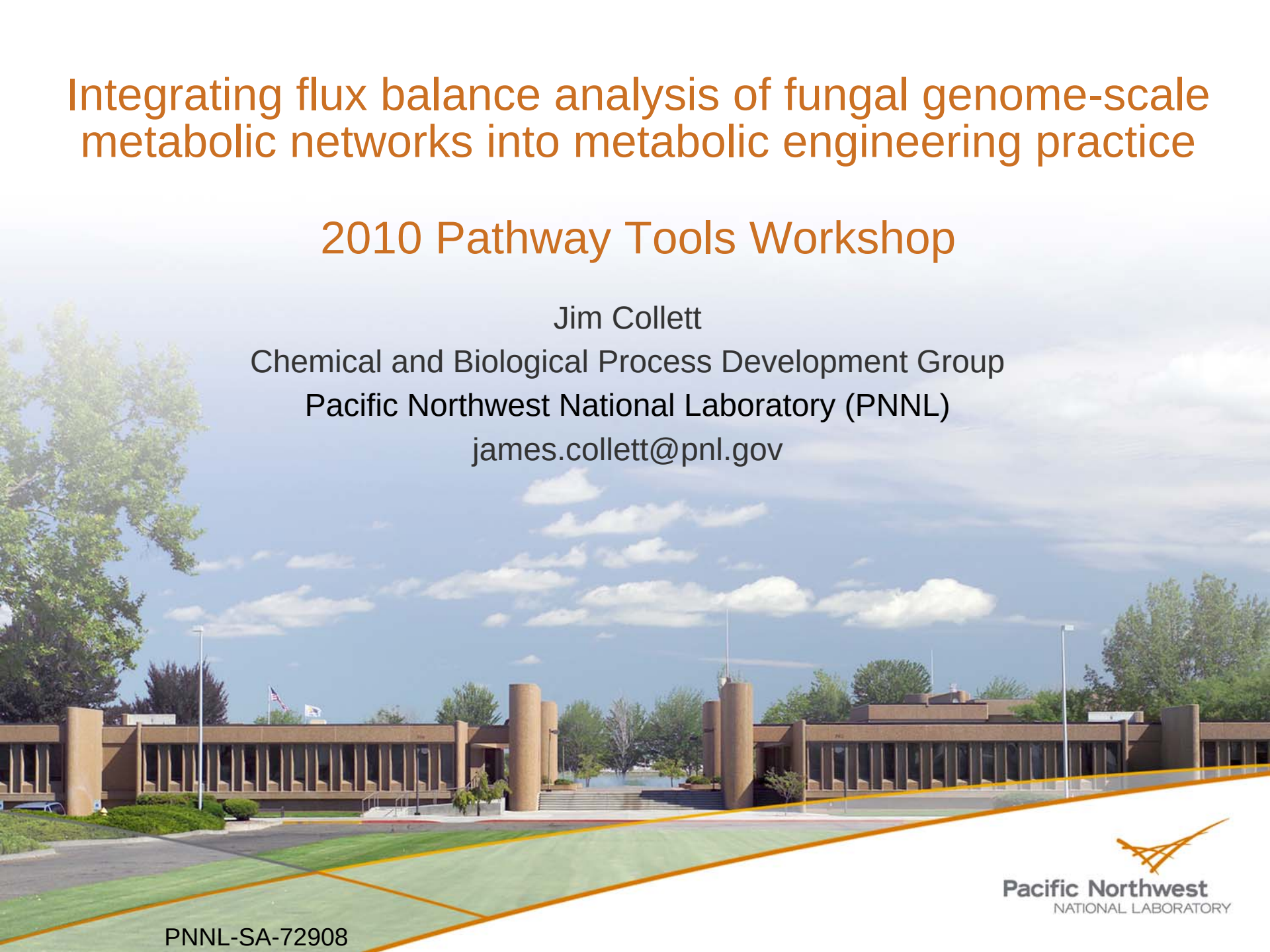
Integrating flux balance analysis of fungal genome-scale metabolic networks into metabolic engineering practice

2010 Pathway Tools Workshop

Jim Collett
Chemical and Biological Process Development Group
Pacific Northwest National Laboratory (PNNL)
james.collett@pnl.gov



Pacific Northwest
NATIONAL LABORATORY

PNNL-SA-72908

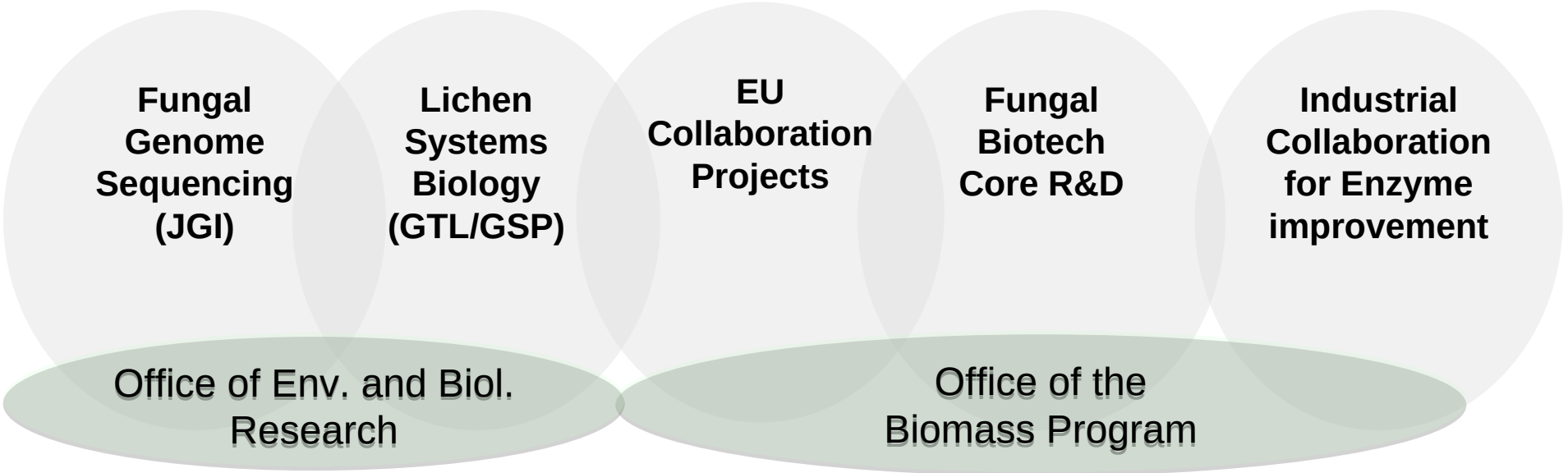
Bioproducts, Sciences, & Engineering Lab at PNNL

WSU & PNNL
BSEL
Bioproducts, Sciences, & Engineering Laboratory

- Thermochemical Conversion
- Biochemical Conversion
- Catalysis and Separations

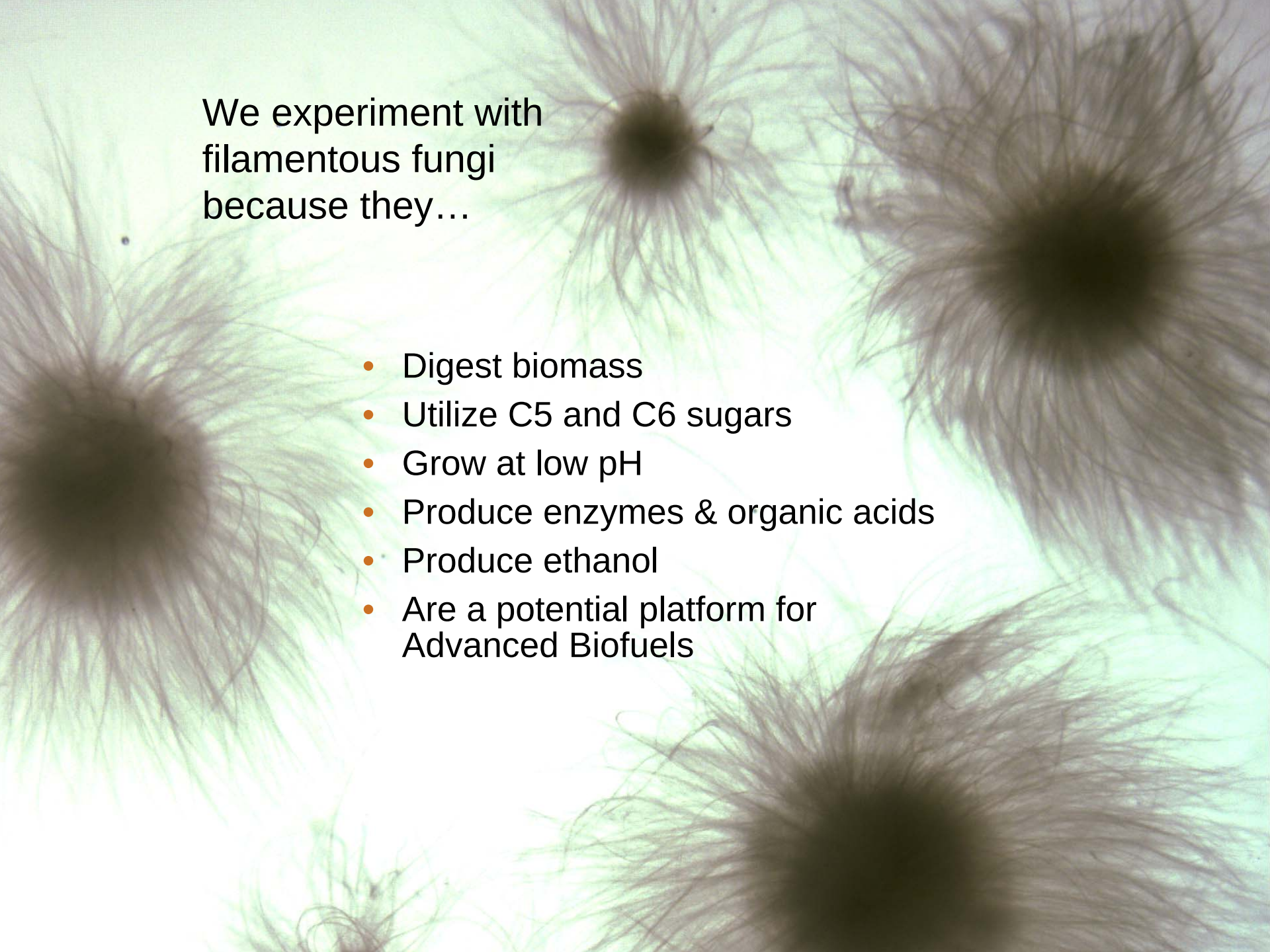
Pacific Northwest
NATIONAL LABORATORY

PNNL fungal research funded by the DOE



Basic Research ⟷ Applied Research

Pacific Northwest NATIONAL LABORATORY



We experiment with filamentous fungi because they…

- Digest biomass
- Utilize C5 and C6 sugars
- Grow at low pH
- Produce enzymes & organic acids
- Produce ethanol
- Are a potential platform for Advanced Biofuels

PNNL/JGI Fungal Genome Sequencing Projects

Aspergillus aculeatus
Aspergillus brasiliensis
Aspergillus carbonarius (2)
Aspergillus niger
Aspergillus tubingensis
Catenaria anguillulae
Cochliobolus heterostrophus
Coemansia reversa
Conidiobolus coronatus
Cryphonectria parasitica
Gonapodya sp.
Neurospora crassa
Orbilia auricolor
Orpinomyces sp.
Phycomyces blakesleeanus
Piromyces sp.
Tremella mesenterica
Trichoderma atroviride
Trichoderma reesei
Trichoderma reesei

Blue = PGDB and curation underway

JGI genome-to-PFF pipeline built by Sebastian Jaramillo-Riveri

Pacific Northwest
NATIONAL LABORATORY

Fungal Genomics Core Research Projects

Genomics: Improved transformation for *A. niger and T. reesei.* Analysis *of A. niger* polyketide synthase (PKS) genes. SNV analysis of highly mutagnenized, cellulse overproducing *T. reesei* strains.

Proteomics: Analysis of *A. niger* mutant strains using an Orbitrap mass spectrometer.

Hyper-productivity and consolidated bioprocesses: Itaconic acid production in *A. terreus.*

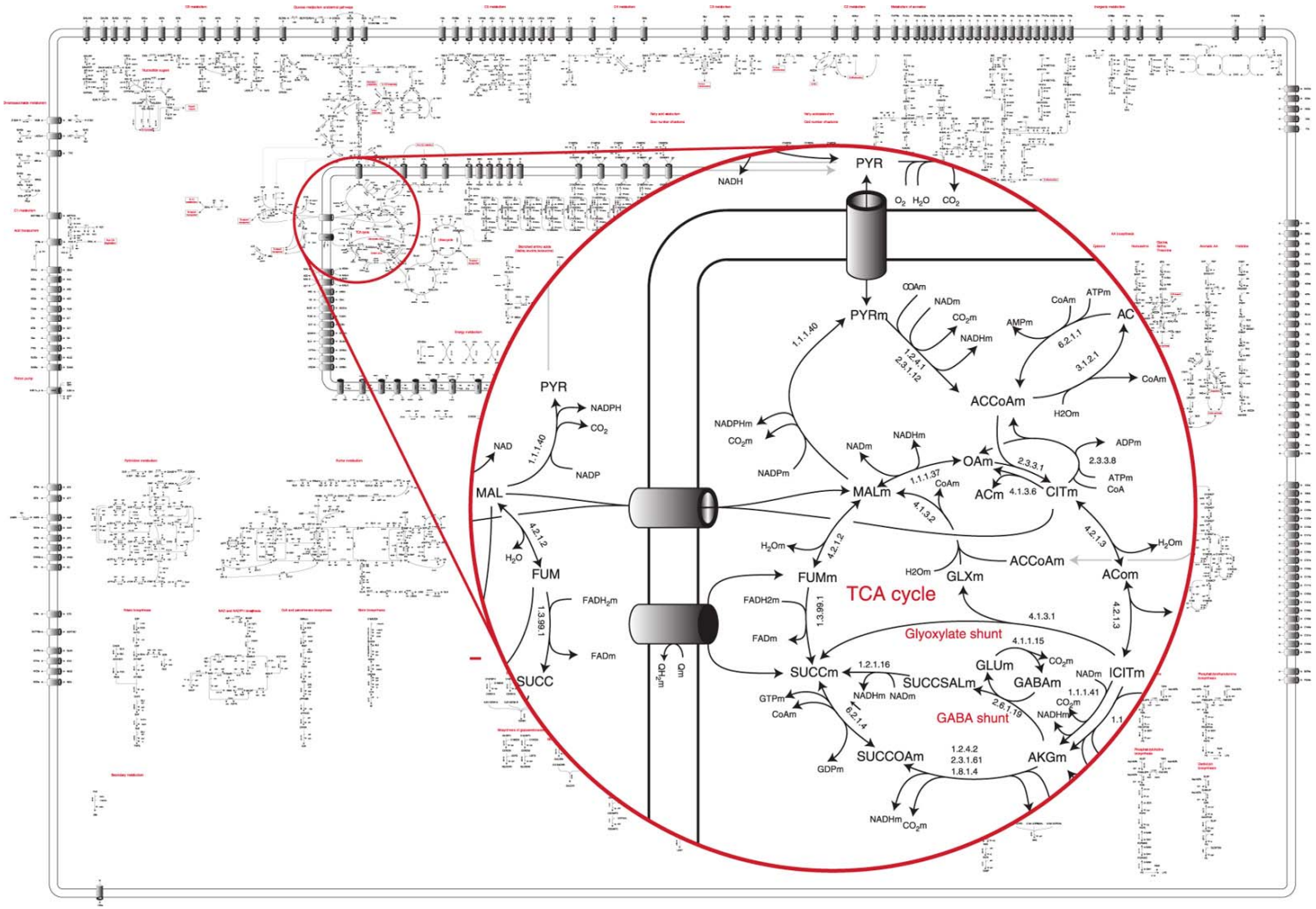
Pentose utilization in filamentous fungal: Study of pentose utilization during *A. oryzae fermentation*.

