Sc2.0 as a driver for the field of genome writing

Leslie Mitchell

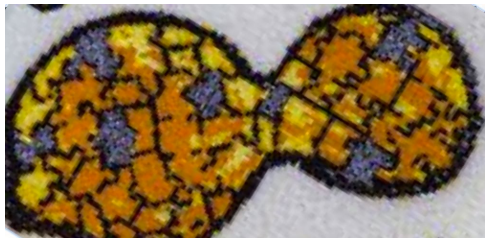
Build-A-Genome Network
July 29 2022

Outline of today's talk

- Intro to the Sc2.0 project
- Project updates
- Sc2.0 as an critical driver of the genome writing field
  - Education
  - Technology development
  - Commercialization

Sc2.0 - the synthetic yeast genome project

- Design & build an entirely synthetic eukaryotic genome to power growth of *S. cerevisiae*
- Improve genetic flexibility and genomic stability while maintaining high fitness



“What would it take to write a designer yeast genome from scratch?”

-- Jef Boeke, PhD, *circa* 2007

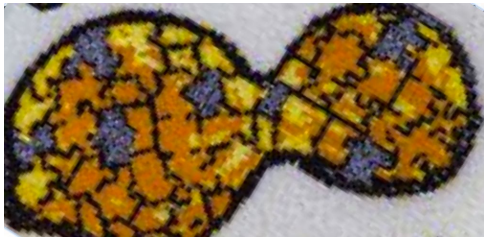
- 12,000,000 bp
- 16 chromosomes
- 6000 genes
- 16 centromeres
- 32 telomeres

JOHNS HOPKINS UNIVERSITY

NYU Langone Health

- Tech dev?
- Team?
- Funding?
- Time?

Sc2.0 design features



Delete

- Repeats (transposons, LTRs, Sub-telomeres)
- Introns

Re-locate

- tRNA genes moved to a new chromosome

Add

- loxPsym sites in 3'UTRs; 'SCRaMbLE'

Re-code

- Watermark sequences (PCRTags)
- 'Stop swaps' (TAG to TAA)

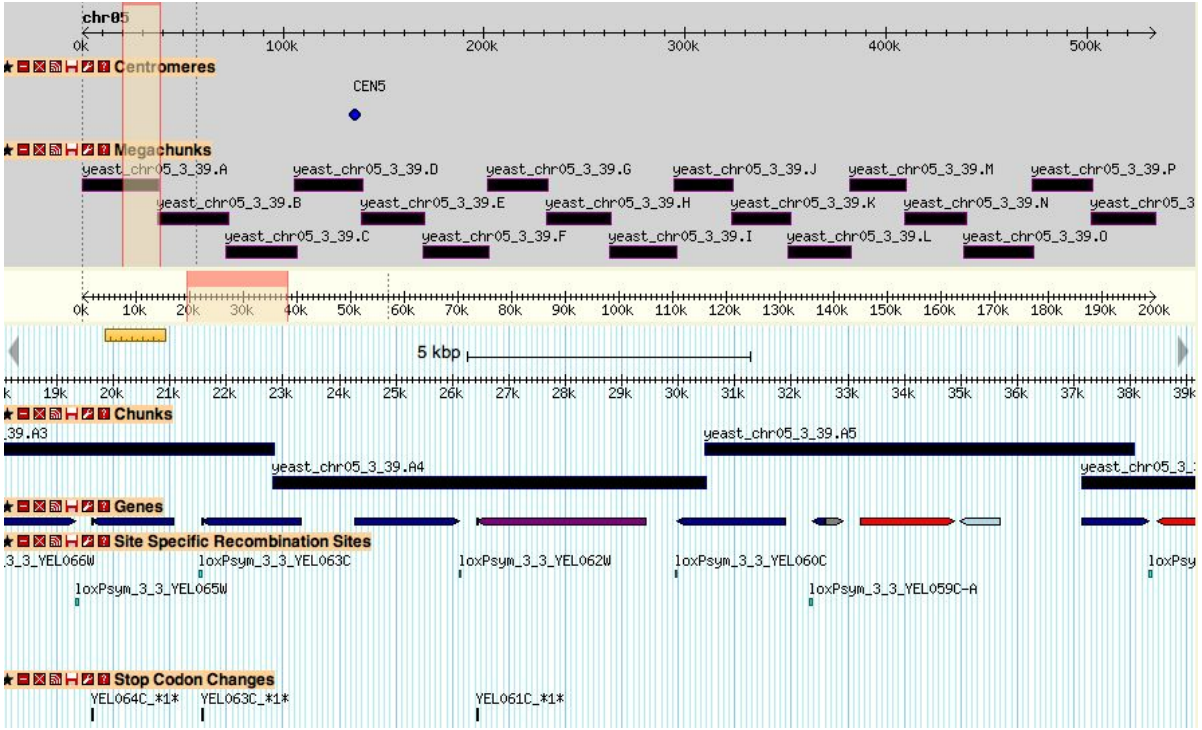
Improved genome stability

Inducible genetic flexibility

- Identification
- Biosecurity
- New protein chemistry

Biostudio, custom design software for Sc2.0

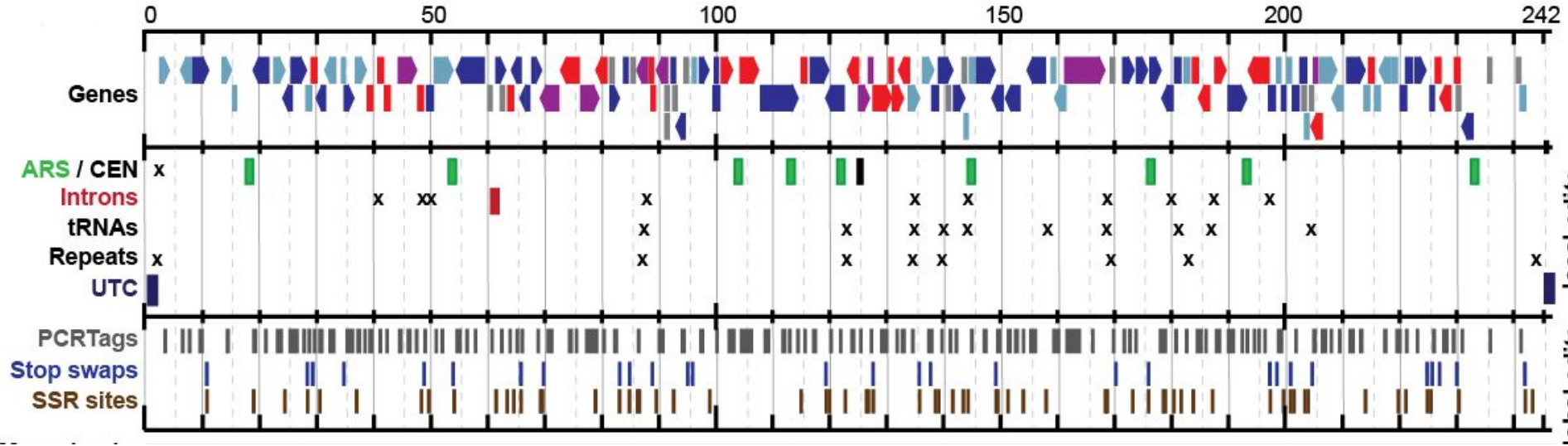
GBrowse “souped up” with editing capabilities (deletion, insertion, substitution)



Richardson et al. 2017

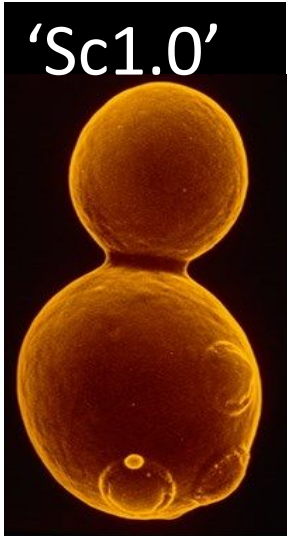
Sc2.0 design applied to chromosome *VI*, *synVI*

- Clusters of designer changes every ~400 bp
- 10% or more shorter in overall length
- Relatively evenly spaced loxPsym sites for SCRaMbLE



