

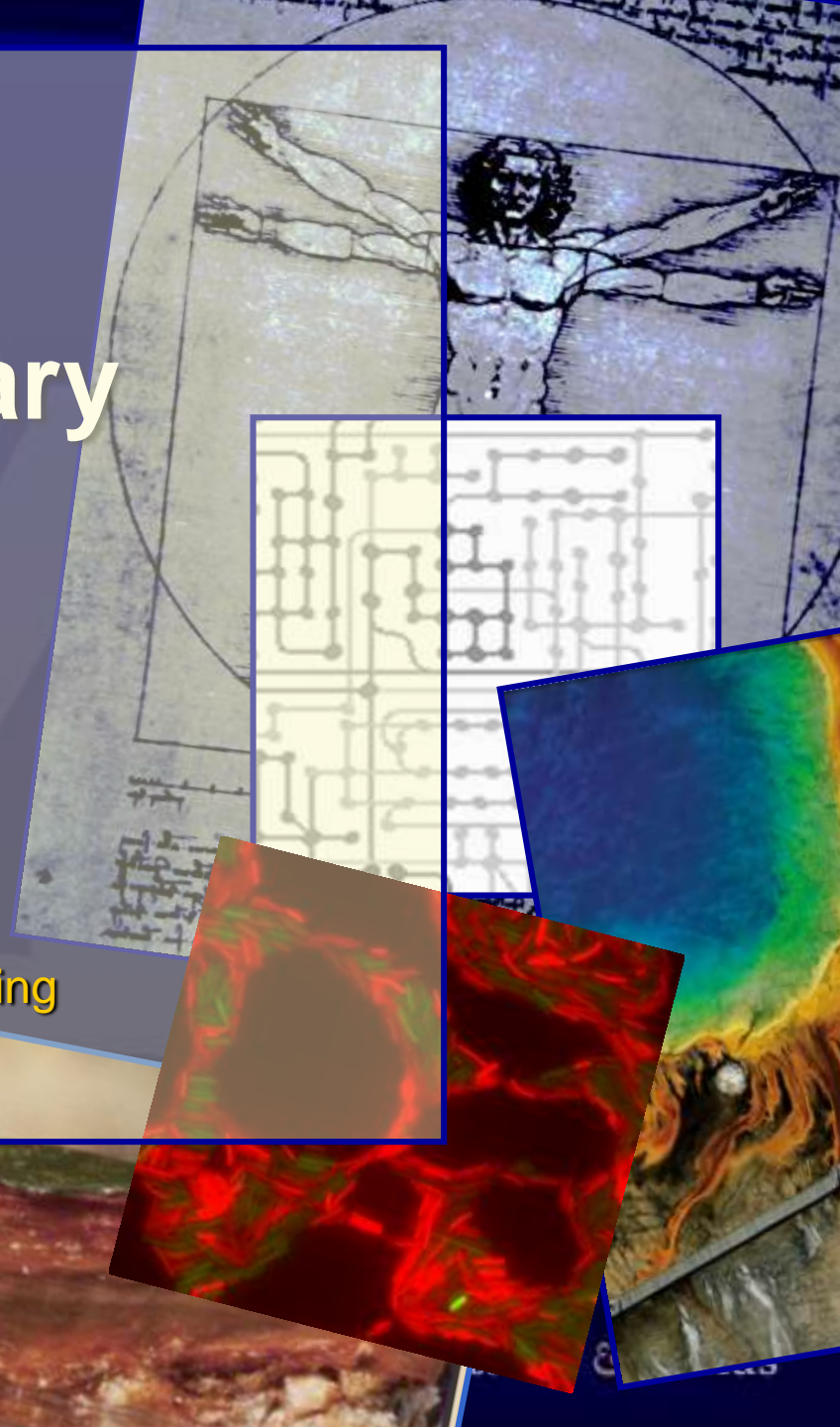
Practical Metabolic Network Analysis Workshop: Elementary Flux Mode Analysis

Ross Carlson

Department of Chemical and Biological Engineering
Montana State University



MONTANA STATE UNIVERSITY



Objective:

- ***Provide a practical toolbox of knowledge and techniques to permit original metabolic network analysis research***



Workshop Contents:

Presentation and Instruction Manual

***programs,
input files,
templates,
exercise files***

Journal Articles



Workshop Outline:

Creating/running models

Analyzing output

Integrating experimental data

Interpreting experimental data

Optional topics:

Microbial community analysis

^{13}C fluxome data interpretation

Recombinant host design



Biochemical Networks

Fundamental chemistries of life are performed by biochemical networks...



MONTANA STATE UNIVERSITY



The Fifth Kingdom, www.pacificcoast.net

Biotechnological processes



Bio-ethanol



NatureWorks™

- PLA



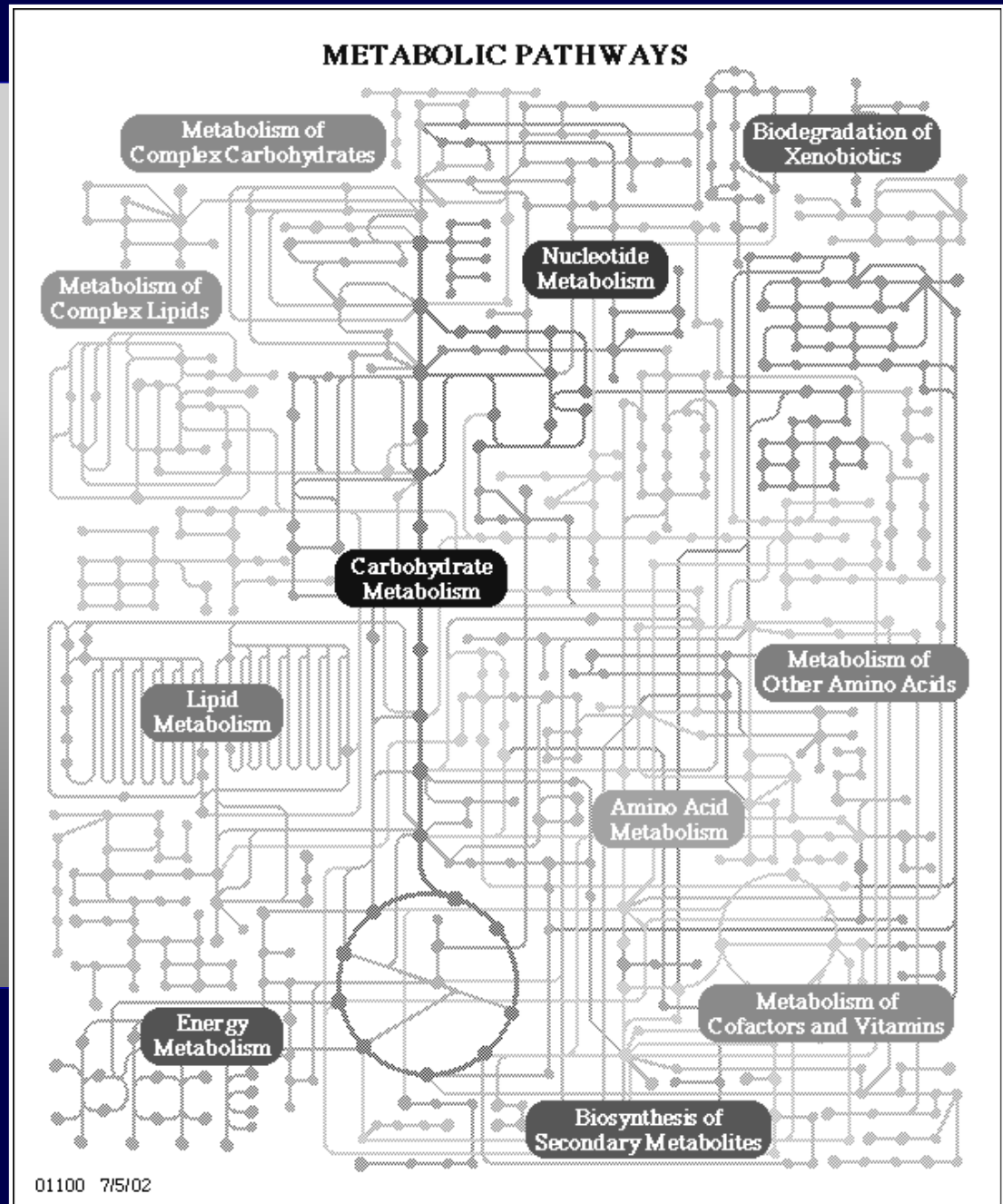
Sorona®

- 1,3-Propanediol



Mountains & Minds

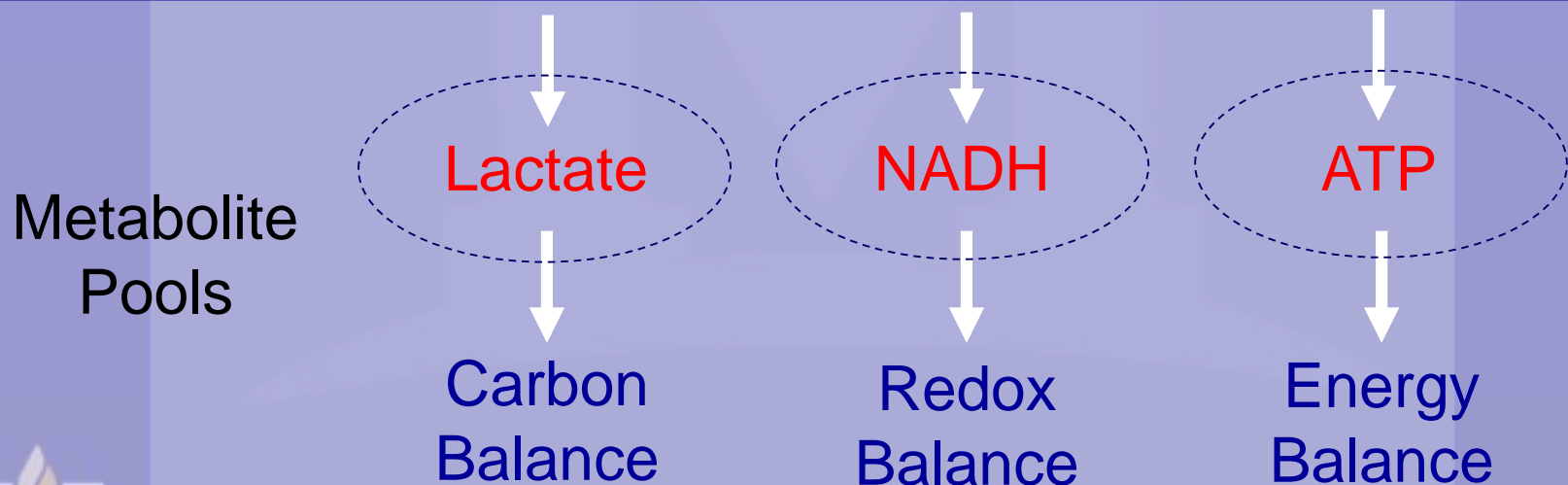
...however,
metabolic
networks
are highly
branched
and highly
coupled.



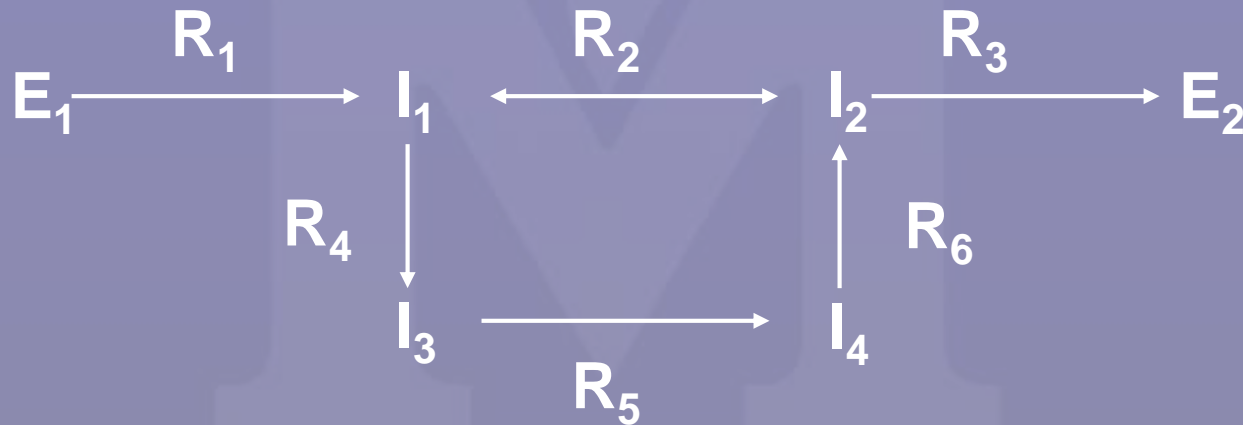
Elementary Mode Analysis

Elementary mode: stoichiometric systems biology approach based on mass and energy balances – identifies mathematically defined ‘pathways’

(Schuster *et al* 1994; 2000)



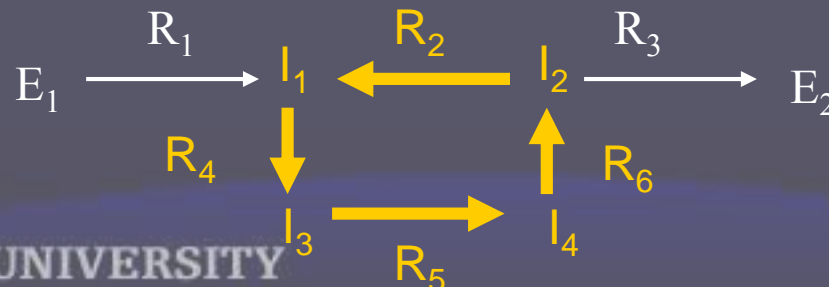
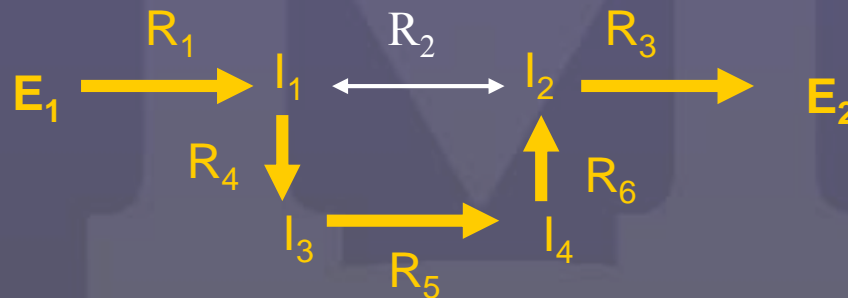
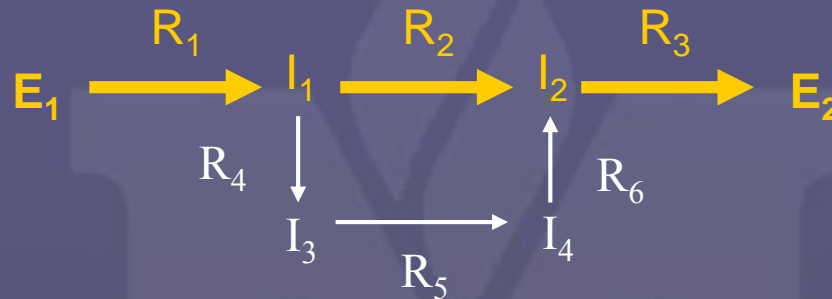
Elementary Flux Modes: Example



Schuster *et al.*, 1996



Elementary Flux Modes: Example



Key Concepts of Systems Biology:

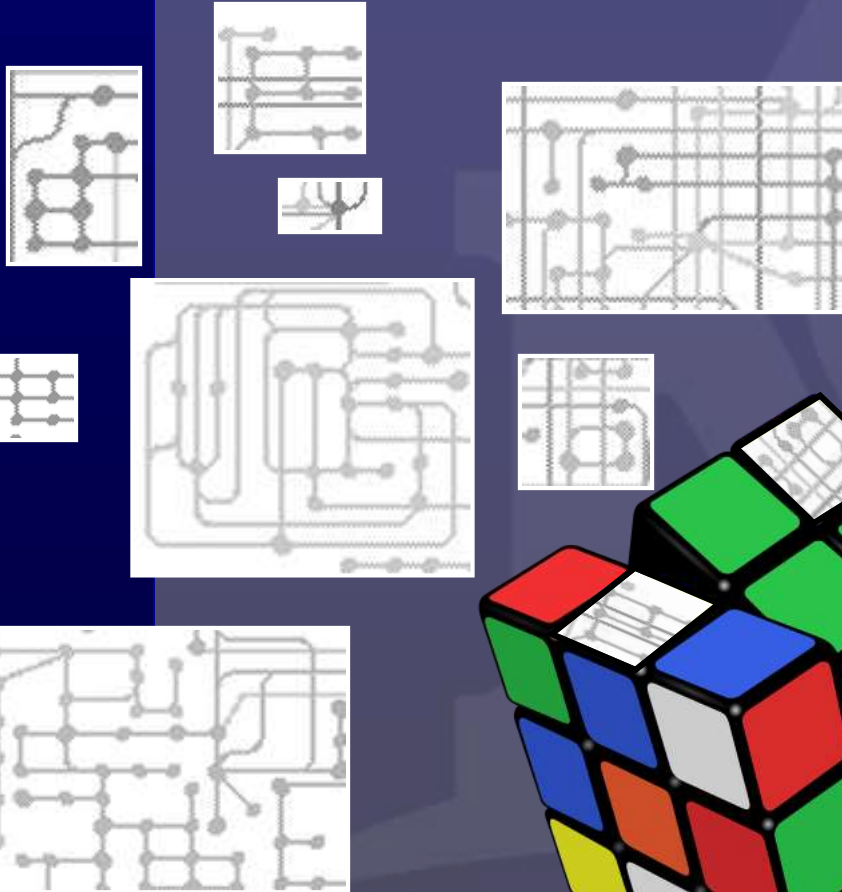
Emergent properties: 'Whole is greater than a sum of parts'

$$\text{Whole} > \sum \text{components}$$

$$\text{Whole} = \sum f(\text{components}, \underline{\text{interactions}})$$

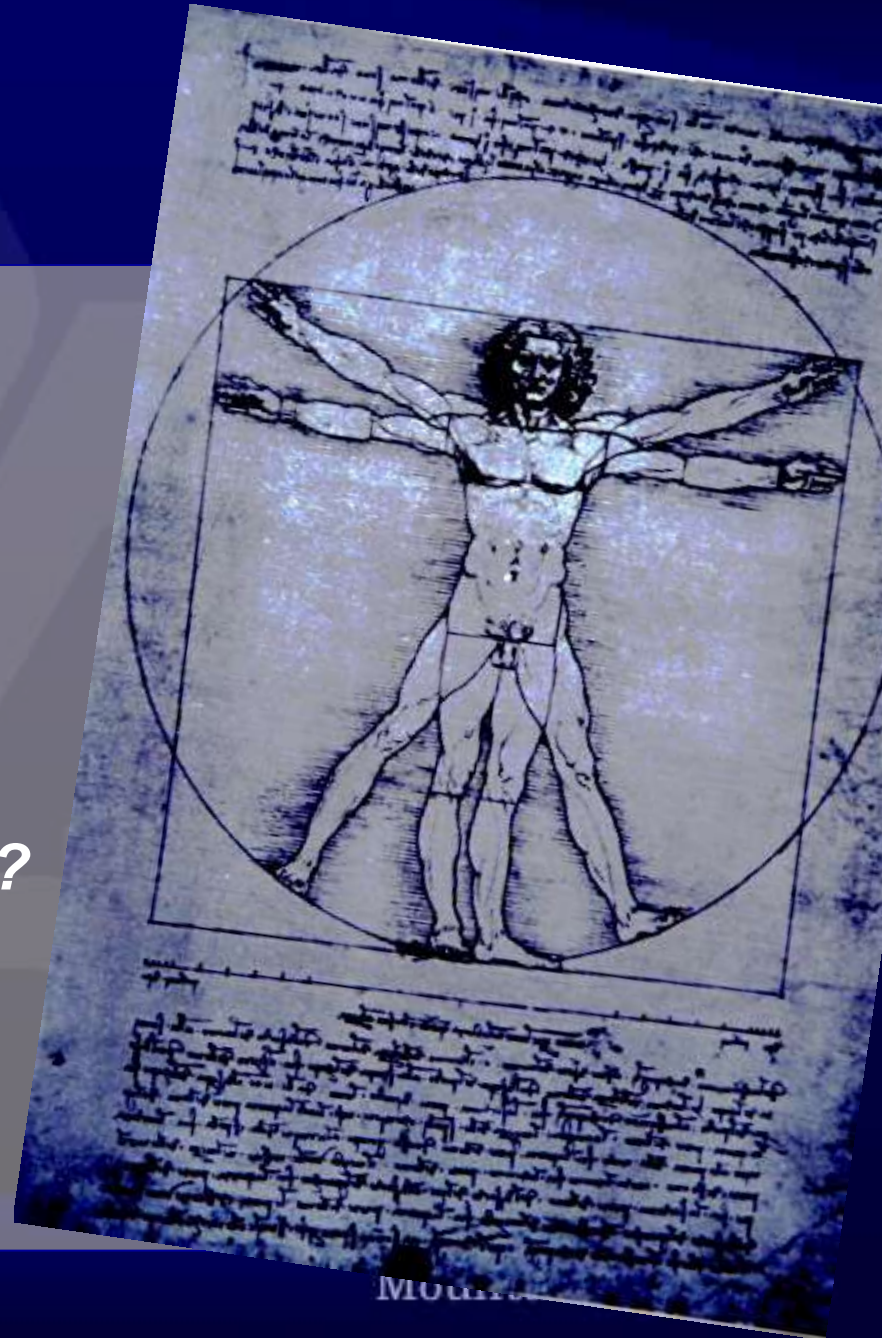


Systems Biology Puzzle

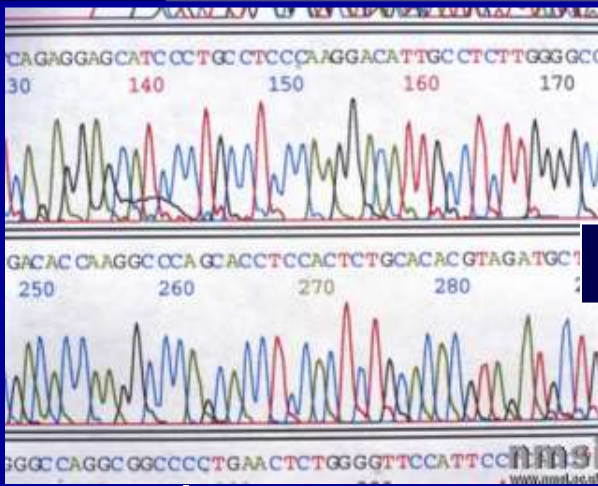


Practical: Sections 1.0-1.3

- *How is a metabolic network constructed?*
- *How is elementary flux mode analysis performed?*
- *What does the output look like?*



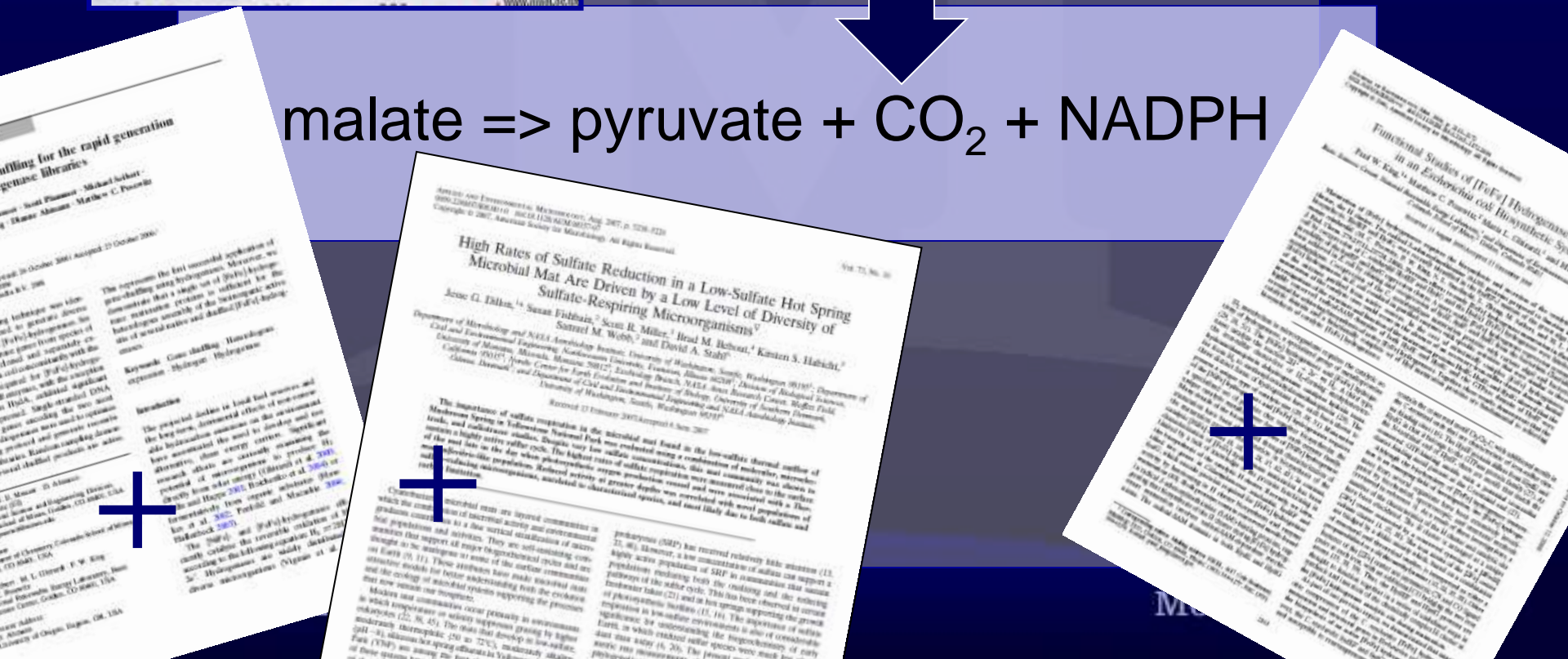
Constructing a model



Sequences producing significant alignments:

	Score (Bits)	E Value
ref NP_416958.1 fused malic enzyme predicted oxidoreductase/...	646 0.0	1e-180
ref YP_408859.1 malic enzyme [Shigella boydii Sb227] >ref YP...	646 0.0	9e-179
ref NP_754870.1 malic enzyme [Escherichia coli CFT073] >ref ...	646 0.0	1e-178
ref NP_311352.1 malic enzyme [Escherichia coli O157:H7 str. ...	646 0.0	1e-178
ref ZP_03064374.1 malate dehydrogenase [Shigella dysenteriae...	646 0.0	1e-178
ref YP_670365.1 malic enzyme [Escherichia coli 536] >gb ABG7...	646 0.0	1e-178
ref YP_002398801.1 putative fused malic enzyme oxidoreductas...	646 0.0	1e-178
ref YP_001881254.1 NADP-dependent malic enzyme [Shigella boy...	644 0.0	1e-178

malate \Rightarrow pyruvate + CO₂ + NADPH



Actual Stoichiometry Matrix

no different than balancing checkbook...



www.darrenbyrne.com

Elementary Flux Mode Analysis Software:

- *METATOOL v4.9, stand-alone executable program*
 - easy set-up and use
 - limited to smaller model sizes
- *CellNetAnalyzer and efmtool*
- *uses different algorithms than METATOOL*
 - requires MATLAB or JAVA
 - can handle larger networks
 - simplifies output processing



References and web addresses in workshop manual

METATOOL Input file:

-ENZREV

Reversible reactions

R2r R5r R6r R7r R8r R11r R12r R13r R14r R15r R23r R26r R27r R28r R29r R53r R54r

-ENZIRREV

Irreversible reactions

R1 R3 R4 R5 RR9 R10 R20 R21 R22 R24 R25 R40 R41 R42 R55 R70 R80 R81 R82 R83

-METINT

Internal metabolites

ATP ADP GLU_6_P FRU_6_P FRU_BIS_P DHAP GA_3P NAD NADH

RIBULOSE_5_P XYL_5_P RIBOSE_5_P SED_7_P ERYTH_4_P PYR PEP CITRATE

OXALO MALATE CoASH ACETYL_CoA FADH FAD AKG ISOCIT ACETATE SUCC

-METEXT

External metabolites

ATP_base GLU_ext ETOH_ext ACETATE_ext CO2_ext LACTATE_ext SUCC_ext

-CAT

Reactions

R1: $GLU_ext + PEP = GLU_6_P + PYR$.

