Polyphasic characterization of four soil-derived phenanthrene-degrading *Acidovorax* strains and proposal of *Acidovorax carolinensis* sp. nov.

David R. Singleton[1,*], **Janice Lee**[2], **Allison N. Dickey**[3], **Aaron Stroud**[2], **Elizabeth H. Scholl**[3], **Fred A. Wright**[3], and **Michael D. Aitken**[1]

[1]Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC 27599-7431.

[2]Department of Biology, University of North Carolina, Chapel Hill, NC, 27599-3280

[3]Bioinformatics Research Center, North Carolina State University, Raleigh, North Carolina, USA
11

Abstract

Four bacterial strains identified as members of the *Acidovorax* genus were isolated from two geographically distinct but similarly contaminated soils in North Carolina, USA, characterized, and their genomes sequenced. Their 16S rRNA genes were highly similar to those previously recovered during stable-isotope probing (SIP) of one of the soils with the polycyclic aromatic hydrocarbon (PAH) phenanthrene. Heterotrophic growth of all strains occurred with a number of organic acids, as well as phenanthrene, but no other tested PAHs. Optimal growth occurred aerobically under mesophilic temperature, neutral pH, and low salinity conditions. Predominant fatty acids were $C_{16:1}\omega7c/C_{16:1}\omega6c$, $C_{16:0}$, and $C_{18:1}\omega7c$, and were consistent with the genus. Genomic G+C contents ranged from 63.6 to 64.2%. A combination of whole genome comparisons and physiological analyses indicated that these four strains likely represent a single species within the *Acidovorax* genus. Chromosomal genes for phenanthrene degradation to phthalate were nearly identical to highly conserved regions in phenanthrene-degrading *Delftia*, *Burkholderia*, *Alcaligenes*, and *Massilia* species in regions flanked by transposable or extrachromosomal elements. The lower degradation pathway for phenanthrene metabolism was inferred by comparisons to described genes and proteins. The novel species *Acidovorax carolinensis* sp. nov. is proposed, comprising the four strains described in this study with strain NA3T as the type strain (=LMG 30136, =DSM 105008).

Keywords

Acidovorax; phenanthrene; polycyclic aromatic hydrocarbons

*Corresponding author: david.r.singleton@duke.edu.
Present address: David Singleton, Pratt School of Engineering, Department of Civil & Environmental Engineering, 121 Hudson Hall, Box 90287, Duke University, Durham NC 27708, Phone: 1-919-668-9104, Fax: 1-919-660-5219

Introduction

Bacteria within the *Acidovorax* genus are found in a diverse range of habitats and have varied functions within those environments. Strains now classified as *Acidovorax* have been associated with soils [11, 45], wastewater treatment plants [23, 46], plants [19, 20, 35, 63], and clinical samples [59]. Mirroring the diversity inherent in this widespread distribution, isolates within this diverse genus have been linked to a variety of phenotypes ranging from plant pathogenicity, to denitrification, to the biodegradation of contaminants.

Environmentally derived 16S rRNA gene sequences from the *Acidovorax* genus have frequently been associated with samples impacted by petroleum contamination which typically contain high concentrations of polycyclic aromatic hydrocarbons (PAHs) [18, 26, 41, 58]. Of particular interest for PAH bioremediation, the isolates *Acidovorax delafieldii* strains P4–1 [44] and TNA921 [47], *Acidovorax temperans* strain GTI-19 [6], and several uncharacterized *Acidovorax* strains [40] have been demonstrated to grow on the three-ring, low-molecular-weight (LMW) PAH phenanthrene. In a prior study of weathered PAH-contaminated soils treated via biostimulation in *ex situ*, lab-scale, slurry-phase bioreactors, we directly linked the degradation of phenanthrene to uncultivated members of the *Acidovorax* genus through stable-isotope probing (SIP) [25, 49]. Additional SIP experiments of bacterial PAH degraders in a similarly contaminated soil from France also directly linked *Acidovorax* organisms to the disappearance of phenanthrene in those samples [38]. We subsequently isolated several strains of phenanthrene-degrading bacteria, including one (designated *Acidovorax* strain NA3) with high 16S rRNA gene sequence identity to SIP-detected *Acidovorax* sequences that could grow on the PAH phenanthrene [50].

A number of *Acidovorax* strains have publicly available draft or finished genomes. While some of these represent *Acidovorax* strains capable of the metabolism of aromatic compounds (e.g., 2-nitrotoluene by *Acidovorax* sp. JS42 [22] or biphenyl and polychlorinated biphenyls by *Acidovorax* sp. KKS102 [27]), no public *Acidovorax* genomes are available from organisms capable of degrading any PAH. In fact, there is little genetic information available for the strains previously shown to grown on phenanthrene, with only a partial 16S rRNA gene sequence for *A. delafieldii* P4–1 available (GenBank accession number DQ282184). The only published work on the genetics of PAH metabolism within the genus was performed on *Acidovorax* strain NA3T using partial sequences recovered from a fosmid library [50]. In that work, the genes likely to be associated with the first three steps of the "upper pathway" of PAH degradation in NA3T (comprising a ring-hydroxylating dioxygenase [RHD], dihydrodiol dehydrogenase, and ring- cleavage dioxygenase) were shown to be highly similar to homologous genes in phenanthrene- degrading members of the *Alcaligenes*, *Delftia*, and *Burkholderia* genera within the Betaproteobacterial order *Bulkholderiales*. Although the genes in strain NA3T were not directly linked to the transformation of phenanthrene, RNA transcripts derived from those genes were induced by the presence of either of the LMW PAHs phenanthrene or naphthalene added to the cell culture. A recent publication analyzing the genome of *Delftia* strain Cs1–4, a strain with putative phenanthrene-degradation genes nearly identical to those in *Acidovorax* NA3T, revealed a genomic "phn island," a 232 kb region of the chromosome flanked by transposable elements containing the *phn* genes thought to be responsible for phenanthrene

degradation [24]. It was suggested in that work that the genomes of organisms harboring highly similar genes, such as phenanthrene-degrading *Acidovorax* spp., may demonstrate a similar gene organization.

In this manuscript, we describe four *Acidovorax* strains isolated from two geographically distinct, weathered PAH-contaminated soils from former manufactured gas plant (MGP) sites in North Carolina, USA. Each isolate was characterized using a variety of physiological tests and the complete genome of each strain was determined. The 16S rRNA genes from all of the strains were highly similar to sequences derived from phenanthrene SIP [25, 49], and all were capable of growth on phenanthrene as a sole source of carbon and energy. The genome of each isolate was examined for genes and genomic islands potentially associated with phenanthrene metabolism. Based on genotypic and physiological differences to characterized *Acidovorax* species, *Acidovorax carolinensis* sp. nov. is proposed to encompass all four strains described herein, with strain NA3 T as the designated type strain.

Methods and Materials

Strain isolation.

Acidovorax sp. strains P3 and P4 were isolated from crystalline phenanthrene- enriched, PAH-contaminated soil collected from the site of a former MGP site in Salisbury, NC, USA that had previously undergone treatment in a lab-scale, slurry-phase aerobic bioreactor [51]. Pure cultures of strains P3 and P4 were obtained by serial dilution of phenanthrene-enriched bioreactor slurry on nutrient agar (NA) plates (Difco, Becton, Dickinson and Company, Franklin Lakes, NJ, USA). Well-separated colonies were screened by PCR using *Acidovorax*-selective primers and conditions as previously described [48]. Strains P3 and P4 specifically were selected for analysis based on genomic differences among 10 PCR-verified *Acidovorax* isolates (see below). Strains NA2 and NA3T were previously isolated in a similar manner, except that the source soil for treatment in the bioreactor was obtained from a former MGP site in Charlotte, NC, USA, and treated under slightly different conditions [50].

Determination of optimal growth conditions.

Strains were generally maintained on either NA or R2A plates (Difco), or in nutrient broth (NB; Difco). For determining optimal growth conditions for each strain, cultures in triplicate 5-mL tubes of NB were tested at 25°C, 27.5°C, 30°C, 32.5°C, and 35°C with shaking at 250 rpm. Optimal exponential growth rates were determined by measuring turbidity at OD_{600} using a DR/3000 Spectrophotometer (Hach, Loveland, CO, USA). Growth at 4°C, 15°C, 37°C, and 40°C was tested by examining NA plates for growth after 2 weeks. Growth under varying pH was determined by buffering triplicate 5-mL NB tubes with 50 mM of (2-(N-morpholino)ethanesulfonic acid) monohydrate (St. Louis, MO; pH 5.5, 6.0, and 6.5), (N-[2-hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]) (Acros Organics, NJ; pH 7.0, 7.5 and 8.0), or Tris·HCl (pH 8.9 and 9.0), and incubating at 32.5°C and 250 rpm for 7 hours. Growth in NB (pH 7.0, 32.5°C) amended with 0%, 1%, 2%, 3%, 4%, and 5% NaCl was also tested. Production of indigo was examined by amending NA plates with 1 mM indole (Aldrich Chemical Co., Milwaukee, WI, USA). Growth under low oxygen conditions was

tested by incubating NA plates in a sealed vessel in the presence of a GasPak™ EZ anaerobe sachet (BD, Franklin Lakes, NJ) for two weeks at room temperature (~23°C). Unless otherwise stated, strains were routinely grown aerobically in NB buffered to pH 7.5 with no additional salt, and incubated 32.5°C.

Cellular morphology.

