Viral attenuation by engineered protein fragmentation

Daniel J. Garry,[1],[†] Andrew D. Ellington,[1] Ian J. Molineux,[2] and James J. Bull[3],*

[1]Department of Molecular Biosciences, Center for Systems and Synthetic Biology, [2]Department of Molecular Biosciences and and [3]Department of Integrative Biology, Institute for Cellular and Molecular Biology, University of Texas, Austin, TX 78712, USA

*Corresponding author: E-mail: bull@utexas.edu

[†]Present address: Eligo Bioscience S.A., Pépinière Paris Santé Cochin, 29 rue du Faubourg Saint-Jacques, 75014 Paris, France.

Abstract

A possible but untested method of viral attenuation is protein fragmentation, engineering wild-type proteins as two or more peptides that self-assemble after translation. Here, the bacteriophage T7 was engineered to encode its essential RNA polymerase as two peptides. Initial fitness was profoundly suppressed. Subjecting the engineered virus to over 100 generations of adaptation by serial transfer resulted in a large fitness increase, still remaining below that of evolved wild-type. The fitness increase was accompanied by three substitutions in the fragmented peptides as well as six mutations in other parts of the genome, but the fragmentation was retained. This study thereby demonstrates the feasibility of using gene fragmentation as a possibly permanent method of attenuation, but the initial fitness of the engineered genome may be a poor measure of its fitness on extended adaptation.

Key words: attenuation; genome engineering; bacteriophage; evolution; adaptation

1. Introduction

Genetic engineering now enables the design of arbitrary genomes almost at will. In the realm of viral vaccine development, this technology means that the bottleneck in generating new vaccines lies chiefly in our imagination and ability to invent functional designs. For live, attenuated vaccines, designs must be compatible with the essential viral functions needed for infection and propagation, but there are otherwise almost limitless possibilities.

The benefit of engineering a live vaccine over old methods is considerable. The old method, one of merely adapting a virus to novel conditions and hoping for reduced growth in the original host, could not only fail to attenuate, but successful attenuations could be fragile, due to few mutations and thus easily evolve back to high virulence during growth in a patient (Fenner and Cairns 1959; Hanley 2011). Engineering allows the possibility of quantitative control of attenuation plus assurance that any reversion is difficult.

A few methods of engineering viral attenuation have been tested, at least *in vitro*: deletion of non-essential elements, wholesale silent codon modifications (which may operate by changing specific dinucleotide frequencies), and genome rearrangement (reviewed in Lauring, Jones, and Andino 2010; Bull 2015; see Kunec and Osterrieder 2016 for a discussion of dinucleotide ratios). All operate by reducing viral growth rate, thus reducing viral load in the host. (Some methods in development alter tissue tropism without overtly suppressing growth rate in other tissues.) Where implemented, these various engineering methods are all successful at suppressing viral growth rate, but they differ in their applicability to different viruses and in their susceptibility to evolutionary reversion (Bull 2015; Bull, Smithson, and Nuismer 2018).

Here we explore a previously untested method of genome modification for viral attenuation: protein 'fragmentation' by encoding formerly single polypeptides as multiple fragments/peptides that assemble inside the infected cell. The expected effect is to alter the stoichiometry of proteins needed for viral
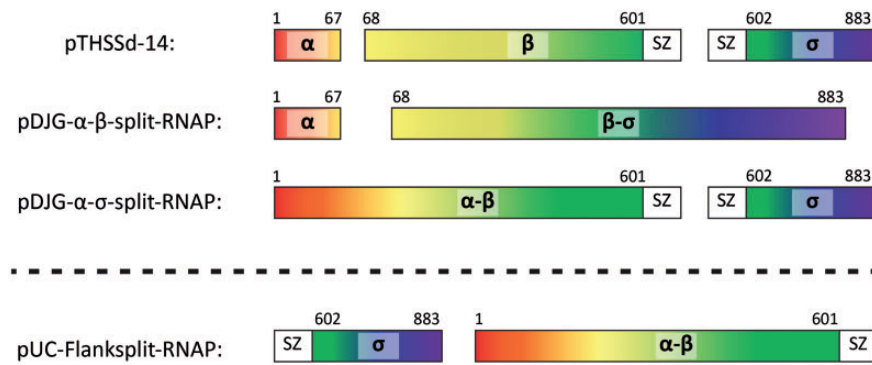
Figure 1. Design of plasmids encoding T7 RNAP fragments, color-coded for ease of identification; plasmid names are given on the left. (Top) A three-fragment RNAP constructed by Segall-Shapiro et al. (2014), showing the breakpoints, boundaries of the α, β, and σ fragments, and positioning of leucine zippers (SZ). (Second and third down) The two possible two-fragment RNAP constructs using the breakpoints from Segall-Shapiro et al. (2014). Both were constructed here, but only the third down complemented a T7 phage lacking its own RNAP. (Bottom) The complementing two-fragment RNAP that was used for recombination into T7Δ1; the σ fragment is 5' of the α-β fragment, to block recombination restoring a single reading frame of the correct orientation in the phage.