Mitchell et al. 2017

Sc2.0 - completely redesigned eukaryotic genome



~750 kb deleted

~215 kb re-coded

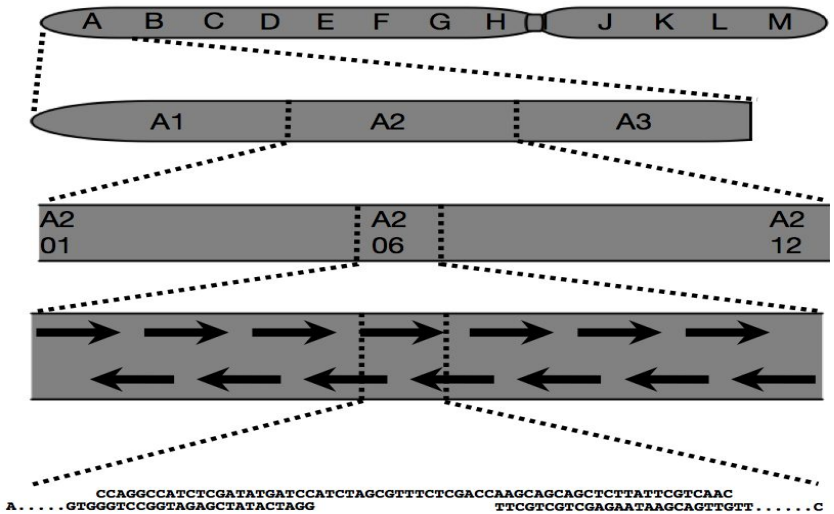
~132 kb inserted

~1.1 Mb changed

~8% shorter than native

Sc2.0 - the major technical challenge

How does one actually construct a “new-to-the-world” sequence that is ~200 kb to >1Mb in length?



