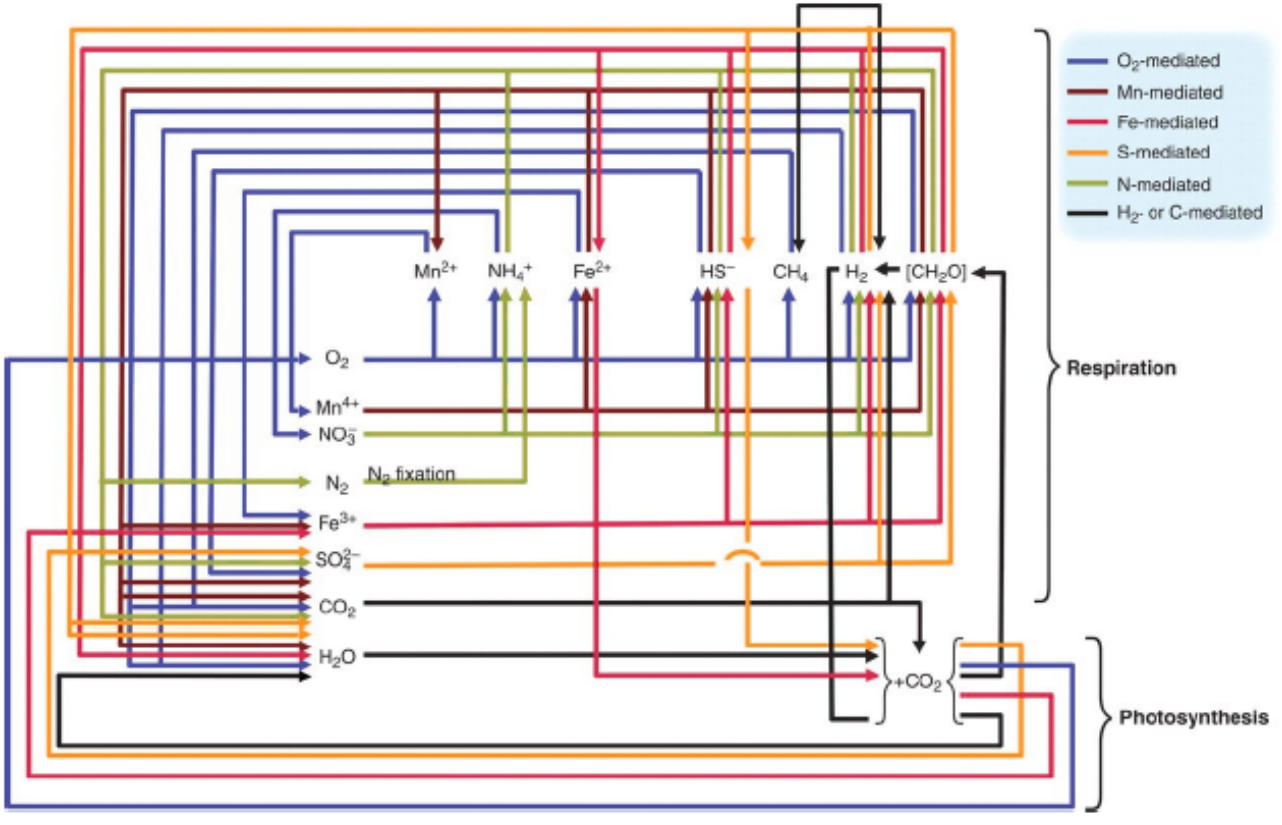
Metabolism: flux balans analysis

Course computational Biology 2024; Paulien Hogeweg;

Theoretical Biology and Bioinformatics Grp Utrecht University

(1) Life *is*..... energy/nutrient cycling



"The individual taxonomic units evolve and go extinct, yet the core machines survive surprisingly unperturbed."

PG Falkowski et al, Science 2008

Metabolic networks: Exploiting constraints

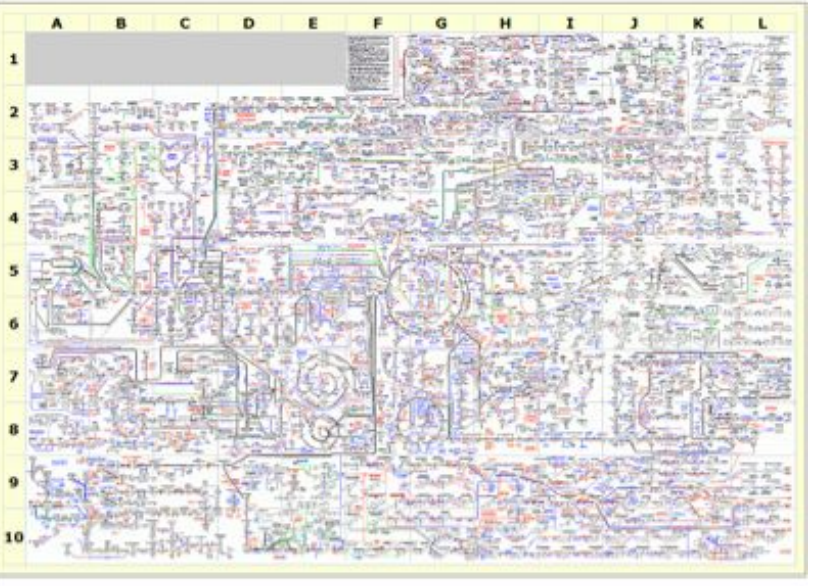
metabolic network are evolved

However metabolic networks many physical/chemical constraints
stochiometry, energetic constraints

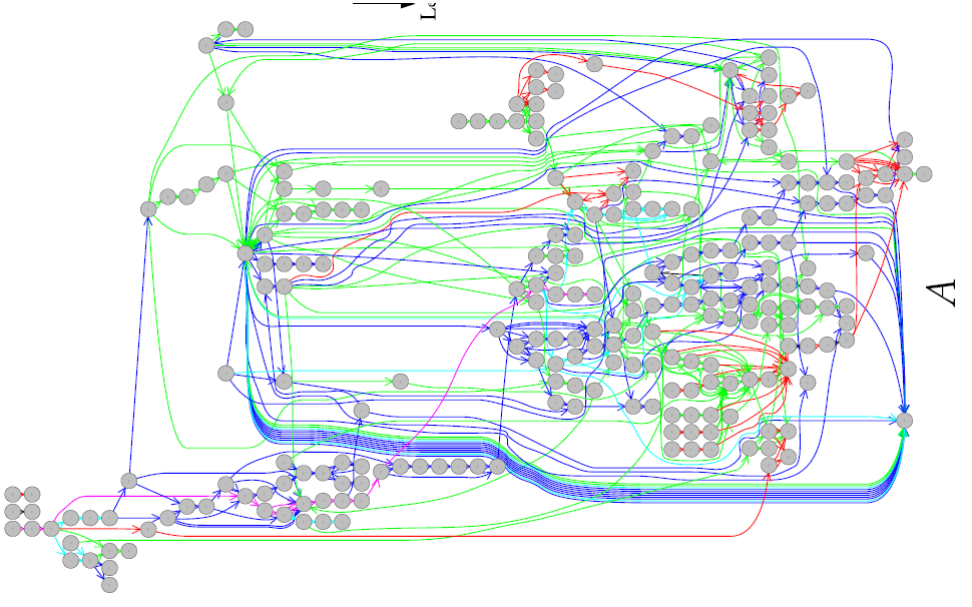
can/should be exploited

allows for model upscaling to complexity of present day organisms

stochiometric constraint, + equillibrium assumption allows calculation of (optimal) flux through large metabolic networks

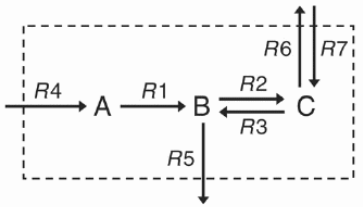


KEGG databse



metabolic network yeast

I. Reaction network formalism



Chemical reactions

Internal	Exchange
$R1$: −1 A → 1 B	$R4$: 1 A
$R2$: −1 B → 1 C	$R5$: −1 B
$R3$: −1 C → 1 B	$R6$: −1 C
	$R7$: 1 C

