Chemical synthesis of the mouse mitochondrial genome

Daniel G Gibson[1], Hamilton O Smith[2], Clyde A Hutchison III[2], J Craig Venter[1,2] & Chuck Merryman[1]

We describe a one-step, isothermal assembly method for synthesizing DNA molecules from overlapping oligonucleotides. The method cycles between *in vitro* recombination and amplification until the desired length is reached. As a demonstration of its simplicity and robustness, we synthesized the entire 16.3-kilobase mouse mitochondrial genome from 600 overlapping 60-mers.

Chemical synthesis of long DNA sequences that encode various genetic functions has steadily improved since the synthesis of the 77-base-pair (bp) gene encoding a yeast alanine transfer RNA[1]. It is now possible to produce 1-kilobase (kb) genes, 30-kb biosynthetic pathways and even entire 1-Mb chromosomes from chemically synthesized DNA[2,3]. Because one has complete control over the sequence of chemically derived DNA molecules, genetic components can be exhaustively optimized. For example, one aim with medicinally important products is to improve heterologous expression in industrial hosts by codon optimization[4,5]. More comprehensively, our goal is to create synthetic organisms whose entire genetic makeup is tailored for efficient production[2,3]. Regardless of the motive, extensive re-engineering of most genetic elements relies on technology that enables the assembly of small synthetic oligonucleotides. Here we describe a one-step *in vitro* reaction for assembling DNA oligos directly into a vector, which can then be cloned in *Escherichia coli*. This method is essentially sequence-independent, fast, largely labor-free and compares favorably to previously described methods for gene synthesis[6], which require additional steps (polymerase cycling assembly, PCR, gel purification, restriction digestion and DNA ligation). Furthermore, error correction procedures are not required in this method because only small fragments are synthesized from oligos, which ensures that error-free molecules are obtained at a reasonable efficiency[4]. To demonstrate these features, we synthesized the mouse mitochondrial genome, to our knowledge the first synthetic genome of an organelle to be made, and estimate that one individual could reconstruct the entire 16.3-kb molecule in just 5 d (**Supplementary Fig. 1**).

We recently described a one-step, isothermal *in vitro* recombination system capable of joining overlapping double-stranded DNA molecules up to hundreds of kilobases long[7]. The assembly reaction mixture in this system contains three separate enzymes (T5 exonuclease, Phusion polymerase and *Taq* ligase) that work in harmony to join multiple DNA fragments. In a typical reaction the assembly is accomplished in as few as 15 min. This method is robust and amenable to automation. For these reasons, we adapted it for assembly beginning at the oligo level. We optimized several parameters including the number of oligos used in a single reaction, their length, the amount of overlap, orientation, oligo concentration in the reaction, reaction temperature and reaction time (**Supplementary Tables 1–9** and **Supplementary Note 1**) to maximize efficiency of oligonucleotide assembly into a linearized pUC19 cloning vector.

We synthesized the entire 16,299-bp mouse mitochondrial genome from 600 overlapping 60-mer oligonucleotides. The overall assembly strategy encompassed four subassembly steps (**Fig. 1**, **Supplementary Fig. 2** and **Supplementary Tables 10** and **11**). We purchased 600 oligos from Integrated DNA Technologies. We made no attempt to select oligos with characteristics that might improve assembly efficiency. Thus, beginning at nucleotide 1, we simply ordered 60-nucleotide oligos as dictated by the sequence of the *Mus musculus* mitochondrial genome (GenBank accession number NC_005089), irrespective of potential to mispair or to form hairpins. Each oligo overlapped its neighbor by 20 nucleotides, and assembly intermediates overlapped by 40 bp. Also, we designed the oligos to include restriction sites and vector overlaps at assembly-stage boundaries. The vector overlaps allowed for direct assembly into a vector or supplied a stage-specific universal primer-binding domain for PCR amplification. The restriction sites (NotI, SbfI, AscI and PmlI) allowed for release of the insert from the vector or non-overlapping sequence before the next stage of assembly.