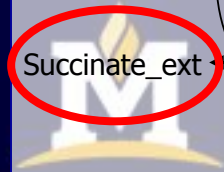
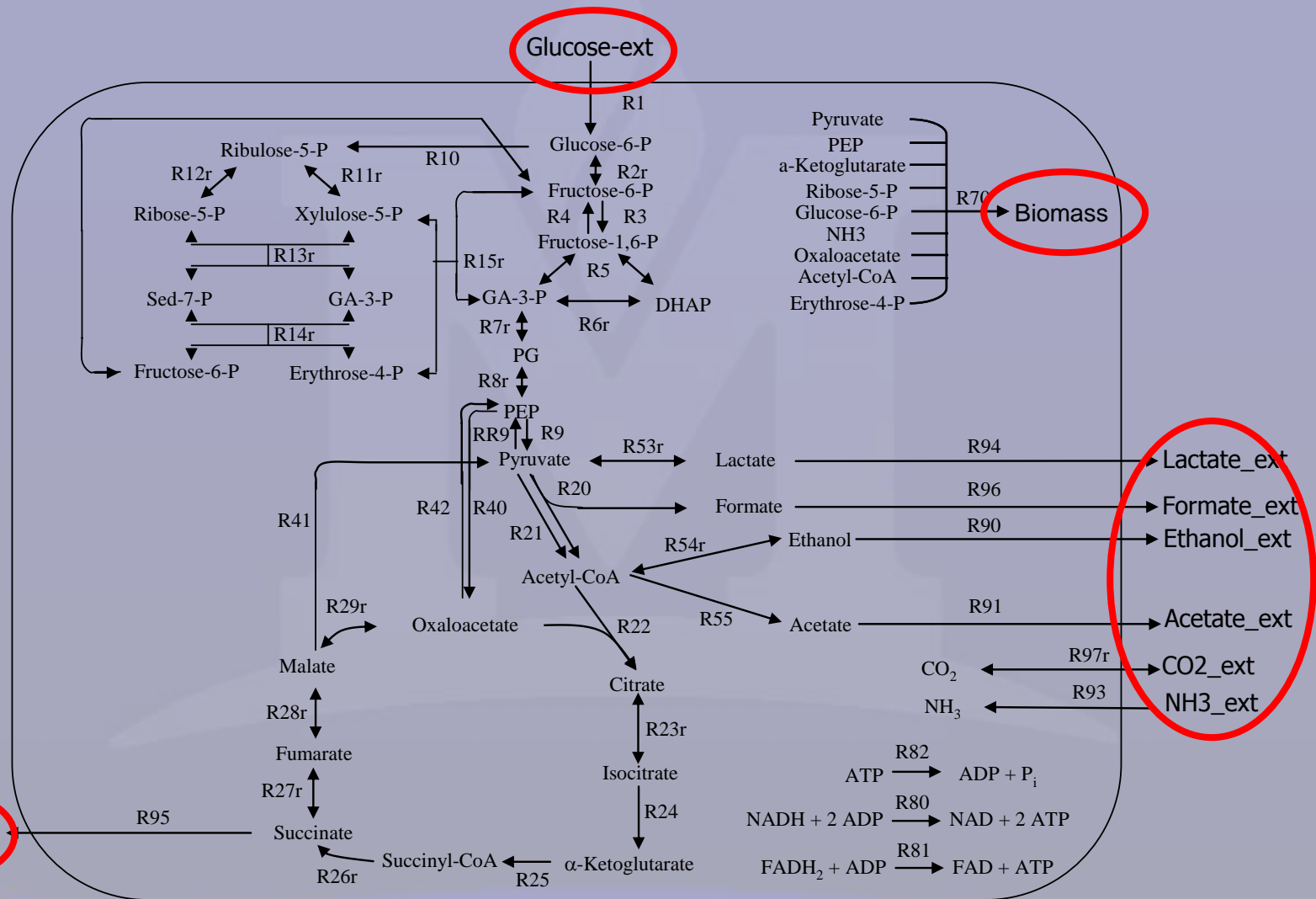
R2r: $GLU_6_P = FRU_6_P$.

R3: $FRU_6_P + ATP = FRU_BIS_P + ADP$.

R4: $FRU_BIS_P = FRU_6_P$.



METATOOL Input file:



Model Biomass Terms

Biomass modeled by considering the draw on 12 central metabolism intermediates at 7 growth rates

Doubling time (min)	Glc-6-P	Rib-5-P	Ery-4-P	PEP	Pyr	AcCoA	α -Kg	Oxalo	CO ₂	NH ₄	NADH	ATP	Total Carbon per Term
200	4	46	31	156	237	72	86	139	-35	731	856	2921	2652
100	4	32	17	91	129	55	47	78	-15	424	487	1674	1554
80	4	22	11	64	87	49	32	53	-10	291	347	1153	1091
60	4	16	7	45	58	44	22	36	-6	198	249	786	777
50	4	14	6	39	50	43	19	31	-4	171	218	678	673
40	4	13	5	34	41	42	16	26	-3	148	192	584	498
30	4	13	5	32	38	41	14	24	-2	139	178	547	471



Elementary Mode Analysis: Output

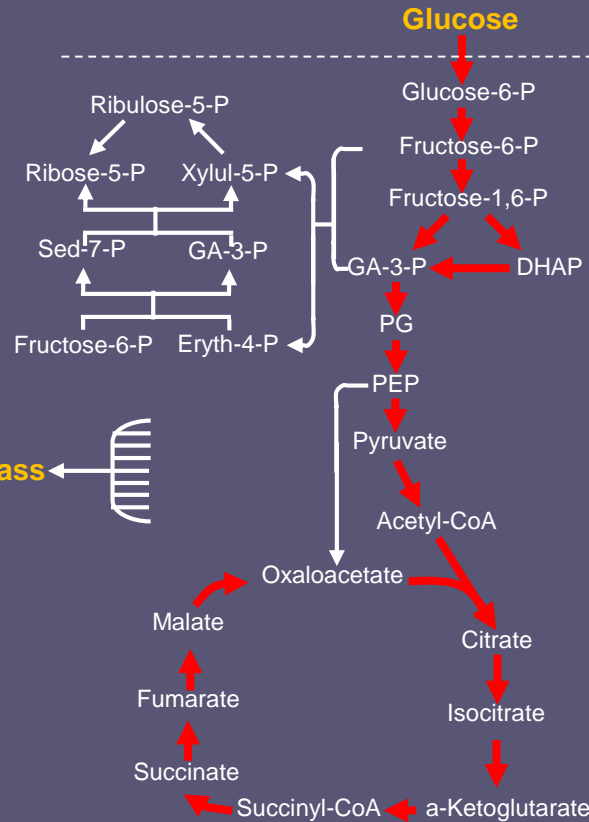
rows = elementary mode

columns = reaction

Sources and sinks

- 1: GLU_ext + 6 C
- 2: GLU_ext + 6 C
- 3: GLU_ext + 6 C
- 4: GLU_ext + 6 C
- 5: GLU_ext + 6 C
- 6: GLU_ext + 6 C
- 7: GLU_ext + 8 C
- 8: 857 GLU_ext +

Biomass



E_ext
TE_ext

ext
_ext
O2_ext
ext + BIOMASS



Central Metabolism Model

Glucose

>3.5 million possible pathways!

91,068 ATP

2,159,153 Biomass



Lactate

Which pathways utilized?

CO₂
Formate
Acetate

Ethanol

O₂

Why?

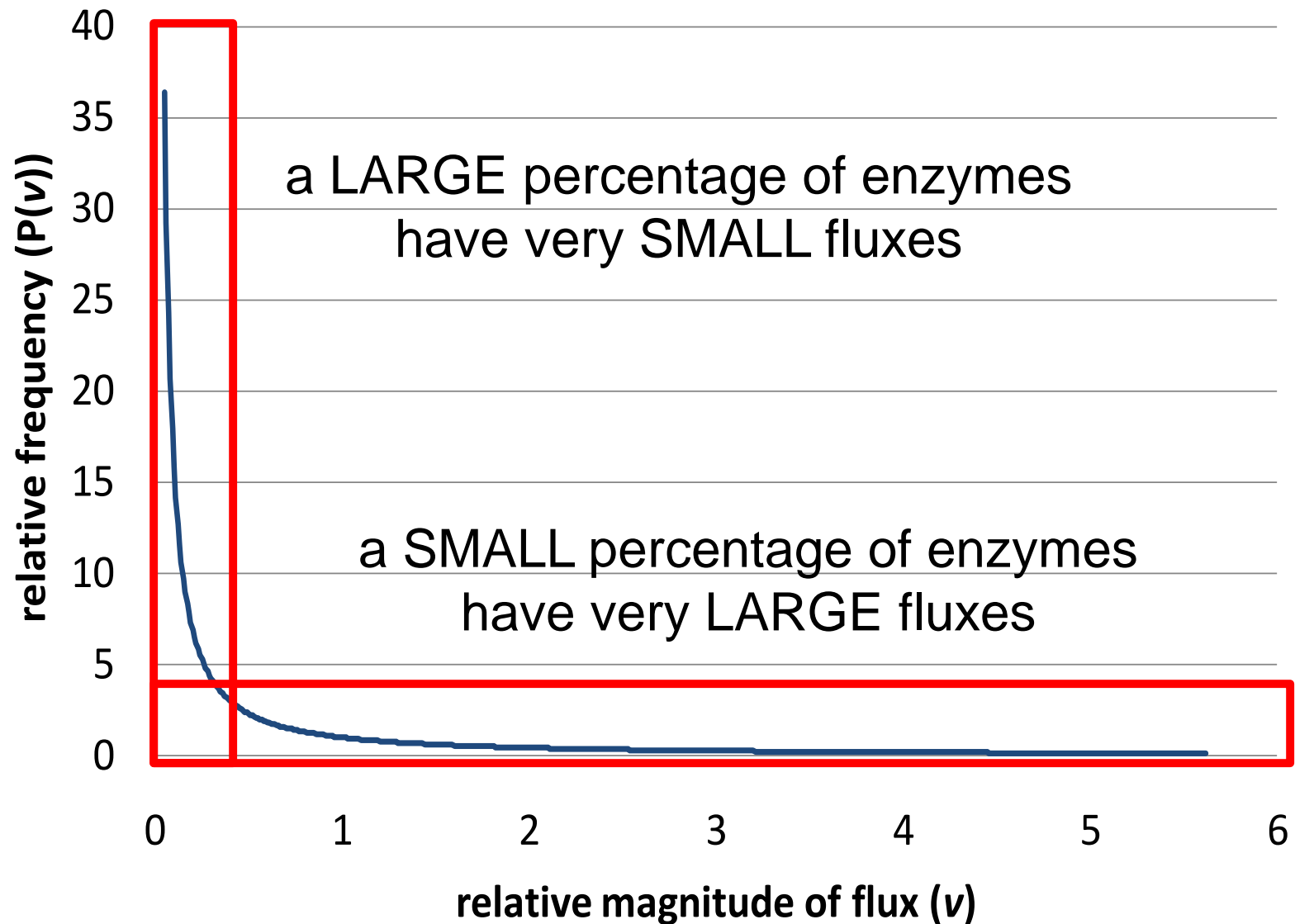
Succinate

NH₃

Glutamate Glutamine

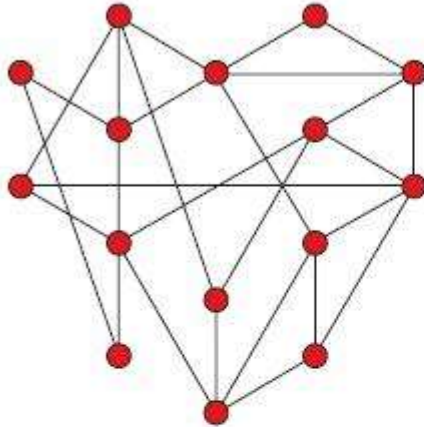


Power Law Relationship



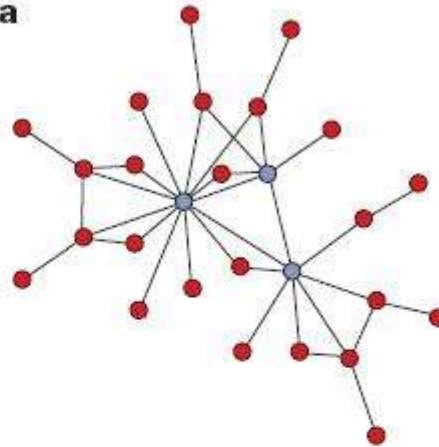
A Random network

Aa

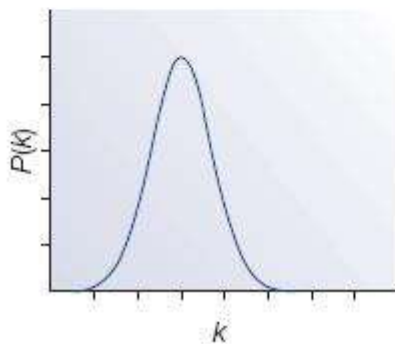


B Scale-free network

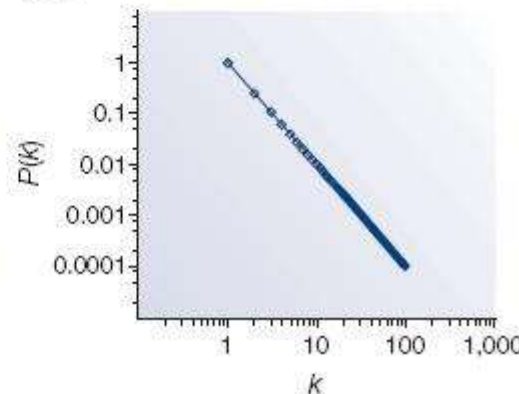
Ba



Ab



Bb



Barabási & Oltvai, 2004

Scale invariance

$$P(k) \sim k^{-\gamma}$$

$$2 < \gamma < 3$$

Other examples:

Fractals

WWW

Airline networks

Social networks

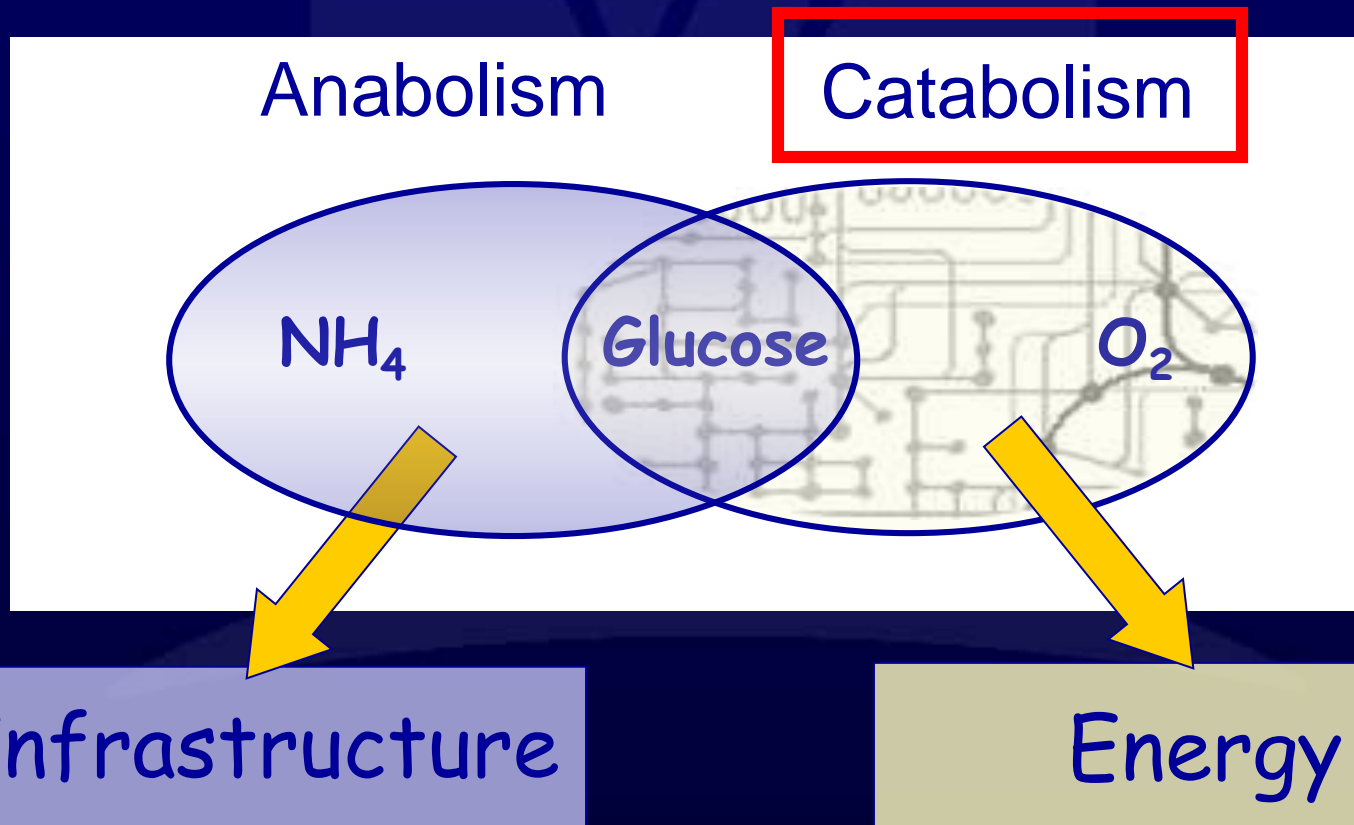
Genes vs. publications

Protein-protein interactions

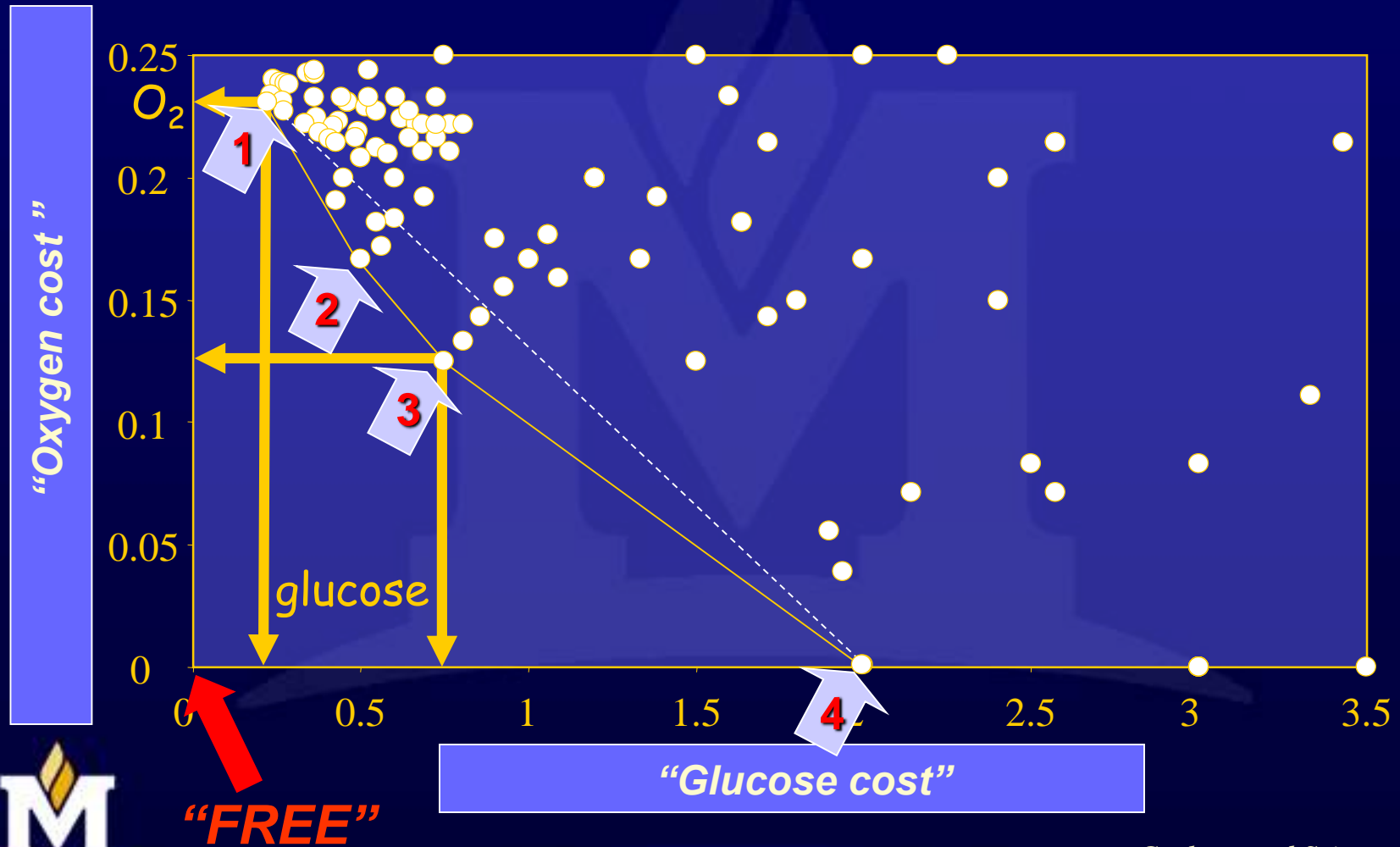
Funding rates vs. investigator



Cell Function Based Analysis



Maintenance Energy Pathways



Analogy:

- *Highly fuel efficient*

*However requires good infrastructure,
good roads, small hills etc*

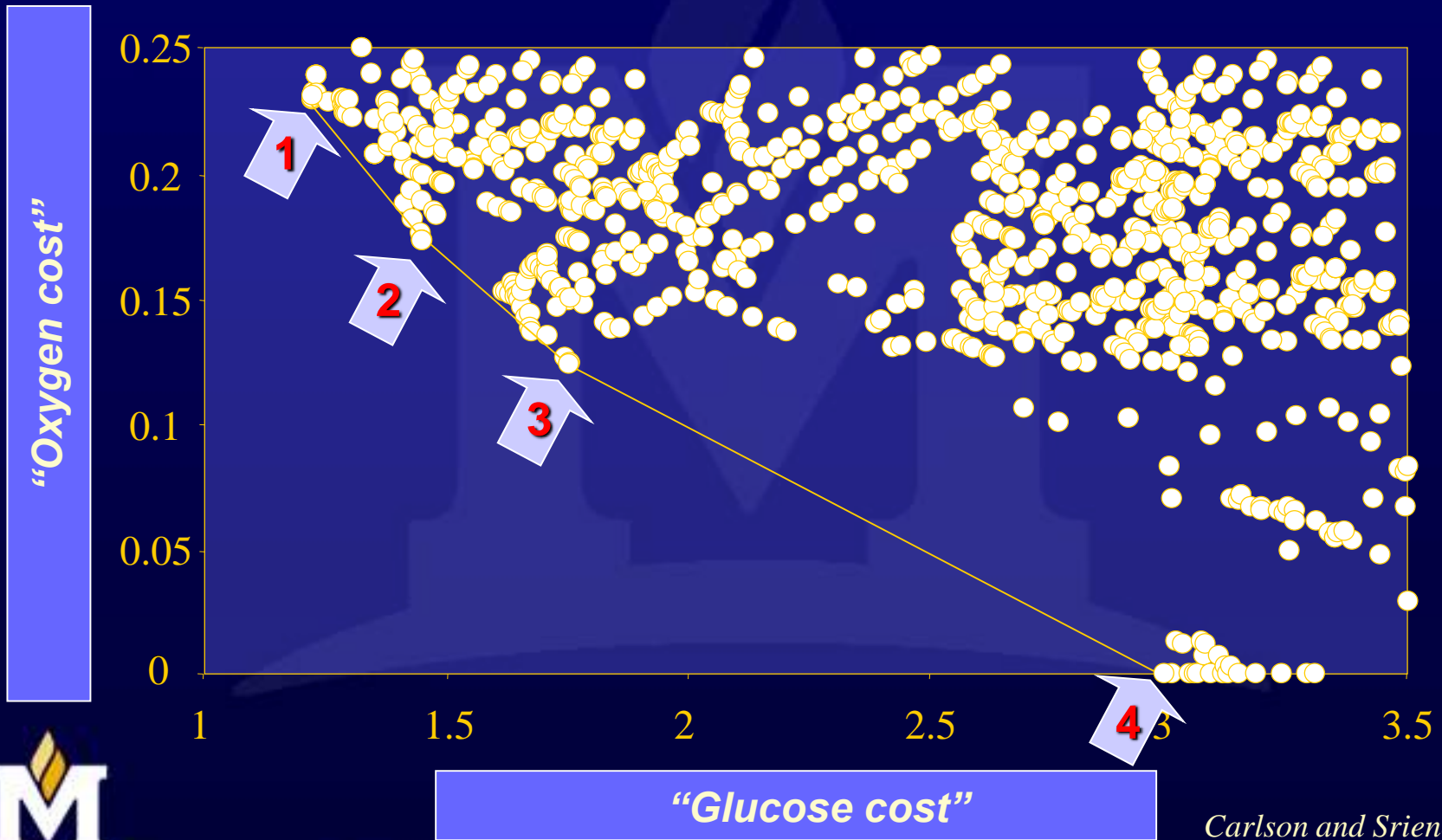


- *Poor fuel efficiency*

*However doesn't require much
infrastructure, what road?, what
river?*



Biomass Pathways

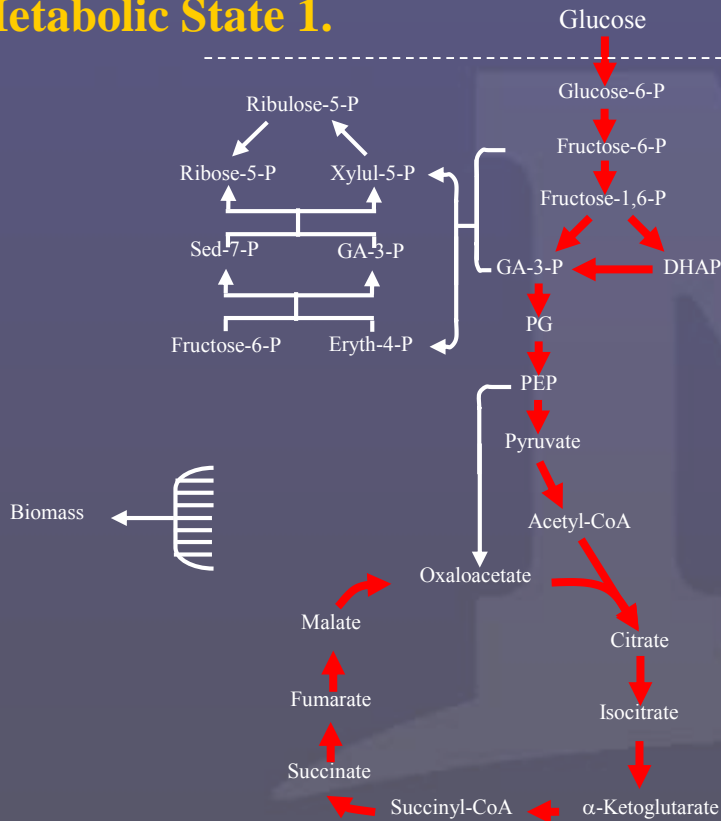


Carlson and Srienc, 2004a

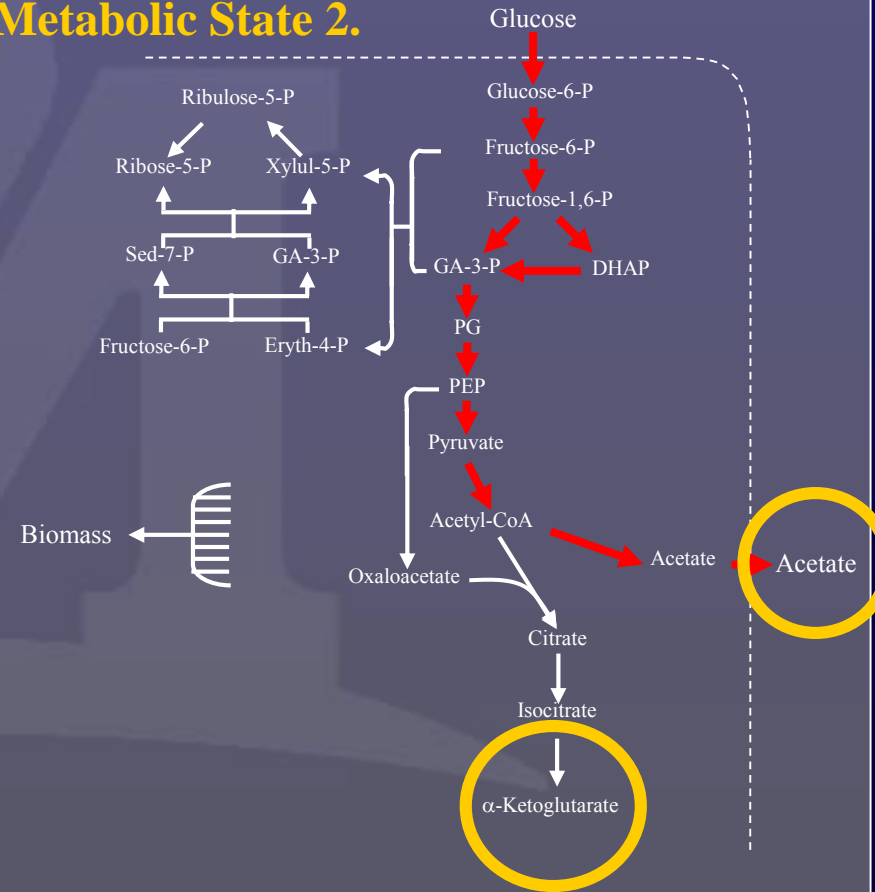


Qualitative Regulation Analysis

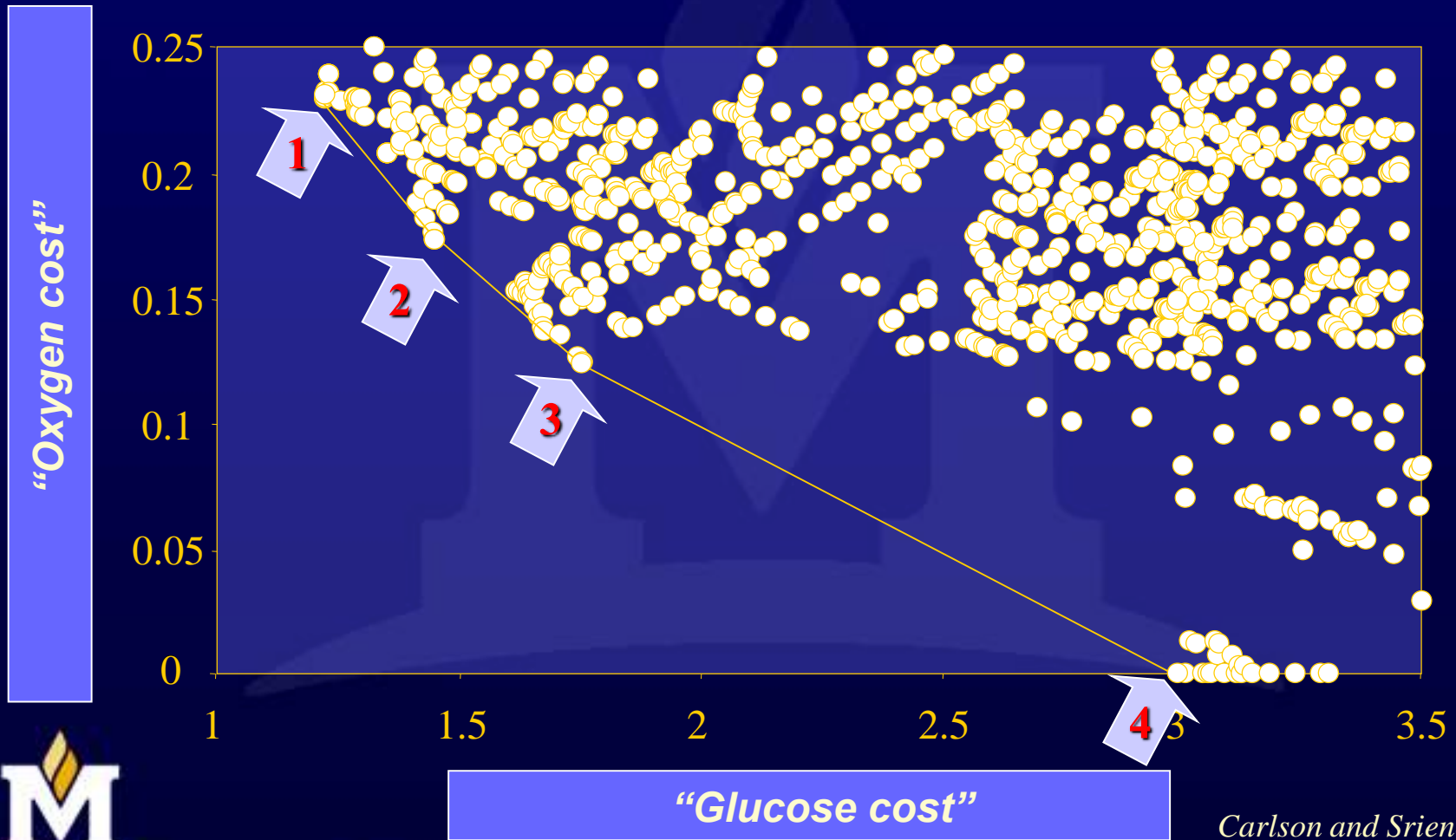
Metabolic State 1.



Metabolic State 2.



Biomass Pathways

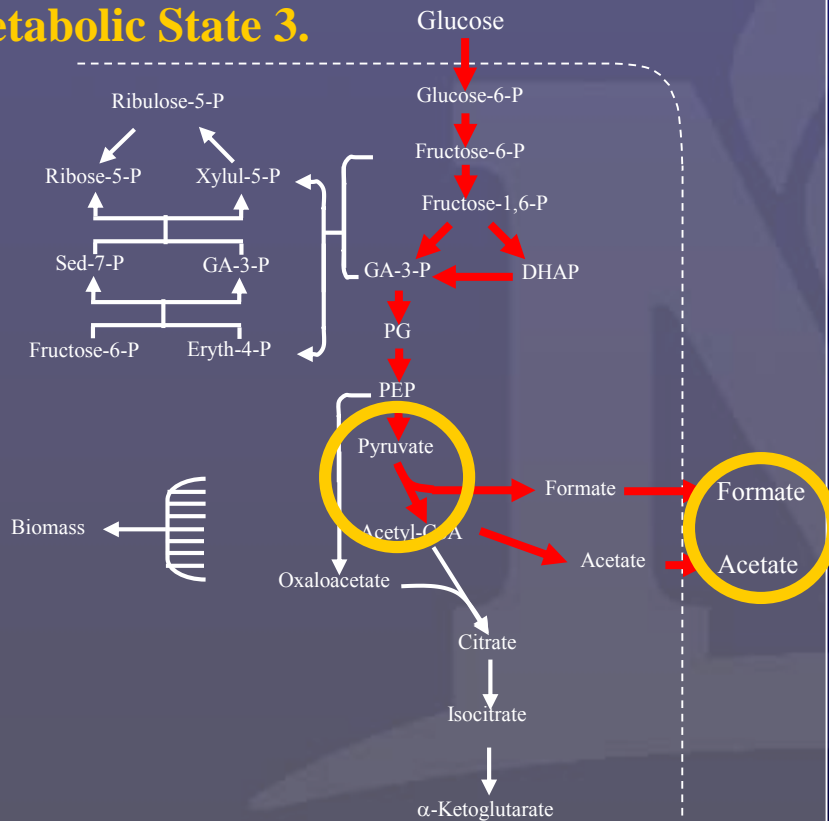


Carlson and Srienc, 2004

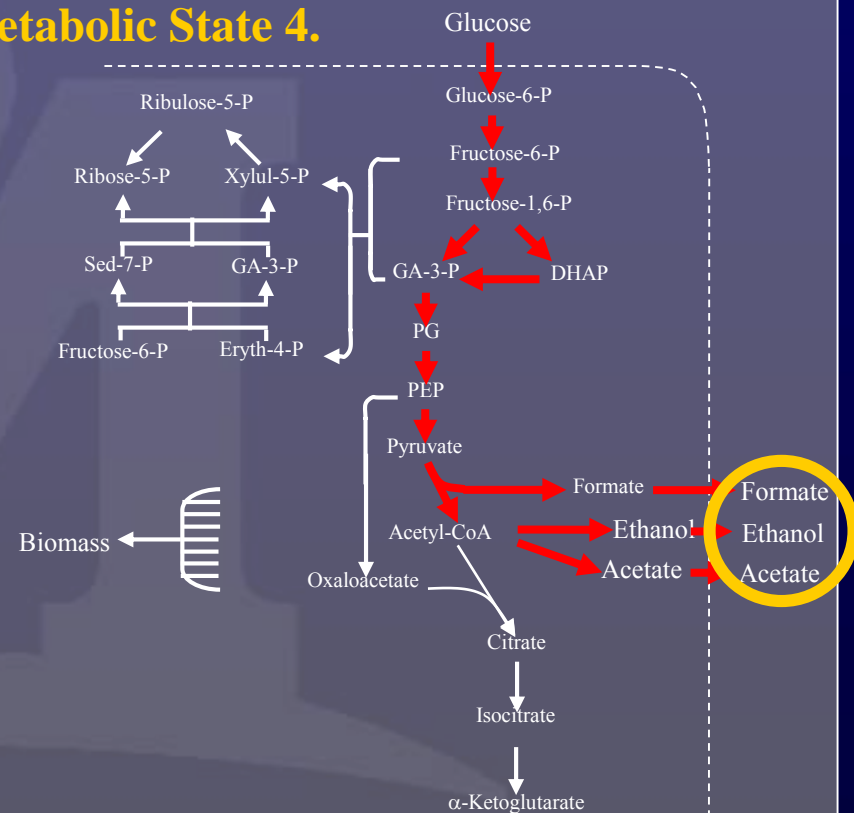


Qualitative Regulation Analysis

Metabolic State 3.



Metabolic State 4.



Practical: Sections 1.4-1.9

- *How is the output file manipulated to identify data of interest?*
- *How is the minimization envelope identified?*
- *How can enzyme usage patterns be identified from the data?*



Quantitative Rate Structure

METABOLIC PATHWAYS

Metabolism of
Complex Carbohydrates

Biodegradation of
Xenobiotics

Network can be described mathematically as

$$\mathbf{R} = \alpha \mathbf{M}$$

'accumulation'

**stoichiometry
matrix**

'velocity'

Lipid
Metabolism

Metabolism of
Other Amino Acids

Scaling factors required to convert pathways
into biologically significant vectors

$$\mathbf{R} = \alpha \left(s_i^x \mathbf{M}_i^x + s_i^{ATP} \mathbf{M}_i^{ATP} \right)$$

Biosynthesis of
Secondary Metabolites

Carbon Mass Balance

Glucose

no different than balancing checkbook...



www.darrenbyrne.com

Ribose
Sed
Fructose
Biomass

Succinate
Succinyl-CoA \leftarrow α -Ketoglutarate

Biomass

Succinate
Succinyl-CoA \rightarrow α -Ketoglutarate

ATP



MONTANA STATE UNIVERSITY

Mountains & Minds

Flux Partitioning

cdw/hr)

80

(rATP)

paid \$20/hr, spend \$8/hour on coffee, how much is left for Twinkies?...



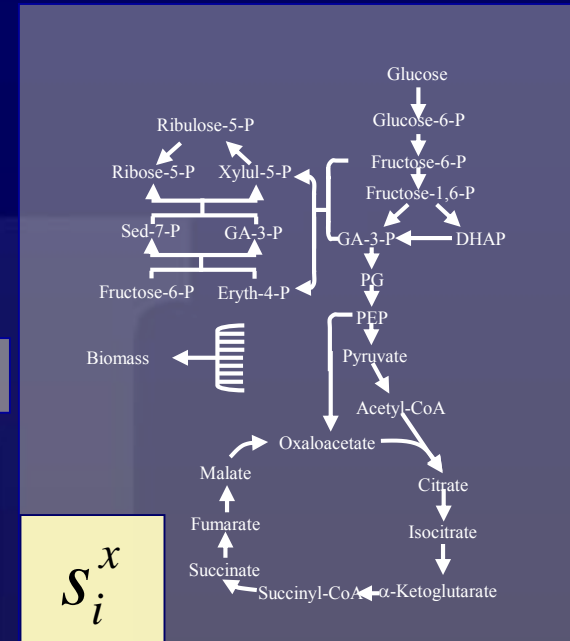
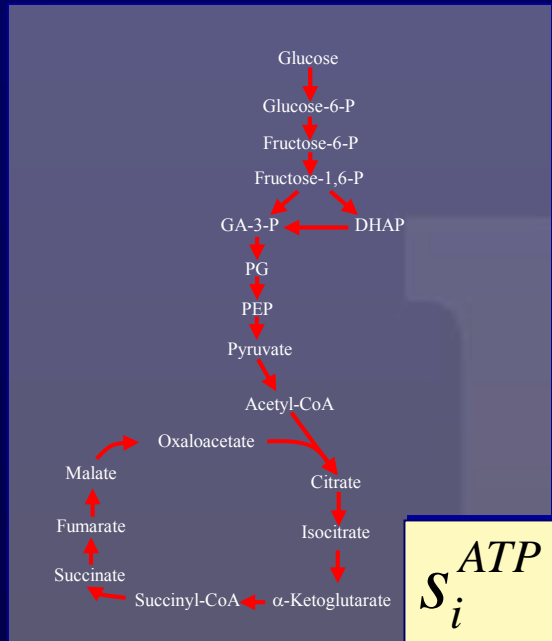
if \$2 / Twinkie, how many Twinkies can you buy per hour...

S_i^{ATP}

S_i^x



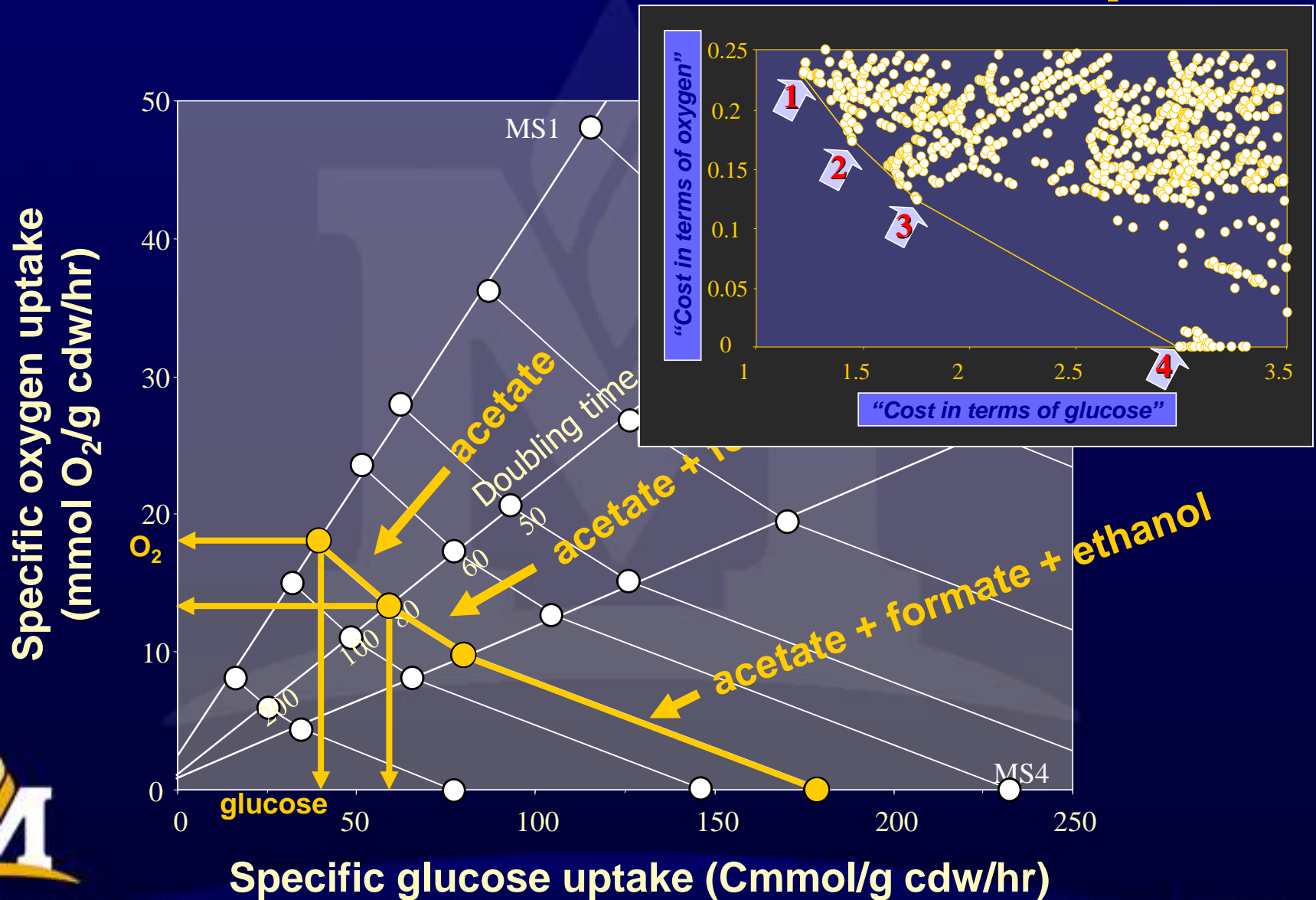
Cellular Rate Structure



Mathematical description of
cellular rates and behavior

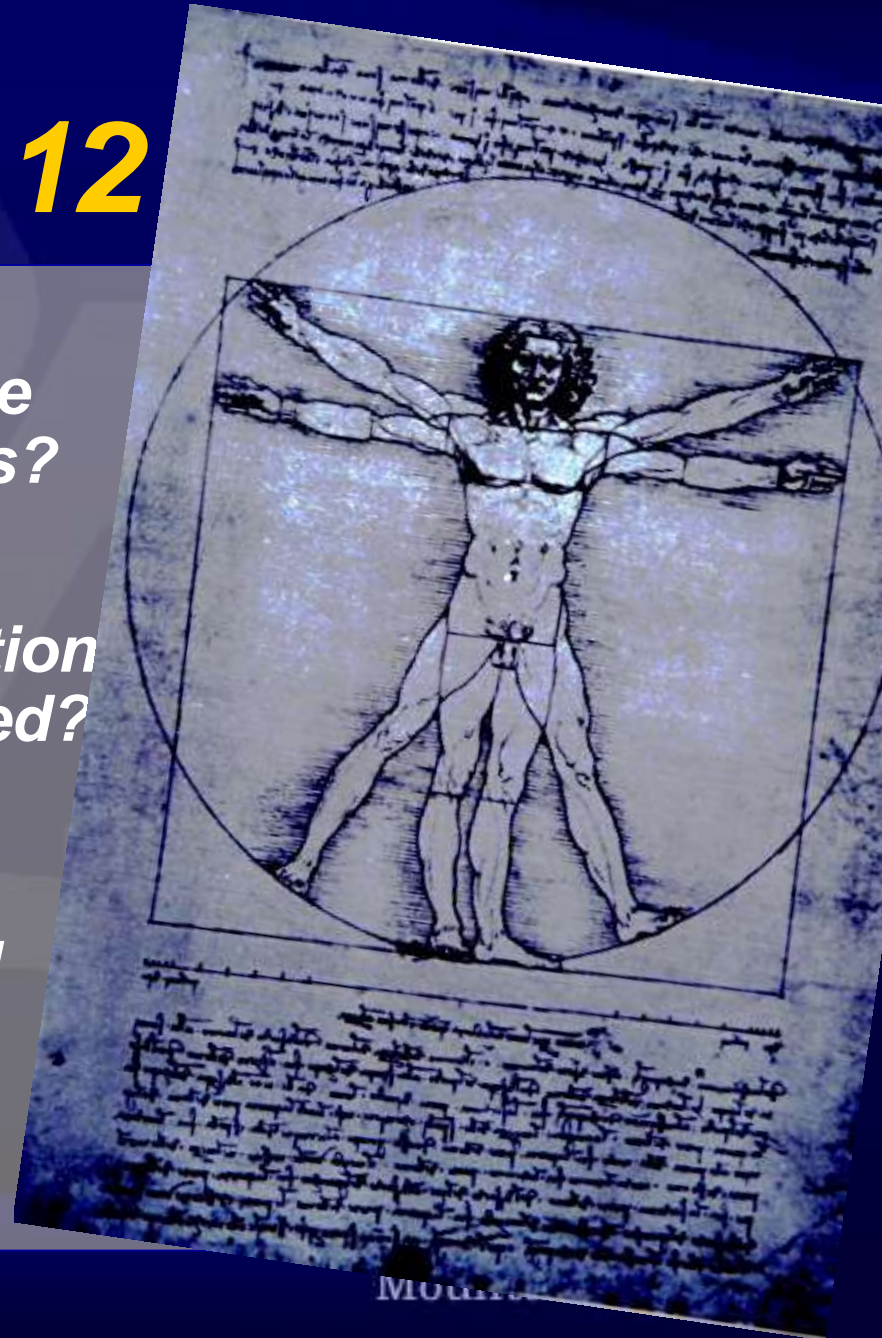


Most Efficient Flux Relationships

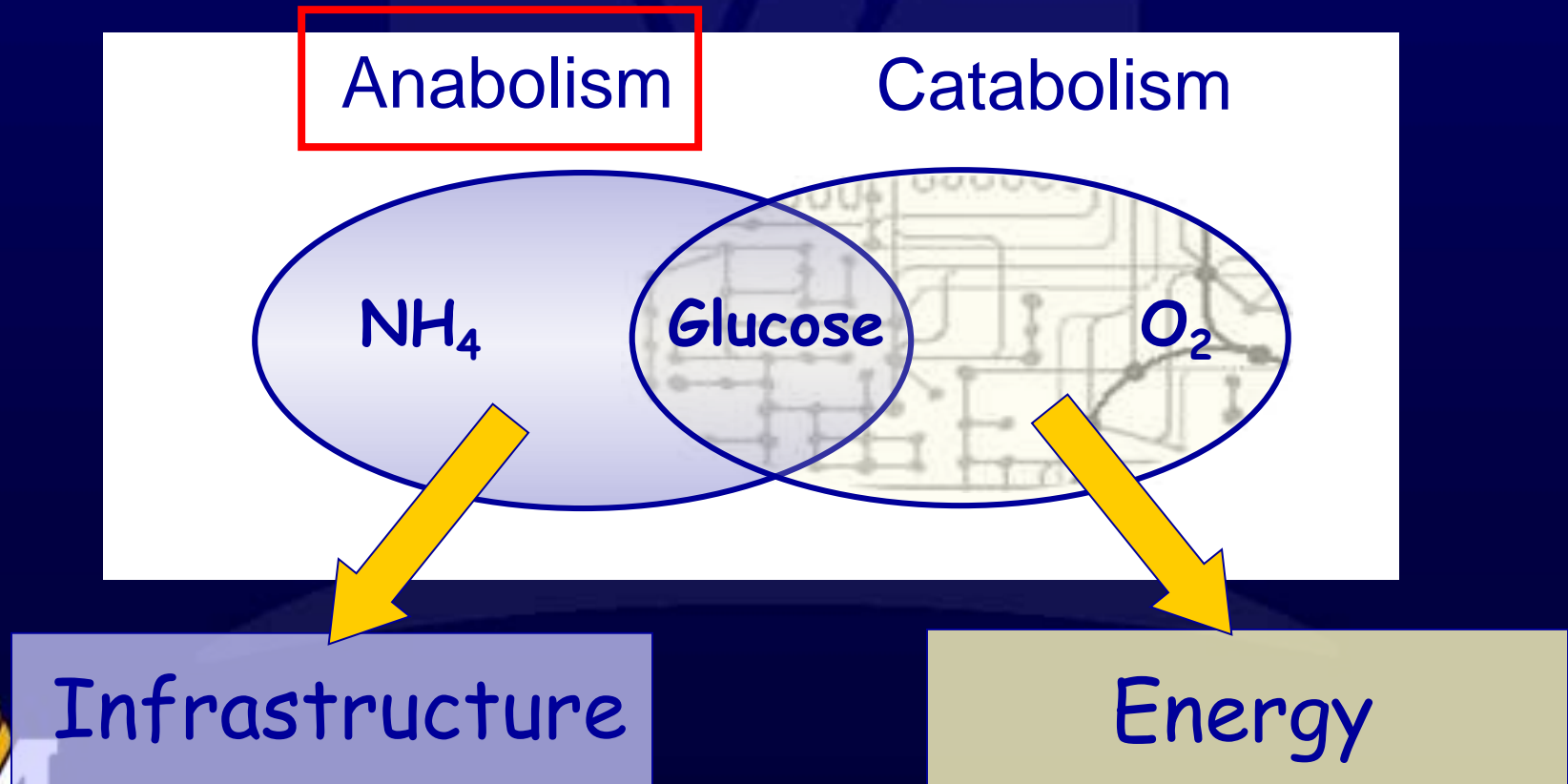


Practical: Sections 1.10-1.12

- *How can experimental data be integrated with pathway models?*
- *How are optimal growth solution spaces defined and constructed?*
- *How are metabolisms not located exactly at a defined node described?*



Pathways, Catabolism and Anabolism



Why Anabolism?

