



SHORT GENOME REPORT

Open Access

Draft genome sequence of *Halomonas lutea* strain YIM 91125^T (DSM 23508^T) isolated from the alkaline Lake Ebinur in Northwest China

Xiao-Yang Gao^{1,11}, Xiao-Yang Zhi², Hong-Wei Li^{2,9}, Yu Zhou¹⁰, Alla Lapidus^{3,4}, James Han⁵, Matthew Haynes⁵, Elizabeth Lobos⁵, Marcel Huntemann⁵, Amrita Pati⁵, Natalia N Ivanova⁵, Konstantinos Mavromatis⁵, Brian J Tindall⁷, Victor Markowitz⁶, Tanja Woyke⁵, Hans-Peter Klenk^{7,12}, Nikos C Kyrpides^{5,8} and Wen-Jun Li^{1,2*}

Abstract

Species of the genus *Halomonas* are halophilic and their flexible adaption to changes of salinity and temperature brings considerable potential biotechnology applications, such as degradation of organic pollutants and enzyme production. The type strain *Halomonas lutea* YIM 91125^T was isolated from a hypersaline lake in China. The genome of strain YIM 91125^T becomes the twelfth species sequenced in *Halomonas*, and the thirteenth species sequenced in *Halomonadaceae*. We described the features of *H. lutea* YIM 91125^T, together with the high quality draft genome sequence and annotation of its type strain. The 4,533,090 bp long genome of strain YIM 91125^T with its 4,284 protein-coding and 84 RNA genes is a part of Genomic Encyclopedia of Type Strains, Phase I: the one thousand microbial genomes (KMG-I) project. From the viewpoint of comparative genomics, *H. lutea* has a larger genome size and more specific genes, which indicated acquisition of function bringing better adaption to its environment. DDH analysis demonstrated that *H. lutea* is a distinctive species, and halophilic features and nitrogen metabolism related genes were discovered in its genome.

Keywords: *Halomonas lutea*, Aerobic, Gram-negative, Chemoorganotrophic, Moderately halophilic, Lake Ebinur

Introduction

Strain YIM 91125^T (= DSM 23508^T = KCTC 12847^T = CCTCC AB 206093^T) is the type strain of *Halomonas lutea* [1]. Currently, there are 83 validly named species in the genus *Halomonas* on the basis of most recent released from LPSN [2] and EzTaxon-e [3]. *Halomonadaceae* comprises the largest number of halophilic and halotolerant bacteria described to date, and *Halomonas* is the largest genus in this family. However, most of the taxa in *Halomonadaceae* have been reclassified in the past due to their heterogeneous features [4-7]. In *Halomonas*, a small group of species has been formally re-located to

Chromohalobacter, *Cobetia* and *Kushneria* by further taxonomic studies. Members of the genus *Halomonas* were usually isolated from saline environments [8-12]. Strain YIM 91125^T was originally isolated from soil sample of Ebinur Lake, which has been a long-term target for the studies of element cycling and microbial biota under extremely high-saline conditions in Xinjiang, Northwest China. As a type strain, it's the original isolate used in species description, which exhibits the relevant phenotypic and genotypic properties cited in the original published taxonomic circumscriptions [13]. This organism grows well across a wide range of salinity and temperature and also participates in nitrogen reduction. In this context, strain YIM 91125^T has been sequenced as a halophilic representative, and becomes a part of Genomic Encyclopedia of Type Strains, Phase I: the one thousand microbial genomes project.

Here, we present a summary classification and a set of features for *H. lutea* strain YIM 91125^T, together with the description of the genomic sequencing and annotation,

* Correspondence: liwenjun@ms.xjb.ac.cn

¹Key Laboratory of Biogeography and Bioresource in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China

²Key Laboratory of Microbial Diversity in Southwest China, Ministry of Education and the Laboratory for Conservation and Utilization of Bio-Resources, Yunnan Institute of Microbiology, Yunnan University, Kunming, China

Full list of author information is available at the end of the article

and provide brief findings of its genome sequence as compared to genomes of other *Halomonas* species. The genomic data will provide insights into its new biotechnological applications, such as sewage treatment. The comprehensive genomes of this genus will facilitate our understanding of the ecological roles that *Halomonas* species play in those hypersaline habitats and their relationships with other halophilic and nonhalophilic microorganisms.

Classification and features

H. lutea YIM 91125^T is a Gram-negative-staining, motile, aerobic and moderately halophilic bacterium, which can reduce nitrate (Table 1). Cells of the strain are short rods, 0.4 to 0.7 μm in diameter and 0.6 to 1.0 μm in length (Figure 1). They are motile by means of single polar flagellum and their colonies are orange, flat, opaque and mucoid with slightly irregular edges (Figure 1). The predominant respiratory quinone found in *H. lutea* YIM 91125^T is Q-9, similar to other members of the genus *Halomonas*. The predominant fatty acids are C_{18:1} ω7c (25.1%), C_{16:0} (17.0%), C_{19:0} cyclo ω8c (13.6%), C_{12:0} 3-OH (10.7%), C_{12:0} (7.9%), C_{10:0} (6.0%) and C_{17:0} cyclo (4.6%) [1]. The profile of major fatty acids in strain YIM 91125^T is also similar to other members of the genus *Halomonas* [14-17].

