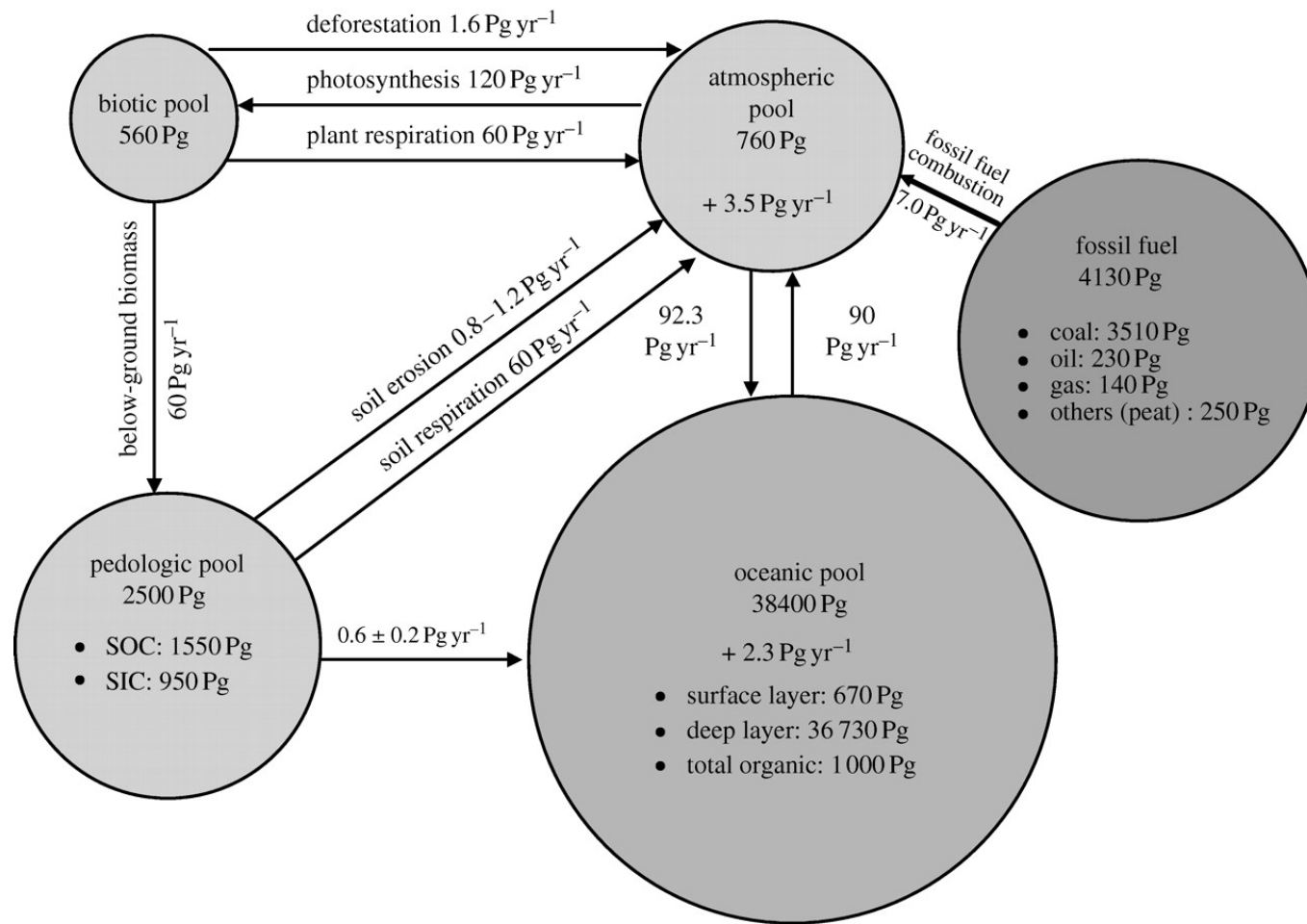


systems and synthetic biology
for biorefineries
to make sustainable
chemicals

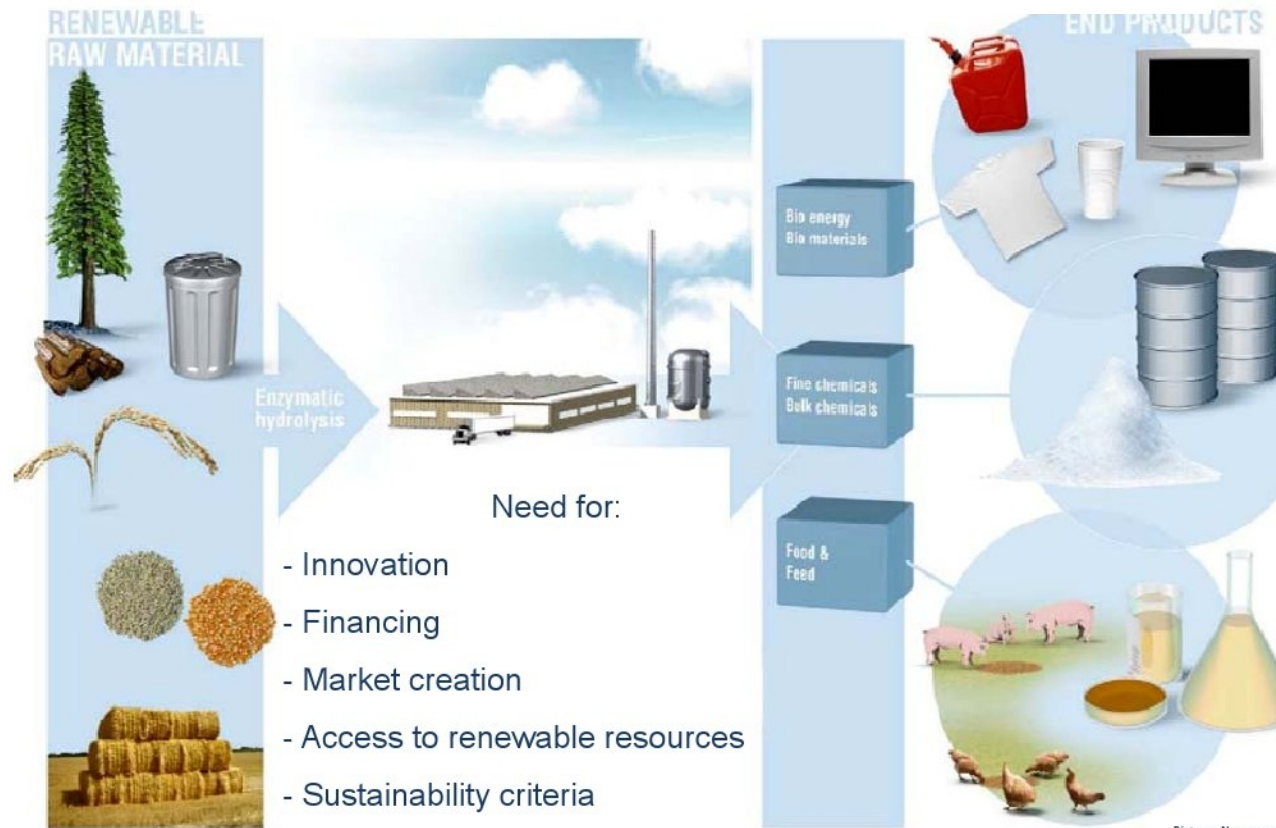
Andrew Tolonen
Genoscope
atolonen@gmail.com



a major scientific challenge is to find way to
balance the global carbon cycle

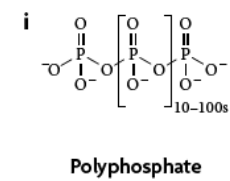
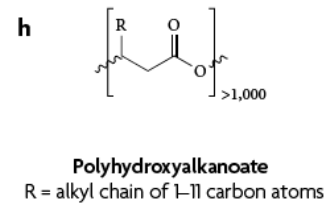
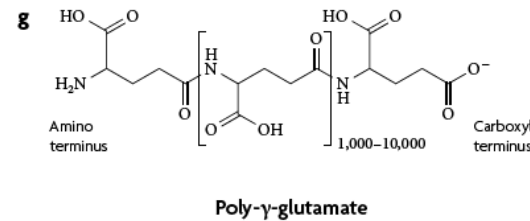
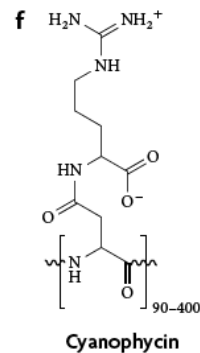
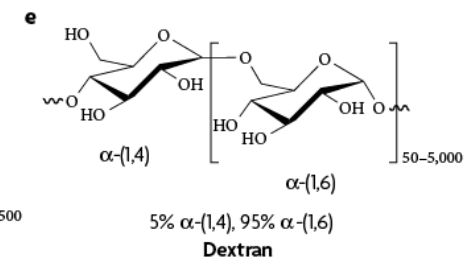
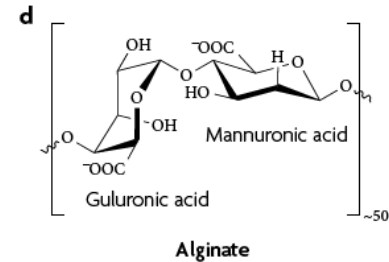
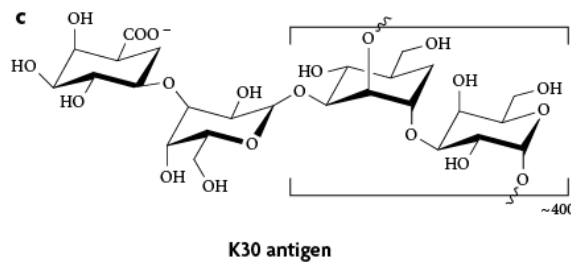
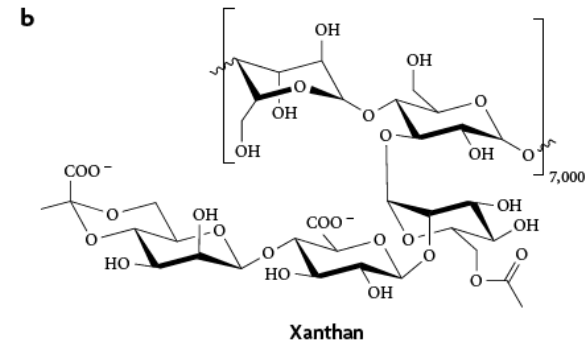
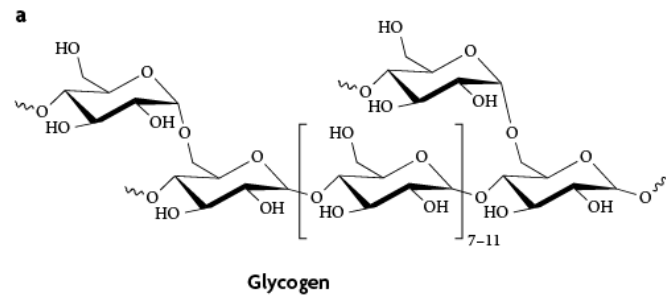