The morphology of overnight cultures was determined by scanning electron microscopy. Weak pellets were obtained by centrifugation for 2 minutes at ~1000 x g and the pellets washed twice with phosphate buffered saline (PBS).Specimens were observed and images acquired as previously described [13].. Motility of the strains was further examined using tubes of nutrient broth containing 0.3% agar (motility agar). Stabs of separate single colonies grown on NA plates were inoculated into triplicate tubes and incubated statically at 32.5°C. Cell suspensions in PBS derived from motility agar tube cultures were stained with Remel™ Flagella Stain (Thermo Scientific, Waltham, MA) according to manufacturer's directions and examined under oil-immersion light microscopy.

Chemotaxonomic and physiological characteristics.

Catalase activity was tested by adding 3% hydrogen peroxide (v/v) solution to cells freshly scraped from the surface of an NA plate. Oxidase activity was determined by adding a few drops of freshly-prepared 1% N',N',N',N'-tetramethyl- *p*-phenylenediamine dihydrochloride (Acros Organics, NJ) to cells scraped from a plate onto filter paper [30]. DNase activity of the strains was tested by streaking each isolate on DNase agar plates with toluidine blue (Remel, Lenaxa, KA, USA). Starch hydrolysis was examined using starch agar plates stained with Gram's iodine solution after 48 hours of growth [32]. Gelatin hydrolysis was tested with a nutrient gelatin medium and examining for liquefaction of the media [15]. Protease activity was analyzed by streaking each of the strains on nutrient agar plates amended with 0.1% casein, as well as skim milk agar plates [2]. Urease activity was examined using slants of Christensen's Urea Agar [12]. Nitrate reduction was determined by growing cells in triplicate tubes of nitrate broth (0.5% peptone, 0.3% beef extract, 0.1% KNO_3), before adding several drops each of nitrate test reagents A and B (Gibson Bioscience, Lexington, KY, USA). Any negative results were confirmed by the addition of a small quantity of elemental zinc dust. Cellulolytic activity was tested on three types of media: NA, R2A, or MSB [53], each amended with 1% cellulose. Lipase activity was tested using plates containing 1% tributyrin.

Growth on polycyclic aromatic hydrocarbons (PAHs) and metabolites of PAH degradation were tested by adding salicylate, phthalate, naphthalene, acenaphthene, anthracene, phenanthrene, fluorene, pyrene, or fluoranthene (2 g·L–1 final concentration) to triplicate 5-mL tubes of either MM2 medium [28] or MSB medium. Tubes were incubated at 32.5°C with shaking at 225 rpm for up to 7 days and examined for turbidity. Uninoculated tubes containing growth substrates were tested as negative controls. Growth on benzene, toluene, ethylbenzene, or a mixture of *o*-, *m*-, and *p*-xylenes (BTEX compounds) was tested by adding each compound(s) to a final concentration of 100 mg·L–1 in triplicate, glass, screw-top tubes of 5-mL MM2 medium containing each strain and incubating for 5 days at 30°C with shaking. Growth was measured spectrophotometrically as described above. Inoculated

MM2 plates without a carbon source were also incubated at room temperature in tightly sealed metal cannisters containing atmospheres with individual BTEX compounds (0.5 mL on a piece of filter paper). Growth was determined by visual examination after three weeks.

The capability of each strain to metabolize a range of carbon sources was tested using Biolog GN2 microplates (Hayward, CA, USA) in duplicate using inocula from freshly grown R2A agar plates (Difco) suspended in GN/GP Inoculating Fluid as described by the manufacturer. Positive results were determined by visual examination and comparison to the water control after 24 and 48 hours of incubation at 30°C. Fatty acid profiles of the fours strains were provided by Microbial ID, Inc. (Newark, DE, USA) using the MIDI Sherlock® Microbial Identification System with cultures prepared on NA plates. Analysis of respiratory quinones and polar lipids was carried out by the Identification Service, DSMZ, Braunschweig, Germany [55–57].

Pulsed-field gel electrophoresis (PFGE).

Ten *Acidovorax* isolates from the Salisbury site (named P1 through P10) were screened for genomic variability by PFGE prior to selection and characterization. Embedded agarose plugs of whole cells grown in NB were prepared according to the CHEF-DR® III Pulsed Field Electrophoresis Systems Instruction Manual and Applications Guide, 2nd revision. Plugs containing DNA from each strain were digested with restriction enzyme SpeI overnight at 37°C (New England Biolabs, Ipswich, MA, USA) before being run on the CHEF- DR® III Pulsed Field Electrophoresis System under the following conditions: 1% agarose gel in 0.5×TBE at 14°C, 20 h run, 6 V/cm, switch times 4s-30s, and a 120° angle. The gel was stained post-run in ethidium bromide and the digest patterns examined for differences. One isolate from each of two observed banding patterns, designated P3 and P4, were selected for further analyses.

Embedded agarose plugs for determination of the presence of plasmids by PFGE were prepared and treated with S1 nuclease [3]. The running conditions of the gel were: 1% agarose gel in 0.5× TBE at 14°C, 18 h run, 6 V/cm, switch times 5s-25s, and a 120° angle. The gel was post- stained in ethidium bromide and plasmid sizes estimated based on comparison to λ-concatemer ladders.

Genomics.

High-quality DNA for genomic sequencing was obtained from overnight cultures grown in NB and extracted using a modification of the FastDNA Spin Kit for Soil protocol (MP Biomedicals, Santa Ana, CA, USA) in which cell suspensions were secured horizontally to a vortexer for 30 seconds at maximum speed for physical cell disruption. DNA quality and size were verified on a 1% agarose gel against size standards and quantified using a NanoDrop-3300 fluorospectrometer (Thermo Scientific, Waltham, MA, USA) with the Quant-iT Picogreen dsDNA Quantification Kit (Invitrogen, Grand Island, NY, USA). High-quality genome sequences of each of the four strains were obtained using the PacBio RSII system (Pacific Biosciences, Menlo Park, CA, USA) at the University of North Carolina High Throughput Sequencing Facility (UNC- HTSF). Data from two to four Single Molecule, Real-Time (SMRT) Cells were used for each strain to obtain the desired level of

coverage. The PacBio reads were *de novo* assembled with SMRT Analysis (version 2.2) using the Hierarchical Genome Assembly Process 3 (HGAP3). In HGAP3, the assembly step is performed using a combination of the Celera Assembler and a custom PacBiounitig consensus caller. The resulting contigs were polished using the Quiver algorithm [10]. The assemblies were visualized using the software Gephi [4] and we observed from the resulting graphs that the longest contigs had circular assemblies for each of the four strains.

To circularize the longest contig from each strain, the contigs were broken into two pieces at the selected origin of replication. To find the origin of replication, the contigs were run through BLAST [1, 64]; the BLAST result with the highest score for all four strains was *Alicycliphilus denitrificans* K601. The region with highest similarity to the replication control/initiator *dna*A gene for *A. denitrificans* K601 was selected to be the origin point for the four strains. BLAST was used to determine the location of best alignment of this gene with the contigs. Minimus2 [52] was used to perform the circularization and the two contig pieces were rejoined such that the start of the circularized contig was located close to the identified origin. Final polishing to reduce the number of insertion, deletion, and substitution errors was performed using Quiver. The resulting consensus sequences had an accuracy of at least 99.999%.

Gene calling and annotation of the genomes were independently performed using the Integrated Microbial Genomes and Microbiome Samples (IMG/MER) system of the U.S. Department of Energy Joint Genomics Institute (DOE/JGI) web portal [37] and the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP) version 4.1. Alignments of select genes and genomes and graphical representations of those results were created using Geneious v.9.1.3 (Biomatters Ltd., Newark NJ, USA). Sequences for rRNA and tRNA genes were identified through RNammer [31] and through the JGI portal. A neighbor-joining phylogenetic tree of 16S rRNA genes was created using ClustalX v.2.1 [54], excluding positions with gaps, and bootstrapping 1000 replicates. The same alignment was used to create a maximum parsimony tree using the PAUP algorithm within Geneious [14]. Putative genes involved in aromatic compound metabolism were identified through analyses using AromaDeg [16] and any homologs between genomes were confirmed using the "Genome Gene Best Homologs" tool within IMG/MER . Whole-genome sequence similarities were determined using both formula 2 of Genome Blast Distance Phylogeny version 2.0 (GBDP) [39] and average nucleotide identity (ANI) (http://enveomics.ce.gatech.edu/g-matrix/) using ANI-Matrix estimates [21]. Publicly available *Acidovorax* genomes for comparison were downloaded from the NCBI GenBank archives using the most recent assemblies as of February 2017.

Complete genomes are available at the JGI-IMG/MER website as well as in GenBank under the Bioproject PRJNA285679. GenBank nucleotide accession numbers for strains NA2, 229 NA3T, P3, and P4 are CP021359–60, CP021361, CP021362–65, and CP021366–70, respectively.

Results and Discussion

Phylogeny.

Each *Acidovorax* strain from this study contained four distinct rRNA operons. The 16S rRNA genes of the four strains (P3, P4, NA2, and NA3T) possessed between 99.6–100% similarity to one another. Strains P3 and P4 had identical 16S rRNA genes, which were distinct from the four identical copies found in strain NA3T. Interestingly, strain NA2 possessed four non- identical 16S rRNA genes with 99.6–99.9% identity to one another and variability observed at 6 of 1521 base positions (at positions 125, 126, 207, 208, 210, and 638; multiple signals were also observed in Sanger sequencing of PCR amplicons at those positions; data not shown). In strain NA2, one copy each of the 16S rRNA gene was identical to P3/P4 and NA3T, and the remaining two copies were 99.7% identical to one another, but distinct from all others.

