Aalto University
School of Chemical
Technology

Synthetic biology

Course CHEM-E8125, spring 2019

Synthetic Yeast 2.0

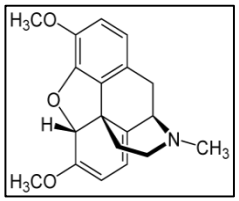
How to design and construct a complete synthetic
”*Saccharomyces cerevisiae*” yeast

Prof. Merja Penttilä

CONTENT

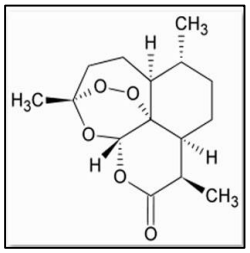
- Why yeast
- Yeast molecular biology tools
- Design and construction of the synthetic yeast
- Group work

Why baker's yeast *Saccharomyces cerevisiae* ?



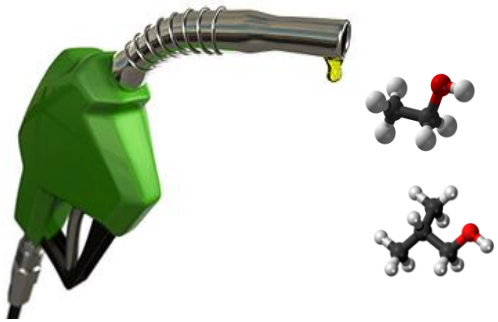
Opiate thebaine

- A companion of man over millennia
- A reliable workhorse



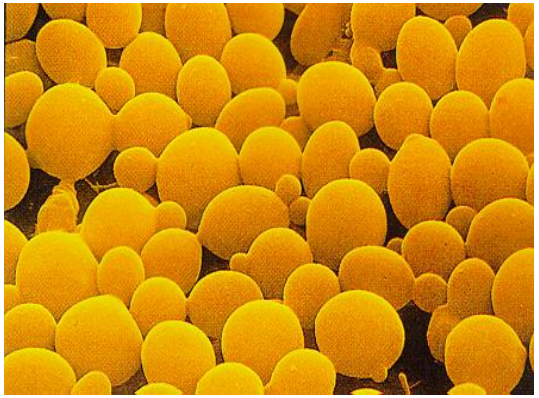
Malaria drug artemisinin

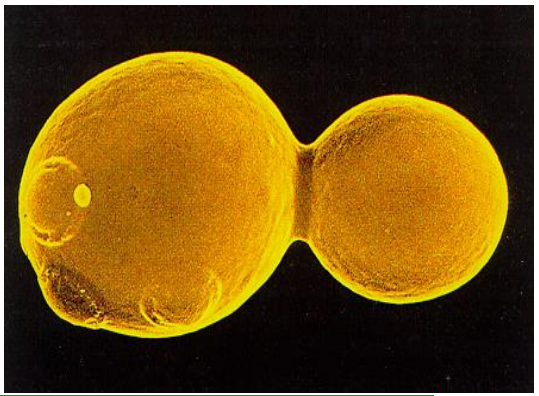
Human insulin



Bioreactors for ethanol production can be over 1000 m^3 (10^6 litres) large

Saccharomyces cerevisiae is a robust process organism



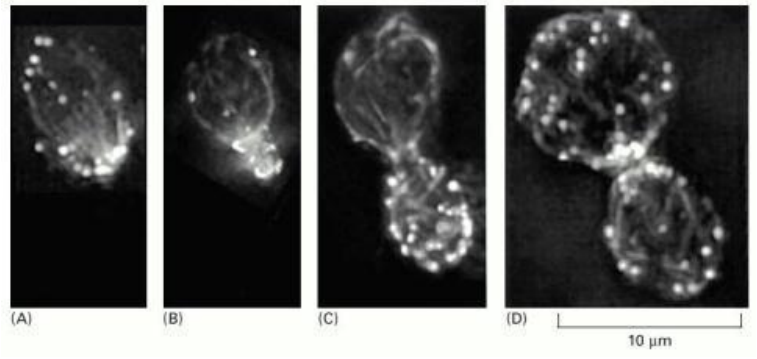


Over 10^8 small cell factories fit in one litre

4 - 5 micrometers

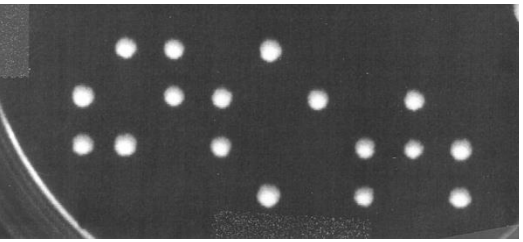
Saccharomyces cerevisiae is a eukaryotic model organism

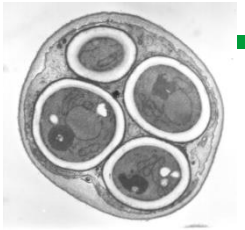
Actin patches and filaments



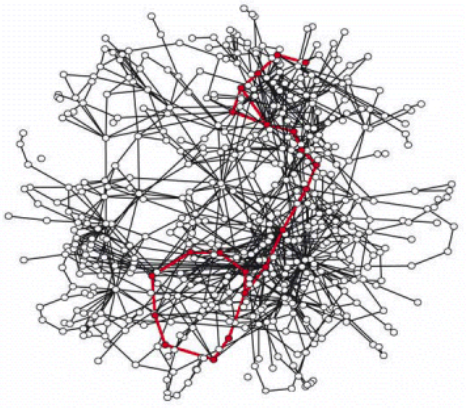
For

- **physiology (anaerobic, aerobic)**
- **genetics (haploid, diploid, polyploid)**
- **cell biology (Nobel prize 2001 to Hartwell for cell cycle regulation)**
- **molecular biology (plasmid transformation (1978) and targeted gene deletion 1983))**
- **genomics (first eukaryotic genome sequenced 1996)**
- **systems biology (-omics methods, metabolic model 2003)**
- **synthetic biology**





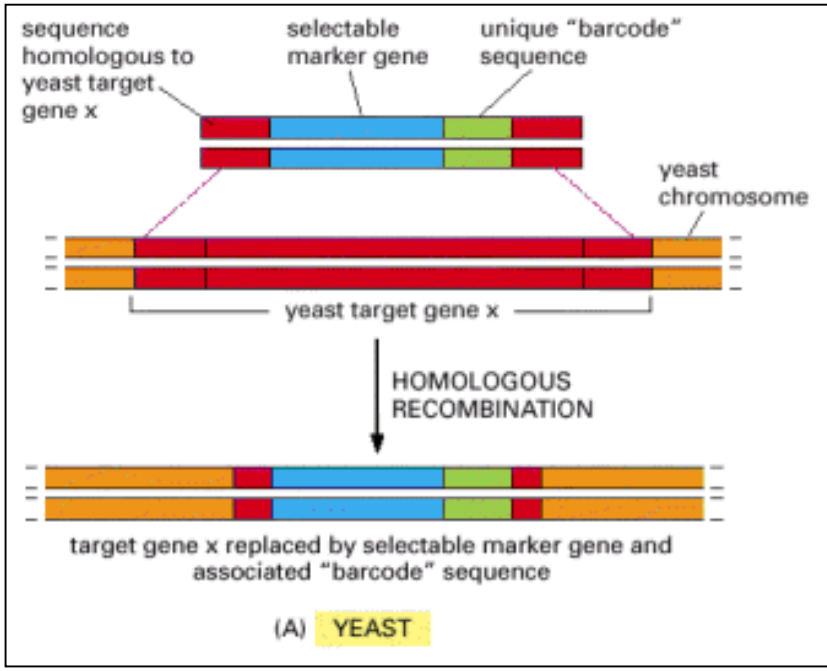
Four spores
In an ascus



Saccharomyces cerevisiae – Molecular tools

- High transformation frequencies, >> 10^6 /ug DNA
- Versatility of vectors (high/lower copy number plasmids, integrative vectors, single copy telomere vectors, YAC (yeast artificial chromosome)
- High frequency of homologous recombination
- Gene deletion libraries available (in both haploid mating types -> heterozygous diploids)
- Tagged versions of genes for localisation studies
- Two-hybrid libraries for protein-protein interaction studies
- Promoter arrays for e.g. chromatin immuno-precipitation experiments