Megachunks: 30-60 kb

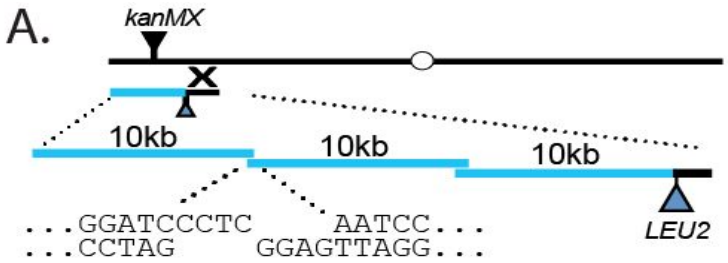
Chunks: 10 kb

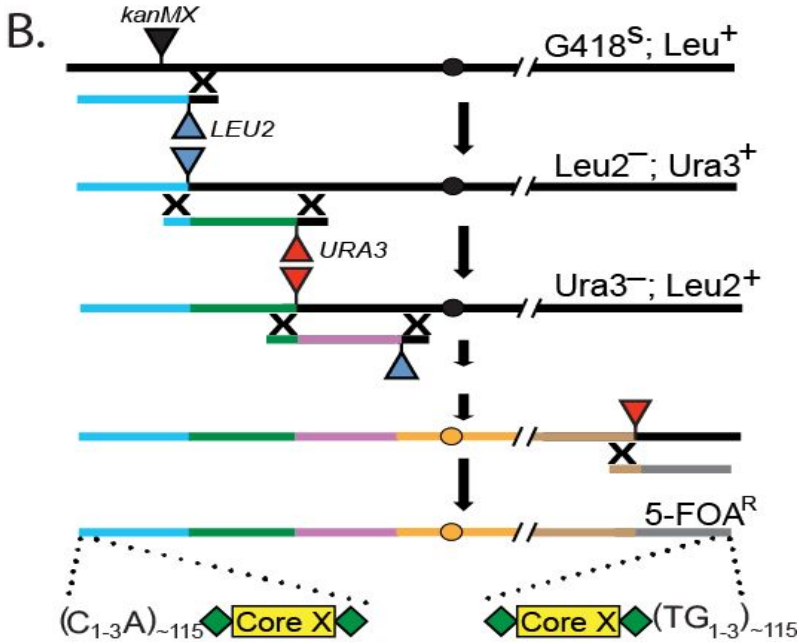
Minichunks: 3kb

Building blocks: 750 bp

Oligos: 70 bp

Sc2.0 - SwAP-In to build synthetic chromosomes

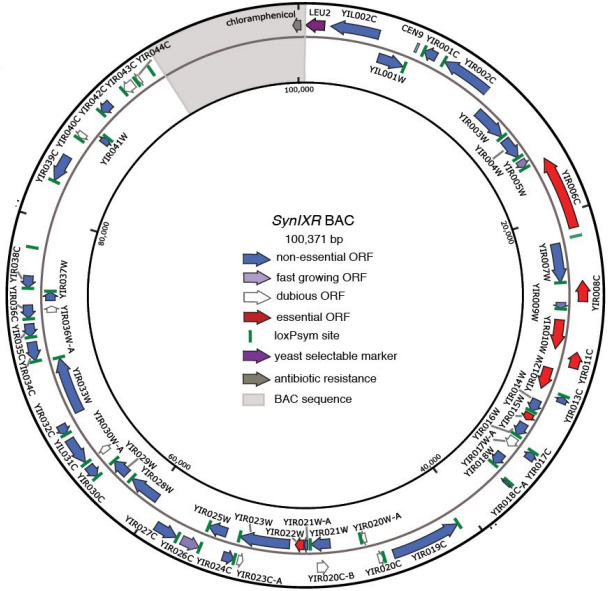




Richardson et al. 2017

synIXR and *synIII* - Sc2.0 POC

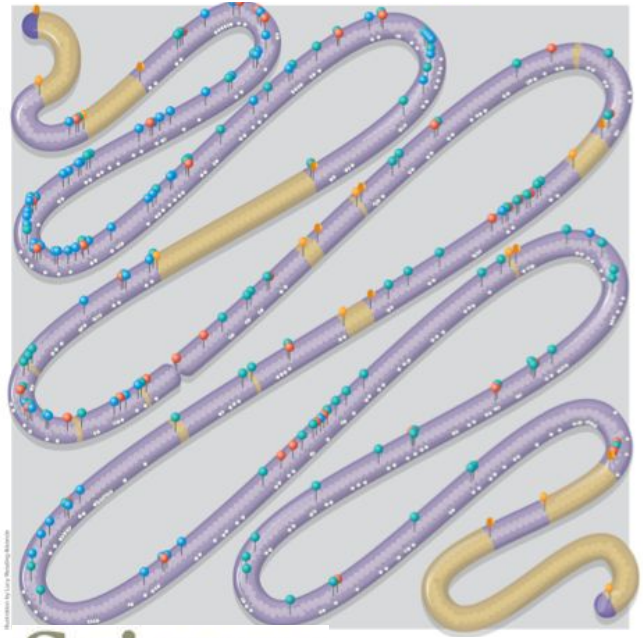
SYNTHETIC CHROMOSOME ARMS FUNCTION IN YEAST & GENERATE PHENOTYPIC DIVERSITY



nature 477, 471–476 (22 September 2011) International weekly journal of science

Dymond et al. 2011

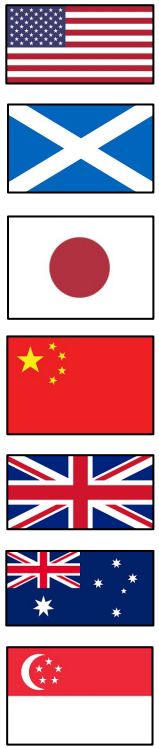
TOTAL SYNTHESIS OF YEAST DESIGNER CHROMOSOME *SYNIII*

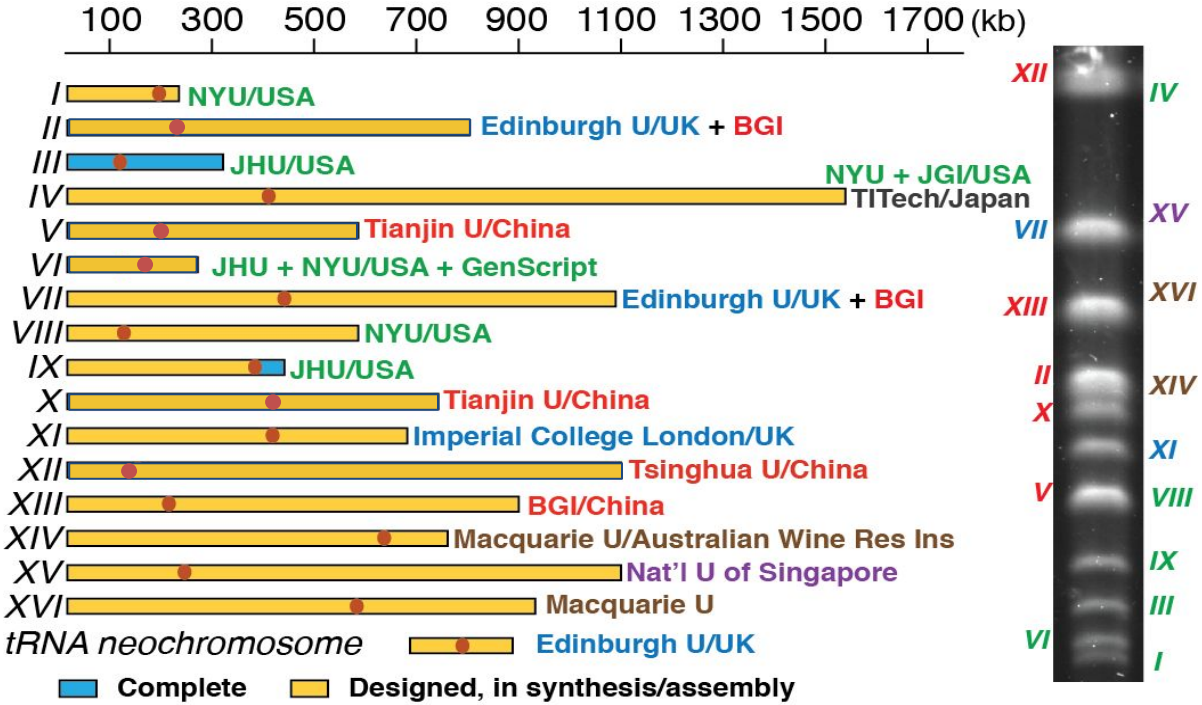


Science

Annaluru et al. 2014

Sc2.0 - an international consortium launched in 2012





Sc2.0 - major milestone in 2017

March 2017

- Design of the Sc2.0 genome
- Assembly & de-bugging of five synthetic chromosomes:
  - *synII* – 770 kb
  - *synV* – 534 kb
  - *synVI* – 240 kb
  - *synX* – 706 kb
  - *synXII* – 980 kb (unique)
- 3D nuclear organization of synthetic chromosomes

~30% of Sc2.0 genome

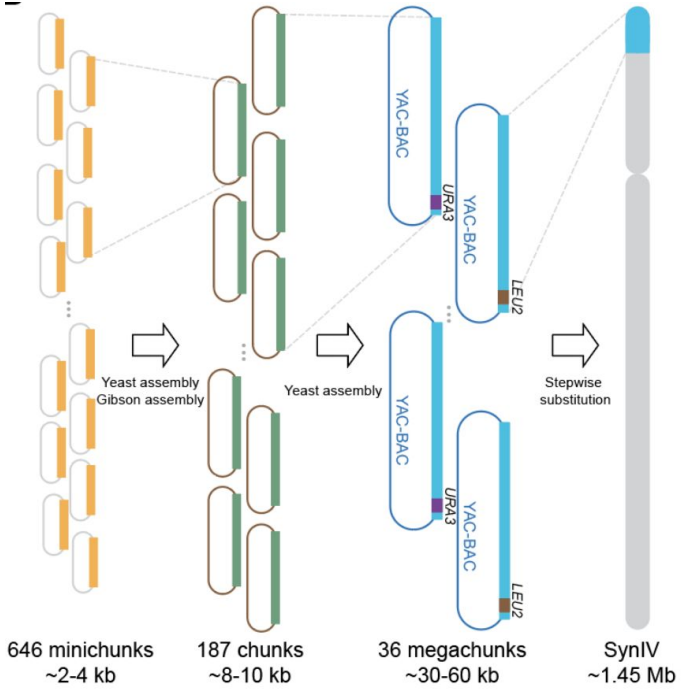
Sc2.0 - updates

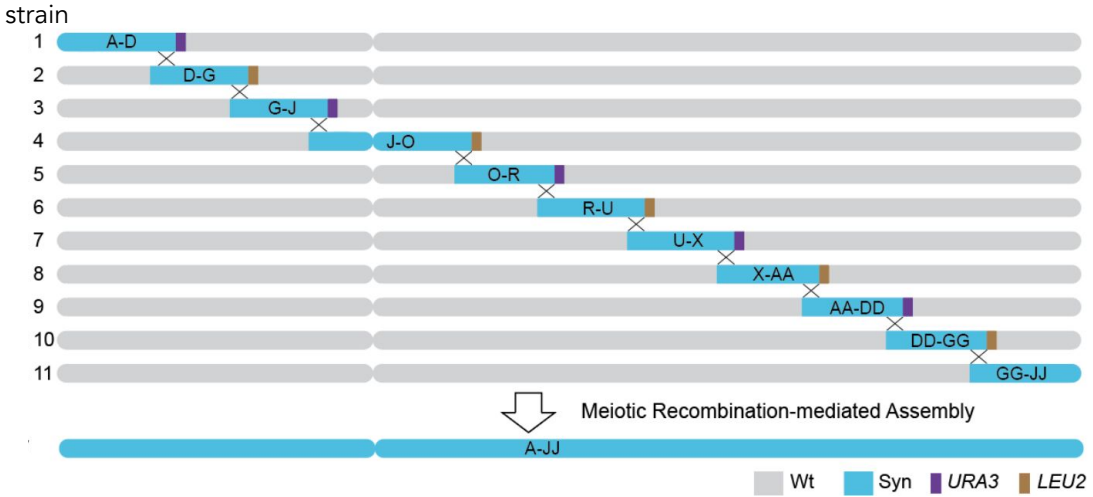
- Assembly & de-bugging of four synthetic chromosomes complete:
  - *synI*
  - *synIV*
  - *synXI*
  - *synXIV*
- Consolidation of 6.5 synthetic chromosomes into one strain

Remainder of chromosomes are complete, manuscripts in prep

Sc2.0 - highlights to date

synIV, the longest Sc2.0 chromosome is now complete (~1.45 Mb)

