Genomics Data 9 (2016) 154–159

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Draft genome sequence of *Pseudomonas moraviensis* strain Devor implicates metabolic versatility and bioremediation potential

Neil T. Miller [a], Danny Fuller [a], M.B. Couger [a], Mark Bagazinski [a], Philip Boyne [a], Robert C. Devor [a], Radwa A. Hanafy [a], Connie Budd [a], Donald P. French [b], Wouter D. Hoff [a], Noha Youssef [a,*]

[a] Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, United States
[b] Department of Integrative Biology, Oklahoma State University, Stillwater, OK, United States

ARTICLE INFO

Article history:
Received 19 July 2016
Accepted 2 August 2016
Available online 4 August 2016

Keywords:
Pseudomonas moraviensis
Draft genome sequence
Detailed analysis
Student Initiated Microbial Discovery (SIMD) project
Metabolic versatility

ABSTRACT

Pseudomonas moraviensis is a predominant member of soil environments. We here report on the genomic analysis of *Pseudomonas moraviensis* strain Devor that was isolated from a gate at Oklahoma State University, Stillwater, OK, USA. The partial genome of *Pseudomonas moraviensis* strain Devor consists of 6016489 bp of DNA with 5290 protein-coding genes and 66 RNA genes. This is the first detailed analysis of a *P. moraviensis* genome. Genomic analysis revealed metabolic versatility with genes involved in the metabolism and transport of fructose, xylose, mannose and all amino acids with the exception of tryptophan and valine, implying that the organism is a versatile heterotroph. The genome of *P. moraviensis* strain Devor was rich in transporters and, based on COG analysis, did not cluster closely with *P. moraviensis* R28-S genome, the only previous report of a *P. moraviensis* genome with a native mercury resistance plasmid.

1. Introduction

The strain Devor was isolated from an outdoor gate located on Oklahoma State University (OSU) campus in Stillwater, OK. The isolation process was part of the Student Initiated Microbial Discovery (SIMD) project at OSU (introduced in [1]). The organism was isolated by an undergraduate student (RCD) during an introductory microbiology lab course, and analyzed by a team of undergraduate (MB and PB) and graduate students (NTM and DF). The genus *Pseudomonas* is known to be phylogenetically and physiologically diverse. Members are ubiquitously found in soil and water systems, with the potential to utilize a broad range of organic compounds [2–12]. Resistance for heavy metal has also been shown for some strains [13]. Genomic analysis of members belonging to the genus *Pseudomonas* can contribute to our understanding of the molecular mechanisms of the biodegradation of organic compounds, including environmental pollutants, and could potentially contribute to bioremediation efforts in multiple environments. Here we report on the draft genomic sequence, and the first detailed analysis of a *Pseudomonas moraviensis* strain.

2. Materials and methods

2.1. Genome project history

Pseudomonas moraviensis strain Devor was isolated with the aim of sequencing its genome as part of an undergraduate project at OSU. The project is funded by the Howard Hughes Medical Institute and aims at improving undergraduate student involvement in authentic research. Isolation of the strain, and analysis of the sequenced genome were performed by undergraduate students in an introductory microbiology course, and an upper division microbial genomics class, respectively. The quality draft assembly and annotation were completed in 2015–2016. Table 1 shows the genome project information.

2.2. Growth conditions and genomic DNA preparation

Pseudomonas moraviensis strain Devor was grown overnight at 30 °C on tryptic soy agar plates, and genomic DNA was isolated using the MPBio PowerSoil® DNA extraction kit according to manufacturer's instructions. Negative stain TEM micrographs were obtained using the services of the Oklahoma State University Microscopy Lab. Briefly, the sample was placed on a carbon film TEM grid and allowed to incubate for 2 min, after which the excess liquid was wicked off. Phosphotungstic acid (PTA; 2% w/v) was then added to the grid followed by a 45-s incubation. Excess PTA was wicked off and the grid was allowed to

* Corresponding author at: 1110 S Innovation way, Stillwater, OK 74074, United States.
E-mail address: noha@okstate.edu (N. Youssef).

http://dx.doi.org/10.1016/j.gdata.2016.08.004

Table 1
Project information.

MIGS ID	Property	Term
MIGS 31	Finishing quality	Draft
MIGS-28	Libraries used	2 × 300 paired end chemistry
MIGS 29	Sequencing platforms	Illumina
MIGS 31.2	Fold coverage	300×
MIGS 30	Assemblers	Velvet
MIGS 32	Gene calling method	Prodigal
	Genbank ID	MAYQ00000000
	GenBank date of release	July 2016
	GOLD ID	Gp0126756
	BIOPROJECT	PRJNA327387
MIGS 13	Project relevance	Environmental

dry before it was visualized using JOEL JEM-2100 transmission electron microscope.

2.3. Genome sequencing and assembly

The genome of *Pseudomonas moraviensis* strain Devor was sequenced at the University of Georgia Genomics Facility using the Illumina MiSeq platform 2 × 300 paired end chemistry. Average library insert size was 700 bp. The short read de Brujin graph assembly program Velvet [14] was used to assemble quality filtered sequence data using a kmer value of 101 bp and a minimum contig coverage value of 7×. The genome project is deposited in GOLD (Genomes On-Line Database) and this Whole Genome Shotgun (WGS) project has been deposited in GenBank under the accession MAYQ00000000. The version described in this paper is version MAYQ01000000.