Alternative renewable fuels from fungi: Polyketide, isoprenoid and fatty acid biosynthesis for advanced hydrocarbon biofuels. NMR analysis of candidate biofuel precursor strains.

Metabolic Process Modeling and Data Integration

Pacific Northwest
NATIONAL LABORATORY

Aspergillus niger genome scale metabloic model from the Nielsen group at DTU/Chalmers



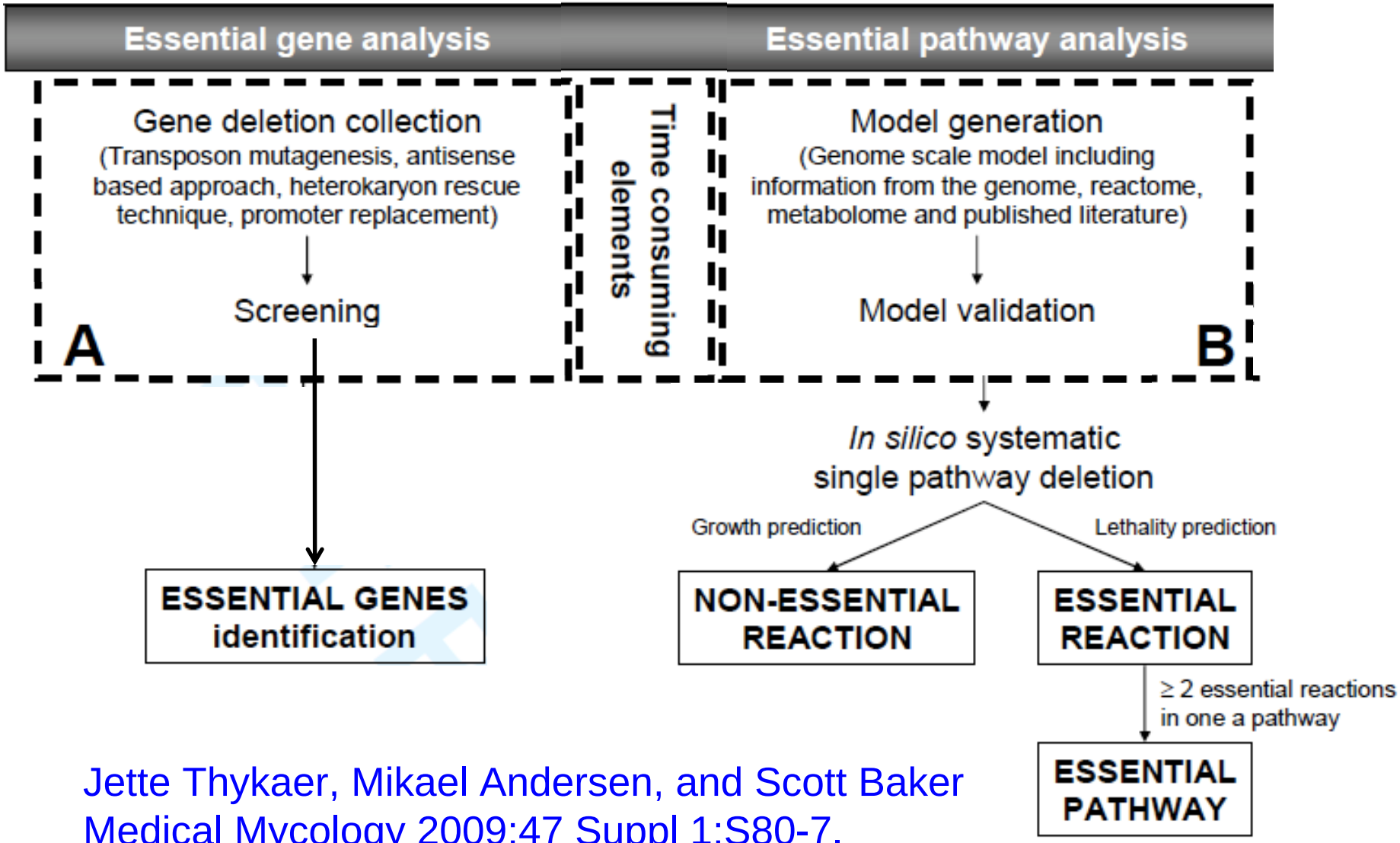
From review of 371 articles

Features:
- 871 ORFs
- 1045 metabolites
- 1190 reactions
- Mitochondrial Compartment

Mikael Rørdam Andersen,[1*] Michael Lynge Nielsen,[1] and Jens Nielsen[1a]
Mol Syst Biol. 2008; 4: 178.

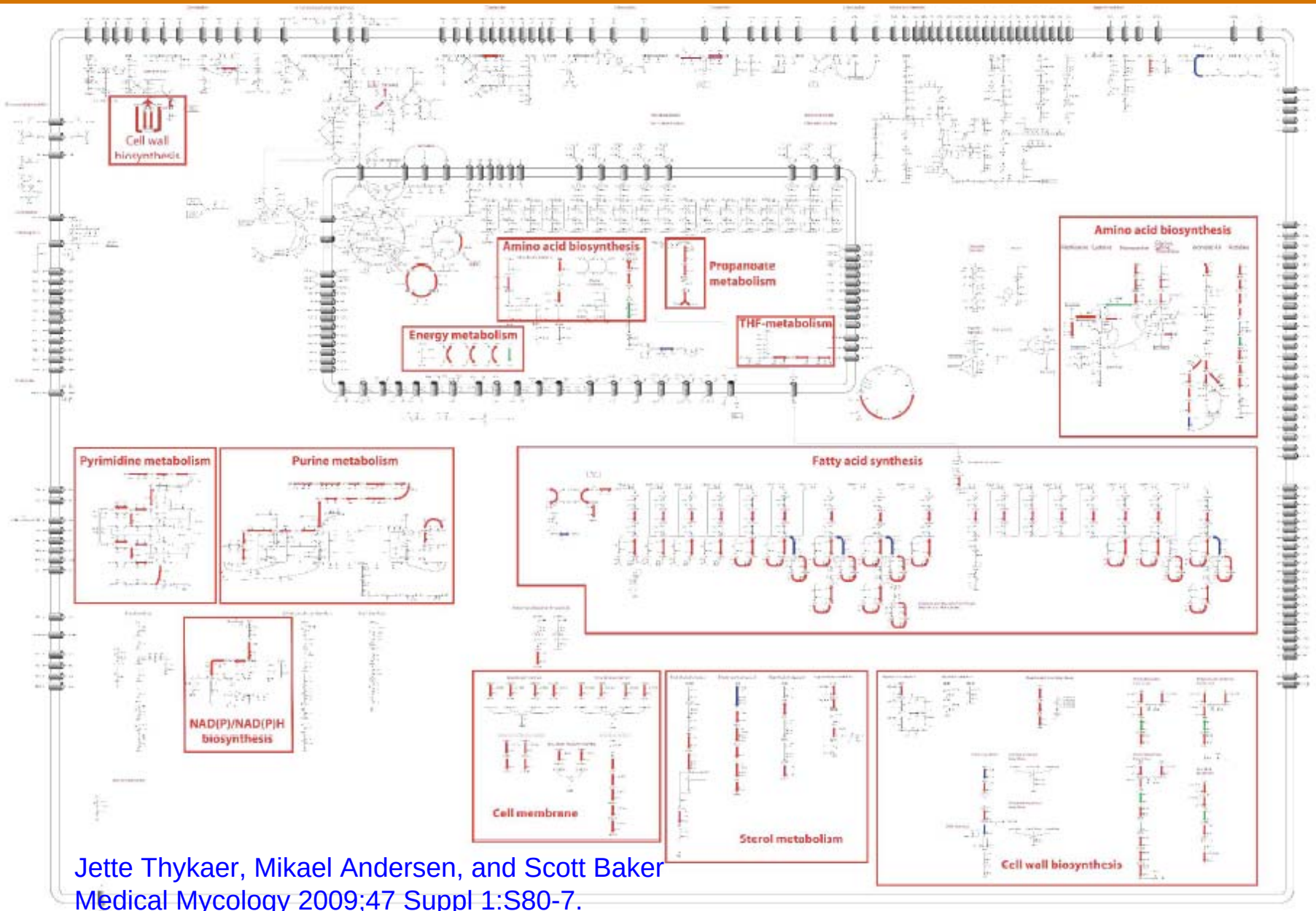
Pacific Northwest NATIONAL LABORATORY

Using Flux Balance Analysis (FBA) in *A. niger* to predict potential antifungal targets in *Aspergillus fumigatus*



Jette Thykaer, Mikael Andersen, and Scott Baker
Medical Mycology 2009;47 Suppl 1:S80-7.

A. niger genes predicted to be essential by FBA were blasted against the *A. fumigatus* and *Homo sapiens* genomes to find possible orthologs



Jette Thykaer, Mikael Andersen, and Scott Baker
Medical Mycology 2009;47 Suppl 1:S80-7.

Predicted antifungal drug targets

Table 1 Potential antifungal targets in *A. fumigatus*

EC no.	Essential enzyme	*A. fumgatus* gene ID	Additonal info	
Amino acids biosynthesis				*
5.4.99.5	Chorismate mutase	Afu5g13130	Aromatic	
2.4.2.18	Anthranilate phosphoribosyl transferase	Afu4g11980	Aromatic	
2.1.1.17	Phosphatidylethanolamine N-methyltransferase	Afu2g15970	Aromatic	
4.2.1.51	Prephenate dehydratase	Afu5g05690	Aromatic	
2.5.1.54	3-deoxy-7-phosphoheptulonate synthase	Afu1g02110	Aromatic	
4.2.1.19	Imidazoleglycerol-phosphate dehydratase	Afu6g04700	His	
3.5.4.19	phosphoribosyl-AMP cyclohydrolase	Afu1g14570	His	
3.1.3.15	Histidinol phosphatase	Afu4g04030	His	
1.1.1.23	Histidinol dehydrogenase	Afu1g17660	His	
2.4.2.17	ATP phosphoribosyltransferase	Afu7g04500	His	
1.1.1.3	Homoserine dehydrogenase	Afu3g11640	Lys, Gly, Ser, Thr	
2.7.2.4	Aspartate kinase	Afu5g05590	Lys, Gly, Ser, Thr	
1.2.1.11	aspartate-semialdehyde dehydrogenase	Afu3g06830	Lys, Gly, Ser, Thr	
2.3.3.14	Homocitrate synthase	Afu4g10460	Lys	
1.1.1.86	Ketol-acid reductoisomerase	Afu3g14490	Val, Leu, Ile	
2.3.1.31	homoserine O-acetyltransferase	Afu5g07210	Met	
2.1.1.13	Methionine synthase	Afu4g07360	Met	
Propanoate metabolism				
4.1.3.30	2-methylisocitrate lyase	Afu6g02860		
4.2.1.79	2-methylcitrate hydrolyase	Afu6g03730		
Fatty acid biosynthesis				*
2.1.1.71	Methylene-fatty-acyl-phospholipid synthase	Afu1g09050		
2.3.1.38	[ACP]acetyltransferase	Afu3g04220		
2.3.1.85	Fatty-acid synthase	Afu3g04210		
Pyrimidine and purine metabolism				
1.8.1.9	Thioredoxin reductase	Afu6g09740		
2.4.2.10	Orotate phosphoribosyltransferase 1	Afu2g11290		
4.1.1.21	Phosphoribosylaminoimidazole carboxylase	Afu4g12600		
Cell wall biosynthesis				*
2.4.1.34	1,3-beta-Glucan synthase	Afu5g05770		
3.1.3.12	Trehalose-phosphatase	Afu3g05650		
Sterol biosynthesis				*
2.5.1.21	Squalene synthase	Afu7g01220		
2.7.4.2	Phosphomevalonate kinase	Afu5g10680		
NADH/NADPH metabolism				
2.4.2.19	Nicotinate mononucleotide pyrophosphorylase	Afu3g05730		
Urea and metabolism of amino groups				
2.3.1.35	Glutamate N-acetyltransferase	Afu5g08120		
Manitol biosynthesis				
1.1.1.17	Mannitol-1-phosphate 5-dehydrogenase	Afu2g10660		