biorefinery: renewable, carbon balanced production of biocommodities



<http://www.abakus.be>

biorefineries will produce many useful polymers



Rehm 2010

examples of renewable biomaterials

poly-lactic
acid



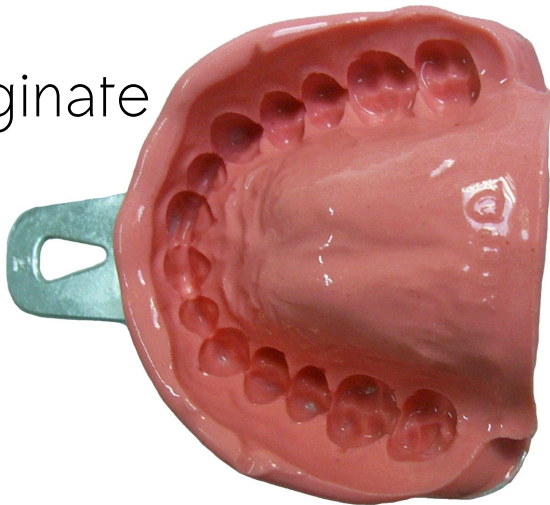
<http://www.enviropack.org.uk>

chitin



<http://www.ecovatedesign.com/>

alginate



www.alibaba.com

spider
silk



hanopolis.com

biorefineries will also produce renewable fuels



© beboy * www.ClipartOf.com/30312



www.qteros.com

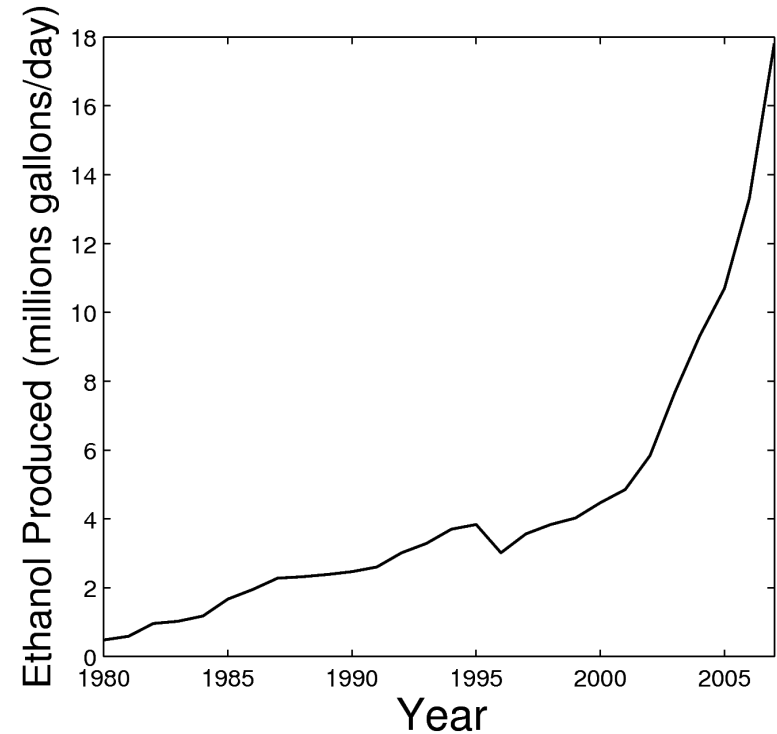
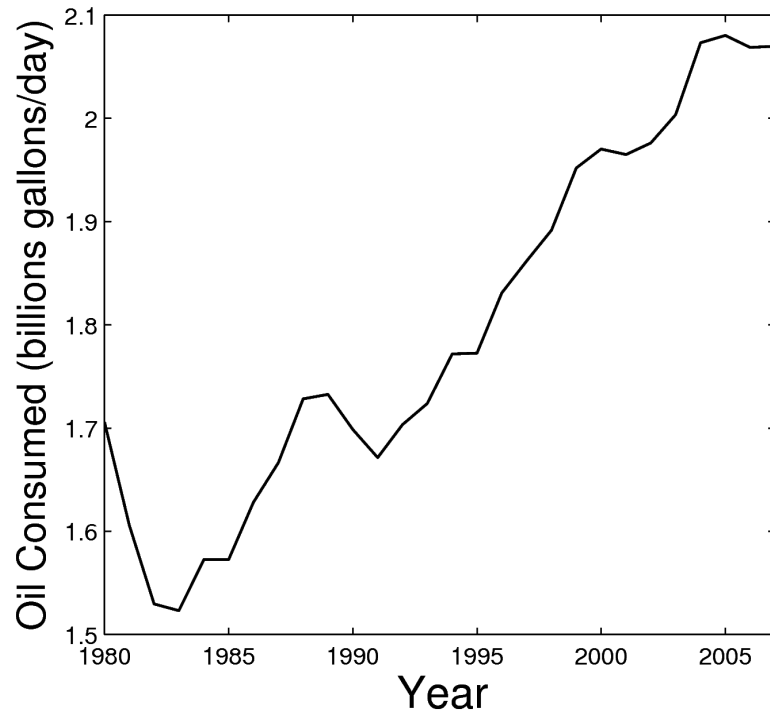


<http://www.energy-enviro.fi>



www.coskata.com

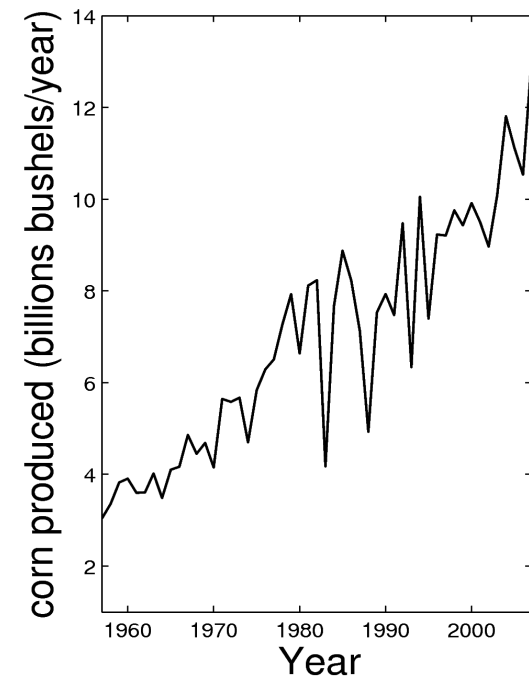
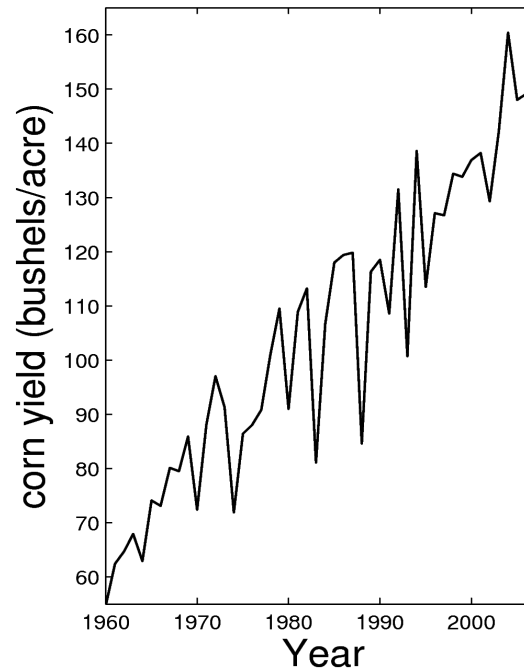
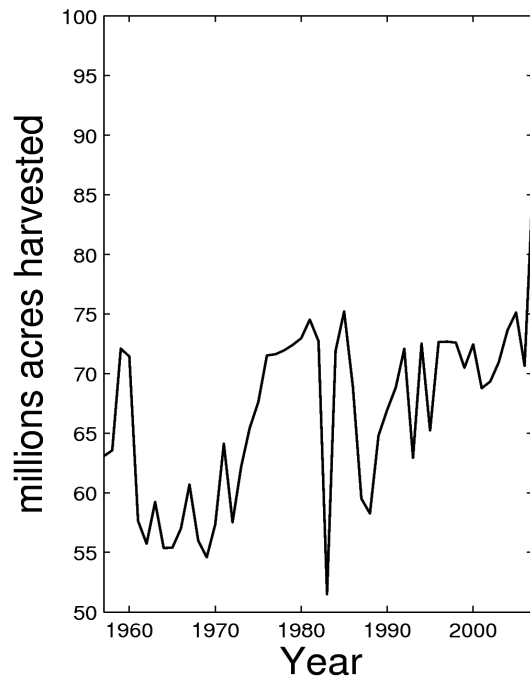
US oil consumption and ethanol production



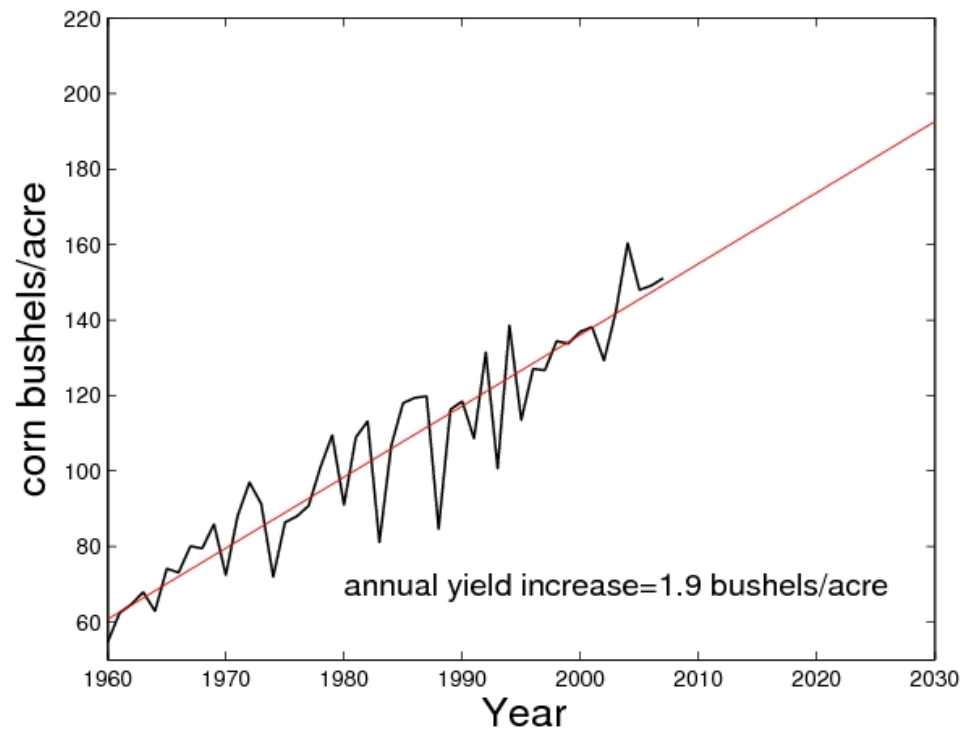
US oil consumption: <http://tonto.eia.doe.gov>

US ethanol production: <http://www.ethanolrfa.org>

US corn production



future corn yields



28% improved yields by 2030

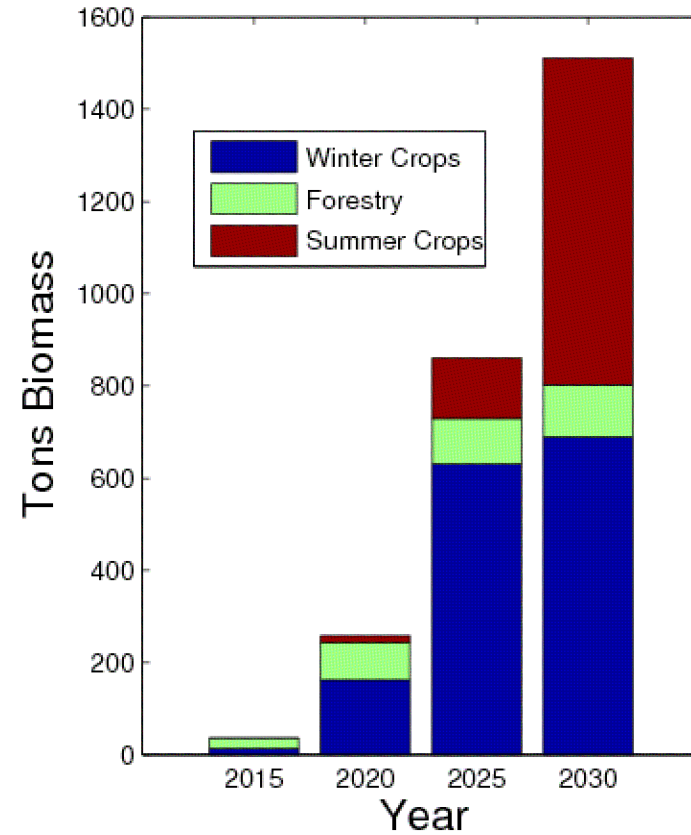
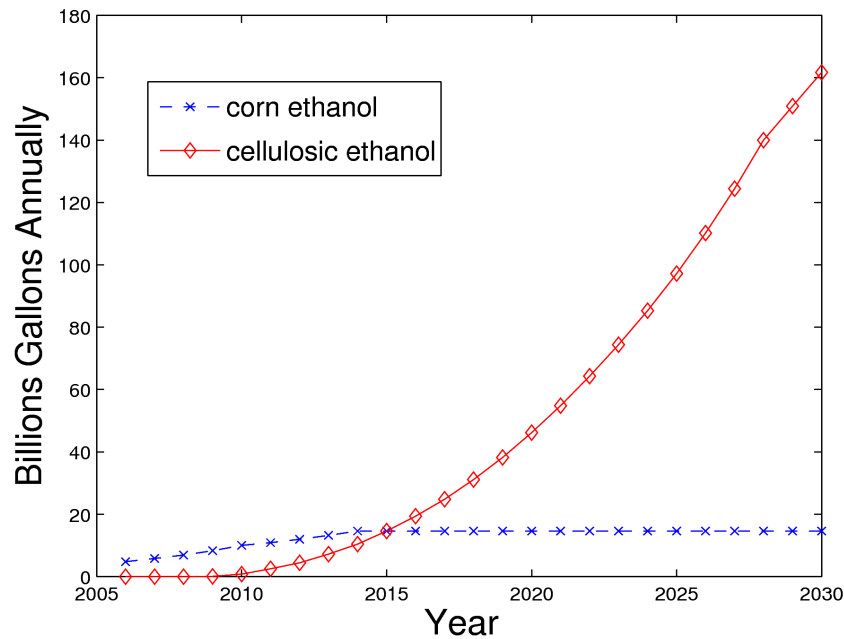
corn data from <http://www.nass.usda.gov>

we cannot grow enough corn to replace
petroleum!