Of the characterized *Acidovorax* species type strains, the most similar 16S rRNA gene sequences were from *A. radicis* sp. N35T (99.0–99.4% 16S rRNA gene similarity) and *A. defluvii* DSM 12644T (98.7–99.1%) (Figure 1). Similarly high 16S rRNA gene sequence similarity values were also observed to several uncharacterized environmental sequences (data not shown). High interspecies 16S rRNA gene identity (>97%) has previously been observed between described species of the *Acidovorax* genus [35]. The 16S rRNA gene sequence similarities of the four strains from this manuscript to a partial gene sequence from *Acidovorax delafieldii* P4–1 (Genbank accession DQ282184), the only other known phenanthrene-degrading *Acidovorax* isolate with a publicly available 16S rRNA gene sequence [44], ranged from 93.1–93.6% similarity over 828 bp. Efforts to obtain a culture of *A. delafieldii* P4–1 for further comparisons were unsuccessful. Comparisons of the RNA polymerase beta subunit (*rpoB*) gene additionally indicated that all four strains from this study comprised a cluster distinct from other characterized species with the closest relative being *Acidovorax* sp. JHL-3 (Appendix Figure A.1).

Characterization.

Isolated *Acidovorax* strains grew well on common, complex growth media including NA and R2A with colonies generally appearing within 24–48 hours. Colonies on NA plates grown for 48 hours for each of the strains were pale yellow to yellow in color (cream to pale yellow on R2A plates), circular and convex with entire margins, and up to 1.0 – 1.5 mm in diameter. Addition of indole to plates resulted in purple colonies for each strain, indicating the probable production of indigo from an active aromatic ring-hydroxylating dioxygenase that catalyzes the initial step of aerobic bacterial polycyclic aromatic hydrocarbon degradation [17].

Growth was observed for all four stains at temperatures tested between 15°C and 35°C, with optimal growth at 30 – 32.5°C for all strains. Weak growth was observed after two weeks of incubation at 4°C, 37°C, and 40°C in regions of the plates with the densest inoculum for all strains but P4, which did not grow at 40ºC. All strains grew at pH 6.0 – 8.0, with optimal growth at either 7.0 or 7.5. In addition, strain NA3T grew weakly at pH 5.5 and strains P3 and P4 at pH 8.5. All strains grew optimally without amended salt but could tolerate 1%

NaCl; no growth was observed after 24 hours in media supplemented with 2–5% NaCl. No growth was observed under anaerobic conditions on the media tested. All subsequent tests were generally conducted in media at a pH of 7.5, a temperature of 30 or 32.5°C, and with no amended salt.

Under scanning electron microscopy, all four strains were observed to be slightly-curved or straight rods without distinguishing external features (Figure 2). The cell sizes ranged from $0.9 - 2.1 \times 0.4 - 0.5$ μm (P3), $0.8 - 1.6 \times 0.4 - 0.6$ μm (P4), $0.8 - 1.5 \times 0.3 - 0.5$ μm (NA2), and $0.9 - 1.9 \times 0.4 - 0.5$ μm (NA3T), with average cell sizes (n = 100) of 1.4×0.5 μm (P3), 1.1×0.4 μm (P4), 1.1×0.4 μm (NA2), and 1.2×0.4 μm (NA3T). No flagella were evident in scanning electron micrographs despite the fact that *Acidovorax* isolates typically exhibit one to three polar flagella [62] and the genomic sequences of these four strains indicated the presence of genes for flagellar synthesis. To examine this apparent discrepancy, each strain was inoculated into motility agar stabs and examined after several days of growth. A feathering of growth radiating from each stab indicated motility for all strains. Single, polar flagella were subsequently detected on cells extracted from those stabs (Appendix Figure A. 2). The lack of flagella in electron micrographs was attributed to physical shearing during sample preparation and/or growth conditions within the preparative media.

The dominant fatty acids of the four strains were identified as summed feature 3 which 283 included $C_{16:1}\omega7c$ and $C_{16:1}\omega6c$ (47–48%), $C_{16:0}$ (26–29%), and $C_{18:1}\omega7c$ (12–16%) (Table 1). Based on the profiles of other *Acidovorax* strains the summed feature was likely to be predominantly $C_{16:1}\omega7c$. Minor fatty acids included $C_{10:0\ 3\text{-OH}}$, $C_{12:0}$, and $C_{14:0}$. In addition to 3-hydroxydecanoic acid ($C_{10:0\ 3\text{-OH}}$), cells from all strains also contained trace amounts of 3-hydroxyoctanoic acid ($C_{8:0}$ 3-OH), both of which are markers for the *Acidovorax* genus [62]. The fatty acid profiles of these strains were consistent with the genus. Major polar lipids in NA3T were phosphatidylethanolamine, diphosphatidylglycerol, phosphatidylglycerol, and two other lipids. The sole detected respiratory quinone in NA3T was Q8 (100%).

Each of the four strains was catalase negative and oxidase positive. All strains were negative for starch and gelatin hydrolysis, as well as DNAse, urease, protease, and extracellular cellulase activity. Conversely, all four strains demonstrated lipase activity and could reduce nitrate to nitrite. All strains used the low-molecular weight polycyclic aromatic hydrocarbon (PAH) phenanthrene as a sole source of carbon and energy, but failed to grow on the PAHs acenaphthene, naphthalene, anthracene, pyrene, fluoranthene, or fluorene. No strain grew on the PAH metabolites salicylate or phthalate. Prior work with strain NA3T indicated that it could mineralize carbon from a variety of 14C-labelled aromatic substrates including naphthalene, chrysene, benz[*a*]anthracene, benzo[*a*]pyrene, and phthalate [50]. No growth was observed for any strain on benzene, toluene, ethylbenzene, o-xylene, p-xylene, or m-xylene when added to liquid medium culture or to plates incubated in the presence of those compounds in the vapor phase.

Metabolism of 95 carbon sources was tested using Biolog GN2 microplates (Table 1). All four stains metabolized Tween 40, Tween 80, pyruvic acid methyl ester, succinic acid mono-methyl ester, β-hydroxybutyric acid, α-ketobutyric acid, α-ketoglutaric acid, α-ketovaleric

acid, DL-lactic acid, succinamic acid, and glycerol. Strains P3 and P4 could additionally metabolize D- galacturonic acid and D-glucuronic acid and strains NA2 and NA3T could utilize 2,3-butanediol. All strains but NA3T could metabolize L-proline. All other GN2 test substrates resulted in negative or inconsistent results.

Genomics.

The genomes of all four strains were determined and their general characteristics analyzed (Table 1). Circular chromosomes of the four strains averaged 4.07 ± 0.06 Mbp with chromosomal G+C contents of $64.1 \pm 0.1\%$. These chromosomes were considerably smaller than the genome of the most closely related described *Acidovorax* species, *A. radicis* N35T (5.5 Mbp) [34]. Each of the four strains in this study displayed similar chromosomal arrangements with highly conserved blocks of genes presumably reflecting homology (Appendix Fig. A.3). Strains NA2 and NA3T each possessed an ~10 kbp region that was absent in P3/P4, containing a small number of hypothetical genes encoding proteins either associated with the inner membrane or of unknown function (Appendix Fig. A.3; Region "a"). In strain NA3T, this region was flanked by a putative plasmid stabilization protein-encoding gene, and in NA2 with putative transposases, suggesting the possibility of chromosomal integration of a plasmid or horizontal gene transfer. Strains P3/P4 additionally possessed a genomic region of ~50 kbp that was absent in NA2/NA3T (Fig. A.3; Region "b"). While several genes in this region were of unknown function, putative genes encoding for a type I restriction modification system, a lactate/2-hydroxyacid dehydrogenase, a phosphate/phosphite/phosphonate ABC transporter, and heavy metal sensor and response proteins were present. Two separate putative integrase genes within this block suggest a possible history of horizontal gene transfer for this region as well.

One, zero, three, and four circular plasmids were present in PFGE analyses of the strains NA2, NA3T, P3 and P4, respectively, and their sequences were determined from the genomes of the strains (Table 1; Appendix Fig. A.4). During assembly, manual inspection of the HGAP3 assemblies indicated missing plasmids from NA2 and P3 that were indicated by PFGE. The assembly for NA2 initially consisted of a single chromosome, however it was observed that two stretches matched segments of the largest P4 plasmid. Together these two pieces were roughly the same size as the missing plasmid for NA2 as determined by PFGE and thus were extracted from the NA2 chromosome and combined to form plasmid pACNA2.1. For strain P3, HGAP3 assembly indicated 2 plasmids. However, the largest P3 plasmid was considerably larger than expected based on PFGE. This large plasmid likely consisted of two separate plasmids that had incorrectly merged during the assembly process. The misassembled plasmid was split at the likely merge position, and the resulting two plasmids were consistent with PFGE results.