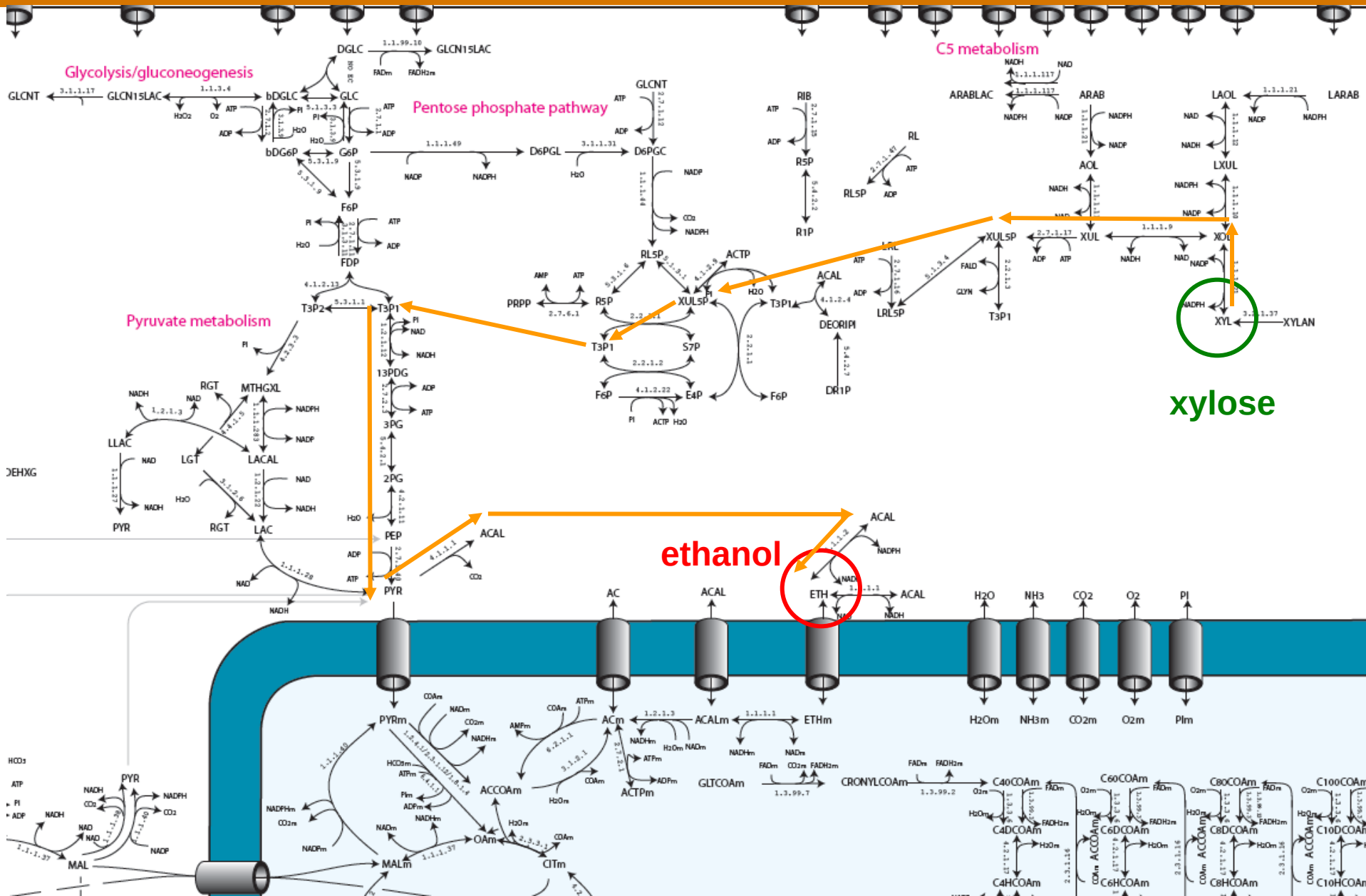
Jette Thykaer, Mikael Andersen, and Scott Baker
Medical Mycology 2009;47 Suppl 1:S80-7.

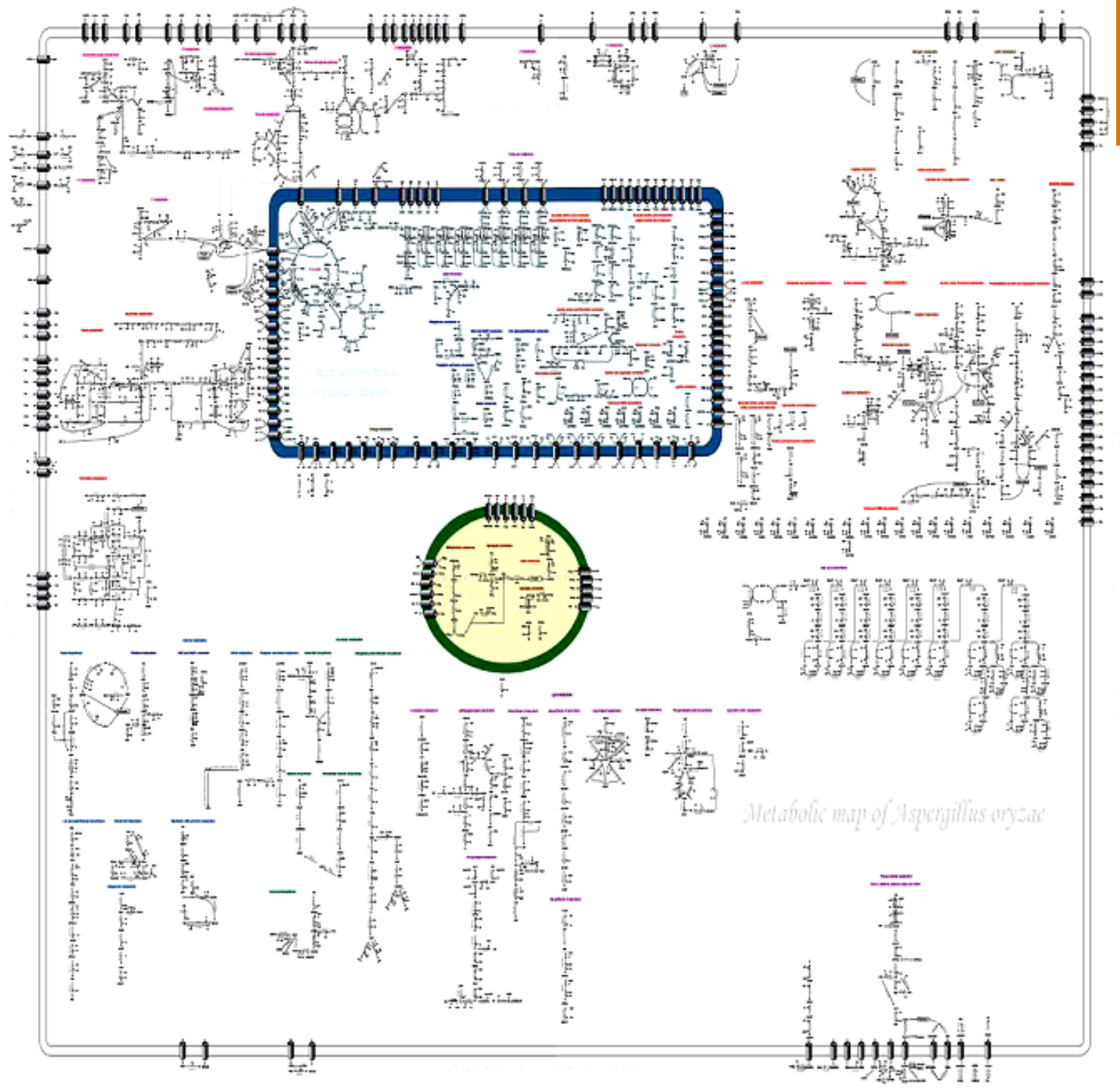
Ethanol overproduction by *Aspergillus oryzae* as a model for pentose utilization in consolidated biofuel production

- *A. oryzae* has been used for over 1000 years to saccharify rice for sake brewing.
- It's the national fungus of Japan!

Pacific Northwest
NATIONAL LABORATORY

Flux balance analysis (FBA) to optimize ethanol production in *A. oryzae*





Aspergillus oryzae RIB 40

Genome-scale metabolic network model

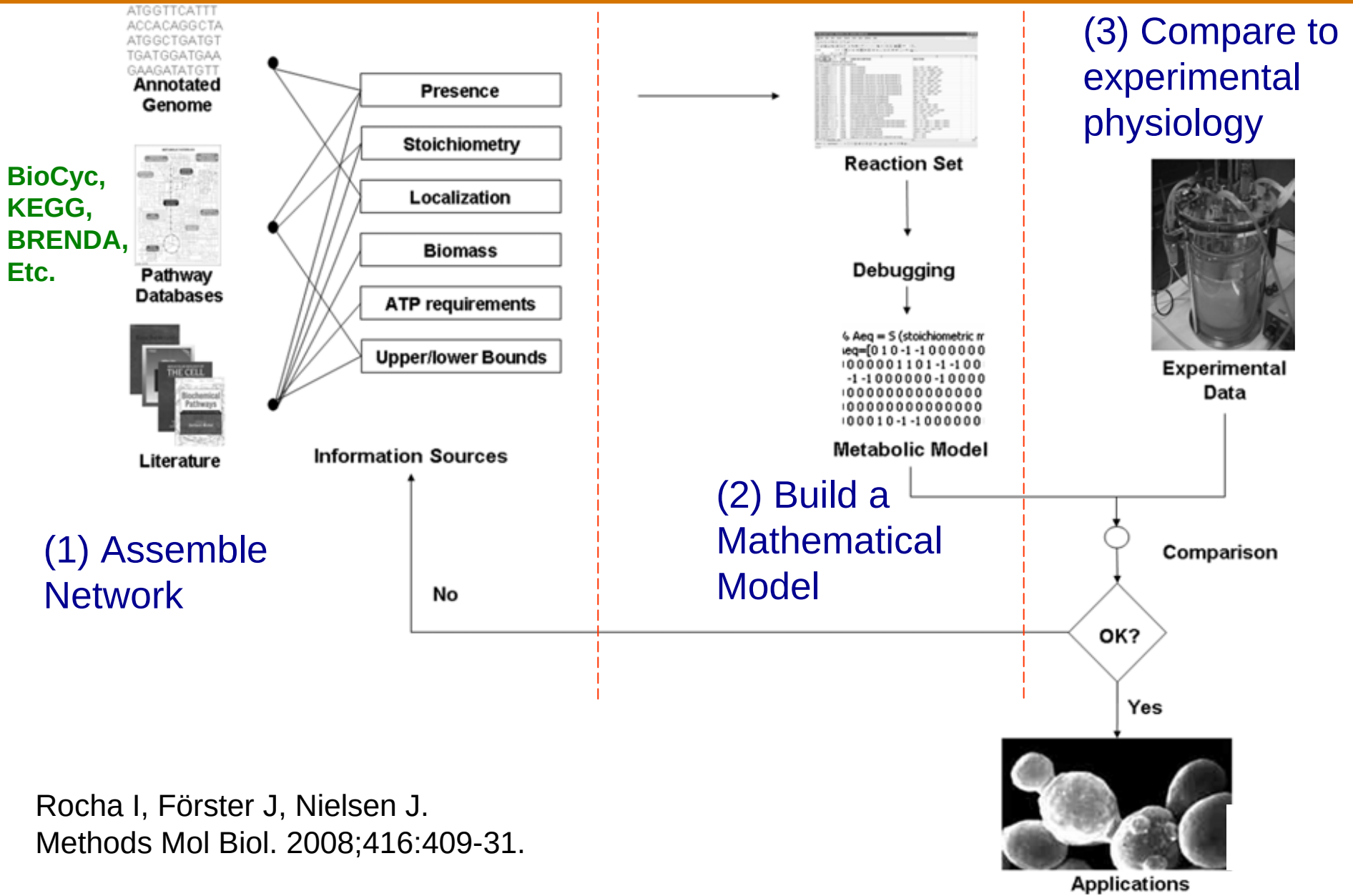
Nielsen group, Chlamers/DTU

- 729 enzymes
- 1314 genes
- 1073 metabolites
- 1846 reactions
- Mitochondrial & Peroxisome Compartments

- Vongsangnak, *et al.* BMC Genomics 2008

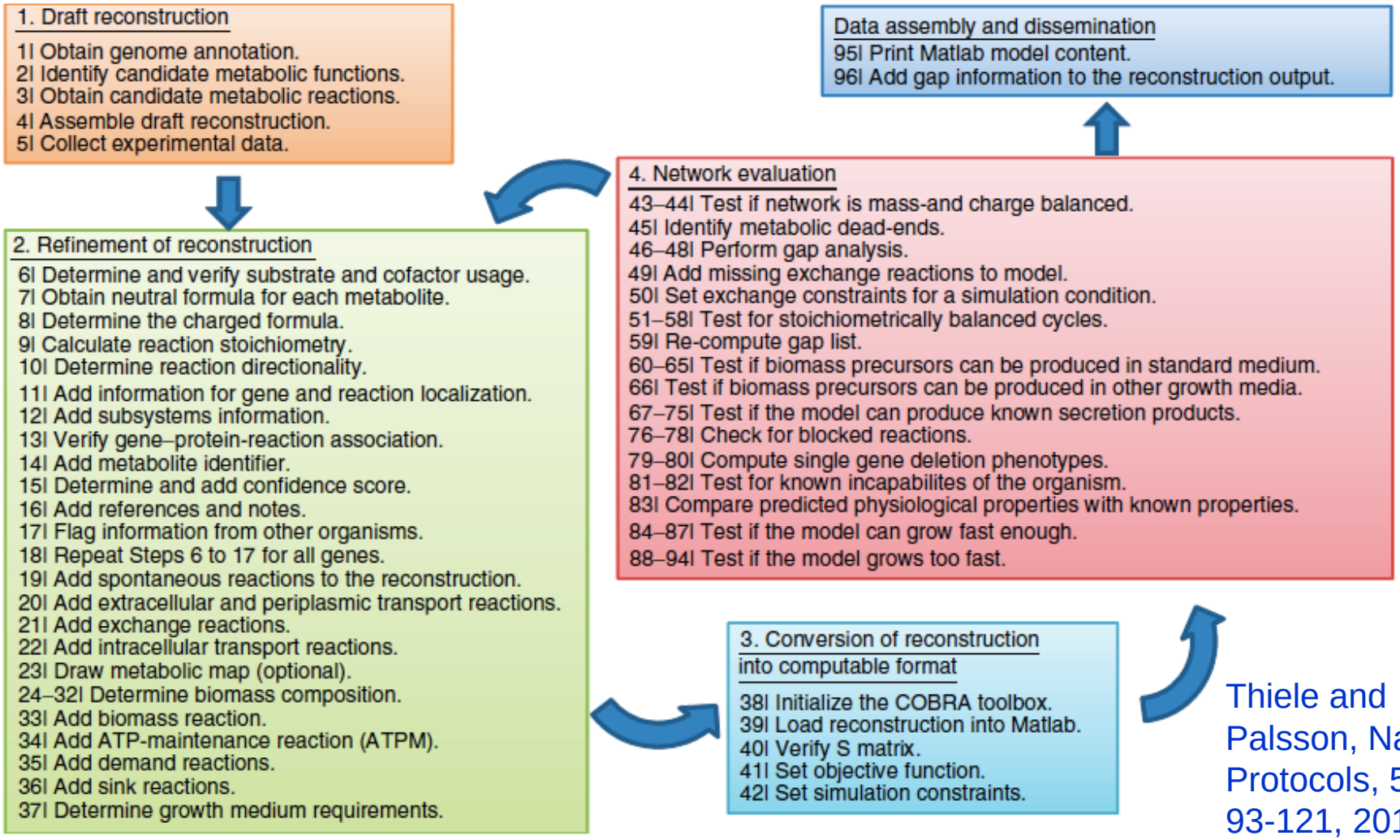
Pacific Northwest NATIONAL LABORATORY

Stoichiometric network reconstruction and analysis



Rocha I, Förster J, Nielsen J.
Methods Mol Biol. 2008;416:409-31.

Stoichiometric network reconstruction and analysis



Thiele and Palsson, Nature Protocols, 5(1): 93-121, 2010.

