

What's New in IMG 3.3

The screenshot shows the homepage of the Integrated Microbial Genomes (IMG) system. At the top left is the JGI logo (DOE Joint Genome Institute, US Department of Energy, Office of Science). To the right is a search bar with a 'GO' button. Below the logo is the 'img' logo and the text 'INTEGRATED MICROBIAL GENOMES'. A navigation menu includes 'IMG Home', 'Find Genomes', 'Find Genes', 'Find Functions', 'Compare Genomes', 'Analysis Cart', and 'My IMG'. On the left side, there is a table titled 'IMG Genomes' with columns for 'finished/draft/perm draft' and 'Total'. The table lists data for Bacteria, Archaea, Eukarya, Plasmids, and Viruses, along with an 'All Genomes' row. Below the table are links for 'Genome by Metadata', 'IMG Statistics', 'Project Map', and 'Content History'. The main content area features a large blue box with the text 'IMG 3.3: What's New' and a paragraph describing the system's purpose. To the right of this text is a vertical stack of three images: a circular micrograph, a cluster of pinkish cells, and a blue micrograph.

Quick Genome Search: **GO**

img INTEGRATED MICROBIAL GENOMES

IMG Home Find Genomes Find Genes Find Functions Compare Genomes Analysis Cart My IMG

	finished/draft/perm draft	Total
Bacteria	1159/1344/4	2507
Archaea	92/8/0	100
Eukarya	19/57/0	76
Plasmids	1155/1/0	1156
Viruses	2584/0/0	2584
All Genomes	5009/1410/4	6423

[Genome by Metadata](#)
[IMG Statistics](#)
[Project Map](#)
[Content History](#)

The Integrated Microbial Genomes (IMG) system ([Nucleic Acids Research, 2010, Vol. 38](#)) serves as a community resource for comparative analysis and annotation of all publicly available genomes from three domains of life in a uniquely integrated...

IMG 3.3: What's New

genome data analysis in IMG usually starts with the definition of an analysis context in terms of selected genomes, functional annotations, and/or genes, followed by the individual or comparative analysis of genomes, functional annotations, or genes.

The content of IMG 3.3 has been updated with new microbial genomes available in RefSeq version 43 (September 05, 2010). IMG 3.3 contains a total of 6,423 genomes consisting of 2,507 bacterial, 100 archaeal, 76 eukaryotic genomes, 2,584 viruses

Table of Contents

IMG 3.3 Content Genomes	3
IMG Statistics	3
IMG 3.3 UI.....	4
New CSS Menu.....	4
New Message Area	4
Updated Question and Comments Form.....	5
Download.....	5
Miscellaneous	6
Protein Expression Studies	7
Find Up/Down Regulated Genes	7
Phylogenetic Distance Tree - Integration	10
Gene Homologs	10
Hierarchical Genome Clustering.....	12
Dot Plot	16
Scaffold Selection	16
New Pathway Data.....	17
Phenotypes	18
Phylogenetic distribution	20
Background.....	20
Phylogenetic Distribution on Microbiome Details Page	20
Phylogenetic Distribution on Compare Genomes Page.....	26
Hide Zeroes in Genome Statistics.....	28
Background.....	28
Procedure	29
Genome Browser	31
Table Configuration	31
Genome Search	33
By Field.....	33
Gene Search	34
Additional Output Columns.....	34
Gene Cart	35
Table Configuration	35
Function Cart	37
Function Search	38
BLAST	39
SEED.....	40
SEED Browser	40
SEED List	42
Transporter Classification.....	43
Transporter Classification Browser	43
Transporter Classification Family List	44
Site Map	45

IMG 3.3 Content Genomes

The content of **IMG 3.3** has been updated with new microbial genomes available in **RefSeq version 43** (September 05, 2010). **IMG 3.3** contains a total of **6,423** genomes consisting of **2,507** bacterial, **100** archaeal, **76** eukaryotic genomes, **2,584** viruses (including bacterial phages), and **1,156** plasmids that did not come from a specific microbial genome sequencing project. Among these genomes, **5,009** are **finished** genomes, and **1,410** are **draft** genomes, and **4** are **permanent draft** (i.e., will never be finished) genomes.

Note that **21** microbial genomes from **IMG 3.2** were **replaced** in **IMG 3.3** because (1) a **draft** genome has been replaced by its **finished** version, or (2) the composition of the genome has changed through the addition of new replicons (plasmids, chromosomes). For replaced genomes, the gene object identifiers (gene OIDs) for the protein-coding genes (CDS) were mapped to their new version in **IMG. 3.3**. See IMG Data Evolution History for details.

IMG 3.3 also contains **proteomic** data from *Arthrobacter chlorophenolicus*1, *Cryptobacterium curtum*, and *Brachybacterium faecium* studies.

IMG Statistics

Various statistics are provided via the **IMG Statistics** link on the home page of IMG, as shown below, including **IMG Total Gene Count** which consists of counting all the genes (protein coding genes, RNA genes) in IMG, except obsolete genes. Compared to **IMG 3.2**, **IMG 3.3** contains **10,151,522 genes**, an increase of **781,122 genes**.

IMG 3.3 UI

New CSS Menu

IMG is now using CSS menu. To display menu you just mouse over text area. See Figure 1

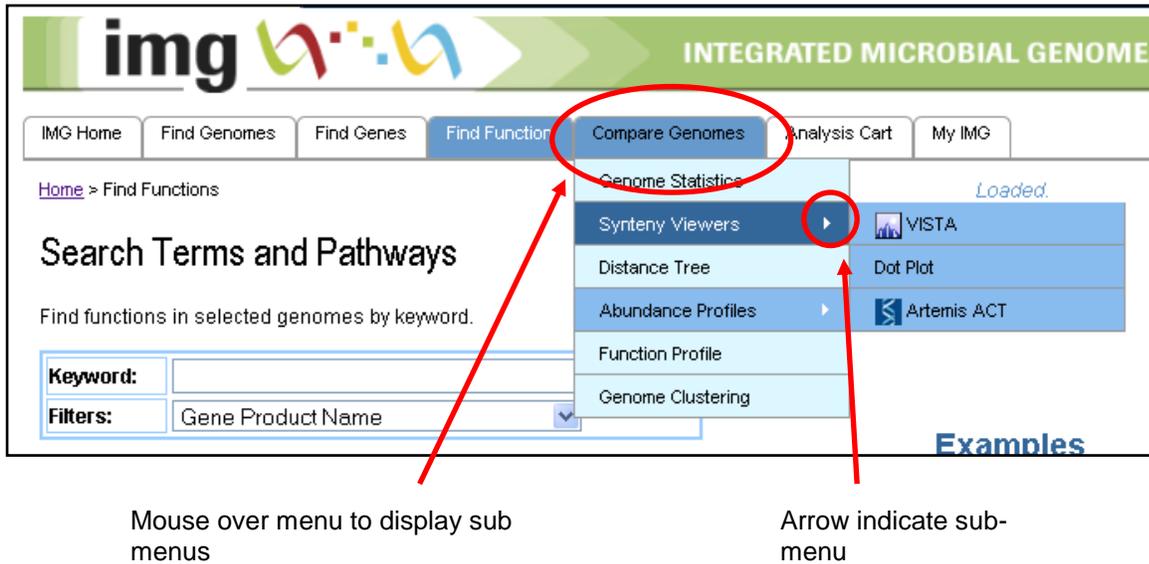


Figure 1 CSS Menu

New Message Area

Message area located below the menu has been updated. See Figure 2. It now has four new features:

1. Navigation breadcrumbs.
2. JavaScript warning message box (when JavaScript is disabled in your browser).
3. New loading message with loading icon and loaded message.
4. Page help indicator image if available.

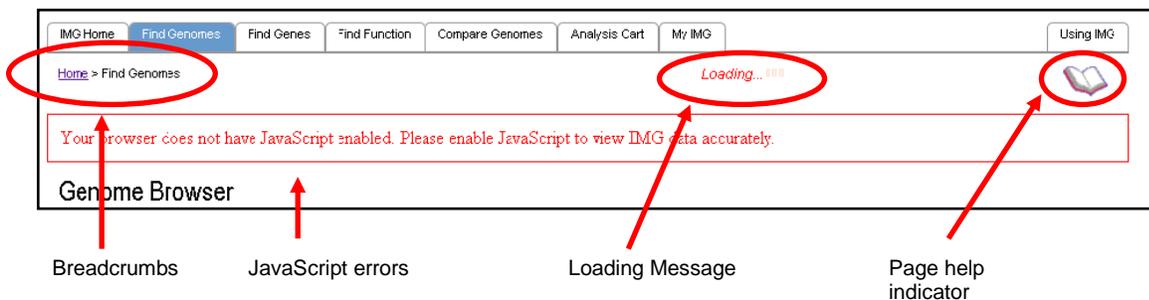


Figure 2 Messages Area

Updated Question and Comments Form

We have updated the Question and Comments Form with Google's reCaptcha as part the submission form to reduce spam and bots. It is a required field. See Figure 3 Question and Comments with Google reCaptcha.

Questions and Comments about IMG

If you request supplementary IMG materials for educational purposes, please provide your full name, affiliation, address, name of course, and course description.

Fields marked with an * are required.

*Name:

*Email:

*Subject:

*Message:

* To prevent spam and abuse, please enter the text shown in the window below. ReCaptcha is required to submit your Questions/Comments.

increased atrago

stop spam. read books.

Submit

Google reCaptcha

Figure 3 Question and Comments with Google reCaptcha

Download

You can now download public genome sequence files used in IMG. Download link is available under menu "Using IMG" (See Figure 4). Currently, it is an ftp site. This is a completely **FREE** service with limited resources, so **PLEASE** be respectful to others and **LIMIT** your ftp connect to **ONE** per session.

What's New in IMG 3.3

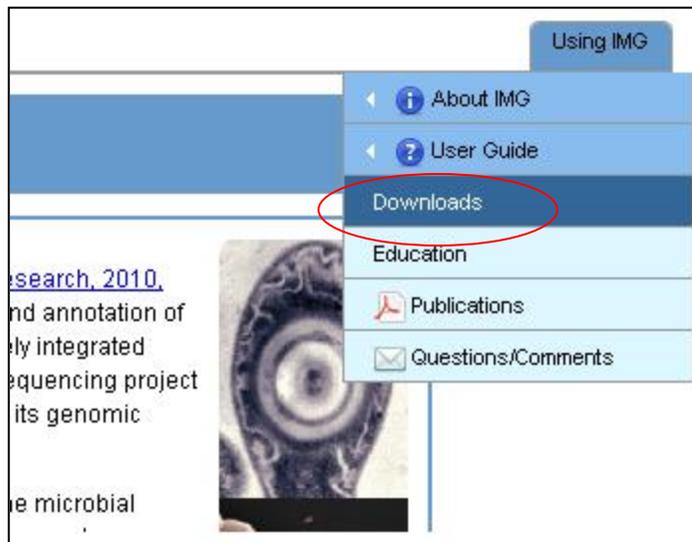


Figure 4 Download menu link



Figure 5 FTP download site.

Miscellaneous

There are several other UI improvements:

1. Increased web page caching.
2. Common BLAST request and results will be cached.
3. Cassette Profiler results cached.
4. Home page and IMG Stats pages will be cached for 24 hours.

Protein Expression Studies

Find Up/Down Regulated Genes

Users can select 2 samples of interest to see which genes are up-regulated and which are down-regulated between the two experimental conditions.

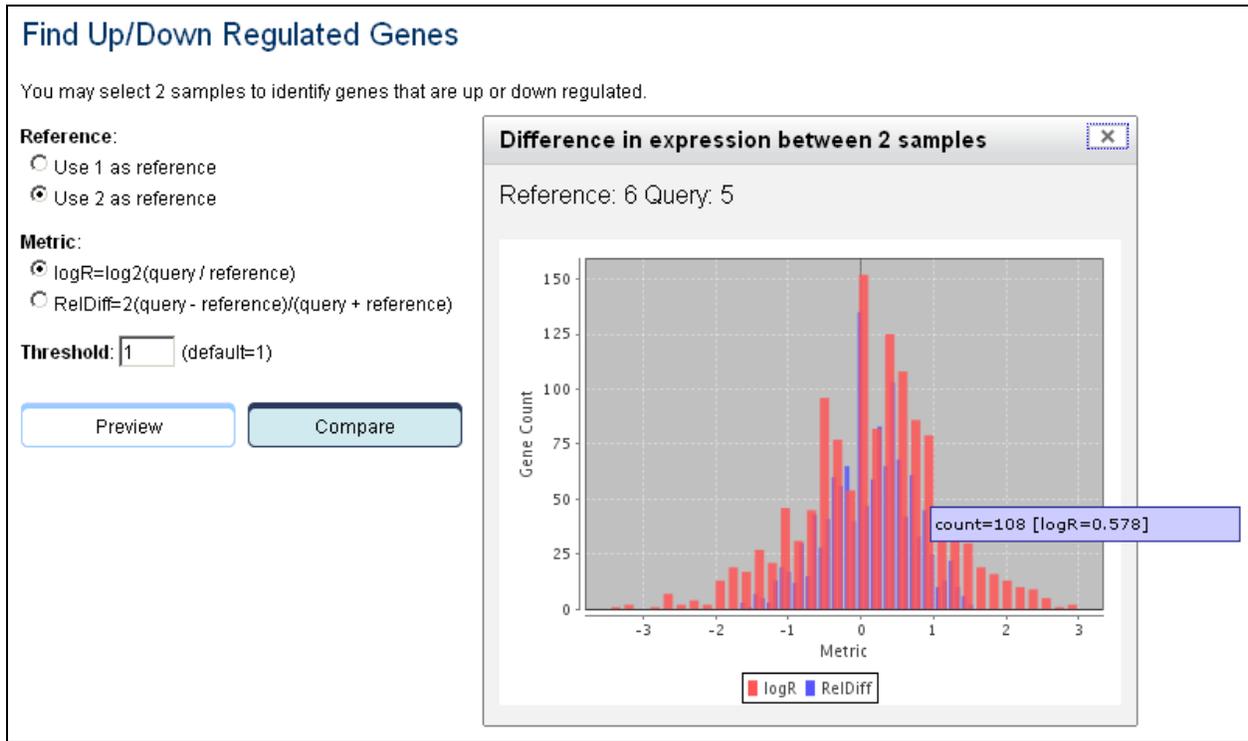


Figure 6 Find Up/Down Regulated Genes - Preview

A user can find all the genes whose expression levels differ by a specified threshold. The difference in expression is computed using either the $\log R = \log_2(\text{query} / \text{reference})$ or the $\text{RelDiff} = 2(\text{query} - \text{reference}) / (\text{query} + \text{reference})$ metric. The user can preview the comparison to see the value spread in a histogram (Figure 6). This helps in choosing which metric and what threshold value to use.