S =

	*R*1	*R*2	*R*3	*R*4	*R*5	*R*6	*R*7
A	−1	0	0	1	0	0	0
B	1	−1	1	0	−1	0	0
C	0	1	−1	0	0	−1	1

II. FBA formulation

Dynamic mass balance

$$\frac{dC}{dt} = \mathbf{S}\boldsymbol{v}$$

C : Concentration
t : Time
$\mathbf{S}$: Stoichiometric matrix
$\boldsymbol{v}$: Flux vector

Steady-state assumption

$$\mathbf{S}\boldsymbol{v} = \mathbf{0}$$

LP formulation

Objective: max $Z = v_5$

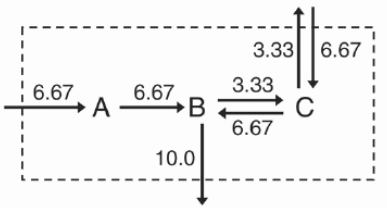
Constraints:

$$\begin{array}{c} \\ \text{A} \\ \text{B} \\ \text{C} \end{array} \begin{array}{c} \begin{array}{ccccccc} R1 & R2 & R3 & R4 & R5 & R6 & R7 \end{array} \\ \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 1 \end{bmatrix} \end{array} \begin{bmatrix} v_1 \\ \vdots \\ v_7 \end{bmatrix} = \mathbf{0} \qquad 0 \le v_1,\ldots,v_7 \le 10$$

III. Hypothetical flux distribution at steady-state

$Z = 10$

$\boldsymbol{v} = [6.67\ 3.33\ 6.67\ 6.6710.0\ 3.33\ 6.67]^{T}$



Flux balance analysis FBA
assume 'automatic' regulation such that flux in equillibrium and maximal growth

---

- FBA solution non-unique:
  use secondary optimization, eq minimal total flux
- stochiometric matrix (nA $->$ mB)
- reactions coupled to enzymes
- set maximum flux (when enzymes are present) (OR, AND reactions))
- however actual flux not proportional to amount of enzymes

examples of flux balance analysis

- How do fluxes (growth) change with change of environment (=input-flux)
- How do fluxes (growth) change with knock-outs?
- How do fluxes (growth) change after endosymbiosis? (cf Pal Papp Nature 2006)
- reconstruction of Ecosystem wide metabolome (cf Bas Dutilh)
- ....
- **How do genomes reduce after whole genome duplication? (cf van Hoek and Hogeweg 2009)**

automatic reconstruction of metabolic networks from annotated genomes cf Henry CS1, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. 2010

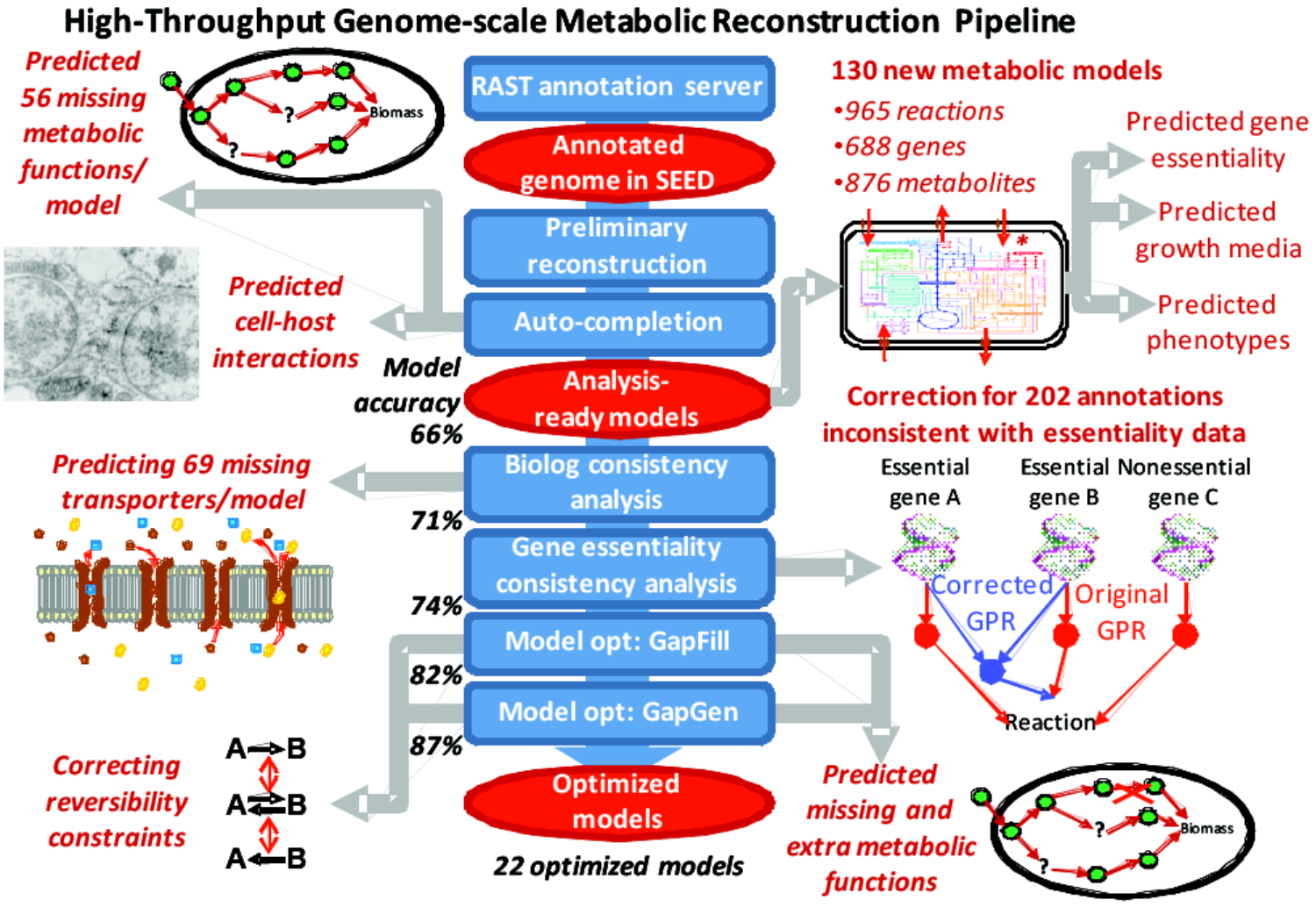
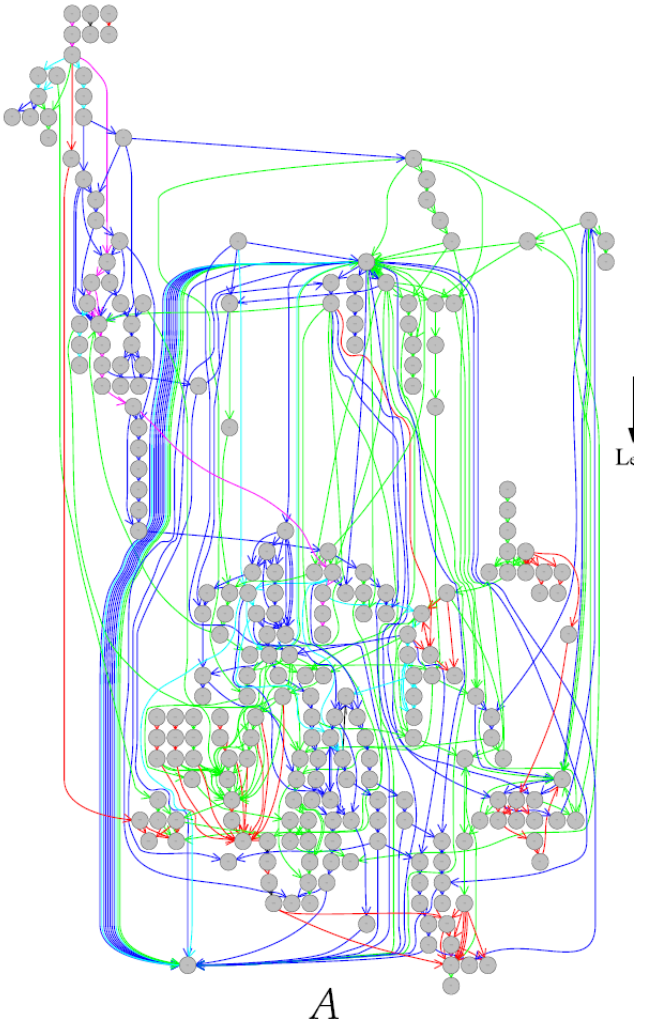


Figure S1: Overview of results and discoveries arising from the Model SEED pipeline.