Figure 1 | Overview of the procedure to iteratively reconstruct metabolic networks. In particular, Stages 2–4 are continuously iterated until model predictions are similar to the phenotypic characteristics of the target organism and/or all experimental data for comparison are exhausted.

Estimated time requirements for constraint-based reconstruction and analysis (COBRA) from Thiele and Palsson

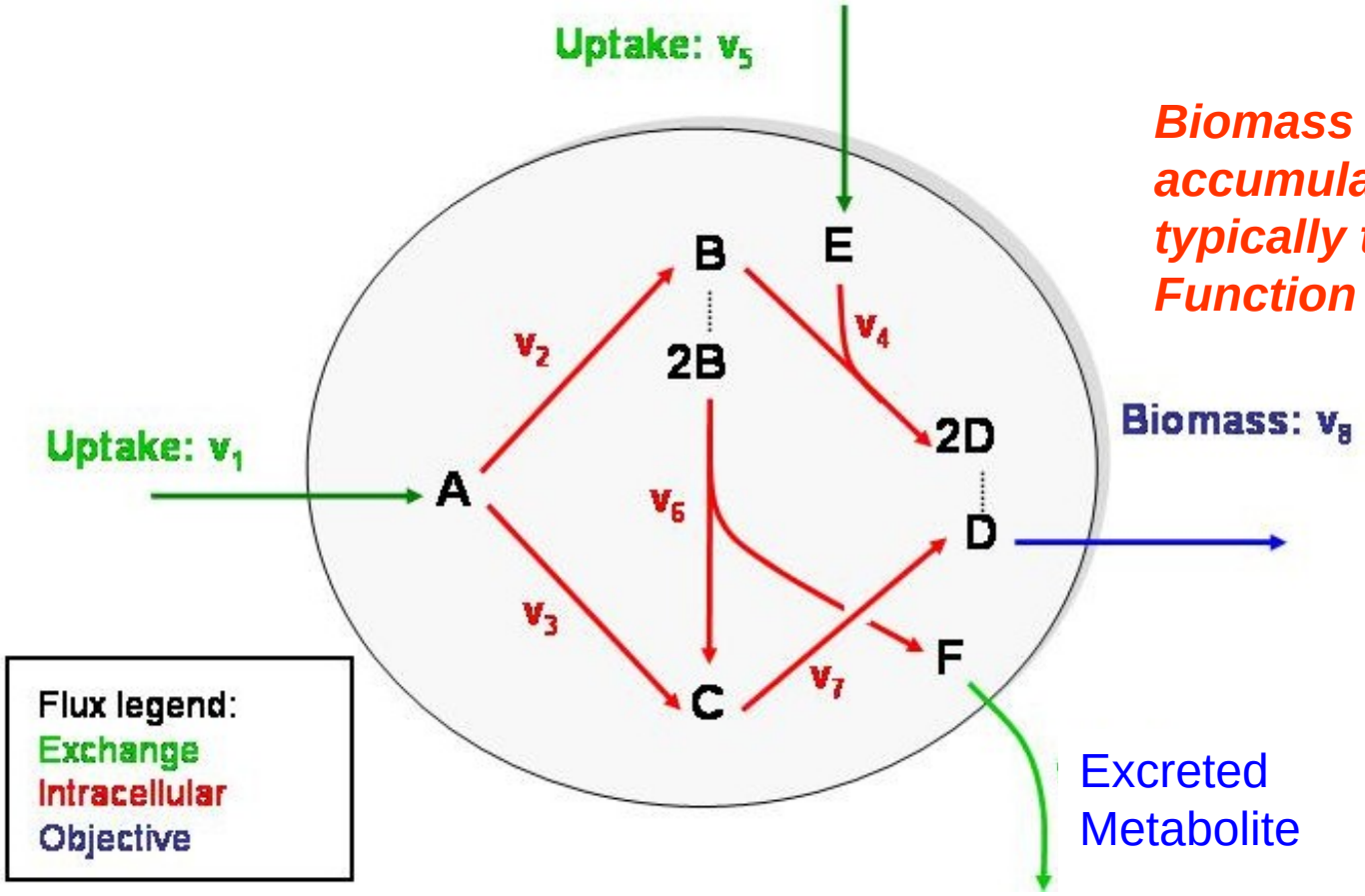
Draft reconstruction	days to weeks
Collect experimental data	ongoing throughout process
Manual reconstruction refinement	months to a year
Determine biomass composition	days to weeks
Mathematical model generation	days to a week
Network evaluation (debugging mode)	week to months
Data assembly and dissemination	days to weeks

Nature Protocols, 5(1): 93-121, 2010.

Pacific Northwest
NATIONAL LABORATORY

Concept of Flux Balance Analysis (FBA)

A steady-state model where all inputs and outputs sum to zero.

