One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome

Daniel G. Gibson[a,1], Gwynedd A. Benders[b], Kevin C. Axelrod[a], Jayshree Zaveri[a], Mikkel A. Algire[a], Monzia Moodie[a], Michael G. Montague[a], J. Craig Venter[a], Hamilton O. Smith[b], and Clyde A. Hutchison III[b,1]

[a]The J. Craig Venter Institute, Synthetic Biology Group, Rockville, MD 20850 and [b]The J. Craig Venter Institute, Synthetic Biology Group, San Diego, CA 92121

Contributed by Clyde A. Hutchison III, October 30, 2008 (sent for review September 11, 2008)

We previously reported assembly and cloning of the synthetic *Mycoplasma genitalium* JCVI-1.0 genome in the yeast *Saccharomyces cerevisiae* by recombination of six overlapping DNA fragments to produce a 592-kb circle. Here we extend this approach by demonstrating assembly of the synthetic genome from 25 overlapping fragments in a single step. The use of yeast recombination greatly simplifies the assembly of large DNA molecules from both synthetic and natural fragments.

in vivo DNA assembly | genome synthesis | combinatorial assembly | yeast transformation | *Mycoplasma genitalium* | synthetic biology

Yeast has long been considered a genetically tractable organism because of its ability to take up and recombine DNA fragments. More than 30 years ago, Hinnen *et al.* (1) reported the restoration of leucine biosynthesis in *Saccharomyces cerevisiae* by transformation of a *leu2*$^-$ strain to *LEU2*+ using a method involving spheroplasts, $CaCl_2$, and PEG. Soon after, Orr-Weaver *et al.* (2) reported mechanistic studies demonstrating that DNA molecules taken up during yeast transformation can integrate into yeast chromosomes through homologous recombination, and that the ends of the linear-transforming DNA are highly recombinogenic and react directly with homologous chromosomal sequences, whereas circular plasmids carrying yeast sequences integrate by a single crossover and only at low frequency. Subsequently, yeast transformation has become an indispensable tool in yeast genetics.

Yeast recombination has since been applied to the construction of plasmids and yeast artificial chromosomes (YACs). In 1987, Ma *et al.* (3) constructed plasmids from two cotransformed DNA fragments containing homologous regions. In another process, called linker-mediated assembly, any DNA sequence can be joined to a vector DNA using short synthetic linkers that bridge the ends (4, 5). Similarly, four or five overlapping DNA pieces can be assembled and joined to vector DNA (4, 6, 7). This work demonstrated that yeast cells can take up multiple pieces of DNA, and that homologous yeast recombination is sufficiently efficient to correctly assemble the pieces into a single recombinant molecule.

The limitations of assembly methods in yeast remain unknown. We recently assembled an entire synthetic *M. genitalium* genome using a combination of in vitro enzymatic recombination in early stages and in vivo yeast recombination in the final stage to produce the complete genome (8). In the first stage, overlapping ≈6-kb DNA cassettes were joined four at a time to form 25 ≈24-kb A-series assemblies. Three A-series assemblies were then joined to make 1/8 genome ≈72-kb B-series assemblies, and then two B-series assemblies were assembled to make each of the ≈145-kb quarter-genome C-series assemblies. We accomplished the final assembly in yeast using three quarter-genome fragments and a fourth quarter fragment that had been cleaved by a restriction enzyme to provide a site for insertion of the vector DNA. Thus, some individual yeast cells have the capacity to simultaneously take up the overlapping genome and vector DNA fragments (a total of six pieces) and recombine them to produce the full 592-kb circular JCVI-1.0 genome. At the time, we wondered whether it would be possible to assemble DNA fragments from earlier stages into a complete genome in a single step in yeast.

The assembly intermediates from our construction of the synthetic JCVI-1.0 genome provided a resource for exploring the limits of yeast uptake and assembly. Here we report the successful assembly of the entire synthetic genome from 25 A-series assemblies in a single step.

Results

Assembly of the *M. genitalium* Genome in Yeast from 25 Overlapping DNA Fragments. Our 25 A-series assemblies comprising the entire *M. genitalium* genome were all synthesized and cloned as bacterial artificial chromosomes (BACs) in *Escherichia coli* as described previously (8) (Table 1). For this study, two variants of assembly A86–89 were produced, with the goal of constructing *M. genitalium* JCVI-1.1 and JCVI-1.9. Each variant has a vector inserted at the same site within a nonessential gene. Each vector contains a histidine auxotrophic marker, a centromere, and an origin of replication for selection and maintenance in yeast. The only difference between JCVI-1.1 and JCVI-1.9 is in the additional vector features. All 25 of these A-series molecules were purified after propagation in *E. coli*. The assemblies were released from their respective BACs by cleavage with restriction enzymes (Table 1).

The presence of yeast propagation elements within one of our assemblies (A86–89) allowed us to investigate whether yeast has the ability to assemble all of these 25 fragments into the *M. genitalium* genome. Fig. 1 illustrates our ambitious question: Can a single yeast cell take up all 25 of these pieces, translocate them to the nucleus, and then recombine them into the complete genome?

After digestion, each of the 25 assemblies contained at least 80 bp of overlapping sequence with the adjacent molecule at either end (Table 1). Equal amounts (≈100 ng or ≈4×10^9 genome equivalents) of each digested piece were pooled together without gel purification to assemble *M. genitalium* JCVI-1.1. Then ≈10^8 yeast spheroplasts were added to this DNA pool, and yeast transformation was performed (9). This amounts to ≈40 *M. genitalium* genome equivalents, or ≈1000 DNA molecules (40×25) per yeast spheroplast. After incubation for 3 days at 30 °C on selective medium, ≈800 colonies were obtained.

