

Genome sequence of *Phaeobacter daeponensis* type strain (DSM 23529^T), a facultatively anaerobic bacterium isolated from marine sediment, and emendation of *Phaeobacter daeponensis*

Marco Dogs¹, Hazuki Teshima², Jörn Petersen³, Anne Fiebig³, Olga Chertkov², Hajnalka Dalingault², Amy Chen⁴, Amrita Pati⁵, Lynne A. Goodwin^{2,5}, Patrick Chain⁵, John C. Detter^{2,5}, Natalia Ivanova⁵, Alla Lapidus⁵, Manfred Rohde⁶, Sabine Gronow³, Nikos C. Kyrpides⁵, Tanja Woyke⁵, Meinhard Simon¹, Markus Göker^{3*}, Hans-Peter Klenk³, Thorsten Brinkhoff¹

¹Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany.

²Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico, USA

³Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany

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

⁵DOE Joint Genome Institute, Walnut Creek, California, USA

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

*Correspondence: Markus Göker (mgo08@dsMZ.du)

Keywords: Marine microbiology, facultative anaerobe, indigoidine, *Rhodobacteraceae*, *Roseobacter* clade

TF-218^T is the type strain of the species *Phaeobacter daeponensis* Yoon *et al.* 2007, a facultatively anaerobic *Phaeobacter* species isolated from tidal flats. Here we describe the draft genome sequence and annotation of this bacterium together with previously unreported aspects of its phenotype. We analyzed the genome for genes involved in secondary metabolite production and its anaerobic lifestyle, which have also been described for its closest relative *Phaeobacter caeruleus*. The 4,642,596 bp long genome of strain TF-218^T contains 4,310 protein-coding genes and 78 RNA genes including four rRNA operons and consists of five replicons: one chromosome and four extrachromosomal elements with sizes of 276 kb, 174 kb, 117 kb and 90 kb. Genome analysis showed that TF-218^T possesses all of the genes for indigoidine biosynthesis, and on specific media the strain showed a blue pigmentation. We also found genes for dissimilatory nitrate reduction, gene-transfer agents, NRPS/ PKS genes and signaling systems homologous to the LuxR/I system.

Introduction

The genus *Phaeobacter* currently is comprised of five species (*P. daeponensis*, *P. gallaeciensis*, *P. inhibens*, *P. arcticus* and *P. caeruleus*) and is a part of the marine *Roseobacter* clade within the *Alphaproteobacteria* [1-5]. The genus name was derived from the dark brownish pigmentation of the type species *P. gallaeciensis* (*phaeos* = dark, brown) [3]. Strain TF-218^T, however, was described as not pigmented. Strain TF-218^T (= KCTC 12794^T = JCM 13606^T = DSM 23529^T) is the type strain of the species *Phaeobacter daeponensis* [1].

It was isolated from tidal flats at Daepo Beach (Yellow Sea), Korea, which led to the species name of *P. daeponensis* [1].

Secondary metabolite production is a well-known feature within the *Roseobacter* clade [6], especially within the *Phaeobacter* cluster, which shows high efficiency for secondary metabolite production [7]. Examples include biosynthesis of the antibiotics tropodithietic acid (TDA) or indigoidine, quorum sensing by *N*-acyl homoserine lactones (AHLs), and presence of genes coding for nonribosomal

peptide synthases (NRPS) and polyketide synthases (PKS) [6-11]. Furthermore, *P. daeponensis* was the first described facultatively anaerobic *Phaeobacter* species, which is capable of nitrate reduction [1].

Here we present the draft genome sequence and annotation of *P. daeponensis* TF-218^T. We analyzed the genome for special features with a focus on secondary metabolite production. Novel aspects of the strain phenotype are also reported.

Classification and features

16S rRNA gene sequence analysis

Figure 1 shows the phylogenetic neighborhood of *P. daeponensis* in a 16S rRNA gene sequence based tree. The sequences of the four 16S rRNA gene copies in the genome of strain DSM 23529^T differ from each other by up to two nucleotides, and differ by up to two nucleotides from the previously published 16S rRNA gene sequence (DQ81486) [Table 1].

A representative genomic 16S rRNA gene sequence of *P. daeponensis* TF-218^T was compared with the Greengenes database for determining the weighted relative frequencies of taxa and (truncated) keywords as previously described [21]. The most frequently occurring genera were *Ruegeria* (31.6%), *Phaeobacter* (28.8%), *Silicibacter* (13.6%), *Roseobacter* (13.3%) and *Nautella* (3.6%) (713 hits in total). Regarding the five hits to sequences from the species, the average identity within HSPs was 99.9%, whereas the average coverage by HSPs was 19.0%. Regarding the 45 hits to sequences from other species of the genus, the average identity within HSPs was 97.8%, whereas the average coverage by HSPs was 18.9%. Among all other species, the one yielding the highest score was *Roseobacter gallaeciensis* (AY881240), which corresponded to an identity of 98.6% and an HSP coverage of 18.8%. (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 sequence was AF253467 (Greengenes short name 'Key aromatic-ring-cleaving enzyme protocatechuate 3,4-dioxygenase ecologically important marine *Roseobacter* lineage d on Indulin seawater'), which showed an identity of 99.8% and an HSP coverage of 18.8%. The most frequently occurring keywords within the labels of all environmental samples which yielded hits were

'microbi' (2.8%), 'marin' (2.7%), 'coral' (2.4%), 'diseas' (1.8%) and 'water' (1.8%) (492 hits in total). The most frequently occurring keywords within the labels of those environmental samples which yielded hits of a higher score than the highest scoring species were 'marin' (17.4%), 'sediment' (8.5%), 'aromatic-ring-cleav, ecolog, enzym, import, indulin, kei, lineag, protocatechu, roseobact, seawat' (4.4%), 'coco, island, near, site' (4.3%) and 'redox-stratifi, reef, sandi' (4.3%) (4 hits in total).

Morphology and physiology

P. daeponensis TF-218^T is a Gram-negative, facultatively anaerobic, mesophilic marine bacterium with an optimal growth temperature of 37°C and an optimal salt-tolerance between 0.1 and 8% (w/v) NaCl. The optimal pH for growth is between 7.0 and 8.0 with pH 5.5 being the lowest possible pH at which growth occurs. Strain TF-218^T possesses oval cells 0.4-0.9 x 0.7-2.0 µm in size (Figure 2) and is motile by means of a single polar flagellum. On marine agar circular, slightly convex, smooth, glistening, yellowish-white colonies 1.5-2.5 mm in diameter are formed [1]. TF-218^T utilizes D-glucose, glycerol, leucine, serine, acetate, citrate and succinate [1].

In addition to the findings reported in [1], we observed that strain DSM 23529^T is able to form blue colonies on YTSS medium, as described for the closely related strain Y4I [11]. This is probably due to the presence of genes for indigoidine biosynthesis in the genome (see below).

The utilization of carbon compounds by *P. daeponensis* was also determined for this study using Generation-III microplates in an OmniLog phenotyping device (BIOLÓG Inc., Hayward, CA, USA). The microplates were inoculated at 28°C with a cell suspension at a cell density of 95-96% turbidity and dye IF-A. Further additives were vitamins, micronutrients and sea-salt solutions. The exported measurement data were further analyzed with the opm package for R [30,31], using its functionality for statistically estimating parameters from the respiration curves and translating them into negative, ambiguous, and positive reactions. The strain was studied in two independent biological replicates, and reactions with a different behavior between the two repetitions were regarded as ambiguous.

For *P. daeponensis* strain DSM 23529^T, positive reactions were observed for pH 6, 1% NaCl, 4% NaCl,

8% NaCl, D-glucose, inosine, glycerol, D-aspartic acid, L-aspartic acid, L-glutamic acid, L-histidine, L-pyroglutamic acid, L-lactic acid, α -keto-glutaric acid, D-malic acid, L-malic acid, lithium chloride, α -hydroxy-butyric acid, β -hydroxy-butyric acid, α -keto-butyric acid, acetoacetic acid, propionic acid, acetic acid and sodium bromated. In contrast, negative reactions were observed for dextrin, D-maltose, D-trehalose, D-cellobiose, β -gentiobiose, sucrose, D-turanose, stachyose, pH 5, D-raffinose, α -D-lactose, D-melibiose, β -methyl-D-galactoside, D-salicin, N-acetyl-D-glucosamine, N-acetyl- β -D-mannosamine, N-acetyl-D-galactosamine, N-acetyl-neuraminic acid, D-mannose, D-fructose, D-galactose, 3-O-methyl-D-glucose, D-fucose, L-fucose, L-rhamnose, fusidic acid, D-serine, D-

sorbitol, D-mannitol, D-arabitol, myo-inositol, D-glucose-6-phosphate, D-fructose-6-phosphate, D-serine, troleandomycin, rifamycin SV, minocycline, gelatin, L-alanine, L-arginine, L-serine, lincomycin, guanidine hydrochloride, niaproof 4, pectin, D-galacturonic acid, L-galactonic acid- γ -lactone, D-glucuronic acid, glucuronamide, mucic acid, quinic acid, D-saccharic acid, vancomycin, tetrazolium violet, methyl pyruvate, D-lactic acid methyl ester, citric acid, bromo-succinic acid, tween 40, aztreonam and butyric acid. Ambiguous results between the replicates were found for 1% sodium lactate, glycyl-L-proline, D-gluconic acid, tetrazolium blue, p-hydroxy-phenylacetic acid, nalidixic acid, potassium tellurite, γ -amino-n-butyric acid and sodium formate.