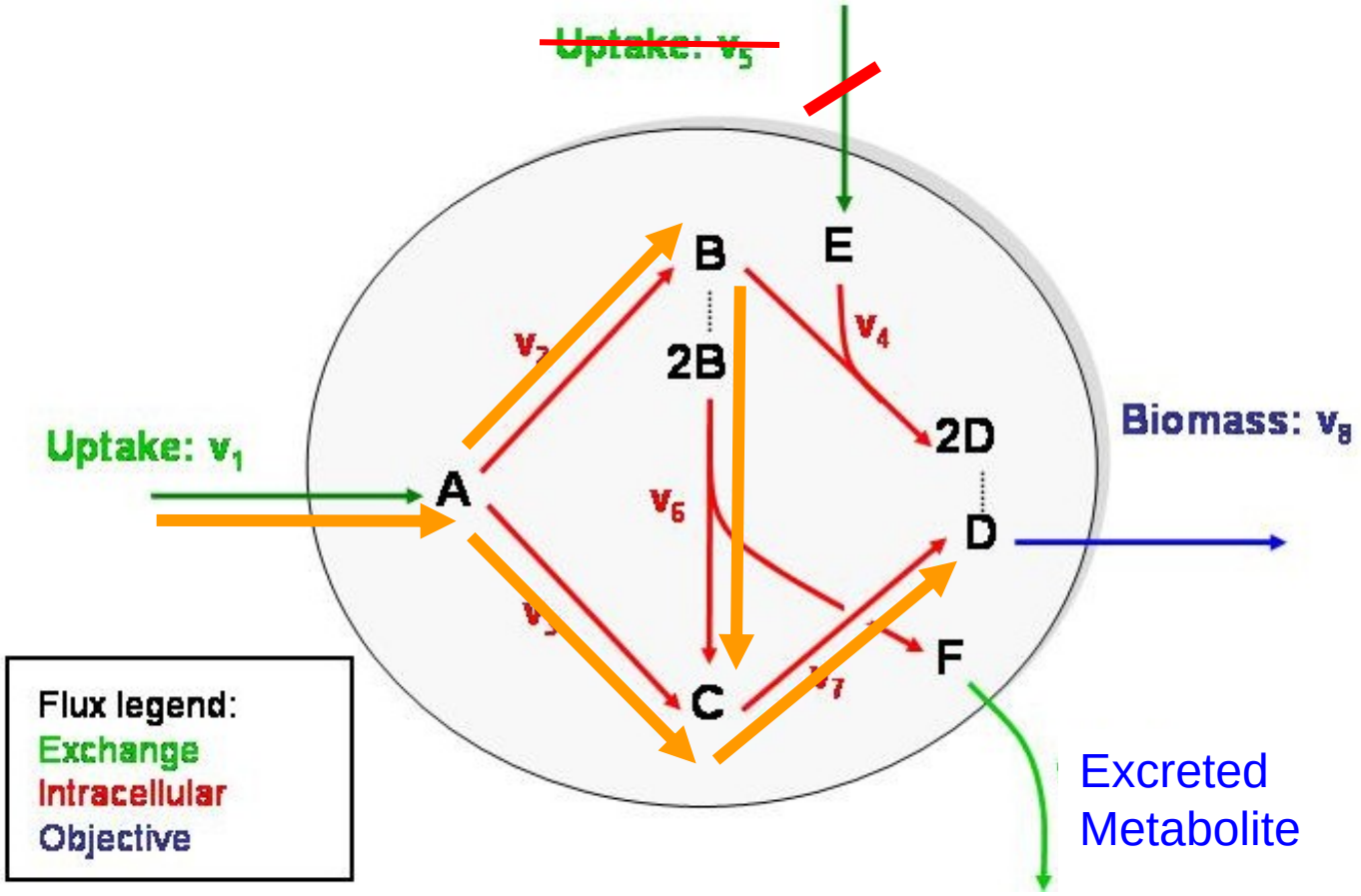


Biomass accumulation is typically the Objective Function for FBA

http://bio.freelogy.org/w/images/1/14/Metabolic-network.JPG

http://bio.freelogy.org/wiki/User:JeremyZucker

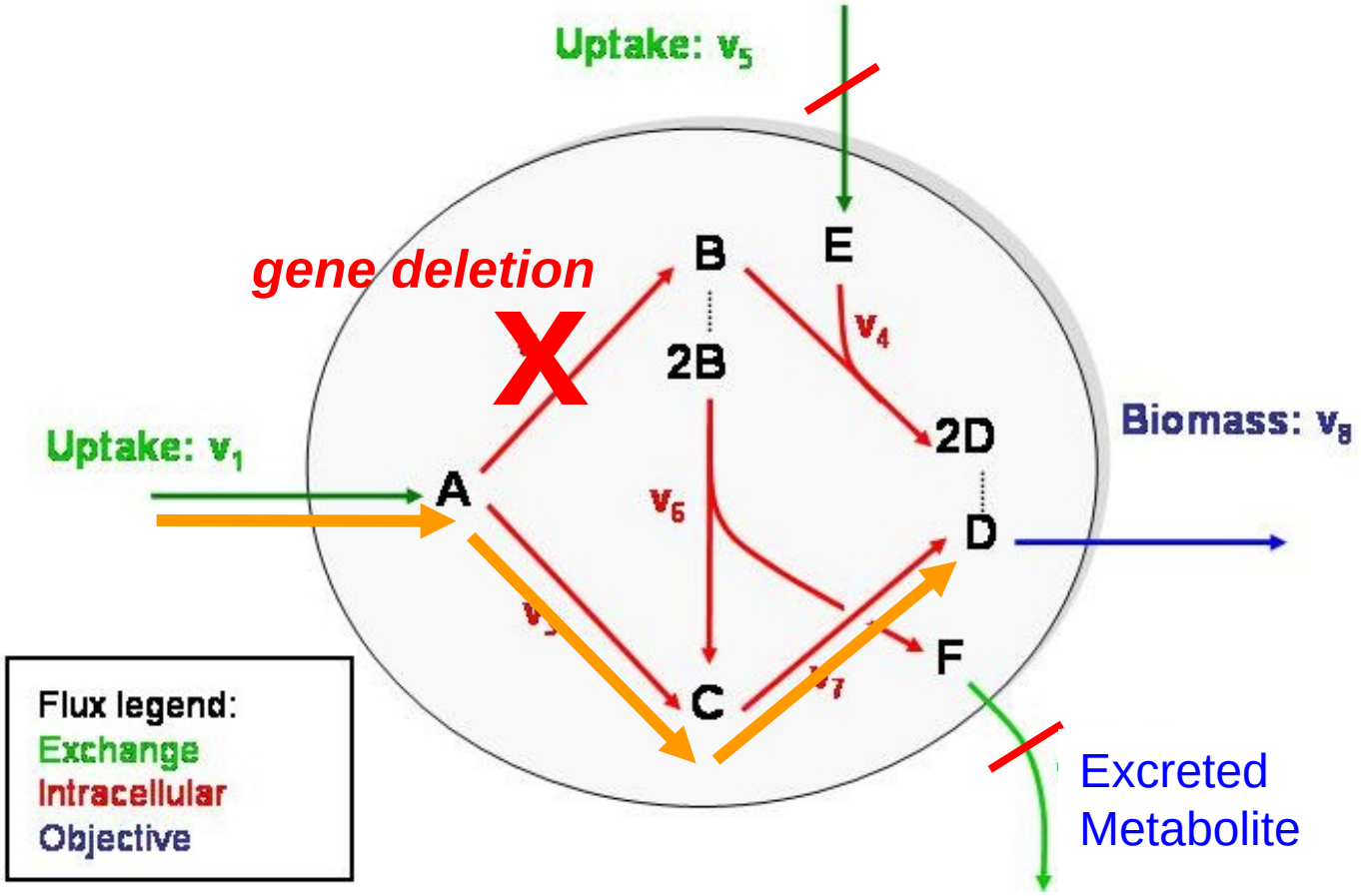
Constraining an uptake flux



http://bio.freelogy.org/w/images/1/14/Metabolic-network.JPG

Pacific Northwest
NATIONAL LABORATORY

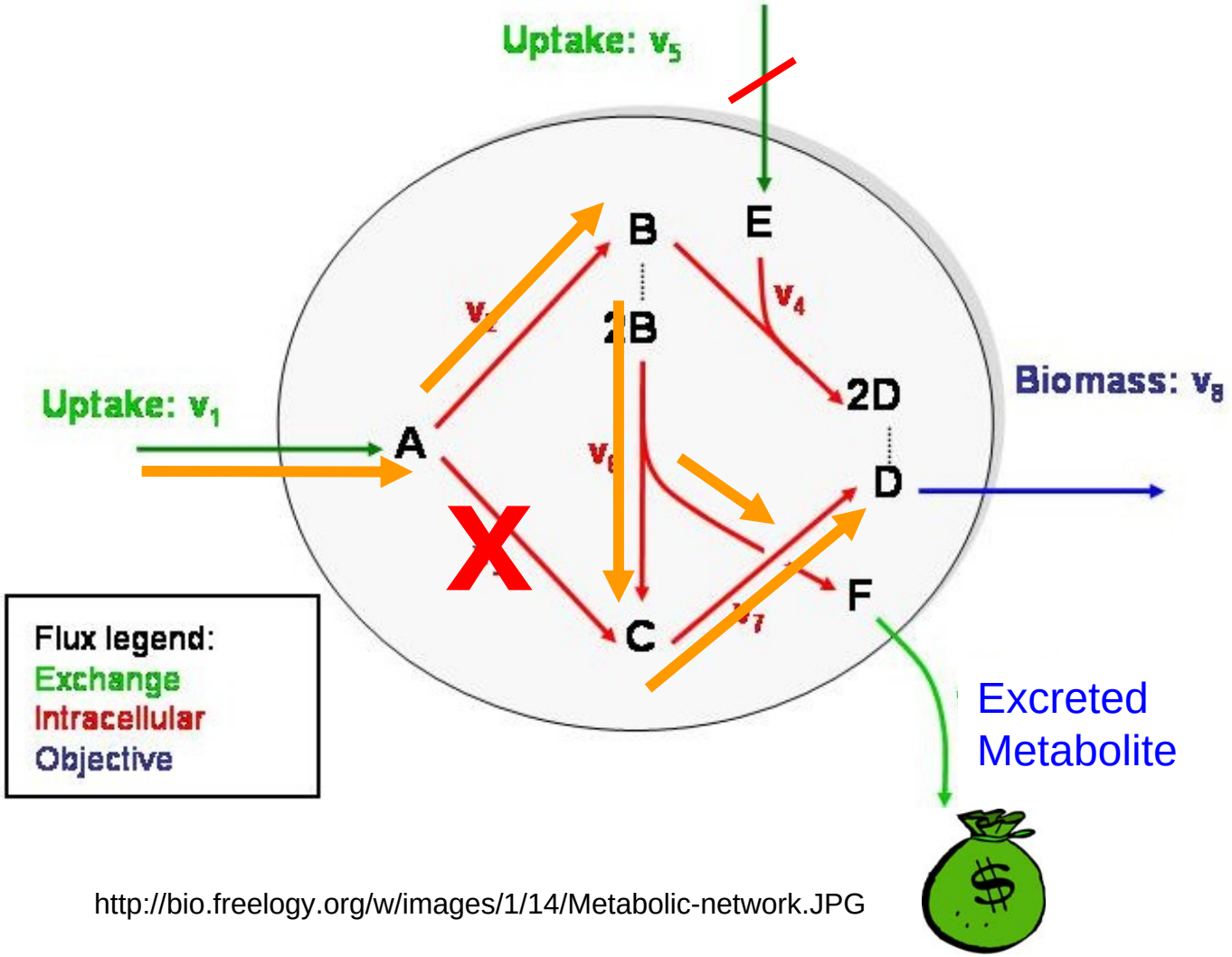
Simulating a gene deletion



http://bio.freelogy.org/w/images/1/14/Metabolic-network.JPG

Pacific Northwest
NATIONAL LABORATORY

Gene deletion to optimize excretion of a specific metabolite



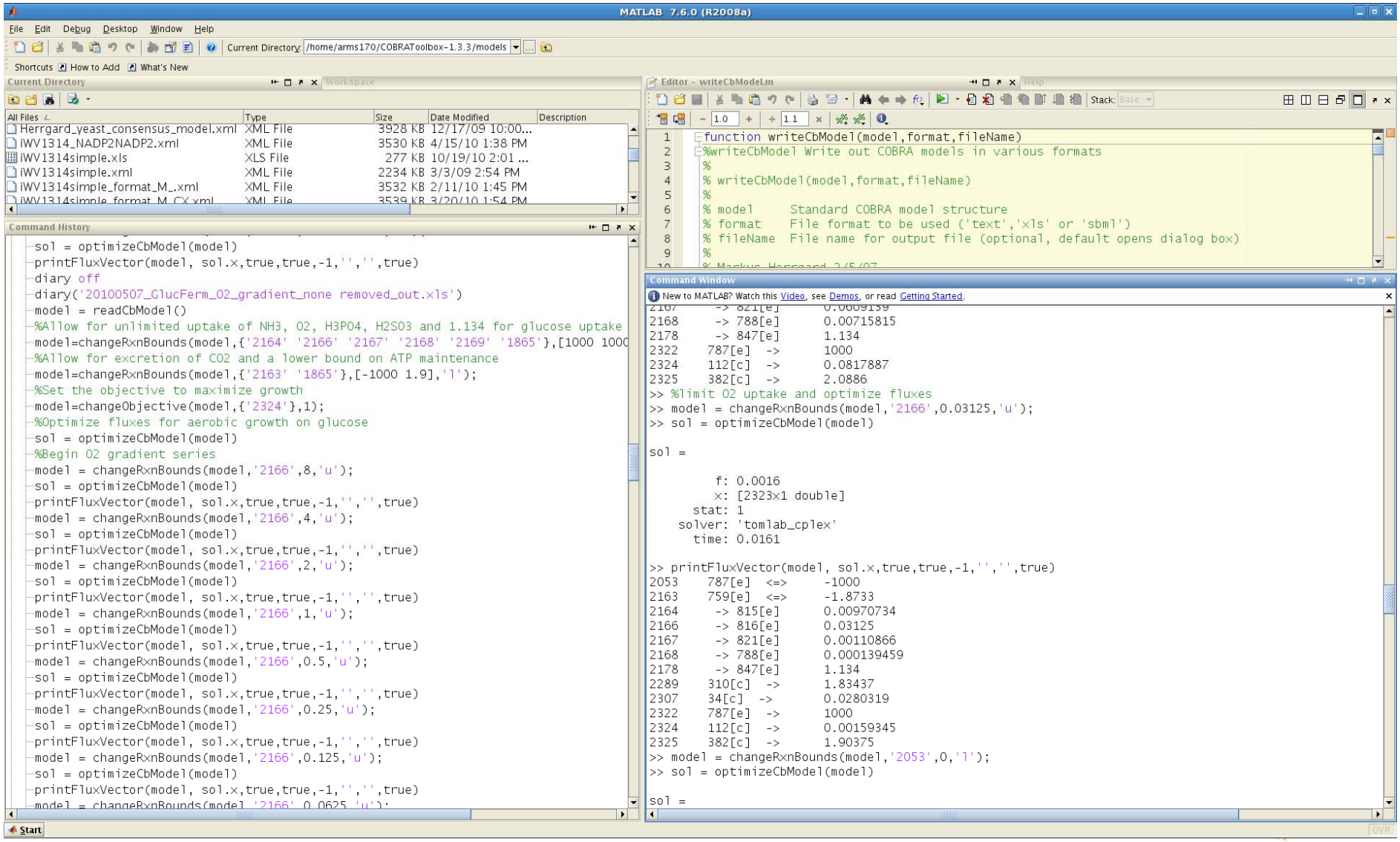
http://bio.freelogy.org/w/images/1/14/Metabolic-network.JPG

Pacific Northwest
NATIONAL LABORATORY

Software packages for FBA and related methods

- COBRA Toolbox (MATLAB)
- CellNetAnalyzer (MATLAB)
- OptFlux (v2.2 Windows; v1.37 Windows, Linux)
- MetaFluxNet (Windows)
- Systems Biology Research Tool (Multi-platform Java)

Using the COBRA Toolbox in MATLAB



Becker SA, *et al*. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. Nature Protocols 2007;2(3):727-38

Pacific Northwest NATIONAL LABORATORY

FBA model structure in COBRA Toolbox/MATLAB

Composed of vectors and matrices for:

- reaction stoichiometry
- genes
- proteins (enzymes)
- Gene-protein-reaction (GPR) associations
- objective function selection
- reaction flux constraints

	GLCt1	HEX1	PGI	PFK	FBP	FBA	TPI	EX_glc
glc-D[e]	−1	0	0	0	0	0	0	−1
glc-D	1	−1	0	0	0	0	0	0
atp	0	−1	0	−1	0	0	0	0
H	0	1	0	1	0	0	0	0
adp	0	1	0	1	0	0	0	0
g6p	0	1	−1	0	0	0	0	0
f6p	0	0	1	−1	1	0	0	0
fdp	0	0	0	1	−1	−1	0	0
pi	0	0	0	0	1	0	0	0
h2o	0	0	0	0	−1	0	0	0
g3p	0	0	0	0	0	1	1	0
dhap	0	0	0	0	0	1	−1	0

First steps of glycolysis pathway

Becker SA, Feist AM, Mo ML, Hannum G, Palsson BØ, Herrgard Mjbased. Nature Protocols 2007;2(3):727-38.

Simulating metabolism under an O_2 uptake gradient to predict optimal ethanol production level in *A. oyrzae*

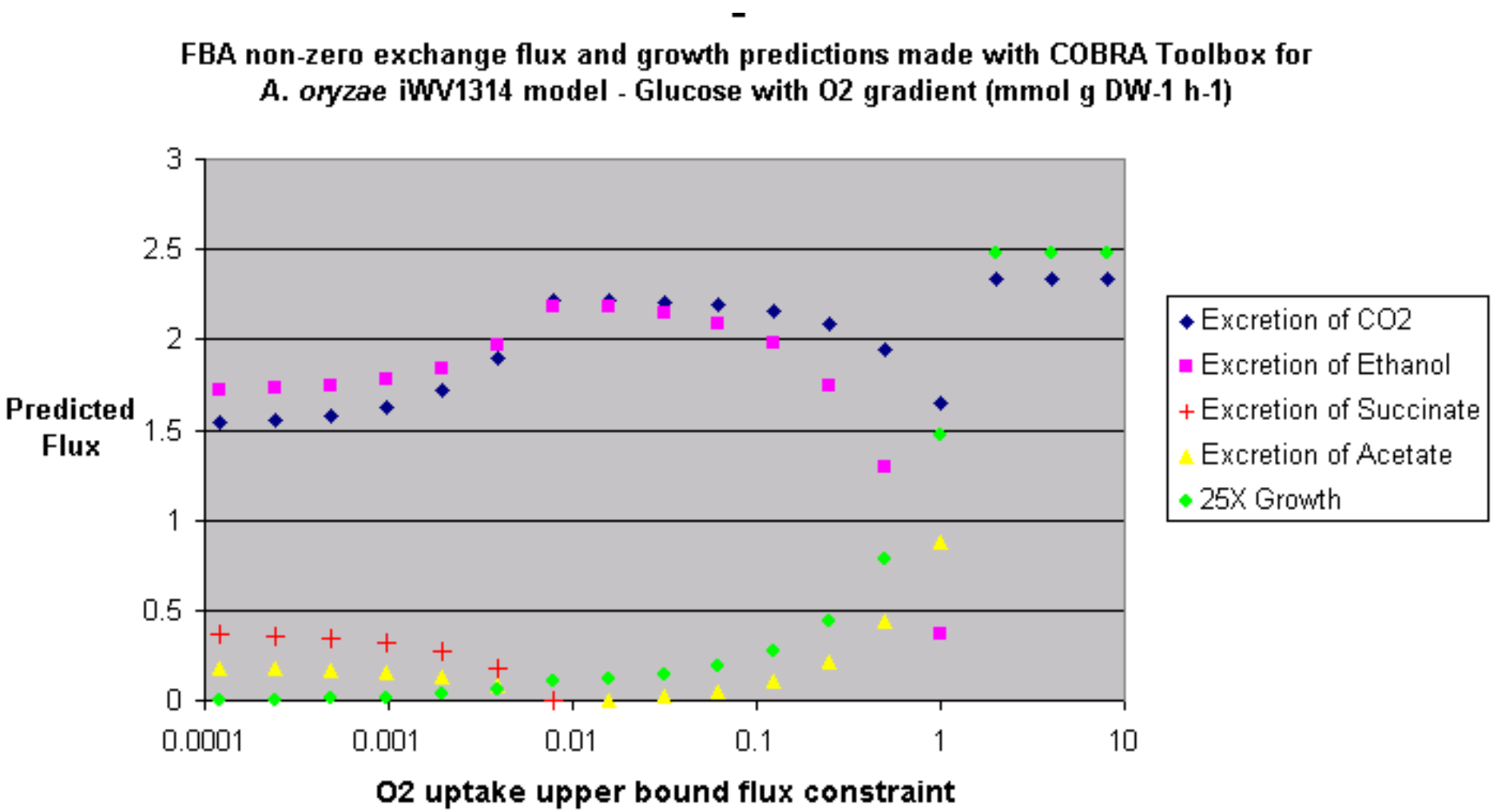
Exchange Flux Constraints (mmol gDW^{-1} hr^{-1})

- NH_3, H_3PO_4, H_2SO_3	Uptake unlimited
- Glucose	Uptake of 1.134
- O_2	Uptake stepwise gradient from 0.0001 to 10
- ATP	Maintain intracellular 1.9

Objective Function

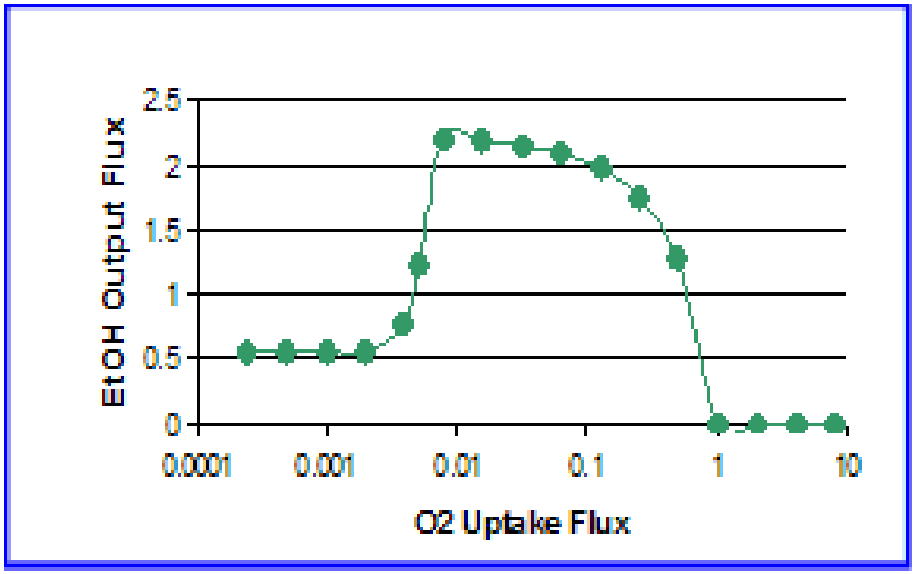
Set as “Growth” to maximize combined fluxes for generating cell biomass constituents (DNA, RNA, amino acids, lipids, carbohydrates, etc.)