What's New in IMG 3.3

Up/Down Regulation

[Arthrobacter chlorophenolicus A6](#)

Impact of Phenolic Substrate and Growth Temperature on the Arthrobacter chlorophenolicus

Reference sample: [Wt: 100 ppm 4-NP; Growth temp \(C\): 28; 0.7 ug protein; Tube 6; \(Replicate 2\)](#)
 Query sample: [Wt: 100 ppm 4-NP; Growth temp \(C\): 28; 13.2 ug protein; Tube 5; \(Replicate 1\)](#)

hint: Click on a tab to select up- or down- regulated genes to add to gene cart
 Difference in expression levels is computed using the *logR* metric
 Expression levels differ by a *threshold*: 1

Up-regulated Genes
Down-regulated Genes

Add Selections To Gene Cart
Select All
Clear All

Search column: logR Search term:

Export Page 1 of 3 << first < prev 1 2 3 next > last >> 100

Column Selector
Select Page
Deselect Page

Select	Gene ID	Locus Tag	Product Name	6	5	logR
<input type="checkbox"/>	643590150	Achl_1740	transcriptional regulator, GntR family	0.0069444	0.0555556	3.00001
<input type="checkbox"/>	643591875	Achl_3469	FAD dependent oxidoreductase	0.0023952	0.0179641	2.90690
<input type="checkbox"/>	643590353	Achl_1949	Cytochrome b/b6 domain protein	0.0053957	0.0341727	2.66296
<input type="checkbox"/>	643589091	Achl_0694	DNA polymerase III, subunits gamma and tau	0.0016639	0.0099834	2.58496
<input type="checkbox"/>	643589253	Achl_0852	transcriptional regulator, CarD family	0.0125000	0.0750000	2.58496
<input type="checkbox"/>	643590398	Achl_1994	phosphopantothenoylcysteine decarboxylase/phosphopantothenate/cysteine ligase	0.0048193	0.0289157	2.58496

Figure 7 Find Up/Down Regulated Genes

When the user clicks “Compare,” the up- and down- regulated genes are presented in separate tabs (Figure 7). The user can then select which genes to add to the Gene Cart.

Cluster Samples

You may select samples and cluster them based on the coverage of the expressed genes. Proximity of grouping indicates the relative degree of similarity of samples to each other.

Clustering Method:

- Pairwise complete-linkage (default)
- Pairwise single-linkage
- Pairwise centroid-linkage
- Pairwise average-linkage

Distance Measure:

- No gene clustering
- Uncentered correlation
- Pearson correlation (default)
- Uncentered correlation, absolute value
- Pearson correlation, absolute value
- Spearman's rank correlation
- Kendall's tau
- Euclidean distance
- City-block distance

Minimum number of samples in which a gene should appear in order to be included: (default=2)

Use only genes from gene cart

Figure 8 Cluster using only genes from gene cart

A user can now go back, select all samples, and cluster them using only the genes in the Gene Cart (Figure 8). Alternatively, the user can use “Describe” to see the Gene Cart genes in the context of the KEGG maps, chromosomal neighborhood, and cassettes that they belong to.

Describe Samples

You may select samples to compare and describe.

Use only genes from gene cart

Figure 9 Cluster using only genes from gene cart

Phylogenetic Distance Tree - Integration

Gene Homologs

“Phylogenetic Distribution” button on the Gene Homologs page now links to the [Archaeopterix applet](#) (Figure 10), used to display the phylogenetic tree. This tree is colored by the count of homologs (Figure 11).

Gene Homolog

[Customized Homolog Display](#)

Homolog Selection: Top IMG Homolog Hits

Top IMG Homolog Hits

Phylogenetic Distribution

Types (T): O = Ortholog, P = Paralog, H - Putative origin of Horizontal Transfer, - = other unidirectional hit.
 Domains(D): B=Bacteria, A=Archaea, E=Eukarya, P=Plasmids, V=Viruses.
 Genome Completion(C): F=Finished, P=Permanent Draft, D=Draft.

Add query gene [638158342](#)

Add Selections To Gene Cart
Select All
Clear All

Search column: Bit Score Search term:

Export Page 1 of 3 << first < prev **1** [2](#) [3](#) next > last >> 100

Column Selector Select Page Deselect Page

Select	Homolog	T	Percent Identity	Length	E-value	Bit Score	Domain	Status	Genome Name
<input type="checkbox"/>	648021755	O	58.43	524aa	5.0e-166	587.0	A	F	Ferroglobus placidus DSM 10642
<input type="checkbox"/>	638187665	O	53.42	538aa	1.0e-147	517.0	A	F	Aeropyrum pernix K1
<input type="checkbox"/>	640820074	O	46.26	523aa	8.0e-119	430.0	B	F	Nitratiruptor sp. SB155-2
<input type="checkbox"/>	647586885	O	51.19	434aa	1.0e-116	423.0	A	D	Aciduliprofundum boonei T469
<input type="checkbox"/>	647586746	-	50.95	434aa	3.0e-116	421.0	A	D	Aciduliprofundum boonei T469

Figure 10 Gene Homologs link to Archaeopterix applet

Hierarchical Genome Clustering

Genome Clustering now also links to the [Archaeopteryx applet](#) (Figure 12).

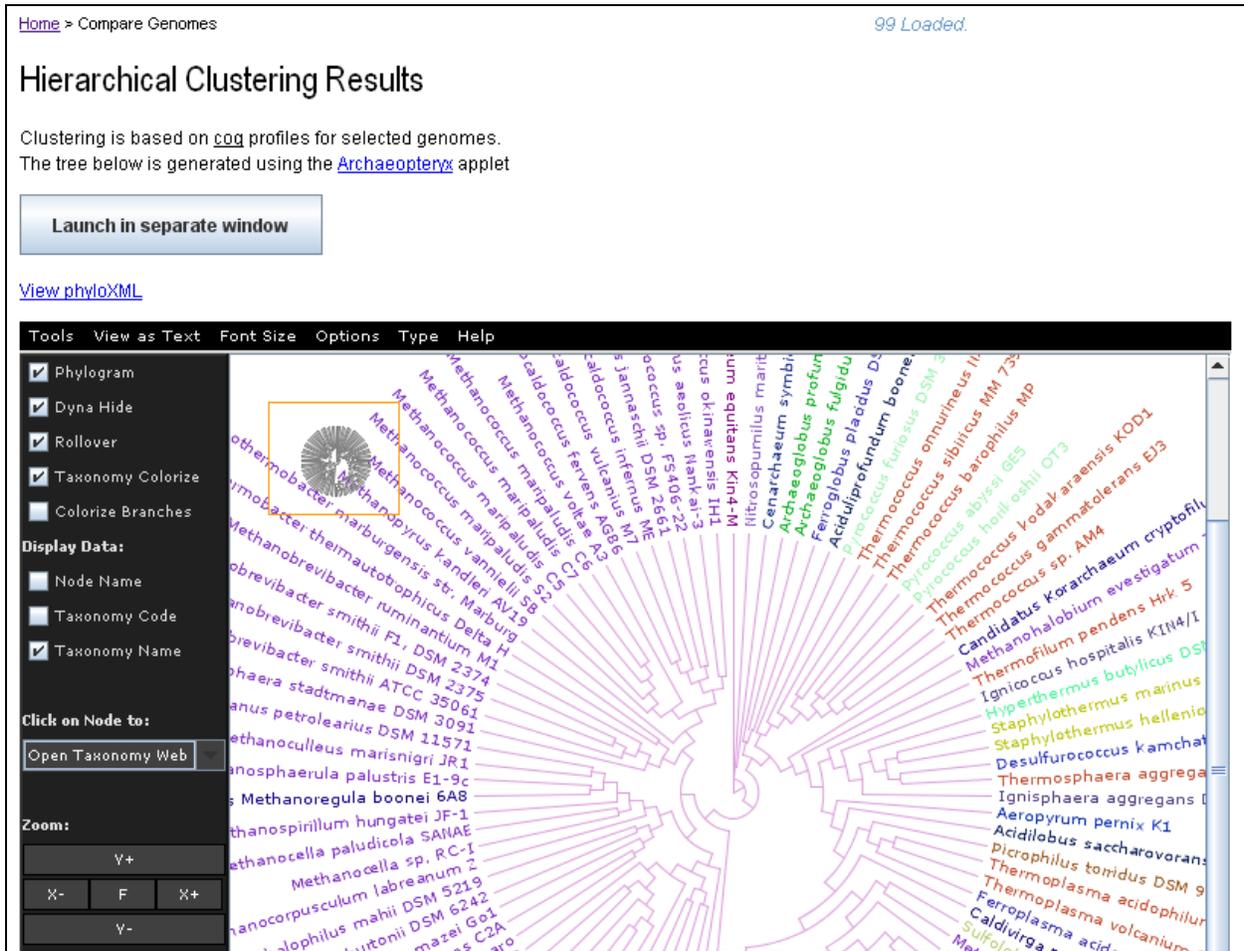


Figure 12 Genome Clustering using the Archaeopteryx applet

ClustalW Alignment

From the Gene Cart, a user can choose to do Sequence Alignment on selected genes. This alignment is displayed using the Jalview applet, which is now embedded (Figure 13). Clicking on the gene in Jalview will open up a Gene Details page in IMG.

ClustalW Alignment of Selected Genes

[Jalview Alignment](#)
[Analyzed Genes](#)
[Phylogenetic Tree](#)

Jalview Alignment

The alignment below is displayed using the [Jalview](#) applet
Click on the id in Jalview to display the gene page in IMG

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V	B	X	Z
Clustal	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
Zappo	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
Taylor	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
Hydrophobicity	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
Helix Propensity	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
Strand Propensity	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
Turn Propensity	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█
Buried Index	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█	█

