Refactoring Bacteriophage T7

Leon Chan[1*], Sriram Kosuri[2*], Drew Endy[2♣]

[1]Department of Biology
[2]Division of Biological Engineering
Massachusetts Institute of Technology, 68-580
77 Massachusetts Avenue, Cambridge, MA 02139
*L.C. and S.K. contributed equally to this work
♣Contact via email [endy@mit.edu]

Summary

Natural biological systems are selected by evolution to continue to exist. Evolution might give rise to complicated systems that are difficult to discover, measure, model, and direct. Here, we redesign the genome of a natural biological system, bacteriophage T7, in order to specify an engineered alternative that is easier to study, understand, and extend. We replaced the left 11,515 base pairs of the wild-type genome with 12,179 base pairs of redesigned DNA. The resulting chimeric genome encodes a viable bacteriophage that maintains key features of the original while being simpler to model and easier to manipulate.

Introduction

Bacteriophage T7 (T7) is an obligate lytic phage that infects *Escherichia coli* (*1,2*). Scientists isolated T7 and other phage to begin to study the physical characteristics and components of biological systems (*3*). Genetics, and then biochemistry, enabled the discovery and characterization of some of the individual elements that participate in T7 development. Sequencing of the T7 genome revealed additional elements, not all of which have obvious functions (supporting online text).

A subsequent synthesis of knowledge of individual parts and mechanisms produced descriptive, system-level models for T7 development, from genome entry to phage particle formation (*4,5*). Two features specific to T7 biology made the construction of system-level models easier. First, compared to other phage, T7 is *relatively* independent of complex host physiology. Second, RNA polymerase pulls most of the T7 genome into the newly infected cell (*6,7*). Polymerase mediated genome entry is a relatively slow process that results in the direct physical coupling of gene expression dynamics to gene position.

More recently, others and we have used computational models of T7 infection to begin to explore questions related to the organization of genetic elements on the T7 genome and the timing and control of gene expression across uncertain physical environments (*8-11*). In using these computational models, some predictions did not agree with experiments (*12*). For example, a mutant phage expected to grow faster than the wild type grew slower (*9*).

Upon inspection, disagreements between model-based prediction and experiment could have arisen for at least three reasons. First, the models could not meaningfully include unknown functions. For example, a non-essential gene with then unknown function, *1.7*, appears to impact phage DNA replication when disrupted (*9*). While differences between expectation and observation can suggest follow-on science, the lack of component-level understanding impaired our system-level analyses. Second, the boundaries of genetic elements on the T7 genome are more complex than our models of the genome. For example, genes *2.8* and *3* are most easily modeled as separable genetic elements even though the actual genes *2.8* and *3* overlap (Figure 1A; Figure S2). Third, a model built with separable parts that encode independent functions can be over-manipulated relative to the physical system. For example, while we could simulate the expected behavior of large sets of permuted genomes, we could not easily move a single open reading frame to another arbitrary position on the actual T7 genome.

While wild-type T7 is a superb organism for discovering the components of a biological system (*13*), is the original T7 isolate also best suited for understanding how the parts of T7 are organized to encode a functioning biological system? Here, we decided to refactor (*14*) the genome of wild-type T7 in order to attempt to specify a new system that might be better suited for our purposes.

Results

We began design of a new T7 genome, which we designated T7.1, by re-annotating the genome of wild-type T7. The wild-type T7 genome is a 39,937 base pair linear double-stranded DNA molecule (*4*). We annotated the genome by specifying the boundaries of the following functional genetic elements: 57 open reading frames with 57 putative ribosome binding sites encoding 60 proteins, and 51 regulatory elements controlling phage gene expression, DNA replication, and genome packaging (supporting online text).

To specify the architecture of T7.1 we organized the functional genetic elements into 73 'parts.' Each part contains one or more elements. While the DNA sequence of elements within parts may overlap, there is no overlap across part boundaries. Next, we organized contiguous parts into 'sections' with section boundaries defined by restriction endonuclease sites found only once in the wild-type sequence. Six sections, *alpha* through *zêta*, make up the T7.1 genome (Figure 2A; Figure S1). Sections were used to compartmentalize changes across the genome and facilitate construction, manipulation, and testing.

To specify the DNA sequence of T7.1 we eliminated sequence overlap across part boundaries. Overlaps were eliminated by exact duplication of the wild-type DNA sequence; subsequent sequence editing produced a single instance of any duplicated element (Figure 1B; Figure S2). All sequence edits within parts were limited to open reading frames, edits within open reading frames maintained the wild-type tRNA specification or, when necessary, specified higher abundance tRNA (*15*). We also added bracketing restriction endonuclease sites to insulate and enable the independent manipulation of each part (Figure 2C, E; Figure S1). Bracketing sites are not used

elsewhere in the sequence of any one section but are reused across sections. The DNA sequence of T7.1 changes or adds 1,424 base pairs to the wild-type genome (supporting online text).

The sections that comprise the T7.1 genome can be built and tested independently. We constructed the first two sections, *alpha* and *beta* (supporting online text). *Alpha* and *beta* contain the first 32 of 73 parts of the T7.1 genome, replacing the left 11,515 base pairs of the wild-type genome with 12,179 base pairs of redesigned DNA, and encoding the entire T7 early region, the primary origins of DNA replication, most of the T7 middle genes, and the control architecture that regulates T7 gene expression. *Alpha* and *beta* also contain the highest density of elements across the genome. We combined *alpha* and *beta* with the remainder of the wild-type (WT) genome to produce three chimeric phage: *alpha*-WT, WT-*beta*-WT, and *alpha*-*beta*-WT.

We first tested and recovered viable chimeric phage by transfection and plating (24). All three chimeric phage are viable. We isolated DNA and performed restriction digests across *alpha* and *beta* to confirm that individual parts could be independently manipulated (24). 30 of 32 parts in sections *alpha* and *beta* can be cut out as designed (Figure 3). We also sequenced *alpha* and *beta*. Sequencing revealed differences between the design of T7.1 and the actual "as-built" sections (24). Relevant sequence differences in section *alpha* include a single base deletion in gene *0.4* and in the *E. coli* terminator TE. Differences in section *beta* include a single amino acid substitution in both genes *1.8* and *2*, a single base deletion in gene *2.5*, and an 82-base truncation in gene *2.8*. All differences were due to errors or limitations in construction (supporting online text).

We characterized some growth properties of the chimeric phage by liquid culture lysis and plating (24). Phage-induced lysis of log-phase 30°C liquid cultures indicated a -20, 1.4, and -22% difference from wild type in the half-lysis times (*16*) of the *alpha*, *beta*, and *alpha-beta* chimeras, respectively (Figure 4A). Plaques were indistinguishable early during plaque growth and at 30°C, but at 37°C, the chimeric phage plaques appeared to stop growing as the bacterial lawns developed (not shown). After 24h at 37°C, plaque sizes relative to the wild type were smaller for each of the chimeric phage, with the *alpha-beta* chimera being smallest (Figure 4B).

Discussion

A system that is partially understood can continue to be studied in hope of exact characterization (*17*). Or, if enough is known about the system, a surrogate can be specified to study, replace, or extend the original. Here, we decided to redesign the genome of a natural biological system, bacteriophage T7, in order to specify an engineered biological system that is easier to study and manipulate. The new genome, T7.1, is based on our incomplete understanding of the information encoded in the wild-type genome and our desire to insulate and independently manipulate known primary genetic elements. We constructed the first two sections of T7.1 and observed that the resulting chimeric phage are viable.

Phage viability demonstrates the following for sections *alpha* and *beta*. First, our parts as chosen can be separated by exogenous DNA sequence. Second, any functions encoded by part overlap are non-essential. Third, our current understanding of T7 is *not insufficient* to specify a viable bacteriophage. Viability does not demonstrate sufficiency because (i) if the chimeric phage had not been viable then our current understanding would have been demonstrably *insufficient*, and (ii) while T7.1 is based on our current understanding, we do *not* have an exact understanding of all functions encoded in the T7.1 genome (e.g., genes of unknown function). Finally, viability, combined with the observed similarities in lysis times, suggest that T7.1 preserves polymerase-mediated genome entry and remains *relatively* independent of host cell physiology.

We constructed sections *alpha* and *beta* manually. Concurrent advances in de novo DNA synthesis technology have recently enabled the rapid automatic synthesis of DNA fragments the size of the T7.1 genome sections (*18*). As genome synthesis (*19*) and engineering (*20-22*) technologies continue to improve, the use of bracketing restriction sites for manipulating each part should become less important, but the physical, and perhaps functional, decoupling of parts by the elimination of sequence overlap will remain useful.

The viability of T7.1 presents a choice. T7.1 is a 'physical model' that can be used to continue to study the wild-type phage. Observed differences in the behavior of T7.1 relative to wild-type highlight relevant gaps in our understanding of the natural system. Or, as T7.1 is simpler and easier to manipulate and appears to retain aspects of wild-type T7 that make it an attractive model system, we can define T7.1 as our new model system; successor phage based on T7.1 can be constructed in order to answer questions of genome organization, regulation, and evolution. More generally, other natural biological systems could be redesigned and built anew in support of scientific discovery and human intention.

References and Notes

1. I. J. Molineux in *Encyclopedia of Molecular Biology*, Ed. T. C. Creighton (Wiley, New York, 1999), pp. 2495-2507.
2. I. J. Molineux in *The Bacteriophages*, R. L. Calendar Ed. (Oxford Univ. Press, New York, *in press*), (available at http://www.thebacteriophages.org/frames_0200.htm)
3. M. Demerec, U. Fano, *Genetics* **30**, 119 (1945).
4. J. J. Dunn, F. W. Studier, *J. Mol. Biol*. **166**, 477 (1983).
5. F. W. Studier, J. J. Dunn, *Cold Spring Symp. Quant. Biol*. **47**, 499 (Pt. 2, 1983).
6. S. K. Zavriev, M. F. Shemyakin, *Nucleic Acids Res*. **10**, 1635 (1982).
7. L. R. Garcia, I. J. Molineux, *J Bacteriol*. **177**, 4066 (1995).
8. D. Endy, D. Kong, J. Yin, *Biotech. Bioeng*. **55**, 375 (1997).
9. D. Endy, L. You, J. Yin, I. J. Molineux, *Proc. Natl. Acad. Sci., U.S.A.* **97**, 5375 (2000).
10. L. You, J. Yin, *J. Bacteriol*. **184**, 1888 (2002).
11. L. You, J. Yin, *Genetics* **160**, 1273 (2002).

12. D. Endy, R. Brent, *Nature* **409**, 391 (2001).
13. F. W. Studier, *Science* **176**, 367 (1972).
14. The goal of refactoring is to improve the internal structure of a system for future use while maintaining external system function (23).
15. T. Ikemura, *J. Mol. Biol.* **151**, 389 (1981).
16. Half lysis times are calculated by first taking the average between the baseline absorbance readings at the start of infection and end of lysis and finding the time the particular culture took to reach that absorbance.
17. B. L. Davis, *Los Anales de Buenos Aires*, año 1, no. 3 (1946) (Written by J. L. Borges and A. B. Casares under pseudonym; http://www.kyb.tuebingen.mpg.de/bu/people/bs/borges.html)
18. H. O. Smith, C. A. Hutchison, C. Pfannkoch, J. C. Venter, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15540 (2003).
19. R. Carlson, *Biosecur. Bioterror* **1**, 203 (2003).
20. D. L. Court, J. A. Sawitzke, L. C. Thomason *Annu. Rev. Genet.* **26**, 261 (2002).
21. D. L. Court *et al.*, *Gene* **315**, 63 (2003).
22. Y. Kang *et al.*, *J. Bacteriol* **186**, 4921 (2004).
23. M. Fowler, K. Beck, J. Brant, W. Opdyke, D. Roberts *Refactoring: Improving the Design of Existing Code* (Addison-Wesley Professional, Boston, 1999).
24. Materials and methods are available as supporting material on Science Online.
25. We thank Ian Molineux, Priscilla Kemp, and Heather Keller for discussions and advice throughout the work. We thank John Dunn and Barbara Lade for the pSCANS-5 vector. We thank Roger Brent, Eric Eisenstadt, Tom Knight and members of the Endy group for additional discussions and sustained encouragement. We thank Austin Che, Heather Keller, Alex Mallet, Kathleen McGinness, Samantha Sutton, and Ty Thomson for comments on the manuscript. We thank Felice Frankel for plaque photography and encouragement. The work was funded by grants to D.E. from the US Office of Naval Research, DARPA, and NIH. S.K. was supported by a NIH MIT BPEC training fellowship. Additional support was provided by MIT.