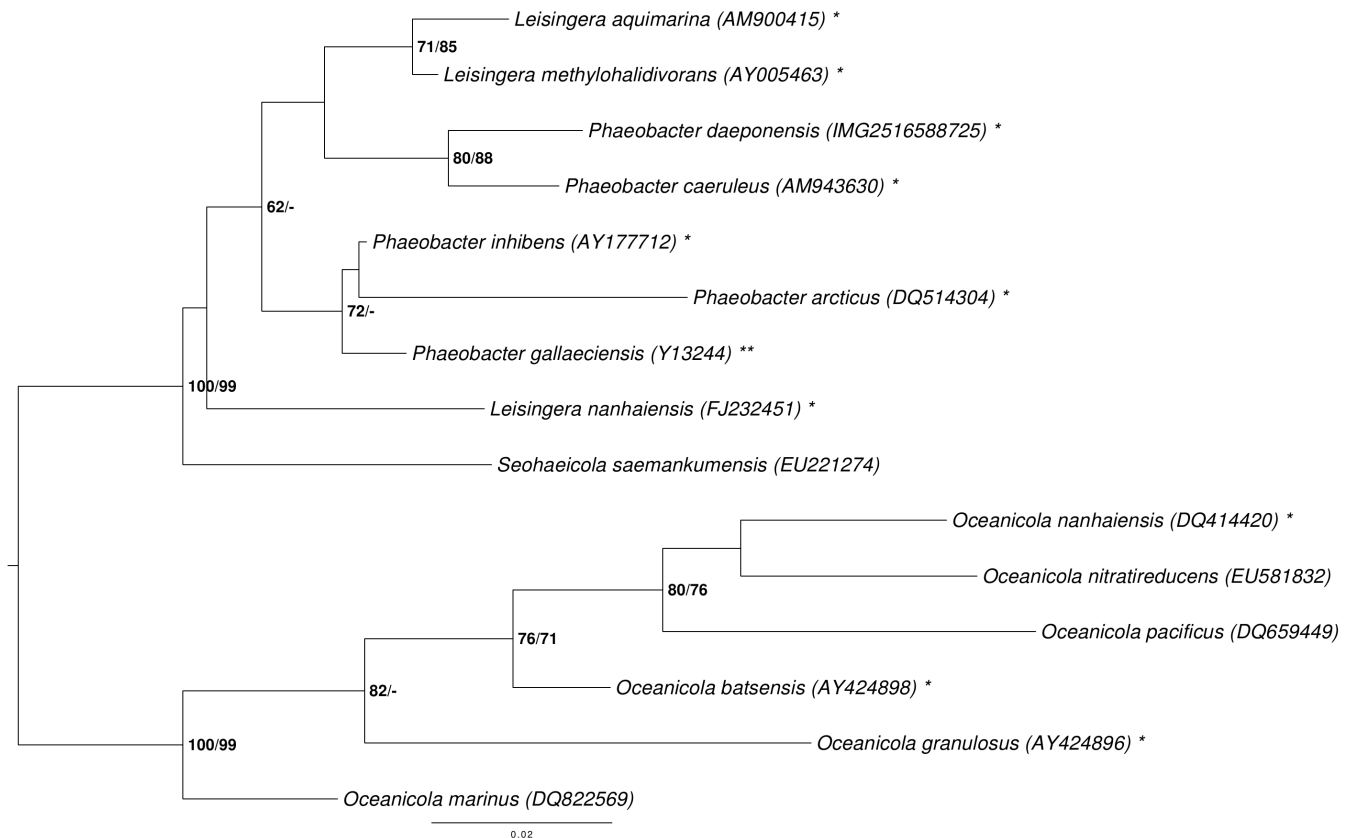


Figure 1. Phylogenetic tree highlighting the position of *P. daeponensis* relative to the type strains of the other species within the genus *Phaeobacter* and the neighboring genera *Leisingera* and *Oceanicola* [1-5,12-20]. The tree was inferred from 1,385 aligned characters of the 16S rRNA gene sequence under the maximum likelihood (ML) criterion as previously described [21]. *Oceanicola* spp. was included in the dataset for use as outgroup taxa. 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 ML bootstrap replicates (left) and from 1,000 maximum-parsimony bootstrap replicates (right) if larger than 60% [21]. Lineages with type strain genome sequencing projects registered in GOLD [22] are labeled with one asterisk, those also listed as 'Complete and Published' with two asterisks [23-25]. The genomes of six more *Leisingera* and *Phaeobacter* species are published in the current issue of *Standards in Genomic Science* [26-28].

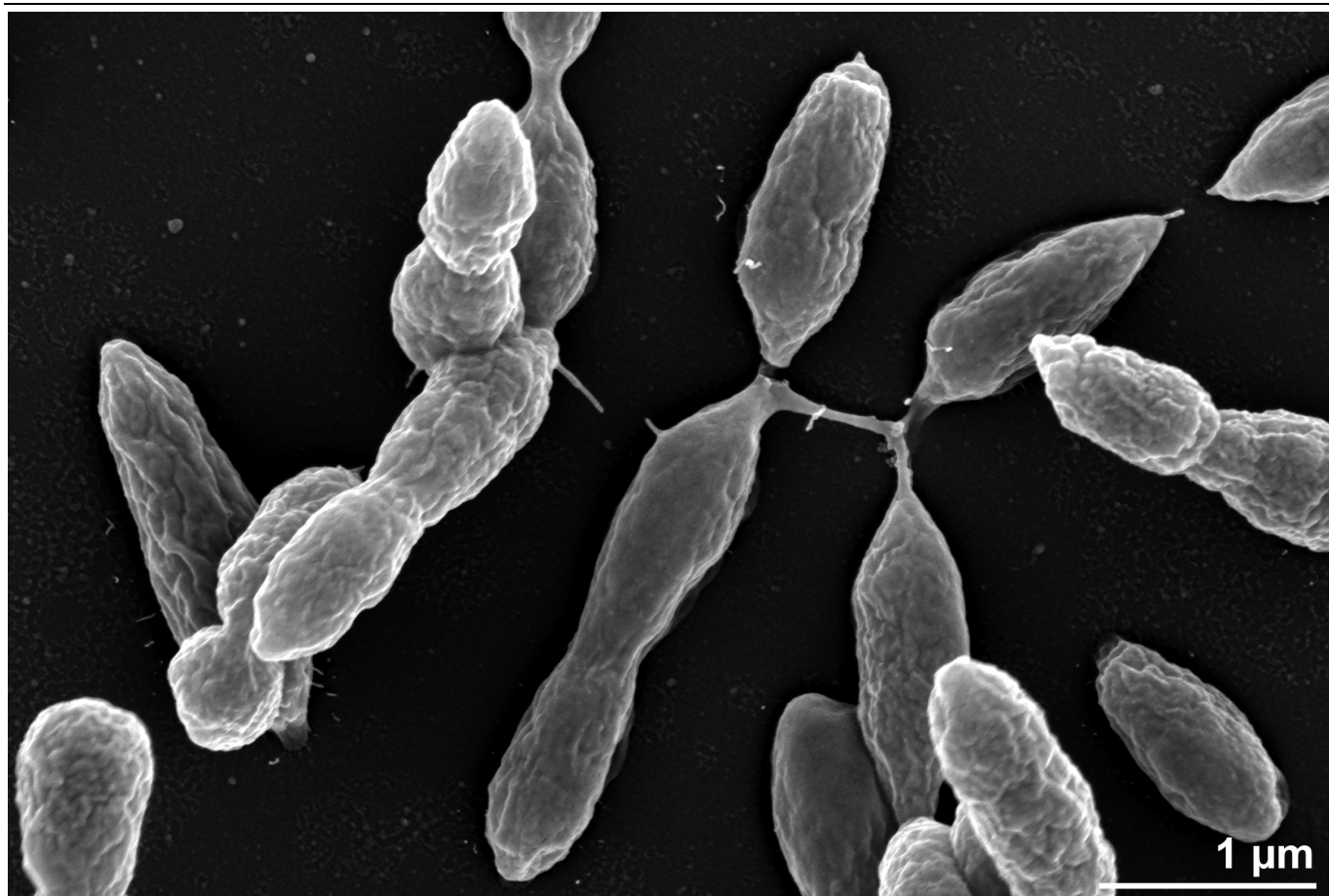


Figure 2. Scanning electron micrograph of *P. daeponensis* DSM 23529^T

Chemotaxonomy

The principal fatty-acid profile of strain TF-128^T consisted of major amounts of unsaturated fatty acid C_{18:1ω7c} (57.7%) and 11-methyl C_{18:1ω7c} (16.6%) in addition to straight-chain fatty acids (12.8%) and hydroxyl fatty acids (9.9%). Apart from the differences in the proportions, the fatty acid profile is similar to those of the type strains of *P. gallaeciensis*, *P. inhibens* and *P. caeruleus*. The major polar lipids of strain TF-218^T are phosphatidylcholine, phosphatidylglycerol, phosphatidylethanolamine, two unidentified lipids and an aminolipid [1].

