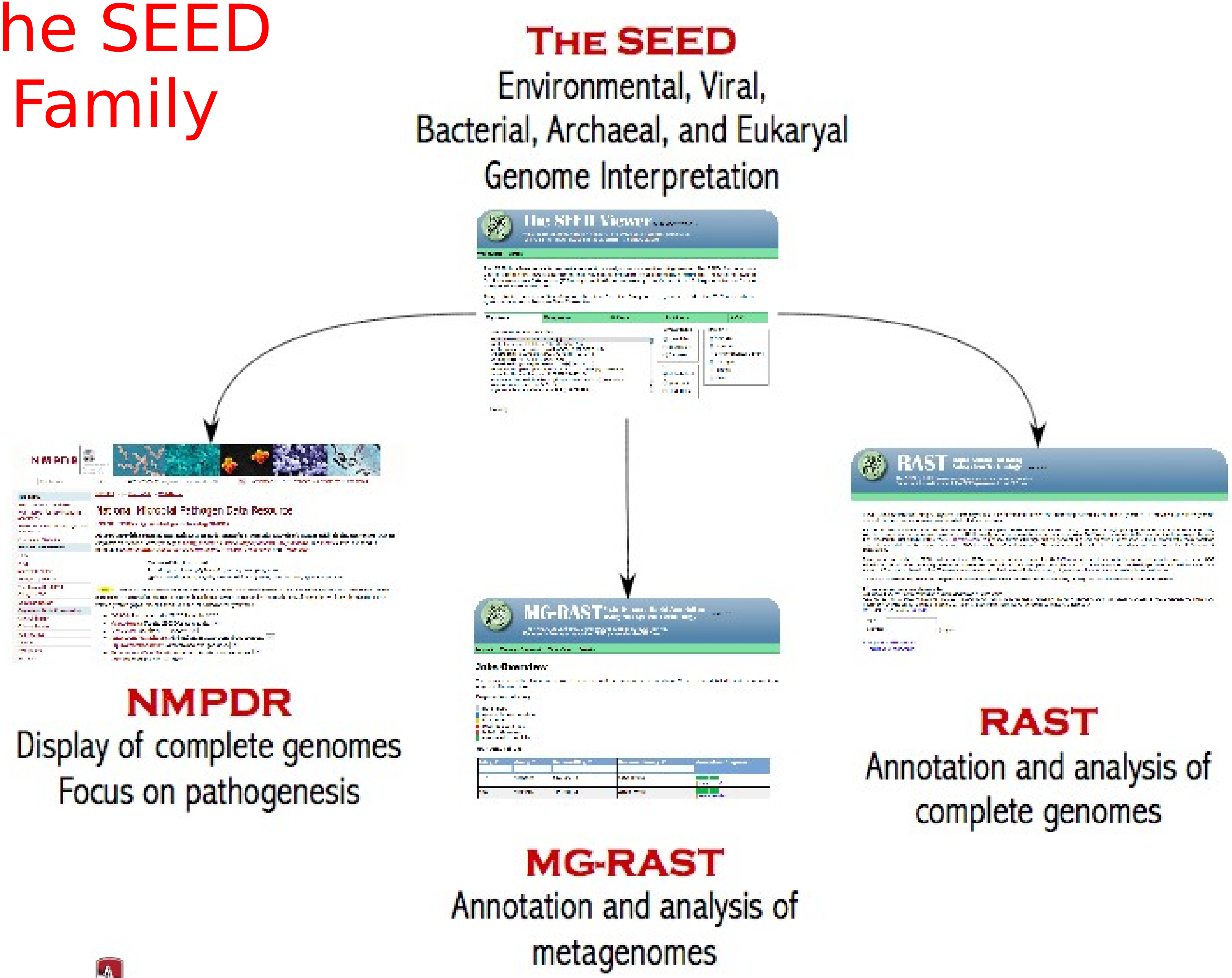
The SEED Family



THE SEED
Environmental, Viral, Bacterial, Archaeal, and Eukaryal Genome Interpretation

NMPDR
Display of complete genomes
Focus on pathogenesis

MG-RAST
Annotation and analysis of metagenomes

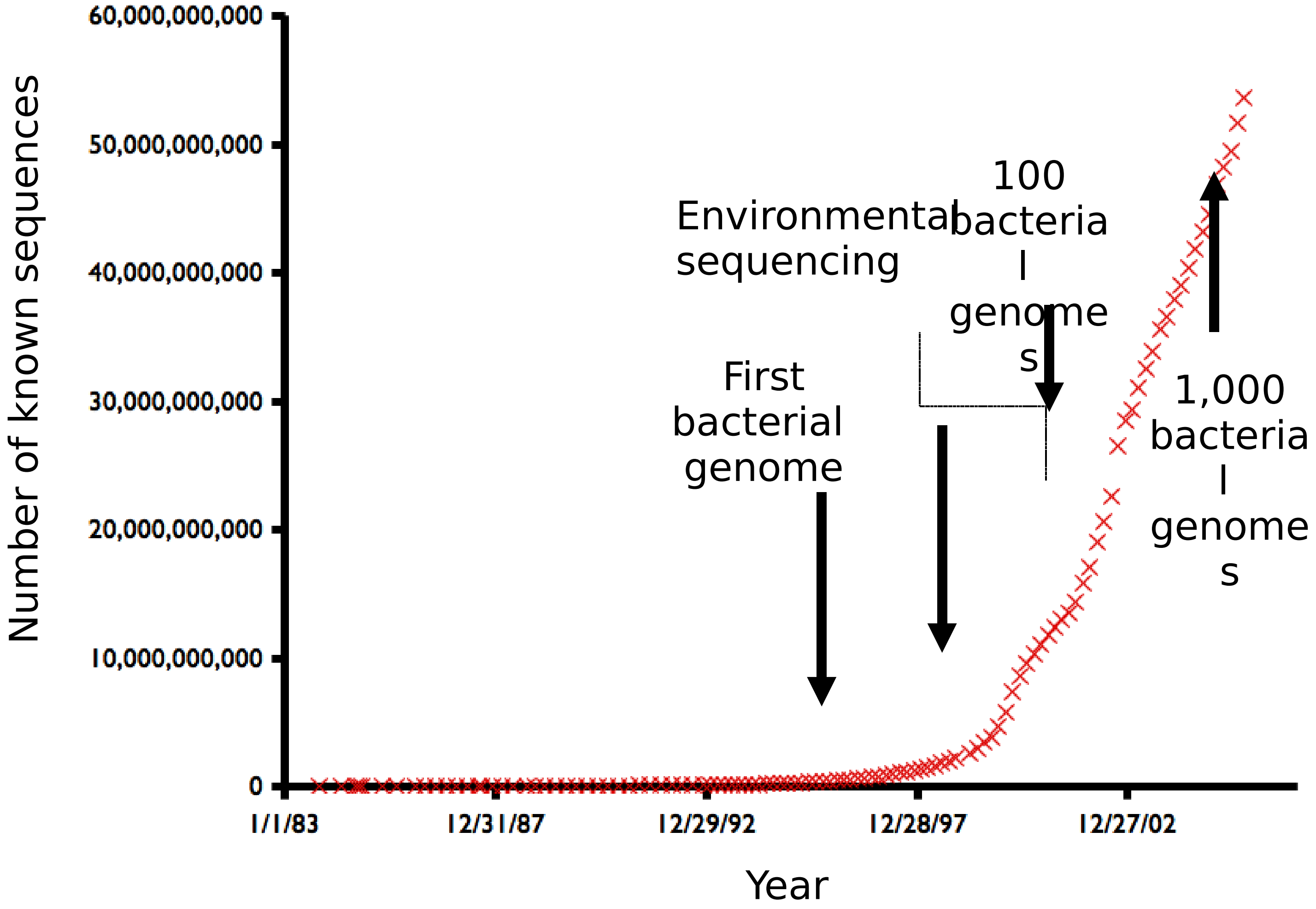
RAST
Annotation and analysis of complete genomes

SAN DIEGO STATE UNIVERSITY

www.nmpdr.org www.theseed.org

NMPDR National Microbial Pathogen Data Resource Center

How much has been sequenced?



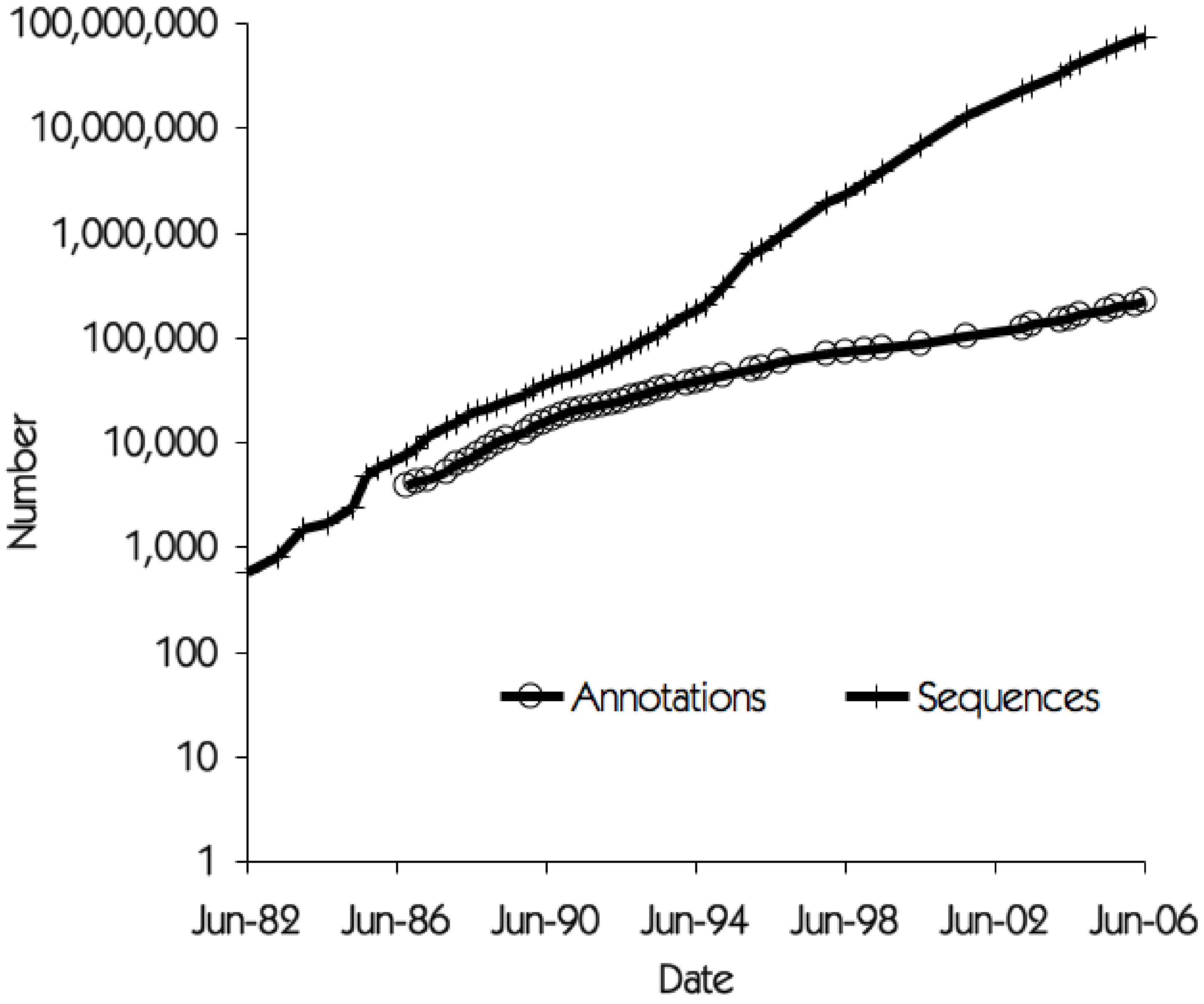
Environmental sequencing

100 bacterial genomes

First bacterial genome

1,000 bacterial genomes

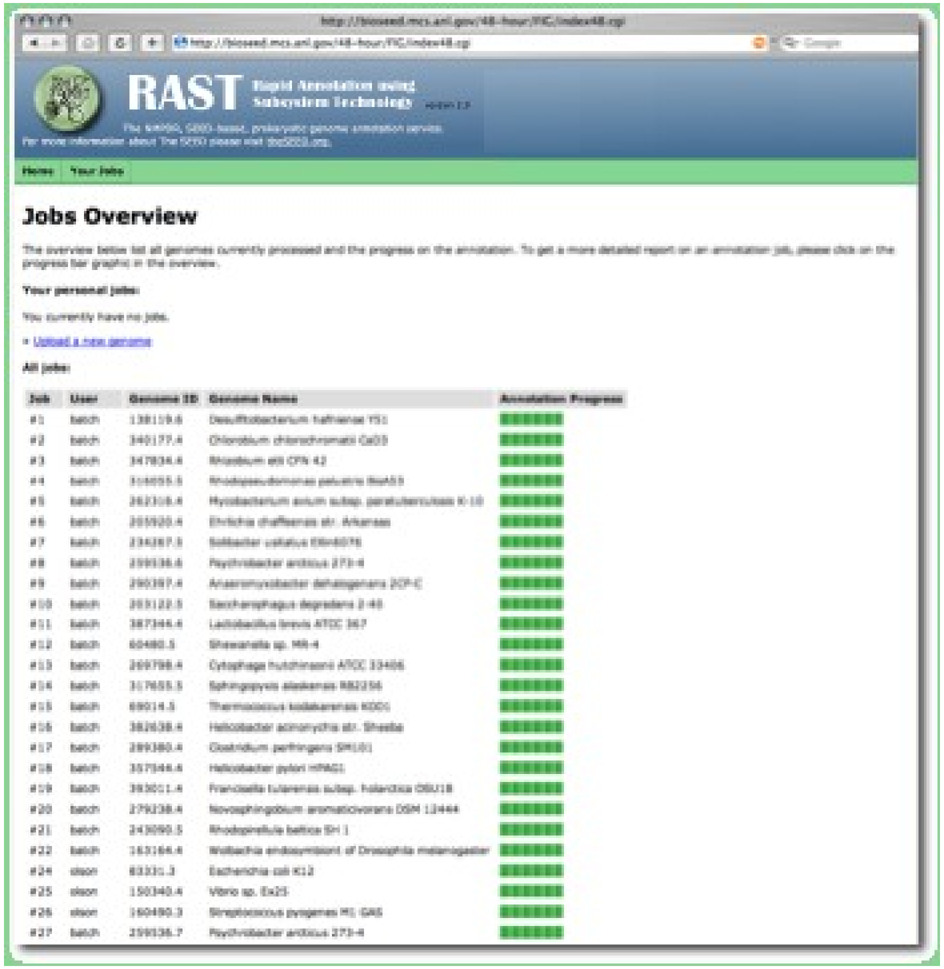
Annotations vs. sequences



Annotation of Complete Genomes

http://rast.nmpdr.org/

- Automated user originated processing
- Takes 1-7 hours depending on size and complexity of the genome
- ~2,000 external submissions, including hundreds of genomes not yet publicly released.
- Reannotation of >500 genomes complete
- 1,000 users, 200 organizations, 25 countries.



