

Mathematical Programming for Cross-Species Comparison of Genome-Scale Metabolic Models

Joshua J. Hamilton and Jennifer L. Reed

Department of Chemical and Biological Engineering
University of Wisconsin-Madison

AIChE Annual Meeting
October 20, 2011



WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON



Outline

- 1 Introduction and Background
- 2 Method
- 3 Identifying Novel Drug Targets

Why models?

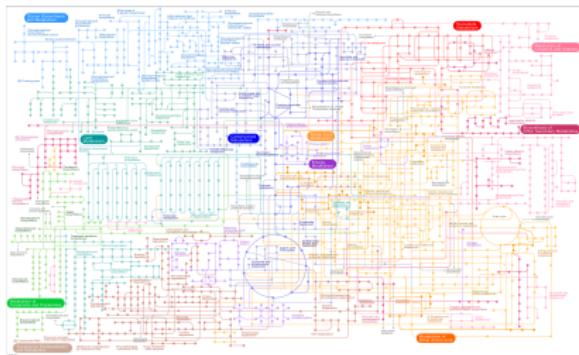
- Understanding cellular behavior
- Evaluate and design metabolic engineering strategies
- Contextualize high-throughput data

Genome-Scale Models & Constraint-Based Modeling

Why models?

- Understanding cellular behavior
- Evaluate and design metabolic engineering strategies
- Contextualize high-throughput data

Metabolism



Cells want to

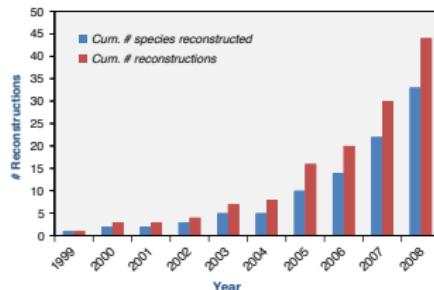
- Maximize growth

But are subject to

- Steady-state mass balance constraints
- Limits on fluxes

Advances in Model Development

- Number of available models growing exponentially ¹

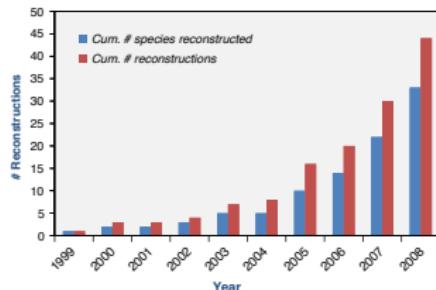


¹Oberhardt et al, Mol Sys Bio 2009

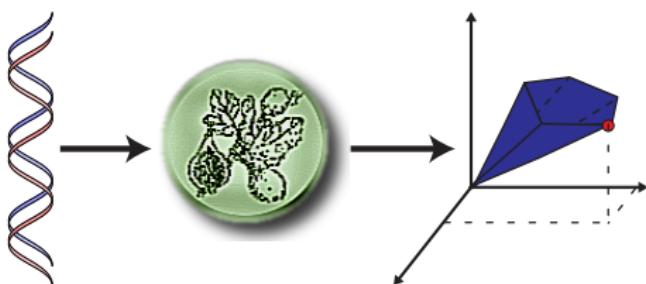
²Henry et al, Nature Biotech 2010

Advances in Model Development

- Number of available models growing exponentially ¹



- New tools enable rapid conversion from genome to model ²



¹Oberhardt et al, Mol Sys Bio 2009

²Henry et al, Nature Biotech 2010



Comparison of Genome-Scale Models

- Existing methods ^{3,4}
 - Are limited to identifying *structural* network differences
 - Involve complex work flows
 - Require mapping of compounds and reactions across networks

³Thiele et al, Mol Sys Bio 2010

⁴Oberhardt et al, PLoS Comp Bio 2011



Comparison of Genome-Scale Models

- Existing methods^{3,4}
 - Are limited to identifying *structural* network differences
 - Involve complex work flows
 - Require mapping of compounds and reactions across networks
- Alignment at the gene level
 - Can be done automatically
 - Enables identification of *functional* network differences

