

Applications of Advanced Synthetic Biology

Genome Synthesis

Lecture Content

In this lecture we'll cover:

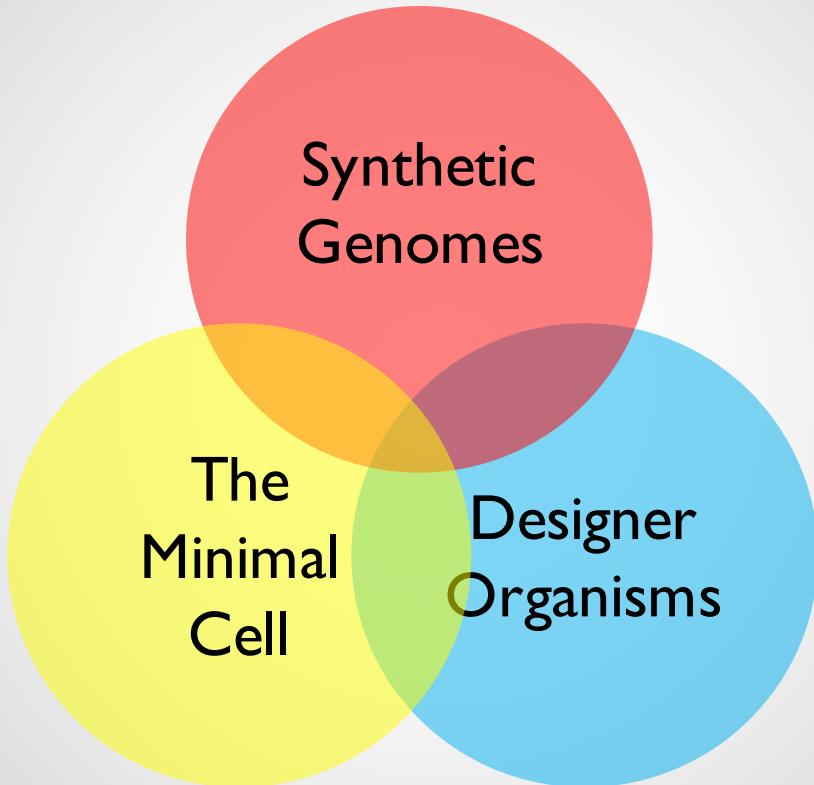
1. Craig Venter's *first cell made with a synthetic genome*
2. The DNA assembly techniques used
3. Whole-cell modelling
4. Sc2.0 – a synthetic yeast genome
5. RE.coli – a recoded E.coli genome
6. Applications of synthetic genomes

Learning Objectives

In this lecture you will learn:

1. Examples of genome synthesis projects
2. DNA Assembly techniques to allow genome-scale work
3. Examples of synthetic changes to make to genomes
4. Applications of synthetic and engineered genomes
5. How whole-cell models may work

The construction of synthetic organisms

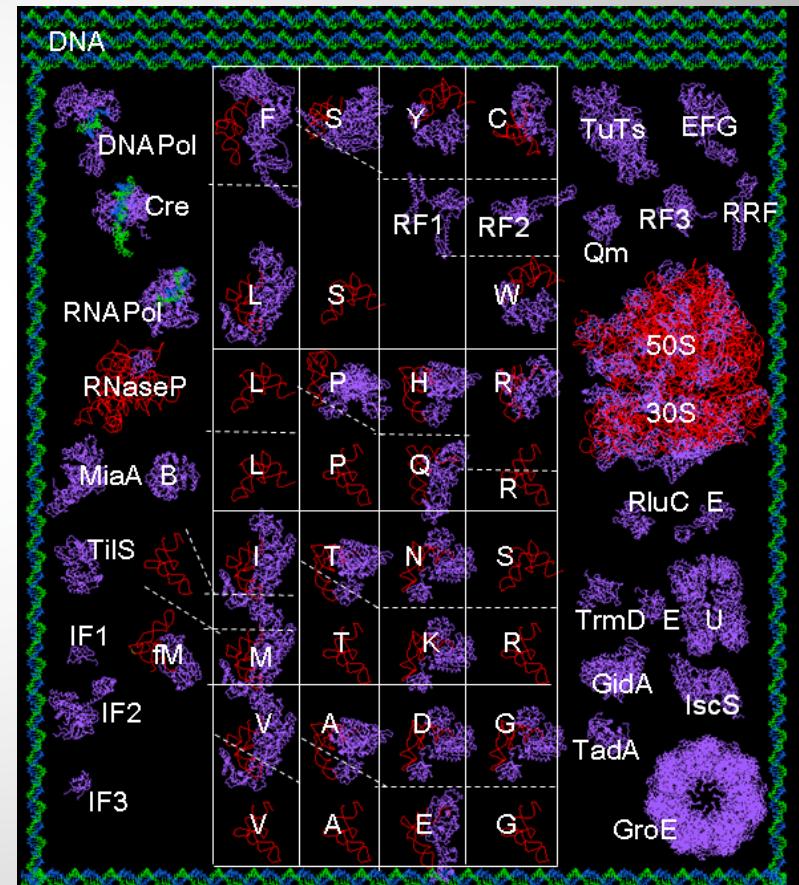
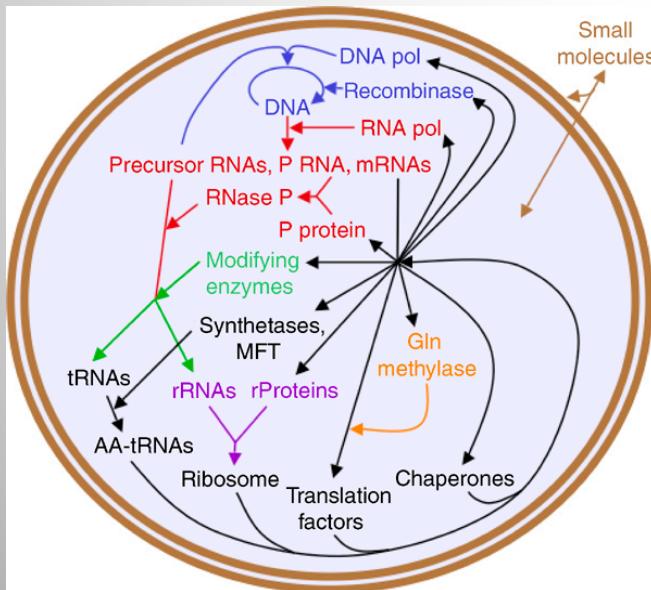


Could you build a minimal cell?

“We know enough about a cell to identify the essential molecules and build our own from scratch”

Biochemistry identifies the essential molecules that make cellular life

DNA → RNA → Protein



Genes required for a minimal cell...

- Estimate: 151 genes = 38 RNAs + 113 proteins (~200 is a better guess)

Basic DNA replication	2	Chaperones	2	Ribosome	63
RNA transcription	I	RNA Processing	3	tRNA set	33
Translation Factors	II	AA-tRNA synthetases	21	tRNA modifiers	15

Would require all metabolites (eg. NTP) to be provided – no metabolism

Would have no control over compartmentalisation – no membrane synthesis

Really minimal cell – fragile *in vitro* system

But... add metabolism, add lipid-synthesis for membranes, add proteins to control cell division, pores and transporters for sugar-import

Working minimal cell – capable of self-evolution

Alternative: Reduction to a minimal cell

Smallest natural genomes = 500 genes 500000 bps of DNA (e.g. *M.genitalium*)
But... not all genes are required for lab-based growth

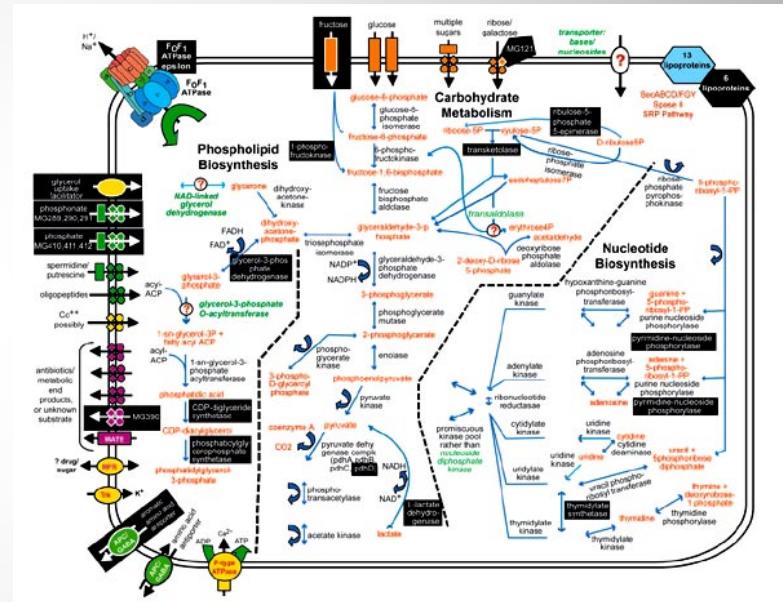
How many essential genes?

- I. Compare DNA throughout nature to identify essential genes

Estimates: 50 to 380

2. Delete genes of small genomes to see what is essential

Estimate: 400



But... Around a quarter of genes identified by these screens have unknown function

How do we really know that a gene is essential and not just playing many roles in a network?

Genome Synthesis – A Synthetic Cell



JCVI synthetic cell

10+ years
30+ people

>\$50 million

Aim: to build a minimal cell by synthesising a reduced genome

The first synthetic organism – JCVI Project

The 1st synthetic organism – life from a chemically synthesized genome

A big project by the J Craig Venter Institute initiated in 2000...

Step 1: Can a complete DNA genome be synthesised from chemicals?

Completed in 2008

Step 2: Can a cleaned DNA genome boot-up a cell?

Completed in 2007

Step 3: Combine Steps 1 and 2 to boot-up a cell from a synthesised genome

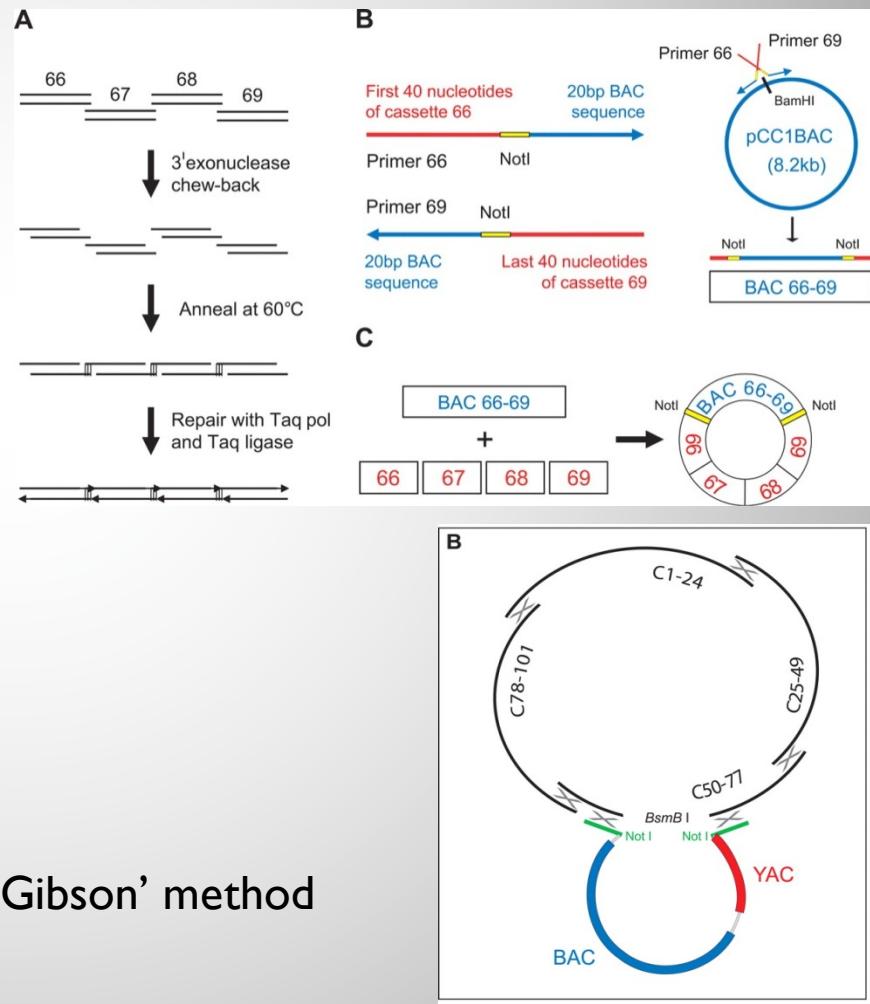
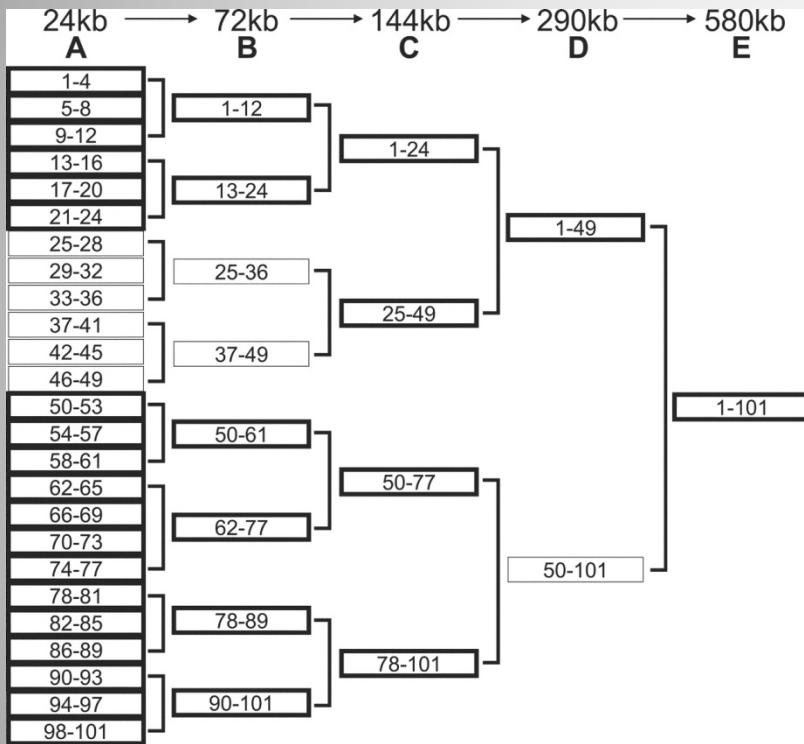
Finally Completed in 2010

The first synthetic organism: (I) Synthesis

AIM: Synthesize $\sim 10^4$ DNA 50-base oligomers and assemble into a complete error-free 582970 bp *M.genitalium* genome
(with watermarks)

1. Companies synthesise 101 pieces of 5 to 7 kb from overlapping oligos (e.g. Blue Heron and GeneArt)
2. 101 pieces recombined using *in vitro* enzymes to make 24 big pieces
3. 24 big pieces maintained in BACs in *E.coli* and recombined to make even bigger pieces
4. Big pieces all inserted into yeast and whole circular genome is made by recombination using native yeast genetics (using a YAC)
5. DNA sequencing used to check fidelity throughout process

The first synthetic organism: (I) Synthesis



101 synthesized fragments with overlap
 Stepwise *in vitro* DNA assembly using a new 'Gibson' method
 Final assembly using yeast as the vector

Gibson Isothermal Assembly (I)

Daniel Gibson's PCR-free method of annealing overlapping DNA sequences

- Overlapping sequences need to be 30 or more bp
- Requires a cocktail of T5 Exonuclease, Taq Ligase and Phusion Polymerase