Genome sequencing and annotation

Genome project history

This organism was selected for sequencing on the basis of the DOE Joint Genome Institute Community Sequencing Program (CSP) 2010, CSP 441 “Whole genome type strain sequences of the genera *Phaeobacter* and *Leisingera* – a monophyletic group of physiologically highly diverse organisms”. The

genome project is deposited in the Genomes On Line Database [22] and the complete genome sequence is deposited in GenBank. Sequencing and annotation were performed by the DOE Joint Genome Institute (JGI) using state-of-the-art sequencing technology [40]. A summary of the project information is shown in Table 2.

Growth conditions and DNA isolation

A culture of DSM 23529^T was grown aerobically in DSMZ medium 514 [41] at 37°C. Genomic DNA was isolated using a Jetflex Genomic DNA Purification Kit (GENOMED 600100) following the standard protocol provided by the manufacturer, but modified by an incubation time of 40 min, the incubation on ice over night on a shaker, the use of an additional 25 μl proteinase K, and the addition of 200 μl protein precipitation buffer. DNA is available from DSMZ through the DNA Bank Network [42].

Table 1. Classification and general features of *P. daeponensis* TF-128^T according to the MIGS recommendations [29].

MIGS ID	Property	Term	Evidence code
	Current classification	Domain <i>Bacteria</i>	TAS [32]
		Phylum <i>Proteobacteria</i>	TAS [33]
		Class <i>Alphaproteobacteria</i>	TAS [34,35]
		Order <i>Rhodobacterales</i>	TAS [35,36]
		Family <i>Rhodobacteraceae</i>	TAS [35,37]
		Genus <i>Phaeobacter</i>	TAS [1,3]
		Species <i>Phaeobacter daeponensis</i>	TAS [1,17]
		Type strain TF-218	TAS [1]
	Gram stain	Negative	TAS [1]
	Cell shape	Egg-shaped	TAS [1]
	Motility	Motile	TAS [1]
	Sporulation	None	TAS [1]
	Temperature range	Mesophile (4°C – 42°C)	TAS [1]
MIGS-6.1	Optimum temperature	37°C	TAS [1]
MIGS-6.3	Salinity	>0 - >9% (w/v)	TAS [1]
MIGS-22	Oxygen requirement	Facultative anaerobic	TAS [1]
	Carbon source	L-malate, pyruvate, D-glucose, glycerol, leucine, serine, acetate, citrate and succinate	TAS [1]
	Energy metabolism	Heterotrophic	TAS [1]
MIGS-6	Habitat	Marine	TAS [1]
MIGS-14	Pathogenicity	None	TAS [1]
MIGS-15	Biotic relationship	Particle associated	TAS [1]
	Biosafety level	1	TAS [38]
MIGS-23.1	Isolation	Tidal flat sediment	TAS [1]
MIGS-4	Geographic location	Daepo Beach (Yellow Sea), Korea	TAS [1]

Evidence 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). Evidence codes are from the Gene Ontology project [39].

Table 2. Genome sequencing project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	permanent draft
MIGS-28	Libraries used	Two Illumina paired-end libraries (221 bp and 9 kb insert size)
MIGS-29	Sequencing platforms	Illumina GAii, PacBio
MIGS-31.2	Sequencing coverage	1,345 × Illumina
MIGS-30	Assemblers	Allpaths version 38445, Velvet 1.1.05, phrap version SPS - 4.24
MIGS-32	Gene calling method	Prodigal 1.4, GenePRIMP
	INSDC ID	AXBD00000000
	GenBank Date of Release	September 30, 2013
	GOLD ID	Gi10859
	NCBI project ID	86087
	Database: IMG	2521172619
MIGS-13	Source material identifier	DSM 23529
	Project relevance	Tree of Life, carbon cycle, sulfur cycle, environmental

Genome sequencing and assembly

The draft genome sequence was generated using Illumina sequencing technology. For this genome, we constructed and sequenced an Illumina short-insert paired-end library with an average insert size of 221 bp, which generated 21,978,034 reads, and an Illumina long-insert paired-end library with an average insert size of 9,327 +/- 1,586 bp, which generated 19,261,756 reads totaling 6,186 Mbp of Illumina data. All general aspects of library construction and sequencing performed can be found at the JGI web site [43]. The initial draft assembly contained 15 contigs in 10 scaffold(s). The initial draft data was assembled with Allpaths [44] and the consensus was computationally shredded into 10 kbp overlapping fake reads (shreds). The Illumina draft data was also assembled with Velvet [45], and the consensus sequences were computationally shredded into 1.5 kbp overlapping fake reads (shreds). The Illumina draft data was assembled again with Velvet using the shreds from the first Velvet assembly to guide the next assembly. The consensus from the second Velvet assembly was shredded into 1.5 kbp overlapping fake reads. The fake reads from the Allpaths assembly, both Velvet assemblies, and a subset of the Illumina CLIP paired-end reads were

assembled using parallel phrap (High Performance Software, LLC) [46]. Possible mis-assemblies were corrected with manual editing in Consed [46]. Gap closure was accomplished using repeat resolution software (Wei Gu, unpublished), and sequencing of bridging PCR fragments with PacBio (Cliff Han, unpublished) technologies. A total of 2 PCR PacBio consensus sequences were completed to close gaps and to raise the quality of the final sequence. The final assembly is based on 6,186 Mbp of Illumina draft data, which provides an average 1,345 × coverage of the genome.

Genes were identified using Prodigal [47] as part of the DOE-JGI genome annotation pipeline [48], followed by a round of manual curation using the JGI GenePRIMP pipeline [49]. 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 [50].

Genome properties

The genome statistics are provided in Table 3 and Figures 3a – 3e. The genome consists of five scaffolds with a total length of 4,642,596 bp and a G+C content of 64.3%. The scaffolds reflect a chromosome that is 3,984,464 bp in length along with four extrachromosomal elements. Of the 4,388

genes predicted, 4,310 were protein-coding genes and 78 RNA genes, including four rRNA operons. The majority of the protein-coding genes (80.7%) 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 3. Genome Statistics

Attribute	Value	% of Total
Genome size (bp)	4,642,596	100.00%
DNA coding region (bp)	4,110,429	88.54%
DNA G+C content (bp)	2,986,366	64.34%
Number of replicons	5	
Extrachromosomal elements	4	
Total genes	4,388	100.00%
RNA genes	78	1.78%
rRNA operons	4	
Protein-coding genes	4,310	98.22%
Pseudo genes	n.a.	n.a.
Genes with function prediction	3,652	83.23%
Genes in paralog clusters	3,523	80.29%
Genes assigned to COGs	3,497	79.69%
Genes assigned Pfam domains	3,714	84.64%
Genes with signal peptides	1,507	34.34%
Genes with transmembrane helices	902	20.56%
CRISPR repeats	0	

