

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,¹ John I. Glass,¹ Carole Lartigue,¹ Vladimir N. Noskov,¹ Ray-Yuan Chuang,¹ Mikkel A. Algire,¹ Gwynedd A. Benders,² Michael G. Montague,¹ Li Ma,¹ Monzia M. Moodie,¹ Chuck Merryman,¹ Sanjay Vashee,¹ Radha Krishnakumar,¹ Nacyra Assad-Garcia,¹ Cynthia Andrews-Pfannkoch,¹ Evgeniya A. Denisova,¹ Lei Young,¹ Zhi-Qing Qi,¹ Thomas H. Segall-Shapiro,¹ Christopher H. Calvey,¹ Prashanth P. Parmar,¹ Clyde A. Hutchison III,² Hamilton O. Smith,² J. Craig Venter^{1,2*}

¹The J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, MD 20850, USA. ²The J. Craig Venter Institute, 10355 Science Center Drive, San Diego, CA 92121, USA.

*To whom correspondence should be addressed. E-mail: jcventer@jcv.org

We report the design, synthesis and assembly of the 1.08-Mbp *Mycoplasma mycoides* JCVI-syn1.0 genome starting from digitized genome sequence information and its transplantation into a *Mycoplasma capricolum* recipient cell to create new *Mycoplasma mycoides* cells that are controlled only by the synthetic chromosome. The only DNA in the cells is the designed synthetic DNA sequence, including “watermark” sequences and other designed gene deletions and polymorphisms, and mutations acquired during the building process. The new cells have expected phenotypic properties and are capable of continuous self-replication.

In 1977, Sanger and colleagues determined the complete genetic code of phage ϕ X174 (1), the first DNA genome to be completely sequenced. Eighteen years later, in 1995, our team was able to read the first complete genetic code of a self-replicating bacterium, *Haemophilus influenzae* (2). Reading the genetic code of a wide range of species has increased exponentially from these early studies. Our ability to rapidly digitize genomic information has increased by more than eight orders of magnitude over the past 25 years (3). Efforts to understand all this new genomic information have spawned numerous new computational and experimental paradigms, yet our genomic knowledge remains very limited. No single cellular system has all of its genes understood in terms of their biological roles. Even in simple bacterial cells, do the chromosomes contain the entire genetic repertoire? If so, can a complete genetic system be reproduced by chemical synthesis starting with only the digitized DNA sequence contained in a computer?

Our interest in synthesis of large DNA molecules and chromosomes grew out of our efforts over the past 15 years to build a minimal cell that contains only essential genes. This work was inaugurated in 1995 when we sequenced the genome from *Mycoplasma genitalium*, a bacterium with the smallest complement of genes of any known organism

capable of independent growth in the laboratory. More than 100 of the 485 protein-coding genes of *M. genitalium* are dispensable when disrupted one-at-a-time (4–6).

We developed a strategy for assembling viral sized pieces to produce large DNA molecules that enabled us to assemble a synthetic *M. genitalium* genome in four stages from chemically synthesized DNA cassettes averaging about 6 kb in size. This was accomplished through a combination of in vitro enzymatic methods and in vivo recombination in *Saccharomyces cerevisiae*. The whole synthetic genome (582,970 bp) was stably grown as a yeast centromeric plasmid (YCp) (7).

Several hurdles were overcome in transplanting and expressing a chemically synthesized chromosome in a recipient cell. We needed to improve methods for extracting intact chromosomes from yeast. We also needed to learn how to transplant these genomes into a recipient bacterial cell to establish a cell controlled only by a synthetic genome. Due to the fact that *M. genitalium* has an extremely slow growth rate, we turned to two faster growing mycoplasma species, *M. mycoides subspecies capri* (GM12) as donor, and *M. capricolum subspecies capricolum* (CK) as recipient.

To establish conditions and procedures for transplanting the synthetic genome out of yeast, we developed methods for cloning entire bacterial chromosomes as centromeric plasmids in yeast, including a native *M. mycoides* genome (8, 9). However, initial attempts to extract the *M. mycoides* genome from yeast and transplant it into *M. capricolum* failed. We discovered that the donor and recipient mycoplasmas share a common restriction system. The donor genome was methylated in the native *M. mycoides* cells and was therefore protected against restriction during the transplantation from a native donor cell (10). However, the bacterial genomes grown in yeast are unmethylated and so are not protected from the single restriction system of the recipient cell. We were able to overcome this restriction

barrier by methylating the donor DNA with purified methylases or crude *M. mycoides* or *M. capricolum* extracts, or by simply disrupting the recipient cell's restriction system (8).

We now have combined all of our previously established procedures and report the synthesis, assembly, cloning, and successful transplantation of the 1.08-Mbp *M. mycoides* JCVI-syn1.0 genome, to create a new cell controlled by this synthetic genome.

