

## Emerging Platforms for Biofuels and Biochemicals: The Role of Metabolic Engineering and Systems Biology

**Ramon Gonzalez** 

Department of Chemical and Biomolecular Engineering
Rice University

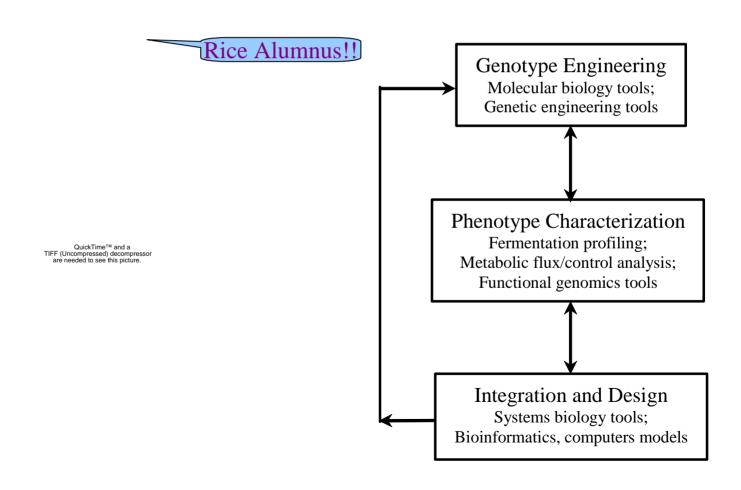
Biomass to Chemicals and Fuels: Science, Technology and Public Policy September 25-26, 2006 Baker Institute International Conference Facility

Rice University, Houston, TX

#### Metabolic Engineering (ME)



Metabolic engineering is the improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology. The oppor-



#### **Systems Biology**



## A relatively new field that aims at systems level understanding of biological processes

#### The two Determinants of Systems Biology

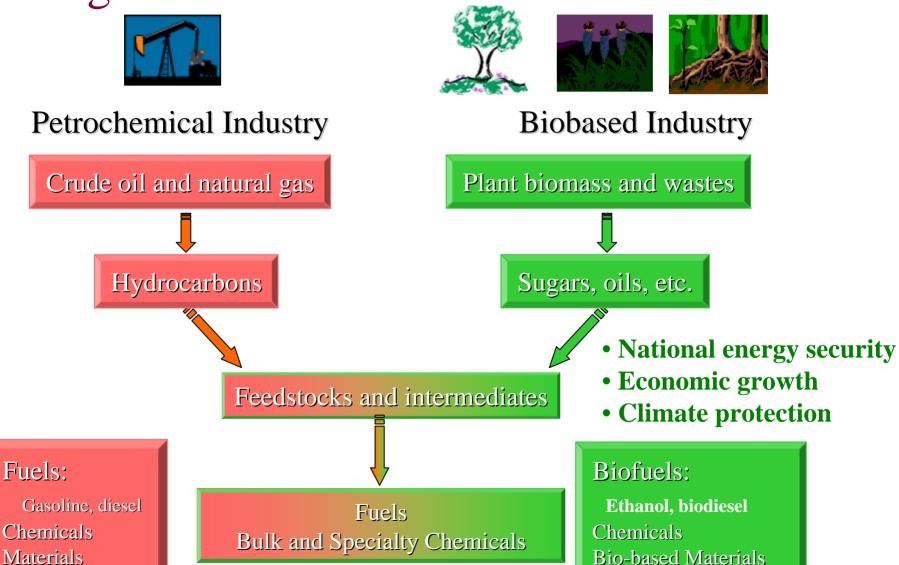
- 1. New technologies for comprehensive, high-throughput, quantitative measurements at system/cellular levels: i.e., functional genomics tools.
- 2. Advances in theory, modeling, softwares, and computational power for data analysis and integration.

QuickTime<sup>TM</sup> and a RFF (Uncompressed) decompress are needed to see this picture

#### Renewables in Our Future?



#### Linking PetroChemical and BioBased Industries



#### Three Main Platforms for Fuels and Threed Main Platforms for Fuels and Chemicals from Biomass



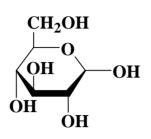
- The sugar platform is based on extracting sugars from plant biomass and using them as substrates (feedstocks) for the production of fuels and chemicals, predominantly via fermentation processes. This is the platform that includes cellulosic ethanol.
- The syngas platform proposes to process the plant biomass thermally (pyrolysis/combustion) to obtain heat, power and a gas mix (syngas or synthesis gas) containing CO, CO<sub>2</sub>, H<sub>2</sub>, and other compounds. Syngas will then be processed via various chemical or fermentation routes to produce ethanol and other chemicals.
- The oil platform is related to biodiesel or bio-distillate production. Vegetable oils or animal fats are used biodiesel (a mixture of fatty acid methyl esters) or process it into bio-distillates via conventional refinery technology.

#### US Feedstocks



#### **Today:** Corn





Glucose

CH<sub>2</sub>OH CH<sub>2</sub>OH OH OH ·OH OH

Maltose

50%, Hemicellulose: 20-40%

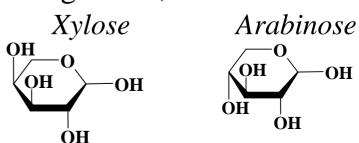
#### **Tomorrow:** Lignocellulosic biomass Lignin: 20-30%, Cellulose: 30-



