Complete genome sequence of *Jonesia denitrificans* type strain (Prevot 55134[T])

Rüdiger Pukall[1], Gabriele Gehrich-Schröter[1], Alla Lapidus[2], Matt Nolan[2], Tijana Glavina Del Rio[2], Susan Lucas[2], Feng Chen[2], Hope Tice[2], Sam Pitluck[2], Jan-Fang Cheng[2], Alex Copeland[2], Elizabeth Saunders[2,3], Thomas Brettin[2,3], John C. Detter[2,3], David Bruce[2,3], Lynne Goodwin[2,3], Amrita Pati[2], Natalia Ivanova[2], Konstantinos Mavrommatis[2], Galina Ovchinnikova[2], Amy Chen[4], Krishna Palaniappan[4], Miriam Land[2,5], Loren Hauser[2,5], Yun-Juan Chang[2,5], Cynthia D. Jeffries[2,5], Patrick Chain[2,6], Markus Göker[1], Jim Bristow[2], Jonathan A. Eisen[2,7], Victor Markowitz[4], Philip Hugenholtz[2], Nikos C. Kyrpides[2], Hans-Peter Klenk[1*], and Cliff Han[2,3]

[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] Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
[6] Lawrence Livermore National Laboratory, Livermore, California, USA
[7] University of California Davis Genome Center, Davis, California, USA

*Corresponding author: Hans-Peter Klenk

Jonesia denitrificans (Prevot 1961) Rocourt *et al.* 1987 is the type species of the genus *Jonesia*, and is of phylogenetic interest because of its isolated location in the actinobacterial suborder *Micrococcineae*. *J. denitrificans* is characterized by a typical coryneform morphology and is able to form irregular nonsporulating rods showing branched and club-like forms. Coccoid cells occur in older cultures. *J. denitrificans* is classified as a pathogenic organism for animals (vertebrates). The type strain whose genome is described here was originally isolated from cooked ox blood. Here we describe the features of this organism, together with the complete genome sequence and annotation. This is the first completed genome sequence of a member of the genus for which a complete genome sequence is described. The 2,749,646 bp long genome with its 2558 protein-coding and 71 RNA genes is part of the **G**enomic **E**ncyclopedia of **B**acteria and **A**rchaea project.

Introduction

Strain Prevot 55134T (= DSM 20603 = ATCC 14870 = CIP 55.134) is the type strain of the species *Jonesia denitrificans*, the type species of the genus *Jonesia* [1]. The isolate originated from cooked ox blood [2]. *J. denitrificans* was originally placed into the genus *Listeria* and described as *L. denitrificans* by Prevot in 1961 [3], even though it differed morphologically from other members of the genus *Listeria.* Later, extensive analysis based on the determination of the G+C content [4], DNA-DNA hybridization [4], peptidoglycan type [5,6], as

well as fatty acids and polar lipid pattern [5,7] confirmed the misclassification of the strain. As a consequence the strain was transferred to the new genus *Jonesia* by Rocourt *et al.* 1987 [1]. Five years later Stackebrandt and Prauser assigned *Jonesia* to the family *Cellulomonadaceae* despite being aware that *Jonesia* differed from members of other genera of the family *e.g.* in its G+C content, peptidoglycan type and its composition of isoprenoid quinones. The rationale to place *Jonesia* into this family was based on the finding that the 16S rRNA

gene sequence analysis showed close relationship to *Promicromonospora* and *Cellulomonas* at that time [8]. Following extensive phylogenetic in 1995, studies *Jonesia* was subsequently excluded from the family *Cellulomonadacea* and placed in the family *Jonesiaceae,* within the suborder *Micrococcineae* [9]. With *Jonesia quinghaiensis*, an environmental isolate from mud of a soda lake in China, a second species of the genus was described by Schumann *et al.* in 2004 [10]. Two additional environmental strains closely related to J. denitrificans, with 98% and 99% 16S rRNA gene sequence similarity have been reported. These organisms were isolated from the microbial community of

feed batch reactors for composting of household biowaste [11]. Here we present a summary classification and a set of features for *J. denitrificans* Prevot 55134ᵀ together with the description of the complete genomic sequencing and annotation.

