OPTIMALITY AND OPTIMIZING DURING EXPERIMENTAL EVOLUTION

CHRISTOPHER MARX

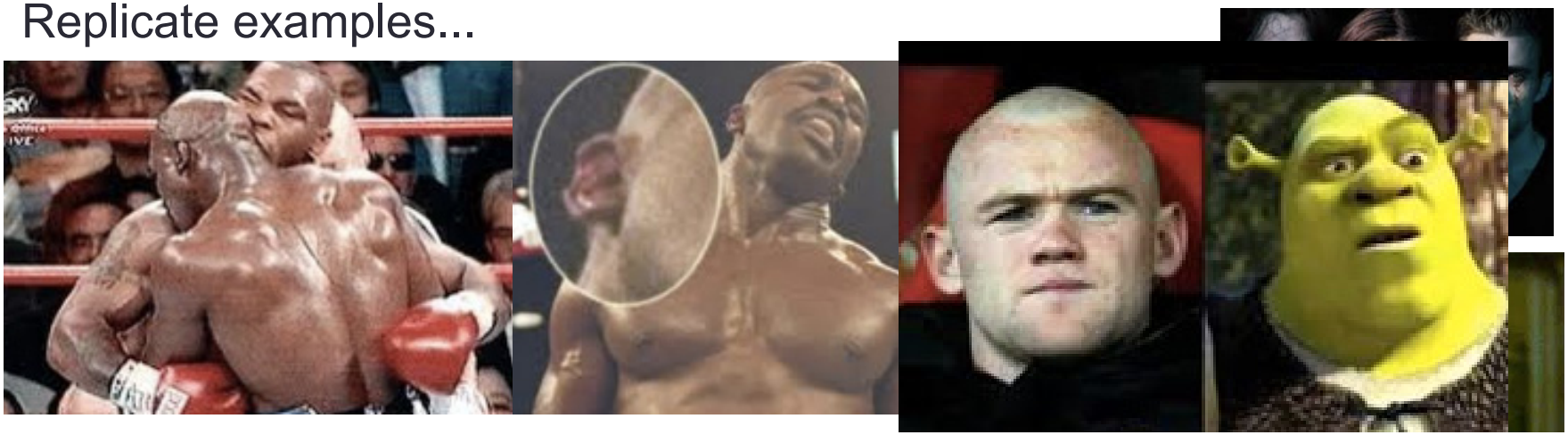
BIOLOGICAL SCIENCES
INSTITUTE FOR BIOINFORMATICS & EVOLUTIONARY STUDIES
UNIVERSITY OF IDAHO
CMARX@UIDAHO.EDU

Cases of known adaptation

Ancestral, primitive behavior

Refined, optimized behavior

Replicate examples...



Cases of known adaptation

Ancestral, primitive behavior

Refined, optimized behavior

Other transformations possible

Cellular economy at many levels

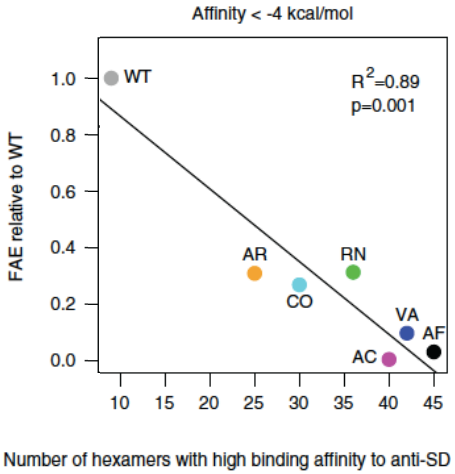
- Selection upon synonymous sites
- Selection to reduce the accessory genome
- Selection can occur with surprisingly little substrate specialization

- Optimizing economy of central metabolism
- Optimizing the use of foreign metabolic pathways
  - Large-scale transcriptome changes scaling with growth rate
  - Optimizing expression of enzymes of a foreign formaldehyde oxidation pathway

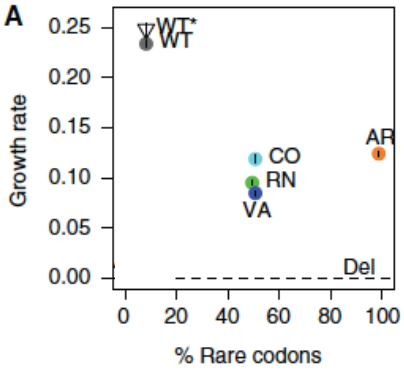
Selection upon synonymous sites

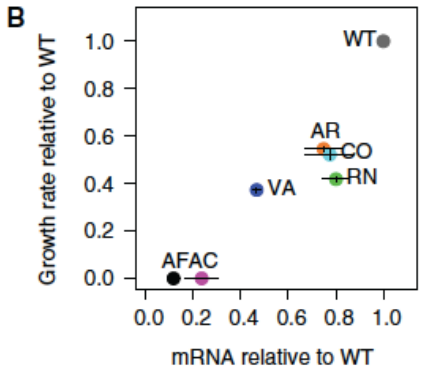
Table 1. Experimental Strains with Synonymous *fae* Alleles.

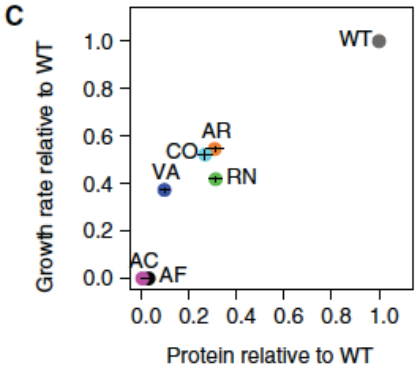
Strain	Codon Distribution: % Frequent	% Rare	Residues with Rare Codons	# Synonymous Mutations	5′ sequence: M A K I T K V Q V G E A L V
WT	71.9	8	Wild type	–	ATGGCAAAAATCACCAAGGTTCAGGTCGGCGAGGCCCTCGTC
AF	100	0	None	46	ATG..C..G.....G.....G.....G..G
AC	89	11	Active	52	ATG..C..G.....G.....G.....T.G..G
RN	49.4	49.4	Random	92	ATG.....A.....G..A..A.....AT.G..G
VA	49.4	50.6	Variable	102	ATG..C..G.....G.....G..A..A...T.G..A
CO	49.4	50.6	Conserved	94	ATG.....A..A..A..A..A.....A..G..G
AR	0	98.8	All	150	ATG.....A..A..A..A..A..A..A..AT.G..A

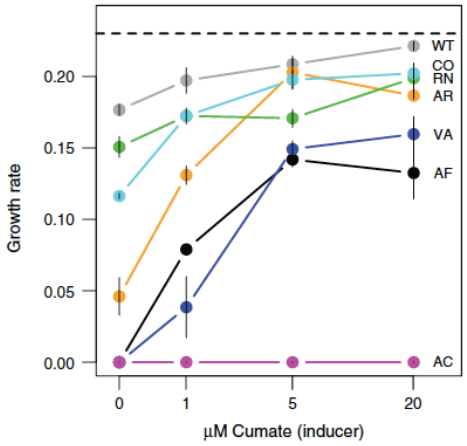


- Synthesized synonymous versions of key enzyme
- Main cause: anti-SDs









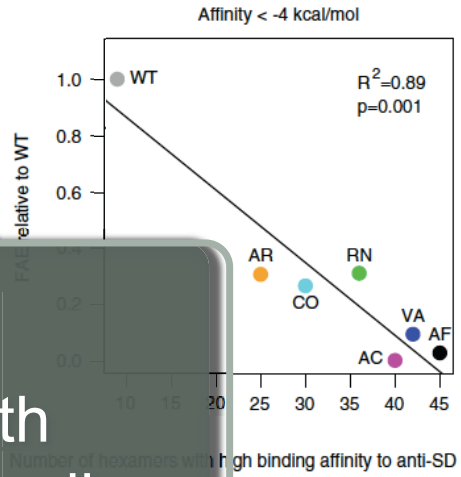
- Massive fitness effects; due to low expression

(Agashe et al., 2013. *Molecular Biology & Evolution*; Agashe et al., in prep)

Selection upon synonymous sites

Table 1. Experimental Strains with Synonymous *fae* Alleles.

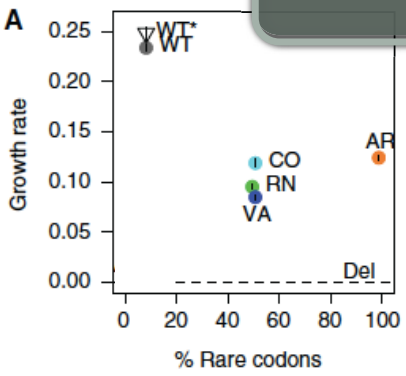
Strain	Codon Distribution: % Frequent	% Rare	Residues with Rare Codons	# Synonymous Mutations	5′ sequence: M A K I T K V Q V G E A L V
WT	71.9	8	Wild type	–	ATGGCAAAAATCACCAAGGTTCAGGTCGGCGAGGCCCTCGTC
AF	100	0	None	46	ATG..C..G.....G.....G.....G..G
AC	89	11	Active	52	ATG..C..G.....G.....T.G..G
RN	49.4	49.4	Random	92	
VA	49.4				
CO	49.4		Conserved	94	
AR	0				

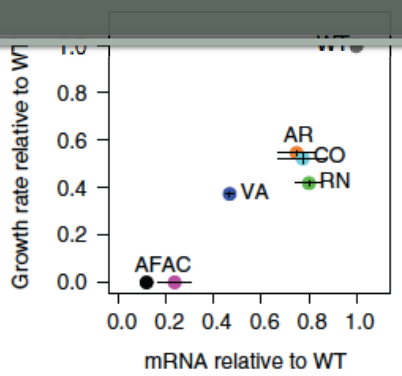


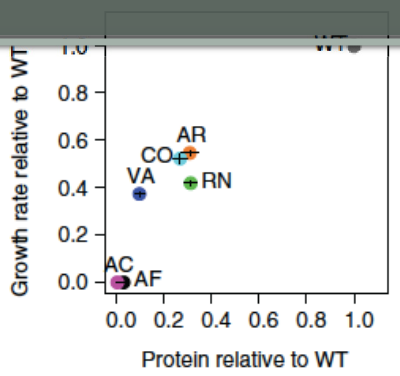
Subsequent adaptation:

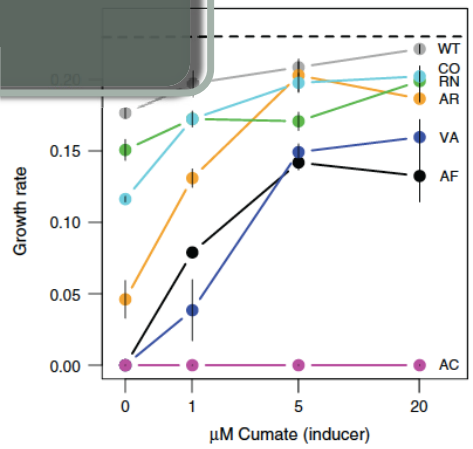
- Single mutations recover WT growth
- Sometimes promoter; sometimes coding
- 5’ end preventing anti-SDs

- Synthesiz[illegible] synonymous versions of key enzyme [illegible] se: anti-SDs





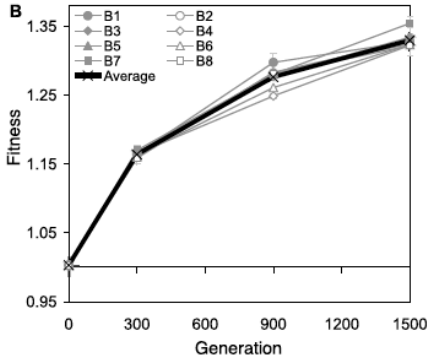


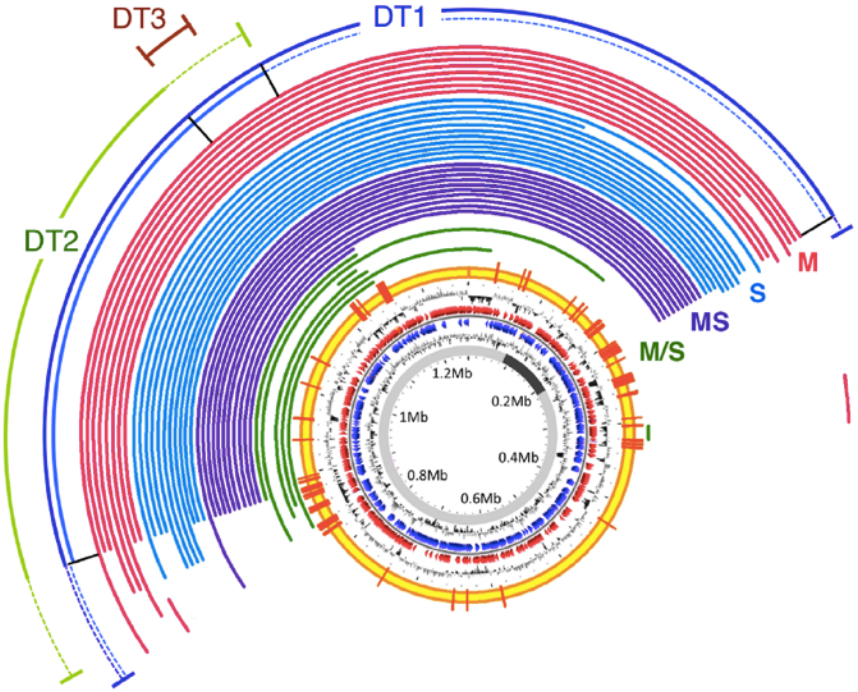


- Massive fitness effects; due to low expression

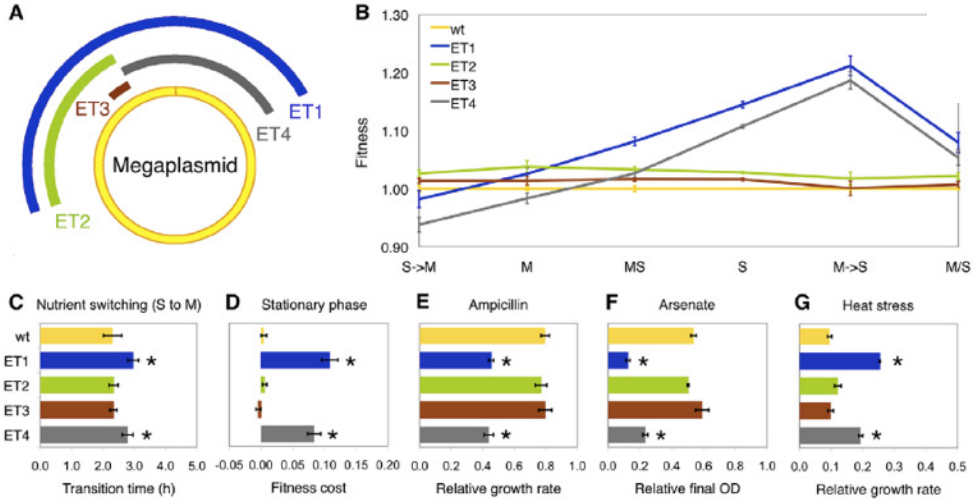
(Agashe et al., 2013. *Molecular Biology & Evolution*; Agashe et al., in prep)

