



DOE Systems Biology Knowledgebase

The DOE Systems Biology Knowledgebase

<http://kbase.science.energy.gov>

KBase Progress and Goals

Adam Arkin

Physical Biosciences Division

Lawrence Berkeley National Laboratory



DOE Office of Science
Office of Biological and Environmental Research



*Plants and Microbes
for Energy and Environment*

DOE Systems Biology Knowledgebase

Components Enabling Its Development

2010 Knowledgebase R&D Project

Year-long effort funded by American Recovery and Reinvestment ACT (ARRA).

Results from this project, which was completed in September 2010, are underpinning Knowledgebase development:

- DOE Systems Biology Knowledgebase Implementation Plan
- ARRA pilot projects

Biological Principles

Metabolism
Integration
Interactions
Data Visualization
Proteins
Mathematics
Algorithms
Gene Expression

Computing

User Community Data and Resources

The Knowledgebase will leverage and establish critical partnerships with other synergistic efforts:

- DOE Joint Genome Institute
- National Center for Biotechnology Information
 - iPlant Collaborative
- DOE Office of Advanced Scientific Computing Research
 - Others

University-Led Projects to Develop Computational Biology and Bioinformatic Methods Enabling the Knowledgebase

In 2010, the DOE Office of Biological and Environmental Research awarded funding to 11 projects for basic research that will support Knowledgebase development in four areas:

- Omic data integration
- Integrated pathway reconstruction
- Genomic annotation
- Whole cellular simulations

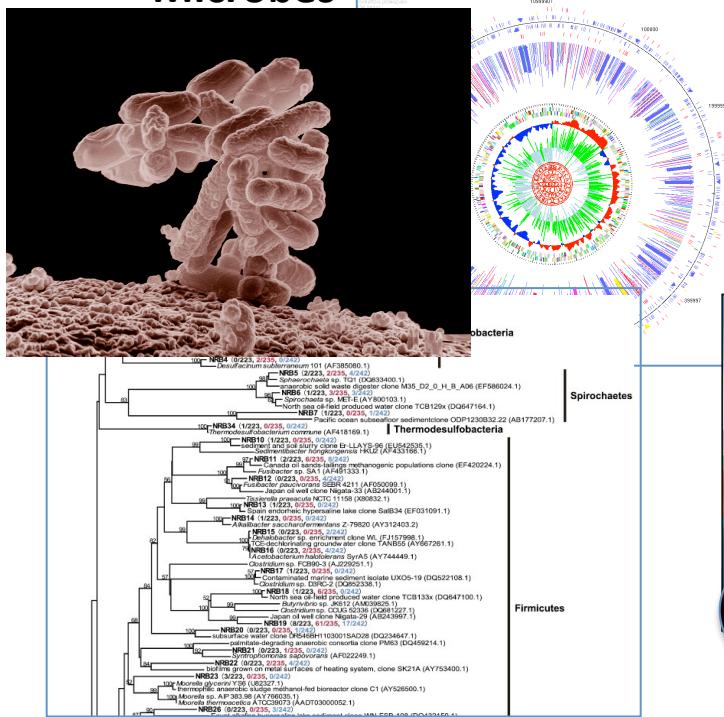
Predictive Understanding

- **DOE mission:** *predict, control and design* the biological components of energetic processes and environmental balance.
- Complex missions with rapidly expanding, intricately related diverse data types require ways to augment scientists' ability to:
 - Filter information
 - Focus attention
 - Ask the right questions
 - Leverage other minds
 - Make the impossible possible. ☺

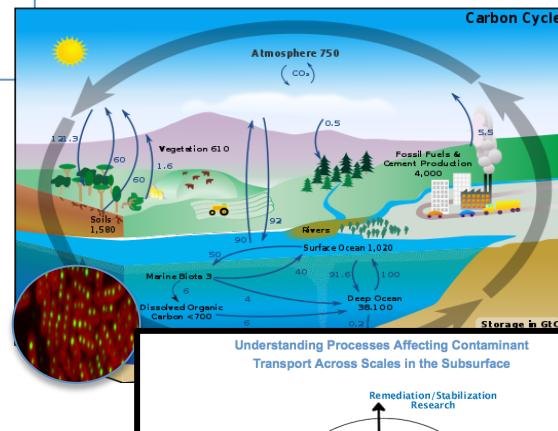
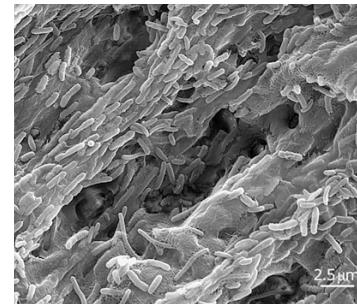


Enable DOE Mission Science

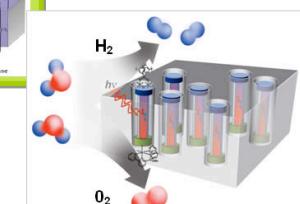
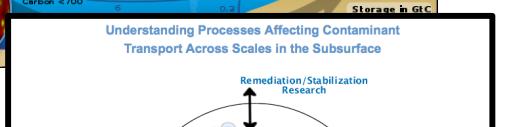
Microbes



Communities



Plants



Before I tell you how...

DID WE MEET SPECIFICATIONS?

Distributed Development

- Four major National Labs and more than 10 collaborating universities/research centers
- 50 people with diverse cultures and expertise and with different institutional and group alliances
- Enthusiasm to create a system to revolutionize understanding and application of biological systems.
- **Deliver version 1 in 18 months and prove it can support DOE scientists!**

Building an airplane in the sky



Community Specifications for KBase: 2011

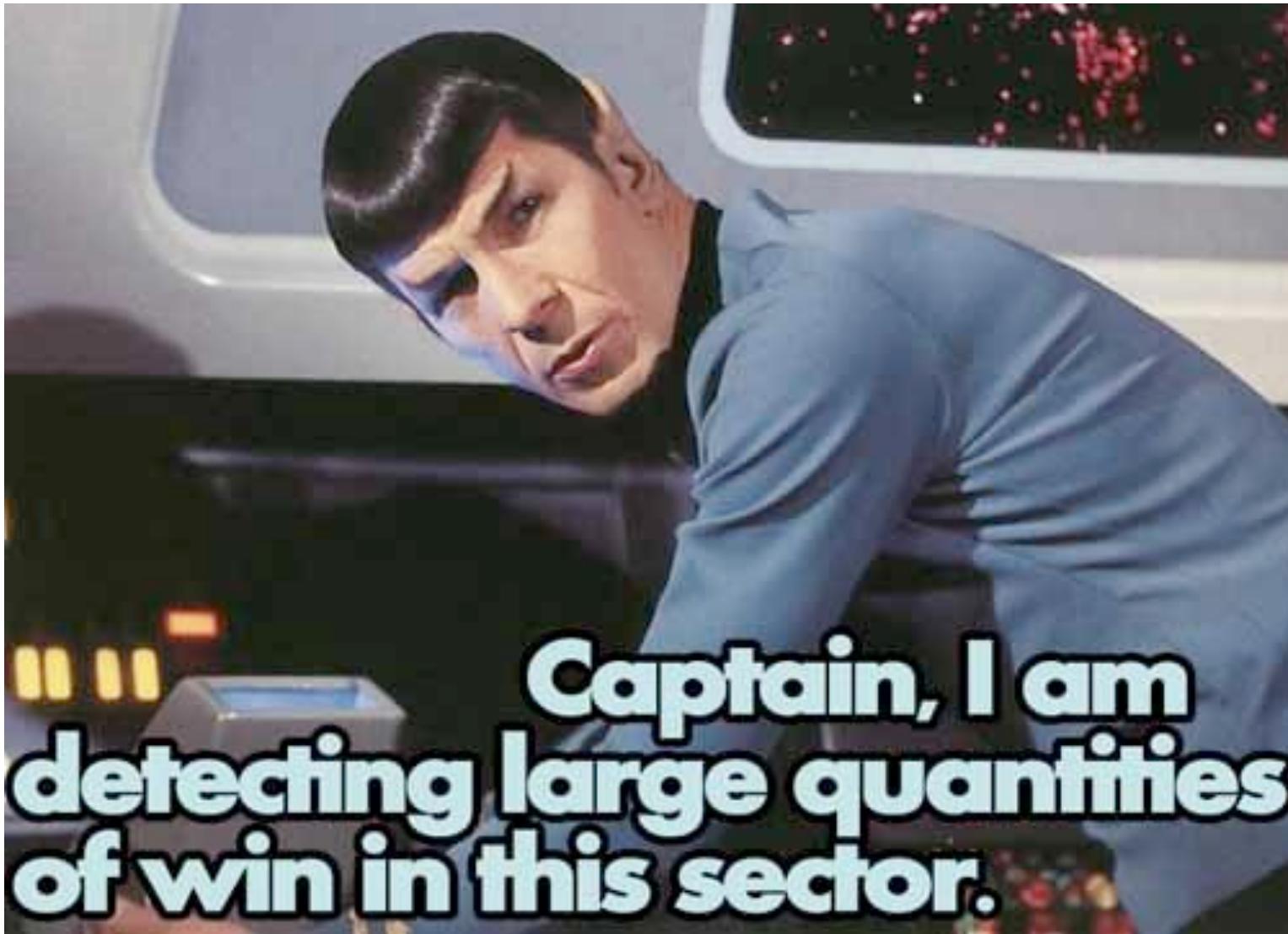
- A Federated infrastructure and 10GBit/s transfer capabilities.
- 2 petabytes data storage and 2000 cores for data processing including interface and between high performance computing and cloud computational resources
- Development of core Knowledgebase integrated data and workflow analysis management tools including Application Programming Interfaces, semantic and user Interfaces.
 - i. For microbial systems, from 100-1000 microbes: Reconstruct and Predict Metabolic and Gene Expression Regulatory Networks to Manipulate Microbial Function
 - ii. For plant systems, for 10 key plants related to DOE missions: Integrate Phenotypic and Experimental Data and Metadata to Predict Biomass Properties from Genotype and Assemble Regulatory Data to Enable Analysis, Cross- Comparisons, and Modeling
 - iii. For Microbial Communities: Model Metabolic Processes within Microbial Communities and Mine Metagenomic Data to Identify Unknown Genes

KBase 2013: 18 Months Deliverables Exceed Spec.

- A Federated infrastructure and 10GBit/s transfer capabilities.
 - Built for high speed data transfer over ESNET using 100 GBit/s rates.
- 2 petabytes data storage and 2000 cores for data processing including interface and between high performance computing and cloud computational resources
 - KBase Magellan has 12,000 cores for data processing via both Open Stack Cloud interfaces and Cluster Services
 - KBase has >3 Petabytes of storage capacity
- Development of core Knowledgebase integrated data and workflow analysis management tools including Application Programming Interfaces, semantic and user Interfaces.
 - Integrated KBase API specified and operational. *Used by third parties to integrate function and build apps.*
 - Integrated data model aware of 925 data types encompassing sequence reads, contigs, genomes, genome features, transcription data, fitness data and more.
 - 40 Interface description documents leading to 821 functions that can be compiled into use for PERL, Python, Java, and R.
 - Prototype Search, Workflow and Novel Narrative/Notebook interfaces for navigating, analyzing and building knowledge in KBase

KBase 2013: 18 Months Deliverables Exceed Spec.

- i. For microbial systems, from 100-1000 microbes: Reconstruct and Predict Metabolic and Gene Expression Regulatory Networks to Manipulate Microbial Function
 - i. Metabolic and regulatory reconstructions for 5534 prokaryotic and 161 archaeal genomes
 - ii. 7830 genome annotations, 23,058,670 features predicted
 - iii. 12,620 regulons with 266,345 protein families inferred
 - iv. 4985 metabolic models including a total of 16,196 compounds and 13,428 reactions
 - v. 6202 growth curves, 1,947,690 strain fitness measurements; 3227 gene expression data sets
 - vi. Services for assembly, annotation, phylogenomics, regulatory and metabolic networks inference, FBA and PROM modeling of metabolism, reconciliation and improvement of models against data
- ii. For plant systems, for 10 key plants related to DOE missions: Integrate Phenotypic and Experimental Data and Metadata to Predict Biomass Properties from Genotype and Assemble Regulatory Data to Enable Analysis, Cross-Comparisons, and Modeling
 - i. Over 175 eukaryotic genomes including many variants of Poplar, Arabidopsis, Sorghum, Chlamydomonas, Brachypodium, Miscanthus, Switchgrass as well as many other algae and fungi.
 - ii. Phenotypes for genome variants of plants and services for calling the genetic variation among individuals.
 - iii. Services for variation calling, mapping genotype-to-phenotype via GWAS style analysis and tools for candidate gene filtering, trait modeling, and pathway enrichment,
 - iv. 731 gene expression experiments in Arabidopsis and Poplar; Plant co-expression network analysis for all.
 - v. Initial plant metabolic modeling
- iii. For Microbial Communities: Model Metabolic Processes within Microbial Communities and Mine Metagenomic Data to Identify Unknown Genes
 - i. Access to 11,000 metagenomes(>21 TBp)
 - ii. Integrated KBase access to QIIME functionality
 - iii. New tools for metagenome sequence quality assessment and experimental design
 - iv. Services for taxonomic and gene identification, abundance, and a host of other functions



The community document is not a specification, though...

WHY DO WE NEED KBASE?

Our goals drive our specification

- We believe it is necessary to solve the grand challenges in biological control of environmental and energy process.
- We believe it is necessary to deal with deep issues in scientific communication and reproducibility
- We need to lower the barriers to computationally efficient use of advanced algorithms and data from diverse producers.
- We believe we need to do this together-- leveraging many minds.

The number of bioinformatics resources is huge

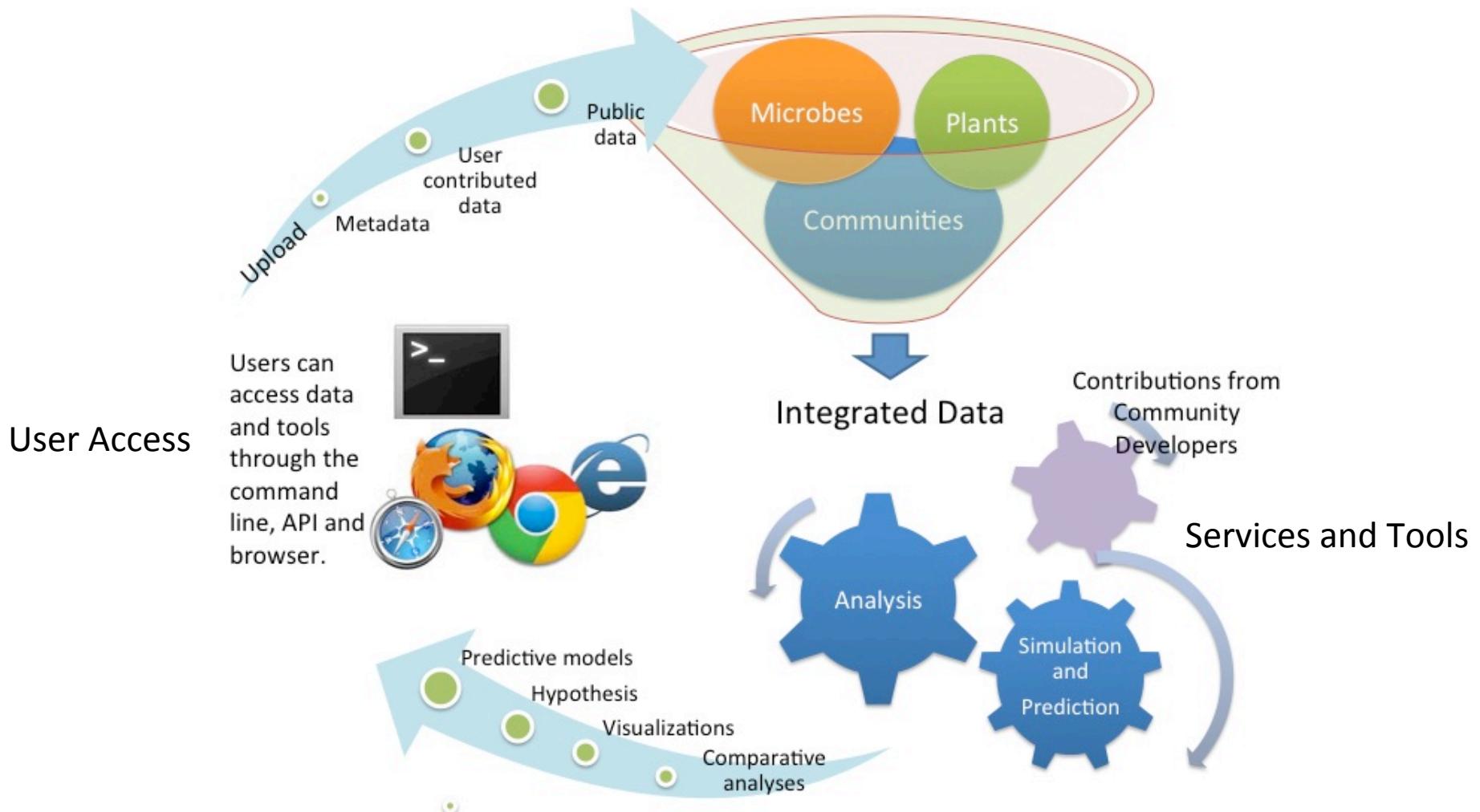
- An increasing number of data warehouses since biology is becoming a big data discipline (e.g., NCBI, Ensembl, etc.)
- Specialized applications and databases for relatively generic analyses (e.g., MG-RAST and MicrobesOnline)
- Evolving libraries of sophisticated computational biology algorithms for use in programming environments (e.g., Bioconductor)
- Workflow tools that allow the chaining of these algorithms together by non-programmers (e.g., Galaxy and Taverna)
- Workflow sharing tools to allow people to use each other's work products
- Open-access publication of journal articles with increasing use of semantic tagging
- Scientific social networks (e.g., ResearchGate, Epernicus, etc.)

So what is missing?!

- KBase drives data through models to predictions and experimental design.
- KBase accelerates reproducible, reusable, and transparent science.
- KBase deeply enables scientists to work together to approach complex biological problems.
- KBase gives credit where it is due and privacy where it is needed.
- KBase is an open software and data environment to which others can contribute and with which others can build.

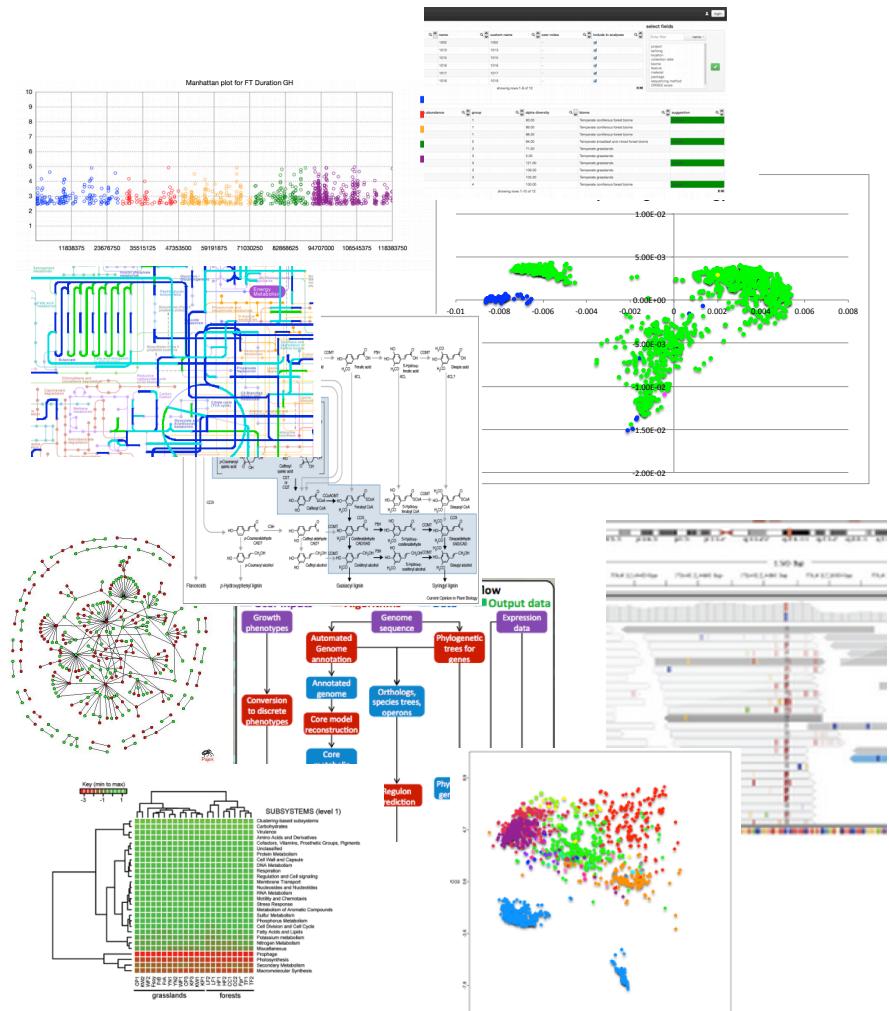


Empowering Users



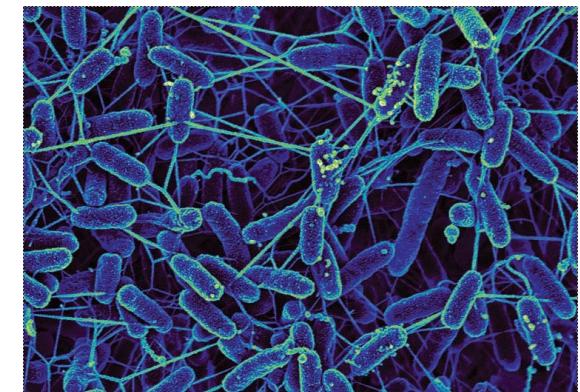
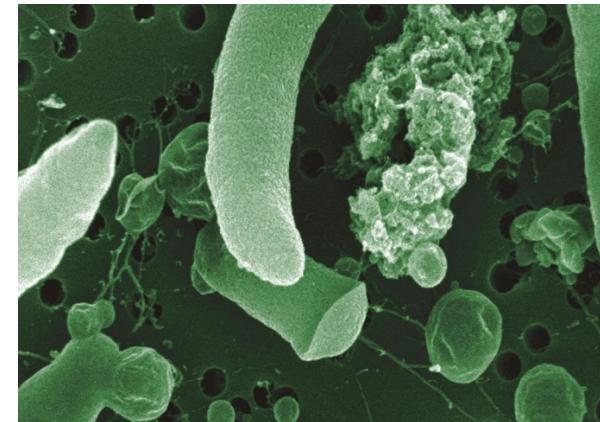
What can we do that was hard before?