Supporting Online Material
www.sciencemag.org
Supporting Online Text
Materials and Methods
Figs. S1, S2
Tables S1, S2

A. Wild-type T7 *2.8-3* elements

```
----------------2.8------------------>
acgcaaagggaggcgacatggcaggttacggcgctaaaggaatccgaaa
     <--3-RBS---><----------------3--------------
```

B. T7.1 parts 28 and 29

```
acgcaaGgggagAcgacaCggcaggttacggcgctaaggatccggccgcaaagggaggcgacatggcaggttacggcgctaaa
----------------2.8-----------------><D28R|D29L><--3RBS------><--------------3----
```

Figure 1. Element decompression and part design. (A) The coding
regions of genes *2.8* and *3* overlap in the wild-type T7 genome.
The ribosome binding site of gene *3* (underlined) is encoded
within gene *2.8*. (B) Distinct genetic parts make up the T7.1
genome. The natural ribosome binding site and start codon
(green) for gene *3* are disrupted by point mutations (capitals);
mutations do not change the amino acid sequence of the 2.8
protein. Parts 28 and 29 are separated by bracketing
restriction sites, BamHI (blue) and EagI (orange). Figure S2
lists all changes from the wild-type T7 genome.
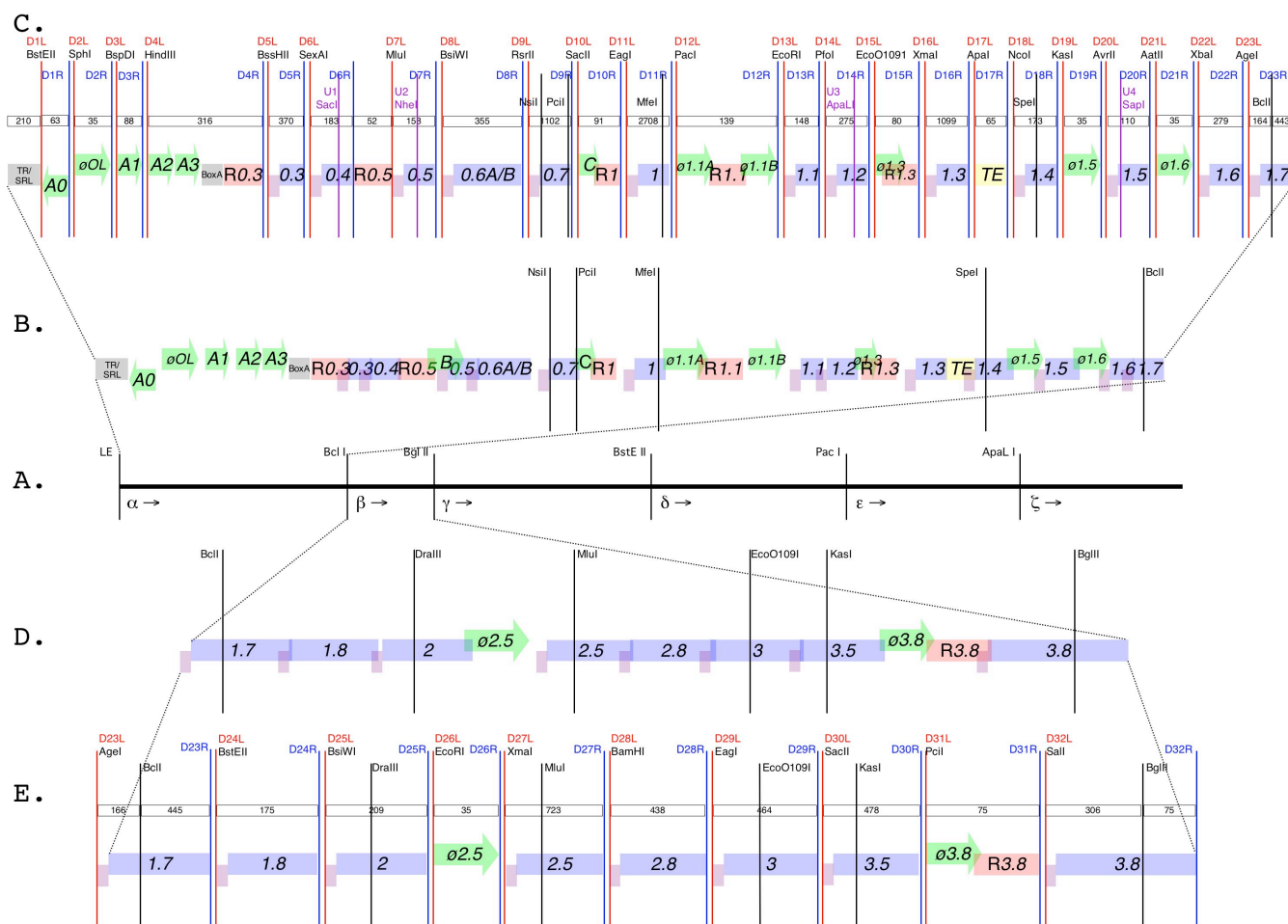
Figure 2. Genome design. (A) We split the wild-type T7 genome into six sections, *alpha* through *zêta*, using five restriction sites unique across the natural sequence. (B) Wild-type section *alpha* genetic elements: protein coding regions (blue), ribosome binding sites (purple), promoters (green), RNase III recognition sites (pink), a transcription terminator (yellow), and others (gray). Elements are labeled by convention (4). Images are not to scale, but overlapping boundaries indicate elements with shared sequence. The five useful natural restriction sites across section *alpha* are shown (black lines). (C) T7.1 section *alpha* parts. Parts are given integer numbers, 1 through 73, starting at the left end of the genome. Unique restriction site pairs bracket each part (red/blue lines, labeled D[part #]L/R). Added unique restriction sites (purple lines, U[part #]) and part length (# base pairs, open boxes) are shown. (D) Wild-type section *beta* genetic elements. (E) T7.1 section *beta* parts.
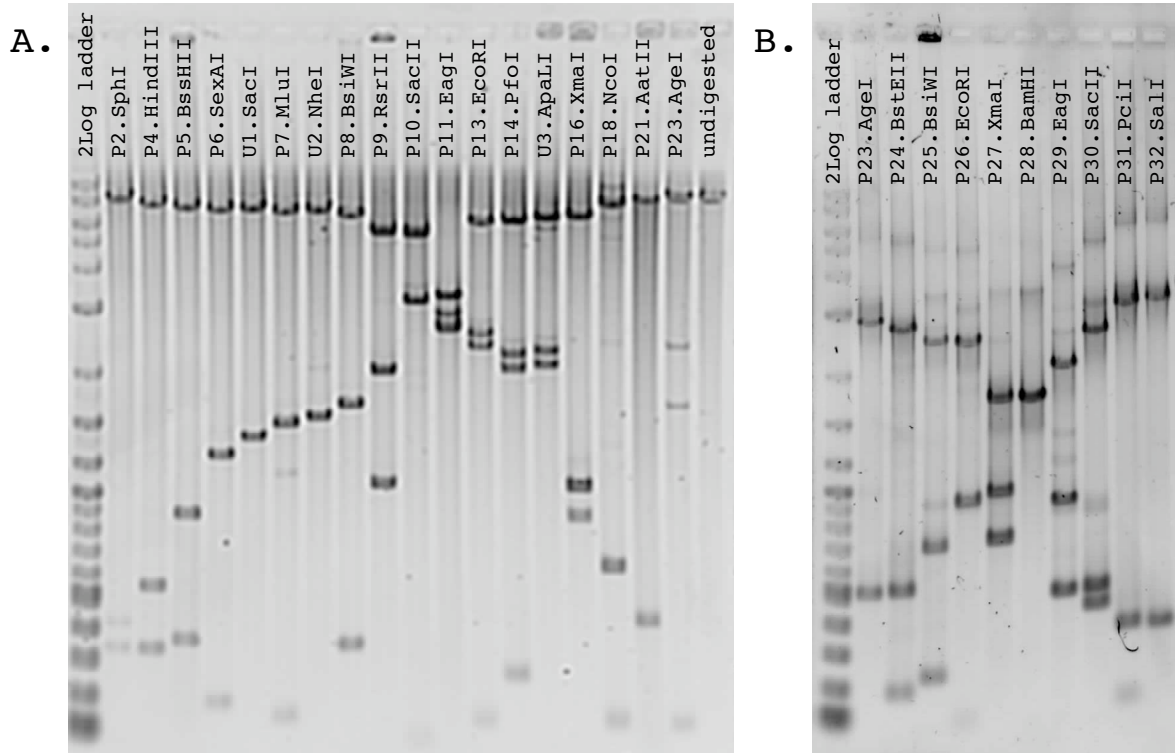
Figure 3. Cutting parts from T7.1 (A) Restriction enzymes specific to the sites that bracket parts (P#.Enzyme) and added unique restriction sites (U#.Enzyme) were used to cut section *alpha* (*24*). A subset of the digests is shown. As built, part 1 cannot be removed. (B) Restriction digests cutting out all parts in section *beta*. As built, part 28 cannot be removed.

A.

B.
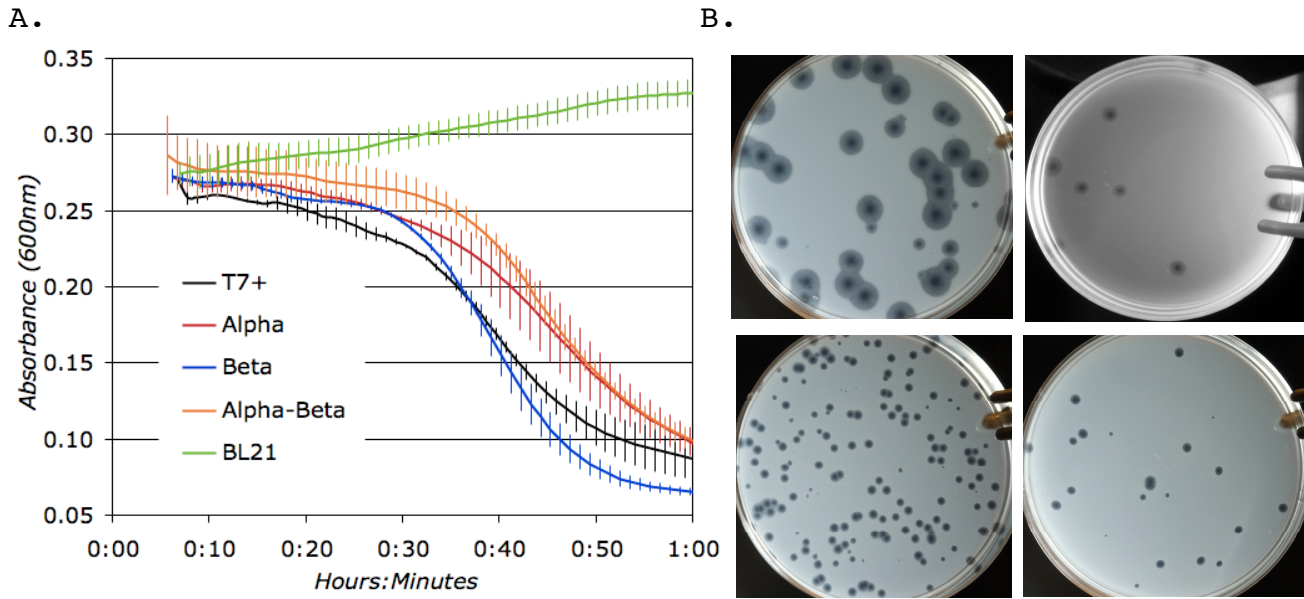
Figure 4. Characterization of T7.1 (A) Lysis of log-phase liquid cultures of *E. coli* BL21 (30˚C) by wild-type T7 (black), *alpha*-WT chimera (red), WT-*beta*-WT chimera (blue), *alpha-beta*-WT chimera (orange); absorbance of 0.275 is ~2E8 cells/ml. Vertical bars show standard deviation at each time point (based on four replicates) (*24*). (B) T7 plaques on *E. coli* BL21 (24 hrs, 37˚C, 10cm Petri dish). Clockwise from top left: wild-type (WT) T7, *alpha*-WT chimera, WT-*beta*-WT chimera, *alpha-beta*-WT chimera (*24*).

Supporting Online Material

Supporting Online Text
Materials and Methods
Figs. S1, S2
Tables S1, S2

Supporting Text

A. Re-Annotation of the Wild-Type T7 Genome
A.1. Summary

We use past experiments and observations to define specific boundaries of functional genetic elements on the bacteriophage T7 genome. We follow the standard naming conventions developed by Studier and Dunn (*1,2*).

A.2. Strain History

Bacteriophage T7 was twice isolated from Ward MacNeal's standard anti-coli-phage mixture (*3,4*). MacNeal's 'mixture' may have been cultured in series – T7 was the only identifiable isolate (*5*). One of the two original T7 isolates was reportedly chosen for future use (*5*) and master cultures of 'wild-type' T7 have been maintained since (*6*).

A.3. Discovery & Mapping of Protein Coding Domains

The T7 protein coding domains were first characterized by the isolation and analysis of randomly generated amber mutants. Nineteen genes were identified by mapping mutants that disrupt T7 DNA synthesis, particle maturation, and culture lysis (*7-9*). Two additional genes, T7 DNA ligase and protein kinase, were isolated via loss of function

and deletion, respectively (*10,11*). The genetic analysis of ligase and kinase mutants was carried out using mutant host strains that do not support the growth of ligase or kinase defective phage (12). Up to 30 T7 proteins have been observed by pulsing phage-infected cells with radioactive amino acids (*13,14*). Further evidence, such as electrophoretic mobility shifts of amber mutants (*15*) provided evidence for 38 proteins. Sequencing of the genome confirmed the previously constructed genetic maps (*1*). Analysis of the complete genome sequence revealed that the protein coding domains revealed by mutagenesis, screening, and mapping were not exhaustive and that additional unidentified open reading frames occupied the remainder of the genome. Additional open reading frames were labeled as protein coding domains by the inferred strengths of their adjacent upstream ribosome binding site and start codon. In all, up to 57 genes encoding 60 potential proteins have been identified (*56*). In the few cases where multiple start codons were postulated, the start codon most upstream was chosen to define the beginning of the protein coding domain.

A.4. Control Elements

Inference of Ribosome Binding Sites
Ribosome binding sites were postulated by analysis of the sequence data upstream of protein coding domain start codons. DNA sequence complementary to the *E. coli* 16S rRNA suggested a functioning RBS. Direct observation of proteins during T7 development confirmed RBS function for a subset of T7 proteins (*13*).