We directly assembled oligos into the pUC19 vector in groups of eight (**Fig. 1**). This produced 75 overlapping subassemblies that we sequenced before hierarchically joining them to produce the desired 16.5-kb product (**Fig. 1**). Owing to the inherent error rate of chemical oligonucleotide synthesis, sequencing is necessary if an error-free version of a large construct is to be obtained at an appreciable rate[8]. To reduce the labor surrounding the preparation and sequencing of the subassemblies, we pooled equal volumes of the 75 first-stage assembly reactions before performing a single *E. coli* transformation. Automated colony picking, plasmid DNA preparation and sequencing[9] of the resulting clones at 8× redundancy (600 clones) deconvoluted the pool so we could

[1]The J. Craig Venter Institute, Rockville, Maryland, USA. [2]The J. Craig Venter Institute, San Diego, California, USA. Correspondence should be addressed to D.G.G. (dgibson@jcvi.org) or C.M. (cmerryman@jcvi.org).

RECEIVED 26 MAY; ACCEPTED 8 SEPTEMBER; PUBLISHED ONLINE 10 OCTOBER 2010; DOI:10.1038/NMETH.1515

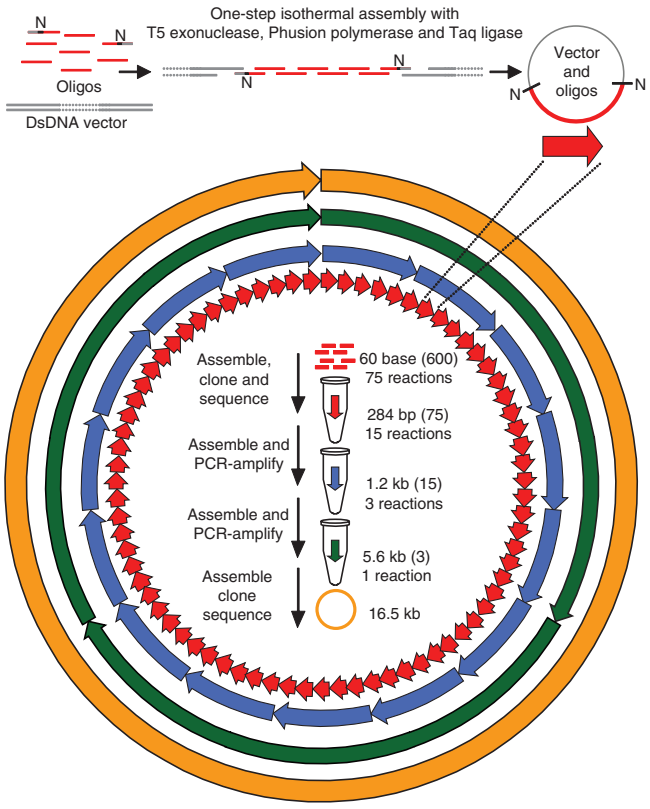

Figure 1 | Schematic demonstrating assembly of the synthetic mouse mitochondrial genome. The 60-base oligos (red lines) were assembled in groups of eight into seventy-five 284-bp cassettes (red arrows). These were joined in sets of five to produce fifteen 1.2-kb assemblies (blue arrows) and then again in sets of five to produce three 5.6-kb assemblies (green arrows). These three fragments were recombined into a complete 16.5-kb genome (orange arrow), which includes a 221-bp repeat. NotI restriction sites (N, black lines) were designed to release the 284-bp cassettes from the pUC19 vector (gray lines).

identify assemblies of perfect sequence (**Fig. 2a**). For comparison, we also individually transformed the 75 assemblies and sequenced eight clones per assembly (**Fig. 2b**). Although this latter approach was more labor intensive, it allowed us to examine collector curves for the two methodologies and validated the utility of the pooled approach. DNA sequencing of the 75-member library resulted in the recovery of a sequence-verified product for 64 out of 75 in the pooled approach (**Fig. 2a**) or 72 out of 75 in the individual approach (**Fig. 2b** and **Supplementary Table 12**) of the first-stage subassemblies. With poor sequencing reads excluded, the overall efficiency for obtaining a correct clone from each individually transformed assembly reaction was 46.8%, with a range of 0–86% (**Supplementary Table 12**). Overall synthesis error rates for each approach shown in **Figure 2** were identical at 0.00308 (1 error per 325 bp). Synthesis error rates for each of the 75 assemblies were 0–0.00880 errors per base pair (**Supplementary Table 12**).

