



SHORT GENOME REPORT

Open Access

High quality draft genome sequence of *Brachymonas chironomi* AIMA4^T (DSM 19884^T) isolated from a *Chironomus* sp. egg mass

Sivan Laviad¹, Alla Lapidus^{2,3}, James Han⁴, Matthew Haynes⁴, TBK Reddy⁴, Marcel Huntemann⁴, Amrita Pati⁴, Natalia N Ivanova⁴, Konstantinos Mavromatis⁴, Elke Lang⁵, Manfred Rohde⁶, Victor Markowitz⁷, Tanja Woyke⁴, Hans-Peter Klenk⁵, Nikos C Kyrpides^{4,8} and Malka Halpern^{1,9*}

Abstract

Brachymonas chironomi strain AIMA4^T (Halpern et al., 2009) is a Gram-negative, non-motile, aerobic, chemoorganotroph bacterium. *B. chironomi* is a member of the *Comamonadaceae*, a family within the class *Betaproteobacteria*. This species was isolated from a chironomid (*Diptera*; *Chironomidae*) egg mass, sampled from a waste stabilization pond in northern Israel. Phylogenetic analysis based on the 16S rRNA gene sequences placed strain AIMA4^T in the genus *Brachymonas*. Here we describe the features of this organism, together with the complete genome sequence and annotation. The DNA GC content is 63.5%. The chromosome length is 2,509,395 bp. It encodes 2,382 proteins and 68 RNA genes. *Brachymonas chironomi* genome is part of the *Genomic Encyclopedia of Type Strains, Phase I: the one thousand microbial genomes (KMG) project*.

Keywords: *Brachymonas chironomi*, *Comamonadaceae*, Chironomid, *Chironomus*, Egg mass, Toxicant

Introduction

Strain AIMA4^T (= LGM 24400^T = DSM 19884^T), is the type strain of *Brachymonas chironomi*, one out of two species in the genus *Brachymonas*. The genus *Brachymonas* was formed by Hiraishi et al. [1] while characterizing rhodoquinone-containing bacteria that had been isolated from soybean crude waste sludge in Japan. Strain AIMA4^T, was isolated from an insect egg mass (*Chironomus* sp.) that was sampled from a waste stabilization pond in northern Israel [2]. Chironomids (*Arthropoda*; *Insecta*; *Diptera*; *Chironomidae*; *Chironomus* sp.) inhabit virtually every type and condition of aquatic habitats. They undergo a complete metamorphosis of four life stages (egg, larva, pupa and adult that emerges into the air) [3]. Eggs are laid in an egg mass at the water's edge. Each egg mass contains hundreds of eggs. Chironomid egg masses were found to harbor *Vibrio cholerae* and

Aeromonas spp. [3-10]. *V. cholerae* degrades chironomid egg masses by the secreted haemagglutinin protease (HAP) [11,12]. Strain AIMA4^T was isolated in the course of a study that investigated endogenous bacterial communities that inhabit chironomid egg masses [2,13,14]. The species epithet *chironomi* was derived from the non-biting midge insect *Chironomus* (*Diptera*; *Chironomidae*), from where this species was isolated. Strain AIMA4^T didn't show the ability to degrade the egg masses like it was found for *V. cholerae*.

Here we describe a summary classification and a set of the features of *Brachymonas chironomi* strain AIMA4^T (DSM 19884^T), together with the genome sequence description and annotation.

