stem

from a cost to discrimination—the T7 that avoided K-12 may have had a slight penalty when growing on C compared to the non-discriminating T7. Thus, when the density of the poor host was low, there was not enough benefit in avoiding that host to overcome this disadvantage.

The optimal foraging model was thus largely successful with T7, and when it failed, the basis of its failure could be understood. The reason for this overall success may lie in the relatively simple genetic basis of discrimination. One gene is involved, and that gene appears to have a single role in the phage life history. Even the domain of the tail fiber protein involved in host recognition does not appear to have contacts with other phage proteins. Thus, the trait being selected is ‘isolated’ from the remainder of the phage genome network, and its evolution does not obviously involve effects on other traits. Under such conditions, phenotype evolution is straightforward.

Predicting evolution: the glass is half full

The motivation behind the studies reviewed here was to use extensive genetic knowledge of a system to predict evolution. The results have been mixed. No adaptations have been without unexplained beneficial mutations. In many studies, up to half the changes can be justified at the genetic level based on interactions known *a priori*, but we usually lack a basis for explaining the effect of the base or residue involved. On the one hand, this level of success is encouraging. Yet it is clear that substantial room for improvement remains. *A priori* predictions are often wrong at a fundamental level about the pathway of compensatory evolution. With the exception of the optimality studies, there is no formal basis for prediction in most of these adaptations, and in the lysis time evolution study, adaptation to low density was enigmatic at both the phenotypic and molecular levels. At a finer level, the changes in many genes involved in an adaptation cannot be rationalized. In some cases, those genes have known functions, but in many cases they do not. The explanation for many of these enigmas may be simply that gene networks do not operate in isolation of each other, so (for example) extreme changes in the DNA metabolism network impact phage assembly and other life cycle functions. We also usually lack an understanding of the fitness effects of many beneficial substitutions. However, we need only to explain substitutions with large effect, those of small effect are not worrisome.

Experimental evolution in free-living organisms

The T7 work and experimental adaptations of a similar nature with bacteria are laying a foundation for a new level of understanding evolution. Although our conjectures about evolution are often wrong, we can at least circumscribe the extent of our misunderstanding and begin to classify the types of errors and successes. Experimental evolution studies of bacteria now rival those of phages, and the understanding derived from both systems should enhance each other. Furthermore, with the realization that cancer is an evolutionary process of individual cells within a multi-celled body

(Frank, 2007), we may expect synergism between studies of adaptation in microbes and work on humans.

Bacteria offer the advantage that a great deal is known about metabolism, at the genetic and biochemical levels, as well as at a system-wide level through metabolic flux (Dykhuizen *et al.*, 1987); consequently, the metabolic bases of evolution have been a focus of study (Maharjan and Ferenci, 2003; Hua *et al.*, 2006, 2007; Maharjan *et al.*, 2006). This multi-level understanding facilitates the interpretation of evolution. It is also now possible to obtain genome-wide sequences of bacteria to identify evolutionary changes (Honisch *et al.*, 2004; Herring *et al.*, 2006; Velicer *et al.*, 2006) and to obtain genome-wide expression levels (Fong *et al.*, 2005b). At the same time, explaining the evolutionary changes can be problematic. For example, adaptations can involve changes in expression levels of thousands of genes (Fong *et al.*, 2005b), and even single-base changes with profound phenotypic effects can be inexplicable in terms of known biology (Fiegna *et al.*, 2006). Furthermore, the course of evolution in bacteria lacking mechanisms for the exchange of DNA is subject to extensive clonal interference, whereby most beneficial changes are lost in competition with the few genomes that happened to get the best mutation (Cooper, 2007; Perfeito *et al.*, 2007). Interpreting the molecular bases of experimental adaptations done in the absence of recombination thus faces the problem that mutations may have evolved and been lost, not just because of their intrinsic effects, but because of other mutations in the same genome. (An interesting irony regarding clonal interference is that, despite the apparent evolutionary advantage of recombination in avoiding clonal interference (Cooper, 2007), the benefit of bacterial transformation—a major mechanism for bacterial recombination—remains elusive in experimental settings (Mongold, 1992; Bacher *et al.*, 2006).) Thus, the ability to predict most of the evolution in an adaptation still eludes us, but the start is encouraging, and this step is a necessary one toward a better theory.

Acknowledgements

The T7 work reviewed here has benefited from consistent funding from National Institutes of Health (GM 57756 to JJB and GM32095 to IJM) and from the technical help of Marty Badgett, Rachael Springman and Priscilla Kemp. JJB received support as the Miescher Regents Professor at the University of Texas.