Figure 13 Jalview Color Scheme for amino acid sequence alignment

What's New in IMG 3.3

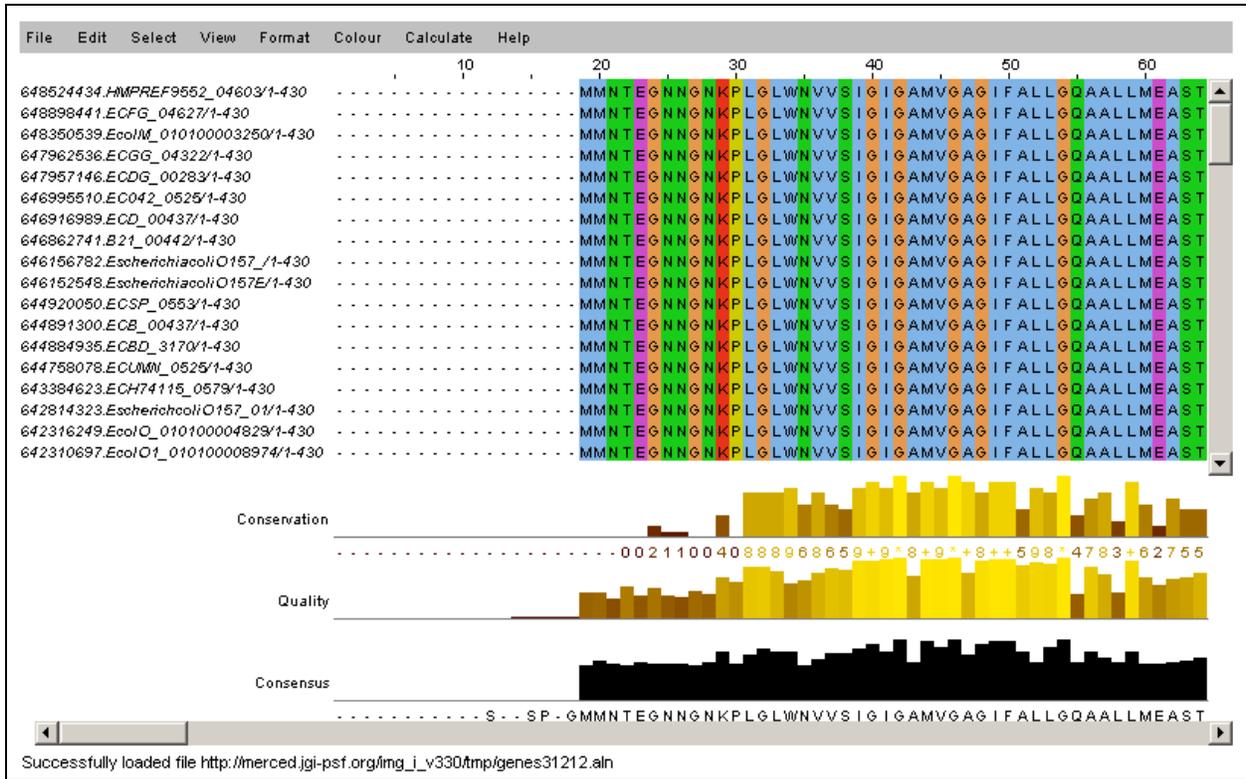


Figure 14 Jalview alignment

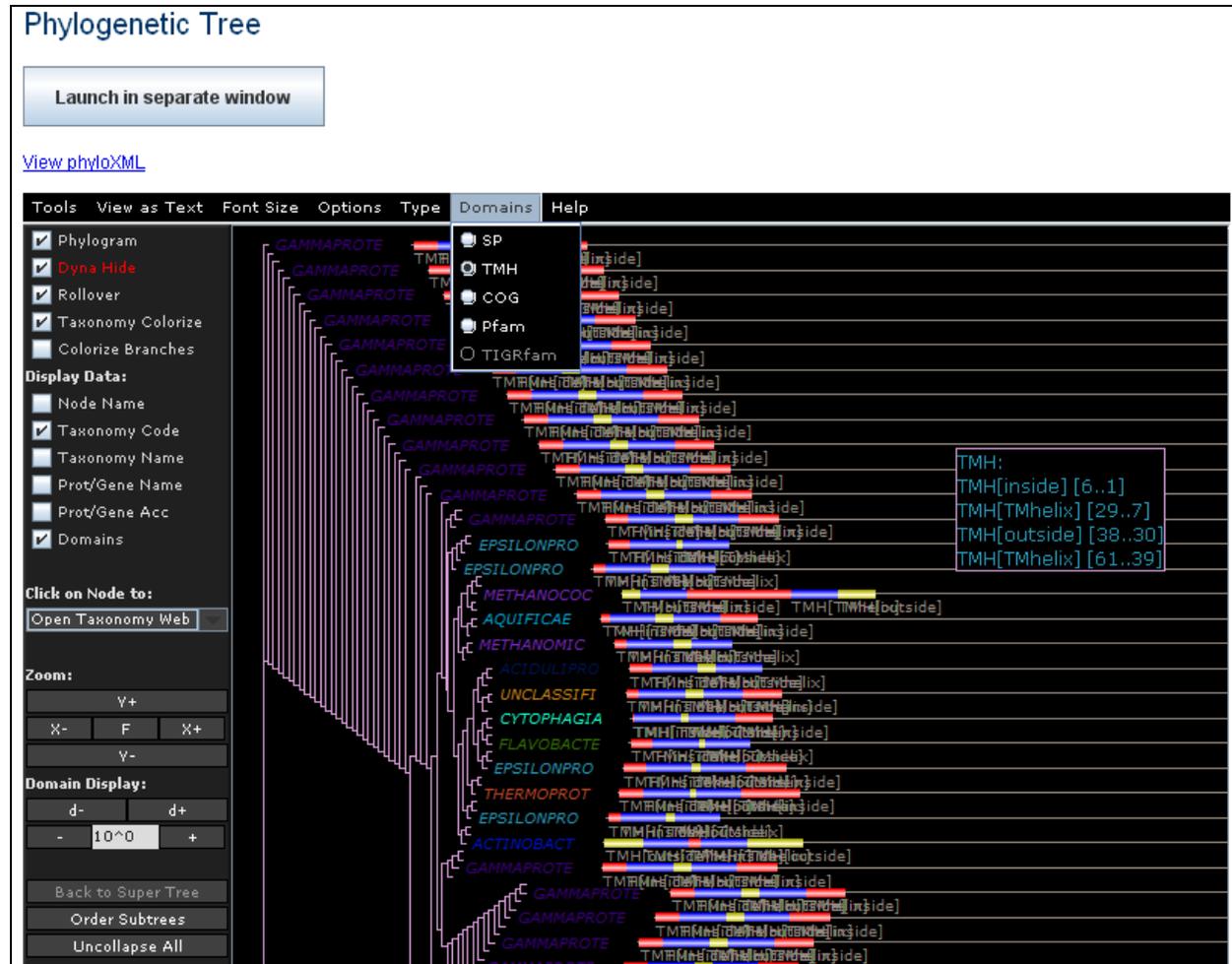


Figure 15 Using the Archaeopterix applet to display domains

Dot Plot

Scaffold Selection

When the number of scaffolds for the 2 genomes selected for Dot Plot is too high, the computation of the Dot Plot may time out. In such a scenario, the user is offered a choice to select a subset of scaffolds or to continue with the calculation using all scaffolds (Figure 16). The current cutoff is when the number of scaffold is >150 for both genomes or >300 for one of the genomes

Home > Compare Genomes > Synteny Viewers Loaded.

Dotplot

The selected genomes are composed of too many scaffolds and plot computation may time-out. You may instead select individual scaffolds for each genome. Otherwise, click **continue** to proceed with the current calculation.

Select Scaffolds

[Select Scaffolds for Bacillus thuringiensis sv berliner ATCC 10792](#)
[Select Scaffolds for Bacillus thuringiensis sv andalousiensis BGSC 4AW1](#)

1. Bacillus thuringiensis sv berliner ATCC 10792

Search column: Scaffold Search term:

Page 1 of 3 << first < prev 1 **2** 3 next > last >> 100

Select	Scaffold	Length (bp)	GC	Type	Topology	No. Genes	Coordinate Range
<input type="checkbox"/>	Bacillus thuringiensis serovar berliner ATCC 10792, unfinished sequence: NZ_ACNF01000001	14600	0.37	chromosome	linear	17	1..14600
<input type="checkbox"/>	Bacillus thuringiensis serovar berliner ATCC 10792, unfinished sequence: NZ_ACNF01000002	48161	0.36	chromosome	linear	51	1..48161
<input type="checkbox"/>	Bacillus thuringiensis serovar berliner ATCC 10792, unfinished sequence:	58203	0.39	chromosome	linear	74	1..58203

<input type="checkbox"/>	Bacillus thuringiensis serovar andalousiensis BGSC 4AW1, unfinished sequence: NZ_ACNG01000099	592	0.35	chromosome	linear		1..592
<input type="checkbox"/>	Bacillus thuringiensis serovar andalousiensis BGSC 4AW1, unfinished sequence: NZ_ACNG01000100	168902	0.37	chromosome	linear	179	1..168902

Page 1 of 3 << first < prev 1 **2** 3 next > last >> 10

Figure 16 Select Scaffolds for Dot Plot

New Pathway Data

In IMG 3.3, new pathway data such as MPW pathways have been added. We added a new UI function to allow users searching through multiple pathway data sources using Enzyme EC numbers and/or free text descriptions. To perform pathway search, select the "All Pathways" menu item under the "Find Functions" tab. A page will show up for users to select the pathway type (IMG Pathways, KEGG Pathways, MetaCyc Pathways and MPW Pathways), EC number, and keyword. Enter search criteria and click the "Search All Pathways" button, a new "Search Pathway Result" page will show up with all pathways in IMG that satisfy the search condition. Click on any Pathway OID to view the detail (see Figure 17).

The screenshot displays the 'Search All Pathways' interface. On the left, a summary table shows the number of pathways in each category: IMG Pathways (780), KEGG Pathways (361), MetaCyc Pathways (1786), and MPW Pathways (6568). Below this is a search form with checkboxes for 'IMG Pathways', 'KEGG Pathways', and 'MetaCyc Pathways'. The 'Enzyme EC' field is empty, and the 'Keyword' field is also empty. A red circle highlights the 'Search All Pathways' button. An inset window titled 'Search Pathway Result' shows a table of pathways. The table has columns for 'Selected', 'Type', 'Pathway OID', and 'Pathway Name'. One pathway, '1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100', is circled in red. A second inset window titled 'MPW Pathway Details' shows the details for this pathway. It includes a 'Reactions' table with columns for 'Reaction Order', 'Reaction Id', 'Condition', 'Gene Group', 'Protein Group', 'Enzyme', 'Gene Count', and 'Genome Count'. The table shows one reaction: '1, 1.1.1.1. Alcohol dehydrogenase class 4 member domain (EC 1.1.1.1)'. The 'Enzyme' column contains 'EC:1.1.1.1'. Below the table is an 'MPW Pathway Image' showing a chemical reaction: Ethanol + NAD+ -> Acetaldehyde + NADH + H+.

Figure 17 Pathway Data

Phenotypes

When we study any organism, one of the major goals is to understand its physiology and how that relates to its adaptation to its environment. When we sequence the organism's genome, we gain further insight into its physiology. Numerous predictions can be made, which can then be tested in the laboratory. For example, many of the pathways for synthesis of essential metabolites are known, and pathways for utilization of many nutrients have been determined. Many physiological functions require several gene products, and they can be grouped into pathways, where genes function in a specific order, or parts lists in which order is not important. Pathways also can only be understood in the context of other pathways within the organism, and this is why we have developed IMG Phenotypes. For example, if an organism degrades cellulose to cellobiose outside the cell, it can only utilize cellulose as a carbon source if it also has a transport pathway for uptake of cellobiose and, within the cell, a metabolic pathway to gain energy from cellobiose. If all three steps are present, then the organism will have the phenotype of cellulose utilization with cellobiose as an intermediate. In some cases the presence or absence of only one pathway is required for a phenotype. There are also cases in which there are multiple possibilities and require multiple combinations of pathways.

The "Phenotypes" menu item under "Find Functions" in IMG 3.3 list all IMG predicated phenotypes and their associated genomes. Users can click on any rule ID to view the phenotype prediction rule. For example, " L-histidine prototroph" is defined by the presence of IMG Pathway 162: L-histidine synthesis.

Click on the number in the "No. of Genomes w/ Phenotype" column to see a phylogenetic tree display of all genomes that are predicted to have this phenotype, or click the small table icon in this column to have the result displayed in a table format. Users can view phenotype prediction inference detail by click on the question mark (?) in the tree display (Figure 18).

What's New in IMG 3.3

Predicted Phenotypes

Phenotypes are broadly defined as an observable characteristic of an organism. The current list of phenotypes in IMG are predicted using a set of rules based on IMG's native collection of pathways and parts lists. Follow the link provided by the **Rule ID** to see the definitions for the prediction rules. Follow the link provided by **No. of Genomes w/ Phenotype** to the list of genomes associated with a specific phenotype. Click the number to view result in tree display, or click to view result in table display.

Search column: Rule ID Search term:

Export Page 1 of 1 << first < prev 1 next > last >> 50

Column Selector

Rule ID	Name	Category	Category Value	Description	No. of Genomes w/ Phenotype
00001	L-histidine prototroph	Metabolism	Prototrophic	Organism is predicted to be able to synthesize L-histidine.	48
00002	Aerobe	Oxygen Requirement	Aerobe	Organism is predicted to be able to grow in the presence of air.	233
00003	L-lysine prototroph	Metabolism	Prototrophic	Organism is predicted to be able to synthesize L-lysine.	318
00004	Denitrifier	Metabolism	Denitrifying	Organism is predicted to reduce nitrate to nitrogen	
00005	Use of nitrate as electron acceptor	Metabolism	Nitrate reducer	Organism is predicted to grow anaerobically with electron acceptor	

Phenotype Rule Information

Rule ID	00001
Name	L-histidine prototroph
Category	Metabolism
Category Value	Prototrophic
Description	Organism is predicted to be able to synthesize L-histidine.
Add Date	2009-04-08
Last Mod Date	2009-11-20
Modified By	IAIN (IJAnderson@tbl.gov)

Rule

([IPWAY:00162](#): L-histidine synthesis)

Genomes with Selected Phenotype

Phenotype Rule (00001): L-histidine prototroph

Domains(D): B=Bacteria, A=Archaea, E=Eukarya.

[Display genomes with selected phenotype in complete tree.](#)

hint: Selections do not take effect until you save them. Click on (?) to see rule prediction detail.

Genome Completion: (F)inished, (P)ermanent Draft, (D)raft

- 01 Bacteria
 - 02 Firmicutes
 - 03 Bacilli
 - 04 Bacillales
 - 05 Bacillaceae
 - 06 Bacillus
 - 08 [Bacillus amyloliquefaciens FZB42](#) [F] -- L-histidine prototroph (?)
 - 08 [Bacillus clausii KSM-K16](#) [F] -- L-histidine prototroph (?)
 - 08 [Bacillus halodurans C-125](#) [F] -- L-histidine prototroph (?)
 - 08 [Bacillus licheniformis ATCC 14580 \(Goettingen\)](#) [F] -- L-histidine prototroph (?)
 - 08 [Bacillus licheniformis ATCC 14580 \(Novozymes\)](#) [F] -- L-histidine prototroph (?)
 - 08 [Bacillus pumilus ATCC 7061](#) [D] -- L-histidine prototroph (?)
 - 08 [Bacillus pumilus SAER-032](#) [F] -- L-histidine prototroph (?)
 - 08 [Bacillus subtilis NCIB 3610](#) [D] -- L-histidine prototroph (?)
 - 08 [Bacillus subtilis subsp. spizizenii ATCC 6633](#) [D] -- L-histidine prototroph (?)
 - 08 [Bacillus subtilis subsp. subtilis str. 168](#) [F] -- L-histidine prototroph (?)
 - 08 [Bacillus subtilis subtilis 168](#) [D] -- L-histidine prototroph (?)

Figure 18 Phenotypes

Phylogenetic distribution

Background

Historically, IMG has had two implementations of PhyloDistribution tools: one accessible from Microbiome Details page and the other accessible through Compare Genomes > Phylogenetic Distribution. In IMG 3.3, the configuration options of both tools have been unified to a vast degree while maintaining a certain degree of individuality.

Phylogenetic Distribution on Microbiome Details Page

Please refer to Figure 19 & Figure 20 for the location of this tool.

What's New in IMG 3.3

The Integrated Microbial Genomes (IMG) system (Vol. 38) serves as a community resource for all publicly available genomes from three domains of life: Bacteria, Archaea, and Eukarya.

Genome Browser

hint: Selections do not take effect until you save them. You must select at least one genome. Go to [Preferences](#) to show or hide plasmids and viruses. Go to home page statistics under [IMG Genomes](#) to select individual phylogenetic domains or all genomes.

Domains(D): B=Bacteria, A=Archaea, E=Eukarya, P=Plasmids, V=Viruses.
Genome Completion(C): F=Finished, P=Permanent Draft, D=Draft.

Click on column name to sort.