Classification and features

Figure 1 shows the phylogenetic neighborhood of *J. denitrificans* strain Prevot 55134ᵀ in a 16S rRNA based tree. The sequences of the five 16S rRNA gene copies in the genome of strain Prevot 55134ᵀ do not differ from each other, and differ by eight nucleotides from the previously published 16S rRNA gene sequence of DSM 20603 (X78420).
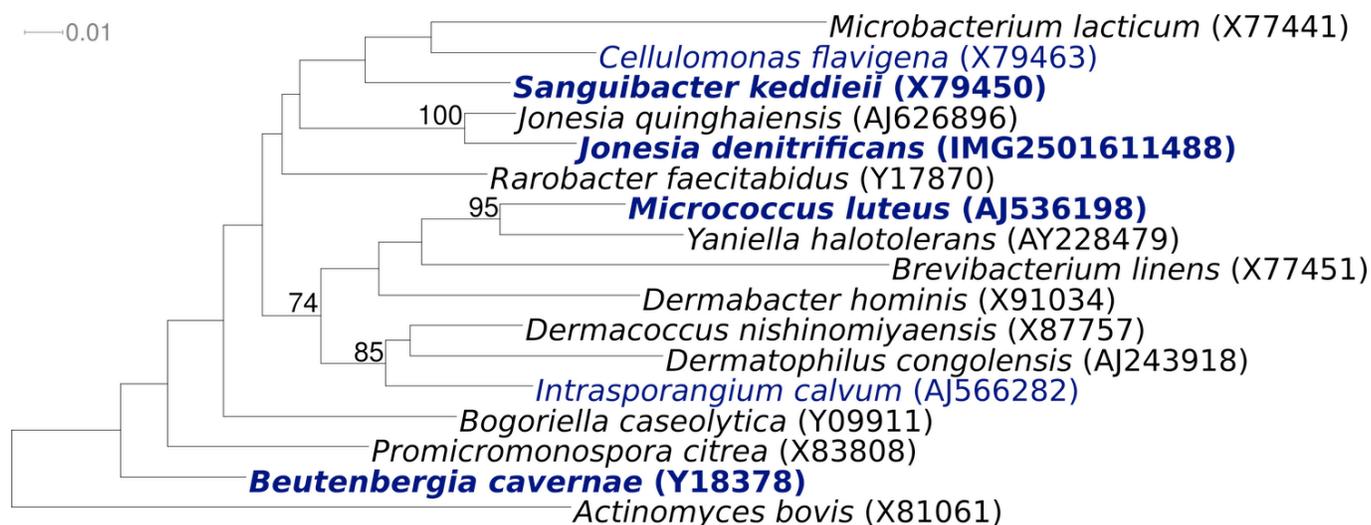
Figure 1. Phylogenetic tree highlighting the position of *J. denitrificans* Prevot 55134ᵀ relative to all type strains of the genus *Jonesia* and the type strains of all families within suborder *Micrococcineae,* inferred from 1,417 aligned characters [19,20] of the 16S rRNA sequence under the maximum likelihood criterion [21]. Rooting was done with the type strain of *Actinomyces bovis*, the type species of *Actinomyces*, which is the type genus of the order *Actinomycetales*. 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 if larger than 60%. Lineages with type strain genome sequencing projects registered in GOLD [22] are shown in blue, published genomes in bold, *e.g.* the GEBA genomes *Beutenbergia cavernae* [23], and *Sanguibacter keddieii* [24].

J. denitrificans type strain cells are Gram-positive, typically coryneform and characterized by irregular rods, 0.3-0.5 μm in diameter and 2-3 μm in length. Coccoid forms occur in older cultures (Table 1),. No endospores are formed. (Figure 2). The cells are motile by means of peritrichous flagella (absent in Figure 2). Colonies range from 0.5 to 1.5 mm in diameter on BHI agar. The optimum temperature for growth is 30°C [16]. The organism is catalase positive, oxidase negative and utilizes D-cellobiose, D-galactose, D-sorbitol, turanose and acetic acid [10]. Cellulose, starch, DNA and RNA

are hydrolyzed. In addition, *J. denitrificans* produces acid from a large variety of sugars, polysaccharides and other compounds as described by Seeliger and Jones in 1986 [16] and is capable of denitrification [25]. The natural habitat of the organism is not known, however, *J. denitrificans* is a known pathogen of rats and mice when injected intraperitoneally [16].