References

- Bacher JM, Metzgar D, de Crecy-Lagard V (2006). Rapid evolution of diminished transformability in *Acinetobacter baylyi*. *J Bacteriol* **188**: 8534–8542.
- Bull JJ (2006). Optimality models of phage life history and parallels in disease evolution. *J Theor Biol* **241**: 928–938.
- Bull JJ, Badgett MR, Molineux IJ (2001). A general mechanism for viral resistance to suicide gene expression. *J Mol Evol* **53**: 47–54.
- Bull JJ, Springman R, Molineux IJ (2007). Compensatory evolution in response to a novel RNA polymerase: orthologous replacement of a central network gene. *Mol Biol Evol* **24**: 900–908.
- Burch CL, Chao L (2000). Evolvability of an RNA virus is determined by its mutational neighbourhood. *Nature* **406**: 625–628.

- Chao L, Tran TT (1997). The advantage of sex in the RNA virus phi6. *Genetics* **147**: 953–959.
- Cooper TF (2007). Recombination speeds adaptation by reducing competition between beneficial mutations in populations of *Escherichia coli*. *PLoS Biol* **5**: e225.
- Dean AM, Dykhuizen DE, Hartl DL (1986). Fitness as a function of beta-galactosidase activity in *Escherichia coli*. *Genet Res* **48**: 1–8.
- Dekel E, Alon U (2005). Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**: 588–592.
- Djordjevic GM, Klaenhammer TR (1997). Bacteriophage-triggered defense systems: phage adaptation and design improvements. *Appl Environ Microbiol* **63**: 4370–4376.
- Dunn JJ, Studier FW (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol* **166**: 477–535.
- Dykhuizen DE, Dean AM, Hartl DL (1987). Metabolic flux and fitness. *Genetics* **115**: 25–31.
- Endy D, You L, Yin J, Molineux IJ (2000). Computation, prediction, and experimental tests of fitness for bacteriophage T7 mutants with permuted genomes. *Proc Natl Acad Sci USA* **97**: 5375–5380.
- Fiegna F, Yu YT, Kadam SV, Velicer GJ (2006). Evolution of an obligate social cheater to a superior cooperator. *Nature* **441**: 310–314.
- Fong SS, Joyce AR, Palsson BO (2005a). The econometrics of evolution. *Nat Chem Biol* **1**: 191–192.
- Fong SS, Joyce AR, Palsson BO (2005b). Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Res* **15**: 1365–1372.
- Fong SS, Marciniak JY, Palsson BO (2003). Description and interpretation of adaptive evolution of *Escherichia coli* K-12 MG1655 by using a genome-scale *in silico* metabolic model. *J Bacteriol* **185**: 6400–6408.
- Frank SA (2007). *Dynamics of Cancer: Incidence, Inheritance, and Evolution*. Princeton University Press: Princeton, NJ.
- Garcia LR, Molineux IJ (1996). Transcription-independent DNA translocation of bacteriophage T7 DNA into *Escherichia coli*. *J Bacteriol* **178**: 6921–6929.
- Grundling A, Blasi U, Young RY (2000). Genetic and biochemical analysis of dimer and oligomer interactions of the lambda S holin. *J Bacteriol* **182**: 6082–6090.
- Hausmann R, Harle E (1971). Expression of the genomes of the bacteriophages T3 and T7. In: Broda E, Locker A, Sringer-Ledere H (eds). *Proceedings of the First European Biophysics Congress*. Weiner Medical Akademie: Vienna. pp 467–498.
- Heineman RH, Bull JJ (2007). Testing optimality with experimental evolution: lysis time in a bacteriophage. *Evolution* **61**: 1695–1709.
- Heineman RH, Springman R, Bull JJ (2008). Optimal foraging by phages through host avoidance. *Am Nat* (in press).
- Herring CD, Raghunathan A, Honisch C, Patel T, Applebee MK, Joyce AR *et al.* (2006). Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* **38**: 1406–1412.
- Honisch C, Raghunathan A, Cantor CR, Palsson BO, van den Boom D (2004). High-throughput mutation detection underlying adaptive evolution of *Escherichia coli*-K12. *Genome Res* **14**: 2495–2502.
- Hua Q, Joyce AR, Fong SS, Palsson BO (2006). Metabolic analysis of adaptive evolution for *in silico*-designed lactate-producing strains. *Biotechnol Bioeng* **95**: 992–1002.
- Hua Q, Joyce AR, Palsson BO, Fong SS (2007). Metabolic characterization of *Escherichia coli* strains adapted to growth on lactate. *Appl Environ Microbiol* **73**: 4639–4647.
- Ibarra RU, Edwards JS, Palsson BO (2002). *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth. *Nature* **420**: 186–189.
- Kemp P, Garcia LR, Molineux IJ (2005). Changes in bacteriophage T7 virion structure at the initiation of infection. *Virology* **340**: 307–317.