The deletion libraries are barcoded, 20bp

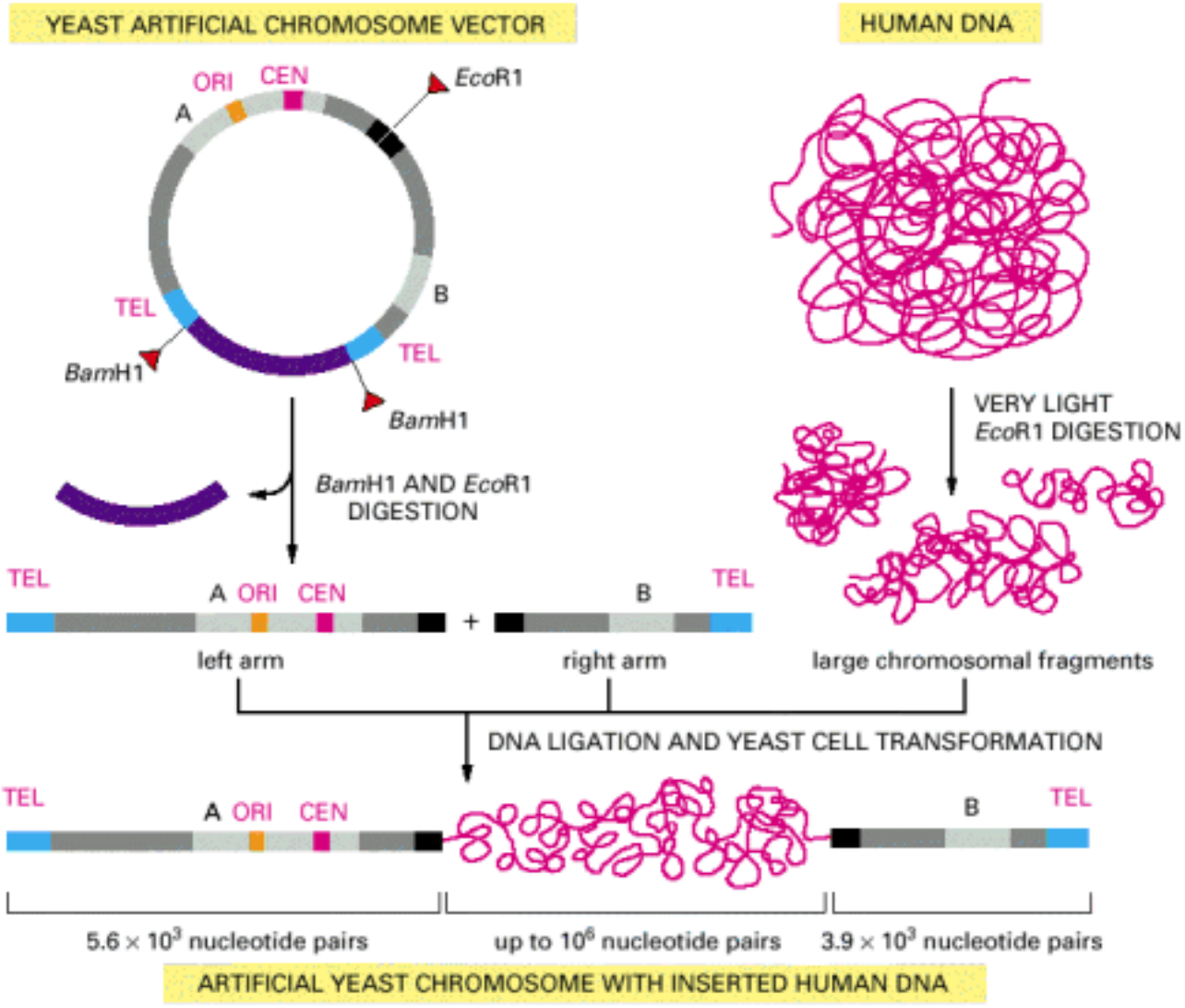


- High through put screens & methods (e.g. disease genes; 70% of mammalian genes with similarity to yeast are functional in yeast)
- Various data bases

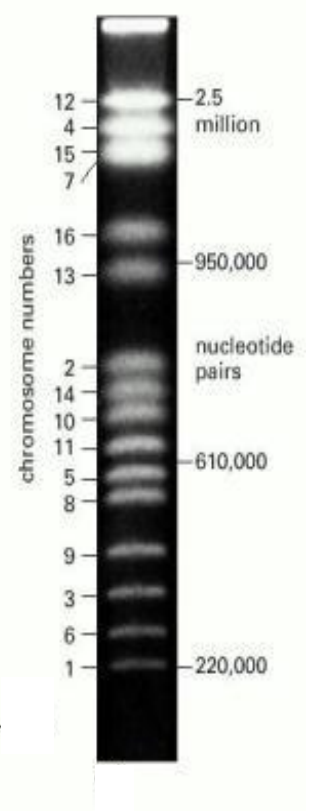
Saccharomyces cerevisiae – artificial chromosomes

- ORI; origin of replication
- CEN; centromere
- TEL; telomere

- Large pieces of DNA can be inserted (>> 50 genes)
- One way to build synthetic genomes
- *Mycoplasma genitalium* genome (0.6Mb) was assembled in yeast



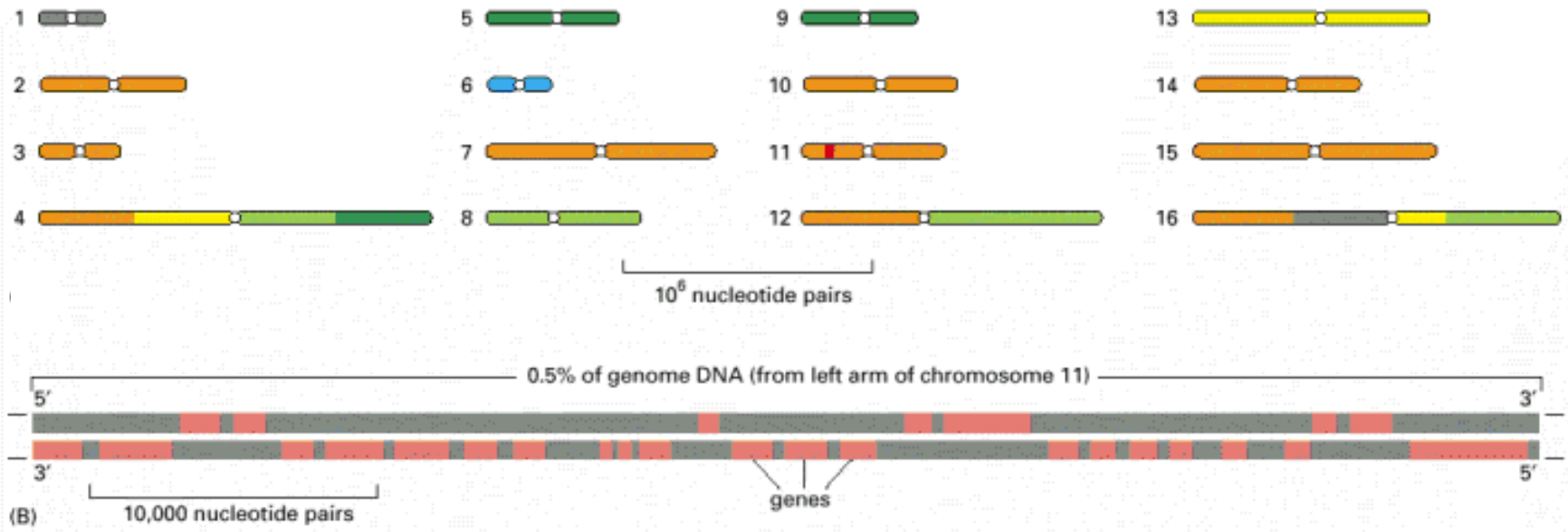
One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. Gibson DG et al. . Proc Natl Acad Sci U S A. 2008 Dec 23;105(51):20404-9