2.4. Genome annotation

Gene models were created using the prokaryotic gene calling software package Prodigal [15]. A total of 5473 gene models were predicted with an average gene size of 1002 bp. Predicted protein sequences were annotated using a combination of NCBI Blast C++ homology search, and HMMER 3.0 [16] hmmscan against the PFAM [17] 26.0 database. Additional functional annotations were carried out through the Integrated Microbial Genomes Expert Review (IMG-ER) platform.

2.5. Comparative genomics

We compared the genome of *Pseudomonas moraviensis* strain Devor to 13 closely related genomes (IMG genome IDs: 2619619019, 2654587541, 2603880217, 2561511156, 2639762619, 2636416187, 2639762638, 2639762633, 2597489942, 2639762506, 2517572232, 2639762618, 2556921015) using the "Genome clustering" function on the IMG-ER analysis platform based on the COG profile. We also used principal component analysis to compare the genomes based on several genomic features including the genome size, the number of genes, the number of transporters identified, the GC content, the number of non-coding bases, the number of genes belonging to COG categories, as well as the number of genes belonging to each COG category. The PCA analysis was conducted using the "princomp" function in the labdsv library of R [18]. The results were visualized using a biplot, where genomes were represented by stars and genomic features or COG categories used for comparison were represented by arrows.

3. Results and discussion

3.1. Classification and features

Cells of *P. moraviensis* strain Devor are Gram negative, motile, aerobic rods that are arranged as singles (Fig. 1). Colonies on TSA agar were orange.

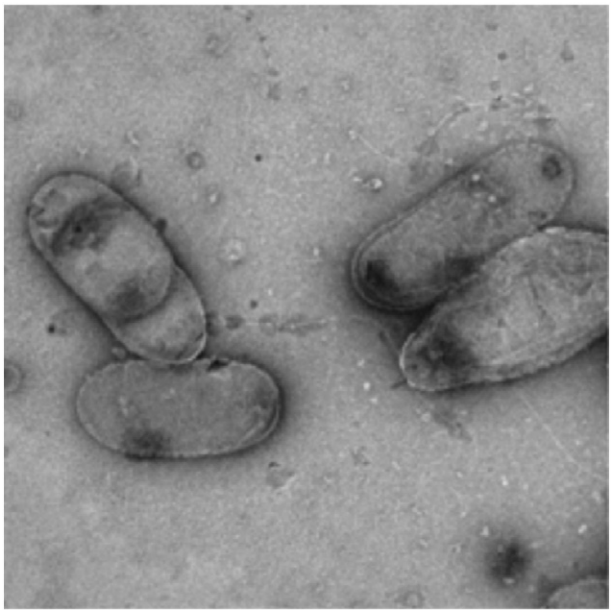

Fig. 1. Negative stain TEM micrograph of *Pseudomonas moraviensis* strain Devor.

Within the genus *Pseudomonas*, 163 species are described with validly published names. Strain Devor shares 77.8%–99.8% 16S rRNA gene identities with other species in the *Pseudomonas* genus as follows: *P. abietaniphilia* (97.7%) type strain ATCC 700689^T, *P. aeruginosa* (94.7%) type strain ATCC 10145^T, *P. aestusnigri* (95.5%) VGXO14^T, *P. agarici* (98%) type strain ATCC 25941^T, *P. alcaligenes* (95.8%) type strain ATCC 14909^T, *P. alcalophilia* (96.5%) type strain AL15-21^T, *P. amygdali* (96.6%) type strain ATCC 33614^T, *P. anguilliseptica* (94.7%) type strain ATCC 33660^T, *P. antarctica* (97.3%) type strain DSM 15318^T, *P. argentinensis* (97.6%) type strain CH01^T, *P. arsenicoxydans* (97.7%) type strain VC-1^T, *P. asplenii* (97.7%) type strain ATCC 23835^T, *P. asturiensis* (97.1%) type strain LPPA 221^T, *P. asuensis* (95.8%) type strain CP 155-2^T, *P. azotifigens* (93.9%) type strain ATCC BAA-1049^T, *P. azotoformans* (98.1%) type strain CCUG 12536^T, *P. baetica* (99.2%) type strain a390^T, *P. balearica* (95%) type strain DSM 6083^T, *P. bauzanensis* (95%) type strain DSM 22558^T, *P. benzenivorans* (96.9%) type strain DSM 8628^T, *P. borbori* (96.1%) type strain DSM 17834, *P. brassicacearum* (97.1%) type strain DSM 13227^T, *P. brenneri* (97.8%) type strain DSM 15294, *P. caeni* (93.8%) type strain HY-14^T, *P. cannabina* (97.8%) type strain CFBP 2341^T, *P. carboxydohydrogena* (77.8%) type strain ATCC 29978^T, *P. caricapapayae* (97.7%) type strain ATCC 33615^T, *P. cedrina* (98.2%) type strain DSM 17516^T, *P. chengduensis* (96.6%) type strain DSM 26382, *P. chloritidismutans* (96.6%) type strain ATCC BAA-443^T, *P. chlororaphis* (97.9%) type strain ATCC 9446^T, *P. cichorii* (96.8%) type strain ATCC 10857^T, *P. citronellosis* (94.4%) type strain ATCC 13674^T, *P. coleopterorum* (97.8%) type strain LMG 28558^T, *P. composti* (96.7%) type strain C2^T, *P. congealans* (97.7%) type strain DSM 14939^T, *P. corrugata* (97.7%) type strain ATCC 29736^T, *P. constantinii* (97.4%) type strain PS 3a^T, *P. cremoricolorata* (97.1%) type strain DSM 17059^T, *P. cuatrocienegasensis* (96.5%) type strain LMG 24676^T, *P. deceptionensis* (97.2%) type strain M1^T, *P. delhiensis* (94.3%) type strain RLD-1^T, *P. duriflava* (94.7%) type strain HR2^T, *P. elongata* (89.3%) type strain ATCC 10144^T, *P. endophytica* (97.3%) type strain BSTT44^T, *P. entomophila* (97.7%) type strain L48^T, *P. extremaustralis* (97.3%) type strain DSM 17835^T, *P. extremorientalis* (97.5%) type strain KMM 3447^T, *P. ficuserectae* (97.6%) type strain ATCC 35104^T, *P. flavescens* (96.5%) type strain ATCC 51555^T, *P. fluorescens* (96.8%) type strain ATCC 13525^T, *P. formosensis* (94.1%) type strain CC-CY503^T, *P. fragi* (97.2%) type strain ATCC 4973^T, *P. frederiksbergensis* (97.3%) type strain ATCC BAA-257^T, *P. fulva* (98.6%) type strain ATCC 31418^T, *P. fuscovaginae* (97.5%) type strain DSM 7231^T, *P. gessardii* (98.2%) type strain CIP 105469^T, *P. graminis*