- Transparently access multiple heterogeneous datasets and bioinformatics tools.
- Efficiently annotate new microbial genomes and infer metabolic and regulatory networks.
- Transform network inferences into metabolic models and map missing reactions to genes using novel data reconciliation tools.
- Design effective sequencing strategies for complex multi-sample metagenomic projects
- Test microbial ecological hypotheses through taxonomic and functional analysis of quality-assessed metagenomic data
- Predict plant gene function and molecular phenotype via navigation and analysis of tissue-specific co-expression networks.
- Discover genetic variations within plant populations and map these to complex organismal traits.
- Share data, analysis tools, workflows and scientific conclusions with the community.

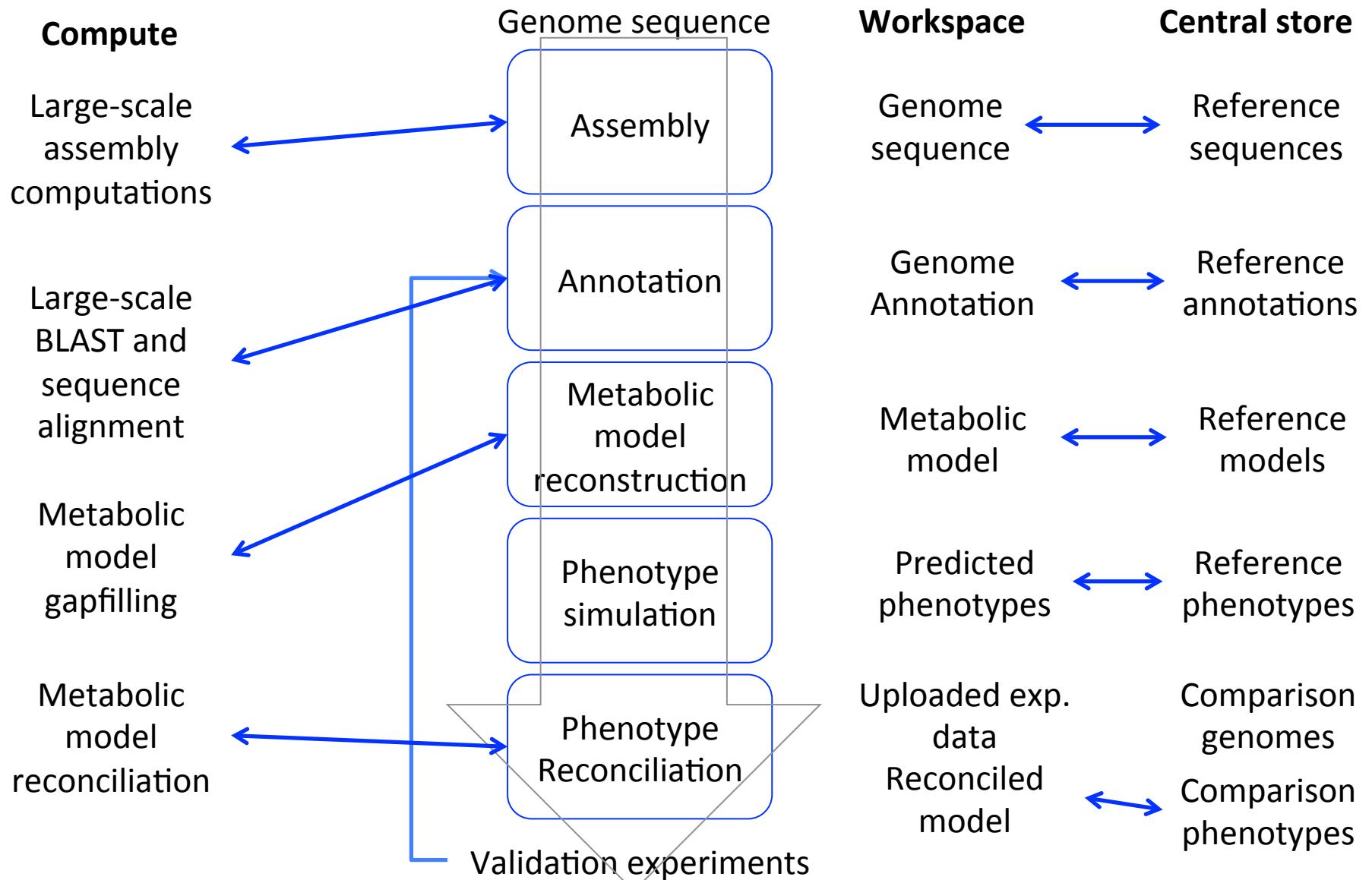


Our overall goals are to:

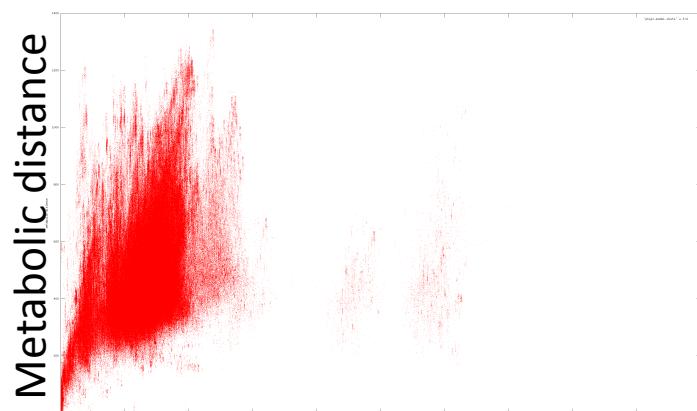
- Reconstruct and predict metabolic and gene expression regulatory networks to manipulate microbial function
- Vastly increase the capability of the scientific community to communicate and utilize their existing data
- Enable the planning of effective experiments and maximize our understanding of microbial system function



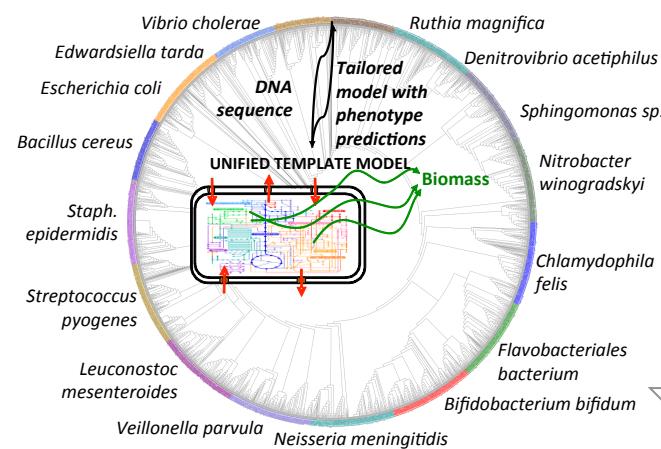
Rapidly building understanding of microbe behavior



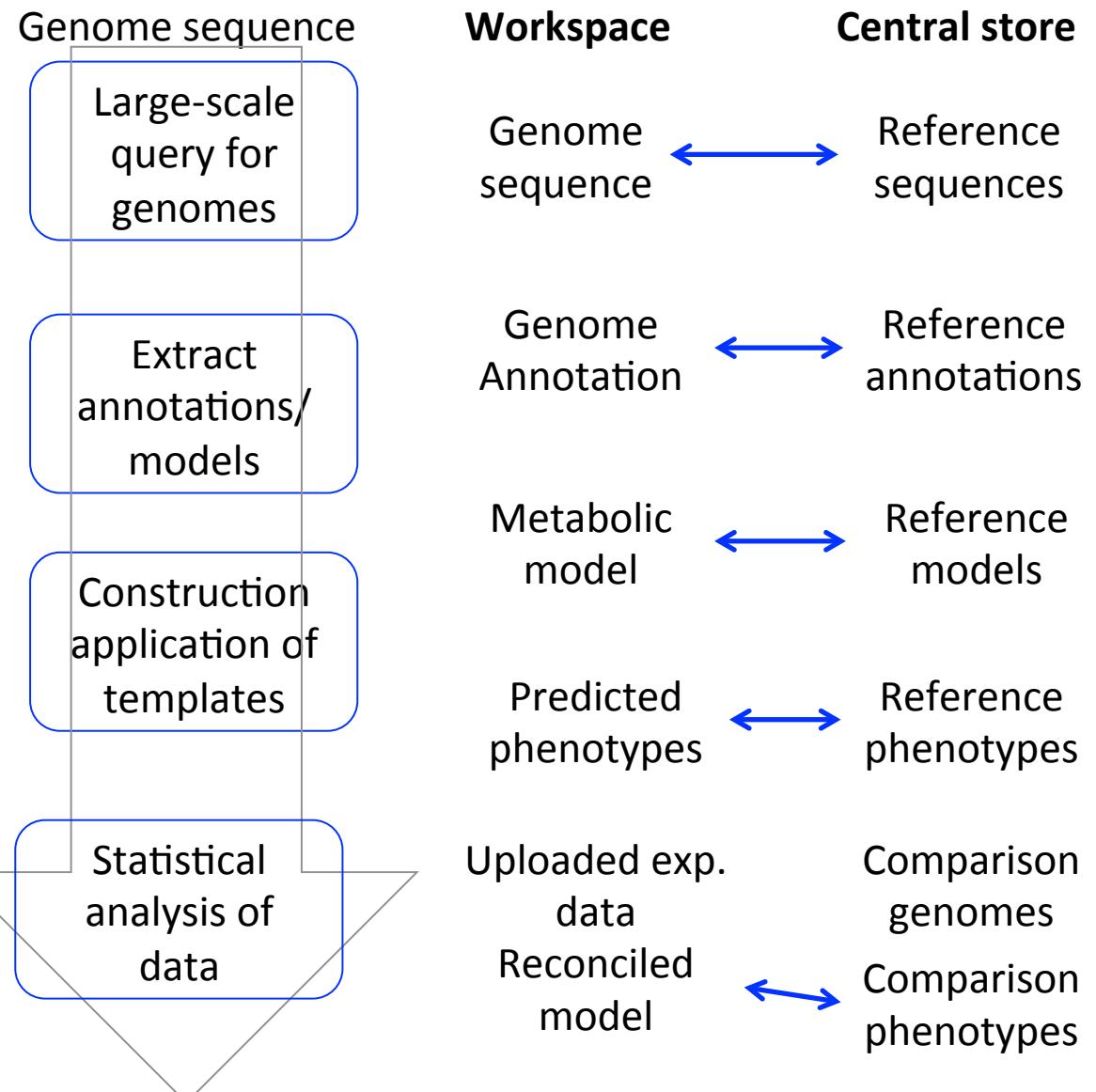
Biological trends across the tree of life



Phylogenetic distance



Assessing “completeness” of our understanding



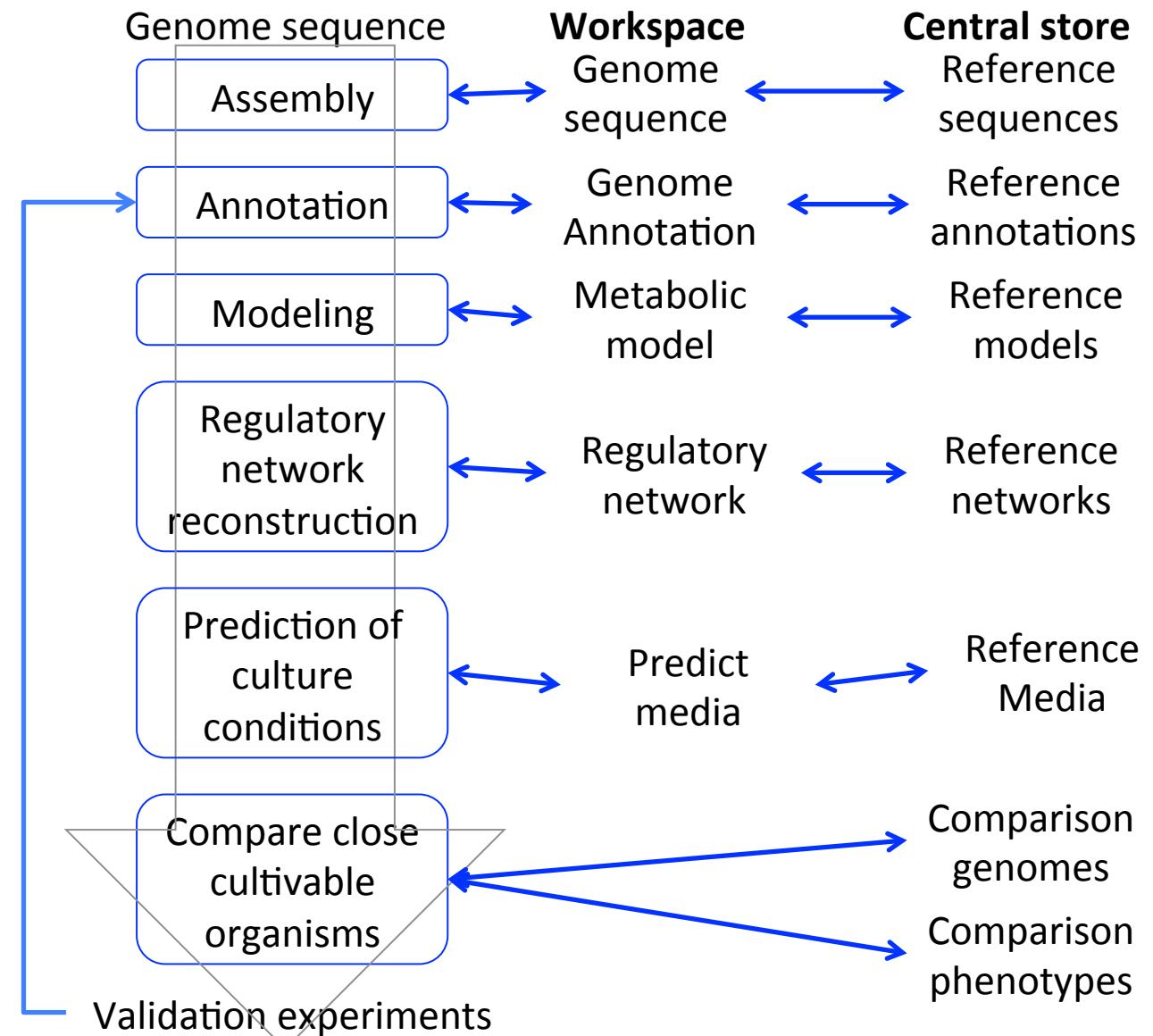
Driving discovery in experimental biology

Predict nutrients required for biomass biosynthesis

Predict potential nutrients that may inhibit growth

Compare models and predicted growth behavior of nearby cultivable microbes with genome of interest

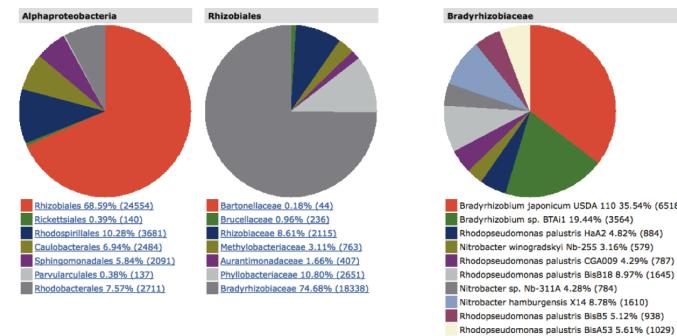
Explore microbe-microbe codependency or microbe-host dependency hypotheses



KBase Communities Supports Massive Metagenomics

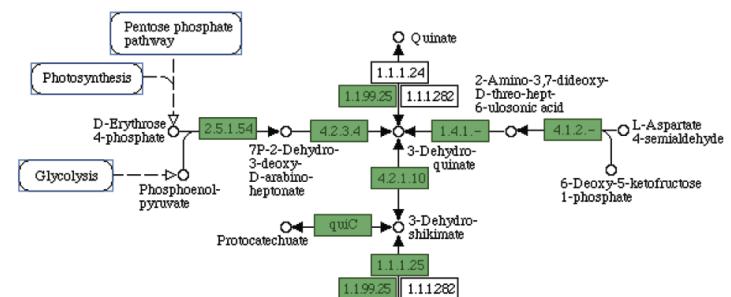
- Automated processing and analysis of metagenomic data (16s/18s, shotgun metagenome, meta transcriptome)
- Incorporation of MG-RAST and QIIME functionality
- Novel sequence QC pipelines (DRSEE)
- Evidence-based design of metagenomic experiments

Who are they?



What are they doing?

PHENYLALANINE, TYROSINE AND TRYPTOPHAN BIOSYNTHESIS



KBase-Plants: Targeted genomes

Poplar



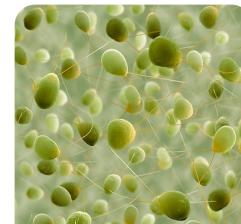
Arabidopsis



Sorghum



Chlamydomonas



Brachypodium



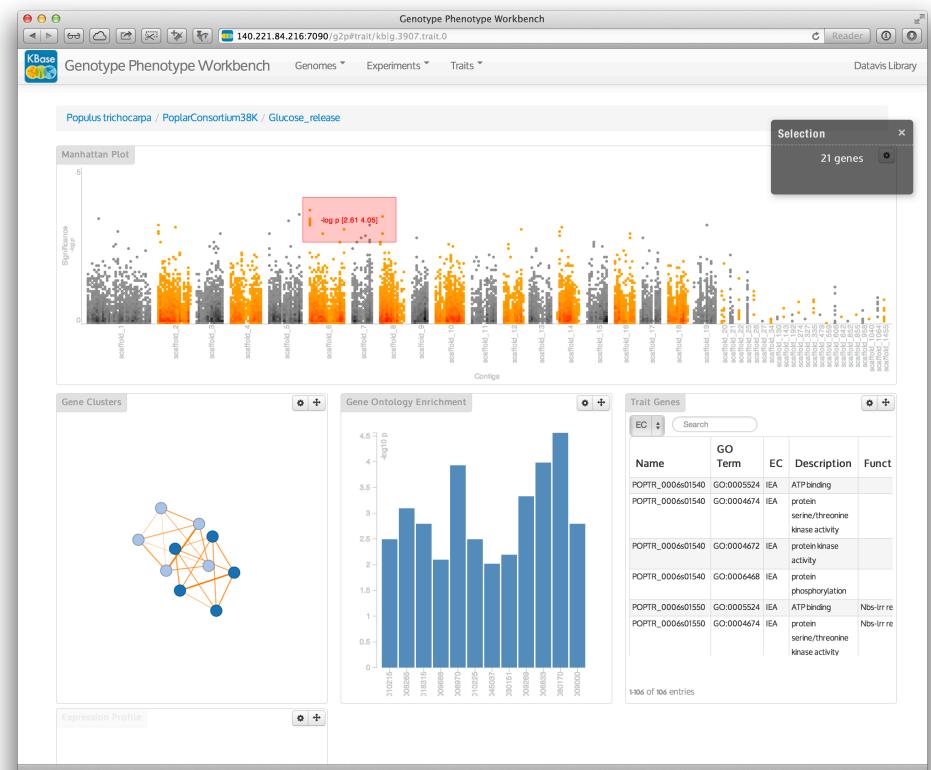
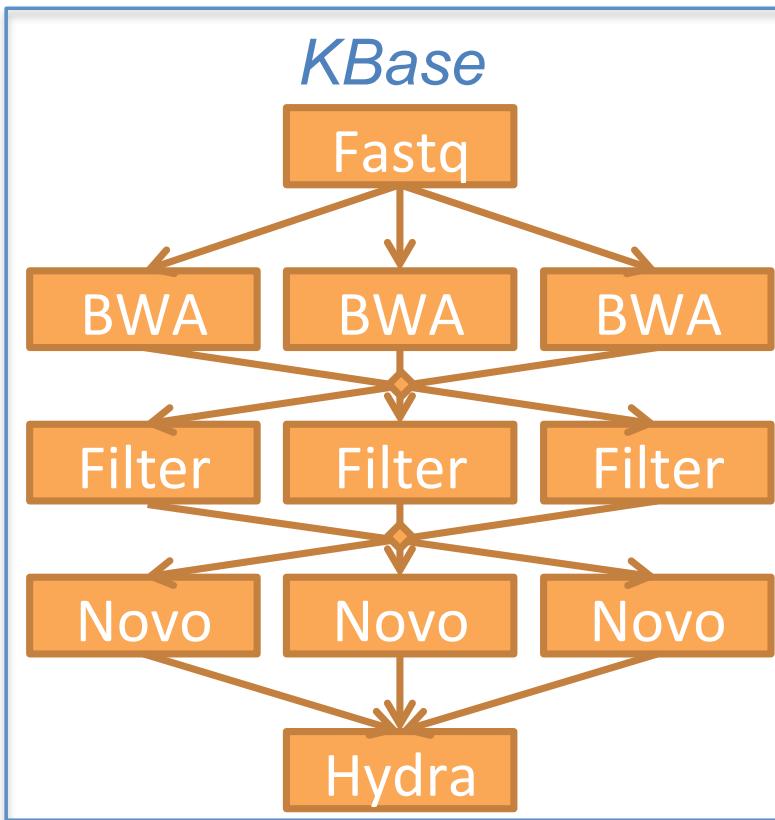
Miscanthus



Switchgrass



From reads through variants to phenotypes



Align and call SNPs from 35M 80bp (14Gbp) reads with maize genome (zmb73v2)

Identified 372k high confidence SNPs

	Serial	Multicore	KBase Cloud
Config	1 core (1 node)	44 core (1 node)	118 cores (15 nodes)
Bowtie2	45 h*	1h 10m	23 m
Sort	2 hr	2 hr	N/A
Samtools	2 hr	2 hr	12 m
End-to-End Speedup	50h* 1x	5h 10m 9.6x	35 m 86x