The largest plasmids from strains P3 and P4, designated pACP3.1 (258 kbp) and pACP4.1 (269 kbp), respectively, possessed >99% sequence similarity (100% query coverage), with only a putative transposase gene within the shared region accounting for the majority of the sequence variation. The second largest plasmids from strains P3 and P4, designated pACP3.2 (157 kbp) and pACP4.2 (164 kbp), were likewise nearly identical (99% query coverage, 99% identity) over their shared region and additionally possessed high similarity

(99% identity) to sections of the *Acidovorax* sp. JS42 genome (Genbank accession CP000539; 27% query coverage) including a region encoding a heavy metal ATPase and a transposase in the Tn3 family. Additional plasmids from strains P3 and P4, pACP3.3 (73 kbp) and pACP4.4 (80 kbp), possessed 99% similarity to one another over the shared region (100% query coverage) and to plasmids from *Acidovorax* sp. JS42 (plasmid pAOVO01; 96% query coverage, 98% identity) and *Alicycliphilus denitrificans* K601 (plasmid pALIDE201; 92% query coverage, 99% identity). These plasmids contained putative genes encoding for (among other things) conjugal transfer proteins, a histidine kinase, as well as a DNA methyltransferase, polymerase, and primase. Plasmid pACP4.3 (115 kbp) was unique to strain P4 among these isolates and possessed only small regions of similarity to multiple elements in the *Acidovorax citrulli* AAC00–1 genome (Genbank accession CP000512; 2% query coverage, 73% identity). As the genomes of strains P3 and P4 were nearly identical, we presume that the fourth plasmid of P4 absent in P3 resulted in the differing PFGE-derived restriction nuclease patterns that distinguished the strains during screening of recovered isolates.

Plasmid pACNA2.1 of strain NA2 (71 kbp), despite being of a comparable size to plasmids pACP3.3 (73 kbp) and pACP4.4 (79 kbp), possessed 99.9% identity to a region spanning 26% of the larger plasmids pACP3.1 and pACP4.1. The largest region of similarity to known sequences in any of the plasmids pACP3.3, pACP4.4, or pACNA2.1 (representing 10% coverage of pACP4.1 and pACP3.1, 36% coverage of pACNA2.1) was to disparate sections of the *Acidovorax avenae* subsp. avenae ATCC 19860T genome (Genbank accession CP002521).

Differentiation of isolated strains from described *Acidovorax* species

As of the time of this work and not counting the strains described in this study, a total of 49 *Acidovorax* genomes representing much of the environmental and phenotypic diversity of the genus were publicly available. To distinguish strains NA2, NA3T, P3, and P4 from other species and strains in the genus, two different genome comparison methods with correlation to DNA-DNA hybridization (DDH) were performed: genome blast distance phylogeny (GBDP) and average nucleotide identity (ANI) (Table 2; Appendix Tables A.1 and A.2). The four strains described here possessed   97.0% ANI to one another, but no higher than 88.1% ANI to any other publicly available *Acidovorax* genome (an ANI value of >95% has been correlated with a DDH value >70% [21]). GBDP using the most conservative estimate predicted DDH values for the four strains from this study to be between 72.4 and 98.3% (with a corresponding 82–98% probability of these four strains comprising the same species); all were values above the 70% DDH threshold for establishing species similarity. Conversely, none of the four strains from this study possessed greater than 34.2% predicted DDH values to any other *Acidovorax* strain with a sequenced genome, including the species with the highest 16S rRNA gene similarity, *A. radicis* N35T (~26% DDH). Among an additional 30 *Acidovorax* strains with publicly available genomes but no species designation, none displayed a DDH value greater than 35.5%. These results indicate that when compared to *Acidovorax* species with sequenced genomes these four strains collectively comprise a new species.

Of described, non-phytopathogenic species of *Acidovorax*, only *A. defluvii* and *A. facilis* did not yet have a strain with a publicly available genome and thus could not be compared to these strains using ANI or GBDP methods. However, the type strains of those genera (DSM 12644T and DSM 649T, respectively) can be differentiated from strain NA3 phenotypically through differences in the utilization of various carbon sources using a standard Biolog GN2 microplate, as well as their catalase reactions (Table 3). Additionally, members of the *A. defluvii* species have been reported as possessing polyhydroxybutyrate granules [46], a phenotype not yet observed in these strains.

Genes putatively involved in aromatic and phenanthrene metabolism

The JGI-assembled and annotated genomes of each of the four strains were subjected to analysis using AromaDeg [15], a predictive algorithm for proteins involved in the metabolism of aromatic compounds. Between 8 and 11 putative oxygenases and dioxygenases were detected in each strain (Table 4). Based on high similarity to characterized genes and gene products in other organisms, several were predicted to encode proteins involved in PAH degradation.

The genes likely responsible for the transformation of phenanthrene, including those encoding for the ring-hydroxylating dioxygenase (RHD) involved in the initial step in the aerobic bacterial metabolism of PAHs, were compared to genes from phenanthrene-degrading isolates from other genera within the *Betaproteobacteria* (Figure 3). The genes encoding enzymes for the transformation of phenanthrene to phthalate, including the four gene components of the initial RHD, were present in a single region with gene organization highly similar to homologous regions in the genomes of the phenanthrene-degrading bacteria *Delftia* sp. Cs1–4 and *Burkholderia* sp. Ch1–1, and to a GenBank sequence derived from the phenanthrene-degrading bacterium *Alcaligenes faecalis* strain AFK2. Only single copies of the four separate genes whose products comprise the subunits of the RHD holoenzyme (*phnAabcd*) were identified for each of the four genomes. For the four *Acidovorax* strains described in this study, the genes and organization of this cluster were nearly identical except for the putative *phnK* gene in strain P3, which was annotated as split into two adjacent genes (data not shown); with this caveat strain NA3T is presented as representative of all four strains from this study (Figure 3). The high similarity of the nucleotide sequence and gene organization of phenanthrene metabolic genes in NA3T to those in other phenanthrene-degrading strains was previously indicated by comparisons of partial *Acidovorax* sp. NA3 T gene sequences to the *Alcaligenes faecalis* AFK2 sequence [50] and more recently by comparisons of NA3T sequences to the complete genomes of *Burkholderia* sp. Ch1–1 and *Delftia* sp. Cs1–4 [24], and is confirmed here with this genomic sequencing effort. Interestingly, this gene cluster was also found in the recently sequenced genomes of *Massilia* sp. WG5 [36] and *Massilia* sp. WF1 [60]. While most representations of this gene cluster have been found on chromosomes, including in the *Acidovorax* strains presented here, the *phn* genes were present on a plasmid in both *Massilia* sp. WG5 (Genbank accession NZ_CP012641) and *Alcaligenes faecalis* AFK2 [29].

The highly conserved *phn* gene cluster present in various members of the genera *Acidovorax*, *Massilia, Delftia*, *Burkholderia*, and *Alcaligenes* strains has been linked to

phenanthrene metabolism through both transcription-based assays [50] and proteomic experiments [24], and has been referred to as the *phn* island or *phn*$_{AFK2}$-type genes (after the *Alcaligenes faecalis* AFK2 Genbank entry first describing the genotype). As has been observed for other organisms possessing this cluster, the genomic region comprising the *phn* genes in these *Acidovorax* strains possessed a significantly lower G+C content than the chromosomal average of the organisms (50% for the *phn* cluster compared to 64% for the complete *Acidovorax* genomes). However, in contrast to *Delftia* sp. Cs1–4 in which the *phn* genes were contained within a large 232 kb genomic island, analyses of the *Acidovorax* strains presented here using the same methodology found no evidence for such a large chromosomal element encompassing the *phn* region. However, a much smaller potential genomic island (~10 kb) in strain NA3T comprising a region between a putative transposase upstream of the *phnAb* gene through the glutathione s-transferase (gst) gene in the *phn* operon was indicated (data not shown).

Unlike *phn* clusters of described *Burkholderia*, *Delftia*, and *Alcaligenes* strains, but similar to the *Massilia* gene clusters, the *phn* genes in these *Acidovorax* strains were uninterrupted by genes encoding for mobile elements (Figure 3). However, the presence of genes related to genetic mobility flanking the *phn* cluster in these and other organisms, in conjunction with the dissimilar G+C content compared to the rest of the genome, supports horizontal gene transfer as an explanation for the presence of extraordinarily highly-conserved degradation genes amongst these phenanthrene-degrading members of the *Betaproteobacteria*. Interestingly, this *phn* gene region from *Acidovorax* and other phenanthrene-degrading Betaproteobacterial strains does not appear to be present in the recently published genome of the phenanthrene-degrading *Burkholderia* sp. HB-1 [42] or in *Burkholderia* sp. RP007 [33], indicating multiple genetic determinants of phenanthrene metabolism within that genus.