Selection to remove up to 10% genome





- Parallel deletions of >600 kb of megaplasmid in 1500 gen. adaptation by *Methylobacterium*

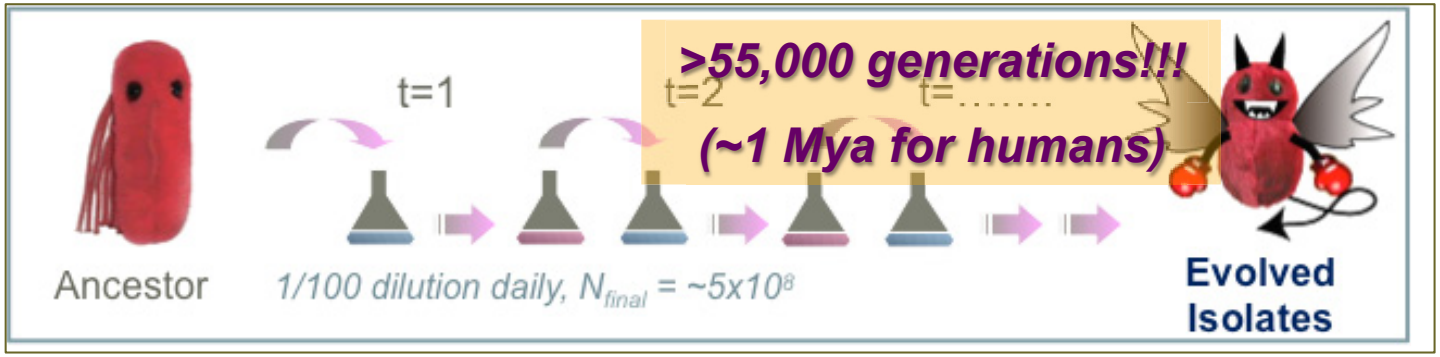


- Benefit due to genes removed, not just DNA
- Caused tradeoffs in other environments

(Lee et al., 2009. *Evolution*; Lee and Marx, 2012. *PLoS Genetics*)

Longest evolution: Lenski long-term lines

- On February 24th, 1988, Rich Lenski started 12 populations of *E. coli* B in minimal glucose medium

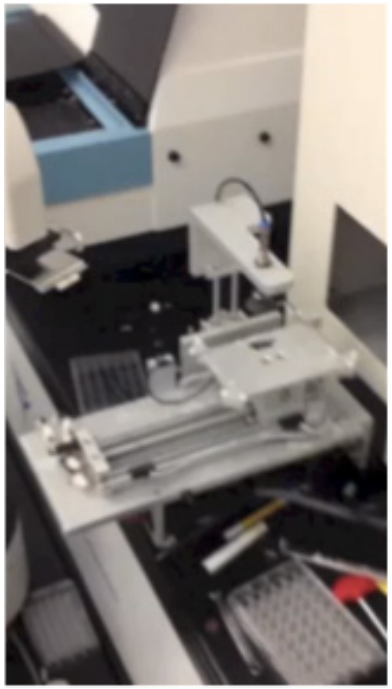


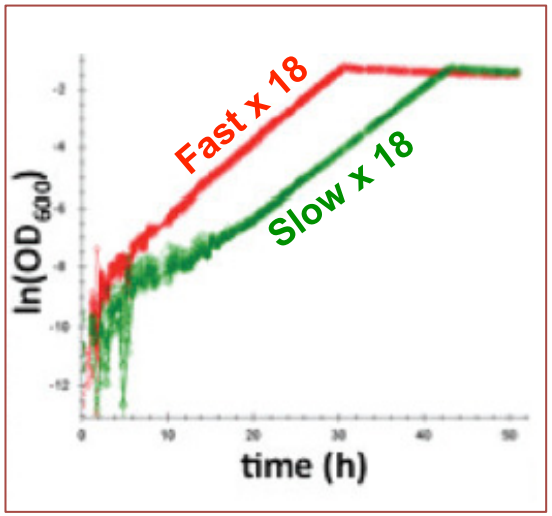
- USA popular culture in February, 1988:

#1 song: "Seasons Change" by Exposé

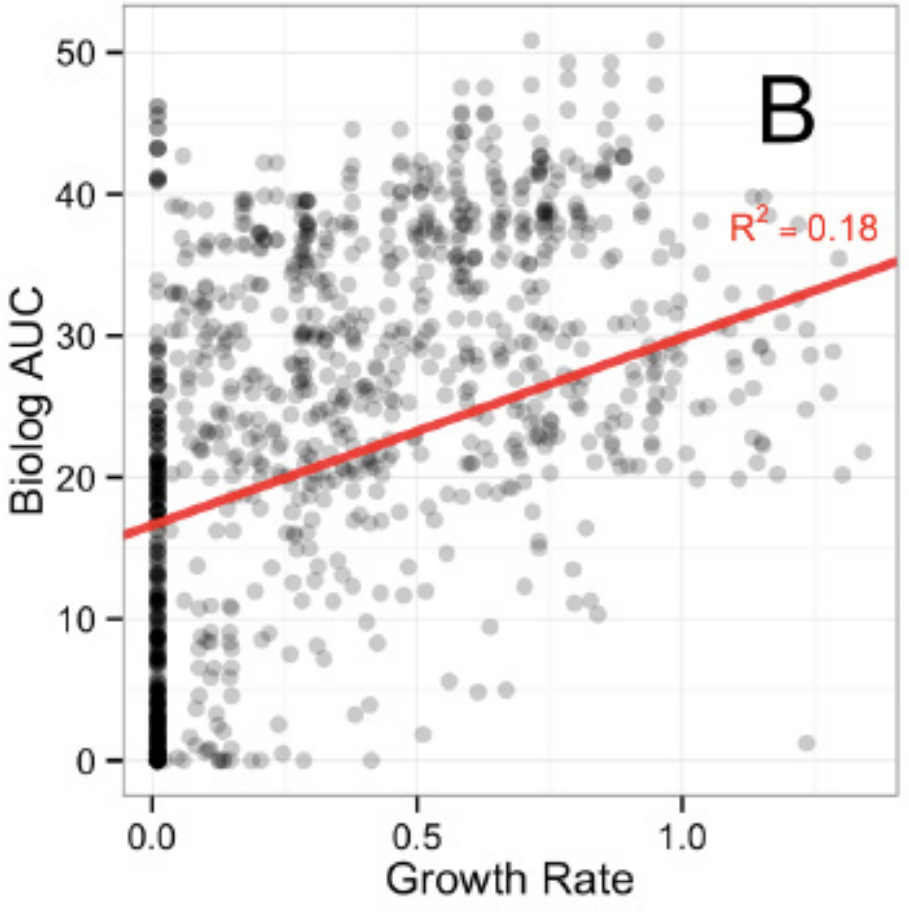
#1 album: Faith, George Michael

Little specialization over 50,000 gen.





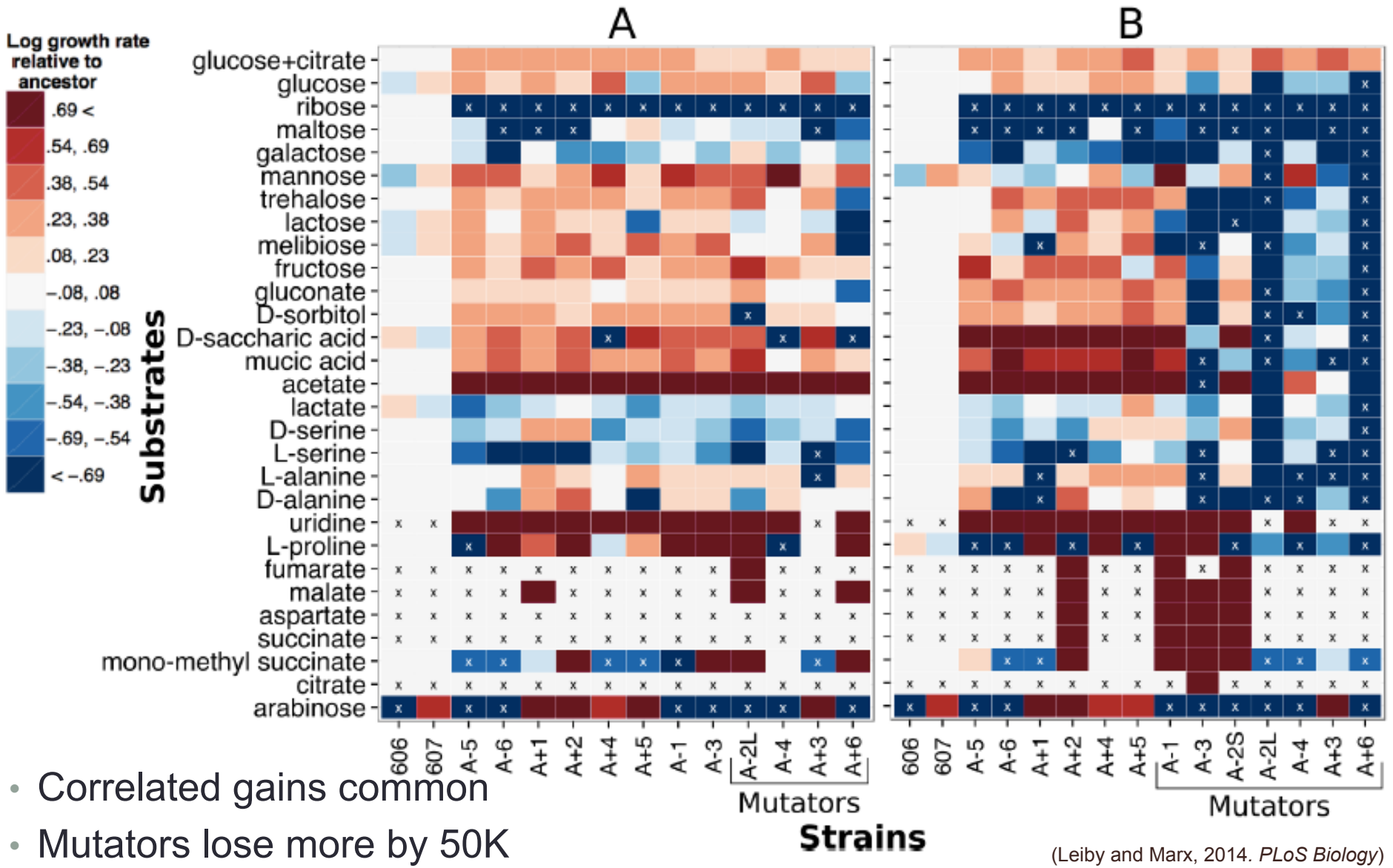
- Automated, robotic growth system



- Biolog ≠ growth

(Delaney et al., 2013. *J. Lab. Automation*; Delaney et al., 2013. *PLoS One*; Leiby and Marx, 2014. *PLoS Biology*)

Little specialization over 50,000 gen.

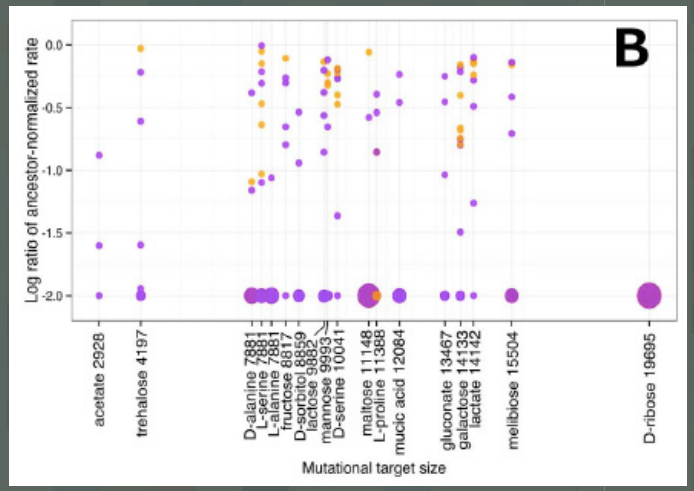


- Correlated gains common
- Mutators lose more by 50K

(Leiby and Marx, 2014. *PLoS Biology*)

Little specialization over 50,000 gen.