Promoters
At least 22 RNA polymerase promoters help to coordinate the dynamic allocation of gene expression resources during T7 infection. The promoters for T7 RNA polymerase and the major and minor host promoters (A1, A2, A3, B, C, and E) were first mapped by *in vitro* transcription studies (*16-23*) and subsequently confirmed by sequencing (*1, 24-31*). The *in vitro* transcription profile corresponds well to *in vivo* transcription data (*12, 32-34*). The individual contributions of the promoters to the timing and level of gene expression during wild-type infection is unclear.
The T7 RNA polymerase promoter structure was determined by sequencing the 23 base pair region common to the late T7 promoters (*26*). Here, we used a 35 base pair region to define T7 promoters; our broader definition of T7 promoter elements hoped to include conserved regions beyond the initial 23 base pairs (*1*).
The structure of an *E. coli* promoter is less well defined relative to a T7 promoter. For the major and minor *E. coli* promoters (A0, A1, A2, A3, B, C, and E) we defined regions of at least 60 bases, ranging from the –50 to +10 relative to the putative transcriptional start-site. A boxA recognition site located between A3 and gene *0.3* is postulated to be involved with causing anti-termination of polymerases coming from the three strong early promoters, A1, A2, and A3 (*35,36*).
The cloning of random sections of the T7 genome into a plasmid that selected for transcription activity from the cloned fragment identified other possible promoters (*37*). Sequence analysis in regions containing these sections identified region of homology to other known promoters (*1*). Any contribution of these additional promoters to wild-type

T7 infection is not now defined. While we annotated these promoters, we did not incorporate them as functional genetic elements of T7.1.

Terminators

Transcription termination plays an important role in regulating T7 development. The first transcription termination site was identified by mapping the endpoints of mRNA starting from *E. coli* promoters (*38*). Later it was shown that termination occurred at the same place *in vivo* and *in vitro* (*39*). The termination site was later mapped precisely, sequenced, and subsequently named 'TE' (*40,41*). A second terminator specific to T7 RNA polymerase was suggested by *in vitro* transcription studies on digested T7 DNA (*18,19*). The terminator, named Tø, was shown to function *in situ* (*41*) and on plasmids (*34*). Both TE and Tø have stem loop structures that are thought to set termination efficiency (*1*). The stem loop and flanking sequence, which includes the poly-uridine sequence, were taken together to form the element used in the refactoring process. Other terminators have been postulated, but the precise location and function, if any, during wild-type infection are tenuous (*1*).

RNaseIII sites

Sites for specific cleavage of RNA by RNaseIII were first shown by in vitro RNA assays and correlated to *in vivo* data (*39*). In time, 10 RNaseIII sites were mapped and their sites of cleavage were identified (reviewed in (*1*)). The sites are thought to stabilize the 3' end of T7 transcripts by providing a stem loop that prevents scanning single stranded RNA degradation enzymes from binding. The RNaseIII sites are often immediately followed by a downstream gene, and thus the element size was kept as short as possible. Minimally, the probable stem loop structures (*1*) were used as the element boundary.

Origins of replication

All putative origins of replications overlap with T7 RNA polymerase promoters. The primary origin was mapped to the dual promoter region downstream of *ø1.1A* and *ø1.1B* by analysis of replication bubbles in electron micrographs (*42,43*) and subsequently sequenced (*44,45*). The secondary origin at *øOL* was identified through deletion studies of the primary origin (*37,44*). Finally, a screen that tested cloned fragments of T7 DNA into plasmids to check for their ability to act as replication origins during T7 infection showed that *øOR* and *ø13* had origin activity (*1*). The precise boundaries of the origins are unknown, though are thought to be correlated to a functioning promoter (*46*). Only the primary origin is annotated and treated as an element. While none of the other replication origins are treated as elements, the promoters that they are associated with are, and thus are possibly conserved in T7.1.

Terminal Regions

Both left and right ends of the T7 genome contain an exact 160 base-pair direct repeats (*47*) that are thought to be involved in concatemer formation during DNA replication (*48*). Adjacent to the direct repeats on both ends is a region that contains 12 regularly arranged highly conserved seven base-pair sequences termed the SRL and SRR (Short-Repeat Left and Right) (*31*). Though the direct repeats, SLR, and SRR are thought to be involved in processes of DNA replication and packaging, the actual mechanisms through

which they act remains unclear. However, their importance to T7 development was assumed from the high degree of homology between the left and right end. Thus, we avoided any changes to the SLR and SRR.

B. Design of T7.1 Genome

B.1. Overview

There were three primary motivations that drove the design of T7.1. First, we wanted to maximize our ability to independently manipulate all the identified functional genetic elements (*Section A*). Second, we wanted system behavior to be as close to wild type as possible. Third, we wanted a system that produced viable phage. The second and third goals acted as constraints on the changes that could be made in pursuit of the first design goal.

The design for T7.1 is broken into six 'sections,' *alpha* through *zêta*. Each section was broken into 'parts' that contain one or more 'functional genetic elements' (Figure S1). Thus, the modification of parts across sections requires a two-stage process. First, parts are manipulated within sections. Second, refactored sections are combined with other natural or refactored sections. Section *alpha* was designed and constructed prior to the design of *beta* through *zêta*. Sections *beta* through *zêta* were subsequently designed with the added intent of avoiding problems encountered during the construction of section *alpha*.

Definitions used throughout:
> *Sections*: the six sections whose boundaries are 1-cutters of the wild-type T7 genome and together define the T7 genome (*alpha - zêta*).
> *Parts*: the set of 73 functional genetic elements that is surrounded by a pair of identical restriction sites.
> *Functional genetic elements*: promoters, genes, ribosome binding sites, etc., that have been defined in the re-annotation (see above).
> *Construct*: any amalgamation of functional genetic elements or parts.
> *#-Cutter*: A restriction enzyme that only cuts a particular DNA sequence # times

B.2. Sections

Sections were chosen to ease construction and to compartmentalize changes across the genome. There were two practical considerations that constrained the boundaries of the sections. First and foremost, the boundaries of the sections had to be compatible with the sparse distribution of 1-cutter sites across the wild-type genome. Using 1-cutter sites for section boundaries allows refactored sections to be combined with sections of wild-type DNA and tested individually or in combination. Second, the number of parts per section was limited by the number of 0-cutters across each wild-type section.

B.3. From Functional Genetic Elements to Parts

Despite our desire to independently manipulate each functional genetic element, it did not always make sense to define each element as an independent part. Certain elements are physically and functionally coupled. For example, the ribosome binding sites of all putative coding domains were grouped together with their cognate coding domains as single parts because separating them would probably disrupt functionality. Also, some functional genetic elements overlap so severely as to prevent efficient separation (e.g., genes *4A*, *4B*, *4.1*, and *4.2*). Other groups of functional genetic elements were short (<150bp) such that variants containing deletions or separations of the individual elements could be easily constructed (e.g., *E. coli* promoter C/ RNase III site R*1*). In all, the T7.1 genome was split into 73 parts. Parts were sequentially number one to 73 starting from the genetic left end.

Parts that overlapped were separated. Separation was carried out by first duplicating the overlapping sequence such that each part independently contains the sequence of its functional genetic elements. If either of the parts contains sequence that retains putative function associated with the functional genetic elements of the other part, that sequence was mutated to eliminate function. These mutations were made in coding domains by silent mutation. The silent mutations involved either no change in the tRNA recognized or, when necessary, changed to a tRNA that is known to be in higher abundance (*49*). All such separations are detailed in Figure S2.

Finally, each part was surrounded by a restriction site not contained elsewhere in the section. Typically, these sites were added between parts but when appropriate, they were mutated into pre-existing sequence to minimize perturbation to the wild-type sequence. In addition, to minimize the total number of bases added to T7.1, we chose adjacent restriction sites to have overlapping sequence with one another.

One of the most significant changes between the design of section *alpha* and the other sections was in the choice of restriction sites used. In section *alpha*, we picked restriction enzymes that did not cut within section *alpha* only. However, as the construction of *alpha* proceeded, and cloning directly into the phage became useful, it was clearly advantageous to use restriction enzymes that did not cut within the entire genome where possible.

B.4. Design Features

Deletion and Insertion
The genome allows for simple deletions of parts. Generally, the section that the part is in can be isolated and digested by the corresponding restriction enzyme. The fragments can then be ligated to reform the section minus the deleted part, and joined to the rest of the genome.

Insertion of new parts is more involved. If there is a pre-existing restriction site between parts due to a previous part deletion, then a new part can be inserted in its place. If no such site exists, another method involves using two restriction enzymes, NgoMIV and BspEI, that are 0-cutters across both the wild-type T7 and all refactored sections.

NgoMIV and BspEI have different recognition sequences but produce the same overhang upon digestion. This allows for ligation of a product into these sites, while simultaneously preventing the restriction sites from being reformed. Thus, a part adjacent to the desired insertion site must be replaced with the same part that has an NgoMIV site appended to it. Then the part to be inserted is amplified with bracketing BspEI sites and inserted into the NgoMIV site. Since neither restriction site is reformed upon insertion, this method can be reused to serially insert parts throughout the genome.

Unstuffing Hooks
Due to the unknown consequences of part separation, the ability to return to the wild-type sequence would be useful for comparison and debugging purposes. As such, where needed, restriction sites that cut only once across the section were mutated into the genome by silent mutation. These restriction sites when combined with other naturally occurring 1-cutters within the section, could be used to replace refactored regions with wild-type sequence. These restriction sites in section *alpha* were labeled U1-U4. In later sections, these extra sites were not added as they were unnecessary because 0-cutters across the genome are used to bracket parts.

Scaffolds
A scaffold was employed to ease assembly of each refactored section. The scaffold is essentially the sequence that remains when all parts are removed from the section. As such, the scaffold contains all the restriction sites required to assemble the parts to form the section. In addition, if a fully refactored phage was not viable, we could use the scaffold to incrementally revert the sequence back to wild type in an attempt to restore function.

B.5. T7.1 Annotation and Sequence

GenBank files for the annotation of T7.1 and the as-built section *alpha* and *beta* are available upon request and are being deposited with NCBI. These three files plus a GenBank file of our re-annotation of wild-type T7 are now available online: http://web.mit.edu/endy/www/ncbi/

C. Construction

C.1. *Alpha*

The *alpha* scaffold included all restriction sites required for assembly of refactored section *alpha* and also included elements that were inexpensive to synthesize or difficult to clone. The scaffold included all functional genetic elements from the left end of T7 through and including R*0.3*, R*0.5*, TE, *ø1.5* and *ø1.6*. This 1334 bp sequence was sent to Blue Heron Biotechnology for synthesis.

The scaffold was delivered by Blue Heron in four fragments with point mutations in each fragment. They are described as follows:

Fragment 1: Single base changes at 89(G-T), 168(A-T), 169(C-A), 245(G-A) and 249(C-A) as well as single base deletions at 138 and 159
Fragment 2: A single base deletion in the -35 box of the *A1* promoter
Fragment 3: A four base deletion between the -35 and -10 boxes of the *A3* promoter
Fragment 4: A single base deletion in the loop of TE

The sheer number of mutations in Fragment 1 rendered it useless for our purposes. The mutations in Fragments 2, 3 and 4 were not as severe and we proceeded to simultaneously correct and utilize these scaffold fragments.

A vector, pREB, was constructed to facilitate the assembly of section alpha. The vector was created from pSB2K3-1, a chimera of pSCANS-5 and pSB1A3-1 (Materials and Methods). The multiple cloning site (MCS) of pSB2K3-1 was replaced with a smaller MCS containing PstI, BstBI and BclI restriction endonuclease sites. Nineteen other restriction sites were also removed by multi-site directed mutagenesis in order to facilitate the cloning of parts into the scaffold. All sites were successfully deleted except for two overlapping EcoO109I sites that we could not remove. pREB retains the inducible copy control of pSCANS-5 that is useful for working with difficult to clone DNA fragments.

The following parts were first cloned into pSB104 as BioBricks (*50*) with their flanking scaffold site: Parts 5, 6, 7, 8, 12, 13, 14, 15, 16, 18, 20, 22 and 24. Part 11 was cloned into pSB2K3. Site directed mutagenesis was then carried out on parts 6, 7, 14 and 20 to introduce the sites U1, U2, U3 and U4 respectively. Site directed mutagenesis on part 20 failed.

A single Eco0109I restriction site was removed by site-directed mutagenesis from vector pUB119BHB carrying scaffold Fragment 4. Part 15 was subsequently cloned into this modified vector. Scaffold Fragment 4 was then transferred by cloning into pREB. The following parts were serially cloned into this vector: 7, 8, 12, 13, 14, 16, 18, 20, 22 and 23. The populated scaffold Fragment 4 was then digested with restriction enzymes NheI and BclI and purified.

Parts 5 and 6 were cloned into pUB119BHB carrying scaffold Fragment 3. This populated Fragment 3 was subsequently used for *in vitro* assembly of a construct spanning from the left end of T7 to part 7. *In vitro* assembly of this construct began with digestion of wild-type T7 genomic DNA with AseI and isolating the 388bp left end fragment and ligating this to scaffold Fragment 2. The correct ligation product was selected by PCR. The mutation in part 3 (A1) was then fixed by PCR ligation by a two-step process. First primers with the corrected sequence for part 3 were used to amplify the two halves of the construct to the left and right of part 3. A subsequent PCR ligation was carried out to join these two constructs. Scaffold Fragment 3 was then added to the above left-end construct once again by PCR ligation as described above. The mutation in part 4 (A2, A3 and R0.3) was repaired similarly to the mutation in part 3. The right most

primer used to amplify the entire construct contained an MluI site on the tail that was then used to ligate on a copy of part 7. The ligation product was again selected by PCR. This populated left-end construct was then digested by NheI and purified.