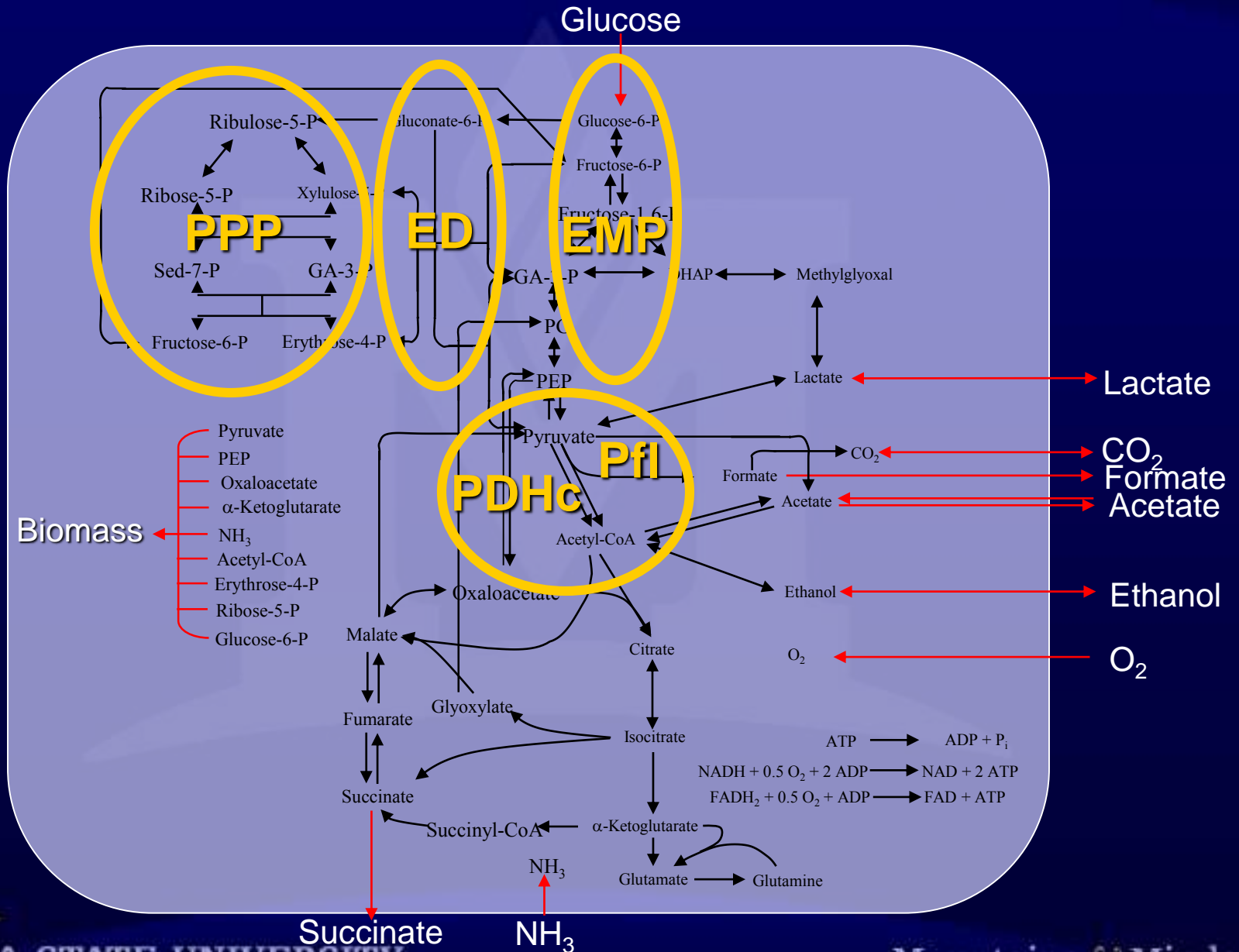
Typical cell: 50-80% protein,
1-3% DNA

Pelagibacter ubique:
optimized DNA operons for
nitrogen investment

(Giovannoni et al., 2005)



Network 'Investment Costs'



Proteome Investment

Pfl: pyruvate → acetyl-CoA + formate

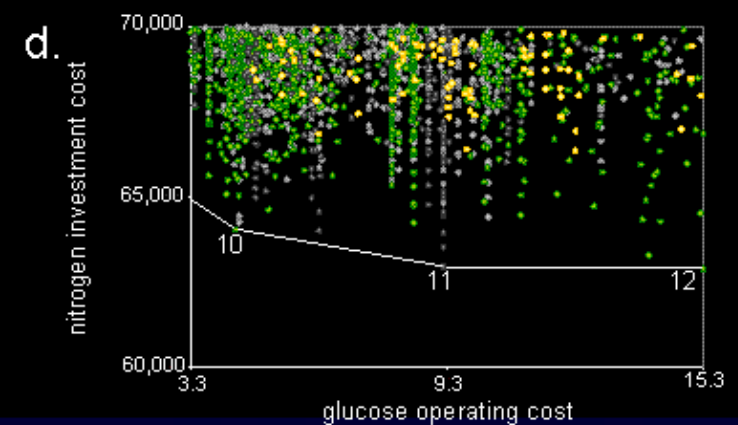
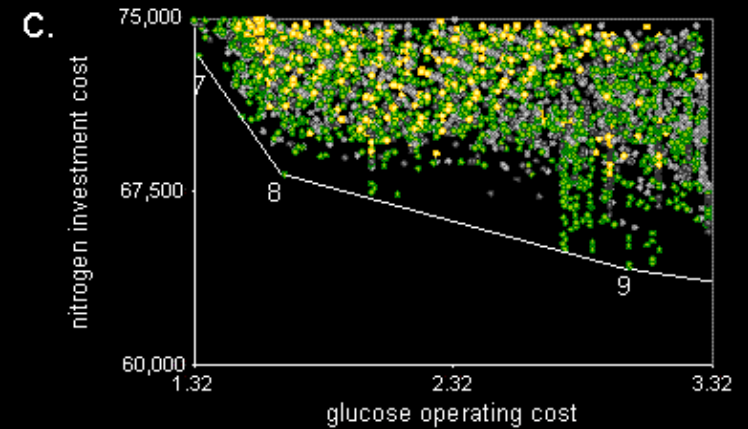
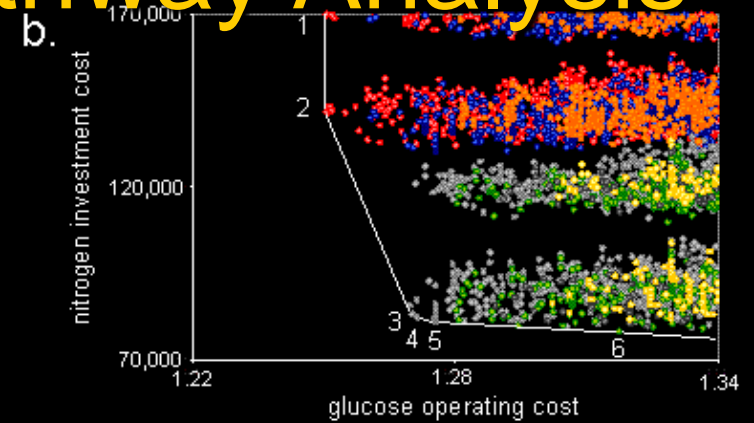
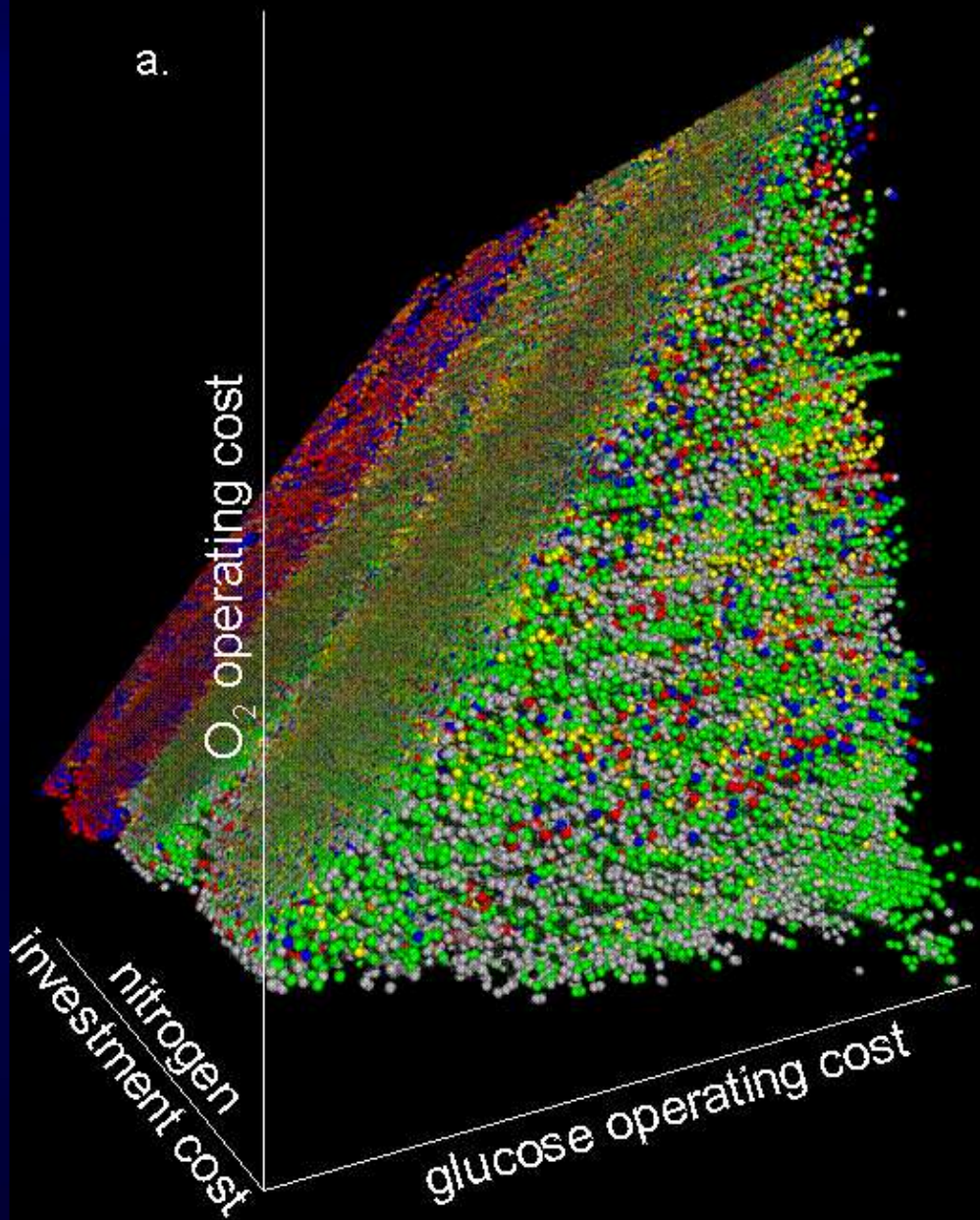
	Carbon	Sulfur	Nitrogen	A.Acids	DNA
2 PflB	7578	74	2048	1520	2280

PDHc: pyruvate → acetyl-CoA + NADH

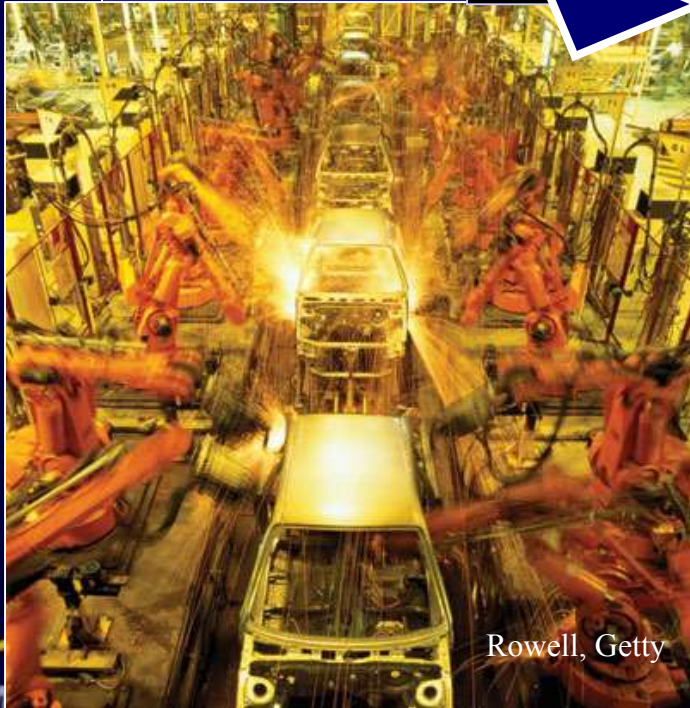
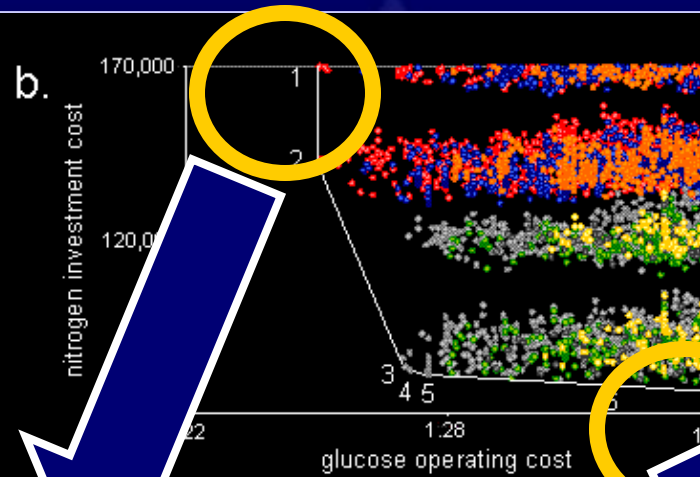
	Carbon	Sulfur	Nitrogen	A.Acids	DNA
24 AceE	106584	672	29232	21288	2661
24 AceF	70320	432	18912	15120	1890
12 LpdA	27192	180	7320	5688	1422
Sum:	204096	1284	55464	42096	5973



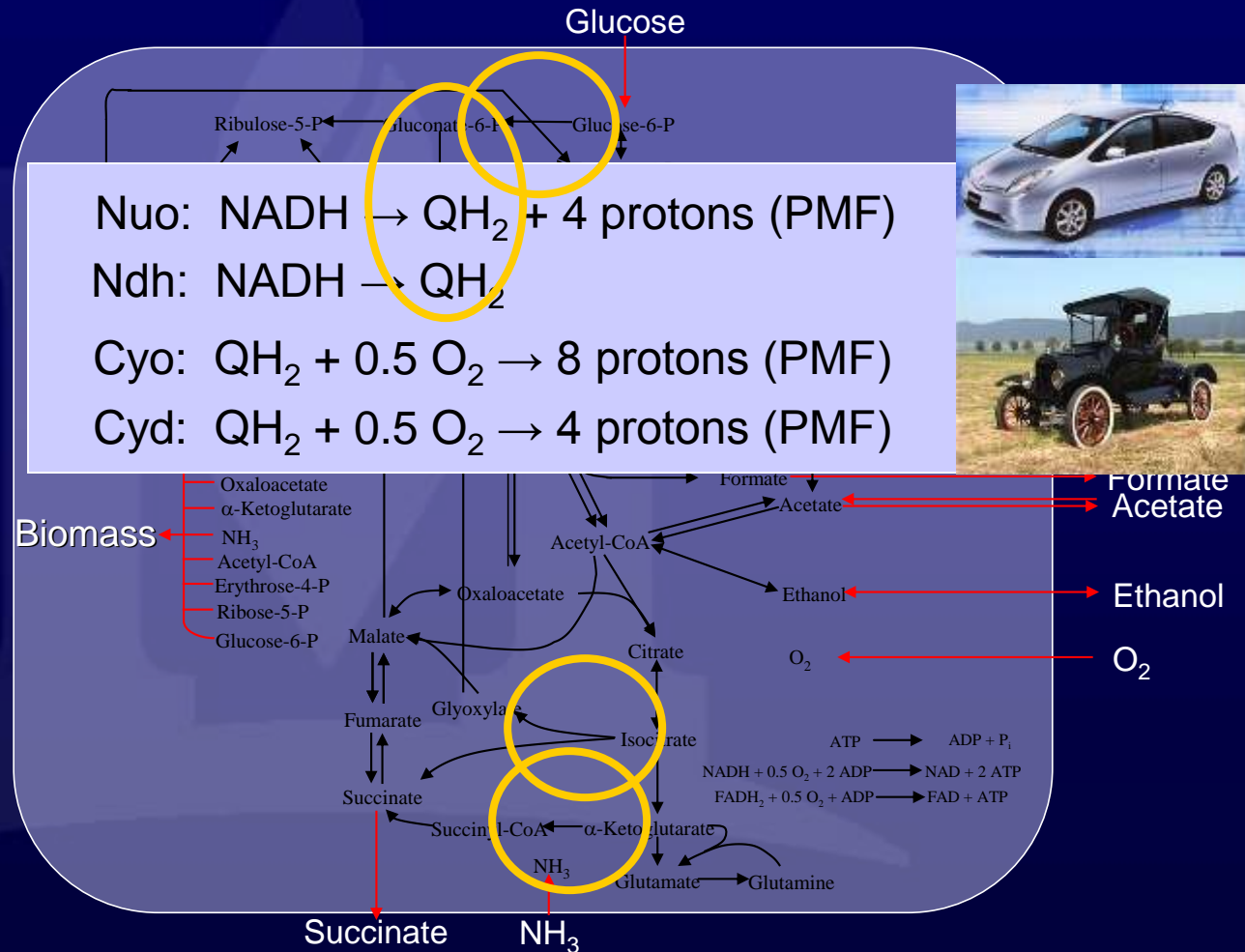
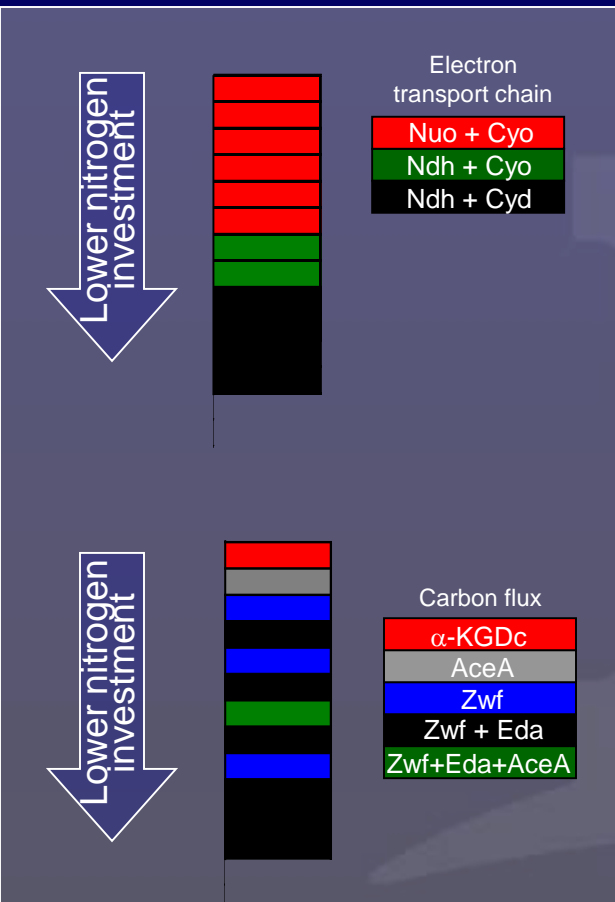
Multi-parameter Pathway Analysis



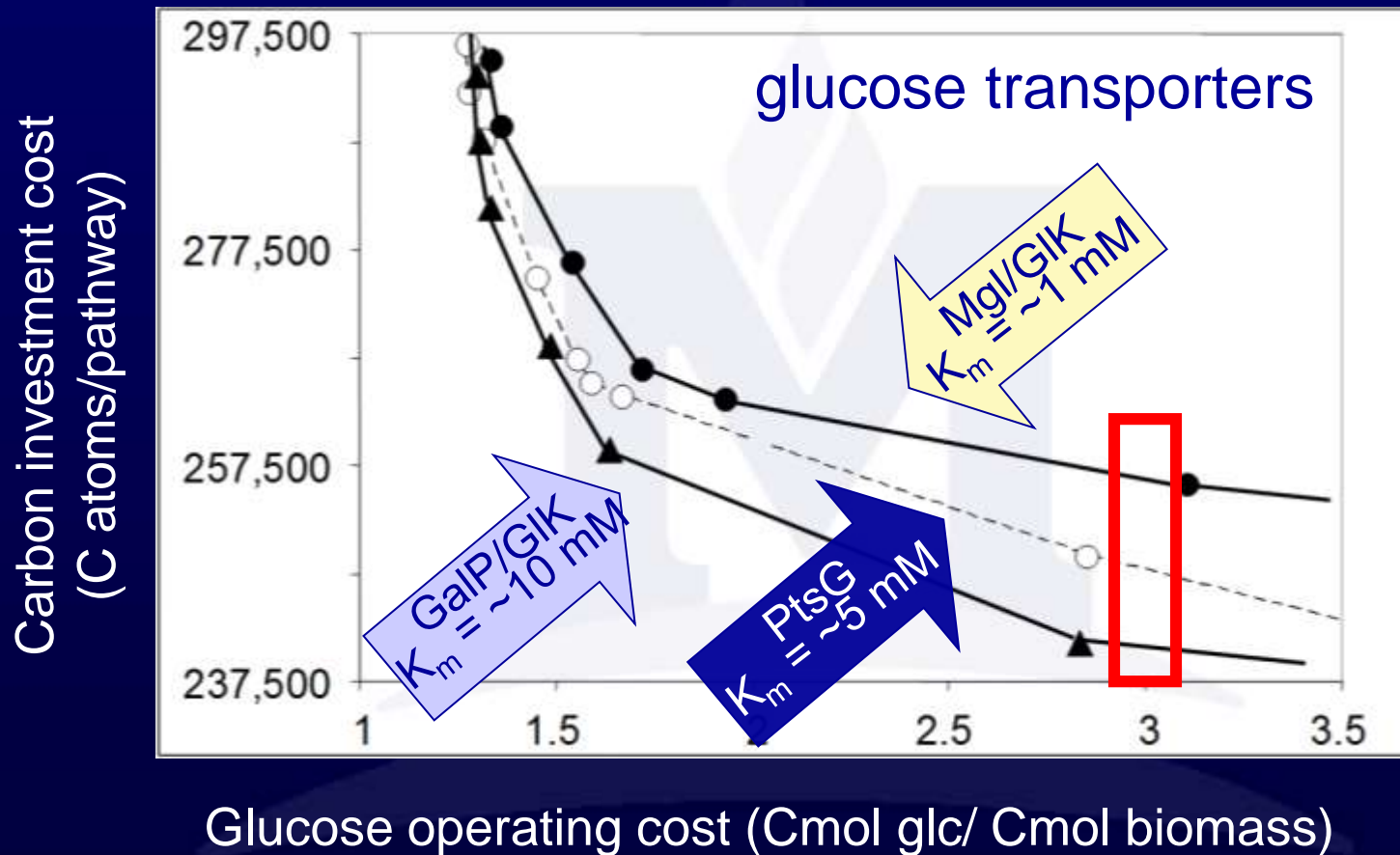
Analogy



Nitrogen Limitation Enzyme Patterns



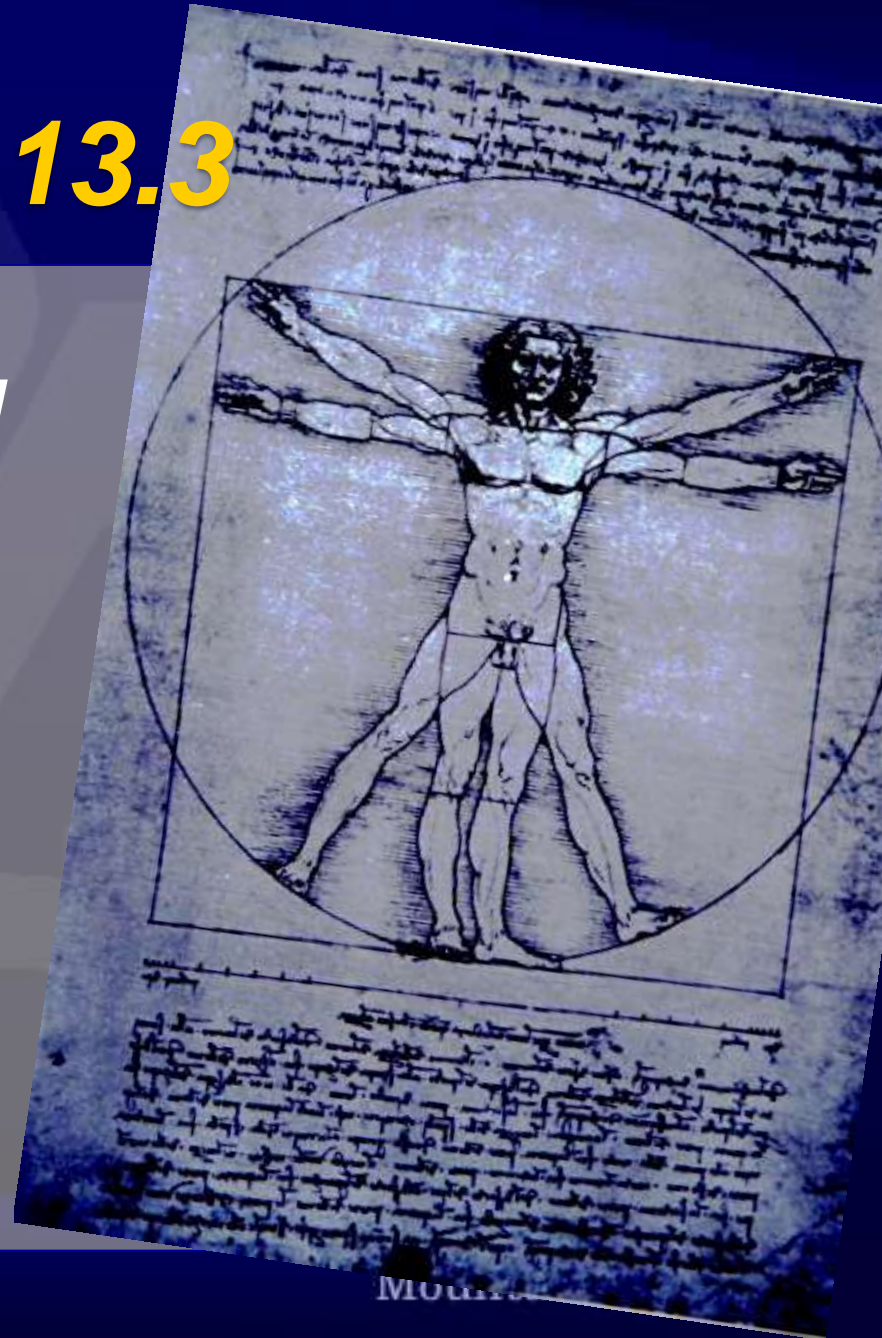
Cost-Benefit: Enzyme Affinity (k strategy)



