Building with DNA 2

Andrew Tolonen
Genoscope et l'Université d'Évry
08 october 2014

atolonen “at” genoscope.cns.fr
@andrew_tolonen
www.tolonenlab.org

Yesterday we talked about ways to assemble DNA building blocks

Cloning with restriction enzymes

- Traditional cloning
- Biobricks
- Golden Gate

Cloning without restriction enzymes

- Gateway
- Circular polymerase extension cloning (CPEC)
- Ligation-independent cloning
- SliCE
- Gibson method (*in vitro*)

Today, we'll discuss large-scale projects to synthesize entire chromosomes

Gibson et al, 2010 “Creation of a bacterial cell controlled by a chemically synthesized genome” *Science* 2010 (PMID 20488990)

This paper is part 3 of 3 in the series

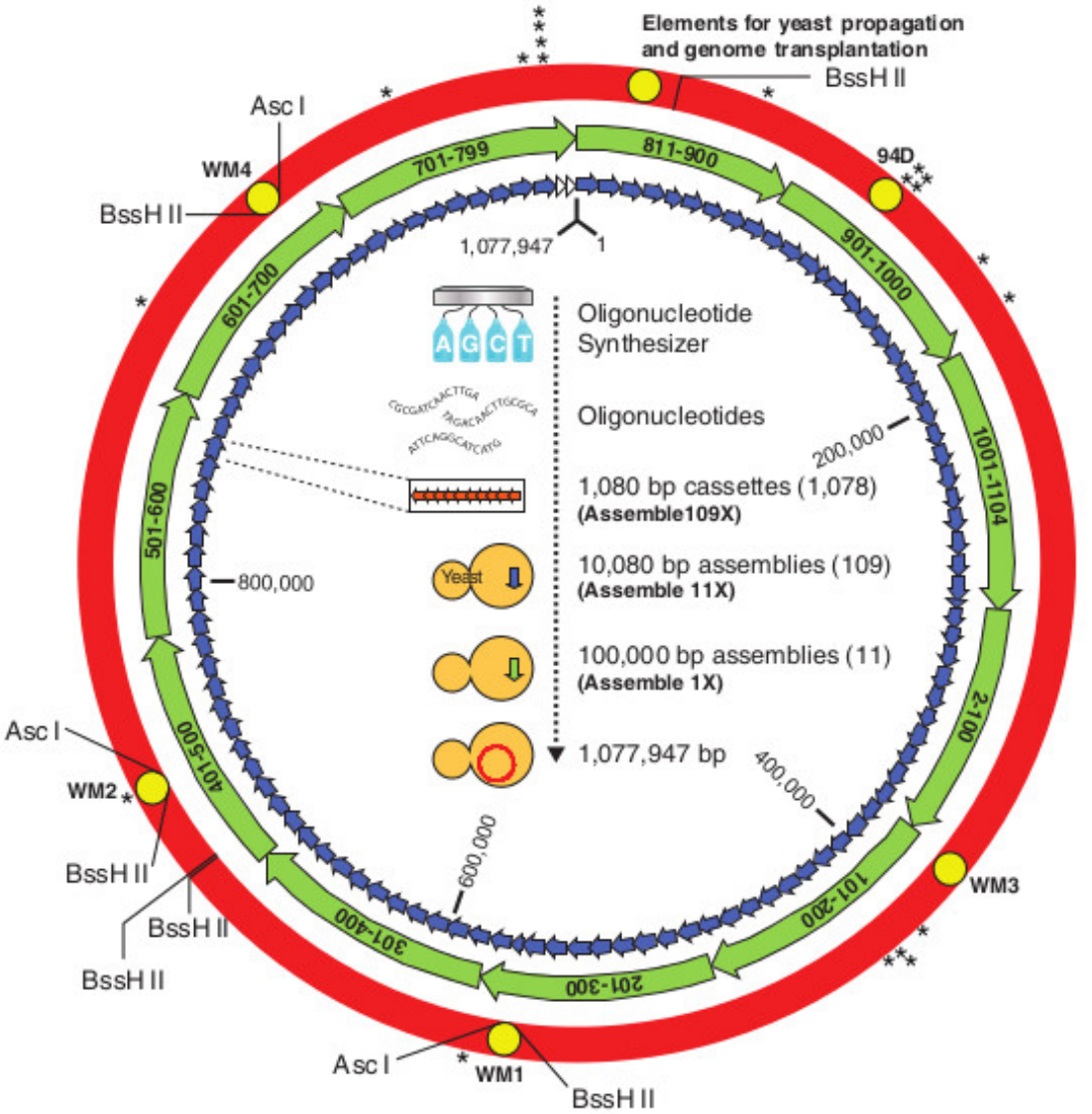
Part 1: Lartigue et al, 2007 *Science* (PMID 17600181) PEG-mediated *Mycoplasma* genome transplantation

Part 2: Gibson et al, 2008 *Science* (PMID 18218864) chemical synthesis of the *Mycoplasma* genome

Part 3: Gibson et al, 2010 Science (PMID 20488990) combine methods from parts 1 and 2 to make 'synethetic' *Mycoplasma mycoides*

assemble 1.08 Mb *Mycoplasma* genome in 3 stages

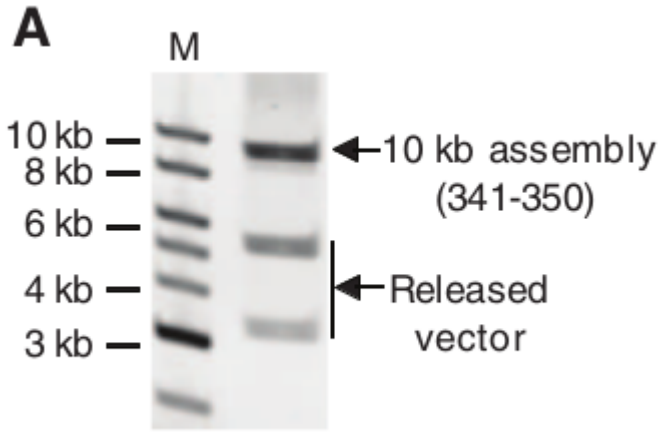
Fig 1



Yellow circles show differences from WT genome:

1. 4 watermarks
2. 4kb deletion
3. yeast propagation elements

stage 1 assembly of 10 kb fragments from 1 kb oligos (Fig 2A)



1. Purchased 1078x1080 bp oligos with 80 bp overlaps
2. Transformed oligos into yeast for *in vivo* recombination (Gibson et al, 2009 PMID 19745056)
3. Transformed plasmids into *E. coli* to improve yields
4. cut plasmids with NotI+SbfI to confirm 109x10 kb inserts

Gibson 2009: use yeast as a 'factory' to assemble oligonucleotides

Fig 1: schema for transformation of ssDNA oligos into yeast spheroplasts for assembly