S.cerevisiae chromosomes on a CHEF gel

The *Saccharomyces cerevisiae* genome

- 13.4 Mb in 16 linear chromosomes + mitochondrial DNA
- About 6000 genes
- Compact genome
- Very few introns, 5% of genome (actin gene intron ~500bp, large)
- ~20% ?? of genes without function still



Sequencing of the *Saccharomyces cerevisiae* genome was a global effort

The nucleotide sequence of chromosome I from *Saccharomyces cerevisiae*. H. Bussey, *et al*., 1995, *PNAS USA* 92:3809-13

Complete DNA sequence of yeast chromosome II. H. Feldmann, *et al*., 1994, *EMBO J* 13(24):5795-809

The complete DNA sequence of yeast chromosome III. S. G. Oliver, *et al*., 1992, *Nature* 357:38-46

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IV. C. Jacq, *et al*. (1997) *Nature* 387(6632 Suppl):75-78

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V. F. S. Dietrich, *et al.* (1997) *Nature* 387(6632 Suppl):78-81

Analysis of the nucleotide sequence of chromosome VI from *Saccharomyces cerevisiae*. Y. Murakami, *et al*., 1995, *Nature Genetics* 10:261-8

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome VII. H. Tettelin, *et al*. (1997) *Nature* 387(6632 Suppl):81-84

Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome VIII. M. Johnston, *et al.*, 1994, *Science* 265:2077-82

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IX. Churcher, *et al.* (1997) *Nature* 387(6632 Suppl):84-87

Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome X. F. Galibert, *et al*., 1996, *EMBO J* 15(9):2031-49

Complete DNA sequence of yeast chromosome XI. B. Dujon, *et al.*, 1994, Nature 369:371-8

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XII. M Johnston, *et al*. (1997) *Nature* 387(6632 Suppl):87-90

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIII. S. Bowman, *et al*. (1997) *Nature* 387(6632 Suppl):90-93

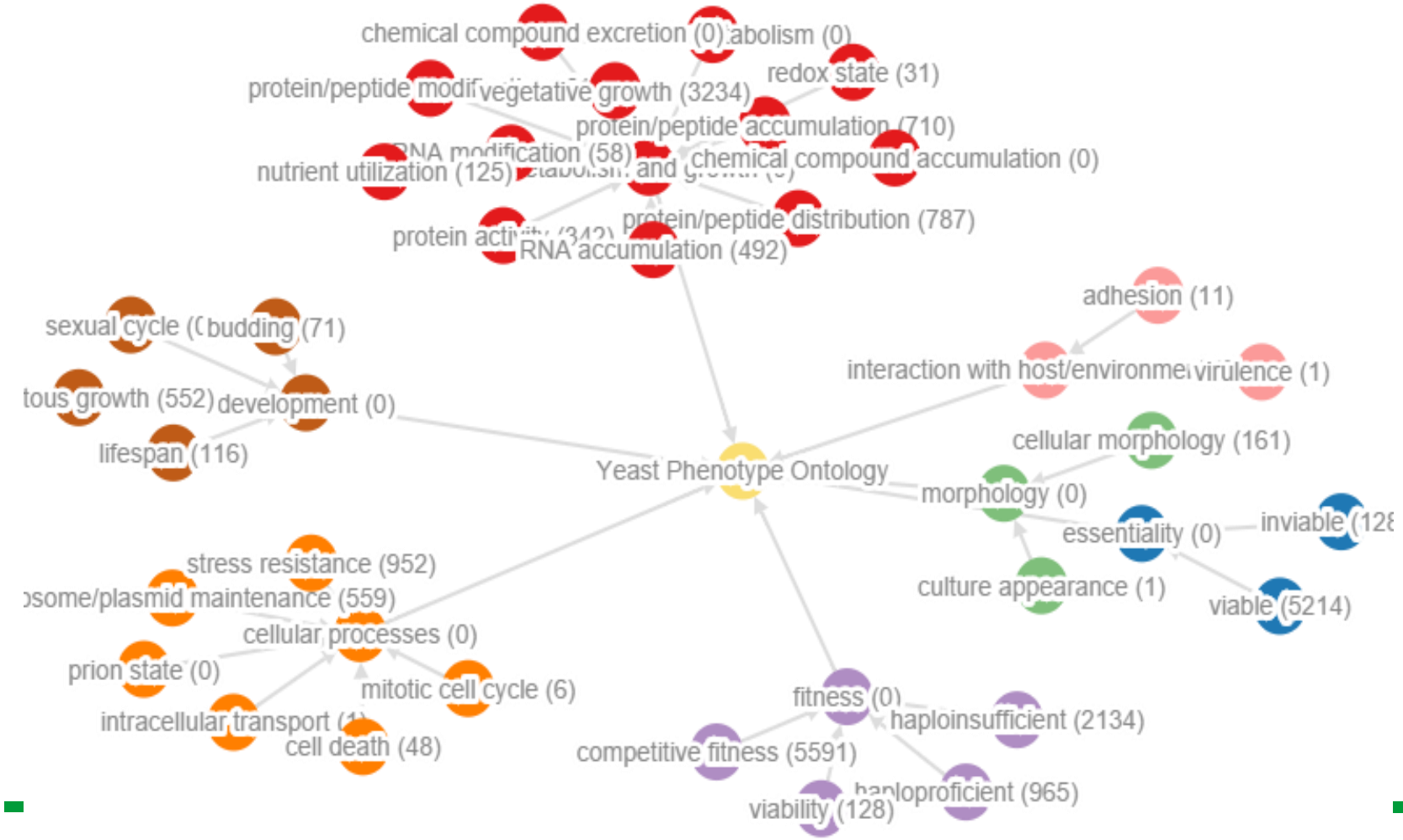
The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIV and its evolutionary implications. P. Philippsen, *et al.* (1997) *Nature* 387(6632 Suppl):93-98

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XV and its evolutionary implications. B. Dujon, *et al.* (1997) *Nature* 387(6632 Suppl):98-102

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XVI and its evolutionary implications. H. Bussey, *et al.* (1997) *Nature* 387(6632 Suppl):103-105

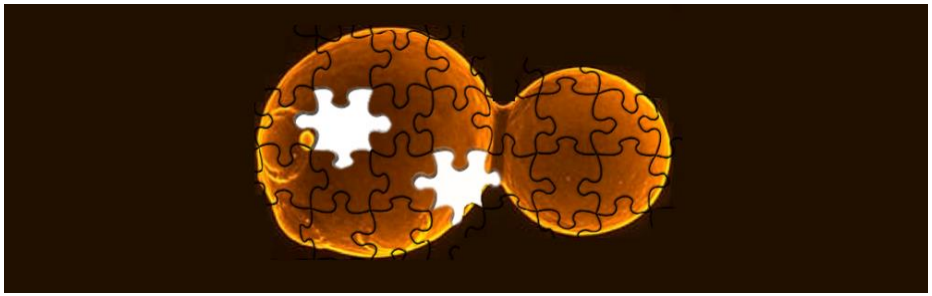
The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae*. F. Foury, *et al.*, 1998, *FEBS Letters* 440:325-31

Saccharomyces genome database (SGD) contains gene, genome and phenotype information



Current Observable Other Observable

SGD 2015-09-08



Yeast 2.0
"Designer genome"