Author contributions: D.G.G., G.A.B., J.C.V., H.O.S., and C.A.H. designed research; M.A.A. and M.G.M. contributed new reagents/analytic tools; D.G.G., G.A.B., K.C.A., J.Z., and M.M. performed research; D.G.G., G.A.B., H.O.S., and C.A.H. analyzed data; and D.G.G., G.A.B., H.O.S., and C.A.H. wrote the paper.

Conflict of interest statement: C.A.H., H.O.S., and J.C.V. have a financial interest in Synthetic Genomics, Inc (SGI). The research reported in this paper is in the field in which SGI is developing products. SGI has funded the research reported in the paper.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: dgibson@jcvi.org or chutchison@jcvi.org.

Table 1. Summary of the 25 A-series assemblies used in constructing the *M. genitalium* synthetic genomes

Assembly	Size, bp	Restriction digest	5′ overlap, bp	3′ overlap, bp
A1–4	23,636	BsmBI	80	80
A5–8	26,166	NotI	80	80
A9–12	23,872	BsmBI	80	80
A13–16	23,653	SalI and NotI	80	80
A17–20	23,312	NotI	80	80
A21–24	23,211	NotI	80	80
A25–28	24,992	NotI	80	80
A29–32	26,257	NotI	80	209
A33–36	20,448	NotI	209	80
A37–41	25,706	NotI	80	257
A42–45	19,223	NotI	257	257
A46–49	19,895	NotI	257	257
A50–53	24,390	NotI	257	80
A54–57	24,789	NotI	80	80
A58–61	24,534	NotI	80	80
A62–65	25,055	NotI	80	80
A66–69	23,128	NotI	80	80
A70–73	23,345	NotI	80	80
A74–77	21,396	NotI	80	165
A78–81	22,320	NotI	165	360
A82–85	23,460	NotI	360	85
A86–89–1.1 variant	34,950	NotI	85	80
A86–89–1.9 variant	31,726	NotI	85	80
A90–93	23,443	NotI	80	80
A94–97	24,511	NotI	80	80
A98–101	17,372	NotI	80	80

Analysis of 10 Clones by Multiplex PCR. We first screened for yeast cells that took up all 25 pieces. Forty primer pairs, producing amplicons ranging in size from 100 bp to 1075 bp, were designed such that they could produce 10 amplicons in each of four individual multiplex PCR reactions (Fig. 2*A*). Multiplex primer sets 1 and 2 were described previously (8). The 40 amplicons are positioned around the *M. genitalium* genome approximately every 15 kb (Fig. 2*B*). All 25 segments must be present to give rise to all 40 amplicons by PCR, thus giving a very good indication that all 25 pieces were incorporated into a yeast cell.

Ten individual colonies were transferred to a single selective plate as small patches. After incubation for 2 days at 30 °C, ≈10^7 cells from each of these 10 patches were scraped into 1 ml of water. Using multiplex primer set 1 and DNA extracted from these 10 clones, multiplex PCR was performed, and the products were analyzed by gel electrophoresis. Clones 1 and 4 produced all 10 amplicons (Fig. 2*C*). Amplicon 1-h was observable when the exposure time was increased. This PCR analysis confirmed the

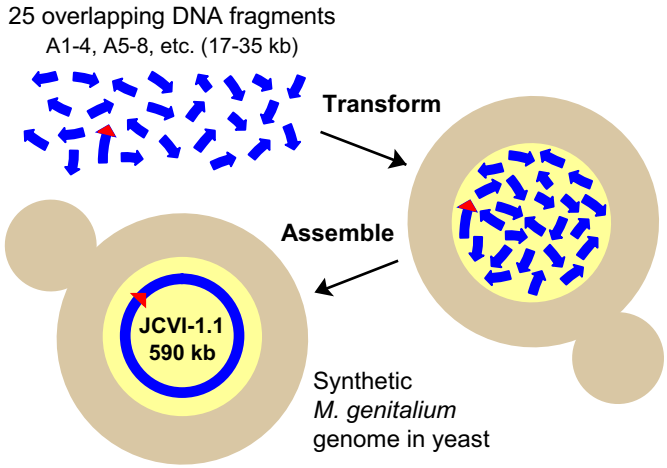


Fig. 1. Construction of a synthetic *M. genitalium* genome in yeast. Yeast cells were transformed with 25 different overlapping A-series DNA segments (blue arrows; ≈17 kb to ≈35 kb each) composing the *M. genitalium* genome. To assemble these into a complete genome, a single yeast cell (tan) must take up at least one representative of the 25 different DNA fragments and incorporate them in the nucleus (yellow), where homologous recombination occurs. This assembled genome, called JCVI-1.1, is 590,011 bp, including the vector sequence (red triangle) shown internal to A86–89. The yeast propagation elements contained within the vector are an origin of replication (*ARSH4*), a centromere (*CEN6*), and a histidine-selectable marker (*HIS3*). In addition to full assembly of the genome as depicted here, some yeast cells may take up fewer than 25 different pieces and produce subassemblies of the genome by a mechanism such as NHEJ (see text). Others may take up more than 25 fragments and produce more than one assembled molecule per cell (not illustrated).