The right arm of a BclI digestion of wild-type T7 genomic DNA was then isolated and ligated to the populated left end construct and the populated Scaffold Fragment 4. The three-way ligation product was then transfected into IJ1127 (Materials and Methods). Plaques were used to create lysates and DNA was purified and digested to screen for desired clones (Materials and Methods).

Part 11 was then cloned into the rebuilt section using the same method of three-way ligation followed by transfection. Cloning of part 9 required the *in vitro* assembly of a construct that spanned part 6 through part 9 because RsrII(D9L/R) cuts wild-type T7 elsewhere. The construct was created by amplifying the region spanning part 5 through part 12 of the refactored genome by PCR. The PCR product was then digested with RsrII and ligated to part 9. The correct ligation product was selected by PCR with a primer on the right end that contained a SacII site in the tail. This PCR product was then digested with SacI and SacII and cloned into the phage as described above. Lastly, the part 10 was cloned into the SacII site of the phage.

C.2. *Beta*

We constructed section *beta* using a process similar to that used with *alpha*. A scaffold with all restriction sites as well as part 26 was made by Klenow extension of overlapping primers. The product was digested with BstBI and cloned into pREB. The following parts were then cloned into this vector: 23, 24, 27, 28, 30, 31 and 32. Part 32 (gene *3.8*) had to be cloned as a truncation since we were unable to clone the full length part probably due to the previously reported toxicity of gene *3.8* product (*35*). The truncated version of part 32 still included the BglII site to allow for assembly of section *beta* into a phage. Parts 25 and 29, which were also previously reported to be toxic, were assembled *in vitro*. To insert the part 25, we amplified a region spanning part 23 through part 27 by PCR. This fragment was then digested with BsiWI and part 25 was ligated to each of these fragments separately and selected for by PCR. These two PCR products were then digested with DraIII, a restriction site internal to part 25, ligated and then selected by PCR. The overall fragment was then digested with BclI and MluI, purified, and ligated to wild-type fragments on the left and right. The same method was used to insert part 29 by using the part 29 internal restriction site EcoO109I and then digesting this overall fragment with MluI and BglII for cloning into a phage. Lastly, these two phage genomes were digested with MluI and the left fragment of the genome containing the refactored region spanning part 23 to 27 was ligated to the right fragment of the genome containing the refactored region spanning from part 27 to 32.

C.3. Synthesis & Construction Errors

Differences from the designed and constructed sections *alpha* and *beta* are detailed in Tables S1 and S2.

Materials and Methods

1. Strains
2. Media Recipes
3. General Protocols
4. Genome Design Tools
5. Genome Construction Protocols

1. Strains

E. coli
BL21: B *hsdS Gal-*
BR3: B *rpoC-E2258K*
D1210: HB101 *lacIq*
DH5alpha: φ80*lacZΔM15 Δ(lacZYA-argF) U169 endA1 recA1 hsdR17 (rk-, mk+) thi-1 gyrA96 relA1 phoA*
DH10B: *mcrA Δ(mrr-hsdRMS-mcrBC)* φ80*lacZΔM15 ΔlacX74 deoR recA1 araD139 Δ(ara leu)7697 galU galK rpsL endA1 nupG*
IJ1126: *E. coli* K-12 *recB21 recC22 sbcA5 endA gal thi* Su+ Δ(*mcrC-mrr*)*102::*Tn*10*
IJ1127: IJ1126 *lacUV5 lacZ::*T7 *gene1-*Knr

Phage
T7+, wild-type bacteriophage T7, was a gift of Dr. Ian Molineux (UT Austin)

2. Media Recipes

L-broth or LB Medium (Luria-Bertani Medium) (*55*)
      10 g Bacto-tryptone
      5 g yeast extract
      10 g NaCl
      distilled water up to 1 L

1.5% T-agar (*54*)
      10 g Bacto-Tryptone
      5 g NaCl
      15 g Bacto-agar
      distilled water up to 1 L

0.7% T-agar (*54*)
      10 g Bacto-Tryptone
      5 g NaCl
      7 g Bacto-agar
      distilled water up to 1 L

50X TAE Electrophoresis Buffer (*55*)
      242 g Tris base
      57.1 ml glacial acetic acid
      100 ml 0.5 EDTA pH 8.0
      distilled water up to 1 L

T7 Buffer (*54*)
      0.1 M Tris-HCl pH 7.5
      1 M NaCl
      1 mM EDTA pH 7.5

TES Buffer (*54*)
      50 mM NaCl
      50 mM Tris-HCl pH 7.5
      5mM EDTA pH 7.5

ρ=1.43 cesium chloride (*54*)
      33 g cesium chloride
      50 ml 10 mM Tris-HCL pH 7.5, 10mM MgCl-2

ρ=1.53 cesium chloride (*54*)
      41 g cesium chloride
      50 ml 10 mM Tris-HCL pH 7.5, 10mM MgCl-2

ρ=1.62 cesium chloride (*54*)
      50 g cesium chloride
      50 ml 10 mM Tris-HCL pH 7.5, 10mM MgCl-2

TE (*55*)
      10mM Tris-Cl pH 8.0
      1mM EDTA pH 8.0

3. General Protocols

Plating of T7
T7 was plated by adding various dilutions of a phage stock to 200uL of saturated BL21 culture and 3mL of molten (46°C) 0.7% T-agar and pouring the mixed contents onto 1.5% T-agar plates. Plaques appeared after 3-5 hours of incubation at 37°C.

Isolation of T7 Genomic DNA From Crude Cell Lysates
T7 genomic DNA was isolated according to the protocol described in Rene Garcia's dissertation (*54*), reproduced here:

    1.   Grow 40 ml of permissive cells to a density of 108-109 cells/ml at 37°C in a rotary shaking water bath. Inoculate the cells with a drop from a master phage stock. Continue to shake cells in the water bath at 37°C until the culture clarifies.

2. Add NaCl to a final concentration of 1 molar. Centrifuge the lysate at 10,000 rpm for 10 min. Discard the cellular debris, and centrifuge the lysate at 24,000 rpm for 90 min in a SW28 rotor (Beckman).

3. Discard the supernatant, and add 1 ml of T7 buffer or TES buffer to the phage pellet.

4. Let the pellet site at 4°C for at least 5 hours. Resuspend the pellet, and quickly spin down the cellular debris. Discard the pellet.

5. To the supernatant, add 0.5 ml of 50 mM Tris-HCl, pH 8 saturated phenol and gently mix the sample until an emulsion forms. Quickly microfuge the sample to separate the layers. Carefully remove the aqueous layer without disturbing the organic layer. Phenol extract with 0.5 ml 50 mM Tris-HCl, pH 8 saturated phenol one more time.

6. To the aqueous layer, add 0.5 ml of 50mM Tris-HCl, pH 8 saturated phenol: chloroform: isoamyl alcohol (25:34:1 by volume) mixture, and gently mix the sample until an emulsion forms. Quickly microfuge the sample to separate layers. Carefully remove the aqueous layer without disturbing the organic layer. Phenol: chloroform: isoamyl alcohol and extract the aqueous layer one more time.

7. Add 3 times volume of 95% ethanol alcohol to the sample. A fibrous precipitate should form. Spin down the precipitate, remove the supernatant, and wash the pellet with ethanol. Dry the pellet. Dissolve the pellet in 500 uL of water or TES buffer. There should be about 109 molecules of phage DNA/uL. Store the DNA at -20°C.

Purification of T7 Particles via CsCl-gradient Centrifugation
T7 particles were purified by cesium chloride gradient centrifugation according to Garcia's protocol (*54*). The protocol is reproduced here:

1. Grow 100ml of permissive cells to a density of 108 to 109 cells/ml at 37°C in a rotary shaking water bath. Inoculate the cells with a drop from a master phage stock. Continue to shake cells in the water bath at 37°C until culture clarifies. [NOTE – As a standard laboratory protocol, T7 stocks have always been propagated at 30°C; however, at this temperature cultures infected with (A1, A2, A3)- T7 mutants take longer to clarify that those infected with (A1, A2, A3)+ phages or with mutants that eject their DNA faster. At 37°C the differences in lysis periods are not as pronounced. Stocks of (A1, A2, A3)- T7 mutants are propagated at 37°C to decrease the growth disadvantage of spontaneous arising mutants that eject their DNA faster. For constancy (A1, A2, A3)+ phages are also grown at this higher temperature.]

2. Add NaCl to the lysate to make the final concentration 1 molar. Centrifuge the lysate at 10,000 rpm for 10 min, Discard the cellular debris, and add 10 grams polyethylene glycol (PEG) m.w. 8000 (10% w/v) to the supernatant. Gently stir the mixture until the PEG has totally dissolved. Keep lysate on ice for 1 hour.

3. Pellet the phage at 5,000 rpm for 15 min. Decant the supernatant, and very gently resuspend the pellet in 3.5 ml of T7 buffer. Centrifuge the lysate at 5,000 rpm for 10 min, and keep the supernatant.

4. Pour a cesium chloride step gradient: ass .5 ml of cesium chloride with a density of 1.6 to the bottom of a centrifuge tube that fits in a SW 40.1 rotor. Gently layer 0.5 ml of cesium chloride ρ=1.5 onto the ρ=1.6 layer. Finally ass 0.5 ml of cesium chloride ρ=1.4 onto the ρ=1.5 layer.

5. Gently layer the phage supernatant onto the cesium chloride step gradient. Centrifuge the phage in a SW 50.1 rotor at 30,000 rpm for 2 to 3 hours. The phage will band at the ρ=1.5 layer.

6. Remove the phage band from the side of the tube with a syringe.

7. Remove the cesium chloride by dialysis against 0.5 to 1 liter of T7 buffer at 4°C.

Purification of T7 Genomic DNA From CsCl-gradient Purified Phage Particles
T7 particles were purified by cesium chloride gradient centrifugation (above). The DNA was then purified by subsequent rounds of phenol and phenol:chloroform extraction as

follows: pH7.8 phenol (55) was added in a 1:1 volume ratio to the sample and the tube was inverted to mix the aqueous and organic phases. The mixture was centrifuged for 10 minutes at 13,000g. The aqueous layer was removed and subjected to an additional round of phenol extraction. This resulting aqueous layer was added in a 1:1 volume ratio to pH7.8 phenol:chloroform:isoamyl alcohol (25:24:1, Sambrook and Russell:A1.23), mixed and centrifuged for 5 minutes. This extraction step was repeated again. The DNA was precipitated by adding a 10% sample volume of 3M sodium acetate and 2-5 sample volumes of cold (4°C) absolute ethanol. The samples were mixed and incubated at -80°C for 1 hour. The DNA was then pelleted by centrifugation for 30 minutes at 13,000g and at 4°C. The DNA pellet was washed once with 80% ethanol, dried and resuspended in TE buffer.

Purification of Restriction Enzyme Digested Fragments
All restriction enzyme digestions were carried out according to the manufacturer's directions. All fragments smaller than 10kb were purified using Qiaquick gel extraction kit (Qiagen). All fragments larger than 10kb were purified by electro-elution as follows: 20ug of digested product was preincubated with 1uL of a 1000X solution of SYBR Gold (Molecular probes) for 15 minutes. 200ng of DNA was loaded into each well of a 0.5% TAE agarose gel and electrophoreised at 1-1.5V/cm and at 4°C for 16-20 hours. Agarose blocks containing desired restriction fragments were excised under UV transillumination and loaded into a dialysis bag (3500 MWCO Snakeskin dialysis tubing, Pierce) containing 1X TAE. Fragments were electro-eluted from the agarose blocks for 1-5 hours at 5V/cm. After the completion of electro-elution was confirmed by UV visualization, the electric field was reversed for 1 minute to aid in elution. The liquid contents of the bag were then subjected to one round of phenol extraction to remove trace amounts of agarose and the DNA was ethanol precipitated and resuspended in TE buffer.

Plating of Phage for Comparative Plaque Analysis
Stocks of cesium chloride purified phage were serially diluted to an appropriate titer. 50, 100 or 200uL of that dilution was mixed with 200uL of saturated BL21 culture, added to 12mL of molten (50°C) 0.7% T-agar and plated directly on Petri dishes. Plaques were allowed to grow for 5-48 hours at 30°C or 37°C.

Measuring Phage Lysis Curves
1mL containing 2x108 cells of BL21 was infected at a MOI of 5 and 200uL of the resulting mixture was loaded per well into a 96 well ViewPlate (Packard) at 30°C. Mineral oil was layered into each well and the OD was monitored at 30°C with agitation by a Wallac Victor2 plate reader (Perkin-Elmer).

Sequencing of DNA
All sequencing of phage DNA was performed using the dideoxy terminator method. All sequencing was performed by the MIT Biopolymers Laboratory using a Perkin Elmer Applied Biosystems Division model 377 DNA sequencer. When long regions of DNA were sequenced, primers were designed at 500-800bp intervals to both sense and antisense strands. All reported sequence represents at least two separate sequence runs with no intervening ambiguities.

Template preparation of phage genomic DNA for sequencing
For phage sequencing, only full length packaged genomic DNA (preparation described above in section 3) was used as template.

Template preparation of cloned parts for sequencing
The preparation of sequencing template for cloned parts was done using Qiaprep spin Miniprep Kit (Qiagen). When the quantity of purified plasmid was insufficient for sequencing, a subsequent Templiphi (Amersham) reaction was used to amplify the sequencing template.

Template preparation of in-vitro constructs for sequencing
In-vitro constructs were amplified using PCR, gel purified and used as template in a sequencing reaction.