Jeff Boeke et al, NYU Medical Centre

1. Design principles

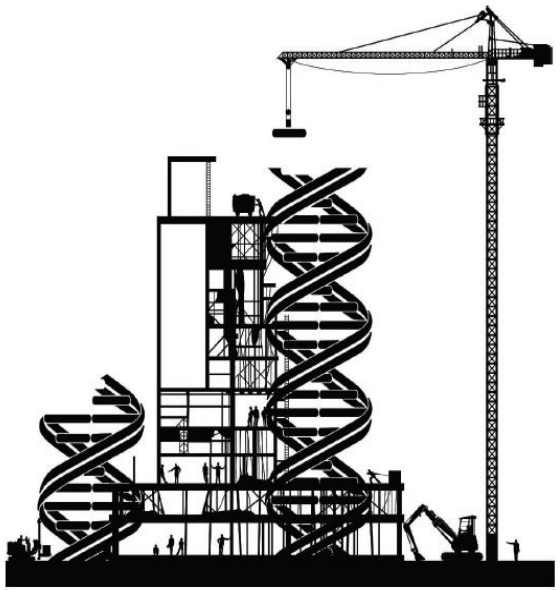
- keep yeastiness
- maintain fitness
- maintain genomic stability
- but increase genetic flexibility; make possible genetic evolution through SCRaMbLE

Questions about species identity and taxonomy

2. Construction principles

- rather conservative
- chromosome by chromosome

A huge construction effort



Building the Ultimate Yeast Genome

(Pennisi 2014)

Design and construction principles of Yeast 2.0

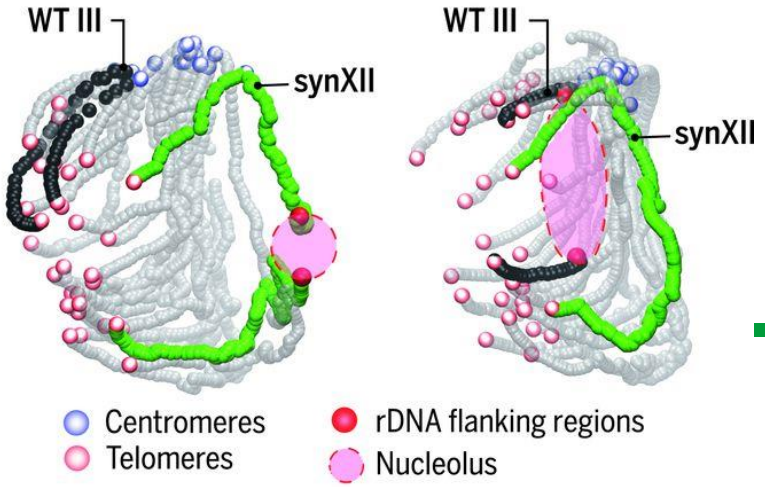
- Rather conservative changes and minimal recoding (= change of DNA sequence) (e.g. due to unknown codon usage or regulatory effects of DNA sequence)
- The synthetic DNA is introduced to the wild-type genome step-wise in 30-60kb pieces replacing the native yeast sequence in the chromosome, “bottom up” approach. This makes it easy to trace back if deleterious changes were introduced.
- Yeast fitness is evaluated in laboratory conditions after each cycle. The cumulative effect of unwanted changes could yield an unviable yeast cell within a relatively low number of cycles, unless monitored.
- Following the success of construction of the first synthetic yeast chromosome, *synIII.*

  N. Annaluru, J. Boeke et al., “Total synthesis of a functional designer eukaryotic chromosome,” *Science*, doi:10.1126/science.1249252, 2014.

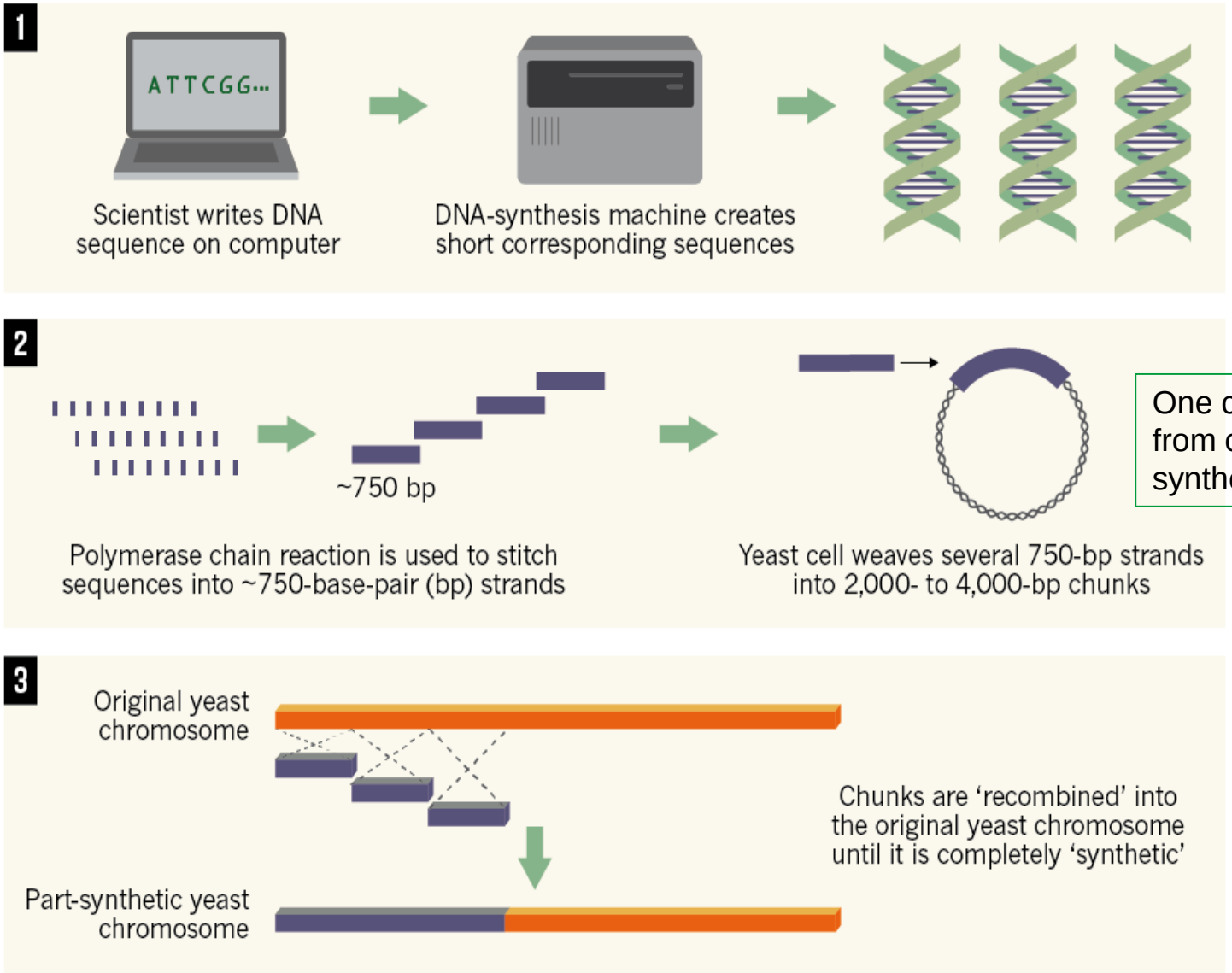
SCIENCE, 10 MARCH 2017, VOL 355, ISSUE 6329

- *synII, synV, synVI, synX, and synXII*
- *30% of yeast DNA synthetic*

1. Design of a synthetic yeast genome
2. Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome
3. “Perfect” designer chromosome V and behavior of a ring derivative
4. Synthesis, debugging, and effects of synthetic chromosome consolidation: synVI and beyond
5. Bug mapping and fitness testing of chemically synthesized chromosome X
6. Engineering the ribosomal DNA in a megabase synthetic chromosome
7. 3D organization of synthetic and scrambled chromosomes



Design and construction principles of Yeast 2.0



One can also order from companies ~10kb synthetic DNA pieces

Design and construction principles of Yeast 2.0

Hierarchial assembly plan

- Building blocks (~750bp) are assembled into minichunks (~3kb), which are assembled into chunks (~10kb), which are subsequently assembled into megachunks (30-50kb). The termini of each 10kb chunk encode restriction enzyme sites that enable directional assembly of chunks into megachunks by *in vitro* ligation.
- The megachunks of ~30-60kb (or 3-6 x 10kb chunks) of synthetic DNA are integrated into the yeast genome at a time to replace the corresponding native chromosomal DNA, exploiting homologous recombination that occurs in yeast.
- Current DNA synthesis technologies can readily produce synthetic DNA fragments in the minichunk and chunk range, and DNA synthesis can be done at companies.