The genes of the phenanthrene degradation pathway likely responsible for the transformation of o-phthalate to protocatechuate in *Acidovorax* strain NA3T were identified in a cluster approximately 46 kb downstream of the *phn* genes on the chromosome, and similar to the *phn* region were nearly identical in sequence and arrangement among the four *Acidovorax* strains (data not shown). Based on gene similarity to other organisms and predicted phenotypes of those genes, the transformation of o-phthalate likely proceeds through 4,5-dihydroxyphthalate using a two-component dioxygenase enzyme similar to that found in *Burkholderia cepacia* [5, 8]. The predicted phthalate 4,5-dioxygenase gene product (*ophA2*) was most similar in protein sequence to predicted products in the Comamonadaceae bacteria *Variovorax* sp. WDL1 (Genbank accession KWT97458; 75% identity) and *Hydrogenophaga* sp. Root209 (accession WP_056272285; 74% identity). Similar to *Variovorax* strain WDL1, the genes in these four *Acidovorax* strains encoding the putative phthalate dioxygense reductase (*ophA1*), phthalate 4,5-dioxygenase (*ophA2*), 4,5-dihydro-4,5-dihydroxyphthalate dehydrogenase (*ophB*), and 4,5-dihydroxyphthalate decarboxylase (*ophC*) were present in a single cluster with a MarR-type transcriptional regulator separating the divergently transcribed reductase/dehydrogenase and decarboxylase/dioxygenase genes. This gene arrangement is distinct from that described for *Delftia* sp. Cs1–4 and *Comamonas testosteroni* KF-1 [24]. No obvious homologs to a *Burkholderia*-type phthalate permease (*ophD*) [9] or an ABC-type phthalate transporter (*ophFGH*) [7] were found in the genomes any of the *Acidovorax* strains. The lack of a phthalate transporter

may explain why none of these described *Acidovorax* strains were capable of growth on that compound as a sole source of carbon and energy despite presumably being able to transform it as part of phenanthrene metabolism. Strains P3 and P4 additionally contained a second, genetically distinct copy of a putative phthalate 4,5- dioxygenase (*ophA2*) located on plasmids pACP3.2 and pACP4.2, respectively (100% identical to one another; ~75.5% identical to the chromosomal *ophA2* genes). However, no other genes predicted to be associated with phthalate metabolism were found in genetic proximity to those isolated copies, and their relevance to the transformation of phthalate or role in phenanthrene metabolism is unknown.

In these *Acidovorax* strains, the transformation of protocatechuate to central metabolism intermediates likely proceeds through a two-component protocatechuate 4,5-dioxygenase (genes *pmdA* and *pmdB*), similarly to that described in *Delftia* sp. Cs1–4 and *Comamonas testosterone* BR6020 [43]. In this pathway, conversion of the semialdehyde product to pyruvate and oxaloacetate occurs through the transformation products of a semialdehyde dehydrogenase (*pmdC*), dicarboxylic acid hydrolase (*pmdD*), hydratase (*pmdE*), and aldolase (*pmdF*). Putative genes encoding these enzymes were located on a single cluster in a region approximately 6 kb upstream of the *phn* gene cluster on the chromosome of *Acidovorax* strain NA3T and contained a divergently transcribed lysR-type regulator, as well as genes with similarities to an isomerase from the 2-methyl citrate pathway, a dehydrogenase with wide representation in bacterial genomes, and two hypothetical genes (data not shown). The entire *pmd* gene region was flanked by putative transposable elements; however, neither the *oph* nor *pmd* gene clusters were predicted to be part of a genomic island and the G+C% of the regions containing the relevant genes were 58.8% and 61.0%, respectively, significantly higher than the 50% of the *phn* gene region and closer to the 64% G+C of the *Acidovorax* chromosomes. Similar to the second copy of *ophA2* on plasmids pACP3.2 and pACP4.2, strains P3 and P4 also harbored a second copy of the *pmdA* and *pmdB* genes on those plasmids with 100% identity to annotated protocatechuate 3,4-dioxygenase subunits from *Acidovorax* sp. CCH12-A4. In contrast to the plasmid-borne *ophA2* genes in P3 and P4, which were not part of an obvious operon, these duplicate copies were also flanked by genes associated with a protocatechuate 4,5-cleavage pathway.

Based on phlyogenetic, phenotypic, and genomic characteristics, we propose that the four strains described in this manuscript constitute a new species within the *Acidovorax* genus. The formal proposal of the new species "*Acidovorax carolinensis* sp. nov." is given in Table 5, which represents information taken from Digital Protologue Taxonumber TA00473.

DIGITAL PROTOLOGUE

TAXONUMBER TA00473

SPECIES NAME Acidovorax phenanthrenivorans

GENUS NAME Acidovorax

SPECIFIC EPITHET phenanthrenivorans

SPECIES STATUS sp. nov.

SPECIES ETYMOLOGY Acidovorax phenanthrenivorans (phen.an´thre.ni.vo´rans. N.L. n. phenanthrenum phenanthrene; L. v. vorare to devour; L. part. adj. vorans devouring, digesting; N.L. part. adj. phenanthrenivorans digesting phenanthrene).

SUBMITTER

E-MAIL OF THE SUBMITTER david.r.singleton@duke.edu

HAS THE TAXON BEEN SUBJECTED TO EMENDATION? NO

DESIGNATION OF THE TYPE STRAIN NA3

STRAIN COLLECTION NUMBERS =LMG 30136, =DSM 105008

16S rRNA GENE ACCESSION NUMBER EU910093.1

GENOME ACCESSION NUMBER [RefSeq] NZ_CP021361.1

GENOME STATUS complete

GENOME SIZE 4122 kbp

GC mol % 64.24

COUNTRY OF ORIGIN USA

REGION OF ORIGIN Charlotte, North Carolina

DATE OF ISOLATION UNKNOWN (< yyyy) 2005

SOURCE OF ISOLATION Contaminated soil

SAMPLING DATE 2002–01-01

GEOGRAPHIC LOCATION Charlotte, NC former manufactured gas plant site

TEMPERATURE OF THE SAMPLE [In Celsius degrees] 20

pH OF THE SAMPLE 7.5

NUMBER OF STRAINS IN STUDY 4

SOURCE OF ISOLATION OF NON-TYPE STRAINS Contaminated soil

GROWTH MEDIUM, INCUBATION Nutrient agar/R2A

CONDITIONS [Temperature, pH, and further 30–32.5 degrees celcius

information] USED FOR STANDARD CULTIVATION pH 7–7.5

CONDITIONS OF PRESERVATION 15% glycerol; -80C

GRAM STAIN NEGATIVE

CELL SHAPE rod

CELL SIZE (length or diameter) 1.2

MOTILITY motile

IF MOTILE flagellar

IF FLAGELLATED single, polar

SPORULATION (resting cells) none

COLONY MORPHOLOGY Cream color to pale yellow,circular, 1.0–1.5mm diameter.

TEMPERATURE RANGE 4–40

LOWEST TEMPERATURE FOR GROWTH 4, nutrient agar plates HIGHEST **TEMPERATURE FOR GROWTH** 40, nutrient agar plates TEMPERATURE OPTIMUM 30–32.5

LOWEST pH FOR GROWTH 5.5, nutrient broth

HIGHEST pH FOR GROWTH 8.0, nutrient broth

pH OPTIMUM 7.0–7.5

pH CATEGORY neutrophile

LOWEST NaCl CONCENTRATION FOR GROWTH 0, nutrient broth

HIGHEST NaCI CONCENTRATION FOR GROWTH 1%, nutrient broth

SALINITY CATEGORY nonhalophile (NaCl inhibitory at <1 % NaCl)

RELATIONSHIP TO O2 aerobe

O2 CONDITIONS FOR STRAIN TESTING aerobiosis; anaerobiosis

CARBON SOURCE USED [class of compounds] organic acids, aromatic compounds

CARBON SOURCE USED [specific compounds] phenanthrene

CARBON SOURCE NOT USED [specific compounds] acenaphthene, anthracene, benzene, ethylbenzene, fluoranthene, fluorene, naphthalene, phthalate, pyrene, salicylate, toluene, xylene

Positive tests with BIOLOG Tween 40, Tween 80, Methyl pyruvate, succinic acid mono-methyl ester, β- hydroxybutyric acid, α-ketobutyric acid, α-ketoglutaric acid, α-ketovaleric acid, DL-lactic acid, succinamic acid, glycerol

Negative tests with BIOLOG α-Cyclodextrin, Dextrin, Glycogen, N-Acetyl-D-galactosamine, N-Acetyl-D- glucosamine, Adonitol, L-Arabinose, D-Arabitol, D-Cellobiose, i-Erythritol, D- Fructose, L-Fucose, D-Galactose, Gentiobiose, D-Glucose, m-Inositol, α-D- Lactose, Lactulose, Maltose, D-Mannitol, D-Mannose, D-Melibiose, β-MethylD- Glucoside, D-Psicose, D-Raffinose, L-Rhamnose, D-Sorbitol, Sucrose, D-Trehalose, Turanose, Xylitol, Acetic Acid, Cis-Aconitic Acid, Citric Acid, Formic Acid, D-Galactonic Acid Lactone, D-Glucosaminic Acid, D-Glucuronic Acid, α-Hydroxy- Butyric Acid, γ-Hydroxy Butyric Acid, p-Hydroxy Phenylacetic Acid, Itaconic Acid, α-Keto Valeric Acid, Malonic Acid, Propionic Acid, Quinic Acid, D-Saccharic Acid, Sebacic Acid, Bromo-Succinic Acid, Glucuronamide, L-Alaninamide, D-Alanine, L-Alanine, L-Alanylglycine, L-Asparagine, L-Aspartic Acid, L-Glutamic Acid, Glycyl-L-Aspartic Acid, Glycyl-L-Glutamic Acid, L-Histidine, Hydroxy-L-Proline, L-Leucine, L-Ornithine, L-Phenylalanine, L-Pyroglutamic Acid, D-Serine, L-Serine, L-Threonine, D,L-Carnitine, γ-Amino Butyric Acid, Urocanic Acid, Inosine, Uridine, Thymidine, Phenyethylamine, Putrescine, 2-Aminoethanol, D,L-α-Glycerol Phosphate, Gluose-1-Phosphate, Glucose-6-Phosphate

Variable tests with BIOLOG D-galacturonic acid, D-glucuronic acid, 2,3-butanediol, L-proline

NITROGEN SOURCE NH4

TERMINAL ELECTRON ACCEPTOR oxygen

ENERGY METABOLISM chemoorganotroph

OXIDASE positive

CATALASE negative

POSITIVE TESTS lipase, nitrate reduction

NEGATIVE TESTS gelatin hydrolysis, cellulase, DNase, skim milk protease, starch hydrolysis, urease

QUINONE TYPE R8

MAJOR FATTY ACIDS C16:1ω7c/C16:1ω6c, C16:0, C18:1ω7c

PHOSPOLIPID PATTERN OR DIAGNOSTIC PHOSPOLIPID
phosphatidylethanolamine, diphosphatidylglycerol, phosphatidylglycerol

BIOSAFETY LEVEL 1

HABITAT terrestrial biome

BIOTIC RELATIONSHIP free-living

KNOWN PATHOGENICITY none

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

[1]. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, Basic local alignment search tool, J. Mol. Biol, 215 (1990) 403–410. [PubMed: 2231712]

[2]. Atlas RM, Handbook of Microbiological Media, Third edition ed., CRC Press, Boca Raton FL, 2004.