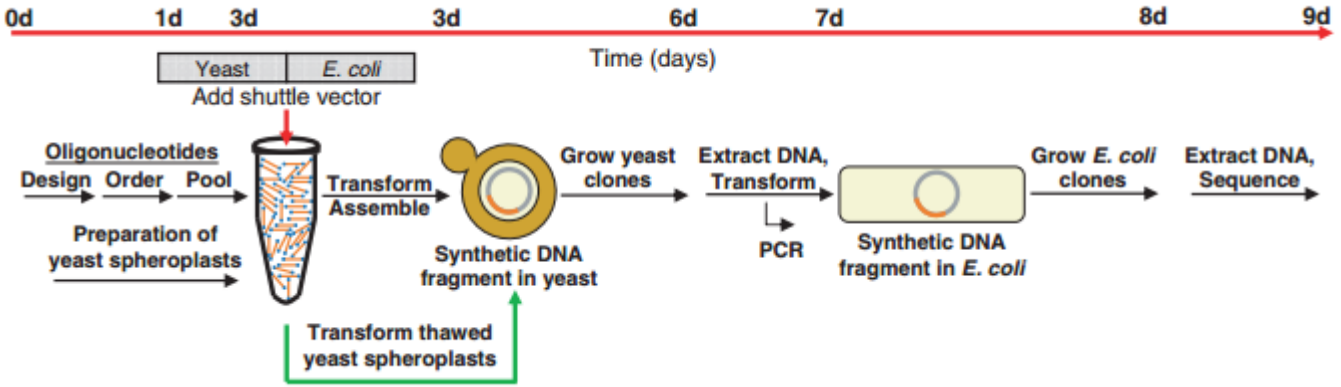
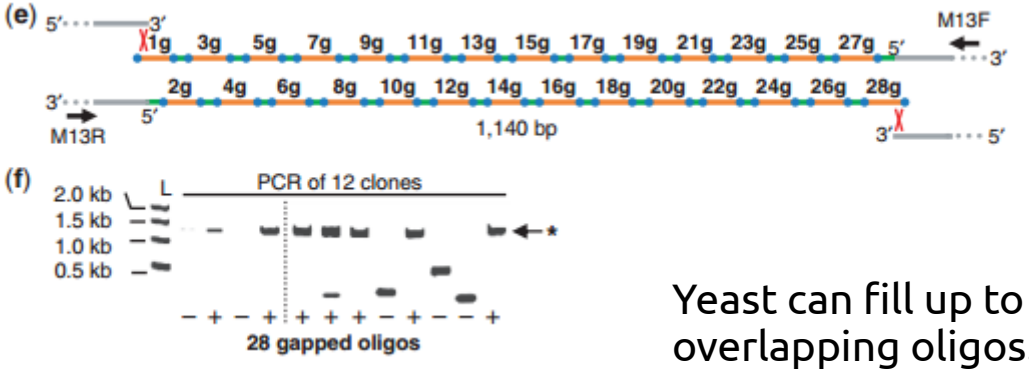


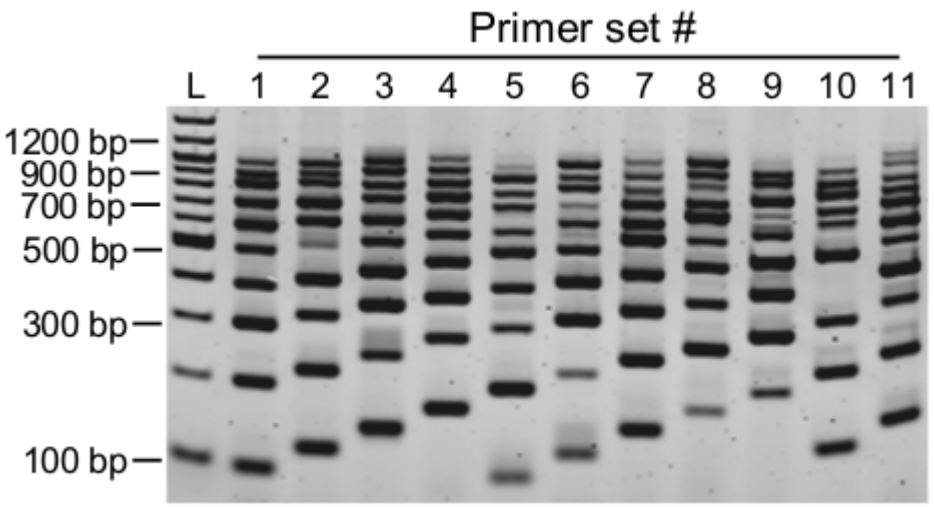
Fig 3: *in vivo* assembly of 28 60-mers with 20 bp overlap (40bp gaps) into a 1140 bp fragment

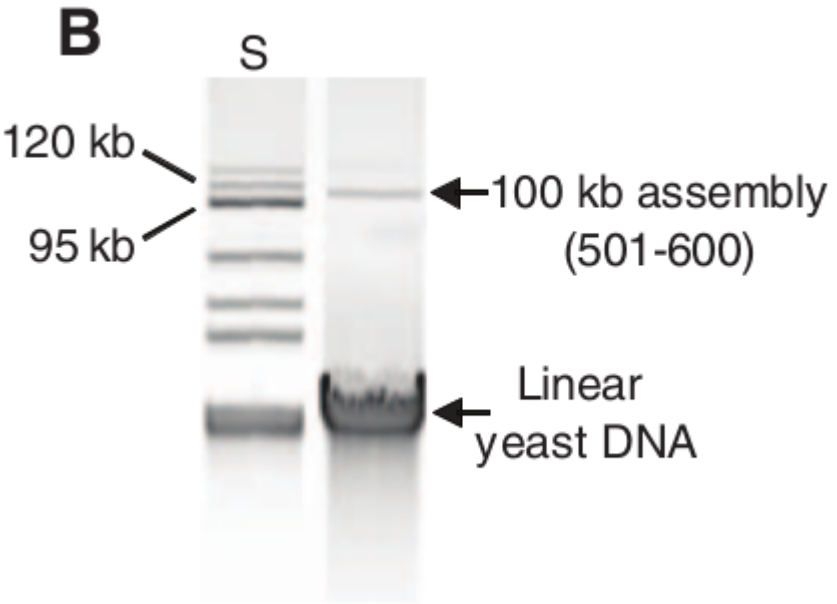


Yeast can fill up to 160 bp gaps between overlapping oligos!

stage 2 assembly of 100 kb fragments (Fig S3,2B-D)

Fig S3





1. Pooled 10x10kb fragments and transformed into yeast for *in vivo* recombination
2. Isolate DNA from yeast
3. Multiplex PCR to confirm that all 10kb segments are present (Fig S3)
4. Confirm 100 kb size on gel (Fig 2B-D)