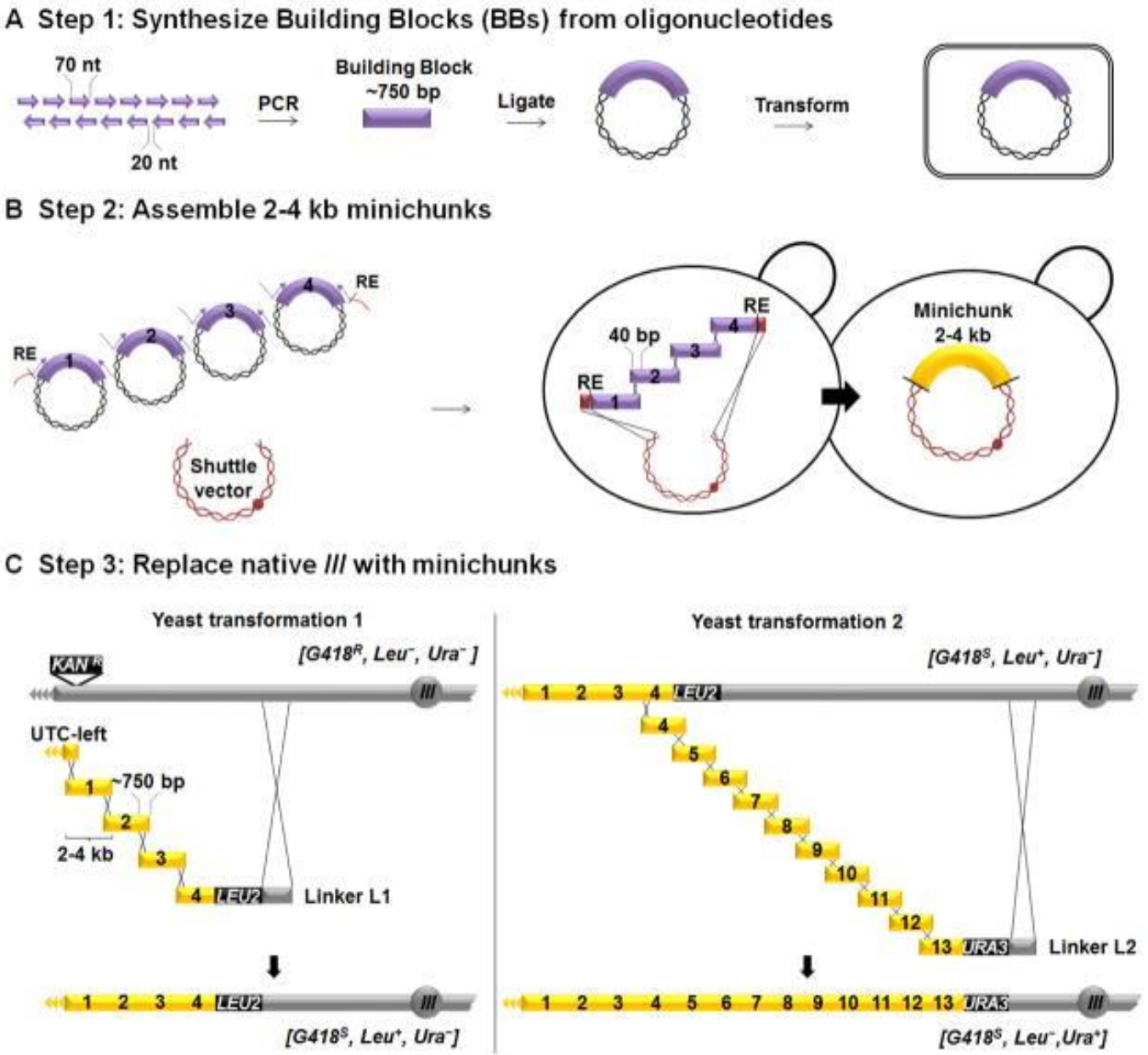
Selection of correct integrations in yeast

- The integration of the leftmost megachunk overwrites a kanMX cassette previously introduced into the native chromosome sequence for negative selection purposes.
- As each subsequent megachunk is incorporated, the auxotrophic marker used in the prior round is eliminated by recombination, and selection is imposed for a second selectable marker (e.g. *LEU2*). In this way, the remainder of the synthetic chromosome can be assembled*in vivo*, with alternating selection for *URA3* and *LEU2* markers.
- In addition to monitoring the loss and gain of auxotrophic markers, the integration of synthetic DNA and loss of native sequence is confirmed with PCRtags.

See Figure next page.

Building synthetic yeast chromosome III (*synIII*)

Fig. 1 *synIII* construction (**A**) Building block (BB) synthesis. 750 bp BBs (purple) were synthesized from oligonucleotides at the Build-AGenome course. (**B**) Assembly of minichunks. 2–4 kb minichunks (yellow) were assembled by homologous recombination in *S. cerevisiae*. Adjacent minichunks were designed to encode overlap of one BB to facilitate downstream assembly steps. Minichunks were flanked by a rare cutting restriction enzyme (RE) site, *Xma*I or *Not*I. (**C**) Direct replacement of native yeast chromosome *III* with pools of synthetic minichunks. Eleven iterative one-step assemblies and replacements of native genomic segments of yeast chromosome *III* were carried out using pools of overlapping synthetic DNA minichunks, encoding alternating genetic markers (*LEU2* or *URA3*), which enabled complete replacement of native *III* with *synIII* in yeast.



Alterations to the Yeast 2.0 genome

PCRTags are incorporated into most open reading frames (ORFs) by recoding a ~20bp segments of the coding regions to a different DNA sequence but encoding the same amino acid sequence. This enables later a quick scan to ascertain that the synthetic substitution of the genome has occurred, and to monitor any changes to the genome later, e.g. after SCRaMbLE (see below).

Synthetic telomere sequences of a simple sequence repeat $(TG_{1\text{-}3})_n$, to reduce chromosome length. These universal telomere caps (UTCs) will be introduced to all chromosome termini.

Removal of introns. There are only about 250-300 introns in the native *S. cerevisiae* genome, and those likely to have no significant impact will be removed. If e.g. small RNA coding regions are present these will be relocated to non-intronic regions.

Removal of non-essential genes for viability/fitness in laboratory conditions, such as mating type genes.

Replacing all TAG stop codons with TAA, allowing for subsequent introduction of new coding schemes with unnatural amino acids (using TAG).