- Klement JF, Moorefield MB, Jorgensen E, Brown JE, Risman S, McAllister WT (1990). Discrimination between bacteriophage T3 and T7 promoters by the T3 and T7 RNA polymerases depends primarily upon a three base-pair region located 10 to 12 base-pairs upstream from the start site. *J Mol Biol* **215**: 21–29.
- Lee I, Date SV, Adai AT, Marcotte EM (2004). A probabilistic functional network of yeast genes. *Science* **306**: 1555–1558.
- Lenski RE, Mongold JA, Sniegowski PD, Travisano M, Vasi F, Gerrish PJ *et al.* (1998). Evolution of competitive fitness in experimental populations of *E. coli*: what makes one genotype a better competitor than another? *Antonie Van Leeuwenhoek* **73**: 35–47.
- Lunzer M, Natarajan A, Dykhuizen DE, Dean AM (2002). Enzyme kinetics, substitutable resources and competition: from biochemistry to frequency-dependent selection in lac. *Genetics* **162**: 485–499.
- MacArthur RH, Pianka ER (1966). On optimal use of a patchy environment. *Am Nat* **100**: 603–609.
- Maharjan R, Seeto S, Notley-McRobb L, Ferenci T (2006). Clonal adaptive radiation in a constant environment. *Science* **313**: 514–517.
- Maharjan RP, Ferenci T (2003). Global metabolite analysis: the influence of extraction methodology on metabolome profiles of *Escherichia coli*. *Anal Biochem* **313**: 145–154.
- Molineux I (2006). The T7 group. In: Calendar R (ed). *The Bacteriophages*. Oxford University Press: New York, NY, USA, pp 277–301.
- Mongold JA (1992). DNA repair and the evolution of transformation in *Haemophilus influenzae*. *Genetics* **132**: 893–898.
- Papadopoulos D, Schneider D, Meier-Eiss J, Arber W, Lenski RE, Blot M (1999). Genomic evolution during a 10,000-generation experiment with bacteria. *Proc Natl Acad Sci USA* **96**: 3807–3812.
- Perfeito L, Fernandes L, Mota C, Gordo I (2007). Adaptive mutations in bacteria: high rate and small effects. *Science* **317**: 813–815.
- Pyke GH, Pulliam HR, Charnov EL (1977). Optimal foraging: a selective review of theory and tests. *Q Rev Biol* **52**: 137–154.
- Riley MS, Cooper VS, Lenski RE, Forney LJ, Marsh TL (2001). Rapid phenotypic change and diversification of a soil bacterium during 1000 generations of experimental evolution. *Microbiology* **147**: 995–1006.
- Rokyta D, Badgett MR, Molineux IJ, Bull JJ (2002). Experimental genomic evolution: extensive compensation for loss of DNA ligase activity in a virus. *Mol Biol Evol* **19**: 230–238.
- Sadowski PD (1974). Suppression of a mutation in gene 3 of bacteriophage T7 (T7 endonuclease I) by mutations in phage and host polynucleotide ligase. *J Virol* **13**: 226–229.
- Springman R, Badgett MR, Molineux IJ, Bull JJ (2005). Gene order constrains adaptation in bacteriophage T7. *Virology* **341**: 141–152.
- Struthers-Schlinke JS, Robins WP, Kemp P, Molineux IJ (2000). The internal head protein Gp16 controls DNA ejection from the bacteriophage T7 virion. *J Mol Biol* **301**: 35–45.
- Sturino JM, Klaenhammer TR (2006). Engineered bacteriophage-defence systems in bioprocessing. *Nat Rev Microbiol* **4**: 395–404.
- Thatcher JW, Shaw JM, Dickinson WJ (1998). Marginal fitness contributions of nonessential genes in yeast. *Proc Natl Acad Sci USA* **95**: 253–257.
- Turner PE, Chao L (1999). Prisoner's dilemma in an RNA virus. *Nature* **398**: 441–443.
- Velicer GJ, Raddatz G, Keller H, Deiss S, Lanz C, Dinkelacker I *et al.* (2006). Comprehensive mutation identification in an

- evolved bacterial cooperator and its cheating ancestor. *Proc Natl Acad Sci USA* **103**: 8107–8112.
- Wang IN (2006). Lysis timing and bacteriophage fitness. *Genetics* **172**: 17–26.
- Wang IN, Dykhuizen DE, Slobodkin LB (1996). The evolution of phage lysis timing. *Evol Ecol* **10**: 545–558.
- Wang IN, Smith DL, Young R (2000). Holins: the protein clocks of bacteriophage infections. *Annu Rev Microbiol* **54**: 799–825.
- Young R (1992). Bacteriophage lysis: mechanism and regulation. *Microbiol Rev* **56**: 430–481.
- Young R (2002). Bacteriophage holins: deadly diversity. *J Mol Microbiol Biotechnol* **4**: 21–36.
- Zhang X, Studier FW (1997). Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme. *J Mol Biol* **269**: 10–27.
- Zhang X, Studier FW (2004). Multiple roles of T7 RNA polymerase and T7 lysozyme during bacteriophage T7 infection. *J Mol Biol* **340**: 707–730.
- Zhong S, Khodursky A, Dykhuizen DE, Dean AM (2004). Evolutionary genomics of ecological specialization. *Proc Natl Acad Sci USA* **101**: 11719–11724.