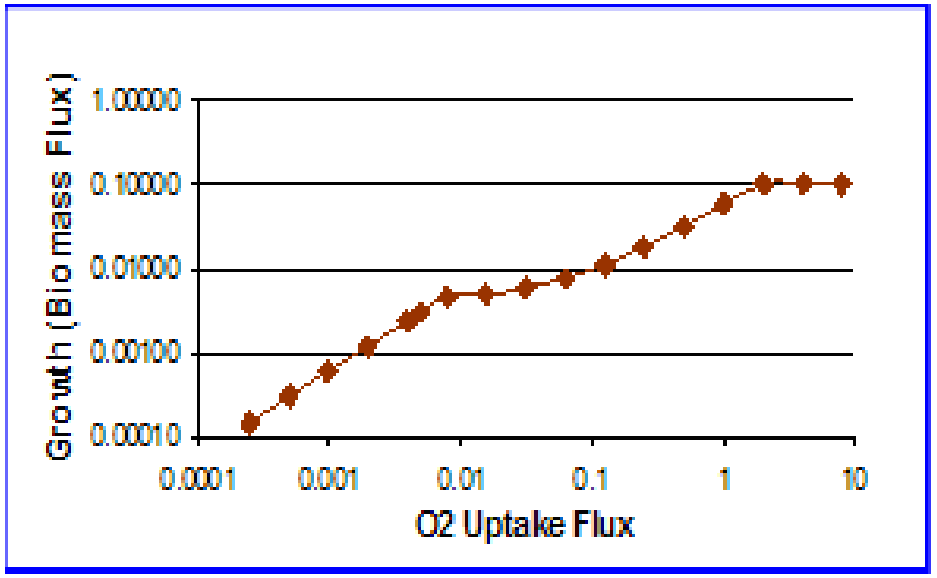
FBA simulation of *A. oryzae* fermentation on glucose



Pacific Northwest
NATIONAL LABORATORY

Predicted ethanol excretion maximum correlates with a plateau in growth in FBA simulation

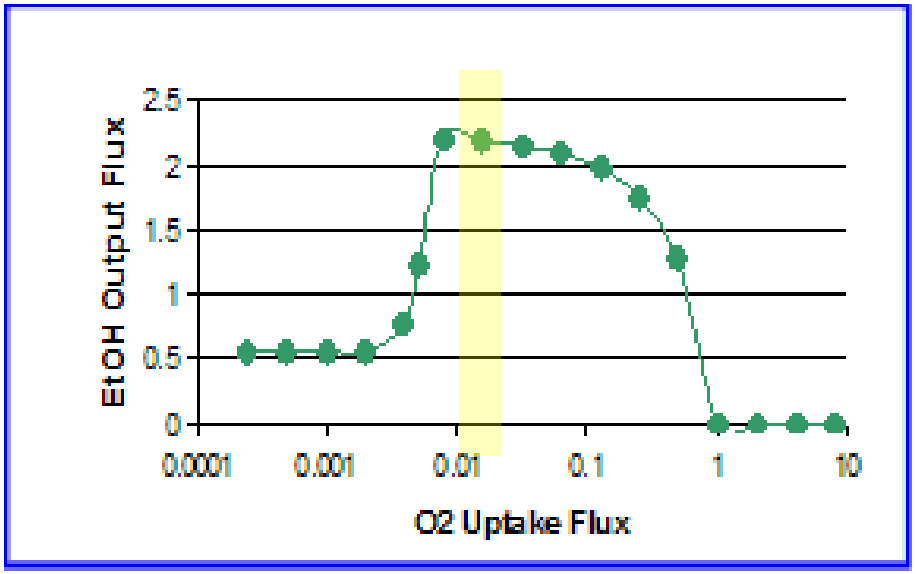


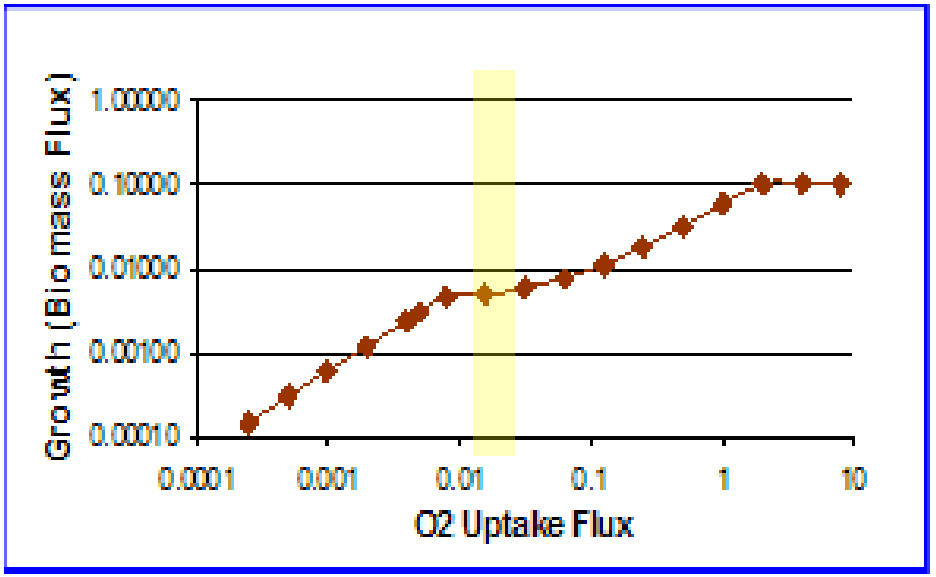


X and Y flux values = in mmol $g(DW)^{-1}$ hr^{-1}

Pacific Northwest
NATIONAL LABORATORY

A genome-wide gene deletion series was conducted under simulated microaerobic conditions (0.02 mmol g_{DW}^{-1} hr^{-1})



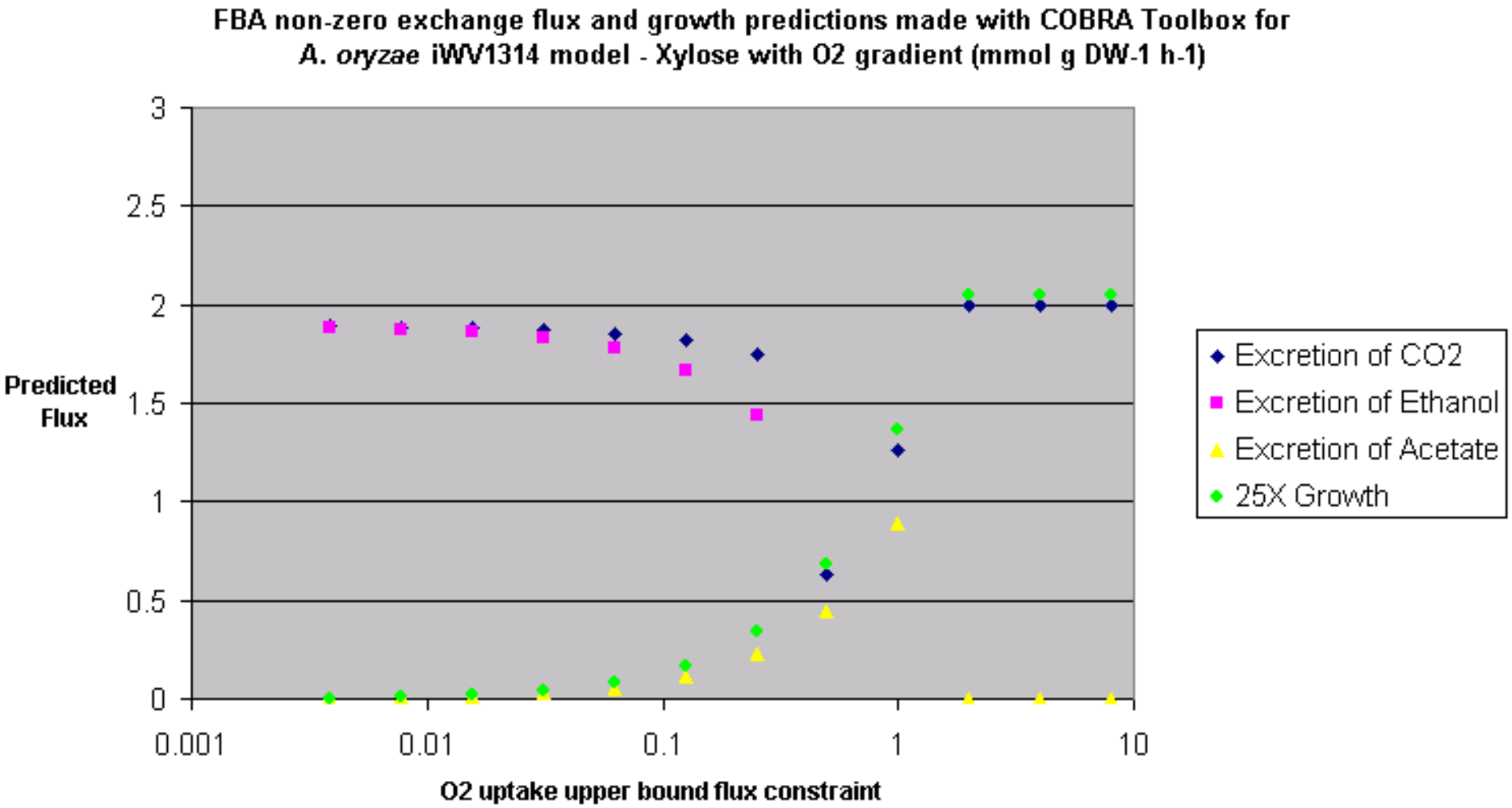


X and Y flux values = in mmol $g(DW)^{-1}$ hr^{-1}

Unconfirmed result: 11 gene deletions were predicted to boost ethanol excretion by 1-5%.

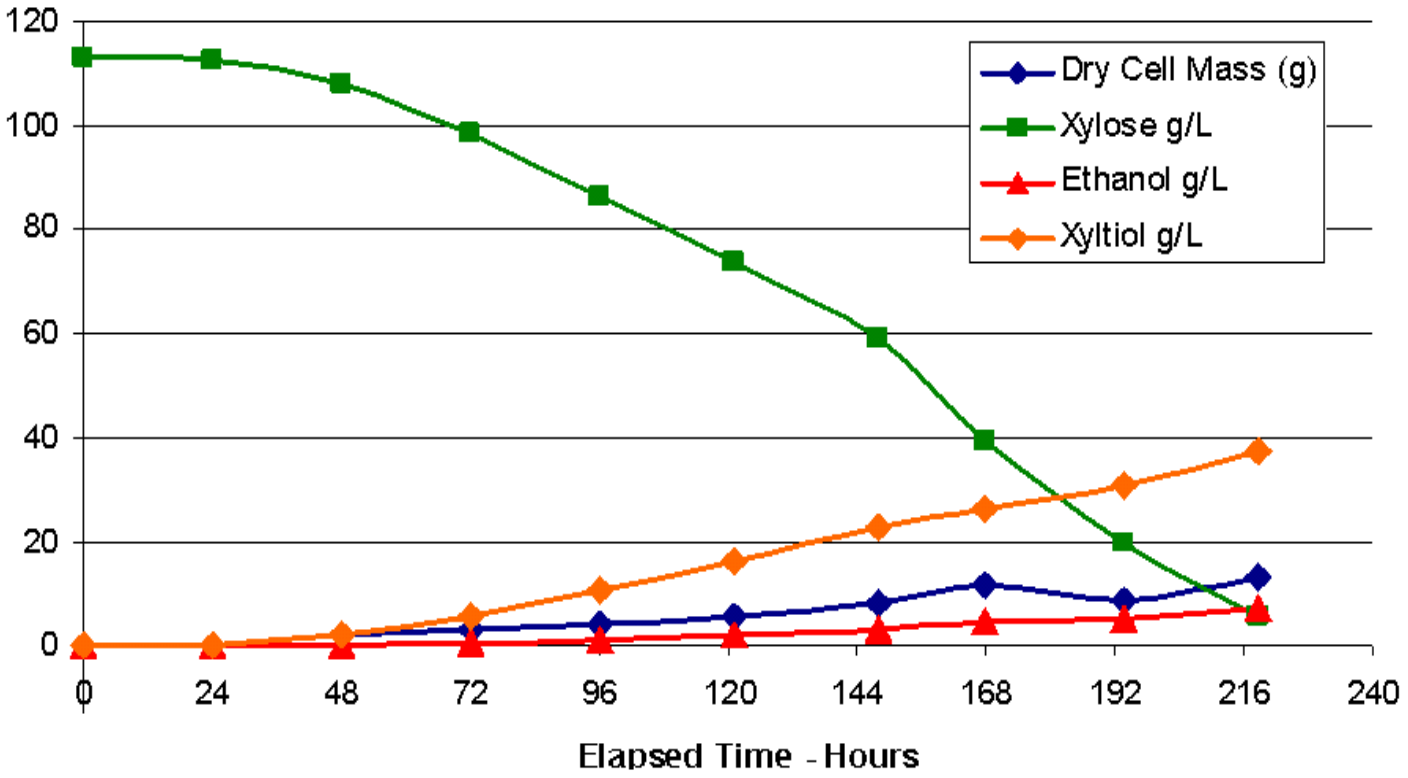
Pacific Northwest
NATIONAL LABORATORY

FBA simulation of *A. oryzae* fermentation on xylose



Pacific Northwest
NATIONAL LABORATORY

A. oryzae fermentation results on xylose



Pacific Northwest
NATIONAL LABORATORY

General “end-user” impressions of currently available FBA models and software

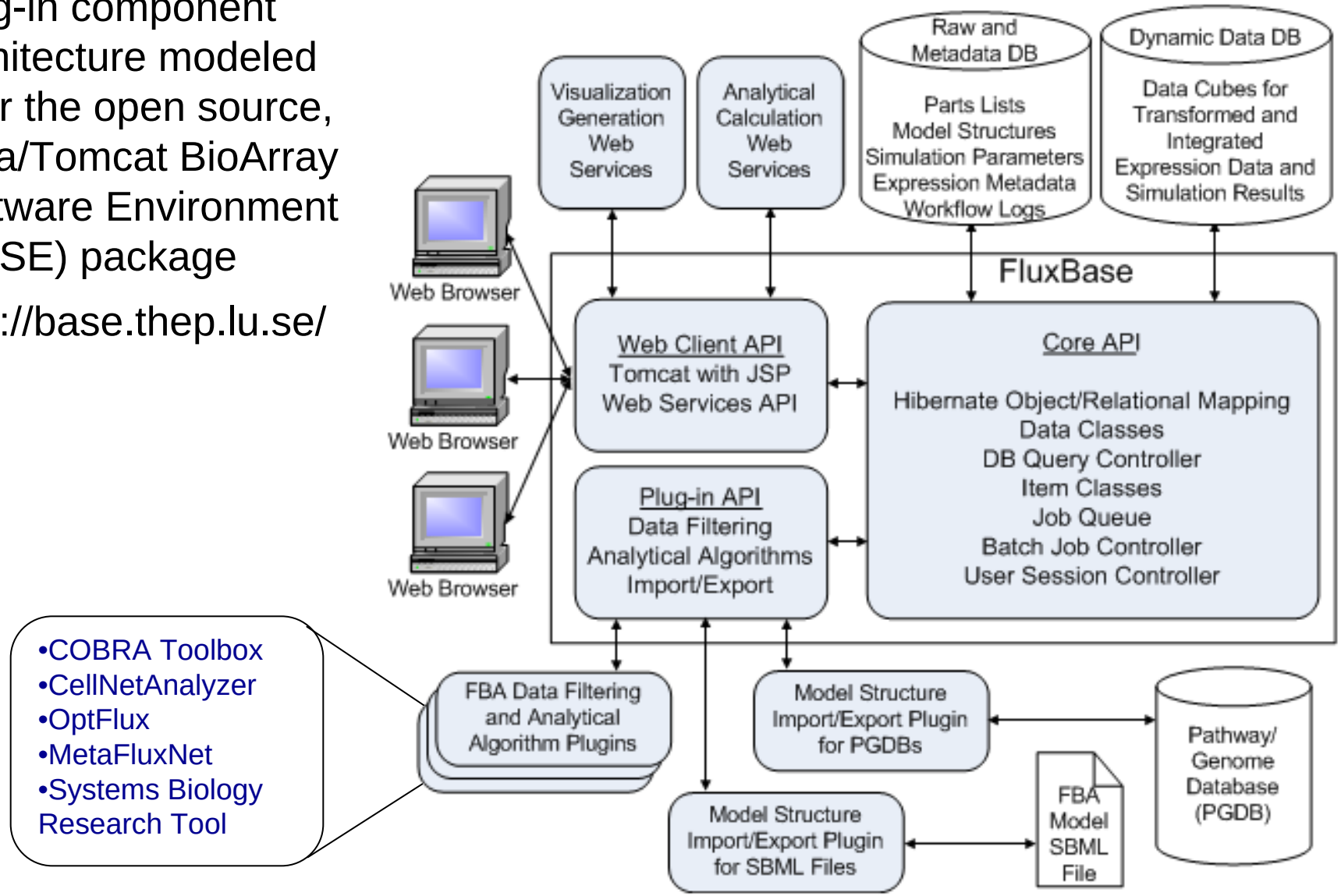
- “Formatted in SBML” != compatible across software packages.
- Model validation by growth rate may not guarantee accurate flux predictions for metabolites of interest.
- More basic research is needed on how to determine the true objective function of organisms under stress, far from idealized growth conditions.
- Metabolic reconstructions should ideally be community projects rather than competing products published by individual labs.
- FBA software should be more like an IDE (i.e., Eclipse) to support the “write-run-debug-run” cycle of model development and refinement.
- More automated tools for diagnosing errors in malfunctioning models are needed.