The NMPDR, SEED-based, prokaryotic genome annotation service.
For more information about The SEED please visit theSEED.org.

Jobs Overview

The overview below list all genomes currently processed and the progress on the annotation. To get a more detailed report on an annotation job, please click on the progress bar graphic in the overview.

Your personal jobs:

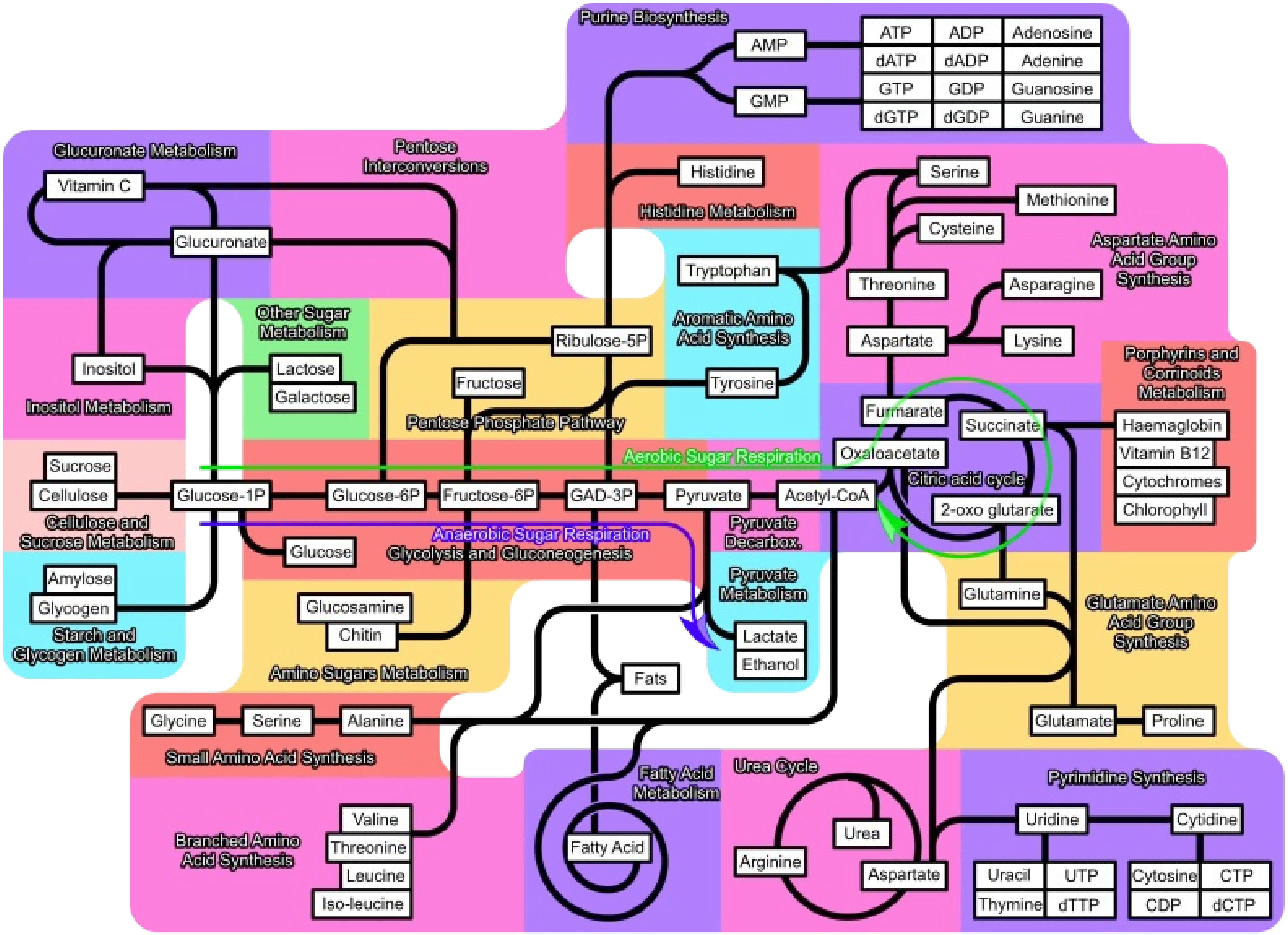
You currently have no jobs.

- Upload a new genome

All jobs:

Job	User	Genome ID	Genome Name	Annotation Progress
#1	batch	138119.6	Desulfitobacterium hafniense Y51	🟩
#2	batch	340177.4	Chlorobium chlorochromatii CaD3	🟩
#3	batch	347834.4	Rhizobium etli CFN 42	🟩
#4	batch	316055.3	Rhodopseudomonas palustris BisA53	🟩
#5	batch	262316.4	Mycobacterium avium subsp. paratuberculosis K-10	🟩
#6	batch	205920.4	Ehrlichia chaffeensis str. Arkansas	🟩
#7	batch	234267.3	Solibacter usitatus Ellin6076	🟩
#8	batch	259536.6	Psychrobacter arcticus 273-4	🟩
#9	batch	290397.4	Anaeromyxobacter dehalogenans 2CP-C	🟩
#10	batch	203122.3	Saccharophagus degradans 2-40	🟩
#11	batch	387344.4	Lactobacillus brevis ATCC 367	🟩
#12	batch	60480.3	Shewanella sp. MR-4	🟩
#13	batch	269798.4	Cytophaga hutchinsonii ATCC 33406	🟩
#14	batch	317655.3	Sphingopyxis alaskensis RB2256	🟩
#15	batch	69014.3	Thermococcus kodakarensis KOD1	🟩
#16	batch	382638.4	Helicobacter acinonychis str. Sheeba	🟩
#17	batch	289380.4	Clostridium perfringens SM101	🟩
#18	batch	357544.4	Helicobacter pylori HPAG1	🟩
#19	batch	393011.4	Francisella tularensis subsp. holarctica OSU18	🟩
#20	batch	279238.4	Novosphingobium aromaticivorans DSM 12444	🟩
#21	batch	243090.3	Rhodopirellula baltica SH 1	🟩
#22	batch	163164.4	Wolbachia endosymbiont of Drosophila melanogaster	🟩
#24	olson	83333.3	Escherichia coli K12	🟩
#25	olson	150340.4	Vibrio sp. Ex25	🟩
#26	olson	160490.3	Streptococcus pyogenes M1 GAS	🟩
#27	batch	259536.7	Psychrobacter arcticus 273-4	🟩

Subsystems Make Up Metabolism



Wikipedia Metabolism
http://en.wikipedia.org/wiki/Portal:Metabolism

Subsystem spreadsheet (conceptually)

	Chaperone	**Subunit**	**Usher**	**Adhesin**
S. enterica Enteritidis	2389	2388	2387	2386
E. coli HS	3068	3067	3066	3065
B. cenocepacia J2315	2604	2603	2602	2601
S. maltophilia	1085	1088	1087	1086

Over 1,000 Subsystems

Three level “hierarchy”

- Amino Acids and Derivatives
  - Alanine, serine, and glycine
    - Serine Biosynthesis

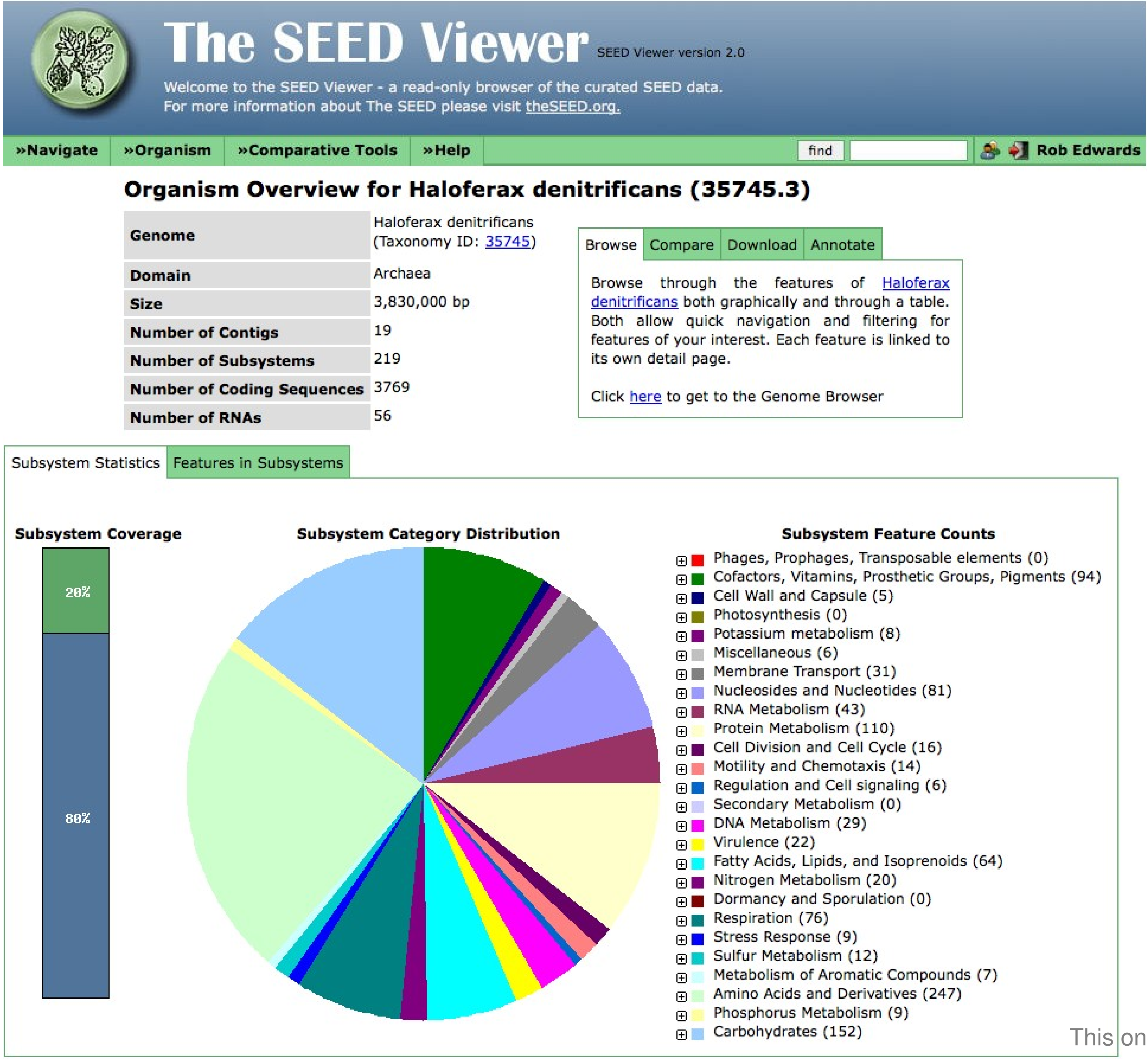
- Amino Acids and Derivatives
  - Lysine, threonine, methionine, and cysteine
    - Methionine Biosynthesis

Make your own subsystems!

Class	# SS	Class	# SS
Amino Acids and Derivatives	56	Nucleosides and Nucleotides	14
Carbohydrates	97	Phosphorus Metabolism	6
Cell Division / Cycle	10	Photosynthesis	9
Cell Wall and Capsule	50	Potassium metabolism	3
Clustering-based ss	193	Protein Metabolism	52
Cofactors, Vitamins, Pigments	43	RNA Metabolism	39
DNA Metabolism	30	Regulation/signaling	23
Fatty Acids, Lipids, and Isoprenoids	22	Respiration	44
Membrane Transport	41	Secondary Metabolism	24
Metabolism of Aromatic Compounds	30	Stress Response	37
Motility and Chemotaxis	8	Sulfur Metabolism	12
Nitrogen Metabolism	11	Virulence	116

The annotation process (complete genomes)

- Find the phylogenetic neighborhood of your genome
- Look for proteins that related organisms have
  - Core proteins
  - Subset of all subsystems
- Use those calls as a training set for critica/glimmer
  - Intrinsic training set!



The SEED Viewer

SEED Viewer version 2.0

Welcome to the SEED Viewer - a read-only browser of the curated SEED data.
For more information about The SEED please visit theSEED.org.

Organism Overview for Haloferax denitrificans (35745.3)

Genome	Haloferax denitrificans (Taxonomy ID: 35745)
Domain	Archaea
Size	3,830,000 bp
Number of Contigs	19
Number of Subsystems	219
Number of Coding Sequences	3769
Number of RNAs	56

Browse through the features of Haloferax denitrificans both graphically and through a table. Both allow quick navigation and filtering for features of your interest. Each feature is linked to its own detail page.

