Chapter 9: Metabolic Networks

9.2 Analysis & Applications

Prof. Yechiam Yemini (YY)
Computer Science Department
Columbia University

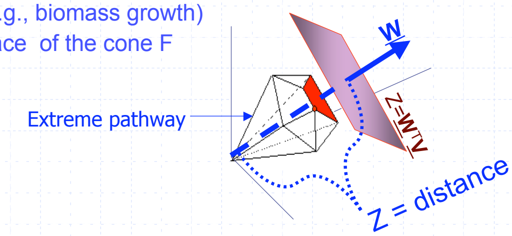
Overview

- Extreme pathways analysis
- Perturbations of metabolic networks
- Re-engineering metabolic networks

2

Summary Of Theory

- A state of a metabolic net is described by a flux vector
- Flux vectors constitute a convex cone of feasible states
  - $F=\{v \mid Sv=0, \underline{\alpha}\le\underline{v}\le\underline{\beta}\}$
  - This cone F describes the feasible modes available to the network
- The flux cone is spanned by its extreme pathways (fluxes)
  - Extreme pathways are subnets corresponding to extreme rays of F
  - A flux vector is a weighted sum of extreme fluxes
- Metabolic nets select fluxes to optimize linear objectives
  - Objective = weighted flux (e.g., biomass growth)
  - Optimal fluxes constitute a face of the cone F



3

Extreme Pathways Analysis

Example

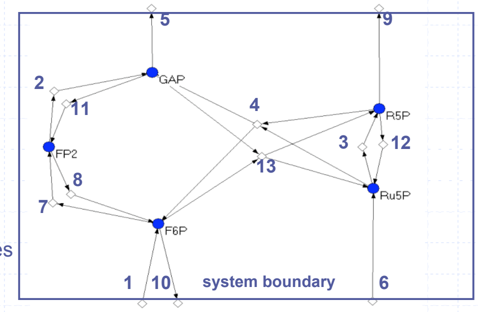
Based on M. Imielinski 2007; www.seas.upenn.edu/~agung/ese680files/imielinski.ppt

Legend

◇ reaction

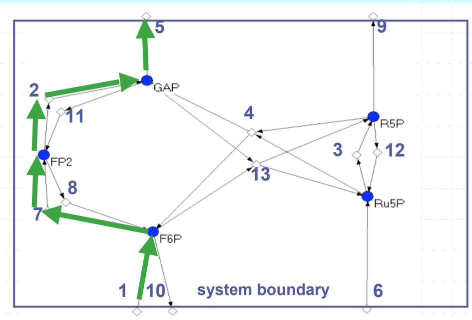
reaction I/O

● biochemical species



	1	2	3	4	5	6	7	8	9	10	11	12	13
Ru5P	0	0	-1	-2	0	1	0	0	0	0	0	1	2
FP2	0	-1	0	0	0	0	1	-1	0	0	1	0	0
F6P	1	0	0	2	0	0	-1	1	0	-1	0	0	-2
GAP	0	2	0	1	-1	0	0	0	0	0	-2	0	-1
R5P	0	0	1	-1	0	0	0	0	-1	0	0	-1	1

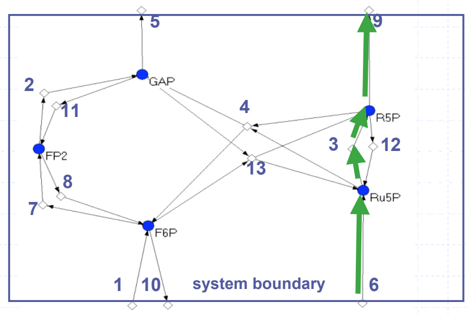
Example



	1	2	3	4	5	6	7	8	9	10	11	12	13
Extreme pathway 1	1	1	0	0	2	0	1	0	0	0	0	0	0
Extreme pathway 2	0	0	1	0	0	1	0	0	1	0	0	0	0
Extreme pathway 3	0	0	1	0	0	0	0	0	0	0	0	1	0
Extreme pathway 4	0	0	2	2	0	6	0	1	0	5	1	0	0

	1	2	3	4	5	6	7	8	9	10	11	12	13
Ru5P	0	0	-1	-2	0	1	0	0	0	0	0	1	2
FP2	0	-1	0	0	0	0	1	-1	0	0	1	0	0
F6P	1	0	0	2	0	0	-1	1	0	-1	0	0	-2
GAP	0	2	0	1	-1	0	0	0	0	0	-2	0	-1
R5P	0	0	1	-1	0	0	0	0	-1	0	0	-1	1

Example

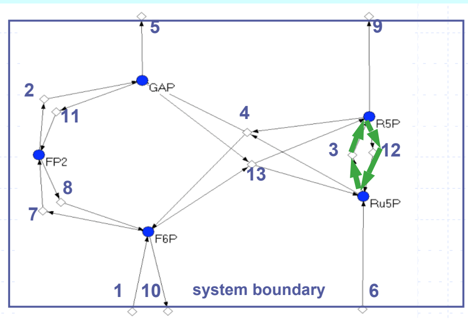


	1	2	3	4	5	6	7	8	9	10	11	12	13
Extreme pathway 1	1	1	0	0	2	0	1	0	0	0	0	0	0
Extreme pathway 2	0	0	1	0	0	1	0	0	1	0	0	0	0
Extreme pathway 3	0	0	1	0	0	0	0	0	0	0	0	1	0
Extreme pathway 4	0	0	2	2	0	6	0	1	0	5	1	0	0