CH<sub>2</sub>OH OH OH OH ÓН

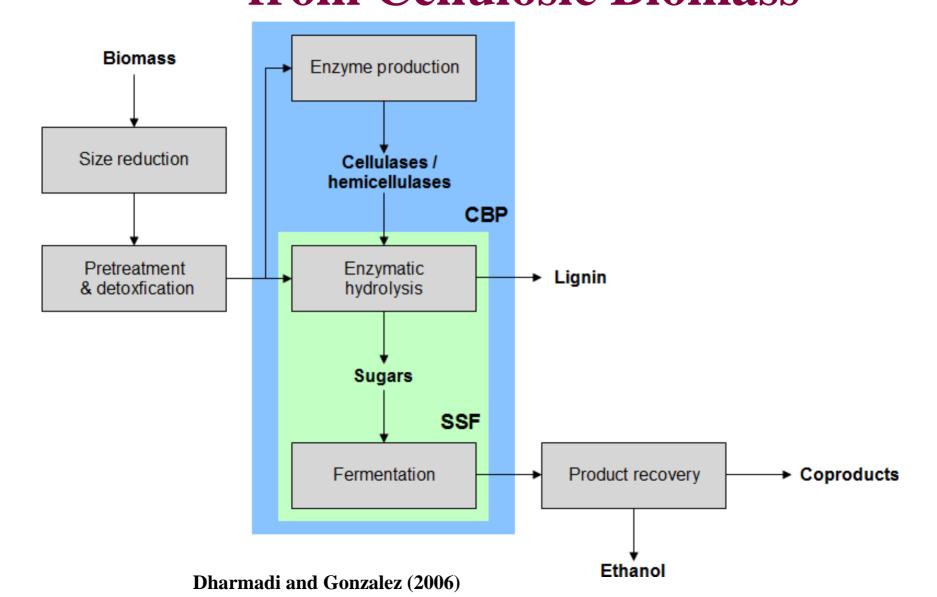
Glucose

Cellulose: Glucose Hemicellulose: Xylose, arabinose, mannose, glucose, etc.



#### EtOH/BuOH...Fuels or Chemicals from Cellulosic Biomass







## Fuels & Chemicals from Biorenewables



Plant Biotechnology Feedstock Deconstruction

Pretreatment and Hydrolysis

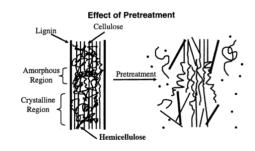
Microbial Fermentation

**Biocatalysis** 

Recovery

Distillation Extraction

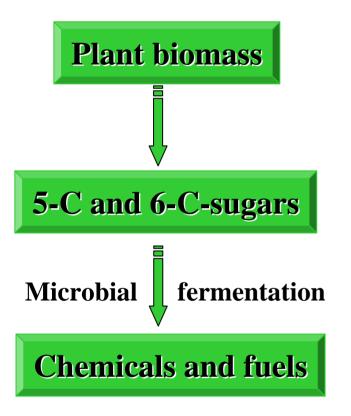




QuickTime\*\* and a TIFF (Uncompressed) decompressor are needed to see this picture. QuickTime™ and a F (Uncompressed) decompressor are needed to see this nicture

### Conversion of Plant Biomass Sugars into Fuels & Chemicals via Fermentation





Escherichia coli, Zymomonas, Lactobacillus, Clostridium, Saccharomyces, Xylose-

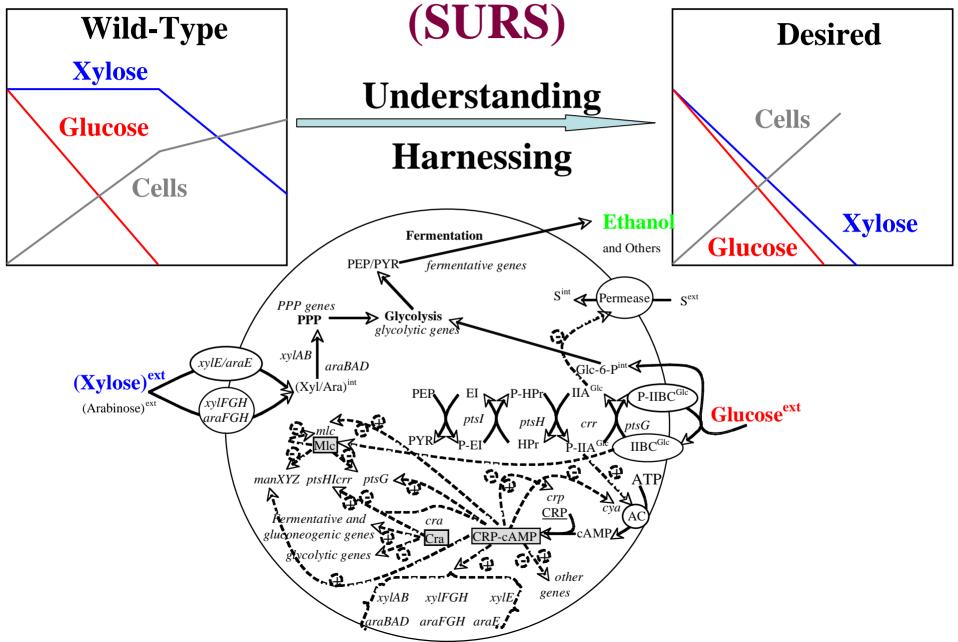
**Assimilating Yeasts** 

#### Search for the ideal biocatalyst

- Broad substrate utilization range (C5/C6)
- Co-fermentation of C5/C6
- High titer, yield, and productivities
- High tolerance to final product
- Minimal nutrients requirements
- Resistance to inhibitors in hydrolysates
- No oxygen requirement
- Available genetic tools and physiological knowledge
- Low fermentation pH
- Use in large-scale fermentations and production on an industrial level

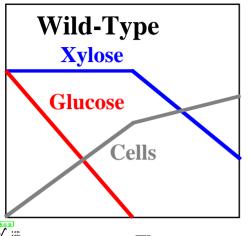
#### Sugar-Utilization Regulatory Systems Sugar-Utilization Regulatory Systems