³Thiele et al, Mol Sys Bio 2010

⁴Oberhardt et al, PLoS Comp Bio 2011



Outline

1 Introduction and Background

2 Method

3 Identifying Novel Drug Targets

CONGA: Comparison of Networks by Gene Alignment

- Bilevel Program

CONGA Formulation⁵

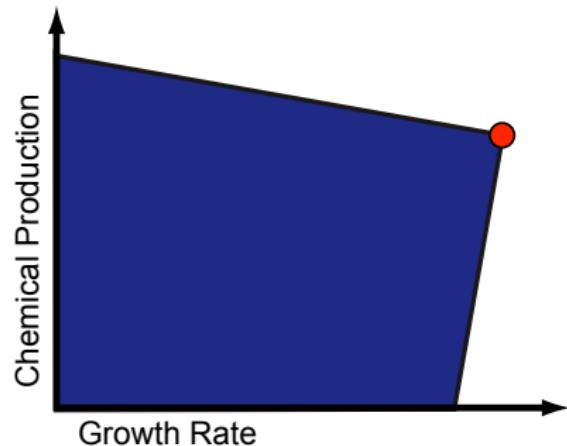
max

s.t. **max** cellular objective

s.t. cellular constraints

max cellular objective

s.t. cellular constraints



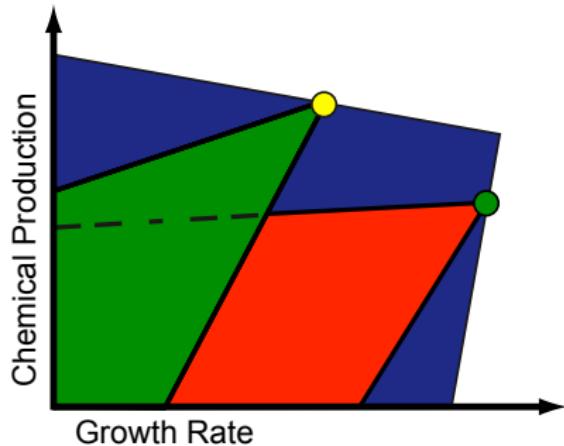
⁵Hamilton and Reed, under review

CONGA: Comparison of Networks by Gene Alignment

- Bilevel Program
 - Identify genetic perturbations

CONGA Formulation ⁵

max
s.t. **max** cellular objective
 s.t. cellular constraints
max cellular objective
 s.t. cellular constraints
number of perturbations \leq limit



⁵Hamilton and Reed, under review

CONGA: Comparison of Networks by Gene Alignment

- Bilevel Program
 - Identify genetic perturbations
 - Giving rise to a difference in phenotype

CONGA Formulation ⁵

max difference in cellular phenotype

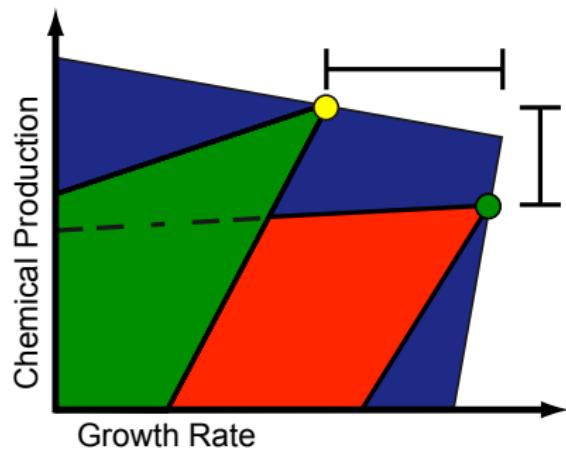
s.t. **max** cellular objective

s.t. cellular constraints

max cellular objective

s.t. cellular constraints

number of perturbations \leq limit



⁵Hamilton and Reed, under review

CONGA: Comparison of Networks by Gene Alignment

- Bilevel Program
 - Identify genetic perturbations
 - Giving rise to a difference in phenotype
 - Which point to genetic differences w.r.t. that phenotype