stage 3 assembly of complete 1.08 Mb genome

Fig 3A

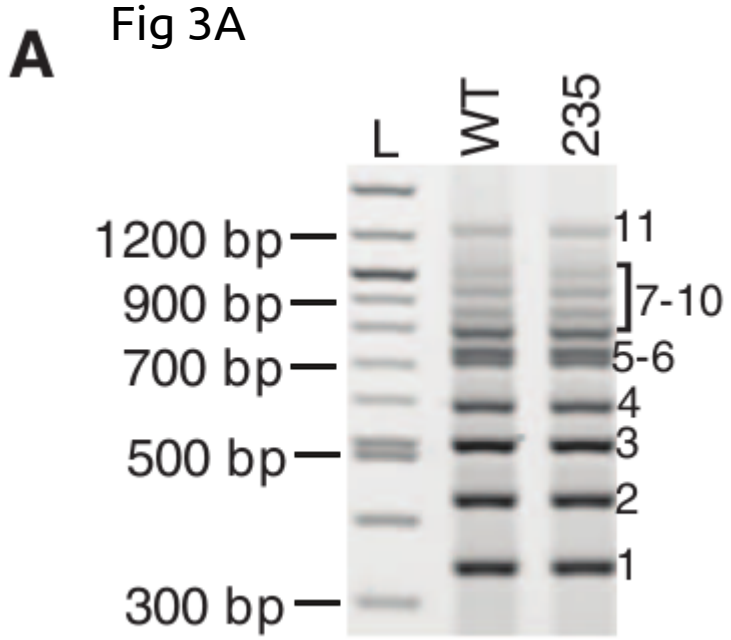
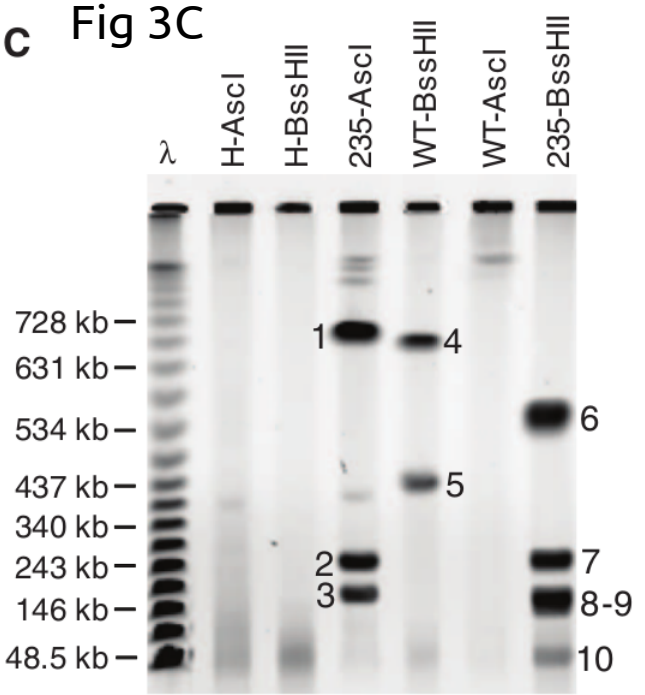
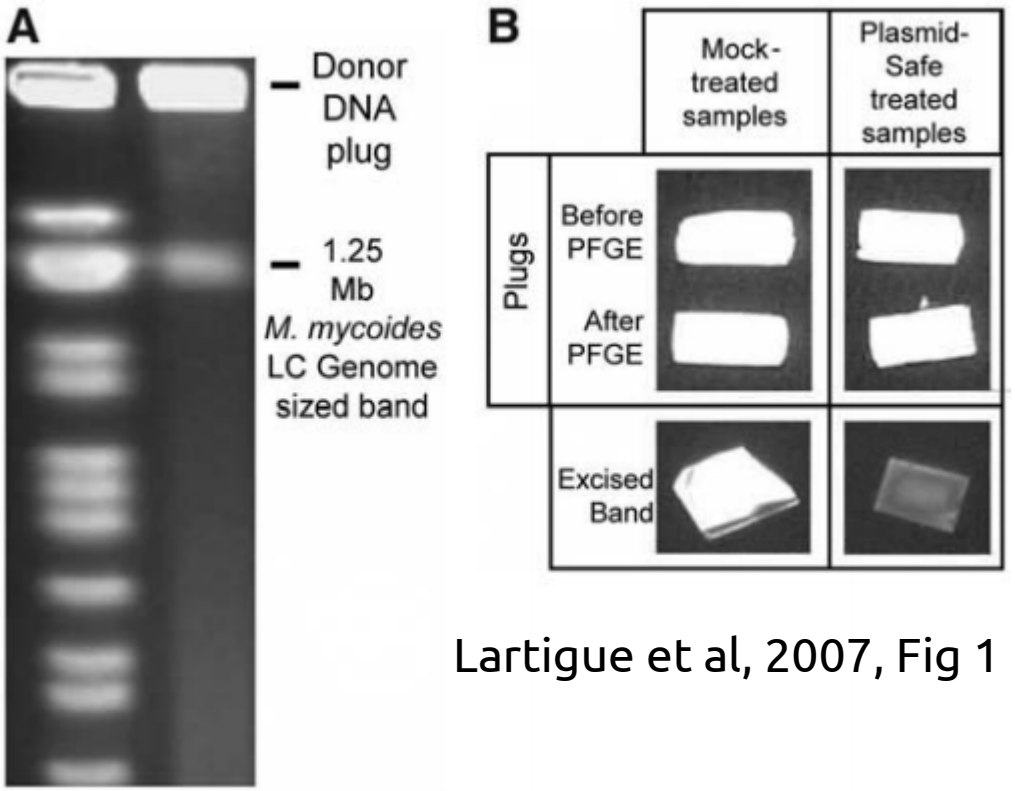


Fig 3C



1. Isolate 100 kb fragments from yeast
2. Remove yeast chromosomal DNA: purify circular DNA by trapping it in agarose
3. Pool 11x100kb fragments and transform into yeast for *in vivo* recombination
4. Multiplex PCR to amplify junction fragments and compare to WT (Fig 3A)
5. Restriction analysis with AscI and BSSHII (Fig 3C).

PEG-mediated genome transplantation



Lartigue et al, 2007, Fig 1

1. Encapsulate circular genome in agarose

2. Transfer genome into PEG-treated Mycoplasma Recipient

show that transplant cells have the synthetic genome

Fig 4A: Multiplex PCR of 4 watermarks

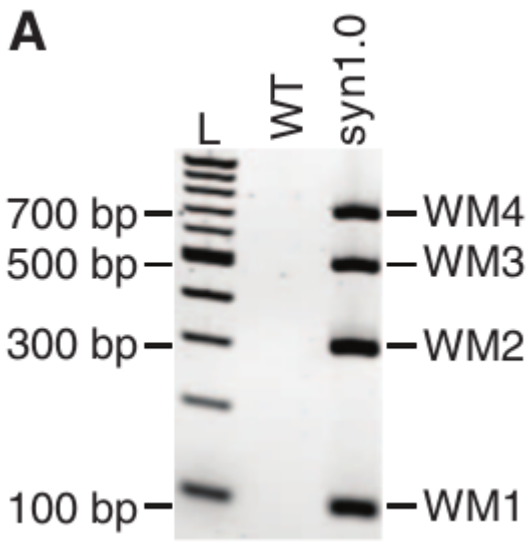
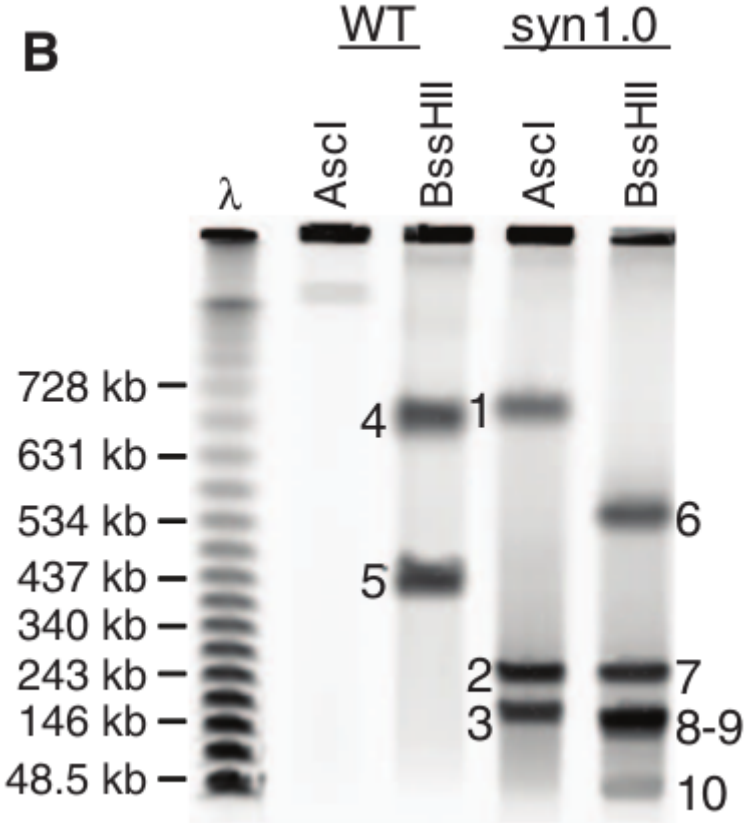


Fig 4B: restriction digest with BssHI and AscI



Sequences of 4 watermarks

Watermark-1, 1246 base pairs

TTAACTAGCTAAGTTCGAATATTTCTATAGCTGTACATATTGTAATGCTGATAACTAATACTGTGCGCTTGACTGTGATCCTGATAAATAACTTCTTCTGTAGGGTAGAGTTTTA
TTTAAGGCTACTCACTGGTTGCAAACCAATGCCGTACATTACTAGCTTGATCCTTGGTCGGTCATTGGGGGATATCTCTTACTAATAGAGCGGCCTATCGCGTATTCTCGCCG
GACCCCCCTCTCCCACACCAGCGGTGTAGCATCACCAAGAAAATGAGGGGAACGGATGAGGAACGAGTGGGGGCTCATTGCTGATCATAATGACTGTTTATATACTAATGC
CGTCAACTGTTTGCTGTGATACTGTGCTTTCGAGGGCGGGAGATTCGTTTTTGACATACATAAATATCATGACAAAACAGCCGGTCATGACAAAACAGCCGGTCATAATAGAT
TAGCCGGTGACTGTGAAACTAAAGCTACTAATGCCGTCAATAAATATGATAATAGCAACGGCACTGACTGTGAAACTAAAGCCGGCACTCATAATAGATTAGCCGGAGTCGT
ATTCATAGCCGGTAGATATCACTATAAGGCCCAGGATCATGATGAACACAGCACCACGTCGTCGTCCGAGTTTTTTTGCTGCGACGTCTATACCACGGAAGCTGATCATAAAT
AGTTTTTTTGCTGCGGCACTAGAGCCGGACAAGCACACTACGTTTGTAAATACATCGTTCCGAATTGTAAATAATTTAATTTCGTATTTAAATTATATGATCACTGGCTATAGTC
TAGTGATAACTACAATAGCTAGCAATAAGTCATATATAACAATAGCTGAACCTGTGCTACATATCCGCTATACGGTAGATATCACTATAAGGCCCAGGACAATAGCTGAACTGA
CGTCAGCAACTACGTTTAGCTTGACTGTGGTCGGTTTTTTTGCTGCGACGTCTATACGGAAGCTCATAACTATAAGAGCGGCACTAGAGCCGGCACACAAGCCGGCACAGT
CGTATTCATAGCCGGCACTCATGACAAAACAGCGGCGCGCCTTAACTAGCTAA