16S rRNA gene sequence of strain YIM 91125^T was compared with the newly released database from the Greengenes [26], using NCBI BLAST [27,28] under default settings (e.g., considering only HSPs from the best 250 hits) and the relative frequencies of taxa were determined, weighted by BLAST scores. The most frequently occurring genera were *Halomonas* (71.4%), *Chromohalobacter* (17.8%), *Bacillus* (3.6%), *Haererehalobacter* (3.6%) and *Modicisalibacter* (3.6%) (228 hits in total). Regarding 186 hits to sequences from members of the genus *Halomonas*, the average identity within HSPs was 95.5%, whereas the average coverage by HSPs was 98.3%. Among all other species, the one yielding the highest score was *Halomonas xinjiangensis*, which corresponded to identity of 99.9% and HSP coverage of 98.0%. (Note that the Greengenes database uses the INSDC (=EMBL/NCBI/DDBJ) annotation, which is not an authoritative source for nomenclature or classification.) The highest scoring environmental sequences were EF157249 and EF157230 (Greengenes short name 'tar pits clone 101-11 k' and 'tar pits clone 101-120 k'), which showed identity of 96.3% and an HSP coverage of 99.6%. The most frequently occurring keywords within the labels of all environmental samples which yielded hits were soil like 'soil', 'seafloor', 'drilling deep-earth'; water like 'groundwater', 'aquatic', 'lake', 'marine'; oil and plant. Environmental samples yielded hits of a higher score than the highest scoring species were not found.

Table 1 Classification and general features of *H. lutea* YIM 91125^T [18]

MIGS ID	Property	Term	Evidence code ^a
	Classification	Domain <i>Bacteria</i>	TAS [19]
		Phylum <i>Proteobacteria</i>	TAS [20]
		Class <i>Gammaproteobacteria</i>	TAS [21,22]
		Order <i>Oceanospirillales</i>	TAS [21,23]
		Family <i>Halomonadaceae</i>	TAS [4]
		Genus <i>Halomonas</i>	TAS [24]
		Species <i>Halomonas lutea</i>	TAS [1]
		Type strain YIM 91125 ^T	TAS [1]
		Gram stain	negative
		Cell shape	short rods
		Motility	motile
		Sporulation	non-sporulating
		Temperature range	4-45°C
		Optimum temperature	37°C
		pH range; Optimum	5.0-9.0
		Carbon source	mono- and polysaccharides
MIGS-6	Habitat	aquatic, fresh water, lake, salinewater	TAS [1]
MIGS-6.3	Salinity	1-20% NaCl (w/v)	TAS [1]
MIGS-22	Oxygen requirement	aerobe	TAS [1]
MIGS-15	Biotic relationship	free living	TAS [1]
MIGS-14	Pathogenicity	none	NAS
MIGS-4	Geographic location	Ebinur Lake (China)	TAS [1]
MIGS-5	Sample collection	2008 or before	NAS
MIGS-4.1	Latitude	45.05	TAS [1]
MIGS-4.2	Longitude	82.977	TAS [1]
MIGS-4.4	Altitude	not reported	

^aEvidence codes – TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [25].

Phylogenetic analyses were carried out with two different algorithms, i.e., neighbor-joining (NJ) and maximum-likelihood (ML). The phylogenetic tree was shown in Figure 2 and Additional file 1: Figure S1, which provides an interesting insight into the nomenclature and classification of members of the genus *Halomonas*, and also indicates the phylogenetic neighborhood of *H. lutea*. The phylogenetic relationships indicate that *H. lutea* YIM 91125^T is most closely to *H. xianhensis* A-1^T with 99% similarity and the sequence of the sole 16S rRNA gene in the genome differs by 10 nucleotides from the previously published 16S rRNA sequence (EF674852).

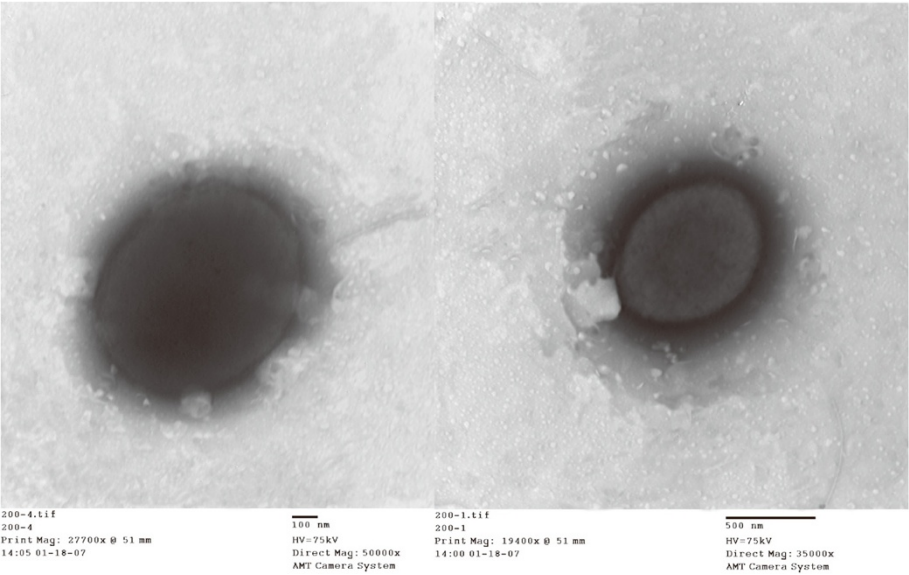


Figure 1 Transmission electron micrograph of *H. lutea* YIM 91125^T.