Using FBA to reconstruct evolution of metabolic network of yeast after WGD

to cope with
genome-size networks:

exploit constraints
and use shortcut:
optimal equillibrium flux



Yeast metab. network

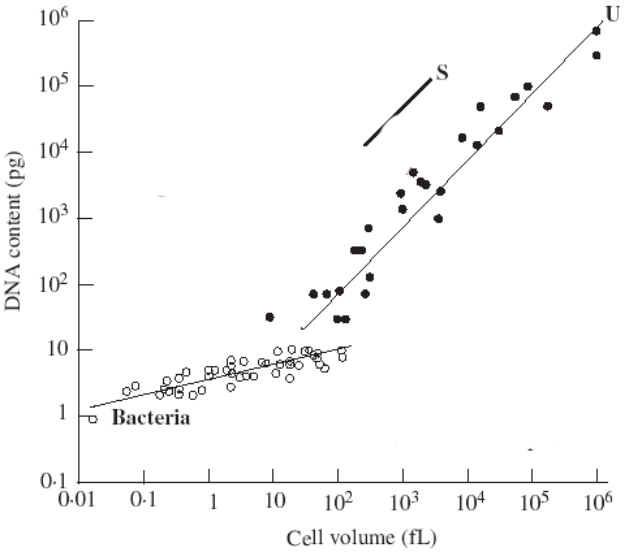
evolution of metabolic flux after WGD
FBA assumptions

- WGD $->$ volume increase (decrease surface/volume ratio)
  volume = depends on genome size
- flux of metabolic reaction depends on gene expression,
  dosis effect: gene copy number
- max flux through each reaction preset to maximum needed for optimal growth in sampled set of realizations of 10 environment types
- enzymes have multiple functions

- reactions need multiple enzymes
  take into account OR, AND (AND/OR) relations
- flux transport reactions: depends on gene expression AND surface/volume ratio
- after gene deletion maxflux reduced accordingly

WGD, cell size and fluxes

cell size scales with amount of DNA
Cavallier Smith (e.g. 2005)
In Yeast diploide cells are:
V=1.89 * haploid cells
surface: 1.56 * haploid cells
$V = N^{.9}$; $A = V^{.7}$
where N number of genes



MaxFlux change as function of area change(α), volume change (β) and gene dosage change γ)

external flux

$$F_{\max}(i) = F_{\max,0}(i)\frac{\alpha\gamma(i)}{\beta}\frac{1+x(i)}{\gamma(i)x(i)+\alpha}$$

internal flux

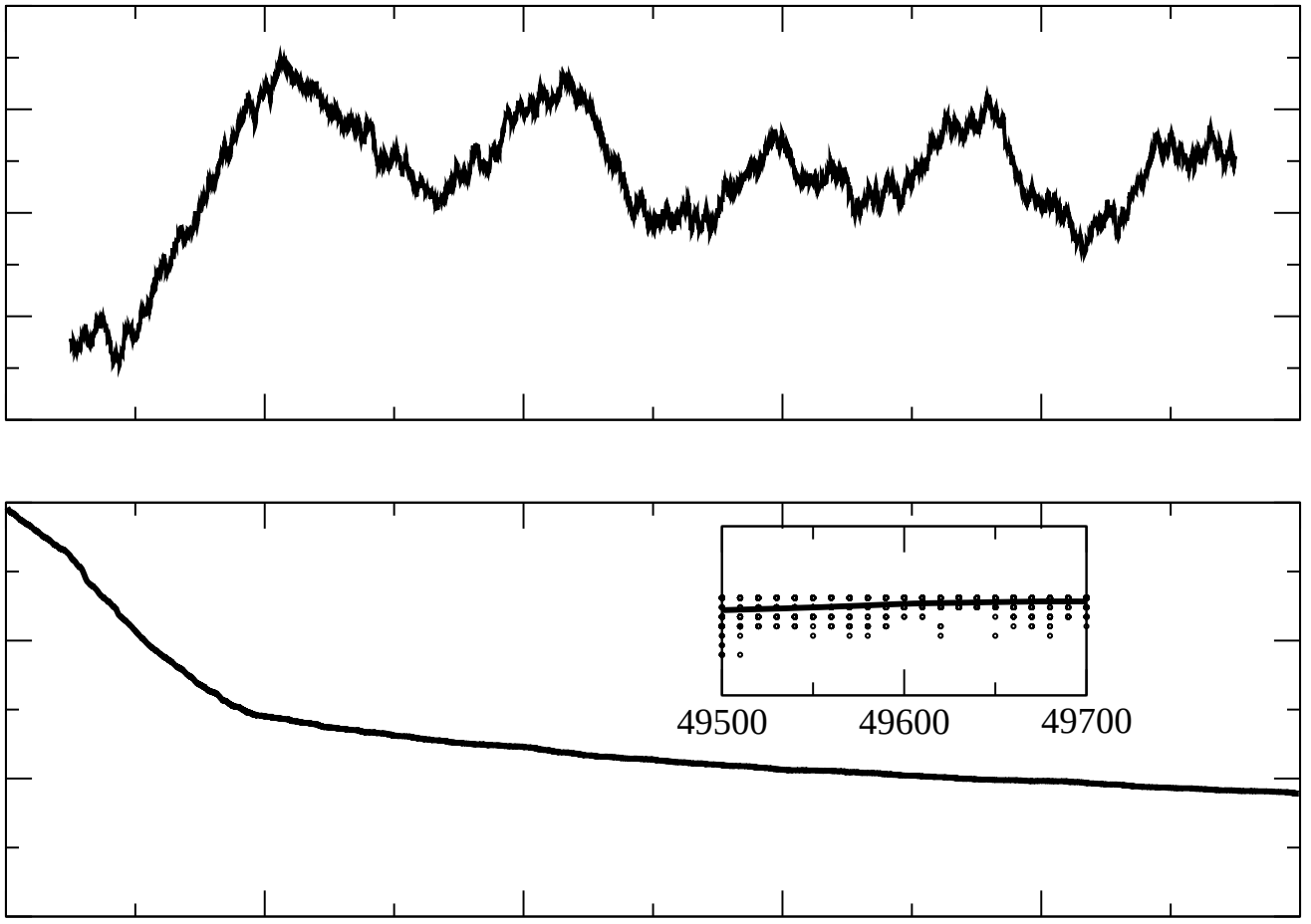
$$F_{\max}(i) = F_{\max,0}(i)\frac{\gamma(i)}{\beta}$$

evolution of metabolic flux after WGD
evolutionary protocol

- 9 types of environments (available nutrients).
  realized in different concentrations
- per generation 1 environmemnt seen
- pop size 100: flux dependent replication
  death: nogrowth + random
- after wgd: only deletions
  or duplication + deletion (max 2 copies)
- no fitness advantage for genome schrinkage smaller than intial volume