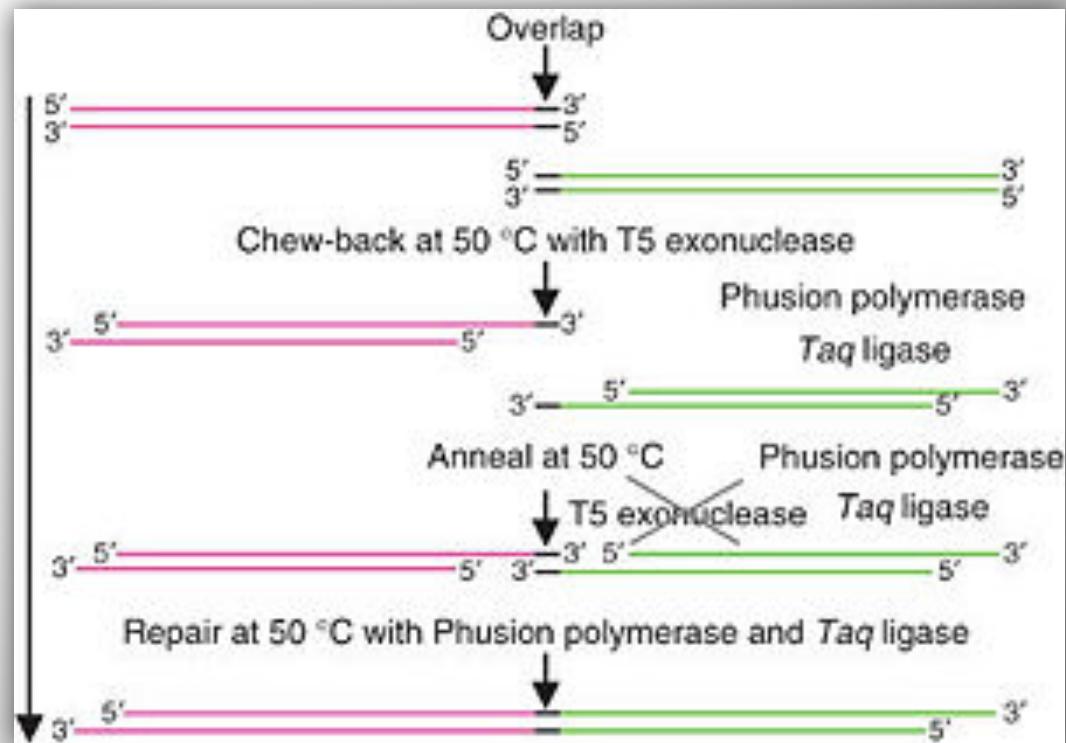
Make and store a master-mix
for months

Takes 30 minutes at 50
degrees C

Very simple and scalable

In vitro method

Requires overlapping DNA



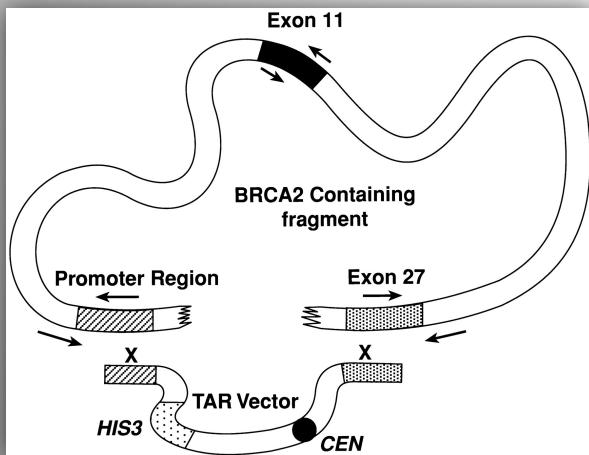
Yeast Assembly (TAR cloning)

Yeast can be transformed with DNA and will assemble overlapping DNA

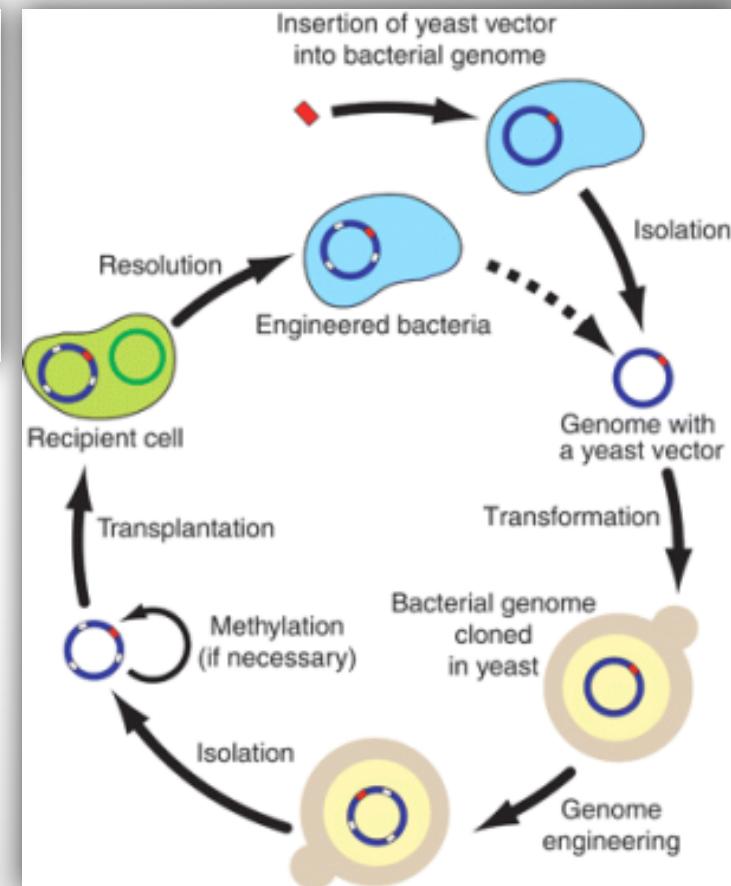
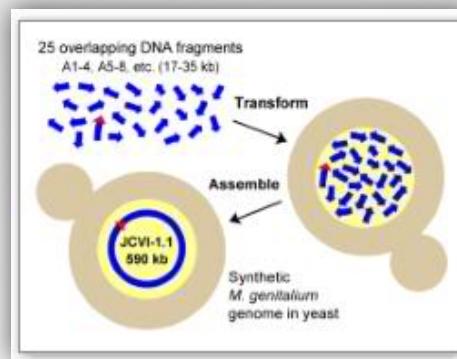
Based on old method of
'TAR Cloning'

Overlaps need to be
about 40 bp or more

Yeast can take >2.5 Mbp



All you
need is
yeast!



The first synthetic organism: (2) Booting-up

AIM: Genome A into Cell B → turn Cell B into Cell A

A: *M.mycoides* B: *M.capricolum* *different but compatible
biology

C Lartigue et al. Science 2007

Comparable to nuclei-switch experiments in *In Vitro* Fertilisation

Genomes are fragile to handle in the lab – maintain in agarose plugs

How to get DNA into cell B? – incredibly inefficient, requires cell fusions
(lucky that Mycoplasma have no cell wall)

What happens to genome of cell B? – doesn't have antibiotic resistance

Verify with sequencing, proteomics and phenotyping – it definitely changed!

Finally making the synthetic cell - 2010

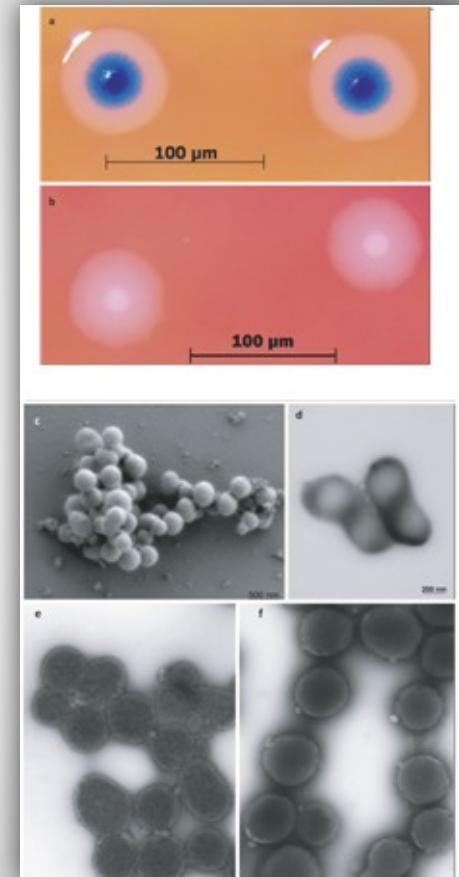
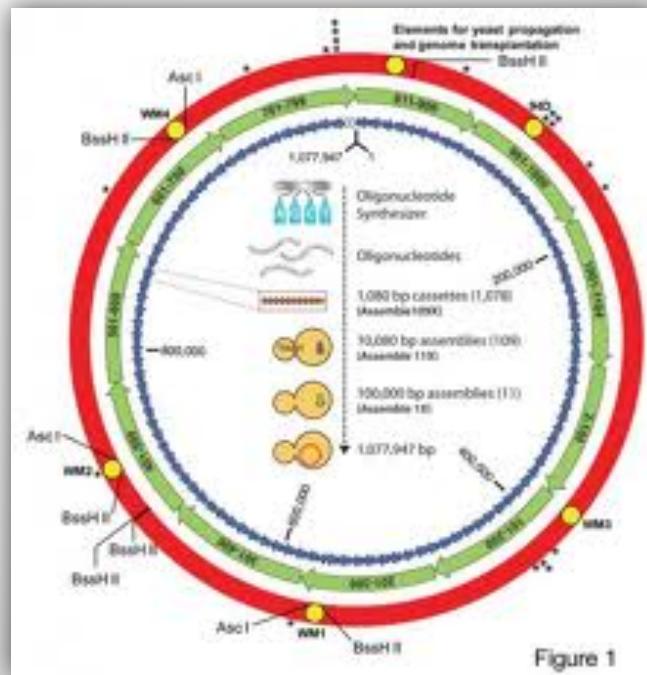
They got it to work: D.G. Gibson *et al* 2010, Science 329 (5987): 52-56

Never achieved for *M.genitalium*.... But....

Switched to a different cell
M. Mycoides
(worked before!)

But..
M. Mycoides not a minimal genome cell

Single-base error in synthesis set them back by months



The first synthetic cell – ‘Synthia’

HOW TO MAKE ARTIFICIAL LIFE

1 Entire DNA of Mycoplasma mycoides, a bug that usually infects goats, is decoded.

2 Researchers buy fragments of DNA from a mail order catalogue. Each of the four bottles of chemicals contains a section of the code.

3 The fragments are put into yeast, which 'stitches' them together, gradually building a synthetic copy of the original DNA.

4 The artificial DNA is put into a recipient bacterium, which then grows and divides, creating two daughter cells, one with the artificial DNA and one with the natural DNA.

5 Antibiotics in the petri dish kill the bacterium with the natural DNA, leaving the one with the synthetic DNA to multiply.

6 Within just a few hours, all traces of the recipient bug are wiped out and bugs with artificial DNA thrive. New life has been created.

7 Possible uses are bugs capable of producing clean fuels and sucking carbon dioxide out of the atmosphere. Also microbes capable of mopping up oil after a spillage or spillages.

Maverick: Dr Craig Venter

Artificial DNA **Natural DNA**

Synthetic DNA code

Entire DNA of Mycoplasma mycoides, a bug that usually infects goats, is decoded.

Researchers buy fragments of DNA from a mail order catalogue. Each of the four bottles of chemicals contains a section of the code.

The fragments are put into yeast, which 'stitches' them together, gradually building a synthetic copy of the original DNA.

The artificial DNA is put into a recipient bacterium, which then grows and divides, creating two daughter cells, one with the artificial DNA and one with the natural DNA.

Antibiotics in the petri dish kill the bacterium with the natural DNA, leaving the one with the synthetic DNA to multiply.

Possible uses are bugs capable of producing clean fuels and sucking carbon dioxide out of the atmosphere. Also microbes capable of mopping up oil after a spillage or spillages.

Alive ... synthetic cell

DAILY Mirror Friday May 21, 2010

SCIENTISTS 'CREATE SIMPLE LIFE FORM'

By DANNY MCKELLAR A BILLIONAIRE biotech tycoon claims to have created some of the first synthetic life forms. And he says his latest milestone could have a major impact on healthcare, ushering in new treatments and drugs. The £100m research programme, using 15 years of computer-aided work, inserted artificial DNA, designed in a computer, and watched it reproduce like crazy. The type is typical of how biotechnology companies try to boost their share price, according to 'Frankenstein marketing'. But Dr Mark Dosez of charity the Society of Biology said: "This is a milestone in biology research and it's great to see scientists continuing to bring new scientific opportunities."

A PIONEERING Craig Venter

Dr Craig Venter says synthetic life forms can benefit humanity. Critics say he is 'playing God'. (Photograph: Polka Photo)

DOC CREATES LIFE 'Frankenstein' grows DNA to bring cell back from dead

By EMMA MORTON
Health and Science Editor

It was last night hailed as a breakthrough that opens the door to exciting new technological advances. But opponents of genetic engi-

Alive ... synthetic cell

neering condemned the moment as dangerous Frankenstein-style tampering with life.

Maverick US biotech tycoon Dr Craig Venter first extracted genes from a bacterium, *Mycoplasma capricolum*. Then he grew strands of artificial DNA and put them into the dead bacterium.

Then he grew strands of artificial DNA and put them into the dead bacterium.

THE Sun

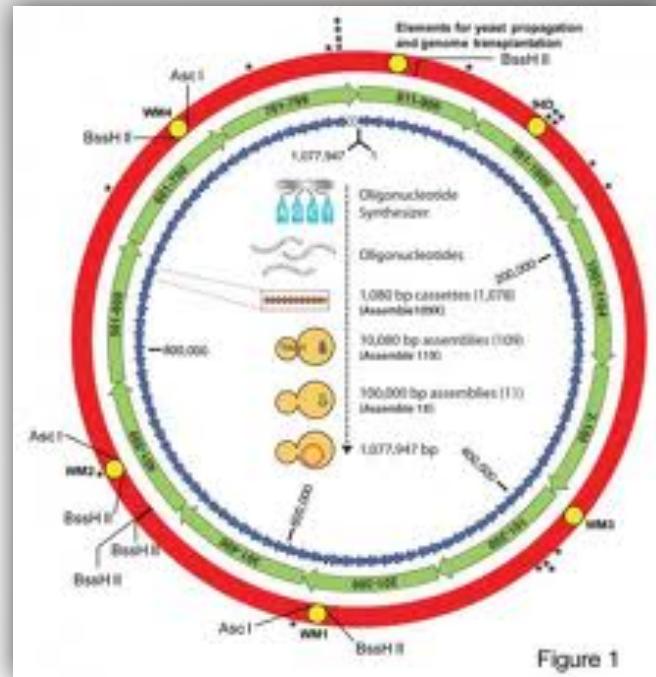
Fri May 21, 2010

But is it a minimal cell?