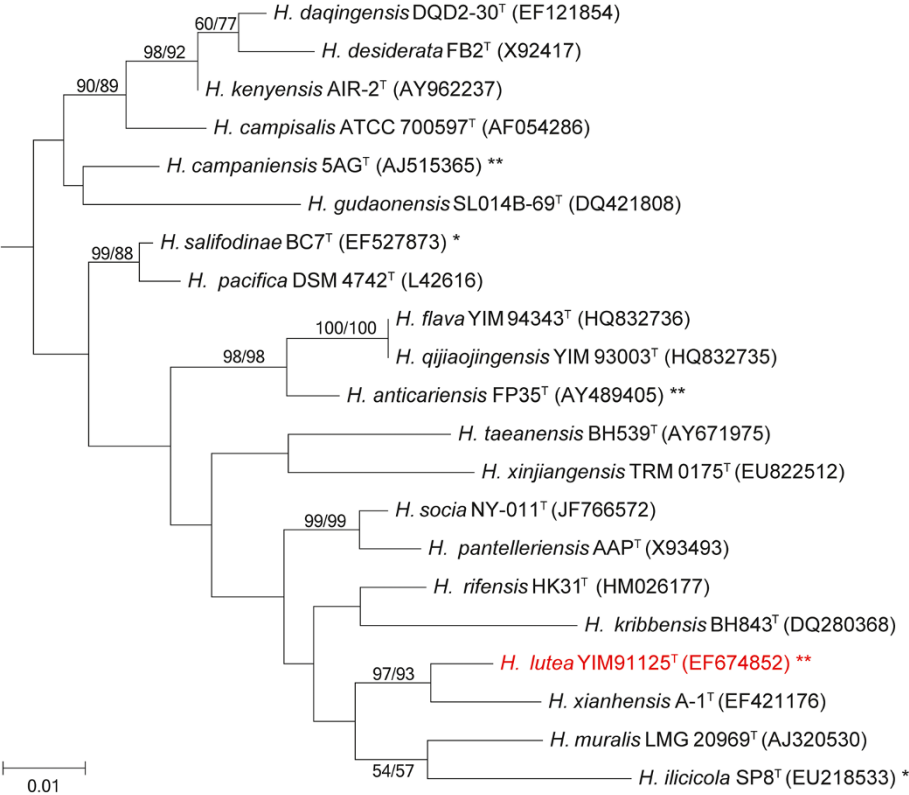


Figure 2 Phylogenetic tree highlighting the position of *H. lutea* relative to the type strains of the other species within *Halomonas*. According to the most recent release of the EzTaxon-e database, all the 16S rRNA gene sequences of the type strains within genus *Halomonas* were retained. The tree was inferred from 1,383 aligned bases [29] under the neighbor-joining (NJ) [30] and maximum-likelihood (ML) [31] methods with 1,000 randomly selected bootstrap replicates using MEGA version 5.2 [32]. The branches are scaled in terms of the expected number of substitutions per site. Numbers adjacent to the branches are support values from 1,000 NJ bootstrap (left) and from 1,000 ML bootstrap (right) replicates [33] if they are larger than 50%. Lineages with type strain genome sequencing projects registered in Genomes OnLine Database (GOLD) [34] are labeled with one asterisk, and those have available genomic data are labeled with two asterisks. Non-type strain LS21 of *H. campani*ensis and *H. elongata* DSM 2581^T listed ‘Complete and Published’ are also labeled with two asterisks.

Genome sequencing and annotation

Genome project history

This organism was selected for sequencing on the basis of its phylogenetic position and biological application importance [35,36], and for a better understand the mechanism of its halophilic adaptation. Sequencing of *H. lutea* YIM 91125^T is part of Genomic Encyclopedia of Type Strains, Phase I: the one thousand microbial genomes (KMG-I) project [37], a follow-up of the GEBA project [38], which aims for increasing the sequencing coverage of key reference microbial genomes. The genome project is deposited in the Genomes OnLine Database (GOLD), and the high quality draft genome sequence is deposited in GenBank. Sequencing, finishing and annotation were performed by the DOE JGI using state of the art sequencing technology [39]. A summary of the project information is shown in Table 2. It presents the project information and in compliance with MIGS version 2.0 compliance [18].

Growth conditions and DNA isolation

H. lutea strain YIM 91125^T (DSM 23508^T), was grown in DSMZ medium 514b (Medium 514 plus additional salt) at 37°C [40]. DNA was isolated from 0.5-1.0 g of cell pasted using Jetflex Genomic DNA Purification Kit (Qiagen, Hilden, Germany), following the standard protocol as recommended by the manufacturer, but with an additional incubation (60 min, 37°C) with 50 µl proteinase K and finally adding 200 µl protein precipitation buffer (PPT). DNA is available through the DNA Bank Network [41].

Table 2 Project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	Improved-High-Quality Draft
MIGS-28	Libraries used	Illumina standard shotgun library
MIGS-29	Sequencing platforms	Illumina HiSeq 2000
MIGS-31.2	Fold coverage	119 ×
MIGS-30	Assemblers	Velvet v. 1.1.04; ALLPATHS v. r41043
MIGS-32	Gene calling method	Prodigal 1.4
	Locus Tag	NZ_ARKK01000000
	Genbank ID	ARKK00000000
	Genbank Date of Release	April 23, 2013
	GOLD ID	Gi11553
	BIOPROJECT	PRJNA199405
	Project relevance	Genomic Encyclopedia of Type Strains, Phase I: the one thousand microbial genomes (KMG-I) project
MIGS-13	Source Material Identifier	<i>Halomonas lutea</i> DSM 23508

Genome sequencing and assembly

The draft genome of strain YIM 91125^T was generated at JGI using Illumina technology [42]. An Illumina standard shotgun library was constructed and sequenced using the Illumina HiSeq 2000 platform which generated 9,251,032 reads totaling 1,387.7 Mb. All general aspects of library construction and sequencing performed at the JGI. All raw Illumina sequence data was passed through DUK, a filtering program developed at JGI, which removes known Illumina sequencing and library preparation artifacts. The following steps were then performed for assembly: (1) filtered Illumina reads were assembled using Velvet version 1.1.04 [43]; (2) 1–3 Kb simulated paired end reads were created from Velvet contigs using Wgsim [44]; (3) Illumina reads were assembled with simulated read pairs using Allpaths-LG [45]. The final draft assembly contained 49 contigs in 42 scaffolds. The total size of the genome is 4.5 Mbp and the final assembly is based on 538.9 Mbp of Illumina data, which provides an average 119.0 × coverage of the genome.