evolution of metabolic flux after WGD

evolutionary dynamics: growth rate and genome reduction

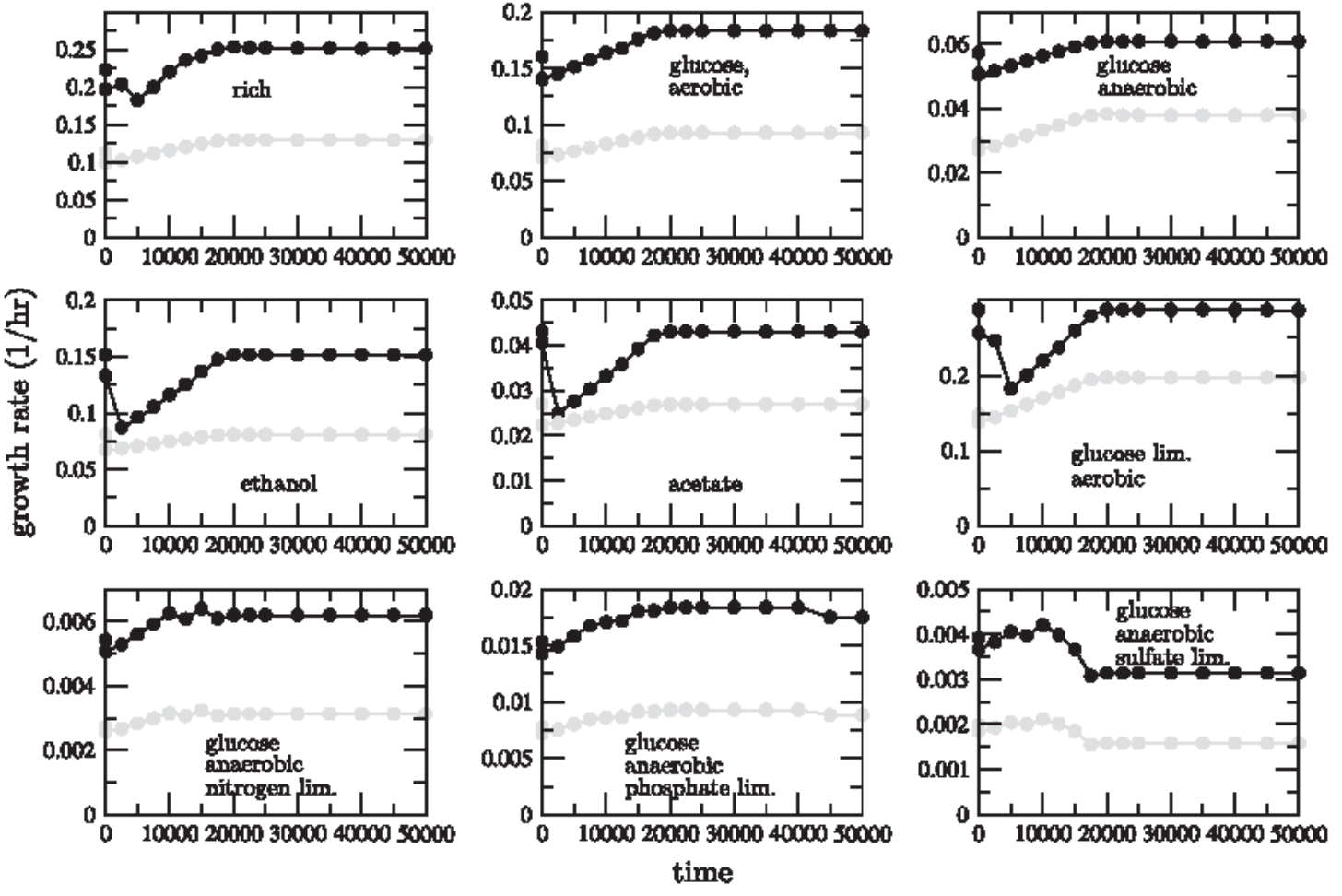


evolution of metabolic flux after WGD

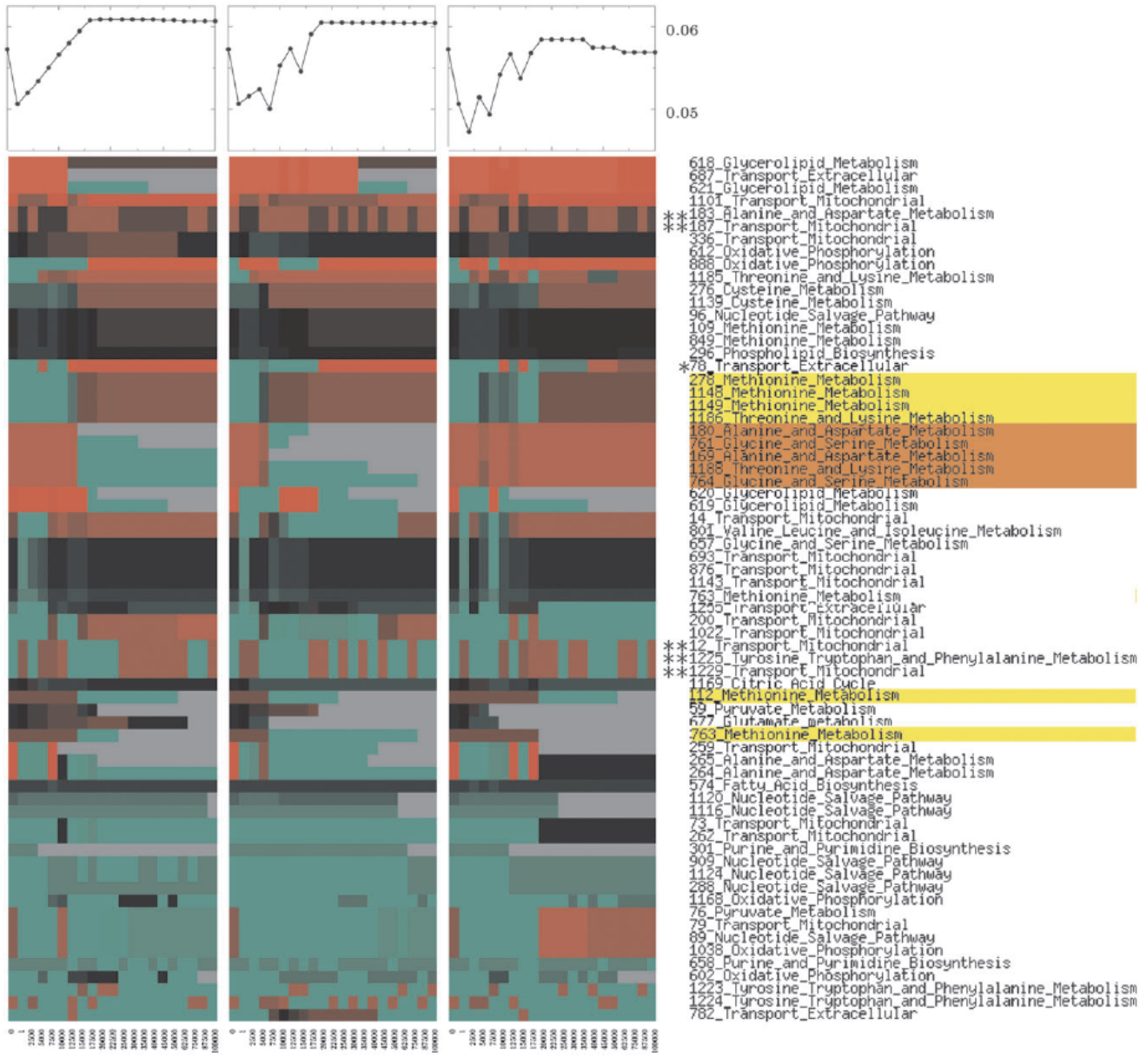
flux in the various environments

(max and mean concentration)

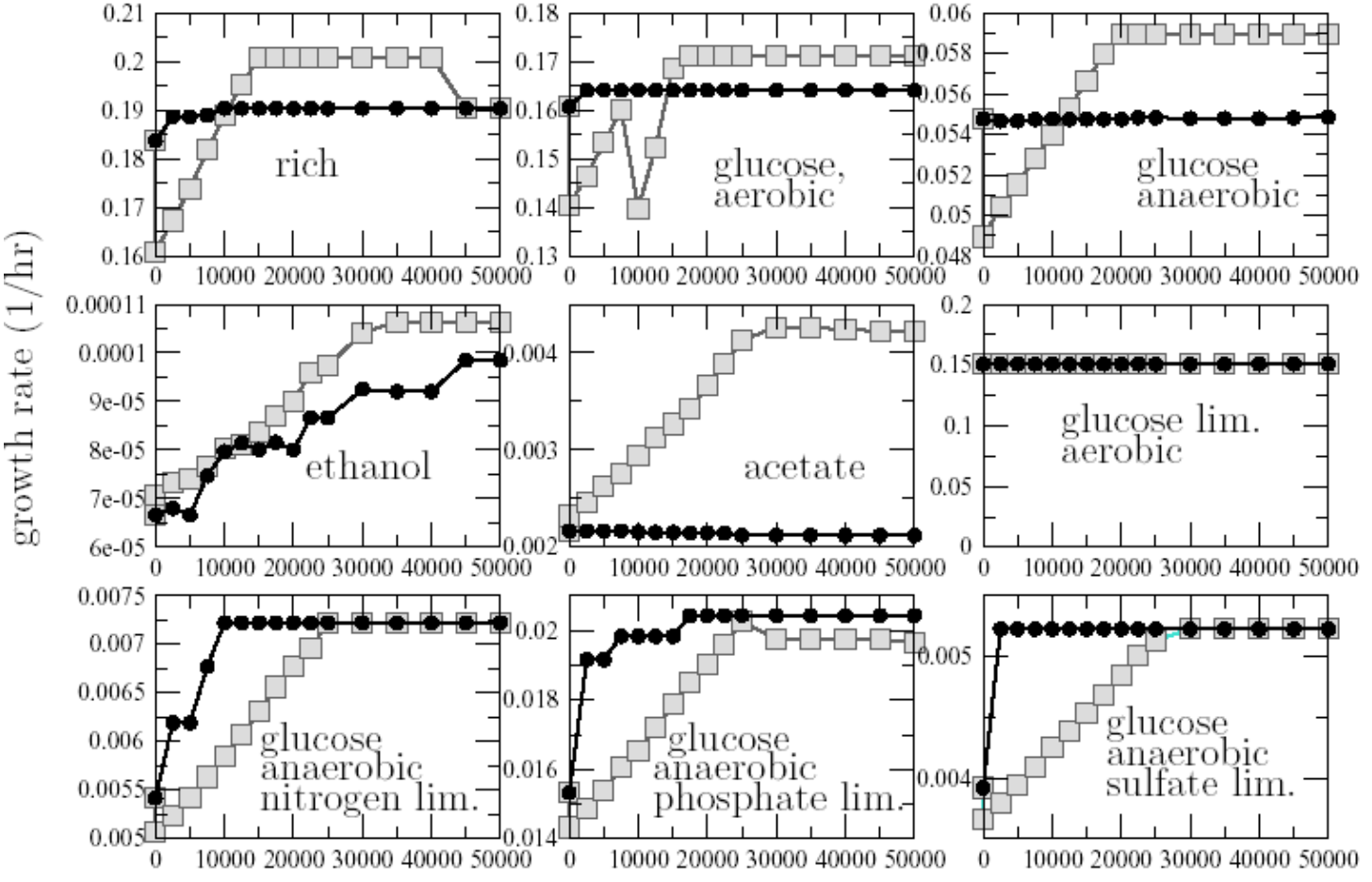
initial decrease – how/when does it happen in evolution