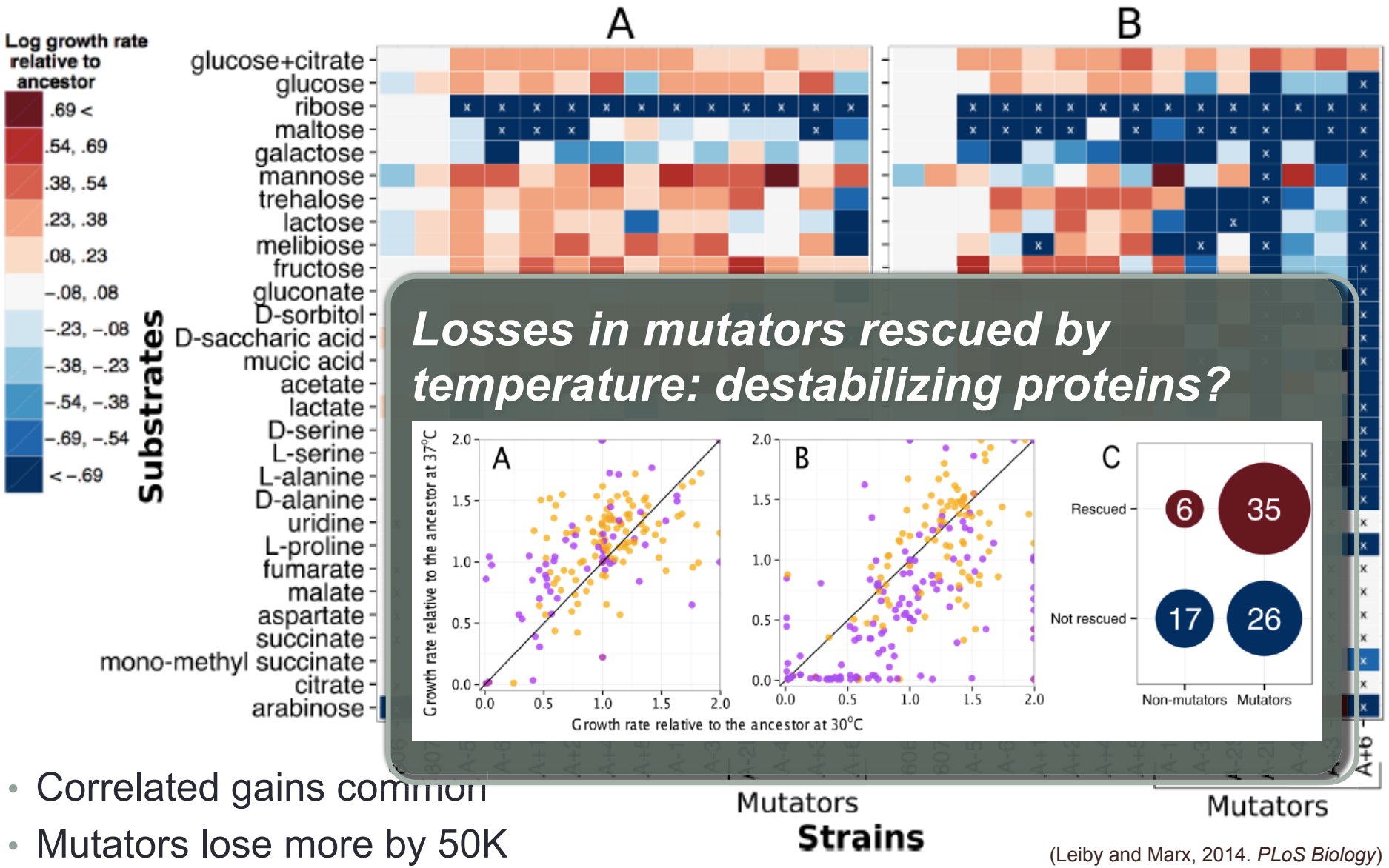
Substrates experiencing decline not correlated with similarity to glucose



- Correlated gains common
- Mutators lose more by 50K

(Leiby and Marx, 2014. *PLoS Biology*)

Little specialization over 50,000 gen.



- Correlated gains common
- Mutators lose more by 50K

(Leiby and Marx, 2014. *PLoS Biology*)

Metabolism: connect phenotype to fitness at two scales

Metabolic control (MCA)

"Changes to which parts would most optimize fitness?"

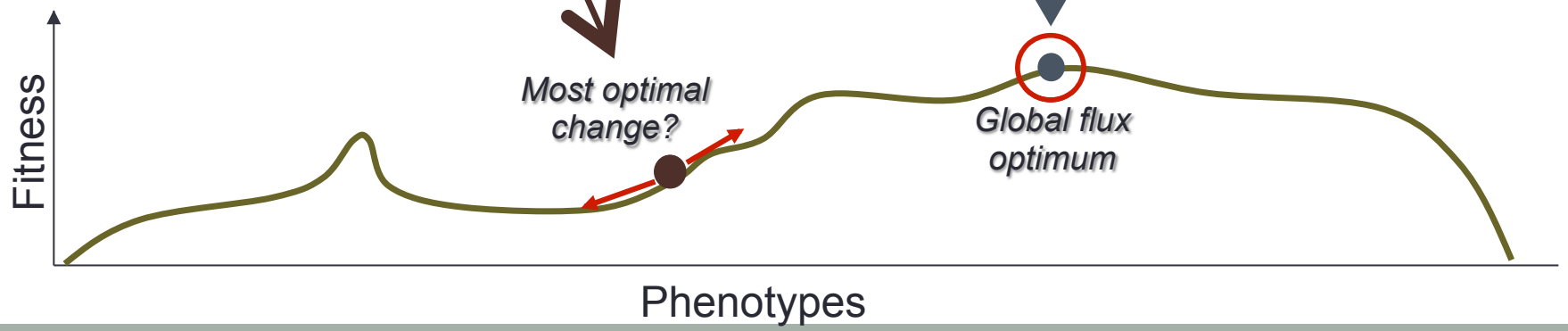
- **Kinetic model of chemical species through time (Michaelis-Menten)**
- **Need to know relevant parameters**
- **Sensitivity of flux to parameter changes ('control' coefficient, c_i)**
- **Most enzymes exert very little control on flux**
- **Assumption: flux α fitness**

Flux balance analysis (FBA)

"What is the optimal flux phenotype?"

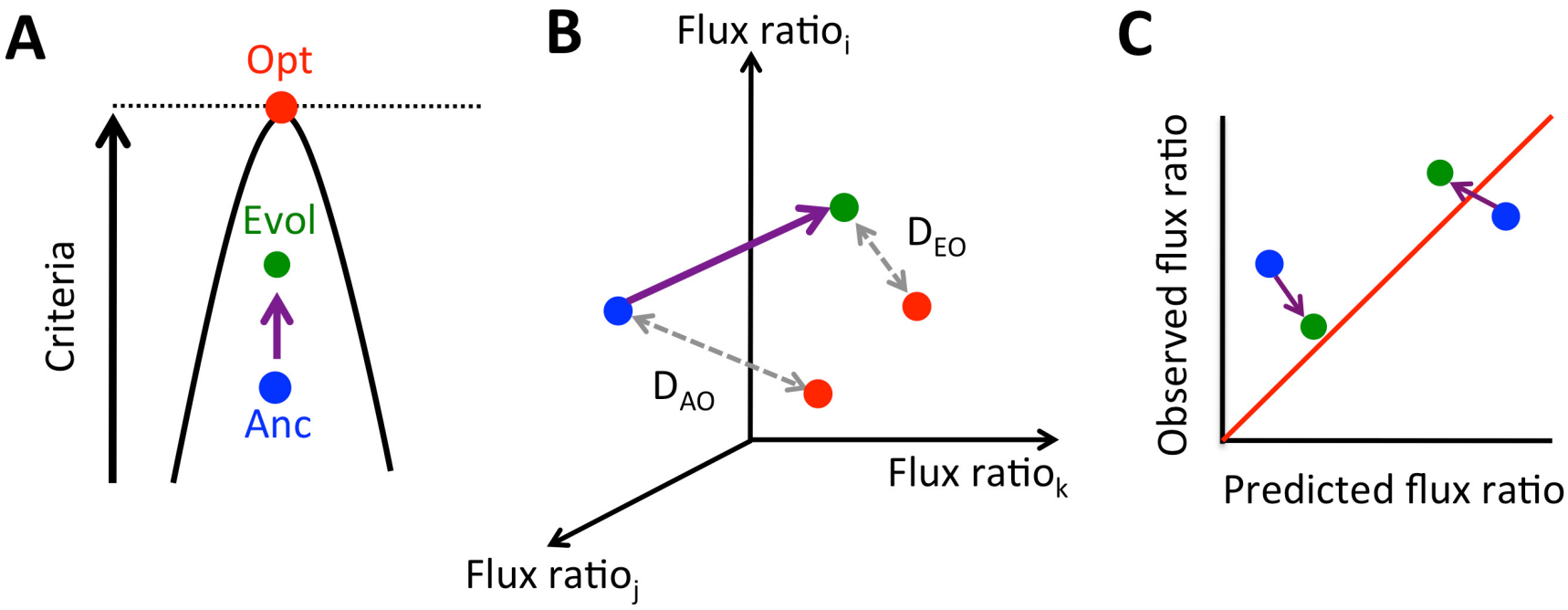
- **Describes network topology via a complete stoichiometric matrix**
- **Defines set of all possible balanced steady-states (w/in constraints)**
- **Choose flux state that is "optimal"**
- **Key assumption: past selection to become perfect for chosen criterion**

Adaptive landscape:



What does it mean to be optimal?

- No previous test if **internal fluxes** evolve to be FBA optimal
- Can compare predictions to measured fluxes in three ways:

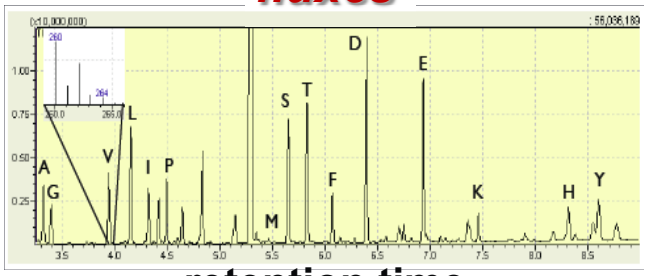


- What should be optimized? Normal = BM/S = **C yield**
- What is selected for in batch culture? Mainly growth **rate**

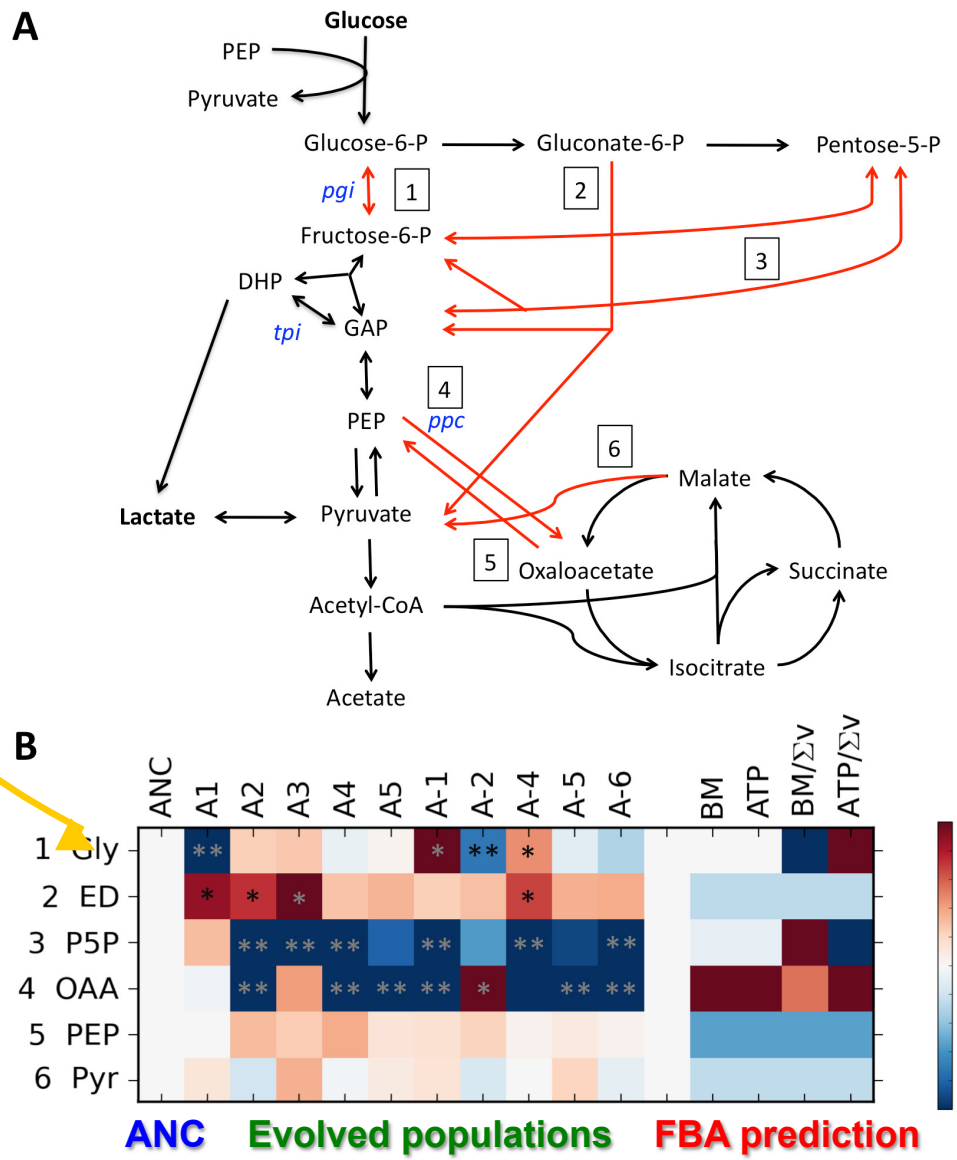
(Harcombe et al., 2013. *PLoS Computational Biology*)

Do fluxes evolve to FBA-predicted optimum?