Genome annotation

Genes were identified using Prodigal [46] as part of the DOE JGI genome annotation pipeline [47], following by a round of manual curation using the JGI GenePRIMP pipeline [48]. The predicted CDSs were translated and used to search the NCBI non-redundant database, UniProt, TIGR-Fam, Pfam, PRIAM, KEGG, COG, and InterPro database. These data sources were combined to assert a product description for each predicted protein. Additional gene prediction analysis and functional annotation were performed within the Integrated Microbial Genomes-Expert Review (IMG-ER) platform [49].

Genome properties

The assembly of the draft genome sequence consists of 42 scaffolds (Figure 3) amounting to 4,533,090 bp, and G+C content is 59.1%. The majority of the protein-coding genes (83.0%) were assigned a putative function while the remaining ones were annotated as hypothetical proteins. 3,325 protein coding genes belong to 422 paralogous families in this genome. The properties and the statistics of the genome are summarized in Table 3. The distribution of genes into COGs functional categories is presented in Table 4.

Insights from the genome sequence

The genomic sequences of twelve *Halomonas* species are available, including *H. lutea* YIM 91125^T. Genome properties of those *Halomonas* species are shown in Table 5, but only *H. elongate* and *H. campaniensis* have complete genome sequences. These *Halomonas* genome sequences exhibit dramatic interspecies variations in size, ranging from 5.34 Mb (*H. titanicae*) to 2.85 Mb

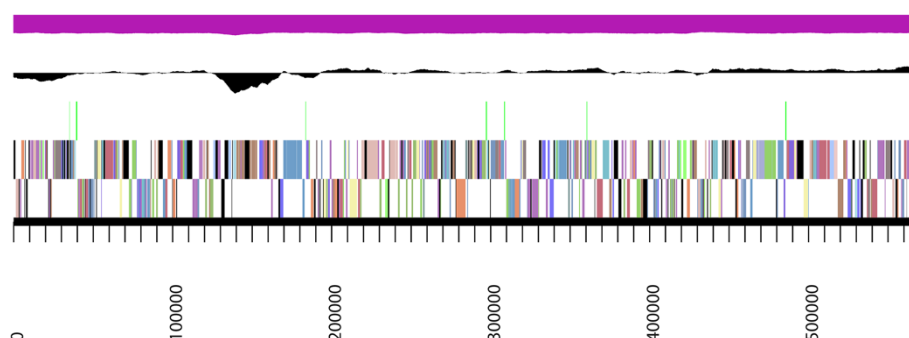


Figure 3 Graphical map of the largest scaffold in *Halomonas lutea* YIM 91125^T. From bottom to the top: Genes on forward strand (colored by COG categories), Genes on reverse strand (colored by COG categories), RNA genes (tRNA green, rRNA red, other RNAs black), GC content, GC skew (purple/olive).

(*H. jeotgali*); and the size of *H. lutea* is larger than the average size, suggesting acquisition of functions may allow better adaption to its environment, e.g., genes coding for tripartite ATP-independent periplasmic (TRAP) transporters for substrate uptake or nitrate degradation [50]. Also, GC contents of those species vary from 52.65% (*H. campaniensis*) to 67.86% (*H. smyrnensis*), and of *H. lutea* (59.05%) is around the average GC content, close to *H. anticariensis* (58.54%). In addition, the distribution of genes into COG categories was not entirely similar in all twelve compared genomes (Figure 4). And *H. lutea* has more specific genes, since proteins with COG only account for 71.18% which is lower than other members. Compared with other *Halomonas* species, the proportions of genes with signal peptide and transmembrane helices of *H. lutea* are respectively 7.46% and 23.65%, close to the corresponding averages. The abundance of transmembrane

Table 4 Number of genes associated with general COG functional categories

Code	Value	% age	Description
J	183	4.66	Translation, ribosomal structure and biogenesis
A	1	0.03	RNA processing and modification
K	278	7.08	Transcription
L	168	4.28	Replication, recombination and repair
B	6	0.15	Chromatin structure and dynamics
D	37	0.94	Cell cycle control, Cell division, chromosome partitioning
V	36	0.92	Defense mechanisms
T	208	5.30	Signal transduction mechanisms
M	210	5.35	Cell wall/membrane biogenesis
N	92	2.34	Cell motility
U	80	2.04	Intracellular trafficking and secretion
O	158	4.02	Posttranslational modification, protein turnover, chaperones
C	291	7.41	Energy production and conversion
G	273	6.95	Carbohydrate transport and metabolism
E	352	8.96	Amino acid transport and metabolism
F	85	2.16	Nucleotide transport and metabolism
H	186	4.74	Coenzyme transport and metabolism
I	133	3.39	Lipid transport and metabolism
P	218	5.55	Inorganic ion transport and metabolism
Q	126	3.21	Secondary metabolites biosynthesis, transport and catabolism
R	467	11.89	General function prediction only
S	339	8.63	Function unknown
-	871	19.94	Not in COGs

The total is based on the total number of protein-coding genes in the annotated genome.

