ACS Synthetic Biology
Research Article
pubs.acs.org/synthbio

Assembly of Large, High G+C Bacterial DNA Fragments in Yeast

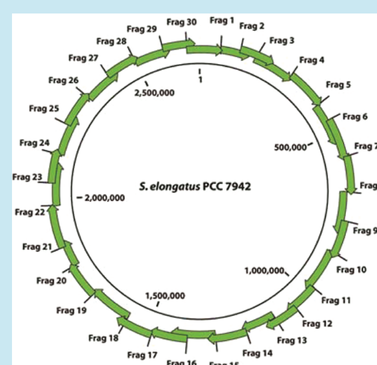
Vladimir N. Noskov,[†] Bogumil J. Karas,[‡] Lei Young,[†] Ray-Yuan Chuang,[†] Daniel G. Gibson,[‡] Ying-Chi Lin,[‡] Jason Stam,[‡] Isaac T. Yonemoto,[‡] Yo Suzuki,[‡] Cynthia Andrews-Pfannkoch,[†] John I. Glass,[†] Hamilton O. Smith,[‡] Clyde A. Hutchison, III,[‡] J. Craig Venter,[‡] and Philip D. Weyman*,[‡]

[†]Department of Synthetic Biology and Bioenergy, J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, Maryland 20850, United States

[‡]Department of Synthetic Biology and Bioenergy, J. Craig Venter Institute, 10355 Science Center Drive, San Diego, California 92121, United States

Supporting Information

ABSTRACT: The ability to assemble large pieces of prokaryotic DNA by yeast recombination has great application in synthetic biology, but cloning large pieces of high G+C prokaryotic DNA in yeast can be challenging. Additional considerations in cloning large pieces of high G+C DNA in yeast may be related to toxic genes, to the size of the DNA, or to the absence of yeast origins of replication within the sequence. As an example of our ability to clone high G+C DNA in yeast, we chose to work with *Synechococcus elongatus* PCC 7942, which has an average G+C content of 55%. We determined that no regions of the chromosome are toxic to yeast and that *S. elongatus* DNA fragments over ~200 kb are not stably maintained. DNA constructs with a total size under 200 kb could be readily assembled, even with 62 kb of overlapping sequence between pieces. Addition of yeast origins of replication throughout allowed us to increase the total size of DNA that could be assembled to at least 454 kb. Thus, cloning strategies utilizing yeast recombination with large, high G+C prokaryotic sequences should include yeast origins of replication as a part of the design process.



KEYWORDS: *yeast, DNA assembly, transformation-associated recombination, Synechococcus, cyanobacteria*

Synthetic biology requires efficient methods for assembly of various DNA parts.[1,2] As the utility of DNA cloning has progressed from single genes to entire biochemical pathways,[3−5] the need for efficient strategies to clone and assemble large pieces of DNA has emerged. Larger fragments of DNA are often cloned into bacterial artificial chromosomes (BACs), fosmids or other phage-derived vectors, or yeast artificial chromosomes (YACs). Baker's yeast (*Saccharomyces cerevisiae*) is a good choice for both cloning and maintaining large DNA pieces because of its efficient recombination-based assembly methods and its ability to stably replicate inserts over 1 Mb.[6,7] YACs were first developed as linear vectors by ligating centromeres and telomeric ends to DNA fragments. More recently, circular YACs have been exploited to maintain molecules as large as the 1.1-Mb *Mycoplasma mycoides* chromosome,[8] the 1.5-Mb *Acholeplasma laidlawii* chromosome,[9] and the 1.6-Mb *Prochlorococcus marinus* chromosome.[10] Circular YACs can be assembled by yeast homologous recombination from over 25 fragments in a single transformation step.[11]

The bacterial chromosomes discussed above that have been cloned as YACs share the common feature of low G+C content. Bacterial chromosomes cloned in yeast to date have a G+C content between 32% and 38%.[8−10] High G+C content may limit the size of insert that can be cloned as a YAC. For example, attempts to clone *Pseudomonas aeruginosa* chromosomal DNA (66% G+C) was limited to fragments smaller than 120 kb,[12] while a *Myxococcus xanthus* (69% G+C) linear YAC library was constructed with insert sizes ranging between 40 and 180 kb with an average insert size of 111 kb.[13] YAC libraries of *Bacillus subtilis*, with a lower G+C content of 44%, contained slightly larger inserts of up to 250 kb.[14]

While YACs of high G+C prokaryotic DNA may be limited in size, YAC clones of eukaryotic DNA can contain inserts greater than 1 Mb.[15] Eukaryotic DNA cloned in yeast may escape this size limitation because it contains origins of replication every 20−60 kb.[16−19] It has been predicted that a strong origin of replication, also called an autonomously replicating sequences (ARS) in yeast, may be able to support 120−300 kb of sequence.[20] By containing ARS sequences that may be recognized by yeast, large eukaryotic DNA fragments can be cloned in yeast using vectors lacking ARS fragments.[21] ARS sequences are also known to be rich in A+T sequence,[22] and prokaryotic sequences with low G+C content may have a statistically greater chance of containing a functional ARS sequence needed for replication in yeast.