M. mycoides is >1 million bp

M. genitalium is ~550,000 bp

Attempts to reduce synthetic *M. mycoides* genome by synthesising versions with less genes have not yet worked

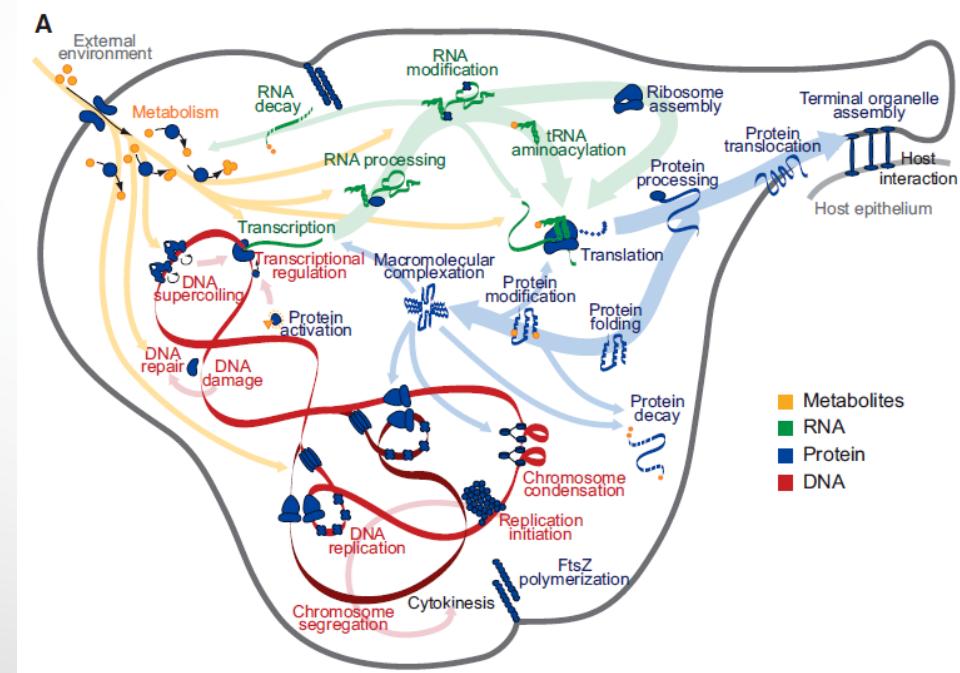
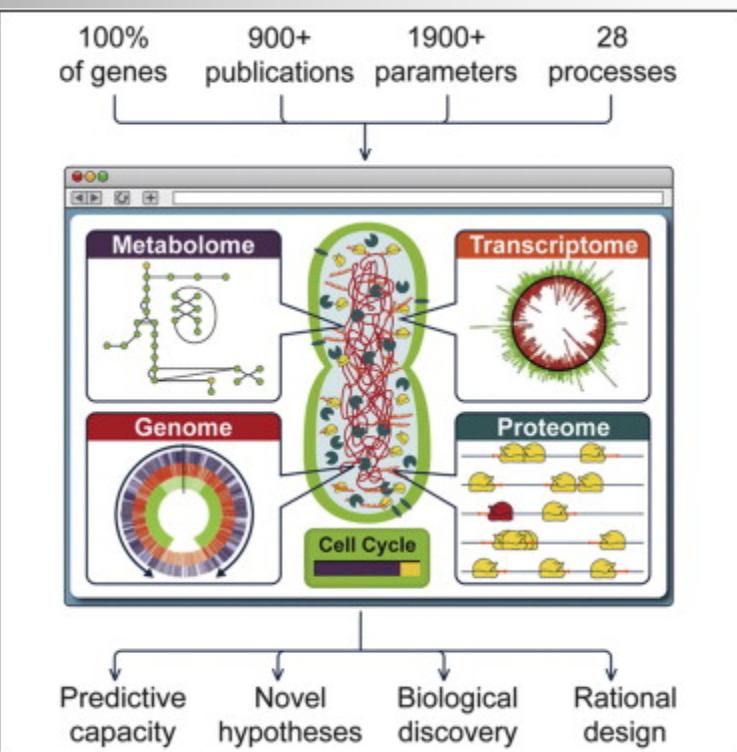


Can a whole-cell model help the project?

A Whole-Cell Computational Model Predicts Phenotype from Genotype

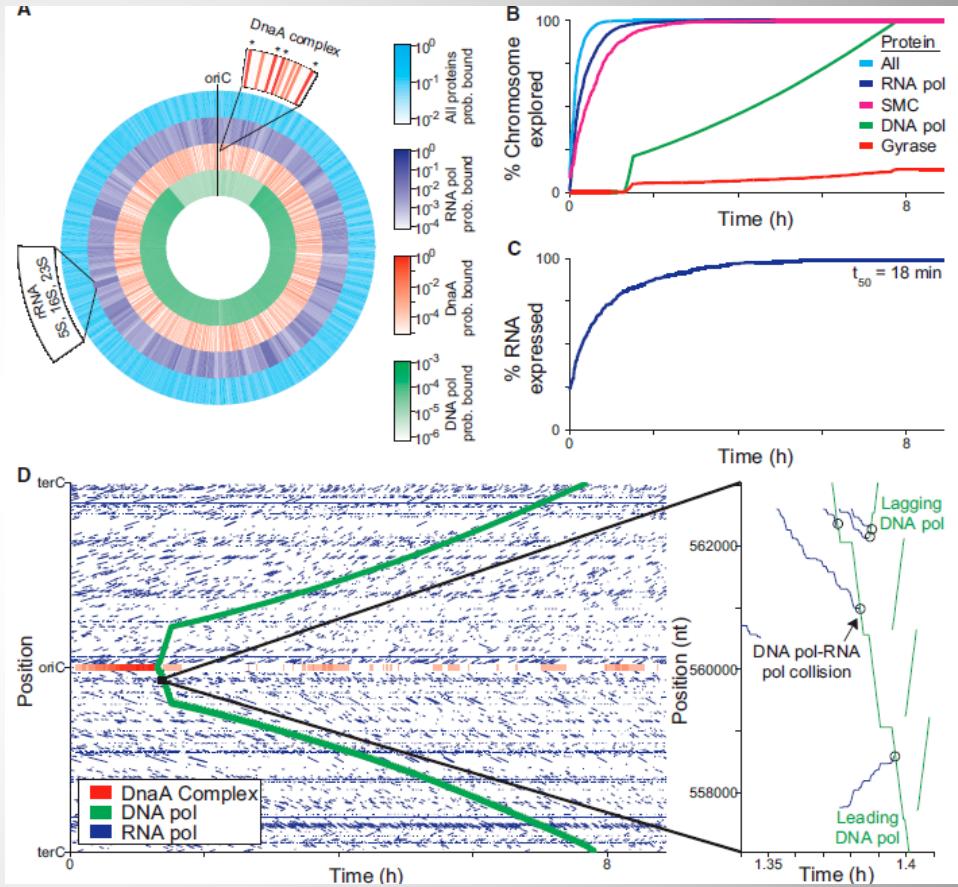
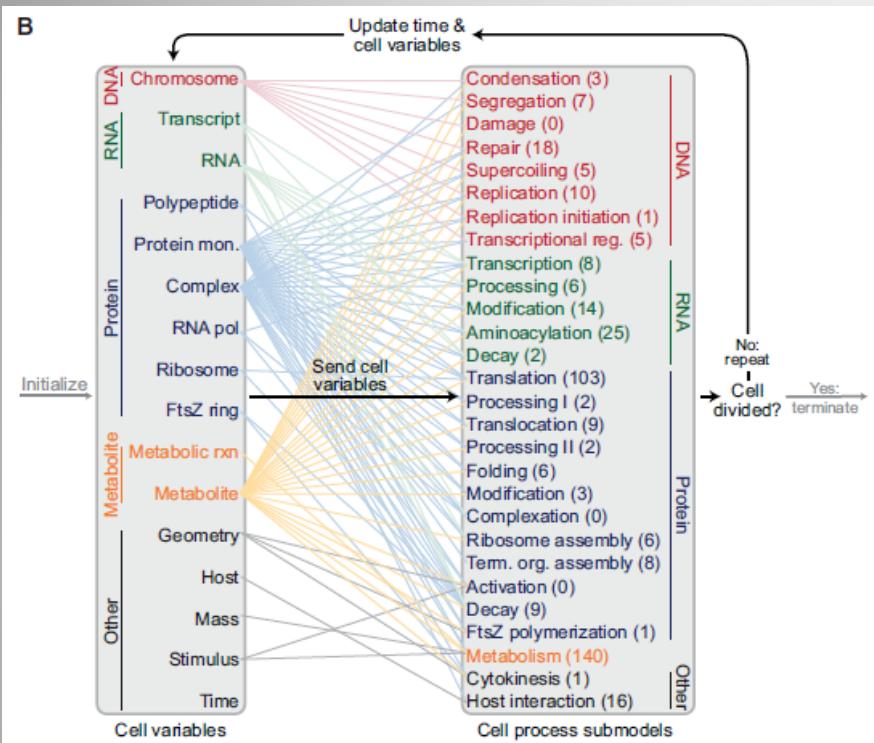


2012



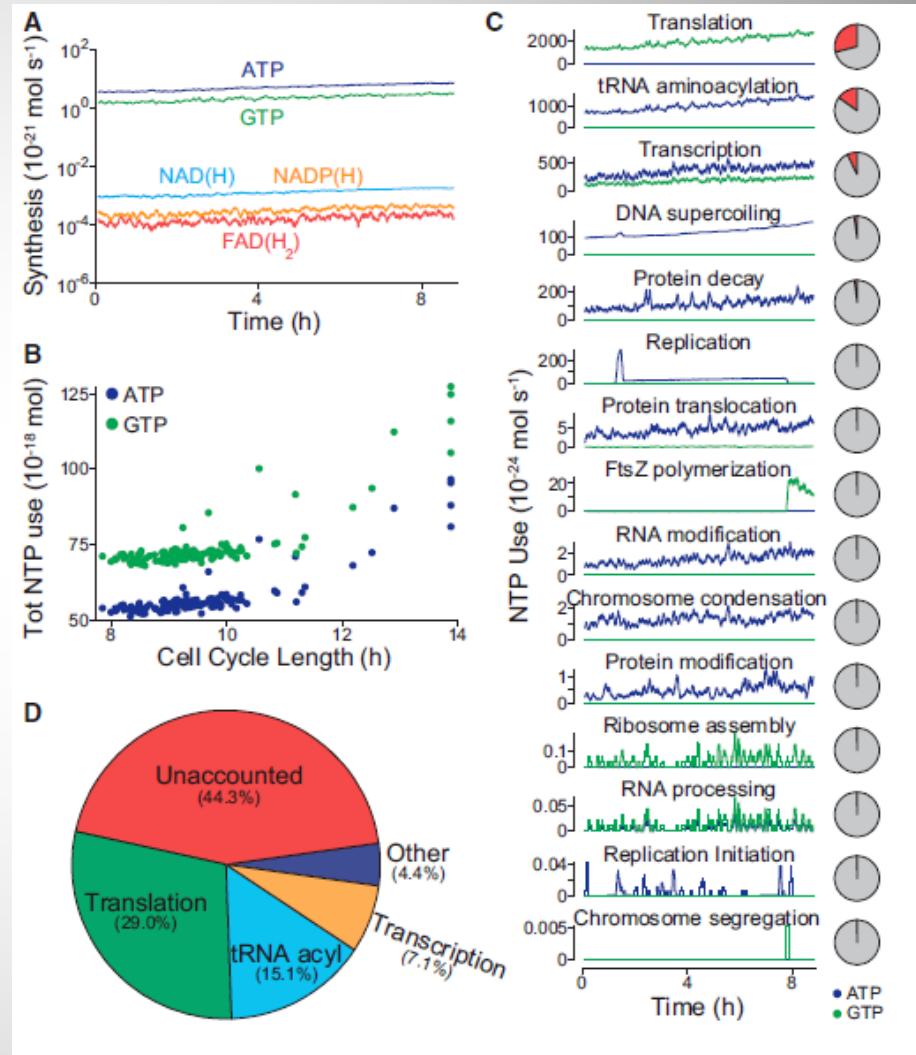
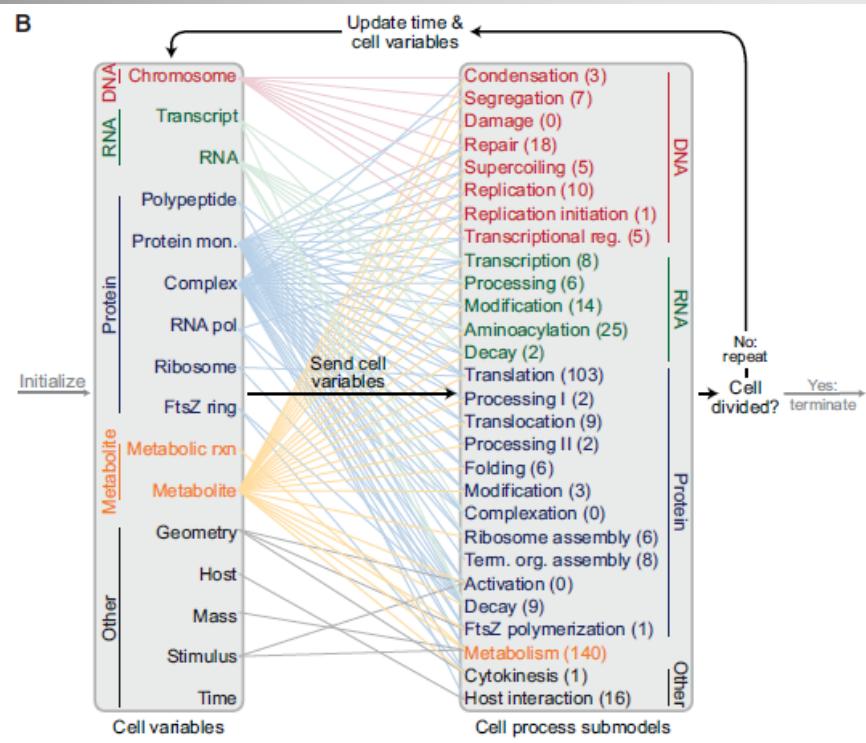
Tying together lots of models into one

<http://www.youtube.com/watch?v=au4sl9CjKFU>



Separate models of processes
 Data taken from large studies
 Can give simulation of a cell cycle

Whole-cell model allows new insights

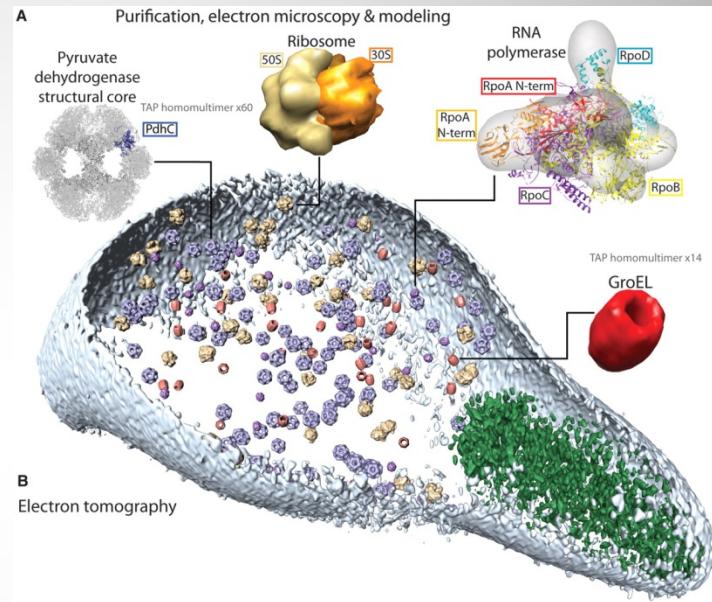
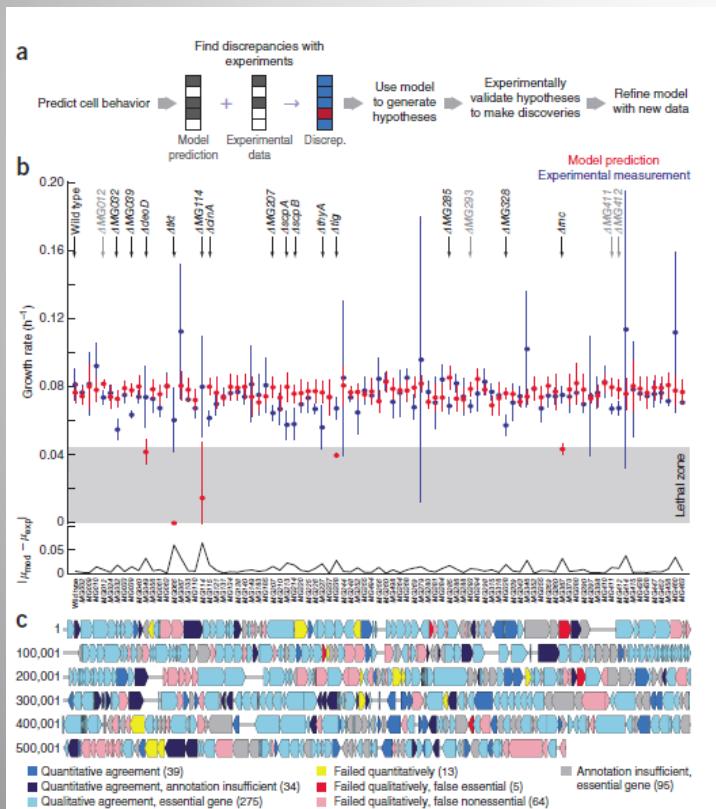


How are energy and resources used by all the different processes?