[3]. Basta T, Keck A, Klein J, Stolz A, Detection and characterization of conjugative degradative plasmids in xenobiotic-degrading *Sphingomonas strains*, J. Bacteriol, 186 (2004) 3862–3872. [PubMed: 15175300]

[4]. Bastian M, Heymann S, Jacomy M, Gephi: An Open Source Software for Exploring and Manipulating Networks, 2009.

[5]. Batie C, LaHaie E, Ballou D, Purification and characterization of phthalate oxygenase and phthalate oxygenase reductase from *Pseudomonas cepacia*, J. Biol. Chem, 262 (1987) 1510–1518. [PubMed: 3805038]

[6]. Bogan BW, Lahner LM, Sullivan WR, Paterek JR, Degradation of straight-chain aliphatic and high-molecular-weight polycyclic aromatic hydrocarbons by a strain of *Mycobacterium austroafricanum*, J. Appl. Microbiol, 94 (2003) 230–239. [PubMed: 12534814]

[7]. Chang HK, Dennis JJ, Zylstra GJ, Involvement of two transport systems and a specific porin in the uptake of phthalate by *Burkholderia* spp, J. Bacteriol, 191 (2009) 4671–4673. [PubMed: 19429613]

[8]. Chang HK, Zylstra GJ, Novel organization of the genes for phthalate degradation from *Burkholderia cepacia* DBO1, J. Bacteriol, 180 (1998) 6529–6537. [PubMed: 9851995]

[9]. Chang HK, Zylstra GJ, Characterization of the phthalate permease OphD from *Burkholderia cepacia* ATCC 17616, J. Bacteriol, 181 (1999) 6197–6199. [PubMed: 10498738]

[10]. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J, Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data, Nat. Meth, 10 (2013) 563–569.

[11]. Choi J-H, Kim M-S, Roh SW, Bae J-W, *Acidovorax soli* sp. nov., isolated from landfill soil, Int. J. Syst. Evol. Microbiol, 60 (2010) 2715–2718. [PubMed: 20061503]

[12]. Christensen WB, Urea decomposition as a means of differentiating *Proteus* and *Paracolon* cultures from each other and from *Salmonella* and *Shigella* types, J. Bacteriol, 52 (1946) 461–466. [PubMed: 16561200]

[13]. Corteselli EM, Aitken MD, Singleton DR, *Rugosibacter aromaticivorans* gen. nov., sp. nov., a novel bacterium within the family *Rhodocyclaceae* isolated from contaminated soil, capable of degrading aromatic compounds, Int. J. Syst. Evol. Microbiol, 67 (2016) 311–318.

[14]. Cummings MP, Hancock JM, Zvelebil MJ, PAUP* (phylogenetic analysis using parsimony (and other methods)), in: Dictionary of Bioinformatics and Computational Biology, John Wiley & Sons, Ltd, 2004.

[15]. dela Cruz TE, Torres JMO, Gelatin hydrolysis test protocol, in, American Society for Microbiology, Washington, DC, 2012.

[16]. Duarte M, Jauregui R, Vilchez-Vargas R, Junca H, Pieper DH, AromaDeg, a novel database for phylogenomics of aerobic bacterial degradation of aromatics, Database (Oxford), 2014 (2014) bau118.

[17]. Eaton RW, Chapman PJ, Formation of indigo and related compounds from indolecarboxylic acids by aromatic acid-degrading bacteria: chromogenic reactions for cloning genes encoding dioxygenases that act on aromatic acids, J. Bacteriol, 177 (1995) 6983–6988. [PubMed: 7592495]

[18]. Eriksson M, Sodersten E, Yu Z, Dalhammar G, Mohn WW, Degradation of polycyclic aromatic hydrocarbons at low temperature under aerobic and nitrate-reducing conditions in enrichment cultures from northern soils, Appl. Environ. Microbiol, 69 (2003) 275–284. [PubMed: 12514005]

[19]. Gardan L, Dauga C, Prior P, Gillis M, Saddler GS, *Acidovorax anthurii* sp. nov., a new phytopathogenic bacterium which causes bacterial leaf-spot of anthurium, Int. J. Syst. Evol Microbiol, 50 (2000) 235–246. [PubMed: 10826809]

[20]. Gardan L, Stead DE, Dauga C, Gillis M, *Acidovorax valerianellae* sp. nov., a novel pathogen of lamb's lettuce [Valerianella locusta (L.) Laterr.], Int. J. Syst. Evol. Microbiol, 53 (2003) 795–800. [PubMed: 12807202]

[21]. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM, DNA-DNA hybridization values and their relationship to whole-genome sequence similarities, Int. J. Syst. Evol. Microbiol, 57 (2007) 81–91. [PubMed: 17220447]

[22]. Haigler BE, Wallace WH, Spain JC, Biodegradation of 2-nitrotoluene by Pseudomonas sp. strain JS42, Appl. Environ. Microbiol, 60 (1994) 3466–3469. [PubMed: 7944378]

[23]. Heylen K, Lebbe L, De Vos P, *Acidovorax caeni* sp. nov., a denitrifying species with genetically diverse isolates from activated sludge, Int. J. Syst. Evol. Microbiol, 58 (2008) 73–77. [PubMed: 18175686]

[24]. Hickey WJ, Chen S, Zhao J, The phn island: a new genomic island encoding catabolism of polynuclear aromatic hydrocarbons, Front. Microbiol, 3 (2012) 125. [PubMed: 22493593]

[25]. Jones MD, Crandell DW, Singleton DR, Aitken MD, Stable-isotope probing of the polycyclic aromatic hydrocarbon-degrading bacterial guild in a contaminated soil, Environ. Microbiol, 13 (2011) 2623–2632. [PubMed: 21564459]

[26]. Jurelevicius D, Alvarez VM, Peixoto R, Rosado AS, Seldin L, Bacterial polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenases (PAH-RHD) encoding genes in different soils from King George Bay, Antarctic Peninsula, Appl. Soil Ecol, 55 (2012) 1–9.

[27]. Kimbara K, Hashimoto T, Fukuda M, Koana T, Takagi M, Oishi M, Yano K, Isolation and characterization of a mixed culture that degrades polychlorinated biphenyls Agric. Biol. Chem, 52 (1988) 2885–2891.

[28]. Kiyohara H, Nagao K, Yana K, Rapid screen for bacteria degrading water-insoluble, solid hydrocarbons on agar plates, Appl. Environ. Microbiol, 43 (1982) 454–457. [PubMed: 16345951]

[29]. Kiyohara H, Takizawa N, Date H, Torigoe S, Yano K, Characterization of a phenanthrene degradation plasmid from *Alcaligenes faecalis* AFK2, J. Ferment. Bioeng, 69 (1990) 54–56.

[30]. Kovacs N, Identification of *Pseudomonas pyocyanea* by the oxidase reaction, Nature, 178 (1956) 703–703. [PubMed: 13369512]

[31]. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW, RNAmmer: consistent and rapid annotation of ribosomal RNA genes, Nucleic Acids Res, 35 (2007) 3100– 3108. [PubMed: 17452365]

[32]. Lal A, Cheeptham N, Starch agar protocol, in: Protocols, American Society for Microbiology, Washington, DC, 2012.

[33]. Laurie AD, Lloyd-Jones G, The *phn* genes of *Burkholderia* sp. strain RP007 constitute a divergent gene cluster for polycyclic aromatic hydrocarbon catabolism, J. Bacteriol, 181 (1999) 585 531–540. [PubMed: 9882674]

[34]. Li D, Rothballer M, Engel M, Hoser J, Schmidt T, Kuttler C, Schmid M, Schloter M, Hartmann A, Phenotypic variation in *Acidovorax radicis* N35 influences plant growth promotion, FEMS Microbiol. Ecol, 79 (2012) 751–762. [PubMed: 22107346]

[35]. Li D, Rothballer M, Schmid M, Esperschütz J, Hartmann A, *Acidovorax radicis* sp. nov., a wheat-root-colonizing bacterium, Int. J. Syst. Evol. Microbiol, 61 (2011) 2589–2594. [PubMed: 21131505]

[36]. Lou J, Gu H, Wang H, An Q, Xu J, Complete genome sequence of *Massilia* sp. WG5, an efficient phenanthrene-degrading bacterium from soil, J. Biotechnol, 218 (2016) 49–50. [PubMed: 26656222]

[37]. Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Pillay M, Ratner A, Huang J, Woyke T, Huntemann M, Anderson I, Billis K, Varghese N, Mavromatis K, Pati A, Ivanova NN, Kyrpides NC, IMG 4 version of the integrated microbial genomes comparative analysis system, Nucleic Acids Res, 42 (2014) D560–D567. [PubMed: 24165883]

[38]. Martin F, Torelli S, Le Paslier D, Barbance A, Martin-Laurent F, Bru D, Geremia R, Blake G, Jouanneau Y, Betaproteobacteria dominance and diversity shifts in the bacterial community of a PAH-contaminated soil exposed to phenanthrene, Environ. Pollut, 162 (2012) 600 345–353.

