

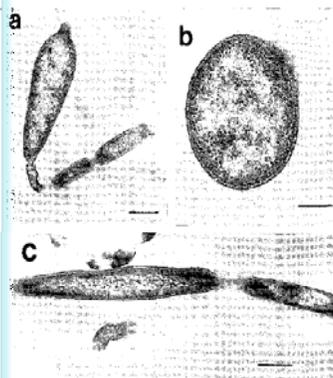
INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY, July 2004, p. 324-340  
 1024-7713/04/\$12.00  
 © 2004, International Union of Microbiological Societies

**Isolation and Characterization of *Halohermophilix orenii* gen. nov., sp. nov., a Halophilic, Thermophilic, Fermentative, Strictly Anaerobic Bacterium**

L. L. CAYO,<sup>1</sup> B. OLLIVIER,<sup>2</sup> B. K. C. DALL'AGLIO,<sup>3</sup> G. PRINSHOR,<sup>4</sup> F. GUZZINETTI,<sup>5</sup> and J. L. GARCÍA<sup>1</sup>

<sup>1</sup>Laboratoire de Microbiologie, ORSTOM, Université de Provence, 13545 Marseille cedex 13, <sup>2</sup>Laboratoire de Microbiologie, Université Blaise Pascal, 63177 Aubière cedex, <sup>3</sup>and Laboratoire de Chimie et Catalyse, Avda. de los Arqueólogos, 20273 Baza, Spain, <sup>4</sup>Faculty of Science and Technology, Griffith University, Brisbane, Queensland 4111, Australia<sup>5</sup>

The occurrence of this obligate, halophilic anaerobic bacterium in the sediments of a Tunisian salted lake was tested in samples collected at 20-cm intervals down to a depth of 1.20 m. A long rod, present only in the 40- to 60-cm layer, was isolated at 60°C in a medium containing 100 g of NaCl per liter and designated strain H168. This strain produced acetate, ethanol, H<sub>2</sub>, and CO<sub>2</sub> from glucose metabolism. Fructose, xylose, ribose, cellobiose, and starch were also oxidized. The optimum temperature for growth was 60°C. No growth was obtained at 42 or 70°C. Strain H168 grew optically in NaCl concentrations ranging from 50 to 100 g per liter, with the upper and lower limits of growth around 200 and -80 g per liter, respectively. The G+C ratio of the DNA was 59.6 mol%. Although halophilic, moderately thermophilic bacteria have been characterized among anaerobes, particularly within methanogens, strain H168 is the first true thermophilic (growing above 60°C) halophilic anaerobic bacterium described so far. The phylogeny, physiology, morphology, lipid content, and high G+C content of strain H168 are sufficiently different from those of genera belonging to the family *Haloanaerobaceae* to justify the definition of a new genus.



**A Genome Analysis Case Study with the Integrated Microbial Genome Expert Review (IMG/ER) System**

**Konstantinos Mavrommatis**

Genome Biology Program  
 Joint Genome Institute

**JGI**  
 DOE JOINT GENOME INSTITUTE  
 UNIVERSITY OF CALIFORNIA

All genomes selected  Quick Genome Search:  **GO**

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

Username:

Password:

Access to your genome  
 + public genomes

(If you are not a registered user, [request an account here](#).)

Version 2.4, December 2007  
[Questions/Comments](#)  
 ©2007 The Regents of the University of California  
[Disclaimer](#)

<http://img.jgi.doe.gov/er>

**JGI**  
DOE JOINT GENOME INSTITUTE  
IN DEPARTMENT OF ENERGY  
OFFICE OF SCIENCE

Quick Genome Search: Halofermothrix **GO**

Hi User, Logout

**img/er** INTEGRATED MICROBIAL GENOMES  
EXPERT REVIEW

IMG Home | **Find Genomes** | Find Genes | Find Functions | Compare Genomes | Analysis Carts | MyIMG | Using IMG

**IMG Genomes**

finished/draft	JGI	Total
Bacteria	152/134	501946
Archaea	14/11	42/14
Eukarya	2/0	192/1
Plasmids	0/0	40/40
Viruses	0/0	166/0
All Genomes	168/142	262798/1
Grand Total	310	3008
User Private	user	4

Version 2.4 December 2007  
[IMG Questions/Comments](#)  
[VISTA Questions/Comments](#)  
 ©2007 The Regents of the University of California  
[Disclaimer](#)

The Integrated Microbial Genomes- Expert Review (IMG/ER) system provides support to individual scientists or group of scientists for functional annotation and curation of their microbial genomes of interest. Often such genomes have not been vet deposited into the public genome sequence archives, and therefore access to them is restricted to specific scientists or groups. Genomes undergoing curation in IMG/ER are integrated with all publicly available genomes in IMG ([Nucleic Acids Research, 2006, Vol. 34, Database Issue D344-D348](#)) and provides the same framework for comparative analysis.

The current version of IMG/ER is based on [IMG 2.3](#). For more details, see [What's New](#) and [Using IMG](#). Please also see [About IMG](#) and [FAQ](#) on IMG.

**Publications**  
[IMG Tutorial](#)

**IMG Companion Systems:** [IMG](#) | [IMG.M](#) | [IMG.EDU](#)

**JGI**  
DOE JOINT GENOME INSTITUTE  
IN DEPARTMENT OF ENERGY  
OFFICE OF SCIENCE

All genomes selected | Quick Genome Search: **GO**

Hi nikos, Logout

**img/er** INTEGRATED MICROBIAL GENOMES  
EXPERT REVIEW

IMG Home | **Find Genomes** | Find Genes | Find Functions | Compare Genomes | Analysis Carts | MyIMG | Using IMG

Genome Browser | Genome Search | Taxonomy Browser | Category Browser

**Genome Browser** 943 genome(s) retrieved. ←

Save Selections | Select All | Clear All | View Phylogenetically

**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, D=Draft.  
 Click on column name to sort.

Select	D	C	Genome Name	Sequencing Center	Gene Count	Genome Size
<input checked="" type="checkbox"/>	A	D	<a href="#">Aciduliprofundum boonei T469</a>	J. Craig Venter Institute	3036	2981805
<input checked="" type="checkbox"/>	A	F	<a href="#">Aeropyrum pernix K1</a>	NITE	2815	1669695
<input checked="" type="checkbox"/>	A	F	<a href="#">Archaeoglobus fulgidus DSM 4304</a>	Univ of Illinois at Urbana-Champaign, TIGR	2519	2178400
<input checked="" type="checkbox"/>	A	D	<a href="#">Caldivaga marquilingensis IC-167</a>	JGI	1986	2077575
<input checked="" type="checkbox"/>	A	D	<a href="#">Cenarchaeum symbiosum</a>	JGI	2065	2045088
<input checked="" type="checkbox"/>	A	D	<a href="#">Ferroplasma acidimanus Fer1</a>	JGI	1740	1865438
<input checked="" type="checkbox"/>	A	F	<a href="#">Halobaculum marismortui ATCC 43049</a>	Institute for Systems Biology	4385	4274642
<input checked="" type="checkbox"/>	A	F	<a href="#">Halobacterium sp. NRC-1</a>	Univ of Washington- Seattle, Univ of Massachusetts	2728	2571010
<input checked="" type="checkbox"/>	A	D	<a href="#">Haloflex volcanii DS2</a>	TIGR	4112	4012900
<input checked="" type="checkbox"/>	A	D	<a href="#">Halogeometricum borinquense DSM 11551</a>	JGI	4060	3921798

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

IMG Home Find Genomes Find Genes Find Functions Compare Genomes Analysis Carts MyIMG About IMG

Genome Browser | Genome Search

### Genome Browser

Save Selections Select All Clear All View Alphabetically

**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.

Genome Completion: [F]inished, [D]raft.

- 01 **Archaea** All None
  - 02 **Crenarchaeota** All None
    - 03 **Thermoprotei** All None
      - 04 **Cenarchaeales** All None
        - 05 **Cenarchaeaceae** All None
          - 08  **Cenarchaeum symbiosum** [D] (JGI)
            - 05 **Syntrophomonadaceae** All None
              - 08  **Syntrophomonas wolfei wolfei** Goettingen [F] (JGI)
                - 05 **unclassified** All None
                  - 08  **Caldicellulosiruptor saccharolyticus** DSM 3903 [F] (JGI)
                    - 08  **Thermosinus carboxydivorans** Nor1 [D] (JGI)
                      - 04 **Halanaerobiales** All None
                        - 05 **Halanaerobiaceae** All None
                          - 08  **Halothermothrix orenii** H 168 (Finished) [F] (JGI)
                            - 08  **Halothermothrix orenii** H 168 (Finished v2) [F] (JGI)
                              - 08  **Halothermothrix orenii** H 168 [D] (JGI)
                                - 04 **Thermoanaerobacteriales** All None
                                  - 08  **Halothermothrix**

Find: Halothermothrix Next Previous Highlight all Match case

**JGI** Organism page All genomes selected Quick Genome Search

Hi nikos, Logout

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

IMG Home Find Genomes Find Genes Find Functions Compare Genomes Analysis Carts MyIMG

### Organism Details

Loaded

[Organism Information](#)  
[Genome Statistics](#)  
[Genome Viewers](#)  
[Export Genome Data](#)

### Organism Information

|                   |  |
|-------------------|--|
| Organism Name     | Halothermothrix orenii H 168 (v3)  |
| Taxon Object ID   | 2300395326   |
| GOLD ID           | G01042   |
| External Links    |  |
| Lineage           | Bacteria; Firmicutes; Clostridia; Halanaerobiales; Halanaerobiaceae; Halothermothrix |
| Sequencing Status | Finished   |
| Sequencing Center | JGI  |
| Funding Agency    | DOE  |
| Is Public         | No   |
| JGI Project ID    | 0  |
| Phenotype         |  |
| Habitat           | Aquatic  |
| Disease           | None   |
| Relevance         | Biotechnological   |
| IMG Release       |  |
| Comment           | JGI's Genome Portal provides sequence files and annotation.                          |

### Genome Statistics

|                                      | Number  | % of Total |
|--------------------------------------|---------|------------|
| <b>DNA, total number of bases</b>    | 2578146 | 100.00%    |
| DNA coding number of bases           | 2281062 | 88.48%     |
| DNA G+C number of bases              | 976684  | 37.89%     |
| <b>DNA scaffolds</b>                 | 1       | 100.00%    |
| <b>Genes total number</b>            | 2446    | 100.00%    |
| Protein coding genes                 | 2361    | 94.52%     |
| rRNA genes                           | 85      | 3.48%      |
| rRNA genes                           | 11      | 0.45%      |
| 5S rRNA                              | 2       | 0.12%      |
| 16S rRNA                             | 4       | 0.16%      |
| 18S rRNA                             | 0       | 0.00%      |
| 23S rRNA                             | 4       | 0.16%      |
| 26S rRNA                             | 0       | 0.00%      |
| tRNA genes                           | 56      | 2.29%      |
| Other RNA genes                      | 18      | 0.74%      |
| Genes with function prediction       | 1785    | 73.39%     |
| Genes without function prediction    | 566     | 23.14%     |
| Pseudo Genes                         | 21      | 1.19%      |
| Genes assigned to enzymes            | 588     | 24.45%     |
| Genes connected to KEGG pathways     | 505     | 20.69%     |
| Genes not connected to KEGG pathways | 1856    | 75.88%     |
| Genes in internal clusters           | 305     | 12.47%     |
| Genes in COGs                        | 1865    | 76.25%     |
| Genes in Pfam                        | 1880    | 76.89%     |
| Genes in TIGRfam                     | 900     | 37.04%     |
| Genes coding signal peptides         | 634     | 25.92%     |
| Genes coding transmembrane proteins  | 710     | 29.39%     |
| Genes in InterPro                    | 0       | 0.00%      |
| Genes with IMG Terms                 | 0       | 0.00%      |
| Genes in IMG Pathways                | 0       | 0.00%      |
| Genes in IMG Parts List              | 0       | 0.00%      |
| Genes with MyIMG Annotation          | 0       | 0.00%      |
| Obsolete Genes                       | 0       | 0.00%      |
| Reversed Genes                       | 0       | 0.00%      |

**Genome Viewers**

- Scaffolds and Contigs
- Chromosome Maps
- Web Artemis

**Compare Gene Annotations**

- Compare Gene Annotations

Gene annotation values are precomputed and stored in a tab delimited file also viewable in Excel.