growth rate (1/hr)
time
rich
glucose, aerobic
glucose anaerobic
ethanol
acetate
glucose lim. aerobic
glucose anaerobic nitrogen lim.
glucose anaerobic phosphate lim.
glucose anaerobic sulfate lim.
0 10000 20000 30000 40000 50000

Genome schrinkage after whole genome duplication dynamics of use of pathways in anaerobic glucose environment (env 3)


0.06
0.05
618_Glycerolipid_Metabolism
687_Transport_Extracellular
621_Glycerolipid_Metabolism
1101_Transport_Mitochondrial
**183_Alanine_and_Aspartate_Metabolism
**187_Transport_Mitochondrial
336_Transport_Mitochondrial
612_Oxidative_Phosphorylation
888_Oxidative_Phosphorylation
1185_Threonine_and_Lysine_Metabolism
276_Cysteine_Metabolism
1139_Cysteine_Metabolism
96_Nucleotide_Salvage_Pathway
109_Methionine_Metabolism
849_Methionine_Metabolism
296_Phospholipid_Biosynthesis
*78_Transport_Extracellular
278_Methionine_Metabolism
1148_Methionine_Metabolism
1149_Methionine_Metabolism
1186_Threonine_and_Lysine_Metabolism
180_Alanine_and_Aspartate_Metabolism
761_Glycine_and_Serine_Metabolism
165_Alanine_and_Aspartate_Metabolism
1188_Threonine_and_Lysine_Metabolism
764_Glycine_and_Serine_Metabolism
620_Glycerolipid_Metabolism
619_Glycerolipid_Metabolism
14_Transport_Mitochondrial
801_Valine_Leucine_and_Isoleucine_Metabolism
657_Glycine_and_Serine_Metabolism
693_Transport_Mitochondrial
876_Transport_Mitochondrial
1143_Transport_Mitochondrial
763_Methionine_Metabolism
1255_Transport_Extracellular
200_Transport_Mitochondrial
1022_Transport_Mitochondrial
**12_Transport_Mitochondrial
**1225_Tyrosine_Tryptophan_and_Phenylalanine_Metabolism
**1229_Transport_Mitochondrial
1169_Citric_Acid_Cycle
112_Methionine_Metabolism
59_Pyruvate_Metabolism
677_Glutamate_Metabolism
763_Methionine_Metabolism
259_Transport_Mitochondrial
265_Alanine_and_Aspartate_Metabolism
264_Alanine_and_Aspartate_Metabolism
574_Fatty_Acid_Biosynthesis
1120_Nucleotide_Salvage_Pathway
1116_Nucleotide_Salvage_Pathway
73_Transport_Mitochondrial
262_Transport_Mitochondrial
301_Purine_and_Pyrimidine_Biosynthesis
909_Nucleotide_Salvage_Pathway
1124_Nucleotide_Salvage_Pathway
288_Nucleotide_Salvage_Pathway
1168_Oxidative_Phosphorylation
76_Pyruvate_Metabolism
79_Transport_Mitochondrial
89_Nucleotide_Salvage_Pathway
1038_Oxidative_Phosphorylation
658_Purine_and_Pyrimidine_Biosynthesis
602_Oxidative_Phosphorylation
1223_Tyrosine_Tryptophan_and_Phenylalanine_Metabolism
1224_Tyrosine_Tryptophan_and_Phenylalanine_Metabolism
782_Transport_Extracellular

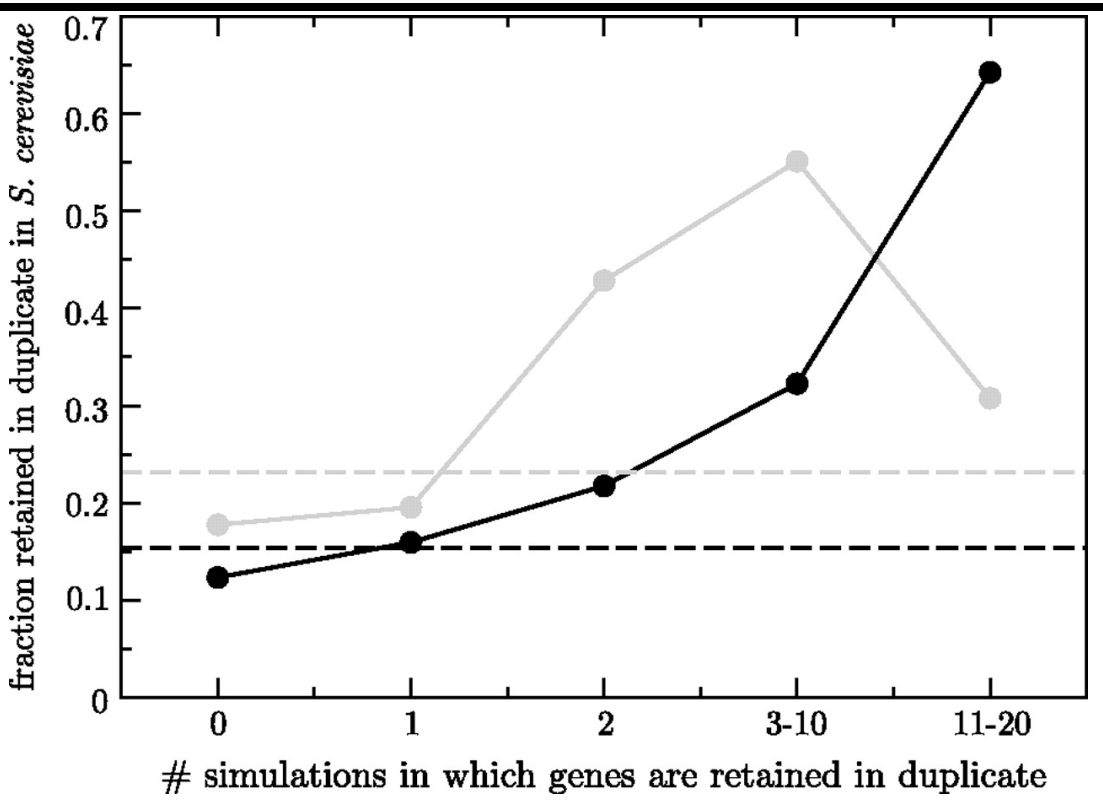
Only in "new" environment - nodirect disadvantage of WGD



BUT single INDELS initially better Exept in ethanol env

WGD mostly better end result than single INDELS

WGD: Simulated evolution and/vs yeast duplication of yeast vs duplication of ancestor of yeast (+hgt)