Given its utility as a model organism for the study of photosynthetic prokaryotes, we chose to work with *Synechococcus elongatus* PCC 7942 to demonstrate the construction of

Received: March 23, 2012
Published: May 8, 2012

 © 2012 American Chemical Society dx.doi.org/10.1021/sb3000194 | ACS Synth. Biol. 2012, 1, 267−273

high G+C prokaryotic assembly procedures in yeast. The genome of *S. elongatus* has been fully sequenced and consists of a single 2.7 Mb chromosome composed of 55% G+C.[23] *S. elongatus* DNA has been cloned as a fosmid library in *Escherichia coli* in fragments of approximately 30–40 kb.[24] Given that the G+C content is significantly higher than that of yeast (38%[25]), this organism is a good example of our ability to clone large pieces (>100 kb) of high G+C DNA. In this paper, we demonstrate methods to assemble pieces of high G+C prokaryotic DNA up to 454 kb and show that the addition of yeast origins of replication optimizes the stability of these assemblies in yeast.

■ RESULTS AND DISCUSSION

Our initial inquiry into cloning *S. elongatus* DNA in yeast was to attempt to clone DNA fragments of different lengths using the transformation-associated recombination (TAR) technique in yeast.[26] Because TAR cloning is more efficient when the sites of recombination are at the ends of DNA fragments, we attempted to clone restriction fragments with lengths between 70 and 662 kb. The maximum size we were successful in cloning was 142 kb (Supplementary Figure 1, Supplementary Table 1). Factors that could have influenced the ability to clone *S. elongatus* DNA were the presence of a "toxic gene" that may have been expressed in yeast, the size of the DNA fragment, and the presence of sequences within the fragment that may have acted as origins of replication.

The presence of prokaryotic sequences that are unable to be cloned in yeast has been encountered previously.[9] To test whether any part of the *S. elongatus* chromosome is unable to be cloned in yeast, we developed a strategy to clone the entire *S. elongatus* chromosome as overlapping, ~100-kb pieces using TAR cloning (Figure 1). To take advantage of the increased efficiency of TAR cloning using linearized ends of target sequence, we designed overlapping pieces of approximately 100 kb using restriction enzymes that cut relatively rarely in the *S. elongatus* chromosome so that at least one side of the fragment to be cloned was freed by restriction enzyme digest. The ~100-kb pieces were between 87 and 134 kb with an average size of 112 kb. The overlaps between adjacent pieces ranged between 4 and 75 kb to ensure complete coverage. All 30 pieces were successfully cloned in yeast and were confirmed as complete clones by PCR across the designed vector junctions and by gel electrophoresis of the cloned pieces as supercoiled plasmids (Figure 1B). Of these 30 pieces, 29 could also be cloned in *E. coli*. Clones of each of the 29 fragments were confirmed by PCR and gel electrophoresis as for the yeast clones (data not shown).

After determining that the entire chromosome could be cloned in 30 pieces and that there were no regions that were toxic to yeast, we proceeded to test how fragment size affected the ability to clone high G+C DNA in yeast. Since the maximum size in our pilot experiment was 142 kb, we tested whether two of the ~100 kb pieces could be assembled using yeast recombination. We chose the pairs 19 and 20 and 22 and 23 which overlapped by 9 and 62 kb, respectively (Figure 2). Fragments 19 and 20 were assembled to make fragment 31 (180-kb insert), and fragments 22 and 23 were assembled to make fragment 32 (167-kb insert). In order to make fragments 31 and 32, the ~100-kb plasmids were purified, and the *S. elongatus* sequence was digested out of the vector using the designed restriction enzyme release sites (Supplementary Table 2). The linearized fragments were then transformed into yeast

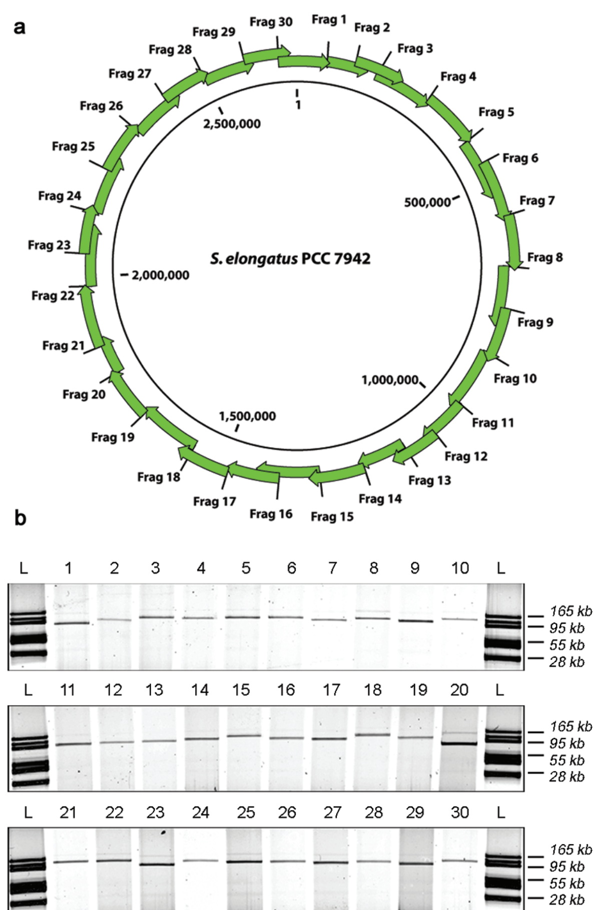


Figure 1. Cloning of the *S. elongatus* chromosome as 30 separate ~100-kb fragments. (A) Diagram of the fragments cloned in yeast. Numbers in the inner circle indicate the position in bp. (B) Each of the 30 plasmids was purified from yeast in supercoiled form and separated by agarose gel electrophoresis next to the BAC-Tracker supercoiled DNA ladder (Epicenter).

with a vector containing homology to sequence at the two ends of final assembled fragments 31 or 32. Resulting yeast colonies were screened by PCR across the junctions of the insert and vector, and the size of the plasmid contained in PCR-positive colonies was determined by gel electrophoresis (Figure 2C). Plasmids of the correct size and with the correct junctions were found in 6/12 and 2/12 colonies for the fragments 31 and 32, respectively. Restriction-enzyme digestion also confirmed that assemblies 31 and 32 contained the correct DNA (Supplemental Figure 2). Fragments 31 and 32 could also be transformed and stably maintained in *E. coli* (Supplemental Figure 2).