Organism information

Classification and features

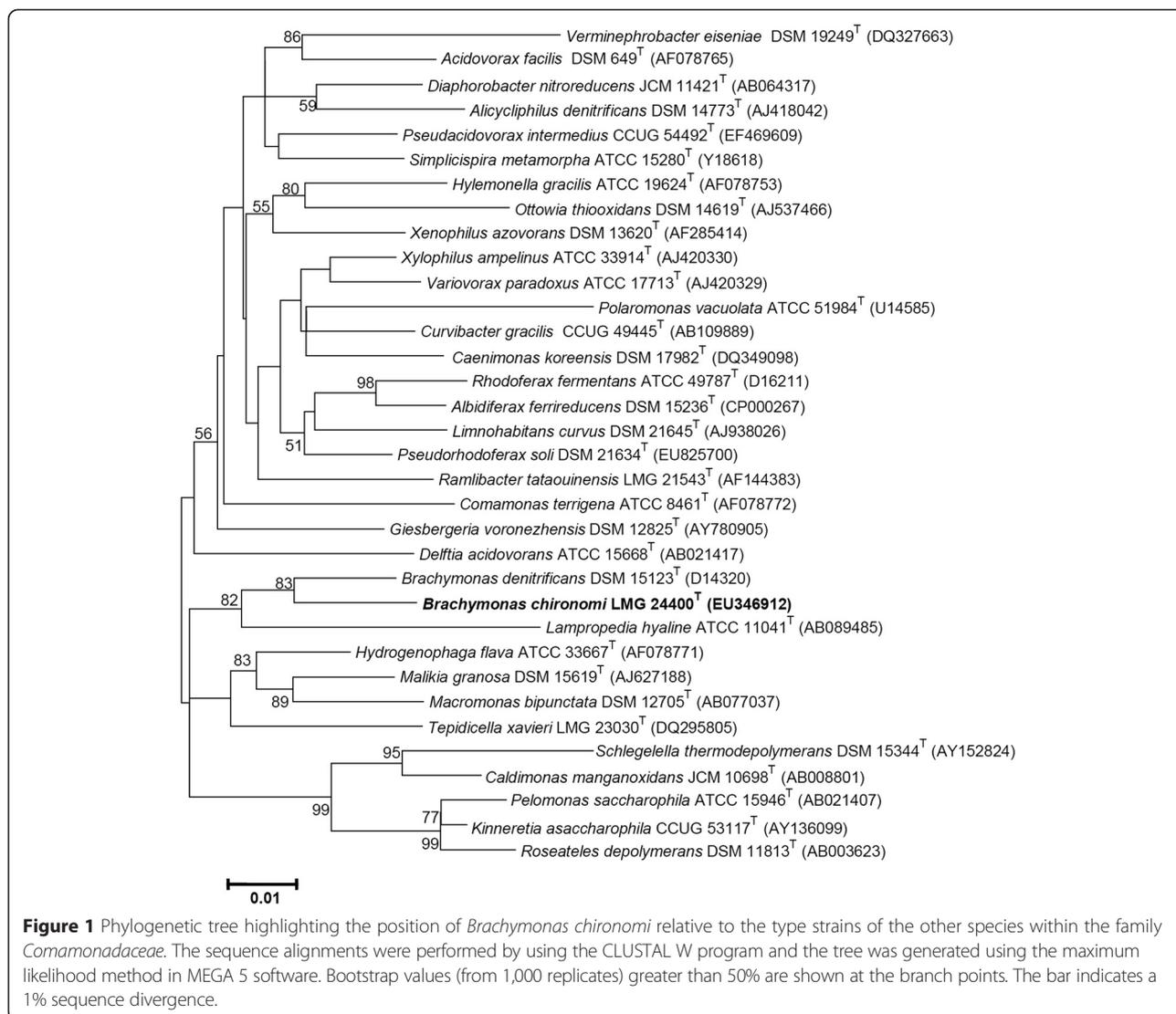
A taxonomic study using a polyphasic approach placed *B. chironomi* strain AIMA4^T in the genus *Brachymonas* within the family *Comamonadaceae* (Figure 1). The family *Comamonadaceae* comprises a larger number of genera

* Correspondence: mhalpern@research.haifa.ac.il

¹Dept. of Evolutionary and Environmental Biology, Faculty of Natural Sciences, University of Haifa, Haifa, Israel

⁹Dept. of Biology and Environment, Faculty of Natural Sciences, University of Haifa, Oranim, Kiryat Tivon, Israel

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



(as shown in Figure 1) and a larger variety of species and phenotypes [15,16].

B. chironomi strain AIMA4^T is a Gram-negative, non-motile coccobacillus or rod (Figure 2). After 48 h incubation on LB agar at 30°C, colonies are beige colored (opaque) that turn to light brown after few days of incubation. Strain AIMA4^T is aerobic, chemoorganotrophic and does not produce acid from carbohydrates (including glucose) [2]. Growth is observed at 18–37°C (optimum 30°C), with 0–2.5% (w/v) NaCl (optimum 0.5% NaCl) and at pH 5.0–9.0 (optimum pH 6.0–8.0) (Table 1). The following enzymatic activities were observed in strain AIMA4^T: catalase and oxidase, alkaline and acid phosphatases, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, trypsin and naphthol-AS-BI-phosphohydrolase. Strain AIMA4^T produces acetoin and reduces nitrate to nitrite [2].