4. Genome Design and Sequence Analysis Tools

Restriction Site Distribution
All design and annotation of DNA constructs was done using Vector NTI (InforMax). All restriction analyses were performed with Vector NTI (InforMax), NEBcutter (*51*), and REBASE (*52*). A perl script was written to search for sites within coding regions where restriction sites could be introduced by silent mutation (http://web.mit.edu/endy/www/software/cuts/).

Sequencing Analysis and Contig Assembly
Sequence analysis and contig assembly was done with AlignX and Contig Express (InforMax).

5. Genome Construction Protocols

Oligonucleotide Synthesis
All oligonucleotides were synthesized by MWG, Invitrogen or using an ABI Model 394 DNA synthesizer (Tom Knight).

Part Amplification
All parts were amplified by PCR using the following reaction mixture: 5uL 10X Thermo Pol Buffer (NEB), 20pM primer1, 20pM primer2, 3-30ng T7 genomic DNA, 1unit Vent polymerase (NEB), 10uM each dNTP and water to 50uL. The mixture was thermocycled (MJ Research PTC-200) as follows: 95°C for 2 minutes, 25-35 cycles of 95°C for 30 seconds, 50°C-60°C for 30 seconds, 72°C for 1-5 minutes, 72°C for 10 minutes.

Part Cloning
All parts and vectors (0.1 - 50pmoles) were restriction enzyme digested according to the manufacturers' directions (NEB, Fermentas). Parts and vectors were then purified by gel electrophoresis (0.5 - 2% TAE agarose gel, 3 - 8 V/cm) and extracted with Qiaquick gel extraction kit (Qiagen). Ligation reactions using T4 DNA ligase (NEB) were carried out

in a 3:1 part:vector molar ratio according to the manufacturer's directions. Ligation products were dialyzed on nitrocellulose membranes (Millipore) against 1000X volume of water for 30 minutes. Ligation products were transformed by electroporation using 1800V across a 1mm gap (Bio-Rad Gene Pulser Xcell) and plated on the appropriate medium. Screening for clones was performed by colony PCR with the following protocol: Colonies were picked and diluted in 100uL of water. 1uL of that cell suspension was added to 1uL 10X Thermo Pol Buffer (NEB), 4pM primer1, 4pM primer2, 0.5U Taq Polymerase (NEB), 2uM each dNTP and water to 10uL. This mixture was thermocycled as follows: 95°C for 6 minutes, 25-35 cycles of 95°C for 30 seconds, 54°C for 30 seconds, 72°C for 1-5 minutes, 72°C for 10 minutes.

Site Directed Mutagenesis of Parts
Site specific changes were performed on the cloned parts using QuickChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's directions. Primers were 5' phosphorylated using Polynucleotide Kinase (NEB) according to the manufacturer's directions.

Construction of pREB Plasmid
pREB was constructed from pSB2K3-1 (*53*), a chimera of pSCANS-5 (gift of John Dunn, Brookhaven National Laboratory) and pSB1A3-1 (*53*). The multiple cloning site of pSB2K3-1 was replaced by a PstI-BstBI-BclI multiple cloning site by primer annealing and cloning. Primer duplexes were prepared using the following steps: the reaction mixture, 100pM each primer, 2uL restriction buffer (NEB) and distilled water to 20uL was incubated as follows: 95°C for 4 minutes, 0.1°C/s ramp to 80°C, 80°C for 4 minutes, 0.1°C/s ramp to 70°C, 70°C for 4 minutes, 0.1°C/s ramp to 60°C, 60°C for 4 minutes, 0.1°C/s ramp to 50°C, 50°C for 4 minutes, 0.1°C/s ramp to 22°C, 22°C for 10 minutes. The annealed duplexes were 5' phosphorylated using Polynucleotide Kinase (NEB). pREB was then cleaned of restriction sites using QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene) and screened by digestion.

Construction and Cloning of the Beta Scaffold
The beta scaffold was constructed by annealing two partially overlapping primers as described above. The overhangs were then filled in using Klenow fragment (NEB) extension according to the manufacturer's directions. The extension product was digested with BstBI and cloned into pREB.

Assembly of Section Fragments in E. coli:
Parts were cloned into the scaffold using the same cloning method as above with the addition of treating the purified cut vector with Antarctic Phosphatase (NEB) according to the manufacturer's directions and using a molar ratio of 6-10:1 insert:vector in the ligation reaction. Screening was again performed by colony PCR but included a second PCR to verify directionality via an internal primer.

Assembly of Section Fragments in vitro
Fragments were assembled *in vitro* using both PCR ligation and traditional T4 DNA ligation. T4 DNA ligation products were subsequently selected by PCR. All

amplification was carried out as described above using either Taq polymerase, Vent Polymerase or a 99:1 Taq:PfuTurbo (Stratagene) enzyme mixture. Selection of ligation products was carried out using various serial dilutions of the ligation products as template.

Ligation of 2 or 3 DNA Fragments for Phage Transfection
$4x10^9$ molecules of each DNA fragment was ligated using T4 DNA ligase (NEB) and incubated at 16°C overnight. In certain cases, up to $4x10^{10}$ molecules of a particular DNA fragment was added to drive the reaction towards the desired outcome.

Preparation of Competent Cells for Phage Transfection
Competent cells were prepared according to Garcia's protocol (*54*) and allowed to rest at 4°C for 20-24 hours prior to transfection. The protocol is reproduced here:

1. Grow 20 ml of cells to a density of 5×10^8 cells/ml in L broth at the desired temperature.
2. Pellet the cells in a centrifuge at 5,000 rpm for 5 min. Remove the supernatant.
3. Resuspend the cell pellet in 10 ml of ice-cold 50mM CaCl2 (half volume of starting culture).
4. Incubate the cells on ice for 30 min.
5. Pellet the cells in a centrifuge at 5,000 rpm for 5 min. Remove the supernatant.
6. Resuspend the cell pellet in 2 ml of ice-cold 50 mM CaCl2 (one-tenth volume of the starting culture). The cells are ready to take up DNA.

Transfection of the Ligation Products
All pipette tips were pre-chilled at -20°C, molten 0.7% T-agar was kept at 46°C and 1.5% T-agar plates were equilibrated to room temperature. The ligation mixture was added to 200uL of cold (4°C) competent cells and incubated in an ice bath for 30 minutes. The mixture was then added to 2.5mL of molten (46°C) 0.7% T-agar, gently mixed for 10 seconds by manual agitation and poured onto a 1.5% T-agar plate. The plates were then incubated at 37°C for 3-5 hours.

Figure S1. Genome design. We split the wild-type T7 genome into six sections, alpha through zêta, using five restriction sites unique across the natural sequence. Each section shown here has a wild-type section with representations of the genetic elements: protein coding regions (blue), ribosome binding sites (purple), promoters (green), RNase III recognition sites (pink), transcription terminators (yellow), and others (gray). Elements are labeled by convention (1). Images are not to scale, but overlapping boundaries indicate elements with shared sequence. The useful natural restriction sites across each section are shown (black lines). T7.1 sections are shown below the wild-type sections. Parts are given integer numbers, 1 through 73, starting at the left end of the genome. Unique restriction site pairs bracket each part (red/blue lines, labeled D[part #]L/R). Added unique restriction sites (purple lines, U[part #]) and part length (# base pairs, open boxes) are shown.

Section *Alpha:*
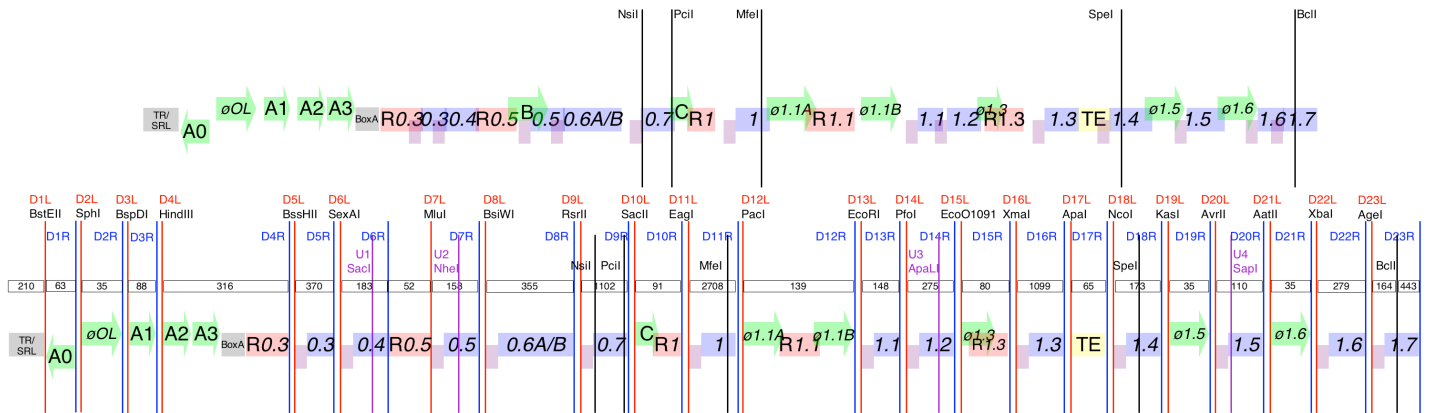
Section *Beta:*
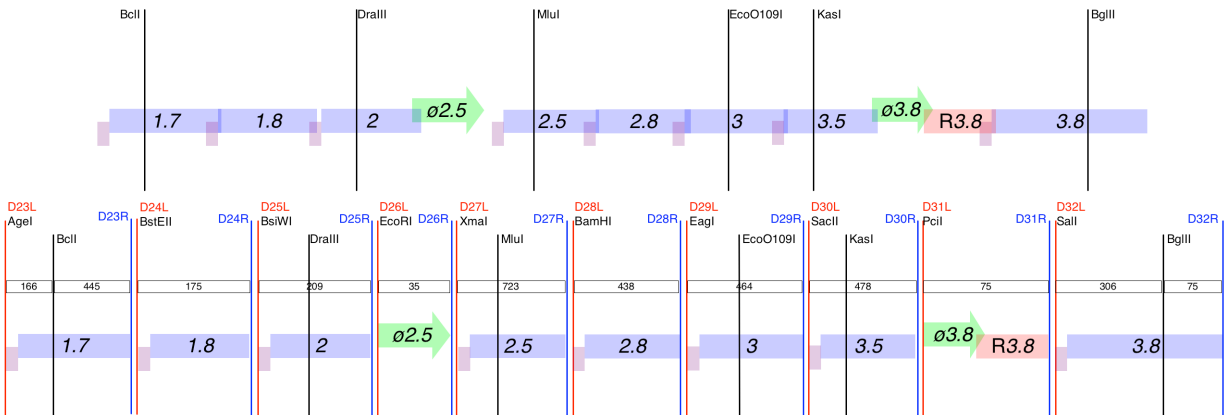
Section *Gamma:*

Section *Delta:*

BstEII DraI BspHI   NsiI RsrII   AseI   NdeI BssHII   EciI AatII   AvaI PacI

| 7.7 | 8 | ø9 | 9 | ø10 | 10A | Tø | 11 | 12 |

D51L PstI   D51R BsiWI   D52L   D52R PvuI   D53L   D53R EcoRI   D54L   D54R SacII   D55L   D55R BamHI   D56L XbaI   D56R   D57L XmaI   D57R ApaI   D58L   D58R   D59L HindIII   D59R

BstEII   DraI BspHI   PvuI   NsiI RsrII   SacII   AseI   XbaI NdeI   BssHII   EciI AatII   AvaI PacI

| 223 | 172 | 1611 | 35 | 1038 | 35 | 464 | 73 | 591 | 2382 | 3 |

| 7.7 | 8 | ø9 | 9 | ø10 | 10A | Tø | 11 | 12 |

Section *Epsilon:*

PacI   NciI AvaI BstBI ScaI ApaLI

| 12 | R*13* ø13 | 13 | 14 | 15 | 16 |

D59L HindIIl   D59R PfoI   D60L   D60R EcoRI   D61L   D61R BsiWI   D62L   D62R PvuI   D63L   D63R EagI   D64L   D64R

PacI   NciI   AvaI   ScaI BstBI   ApaLI

| 2382 | 3 | 62 | 417 | 591 | 2244 | 3520 | 437 |

| 12 | R*13* ø13 | 13 | 14 | 15 | 16 |

Section *Zêta:*

ApaLI XbaI   FspI BglI KasI   BlpI   BspDI   DraIII   AcvI Acil   KpnI

| 16 | ø17 | 17 | 17.5 | 18 | R*18.5* | E | 18.5/18.7 | 19/19.2/19.3 | øOR | 19.5 | SRR/TR |

D64L EagI   D64R SacII   D65L   D65R HindIII   D66L   D66R BamHI   D67L XmaI   D67R   D68L   D68R ApaI   D69L EcoRI   D69R   D70L   D70R BsiWI   D71L   D71R PstI   D72L SalI   D72R   D73L   D73R

ApaLI XbaI   FspI KasI   BglI   BlpI   BspDI   DraIII   AcvI AciI   KpnI

| 3520 | 437 | 35 | 1662 | 204 | 270 | 53 | 460 | 1761 | 35 | 150 | 160 |

| 16 | ø17 | 17 | 17.5 | 18 | R*18.5* | E | 18.5/18.7 | 19/19.2/19.3 | øOR | 19.5 | SRR/TR |

Figure S2. Differences between wild-type T7 and T7.1. A listing from left to right on the T7 genome of the changes made during design of T7.1. Changes are shown by comparison of the annotated wild-type T7 (above) and T7.1 (below) sequences. Point mutations are capitalized. The natural ribosome binding sites are underlined. The bracketing restriction sites surrounding parts are orange (left cutter) and blue (right cutter), with overlaps of neighboring bracketing restriction sites in light green. Other features include start codons (green), stop codons (red), and overlaps of start and stop codons (purple).