**Export Genome Data**

Download sequences and gene information for this genome.

- [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 file for Excel with gene information](#)

Generate Genbank File

**hint:** Right click on link to see menu for saving link contents to target file. Please be patient during download.

**Chromosome Viewer**

Scaffolds and contigs for *Halothermothrix orenii* H168 (Finished v2)

User Selectable Coordinates

| Scaffold   | Length (bp) | GC   | No. Genes | Coordinate Range   |
|--|-------------|------|-----------|--|
| Halothermothrix orenii H168 (Finished v2) : haloTherm168v2_haloOrH168_Contig4393 | 2578146     | 0.38 | 2446      | 1-500000<br>500001-1000000<br>1000001-1500000<br>1500001-2000000<br>2000001-2500000<br>2500001-2578146 |

User Enterable Coordinates

Halothermothrix orenii H168 (Finished v2) : haloTherm168v2 haloOrH168 Cov: Start End

Go Reset

**hint:** WARNING: Some browsers may be overwhelmed by a large coordinate range.

**Chromosome Viewer**

Switch coloring to: GC percentage

Halothermothrix orenii H168 (v3) : HaloTherm168v2\_haloOrH168\_Contig4393 (2578146bp gc=0.38) (coordinates 1-500000)

**hint:** Mouse over a gene to see details. Query gene in red. RNAs in black. Pseudo genes in white.

**COG Coloring Selection**

Color code of function category for top COG hit is shown below. You may select a subset to view specific categories.

| Show Color                          | COG Code | COG Function Definition                                       |
|-------------------------------------|----------|---|
| <input checked="" type="checkbox"/> | [A]      | RNA processing and modification                               |
| <input checked="" type="checkbox"/> | [B]      | Chromatin structure and dynamics                              |
| <input checked="" type="checkbox"/> | [C]      | Energy production and conversion                              |
| <input checked="" type="checkbox"/> | [D]      | Cell cycle control, cell division, chromosome partitioning    |
| <input checked="" type="checkbox"/> | [E]      | Amino acid transport and metabolism                           |
| <input checked="" type="checkbox"/> | [F]      | Nucleotide transport and metabolism                           |
| <input checked="" type="checkbox"/> | [G]      | Carbohydrate transport and metabolism                         |
| <input checked="" type="checkbox"/> | [H]      | Coenzyme transport and metabolism                             |
| <input checked="" type="checkbox"/> | [I]      | Lipid transport and metabolism                                |
| <input checked="" type="checkbox"/> | [J]      | Translation, ribosomal structure and biogenesis               |
| <input checked="" type="checkbox"/> | [K]      | Transcription   |
| <input checked="" type="checkbox"/> | [L]      | Replication, recombination and repair                         |
| <input checked="" type="checkbox"/> | [M]      | Cell wall/membrane/envelope biogenesis                        |
| <input checked="" type="checkbox"/> | [N]      | Cell motility   |
| <input checked="" type="checkbox"/> | [O]      | Posttranslational modification, protein turnover, chaperones  |
| <input checked="" type="checkbox"/> | [P]      | Inorganic ion transport and metabolism                        |
| <input checked="" type="checkbox"/> | [Q]      | Secondary metabolites biosynthesis, transport and catabolism  |
| <input checked="" type="checkbox"/> | [R]      | General function prediction only                              |
| <input checked="" type="checkbox"/> | [S]      | Function unknown  |
| <input checked="" type="checkbox"/> | [T]      | Signal transduction mechanisms                                |
| <input checked="" type="checkbox"/> | [U]      | Intracellular trafficking, secretion, and vesicular transport |
| <input checked="" type="checkbox"/> | [V]      | Defense mechanisms  |
| <input checked="" type="checkbox"/> | [W]      | Extracellular structures                                      |
| <input checked="" type="checkbox"/> | [Y]      | Nuclear structure   |
| <input checked="" type="checkbox"/> | [Z]      | Cytoskeleton  |

Save Selections Select All Clear All

**hint:** Saving with no selections defaults to show all colors.

**Genome Viewers**

[Scaffolds and Contigs](#)

[Chromosome Maps](#)

[Web Artemis](#)

**Compare Gene Annotations**

[Compare Gene Annotations](#)

Gene annotation values are precomputed and stored in a tab deli also viewable in Excel.

**Export Genome Data**

Download sequences and gene information for this genome.

[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 file for Excel with gene information](#)

[Generate Genbank File](#)

**hint:** Right click on link to see menu for saving link contents. Please be patient during download.

*Advancing Science with DNA Sequence*

**Map of chromosome**

Download [publication quality image](#)

**Legend**

- From outside to the center
- Genes on forward strand (color by COG categories)
- Genes on reverse strand (color by COG categories)
- RNA genes (tRNAs green, sRNAs red)
- GC content
- GC skew

DOE JOINT GENOME INSTITUTE  
US DEPARTMENT OF ENERGY  
OFFICE OF SCIENCE

All genomes selected

Quick Genome Search:  **GO**

Hi nikos, Logout

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

[IMG Home](#) | [Find Genomes](#) | [Find Genes](#) | [Find Functions](#) | [Compare Genomes](#) | [Analysis Carts](#) | [MyIMG](#) | [Using IMG](#)

[Genome Browser](#) | [Genome Search](#) | [Taxonomy Browser](#) | [Category Browser](#)

# IMG curation

# Gene annotation curation

## Genome Viewers

- Scaffolds and Contigs
- Chromosome Maps
- Web Artemis

## Compare Gene Annotations

Compare Gene Annotations

Gene annotation views are precomputed and stored in a tab delimited file also viewable in Excel.

## Export Genome Data

Download sequences and gene information for this genome.

- [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 file for Excel with gene information](#)

Generate Genbank File

**hint:** Right click on link to see menu for saving link contents to your computer. Please be patient during download.

Compare Gene Annotations 2372 Loaded

View annotations for *Halobacterium salinarum* H168 (v3)

Select Filter:

Current filter: **No Product Name/With Evidence**

| Gene Object ID | Locus Tag | Source         | Cluster Annotation                                | Gene Annotation | E-value  |
|----------------|-----------|----------------|---|-----------------|----------|
| 23042936       | Hom_00010 | COG0593        | ATPase involved in DNA replication initiation     |                 | 2.0e-129 |
| 23042936       | Hom_00010 | pflm0303       | Bac_DnaA  |                 | 1.0e-130 |
| 23042936       | Hom_00010 | product_name   | chromosomal replication initiator protein DnaA    |                 | 0.0e+00  |
| 23042936       | Hom_00010 | DNA_length     |   | 1303bp          |          |
| 23042936       | Hom_00010 | Protein_length |   | 407aa           |          |
| 23042937       | Hom_00000 | COG0592        | DNA polymerase alpha clamp subunit (PCNA homolog) |                 | 2.0e-67  |
| 23042937       | Hom_00000 | pflm0712       | DNA_pol3_beta                                     |                 | 4.0e-17  |
| 23042937       | Hom_00000 | pflm0787       | DNA_pol3_beta_2                                   |                 | 4.0e-19  |
| 23042937       | Hom_00000 | pflm0768       | DNA_pol3_beta_3                                   |                 | 5.0e-20  |
| 23042937       | Hom_00000 | T1180663       | DNA polymerase III, beta subunit                  |                 | 2.3e-95  |
| 23042937       | Hom_00000 | product_name   | DNA polymerase III, beta subunit                  |                 |          |
| 23042937       | Hom_00000 | DNA_length     |   | 1104bp          |          |
| 23042937       | Hom_00000 | Protein_length |   | 367aa           |          |
| 23042938       | Hom_00000 | COG2601        | Uncharacterized conserved protein                 |                 | 4.0e-13  |
| 23042938       | Hom_00000 | pflm0789       | Sa  |                 | 1.0e-03  |
| 23042938       | Hom_00000 | product_name   | BNA-binding S4 domain protein                     |                 |          |
| 23042938       | Hom_00000 | DNA_length     |   | 218bp           |          |
| 23042938       | Hom_00000 | Protein_length |   | 73aa            |          |

**JGI Gene page** All genomes selected Quick Genome Search

**img/er** INTEGRATE EXPERT RE

Gene Detail

Gene Information

|                 |   |
|-----------------|---|
| Gene Object ID  | 2500344848  |
| Gene Symbol     |   |
| Locus Tag       | Hom_0001  |
| Product Name    | chromosomal replication initiator protein DnaA  |
| Gene            | <a href="#">Halobacterium salinarum H168 (Finished v3)</a>                            |
| DNA Coordinates | 593..1924 (+)X11827bp   |
| Scaffold Source | <a href="#">Halobacterium salinarum H168 (Finished v3) - haloTherm168 (2578148bp)</a> |
| IMG ORF Type    |   |
| GC Content      | 0.39  |
| External Links  |   |

Protein Information

|                            |   |
|----------------------------|---|
| Amino Acid Sequence Length | 463aa   |
| COG                        | COG0593 - ATPase involved in DNA replication initiation |
| IMG Term                   |   |
| Families                   | - T1GR00361 chromosomal replication initiator protein D |
| Transmembrane Helices      | No  |
| Signal Peptide             | No  |
| Statistics                 | <a href="#">ppp46c</a>                                  |