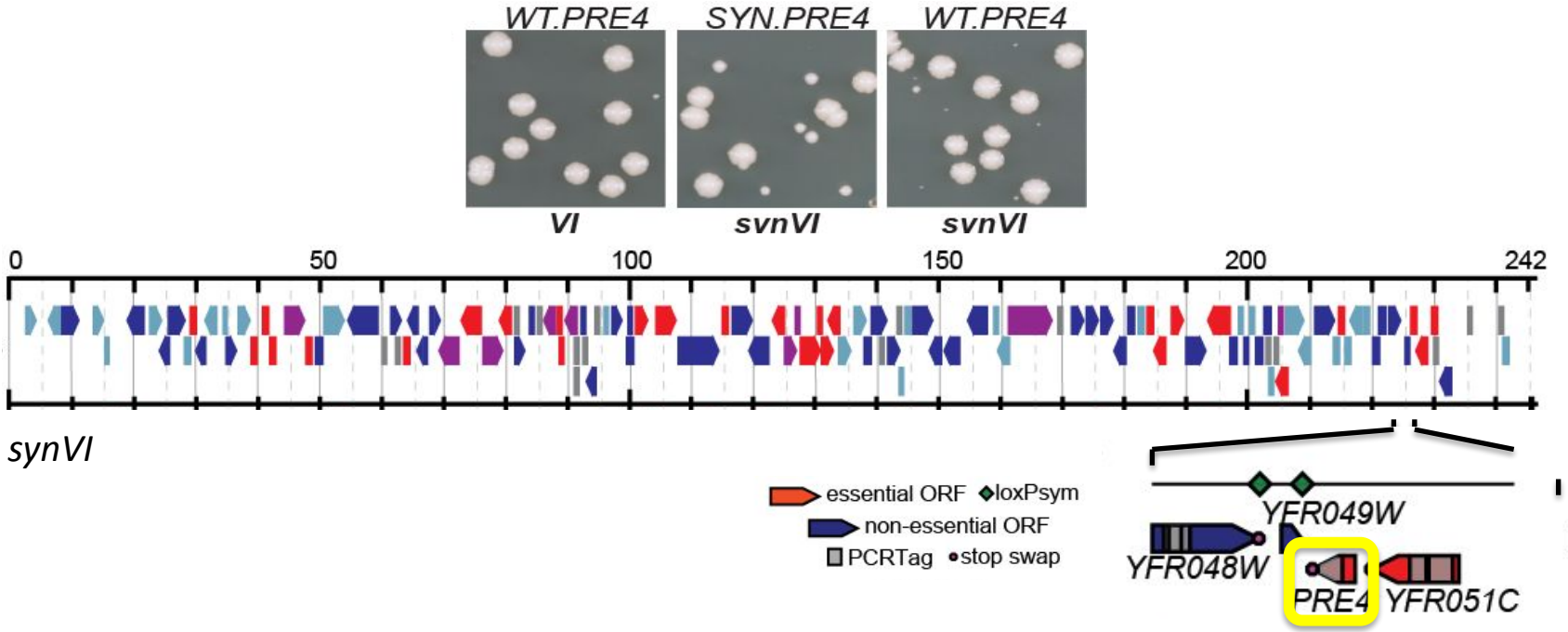




Zhang et al. 2022, bioRxiv

Sc2.0 - highlights to date

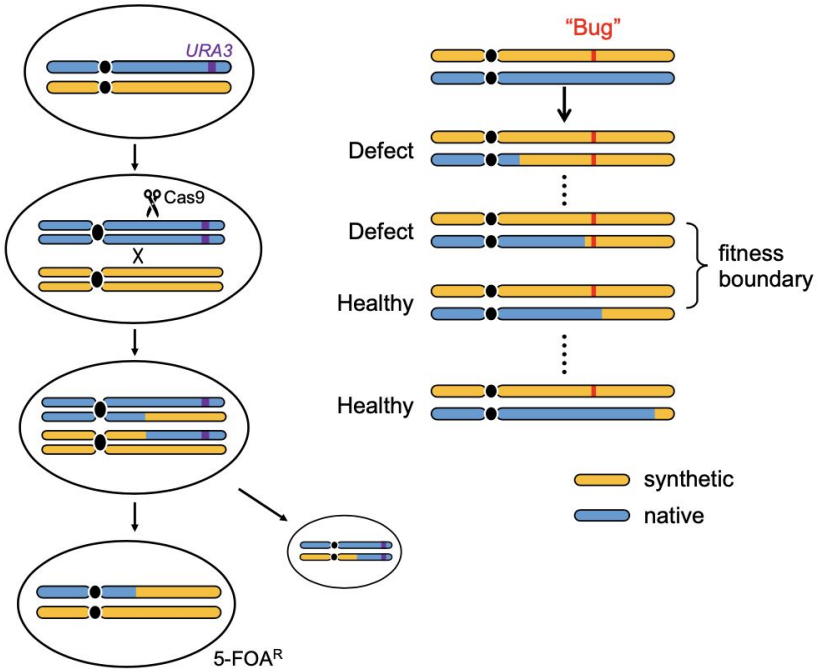
Rare ‘bugs’ represent an opportunity for biological discovery



Mitchell et al. 2017

Sc2.0 - highlights to date

CRISPR D-BUGS to map fitness defects

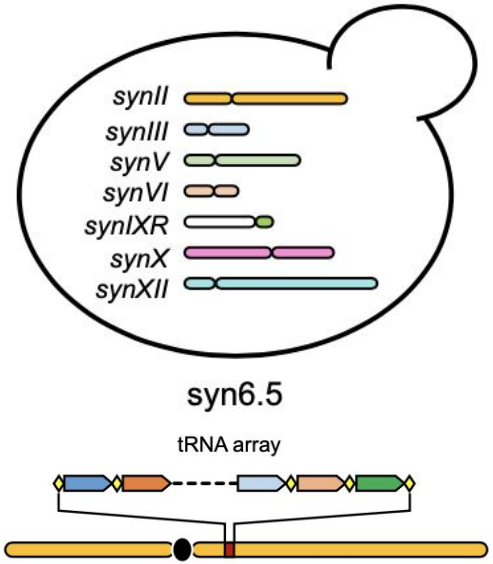


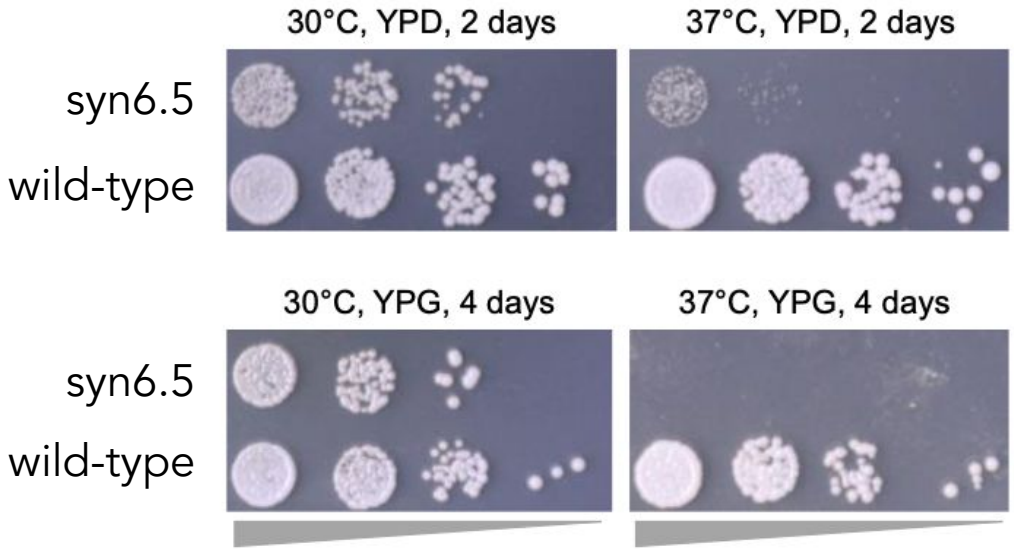
Zhao et al. 2022, BioRxiv

Sc2.0 - highlights to date

6.5 synthetic chromosomes are consolidated into a single strain

91 introns deleted | 97 tRNAs relocated | 444 stop codons swapped | 1181 loxPsym sites | 4814 PCRTag recoded | 13 universal telomere caps