function and thus to slow viral replication and assembly, if not also to reduce progeny production. We test the method by encoding T7 RNA polymerase (RNAP) as two peptides expressed from the viral genome and measure viral growth rate under defined conditions. To evaluate the evolutionary stability of the design, the modified genome is selected for higher fitness. Although the genome engineering is stable, the fitness is not: the initial fitness defect is greatly ameliorated on subsequent adaptation.

2. Genome design and implementation

In the wild-type genome, T7 RNAP is encoded as a single 883 amino acid protein (Dunn and Studier 1983). The ability of T7 RNAP to function when 'fragmented,' i.e. encoded as multiple, simultaneously expressed peptides, was demonstrated by Segall-Shapiro et al. (2014). Using a clever transposon-based gene disruption method, they identified breakpoints in the protein that preserved function when the different segments were translated in the same cell. Functional assays of the fragmented RNAPs were limited to transcription of a fluorescent protein expressed from a T7 promoter on a separate plasmid. Thus, the effect on these fragmented RNAPs on the intracellular life cycle of bacteriophage T7 was unknown.

Our goal was to introduce any of those constructs into the genome of bacteriophage T7 such that T7 growth was supported by the fragmented RNAP. Neither their 3-fragment nor their 4-fragment RNAP constructs complemented a T7 genome lacking its own RNAP gene (Fig. 1 shows the 3-fragment RNAP tested), even though the growth of such deletion phages is easily supported by complementing plasmids that encode the non-fragmented, wild-type RNAP (Bull, Springman, and Molineux 2007). This complementation failure suggests either that some essential function of the RNAP other than transcription (Zhang and Studier 1995, 2004) was abolished by the fragmentation, or that the level of transcription by the fragmented RNAP was too low to support detectable phage growth.

A two-fragment T7 RNAP, using the β-σ breakpoint identified by Segall-Shapiro et al. (2014) was engineered into a plasmid behind a P_{tac} promoter (plasmid pDJG-α-σ-split-RNAP in Fig. 1). The short σ fragment contains the specificity loop, and both fragments were shown to contain elements essential to transcription (Segall-Shapiro et al. 2014). As per that previous study, complementary binding domains from a leucine zipper were

encoded on each fragment to enhance the association of both peptides. A T7 phage deleted of its RNAP gene (T7Δ1) formed plaques on hosts expressing this plasmid, indicating that all essential functions of RNAP were maintained. This fragmentation was then encoded in reverse order (the σ fragment 5' of the α-β fragment) to ensure that no simple deletion could restore the two RNAP fragments into a single reading frame with the wild-type orientation. To enable recombination into T7Δ1, the reverse-order, fragmented RNAP gene was encoded on a plasmid flanked with DNA sequences matching those on each side of the RNAP deletion in T7Δ1 (plasmid pUC-Flanksplit-RNAP in Fig. 1, the flanking sequences omitted from the illustration). Plating of this phage on hosts carrying the plasmid yielded several recombinants that grew without complementation. A single recombinant (JB 13-78) was used for all subsequent studies.

3. Results

A viral isolate encoding the rearranged RNAP gene was created. This isolate was then used to found a population that was subjected to serial transfer adaptation which favored increased growth rate. Fitness assays and whole-genome sequencing were conducted of the initial isolate, of an intermediate population at 26 h of adaptation and of the endpoint population at 38 h of adaptation.