Watermark-2 1081 base pairs

TTAACTAGCTAACAACTGGCAGCATAAAACATATAGAACTACCTGCTATAAGTGATACAACTGTTTTCATAGTAAAACATACAACGTTGCTGATAGTACTCCTAAGTGATAGCTT
AGTGCGTTTAGCATATATTGTAGGCTTCATAATAAGTGATATTTTAGCTACGTAACTAAATAAACTAGCTATGACTGTACTCCTAAGTGATATTTTCATCCTTTGCAATACAATAA
CTACTACATCAATAGTGCGTGATATGCCTGTGCTAGATATAGAACACATAACTACGTTTGCTGTTTTCAGTGATATGCTAGTTTCATCTATAGATATAGGCTGCTTAGATTCCCT
ACTAGCTATTTCTGTAGGTGATATACGTCCATTGCATAAGTTAATGCATTTAACTAGCTGTGATACTATAGCATCCCCATTCCTAGTGCATATTTTCATCCTAGTGCTACGTGAT
ATAATTGTACTAATGCCTGTAGATAATTTAATGCCTGGCTCGTTTGTAGGTGATAATTTAGTGCCTGTAAAACATATACCTGAGTGCTCGTTGCGTGATAGTTCGTTCATGCAT
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
CTATAGATTTCCTGCTAAGTGATAAGTCTACTGATTTACTAATGAATAGCTTGGTTTGGCATACACTGTGCGCTGCACTGGTGATAGCTTTTCGTTGATGAATAATTTCCCTA
GCACTGTGCGTGATATGCTAGATTCTGTAGATAGGCTAAATTCGTCTACGTTTGTAGGTGATAGTTTAGTTGCTGTAACTAATATTATCCTGTGCCGTTGCTAAGCTGTGATA
TCATAGTGCTGCTAGATATGATAAGCAAACTAATAGAGTCGAGGGGGAGTCTCATAGTGAATACTGATATTTAGTGCTGCCGTTGAATAAGTTCCCTGAACATTGTGATACT
GATATTTAGTGCTGCCGTTGAATATCCTGCATTTAACTAGCTTGATAGTGCATTCGAGGAATACCCATACTACTGTTTTCATAGCTAATTATAGGCTAACATTGCCAATAGTGC
GGCGCGCCTTAACTAGCTAA

Cracking the watermark code

A new triplet code to include all 26 letters and punctuation

TAG = a	GCA = k	TCC = u	AGA = 4	CAC = /
AGT = b	AAC = l	TTG = v	GCG = 5	CCA = =
TTT = c	CAA = m	GTC = w	GCC = 6	CGA = .
ATT = d	TGC = n	GGT = x	TAT = 7	GAG = !
TAA = e	CGT = o	CAT = y	CGC = 8	CAG = :
GGC = f	ACA = p	TGG = z	GTA = 9	GGA = "
TAC = g	TTA = q	TCT = 0	ATA = space	GTG = ,
TCA = h	CTA = r	CTT = 1	GGG = chr(10)	TCG = @
CTG = i	GCT = s	ACT = 2	AGC = >	CCC = -
GTT = j	TGA = t	AAT = 3	CGG = <	

http://spth.virii.lu/InfectingDNA.txt

Cracking the watermark code

http://genomevolution.org/wiki/index.php/Mycoplasma_mycoides_JCVI-syn1.0_Decoded

Watermark 1

J. CRAIG VENTER INSTITUTE 2009 ABCDEFGHIJKLMNOPQRSTUVWXYZ
0123456789?@??-??=/:<?>??????"??!'., SYNTHETIC GENOMICS, INC.
<!DOCTYPE HTML><HTML><HEAD><TITLE>GENOME TEAM</TITLE>
</HEAD><BODY><A HREF="HTTP://WWW.JCVI.ORG/">THE JCVI</A><P>
PROVE YOU'VE DECODED THIS WATERMARK BY EMAILING US
<A HREF="MAILTO:XXXXXXXX@JCVI.ORG">HERE!</A></P></BODY></HTML>

From: "Montague, Michael" <MMontague@jcvi.org>
Subject: RE: A bit late
Date: March 23, 2012 7:53:28 AM MST
To: Eric Lyons <elyons.uoa@gmail.com>

Hi Eric,

You are the 83rd person or group to decode the watermarks.
The first decoder was a recently graduated student from U. Penn.
named Andrew Ettenger only 3 hours and 13 minutes after the watermark sequences were released.

--Mike Montague

characterize phenotype of cells with synthetic genome

Streak cells on X-gal plates: WT genome is white and synthetic genome is blue (contains *lacZ* gene)

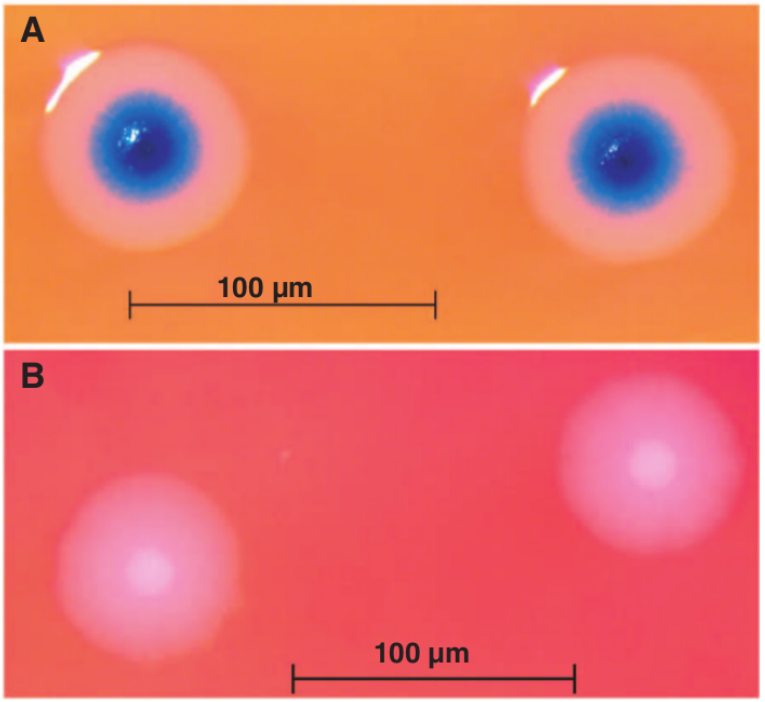
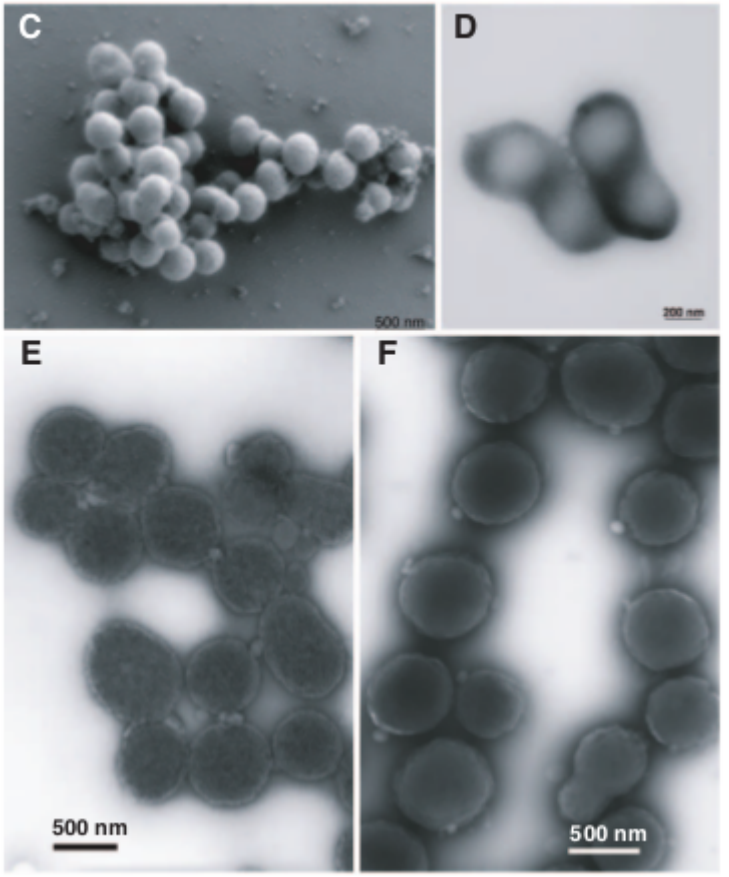