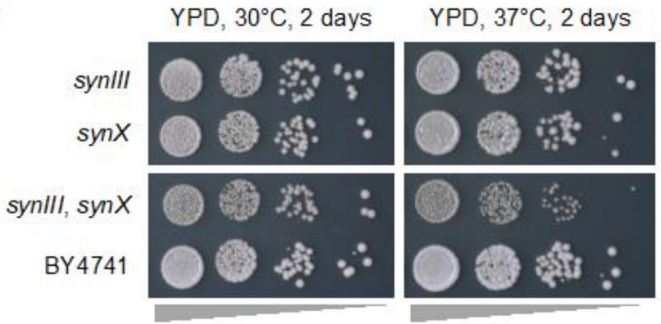


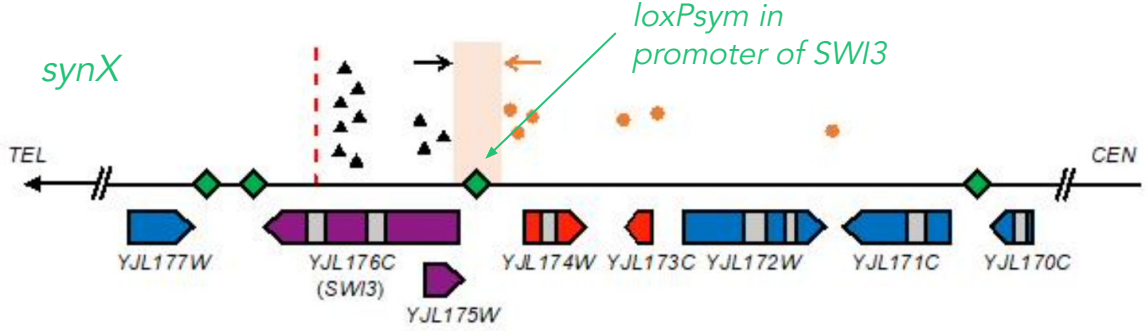


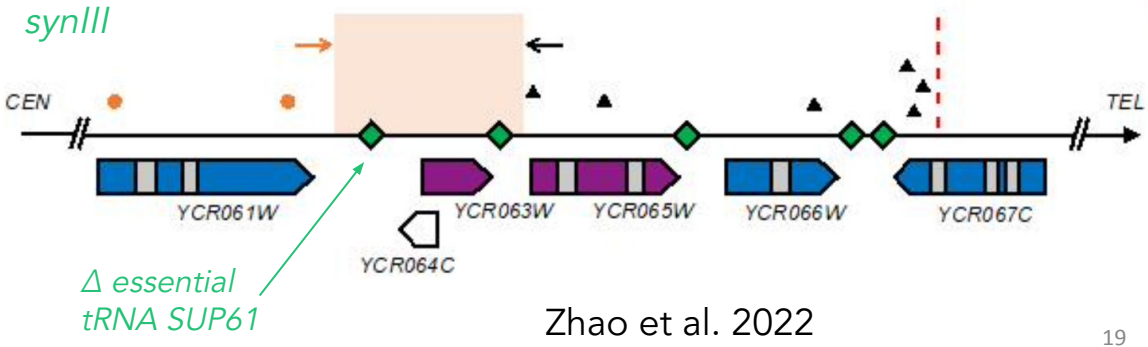
Zhao et al. 2022

Sc2.0 - highlights to date

Combinatorial growth defects arise - *synIII* x *synX*





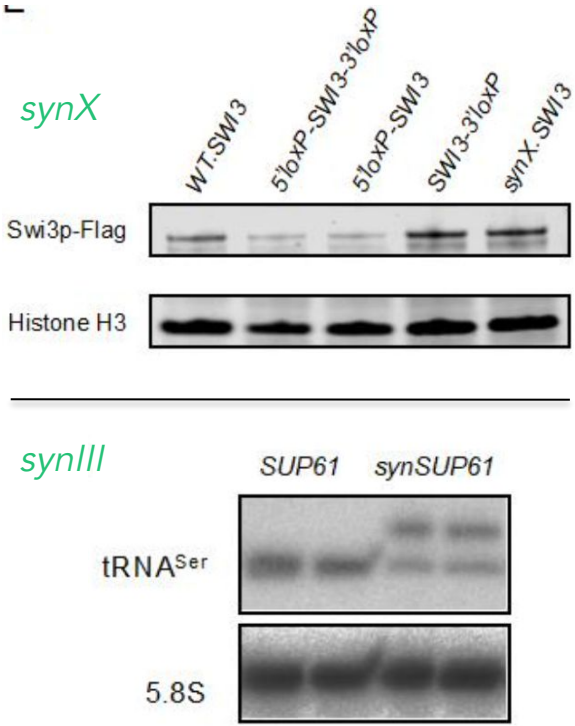


- loxPsym
- Essential
- Req'd for fast growth
- Nonessential

Zhao et al. 2022

Sc2.0 - highlights to date

Combinatorial growth defects arise - *synIII* x *synX*



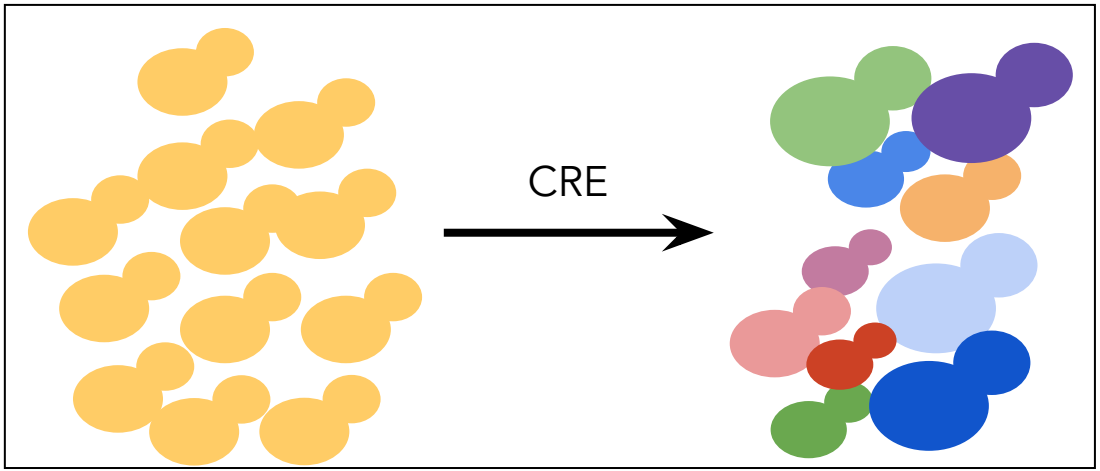
- Single-copy *SUP61* produces the only tRNA decoding the rare UCG serine codon
- *SWI3* has an above average usage of UCG for serine, including two tandem UCG codons
- Hypothesis: reduced abundance of *synSUP61* further reduces expression of *SWI3* below already low level
- Two buggy components restored to wild-type form results in wild-type fitness
- All known bugs repaired in *syn6.5* strain led near wild-type fitness with only residual growth defects at high temperature

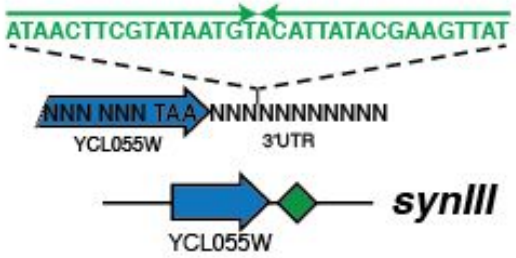
Zhao et al. 2022

Sc2.0 - highlights to date

SCRaMbLE is a powerful tools to study genome structure-function

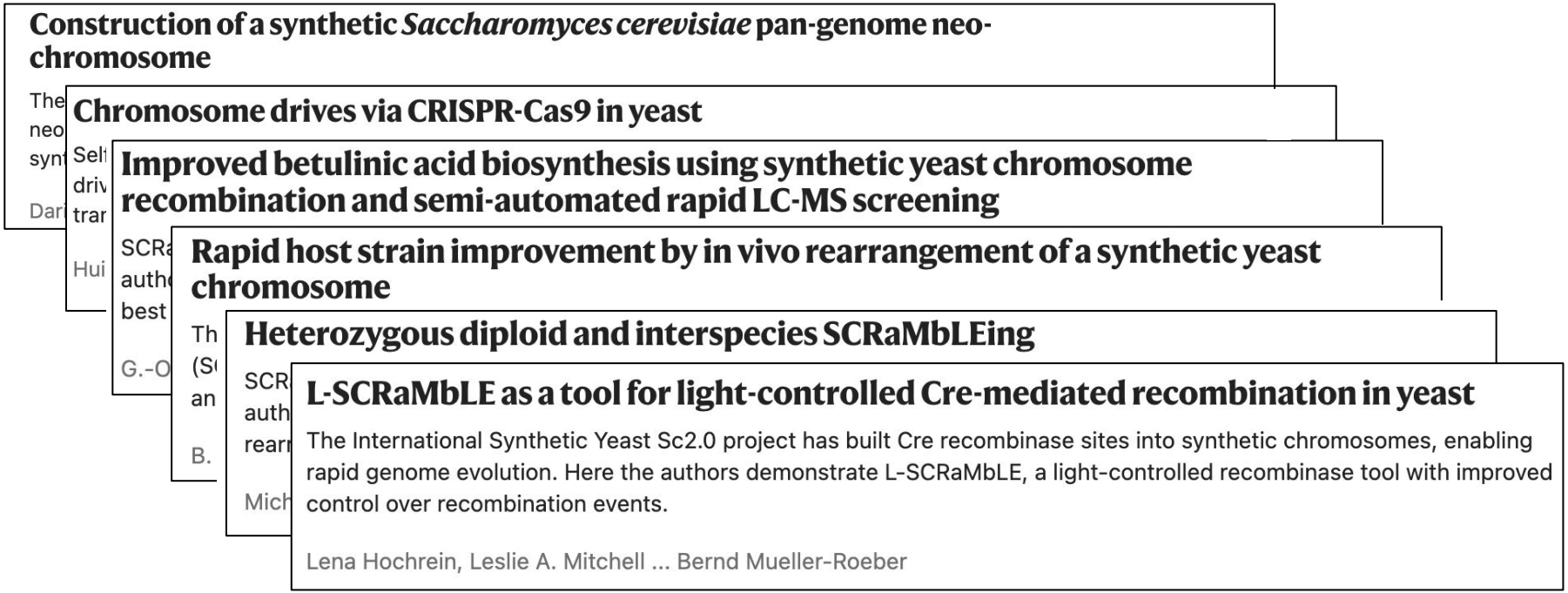
- Synthetic Chromosome Rearrangement and Modification by LoxP-Mediated Evolution
- LoxPsym site encoded downstream of non-essential genes on syn chrs
- Express Cre recombinase to cause deletions/inversions/duplications





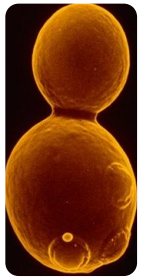
Sc2.0 - highlights to date

Nature Communications “special collection”



“What would it take to write a designer yeast genome from scratch?”