CONGA Formulation ⁵

max difference in cellular phenotype

s.t. **max** cellular objective

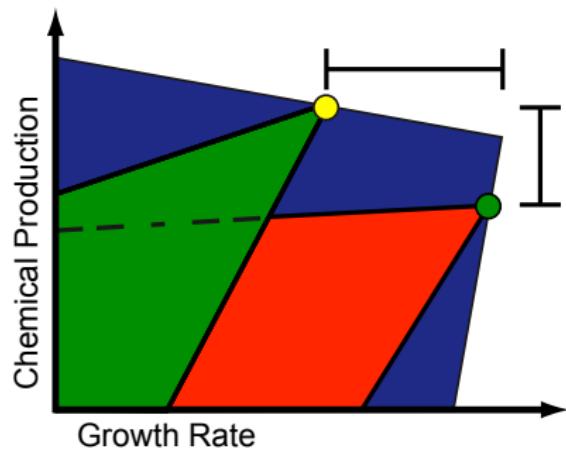
s.t. cellular constraints

max cellular objective

s.t. cellular constraints

number of perturbations \leq limit

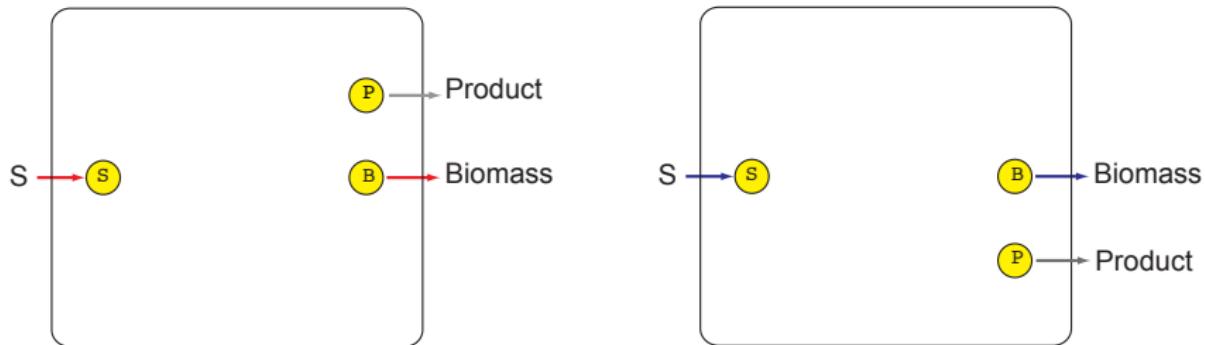
orthologs perturbed in common



⁵Hamilton and Reed, under review

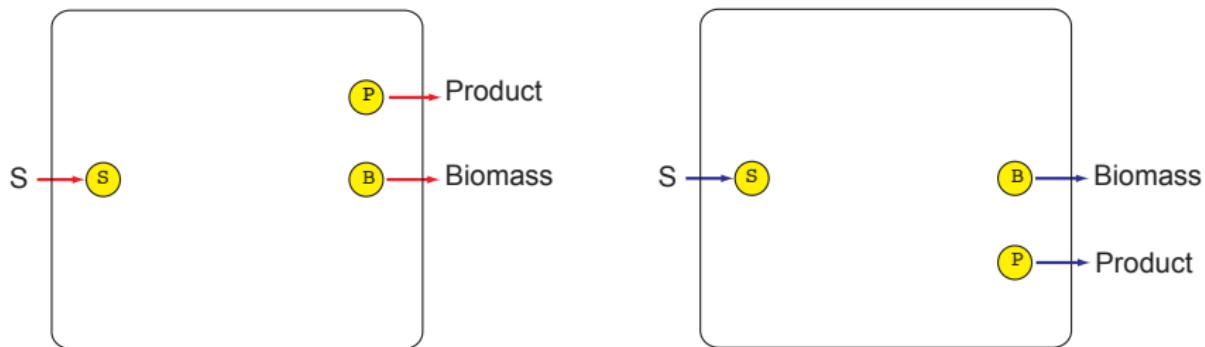
An Illustrative Example

- Two networks capable of producing product P from substrate S



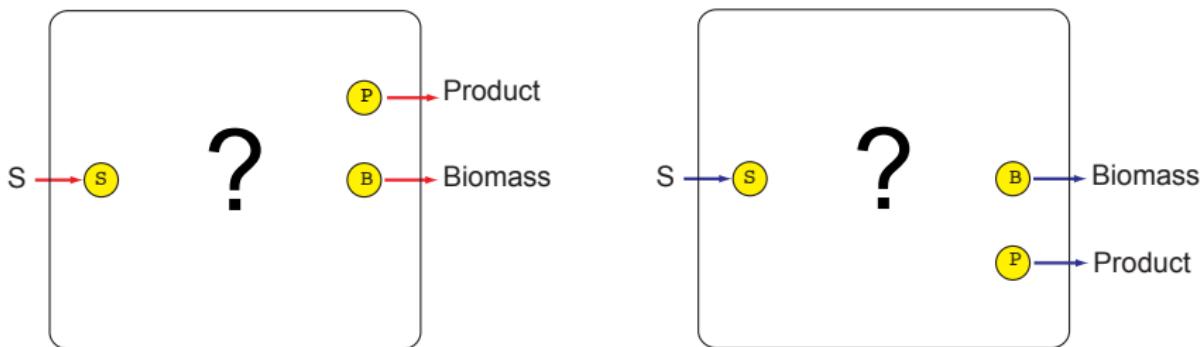
An Illustrative Example