Fig 5A,B

Electron microscopy show expected morphology



They sequenced the genome the synthetic strain, JCVI-syn1.0

-8 single base pair changes

-an *E. coli* IS1 transposon insertion

-85 bp duplication

Discussion questions

Question 1: Why did they choose to synthesize the genome of *Mycoplasma*? What genome would you do next?

Question 2: The entire *Mycoplasma* genome was first cloned as a centromeric plasmid in yeast, which was then transferred into a *Mycoplasma* cell. What difficulties did they encounter in transferring the genome into *Mycoplasma*?

Question 3: Did they creat a 'synthetic cell'? What is their supporting rationale?

Question 4: Was this project worth $40 million?

Now we'll discuss the first synthesis of a eukaryotic chromosome

Annaluru et al ,2014 “Total synthesis of a functional designer eukaryotic chromsome” *Science* PMID 24674868

Saccharomyces cerevisiae genome

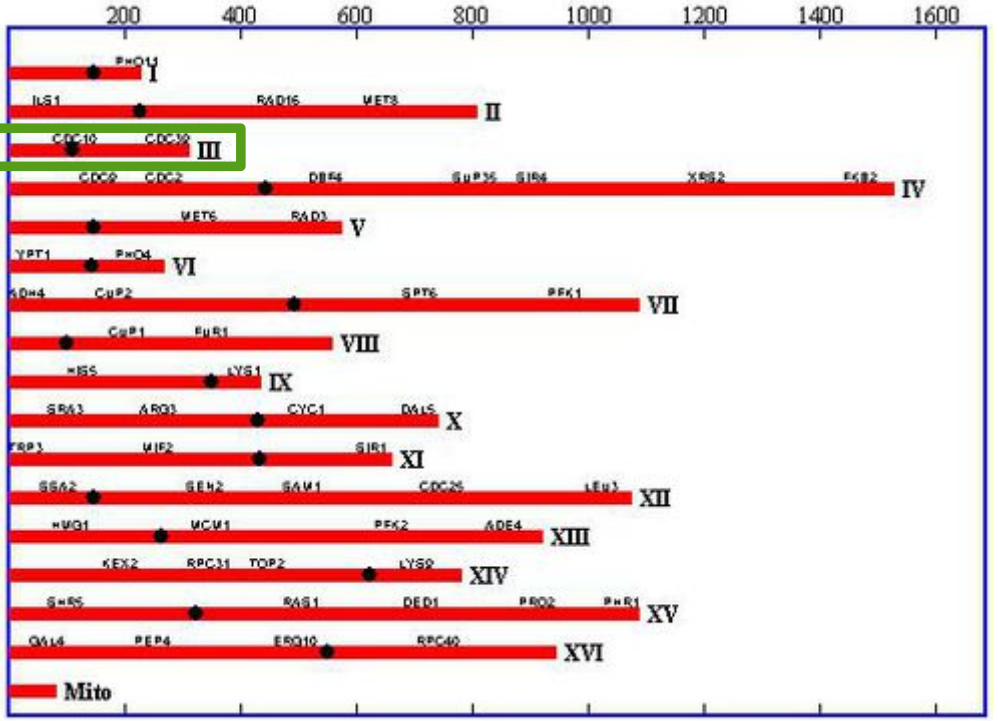
- first eukaryotic genome sequenced (1996)

-16 chromosomes

-6,275 genes (~5000 individually non-essential)

-12,156,677 bp

chromosome III is the 3rd smallest



strategy for SynIII chromosome synthesis

This was done by students (L2, L3) in a Build-a-Genome course!

