www.sciencemag.org/cgi/content/full/science.1190719/DC1

Supporting Online Material for

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

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

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

Published 20 May 2010 on *Science* Express
DOI: 10.1126/science.1190719

This PDF file includes:

Supporting Online Material

MATERIALS AND METHODS

Design of cassettes for assembly of the synthetic *Mycoplasma mycoides* genome

The synthetic *M. mycoides* genome described in this paper was constructed from 1,078 overlapping synthetic DNA cassettes. A Not I cleavage site (GCGGCCGC) is included at the beginning and the end of each cassette sequence. Our original intention was to build the complete sequence of naturally occurring *M. mycoides*, with several watermark sequences added at places where insertion of additional sequence would not interfere with viability. During construction of the synthetic genome we learned that certain sequences could be deleted without affecting viability, and some of these deletions were incorporated into the synthetic genome design. The following describes the process of designing the genome, and the cassettes to assemble it, in some detail.

Design of the synthetic *M. mycoides* genome was based on sequences of two laboratory strains of *M. mycoides* determined by our group:

GenBank accession CP001621 – This is the sequence of the *M. mycoides* strain used as the genome donor in our first paper on genome transplantation (*1*), with a length of 1,089,202 bp.

GenBank accession CP001668 – This is the sequence of an *M. mycoides* strain engineered in yeast by deleting the gene for a Type III restriction endonuclease, then transplanted to produce a mycoplasma (*2*). This sequence has a length of 1,084,586 bp.

Design of our synthetic *M. mycoides* genome was begun using a draft genome sequence from the project that produced the above sequence CP001621, with a length of 1,114,292 bp. This draft sequence is referred to as mmycDRAFT, and incorrectly contains a large duplication. First the mmycDRAFT sequence was divided into cassettes 1080 bp in length, with 80-bp overlaps, and a Not I restriction site was added at each end, as follows:
mmyc0 = bp 1-1080 preceded and followed by GCGGCCGC
mmyc1 = bp 1001-2080 preceded and followed by GCGGCCGC
mmyc2 = bp 2001-3080 preceded and followed by GCGGCCGC
.
.
mmyc1113 = bp 1113001-1114080 preceded and followed by GCGGCCGC
mmyc1114 = bp 1114001-1114292 preceded and followed by GCGGCCGC

Designed cassettes located within parts of the mmycDRAFT sequence that were considered reliable at that time were ordered for synthesis by Blue Heron. This initial order of 1,072 cassettes included all 1,115 designed cassettes except mmyc0-mmyc14 (15 cassettes), mmyc835-mmyc852 (18 cassettes), and mmyc1105-mmyc1114 (10 cassettes).

Assembly of these 1,072 cassettes would yield two contiguous stretches of DNA, with two gaps corresponding to regions where the sequence was uncertain.

Design of the remaining cassettes was based on sequence CP001668, after it was completely finished, because that genome was known to be stably maintained in yeast and to be transplantable from yeast to produce a viable mycoplasma. First cassettes were designed to fill the two largest gaps. "a" designates cassettes that were not part of the initial order from Blue Heron.

Cassettes to fill gap 1. Cassettes mmyc835a-mmyc850a produce a sequence that matches CP001668 and fill the gap left by cassettes mmyc835-mmyc852, which were omitted from the initial Blue Heron order.

Cassettes to fill gap 2. Cassettes mmyc2a-mmyc12ax produce a sequence that matches CP001668 and fill the gap left by cassettes mmyc1105-mmyc1114 and mmyc0-mmyc14, which were omitted from the initial Blue Heron order. This sequence is considerably shorter than the corresponding region of the draft sequence, mainly due to problems with assembly of that sequence. The sequence of mmyc12ax is longer than most, hence "x" for eXtended.

Cassettes to make the synthetic genome match CP001668 near the origin of replication. There were many differences between the mmycDRAFT sequence and CP001668 in this region.

mmyc799a (overlaps mmyc798 by 80 bp at one terminus and spans to bp 1,084,586 (the end of CP001668)
mmyc811a (bp 80 to 1074 of CP001668)
mmyc812a (bp 995 to 2074 of CP001668)

Cassettes synthesized to match CP001668

mmyc56a (344,335 to 345,333 of CP001668)
mmyc58a (345,252 to 345,926 of CP001668)
mmyc938.1a (lies between mmyc938 and mmyc939)

Cassettes "fixed" by oligonucleotide mutagenesis to match CP001668. "f" cassettes had small numbers of sequence differences that could be conveniently fixed by oligonucleotide mutagenesis to exactly match CP001668. See below for details on how these cassettes were fixed. The following cassettes were fixed: mmyc1011f, mmyc1028f, mmyc247f, mmyc248f, mmyc342f, mmyc399f, mmyc400f, mmyc528f, mmyc529f, mmyc578f, mmyc579f, mmyc632f, mmyc642f , mmyc759f, and mmyc874f. In addition, for constructing genomes with the "94D" deletion, which is missing mmyc936-mmyc939, cassette mmyc940f was produced to contain a 5′ overlap with cassette mmyc935. This region was demonstrated to be dispensable by deleting it from the natural genome in yeast, then transplanting to produce a viable mycoplasma.

Cassettes containing watermark sequences.
mmycWM1: replaces mmyc282-287 (5′ 80-bp overlap from mmyc282; 3′ 80-bp overlap from mmyc287)
mmycWM2b: replaces mmyc447 (both 80-bp overlaps are from mmyc447)

mmycWM3: replaces mmyc106 (both 80-bp overlaps are from mmyc106)
mmycWM4: replaces mmyc680 (both 80-bp overlaps are from mmyc680)

Nucleotide 1 is at the initiation codon of the dnaA gene. Both *M. mycoides* sequences CP001621 and CP001668 use a numbering system in which nucleotide 1 is at the initiation codon of the dnaA gene, in the vicinity of the DNA replication origin. However, the mmycDRAFT sequence used for the initial cassette design was numbered from a different origin (but, mercifully, in the same direction). Consequently, nucleotide 1 of the synthetic genome sequence is located in cassette mmyc811a.

Differences between cassette designs and CP001668 that were not fixed. There are 19 differences between cassette designs and the CP001668 sequence that appeared harmless, which we opted not to fix. These were all differences in lengths of homopolymer or dinucleotide runs and were located between annotated genes. However, some may really lie within genes for which the start sites were incorrectly annotated. Also, some may affect the spacing between the -35 and -10 sequences of promoters, and so could affect gene expression. Table S1 gives details of these 19 differences. They provide more sequence polymorphisms to distinguish the synthetic genome from the natural *M. mycoides* genome that we have cloned in yeast, in addition to the watermark sequences.

DNA cassette sequence corrections
The synthetic cassettes comprising the synthetic genome were ordered based on an imperfect draft sequence. Because of this, there were small differences between the ordered synthetic cassette sequences and the completed *M. mycoides* subspecies *capri* genome sequence. Forty-three cassettes contained differences. As mentioned above, many of the differences were small insertions or deletions in homopolymer or dinucleotide runs not predicted to be located in genes. These changes were thought to be benign and were not fixed. Sixteen cassettes were fixed. The differences were all single base changes except for a deletion of a 12-bp repeat unit in one cassette.

Two strategies were employed to change the synthetic cassette sequences to match the completed *M. mycoides* subspecies *capri* genome sequence CP001668. All fixes were performed on the 1,080-bp cassette clones. The QuikChange II site directed mutagenesis kit (Stratagene) was used for some of the single base changes. The remainder of the changes was performed by a combination of PCR and in vitro recombination (Fig. S5) (*3*).

Primers BH pUC bckbn For1 (5′-CGTCAAAGCAACCATAGTACGCGCCCTGTAG-3′) and BH pUC bckbn Rev1 (5′-CTGACTCGCTGCGCTCGGTCGTTCGGC-3′) were used to amplify the plasmid backbone of a cassette clone. This amplicon (BH backbone) is ~ 2,600 bp and contains the ampicillin resistance marker and an origin of replication. Primers BH insert For1 (5′-GCCGAACGACCGAGCGCAGCGAGTCAG-3′) and BH insert Rev1 (5′-CTACAGGGCGCGTACTATGGTTGCTTTGACG-3′) are the reverse complements of BH pUC bckbn For1 and BH pUC bckbn Rev1 and are used as the forward and reverse primers with the correction primers to amplify the cassette insert and create homology regions with the BH backbone for in vitro recombination.

To change a single nucleotide in a cassette, two oligonucleotides are required. The first contains the corrected nucleotide sequence flanked by approximately 20-25 bases and the second is the reverse complement of the first. Examples of oligonucleotides used to change a nucleotide to a G are as follows, with the nucleotide to be changed in bold : cassette Fix For1 (5′-GAACTGAAAATTATATATCAGGTA**G**ATATGAATAGGAAATAGTATGTC-3′) and cassette Fix Rev1 (5′-GACATACTATTTCCTATTCATAT**C**TACCTGATATATAATTTTCAGTTC-3′). PCR with these primers is used to change the selected base. Primers BH insert For1 and Cassette Fix Rev1 are used to amplify part of a corrected insert and primers BH insert Rev1 and Cassette Fix For1 are for amplifying the remaining insert also with the correction. Because the Cassette Fix For1 and Cassette Fix Rev1 primers are reverse complements of each other, this creates a homology region for in vitro recombination between the amplicons. As mentioned previously, the BH insert primers create homology with the BH backbone piece, allowing a three piece in vitro recombination reaction to create the corrected insert plasmid.

One synthetic cassette contained five repeats of a 12-bp sequence, whereas the completed genome sequence showed six repeats of this sequence. PCR and in vitro recombination were used as above to correct the 12-bp deletion with minor changes. The forward primer used to add the 12-bp repeat unit contained four repeat units on the 5′ end followed by 27 bases of non-repeat sequence on the 3′ end. This primer was used with primer BH insert Rev1 to create an amplicon with four repeat units at the 5′ end. The reverse primer used to add the 12-bp repeat unit contained the reverse complement of four repeat units on the 5′ end followed by 27 bases of non-repeat sequence on the 3′ end. This primer was used with primer BH insert For1 to create an amplicon with four repeat units at the 3′ end. Following in vitro recombination, clones were sequenced to identify the clones that recombined between two repeat units on the 3′ end and two units on the 5′ end of the amplicons, resulting in a cassette insert sequence with six 12-bp repeat units.

A further change was made in cassette 940 that was not related to differences between the synthetic and native sequences. Cassettes 936-939 were difficult to assemble. Further work demonstrated that cassettes 936-939 were not essential for the viability of *M. mycoides* subspecies *capri* and could be deleted. To delete this region during construction of the synthetic genome, the 5′ 80-bp overlap of cassette 940 was changed to match the 80-bp 3′ overlap of cassette 935. Upon construction, cassette 935 would then join with cassette 940, deleting cassettes 936-939. To change the overlap region, cassette 940 was amplified with BH insert Rev1 and a forward primer that binds adjacent to the 5′ 80-bp overlap region. The 5′ overlap from cassette 936 (which is identical to the 3′ 80-bp overlap of cassette 935) was obtained by amplifying cassette 936 with BH insert For1 and a reverse primer that binds the 3′ 40-bp end of the overlap with a 5′ extension that is the complement of the forward primer used to amplify cassette 940. In vitro recombination was then used with the BH backbone piece to generate cassette 940 with a 5′ 80-bp overlap to cassette 935.

Assembly vector preparation

Polymerase chain reaction (PCR) amplification was used to produce a unique vector for the cloning of each assembly. The amplified vectors contain terminal overlaps to the ends of the assembly. The strategy for assembly vector preparation has previously been described (*4*). Each PCR primer includes a 20-bp overlap with one end of the vector, a Not I restriction site, and a 40-bp overlap with one end of the cassette assembly. For the first stage of assembly, a yeast/ *Escherichia coli* shuttle vector, termed pCC1BAC-LCYEAST, was produced for template DNA in PCR. This vector was constructed by in vitro assembly (*3*) of Afe I-digested pCC1BAC (Epicentre) with a PCR product consisting of 40-bp overlaps to the Afe I restriction fragment, a histidine selectable marker, a centromere, and an origin of replication (derived from pRS313 (*5*)). For the second stage of assembly, pRS314 (*5*) was used as template DNA in PCR (with the exception of the assembly of 831-840, which used pCC1BAC as template DNA). Unique assembly vectors were generated by PCR using the Phusion Hot Start High-Fidelity DNA polymerase with HF buffer (New England Biolabs; NEB) according to the manufacturer's instructions except reactions were supplemented with 1 mM additional $MgCl_2$ and products were annealed at 60 °C and extended for 1 min per kilobase. PCR products were extracted from agarose gels after electrophoresis and purified using the QIAquick Gel Extraction kit (Qiagen) according to the manufacturer's instructions. Although we could have used the same second-stage vectors for cloning natural and synthetic fragments, an additional vector sequence was constructed for cloning natural 100-kb fragments. This vector sequence is named pCC1BAC-URA and was construct in the same way that pCC1BAC-LCYEAST was constructed except pRS316 was used instead of pRS313. The primers used to produce the PCR-amplified assembly vectors are listed in Tables S5-S7.

Preparation of cassette DNA for yeast transformation to produce 10-kb assemblies

Approximately 500 ng uncut cassette DNA was pooled in sets of 10, digested with Not I, and electrophoresed on a 1% low-melting point agarose gel for 90 min at 96 V. One-kb DNA fragments were cut from the gels and the mass of the gel slice was measured. A 1:10 solution of 10X TAE buffer containing 3M sodium acetate was added to the gel slice, and the agarose gel was melted at 65 °C For 10 min. Following incubation at 42 °C for 15 min, β-agarase (NEB) was added 1:50 and the incubation was continued for 1 h longer. Following a phenol extraction and ethanol precipitation (in the presence of 1 µl glycoblue (Ambion), DNA was resuspended in 40 µl TE (pH 8.0). Ten microliters was used in transformation experiments.

Yeast 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 (*6*) with one modification: cells were grown to an OD_{600} of 0.5 (~10^7 cells/ml) prior to the preparation of yeast spheroplasts. We have found this optical density to be optimal for the assembly of multiple overlapping fragments in yeast. The Not I-digested fragments were pooled, mixed with 40 ng unique PCR-amplified assembly vector (except for the final stage of assembly), and then added to ~2 X 10^8 yeast spheroplasts. After transformation, yeast spheroplasts were regenerated and selected on complete supplemental medium without histidine (CSM-His; 10-kb assemblies, 811-900, and complete genome) or without tryptophan (CSM-Trp; 100-kb assemblies with the exception of 811-900) and 1M sorbitol agar plates for 3 days at 30

ºC. Primary transformants were then transferred onto selective plates as small patches and incubated overnight at 30 °C.

10-kb assembly screening
DNA was extracted from ~10^7 yeast cells (from patches) using a QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions but with one modification: buffer P1 was supplemented with 1:1000 β-mercaptoethanol and 1:100 20mg/ml Zymolyase-100T (USB), and cells were incubated at 37 ºC for 30-60 min prior to the addition of buffer P2. Samples (up to 3 µl) of the yeast extracted DNA were transformed into 30 µl EPI300 (Epicentre) electrocompetent *E. coli* cells in a 1-mm cuvette (BioRad) at 1,200 V, 25 µF and 200 Ω using a Gene Pulser Xcell electroporation system (BioRad). Cells were allowed to recover at 37 °C for 1.5 h in 1 ml SOC medium, then plated onto LB medium containing 12.5 µg/ml chloramphenicol. After incubation at 37 °C for 16-24 h, individual colonies were selected and grown in 3 ml LB medium with 12.5 µg/ml chloramphenicol overnight at 37 °C. DNA was prepared from these cells by alkaline lysis using the P1, P2 and P3 buffers (Qiagen), followed by isopropanol precipitation. DNA pellets were dissolved in TE buffer (pH 8.0) containing RNase A and RNase T1 (Ambion). Alternatively, DNA was prepared from the QIAprep Spin Miniprep Kit according to the instructions provided. Following purification, DNA was digested with Not I (and occasionally Sbf I) to release the insert from the vector and sized by gel electrophoresis on 0.8% E-gels (Invitrogen) for 30-60 min. Bands were visualized using a GE Typhoon 9410 Imager.