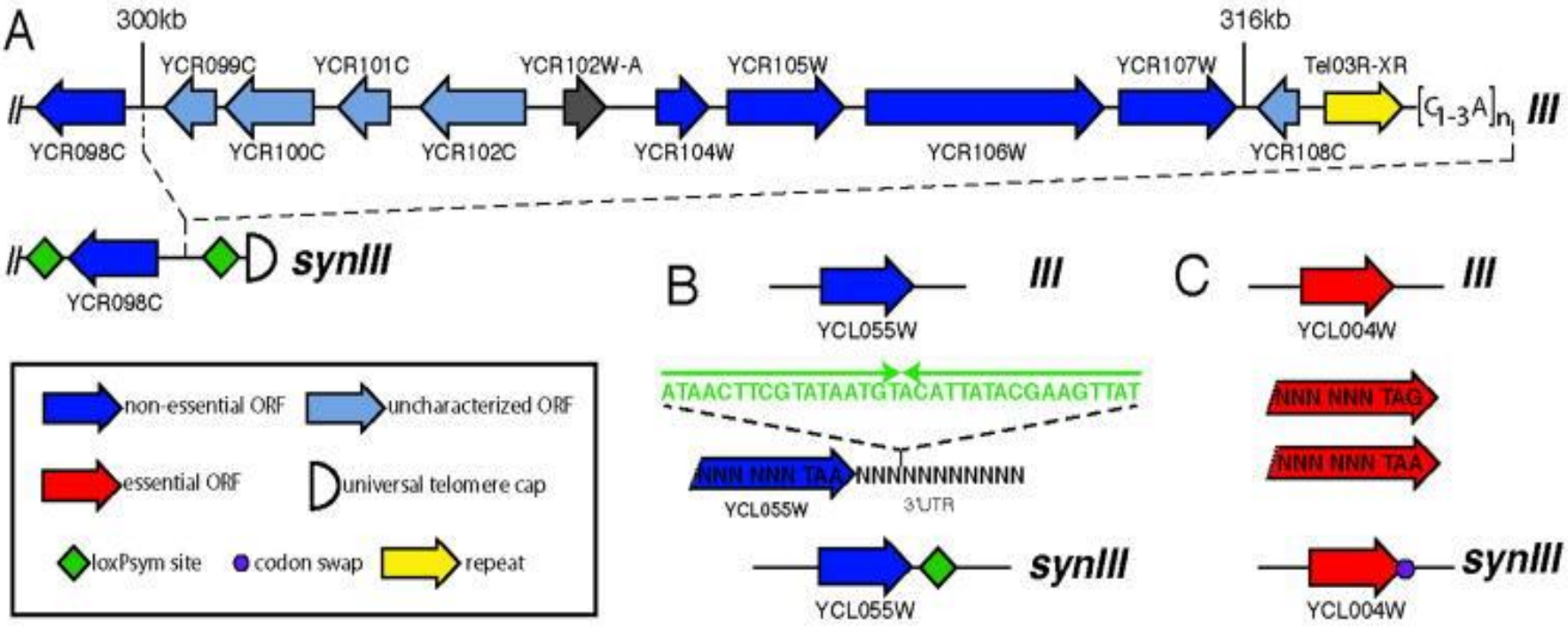
Alterations to the Yeast 2.0 genome

Deletion of transposons to increase the stability of the genome, as well as deletion of *SPT3* gene, required for Ty1 transposon transcription.

Relocation of tRNA genes that are hotspots for genomic instability onto a *separate* centromeric plasmid, a “neochromosome”.

SCRaMbLE. Site-specific recombination sequences, loxP sites, allowing subsequent evolution of the yeast strains will be introduced in all non-essential genes at the 3′ UTRs, as well as points where deletions are made (e.g. tRNA genes, transposons), ultimately resulting in ~5000 sites genome-wide. The loxP spesific recombinase will be expresssed conditionally, e.g. from galactose inducible promoter. This is a powerful tool for evolution and addressing what could constitute a minimal yeast genome (in particular conditions).

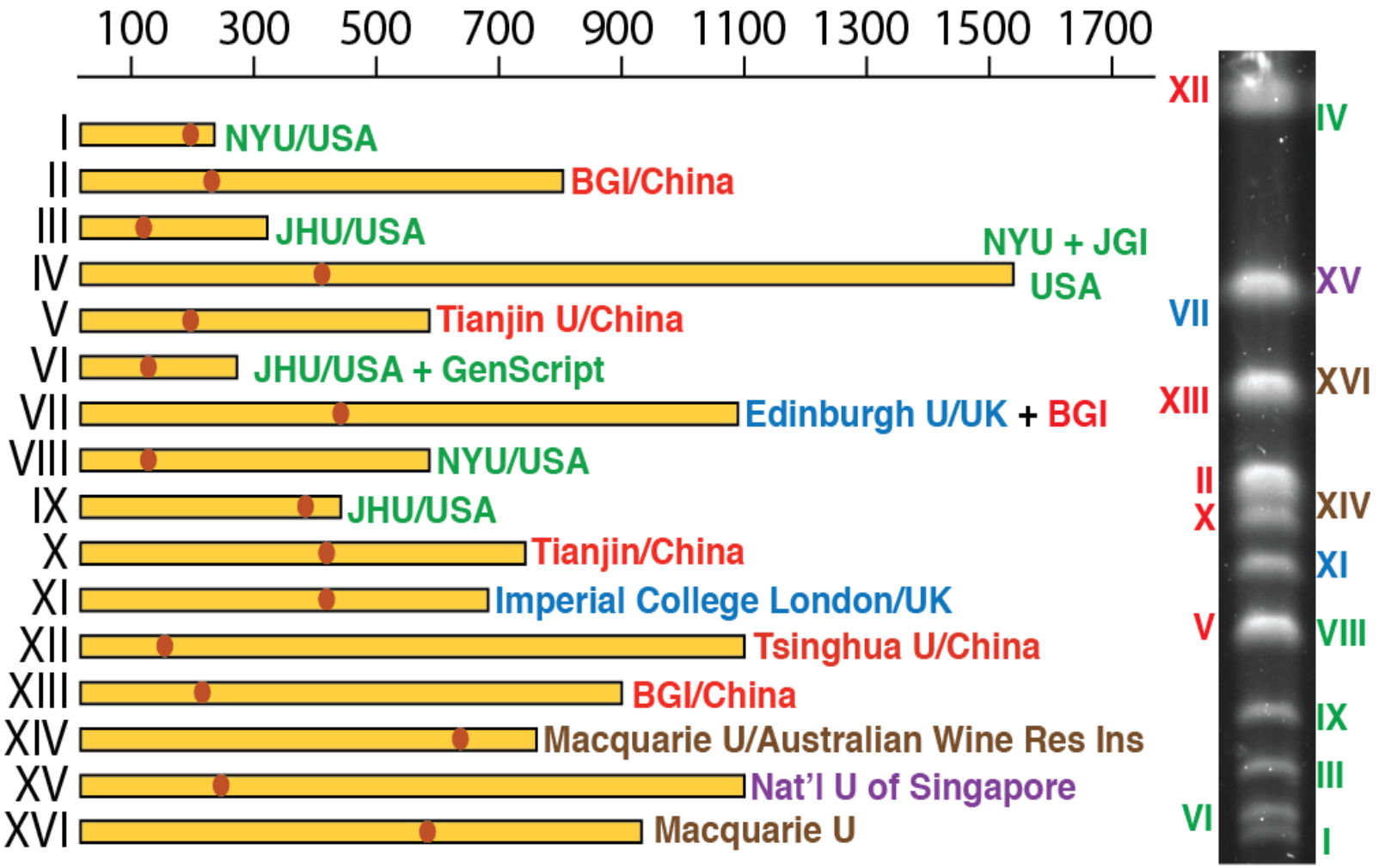
Building synthetic yeast chromosome III (*synIII*)



Representative *synIII* design segments for loxPsym site insertion (A & B) and stop codon TAG to TAA editing (C) are shown. Green diamonds represent loxPsym sites embedded in the 3′ UTR of non-essential genes and at several other landmarks. Fuchsia circles indicate synthetic stop codons (TAG recoded to TAA).

YCR099C: Y (yeast), C (chr.III), R (right from centromere), ORF 99, C/W (Crick/Watson coding strand)

A global consortium is needed for Yeast 2.0



Aalto University
School of Chemical Technology

http://syntheticyeast.org/collaborators/

Production host considerations

Bioreactors for ethanol production can be over 1000 m^3 (10^6 litres) large. Over 10^8 small cell factories fit in one litre.