Chemotaxonomy

The cell wall of strain Prevot 55134ᵀ contains murein of type A4α, composed of L-Lys-L-Ser-D-

Glu only [5,6], type A11.48 according to the DSMZ catalogue of strains (http://www.dsmz.de/ microorganisms/main.php?content id=35). In addition to the amino sugars muramic acid and glucosamine, galactosamine was detectable in the hydrolysate of the cell walls of *J. denitrificans* [5]. 12-Methyl-tetradonic acid (ai-C15:0) and hexadonic acid (C16:0) constituted the major cellular fatty acid, and minor amounts of 14-methyl-hexadonic acid (ai-C17:0) and tetradecanoic acid (C14:0) were also detected [6,7]. Diphosphatidylglycerol (DPG) and phosphatidylinositol (PI) were identified by TLC as the polar lipids [6] and menaquinone of the MK-9 type was detected as the major component.
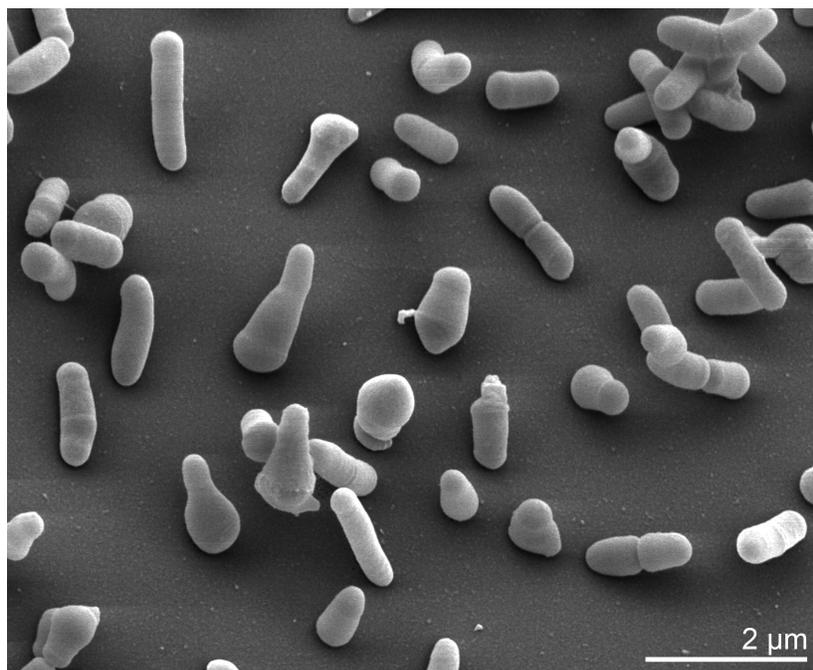
Figure 2. Scanning electron micrograph of *J. denitrificans* Prevot 55134[T] (Manfred Rohde, Helmholtz Centre for Infection Research (HZI), Braunschweig)

Table 1. Classification and general features of *J. denitrificans* Prevot 55134[T] in accordance with the MIGS recommendations [12]

MIGS ID	Property	Term	Evidence code
		Domain *Bacteria*	TAS [13]
		Phylum *Firmicutes*	TAS [14]
		Class *Actinobacteria*	TAS [15]
		Subclass *Actinobacteridae*	TAS [15]
		Order *Actinomycetales*	TAS [15]
		Suborder *Micrococcineae*	TAS [15]
		Family *Jonesiaceae*	TAS [15]
		Genus *Jonesia*	TAS [1]
		Species *Jonesia denitrificans*	TAS [1]
		Type strain Prevot 55134	TAS [16]
	Gram stain	positive	TAS [16]
	Cell shape	irregular rods, coccoid	TAS [16]
	Motility	motile	TAS [16]
	Sporulation	nonsporulating	TAS [16]
	Temperature range	mesophile	TAS [16]
	Optimum temperature	30-37°C	TAS [16]
	Salinity	5% NaCl	TAS [16]