Practical:

Sections 1.13.1-1.13.3

- *How are protein amino acid sequences converted into investment criteria?*
- *How is pathway data translated into investment costs?*



B 100% [$1\text{-}^{13}\text{C}$]glucose

PP pathway Glycolysis ED pathway

100% unlabeled 50% [$3\text{-}^{13}\text{C}$] 50% unlabeled 50% [$1\text{-}^{13}\text{C}$] 50% unlabeled

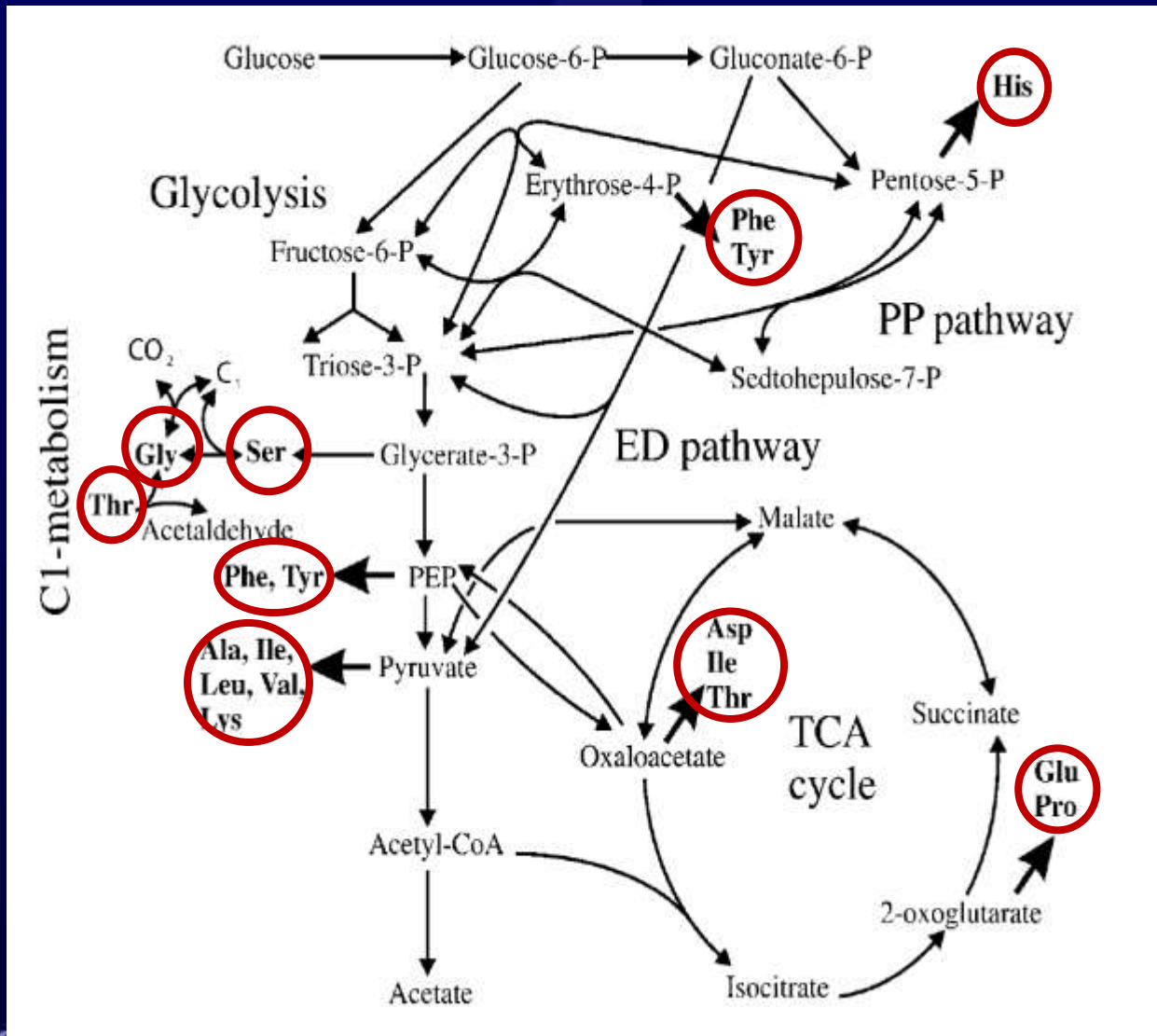
Alanine (C1-C3) Alanine (C2-C3)

Fragmentation

The diagram illustrates the metabolic pathways for the conversion of 100% [$1\text{-}^{13}\text{C}$]glucose into alanine. The starting material is glucose, with the C1 position labeled (indicated by a black dot and the number 1). The three pathways shown are the PP pathway, Glycolysis, and the ED pathway. The PP pathway leads to 100% unlabeled alanine. The Glycolysis pathway leads to 50% [$3\text{-}^{13}\text{C}$] and 50% unlabeled alanine. The ED pathway leads to 50% [$1\text{-}^{13}\text{C}$] and 50% unlabeled alanine. The alanine molecule is shown with its three carbon positions (C1, C2, C3) and the corresponding mass-to-charge ratio (m/z) for the unlabeled and labeled forms. A fragmentation box indicates the fragmentation of the alanine molecule into three fragments: C1, C2, and C3. The fragmentation of alanine (C1-C3) results in three fragments: C1 (m/z 0), C2 (m/z +1), and C3 (m/z +2). The fragmentation of alanine (C2-C3) results in two fragments: C2 (m/z 0) and C3 (m/z +2).

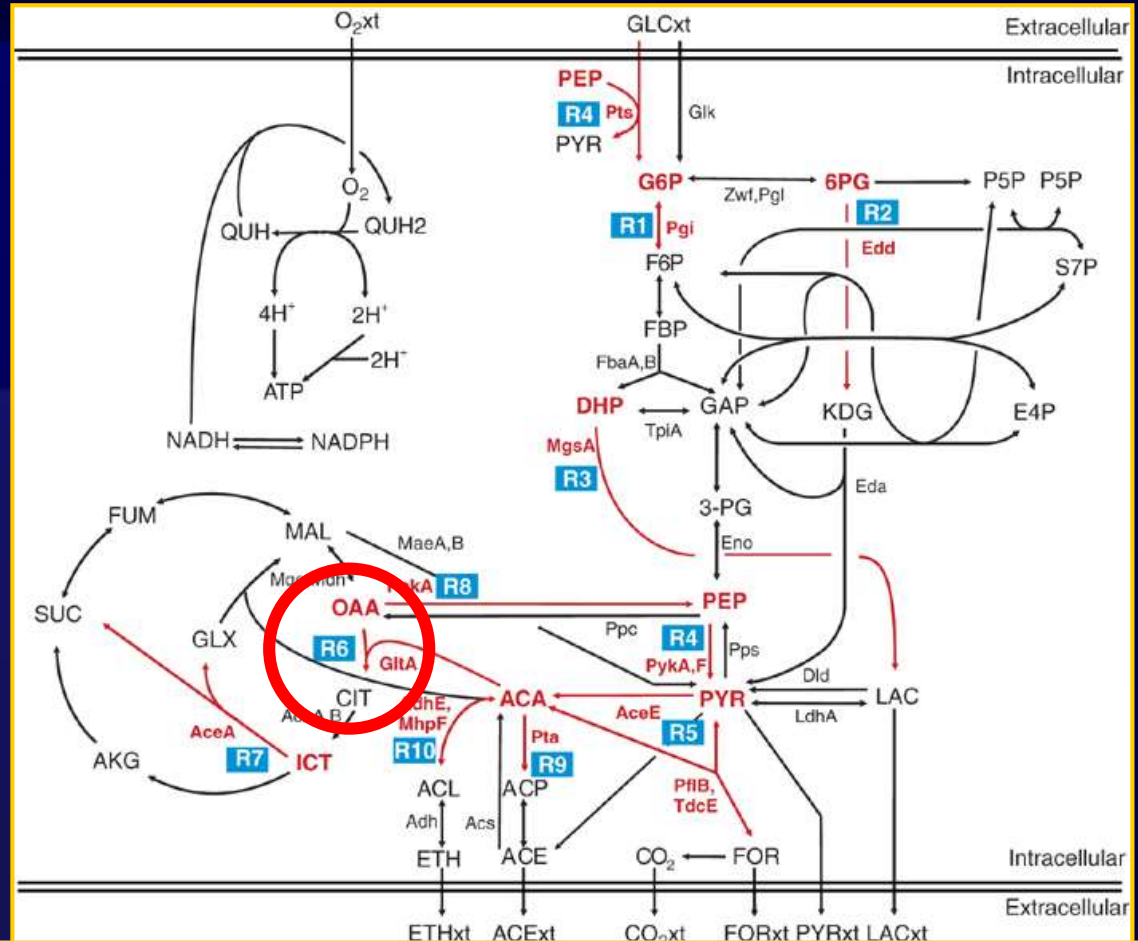


Amino acids and the central metabolism



Fluxomic data:

- Experimental data
- Why are there such different fluxes?
- Alternative approach

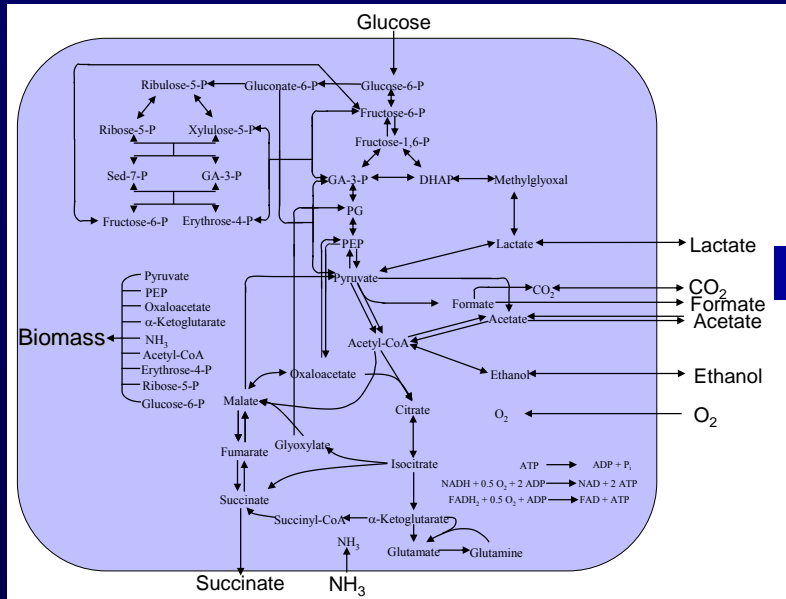


Schuetz, Kuepfer & Sauer, 2007

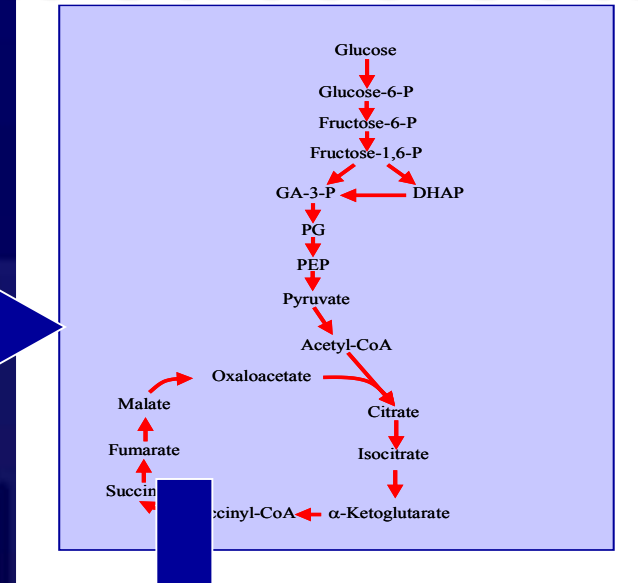
	Batch	Chemostat		
	0.6 h ⁻¹ Aerobe ^b	0.1 h ⁻¹ C-limited ^c	0.4 h ⁻¹ C-limited ^c	0.4 h ⁻¹ N-limited ^d
Growth rate				
Split ratio ^a				
R1 (flux into glycolysis)	0.70 ± 0.02	0.69 ± 0.12	0.64 ± 0.05	0.96 ± 0.14
R2 (flux into Entner-Doudoroff pathway)	0.13 ± 0.06	0.23 ± 0.20	0.19 ± 0.11	0.00 ± 0.05
R3 (flux into methylglyoxal pathway)	0.00 ± 0.05	0.00 ± 0.05	0.00 ± 0.05	0.00 ± 0.05
R4 (PEP to pyruvate flux)	0.78 ± 0.02	0.84 ± 0.14	0.70 ± 0.06	0.72 ± 0.10
R5 (pyruvate to acetylcoenzyme A flux)	0.61 ± 0.05	0.91 ± 0.21	0.84 ± 0.10	0.90 ± 0.15
R6 (flux into TCA cycle)	0.24 ± 0.02	0.64 ± 0.13	0.85 ± 0.09	0.50 ± 0.06
R7 (flux into glyoxylate shunt)	0.00 ± 0.05	0.40 ± 0.15	0.00 ± 0.05	0.00 ± 0.05
R8 (oxaloacetate to PEP flux)	0.00 ± 0.05	0.35 ± 0.08	0.12 ± 0.03	0.01 ± 0.01
R9 (acetate secretion)	0.58 ± 0.03	0.00 ± 0.05	0.00 ± 0.05	0.04 ± 0.01
R10 (ethanol secretion)	0.00 ± 0.05	0.00 ± 0.05	0.00 ± 0.05	0.00 ± 0.05



Ecologically Relevant Stress Sets



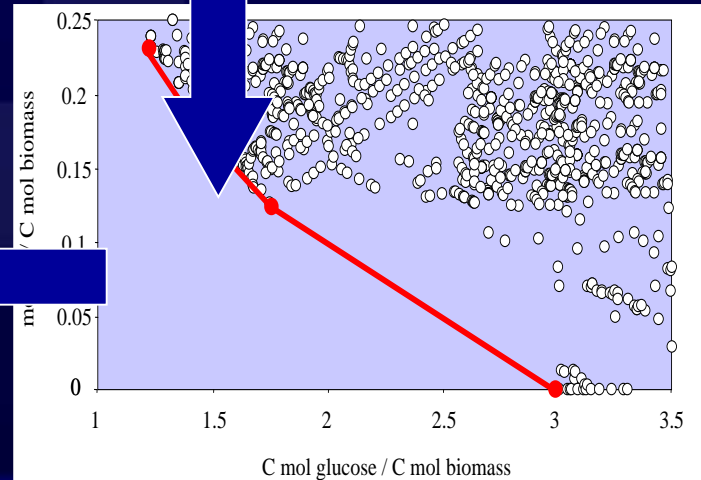
Pathway
analysis



Cost-benefit analysis

023	0.000489	0	0	0.002535	0.002535	0	0.005069	0.002535	0.002535	0	0	0.037439	0.052994	0
241	0.000667	0	0	0.007574	0.007574	0	0.015148	0.007574	0.007574	0	0	0	0.073024	0.025151
858	0.000769	0	0	0.005089	0.005089	0.010177	0	0.005089	0.005089	0	0	0	0.077456	0.036183
858	0.000769	0	0	0.005089	0.005089	0.006849	0	0.005089	0.005089	0	0	0	0.077456	0.036183
075	0.000631	0	0	0.001444	0.001444	0.002888	0	0.001444	0.001444	0	0	0	0.055349	0.026952
011	0.000611	0	0	0.000611	0.000611	0.001222	0	0.000611	0.000611	0	0	0	0.052717	0.024259
285	0.00449	0.003705	0.003705	0.005538	0.005538	0.011076	0	0.005538	0.005538	0	0	0	0.075015	0.034757
568	0.000568	0	0	0	0	0	0	0	0	0.000939	0	0	0.046794	0.023397
481	0.000548	0	0	0.000548	0.000548	0.001096	0	0.000548	0.000548	0	0	0	0.055595	0
437	0.000437	0	0	0.000437	0.000437	0.000874	0	0.000437	0.000437	0	0	0	0.029012	0
041	0.00041	0	0	0	0	0	0	0	0	0.000939	0	0	0	0
033	0.00033	0	0	0	0	0	0	0	0	0.000756	0	0	0	0
146	0.000428	0	0	0.000428	0.000428	0.000856	0	0.000428	0.000428	0	0	0	0.025512	0
784	0.006784	0.005998	0.005998	0.005998	0.005998	0.011996	0	0.005998	0.005998	0.001799	0.000799	0	0.074157	0.034178
007	0.007007	0.006222	0.006222	0.006222	0.006222	0.012444	0	0.006222	0.006222	0.001798	0.001891	0	0.073687	0.033969
624	0.006624	0.005838	0.005838	0.005838	0.005838	0.011676	0	0.005838	0.005838	0	0	0	0.074495	0.034329
624	0.006624	0.005838	0.005838	0.005838	0.005838	0.011676	0	0.005838	0.005838	0	0	0	0.074495	0.034329
169	0.000169	0	0	0	0	0	0	0	0	0.000387	0	0	0	0
299	0.000299	0	0	0	0	0	0	0	0	0.000684	0	0	0.012258	0
124	0.000124	0	0	0	0	0	0	0	0	0.000283	0	0	0	0
522	0.000522	0	0	0	0	0	0	0	0	0.001195	0.005881	0	0.038611	0.020041

38 'Stress'
Vectors



[illegible][illegible]

R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
0	0	0	0.784054	0.274671	0.466706	0.838384	0	0	0
0	0	0	0.942606	0.767511	0.466706	0.838384	0.254572	0	0
0	0	0	0.961518	0.812315	0.466706	0.838384	0.254572	0	0
0	0	0	0.866046	0.154255	0.466706	0.838384	0.254572	0	0
0	0	0	0.962766	0.852336	0.466706	0.838384	0.279367	0	0
0	0	0	0.962766	0.852336	0.466706	0.838384	0.279367	0	0
0	0	0	0.962766	0.852336	0.466706	0.838384	0.279367	0	0
0	0	0	0.962766	0.852336	0.466706	0.838384	0.279367	0	0
0.339391	0	0	0.866723	0.877814	0.966001	0	0	0.823236	0
0.202465	0.956584	0	0.978545	0.939179	0.487555	0.942172	0.429751	0	0
0.661282	0	0	0.979746	0.921067	0.122714	0.695962	0	0.721635	0
0.682197	0	0	0.909864	0.921067	0.03731	0	0	0.892443	0
0.743915	0	0	0.929367	0.939179	0.7982	0	0	0.919332	0
0.378148	0.956584	1	0.967727	0.954171	0.0087	0.957863	0.432263	0	0
0.837653	0	0	0.815934	0.823171	0.0075	0	0	0	0
0.924906	0	0	0.819236	0.819171	0.819171	0	0	0	0
0.774937	0	0	0.813491	0.813424	0.813424	0	0	0	0
0.774937	0	0	0.813491	0.817824	0.817399	0	0	0	0
0.869018	0	0	0.965938	0.971198	0.012551	0	0	0.412341	0.551477
0.565595	0.956584	0	0.922312	0.948554	0.023614	0	0	0.931927	0
0.904061	0	0	0.975445	0.121649	0.071588	0	0	0	0.793624
0.982291	0	0	0.888451	0.900235	0.048249	0	0	0.860908	0
0.920096	0	0	0.885499	0.897254	0.049855	0	0	0.856276	0
0.579852	0.914156	0	0.923184	0.948062	0.032851	0	0	0.931242	0
0.986723	0	0	0.918876	0.929847	0.032847	0	0	0.905309	0
0.993811	0	0	0.963434	0.97062	0.013178	0	0	0.482974	0.479036
0.994111	0	0	0.932345	0.970796	0.013097	0	0	0.473857	0.488386
0.594204	0.868132	0	0.924019	0.947553	0.024098	0	0	0.93053	0
0.651132	0.868132	0	0.937508	0.955835	0.020117	0	0	0.942007	0
0.77653	0.868132	0	0.963458	0.972931	0.012113	0	0	0.467731	0.49735
0.593462	0.870791	0	0.923988	0.94759	0.02408	0	0	0.930583	0
0.842909	0.868132	0	0.975445	0.043154	0.187683	0	0	0	0.458944
0.650452	0.870791	0	0.937483	0.955859	0.020106	0	0	0.942039	0

Experimental Split Set

Nonnegative Least Squares

Split Ratio Sets

Analysis



Minimize $\|b-Ax\|$

'x' represents best fit of ecologically selected vectors to experimental data

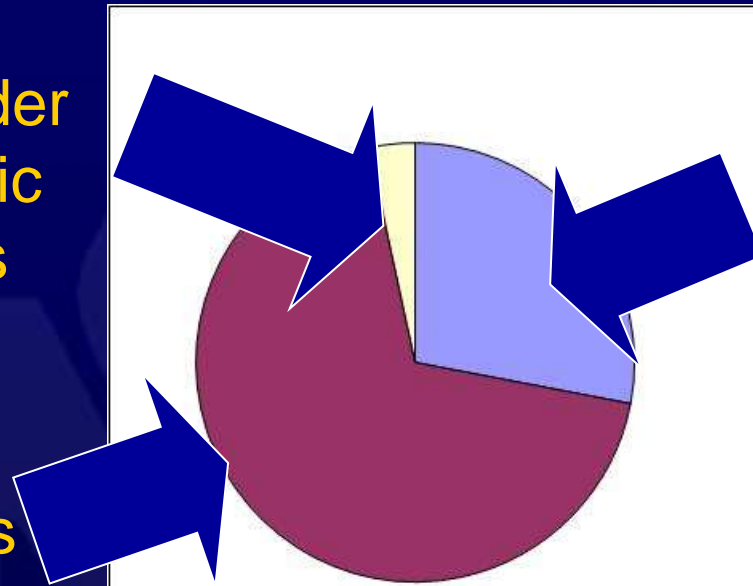
Results: Why

- Batch

First order
anabolic
stress

O₂ stress

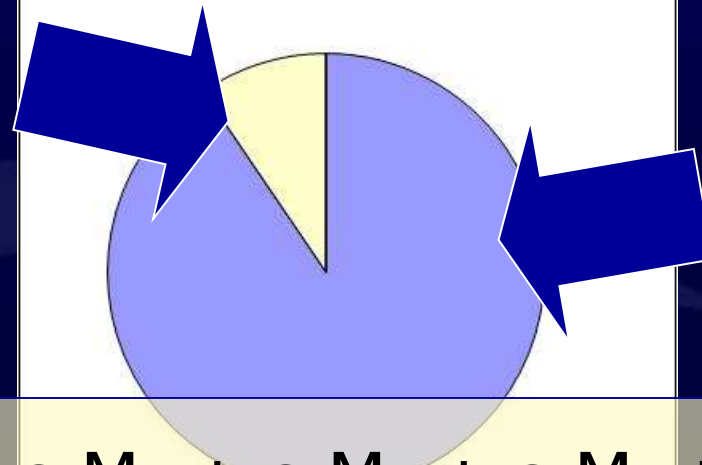
Optimal
growth



- Chemostat (C-limited)

First order
anabolic
stress

Optimal
growth



$$F = s_1 M_1 + s_2 M_2 + s_3 M_3 + \dots$$



Comparison to other methods

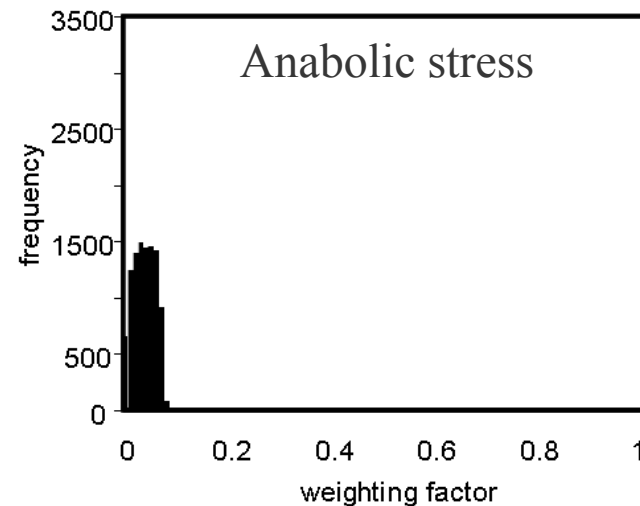
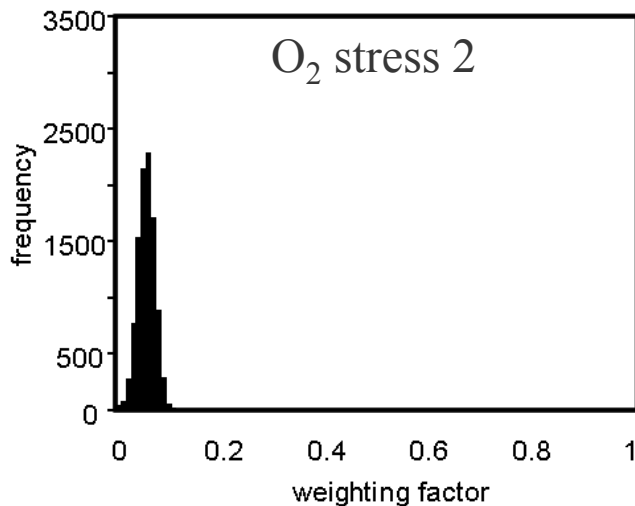
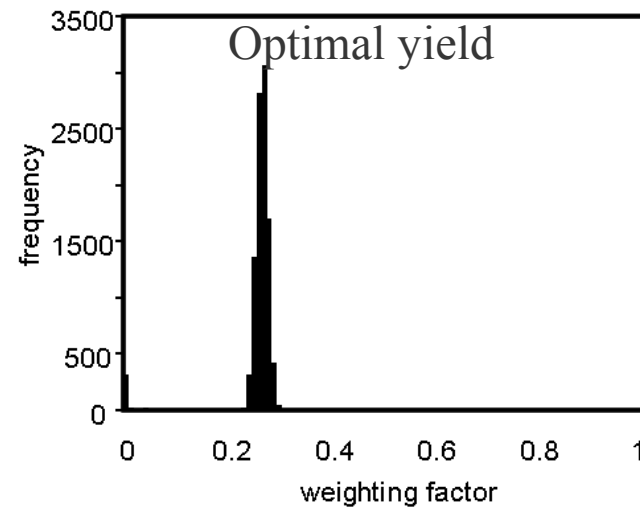
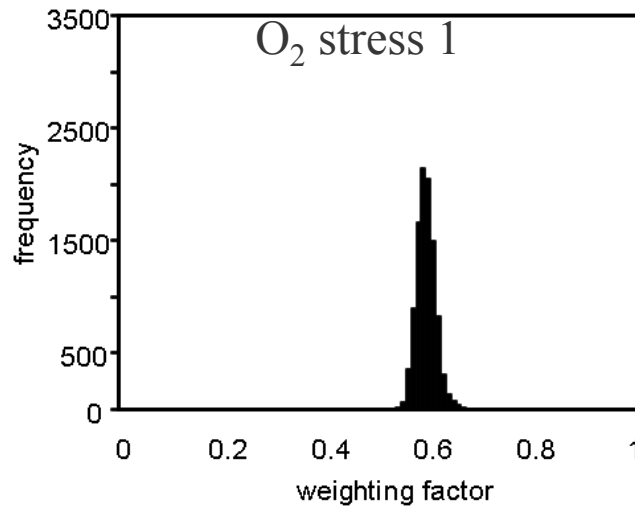
Euclidean distance		
growth conditions:	ecological costs based:	LP based:
batch growth	0.152	0.478
glucose limited chemostat (D = 0.1/hr)	0.267	0.700
glucose limited chemostat (D = 0.4/hr)	0.196	0.287
nitrogen limited chemostat (D=0.4/hr)	0.251	0.440



Significance of Results

growth conditions:	p-value	average no. pathways, random	no. pathways, costs-based
batch growth	0.006	8.4 ± 1.1	4
glucose limited chemostat (D = 0.1/hr)	0.029	7.1 ± 1.2	3
glucose limited chemostat (D = 0.4/hr)	0.052	8.0 ± 1.3	4
nitrogen limited chemostat (D=0.4/hr)	0.017	7.5 ± 1.2	4
all four conditions	≤ 0.0001	n.a.	n.a.