10-kb DNA extraction and preparation for transformation to produce 100-kb assemblies
Positive clones were propagated in 10 ml LB medium containing 12.5 µg/ml chloramphenicol and 1:1000 induction solution (Epicentre) and incubated overnight at 37 °C. DNA was purified from these cultures using the QIAprep Spin Miniprep Kit according to the manufacturer's instructions. Assembly 211-220 was unstable in the Epi300 strain so was transferred to the Stbl4 strain (Invitrogen), where it could be stably maintained. This assembly cannot be induced to higher copy levels in this strain and was not column-purified as above. Instead, this clone was propagated in 50 ml LB medium containing 12.5 µg/ml chloramphenicol. After cell lysis with sodium hydroxide and sodium dodecyl sulfate, followed by neutralization with potassium acetate, these DNA molecules were centrifuged and then precipitated with isopropanol. DNA pellets were dissolved in TE buffer (pH 8.0) then RNase treated, phenol-chloroform extracted, and ethanol precipitated. DNA pellets were dissolved in TE buffer. For each 10-kb assembly, Not I-digested DNA was quantified by gel electrophoresis alongside known DNA standards. Approximately 125 ng uncut DNA was pooled in sets of 10, digested with Not I (and usually also with Sbf I to provide better separation between the vector bands and insert) and electrophoresed on a 1% low-melting point agarose gel for 90 min at 96 V. Ten-kb DNA fragments were cut from the gels and extracted following β–agarase treatment as described above. DNA was resuspended in 20 µl TE (pH 8.0). Ten microliters was used in transformation experiments with 40 ng vector.

100-kb assembly multiplex PCR screening
DNA was extracted from patches as above. Multiplex PCR was performed using a Qiagen Multiplex PCR Kit. A 1/50 volume (1 µl) of the DNA extract and 1 µl of a 10X primer stock containing 20 oligos at 2.5-5 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, 57-60 °C for 90 s, and 72 °C for 90s, 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 a Typhoon 9410 Imager.

Small-scale isolation of 100-kb supercoiled assemblies from yeast and analysis

Supercoiled, circular assemblies were isolated from a 5 ml saturated yeast culture grown in selective medium. The procedure is based on a previously described method (*7*) but scaled down and without emphasis on removing yeast chromosomal DNA. Harvested yeast cells were transferred to a microfuge tube with 1 ml water, and resuspended in 1 ml Pretreatment Buffer (1.2M sorbitol, 200mM Tris-Cl, 100mM EDTA, pH 9.1) with 8 µl 14 M β-mercaptoethanol. Following a 10 min incubation at room temperature, cells were harvested, washed twice with 1 ml SCE (1M sorbitol, 60mM EDTA,100mM sodium citrate, pH 5.75), and resuspended in 1 ml SCE plus 10 µl Zymolyase-100T solution (20 mg/ml Zymolyase-100T [USB], 50% glycerol, 2.5% glucose, 50 mM Tris-Cl, pH 8.0). Following a 1 hour incubation at 37 °C, spheroplasts were harvested and then resuspended in 25 µl Tris/sucrose buffer (50mM Tris-Cl pH 8.0, 25% sucrose) and 20 µl proteinase K solution (10mg/ml proteinase K [Sigma], 1 mM calcium chloride, 50 mM Tris-Cl pH 8.0). Next, 475 µl Lysis Buffer (20mM EDTA, 50mM Tris-Cl, 1% SDS, pH 12.45) was added and mixed by pipetting up and down. Following a 30 min incubation at 37 °C, 100 µl 2M Tris-Cl pH 7.0 was added and mixed. Next, 100 µl 4M NaCl was added and mixed. Following a 30 min incubation at room temperature, 70 µl 3M sodium acetate was added and mixed. A standard phenol-chloroform extraction, isopropanol precipitation, and 70% ethanol wash was carried out and DNA pellets were resuspended in 100 µl TE pH 8.0 containing RNase A. Following a one hour incubation at 37 °C, 10 µl DNA was loaded onto a 1% agarose gel in 1X TAE buffer and electrophoresis was carried out for 3 hours at 4.5V/cm. After electrophoresis, the gel was stained with SYBR Gold and scanned with a Typhoon 9410 imager.

Large-scale isolation of 100-kb assemblies from yeast and analysis

A pre-culture (5 to 10 ml) of each VL6-48N strain harboring one of the eleven 100-kb assemblies was grown overnight to saturation in selective medium and then inoculated into 400 ml of selective medium. Once the culture reached an OD_{600} of 1.5, cells were harvested (2,205 rcf, 3 min). Pellets were resuspended in 100 ml water and transferred to two 50 ml conical tubes and harvested as above. Cell pellets were resuspended in 40 ml SPE (1 M Sorbitol, 10 mM Na_2EDTA, 0.01 M Na phosphate, pH 7.5) containing 0.4 ml Zymolyase-20T (10 mg/ml) and 40 µl 14 M β-mercaptoethanol. The cell suspension was incubated at 37 ºC for 60 min. Spheroplasts were harvested (1,125 rcf for 5 min) and resuspended in 1 ml 1 M Sorbitol, and then 20 ml of lysis Buffer (0.05 M Tris-HCl, 0.02 M EDTA, 1% SDS, pH 12.8) was added and the tubes were inverted 10 times. Cell lysates were incubated at 37 ºC for 30 min and then extracted with 20 ml phenol/chloroform/isoamyl-alcohol (Invitrogen) by 20 gentle inversions and centrifugation at 3,645 rcf for 20 min. The aqueous phase was transferred to a new 50 ml conical tube. DNA precipitation was carried out by adding 2 ml 3M NaAc (pH 5.2) and 20 ml of isopropanol, followed by centrifugation at 3,645 rcf for 30 min at 4 ºC. The pellets were resuspended in 1 ml TE (pH 8.0) containing RNase A (30 µg/ml) and incubated at 37 ºC for 30 min. At this point, the two DNA samples were combined. The circular DNA was further purified by the Qiagen Large-Construct Kit, according to the manufacturer's instructions with

some minor modifications. Ten ml of EX Buffer was mixed with the DNA solution, followed by the addition of 200 µl exonuclease and 300 µl 100 mM ATP. After a 45 min incubation at 37 ºC, the reaction was stopped by adding 12 ml of QS buffer. This solution was then applied to the Qiagen-tip 500 column. The column was washed with 30 ml of QF buffer, and DNA was eluted with 12 ml QF buffer, pre-warmed at 55 ºC. DNA was then precipitated overnight at -20 ºC by the addition of 2 volumes of ethanol, in the presence of 1.2 ml 3M NaAc (pH 5.2), 15 ul of GlycoBlue, and 15 ul of yeast total tRNA (Sigma). Alternatively, DNA was precipitated by the addition of 1 volume isopropanol. The precipitated DNA was recovered by centrifugation at 3,645 rcf for 1 hr at 4 ºC. DNA pellets were washed with 70% ethanol and resuspended in 150 µl TE (pH 8.0). A sample (0.75 µl) of each was digested with Not I and analyzed by FIGE. The FIGE analysis was performed on a 1% agarose gel (BioRad, catalog # 161-3016) in 1X TAE buffer without circulation and the parameters were forward 90 V, initial switch 5.0 sec, final switch 30 sec, with linear ramp, and reverse 60 V, initial switch 5.0 sec, final switch 30 sec, with linear ramp. After electrophoresis, the gel was stained with SYBR Gold and scanned with a Typhoon 9410 imager.

Topological trapping and analysis

Twenty microliters of each uncut 100-kb assembly were pooled and equilibrated to 50 °C. One volume (220 µl) of 2% low melting point agarose, also equilibrated to 50 °C, was mixed with the pooled DNA. Approximately 85 µl of this mixture was added to agarose plug molds (Bio-Rad), which were kept cold on ice. Following a 30 min incubation on ice, agarose plugs were added to the wells of a 1% agarose gel (1X TAE buffer) and electrophoresis was carried out at 4.5 V/cm for 2 hours. Plugs were removed from the agarose gel and washed by inverting in 5 ml 0.1X Wash Buffer (Bio-Rad CHEF Genomic DNA Plug Kit) for 1 hour at room temperature. This buffer was removed and 5 ml 1X Buffer 3 (NEB) was added. Following a 1 hour incubation/inversion at room temperature, the buffer was removed and fresh 1X Buffer 3 (2.5 ml) with 250 units Not I was added. Not I digestion was carried out by incubating the agarose plugs overnight at 37 °C. Agarose plugs were then inverted in 5 ml 1X TAE/0.3M sodium acetate solution for 1 hour, and then moved to a microfuge tube. A solution of 10X TAE buffer containing 3M sodium acetate was added 1:10 (~40 µl) to the gel slice, and the agarose gel was melted at 68 °C For 7 min after an initial incubation at 50 °C for 15 min. Following a 15 min incubation at 42 °C, β-agarase (NEB) was added 1:50 (~8 µl) and the incubation was continued for 1 h longer. Following a gentle phenol extraction (by slowly inverting the tube for 10 min), and a standard ethanol precipitation (in the presence of 1 µl glycoblue), DNA was resuspended in 20 µl TE (pH 8.0). All but 0.5 µl was used to transform yeast. This 0.5 µl sample was analyzed by FIGE as above.

Yeast Agarose Plugs

Yeast cultures (50 ml) were grown in CSM-His plus adenine medium (Teknova) to an OD_{600} of 1.0. Cultures were harvested and then washed with 50 ml water. Next, the cultures were harvested and then washed with 10 ml EDTA pH 8.0. The cultures were harvested and then transferred to microfuge tubes with 750 µl cell resuspension buffer (Bio-Rad CHEF Genomic DNA Plug Kit). Cell pellets were then resuspended in 150 µl cell resuspension buffer. This mixture was equilibrated to 50 °C and then mixed with 85 µl Zymolyase-100T solution (20 mg/ml Zymolyase-100T [USB], 50% glycerol, 2.5% glucose, 50 mM Tris-Cl, pH 8.0) and 225

µl 2% low melting point agarose. Approximately 85 µl of this mixture was added to agarose plug molds (Bio-Rad), which were kept cold on ice. Following a 30 min incubation on ice, plugs were added to 5 ml lyticase buffer (10 mM Tris-Cl pH 7.5, 50 mM EDTA pH 8.0) containing 500 µl Zymolyase-100T solution. Following incubation at 37 °C for 2 hours, plugs were washed with 25 ml water. The Bio-Rad CHEF Genomic DNA Plug Kit was used to carry out the Proteinase K incubation and wash steps, which are described in the manual that is provided with the kit.

Restriction analysis of *M. mycoides* genomes propagated in yeast
Yeast agarose plugs were washed two times for 1hr in 1ml of 0.1X Wash Buffer (Bio-Rad CHEF Genomic DNA Plug Kit) and equilibrated for 1hr in 1ml of 1X Buffer 2 supplemented with BSA (NEB). The yeast chromosomal DNA was digested overnight with 50 units each of the restriction enzymes AsiSI and RsrII. These enzymes do not digest the *M. mycoides* genome. Plugs were then loaded onto a 1% TAE agarose gel to run digested yeast genomic DNA fragments out of the plugs (the circular *M. mycoides* genomes remain in the plug). The agarose gel was electrophoresed for 3 hours at 6V/cm. After electrophoresis, the agarose plugs were removed from the wells and were washed two times for 1hr in 1ml of 0.1X Wash Buffer and equilibrated for 1 hr in 1 ml of 1X Buffer 3 (NEB; for BssH II digests) or 1X Buffer 4 (NEB; for Asc I digests). The *M. mycoides* genomic DNA was digested overnight with 50 units of the restriction enzyme BssH II or 50 units of Asc I. Following incubation, all plugs were washed for 1 hour with 1 ml 0.1X Wash Buffer at room temperature and loaded on an agarose gel and subjected to pulsed-field gel electrophoresis. All restriction enzymes were purchased from NEB.

Pulsed-field gel electrophoresis
Yeast agarose plugs and *M. mycoides* agarose plugs were subjected to pulsed-field electrophoresis in a 1% agarose gel in 1X TAE with 0.5 µg/ml ethidium bromide with circulation at 14 °C, with a contour-clamped homogeneous electric field (CHEF DR III; Bio-Rad). Pulse times were ramped from 60 to 120 s for 20-24 hr at 4.0 V/cm.

Verification of cassettes, assembly intermediates, and *M. mycoides* JCVI-syn1.0
Just as the construction scheme for the synthetic cell genome followed a tiered hierarchy (1-kb cassettes, 10-kb subassemblies, 100-kb subassemblies, and the completed synthetic genome), so did the quality control efforts to ensure that the DNA being assembled had the sequence that it was designed to have. Because of variable cost constraints, variable time constraints, and the intrinsic differences in sequencing (many small regions of DNA with overlapping sequence from sequencing one large region of largely non-repetitious DNA), different strategies of sequencing were applied to each tier. The easiest, most economical, and fastest methods of sequence validation were substantially different at each of the different tiers of the construction effort. Also, the variable limitations of these methods caused a minority of mutations introduced at the cassette level to persist all the way to the 100-kb tier. Also, every tier of assembly was associated with the creation of some mutations. This demonstrates the need for quality control at all levels of synthetic biology efforts that endeavor to construct large regions of DNA.

1-kb cassette verification. Cassettes were purchased from Blue Heron. Blue Heron also provided sequence trace files confirming the sequences of the delivered DNA. These trace files

were processed by us in a semi-automated manner. In-house base-calling software was used to extract sequence from the trace files into FASTA format. The resulting sequence files were aligned to the target reference sequence using ClustalW (*8*) with default settings. The resulting pair-wise alignments were parsed with PERL scripts written for the purpose and sequence positions which did not match the target reference sequence were identified. Because each 1-kb cassette was covered by two or more Blue Heron-provided trace files, which in turn gave rise to as many pair-wise alignments, it was necessary to reconcile the list of mismatched positions from all pair-wise alignments covering a particular 1-kb cassette with each other. This was done by another custom PERL script, and cassette-positions that differed from the target sequence when all reads were taken into account were then manually analyzed. Of 1031 cassettes, 578 passed wholly automated screening. Of the remaining 453 cassettes, 370 had discrepancies that fit into categories that were judged not to require direct manual analysis of the trace files. There were two such categories: 1) The apparent sequence deviation was in the Not I site flanking the cassette. This was very common due to the poor performance of Sanger sequencing immediately following the sequencing primer, but was not of any concern because the intactness of the Not I site was immediately testable in the normal course of the processing of the cassette which involved cutting the cassette out of the cloning vector with Not I, and gel purifying it and 2) Some of the Blue Heron provided trace files contained unusually broad C peaks as is sometimes associated with the oxidation of the C-big-dye terminator reagent. As a result, some of the trace files gave rise to incorrect base calls of C. Eighty-three of the 453 cassettes that did not pass automated verification had their trace files inspected by eye; 62 were deemed to not require in-house re-sequencing, and 21 were re-sequenced. Three of the 1-kb cassettes were, in the final analysis, determined to have mutations (two single nucleotide deletions and one single nucleotide substitution), but only one of those was found during the screening of Blue Heron-provided trace files. The failure to detect the other two errors at the 1-kb tier was largely due to the poor quality of a minority of reads (those with unusually broad C peaks) and, in one case, the incorrect clone having been delivered. All 1-kb cassettes that passed this screen were used to create 10-kb subassemblies.

10-kb assembly intermediate verification. The 10-kb subassemblies were pooled and verified via 454 sequencing using non-paired-end reads. Initially 116 10-kb assemblies were screened. These 116 include some 10-kb subassemblies that were duplicates or near-duplicates. This was the case because alternate clones of certain troublesome 10-kb subassemblies were tested in parallel. Also, alternate versions of certain 10-kb subassemblies that were identical except for certain designed changes such as the presence of the watermarks were tested in parallel. Because of the presence of duplicate or near-duplicate sequences that needed to be read independently, as well as the existence of duplicate regions in contiguous 10-kb subassemblies due to the 80 base pair overlaps, the 116 10-kb subassemblies were pooled into 4 separate pools that prevented any 10-kb subassembly from being in the same pool as either of the two 10-kb subassemblies that shared 80 base pairs with its two ends. All alternate clones and alternate versions of specific 10-kb segments were separated from each other. Also, certain 10-kb subassemblies which had highly similar sequence to one another, such as those that contained IS1296 elements, were separated as much as possible from one another among the four pools. The four pools were bar-coded and then pooled into one pool which was sequenced by 454. The reads from the four bar-coded pools were then separated and analyzed using the CLC Workbench software package's

high-throughput reference assembly tool at default settings. The reads from each pool were assembled against an artificially constructed "reference sequence" which consisted of the 29 10-kb target sequences for that pool separated by runs of 500 N's. After assembly to this reference sequence, mutations were detected with a mixture of manual observation, the CLC SNP detection tool, and the CLC DIP detection tool. Oddly, CLC has no automated way of detecting differences between the reference and consensus sequences. To bypass this defect, the following non-default settings on the SNP and DIP tools were used: Minimum Coverage=1, Maximum Coverage=999999, Minimum Variant Frequency=50% or count=999999. All other settings were at their defaults. Initially 18 of the 116 10-kb subassemblies were found to contain variations from their target sequences. Eventually, three of these were determined to be caused by sequence similarity between the markers in the synthetic genome and the markers in the plasmid backbone of the 10-kb subassembly clones with backbone sequence contaminating the 10-kb sequences. Several of the remainder were found to be associated with the plasmid induction of the 10-kb clones in *E. coli*. This was determined after targeted PCR and sequencing of mutation regions of selected clones isolated directly from yeast, from *E. coli* before induction, and *E. coli* after induction. One class of mutation that was observed 3 times was the introduction of a GC at the sight of an overlap. We postulate that this might be a remnant of the Not I site flanking the cassette overlap sequences. Alternate assemblies of the 10-kb segments that were mutated in the first attempt to make them were then sequenced as before, except this time in only two bar-coded pools. One of the three mutations in the 1-kb cassettes was detected at the 10-kb level. All the other mutations that were not caused by induction *E. coli* originated in construction.