3.1 Fitness

Initial fitness of the recombinant phage was low, approximately 15 doublings/hr (Fig. 2). For comparison, wild-type T7 fitness is near 35 (Heineman, Molineux, and Bull 2005), and adaptation to the constant host and culture conditions used here results in a wild-type fitness in the low-mid 40s (Bull, Heineman, and Wilke 2011). Since these measures use a log scale, the fitness difference of almost 20 between non-adapted wild-type T7 and the engineered genome is large, approximately a million-fold descendants per hour. Thus, a single phage with a fitness of 15 produces nearly 3.3×10^4 descendants in an hour, whereas with a fitness of 35, the number of descendants is just over 3.4×10^{10}. The fragmentation of this one protein thus had an initially profound effect on fitness.

Attenuated viruses are live (infectious) and thus can potentially evolve when growing within a patient. Rapid reversal of attenuation is known for the oral polio vaccine (Burns et al. 2014) and is a drawback of that vaccine—for one of the
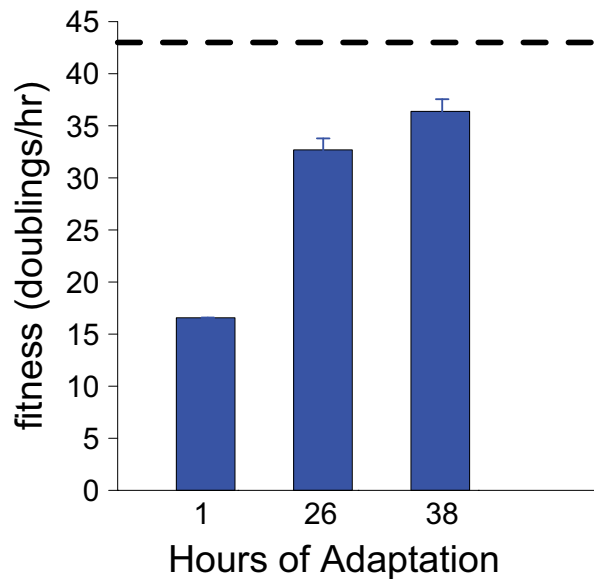
Figure 2. Fitness evolution of T7 carrying the fragmented RNAP. The initial fitness was taken at the end of ~1 h of growth, due to the fact that the phage is first identified from a plaque (which grows for at least an hour). For phages of very low fitness, an unknown amount of adaptation may occur during formation of the initial plaque, so the true initial fitness is possibly less than was measured. Fitness was also measured at 26 and 38 h of transfer. The improvement between beginning and end is over 20 doublings/hour. Fitness is measured as doublings of the phage population per hour when grown under defined conditions with hosts in excess. The dashed line represents fitness of wild-type T7 adapted to these growth conditions (Bull, Heineman, and Wilke 2011).

serotypes, the attenuation is due to just two mutations, so rapid reversal is an easily understood outcome (Nathanson and Kew 2010). Our design was intended to block evolution that could restore the wild-type RNAP peptide, but even with the engineered defect maintained, there is the possibility of compensatory evolution to mitigate the effect of the engineering (Berkhout et al. 1999; Rokyta et al. 2002; Springman et al. 2005; Harcombe, Springman, and Bull 2009). To test stability of attenuation, the engineered phage was subjected to serial transfer adaptation for a combined duration of ~38 h, or just over 100 generations.

Fitness increased markedly, reaching the mid-high 30s (Fig. 2). The final fitness of ~37 is still well below that attained with adaptations of wild-type (Bull, Heineman, and Wilke 2011), but the magnitude of fitness recovery was approximately three fourths of the difference between the initial fitness and that of evolved wild-type (on a log scale). The recovery of >20-fold doublings per hour translates into an increase of more than a million-fold descendants per hour. Further adaptation might have recovered even higher fitness, but much of our interest lies in the large magnitude of recovery observed.