Table 3 Genome statistics

Attribute	Value
Genome size (bp)	4,533,090
DNA coding (bp)	3,982,279
DNA G + C (bp)	2,676,712
DNA scaffolds	42
Total genes	4,368
Protein-coding genes	4,284
RNA genes	84
Pseudo genes	51
Genes in internal clusters	3,325
Genes with function prediction	3,625
Genes assigned to COGs	3,497
Genes with Pfam domains	3,674
Genes with signal peptides	326
Genes with transmembrane helices	1,033
CRISPR repeats	1

Table 5 Comparison of genome features of Halomonas species

Species	Genome size (Mb)	GC content (%)	Gene count
<i>H. anticariensis</i> FP35 ^T	5.07	58.54	4817
<i>H. boliviensis</i> LC1 ^T	4.14	54.68	3915
<i>H. campaniensis</i> LS21	4.07	52.65	3665
<i>H. elongata</i> DSM 2581 ^T	4.06	63.61	3556
<i>H. halocynthiae</i> DSM 14573 ^T	2.88	53.80	2773
<i>H. halodenitrificans</i> DSM 735 ^T	3.47	63.95	3256
<i>H. jeotgali</i> Hwa ^T	2.85	62.92	2636
<i>H. lutea</i> YIM 91125 ^T	4.53	59.05	4368
<i>H. smyrnensis</i> AAD6 ^T	3.56	67.86	3326
<i>H. stevensii</i> S18214 ^T	3.69	60.25	3523
<i>H. titanicae</i> BH1 ^T	5.34	54.58	2908
<i>H. zhanjiangensis</i> DSM 21076 ^T	4.06	54.48	3739

helices related genes indicates the important role in metabolism process of *Halomonas*.

DNA-DNA hybridization is considered as a gold-standard of distinguishing species [51]. Digital DDH similarities between genome of *H. lutea* and those of other *Halomonas* species were calculated using GGDC web server version 2.0 under recommend setting [52,53]. The probabilities of DDH value > 70% assessed via logistic regression under three formulae indicate that *H. lutea* is different from other species of the genus (Table 6). The inter-genome distances under formula 2 between *H. lutea* and *H. anticariensis*, *H. boliviensi*, *H. campaniensis*, *H. elongata*, *H. halocynthiae*, *H. halodenitrificans*, *H. jeotgali*, *H. smyrnensis*, *H. stevensii*, *H. titanicae* and *H. zhanjiangensis* are about 0.22, the corresponding DDH estimates below the 70% threshold under formula 2 are: 19.5% (± 2.29), 20.2% (± 2.31), 21.1% (± 2.33), 20.1% (± 2.31), 19.2% (± 2.29), 19.4% (± 2.29), 19.9% (± 2.30), 20.3% (± 2.32), 20.4% (± 2.32), 20.5% (± 2.32), 18.9% (± 2.28), respectively. The standard deviations indicate the

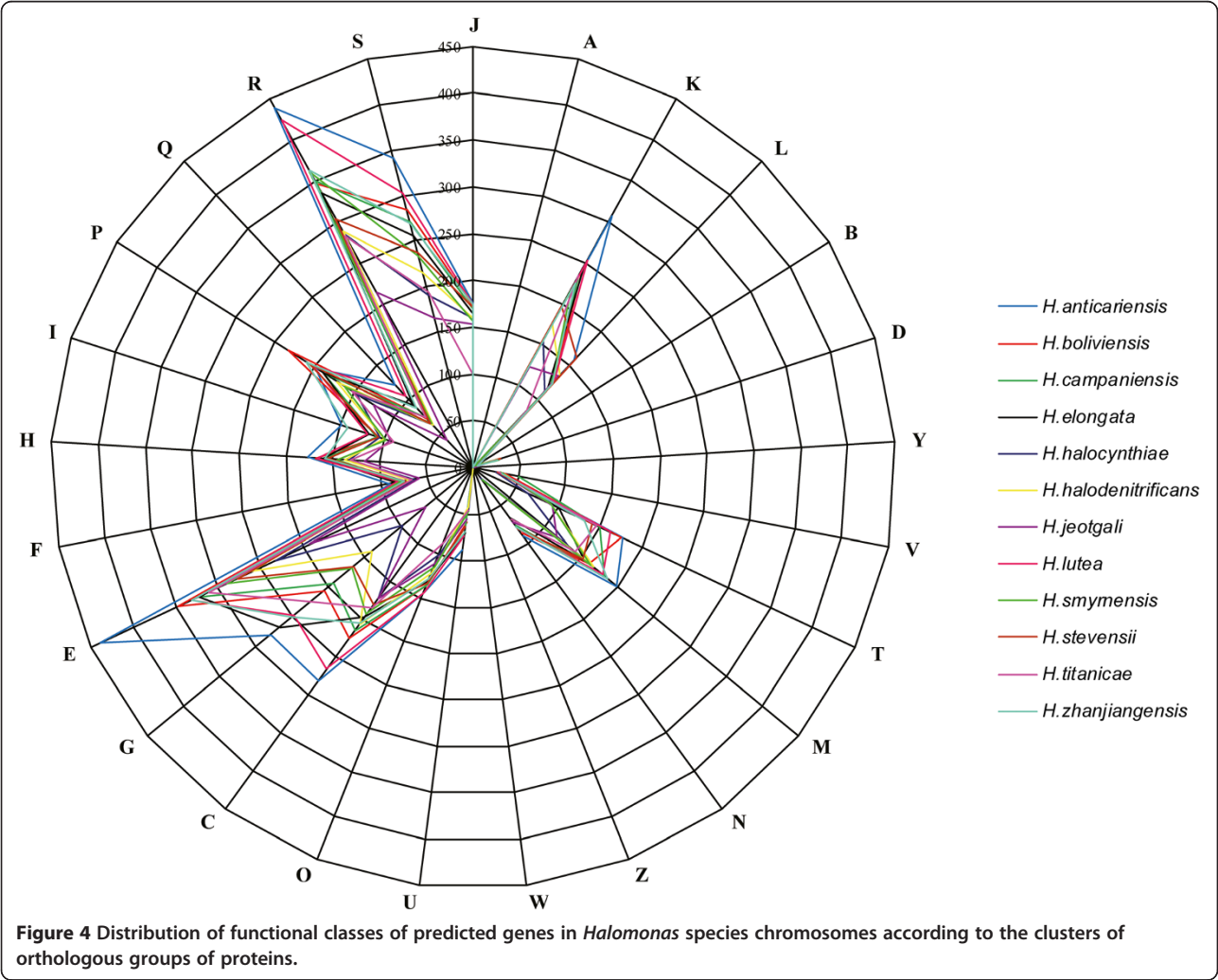


Table 6 Digital DDH similarities between *H. lutea* DSM 23529^T and the other *Halomonas* species