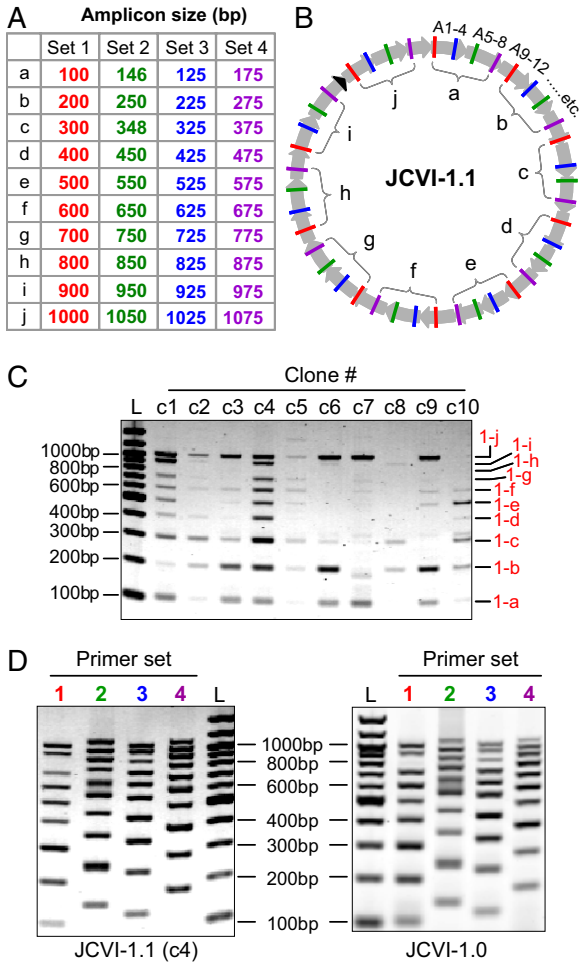


Fig. 2. Multiplex PCR analysis to screen for yeast cells that took up all 25 segments. (*A*) Forty amplicons were designed such that 10 products could be produced in four separate multiplex PCR reactions (set 1 [red], set 2 [green], set 3 [blue], set 4 [purple]). PCR products ranged in size from 100 bp to 1075 bp and could be easily separated by electrophoresis on 2% agarose gels (*C* and *D*). (*B*) The 40 sets of primers amplified a small portion of the *M. genitalium* genome approximately once every 15 kb. Each of the 25 fragments (gray arrows) provided primer-binding sites for at least 1 of the 40 amplicons (red, green, blue, and purple lines). The vector sequence is represented by a black triangle. (*C*) DNA was extracted from 10 yeast clones, and multiplex PCR, with primer set 1, was performed. Clones 1 and 4 (c1 and c4) efficiently generated 9 of the 10 predicted amplicons. Clone 4 was selected for further analysis. (*D*) Multiplex PCR was performed on clone 4 and JCVI-1.0 as a positive control using all four primer sets. With the exception of amplicon 1-h, all 40 amplicons were efficiently generated from these clones. In the lanes labeled "L," the 100-bp ladder (New England Biolabs) was loaded; sizes are indicated.

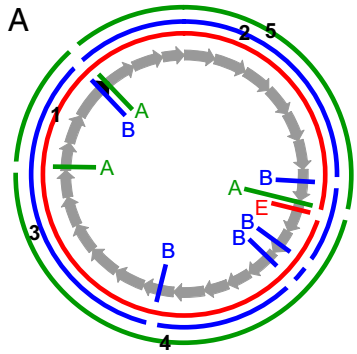
presence of 10 of the 25 assemblies in clones 1 and 4 (compare with Fig. 2*B*).

Detailed Analysis of Clone 4. Individual colonies, obtained by restreaking the patches onto selective medium, were analyzed (see *Materials and Methods*). Using multiplex primer sets 1–4, multiplex PCR was performed, and the products were visualized using gel electrophoresis (Fig. 2*D*). With the exception of amplicon 1-h, all 40 PCR products were efficiently generated from DNA extracted from clone 4. This confirms the presence of all 25 pieces (compare Fig. 2*B*). A70–73 was represented by amplicons 1-h and 3-h. Although amplicon 1-h was generated at lower levels than the other amplicons, amplicon 3-h was produced efficiently, providing strong evidence for the presence of A70–73 in clone 4. Furthermore, amplicon 1-h was not efficiently generated even in JCVI-1.0, the sequence-confirmed positive control (Fig. 2*D*) (8).

To determine whether all 25 pieces were assembled into the complete genome in clone 4, DNA was prepared from these yeast cells in agarose plugs for restriction analyses. As a negative control, DNA also was extracted from the untransformed host strain. To enrich for the circular genome in the plugs, most of the linear yeast chromosomal DNA was removed by constant voltage electrophoresis. Digestion was done with EagI, BssHII, and AatII (Fig. 3*A* and *B*). Five restriction fragments predicted by these digests for a complete genome assembly were observed when the digested plugs were subjected to field-inversion gel electrophoresis (FIGE) (Fig. 3*C*). The host yeast control showed none of these bands, and only a smear of DNA smaller than ≈150 kb was observed. This smear obscured the other predicted fragments of the assembled genome. Together with the multiplex PCR analyses (Fig. 2), these results indicate that clone 4 had an assembled JCVI-1.1 synthetic genome, and thus yeast cells have the capacity to take up at least 25 overlapping segments of DNA and recombine them in vivo.

Clone 1 exhibited a pattern similar to that of clone 4 when multiplex PCR products, amplified using primer set 1, were analyzed (Fig. 2*C*). Therefore, we analyzed clone 1 further. Using the four multiplex PCR primer sets shown in Fig. 2, all 40 amplicons were observed for clone 1, and after an EagI digest (as shown in Fig. 3), the predicted 590-kb band was present (data not shown). Thus, 2 out of 10 yeast clones were confirmed to contain a complete JCVI-1.1 genome after this single transformation experiment.