^{13}C-method to assay fluxes

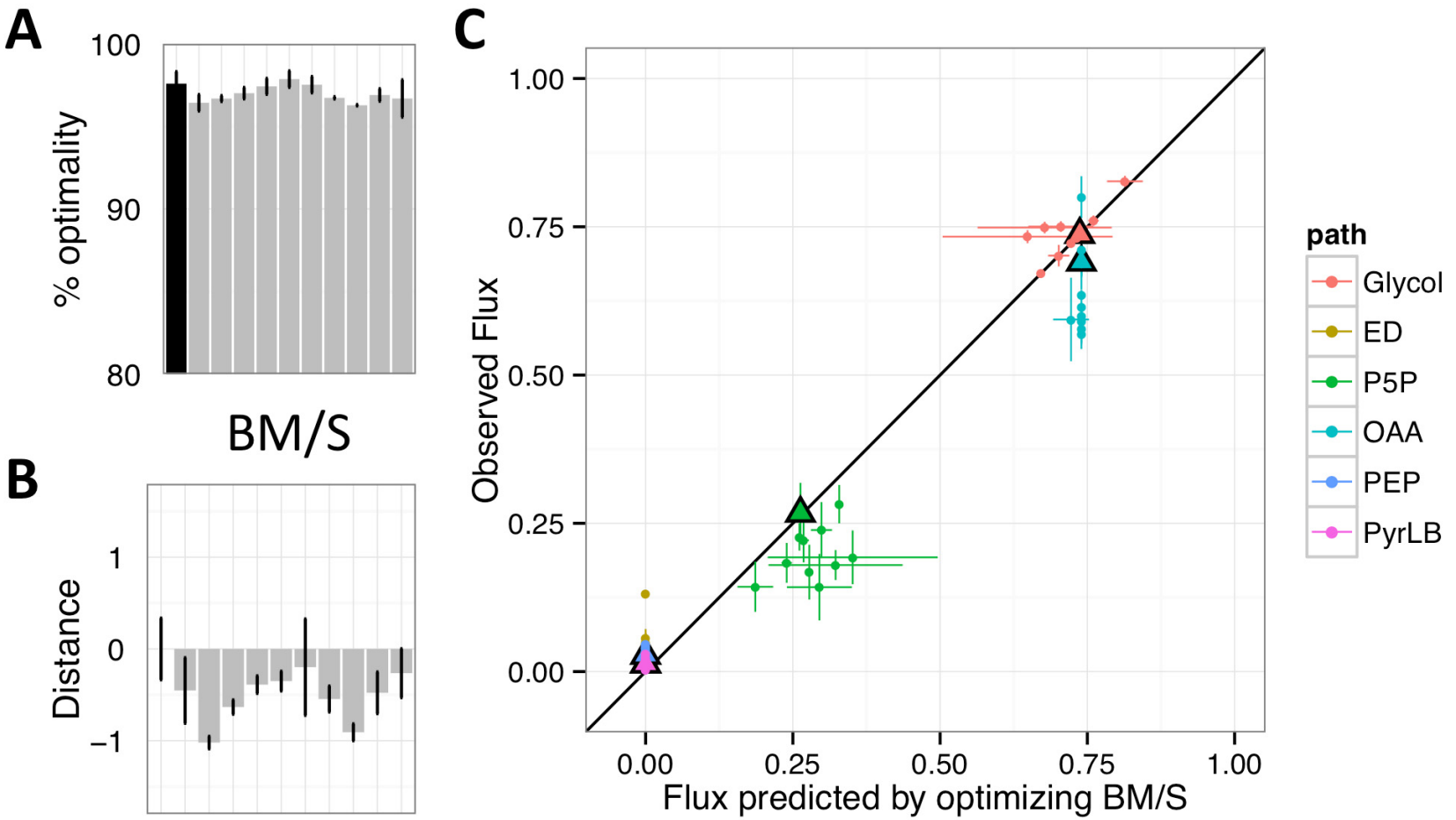


1. Harvest total protein
2. Hydrolyze amino acids
3. Derivatize
4. GC-MS for mass dist.
5. Analyze via network



(Harcombe et al., 2013. *PLoS Computational Biology*)

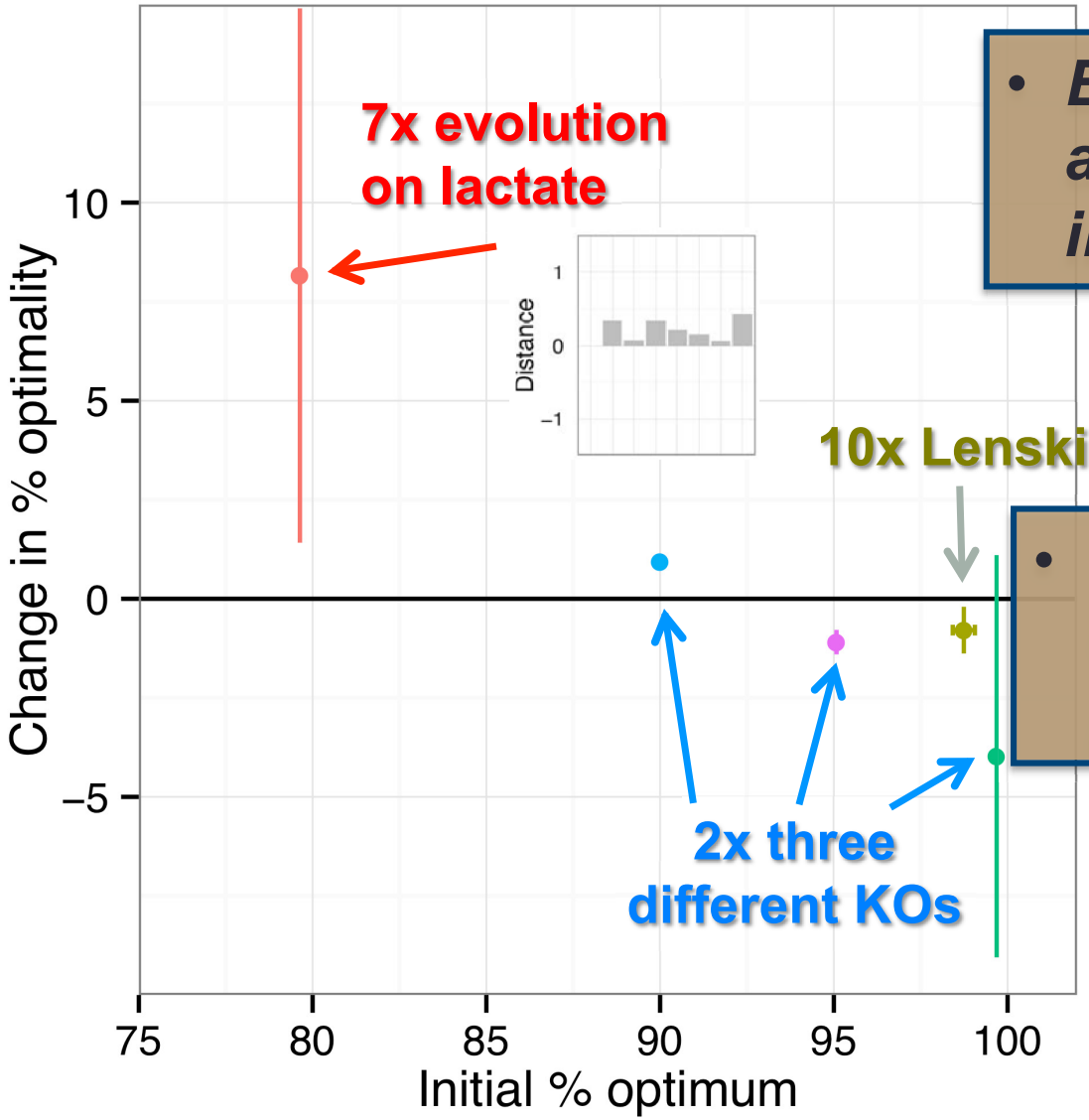
Do fluxes evolve to FBA-predicted optimum?



- ***Yield (BM/S) decreased slightly; fluxes consistently moved a bit away from optimum***

(Harcombe et al., 2013. *PLoS Computational Biology*)

Predictability depends on distance to optimum



- ***Either rate increased and slight drop, or both increased together***

- ***Take home: FBA optimizing for BM/S not perfect, but not bad...***

(Harcombe et al., 2013. *PLoS Computational Biology*)

Model system: C_1 in *Methylobacterium*

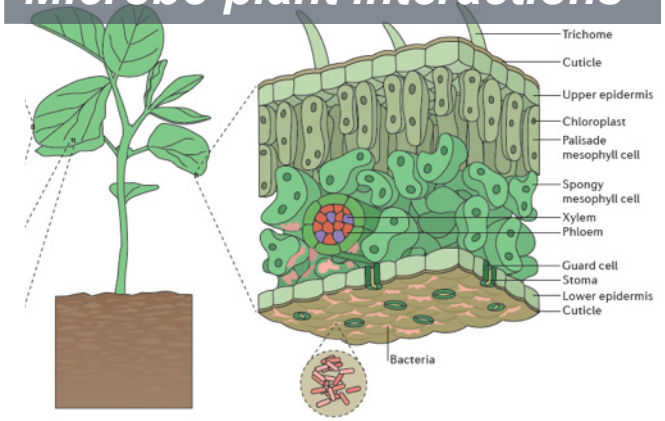
- *M. extorquens* is a plant epiphyte
- Model for C_1 metabolism (>50 yrs)

Leaf print on methanol plate.

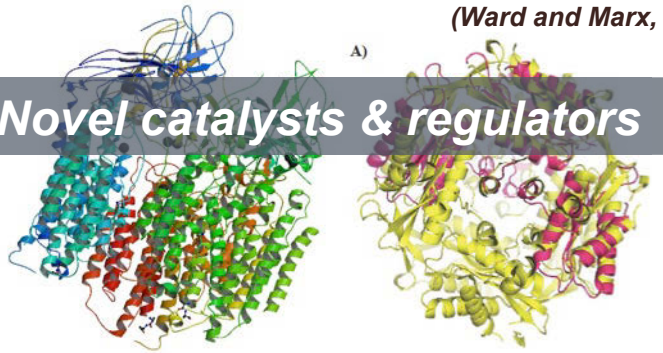
Plants release methanol during cell wall growth…

(Ward and Marx, unpublished)

Microbe-plant interactions



Novel catalysts & regulators



Soil ecology & global warming

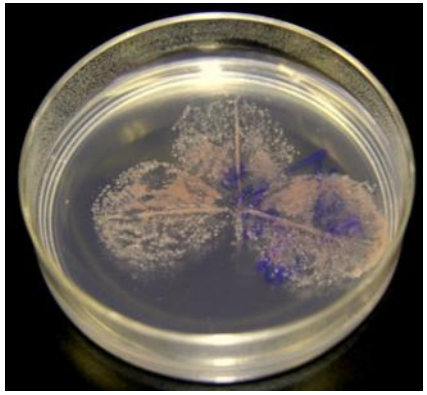
Methanol-based biotechnology

Model system: C_1 in *Methylobacterium*

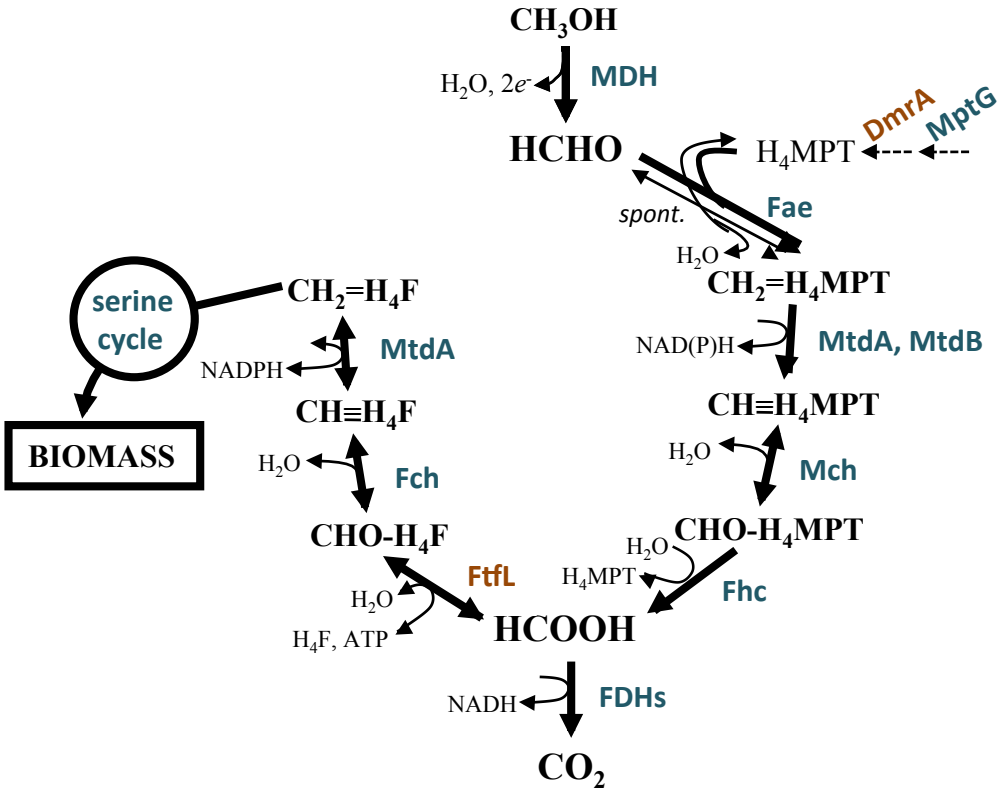
- *M. extorquens* is a plant epiphyte
- Model for C_1 metabolism (>50 yrs)

Leaf print on methanol plate.