Table 1. Classification and general features of *J. denitrificans* Prevot 55134[T] in accordance with the MIGS recommendations (cont.) [12]

MIGS ID	Property	Term	Evidence code
MIGS-22	Oxygen requirement	facultatively anaerobic	TAS [16]
	Carbon source	unknown	
	Energy source	unknown	
MIGS-6	Habitat	natural habitat not known	TAS [16]
MIGS-15	Biotic relationship	free living	NAS
MIGS-14	Pathogenicity	pathogenic for animals	TAS [16]
	Biosafety level	2	TAS [17]
	Isolation	cooked ox blood	TAS [2]
MIGS-4	Geographic location	France	NAS
MIGS-5	Sample collection time	not reported	
MIGS-4.1 MIGS-4.2	Latitude – Longitude	not reported	
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 the Gene Ontology project [18]. If the evidence code is IDA, then the property was directly observed for a living isolate by one of the authors or an expert mentioned in the acknowledgements.

Genome sequencing information
Genome project history

This organism was selected for sequencing on the basis of its phylogenetic position, and is part of the *Genomic Encyclopedia of Bacteria and Archaea* project. The genome project is deposited in the Genomes OnLine Database [22] and the complete genome sequence 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.

Table 2. Genome sequencing project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	Finished
MIGS-28	Libraries used	Two Sanger libraries –8 kb pMCL200 and fosmid pcc1Fos
MIGS-29	Sequencing platforms	ABI3730
MIGS-31.2	Sequencing coverage	9.5x Sanger
MIGS-30	Assemblers	phrap
MIGS-32	Gene calling method	Prodigal, GenePRIMP
	INSDC ID	CP001706
	Genbank Date of Release	August 27, 2009
	GOLD ID	Gc01092
	NCBI project ID	20833
	Database: IMG-GEBA	2501533218
MIGS-13	Source material identifier	DSM 20603
	Project relevance	Tree of Life, GEBA

Growth conditions and DNA isolation

J. denitrificans strain Prevot 55134[T], DSM 20603, was grown in DSMZ medium 215 (BHI broth) [26], at 37°C. DNA was isolated from 0.5-1 g of cell paste using the JGI CTAB-Protocol with a modified protocol for cell lysis (ALM), according to Wu *et al.* [27].

Genome sequencing and assembly

The genome was sequenced using only the Sanger platform. All general aspects of library construction and sequencing performed at the JGI can be found at the JGI website (http://www.jgi.doe.gov/). All reads were assembled using the phrap assembler. Possible mis-assemblies were corrected with Dupfinisher or transposon bombing of bridging clones [28]. Gaps between contigs were closed by editing in Consed, custom primer walk or PCR amplification. A total of 653 Sanger finishing reads were produced. The error rate of the completed genome sequence is less than 1 in 100,000. Together all sequenced reads provided 9.5x coverage of the genome. The final assembly consists of 35,028 Sanger 454 reads.

Genome annotation

Genes were identified using Prodigal [29] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline (http://geneprimp.jgi-psf.org) [30]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) non-redundant 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 [31].

Genome properties

The single replicon genome is 2,749,646 bp long with a 58.4% GC content (Table 3 and Figure 3). Of the 2,629 genes predicted, 2,558 were protein coding genes, and 71 RNAs. In addition, 47 pseudogenes were also identified. The majority of the genes (68.3%) were assigned with a putative function, while the remaining ones are annotated as hypothetical proteins. The distribution of genes into COGs functional categories is presented in Table 4.