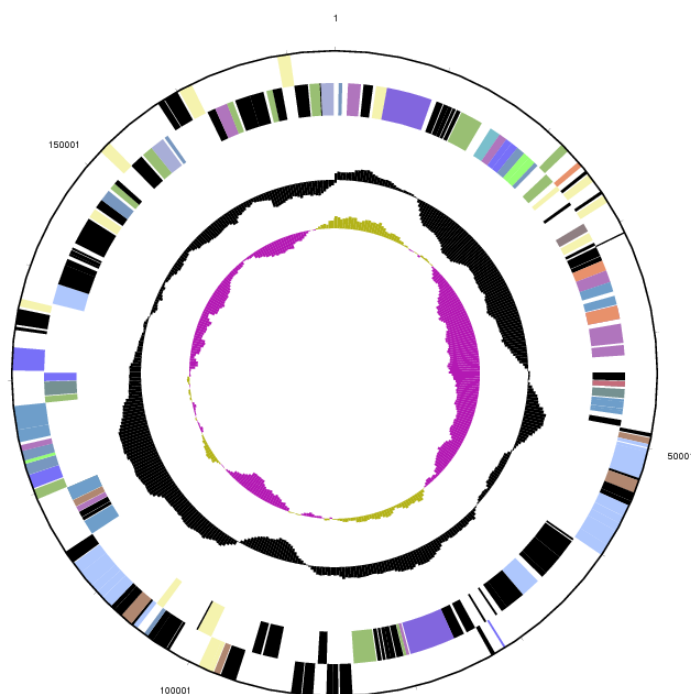


Figure 3a. Graphical map of the extrachromosomal element pDaep_B174 in strain TF-218^T. From margin to center: genes on forward strand (color by COG categories), genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew. The genome of *P. daeponensis* DSM 23529^T consists of four extrachromosomal elements (pDaep_B174; Figure 3b, pDaep_A276; Figure 3c, pDaep_C117; Figure 3d, pDaep_D91) and one chromosome (Figure 3e, cDaep_3984), as evidenced by their replication initiation system (see below).

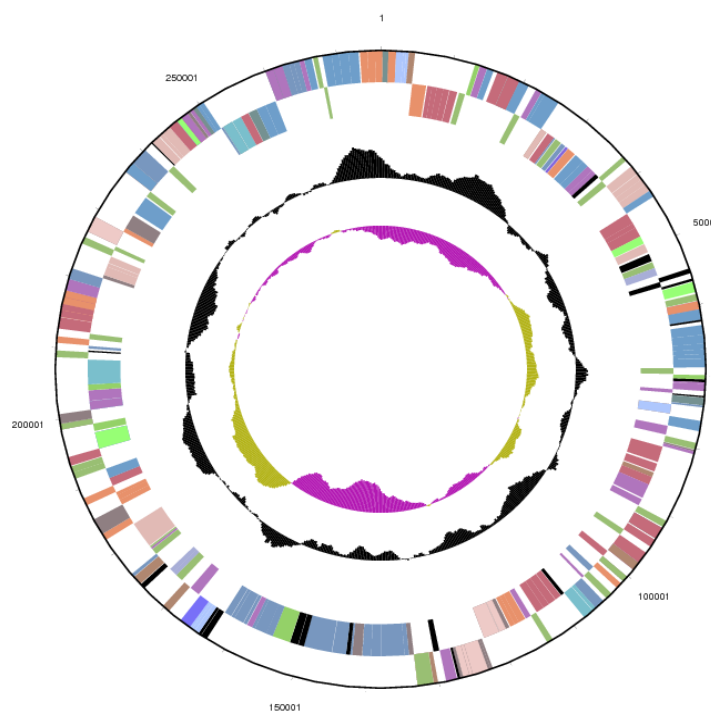


Figure 3b. Graphical map of the extrachromosomal element pDaep_A276 in strain TF-218^T. From margin to center: genes on forward strand (color by COG categories), genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

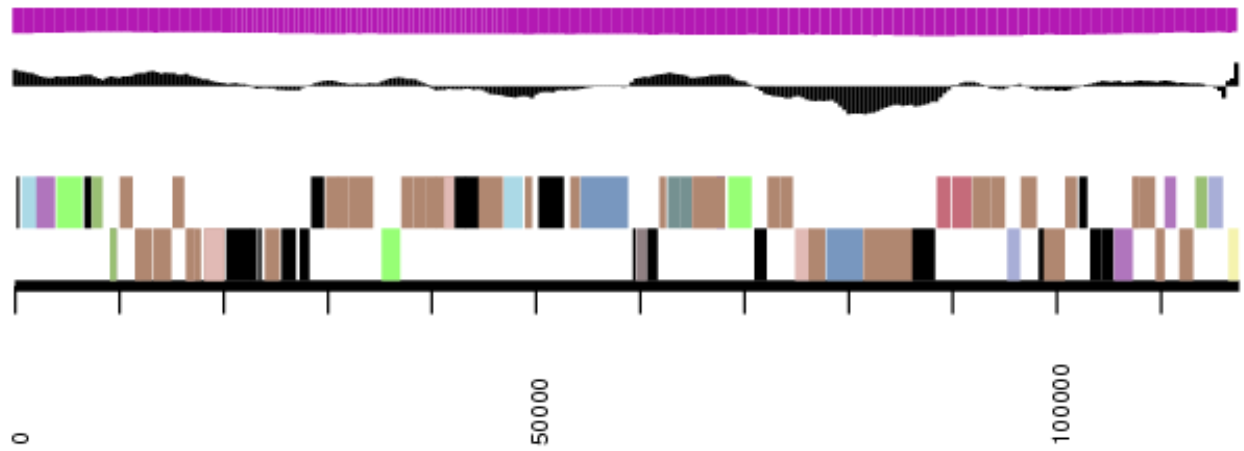


Figure 3c. Graphical map of the extrachromosomal element pDaep_C117 in strain TF-218^T. From bottom to top: genes on forward strand (color by COG categories), genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

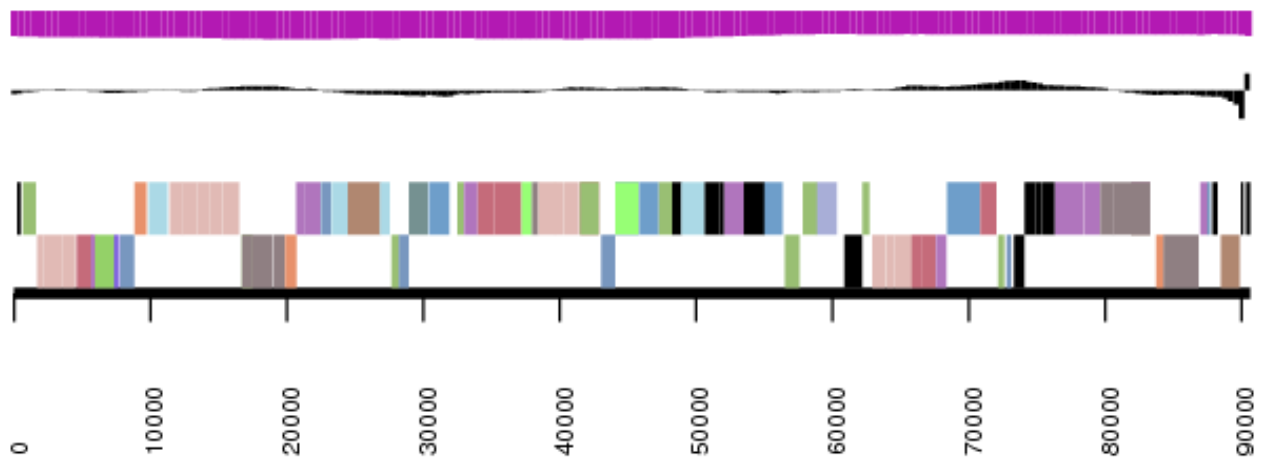


Figure 3d. Graphical map of the extrachromosomal element pDaep_D91 in strain TF-218^T. From margin to center: genes on forward strand (color by COG categories), genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

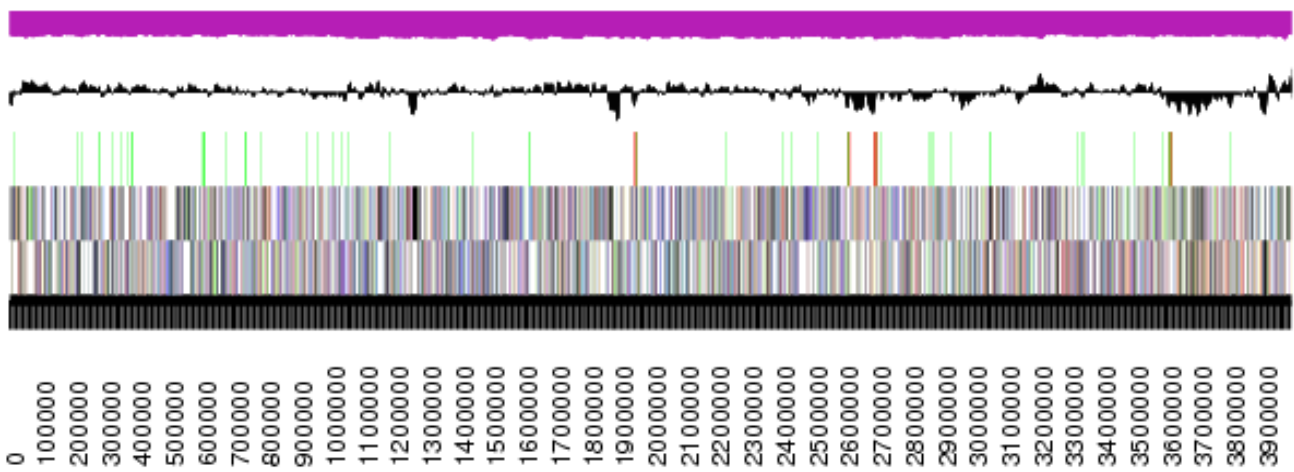


Figure 3e. Graphical map of the chromosome (cDaep_3984) in strain TF-218^T. From bottom to top: genes on forward strand (color by COG categories), genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