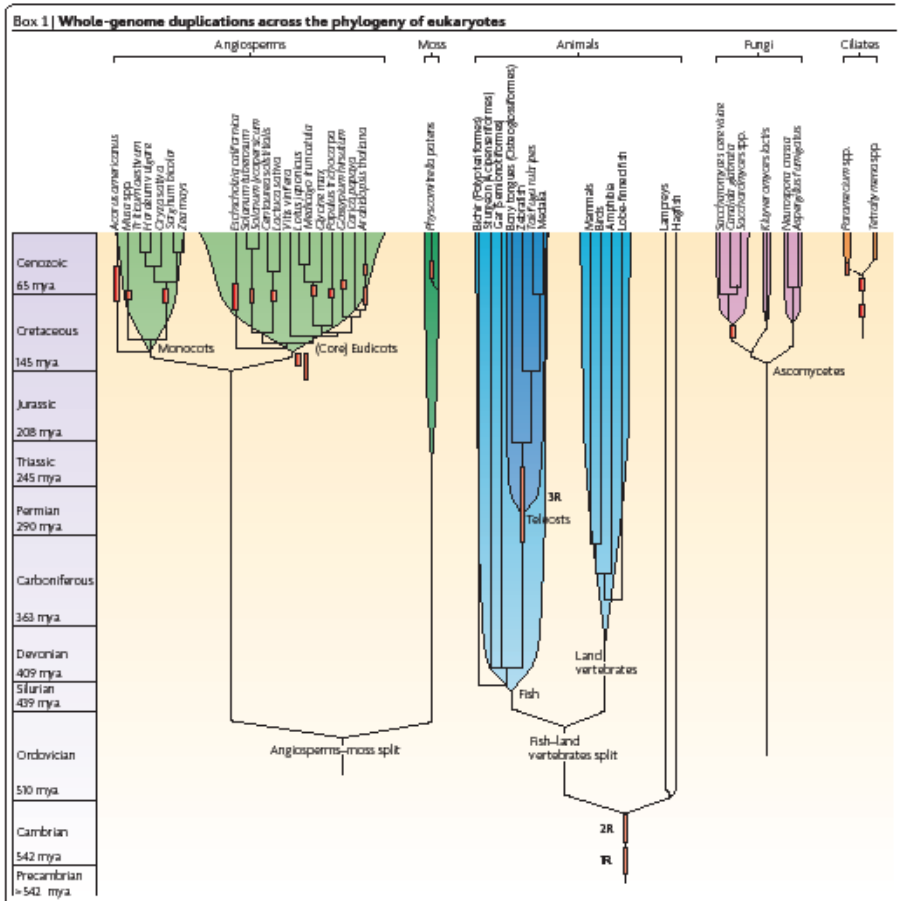
Preferential retained genes: Glycolysis pathway and Transporters

Evolution predictable!

conclusions
supervised vs non-supervised modeling of WGD in Yeast

"Supervised" Conan & Wolfe (2007)	**"Non Supervised"** van Hoek & H. (2009)
find genes preferentially retained *glycolysis pathway* Model glycolysis pathway assuming dosis effect of duplicated genes demonstrate WGD can lead to increased glycolic flux	take known interactions metab. net + DNA-volume relation model evolution find preferentially retained genes *glycolysis & transport* WGD mostly disadvantaeous initially except in "new" environments seldom better than single INDELS evolutionary outcome "deterministic"
WGD enabled to exploit high glucose resource during emergence of angiosperms	WGD enabled to exploit high glucose resource during emergence of angiosperms
observed outcome of WGD	**expected outcome of WGD**

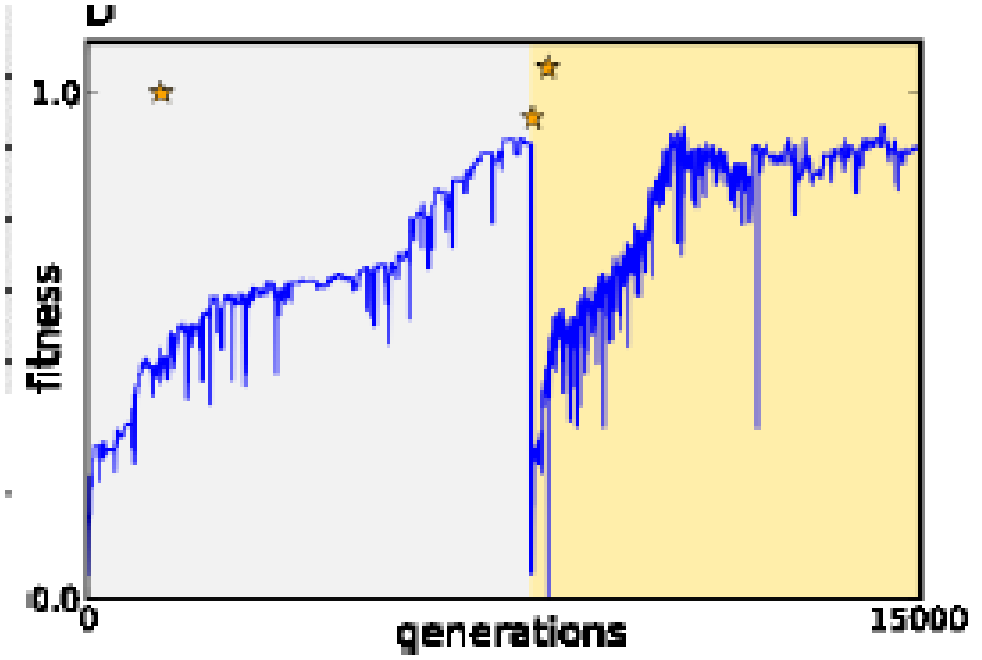
WGD observed in phylogeny at times of environmental shifts



van der Peer et al 2009, Nature genetic reviews

WGD observed in virtual cell model at times of environmental shifts

WGD ongoing mutation,
but only fixed in population EARLY in evolution
OR after SOME (severe?) environmental changes



and WGD leads to high fitness much later

Cuypers & Hogeweg 2014

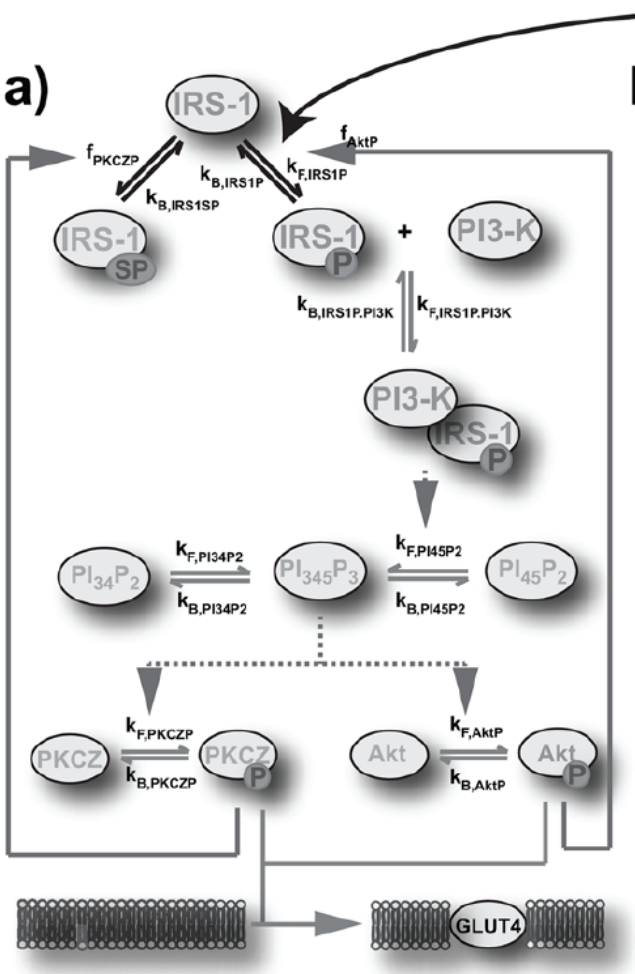
Neutral Paths, Causal Drift, Robust Signaling, and Complex Disease Andreas Wagner 2015 PLONE