Table 3. Genome Statistics

Attribute	Value	% of Total
Genome size (bp)	2,749,646	100.00%
DNA coding region (bp)	2,530,061	92.01%
DNA G+C content (bp)	1,606,252	58.42%
Number of replicons	1	
Extrachromosomal elements	0	
Total genes	2,629	100.00%
RNA genes	71	2.92%
rRNA operons	5	
Protein-coding genes	2,558	97.08%
Pseudo genes	47	1.78%
Genes with function prediction	1,796	68.31%
Genes in paralog clusters	223	8.46%
Genes assigned to COGs	1,830	69.45%
Genes assigned Pfam domains	1,849	70.17%
Genes with signal peptides	615	23.34%
Genes with transmembrane helices	682	25.88%
CRISPR repeats	0	
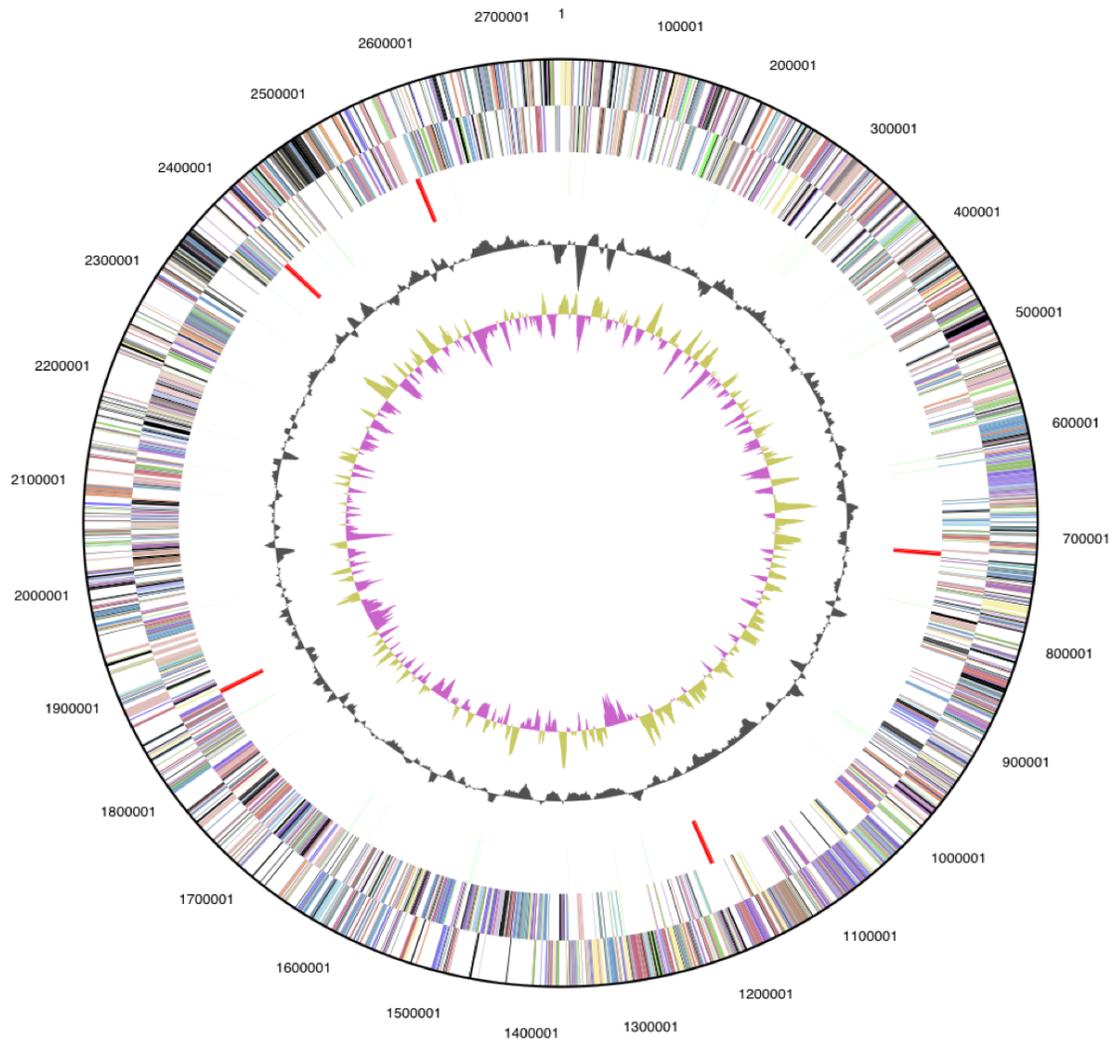
Figure 3. Graphical circular map of the genome. From outside to the center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

