

Complete genome sequence of *Deinococcus maricopensis* type strain (LB-34^T)

Rüdiger Pukall¹, Ahmet Zeytun^{2,3}, Susan Lucas², Alla Lapidus², Nancy Hammon², Shweta Deshpande², Matt Nolan², Jan-Fang Cheng², Sam Pitluck², Konstantinos Liolios², Ioanna Pagani², Natalia Mikhailova², Natalia Ivanova², Konstantinos Mavromatis², Amrita Pati², Roxane Tapia^{2,3}, Cliff Han^{2,3}, Lynne Goodwin^{2,3}, Amy Chen⁴, Krishna Palaniappan⁴, Miriam Land^{2,5}, Loren Hauser^{2,5}, Yun-Juan Chang^{2,5}, Cynthia D. Jeffries^{2,5}, Evelyne-Marie Brambilla¹, Manfred Rohde⁶, Markus Göker¹, J. Chris Detter^{2,3}, Tanja Woyke², James Bristow², Jonathan A. Eisen^{2,7}, Victor Markowitz⁴, Philip Hugenholtz^{2,8}, Nikos C. Kyrpides², and Hans-Peter Klenk^{1*}

¹ DSMZ – German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany

² DOE Joint Genome Institute, Walnut Creek, California, USA

³ Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico USA

⁴ Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, California, USA

⁵ Lawrence Livermore National Laboratory, Livermore, California, USA

⁶ HZI – Helmholtz Centre for Infection Research, Braunschweig, Germany

⁷ University of California Davis Genome Center, Davis, California, USA

⁸ Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Australia

*Corresponding author: Hans-Peter Klenk

Keywords: aerobic, non-motile, Gram-positive, radiation-resistant, mesophilic, chemoorganotrophic, *Deinococcaceae*, GEBA

Deinococcus maricopensis (Rainey and da Costa 2005) is a member of the genus *Deinococcus*, which is comprised of 44 validly named species and is located within the deeply branching bacterial phylum *Deinococcus–Thermus*. Strain LB-34^T was isolated from a soil sample from the Sonoran Desert in Arizona. Various species of the genus *Deinococcus* are characterized by extreme radiation resistance, with *D. maricopensis* being resistant in excess of 10 kGy. Even though the genomes of three *Deinococcus* species, *D. radiodurans*, *D. geothermalis* and *D. deserti*, have already been published, no special physiological characteristic is currently known that is unique to this group. It is therefore of special interest to analyze the genomes of additional species of the genus *Deinococcus* to better understand how these species adapted to gamma- or UV ionizing-radiation. The 3,498,530 bp long genome of *D. maricopensis* with its 3,301 protein-coding and 66 RNA genes consists of one circular chromosome and is a part of the **Genomic Encyclopedia of Bacteria and Archaea** project.

Introduction

Strain LB-34^T (= DSM 21211 = NRRL B-23946 = LMG 22137) is the type strain of *Deinococcus maricopensis* [1]. In addition to the type strain LB-34^T, two more strains of this species, KR 1 and KR 23, were characterized by Rainey *et al.* [1]. The generic name derives from the Greek words ‘deinos’ meaning ‘strange or unusual’ and ‘coccus’ meaning ‘a grain or berry’ [2]. The species epithet

is derived from the Neo-Latin word ‘maricopensis’ referring to the Maricopa Nation, a native tribe in Arizona [1]. Strain LB 34^T was isolated from desert soil in Arizona and described by Rainey *et al.* in 2005 [1]. The genus *Deinococcus* was proposed in 1981 by Brooks and Murray [2] to separate the distinct radiation-resistant species from the genus *Micrococcus* in which those species

were originally classified. With the description of *Deinobacter grandis* by Oyaizu *et al.* [3], a second genus was placed to the family *Deinococcaceae*, and in 1997 Rainey *et al.* proposed to transfer *Deinobacter* to the genus *Deinococcus*, based on investigations of the phylogenetic diversity of the *Deinococci* as determined by 16S rRNA gene sequence analysis. In conclusion, an emended description of the genus *Deinococcus* was published, showing that the cells can be spherical or rod-shaped [4]. Members of the genus *Deinococcus* were isolated from various environmental habitats including air [5-7], arid soil [1,8-12], water and activated sludge [13-15], alpine environments [16], rhizosphere [17], Antarctica [18], hot springs [19], aquifer [20], marine fish [21] and radioactive sites [22]. Here we present a summary classification and a set of features for *D. maricopensis* LB-34^T, together with the description of the complete genomic sequencing and annotation.

Classification and features

A representative genomic 16S rRNA sequence of strain LB-34^T was compared using NCBI BLAST under default settings (e.g., considering only the high-scoring segment pairs (HSPs) from the best 250 hits) with the most recent release of the Greengenes database [23] and the relative frequencies, weighted by BLAST scores, of taxa and keywords (reduced to their stem [24]) were determined. The single most frequent genus was *Deinococcus* (100.0%) (114 hits in total). Regarding the three hits to sequences from members of the species, the average identity within HSPs was 99.9%, whereas the average coverage by HSPs was 97.6%. Regarding the 77 hits to sequences from other members of the genus, the average identity within HSPs was 91.5%, whereas the average coverage by HSPs was 60.5%. Among all other species, the one yielding the highest score was *D. radiodurans*, which corresponded to an identity of 91.2% and an HSP coverage of 88.0%. The highest-scoring environmental sequence was AY905380 ('Extensive ionizing-radiation-resistant recovered sonoran and description nine new species genus *Deinococcus* obtained single mixed agricultural/open desert soil clone L14-471'), which showed an identity of 98.1% and a HSP coverage of 70.2%. The five most frequent keywords within the labels of environmental samples which yielded hits were 'skin' (7.7%), 'litholog/stream' (2.8%), 'fossa' (2.4%), 'microbi' (2.4%) and 'forearm' (2.1%) (136 hits in total).