To demonstrate that *S. elongatus* fragments larger than ~200 kb could be maintained in yeast, we attempted to assemble the equivalent of five of the ~100-kb *S. elongatus* DNA fragments into a single 454-kb plasmid spanning fragments 19–23. This was performed as an assembly of fragments 31 (19 + 20 combined), 21, and 32 (22 + 23 combined) (Figure 3A and B). The *S. elongatus* sequence was digested out of the vectors using the designed vector release sites and purified, and the linearized pieces were separated by agarose gel electrophoresis to confirm the correct size (Figure 3C). Purified fragments 31, 21, and 32 were transformed into yeast with a vector containing homology to the 5′ end of fragment 31 and the 3′ end of fragment 32. When the resulting colonies were screened by PCR across the

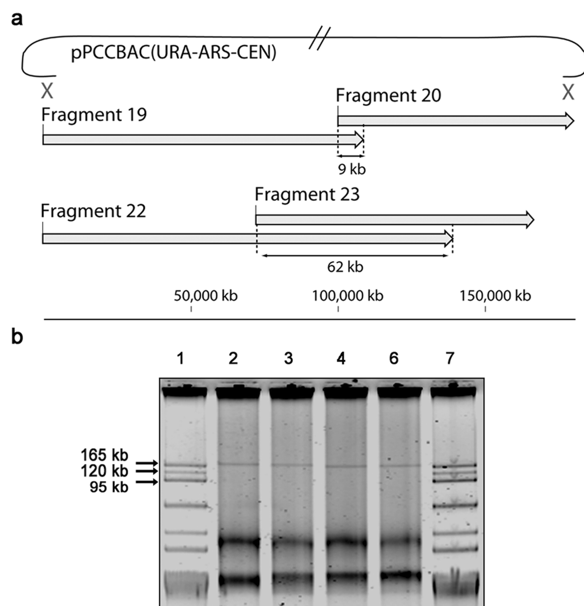


Figure 2. Assembly of *S. elongatus* fragments 19 + 20 and 22 + 23 in yeast. (A) Fragment 31 was assembled from fragments 19 and 20 with a 9-kb overlap. Fragment 32 was assembled from fragments 22 and 23 with a 62-kb overlap. All objects except the vector are drawn to scale. (B) Plasmids from two of the positive clones from each of assemblies 31 and 32 were purified in supercoiled form and separated by agarose gel electrophoresis next to the BAC-Tracker supercoiled DNA ladder (Epicenter).

junctions of the DNA pieces to be assembled, no clones were positive for all of the junctions. We reasoned that since we were unable to clone fragments over ~150 kb in our initial TAR cloning attempts (Figure 1), assembly of the 454-kb plasmid may not be stably maintained in yeast because the distance between origins of replication was too great and plasmid replication was delayed.

To circumvent the lack of yeast ARS sequence in our 454-kb assembly, we inserted a yeast origin of replication in each of fragments 21, 31, and 32 and attempted the assembly in yeast (Figure 3B). Positive yeast colonies from the 454-kb/3ARS assembly were identified by PCR screening of the expected junctions of assembly. When gel electrophoresis was performed on the plasmids purified from positive yeast strains, a supercoiled plasmid of the expected size of the 454-kb/3ARS plasmid was observed in addition to a second, smaller 250-kb plasmid (Figure 4A). When DNA extracted from a yeast strain containing both the 454-kb and 250-kb plasmids was used to transform *E. coli* by electroporation, the resulting *E. coli* colonies were found to have either a 454-kb or a 250-kb plasmid, and no *E. coli* colonies contained both plasmids (Figure 4B). The 454-kb/3ARS plasmid was stably maintained in *E. coli*. When one of the 454-kb/3ARS plasmids purified from *E. coli* cells was transformed back into yeast, the resulting yeast colonies again were found to have plasmids of both 250 kb and 454 kb, suggesting that a smaller, recombined version of the plasmid may have been generated by the yeast (Figure 4C).