*estimated time

Maize Population Analysis

Align and call SNPs from 131 maize samples
1 TB fastq / 408Gbp input data

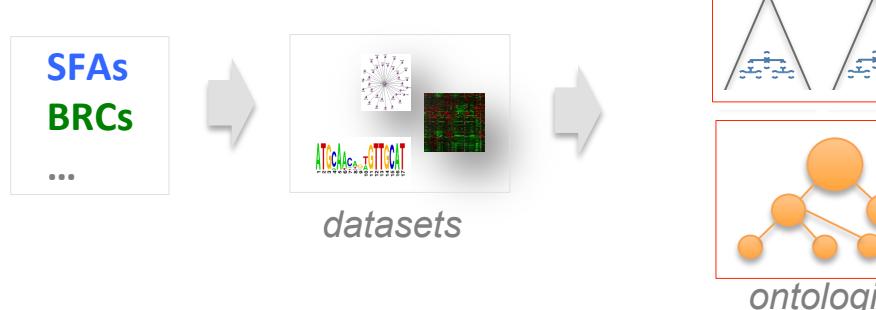
	Serial	KBase cloud (small)	KBase Cloud (large)
Config	1 core (1 node)	210 cores (15 nodes)	854 cores (61 nodes)
Bowtie2	1311 hr*	19.5 hr	5 hr
Sort	58 hr*	N/A	N/A
Samtools	58 hr*	3.5 hr	1.5 hr
End-to-End Speedup	1427 hr* 1x	23 hr 62x	6.5 hr 219x

*estimated time

Networks: Cross-cut example

I. Networks-based knowledge discovery

- Networks-based ontologies



- Gene-ontologies matcher

Algorithms and UI tools to search and analyze best matched network components for a user specified gene set

II. Facilitate building predictive models

- Networks-based reliable orthologs

Plant genomes

co-expression based functional orthologs
(Mutwil et al. Plant Cell 2011)

Microbial genomes

co-localization based reliable orthologs

III. Infrastructure for scientific social networks

- Networks of scientific communities and projects

Support narrative interface. Social networks connecting users and joint projects

- Networks of algorithms similarity



Allows KBase to diversify its menu of algorithms and avoid algorithms producing nearly identical results

- Networks quality control

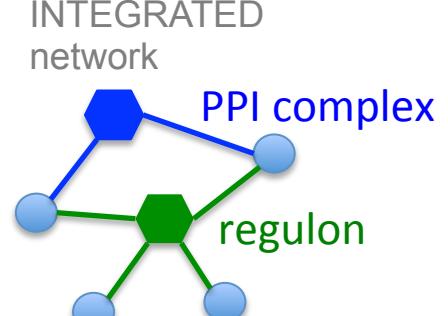
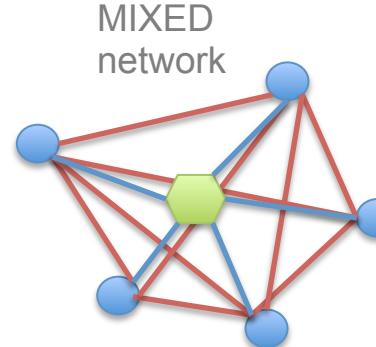
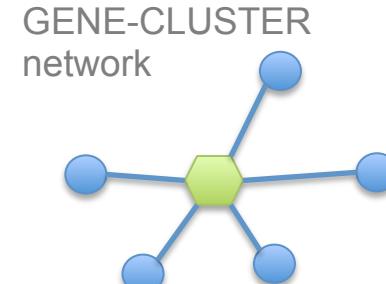
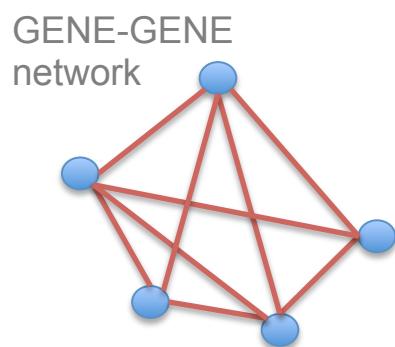
Assign quality measures to KBase networks to allow healthy competition between algorithms and teams

Datasets and Networks API

I. Datasets: collected, processed, classified

KBase domain	Dataset sources	Datasets	Network types	
Microbes	7	5099	REGULATORY_NETWORK	CO_FITNESS
Plants	6	46	PROT_PROT_INTERACTION	CO_EXPRESSION
Communities	1	1	METABOLIC_SUBSYSTEM	FUNCTIONAL_ABUNDANCE
			FUNCTIONAL_ASSOCIATION	PHYLOTYPE_ABUNDANCE

II. Networks API: provide heterogeneous networks in unified format



Networks build methods

```
Network buildFirstNeighborNetwork(  
    list<string> datasetIds,  
    list<string> entityIds,  
    list<EdgeType> edgeTypes)
```

```
Network buildInternalNetwork(  
    list<string> datasetIds,  
    list<string> genelids,  
    list<EdgeType> edgeTypes)
```

Datasets management methods

```
list<Dataset> allDatasets()  
list<DatasetSource> allDatasetSources()  
list<NetworkType> allNetworkTypes()  
...
```

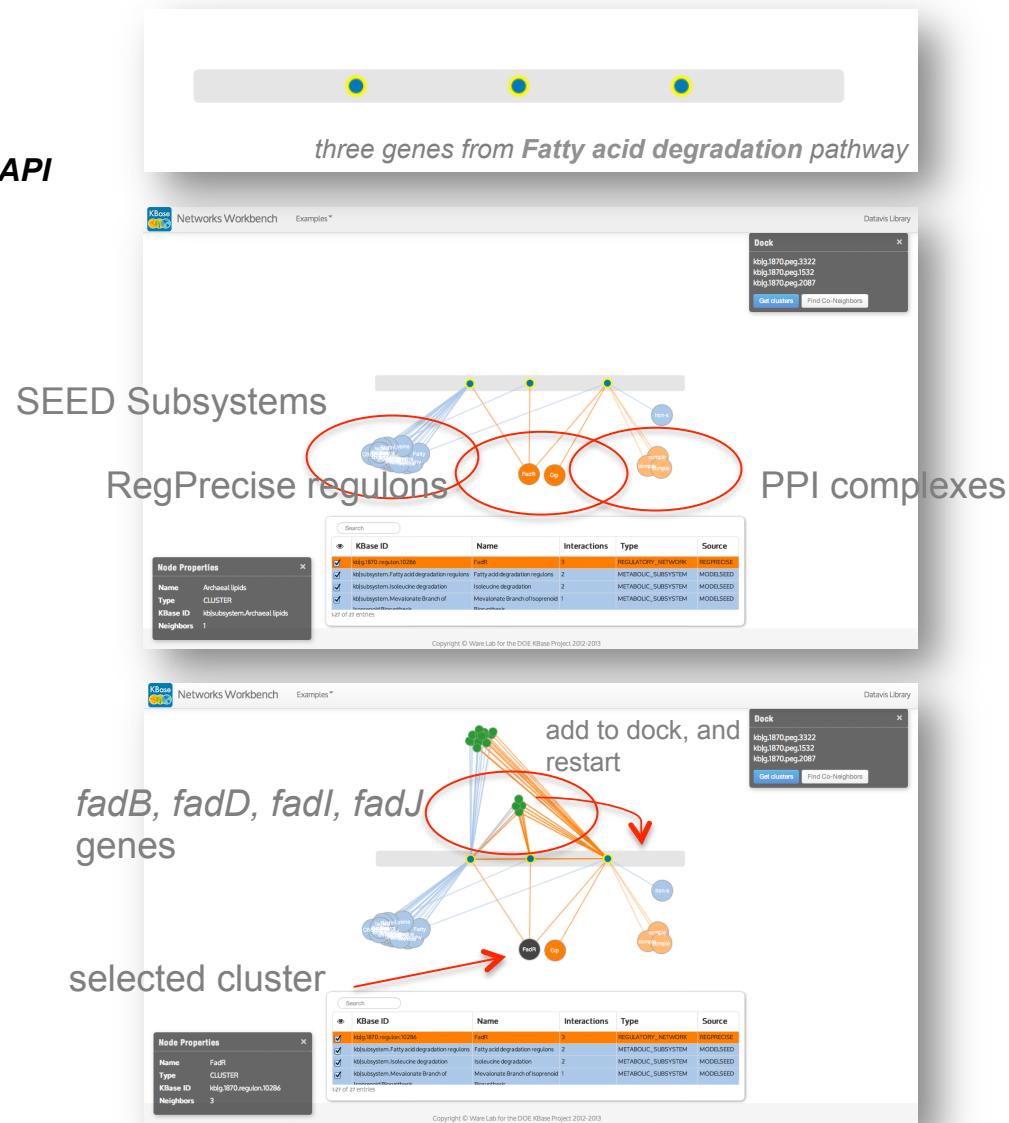
Networks Workbench

Iterative exploration and building network components

powered by the developed **Networks API**

Gene clusters associated with genes from dock panel

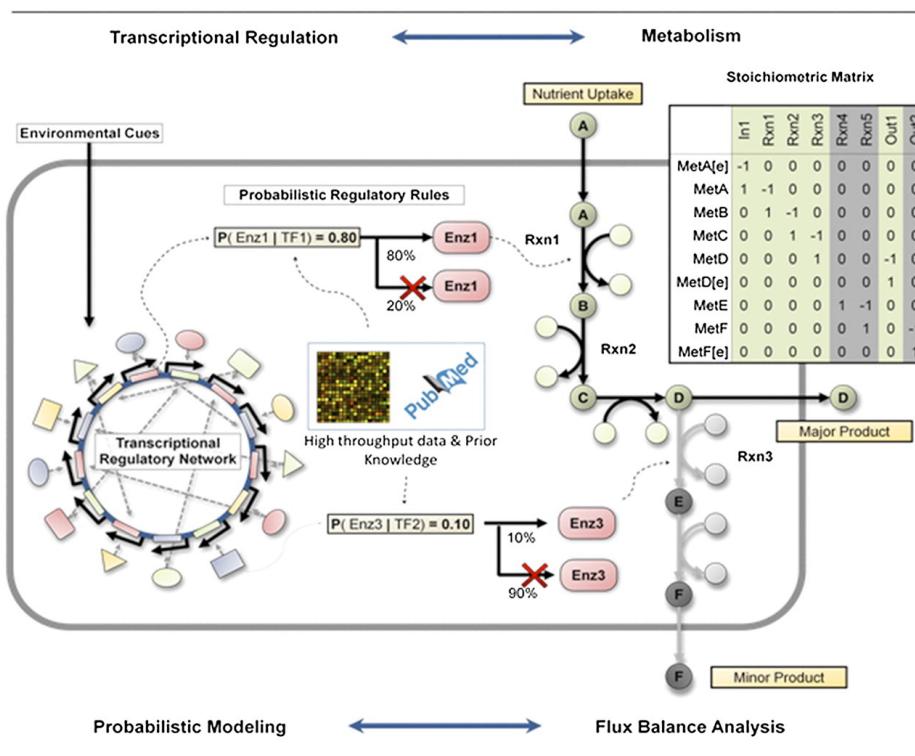
Putative members of the same network component (docked genes)



Working with collaborators early

ENSURING IMPACT

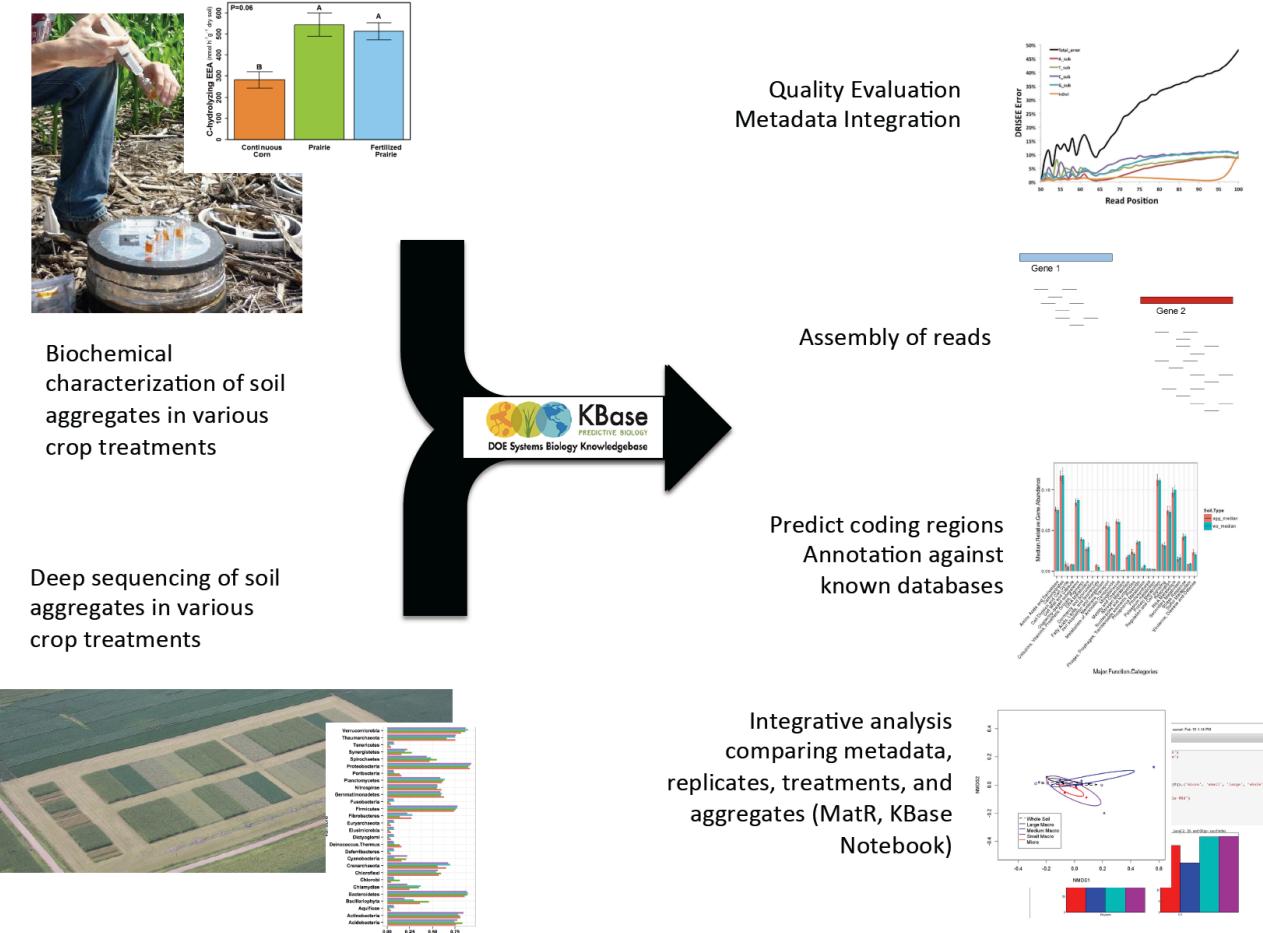
Using KBase framework to enable PROM



- Nathan Price's probabilistic regulation of metabolism (PROM) integrates metabolic models and functional data to make better predictions of growth given genetic or transcription variation.
- Difficult to access and for people to use.
- Previously applied to just two organisms.
- Now a KBase service that can be applied to any genome for which there is expression or variation data.
- Testing on *Shewanella oneidensis* MR-1 with transcription, TF knockout and growth/fitness data available in KBase.



KBase Microbial Communities Analysis Tools to Discover Niche Selection in Crop Soil Aggregates

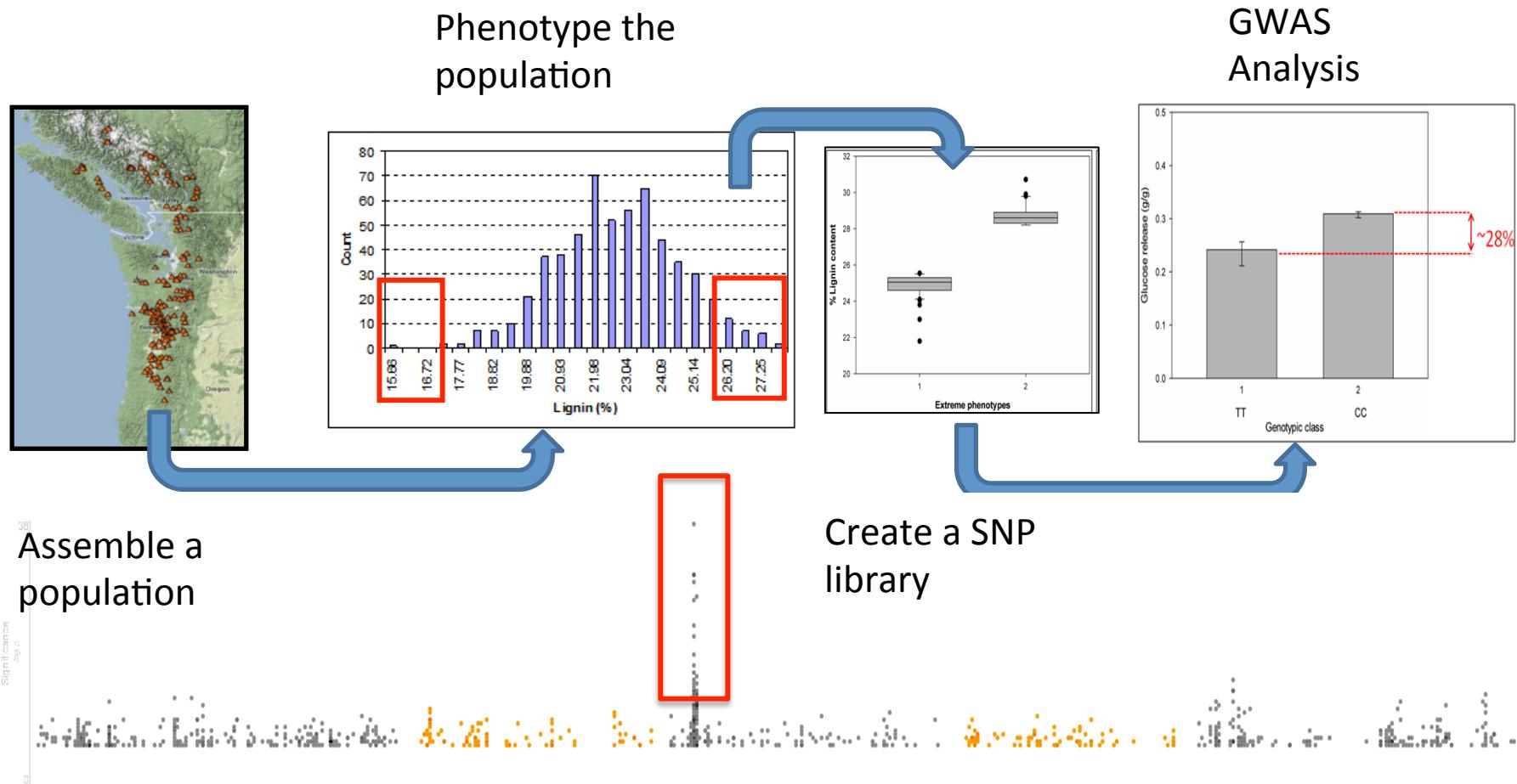


Kirsten Hofmockel, Iowa State, used KBase services to analyze and compare metagenomes from different sized soil aggregates across different crop treatments.

Work with Tuskan (ORNL) & DiFazio (WVU) Identifies important SNPs

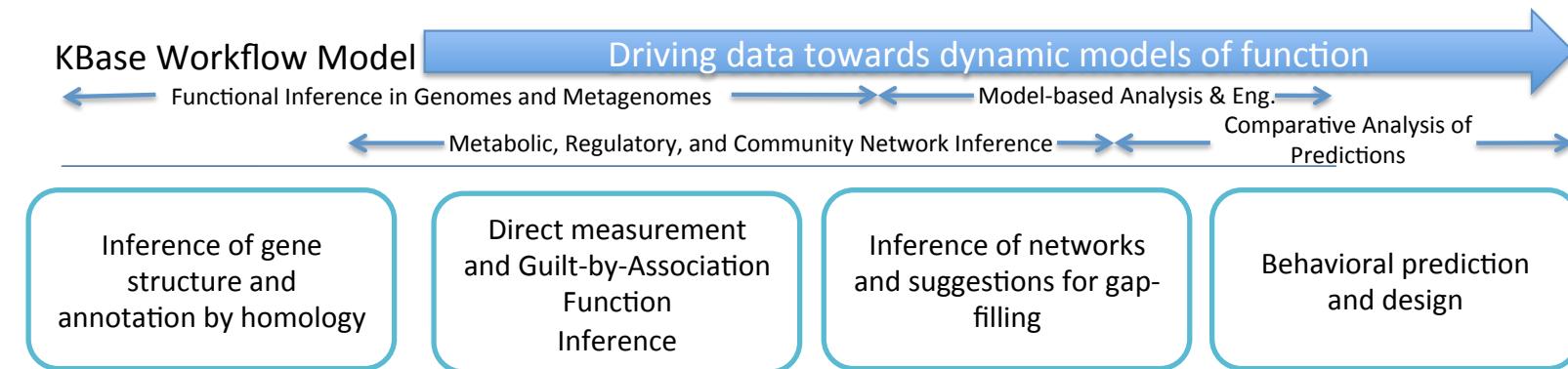
Variation in Lignin Composition and Content

Extreme variation contained in native populations of *Populus* detected in common garden experiments linked to genes using Association Genetics

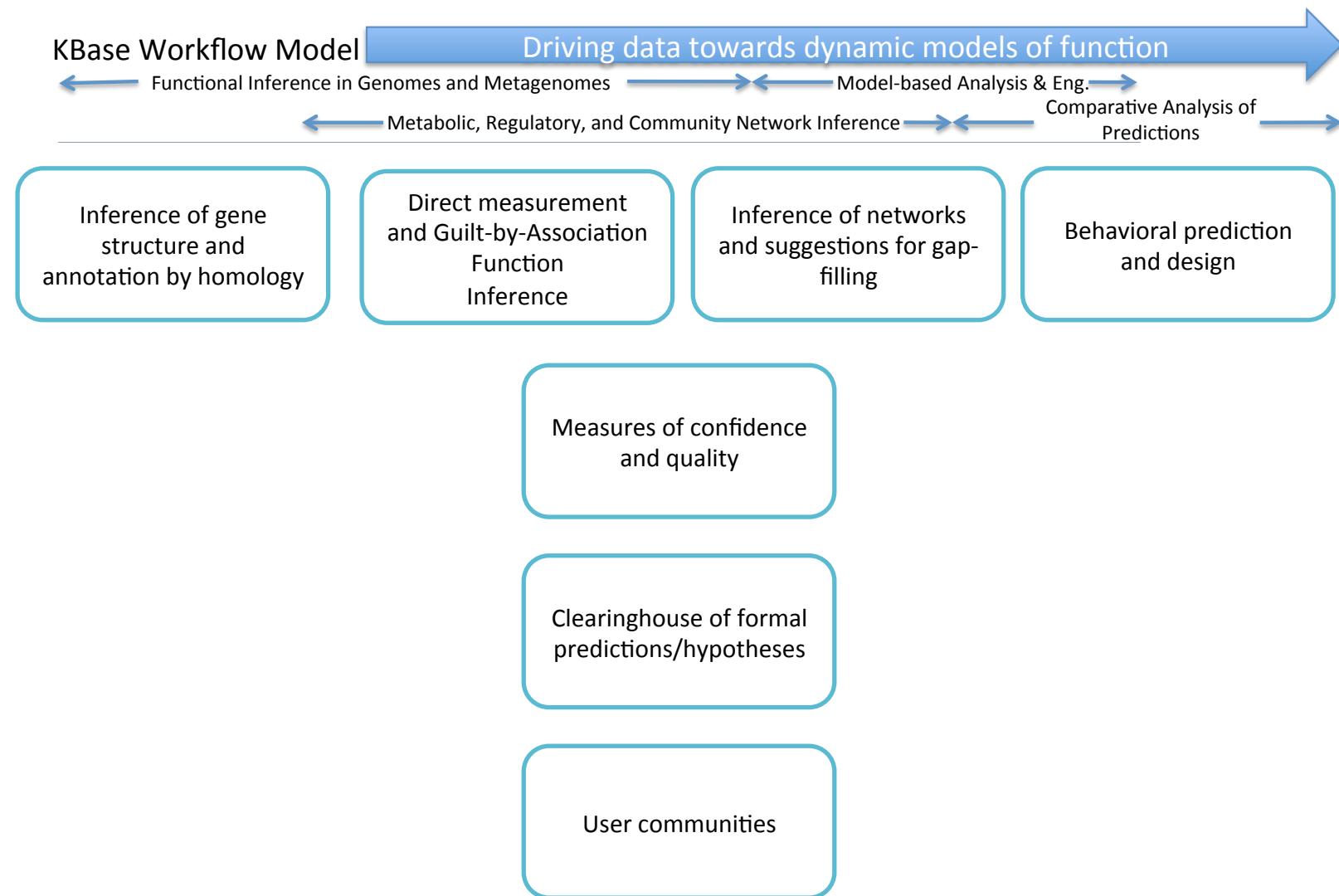


HOW ARE WE BUILDING KBASE?