Environmental samples which yielded hits of a higher score than the highest scoring species were not found.

Figure 1 shows the phylogenetic neighborhood of *D. maricopensis* LB-34^T in a 16S rRNA based tree. The sequences of the four identical 16S rRNA gene copies in the genome differ by one nucleotide from the previously published 16S rRNA sequence (AY743274).

The cells of *D. maricopensis* are rod-shaped, up to 6 µm in length and 2.0 µm wide (Figure 2). *D. maricopensis* is a Gram-positive, non-spore-forming bacterium (Table 1). Colonies on Rich medium are orange to pink. The cells are non-motile. The organism is chemoorganotrophic [1]. The temperature range for growth is 10° to 45°C, with an optimum at 40°C [1]. Cytochrome oxidase and catalase activity have been observed [1]. Strains may utilize L-arabinose, cellobiose, galactose, glucose, mannose, maltose, sucrose, trehalose, glucosamine, glycerol, malate, asparagine, aspartate, glutamate, L-glutamine, ornithine and proline. Fructose can be used by strain KR23, but not by strain LB-34^T [1]. Strain LB-34^T showed similar levels of desiccation tolerance of up to four weeks as compared to *D. radiodurans* strain R1^T. Strain LB-34^T is resistant to > 10kGy, but more sensitive to ionizing radiation than strain *D. radiodurans* R1^T [1].

Chemotaxonomy

The major cellular fatty acids of the strain LB-34^T were identified as *iso*-C_{15:0}, *iso*-C_{17:0} and C_{16:0}. Menquinone 8 (MK-8) was determined as the major respiratory quinone of the strain. Phosphoglycolipid and glycolipid pattern are similar to those of other *Deinococcus* species [1]. No data are available for strain LB-34^T showing the peptidoglycan type of the cell wall.

Genome sequencing and annotation

Genome project history

This organism was selected for sequencing on the basis of its phylogenetic position [45], and is part of the *Genomic Encyclopedia of Bacteria and Archaea* project [46]. The genome project is deposited in the Genomes On Line Database [29] and the complete genome sequence is deposited in GenBank. Sequencing, finishing and annotation were performed by the DOE Joint Genome Institute (JGI). A summary of the project information is shown in Table 2.