3.2 Molecular evolution

Sequencing of the engineered phage population at the start, at 26 h and at the 38 h endpoint revealed the changes contributing to these fitness gains. The consensus changes observed are given in Table 1 (the triplet insert was only at low frequency at 26 h but is listed at this time point to indicate that its ascent began before 26 h). All mutations but two were single base changes, the exceptions being a triplet TAC (tyrosine) insert in gene 0.3 (anti-type I restriction defense) and, only at 26 h, a deletion that fused most of 0.3 with the leucine zipper. (The base changes in this table use the GenBank reference for JB 13-78

(MG833025), although the protein changes are the same as for T7 wild-type, except for gene 1.)

Three substitutions occurred in the RNAP subunits: V133A (V685 in the full-length, wild-type T7 RNAP), N145S (N697), and T530I (T596). The first two are in the σ fragment (yellow ribbon in Fig. 3), the third of these is very near the carboxy end of the α-β fragment (green ribbon). The V133A (V685A) mutation has been reported in a T7 RNAP variant with increased ability to incorporate 2'-O-methyl-modified NTPs (Meyer et al. 2015). The T530I (T596) mutation is in a flexible loop for which half of the loop is from one sub-unit and the other half is from another, strongly suggesting that it is a stabilizing mutation.

The 10B mutation has been seen in several adaptations of the phage (e.g. Bull, Springman, and Molineux 2007; Bull, Molineux, and Wilke 2012) and thus likely has no role specific to the fragmentation. In contrast, there is genetic evidence of an interaction between gp19 and T7 RNAP in the roles of maturation and DNA packaging (Zhang and Studier 2004), so the change in 19 may well be in response to the fragmentation. No change was found in lysozyme (gene 3.5), whose product is also known to associate with RNAP.

An interesting deletion mutant was observed at the 26-h analysis point. This deletion removed the 3' 99 bases of gene 0.3, all of 0.4 and through the first 11 codons of the leucine zipper (genes 0.5–0.7 are absent in the ancestral T7Δ1). The high fitness of this mutant is evident from its mid-range frequency in the phage population, but it was no longer detectable by 38 h (it's absence at the final time point was indicated by breseq analysis and also by a direct search of the fastq read file for the sequence of the deletion junction). Deletions in the early region are commonly observed in T7 grown in rich media (Studier 1973; Cunningham et al. 1997); what is interesting in this case is its eventual loss. However, the deletion may have risen to high frequency before the ascent of other beneficial mutations, those other mutations benefitting from elements absent from the deletion.

In the final population, therefore, the profound fitness recovery can be attributed to no more than seven-point mutations (the triplet insert is too rare at 26 h to explain the high fitness, and the 10B mutation likely having at best a small effect independent of the RNAP fragmentation).

4. Discussion

In phage T7, fragmentation of the essential RNAP gene to encode the protein as two peptides led to a viable phage but a profound fitness reduction. The choice of the fragmentation site was based on an earlier study that identified three fragmentation sites that were compatible with transcription (Segall-Shapiro et al. 2014), but here we found that only one of those sites was compatible with phage viability. This mechanism of attenuation should be feasible with other proteins and other viruses.

A priori, the expected basis of fitness reduction was an altered stoichiometry of proteins essential for phage growth. Yet high fitness was recovered on subsequent adaptation without restoring the wild-type stoichiometry, so some other mechanism must have operated for most of the initial fitness loss. It is likely that the fragmentation disrupted but did not irreversibly abolish some function of RNAP, and adaptation merely improved the ability of the RNAP complex to perform that function. Several mutations in T7 RNAP are known that render the phage inviable while maintaining transcription (Zhang and Studier 1995, 2004), so this explanation has precedent.

Table 1. Changes evolved in the phage genome with the fragmented RNAP (MG833025).

Base	DNA Change	Gene	Protein change[a]	When observed (hr)
1,074	TAC insertion	0.3	Y ins after AA 47	26[b], 38
1180-1534	deleted	partial 0.3, all of 0.4	fuse 0.3 with leucine zipper	26[c]
1,486	T→G	intergenic		26, 38
1,893	T→C	1	V133A	26, 38
1,929	A→G	1	N145S	26, 38
4,297	C→T	1	T530I	26, 38
17,498	C→A	6.7	A18A	26, 38
22,666	G→A	10B	E375K	26, 38
33,991	G→A	17	D263N	26, 38
36,465	C→T	19	A172V	26, 38

[a]a protein residue is one less than the codon number when the initiating methionine is absent from the mature protein (Dunn and Studier 1983).
[b]In far <50% of the population at 26 h.
[c]In ~50% of the population at 26 h.