Changes in Section *Alpha*

```
LeftEnd-Part1: TR-SRL/A0
T7:      gagtgtctctctgtgtccctatctgttacagtctcctaaagtatcctcct
         ------TR---------->               <-------SRL-----
T7.1:    gagtgtctctctgtgtccctatcGgttacCgtctcctaaagtatcctcct
         ------TR---------->     <-D1L->   <-------SRL-----


Part1-Part2: SRL/A0-øOL
T7:      acctaaagacgccttgttgttagccataaagtgataacctttaatcattgtctttattaa
         --SRL-->                               <-----øOL----
T7.1:    acctaaagGTTaccgcatgcttgttgttagccataaagtgataacctttaatcattgtctttattaa
         --SRL-->        <D2L->                              <-----øOL----
                 <-D1R->


Part2-Part3: øOL-A1
T7:      aaggagagacaacttaaagagacttaaaagattaatttaaaatttatcaaaaag
         ---ø0L-->                               <---A1---
T7.1:    aaggagagacaacttaaagagCAtgcttaaaagattaatcgattaaaatttatcaaaaag
         ---øOL-->            <D2R->        <D3L->     <---A1---


Part3-Part4: A1-A2/A3/BoxA/R0.3
T7:      gagagggacacggcgaatagccatcccaatcgacaccgggggtcaaccggataagtagacagcctgataagtcgcacgaaaaacagg
         ----A1---->                                          <----A2----
T7.1:    gagagggacacggcgaatagccatcccaatcgaTaccgggggtcaaccggataagtagaAagcTtgataagtcgcacgaaaaacagg
         ----A1---->                  <D3R->             <D4L->         <----A2----


Part4-Part5: A2/A3/BoxA/R0.3-0.3
T7:      gatattcactaataactgcacgaggtaacacaagatggctatgtctaaca
         ---R0.3--->          <-0.3RBS->     <------0.3------
T7.1:    gatattcactaagcttgcgcgctgcacgaggtaacacaagatggctatgtctaaca
         ---R0.3--->      <D5L->  <-0.3RBS->       <------0.3------
                  <D4R->


Part5-Part6: 0.3-0.4
T7:      cgaggagtacgaggaggatgaagagtaatgtctactacc
         -------------0.3----------->
                    <-0.4RBS->           <----0.4----
T7.1:    cgaggagtacgaggaggaCgaagagtaagcgcgcaccaggtcgaggagtacgaggaggatgaagagtaatgtctactacc
         -------------0.3-----------><D5R-><D6L->           <-0.4RBS->        <----0.4----


U1-R0.5: U1/0.4-R0.5
T7:      caaagaactgtacgaaaacaacaaggcaatagctttagaatctgctgagtgatagactcaaggtc
         -------------------------------------------0.4---->
                                                   <-------R0.5-------
T7.1:    caaagaGctCtacgaaaacaacaaggcaatagctttagaatctgctgagtgaaccaggtgagtgatagactcaaggtc
         -------------------------------------------0.4----><-D6R-><-------R0.5-------
             <-U1->


R0.5-Part7: R0.5-0.5
T7:      gcctttatgattatcactttacttatgagggagtaatgtatatgctt
         -------R0.5------->     <-0.5RBS-><----0.5----
T7.1:    gcctttatgattatcacttacgcgtcttatgagggagtaatgtatatgctt
         -------R0.5-------><D7L->     <-0.5RBS-><----0.5----


U2
T7:      gctctaggtctagctgtaggtgcatcc
         ---------------0.5---------
T7.1:    gctctaggGctagctgtaggtgcatcc
         ---------------0.5---------
                 <-U2->


Part7-Part8: 0.5-0.6A/B
T7:      catcaaagqqqgcactacgcaaatgatgaagcac
         ---------0.5------------>
             <-0.6RBS->           <----0.6----
T7.1:    catcaaagqCgcactacgcaaatAaacgcgtacgcaaagqqgcactacgcaaatgatgaagcac
         ---------0.5------------><D7R->   <-0.6RBS->        <----0.6----
                                      <D8L->


Part8-Part9: 0.6A/B-0.7
```

```
T7:
          aacaggcactagccaacacactgaacgctatctcataacgaacataaaaggacacaatgcaatgaacattacc
          ----0.6---->                              <-0.7RBS->      <----0.7----
T7.1:
          aacaggcactagcgtacggtccgcgaacataaaggacacaatgcaatgaacattacc
          ----0.6----><D8R>           <-0.7RBS->      <----0.7----
                          <-D9L->
```

Part9–Part10: *0.7*–C/R*1*
```
T7:       caacattgataagcaacttgacgcaatgttaatgggctgatagtcttatct
          ------------------0.7-----------------> <----R1---
             <------------------------C--------------------
T7.1:     caacattgataagcaacttgacgcaatgttaatgggctgacggtccgccgcggattgataagcaacttgacgcaatgttaatgggctgatagtcttatct
          ------------------0.7----------------->< -D9R-><D10L><------------------------C--------------------
                                                                                                <----R1---
```

Part10–Part11: C/R*1–1*
```
T7:       ataggtacgatttactaactggaagaggcactaaatgaacacgatt
          ----R1----->            <-1RBS->    <-----1-----
T7.1:     ataggtacgatttactaaccgcggccgctggaagaggcactaaatgaacacgatt
          ----R1----->       <D10R>    <-1RBS->   <-----1-----
                              <D11L>
```

Part11–Part12: *1*–ø*1.1*A/R*1.1*/ø*1.1*B
```
T7:       gcgttcgcgtaacgccaaatcaatacgactcactatagagggacaaac
          -----1----->                      <---R1.1----
               <---------------ø1.1A---------------->
T7.1:     gcgttcgcgtaacggccggttaattaaaacgccaaatcaatacgactcactatagagggacaaac
          -----1-----><D11R><-D12L-><---------------ø1.1A---------------->
                                                     <---R1.1----
```

Part12–Part13: ø*1.1*A/R*1.1*/ø*1.1*B–*1.1*
```
T7:       tataggagaaccttaaggtttaactttaagacccttaagtgttaattagagatttaaattaaagaattactaagagaggactttaagtatgcgtaacttc
          ---ø1.1B--->                                                    <-1.1RBS->     <----1.1----
T7.1:     tataggagaaccttaaggtttaactttaagacccttaagtgttaattaAagatttaaattaaagaattCctaagagaggactttaagtatgcgtaacttc
          ---ø1.1B--->                                 <-D12R->           <D13L>  <-1.1RBS->     <----1.1----
```

Part13–Part14: *1.1–1.2*
```
T7:       ctgggagggtcagtaagatggggacgttta
          ------1.1------> <----1.2----
          <-1.2RBS->
T7.1:     ctgggagggtcagtaagaattccaggactgggagggtcagtaagatggggacgttta
          ------1.1------><D13R>      <-1.2RBS->       <----1.2----
                             <D14L>
```

U3
```
T7:       gacgaggacgttctgttcaatatgtgtactgattggttgaaccat
          ----------1.2--------------------------------
T7.1:     gacgaggacgttctgttcaatatgtgCactgattggttgaaccat
          ----------1.2--------------------------------
                              <-U3->
```

Part14–Part15: *1.2*–ø*1.3*/R*1.3*
```
T7:       gttgaaggactggaagtaatacgactcagtatagggacaa
          --------1.2------->          <---R1.3---
               <--------------ø1.3---------------
T7.1:     gttgaaggactggaagtaatccaggacccggactggaagtaatacgactcagtatagggacaa
          --------1.2-------><D14R->   <--------------ø1.3---------------
                              <D15L->                    <---R1.3---
```

Part15–Part16: ø*1.3*/R*1.3–1.3*
```
T7:       atttaaccaataggagataaacattatgatgaacatt
          ----R1.3--->         <----1.3----
                <-1.3RBS->
T7.1:     atttaaccaataggaggacccgggccaataggagataaacattatgatgaacatt
          ----R1.3--->       <D16L>  <-1.3RBS->     <----1.3----
                          <D15L->
```

Part16–Part17: *1.3*–TE
```
T7:       agagaaatgtaatcacactggctcaccttcgggtgggcctt
          -----1.3---->   <----------TE------------
T7.1:     agagaaatgtaacccgggcccaaaatgtaatcacactggctcaccttcgggtgggcctt
          -----1.3----><D16R>          <----------TE------------
                          <D17L>
```

Part17–Part18: TE–*1.4*
```
T7:       gcctttctgcgtttataaggagacactttatgtttaagaag
          -TE-->            <1.4RBS >     <----1.4----
T7.1:     gcctttcagggaaaacgggcccatggttataaggagacactttatgtttaagaag
          -TE-->            <D17R>    <1.4RBS >    <----1.4----
                              <D18L>
```

Part18–Part19: *1.4*–ø*1.5*
```
T7:       cgtgtggagtatagttaactggtaatacgactcactaaagg
          -----------1.4---------->
                   <---------ø1.5--------------
T7.1:     cgtgtggagtatagttaactggtaatccatggcgccgttaactggtaatacgactcactaaagg
          -----------1.4----------> <D18R>   <---------ø1.5--------------
```

Part19-Part20: *ø1.5-1.5*

```
T7:      taaaaggaggtacacaccatgatgtactta
         ----ø1.5--->      <----1.5---->
          <-1.5RBS->
T7.1:    taaaggaggtacggcgcctaggcactaaaaggaggtacacaccatgatgtactta
         ----ø1.5--->&lt;D19R&gt;         <-1.5RBS->      <----1.5---->
                         &lt;D20L&gt;
```

U4

```
T7:      gtcattgtaggatgccttgcgctccactgtagcgatgat
         -------1.5-----------------------------
T7.1     gtcattgtaggatgccttgcTctTcactgtagcgatgat
         -------1.5-----------------------------
                        <--U4-->
```

Part20-Part21: *1.5-ø1.6*

```
T7:      tgccagatggtcacgcttaatacgactcact
         --------1.5-------->
               <--------ø1.6----------
T7.1:    tgccagatggtcacgcttaacctaggacgtctggtcacgcttaatacgactcact
         --------1.5-------->&lt;D20L&gt;      <--------ø1.6----------
                                      &lt;D21L&gt;
```

Part21-Part22: *ø1.6-1.6*

```
T7:      taaaaggagacactatatgtttcgactt
         ----ø1.6--->   <----1.6---->
         <-1.6RBS->
T7.1:    taaaggagacacgacgtctagacactaaaggagacactatatgtttcgactt
         ----ø1.6--->&lt;D21R&gt;      <-1.6RBS->      <----1.6---->
                           &lt;D22L&gt;
```

Par22-Part23: *1.6-1.7*

```
T7:      tcaaggaggtgttctgatgggactgtta
         -------1.6------>
          <--1.7RBS-->   <----1.7---->
T7.1:    tcaaAgaAgtgttctgatctagaaccggttcaaggaggtgttctgatgggactgtta
         -------1.6------>&lt;D22R&gt;&lt;D23L&gt; <--1.7RBS-->   <----1.7---->
```

Changes in Section *Beta*

Part23-Part24: *1.7-1.8*

```
T7:      gaactctttgagaaacataaggataaatgttatgcataacttcaagtca
         -----------------1.7------------------>
                       <-1.8RBS->      <------1.8--------
T7.1:    gaactctttgagaaacataaAgataaatgttaCgcataaccggtgacctctttgagaaacataaggataaatgttatgcataacttcaagtca
         -----------------1.7------------------> &lt;D24L-&gt;          <-1.8RBS->      <-----1.8---------
                                              &lt;D23R&gt;
```

Part24-Part25: *1.8-2*

```
T7:      gaactttggaaatcgagaggtcaatgactatgtcaaacgta
         ------------1.8-----------> <-----2----->
                        <--RBS-->
T7.1:    gaactttggaaatcgagaggtcaatgaggtcaccgtacgttggaaatcgagaggtcaatgactatgtcaaacgta
         ------------1.8----------->&lt;D24R&gt;          <--2RBS-->   <-----2----->
                                 &lt;D25L&gt;
```

Part25-Part26: *2-ø2.5*

```
T7:      ttgtgtagcaccgaagtaatacgactcactat
         ---------2-------->
               <-----------ø2.5----------
T7.1:    ttgtgtagcaccgaagtaacgtacgaattcagcaccgaagtaatacgactcactat
         ---------2-------->&lt;D25R&gt;      <------------ø2.5---------
                             &lt;D26L&gt;
```

Part26-Part27: *ø2.5-2.5*

```
T7:      cactattagggagactccctctgagaaaccaaacgaaacctaaaggagattaacattatggctaagaag
         -----ø2.5----->                          <-2.5RBS->      <----2.5---->
T7.1:    cactattagggagagaattcccgggcgaaacctaaaggagattaacattatggctaagaag
         -----ø2.5----->&lt;D26R&gt;         <-2.5RBS->      <----2.5---->
                          &lt;D27L&gt;
```

Part27-Part28: *2.5-2.8*

```
T7:      gcagacgaagacggagacttctaagtggaactgcgg
         ---------2.5----------->&lt;----2.8---->
                <-2.8RBS->
T7.1:    gcagacgaagacggGgacttctaacccgggatccgaagacggagacttctaagtggaactgcgg
         ---------2.5----------->&lt;D27R&gt;      <-2.8RBS->    <----2.8---->
                                      &lt;D28L&gt;
```