Reference species	Formula 1	Formula 2	Formula 3
<i>H. anticariensis</i>	14.9 ± 3.14	19.5 ± 2.29	15.0 ± 2.67
<i>H. boliviensis</i>	13.0 ± 2.99	20.2 ± 2.31	13.4 ± 2.56
<i>H. campaniense</i>	13.0 ± 2.99	21.1 ± 2.33	13.3 ± 2.56
<i>H. elongata</i>	15.6 ± 3.19	20.1 ± 2.31	15.6 ± 2.70
<i>H. halocynthiae</i>	13.0 ± 2.99	19.2 ± 2.29	13.3 ± 2.56
<i>H. halodenitrificans</i>	14.5 ± 3.11	19.4 ± 2.29	14.6 ± 2.65
<i>H. jeotgali</i>	13.5 ± 3.03	19.9 ± 2.30	13.8 ± 2.59
<i>H. smyrnensis</i>	15.5 ± 3.18	20.3 ± 2.32	15.5 ± 2.70
<i>H. stevensii</i>	13.5 ± 3.04	20.4 ± 2.32	13.8 ± 2.59
<i>H. titanicae</i>	13.0 ± 2.99	20.5 ± 2.32	13.3 ± 2.56
<i>H. zhanjiangensis</i>	13.2 ± 3.01	18.9 ± 2.28	13.5 ± 2.57

GenBank accession numbers for the reference genomes: *H. anticariensis* (NZ_ASTJ000000000), *H. boliviensis* (NZ_AGQZ000000000), *H. campaniense* (CP007757), *H. elongata* (NC_014532), *H. halocynthiae* (AUDZ000000000), *H. halodenitrificans* (JHVH000000000), *H. jeotgali* (NZ_AMQY000000000), *H. smyrnensis* (NZ_AJKS000000000), *H. stevensii* (NZ_AJTS000000000), *H. titanicae* (NZ_AOPO000000000), *H. zhanjiangensis* (NZ_ARIT000000000).

inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets. Given that the low degree of DNA-DNA similarity among *Halomonas* species, it appears justified to assume that these strains represent different species. For better understanding of the relationships between *H. lutea* and other *Halomonas* members, availability of more genome sequences of representatives are needed to implement phylogenomic inference.

As a halophilic bacterium, the genome of *H. lutea* also shows properties related to solute and ion transport, 203 genes related ion transport and metabolism, 60 genes related TRAP-type C4-dicarboxylate transport system which is a crucial family of solute transporters. Moreover, nitrate reduction was tested using API 20NE system and 57 genes were predicted to participate in the nitrogen metabolism. PTS IIA-like nitrogen-regulatory protein, nitrate and sulfonate transport systems related genes were also detected in its genome.

Conclusions

The genome sequence and annotation of *H. lutea* YIM 91125^T were presented. The genome comprises 42 scaffolds which together represent the organism of approximately 4.53 Mb. It encodes for key genes and pathways involved in the compatible solutes production and nitrogen degradation. This provides clues to discover novel genes and functions, and leads to an improved understanding of halophilic microbial evolution and function in the extremely salty conditions. YIM 91125^T participates in nitrogen cycling, although the process of reducing nitrogen needs further studies to fully understand the related pathways. The genome sequencing of *H. lutea* marks an

important step toward a comprehensive genomic catalog and the metabolic diversity of halophilic bacteria. It may contribute to further studies on important process for *Halomonas*, such as quorum-sensing regulatory and osmoadaptation. Combining with genomes of other members in *Halomonas*, will make an important advance in understanding of the ecological roles that *Halomonas* species play in those hypersaline environments and their relationships with other halophilic and nonhalophilic microorganisms.

Additional file

Additional file 1: Figure S1. Phylogenetic tree of the genus *Halomonas*.

Abbreviations

DDH: DNA-DNA hybridization; HSP: High-scoring segment pair.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

WJL and HPK conducted the study. XYG performed the data analyses, genome comparison, and wrote the manuscript. XYZ, HWL, YZ, AL, HPK, NCK and WJL participated in writing the manuscript. JH, MH, EL, MH, AP, NNI, KM, BJT, VM and TW performed genome sequencing, assembly and annotation. All authors read and approved the final manuscript.

Acknowledgements

Susanne Schneider is gratefully acknowledged the assistance for growing *H. lutea* cultures. We also thank Evelyne-Marie Brambilla for DNA extraction and quality control (both at the DSMZ). This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under Contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344. A. L. was supported in part by Russian Ministry of Science Mega-grant no. 11. G3431.0068 (Dr. Stephen J O'Brien Principal Investigator). W.-J. Li was supported by 'Hundred Talents Program' of the Chinese Academy of Sciences.

Author details

¹Key Laboratory of Biogeography and Bioresource in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China. ²Key Laboratory of Microbial Diversity in Southwest China, Ministry of Education and the Laboratory for Conservation and Utilization of Bio-Resources, Yunnan Institute of Microbiology, Yunnan University, Kunming, China. ³Theodosius Dobzhansky Center for Genome Bionformatics, St. Petersburg State University, St. Petersburg, Russia. ⁴Algorithmic Biology Lab, St. Petersburg Academic University, St. Petersburg, Russia. ⁵DOE Joint Genome Institute, Walnut Creek, California, USA. ⁶Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, California, USA. ⁷Leibniz-Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. ⁸Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia. ⁹The First Hospital of Qujing City, Qujing Affiliated Hospital of Kunming Medical University, Qujing, China. ¹⁰State Key Laboratory Breeding Base for Zhejiang Sustainable Plant Pest Control, Institute of Quality and Standard for Agro-products, Zhejiang Academy of Agricultural Sciences, Hangzhou, Zhejiang, China. ¹¹University of Chinese Academy of Sciences, Beijing, China. ¹²School of Biology, Newcastle University, Newcastle upon Tyne, UK.