Chemotaxonomic data

The dominant cellular fatty acids are C_{16:1} ω7c, C_{16:0} and C_{18:1} ω7c. The main isoprenoid quinone is Q-8. Phosphatidylglycerol, phosphatidylethanolamine and phosphatidylserine occur as polar lipids [2].

Genome sequencing and annotation

Genome project history

This organism was selected for sequencing on the basis of its phylogenetic position [17–19]. Sequencing of *B. chironomi* strain AIMA4^T is part of Genomic Encyclopedia of Type Strains, Phase I: the one thousand microbial genomes project [20] which aims in increasing the sequencing coverage of key reference microbial genomes [21]. The genome project is deposited in the Genomes On-Line Database [22] and the permanent draft genome sequence is deposited in GenBank. Sequencing, finishing and

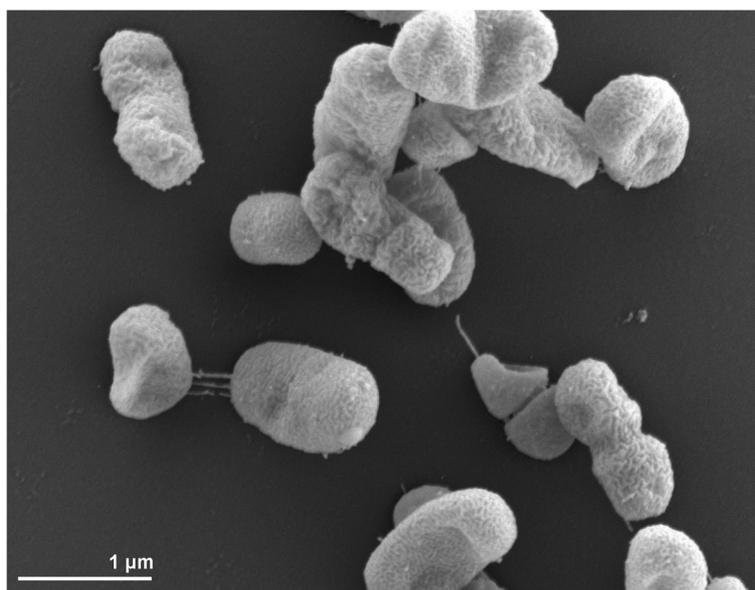


Figure 2 Scanning electron micrograph of *B. chironomi* AIMA4^T.

annotation were performed by the DOE Joint Genome Institute (JGI) using state of the art sequencing technology [23]. A summary of the project information is shown in Table 2.

Growth conditions and genomic DNA preparation

B. chironomi strain AIMA4^T, DSM 19884^T, was grown in DSMZ medium 1 (Nutrient Agar), at 28°C [24]. DNA was isolated from 0.5-1 g of cell paste using JetFlex Genomic DNA Purification Kit (GENOMED) following the standard protocol as recommended by the manufacturer, however with additional 50 μl protease K (20 mg/ml) during digest for 60 min. at 58°C. Protein precipitation was done with additional 200 μl Protein Precipitation Buffer, followed by over night incubation on ice. DNA is available through the DNA Bank Network [25].

Genome sequencing and assembly

The draft genome of *B. chironomi* strain AIMA4^T was generated using the Illumina technology [23,26]. An Illumina standard shotgun library was constructed and sequenced using the Illumina HiSeq 2000 platform which generated 14,014,260 reads totaling 2,102.1 Mb. All general aspects of library construction and sequencing performed at the JGI can be found at the institute website [27]. All raw Illumina sequence data was passed through DUK, a filtering program developed at JGI, which removes known Illumina sequencing and library preparation artifacts [28]. Following steps were then performed for assembly: (1) filtered Illumina reads were assembled using Velvet [29], (2) 1–3 Kbp simulated paired end reads were created from Velvet Contigs using wgsim [30], (3) Illumina reads were