(98%) type strain DSM 11363^T, *P. granadensis* (99.4%) type strain DSM 28040^T, *P. grimontii* (97.4%) type strain ATCC BAA-140^T, *P. guangdongensis* (95.1%) type strain SgZ-6^T, *P. guariconensis* (97.2%) type strain PCAVU11^T, *P. guineae* (95.5%) type strain M8^T, *P. helmanticensis* (98.8%) type strain OHA11^T, *P. hussainii* (93.8%) type strain CC-AMH-11^T, *P. indica* (94.1%) type strain DSM 14015^T, *P. japonica* (96.9%) type strain WL^T, *P. jessennii* (99.3%) type strain ATCC 700870^T, *P. jinjuensis* (95%) type strain Pss 26^T, *P. kilonensis* (97.7%) type strain DSM 13647^T, *P. knackmussii* (94.5%) type strain DSM 6978^T, *P. koreensis* (99.8%) type strain Ps 9-14^T, *P. kunmingensis* (96.3%) type strain DSM 25974^T, *P. kuykendallii* (97.7%) type strain H2^T, *P. libanensis* (98.2%) type strain CCUG 43190^T, *P. lini* (97.7%) type strain DLE411J^T, *P. linyingensis* (95%) type strain LYBRD3-7^T, *P. litoralis* (94.8%) type strain 2SM5^T, *P. lundensis* (97.5%) type strain ATCC 49968^T, *P. lurida* (97.4%) type strain DSM 15835^T, *P. lutea* (98.2%) type strain OK2^T, *P. luteola* (95.6%) type strain ATCC 43273^T, *P. mandelii* (97.8%) type strain ATCC 700871^T, *P. marginalis* (97.2%) type strain ATCC 10844^T, *P. marincola* (96%) type strain JCM 14761^T, *P. matsuisoli* (93.8%) type strain CC-MHH0089, *P. mediterranea* (97%) type strain CFBP 5447^T, *P. mendocina* (96.2%) type strain ATCC 25411^T, *P. meridiana* (97.1%) type strain DSM 15319^T, *P. migulae* (98.1%) type strain CCUG 43165^T, *P. mohnii* (99.2%) type strain DSM 18327^T, *P. monteilii* (97.8%) type strain ATCC 700476^T, *P. moorei* (99.3%) type strain DSM 12647^T, *P. moraviensis* (99.4%) type strain DSM 16007^T, *P. mosselii* (97.6%) type strain ATCC BAA-99^T, *P. mucidolens* (98.2%) type strain ATCC 4685^T, *P. multiresinivorans* (94.8%) type strain ATCC 700690^T, *P. nitritireducens* (94.8%) type strain WZBFD3-5A2^T, *P. nitroreducens* (94.9%) type strain ATCC 33634^T, *P. oleovorans* (94.4%) type strain ATCC 8062^T, *P. orientalis* (97.1%) type strain DSM 17489^T, *P. oryzihabitans* (97.3%) type strain ATCC 43272^T, *P. otitidis* (95%) type strain ATCC BAA-1130^T, *P. pachastrellae* (95.2%) type strain CCUG 46540^T, *P. palleroniana* (97.4%) type strain CFBP 4389^T, *P. panacis* (97.7%) type strain CG20106^T, *P. panipatensis* (94.5%) type strain Esp-1, *P. parafulva* (98.6%) type strain DSM 17004^T, *P. pelagia* (94.3%) CL-AP6^T, *P. peli* (95.5%) type strain DSM 17833^T, *P. pertucinogena* (93.3%) type strain ATCC 190^T, *P. pictorum* (84.4%) type strain ATCC 23328^T, *P. plecoglossicida* (97.7%) type strain ATCC 700383^T, *P. poae* (97.2%) type strain DSM 14936^T, *P. pohangensis* (95.7%) type strain DSM 17875^T, *P. prosekii* (97.8%) type strain AN/28/1^T, *P. protegens* (97%) type strain DSM 19095^T, *P. proteolytica* (97.4%) type strain DSM 15321^T, *P. pseudoalcaligenes* (96.9%) type strain ATCC 17440^T, *P. psychrophilia* (97.1%) type strain E-3^T, *P. psychrotolerans* (94.7%) type strain DSM 15758^T, *P. punonensis* (97.7%) type strain LMT03^T, *P. putida* (97.2%) type strain ATCC 12633^T, *P. reinekei* (99.5%) type strain DSM 18361^T, *P. resinovorans* (93.6%) type strain ATCC 14235^T, *P. rhizosphaerae* (97.8%) type strain IH5^T, *P. rhodesiae* (97.4%) type strain DSM 14020^T, *P. sabulinigri* (94.9%) type strain J64^T, *P. sagittaria* (94.8%) type strain CC-OPY-1^T, *P. salegens* (94%) type strain GBPy5^T, *P. salina* (94.3%) type strain XCD-X85^T, *P. salomonii* (96.9%) type strain CFBP 2022^T, *P. saponiphilia* (96.5%) type strain DSM 9751^T, *P. savastanoi* (97.6%) type strain ATCC 13522^T, *P. segetis* (96.3%) type strain FR1439^T, *P. seleniipraecipitans* (97%) type strain CA5^T, *P. simiae* (97.5%) type strain OLi^T, *P. soli* (97.8%) type strain DSM 28043^T, *P. straminea* (97%) type strain ATCC 33636^T, *P. stutzeri* (96.4%) type strain ATCC 17588^T, *P. synxantha* (98.1%) type strain ATCC 9890^T, *P. syringae* (97.6%) type strain ATCC 19310^T, *P. taenensis* (96.9%) type strain MS-3^T, *P. taetrolans* (97.1%) type strain ATCC 4683^T, *P. taiwanensis* (97.8%) type strain DSM 21245^T, *P. testosteroni* (80.8%) type strain ATCC 11996^T, *P. thivervalensis* (97.6%) type strain DSM 13194^T, *P. tolaasii* (97.4%) type strain ATCC 33618^T, *P. toyotomiensis* (96.5%) type strain HT-3^T, *P. tremae* (97.8%) type strain CFBP 3229^T, *P. trivialis* (97.1%) type strain DSM 14937^T, *P. tuomuerensis* (95.2%) type strain 78-123, *P. usmongensis* (98.9%) type strain Ps 3-10^T, *P. vancouverensis* (99.3%) type strain ATCC 700688^T, *P. veronii* (97.6%) type strain ATCC 700272/700474^T, *P. viridiflava* (96.2%) type strain ATCC 13223^T, *P. vranovensis* (96.7%) type strain DSM 16006^T, *P. xanthomarina* (96.5%) type strain CCUG 46543^T, *P. xiamenensis* (93.5%) type strain C10-2^T, *P. xinjiangensis* (94.5%) type strain S3-3^T, *P. yamanorum* (98%) type strain DSM 26522^T, *P. zeshuii* (95.4%) type strain BY-1^T, *P. zhaodongensis* (96.6%) type strain DSM 27559^T.