Save Selections Select All Clear All View Phylogenetically

2683 of 2683 rows selected

Search column: Domain Search term:

Export Page 1 of 27 << first < prev 1 2 3 4 5 6 7 8 9 10 next > last >> 100

Select	Domain	Status	Genome Name	Sequencing Center	Gene Count	Genome Size
<input checked="" type="checkbox"/>	A	F	Acidilobus saccharovorans 345-15	Genentech Bioengineering	1547	1496453
<input checked="" type="checkbox"/>	A	D	Aciduliprofundum boonei T469	J. Craig Venter Institute, Portland State Univ	3043	2981805
<input checked="" type="checkbox"/>	A	F	Aciduliprofundum boonei T469	DOE Joint Genome Institute	1587	1486778

Figure 19: IMG Home > Genome Browser

What's New in IMG 3.3

The image shows a screenshot of the IMG 3.3 web interface. The top navigation bar includes links for 'IMG Home', 'Find Genomes', 'Find Geneses', 'Find Functions', and 'Compare Ge...'. The main content area is divided into two sections. The left section, titled 'Organism Details', contains a list of links: 'Organism Information', 'Genome Statistics', 'Phylogenetic Distribution of Genes', 'Putative Horizontally Transferred Genes', 'Genome Viewers', 'Compare Gene Annotations', 'Download Gene Information', and 'Export Genome Data'. The 'Phylogenetic Distribution of Genes' link is highlighted with a red box. The right section, titled 'Phylogenetic Distribution of Genes', contains three buttons: 'Distribution by BLAST percent identities', 'Putative Horizontally Transferred', and 'Genome Viewers'. The 'Distribution by BLAST percent identities' button is also highlighted with a red box. A red arrow points from the 'Phylogenetic Distribution of Genes' link in the left section to the 'Distribution by BLAST percent identities' button in the right section. Below this, a larger screenshot shows the full interface for the 'Phylogenetic Distribution of Genes' tool. The header includes the JGI logo and 'img INTEGRATED MICROBIAL GENOMES'. The navigation bar includes 'IMG Home', 'Find Genomes', 'Find Geneses', 'Find Functions', 'Compare Genomes', 'Analysis Cart', and 'My IMG'. The main content area is titled 'Phylogenetic Distribution of Genes' and 'Acidilobus saccharovorans 345-15'. It includes a 'Quick Genome Search' box with 'ALL Genomes' and a 'GO' button. The tool configuration options are as follows:

- Percent Identity:**
 - Successive (30% to 59%, 60% to 89%, 90%+)
 - Cumulative (30%+, 60%+, 90%+)
 - Display hit genome count (slow)
- Gene count:**
 - Gene count
 - Estimated gene copies
- Display Options:**
 - Show percentage column.
 - Show histogram column.

At the bottom, there are 'Go' and 'Reset' buttons.

Figure 20: Organism Details >> Phylogenetic Distribution Configuration Options

What's New in IMG 3.3

Previous versions of this tool were non-configurable. In IMG 3.3, the UI has been extended to provide configuration options as seen in Figure 20. Selecting the required options and clicking the “Go” button to display the final table. The configuration options available in this tool are:

1. Percent identity
2. Gene count and Estimated gene copies
3. Display Options

Percent Identity: This configuration section provides the following mutually exclusive configurable options.

- a) Successive (30% to 59%, 60% to 89%, 90%+)
- b) Cumulative (30%+, 60%+, 90%+)
 - i) Display hit genome count (slow)

Selecting the “Successive” radio button results in the BLAST hits being successively displayed in 3 separate sets of columns – hits 30% - 59%, hits 60% - 89%, and hits 90% and above (see Figure 21). Relevant column headings provide explanatory tooltips when mouse over.

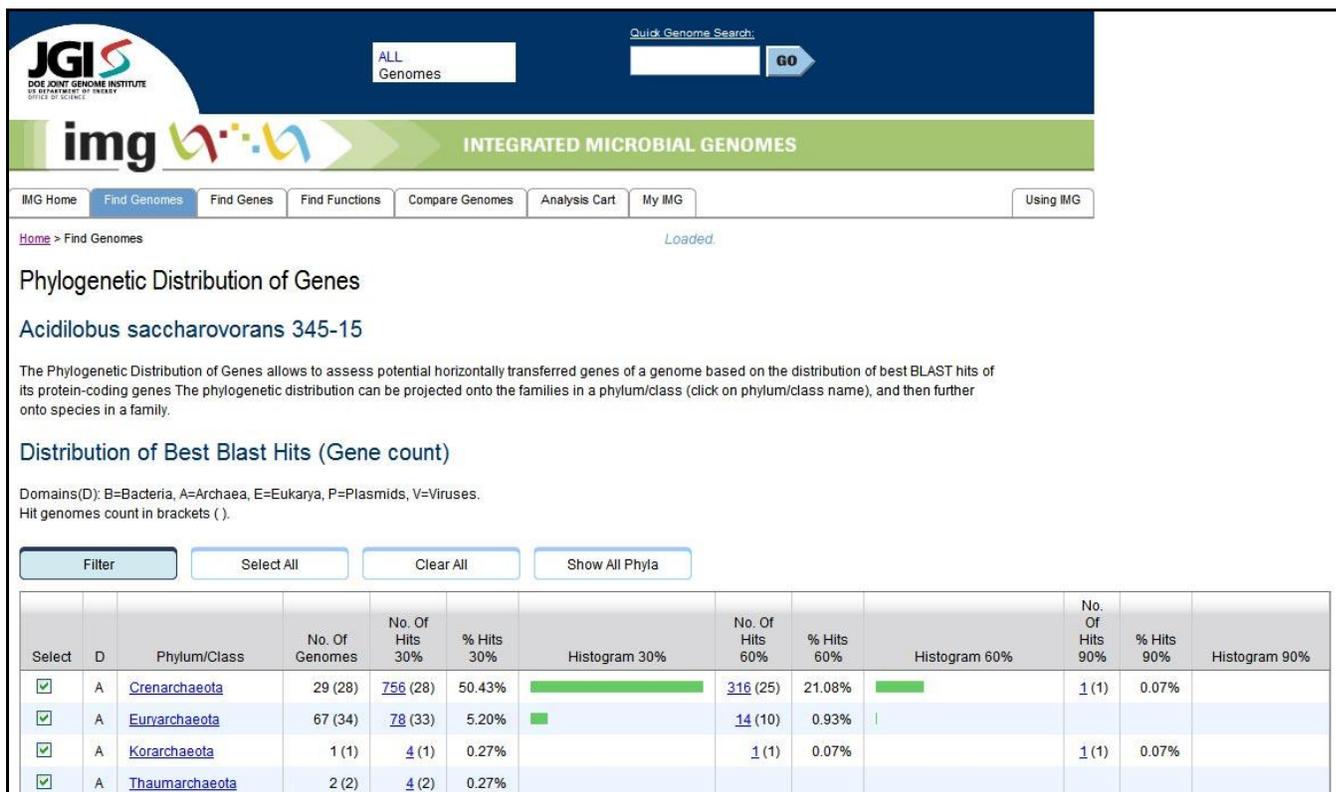


Figure 21: Phylogenetic Distribution - Successive

What's New in IMG 3.3

Selecting the “Cumulative” radio button results in the BLAST hits again being displayed cumulatively in 3 sets of columns showing hit for 30% and above, 60% and above, and 90% and above. The column headings are identified by a “+” sign after the percent symbol as in 30%+, 60%+, and 90%+ to distinguish between the Successive and Cumulative results (see Figure 22).

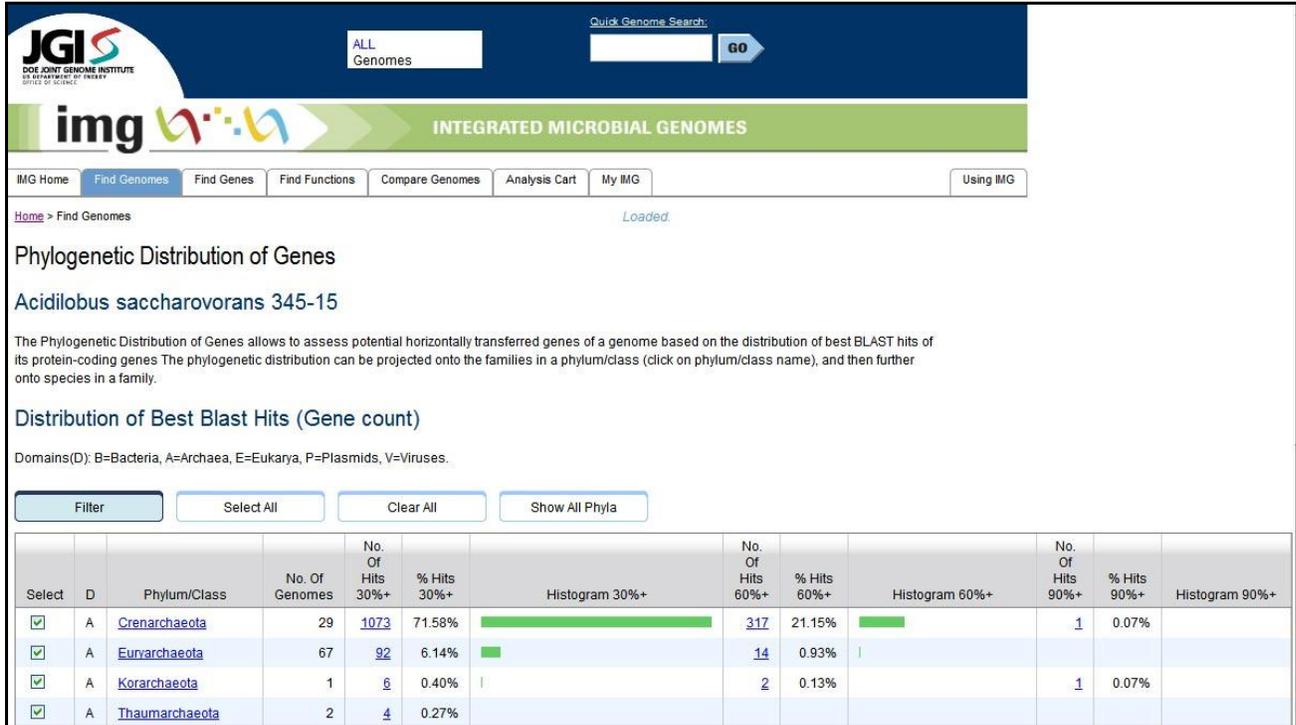


Figure 22: Phylogenetic Distribution – Cumulative (without hit genome count)

By default, the Cumulative option does not display the hit genome count in parenthesis -- unlike the Successive option due to performance considerations. However, if viewing the hit genome count is essential in the cumulative mode, the checkbox “**Display hit genome count (slow)**” can be checked. The result table now shows hit genome count in parenthesis in the relevant columns (see Figure 23). It should be noted that this option incurs additional computations and will hence take longer to display the result table.

What's New in IMG 3.3

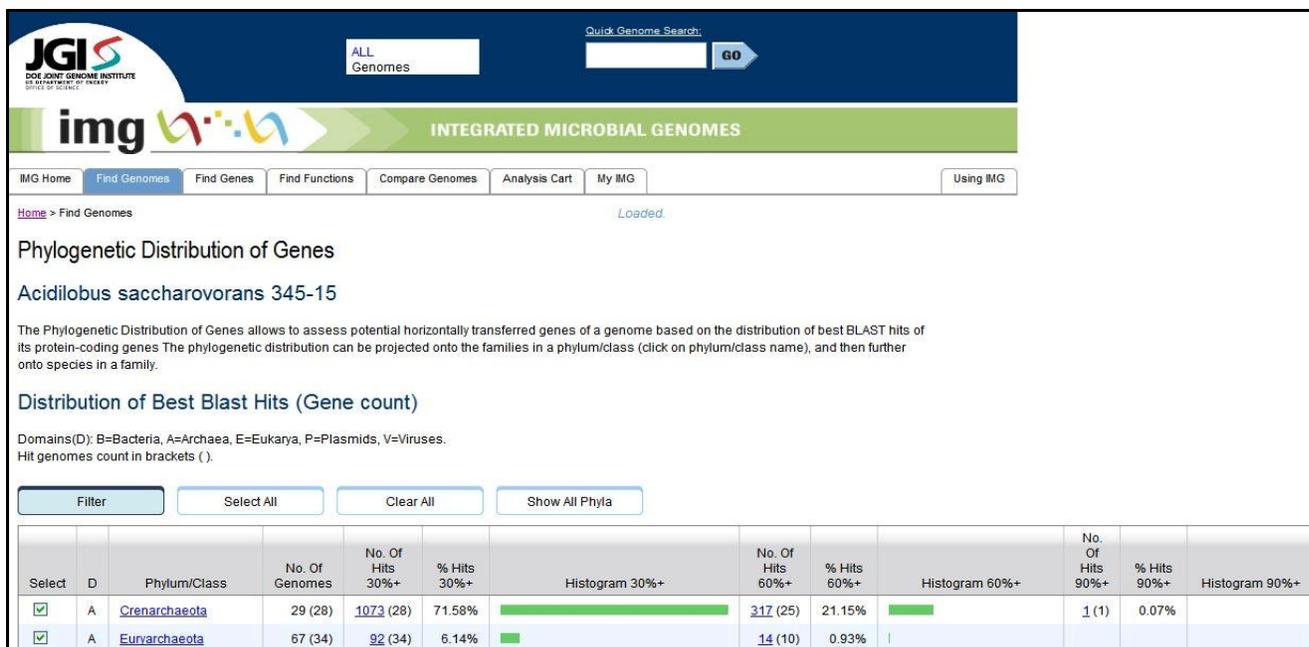


Figure 23: Phylogenetic Distribution – Cumulative (with hit genome count)

Gene count and Estimated gene copies: These options are mutually exclusive. Selecting “**Gene count**” uses the actual gene count to show hits. Selecting “**Estimated gene copies**” uses gene count multiplied by read depth (when such information is available) to show hits. The option selected is identified in the sub heading of the report - “Distribution of Best Blast Hits.”

Display Options: These options are available as checkboxes and can be selected independently of each other. Both options are selected by default. They are:

- Show percentage column.
- Show histogram column.