assembled with simulated read pairs using Allpaths-LG [31]. Parameters for assembly steps were: (1) Velvet (velveth: 63 –shortPaired and velvetg: –very clean yes –export-Filtered yes –min contig lgth 500 –scaffolding no –cov cutoff 10) (2) wgsim (–e 0 –1 100 –2 100 –r 0 –R 0 –X 0) (3) Allpaths-LG (PrepareAllpathsInputs: PHRED 64 = 1 PLOIDY = 1 FRAG COVERAGE = 125 JUMP COVERAGE = 25 LONG JUMP COV = 50, RunAllpathsLG: THREADS = 8 RUN = std shredpairs TARGETS = standard VAPI WARN ONLY = True OVERWRITE = True). The final draft assembly contained 36 contigs in 36 scaffolds. The total size of the genome is 2.5 Mbp and the final assembly is based on 249.2 Mbp of Illumina data, which provides an average 99.6 × coverage of the genome.

Genome annotation

Genes were identified using Prodigal [32] as part of the DOE-JGI genome annotation pipeline [33,34], following by a round of manual curation using the JGI GenePRIMP pipeline [35]. The predicted CDSs were translated and searched against the Integrated Microbial Genomes (IMG) non-redundant database, UniProt, TIGERfam, Pfam, PRIAM, KEGG, COG, and InterPro databases. These data sources were combined to assert a product description for each predicted protein. Additional gene prediction analysis and functional annotation was performed within the Integrated Microbial Genomes-Expert Review (IMG-ER) platform [36].

Genome properties

The assembly of the draft genome sequence consists of 36 scaffolds amounting to 2,509,395 bp, and the G + C

Table 1 Classification and general features of *Brachymonas chironomi* strain AIMA4^T according to the MIGS recommendations [40], published by the Genome Standards Consortium [41] and the Names for Life database [42]

MIGS ID	Property	Term	Evidence code ^a
	Classification	Domain <i>Bacteria</i>	TAS [43]
		Phylum <i>Proteobacteria</i>	TAS [44]
		Class <i>Betaproteobacteria</i>	TAS [45]
		Order <i>Burkholderiales</i>	TAS [46]
		Family <i>Comamonadaceae</i>	TAS [47]
		Genus <i>Brachymonas</i>	TAS [1]
		Species <i>Brachymonas chironomi</i>	TAS [2]
		Type strain AIMA4 ^T	TAS [2]
	Gram stain	Negative	TAS [2]
	Cell shape	Cocci/bacilli or rods	TAS [2]
	Motility	Non-motile	TAS [2]
	Sporulation	Non-sporulating	IDS
	Temperature range	18–37°C	TAS [2]
	Optimum temperature	30°C	TAS [2]
	pH range; Optimum	5.0–9.0; 6.0–8.0	TAS [2]
	Carbon source ^b	phenylacetic acid	TAS [2]
MIGS-6	Habitat	Aquatic/Insect host	TAS [2]
MIGS-6.3	Salinity	0–2.5% NaCl (w/v)	TAS [2]
MIGS-22	Oxygen requirement	Aerobic	TAS [2]
MIGS-15	Biotic relationship	Commensal (Insect, chironomid)	TAS [2]
MIGS-14	Pathogenicity	Non-pathogen	NAS
MIGS-4	Geographic location	Israel	TAS [2]
MIGS-5	Sample collection	July, 2006	TAS [2]
MIGS-4.1	Latitude	32.669167	TAS [2]
MIGS-4.2	Longitude	35.128639	TAS [2]
MIGS-4.4	Altitude	40 m	TAS [2]

^aEvidence codes - IDA: Inferred from Direct Assay; 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). Evidence codes are from the Gene Ontology project [48].

^bThe only carbon source that was positive for this strain, out of all carbon sources that were tested (strain AIMA4^T does not use carbohydrates, not even glucose) [2].

content is 63.5% (Table 3). Of the 2,450 genes predicted, 2,382 were protein-coding genes, and 68 RNAs. The majority of the protein-coding genes (85.5%) were assigned a putative function while the remaining ones were annotated as hypothetical proteins. The distribution of genes into COGs functional categories is presented in Table 4.