	1	2	3	4	5	6	7	8	9	10	11	12	13
Ru5P	0	0	-1	-2	0	1	0	0	0	0	0	1	2
FP2	0	-1	0	0	0	0	1	-1	0	0	1	0	0
F6P	1	0	0	2	0	0	-1	1	0	-1	0	0	-2
GAP	0	2	0	1	-1	0	0	0	0	0	-2	0	-1
R5P	0	0	1	-1	0	0	0	0	-1	0	0	-1	1

7

Example

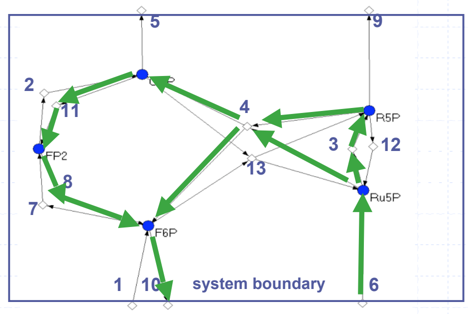


	1	2	3	4	5	6	7	8	9	10	11	12	13
Extreme pathway 1	1	1	0	0	2	0	1	0	0	0	0	0	0
Extreme pathway 2	0	0	1	0	0	1	0	0	1	0	0	0	0
Extreme pathway 3	0	0	1	0	0	0	0	0	0	0	0	1	0
Extreme pathway 4	0	0	2	2	0	6	0	1	0	5	1	0	0

	1	2	3	4	5	6	7	8	9	10	11	12	13
Ru5P	0	0	-1	-2	0	1	0	0	0	0	0	1	2
FP2	0	-1	0	0	0	0	1	-1	0	0	1	0	0
F6P	1	0	0	2	0	0	-1	1	0	-1	0	0	-2
GAP	0	2	0	1	-1	0	0	0	0	0	-2	0	-1
R5P	0	0	1	-1	0	0	0	0	-1	0	0	-1	1

8

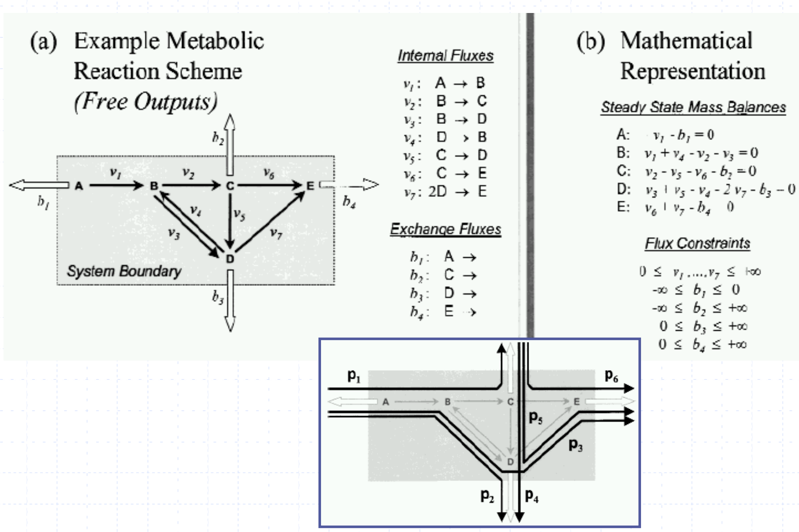
Example



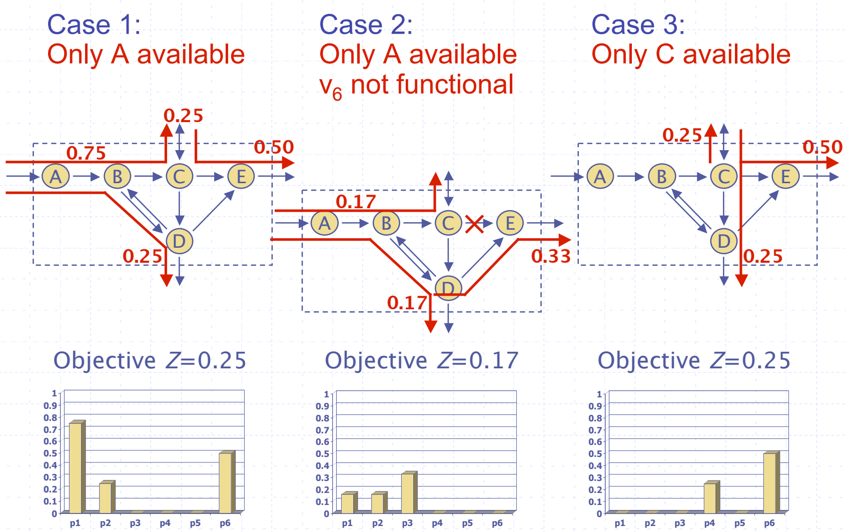
	1	2	3	4	5	6	7	8	9	10	11	12	13
Extreme pathway 1	1	1	0	0	2	0	1	0	0	0	0	0	0
Extreme pathway 2	0	0	1	0	0	1	0	0	1	0	0	0	0
Extreme pathway 3	0	0	1	0	0	0	0	0	0	0	0	1	0
Extreme pathway 4	0	0	2	2	0	6	0	1	0	5	1	0	0