Phylogenetic analysis based on the 16S rRNA gene identified strain WW1 to be the closest taxonomic relative of *Pseudomonas moraviensis* strain Devor (Table 2, and Fig. 2).

Comparison to other *Pseudomonas moraviensis* sequenced genomes revealed that strain Devor shares 99% 16S rRNA sequence similarity to strain R28-S (Genbank accession number AYMZ00000000) [5].

3.2. Genome properties

The genome assembly process produced a contig N50 of 3,923,117 bp and a total genome size of 6,016,489 bp. The GC content was 60.0%. One hundred and eighty three RNA genes were identified including 3 ribosomal RNA and 63 tRNA genes. The ribosomal RNA operon showed an atypical organization, where 2 copies of 5S rRNA were present along with one copy of the 23S rRNA gene. Two copies each of the tRNA^{Ala} and the tRNA^{Ile} were present in the operon. The operon lacked a 16S rRNA gene. Of the 5473 genes detected, 5290 were protein-coding, of which 79.8% had a function prediction, 72.7% represented a COG functional category, and 11.3% were predicted to have a signal peptide. Using Psort [19], proteins were classified as 44.12% cytoplasmic, 1.06% extracellular, and 25.25% associated with the membrane. Based on the presence of 139 single copy genes [20], the genome is predicted to be 83.45% complete. Genome statistics are shown in Table 3. The distribution of genes into COG functional categories is shown in Table 4.