Table 2 Genome sequencing project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Level 2: High-Quality Draft
MIGS-28	Libraries used	Illumina Std. shotgun library
MIGS 29	Sequencing platforms	Illumina HiSeq 2000
MIGS 31.2	Fold coverage	99.6x
MIGS 30	Assemblers	Velvet v. 1.1.04, ALLPATHS v. R37654
MIGS 32	Gene calling method	Prodigal 2.5
	Locus Tag	C513
	GenBank ID	ARGE000000000
	GenBank Date of Release	September 16, 2013
	GOLD ID	Gp0013605
	BIOPROJECT	174982
MIGS 13	Source Material Identifier	DSM 19884 ^T
	Project relevance	Tree of Life, GEBA-KMG

Insights from the genome sequence

Strain AIMA4^T was isolated from chironomid egg masses. Using pyrosequencing method, we have recently shown that the prevalence of *Brachymonas* in the endogenous bacterial communities of chironomid egg masses and larva was 0.04% and 0.006%, respectively [37]. Chironomid tolerance towards pollution is well documented [38]. Senderovich and Halpern [37,39], demonstrated by using Koch's postulates that endogenous bacteria in chironomids have a role in protecting the insect from toxicants. Although *B. chironomi* was isolated from chironomid egg masses, its features regarding its protective potential have never been examined. Nevertheless, its genome reveals the potential

Table 3 Genome statistics

Attribute	Value	% of Total
Genome size (bp)	2,509,395	100.00%
DNA coding (bp)	2,294,427	91.43%
DNA G + C (bp)	1,593,935	63.52%
DNA scaffolds	36	100.00%
Total genes	2,450	100.00%
Protein coding genes	2,382	97.22%
RNA genes	68	2.78%
Pseudo genes	0	0
Genes in internal clusters	1,788	72.98%
Genes with function prediction	2,095	85.51%
Genes assigned to COGs	1,829	74.65%
Genes with Pfam domains	2,129	86.90%
Genes with signal peptides	171	6.98%
Genes with transmembrane helices	505	20.61%
CRISPR repeats	0	0

Table 4 Number of genes associated with the general COG functional categories

Code	Value	% age	Description
J	149	7.44	Translation, ribosomal structure and biogenesis
A	1	0.05	RNA processing and modification
K	104	5.19	Transcription
L	106	5.29	Replication, recombination and repair
B	1	0.05	Chromatin structure and dynamics
D	26	1.30	Cell cycle control, cell division, chromosome partitioning
V	32	1.60	Defense mechanisms
T	60	3.00	Signal transduction mechanisms
M	122	6.09	Cell wall/membrane/envelope biogenesis
N	15	0.75	Cell motility
U	60	3.00	Intracellular trafficking, secretion, and vesicular transport
O	95	4.75	Posttranslational modification, protein turnover, chaperones
C	137	6.84	Energy production and conversion
G	66	3.30	Carbohydrate transport and metabolism
E	182	9.09	Amino acid transport and metabolism
F	54	2.70	Nucleotide transport and metabolism
H	113	5.64	Coenzyme transport and metabolism
I	103	5.14	Lipid transport and metabolism
P	115	5.74	Inorganic ion transport and metabolism
Q	52	2.60	Secondary metabolites biosynthesis, transport and catabolism
R	227	11.34	General function prediction only
S	180	8.99	Function unknown
-	621	25.35	Not in COGs

of this species to protect its host in polluted environments. Genes encoding arsenate detoxification are present in *B. chironomi* strain AIMA4^T. These genes include an arsenical resistance gene cluster with candidates for transcriptional regulator, ArsR; arsenical resistance operon trans-acting repressor, ArsD; arsenite efflux ATP-binding protein, ArsA and a hypothetical arsenic resistance protein (ACR3 family). A gene for arsenate reductase (ArsC family) is present in a different operon. More genes which may indicate the potential of this bacterium to tolerate or detoxify metals are: copper resistance protein D, CopD; copper chaperone, copper-resistance protein, CopA; copper (or silver) translocating P-type ATPase; uncharacterized lipoprotein NlpE involved in copper resistance; magnesium Mg(2+) and cobalt Co(2+) transport protein, CorA. Moreover, two genes encoding ABC-type transport system involved in resistance to organic solvents, auxiliary and periplasmic components are also present.