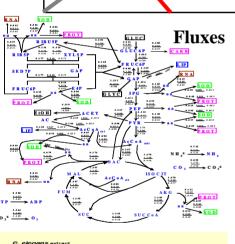


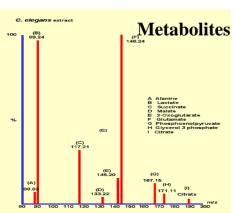
#### Systems Biology-Based Approach

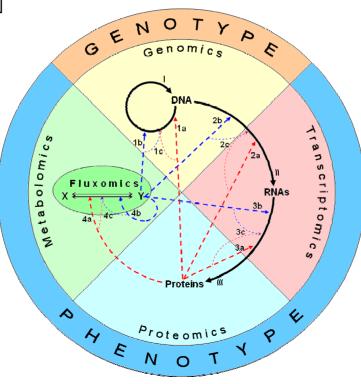


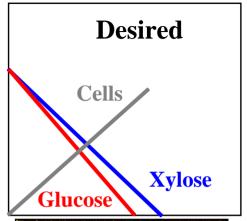


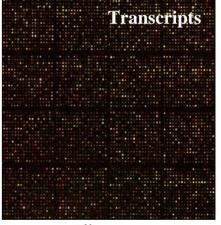
Systems Biology
Approach to
Understand SURS

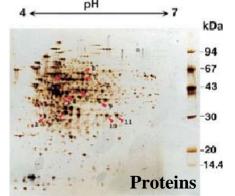








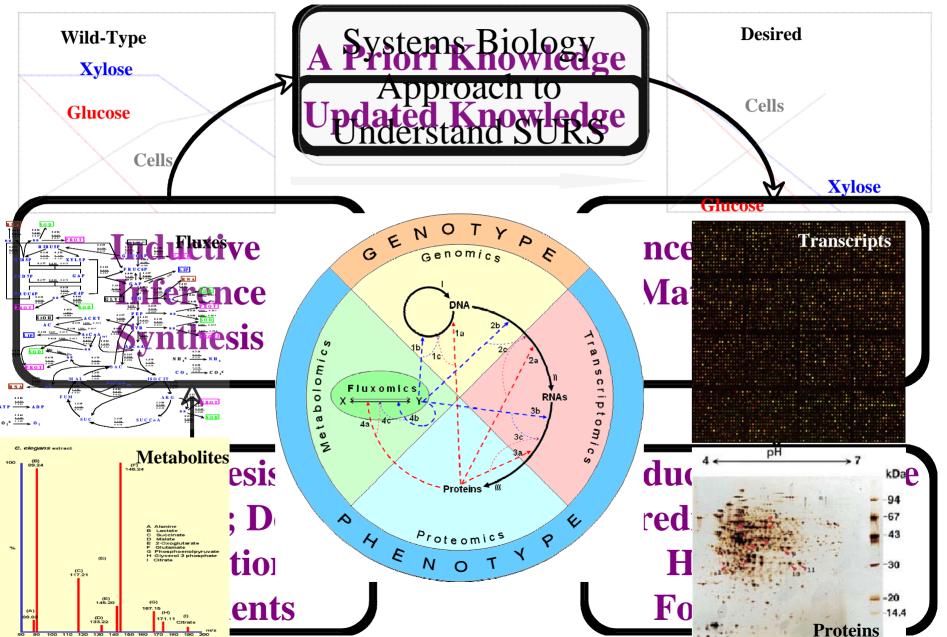




QuickTime™ and a FF (Uncompressed) decompressor are needed to see this picture.

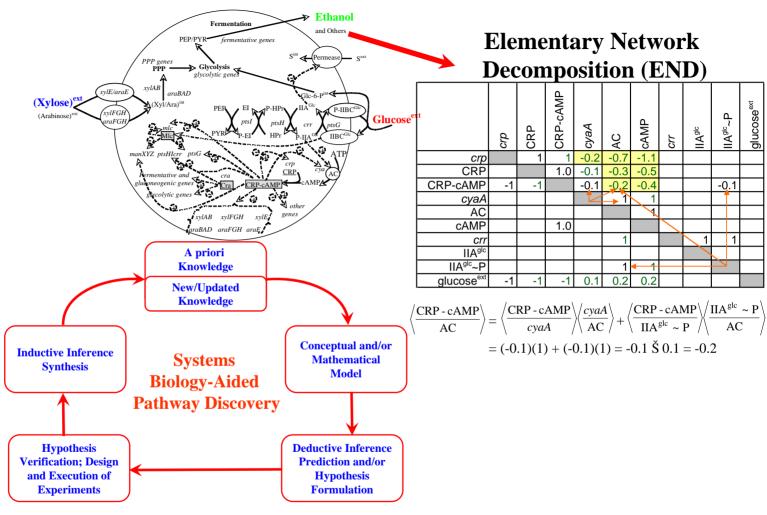
#### Systems Biology-Based Approach





#### **SURS** and **Ethanol Tolerance**





#### Elementary Network Decomposition (END)



#### To elucidate the topology of a regulatory network

- The network is broken down to the smallest elements: genes, proteins, metabolites, phoshorylated/unphosphorylated species, allosteric effectors, transcriptional repressors/activators, etc.
- By arrangement of these n elements into a  $n \times n$  table, the network is then rebuilt from the ground up through complete enumeration of all possible n(n-1) interactions.
- This is easily done using a spreadsheet, which also prevents circular reference in building inferences, and accommodates network expansion/reduction by simple row and column insertion/deletion.