Selecting or deselecting “**Show percentage column**” displays or hides the percent columns for 30%, 60%, and 90% hits.

Selecting or deselecting “**Show histogram column**” displays or hides the histogram columns for 30%, 60%, and 90% hits.

Phylogenetic Distribution on Compare Genomes Page

Please refer to Figure 24 for the location of this tool. The primary difference between Phylogenetic Distribution on Microbiome Details Page and Phylogenetic Distribution on Compare Genomes Page is that the latter allows for selecting multiple metagenomes (see Figure 24).

Configuration options are similar to Phylogenetic Distribution on Microbiome Details Page. Choosing the desired options, followed by selecting the required metagenomes, and clicking the “Go” button displays the final table. The configurations options available in this tool are:

1. Percent identity
2. Gene count and Estimated gene copies
3. Display Options

Percent Identity: This configuration provides the following mutually exclusive configurable options in the form of a drop down menu:

- 30% - 59%
- 60% - 89%
- 90%+
- 30%+
- 60%+

For an explanation of these options, please refer to “Phylogenetic Distribution on Microbiome Details Page” above.

For an explanation on Gene count and Estimated gene copies as well as Display Options, please refer to “Phylogenetic Distribution on Microbiome Details Page” above.

What's New in IMG 3.3

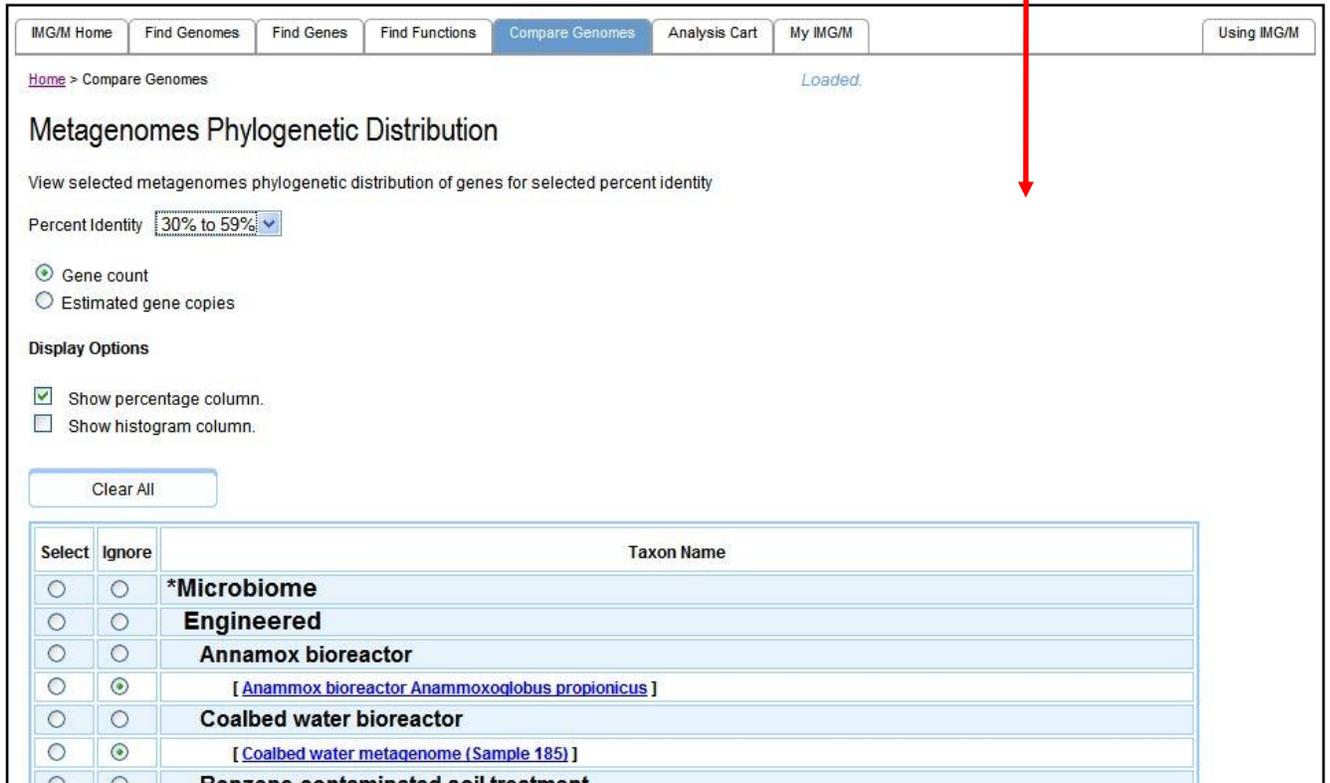
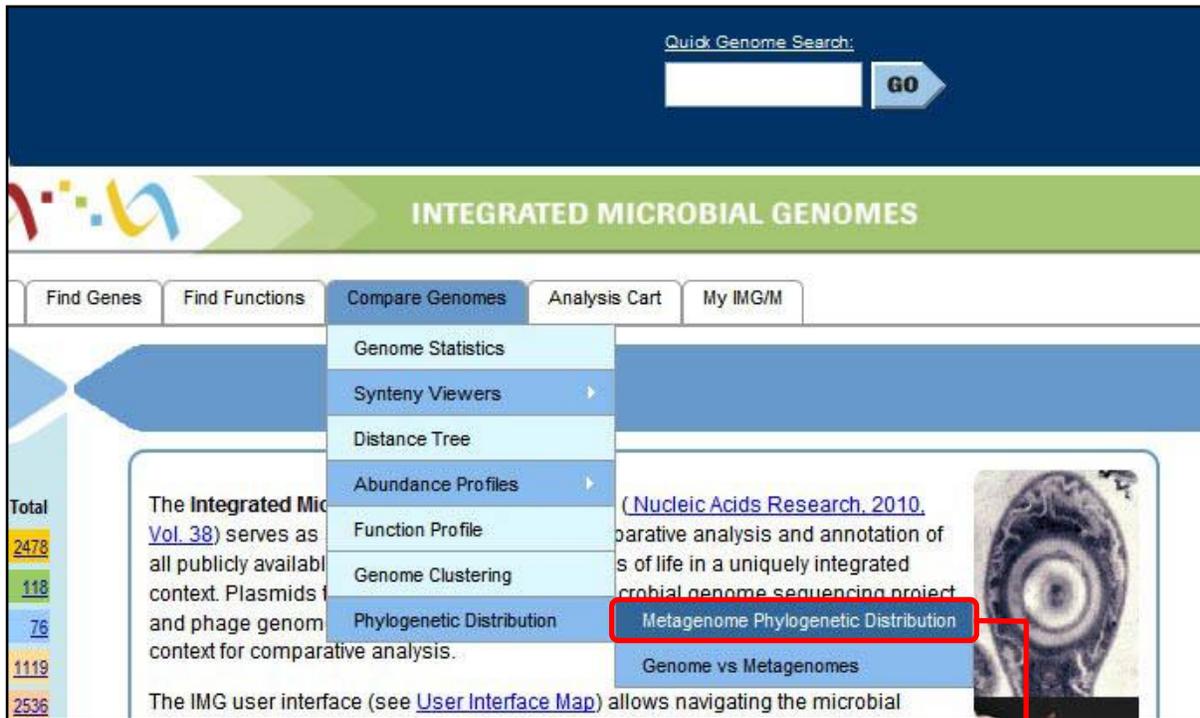


Figure 24: Phylogenetic Distribution on Compare Genomes Page

Hide Zeroes in Genome Statistics

Background

Previously, whenever genome statistics were viewed within IMG, all rows were displayed regardless of whether numbers in those rows were zero or not. If several of the rows were just zeroes in a genome, this would take up valuable screen real estate. The solution was to suppress all rows where the numbers were just zeroes. At the same time, users may want confirmation that particular elements of genome statistics were in fact zero. As a result the IMG user interface has been extended to include a new user preference named “Hide Zeroes in Genome Statistics”. See Figure 25 for the location of this preference.

The screenshot shows the IMG 3.3 user interface. At the top, there is a navigation bar with the JGI logo and a search box. Below the navigation bar, there is a menu with options: IMG Home, Find Genomes, Find Genes, Find Functions, Compare Genomes, Analysis Cart, and My IMG. The 'My IMG' menu is open, showing 'MyIMG Home' and 'Preferences'. The 'Preferences' page is displayed, showing a table of parameters and their current settings. The 'Hide Zeroes in Genome Statistics' parameter is highlighted with a red box, and its current setting is 'Yes'.

Parameter	Current Setting
Max. Paralog Groups	500
Max. Gene List Results	1000
Max. Homolog Results	200
Max. Taxon Gene Neighborhoods	15
Default Gene Page Homolog List	None
Min. Homolog Percent Identity	30
Hide Viruses From Genome Lists	Yes
Hide Plasmids From Genome Lists	Yes
Hide Zeroes in Genome Statistics	Yes

Figure 25: Hide Zeroes in Genome Statistics

By default, “Hide Zeroes in Genome Statistics” is set to “Yes.” This means that display of all rows in the genome statistics table that are zeroes will be suppressed. Changing this setting to “No” reveals all rows regardless of whether their values are zero or not.

Procedure

To see this preference in action, carry out the following steps:

1. From the IMG Home page, click on the “Genome Browser” tab.
2. This will bring up the list of all available genomes.
3. In the resulting table, clicked on any desired genome under the column titled “Genome Name.”
4. This will bring up the Organism/Microbiome Details page.
5. Click on the anchor titled “Genome Statistics.”

Assuming that the “Hide Zeroes in Genome Statistics” parameter is at its default setting of “Yes,” the Genome Statistics table will have no rows with numbers that are zero (see Figure 26).

Genome Statistics

hint: To view rows that are zero, go to [MyIMG preferences](#) and set “Hide Zeroes in Genome Statistics” to “No”.

	Number	% of Total
DNA, total number of bases	1496453	100.00%
DNA coding number of bases	1356593	90.65%
DNA G+C number of bases	856231	57.22% ¹
DNA scaffolds	1	100.00%
CRISPR Count	6	
Genes total number	1547	100.00%
Protein coding genes	1499	96.90%
RNA genes	48	3.10%
rRNA genes	3	0.19%
5S rRNA	1	0.06%
16S rRNA	1	0.06%
23S rRNA	1	0.06%
tRNA genes	45	2.91%
Protein coding genes with function prediction	981	63.41%
Protein coding genes without function prediction	518	518.00%
Genes w/o function with similarity	437	28.25%
Genes w/o function w/o similarity	81	5.24%
not connected to SwissProt Protein Product	1499	96.90%

Figure 26: Genome Statistics with “Hide Zeroes in Genome Statistics” set to “Yes”

What's New in IMG 3.3

Now change the “Hide Zeroes in Genome Statistics” parameter to “No” (see Figure 25). Repeat steps 1 through 5 above (or do a browser Reload/Refresh of the Genome Statistics page). The result table will now reveal all statistics rows that are zero (see Figure 27). Note the additional rows with zeroes in the table. This setting stays in effect until the “Hide Zeroes in Genome Statistics” parameter is manually changed or the browser session is terminated.

Genome Statistics

hint: To view rows that are zero, go to [MyIMG preferences](#) and set **"Hide Zeroes in Genome Statistics"** to **"No"**.

	Number	% of Total
DNA, total number of bases	1496453	100.00%
DNA coding number of bases	1356593	90.65%
DNA G+C number of bases	856231	57.22% ¹
DNA scaffolds	<u>1</u>	100.00%
CRISPR Count	<u>6</u>	
Genes total number	1547	100.00%
Protein coding genes	<u>1499</u>	96.90%
Pseudo Genes	0	0.00% ²
RNA genes	<u>48</u>	3.10%
rRNA genes	<u>3</u>	0.19%
5S rRNA	<u>1</u>	0.06%
16S rRNA	<u>1</u>	0.06%
18S rRNA	0	0.00%
23S rRNA	<u>1</u>	0.06%
28S rRNA	0	0.00%
tRNA genes	<u>45</u>	2.91%
Other RNA genes	0	0.00%
Protein coding genes with function prediction	<u>981</u>	63.41%
Protein coding genes without function prediction	518	518.00%
Genes w/o function with similarity	<u>437</u>	28.25%
Genes w/o function w/o similarity	<u>81</u>	5.24%
Protein coding genes connected to SwissProt Protein Product	0	0.00%
not connected to SwissProt Protein Product	<u>1499</u>	96.90%

Figure 27: Genome Statistics with “Hide Zeroes in Genome Statistics” set to “No”

Genome Browser

Table Configuration

By default, the alphabetical list of genomes includes information on domain, completion status, sequencing center, gene count, and genome size. However, users can configure the displayed columns by using the **Table Configuration** selector at the bottom of the **Gene Browser** page, as shown in Figure 28.

