# Applied mutagenesis: pathway engineering and synthetic biology

Increase biological production of useful molecules

- Random screening for overproducing strains (genome shuffling)
- Rational engineering of pathways and organisms

## Guide to readings

1) Metabolic engineering short review 2010.

2) Original genome shuffle. A paper reporting a recombination approach for improvement of antibiotic production

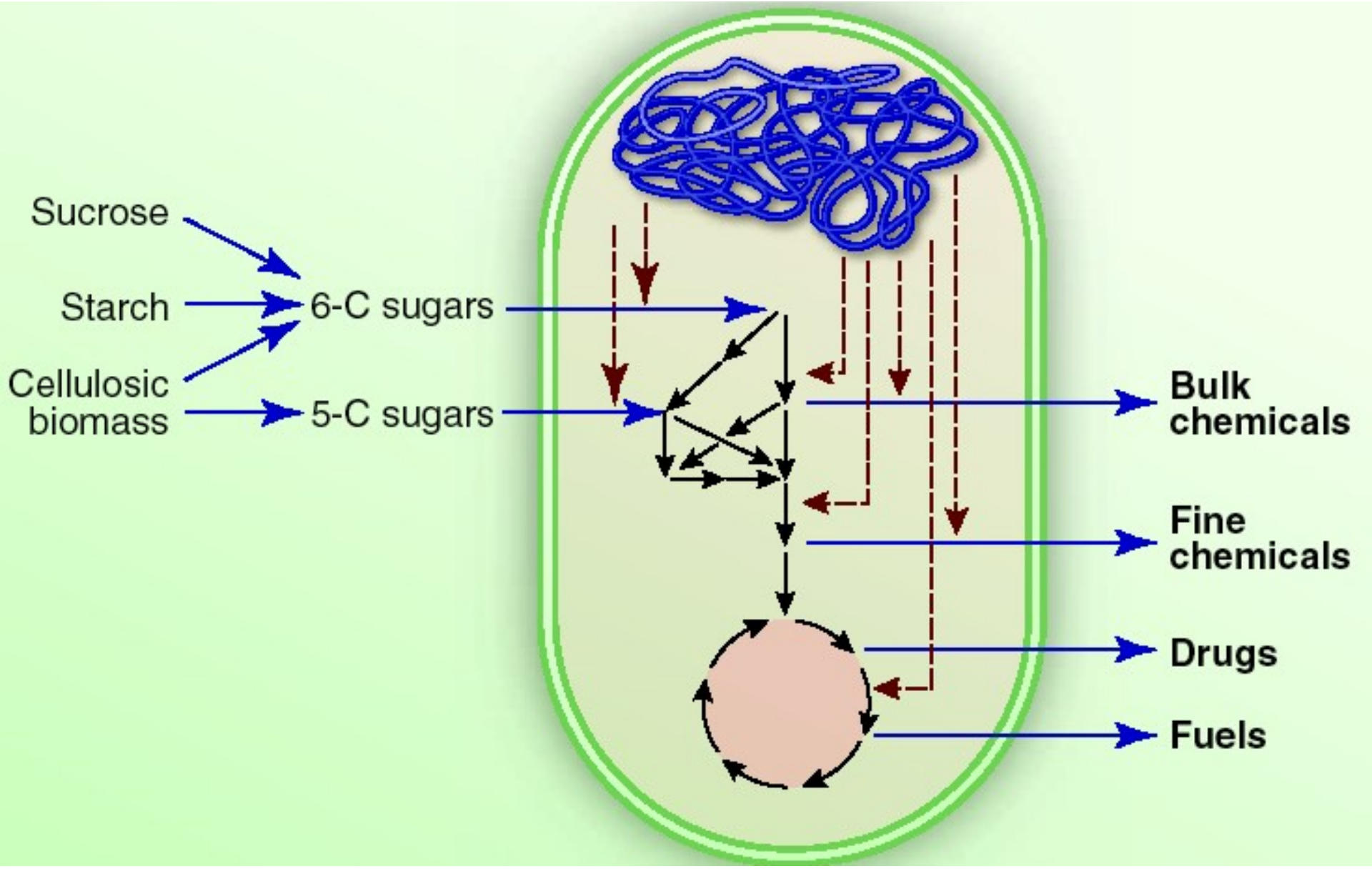
3) MAGE 2009. A paper describing a method to rapidly introduce genetic variation to targeted regions of the genome

4) Synthetic artemisenin. A paper describing metabolic engineering of *S. cerevisiae* to produce the antimalarial compound artemisinin.

5) Implications of synthetic genomes (2010). Perspective on the first completely synthesized bacterial genome.

6) Biocontainment of genetically modified bacteria (2015)

# The cell as a bioreactor




INPUT → PROCESS → OUTPUT

# Biologically derived molecules for sale:

- Antibiotics
- Vitamins
- Metabolic by-products (ethanol, lactic acid)
- Amino acids and derivatives (indigo, aspartame)
- “ secondary metabolites” from plants – e.g. alkaloids (caffeine, theobromine, etc.)
- Hydrocarbons for fuel

- Synthesis often requires multiple steps and enzymes, making these molecules difficult to synthesize chemically

# Increased production of antibiotics: Classical Strain Improvement

1) Obtain organism that produces the compound of interest -- the original strain of *Penicillium* mold made penicillin at micrograms per liter of culture

2) Random mutagenesis followed by screen for increased production.

3) With top producer, repeat mutagenesis and screen

4) Outcome: milligrams of penicillin per liter of culture (1000-fold increase in production)

- Time consuming and expensive process!

# Genome shuffling: an alternative to Classical Strain Improvement

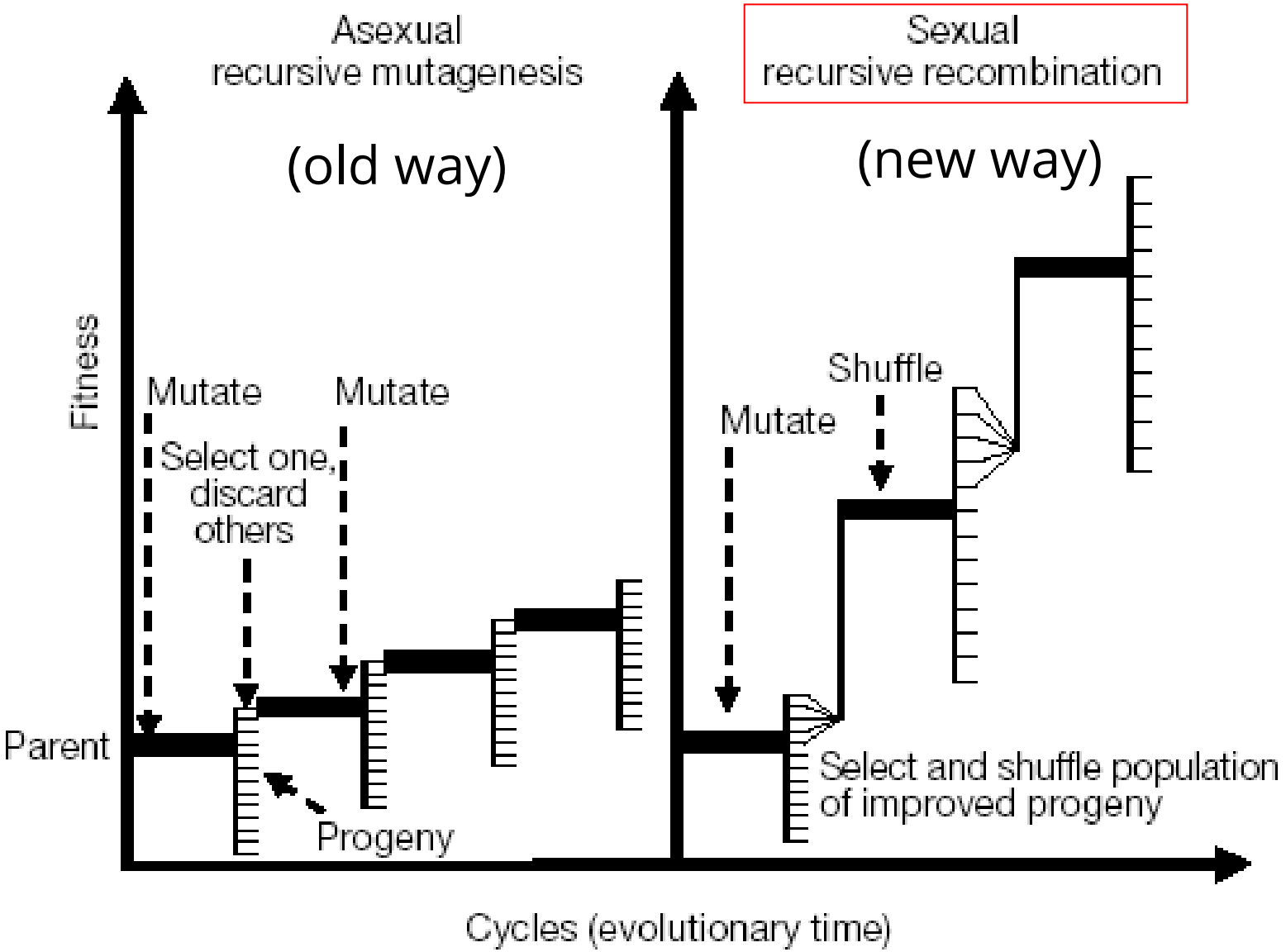
## Genome shuffling leads to rapid phenotypic improvement in bacteria

**Ying-Xin Zhang*, Kim Perry*, Victor A. Vinci†, Keith Powell*, Willem P. C. Stemmer* & Stephen B. del Cardayré***

* *Maxygen, 515 Galveston Drive, Redwood City, California 94063, USA*
† *Eli Lilly and Company, Lilly Corporate Center, Indianapolis, Indiana 46285, USA*

NATURE | VOL 415 | 7 FEBRUARY 2002 | www.nature.com

# The shuffling advantage: simultaneous recombination of entire genomes (breeding) with multiple parents



# Testing recursive shuffling

- Compare classical strain improvement (CSI) to genome shuffling

- *Streptomyces sp.*: produce polyketide antibiotics

- Induce recombination by recursive protoplast fusion:
  - Fuse protoplasts
  - Regenerate cell walls, grow as a population (F1)
  - Make protoplasts with F1, repeat until F4

- Test with 4 auxotrophy markers (next page)

- Test for increased antibiotic production

# Test of recursive shuffling

4 parental strains

| Strain | Genotype | Supplements required: |
|---|---|---|
| *S. coelicolor* 2684 | *proA1 argA1 cys$^{wt}$ uraA1* | pro, arg, ura (not cys) |
| *S. coelicolor* 2685 | *proA1 arg$^{wt}$ cysD18 uraA1* | pro, cys, ura (not arg) |
| *S. coelicolor* 2686 | *pro$^{wt}$ argA1 cysD18 uraA1* | arg, cys, ura (not pro) |
| *S. coelicolor* M124 | *proA1 argA1 cysD18 ura$^{wt}$* | pro, arg, cys (not ura) |

" Shuffle" (recombine) all 4 strains

Can progeny be isolated that can grow without pro, arg, ura, and cys supplementation (indicating progeny with all 4 genes wild type)?