Define over research “paths”



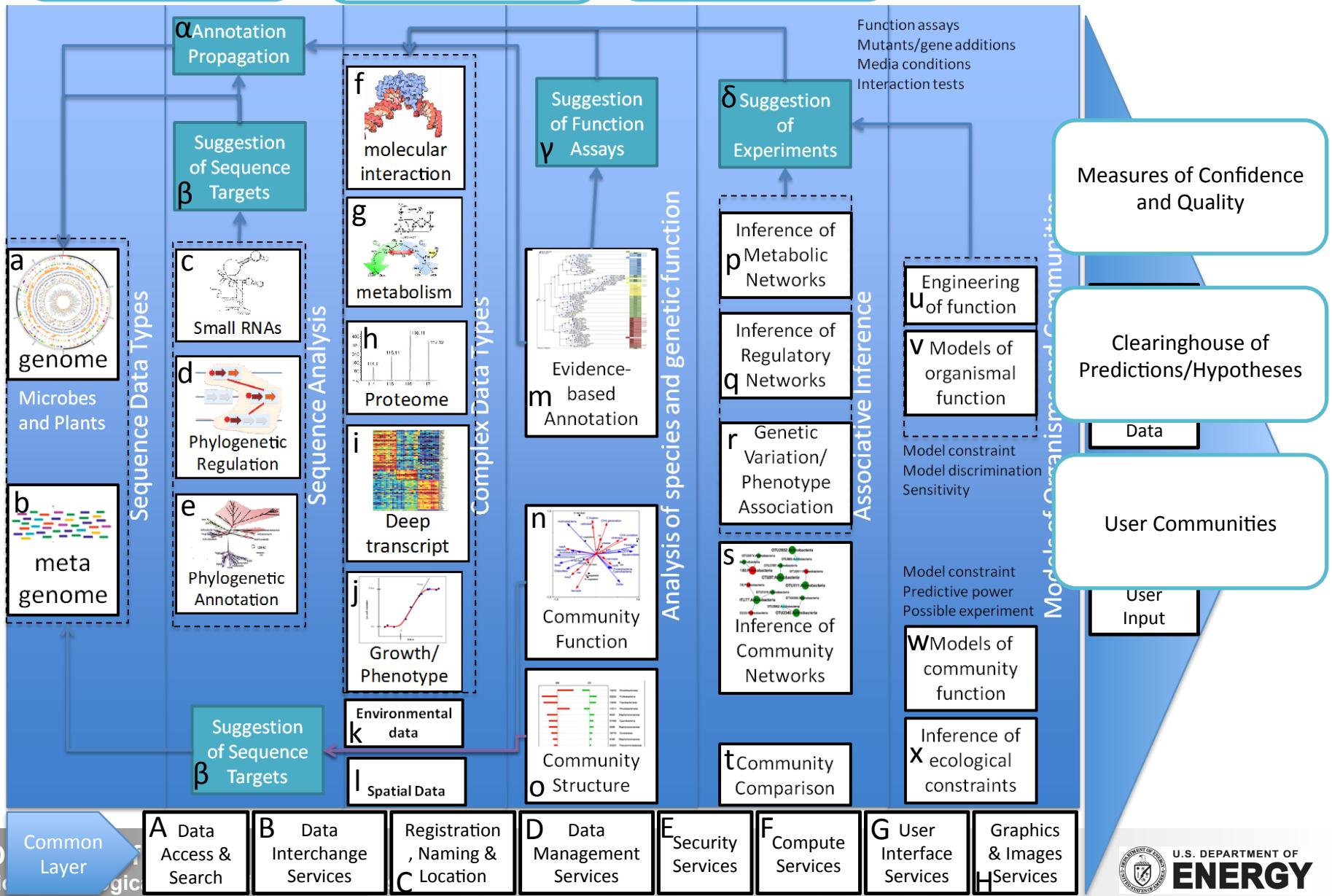
Link design to KBase Concepts



Inference of gene structure and annotation by homology

Direct measurement and Guilt-by-Association Function Inference

Inference of networks and suggestions for gap-filling





Three “Products” highlighted on Home Page

The Department of Energy Systems Biology Knowledgebase (**KBase**) is an emerging software and data environment designed to enable researchers to collaboratively generate, test and share new hypotheses about gene and protein functions, perform large-scale analyses on a scalable computing infrastructure, and model interactions in microbes, plants, and their communities. KBase provides an open, extensible framework for secure sharing of data, tools, and scientific conclusions in predictive and systems biology.

Try KBase Now
Use a web-based command-line interface—no installation necessary

Download the Tools
Install and run KBase command-line tools on your computer

Visit KBase Labs
Sneak a peek at KBase applications in development

KBase includes

- 5695 prokaryotic genomes
- 175 eukaryotic genomes
- 4985 models
- 12 services

Search the database:

[Advanced search](#)

What can KBase do?

- Efficiently annotate new microbial genomes and infer metabolic and regulatory networks.
- Transform network inferences into metabolic models and map missing reactions to genes using novel data reconciliation tools.
- Test microbial ecological hypotheses through taxonomic and functional analysis of quality-assessed metagenomic data
- Discover genetic variations within plant populations and map these to complex organismal traits.

Glimpse the future

[Sign up for a KBase account](#)

KBase is sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research

U.S. DEPARTMENT OF ENERGY
Office of Science

[Acknowledgements](#)
[Privacy and Security](#)

Search this site

[f](#) [t](#) [r](#)

Latest News

KBase Developer's Minor Release
Posted by niharris Jan 18, 2013

KBase at International Plant and Animal Genome XXI
Posted by salazar Jan 09, 2013

KBase Team at Argonne for November Build
Posted by salazar Nov 30, 2012

[View news](#)

Upcoming Events

2013-02-18
[BERAC Presentations](#)

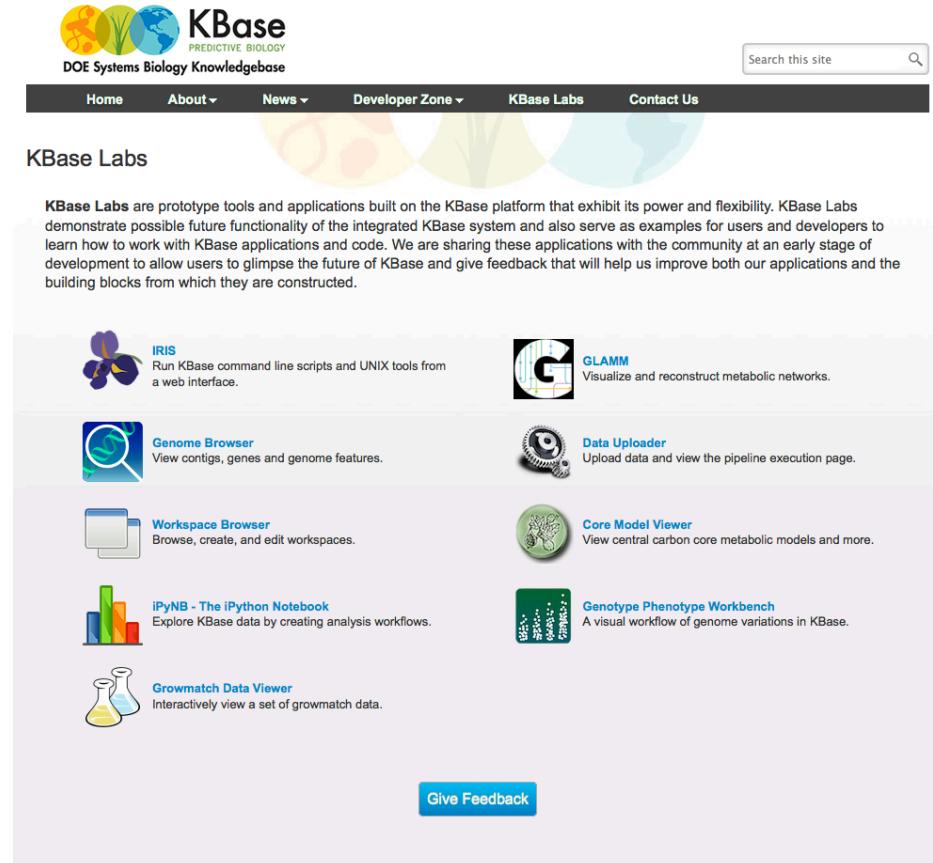
2013-02-22
[Microbes Webinar](#)

2013-02-24
[DOE/NIFA Plant Feedstocks Genomics for Bioenergy](#)

2013-02-24
[Genomic Science Contractors-Grantees Meeting](#)

[View calendar](#)

- Labs are prototypes and applications that demonstrate KBase present and possible future function.
- Early stage functionality both provides utility and a way for the community to provide feedback.



The screenshot shows the KBase Labs page. At the top, there's a navigation bar with links for Home, About, News, Developer Zone, KBase Labs (which is highlighted), and Contact Us. A search bar is also present. Below the navigation, there's a brief introduction about KBase Labs being prototypes built on the KBase platform. The main content area displays seven prototype tools, each with an icon and a brief description:

- IRIS**: Run KBase command line scripts and UNIX tools from a web interface.
- GLAMM**: Visualize and reconstruct metabolic networks.
- Genome Browser**: View contigs, genes and genome features.
- Data Uploader**: Upload data and view the pipeline execution page.
- Workspace Browser**: Browse, create, and edit workspaces.
- Core Model Viewer**: View central carbon core metabolic models and more.
- iPyNB - The iPython Notebook**: Explore KBase data by creating analysis workflows.
- Genotype Phenotype Workbench**: A visual workflow of genome variations in KBase.
- Growmatch Data Viewer**: Interactively view a set of growmatch data.

A "Give Feedback" button is located at the bottom right of the tool list.

IRIS: Command-Line On the Web

Metabolic workflow in IRIS

```

>/ kbfba-loadgenome kb|g.422
Genome successfully loaded to workspace:
>/ kbfba-buildfbamodel kb|g.422
Model successfully generated in workspace:
>/ kbfba-exportfbamodel kb|g.422.fbaml.2 html > model.html
Command completed.

>/ kbfba-addmedia C-D-Mannitol "D-Mannitol;NH4+;Phosphate;Sulfate;Al;Boron;Ca2+;Cl-
;Co2+;Cu2+;Fe2+;Fe3+;H+;H2O;K+;Mg;Mn2+;Ni2+;Molybdate;Na+;O2;Zn2+" -d -type Minimal -e
Successfully added media to workspace:
>/ kbfba-runfba kb|g.422.fbaml.2 -m C-D-Mannitol
Flux balance analysis successful:
>/ kbfba-exportfba kb|g.422.fbaml.2.fba.2 html > fba.html
Command completed.

>/ kbfba-importpheno kb|g.422 WT-phenotypes.txt
Command completed.
>/ kbfba-simpheno kb|g.422.fbaml.2 kb|g.422.phenos.0
Command completed.
>/ kbfba-exportphenosim kb|g.422.fbaml.2.simpheno.1 html > phenosim.html
Command completed.

>/ kbfba-gapfill kb|g.422.fbaml.2 -m C-D-Mannitol
Gapfilling job queued:

```

Directory of services

Command and result history

User enters commands

User data upload

Upload a file

IRIS is a web-based terminal to give you access to the full KBase command line without the need to install anything!

- Command-line environment for programmers
- Available via downloadable Macintosh DMG or Ubuntu (linux) image
- Installs clean environment for KBase application development

Populus reads to SNPs in Five Easy Steps

1. Identify reference genome
`$ all_entities_Genome -f scientific_name | grep -i 'Populus'`
2. Upload Reads to KBase cloud
`$ jk_fs_put_pe populus.1.fq.gz populus.2.fq.gz populus`
3. Align Reads with Bowtie2
`$ jk_compute_bowtie -in=populus.pe -org=populus -out=populus_align`
4. Call SNPs with SAMTools
`$ jk_compute_samtools.snp -in=populus_align -org=populus -out=populus_snps`
5. Merge and Download VCF files
`$ jk_compute_vcf_merge -in=populus_snps --alignments=populus_align -out=populus.vcf`
`$ jk_fs_get populus.vcf`



DOE Systems Biology Knowledgebase

Prototype collaborative workspaces

View Objects Produced by Analysis in an Online Collaborative Workspace



DOE Systems Biology Knowledgebase

chenry
[logout](#)

Workspaces ▾	
19	chenrydemo
~ 10	edwardsCollaboration
~ 4	kbase
~ 98	KBaseCDMGenomes
~ 98	KBaseCDMMModels
~ 0	liberibacter
~ 8	ModelingBootcamp
~ 1065	PamGenomes
129	phenotypeDemo
~ 4	PublishedFBAModels
~ 1	SEEDGenomes
~ 10	TaylorCollaboration
~ 4	test
+ create	

10 ▾ records per page						
Workspace	ID	Type	Command	Owner	Modified	
chenrydemo	kb g.422.fbamdl.2.fba.3 (0)	FBA	runfba	chenry	2013-02-19T01:58:45	
chenrydemo	kb g.422.fbamdl.2.fba.0 (0)	FBA	runfba	chenry	2013-02-19T01:58:45	
chenrydemo	kb g.2793.fbamdl.0.fba.4 (0)	FBA	runfba	chenry	2013-02-19T01:58:45	
chenrydemo	kb g.422.fbamdl.2.fba.2 (0)	FBA	runfba	chenry	2013-02-19T01:58:45	
chenrydemo	kb g.2793.fbamdl.0.fba.2 (0)	FBA	runfba	chenry	2013-02-19T01:58:45	
chenrydemo	kb g.2793.fbamdl.0.fba.0 (0)	FBA	runfba	chenry	2013-02-19T01:58:45	
chenrydemo	kb g.2793.fbamdl.0.fba.3 (0)	FBA	runfba	chenry	2013-02-19T01:58:45	
chenrydemo	kb g.2793 (1)	Genome	genome_to_workspace	chenry	2013-02-19T01:58:45	
chenrydemo	83333.1 (0)	Genome	genome_to_workspace	chenry	2013-02-19T01:58:45	
chenrydemo	kb g.422 (3)	Genome	genome_to_workspace	chenry	2013-02-19T01:58:45	

- Object model supports provenance, meaning all previous versions of the objects can be retrieved
- Workspace browser provides direct links to view any object contained in the workspace

Visualizing Metabolic Models and FBA in KBase

All views may be used for all public and private models and FBA solutions

Overview Compartments Compounds Reactions Biomass reactions Gapfilling Gapgen

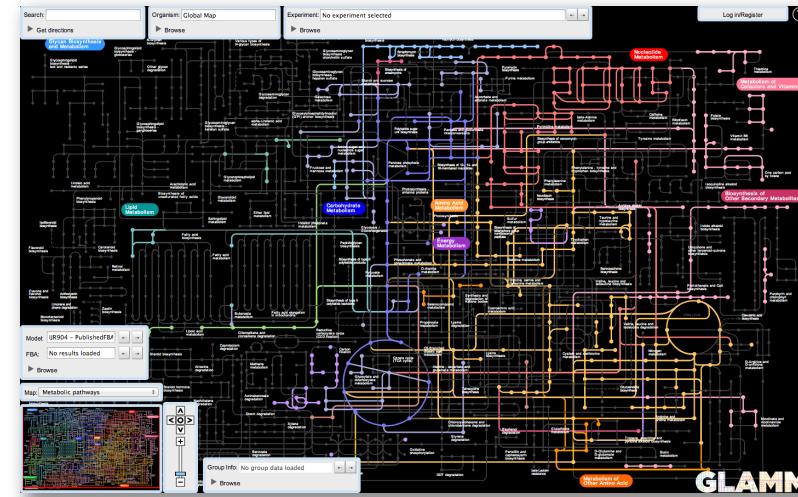
Biomass

	DNA	RNA	Protein	Cellwall	Lipid	Cofactor	Energy
bio1	[0.031]	0.21	0.563	0.177	0.093	0.039	40

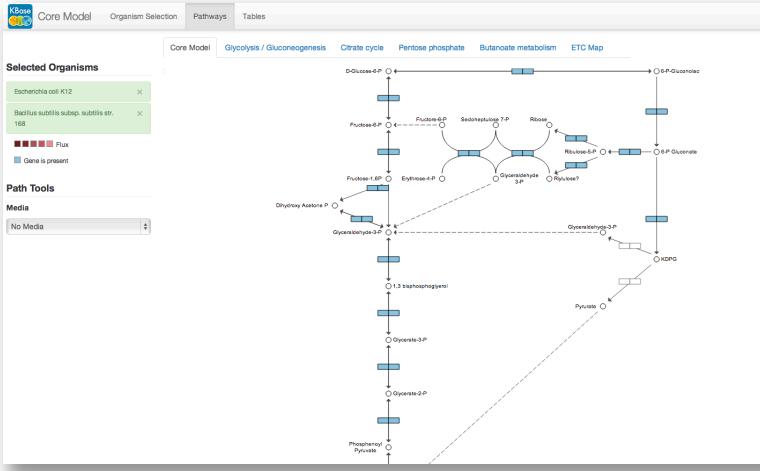
Biomass

	Model compound	Name	Compartment	Coefficient
bio1	cpx00063_c0	Ca2+-c0	c0	-7.78132482043096e-06
bio1	cpx00156_c0	L-Valine_c0	c0	-0.0003528894949968271
bio1	cpx00300_c0	Mn2+-c0	c0	-7.78132482043096e-06
bio1	cpx00301_c0	K+-c0	c0	-7.78132482043096e-06
bio1	cpx00315_c0	L-Alanine_c0	c0	1.8004173289153499
bio1	cpx00360_c0	L-Methionine_c0	c0	-0.0001280371599737
bio1	cpx00366_c0	L-Phenylalanine_c0	c0	-0.0001548076087483
bio1	cpx00317_c0	S-Adenosyl-L-methionine_c0	c0	-7.78132482043096e-06
bio1	cpx00010_c0	CoA_c0	c0	-7.78132482043096e-06
bio1	cpx15665_c0	Peptidoglycan polymer (n subunits)_c0	c0	-6.09084652443221e-05
bio1	cpx00052_c0	dCTP_c0	c0	-8.41036156544863e-05
bio1	cpx10516_c0	fe3_c0	c0	-7.78132482043096e-06
bio1	cpx00099_c0	Cl-_c0	c0	-7.78132482043096e-06
bio1	cpx00356_c0	dCTP_c0	c0	-3.2024793384298e-05
bio1	cpx10515_c0	Fe2+-c0	c0	-7.78132482043096e-06
bio1	cpx00254_c0	Mg_c0	c0	-7.78132482043096e-06
bio1	cpx00058_c0	L-Glutamine_c0	c0	-0.000324939141304
bio1	cpx00058_c0	Cu2+-c0	c0	-7.78132482043096e-06
bio1	cpx00149_c0	Co2+-c0	c0	-7.78132482043096e-06
bio1	cpx00041_c0	L-Aspartate_c0	c0	-0.000201205267995816
bio1	cpx10743_c0	RNA transcription_c0	c0	-1
bio1	cpx00023_c0	L-Glutamate_c0	c0	-0.00021949655995436
bio1	cpx00053_c0	L-Glutamine_c0	c0	-0.00021949655995436
bio1	cpx00107_c0	L-Leucine_c0	c0	-0.00037608782528765
bio1	cpx00220_c0	Riboflavin_c0	c0	-7.78132482043096e-06
bio1	cpx00054_c0	L-Serine_c0	c0	-0.000179790960093822
bio1	cpx00065_c0	L-Tryptophan_c0	c0	-4.72899299502361e-05
bio1	cpx00065_c0	phosphate_c0	c0	-6.09084652443221e-05