	1	2	3	4	5	6	7	8	9	10	11	12	13
Ru5P	0	0	-1	-2	0	1	0	0	0	0	0	1	2
FP2	0	-1	0	0	0	0	1	-1	0	0	1	0	0
F6P	1	0	0	2	0	0	-1	1	0	-1	0	0	-2
GAP	0	2	0	1	-1	0	0	0	0	0	-2	0	-1
R5P	0	0	1	-1	0	0	0	0	-1	0	0	-1	1

Input Control With Extreme Pathways

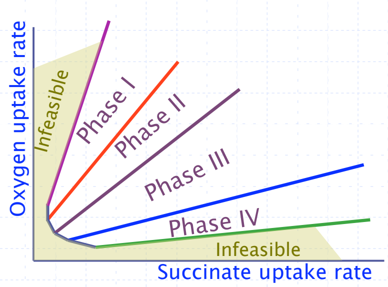


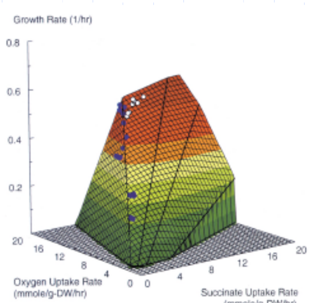
Input Control of Flux Distribution



Extreme Modes Describe Operational Phases

- Recall the E.coli FBA
- Consider the projections of extreme pathways onto the input space
- The regions bounded by EPs represent operational phases
  - A phase represents a subnet selected to process given inputs (optimally?)
  - E.g., diauxic shift
  - Subnet = weighted sum of respective EPs
  - A “phase” is defined by enzymes activating subnets (EPs)
  - The regulatory network selects a phase based on environmental input

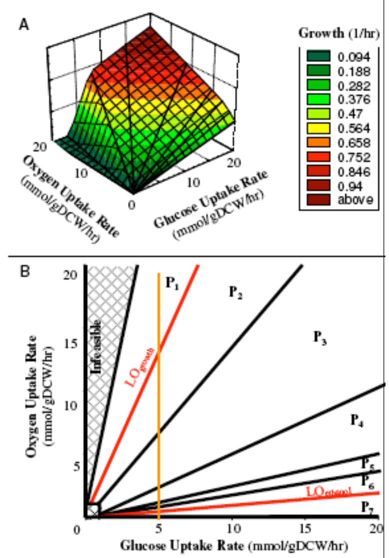




Input Control of Yeast Metabolism

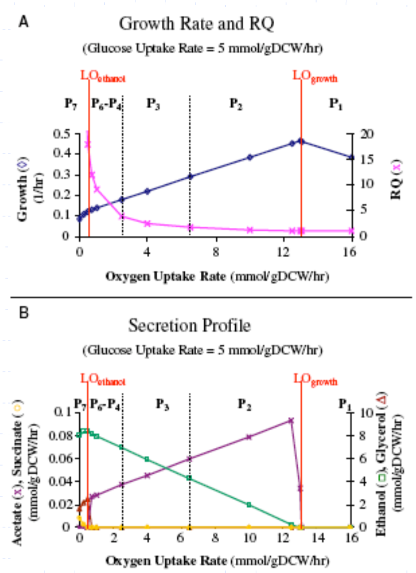
Duare et al. BMC Genomics 2004 5:63

- Challenge: genome-scale analysis
- Use phenotypical phases to simplify analysis
- Consider lines of optimality (LO)
  - Growth, ethanol production…

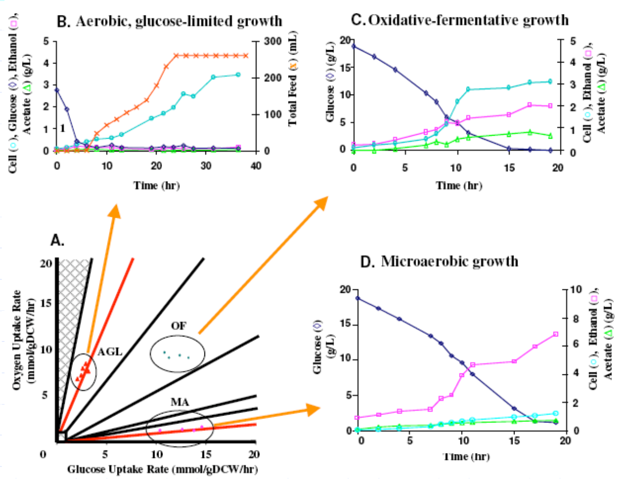


Growth Control Through Oxygen Supply

Simulation of metabolic behavior for optimal cellular growth as a function of oxygen availability, ranging from completely anaerobic fermentation to completely aerobic growth in *S. cerevisiae*. The range of oxygen uptake rates used in the simulations (orange line, Fig. 1) allows for the characterization of the PhPP's seven phases (P_1 - P_7) and two lines of optimality (LO_{growth}, $LO_{ethanol}$). **(a)** Growth rate and respiratory quotient (RQ). **(b)** Secretion profile for acetate, succinate, ethanol, and glycerol.