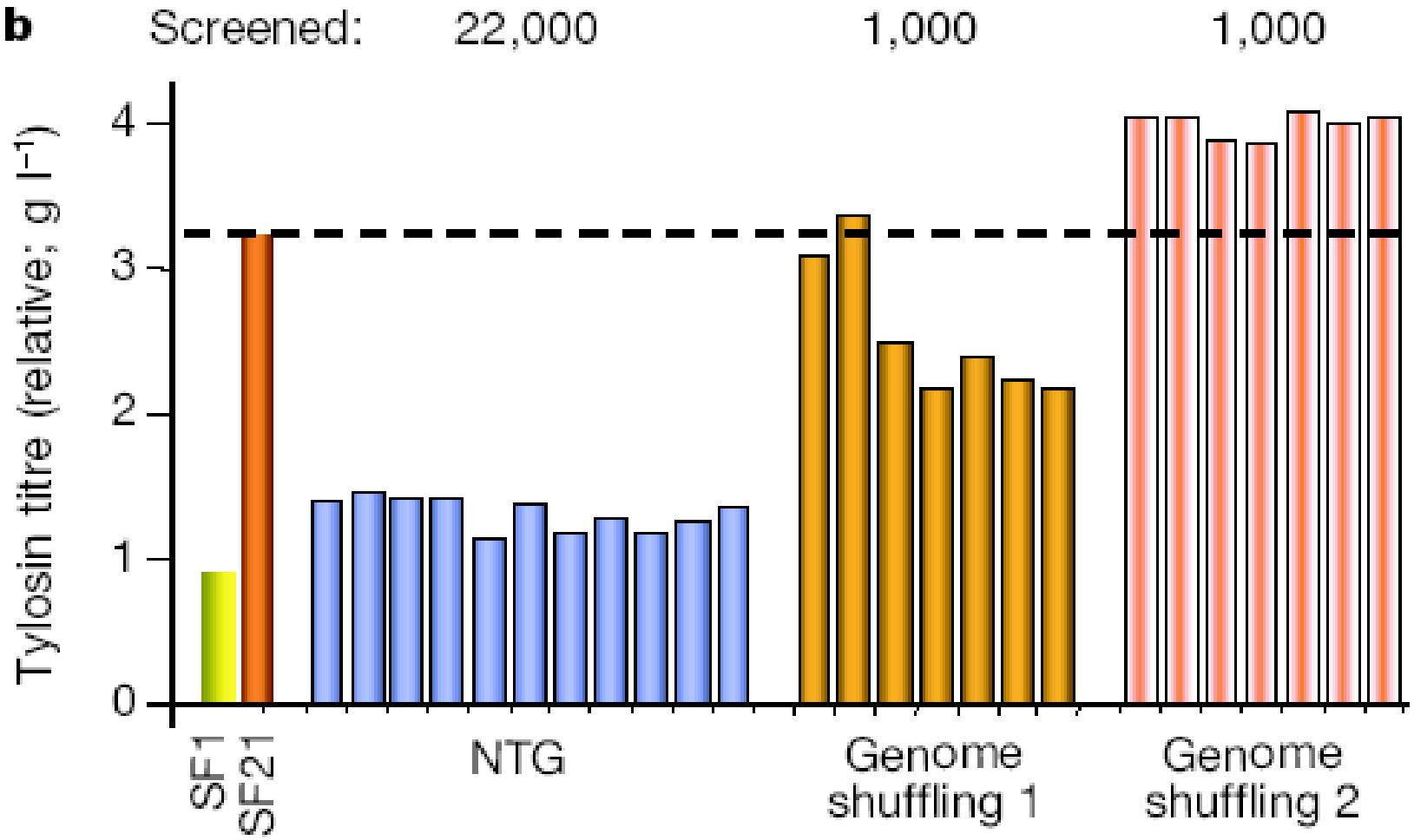
Shuffling: increased efficiency of recombinati

**Table 1 Distribution of phenotypes in a four-strain cross of *S. coelicolor***

| Phenotype | Single fusion* | Recursive fusion† |
|---|---|---|
| Two markers | 8.4% | 60% |
| Three markers | 0.73% | 17% |
| Four markers | 0.000045% | 2.5% |

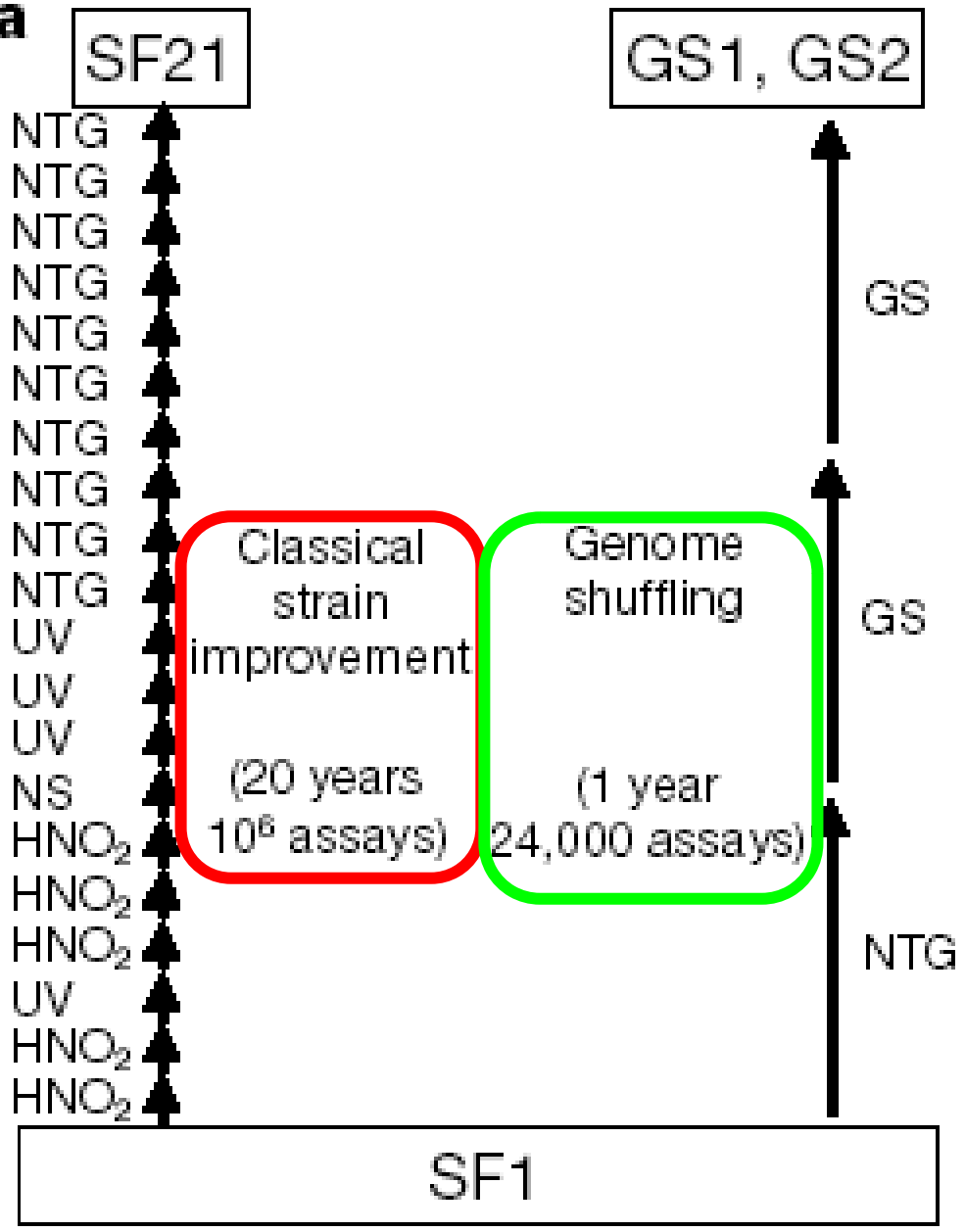
The distribution of phenotypes from each fusion is reported in Supplementary Information.
* Phenotypes were determined from colony counts on defined medium containing 16 combinations of the four supplements (See Methods). Each phenotype is corrected for dilution and the presence of prototrophic markers, and divided by the total colonies growing on completely supplemented medium. The value shown represents the sum of the frequencies from each phenotypic class.
† The distribution of 483 individual colonies characterized for marker phenotype.