Click here to get to the Genome Browser

Subsystem Feature Counts

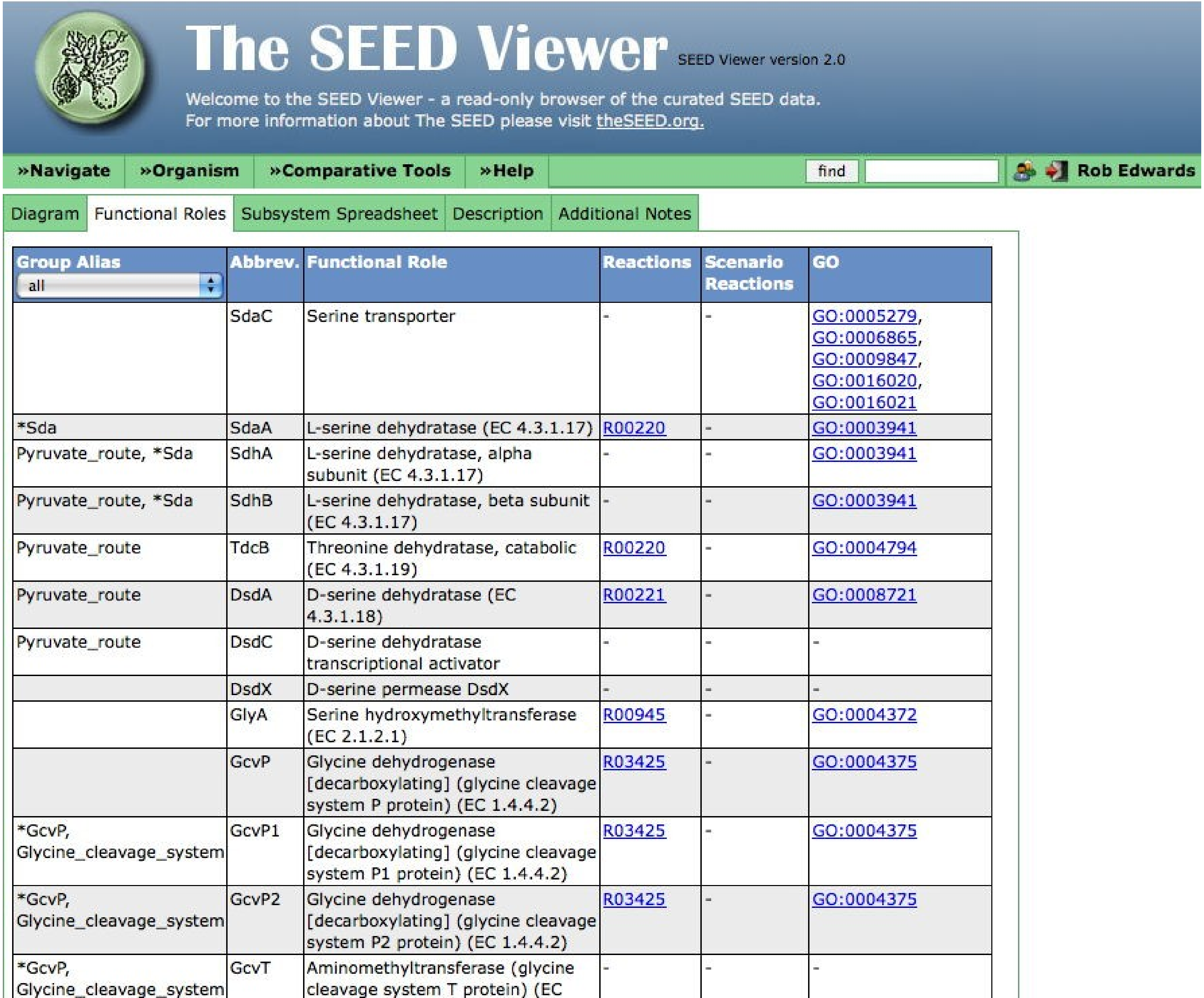
- ⊞ Phages, Prophages, Transposable elements (0)
- ⊞ Cofactors, Vitamins, Prosthetic Groups, Pigments (94)
- ⊞ Cell Wall and Capsule (5)
- ⊞ Photosynthesis (0)
- ⊞ Potassium metabolism (8)
- ⊞ Miscellaneous (6)
- ⊞ Membrane Transport (31)
- ⊞ Nucleosides and Nucleotides (81)
- ⊞ RNA Metabolism (43)
- ⊞ Protein Metabolism (110)
- ⊞ Cell Division and Cell Cycle (16)
- ⊞ Motility and Chemotaxis (14)
- ⊞ Regulation and Cell signaling (6)
- ⊞ Secondary Metabolism (0)
- ⊞ DNA Metabolism (29)
- ⊞ Virulence (22)
- ⊞ Fatty Acids, Lipids, and Isoprenoids (64)
- ⊞ Nitrogen Metabolism (20)
- ⊞ Dormancy and Sporulation (0)
- ⊞ Respiration (76)
- ⊞ Stress Response (9)
- ⊞ Sulfur Metabolism (12)
- ⊞ Metabolism of Aromatic Compounds (7)
- ⊞ Amino Acids and Derivatives (247)
- ⊞ Phosphorus Metabolism (9)
- ⊞ Carbohydrates (152)

This one’s for Gary

Automatic metabolic reconstruction

- Subsystem, GO, and KEGG connections
  - KEGG EC numbers
  - KEGG reaction numbers
  - SEED reaction numbers (Chris Henry)

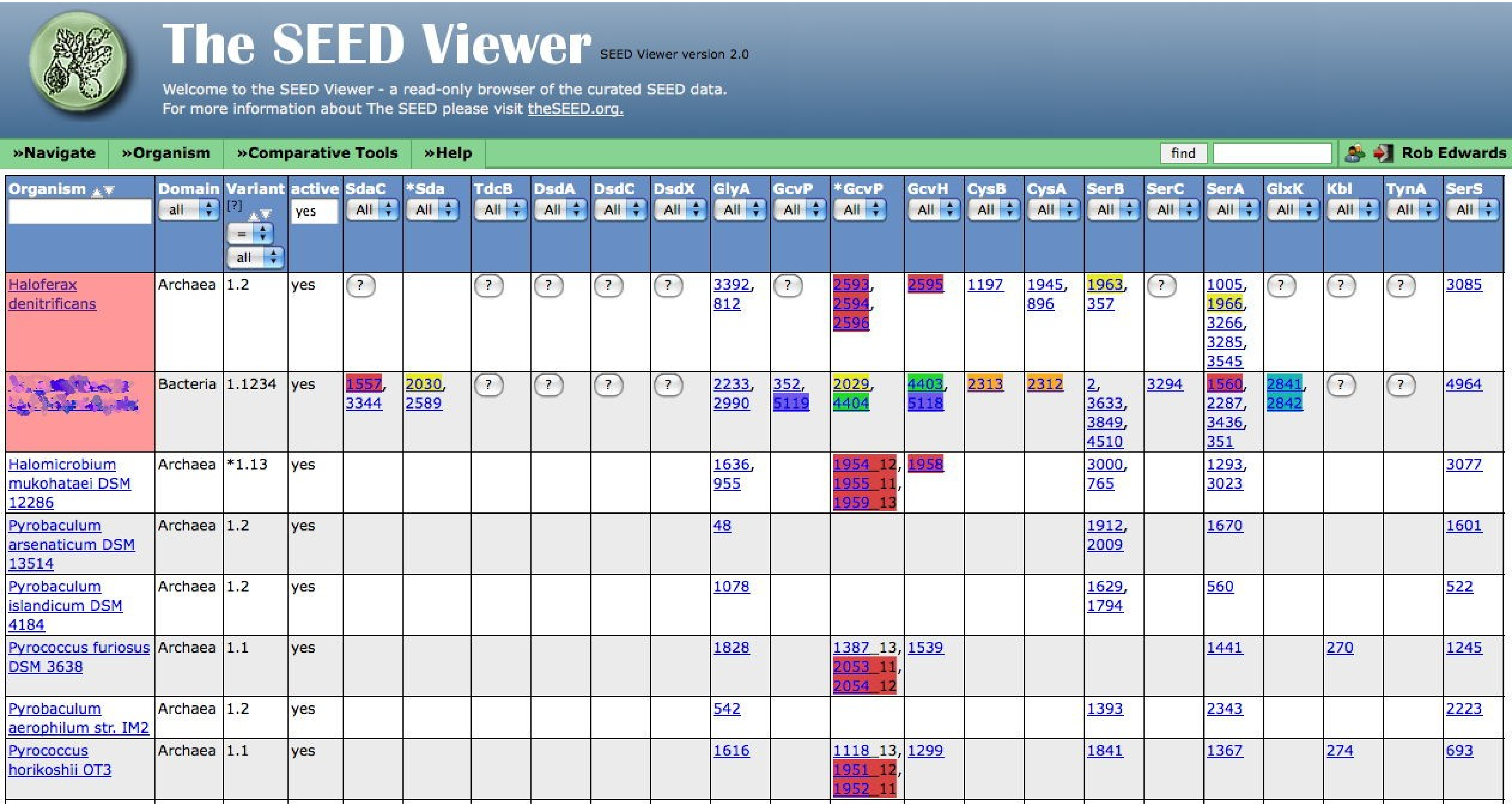
- Metabolic flux models
  - Automatically generate FBA matrices (Aaron Best/Matt DeJongh; Hope College)



Welcome to the SEED Viewer - a read-only browser of the curated SEED data.
For more information about The SEED please visit theSEED.org.

Group Alias all	Abbrev.	Functional Role	Reactions	Scenario Reactions	GO
	SdaC	Serine transporter	-	-	GO:0005279, GO:0006865, GO:0009847, GO:0016020, GO:0016021
*Sda	SdaA	L-serine dehydratase (EC 4.3.1.17)	R00220	-	GO:0003941
Pyruvate_route, *Sda	SdhA	L-serine dehydratase, alpha subunit (EC 4.3.1.17)	-	-	GO:0003941
Pyruvate_route, *Sda	SdhB	L-serine dehydratase, beta subunit (EC 4.3.1.17)	-	-	GO:0003941
Pyruvate_route	TdcB	Threonine dehydratase, catabolic (EC 4.3.1.19)	R00220	-	GO:0004794
Pyruvate_route	DsdA	D-serine dehydratase (EC 4.3.1.18)	R00221	-	GO:0008721
Pyruvate_route	DsdC	D-serine dehydratase transcriptional activator	-	-	-
	DsdX	D-serine permease DsdX	-	-	-
	GlyA	Serine hydroxymethyltransferase (EC 2.1.2.1)	R00945	-	GO:0004372
	GcvP	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P protein) (EC 1.4.4.2)	R03425	-	GO:0004375
*GcvP, Glycine_cleavage_system	GcvP1	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P1 protein) (EC 1.4.4.2)	R03425	-	GO:0004375
*GcvP, Glycine_cleavage_system	GcvP2	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)	R03425	-	GO:0004375
*GcvP, Glycine_cleavage_system	GcvT	Aminomethyltransferase (glycine cleavage system T protein) (EC	-	-	-

The Populated Subsystem

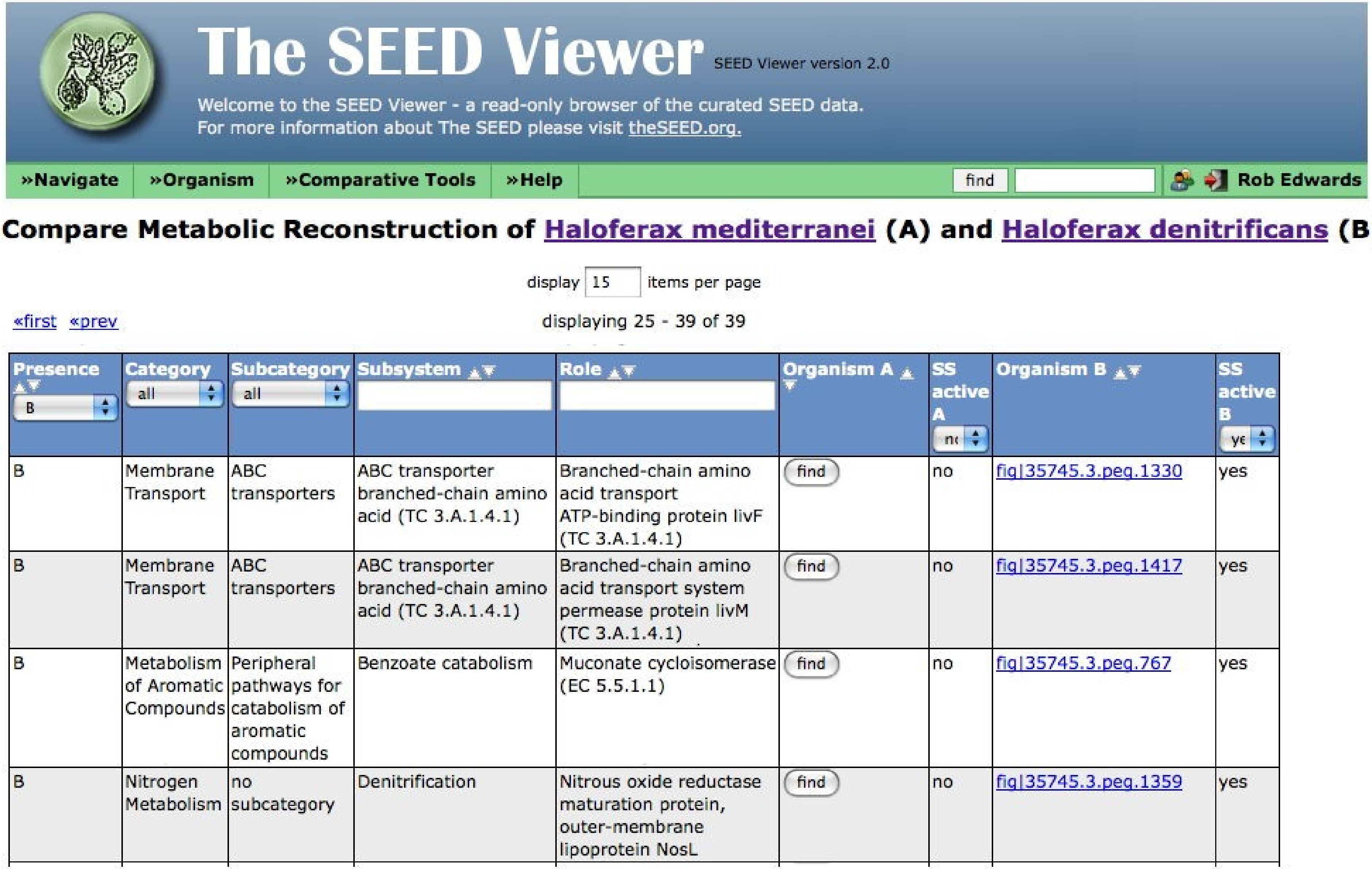


The SEED Viewer SEED Viewer version 2.0

Welcome to the SEED Viewer - a read-only browser of the curated SEED data.
For more information about The SEED please visit theSEED.org.

Organism	Domain	Variant [?]	active	SdaC	*Sda	TdcB	DsdA	DsdC	DsdX	GlyA	GcvP	*GcvP	GcvH	CysB	CysA	SerB	SerC	SerA	GlxK	Kbl	TynA	SerS
	all	= all	yes	All	All	All	All	All	All	All	All	All	All	All	All	All	All	All	All	All	All	All
Haloferax denitrificans	Archaea	1.2	yes	?		?	?	?	?	3392, 812	?	2593, 2594, 2596	2595	1197	1945, 896	1963, 357	?	1005, 1966, 3266, 3285, 3545	?	?	?	3085
[redacted]	Bacteria	1.1234	yes	1557, 3344	2030, 2589	?	?	?	?	2233, 2990	352, 5119	2029, 4404	4403, 5118	2313	2312	2, 3633, 3849, 4510	3294	1560, 2287, 3436, 351	2841, 2842	?	?	4964
Halomicrobium mukohataei DSM 12286	Archaea	*1.13	yes							1636, 955		1954_12, 1955_11, 1959_13	1958			3000, 765		1293, 3023				3077
Pyrobaculum arsenaticum DSM 13514	Archaea	1.2	yes							48						1912, 2009		1670				1601
Pyrobaculum islandicum DSM 4184	Archaea	1.2	yes							1078						1629, 1794		560				522
Pyrococcus furiosus DSM 3638	Archaea	1.1	yes							1828		1387_13, 2053_11, 2054_12	1539					1441		270		1245
Pyrobaculum aerophilum str. IM2	Archaea	1.2	yes							542						1393		2343				2223
Pyrococcus horikoshii OT3	Archaea	1.1	yes							1616		1118_13, 1951_12, 1952_11	1299			1841		1367		274		693

Automatically compare metabolic reconstructions



Welcome to the SEED Viewer - a read-only browser of the curated SEED data.
For more information about The SEED please visit theSEED.org.