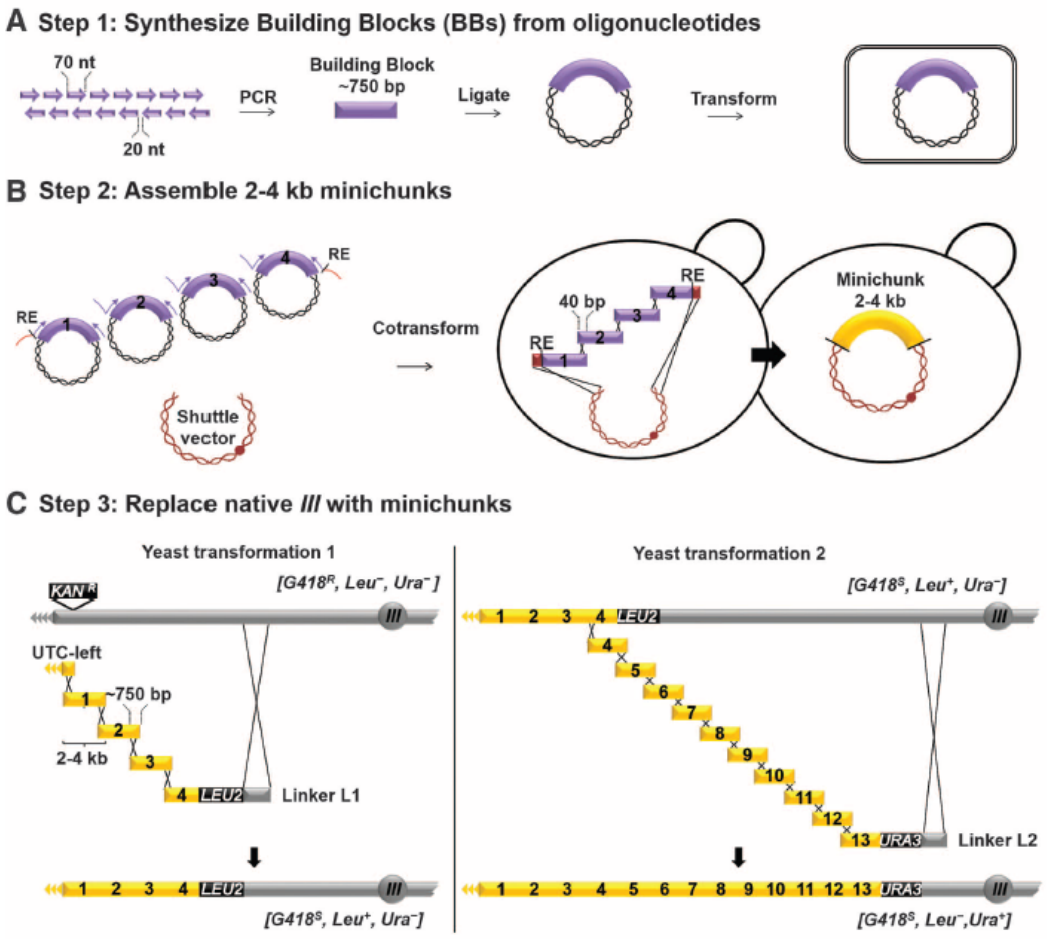


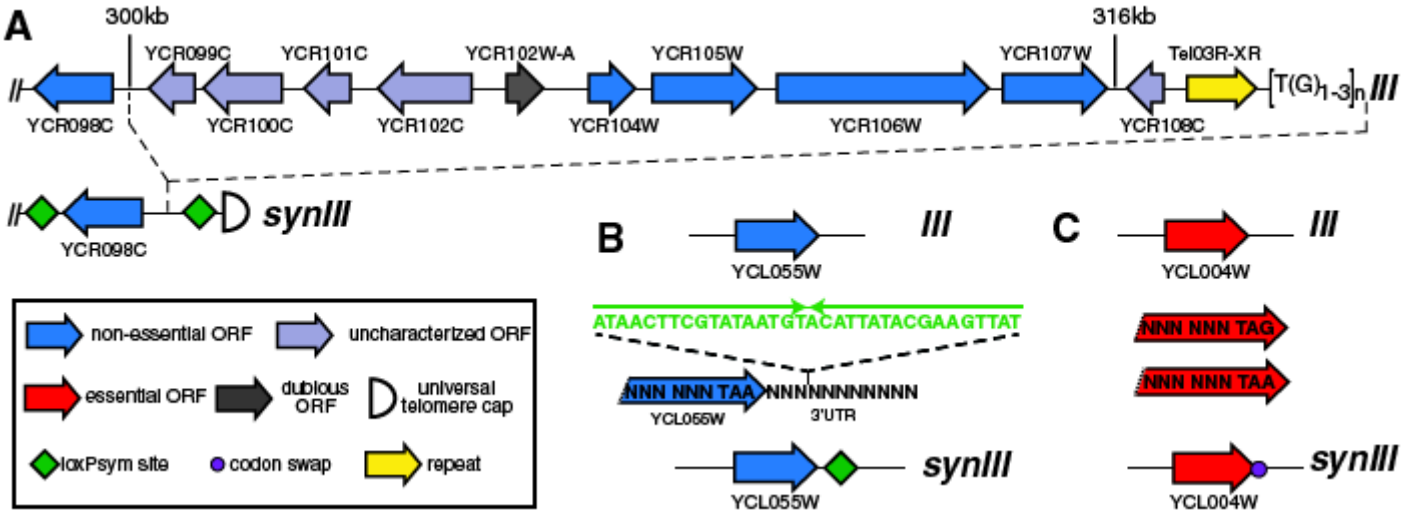
Fig 2

overview of synthetic yeast chromosome III (synIII)

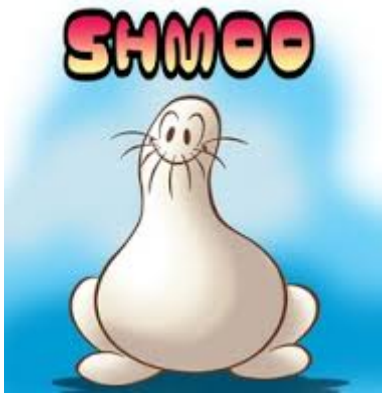
-removed ~50,000 bp (316,617->278,871 bp): sub-telomeres, introns, TRNAs, transposons, silent mating loci

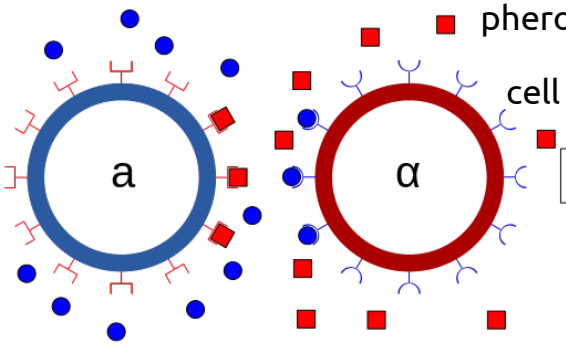
-changed stop codons TAG->TAA (free up a codon)

-inserted 'scrambling sites' (loxPsym) around non-essential genes

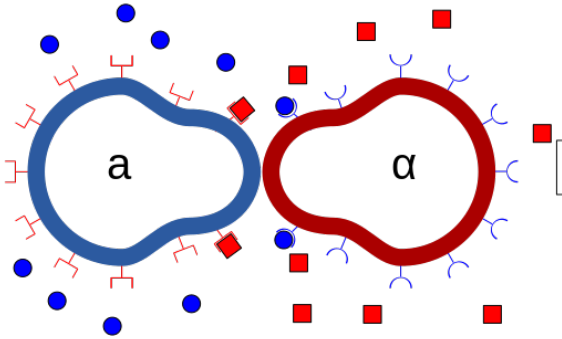


chromosome III contains the MAT locus (determines mating type)

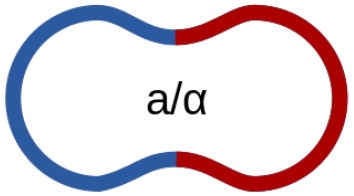




1 Yeast have 2 haploid mating types (a and alpha) that produce different pheromones and receptors.



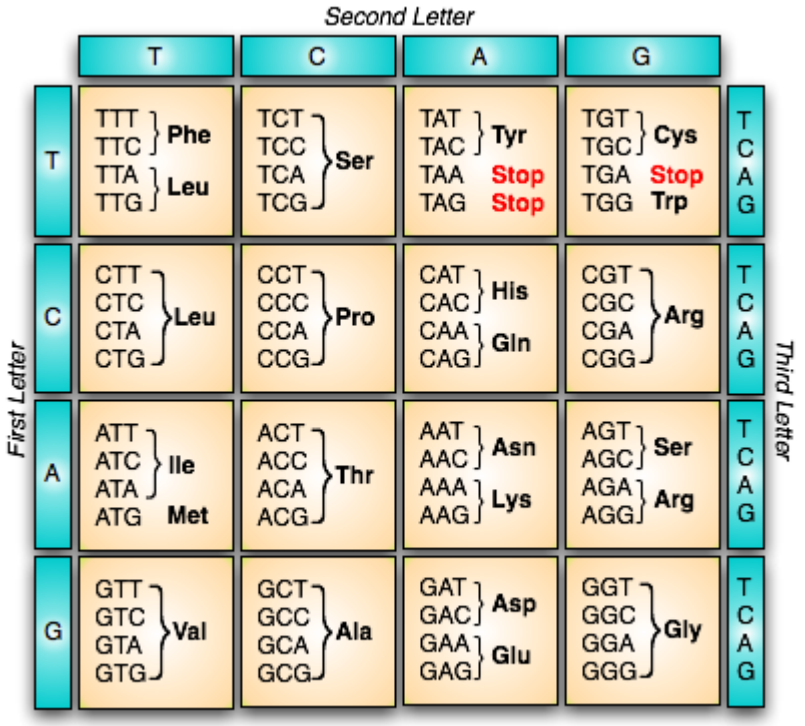
2 Haploid cells respond to pheromones by forming a 'shmoo'