Pacific Northwest
NATIONAL LABORATORY

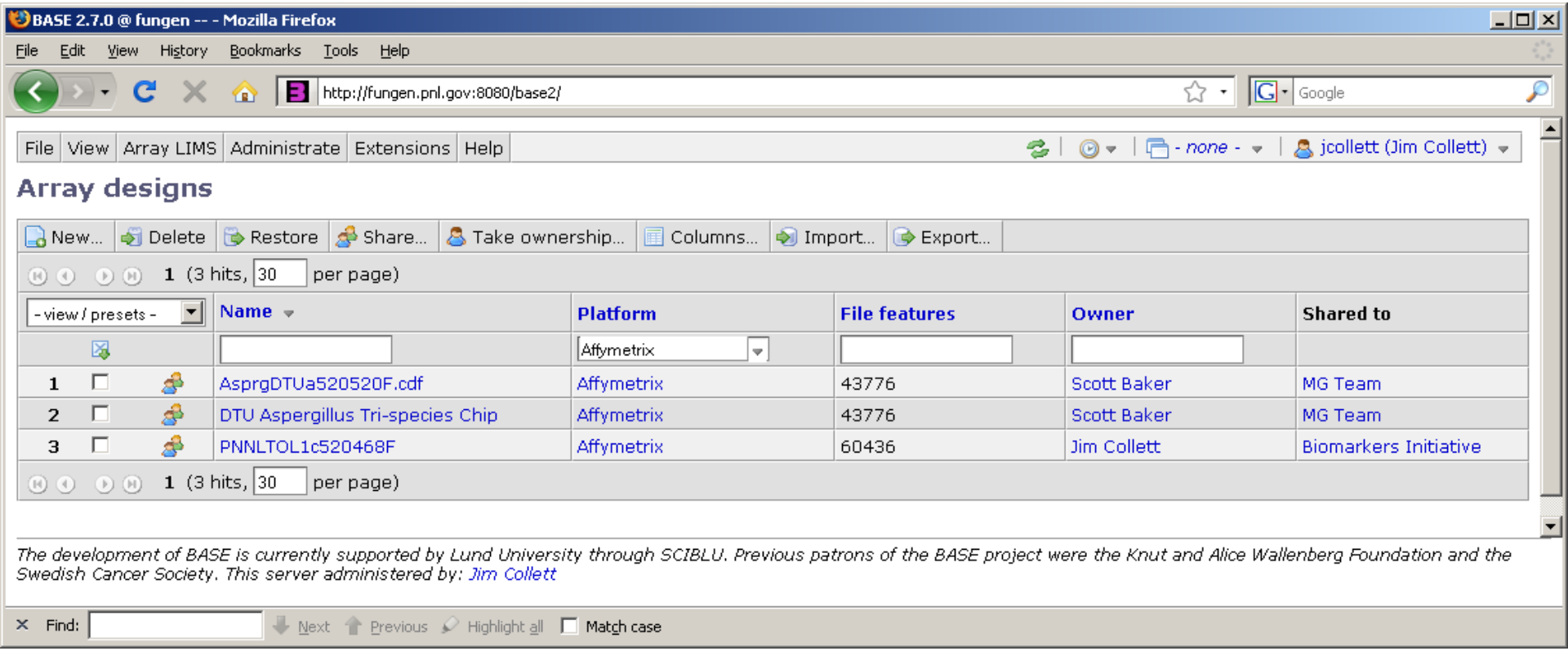
Suggested architecture for a collaborative metabolic network reconstruction & analysis and PGDB data management system

Plug-in component architecture modeled after the open source, Java/Tomcat BioArray Software Environment (BASE) package

http://base.thep.lu.se/



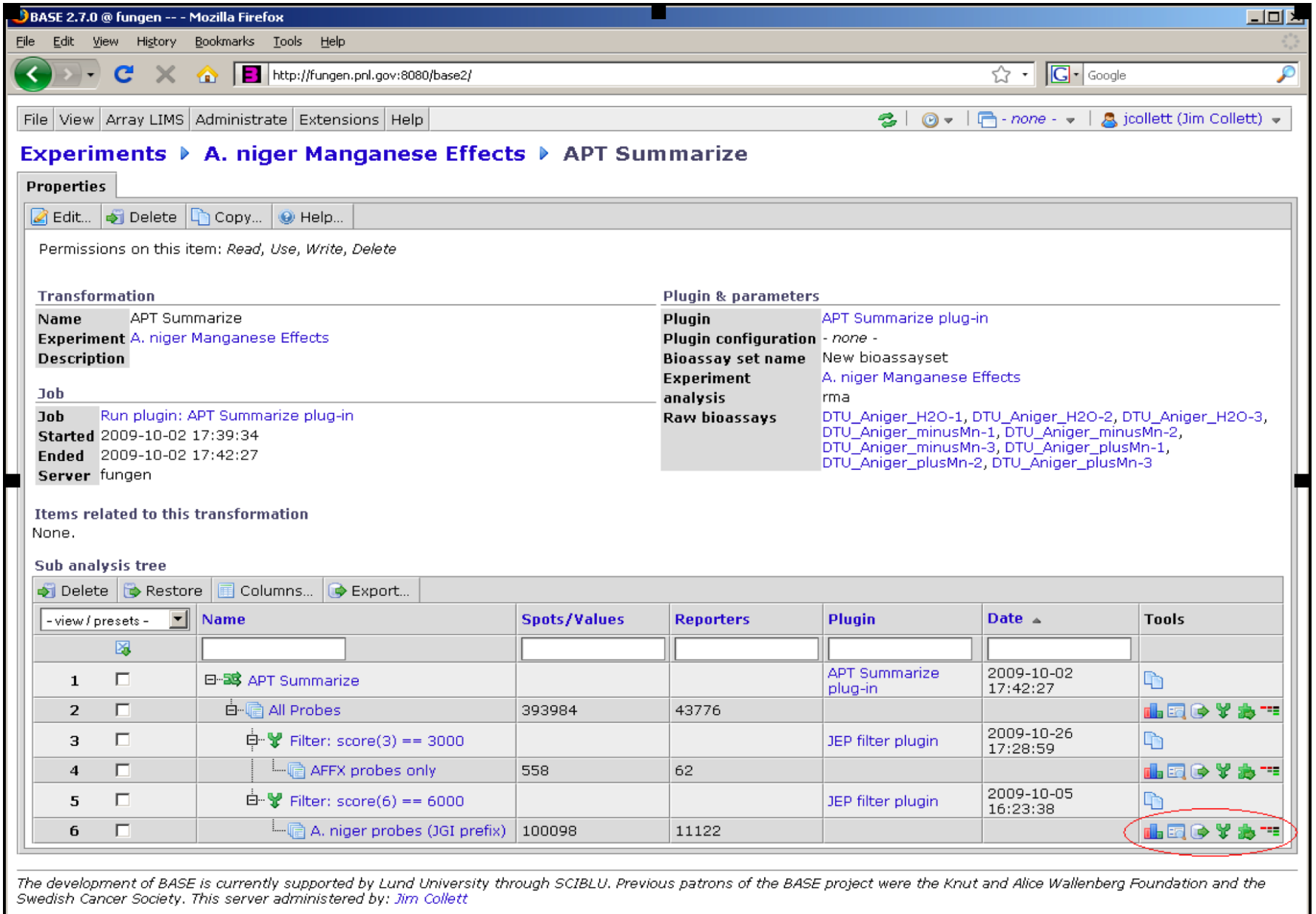
Data management features in BASE that would be useful in a collaborative FBA/PGDB computing environment



User- and group-level permissions and item ownership facilitate provenance control in projects with very large datasets and complex analytical workflows.

Pacific Northwest
NATIONAL LABORATORY

Analytical workflow features in BASE that would be useful in a collaborative FBA/PGDB computing environment



Collaboration with EU partners and JGI

Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing

Stéphane Le Crom[a,b,c,1], Wendy Schackwitz[d,1], Len Pennacchio[d], Jon K. Magnuson[e], David E. Culley[e], James R. Collett[e], Joel Martin[d], Irina S. Druzhinina[f], Hugues Mathis[g], Frédéric Monot[g], Bernhard Seiboth[f], Barbara Cherry[h], Michael Rey[h], Randy Berka[h], Christian P. Kubicek[f], Scott E. Baker[d,e,2], and Antoine Margeot[g,2]

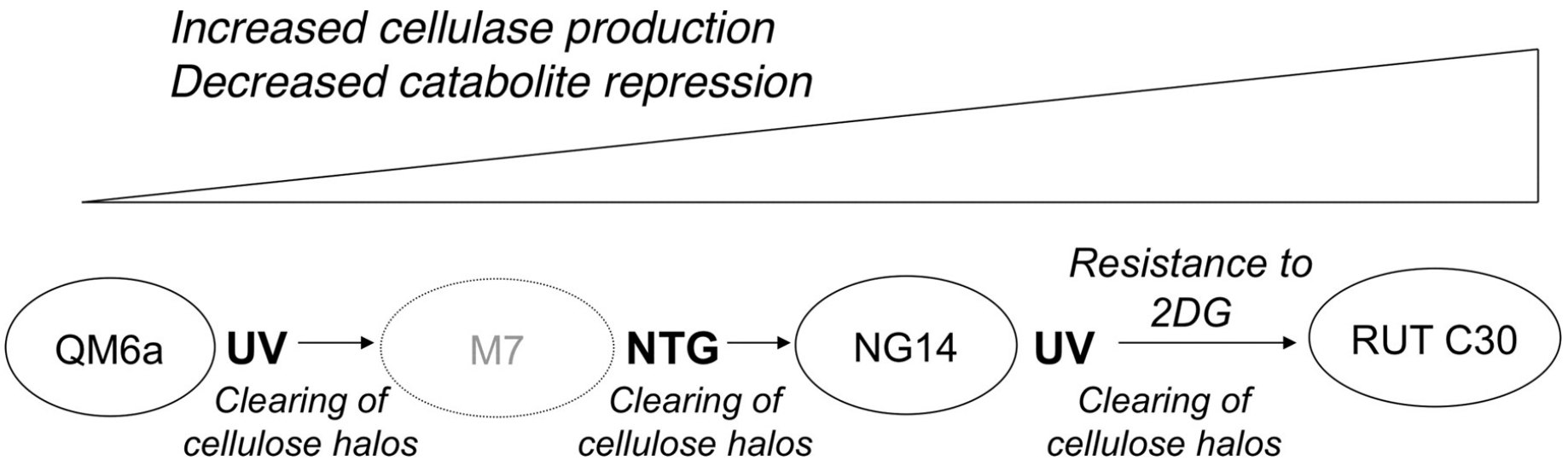
[a]Institut National de la Santé et de la Recherche Médicale, U784, 46 rue d'Ulm, 75230 Paris Cedex 05, France; [b]Institut Fédératif de Recherche 36, Plate-forme Transcriptome, 46 rue d'Ulm, 75230 Paris Cedex 05, France; [c]École Normale Supérieure, 46 rue d'Ulm, 75230 Paris Cedex 05, France; [d]Department of Energy Joint Genome Institute, 2800 Mitchell Avenue, Walnut Creek, CA 94598; [e]Pacific Northwest National Laboratory, P.O. Box 999, Richland, WA 99352; [f]Institute of Chemical Engineering, Technische Universitat Wien, Getreidemarkt 9/166, A-1060 Vienna, Austria; [g]IFP, Département Biotechnologie, Avenue de Bois-Préau, 92852 Rueil-Malmaison Cedex, France; and [h]Novozymes, Inc., 1445 Drew Avenue, Davis, CA 95618

Edited by Joan Wennstrom Bennett, Rutgers University, New Brunswick, NJ, and approved July 27, 2009 (received for review May 28, 2009)