3.3. Insights from the genome sequence

Genome analysis of *Pseudomonas moraviensis* strain Devor identified a Gram negative microorganism with typical cell wall structure, a peptidoglycan layer lacking pentaglycine bridges with diaminopimelic

Table 2
Classification and general features of *Pseudomonas moraviensis* strain Devor [28].

MIGS ID	Property	Term	Evidence code[a]
	Classification	Domain *Bacteria*	TAS [9]
		Phylum *Proteobacteria*	TAS [9]
		Class *Gammaproteobacteria*	TAS [9]
		Order *Pseudomonadales*	TAS [9]
		Family *Pseudomonadaceae*	TAS [9]
		Genus *Pseudomonas*	TAS [9]
		Species *moraviensis*	TAS [9]
		(Type) strain: *Devor*	
	Gram stain	Negative	TAS [9]
	Cell shape	Rod	TAS [9]
	Motility	Motile	TAS [9]
	Sporulation	Non-spore forming	TAS [9]
	Temperature range	4–35 °C	TAS [9]
	Optimum temperature	28–35 °C	TAS [9]
	pH range; optimum	Unknown	
	Carbon source	glycerol, L-arabinose, D-glucose, D-fructose, D-trehalose, D-lyxose, gluconate and 2-ketogluconate	TAS [9]
MIGS-6	Habitat	Soil/Gate	IDA
MIGS-6.3	Salinity	Unknown	
MIGS-22	Oxygen requirement	Aerobic	IDA
MIGS-15	Biotic relationship	Free living	IDA
MIGS-14	Pathogenicity	Non-pathogenic	NAS
MIGS-4	Geographic location	Stillwater, OK, USA	IDA
MIGS-5	Sample collection	March 2016	IDA
MIGS-4.1	Latitude	36.1157	IDA
MIGS-4.2	Longitude	−97.0586	IDA
MIGS-4.4	Altitude	1 M	

[a] Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [29].

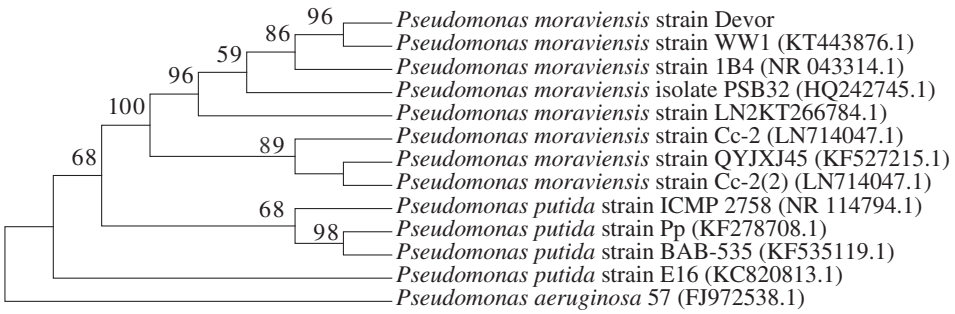

Fig. 2. A maximum likelihood phylogenetic tree constructed using multiple sequence alignments of 16S rRNA genes. "*Pseudomonas moraviensis* strain Devor" sequence is shown with in bold. Reference sequences are also shown and Genbank accession numbers are given in parentheses. The tree was obtained under "K2 + G" model with a variable site γ shape parameter of 0.05. "*Pseudomonas aeruginosa 57 16S ribosomal RNA gene partial sequence*" was used as the outgroup. Bootstrap values, in percent, are based on 100 replicates and are shown for branches with >50% bootstrap support. Multiple sequence alignment, model selection, and maximum likelihood analysis were carried out in Mega [30].

acid as the second amino acid in the peptide linkage. Genes encoding enzymes for the biosynthesis of the glycerophospholipids phosphatidylcholine, phosphatidyl-glycerophosphate, phosphatidyl-serine, and phosphatidyl ethanolamine were identified in the genome. The analysis also revealed the presence of genes encoding a complete flagellar assembly in line with the evidence for the presence of flagella in electron micrographs of the organism (Fig. 1). Extracellular structures including type IV pili and the Flp system, and some structural genes for the production of exopolysaccharides including a partial Yjb operon [21], as well as evidence for the biosynthesis of multiple nucleotide sugars were also found in the genome.

Further analysis suggested the capability to utilize xylose, fructose, mannose, glucose, all amino acids except valine and tryptophan, and fatty acids as carbon and energy sources. Almost complete to complete catabolic KEGG pathways were identified in the genome for each of the above carbon sources. A complete TCA cycle and electron transport chain with F-type ATPase subunits confirmed the aerobic nature of the microorganism. Facultative fermentation capability was also identified in the genome where genes for lactate and acetate fermentation were present. Comparison of the protein-coding genes against the transporter database [22] identified ABC and secondary transporters for several amino acids.

When compared against the virulence factor database [23], the genome of *Pseudomonas moraviensis* strain showed 1733 virulence factor hits (32.8% of the protein-coding genes). These included secretion systems (type I, type II, and type IV secretion systems) among others.

The Devor genome also encoded several proteins that suggest the potential for using this strain in bioremediation efforts. These included genes encoding for 4-hydroxybenzoate degradation to succinate, as well as genes encoding for the detoxification of the plant-secreted toxin nitronate and the industry-applied chelating agent nitriloacetate [24–27]. Evidence for the capability of salvaging S from organo-S-compounds (e.g. taurine and alkanesulfonates) in cases of limiting inorganic S were also identified in the genome.