Figure 3. Protein structure ribbon diagram of intact T7 RNAP (PDB: 1QLN). Yellow indicates the portion of the molecule corresponding to the σ fragment, cyan the α-β fragment. Bound DNA is indicated by orange. Red marks the three substitutions evolved during adaptation of the recombinant phage genome. One substitution is at the bottom of the figure in the thin strand, barely evident.

Comprehensive understanding of the initial defect may prove challenging because it could involve any of several different functions of the RNAP (terminator read-through, interaction with lysozyme, interaction with terminase, primer synthesis for the DNA polymerase, or rate of abortive vs productive transcription initiation). Furthermore, because phage fitness is strongly affected by phage generation time, part of the initial fitness loss may be due merely to the increased time required for the fragments to assemble into a functional RNAP complex rather than defects with the assembled complex.

Adaptation of the attenuated virus led to a vast fitness improvement despite the fragmentation being maintained. Most of the improvement can be attributed to no more than seven-point mutations, three of which were in the RNAP fragments themselves. Many methods for attenuation by engineering are susceptible to partial evolutionary recovery in the short term despite retaining the engineered changes, and in this respect, attenuation by fragmentation fits the pattern. For example, in T7 deleted of important but otherwise non-essential genes, the fitness impact was reduced through compensatory evolution in other parts of the genome without restoring the deletion (Rokyta et al. 2002; Heineman et al. 2005; Harcombe, Springman,

and Bull 2009). Adaptation of T7 constructs with genome rearrangements has also exhibited partial fitness recovery while retaining the new genome order (Cecchini et al. 2013). An exception to this pattern has been observed in some RNA viruses attenuated by deletions, where fitness recovery evolved via reconstitution of the deleted elements (Whatmore et al. 1995; Olsthoorn and van Duin 1996; Berkhout et al. 1999). Despite partial fitness recovery with most engineered attenuations, some methods have maintained substantial fitness suppression over the long term (Bull 2015).

Although it is premature to generalize from a single fragmentation and a single adaptation, the fact that we could *a priori* design a fragmentation that did not revert and that fitness was not fully recovered may mean that fragmentation offers a means of enforcing a 'permanent' fitness loss to a virus, albeit that the magnitude of permanent loss will not be evident until after a period of adaptation. The possible evolutionary permanence of fitness suppression by fragmentation stands in contrast to the recoveries noted above for deletions in some RNA viruses and for engineered codon deoptimizations (Bull 2015). An attenuated vaccine designed by protein fragmentation should maintain wild-type antigenicity unless the fragmentation points lie within peptide epitopes displayed by the immune system. These designs should also be free of novelties arising by recombination with wild-type: when designed in reverse order, as here, the only viable recombinants would be those that completely replaced the fragmented proteins, thus merely reconstituting the wild-type genome that was the partner in recombination.

If the long-term fitness effect of a single fragmentation is too slight to produce the necessary level of attenuation for vaccine purposes, a more profound long-term suppression may be achieved by introducing multiple fragmentations within a genome. The main challenge with introducing multiple fragmentations into a single genome simultaneously would be that the initial fitness might be too low for detectable growth; this difficulty can be overcome by stepwise introductions of fragmentations, each followed by evolving the virus (*in vitro*) before introducing the next fragmentation (a similar argument was proposed for attenuation by multiple deletions, Berkhout et al. 1999), with the additional benefit that fitness would be relatively stable when the engineering was complete. Alternatively, engineering the fragments without a leucine zipper might limit long-term fitness adequately for single fragmentations.