The genome of *B. chironomi* strain AIMA4^T reveals the potential of the species to produce a polysaccharide

capsule. It includes two gene clusters with candidates for capsule polysaccharide export protein, periplasmic protein involved in polysaccharide export, ABC-type polysaccharide/polyol phosphate transport system, ATPase component, ABC-type polysaccharide/polyol phosphate export systems, permease component and predicted glycosyltransferase involved in capsule biosynthesis. Another feature that is found in the genome of *B. chironomi* AIMA4^T is its potential to produce a pilus (or pili). The following predicted genes indicate this ability; type IV pilus assembly protein PilB; type IV pilus secretin PilQ; Tfp pilus assembly proteins PilP, PilO and PilV; type IV prepilin peptidase; prepilin-type N-terminal cleavage/methylation domain and pilus retraction ATPase PilT (indicating the ability of twitching motility).

Tolerance of 2.5% NaCl was described for strain AIMA4^T by Halpern et al. [2]. The presence of ABC-type proline/glycine betaine transport system in the genome may explain the way this species can tolerate high salt concentrations. In respect to the ampicillin (beta-lactam) antibiotic resistance, the genome encodes one beta-lactamase class B and a negative regulator of beta-lactamase expression. Three genes encoding two component transcriptional regulators (LuxR family), can be found in the genome of strain AIMA4^T and demonstrate quorum sensing skills.

Conclusions

In the current study, we characterized the genome of *B. chironomi* strain AIMA4^T that was isolated from a chironomid egg mass [2]. *B. chironomi* belongs to the family *Comamonadaceae* (order *Bukholderiales*; class *Betaproteobacteria*) (Figure 1). Members of this family are known for their ability to cope with harsh environmental condition such as high concentration of toxic metals and other pollutants like aromatic compounds or polymers [e.g. poly (3-hydroxybutyrate-co-3-hydroxyvalerate)] [16]. Likewise, the genome of strain AIMA4^T reveals the potential of this species to cope with toxic metals. These demonstrate that *B. chironomi* may have a role in protecting its aquatic host (chironomids) in polluted environments.

Abbreviations

KMG: One thousand microbial genomes; PHBV: Poly(3-hydroxybutyrate-co-3-hydroxyvalerate).

Competing interest

The authors declare that they have no competing interests.

Authors' contributions

MH (Halpern) isolated and characterized strain *B. chironomi* AIMA4^T; SL, MH (Halpern), HPK and NCK drafted the manuscript. AL, JH, MH (Haynes), TBKR, MH (Huntemann), AP, NNI, KM, VM and TW sequenced, assembled and annotated the genome. EL provided the biomass for DNA extraction and collected data about the organism. MR performed electron microscopy. All authors read and approved the final manuscript.

Acknowledgements

We would like to gratefully acknowledge the help of Nicole Reimann for growing *B. chironomi* cultures, and Evelyne-Marie Brambilla for DNA extraction

and quality control (both at 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 Genome analysis was supported by the National Basic Research Program of China (No. 2010CB833801). A.L. was supported in part by Russian Ministry of Science Mega-grant no.11.G34.31.0068 (PI Dr Stephen J O'Brien). M. H. was supported in part by a grant from the US Civilian Research and Development Foundation (CRDF grant no. ILB1-7045-HA).

Author details

¹Dept. of Evolutionary and Environmental Biology, Faculty of Natural Sciences, University of Haifa, Haifa, Israel. ²Theodosius Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, St. Petersburg, Russia. ³Algorithmic Biology Lab, St. Petersburg Academic University, St. Petersburg, Russia. ⁴DOE Joint Genome Institute, Walnut Creek, CA, USA. ⁵Leibniz-Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. ⁶Helmholz Centre for Infection Research, Braunschweig, Germany. ⁷Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ⁸Dept. of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia. ⁹Dept. of Biology and Environment, Faculty of Natural Sciences, University of Haifa, Oranim, Kiryat Tivon, Israel.