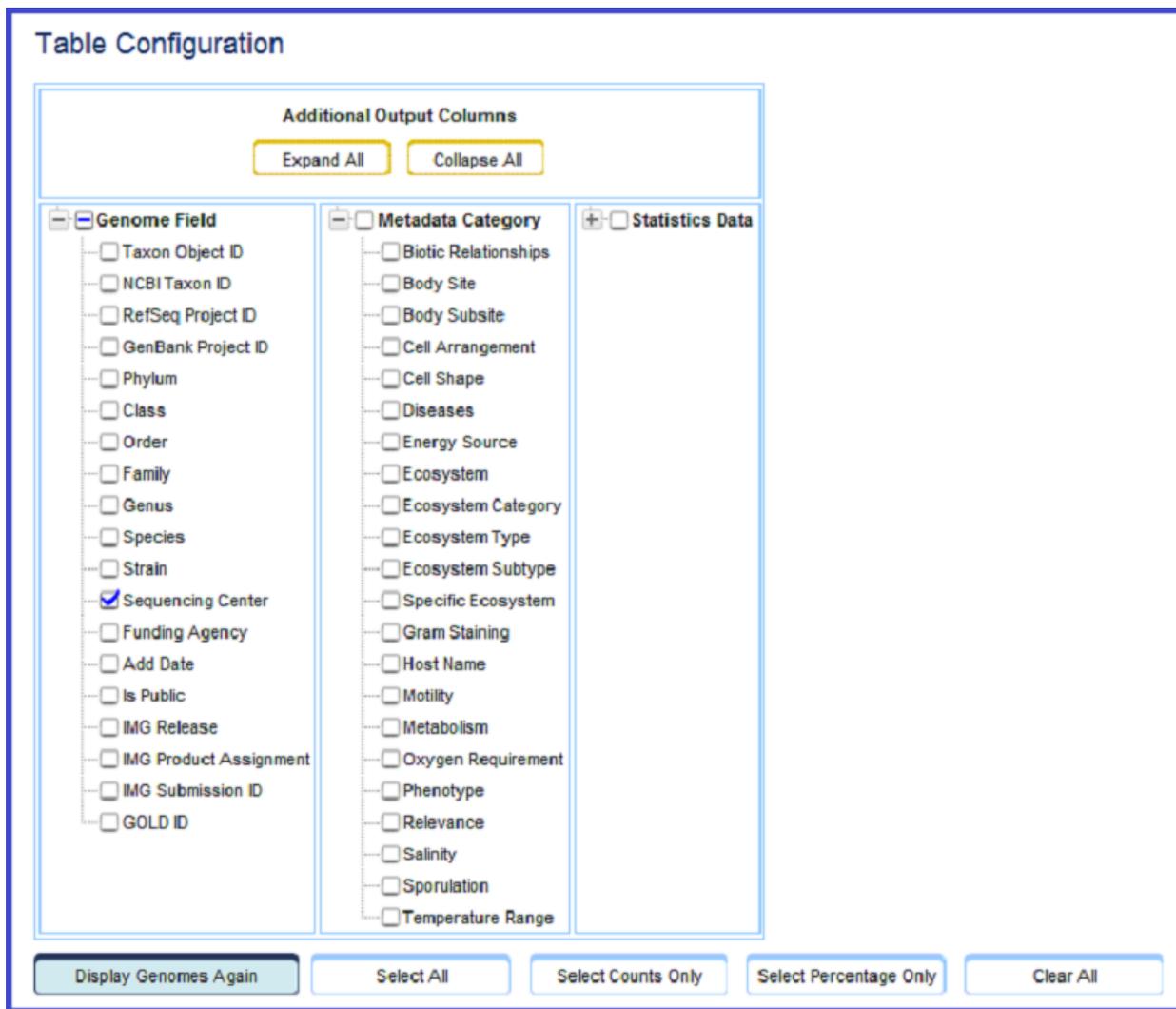


Figure 28: Table Configuration selector.

The **Table Configuration** selector lists three groups of “Additional Output Columns”:

1. Genome Field:
Taxon Object ID; NCBI Taxon ID; RefSeq Project ID; GenBank Project ID; Phylum; Class; Order; Family; Genus; Species; Strain; Sequencing Center; Funding Agency;

What's New in IMG 3.3

Add Date; Is Public; IMG Release; IMG Product Assignment; IMG Submission ID; GOLD ID.

2. Metadata Category:

Biotic Relationships; Body Site; Body Subsite; Cell Arrangement; Cell Shape; Diseases; Energy Source; Ecosystem; Ecosystem Category; Ecosystem Type; Ecosystem Subtype; Specific Ecosystem; Gram Staining; Host Name; Motility; Metabolism; Oxygen Requirement; Phenotype; Relevance; Salinity; Sporulation; Temperature Range.

3. Statistics Data:

Scaffold Count; CRISPR Count; GC Count; GC %; Coding Base Count; Genome Size; Gene Count; CDS Count; CDS %; RNA Count; rRNA Count; 5S rRNA Count; 16S rRNA Count; 18S rRNA Count; 23S rRNA Count; 28S rRNA Count; tRNA Count; Other RNA Count; Fused Count; Fused %; Fusion Component Count; Fusion component %; Pseudo Count; Pseudo %; Unchar Count; Unchar %; Obsolete Count; Obsolete %; Revised Count; Revised %; w/ Func Pred Count; w/ Func Pred %; w/o Func Pred Sim Count; w/o Func Pred Sim %; w/o Func Pred No Sim Count; w/o Func Pred No Sim %; Signal Peptide Count; Signal Peptide %; Transmembrane Count; Transmembrane %; SwissProt Count; SwissProt %; Not SwissProt Count; Not SwissProt %; SEED Count; SEED %; Not SEED Count; Not SEED %; COG Count; COG %; Pfam Count; Pfam %; TIGRfam Count; TIGRfam %; COG Cluster Count; Pfam Cluster Count; TIGRfam Cluster Count; InterPro Count; InterPro %; Enzyme Count; Enzyme %; TC Count; TC %; KEGG Count; KEGG %; Not KEGG Count; Not KEGG %; KO Count; KO %; Not KO Count; Not KO %; MetaCyc Count; MetaCyc %; Not MetaCyc Count; Not MetaCyc %; IMG Term Count; IMG Term %; IMG Pathway Count; IMG Pathway %; IMG Parts List Count; IMG Parts List %; Genome Property Count; Genome Property %; Chromosomal Cassette Gene Count; Chromosomal Cassette Gene %; Chromosomal Cassette Count.

Users can use “Expand All” and “Collapse All” buttons to show or hide the entire columns, or click on the top “+/-” symbols to display or hide columns in each group.

To select/deselect columns in **Table Configuration** selector, a user simply needs to select/deselect the checkbox next to the column name. The user can also click on the checkbox next to the top group name to select/deselect the whole group. There are also “Select All”, “Select Counts Only”, “Select Percentage Only”, “Clear All” buttons at the bottom to facilitate such selections/de-selections.

To add/remove displayed columns in table, a user first selects/deselects the column names in the **Table Configuration** selector, and then clicks the “Display Genomes Again” button to display.

Genome Search

By Field

The “Scaffold Object ID” is the newly added filter in IMG 3.3, as shown in Figure 29.

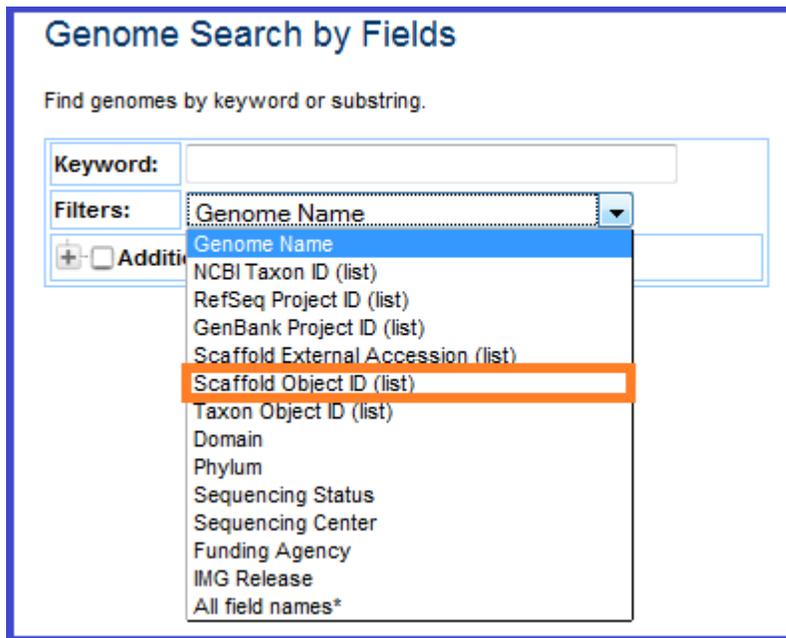


Figure 29: Scaffold Object ID as Search Filter.

Gene Search

Additional Output Columns

Users are also provided with the capability to select and display “Additional Output Columns” in result page, as shown in Figure 30:

Keyword:

Filters: Gene Product Name (inexact) ▼

Additional Output Columns

- Gene Symbol
- GenBank Accession
- Chromosome
- Start Coord
- End Coord
- Strand
- DNA Sequence Length
- AA Sequence Length
- Locus Type
- Is Pseudogene
- Is Obsolete
- Add Date
- Scaffold Object ID

(not applied to 'Pfam Domain', 'Protein Regular Expression Pattern' searches)

Figure 30: Additional Output Columns for Selection and Display.

These additional display columns are arranged into the following five groups:

1. Gene Symbol; GenBank Accession
2. Chromosome; Start Coord; End Coord; Strand; DNA Sequence Length; AA Sequence Length; Locus Type
3. Is Pseudogene; Is Obsolete
4. Add Date
5. Scaffold

Gene Cart

Table Configuration

Gene Cart table is added with the functionality of **Table Configuration**, as shown in Figure 31.

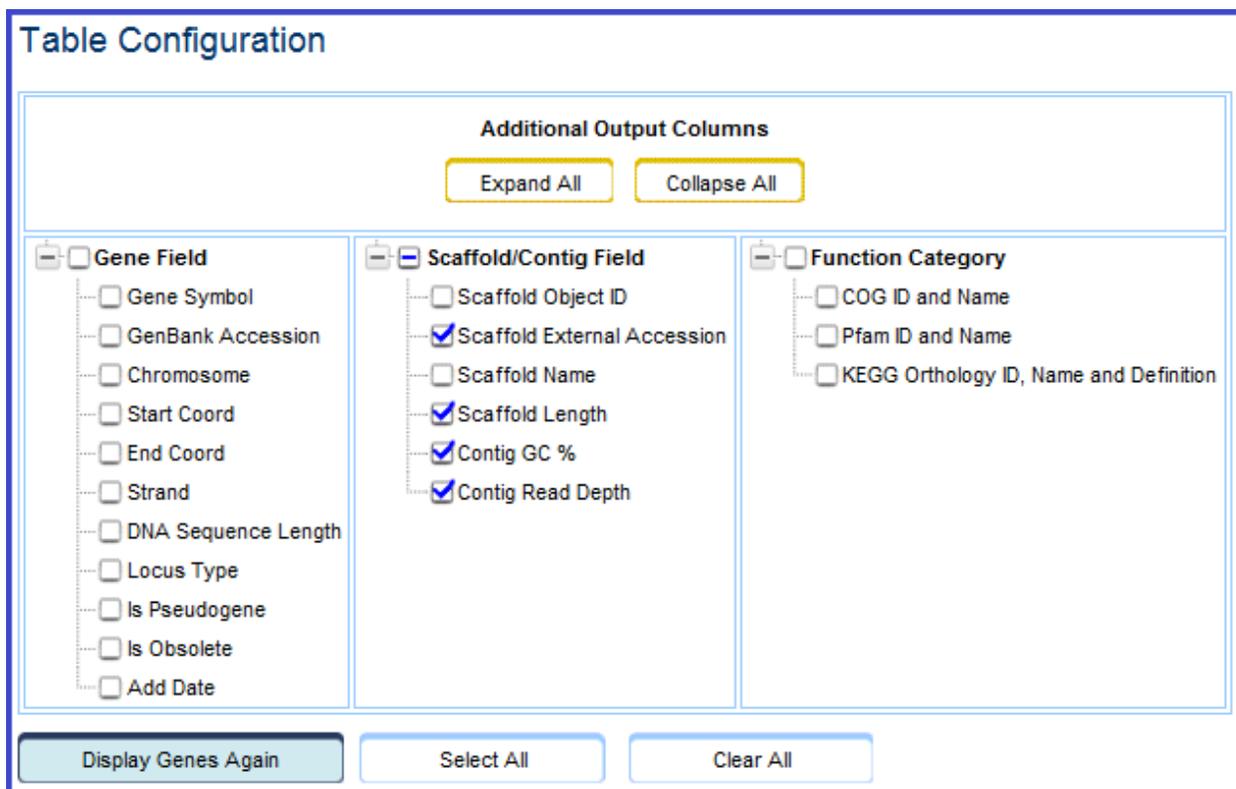


Figure 31: Table Configuration.

The **Table Configuration** selector lists three groups of “Additional Output Columns”:

1. Gene Field:
Gene Symbol; GenBank Accession; Chromosome; Start Coord; End Coord; Strand;
DNA Sequence Length; Locus Type; Is Pseudogene; Is Obsolete; Add Date
2. Scaffold/Contig Field:
Scaffold Object ID; Scaffold External Accession; Scaffold Name; Scaffold Length;
Contig GC %; Contig Read Depth
3. Function Category:
COG ID and Name; Pfam ID and Name; KEGG Orthology ID, Name and Definition

Users can use “Expand All” and “Collapse All” buttons to show or hide these columns, or click on the top “+/-” symbols to display or hide columns in each group.

What's New in IMG 3.3

To select/deselect columns in **Table Configuration** selector, a user simply needs to select/deselect the checkbox next to the column name. The user can also click on the checkbox next to the top group name to select/deselect the whole group. The “Select All” and “Clear All” buttons at the bottom are provided to facilitate such selections/de-selections. After choosing the desired columns, the user needs to click the “Display Genomes Again” button to redisplay.

Moreover, the “Add to Function Cart” button also has the capability to add “GO,” “InterPro” and “Transporter Classification” into the Function Cart, as shown in Figure 32.

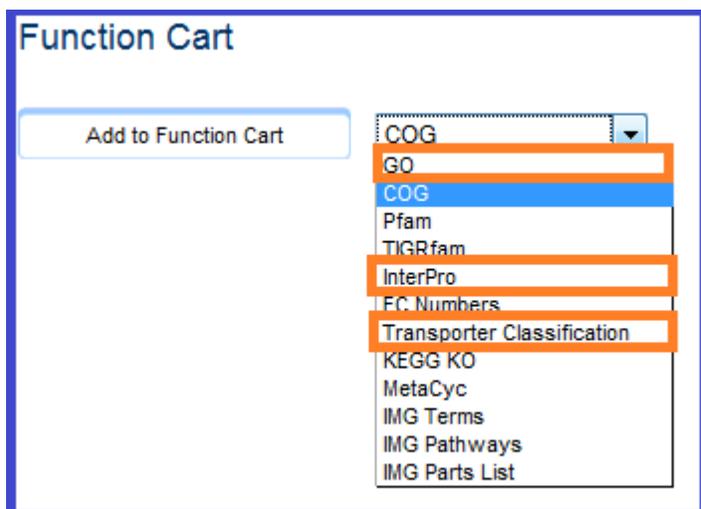


Figure 32: GO, InterPro and Transporter Classification in “Add to Function Cart”.