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

Code	Value	%age	Description
J	178	4.65	Translation, ribosomal structure and biogenesis
A	0	0	RNA processing and modification
K	298	7.78	Transcription
L	139	3.63	Replication, recombination and repair
B	3	0.08	Chromatin structure and dynamics
D	40	1.04	Cell cycle control, cell division, chromosome partitioning
Y	0	0	Nuclear structure
V	42	1.10	Defense mechanisms
T	208	5.43	Signal transduction mechanisms
M	219	5.72	Cell wall/membrane biogenesis
N	57	1.49	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	74	1.93	Intracellular trafficking and secretion, and vesicular transport
O	140	3.66	Posttranslational modification, protein turnover, chaperones
C	248	6.48	Energy production and conversion
G	161	4.20	Carbohydrate transport and metabolism
E	420	10.97	Amino acid transport and metabolism
F	89	2.32	Nucleotide transport and metabolism
H	180	4.70	Coenzyme transport and metabolism
I	153	4.00	Lipid transport and metabolism
P	195	5.09	Inorganic ion transport and metabolism
Q	130	3.40	Secondary metabolites biosynthesis, transport and catabolism
R	455	11.88	General function prediction only
S	400	10.45	Function unknown
-	891	20.31	Not in COGs

Insights into the genome

Genome sequencing of *P. daeponensis* DSM 23529^T revealed the presence of four plasmids with sizes between 91 kb and 276 kb (Table 5). The circular conformation of the two largest extrachromosomal elements was experimentally validated using PCR. The plasmids contain characteristic replication modules of the RepABC-, RepA- and RepB-type comprising a replicase as well as the *parAB* partitioning operon [51]. The respective replicases that mediate the initiation of replication are designated according to the established plasmid classification scheme [52]. The different numbering of the replicases (e.g., RepC-8, RepC-9a and RepC-9b) from RepABC-type [53,54] plasmids corresponds to specific plasmid compatibility groups that are required for a stable coexistence of the replicons within the same cell [56; unpublished results].

The 276 kb RepC-8 type replicon pDaep_A276 contains an additional DnaA-like I replicase gene (Daep_04147), but the *parAB* partitioning operon is lacking (Table 6). This distribution may be the result of a plasmid fusion and a functional inactivation of one replication module. This explanation is in agreement with the presence of two post-segregational killing systems (PSK) each consisting of a typical operon with two small genes encoding a stable toxin and an unstable antitoxin [55]. Moreover, this RepC-8 type plasmid contains a large type-VI secretion system (T6SS) with a size of about 30 kb. The role of this export system has first been described in the context of bacterial pathogenesis, but recent findings indicate a more general physiological role in defense against eukaryotic cells and

other bacteria in the environment [56-58]. We found T6S systems also on DnaA-like I type plasmids of *P. caeruleus* DSM 24564^T (pCaer_C109), *L. methylohalidivorans* DSM 14336^T (pMeth_A285) and *L. aquimarina* DSM 24565^T (pAqui_F126).

The 174 kb plasmid pDaep_B174 contains two RepABC-9 type replication modules (Figure 3a). Both of them harbor a specific perfect palindrome sequence (5'-ATCCGCG' [RepABC-9a]; 5'-TTGCACG' [RepABC-9b]) that may represent the functional cis-acting anchor for plasmid partitioning [59]. This composite replicon may have either originated from a plasmid fusion or from a horizontal recombination. The latter explanation is supported by two site-specific XerC recombinase genes (Daep_04383, Daep_04398) that are located head-to-head adjacent to the two replicases *repC9-a* and *repC9-b*.

This plasmid contains many transposases and putative phage-derived components including a DNA-primase (Daep_04238) and an RNA-directed DNA polymerase (Daep_04390). The general operon structure of this plasmid seems to be scrambled by transposition or recombination events, as illustrated by the type-IV secretion system. pDaep_B174 contains two copies of the characteristic *virD*-operon comprising the relaxase VirD2 and the coupling protein VirD4 (Table 6). Moreover, the operon contains a complete, as well as a partial, *virB* gene cluster for the transmembrane channel [57]. The first four genes in the partial cluster are missing, and the truncated *virB4* pseudogene (Daep_04339) is flanked by a transposase. But plasmid stability is probably ensured by a PSK system (Table 6).

Table 5. General genomic features of the chromosome and extrachromosomal replicons

Replicon	Scaffold	Replicase	Length (bp)	GC (%)	Topology	No. Genes [#]
cDaep_3984	1	DnaA	3,984,464	64	linear*	3,812
pDaep_A276	2	RepC-8 DnaA-like I	275,981	66	circular	245
pDaep_B174	3	RepC-9a RepC-9b	174,096	60	circular	168
pDaep_C117	4	RepA-I	117,447	66	linear*	78
pDaep_D91	5	RepB-I	90,608	67	linear*	85

*Circularity not experimentally validated

#Deduced from automatic annotation

Finally, the most conspicuous finding on this plasmid is the presence of a complete or nearly complete phenylacetate catabolon (Daep_04356 to Daep_04367), containing *paa* genes for the following proteins: PaaJ, PaaA, PaaB, PaaC, PaaD, PaaE, PaaZ, PaaY, PaaK, PaaF. The extrachromosomal localization of this catabolon has previously been shown for *Silicibacter* sp.

TM1040, *Jannaschia* sp. CCS1 and *Dinoroseobacter shibae* DSM 16493^T [60,61], which also belong to the *Roseobacter* clade.

The 117 kb RepA-I type replicon pDaep_C117 contains a LuxR-type two-component transcriptional regulator (Daep_03918) and a complete rhamnose operon [62] and is dominated by genes that are required for polysaccharide biosynthesis.

Table 6. Integrated Microbial Genome (IMG) locus tags of *P. daeponensis* DSM 23529^{T†}

Replicon	Replication Initiation Replicase	Plasmid Stability Locus Tag	Type IV Secretion Toxin	Replicon Antitoxin	Replication Initiation VirB4	Plasmid Stability VirD4
cDaep_3984	DnaA	Daep_02705	-	-	-	-
pDaep_A276	RepC-8 DnaA-like I	Daep_04038, Daep_04147	Daep_04069, Daep_04151	Daep_04068, Daep_04152	-	-
pDaep_B174	RepC-9a RepC-9b	Daep_04399, Daep_04384	Daep_04312	Daep_04313	Daep_04288, Daep_04339 [‡]	Daep_043022, Daep_04371 [‡]
pDaep_C117	RepA-I	Daep_03389	-	-	-	-
pDaep_D91	RepB-I	Daep_03883	-	-	-	-

[†]Genes for the initiation of replication, toxin/antitoxin modules and type IV secretion systems (T4SS) that are required for conjugation. The locus tags are accentuated in blue.

[‡]solitary replicase without partitioning module; [§]presence of adjacent DNA relaxase VirD2; [¶]partial pseudogene.

P. daeponensis was described as a facultatively anaerobic bacterium that uses nitrate as electron acceptor [1]. We found genes involved in nitrogen metabolism scattered over the chromosome, involved in the pathways of the assimilatory and the dissimilatory nitrate reduction to ammonia (Daep_03263, _03264 and _03265; Daep_03099, _03100, _03263 and _03264) [63-65]. Furthermore, we detected all genes necessary for the dissimilatory nitrate reduction to nitrogen, including a cluster for the nitrate reductase (Daep_03099, _03100), the nitrite reductase (Daep_02798), the nitric oxide reductase (Daep_00020, _00021) and the nitrous oxide reductase (Daep_03697) [64].

P. daeponensis encodes a gene transfer agent (GTA), a virus-like particle that mediates the transfer of genomic DNA between prokaryotes [66]. The GTA cluster has a length of ~17 kb (Daep_01107 - Daep_01126) and has a high homology to GTAs of other *Phaeobacter* species, e.g. the *P. inhibens* strains DSM 17395, 2.10 and

T5^T [28,67]. Screenings for genes coding for phage-related proteins gave hits for a phage integrase (Daep_00002, _00008 and _01212) and a phage-related gene (Daep_02906), but no complete prophage genomes were detected.