Multiplex PCR performed on the *E. coli* clones containing either the 454-kb/3ARS or the 250-kb plasmids indicated that the fragment missing was found in the 208-kb region between two of the inserted ARS sequences (Figure 4D). Further analysis by PCR indicated that recombination had occurred between ARS2 and ARS3 leading to a truncated insertion (data not shown). We reasoned that the 208-kb region that was recombined out by yeast may have been too long to be replicated in a timely manner and that there was strong

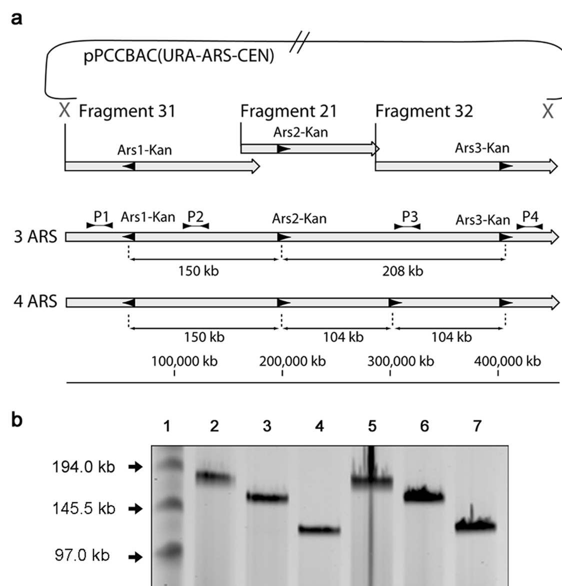


Figure 3. Assembly of *S. elongatus* fragments to make 454-kb insert in yeast. A. Diagram of the assembly of fragments 31, 21, and 32. The positions and orientations of the ARS sequences is indicated by arrows. Also indicated are the locations of screening primers (P1−4) and distances between ARS insertions. All objects except the vector are drawn to scale. (B) Sizes of fragments before assembly. After restriction enzyme digest to remove vector sequence, fragments were separated by agarose gel electrophoresis.

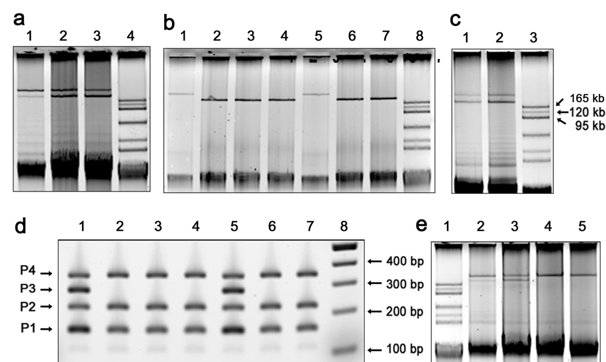


Figure 4. Analysis of 454-kb fragment cloned in yeast and *E. coli.* (A) Yeast colonies containing positive assemblies of 454-kb/3ARS were identified by PCR, and plasmids from three positive yeast colonies were purified and separated by agarose gel electrophoresis (lanes 1−3) with the BAC-Tracker supercoiled DNA ladder (Epicenter) (lane4). (B) Plasmids purified from colony 1 were transformed into *E. coli*, and plasmids from the resulting *E. coli* colonies were purified and separated by agarose gel electrophoresis. (C) Plasmid from *E. coli* colony 5 was transformed back into yeast and screened by PCR. Plasmids from two colonies that were positive for all PCR tests were purified and separated (lanes 1 and 2). (D) Multiplex PCR performed on the same seven colonies from part B showing that the region missing from the smaller, 250-kb plasmids corresponds to the 208-kb region between two ARS fragments. (E) Comparison of plasmid stability before (454-kb/3ARS, lanes 2 and 3) and after (454-kb/4ARS, lanes 4 and 5) insertion of a fourth ARS in the middle of the 208-kb region.

selection for deletions of this piece. To remedy this problem, we inserted a fourth ARS sequence equidistant between these two ARS sequences so that the distance between the two ARS sequences was now 104 kb (Figure 3B). When this new 454-kb/4ARS plasmid was transformed into yeast, it was found to be stably maintained (Figure 4E). Analysis by multiplex PCR indicated that only 1/50 colonies of the 454-kb/4ARS plasmid contained a rearrangement (Table 1). This was comparable to

Table 1. Frequency of Rearrangement of 454-kb/4ARS Plasmid

YAC	colonies tested	deletions/ rearrangements	% incorrect
454-kb/4ARS, clone1	50	1	2
454-kb/4ARS, clone2	50	0	<1
582-kb *M. genitalium* whole genome	50	1	2

the frequency of rearrangement in the maintenance of the 582-kb *Mycoplasma genitalium* genome in yeast (Table 1). To confirm that the junctions had assembled seamlessly, we sequenced the entire junction between fragments 19–20 (8.9 kb) and 21–22 (4.1 kb). The assembled sequence matched the original *S. elongatus* genome sequence at every nucleotide position (data not shown).

Given this result, design of DNA assemblies involving large, high G+C prokaryotic DNA optimally includes origins of replication approximately every 100 kb. The 4 ARS sequences inserted into the fragments were identical and contained adjacent sequence for a bacterial antibiotic cassette. Since the structure of a yeast ARS consists of several sequence components, the exact requirements for a given sequence to function as an ARS in yeast are still poorly understood. The best-studied component of the ARS consensus sequence (ACS) is the binding site for the origin recognition complex and has the A+T-rich consensus sequence (A/T)TTTAT(A/G)TTT-(A/T).[22] Searching for this consensus sequence in the *S. elongatus* chromosome returns seven hits, while a similar analysis in the *P. marinus* MED4 chromosome yielded 328 ACS sites.[10] There are approximately 12,000 ACS sequences in the yeast genome, but only 337 characterized origins, and thus the presence of an ACS does not necessarily denote origin function.[18] While the ACS is not sufficient for initiation of replication, it is necessary as single point mutations along this sequence can abolish replication or lead to plasmid instability.[27,28] Additional elements making up a functional ARS are located 3′ of the ACS sequence and aid in binding of the origin recognition complex and unwinding of DNA by helicases.[29] The presence of a functional ARS in prokaryotic DNA is a matter of chance, and organisms enriched for A+T sequence may have a greater probability of containing a functional ARS.[10]