100-kb assembly intermediate verification. Rather than sequence each of the 100-kb subassemblies, semi-synthetic genomes were constructed. These semi-synthetic genomes contained between two and ten of the eleven 100-kb subassemblies, and the rest of the genome was derived from natural *M. mycoides* genomes. These semi-synthetic genomes were transplanted, validating that no lethal mutations were in the synthetic fraction of the genome. It rapidly became apparent that only one of the 100-kb subassemblies was not viable: the one corresponding to cassettes 811a-900. Because this 89-kb segment was not available conveniently from yeast in quantities that would make direct genomic walking or 454 sequencing easy or cost effective, overlapping PCR amplicons were generated and sequenced with a sequencing primer every 200 base pairs along each amplicon, alternating between strands. This process was done for both the synthetic 811a-900 and the corresponding region of the natural *M. mycoides* genome that had been shown to be viable in the semi-synthetic experiments. Re-sequencing of the natural fragment was to guard against the possibility that the 811a-900 region had been designed with a deviation that would prevent viability upon transplantation. Whereas the sequencing of the synthetic 811a-900 was to guard against mutations away from the designed sequence, one such mutation away from the design was detected in the 811a-900 sequence. It was a one base pair deletion that was eventually determined to have existed in the 812a 1-kb cassette. This mutation was not captured at the 1-kb level due to the wrong 1-kb clone having been sent by Blue Heron. It was not detected at the 10-kb level because, coincidentally, many of the 454 reads happened to end or start in its immediate vicinity, causing mutation to seem to be poorly supported across the population of read, and thus ignored by the automated SNP and DIP detection tools.

Sequence analysis of the *M. mycoides* sMmYCp235-1 genome isolated from *M. mycoides* sMmYCp235 cells. Cells were grown in SP4 medium containing 10 mg/l tetracycline and genomic DNA was extracted using a Promega Genomic DNA Extraction Kit. Genomic sequencing by a small insert shotgun Sanger library generated 14,319 successful reads. PCRs and genomic walks for known repeats and problematic areas generated 551 additional Sanger reads. The combined sequencing data was assembled using Celera Assembler (Meyers, 2000) resulting in 7 contigs in 3 scaffolds, with a contig N50 of 279215 bp. Average coverage was 11.5X. An additional 133 Sanger clone primer walks and genomic walks were generated to close the remaining gaps and cover low coverage regions. A completely finished genome was produced in 19 days. The resulting *M. mycoides* sMmYCp235 assembly contains eight SNP discrepancies and two insertions compared to the designed *M. mycoides* JCVI-syn1.0 genome. Differences are summarized in Table S1.

Additional content regarding the assembly of 10-kb synthetic intermediates
DNA cassettes were supplied at about equal concentrations and contained in an *E. coli* cloning vector. Equal amounts of these cassettes were pooled in sets of 10, digested with Not I to release the inserts, gel-purified, and then mixed with a unique yeast/*E. coli* shuttle vector. This vector is a bacterial artificial chromosome (BAC; pCC1BAC) with an inserted histidine auxotrophic marker, centromere, and origin of replication for selection and propagation in yeast *(15)*. Yeast propagation elements were already designed into assembly 831-840. For this reason, the cloning vector for this assembly only included the BAC elements (pCC1BAC). Unique first-stage assembly vectors were produced by PCR-amplification with primers that contained 40 bp overlapping sequence to the ends of the cassettes that were to be assembled (*4*). Also included in these primers were Not I restriction sites, which allowed the assembled cassettes to be released intact from the vector. The vector/cassette mix was then transformed into yeast and incubated on selective plates for several days. Plasmid DNA was extracted from individual yeast clones and transformed into *E. coli*, a more suitable host for propagation of the assembled cassettes. Plasmid DNA was then isolated from individual *E. coli* clones and digested with Not I to screen for cells containing a vector with an assembled 10kb insert. 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%. One assembly, 791-799, could not be produced by homologous recombination in yeast. This fragment was divided into two parts, 791-795 and 796-799, which were individually assembled by in vitro recombination (*3*). With the exception of 211-220, all first stage assemblies were propagated in Epi300 *E. coli* cells. This cell line allows for induction of the cloning vector from single-copy to high-copy numbers. Assembly 211-220 was unstable in Epi300 cells and so was transferred to Stbl4 cells, where it could be stably maintained.

All of the first-stage intermediates were sequenced. Nineteen out of 111 assemblies contained errors. Our sequencing analysis revealed that assemblies 81-90 and 811-820 each contained a single error originating from cassettes 82 and 812, respectively. Cassette 82 was corrected at the 1-kb level and reassembled to produce an error-free clone of 81-90. The mutation in cassette 812 occurred in the 811 overlap, which does not contain an error. Therefore, when additional 811-820 clones were sequenced, some of these did not contain errors. One sequence-verified 811-820 clone was used in subsequent assembly reactions. We opted not to correct an error that was present in 121-130 since it was a synonymous mutation in a non-essential gene. This

mutation served as an additional variation to further distinguish the synthetic genome from a natural one. One mutation could be avoided by maintaining the assembled fragments at single-copy number during propagation in *E. coli*. One error was produced by incomplete removal of the Not I restriction site at one of the cassette junctions. Four errors likely originated from the primers used to PCR-amplify the cloning vectors. The remaining errors were produced during propagation in yeast. Alternate clones of 15 assemblies were selected and sequence-verified.

Supporting Online Material References

1. C. Lartigue *et al.*, *Science* **317**, 632 (Aug 3, 2007).
2. C. Lartigue *et al.*, *Science* **325**, 1693 (Sep 25, 2009).
3. D. G. Gibson *et al.*, *Nat Methods* **6**, 343 (May, 2009).
4. D. G. Gibson *et al.*, *Science* **319**, 1215 (Feb 29, 2008).
5. R. S. Sikorski, P. Hieter, *Genetics* **122**, 19 (May, 1989).
6. N. Kouprina, V. Larionov, *Nat Protoc* **3**, 371 (2008).
7. S. H. Leem, Y. H. Yoon, S. I. Kim, V. Larionov, *Genome* **51**, 155 (Feb, 2008).
8. M. A. Larkin *et al.*, *Bioinformatics* **23**, 2947 (Nov 1, 2007).
9. A. W. Rodwell, in *Methods in mycoplasmology, vol. 1,* S. Razin, J. G. Tully, Eds. (Academic Press, Inc., Newe York, 1983), pp. 163-172.

Watermark-1, 1246 base pairs

TTAACTAGCTAAGTTCGAATATTTCTATAGCTGTACATATTGTAATGCTGATAACTAATACTGTGCGCTTGACTGTGATCCTGATAAATAACTTCTTCTGTAGGGTAGAGTTTTA
TTTAAGGCTACTCACTGGTTGCAAACCAATGCCGTACATTACTAGCTTGATCCTTGGTCGGTCATTGGGGGATATCTCTTACTAATAGAGCGGCCTATCGCGTATTCTCGCCG
GACCCCCCTCTCCCACACCAGCGGTGTAGCATCACCAAGAAAATGAGGGGAACGGATGAGGAACGAGTGGGGGCTCATTGCTGATCATAATGACTGTTTATATACTAATGC
CGTCAACTGTTTGCTGTGATACTGTGCTTTCGAGGGCGGGAGATTCGTTTTTGACATACATAAATATCATGACAAAACAGCCGGTCATGACAAAACAGCCGGTCATAATAGAT
TAGCCGGTGACTGTGAAACTAAAGCTACTAATGCCGTCAATAAATATGATAATAGCAACGGCACTGACTGTGAAACTAAAGCCGGCACTCATAATAGATTAGCCGGAGTCGT
ATTCATAGCCGGTAGATATCACTATAAGGCCCAGGATCATGATGAACACAGCACCACGTCGTCGTCCGAGTTTTTTTGCTGCGACGTCTATACCACGGAAGCTGATCATAAAT
AGTTTTTTTGCTGCGGCACTAGAGCCGGACAAGCACACTACGTTTGTAAATACATCGTTCCGAATTGTAAATAATTTAATTTCGTATTTAAATTATATGATCACTGGCTATAGTC
TAGTGATAACTACAATAGCTAGCAATAAGTCATATATAACAATAGCTGAACCTGTGCTACATATCCGCTATACGGTAGATATCACTATAAGGCCCAGGACAATAGCTGAACTGA
CGTCAGCAACTACGTTTAGCTTGACTGTGGTCGGTTTTTTTGCTGCGACGTCTATACGGAAGCTCATAACTATAAGAGCGGCACTAGAGCCGGCACACAAGCCGGCACAGT
CGTATTCATAGCCGGCACTCATGACAAAACAGCGGCGCGCCTTAACTAGCTAA

Watermark-2 1081 base pairs

TTAACTAGCTAACAACTGGCAGCATAAAACATATAGAACTACCTGCTATAAGTGATACAACTGTTTTCATAGTAAAACATACAACGTTGCTGATAGTACTCCTAAGTGATAGCTT
AGTGCGTTTAGCATATATTGTAGGCTTCATAATAAGTGATATTTTAGCTACGTAACTAAATAAACTAGCTATGACTGTACTCCTAAGTGATATTTTCATCCTTTGCAATACAATAA
CTACTACATCAATAGTGCGTGATATGCCTGTGCTAGATATAGAACACATAACTACGTTTGCTGTTTTCAGTGATATGCTAGTTTCATCTATAGATATAGGCTGCTTAGATTCCCT
ACTAGCTATTTCTGTAGGTGATATACGTCCATTGCATAAGTTAATGCATTTAACTAGCTGTGATACTATAGCATCCCCATTCCTAGTGCATATTTTCATCCTAGTGCTACGTGAT
ATAATTGTACTAATGCCCTGTAGATAATTTAATGCCTGGCTCGTTTGTAGGTGATAATTTAGTGCCTGTAAAACATATACCTGAGTGCTCGTTGCGTGATAGTTCGTTCATGCAT
ATACAACTAGGCTGCTGTGATATGGTCACTGCCCTTACTGTGCTACATATTACTGCGAGGGGGATGACGTATAAACCTGTTGTAAGTGATATGACGTATATAACTACTAGTGA
TATGACGTATAGGCTAGAACAACGTGATATGACGTATATGACTACTGTCCCAAACATCAGTGATATGACGTATACTATAATTTCTATAATAGTGATAAATAAACCTGGGCTAAA
TACGTTCCTGAATACGTGGCATAAACCTGGGCTAACGAGGAATACCCATAGTTTAGCAATAAGCTATAGTTCGTCATTTTTAAGGCGCGCCTTAACTAGCTAA

Watermark-3 1109 base pairs

TTAACTAGCTAATTTAACCATATTTAAATATCATCCTGATTTTCACTGGCTCGTTGCGTGATATAGATTCTACTGTAGTGCTAGATAGTTCTGTACTAGGTGATACTATAGATTTC
ATAGATAGCACTACTGGCTTCATGCTAGGCATCCCAATAGCTAGTGATAGTTTAGTGCATACAACGTCATGTGATACAACGTTGCTGGCTGTAGATACAACGTCGTATTCTGT
AAGTGATACAATAGCTATTGCTGTGCATAGGCCTATAGTGGCTGTAACTAGTGATATCACGTAACAACCATATAAGTTAGATTTAATGCCCCTGACTGAACGCTCGTTGCGTG
ATAGTTTAGGCTCGTTGCATACAACTGTGATTTTCATAAAACAACGTGATAATTTAGTGCTAGATAAGTTCCGCTTAGCAAGTGATAGTTTCCGCTTGACTGTGCATAGTTCGT
TCATGCGCTCGTTGCGTGATAAACTAGGCAGCTTCACAACTGATAATTTAATTGCTGATATTGCTGGCTGTCTAGTGCTAGTGATCATAGTGCGTGATAGTTTAAGCTGCTCT
GTTTTAGATATCACGTGCTTGATAATGAAACTAACTAGTGATACTACGTAGTTAACTATGAATAGGCCTACTGTAAATTCAATAGTGCGTGATATTGAACTAGATTCTGCAACTG
CTAATATGCCGTGCTGCACGTTTGGTGATAGTTTAGCATGCTTCACTATAATAAATATGGTAGTTGTAACTACTGCGAATAGGGGGAGCTTAATAAATATGATCACTGTGCTAC
GCTATATGCCGTTGAATATAGGCTATATGATCATAACATATATAGCTATAAGTGATAAGTTCCTGAATATAGGCTATATGATCATAACATATACAACTGTACTCATGAATAAGTT
AACGAGGATTAACTAGCTAA

Watermark-4 1222 base pairs

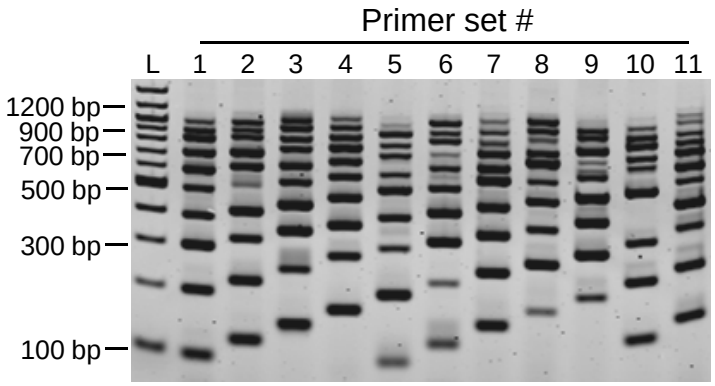
TTAACTAGCTAATTTCATTGCTGATCACTGTAGATATAGTGCATTCTATAAGTCGCTCCCACAGGCTAGTGCTGCGCACGTTTTTCAGTGATATTATCCTAGTGCTACATAACA
TCATAGTGCGTGATAAACCTGATACAATAGGTGATATCATAGCAACTGAACTGACGTTGCATAGCTCAACTGTGATCAGTGATATAGATTCTGATACTATAGCAACGTTGCGT
GATATTTTCACTACTGGCTTGACTGTAGTGCATATGATAGTACGTCTAACTAGCATAACTAGTGATAGTTATATTTCTATAGCTGTACATATTGTAATGCTGATAACTAGTGATA
TAATCCAACTAGATAGTCCTGAACTGATCCCTATGCTAACTAGTGATAAACTAACTGATACATCGTTCCTGCTACGTGATAGCTTCACTGAGTTCCATACATCGTCGTGCTTAA
ACATCAGTGATAACACTATAGAGTTCATAGATACTGCATTAACTAGTGATATGACTGCAAATAGCTTGACGTTTTGCAGTCTAAAACAACGTGATAATTCTGTAGTGCTAGATA
CTATAGATTTCCTGCTAAGTGATAAGTCTACTGATTTACTAATGAATAGCTTGGTTTTGGCATACACTGTGCGCTGCACTGGTGATAGCTTTTCGTTGATGAATAATTTCCCTA
GCACTGTGCGTGATATGCTAGATTCTGTAGATAGGCTAAATTCGTCTACGTTTGTAGGTGATAGTTTAGTTGCTGTAACTAATATTATCCCTGTGCCGTTGCTAAGCTGTGATA
TCATAGTGCTGCTAGATATGATAAGCAAACTAATAGAGTCGAGGGGGAGTCTCATAGTGAATACTGATATTTTAGTGCTGCCGTTGAATAAGTTCCCTGAACATTGTGATACT
GATATTTTAGTGCTGCCGTTGAATATCCTGCATTTAACTAGCTTGATAGTGCATTCGAGGAATACCCATACTACTGTTTTCATAGCTAATTATAGGCTAACATTGCCAATAGTGC
GGCGCGCCTTAACTAGCTAA