Function Cart

The Function Cart table now includes “GO,” “InterPro” and “Transporter Classification” as shown in Figure 33. Their underneath genes for specific genomes can also be added into Gene Cart through the “Add to Gene Cart” button.

Function List

3 function(s) in cart

Remove Selected Select All Clear All Toggle Selected

3 of 3 rows selected

Search column: Batch Search term:

Export Page 1 of 1 << first < prev 1 next > last >> All

Column Selector Select Page Deselect Page

Selection	Function ID	Name	Batch ¹
<input checked="" type="checkbox"/>	3.D.9	The H ⁺ -translocating F ₄₂₀ H ₂ Dehydrogenase (F ₄₂₀ H ₂ DH) Family	2
<input checked="" type="checkbox"/>	GO:0004765	shikimate kinase activity	2
<input checked="" type="checkbox"/>	IPR010189	Shikimate kinase, archaea	2

Export Page 1 of 1 << first < prev 1 next > last >> All

3 of 3 rows selected

1 - Each time a set of functions is added to the cart, a new distinguishing batch number is generated for the set.

Figure 33: GO, InterPro and Transporter Classification in Function Cart.

Function Search

We have added “Transporter Classification (list)” and “MetaCyc (list)” to the function search filters as shown in Figure 34. These two new options are included into the “All function names*” search too.

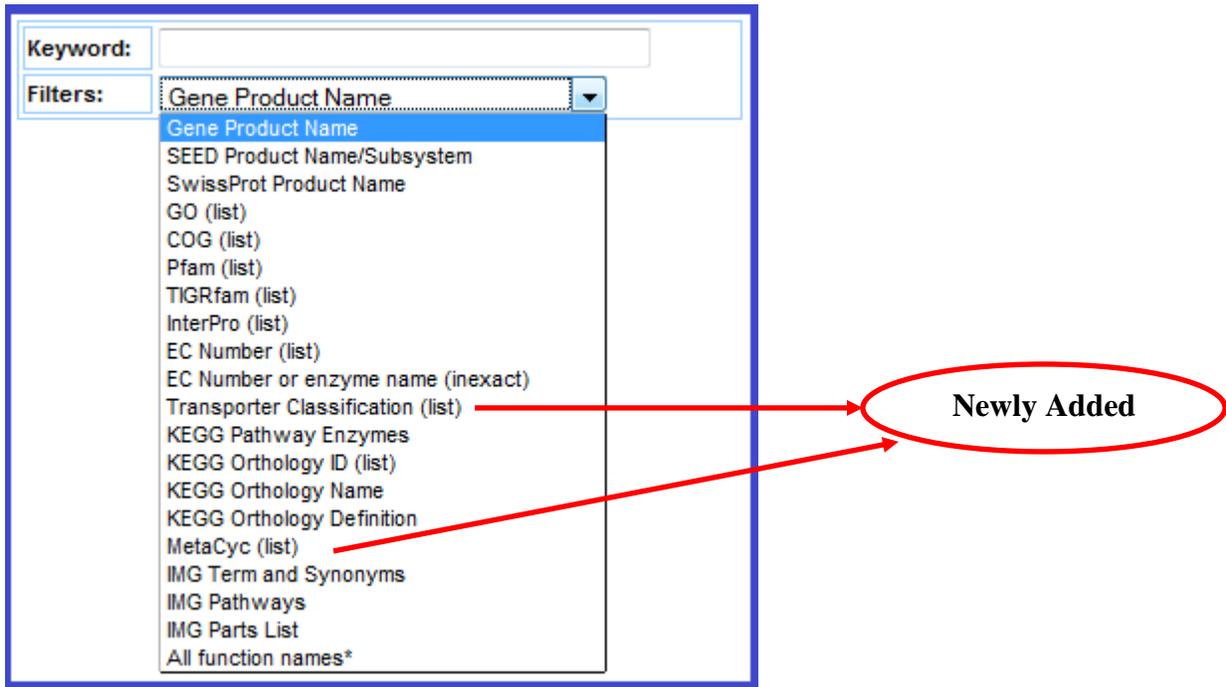


Figure 34: Function Search Filters.

BLAST

The **Genome filter** is applied to **Blast**, and contributes to the new look-and-feel as shown in Figure 35.

The behavior of **Genome Filter** in **Blast** is slightly different from that in **Find Genes**, **Find Functions**, etc.. Even if there's genome selection, the **Genome Filter** in **Blast** is still defaulted to "All IMG Genomes," and the search is conducted on "All IMG Genomes." However, user can override the selection via the **Genome Filter** provided in the page, if they have selected and saved genomes.

The screenshot shows the BLAST search interface. At the top, it says "BLAST" and "Find matches in selected genomes ('All IMG genomes' if none are selected). If there are selections in Genome Browser, please use 'Show All/Selected Genomes' buttons to access 'All IMG genomes' or 'Currently selected genomes from Genome Browser'." Below this is a large text area labeled "Sequence". Underneath the sequence area are two dropdown menus: "Program" set to "blastp (Protein vs. Protein)" and "E-value" set to "1e-5".

The "Genome Filter" section is below. It includes instructions: "Domains: (B)acteria, (A)rchaea, (E)ukarya, (P)lasmids, (V)iruses. Genome Completion: [F]inished, [P]ermanent Draft, [D]raft." There are two dropdown menus: "Seq. Status" set to "All Finished, Permanent Draft and Draft" and "Domain" set to "All". Below these are radio buttons for "List" (selected) and "Tree". A scrollable list of genomes is shown, including:

- Acidilobus saccharovorans 345-15 (A)[F]
- Aciduliprofundum boonei T469 (A)[D]
- Aciduliprofundum boonei T469 (A)[F]
- Aeropyrum pernix K1 (A)[F]
- Archaeoglobus fulgidus DSM 4304 (A)[F]
- Archaeoglobus profundus DSM 5631 (A)[F]
- Caldivirga maquilingensis IC-167 (A)[F]
- Candidatus Korarchaeum cryptofilum OPF8 (A)[F]
- Candidatus Methanoregula boonei 6A8 (A)[F]
- Cenarchaeum symbiosum A. (A)[F]

At the bottom are two buttons: "Run BLAST" and "Reset".

Figure 35: BLAST.

SEED

SEED Browser

IMG 3.3 provides **SEED Browser** for users to browse SEED product names and subsystems. The function can be accessed through **SEED** on the second-level menu of **Find Functions** as shown in Figure 36.

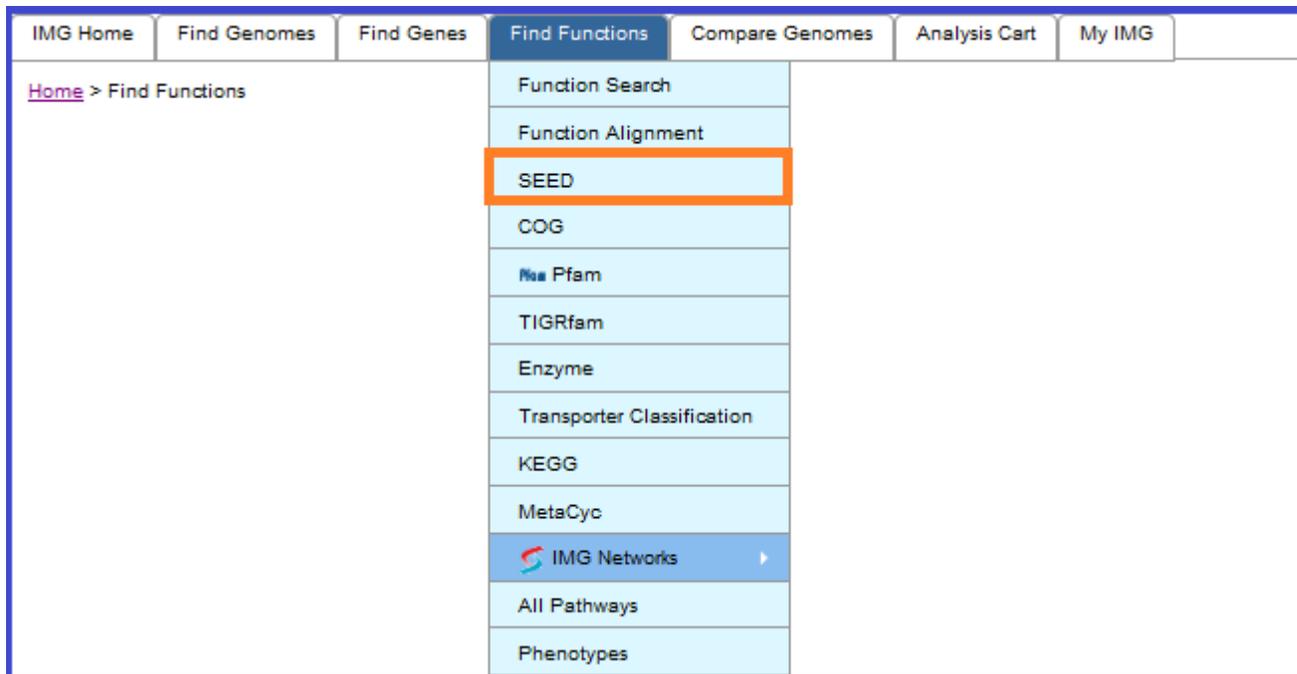


Figure 36: Access SEED Browser.

The page of **SEED Browser**, as shown in Figure 37, displays a complete list of SEED product names and subsystems. Clicking the link for “SEED product name” results in a table display of that particular SEED product (see Figure 38).

What's New in IMG 3.3

[IMG Home](#) | [Find Genomes](#) | [Find Genes](#) | **Find Functions** | [Compare Genomes](#) | [Analysis Cart](#) | [My IMG](#)

Home > Find Functions 14382 SEED Product (1094 Subsystem) Loaded.

SEED Browser

[SEED list](#)

01 Amino Acids and Derivatives

02 Alanine, serine, and glycine

Alanine_biosynthesis

- [Alanine racemase \(EC 5.1.1.1\)](#)
- [Alanine racemase, biosynthetic \(EC 5.1.1.1\)](#)
- [Alanine racemase, catabolic \(EC 5.1.1.1\)](#)
- [Branched-chain amino acid aminotransferase \(EC 2.6.1.42\)](#)
- [Cysteine desulfurase \(EC 2.8.1.7\)](#)
- [Cysteine desulfurase \(EC 2.8.1.7\), lscS subfamily](#)
- [Cysteine desulfurase \(EC 2.8.1.7\), NifS subfamily](#)
- [Cysteine desulfurase \(EC 2.8.1.7\), SufS subfamily](#)
- [Cysteine desulfurase CsdA-CsdE \(EC 2.8.1.7\), main protein CsdA](#)
- [Cysteine desulfurase, mitochondrial precursor \(EC 2.8.1.7\)](#)
- [Glutamate-pyruvate aminotransferase \(EC 2.6.1.2\)](#)
- [HTH-type transcriptional regulator IlvY](#)
- [Probable valine-pyruvate aminotransferase \(EC 2.6.1.66\)](#)
- [Valine--pyruvate aminotransferase \(EC 2.6.1.66\)](#)

Glycine_Biosynthesis

- [2-amino-3-ketobutyrate coenzyme A ligase \(EC 2.3.1.29\)](#)

Glycine_riboswitch

- [L-threonine 3-dehydrogenase \(EC 1.1.1.103\)](#)
- [Low-specificity L-threonine aldolase \(EC 4.1.2.5\)](#)
- [Serine hydroxymethyltransferase \(EC 2.1.2.1\)](#)

Glycine_and_Serine_Utilization

- [2-amino-3-ketobutyrate coenzyme A ligase \(EC 2.3.1.29\)](#)
- [Aminomethyltransferase \(glycine cleavage system T protein\) \(EC 2.1.2.10\)](#)
- [Cystathionine beta-synthase \(EC 4.2.1.22\)](#)
- [Cystathionine gamma-lyase \(EC 4.4.1.1\)](#)
- [D-3-phosphoglycerate dehydrogenase \(EC 1.1.1.95\)](#)
- [D-serine dehydratase \(EC 4.3.1.18\)](#)
- [D-serine dehydratase transcriptional activator](#)
- [D-serine permease DsdX](#)
- [D-serine/D-alanine/glycine transporter](#)
- [Glycerate kinase \(EC 2.7.1.31\)](#)

Figure 37: SEED Browser.

[IMG Home](#) | [Find Genomes](#) | [Find Genes](#) | **Find Functions** | [Compare Genomes](#) | [Analysis Cart](#) | [My IMG](#)

Home > Find Functions 1 SEEDs retrieved

SEEDs

Select All | Clear All

Search column: SEED Product Name | Search term:

Export | Page 1 of 1 | << first < prev 1 next > last >> | All

Column Selector | Select Page | Deselect Page

Select	SEED Product Name	SEED Subsystem	Level 2	Level 1	Gene Count	Genome Count
<input type="checkbox"/>	Alanine racemase (EC 5.1.1.1)	Alanine_biosynthesis	Alanine, serine, and glycine	Amino Acids and Derivatives	854	771

Export | Page 1 of 1 | << first < prev 1 next > last >> | All

Figure 38: Table Display of individual SEED.

SEED List

Furthermore, a user can click the “SEED List” link to have a table list of all the SEEDs as shown in Figure 39. The user can select and export the data of desired rows by using the provided “Select All”/“Clear All” buttons.