After automated sequencing and recovery of error-free clones, we pooled the 284-bp assemblies (340 bp minus two 20-bp vector overlaps and minus two 8-bp NotI sites) in groups of five, NotI-digested them and then joined them to form 1.2-kb assembly intermediates (**Fig. 1** and **Supplementary Fig. 3a**). Instead of propagating the assemblies in a host organism, we PCR-amplified the assembly products (**Supplementary Fig. 3b**) and pooled these 15 PCR products in groups of five, digested them with SbfI and then joined them to form 5.6-kb assembly intermediates (**Supplementary Fig. 3c**). We PCR-amplified the assembly products again and pooled the three PCR products before digestion with AscI and then let them join to form the complete synthetic mouse mitochondrial genome (**Supplementary Fig. 3d**,**e**). To produce a pure clone containing the genome, we carried out the assembly reaction in the presence of a bacterial artificial chromosome (BAC) and transformed it into *E. coli*. We obtained approximately 20,000 colonies and screened ten. Four of these had the expected restriction pattern, and we sequenced them in both directions with the 60-mer oligos used to build the synthetic genome. One of these contained the intended sequence (**Supplementary Table 13**). The vector was flanked by a synthetic, 221-bp repeat (**Supplementary Fig. 2** and **Supplementary Table 10**). Therefore, when we released the genome insert from the vector by PmlI digestion, we circularized it to form the natural mitochondrial genome sequence without the intervening BAC (**Supplementary Fig. 4**).

The use of PCR to generate the various assembly intermediates was exceptionally fast and reduced labor. However, PCR can introduce errors. In the production of the error-free genome described above, there were 45 cycles of amplification. From sequence-confirmed initial assemblies to the full-length synthetic genome, the synthesis error-rate was 3.78×10^{-4} errors per base pair (1 error per 2,643 bp; range, 0.0 to 6.14×10^{-4} errors per base pair). These error rates explain why only one of four clones was error-free, and the others had six or more errors. Resequencing after synthesis of the 6-kb fragments may be a more appropriate route in the future. Alternatively, to avoid errors introduced by PCR, the assembly intermediates can be cloned in *E. coli*.

DNA fragments were synthesized from imperfect oligos. We relied on a time-consuming *E. coli* cloning step to filter out errors.

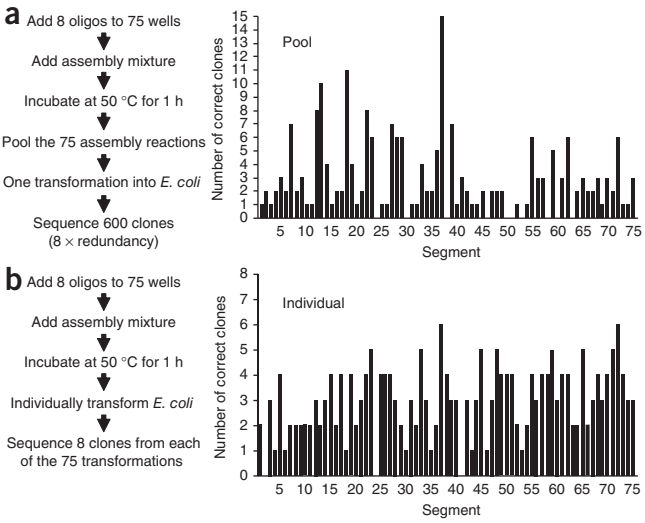

Figure 2 | Summary of results for obtaining the 75 sequence-verified first-stage assemblies. (**a**,**b**) The 75 reactions of eight oligos each were pooled and transformed into *E. coli* (**a**) or individually transformed (**b**). The number of correct clones obtained for each segment (1–75) is shown.

However, it should be possible to amplify single, assembled molecules *in vitro* that can be sequence-verified and used in subsequent assembly reactions. The oligos were synthesized at the smallest possible scale, yet only 0.0025% of the total volume from each oligo suspension were needed in the assembly reactions. Our method now does not address this inefficient oligo use, which we estimate is 70% of the total DNA synthesis costs. We discuss the limitations and potential improvements of the method, including the use of PCR and restriction sites in **Supplementary Figures 5** and **6**, and **Supplementary Note 1**.