#### **Elementary Network Decomposition**



(END)

	crp	CRP	CRP-cAMP	cyaA	AC	сАМР	crr	IIA <sup>glc</sup>	IIA <sup>glc</sup> ∼P	glucose <sup>ext</sup>
crp		1	1	-0.2	-0.7	-1.1				
CRP			1.0	-0.1	-0.3	-0.5				
CRP-cAMP	-1	-1		-0.1	-0.2	-0.4			-0.1	
cyaA					个	1				
AC						7				
cAMP			1.0							
crr					1			1	1	
IIA <sup>glc</sup>		·		·	·	·				
IIA <sup>glc</sup> ~P					1	1				
glucose <sup>ext</sup>	-1	-1	-1	0.1	0.2	0.2				

$$\left\langle \frac{\text{CRP-cAMP}}{\text{AC}} \right\rangle = \left\langle \frac{\text{CRP-cAMP}}{\text{cyaA}} \right\rangle \left\langle \frac{\text{cyaA}}{\text{AC}} \right\rangle + \left\langle \frac{\text{CRP-cAMP}}{\text{IIA}^{\text{glc}}} \right\rangle \left\langle \frac{\text{IIA}^{\text{glc}} \sim \text{P}}{\text{AC}} \right\rangle$$
$$= (-0.1)(1) + (-0.1)(1) = -0.1 - 0.1 = -0.2$$

#### "END" Features



- (1) All knowledge about regulatory interactions in the network can be catalogued/summarized/integrated within a compact representation.
- (2) Construction of END requires no aesthetic effort in spatial arrangement of the network elements.
- (3) New elements and knowledge can be easily appended, and conversely, smaller subsets of the network can be easily extracted.
- (4) Correlations between network elements are self-consistent and explicitly presented.
- (5) Once mechanistic interactions are set as highest-level entries, other interactions can be inferred in a self-consistent manner through simple rules of arithmetic multiplication and addition.
- (6) END can show how exertion of regulatory control can propagate through cooperativity with lower-level interactions.
- (7) END is potentially useful as discovery tool for uncovering hidden regulatory structures: emerging properties.

#### **END of SURS: Role of Mlc**



	crp	CRP	CRP-cAMP	cyaA	AC	сАМР	crr	IIA <sup>glc</sup>	d∼ <sub>ɔlô</sub> ∀II	glucose <sup>ext</sup>	ptsl	EI	El~P	PEP	Pyruvate	hstd	HPr	HPr∼P	ptsG	IIB <sup>glc</sup>	IIB <sup>glc</sup> ∼P	G6P	xyl	lacY	LacY	mlc	MIc
crp		1	1	-0.2	-0.7	-1.1																					
CRP			1	-0.1	-0.3	-0.5																					
CRP-cAMP	-1	-1		-0.1	-0.2	-0.4			-0.1														1	1			
cyaA					1	1																					
AC						1																	1				
cAMP			1																				1				
crr					1	1		1	1																		
IIA <sup>glc</sup>									-1											-1	1				-1		
IIA <sup>glc</sup> ~P					1	1		-1									1	-1					1				
glucose <sup>ext</sup>	-1	-1	-1	0.1	-0.8	-1.8		1	-1											-1	1	-1	-1	-1	-2		
ptsl			1		1	1			1			1	1														
EI													-1				-1	1									
EI~P									1			-1		-1	1			-1									
PEP																											
Pyruvate												-1	1	-1													
ptsH			1		1	1			1								1	1									
HPr								-1	1									-1									
HPr~P									-1			1	-1				-1										
ptsG			-1		-1	-1			-1											1	1		-2				
IIB <sup>glc</sup>										-1											-1	1					
IIB <sup>glc</sup> ∼P								1	-1											-1							
G6P										-1										1	-1						
xyl																											
lacY																									1		
LacY																											
mlc																							2				1
Mlc																			-1				2				

#### **END of SURS: Role of Mlc**

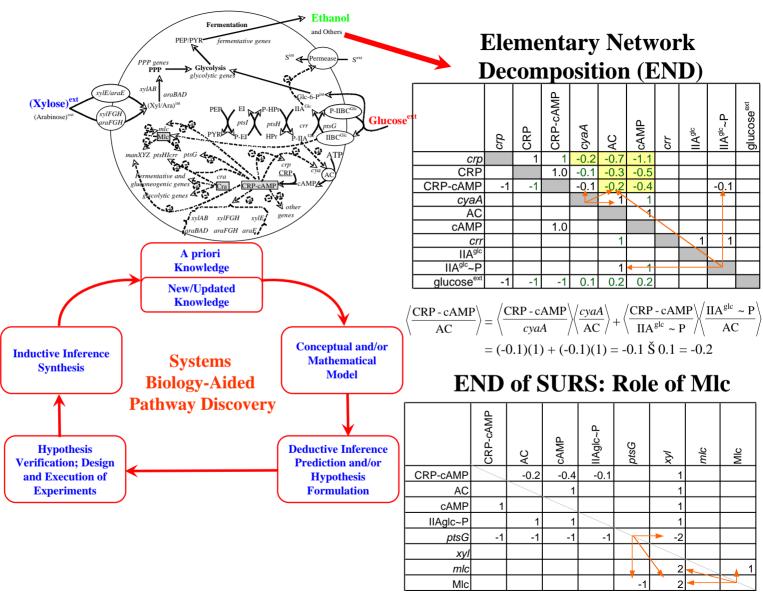


	CRP-cAMP	AC	сАМР	IIAglc~P	ptsG	xyl	mlc	MIc
CRP-cAMP		-0.2		-0.1		1		
AC			1			1		
cAMP	1					1		
IIAglc~P		1	1			1		
ptsG		-1	-1	-1		-2		
xyl								
mlc						2		1
Mlc					-1	2		

$$\left\langle \frac{\Delta mlc}{xyl} \right\rangle = \left\langle \frac{\Delta mlc}{mlc} \right\rangle \left\langle \frac{mlc}{Mlc} \right\rangle \left\langle \frac{Mlc}{ptsG} \right\rangle \left\langle \frac{ptsG}{xyl} \right\rangle = (-1)(1)(-1)(-2) = -2$$