Evidence For Function Prediction

Neighborhood

red = Current Gene  
green = Positional Cluster Gene in the same KEGG Pathway as the Current Gene  
cyan = High scoring genes with MyIMG EC number assignments  
blue = CEPS array  
Show neighborhood regions with the same top COG hit  
Show in the interactive viewer

| COG ID  | Consensus Sequence Length | Description  | Percent Identity | Alignment On Query Gene | E-value  | Bit Score |
|---------|---------------------------|--|------------------|-------------------------|----------|-----------|
| COG0593 | 408                       | [L] <a href="#">Replication, recombination and repair</a><br>ATPase involved in DNA replication initiation | 48.42            |                         | 2.0e-129 | 456       |

Plan

| Plan Domain | COG Plan Hit | Description            | Percent Identity | Alignment On Query Gene | E-value  | Bit Score |
|-------------|--------------|------------------------|------------------|-------------------------|----------|-----------|
| Bac_DnaA    | pflm0303     | Bacterial dnaA protein | 62.97            |                         | 1.0e-130 | 459       |

External Sequence Search

NCBI BLAST  
EBI InterPro Scan  
Protein Data Bank BLAST

IMG Sequence Search

IMG Genome BLAST

Homolog Display

Customized Homolog Display

Homolog Selection

**JGI**  
DOE JOINT GENOME INSTITUTE  
AN OAK RIDGE NATIONAL LABORATORY PROJECT

All genomes selected Quick Genome Search **GO**

**img/er** INTEGRATED MICROBIAL GENOMES  
EXPERT REVIEW

IMG Home Find Genomes Find Genes Find Functions Compare Genomes Analysis Carts MyIMG About IMG

Chosen Interacts: Genes Functions Curation

**Gene Cart** 1 gene(s) in cart

Gene List  
[Upload & Export](#)  
[Companion Tools](#)  
[Chromosome Map](#)  
[Sequence Alignment](#)  
[Gene Neighborhoods](#)  
[Profile Tools](#)  
[Gene Profile](#)  
[Occurrence Profile](#)

**Gene List** 1 gene(s) in cart

Remove Selected Select All Clear All

| Selection                           | Gene Object ID | Locus Tag | Product Name                                   | AA Seq. Length | Genome                                    | Batch <sup>1</sup> |
|-------------------------------------|----------------|-----------|--|----------------|---|--------------------|
| <input checked="" type="checkbox"/> | 2500344318     | Hore_0001 | chromosomal replication initiator protein DnaA | 463aa          | Halothermothrix orenus HI68 (Finished v2) | 1                  |

<sup>1</sup> - Each time a set of genes is added to the cart, a new distinguishing batch number is generated for the set.

**Gene Term Associations**

Associate genes with IMG terms. Search for IMG terms. Enter new IMG terms or upload file with gene / term associations.

Gene Term Associations File Upload

**Enter MyIMG Annotation**

You may enter, update or delete your product name, function, EC number, PUBMED ID, etc. for the selected genes.

Annotate Selected Genes

**JGI**  
DOE JOINT GENOME INSTITUTE  
AN OAK RIDGE NATIONAL LABORATORY PROJECT

All genomes selected Quick Genome Search **GO**

**img/er** INTEGRATED MICROBIAL GENOMES  
EXPERT REVIEW

IMG Home Find Genomes Find Genes Find Functions Compare Genomes Analysis Carts MyIMG About IMG

MyIMG Home My Genomes Annotations Preferences Logout

**MyIMG Annotation for Selected Genes**

1 gene(s) selected

| Select                              | Gene Object ID | Locus Tag | Original Product Name                          | Annotated Product Name | AA Seq. Length | Genome                                    | Batch |
|-------------------------------------|----------------|-----------|--|------------------------|----------------|---|-------|
| <input checked="" type="checkbox"/> | 2500344318     | Hore_0001 | chromosomal replication initiator protein DnaA |                        | 463aa          | Halothermothrix orenus HI68 (Finished v2) | 1     |

**MyIMG Annotation**

Upload existing MyIMG annotation for gene:  Load Gene Annotation

Enter or update MyIMG annotation for selected gene(s).

- Product Name cannot be blank. Use ':' to separate multiple product names.
- EC Number can contain multiple EC numbers separated by blank or ',' (e.g., EC:1.2.3.4 EC:4.3.-).
- PUBMED ID can contain multiple ID values separated by blank or ','.

|                         |  |
|-------------------------|--|
| <b>Product Name</b>     | DnaB   |
| <b>Prot Description</b> | Any description  |
| <b>EC Number</b>        | EC:1.1.1.1   |
| <b>PUBMED ID</b>        |  |
| <b>Inference</b>        | experimental   |
| <b>Notes</b>            | This is a bogus annotation. Made up for this presentation. Applause !! |
| <b>Is Pseudo Gene?</b>  | Yes<br>No<br>Unknown   |

Update Annotation Delete Annotation Reset Cancel

**JGI** DOE JOINT GENOME INSTITUTE  
OFFICE OF SCIENCE

Curation check All genomes selected  Quick Genome Search:

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

IMG Home | Find Genomes | Find Genes | Find Functions | Compare Genomes | Analysis Carts | MyIMG | About IMG

Search | COG | Pfam | **KEGG** | Enzyme | TIGRfam | IMG Network

### KEGG Pathways

201 KEGG pathways retrieved.

**Amino Acid Metabolism**

- [Alanine and aspartate metabolism](#)
- [Arginine and proline metabolism](#)
- [Cysteine metabolism](#)
- [Glutamate metabolism](#)
- [Glycine, serine and threonine metabolism](#)
- [Histidine metabolism](#)
- [Lysine biosynthesis](#)
- [Lysine degradation](#)
- [Methionine metabolism](#)
- [Phenylalanine metabolism](#)
- [Phenylalanine, tyrosine and tryptophan biosynthesis](#) ←
- [Tryptophan metabolism](#)
- [Tyrosine metabolism](#)
- [Urea cycle and metabolism of amino groups](#)
- [Valine, leucine and isoleucine biosynthesis](#)
- [Valine, leucine and isoleucine degradation](#)

**Biosynthesis of Polyketides and Nonribosomal Peptides**

- [Biosynthesis of 12-, 14- and 16-membered macrolides](#)
- [Biosynthesis of ansamycins](#)
- [Biosynthesis of siderophore group nonribosomal peptides](#)
- [Biosynthesis of type II polyketide backbone](#)
- [Biosynthesis of type II polyketide products](#)
- [Biosynthesis of vancomycin group antibiotics](#)
- [Polyketide sugar unit biosynthesis](#)

**JGI** DOE JOINT GENOME INSTITUTE  
OFFICE OF SCIENCE

All genomes selected  Quick Genome Search:

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

IMG Home | Find Genomes | Find Genes | Find Functions | Compare Genomes | Analysis Carts | MyIMG | About IMG

Search | COG | Pfam | KEGG | Enzyme | TIGRfam | IMG Network

### KEGG Pathway Details

Loaded

Details for *Phenylalanine, tyrosine and tryptophan biosynthesis*.

**Enzymes in Pathway**

Add Selected to Function Cart | Select All | Clear All

Click on column name to sort.

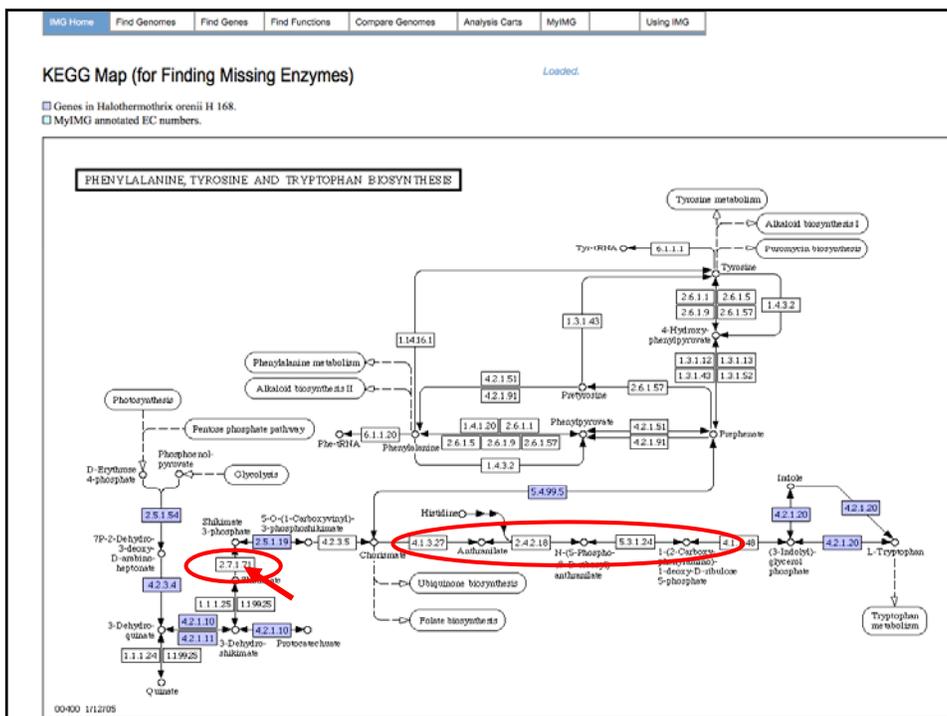
| Select                   | EC Number    | Enzyme Name                                      | Genome Count |
|--------------------------|--------------|--|--------------|
| <input type="checkbox"/> | EC:1.1.1.24  | Quinate dehydrogenase                            | 6            |
| <input type="checkbox"/> | EC:1.1.1.25  | Shikimate dehydrogenase                          | 430          |
| <input type="checkbox"/> | EC:1.1.99.25 | Quinate dehydrogenase (pyrroloquinoline-quinone) | 17           |
| <input type="checkbox"/> | EC:1.14.16.1 | Phenylalanine 4-monooxygenase                    | 155          |
| <input type="checkbox"/> | EC:1.3.1.12  | Prephenate dehydrogenase                         | 287          |
| <input type="checkbox"/> | EC:1.3.1.13  | Prephenate dehydrogenase (NADP(+))               | 19           |
| <input type="checkbox"/> | EC:1.3.1.43  | Arogenate dehydrogenase                          | 16           |
| <input type="checkbox"/> | EC:1.3.1.52  | 2-methyl-branched-chain-enoyl-CoA reductase      | 0            |

**View Pathway Map**

View KEGG map for one genome.  
Domains: (B)acteria, (A)rchaea, (E)ukarya, (P)lasmiids, (V)iruses.  
Genome Completion [F]inished, [D]raft.