Overall, this research allows strategies to be developed to assemble and manipulate larger DNA fragments in many medically or industrially relevant bacteria that have genomes composed of high G+C. For example, sequences from *E. coli* (50% G+C), *Burkholdaria* spp. (66–68% G+C), *Mycobacterium tuberculosis* (65% G+C), *Ralstonia eutropha* (64% G+C), *Bradyrhizobium japonicum* (64% G+C), and *Geobacter metallireducens* (60% G+C) would all have difficulty in the assembly of large DNA fragments in yeast. Our current test case is limited to *S. elongatus* DNA at 55% G+C, and although we do not anticipate additional assembly problems for DNA with even higher G+C content beyond the need to insert origins of replication every 100 kb, this remains to be tested.

The ability to use yeast to perform sequence manipulations may greatly accelerate the progress of research with slow-growing cyanobacteria. Interest in photosynthetically produced bioproducts has been recently renewed by rising commodity prices and projected resource scarcity to fulfill the needs of a growing population. Photosynthetic prokaryotes such as cyanobacteria are ideal candidates for industrial biosynthetic approaches due to their autotrophic lifestyles and ability to be genetically manipulated. While many cyanobacteria have well established genetic systems, the slow growth rate of these organisms is often a factor that prohibits rapid experimental progress and multiple serial manipulations. Therefore, manipulation of large regions of cyanobacterial DNA in yeast and *E. coli* in a single step may allow more rapid progress in the development of these platforms for photosynthetically produced bioproducts, such as oils and proteins. Once the DNA is cloned in an alternative fast growing host, multiple serial manipulations can be easily made and the DNA can be returned to the cyanobacterium. Future developments in genetic manipulations of model photosynthetic prokaryotes will speed maturation of production strains.

METHODS

Strains, Media, and Growth Conditions. A highly transformable *S. cerevisiae* strain VL6-48 (ATCC MYA-3666: MATα his3-Δ200 trp1-Δ1 ura3-52 lys2 ade2-1 met14 cir^0) was used as a host for recombination-based cloning experiments. Yeast were grown in standard rich (YEPD) or complete minimal (CM) medium.[8] Ground drop-out powders were used for CM minus HIS and CM minus URA media preparation (Teknova). Media used for spheroplast transformation were supplemented with 1 M D-sorbitol (Sigma). *E. coli* strains were grown on Luria broth or agar supplemented with antibiotics at the following concentrations: chloramphenicol 25 mg L^{-1}, kanamycin 25 mg L^{-1}, spectinomycin 50 mg L^{-1}. Plasmids were electroporated into DH10b or Stbl4 electrocompentent cells (Invitrogen) according to the manufacturer's instructions. *S. elongatus* PCC 7942 genome sequence (Genbank Accession CP000100) was used for the design of all experiments. *S. elongatus* was growth in Bg11 medium at 28 °C under constant illumination with cool white fluorescent lights (50 μE m^{-2} s^{-1}).

Preparation of *S. elongatus* Genomic DNA for TAR Cloning. To isolate genomic DNA from *S. elongatus* in agarose plugs, 50 mL of mid-log phase (10^8 cells mL^{-1}) *S. elongatus* culture was centrifuged at 4000*g* for 15 min at 10 °C. Cells were washed twice with 25 mL of 10 mM Tris, 0.5 M sucrose, pH 6.5 and were resuspended in 0.5 mL of 10 mM Tris (pH 6.5), 0.5 M sucrose buffer. The cell suspension was incubated for 5 min at 50 °C and mixed with an equal volume of 2.2% low-melting-point agarose in 1xTAE buffer (40 mM Tris, 20 mM acetic acid, and 1 mM EDTA), which was also equilibrated at 50 °C. Aliquots of 100 μL were transferred into plug molds (Biorad, catalog no. 170-3713) and allowed to solidify for 10 min at 4 °C. After cooling, plugs were removed from the molds into 50-mL conical tubes containing 5 mL of lysozyme buffer (10 mM Tris, pH 7.2, 50 mM NaCl, 0.2% sodium deozycholate, 0.5% sodium lauryl sarcosine, 2 mg/mL lysozyme) and incubated for 2 h at 37 °C. After lysozyme treatment, plugs were washed with 25 mL of wash buffer (20 mM Tris, 50 mM EDTA, pH 8.0) for 1 h at RT with gentle agitation, followed by a wash with 6 mL of Proteinase K buffer

(100 mM EDTA (pH 8.0), 0.2% sodium deoxycholate, and 1% sodium lauryl sarcosine, 1 mg mL^{-1} Proteinase K) for 24 h at 50 °C. After Proteinase K treatment, plugs were washed twice with 25 mL of wash buffer for 2 h at RT, followed by incubation in wash buffer containing 1 mM PMSF for 2 h at RT, and a final wash for 2 h at RT with wash buffer.

For restriction digestion plugs were washed additionally with 0.1X wash buffer for 1 h followed by a 1-h wash with 1X restriction buffer (specific for each enzyme used) for 1 h at RT. Finally, each plug was digested in 25−50 units mL^{-1} of restriction enzyme in 1 mL for 3−12 h at the recommended temperature for each enzyme. Following the digest, plugs were washed for 1 h with TE buffer (pH 8). Next each 100-μL plug was moved to a 1.5-mL centrifuge tube to which 100 μL of TE buffer (pH 8) was added, and plugs were melted for 10 min at 65 °C. Then the solution was equilibrated at for 10 min at 42 °C before adding 2 μL of β-agarase (NEB). Finally the solution was incubated for 3−12 h to allow the agarose to be digested. On average each plug contained 50−100 ng DNA μL^{-1}. Each transformation used 0.5−2 μg of genomic DNA from melted, digested plugs.