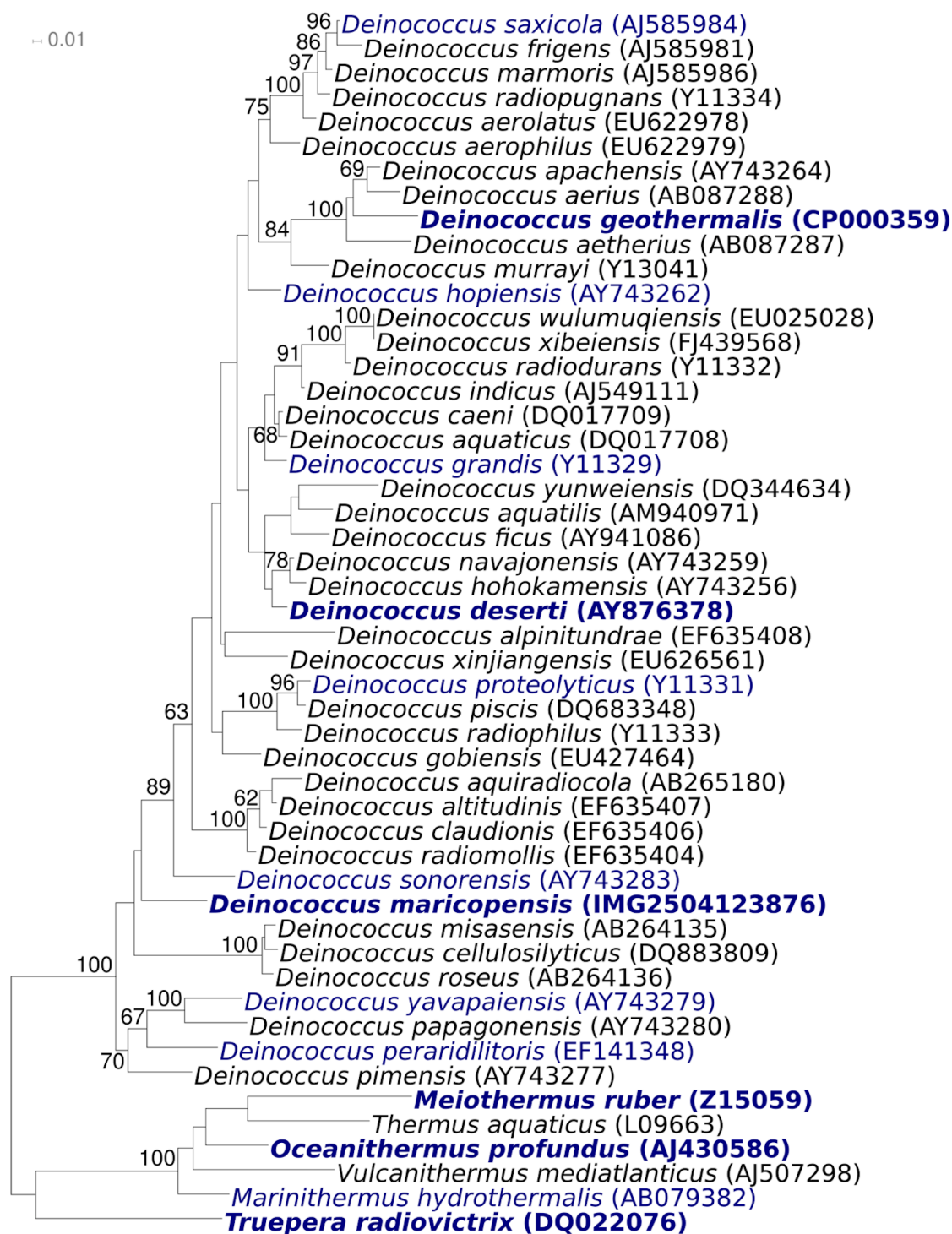


Figure 1. Phylogenetic tree highlighting the position of *D. maricopensis* relative to the other type strains within the family *Deinococcaceae*. The tree was inferred from 1,382 aligned characters [25,26] of the 16S rRNA gene sequence under the maximum likelihood criterion [27] and rooted in accordance with the current taxonomy. The branches are scaled in terms of the expected number of substitutions per site. Numbers above branches are support values from 1,000 bootstrap replicates [28] if larger than 60%. Lineages with type strain genome sequencing projects registered in GOLD [29] are shown in blue, and published genomes in bold [30-34]. The genome of *D. radiodurans* published by White *et al.* in 1999 [35] later turned out not to be from the type strain [36].

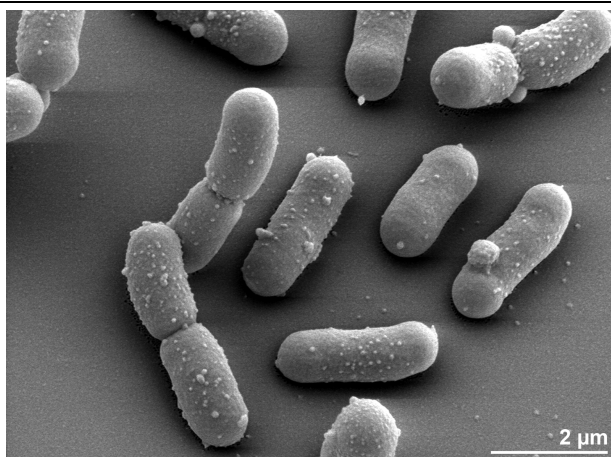


Figure 2. Scanning electron micrograph of *D. maricopensis* LB-34^T

Table 1. Classification and general features of *D. maricopensis* LB-34^T according to the MGS recommendations [37].

MIGS ID	Property	Term	Evidence code
		Domain <i>Bacteria</i>	TAS [38]
		Phylum <i>Deinococcus-Thermus</i>	TAS [39]
		Class <i>Deinococci</i>	TAS [40,41]
	Current classification	Order <i>Deinococcales</i>	TAS [4]
		Family <i>Deinococcaceae</i>	TAS [2,4]
		Genus <i>Deinococcus</i>	TAS [2,4]
		Species <i>Deinococcus maricopensis</i>	TAS [1,42]
		Type strain LB-34	TAS [1]
	Gram stain	positive	TAS [1]
	Cell shape	rods	TAS [1]
	Motility	non-motile	TAS [1]
	Sporulation	none	TAS [1]
	Temperature range	mesophile, 10°C–45°C	TAS [1]
	Optimum temperature	40°C	TAS [1]
	Salinity	not reported	
MIGS-22	Oxygen requirement	aerobic	TAS [1]
	Carbon source	carbohydrates	TAS [1]
	Energy metabolism	chemoorganotroph	TAS [1,2]
MIGS-6	Habitat	soil	TAS [1]
MIGS-15	Biotic relationship	free-living	NAS
MIGS-14	Pathogenicity	none	NAS
	Biosafety level	1	TAS [43]
	Isolation	soil	TAS [1]
MIGS-4	Geographic location	Sonoran Desert, Arizona, USA	TAS [1]
MIGS-5	Sample collection time	1999	NAS
MIGS-4.1	Latitude	32.93	NAS
MIGS-4.2	Longitude	-112.30	NAS
MIGS-4.3	Depth	not reported	
MIGS-4.4	Altitude	not reported	

Evidence codes - IDA: Inferred from Direct Assay (first time in publication); 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 of the Gene Ontology project [44]. If the evidence code is IDA, then the property was directly observed by one of the authors or an expert mentioned in the acknowledgements.