- Two networks capable of producing product P from substrate S
- Genetic engineering required to couple P to biomass



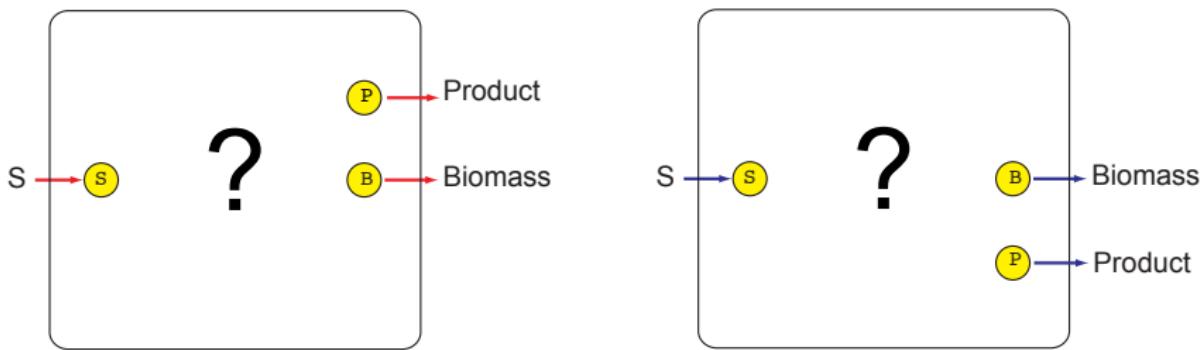
An Illustrative Example

- Two networks capable of producing product P from substrate S
- Genetic engineering required to couple P to biomass



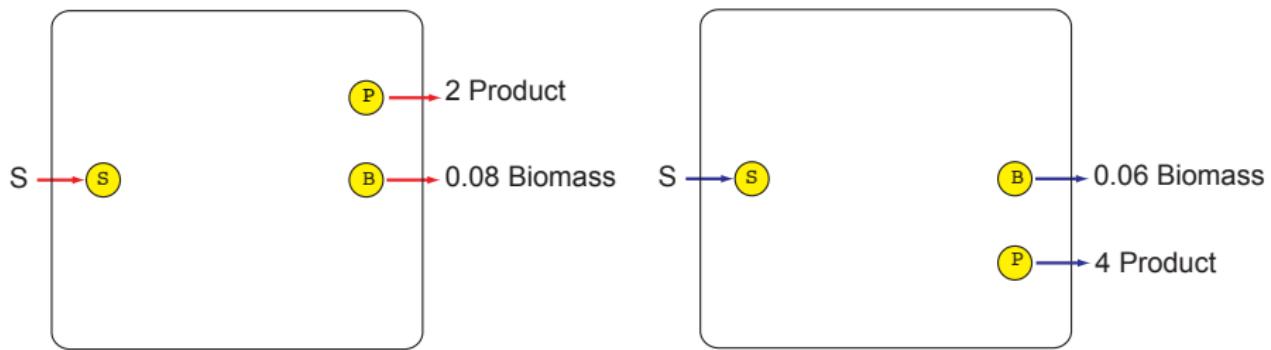
An Illustrative Example

- Two networks capable of producing product P from substrate S
- Genetic engineering required to couple P to biomass
- Outer objective: $\max v_{P,blue} - v_{P,red}$