Assembly Vector Preparation. Two plasmids, pCC1BAC-HIS3 and pCC1BAC-URA3, were used for preparation of DNA assembly vectors.[8] Vectors were amplified by PCR using primers that contained a 20-bp overlap with one end of the linearized plasmid and a 48−60-bp overlap with one end of the DNA to be assembled (Supplementary Table 3). PCR amplification of the plasmid was performed using the Phusion Hot Start High-Fidelity DNA polymerase (NEB) as described previously.[8] Vector PCR products were separated by agarose gel electrophoresis, cut from the gels, and purified from the agarose using the QIAquick Gel Extraction kit (Qiagen). For each transformation 20−400 μg of vector DNA was used.

Assembly of ∼100-kb DNA Pieces in Yeast. Assembly of several 100-kb *S. elongatus* DNA pieces was performed by homologous recombination in yeast. Vectors were prepared by PCR amplification as described above using primers Con19R and Con20F for assembly of fragment 31 and Con22R and Con23F for assembly of fragment 32 (Supplementary Table 3). In each assembly experiment 100 ng of each DNA piece and 10 ng of cloning vector were mixed with 200 μL of freshly prepared yeast spheroplasts as previously described.[30] Transformants were selected on -His or -Ura plates, and colonies were restreaked as small patches onto selective media. DNA was isolated from the patches as previously described.[31] Colonies were screened by diagnostic PCR using pairs of primers that can detect junctions between the vector and 5′-end of the assembly and the vector and 3′-end of the assembly.

Isolation of YAC DNA from Yeast. Small-scale isolation of supercoiled DNA (for DNA < 200 kb) was performed from 4 mL of yeast culture grown overnight in selective medium. Cells were centrifuged and resuspended in 1 mL of SPE buffer,[32] containing 10 μL of Zymolyase-20T (10 mg mL^{-1}) and 1 μL of 14 M β-mercaptoethanol. Following incubation at 37 °C for 1 h, cells were harvested at 5,000*g* for 5 min. Spheroplasts were resuspended in 100 μL of 1 M sorbitol by pipetting up and down, combined with 500 μL of lysis buffer,[32] and the tube was inverted several times. Following incubation at 37 °C for 30 min, 400 μL of buffer-saturated phenol (Invitrogen) was added, and the tube was inverted several times. After centrifugation at 14,000*g* for 5 min, the aqueous phase (400 μL) was transferred to a new microfuge tube, and DNA was precipitated by adding 400 μL of isopropyl alcohol and 40 μL of 3 M sodium acetate. The DNA pellet was washed once with 70% ethanol, dried, and resuspended in 100 μL of TE buffer (pH 8.0). To estimate the size of the YAC, 10 μL of the DNA solution was separated in a 1% agarose gel in 1xTAE buffer by constant voltage electrophoresis (3 h at 4.5 V cm^{-1}). After electrophoresis, the gel was stained with SYBR Gold and scanned with a Typhoon 9410 imager (GE Healthcare Life Sciences). For isolation YACs larger than 200 kb, we used the large-scale protocol described previously.[32]

Isolation of BAC DNA from *E. coli*. Minipreps of SC BAC DNA were performed from 2−4 mL of overnight culture as described previously[33,34] and were used to determine a size of the clone by gel electrophoresis. Large-scale BAC isolation was performed from 200 mL of culture using the Large Construct Kit (Qiagen) according to the manufacturer's instructions. DNA isolated by this method was used for preparation of fragments for DNA assembly.

Preparation of *S. elongatus* DNA Fragments Propagated in *E.coli* for Assembly. Fragments were prepared by digesting 10 μg of BAC DNA in 200 μL total reaction volume with a restriction enzyme that released the fragment from the vector. Following a 2 h incubation DNA was precipitated with 200 μL of isopropyl alcohol and 20 μL of 3 M sodium acetate. The resulting DNA pellet was washed with 70% ethanol, dried, and resuspended in 50 μL of TE buffer pH 8.0. DNA concentration was estimated by NanoDrop Spectrophotometer (Thermo Scientific). To verify complete DNA digestion, 100 ng of DNA was separated on a 1% agarose gel in 1xTAE buffer by field inversion gel electrophoresis (FIGE, Biorad).[32]

Lambda Red Insertion of ARS Sequence. ARS sequences were inserted into defined locations of cloned *S. elongatus* DNA fragments as antibiotic resistance cassettes linked to the ARS sequence using lambda red recombineering.[35] To insert ARS sequences into *S. elongatus* DNA fragments 19, 21, and 23, a plasmid containing the kanamycin resistance gene from pACYC177 and the ARSH3 was constructed using the one-step isothermal assembly method.[1] The KanR-ARS piece was amplified with primers that included 35−45 nt of homology to targets in the fragments for insertion using lambda red recombineering and the following primers: fragment 19 used primers 19F and 19R2, fragment 21 used primers 21F2 and 21R2, and fragment 23 used primers 23F2 and 23R2 (Supplementary Table 4). Resulting *E. coli* colonies were screened using the following primers: fragment 19 used 19screenF and 19screenR2, fragment 21 used 21screenF2 and 21screenR2, and fragment 23 used 23screenF and 23screenR2. Colonies that were positive by PCR were grown in liquid culture and plasmids were purified, confirmed to have the correct size by agarose gel electrophoresis, and used in reassembly.