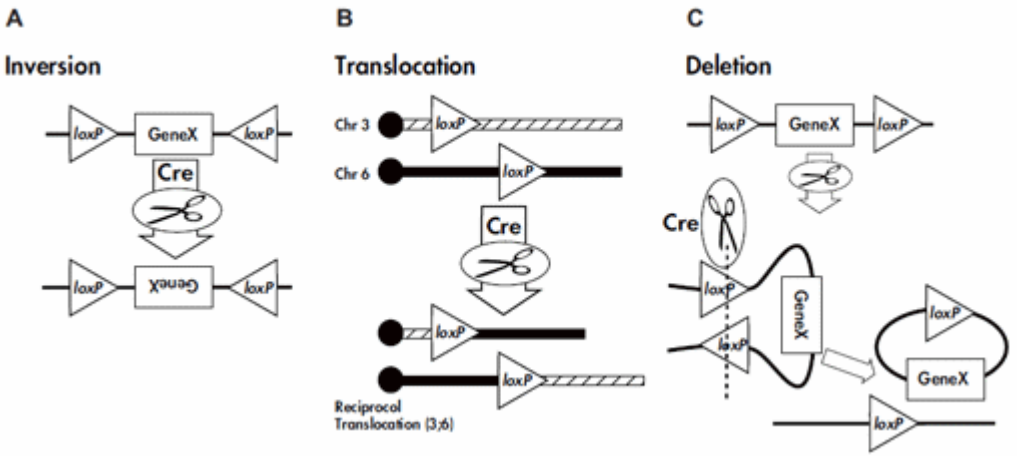
3 Diploid cell can later form haploid spores by meoisis

http://en.wikipedia.org/wiki/Mating_of_yeast

TAG->TAA frees up a TAA codon to use unnatural amino acids

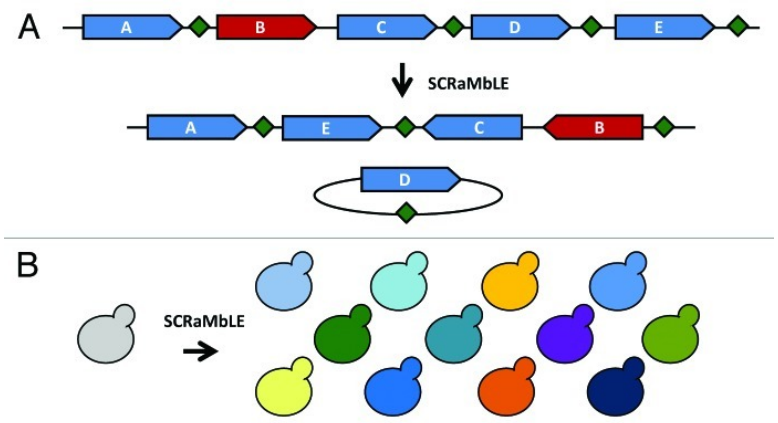


SCRaMbLE genes on synIII using Cre-lox



Cre recombinase recognizes lox sites to mediate inversions, translocations, and deletions.

SCRaMbLE

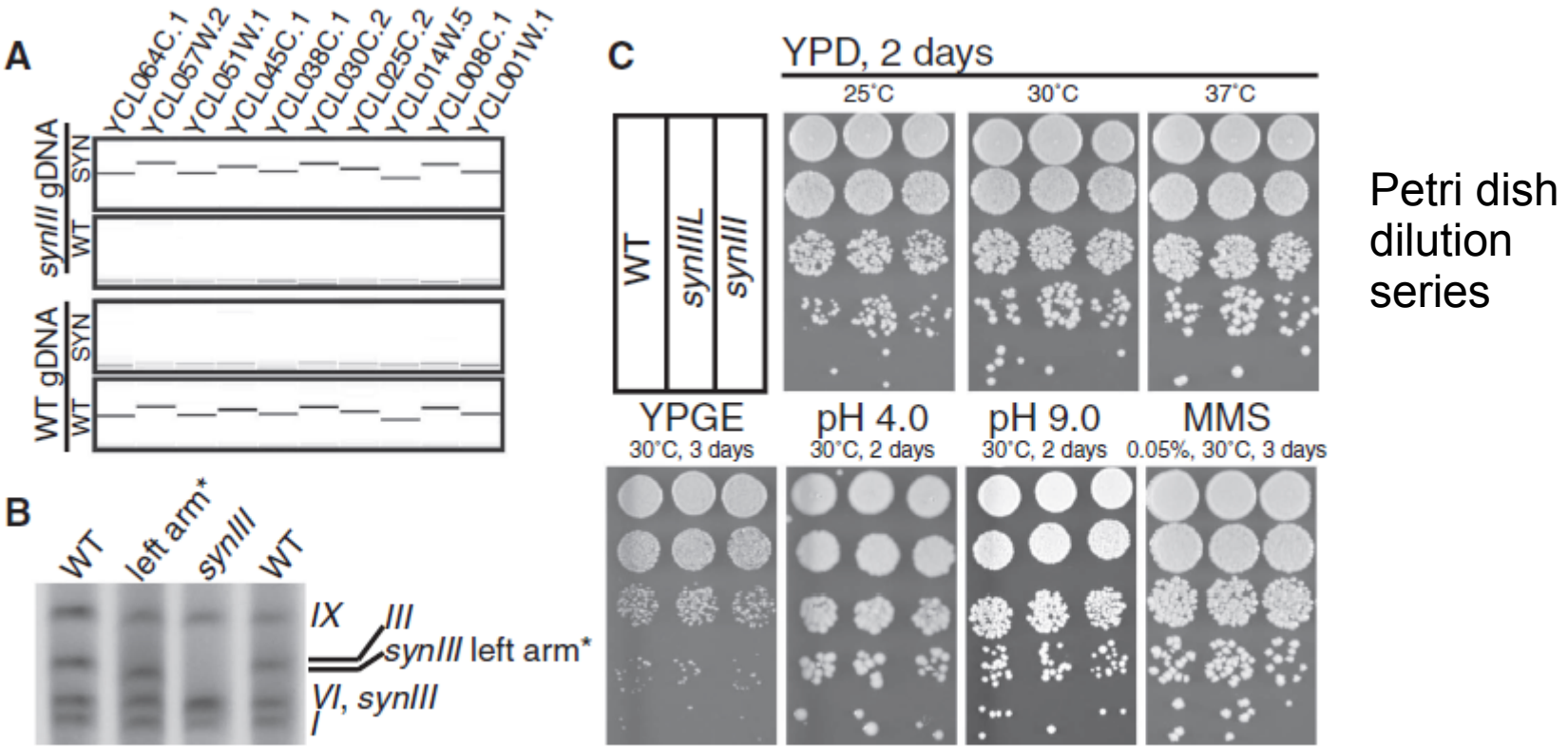


Put lox sites in 3' UTR of non-essential genes.

Use Cre recombinase to reorder and remove genes.