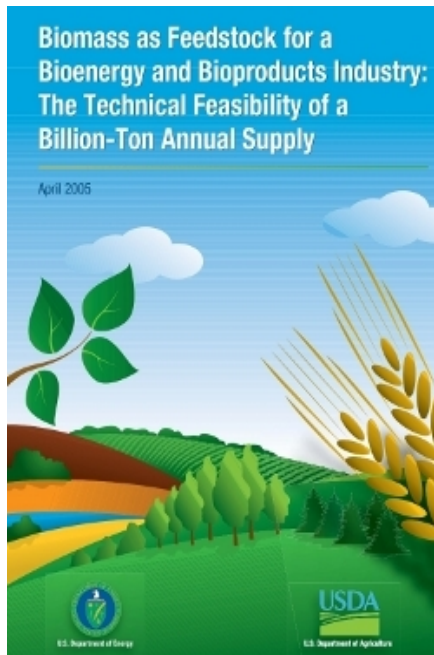
<http://www.chicagoboyz.net>

cellulosic biomass can replace corn to make biofuels at larger scale



data from <http://www.khoslaventures.com>

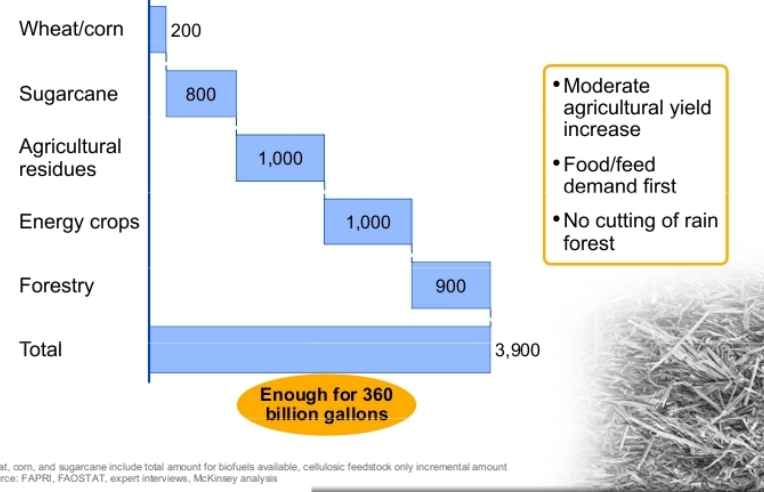
cellulosic feedstocks are abundant, inexpensive and reduce GHG emissions



Perlack et al., 2005 DOE/USDA report

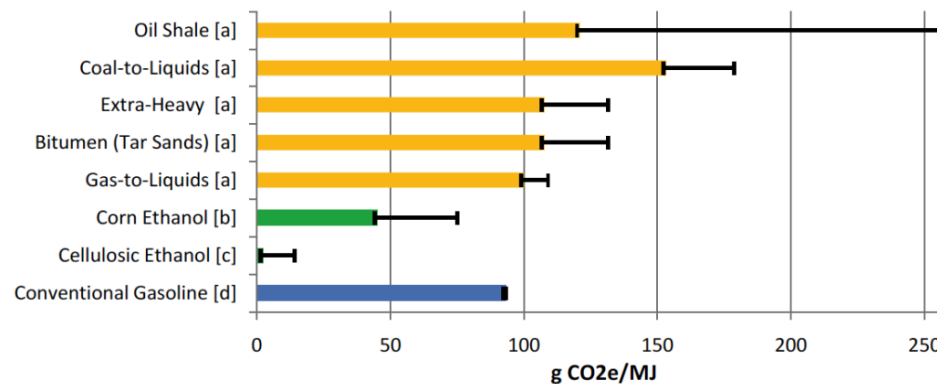
Enough biofeedstock to replace 50% of fuel

INCREMENTAL FEEDSTOCK POTENTIAL 2020*,
Million tons



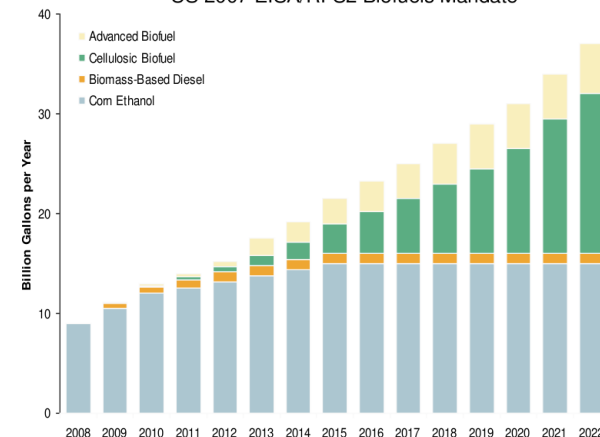
McKinsey and Co.

Lifecycle GHG Emissions

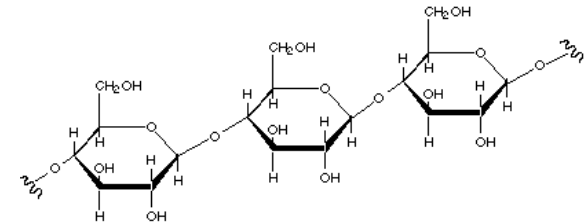


Data from Renewable Fuels Assn

US 2007 EISA/RFS2 Biofuels Mandate

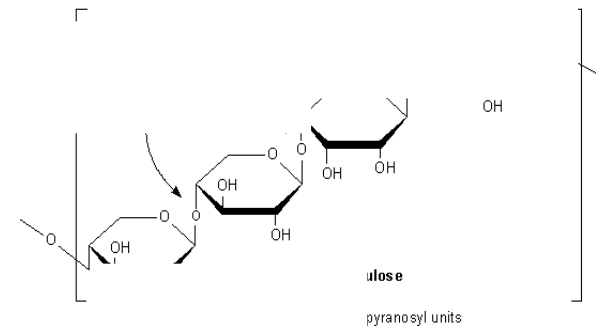


however, plant biomass is tough for microbes to eat

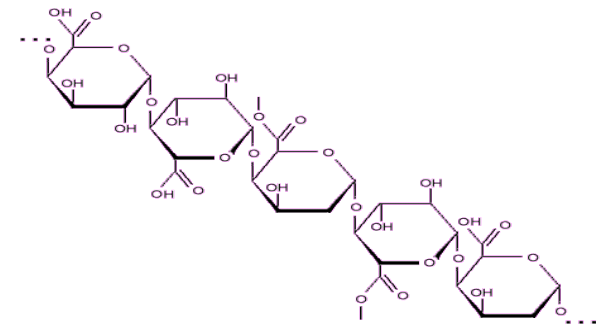


D-glucopyranoside)

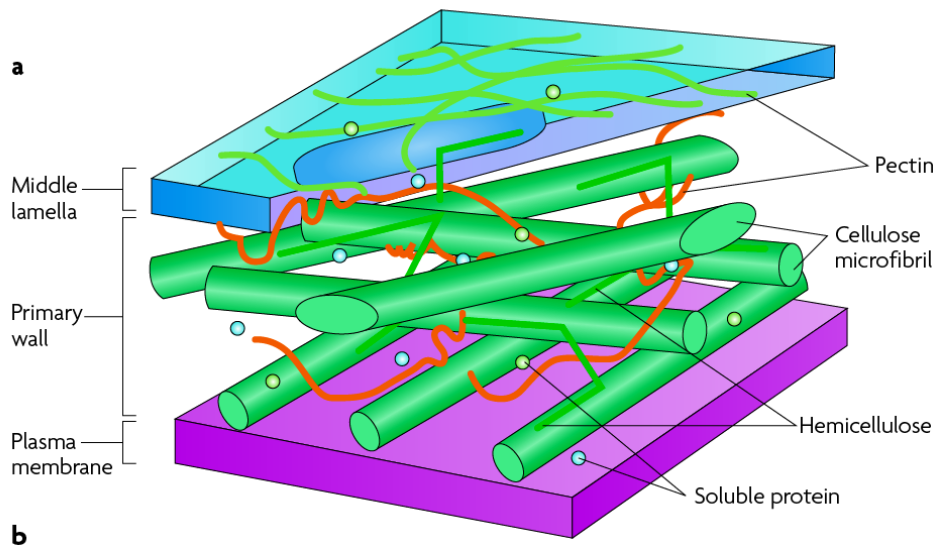
cellulose (1,4-beta-D-glucopyranoside)
35-50% of biomass



hemicellulose (1,4-beta-D-xylopyranoside)
20-30% biomass

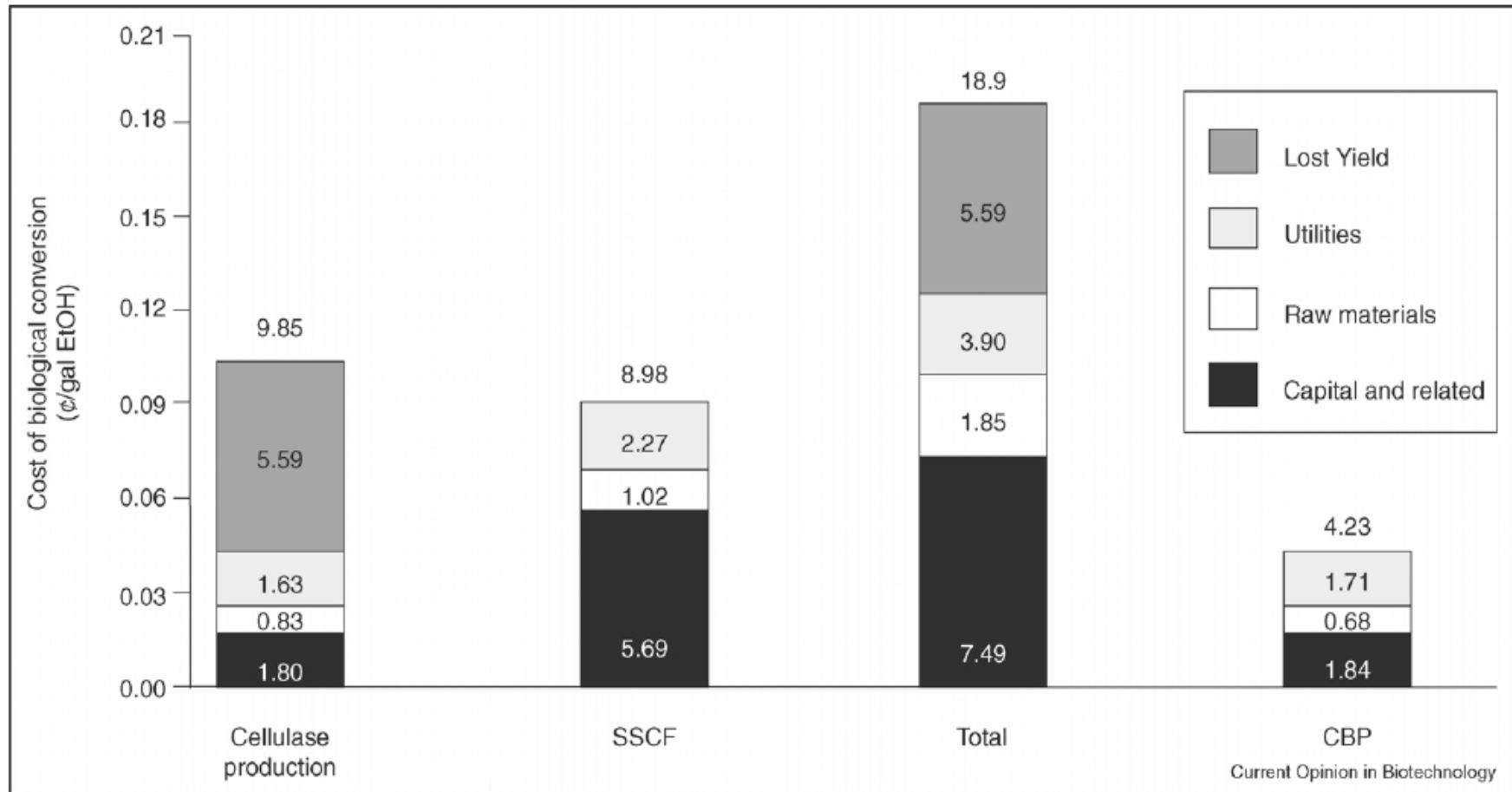


pectin α -(1-4) galacturonic acid ~ 10%
biomass



Stricklen et al, 2008

a microbe that can convert biomass to fuel in a single step is ideal



Clostridium phytofermentans: a new opportunity for cellulosic ethanol



International Journal of Systematic and Evolutionary Microbiology (2002), 52, 1155–1160

DOI: 10.1099/ijs.0.02125-0

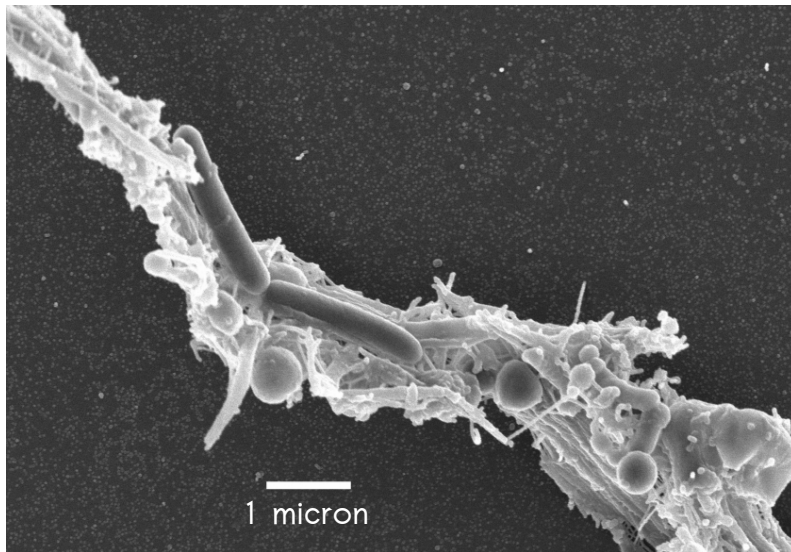
***Clostridium phytofermentans* sp. nov., a cellulolytic mesophile from forest soil**

Department of
Microbiology, University of
Massachusetts, Amherst,
MA 01003, USA

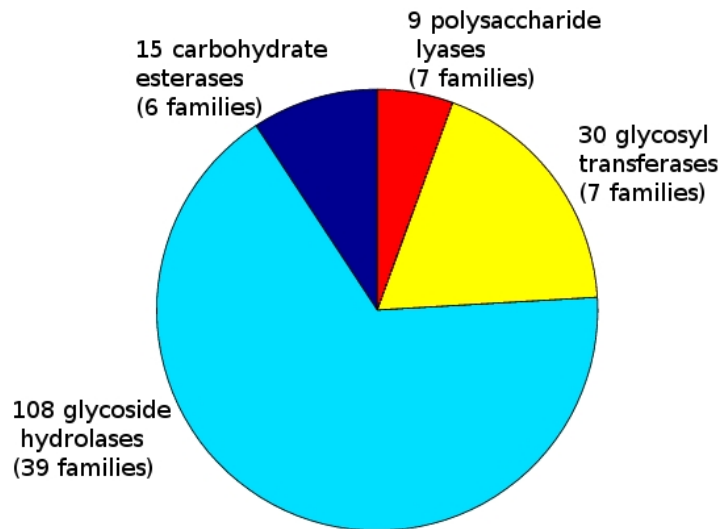
Thomas A. Warnick, Barbara A. Methé and Susan B. Leschine