Perturbation Analysis



Principle Component Analysis

First 22 principle components capture
99.9% of variance

22 principle components are distinct from the
ecological vector set

18 principle components required to out
compete 12 ecological vectors



Community analysis

Identifying general ecological principles

Organize

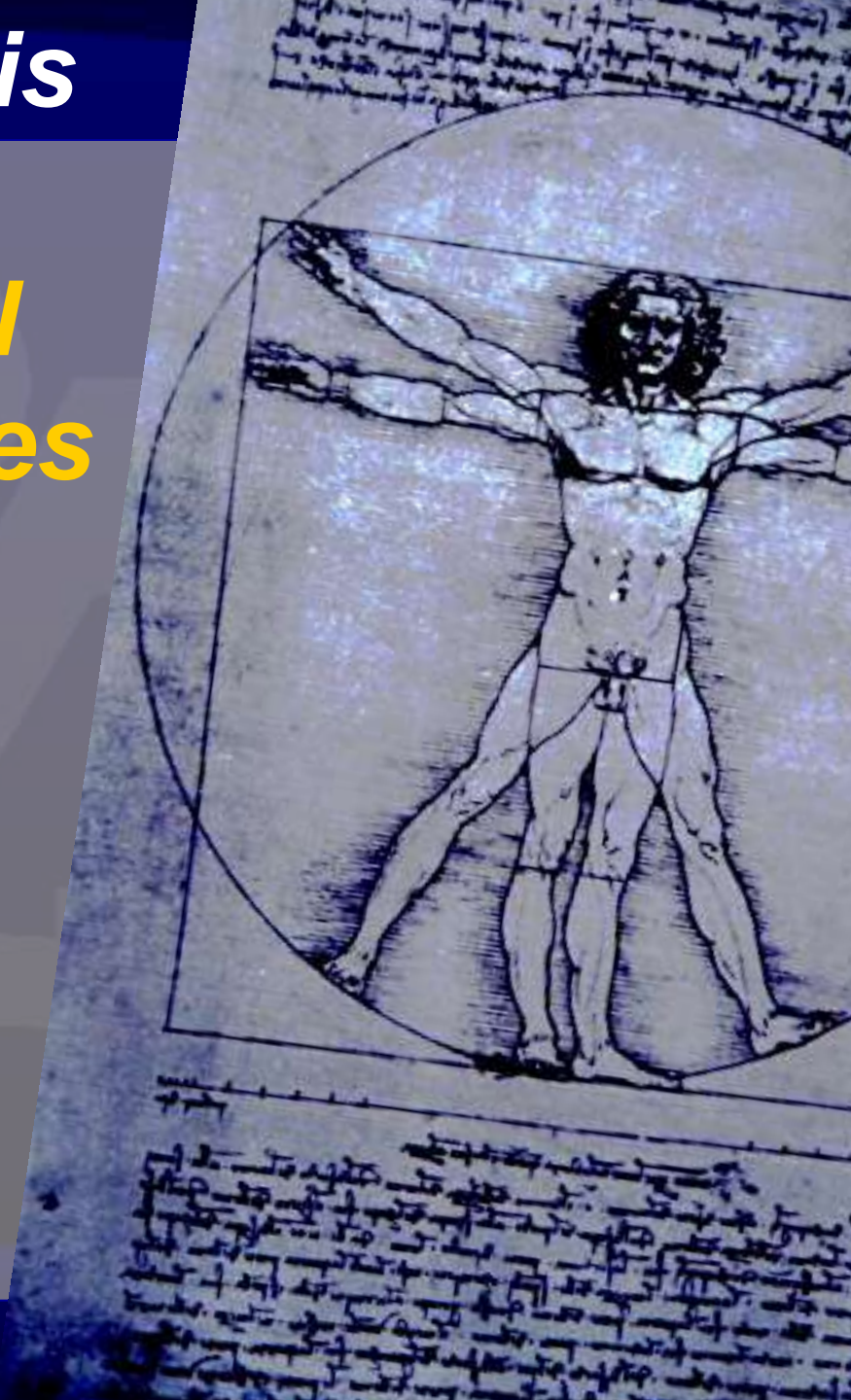
Explain

Predict

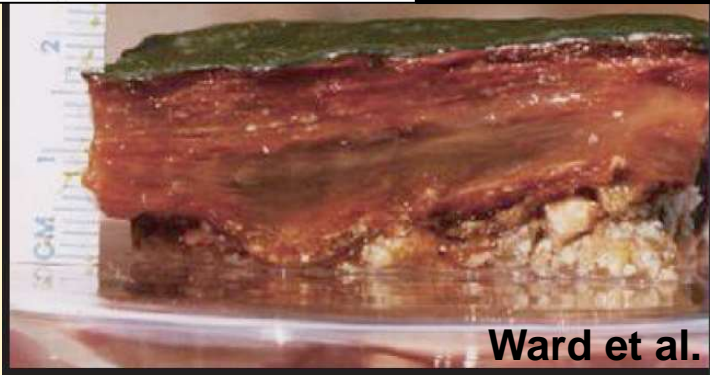
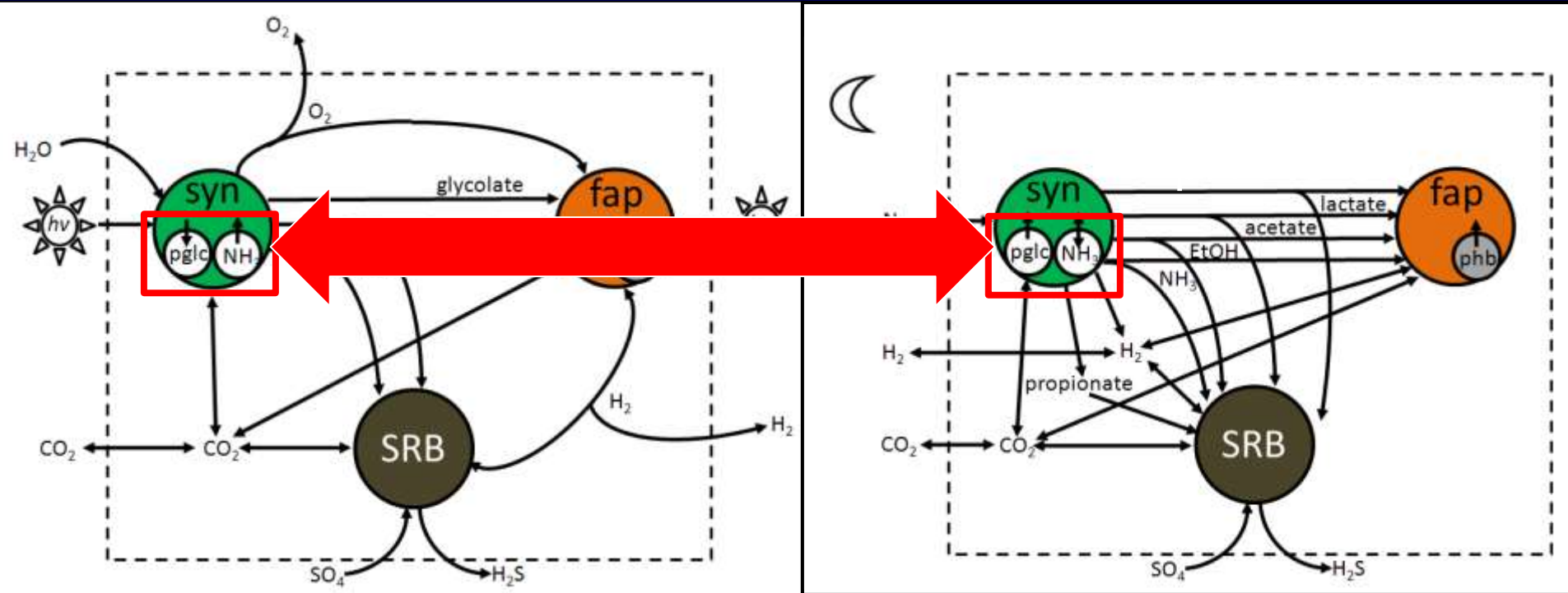
Control



MONTANA STATE UNIVERSITY



Test Case: Octopus Spring



3 representative guilds, 2 diel phases

Organize: Representative Models



Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function

David M. Ward^{1,2}, Mary M. Bonner¹, Alex Kopyov¹, Michael Krumholz¹, Andrew Whitford^{1,2}, Michael J. Farris¹, Michael Krumholz¹, Frederick M. LaRue¹

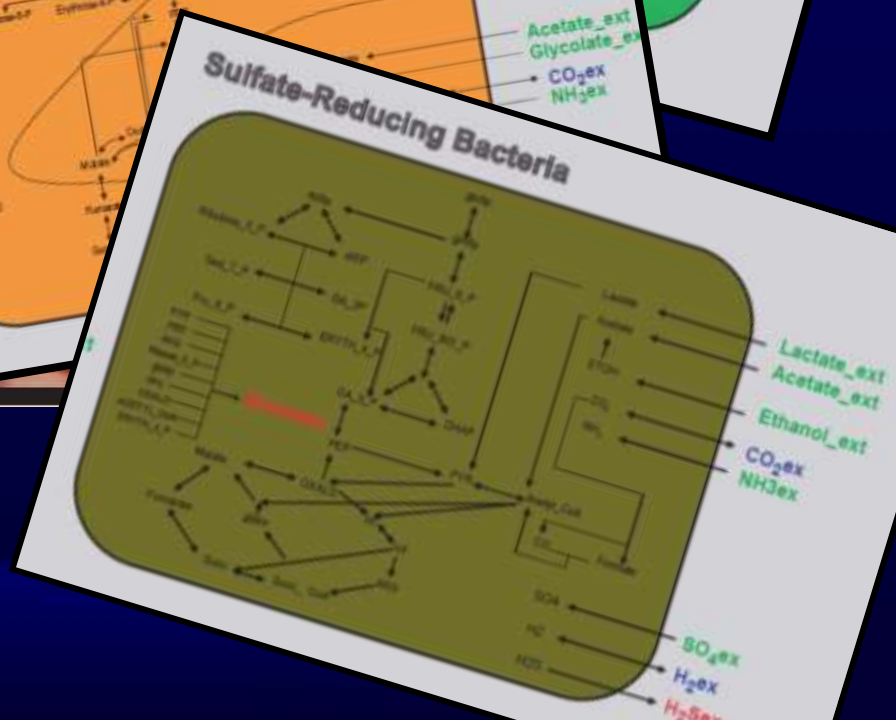
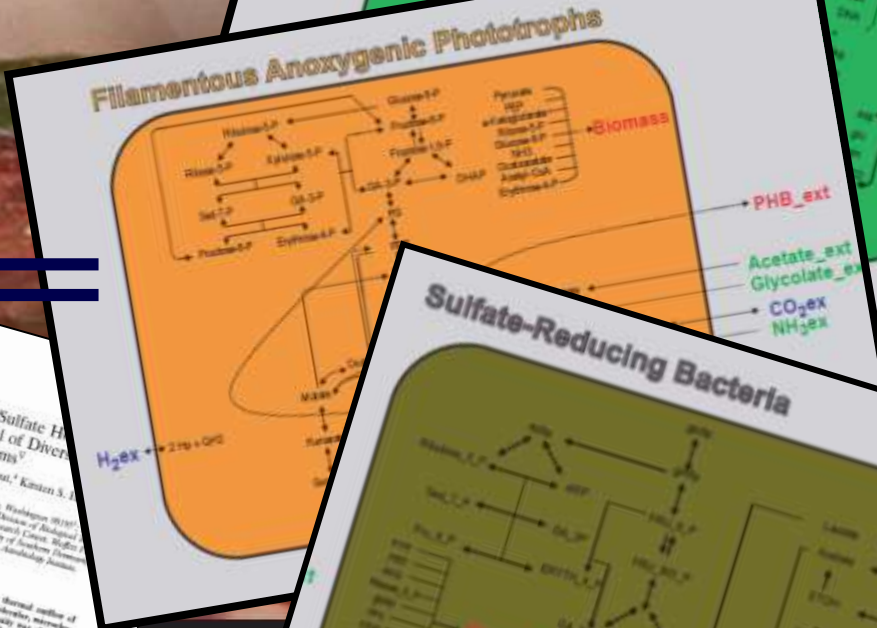
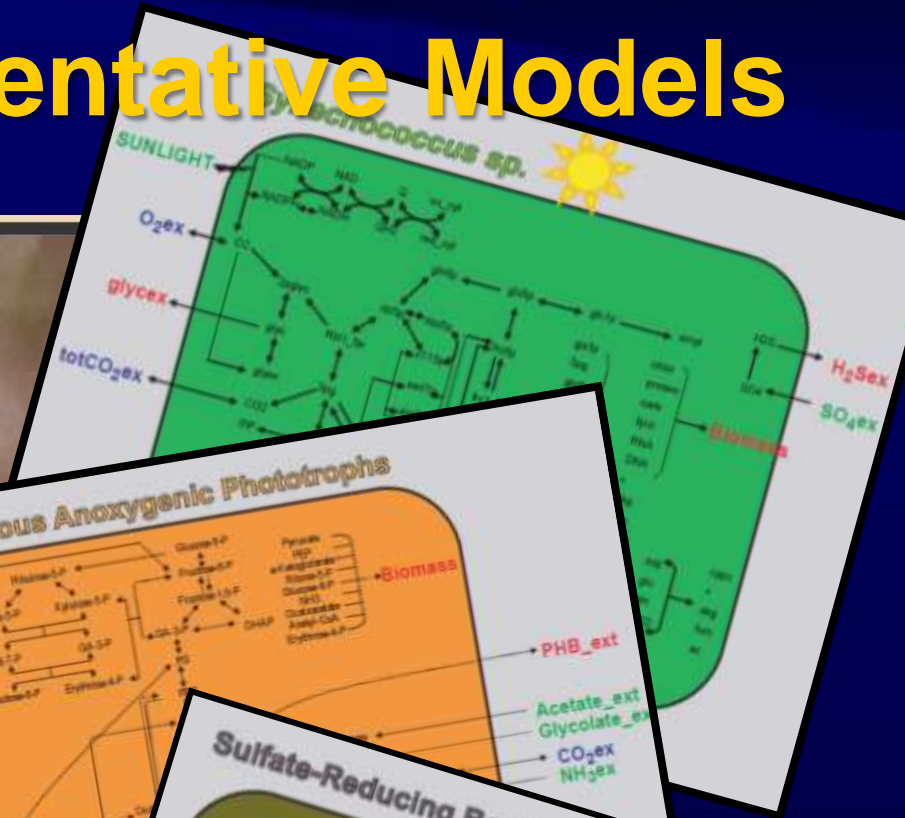
¹Department of Geological Engineering and Science, Stanford University, Stanford, California 94305-5080; ²Department of Earth and Planetary Sciences, University of California, Berkeley, California 94720-1767

Abstract

Microbial mats are complex, multi-layered communities of diverse microorganisms that develop in a variety of environments. They are important components of the Earth's microbial diversity and play a role in the biogeochemical cycle of carbon and other elements. In this study, we used DNA sequencing to identify and characterize the microbial community structure and function of cyanobacterial ecotypes in the microbial mat community of Mushroom Spring, Yellowstone National Park, Wyoming. We found that the community is composed of several distinct ecotypes, each with a unique set of genes and metabolic capabilities. These ecotypes are linked by their shared metabolic pathways, which allow them to function as a single unit. Our results suggest that the microbial mat community of Mushroom Spring is a highly organized and functional system, with each ecotype playing a specific role in the overall community structure and function.

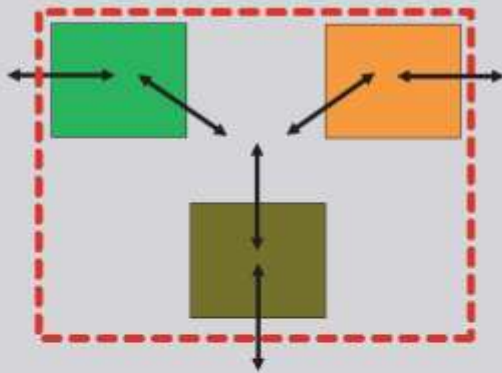
Introduction

Microbial mats are complex, multi-layered communities of diverse microorganisms that develop in a variety of environments. They are important components of the Earth's microbial diversity and play a role in the biogeochemical cycle of carbon and other elements. In this study, we used DNA sequencing to identify and characterize the microbial community structure and function of cyanobacterial ecotypes in the microbial mat community of Mushroom Spring, Yellowstone National Park, Wyoming. We found that the community is composed of several distinct ecotypes, each with a unique set of genes and metabolic capabilities. These ecotypes are linked by their shared metabolic pathways, which allow them to function as a single unit. Our results suggest that the microbial mat community of Mushroom Spring is a highly organized and functional system, with each ecotype playing a specific role in the overall community structure and function.



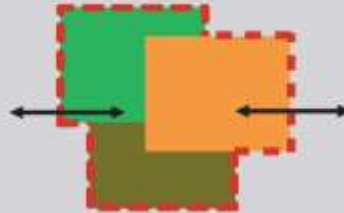
Community Analysis Tools

‘Compartmentalized’



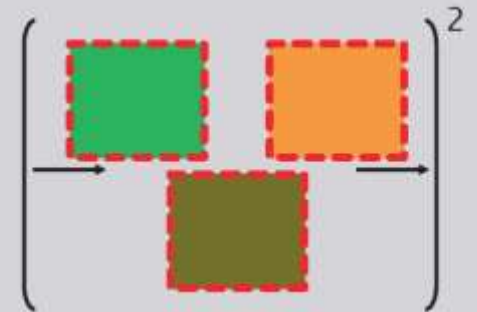
- Guilds maintain individual pathway functionalities
- Linked by the exchange of internal “pool” metabolites
- Partitioning of metabolites into external and internal pools

‘SOUP’



- Cell functions combined into one “soup”
- Maximum metabolic potential of the microbial community
- Community without cell boundaries

‘MODE²’

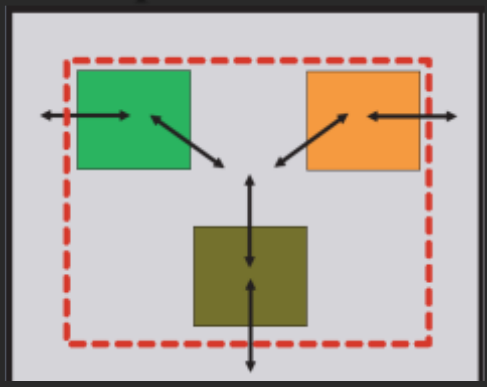


- Modes of interest are mined from individual model data
- Modes are combined as inputs for further analysis
- Individual vs. community benefits

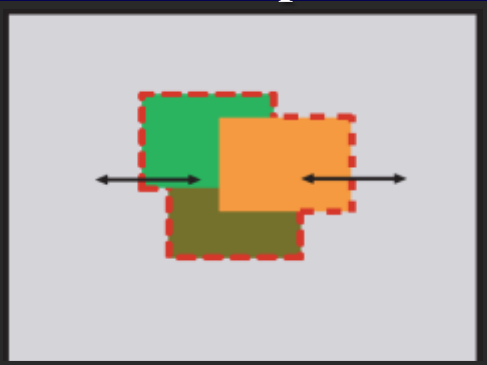


Community Analysis Tools

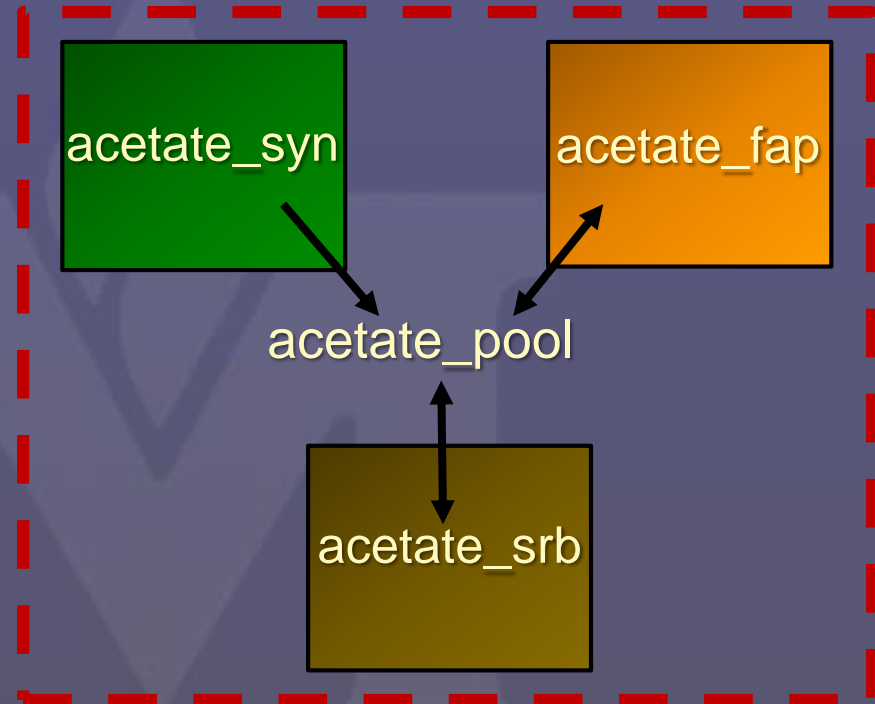
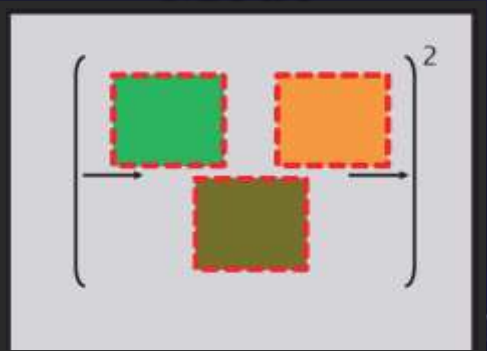
‘Compartmentalized’



‘soup’



‘Mode²’



$$\text{acetyl_CoA_syn} = \text{acetate_syn} + \text{ATP_syn}$$

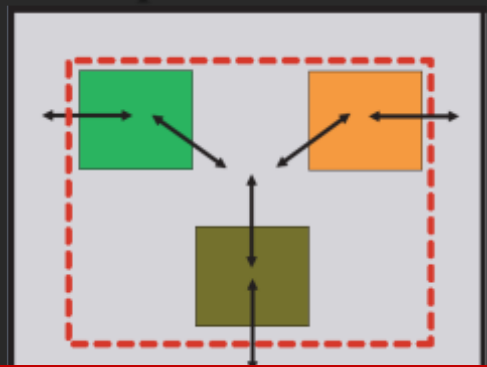
$$\text{acetate_syn} = \text{acetate_pool}$$

$$\text{acetate_pool} = \text{acetate_fap}$$

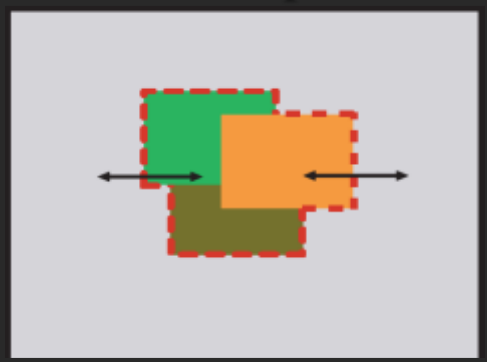
$$\text{acetate_fap} + 2 \text{ ATP_fap} = \text{acetyl_CoA_fap}$$

Community Analysis Tools

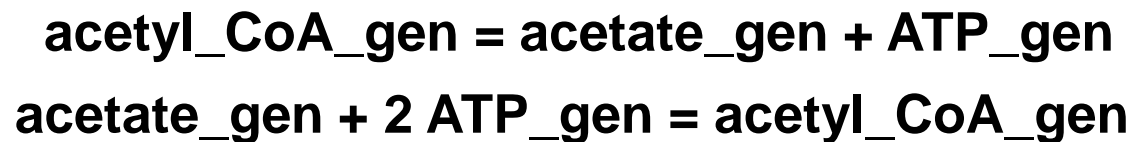
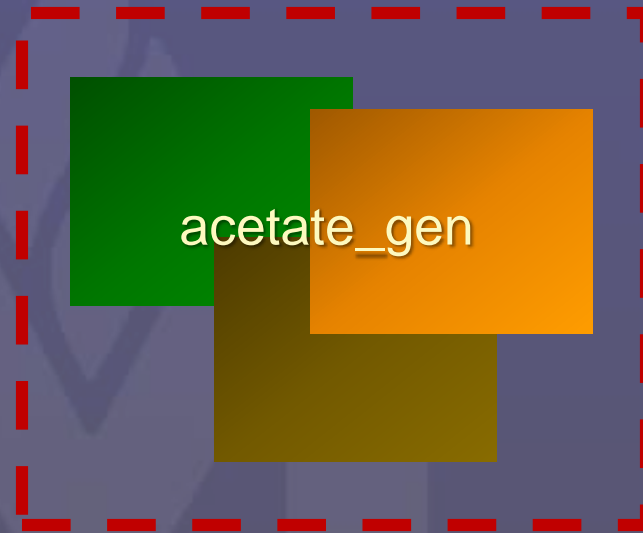
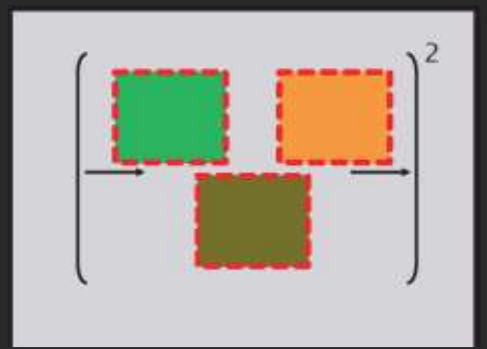
‘Compartmentalized’