Plants release methanol during cell wall growth…

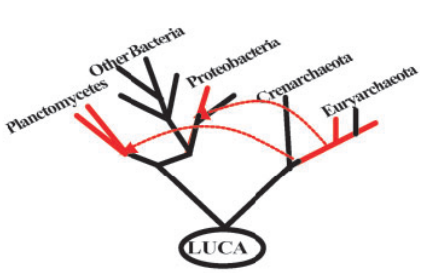


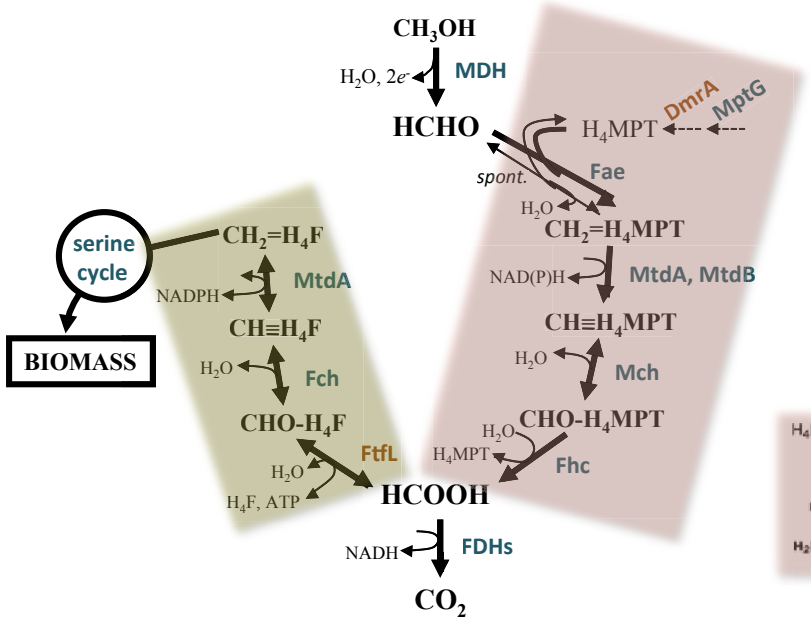
(Ward and Marx, unpublished)

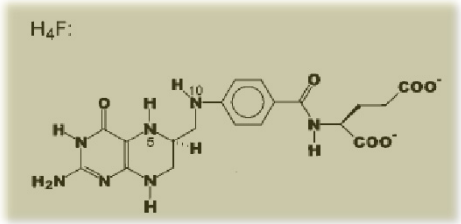


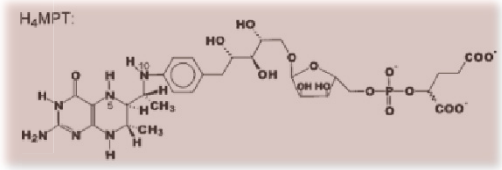
C_1 and HGT: new model for evolution

- C_1 genes transferred between bacteria; maybe from archaea
- **How are new functions incorporated?**
- Discovered one new gene (DmrA) that was *Methylobacterium*-specific innovation









DHFR

F → H_2F → H_4F

NADPH NADPH

DmrA

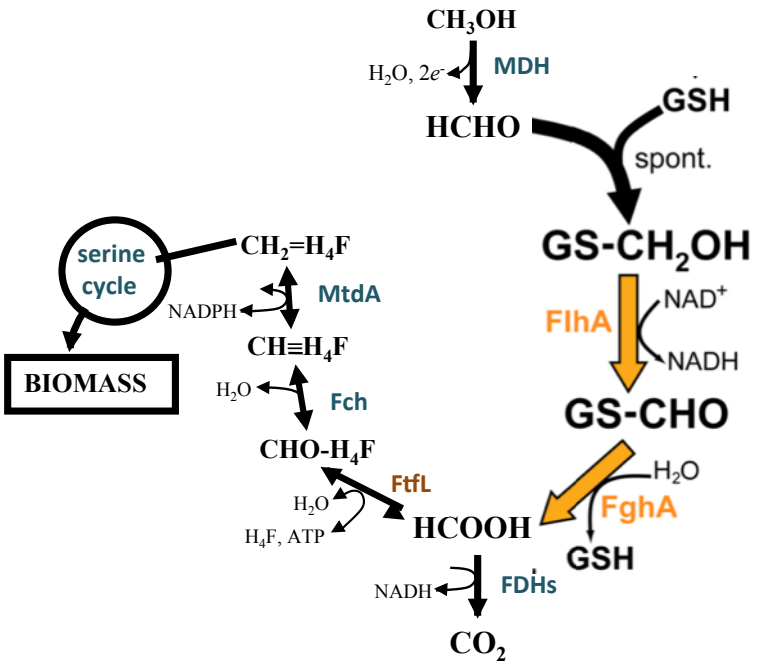
MPT → H_2MPT → H_4MPT

NADPH NADPH

(Marx et al., 2003a, 2003c. *J. Bacteriology*; Chistoserdova et al., 2004. *Mol. Biol. Evol.*; Kalyuzhnaya et al., 2005. *J. Bacteriology*)

Innovative combos not always initially fit...

C_1 and HGT: new model for evolution



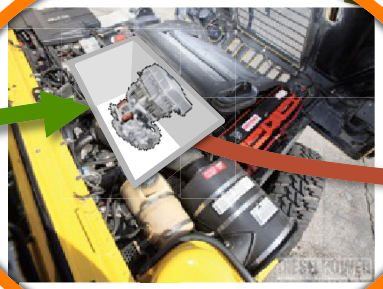
- Replaced existing formaldehyde oxidation pathway
  - Foreign GSH-dependent pathway from *Paracoccus denitrificans*
  - Single transcript, on plasmid, strong promoter
- ***Recovered growth, but 3x slower***

(Marx et al., 2003a, 2003b. *J. Bacteriology*; Chistoserdova et al., 2004. *Mol. Biol. Evol.*; Kalyuzhnaya et al., 2005. *J. Bacteriology*)

C_1 and HGT: new model for evolution

- An analogous scenario:

IN



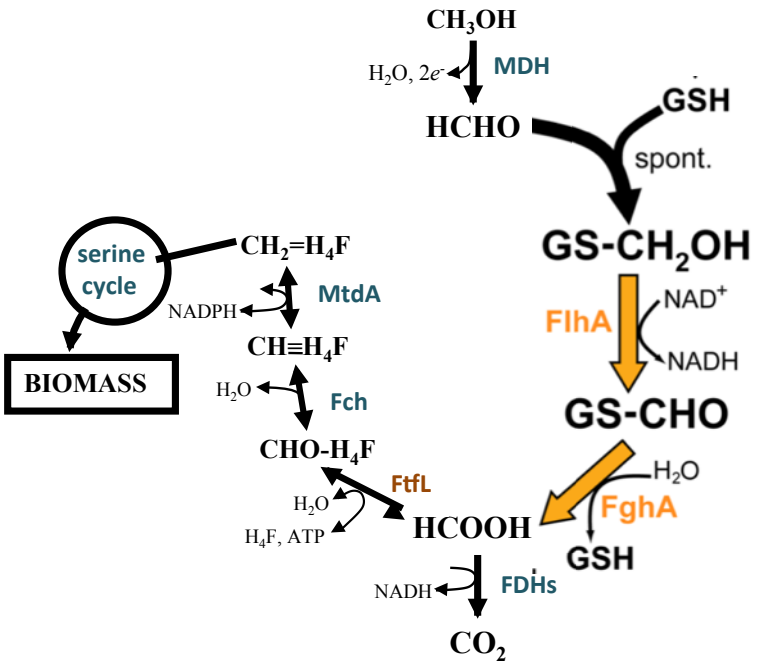
OUT

evolve…

+

???

C_1 and HGT: new model for evolution



- Replaced existing formaldehyde oxidation pathway
  - Foreign GSH-dependent pathway from *Paracoccus denitrificans*
  - Single transcript, on plasmid, strong promoter
- ***Recovered growth, but 3x slower***

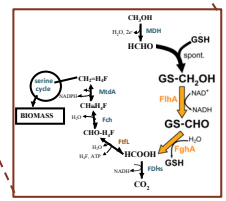
- ***Hypothesis, v.1:*** *Biggest benefit mutations directly involved with C_1 replacement*

- ***Hypothesis, v.2:*** *Metabolic model of C_1 catalysis can predict targets/interactions*

(Marx et al., 2003a, 2003b. *J. Bacteriology*; Chistoserdova et al., 2004. *Mol. Biol. Evol.*; Kalyuzhnaya et al., 2005. *J. Bacteriology*)

Evolution with engineered HGT

Engineered ancestor

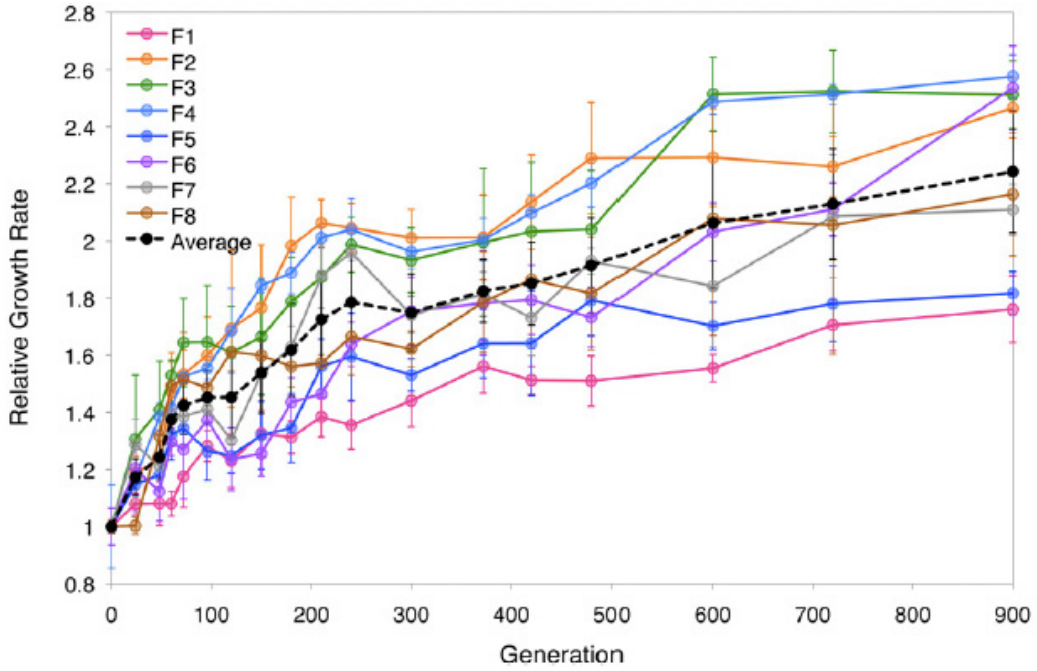


Evolved in methanol

900 gen.
$N_e \approx 10^8$

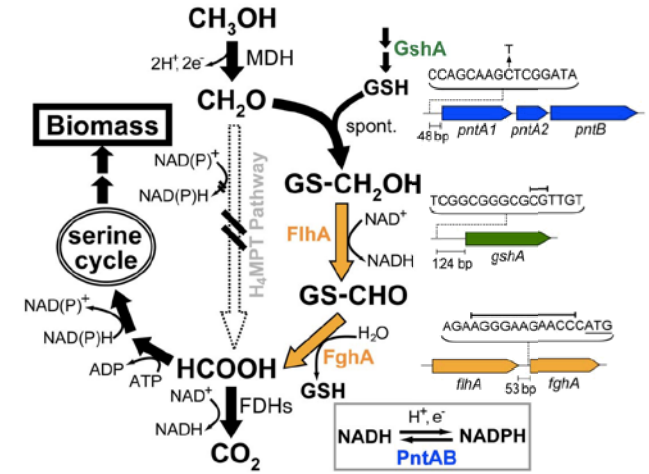
-80 °C

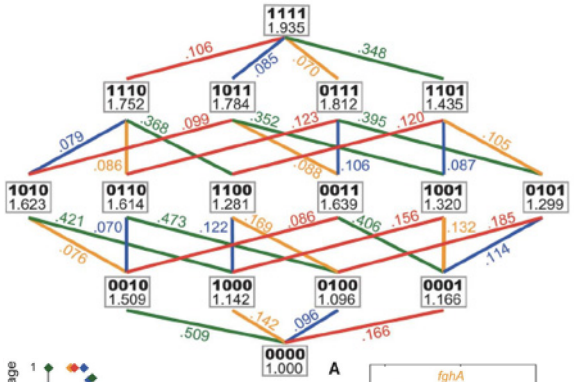
- Dramatic (>2X) but varied improvement
- Almost entirely specific to methanol
- *What changed?*

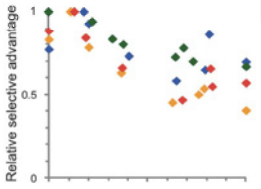


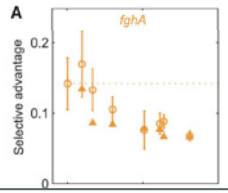
(Chou et al., 2011. *Science*; Lee and Marx, 2013. *Genetics*)

Beneficial mutations outside pathway



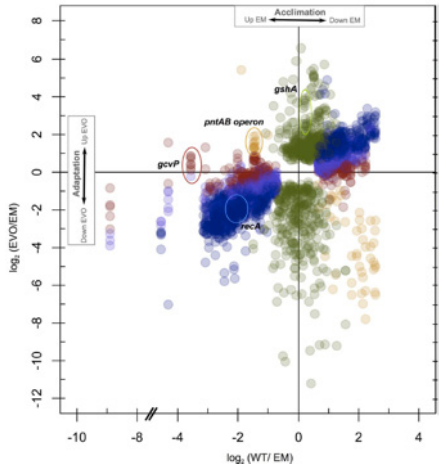


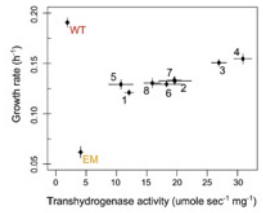


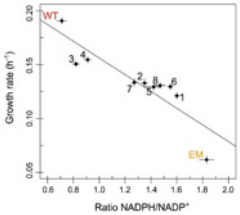


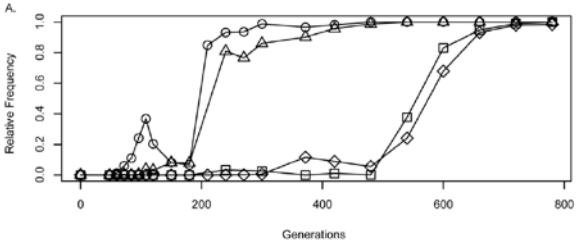
$$W_{\text{mutant}} = \prod_{i \in \text{Alleles}} \lambda_i b_0 - \prod_{i \in \text{Alleles}} \theta_i c_0$$

- Most mutations in or directly related to new C_1 pathway
- Uncovered general pattern of diminishing returns that decelerates adaptation
- Could predict epistasis at rough level
- Broad physiological disturbance and recovery (global mRNA; NAD(P)H levels)
- Uncovered order and dynamics of alleles