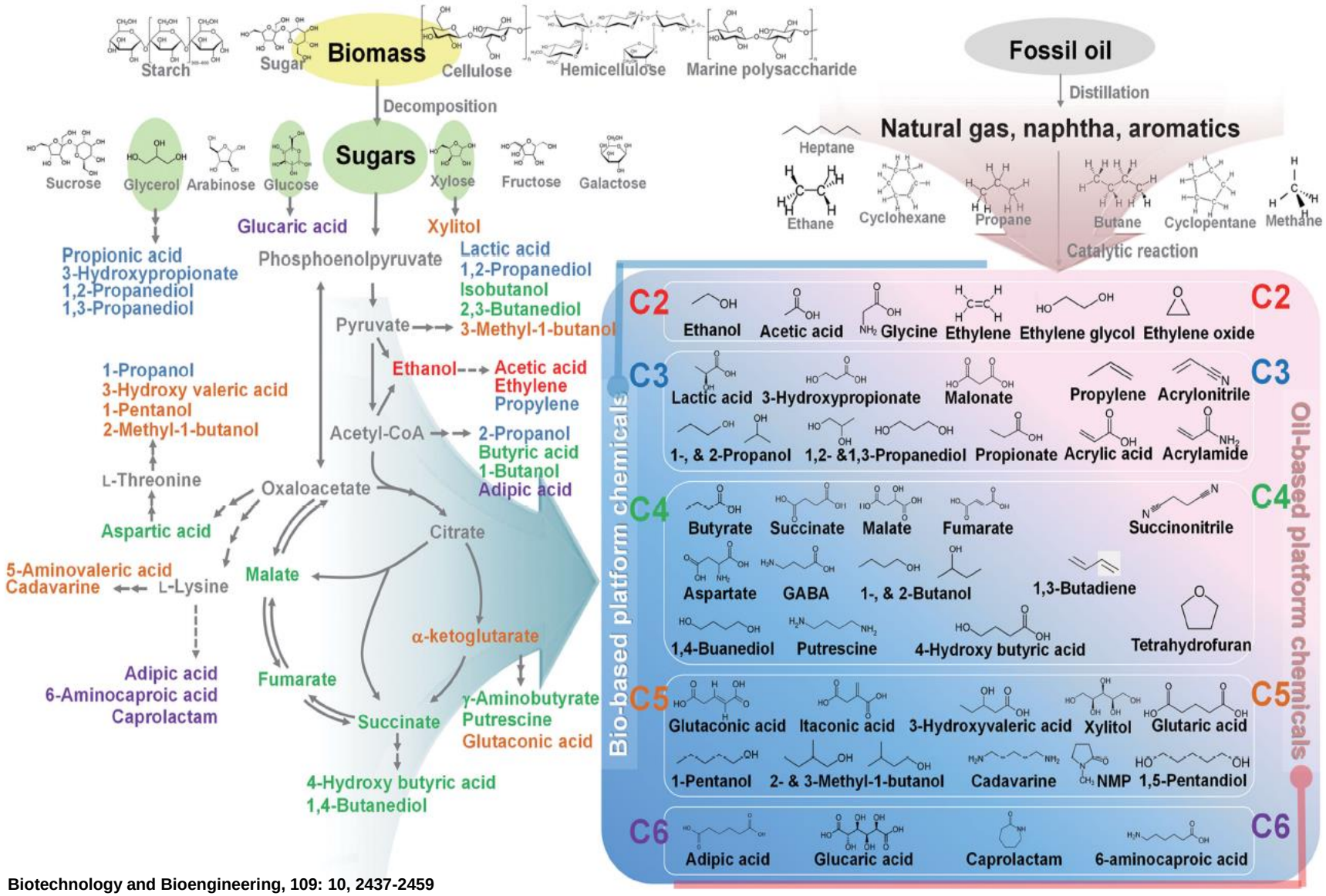
Synthetic chassis vs. use of synbio tools in “natural” hosts?

- Process robustness is important – traditionally difficult to engineer
  - Low pH, T, raw material or product tolerance, pressure, oxygen variation, growth rate
- Natural organisms, even non-conventional ones may provide beneficial features and natural biodiversity (e.g. lipid production, acid tolerance, difficult to engineer pathways)
- Host is critical for achieving high production yields, rates and titres

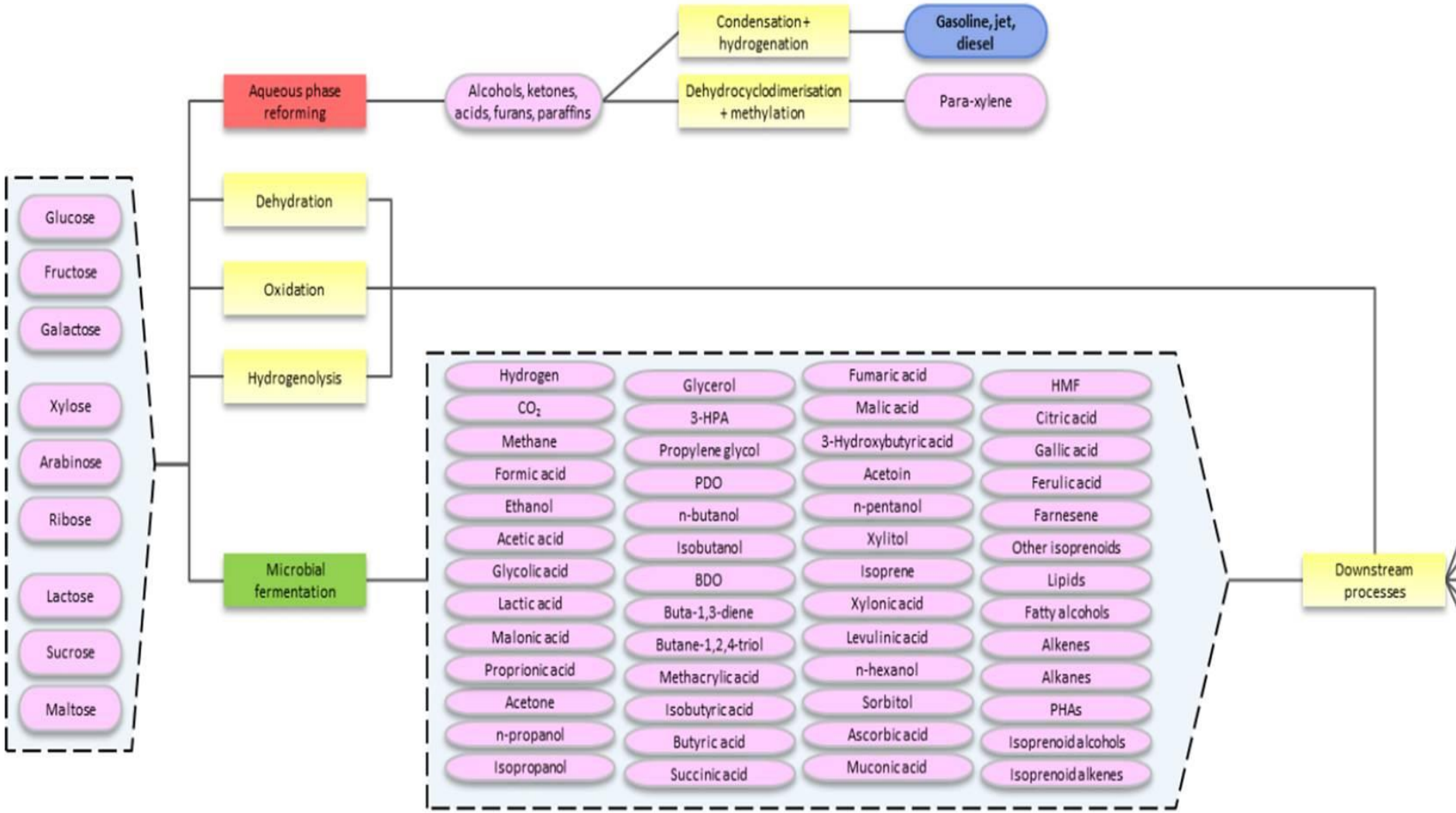
A difficult question: Synthetic chassis, favoured host, or a new natural one?

Does the Yeast 2.0 make a difference?