What is the function of all the genes?

Model can help predict but not yet perfect...

Sanghvi et al. Accelerated discovery via a whole-cell model (Nature Methods 2013) > M. genitalium



M. pneumoniae is the best-studied small cell
Science: 27th November 2009
Guell et al. – Systems biology study
Yus et al. - Metabolism study
Kunhner et al. - Proteome study
Data now being used to make whole-cell simulation

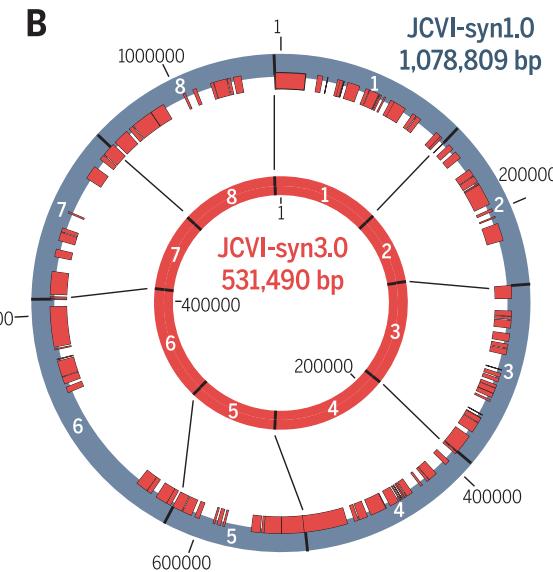
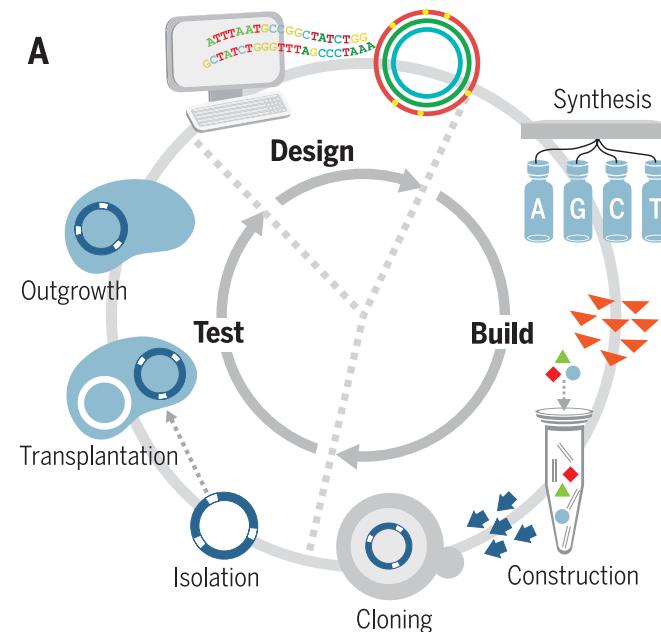
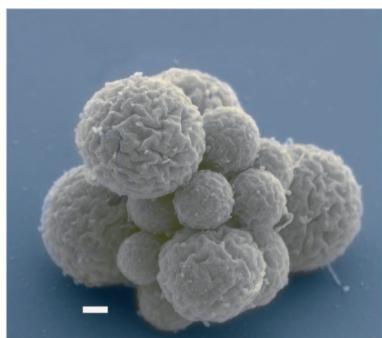
2016 Update: A Minimised Genome!

RESEARCH ARTICLE SUMMARY

SYNTHETIC BIOLOGY

Design and synthesis of a minimal bacterial genome

Clyde A. Hutchison III,^{*†} Ray-Yuan Chuang,[†] Vladimir N. Noskov, Nacyra Assad-Garcia, Thomas J. Deerinck, Mark H. Ellisman, John Gill, Krishna Kannan, Bogumil J. Karas, Li Ma, James F. Pelletier, Zhi-Qing Qi, R. Alexander Richter, Elizabeth A. Strychalski, Lijie Sun, Yo Suzuki, Billyana Tsvetanova, Kim S. Wise, Hamilton O. Smith, John I. Glass, Chuck Merryman, Daniel G. Gibson, J. Craig Venter^{*}



Assembly and Design of a 500,000 bp genome

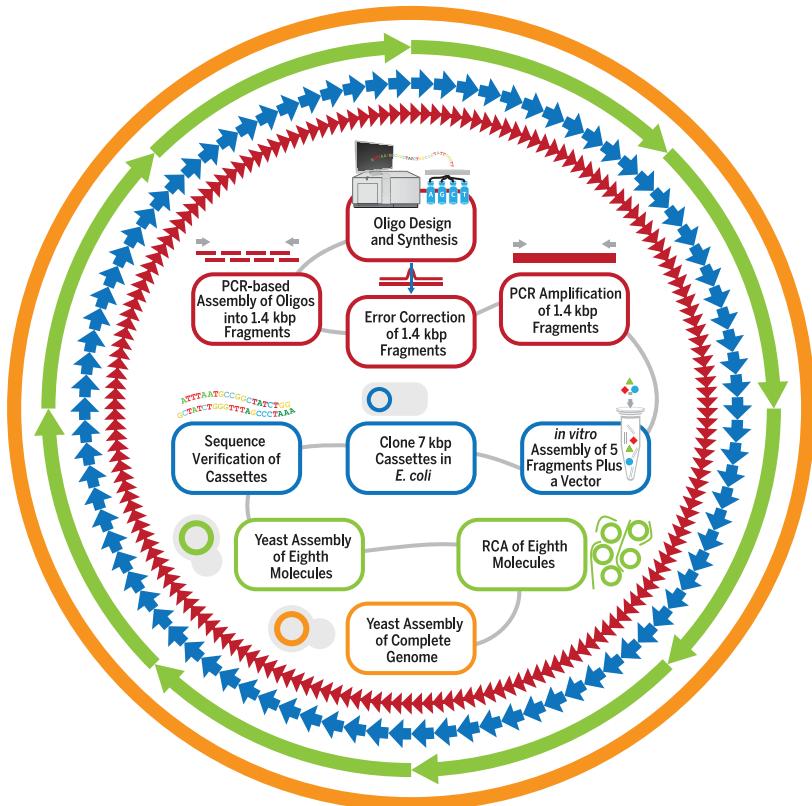


Fig. 2. Strategy for whole-genome synthesis. Overlapping oligonucleotides (oligos) were designed, chemically synthesized, and assembled into 1.4-kbp fragments (red). After error correction and PCR amplification, five fragments were assembled into 7-kbp cassettes (blue). Cassettes were sequence-verified and then assembled in yeast to generate one-eighth molecules (green). The eight molecules were amplified by RCA and then assembled in yeast to generate the complete genome (orange).

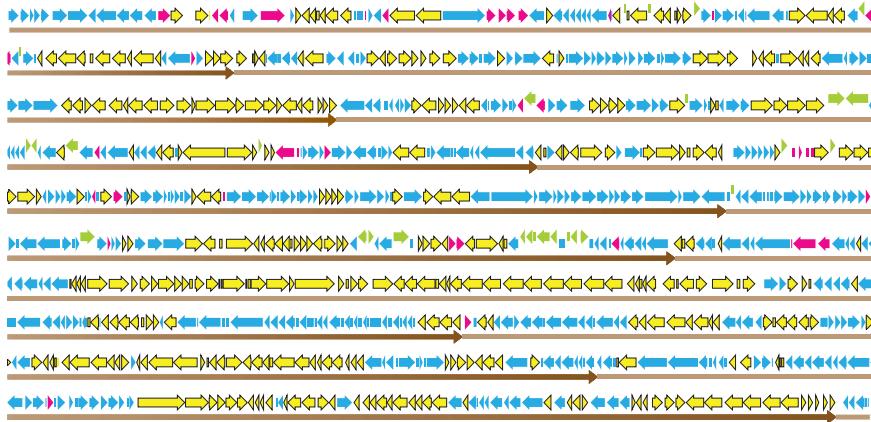
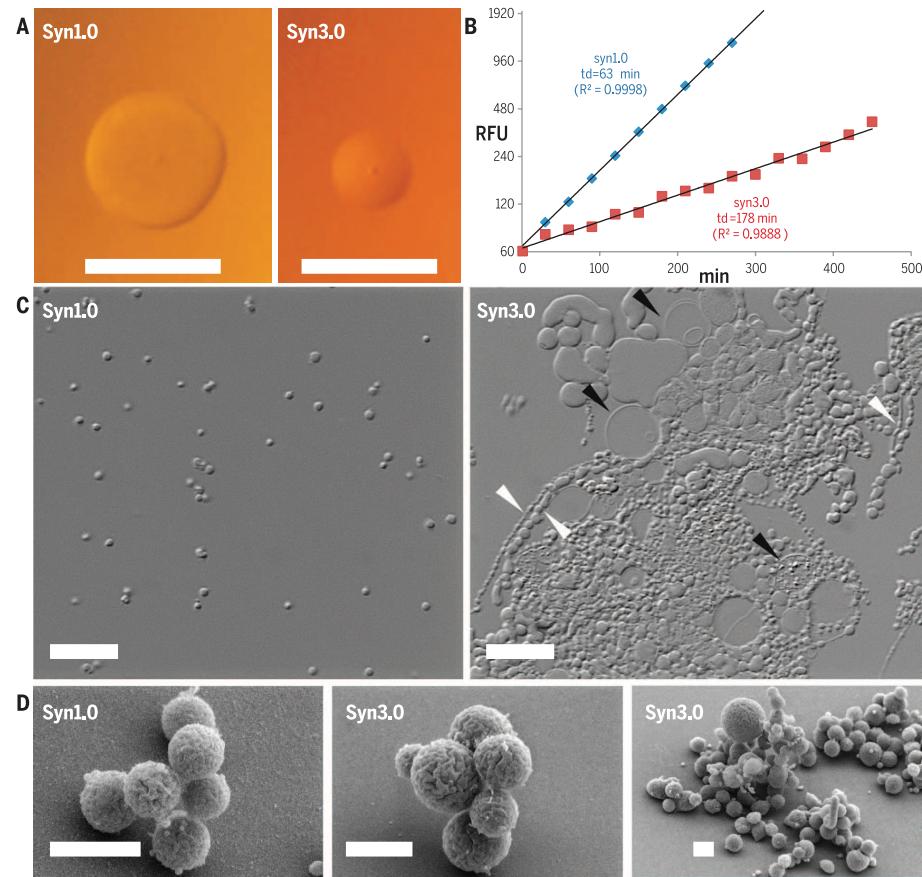


Fig. 4. The three DBT cycles involved in building syn3.0. This detailed map shows syn1.0 genes that were deleted or added back in the various DBT cycles leading from syn1.0 to syn2.0 and finally to syn3.0 (compare with fig. S7). The long brown arrows indicate the eight NotI assembly segments. Blue arrows represent genes that were retained throughout the process. Genes that were deleted in both syn2.0 and syn3.0 are shown in yellow. Green arrows (slightly offset) represent genes that were added back. The original RGD1 design was not viable, but a combination of syn1.0 segments 1, 3, 4, and 5 and designed segments 2, 6, 7, and 8 produced a viable cell, referred to as RGD2678. Addition of the genes shown in green resulted in syn2.0, which has eight designed segments. Additional deletions, shown in magenta, produced syn3.0 (531,560 bp, 473 genes). The directions of the arrows correspond to the directions of transcription and translation.

Deletion of genes leads to worse growth

Table 1. Syn1.0 genes listed by functional category and whether they were kept or deleted in syn3.0. Categories with asterisks are mostly kept in syn3.0, whereas those without are depleted in syn3.0. Vector sequences, for selection of the genome and for propagation in other hosts, are not included in these gene tallies.

Functional category	Kept	Deleted
Glucose transport and glycolysis*	15	0
Ribosome biogenesis*	14	1
Protein export*	10	0
Transcription*	9	0
RNA metabolism*	7	0
DNA topology*	5	0
Chromosome segregation*	3	0
DNA metabolism*	3	0
Protein folding*	3	0
Translation*	89	2
RNA (rRNAs, tRNAs, small RNAs)*	35	4
DNA replication*	16	2
Lipid salvage and biogenesis*	21	4
Cofactor transport and salvage*	21	4
rRNA modification*	12	3
tRNA modification*	17	2
Efflux*	7	3
Nucleotide salvage	19	8
DNA repair	6	8
Metabolic processes	10	10
Membrane transport	31	32
Redox homeostasis	4	4
Proteolysis	10	11
Regulation	9	10
Unassigned	79	134
Cell division	1	3
Lipoprotein	15	72
Transport and catabolism of nonglucose carbon sources	2	34
Acyl/glycerol breakdown	0	4
Mobile elements and DNA restriction	0	73
Total	473	428



Summary of JCVI Genomes

Table 2. Reduced genomes resulting from the DBT cycles, ultimately leading to syn3.0. Column 1 indicates the round of genome design (dashes indicate the starting genome, syn1.0), column 2 gives the size of the designed genome (in kilobase pairs), and column 3 gives the number of mycoplasma genes in the design. Column 4 shows the genome composition for key viable cell strains; for nonviable designs, a viable strain with the highest number of segments from the design is shown, as well as a more robust alternative for RGD1.0 (fourth row) and a smaller derivative for RGD2.0 (sixth row, syn2.0). Column 5 gives the size of the genome corresponding to column 3, and column 6 shows a quantitative or qualitative estimate of the growth rate of cells with the genome described in column 4.

1. Genome design	2. Design size	3. Number of design genes	4. Cellular genome segment composition for key viable strains	5. Cellular genome size	6. Growth rate
-		901	syn1.0: all eight syn1.0 segments	1079 kbp	Doubling time, td = 60 min
HMG	483 kbp	460	HMG segment 2 + 7/8 syn1.0	1003 kbp	Slow-growing
RGD1.0	544 kbp	483	RGD1.0 segments 2,6,7,8 + syn1.0 segments 1,3,4,5	758 kbp	Slow-growing
"	"		RGD1.0 segments 1,2,4,6,8 + syn1.0 segments 3,5,7	718 kbp	Slow-growing
RGD2.0	575 kbp	512	RGD2.0 segments 1,2,3,4,6,7,8 + syn1.0 segment 5	617 kbp	?
"	"		syn2.0: RGD2.0 segments 1,2,3,4,6,7,8 + syn1.0 segment 5 with genes <i>MMSYN1_0454</i> to <i>-0474</i> and <i>MMSYN1_0483</i> to <i>-0492</i> deleted	576 kbp	td = 92 min
RGD3.0	531 kbp	473	syn3.0: all eight segments of RGD3.0	531 kbp	td = 180 min

What about a minimal Eukaryotic Cell?

Synthesising a small eukaryote genome (11 million bp)

Project = Synthetic Yeast (Sc2.0) <http://syntheticyeast.org/>

Undergrad course & iGEM 2008 = JHU Build-A-Genome Project
Now international (China, Singapore, USA, Australia and UK)



Team:Johns Hopkins-BAG/B-A-G course



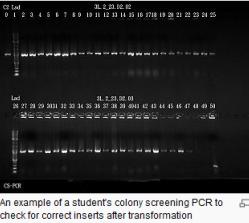
JHU Build A Genome

Moodle Synthetic Yeast Wiki Protocols References Acknowledgements Contact

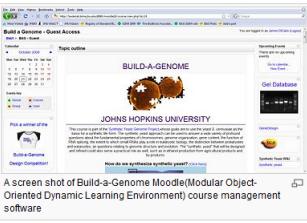
B-A-G Course

The Build-a-Genome course offers a cool environment in which undergraduate students can participate in cutting-edge interdisciplinary research. In pursuing the synthesis of *Saccharomyces cerevisiae* Sc2.0 chromosomes, students gain hands-on experience in synthetic biology research. The *de novo* synthesis of genetic information offers the promise of a deeper understanding of chromosome structure, gene function, and gene order. The ultimate goal of this class is the development of students into skilled, independent researchers with well-developed troubleshooting skills and experience and familiarity with molecular biology and computational approaches to genomics problems.