Results

Synthetic genome design

Design of the *M. mycoides* JCVI-syn1.0 genome was based on the highly accurate finished genome sequences of two laboratory strains of *M. mycoides* subspecies *capri* GM12 (8, 9) (11). One was the genome donor used by Lartigue et al. [GenBank accession CP001621] (10). The other was a strain created by transplantation of a genome that had been cloned and engineered in yeast, YCpMmyc1.1-*AtypeIIIres*, [GenBank accession CP001668] (8). This project was critically dependent on the accuracy of these sequences. Although we believe that both finished *M. mycoides* genome sequences are reliable, there are 95 sites at which they differ. We began to design the synthetic genome before both sequences were finished. Consequently, most of the cassettes were designed and synthesized based upon the CP001621 sequence (11). When it was finished, we chose to use the sequence of the genome successfully transplanted from yeast (CP001668) as our design reference (except that we kept the intact *typeIIIres* gene). All differences that appeared biologically significant between CP001668 and previously synthesized cassettes were corrected to match it exactly (11). Sequence differences between our synthetic cassettes and CP001668 that occurred at 19 sites appeared harmless, and so were not corrected. These provide 19 polymorphic differences between our synthetic genome (JCVI-syn1.0) and the natural (non-synthetic) genome (YCpMmyc1.1) that we have cloned in yeast and use as a standard for genome transplantation from yeast (8). To further differentiate between the synthetic genome and the natural one, four watermark sequences (fig. S1) were designed to replace one or more cassettes in regions experimentally demonstrated (watermarks 1 [1246 bp] and 2 [1081 bp]) or predicted (watermarks 3 [1109 bp] and 4 [1222 bp]) to not interfere with cell viability. These watermark sequences encode unique identifiers while limiting their translation into peptides. Table S1 lists the differences between the synthetic genome and this natural standard. Figure S2 shows a map of the *M. mycoides* JCVI-syn1.0 genome. Cassette and assembly intermediate boundaries, watermarks, deletions, insertions, and genes of the *M. mycoides* JCVI syn1.0 are shown in fig. S2, and the sequence of the transplanted mycoplasma clone

sMmYCp235-1 has been submitted to GenBank (accession # CP002027).

pSynthetic genome assembly strategy

The designed cassettes were generally 1,080 bp with 80-bp overlaps to adjacent cassettes (11). They were all produced by assembly of chemically synthesized oligonucleotides by Blue Heron; Bothell, Washington. Each cassette was individually synthesized and sequence-verified by the manufacturer. To aid in the building process, DNA cassettes and assembly intermediates were designed to contain Not I restriction sites at their termini, and recombined in the presence of vector elements to allow for growth and selection in yeast (7) (11).

pA hierarchical strategy was designed to assemble the genome in 3 stages by transformation and homologous recombination in yeast from 1,078 one-kb cassettes (Fig. 1) (12, 13).

Assembly of 10-kb synthetic intermediates. In the first stage, cassettes and a vector were recombined in yeast and transferred to *E. coli* (11). Plasmid DNA was then isolated from individual *E. coli* clones and digested to screen for cells containing a vector with an assembled 10-kb insert. One successful 10-kb assembly is represented (Fig. 2a). In general, at least one 10-kb assembled fragment could be obtained by screening 10 yeast clones. However, the rate of success varied from 10-100%. All of the first-stage intermediates were sequenced. Nineteen out of 111 assemblies contained errors. Alternate clones were selected, sequence-verified, and moved on to the next assembly stage (11).

Assembly of 100-kb synthetic intermediates. The pooled 10-kb assemblies and their respective cloning vectors were transformed into yeast as above to produce 100-kb assembly intermediates (11). Our results indicated that these products cannot be stably maintained in *E. coli* so recombined DNA had to be extracted from yeast. Multiplex PCR was performed on selected yeast clones (fig. S3 and table S2). Because every 10-kb assembly intermediate was represented by a primer pair in this analysis, the presence of all amplicons would suggest an assembled 100-kb intermediate. In general, 25% or more of the clones screened contained all of the amplicons expected for a complete assembly. One of these clones was selected for further screening. Circular plasmid DNA was extracted and sized on an agarose gel alongside a supercoiled marker. Successful second-stage assemblies with the vector sequence are approximately 105 kb in length (Fig. 2b). When all amplicons were produced following multiplex PCR, a second-stage assembly intermediate of the correct size was usually produced. In some cases, however, small deletions occurred. In other instances, multiple 10-kb fragments were assembled, which produced a larger second-stage assembly intermediate. Fortunately, these differences could easily be

detected on an agarose gel prior to complete genome assembly.