Table 2. Genome sequencing project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	Finished
MIGS-28	Libraries used	Three genomic libraries: one 454 pyrosequence standard library, one 454 PE library (7 kb insert size), one Illumina library
MIGS-29	Sequencing platforms	Illumina GAii, 454 GS FLX Titanium
MIGS-31.2	Sequencing coverage	170.9 × Illumina; 75.4 × pyrosequence
MIGS-30	Assemblers	Newbler version 2.3-PreRelease-10-21-2009-gcc-4.1.2-threads, Velvet version 0.7.63, phrap
MIGS-32	Gene calling method	Prodigal 1.4, GenePRIMP
	INSDC ID	CP002454
	Genbank Date of Release	January 20, 2011
	GOLD ID	Gc01597
	NCBI project ID	43461
	Database: IMG-GEBA	2503982045
MIGS-13	Source material identifier	DSM 21211
	Project relevance	Tree of Life, GEBA

Growth conditions and DNA isolation

D. maricopenis LB-34^T, DSM 21211, was grown in DSMZ medium 736 (Rich Medium) [47] at 28°C. DNA was isolated from 0.5-1 g of cell paste using MasterPure Gram-positive DNA purification kit (Epicentre MGP04100) following the standard protocol as recommended by the manufacturer, with a modification in cell lysis by adding 20 µl lysozyme (100 mg/µl), and 10 µl mutanolysine, achromopeptidase and lysosthphine, each, for 40 min at 37°C, followed by one hour incubation on ice after the MPC step. DNA is available through the DNA Bank Network [48,49].

Genome sequencing and assembly

The genome was sequenced using a combination of Illumina and 454 sequencing platforms. All general aspects of library construction and sequencing can be found at the JGI website [50]. Pyrosequencing reads were assembled using the Newbler assembler version 2.3 (Roche). The initial Newbler assembly consisting of 58 contigs in two scaffolds was converted into a phrap assembly by [51] making fake reads from the consensus, to collect the read pairs in the 454 paired end library. Illumina GAii sequencing data (957.8 Mb) were assembled with Velvet version 0.7.63 [52] and the consensus sequences were shredded into 1.5 kb overlapped fake reads and assembled together with the 454 data. The 454 draft assembly was based on 234.5 Mb 454 draft data and all of the 454 paired end data. Newbler parameters are -consed -a 50 -l 350 -g -m -ml 20. The Phred/Phrap/Consed software package [51] was used for sequence assembly and quality

assessment in the subsequent finishing process. After the shotgun stage, reads were assembled with parallel phrap (High Performance Software, LLC). Possible mis-assemblies were corrected with gap-Resolution [50], Dupfinisher [53], or sequencing cloned bridging PCR fragments with subcloning or transposon bombing (Epicentre Biotechnologies, Madison, WI). Gaps between contigs were closed by editing in Consed, by PCR and by Bubble PCR primer walks (J.-F.Chang, unpublished). A total of 255 additional reactions were necessary to close gaps and to raise the quality of the finished sequence. Illumina reads were also used to correct potential base errors and increase consensus quality using a software Polisher developed at JGI [54]. The error rate of the completed genome sequence is less than 1 in 100,000. Together, the combination of the Illumina and 454 sequencing platforms provided 246.3 × coverage of the genome. The final assembly contained 872,337 pyrosequence and 16,604,657 Illumina reads.

Genome annotation

Genes were identified using Prodigal [55] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [56]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant database, UniProt, TIGR-Fam, Pfam, PRIAM, KEGG, COG, and InterPro databases. Additional gene prediction analysis and functional annotation was performed

within the Integrated Microbial Genomes - Expert Review (IMG-ER) platform [57].

Genome properties

The genome consists of a 3,498,530 bp long chromosome with a G+C content of 69.8% (Figure 3 and Table 3). Of the 3,367 genes predicted, 3,301 were protein-coding genes, and 66 RNAs; 37

pseudogenes were also identified. The majority of the protein-coding genes (70.3%) were assigned with 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.