The Build-a-Genome course consists of lectures, "Molecular Biology Boot Camp", and eventually, independent research. The lectures offered in the course reflect the many-faceted underpinnings of synthetic biology, ranging in topics from fundamentals of genetics (such as nucleic acid structure and function, chromosome structure, and genome organization) to bioinformatics, to central concepts of synthetic biology (such as recombinant DNA technology, gene synthesis, synthetic circuitry and of course iGEM). After introductory lectures are complete, students go through eight sessions of molecular biology "boot camp" that serve as a period to review lecture topics, master lab techniques, and learn the methods used in this project. Graduation from the boot camp requires students to submit assignments that verify proficiency in each step of the gene synthesis protocol – such as PCR, agarose gel electrophoresis, molecular cloning, sequence analysis, etc.

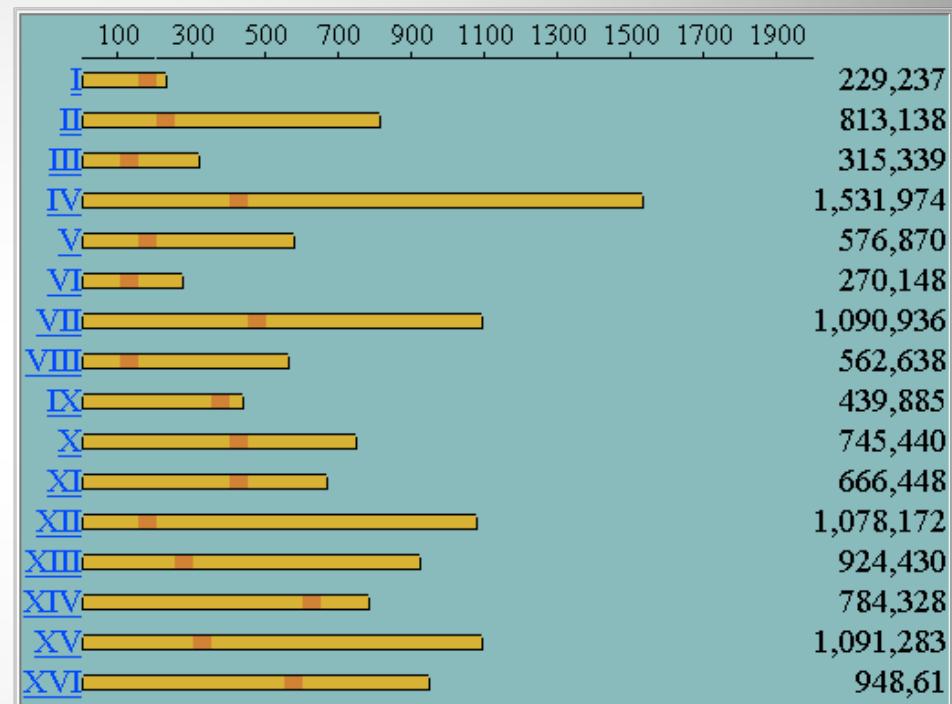


An example of a student's colony screening PCR to check for correct inserts after transformation



A screen shot of Build-a-Genome Moodle/Modular Object-Oriented Dynamic Learning Environment course management software.

Saccharomyces cerevisiae complete genome



Undergrads working on building synthetic parts of an 11 million bp genome with 16 chromosomes

Dymond et al 2011 - 2 synthetic chromosome arms in yeast

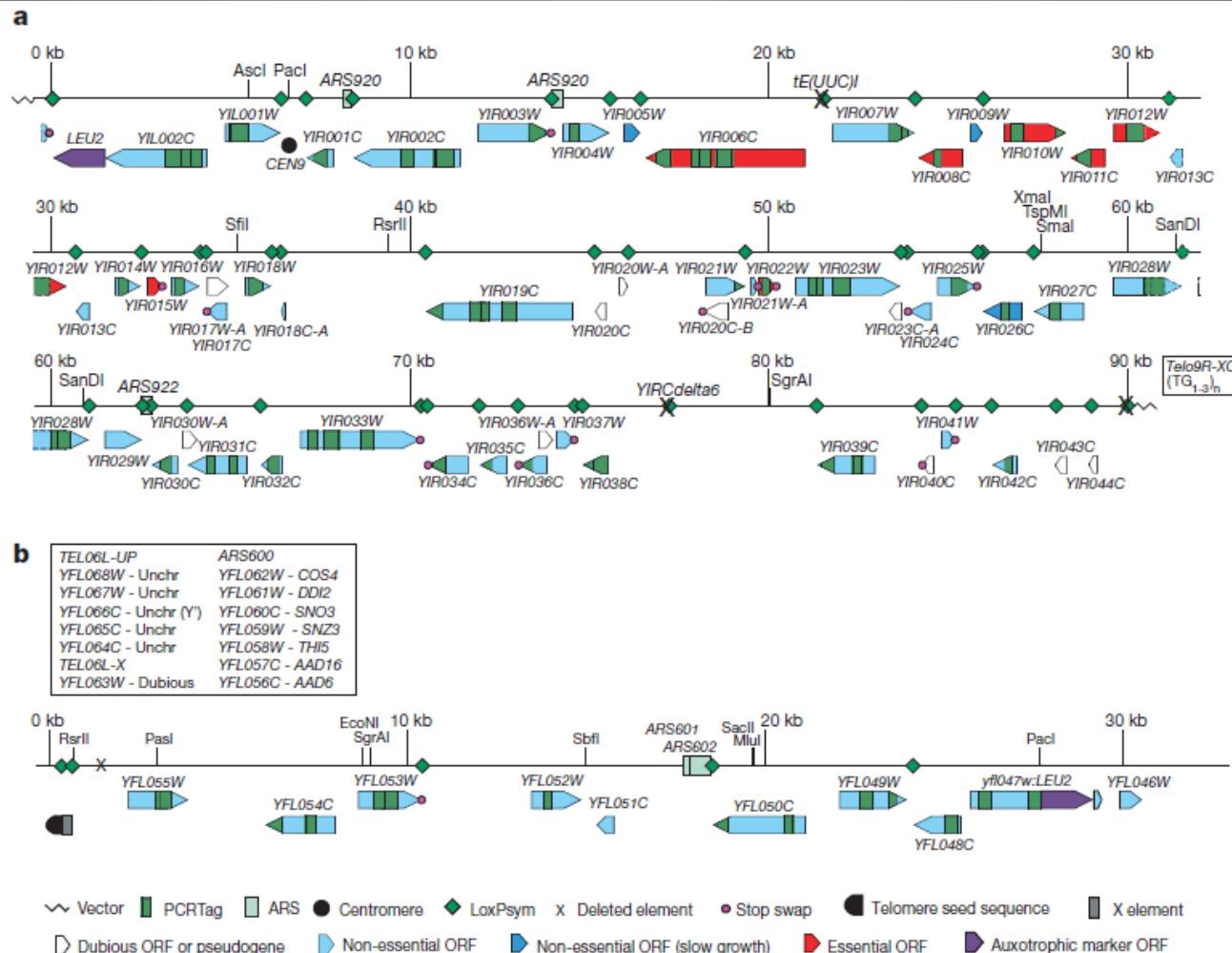


Figure 1 | Maps of synIXR and semi-synVIL. Boxed text indicates elements deleted in the synthetic chromosomes. Vertical green bars inside ORFs indicate PCRTag amplicons; only sequences at the outside edges of these are recoded.

ARS, autonomously replicating sequence. **a**, SynIXR. Vector is circular. **b**, Semi-synVIL.

Two chromosome arms working by 2011

2011 Nature Paper

(make sure to also look at the Supplementary materials)

Jessica S. Dymond et al. Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. Nature, 2011; DOI: 10.1038/nature10403

Same basic idea as the iGEM 2009 project and the Build-A-Genome Course

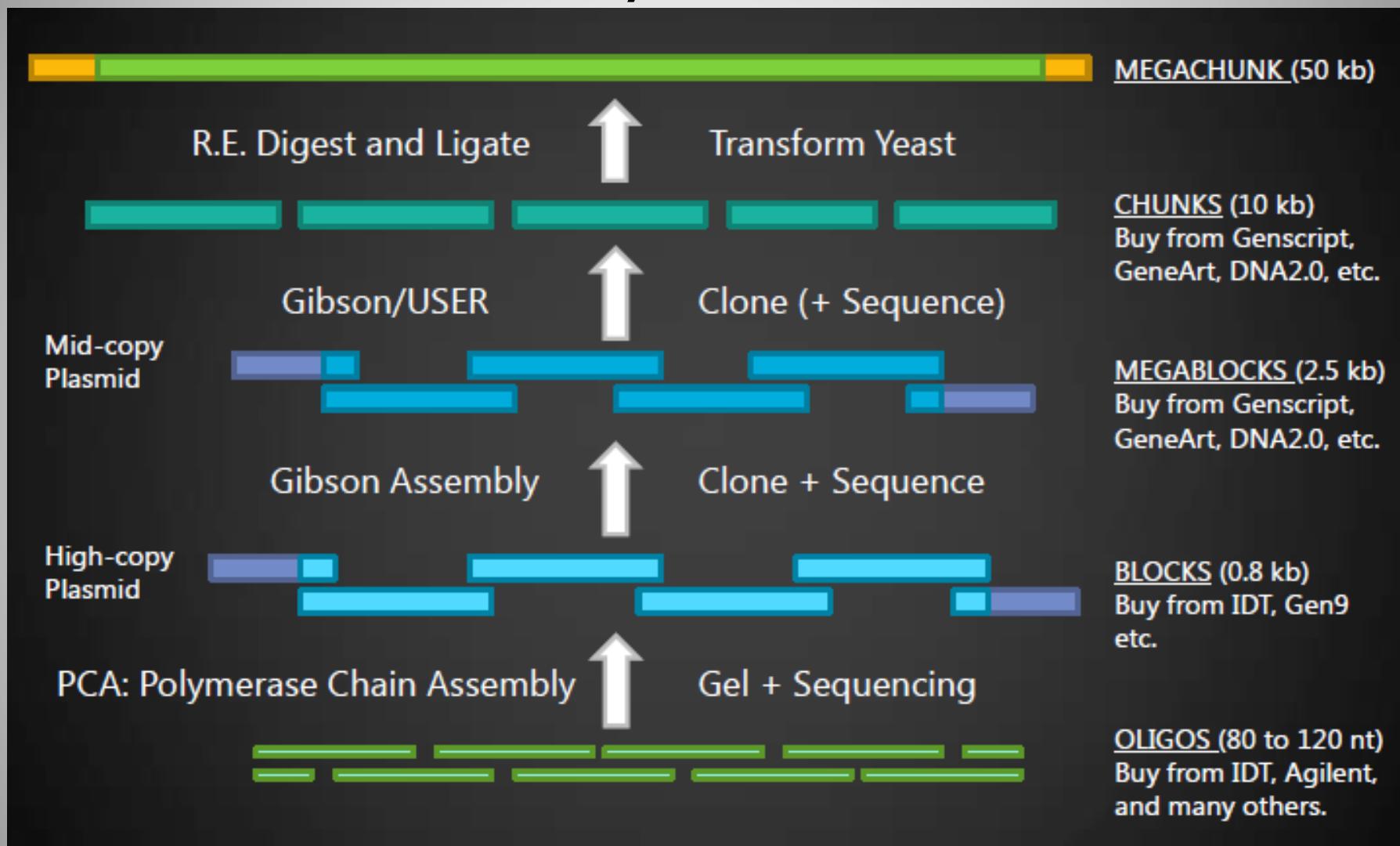
1. synIXR Chromosome 9 right arm added in one go

made by commercial gene synthesis and assembled into a BAC (bacterial artificial chromosome) by the company Codon Devices.

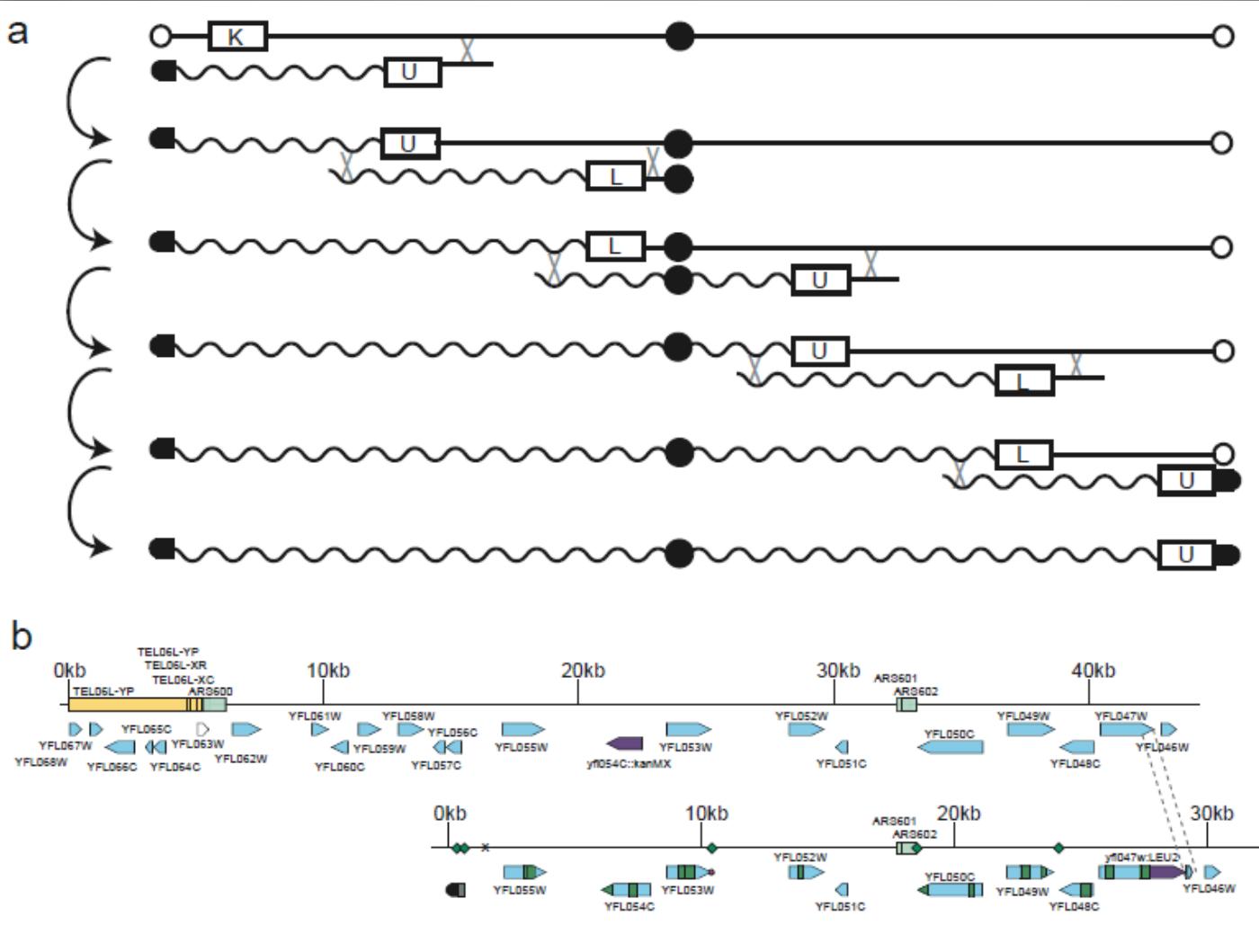
2. semi-synVIL Chromosome 6 left arm 4 chunks added serially

4 big chunks (up to 10,000 bp) were assembled from synthetic oligos by Epoch Biolabs

DNA Assembly at different scales



How to replace native sequence with synthetic in yeast



2014 – First Chromosome Finished!