Table 3
Genome statistics.

Attribute	Value	% of Total
Genome size (bp)	6,016,489	83.4
DNA coding (bp)	5,314,299	88.3
DNA G + C (bp)	3,609,956	60.0
DNA scaffolds	7	100
Total genes	5473	100
Protein coding genes	5290	96.6
RNA genes	183	3.3
Pseudo genes	0	0
Genes in internal clusters	1440	26.3
Genes with function prediction	4220	77.1
Genes assigned to COGs	3846	70.3
Genes with Pfam domains	4512	82.4
Genes with signal peptides	598	10.9
Genes with transmembrane helices	1208	22.1
CRISPR repeats	0	0

3.4. Insights from comparative genomics

We used the COG profile to compare the genome of *Pseudomonas moraviensis* strain Devor to 13 closely related genomes. The genome clustered with *Pseudomonas moraviensis* strain BS3668 (Fig. 3A), which was in agreement with the phylogenetic position of the isolate as *Pseudomonas moraviensis* (Fig. 2). We used genomic features including the genome size, the number of genes, the number of transporters identified, the GC content, the number of non-coding bases, the number of genes belonging to COG categories, as well as the number of genes belonging to each COG category to cluster *Pseudomonas moraviensis* strain Devor's genome in comparison to the 13 other closely related genomes. Results are shown in Fig. 3B. The genome *Pseudomonas moraviensis* strain Devor clustered close to the two *P. moraviensis* genomes, and the *P. koorensis* genome as well. This is consistent with its phylogeny as a *P. moraviensis* and as a member of the *P. koorensis* clade [9]. These genomes appear to be enriched in the number of transporters identified in the genome.

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

Code	Value	% age	Description
J	236	5.40	Translation, ribosomal structure and biogenesis
A	2	0.05	RNA processing and modification
K	393	9.00	Transcription
L	124	2.84	Replication, recombination and repair
B	2	0.05	Chromatin structure and dynamics
D	40	0.92	Cell cycle control, cell division, chromosome partitioning
V	96	2.20	Defense mechanisms
T	286	6.55	Signal transduction mechanisms
M	270	6.18	Cell wall/membrane biogenesis
N	123	2.82	Cell motility
U	89	2.04	Intracellular trafficking and secretion
O	157	3.60	Posttranslational modification, protein turnover, chaperones
C	276	6.32	Energy production and conversion
G	230	5.27	Carbohydrate transport and metabolism
E	447	10.24	Amino acid transport and metabolism
F	96	2.20	Nucleotide transport and metabolism
H	206	4.72	Coenzyme transport and metabolism
I	213	4.88	Lipid transport and metabolism
P	253	5.79	Inorganic ion transport and metabolism
Q	121	2.77	Secondary metabolites biosynthesis, transport and catabolism
R	387	8.86	General function prediction only
S	263	6.02	Function unknown
–	1640	29.97	Not in COGs