#### **SURS** and **Ethanol Tolerance**

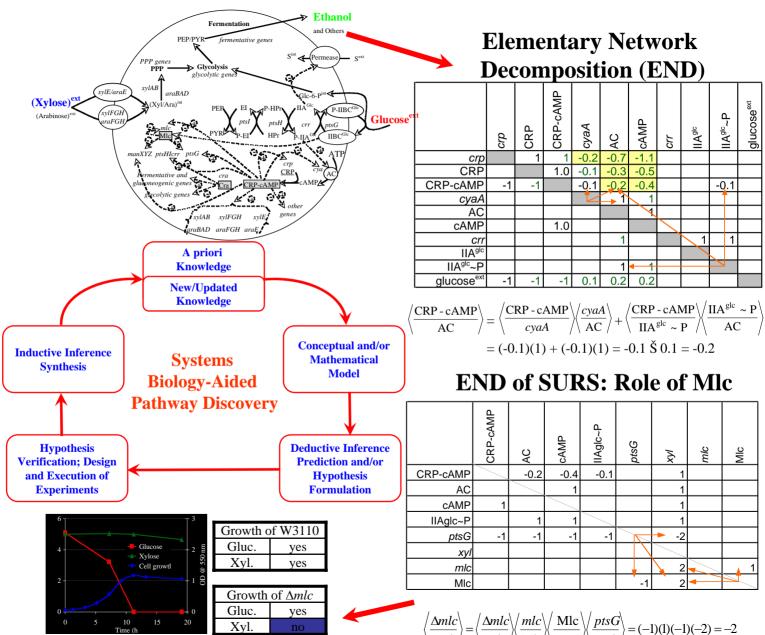




$$\left\langle \frac{\Delta mlc}{xyl} \right\rangle = \left\langle \frac{\Delta mlc}{mlc} \right\rangle \left\langle \frac{mlc}{Mlc} \right\rangle \left\langle \frac{mlc}{ptsG} \right\rangle \left\langle \frac{ptsG}{xyl} \right\rangle = (-1)(1)(-1)(-2) = -2$$

#### **SURS** and **Ethanol Tolerance**



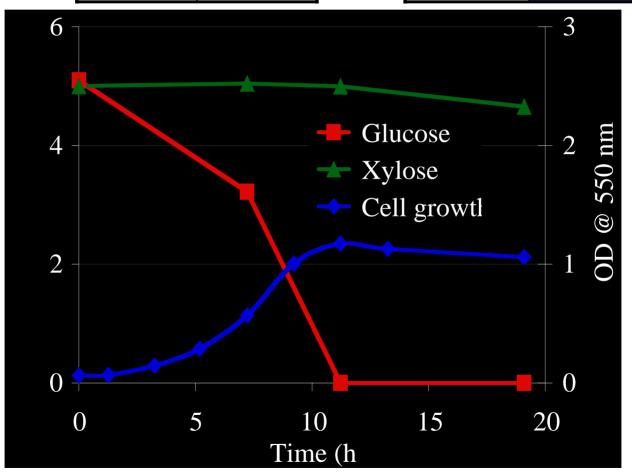


#### Role of Mlc on Xylose Utilization



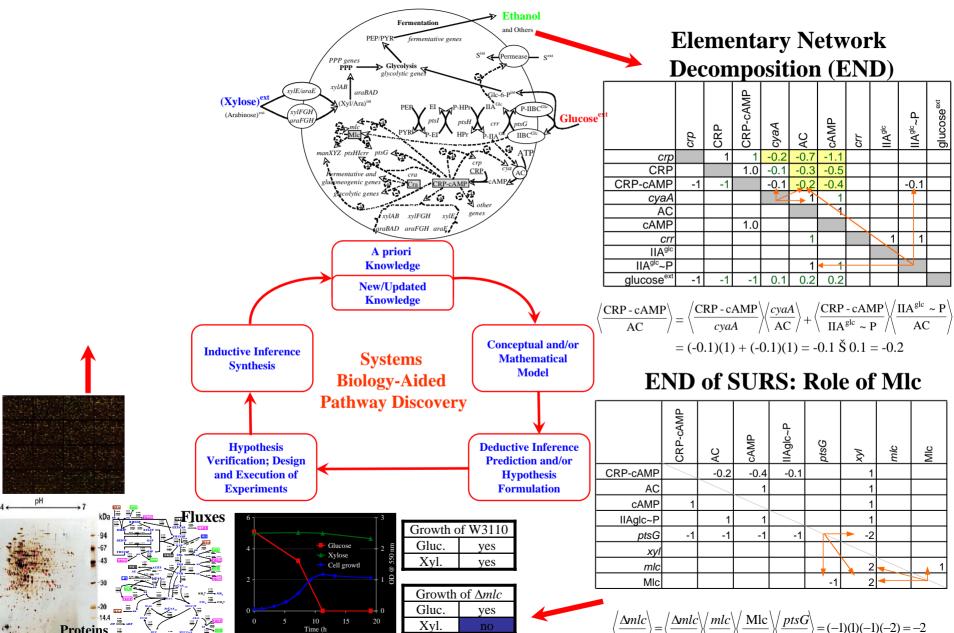
Growth of W3110							
Glucose	yes						
Xylose	yes						