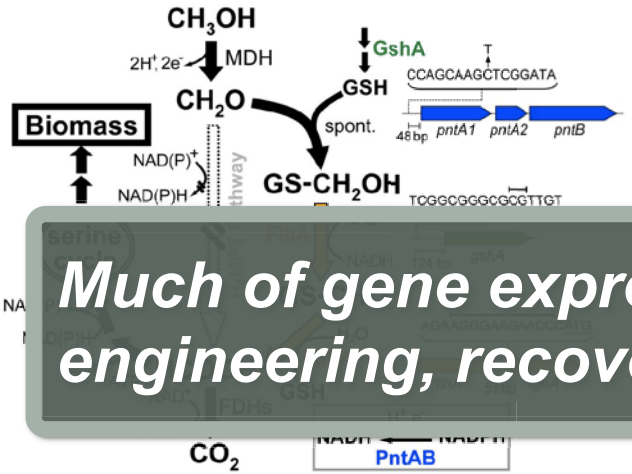






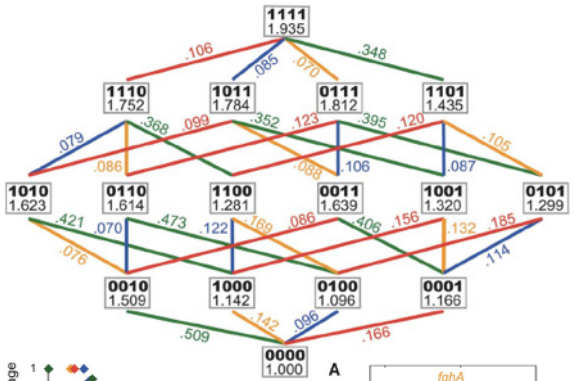
(Chou et al., 2011. *Science*; Chubiz et al., 2012. *PLoS One*; Lee and Marx, 2013. *Genetics*; Carroll and Marx, 2013. *PLoS Genetics*)

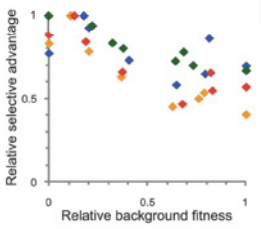
Beneficial mutations outside pathway

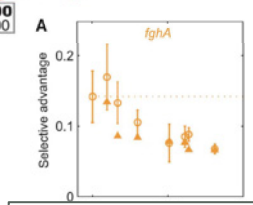


- Most mutations in or directly related to new C_1 pathway
- Uncovered general pattern of diminishing returns that decelerates adaptation
- Could predict mutations at gene level
- Broad physiological disturbance and recovery (global mRNA; NAD(P)H levels)
- Uncovered order and dynamics of alleles

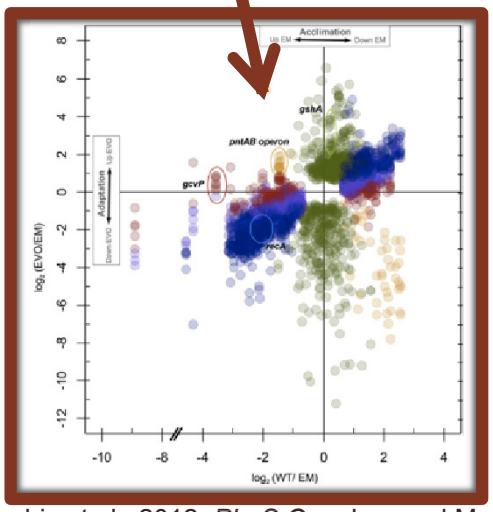
Much of gene expression changed upon engineering, recovered in parallel with growth rate

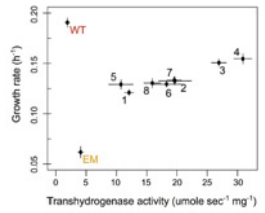


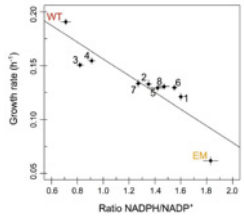


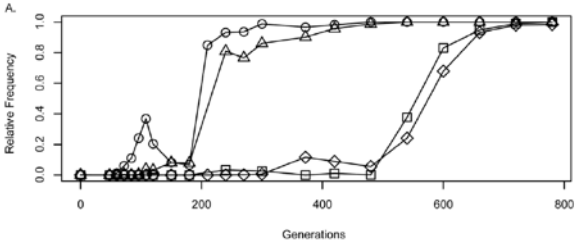


$$W_{\text{mutant}} = \prod_{i \in \text{Alleles}} \lambda_i b_0 - \prod_{i \in \text{Alleles}} \theta_i c_0$$





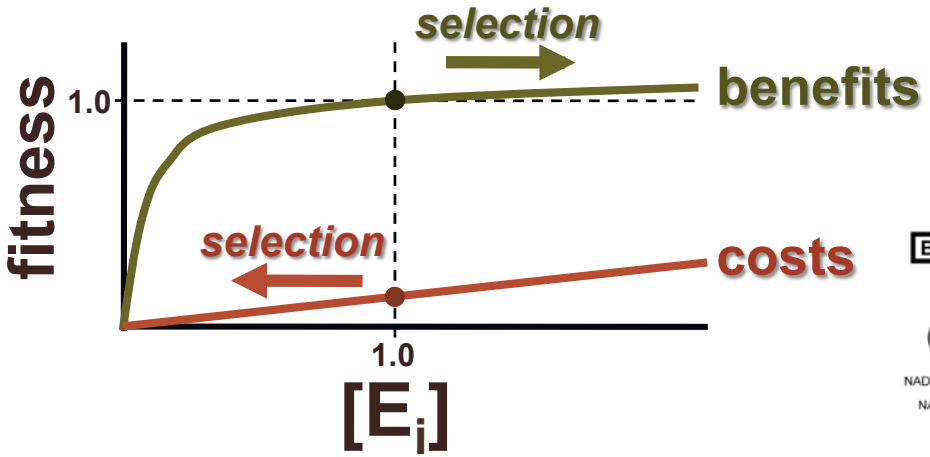


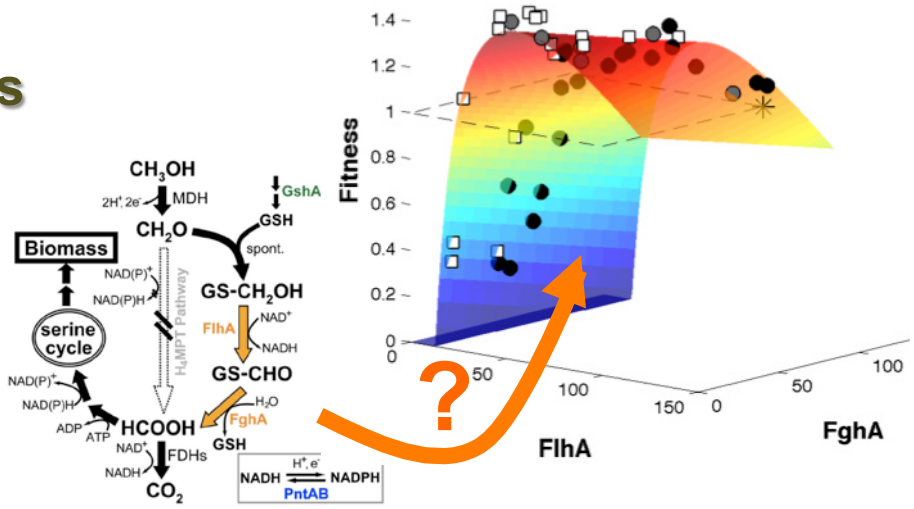


(Chou et al., 2011. *Science*; Chubiz et al., 2012. *PLoS One*; Lee and Marx, 2013. *Genetics*; Carroll and Marx, 2013. *PLoS Genetics*)

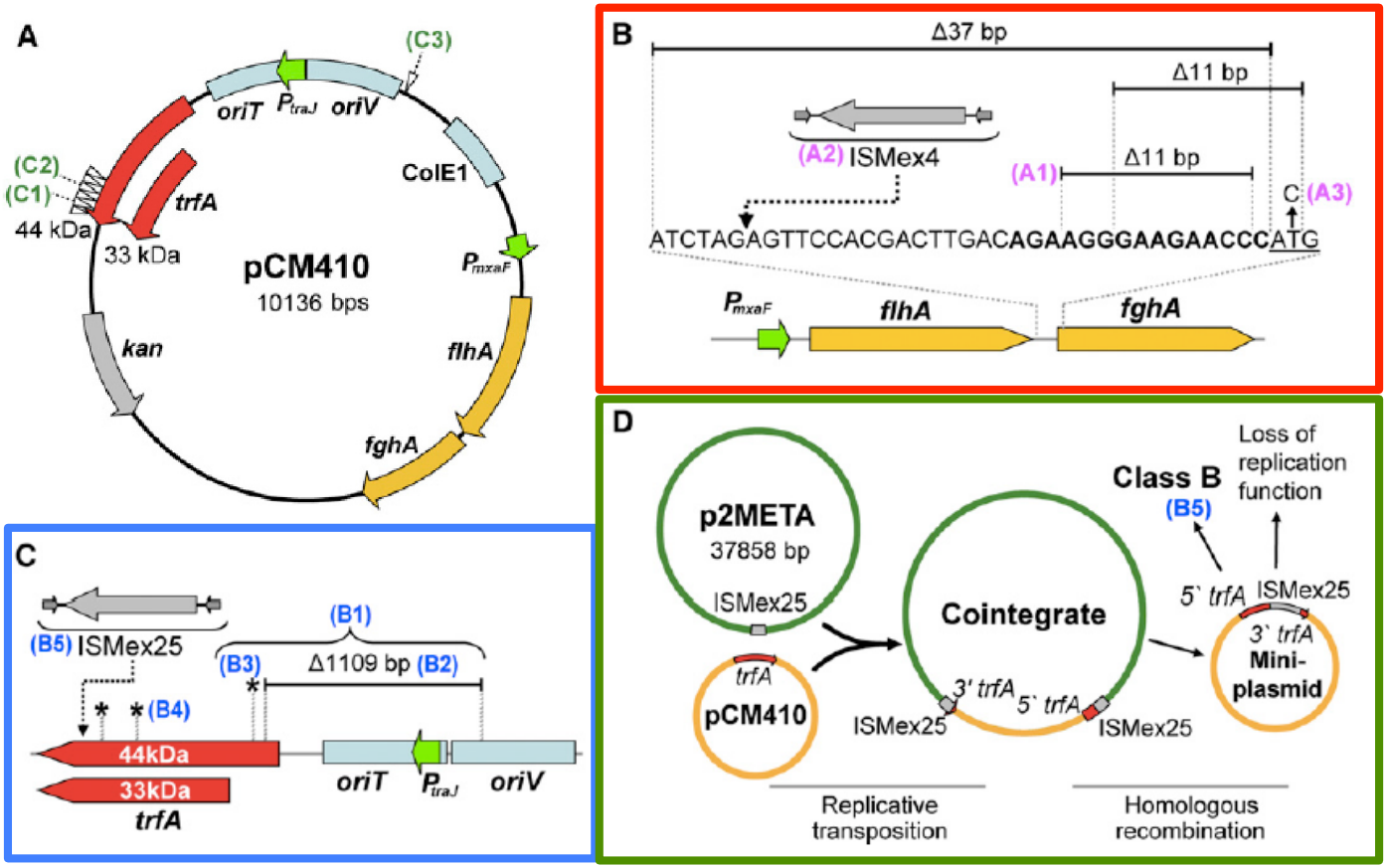
Optimize (benefits–costs) for pathway

- Need enzymes for catalysis, but they are costly to make.
- What is the optimal amount across a pathway?
- How do populations evolve toward this?
- How well do expression-altering mutations interact?
- *Predict this based upon mapping phenotypes to fitness?*





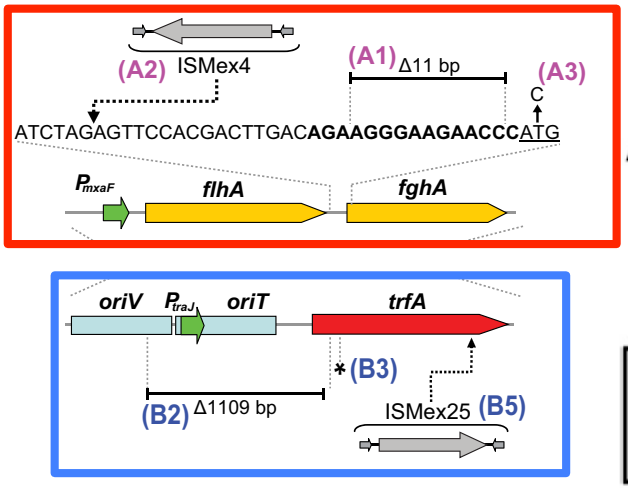
Beneficial mut. affecting pathway exp.