# Test case: increase tylosin production by *S. fradiae*




SF1 was treated with NTG, 11 strains selected (22,000 screened), those 11 strains were shuffled once (GS1) and then again (GS2)

# Comparison: CSI versus genome shuffling

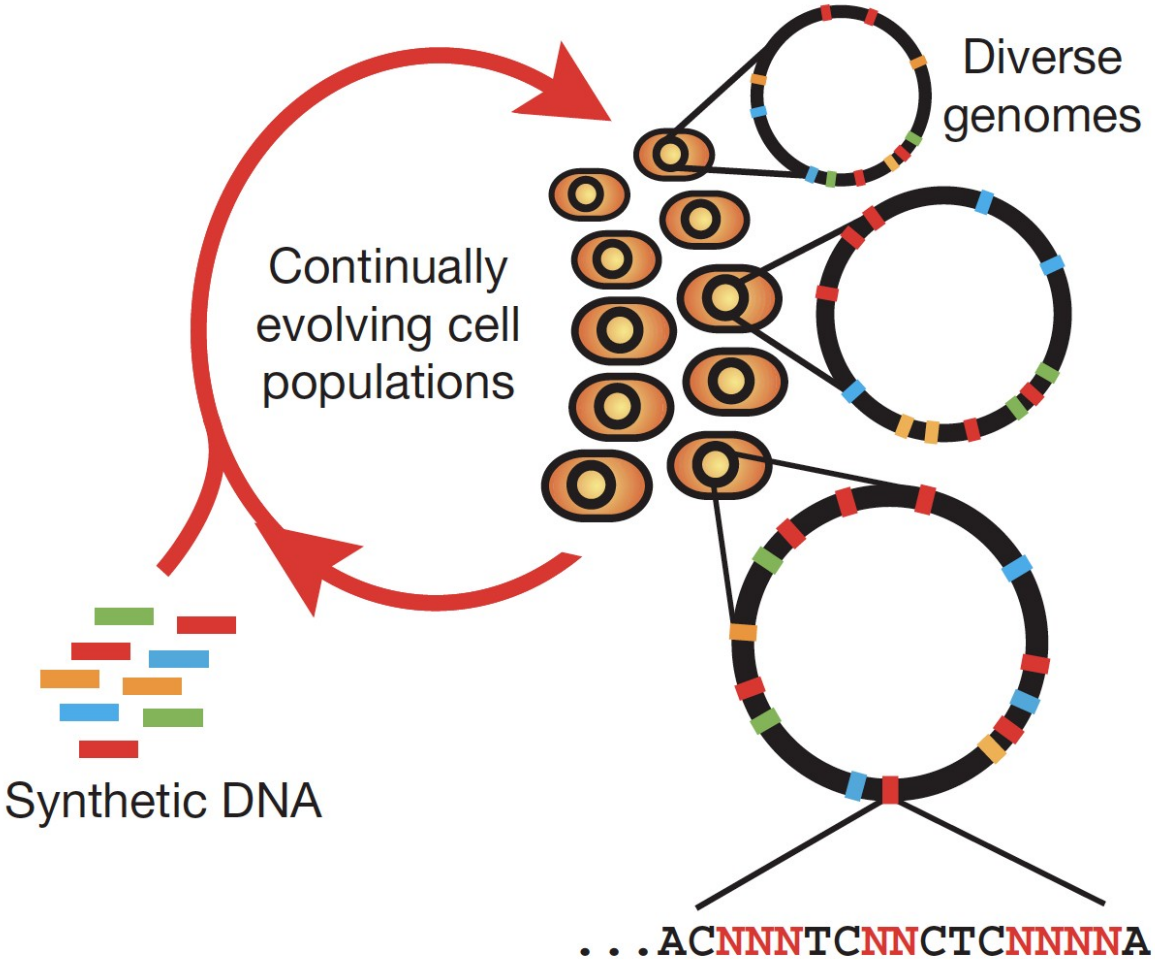



**c**

| Strain | Titre (rel. g $l^{-1}$) |
|---|---|
| SF1 | 1.0 ± 0.1 |
| SF21 | 6.2 ± 2.4 |
| GS1 | 8.1 ± 1.2 |
| GS2 | 6.2 ± 1.2 |

Similar results &
much faster with GS

# Targeted genome evolution by MAGE (Multiplex Automated Genome Engineering)




nature
Vol 460|13 August 2009|doi:10.1038/nature08187

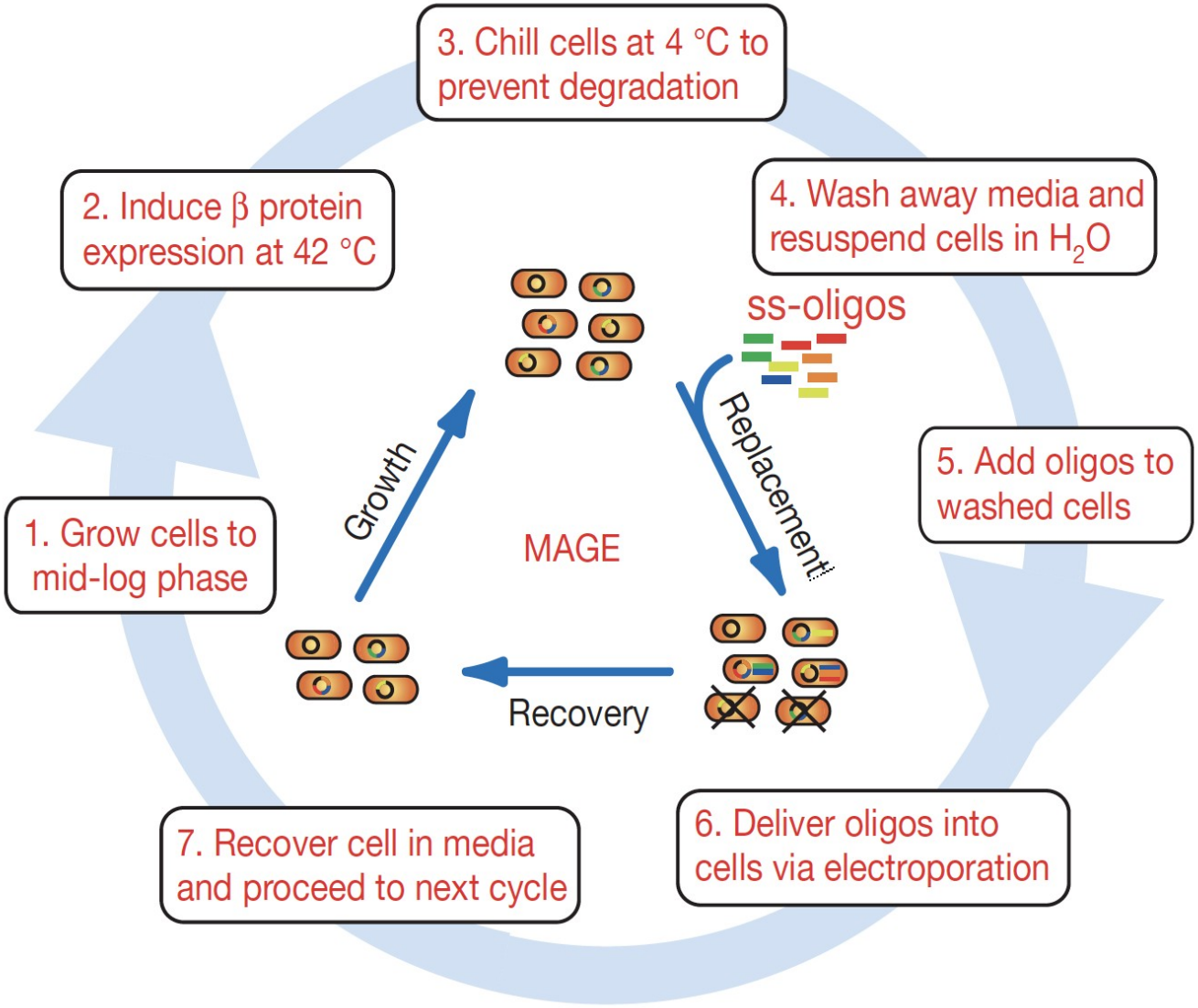
LETTERS

**Programming cells by multiplex genome engineering and accelerated evolution**

Harris H. Wang[1,2,3]*, Farren J. Isaacs[1]*, Peter A. Carr[4,5], Zachary Z. Sun[6], George Xu[6], Craig R. Forest[7] & George M. Church[1]

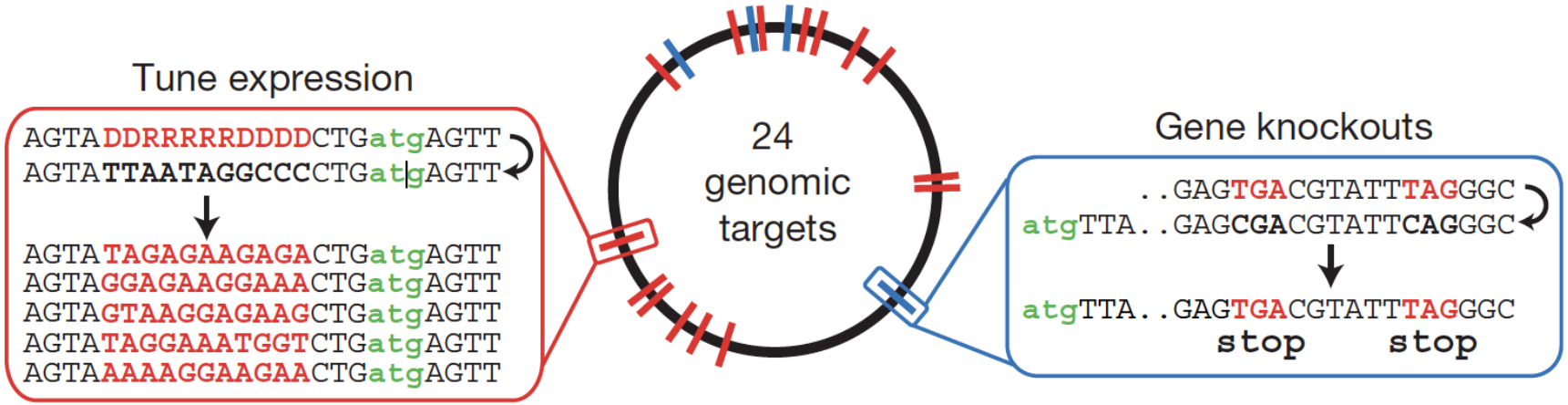
Cells are repeatedly transformed with synthetic oligonucleotides that recombine with genome and increase sequence diversity