Separating the reading frames of different fragments into different parts of the genome may also help limit fitness recovery. In the present system, the σ and α-β fragments were encoded adjacently, on the same transcript, and so would have been translated as adjacent proteins on the ribosome. This adjacency may thus have disposed their proteins to rapid association. Separating them to different parts of the genome may reduce their rate of association. Indeed, a mere reordering of genes in the T7 genome has lasting and large fitness effects (Cecchini et al. 2013), so combining protein fragmentation with reordering the peptides may impose even stronger, irreversible fitness costs. When fragmenting T7 RNAP, a basic constraint is that all parts must be located in the genome early region (transcribed by *Escherichia coli* RNAP), or the complete T7 RNAP will not be made. This constraint does not apply to other T7 genes.

The study here fragmented the RNAP, an essential protein but one required in low levels. Some proteins are required in high amounts, such as the major capsid protein and the scaffold protein. Fragmentation of high-copy proteins might lead to more substantial fitness suppression in the long term, but there would still be the potential for adaptation to greatly improve fitness beyond the initial level. With slight modifications, the transposon method of (Segall-Shapiro et al. 2014) applied to entire genomes provides a way to identify potentially many viable protein breakpoints, ultimately allowing the test of other fragmentations.

There are conceptual parallels with the approach used here and viral polyproteins (polyproteins are reviewed in Yost and Marcotrigiano 2013; Crépin et al. 2015). In many viruses, polyproteins are generated as single peptides that are then cleaved into fragments that perform different functions. Our approach also involved a form of 'splitting' single peptides, but the fragments were encoded separately. Those fragments did not knowingly perform separate functions and were instead required to assemble to perform a common essential role in promoting phage growth. In applying our design for viral attenuation, the hope was that the assembly requirement would delay the viral life cycle if not also impair function. Another difference between engineered fragmentation and polyproteins is that polyproteins force the ratios of their constituent proteins to obey specific rules, whereas encoding single proteins as multiple fragments in principle allows arbitrary relative abundances of the different fragments.

The directed evolution approach used here—of selecting a viral genome in response to an imperfect genome modification—can be used to improve upon engineering for biotechnical ends. In the present system, evolved changes to the engineered RNAP gene itself (of which there were three) may improve the function of the engineered protein and could be incorporated in any application using the RNAP outside the phage genome. Likewise, a phage genome encoding only one of the two RNAP fragments could be adapted to accommodate any of a range of partner fragments encoded in *trans* by the host. By placing selection on the engineered phage genome, even slight improvements will be easily amplified to high frequency over extended transfers, outpacing almost any screening method that might be used to identify beneficial mutants. The main possible drawback of using a whole-genome approach to improve the performance of individual genes/proteins is that fitness improvements may stem from mutations outside of the engineered portion. If beneficial mutations that improve the engineered protein indeed exist, independent replicate adaptations should at least sometimes recover beneficial mutations within the engineered regions.

5. Methods

5.1 Media

LB broth (10 g Bacto tryptone, 10 g NaCl, 5 g Bacto yeast extract per liter) was used for throughout. Media for plates consisted of LB with 15 g/l of Bacto agar; media for overlays used 7 g/l agar.

5.2 Strains

IJ1133 (an *E. coli* K-12 strain) was the host used for all phage growth (Bull, Heineman, and Wilke 2011). T7Δ1 is a derivative of wild-type T7 with a deletion that removes genes *0.5, 0.6A, 0.6B, 0.7, 1* and extends into part of the T7 promoter ø1.1 A (the deletion lies in the 'early' region). The genome of the T7 carrying the fragmented RNAP gene, JB 13-78, is provided in GenBank file MG833025.

5.3 Plasmids

Plasmids pTHSSd-14 (three-fragment T7 RNAP) and pTHSSd-18 (four-fragment T7 RNAP) were created by Segall-Shapiro et al. (2014) and obtained from Addgene. Plasmid pDJG-beta-sigma-split RNAP, used for complementation, consisted of the pTHSSd backbone with a P_{Tac} promoter and the α-β and σ fragments of RNAP in that order. Plasmid pUC19-FlanksplitRNAP, used for recombination over T7Δ1, consisted of a pUC backbone with the σ and α-β fragments of RNAP in that order. The two fragments were separated by 19 bases that included an extra stop codon at the end of the σ fragment, thus ensuring that the two open reading frames were not encoded as a single peptide.