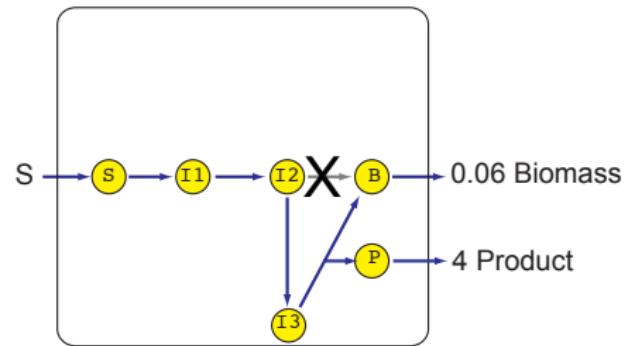
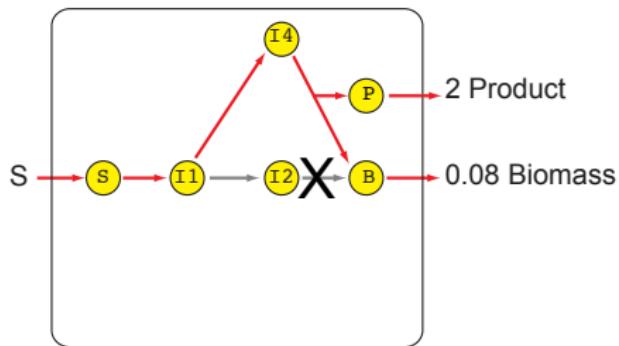
An Illustrative Example

- Two networks capable of producing product P from substrate S
- Genetic engineering required to couple P to biomass
- Outer objective: $\max v_{P,blue} - v_{P,red}$



An Illustrative Example

- Two networks capable of producing product P from substrate S
- Genetic engineering required to couple P to biomass
- Outer objective: $\max v_{P,blue} - v_{P,red}$
- Examination of the network reveals a unique synthesis pathways





Outline

- 1 Introduction and Background
- 2 Method
- 3 Identifying Novel Drug Targets



Identifying Novel Drug Targets

- *S. aureus*⁶ and *M. tuberculosis*⁷ are deadly human pathogens with well-studied metabolic networks

⁶Becker and Palsson, BMC Microbio 2005

⁷Jamshidi and Palsson, BMC Sys Bio 2007



Identifying Novel Drug Targets

- *S. aureus*⁶ and *M. tuberculosis*⁷ are deadly human pathogens with well-studied metabolic networks
- CONGA objective: maximize difference in biomass

⁶Becker and Palsson, BMC Microbio 2005

⁷Jamshidi and Palsson, BMC Sys Bio 2007



Identifying Novel Drug Targets

- *S. aureus*⁶ and *M. tuberculosis*⁷ are deadly human pathogens with well-studied metabolic networks
- CONGA objective: maximize difference in biomass
- Solutions will point to uniquely essential metabolic reactions

⁶Becker and Palsson, BMC Microbio 2005

⁷Jamshidi and Palsson, BMC Sys Bio 2007



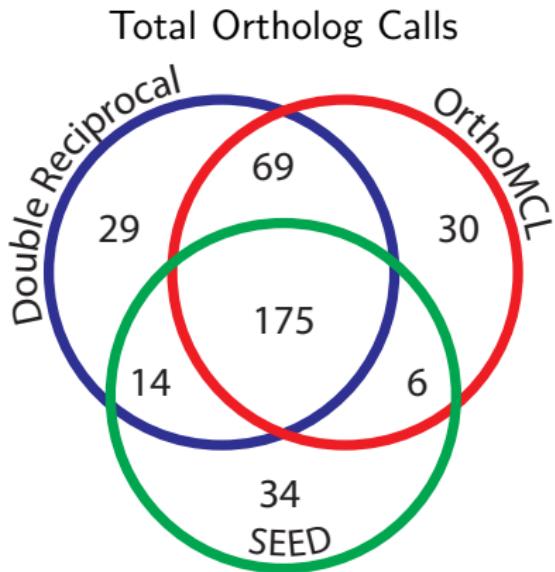
Identifying Novel Drug Targets

- *S. aureus*⁶ and *M. tuberculosis*⁷ are deadly human pathogens with well-studied metabolic networks
- CONGA objective: maximize difference in biomass
- Solutions will point to uniquely essential metabolic reactions
- Unique reactions may be targets of new antibiotics