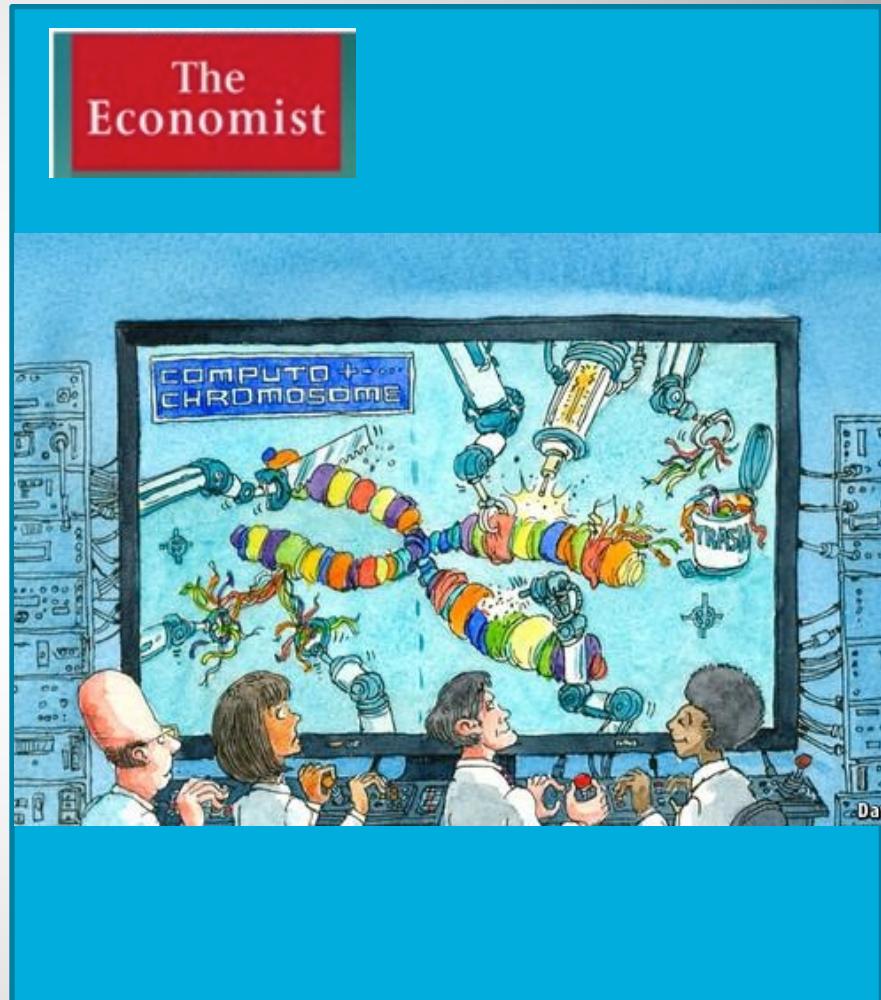


Resea

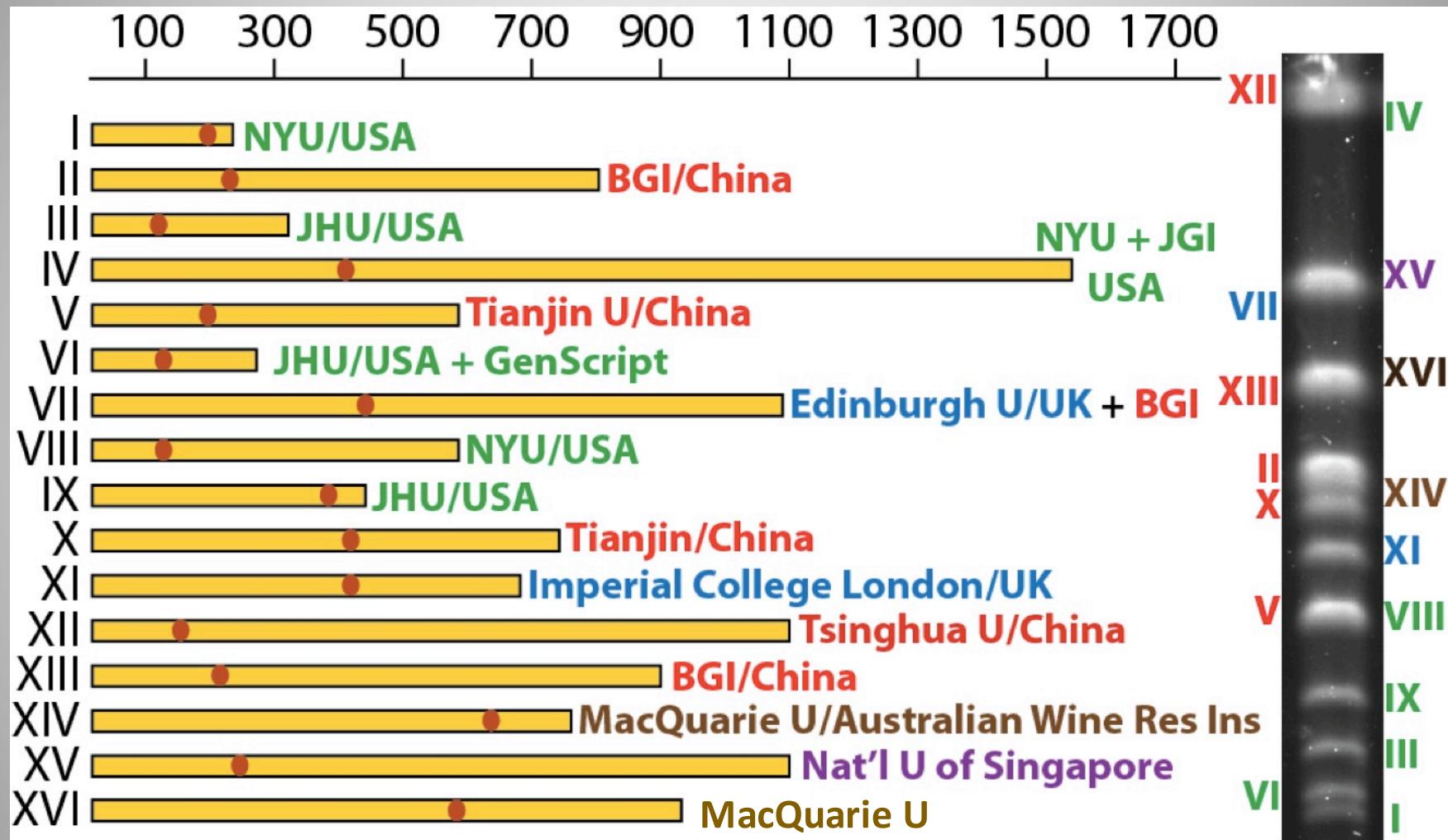
EMBARGOED UNTIL 2:00 PM US ET THURSDAY, 27 MARC

Total Synthesis of a Functional Designer Eukaryotic Chromosome

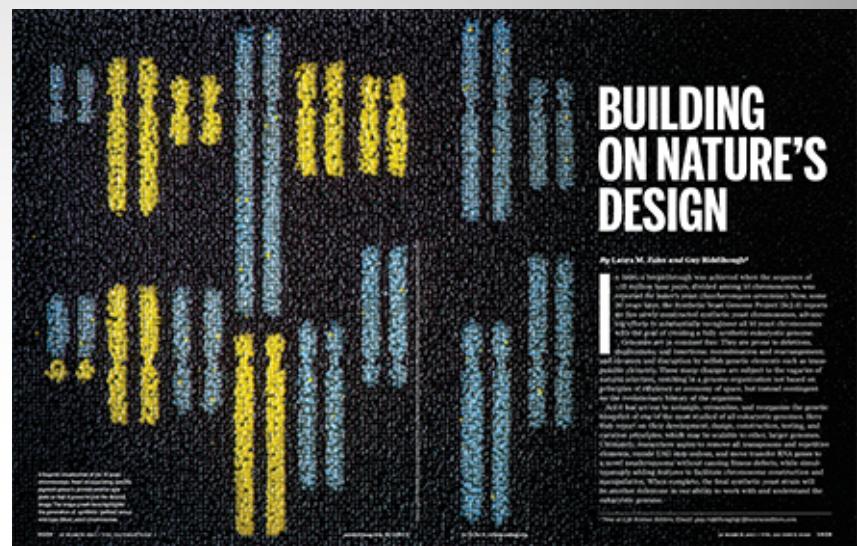
Narayana Annaluru,^{1*} Héloïse Muller,^{1,2,3,4*} Leslie A. Mitchell,² Sivaprakash Ramalingam,¹ Giovanni Stracquadanio,^{2,5} Sarah M. Richardson,⁵ Jessica S. Dymond,^{2,6} Zheng Kuang,² Lisa Z. Scheifele,^{2,7} Eric M. Cooper,² Yizhi Cai,^{2,8} Karen Zeller,² Neta Agmon,² Jeffrey S. Han,⁹ Michalis Hadjithomas,¹⁰ Jennifer Tullman,⁵ Katrina Caravelli,¹ Kimberly Cirelli,¹ Zheyuan Guo,¹ Viktoriya London,¹ Apurva Yeluru,¹ Sindurathy Murugan,⁵ Karthikeyan Kandavelou,^{1,11} Nicolas Agier,^{12,13} Gilles Fischer,^{12,13} Kun Yang,^{2,5} J. Andrew Martin,² Murat Bilgel,¹ Pavlo Bohutski,¹ Kristin M. Boulier,¹ Brian J. Capaldo,¹ Joy Chang,¹ Kristie Charoen,¹ Woo Jin Choi,¹ Peter Deng,¹ James E. DiCarlo,¹ Judy Doong,¹ Jessilyn Dunn,¹ Jason I. Feinberg,¹ Christopher Fernandez,¹ Charlotte E. Floria,¹ David Gladowski,¹ Pasha Hadidi,¹ Isabel Ishizuka,¹ Javaneh Jabbari,¹ Calvin Y. L. Lau,¹ Pablo A. Lee,¹ Sean Li,¹ Denise Lin,¹ Matthias E. Linder,¹ Jonathan Ling,¹ Jaime Liu,¹ Jonathan Liu,¹ Mariya London,¹ Henry Ma,¹ Jessica Mao,¹ Jessica E. McDade,¹ Alexandra McMillan,¹ Aaron M. Moore,¹ Won Chan Oh,¹ Yu Ouyang,¹ Ruchi Patel,¹ Marina Paul,¹ Laura C. Paulsen,¹ Judy Qiu,¹ Alex Rhee,¹ Matthew G. Rubashkin,¹ Ina Y. Soh,¹ Nathaniel E. Sotuyo,¹ Venkatesh Srinivas,¹ Allison Suarez,¹ Andy Wong,¹ Remus Wong,¹ Wei Rose Xie,¹ Yijie Xu,¹ Allen T. Yu,¹ Romain Koszul,^{3,4} Joel S. Bader,^{2,5} Jef D. Boeke,^{2,10,14†} Srinivasan Chandrasegaran^{1†}

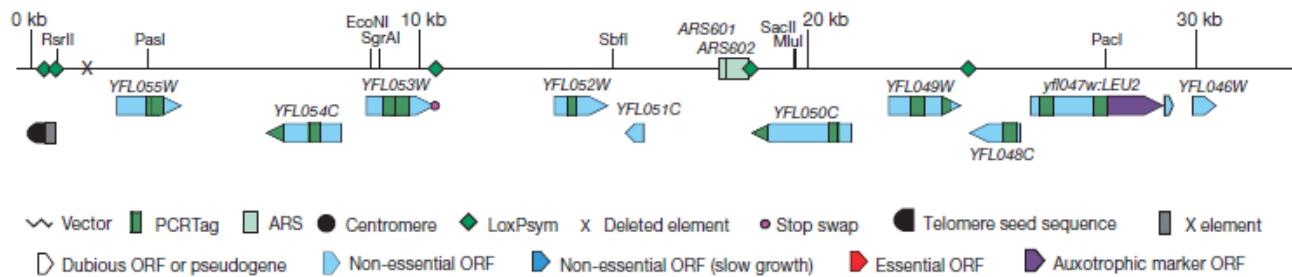


Synthetic Yeast Goes Global



2017: Completed 5 more Syn Chromosomes





BOX I

Modifications in synthetic sequence

Elements removed

Retrotransposons: The *S. cerevisiae* genome contains both active retrotransposons and retrotransposon-derived sequences. These highly repetitive sequences are known to contribute to genome instability²². Because retrotransposons are presumed to be nonessential in yeast, we are eliminating these sequences from the synthetic genome.

Subtelomeric repeats: Two major types of subtelomeric repeats, Y' and X elements, reside in the genome. Y' elements are of unknown function, and are present at some, but not all, *S. cerevisiae* chromosome ends²³. In contrast, X elements are present in a single copy at all *S. cerevisiae* chromosome ends; they are more highly divergent, and function in telomeric silencing and possibly in chromosome segregation²³. To create a more streamlined genome, all Y' elements will be deleted from the synthetic genome; extant X elements will be replaced with the consensus core X-element sequence, as in semi-synVIL.

Introns: The yeast genome is estimated to contain approximately 285 introns. Based on a previous intron-deletion study²⁴ we do not anticipate that removal of introns will result in fitness defects; however, in some cases these introns house small non-coding RNAs (snoRNAs) that can be expressed ectopically in the synthetic yeast.

Elements relocated to extrachromosomal array

tRNA genes: tRNA genes (tDNAs) are highly redundant, with 275 nuclear tDNAs encoding only 42 tRNA species²⁵. In addition, these genes are known regions of genome instability^{8,9}. They will therefore be relocated to a dedicated chromosome to contain any instability resulting from their presence.

Elements replaced

TAG stop codons replaced by TAA: Removal of the TAG stop codon from the synthetic genome will allow future genetic code manipulation. The 'free' codon may be used to incorporate artificial amino acids^{11,12}; alternatively, the TAG codon may be placed in essential genes, and, exploiting an engineered orthogonal synthetase/tRNA pair, specify a non-genetically encoded amino acid, thereby providing a mechanism of reproductive isolation and an additional level of control over the synthetic yeast.

Individual synonymous codons: The synthetic genome is fabricated in fragments as small as 750 bp²⁶. Unique restriction sites are necessary within the synthetic fragment to facilitate construction of these building blocks into large contigs of up to 100 kb. Short stretches of fewer than four codons may therefore be synonymously recoded to introduce or eliminate restriction sites.

Strings of synonymous codons: Although several modifications exist between the native and synthetic genomes, the presence of a dedicated mechanism to distinguish between the two sequence types is invaluable. Short stretches of fewer than ten codons are therefore recoded to generate 'PCRTags', synonymous sequences used as the basis for PCR primer design to amplify selectively from wild-type or synthetic genomes.

Elements introduced

LoxPsym sites: Symmetrical loxP sites¹³ are inserted in the 3' UTR of all non-essential genes, as well as at synthetic landmarks. LoxPsym sites lack the directionality of canonical loxP sites, and can therefore align in two orientations. As a result, both inversions and deletions are predicted at equal probability. These loxPsym sites and an inducible Cre recombinase¹⁵ form the basis of the SCRaMbLE toolkit.

Elements not changed

Gene order: Gene order is preserved in the synthetic yeast to prevent incorporation of a non-permissible configuration in the design phase. Induction of SCRaMbLE results in changes in gene order and chromosome structure; all recovered SCRaMbLEd yeast have viable genome structures.