Complete genome assembly. In preparation for the final stage of assembly, it was necessary to isolate microgram quantities of each of the 11 second-stage assemblies (11). As reported (14), circular plasmids the size of our second-stage assemblies could be isolated from yeast spheroplasts after an alkaline-lysis procedure. To further purify the 11 assembly intermediates, they were exonuclease-treated and passed through an anion-exchange column. A small fraction of the total plasmid DNA (1/100th) was digested with Not I and analyzed by field-inversion gel electrophoresis (FIGE) (Fig. 2c). This method produced ~1 µg of each assembly per 400 ml yeast culture (~10¹¹ cells).

The method above does not completely remove all of the linear yeast chromosomal DNA, which we found could significantly decrease the yeast transformation and assembly efficiency. To further enrich for the eleven circular assembly intermediates, ~200 ng samples of each assembly were pooled and mixed with molten agarose. As the agarose solidifies, the fibers thread through and topologically “trap” circular DNA (15). Untrapped linear DNA can then be electrophoresed out of the agarose plug, thus enriching for the trapped circular molecules. The eleven circular assembly intermediates were digested with Not I so that the inserts could be released. Subsequently, the fragments were extracted from the agarose plug, analyzed by FIGE (Fig. 2d), and transformed into yeast spheroplasts (11). In this third and final stage of assembly, an additional vector sequence was not required since the yeast cloning elements were already present in assembly 811-900.

To screen for a complete genome, multiplex PCR was carried out with 11 primer pairs, designed to span each of the eleven 100-kb assembly junctions (table S3). Of 48 colonies screened, DNA extracted from one clone (sMmYcP235) produced all 11 amplicons. PCR of the wild type (WT) positive control (YcPmMyc1.1) produced an indistinguishable set of 11 amplicons (Fig. 3a). To further demonstrate the complete assembly of a synthetic *M. mycooides* genome, intact DNA was isolated from yeast in agarose plugs and subjected to two restriction analyses; Asc I and BssH II (11). Because these restriction sites are present in three of the four watermark sequences, this choice of digestion produces restriction patterns that are distinct from the natural *M. mycooides* genome (Figs. 1 and 3b). The sMmYcP235 clone produced the restriction pattern expected for a completely assembled synthetic genome (Fig. 3c).

pSynthetic genome transplantation

Additional agarose plugs used in the gel analysis above (Fig. 3c) were also used in genome transplantation experiments (11). Intact synthetic *M. mycooides* genomes from the

sMmYcP235 yeast clone were transplanted into restriction-minus *M. capricolum* recipient cells, as described (8). Results were scored by selecting for growth of blue colonies on SP4 medium containing tetracycline and X-gal at 37 °C. Genomes isolated from this yeast clone produced 5-15 tetracycline-resistant blue colonies per agarose plug. This was comparable to the YcPmMyc1.1 control. Recovery of colonies in all transplantation experiments was dependent on the presence of both *M. capricolum* recipient cells and an *M. mycooides* genome.

Semi-synthetic genome assembly and transplantation

To aid in testing the functionality of each 100-kb synthetic segment, semi-synthetic genomes were constructed and transplanted. By mixing natural pieces with synthetic ones, the successful construction of each synthetic 100-kb assembly could be verified without having to sequence these intermediates. We cloned 11 overlapping natural 100-kb assemblies in yeast by using a previously described method (16). In 11 parallel reactions, yeast cells were co-transformed with fragmented *M. mycooides* genomic DNA (YcPmMyc 1.1) that averaged ~100 kb in length and a PCR-amplified vector designed to overlap the ends of the 100-kb inserts. To maintain the appropriate overlaps so that natural and synthetic fragments could be recombined, the PCR-amplified vectors were produced via primers with the same 40-bp overlaps used to clone the 100-kb synthetic assemblies. The semi-synthetic genomes that were constructed contained between two and ten of the eleven 100-kb synthetic subassemblies (Table 1). The production of viable colonies produced after transplantation, ionfirmed that the synthetic fraction of each genome contained no lethal mutations. Only one of the 100-kb subassemblies, 811-900, was not viable.

Initially, an error-containing 811-820 clone was used to produce a synthetic genome that did not transplant. This was expected since the error was a single base pair deletion that creates a frameshift in *dnaA*, an essential gene for chromosomal replication. We were previously unaware of this mutation. By using a semi-synthetic genome construction strategy, we were able to pinpoint 811-900 as the source for failed synthetic transplantation experiments. Thus, we began to reassemble an error-free 811-900 assembly, which was used to produce the sMmYcP235 yeast strain. The *dnaA*-mutated genome only differs by one nucleotide from the synthetic genome in sMmYcP235. This genome served as a negative control in our transplantation experiments. The *dnaA* mutation was also repaired at the 811-900 level by genome engineering in yeast (17). A repaired 811-900 assembly was used in a final stage assembly to produce a yeast clone with a repaired genome. This yeast clone is named sMmYcP142 and could be transplanted. A complete list of genomes that

have been assembled from 11 pieces and successfully transplanted is provided in Table 1.

