

Trouble Shooting GENI-ACT

The GENI-ACT site is down. What can I do?

Unfortunately the GENI-ACT site goes offline at times without warning. We know that you are often restricted in the timing you can work with your students and we have developed the following “work around” to help you salvage your work session with your students.

If you have not done so already, it is a good idea to have your students keep a copy of the FASTA formatted amino acid sequence of the gene on which you are working in a Word or text file saved on their computer or a flash drive. Teachers can access the basic information page of all their student’s genes to keep a master list as well. At the very least the locus tag of genes students are working on should be available. The procedure below provides directions for you to retrieve an amino acid sequence on which to work. If you already have your FASTA formatted sequence available to you, skip to step 6 below.

Procedures

1. Log into IMG/edu using the following link:
<http://img.jgi.doe.gov/cgi-bin/edu/main.cgi>.
2. Click on the Find Genes tab at the top of the page as indicated by the arrow in Figure 1 below and select the Gene Search option from the pull down menu.

The screenshot shows the IMG/M Home page. At the top, there is a navigation bar with tabs: Home, Find Genomes, Find Genes (indicated by a black arrow), Find Functions, Compare Genomes, OMICS, My IMG, Data Marts, and Help. Below the navigation bar, there is a section for 'IMG Content' with a table of datasets. To the right, there is a 'Quick Genome Search' box and a 'Go' button. Below that, there is a section for 'IMG Statistics' with a table showing the distribution of metagenome and metatranscriptome datasets. At the bottom, there is a table showing the distribution of IMG datasets by type and environment.

IMG Content		
Datasets	JGI	All
Bacteria	5763	39296
Archaea	269	773
Eukarya	25	220
Plasmids		1192
Viruses		3907
Genome Fragments		1192
Metagenome & Metatranscriptome	4322	5699
Total Datasets		52279

Metagenome and Metatranscriptome dataset distribution:						
Sequenced at:	Engineered		Environmental		Host-associated	
	JGI	All	JGI	All	JGI	All
Metagenome	365	476	2667	2912	420	1419
Metatranscriptome	120	134	649	655	100	102

IMG contains 250 public studies, 5699 public metagenome datasets (5194 unique samples) distributed as follows: (Public Metagenome count / Public Metatranscriptome count)						
Engineered	476 / 134	Environmental	2912 / 655	Host-associated	1419 / 102	
Bioreactor	15 / 4	Air	31 / 0	Algae	52 / 0	

Figure 1. The IMG/M Home page. The Find Genes tab is indicated by the arrow.

3. In the Gene Search window paste the locus tag for your gene in the keyword box, select Locus Tag (list) from the filters pull down menu (Figure 2) and click Go.
 - 3.1. The locus tags for all of our student's genes are on the set-up document sent to you by way of email when your gene assignments were set up.
 - 3.2. It is a good idea to keep the document at hand when working with your students.
 - 3.3. For those of you with group assignments, you should also keep a master list of which gene locus tag was assigned by you to each student in your group.
 - 3.4. If you have misplaced the document please notify Dr. Koury and he will resend it to you.

JGI  **IMG/M** 
INTEGRATED MICROBIAL GENOMES & MICROBIOME SAMPLES

Quick Genome Search:

My Analysis Carts**: 0 Genomes | 0 Scaffolds | 0 Functions | 0 Genes

Home Find Genomes Find Genes Find Functions Compare Genomes OMICS My IMG Data Marts Help

Home > Find Genes loaded. 

Gene Search

Find genes in selected genomes by keyword. It's required to add selections into "Selected Genomes" unless blocked.
*MER-FS Metagenome supported search filters.

Keyword

Filters

Figure 2. The Gene Search window at IMG/M.

- 4. The page that will appear is called the Gene Detail page in the IMG/edu database (Figure 3). It contains a wealth of information known about the gene.