- Haemophilus influenzae Rd KW20 (B)[F]
- Haemophilus somnus 129PT (B)[F]
- Haemophilus somnus 2336 (B)[D]
- Halobacterium salinarum R1 (B)[F]
- Halobacterium salinarum R2 (B)[F]
- Halobacterium salinarum R3 (B)[F]
- Halobacterium salinarum R4 (B)[F]
- Halobacterium salinarum R5 (B)[F]
- Halobacterium salinarum R6 (B)[F]
- Halobacterium salinarum R7 (B)[F]
- Halobacterium salinarum R8 (B)[F]
- Halobacterium salinarum R9 (B)[F]
- Halobacterium salinarum R10 (B)[F]
- Halobacterium salinarum R11 (B)[F]
- Halobacterium salinarum R12 (B)[F]
- Halobacterium salinarum R13 (B)[F]
- Halobacterium salinarum R14 (B)[F]
- Halobacterium salinarum R15 (B)[F]
- Halobacterium salinarum R16 (B)[F]
- Halobacterium salinarum R17 (B)[F]
- Halobacterium salinarum R18 (B)[F]
- Halobacterium salinarum R19 (B)[F]
- Halobacterium salinarum R20 (B)[F]
- Halobacterium salinarum R21 (B)[F]
- Halobacterium salinarum R22 (B)[F]
- Halobacterium salinarum R23 (B)[F]
- Halobacterium salinarum R24 (B)[F]
- Halobacterium salinarum R25 (B)[F]
- Halobacterium salinarum R26 (B)[F]
- Halobacterium salinarum R27 (B)[F]
- Halobacterium salinarum R28 (B)[F]
- Halobacterium salinarum R29 (B)[F]
- Halobacterium salinarum R30 (B)[F]
- Halobacterium salinarum R31 (B)[F]
- Halobacterium salinarum R32 (B)[F]
- Halobacterium salinarum R33 (B)[F]
- Halobacterium salinarum R34 (B)[F]
- Halobacterium salinarum R35 (B)[F]
- Halobacterium salinarum R36 (B)[F]
- Halobacterium salinarum R37 (B)[F]
- Halobacterium salinarum R38 (B)[F]
- Halobacterium salinarum R39 (B)[F]
- Halobacterium salinarum R40 (B)[F]
- Halobacterium salinarum R41 (B)[F]
- Halobacterium salinarum R42 (B)[F]
- Halobacterium salinarum R43 (B)[F]
- Halobacterium salinarum R44 (B)[F]
- Halobacterium salinarum R45 (B)[F]
- Halobacterium salinarum R46 (B)[F]
- Halobacterium salinarum R47 (B)[F]
- Halobacterium salinarum R48 (B)[F]
- Halobacterium salinarum R49 (B)[F]
- Halobacterium salinarum R50 (B)[F]
- Halobacterium salinarum R51 (B)[F]
- Halobacterium salinarum R52 (B)[F]
- Halobacterium salinarum R53 (B)[F]
- Halobacterium salinarum R54 (B)[F]
- Halobacterium salinarum R55 (B)[F]
- Halobacterium salinarum R56 (B)[F]
- Halobacterium salinarum R57 (B)[F]
- Halobacterium salinarum R58 (B)[F]
- Halobacterium salinarum R59 (B)[F]
- Halobacterium salinarum R60 (B)[F]
- Halobacterium salinarum R61 (B)[F]
- Halobacterium salinarum R62 (B)[F]
- Halobacterium salinarum R63 (B)[F]
- Halobacterium salinarum R64 (B)[F]
- Halobacterium salinarum R65 (B)[F]
- Halobacterium salinarum R66 (B)[F]
- Halobacterium salinarum R67 (B)[F]
- Halobacterium salinarum R68 (B)[F]
- Halobacterium salinarum R69 (B)[F]
- Halobacterium salinarum R70 (B)[F]
- Halobacterium salinarum R71 (B)[F]
- Halobacterium salinarum R72 (B)[F]
- Halobacterium salinarum R73 (B)[F]
- Halobacterium salinarum R74 (B)[F]
- Halobacterium salinarum R75 (B)[F]
- Halobacterium salinarum R76 (B)[F]
- Halobacterium salinarum R77 (B)[F]
- Halobacterium salinarum R78 (B)[F]
- Halobacterium salinarum R79 (B)[F]
- Halobacterium salinarum R80 (B)[F]
- Halobacterium salinarum R81 (B)[F]
- Halobacterium salinarum R82 (B)[F]
- Halobacterium salinarum R83 (B)[F]
- Halobacterium salinarum R84 (B)[F]
- Halobacterium salinarum R85 (B)[F]
- Halobacterium salinarum R86 (B)[F]
- Halobacterium salinarum R87 (B)[F]
- Halobacterium salinarum R88 (B)[F]
- Halobacterium salinarum R89 (B)[F]
- Halobacterium salinarum R90 (B)[F]
- Halobacterium salinarum R91 (B)[F]
- Halobacterium salinarum R92 (B)[F]
- Halobacterium salinarum R93 (B)[F]
- Halobacterium salinarum R94 (B)[F]
- Halobacterium salinarum R95 (B)[F]
- Halobacterium salinarum R96 (B)[F]
- Halobacterium salinarum R97 (B)[F]
- Halobacterium salinarum R98 (B)[F]
- Halobacterium salinarum R99 (B)[F]
- Halobacterium salinarum R100 (B)[F]

EC equivalents and genes in selected genome.  
 Relative abundance of genes in selected genome.  
 Find missing enzymes.



**Find Candidate Genes for Missing Function**

Genome: *Halothermothrix orenii* H 168  
 Bacteria | Firmicutes | Clostridia | Halanarcales | Halanarcales | *Halothermothrix orenii*

Function: (EC:2.7.1.71) Shikimate kinase

Using Homology or Orthology  
 Using PRIAM  
 Using Both

Go

Using Homology or Orthology  
 This tool allows you to find genes associated with enzymes homologous from the query genome (*Halothermothrix orenii*) has homology to other genomes associated with EC:2.7.1.71. These homologies have a reciprocal hit in the query genome.

**Candidate Genes for Missing Function**

Genome: *Halothermothrix orenii* H 168  
 Function: (EC:2.7.1.71) Shikimate kinase.

Domains(D): B-Bacteria, A-Archaea, E-Eukarya, P-Plasmid, V-Viruses.  
 Genome Completion(C): P-Finished, D-Draft.  
 Click on column name to sort.

| Select                   | Candidate Gene | Candidate Gene Product | Enzyme for Candidate Gene | Homolog Gene     | Homolog Gene Product (IMG Term) | Enzyme for Homolog Gene | D | G                                   | Genome | Percent Ident | Alignment On Candidate | Alignment On Homolog | Evalue | Bit Score | Confirmed by PRIAM? | PRIAM Percent Ident | PRIAM Alignment On Candidate | PRIAM Bit Score |
|--------------------------|----------------|------------------------|---------------------------|------------------|---------------------------------|-------------------------|---|-------------------------------------|--------|---------------|------------------------|----------------------|--------|-----------|---------------------|---------------------|------------------------------|-----------------|
| <input type="checkbox"/> | 639072658      | Shikimate kinase       | 637262811                 | Shikimate kinase | EC:2.7.1.71                     | B                       | F | <i>Halothermothrix orenii</i> H 168 | 46.25  | 100%          | 100%                   | 5.00e-30             | 131    | Yes       | 43.37               | 100%-41             | 162                          |                 |
| <input type="checkbox"/> | 639071182      | Homocysteine kinase    |                           |                  |                                 |                         |   |                                     |        |               |                        |                      |        | Yes       | 21.15               | 4.00e-13            | 68                           |                 |

Add to MyIMG Annotation    Select All    Clear All

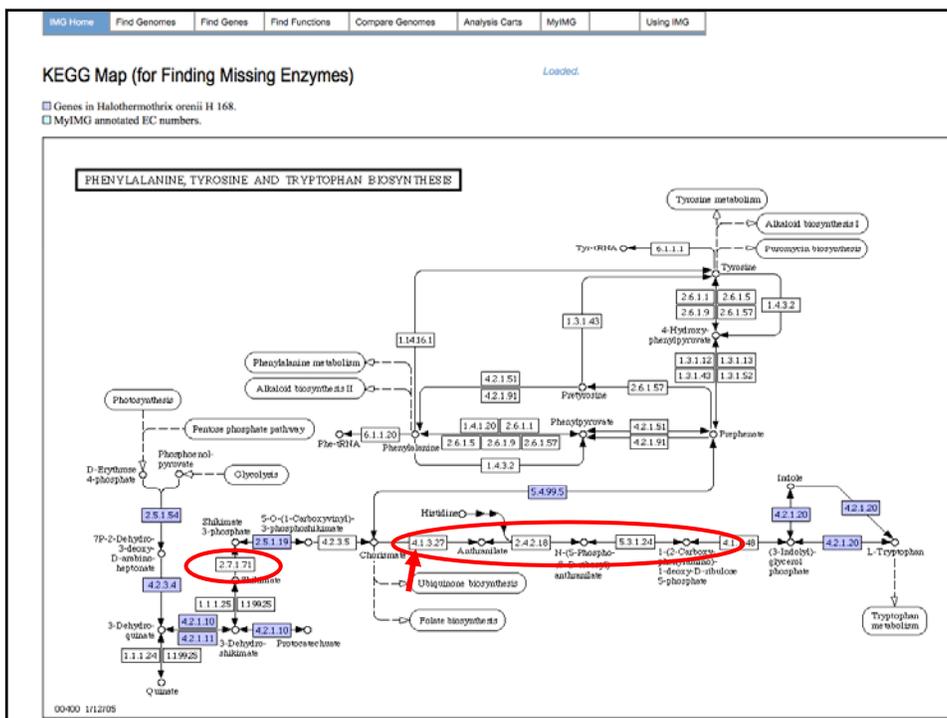
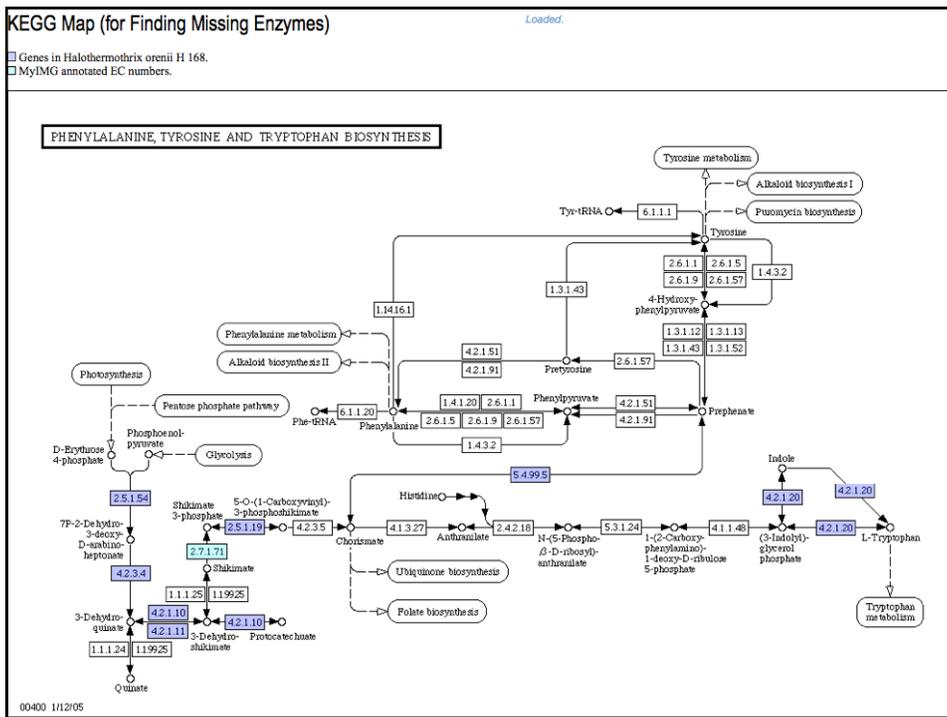
---

**Add Enzyme to Candidate Gene(s) in MyIMG Annotation**

| Select                              | Gene Object ID | Gene Display Name | Genome                              | Old MyIMG Enzyme(s) | New MyIMG Enzyme | Add/Replace   |
|-------------------------------------|----------------|-------------------|-------------------------------------|---------------------|------------------|---|
| <input checked="" type="checkbox"/> | 639072658      | Shikimate kinase  | <i>Halothermothrix orenii</i> H 168 |                     | EC:2.7.1.71      | <input checked="" type="radio"/> Add<br><input type="radio"/> Replace |

Note: Click the 'Update MyIMG Annotation' button will save your change(s) to the database. Only selected gene-enzyme associations will be added to MyIMG annotation.