Input Control of Growth Modes



Growth experiments shown on the PhPP. (a) The three groups of experimental data displayed on the *S. cerevisiae* PhPP were used as an index for the time course profiles in panels (b), (c) and (d). **(b)** Aerobic glucose-limited growth controlled by fed-batch operation. **(c)** Oxidative-fermentative growth with unlimited glucose and oxygen availability. **(d)** Microaerobic growth with unlimited glucose and very low oxygen availability. The AGL (b) and MA (d) data sets are located on lines of optimality and as a result are stable metabolic states with only one degree of freedom (glucose for AGL and oxygen for MA). OF (c) is an unstable metabolic state with two degrees of freedom (glucose and oxygen), making it more difficult to control this type of growth condition. By perturbing the environmental conditions, cells in OF can be shifted to either AGL or MA (unpublished results).

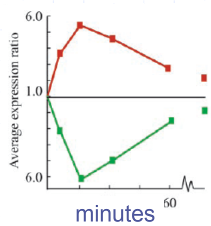
15

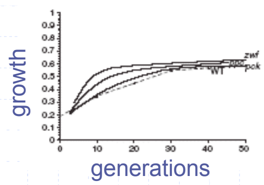
Perturbations of Metabolic Networks

In Silico Perturbations of Network

- Key idea: investigate impact of change on flux
  - Metabolic network can adapt to changes
  - Changes: availability of nutrients and enzymes

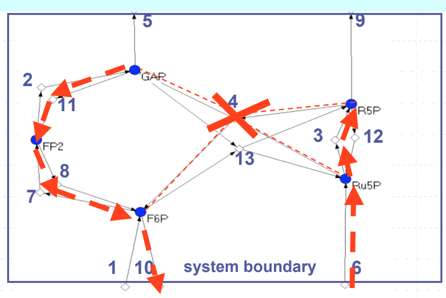
- Can in-silico models explain adaptation?
  - Change in inputs: limit the nutrients provided by medium
  - Change in network: delete reactions (genes)





17

In-Silico Gene Knock-Out



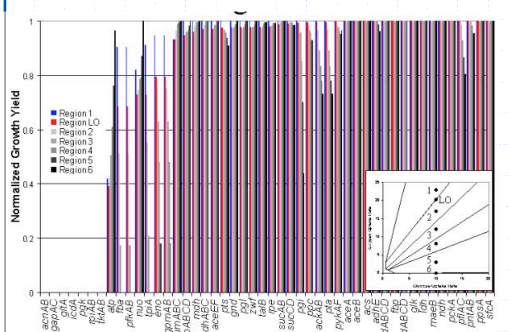
	1	2	3	4	5	6	7	8	9	10	11	12	13
Extreme pathway 1	1	1	0	0	2	0	1	0	0	0	0	0	0
Extreme pathway 2	0	0	1	0	0	1	0	0	1	0	0	0	0
Extreme pathway 3	0	0	1	0	0	0	0	0	0	0	0	1	0
~~Extreme pathway 4~~	~~0~~	~~0~~	~~2~~	~~2~~	~~0~~	~~6~~	~~0~~	~~1~~	~~0~~	~~5~~	~~1~~	~~0~~	~~0~~

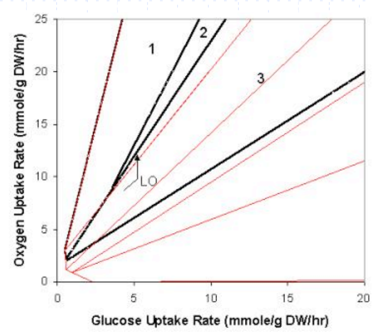
- Knock out the enzyme (gene) of reaction 4
  - Gene→ enzyme →reaction
- EPs using reaction 4 are lost (e.g., EP 4)
- Resulting in loss of dependent subnets and fluxes (e.g., reaction 10)

18

Analysis of E.Coli

Edwards et al. BMC Bioinformatics 2000 1:1 doi:10.1186/1471-2105-1-1





- Phase analysis of E.coli metabolism
- Consider impact of deleting genes on optimal growth

19

In-Silico Gene Knock-Out

The Escherichia coli MG1655 in silico metabolic genotype..
J. S. Edwards, and B. O. Palsson PNAS 2000;97;5528-5533