-- Jef Boeke, PhD, *circa* 2007



- 12,000,000 bp
- 16 chromosomes
- 6000 genes
- 16 centromeres
- 32 telomeres

JOHNS HOPKINS UNIVERSITY

NYU Langone Health

- ~15 years
- ~20 institutions
- ~50(+) scientists
- > $50 M?
- Unlimited new learnings

Seven consecutive papers in a single issue of Science.

Sc2.0: the world's first designed eukaryotic genome synthesized from scratch

- Centralized design with distributed collaboration model including labs across 4 continents
- Functional demonstration of the entire genome in yeast cells, bugs leads to new biological discoveries
- Formalized technology stack focused on eukaryotic genome-scale synthetic biology

Sc2.0 has budded new generations of genome writers

- 100's of new trainees:
  - Build-A-Genome: JHU, Tianjin U, NYU
  - Build-A-Genome Network: Lisa Scheifele, Eric Cooper, Rob Newman
- 10's of new PIs:
  - Establishing Sc2.0 and other genome writing projects globally
- 1 new and fast-growing community
  - Annual Sc2.0 meetings
  - GP-Write
  - Increasingly ambitious mammalian genome writing projects

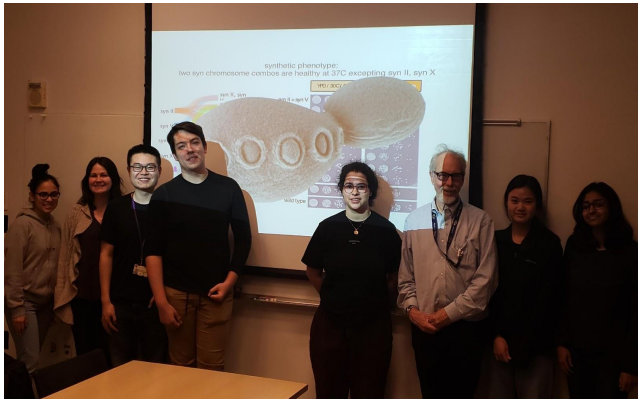
National Science Foundation
WHERE DISCOVERIES BEGIN

NYU Build-A-Genome class

- 18 trainees to date
- 2020/2021 - synthetic chromosome consolidation (endoreduplication intercross)
- Spring 2022 - next-gen synthetic chromosome consolidation (chromosome swapping, aka KAR crosses)
- Spring 2023 - big-DNA assembly (mouse, human) for delivery to cells in culture to study disease and discover therapies

Yu "Jeremy" Zhao
Postdoc
Boeke lab

Stephanie Lauer
Postdoc
Boeke lab



Spring 2020 NYU BAG course

Expanding the scope of genome writing beyond yeast

Technological challenges and milestones for writing genomes

Synthetic genomics requires improved technologies

By **Nili Ostrov[1], Jacob Beal[2], Tom Ellis[3], D. Benjamin Gordon[4], Bogumil J. Karas[5], Henry H. Lee[1], Scott C. Lenaghan[6], Jeffery A. Schloss[7], Giovanni Stracquadanio[8], Axel Trefzer[9], Joel S. Bader[10], George M. Church[1,11], Cintia M. Coelho[12], J. William Efcavitch[13], Marc Güell[14], Leslie A. Mitchell[15], Alec A. K. Nielsen[16], Bill Peck[17], Alexander C. Smith[18], C. Neal Stewart Jr.[19], Hille Tekotte[20]**

SCIENCE sciencemag.org 18 OCTOBER 2019 • VOL 366 ISSUE 6463

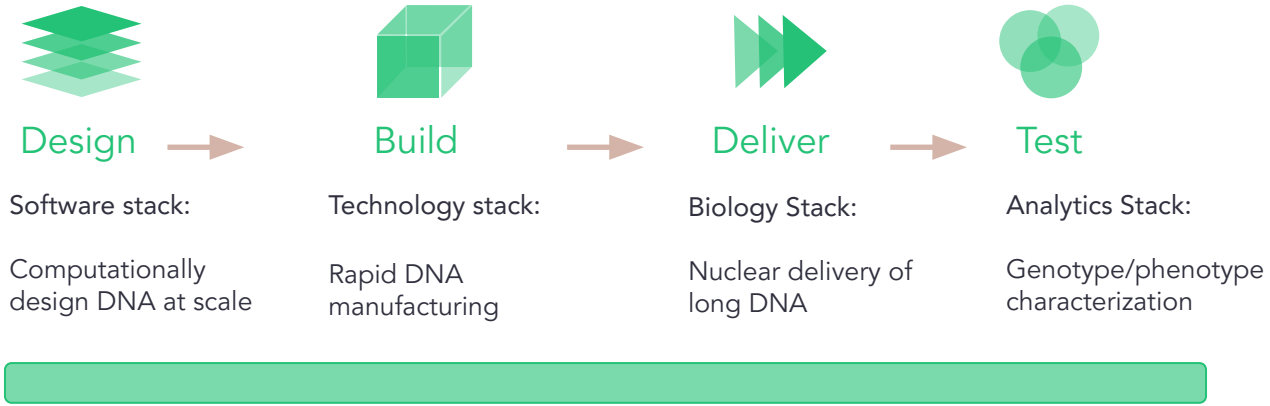
Sc2.0 lessons learned applied to mammalian genomes

- Scalable DNA assembly methods
- Delivery, delivery, delivery

KEY TECHNOLOGY DEVELOPMENT TARGET	EXAMPLE OF DESIRED MILESTONES	ESTIMATED TIME (YEARS)
Genome design		
Develop tools for genome-scale design, visualization, and quality control.	Design a virus-proof mammalian chromosome.	3
Integrate structural information (2D and 3D) into genome design software	Predict the conformation of a synthetic yeast chromosome.	5
Develop sequence-to-phenotype whole-cell modeling.	Optimize metabolic profile, accurate to within twofold, for 100 key gene products of a synthetic virus-proof chromosome.	10
DNA synthesis		
Increase coupling efficiency for oligonucleotide synthesis.	Synthesize high-fidelity oligonucleotides longer than 500 nucleotides.	3
Increase efficiency of in vitro DNA assembly for fragments >20 kb.	Assemble 20 kb with >50% yield.	4
Develop methods for synthesis of difficult sequences, including homopolymers, high-GC content, and secondary structure.	Synthesize a centromere.	5
Develop enzymatic methods for direct synthesis of multikilobase DNA fragments.	Synthesize a 10-kb fragment (without assembly).	7
Decrease cost of DNA synthesis by 1000-fold.	Synthesize and assemble DNA for one haploid human genome (i.e., 3.2×10^9 bases) for $1000.	10
Chromosome construction		
Develop methods for temporal and spatial control of single chromosomes, such as chromatin state.	Engineer segregating, stable human artificial chromosomes (HAC).	2
Develop specialized host cells with high efficiency for DNA assembly, particularly for difficult-to-assemble sequences.	Establish in vivo chromosome assembly methods in the host *Streptomyces coelicolor* (72% GC content).	5
Develop efficient, inexpensive methods for routine and automated delivery of entire chromosomes into cells.	Demonstrate routine, device-based chromosome delivery in mammalian cells by cell fusion.	3
Develop methods for assembly and testing of Mb-size chromosomes.	Assemble a synthetic recoded human chromosome 21 from DNA fragments.	10

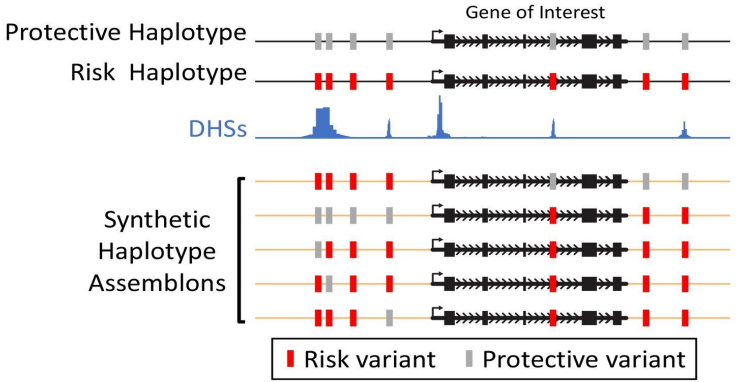
Genome-scale cell engineering tech stack