5.4 Sequences

Sequences were generated on an Illumina MiSeq platform by the IBEST Genomics Resources Core at the University of Idaho. Sequences reads were 300 bases long, generated as paired ends, but because of the small genome, mapping was limited to a single fastq file and did not use the pairings (more information about sequence methodology is provided with the online sequences—see below). The fastq files were mapped onto the template of JB 13-78 (Genbank MG833025) by breseq (Deatherage and Barrick 2014; Barrick et al. 2014).The T7 RNAP crystal structure (PDB: 1QLN) was visualized using the PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.

5.5 Fitness and adaptation

Adaptation of JB 13-78 was carried out as in many previous studies (e.g. Bull, Heineman, and Wilke 2011). Briefly, IJ1133 cells were grown for an hour at 37°C in 10 ml LB to $\sim10^8$/ml; phage were added at low multiplicity and grown for a fixed time, then transferred to another culture of cells. The volume and timing of transfers were adjusted to sometimes allow lysis of the culture (ensuring high multiplicity and recombination), and sometimes not (to maintain selection on phage growth rate). Fitness assays were conducted under the same conditions as adaptation, except that high densities of phage were avoided. Fitness was measured as the $\log_2$ net expansion of phage density across an hour of growth, accounting for dilutions at transfer.

Data availability

Fastq sequence files are deposited in the NCBI Sequence Read Archive (BioProject PRJNA436551).

References

Barrick, J. E., Colburn, G., Deatherage, D. E., Traverse, C. C., Strand, M. D., Borges, J. J., Knoester, D. B., Reba, A., and Meyer, A. G. (2014) 'Identifying Structural Variation in Haploid Microbial Genomes from Short-Read Resequencing Data Using Breseq', *BMC Genomics*, 15: 1039.

Berkhout, B., Verhoef, K., van Wamel, J. L., et al. (1999) 'Genetic Instability of Live, Attenuated Human Immunodeficiency Virus Type 1 Vaccine Strains', *Journal of Virology*, 73: 1138–45.

Bull, J. J. (2015) 'Evolutionary Reversion of Live Viral Vaccines: Can Genetic Engineering Subdue It?', *Virus Evolution*, 1: vev005.

——, Heineman, R. H., and Wilke, C. O. (2011) 'The Phenotype-Fitness Map in Experimental Evolution of Phages', *PloS One*, 6: e27796. 6.

——, Molineux, I. J., and —— (2012) 'Slow Fitness Recovery in a Codon-Modified Viral Genome', *Molecular Biology and Evolution*, 29: 2997–3004.

——, Smithson, M. W., and Nuismer, S. L. (2018) 'Transmissible Viral Vaccines', *Trends in Microbiology*, 26: 6–15.

——, Springman, R., and Molineux, I. J. (2007) 'Compensatory Evolution in Response to a Novel RNA Polymerase: Orthologous Replacement of a Central Network Gene', *Molecular Biology and Evolution*, 24: 900–8.

Burns, C. C., Diop O. M. et al. (2014) 'Vaccine-Derived Polioviruses", *The Journal of Infectious Diseases*, 210: S283–93.

Cecchini, N., Schmerer, M., Molineux, I. J., Springman, R., et al. (2013) 'Evolutionarily Stable Attenuation by Genome Rearrangement in a Virus', *G3 (Bethesda, Md.)*, 3: 1389–97.

Crépin, T., Swale, C., Monod, A., Garzoni, F., Chaillet, M., and Berger, I. (2015) 'Polyproteins in Structural Biology', *Current Opinion in Structural Biology*, 32, 139–46.

Cunningham, C. W., Jeng, K., Husti, J., Badgett, M., Molineux, I. J., Hillis, D. M., and Bull, J. J. (1997) 'Parallel Molecular Evolution of Deletions and Nonsense Mutations in Bacteriophage T7', *Molecular Biology and Evolution*, 1: 113–6.

Deatherage, D. E., and Barrick, J. E. (2014) 'Identification of Mutations in Laboratory-Evolved Microbes from Next-Generation Sequencing Data Using Breseq', *Methods in Molecular Biology (Clifton, N.J.)*, 1151: 165–88.

Dunn, J. J., and Studier, F. W. (1983) 'Complete Nucleotide Sequence of Bacteriophage T7 DNA and the Locations of T7 Genetic Elements', *Journal of Molecular Biology*, 166: 477–535.