Accurate Assembly of the 25 Overlapping Segments. Many organisms, including *S. cerevisiae*, perform nonhomologous end joining (NHEJ) as a way to maintain the integrity of the genome in the presence of double-strand breaks (10). In contrast to homologous recombination, this mechanism for joining DNA does not depend on sequence identity. Thus, NHEJ possibly could be responsible for joining one or more of the A-series molecules. Moreover, because most of the fragments contain NotI cohesive ends, some joining could have occurred by cohesive-end ligation. Either mechanism could result in a genome containing more than or fewer than one each of the 25 segments; for example, one segment may not participate in assembly and be replaced by a duplicate A-series molecule. The multiplex PCR results demonstrating the presence of all 25 pieces could then be explained by multiple subassemblies within the same cell. To address this possibility, the linearized JCVI-1.1 (C4) genome fragment was gel-purified along with JCVI-1.0 as a positive control (Fig. 4*A*). These DNA molecules were then used as template for multiplex PCR using the four sets of primers described in Fig. 2. All 40 amplicons were detected from the gel-purified JCVI-1.0 and JCVI-1.1 genomes (Fig. 4*B*), confirming the presence of all 25 segments within the linearized JCVI-1.1 genome. This analysis does not verify that all 25 pieces were assembled in the correct order, however. To validate a correctly assembled genome, PCR primers were designed to produce an amplicon across each of the 25 junctions (Fig. 4*C*). All 25 amplicons were produced after PCR, and the sizes exactly matched those



B

Fragment	1 EagI	2 BssHII	3 BssHII	4 AatII	5 AatII
Size (bp)	590,011	225,833	203,565	273,856	241,248

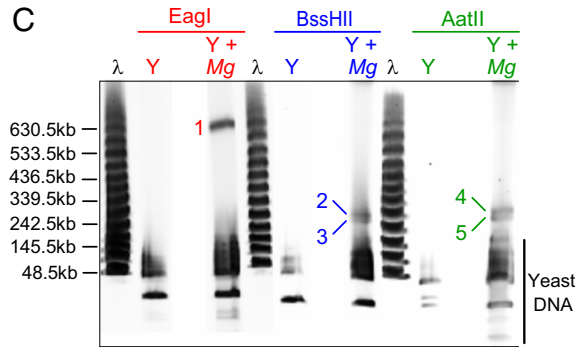
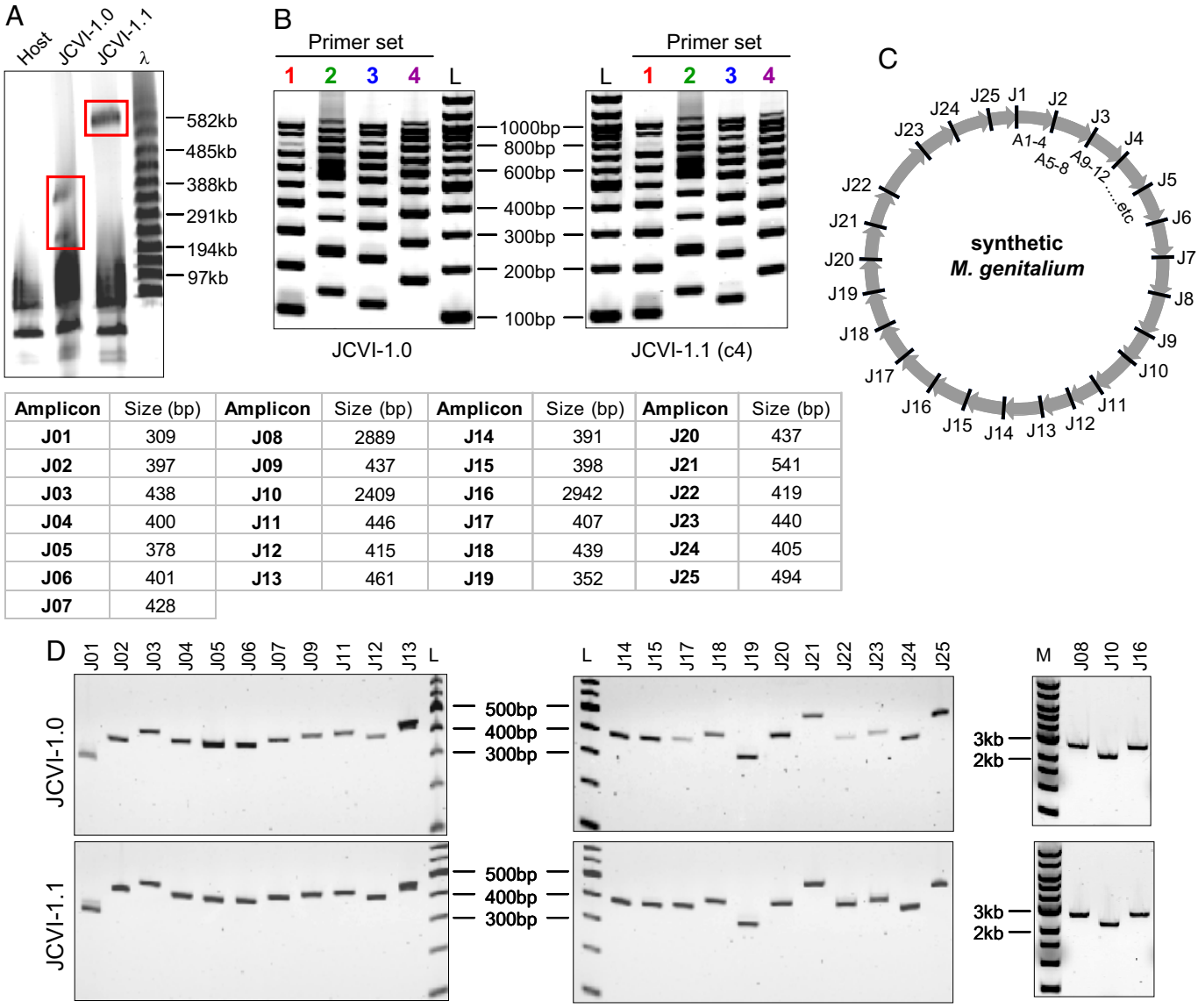