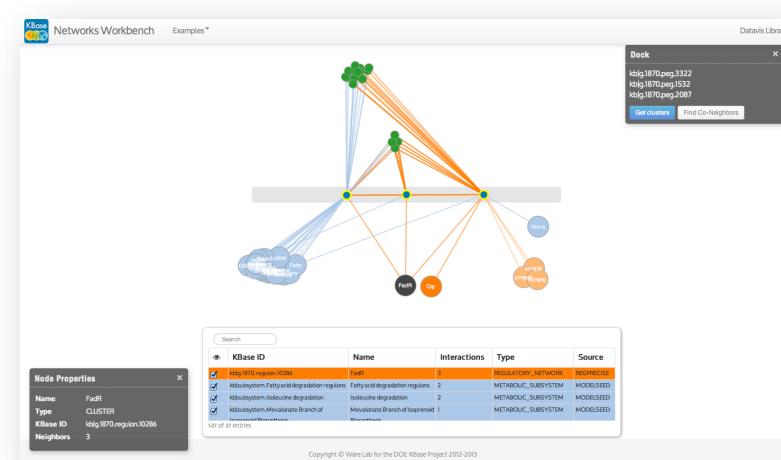
Simple Tabular Views



Interactive metabolic maps and models

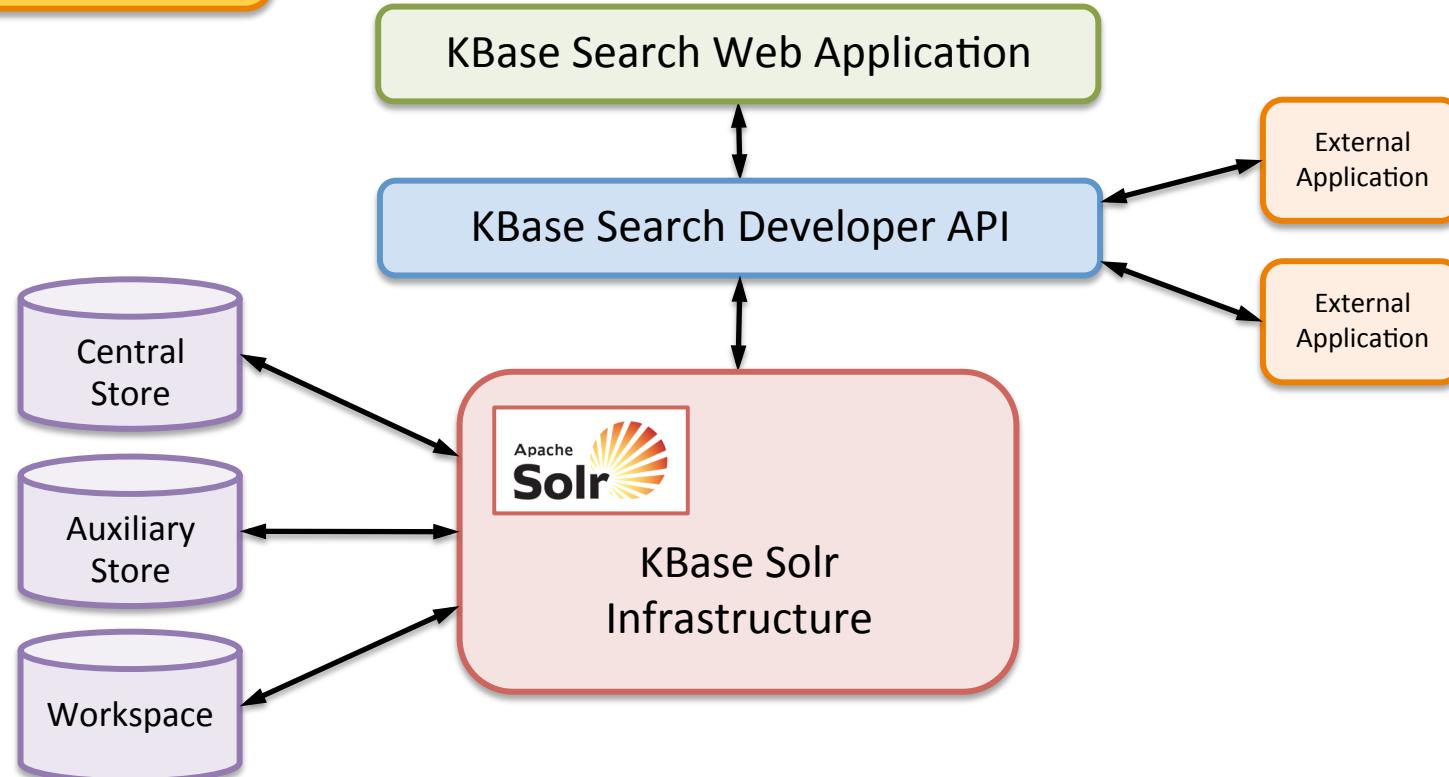


Comparative Model Viewer



Network Analysis Workbench

Search Architecture



Search with keyword: **thermocellum**

clostridium

Search Advanced

KBase Search

thermocellum

Search

Advanced

results organized
using central data
model

Searched for **thermocellum** found **28** results.
Now showing 10 results.

Results

Literature

Pubmed (20)

Genomes

Bacteria (7)

Viruses (0)

Eukaryota (0)

Archaea (0)

Function

Genes (1)

Locus (0)

Prophage (0)

Pseudo Gene (0)

Regulation

Promoter (0)

Operator (0)

Protein Binding Sites (0)

1. Thermostable chaperonin from Clostridium thermocellum.

Literature Published Date: Mar 12, 2012 Pubmed Link: [PUBMED](#)

Abstract →

2. Nucleotide sequence of the cellulase gene celF of Clostridium thermocellum.

Literature Published Date: Mar 11, 2012 Pubmed Link: [PUBMED](#)

Abstract →

3. Nucleotide sequence and deletion analysis of the cellulase-encoding gene celH of Clostridium thermocellum.

Literature Published Date: Mar 11, 2012 Pubmed Link: [PUBMED](#)

Abstract →

...

link to external
data in PUBMED



DOE Systems Biology Knowledgebase

Genome page

KBase Search

thermocellum

Search

Advanced

Clostridium thermocellum ATCC 27405

Bacteria

RNA Features

64

Protein Encoding Genes

3466

Contigs

1

GC Content

38.9865899132022

Genome Size

3843301

Complete

✓

links to
features

genome
details

Subsystem Contigs Genes RNAs

Subsystem

#	Function In Cell	Version
1	Restriction-Modification System	2.x
2	Glutathione-regulated potassium-efflux system and associated functions	9.71
3	Queuosine-Archeosine Biosynthesis	1.111
4	Experimental tye	4
5	Methylglyoxal Metabolism	9.0
6	Ribospme LSU Symbiont	1
7	Synthesis of osmoregulated periplasmic glucans	1.x

...

Gene page

KBase Search

thermocellum

Carbon storage regulator

Feature

Genome
Clostridium thermocellum ATCC 27405

KBase Id
kb|g.222.peg.1005

DNA Size
228

Location
kb|g.222.c.0.2667730,-228

[Literature](#) [Annotation](#) [Subsystem](#) [Co-expressed](#) [Co-occurrences](#) [DNA Sequence](#) [Protein Sequence](#)

links to related content

gene details

Subsystem

1. Carbon storage regulator
Carbon storage regulator

Co-occurrences

#	Id 1	Id 2	Occurrences
1	kb g.222.peg.1005	kb g.222.peg.1459	26
2	kb g.222.peg.1005	kb g.222.peg.1214	172
3	kb g.222.peg.1005	kb g.222.peg.1848	49
4	kb g.222.peg.1005	kb g.222.peg.1957	164
5	kb g.222.peg.1005	kb g.222.peg.1919	164

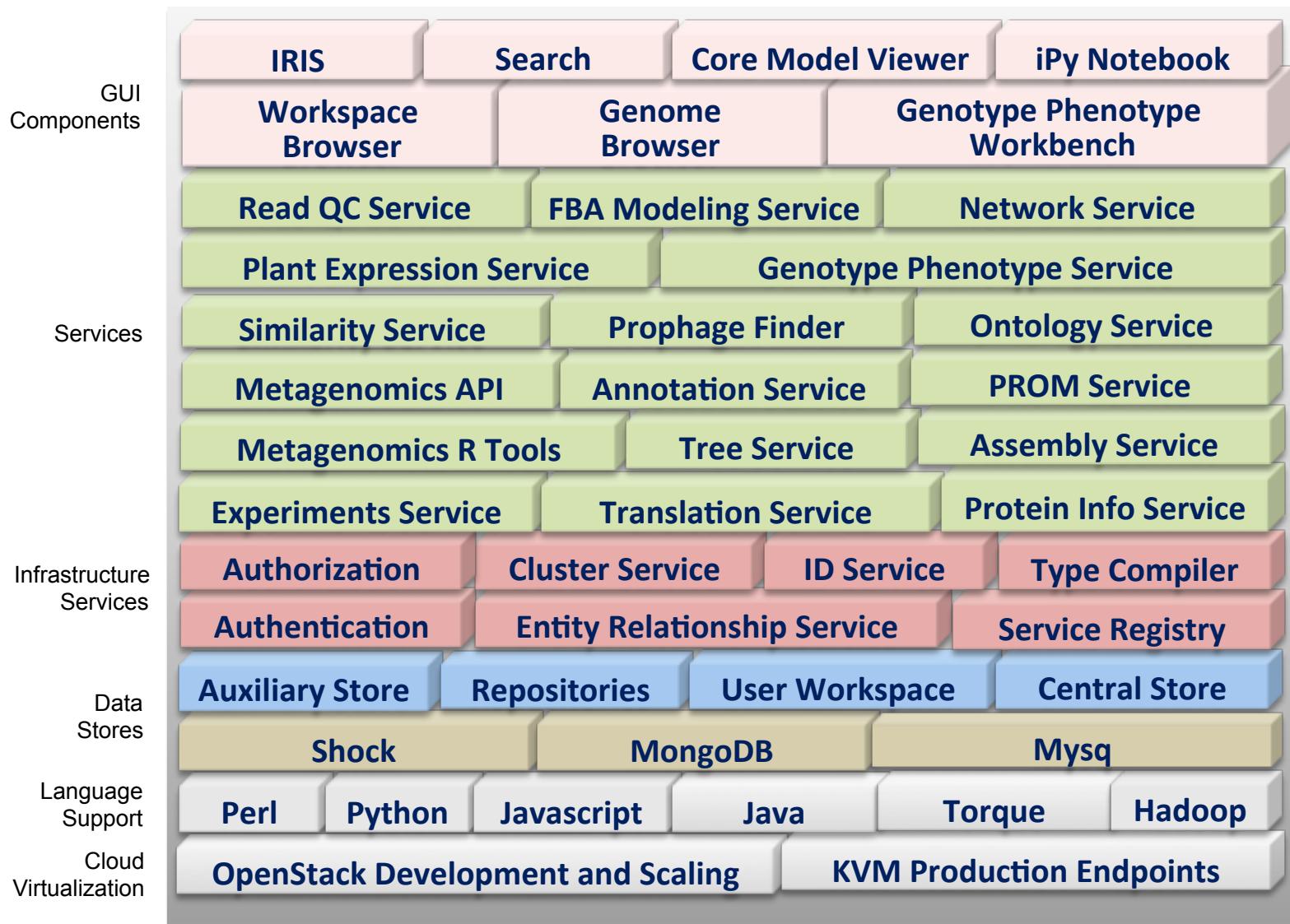
DNA Sequence

GTGCTGTTCTGACGAAAGAAAAATGAATCCATAATTAAATGACAATATTGAGATTACTGTTGACATTCAAGGGAGC
AGGTGCGTATCGGCATAATGCCCGAAAGCATATCTATTTACAGGAAGGAGATAACCTGAAATACAGGCTGAAAACAAAA
AGGCTGCGAAATTAAAGATGTGGATCTGAAGGAGGACTTAAGGATTGGAGTAA

Protein Sequence

MLVLTRKKNESIINDNIEITVVDIQLGEQVRGINAPKSISIYRKEIYLEIQAENKKAAEIKNVDLKEDLKDFLK

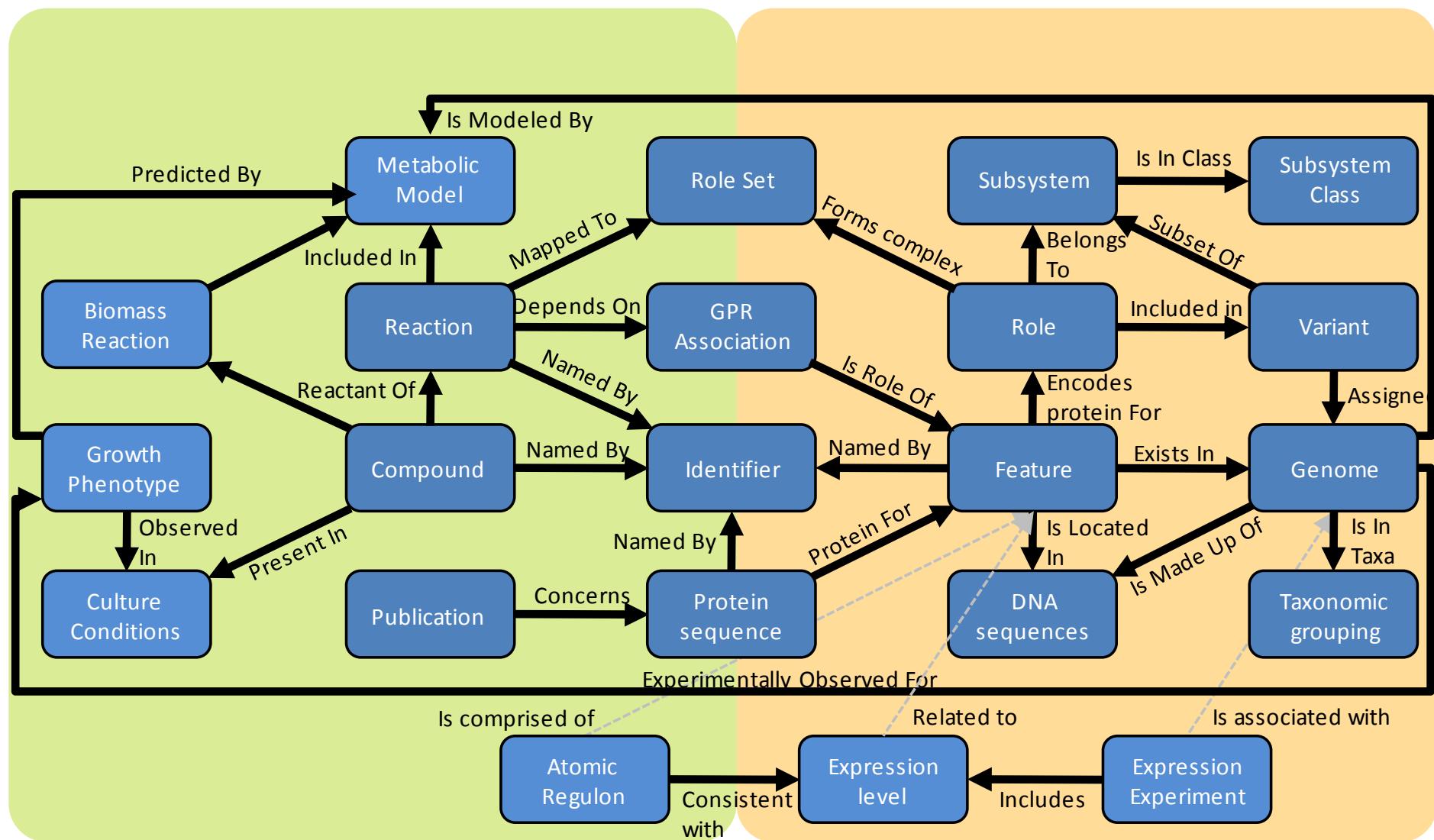
Initial Platform Modules



Initial KBase Services

- IRIS
- KBase Notebook
- Auxiliary Store Service
- Communities API
- Metagenomics Analysis Tools R
- QC Service
- Annotation Service
- Central Store
- Core Model Viewer
- Experiment data
- FBA Modeling
- Metabolic Map Viewer
- Microbes Model Builder
- Phispy
- Probabilistic Annotation
- Protein Info service
- Regulation Service
- Similarity Service
- Translation Service
- Workspace Service
- KB Model Seed
- Tree Service
- Assembly Service
- Authorization Service
- Network Service
- Genotype Phenotype Service
- Genotyping Service
- Ontology Service Plant
- Expression Service
- Authentication and Authorization Client
- Cluster Service
- ERDB service
- File Type Service
- ID Service
- Registry
- Type Compiler

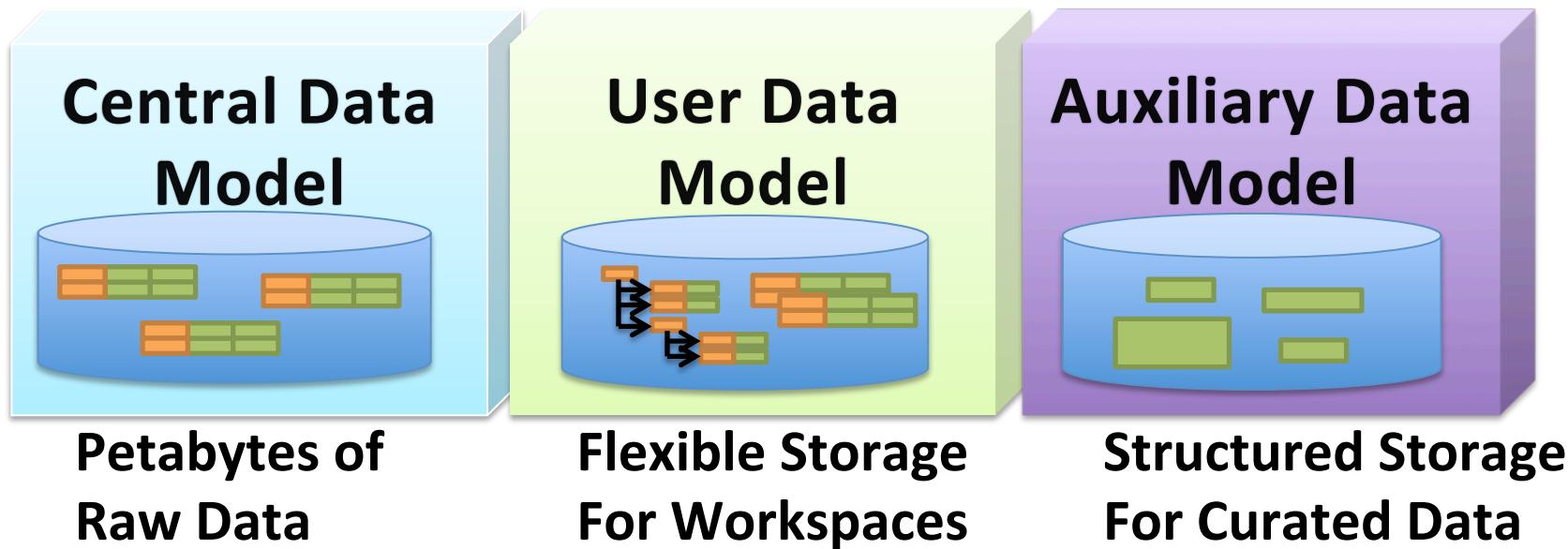
KBase Central Data Model



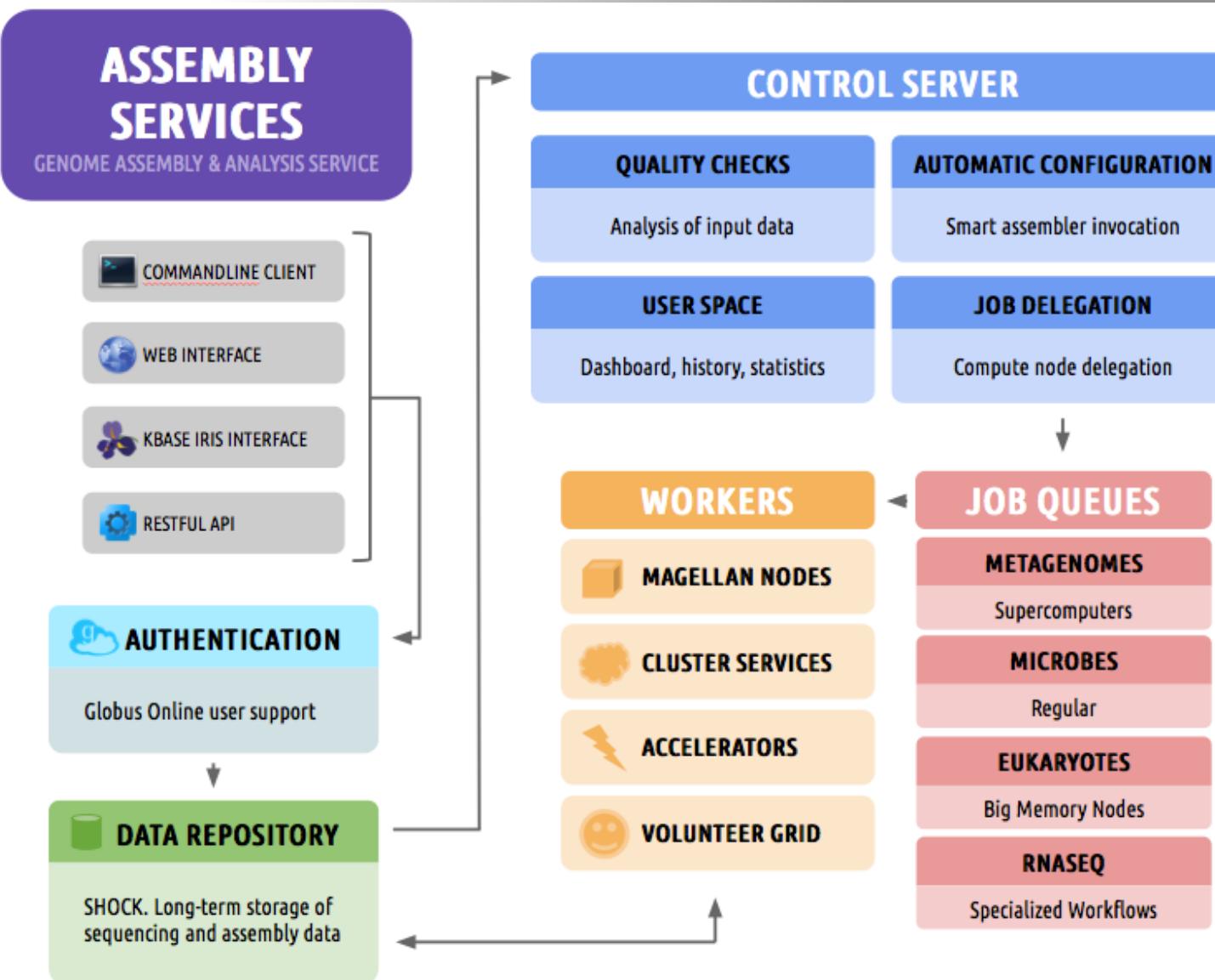
The KBase Data Stores

Data Stores: Storing a diverse range of biological data requires multiple approaches:

- Highly structured data in relational databases
- Frequently changing user data
- Large bulk data

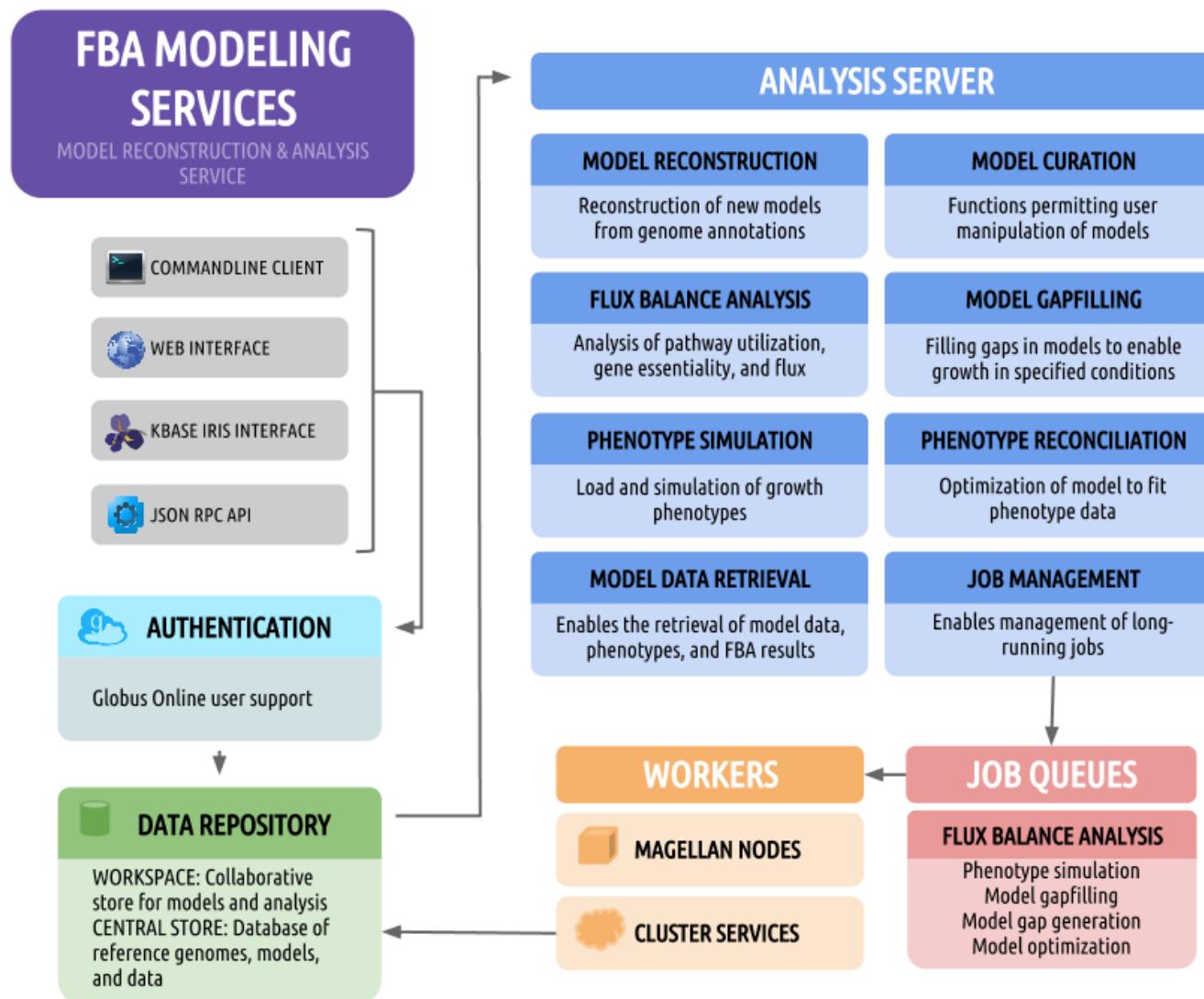


The KBase Assembly Service





Flux Balance Analysis Modeling Services





Helping users and developers manage complexity

Website and Outreach

KBase Training Resources

Several types of training are being offered, either as in-person sessions or as web-based resources:

- **User Manuals** are online procedural documents describing each of the services with detailed information about functionality and usage.
- **Tutorials** are both online training resources and in-person sessions* focused on using specific aspects of KBase. Online resources are in the form of written walk-throughs and video tutorials. In-person tutorials will typically be 2-3 days, and may be immediately followed by a 1-day workshop as part of the same agenda. In-person tutorials are bimonthly.
- **Workshops** are in-person sessions* focused on scientific problems. We will guide you through the use of KBase to accomplish tasks that are of interest to you. Currently, workshops will be single-day sessions that are paired with relevant tutorials (i.e., offered at the end of a tutorial session). Later in the project, workshops will be stand-alone, multiple-day events held annually.
- **Bootcamps** are in-person sessions focused on developing computational tools in KBase. Join our developers and help build KBase.

plus ... Monthly webinars



When and Where

Scheduled KBase Events can be found at <http://kbase.us>

Calendar

KBase Public Calendar

Today [◀](#) [▶](#) February 2013 [▼](#)

[Print](#)

Sun	Mon	Tue	Wed	Thu	Fri	Sat
27	28	29	30	31	Feb 1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
BERAC Presentations					10am Microbes	
24	25	26	27	28	Mar 1	2
DOE/NIFA Plant Feedstocks Genomics for Bioenergy Genomic Science Contractors-Grantees Meeting				10am Microbes		

Events shown in time zone: Pacific Time [+ Google Calendar](#)

Upcoming Training Sessions

Bootcamps:

February 12-13, 2013: Full
March 6-7, 2013: Open
April 24-25, 2013: Open

Tutorials:

February 11-13, 2013: Full
March 11-13, 2013: Full
May 6-8, 2013: Open, topics include IRIS for Microbes, IRIS for Plants, and matR for Communities
July 15-17, 2013: Open, topics TBD

Workshops:

Requires concurrent registration in a Tutorial Session.
Feb 14, 2013: Full
March 14, 2013: Full
May 9, 2013: Open, topics TBD
July 18, 2013: Open, topics TBD

Contact us at outreach@kbase.us to register for a session.

Online Resources

- User manuals and tutorials
- Video tutorials (coming soon)
- Developer documentation
- Calendar of events
- FAQs
- Press

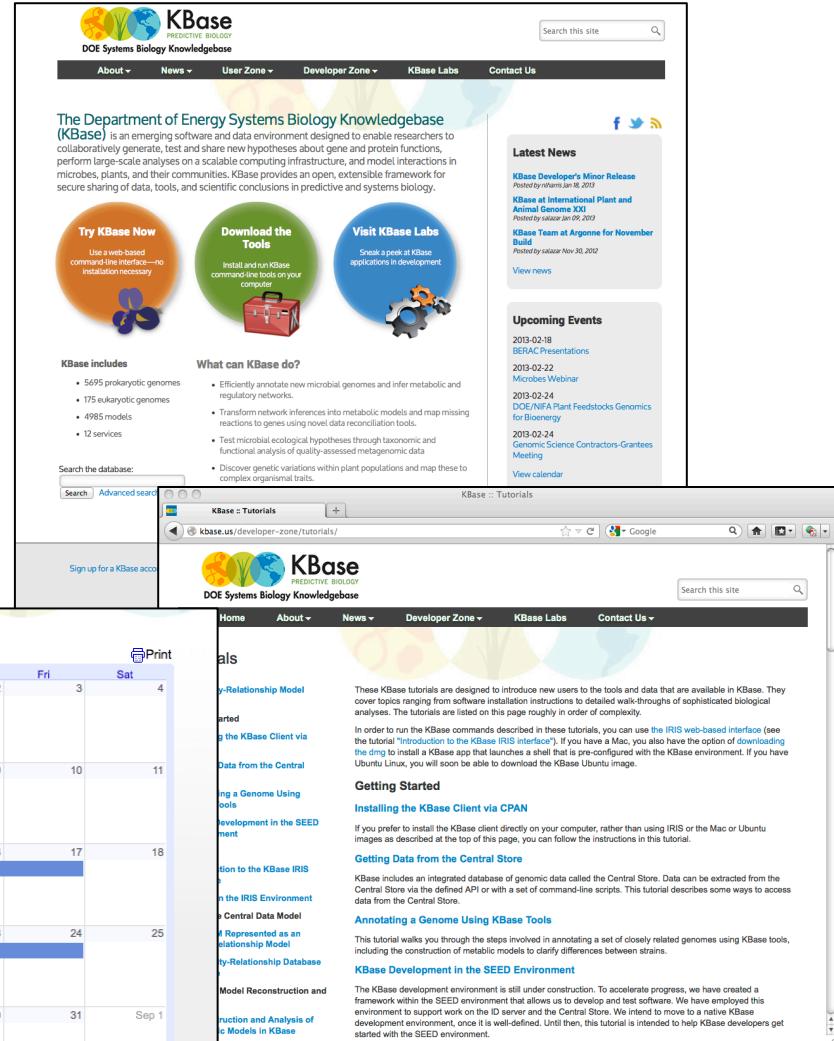
help@kbase.us

KBase Public Calendar

Today | [◀](#) [▶](#) August 2012 ▾

Sun	Mon	Tue	Wed	Thu	Fri	Sat
29	30	31	Aug 1	2	3	4
Sergei Maslov at the DOE Office of Science Graduate Fellow						
5	6	7	8	9	10	11
12	13	14	15	16	17	18
KBase August Build						
19	20	21	22	23	24	25
Paramvir Dehal and Dylan Chivian at ISME-14						
26	27	28	29	30	31	Sep 1
Events shown in time zone: Pacific Time						

[Print](#) [Google Calendar](#)



The screenshot shows the KBase website with several key features highlighted:

- Homepage:** Features a banner about the Department of Energy Systems Biology Knowledgebase (KBase), a "Try KBase Now" button, a "Download the Tools" button, and a "Visit KBase Labs" button.
- Latest News:** A sidebar with news items like "KBase Developer's Minor Release" and "KBase at International Plant and Animal Genome XXI".
- Upcoming Events:** A sidebar with events such as "BERAC Presentations" and "Microbes Webinar".
- Tutorial Page:** Shows a detailed "KBase Tutorials" page with sections like "Getting Started", "Installing the KBase Client via CPAN", "Getting Data from the Central Store", and "Annotating a Genome Using KBase Tools".



New Site for Easy Navigation

The Department of Energy Systems Biology Knowledgebase (KBase) is an emerging software and data environment designed to enable researchers to collaboratively generate, test and share new hypotheses about gene and protein functions, perform large-scale analyses on a scalable computing infrastructure, and model interactions in microbes, plants, and their communities. KBase provides an open, extensible framework for secure sharing of data, tools, and scientific conclusions in predictive and systems biology.

Try KBase Now
Use a web-based command-line interface—no installation necessary

Download the Tools
Install and run KBase command-line tools on your computer

Visit KBase Labs
Sneak a peek at KBase applications in development

KBase includes

- 5695 prokaryotic genomes
- 175 eukaryotic genomes
- 4985 models
- 12 services

Search the database:

[Advanced search](#)

What can KBase do?

- Efficiently annotate new microbial genomes and infer metabolic and regulatory networks.
- Transform network inferences into metabolic models and map missing reactions to genes using novel data reconciliation tools.
- Test microbial ecological hypotheses through taxonomic and functional analysis of quality-assessed metagenomic data
- Discover genetic variations within plant populations and map these to complex organismal traits.

Glimpse the future

[Sign up for a KBase account](#)

KBase is sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research

U.S. DEPARTMENT OF ENERGY Office of Science

Acknowledgements
[Privacy and Security](#)

Latest News

KBase Developer's Minor Release
Posted by niharris Jan 18, 2013

KBase at International Plant and Animal Genome XXI
Posted by salazar Jan 09, 2013

KBase Team at Argonne for November Build
Posted by salazar Nov 30, 2012

[View news](#)

Upcoming Events

2013-02-18
[BERAC Presentations](#)

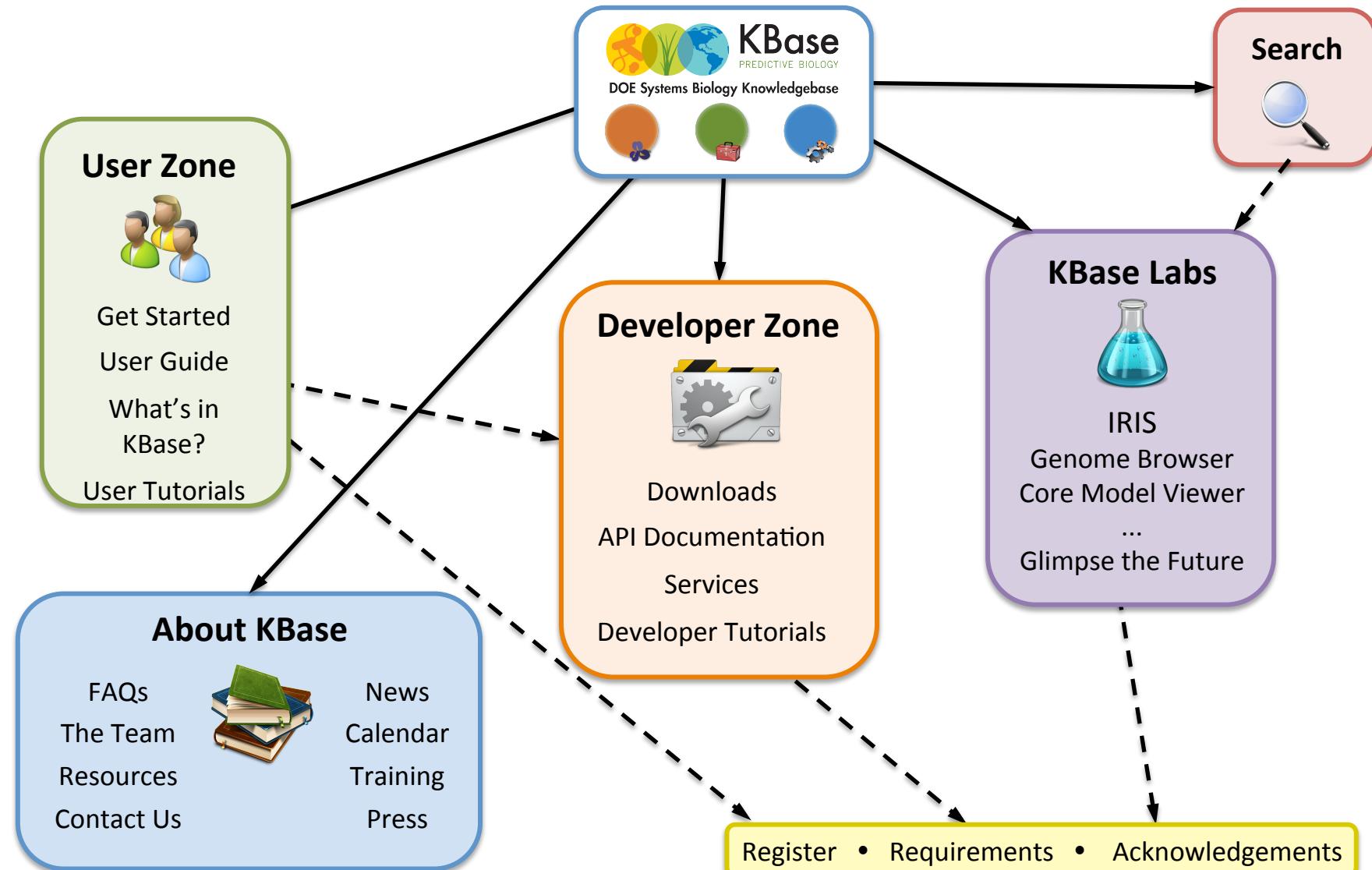
2013-02-22
[Microbes Webinar](#)

2013-02-24
[DOE/NIFA Plant Feedstocks Genomics for Bioenergy](#)

2013-02-24
[Genomic Science Contractors-Grantees Meeting](#)

[View calendar](#)

Sitemap for revamped kbase.us website





Extensive Tutorials and Documentation

KBase
PREDICTIVE BIOLOGY
DOE Systems Biology Knowledgebase

Search this site

Home About ▾ News ▾ Developer Zone ▾ KBase Labs Contact Us

Tutorials

[CDM Entity-Relationship Model Browser](#)

Getting Started

[Installing the KBase Client via CPAN](#)

[Getting Data from the Central Store](#)

[Annotating a Genome Using KBase Tools](#)

[KBase Development in the SEED Environment](#)

Iris

[Introduction to the KBase IRIS interface](#)

[RAST2 in the IRIS Environment](#)

The KBase Central Data Model

[The CDM Represented as an Entity-Relationship Model](#)

[The Entity-Relationship Database Package](#)

Metabolic Model Reconstruction and Analysis

[Reconstruction and Analysis of Metabolic Models in KBase](#)

[Command Line Scripts](#)

[Accessing Central Store Data](#)

[Getting What You Need from the](#)

These KBase tutorials are designed to introduce new users to the tools and data that are available in KBase. They cover topics ranging from software installation instructions to detailed walk-throughs of sophisticated biological analyses. The tutorials are listed on this page roughly in order of complexity.

In order to run the KBase commands described in these tutorials, you can use [the IRIS web-based interface](#) (see the tutorial "[Introduction to the KBase IRIS interface](#)"). If you have a Mac, you also have the option of [downloading the dmg](#) to install a KBase app that launches a shell that is pre-configured with the KBase environment. If you have Ubuntu Linux, you will soon be able to download the KBase Ubuntu image.

Getting Started

[Installing the KBase Client via CPAN](#)

If you prefer to install the KBase client directly on your computer, rather than using IRIS or the Mac or Ubuntu images as described at the top of this page, you can follow the instructions in this tutorial.

[Getting Data from the Central Store](#)

KBase includes an integrated database of genomic data called the Central Store. Data can be extracted from the Central Store via the defined API or with a set of command-line scripts. This tutorial describes some ways to access data from the Central Store.

[Annotating a Genome Using KBase Tools](#)

This tutorial walks you through the steps involved in annotating a set of closely related genomes using KBase tools, including the construction of metabolic models to clarify differences between strains.

[KBase Development in the SEED Environment](#)

The KBase development environment is still under construction. To accelerate progress, we have created a framework within the SEED environment that allows us to develop and test software. We have employed this environment to support work on the ID server and the Central Store. We intend to move to a native KBase development environment, once it is well-defined. Until then, this tutorial is intended to help KBase developers get started with the SEED environment.

[IRIS: a Browser-based Framework for Interacting with KBase](#)

[Introduction to the KBase IRIS interface](#)



DOE Systems Biology Knowledgebase

Extensive Tutorials and Documentation



DOE Systems Biology Knowledgebase

Search this site



Home

About ▾

News ▾

Developer Zone ▾

KBase Labs

Contact Us

API and Command-Line Documentation

[CDM Command Line Scripts](#)

[CDM API](#)

[CDMI Entity-Relationship
Command Line Scripts](#)

[CDMI Entity-Relationship API](#)

[ID Server Command Line Scripts](#)

[ID Server API](#)

[Workspace Service](#)

[Plant Expression Service](#)

[Tree Service](#)

[Ontology Service](#)

[Protein Info Service](#)

[Similarity Service](#)

[Communities API](#)

[QC Service](#)

[Authorization Client](#)

[Genome Annotation Service](#)

[Translation Service](#)

[Networks Service](#)

Reference Documentation

Each of the [services](#) provided by the KBase project includes developer reference documentation. The list of currently available reference documents is shown on the left side of this page.

[CDM Command Line Scripts](#)

The CDM command line scripts allow users to access data in the Central Store. These scripts generally implement well trodden paths. These are commands that are useful under many circumstances and provide the basic functionality required for common analyses.

[CDM API](#)

The Central Data Model API allows users to query instances of the CDM. These include the Central Store which is the primary repository of genomic and modeling data, and Tiny Stores, which support per-user storage of genomic and modeling data.

[CDMI Entity-Relationship Command Line Scripts](#)

The CDMI Entity-Relationship command line scripts enable direct access to entities and relationships in the Central Store. Accessing relationships allows an user to easily navigate between entities in the central store, providing the ability to join data represented in more than one entity.

[CDMI Entity-Relationship API](#)

The CDMI Entity-Relationship API provides access to all entities and relationships in the CDM entity-relationship model. It is the API on which the CDM Entity-Relationship command scripts are built.