Noncoding regions: Except where noted, noncoding regions have not been modified. The yeast genome is well annotated; however, it is of paramount importance that the synthetic yeast be as fit as wild type until SCRaMbLE is induced. We therefore eschewed changes of noncoding regions to avoid disrupting unannotated critical elements. The few modifications that are made in noncoding sequence are kept to a minimum.

Genome Synthesis: what changes to make?

Venter – added watermark sequences at intergenic sites using a cryptic code
scientists' names, famous quotes, email address (rest stays the same)

SC2.0 – many more changes...

1. Remove unwanted elements

Retrotransposons, subtelomeric repeats, introns (285 in yeast)

2. Relocate essential elements

Move tRNA genes to a dedicated chromosome

3. Introduce new elements

Symmetrical loxP sites inserted in the 3'UTR of all non-essential genes, and at synthetic landmarks. This generates the SCRaMbLE toolkit

4. Recode existing elements

With DNA synthesis it is possible to 'silently' change protein coding sequence by using synonymous codons, so you can (i) incorporate unique sequence tags for PCR checking (ii) add or remove some restriction sites and (iii) **change all TAG stop codons to TAA codons**

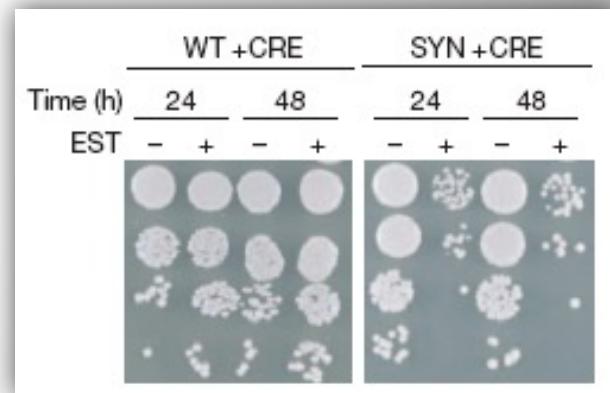
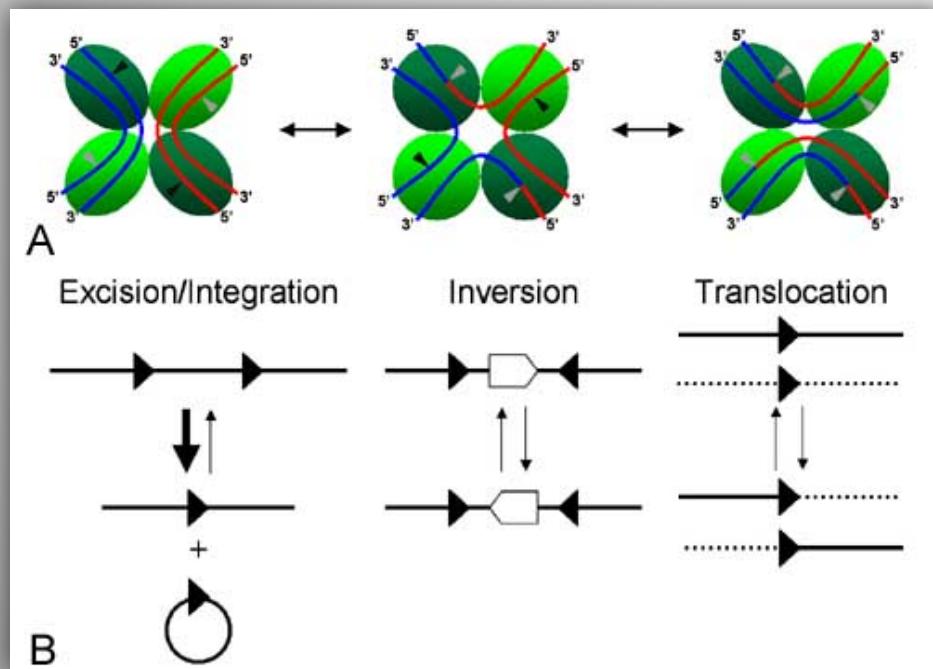
SCRaMbLE toolkit: Automated genome minimisation

SC2.0

Symmetrical loxP sites inserted in the 3'UTR of all non-essential genes, and at synthetic landmarks. This generates the SCRaMbLE toolkit.

LoxPsym sites are cut and moved around by Cre recombinase

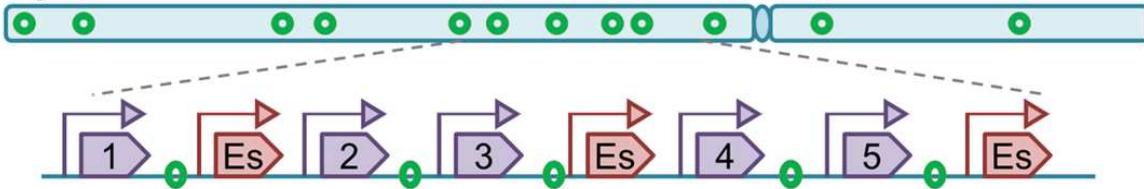
SC2.0 has inducible Cre expression. Add oestradiol = whole genome shuffle



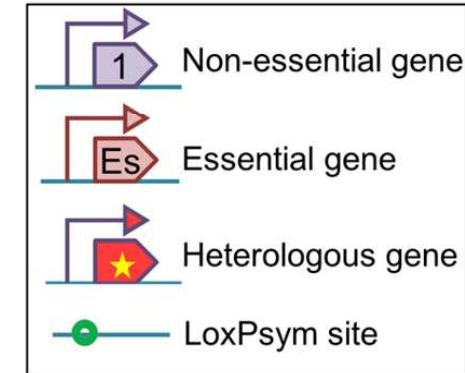
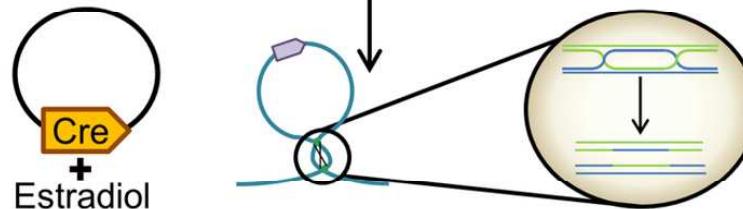
Automatic refactoring of genome – un-needed genes will be lost

SCRaMbLEing the yeast genome

SynIII



Transform Cre plasmid into semi-synthetic strain & induce with estradiol



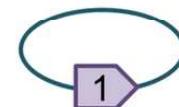
Homologous recombination between loxPsym sites catalyses rearrangements

Translocation

Inversion

Deletion

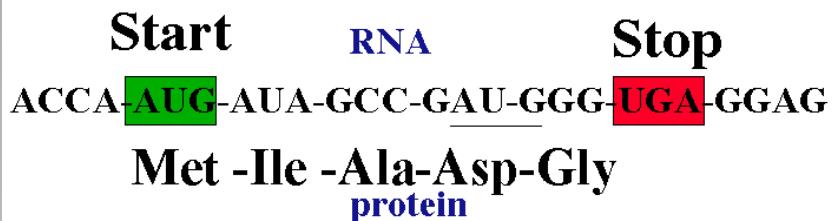
Insertion



Scrambled
SynIII



Swap all TAG codons to TAA: Freeing-up a codon



The start codon is AUG and it also codes for Methionine

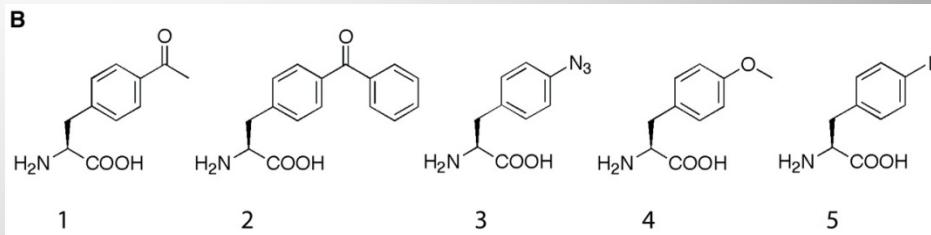
There are three stop codons UGA, UAA, UAG

The rarest codon in most genomes is the TAG stop codon

Swapping all of these for TAA codon will ‘free-up’ an unused codon

Previous work by Peter Schultz's Lab has shown that unnatural amino acids can be programmed into yeast proteins by hijacking the TAG codon.

An Expanded Eukaryotic Genetic Code
Jason W. Chin et al. *Science* 2003



Codon Idea taken from a 3rd Project - RE.Coli

Genome Engineering of *E. coli*

George Church Lab – Harvard

Remove all TAG stop codons on the natural genome (326)

Free-up further codons in the future

NOT using DNA synthesis

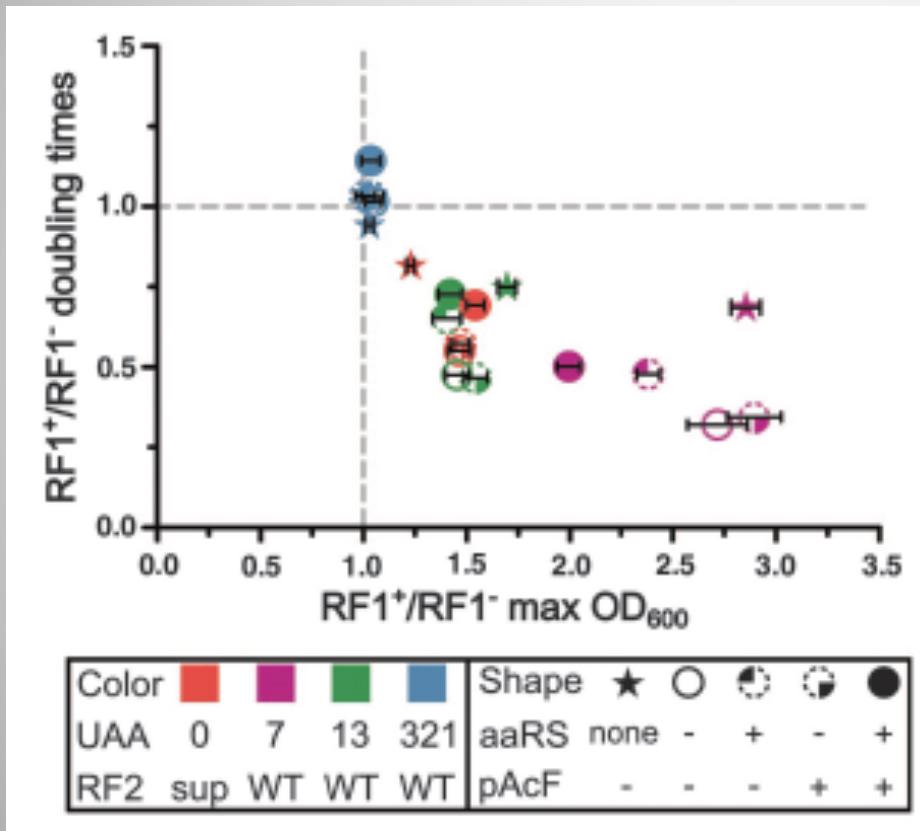
Done by careful mutation

TTT	30362	TCT	11495	TAT	21999	TGT	7048
TTC	F 22516	TCC	S 11720	TAC	Y 16601	TGC	C 8816
TTA	4 18932	TCA	S 9783	TAA	STOP 2703	TGA	STOP 1256
TTG	L 18602	TCG	12166	TAG	STOP 1 326	TGG	W 20683
CTT	15002	CCT	9559	CAT	17613	CGT	28382
CTC	15077	CCC	7485	CAC	H 13227	CGC	29898
CTA	5314	CCA	P 11471	CAA	20888	CGA	R 4859
CTG	L 71553	CCG	31515	CAG	Q 39188	CGG	7399
ATT	41309	ACT	12198	AAT	24159	AGT	11970
ATC	I 34178	ACC	T 31796	AAC	N 29385	AGC	S 3 21862
ATA	5967	ACA	T 9670	AAA	45687	AGA	2 2896
ATG	M 37915	ACG	19624	AAG	K 14029	AGG	R 2 1692
GTT	24858	GCT	20762	GAT	43719	GGT	33622
GTC	V 20753	GCC	34695	GAC	D 25918	GGC	40285
GTA	V 14822	GCA	A 27418	GAA	53641	GGA	G 10893
GTG	35918	GCG	45741	GAG	E 24254	GGG	15090

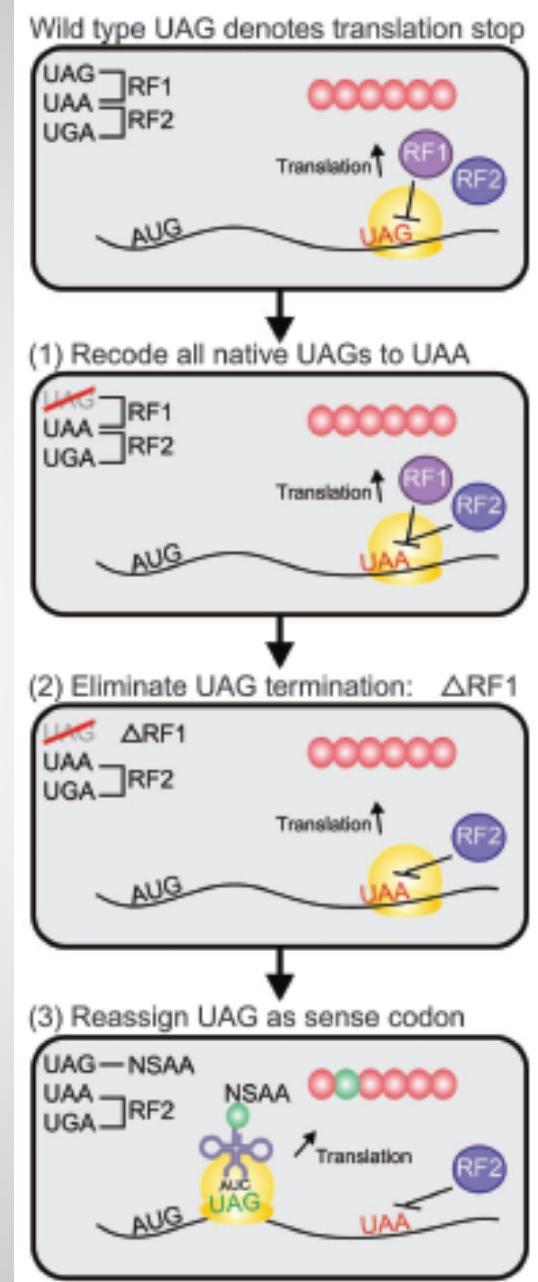
Total number of codons in *E. coli* genome

RE.Coli completed in late 2013!