Fig. 3. Validation of an intact *M. genitalium* genome by restriction analysis. (*A*) Diagram of the EagI (red), BssHII (blue), and AatII (green) restriction fragments expected for a complete and proper assembly of the synthetic JCVI-1.1 *M. genitalium* genome. The 25 pieces are represented by gray arrows. (*B*) The sizes of the five restriction fragments indicated in (*A*). (*C*) Total DNA from clone 4 (Y + *Mg*) and its host strain alone (Y) were isolated from yeast cells embedded in agarose. Most of the linear DNA was electrophoresed out of the agarose plugs. These plugs were then digested with EagI, BssHII, or AatII and analyzed by FIGE on 1% agarose gels. Restriction fragments corresponding to the correct sizes are indicated by the fragment numbers shown in (*A*) and (*B*). In the lanes labeled "λ," the lambda ladder (New England Biolabs) was loaded; sizes are indicated.

predicted (Fig. 4*D*). This convincingly demonstrates that the JCVI-1.1 genome was assembled in vivo by homologous recombination. We cannot rule out the possibility that one or more additional molecules were assembled within the same cell.

Reproducibility of the Single-Step, 25-Piece Assembly. We investigated whether a third version of the synthetic *M. genitalium* genome could be assembled in an independent experiment. Transformation of the 25 A-series fragments into yeast yielded 150 colonies. DNA was extracted from 10 of these colonies (c11–c20), and then multiplex PCR, using primer set 3, was performed. DNA from clone 11 produced all 10 amplicons predicted for a complete assembly (Fig. 5*A*), so this clone was further analyzed using all four multiplex primer sets (Fig. 5*B*). All 40 amplicons were efficiently generated from this clone, suggesting the presence of all 25 A-series molecules within these cells. To determine whether this genome, termed JCVI-1.9, was completely assembled, DNA from this clone was prepared in agarose plugs as shown in Fig. 3 and then analyzed by contour-clamped homogeneous electric field (CHEF) gel electrophoresis alongside a host-only negative control (strain VL6–48N). A fragment of the correct size for JCVI-1.9 (≈587 kb) is evident (Fig. 5*C*). To determine whether this genome was assem-



Amplicon	Size (bp)	Amplicon	Size (bp)	Amplicon	Size (bp)	Amplicon	Size (bp)
J01	309	**J08**	2889	**J14**	391	**J20**	437
J02	397	**J09**	437	**J15**	398	**J21**	541
J03	438	**J10**	2409	**J16**	2942	**J22**	419
J04	400	**J11**	446	**J17**	407	**J23**	440
J05	378	**J12**	415	**J18**	439	**J24**	405
J06	401	**J13**	461	**J19**	352	**J25**	494
J07	428						

Fig. 4. PCR analysis of linearized and gel-purified JCVI-1.1. (*A*) Total DNA from JCVI-1.1 (c4), its host strain alone as a negative control, and JCVI-1.0 as a positive control were prepared and then analyzed after digestion with EagI as shown in Fig. 3. JCVI-1.0 has two additional EagI sites within its vector sequence, producing two genome fragments of ≈233 kb and ≈350 kb as well as a released ≈9-kb vector. The linearized JCVI-1.0 and JCVI-1.1 genome fragments were cut out of the gel, as indicated by red rectangles. (*B*) DNA from the gel slices in (*A*) was extracted and purified, then used as template for multiplex PCR, as shown in Fig. 2. (*C*) Twenty-five amplicons (J1-J25, black lines) are produced from 25 sets of primers that span each of the 25 junctions at which joining of the A-series assemblies (gray arrows) occurs. The predicted amplicon sizes for correctly assembled molecules are indicated. It was necessary to design larger amplicons for J08, J10, and J16 to ensure unique primer binding sites due to MgPa repeats at these junctions. (*D*) Purified JCVI-1.0 and JCVI-1.1 DNA extracted from the gel slices in (*A*) produced all 25 expected fragments after PCR. Products were analyzed on 2% agarose gels except for J08, J10, and J16, which were analyzed on 0.8% agarose gels. In the lanes labeled ''M,'' the 1-kb ladder (New England Biolabs) was loaded; sizes are indicated.

bled in the correct order by homologous recombination, this linearized genome fragment was gel-purified and subjected to PCR at each junction, as shown in Fig. 4. All predicted amplicons were produced for JCVI-1.9, and the PCR pattern exactly matched that for the positive control, JCVI-1.0 (Fig. 5*D*; compare with Fig. 4*D*), demonstrating construction of a second accurately assembled synthetic genome by an independent transformation event of 25 A-series molecules.