Compare Metabolic Reconstruction of Haloferax mediterranei (A) and Haloferax denitrificans (B)

display 15 items per page

«first «prev

displaying 25 - 39 of 39

Presence	Category	Subcategory	Subsystem	Role	Organism A	SS active A	Organism B	SS active B
B	Membrane Transport	ABC transporters	ABC transporter branched-chain amino acid (TC 3.A.1.4.1)	Branched-chain amino acid transport ATP-binding protein livF (TC 3.A.1.4.1)	find	no	fig|35745.3.peg.1330	yes
B	Membrane Transport	ABC transporters	ABC transporter branched-chain amino acid (TC 3.A.1.4.1)	Branched-chain amino acid transport system permease protein livM (TC 3.A.1.4.1)	find	no	fig|35745.3.peg.1417	yes
B	Metabolism of Aromatic Compounds	Peripheral pathways for catabolism of aromatic compounds	Benzoate catabolism	Muconate cycloisomerase (EC 5.5.1.1)	find	no	fig|35745.3.peg.767	yes
B	Nitrogen Metabolism	no subcategory	Denitrification	Nitrous oxide reductase maturation protein, outer-membrane lipoprotein NosL	find	no	fig|35745.3.peg.1359	yes

Find and suggest candidate functions

- Rapidly correct missing annotations

- Add more members to subsystems

- Improves future genome annotations!

  (especially with new subsystems)

The Live ASM Test

Philadelphia, 2009

- 10 genomes submitted on Thursday at 6 pm
- First annotation complete before 8 am Friday
- Remaining annotations completed Friday before noon
- (there were others in the pipeline too!)
- Presentation ASM 2009 Tuesday, 8pm

Subsystems coverage of sequenced Archaea

Genome	Percent of Proteins in Subsystems
Haloferax denitrificans	20%
Haloferax mediterranei	19%
Haloferax sulfurifontis	19%
Haloferax volcanii DS2	19%
Haloarcula sp 33800	19%
Haloarcula sp 33799	18%

Prophages

PHANTOME

Mya Breitbart, Matt Sullivan, Jeff Elhai, Rob Edwards

NSF

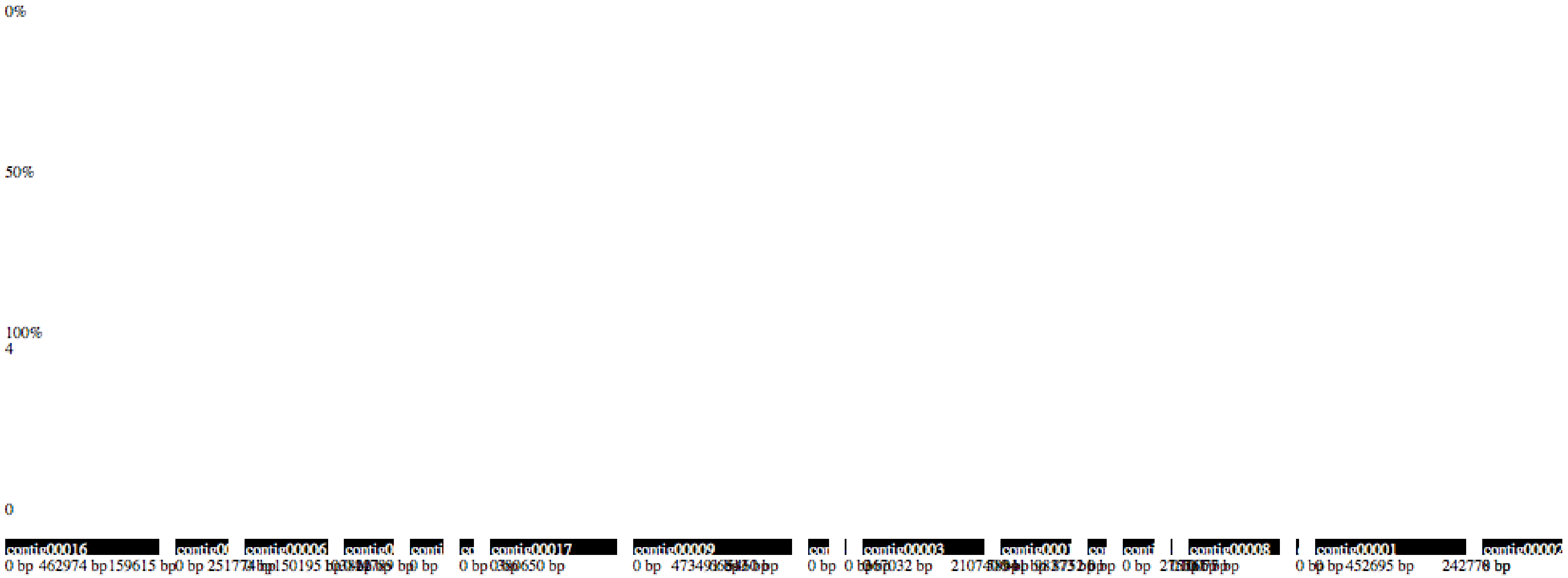
Haloferax sulfurifontis prophage

Comparing complete genomes to metagenomes

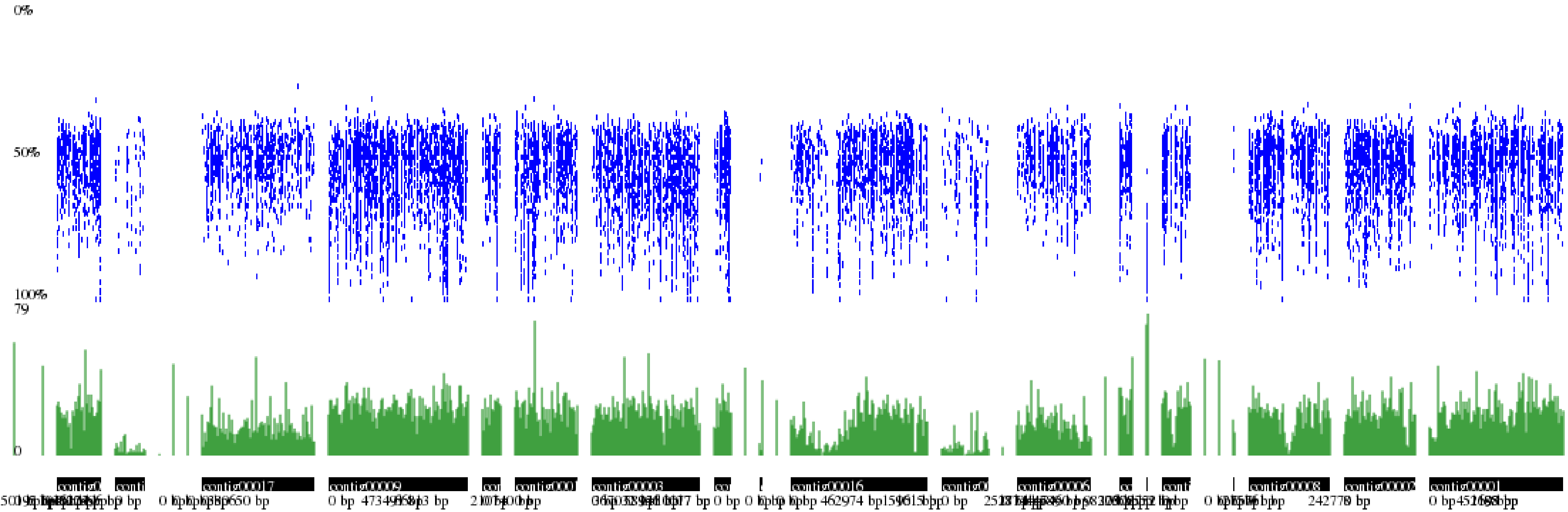
Metagenomics RAST has 300 public metagenomes

Compared using tblastx

Human Poop


0%
50%
100%
4
0
contig00016
0 bp
462974 bp
159615 bp
contig00
0 bp
251774 bp
contig00006
150195 bp
contig0
conti
0 bp
co
0 bp
contig00017
390650 bp
contig00009
0 bp
473493 bp
con
0 bp
contig00003
367032 bp
210745 bp
contig0001
contig00008
contig00001
0 bp
452695 bp
242778 bp
contig00002
0 bp

High Salinity Salterns
San Diego, July 2004

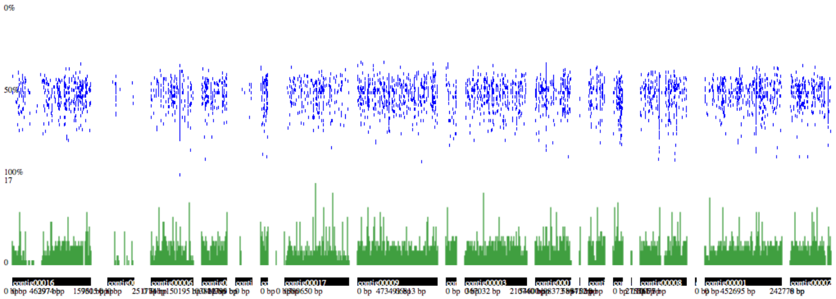


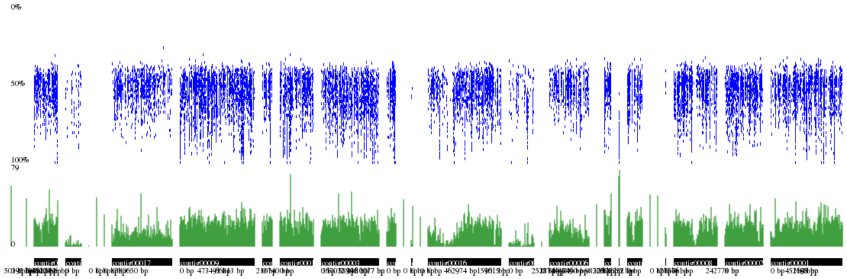
Thanks Nick Celms, Beltran Rodriguez-Mueller, Mya Breitbart, & Forest Rohwer

Low salinity salterns

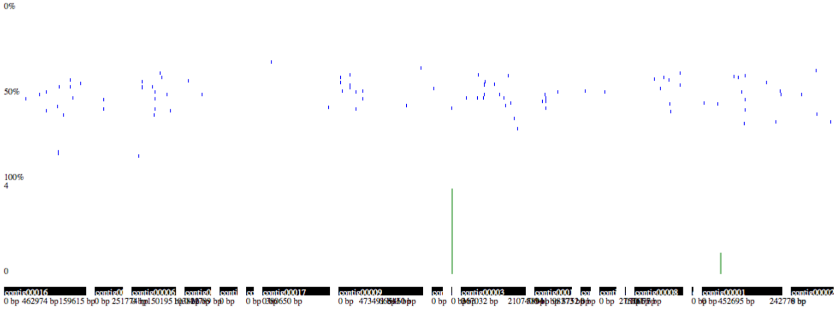
High salinity salterns

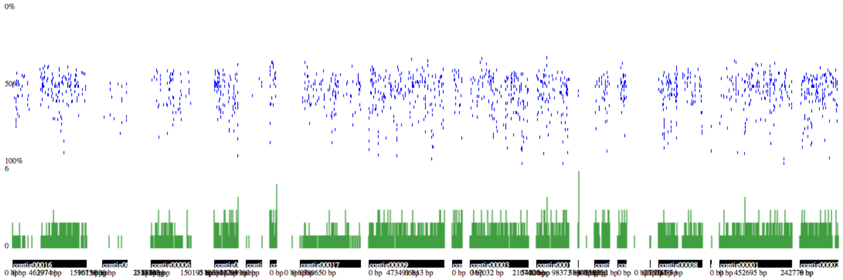
July 2004



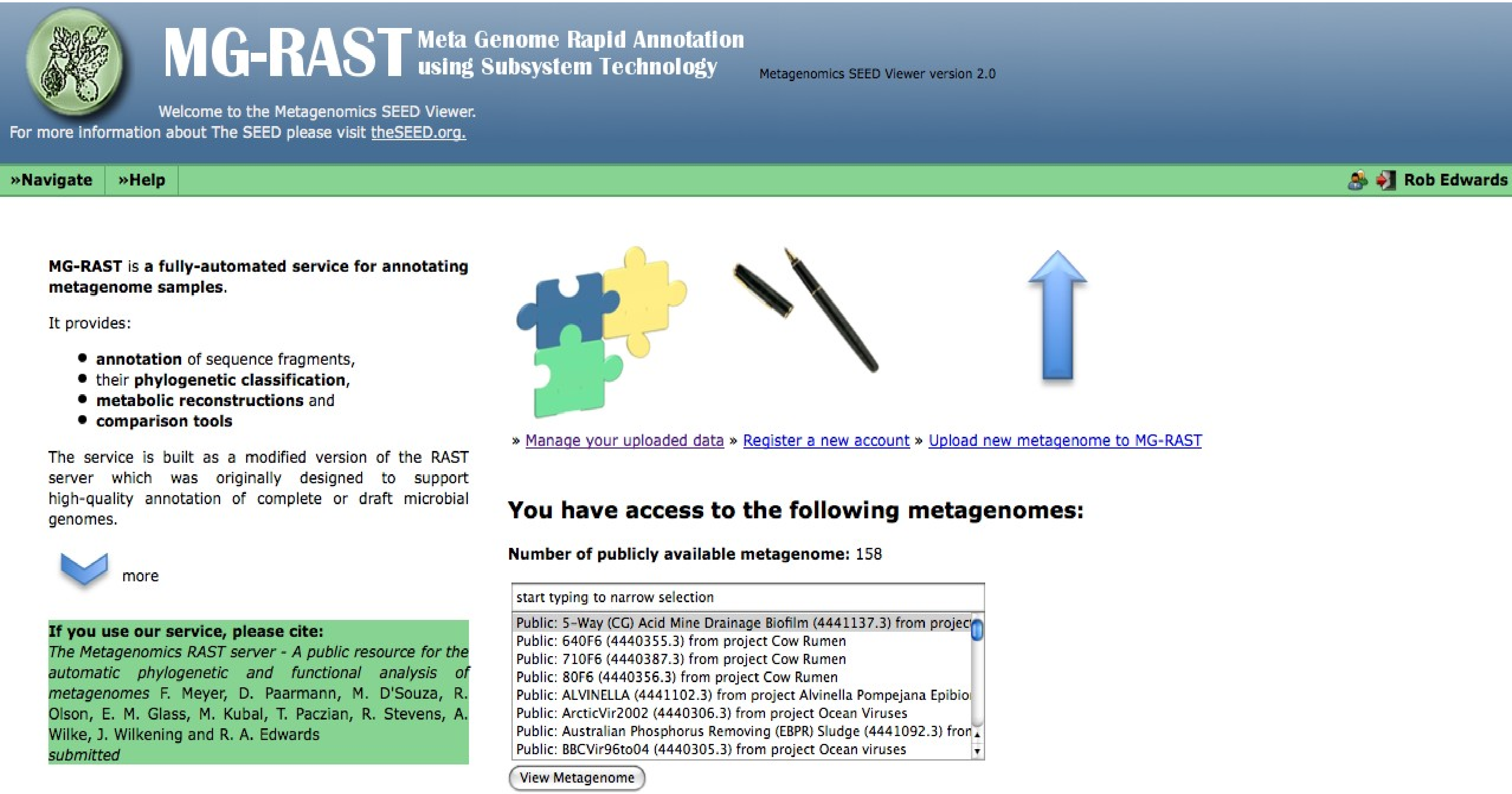


Nov 2005





The metagenomics RAST server



MG-RAST Meta Genome Rapid Annotation using Subsystem Technology

Metagenomics SEED Viewer version 2.0

Welcome to the Metagenomics SEED Viewer.
For more information about The SEED please visit theSEED.org.