Characterization of the synthetic transplants

To rapidly distinguish the synthetic transplants from *M. capricolum* or natural *M. mycooides*, two analyses were performed. First, four primer pairs that are specific to each of the four watermarks were designed such that they produce four amplicons in a single multiplex PCR reaction (table S4). All four amplicons were produced by transplants generated from sMmYCp235, but not YCpMmyc1.1 (Fig. 4a). Second, the gel analysis with Asc I and BssH II, described above (Fig. 3d), was performed. The restriction pattern obtained was consistent with a transplant produced from a synthetic *M. mycooides* genome (Fig. 4b).

A single transplant originating from the sMmYCp235 synthetic genome was sequenced. We refer to this strain as *M. mycooides* JCVI-syn1.0. The sequence matched the intended design with the exception of the known polymorphisms, 8 new single nucleotide polymorphisms, an *E. coli* transposon insertion, and an 85-bp duplication (table S1). The transposon insertion exactly matches the size and sequence of IS1, a transposon in *E. coli*. It is likely that IS1 infected the 10-kb sub-assembly following its transfer to *E. coli*. The IS1 insert is flanked by direct repeats of *M. mycooides* sequence suggesting that it was inserted by a transposition mechanism. The 85-bp duplication is a result of a non-homologous end joining event, which was not detected in our sequence analysis at the 10-kb stage. These two insertions disrupt two genes that are evidently non-essential. We did not find any sequences in the synthetic genome that could be identified as belonging to *M. capricolum*. This indicates that there was a complete replacement of the *M. capricolum* genome by our synthetic genome during the transplant process.

The cells with only the synthetic genome are self replicating and capable of logarithmic growth. Scanning and transmission electron micrographs (EM) of *M. mycooides* JCVI-syn1.0 cells show small, ovoid cells surrounded by cytoplasmic membranes (Fig. 5c-5f). Proteomic analysis of *M. mycooides* JCVI-syn1.0 and the WT control (YCpMmyc1.1) by two-dimensional gel electrophoresis revealed almost identical patterns of protein spots (fig. S4) and these were clearly different from those previously reported for *M. capricolum* (10). Fourteen genes are deleted or disrupted in the *M. mycooides* JCVI-syn1.0 genome, however the rate of appearance of colonies on agar plates and the colony morphology are similar (compare Fig. 5a and b). We did observe slight differences in the growth rates in a color changing unit assay, with the JCVI-syn1.0 transplants growing slightly faster than the MmcyYCp1.1 control strain (fig. S6).

Discussion

In 1995, the quality standard for sequencing was considered to be one error in 10,000 bp and the sequencing of a microbial genome required months. Today, the accuracy is substantially higher. Genome coverage of 30-50X is not unusual, and sequencing only requires a few days. However, obtaining an error-free genome that could be transplanted into a recipient cell to create a new cell controlled only by the synthetic genome was complicated and required many quality control steps. Our success was thwarted for many weeks by a single base pair deletion in the essential gene *dnaA*. One wrong base out of over one million in an essential gene rendered the genome inactive, while major genome insertions and deletions in non-essential parts of the genome had no observable impact on viability. The demonstration that our synthetic genome gives rise to transplants with the characteristics of *M. mycooides* cells implies that the DNA sequence upon which it is based is accurate enough to specify a living cell with the appropriate properties.

Our synthetic genomic approach stands in sharp contrast to a variety of other approaches to genome engineering that modify natural genomes by introducing multiple insertions, substitutions, or deletions (18–22). This work provides a proof of principle for producing cells based upon genome sequences designed in the computer. DNA sequencing of a cellular genome allows storage of the genetic instructions for life as a digital file. The synthetic genome described in this paper has only limited modifications from the naturally occurring *M. mycooides* genome. However, the approach we have developed should be applicable to the synthesis and transplantation of more novel genomes as genome design progresses (23).

We refer to such a cell controlled by a genome assembled from chemically synthesized pieces of DNA as a “synthetic cell”, even though the cytoplasm of the recipient cell is not synthetic. Phenotypic effects of the recipient cytoplasm are diluted with protein turnover and as cells carrying only the transplanted genome replicate. Following transplantation and replication on a plate to form a colony (>30 divisions or >10⁹ fold dilution), progeny will not contain any protein molecules that were present in the original recipient cell (10, 24). This was previously demonstrated when we first described genome transplantation (10). The properties of the cells controlled by the assembled genome are expected to be the same as if the whole cell had been produced synthetically (the DNA software builds its own hardware).

The ability to produce synthetic cells renders it essential for researchers making synthetic DNA constructs and cells to clearly watermark their work to distinguish it from naturally occurring DNA and cells. We have watermarked the synthetic chromosome in this and our previous study (7).