Genomically Recoded Organisms Expand Biological Functions *Science* 342, 357 (2013); Marc J. Lajoie *et al.*



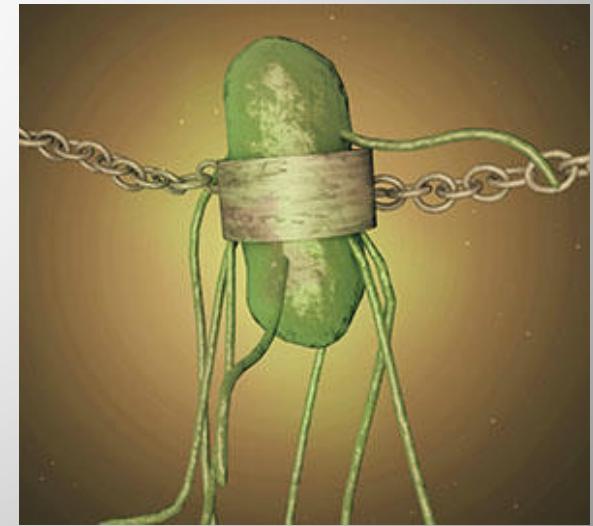
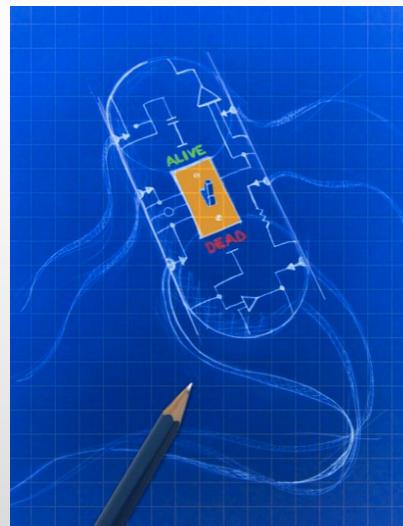
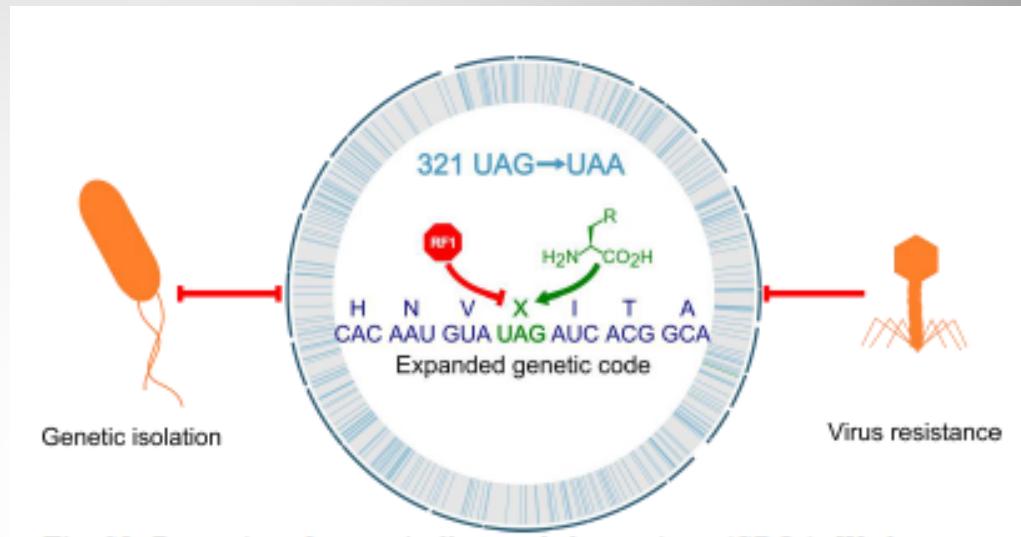
Growth in NSAA media is healthy – orthogonal!
(NSAA = non-standard amino acid)



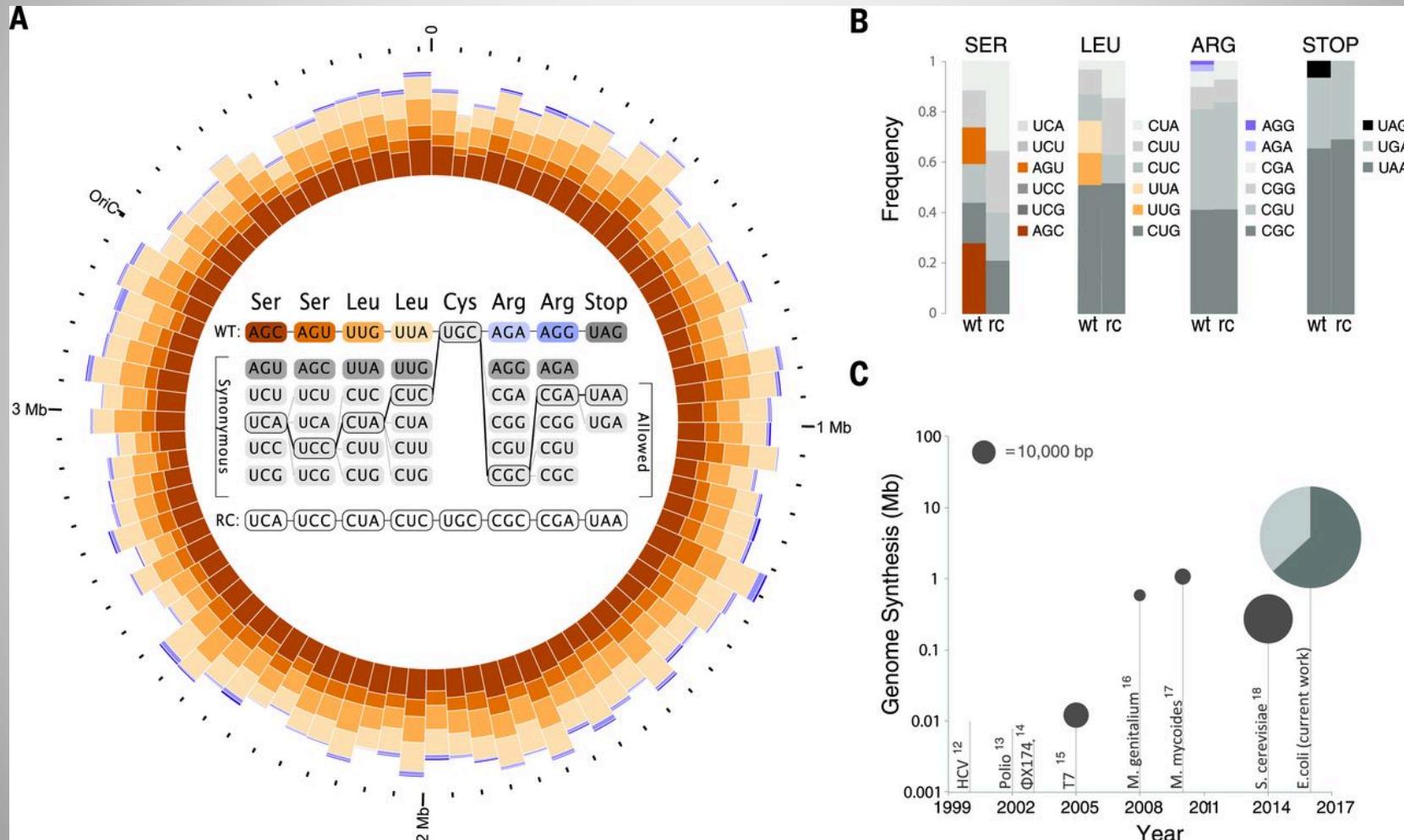
RE.Coli is orthogonal biology: ‘xenobiology’?

RE.coli GRO
(321 codon changes)

- Allows precise non-standard amino acid incorporation into proteins
 - e.g. novel enzymes
- More resistant to phage attack due to orthogonality
 - Phage genome is misread
- Can be made to be entirely dependent on non-standard amino acids for growth
 - In-built biosecurity to halt growth in the wild



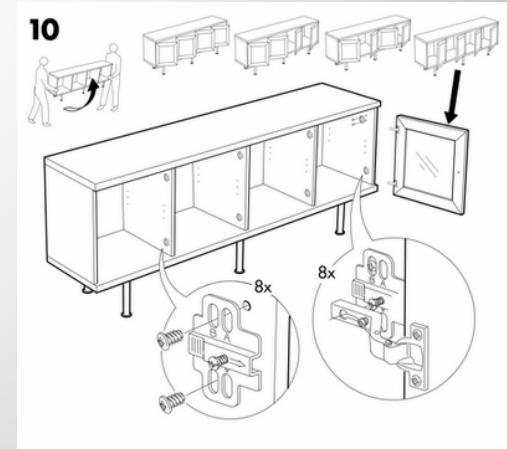
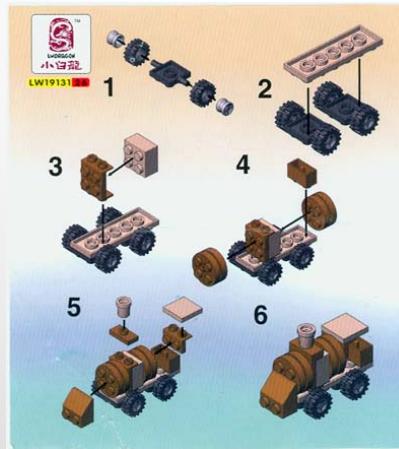
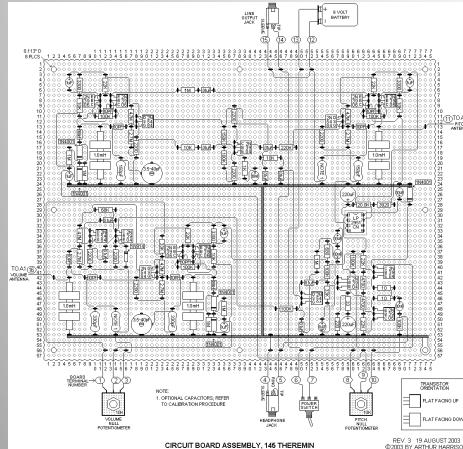
RE.Coli version 2: 57 codons



Why do genome synthesis?

I. Bottom-up synthetic biology

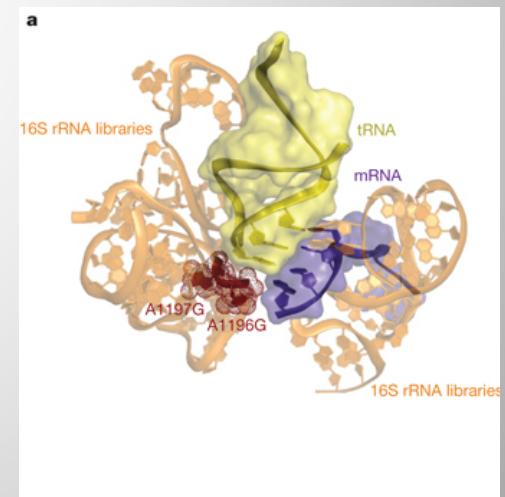
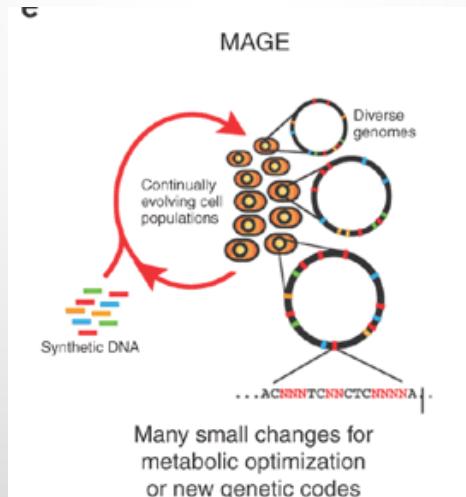
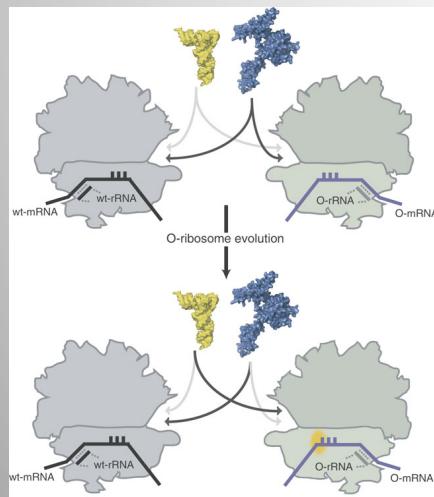
- Adding genes and devices should be more predictable
- Creating a whole-cell model should be easier and allow better predictions of behaviour – key to future synthetic biology
- Provides a route to designing the chassis cell fit for a specific application
- Removal of unstable / recombination elements



Why do genome synthesis?

2. Provides for safer synthetic biology

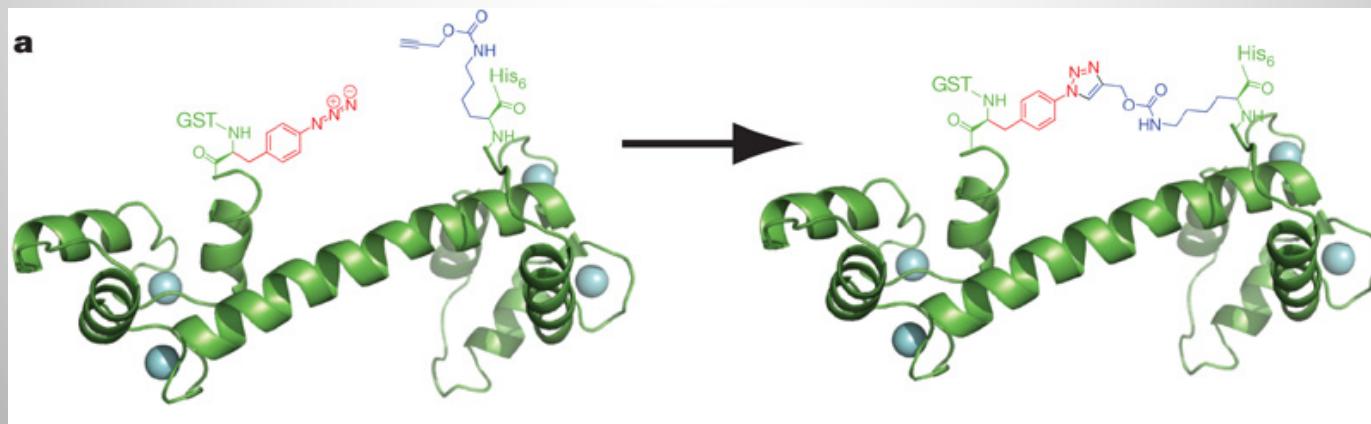
- Cell can be designed to only survive in lab conditions
- Cell could be made orthogonal so that its biology doesn't interact with nature – a.k.a *xenobiology*
 - examples: change codon usage or change stereochemistry
- Better predictability from bottom-up design



Why do genome synthesis?

3. Custom synthesis of products

- Cells could be designed to produce non-natural proteins and sugars using synthetic building blocks
- Minimal cells would only use resources to make the desired products and so be more efficient
- Very cheap production of DNA could be engineered
- Synthesis of molecules that are toxic to produce in normal cells
- Recoded cells may not be attacked by viruses and phages



What you should now know (and read up on)

You could get exam questions on...

1. How JCVI made the first cell with a synthetic genome
2. The DNA assembly methods for assembling a genome
3. How to model a whole cell and why
4. How to make Synthetic Yeast chromosomes
5. Re-factoring genomes rationally vs. automatically
6. The uses of a recoded *E. coli* genome
7. Applications for modelled cells and engineered genomes

Reading – 4 Key Papers

Complete Chemical Synthesis, Assembly, and Cloning of a *Mycoplasma genitalium* Genome

- DG Gibson et al. Science Vol. 319 no. 5867 (29 February 2008), pp. 1215-1220.

Synthetic chromosome arms function in yeast and generate phenotypic diversity by design

- JS Dymond et al. Nature 477 (22 September 2011) pp. 471–476

ALSO take a look at the Supplementary Materials for this paper

And also there is: <http://syntheticyeast.org/>

Total synthesis of a functional designer eukaryotic chromosome

- N Annaluru et al. Science Vol. 344 no. 6179 (4 April 2014), pp. 55-58

A Whole-Cell Computational Model Predicts Phenotype from Genotype

- JR Karr et al. Cell Vol. 150 (20 July 2012) pp. 389-401

<http://www.youtube.com/watch?v=au4sl9CjKFU>

Reading – Further Research Papers

RE. COLI:

Genomically Recoded Organisms Expand Biological Functions

Science **342**, 357 (2013); Marc J. Lajoie et al.

Biocontainment of genetically modified organisms by synthetic protein design

Nature **518**, pp. 55-60 (2014); Daniel J. Mandell et al

METHODS FOR GENOME ASSEMBLY AND TRANSPLANTATION BY VENTER GROUP

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

- DG Gibson et al. *Science* Vol. 329 no. 5987 (2 July 2010), pp. 52-56

Enzymatic assembly of DNA molecules up to several hundred kilobases

- DG Gibson et al. *Nature Methods* 6, 343 - 345 (2009)

Genome Transplantation in Bacteria: Changing One Species to Another

- C Lartigue et al. *Science* Vol. 317 no. 5838 (3 August 2007), pp. 632-638