⁶Becker and Palsson, BMC Microbio 2005

⁷Jamshidi and Palsson, BMC Sys Bio 2007

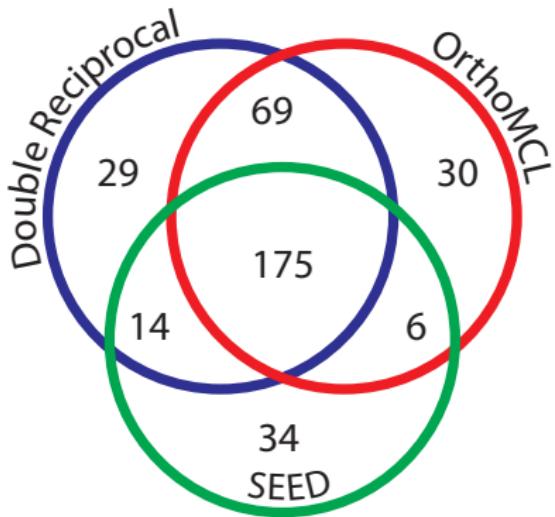
How To Assign Orthologs



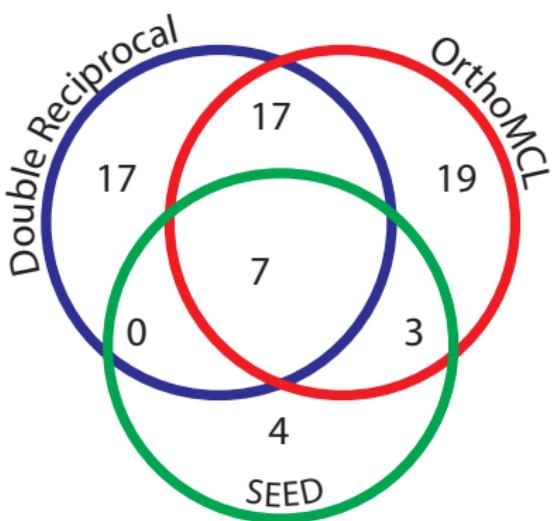
- BLAST and OrthoMCL call the most orthologs