Growth of $\Delta mlc$						
Glucose	yes					
Xylose	no					



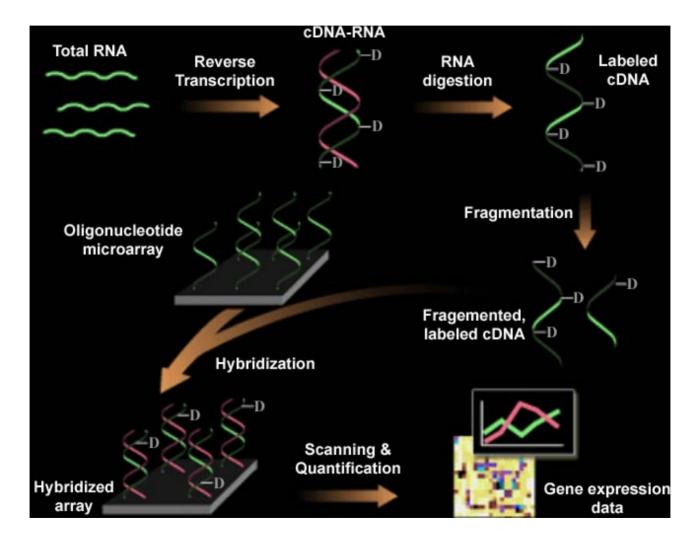
#### **SURS** and **Ethanol Tolerance**





#### DNA Microarrays: Experimental procedure



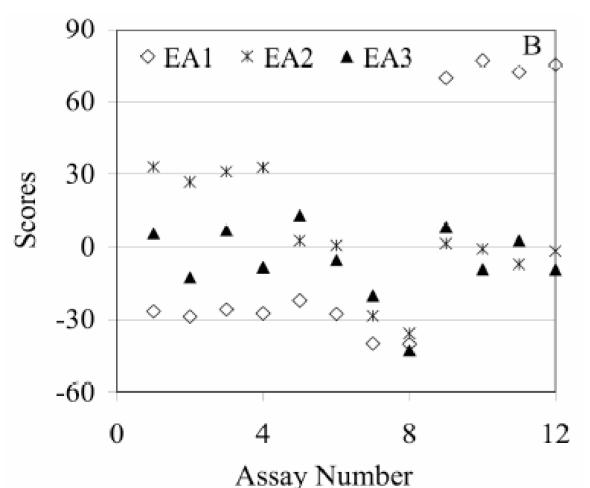


#### A Novel Data Mining Method to Identify Assay-Specific Signatures in Functional Genomic Studies

	Assay characteristics								
Assay	Strain	Sugar	1% EtOH in the initial medium	2% EtOH challenge					
1	KO11	Glucose	NO	NO					
2	LY01	Glucose	NO	NO					
3	KO11	Xylose	NO	NO					
4	LY01	Xylose	NO	NO					
5	KO11	Glucose	YES	NO					
6	LY01	Glucose	YES	NO					
7	KO11	Xylose	YES	NO					
8	LY01	Xylose	YES	NO					
9	KO11	Glucose	NO	YES					
10	LY01	Glucose	NO	YES					
11	KO11	Xylose	NO	YES					
12	LY01	Xylose	NO	YES					

## Assay Contribution: Identifying Associating Between Assays and Principal Components

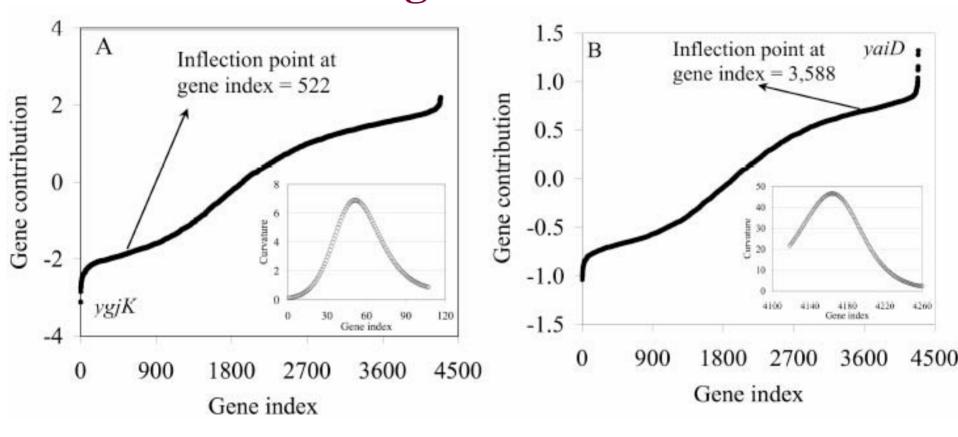




The EAs efficiently identify signatures corresponding to ethanoland non-ethanol-challenged cultures, presence and absence of ethanol in the initial culture medium, and a strain-specific signature

## Identifying Response-to-Ethanol Signature





Complete gene signatures and subset of genes contributing the most to each signature are identified by the existence of points of inflection and maximum curvature, respectively.