Update MyIMG Annotation    Reset



MG Home | Find Genomes | Find Genes | Find Functions | Compare Genomes | Analysis Carts | MyMG | Using MG

Home | Genes | Functions | Custom

### Find Candidate Genes for Missing Function

Genome: **Halothermothrix orenii H 168**

Bacteria; Firmicutes; Clostridia; Halantracrobiales; Halantracrobaceae; Halothermothrix orenii

Function: (EC:4.1.3.27) Anthranilate synthase.

Using Homologs or Orthologs  
 Using PRAM  
 Using Both

Go

---

**Using Homologs or Orthologs**

This tool allows you to find genes associated with enzymes through homologs in other genomes. Homologs from the query genome *Halothermothrix orenii H 168* has homologs in other genomes associated with (EC:4.1.3.27). These homologs have a reciprocal hit in the query genome, which are listed as candidates for gene-enzyme association.

**Database Search Options:**

Currently selected genomes (list)  
 Whole database (slow)  
 Bacteria  
 Firmicutes  
 Clostridia  
 Halantracrobiales  
 Halantracrobaceae

**Restrictions:**

Orthologs (N-directional best hits) (fast)  
 Homologs (slow for whole database)

**Percent Identity Cutoff:** 10

**E-value cutoff:** 1e-2

**Maximum Homology:** 100

---

**Using PRAM**

**Percent Identity Cutoff:** 10

**E-value cutoff:** 1e-30

**MS score cutoff:** 0

**Percent Alignment Cutoff:** 10

MG Home | Find Genomes | Find Genes | Find Functions | Compare Genomes | Analysis Carts

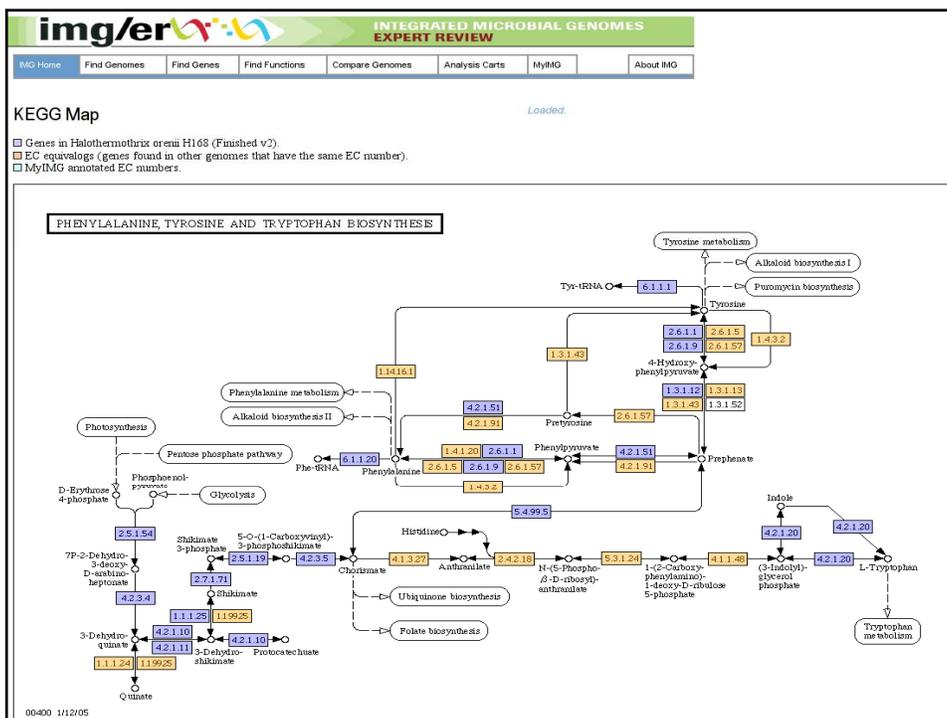
Home | Genes | Functions | Custom

### Candidate Genes for Missing Function

Genome: **Halothermothrix orenii H 168**

Function: (EC:4.1.3.27) Anthranilate synthase.

No candidates found.



**JGI** Curation check All genomes selected Quick Genome Search GO

**img/er** INTEGRATED MICROBIAL EXPERT REVIEW

[IMG Home](#) | [Find Genomes](#) | [Find Genes](#) | [Find Functions](#) | [Compare Genomes](#) | [Analysis Carts](#)  
[Search](#) | [COG](#) | [in](#) | [HOG](#) | [Enzyme](#) | [TIGRfam](#) | [IMG Network](#)

### COG Browser

**COG list**

- Amino acid transport and metabolism [E]**
  - Arginine biosynthesis
  - Glutamine biosynthesis
  - Isoleucine biosynthesis
  - Leucine biosynthesis
  - Methionine biosynthesis
  - Phenylalanine/tyrosine biosynthesis
  - Proline biosynthesis
  - Threonine biosynthesis
  - Tryptophan biosynthesis
  - Valine biosynthesis
- Carbohydrate transport and metabolism [G]**
  - Emmer-Doudoroff pathway
  - Glycolysis
  - Pentose phosphate pathway
- Secondary metabolites biosynthesis, transport and catabolism [Q]**
- Signal transduction mechanisms [T]**
- Transcription [K]**
  - Basal transcription factors
  - RNA-dependent RNA polymerase subunits
  - Transcriptional regulators
- Translation, ribosomal structure and biogenesis [J]**
  - Aminocyl-RNA synthetases and alternate systems for amino acid activation
  - Ribosomal proteins - large subunit
  - Ribosomal proteins - small subunit
  - Translation factors and enzymes involved in translation

**COG Table:**

| Select                   | COG ID  | COG Name   | Genome Count |
|--------------------------|---------|--|--------------|
| <input type="checkbox"/> | COG0048 | Ribosomal protein S12  | 852          |
| <input type="checkbox"/> | COG0049 | Ribosomal protein S7   | 862          |
| <input type="checkbox"/> | COG0051 | Ribosomal protein S10  | 846          |
| <input type="checkbox"/> | COG0052 | Ribosomal protein S2   | 882          |
| <input type="checkbox"/> | COG0092 | Ribosomal protein S3   | 865          |
| <input type="checkbox"/> | COG0098 | Ribosomal protein S8   | 872          |
| <input type="checkbox"/> | COG0099 | Ribosomal protein S5   | 877          |
| <input type="checkbox"/> | COG0099 | Ribosomal protein S13  | 873          |
| <input type="checkbox"/> | COG0100 | Ribosomal protein S9   | 876          |
| <input type="checkbox"/> | COG0103 | Ribosomal protein S9   | 875          |
| <input type="checkbox"/> | COG0184 | Ribosomal protein S19P/S13E  | 873          |
| <input type="checkbox"/> | COG0185 | Ribosomal protein S19  | 838          |
| <input type="checkbox"/> | COG0186 | Ribosomal protein S17  | 849          |
| <input type="checkbox"/> | COG0192 | Ribosomal protein S14  | 822          |
| <input type="checkbox"/> | COG0228 | Ribosomal protein S16  | 781          |
| <input type="checkbox"/> | COG0238 | Ribosomal protein S18  | 792          |
| <input type="checkbox"/> | COG0268 | Ribosomal protein S20  | 740          |
| <input type="checkbox"/> | COG0360 | Ribosomal protein S6   | 808          |
| <input type="checkbox"/> | COG0522 | Ribosomal protein S4 and related proteins                            | 883          |
| <input type="checkbox"/> | COG0539 | Ribosomal protein S11  | 796          |
| <input type="checkbox"/> | COG0828 | Ribosomal protein S21  | 752          |
| <input type="checkbox"/> | COG1099 | Predicted RNA binding protein (contains ribosomal protein S1 domain) | 142          |
| <input type="checkbox"/> | COG1358 | Ribosomal protein H56-type (S12L30L7a)                               | 242          |
| <input type="checkbox"/> | COG1471 | Ribosomal protein S4E  | 84           |
| <input type="checkbox"/> | COG1890 | Ribosomal protein S3AE   | 85           |
| <input type="checkbox"/> | COG1998 | Ribosomal protein S7AE   | 84           |
| <input type="checkbox"/> | COG2004 | Ribosomal protein S2AE   | 84           |
| <input type="checkbox"/> | COG2007 | Ribosomal protein S8E  | 84           |
| <input type="checkbox"/> | COG2051 | Ribosomal protein S7E  | 82           |
| <input type="checkbox"/> | COG2053 | Ribosomal protein S2E/S33  | 90           |
| <input type="checkbox"/> | COG2125 | Ribosomal protein S6E (S10)  | 85           |

### Function Cart

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

#### Upload Function Cart from File

Upload function cart from a tab-delimited file.

#### Export Functions

Function information may be exported for reloading through Upload Function Cart later.

#### Function Profile

View selected function(s) against selected genomes. Please select 1 to 50 genome(s).