Discussion

Statistics for Yeast DNA Uptake. An individual yeast cell must take up at least one each of the 25 DNA pieces to assemble the complete genome. If a cell randomly takes up exactly 25 pieces, then the probability that it takes up one of each is $25!/25^{25} = 1.7 \times 10^{-10}$—a very rare occurrence. Thus, the average yeast cell likely took up substantially more than 25 pieces, given the results we observed. If a yeast cell randomly takes up $n > 25$ equally represented pieces, what is the probability of getting all 25 different pieces? This is an example of the "collector's problem" (11). An approximate calculation of the probabilities for different n can be readily obtained by computer simulation. Fig. 6 shows that if a cell takes up ≈90 pieces, then it has about a 50% chance of taking up all 25 different pieces. In a typical experiment, 10^8 yeast cells are exposed to DNA. Therefore, even if the average cell were to take up only 40 pieces, about 1 in 1000 would get a complete set of pieces. The efficiency of DNA uptake and assembly in yeast is unknown. If recombination is highly efficient, then only a few yeast cells in the population might need to take up a complete set of overlapping pieces. If the

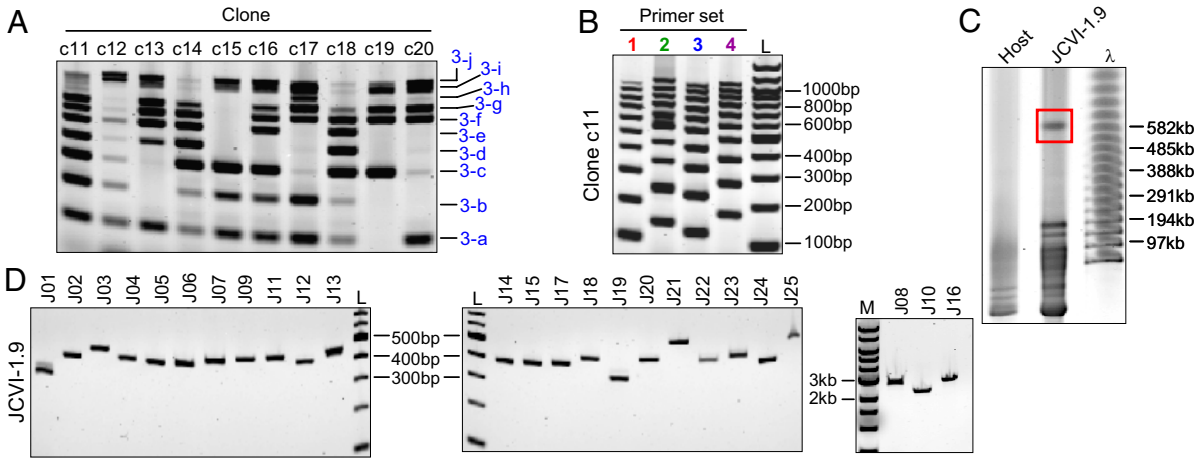


Fig. 5. The construction of JCVI-1.9 by an independent 25-piece assembly. (*A*) DNA was extracted from 10 yeast clones (c11–c20), and multiplex PCR, with primer set 3, was performed. Clone 11 generated all 10 predicted amplicons and thus was selected for further analysis. (*B*) Multiplex PCR was performed on clone 11 using all four primer sets. All 40 amplicons were efficiently generated from this clone. (*C*) DNA was prepared from clone 11 in agarose plugs, then digested with EagI, as shown in Fig. 3. The linearized genome (≈587 kb) was separated on a 1% agarose gel in 0.5X Tris-borate-EDTA buffer at 14 °C on a Bio-Rad Mapper XA CHEF system. The run time was 16 h at 6 V/cm with a 50–90 s switch time ramp at an included angle of 120°. The linearized JCVI-1.9 genome fragment was cut out of the gel as indicated by the red rectangle. (*D*) DNA from the gel slice in (*C*) was extracted and purified, then used as template for PCR at the 25 junctions shown in Fig. 4. All 25 predicted amplicons were observed after analysis on 2% and 0.8% agarose gels.

efficiency is low, then relatively more yeast cells would require a complete set. Independent measurements of the quantity of DNA taken up per cell would help answer some of these questions.

Exploring the Capacity of DNA Assembly in Yeast. We still do not know the limits of DNA assembly in yeast. Once we assembled the synthetic *M. genitalium* genome from six pieces, we wondered whether we could have assembled it from 25. Now we wonder whether this genome could have been assembled from cotransformation of all 101 DNA cassettes (8). We also do not yet know the size limit for DNA assembly in yeast; however, YACs as large as 2.3 Mb have been cloned (12). Presumably, any DNA construct can be assembled inside a yeast cell as long as the overlapping subfragments can get into the nucleus, and the constructed DNA segment is completely and accurately replicated and not toxic to the host cell. DNA replication of the assembled molecule in yeast depends on the presence of the highly conserved and essential autonomously replicating sequence (ARS) consensus sequence, $5'$-$^{T}_{A}$TTTA$^{TA}_{CG}$T-TT$^{T}_{A}$-$3'$ (13). Statistically, this sequence should be present once every ≈262 kb in an assembled construct when all four bases are equally represented; however, these elements should be present more frequently in the *M. genitalium* genome given its high A+T composition. Indeed, 20 instances of the ARS consensus sequence can be found in our assembled genome, including 2 from the vector. This is an important consideration, because a 170-kb YAC lacking efficient origins of replication can invoke a cell-cycle checkpoint response in yeast (14).