## Identifying Response-to-Ethanol Signature

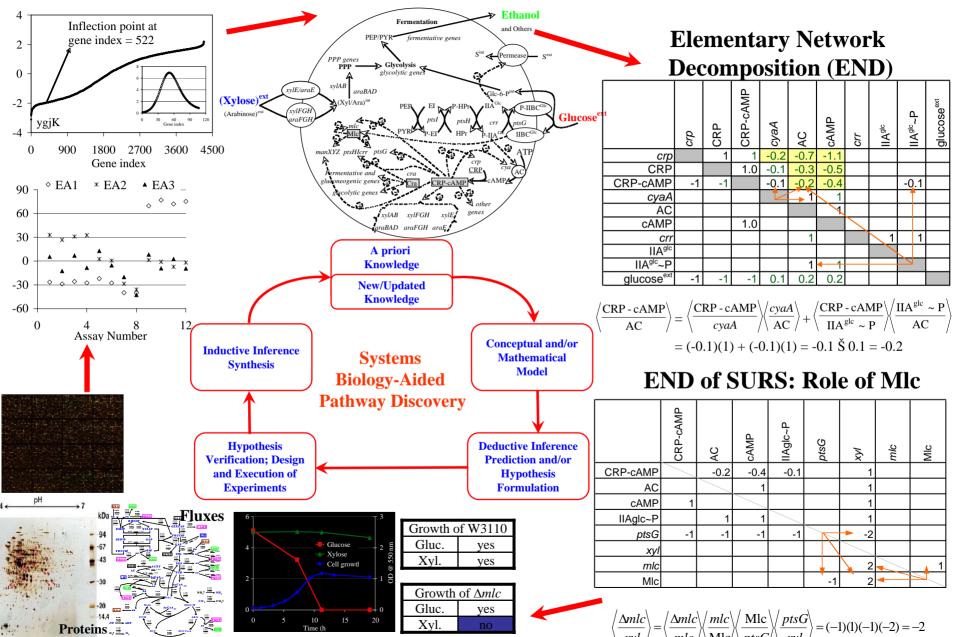


Gene Name	Rank	Expression Ratio	Gene Name	Rank	Expression Ratio
yaiD	1	298.3	ygjK	1	-20.9
argH	2	170.0	tktA	2	-5.2
mngA	3	85.6	dsbC	3	-8.5
plsC	4	76.9	cvrA	4	-2.3
caiA	5	300.3	nrfE	5	-11.9
yebU	6	51.6	yehI	6	-19.2
ylbF	7	29.9	ybbA	7	-4.1
nrfG	8	46.8	evgS	8	-16.2
yaiY	9	27.6	<i>ynfE</i>	9	-14.4
pnp	10	4.9	pqiB	10	-3.7

Many of the top-ranked genes encode functions that one would expect to be involved in the cellular response to an ethanol challenge such as the metabolism and transport of osmolytes (mngA, cvrA, and caiA), the biosynthesis of phospholipids (plsC) which are major constituents of the cell membrane, and the repairing of misfolded proteins (dsbC). In fact, increased tolerance to ethanol in certain E.coli strains is related to the increased availability of osmolytes like betaine and trehalose.

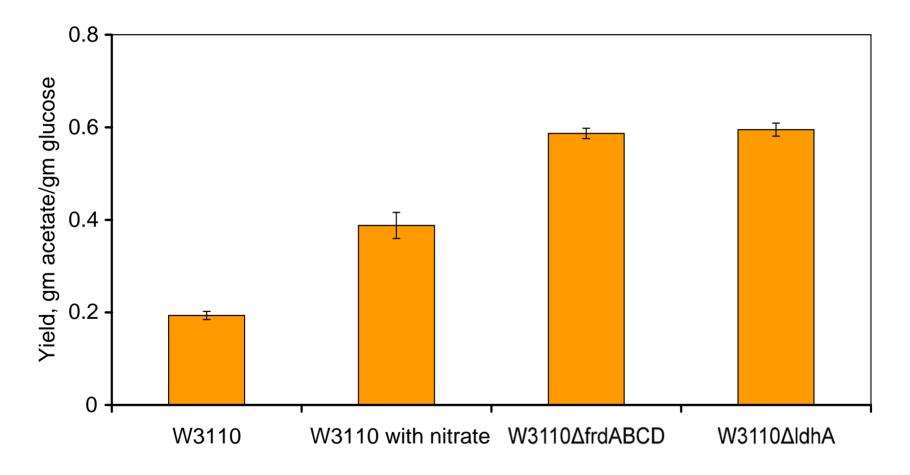
#### **SURS** and **Ethanol** Tolerance





#### Biorefinery Production of Acetate: Yields

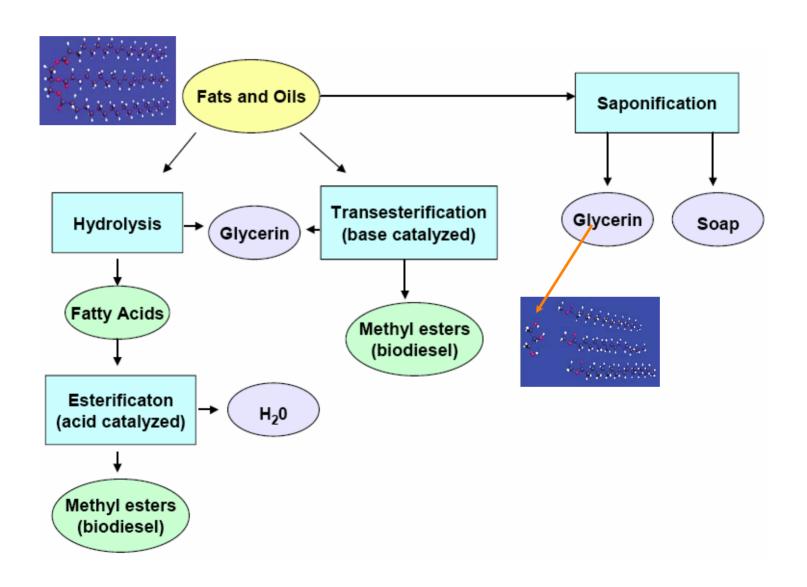




Obtained acetate yields ~90% of the maximum theoretical value Smith et al., (2006). (in prep.)