If the methods described here can be generalized, design, synthesis, assembly, and transplantation of synthetic chromosomes will no longer be a barrier to the progress of synthetic biology. We expect that the cost of DNA synthesis will follow what has happened with DNA sequencing and continue to exponentially decrease. Lower synthesis costs combined with automation will enable broad applications for synthetic genomics.

We have been driving the ethical discussion concerning synthetic life from the earliest stages of this work (25, 26). As synthetic genomic applications expand, we anticipate that this work will continue to raise philosophical issues that have broad societal and ethical implications. We encourage the continued discourse.

References and Notes

1. F. Sanger *et al.*, *Nature* **265**, 687 (Feb 24, 1977).
2. R. D. Fleischmann *et al.*, *Science* **269**, 496 (Jul 28, 1995).
3. J. C. Venter, *Nature* **464**, 676 (Apr 1).
4. C. A. Hutchison *et al.*, *Science* **286**, 2165 (Dec 10, 1999).
5. J. I. Glass *et al.*, *Proc Natl Acad Sci U S A* **103**, 425 (Jan 10, 2006).
6. H. O. Smith, J. I. Glass, C. A. Hutchison III, J. C. Venter, in *Accessing Uncultivated Microorganisms: From the Environment to Organisms and Genomes and Back* K. Zengler, Ed. (ASM Press, Washington, 2008), pp. 320.
7. D. G. Gibson *et al.*, *Science* **319**, 1215 (Feb 29, 2008).
8. C. Lartigue *et al.*, *Science* **325**, 1693 (Sep 25, 2009).
9. G. A. Benders *et al.*, *Nucleic Acids Res*, (Mar 7, 2010).
10. C. Lartigue *et al.*, *Science* **317**, 632 (Aug 3, 2007).
11. Supplementary information is available on *Science Online*.
12. D. G. Gibson, *Nucleic Acids Res* **37**, 6984 (Nov, 2009).
13. D. G. Gibson *et al.*, *Proc Natl Acad Sci U S A* **105**, 20404 (Dec 23, 2008).
14. R. J. Devenish, C. S. Newlon, *Gene* **18**, 277 (Jun, 1982).
15. W. W. Dean, B. M. Dancis, C. A. Thomas, Jr., *Anal Biochem* **56**, 417 (Dec, 1973).
16. S. H. Leem *et al.*, *Nucleic Acids Res* **31**, e29 (Mar 15, 2003).
17. V. N. Noskov, T. H. Segall-Shapiro, R. Y. Chuang, *Nucleic Acids Res* **38**, 2570 (May 1).
18. M. Itaya, K. Tsuge, M. Koizumi, K. Fujita, *Proc Natl Acad Sci U S A* **102**, 15971 (Nov 1, 2005).
19. M. Itaya, *FEBS Lett* **362**, 257 (Apr 10, 1995).
20. H. Mizoguchi, H. Mori, T. Fujio, *Biotechnol Appl Biochem* **46**, 157 (Mar, 2007).
21. J. Y. Chun *et al.*, *Nucleic Acids Res* **35**, e40 (2007).
22. H. H. Wang *et al.*, *Nature* **460**, 894 (Aug 13, 2009).
23. A. S. Khalil, J. J. Collins, *Nat Rev Genet* **11**, 367 (May).
24. A mycoplasma cell, with a cell mass of about 10^{-13} g, contains fewer than 10^6 molecules of protein. (If it contains 20% protein this is 2×10^{-14} g protein per cell. At

a molecular weight of 120 Daltons per amino acid residue each cell contains $(2 \times 10^{-14})/120 = 1.7 \times 10^{-16}$ moles of peptide residues. This is $1.7 \times 10^{-16} \times 6 \times 10^{23} = 1 \times 10^8$ residues per cell. If the average size of a protein is 300 residues then a cell contains about 3×10^5 protein molecules.) After 20 cell divisions the number of progeny exceeds the total number of protein molecules present in the recipient cell. So, following transplantation and replication to form a colony on a plate, most cells will contain no protein molecules that were present in the original recipient cell.

25. M. K. Cho, D. Magnus, A. L. Caplan, D. McGee, *Science* **286**, 2087 (Dec 10, 1999).
26. M. S. Garfinkel, D. Endy, G. E. Epstein, R. M. Friedman. (2007).
27. D. G. Gibson *et al.*, *Nat Methods* **6**, 343 (May, 2009).
28. We thank Synthetic Genomics, Inc. for generous funding of this work. We thank J. B. Hostetler, D. Radune, N. B. Fedorova, M. D. Kim, B. J. Szczypinski, I. K. Singh, J. R. Miller, S. Kaushal, R. M. Friedman, and J. Mulligan for their contributions to this work. Electron micrographs were generously provided by T. Deerinck and M. Ellisman of the National Center for Microscopy and Imaging Research at the University of California at San Diego. J.C.V. is Chief Executive Officer and Co-Chief Scientific Officer of SGI. H.O.S. is Co-Chief Scientific Officer and on the Board of Directors of SGI. C.A.H. is Chairman of the SGI Scientific Advisory Board. All three of these authors and JCVI hold SGI stock. JCVI has filed patent applications on some of the techniques described in this paper.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1190719/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S7
References

9 April 2010; accepted 13 May 2010

Published online 20 May 2010; 10.1126/science.1190719

Include this information when citing this paper.