The screenshot shows the top portion of the IMG Gene Detail page. At the top left is the JGI logo and the IMG/M logo with the tagline 'INTEGRATED MICROBIAL GENOMES & MICROBIOME SAMPLES'. To the right is a 'Quick Genome Search' box with a 'Go' button. Below this is a navigation bar with links for 'Home', 'Find Genomes', 'Find Genes', 'Find Functions', 'Compare Genomes', 'OMICS', 'My IMG', 'Data Marts', and 'Help'. A status bar indicates 'My Analysis Carts**': 0 Genomes | 0 Scaffolds | 0 Functions | 0 Genes. The main heading is 'Gene Detail' with a 'Loaded.' status. A list of links includes 'Gene Information', 'Find Candidate Product Name', 'Evidence For Function Predictions', 'Sequence Search', 'External Sequence Search', 'IMG Sequence Search', and 'Homolog Display'. Below this is a 'Gene Information' table.

Gene Information	
Gene ID	644990317
Gene Symbol	dnaA
Locus Tag	Ksed_00010
IMG Product Name	chromosomal replication initiator protein DnaA
Original Gene Product Name	chromosomal replication initiator protein DnaA
IMG Product Source	
SwissProt Protein Product	
SEED	[Chromosomal replication initiator protein DnaA] figl478801.4.peg.1954 DNA-replication
IMG Term	
Description	PFAM: Bacterial dnaA protein helix-turn-helix domain; Bacterial dnaA protein; TIGRFAM: chromosomal replication initiator protein DnaA
Genome	Kytococcus sedentarius 541, DSM 20547
DNA Coordinates	209..1729 (+)(1521bp)

Figure 3. The IMG Gene Detail Page. Only the uppermost portion of the page is shown.

5. Scrolling further down the page you will lead you to a section entitled “amino acid sequence length” with a hyperlink in the form of the number of amino acids in the sequence (arrow, Figure 4).

Gene Information

Gene Information	
Gene ID	644990317
Gene Symbol	
Locus Tag	Ksed_00010
Product Name	chromosomal replication initiator protein DnaA
SwissProt Protein Product	
SEED	[Chromosomal replication initiator protein DnaA] fig 478801.4.peg.1954 DNA-replication
IMG Term	
Description	PFAM: Bacterial dnaA protein helix-turn-helix domain; Bacterial dnaA protein; TIGRFAM: chromosomal replication initiator protein DnaA
Genome	Kytococcus sedentarius 541_DSM 20547
DNA Coordinates	209..1729 (+)(1521bp)
Scaffold Source	Kytococcus sedentarius DSM 20547: NC_013169 (2785024bp)
IMG ORF Type	
GC Content	0.68
Accession	YP_003147866
External Links	GI:256823906 ; GeneID:8374208 ; UniProtKB/TrEMBL:C7NID0
Fused Gene	No
GO Terms	GO:0005524 - ATP binding [evidence=IEA] GO:0005524 - ATP binding [evidence=IEA] GO:0003677 - DNA binding [evidence=IEA] GO:0003677 - DNA binding [evidence=IEA] GO:0003688 - DNA replication origin binding [evidence=IEA] GO:0003688 - DNA replication origin binding [evidence=IEA] GO:0017111 - nucleoside-triphosphatase activity [evidence=IEA] GO:0000166 - nucleotide binding [evidence=IEA] GO:0000166 - nucleotide binding [evidence=IEA] GO:0043565 - sequence-specific DNA binding [evidence=IEA]
Protein Information	
Amino Acid Sequence Length	506aa
COG	COG0593 - ATPase involved in DNA replication initiation
KOG	

Figure 4. The Gene Detail Page at IMG/edu showing where the amino acid sequence of the protein can be found. The arrow indicates a hyperlink that will lead you to the amino acid sequence.

6. Clicking on the Amino Acid Sequence Length hyperlink indicated by the arrow in figure 4 will open a FASTA formatted sequence, as shown in Figure 5. The sequence plus the FASTA header should be copied and pasted into a Word document or text editor. The FASTA header has a gene identifier number from the IMG database that is different from the locus tag we normally use. For consistency, replace the FASTA header from IMG/edu with the locus tag of your gene as described in the Basic Information Module Instructions and in the caption of figure 5.