Author for correspondence: Susan B. Leschine. Tel: +1 413 545 0673. Fax: +1 413 545 1578.
e-mail: suel@microbio.umass.edu

Clostridium phytofermentans (Cphy) ferments cellulosic biomass to ethanol and hydrogen

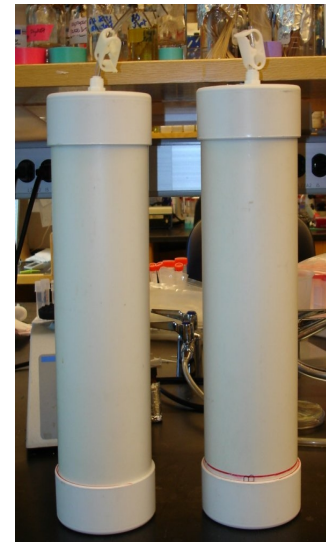
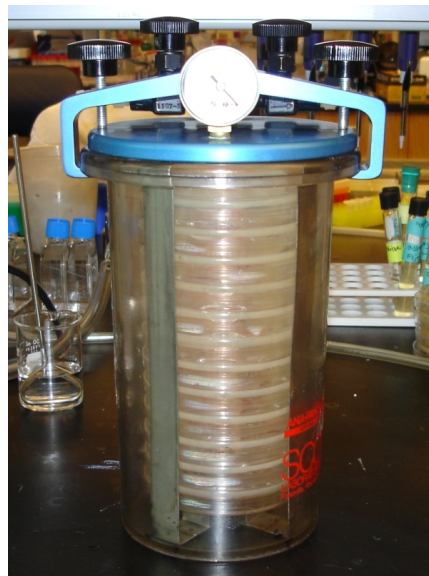


Cphy genome has 161 carbohydratases (CAZy)

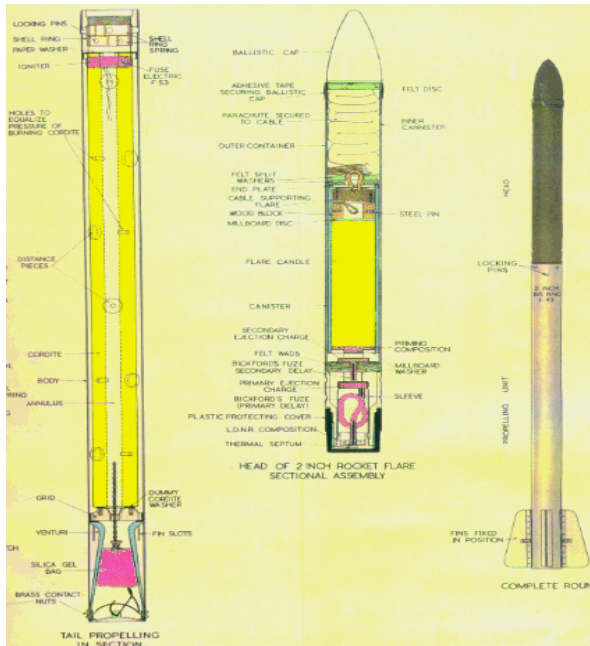


Cphy is an obligate anaerobe

"la vie sans l'air" -Louis Pasteur 1862



Chaim Weizmann: clostridia, cordite, and the founding of Israel



<http://www.cyber-heritage.co.uk>



<http://www.lesjones.com>



<http://www.wereldoorlog1418.nl>

"There is only one thing I want: A national home for my people." -Chaim Weizmann



● Clostridial fermentation plant <http://www.world-atlas.us>



<http://www.trumanlibrary.org>



<http://www.lonklab.ac.uk>

my research focus: new technologies

1. Systems biology: identify key enzymes/pathways
2. Synthetic biology: genome engineering



new generation
biorefineries



fuels

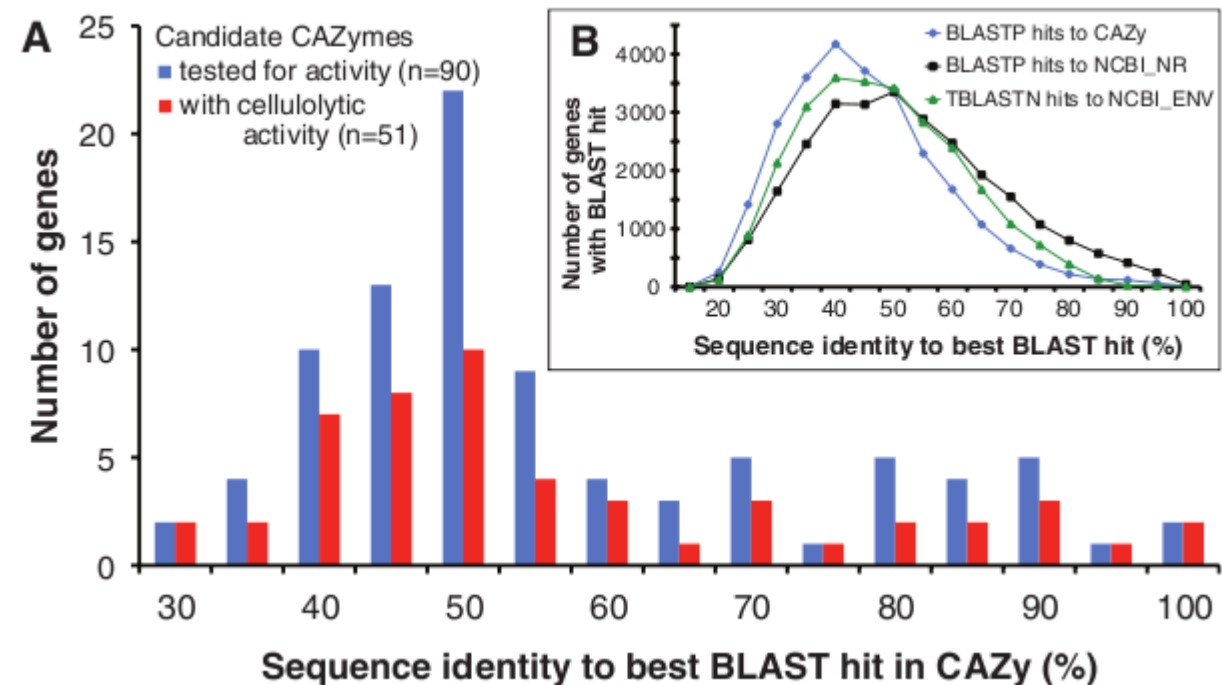
biocommodities

metagenomics for new enzymes from the cow rumen

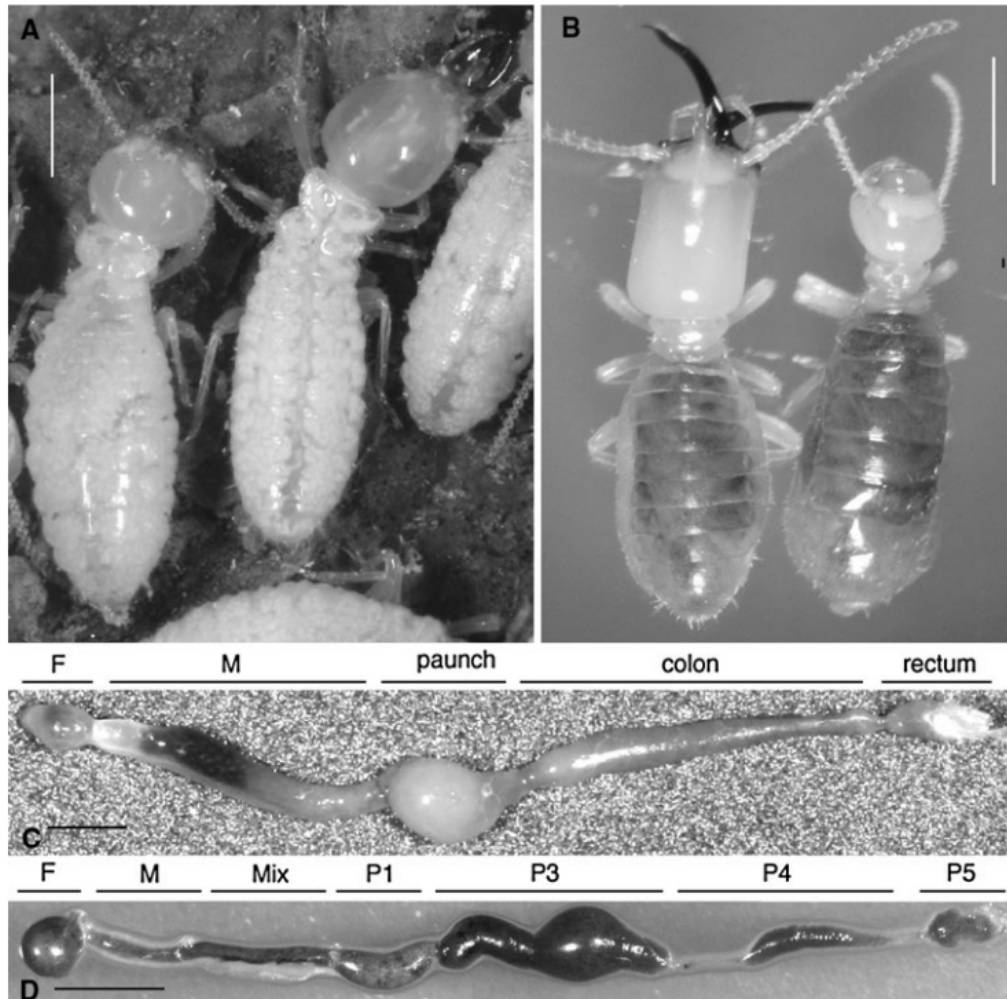


Hess et al, 2011

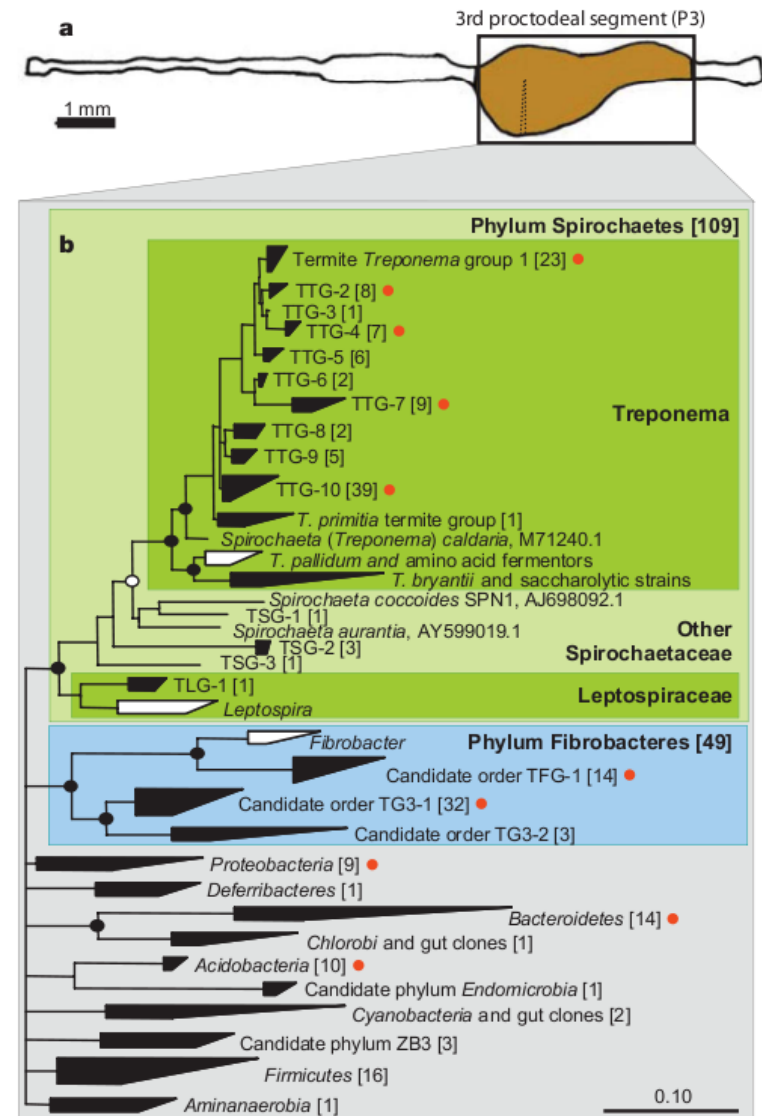
Identified 27,755 putative carbohydrate-active genes



...and also metagenomics of the termite gut

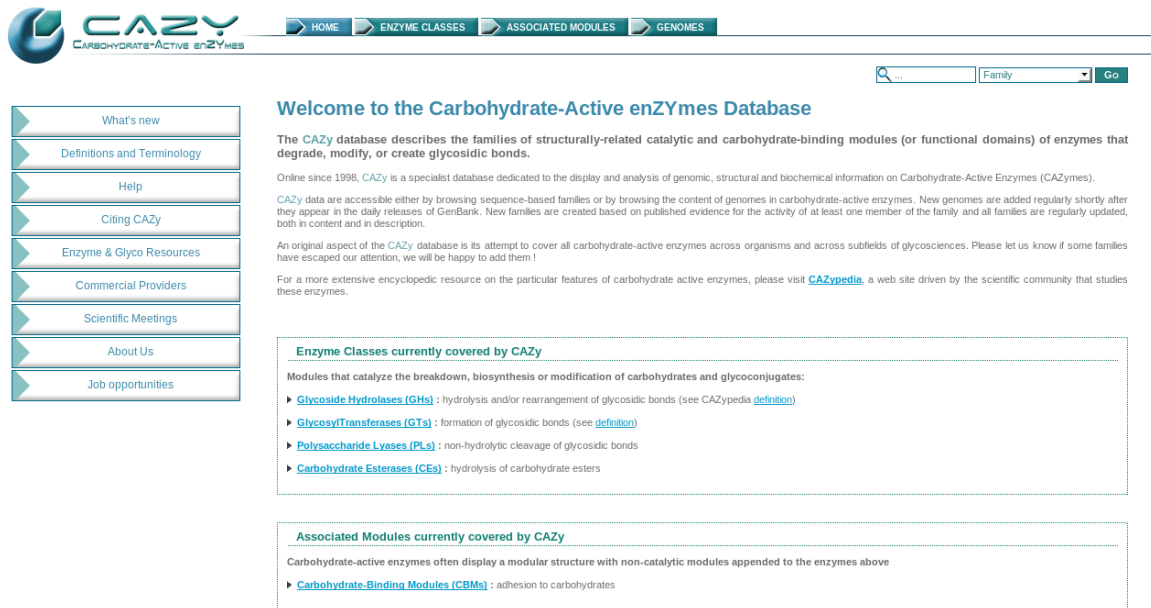


Hongoh 2011



Warnecke et al, 2007

many putative cellulolytic enzymes sequenced, but what are their substrates?