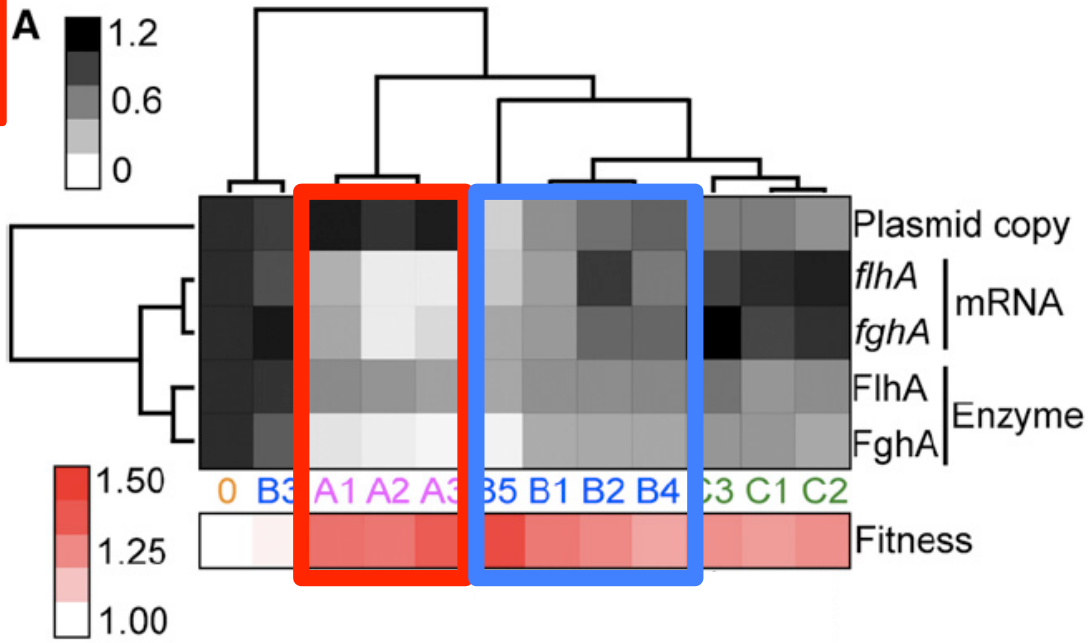


- Mutations of many types in different populations; 25-45% benefit; affect expression differently

(Chou and Marx, 2012. *Cell Reports*)

Expression per copy vs. copy#

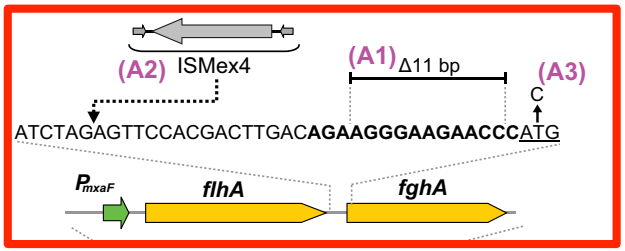


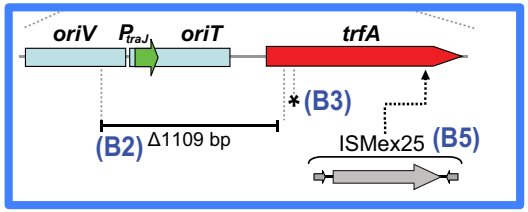


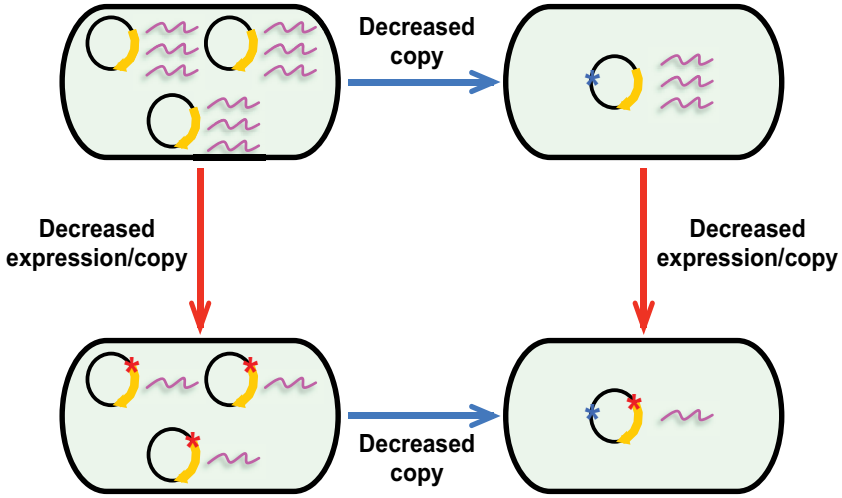
- Distinct, independent mechanisms to **reduce** expression of the GSH pathway enzymes
- But no promoter mutations?...

(Chou and Marx, 2012. *Cell Reports*)

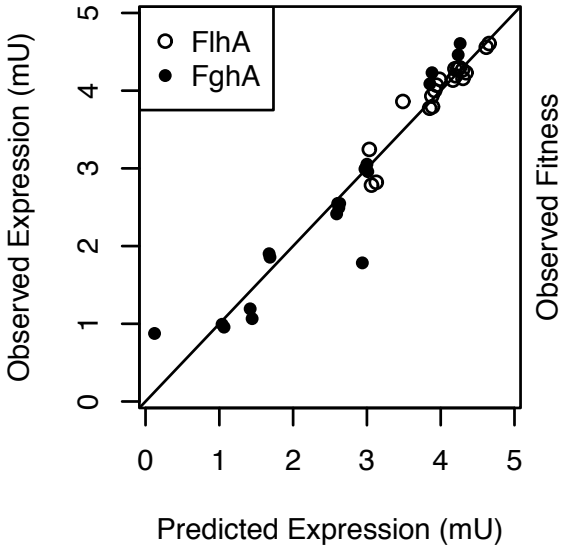
Independent effects upon expression

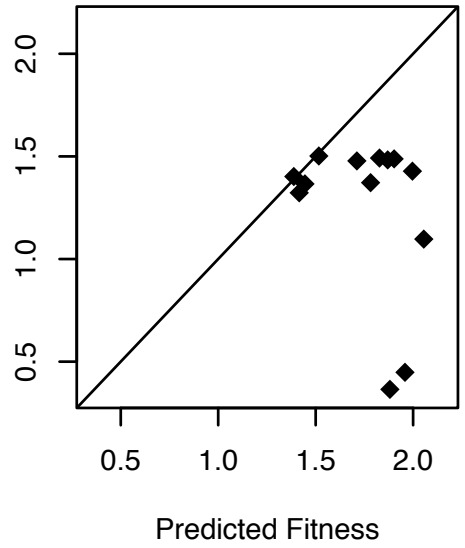






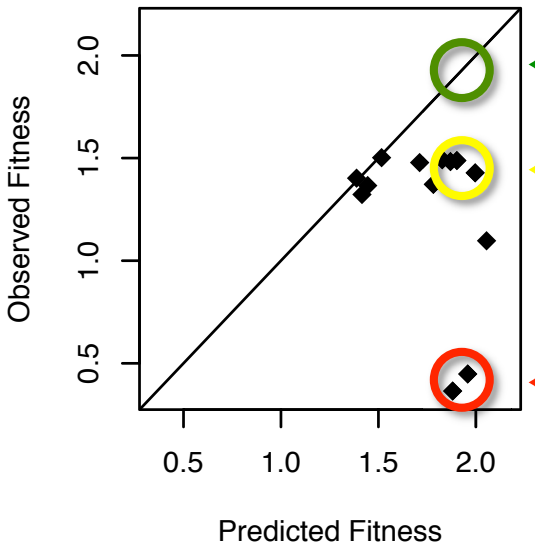
- Two classes should interact independently:
  - $E_{AB} = E_A \times E_B$
  - Yes
- Indep. upon fitness?
  - **No.**





(Chou et al., *PLoS Genetics*, 2014)

Why antagonism and sign epistasis?

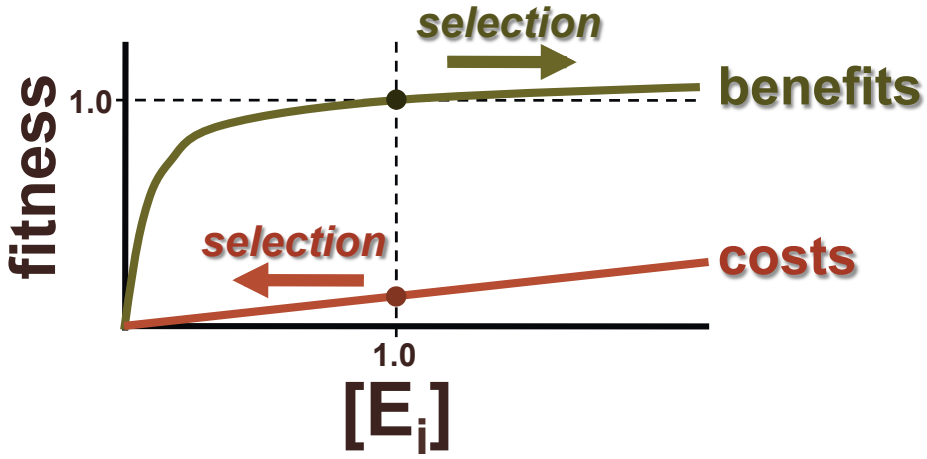


$W_{AB} = W_A \times W_B$; both provide benefit when together

$W_{AB} \approx \max(W_A, W_B)$; pair same as best individual

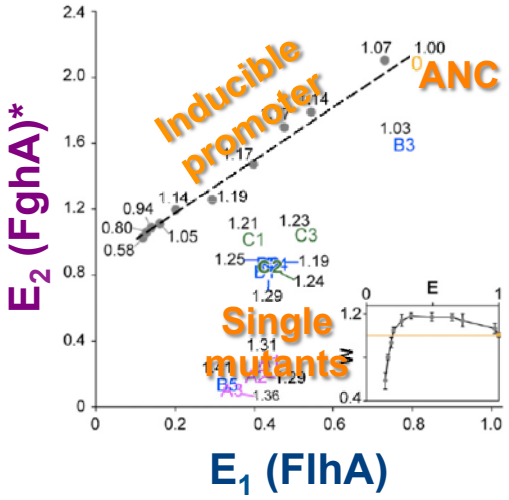
$W_{AB} < W_{ANC}$; pair much worse than individual mut.

- Caused by tension between benefits and costs?
  - Metabolic Control Analysis (MCA)



(Chou et al., *PLoS Genetics*, 2014)

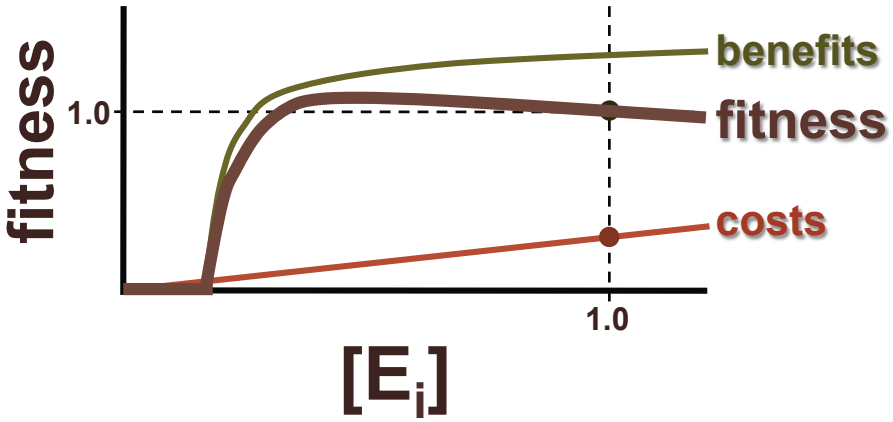
Map expression to fitness via model



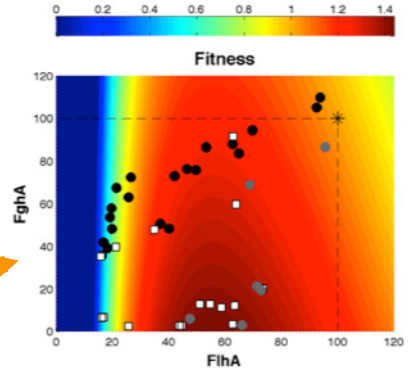
*FghA necessary for catalysis but not in benefit term because all data higher than threshold

$$W = \textit{Flux above threshold} - \textit{enzyme costs}$$

$$W = (v_{max} \times \mathbf{E_1}/(\mathbf{E_1} + E_{½\,max}) - v_T) - a \times \mathbf{E_1} - b \times \mathbf{E_2}$$

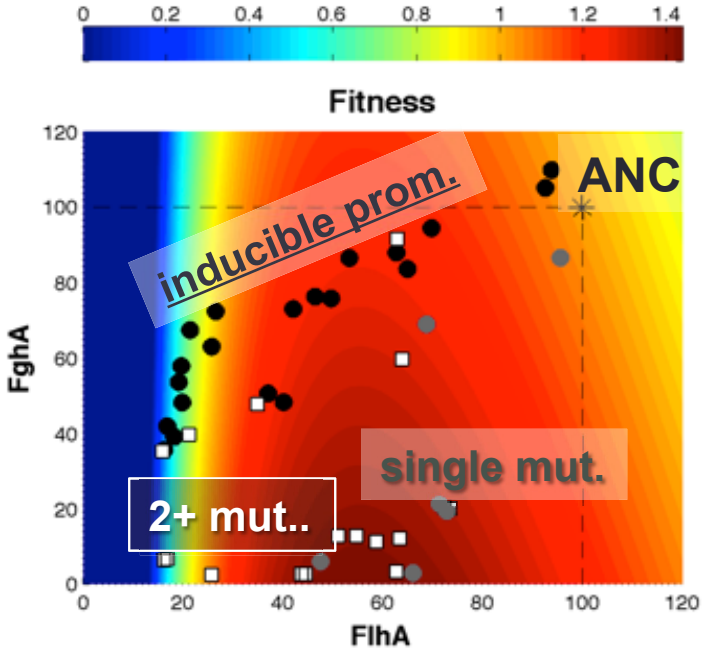


- ANC, single mutants, inducible promoter constructs (27 data points) to fit parameters
  - Try to predict 17 mutational combinations

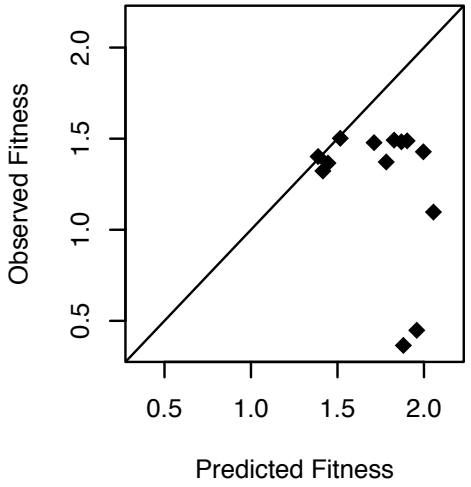


(Chou and Marx, 2012. *Cell Reports*; Chou et al., *PLoS Genetics*, 2014)