# One automated cycle (takes ~2.5 hours )



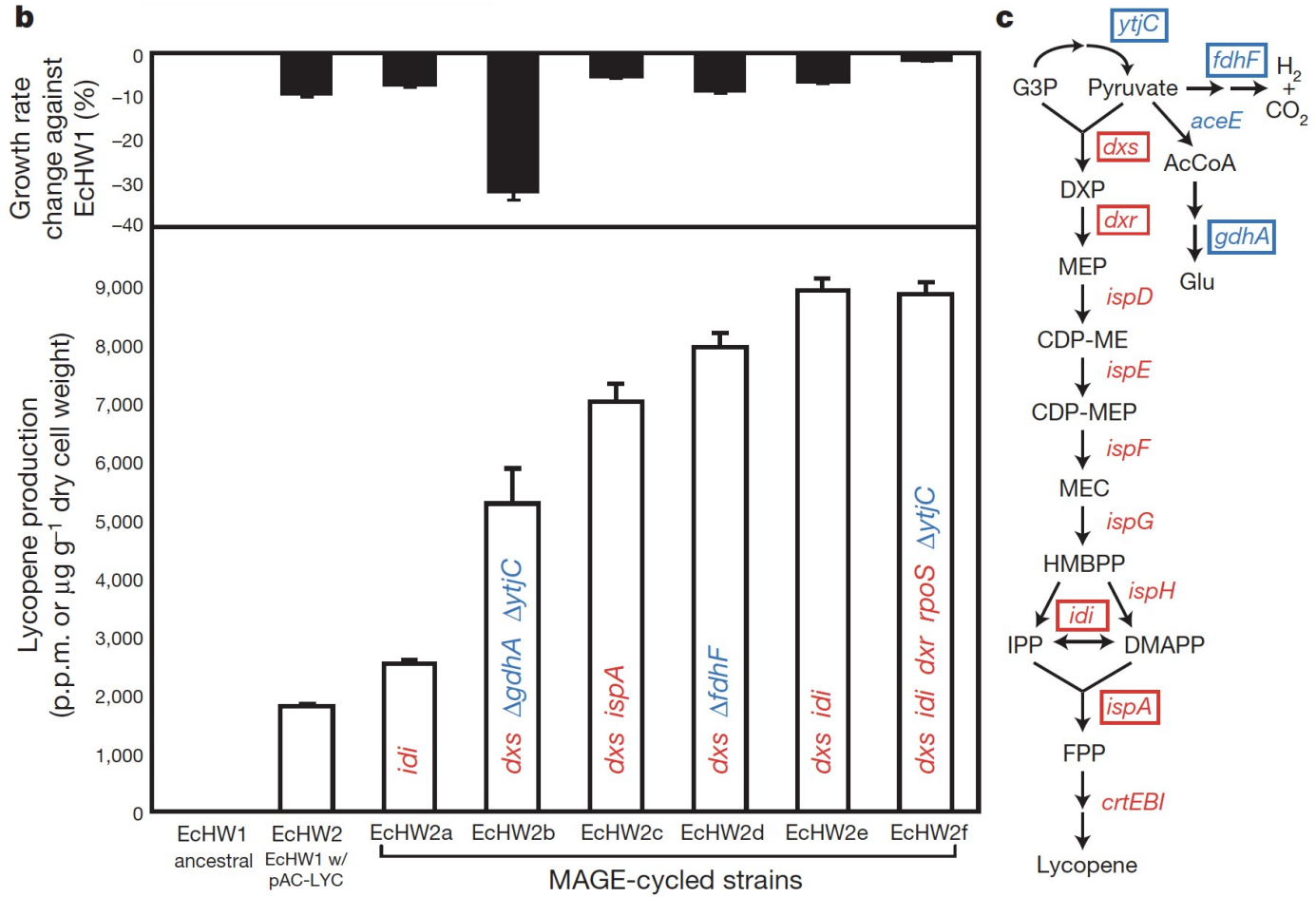

After 5 cycles, measured average of 3.1 base changes per cell

# Test of method: lycopene production




- 20 genes in *E. coli* known to enhance lycopene production, and 4 genes known to divert resources away from the lycopene pathway
- DNA was added to randomize the 20 genes' ribosome binding sites, as well as to knock out the diverting genes
- MAGE cycling was done and colonies were selected for analysis on the basis of increased red color (lycopene)

# Optimized translation of a subset of the genes in pathway, and knock out of 3 of the 4 diverting genes, provided up to 5X increases in lycopene production



# Increasing production of a biological compound: rational design

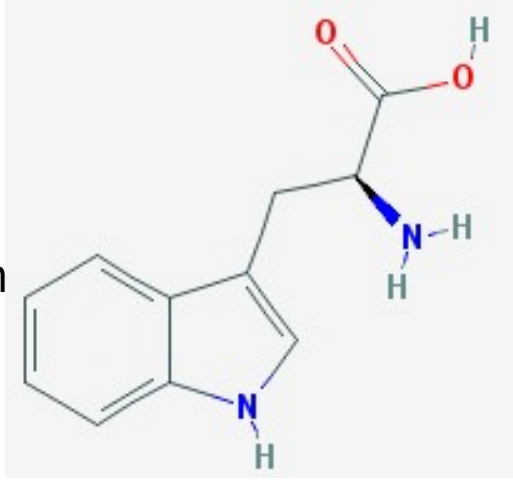
1) Increase production of a naturally produced commercial compound
   - Modify existing genes

2) Obtain a new organism that can convert an existing compound into a commercial compound
   - Introduce  new genes
   - Modify existing genes

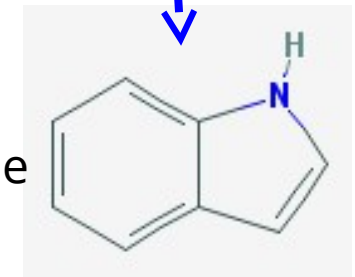
# natural source of indigo: woad [*Isatis tinctoria*]

woad

tryptophan

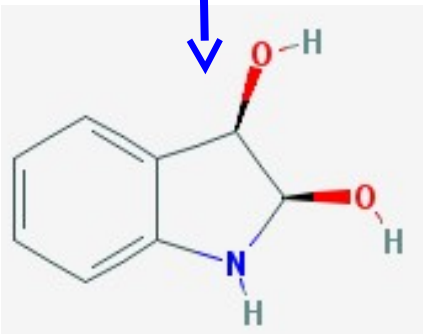


Tryptophanase (*E. coli*)



indole

Napthalene dioxygenase (cloned)

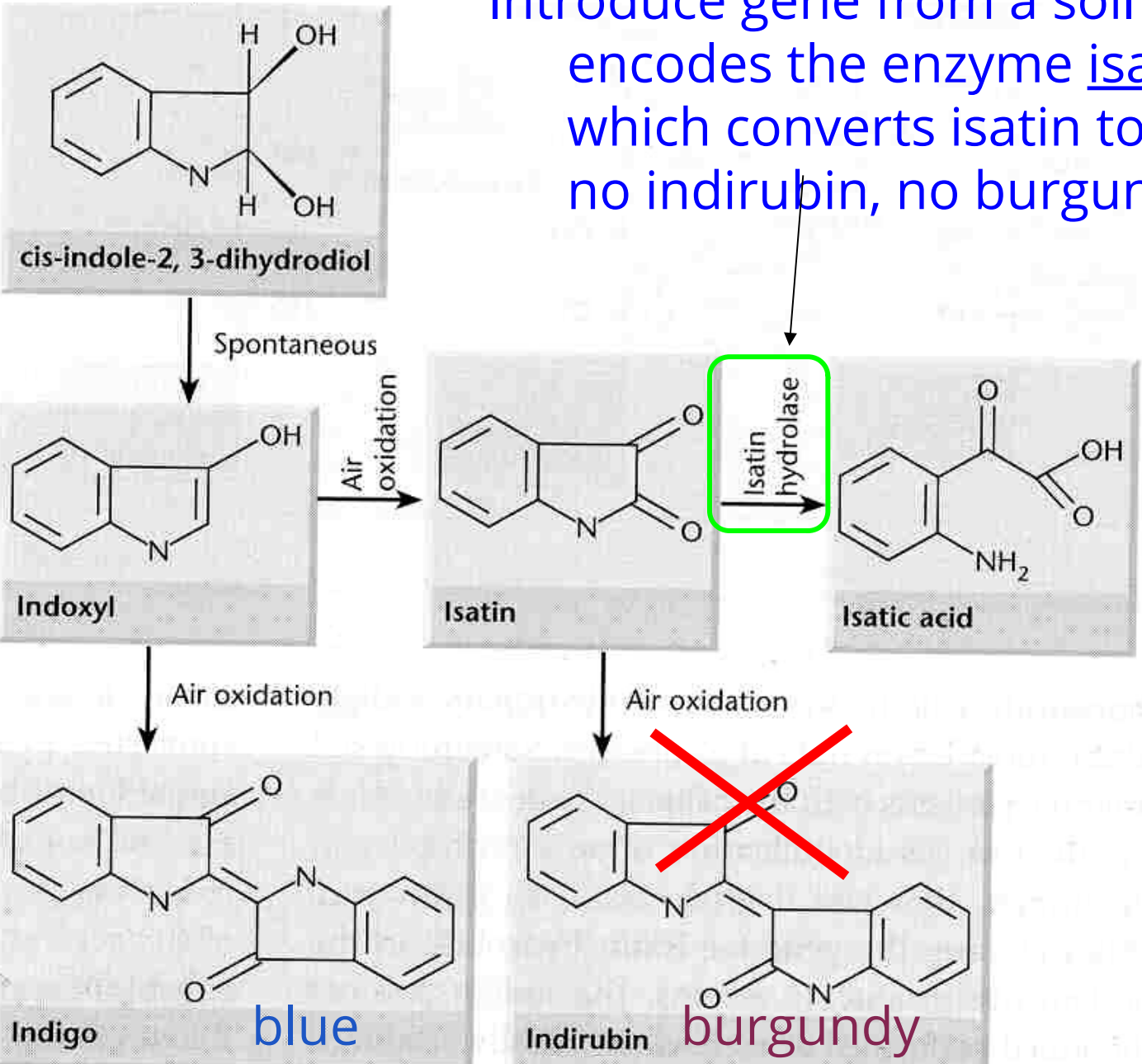


cis-indole-2,3-dihydrodiol

## Engineering *E. coli* to produce indigo

- Mutate tryptophan synthase complex to release indole
- Introduce napthalene dioxygenase (from *Pseudomonas putida*)