Further genome analysis of *P. daeponensis* also revealed genes related to secondary metabolism. We found genes coding for a non-ribosomal peptide synthase (Daep_00048, _01832, _01834, _01837, _02357 and _03495) and a polyketide synthase (Daep_00050). Two homologs to the *luxRI* quorum sensing system [68] were also determined (Daep_01951 and _01952; Daep_03917 and _03918). Genes coding for biosynthesis of tropodithietic acid and siderophores, as described for the *P. inhibens* strains DSM 17395, 2.10 and T5^T [66,67], were not detected.

P. daeponensis was described as a yellowish white colony forming bacterium on Marine Agar (MA; Difco) [1]. Here we could show that *P. daeponensis* forms blue-framed colonies when grown on YTSS

broth [11]. In the genome we found genes probably encoding indigoidine biosynthesis [11]. The respective operon (Daep_03493, _03494, _03495, _03496, _03497 and _03498) is similar to the operon recently described for the closely related strain *Phaeobacter* sp. Y4I [11]. The *luxRI* genes and the gene Daep_01773 show homology to the quorum-sensing systems and the *clpA* gene of *Phaeobacter* sp. strain Y4I, respectively. Strain Y4I lost its pigmentation by transposon insertions in each of the two *luxRI* quorum-sensing systems, revealing that pigment production in strain Y4I is regulated via quorum sensing [11]. Transposon insertion in gene *clpA* of strain Y4I, coding for a universal regulatory chaperone protein ClpA, which degrades abnormal and regulatory proteins, led to a higher pigment production. The presence of the biosynthesis operon and the regulatory systems indicates that *P. daeponensis* is also able to produce indigoidine in a similar way as strain Y4I.

Phylogenetic analysis shows that *P. daeponensis* and *P. caeruleus* form a cluster together with the

Leisingera species *L. methylohalidivorans* and *L. aquimarina* (Figure 1). The cluster is set apart from the clade comprising *P. gallaeciensis*, *P. inhibens* and *P. arcticus*, but the backbone of the 16S rRNA gene tree shown in Figure 1 is rather unresolved. Using the Genome-to-Genome Distance Calculator (GGDC) [69-71], we performed a preliminary phylogenomic analysis of the draft genomes of the type strains of the genera *Leisingera* and *Phaeobacter* and the finished genomes of the *P. inhibens* strains DSM 17395 and 2.10. Table 7 shows the results of the *in-silico* calculated DNA-DNA hybridization (DDH) similarities of *P. daeponensis* to other *Phaeobacter* and *Leisingera* species. The highest values were obtained for *P. caeruleus*, *L. aquimarina* and *L. methylohalidivorans*, thus confirming the 16S rRNA gene analysis. A reclassification of *P. daeponensis* and *P. caeruleus* as species of the genus *Leisingera* is one possible solution to taxonomically better represent the genomic data.

Table 7. Digital DDH similarities between *P. daeponensis* DSM 23529^T and the other *Phaeobacter* and *Leisingera* species[†]

Reference strain (type strain unless indicated)	formula 1	formula 2	formula 3
<i>P. arcticus</i> (AXBF00000000)	17.00±3.27	21.00±2.33	16.90±2.77
<i>P. caeruleus</i> (AXBI00000000)	62.50±3.67	40.30±2.51	57.80±3.18
<i>P. inhibens</i> (AXBB00000000)	19.90±3.39	21.20±2.34	19.20±2.86
<i>P. gallaeciensis</i> (AOQA01000000)	19.10±3.36	21.40±2.34	18.70±2,84
<i>P. inhibens</i> DSM 17395 (CP002976, CP002977, CP002978, CP002979)	19.70±3.38	21.40±2.34	19.10±2.86
<i>P. inhibens</i> 2.10 (NC_018286)	19.80±3.39	21.10±2.33	19.20±2.86
<i>L. aquimarina</i> (AXBE00000000)	47.30±3.42	27.90±2.43	41.30±3.01
<i>L. methylohalidivorans</i> (CP006773, CP006774, CP006775)	48.70±3.43	26.90±2.42	41.90±3.01
<i>L. nanhaiensis</i> (AXBG00000000)	14.70±3.13	19.60±2.30	14.80±2.66

[†]Including the genome-sequenced type strains and *P. inhibens* strains DSM 17395 and 2.10) calculated *in silico* with the GGDC server version 2.0 [69]. The standard deviations indicate the inherent uncertainty in estimating DDH values from inter-genomic distances based on models derived from empirical test data sets (which are always limited in size); see [69] for details. The distance formulas are explained in [70]. The numbers in parentheses are GenBank accession numbers identifying the underlying genome sequences.

Even though discrepancies between the current classification of the group and the genomic data apparently exist, it is also obvious that *P. caeruleus*, which forms blue colonies [5], is the closest known relative of *P. daeponensis* (Table 7). For this reason, the formation of blue colonies by *P. daeponensis* DSM 23529^T on YTSS medium [11] observed in this study, confirmed by the presence of genes for indigoidine biosynthesis in the genome, is probably of taxonomic relevance. This

warrants an update of the taxonomic description of *P. daeponensis*.

Emended description of the species

Phaeobacter daeponensis Yoon *et al.* 2007

The description of the species *Phaeobacter daeponensis* is the one given by Yoon *et al.* 2007 [1], with the following modification. Forms blue colonies when cultivated on YTSS medium.

Acknowledgements

The authors would like to gratefully acknowledge the assistance of Iljana Schröder for growing *P. daeponensis* cultures and Evelyne-Marie Brambilla for DNA extraction and quality control (both at DSMZ). The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the U.S. Department of Energy under contract

No. DE-AC02-05CH11231; the work conducted by members of the Roseobacter consortium was supported by the German Research Foundation (DFG) Transregio-SFB 51. We also thank the European Commission which supported phenotyping *via* the Microme project 222886 within the Framework 7 program.

References

1. Yoon JH, Kang SJ, Lee SY, Oh TK. *Phaeobacter daeponensis* sp. nov., isolated from a tidal flat of the Yellow Sea in Korea. *Int J Syst Evol Microbiol* 2007; **57**:856-861. [PubMed](#)
<http://dx.doi.org/10.1099/ijs.0.64779-0>
2. Ruiz-Ponte C, Cilia V, Lambert C, Nicolas JL. *Roseobacter gallaeciensis* sp. nov., a new marine bacterium isolated from rearings and collectors of the scallop *Pecten maximus*. *Int J Syst Bacteriol* 1998; **48**:537-542. [PubMed](#)
<http://dx.doi.org/10.1099/00207713-48-2-537>
3. Martens T, Heidorn T, Pukall R, Simon M, Tindall BJ, Brinkhoff T. Reclassification of *Roseobacter gallaeciensis* Ruiz-Ponte *et al.* 1998 as *Phaeobacter gallaeciensis* gen. nov., comb. nov., description of *Phaeobacter inhibens* sp. nov., reclassification of *Ruegeria algicola* (Lafay *et al.* 1995) Uchino *et al.* 1999 as *Marinovum algicola* gen. nov., comb. nov., and emended descriptions of the genera *Roseobacter*, *Ruegeria* and *Leisingera*. *Int J Syst Evol Microbiol* 2006; **56**:1293-1304. [PubMed](#)
<http://dx.doi.org/10.1099/ijs.0.63724-0>
4. Zhang DC, Li HR, Xin YH, Liu HC, Chi ZM, Zhou PJ, Yu Y. *Phaeobacter arcticus* sp. nov., a psychrophilic bacterium isolated from the Arctic. *Int J Syst Evol Microbiol* 2008; **58**:1384-1387. [PubMed](#)
<http://dx.doi.org/10.1099/ijs.0.65708-0>
5. Vandecastelaere I, Segaeert E, Mollica A, Faimali M, Vandamme P. *Phaeobacter caeruleus* sp. nov., a blue-coloured, colony-forming bacterium isolated from a marine electroactive biofilm. *Int J Syst Evol Microbiol* 2009; **59**:1209-1214. [PubMed](#)
<http://dx.doi.org/10.1099/ijs.0.002642-0>
6. Martens T, Gram L, Grossart HP, Kessler D, Müller R, Simon M, Wenzel SC, Brinkhoff T. Bacteria of the *Roseobacter* clade show potential for secondary metabolite production. *Microb Ecol* 2007; **54**:31-42. [PubMed](#)
<http://dx.doi.org/10.1007/s00248-006-9165-2>
7. Bruhn JB, Nielsen KF, Hjelm M, Hansen M, Bresciani J, Schulz S, Gram L. Ecology, inhibitory activity and morphogenesis of a potential marine fish larvae probiotic bacteria, *Roseobacter* strain 27-4. *Appl Environ Microbiol* 2005; **71**:7263-7270. [PubMed](#)
<http://dx.doi.org/10.1128/AEM.71.11.7263-7270.2005>
8. Brinkhoff T, Bach G, Heidorn T, Liang L, Schlingloff A, Simon M. Antibiotic production by a *Roseobacter* clade-affiliated species from the German Wadden Sea and its antagonistic effects on indigenous isolates. *Appl Environ Microbiol* 2004; **70**:2560-2565. [PubMed](#)
<http://dx.doi.org/10.1128/AEM.70.4.2560-2565.2003>
9. Rao D, Webb JS, Kjelleberg S. Competitive interactions in mixed-species biofilms containing the marine bacterium *Pseudoalteromonas tunicata*. *Appl Environ Microbiol* 2005; **71**:1729-1736. [PubMed](#)