The mouse mitochondrial genome has been previously reconstructed from overlapping, nonsynthetic PCR products[10]. Now we synthesized the complete genome beginning from chemically synthesized oligos, which allows for design beginning at the nucleotide level. Such a synthetic genomics approach could be used to restore respiratory competence to mammalian cells with mitochondrial deficiencies. Mutations in mitochondrial DNA are associated with the pathogenesis of neurodegenerative disease, blindness, deafness, various cancers, type II diabetes and the aging process[11]. Here we offer powerful tools for the design and the synthesis of complete mitochondrial chromosomes, which may be expressed once installed in mitochondria. Reiterative synthesis and transplantation of many different versions of mitochondrial genomes should help accelerate progress in determining the genetic defects associated with mitochondrial diseases (**Supplementary Note 1**).

The technology used to synthesize DNA is rapidly progressing. The automation of DNA synthesis has already been reported[4,12] and is commonly used in commercial production. In short time, DNA synthesis will be completely automated beginning at the nucleotide level, and digitized sequence will be converted to genes and genomes without human intervention. The simplicity and robustness of the method described here should help make this possible. Additional improvements in oligonucleotide cost, quality and scale[13] would dramatically accelerate progress in synthetic biology (**Supplementary Note 1**). Synthetic biologists are synthesizing and expressing genetic elements to provide a sustainable means for producing desirable products such as new and improved drugs, vaccines, biosensors, biofuels, bioremediation tools, food ingredients, cosmetics and industrial compounds[14]. Here we provide widely applicable procedures for constructing these genetic elements from synthetic DNA oligonucleotides.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank Synthetic Genomics, Inc. for funding this work, J. Glass, Y.-H. Rogers, J. Gill, M. Frazier, E. Eisenstadt and M. Gibson for helpful discussions, and M. Algire and J. Zaveri for technical assistance.

AUTHOR CONTRIBUTIONS

D.G.G. and C.M. designed research, performed research, analyzed data and wrote the paper. H.O.S., C.A.H. III and J.C.V. designed research and analyzed data.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturemethods/.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Agarwal, K.L. *et al. Nature* **227**, 27–34 (1970).
2. Gibson, D.G. *et al. Science* **319**, 1215–1220 (2008).
3. Gibson, D.G. *et al. Science* **329**, 52–56 (2010).
4. Kodumal, S.J. *et al. Proc. Natl. Acad. Sci. USA* **101**, 15573–15578 (2004).
5. Welch, M. *et al. PLoS ONE* **4**, e7002 (2009).
6. Xiong, A.S. *et al. FEMS Microbiol. Rev.* **32**, 522–540 (2008).
7. Gibson, D.G. *et al. Nat. Methods* **6**, 343–345 (2009).
8. Tian, J. *et al. Nature* **432**, 1050–1054 (2004).
9. Meldrum, D. *Genome Res.* **10**, 1081–1092 (2000).
10. Itaya, M. *et al. Nat. Methods* **5**, 41–43 (2008).
11. Taylor, R.W. & Turnbull, D.M. *Nat. Rev. Genet.* **6**, 389–402 (2005).
12. Cox, J.C. *et al. Protein Sci.* **16**, 379–390 (2007).
13. Lee, C.C., Snyder, T.M. & Quake, S.R. *Nucleic Acids Res.* **38**, 2514–2521 (2010).
14. Khalil, A.S. & Collins, J.J. *Nat. Rev. Genet.* **11**, 367–379 (2010).

ONLINE METHODS

Oligonucleotide design. We broke down the 16,299-bp mitochondrial genome plus 221 nucleotides of repeated sequence (16,520 bp) into 600 overlapping 60-base oligos (**Supplementary Table 10**). We designed single-stranded DNA oligos to overlap by 20 bases and dsDNA assemblies to overlap by 40 bp. We designed first-stage assemblies to include 20 bp of pUC19 overlapping sequence and NotI sites to release the inserts. We designed second-stage assemblies to include 40 bp of pBR322 overlapping sequence and AscI sites to release the inserts. We designed third-stage assemblies to include 40 bp of pSMART-VC (Lucigen) overlapping sequence and SbfI sites to release the inserts. We designed the final-stage assembly of the complete genome to include 40 bp of pCC1BAC (Epicentre) overlapping sequence and PmlI sites to release the insert. We designed all overlaps, including those to the vector sequences and restriction sites, in the 600 oligos (**Supplementary Table 10**). Our original intention was to assemble the respective intermediates into pBR322 or pSMART-VC and clone by *E. coli* transformation. Instead, we used these 40-bp regions as universal primer-binding domains for PCR amplification.