Once individual clones have been identified as containing a complete genome, instability within yeast has not been observed (8). This is a common concern, however; deletions ranging in size from several kb to 250 kb during the transformation process and cell growth of the transformants have been reported (15). These deletions can result from recombination between repetitive sequences within the genome, which can be suppressed by using a recombination-deficient yeast strain (16).

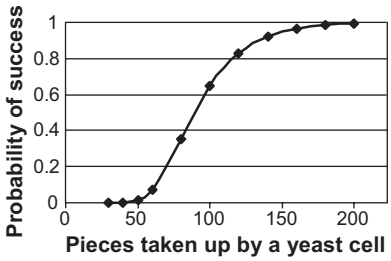


Fig. 6. Plot of the probability that an entire set of 25 DNA pieces is taken up (success) versus the number of DNA pieces taken up by an individual yeast cell. Data are from a computer simulation. N draws were made from a bag of 25 objects (pieces of DNA), replacing the object after each draw. The trial was a success if all 25 objects were drawn at least once. A million trials were conducted for each N to obtain a good approximation to the true probability of success.

Combinatorial Assembly. It should be possible to use the method described here to create many combinations of DNA molecules in yeast. Genetic pathways can be constructed in a combinatorial fashion such that each member in the combinatorial library has a different combination of gene variants. For example, a combinatorial library of 6^5 variants can be constructed from 30 individual DNA elements (6 variants of each of 5 overlapping elements in the pathway). The elements can be genes, transcriptional promoters and terminators, and appropriate signals for initiation of protein synthesis. Through screening and selection methods, yeast cells bearing the best pathways with the highest yields can be obtained. The yeast then can become factories for the manufacture of the products specified by the designed pathways. Alternatively, if yeast is not the intended host organism, then DNA can be extracted from a pool of individual yeast clones, each containing a variant of the pathway, and transformed into the desired host, assuming that the required vector elements are present.

Rapid Construction of Large DNA Molecules. We have made use of yeast's remarkable capacity to take up and recombine numerous overlapping DNA molecules to assemble a complete genome. Assembling large, genome-size molecules in host organisms by in vivo recombination is not a new concept. Holt *et al.* (17) proposed using the lambda Red recombination system to assemble an 1830-kb *Haemophilus influenzae* genome within an *E. coli* cell. Itaya *et al.* (18) and Yonemura *et al.* (19) developed methods for assembling large DNA segments in *Bacillus subtilis*. But these DNA molecules are built in the host organism only after stepwise addition

of subfragments. Genome assembly in yeast, as we have described it, is accomplished not by the addition of overlapping segments one at a time, but rather by cotransformation of 25 different pieces at once. Thus, large DNA molecules can be assembled much more rapidly from synthetic or naturally occurring subfragments than with any other system described previously. Our methods should accelerate research projects, particularly in the emerging field of synthetic biology.

Materials and Methods

Production of A-Series Assemblies. Each of the 25 A-series assemblies (Table 1) is contained in pCC1BAC (Epicentre) and was propagated in the Epi300 *E. coli* strain (Epicentre). This cloning system permits induction to 10 or more copies of these BACs per cell. An *E. coli* strain carrying each of these 25 pieces was inoculated into 150 ml of LB plus 12.5 μg/ml of chloramphenicol and 1X induction solution (Epicentre) and incubated at 37 °C for 16 h. The cultures were harvested, and the BACs were purified using a Qiagen HiSpeed Plasmid Maxi Kit. DNA was eluted using 500 μl of Tris-EDTA (TE) buffer. Each insert was excised by restriction digestion at the recommended temperature for 16 h and then terminated by phenol-chloroform extraction and ethanol precipitation. The products were dissolved in TE buffer to a final concentration of 120 ng/μl, as determined by gel electrophoresis with standards.

Yeast Spheroplast Transformation. In the yeast spheroplast transformation procedure, cells are treated with zymolyase to remove the cell wall and then made competent to take up foreign DNA by treatment with PEG and $CaCl_2$. This procedure was carried out using a previously published protocol with the VL6–48N yeast strain (9). Cells were grown to an OD_{600} of 1.3 ($\approx 5 \times 10^7$ cells/ml) before preparation of yeast spheroplasts. The 25 digested fragments were pooled by adding 96 ng (0.8 μl of 120 ng/μl) of each, then mixed with $\approx 10^8$ yeast spheroplasts. After transformation, yeast spheroplasts were regenerated and selected on complete supplemental medium without histidine (CSM-His) and 1 M sorbitol agar plates for 3 days at 30 °C. Primary transformants were then transferred onto CSM-His agar plates as small patches and incubated for 2 days at 30 °C. After screening and confirmation of the desired assembly, individual colonies for analysis were obtained by restreaking these patches onto selective plates. Only 10%–20% of the restreaked colonies contained an intact genome. These were used for further analysis.

Multiplex PCR Analysis. DNA was extracted from $\approx 10^7$ yeast cells using an Invitrogen ChargeSwitch Plasmid Yeast Mini Kit according to the manufacturer's instructions. Multiplex PCR was done using a Qiagen Multiplex PCR Kit. A 1/50 volume (2 μl) of the DNA extract and 1 μl of a 10X primer stock containing 20 oligos at 5 μM each were included in each 10-μl reaction. Cycling parameters were 94 °C for 15 min, then 35 cycles of 94 °C for 30 s, 52 °C for 90 s, and 72 °C for 90 s, followed by a single 3-min incubation at 72 °C. Then 2 μl of each reaction was loaded onto a 2% E-gel (Invitrogen), and 72 V was applied for 30 min. Bands were visualized using an Amersham Typhoon 9410 Fluorescence Imager.