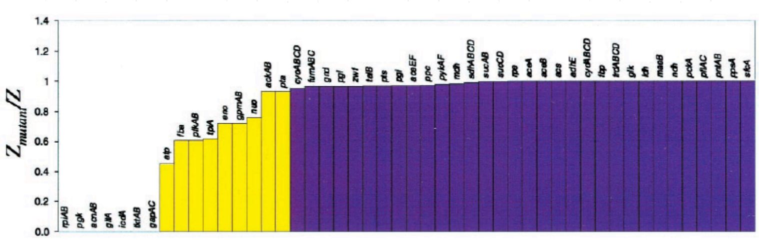


Fig. 2. Gene deletions in *E. coli* MG1655 central intermediary metabolism; maximal biomass yields on glucose for all possible single gene deletions in the central metabolic pathways. The optimal value of the mutant objective function (Z_{mutant}) compared with the "wild-type" objective function (Z), where Z is defined in Eq. 3. The ratio of optimal growth yields (Z_{mutant}/Z). The results were generated in a simulated aerobic environment with glucose as the carbon source. The transport fluxes were constrained as follows: $\beta_{glucose}$ = 10 mmol/g-dry weight (DW) per h; β_{oxygen} = 15 mmol/g-DW per h. The maximal yields were calculated by using FBA with the objective of maximizing growth. The biomass yields are normalized with respect to the results for the full metabolic genotype. The yellow bars represent gene deletions that reduced the maximal biomass yield to less than 95% of the *in silico* wild type.

- Impact of gene deletions on growth is computed

20

In Silico Gene Knock-Out

Results are scored as + or − meaning growth or no growth determined from *in vivo*/*in silico* data. The ± indicates that suppressor mutations have been observed that allow the mutant strain to grow. In 68 of 79 cases the *in silico* behavior is the same as the experimentally observed behavior. glc, glucose; ac, acetate; gl, glycerol; succ, succinate.

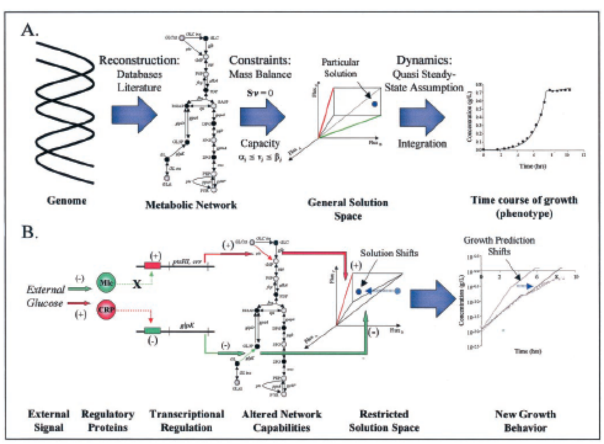
Gene	glc	gl	succ	ac	Reference
aceA	+/+		+/+	−/−	(58)
aceB				−/−	(58)
*aceEF**	−/+				(60)
ackA				+/+	(61)
acn	−/−			−/−	(58)
acs				+/+	(61)
cyd	+/+				(62)
cyo	+/+				(62)
eno†	−/+	−/+	−/−	−/−	(30)
fba‖	−/+				(30)
fbp	+/+	−/−	−/−	−/−	(30)
frd	+/+		+/+	+/+	(60)
gap	−/−	−/−	−/−	−/−	(30)
glk	+/+				(30)
gltA	−/−			−/−	(58)
gnd	+/+				(30)
idh	−/−			−/−	(58)
mdh††	+/+	+/+	+/+		(63)
ndh	+/+	+/+			(59)
nuo	+/+	+/+			(59)
pfk†	−/+				(30)
pgi‡	+/+	+/−	+/−		(30)
pgk	−/−	−/−	−/−	−/−	(30)
pgl	+/+				(30)
pntAB	+/+	+/+	+/+		(29)
ppc§	±/+	−/+	+/+		(63, 64)
pta				+/+	(61)
pts	+/+				(30)
pyk	+/+				(30)
rpi	−/−	−/−	−/−	−/−	(30)
sdhABCD	+/+		−/−	−/−	(58)
sucAB	+/+		−/+	−/+	(60)
tktAB	−/−				(30)
*tpi***	−/+	−/−	−/−	−/−	(30)
unc	+/+		±/+	−/−	(66–68)
zwf	+/+	+/+	+/+		(30)

- 86% predictions
- What is the source of errors?
  - FBA ignores regulatory changes

21

Incorporating Regulatory Interactions Model

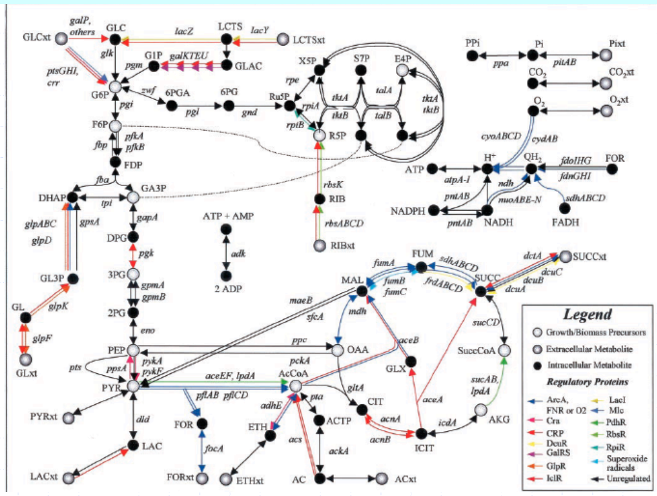
Covert & Palsson. J. of Bio Chem, vol 277 (31) 2002