How To Assign Orthologs

Total Ortholog Calls

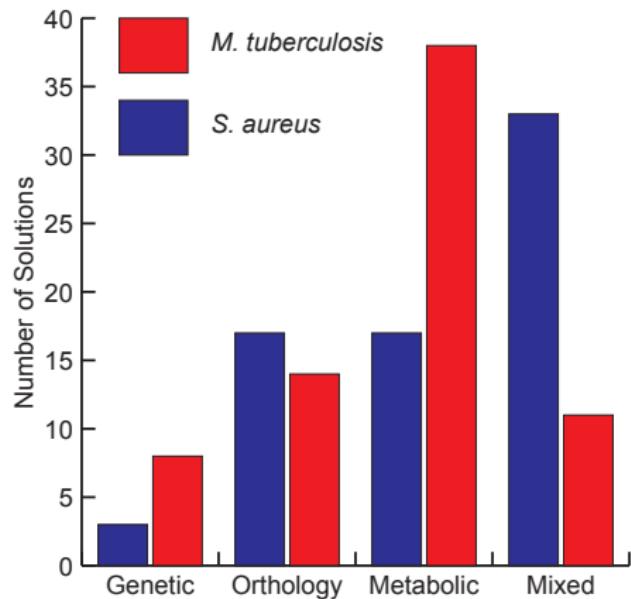


Incorrect Ortholog Calls



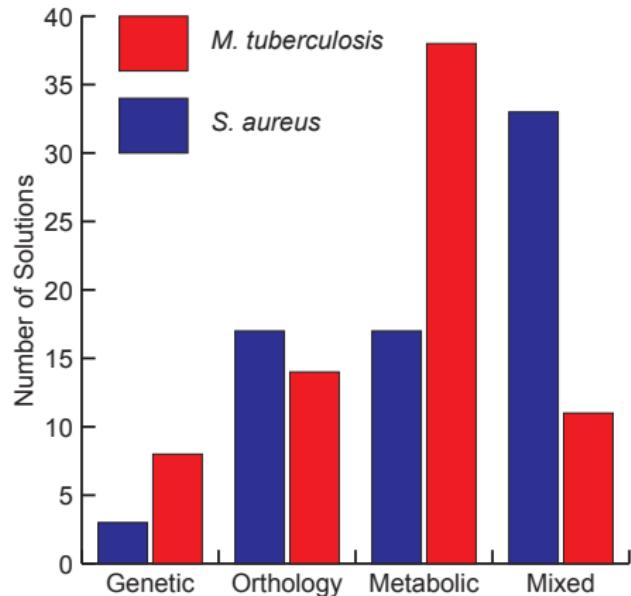
- BLAST and OrthoMCL call the most orthologs
- SEED makes the fewest incorrect calls

Types of CONGA Solutions



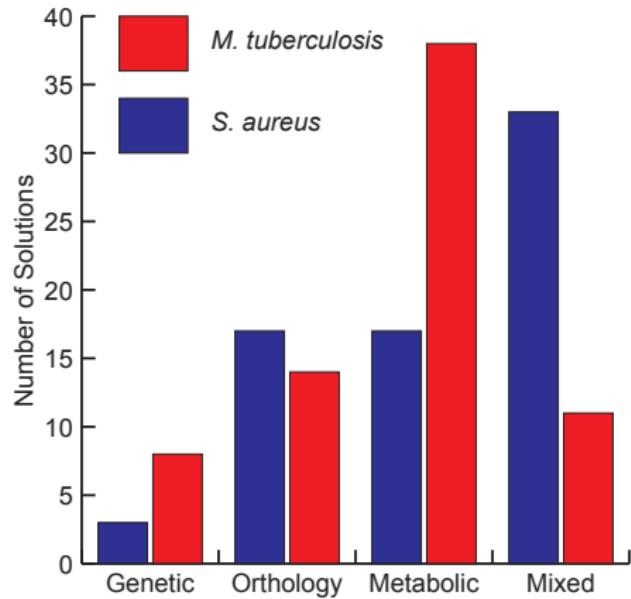
- Genetic. Gene-protein-reaction annotations differ between organisms.

Types of CONGA Solutions



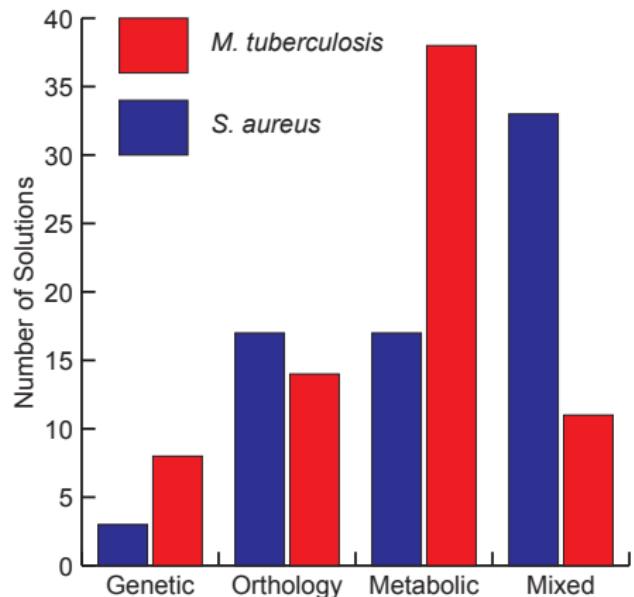
- Genetic. Gene-protein-reaction annotations differ between organisms.
- Orthology. Genes encoding proteins with the same function do not get assigned as orthologs.

Types of CONGA Solutions



- Genetic. Gene-protein-reaction annotations differ between organisms.
- Orthology. Genes encoding proteins with the same function do not get assigned as orthologs.
- Metabolic. Differences in reaction capabilities.

Types of CONGA Solutions



- Genetic. Gene-protein-reaction annotations differ between organisms.
- Orthology. Genes encoding proteins with the same function do not get assigned as orthologs.
- Metabolic. Differences in reaction capabilities.
- Mixed. More than one of the above.