MG-RAST is **a fully-automated service for annotating metagenome samples**.

It provides:

- **annotation** of sequence fragments,
- their **phylogenetic classification**,
- **metabolic reconstructions** and
- **comparison tools**

The service is built as a modified version of the RAST server which was originally designed to support high-quality annotation of complete or draft microbial genomes.

If you use our service, please cite:
The Metagenomics RAST server - A public resource for the automatic phylogenetic and functional analysis of metagenomes F. Meyer, D. Paarmann, M. D'Souza, R. Olson, E. M. Glass, M. Kubal, T. Paczian, R. Stevens, A. Wilke, J. Wilkening and R. A. Edwards
submitted

» Manage your uploaded data » Register a new account » Upload new metagenome to MG-RAST

You have access to the following metagenomes:

Number of publicly available metagenome: 158

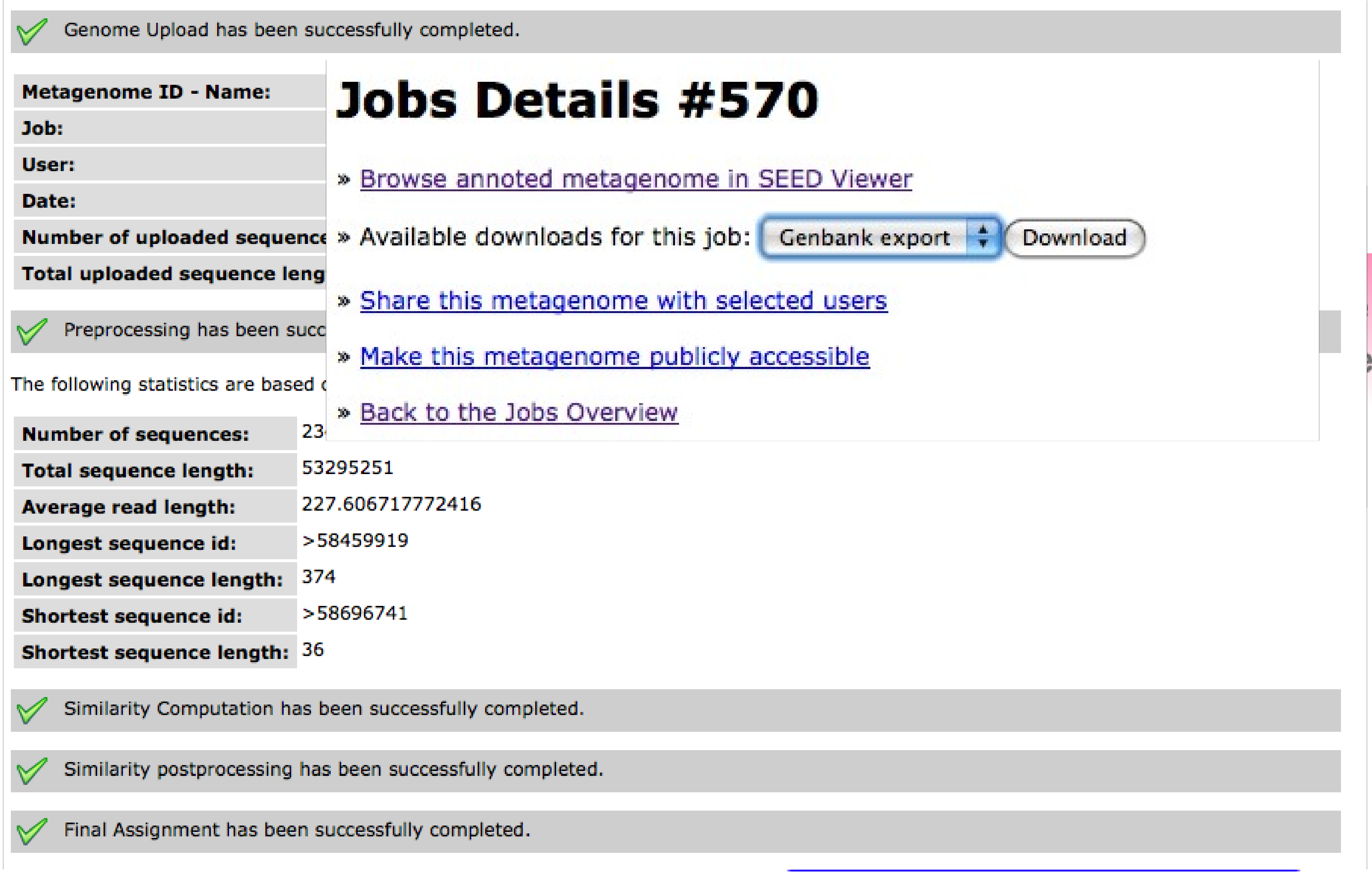
start typing to narrow selection

Public: 5-Way (CG) Acid Mine Drainage Biofilm (4441137.3) from projec
Public: 640F6 (4440355.3) from project Cow Rumen
Public: 710F6 (4440387.3) from project Cow Rumen
Public: 80F6 (4440356.3) from project Cow Rumen
Public: ALVINELLA (4441102.3) from project Alvinella Pompejana Epibio
Public: ArcticVir2002 (4440306.3) from project Ocean Viruses
Public: Australian Phosphorus Removing (EBPR) Sludge (4441092.3) fron
Public: BBCVir96to04 (4440305.3) from project Ocean viruses

Automated Processing

User uploads
DNA sequences

Pre-processing:



✓ Genome Upload has been successfully completed.

Jobs Details #570

» Browse annoted metagenome in SEED Viewer

» Available downloads for this job: Genbank export Download

» Share this metagenome with selected users

» Make this metagenome publicly accessible

» Back to the Jobs Overview

Metagenome ID - Name:	
Job:	
User:	
Date:	
Number of uploaded sequence	
Total uploaded sequence leng	

✓ Preprocessing has been succ

The following statistics are based c

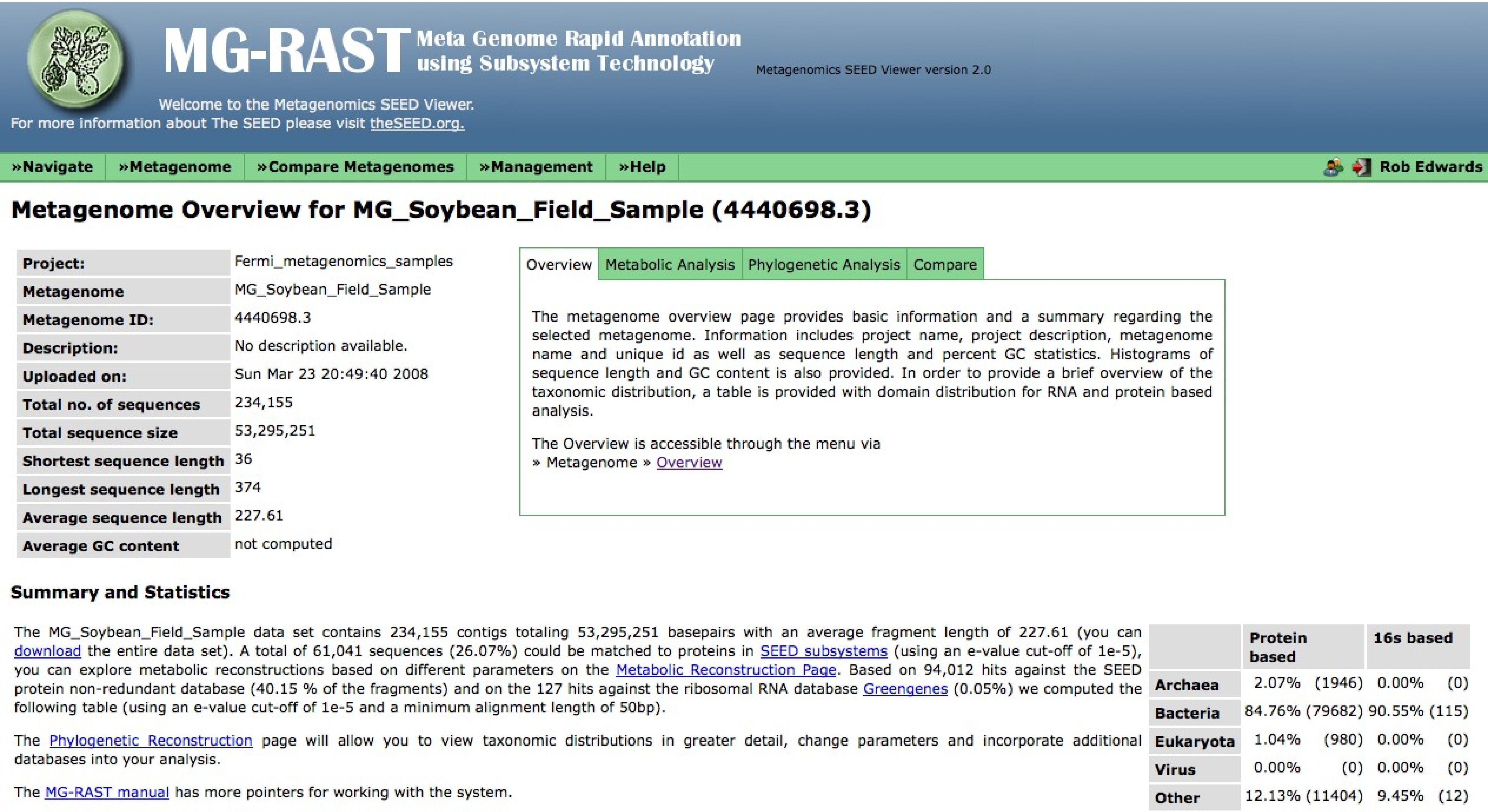
Number of sequences:	23
Total sequence length:	53295251
Average read length:	227.606717772416
Longest sequence id:	>58459919
Longest sequence length:	374
Shortest sequence id:	>58696741
Shortest sequence length:	36

✓ Similarity Computation has been successfully completed.

✓ Similarity postprocessing has been successfully completed.

✓ Final Assignment has been successfully completed.

Summary View



Welcome to the Metagenomics SEED Viewer.
For more information about The SEED please visit theSEED.org.

Metagenome Overview for MG_Soybean_Field_Sample (4440698.3)

Project:	Fermi_metagenomics_samples
Metagenome	MG_Soybean_Field_Sample
Metagenome ID:	4440698.3
Description:	No description available.
Uploaded on:	Sun Mar 23 20:49:40 2008
Total no. of sequences	234,155
Total sequence size	53,295,251
Shortest sequence length	36
Longest sequence length	374
Average sequence length	227.61
Average GC content	not computed

The metagenome overview page provides basic information and a summary regarding the selected metagenome. Information includes project name, project description, metagenome name and unique id as well as sequence length and percent GC statistics. Histograms of sequence length and GC content is also provided. In order to provide a brief overview of the taxonomic distribution, a table is provided with domain distribution for RNA and protein based analysis.

The Overview is accessible through the menu via
» Metagenome » Overview

Summary and Statistics

The MG_Soybean_Field_Sample data set contains 234,155 contigs totaling 53,295,251 basepairs with an average fragment length of 227.61 (you can download the entire data set). A total of 61,041 sequences (26.07%) could be matched to proteins in SEED subsystems (using an e-value cut-off of 1e-5), you can explore metabolic reconstructions based on different parameters on the Metabolic Reconstruction Page. Based on 94,012 hits against the SEED protein non-redundant database (40.15 % of the fragments) and on the 127 hits against the ribosomal RNA database Greengenes (0.05%) we computed the following table (using an e-value cut-off of 1e-5 and a minimum alignment length of 50bp).