Part28-Part29: *2.8-3*

```
T7:      acgcaaaggggaggcgacatggcaggttacggcgctaaaggaatccgaaa
         ----------------2.8----------------->
             <--3RBS-->   <----------------3--------------
T7.1:    acgcaaGgggagAcgacaCggcaggttacggcgctaaggatccggccgcaaaggggaggcgacatggcaggttacggcgctaaaggaatccgaaa
         ----------------2.8----------------->&lt;D28R&gt;      <--3RBS-->   <---------------3---------------
```

Part29–Part30: *3–3.5*

```
T7:       gattaaaaaggaaaggaggaaagaaataatggctcgtgta
          ------------3-------------->
                   <-3.5RBS->           <----3.5----
T7.1:     gattaaaaCgCaaGggGggGaagaaataacggccgccgcggaaaggaaaggaggaaagaaataatggctcgtgta
          ------------3--------------><D29R><D30L>     <-3.5RBS->          <----3.5----
```

Part30–Part31: *3.5–ø3.8*/R*3.8*

```
T7:       tctgaccgtggataattaattgaactcactaaag
          ------3.5----->
                   <------------ø3.8-----------
T7.1:     tctgaccgtggataaccgcggacatgtcgtggataattaattgaactcactaaag
          ------3.5-----><D30R><D31R><-----------ø3.8-----------
```

Part31–Part32: *ø3.8*/R*3.8–3.8*

```
T7:       tttccctttgttcgcattggaggtcaaataatgcgcaagtct
          -------R3.8------>           <----3.8----
                         <-3.8RBS->
T7.1:     tttccctttgttcgcattggaggtcaaataatacatgtcgacgaggtcaaataatgcgcaagtct
          -------R3.8------>                <D31R> <-3.8RBS-> <----3.8----
                                                      <D32L>
```

Changes in Section *Gamma*

Part32–Part33: *3.8–4A/4B/4.1/4.2*

```
T7:       tagaactaggagggaattgcatggacaattcgcacgattccgatagtgt
          -------------------3.8---------------------->
                  <-4ARBS->    <------------4A-------------
T7.1:     tagaactagAagAgaattgcaCggacaattcgcacgattccgataggtcgacgtacgctaggagggaattgcatggacaattcgcacgattccgatagtgt
          -------------------3.8----------------------><D32R>     <-4ARBS->    <------------4A--------------
                                                       <D33L>
```

Part33–Part34: *4A/4B/4.1/4.2–ø4.3*

```
T7:       ggagagtcccattctaatacgactcactaaa
          -------4.2------>
               <------------ø4.3----------
T7.1:     ggagagtcccattctaacgtacggccgagtcccattctaatacgactcactaaa
          -------4.2------><D33R>    <------------ø4.3----------
                         <D34L>
```

Part34–Part35: *ø4.3–4.3*

```
T7:       ctaaaggagacacaccatgttcaaactg
          ----ø4.3----> <-----4.3----
           <-4.3RBS->
T7.1:     ctaaaggagacacaccggccggtggcgcgcctaaaggagacacaccatgttcaaactg
          ----ø4.3----> <D34R>   <D35L> <-4.3RBS->    <-----4.3----
```

Part35–Part36: *4.3–4.5*

```
T7:       ttctttgagtaatcaaacaggagaaaccattatgtctaacgta
          ----4.3----> <-4.5RBS->   <----4.5----
T7.1:     ttctttgagtaatggcgcgccaccggcgaaacaggagaaaccattatgtctaacgta
          ----4.3----> <D35R><D36L> <-4.5RBS->    <----4.5----
```

Part36–Part37: *4.5–R4.7/ø4.7*

```
T7:       attgataactaagagtggtatcct
          ----4.5----><----R4.7---
T7.1:     attgataactaacaccggcgaagcttaagagtggtatcct
          ----4.5----><-D36R><D37L> <----R4.7---
```

Part37–Part38: R*4.7*/ø*4.7–4.7*

```
T7:       ctataggagatattaccatgcgtgaccct
          ----ø4.7----> <----4.7----
           <-4.7RBS->
T7.1:     ctataggagatattaccaagcttcctggactataggagatattaccatgcgtgaccct
          ----ø4.7----> <D37R>      <-4.7RBS->   <----4.7----
                         <D38L->
```

Part38–Part39: *4.7–5*

```
T7:       aagtcacgataatcaataggagaaatcaatatgatcgtttct
          ----4.7----> <-5RBS->   <-----5-----
T7.1:     aagtcacgataatcctggagctagcatcaataggagaaatcaatatgatcgtttct
          ----4.7----><D38R><D39L>     <-5RBS->        <-----5-----
```

Part39–Part40: *5–5.3*

```
T7:       atttgccactgatacaggaggctactcatgaacgaaaga
          -----5-----> <-5.3RBS->    <----5.3----
T7.1:     atttgccactgagctagcatgctacaggaggctactcatgaacgaaaga
          -----5-----><D39R>    <-5.3RBS->   <----5.3----
                         <D40L>
```

Part40–Part41: *5.3–5.5*

```
T7:       ataaaactataggagaaattattatggctatgaca
          ----5.3---->           <----5.5----
              <-5.5RBS->
T7.1:     ataaaactatagcatgccatggtataggagaaattattatggctatgaca
          ----5.3----> <D41L> <-5.5RBS->    <----5.5----
                    <D40R>
```

Part41-Part42: *5.5-5.7*

```
T7:     acgggaggtgttctgatgtctgactac
        ------5.5------>
         <-5.7RBS->    <----5.7----
T7.1:   acgCgaggtgttctgaccatggatccgggaggtgttctgatgtctgactac
        ------5.5------><41R>    <-5.7RBS->    <----5.7----
                           <42L>
```

Part42-Part43: *5.7-5.9*

```
T7:     tgggaggatgtgtctaatgtctcgtgac
        -------5.7------>
        <-5.9RBS->    <----5.9----
T7.1:   tgggCggGtgtgtctaaggatccgcggcgaatgggaggatgtgtctatgtctcgtgac
        -------5.7------><D42R>    <-5.9RBS->    <----5.9----
                                <D43L>
```

Part43-Part44: *5.9-6*

```
T7:     ctagaggagaaacttaatggcacttcttgacc
        ---------------5.9----------->
        <-6RBS->    <--------6------
T7.1:   ctagaAgaAaaacttaaCggcacttcttgaccgcgggcccggaactagaggagaaacttaatggcacttcttgacc
        ---------------5.9-----------><D43R>    <-6RBS->    <--------6------
                                          <D44L>
```

Part44-Part45: *6-6.3*

```
T7:     gacaaggagatttacctgtggagaccgtagcgt
        --------------6-------------->
          <-6.3RBS->    <------6.3------
T7.1:   gacaaggaAatttacctCtggagaccgtaggcccgggcaaggagatttacctgtggagaccgtagcgt
        --------------6--------------> <D45L><-6.3-RBS->    <------6.3------
                                          <D44R>
```

Part45-Part46: *6.3-R6.5/ø6.5*

```
T7:     gacactaagtgataaact
        ----6.3---->
             <---R6.5----
T7.1:   gacactaaAtAacccggggagctcactaagtgataaact
        ----6.3----><D45R>    <---R6.5----
                         <D46L>
```

Part46-Part47: *R6.5/ø6.5-6.5*

```
T7:     cgattattactttaagatttaactctaagaggaatctttattatgttaacacct
        ----R6.5---->        <-6.5RBS->    <----6.5----
T7.1:   cgattattactttaagatttaagagctcgagtaagaggaatctttattatgttaacacct
        ----R6.5---->        <D46R>    <-6.5RBS->    <----6.5----
                                 <D47L>
```

Part47-Part48: *6.5-6.7*

```
T7:     tgatggggaggattgacactatgtgtttctca
        ------6.5------>    <----6.7----
         <-6.7RBS->
T7.1:   tgatggCgaAgattgactcgagaattctgatggggaggattgacactatgtgtttctca
        ------6.5------><D47R>    <-6.7RBS->    <----6.7----
                            <D48L>
```

Part48-Part49: *6.7-7*

```
T7:     tttggaggtaagaagtgatgtctgagttc
        ------6.7-------->
         <-7RBS->    <-----7-----
T7.1:   tttggaggtaagaagtgagaattcgatcgcatttggaggtaagaagtgatgtctgagttc
        ------6.7--------><D48R>    <-7RBS->    <-----7-----
                             <D49L>
```

Part49-Part50: *7-7.3*

```
T7:     tttaaggaggtataagttatgggtaagaaa
        -------7------>   <----7.3----
         <-7.3RBS->
T7.1:   tttaaggaggtataacgatcggtccgctttaaggaggtataagttatgggtaagaaa
        -------7------><D49L>    <-7.3RBS->    <----7.3----
                           <D50R->
```

Part50-Part51: *7.3-7.7*

```
T7:     atcaacatttaatcaggaggttatcgtggaagactgc
        ----7.3----> <-7.7RBS->   <----7.7----
T7.1:   atcaacatttaacggtccgctgcagtcaggaggttatcgtggaagactgc
        ----7.3----><D50R-><D51L> <-7.7RBS->   <----7.7----
```

Changes in Section *Delta*

Part51-Part52: *7.7-8*

```
T7:     gacatggagacacatttaatggctgagaaa
        -------7.7-------->
         <-8RBS->       <-----8-----
T7.1:   gacatggagacacatttaactgcagcgtacgagacatggagacacatttaatggctgagaaa
        -------7.7--------><D51R><D52L>    <-8RBS->    <-----8-----
```

Part52-Part53: *8-ø9*

```
T7:      cagccgggaatttaatacgactcactatag
         -------8------>
          <------------ø9-------------
T7.1:    cagccgggaatttaacgtacgatcgccgggaatttaatacgactcactatag
         -------8------><D52L>  <------------ø9-------------
                         <D53L>
```

Part53–Part54: *ø9–9*
```
T7:      tagggagacctcatctttgaaatgagcgatgacaagaggttggagtcctcggtcttcctgtagttcaactttaaggagacaataataatggctgaat
         ----ø9--->                                                       <-9RBS->            <----9----
T7.1:    tagggagacctcatctttgaaatgagcgatcgacaagaggttggagtcctcggtcttcctgtagaattcaactttaaggagacaataataatggctgaat
         ----ø9--->                  <D53R>                              <D54L>   <-9RBS->            <----9----
```

Part54–Part55: *9–ø10*
```
T7:      tcgaacttctgatagacttcgaaatta
         -----9----->  <----ø10----
T7.1:    tcgaacttctgatagaattccgcggacttcgaaatta
         -----9-----> <D54R>    <----ø10----
                        <D55L>
```

Part55–Part56: *ø10–10A/B*
```
T7:      tatagggagaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatatacatatggctagcatg
         ----ø10--->                                        <-10RBS->         <----10-----
T7.1:    tatagggagaccacaaccgcggatccctctagaaataattttgtttaactttaagaaggagatatacatatggctagcatg
         ----ø10---> <D55R>                                 <-10RBS->         <----10-----
                        <D56L>
```

Part56–Part57: *10A/B–Tø*
```
T7:      gctgagcaataactagcataacccccttggggcct
         ----10----->   <-------Tø-------
T7.1:    gctgagcaataaggatcccgggctagcataacccccttggggcct
         ----10-----><D56R>       <-------Tø-------
                        <D57L>
```

Part57–Part58: *Tø–11*
```
T7:      gttttttgctgaaaaggaggaactatatgcgctcata
         --Tø-->  <-11RBS->     <----11----
T7.1:    gttttttgctgacccgggccctgaaaaggaggaactatatgcgctcata
         --Tø-->     <D57R>  <-11RBS->     <----11----
                         <D58L>
```

Part58–Part59: *11–12*
```
T7:      tgactcgctaacattaataaataaggaggctctaatggcactcat
         ----11---->         <-12RBS->       <----12----
T7.1     tgactcgctaagggcccaagcttaataaataaggaggctctaatggcactcat
         ----11----><D58R><D59L>    <-12RBS->       <----12----
```

Changes in Section *Epsilon*
Part59–Part60: *12*–R*13*/ø*13*
```
T7:      ccggtatttaataaatattctccctgtgg
         ----12---->   <----R13----
T7.1:    ccggtatttaattaaagcttccggattctccctgtgg
         ----12----> <D59L>    <----R13----
                        <D60R->
```

Part60–Part61: R*13*/ø*13–13*
```
T7:      tatagggagaacaatacgactacgggagggtttcttatgatgactat
         ----ø13--->         <-13RBS->     <----13----
         ----R13--->
T7.1:    tatagggagaacaatactcccggaattctacgggaggggtttcttatgatgactat
         ----ø13--->       <D60R->  <-13RBS->     <----13----
         ----R13--->           <D61L>
```

Part61–Part62: *13–14*
```
T7:      acgaaaggaggataaccatatgtgttgggc
         ------13------>   <----14----
                           <-14RBS->
T7.1:    acgaaaggCggCtaagaattcgtacgcacgaaaggaggataaccatatgtgttgggc
         ------13------><D61R>        <-14RBS->       <----14----
                        <D62L>
```

Part62–Part63: *14–15*
```
T7:      gacgggaggtaatgagctatgagtaaaat
         -----14----->   <----15----
          <-15RBS->
T7.1:    gacCggCaggtaacgtacgatcgccaagacgggaggtaatgagctatgagtaaaat
         -----14-----><D62R>        <-15RBS->       <----15----
                        <D63L>
```