ARS4 was inserted by constructing a plasmid containing a spectinomycin resistance cassette from pAM1414[36] and the ARSH3 sequence. The SpR-ARS piece was amplified using primers MpDr-1 and MpDr2 and inserted fragment 32 by lambda red recombineering.[35] Resulting *E. coli* colonies were screened by PCR using primers MpDr-3 and MpDr-4 and were used as above for other ARS insertions.

Estimation of 454-kb *S. elongatus* Clone Stability. Yeast colonies carrying the 454-kb/4ARS fragment on centromeric vector were inoculated into CM-Ura liquid medium (100 cells per ml) and grown for 2 days to stationary phase. The yeast suspension then was diluted 4 orders of magnitude, and 100 μL was plated onto CM-Ura solid medium.

Individual colonies were picked up and analyzed for completeness by Multiplex PCR.

Multiplex PCR was performed using Multiplex PCR kit (Qiagen) according to the Manufacturer's instructions. The primer set used in this procedure produced 4 amplicons (P1F + P1R: 160 bp, P2F + P2R: 230 bp, P3F + P3R: 280bp, and P4F + P4R: 350 bp) distributed along the 454-kb DNA fragment. Primer positions along the 454-kb assembly are shown in Figure 3B, and sequences are given in Supplementary Table 4.

■ ASSOCIATED CONTENT

Supporting Information

Two additional figures and four tables. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 858-200-1815. Fax: 858-200-1880. E-mail: pweyman@jcvi.org.

Author Contributions

V.N.N., B.J.K., L.Y., Y.C.L., J.S., I.T.Y., and P.D.W. performed the experiments. V.N.N., B.J.K., L.Y., J.S., I.T.Y., P.D.W., R.Y.C., D.G.G., Y.C.L., J.S., I.T.Y., Y.S., C.P., J.I.G., H.O.S., C.A.H., J.C.V., and P.D.W. designed experiments and interpreted results, and V.N.N., B.J.K., and P.D.W. wrote the paper.

Notes

The authors declare the following competing financial interest(s): J.C.V. is Chief Executive Officer and Co-Chief Scientific Officer of Synthetic Genomics, Inc. H.O.S. is Co-Chief Scientific Officer and on the Board of Directors of Synthetic Genomics, Inc. C.A.H. is Chairman of the Synthetic Genomics, Inc. Scientific Advisory Board. D.G.G. is a Principal Scientist at SGI. All four of these authors and the J. Craig Venter Institute hold Synthetic Genomics, Inc. stock.

■ ACKNOWLEDGMENTS

We thank Christian Tagwerker, Mikkel Algire, Chuck Merryman, and Sanjay Vashee for helpful discussions and suggestions. This work was supported by Synthetic Genomics, Inc. (SGI). All authors were supported in part by SGI, and B.J.K was supported by National Science and Engineering Research Council of Canada (NSERC), PDF fellowship, and by SGI.

■ REFERENCES

(1) Gibson, D. G. (2011) Enzymatic assembly of overlapping DNA fragments. *Method Enzymol. 498*, 349−361.

(2) Canton, B., Labno, A., and Endy, D. (2008) Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol. 26*, 787−793.

(3) Watanabe, K., Hotta, K., Praseuth, A. P., Koketsu, K., Migita, A., Boddy, C. N., Wang, C. C., Oguri, H., and Oikawa, H. (2006) Total biosynthesis of antitumor nonribosomal peptides in *Escherichia coli*. *Nat. Chem. Biol. 2*, 423−428.

(4) Wargacki, A. J., Leonard, E., Win, M. N., Regitsky, D. D., Santos, C. N. S., Kim, P. B., Cooper, S. R., Raisner, R. M., Herman, A., Savitz, A. B., Lakshmanaswamy, A., Kashiyama, Y., Baker, D., and Yoshikuni, Y. (2012) An engineered microbial platform for direct biofuel production from brown macroalgae. *Science 335*, 308−313.

(5) Weyman, P. D., Vargas, W. a, Chuang, R.-Y., Chang, Y., Smith, H. O., and Xu, Q. (2011) Heterologous expression of *Alteromonas macleodii* and *Thiocapsa roseopersicina* [NiFe] hydrogenases in *Escherichia coli*. *Microbiology (Reading, England) 157*, 1363−1374.

(6) Burke, D. T., Carle, G. F., and Olson, M. V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science 236*, 806−812.

(7) Larin, Z., Monaco, a P., and Lehrach, H. (1991) Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proc. Natl. Acad. Sci. U.S.A. 88*, 4123−4127.

(8) Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R.-Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z.-Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A., III, Smith, H. O., and Venter, J. C. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science 329*, 52−56.

(9) Karas, B. J., Tagwerker, C., Yonemoto, I. T., Hutchison, C. A., III, and Smith, H. O. (2012) Cloning the *Acholeplasma laidlawii* PG-8A genome in *Saccharomyces cerevisiae* as a yeast centromeric plasmid. *ACS Synth. Biol. 1*, 22−28.

(10) Tagwerker, C.; Dupont, C.; Smith, H. O.; Hutchison 3rd, C. A. Cloning the complete 1.66 Mb *Prochlorococcus marinus* MED4 genome in yeast. Personal Communication.