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

Code	value	%age	Description
J	144	5.6	Translation, ribosomal structure and biogenesis
A	1	0.0	RNA processing and modification
K	160	6.3	Transcription
L	100	3.9	Replication, recombination and repair
B	1	0.0	Chromatin structure and dynamics
D	22	0.9	Cell cycle control, mitosis and meiosis
Y	0	0.0	Nuclear structure
V	53	2.1	Defense mechanisms
T	89	3.5	Signal transduction mechanisms
M	93	3.6	Cell wall/membrane biogenesis
N	46	0.0	Cell motility
Z	0	0.0	Cytoskeleton
W	0	0.0	Extracellular structures
U	42	1.6	Intracellular trafficking and secretion

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

Code	value	%age	Description
O	79	3.1	Posttranslational modification, protein turnover, chaperones
C	110	4.3	Energy production and conversion
G	210	8.2	Carbohydrate transport and metabolism
E	162	6.3	Amino acid transport and metabolism
F	68	2.7	Nucleotide transport and metabolism
H	93	3.6	Coenzyme transport and metabolism
I	54	1.1	Lipid transport and metabolism
P	121	4.7	Inorganic ion transport and metabolism
Q	23	0.9	Secondary metabolites biosynthesis, transport and catabolism
R	207	8.1	General function prediction only
S	141	5.5	Function unknown
-	728	28.5	Not in COGs

Acknowledgements

References

1. Rocourt J, Wehmeyer U, Stackebrandt E. Transfer of *Listeria denitrificans* to a new genus, *Jonesia* gen. nov., as *Jonesia denitrificans* comb. nov. *Int J Syst Bacteriol* 1987; **37**:266-270.

2. Prévot S. Traité de bactériologie systématique, Dunod, Paris, 1961

3. Sohier R, Benazet F, Piechaud M. Sur un germe du genre *Listeria* apparemment non pathoène. *Ann Inst Pasteur (Paris)* 1948; **74**:54-57.

4. Stuart SE, Welshimer HJ. Intrageneric relatedness of *Listeria* Pirie. *Int J Syst Bacteriol* 1973; **23**:8-14.

5. Fiedler F, Seger J. The murein types of *Listeria grayi, Listeria murrayi* and *Listeria denitrificans*. *Syst Appl Microbiol* 1983; **4**:444-450.

6. Collins MD, Feresu S, Jones D. Cell wall, DNA base composition and lipid studies on *Listeria denitrificans* (Prévot). *FEMS Microbiol Lett* 1983; **18**:131-134. doi:10.1111/j.1574-6968.1983.tb00464.x

7. Bernard KA, Bellefeuille M, Ewan EP. Cellular fatty acid composition as an adjunct to the identification of asporogenous, aerobic gram-positive rods. *J Clin Microbiol* 1991; **29**:83-89. PubMed

8. Stackebrandt E, Prauser H. The family *Cellulomonadaceae. In:* The Prokaryotes, 2nd edition, Springer Verlag 1992, Vol. **2**:1323-1345.

9. Rainey FA, Weiss N, Stackebrandt E. Phylogenetic analysis of the genera *Cellulomonas, Promicromonospora*, and *Jonesia* and proposal to exclude the genus *Jonesia* from the family *Cellulomonadaceae. Int J Syst Bacteriol* 1995; **45**:649-652. PubMed

10. Schumann P, Cui X, Stackebrandt E, Kroppenstedt RM, Xu L, Jiang C. *Jonesia quinghaiensis* sp. nov., a new member of the suborder *Micrococcineae. Int J Syst Evol Microbiol* 2004; **54**:2181-2184. PubMed doi:10.1099/ijs.0.63223-0