- <http://dx.doi.org/10.1128/AEM.71.4.1729-1736.2005>
10. Ruiz-Ponte C, Samain JF, Sanchez JL, Nicolas JL. The benefit of a *Roseobacter* species on the survival of scallop larvae. *Mar Biotechnol (NY)* 1999; **1**:52-59. [PubMed](#) <http://dx.doi.org/10.1007/PL00011751>
 11. Cude WN, Mooney J, Tavanaei AA, Hadden MK, Frank AM, Gulvik CA, May AL, Buchan A. Production of the antimicrobial secondary metabolite indigoidine contributes to competitive surface colonization by the marine roseobacter *Phaeobacter* sp. strain Y4I. *Appl Environ Microbiol* 2012; **78**:4771-4780. [PubMed](#) <http://dx.doi.org/10.1128/AEM.00297-12>
 12. Gu J, Guo B, Wang YN, Yu SL, Inamori R, Qu R, Ye YG, Wu XL. *Oceanicola nanhaiensis* sp. nov., isolated from sediments of the South China Sea. *Int J Syst Evol Microbiol* 2007; **57**:157-160. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.64532-0>
 13. Lin KY, Sheu SY, Chang PS, Cho JC, Chen WM. *Oceanicola marinus* sp. nov., a marine alphaproteobacterium isolated from seawater collected off Taiwan. *Int J Syst Evol Microbiol* 2007; **57**:1625-1629. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.65020-0>
 14. Schaefer JK, Goodwin KD, McDonald IR, Murrell JC, Oremland RS. *Leisingera methylohalidivorans* gen. nov., sp. nov., a marine methylotroph that grows on methyl bromide. *Int J Syst Evol Microbiol* 2002; **52**:851-859. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.01960-0>
 15. Sun F, Wang B, Liu X, Lai Q, Du Y, Li G, Luo J, Shao Z. *Leisingera nanhaiensis* sp. nov., isolated from marine sediment. *Int J Syst Evol Microbiol* 2010; **60**:275-280. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.010439-0>
 16. Thrash JC, Cho JC, Vergin KL, Giovannoni SJ. Genome sequences of *Oceanicola granulosus* HTCC2516^T and *Oceanicola batsensis* HTCC2597^T. *J Bacteriol* 2010; **192**:3549-3550. [PubMed](#) <http://dx.doi.org/10.1128/JB.00412-10>
 17. Vandecastelaere I, Segaeert E, Mollica A, Faimali M, Vandamme P. *Leisingera aquimarina* sp. nov., isolated from a marine electroactive biofilm, and emended descriptions of *Leisingera methylohalidivorans* Schaefer et al. 2002, *Phaeobacter daeponensis* Yoon et al. 2007 and *Phaeobacter inhibens* Martens et al. 2006. *Int J Syst Evol Microbiol* 2008; **58**:2788-2793. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.65844-0>
 18. Yoon JH, Kang SJ, Lee SY, Oh KH, Oh TK. *Seohaecicola saemankumensis* gen. nov., sp. nov., isolated from a tidal flat. *Int J Syst Evol Microbiol* 2009; **59**:2675-2679. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.011312-0>
 19. Yuan J, Lai Q, Wang B, Sun F, Liu X, Du Y, Li G, Gu L, Zheng T, Shao Z. *Oceanicola pacificus* sp. nov., isolated from a deep-sea pyrene-degrading consortium. *Int J Syst Evol Microbiol* 2009; **59**:1158-1161. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.003400-0>
 20. Zheng Q, Chen C, Wang YN, Jiao N. *Oceanicola nitrareducens* sp. nov., a marine alphaproteobacterium isolated from the South China Sea. *Int J Syst Evol Microbiol* 2010; **60**:1655-1659. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.016311-0>
 21. Göker M, Cleland D, Saunders E, Lapidus A, Nolan M, Lucas S, Hammon N, Deshpande S, Cheng JF, Tapia R, et al. Complete genome sequence of *Isosphaera pallida* type strain (IS1B^T). *Stand Genomic Sci* 2011; **4**:63-71. [PubMed](#) <http://dx.doi.org/10.4056/sigs.1533840>
 22. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyrpides NC. The Genomes OnLine Database (GOLD) in 2009:status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* 2010; **38**:D346-D354. [PubMed](#) <http://dx.doi.org/10.1093/nar/gkp848>
 23. Riedel T, Teshima H, Petersen J, Fiebig A, Davenport K, Dalingault H, Erkkila T, Gu W, Munk C, Xu Y, et al. Genome sequence of the *Leisingera aquimarina* type strain (DSM 24565^T), a member of the *Roseobacter* clade rich in extrachromosomal elements. *Stand Genomic Sci* 2013; **8**:389-402. <http://dx.doi.org/10.4056/sigs.3858183>
 24. Beyersmann PG, Chertkov O, Petersen J, Fiebig A, Chen A, Pati A, Ivanova N, Lapidus A, Goodwin LA, Chain P, et al. Genome sequence of *Phaeobacter caeruleus* type strain (DSM 24564^T), a surface-associated member of the marine *Roseobacter* clade. *Stand Genomic Sci* 2013; **8**:403-419. <http://dx.doi.org/10.4056/sigs.3927626>
 25. Freese H, Dalingault H, Petersen J, Pradella S, Fiebig A, Davenport K, Teshima H, Chen A, Pati A, Ivanova N, et al. Genome sequence of the plasmid and phage-gene rich marine *Phaeobacter arcticus* type strain (DSM 23566^T). *Stand Genomic Sci* 2013; **8**:450-464. <http://dx.doi.org/10.4056/sigs.383362>