Restriction Analysis of the Synthetic *M. genitalium* Genomes. DNA was prepared from yeast clones in 1% agarose using a Bio-Rad Yeast DNA Plug Kit, following the manufacturer's instructions. After this protocol was completed, plugs were submerged in 1X Tris-acetate-EDTA (TAE) buffer in a horizontal submarine electrophoresis tray, and 5.4 V/cm was applied for 2 h. The plugs were then removed and stored in TE buffer. Plugs were digested according to the Bio-Rad manual with EagI, BssHII, or AatII; restriction enzymes and buffers were supplied by New England Biolabs. After incubation at 37 °C for 16 h, equal amounts of each plug ($\approx$20 μl) were loaded onto 1% BioRad Ready Agarose Mini Gels and subjected to FIGE for 14 h at 23 °C in a Hoeffer HE33 mini horizontal submarine electrophoresis tray using 1X TAE buffer with 0.5 μg/ml of ethidium bromide without circulation. The FIGE parameters were forward: 90 V; initial switch, 5.0 s; final switch, 30 s, with linear ramp, and reverse: 60 V; initial switch, 5.0 s; final switch, 30 s, with linear ramp. Bands were visualized using an Amersham Typhoon 9410 Fluorescence Imager with 532-nm excitation and 610-nm emission wavelengths.

Analysis of Linearized Synthetic Genomes. Gel slices containing the linearized synthetic genomes were excised, and the DNA was purified using a Qiagen QIAquick Gel Extraction Kit according to the manufacturer's instructions. A fraction of the eluted DNA (1/100) was used in multiplex PCR as above and in PCR at the 25 junctions as follows. Amplicon J13 was generated using a Takara LA PCR kit, version 2.1, according to the manufacturer's instructions at an annealing temperature of 60 °C. All other amplicons were generated using a New England Biolabs Phusion Hot Start High-Fidelity DNA polymerase with HF buffer according to the manufacturer's instructions at an annealing temperature of 57 °C. Primers for amplification were designed to be positioned outside the overlaps of the 25 A-series assemblies and to specifically generate a single product.

ACKNOWLEDGMENTS. We thank John I. Glass and Vladimir N. Noskov for critically reading the manuscript and Radha Krishnakumar for providing technical assistance. We also thank Fred Blattner for suggesting additional important experiments that greatly improved the manuscript. This work was supported by Synthetic Genomics, Inc.

1. Hinnen A, Hicks JB, Fink GR (1978) Transformation of yeast. *Proc Natl Acad Sci U S A* 75:1929–1933.
2. Orr-Weaver TL, Szostak JW, Rothstein RJ (1981) Yeast transformation: A model system for the study of recombination. *Proc Natl Acad Sci U S A* 78:6354–6358.
3. Ma H, Kunes S, Schatz PJ, Botstein D (1987) Plasmid construction by homologous recombination in yeast. *Gene* 58:201–216.
4. Raymond CK, Pownder TA, Sexson SL (1999) General method for plasmid construction using homologous recombination. *Biotechniques* 26:134–141.
5. Raymond CK, Sims EH, Olson MV (2002) Linker-mediated recombinational subcloning of large DNA fragments using yeast. *Genome Res* 12:190–197.
6. Marykwas DL, Passmore SE (1995) Mapping by multifragment cloning in vivo. *Proc Natl Acad Sci U S A* 92:11701–11705.
7. Ebersole T, *et al.* (2005) Rapid generation of long synthetic tandem repeats and its application for analysis in human artificial chromosome formation. *Nucleic Acids Res* 33:e130.
8. Gibson DG, *et al.* (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 319:1215–1220.
9. Kouprina N, 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.
10. Daley JM, Palmbos PL, Wu D, Wilson TE (2005) Nonhomologous end joining in yeast. *Annu Rev Genet* 39:431–451.
11. Feller W (1957) *An Introduction to Probability Theory and Its Applications* (Wiley, New York), pp 47–50.
12. Marschall P, Malik N, Larin Z (1999) Transfer of YACs up to 2.3 Mb intact into human cells with polyethylenimine. *Gene Ther* 6:1634–1637.
13. Bell SP (1995) Eukaryotic replicators and associated protein complexes. *Curr Opin Genet Dev* 5:162–167.
14. van Brabant AJ, Buchanan CD, Charboneau E, Fangman WL, Brewer BJ (2001) An origin-deficient yeast artificial chromosome triggers a cell cycle checkpoint. *Mol Cell* 7:705–713.
15. Kouprina N, Eldarov M, Moyzis R, Resnick M, Larionov V (1994) A model system to assess the integrity of mammalian YACs during transformation and propagation in yeast. *Genomics* 21:7–17.
16. Larionov V, *et al.* (1994) Transformation-associated recombination between diverged and homologous DNA repeats is induced by strand breaks. *Yeast* 10:93–104.
17. Holt RA, Warren R, Flibotte S, Missirlis PI, Smailus DE (2007) Rebuilding microbial genomes. *Bioessays* 29:580–590.
18. Itaya M, Tsuge K, Koizumi M, Fujita K (2005) Combining two genomes in one cell: Stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc Natl Acad Sci U S A* 102:15971–15976.
19. Yonemura I, *et al.* (2007) Direct cloning of full-length mouse mitochondrial DNA using a *Bacillus subtilis* genome vector. *Gene* 391:171–177.