- Key idea: consider quasi-static Boolean regulation
  - Compute flux distribution for a regulatory state

22

Metabolic/Regulatory Interactions in E.Coli



- Extend FBA with regulatory model of gene expression
- Use simulation to study quasi-static dynamics

Mutants Analysis

Input metabolite (glucose, glycerol, succinate…)

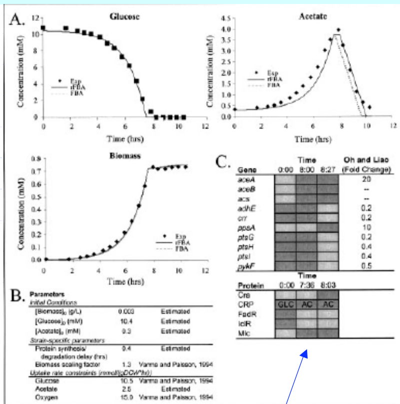
gene

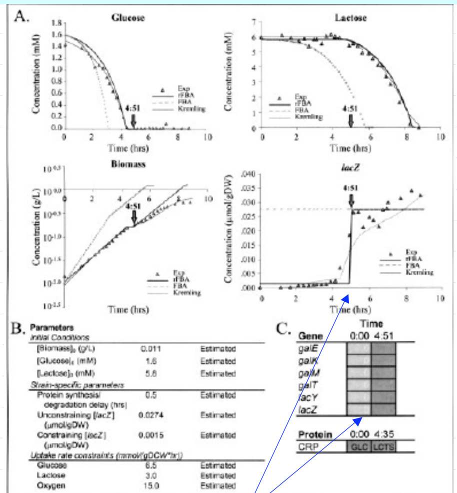
Growth predictions:
In-vivo/FBA/rFBA

	glc	gl	suc	ac	rib	glc ($-O_2$)	Dual Substrates	Ref
aceA	+/+/+		+/+/+	-/-/-		+/+/+		21
aceB				-/-/-				22
aceEF	-/+/-		-/+/-	+/+/+		+/+/+	(glc-ac) +/+/+	23
ackA				+/+/+				24
ackA + *pta* + *acs*				-/-/-				24
acnA	+/+/+	+/+/+	+/+/+	+/+/+		+/+/+		22,25
acnB"	+/+/+	+/+/+	+/+/+	-/+/+		+/+/+		25
acnA + *acnB*	-/-/-	-/-/-	-/-/-	-/-/-		-/-/-		25
acs				+/+/+				24
adh[†]	+/+/+					-/+/+		26
cyd	+/+/+							27
cyo	+/+/+							27
eno	-/-/-	-/-/-	-/-/-				(gl-suc) +/+/+	28

FBA has some errors
But rFBA gets 100% hits

Quasi-Static Dynamic Interactions





Regulatory state

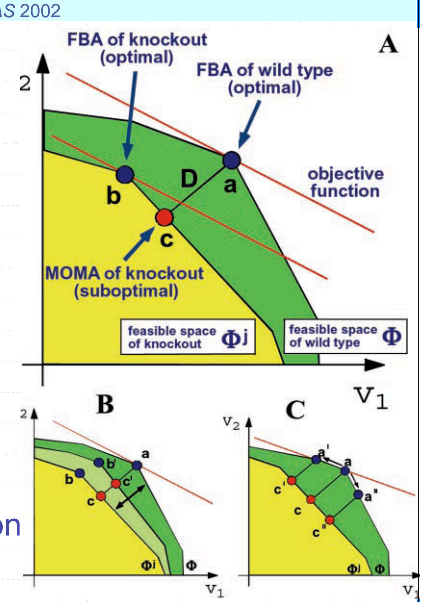
Diauxic Shift

25

Minimization of Metabolic Adjustment (MOMA)

Segre et al *PNAS* 2002

- FBA models result of gene deletion as re-optimizing
  - Network handles deletion by computing new optimal flux
- MOMA considers an alternate model for handling deletions
  - Network tries to operate "close" to the wild-type flux
  - "Close" = minimize Euclidian distance from wild-type flux
- Leads to a quadratic optimization

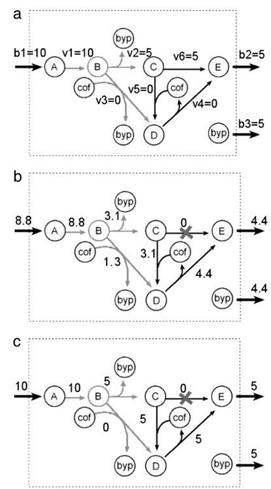


26

Regulatory On-Off Minimization (ROOM)