Domains: (B)acteria, (A)rchaea, (E)ukarya, (P)lasmids, (V)iruses  
 Genome Completion: [F]inished, [D]raft

Haemophilus somnus 129P1 (B)[F]  
 Haemophilus somnus 2396 (B)[D]  
 Haella chujuensis KCTC 2396 (B)[F]  
 Halothelodospira halophila S11 (B)[F]  
 Halothelodospira halophila S12 (B)[D]  
 Halothelodospira halophila S13 (B)[F]  
 Halothelodospira halophila S14 (B)[F]  
 Halothelodospira halophila S15 (B)[F]  
 Halothelodospira halophila S16 (B)[F]  
 Halothelodospira halophila S17 (B)[F]  
 Halothelodospira halophila S18 (B)[F]  
 Halothelodospira halophila S19 (B)[F]  
 Halothelodospira halophila S20 (B)[F]  
 Halothelodospira halophila S21 (B)[F]  
 Halothelodospira halophila S22 (B)[F]  
 Halothelodospira halophila S23 (B)[F]  
 Halothelodospira halophila S24 (B)[F]  
 Halothelodospira halophila S25 (B)[F]  
 Halothelodospira halophila S26 (B)[F]  
 Halothelodospira halophila S27 (B)[F]  
 Halothelodospira halophila S28 (B)[F]  
 Halothelodospira halophila S29 (B)[F]  
 Halothelodospira halophila S30 (B)[F]  
 Halothelodospira halophila S31 (B)[F]  
 Halothelodospira halophila S32 (B)[F]  
 Halothelodospira halophila S33 (B)[F]  
 Halothelodospira halophila S34 (B)[F]  
 Halothelodospira halophila S35 (B)[F]  
 Halothelodospira halophila S36 (B)[F]  
 Halothelodospira halophila S37 (B)[F]  
 Halothelodospira halophila S38 (B)[F]  
 Halothelodospira halophila S39 (B)[F]  
 Halothelodospira halophila S40 (B)[F]  
 Halothelodospira halophila S41 (B)[F]  
 Halothelodospira halophila S42 (B)[F]  
 Halothelodospira halophila S43 (B)[F]  
 Halothelodospira halophila S44 (B)[F]  
 Halothelodospira halophila S45 (B)[F]  
 Halothelodospira halophila S46 (B)[F]  
 Halothelodospira halophila S47 (B)[F]  
 Halothelodospira halophila S48 (B)[F]  
 Halothelodospira halophila S49 (B)[F]  
 Halothelodospira halophila S50 (B)[F]

Max. E-value: 0.1  
 Min. Percent Identity: 10

#### Function Profile

hint: Mouse over genome abbreviation to see genome name.  
 (Cell coloring is highlighting of gene counts: white = 0, blue = 1-4, yellow = 5)  
 Click on column name to sort

| function ID | Name   | Clu | Hal | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
|             |  | ACE | BD  | NEW |
|             |  | 824 | 352 | NT1 | 134 | SM1 | EB8 | Z75 | V21 |     |
| COG0048     | Ribosomal protein S12  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0049     | Ribosomal protein S7   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0051     | Ribosomal protein S10  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0052     | Ribosomal protein S2   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0092     | Ribosomal protein S3   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0098     | Ribosomal protein S5   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0099     | Ribosomal protein S13  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0100     | Ribosomal protein S9   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0103     | Ribosomal protein S9   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 2   |
| COG0184     | Ribosomal protein S19P/S13E  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0185     | Ribosomal protein S19  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0186     | Ribosomal protein S17  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0192     | Ribosomal protein S14  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0228     | Ribosomal protein S16  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0238     | Ribosomal protein S18  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0268     | Ribosomal protein S20  | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0360     | Ribosomal protein S6   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG0522     | Ribosomal protein S4 and related proteins                            | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 2   |
| COG0539     | Ribosomal protein S11  | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 2   |
| COG0828     | Ribosomal protein S21  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| COG1099     | Predicted RNA binding protein (contains ribosomal protein S1 domain) | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   | 1   |
| COG1358     | Ribosomal protein H56-type (S12L30L7a)                               | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 2   | 2   |
| COG1471     | Ribosomal protein S4E  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| COG1890     | Ribosomal protein S3AE   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| COG1998     | Ribosomal protein S7AE   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| COG2004     | Ribosomal protein S2AE   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| COG2007     | Ribosomal protein S8E  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| COG2051     | Ribosomal protein S7E  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| COG2053     | Ribosomal protein S2E/S33  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| COG2125     | Ribosomal protein S6E (S10)  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |

hint: - Hold down control key (or command key in the case of the Mac) to select multiple genomes.  
 - Drag down list to select all genomes.  
 - More genome and function selections result in slower query.

**JGIS**  
DOE JOINT GENOME INSTITUTE  
AT BROTHERHOOD UNIVERSITY

Quick Genome Search: All genomes selected

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

IMG Home Find Genomes Find Genes Find Functions Compare Genomes Analysis Carts MyIMG About IMG

Current Selection: Genes [ Functions ] Curation

**Profile Genes**

Add Selected to Gene Cart

2500345021 ribosomal prote...  
2500345120 ribosomal prote...

**Gene Cart**

**Gene Neighborhoods** Loaded

Neighborhoods of selected genes in cart.  
Genes of the same color (except light yellow) are from the same orthologous group (top COG hit).  
Light yellow = no COG assignment.

hint: Mouse over a gene to see details (once page has loaded).

Halothermothrix orenii H168 (Finished v2) | haloTherm168v2\_haloOrH168\_Contig1393

Halothermothrix orenii H168 (Finished v2) | haloTherm168v2\_haloOrH168\_Contig1393

**JGIS**  
DOE JOINT GENOME INSTITUTE  
AT BROTHERHOOD UNIVERSITY

Quick Genome Search: All genomes selected

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

IMG Home Find Genomes Find Genes Find Functions Compare Genomes Analysis Carts MyIMG About IMG

Current Selection: [ Genes ] Functions Curation

**Sequence Alignments**

Protein  
DNA  bp upstream

Color Alignments with Mview

**Do Alignment**

**Gene Neighborhoods**

5'-3' direction of each selected gene is  
5'-3' direction of plus strand is always

Show Neighborhoods

**ClustalW Alignment** Loaded

Identities computed with respect to: (1) 2500345021  
Colored by: consensus/80% and property

```

1 2500345021  ATGGCGGAGAGGTAACAATCTGGGTACCGGTAGAGTAAATTTAACTGCCAGGGTCTGTTTAGTACCCGGTAGTGGTAAGATTTTAAATAATGGAAGAGATGTTA
2 2500345120  ATGGCGGAGAGGTAACAATCTGGGTACCGGTAGAGTAAATTTAACTGCCAGGGTCTGTTTAGTACCCGGTAGTGGTAAGATTTTAAATAATGGAAGAGATGTTA
    
```

Mview 1.47.3, Copyright (c) Nigel P. Brown 1997-2002

| Gene Object ID | Product Name         | Genome                                    | Scaffold ID                          |
|----------------|----------------------|---|--------------------------------------|
| 2500345021     | ribosomal protein S9 | Halothermothrix orenii H168 (Finished v2) | haloTherm168v2_haloOrH168_Contig1393 |
| 2500345120     | ribosomal protein S9 | Halothermothrix orenii H168 (Finished v2) | haloTherm168v2_haloOrH168_Contig1393 |

**Jalview alignment editor**

File Edit Font View Colour Calculate Align Help

JalView button

JalView

```

2500345021/1-396 190 200 210 220 230 240 250 260 270 280
2500345120/1-396 190 200 210 220 230 240 250 260 270 280
    
```

Version 2.4.0  
Questions/Con...  
©2007 The Re...  
Disclaimer

**JGI**  
DOE JOINT GENOME INSTITUTE  
AN OFFICE OF SCIENCE

All genomes selected  Quick Genome Search:  **GO**

Hi nikos, Logout

**img/er** INTEGRATED MICROBIAL GENOMES  
EXPERT REVIEW

IMG Home Find Genomes Find Genes Find Functions Compare Genomes Analysis Carts MyIMG Using IMG

Genome Browser | Genome Search | Taxonomy Browser | Category Browser

# IMG directed searches

## Gene searches

*Microbiology* (2002), 148, 2343–2349

Printed in Great Britain

### Cloning, sequencing and expression of an $\alpha$ -amylase gene, *amyA*, from the thermophilic halophile *Halothermothrix orenii* and purification and biochemical characterization of the recombinant enzyme

Benjamin N. Mijts and Bharat K. C. Patel

Author for correspondence: Bharat Patel. Tel: +61 438 173 185. Fax: +61 7 3875 7800.  
e-mail: bharat@genomes.sci.gu.edu.au

Microbial Discovery Research  
Unit, School of Biomolecular  
and Biomedical Sciences,  
Faculty of Science, Griffith  
University, Brisbane,  
Queensland 4111, Australia

A recombinant clone expressing an amylase was identified from an *Escherichia coli* generated genomic library of the thermophilic, moderately halophilic, anaerobic bacterium *Halothermothrix orenii* by activity screening, and the gene encoding the enzyme was designated AmyA. The *amyA* gene was 1545 bp long, and encoded a 515 residue protein composed of a 25 amino acid putative signal peptide and a 490 amino acid mature protein. It possessed the five consensus regions characteristic of the  $\alpha$ -amylase family and showed the greatest homology to the *Bacillus megaterium* group of  $\alpha$ -amylases. The *amyA* gene was expressed in *E. coli* as a hexahistidine-tagged enzyme and purified. The purified recombinant enzyme was optimally active at 65 °C in 5% (w/v) NaCl at pH 7.5, with significant activity retained in the presence of up to 25% (w/v) NaCl. It had a specific activity of 2232 U mg<sup>-1</sup> and required NaCl and CaCl<sub>2</sub> for optimum activity and thermostability. The relatively high proportion of acidic amino acids typically observed for many enzymes from halophiles was absent in *H. orenii* AmyA.

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

Gene Search | BLAST | Phylogenetic Profiler

Find genes in selected genomes by keyword

Keyword:  Filters:

Assume current genome selections or select from Domains: (B)acteria, (A)rchaea, (E)ukarya, (P)lan... Genome Completion: (F)inished, (D)raft.

Go

**Gene Product Name Results** 9 gene(s) retrieved.

Search by 'Product Name' with expression 'amylase'

Add Selected to Gene Cart Select All Clear All

| Selection                | Gene Object ID | Match Text                                     | Genome                                    |
|--------------------------|----------------|--|---|
| <input type="checkbox"/> | 2500345114     | alpha amylase ( EC:3.2.1.98 )                  | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500345844     | alpha amylase catalytic region ( EC:3.2.1.10 ) | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500345868     | alpha amylase ( EC:3.2.1.1 )                   | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500346385     | alpha amylase ( EC:3.2.1.10 )                  | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500347253     | alpha amylase ( EC:3.2.1.1 )                   | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500346742     | alpha amylase ( EC:3.2.1.98 )                  | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500346791     | alpha amylase ( EC:3.2.1.98 )                  | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500347186     | alpha amylase catalytic region ( EC:3.2.1.10 ) | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500346534     | alpha amylase catalytic region ( EC:3.2.1.54 ) | Halothermothrix orenii H168 (Finished v2) |

**img/er** INTEGRATED MICROBIAL GENOMES EXPERT REVIEW

Search | COG | Pfam | KEGG | Enzyme

**Function Search Results** 7 functions retrieved.

The number of genes is shown in parentheses.