Figure S1


ribose galactose
yeast vector sequence
phosphonate
oligopeptide
carbohydrate/maltose
spermidine/putrescine
phosphate
oligopeptide
cobalt
alkylphosphonate
Type II R/M - GCATC
Type II R/M - CCATC
Type II R/M - GANTC
Type II R/M - CCTTC
Type III R/M
40 kb
80 kb
120 kb
160 kb
200 kb
240 kb
280 kb
320 kb
360 kb
400 kb
440 kb
480 kb
520 kb
560 kb
600 kb
640 kb
680 kb
720 kb
760 kb
800 kb
840 kb
880 kb
920 kb
960 kb
1000 kb
1040 kb
Biosynthesis of cofactors, & carriers
Transport & binding proteins
Fatty acid & phospholipid metabolism
Central intermediary metabolism
Purines, pyrimidines, nucleosides, & nucleotides
Regulatory functions
Energy metabolism
Amino acid biosynthesis
Mobile element functions
Hypothetical proteins
Unknown function
Transcription
Unclassified
RNA gene
Protein fate
Cellular processes
Cell envelope
DNA metabolism
Protein synthesis
WM#
Watermarks
100 Kb segment border
10 Kb segment border
Designed polymorphism
Unexpected polymorphism
E. coli IS element
85 bp insertion

Figure S2

A



B

Amplicon	a	b	c	d	e	f	g	h	i	j	k
Set 1 sizes (bp)	100	200	300	400	500	600	700	800	900	1000	
Set 2 sizes (bp)	125	225	325	425	525	625	725	825	925	1025	
Set 3 sizes (bp)	150	250	350	450	550	650	750	850	950	1050	
Set 4 sizes (bp)	175	275	375	475	575	675	775	875	975	1075	
Set 5 sizes (bp)	100	200	300	400	500	600	700	800	900	1000	
Set 6 sizes (bp)	125	225	325	425	525	625	725	825	925	1025	
Set 7 sizes (bp)	150	250	350	450	550	650	750	850	950	1050	
Set 8 sizes (bp)	175	275	375	475	575	675	775	875	975	1075	
Set 9 sizes (bp)		200	300	400	500	600	700	800	900	1000	
Set 10 sizes (bp)	125	225	325	425	525	625	725	825	925	1025	
Set 11 sizes (bp)	150	250	350	450	550	650	750	850	950	1050	1150

Figure S3

Figure S4

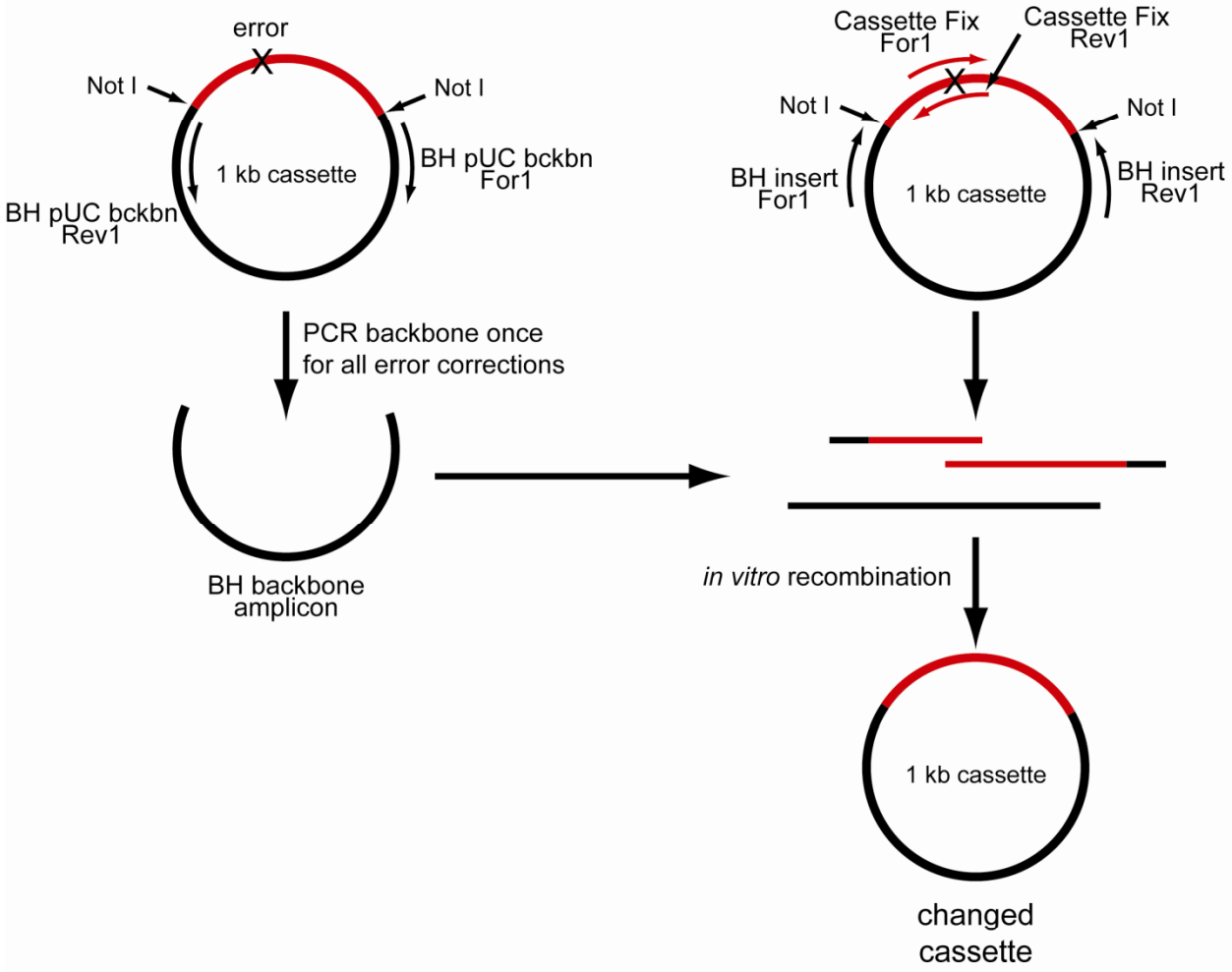

error
Not I
Not I
1 kb cassette
BH pUC bckbn For1
BH pUC bckbn Rev1
PCR backbone once for all error corrections
BH backbone amplicon
Cassette Fix For1
Cassette Fix Rev1
Not I
Not I
BH insert For1
1 kb cassette
BH insert Rev1
in vitro recombination
1 kb cassette
changed cassette

Figure S5

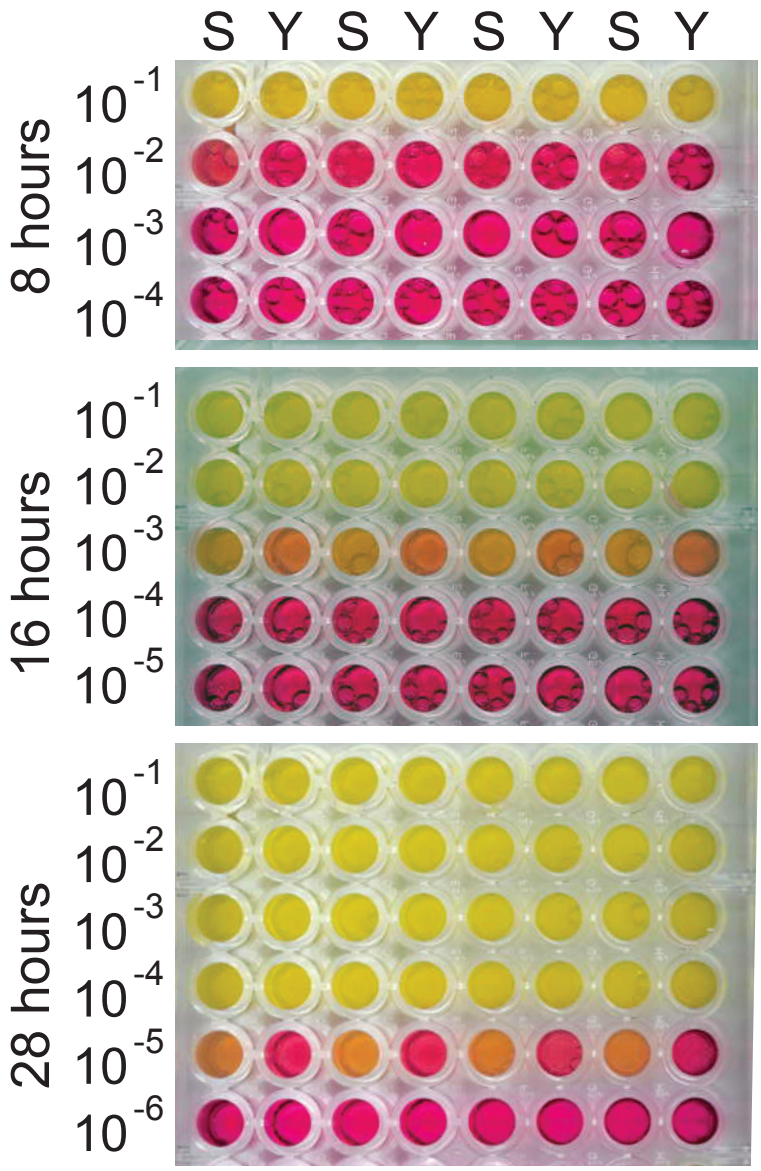

S Y S Y S Y S Y
8 hours
10-1
10-2
10-3
10-4
16 hours
10-1
10-2
10-3
10-4
10-5
28 hours
10-1
10-2
10-3
10-4
10-5
10-6

Figure S6

Supplementary Figure Legends

Figure S1. Synthetic genome watermarks. A 6-frame stop sequence is in black on green. Asc I restriction sites are in red on grey.

Figure S2. Genome map of *M. mycoides* JCVI-syn1.0. Genes, structural RNAs, watermarks, polymorphisms relative to natural *M. mycoides capri* GM12, and the coordinates of the synthetic DNA cassettes used to construct the genome are shown.

Figure S3. Multiplex PCR patterns used to identify 100-kb intermediates containing all first-stage assemblies. (a) To demonstrate the expected PCR patterns for each of the 100-kb assemblies, the multiplex PCR primers listed in **Table S3** were used to screen YCpMmyc1.1, which was extracted from yeast. **(b)** PCR amplicons are spaced 100 bp apart. With the exception of amplicon 10d (shaded in yellow; the region corresponding to 931-940), and assembly 796-799, all first-stage assemblies can be accounted for by this PCR analysis. Downstream analyses were used to further confirm the 701-799 and 901-100 intermediates.

Figure S4. Proteomic Analysis. Two-dimensional gels were run using cell lysates from *M. mycoides* YCpMmyc1.1 (WT) and *M. mycoides* JCVI-syn1.0. Standard conditions were used for the separation of protein spots. The first dimension separated proteins on immobilizing pH gradient strips (pH range 4 to 7 from left to right in the images). In the second dimension, the proteins are separated in 9-17% acrylamide gradient gels that resolve the proteins in an M_r range from around 130 kDa to 10 kDa (from top to bottom in images). The gels were stained with Coomassie Brilliant Blue G-250.

Figure S5. Error correction using PCR and in vitro recombination. Primers (BH pUC bckbn For1 and Rev1) are used to amplify the plasmid backbone for use in recombination reactions. Complements of backbone primers (BH insert For1 and Rev1) are used in conjunction with error-correcting primers (Cassette Fix For1 and Rev1) to produce amplicons with regions of homology to each other and the BH backbone amplicon. The three PCR products are used in in vitro recombination to generate the corrected cassette.

Figure S6. Comparison of the growth rates for *M. mycoides* JCVI-syn1.0 and the WT control strain YCpMmyc1.1. Growth rates were compared using a color changing unit (CCU) assay to measure the rates at which serial dilutions of each cell type acidified the culture media. This causes the phenol red in the SP4 culture media to change from a red to a yellow culture. This assay is a standard method used by mycoplasmologists to determine culture titers (*9*). We took aliquots of both the JCVI-syn1.0 (S) and YCpMmyc1.1 M. mycoides (Y) strains that were just beginning to turn from red to orange. These strains were compared because they are the most similar organisms we have in our library. Both organisms have the YAC vector sequence, although there are several other genes absent in the JCVI-syn1.0 genome that could alter the growth rates. We made serial 10 fold dilutions in a 96-well plate of each culture and incubated the cells at 37 °C. Every hour we photographed the cultures to determine whether the cultures were red (early log phase), orange (late log phase) or yellow (senescent). After 8 hours of growth, all the JCVI-syn1.0 and YCpMmyc1.1 cultures at each dilution are essentially the same color, indicating the wells at each dilution contain similar numbers of organisms at each dilution. After 28 hours, the JCVI-syn1.0 culture in the 10^{-5} dilution is much more acidic and turbid than the YCpMmcy1.1 wells, which are still red. In the figure, we cropped dilutions that had not begun to change color and one well in the 10^{-7} dilution that was apparently contaminated.

1	2	3	4	5	6	7
designed difference	**coordinate on JCVI-syn1**	**JCVI-syn1**	**YCpMmyc1.1**	**affected cassettes**	**Search string following difference**	**Comment**
homopolymer run	59700	A15	A16	mmyc872	TAATATAATTTGCATACTATAAAAA	intergenic
homopolymer run	118608	A16	A17	mmyc931	GCAAAAAAACTAGTAAATATCGTAT	intergenic
homopolymer run	120703	A18	A17	mmyc933	CAAAATAGCTTAGTACATATCGCAT	intergenic
homopolymer run	122656	A17	A16	mmyc935	GCAACAAGGTGTGATCGCACCATTT	intergenic
deletion "94D"	123020	deletion	4130 bp	mmyc936-939	TATTTTGTAAAACAGTTTCATCAAGTT	putative lipoprotein genes
homopolymer run	123429	A18	A21	mmyc940f	GCAAACAAACCTAGTACACATCGTA	intergenic
dinucleotide run	159343	(TA)10	(TA)9	mmyc976	ATGGAATTAACTATATACACTAGGT	intergenic
homopolymer run	174542	T16	T15	mmyc991	AATAAAAAGACTCACATAAGTGAGT	intergenic
watermark	388631-389563	WM3	1000 bp	mmyc106	TGATGCACTGATGTAACAGAGTTTAAA	coded watermark sequence
snp	405381	G	A	mmyc122	CCAATGCTAAAAAATCAAAATGCTCATT	A->G, Gly synonymous, CDS: putative lipoprotein
homopolymer run	420496	A22	A18	mmyc137	CTAAAAACAAACTCTATTTCTTAAA	intergenic
homopolymer run	421365	A19	A20	mmyc138	CTAAAAACAAACTCTATTTCTTAAA	intergenic
homopolymer run	426516	T16	T15	mmyc143	GTAAAAATAAAAGTAAGATTTGGAT	intergenic
watermark	564644-565713	WM1	6000 bp	mmyc282-287	AATTACCAAAAGGTGTAGTTTCAGTTC	coded watermark sequence
homopolymer run	577948	T18	T17	mmyc300	GCTAAAAAGCCTATGTAAAGCCTTT	frameshift, CDS: hypothetical protein
watermark	724794-725698	WM2	1000 bp	mmyc447	TATCAAGATACTTAGTAATGCTAGTTTCTCC	coded watermark sequence
dinucleotide run	727553	(TC)12	(TC)10	mmyc449	GGCTACTATCAAGTTCTTAAGTGCATCA	intergenic
homopolymer run	900703	T21	T19	mmyc622/623	ACAAGCAATTTAATACAAACTTGTA	intergenic
watermark	957779-958824	WM4	1000 bp	mmyc680	CAAATTCAATGCTAGACAAAGCACTT	coded watermark sequence
homopolymer run	1014079	A16	A17	mmyc736	TAAATCCAGCTTAGTACTCATCAAA	intergenic
homopolymer run	1064216	T18	A17	mmyc786	GCAGAGCTCGTATTATTCTTTTCTT	intergenic
dinucleotide run	1069129	(AT)8	(AT)10	mmyc791	TATAATGTGTCTTACAAAAGATAAA	intergenic
dinucleotide run	1069967	(TA)15	(TA)12	mmyc792	ATGGTTATTGACAGGAAAAATATAT	intergenic
dinucleotide run	1070891	(TA)13	(TA)14	mmyc793	ATGTGTCTTACCAAATTATAAACAA	intergenic
dinucleotide run	1071855	(TA)12	(TA)11	mmyc793/794	ATGATTATTGATATGAAAAATATAT	intergenic
not designed		**sMmYCp235-1**				
snp	48020	G	C	mmyc861	CAAATAAAGTACCAAGTACATAACCTA	C -> G, Gly -> Ala, CDS: GnsA/GnsB family
snp	189893	C	G	mmyc1006	CAAATAAAGCGTTTGCTCATTATCGTT	G -> C, Ala -> Pro, CDS: rpsG
insertion	308736	777 bp insertion	"-"	mmyc24	ATGTTGGGTTAAAATCATATGATTATATAA	E. coli IS1 transposable element inserted in peptidase, S41 family
insertion	338117	85 bp insertion	"-"	mmyc53-54	AAGTATAATCTTGATAATCTTCAAATTTTA	GGCCGC (from NotI) plus 79 bp duplication Disrupts CDS: thiamine pyrophosphokinase
snp	347965	A	T	mmyc65	AAAAATGTTGCTAATGAGTAATAAAAAGT	T -> A, Leu -> Phe, CDS: efflux ABC transporter,_permease protein
snp	678748	G	A	mmyc400f/401	GTTGAGTTGATTTTTTAGTATTTTCAGC	A -> G, Ser -> Gly, CDS: hypothetical protein
snp	735321	A	G	mmyc457	AGATAAAGAACAAATGCCACCAACTAAT	G -> A, Gly -> Glu, CDS: hypothetical protein
snp	858167	A	G	mmyc580	AAAAACAGCAACAAGTAGTGATGCTGA	G -> A, Gly -> Glu, CDS: tpiA
snp	858175	T	G	mmyc580	CAACAAGTAGTGATGCTGAAGAAGTATG	G -> T, Ala -> Ser, CDS: tpiA
snp	1036988	G	C	mmyc759f	TTATTCATTCATTTTTTCTTTAACAGAT	intergenic
Table S1						