[ID Server Command Line Scripts](#)

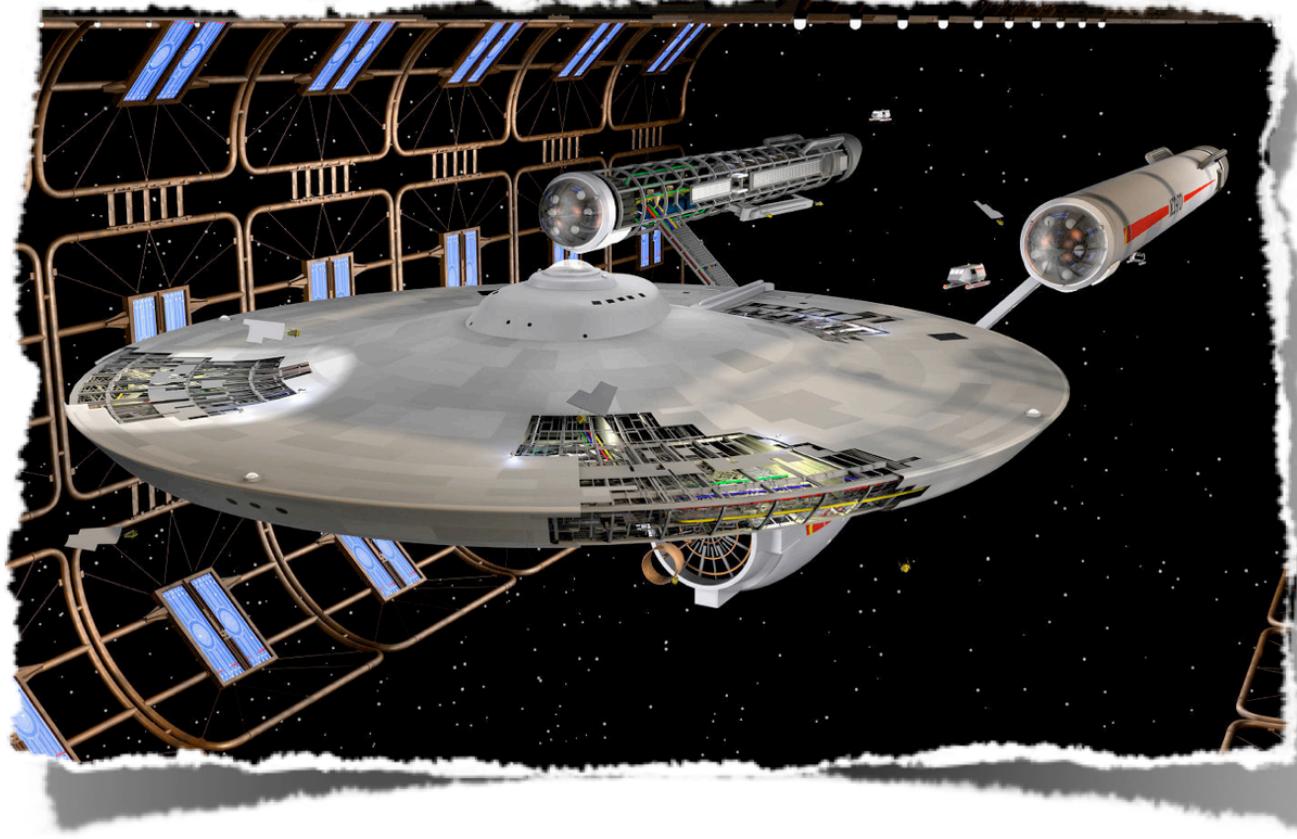
The ID Server command line scripts allow ID allocation and lookup services. This is the core service for generating

- Coordination and shared milestones with the BRCs
- Infrastructure collaborations with JGI, EMSL, NERC.
- Scientific collaborations with ENIGMA, PMI, and other SFAs
- MOAs with NSF for iPLANT and pending for USDA
- Preparation for collaboration and support for future DOE calls
- Needs survey, bootcamps, and scientific discussions with >70 community scientists.



KBase
PREDICTIVE BIOLOGY

DOE Systems Biology Knowledgebase



What next?

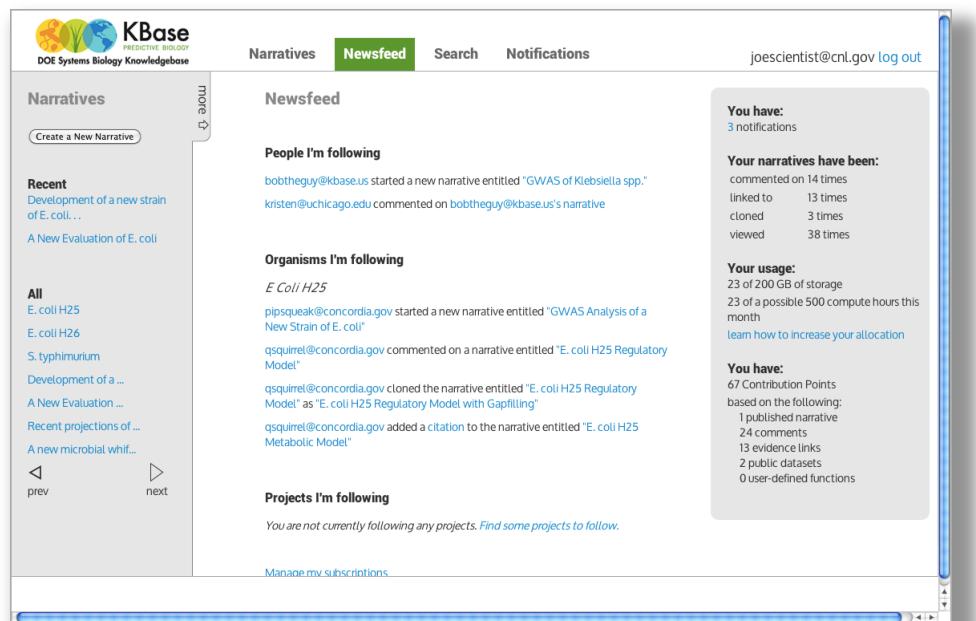
Building a starship in deep space?

- The data model must evolve to support the modeling mission
- A framework for turning bioinformatics algorithms' output into models needs further development
- **Theory for integrative, cross-scale predictive biology under development**
- Much better ontologies for nearly everything
- Building the social tools.
- Need to improve data import, quality assessment and metadata
- Efficient incorporation of new third party algorithms and support with scalable computing.
- Growing a strong external development community while maintaining quality, stability and vision.
- Launching the KBase Foundation to ease licensing and growth of KBase user participation.



A more concrete view of the “Narrative” interface

- **Users can share knowledge at multiple levels of granularity**
 - Cut and paste parts of narratives to reuse workflows
 - Cross-citation and branching of narratives
- **Project management**
 - Creation of teams allows management of projects
 - Electronic lab notebook for computational researchers
 - Projects can track progress
- **Possible publication model**
 - Finalized Narratives are reviewed, assigned DOI numbers, and accepted in an appropriate journal
 - Citation metrics
- **Metrics of Research Efficiency**
 - Times to completion of narratives
 - Times from hypothesis to confirming data/narratives
- **User, data, and algorithm ratings** by how many times the products appear in or are cited by narratives.
 - Can be aggregated by user, team, location, agency
 - People networks can be inferred by looking at team, citation, and comment structures.



The screenshot shows the KBase Newsfeed interface. At the top, there's a navigation bar with tabs for 'Narratives', 'Newsfeed' (which is active and highlighted in green), 'Search', and 'Notifications'. On the far right of the top bar, it shows the user's email (joescientist@cnl.gov) and a 'log out' link. Below the navigation, there's a sidebar on the left labeled 'Recent' which lists 'Development of a new strain of E. coli...' and 'A New Evaluation of E. coli'. The main content area is titled 'Newsfeed' and shows a list of recent activities. One activity is from 'bobtheguy@kbase.us' starting a new narrative titled "GWAS of Klebsiella spp.". Another is from 'kristen@uchicago.edu' commenting on 'bobtheguy@kbase.us's narrative. To the right of the newsfeed, there's a sidebar with various metrics: 'You have:' (3 notifications), 'Your narratives have been:' (commented on 14 times, linked to 13 times, cloned 3 times, viewed 38 times), 'Your usage:' (23 of 200 GB of storage, 23 of a possible 500 compute hours this month, with a link to 'learn how to increase your allocation'), and 'You have:' (67 Contribution Points based on publishing, commenting, linking, etc.).

Narratives

[Create a New Narrative](#)

more →

Recent

Development of a new strain of E. coli...

A New Evaluation of E. coli

All

E. coli H25

E. coli H26

S. typhimurium

Development of a ...

A New Evaluation ...

Recent projections of ...

A new microbial whif...



prev



next

Newsfeed

People I'm following

bobtheguy@kbase.us started a new narrative entitled "GWAS of Klebsiella spp."

kristen@uchicago.edu commented on bobtheguy@kbase.us's narrative

Organisms I'm following

E Coli H25

pipsqueak@concordia.gov started a new narrative entitled "GWAS Analysis of a New Strain of E. coli"

qsquirrel@concordia.gov commented on a narrative entitled "E. coli H25 Regulatory Model"

qsquirrel@concordia.gov cloned the narrative entitled "E. coli H25 Regulatory Model" as "E. coli H25 Regulatory Model with Gapfilling"

qsquirrel@concordia.gov added a citation to the narrative entitled "E. coli H25 Metabolic Model"

Projects I'm following

You are not currently following any projects. [Find some projects to follow.](#)

[Manage my subscriptions](#)

You have:

3 notifications

Your narratives have been:

commented on 14 times

linked to 13 times

cloned 3 times

viewed 38 times

Your usage:

23 of 200 GB of storage

23 of a possible 500 compute hours this month

[learn how to increase your allocation](#)

You have:

67 Contribution Points

based on the following:

1 published narrative

24 comments

13 evidence links

2 public datasets

0 user-defined functions

[Narratives](#)[Newsfeed](#)[Notifications](#)[Search](#)janescientist@mu.edu [log out](#)

Data	+ add data
In narrative	
In workspace	
Escherichia coli TW10828	

Escherichia coli E482

Escherichia coli 55989

Functions

- annotation
- browse genome
- metabolic reconstruction
- build FBA model
- FBA
- gapfilling
- geno2pheno
- variation

Untitled

Jane Scientist, Midwestern University

[Sharing](#) [Add Comment](#) [Add Evidence](#) [Copy URL](#)

Untitled

[Narratives](#)[Newsfeed](#)[Notifications](#)[Search](#)janescientist@mu.edu [log out](#)

Data	+ add data
In narrative	In workspace

Escherichia coli TW10828

Escherichia coli E482

Escherichia coli 55989

Functions

- annotation
- browse genome
- metabolic reconstruction
- build FBA model
- FBA
- gapfilling
- geno2pheno
- variation

Untitled

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

[Narratives](#)[Newsfeed](#)[Notifications](#)[Search](#)janescientist@mu.edu [log out](#)

Data	+ add data
In narrative	In workspace
Escherichia coli TW10828	
Escherichia coli E482	

Escherichia coli 55989

Functions

annotation
browse genome
metabolic reconstruction
build FBA model
FBA
gapfilling
geno2pheno
variation

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#) [Add Comment](#) [Add Evidence](#) [Copy URL](#)[Refining the Metabolic Model for Escherichia coli F11](#)

[Narratives](#)[Newsfeed](#)[Notifications](#)[Search](#)janescientist@mu.edu [log out](#)

Data	+ add data
In narrative	
In workspace	

Escherichia coli TW10828

Escherichia coli E482

Escherichia coli 55989

Functions

- annotation
- browse genome
- metabolic reconstruction
- build FBA model
- FBA
- gapfilling
- geno2pheno
- variation

[Refining the Metabolic Model for Escherichia coli F11](#)

Refining the Metabolic Model for *Escherichia coli* F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

Click to add text

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

[Add Comment](#) [Add Evidence](#) [Copy URL](#)**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)[Refining the Metabolic Model for Escherichia coli F11](#)

[Narratives](#)[Newsfeed](#)[Notifications](#)[Search](#)janescientist@mu.edu [log out](#)

Data	+ add data
In narrative	
In workspace	
Escherichia coli TW10828	

Escherichia coli E482

Escherichia coli 55989

Functions

- annotation
- browse genome
- metabolic reconstruction
- build FBA model
- FBA
- gapfilling
- geno2pheno
- variation

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

Refining the Metabolic Model for Escherichia coli F11

Data [+ add data](#)[In narrative](#) [In workspace](#)

Escherichia coli TW10828

Escherichia coli E482

Escherichia coli 55989

Functions

annotation

browse genome

metabolic reconstruction

build FBA model

FBA

gapfilling

geno2pheno

variation

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

Add a Dataset to your Narrative

- [Upload a file](#)

-
- [Search the Central Data Store](#)

[Refining the Metabolic Model for Escherichia coli F11](#)

Search KBase

Searched for **Escherichia coli** found 11385 results.

Now showing 10 results.

[add selected datasets to: ▾](#)

Literature

[Pubmed \(6333\)](#)

Genomes

[Bacteria \(272\)](#)

[Viruses \(0\)](#)

[Eukaryota \(0\)](#)

[Archaea \(0\)](#)

Function

[Genes \(4780\)](#)

[Locus \(0\)](#)

[Prophage \(0\)](#)

[Pseudo Gene \(0\)](#)

Regulation

1. **Escherichia coli TW10828**

Bacteria Contigs: 227 RNA Features: 122 Protein Encoding Genes: 5139

— Genome Size: 5280267 bp GC Content: 50.64 %

[Details](#)

2. **Escherichia coli 55989**

Bacteria Contigs: 2 RNA Features: 116 Protein Encoding Genes: 5175

— Genome Size: 5227344 bp GC Content: 50.59 %

[Details](#)

3. **Escherichia coli E482**

Bacteria Contigs: 94 RNA Features: 94 Protein Encoding Genes: 4752

— Genome Size: 4833169 bp GC Content: 50.70 %

[Details](#)

Search KBase

Searched for **Escherichia coli** found 11385 results.

Now showing 10 results.

▾

Pseudo Gene (0)	— Genome Size: 4855169 bp GC Content: 50.70 %	Details
Regulation		
Promoter (0)	—	Details
Operator (0)	—	Details
Protein Binding Sites (0)	—	Details
Binding Sites (0)	—	Details
Riboswitch (0)	—	Details
RNA		
RNAs (0)	—	Details
Transposons (0)	—	Details
mRNAs (0)	—	Details
Small RNAs (0)	—	Details
4. Escherichia coli F11	Bacteria Contigs: 119 RNA Features: 96 Protein Encoding Genes: 5102	Details
— Genome Size: 5215961 bp GC Content: 50.48 %		
5. Escherichia coli 9.1649	Bacteria Contigs: 33 RNA Features: 112 Protein Encoding Genes: 5037	Details
— Genome Size: 5102049 bp GC Content: 50.64 %		
6. Escherichia coli W	Bacteria Contigs: 1 RNA Features: 0 Protein Encoding Genes: 6	Details
— Genome Size: 5380 kb GC Content: 48.18 %		

Search KBase

Searched for **Escherichia coli** found 11385 results.

Now showing 10 results.

Pseudo Gene (0)	— Genome Size: 4055109 bp GC Content: 50.70 %	Details
Regulation		
Promoter (0)	<input checked="" type="checkbox"/>	4. Escherichia coli F11
Operator (0)		Bacteria Contigs: 119 RNA Features: 96 Protein Encoding Genes: 5102
Protein Binding Sites (0)		— Genome Size: 5215961 bp GC Content: 50.48 %
Binding Sites (0)		Details
Riboswitch (0)		
RNA		
RNAs (0)	<input type="checkbox"/>	5. Escherichia coli 9.1649
Transposons (0)		Bacteria Contigs: 33 RNA Features: 112 Protein Encoding Genes: 5037
mRNAs (0)		— Genome Size: 5102049 bp GC Content: 50.64 %
Small RNAs (0)		Details
6. Escherichia coli W		
		Bacteria Contigs: 1 RNA Features: 0 Protein Encoding Genes: 6
		— Genome Size: 5380 bp GC Content: 48.18 %

Search KBase

Searched for **Escherichia coli** found 11385 results.

Now showing 10 results.

add selected datasets to: ▾

- a new narrative
- the current narrative
- the current workspace

Promoter (0)

Operator (0)

Protein Binding Sites (0)

Binding Sites (0)

Riboswitch (0)

RNA

RNAs (0)

Transposons (0)

mRNAs (0)

Small RNAs (0)

— Genome Size: 4855169 bp GC Content: 50.70 %

Details

4. **Escherichia coli F11**

Bacteria

Contigs: 119 RNA Features: 96 Protein Encoding Genes: 5102

— Genome Size: 5215961 bp GC Content: 50.48 %

Details

5. **Escherichia coli 9.1649**

Bacteria

Contigs: 33 RNA Features: 112 Protein Encoding Genes: 5037

— Genome Size: 5102049 bp GC Content: 50.64 %

Details

6. **Escherichia coli W**

Bacteria

Contigs: 1 RNA Features: 0 Protein Encoding Genes: 6

— Genome Size: 5380 bp GC Content: 48.18 %

Search KBase

Searched for **Escherichia coli** found 11385 results.

Now showing 10 results.

add selected datasets to: ▾

- a new narrative
- the current narrative** 
- the current workspace

Promoter (0)

Operator (0)

Protein Binding Sites (0)

Binding Sites (0)

Riboswitch (0)

RNA

RNAs (0)

Transposons (0)

mRNAs (0)

Small RNAs (0)

— Genome Size: 4055109 bp GC Content: 50.70 %

[Details](#)

4. **Escherichia coli F11**

Bacteria

Contigs: 119 RNA Features: 96

Protein Encoding Genes: 5102

— Genome Size: 5215961 bp GC Content: 50.48 %

[Details](#)

5. **Escherichia coli 9.1649**

Bacteria

Contigs: 33 RNA Features: 112

Protein Encoding Genes: 5037

— Genome Size: 5102049 bp GC Content: 50.64 %

[Details](#)

6. **Escherichia coli W**

Bacteria

Contigs: 1 RNA Features: 0

Protein Encoding Genes: 6

— Genome Size: 5380 bp GC Content: 48.18 %

[Narratives](#)[Newsfeed](#)[Notifications](#)[Search](#)janescientist@mu.edu [log out](#)

Data	+ add data
In narrative	In workspace

Escherichia coli TW10828

Escherichia coli E482

Escherichia coli 55989

Escherichia coli F11 Genome, annotated

Functions

annotation

browse genome

metabolic reconstruction

build FBA model

FBA

gapfilling

geno2pheno

variation

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

Escherichia coli F11 Genome, annotated

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB|g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

[Refining the Metabolic Model for Escherichia coli F11](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)[Refining the Metabolic Model for Escherichia coli F11](#)

Refining the Metabolic Model for *Escherichia coli* F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of *E. coli* F11 from the KBase Central Data Store.

[Escherichia coli F11 Genome, annotated](#)

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB|g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

[Click to add text](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)[Refining the Metabolic Model for Escherichia coli F11](#)

Refining the Metabolic Model for *Escherichia coli* F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of *E. coli* F11 from the KBase Central Data Store.

[Escherichia coli F11 Genome, annotated](#)

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB|g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

Let's take a look at this genome.

Data [+ add data](#)

In narrative	In workspace
--------------	--------------

Escherichia coli TW10828

Escherichia coli E482

Escherichia coli 55989

Escherichia coli F11 Genome, annotated

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)**Escherichia coli F11 Genome, annotated**

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB|g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

Complete genome

GC Content: 50.64%

Let's take a look at this genome.

Genome Information Browser

ID: kb|g.961

Scientific Name: Escherichia coli

F11

Domain: Bacteria

Complete? Yes

Genome Size: 5,215,961 bp

GC Content: 50.48%

Number of Contigs: 119

Number of Features: 5198

Number of Protein Encoding

Genes: 5102

Number of RNA Features: 96

Source: SEED: 340197.5

Contig: kb|g.961.c.0 - 465383 bp



Feature ID

kb|g.961.peg.3915

Function

Length

2007 bp

[View in Separate Window](#)

Refining the Metabolic Model for Escherichia coli F11

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

Escherichia coli F11 Genome, annotated

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB|g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

Let's take a look at this genome.

▼ Genome Information Browser**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)[Refining the Metabolic Model for Escherichia coli F11](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

**Escherichia coli F11 Genome,
annotated**

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB|g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

Let's take a look at this genome.

▼ Genome Information Browser**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)[Refining the Metabolic Model for Escherichia coli F11](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

Escherichia coli F11 Genome, annotated

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB|g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

Functions[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)

Let's take a look at this genome.

▼ Genome Information Browser

Now I'm going to run a metabolic reconstruction on this genome. Let's see how

[Add Comment](#)[Add Evidence](#)[Copy URL](#)

Refining the Metabolic Model for Escherichia coli F11

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

Escherichia coli F11 Genome, annotated

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB|g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

Let's take a look at this genome.

▼ Genome Information Browser

Now I'm going to run a metabolic reconstruction on this genome. Let's see how the initial one comes out.

This is the first time I've done this for this particular strain, so it will be interesting to see which roles are found.

Functions[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)[Refining the Metabolic Model for Escherichia coli F11](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

Run metabolic reconstruction**Dataset:**[E. coli F11 Genome](#)[add a dataset](#)**Other Options:**[run](#) [cancel](#)

gs: 2
pairs: 5,280,267
232
content: 50.64%

the initial one comes out.
to see which roles are found.

Functions[annotation](#)
[browse genome](#)
[metabolic reconstruction](#)
[build FBA model](#)
[FBA](#)
[gapfilling](#)
[geno2pheno](#)
[variation](#)[Refining the Metabolic Model for Escherichia coli F11](#)

Data	+ add data
In narrative	
In workspace	
Escherichia coli TW10828	
Escherichia coli E482	
Escherichia coli 55989	
Escherichia coli F11 Genome, annotated	
Escherichia coli F11 Metabolic Reconstruction	

Functions

annotation
 browse genome
 metabolic reconstruction
 build FBA model
 FBA
 gapfilling
 geno2pheno
 variation

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

I'll start by getting the genome of *E. coli* F11 from the KBase Central Data Store.

Escherichia coli F11 Genome, annotated

KBase ID: KB1g.691

Source: Joe Scientist

Contigs: 2

Version: Mar 27, 2013 08:13:00 PST

Base pairs: 5,280,267

Domain: Bacteria

PEGs: 232

Taxon: gamma-proteobacteria

Complete genome

GC Content: 50.64%

Let's take a look at this genome.

Genome Information Browser

Now I'm going to run a metabolic reconstruction on this genome. Let's see how the initial one comes out.
 This is the first time I've done this for this particular strain, so it will be interesting to see which roles are found.

Run Metabolic Reconstruction on *E. coli* F11 Genome

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm

**Escherichia coli F11 Metabolic Reconstruction**

Source: calculated

Roles: 29,294

Version: Mar 27, 2013 08:13:00 PST

Refining the Metabolic Model for Escherichia coli F11

Data	+ add data
In narrative	In workspace
Escherichia coli TW10828	
Escherichia coli E482	
Escherichia coli 55989	
Escherichia coli F11 Genome, annotated	
Escherichia coli F11 Metabolic Reconstruction	

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

Escherichia coli F11 Genome, annotated

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB1g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

Let's take a look at this genome.

▼ Genome Information Browser

Now I'm going to run a metabolic reconstruction on this genome. Let's see how the initial one comes out. This is the first time I've done this for this particular strain, so it will be interesting to see which roles are found.

Run Metabolic Reconstruction on E. coli F11 Genome

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm



Escherichia coli F11 Metabolic Reconstruction

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

Roles: 29,294

[1 Comment](#)

[Add Comment](#) [Add Evidence](#) [Copy URL](#)

Refining the Metabolic Model for Escherichia coli F11

Data	+ add data
In narrative	In workspace
Escherichia coli TW10828	
Escherichia coli E482	
Escherichia coli 55989	
Escherichia coli F11 Genome, annotated	
Escherichia coli F11 Metabolic Reconstruction	

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)
[Add Comment](#)
[Add Evidence](#)
[Copy URL](#)

Escherichia coli F11 Genome, annotated

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB|g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

Let's take a look at this genome.