Add Selected to Function Cart Select All Clear All

pfam00128 - Alpha-amylase - Alpha amylase, catalytic domain. Alpha amylase is classified as family 13 of the glycosyl hydrolases. The structure is an 8 stranded alpha/beta barrel containing the active site, interrupted by a ~70 aa. calcium-binding domain protruding between beta strand 3 and alpha helix 3, and a carboxyl-terminal Greek key beta-barrel domain (1).

hint: The function

**Genes In pfam00128** 12 genes retrieved.

Add Selected to Gene Cart Select All Clear All

|                          |            |   |   |
|--------------------------|------------|---|---|
| <input type="checkbox"/> | 2500345114 | alpha amylase (EC:3.2.1.98)                     | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500345579 | 1,4-alpha-glucan branching enzyme (EC:2.4.1.18) | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500345844 | alpha amylase catalytic region (EC:3.2.1.10)    | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500345868 | alpha amylase (EC:3.2.1.1)                      | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500346292 | pullulanase, type 1 (EC:3.2.1.41)               | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500345835 | alpha amylase (EC:3.2.1.10)                     | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500346534 | alpha amylase catalytic region (EC:3.2.1.54)    | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500346742 | alpha amylase (EC:3.2.1.98)                     | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500346791 | alpha amylase (EC:3.2.1.98)                     | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500346725 | pullulanase, type 1 (EC:3.2.1.41)               | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500347186 | alpha amylase catalytic region (EC:3.2.1.10)    | Halothermothrix orenii H168 (Finished v2) |
| <input type="checkbox"/> | 2500347253 | alpha amylase (EC:3.2.1.1)                      | Halothermothrix orenii H168 (Finished v2) |

IMG Home Find Genomes Find Genes Find Functions Compare Genomes **Analyze Clust** MyIMG About IMG

Current Selection: [ Genes ] Functions Clustion

### Gene Cart

12 gene(s) in cart

Gene List  
[Upload & Export](#)  
[Comparison Tools](#)  
[Chromosome Map](#)  
[Sequence Alignment](#)  
[Gene Neighborhoods](#)  
[Profile Tools](#)  
[Gene Profile](#)  
[Document Profile](#)

### Gene List

12 gene(s) in cart

Remove Selected Select All Clear All

| Selection                           | Gene Object ID | LOCUS Tag | Product Name                                    | AA Seq. Length | Genome                                    | Batch <sup>1</sup> |
|-------------------------------------|----------------|-----------|---|----------------|---|--------------------|
| <input checked="" type="checkbox"/> | 2500345114     | Hore_0241 | alpha amylase (EC:3.2.1.98)                     | 364aa          | Halothermothrix orenii H168 (Finished v2) | 2                  |
| <input checked="" type="checkbox"/> | 2500345868     | Hore_0970 | alpha amylase (EC:3.2.1.1)                      | 426aa          | Halothermothrix orenii H168 (Finished v2) | 2                  |
| <input checked="" type="checkbox"/> | 2500346791     | Hore_1873 | alpha amylase (EC:3.2.1.98)                     | 442aa          | Halothermothrix orenii H168 (Finished v2) | 2                  |
| <input checked="" type="checkbox"/> | 2500346385     | Hore_1474 | alpha amylase (EC:3.2.1.10)                     | 515aa          | Halothermothrix orenii H168 (Finished v2) | 2                  |
| <input checked="" type="checkbox"/> | 2500347186     | Hore_2253 | alpha amylase catalytic region (EC:3.2.1.10)    | 562aa          | Halothermothrix orenii H168 (Finished v2) | 2                  |
| <input checked="" type="checkbox"/> | 2500345844     | Hore_0946 | alpha amylase catalytic region (EC:3.2.1.10)    | 563aa          | Halothermothrix orenii H168 (Finished v2) | 2                  |
| <input checked="" type="checkbox"/> | 2500346534     | Hore_1616 | alpha amylase catalytic region (EC:3.2.1.54)    | 582aa          | Halothermothrix orenii H168 (Finished v2) | 2                  |
| <input checked="" type="checkbox"/> | 2500346742     | Hore_1824 | alpha amylase (EC:3.2.1.98)                     | 623aa          | Halothermothrix orenii H168 (Finished v2) | 2                  |
| <input type="checkbox"/>            | 2500345579     | Hore_0693 | 1,4-alpha-glucan branching enzyme (EC:2.4.1.18) | 630aa          | Halothermothrix orenii H168 (Finished v2) | 1                  |
| <input type="checkbox"/>            | 2500346297     | Hore_1388 | pullulanase, type I (EC:3.2.1.41)               | 640aa          | Halothermothrix orenii H168 (Finished v2) | 1                  |
| <input checked="" type="checkbox"/> | 2500347262     | Hore_2320 | alpha amylase (EC:3.2.1.1)                      | 654aa          | Halothermothrix orenii H168 (Finished v2) | 2                  |

**Sequence Alignments**

Protein  DNA [-0 bp upstream, +0 bp downstream]

Color Alignments with Mview

←

img/er **INTEGRATED MICROBIAL GENOMES**  
EXPERT REVIEW

IMG Home Find Genomes Find Genes Find Functions Compare Genomes **Analyze Clust** MyIMG About IMG

Current Selection: [ Genes ] Functions Clustion

### ClustalW Alignment

Identities computed with respect to: (1) 2500346385  
 Colored by: consensus/80% and property

1 2500346385  
 2 2500347253  
 3 2500347186  
 4 2500345844  
 5 2500346297  
 6 2500346875  
 7 2500345114  
 8 2500346791  
 9 2500346742  
 10 2500345868  
 11 2500346534  
 12 2500345579

consensus/100%  
 consensus/90%  
 consensus/80%  
 consensus/70%

Mview 1.47.3, Copyright (c) Nigel P. Brown 1997-2002

JalView button may take a moment to load.

Neighbour Joining tree using PID

Font size: 12 Show distances Close Output

2500346385/1-1000  
 2500347253/1-1000  
 2500347186/1-1000  
 2500345844/1-1000  
 2500346297/1-1000  
 2500346875/1-1000  
 2500345114/1-1000  
 2500346791/1-1000  
 2500346742/1-1000  
 2500345868/1-1000  
 2500346534/1-1000  
 2500345579/1-1000

Quality/1-1000

The screenshot shows the top navigation bar of the IMG/ER website. It includes the JGI logo, a search bar with 'All genomes selected', and a 'GO' button. Below the navigation bar is a green banner with the text 'img/er' and 'INTEGRATED MICROBIAL GENOMES EXPERT REVIEW'. A secondary navigation bar contains links like 'IMG Home', 'Find Genomes', 'Find Genes', 'Find Functions', 'Compare Genomes', 'Analysis Carts', 'MyIMG', and 'Using IMG'. The main content area is a large white space with the text 'IMG discovery' centered in a large, black, sans-serif font.

The screenshot displays the 'Profile' page in the IMG/ER interface. On the left, there is a 'Phylogenetic Profiler' section with a description and a red arrow pointing towards the main table. The main part of the page is a table with columns for 'Find Genes In\*', 'With Homologs In', 'Without Homologs In', 'Ignoring', and 'Taxon Name'. The table lists various bacterial species, including *Caldicellulosiruptor saccharolyticus*, *Thermosinus*, *Haloethermothrix*, *Thermoanaerobacter*, *Proteobacteria*, *Idiomarina*, *Thermotogae*, *Fervidobacterium*, *Thermosipho*, and *Thermotoga*. Below the table is a 'Similarity Cutoffs' section with several dropdown menus for 'Max. E-value', 'Min. Percent Identity', 'Algorithm', 'Min. Taxon Percent With Homologs', and 'Min. Taxon Percent Without Homologs'. At the bottom, there are 'Go' and 'Reset' buttons.

| Find Genes In*           | With Homologs In         | Without Homologs In      | Ignoring                 | Taxon Name  |
|--------------------------|--------------------------|--------------------------|--------------------------|---|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Caldicellulosiruptor saccharolyticus DSM 8903</a> [F] |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <b>Thermosinus</b>  |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Thermosinus carboxydovorans Nor1</a> [D]              |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <b>Haloethermothrix</b>   |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Haloethermothrix oreum H 168</a> [D]                  |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <b>Thermoanaerobacter</b>   |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Thermoanaerobacter ethanolicus ATCC 33223</a> [D]     |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Thermoanaerobacter ethanolicus X514</a> [D]           |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Thermoanaerobacter tengcongensis MB4</a> [F]          |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <b>Proteobacteria</b>   |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <b>Idiomarina</b>   |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Idiomarina ballica OS145</a> [D]                      |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <b>Thermotogae</b>  |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <b>Fervidobacterium</b>   |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Fervidobacterium nodosum Rt17-B1</a> [D]              |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <b>Thermosipho</b>  |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Thermosipho melanesiensis B1420</a> [D]               |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <b>Thermotoga</b>   |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Thermotoga maritima MSB8</a> [F]                      |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <a href="#">Thermotoga petrophila RJU-1</a> [F]                   |

### Genes with homologs in Gram + (Firmicutes) absent in Gram -

| Select                   | Result Row | Gene Object ID | Locus Tag | Gene Name  | Length | COG     | Enzyme     | Pfam      | InterPro |
|--------------------------|------------|----------------|-----------|--|--------|---------|------------|-----------|----------|
| <input type="checkbox"/> | 1          | 2500344857     | Hore_0010 | GerA spore germination protein                   | 531aa  | -       | -          | pfam03323 | -        |
| <input type="checkbox"/> | 2          | 2500344953     | Hore_0085 | UvrB/UvrC protein                                | 170aa  | COG3880 | -          | pfam02151 | -        |
| <input type="checkbox"/> | 3          | 2500344954     | Hore_0086 | Arginine kinase                                  | 349aa  | COG3869 | EC:2.7.3.3 | pfam00217 | -        |
| <input type="checkbox"/> | 4          | 2500345390     | Hore_0514 | SCP-like extracellular                           | 258aa  | COG2340 | -          | -         | -        |
| <input type="checkbox"/> | 5          | 2500345430     | Hore_0553 | cell envelope-related transcriptional attenuator | 252aa  | COG1316 | -          | pfam03816 | -        |
| <input type="checkbox"/> | 6          | 2500345474     | Hore_0589 | transcriptional regulator, AbrB family           | 80aa   | COG2002 | -          | pfam04014 | -        |
| <input type="checkbox"/> | 7          | 2500345521     | Hore_0636 | transcriptional regulator, AbrB family           | 86aa   | COG2002 | -          | pfam04014 | -        |
| <input type="checkbox"/> | 8          | 2500345586     | Hore_0700 | anti-sigma-factor antagonist spoIIAA             | 120aa  | COG1366 | -          | pfam01740 | -        |
| <input type="checkbox"/> | 9          | 2500345592     | Hore_0706 | Stage V sporulation AD protein                   | 343aa  | -       | -          | pfam07451 | -        |
| <input type="checkbox"/> | 10         | 2500345593     | Hore_0707 | Sporulation stage V protein AE                   | 117aa  | -       | -          | pfam03862 | -        |
| <input type="checkbox"/> | 11         | 2500346169     | Hore_1261 | putative stage IV sporulation YqD                | 380aa  | -       | -          | pfam06898 | -        |
| <input type="checkbox"/> | 12         | 2500346248     | Hore_1339 | cell envelope-related transcriptional attenuator | 405aa  | COG1316 | -          | pfam03816 | -        |
| <input type="checkbox"/> | 13         | 2500346249     | Hore_1340 | metal dependent phosphohydrolase                 | 189aa  | COG1713 | -          | pfam01966 | -        |
| <input type="checkbox"/> | 14         | 2500346504     | Hore_1586 | protein of unknown function DUF199               | 313aa  | COG1481 | -          | pfam02650 | -        |
| <input type="checkbox"/> | 15         | 2500346716     | Hore_1798 | protein of unknown function DUF1385              | 293aa  | COG3872 | -          | pfam07136 | -        |
| <input type="checkbox"/> | 16         | 2500346956     | Hore_2035 | PSP1 domain protein                              | 288aa  | COG1774 | -          | pfam04468 | -        |
| <input type="checkbox"/> | 17         | 2500347050     | Hore_2119 | transcriptional regulator, AbrB family           | 185aa  | COG2002 | -          | pfam04014 | -        |