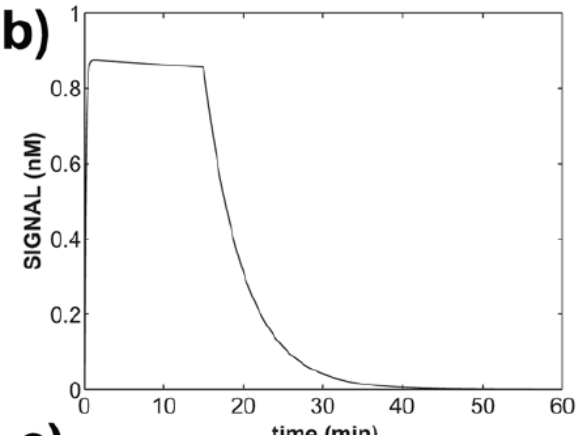
Explicit model of Insuline signaling pathway

Random sampling of 15 kinetic parameters $10^{-3} - 10^{3}$
and evolving populations by mutating these parameters

Generate many "healthy" and "sick" individuals (pathway instantiations)

Classifying behavior as "normal" $V = 0.076x10^{-4}$
or "deseased" $V = 0.33x10^{-4}$
(based on glucase uptake-curve in time) Determine sensitivity of parameters in different populations and during evolution.





c)

The model's first three equations are

$$\frac{d[IRS1]}{dt} = -\mathbf{k}_{\mathbf{F,IRS1P}}[IRS1][SIGNAL] + \mathbf{f}_{\mathbf{AktP}}\mathbf{k}_{\mathbf{B,IRS1P}}[Akt][IRS1P] + \mathbf{k}_{\mathbf{B,IRS1SP}}[IRS1SP] - \mathbf{f}_{\mathbf{PKCZP}}[PKCZP][IRS1]$$

$$\frac{d[IRS1P]}{dt} = \mathbf{k}_{\mathbf{F,IRS1P}}[IRS1][SIGNAL] - \mathbf{f}_{\mathbf{AktP}}\mathbf{k}_{\mathbf{B,IRS1P}}[Akt][IRS1P] + \mathbf{k}_{\mathbf{B,IRS1P.PI3K}}[IRS1P.PI3K] - \mathbf{k}_{\mathbf{F,IRS1P.PI3K}}[PI3K][IRS1P]$$

$$\frac{d[IRS1SP]}{dt} = \mathbf{f}_{\mathbf{PKCZP}}[PKCZP][IRS1] - \mathbf{k}_{\mathbf{B,IRS1SP}}[IRS1SP]$$

Number of parameter sets

Fig 1. Insulin signaling model, input and output. a) Molecular interactions in the signaling pathway modeled here. Briefly, extracellular insulin leads to phosphorylation of the insulin receptor, which promotes the phosphorylation of *IRS*1 to yield *IRS*1*P*. The latter molecule associates with *PI3K* in a complex that triggers production of the second messenger $PI_{345}P_3$, which activates the protein kinases *Akt* and *PKCZ*. These kinases then promote the translocatior of the glucose transporter *GLUT*4 to the membrane, where it helps import glucose into the cell. Mass-action parameters that determine the rates of the respective reactions are indicated by a '**k**' followed by a subscript. Activated *PKCZ* and *Akt* exert feedback on the production of two different phosphorylated forms of *IRS*1 (*IRS*1*SP* and *IRS*1*P*). The strength of this feedback is encapsulated by parameters f_{PKCZP} and f_{AktP}, respectively. See Methods for details. **b)**

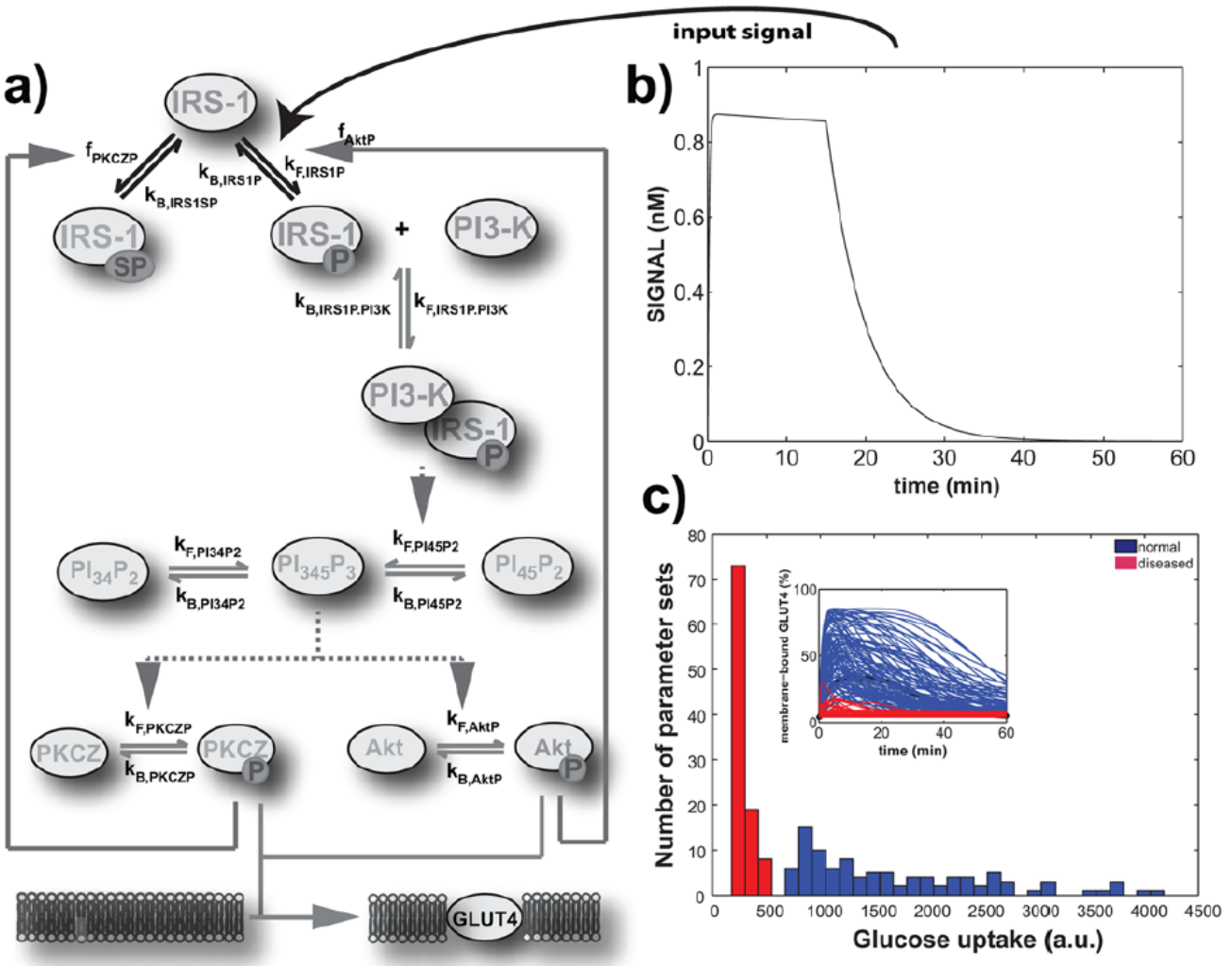
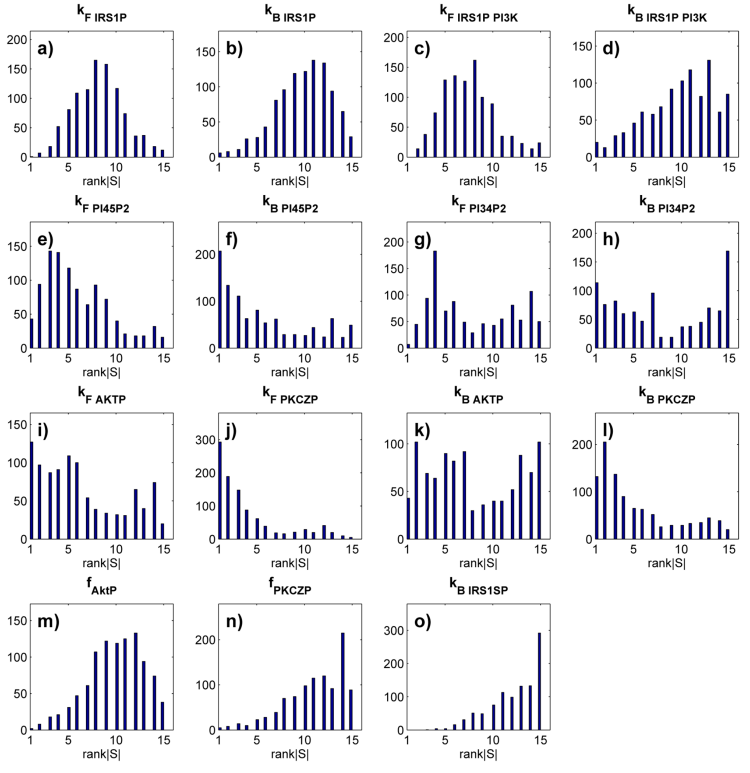
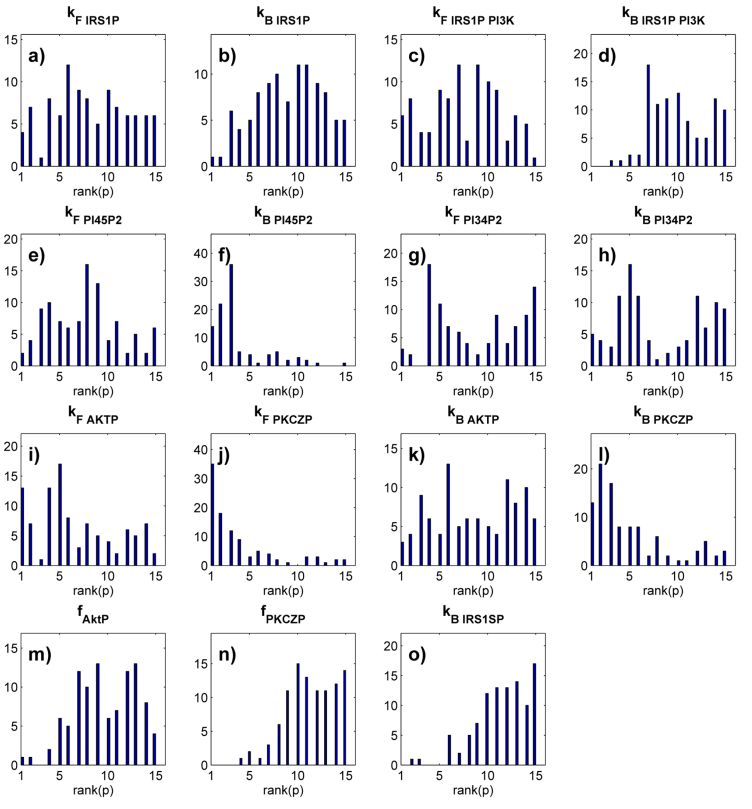