Fig. 1. The assembly of a synthetic *M. mycoides* genome in yeast. A synthetic *M. mycoides* genome was assembled from 1,078 overlapping DNA cassettes in three steps. In the first step, 1,080-bp cassettes (orange arrows), produced from overlapping synthetic oligonucleotides, were recombined in sets of 10 to produce one hundred nine ~10-kb assemblies (blue arrows). These were then recombined in sets of 10 to produce eleven ~100-kb assemblies (green arrows). In the final stage of assembly, these eleven fragments were recombined into the complete genome (red circle). With the

exception of 2 constructs that were enzymatically pieced together *in vitro* (27) (white arrows), assemblies were carried out by *in vivo* homologous recombination in yeast. Major variations from the natural genome are shown as yellow circles. These include 4 watermarked regions (WM1-WM4), a 4-kb region that was intentionally deleted (94D), and elements for growth in yeast and genome transplantation. In addition, there are 20 locations with nucleotide polymorphisms (asterisks). Coordinates of the genome are relative to the first nucleotide of the natural *M. mycoides* sequence. The designed sequence is 1,077,947 bp. The locations of the Asc I and BssH II restriction sites are shown. Cassettes 1 and 800-810 were unnecessary and removed from the assembly strategy (11). Cassette 2 overlaps cassette 1104 and cassette 799 overlaps cassette 811.

Fig. 2. Analysis of the assembly intermediates. (a) Not I and Sbf I double restriction digestion analysis of assembly 341-350 purified from *E. coli*. These restriction enzymes release the vector fragments (5.5 kb and 3.4 kb) from the 10-kb insert. Insert DNA was separated from the vector DNA on a 0.8% E-gel (Invitrogen). M indicates the 1-kb DNA ladder (New England Biolabs; NEB). (b) Analysis of assembly 501-600 purified from yeast. The 105-kb circles (100-kb insert plus 5-kb vector) were separated from the linear yeast chromosomal DNA on a 1% agarose gel by applying 4.5 V/cm for 3 hours. S indicates the BAC-Tracker supercoiled DNA ladder (Epicentre). (c) Not I restriction digestion analysis of the eleven ~100-kb assemblies purified from yeast. These DNA fragments were analyzed by FIGE on a 1% agarose gel. The expected insert size for each assembly is indicated. λ indicates the lambda ladder (NEB). (d) Analysis of the 11 pooled assemblies shown in (c) following topological trapping of the circular DNA and Not I digestion. One fortieth of the DNA used to transform yeast is represented.

Fig. 3. Characterization of the synthetic genome isolated from yeast. (a) Yeast clones containing a completely assembled synthetic genome were screened by multiplex PCR with a primer set that produces 11 amplicons; one at each of the 11 assembly junctions. Yeast clone sMmYCp235 (235) produced the 11 PCR products expected for a complete genome assembly. For comparison, the natural genome extracted from yeast (WT) was also analyzed. PCR products were separated on a 2% E-gel (Invitrogen). L indicates the 100-bp ladder (NEB). (b) The sizes of the expected Asc I and BssH II restriction fragments for natural (WT) and synthetic (Syn235) *M. mycoides* genomes. (c) Natural (WT) and synthetic (235) *M. mycoides* genomes were isolated from yeast in agarose plugs. In addition, DNA was purified from the host strain alone (H). Agarose plugs were digested with Asc I or BssH II and fragments were separated by clamped