11. Hiraishi A, Narihiro T, Yamanaka Y. Microbial community dynamics during start-up operation of flowerpot-using fed-batch reactors for composting of household biowaste. *Environ Microbiol* 2003; **5**:765-776. PubMed doi:10.1046/j.1462-2920.2003.00473.x

12. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, *et al*. Towards a richer description of our complete collection of genomes and metagenomes: the "Minimum Information about a Genome Sequence" (MIGS) specification. *Nat Bio-*

technol 2008; **26**:541-547. PubMed
doi:10.1038/nbt1360

13. 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 doi:10.1073/pnas.87.12.4576

14. Garrity GM, Holt JG. Taxonomic Outline of the Archaea and Bacteria. *Bergey's Manual of Systematic Bacteriology* 2nd ed. (D.R. Boone and R.W. Castenholz, eds.), Springer-Verlag, New York 2001; **1:**155-166.

15. Stackebrandt E, Rainey FA, Ward-Rainey NL. Proposal for a new hierachic classification system, *Actinobacteria* classis nov. *Int J Syst Bacteriol* 1997; **47**:479-491.

16. Seeliger HPR, Jones D. The genus *Listeria. In* Bergey's Manual of Systematic Bacteriology, Williams and Wilkins 1986, Vol. **2**:1235-1245.

17. Anonymous. Biological Agents: Technical rules for biological agents www.baua.de TRBA 466.

18. 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. The Gene Ontology Consortium. *Nat Genet* 2000; **25**:25-29. PubMed doi:10.1038/75556

19. 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

20. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000; **17**:540-552. PubMed

21. 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

22. Liolios K, Mavromatis K, Tavernarakis N, Kyrpides NC. The Genomes OnLine Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. *Nucleic*

Acids Res 2008; **36**:D475-D479. PubMed doi:10.1093/nar/gkm884

23. Land M, Pukall R, Abt B, Göker M, Rohde M, Glavina Del Rio T, Tice H, Copeland A, Cheng JF, Lucas S, *et al*. Complete genome of *Beutenbergia cavernae* type strain (HKI 0122[T]). *Stand Genomic Sci* 2009; **1**:21-28. doi:10.4056/sigs.1162

24. Ivanova N, Sikorski J, Sims D, Brettin T, Detter JC, Han C, Lapidus A, Copeland A, Glavina Del Rio T, Nolan M, *et al*. Complete genome of *Sanguibacter keddieii* type strain (ST-74[T]). *Stand Genomic Sci* 2009; **1**:110-118.. doi:10.4056/sigs.16197

25. Zumft WG. The denitrifying Prokaryotes. *In:* The Prokaryotes, Springer Verlag 1992, Vol. **1**:554-582.

26. List of growth media used at DSMZ: http://www.dsmz.de/microorganisms/ media_list.php

27. Wu M, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova N, Kunin V, Goodwin L, Wu M, Tindall BJ, *et al*. A phylogeny-driven genomic encyclopedia of Bacteria and Archaea. (In press).

28. Sims D, Brettin T, Detter JC, Han C, Lapidus A, Copeland A, Glavina Del Rio T, Nolan M, Chen F, Lucas S, *et al*. Complete genome of *Kytococcus sedentarius* type strain (541[T]). *Stand Genomic Sci* 2009; **1**:12-20. doi:10.4056/sigs.761

29. Anonymous. Prodigal Prokaryotic Dynamic Programming Genefinding Algorithm. Oak Ridge National Laboratory and University of Tennessee 2009. http://compbio.ornl.gov/prodigal/

30. Pati A, Ivanova N, Mikhailova N, Ovchinikova G, Hooper SD, Lykidis A, Kyrpides NC. GenePRIMP: A Gene Prediction Improvement Pipeline for microbial genomes. (Submitted) 2009

31. Markowitz VM, Mavromatis K, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. Expert IMG ER: A system for microbial genome annotation expert review and curation. *Bioinformatics* 2009; **25**:2271-2278. PubMed doi:10.1093/bioinformatics/btp393