Le Crom, Schackwitz, et al. 2009. PNAS 106 (38): 16151-6

Pacific Northwest NATIONAL LABORATORY

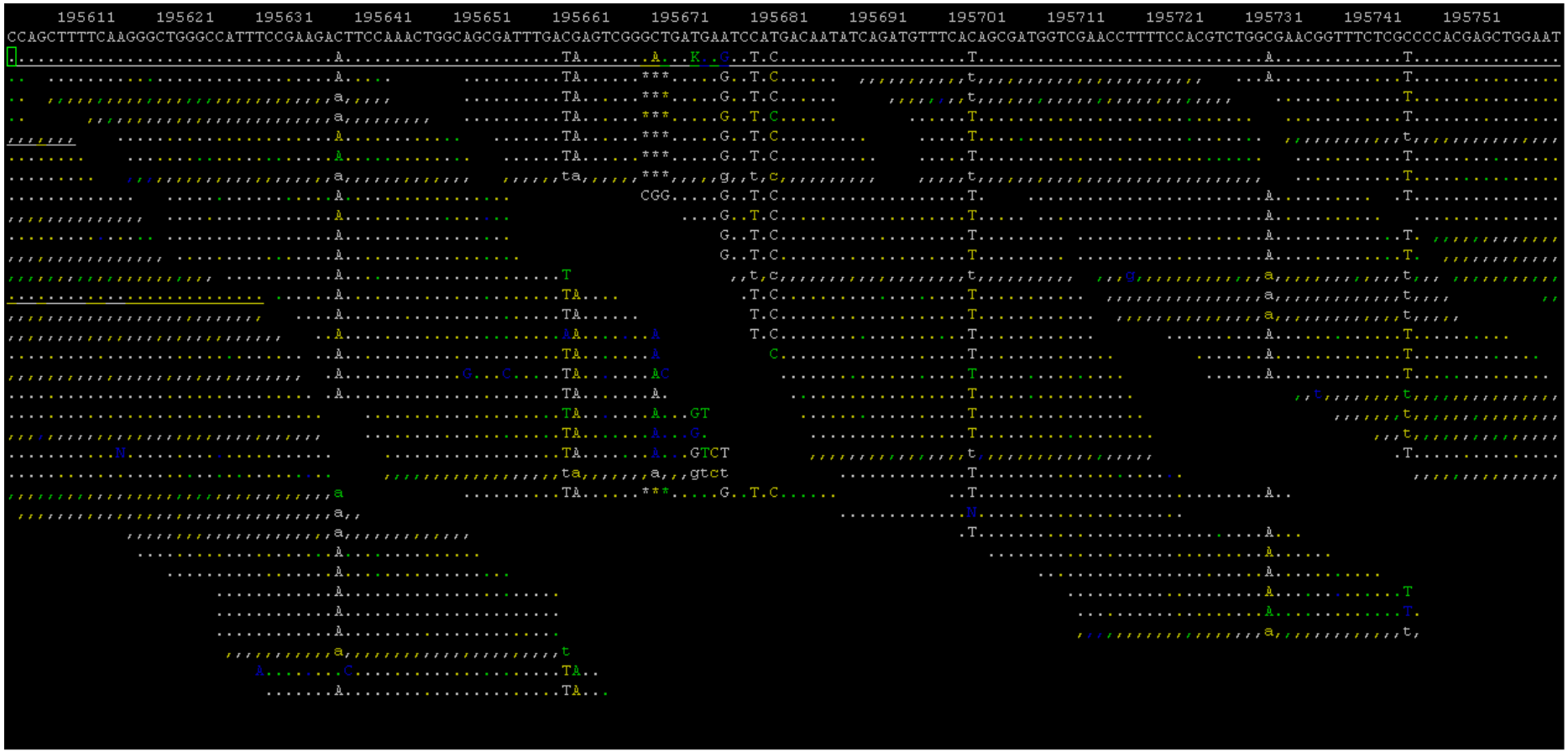
Genealogy of mutagenized *T. reesei* strains



Le Crom S et al. PNAS 2009;106:16151-16156

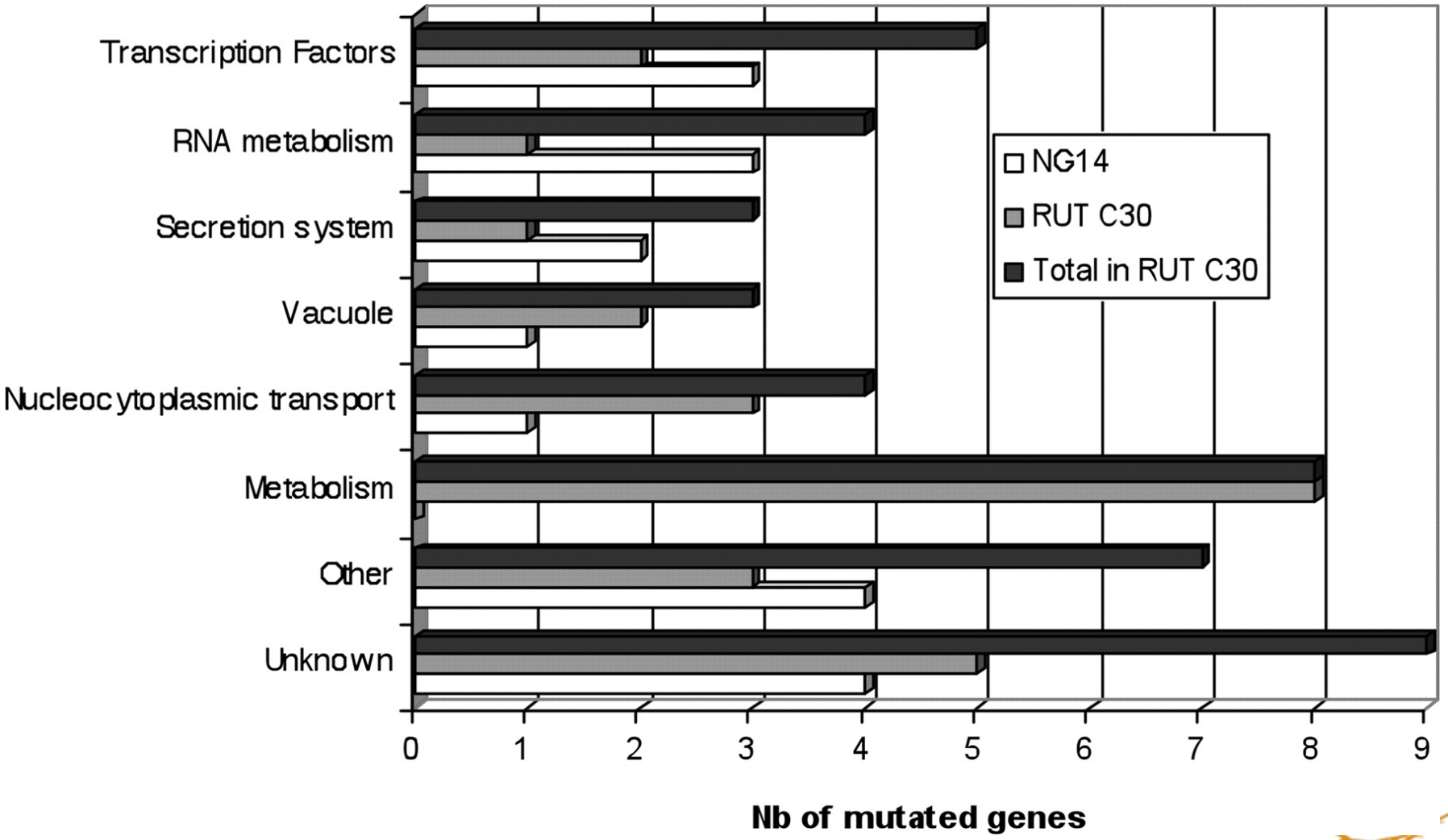
Pacific Northwest NATIONAL LABORATORY

Reads from *T. reesei* strains NG14 and RUT C30 aligned with QM6a to identify SNVs and indels



Pacific Northwest
NATIONAL LABORATORY

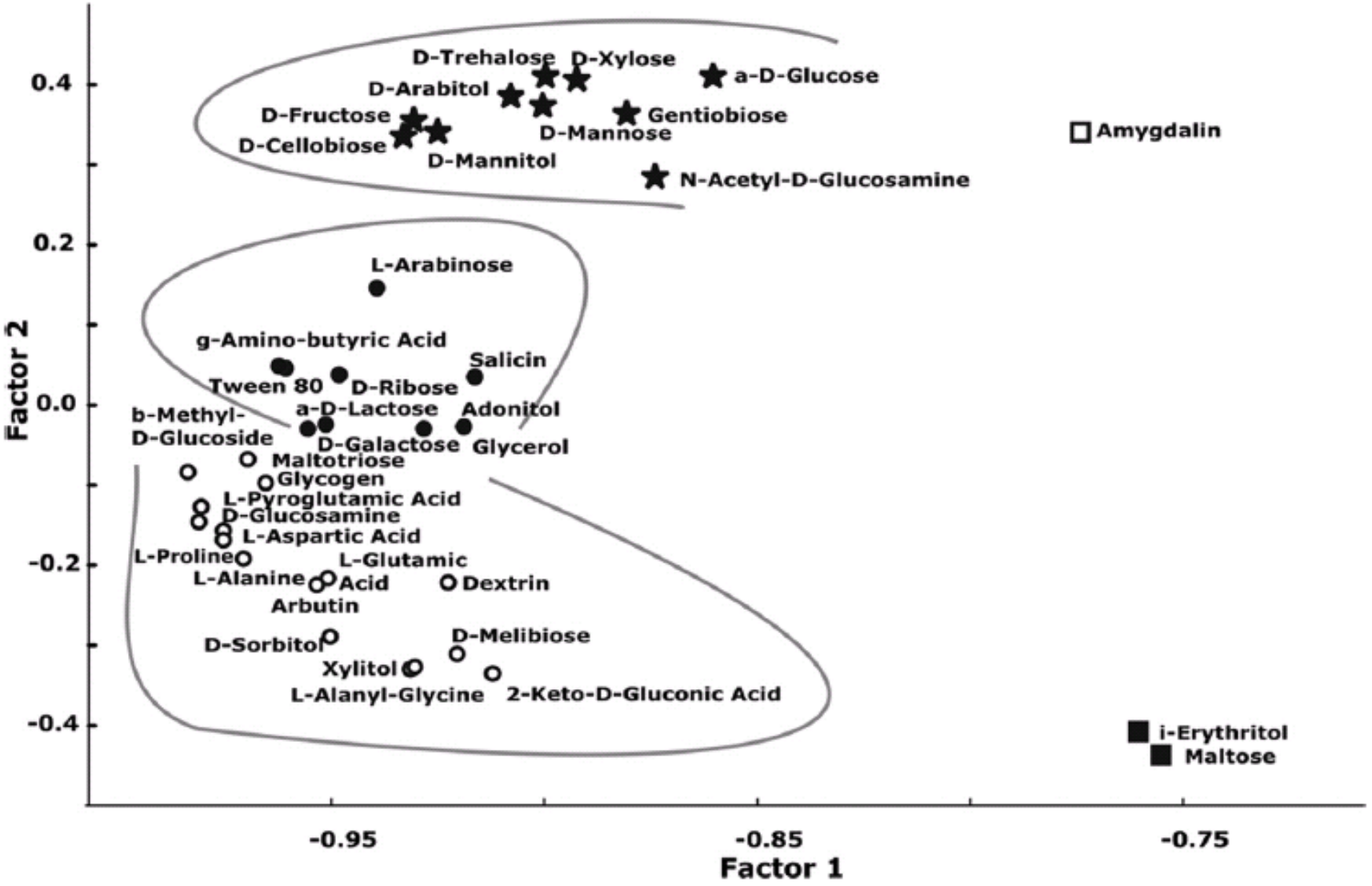
Gene categories of mutagenic events



Pacific Northwest
NATIONAL LABORATORY

Le Crom S et al. PNAS 2009;106:16151-16156

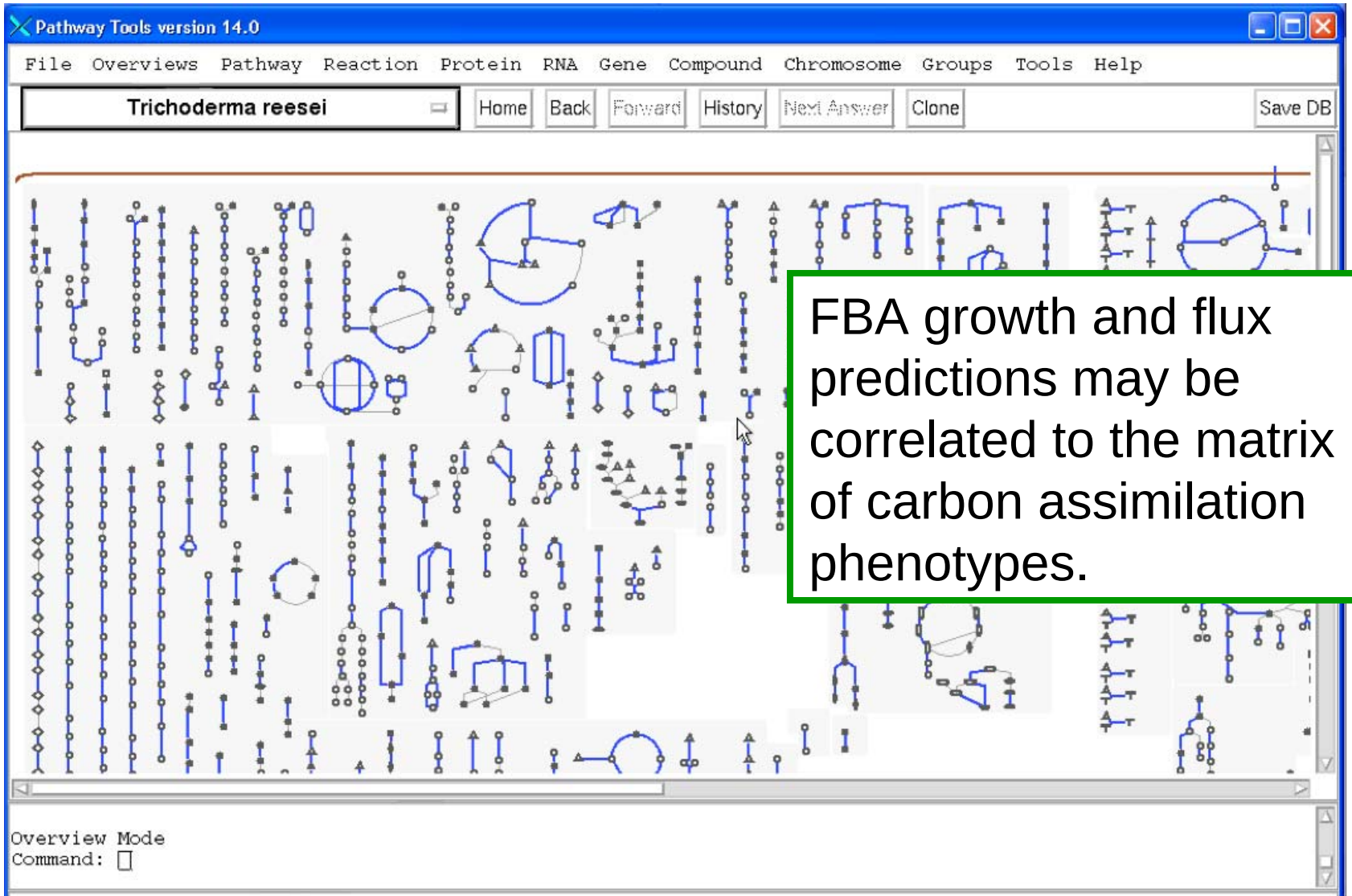
Biomass growth profiling on 95 carbon substrates using the Biolog phenotyping system



Pacific Northwest NATIONAL LABORATORY

Le Crom S et al. PNAS 2009;106:16151-16156

Plans for using P-Tools 14. 5+ to correlate SNVs with KO experiments, and to help generate FBA models



FBA growth and flux predictions may be correlated to the matrix of carbon assimilation phenotypes.

Acknowledgements

PNNL Fungal Biotech Team
Scott Baker (Genomics PM), Deanna Auberry, Ken Bruno, Mark Butcher, Dave Culley, Ziyu Dai, Shuang Deng, Beth Hofsted, Sue Karagiosis, Debbie Lee, John Magnuson, Iva Jovanovic, Ellen Panisko, Andy Zwoster + Sebastian Jaramillo-Riveri.
Special thanks to our EU and JGI collaborators.

Pacific Northwest
NATIONAL LABORATORY