Preparation of assembly vectors for cloning the mitochondrial genome or subassemblies. We used the pUC19 plasmid to clone the 75 first-stage assemblies into *E. coli*. We used the pCC1BAC vector to clone the assembled mitochondrial genome into *E. coli*. To prepare the assembly vectors, we linearized these plasmids by restriction digestion with BamHI then extracted them from an agarose gel following electrophoresis. We then PCR-amplified the linearized vectors using Phusion High-Fidelity DNA polymerase (New England Biolabs (NEB)) with primers 'puc19Univ-synthesisNONot1For' and 'puc19Univ-synthesisNONot1Rev' (**Supplementary Table 14**) for cloning into pUC19 or 'pCC1BAC Assembly For' and 'pCC1BAC Assembly Rev' (**Supplementary Table 14**) for cloning into pCC1BAC. We set up 100-µl reactions using 0.04 ng µl^{-1} linearized vector, 500 nM each primer, 200 µM each dNTP, 1× high-fidelity (HF) buffer (NEB) with 0.5 mM additional $MgCl_2$ and 20 U ml^{-1} enzyme. Cycling parameters were 98 °C for 30 s, then 30 cycles of 98 °C for 10 s, 60 °C for 30 s, and 72 °C for 4 min (pUC19) or 7 min (pCC1BAC), followed by a single 72 °C incubation for 5 min. We extracted the PCR products from agarose gels after electrophoresis using the QIAquick Gel Extraction kit (Qiagen) according to the instructions provided. We quantified the PCR-amplified vectors and diluted them to 200 ng µl^{-1} with Tris-EDTA buffer pH 8.0 (TE buffer).

Preparation of the oligonucleotides used to synthesize the mitochondrial genome. We purchased the oligos from Integrated DNA Technologies unmodified with standard desalting, at the 10 nmol scale, and suspended them to 50 µM with TE buffer. We pooled equal volumes of each oligonucleotide in groups of eight then diluted them in TE buffer to a per-oligo concentration of 180 nM. For cassettes that were inefficiently assembled (for example, segments 24 and 46), we found that results could be improved by heating a 5 µl sample of the oligo pool to 95 °C for 1 min then slow cooling at 0.1 °C s^{-1} to 4 °C before adding 15 µl of the enzyme-reagent mix (see below). Optimal concentrations for assembling more than eight oligos at a time are listed in **Supplementary Table 1**.

One-step-isothermal DNA assembly. We assembled the oligonucleotides into PCR-amplified pUC19 by using our previously published isothermal assembly protocol[7]. Briefly, we added 5 µl of diluted oligo DNA (180 nM each) to a thawed 15 µl sample of an assembly mixture already containing the PCR-amplified pUC19 vector. We then incubated the reaction at 50 °C for 1 h. We prepared the vector-enzyme-reagent mixture by combining 320 µl 5× isothermal (ISO) reaction buffer, 20 µl PCR-amplified pUC19 (200 ng µl^{-1}), 0.64 µl of 10 U µl^{-1} T5 exo (Epicentre), 20 µl of 2 U µl^{-1} Phusion polymerase, 160 µl of 40 U µl^{-1} Taq ligase (NEB) and water up to a final volume of 1.2 ml. We aliquoted 15 µl of this enzyme-reagent mix and stored the aliquots at −20 °C. We prepared the 5× ISO reaction buffer by combining 3 ml of 1 M Tris-HCl (pH 7.5), 150 µl of 2 M $MgCl_2$, 60 µl of 100 mM dGTP, 60 µl of 100 mM dATP, 60 µl of 100 mM dTTP, 60 µl of 100 mM dCTP, 300 µl of 1 M DTT, 1.5 g PEG-8000 (United States Biochemical) and 300 µl of 100 mM β-micotinamide adenine dinucleotide (NAD). This will produce 6 ml, which can be aliquoted and stored at −20 °C.