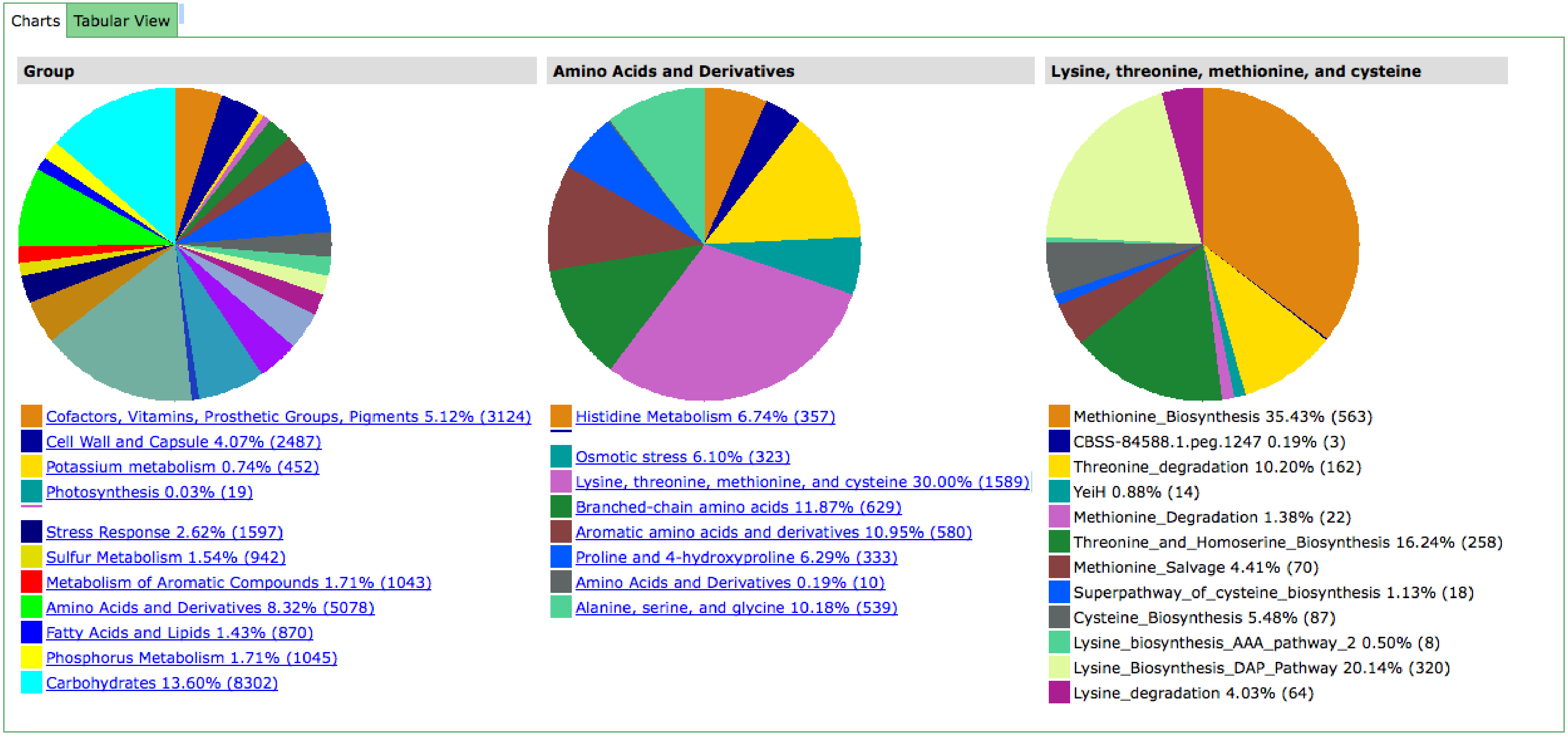
The Phylogenetic Reconstruction page will allow you to view taxonomic distributions in greater detail, change parameters and incorporate additional databases into your analysis.

The MG-RAST manual has more pointers for working with the system.

	Protein based	**16s based**
Archaea	2.07% (1946)	0.00% (0)
Bacteria	84.76% (79682)	90.55% (115)
Eukaryota	1.04% (980)	0.00% (0)
Virus	0.00% (0)	0.00% (0)
Other	12.13% (11404)	9.45% (12)

www.nmpdr.org www.theseed.org

Metagenomics Tools
Annotation & Subsystems



Group

- Cofactors, Vitamins, Prosthetic Groups, Pigments 5.12% (3124)
- Cell Wall and Capsule 4.07% (2487)
- Potassium metabolism 0.74% (452)
- Photosynthesis 0.03% (19)
- Stress Response 2.62% (1597)
- Sulfur Metabolism 1.54% (942)
- Metabolism of Aromatic Compounds 1.71% (1043)
- Amino Acids and Derivatives 8.32% (5078)
- Fatty Acids and Lipids 1.43% (870)
- Phosphorus Metabolism 1.71% (1045)
- Carbohydrates 13.60% (8302)

Amino Acids and Derivatives

- Histidine Metabolism 6.74% (357)
- Osmotic stress 6.10% (323)
- Lysine, threonine, methionine, and cysteine 30.00% (1589)
- Branched-chain amino acids 11.87% (629)
- Aromatic amino acids and derivatives 10.95% (580)
- Proline and 4-hydroxyproline 6.29% (333)
- Amino Acids and Derivatives 0.19% (10)
- Alanine, serine, and glycine 10.18% (539)

Lysine, threonine, methionine, and cysteine

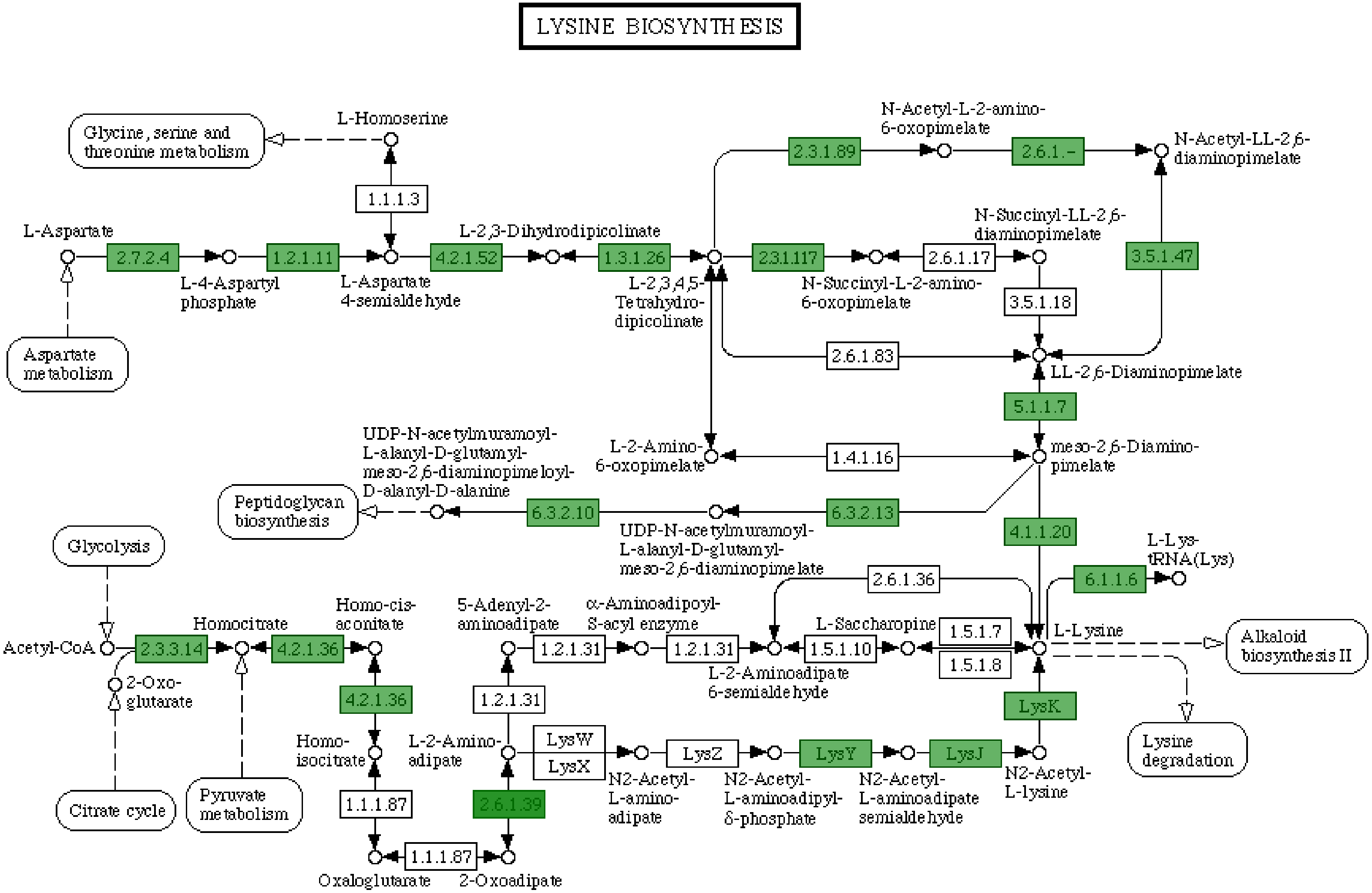
- Methionine_Biosynthesis 35.43% (563)
- CBSS-84588.1.peg.1247 0.19% (3)
- Threonine_degradation 10.20% (162)
- YeiH 0.88% (14)
- Methionine_Degradation 1.38% (22)
- Threonine_and_Homoserine_Biosynthesis 16.24% (258)
- Methionine_Salvage 4.41% (70)
- Superpathway_of_cysteine_biosynthesis 1.13% (18)
- Cysteine_Biosynthesis 5.48% (87)
- Lysine_biosynthesis_AAA_pathway_2 0.50% (8)
- Lysine_Biosynthesis_DAP_Pathway 20.14% (320)
- Lysine_degradation 4.03% (64)

SAN DIEGO STATE UNIVERSITY

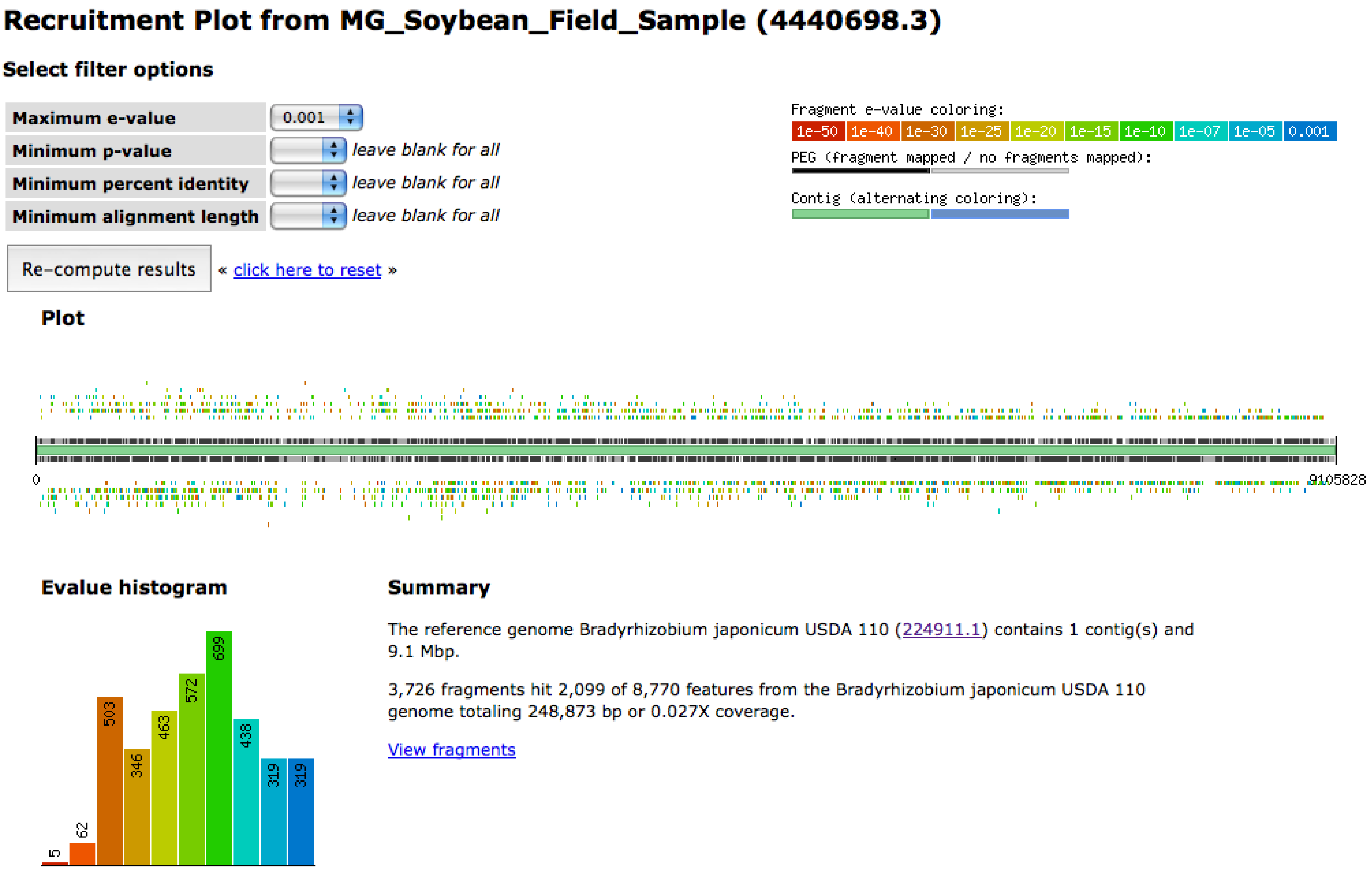
www.nmpdr.org www.theseed.org

NMPDR National Microbial Pathogen Data Resource Center

Metagenomics Tools
Annotation & KEGG maps



Metagenomics Tools
Recruitment Plots



Recruitment Plot from MG_Soybean_Field_Sample (4440698.3)

Select filter options

Maximum e-value	0.001	
Minimum p-value		*leave blank for all*
Minimum percent identity		*leave blank for all*
Minimum alignment length		*leave blank for all*

Fragment e-value coloring:

1e-50 | 1e-40 | 1e-30 | 1e-25 | 1e-20 | 1e-15 | 1e-10 | 1e-07 | 1e-05 | 0.001

PEG (fragment mapped / no fragments mapped):

Contig (alternating coloring):

Re-compute results « click here to reset »

Plot

Evalue histogram

Summary

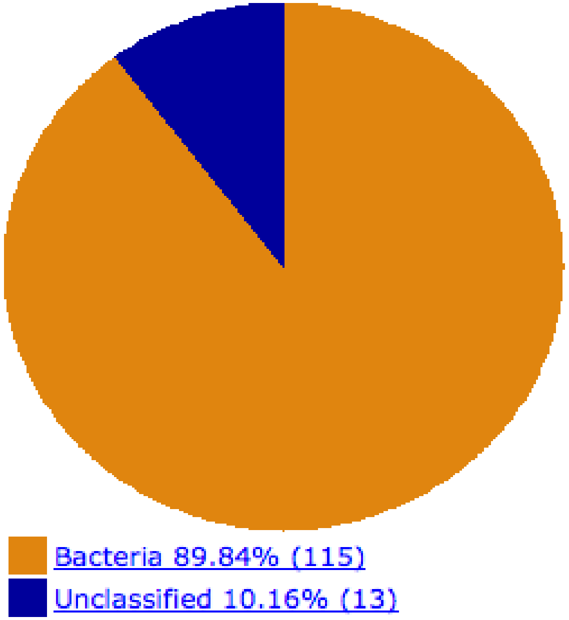
The reference genome Bradyrhizobium japonicum USDA 110 (224911.1) contains 1 contig(s) and 9.1 Mbp.

3,726 fragments hit 2,099 of 8,770 features from the Bradyrhizobium japonicum USDA 110 genome totaling 248,873 bp or 0.027X coverage.