Metabolic Differences Point to Potential Drug Targets

- Consensus set of minimal targets from a variety of conditions

Metabolic Differences Point to Potential Drug Targets

- Consensus set of minimal targets from a variety of conditions
- Orthology calls correctly identify cell walls as being structurally unique

Metabolic Differences Point to Potential Drug Targets

- Consensus set of minimal targets from a variety of conditions
- Orthology calls correctly identify cell walls as being structurally unique
- Identify known antibiotic targets

Metabolic Differences Point to Potential Drug Targets

- Consensus set of minimal targets from a variety of conditions
- Orthology calls correctly identify cell walls as being structurally unique
- Identify known antibiotic targets
- Propose new targets
 - Unique proteins in fatty acid synthesis
 - *FabH* in *S. aureus*
 - *TesA* in *M. tuberculosis*
 - Chorismate lyase in *M. tuberculosis* synthesizes glycolipid components
 - Synthesis and scavenging of other biomass precursors



Other Applications

- Developed a metabolic model for *Synechococcus* sp. PCC 7002⁵

	Before CONGA	After CONGA
Genes	542	611
Reactions	491	552

⁵Hamilton and Reed, under review

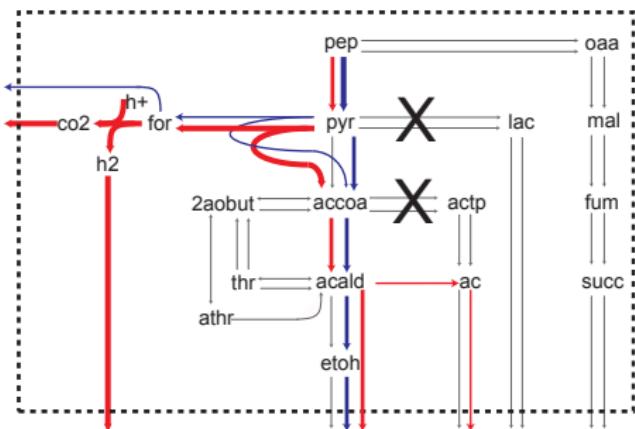
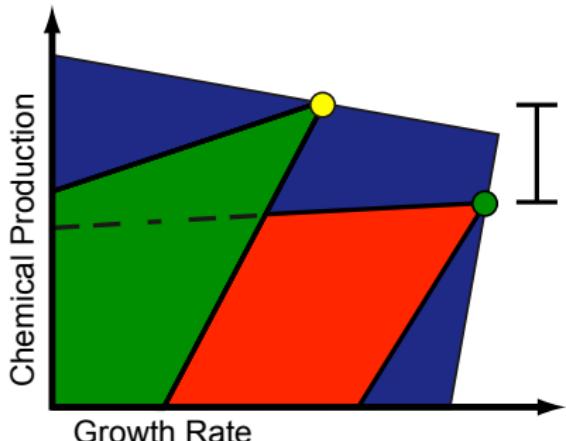
Other Applications

- Developed a metabolic model for *Synechococcus* sp. PCC 7002⁵

Before CONGA After CONGA

Genes	542	611
Reactions	491	552

- Comparisons for metabolic engineering⁵



⁵Hamilton and Reed, under review



Conclusions

- Developed a mathematical programming method for comparison of genome-scale network models



Conclusions

- Developed a mathematical programming method for comparison of genome-scale network models
 - Identifies structural network differences relevant in a particular context
 - Examine impact of functional differences on organismal fitness
 - Can be applied to other perturbation strategies



Conclusions

- Developed a mathematical programming method for comparison of genome-scale network models
 - Identifies structural network differences relevant in a particular context
 - Examine impact of functional differences on organismal fitness
 - Can be applied to other perturbation strategies
- Motivate a shift from *identifying* to *understanding* impact of differences between organisms

Acknowledgements

Prof. Jennifer Reed

- Dr. Dave Baumler
- Camo Cotten
- Joonhoon Kim
- Wai Kit Ong
- Chris Tervo
- Trang Vu
- Xiaolin Zhang

Funding

- National Science Foundation: Graduate Research Fellowship Program



- Department of Energy: Genomic Science Program