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

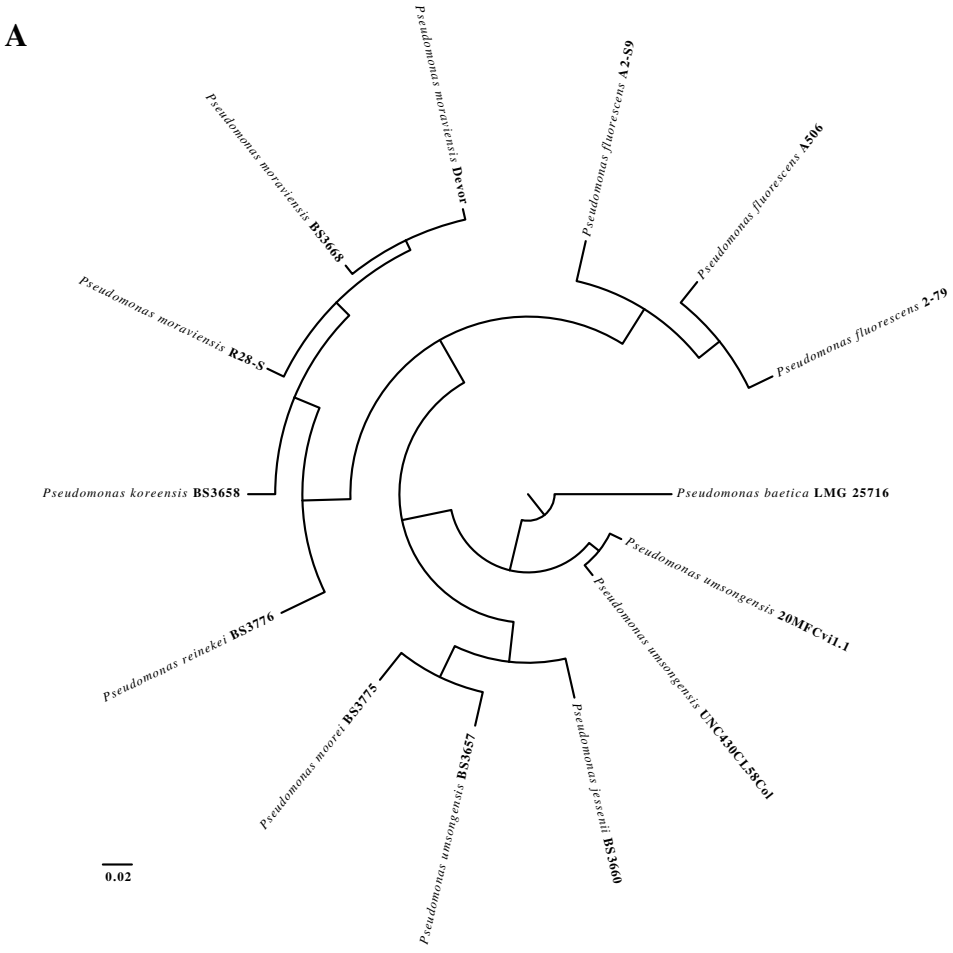

Fig. 3. Comparative genomics of *Pseudomonas moraviensis* strain Devor and 13 closely related genomes. (A) COG profile clustering of the genomes compared in this study. (B) Principal component analysis biplot of the genomic features and COG category distribution in the genomes compared. Genomes are represented by stars. Arrows represent genomic features or COG categories used for comparison. The arrow directions follow the maximal abundance, and their lengths are proportional to the maximal rate of change between genomes. The first two components explained 75% of variation. Color coding: *P. jessenei*, brown; *P. baetica*, orange; *P. umsongenesis*, purple; *P. morviensis*, red; *P. reineke*, grey; *P. fluorescens*, green; *P. koorensis*, black; *P. moorei*, blue.

4. Conclusions

This study presents the genome sequence and annotation of *Pseudomonas moraviensis* strain Devor. The genome revealed metabolic versatility with suggested capability to degrade fatty acids, most amino acids, fructose, mannose, and xylose. Comparison to the virulence factor database identified 1733 genes in the genome with potential virulence-associated function including type I, type II, and type IV secretion systems. The genome also suggests the potential use in bioremediation efforts with evidence for degradation of 4-hydroxybenzoate, nitronate, and nitriloacetate. Comparative genomics using general genomic features and the COG function profile are in line with the phylogenetic position predicted based on the 16S rRNA gene sequence placing strain Devor with other *P. moraviensis* representatives.

Competing interests

All authors declare no competing interests.

Funding

This is Draft Genome #7 in the SIMD project supported in part by a grant from the Howard Hughes Medical Institute (1554854) through the Science Education Program. WDH acknowledges support by NSF grants MCB-1051590, MRI-1338097, and CHE-1412500.

Authors' contributions

NTM, DF, MB, PB, MBC, and NY contributed to the analysis. NTM, DF, WDH, DPF, and NY wrote the manuscript. RCD, CB, AR, and RAH performed the lab experiments.

Transparency document

The Transparency document associated with this article can be found, in the online version.