Primer name	Primer sequence	Amplicon size (bp)
C3a F Set 1(1)	ACTTATCTCGTAGACAAGATATTTATCAAAAC	100
C3a R Set 1(1)	GCAACCTCCACAAGCTTCTTGT	
C16 F Set 1(2)	ATCAATATAATGTTAATTTTTCTTCATTATATATG	100
C16 R Set 1(2)	ATAAGCATCAAATAACATTATCTCCAACTTTTATA	
C25 F Set 1(3)	AAAACTATTACAGTATCTGATAATGACTTTTTTAC	100
C25 R Set 1(3)	TAATAATCGACTTTCCAAGTGTTTTACT	
C35 F Set 1(4)	TAAGGTTAAACAAACTATAGAGTAAAAACAATAAT	400
C35 R Set 1(4)	ATTTTGTTTTTGAGTCTTTTTGAAATC	
C45 F Set 1(5)	TAAAATTAGTTTGCTCTAGTTTTAAAAAAGC	500
C45 R Set 1(5)	AAATGATATTTTATTACAAGTTATTAAATTTTATGATTATC	
C55 F Set 1(6)	ACATCAACATTAAATATCAATTTTTTTTAATATC	600
C55 R Set 1(6)	TTTACAGGACAAACTGGAGCTG	
C63 F Set 1(7)	CCAATCAGTTAATCTTGGTCTTG	700
C63 R Set 1(7)	AGCAATTGATCAAATTATTGGTGATC	
C75 F Set 1(8)	TGCTTCAAGATATTTTAAAAATCCTTTAG	800
C75 R Set 1(8)	TTAAAAAGCAGCTGTATTTGTTTTAAA	
C85 F Set 1(9)	AAGTATAGTGTTCTTTTCATATATCTCTCTTATTT	900
C85 R Set 1(9)	CACCAATTAGGGATCGAAGTT	
C95 F Set 1(10)	TTACATTAGCAGCCCCTACTCTCTCAT	1000
C95 R Set 1(10)	ATTTAAATAGTTTAAATATAGCAAGACAAAAAAATT	
C105 F Set 2(1)	TATTTTCTTTTAAAACTTCACTACCAGTTG	125
C105 R Set 2(1)	TGATTTTATAGATAAAAAGGGTATGGAA	
C115 F Set 2(2)	TAAACACTAGTGTTGGTACTTTAAAAAAAG	225
C115 R Set 2(2)	CAATTGCTATAATTGTTGGATATTTTAAATC	
C125 F Set 2(3)	TGAATATACAAAAAAAATCGTGTGAATAAAC	325
C125 R Set 2(3)	TTTTGATTTTGAATCATCCATTAATAAAA	
C135 F Set 2(4)	AATTTTATAATGCATCCAAATATGAGTTT	425
C135 R Set 2(4)	ACAGTACCAAAAATTGCAATTTTCAT	
C145 F Set 2(5)	ATATTGATATGATAAAAAAGCTTTAGAAGATAAAAC	525
C145 R Set 2(5)	TCATATTTTTTGAATGTATAAAAGTGTTGTT	
C155 F Set 2(6)	AGATTATATGGTTAAAGCTCAACAACAACTAG	625
C155 R Set 2(6)	AACAAAATAATAACATCAGCTGTTCAATA	
C165 F Set 2(7)	AGATGTATCAGGTTTAACAAAATATCAGG	725
C165 R Set 2(7)	CGCCTAATACTACATGACTTCCTT	
C175 F Set 2(8)	ACTTGACTATATTAATATTAAAGCTTGGTCAAA	825
C175 R Set 2(8)	ATAATAATGAATATCAGTTGTGATTAGCTTTC	
C185 F Set 2(9)	AGTGTTAGAAACTCAAATGGATCA	925
C185 R Set 2(9)	CATATTTTAATATCTAAACTTAGATTTAAAAAATTATTAGC	
C195 F Set 2(10)	CAAGGAATAGTTGTTAAAACTTTAGATAAAGC	1025
C195 R Set 2(10)	TAAAAATTTTATCTTCATTAATATTTAAAATCTTGTAAT	
C205 F Set 3(1)	GAATATAAATAGCCTTTATAACTTGCTTCATAATC	150
C205 R Set 3(1)	TGATTCAAAAGAATCACAAGAAAATG	
C215 F Set 3(2)	AAGTTTTTATTAAATAAGTATTATCAAAAAATAAAAGTAAAAAAC	250
C215 R Set 3(2)	AAAGATCAAAACAAATAAACTAAAAAAATAGT	
C225 F Set 3(3)	AGAAAAAAACAACCCACATATTGTAATT	350
C225 R Set 3(3)	GTATTAAACTTTGCTCTAGTTCTTTGTTTTAG	
C235 F Set 3(4)	AGGGCAACTAAAGACTTTTCTCCTC	450
C235 R Set 3(4)	AAATTACTGAAGAAACCGCATTAAC	
C245 F Set 3(5)	ATATATGTTATTTAAAAAAATAACAAATAATGTTCAG	550
C245 R Set 3(5)	TTTAATAGCATCATAACTTAAACTAATGATAATTG	
C255 F Set 3(6)	GATATAAATGATCACATTATTCATAATGCTAAT	650
C255 R Set 3(6)	GTTTATGAGCAATTACAGTTCAACAAG	
C265 F Set 3(7)	TTGATTAGAAAAAACTGAAATTCCAA	750
C265 R Set 3(7)	AAGTTTTTATTATTATAAATTTGATCTAAATTAGTTGC	
C275 F Set 3(8)	ATAAAATTACTTTTAATGATAAAAGTTTTATTTATTTATGAG	850
C275 R Set 3(8)	CAGCAATTCCAGTAGCTCCAT	
C285 F Set 3(9)	AATTACCAAAAGGTGTAGTTTCAGTTC	950
C285 R Set 3(9)	GTCTTTAACTTTTGAGCACTATTTAATTTATC	
C295 F Set 3(10)	TCAACTTCTCGGGTTTTGAAAATATC	1050
C295 R Set 3(10)	GTGATATTGATCAAAACAAATTTATGAGT	
C305 F Set 4(1)	TGATCAACTCCAATCATTGC	175
C305 R Set 4(1)	TAGTGACTTATGCTAGTGAATTAGATGATG	
C315 F Set 4(2)	TAATTTAAAAATTTTATTTTTTAATAATTTCAGCC	275
C315 R Set 4(2)	ACAAGATATTTATTTGCTATTTGATATATAAAAATC	
C325 F Set 4(3)	TTTAGTATAACCTCTTAAAGTTGTACTTATTAAATCA	375
C325 R Set 4(3)	ATGTGTTATGGCACTAGTTTTAGGTGG	
C335 F Set 4(4)	TTTTTTCAGCAAATACTTGATATGAATTA	475
C335 R Set 4(4)	TATAATTGAGTTAAAAAGCTGTTAGCA	
C345 F Set 4(5)	TTGCTAGGTTTGGATATTTTCATTC	575
C345 R Set 4(5)	AAGAACAACAAGAAGTAGCGAATTTG	
C355 F Set 4(6)	ACATTTAAAACAAATTTTTAACAAAAGTAGT	675
C355 R Set 4(6)	TTATAGTTTATTYCAAGTAGTTAATCATTATAAACAATT	
C365 F Set 4(7)	AAAACAGAACCAATAACAGCATTTC	775
C365 R Set 4(7)	ACCATATGATAGAAATGTAATTACTGAAGC	
C375 F Set 4(8)	GAAATTAACTCTACAACTAAAATTCCTGTTG	875
C375 R Set 4(8)	GTTGTTAATGAAATTCAAAAATCTAATGG	
C385 F Set 4(9)	TTTAAAATCTCTTGTTATTATAAATCTGATAAAC	975
C385 R Set 4(9)	TGTTATTAGCATTTTTATAAAACTCTTTAATCAC	
C395 F Set 4(10)	GCCAAAAAGAGAATTATTGAAAAAATT	1075
C395 R Set 4(10)	AGCTTGTGTTTTAACTAATTGCTGTTTAG	
C405 F Set 5(1)	AGAAAATCCAATTGGATTAAAGAGTT	100
C405 R Set 5(1)	ATCTTCATATTCTTCGATTAATTTTTAATTCC	
C415 F Set 5(2)	TATGTAATGCTATAGGTTTAAGCACAGTT	200
C415 R Set 5(2)	TTTAATTTGTAAGAATTGTAAGGCTCTA	
C425 F Set 5(3)	GTGAATAAGTAAAAATATATATAAATGCAGAAATACTAG	300
C425 R Set 5(3)	TTTCATCTCTTTTTGTATCAATGTCAT	
C435 F Set 5(4)	CAACACCACTAATATAAATTTTTCTTGTT	400
C435 R Set 5(4)	TTAAAGAAGAAGAAAAATAAAATTTAACATTAGTATTAG	
C445 F Set 5(5)	TGTTATTTTTCTAGTTCTATGTGATGT	500
C445 R Set 5(5)	ATTCAAGCAGTAAAATACTGTATTTTTTAAAAAGAAG	
C455 F Set 5(6)	AACAATTCTTTTTATAAAACTAAAATTTATACTCAAT	600
C455 R Set 5(6)	TATTTTAGTATCAATATTTTGTTCATTAATAGTTG	
C465 F Set 5(7)	CCAGTAATTTTAAATAAAAATCTTAATAAGTAATTAAA	700
C465 R Set 5(7)	CGTTCTTTTGAATTTTATTAAAAATAGAATAAAC	
C475 F Set 5(8)	CTTTTATATTTGGATCTAAAATAGTTTTTTCTAAAGC	800
C475 R Set 5(8)	TAGTAAAGTTCATACTATTTTTAATCATACTTTAACAAATAC	
C485 F Set 5(9)	AACAGTTTATCAAAATAATGATCTTGC	900
C485 R Set 5(9)	GTGTGAAAAAAATGGTATTAAGCTAAGTC	
C495 F Set 5(10)	AGTATTGTATAAGTAGATTTTTTTATTTTTATTTAGTATTACAG	1000
C495 R Set 5(10)	TGATGGTTCTTATATGAAAAAATTATTAAGTGTTTTAG	
C505 F Set 6(1)	TCTTATGATTTTATATTAATAAATAAATTCATTAACAAATTG	125
C505 R Set 6(1)	AACTTTTTCATCAATTCAAGCAAAAC	
C515 F Set 6(2)	TCCAAGGTAAAATGCGGATG	225
C515 R Set 6(2)	TTTTGAACATACTATTTTAATTACAAAAGATGG	
C525 F Set 6(3)	TTTCCTAAAAGTTGAACTGAATGAACC	325
C525 R Set 6(3)	AACAAAAATACCGTATTATTGACTTTAAAAGAAAT	
C535 F Set 6(4)	GAACCCGTGTGTCCACGGAT	425
C535 R Set 6(4)	CAATGAATGCTATGTTTTCTGGTGC	
C545 F Set 6(5)	ATTTTTACATAGCATTGTTTCTTTATTAGTTTTAA	525
C545 R Set 6(5)	TTAGGTAATGTTTTTAGTTTTTGATGAGTTTTTAGAT	
C555 F Set 6(6)	TTTAAATGTTCTAAATGTTTAATTTATTATATCATTTAAATAAAC	625
C555 R Set 6(6)	TTATTCTGCACAATATGCATAGAGTTGAAGTTAG	
C565 F Set 6(7)	TTAGGTTTAAAGTTTGAATTATATGAATAGTAGC	725
C565 R Set 6(7)	TGTTAAAAAATATTAAGTTTTTAAAATATTAAAAATCTAAATTATTAAC	
C575 F Set 6(8)	AACTGTATGGGGTATTCTTTATCTTAC	825
C575 R Set 6(8)	CCAAATAATGAAAATGAATCTTCAATAATT	
C585 F Set 6(9)	TAACTCCACCAGCAGCTTTAACC	925
C585 R Set 6(9)	ATATGAAAAATATGGGATTATTACACAACTACTTATAATTTAAAC	
C595 F Set 6(10)	GGTAATTTTAAATGCACCAACTAATAACC	1025
C595 R Set 6(10)	TTAAAGCTGATGATTCAAATATTGGAG	
C605 F Set 7(1)	ATAGAAGTTATATTGTCATAAACTACGTATGATACAC	150
C605 R Set 7(1)	AGAACAAATTAAAAGAGAAATATGAAATTTTAAATTAAAG	
C615 F Set 7(2)	GCTCAGCTTCAATTATCAATTAATG	250
C615 R Set 7(2)	ACCCACAAAACACTATATGATATAGCCTATAGATAT	
C625 F Set 7(3)	GTTTGTTTACGATTCTGGTCCTACTG	350
C625 R Set 7(3)	ACAAGAATTGGATGATTAGATTTAGGTG	
C635 F Set 7(4)	ATTTAAAACAATATCTAATTACAAAAGATCTAGAAAA	450
C635 R Set 7(4)	TATTTTTAACAATTAAATGTTTCAGACATAAAAAC	
C645 F Set 7(5)	AATAGTTTGGGAAAACCTTCTAACGC	550
C645 R Set 7(5)	TGCTAAACCTGATGTAATCTTTATGC	
C655 F Set 7(6)	GATCCATAAACGTTATTTGGAAAC	650
C655 R Set 7(6)	GTTGATGCTTTAATAAAAAGTTTAAAAGATAAAG	
C665 F Set 7(7)	GATTTGATTAAAGTAAGTTTCCACAC	750
C665 R Set 7(7)	AAACAGGGTTGTTTTAGACAAAACTTATTAG	
C675 F Set 7(8)	GATTTGAGTGATTTATCATAAACATCTATTAAT	850
C675 R Set 7(8)	AAGACAAGCATATTTAAATTTATAAGTTTTAATATG	
C685 F Set 7(9)	AACAAAAAAATAAATTAGAGGTTGAATTTAATAAAC	950
C685 R Set 7(9)	CAACCAGTAGCTACACCACTTCTAATTAC	
C695 F Set 7(10)	CAAGTGATTGGAAGATGAGGTAAT	1050
C695 R Set 7(10)	TATTTTATTAAATGATGATATATTGGTCTCTATAAGTTTATTAG	
C705 F Set 8(1)	ATCATGTTGTATATGATAAATATAATAAACTTG	
C705 R Set 8(1)	TGCATATTTAGAAGATTTGAAATCTTTAAATAT	175
C715 F Set 8(2)	ATACTCATTTCATATTTTTAATATCAAAATAAAC	
C715 R Set 8(2)	ACGCTCAAACTGTTGTTGATTTAC	275
C725 F Set 8(3)	AATTGCATTTGTAGTTTATCTAAAACTTG	
C725 R Set 8(3)	GATTGGTTAAACATCCCTCTCATAG	375
C735 F Set 8(4)	ATATAAAGATAATGAATCTGTTTAAATTGGATATT	
C735 R Set 8(4)	AGAAATTTATAATGCTTTAAATATTAAAGTTAAAAAG	475
C745 F Set 8(5)	AGCTTCTGATTGAGTAACTAGAAATTCAG	
C745 R Set 8(5)	ATTAAATATAGGAATCTCTAATTTGAAATGACTTAAAATATAAATC	575
C755 F Set 8(6)	TTTTAAATATATTTATTTTAATTATCCTGATAGTTTAATTTATC	
C755 R Set 8(6)	CACCTAGTGGGACTATAAAGTTAATAATC	675
C765 F Set 8(7)	AAAAAAGAGCAAAACATCAAGATGTAG	
C765 R Set 8(7)	ATTGCTTTTTAAATAAAATAATAGCAATTTG	775
C775 F Set 8(8)	AATCTATTTACTGCTGGACAAATTAATG	
C775 R Set 8(8)	TTTTTGCAGATCTACTAAATAAAGTCTCTAAAAC	875
C785 F Set 8(9)	CAGTATTTAACTATTCAAATTTTAAACCTTTTTC	
C785 R Set 8(9)	TAATTTGACTATTTGATCAATTCTTAAAGC	975
C795 F Set 8(10)	AACAAAAAAATAAATTAGAGGTTGAATTTAATAAAC	
C795 R Set 8(10)	ATTAGGCTCTGTTGGGCTAGTT	1075
C805 F Set 9(1)	CTAATAGTAATGATCTAAGAGAAGTGGGAAGAG	
C805 R Set 9(1)	TAGTAATTTTTAAATTTATATAATCTAAATATAAAGATTTTAGATATTAA	100
C815 F Set 9(2)	GTAGAAATGTTGAAAAACTAAAGAAAATC	
C815 R Set 9(2)	ATCCAAATGTGAAAAGGGTTTTTATC	200
C825 F Set 9(3)	ATTGGACTAGAAGTTTTAATATAAAAGTTAGACCAG	
C825 R Set 9(3)	CTTAAATTCCACTGATAAAAAGTTGGG	300
C835 F Set 9(4)	CTAGAATACCATTGTAATCATATCTAACTG	
C835 R Set 9(4)	GGGCCGAAGCAGGCGTTGTTG	400
C845 F Set 9(5)	CCATGTGCCTCTCTTCCGCG	
C845 R Set 9(5)	TCTCTATATATCTTTAATATATCTTTATTGTATCATCTATATCAC	500
C855 F Set 9(6)	TGAGGATAGTAGGAAGGGTATAGTTTTCATG	
C855 R Set 9(6)	TCAAATTTTCTATCTTTAAAATAGTTCTTAAAATAG	700
C865 F Set 9(7)	AAATATCTCTTAGTTTTTTAGGAATCAGAACTC	
C865 R Set 9(7)	AGTAAAATGAAGATGGAGCTATTTAATAAC	800
C875 F Set 9(8)	ATAGCTGGATAATTGGAAATCC	
C875 R Set 9(8)	ATATTTCTAGCACTTCCTGATAAGCAC	900
C885 F Set 9(9)	TTGATGAAATTATTGAATCAATTAGAACA	
C885 R Set 9(9)	TTATCAAAAAGTTTGCCTCAAGC	1000
C895 F Set 9(10)	ACTAAACATTTATTTGAACTCAAGAAAATC	
C905 F Set 10(1)	CTTCGATAGTTTTTGGTGAAAAAATG	125
C905 R Set 10(1)	TAGGTATAATATATTTTAAAAACAAGGGAGA	
C915 F Set 10(2)	ACCAATTACAGATGACCCATACC	225
C915 R Set 10(2)	TATAATTTATGACAAACTATAAACTTTCATTTATTG	
C925 F Set 10(3)	TTTTTTCAAAATTTTATAACTTTTATTGTTTTAC	325
C925 R Set 10(3)	CGGTCATGGTTTCTTGCC	
C935 F Set 10(4)	TTGTTTCAAATCTTTGGTAGTTTTCTAATATTAG	425
C935 R Set 10(4)	TTCAAATAGATCAACTACTCCAGTTGATC	
C945 F Set 10(5)	ATGTTTTAGTTTTTCTGGGATCATTGAT	525
C945 R Set 10(5)	GGAATTTATCTACACAATATGGCACAAAAC	
C955 F Set 10(6)	ATTTTATAAATTGATCTAAAATATTTATAGATTTAGAAGC	625
C955 R Set 10(6)	AAGATATTAGTAATTAGAAGCTGAAATATAAATTTCATTAC	
C965 F Set 10(7)	GATAAATACGATTGTTTTGTTCTTTG	725
C965 R Set 10(7)	ATAATTCCAAATATAACTATTGGTGTGGTTGTAC	
C975 F Set 10(8)	AGCTGTTGCTTTAGGATATGAGTTTATTTC	825
C975 R Set 10(8)	ATAAACAGTTCCTCGAGCCATAG	
C985 F Set 10(9)	TATTAAAGAAAAAATTTCTCTATTCTGTAAGTGTCC	925
C985 R Set 10(9)	AACAATTGCTTTAATAAATCCTCCTAAATAAAG	
C995 F Set 10(10)	AACCTTTCGTTCTAAAAATCCTAGTTTAAATAAAG	1025
C995 R Set 10(10)	TTTTTCTAATTGGTTCTATGTGAAGTTCTATTTG	
C1005 F Set 11(1)	AAAAATTAAGTTATTAGAATATGGGGTGTCAA	150
C1005 R Set 11(1)	CCAATATTAATTCCACCAATATAACCA	
C1015 F Set 11(2)	TAATAAATTTAATATATAAATGTTGATGTAGATTCTCTTTTTAGTAT	250
C1015 R Set 11(2)	TTGTAACATTCTTGCAAAGTTTACAAAC	
C1025 F Set 11(3)	GTTGTAATAAGACCGTCTCTTTTTTGG	350
C1025 R Set 11(3)	AAAGATTTTAATAAAGGAATTCAAAGTATTTTAGG	
C1035 F Set 11(4)	ATAATAAACTAAACTAGCTCAAGAATATGAATCAG	450
C1035 R Set 11(4)	CTCTAGCAATACCAATTCTTTGACG	
C1045 F Set 11(5)	ATGCTTATTATAATTGCAACTTAAATATATACATAATAGC	550
C1045 R Set 11(5)	TAAAAATATTTTAGTAACTATTAAAGATAAAAAATTTACTTTGG	
C1055 F Set 11(6)	GAACAACAGTTATGTCGCTTGTTG	650
C1055 R Set 11(6)	TTTTTCAGATTCGTAAGTTGTTCTAAATATTTC	
C1065 F Set 11(7)	ATTGGTATAAATTATGGATCTAGTTTAAATCATAC	750
C1065 R Set 11(7)	AGTACCTTGCATTAAAAAGAAAGTTGTTAG	
C1075 F Set 11(8)	GATAATGAAGAGTATTCAATTTCTACTTTATTTTC	850
C1075 R Set 11(8)	ATTTAAATCATAAAAGAGAAGTAATTATTGACAAATTTAC	
C1085 F Set 11(9)	TTTTATCTCCTAATACCATCAGAATCCATATCTAAC	950
C1085 R Set 11(9)	TTAATAAAATAGATAAAAAAGATCAAAGAAGTTTAGAAGTAG	
C1095 F Set 11(10)	TACTAATAGTATTATAAGGGTTTATGATTTTAGTAAATATTTAAAATTA	1050
C1095 R Set 11(10)	GATAAGATGACACAACCATAAAATCC	
C1102 F Set 11(11)	ACAGCTGAAATGTAAGACAAATTAAGAAT	1150
C1102 R Set 11(11)	TAATATCTTAATTGAAGTTCTTGATTATATTCT	