26. Breider S, Teshima H, Petersen J, Fiebig A, Chertkov O, Dalingault H, Chen A, Pati A, Ivanova N, Lapidus A, et al. Genome sequence of *Leisingera nanhaiensis* strain DSM 23252^T isolated from marine sediment. *Stand Genomic Sci* 2013; **8**:389-402.
27. Buddruhs N, Chertkov O, Fiebig A, Petersen J, Chen A, Pati A, Ivanova N, Lapidus A, Goodwin LA, Chain P, et al. Complete genome sequence of the marine methyl-halide oxidizing *Leisingera methylohalidivorans* type strain (DSM 14336^T), a member of the *Roseobacter* clade. *Stand Genomic Sci* 2013; In press.
28. Dogs M, Voget S, Teshima H, Petersen J, Fiebig A, Davenport K, Dalingault H, Chen A, Pati A, Ivanova N, et al. Genome sequence of *Phaeobacter inhibens* strain T5^T, a secondary-metabolite producing member of the marine *Roseobacter* clade, and emendation of the species *Phaeobacter inhibens*. *Stand Genomic Sci* 2013; In press.
29. 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](http://dx.doi.org/10.1038/nbt1360) <http://dx.doi.org/10.1038/nbt1360>
30. Vaas LAI, Sikorski J, Michael V, Göker M, Klenk HP. Visualization and curve-parameter estimation strategies for efficient exploration of phenotype microarray kinetics. *PLoS ONE* 2012; **7**:e34846. [PubMed](http://dx.doi.org/10.1371/journal.pone.0034846) <http://dx.doi.org/10.1371/journal.pone.0034846>
31. Vaas LAI, Sikorski J, Hofner B, Buddruhs N, Fiebig A, Klenk HP, Göker M. opm: An R package for analysing OmniLog® Phenotype MicroArray Data. *Bioinformatics* 2013; **29**:1823-1824. [PubMed](http://dx.doi.org/10.1093/bioinformatics/btt291) <http://dx.doi.org/10.1093/bioinformatics/btt291>
32. 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](http://dx.doi.org/10.1073/pnas.87.12.4576) <http://dx.doi.org/10.1073/pnas.87.12.4576>
33. Garrity GM, Bell JA, Lilburn T. Phylum XIV. *Proteobacteria* phyl. nov. In: DJ Brenner, NR Krieg, JT Staley, GM Garrity (eds), *Bergey's Manual of Systematic Bacteriology*, second edition, vol. 2 (The *Proteobacteria*), part B (The *Gammaproteobacteria*), *Springer, New York*, 2005, p. 1.
34. Garrity GM, Bell JA, Lilburn T. Class I. *Alphaproteobacteria* class. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 2, Part C, *Springer, New York*, 2005, p. 1.
35. Validation List No. 107. List of new names and new combinations previously effectively, but not validly, published. *Int J Syst Evol Microbiol* 2006; **56**:1-6. [PubMed](http://dx.doi.org/10.1099/ijs.0.64188-0) <http://dx.doi.org/10.1099/ijs.0.64188-0>
36. Garrity GM, Bell JA, Lilburn T. Order III. *Rhodobacterales* ord. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 2, Part C, *Springer, New York*, 2005, p. 161.
37. Garrity GM, Bell JA, Lilburn T. Family I. *Rhodobacteraceae* fam. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 2, Part C, *Springer, New York*, 2005, p. 161.
38. BAuA. Classification of *Bacteria* and *Archaea* in risk groups. *TRBA* 2010; **466**:93.
39. 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](http://dx.doi.org/10.1038/75556) <http://dx.doi.org/10.1038/75556>
40. Mavromatis K, Land ML, Brettin TS, Quest DJ, Copeland A, Clum A, Goodwin L, Woyke T, Lapidus A, Klenk HP, et al. The fast changing landscape of sequencing technologies and their impact on microbial genome assemblies and annotation. *PLoS ONE* 2012; **7**:e48837. [PubMed](http://dx.doi.org/10.1371/journal.pone.0048837) <http://dx.doi.org/10.1371/journal.pone.0048837>
41. List of growth media used at the DSMZ: <http://www.dmsz.de/catalogues/catalogue-microorganisms/culture-technology/list-of-media-for-microorganisms.html>.
42. 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. *Biopreserv Biobank* 2011; **9**:51-55. <http://dx.doi.org/10.1089/bio.2010.0029>
43. The DOE Joint Genome Institute. www.jgi.doe.gov
44. Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, Nusbaum C, Jaffe DB. ALLPATHS: de novo assembly of whole-genome shotgun microreads. *Genome Res* 2008; **18**:810-

820. [PubMed](#)
<http://dx.doi.org/10.1101/gr.7337908>
45. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008; **18**:821-829. [PubMed](#)
<http://dx.doi.org/10.1101/gr.074492.107>
46. Phrap and Phred for Windows, MacOS, Linux, and Unix. <http://www.phrap.com>
47. 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](#)
<http://dx.doi.org/10.1186/1471-2105-11-119>
48. 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**:63-67. [PubMed](#)
<http://dx.doi.org/10.4056/sigs.632>
49. 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](#)
<http://dx.doi.org/10.1038/nmeth.1457>
50. 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](#)
<http://dx.doi.org/10.1093/bioinformatics/btp393>
51. Petersen J. Phylogeny and compatibility: plasmid classification in the genomics era. *Arch Microbiol* 2011; **193**:313-321. [PubMed](#)
52. Petersen J, Brinkmann H, Berger M, Brinkhoff T, Päuker O, Pradella S. Origin and evolution of a novel DnaA-like plasmid replication type in *Rhodobacterales*. *Mol Biol Evol* 2011; **28**:1229-1240. [PubMed](#)
<http://dx.doi.org/10.1093/molbev/msq310>
53. Cevallos MA, Cervantes-Rivera R, Gutierrez-Rios RM. The repABC plasmid family. *Plasmid* 2008; **60**:19-37. [PubMed](#)
<http://dx.doi.org/10.1016/j.plasmid.2008.03.001>
54. Castillo-Ramírez S, Vazque-Castellanos JF, Gonzalez V, Cevallos MA. Horizontal gene transfer and diverse functional constraints within a common replication-partitioning system in *Alphaproteobacteria*: the repABC operon. *BMC Genomics* 2009; **10**:536. [PubMed](#)
<http://dx.doi.org/10.1186/1471-2164-10-536>
55. Zielenkiewicz U, Ceglowski P. Mechanisms of plasmid stable maintenance with special focus on plasmid addiction systems. *Acta Biochim Pol* 2001; **48**:1003-1023. [PubMed](#)
56. Schwarz S, Hood RD, Mougous JD. What is type VI secretion doing in all those bugs? *Trends Microbiol* 2010; **18**:531-537. [PubMed](#)
<http://dx.doi.org/10.1016/j.tim.2010.09.001>
57. del Solar G, Giraldo R, Ruiz-Echevarria MJ, Espinosa M, Diaz-Orejes R. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* 1998; **62**:434-464. [PubMed](#)
58. Cascales E, Christie PJ. The versatile bacterial type IV secretion systems. *Nat Rev Microbiol* 2003; **1**:137-149. [PubMed](#)
<http://dx.doi.org/10.1038/nrmicro753>
59. Petersen J, Brinkmann H, Pradella S. Diversity and evolution of repABC type plasmids in *Rhodobacterales*. *Environ Microbiol* 2009; **11**:2627-2638. [PubMed](#)
<http://dx.doi.org/10.1111/j.1462-2920.2009.01987.x>
60. Moran MA, Belas R, Schell MA, González JM, Sun F, Sun S, Binder BJ, Edmonds J, Ye W, Orcutt B, et al. Ecological genomics of marine *Roseobacters*. *Appl Environ Microbiol* 2007; **73**:4559-4569. [PubMed](#)
<http://dx.doi.org/10.1128/AEM.02580-06>
61. Wagner-Döbler I, Ballhausen B, Berger M, Brinkhoff T, Buchholz I, Bunk B, Cypionka H, Daniel R, Drepper T, Gerds G, et al. The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a hitchhiker's guide to life in the sea. *ISME J* 2010; **4**:61-77. [PubMed](#)
<http://dx.doi.org/10.1038/ismej.2009.94>
62. Giraud MF, Naismith JH. The rhamnose pathway. *Curr Opin Struct Biol* 2000; **10**:687-696. [PubMed](#)
[http://dx.doi.org/10.1016/S0959-440X\(00\)00145-Z](http://dx.doi.org/10.1016/S0959-440X(00)00145-Z)
63. Stolz JF, Basu P. Evolution of nitrate reductase: molecular and structural variations on common function. *ChemBioChem* 2002; **3**:198-206. [PubMed](#)
[http://dx.doi.org/10.1002/1439-7633\(20020301\)3:2/3<198::AID-CBIC198>3.0.CO;2-C](http://dx.doi.org/10.1002/1439-7633(20020301)3:2/3<198::AID-CBIC198>3.0.CO;2-C)
64. Morozkina EV, Zvyagil'skaya RA. Nitrate reductase: structure, functions, and effect of stress factors. *Biochemistry (Mosc)* 2007; **72**:1151-1160. [PubMed](#)
<http://dx.doi.org/10.1134/S0006297907100124>

65. Cabello P, Roldan MD, Moreno-Vivian C. Nitrate reduction and the nitrogen cycle in archaea. *Microbiology* 2004; **150**:3527-3546. [PubMed](#) <http://dx.doi.org/10.1099/mic.0.27303-0>
66. Lang AS, Beatty JT. The gene transfer agent of *Rhodobacter capsulatus* and "constitutive transduction" in prokaryotes. *Arch Microbiol* 2001; **175**:241-249. [PubMed](#) <http://dx.doi.org/10.1007/s002030100260>
67. Thole S, Kalhoefer D, Voget S, Berger M, Engelhardt T, Liesegang H, Wollherr A, Kjelleberg S, Daniel R, Simon M, et al. *Phaeobacter gallaeciensis* genomes from globally opposite locations reveal high similarity of adaption to surface life. *ISME J* 2012; **6**:2229-2244. [PubMed](#) <http://dx.doi.org/10.1038/ismej.2012.62>
68. Fuqua C, Winans SC, Greenberg EP. Census and consensus in bacterial ecosystems: the LuxR-LuxI family of quorum-sensing transcriptional regulators. *Annu Rev Microbiol* 1996; **50**:727-751. [PubMed](#) <http://dx.doi.org/10.1146/annurev.micro.50.1.727>
69. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013; **14**:60. [PubMed](#) <http://dx.doi.org/10.1186/1471-2105-14-60>
70. 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**:142-148. [PubMed](#) <http://dx.doi.org/10.4056/sigs.541628>
71. 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**:117-134. [PubMed](#) <http://dx.doi.org/10.4056/sigs.531120>