Cellular chemistry can be harnessed to produce platform chemicals and fuels that can replace those currently made from fossil resources – but also new chemicals for novel uses



Biotechnology and Bioengineering, 109: 10, 2437-2459



Downstream process options from sugars (the majority of which are fermentation based)

Needs in industrial production

- Replacement of fossil resources with renewable ones (plant biomass, photosynthesis) in production of chemicals, materials and fuels
  - Engineering of substrate utilisation pathways & photosynthetic organisms
- Equivalent products to petrochemicals by microbial fermentation
  - Metabolic engineering, heterologous pathway expression
- Novel, better products through biotechnology (materials, drugs etc)
  - Combinatorial pathways, novel enzyme catalysts
- Efficiency of production (titer, rate, yield)
  - Cut-off side reactions, increase flux, engineer cellular energetics & redox; predictive cellular modelling,.. thermodynamics, chemical biology etc
- Improve process robustness
  - Mutagenesis, product efflux, stress biomarkers,

Synthetic biology targets

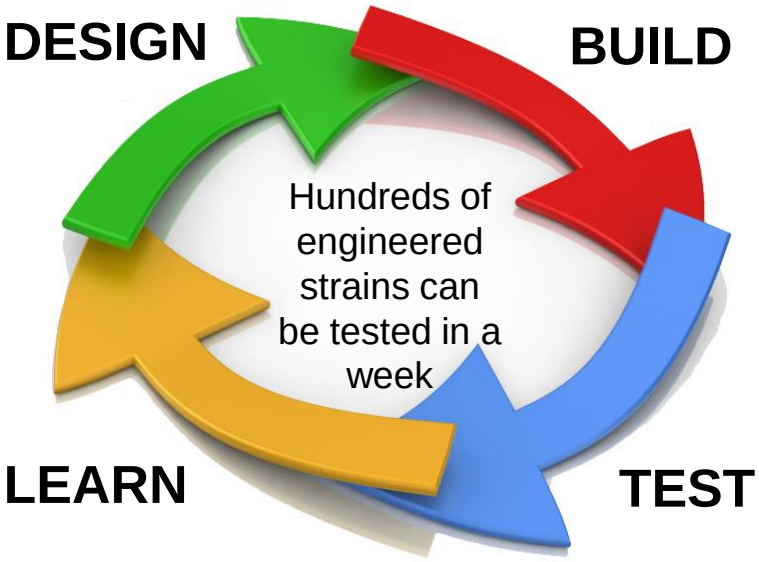
- Host strains that have predictable behaviour and are easy to manipulate (“minimalistic” chassis)
- New product pathways (balanced redox and energy, minimal carbon loss = carbon economy)
- Controllable and efficient expression (expression modules and circuits with synthetic designed elements)
- Novel chemistry (protein engineering, combinatorial biochemistry)
- Control of process robustness (intracellular sensors and control loops)

Living Factories

Automated strain engineering

Computational recepies for the robot to carry out Build and Test phases

Full automation of strain construction and cultivation



Design
Production strains and their parts are designed using computational tools

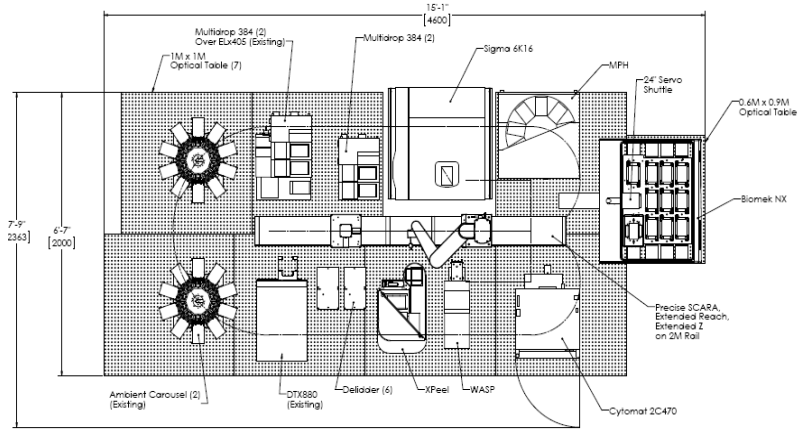
Construction of production strains
Synthetic DNA is delivered to the cells using genome editing tools such as CRISPR.

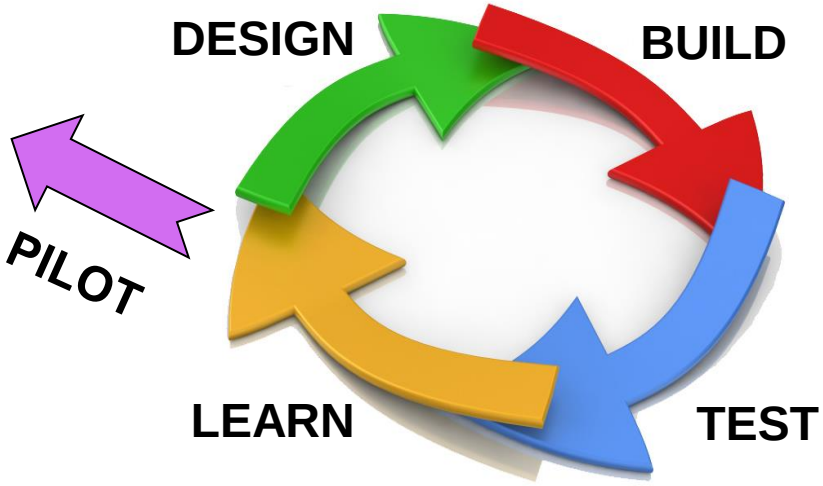
Analysis and decisions
Machine learning algorithms can help the researcher to analyse and understand measured data.

Cultivation and measurement
Robots are cultivating the strains and carry out measurements. The results are automatically stored in databases.

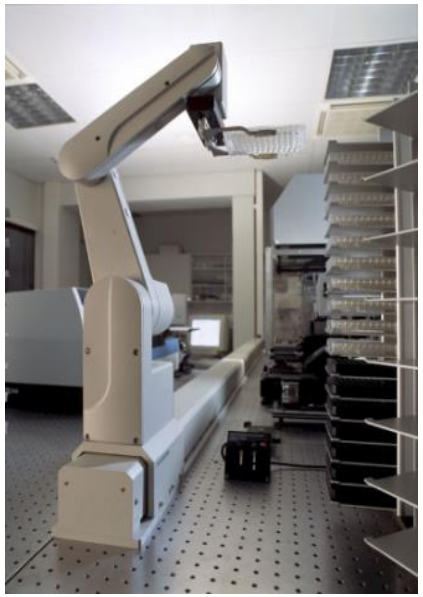
Aalto-VTT national Bioeconomy infrastructure: From synthetic biology to piloting

A versatile computing platform for design, prediction and analysis

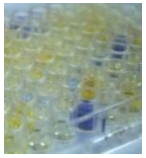




A robotic platform for efficient DNA assembly, transformation and strain screening



Controlled parallel bioreactor systems with automated sampling and analytics



Reading

N. Annaluru, J. Boeke et al., “Total synthesis of a functional designer eukaryotic chromosome,” *Science*, 2014. doi:10.1126/science.1249252,

http://syntheticyeast.org/

http://www.yeastgenome.org/
SGD Saccharomyces genome database

YeastBook: An Encyclopedia of the Reference Eukaryotic Cell

Genome engineering: Synthetic genome technology for yeast.
Casci, Tanita // Nature Reviews Genetics; Nov2011,
Vol. 12 Issue 11, p742

Group work

Prepare together a **15+5 min presentation** that contains

A synthetic design of a ~30kb region (”megachunk”) of a selected chromosome.

- Why did you choose this region?
- What would you include or omit from the design? (No need to go for a single nucleotide level). Illustrate as in slide 19.
- Which computer programs would you use/need?
- Brief explanation of the wet lab construction procedure

Answers to the questions:

- Why is the Yeast 2.0 synthetic?
- What is the significance and impact of the yeast 2.0?
- What would you use the yeast for or develop further? How?

You may also point out the possible problems you encounted in your group work.

Special points to be addressed:
Group 1: Explain how syn chromosomes can be combined into one strain
Group 2: Explain the Scramble mechanism
Group 3: How would you like to make the strain a better chassis (may be a basis for choosing the chromosomal region)