Table S2

Amplicon #	Name	Primer sequence	Amplicon Size (bp)
1	100F	GTAATTGAACCTAATTCTTTTTCTAATC	337
	101R	GGACTTGGTGGAATTAGACATC	
2	200F	TCAGCTTATTTAGCTACAAATTCTG	429
	201R	GAAGAAGATACTTCATGAACAAATG	
3	300F	AGAAGATATTGCAGATGCAGAAG	514
	301R	AGTTGCATTGCTTGAACTAGTTG	
4	400F	AAACTAGACAAAATGAAGATGGAAG	589
	401R	CTTCATCATCTTCATATCAAGGAC	
6	500F	AACCATCTGCACCAGATAGTTC	726
	501R	AGTGGTATATTTAGTTTAGCAAAACC	
8	600F	TAGCTGTTTGCTTGCTAAGGTC	846
	601R	TGGGTTTGTATTTAGTAGTAGTGC	
9	700F	GAAGTTAGTGACGGATATGCAAG	952
	701R	ATTATATCCAGTTGGAACTCCTG	
5	799aF-2	GTCAAGTTCTTTTCATACCACTAC	685
	811aR	TTGTTCATTACTTGCACCGATTAC	
7	900F-2	GAAGTATGATTTCCAGAACAAAAC	784
	901R-2	AACTAGCTCCGTGTTGCTTTG	
10	1000F-2	TGTAGATCTGCCAAGTAAGTCTC	1010
	1001R-2	CCTGTAATTTGTTTGATTGCTTG	
11	1104F-2	AGGTGTGTTATCTTTGTGATTAGC	1149
	2aR-2	ATGGTCAAGTTCCTGTAGCTAC	

Table S3

Amplicon name	Primer name	Primer sequence	Amplicon size (bp)
WM1	MLC WM1 for	GCTCATTGCTGATCATAATGACTGTTTATATAC	100
	MLC WM1 rev	ATGTATGTCAAAAACGAATCTCCCG	
WM2	MLC WM2 for	GATATTTTCATCCTTTGCAATACAATAACTACTACATC	300
	MLC WM2 rev	CGAGCCAGGCATTAAATTATCTACAG	
WM3	MLC WM3 for	GCTGTAGATACAACGTCGTATTCTGTAAGTG	500
	MLC WM3 rev	GTGCAGCACGGCATATTAGCAG	
WM4	MLC WM4 for	TCTGATACTATAGCAACGTTGCGTGATATTT	700
	MLC WM4 rev	AATGTTCAGGGAACTTATTCAACGG	