Cloning the DNA assembly products. To clone the assembled products, we transformed 1-µl samples from the assembly reactions into 30 µl Epi300 *E. coli* cells (Epicentre) in a 1-mm cuvette (BioRad) at 1,200 V, 25 µF and 200 Ω using a Gene Pulser Xcell Electroporation System (BioRad). We allowed the cells to recover at 37 °C for 1.5 h in 1 ml of SOC medium. We plated 100 µl of the individual 75 subassembly transformations onto LB medium with 100 µg ml^{-1} carbenicillin. For the transformation of the pooled assemblies, we plated 1 ml onto a Genetix Bioassay tray containing LB medium with 100 µg ml^{-1} carbenicillin. We plated complete genome assembly transformations on LB medium with 12.5 µg ml^{-1} chloramphenicol. After incubation at 37 °C for 16 h, we grew individual colonies in 1 ml of LB medium with 100 µg ml^{-1} carbenicillin (pUC19) or 12.5 µg ml^{-1} chloramphenicol (pCC1BAC) and incubated them overnight at 37 °C.

Screening for full-length cassettes assembled from oligos. To screen for full-length first-stage subassemblies not initially found by DNA sequencing, we performed colony PCR. We diluted a single colony in 100 µl water and then used this mixture as template in a PCR with the M13F and M13R primers (**Supplementary Table 14**). We set up 20 µl reactions using 1 µl template, 500 nM each primer, 200 µM each dNTP, 2 µl 10× ThermoPol Reaction buffer (NEB) and 0.2 µl Taq polymerase (NEB). Cycling parameters were 95 °C for 5 min, then 30 cycles of 95 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s, followed by a single 72 °C incubation for 5 min. We analyzed products on 2% E-gels (Invitrogen) alongside the 100-bp DNA ladder (NEB).

Preparation and assembly of 284-bp cassettes. For genome assembly after 45 PCR cycles, the same DNA samples used for DNA sequencing were also used for DNA assembly. We pooled 20 µl of each in sets of five and digested with NotI in a 200 µl reaction, at 37 °C, for 3 h. We purified the 15 reactions using the QIAquick PCR purification kit (Qiagen) according to the instructions provided, except we eluted products with 100 µl TE Buffer. We assembled 5 µl of this DNA with 15 µl of the enzyme-reagent mix, as above but without the pUC19 vector. For genome assembly after 80 PCR cycles, we PCR-amplified the 75 cassettes using Phusion High-Fidelity DNA polymerase with primers 'pUC19

doi:10.1038/nmeth.1515

CPCR For' and 'pUC19 CPCR Rev' (**Supplementary Table 14**). We set up 100 µl reactions using 2 µl template, 500 nM each primer, 200 µM each dNTP, 1× HF buffer and 20 U ml^{-1} enzyme. Cycling parameters were 98 °C for 30 s, then 30 cycles of 98 °C for 10 s, 55 °C for 30 s and 72 °C for 30 s, followed by a single 72 °C incubation for 5 min. We purified PCR products using the QIAquick PCR purification lit as above. We quantified each of the 75 fragments by spectrophotometry on a ND-1000 NanoDrop (Thermo Scientific) and determined that they were about 40 ng $µl^{-1}$. As above, we pooled 20 µl of each fragment, digested them with NotI, column-purified them and then assembled the DNA.

Preparation and assembly of 1.2-kb intermediates. We amplified second-stage assembly products with Phusion polymerase with primers 'pBR322 CPCR 20bp-Clone FOR' and 'pBR322 CPCR 20bp-Clone REV' (**Supplementary Table 14**). We set up 100 µl reactions using 2 µl template, 500 nM each primer, 200 µM each dNTP, 1× HF buffer and 20 U ml^{-1} enzyme. Cycling parameters were 98 °C for 30 s, then 25 cycles (or 30 cycles for genomes constructed following 80 PCR cycles) of 98 °C for 5 s, 60 °C for 30 s, and 72 °C for 20 s, followed by a single 72 °C incubation for 5 min. We purified PCR products using the QIAquick PCR purification kit as above. We quantified each of the 15 PCR products by NanoDrop spectrophotometry and determined that they were about 10 ng $µl^{-1}$ (25 cycles) or 60 ng $µl^{-1}$ (30 cycles). We pooled 20 µl of each in sets of five and digested with AscI in a 200 µl reaction, at 37 °C, for 3 h. We purified the three reactions using the QIAquick PCR purification kit as above. We assembled 5 µl of this DNA with 15 µl of the enzyme-reagent mix, as above.