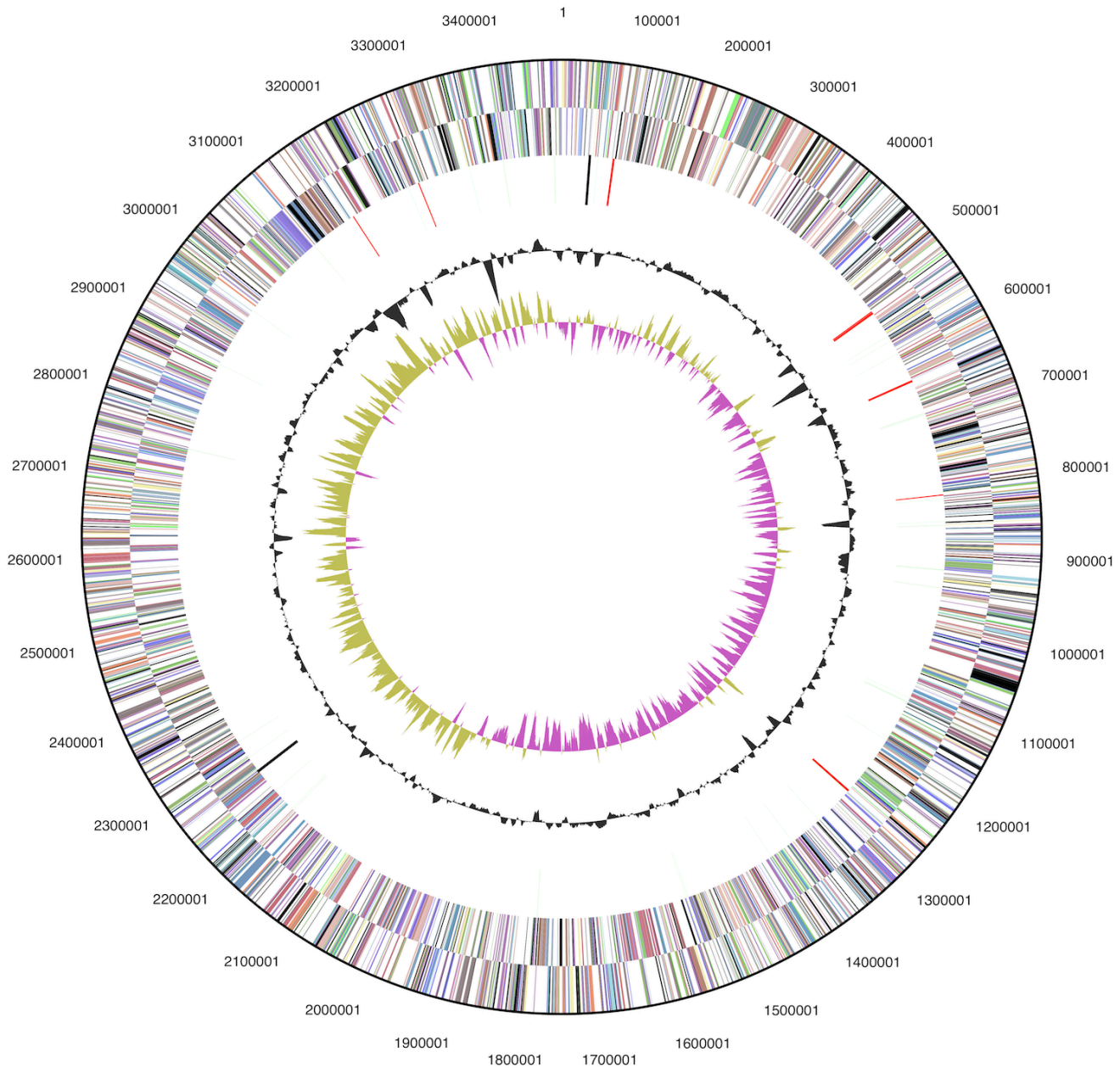


Figure 3. Graphical circular map of the chromosome. From outside to the center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

Table 3. Genome Statistics

Attribute	Value	% of Total
Genome size (bp)	3,498,530	100.00%
DNA coding region (bp)	3,127,041	89.38%
DNA G+C content (bp)	2,442,849	69.83%
Number of replicons	1	
Extrachromosomal elements	0	
Total genes	3,367	100.00%
RNA genes	66	1.96%
rRNA operons	4	
Protein-coding genes	3,301	98.04%
Pseudo genes	37	1.10%
Genes with function prediction	2,366	70.27%
Genes in paralog clusters	368	10.93%
Genes assigned to COGs	2,412	71.64%
Genes assigned Pfam domains	2,495	74.10%
Genes with signal peptides	1,005	29.85%
Genes with transmembrane helices	662	19.66%
CRISPR repeats	0	

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

Code	value	%age	Description
J	160	6.0	Translation, ribosomal structure and biogenesis
A	0	0.0	RNA processing and modification
K	188	7.1	Transcription
L	109	4.1	Replication, recombination and repair
B	2	0.1	Chromatin structure and dynamics
D	29	1.1	Cell cycle control, cell division, chromosome partitioning
Y	0	0.0	Nuclear structure
V	45	1.7	Defense mechanisms
T	195	7.3	Signal transduction mechanisms
M	137	5.2	Cell wall/membrane/envelope biogenesis
N	15	0.6	Cell motility
Z	1	0.0	Cytoskeleton
W	0	0.0	Extracellular structures
U	43	1.6	Intracellular trafficking, secretion, and vesicular transport
O	113	4.3	Posttranslational modification, protein turnover, chaperones
C	125	4.7	Energy production and conversion
G	205	7.7	Carbohydrate transport and metabolism
E	237	8.9	Amino acid transport and metabolism
F	77	2.9	Nucleotide transport and metabolism
H	119	4.5	Coenzyme transport and metabolism
I	105	4.0	Lipid transport and metabolism
P	121	4.6	Inorganic ion transport and metabolism
Q	60	2.3	Secondary metabolites biosynthesis, transport and catabolism
R	334	12.6	General function prediction only
S	238	9.0	Function unknown
-	955	28.4	Not in COGs

Acknowledgements

We would like to gratefully acknowledge the help of Gabriele Gehrich-Schröter (DSMZ) for growing *D. maricopensis* cultures. This work was performed under the auspices of the US Department of Energy 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, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396, UT-Battelle and Oak Ridge National Laboratory under contract DE-AC05-00OR22725, as well as German Research Foundation (DFG) INST 599/1-2.