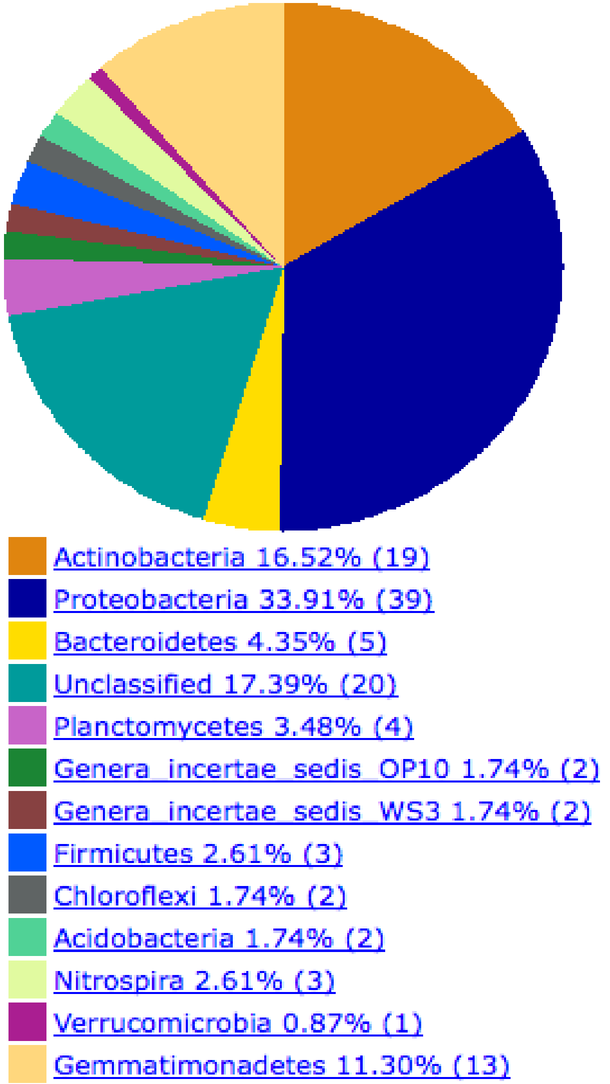
View fragments

Metagenomics Tools
Phylogenetic Reconstruction

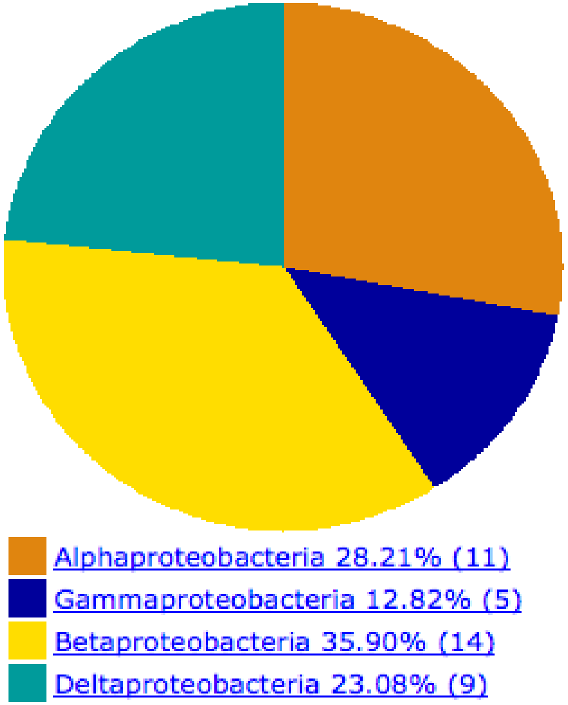
Domain



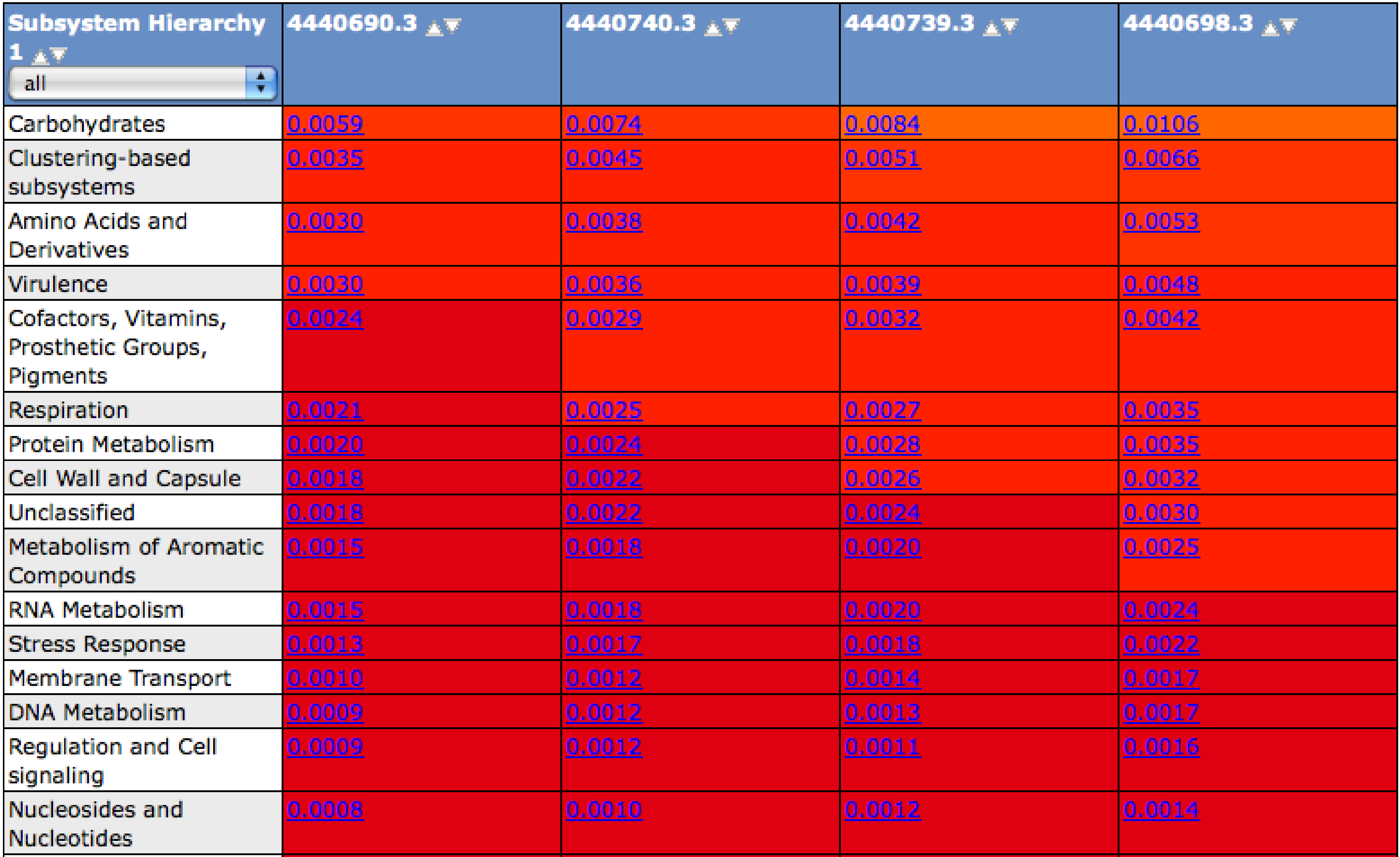
Bacteria



Proteobacteria



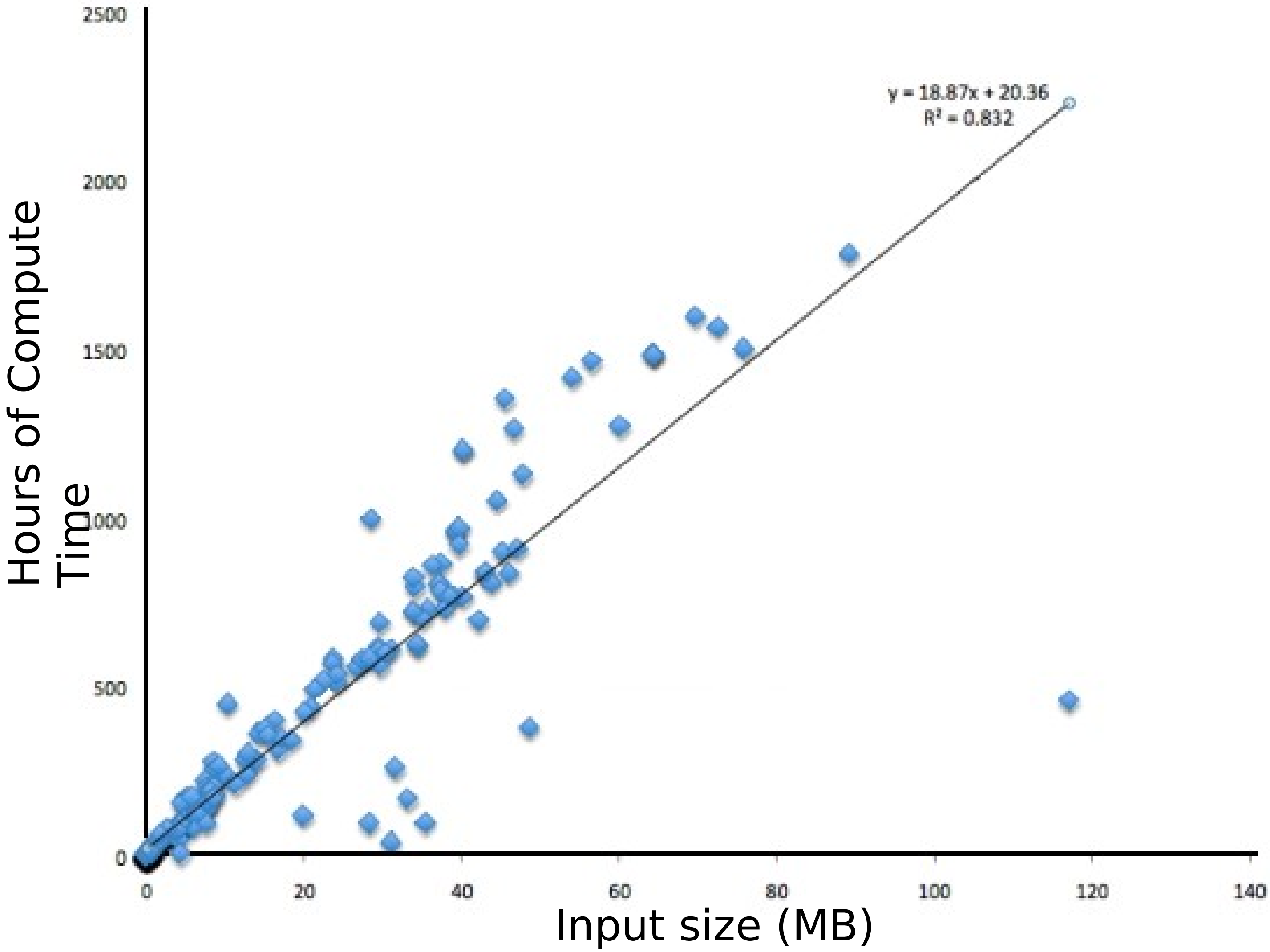
Metagenomics Tools
Comparative Tools



Subsystem Hierarchy 1 ▲▼ all	4440690.3 ▲▼	4440740.3 ▲▼	4440739.3 ▲▼	4440698.3 ▲▼
Carbohydrates	0.0059	0.0074	0.0084	0.0106
Clustering-based subsystems	0.0035	0.0045	0.0051	0.0066
Amino Acids and Derivatives	0.0030	0.0038	0.0042	0.0053
Virulence	0.0030	0.0036	0.0039	0.0048
Cofactors, Vitamins, Prosthetic Groups, Pigments	0.0024	0.0029	0.0032	0.0042
Respiration	0.0021	0.0025	0.0027	0.0035
Protein Metabolism	0.0020	0.0024	0.0028	0.0035
Cell Wall and Capsule	0.0018	0.0022	0.0026	0.0032
Unclassified	0.0018	0.0022	0.0024	0.0030
Metabolism of Aromatic Compounds	0.0015	0.0018	0.0020	0.0025
RNA Metabolism	0.0015	0.0018	0.0020	0.0024
Stress Response	0.0013	0.0017	0.0018	0.0022
Membrane Transport	0.0010	0.0012	0.0014	0.0017
DNA Metabolism	0.0009	0.0012	0.0013	0.0017
Regulation and Cell signaling	0.0009	0.0012	0.0011	0.0016
Nucleosides and Nucleotides	0.0008	0.0010	0.0012	0.0014

Computational Requirements

~19 hours of compute per input megabyte



How much so far

Total:

2,854 metagenomes
265,126,281 sequences
68,490,338,939 bp (68 Gbp)

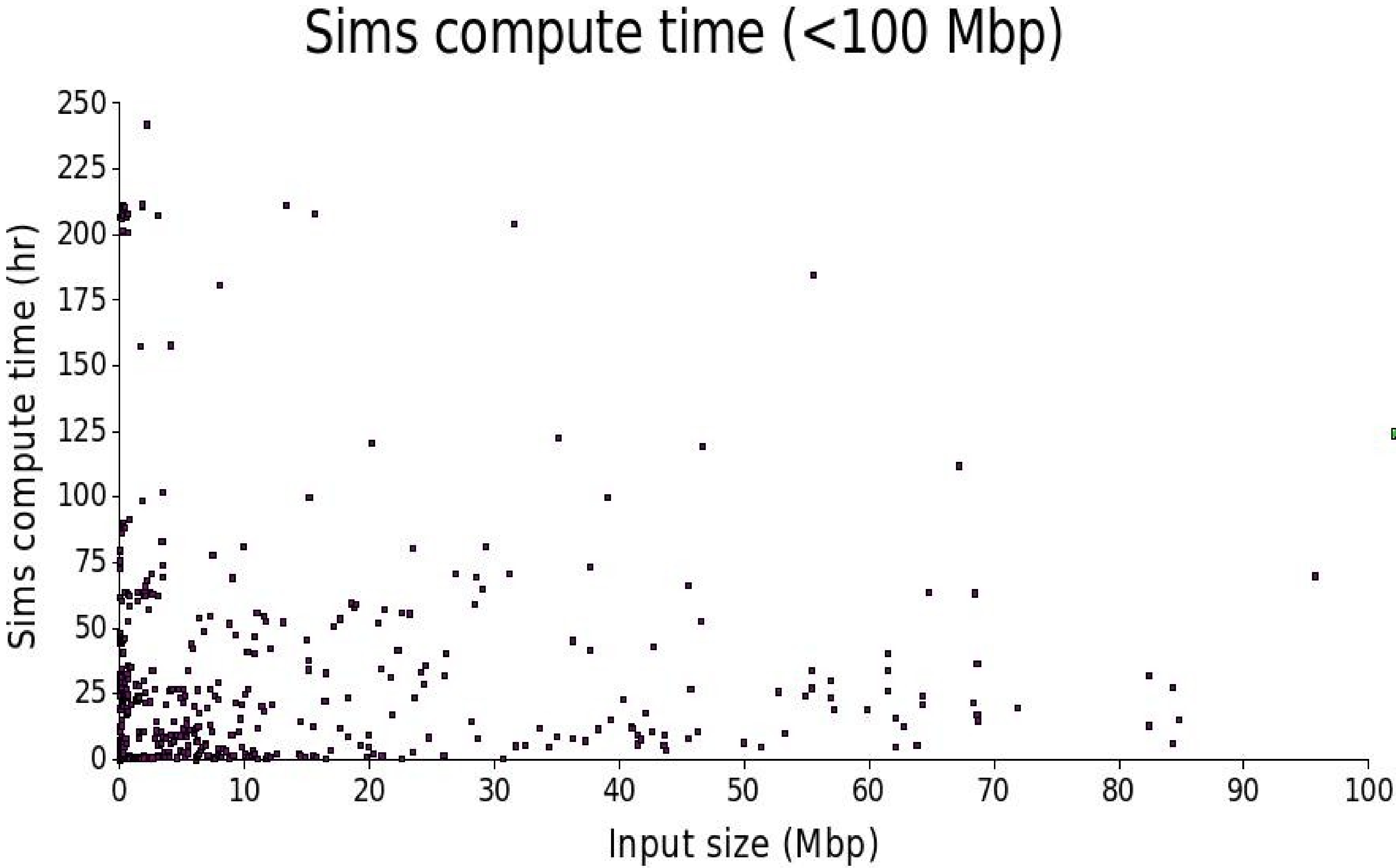
Public:

299 Metagenomes
45,445,163 sequences
19,341,509,132 bp (19 Gbp)

Compute time (on a single CPU):

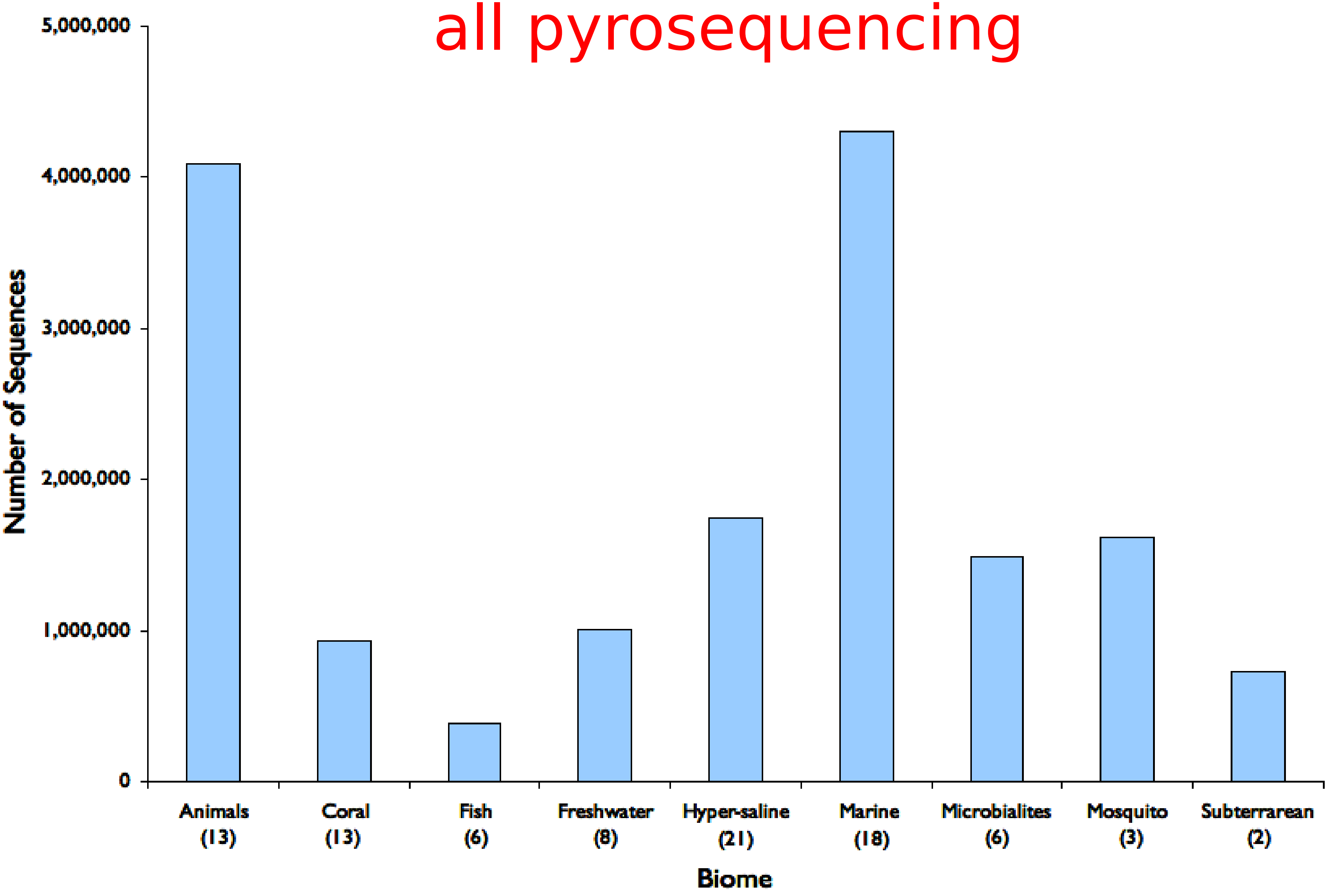
1,292,000 hours = 53,833 days = 147 years

Lots of computers, no pattern

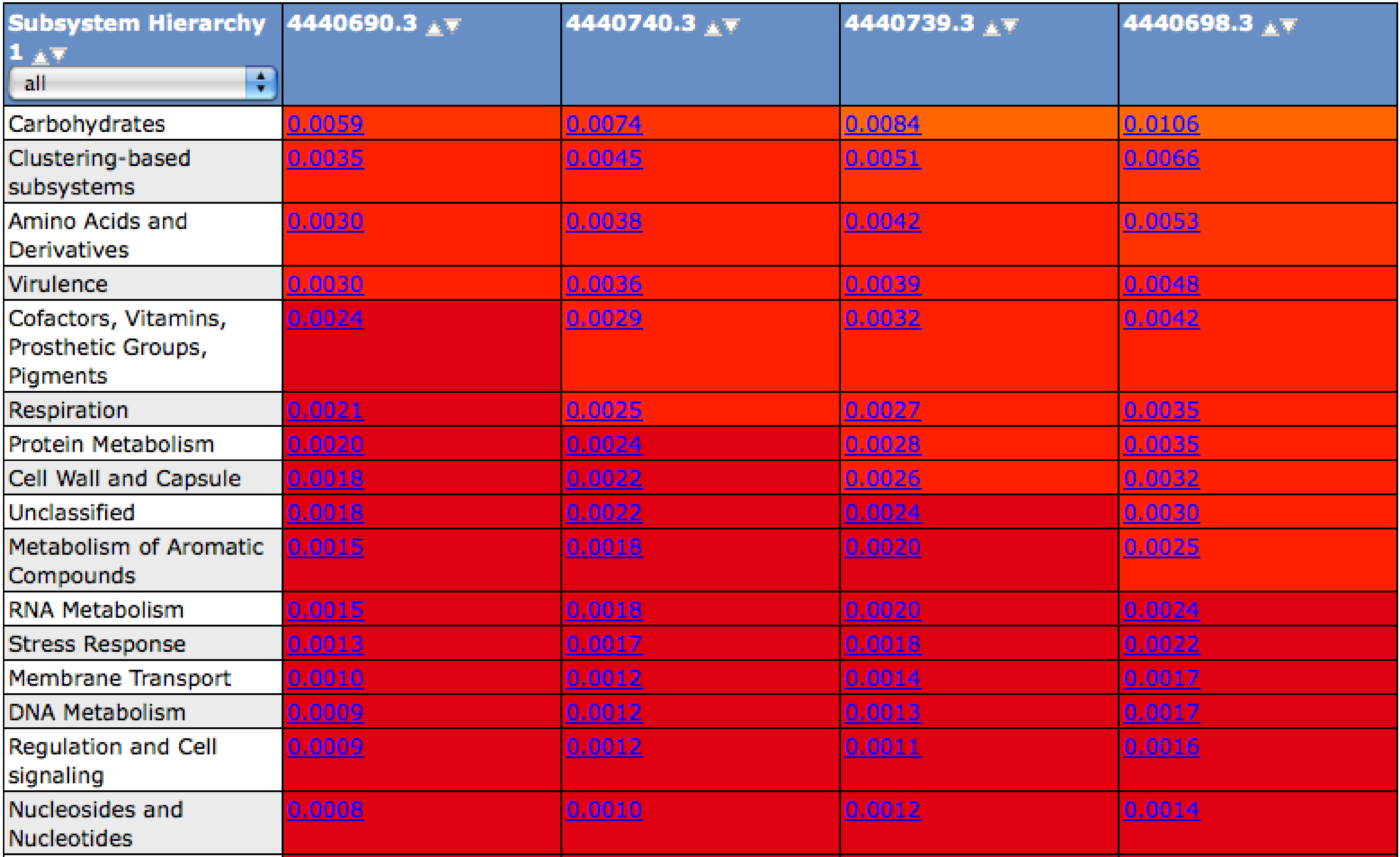


Does it work?

Lots of sequences
all pyrosequencing

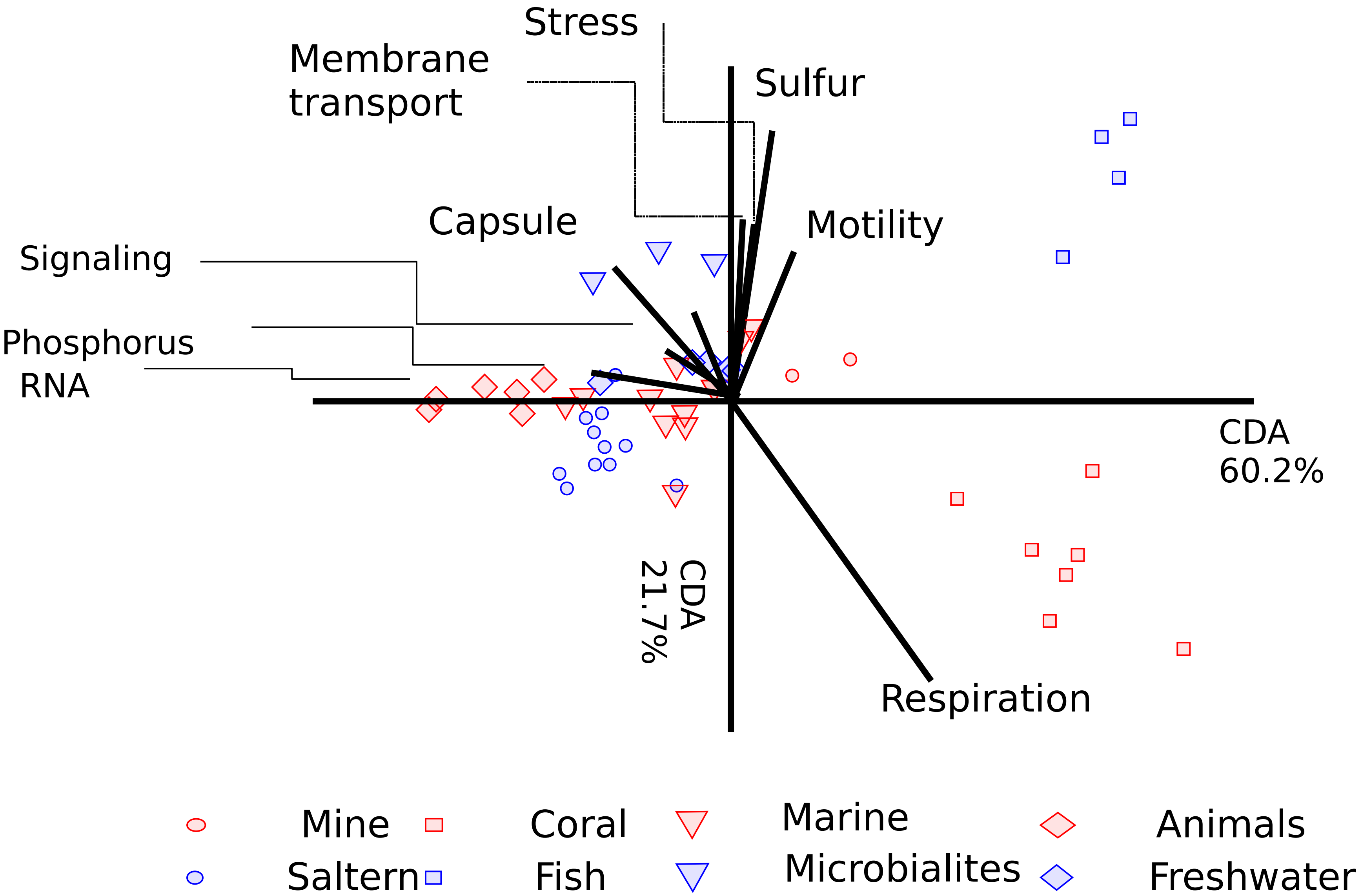


Metagenomics Tools
Functional Heat Maps



Subsystem Hierarchy 1 ▲▼ all	4440690.3 ▲▼	4440740.3 ▲▼	4440739.3 ▲▼	4440698.3 ▲▼
Carbohydrates	0.0059	0.0074	0.0084	0.0106
Clustering-based subsystems	0.0035	0.0045	0.0051	0.0066
Amino Acids and Derivatives	0.0030	0.0038	0.0042	0.0053
Virulence	0.0030	0.0036	0.0039	0.0048
Cofactors, Vitamins, Prosthetic Groups, Pigments	0.0024	0.0029	0.0032	0.0042
Respiration	0.0021	0.0025	0.0027	0.0035
Protein Metabolism	0.0020	0.0024	0.0028	0.0035
Cell Wall and Capsule	0.0018	0.0022	0.0026	0.0032
Unclassified	0.0018	0.0022	0.0024	0.0030
Metabolism of Aromatic Compounds	0.0015	0.0018	0.0020	0.0025
RNA Metabolism	0.0015	0.0018	0.0020	0.0024
Stress Response	0.0013	0.0017	0.0018	0.0022
Membrane Transport	0.0010	0.0012	0.0014	0.0017
DNA Metabolism	0.0009	0.0012	0.0013	0.0017
Regulation and Cell signaling	0.0009	0.0012	0.0011	0.0016
Nucleosides and Nucleotides	0.0008	0.0010	0.0012	0.0014

From Sequences To Environments



Dinsdale et al, Nature 2008

BACK!