The screenshot shows the CAZY database homepage. At the top is the CAZY logo and navigation links: HOME, ENZYME CLASSES, ASSOCIATED MODULES, and GENOMES. Below the logo is a search bar with a 'Go' button. A left sidebar contains links: What's new, Definitions and Terminology, Help, Citing CAZY, Enzyme & Glyco Resources, Commercial Providers, Scientific Meetings, About Us, and Job opportunities. The main content area has a heading 'Welcome to the Carbohydrate-Active enZYmes Database' followed by a description of the database. It lists enzyme classes covered: Glycoside Hydrolases (GHs), Glycosyltransferases (GTs), Polysaccharide Lyases (PLs), and Carbohydrate Esterases (CEs). It also lists associated modules: Carbohydrate-Binding Modules (CBMs).

Welcome to the Carbohydrate-Active enZYmes Database

The CAZY database describes the families of structurally-related catalytic and carbohydrate-binding modules (or functional domains) of enzymes that degrade, modify, or create glycosidic bonds.

Online since 1996, CAZY is a specialist database dedicated to the display and analysis of genomic, structural and biochemical information on Carbohydrate-Active Enzymes (CAZymes).

CAZY data are accessible either by browsing sequence-based families or by browsing the content of genomes in carbohydrate-active enzymes. New genomes are added regularly shortly after they appear in the daily releases of GenBank. New families are created based on published evidence for the activity of at least one member of the family and all families are regularly updated, both in content and in description.

An original aspect of the CAZY database is its attempt to cover all carbohydrate-active enzymes across organisms and across subfields of glycosciences. Please let us know if some families have escaped our attention, we will be happy to add them !

For a more extensive encyclopedic resource on the particular features of carbohydrate active enzymes, please visit [CAZypedia](#), a web site driven by the scientific community that studies these enzymes.

Enzyme Classes currently covered by CAZY

Modules that catalyze the breakdown, biosynthesis or modification of carbohydrates and glycoconjugates:

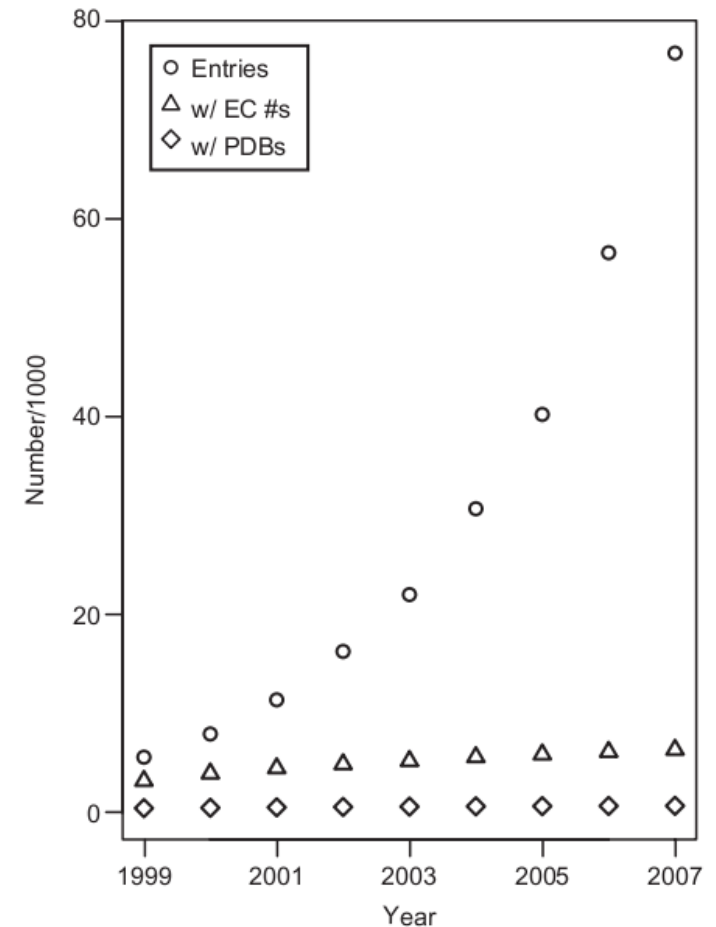
- ▶ [Glycoside Hydrolases \(GHs\)](#) : hydrolysis and/or rearrangement of glycosidic bonds (see CAZypedia [definition](#))
- ▶ [Glycosyltransferases \(GTs\)](#) : formation of glycosidic bonds (see [definition](#))
- ▶ [Polysaccharide Lyases \(PLs\)](#) : non-hydrolytic cleavage of glycosidic bonds
- ▶ [Carbohydrate Esterases \(CEs\)](#) : hydrolysis of carbohydrate esters

Associated Modules currently covered by CAZY

Carbohydrate-active enzymes often display a modular structure with non-catalytic modules appended to the enzymes above

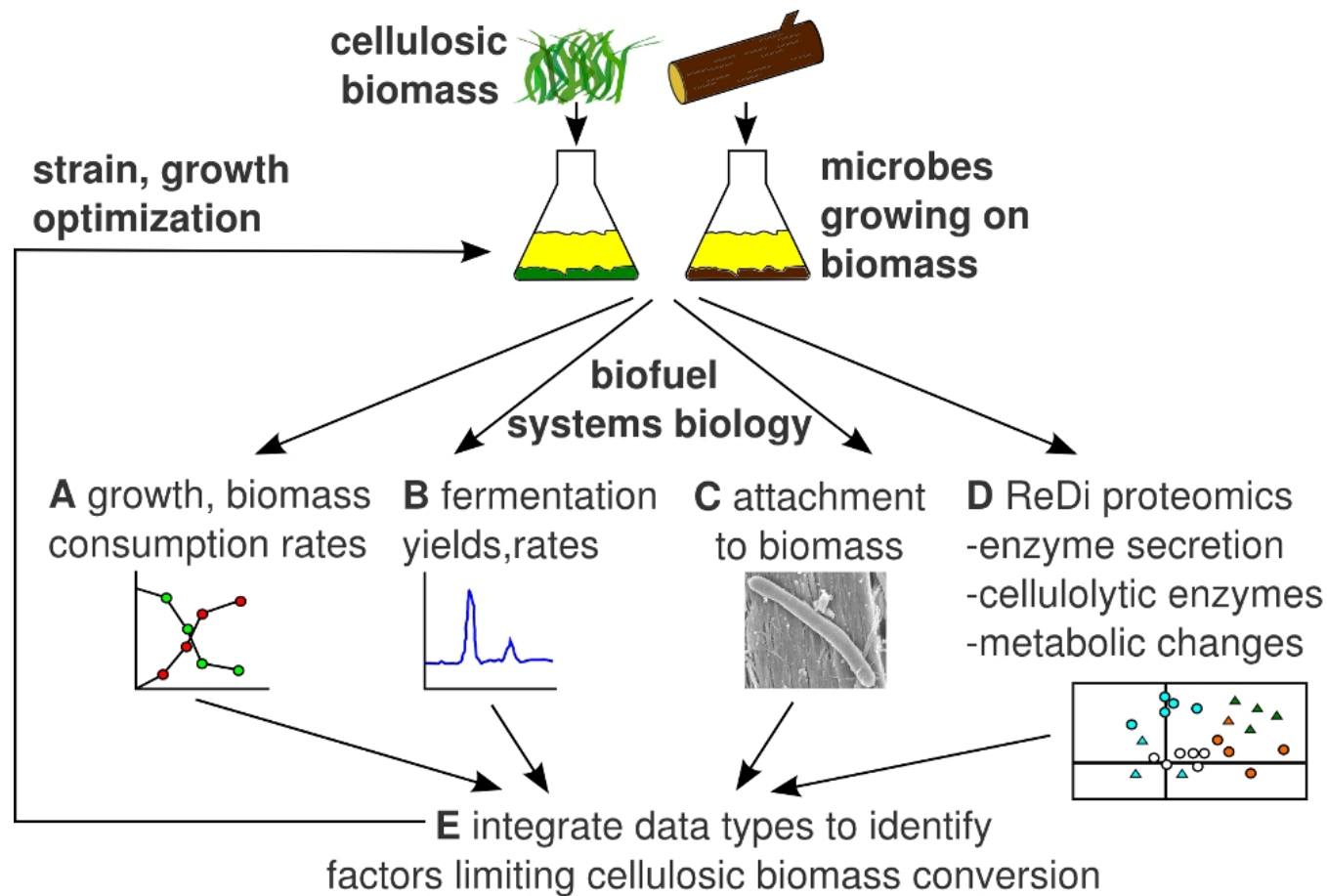
- ▶ [Carbohydrate-Binding Modules \(CBMs\)](#) : adhesion to carbohydrates

www.cazy.org



Cantarel et al, 2009

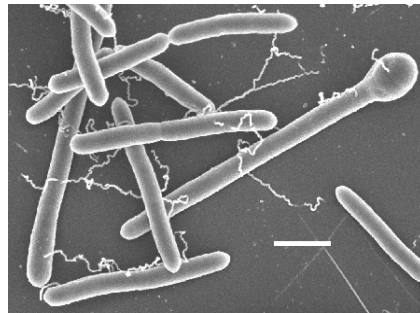
systems biology of cellulosic bioconversion



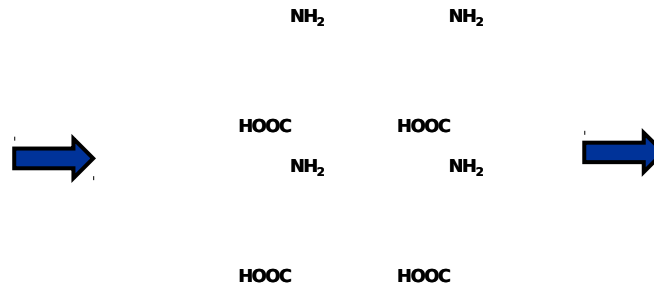
Tolonen et al, 2011

how proteins are identified by mass spectrometry

Sample Preparation



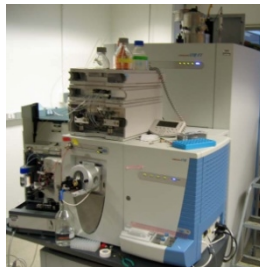
Cphy culture



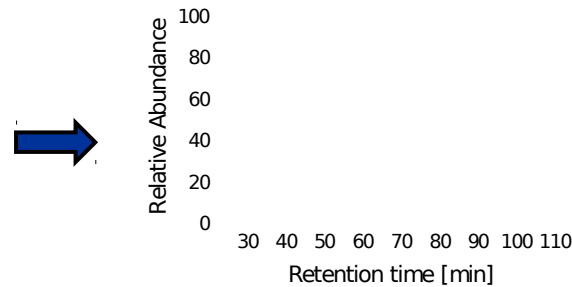
protein lysate by french press

proteins $\xrightarrow{\text{trypsin}}$ peptides

Data Acquisition



LC-MS/MS



preparation of Peptides

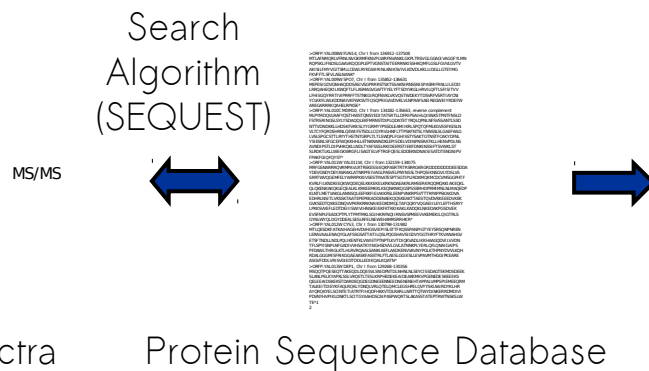
Survey MS

MS Spectrum

MS/MS

MS/MS Spectra

Data Analysis



MS/MS Spectra

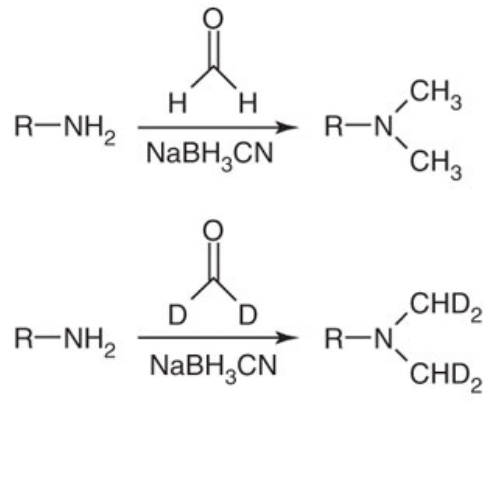
Protein Sequence Database

Peptide Identifications

Protein Identifications

quantify protein differences between treatments by reduction dimethylation (ReDi)