References

[1] M.B. Couger, A. Hurlbut, C.L. Murphy, C. Budd, D.P. French, W.D. Hoff, M.S. Elshahed, N.H. Youssef, Draft genome sequence of the environmental isolate *Chryseobacterium* sp. Hurlbut01. Genome Announc 3 (5) (2015).
[2] R.J. Ash, B. Mauck, M. Morgan, Antibiotic resistance of gram-negative bacteria in rivers, United States. Emerg. Infect. Dis. 8 (7) (2002) 713–716.
[3] J.C. Cho, J.M. Tiedje, Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. Appl. Environ. Microbiol. 66 (12) (2000) 5448–5456.
[4] S. Grobe, J. Wingender, H.G. Truper, Characterization of mucoid *Pseudomonas aeruginosa* strains isolated from technical water systems. J. Appl. Bacteriol. 79 (1) (1995) 94–102.
[5] S.S. Hunter, H. Yano, W. Loftie-Eaton, J. Hughes, L. De Gelder, P. Stragier, P. De Vos, M.L. Settles, E.M. Top, Draft genome sequence of *Pseudomonas moraviensis* R28-S. Genome Announc 2 (1) (2014).
[6] A. Kumar, A. Kumar, Isolation of a *Pseudomonas aeruginosa* strain capable of degrading acrylamide. J. Microbiol. Biotechnol. 8 (4) (1998) 347–352.
[7] A. Pandey, L.M.S. Palni, Isolation of *Pseudomonas corrugata* from Sikkim Himalaya. World J. Microbiol. Biotechnol. 14 (3) (1998) 411–413.
[8] A. Pandey, P. Trivedi, B. Kumar, L.M. Palni, Characterization of a phosphate solubilizing and antagonistic strain of *Pseudomonas putida* (B0) isolated from a sub-alpine location in the Indian Central Himalaya. Curr. Microbiol. 53 (2) (2006) 102–107.
[9] L. Tvrzova, P. Schumann, C. Sproer, I. Sedlacek, Z. Pacova, O. Sedo, Z. Zdrahal, M. Steffen, E. Lang, Pseudomonas moraviensis sp. nov. and *Pseudomonas vranovensis* sp. nov., soil bacteria isolated on nitroaromatic compounds, and emended description of *Pseudomonas asplenii*. Int. J. Syst. Evol. Microbiol. 56 (Pt 11) (2006) 2657–2663.
[10] S. Verhille, N. Baida, F. Dabboussi, D. Izard, H. Leclerc, Taxonomic study of bacteria isolated from natural mineral waters: proposal of *Pseudomonas jessenii* sp. nov. and *Pseudomonas mandelii* sp. nov. Syst. Appl. Microbiol. 22 (1) (1999) 45–58.
[11] W. Yang, H. Cao, L. Xu, H. Zhang, Y. Yan, A novel eurythermic and thermostable lipase LipM from *Pseudomonas moraviensis* M9 and its application in the partial hydrolysis of algal oil. BMC Biotechnol. 15 (2015) 94.
[12] K.-P. Yoon, Isolation and characterization of *Pseudomonas* sp. KM10, a cadmium- and mercury-resistant, and phenol-degrading bacterium. J. Microbiol. Biotechnol. 8 (4) (1998) 388–398.
[13] L.C. Staicu, C.J. Ackerson, P. Cornelis, L. Ye, R.L. Berendsen, W.J. Hunter, S.D. Noblitt, C.S. Henry, J.J. Cappa, R.L. Montenieri, et al., *Pseudomonas moraviensis* subsp. stanleyae, a bacterial endophyte of hyperaccumulator *Stanleya pinnata*, is capable of efficient selenite reduction to elemental selenium under aerobic conditions. J. Appl. Microbiol. 119 (2) (2015) 400–410.
[14] D.R. Zerbino, E. Birney, Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18 (5) (2008) 821–829.
[15] D. Hyatt, G.L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer, L.J. Hauser, Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf. 11 (2010) 119.
[16] J. Mistry, R.D. Finn, S.R. Eddy, A. Bateman, M. Punta, Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 41 (12) (2013), e121.
[17] R.D. Finn, P. Coggill, R.Y. Eberhardt, S.R. Eddy, J. Mistry, A.L. Mitchell, S.C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, et al., The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44 (D1) (2016) D279–D285.
[18] D. Roberts, labdsv: ordination and multivariate analysis for ecology. 2007.
[19] P. Horton, K.-J. Park, T. Obayashi, N. Fujita, H. Harada, C.J. Adams-Collier, K. Nakai, WoLF PSORT: protein localization predictor. Nucleic Acids Res. 35 (2007) W585–W587 Web Server issue.
[20] C. Rinke, P. Schwientek, A. Sczyrba, N.N. Ivanova, I.J. Anderson, J.F. Cheng, A. Darling, S. Malfatti, B.K. Swan, E.A. Gies, et al., Insights into the phylogeny and coding potential of microbial dark matter. Nature 499 (7459) (2013) 431–437.
[21] L. Ferrieres, S.N. Aslam, R.M. Cooper, D.J. Clarke, The yjbEFGH locus in *Escherichia coli* K-12 is an operon encoding proteins involved in exopolysaccharide production. Microbiology 153 (Pt 4) (2007) 1070–1080 Reading, England.
[22] M.H. Saier Jr., V.S. Reddy, D.G. Tamang, A. Vastermark, The transporter classification database. Nucleic Acids Res. 42 (Database issue) (2014) D251–D258.
[23] L. Chen, J. Yang, J. Yu, Z. Yao, L. Sun, Y. Shen, Q. Jin, VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 33 (2005) D325–D328 Database issue.
[24] K. Francis, C. Smitherman, S.F. Nishino, J.C. Spain, G. Gadda, The biochemistry of the metabolic poison propionate 3-nitronate and its conjugate acid, 3-nitropropionate. IUBMB life 65 (9) (2013) 759–768.
[25] F. Salvi, J. Agniswamy, H. Yuan, K. Vercammen, R. Pelicaen, P. Cornelis, J.C. Spain, I.T. Weber, G. Gadda, The combined structural and kinetic characterization of a bacterial nitronate monooxygenase from *Pseudomonas aeruginosa* PAO1 establishes NMO class I and II. J. Biol. Chem. 289 (34) (2014) 23764–23775.
[26] Y. Xu, M.W. Mortimer, T.S. Fisher, M.L. Kahn, F.J. Brockman, L. Xun, Cloning, sequencing, and analysis of a gene cluster from *Chelatobacter heintzii* ATCC 29600 encoding nitrilotriacetate monooxygenase and NADH:flavin mononucleotide oxidoreductase. J. Bacteriol. 179 (4) (1997) 1112–1116.
[27] Y. Zhang, T.E. Edwards, D.W. Begley, A. Abramov, K.B. Thompkins, M. Ferrell, W.J. Guo, I. Phan, C. Olsen, A. Napuli, et al., Structure of nitrilotriacetate monooxygenase component B from *Mycobacterium thermoresistibile*. Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun. 67 (Pt 9) (2011) 1100–1105.
[28] D. Field, G. Garrity, T. Gray, N. Morrison, J. Selengut, P. Sterk, T. Tatusova, N. Thomson, M.J. Allen, S.V. Angiuoli, et al., The minimum information about a genome sequence (MIGS) specification. Nat. Biotechnol. 26 (5) (2008) 541–547.
[29] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, et al., Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25 (1) (2000) 25–29.
[30] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7) (2016) 1870–1874.