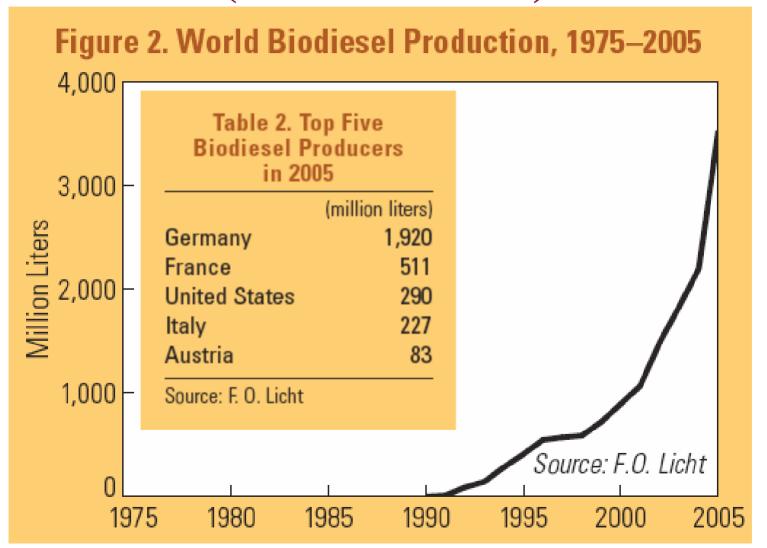
### (8)

#### The "Oil" Platform



## World Production of Biodiesel (Million liters)





Biofuels For Transportation: Global Potential and Implications for Sustainable Agriculture and Energy in the 21<sup>st</sup> Century. Worldwatch Institute and German Agency for Technical Cooperation: www.worldwatch.org/pubs/biofuels.

QuickTime™ and a FF (Uncompressed) decompressor are needed to see this picture

#### It is Not a Garage Industry!!



QuickTime™ and a TIFF (LZW) decompressor are needed to see this picture.

Holmes, E. T. (Holmes Associates LLC). (2005). Biodiesel Southeast Perspective. Biofuels Workshop & Trade Show in Atlanta, Georgia, October 11, 2005

QuickTime™ and a FF (Uncompressed) decompressor are needed to see this picture

#### Biodiesel Process Flow (Lurgi PSI)



95-100 lb

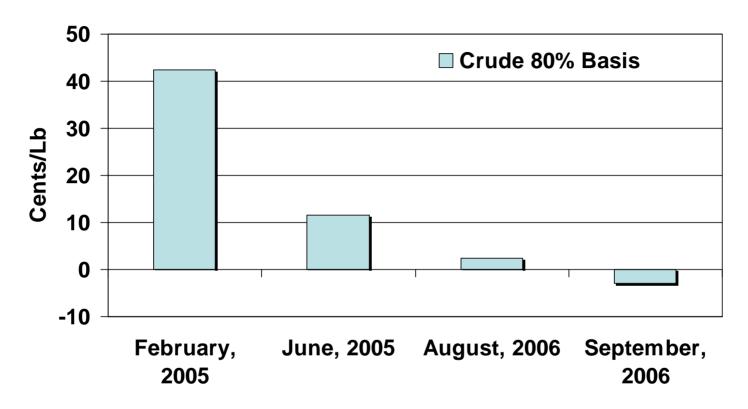
QuickTime<sup>™</sup> and a TIFF (LZW) decompressor are needed to see this picture.

10 lb

Holmes, E. T. (Holmes Associates LLC). (2005). Biodiesel Southeast Perspective. Biofuels Workshop & Trade Show in Atlanta, Georgia, October 11, 2005

#### **Crude Glycerin Prices**





QuickTime™ and a TIFF (Uncompressed) decompressor are needed to see this picture.

Since 1865, The Jacobsen has been recognized as the most insightful and accurate resource for commentary and daily spot market prices for the agricultural by-products industry.



# hnical Repor

#### Biomass Oil Analysis: Research Needs and Recommendations

K. Shaine Tyson, Joseph Bozell, Robert Wallace, Eugene Petersen, and Luc Moens



Operated for the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy by Midwest Research Institute • Battelle

Contract No. DE-AC36-99-GO10337







#### **New Glycerol Platforms**

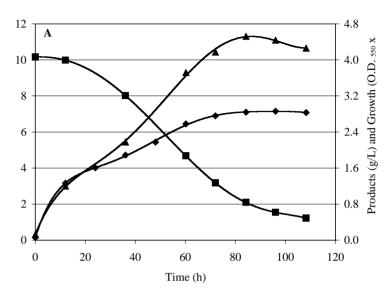
QuickTime<sup>™</sup> and a TIFF (LZW) decompressor are needed to see this picture.

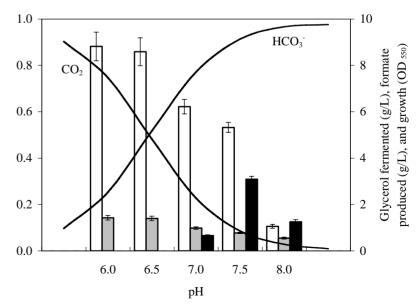
#### Anaerobic Fermentation of Glycerol in *E. coli*



For nearly 80 years, researchers have believed that the metabolism of glycerol in *E. coli* is restricted to respiratory conditions: i.e., an external electron acceptor is needed.

- Quastel JH, Stephenson M (1925). Biochem J. 19:660.
- Quastel JH, Stephenson M, Whetham MD (1925). Biochem J. 14:304.
- Lin EC (<u>1976</u>). Annu Rev Microbiol. 30:535-78.
- Booth I (2005). In Neidhardt FC, Curtiss III R, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (ed.), *E. coli* and Salmonella: cellular and molecular biology, Web Edition. ASM Press, Washington, DC. Available at www.ecosal.org.





#### Acknowledgements



Graduate Students: Undergraduate Students:

Y. Dharmadi K. Smith

A. Gupta M. Yu

B. A. Murarka Collaborators:

V. Rigou

J.V. Shanks, K-Y. San, M. Lamn,

R. Sun D. Rollins

CSREES

INCLUME

USDA

STREET

USDA - National Research Initiative of the USDA Cooperative State Research, Education and Extension Service (2005-35504-15222/16698)

NSF - Division of Bioengineering and Environmental Systems (BES-



0331388/0601549)



US Department of Energy (DEFG3603GO13159)