Shlomi et al *PNAS* 2005

- An alternate model
- Assume post-deletion flux is trying to approximate wild-type flux
- Measure of “closeness” is # of regulatory changes



27

ROOM➔ MILP

- ROOM Optimization

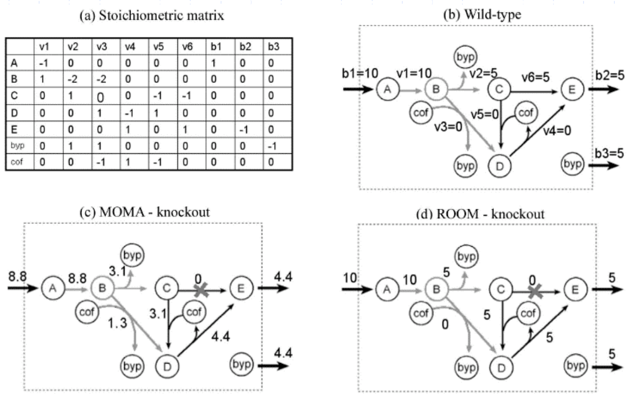
$y_i = 1$ ↔ Flux v_i change from wild-type

Min $\sum y_i$	- minimize changes
s.t	
$v - y\,(v_{max} - w) \le w$	- distance constraints
$v - y\,(v_{min} - w) \ge w$	- distance constraints
$S \cdot v = 0$,	- mass balance constraints
$v_j = 0,\ j \in G$	- knockout constraints

- Mixed Integer Linear Programming (MILP)
  - This is NP Hard
  - Relax constraints solve LP to approximate solution

28

Example: FBA/MOMA/ROOM

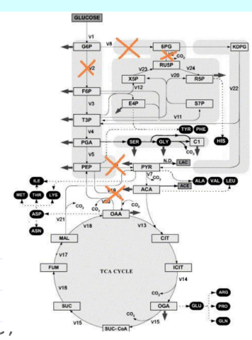


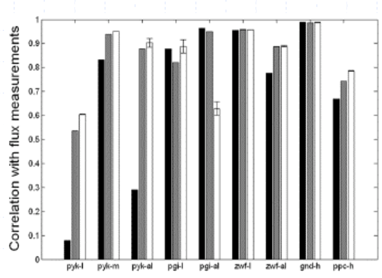
(a) Stoichiometric matrix

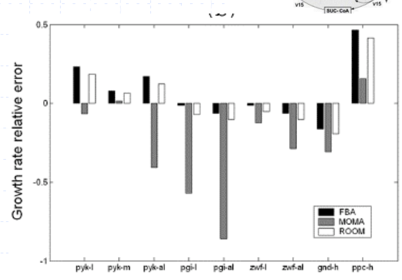
	v1	v2	v3	v4	v5	v6	b1	b2	b3
A	-1	0	0	0	0	0	1	0	0
B	1	-2	-2	0	0	0	0	0	0
C	0	1	0	0	-1	-1	0	0	0
D	0	0	1	-1	1	0	0	0	0
E	0	0	0	1	0	1	0	-1	0
byp	0	1	1	0	0	0	0	0	-1
cof	0	0	-1	1	-1	0	0	0	0

ROOM>FBA≥MOMA

- Intracellular fluxes measured by tracing ^{13}C
- Knock outs: pyk, pgi, zwf, ppc
- FBA gets over 90% wild-type predictions





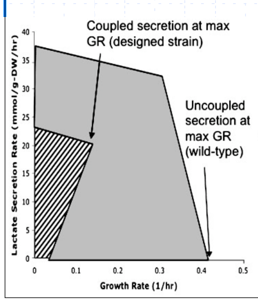


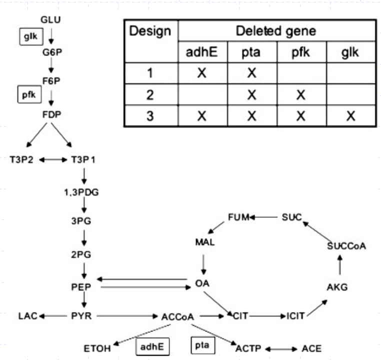
Metabolic Networks Re-Engineered

Optimizing Lactate Production

Fong, S.S. et al, Biotechnology and Bioengineering, 91(5):643-648 (2005).