(11) Gibson, D. G., Benders, G. A., Axelrod, K. C., Zaveri, J., Algire, M. A., Modie, M., Moontague, M. G., Venter, J. C., Smith, H. O., and Hutchison, C. A. (2008) One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc. Natl. Acad. Sci. U.S.A. 105*, 20404−20409.

(12) Heuer, T., Bürger, C., Maass, G., and Tümmler, B. (1998) Cloning of prokaryotic genomes in yeast artificial chromosomes: application to the population genetics of *Pseudomonas aeruginosa*. *Electrophoresis 19*, 486−494.

(13) Kuspa, A., Vollrath, D., Cheng, Y., and Kaiser, D. (1989) Physical mapping of the *Myxococcus xanthus* genome by random cloning in yeast artificial chromosomes. *Proc. Natl. Acad. Sci. U.S.A. 86*, 8917−8921.

(14) Azevedo, V., Alvarez, E., Zumstein, E., Damiani, G., Sgaramella, V., Ehrlich, S. D., and Serror, P. (1993) An ordered collection of *Bacillus subtilis* DNA segments cloned in yeast artificial chromosomes. *Proc. Natl. Acad. Sci. U.S.A. 90*, 6047−60451.

(15) Albertsen, H. M., Abderrahim, H., Cann, H. M., Dausset, J., Le Paslier, D., and Cohen, D. (1990) Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. *Proc. Natl. Acad. Sci. U.S.A. 87*, 4256−4260.

(16) Newlon, C. (1988) Yeast chromosome replication and segregation. *Microbiol. Mol. Biol. Rev. 52*, 568−601.

(17) Shirahige, K., Iwasaki, T., Rashid, M. B., Ogasawara, N., and Yoshikawa, H. (1993) Location and characterization of autonomously replicating sequences from chromosome VI of *Saccharomyces cerevisiae*. *Mol. Cell. Biol. 13*, 5043−5056.

(18) Nieduszynski, C. A., Hiraga, S., Ak, P., Benham, C. J., and Donaldson, A. D. (2007) OriDB: a DNA replication origin database. *Nucleic Acids Res. 35*, D40−46.

(19) Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. U.S.A. 89*, 8794−8797.

(20) Dershowitz, A., and Newlon, C. S. (1993) The effect on chromosome stability of deleting replication origins. *Mol. Cell. Biol. 13*, 391−398.

(21) Kouprina, N., and Larionov, V. (2006) TAR cloning: insights into gene function, long-range haplotypes and genome structure and evolution. *Nat. Rev. Genet. 7*, 805−812.

(22) Newlon, C. S., and Theis, J. F. (1993) The structure and function of yeast ARS elements. *Curr. Opin. Genet. Dev. 3*, 752−758.

(23) http://genome.jgi-psf.org/synel/synel.home.html.

(24) Holtman, C. K., Chen, Y., Sandoval, P., Gonzales, A., Nalty, M. S., Thomas, T. L., Youderian, P., and Golden, S. S. (2005) High-throughput functional analysis of the *Synechococcus elongatus* PCC 7942 genome. *DNA Res. 12*, 103−115.

(25) NCBI Yeast Genome; http://www.ncbi.nlm.nih.gov/genome/15/?project_id=128.
(26) Noskov, V. N., Kouprina, N., Leem, S.-H., Ouspenski, I., Barrett, J. C., and Larionov, V. (2003) A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast. *BMC Genomics 4*, 16.
(27) Van Houten, J. V., and Newlon, C. S. (1990) Mutational analysis of the consensus sequence of a replication origin from yeast chromosome III. *Mol. Cell. Biol. 10*, 3917−3925.
(28) Kipling, D., and Kearsey, S. E. (1990) Reversion of autonomously replicating sequence mutations in *Saccharomyces cerevisiae*: creation of a eucaryotic replication origin within procaryotic vector DNA. *Mol. Cell. Biol. 10*, 265−272.
(29) DePamphilis, M. L. (1999) Replication origins in metazoan chromosomes: fact or fiction? *BioEssays 21*, 5−16.
(30) Kouprina, N., and Larionov, V. (2008) Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat. Protoc. 3*, 371−377.
(31) Kouprina, N., Noskov, V. N., and Larionov, V. (2006) Selective isolation of large chromosomal regions by transformation-associated recombination cloning for structural and functional analysis of mammalian genomes. *Methods Mol. Biol. 349*, 85.
(32) Noskov, V. N., Chuang, R.-Y., Gibson, D. G., Leem, S.-H., Larionov, V., and Kouprina, N. (2010) Isolation of circular yeast artificial chromosomes for synthetic biology and functional genomics studies. *Nat. Protoc. 6*, 89−96.
(33) Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R.-Y., Algire, M. a, Benders, G. a, Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z.-Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A., III, Smith, H. O., and Venter, J. C. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science 329*, 52−56.
(34) Noskov, V. N., Lee, N. C., Larionov, V., and Kouprina, N. (2011) Rapid generation of long tandem DNA repeat arrays by homologous recombination in yeast to study their function in mammalian genomes. *Biol. Proced. Online 13*, 8.
(35) Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A. 97*, 6640−6645.
(36) Andersson, C. R., Tsinoremas, N. F., Shelton, J., Lebedeva, N. V., Yarrow, J., Min, H., and Golden, S. S. (2000) Application of bioluminescence to the study of circadian rhythms in cyanobacteria. *Methods Enzymol. 305*, 527−542.