Received: 15 September 2014 Accepted: 16 April 2015

Published online: 27 May 2015

References

- Hiraishi A, Shin YK, Sugiyama J. *Brachymonas denitrificans* gen. nov., sp. nov., an aerobic chemoorganotrophic bacterium which contains rodoquinones, and evolutionary relationships of rodoquinone producers to bacterial species with various quinone classes. *J Gen Appl Microbiol*. 1995;41:99–117.
- Halpern M, Shaked T, Schumann P. *Brachymonas chironomi* sp. nov., isolated from a chironomid egg mass, and emended description of the genus *Brachymonas*. *Int J Syst Evol Microbiol*. 2009;59:3025–9.
- Halpern M, Landsberg O, Raats D, Rosenberg E. Culturable and VBNC *Vibrio cholerae*: interactions with chironomid egg masses and their bacterial population. *Microb Ecol*. 2007;53:285–93.
- Broza M, Halpern M. Chironomid egg masses and *Vibrio cholerae*. *Nature*. 2001;412:40.
- Halpern M, Broza YB, Mittler S, Arakawa E, Broza M. Chironomid egg masses as a natural reservoir of *Vibrio cholerae* non-O1 and non-O139 in freshwater habitats. *Microb Ecol*. 2004;47:341–9.
- Halpern M, Raats D, Lavion R, Mittler S. Dependent population dynamics between chironomids (non-biting midges) and *Vibrio cholerae*. *FEMS Microbiol Ecol*. 2006;55:98–104.
- Senderovich Y, Gershtein Y, Halewa E, Halpern M. *Vibrio cholerae* and *Aeromonas*; do they share a mutual host? *ISME J*. 2008;2:276–83.
- Figueras MJ, Beaz-Hidalgo R, Senderovich Y, Laviad S, Halpern M. Re-identification of *Aeromonas* isolates from chironomid egg masses as the potential pathogenic bacteria *Aeromonas aquariorum*. *Environ Microbiol Rep*. 2011;3:239–44.
- Beaz-Hidalgo R, Shakèd T, Laviad S, Halpern M, Figueras M. Chironomid egg masses harbour the clinical species *Aeromonas taiwanensis* and *Aeromonas sanarellii*. *FEMS Microbiol Lett*. 2012;337:48–54.
- Halpern M. Chironomids and *Vibrio cholerae*. In: Rosenberg E, Gophna U, editors. *Beneficial Microorganisms in Multicultural Life Forms*. Berlin Heidelberg: Springer; 2011. p. 43–56.
- Halpern M, Gancz H, Broza M, Kashi Y. *Vibrio cholerae* hemagglutinin/protease degrades chironomid egg masses. *Appl Environ Microbiol*. 2003;69:4200–4.
- Halpern M. Novel insights into hemagglutinin (HAP) gene regulation in *Vibrio cholerae*. *Mol Ecol*. 2010;19:4108–12.
- Raats D, Halpern M. *Oceanobacillus chironomi* sp. nov., a halotolerant and facultative alkaliphilic species isolated from a chironomid egg mass. *Int J Syst Evol Microbiol*. 2007;57:255–9.
- Halpern M, Senderovich Y, Snir S. *Rheinheimeria chironomi* sp. nov., isolated from a chironomid (*Diptera*; *Chironomidae*) egg mass. *Int J Syst Evol Microbiol*. 2007;57:1872–5.
- Willems A, Pot B, Falsen E, Vandamme P, Gillis M, Kersters K, et al. Polyphasic taxonomic study of the emended genus *Comamonas*: relationship to *Aquaspirillum aquaticum*, E. Falsen group 10, and other clinical isolates. *Int J Syst Bacteriol*. 1991;41:247–244.
- Weiss M, Kesberg A, LaButti K, Pitluck S, Bruce D, Hauser L, et al. Permanent draft genome sequence of *Comamonas testosteroni* KF-1. *Stand Genomic Sci*. 2013;8:2.
- Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, et al. A phylogeny-driven Genomic Encyclopaedia of *Bacteria* and *Archaea*. *Nature*. 2009;462:1056–60.
- Klenk HP, Göker M. *En route* to a genome - based classification of *Archaea* and *Bacteria*? *Syst Appl Microbiol*. 2010;33:175–82.
- Göker M, Klenk HP. Phylogeny-driven target selection for large-scale genome-sequencing (and other) projects. *Stand Genomic Sci*. 2013;8:360–74.
- Kyrpides NC, Woyke T, Eisen JA, Garrity G, Lilburn TG, Beck BJ, et al. Genomic encyclopaedia of type strains, phase I: the one thousand microbial genomes (KMG-I) project. *Stand Genomic Sci*. 2013;9:628–6234.
- Kyrpides NC, Hugenholtz P, Eisen JA, Woyke T, Göker M, Parker CT, et al. Genomic Encyclopaedia of *Bacteria* and *Archaea*: sequencing a myriad of type strains. *PLoS Biol*. 2014;8:e1001920.
- Pagani I, Liolios K, Jansson J, Chen IM, Smirnova T, Nosrat B, et al. The genomes OnLine database (GOLD) v. 4: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res*. 2012;40:D571–9.
- Mavromatis K, Land ML, Brettin TS, Quest DJ, Copeland A, Clum A, et al. 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-microorganisms.html>]
- Gemeinholzer B, Dröge G, Zetzsche H, Haszprunar G, Klenk HP, Güntsch A, et al. The DNA bank network: the start from a German initiative. *Biopreserv Biobank*. 2011;9:51–5.
- Bennett S. Solexa Ltd. *Pharmacogenomics*. 2004;5:433–8.
- JGI web page [<http://www.jgi.doe.gov>]
- Mingkun L, Copeland A, Han J. DUK, unpublished, 2011
- Zerbino D, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 2008;18:821–9.
- <https://github.com/lh3/wgsim>
- Gnerre S, MacCallum I. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A*. 2011;108:1513–8.
- 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.
- 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.
- Chen IM, Markowitz VM, Chu K, Anderson I, Mavromatis K, Kyrpides NC, et al. Improving microbial genome annotations in an integrated database context. *PLoS One*. 2013;8:e54859.
- Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, et al. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods*. 2010;7:455–7.
- Markowitz VM, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics*. 2009;25:2271–8.
- Senderovich Y, Halpern M. The protective role of endogenous bacterial communities in chironomid egg masses and larvae. *ISME J*. 2013;7:2147–58.
- Armitage P, Cranston PS, Pinder LCV. *The Chironomidae: The Biology and Ecology of Non-biting Midges*. London: Chapman and Hall; 1995.
- Senderovich Y, Halpern M. Bacterial community composition associated with chironomid egg masses. *J Insect Sci*. 2012;12:148.
- Field D, Garrity GM, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol*. 2008;26:541–7.
- Field D, Amaral-Zettler L, Cochrane G, Cole JR, Dawyndt P, Garrity GM, et al. The Genomic Standards Consortium. *PLoS Biol*. 2011;9:e1001088.
- Garrity GM. Namesforlife browser tool takes expertise out of the database and puts it right in the browser. *Microbiol Today*. 2010;37:9.
- Woese CR, Kandler O, Weelis ML. Towards a natural system of organisms. Proposal for the domains *Archaea* and *Bacteria*. *Proc Natl Acad Sci U S A*. 1990;87:4576–9.
- Garrity GM, Bell JA, Lilburn T. Phylum XIV. *Proteobacteria* phyl nov. In: Brenner DJ, Krieg NR, Stanley JT, Garrity GM, editors. *Bergey's Manual of*