Introduce gene from a soil microbe that encodes the enzyme isatin hydrolase which converts isatin to isatic acid (so no indirubin, no burgundy color)




# Metabolic engineering: genetic manipulations to maximize synthesis of biological molecules

- Remove transcription, translation repressors
- Defeat enzyme feedback control
- Speed up rate-limiting enzymes
- Block competing pathways
- Funnel carbon to the pathway of interest
- Increase transport of compound out of cell

# How to overproduce phenylalanine in *E. coli*

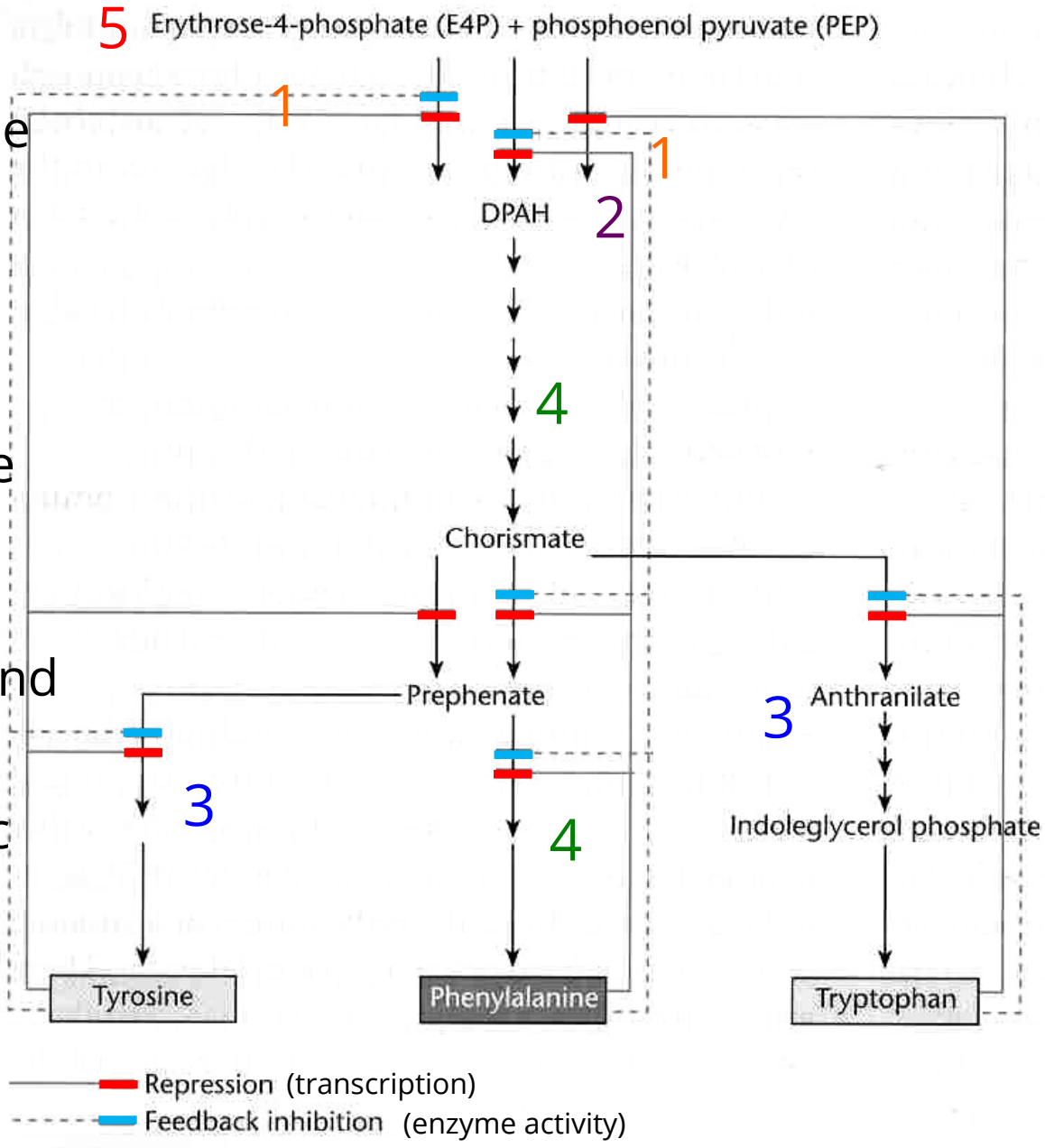
1) Remove feedback inhibition (select strains resistant to phenylalanine analogue feedback inhibitors)

2) Avoid transcriptional repression (place genes under control of non-phe controlled promoters)

3) Remove pathway competition (delete tyr and trp specific genes)

4) Overexpress phe-specific genes

5) Increase E4P and PEP synthesis



# *S. cerevisiae* engineered to produce artemisinic acid

- Artemisinin is a primary medicine for treating malaria, which caused 660,000 deaths worldwide in 2010

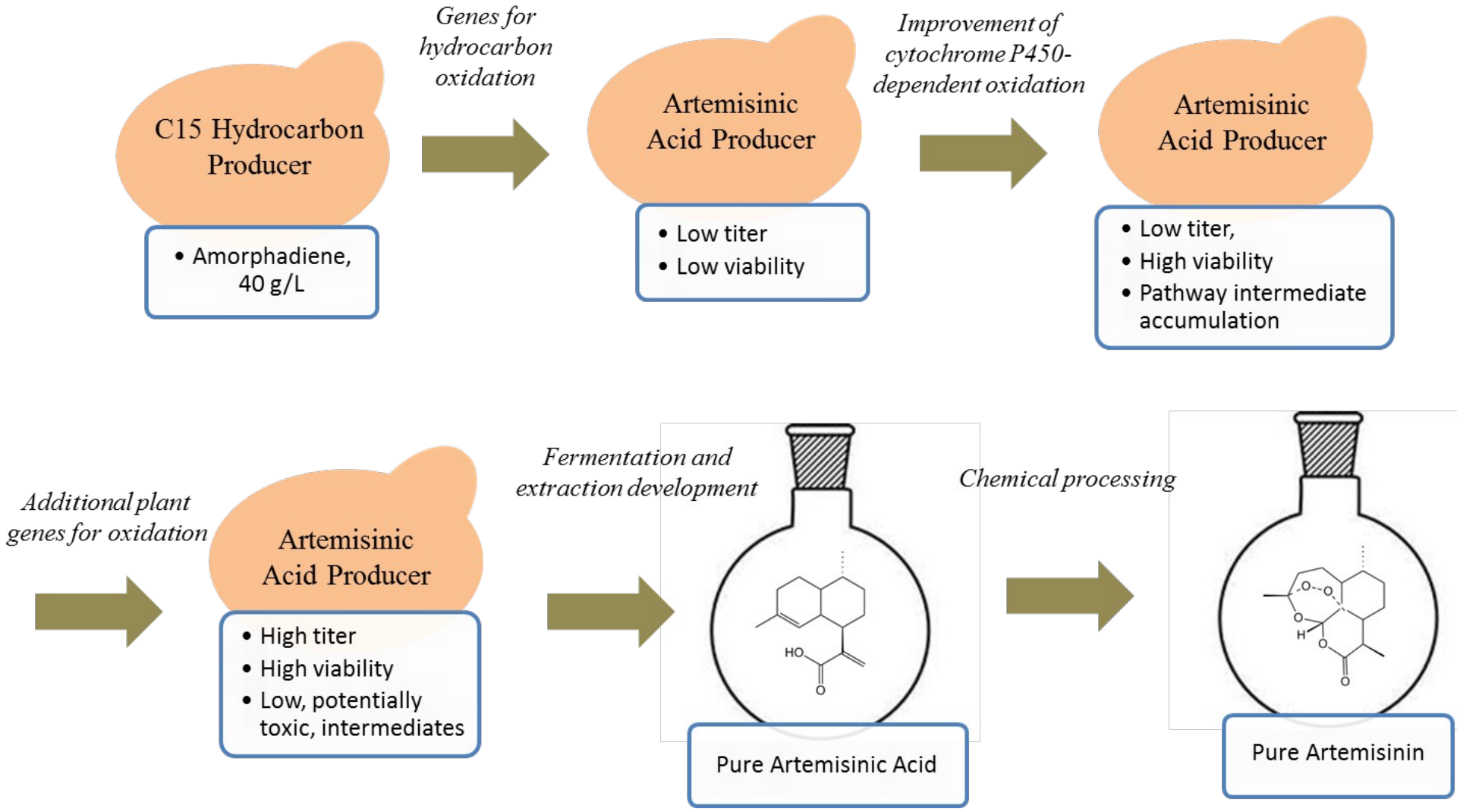
- Production: the plant *Artemisia annua*