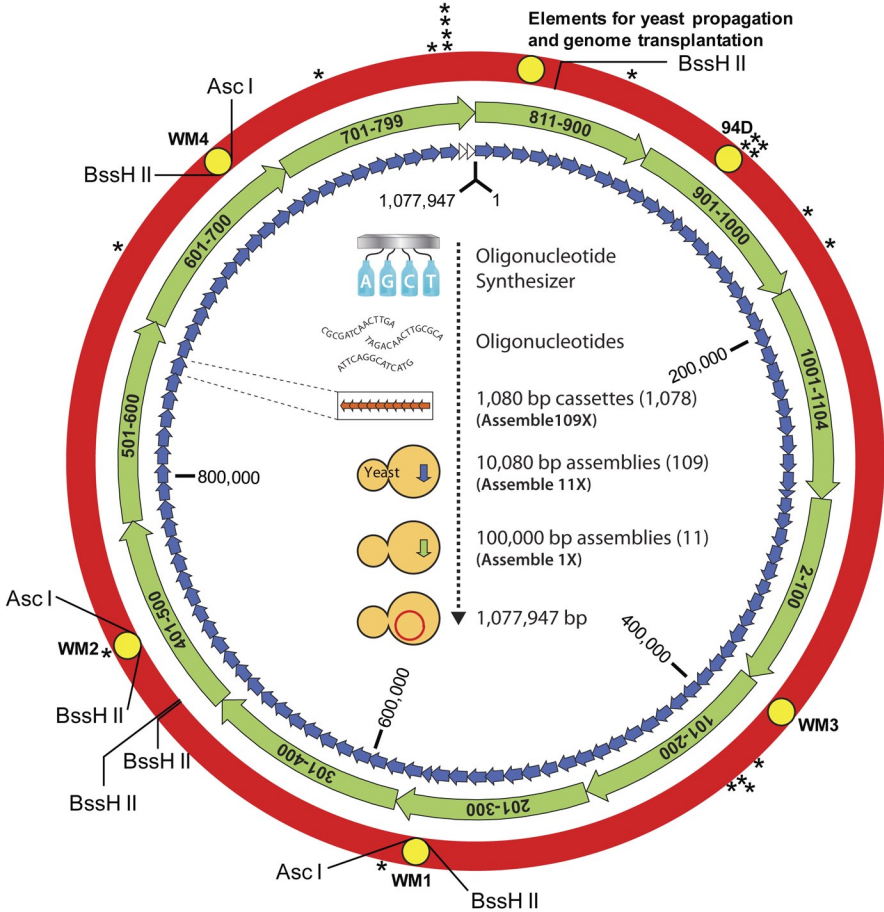
homogeneous electrical field (CHEF) gel electrophoresis. Restriction fragments corresponding to the correct sizes are indicated by the fragment numbers shown in (b).

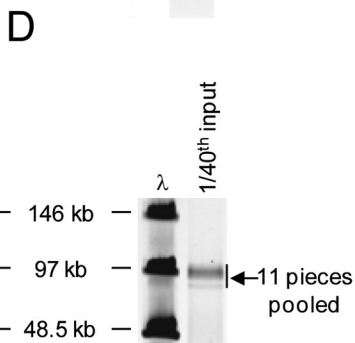
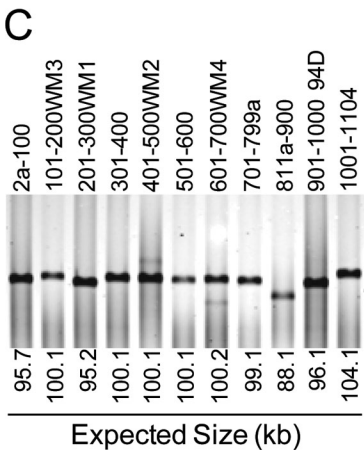
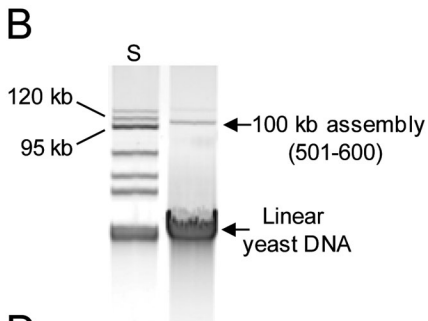
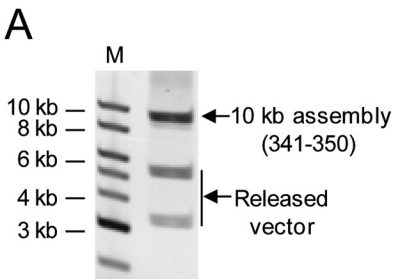
Fig. 4. Characterization of the transplants. (a) Transplants containing a synthetic genome were screened by multiplex PCR with a primer set that produces 4 amplicons; one internal to each of the four watermarks. One transplant (syn1.0) originating from yeast clone sMmYCp235 was analyzed alongside a natural, non-synthetic genome (WT) transplanted out of yeast. The transplant containing the synthetic genome produced the 4 PCR products whereas the WT genome did not produce any. PCR products were separated on a 2% E-gel (Invitrogen). (b) Natural (WT) and synthetic (syn1.0) *M. mycoides* genomes were isolated from *M. mycoides* transplants in agarose plugs. Agarose plugs were digested with Asc I or BssH II and fragments were separated by CHEF gel electrophoresis. Restriction fragments corresponding to the correct sizes are indicated by the fragment numbers shown in Fig. 3b.