- Systematic Bacteriology*, volume 2 (The Proteobacteria part B The Gammaproteobacteria). 2nd ed. New York: Springer; 2005. p. 1.
45. Garrity GM, Bell JA, Lilburn T. Class II. *Betaproteobacteria* class. nov. In: Brenner DJ, Krieg NR, Stanley JT, Garrity GM, editors. *Bergey's Manual of Systematic Bacteriology*. Volume 2 (the Proteobacteria part C the alpha-, beta-, delta-, and Epsilonproteobacteria). 2nd ed. New York: Springer; 2005. p. 575.
 46. Garrity GM, Bell JA, Lilburn T. Order I. *Burkholderiales* ord. nov. In: Brenner DJ, Krieg NR, Stanley JT, Garrity GM, Brenner DJ, Krieg NR, Stanley JT, Garrity GM, editors. *Bergey's Manual of Systematic Bacteriology. Volume 2 (The Proteobacteria part C The Alpha-, Beta-, Delta-, and Epsilonproteobacteria)*. 2nd ed. New York: Springer; 2005. p. 575.
 47. Willems A, De Ley J, Gillis M, Kersters K. *Comamonadaceae*, a new family encompassing the acidovorans rRNA complex, including *Variovorax paradoxus* gen. nov., comb. nov., for *Alcaligenes paradoxus* (Davis 1969). *Int J Syst Bacteriol.* 1991;41:445–50.
 48. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *Gene Ontology Consortium.* *Nat Genet.* 2000;25:25–9.

**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