label amines with
methyl isotopes



glucose

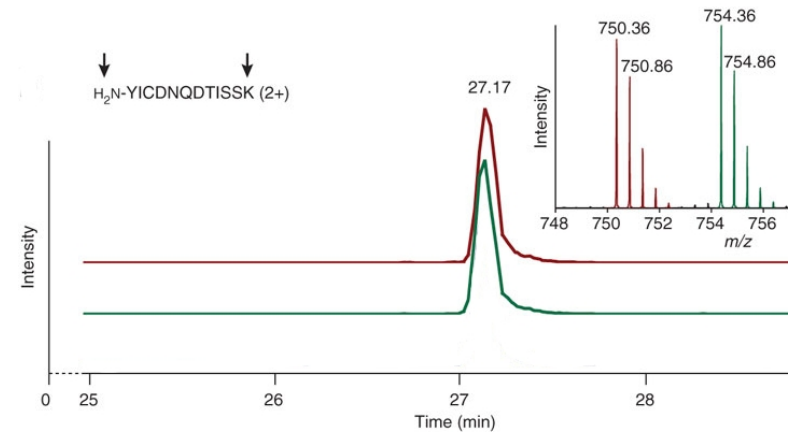


Mix 1:1

cellulose

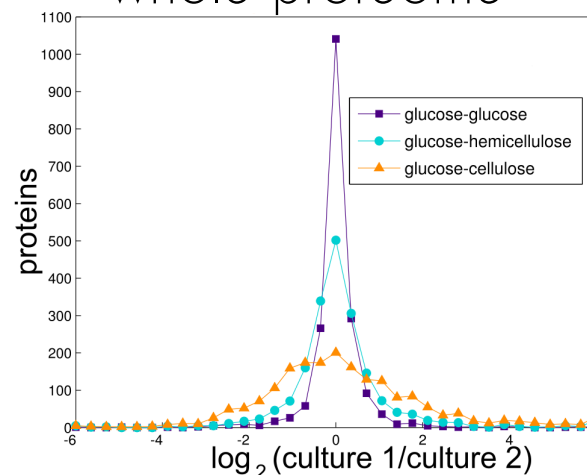


quantify expression changes
using MS1 peak area ratios

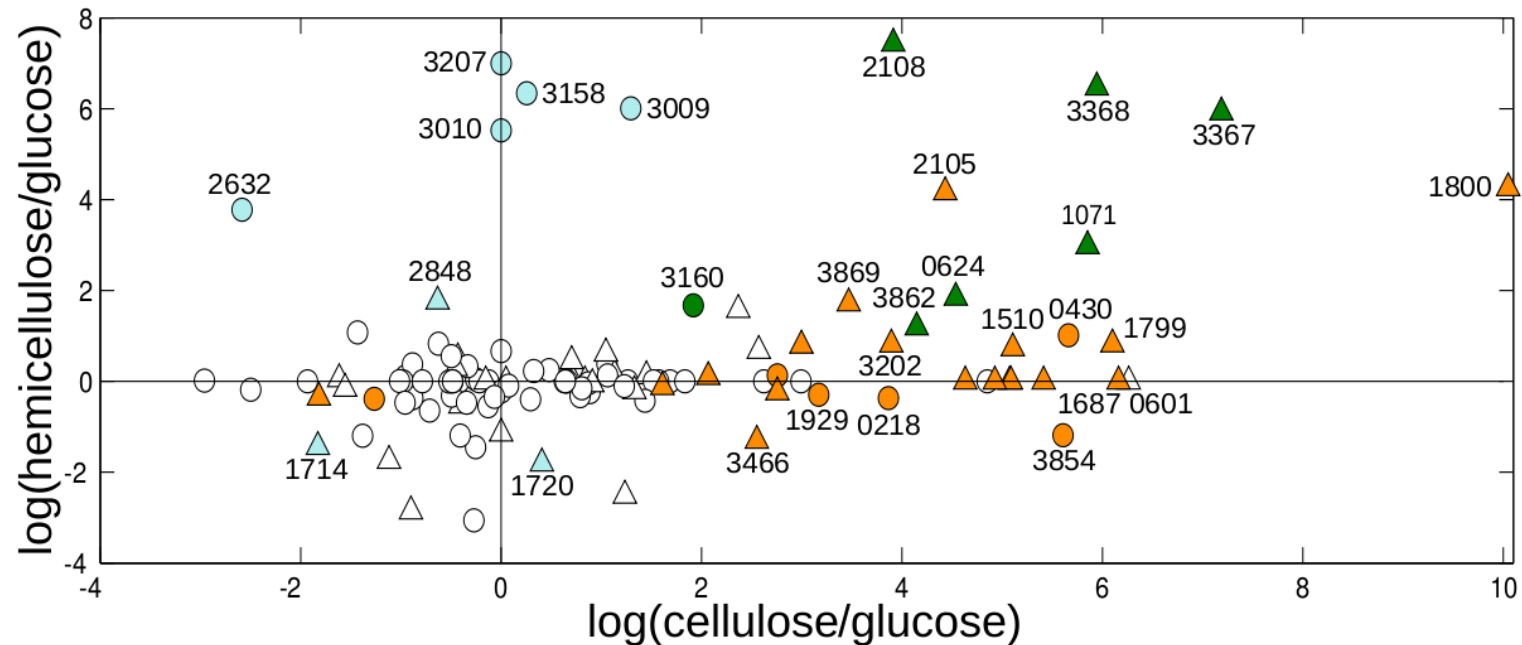


Boersema et al, 2009

whole proteome



distinct CAZy repond to hemicellulose and cellulose



Tolonen et al, 2011

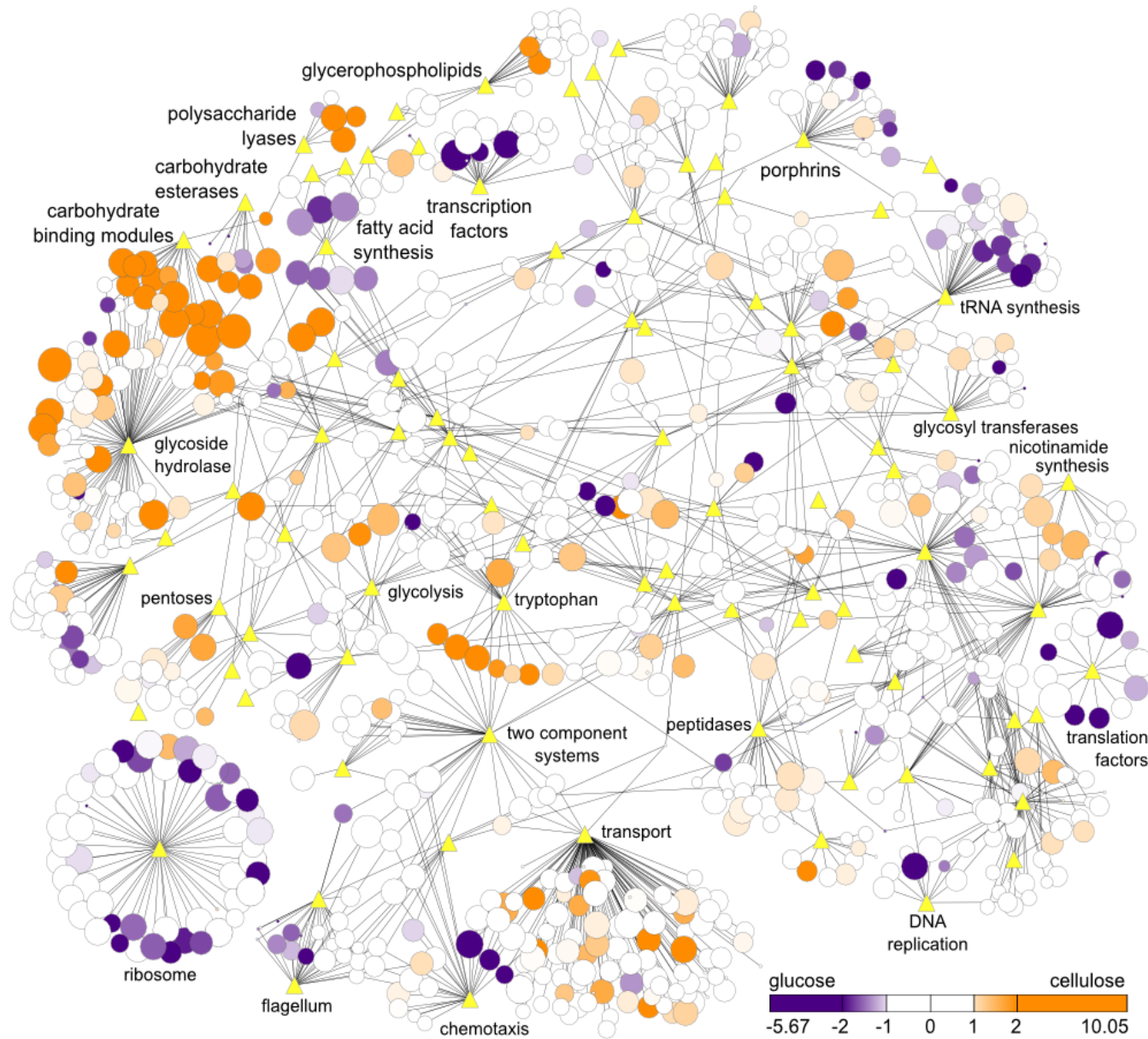
significant differential expression

cellulose=orange, hemicellulose=cyan, both=green

localization

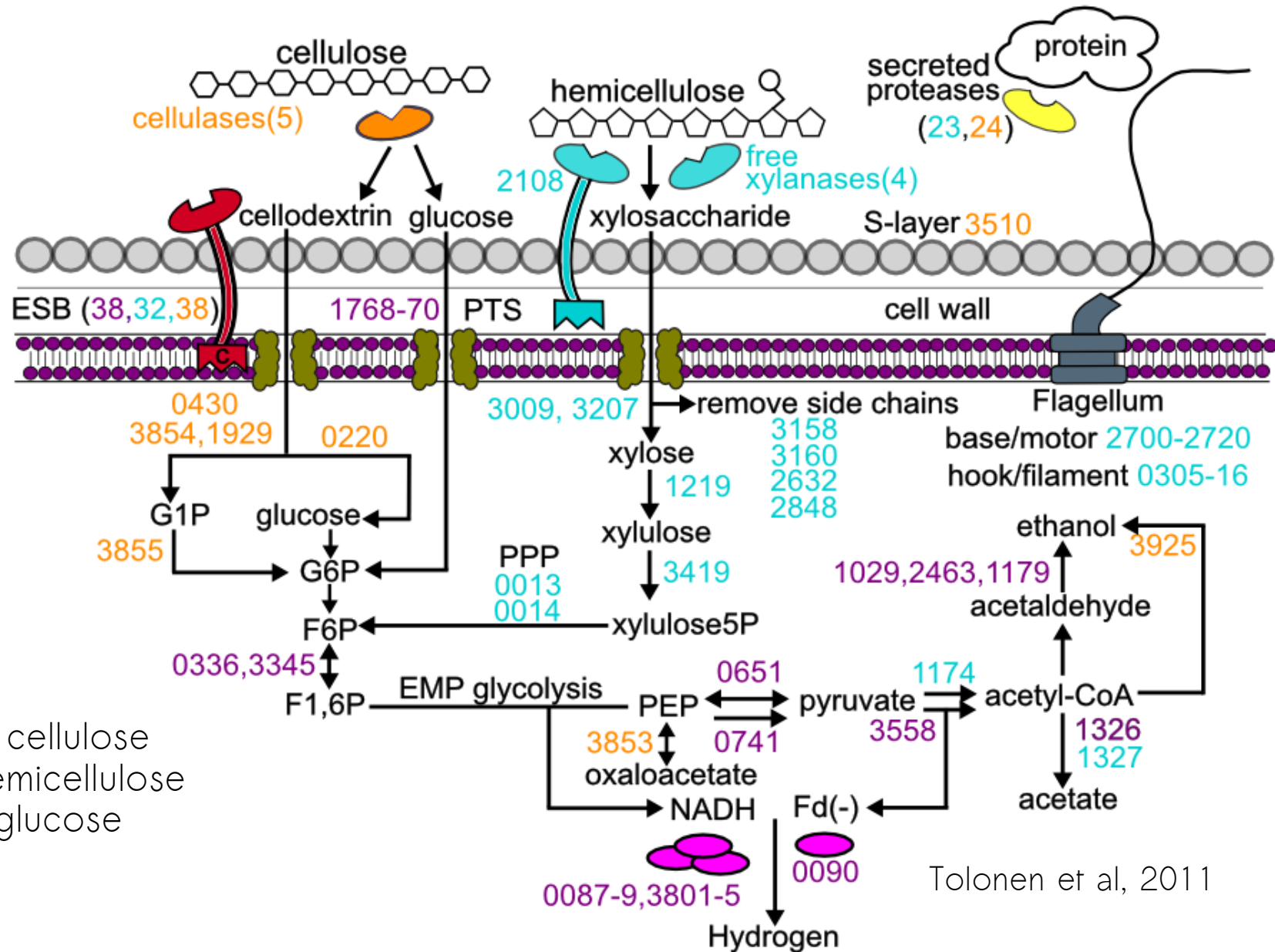
secreted=triangle, intracellular=circle

interaction map of proteome expression changes



Tolonen et al, 2011

deconstruction and fermentation model

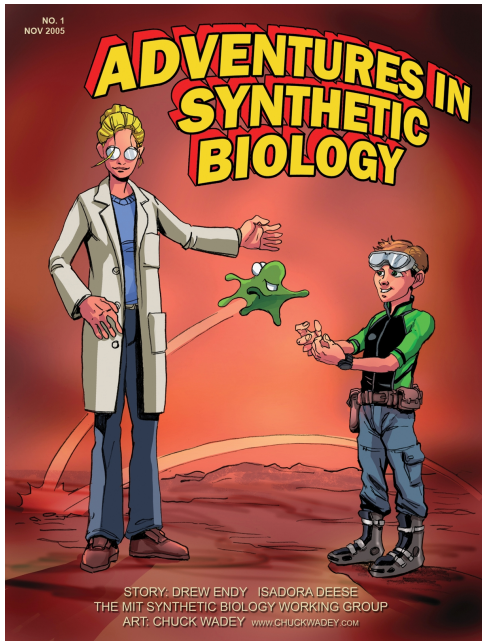


Tolonen et al, 2011

Synthetic biology: once we know the enzymes, we can engineer the microbes

de novo construct
biocatalysts from model microbes

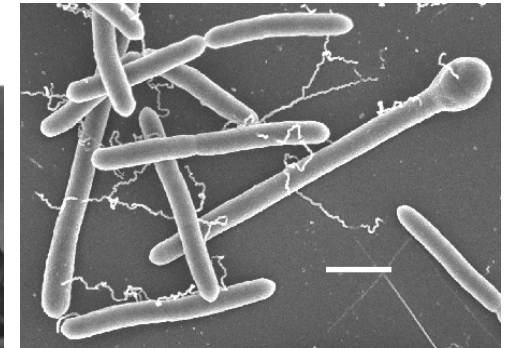