Fig 1. Insulin signaling model, input and output. a) Molecular interactions in the signaling pathway modeled here. Briefly, extracellular insulin leads to phosphorylation of the insulin receptor, which promotes the phosphorylation of *IRS*1 to yield *IRS*1*P*. The latter molecule associates with *PI3K* in a complex that triggers production of the second messenger $PI_{345}P_3$, which activates the protein kinases *Akt* and *PKCZ*. These kinases then promote the translocation of the glucose transporter *GLUT*4 to the membrane, where it helps import glucose into the cell. Mass-action parameters that determine the rates of the respective reactions are indicated by a '**k**' followed by a subscript. Activated *PKCZ* and *Akt* exert feedback on the production of two different phosphorylated forms of *IRS*1 (*IRS*1*SP* and *IRS*1*P*). The strength of this feedback is encapsulated by parameters f_{PKCZP} and f_{AktP}, respectively. See Methods for details. **b)**

Importance of parameter varies greatly depemdemt om parameter set (= genetic background)

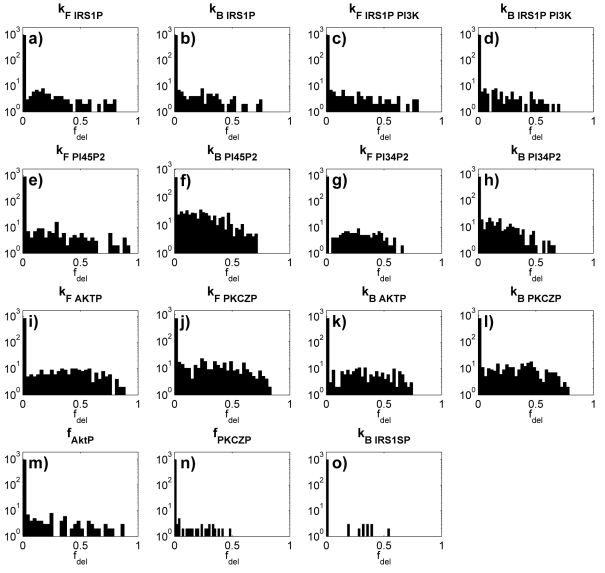


sensitivity of parameter

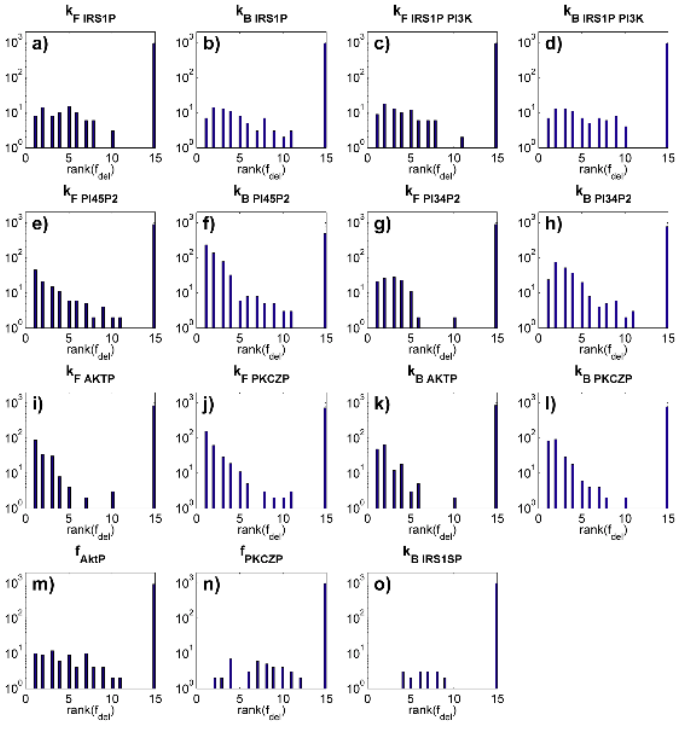


discrimination healthy/sick

very high neutrality of 'gene'deletions but very different in different parameter sets (instantiations).



neutrality of deletions



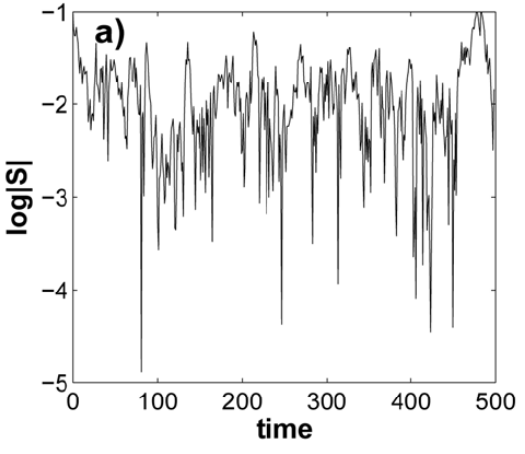
likelyhood of deleterious effects

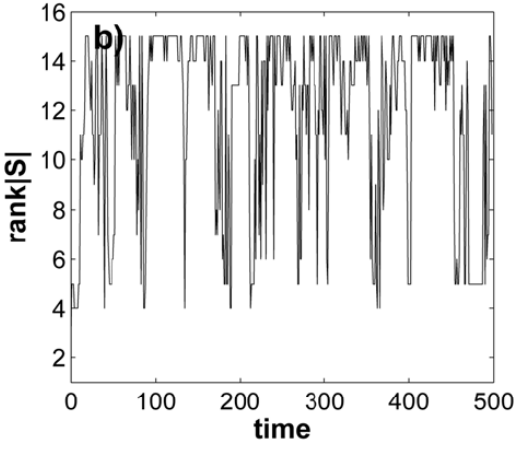
Rapid "Causal drift"

raplid change of sensitivity to parameter changes (mutations) due to neutral drift

"genetic backgrour

"cause of desease"





cf GWAS studies
50% "explained"

Mouse models