Table S4

Assembly	Forward Primer Sequence	Reverse Primer Sequence
2-10	GTTTTTAATTTAACTAATATTTATTACAAATAAAAAACTTgcggccgcgatcctctagagtcgacctg	TTTTCCAGATATAAATGCCGGAAATATTGGTTATAAGATTgcggccgccgggtaccgagctcgaattc
11-20	ACAAAGATATCTGGAGTTTGTTTAGTTAATAAACAGTTTTgcggccgcgatcctctagagtcgacctg	CCATATTCAGAAACTTTAACTTTAGCTCTTTCTTTACCATgcggccgccgggtaccgagctcgaattc
21-30	TCAACAATTACAAGAAAAACAAAAAGTTAAATTTATGTATgcggccgcgatcctctagagtcgacctg	ATGAAGCTAGTAAAAAAGATAAACAAAAAATAACATCATTgcggccgccgggtaccgagctcgaattc
31-40	CAATAGTAGCAATTTGAGAATTTAAATTATCAATTTCGTTgcggccgcgatcctctagagtcgacctg	TCTAAAGACTTAAACTCATATAGTCACATTAATCCTTGGTgcggccgccgggtaccgagctcgaattc
41-50	TTTTAGTTCTCTTAGTCTATAAATTTCAATCATTTCTTATgcggccgcgatcctctagagtcgacctg	AATATCACTTGATTCTTCATATTTGTTTGTTATAAAACTAgcggccgccgggtaccgagctcgaattc
51-60	GTAGATGAATTATTAAACATATTAAAAAACACTGATTTTAgcggccgcgatcctctagagtcgacctg	GTCAATAATTTAGAACAAACTTATAAAATAGCAAAAAAAAgcggccgccgggtaccgagctcgaattc
61-70	TTTAACTTTCATATATAACCTCATTTACTTTAATTATAAAgcggccgcgatcctctagagtcgacctg	AAAAAATACGGAAAAGAATTTATTGATCTTTTAACATCTTgcggccgccgggtaccgagctcgaattc
71-80	AATTAAATCAAAATTCATTTTGTTTATTTTTTACTAATTCTgcggccgcgatcctctagagtcgacctg	ATAAAATTGATTTACCTAGTGCTGATGTTGATAAAGTTAAgcggccgccgggtaccgagctcgaattc
81-90	TAATTACTGGAATTATTTCTAAATTACTATCAATTGCTAAgcggccgcgatcctctagagtcgacctg	ATTATTTTCTTTAACTAATAGATCTTCATCTCTAGTTTTTgcggccgccgggtaccgagctcgaattc
91-100	AGAACTGATAGAGTTGGTAACTTTAGAAAATATGAAAGAGgcggccgcgatcctctagagtcgacctg	TAACATTTATCTTTAATTTTTCATTAACATTAGCAATTTgcggccgccgggtaccgagctcgaattc
101-110WM3	TCAAACAGCAAAGCAAATGTATGGAGAAATGCTTCCAGAAgcggccgcgatcctctagagtcgacctg	CATTAGAGTTTTTATATTTATATTGAGTTTGTCAAATTCCgcggccgccgggtaccgagctcgaattc
111-120	TAATAGTTGATGCTATTATTAAAACACAAATTAATGATAAgcggccgcgatcctctagagtcgacctg	ATCTTTATTAAAAGTACTAAAGAGTTTAATAAATTAAAATgcggccgccgggtaccgagctcgaattc
121-130	TTCTGTATTATAAATTTGACGTTTTTGTATTGAGTCTCACgcggccgcgatcctctagagtcgacctg	GTTCTACTTTAACTATTTTACTTTTTTTAAGTAGTTATTTgcggccgccgggtaccgagctcgaattc
131-140	TTGAAAACGATTTTAAAATTGGTGAATCAATATAATTTGTgcggccgcgatcctctagagtcgacctg	GCTTTATTTGCTTTTATAGGTAGTTTATTACAAAAAAAGAgcggccgccgggtaccgagctcgaattc
141-150	TACTAATATAGCCGGTGTTGAAATAAATTCTGTAAAGAAAgcggccgcgatcctctagagtcgacctg	TAGTTTATATAGTTTAGATTTTAGTGTTGATGCTAAAAAGgcggccgccgggtaccgagctcgaattc
151-160	GTTGGATCTGTTATTTTGTTACTGTTATTTAGGTATTTTTgcggccgcgatcctctagagtcgacctg	CTACAGTATAAAACCAGGTATGATGCCATCTACAACTCCgcggccgccgggtaccgagctcgaattc
161-170	TAGAAGAATTTGAATCTTTAAATTTAGGTAATTTTGAAGAgcggccgcgatcctctagagtcgacctg	ATTTGATGATCAAATTAAATTAGATCATAATATTACTTATgcggccgccgggtaccgagctcgaattc
171-180	AATTTTCTATCTATCATAATGAACTAATATCATTACCTTgcggccgcgatcctctagagtcgacctg	TAATCAAGAACTAGAAATAAGTTATATTCCAACTGATTTTgcggccgccgggtaccgagctcgaattc
181-190	ATTAGTTTTAACGCTTTGTCAAATTTATTATTATCAATTAgcggccgcgatcctctagagtcgacctg	TTATTTATATTCAACAAGTTGTTTTGTAAATGATATTAATgcggccgccgggtaccgagctcgaattc
191-200	CTTTGATCAGATGTACTACATACATATAAAATCTTTTTTTgcggccgcgatcctctagagtcgacctg	CTGCTGTTAAACTTGGTATAGGAATTGATTATAAATACCCgcggccgccgggtaccgagctcgaattc
201-210	ATTTATAATGATTATCCTTTAGATGTTTTGGTACATTATAgcggccgcgatcctctagagtcgacctg	GTTTCAAATTCTCCTTTAACAAAACCAACGATTAATTATTgcggccgccgggtaccgagctcgaattc
211-220	ACCTTTTGAATAATTATATTCATTTGCATATTTAATACTTgcggccgcgatcctctagagtcgacctg	ACTACTATTAATAAACTTGTTGAACTAACTAAAATTTCAGgcggccgccgggtaccgagctcgaattc
221-230	TTCAACTAAAATTTCATTATTAAGATCAAAAATAAAATTTgcggccgcgatcctctagagtcgacctg	AGTAATAGTTTTAGTAAGTTCACTAGTAGTTATTTTTAGAgcggccgccgggtaccgagctcgaattc
231-240	AAAATTAATGTTAAAGCTAAAGCAACAAAATAAGTTCATTgcggccgcgatcctctagagtcgacctg	ATTTAGATAGATATGAGTAATTTTTTTGAAGCTGTTAAAGAgcggccgccgggtaccgagctcgaattc
241-250	CATTTAGAAATGTAGTTTCATTTAATATTACTAAAGCTGAgcggccgcgatcctctagagtcgacctg	GTGGTAGCAACCTATTAGATTAGGTGCTTAAATAAGGGCTgcggccgccgggtaccgagctcgaattc
251-260	TTCATTAAGTTTTATCTCTCCATGGTTCTTAAAGATATgcggccgcgatcctctagagtcgacctg	TTATCTTAGCATCTTGACCTCTAAAAATAGTATCAAATCTgcggccgccgggtaccgagctcgaattc
261-270	AAGCAAATACTAGTATAAATAATAAAGCAGAATTCAAGGAgcggccgcgatcctctagagtcgacctg	TTTAGAACAAAAAATAGGAGATATAGTTCTTATAAAGGAAgcggccgccgggtaccgagctcgaattc
271-280	TTTAATACACTGAACTCCAAGCTCTTTAGTTAACCTTAAAgcggccgcgatcctctagagtcgacctg	TTGTAACATGTAAACTCCTAAAACTAAAATAGCTAATAAGgcggccgccgggtaccgagctcgaattc
281-290WM1	ATGGGGTTATCTTCATATGTAAACTATTTAAAAAACGTGTgcggccgcgatcctctagagtcgacctg	AGATGATTGATTCTTATACCCAGTAATAACTCCACACAATgcggccgccgggtaccgagctcgaattc
291-300	TCATTATTAAATGTTGATCGTTCTGCTGCATTATTTGGTTgcggccgcgatcctctagagtcgacctg	CAAATGAAGATTGAGAAGATTTTATTGCTTTAGATGCTCTgcggccgccgggtaccgagctcgaattc
301-310	CTTTTCTACTAATTTCATCAAATTTATGATTATGTCTTTCgcggccgcgatcctctagagtcgacctg	TGGTTTTGTTTTCTAAAGATTTTAATTTAGCTTCTAATCCgcggccgccgggtaccgagctcgaattc
311-320	ACTAAATTTAATCTATCTCTTCTAACTAAAAACAAAGGCTgcggccgcgatcctctagagtcgacctg	AAATATGTTTTATATAAATAAAAACGGAGAAGAACAAGATTgcggccgccgggtaccgagctcgaattc
321-330	TAAAAACTTTAAATTATTATTTGTAGGTGCTTCTAAATACgcggccgcgatcctctagagtcgacctg	CTAAATTAATTTCTTTCATAATAATACCTCCTATTAAGGTgcggccgccgggtaccgagctcgaattc
331-340	AAAATACAAAAGAAATTATTGGCGGAGCTGGAGTTAGTGGGcggccgcgatcctctagagtcgacctg	AGATTTTGATAAACTTCTTCAATTGTATGAACTAAAATTGgcggccgccgggtaccgagctcgaattc
341-350	AAATGATAATAATAAAGTTGAATCAGTTCATTTATTAGATgcggccgcgatcctctagagtcgacctg	TATAATCTAGCAACTTCTTCACATAAATCATTTTTACTAGgcggccgccgggtaccgagctcgaattc
351-360	TTCATATGATAAAATTGATGAAGTTCCACTAAGCTTTACTgcggccgcgatcctctagagtcgacctg	ATTAGGTTTTACCCTAACTTCAACCTGCTCATGGCTAGATgcggccgccgggtaccgagctcgaattc
361-370	GGAGGACCGAACCAGTATTCGTTGAAAAGACTTTGGATGAgcggccgcgatcctctagagtcgacctg	AAAGTAACTTCAATTTGTGGAATACCTCTAGGAGCTGGTTgcggccgccgggtaccgagctcgaattc
371-380	TGAAATTGATGCTAATGGTATTGTAAGTGTTTCAGCAAAAgcggccgcgatcctctagagtcgacctg	GTTTGATTATTCTAGAAATCAAAAAGATGCTAAGTTTAAAgcggccgccgggtaccgagctcgaattc
381-390	ATATATTTAGTTGTAATCTTATCTTTATTATTTATAAATTgcggccgcgatcctctagagtcgacctg	TATGTTTTTTGAAGCAACTTCATTTAACCAAGATATATCAgcggccgccgggtaccgagctcgaattc
391-400	TATGACATATCAGTTACATTTGAAACATTTCATTTACCAAgcggccgcgatcctctagagtcgacctg	GTTGATTTTTTAGTATTTTCAGCAACTATCATCTTATCAAgcggccgccgggtaccgagctcgaattc
401-410	TCAACTAAAATCTTTTCAAAAAAGATTTTGTGTTTTTTTgcggccgcgatcctctagagtcgacctg	TTATAAATAAAAATTCGATTTCTCAAAATGTTCAAATTGAgcggccgccgggtaccgagctcgaattc
411-420	AACTCTTTGCTAAATCTTCTTTTTCTTTAGCTTTTTGTTTgcggccgcgatcctctagagtcgacctg	AGTGGTTTTTTGTTTAAATCTGGTTGATTTGGTGGTTGgcggccgccgggtaccgagctcgaattc
421-430	AGATCCAACTGAAAGAAGAGATCAAAACTCAGTTGAATATgcggccgcgatcctctagagtcgacctg	TCCAATTTGATATGATAAGAAATTTCCTTTTTCTAATTCCgcggccgccgggtaccgagctcgaattc
431-440	TATAGATCATTTATTGATAAACCTGGATTAACTGATGGATgcggccgcgatcctctagagtcgacctg	CCGATTCTATGAACTGCTAAAAATGCATCAGTTAGTCCTGgcggccgccgggtaccgagctcgaattc
441-450WM2	TAATGATTTATACACACTTTATGATAAAAATAAAAAAGAAgcggccgcgatcctctagagtcgacctg	TGAATTTAATTGTTCTGATTTCATTTCACACATTACACTTgcggccgccgggtaccgagctcgaattc
451-460	CATAATAAAAAGAAAAATAGTGATCATTATAAAAAATTAGgcggccgcgatcctctagagtcgacctg	AATTCAAAAGACAAAAATTTAGGTAATTTTGATAATATTGgcggccgccgggtaccgagctcgaattc
461-470	AAGAACTTGATCAACATTAGTTTTTCTGAAGCTAATTTCTgcggccgcgatcctctagagtcgacctg	CCTTTTAAATGAGGTAATACTAAGCTAACTGCTTTAGCAGgcggccgccgggtaccgagctcgaattc
471-480	TAAATTAGATGGATATGCATTACGTGTGCCAACTATTACTgcggccgcgatcctctagagtcgacctg	TAGTATCTTTATTAATTTGAGTATTTGGTAAAGAGTTTAAgcggccgccgggtaccgagctcgaattc
481-490	AAGTAATTGATGCTTTAATTGAAGTAAAAGCTGCTAGTTCgcggccgcgatcctctagagtcgacctg	GTTTTTACATCTAATATTACATTAACTGCAATAGTTCCAAgcggccgccgggtaccgagctcgaattc
491-500	TTTAGATGTTTCTAATGGTATTATTCAAACAATTGAAAAAgcggccgcgatcctctagagtcgacctg	AAAATCAGATAGTTTTATAGTAAATAAATTCTTAAGATTTgcggccgccgggtaccgagctcgaattc
501-510	ATTAAATTAGAAAATAAACCAAATTTTGGTTCTGATTATTgcggccgcgatcctctagagtcgacctg	AAATAATTAATCCATTTGTTAATTGGATCATTGTTGATTTgcggccgccgggtaccgagctcgaattc
511-520	CTGAAACTGGTCAAATTATAGTTGGTGATTATGCTATTCCgcggccgcgatcctctagagtcgacctg	TGGGAAAATAATTTGTTCTTTAATTCCAGTTGTAAAATTTgcggccgccgggtaccgagctcgaattc
521-530	GAAGTAGATTATGATAAAGTTATTAGATTACGTGGTATGGgcggccgcgatcctctagagtcgacctg	TGAATCATTACAGCATTATACATTGTATGTAATCCAGTAAgcggccgccgggtaccgagctcgaattc
531-540	AAACACAGCAACAATTGGTGGTTCATTTATTTTCCCAATTgcggccgcgatcctctagagtcgacctg	TCTAAACTAGTTAAAATAATATTAGCATCTTCACTAGTTAgcggccgccgggtaccgagctcgaattc
541-550	AATGACTAATTTTTTTGAAAAAACTATTAAATCAACTACTgcggccgcgatcctctagagtcgacctg	AAAAATGCCATAGCATCTTTAATTACTTTTCTTTCTCTAAgcggccgccgggtaccgagctcgaattc
551-560	AAAAATGATTTCAATTAAAGATGACTTGTCATCTCAAAGAgcggccgcgatcctctagagtcgacctg	TTAACAAATTGTACTAGTTGTTCATTAAAATCGTGTTTTTgcggccgccgggtaccgagctcgaattc
561-570	TAAAGATTCACATAACTATATAACTTTAGATAGTTTAATgcggccgcgatcctctagagtcgacctg	ATATTAAACAATTTTATTTATTAGATGATAATTGCggccgccgggtaccgagctcgaattc
571-580	TATCTTTACATCATCTTCTTTCATGAGTTTTTGATCTATTgcggccgcgatcctctagagtcgacctg	AGGATATGTTGATGATTTAAAAACTGTAATAACTAATAAAgcggccgccgggtaccgagctcgaattc
581-590	CAAATTTTAATATCACTATTTTTATATTTATCACTAGTTAgcggccgcgatcctctagagtcgacctg	GTATTCAATCCAATGTTTAAAACCTGCTTAAAATTTTCAAgcggccgccgggtaccgagctcgaattc
591-600	TAATGAATTTGAAAATATAGATGATCTTGATTTTTTAACTgcggccgcgatcctctagagtcgacctg	AAAATATCTCTAATTACTAAATCTTTATCTAATTCATTTAgcggccgccgggtaccgagctcgaattc
601-610	AAAATTTATGTAATTTATTAATTTTTATCTTTATAATATAgcggccgcgatcctctagagtcgacctg	TAAACGCATCATAATGTCCAGTAGTATGAGCTATTTTATCgcggccgccgggtaccgagctcgaattc
611-620	TTAAAAAACAAATTCCAGCTAAAAAAATAAAACTTTATTgcggccgcgatcctctagagtcgacctg	ACCAAATACTAATGCAGTTTTATTAAATGGCCCTCCTAAAgcggccgccgggtaccgagctcgaattc
621-630	ACTGTTGTGTTTTATCAATCATTAGCTCAAGTGTTATCAAgcggccgcgatcctctagagtcgacctg	GGTTCAATATAGTTTTTTAAATATAATTCAGCTGCTTTAAgcggccgccgggtaccgagctcgaattc
631-640	TAGAATGAAAAAATTTTCTAGTCTAAAAGAAAGATTAGATgcggccgcgatcctctagagtcgacctg	TTTCTTCCTTCAACTAATTGATATTATGGATTAAGTTTAGgcggccgccgggtaccgagctcgaattc
641-650	TAATAAATCAACCTTACTTAACTCATCACATAATTGCTCTgcggccgcgatcctctagagtcgacctg	TTAAAGGAGTAGGTAGCTCTCTATCTTGTTTATCTACTAAgcggccgccgggtaccgagctcgaattc
651-660	AACTAGCATAATAAAATAAAAACCTATACTAGTTCTAAATgcggccgcgatcctctagagtcgacctg	ATCCAGATGAACCTCCATCTAAAGCTTTATGATCTTCATTgcggccgccgggtaccgagctcgaattc
661-670	CAATGGCTATTGATTCTTCATTTAATTTAATTGGTATTAAgcggccgcgatcctctagagtcgacctg	TCTTTTTTAACTCATTTATAAATACAGCTTGCATGATTGTgcggccgccgggtaccgagctcgaattc
671-680WM4	TAAAAAAAGGTTGTTTATATTTGAAGATGATCTTAGTTTTgcggccgcgatcctctagagtcgacctg	TACAAATGGTTTGTTAATTCACTCTGATCAGGGATTTCATgcggccgccgggtaccgagctcgaattc
681-690	TTTTTAACCTTTTTAAGTGCTTTGTCTAGCATTGAATTTGgcggccgcgatcctctagagtcgacctg	AATATAGTAATAATCACATAAAATTCTAAATTTTTGTCATgcggccgccgggtaccgagctcgaattc
691-700	GATCTAAACACTCCTTTAAAAGATTTGACTCAAAAACAAAgcggccgcgatcctctagagtcgacctg	ATGCTTTAGCTGACATTGTTTCAGTTTTAAAAGTTGATGAgcggccgccgggtaccgagctcgaattc
701-710	TAGAAAAACTCATAGCTACACCTAAAACAAAACGTTCAGCgcggccgcgatcctctagagtcgacctg	ATTTTGCTAAAAAAGATAAAAACGGTTCTATAAGTTTTGAgcggccgccgggtaccgagctcgaattc
711-720	TAAAAGTAGCATATTCTCCATAATCTTTAATTGCACCAAAgcggccgcgatcctctagagtcgacctg	ATTTATGATGGAGAGGTTCTAGGTTTAGTTGGTGAGTCAGgcggccgccgggtaccgagctcgaattc
721-730	ATTTAAATTTAGATCTTTAATAGCAGTAAACTCTTTTGCAgcggccgcgatcctctagagtcgacctg	AAATTCATCTTGGAGTCACTGATGATGGTCATATTGACAAgcggccgccgggtaccgagctcgaattc
731-740	CTCCACCTTTAGTATTTAAAAAAGCTACAATTTCTCTTTTgcggccgcgatcctctagagtcgacctg	CGTTTTTTCTAGCATAAATAATATCTCTAGATCTAGAGTTgcggccgccgggtaccgagctcgaattc
741-750	CTAATGACATTAACACTAATTTTGTTGAAGATGAAAACGAgcggccgcgatcctctagagtcgacctg	TCTTCATAGTTCATATTATTTACTCTCTTTTAAAATTTGTgcggccgccgggtaccgagctcgaattc
751-760	ACTTGAAATTGGAGATATTATTGAATTAAAAAAACCTCATgcggccgcgatcctctagagtcgacctg	GAAATTCATGGAAATAATAATAAGCTTAATGATGTTGGAAgcggccgccgggtaccgagctcgaattc
761-770	TGGTTTAGAATCTAATACTTTAGCTAATGAGATAAGAATAgcggccgcgatcctctagagtcgacctg	ATAGTTTCTAAGTTATAACCATAAACATAAAAACTGTTGTgcggccgccgggtaccgagctcgaattc
771-780	GTTGGCTGAAAAAATTCATGCAATCAAAGAATTAAGCTTGgcggccgcgatcctctagagtcgacctg	AATTTCACCATTAAAAAAAGTATGATTTGCATCATTTTTAgcggccgccgggtaccgagctcgaattc
781-790	AGAAAACTTCAAGATAAGTTAGTTTTATCAAATACTAAACgcggccgcgatcctctagagtcgacctg	TAATAACCTCTAAATTAACGTCACATTCTAATTTAATTTTgcggccgccgggtaccgagctcgaattc
791-795	TGAAAAAATTAGTGGCATTACTTGCTGCTATTAGTGTGTTgcggccgcgatcctctagagtcgacctg	TACTCTTCCCATTTCTCTTAGATCATTACTATTAGTTCTAgcggccgccgggtaccgagctcgaattc
796-799	ATTTCAACAGTTTTACAATGCCTAGATAATTTAAATAAAAgcggccgcgatcctctagagtcgacctg	TTTTAACAAGTGTTTAACTATAATATTTTTGGAGACAAATgcggccgccgggtaccgagctcgaattc
811-820	ATGTGGAAAACGTGGAAAAAATCCTTATAACATAGATATAgcggccgcgatcctctagagtcgacctg	CTCTTGATGTCTTACTAAATGCAACCATAGCTAGTATTGTgcggccgccgggtaccgagctcgaattc
821-830	CACCAGCTGCAGTTGGTGTACCTTTTGATAAAGCAAAAAGgcggccgcgatcctctagagtcgacctg	TTACAGTAATAAGAGTGAAGAAAATAATACAATTATAGCGgcggccgccgggtaccgagctcgaattc
831-840	TCTAATTTTAGTGACTGTGTATCTCTATATATATGATTCTgcggccgcgatcctctagagtcgacctg	TGATTACGAATAAATTTTATTCATCAACACATCGAGGTCCgcggccgccgggtaccgagctcgaattc
841-850	CTTCTATAAGGTTATCAATCCCTATATTGTTTTTTGCACTgcggccgcgatcctctagagtcgacctg	TTCATCAGTGATTTTGAATCCTTGGTTTGATTCAGATTGTgcggccgccgggtaccgagctcgaattc
853-860	TATTAAAAAAATTGTTGTTTGTTGTGAAGCAGGAGTAGGAgcggccgcgatcctctagagtcgacctg	AATTGATAATTCATAGTTTGGATCACTTGATGATAATCCAgcggccgccgggtaccgagctcgaattc
861-870	GAAAATAAAAAAGATGCAGTTTCTATTTTAAATGTCTTTAgcggccgcgatcctctagagtcgacctg	ACTAAATGAGAAGATCAACCAATTCAAATTATTAATAATAgcggccgccgggtaccgagctcgaattc
871-880	ATGTTTATTTACATATTCAGTTTTTAAATAATCTGAATCAgcggccgcgatcctctagagtcgacctg	AAATACATAAAAAAGTGCCTTGTATTTGTAGTCAAACTATgcggccgccgggtaccgagctcgaattc
881-890	GTTCAAAAACAACATCTGCTTGTTGGTATCAAGATTTATTgcggccgcgatcctctagagtcgacctg	AGTATTTAGAATTCCAATTACATTAAATGATGAATCAATAgcggccgccgggtaccgagctcgaattc
891-900	GAAGTTCAAGATGTTCCTGGAAGTAAATATATTTCAATTCgcggccgcgatcctctagagtcgacctg	TATGCCAATTATTAAATTTAGTGGACTAGATAAAGAACAAgcggccgccgggtaccgagctcgaattc
901-910	ATATTATTCTTCCTTTTTTCTATGTAATTTTATTACAAAAgcggccgcgatcctctagagtcgacctg	TTTATTAGACCAAACAATAGAAATATCTGGTGTGTTTAACgcggccgccgggtaccgagctcgaattc
911-920	AACTTATTGTTAAAGTCATTAACATAAAACTTAATATTTTgcggccgcgatcctctagagtcgacctg	AATTGGAATCAAATTATCATTGATTTTAAATTCTAGTTCTgcggccgccgggtaccgagctcgaattc
921-930	AAATTTAAAACTAGTGACAATATAACAATTTCAGCACTTAgcggccgcgatcctctagagtcgacctg	ATCTCTTTTCTTCTTTTTCAATTTGTTCTTTAGTTGGATTgcggccgccgggtaccgagctcgaattc
931-940 (94D)	TAGAAGATTGAGTTAAGGAATACCTAGATAAAAATAAAAGgcggccgcgatcctctagagtcgacctg	AGATTTTGTTATTTCTGTAGTTTTTGTAGCggccgccgggtaccgagctcgaattc
941-950	TTAGAACCACTTAATGCAGAAAAAAACAAAATTGACAAACgcggccgcgatcctctagagtcgacctg	TGTTGGAAATGTTGATTTTAATAAAAATGTATATCAAGCAgcggccgccgggtaccgagctcgaattc
951-960	GGAAGATTAAATTTTCCTTGTTGATTAAATCTAATCATATgcggccgcgatcctctagagtcgacctg	TAACACTTCCGACTAAATAATTATTTTGTTTATTCATACTgcggccgccgggtaccgagctcgaattc
961-970	ATGAAGCTATTAGTTATAAAGCAAATACTTTTATGATCTTgcggccgcgatcctctagagtcgacctg	AAAATCTGAGCTTATGTTAAATCCGATGAGTGATTTATCTgcggccgccgggtaccgagctcgaattc
971-980	ATTCCAAATCTTGAAAGTGTTGCTCTAATTTCAGGATCTAgcggccgcgatcctctagagtcgacctg	TTTTAGTGATATAGATTTGTTGTAATTTAAAAAAAACTTgcggccgccgggtaccgagctcgaattc
981-990	ACAAAATAAAAGTTCAGATTTAATCTGAACTTTATAATTAgcggccgcgatcctctagagtcgacctg	CAACTAAAGAAAATAAATTTCCCTTTATTGCAAAACTTACgcggccgccgggtaccgagctcgaattc
991-1000	GAACTGGTAAGTTGAAATTATATGTATAATAAGCTTTTGTgcggccgcgatcctctagagtcgacctg	ATTTCTTTACAAAATATTACTTTTTTTGAATATGAAAAAAgcggccgccgggtaccgagctcgaattc
1001-1010	ACTTGTAACAAACATTAAAAAGATTTGTACAAAAATAATTgcggccgcgatcctctagagtcgacctg	TTATAATAAATATTACAATTTTAATAAATTACCACAAGCAgcggccgccgggtaccgagctcgaattc
1011-1020	ATTGTTGCACCAACAAATCCCATTATAATTCCACCAATGgcggccgccgcgatcctctagagtcgacctg	GAAAACTTTGAAATGCCAGCATTTATGATCAACAATGCATgcggccgccgggtaccgagctcgaattc
1021-1030	TTCGTTTTTTTCTGCGTTTTTTCAAACCTAGTAGTATTTAgcggccgcgatcctctagagtcgacctg	GTACATATCTTTAGAAATGCTGGAGTAAGAATTATAAGAAgcggccgccgggtaccgagctcgaattc
1031-1040	ATAATAAGTTTTATTTGAACTAGTTCCTTTAGCTAAAGCAgcggccgcgatcctctagagtcgacctg	ATATCTTTTATGACCTAAATTAGCAACAAAATCAACTTCAgcggccgccgggtaccgagctcgaattc
1041-1050	TACATTCAATATGCAGATAATATTAATACTAAAGAAAAAAgcggccgcgatcctctagagtcgacctg	CGAAATTTCAGGATTCCCATTAGCAAATCTTTCTGGAAAAgcggccgccgggtaccgagctcgaattc
1051-1060	CCAAAAAATGTTTTAGCTTGAAATTCAGAAGCTCCAGCTGgcggccgcgatcctctagagtcgacctg	AATTACTATTAGATAATCAAAGAGACTGAACAATGGCTACgcggccgccgggtaccgagctcgaattc
1061-1070	GTGGTAATAAAACATCACCAAATGGTCCTATAAATGATCAgcggccgcgatcctctagagtcgacctg	TCAATATTAAAAAGCTTTTTAATTATTTAAAATAGATTCAgcggccgccgggtaccgagctcgaattc
1071-1080	TTTTAACAATAATATTAAACTTTATAATACTGATTAAATTgcggccgcgatcctctagagtcgacctg	AGCACGTGCTTCTTTCTTAAACGTTAATATTATATATATTgcggccgccgggtaccgagctcgaattc
1081-1090	CACCATTAGCTATTAGTGAAATGGTAATAACTTTTTTTATgcggccgcgatcctctagagtcgacctg	CAACAGTTAACCAATTAGTAGTTGGATTTGTTATATTAATgcggccgccgggtaccgagctcgaattc
1091-1100	GTTTTAGACCAATTATTAAAAACGGTGAAATAATTTTTTCgcggccgcgatcctctagagtcgacctg	GTTTTAGGACATATTCAACGTGGTGGTAGACCAACTGCTAgcggccgccgggtaccgagctcgaattc
1101-1104	AGTTGCTTTAGTTTCATAACCACTAATTTTTTTGAATTTTTgcggccgcgatcctctagagtcgacctg	CAAATCTGTAATTGTATTTAAAAACTCTTAAAAAACTAGA *gcggccgccgggtaccgagctcgaattc*