study, modify natural isolates



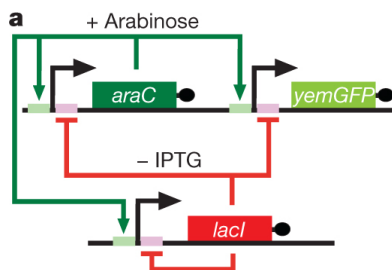
nature.com



www.trumanlibrary.org

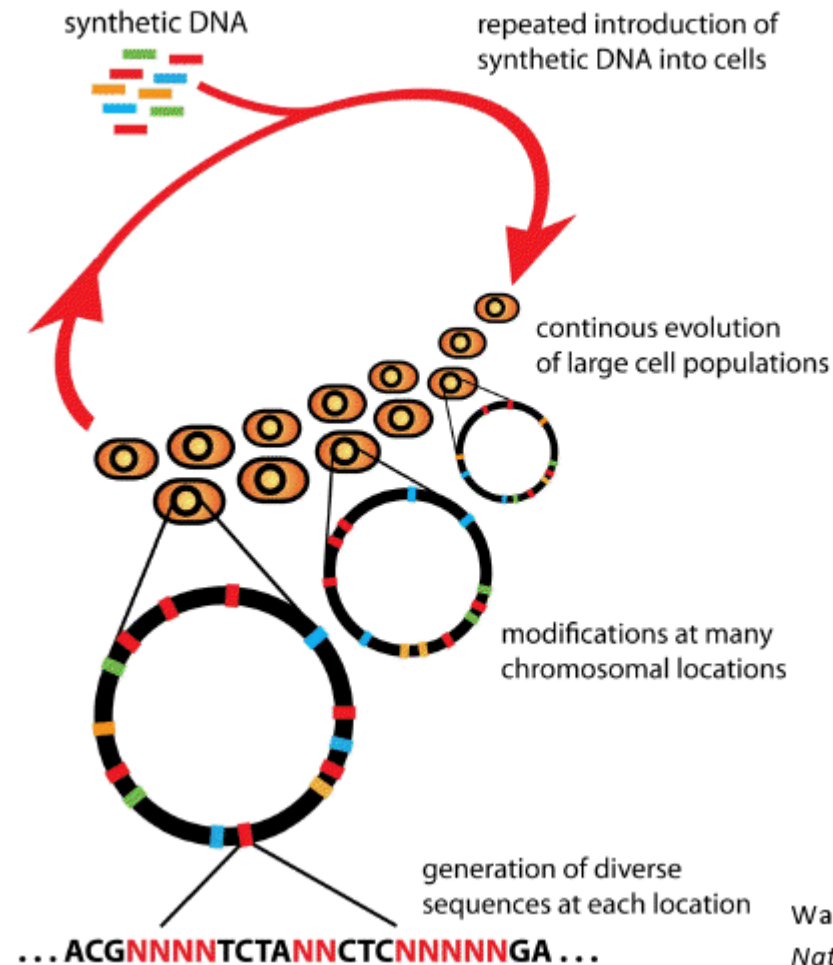


<http://www.lesjones.com>



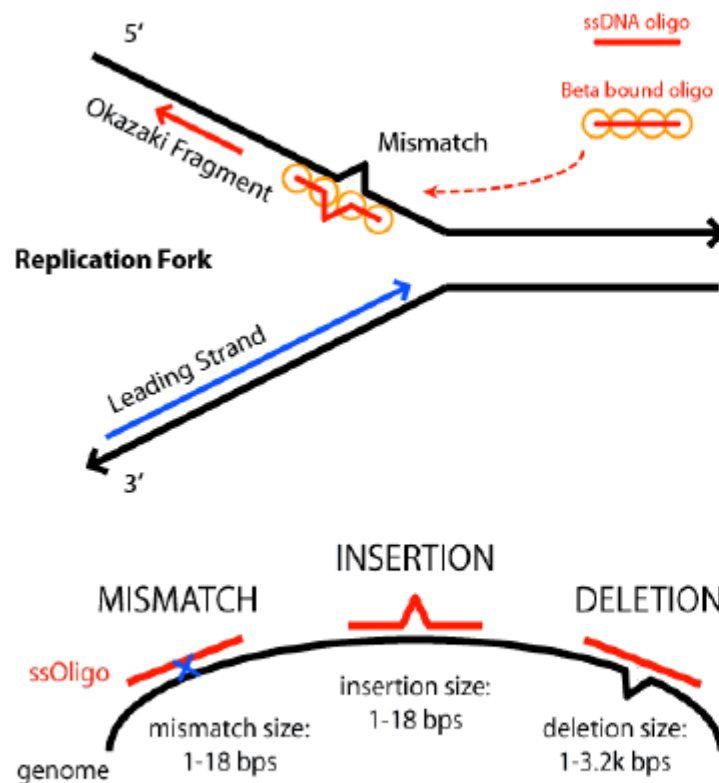
<http://ucsdnews.ucsd.edu>

MAGE is a method to make many small changes to the *E. coli* genome



Wang, H.H., Isaacs, F.J., et al.
Nature (2009) 460, 894-898

How MAGE works



Expediting the design & evolution of organisms with new & improved properties

- Simultaneously targets many genetic locations
- Automatable
- Highly efficient genetic changes (>30%)
- Integrate engineering & evolution
- Combinatorial genomic diversity

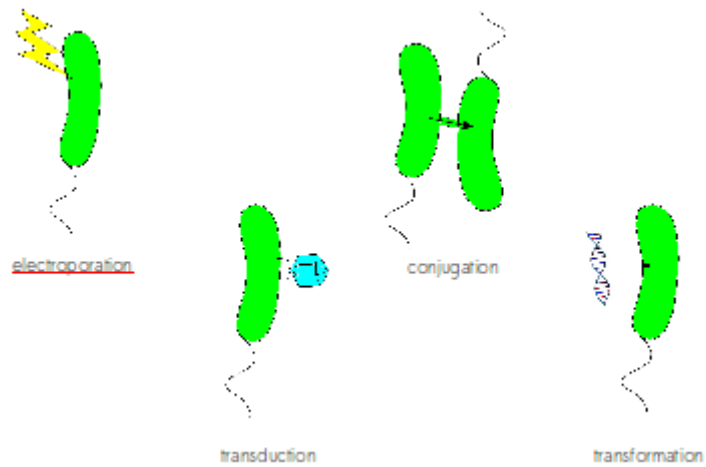
Needs

- Fewer clones with no changes
- More simultaneous changes

Wang, H.H., Isaacs, F.J., et al. *Nature* (2009) **460**, 894-898

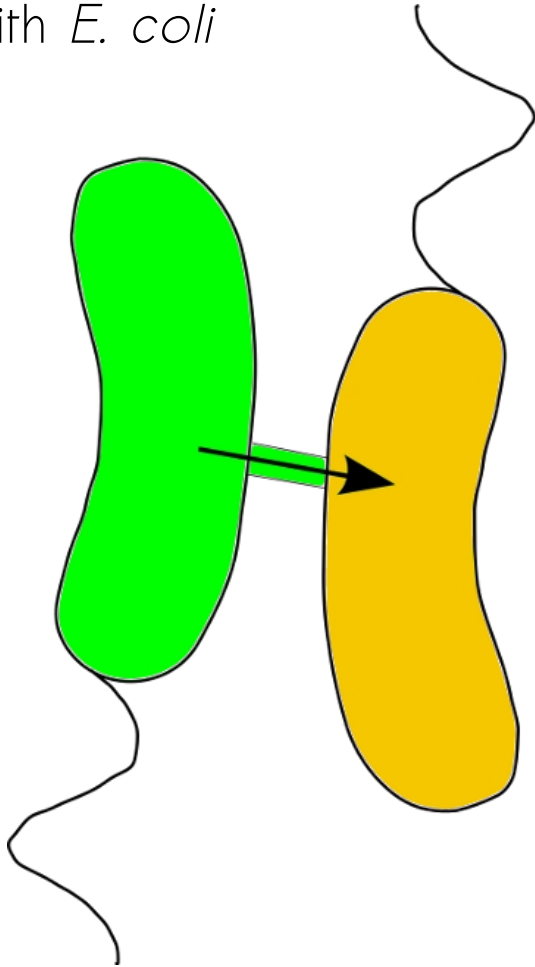
What about engineering novel
bacteria? (ie *Clostridium*
phytofermentans)