Received: 28 July 2014 Accepted: 6 November 2014
Published: 20 January 2015

References

- Wang Y, Tang SK, Lou K, Mao PH, Jin X, Jiang CL, Xu LH, Li WJ. *Halomonas lutea* sp. nov., a moderately halophilic bacterium isolated from a salt lake. *Int J Syst Evol Microbiol*. 2008; **58**:2065–69.
- Parte AC. LPSN-List of Prokaryotic names with Standing in nomenclature. *Nucl Acids Res*. 2013; **41**:1–4.
- Kim OS, Cho YJ, Lee K, Yoon SH, Kim M, Na H, Park SC, Jeon YS, Lee JH, Yi H, Won S, Chun J. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol*. 2012; **62**:716–21.
- Franzmann PD, Wehmeyer U, Stackebrandt E. *Halomonadaceae* fam. nov., a New Family of the Class *Proteobacteria* to Accommodate the Genera *Halomonas* and *Deleya*. *Syst Appl Microbiol*. 1988; **11**:16–9.
- Anzai Y, Kim H, Park JY, Wakabayashi H, Oyaizu H. Phylogenetic affiliation of the pseudomonads based on 16S rRNA sequence. *Int J Syst Evol Microbiol*. 2000; **50**:1563–89.
- Arahal DR, Ludwig W, Schleifer KH, Ventosa A. Phylogeny of the family *Halomonadaceae* based on 23S and 16S rDNA sequence analyses. *Int J Syst Evol Microbiol*. 2002; **52**:241–49.
- Mellado E, Moore ERB, Nieto JJ, Ventosa A. Phylogenetic Inferences and Taxonomic Consequences of 16S Ribosomal DNA Sequence Comparison of *Chromohalobacter marismortui*, *Volcaniella eurihalina*, and *Deleya salina* and Reclassification of *V. eurihalina* as *Halomonas eurihalina* comb. nov. *Int J Syst Evol Microbiol*. 1995; **45**:712–16.
- Cabrera A, Aguilera M, Fuentes S, Incerti C, Russell NJ, Ramos-Cormenzana A, Montalvo-Sánchez M. *Halomonas indalini* sp. nov., a moderately halophilic bacterium isolated from a solar saltern in Cabo de Gata, Almería, southern Spain. *Int J Syst Evol Microbiol*. 2007; **57**:376–80.
- Kim KK, Jin L, Yang HC, Lee ST. *Halomonas gomseomensis* sp. nov., *Halomonas janggokensis* sp. nov., *Halomonas salaria* sp. nov. and *Halomonas denitrificans* sp. nov., moderately halophilic bacteria isolated from saline water. *Int J Syst Evol Microbiol*. 2007; **57**:675–81.
- Soto-Ramírez N, Sánchez-Porro C, Rosas S, González W, Quiñones M, Ventosa A, Montalvo-Rodríguez P. *Halomonas avicenniae* sp. nov., isolated from the salty leaves of the black mangrove *Avicennia germinans* in Puerto Rico. *Int J Syst Evol Microbiol*. 2007; **57**:900–05.
- Ventosa JJ, Nieto JJ, Oren A. Biology of moderately halophilic aerobic bacteria. *Microbiol Mol Biol Rev*. 1998; **62**:504–44.
- Wang YN, Cai H, Chi CQ, Lu AH, Lin XG, Jiang ZF, Wu XL. *Halomonas shengliensis* sp. nov., a moderately halophilic, denitrifying, crude-oil-utilizing bacterium. *Int J Syst Evol Microbiol*. 2007; **57**:1222–26.
- Kyrpides NC, Hugenholtz P, Eisen JA, Woyke T, Goker M, Parker CT, Amann R, Beck BJ, Chain PS, Chun J, Colwell RR, Danchin A, Dawyndt P, Dedeurwaerdere T, DeLong EF, Detter JC, De Vos P, Donohue TJ, Dong XZ, Ehrlich DS, Fraser C, Gibbs R, Gilbert J, Gilna P, Glöckner FO, Jansson JK, Keasling JD, Knight R, Labeda D, Lapidus A, et al. Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. *PLoS Biol*. 2014; **12**:e1001920.
- Berendes F, Gottschalk G, Heine-Dobbernack E, Moore ERB, Tindall BJ. *Halomonas desiderata* sp. nov., a new alkaliphilic, halotolerant and denitrifying bacterium isolated from a municipal sewage works. *Syst Appl Microbiol*. 1996; **19**:158–67.
- Romano I, Gottschalk G, Heine-Dobbernack E, Moore ERB, Tindall BJ. Characterization of a haloalkaliphilic strictly aerobic bacterium, isolated from Pantelleria island. *Syst Appl Microbiol*. 1996; **19**:326–33.
- Heyman J, Balcaen A, De Vos P, Swings J. *Halomonas muralis* sp. nov., isolated from microbial biofilms colonizing the walls and murals of the Saint-Catherine chapel (Castle Herberstein, Austria). *Int J Syst Evol Microbiol*. 2002; **52**:2049–54.
- Jeon CO, Lim JM, Lee JR, Lee GS, Park DJ, Lee JC, Oh HW, Kim CJ. *Halomonas kribbensis* sp. nov., a novel moderately halophilic bacterium isolated from a solar saltern in Korea. *Int J Syst Evol Microbiol*. 2007; **57**:2194–98.
- Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, Ashburner M, Axelrod N, Baldauf S, Ballard S, Boore J, Cochrane G, Cole J, Dawyndt P, De Vos P, DePamphilis C, Edwards R, Faruque N, Feldman R, Gilbert J, Gilna P, Glöckner FO, Goldstein P, Guralnick R, Haft D, Hancock D, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol*. 2008; **26**:541–47.
- Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A*. 1990; **87**:4576–79.
- Garrity GM, Bell JA, Lilburn T. Phylum XIV. *Proteobacteria* phyl. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, Volume 2. 2nd ed. New York: Springer; 2005.
- Garrity A. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. *Int J Syst Evol Microbiol*. 2005; **55**:2235–38.
- Garrity GM, Bell JA, Lilburn T. Class III. *Gammaproteobacteria* class. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, Volume 2. 2nd ed. New York: Springer; 2005.
- Garrity GM, Bell JA, Lilburn T. Order VIII. *Oceanospirillales* ord. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, Volume 2. 2nd ed. New York: Springer; 2005.
- Vreeland R, Litchfield CD, Martin EL, Elliot E. *Halomonas elongata*, a new genus and species of extremely salt-tolerant bacteria. *Int J Syst Evol Microbiol*. 1980; **30**:485–95.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. *Nat Genet*. 2000; **25**:25–9.
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. 2006; **72**:5069–72.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990; **215**:403–10.
- Korf I, Yandell M, Bedell J. BLAST. Sebastopol, CA: O'Reilly Press; 2003.
- Lee C, Grasso C, Sharlow MF. Multiple sequence alignment using partial order graphs. *Bioinformatics*. 2002; **18**:452–64.
- Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987; **4**:406–25.
- Strimmer K, von Haeseler A. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol Biol Evol*. 1996; **13**:964–69.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011; **28**:2731–39.
- Pattengale ND, Alipour M, Bininda-Emonds OR, Moret BM, Stamatakis A. How many bootstrap replicates are necessary? *J Comput Biol*. 2010; **17**:337–54.
- Pagani I, Liolios K, Jansson J, Chen IM, Smirnova T, Nosrat B, Markowitz VM, Kyrpides NC. The Genomes OnLine Database (GOLD) v. 4: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res*. 2012; **40**:D571–79.
- Göker M, Klenk HP. Phylogeny-driven target selection for genome-sequencing (and other) projects. *Standards in Genomic Sci*. 2013; **8**:360–74.
- Klenk HP, Göker M. En route to a genome-based classification of Archaea and Bacteria? *Syst Appl Microbiol*. 2010; **33**:175–82.
- Kyrpides NC, Woyke T, Eisen JA, Garrity G, Lilburn TG, Beck BJ, Whitman WB, Hugenholz P, Klenk HP. Genomic Encyclopedia of Type Strains, Phase I: the one thousand microbial genomes (KMG-I) project. *Stand Genomic Sci*. 2013; **9**:628–34.
- Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ, Hooper SD, Pati A, Lykidis A, Spring S, Anderson IJ, D'haeseleer P, Zemla A, Singer M, Lapidus A, Nolan M, Copeland A, Han C, Chen F, Cheng JF, Lucas S, Kerfeld C, Lang E, Gronow S, Chain P, Bruce D, et al. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature*. 2009; **462**:1056–60.
- Mavromatis K, Land ML, Brettin TS, Quest DJ, Copeland A, Clum A, Goodwin L, Woyke T, Lapidus A, Klenk HP, Cottingham RW, Kyrpides NC. The fast changing landscape of sequencing technologies and their impact on microbial genome assemblies and annotation. *PLoS ONE*. 2012; **7**:e48837.
- List of growth media used at DSMZ [http://www.dsmz.de/catalogues/catalogue-microorganisms/culture-technology/list-of-media-for-micro-organisms.html]
- Gemeinhöfer B, Dröge G, Zetzsche H, Haszprunar G, Klenk HP, Güntsch A, Berendsohn WG, Wägele JW. The DNA Bank Network: the start from a German initiative. *Biopreserv Biobank*. 2011; **9**:51–5.
- Bennett S. Solexa Ltd. *Pharmacogenomics*. 2004; **5**:433–38.
- Zerbino D, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 2008; **18**:821–29.
- Wgsim [https://github.com/lh3/wgsim]