- The supply line depends on robust crop yields, so off- years could cause drug shortages

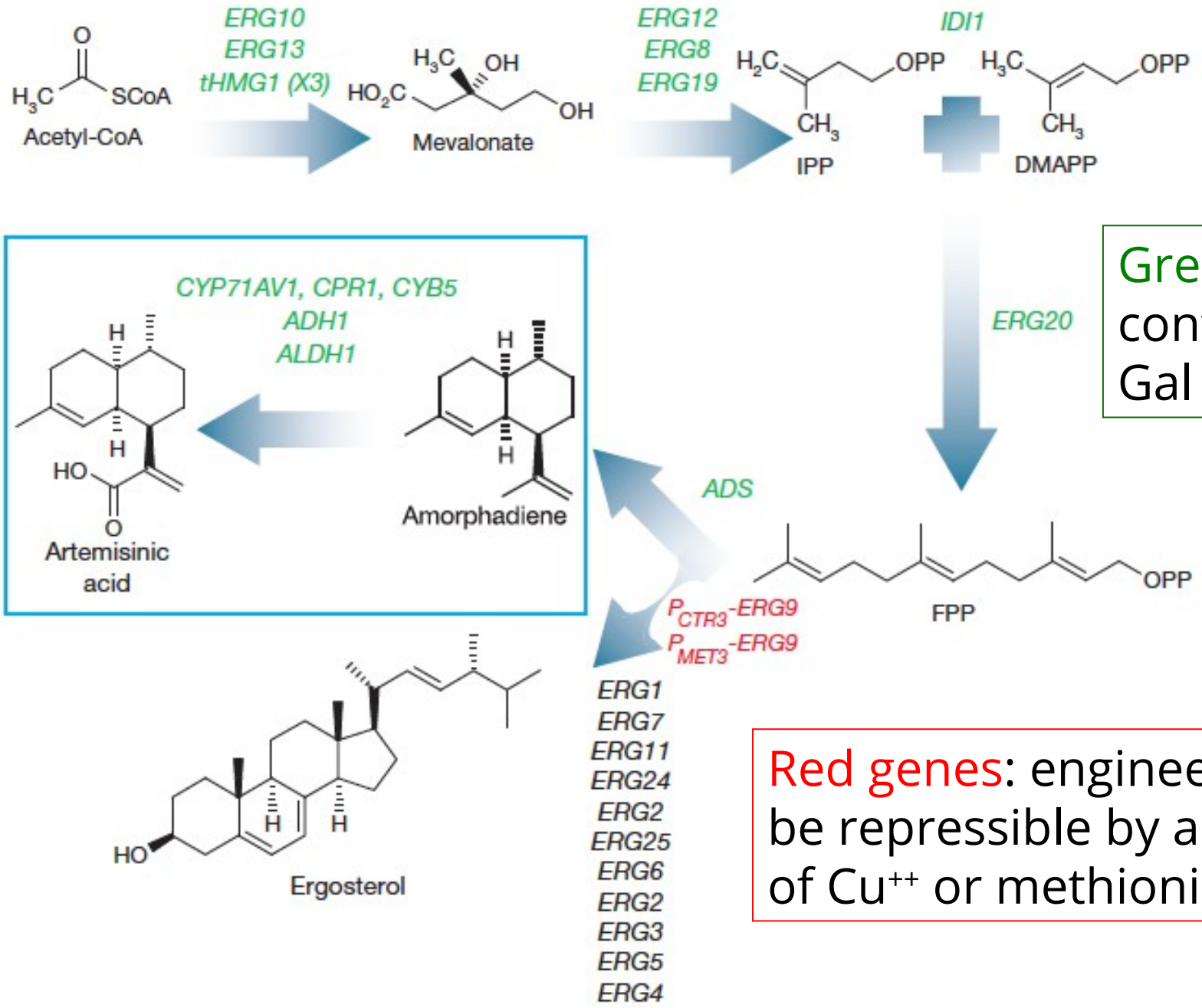
- The “semi-synthetic” production of artemisinin through engineering of yeast was recently reported

- Expected production of artemisinin by this approach: 50-60 tons/year, or 80-150 million doses, produced relatively cheaply

# Synthesis of the antimalarial drug artemisenin (originally produced only by the wormwood plant, *Artemisia annua*)

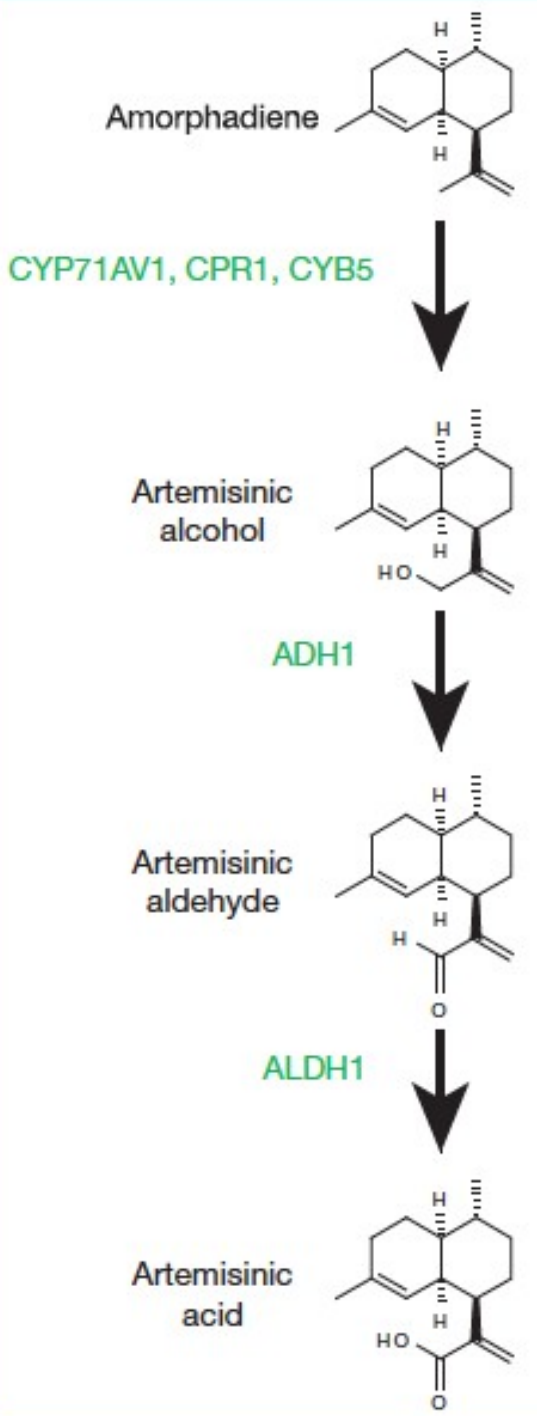


# Pathway for synthesis of artemisinic acid




Green genes: controlled by Gal induction

Red genes: engineered to be repressible by addition of $Cu^{++}$ or methionine




Genes isolated and cloned from *Artemisia annua* were the key

However: farmers who grow *Artemisia* (in Asia and Africa) could lose buyers – can they adapt?

Other potentially disruptive synthetic biology projects on the horizon:

- Vanilla
- Vetiver
- Patchouli
- Rubber
- Coconut
- Saffron
- Opioids

http://www.guardian.co.uk/global-development/poverty-matters/2013/apr/12/synthetic-malaria-compound-artemisia-farmers

# Rational metabolic engineering

- Requires at least some knowledge of the biochemical pathway required for compound synthesis
- High degree of control
- Trial and error approach can be time consuming
  - try something
  - see if it works
  - find out where the new block to production is
  - change it, too
  - and so on…

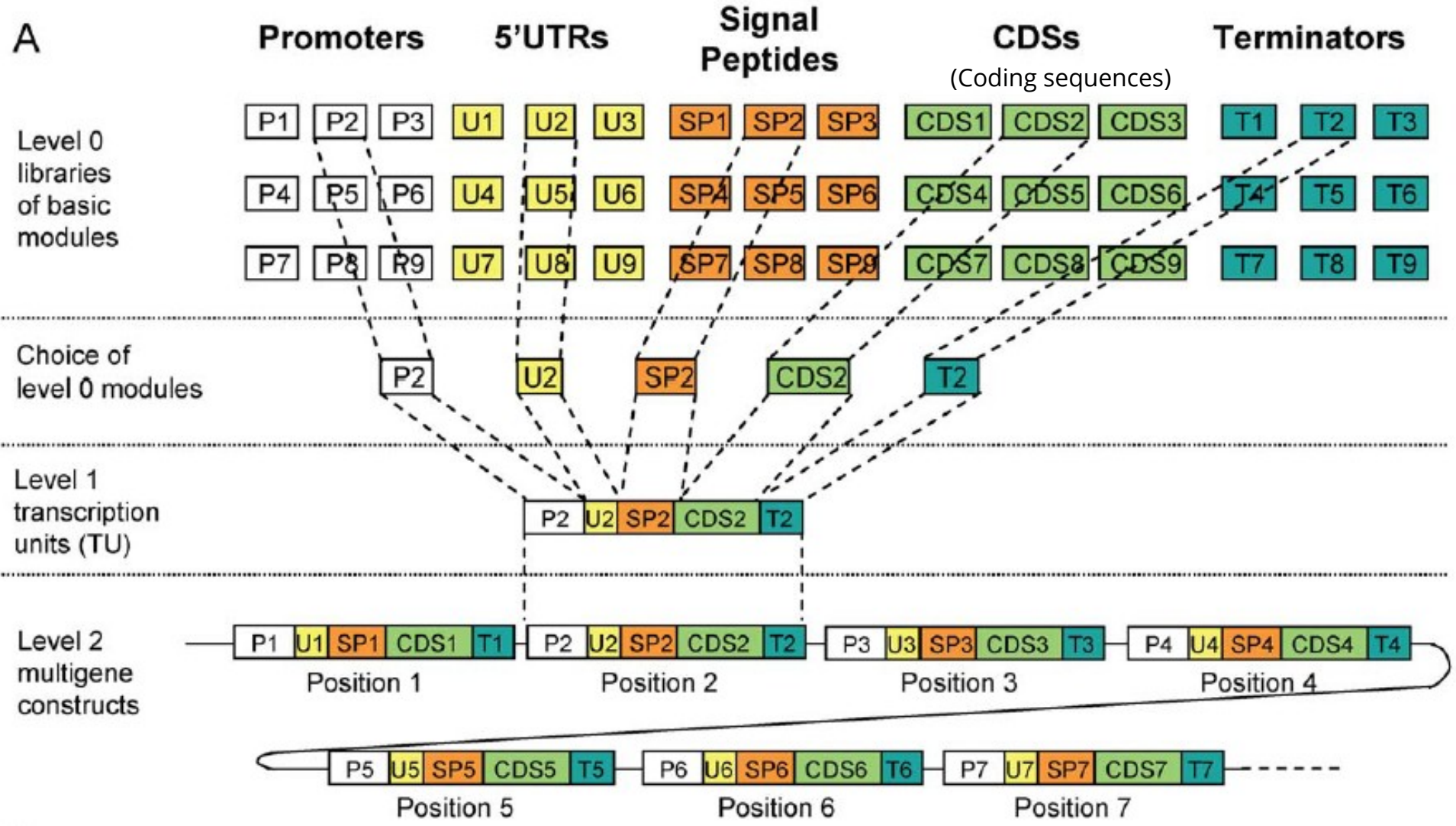
## Synthetic biology:

Construction of cellular systems from component parts to reprogram an organism, or to create a new organism

- engineering principles are followed
- synthesis of modules, followed by ordered assembly
- complete, synthetic genome construction is possible

However: it can be difficult to predict how the modules will interact with each other or with the environment in a living system

# The biobrick approach: mix and match modules




## A Modular Cloning System for Standardized Assembly of Multigene Constructs

PLoS ONE | www.plosone.org

**Ernst Weber, Carola Engler, Ramona Gruetzner, Stefan Werner, Sylvestre Marillonnet***

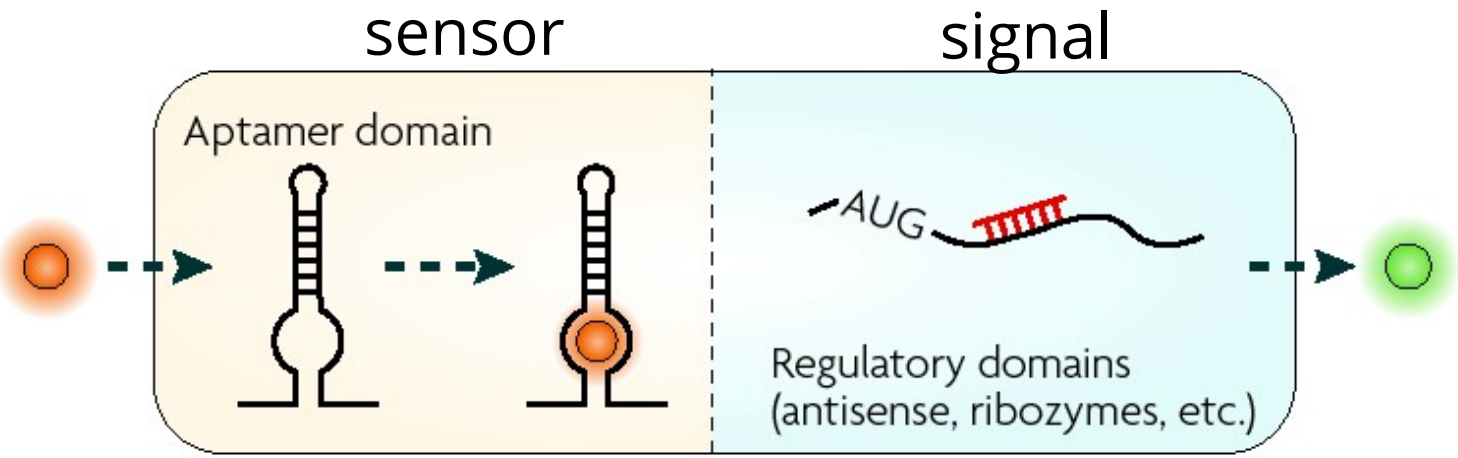
Icon Genetics GmbH, Halle/Saale, Germany

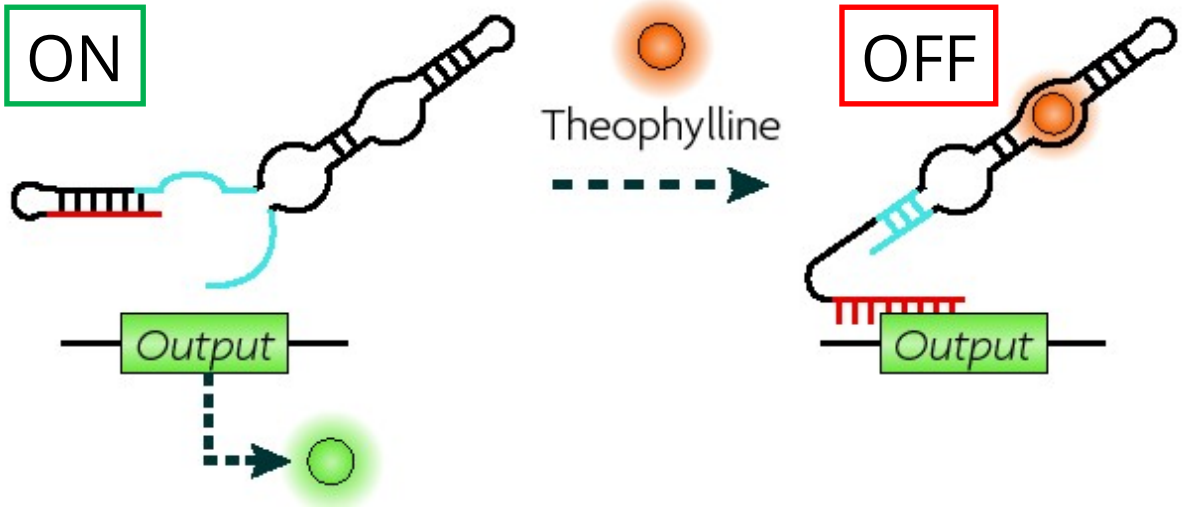
February 2011 | Volume 6 | Issue 2 | e16765

**Synthetic biology module:**

An engineered genetic element that performs a specific task

BIOSENSOR: a genetic control module. Presence of a small molecule shuts off signal

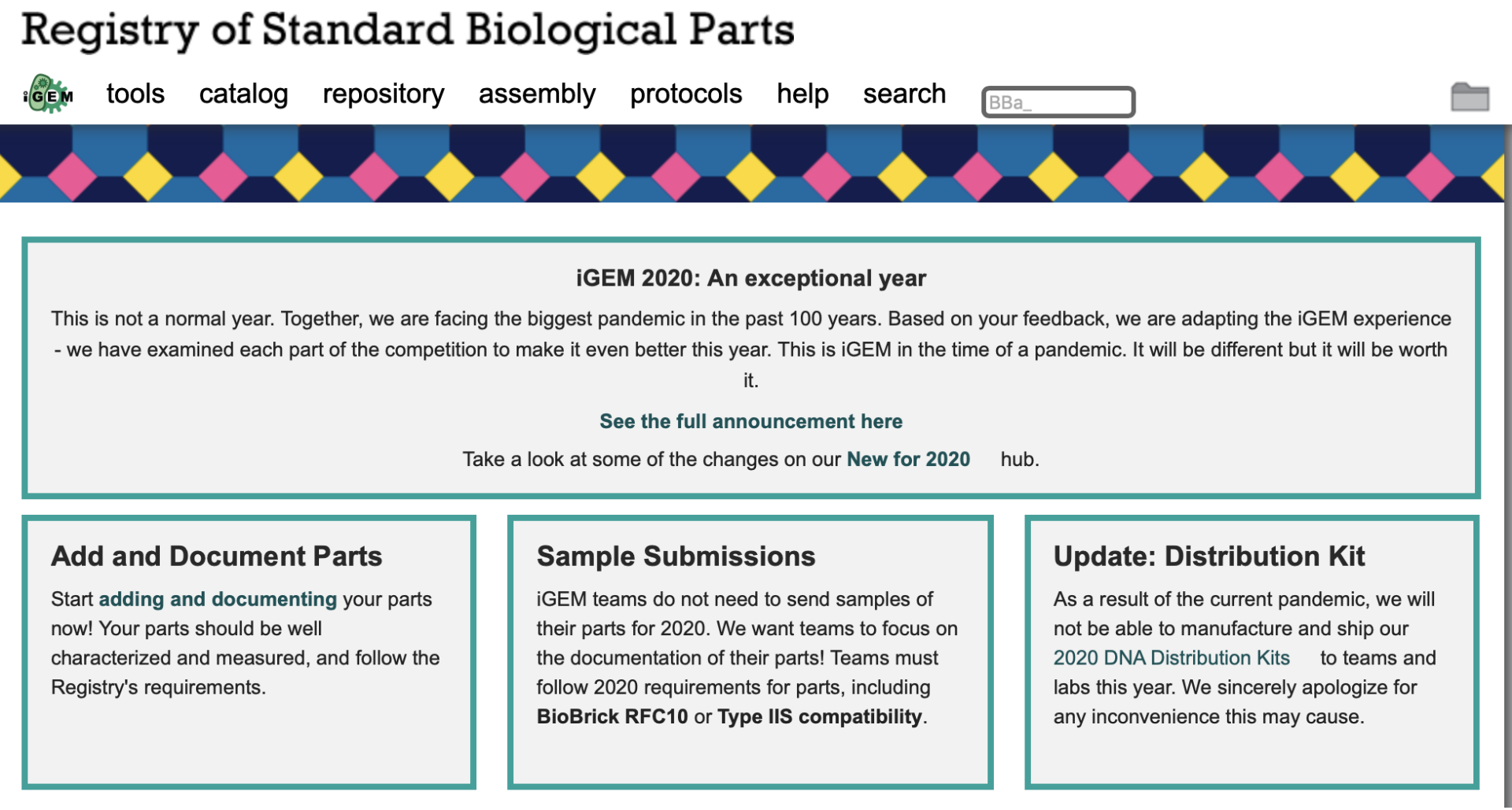



A state change in the RNA is induced by theophylline, which causes a gene to be turned off