References

1. Rainey FA, Ray K, Ferreira M, Gatz BZ, Nobre MF, Bagaley D, Rash BA, Park NJ, Earl AM, Shank NC, et al. Extensive diversity of ionizing-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus *Deinococcus* obtained from a single soil sample. *Appl Environ Microbiol* 2005; **71**:5225-5235. [PubMed](#) [doi:10.1128/AEM.71.9.5225-5235.2005](https://doi.org/10.1128/AEM.71.9.5225-5235.2005)
2. Brooks BW, Murray RGE. Nomenclature for "Micrococcus radiodurans" and other radiation-resistant cocci: *Deinococcaceae* fam. nov. and *Deinococcus* gen. nov., including five species. *Int J Syst Bacteriol* 1981; **31**:353-360. [doi:10.1099/00207713-31-3-353](#)
3. Oyaizu H, Stackebrandt E, Schleifer KH, Ludwig W, Pohla H, Ito H, Hirata A, Oyaizu Y, Komagata K. A radiation-resistant rod-shaped bacterium, *Deinobacter grandis* gen. nov., sp. nov., with peptidoglycan containing ornithine. *Int J Syst Bacteriol* 1987; **37**:62-67. [doi:10.1099/00207713-37-1-62](#)
4. Rainey FA, Nobre MF, Schumann P, Stackebrandt E, da Costa MS. Phylogenetic diversity of the *Deinococci* as determined by 16S ribosomal DNA sequence comparison. *Int J Syst Bacteriol* 1997; **47**:510-514. [PubMed](#) [doi:10.1099/00207713-47-2-510](https://doi.org/10.1099/00207713-47-2-510)
5. Weon HY, Kim BY, Schumann P, Son JA, Jang J, Go SJ, Kwon SW. *Deinococcus cellulosilyticus* sp. nov., isolated from air. *Int J Syst Evol Microbiol* 2007; **57**:1685-1688. [PubMed](#) [doi:10.1099/ijs.0.64951-0](https://doi.org/10.1099/ijs.0.64951-0)
6. Yang Y, Itoh T, Yokobori S, Itahashi S, Shimada H, Satoh K, Ohba H, Narumi I, Yamagishi A. *Deinococcus aerius* sp. nov., isolated from the high atmosphere. *Int J Syst Evol Microbiol* 2009; **59**:1862-1866. [PubMed](#) [doi:10.1099/ijs.0.007963-0](https://doi.org/10.1099/ijs.0.007963-0)
7. Yoo SH, Weon HY, Kim SJ, Kim YS, Kim BY, Kwon SW. *Deinococcus aerolatus* sp. nov. and *Deinococcus aerophilus* sp. nov., isolated from air samples. *Int J Syst Evol Microbiol* 2010; **60**:1191-1195. [PubMed](#) [doi:10.1099/ijs.0.016030-0](https://doi.org/10.1099/ijs.0.016030-0)
8. Rainey FA, Ferreira M, Nobre MF, Ray K, Bagaley D, Earl AM, Battista JR, Gómez-Silva B, McKay CP, da Costa MS. *Deinococcus peraridilitoris* sp. nov., isolated from a coastal desert. *Int J Syst Evol Microbiol* 2007; **57**:1408-1412. [PubMed](#) [doi:10.1099/ijs.0.64956-0](https://doi.org/10.1099/ijs.0.64956-0)
9. de Groot A, Chapon V, Servant P, Christen R, Saux MF, Sommer S, Heulin T. *Deinococcus deserti* sp. nov., a gamma-radiation-tolerant bacterium isolated from the Sahara Desert. *Int J Syst Evol Microbiol* 2005; **55**:2441-2446. [PubMed](#) [doi:10.1099/ijs.0.63717-0](https://doi.org/10.1099/ijs.0.63717-0)
10. Peng F, Zhang L, Luo X, Dai J, An H, Tang Y, Fang C. *Deinococcus xinjiangensis* sp. nov., isolated from desert soil. *Int J Syst Evol Microbiol* 2009; **59**:709-713. [PubMed](#) [doi:10.1099/ijs.0.004564-0](https://doi.org/10.1099/ijs.0.004564-0)
11. Yuan M, Zhang W, Dai S, Wu J, Wang Y, Tao T, Chen M, Lin M. *Deinococcus gobiensis* sp. nov., an extremely radiation-resistant bacterium. *Int J Syst Evol Microbiol* 2009; **59**:1513-1517. [PubMed](#) [doi:10.1099/ijs.0.004523-0](https://doi.org/10.1099/ijs.0.004523-0)
12. Wang W, Mao J, Zhang Z, Tang Q, Xie Y, Zhu J, Zhang L, Liu Z, Shi Y, Goodfellow M. *Deinococcus wulumuqiensis* sp. nov., and *Deinococcus xibeiensis* sp. nov., isolated from radiation-polluted soil. *Int J Syst Evol Microbiol* 2010; **60**:2006-2010. [PubMed](#) [doi:10.1099/ijs.0.015917-0](https://doi.org/10.1099/ijs.0.015917-0)
13. Im WT, Jung HM, Ten LN, Kim MK, Bora N, Goodfellow M, Lim S, Jung J, Lee ST. *Deinococcus aquaticus* sp. nov., isolated from fresh water, and *Deinococcus caeni* sp. nov., isolated from activated sludge. *Int J Syst Evol Microbiol* 2008; **58**:2348-2353. [PubMed](#) [doi:10.1099/ijs.0.64082-0](https://doi.org/10.1099/ijs.0.64082-0)
14. Kämpfer P, Lodders N, Huber B, Falsen E, Busse HJ. *Deinococcus aquatilis* sp. nov., isolated from water. *Int J Syst Evol Microbiol* 2008; **58**:2803-2806. [PubMed](#) [doi:10.1099/ijs.0.2008/001206-0](https://doi.org/10.1099/ijs.0.2008/001206-0)
15. Asker D, Awad TS, Beppu T, Ueda K. *Deinococcus aquiradiocola* sp. nov., isolated from a radioactive site in Japan. *Int J Syst Evol Microbiol* 2009; **59**:144-149. [PubMed](#) [doi:10.1099/ijs.0.65762-0](https://doi.org/10.1099/ijs.0.65762-0)
16. Callegan RP, Nobre MF, McTernan PM, Battista JR, Navarro-González R, McKay CP, da Costa MS, Rainey FA. Description of four novel psychrophilic, ionizing radiation-sensitive *Deinococcus* species from