[39]. Meier-Kolthoff JP, Auch AF, Klenk HP, Goker M, Genome sequence-based species delimitation with confidence intervals and improved distance functions, BMC Bioinformatics, 14 (2013) 60. [PubMed: 23432962]

[40]. Meyer S, Moser R, Neef A, Stahl U, Kämpfer P, Differential detection of key enzymes of polyaromatic-hydrocarbon-degrading bacteria using PCR and gene probes, Microbiology, 145 (Pt 606 7) (1999) 1731–1741. [PubMed: 10439412]

[41]. Mizzouri NS, Shaaban MG, Kinetic and hydrodynamic assessment of an aerobic purification system for petroleum refinery wastewater treatment in a continuous regime, Int. Biodeterior. Biodegrad, 83 (2013) 1–9.

[42]. Ohtsubo Y, Moriya A, Kato H, Ogawa N, Nagata Y, Tsuda M, Complete genome sequence of a phenanthrene degrader, *Burkholderia* sp. HB-1 (NBRC 110738), Genome Announcements, 3 (2015).

[43]. Providenti MA, Mampel J, MacSween S, Cook AM, Wyndham RC, *Comamonas testosteroni* BR6020 possesses a single genetic locus for extradiol cleavage of protocatechuate, Microbiology, 147 (2001) 2157–2167. [PubMed: 11495993]

[44]. Samanta SK, Chakraborti AK, Jain RK, Degradation of phenanthrene by different bacteria: evidence for novel transformation sequences involving the formation of 1-naphthol, Appl. Microbiol. Biotechnol, 53 (1999) 98–107. [PubMed: 10645629]

[45]. Schatz A, Bovell C, Growth and hydrogenase activity of a new bacterium, *Hydrogenomonas facilis*, J. Bacteriol, 63 (1952) 87–98. [PubMed: 14927551]

[46]. Schulze R, Spring S, Amann R, Huber I, Ludwig W, Schleifer K-H, Kämpfer P, Genotypic diversity of *Acidovorax* strains isolated from activated sludge and description of *Acidovorax defluvii* sp. nov, Syst. Appl. Microbiol, 22 (1999) 205–214. [PubMed: 10390871]

[47]. Shuttleworth KL, Cerniglia CE, Bacterial degradation of low concentrations of phenanthrene and inhibition by naphthalene, Microb. Ecol, 31 (1996) 305–317. [PubMed: 8661535]

[48]. Singleton DR, Hunt M, Powell SN, Frontera-Suau R, Aitken MD, Stable-isotope probing with multiple growth substrates to determine substrate specificity of uncultivated bacteria, J. Microbiol. Methods, 69 (2007) 180–187. [PubMed: 17267058]

[49]. Singleton DR, Powell SN, Sangaiah R, Gold A, Ball LM, Aitken MD, Stable-isotope probing of bacteria capable of degrading salicylate, naphthalene, or phenanthrene in a bioreactor treating contaminated soil, Appl. Environ. Microbiol, 71 (2005) 1202–1209. [PubMed: 15746319]

[50]. Singleton DR, Ramirez LG, Aitken MD, Characterization of a polycyclic aromatic hydrocarbon degradation gene cluster in a phenanthrene-degrading *Acidovorax* strain, Appl. Environ. Microbiol, 75 (2009) 2613–2620. [PubMed: 19270134]

[51]. Singleton DR, Richardson SD, Aitken MD, Pyrosequence analysis of bacterial communities in aerobic bioreactors treating polycyclic aromatic hydrocarbon-contamined soil Biodegradation, 22 (2011) 1061–1073. [PubMed: 21369833]

[52]. Sommer DD, Delcher AL, Salzberg SL, Pop M, Minimus: a fast, lightweight genome assembler, BMC Bioinformatics, 8 (2007) 64.

[53]. Stanier RY, Palleroni NJ, Doudoroff M, The aerobic Pseudomonads: a taxonomic study, Journal of General Microbiology, 43 (1966) 159–271. [PubMed: 5963505]

[54]. Thompson JD, Gibson TJ, Higgins DG, Multiple sequence alignment using ClustalW and ClustalX, Curr. Protoc. Bioinformatics, Chapter 2 (2002) Unit 2 3.

[55]. Tindall BJ, A comparative study of the lipid composition of *Halobacterium saccharovorum* from various sources, Syst. Appl. Microbiol, 13 (1990) 128–130.

[56]. Tindall BJ, Lipid composition of *Halobacterium lacusprofundi*, FEMS Microbiol. Lett, 66 (1990) 199–202.

[57]. Tindall BJ, Sikorski J, Smibert RM, Kreig NR, Phenotypic characterization and the principles of comparative systematics., in: Reddy CA, Beveridge TJ, Breznak JA, Marzluf G, Schmidt TM, Snyder LR (Eds.) Methods for General and Molecular Microbiology, ASM Press, Washington DC, USA, 2007, pp. 330–393.

[58]. Tischer K, Kleinsteuber S, Schleinitz KM, Fetzer I, Spott O, Stange F, Lohse U, Franz J, Neumann F, Gerling S, Schmidt C, Hasselwander E, Harms H, Wendeberg A, Microbial communities along biogeochemical gradients in a hydrocarbon-contaminated aquifer, Environ. Microbiol, 15 (2013) 2603–2615. [PubMed: 23809669]

[59]. Vaneechoutte M, Janssens M, Avesani V, Delmée M, Deschaght P, Description of *Acidovorax wautersii* sp. nov. to accommodate clinical isolates and an environmental isolate, most closely related to *Acidovorax avenae*, Int. J. Syst. Evol. Microbiol, 63 (2013) 2203–2206. [PubMed: 23148096]

[60]. Wang H, Lou J, Gu H, Luo X, Yang L, Wu L, Liu Y, Wu J, Xu J, Efficient biodegradation of phenanthrene by a novel strain *Massilia* sp. WF1 isolated from a PAH-contaminated soil, Environ. Sci. Pollut. Res. Int, (2016).

[61]. Willems A, Falsen E, Pot B, Jantzen E, Hoste B, Vandamme P, Gillis M, Kersters K, De Ley J, *Acidovorax*, a new genus for *Pseudomonas facilis*, *Pseudomonas delafieldii*, E. Falsen (EF) group 13, EF group 16, and several clinical isolates, with the species *Acidovorax facilis* comb. nov., *Acidovorax delafieldii* comb. nov., and *Acidovorax temperans* sp. nov, Int. J. Syst. Evol. Microbiol, 40 (1990) 384–398.

[62]. Willems A, Gillis M, Genus II Acidovorax Willems, Falsen, Pot, Jantzen, Hoste, Vandamme, Gillis, Kersters and De Ley 1990, 394VP, in: Brenner DJ, Krieg NR, Staley JT, Garrity GM (Eds.) Bergey's Manual of Systematic Bacteriology, Springer, New York, 2005, pp. 696–703.

[63]. Willems A, Goor M, Thielemans S, Gillis M, Kersters K, De Ley J, Transfer of several phytopathogenic *Pseudomonas* species to *Acidovorax* as *Acidovorax avenae* subsp. *avenae* subsp. nov., comb. nov., *Acidovorax avenae* subsp. *citrulli*, *Acidovorax avenae* subsp. *cattleyae*, and 674*Acidovorax konjaci*, Int. J. Syst. Bacteriol, 42 (1992) 107–119. [PubMed: 1371056]