**Unknown:**  
sporulation not observed in *Halothermothrix*

|                                     |            |            |                      |      |                                   |   |
|-------------------------------------|------------|------------|----------------------|------|-----------------------------------|---|
| <input checked="" type="checkbox"/> | 2500423182 | Hore_11980 | hypothetical protein | 97aa | Halothermothrix orenii H 168 (v3) | 3 |
| <input checked="" type="checkbox"/> | 2500423503 | Hore_15080 | hypothetical protein | 97aa | Halothermothrix orenii H 168 (v3) | 2 |
| <input checked="" type="checkbox"/> | 2500423360 | Hore_13790 | hypothetical protein |      | Halothermothrix orenii H 168 (v3) |   |

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

#### Gene Term Associations

Associate genes with IMG terms. Search for IMG terms. Enter new IMG terms or upload file with gene / term associations.

Enter MyIMG Annotation

You may enter, update or delete your product name, function, EC number, PUBMED ID, etc. for the set.

Upload Gene Cart from File

Upload gene cart from a tab-delimited file.

Export Genes

Select genes to export and one of the following export options.

FASTA Amino Acid format.

FASTA Nucleic Acid format.  bp upstream.  bp downstream

Gene information in tab delimited format to Excel.

Chromosome Map

View selected genes against the entire chromosome.

INTEGRATED MICROBIAL GENOMES  
EXPERT REVIEW

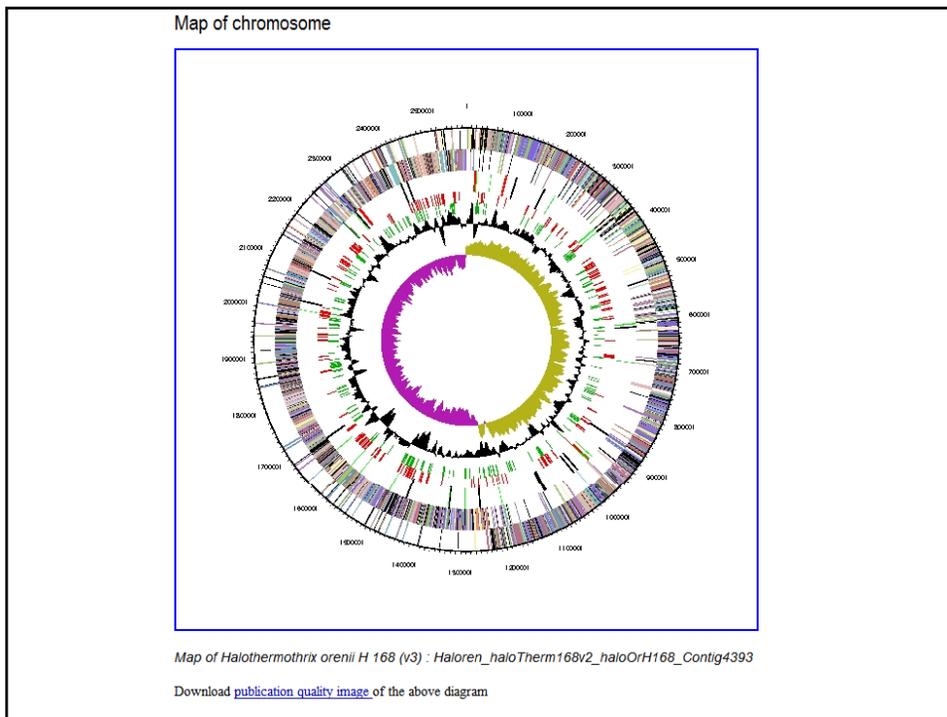
[IMG Home](#)
[Find Genomes](#)
[Find Genes](#)
[Find Functions](#)
[Compare Genomes](#)
[Analysis Tools](#)
[MyIMG](#)
[Using IMG](#)

[Cancel](#)
[Tables](#)
[Functions](#)
[Queries](#)

#### Chromosome Map

| Gene Object ID | Product Name   | Start   | End     | Strand | Locus tag  | Scaffold                          | Batch | Band 1                              | Band 2                              | Band 3                   | Band 4                   |
|----------------|--|---------|---------|--------|------------|-----------------------------------|-------|-------------------------------------|-------------------------------------|--------------------------|--------------------------|
| 2500423971     | hypothetical protein   | 28875   | 29462   | +      | Hore_00200 | Halothermothrix orenii H 168 (v3) | 3     | <input type="checkbox"/>            | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500423974     | hypothetical protein   | 30613   | 30894   | +      | Hore_00200 | Halothermothrix orenii H 168 (v3) | 3     | <input type="checkbox"/>            | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500423976     | hypothetical protein   | 31940   | 32761   | +      | Hore_00250 | Halothermothrix orenii H 168 (v3) | 3     | <input type="checkbox"/>            | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500423977     | hypothetical protein   | 32974   | 33846   | +      | Hore_00260 | Halothermothrix orenii H 168 (v3) | 2     | <input checked="" type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500423978     | ribonucleic acid binding protein containing TIR-like domain                      | 34102   | 34926   | +      | Hore_00270 | Halothermothrix orenii H 168 (v3) | 3     | <input type="checkbox"/>            | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500423982     | Predicted nucleic acid binding protein, contains PCY domain                      | 37748   | 38152   | -      | Hore_00310 | Halothermothrix orenii H 168 (v3) | 3     | <input type="checkbox"/>            | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500423983     | hypothetical protein   | 38149   | 38397   | -      | Hore_00320 | Halothermothrix orenii H 168 (v3) | 3     | <input type="checkbox"/>            | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500423980     | extracellular solute-binding protein family 1                                    | 42408   | 43676   | +      | Hore_00380 | Halothermothrix orenii H 168 (v3) | 2     | <input checked="" type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500426340     | binding protein-dependent peroxidase component                                   | -       | -       | -      | -          | Halothermothrix orenii H 168 (v3) | -     | <input type="checkbox"/>            | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500426340     | ABC-type transport system, involved in lipopeptide release, peroxidase component | 2544887 | 2546146 | -      | Hore_23290 | Halothermothrix orenii H 168 (v3) | 2     | <input checked="" type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500426341     | ABC-type transport system, involved in lipopeptide release, peroxidase component | 2546158 | 2547393 | -      | Hore_23270 | Halothermothrix orenii H 168 (v3) | 2     | <input checked="" type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="checkbox"/> |
| 2500426354     | TPP repeat-containing protein  | 2537726 | 2538217 | +      | Hore_23400 | Halothermothrix orenii H 168 (v3) | 2     | <input checked="" type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="checkbox"/> |

18



### Search for *Halothermothrix* genes, absent in Gram + (Firmicutes)

Add Selected to Gene Cart    Select All    Clear All

| Select                   | Result Row | Gene Object ID             | Locus Tag | Gene Name   | Length | COG     | Enzyme     | Pfam                   | InterPro |
|--------------------------|------------|----------------------------|-----------|---|--------|---------|------------|------------------------|----------|
| <input type="checkbox"/> | 1          | <a href="#">2500346572</a> | Hore_1654 | flagellar P-ring protein  | 383aa  | COG1706 | -          | pfam02119              | -        |
| <input type="checkbox"/> | 2          | <a href="#">2500346573</a> | Hore_1655 | flagellar L-ring protein  | 192aa  | COG2063 | -          | pfam02107              | -        |
| <input type="checkbox"/> | 3          | <a href="#">2500346673</a> | Hore_1657 | UDP-3-O-(3-hydroxytristatyl)-N-acetylglucosamine deacetylase    | 388aa  | COG0774 | EC:3.5.1.- | pfam03331              | -        |
| <input type="checkbox"/> | 4          | <a href="#">2500346808</a> | Hore_1889 | Three-deoxy-D-manno-octulosonic-acid transferase domain protein | 779aa  | COG1519 | -          | pfam04413<br>pfam01075 | -        |

**LIPOLYSACCHARIDE BIOSYNTHESIS**

**known:** from cell envelope characterization, consistent with Gram – ultrastructure

**RESULT:** *Halothermothrix* genes, absent in Gram +ves (Firmicutes) but present in Gram -ves

# Abundance comparison

**Abundance Profile Tools**

The following tools operate on functional profiles of multiple genomes.

| Tool  | Description   |
|---|---|
| <a href="#">Abundance Profile Viewer</a>      | View selected functional profiles for selected genomes.                       |
| <a href="#">Abundance Profile Search</a>      | Search for functions based on over or under abundance in one or more genomes. |
| <a href="#">Function Category Comparisons</a> | Fairwise functional category abundance comparisons.                           |
| <a href="#">Abundance Toolkit</a>             | Customize and export abundance measurements.                                  |

- Highest density of carbohydrate ABC transporters except for members of the Rhizobiaceae, 3% the genome
- The phosphotransferase system appears to be in the process of being removed from the genome.
  - Enzyme I (Hore\_20150) has a frameshift that does not occur at a slippery site,
  - IIA subunit
  - IIC subunit
  - NO IIB subunit

Adaptation to low nutrient environment

# Horizontally transferred genes.

**Chromosome Viewer**

Switch coloring for GC content

Horizontally transferred genes (HGT) are shown in black.

GC Coloring

| Show Color                          | Color      | Description                      |
|-------------------------------------|------------|----------------------------------|
| <input checked="" type="checkbox"/> | [ +20% ]   | cgc = 20%                        |
| <input checked="" type="checkbox"/> | [ -10% ]   | cgc = 10%                        |
| <input checked="" type="checkbox"/> | [ +5% ]    | cgc = 5%                         |
| <input checked="" type="checkbox"/> | [ -2% ]    | cgc = 2%                         |
| <input checked="" type="checkbox"/> | [ +2 -2% ] | 38% characteristic GCs (cgc) ±2% |
| <input checked="" type="checkbox"/> | [ -2% ]    | cgc = 2%                         |
| <input checked="" type="checkbox"/> | [ -5% ]    | cgc = 5%                         |
| <input checked="" type="checkbox"/> | [ -10% ]   | cgc = 10%                        |
| <input checked="" type="checkbox"/> | [ -20% ]   | cgc = 20%                        |