Home > Find Functions 14382 SEEDs retrieved

SEEDs

Select All Clear All

Search column: SEED Product Name Search term:

Export Page 1 of 144 << first < prev 1 2 3 4 5 6 7 8 9 10 next > last >> 100

Column Selector Select Page Deselect Page

Select	SEED Product Name	SEED Subsystem	Level 2	Level 1	Gene Count	Genome Count
<input type="checkbox"/>	"phi-Carotenoid synthase" (EC 1.3.-.- and EC 2.1.1.-)	Carotenoids	Isoprenoids	Fatty Acids, Lipids, and Isoprenoids	13	13
<input type="checkbox"/>	(3R)-hydroxymyristoyl-[ACP] dehydratase (EC 4.2.1.-)	Phospholipid_and_Fatty_acid_biosynthesis_related_cluster	Fatty acids	Fatty Acids, Lipids, and Isoprenoids	98	97
<input type="checkbox"/>	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)	Fatty_Acid_Biosynthesis_FASII	Fatty acids	Fatty Acids, Lipids, and Isoprenoids	1241	1157
<input type="checkbox"/>	(GlcNAc)2 ABC transporter, ATP-binding component 1	(GlcNAc)2_Catabolic_Operon	Aminosugars	Carbohydrates	52	52
<input type="checkbox"/>	(GlcNAc)2 ABC transporter, ATP-binding component 2	(GlcNAc)2_Catabolic_Operon	Aminosugars	Carbohydrates	45	45
<input type="checkbox"/>	(GlcNAc)2 ABC transporter, periplasmic substrate-binding protein	(GlcNAc)2_Catabolic_Operon	Aminosugars	Carbohydrates	54	54
<input type="checkbox"/>	(GlcNAc)2 ABC transporter, permease component 1	(GlcNAc)2_Catabolic_Operon	Aminosugars	Carbohydrates	52	52
<input type="checkbox"/>	(GlcNAc)2 ABC transporter, permease component 2	(GlcNAc)2_Catabolic_Operon	Aminosugars	Carbohydrates	54	54
<input type="checkbox"/>	(Pyruvate) Oxoisovalerate Dehydrogenase Alpha & Beta Fusion like	Dehydrogenase_complexes	Central carbohydrate metabolism	Carbohydrates	20	20

Figure 39: Table Display of All SEEDs.

Transporter Classification

Transporter Classification Browser

Transporter Classification Browser provides a way to browsing Transporter Classification families. The function can be accessed through **Transporter Classification** on the second-level menu of **Find Functions** as shown in Figure 40.

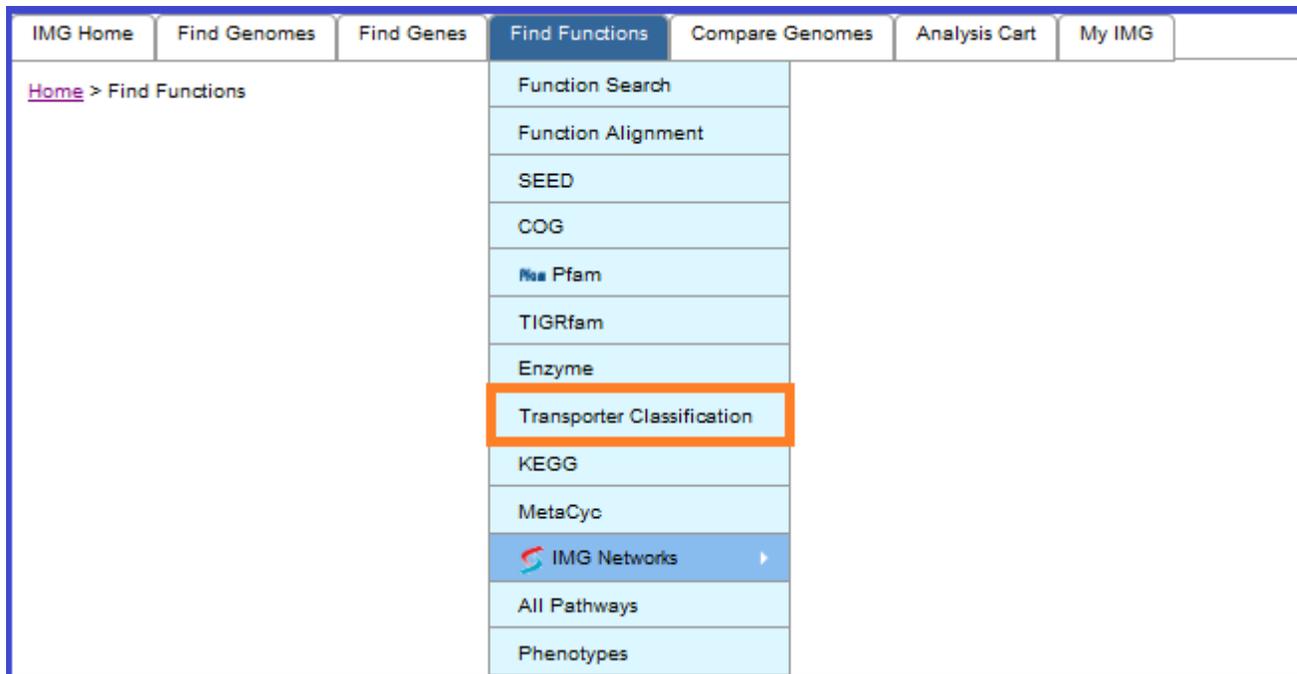


Figure 40: Access Transporter Classification Browser.

The page of **Transporter Classification Browser**, as shown in **Error! Reference source not found.**, displays a complete list of Transporter Classification families.

What's New in IMG 3.3

Figure 41 Transporter Classification Browser

Clicking the link for “Transporter Classification family number” leads to the website of Transporter Classification Database. On the other hand, clicking the link for “Transporter Classification family name” results in a table display of that Transporter Classification family as shown in Figure 42.

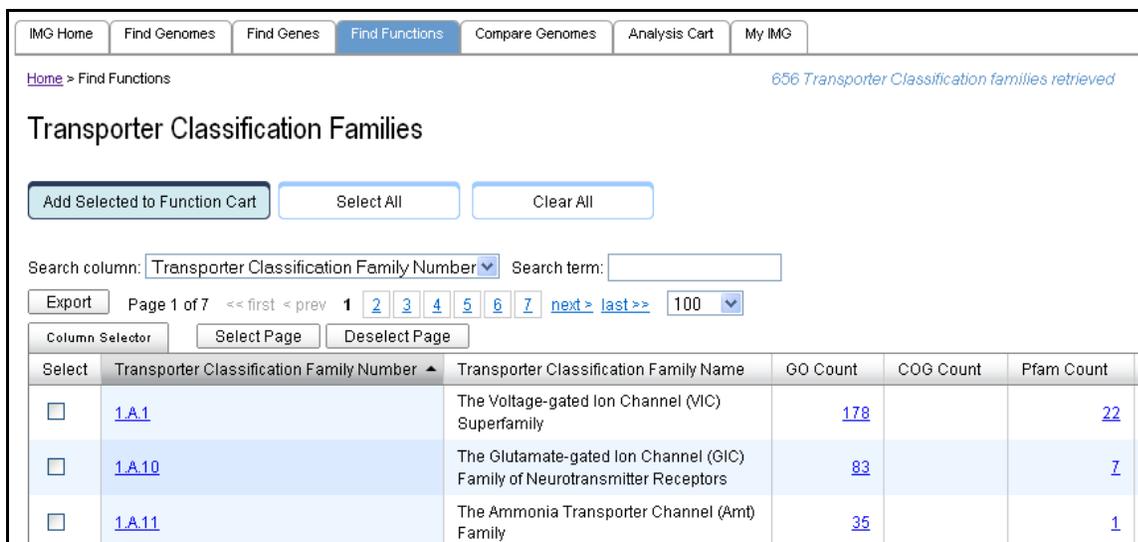
Select	Transporter Classification Family Number	Transporter Classification Family Name	GO Count	COG Count	Pfam Count	IMG
<input type="checkbox"/>	1.A.30	The H ⁺ - or Na ⁺ -translocating Bacterial Flagellar Motor/ExbBD Outer Membrane Transport Energizer (Mot/Exb) Superfamily	16	1	3	

Figure 42: Table Display of individual Transporter Classification family.

Transporter Classification Family List

Users can click the link for “Transporter Classification Family List” to have a table list of all the Transporter Classification families as shown in Figure 43. Users not only can use the provided “Select All”/“Clear All” buttons to select and export the data of desired rows, but also can add them into the Function Cart.

What's New in IMG 3.3



Select	Transporter Classification Family Number	Transporter Classification Family Name	GO Count	COG Count	Pfam Count
<input type="checkbox"/>	1.A.1	The Voltage-gated Ion Channel (VIC) Superfamily	178		22
<input type="checkbox"/>	1.A.10	The Glutamate-gated Ion Channel (GIC) Family of Neurotransmitter Receptors	83		7
<input type="checkbox"/>	1.A.11	The Ammonia Transporter Channel (Amt) Family	35		1

Figure 43: Table Display of All Transporter Classification families.

Site Map

The new **Site Map** not only helps user with easy site navigation, but also includes the information of “**What’s New**” now. It can be accessed on the third-level menu of **Using IMG** as shown in Figure 44.

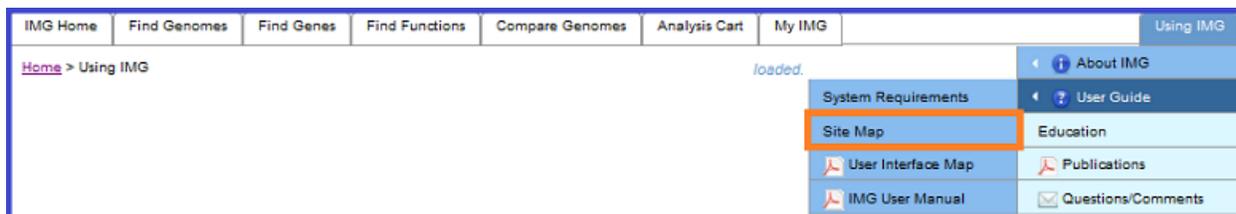


Figure 44: Access Site Map.

The **Site Map** consists of the following sections:

1. **What’s new**

Provide the link to the current release note as shown in Figure 45.



Figure 45: What’s New.

2. **Navigation Menus**

What's New in IMG 3.3

The navigation menus not only list menu and useful links to the tool and documentation, but also record the information of **What's New** through the “updated” and “new” icons as shown in Figure 46.

Menu	Description	Document
IMG Home	IMG home page	
Find Genomes		
Genome Browser updated		
 Genome Search	By Fields By Metadata	
Find Genes		
 Gene Search updated	Find genes in selected genomes by keyword.	
 BLAST updated	Find sequence similarity in IMG database.	
Phylogenetic Profilers	Find genes in genome (bin) of interest qualified by similarity to sequences in other genomes (based on BLASTP alignments). Only user-selected genomes appear in the profiler.	
Find Functions		
Function Search updated	Find functions in selected genomes by keyword.	
Function Alignment	Find alignments in selected functions and genomes by keyword.	
SEED NEW	List of SEED product names and subsystems.	
COG	List of COGs	
 Pfam	Pfam list.	
TIGRfam	TIGRfam roles and list.	
Enzyme	List of Enzymes, EC numbers.	
Transporter Classification NEW	List of Transporter Classification families.	
KEGG	KEGG Orthology (KO) Terms and Pathways KO Term Distribution KEGG Orthology (KO) Terms	
MetaCyc	MetaCyc Pathways	
 IMG Network	IMG Network Browser IMG Parts List IMG Pathways IMG Terms	
 IMG Network Browser		
 IMG Parts List		
 IMG Pathways		
 IMG Terms		
All Pathways		
Phenotypes		

Figure 46: Navigation Menus.

3. Sub-Pages and Components

Sub-Pages and Components page lists important sub-pages and components that are not covered by navigation menus as shown in Figure 47.

What's New in IMG 3.3

sub-Pages and Components		
List of important sub-pages and components that are not covered by navigation menus.		
Page	Description	Document
Genome Detail Page		
Organism Information	[see example]	
Genome Statistics	[see example]	
Phylogenetic Distribution of Genes	[see example]	
Putative Horizontally Transferred Genes	[see example]	
Genome Viewers		
Scaffolds and Contigs	[see example]	
Chromosome Maps	[see example]	
Web Artemis	[see example]	
Compare Gene Annotations		
Download Gene Information		
Export Genome Data		
FASTA	FASTA nucleic acid file for all scaffolds FASTA amino acid file for all proteins FASTA nucleic acid file for all genes FASTA intergenic sequences	
Tab Delimited		
Genbank		
Generate GenBank File		
Gene Detail Page		
Gene Information	[see example]	
Add to Gene Cart	Click on "Add To Gene Cart" button under Gene Information section to add this gene to the gene cart	
Find Candidate Enzymes	Click on "Find Candidate Enzymes" button under Gene Information section to find candidate enzymes for this gene	
Find Candidate Product Name	Find candidate product name for this gene. [see example]	
Evidence For Function Predictions		
Neighborhood	Sequence Viewer For Alternate ORF Search [see example] Chromosome Viewer (colored by COG, GC, KEGG, Pfam TIGRfam, Expression) [see example]	
Conserved Neighborhood	Ortholog Neighborhood Viewer Chromosomal Cassette Viewer (COG, IMG Ortholog Cluster, Pfam)	
External Sequence Search	[see example]	

Figure 47: Sub-Pages and components.

4. Archive of past What's New

Archive provides links to past documents of "What's New" as shown in Figure 48.

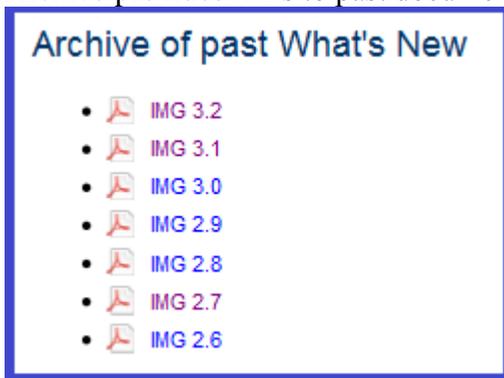


Figure 48: Archive of past What's New.