‘Soup’

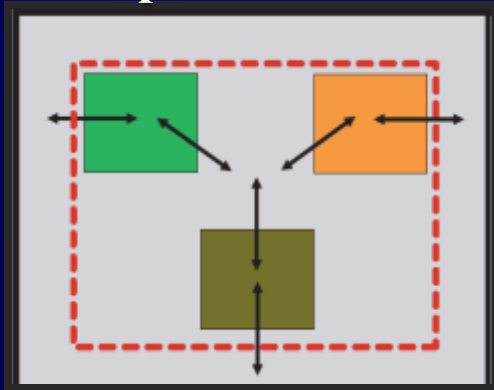


‘Modular’

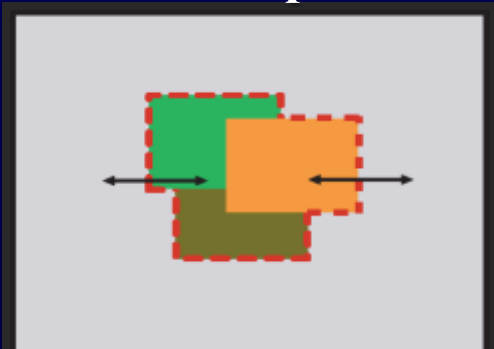


Community Analysis Tools

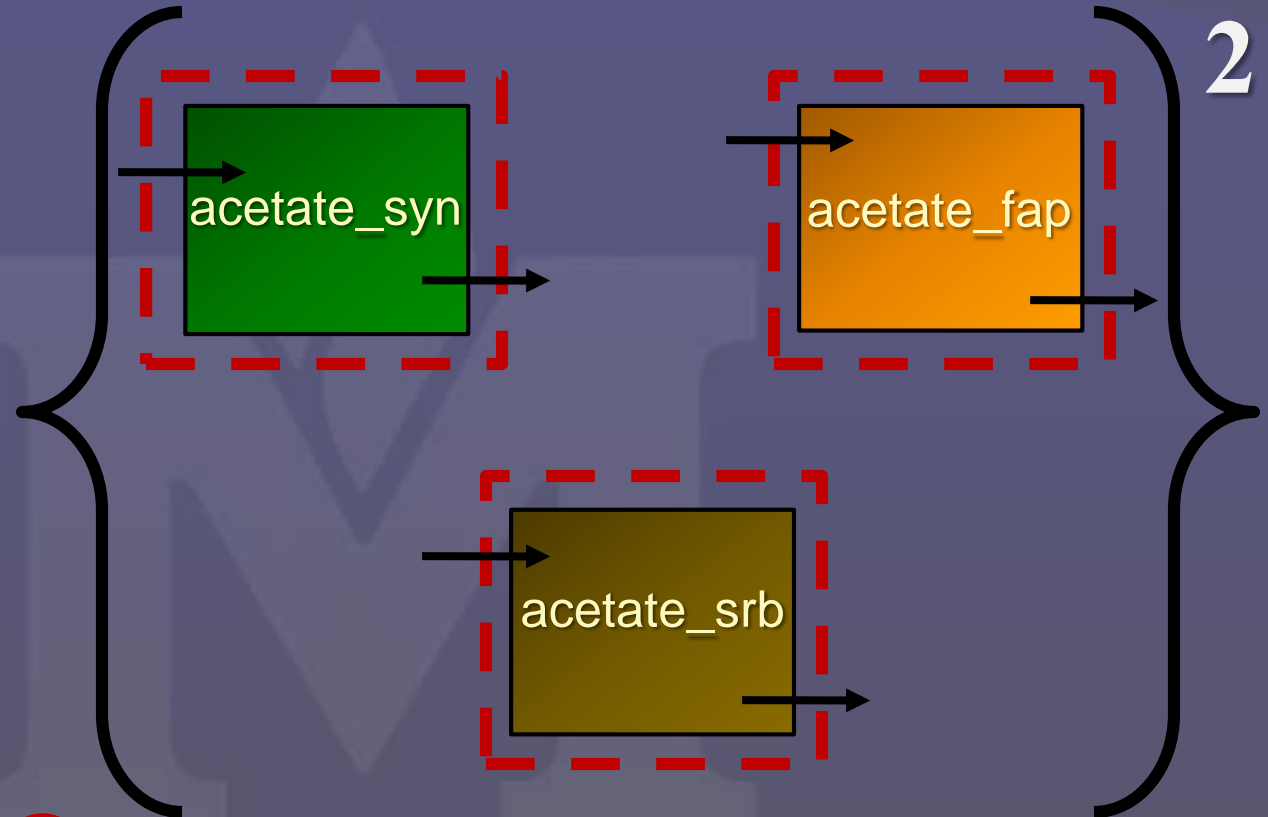
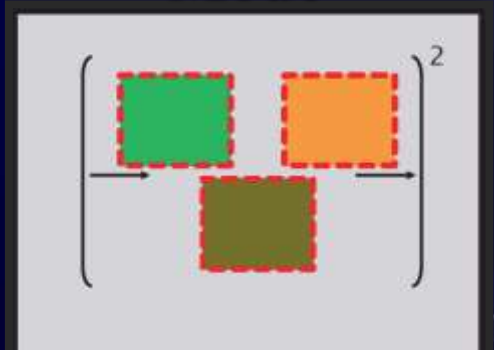
‘Compartmentalized’



‘Soup’



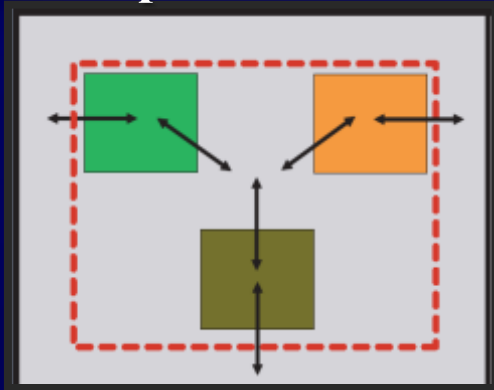
‘Mode²’



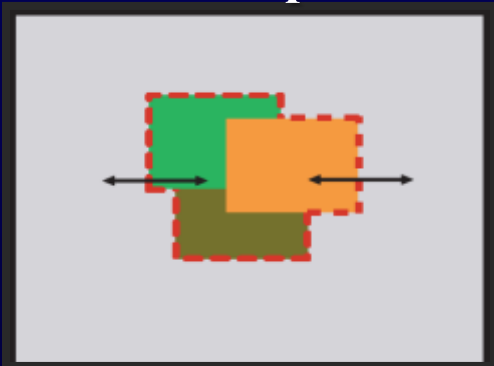
$$\begin{aligned}
 &2x \quad 12 \text{ } hv_ext + 2 \text{ } CO_2_ext = \text{acetate_ext} + 4 \text{ } O_2_ext \\
 &+ \\
 &1x \quad \text{acetate_ext} + 6 \text{ } O_2_ext = \text{biomass_fap} + 3 \text{ } CO_2_ext
 \end{aligned}$$

$$= 24 \text{ } hv_ext + CO_2_ext = \text{biomass_fap} + 2 \text{ } O_2_ext$$

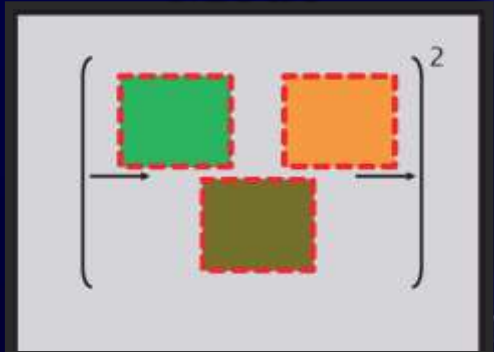
‘Compartmentalized’



‘Soup’



‘Mode²’



Output Parameters



74,507

14,004



Which pathways utilized?



38,216

36,233



Why?

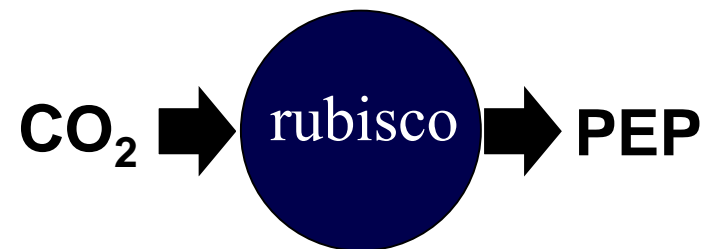
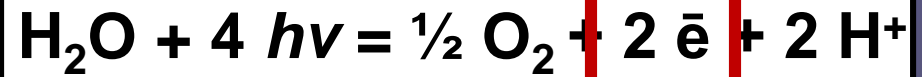
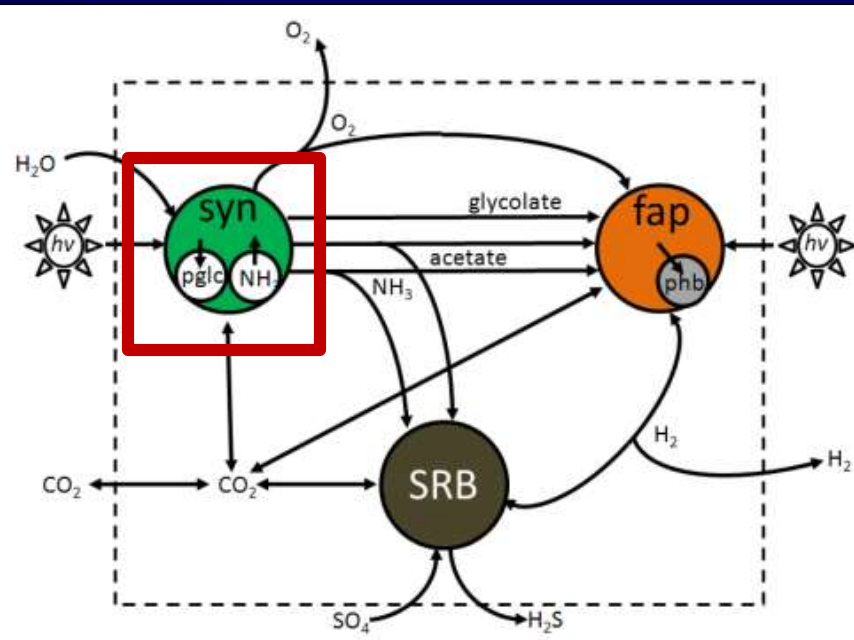


428

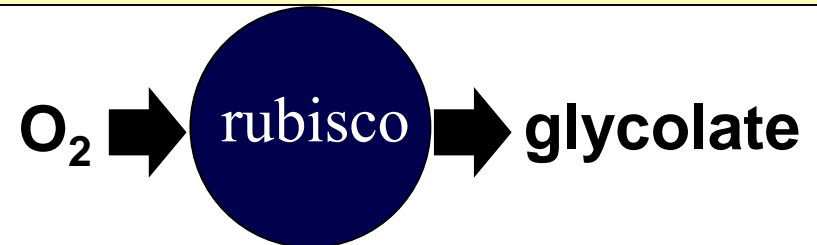
46,038



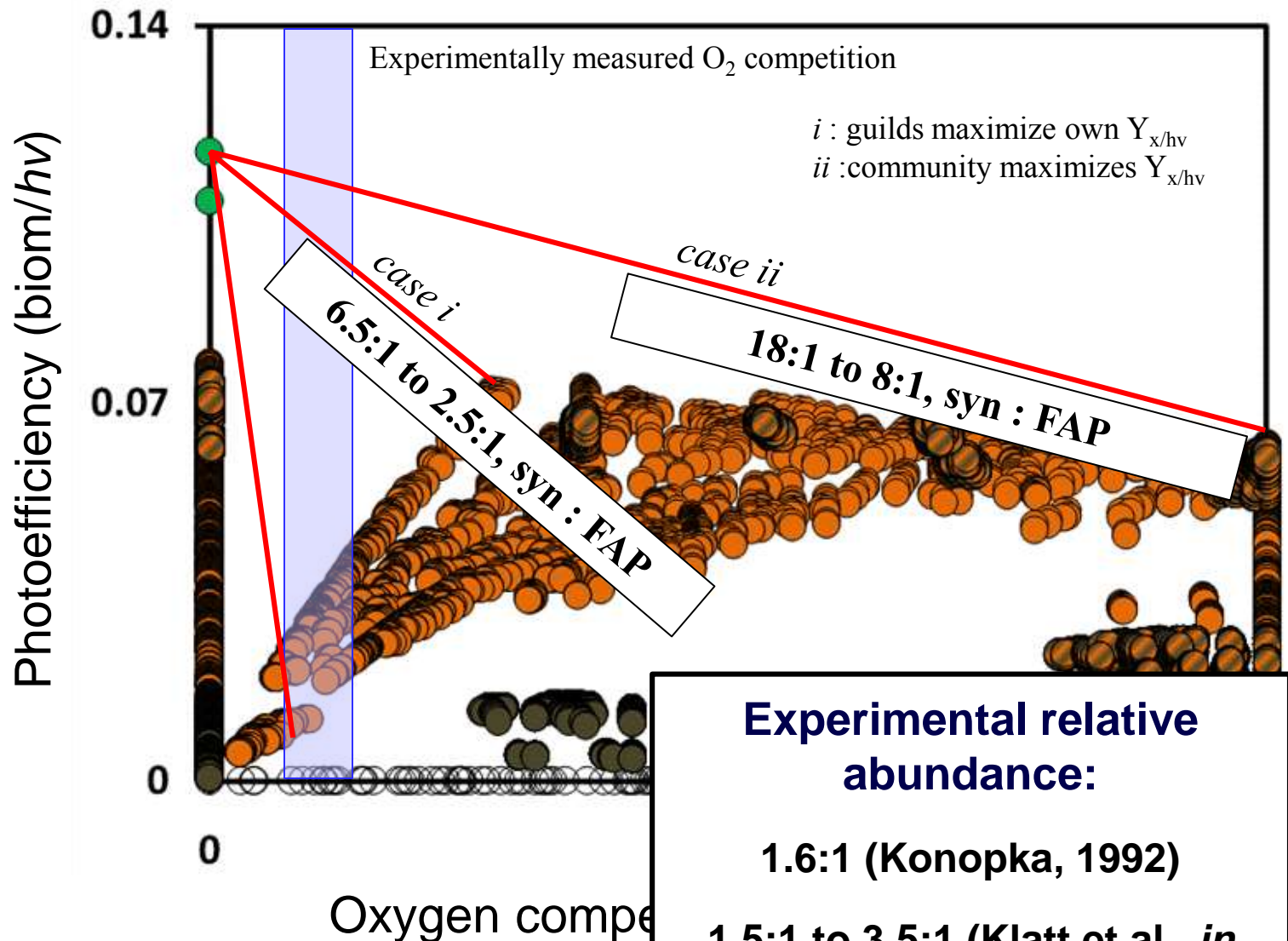
Explain/Predict: Community Photoefficiency



~20 CO_2 : 1 O_2 (Bateson and Ward, 1988)

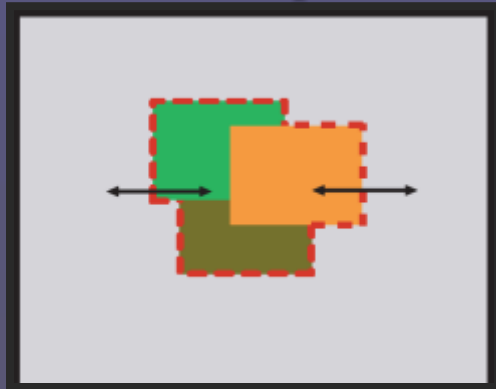


Community Photoefficiency

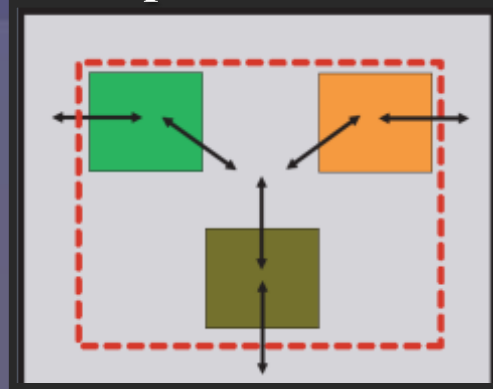


Comparison of Approaches: photoefficiency

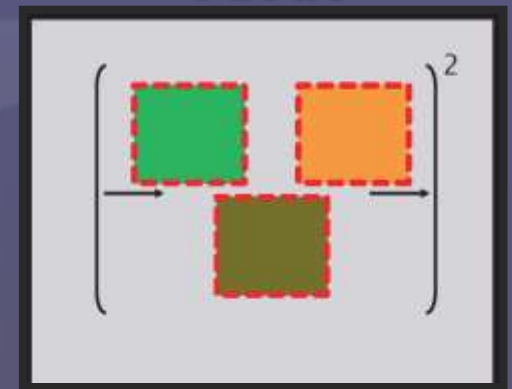
‘Soup’



‘Compartmentalized’



‘Mode²’



0.128

0.117

0.117

0.029

0.027

0.025

Cmole biomass / photon



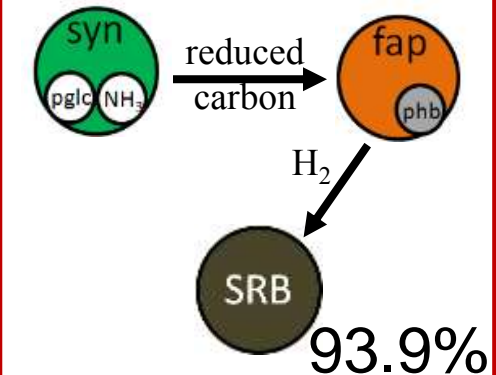
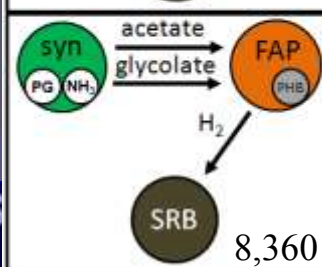
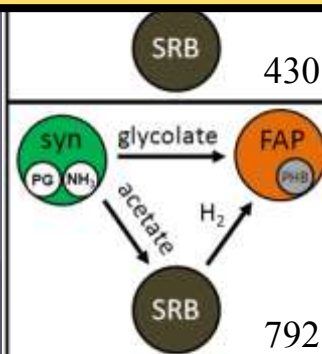
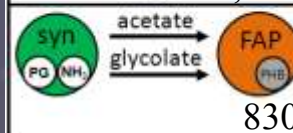
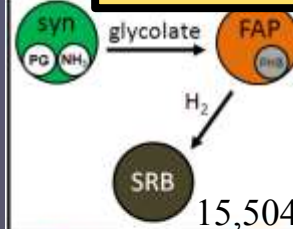
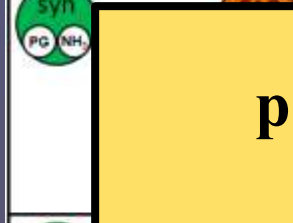
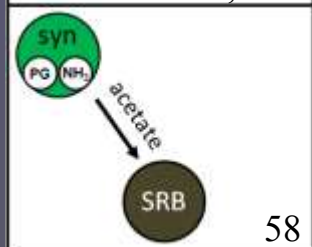
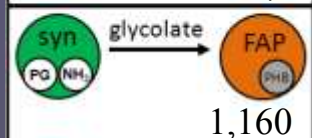
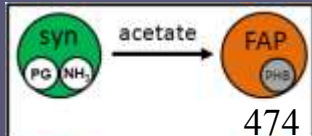
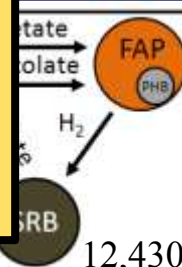
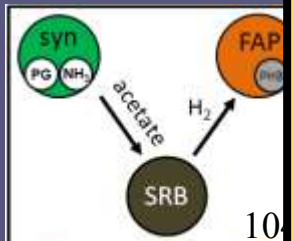
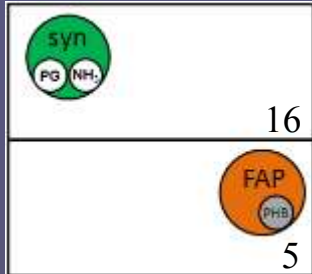
Community Interactions: Food Web

74,507 total strategies

associated
yields and
enzymes

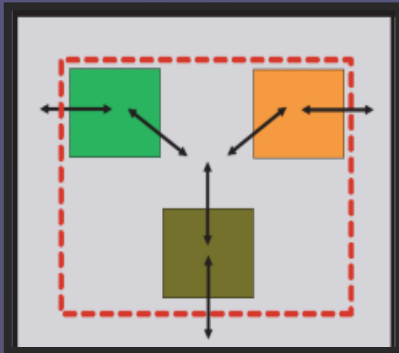
prediction = f(knowledge)

Common
Genes-to-Function
Motif

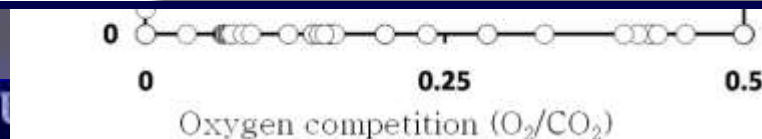
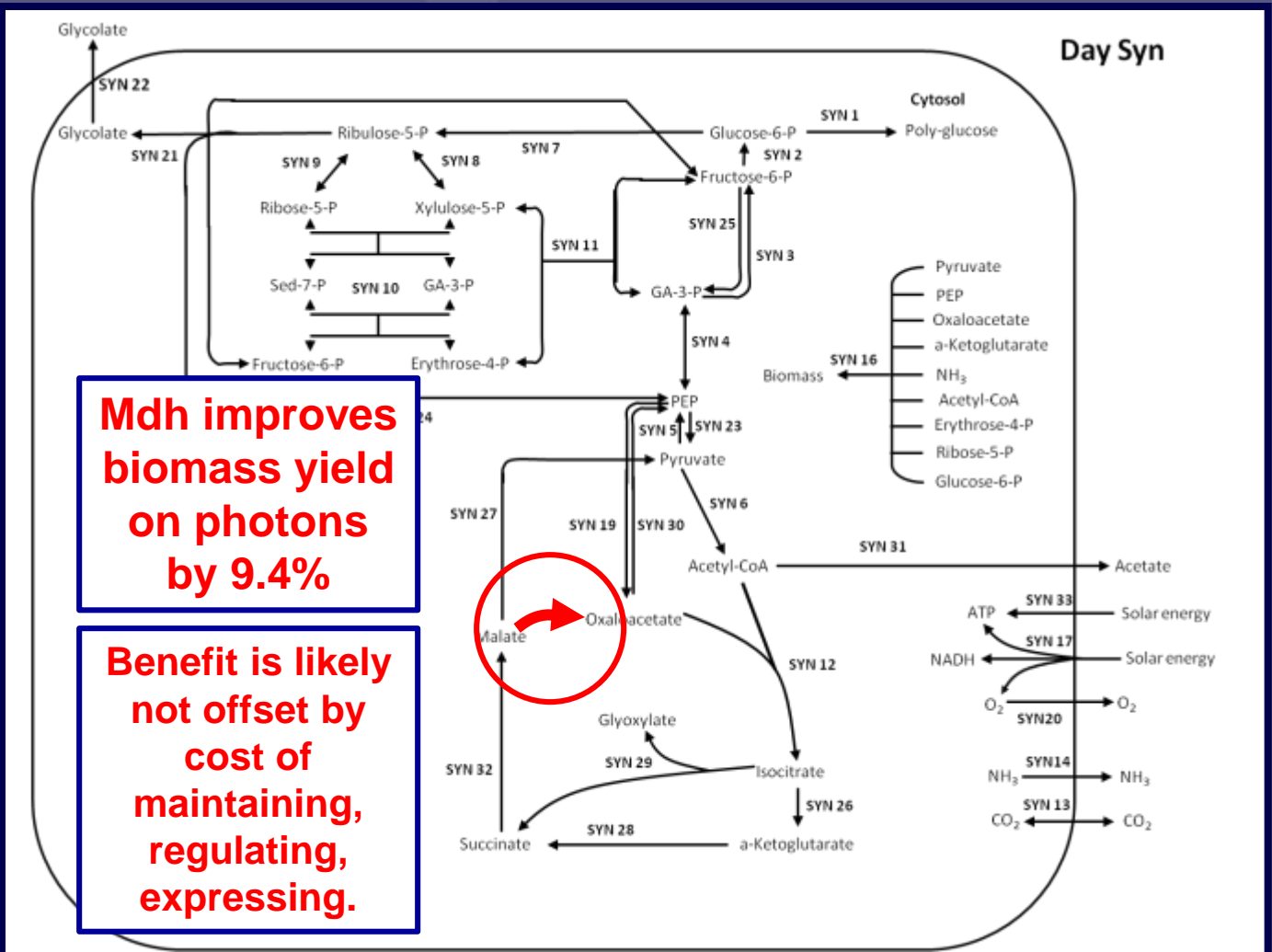
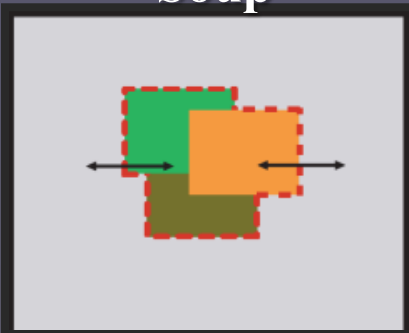


Evolutionary Cost Benefit: Compart. vs. 'Soup'

'Compartmentalized'



'Soup'