The screenshot shows the JGI IMG/M website interface. At the top, there is a navigation bar with the JGI logo and the text "IMG/M INTEGRATED MICROBIAL GENOMES & MICROBIOME SAMPLES". A search bar is present with the text "Quick Genome Search:" and a "Go" button. Below the search bar, there is a status bar showing "My Analysis Carts**:" followed by counts for Genomes, Scaffolds, Functions, and Genes. A main navigation menu includes links for Home, Find Genomes, Find Genes, Find Functions, Compare Genomes, OMICS, My IMG, Data Marts, and Help. The breadcrumb trail shows "Home > Find Genes" and the word "loaded." is displayed. The main content area displays a FASTA formatted sequence for the gene Ksed_00010. The header is highlighted in blue and reads: ">644990317 YP_003147866 chromosomal replication initiator protein DnaA [Kytococcus sedentarius DSM 20547: NC_013169]". The sequence itself is a single line of amino acid letters. At the bottom of the page, there is a footer section with contact information, social media icons, and logos for GOLD and U.S. DEPARTMENT OF ENERGY Office of Science. The footer also includes the text: "JGI is a DOE Office of Science User Facility managed by Lawrence Berkeley National Laboratory. © 1997-2016 The Regents of the University of California. Version 4.560 Mar. 2016 gpweb36.nersc.gov gemini2_shared 5.016000 2016-06-30-11.20.13 128.205.95.195".

Home > Find Genes loaded.

```
>644990317 YP_003147866 chromosomal replication initiator protein DnaA [Kytococcus sedentarius DSM 20547: NC_013169]
MSQTPDDHATAIWQEMVHLQAGLAPRDIGVLRRLATLVGLLEGTALLAV
KYDHVKDAVEGHLREDVSTALAEVLDRLIRLAVSVDPDAVSAAQEEAAPP
APSPAEDDDPATGEGPLSTAVDGAVEKHEGSSPARAGESVAPATTASLTA
TNSSPGVERDYSALNHKYTFDTFVLGSSNRFAAAAATAVAEAPARAYNPL
FIYGGSGLGKTHLLHAIGHYARTLDSSVRVKYVNSEEFTNQFINAVSAGQ
ANAFQRQYRDVVDLLIDDIQFLQKQETMBEFPFHTFNTLHNSKQIVITS
DQPPKLLSGFAERMRSRFEWGLLTDVQPPDLETRIALLRKAAADKLDIP
DDVHLHLIASKISSNIRELEGALTRVTAFASLSGSPLDYLRARTVLKDVMP
GGDSGQITPMTILEETAGYFVISVEEIQGASRSRNLTRARQIAMYLCREL
TDLSLPKIGKEFGGRDHTVMHAERKIKQLLGEDRRVYDEVSELTSIIRK
KAARGR
```

Contact us
Accessibility/Section 508
Disclaimer

JGI is a DOE Office of Science User Facility managed by Lawrence Berkeley National Laboratory.
© 1997-2016 The Regents of the University of California

Version 4.560 Mar. 2016
gpweb36.nersc.gov gemini2_shared 5.016000 2016-06-30-11.20.13 128.205.95.195

Figure 5. The FASTA formatted sequence of Ksed_00010 as obtained from the IMG gene details page. The FASTA header in blue in the image above should be replaced with the locus tag of your gene. In this case the edited header would read >Ksed_00010 amino acid sequence, as is described in the Basic Information Module instructions

7. Once you have the FASTA formatted sequence of your protein, from either a saved file or from IMG/edu as described above, you can use the online module instructions to perform the annotation exercises.
8. You can access the online module instructions at:
<http://ubwp.buffalo.edu/wnygirahcp/educational-resources/>
9. Select the instructions for whichever module on which you are planning to work to see what happens in that module. Your students will then be able to find the bioinformatics tools on the Web and know how to use them. The module instructions will tell your students what data to save.
10. Your students can save their work in a Word or PowerPoint document and then upload it to the GENI-ACT notebook when it becomes available. Below is a link to a webpage that will have the Word document (blank_GENIACT_notebook_template.doc) that has all the GENI-ACT notebook sections in it, active links to the GENI-ACT tools and active links to instructions for each module.

The link is:

<https://sites.google.com/site/ubbclsgenomeannotation/educational-resources>