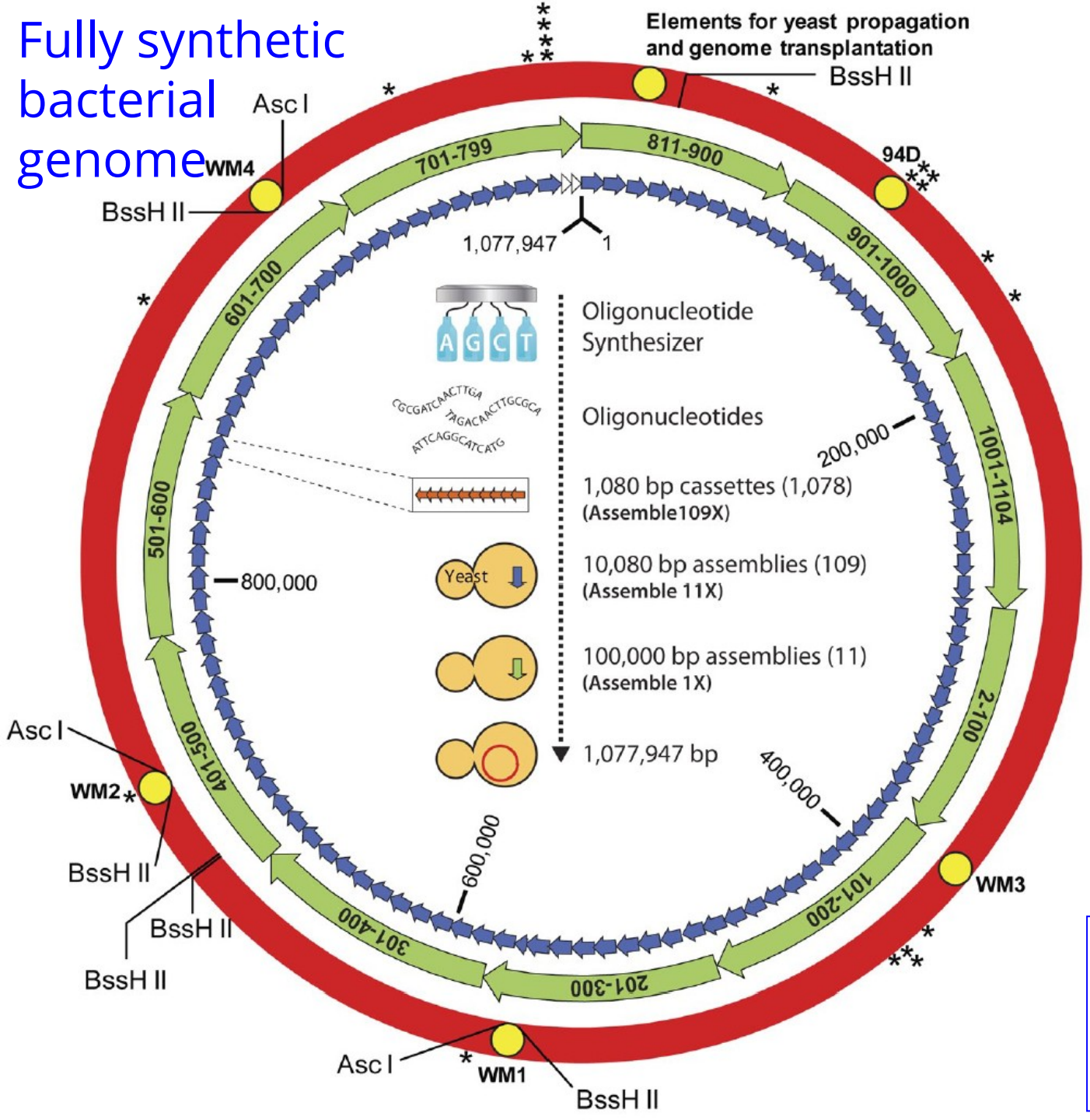
# Registry of biological parts:

http://parts.igem.org/Main_Page




iGEM: International Genetically Engineered Machine

# Fully synthetic bacterial genome



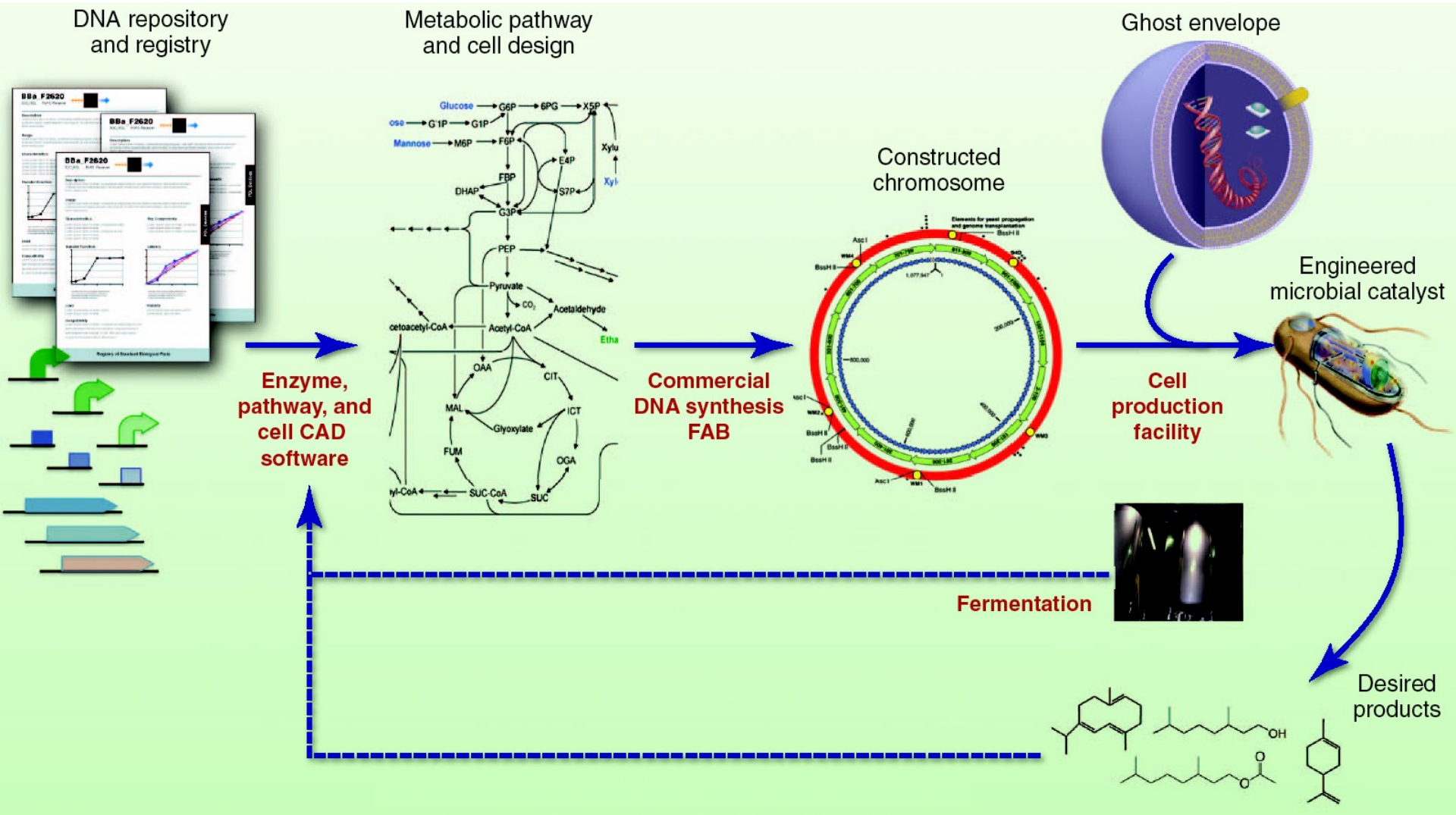

1080 bp synthetic cassettes assembled into sets (10 per set, 109 sets total) by recombination in yeast

10,800 bp fragments assembled into sets (10 per set, 11 sets total)

11 ~100 kb fragments assembled into complete genome

~ $40M
20 people
>10 years

# The future of bioengineering?



## Risks? How to manage them?

How do synthetic life forms interact with various potential environments?

How long do synthetic organisms persist, and how well do they compete with non-synthetic organisms?

How quickly can the synthetic organism evolve?

Can synthetic organisms transfer their DNA to natural organisms and vice versa?

“Once released, synthetic organisms cannot be retrieved”

# Genetic engineering for bio-containment

- Make the organism dependent on one or more non-canonical amino acids
  - Genetic code is re engineered
  - Requires feeding with non-canonical amino acid
  - Genetic code differs from other organisms: genetic exchange can't readily occur

- Make the organism dependent on added small molecules not found in nature
  - 'riboregulators' based on riboswitches controlling many essential genes in an organism
  - Addiction modules encode nucleases that kill the cell if the antidote isn't made, and the antidote is controlled by a small molecule

# Applied mutagenesis: pathway engineering and synthetic biology

Increase biological production of useful molecules

- Random screening for overproducing strains (genome shuffling)
- Rational engineering of pathways and organisms