Control: Building a system from components



•Control community functioning
through understanding

Bioprocess Applications

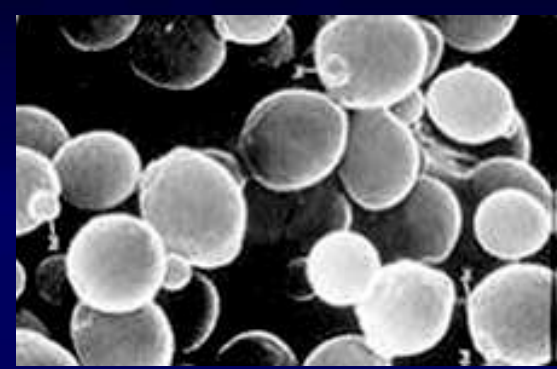


 **NatureWorks**
Real. Right. Renewable.



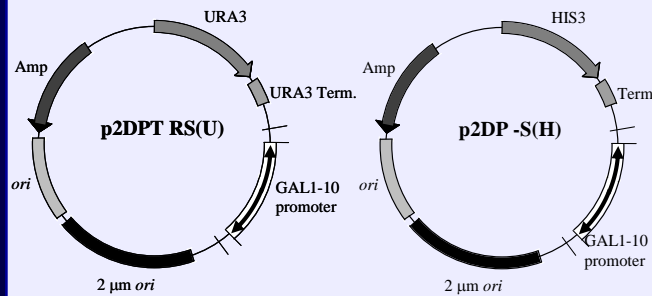
MONTANA STATE UNIVERSITY

Cell Factories: Network Structure Analysis



Fifth Kingdom

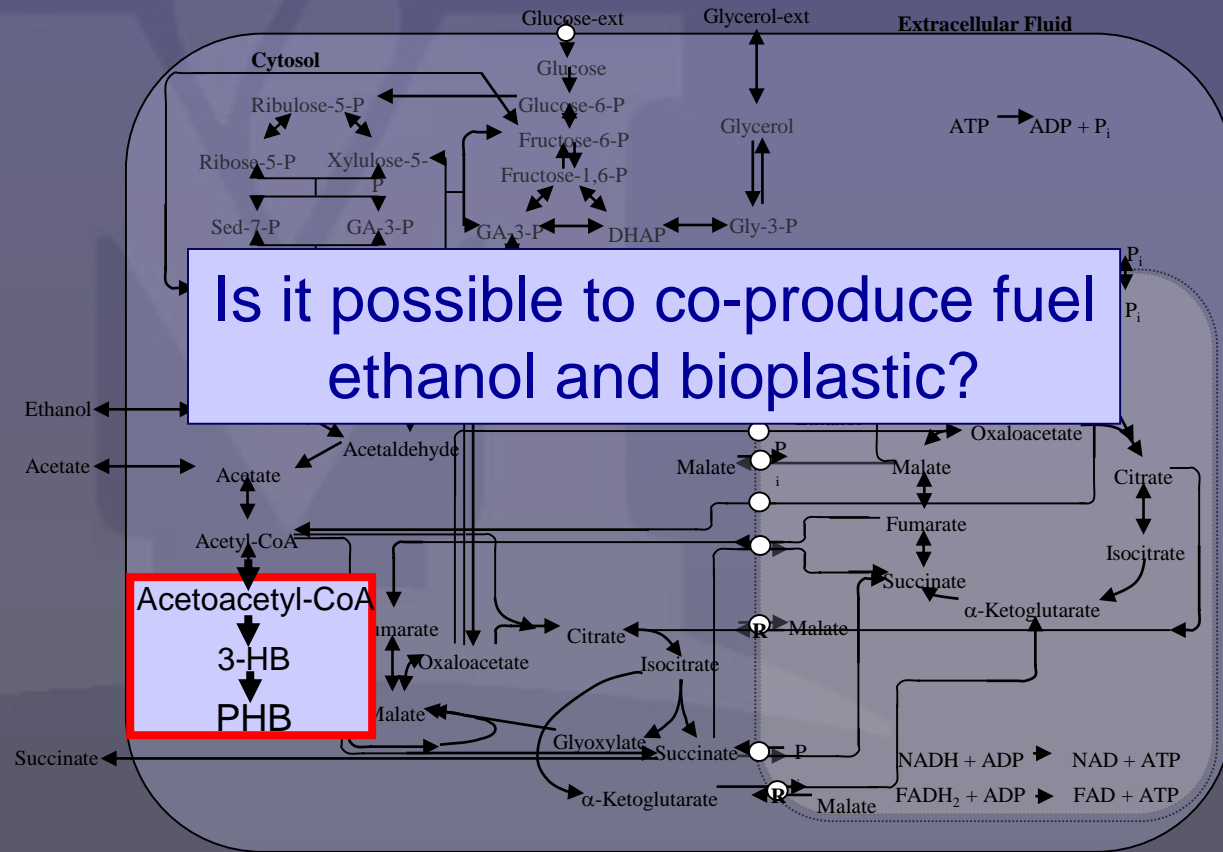
- Brewer's yeast :



- Genetically engineered to produce a bioplastic

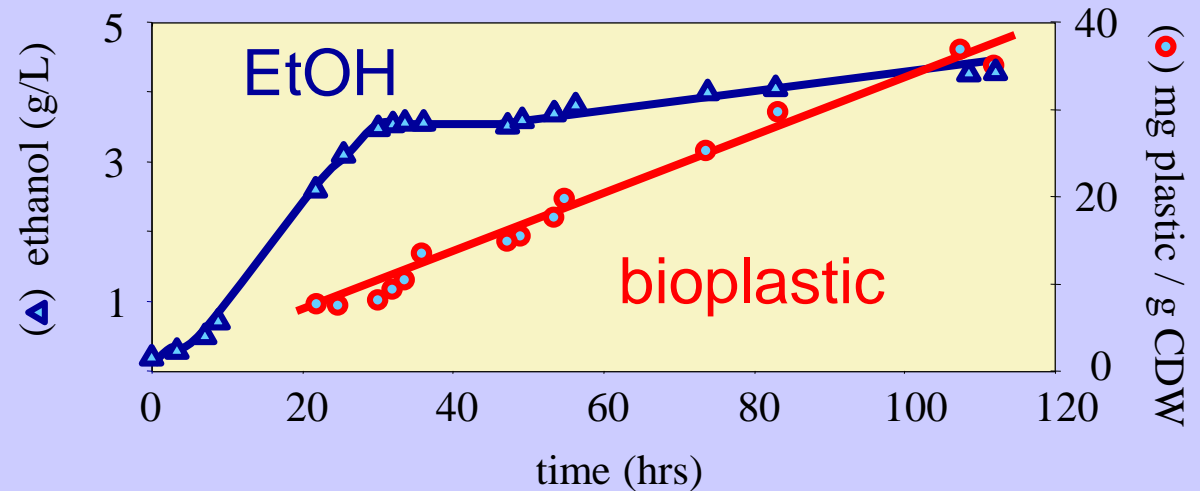
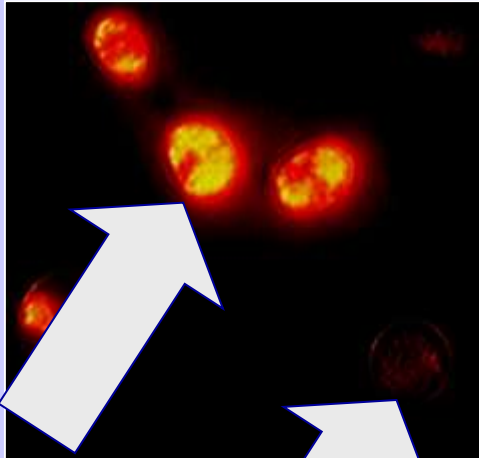
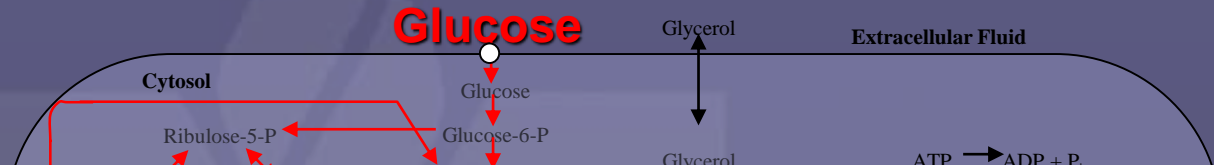
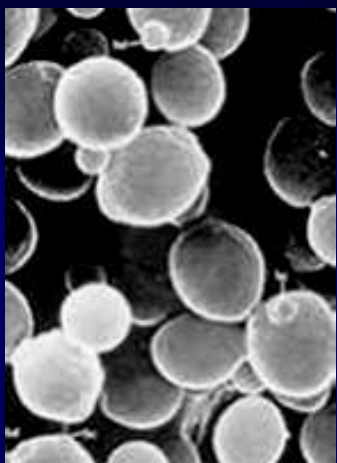


MONTANA STATE UNIVERSITY

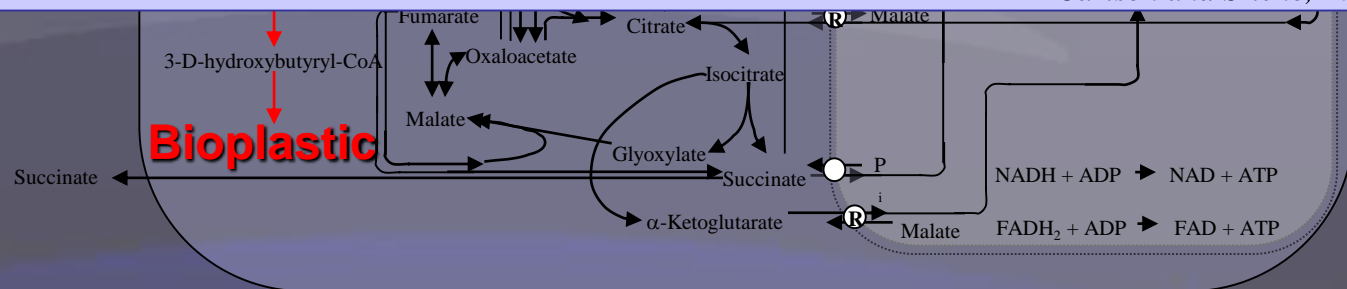


Mountains & Minds

Cell Factories: Network Structure Analysis



Carlson and Srienc, 2006



MONTANA STATE UNIVERSITY

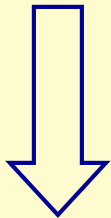
Mountains & Minds

Carlson et al., 2002

Network Engineering

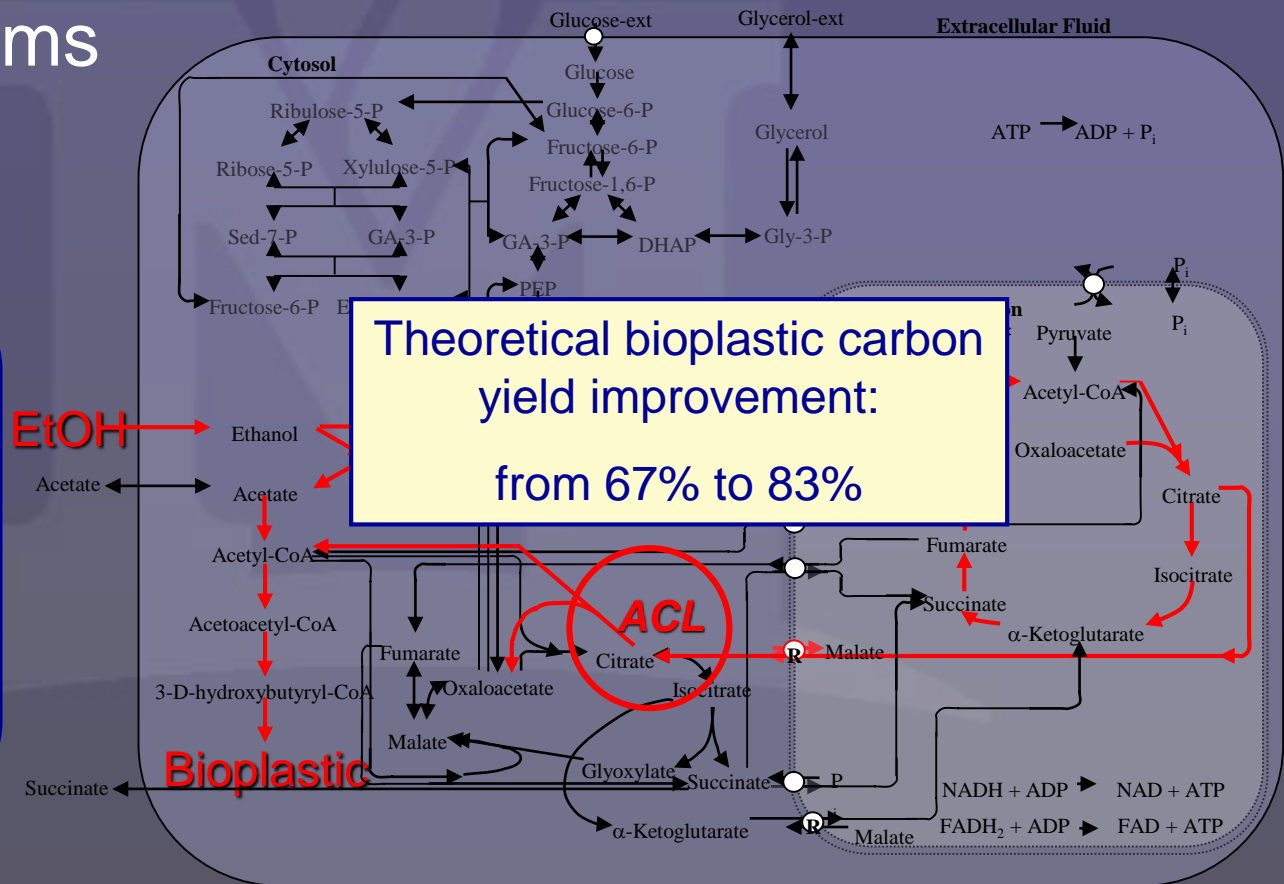
- How can we engineer systems to improve function?

Citrate + ATP

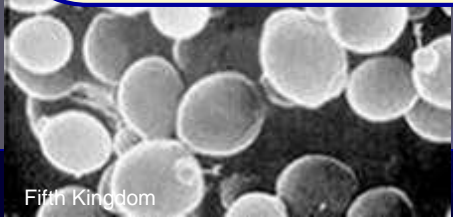


ACL

Acetyl-CoA +
Oxaloacetate



Carlson et al., 2002



Cell Factories: Predicting Recombinant Behavior

- Minimization of Metabolic Adjustment (Segre *et al.*, 2002)
- Euclidean distance between wild-type and perturbed state



98 anaerobic *E. coli* bioplastic pathways

What is the likely behavior?

$$D_j(M_j^{PHB}, M_4^{ATP}) = \sqrt{\sum_{i=1}^R (m_{i,j}^{PHB} - m_{i,4}^{ATP})^2}$$

j=1,98





Pathway:

3 glucose = 6 ATP + 2 EtOH + 6 formate + 2 PHB.

glucose + 2 acetate = 2 formate + 2 PHB

3 glucose + acetate = 4 EtOH + 5 formate

3 glucose + acetate = 2 EtOH + 6 formate

3 glucose + acetate + CO₂ = succinate

3 glucose = 3 ATP + 2 EtOH + 6 formate

Euclidean
distance:

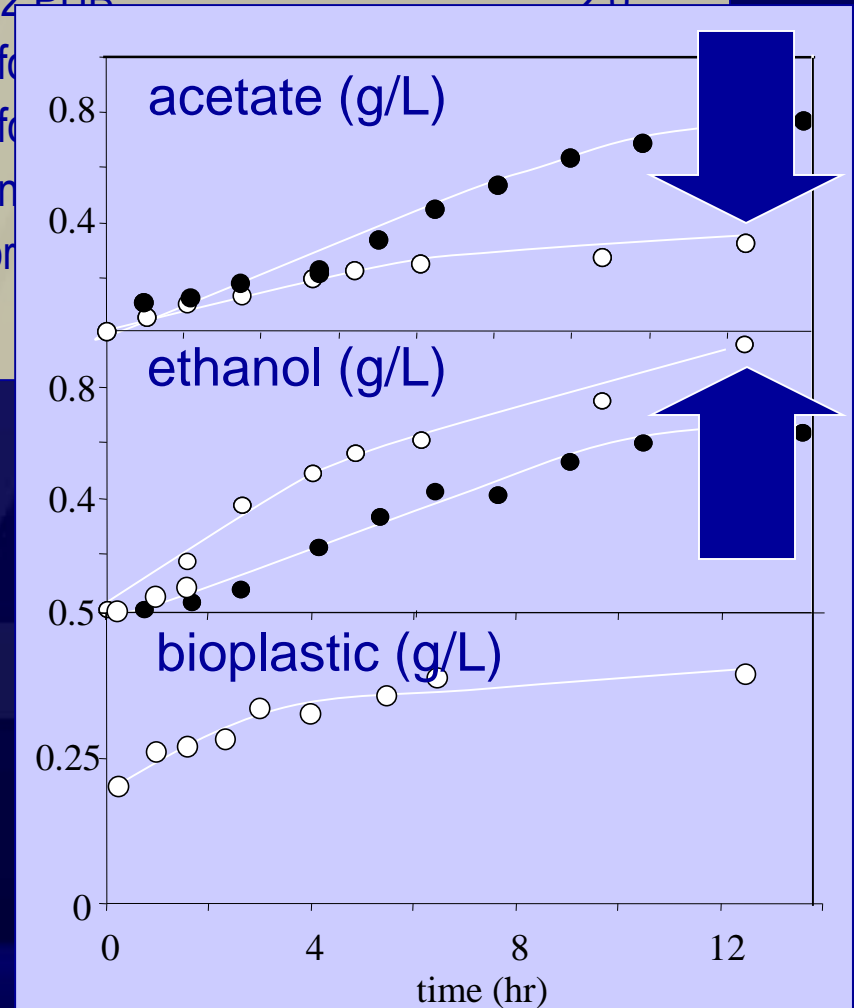
1.1

2.0

[acetate]



[ethanol]



MONTANA STATE UNIVERSITY

Mountains & Minds

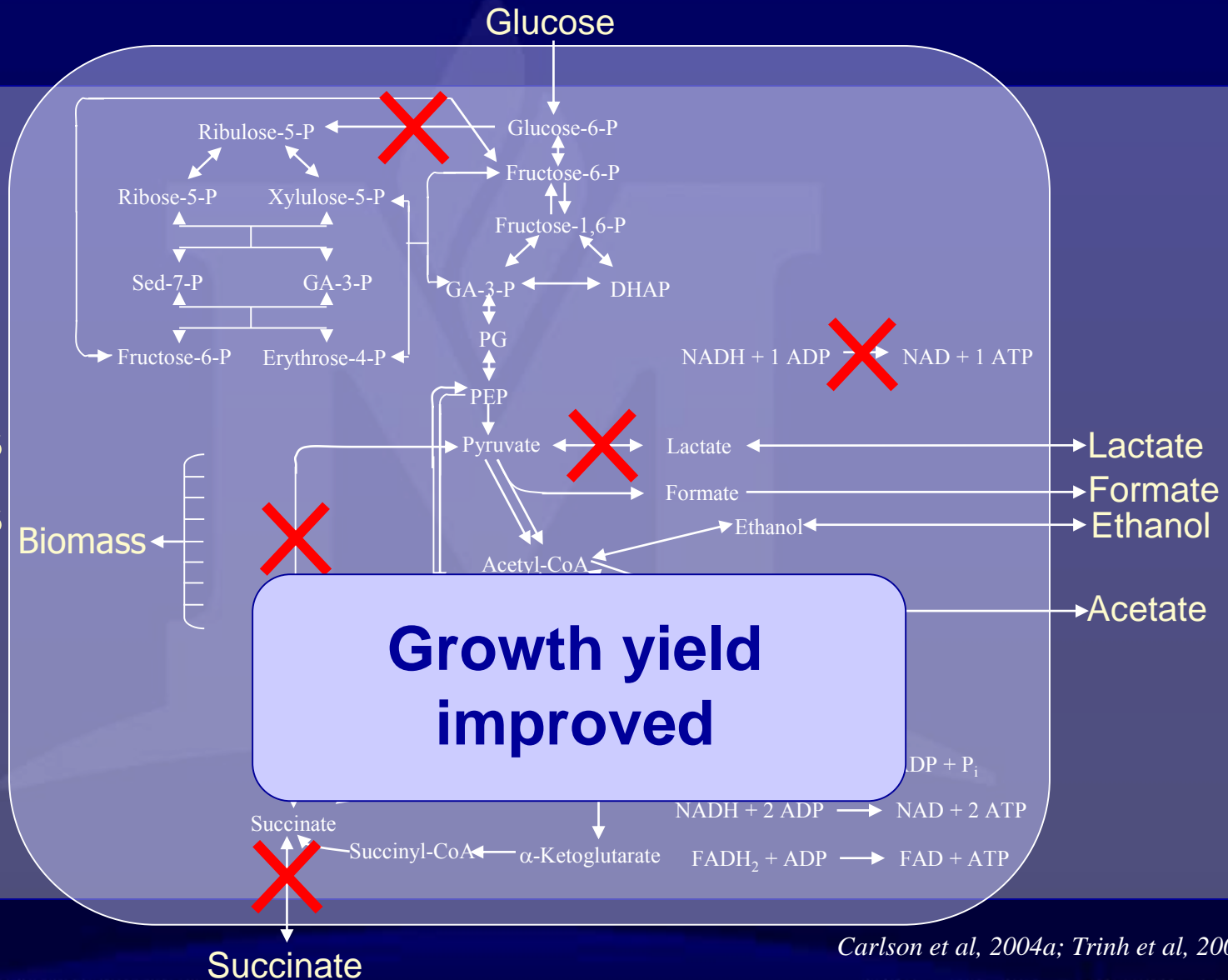
Carlson et al, 2005



Cell Factories: Engineering Network Fluxes

~~852~~ modes
~~189~~ modes
~~182~~ modes

4 modes



Growth yield improved

Carlson et al, 2004a; Trinh et al, 2006

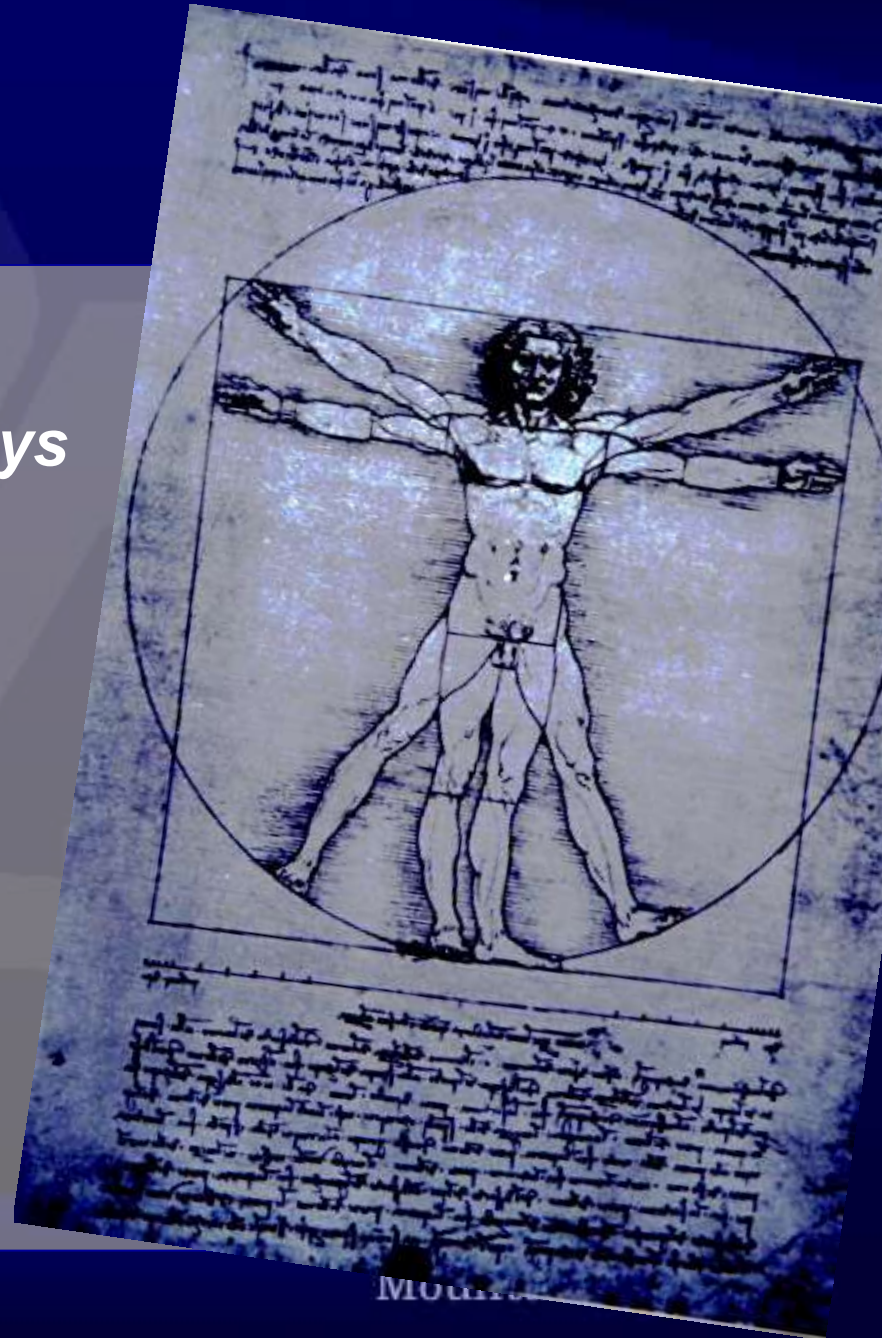


MONTANA STATE UNIVERSITY

Mountains & Minds

Practical: Supplemental

- *How are recombinant pathways added to metabolic model?*
- *How can we predict likely recombinant cell behavior?*

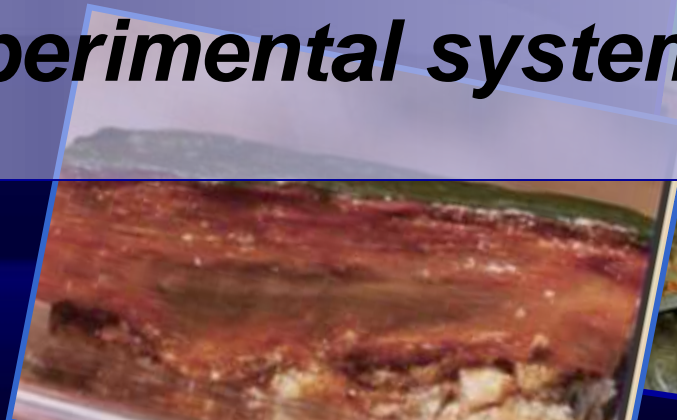


Conclusions

- *Acquired basic understanding and computational toolbox for original network analysis*
- *Define network capabilities*
- *Sort pathways based on cell function*
- *Identify competitive network strategies*
- *Applied to experimental systems*



MONTANA STATE UNIVERSITY





Acknowledgements

CENTER FOR
BIOFILM
ENGINEERING

Reed Taffs, James Folsom, Kristopher Hunt

NIH/NIBIB, IMAG, MSM
Idaho National Laboratory
EPSCoR
NSF IGERT Geobiological Systems
Center for Biofilm Engineering
Thermal Biology Institute



STATE UNIVERSITY

⚡ THERMAL BIOLOGY INSTITUTE