- alpine environments. *Int J Syst Evol Microbiol* 2008; **58**:1252-1258. [PubMed](#) [doi:10.1099/ijs.0.65405-0](https://doi.org/10.1099/ijs.0.65405-0)
17. Lai WA, Kämpfer P, Arun AB, Shen FT, Huber B, Rekha PD, Young CC. *Deinococcus ficus* sp. nov., isolated from the rhizosphere of *Ficus religiosa* L. *Int J Syst Evol Microbiol* 2006; **56**:787-791. [PubMed](#) [doi:10.1099/ijs.0.64007-0](https://doi.org/10.1099/ijs.0.64007-0)
 18. Hirsch P, Gallikowski CA, Siebert J, Peissl K, Kropfenstedt R, Schumann P, Stackebrandt E, Anderson R. *Deinococcus frigens* sp. nov., *Deinococcus saxicola* sp. nov., and *Deinococcus marmoris* sp. nov., low temperature and draught-tolerating, UV-resistant bacteria from continental Antarctica. *Syst Appl Microbiol* 2004; **27**:636-645. [PubMed](#) [doi:10.1078/0723202042370008](https://doi.org/10.1078/0723202042370008)
 19. Ferreira AC, Nobre MF, Rainey FA, Silva MT, Wait R, Burghardt J, Chung AP, da Costa MS. *Deinococcus geothermalis* sp. nov. and *Deinococcus murrayi* sp. nov., two extremely radiation-resistant and slightly thermophilic species from hot springs. *Int J Syst Bacteriol* 1997; **47**:939-947. [PubMed](#) [doi:10.1099/00207713-47-4-939](https://doi.org/10.1099/00207713-47-4-939)
 20. Suresh K, Reddy GS, Sengupta S, Shivaji S. *Deinococcus indicus* sp. nov., an arsenic-resistant bacterium from an aquifer in West Bengal, India. *Int J Syst Evol Microbiol* 2004; **54**:457-461. [PubMed](#) [doi:10.1099/ijs.0.02758-0](https://doi.org/10.1099/ijs.0.02758-0)
 21. Shashidhar R, Bandekar JR. *Deinococcus piscis* sp. nov., a radiation-resistant bacterium isolated from a marine fish. *Int J Syst Evol Microbiol* 2009; **59**:2714-2717. [PubMed](#) [doi:10.1099/ijs.0.003046-0](https://doi.org/10.1099/ijs.0.003046-0)
 22. Asker D, Awad TS, Beppu T, Ueda K. *Deinococcus misasensis* and *Deinococcus roseus*, novel members of the genus *Deinococcus*, isolated from a radioactive site in Japan. *Syst Appl Microbiol* 2008; **31**:43-49. [PubMed](#) [doi:10.1016/j.syapm.2007.10.002](https://doi.org/10.1016/j.syapm.2007.10.002)
 23. 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**:5069-5072. [PubMed](#) [doi:10.1128/AEM.03006-05](https://doi.org/10.1128/AEM.03006-05)
 24. Porter MF. An algorithm for suffix stripping. Program: electronic library and information systems 1980; **14**:130-137.
 25. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000; **17**:540-552. [PubMed](#)
 26. Lee C, Grasso C, Sharlow MF. Multiple sequence alignment using partial order graphs. *Bioinformatics* 2002; **18**:452-464. [PubMed](#) [doi:10.1093/bioinformatics/18.3.452](https://doi.org/10.1093/bioinformatics/18.3.452)
 27. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 2008; **57**:758-771. [PubMed](#) [doi:10.1080/10635150802429642](https://doi.org/10.1080/10635150802429642)
 28. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. How many bootstrap replicates are necessary? *Lect Notes Comput Sci* 2009; **5541**:184-200. [doi:10.1007/978-3-642-02008-7_13](#)
 29. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyripides NC. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* 2009; **38**:D346-D354. [PubMed](#) [doi:10.1093/nar/gkp848](https://doi.org/10.1093/nar/gkp848)
 30. de Groot A, Dulermo R, Ortet P, Blanchard L, Guerin P, Fernandez B, Vacherie B, Dossat C, Jolivet E, Siguier P, et al. Alliance of proteomics and genomics to unravel the specificities of Sahara bacterium *Deinococcus deserti*. *PLoS Genet* 2009; **5**:e1000434. [PubMed](#) [doi:10.1371/journal.pgen.1000434](https://doi.org/10.1371/journal.pgen.1000434)
 31. Makarova KS, Omelchenko MV, Gaidamakova EK, Matrosova VY, Vasilenko A, Zhai M, Lapidus A, Copeland A, Kim E, Land M, et al. *Deinococcus geothermalis*: the pool of extreme radiation resistance genes shrinks. *PLoS ONE* 2007; **2**:e955. [PubMed](#) [doi:10.1371/journal.pone.0000955](https://doi.org/10.1371/journal.pone.0000955)
 32. Ivanova N, Rhode C, Munk C, Nolan M, Lucas S, Glavina Del Rio T, Tice H, Deshpande S, Cheng JF, Tapia R, et al. Complete genome sequence of *Truopera radiovictrix* type strain (RQ-24^T). *Stand Genomic Sci* 2011; **4**:91-99. [PubMed](#) [doi:10.4056/sigs.1563919](https://doi.org/10.4056/sigs.1563919)
 33. Tindall BJ, Sikorski J, Lucas S, Goltsman E, Copeland A, Glavina Del Rio T, Nolan M, Tice H, Cheng JF, Han C, et al. Complete genome sequence of *Meiothermus ruber* type strain (21^T). *Stand Genomic Sci* 2010; **3**:26-36. [PubMed](#) [doi:10.4056/sigs.1032748](https://doi.org/10.4056/sigs.1032748)
 34. Pati A, Zhang X, Lapidus A, Nolan M, Lucas S, Glavina Del Rio T, Tice H, Cheng JF, Tapia R, Han C, et al. Complete genome sequence of *Oceanithermus profundus* type strain (506^T). *Stand Genomic Sci* 2011; **4**:210-220. [doi:10.4056/sigs.1513795](#)
 35. White O, Eisen JA, Heidelberg JF, Hickey EK, Peterson JD, Dodson RJ, Haft DH, Gwinn ML, Nelson WC, Richardson DL, et al. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science* 1999; **286**:1571-1577. [PubMed](#) [doi:10.1126/science.286.5444.1571](https://doi.org/10.1126/science.286.5444.1571)