A combination of software and wetlab processes to enable DNA reprogramming inside cells at scale



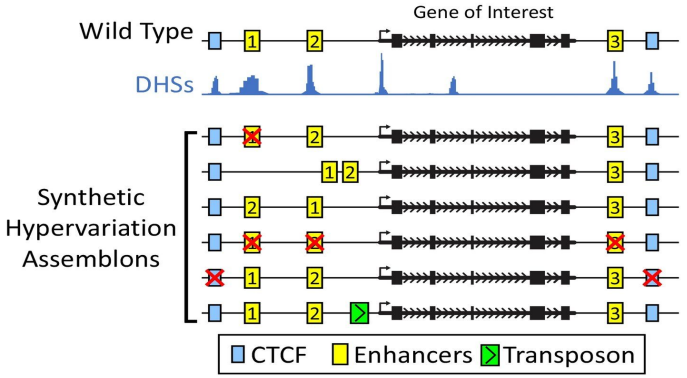
Locus-scale mammalian genome engineering

Synthetic haplotypes:



Complex haplotypes with many variants
(SNPs, DNAse-hypersensitive sites)

Synthetic hyper-variation:

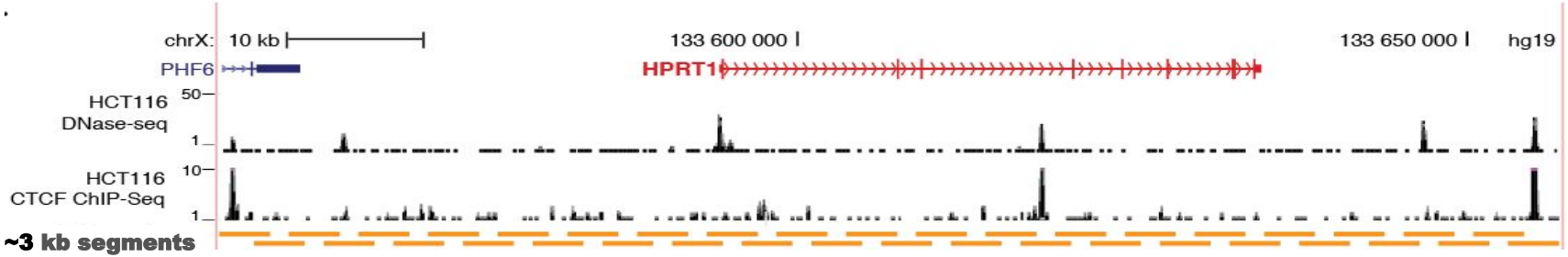


Loci with many regulatory sites
(enhancers, CTCF sites, transposons, ncRNAs)

- Design: combinatorial manipulation of elements of interest *in silico*
- Build: 100's of variant assemblons '*in yeasto*'
- Test: deliver variants to precise location in genome & survey effects on expression & chromatin state
- Model effects of sequence variation on gene expression

Transplanting the 102 kb human HPRT1 gene into mouse cells

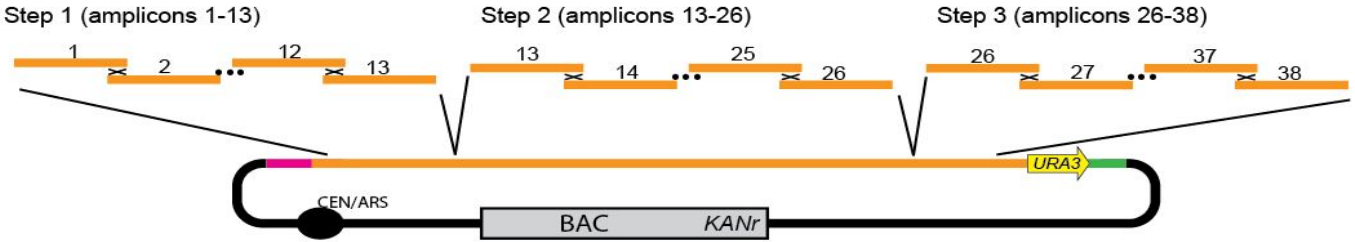
- Human HPRT1 (hHPRT1) encodes an enzyme with a key role in purine salvage
- Sequence: ~102 kb in length (40 kb gene body, 60 kb 'regulatory sequence')
  - Regulatory landscape includes DNase hypersensitive sites and CTCF binding sites



- Isolate the 102 kb *hHPRT* sequence from human genome
- Re-design sequence to desired specifications, introduce designer elements to enable downstream delivery
- Sub-divide locus into ~3-5 kb segments with terminal homology optimized for highly efficient *in yeasto assembly*
- Source ~3-5 kb segments from commercial synthesis provider or by PCR

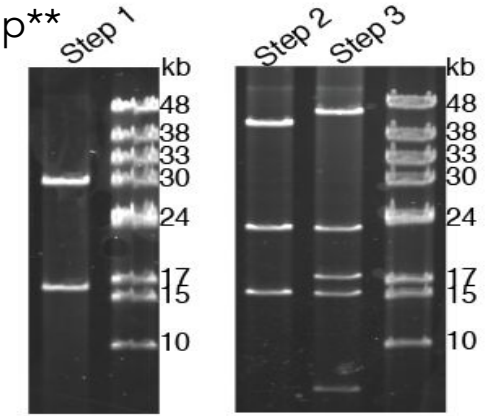
Mitchell et al. 2021

Transplanting the 102 kb human HPRT1 gene into mouse cells



- Build 102 kb hHPRT1 gene in three steps in yeast
  we are now building 100 kb assemblons in a single step
- Transfer assemblons to *E. coli* for digestion & sequence verification, and to produce DNA for delivery

	Amplicons	Selection	Length	PacI digestion
Step 1	1-13	Ura+	47 kb	31,16kb
Step 2	13-26	Leu+ Ura–	73 kb	43, 24,16kb
Step 3	26-38	Ura+ Leu–	102 kb	46, 24,19,16,7kb

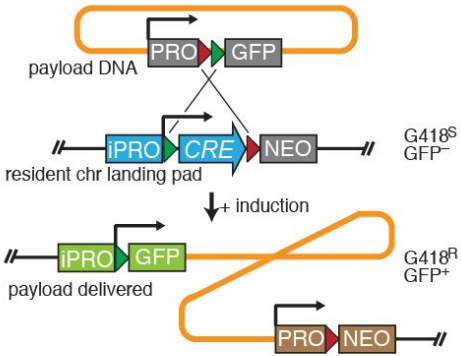


FIGE: PacI digestion

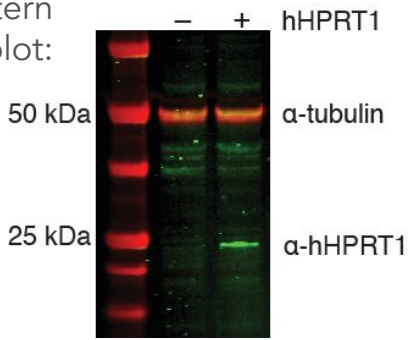
Transplanting the 102 kb human HPRT1 gene into mouse cells

- DNA delivery by "Inducible Cassette Exchange" (ICE)
- Select GFP+/G418^R clones, validate expected genotype by PCR
- Show functional expression of protein by Western blot with *hHPRT1* antibody
- Evaluate overall genome integrity of mESCs post-delivery

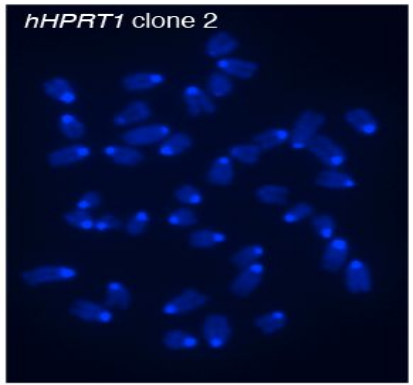
ICE:



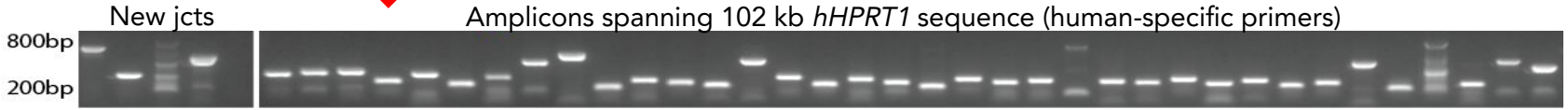
Western blot:



DAPI stained metaphase spread:

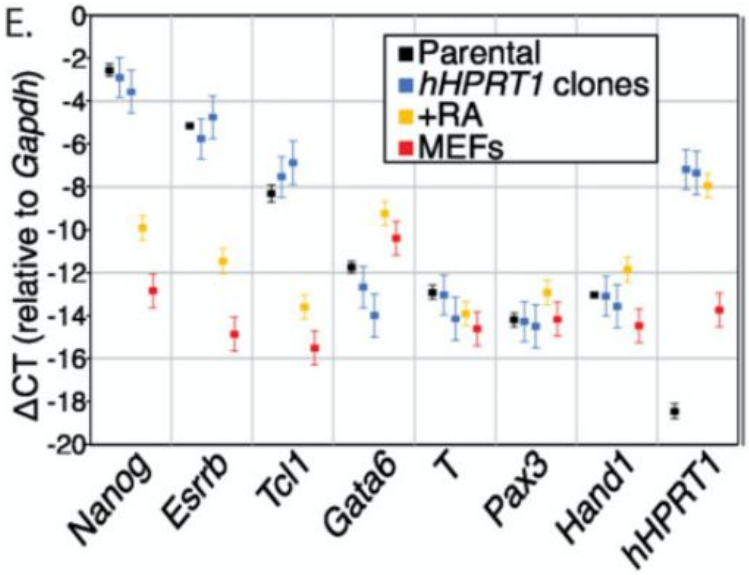


Genotyping PCR:



Transplanting the 102 kb human HPRT1 gene into mouse cells

- Modified mESCs retain stem cell potential



The Center for Synthetic Regulatory Genomics (SyRGe) at NYU Langone Health

neochromosome, inc.

The genome-scale cell engineering company

Division of opentrons™

Genomes are the basis for biotechnology products

Naturally-occurring genomes are evolved not designed.

Evolution has driven them to be big, repetitive, disorganized, and species-specific.

Genome writing technology means we don't have to start with extant genomes as substrates for biotech product development.

1

Viral genomes

- Gene therapy
- Gene delivery
- Vaccines
- Protein production
- Anti-infectives

2

Microbial genomes

- Food & feed
- Biomanufacturing
  - Biologics
  - Biofuels
  - Small molecules
- Agriculture
- Therapeutics

3

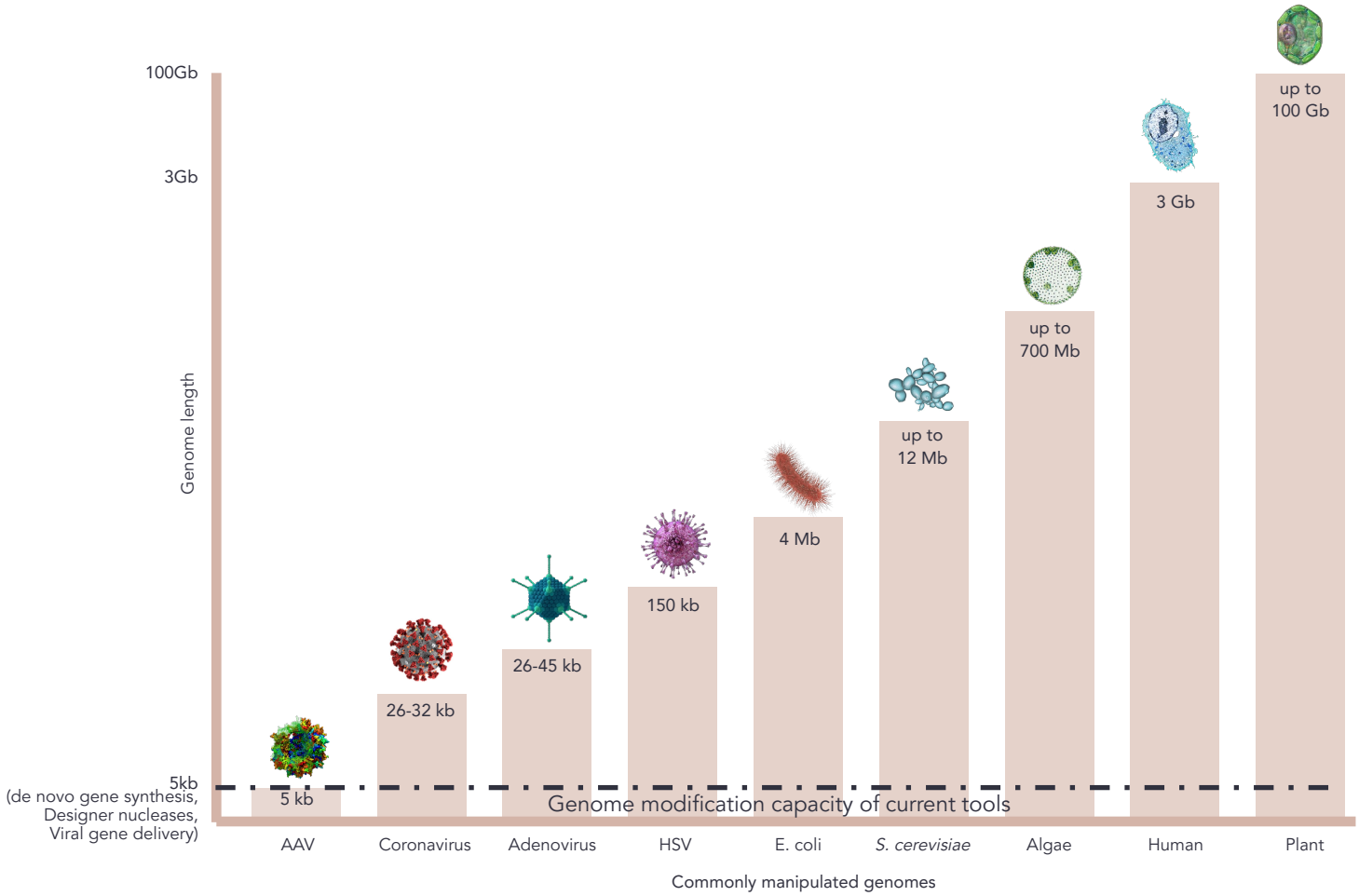
Mammalian genomes

- Cell therapies
- Regenerative medicines
- Biologics
- Animal models
- Drug discovery

PROBLEM

Most genome lengths far exceed current tools' capacity

Commonly used tools are limited to ~5 kb genome modifications, representing a very tiny fraction of the length of most genomes



Common genome engineering tools target ~5kb

The state-of-the-art in biotech is to use a handful of genome engineering tools in the design-build-test-deliver cycle to develop new products.

These tools propel an engineering cycle limited to 5 kb and reinforce the ubiquitous mindset of gene-at-a-time editing.

Designer nucleases

Base pair changes and gene deletions in existing genomes

Synthetic DNA

Max 5kb fragments readily sourced commercially

Viral gene delivery

Limited to 3-5kb payload capacity

Neochromosome

Unlocking next-generation products through genome-scale cell engineering

Ecosystem

Support & grow the genome-scale cell engineering ecosystem

Cell platforms

Build novel cell platforms for biomanufacturing in the cell and gene therapy space

Growing the ecosystem: unblocking long and complex DNA

Neochromosome is offering our capability to manufacture DNAs up to ~10kb long with virtually no sequence complexity restrictions as a service

info@neochromosome.com

Building novel cell platforms

Neochromosome is building cellular technology to manufacture new-to-nature proteins

info@neochromosome.com

We are a team of ~40 and growing, based in NYC

PARENT COMPANY

Neo is a wholly-owned subsidiary of Opentrons Labworks

Using biology to solve humanity's grand challenges

- Overall business is highly profitable
- Over $200M raised, 600+ employees globally
- Three independent business units with game-changing technologies

Robotics

neo

Cell Engineering

PRL

Lab Diagnostics

www.opentrons.com/jobs

All the contributors

Thank you!

Sc2.0

Jef Boeke
Joel Bader
Weimin Zhang
Yu Zhao
…

Tech Dev

Jef Boeke
Matt Maurano
Ran Brosh
Jon Laurent
Sudarshan Pinglay
…

Neochromosome

Jef Boeke
Joel Bader
Jon Brennan-Badal
Neta Agmon
Megan Hogan
John Anderson
Alfie Umbhau
Albert Serra
Mike Hammerling
Steve Riedmuller
…