- Key idea: knock-out genes to optimize lactate flux
- Compute optimal genes to max lactate flux (Optknock)
- Evolve mutants in media with inputs to optimize production

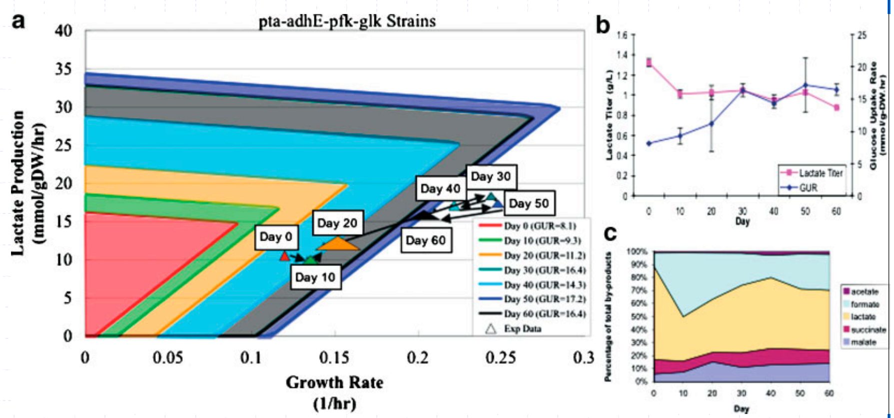




Design	Deleted gene			
	adhE	pta	pfk	glk
1	X	X		
2		X	X	
3	X	X	X	X

32

Lactate Production



Adaptation Effects

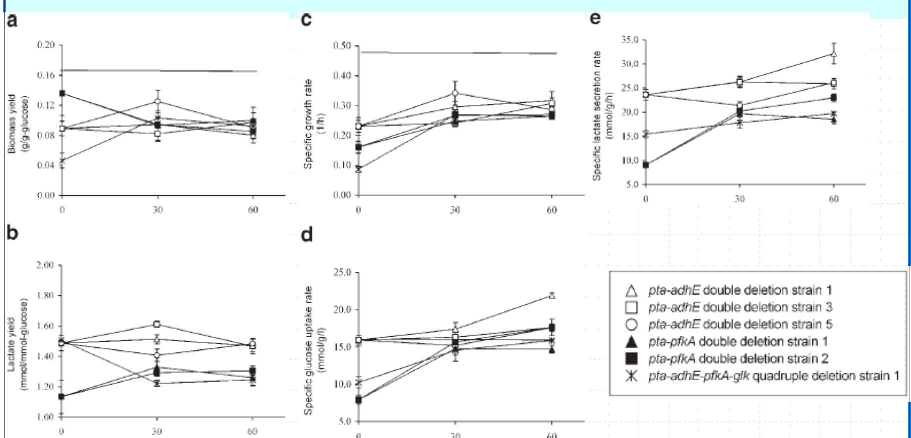


Figure 2. Physiological data for unevolved and evolved lactate-producing strains. Each plot depicts the change in various physiological characteristics over adaptive evolutionary time. Each line is associated with individual OptKnock-designed lactate-producing strains. The measured parameters shown in the plots are as follows: (**a**) Biomass yield; (**b**) Lactate yield; (**c**) Specific growth rate; (**d**) Specific glucose uptake rate; (**e**) Specific lactate production rate. The straight line shown in plots (a), (c), and (d) indicates the corresponding physiological parameter for the unevolved wild-type strain.

Regulatory Adaptation

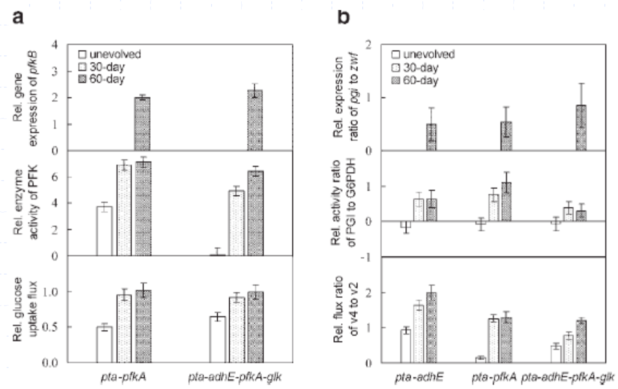


Figure 5. Averaged gene expression data, in vitro enzyme data and flux data relative to wild-type strain ($\log_2$ ratio, except for flux data) for unevolved and evolved lactate-producing strains. The expression state was assayed at Day 60 of evolution and compared to gene expression from wild-type anaerobic glucose culture. Enzyme activities and fluxes were assayed at three time points (unevolved, 30-day, 60-day) and compared to the data from the wild-type culture. **a**: Relative gene expression for the *pfkB* gene (**upper panel**), relative enzyme activity for the phosphofructokinase (**mid panel**), and relative glucose uptake flux (**lower panel**). **b**: Relative ratio of gene expression (*pgi* to *zwf*) (**upper panel**), relative ratio of enzyme activity (phosphoglucose isomerase to glucose 6-phosphate dehydrogenase) (**mid panel**), and relative ratio of flux (glycolytic flux to oxidative PP pathway flux) (**lower panel**).

35

Concluding Notes

Conclusions

- Metabolic net analysis can be reduced to convex analysis
- Genome scale analysis can be simplified through
  - By representing problems in terms of extreme pathways
  - By considering input-modes in terms of phenotypical phase plan
- Metabolic network can be controlled
  - By controlling input supply and using adaptive evolution of net
  - By knocking-out selective genes to optimize flux of interest
- Theoretical predictions provide good approximations
- Grand challenge: integrate with regulatory & signaling nets