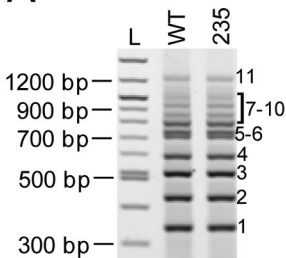
Fig. 5. Images of *M. mycoides* JCVI-syn1.0 and WT *M. mycoides*. To compare the phenotype of the JCVI-syn1.0 and non-YCp WT strains, we examined colony morphology by plating cells on SP4 agar plates containing X-gal. Three days after plating, the JCVI-syn1.0 colonies are blue because the cells contain the *lacZ* gene and express beta-galactosidase, which converts the X-gal to a blue compound (a). The WT cells do not contain *lacZ* and remain white (b). Both cell types have the fried egg colony morphology characteristic of most mycoplasmas. EMs were made of the JCVI-syn1.0 isolate using two methods. (c) For scanning EM, samples were post-fixed in osmium tetroxide, dehydrated and critical point dried with CO₂, and visualized using a Hitachi SU6600 SEM at 2.0 keV. (d) Negatively stained transmission EMs of dividing cells using 1% uranyl acetate on pure carbon substrate visualized using JEOL 1200EX CTEM at 80 keV. To examine cell morphology, we compared uranyl acetate stained EMs of *M. mycoides* JCVI-syn1.0 cells (e) with EMs of WT cells made in 2006 that were stained with ammonium molybdate (f). Both cell types show the same ovoid morphology and general appearance. EMs were provided by Tom Deerinck and Mark Ellisman of the National Center for Microscopy and Imaging Research at the University of California at San Diego.

Table 1. Genomes that have been assembled from 11 pieces and successfully transplanted. Assembly 2-100 = 1, assembly 101-200 = 2, assembly 201-300 = 3, assembly 301-400 = 4, assembly 401-500 = 5, assembly 501-600 = 6, assembly 601-700 = 7, assembly 701-799 = 8, assembly 811-900 = 9, assembly 901-1000 = 10, assembly 1001-1104 = 11. WM indicates watermarked assembly.

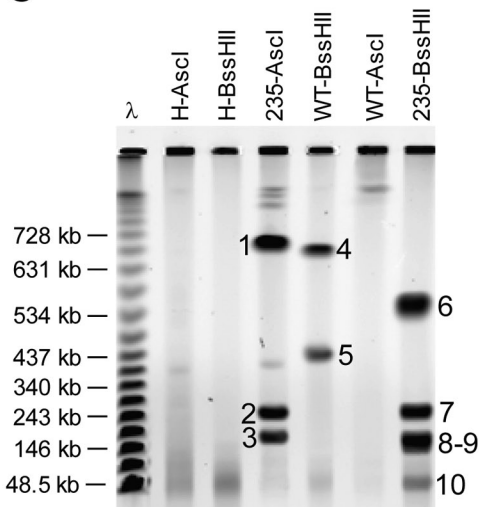
| Genome Assembly | Synthetic Fragments | Natural Fragments |
|---|---------------------------------------|-------------------|
| Reconstituted natural genome | None | 1-11 |
| 2/11 semi-synthetic genome with 1 watermark | 5WM, 10 | 1-4, 6-9, 11 |
| 8/11 semi-synthetic genome without watermarks | 1-4, 6-8, 11 | 5, 9, 10 |
| 9/11 semi-synthetic genome without watermarks | 1-4, 6-8, 10-11 | 5, 9 |
| 9/11 semi-synthetic genome with 3 watermarks | 1, 2WM, 3WM, 4, 6, 7WM, 8, 10-11 | 5, 9 |
| 10/11 semi-synthetic genome with 3 watermarks | 1, 2WM, 3WM, 4, 5WM, 6, 7WM, 8, 10-11 | 9 |
| 11/11 synthetic genome, 811-820 correction of <i>dnaA</i> | 1, 2WM, 3WM, 4, 5WM, 6, 7WM, 8, 9-11 | None |
| 11/11 synthetic genome, 811-900 correction of <i>dnaA</i> | 1, 2WM, 3WM, 4, 5WM, 6, 7WM, 8, 9-11 | None |

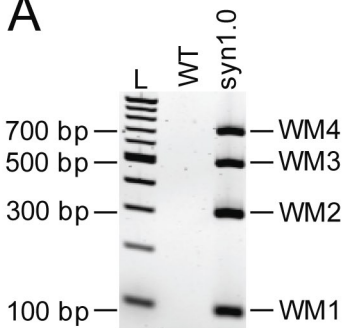




A**B**

| Strain | Digest | Fragment # and size (kb) |
|--------|--------|--|
| WT | Ascl | No sites |
| WT | BssHII | (4) 668 (5) 419 |
| Syn235 | Ascl | (1) 685 (2) 233 (3) 160 |
| Syn235 | BssHII | (6) 533 (7) 233 (8) 152 (9) 126 (10) 34 |

C

A**B**