Fenner, F. and Cairns J., 1959. "Variation in Virulence in Relation to Adaptaton to New Hosts", in F. M., Burnet and W. M., Stanley (eds.) *The Viruses, Vol. 3. Animal Viruses*, Vol. 3, pp. 225–49. New York, NY: Academic Press.

Hanley, K. A. (2011) 'The Double-Edged Sword: How Evolution Can Make or Break a Live-Attenuated Virus Vaccine', *Evolution: Education and Outreach*, 4: 635–43.

Harcombe, W. R., Springman, R., and Bull, J. J. (2009) 'Compensatory Evolution for a Gene Deletion Is Not Limited to Its Immediate Functional Network', *BMC Evolutionary Biology*, 9: 106.

Heineman, R. H., Molineux, I. J., and —— (2005) 'Evolutionary Robustness of an Optimal Phenotype: Re-Evolution of Lysis in a Bacteriophage Deleted for Its Lysin Gene', *Journal of Molecular Evolution*, 61: 181–91.

Kunec, D., and Osterrieder, N. (2016) 'Codon Pair Bias Is a Direct Consequence of Dinucleotide Bias', *Cell Reports*, 14: 55–67.

Lauring, A. S., Jones, J. O., and Andino, R. (2010) 'Rationalizing the Development of Live Attenuated Virus Vaccines', *Nature Biotechnology*, 28: 573–9.

Meyer, A. J., Garry, D. J., Hall, B., Byrom, M. M., McDonald, H. G., Yang, X., Yin, Y. W., and Ellington, A. D. (2015) 'Transcription Yield of Fully 2'-Modified RNA Can Be Increased by the Addition of Thermostabilizing Mutations to T7 RNA Polymerase Mutants', *Nucleic Acids Research*, 43: 7480–8.

Nathanson, N., and Kew, O. M. (2010) 'From Emergence to Eradication: The Epidemiology of Poliomyelitis Deconstructed', *American Journal of Epidemiology*, 172: 1213–29.

Olsthoorn, R. C., and van Duin, J. (1996) 'Evolutionary Reconstruction of a Hairpin Deleted from the Genome of an RNA Virus', *Proceedings of the National Academy of Sciences of the United States of America*, 93: 12256–61.

Rokyta, D., Badgett, M. R., Molineux, I. J., et al. (2002) 'Experimental Genomic Evolution: Extensive Compensation for Loss of DNA Ligase Activity in a Virus', *Molecular Biology and Evolution*, 19: 230–8.

Segall-Shapiro, T. H., Meyer, A. J., Ellington, A. D., Sontag, E. D., and Voigt, C. A. (2014) 'A 'Resource Allocator' for Transcription Based on a Highly Fragmented T7 RNA Polymerase', *Molecular Systems Biology*, 10: 742.

Springman, R., Badgett, M. R., Molineux, I. J., and Bull, J. J. (2005) 'Gene Order Constrains Adaptation in Bacteriophage T7', *Virology*, 341: 141–52.

Studier, F. W. (1973) 'Genetic Analysis of Non-Essential Bacteriophage T7 Genes', *Journal of Molecular Biology*, 79: 227–36.

Whatmore, A. M., Cook, N., Hall, G. A., Sharpe, S., Rud, E. W., and Cranage, M. P. (1995) 'Repair and Evolution of Nef in Vivo Modulates Simian Immunodeficiency Virus Virulence', *Journal of Virology*, 69: 5117–23.

Yost, S. A., and Marcotrigiano, J. (2013) 'Viral Precursor Polyproteins: Keys of Regulation from Replication to Maturation', *Current Opinion in Virology*, 3: 137–42.

Zhang, X., and Studier, F. W. (1995) 'Isolation of Transcriptionally Active Mutants of T7 RNA Polymerase That Do Not Support Phage Growth', *Journal of Molecular Biology*, 250: 156–68.

——, and —— (2004) 'Multiple Roles of T7 RNA Polymerase and T7 Lysozyme during Bacteriophage T7 Infection', *Journal of Molecular Biology*, 340: 707–30.