Table S5

Assembly	Primer Name	Primer Sequence
2-100	pRS-2aF	GTTTTTAATTTAACTAATATTTATTACAAATAAAAAACTTgcggccgcgcaaggcgattaagttggg
	pRS-100R	TAACATTTATCTTTTAATTTTTCATTAACATTAGCAATTTgcggccgcacatccccccttcgccagc
101-200WM3	pRS-101F	TCAAACAGCAAAGCAAATGTATGGAGAAATGCTTCCAGAAgcggccgcgcaaggcgattaagttggg
	pRS-200R	CTGCTGTTAAACTTGGTATAGGAATTGATTATAAATACCCgcggccgcacatccccccttcgccagc
201-300WM1	pRS-201F	ATTTATAATGATTATCCTTTAGATGTTTTGGTACATTATAgcggccgcgcaaggcgattaagttggg
	pRS-300R	CAAATGAAGATTGAGAAGATTTTATTGCTTTAGATGCTCTgcggccgcacatccccccttcgccagc
301-400	pRS-301F	CTTTTCTACTAATTTCATCAAATTTATGATTATGTCTTTCgcggccgcgcaaggcgattaagttggg
	pRS-400R	GTTGATTTTTTTAGTATTTTCAGCAACTATCATCTTATCAAgcggccgcacatccccccttcgccagc
401-500WM2	pRS-401F	TCAACTAAAATCTTTTCAAAAAAGATTTTGTGTTTTTTTTTgcggccgcgcaaggcgattaagttggg
	pRS-500R	AAAATCAGATAGTTTTATAGTAAATAAATTCTTAAGATTTgcggccgcacatccccccttcgccagc
501-600	pRS-501F	ATTAAATTAGAAAATAAACCAAATTTTGGTTCTGATTATTgcggccgcgcaaggcgattaagttggg
	pRS-600R	AAAATATCTCTAATTACTAAATCTTTATCTAATTCATTTAgcggccgcacatccccccttcgccagc
601-700WM4	pRS-601F	AAAATTTATGTAATTTATTAATTTTTATCTTTATAATATAgcggccgcgcaaggcgattaagttggg
	pRS-700R	ATGCTTTAGCTGACATTGTTTCAGTTTTAAAAGTTGATGAgcggccgcacatccccccttcgccagc
701-799	pRS-701F	TAGAAAAACTCATAGCTACACCTAAAACAAAACGTTCAGCgcggccgcgcaaggcgattaagttggg
	pRS-799aR	TTTTAACAAGTGTTTAACTATAATATTTTTGGAGACAAATgcggccgcacatccccccttcgccagc
811-900	Mmyc811aF	ATGTGGAAAACGTGGAAAAAATCCTTATAACATAGATATAgcggccgcgatcctctagagtcgacctg
	Mmyc900R	TATGCCAATTATTAAATTTAGTGGACTAGATAAAGAACAAgcggccgccgggtaccgagctcgaattc
901-1000 (94D)	pRS-901F	ATATTATTCTTCCTTTTTTCTATGTAATTTTATTACAAAAgcggccgcgcaaggcgattaagttggg
	pRS-1000R	ATTTCTTTACAAAATATTACTTTTTTTGAATATGAAAAAAgcggccgcacatccccccttcgccagc
1001-1104	pRS-1001F	ACTTGTAACAAACATTAAAAAGATTTGTACAAAAATAATTgcggccgcgcaaggcgattaagttggg
	pRS-1104R	CAAATCTGTAATTGTATTTAAAAACTCTTAAAAAACTAGAgcggccgcacatccccccttcgccagc

Table S6

Assembly	Primer Name	Primer Sequence
2-100	Mmyc2aF	GTTTTTAATTTAACTAATATTTATTACAAATAAAAAACTTgcggccgcgatcctctagagtcgacctg
	Mmyc100R	TAACATTTATCTTTTAATTTTTCATTAACATTAGCAATTTgcggccgccgggtaccgagctcgaattc
101-200WM3	Mmyc101F	TCAAACAGCAAAGCAAATGTATGGAGAAATGCTTCCAGAAgcggccgcgatcctctagagtcgacctg
	Mmyc200R	CTGCTGTTAAACTTGGTATAGGAATTGATTATAAATACCCgcggccgccgggtaccgagctcgaattc
201-300WM1	Mmyc201F	ATTTATAATGATTATCCTTTAGATGTTTTGGTACATTATAgcggccgcgatcctctagagtcgacctg
	Mmyc300R	CAAATGAAGATTGAGAAGATTTTATTGCTTTAGATGCTCTgcggccgccgggtaccgagctcgaattc
301-400	Mmyc301F	CTTTTCTACTAATTTCATCAAATTTATGATTATGTCTTTCgcggccgcgatcctctagagtcgacctg
	Mmyc400R	GTTGATTTTTTAGTATTTTCAGCAACTATCATCTTATCAAgcggccgccgggtaccgagctcgaattc
401-500WM2	Mmyc401F	TCAACTAAAATCTTTTCAAAAAAGATTTTGTGTTTTTTTTgcggccgcgatcctctagagtcgacctg
	Mmyc500R	AAAATCAGATAGTTTTATAGTAAATAAATTCTTAAGATTTgcggccgccgggtaccgagctcgaattc
501-600	Mmyc501F	ATTAAATTAGAAAATAAACCAAATTTTGGTTCTGATTATTgcggccgcgatcctctagagtcgacctg
	Mmyc600R	AAAATATCTCTAATTACTAAATCTTTATCTAATTCATTTAgcggccgccgggtaccgagctcgaattc
601-700WM4	Mmyc601F	AAAATTTATGTAATTTATTAATTTTTATCTTTATAATATAgcggccgcgatcctctagagtcgacctg
	Mmyc700R	ATGCTTTAGCTGACATTGTTTCAGTTTTAAAAGTTGATGAgcggccgccgggtaccgagctcgaattc
701-799	Mmyc701F	TAGAAAAACTCATAGCTACACCTAAAACAAAACGTTCAGCgcggccgcgatcctctagagtcgacctg
	Mmyc799aR	TTTTAACAAGTGTTTAACTATAATATTTTTGGAGACAAATgcggccgccgggtaccgagctcgaattc
811-900	Mmyc811aF	ATGTGGAAAACGTGGAAAAAATCCTTATAACATAGATATAgcggccgcgatcctctagagtcgacctg
	Mmyc900R	TATGCCAATTATTAAATTTAGTGGACTAGATAAAGAACAAgcggccgccgggtaccgagctcgaattc
901-1000 (94D)	Mmyc901F	ATATTATTCTTCCTTTTTTCTATGTAATTTTATTACAAAAgcggccgcgatcctctagagtcgacctg
	Mmyc1000R	ATTTCTTTACAAAATATTACTTTTTTTGAATATGAAAAAAgcggccgccgggtaccgagctcgaattc
1001-1104	Mmyc1001F	ACTTGTAACAAACATTAAAAAGATTTGTACAAAAATAATTgcggccgcgatcctctagagtcgacctg
	Mmyc1104R	CAAATCTGTAATTGTATTTAAAAACTCTTAAAAAACTAGAgcggccgccgggtaccgagctcgaattc

Table S7

Supplementary Table Legends

Table S1. Differences between *M. mycoides* JCVI-syn1.0 and the natural genome YCpMmyc1.1. The differences are divided into two groups: 1) "designed differences" - 25 differences between the synthetic genome design and the natural YCpMmyc1.1 genome, and 2) "not designed" - 10 observed differences between the sequenced genome arising from the transplanted clone sMmYCp235-1 and YCpMmyc1.1. The differences are classified by type (column 1). The coordinate of each difference on the designed *M. mycoides* JCVI-syn1.0 sequence is indicated (column 2). The actual sequence differences on the synthetic and the natural genomes are listed for snps, and for homopolymer and dinucleotide runs; for watermarks the name of the watermark is given and the length of the substituted *M. mycoides* sequence is indicated; the size is of deletions and insertions given (columns 3 and 4). The cassettes affected by the difference and a search string to locate each difference is also shown (columns 5 and 6). The comment in column 7 indicate the gene(s) affected by the difference.

Table S2. Multiplex PCR primers used to identify 100-kb intermediates containing all first-stage assemblies.

Table S3. Multiplex PCR primers used to identify complete assembled genomes. We refer to this primer set as TSS2. Each primer set crosses one of the eleven 100-kb junctions.

Table S4. Multiplex PCR primers used to identify genome assemblies containing watermark sequences. We refer to this primer set as WM1-4. Each primer set produces an amplicon in one of the four watermark sequences.

Table S5. Primers used to produce unique first-stage assembly vectors. Overlaps to the ends of the cassette sequences are shown in upper case.

Table S6. Primers used to produce unique synthetic second-stage assembly vectors. Overlaps to the ends of the 10-kb assembly sequences are shown in upper case.

Table S7. Primers used to produce unique vectors for cloning 100-kb natural *M. mycoides* fragments. Overlaps to the ends of the 100-kb natural sequences are shown in upper case.