Delivery of foreign DNA into bacteria

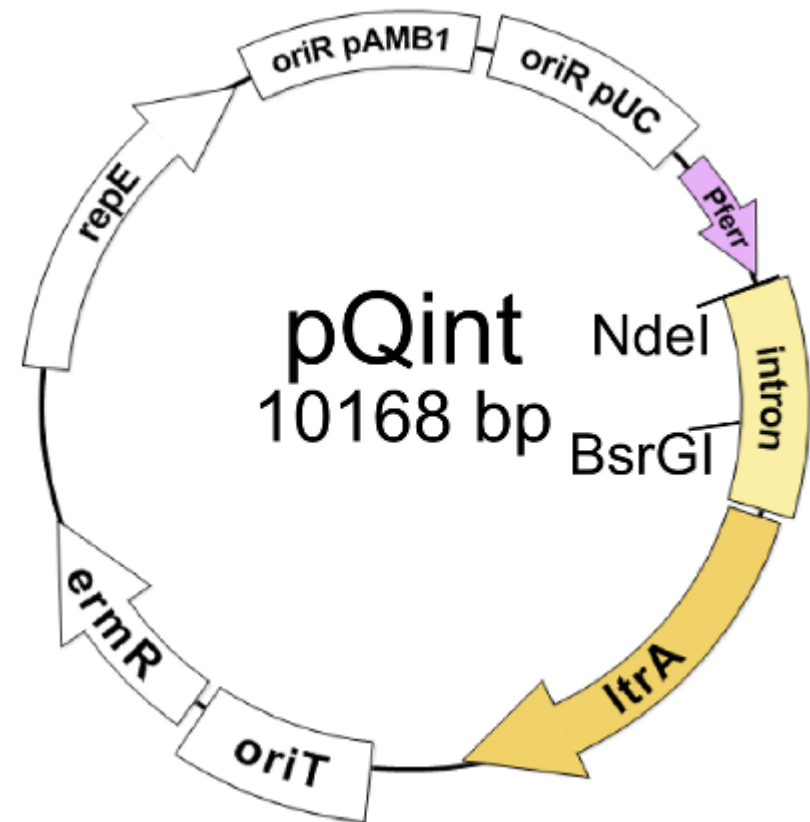


Plasmid DNA transfer to bacteria

DNA delivery by conjugation
with *E. coli*



Plasmid for targeted chromosomal
insertions

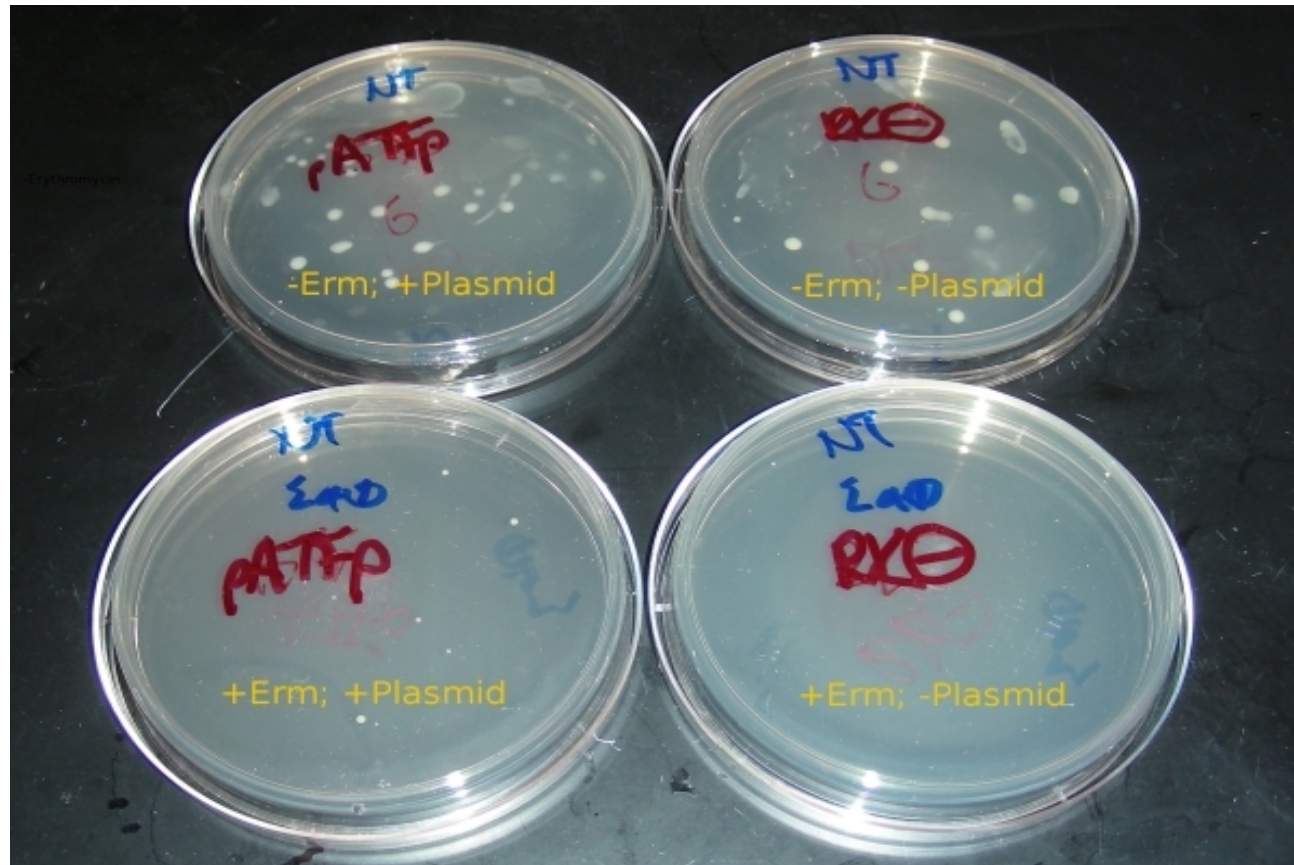


Tolonen et al, 2009

we need antibiotics to determine which cells got the plasmid DNA

$10^{-6} \times$
dilution

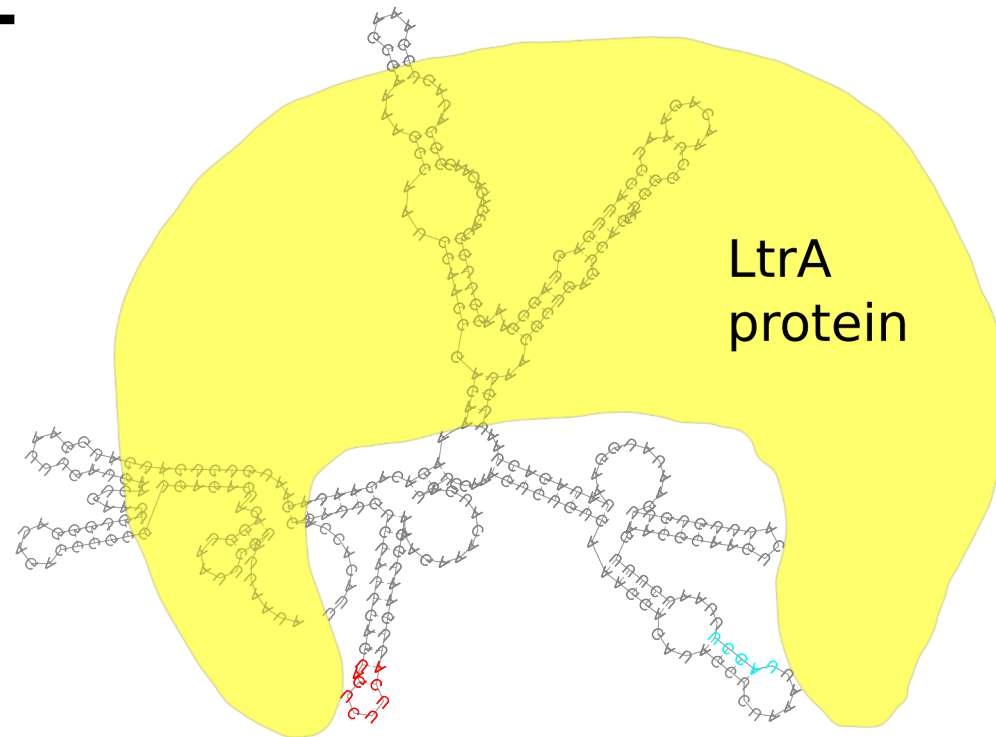
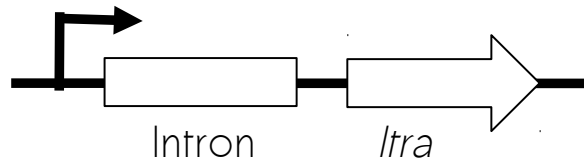
1x
dilution



Now we have DNA in the cell. How do we get it into the genome?

Homologous recombination

group II introns are more efficient than recombination

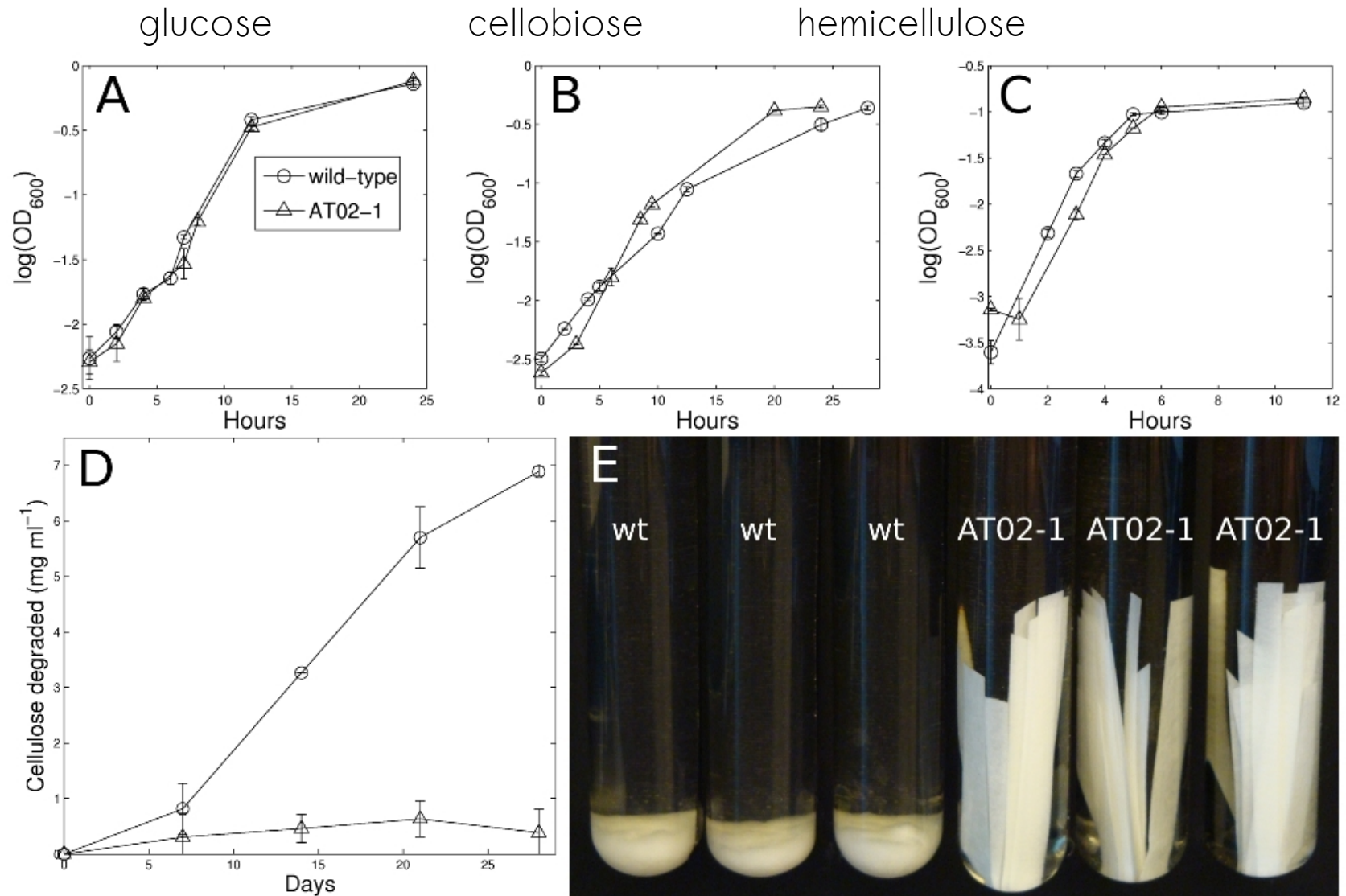


5' atgaaaaagataataagtccttttattagtgataaac**aCttctgat**atccatggcaccatcgaaagctgacgcagcg3'
3' tactttttctattatttcagaaaataatcactattgt**Gaagacta**taggtaccgtggtagctttcgactgcgtcgc5'

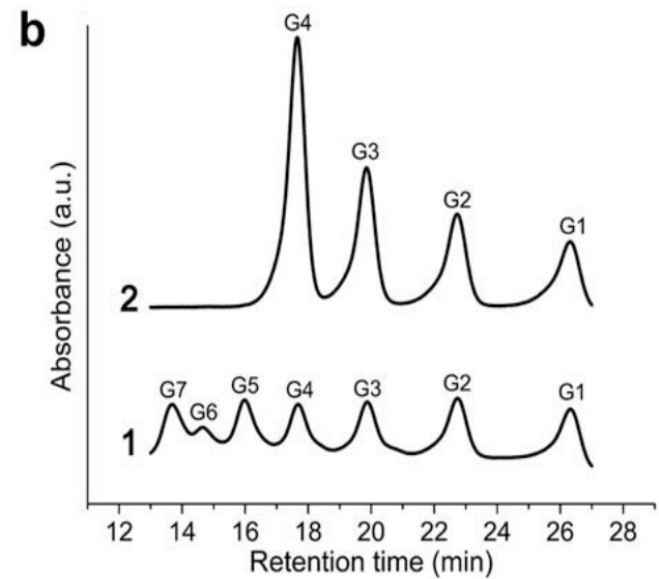
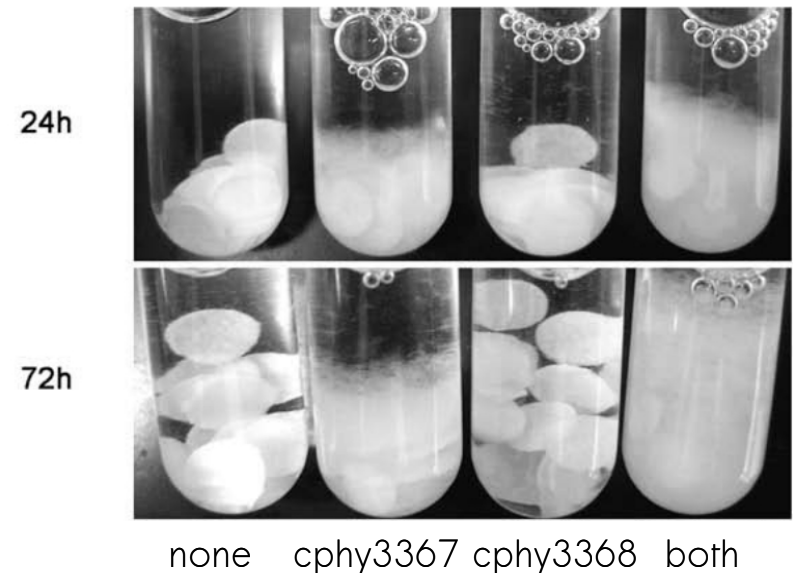
LtrA EBS1 ESB2 LtrA

↑
Insert site

a single cellulase is required for cellulose degradation



Cphy3367: a "Rambo" cellulase



images from Zhang et al., 2010 (except Rambo)

Now we can efficiently modify the *Cphy* genome, our future directions are to use systems biology to guide genome engineering