PMID 22572789

confirmation that synIII is correct



Petri dish dilution series

Fig 3

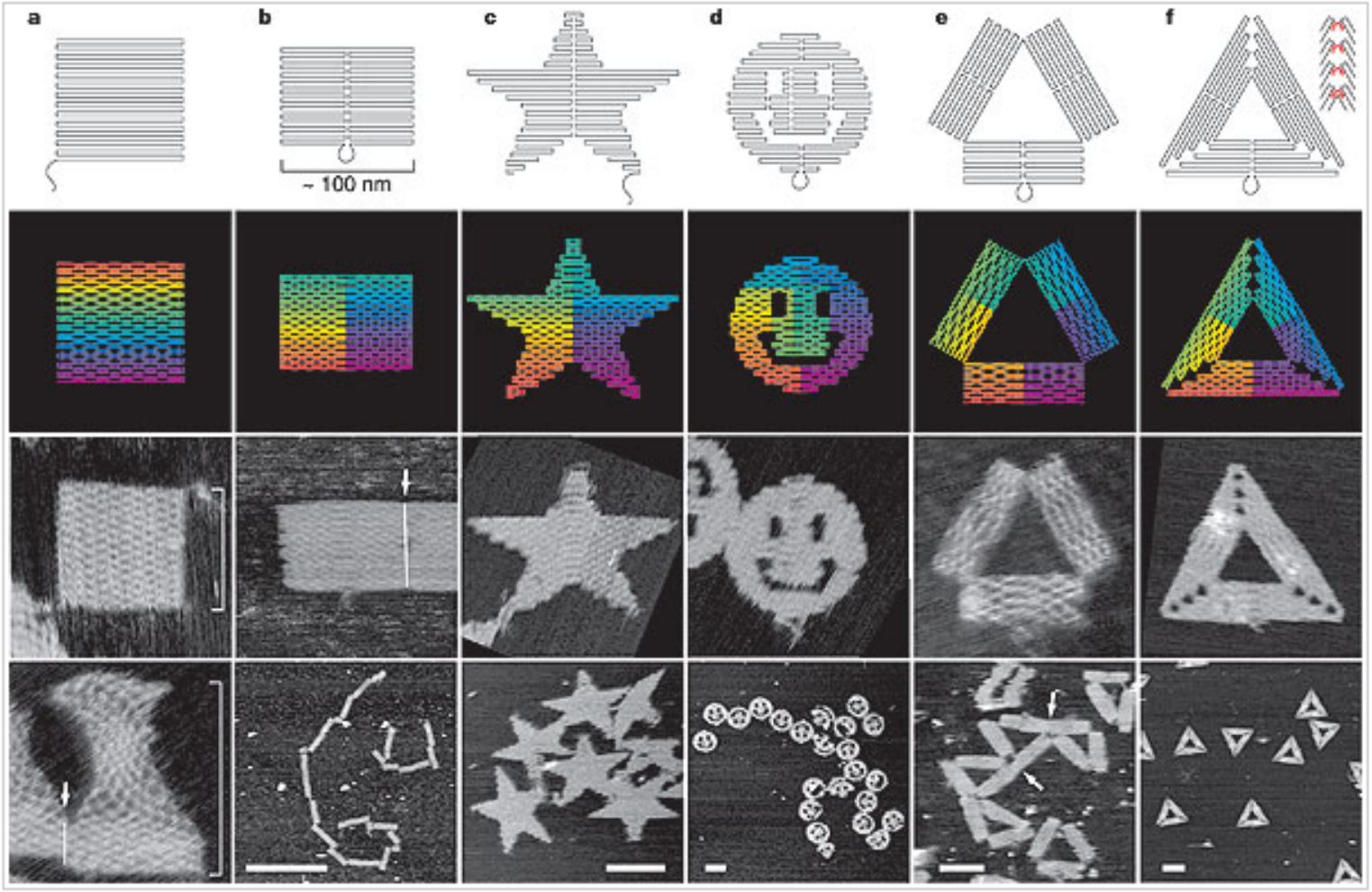
A. amplify PCR tags specific to SynIII. B gel to show that SynIII is shorter.
C. Cells containing synIII grow the same as WT.

The next step is synthetic yeast 2.0 to make the rest of the genome

www.syntheticyeast.org

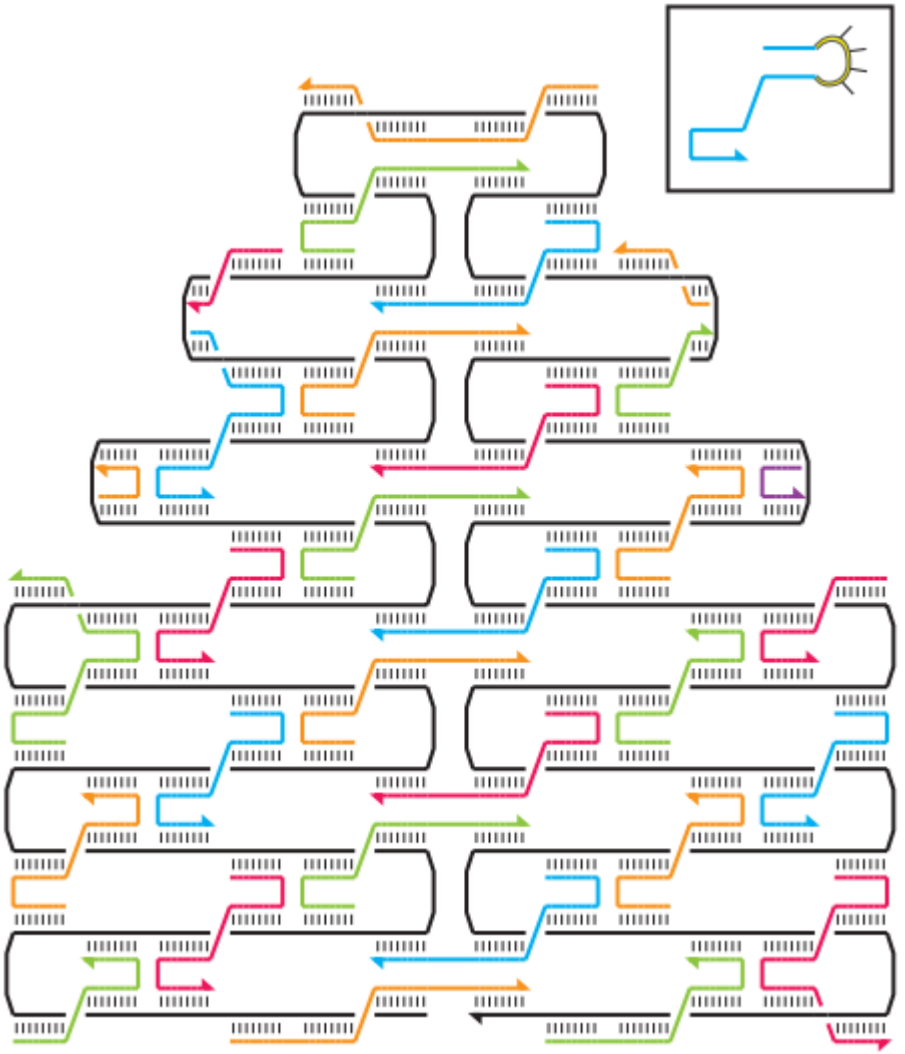
A few other futuristic applications for “Building with DNA”

DNA origami



http://www.nature.com/scitable/blog/bio2.0/dna_origami

DNA origami: fold ssDNA using oligo staples



http://bsclarified.wordpress.com

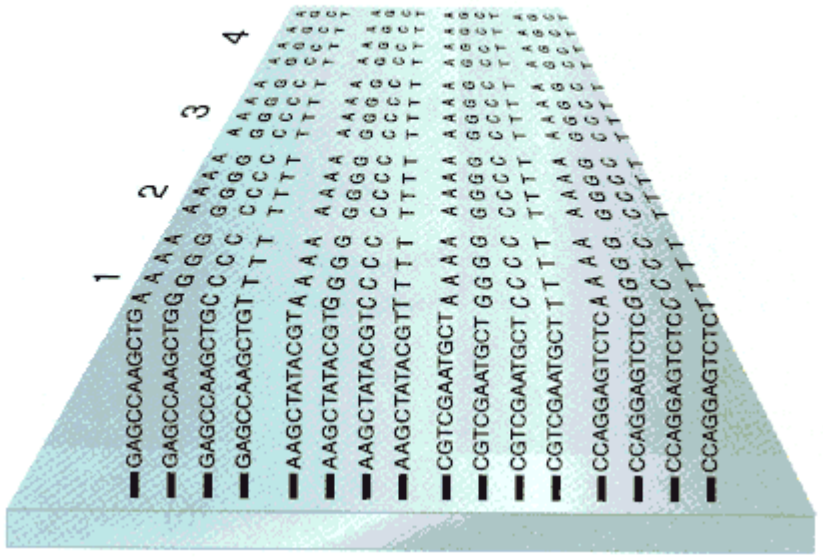
What are applications for DNA origami?

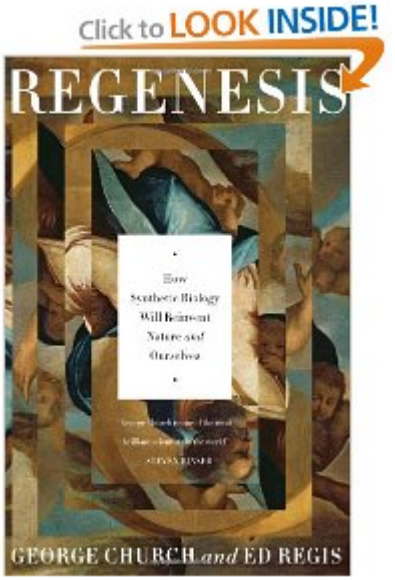
DNA as storage material

Next-Generation Digital Information Storage in DNA

(Griffiths et al. 1999)

Oligonucleotide array





http://www.colbertnation.com/the-colbert-report-videos/419824/october-04-2012/george-church