▼ Genome Information Browser

Now I'm going to run a metabolic reconstruction on this genome. Let's see how the initial one comes out. This is the first time I've done this for this particular strain, so it will be interesting to see which roles are found.

Run Metabolic Reconstruction on E. coli F11 Genome

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm


[X close](#)

Escherichia coli F11 Metabolic Reconstruction

 So
Ve

 Comment by [bobtheguy@kbase.us](#):

I've grown this without serine, but the genes for serine are missing in this automated reconstruction. I think you'll need to add [kb|g.362.peg.287](#), [kb|g.382.peg.123](#), and [kb|g.382.peg.898](#).

[Add Comment](#)
[Add Evidence](#)
[Copy URL](#)
[Refining the Metabolic Model for Escherichia coli F11](#)

Data	+ add data
In narrative	In workspace
Escherichia coli TW10828	
Escherichia coli E482	
Escherichia coli 55989	
Escherichia coli F11 Genome, annotated	
Escherichia coli F11 Metabolic Reconstruction	

Functions
annotation
browse genome
metabolic reconstruction
build FBA model
FBA
gapfilling
geno2pheno
variation

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

Escherichia coli F11 Genome, annotated

Domain: Bacteria

Taxon: gamma-proteobacteria

KBase ID: KB1g.691

Source: Joe Scientist

Version: Mar 27, 2013 08:13:00 PST

Complete genome

Contigs: 2

Base pairs: 5,280,267

PEGs: 232

GC Content: 50.64%

Let's take a look at this genome.

▼ Genome Information Browser

Now I'm going to run a metabolic reconstruction on this genome. Let's see how the initial one comes out. This is the first time I've done this for this particular strain, so it will be interesting to see which roles are found.

Run Metabolic Reconstruction on E. coli F11 Genome

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm



Escherichia coli F11 Metabolic Reconstruction

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

Roles: 29,294

[1 Comment](#)

[Add Comment](#) [Add Evidence](#) [Copy URL](#)

Refining the Metabolic Model for Escherichia coli F11

Data	+ add data
In narrative	In workspace
Escherichia coli TW10828	
Escherichia coli E482	
Escherichia coli 55989	
Escherichia coli F11 Genome, annotated	
Escherichia coli F11 Metabolic Reconstruction	

Functions
annotation
browse genome
metabolic reconstruction
build FBA model
FBA
gapfilling
geno2pheno
variation

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)
[Add Comment](#)
[Add Evidence](#)
[Copy URL](#)

I'll start by getting the genome of E. coli F11 from the KBase Central Data Store.

Run build FBA model

Dataset: [add a dataset](#)

Other Options:

Escherichia coli F11 Metabolic Reconstruction

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

Roles: 29,294

[1 Comment](#)
[Refining the Metabolic Model for Escherichia coli F11](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)[Escherichia coli F11 Metabolic Reconstruction](#)[Escherichia coli F11 FBA Model](#)**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

Run Metabolic Reconstruction on E. coli F11 Genome

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm



[Escherichia coli F11 Metabolic Reconstruction](#)

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

Roles: 29,294

[1 Comment](#)

Run Build FBA Model on E. coli F11 Genome

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm



[Escherichia coli F11 FBA Model](#)

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

Reactions: 29,294

Compounds: 2098

Annotations: 234

[Refining the Metabolic Model for Escherichia coli F11](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)[Escherichia coli F11 Metabolic Reconstruction](#)[Escherichia coli F11 FBA Model](#)**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

Run Metabolic Reconstruction on E. coli F11 Genome

[Run FBA](#)**Dataset:**[E. coli F11 Genome FBA Model](#) [+ add a dataset](#)**Other Options:**[run](#) [cancel](#)[Run](#)[Star](#)[Finish](#)**Escherichia coli F11 FBA Model**

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

Reactions: 29,294

Compounds: 2098

Annotations: 234

[Refining the Metabolic Model for Escherichia coli F11](#)

Data	+ add data
In narrative	In workspace
Escherichia coli TW10828	
Escherichia coli E482	
Escherichia coli 55989	
Escherichia coli F11 Genome, annotated	
Escherichia coli F11 Metabolic Reconstruction	
Escherichia coli F11 FBA Model	
Escherichia coli F11 Flux Balance Analysis	

Functions
annotation
browse genome
metabolic reconstruction
build FBA model
FBA
gapfilling
geno2pheno
variation

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)

Run Build FBA Model on E. coli F11 Genome

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm



Escherichia coli F11 FBA Model

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

Reactions: 29,294

Compounds: 2098

Annotations: 234

Run FBA on E. coli F11 Metabolic Model

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm



Escherichia coli F11 FBA Result

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

P(Growth): 0.98

Refining the Metabolic Model for Escherichia coli F11

Data [+ add data](#)

In narrative In workspace

Escherichia coli TW10828
Escherichia coli E482
Escherichia coli 55989
Escherichia coli F11 Genome, annotated
Escherichia coli F11 Metabolic Reconstruction
Escherichia coli F11 FBA Model
Escherichia coli F11 Flux Balance Analysis

Functions

annotation
browse genome
metabolic reconstruction
build FBA model
FBA
gapfilling
geno2pheno
variation

Refining the Metabolic Model for Escherichia coli F11

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#) [Add Comment](#) [Add Evidence](#) [Copy URL](#)

Run Build FBA Model on E. coli F11 Genome

Star
Finish

Add a Dataset to your Narrative

- [Upload a file](#)

- [Search the Central Data Store](#)

[Search](#)

Escherichia coli F11 FBA Result

Source: calculated
Version: Mar 27, 2013 08:13:00 PST

P(Growth): 0.98

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)[Escherichia coli F11 Metabolic Reconstruction](#)[Escherichia coli F11 FBA Model](#)[Escherichia coli F11 Flux Balance Analysis](#)**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)**Upload a File to your Narrative****File:**  ecoli_F11_growth021213.txt**Dataset Type:****Dataset Name:**

Reactions: 29,294

Compounds: 2098

Enzymes: 234

Escherichia coli F11 FBA Result

Source: calculated

P(Growth): 0.98

Version: Mar 27, 2013 08:13:00 PST

[Refining the Metabolic Model for Escherichia coli F11](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)[Escherichia coli F11 Metabolic Reconstruction](#)[Escherichia coli F11 FBA Model](#)[Escherichia coli F11 Flux Balance Analysis](#)**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)**Upload a File to your Narrative****File:**  ecoli_F11_growth021213.txt**Dataset Type:**

- select...
 - Abundance Profile
 - Genome sequence, FASTA
 - Genome sequence, Key-Value
 - Gene Sequence, FASTA
 - Gene Sequence, Key-Value
 - Protein sequence, FASTA
 - Protein Sequence, Key-Value
 - Sequence Alignments, .aln
 - Sequence Alignments, .msf
 - Sequence Mapping
 - Expression Profile
- Growth/No Growth
- Metabolic Model
- Flux Distribution
- Regulatory Model

Escherichia coli F11 FBA Result

Source: calculated

P(Growth): 0.98

Version: Mar 27, 2013 08:13:00 PST

[Refining the Metabolic Model for Escherichia coli F11](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)[Escherichia coli F11 Metabolic Reconstruction](#)[Escherichia coli F11 FBA Model](#)[Escherichia coli F11 Flux Balance Analysis](#)**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)**Upload a File to your Narrative****File:**  ecoli_F11_growth021213.txt**Dataset Type:****Dataset Name:**

Escherichia coli F11 Growth

actions: 29,294

compounds: 2098

rotations: 234

Escherichia coli F11 FBA Result

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

P(Growth): 0.98

[Refining the Metabolic Model for Escherichia coli F11](#)

Data [+ add data](#)[In narrative](#) [In workspace](#)[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)[Escherichia coli F11 Metabolic Reconstruction](#)[Escherichia coli F11 FBA Model](#)[Escherichia coli F11 Flux Balance Analysis](#)[Escherichia coli F11 Growth](#)**Functions**[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[geno2pheno](#)[variation](#)

Refining the Metabolic Model for Escherichia coli F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)**Escherichia coli F11 FBA Model**

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

Reactions: 29,294

Compounds: 2098

Annotations: 234

Run FBA on E. coli F11 Metabolic Model

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm

**Escherichia coli F11 FBA Result**

Source: calculated

Version: Mar 27, 2013 08:13:00 PST

P(Growth): 0.98

[Refining the Metabolic Model for Escherichia coli F11](#)**Escherichia coli F11 Growth**

Source: user uploaded

Version: Mar 27, 2013 08:13:00 PST

growth x nickel abundance

Data [+ add data](#)
[In narrative](#) [In workspace](#)
[Escherichia coli TW10828](#)[Escherichia coli E482](#)[Escherichia coli 55989](#)[Escherichia coli F11 Genome, annotated](#)[Escherichia coli F11 Metabolic Reconstruction](#)[Escherichia coli F11 FBA Model](#)[Escherichia coli F11 Flux Balance Analysis](#)[Escherichia coli F11 Growth](#)
Functions
[annotation](#)[browse genome](#)[metabolic reconstruction](#)[build FBA model](#)[FBA](#)[gapfilling](#)[visualize in GLAMM](#)[geno2pheno](#)[variation](#)
[Refining the Metabolic Model for Escherichia coli F11](#)

Refining the Metabolic Model for *Escherichia coli* F11

Jane Scientist, Midwestern University

[Sharing](#)[Add Comment](#)[Add Evidence](#)[Copy URL](#)**Model**

Version: Mar 27, 2013 08:13:00 PST

Compounds: 2098

Annotations: 234

Run gapfilling**Model Dataset:**
[E. coli F11 Genome FBA Model](#) [+ add a dataset](#)
Phenotype Dataset:
[E. coli F11 Growth Data](#) [+ add a dataset](#)
Other Options:
[run](#) [cancel](#)
***Escherichia coli* F11 Growth**

Source: user uploaded

Version: Mar 27, 2013 08:13:00 PST

growth x nickel abundance

Data [+ add data](#)

In narrative In workspace

[Escherichia coli TW10828](#)

[Escherichia coli E482](#)

[Escherichia coli 55989](#)

[Escherichia coli F11 Genome, annotated](#)

[Escherichia coli F11 Metabolic Reconstruction](#)

[Escherichia coli F11 FBA Model](#)

[Escherichia coli F11 Flux Balance Analysis](#)

[Escherichia coli F11 Gapfilled FBA Model](#)

Functions

[annotation](#)

[browse genome](#)

[metabolic reconstruction](#)

[build FBA model](#)

[FBA](#)

[gapfilling](#)

[visualize in GLAMM](#)

[geno2pheno](#)

[variation](#)

[Refining the Metabolic Model for Escherichia coli F11](#)

Refining the Metabolic Model for *Escherichia coli* F11

Jane Scientist, Midwestern University

[Sharing](#)

[Add Comment](#)

[Add Evidence](#)

[Copy URL](#)

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm

Escherichia coli F11 FBA Result

Source: calculated

P(Growth): 0.98

Version: Mar 27, 2013 08:13:00 PST

Run Gapfilling on *E. coli* F11 Metabolic Model

Started: 9/27/2012 4:15:32pm

Finished: 9/27/2012 4:15:33pm

Escherichia coli F11 Gapfilled FBA Model

Source: calculated

Reactions Added: 4

Version: Mar 27, 2013 08:27:00 PST

growth on serine- media

Data [+ add data](#)
[In narrative](#) [In workspace](#)

- [Escherichia coli TW10828](#)
- [Escherichia coli E482](#)
- [Escherichia coli 55989](#)
- [Escherichia coli F11 Genome, annotated](#)
- [Escherichia coli F11 Metabolic Reconstruction](#)
- [Escherichia coli F11 FBA Model](#)
- [Escherichia coli F11 Flux Balance Analysis](#)
- [Escherichia coli F11 Growth](#)
- [Escherichia coli F11 Gapfilled FBA Model](#)

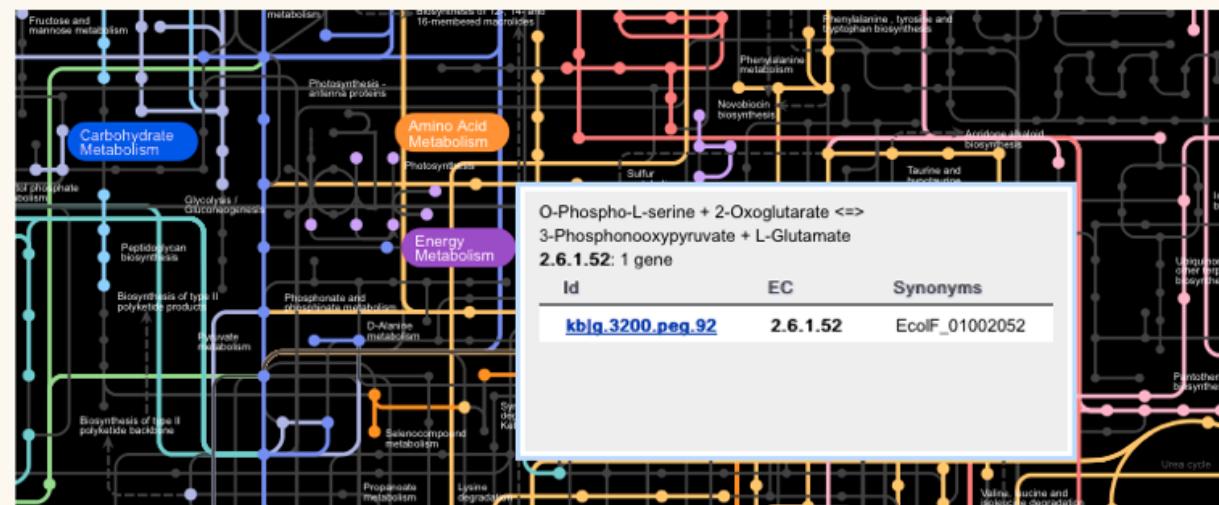
Functions

- [annotation](#)
- [browse genome](#)
- [metabolic reconstruction](#)
- [build FBA model](#)
- [FBA](#)
- [gapfilling](#)
- [visualize in GLAMM](#)
- [geno2pheno](#)
- [variation](#)

[Refining the Metabolic Model for Escherichia coli F11](#)

Refining the Metabolic Model for *Escherichia coli* F11

Jane Scientist, Midwestern University

[Sharing](#)
[Add Comment](#)
[Add Evidence](#)
[Copy URL](#)
VERSION: May 27, 2013 08:27:00 PST
grown on serine- media
▼ Visualize in GLAMM

[View in Separate Window](#)

Narratives

[Create a New Narrative](#)

Recent

Development of a new strain
of *E. coli*...

A New Evaluation of *E. coli*

All

E. coli H25

E. coli H26

S. typhimurium

Development of a ...

A New Evaluation ...

Recent projections of ...

A new microbial whif...



prev



next

Newsfeed

People I'm following

[bobtheguy@kbase.us](#) started a new narrative entitled "[GWAS of Klebsiella spp.](#)"

Organisms I'm following

E Coli F11

[janescientist@mu.edu](#) published a narrative entitled "[Refining the Metabolic Model for Escherichia coli F11](#)"

Projects I'm following

You are not currently following any projects. [Find some projects to follow.](#)

[Manage my subscriptions](#)

You have:

3 notifications

Your narratives have been:

commented on 14 times

linked to 13 times

cloned 3 times

viewed 38 times

Your usage:

23 of 200 GB of storage

23 of a possible 500 compute hours this month

[learn how to increase your allocation](#)

You have:

67 Contribution Points

based on the following:

1 published narrative

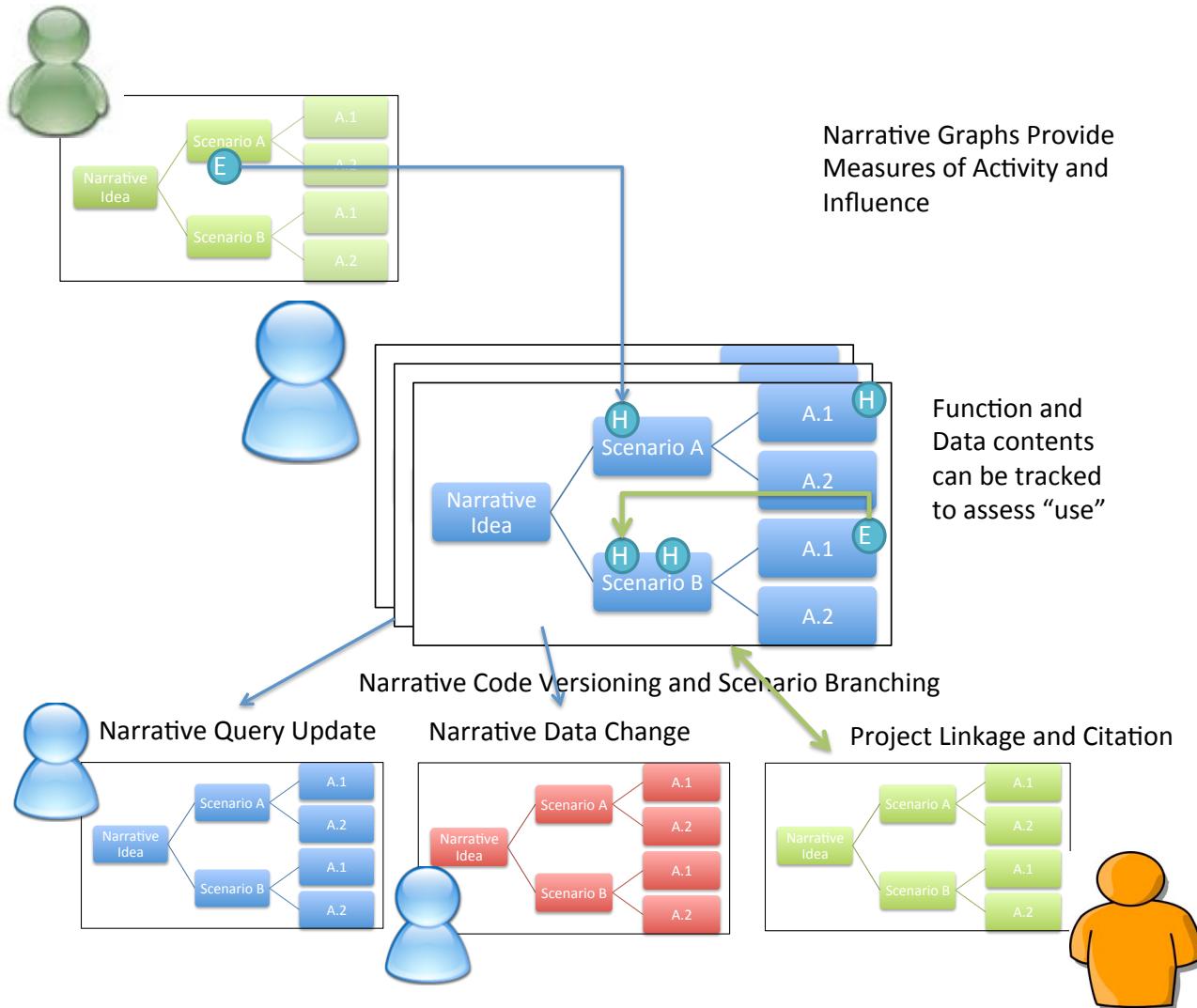
24 comments

13 evidence links

2 public datasets

0 user-defined functions

Concept: Interactive community knowledge



Into the future

- The data model must evolve to support the modeling mission
- A framework for turning bioinformatics algorithms' output into models needs further development
- **Theory for integrative, cross-scale predictive biology under development**
- Much better ontologies for nearly everything
- Building the social tools.
- Need to improve data import, quality assessment and metadata
- Efficient incorporation of new third party algorithms and support with scalable compute.
- Growing a strong external development community while maintaining quality, stability and vision.
- Launching the KBase Foundation to ease licensing and growth of KBase user participation.



A more concrete view of the “Narrative” interface

KBase Team

Berkeley

Adam Arkin	Matt Henderson
Dylan Chivian	Marcin Joachimiak
Paramvir Dehal	Keith Keller
Nomi Harris	Pavel Novichkov
Jason Baumohl	Sarah Poon
Ben Bowen	Gavin Price
Steven Brenner	Bill Riehl
Shane Canon	Michael Sneddon
Stephen Chan	Gwyneth Terry
John-Marc Chandonia	Cary Whitney
Shreyas Cholia	
Annette Greiner	

Yale

Mark Gerstein
Gang Fang
Lucas Lochovsky
Daifeng Wang

Communities	Microbes
Infrastructure	Plants
Management	Outreach



UIUC

Gary Olsen

UC Davis

Pamela Ronald
Taeyun Oh

Argonne
NATIONAL LABORATORY

Brookhaven

Sergei Maslov
Fei He
Shinjae Yoo
Dantong Yu

Cold Spring Harbor

Doreen Ware
James Gurtowski
Sunita Kumari
Shiran Pasternak
Michael Schatz
James Thomason

Hope College

Matt DeJongh
Aaron Best
Nathan Tintle

Univ. TN

Priya Ranjan

Argonne

Rick Stevens	Adina Howe
Tom Brettin	Kevin Keegan
Elizabeth Glass	Silvia Mulligan
Chris Henry	Bob Olson
Folker Meyer	Dan Olson
Jennifer Salazar	Ross Overbeek
Jared Bischof	Tobias Paczian
Neal Conrad	Bruce Parello
Narayan Desai	Sam Seaver
Scott Devoid	Will Trimble
Terry Disz	Andreas Wilke
Paul Frybarger	Jared Wilkening
India Gordon	Fangfang Xia
Travis Harrison	

Oak Ridge

Bob Cottingham	Michael Leuze
Brian Davison	Steve Moulton
Dave Weston	Tony Palumbo
Meghan Drake	Mustafa Syed
Guru Kora	
Miriam Land	

Leadership Team



Adam Arkin



Rick Stevens
Project Management



Bob Cottingham



Sergei Maslov

Microbes



Paramvir Dehal

Communities



Folker Meyer

Public Relations



Nomi Harris

Plants



Doreen Ware

Infrastructure



Tom Brettin

Outreach



Elizabeth Glass

Chris Henry



Dylan Chivian



Jennifer Salazar



Dave Weston

Shane Canon



Brian Davison



BERKELEY LAB



Thank you!

Visit us at kbase.science.energy.gov

Contact us at outreach@kbase.us