45. Gnerre S, MacCallum I. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A*. 2011; **108**:1513–18.
46. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 2010; **11**:119.
47. Mavromatis K, Ivanova NN, Chen IM, Szeto E, Markowitz VM, Kyrpides NC. The DOE-JGI Standard operating procedure for the annotations of microbial genomes. *Stand Genomic Sci*. 2009; **1**:63–7.
48. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods*. 2010; **7**:455–57.
49. Markowitz VM, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics*. 2009; **25**:2271–78.
50. Koehler S, Plewniak F, Barbe V, Battaglia-Brunet F, Jost B, Joulain C, Philipps M, Vicaire S, Vincent S, Ye T, Bertin PN. Genome Sequence of *Halomonas* sp. Strain A3H3, Isolated from Arsenic-Rich Marine Sediments. *Genome Announc*. 2013; **1**:e00819–13.
51. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A*. 2009; **106**:19126–31.
52. Auch AF, von Jan M, Klenk HP, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci*. 2010; **2**:117–34.
53. Auch AF, Klenk HP, Göker M. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. *Stand Genomic Sci*. 2010; **2**:142–48.

doi:10.1186/1944-3277-10-1

Cite this article as: Gao et al.: Draft genome sequence of *Halomonas lutea* strain YIM 91125^T (DSM 23508^T) isolated from the alkaline Lake Ebinur in Northwest China. *Standards in Genomic Sciences* 2015 **10**:1.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