36. Corrections and Clarifications. [Erratum: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.]. *Science* 2004; **303**:766b. [doi:10.1126/science.303.5659.766b](https://doi.org/10.1126/science.303.5659.766b)
37. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547. [PubMed](https://pubmed.ncbi.nlm.nih.gov/16811111/) [doi:10.1038/nbt1360](https://doi.org/10.1038/nbt1360)
38. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. [PubMed](https://pubmed.ncbi.nlm.nih.gov/1712100/) [doi:10.1073/pnas.87.12.4576](https://doi.org/10.1073/pnas.87.12.4576)
39. Weisburg WG, Giovannoni SJ, Woese CR. The *Deinococcus-Thermus* phylum and the effect of rRNA composition on phylogenetic tree construction. *Syst Appl Microbiol* 1989; **11**:128-134. [PubMed](https://pubmed.ncbi.nlm.nih.gov/1712100/)
40. List Editor. Validation List no. 85. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. *Int J Syst Evol Microbiol* 2002; **52**:685-690. [PubMed](https://pubmed.ncbi.nlm.nih.gov/12111111/) [doi:10.1099/ijs.0.02358-0](https://doi.org/10.1099/ijs.0.02358-0)
41. Garrity GM, Holt JG. Class I. *Deinococci* class. nov. In: Garrity GM, Boone DR, Castenholz RW (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 1, Springer, New York, 2001, p. 395.
42. List Editor. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. List no. 106. *Int J Syst Evol Microbiol* 2005; **55**:2235-2238. [PubMed](https://pubmed.ncbi.nlm.nih.gov/16111111/) [doi:10.1099/ijs.0.64108-0](https://doi.org/10.1099/ijs.0.64108-0)
43. Classification of bacteria and archaea in risk groups. TRBA 466, <http://www.baua.de>
44. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene Ontology: tool for the unification of biology. *Nat Genet* 2000; **25**:25-29. [PubMed](https://pubmed.ncbi.nlm.nih.gov/10111111/) [doi:10.1038/75556](https://doi.org/10.1038/75556)
45. Klenk HP, Göker M. En route to a genome-based classification of *Archaea* and *Bacteria*? *Syst Appl Microbiol* 2010; **33**:175-182. [PubMed](https://pubmed.ncbi.nlm.nih.gov/20111111/) [doi:10.1016/j.syapm.2010.03.003](https://doi.org/10.1016/j.syapm.2010.03.003)
46. Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ, et al. A phylogeny-driven genomic encyclopaedia of *Bacteria* and *Archaea*. *Nature* 2009; **462**:1056-1060. [PubMed](https://pubmed.ncbi.nlm.nih.gov/19111111/) [doi:10.1038/nature08656](https://doi.org/10.1038/nature08656)
47. List of growth media used at DSMZ: http://www.dsmz.de/microorganisms/media_list.php.
48. Gemeinholzer B, Dröge G, Zetsche H, Haszprunar G, Klenk HP, Güntsch A, Berendsohn WG, Wägele JW. The DNA Bank Network: the start from a German initiative. *Biopreservation and Biobanking* (In press).
49. DNA Bank Network. <http://www.dnabank-network.org>
50. DOE Joint Genome Institute. <http://www.jgi.doe.gov>
51. Phrap and Phred for Windows, MacOS, Linux, and Unix. <http://www.phrap.com>
52. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008; **18**:821-829. [PubMed](https://pubmed.ncbi.nlm.nih.gov/18111111/) [doi:10.1101/gr.074492.107](https://doi.org/10.1101/gr.074492.107)
53. Han C, Chain P. 2006. Finishing repeat regions automatically with Dupfinisher. in Proceeding of the 2006 international conference on bioinformatics & computational biology. Edited by Hamid R. Arabnia & Homayoun Valafar, CSREA Press. June 26-29, 2006: 141-146.
54. Lapidus A, LaButti K, Foster B, Lowry S, Trong S, Goltsman E. POLISHER: An effective tool for using ultra short reads in microbial genome assembly and finishing. AGBT, Marco Island, FL, 2008.
55. 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. [PubMed](https://pubmed.ncbi.nlm.nih.gov/20111111/) [doi:10.1186/1471-2105-11-119](https://doi.org/10.1186/1471-2105-11-119)
56. 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**:455-457. [PubMed](https://pubmed.ncbi.nlm.nih.gov/20111111/) [doi:10.1038/nmeth.1457](https://doi.org/10.1038/nmeth.1457)
57. 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**:2271-2278. [PubMed](https://pubmed.ncbi.nlm.nih.gov/19111111/) [doi:10.1093/bioinformatics/btp393](https://doi.org/10.1093/bioinformatics/btp393)