Preparation and assembly of 5.6-kb intermediates. We amplified third-stage assembly products with Phusion polymerase with primers 'pSMART-VNTI-CPCR-For' and 'pSMART-VNTI-CPCR-Rev' (**Supplementary Table 14**). We setup 100 µl reactions using 1 µl template, 500 nM each primer, 200 µM each dNTP, 1× HF buffer and 20 U ml^{-1} enzyme. Cycling parameters were 98 °C for 30 s, then 20 cycles of 98 °C for 5 s, 60 °C for 30 s and 72 °C for 1.5 min, followed by a single 72 °C incubation for 5 min. We pooled equal volumes of PCR products and then purified them using the QIAquick PCR purification kit as above. We digested the PCR products with SbfI in a single 200 µl reaction, at 37 °C, for 3 h. We purified the reaction using the QIAquick PCR purification kit as above. We assembled 50 ng of each 5.6-kb fragment and 60 ng of the PCR-amplified pCC1BAC vector by adding 15 µl of the enzyme-reagent mix, as above.

Additional notes on cycling between DNA assembly and PCR. Results could be improved by gel purifying restriction-digested PCR products before DNA assembly. However, this was not necessary. Instead, we column-purified the reactions. However, we found that PCRs did not need to be cleaned before setting up an assembly reaction. We also found that PCR products do not necessarily need to be digested. This should make future DNA synthesis efforts more amenable to automation.

Screening for complete mitochondrial genomes and subassemblies. Approximately 20,000 colonies were produced per 1-µl assembly reaction that we transformed. We isolated BACs potentially containing complete mitochondrial genomes from *E. coli* by alkaline lysis using the P1, P2 and P3 buffers supplied by Qiagen followed by isopropanol precipitation. We dissolved DNA pellets in TE buffer containing RNase. We analyzed restriction patterns after digestion with PmlI or EcoRI. We visualized patterns on 0.8% E-gels alongside the 1-kb DNA ladder (NEB).

Purification of mitochondrial genomes from *E. coli*. The synthetic mitochondrial genomes were contained in pCC1BAC and propagated in the Epi300 *E. coli* strain. This cloning system permitted induction to ten copies or more of these BACs per cell. We inoculated the strains containing the cloned genomes into 150 ml LB plus 12.5 µg ml^{-1} chloramphenicol and 1× induction solution (Epicentre) and incubated them at 37 °C for 16 h. We collected the cultures and purified the BACs using a HiSpeed Plasmid Maxi kit (Qiagen). We eluted the DNA with 500 µl TE buffer and recovered ~100 µg.

Automation. We submitted a Genetix Bioassay tray containing ~1,000 *E. coli* clones for colony picking (QPix, Genetix), plasmid DNA preparation (DNA purification Robotic Workstation, Thermo CRS), sequencing reaction setup (Biomek FX, Beckman Coulter) and sequencing (3730xl DNA Analyzer, Applied Biosystems). We automated each of these steps.

DNA sequencing analysis. We sequenced first-stage cassettes in only one direction with the M13R primer. We carried out standard sequencing reactions as above. We aligned trace files with the reference sequences (**Supplementary Table 10**) with ClustalW Multiple alignment[15], contained within the BioEdit Sequence Alignment Editor software. The sequence data contained 384-well plate coordinates of template DNA and *E. coli* glycerol stocks of the clones. We sequenced complete synthetic mitochondrial genomes (**Supplementary Tables 13** and **15**) with 48 of the 60-base oligos used to build the DNA molecule (**Supplementary Table 10**), and 'Seq pCC1BAC Ass Vector For' and 'Seq pCC1BAC Ass Vector Rev' (**Supplementary Table 14**). Twenty-five oligos annealed to the top strand and 25 of them annealed to the bottom strand. We analyzed sequence as above. We determined synthesis error rates by dividing the number of errors in the complete insert clones by the total number of bases synthesized.

15. Thompson, J.D., Higgins, D.G. & Gibson, T.J. *Nucleic Acids Res.* **22**, 4673–4680 (1994).