Part63–Part64: *15–16*
```
T7:      gtaaggagtaactaaaggctacataaggaggccctaaatggataagta
         ----15---->          <-16RBS->     <----16----
T7.1:    gtaaggagtaacgatcggccgtaaaggctacataaggaggccctaaatggataagta
         ----15----><D63R>         <-16RBS->     <----16----
                        <D64L>
```

Changes in Section *Zêta*

Part64-Part65: *16-ø17*

```
T7:      gggagcgtaggaaataatacgactcactatag
         -------16------->
            <--------------ø17----------
T7.1:    gggagcgtaggaaataacggccgcggcgtaggaaataatacgactcactatag
         -------16-------><D64R>   <--------------ø17----------
                            <D65L>
```

Part65-Part66: *ø17-17*

```
T7:      gggagaggcgaaataatcttctccctgtagtctcttagatttactttaaggaggtcaaatggctaacgt
         --ø17-->                                    <-17RBS->   <----17----
T7.1:    gggagaggcgaaataatcttctccgcggtgtaagcttcttagatttactttaaggaggtcaaatggctaacgt
         --ø17-->              <D65R>   <D66L>            <-17RBS->   <----17----
```

Part66-Part67: *17-17.5*

```
T7:      aacgagtaattggtaaatcacaaggaaagacgtgtagtccacggatggactctcaaggaggtacaaggtgctatca
         ---17--->                                            <-17.5RBS->   <--17.5--
T7.1:    aacgagtaattggtaaatcacaagcttgaaagacgtgtagtccacggatccggactctcaaggaggtacaaggtgctatca
         ---17--->              <D66R>            <D67L>        <-17.4RBS->   <--17.5--
```

Part67-Part68: *17.5-18*

```
T7:      caataaggagtgatatgtatggaaaagga
         ----17.5---->     <----18----
          <-18RBS->
T7.1:    caataaAgaAtgaggatcccgggaataaggagtgatatgtatggaaaagga
         ----17.5----><D67R>    <-18RBS->        <----18----
                      <D68L>
```

Part68-Part69: *18-R18.5*

```
T7:      cattacagtgatatactcaa
         ----18---->
            <-----R18.5-----
T7.1:    cattacaatagcccgggcccacattacagtgatatactcaa
         ----18----><D68R>        <-----R18.5-----
                      <D69L>
```

Part69-Part70: *R18.5-E/18.5/18.7*

```
T7:      gtcattgtctatacgagatgctcctacgtgaaatctgaaagttaacgggaggcattatgctagaatt
         -R18.5->                   <------------------E-----------------
                                              <-18.5RBS-> <---18.5---
T7.1:    gtcattgtctatacgagatgctgggccctacgtgaattctgaaagttaacgggaggcattatgctagaatt
         -R18.5->              <D69R>      <D70L>       <-18.5RBS-> <---18.5---
```

Part70-Part71: *E/18.5/18.7-19/19.2/19.3*

```
T7:      aacgtaagtaggaaatcaagtaaggaggcaatgtgtctactca
         ---18.5--->       <-19RBS->     <----19----
T7.1:    aacgtaagtaggaattcgtacgaagtaaggaggcaatgtgtctactca
         ---18.5---><D70R>       <-19RBS->     <----19----
                     <D71L>
```

Part71-Part72: *19/19.2/19.3-øOR*

```
T7:      ggtgatttatgcattaggactgcatagggatgcactatagaccacggatggtcagttctttaagttactgaaaagacacgat
         -19->                                                              <-øOR-
T7.1:    ggtgatttatgcattaggactgcatagggatgcactatagaccacgtacgatggtcagttctttaagttactgcagaaaagacacgat
         -19->                                     <D71R>             <D72L>        <-øOR-
```

Part72-Part73: *øOR-19.5*

```
T7:      gagagga...94nt... gattatattgtattagtatcaccttaacttaaggaccaacataaagggaggagagactcatgttcc
         -øOR->                                            <-19.5RBS->     <-19.5-
T7.1:    gagagga...94nt... gattatattgtattagtatcaccttaactcgcagtcgaccaacataaagggaggagagactcatgttcc
         -øOR->                                   <D72R>      <-19.5RBS->     <-19.5-
                                                   <D73L>
```

Part73-SRR/TR: *19.5-SRR/TR*

```
T7:      cgattaggtcttcctgaccgactgatggctcaccgagggattcagcggtatgattgcatcacaccacttcatccctata
         -19.5->                                                            <-SRR-
T7.1:    cgattaggtcgacttcctgaccgactgatggctcaccgagggattcagcggtatgattgcatcacaccacttcatccctata
         -19.5-> <D73R>                                                     <-SRR-
```

Table S1: Errors in synthesis of Section *Alpha*

Location on T7.1 (Genome Position)	Nature of Difference	Probable Reason for Difference	Expected Outcome
D1L (164-170), D1R & D2L (338-350)	Restriction sites were not added in construction	Difficulties in manipulating left end of genome resulted in using wild-type	Loss of manipulability in part 1 (containing A0)
gene *0.4* (1418)	Single base deletion	Unknown	Frameshift after 27th amino acid followed by early termination of gene *0.4*
D6L (1304-1310) D6R (1494-1500)	Restriction sites appear twice.	Inefficiency of digestion of scaffold	No expected change
gene *0.6B*	Single base addition	Error is known to be in stock of wild-type genome	Dependent upon nature of putative translational slippage in formation of gene *0.6B*
D11L (3302-3307)	Restriction site appears twice	Inefficiency of digestion of scaffold	No expected change
gene *1* (4877)	Single base mutation	Error in PCR or within wild-type genome	Silent mutation, no expected change
gene *1* (5159)	Single base mutation	Error in PCR or within wild-type genome	Silent mutation, no expected change
gene 1 (5399)	Single base mutation	Error in PCR or within wild-type genome	Silent mutation, no expected change
D14R (6591-6597)	Restriction site appears twice	Inefficiency of digestion of scaffold	No expected change
TE (7827)	Single base deletion	Primer synthesis error	Possible loss of function of transcriptional terminator.
D20L (8082-8086)	Restriction site appears twice	Inefficiency of digestion of scaffold	No expected change
U4 (8153-8159)	Restriction site was not added in construction	Failure in site-directed mutagenesis	Loss of manipulability of overlap in parts 18 and 19
D22L (8247-8253)	Restriction site appears twice	Inefficiency of digestion of scaffold	No expected change

Table S2: Errors in synthesis of Section *Beta*

Location on T7.1	Nature of Difference	Probable Reason for Difference	Expected Outcome
gene *1.7* (**8794**)	Single base silent mutation	Error during PCR or within wild-type genome	No expected change
gene *1.8* (**9245**)	Single base mutation	Error during PCR or within wild-type genome	Amino acid change in gene *1.8* from Asp to Gly
gene *2.0* (**9447**)	Single base mutation	Error during PCR or within wild-type genome	Amino acid change in gene *2.0* from Glu to Val
gene *2.5* (**10351**)	Singe base deletion	Error during primer design	deletion in stop codon; read-through adding on 8AA
gene *2.8* (**10627**)	Single base mutation	Error during PCR	Amino acid change in gene *2.8* from Asp to Gly
gene *2.8* (**10717-10803**)	82 base deletion	Error in cloning of part	Loss of function in gene *2.8* in addition to unknown effect on translation of *3.0* due to read-through
gene *3.0* (**10926**)	Single base silent mutation	Error during PCR or within wild-type genome	No expected change

Supporting References and Notes:

S1. J. J. Dunn, F. W. Studier, *J. Mol. Biol.* **166**, 477 (1983).

S2. F. W. Studier, J. J. Dunn, *Cold Spring Symp. Quant. Biol.* **47**, 499 (Pt. 2, 1983).

S3. M. Demerec, U. Fano, *Genetics* **30**, 119 (1945).

S4. M. Delbrück, *Biol. Rev. Cambridge Phil. Soc.* **21**, 30 (1946).

S5. F. W. Studier, *Virology* **95**, 70 (1979).

S6. The T7 stocks used in this study come from stocks maintained by the following labs in reverse chronological order: I. J. Molineux, F. W. Studier, R. L. Sinsheimer, M. Delbrück.

S7. F. W. Studier, *Virology* **39**, 562 (1969).

S8. R. Hausmann, B. Gomez, *J. Virol.* **1**, 779 (1967).

S9. R. Hausmann, K. LaRue, *J. Virol.* **3**, 278 (1969).

S10. Y. Masamune, G. D. Frenkel, C. C. Richardson, *J. Biol. Chem.* **246**, 6874 (1971).

S11. D. A. Richie, F. E. Malcolm, *J. Gen. Virol.* **9**, 35 (1970).

S12. F. W. Studier, *J. Mol. Biol.* **79**, 227 (1973).

S13. F. W. Studier, J. V. Maizel, *Virology* **39**, 575 (1969).

S14. F. W. Studier, *J. Mol. Biol.* **29**, 237 (1973).

S15. F. W. Studier, *J. Mol. Biol.* **153**, 493 (1981).

S16. R. W. Davis, R. W. Hyman, *Cold Spr. Harb. Symp. Quant. Biol.* **35**, 1 (1970).

S17. E. G. Minkley, D. Pribnow, *J. Mol. Biol.* **77**, 255 (1973).

S18. M. Golumb, M. Chamberlin, *Proc. Natl. Acad. Sci., U.S.A.* **71**, 760 (1974).

S19. E. G. Niles, R. C. Condit, *J. Mol. Biol.* **98**, 57 (1975).

S20. W. T. McAllister, R. J. McCarron, *Virology* **82**, 288 (1977).

S21. S. J. Stahl, M. J. Chamberlin, *J. Mol. Biol.* **112**, 577 (1977).

S22. G. A. Kassavetis, M. J. Chamberlin, *J. Virol.* **29**, 196 (1979).

S23. N. Panayotatos, R. D. Wells, *Nature* **280**, 35 (1979).

S24. J. L. Oakley, J. E. Coleman, *Proc. Natl. Acad. Sci., U.S.A.* **74**, 4266 (1977).

S25. J. C. Boothroyd, R. S. Hayward, *Nucleic Acids Res.* **7**, 1931 (1979).

S26. M. D. Rosa, *Cell* **16**, 815 (1979).

S27. M. D. Rosa, *J. Mol. Biol.* **147**, 55 (1981).

S28. M. D. Rosa, *J. Mol. Biol.* **147**, 199 (1981).

S29. H. L. Osterman, J. E. Coleman, *J. Mol. Biol.* **77**, 255 (1981).

S30. A. D. Carter, W. T. McAllister, *J. Mol. Biol.* **153**, 825 (1981).

S31. J. J. Dunn, F. W. Studier, *J. Mol. Biol.* **148**, 303 (1981).

S32. W. C. Summers, I. Brunovskis, R. W. Hyman, *J. Mol. Biol.* **74**, 291 (1973).

S33. W. T. McAllister, H. L. Wu, *Proc. Natl. Acad. Sci., U.S.A.* **75**, 804 (1978).

S34. W. T. McAllister, C. Morris, A. H. Rosenberg, F. W. Studier, *J. Mol. Biol.* **153**, 527 (1981).

S35. I. J. Molineux, personal communication

S36. E. R. Olson, E. L. Flamm, D. I. Friedman, *Cell* **31**, 61 (1982)

S37. F. W. Studier, A. H. Rosenburg, *J. Mol. Biol.* **153**, 503 (1981).

S38. F. W. Studier, *Science* **176**, 367 (1972).

S39. J. J. Dunn, F. W. Studier, *Proc. Natl. Acad. Sci., U.S.A.* **70** 1559 (1973).

S40. F. W. Studier, A. H. Rosenburg, M. N. Simon, J. J. Dunn, *J. Mol. Biol.* **135**, 917 (1979).

S41. J. J. Dunn, F. W. Studier, *Nucleic Acids Res.* **8**, 2119 (1980).

S42. D. Dressler, J. Wolfson, M. Magazin, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 998 (1972).

S43. J. Wolfson, D. Dressler, M. Magazin, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 499 (1972).

S44. F. Tamanoi, H. Saito, C. C. Richardson, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2656 (1980).

S45. H. Saito, S. Tabor, F. Tamanoi, C. C. Richardson, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3917 (1980).

S46. X. Zhang, F. W. Studier, *J. Mol. Biol.* **340**, 707 (2004).

S47. D. A. Richie, C. A. Thomas Jr. , L. A. MacHattie, P. C. Wensinnk, *J. Mol. Biol.* **135**, 907 (1967).

S48. T. J. Kelly, C. A. Thomas Jr., *J. Mol. Biol.* **44**, 459 (1969).

S49. T. Ikemura, *J. Mol. Biol.* **151**, 389 (1981).

S50. T. Knight, "Idempotent Vector Design for Standard Assembly of Biobricks" (MIT Synthetic Biology Working Group Technical Report 0, 2002; http://web.mit.edu/synbio/release/docs/biobricks.pdf).

S51. T. Vincze, J. Posfai, R. J. Roberts, *Nucleic Acids Res.* **31**, 3688 (2003).

S52. R. J. Roberts, T. Vincze, J. Posfai, D. Macelis, *Nucl. Acids Res.* **31**, 418 (2003).

S53. Vectors were obtained from the Registry of Standard Biological Parts (http://parts.mit.edu).

S54. L. R. Garcia, thesis, The University of Texas at Austin (1996).

S55. J. Sambrook, D. W. Russell, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor
     Laboratory Press, New York, ed. 3, 2001).

S56. I. J. Molineux in The Bacteriophages, R. L. Calendar Ed. (Oxford Univ. Press, New York, in press),
     (available at http://www.thebacteriophages.org/frames_0200.htm)