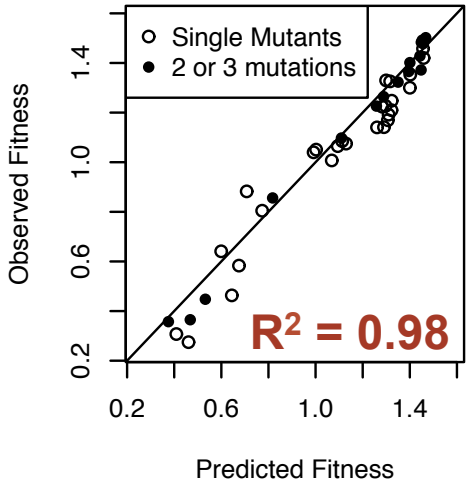
Model predicts mutational combos



$$W_{AB} = W_A \times W_B$$



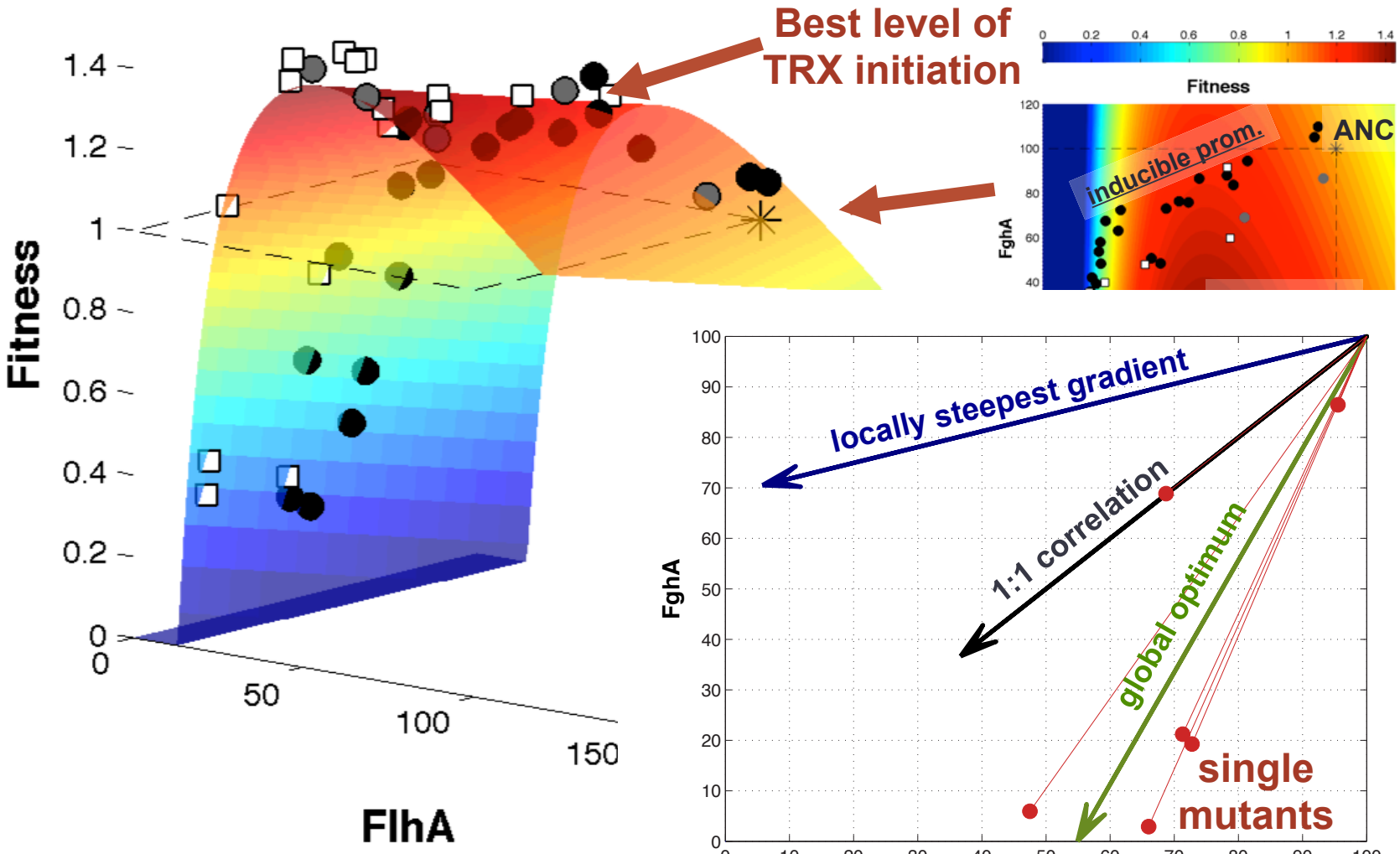
$$W = (v_{max} \times \mathbf{E_1}/(\mathbf{E_1} + E_{½\ max}) - v_T) - a \times \mathbf{E_1} - b \times \mathbf{E_2}$$



- Mechanistic model works quite well

(Chou et al., *PLoS Genetics*, 2014)

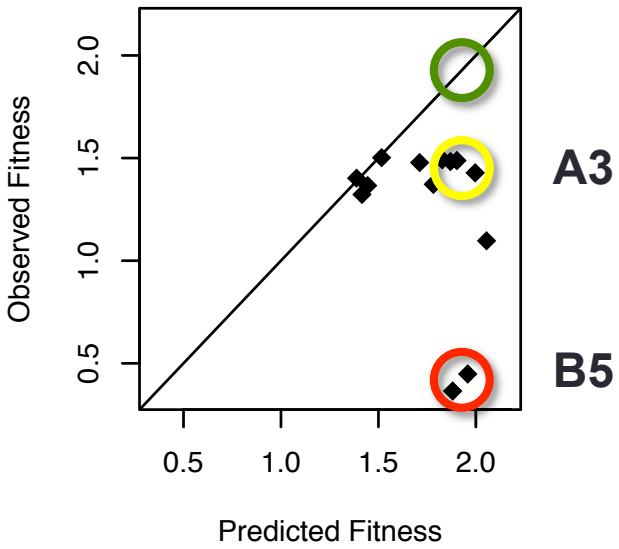
Interpret adaptation in light of model

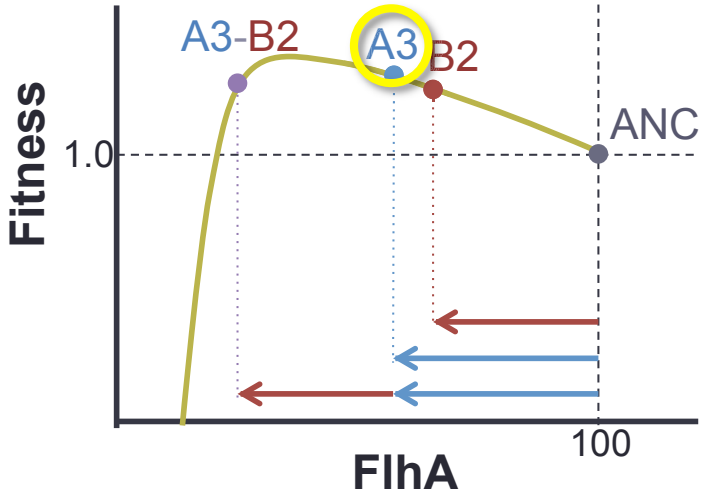


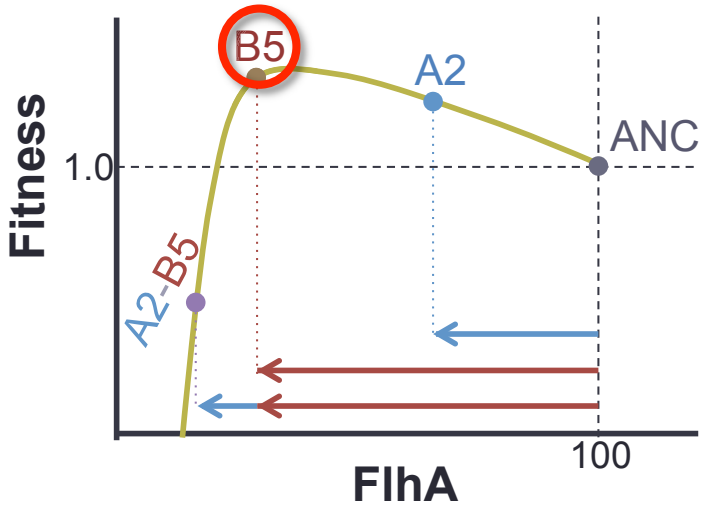
- *What direction was phenotypic change?*
- *How far to optimum was achieved?*

(Chou et al., *PLoS Genetics*, 2014)

Interpret interactions in light of model



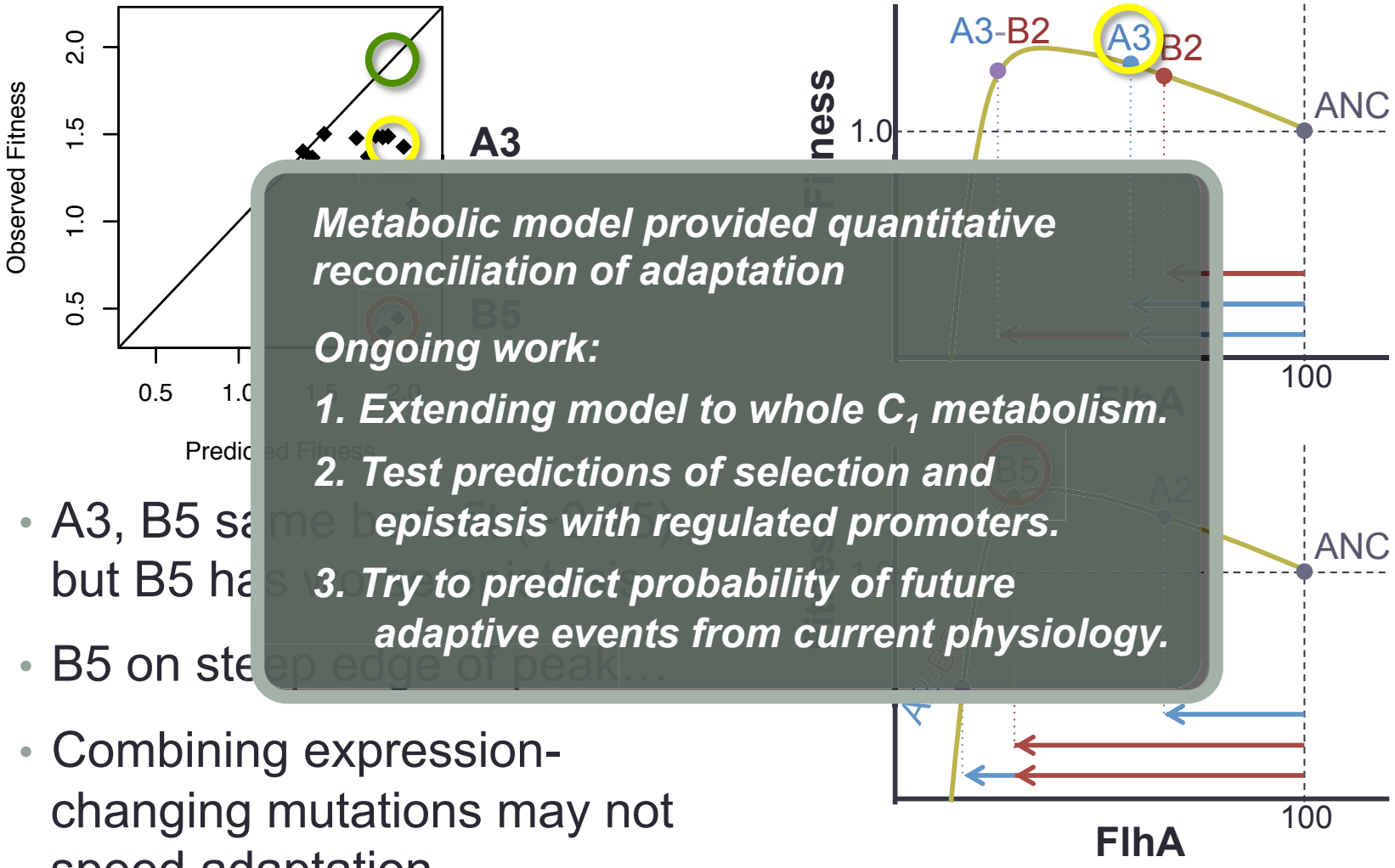




- A3, B5 same benefit (~0.45), but B5 has worse epistasis
- B5 on steep edge of peak…
- Combining expression-changing mutations may not speed adaptation

(Chou et al., *PLoS Genetics*, 2014)

Interpret interactions in light of model



Metabolic model provided quantitative reconciliation of adaptation

Ongoing work:

1. ***Extending model to whole C_1 metabolism.***
2. ***Test predictions of selection and epistasis with regulated promoters.***
3. ***Try to predict probability of future adaptive events from current physiology.***

- A3, B5 same benefit (~0.5), but B5 has worse epistasis
- B5 on steep edge of peak…
- Combining expression-changing mutations may not speed adaptation

(Chou et al., *PLoS Genetics*, 2014)

Acknowledgements

Collaborators

- Daniel Segrè (Boston U.)
- Nate Cady (U. Albany)
- Colleen Hansel (WHOI)
- Trent Northern (LBNL)
- Victoria Orphan (CalTech)
- Claus Wilke, Jeff Barrick (UT Austin)
- Josh Plotkin (U. Penn), Jeremy Draghi (UBC)
- Yousif Shamoo (Rice U.)
- Marty Ytreberg (U. Idaho)
- Stéphane Vuilleumier, Françoise Bringel (U. Strasbourg)
- Owen Leiser, Paul Keim (Northern Arizona U.)
- Zachary Blount, Rich Lenski (Michigan State U.)
- Susi Remold (U. Louisville)
- Ben Kerr (U. Washington)
- Dan Davidi, Ron Milo (Weizmann)
- Mike Springer, Ann Pearson (Harvard U.)

Funding

- NIH (R01 GM078209)
- NSF (EAR-1024723)
- NSF (DEB-0845893, CAREER Award)
- NSF (IOB-0612591)
- DOE (DE-SC0006731)
- DOE (PI: Segrè, DE-SC0004962)
- DoD (PI: Wilke, 61789-MA-MUR)
- BEACON (PI: Barrick)
- William H. Milton Fund
- Clark/Cooke Fund

(Sketch: Maria Marx)


University of Idaho
Siavash Riazi
Josh Michener
Ozzy Baltekin
José Rojas Echenique
Nigel Delaney
Lon Chubiz
Sarah Douglas
Will Harcombe
Alex Bradley
Sean Carroll
Nick Leiby
Paige Swanson
Dipti Nayak
Deepa Agashe
Jessica Purswani
Maryska Kaczmarek

View from our living room