[64]. Zhang Z, Schwartz S, Wagner L, Miller W, A greedy algorithm for aligning DNA sequences, J. Comput. Biol, 7 (2000) 203–214. [PubMed: 10890397]
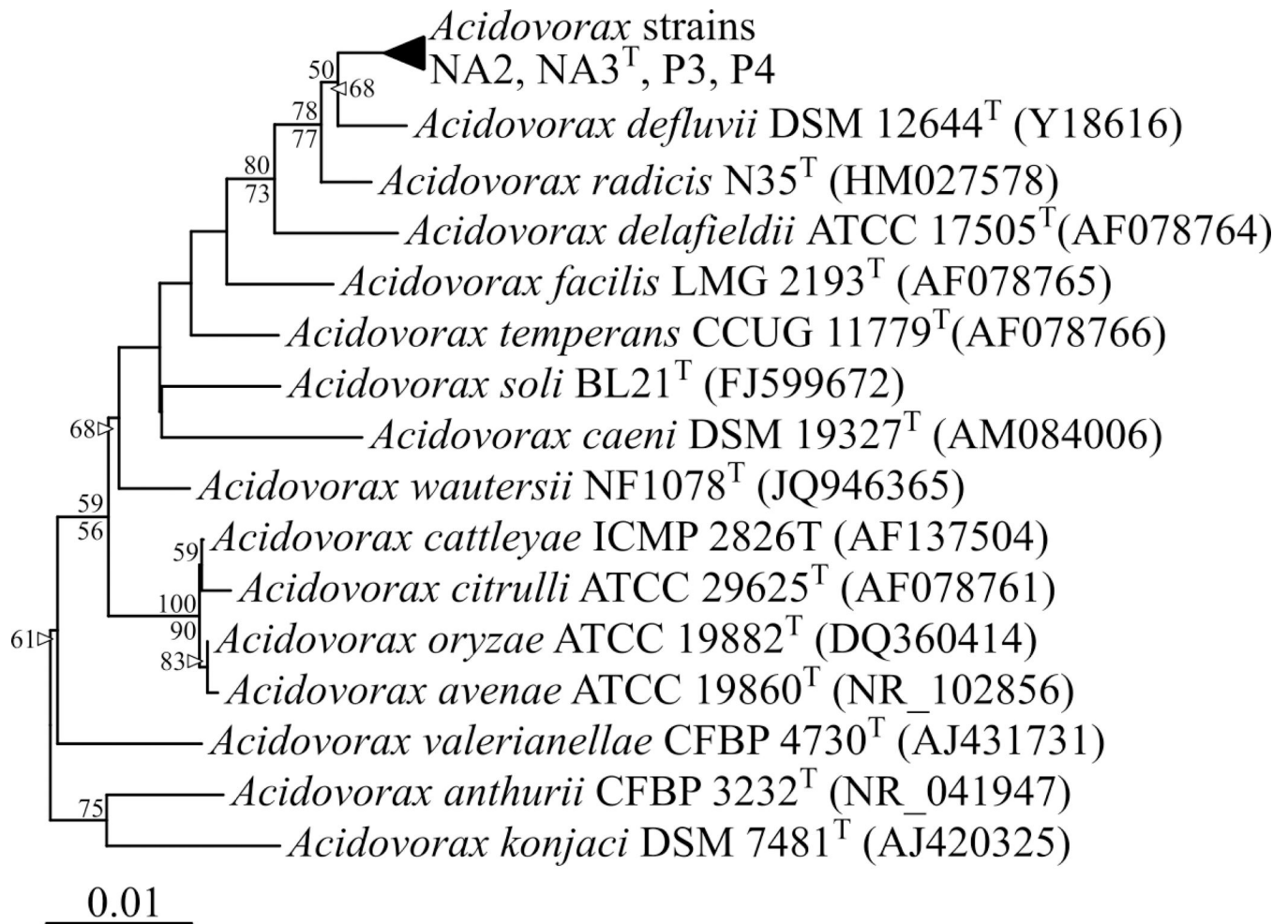
Figure 1.
Phylogenetic tree of 16S rRNA genes from *Acidovorax* species type strains compared to strains from this work. Labels at nodes indicate bootstrap support 50 (as a percentage of 1000 iterations) for neighbor-joining (above nodes) and maximum-parsimony (below nodes) algorithms. Genbank accession numbers are shown in parentheses. The tree was rooted with *Variovorax paradoxus* CBF3 (JN990697; not shown). The scale bar corresponds to 0.01 substitutions per nucleotide position.
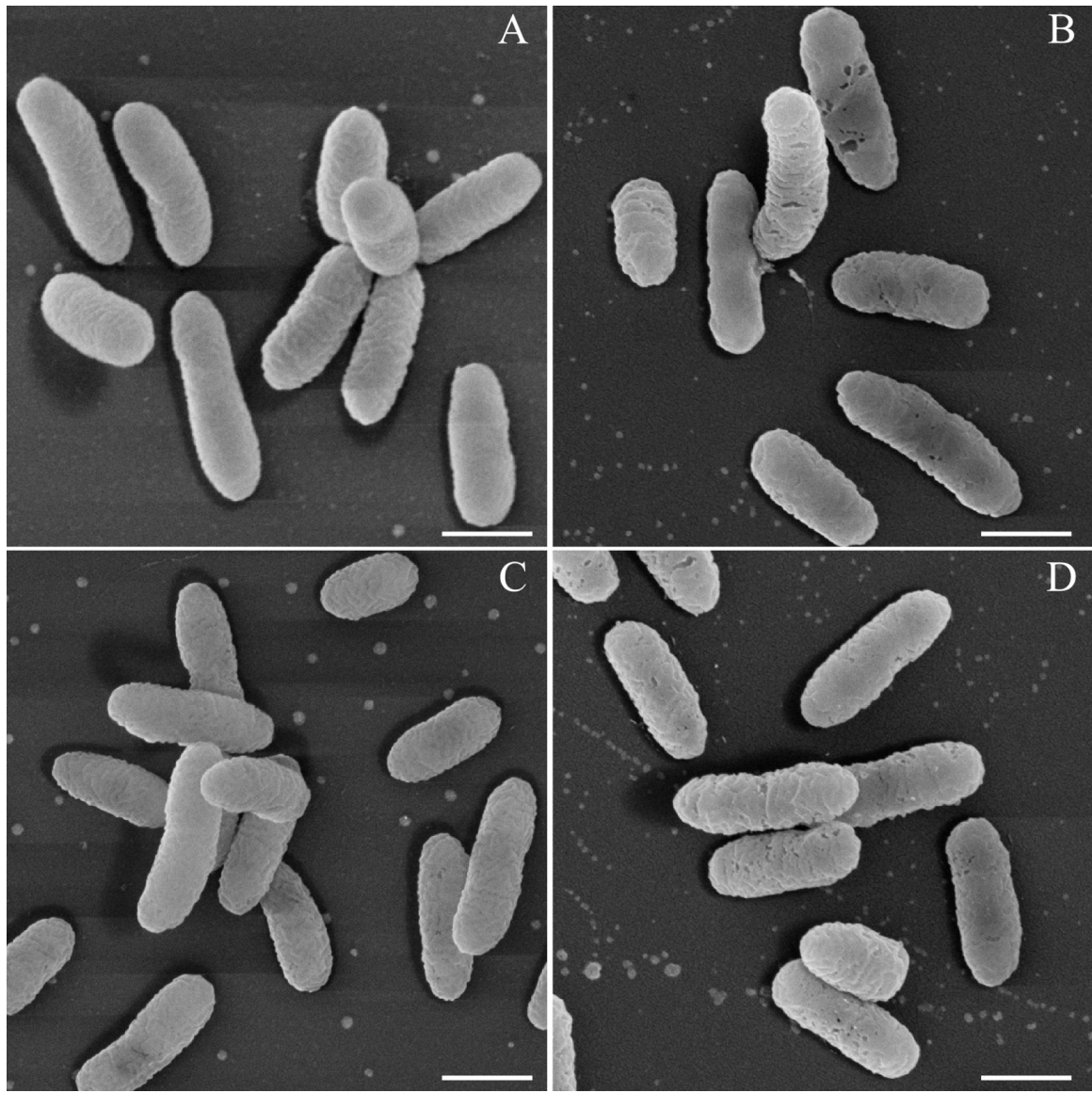
Figure 2.
Scanning electron micrograph of the four *Acidovorax* strains from this work: (A) P3, (B) P4, (C) NA2, (D) NA3T. The scale bar for all panels is 0.5 μm.
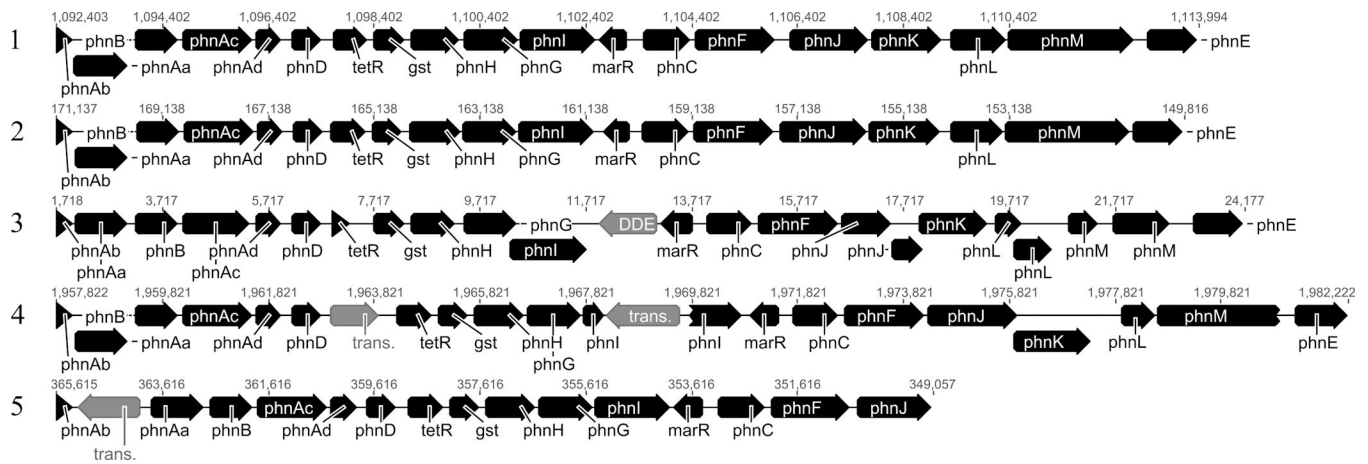
Figure 3.
Alignment of highly conserved genomic regions containing *phn* genes from (1) *Acidovorax* sp. NA3 [representing all strains from this work, except as noted in the text], (2) *Massilia* sp. WG5 [Genbank accession NZ_CP012641], also representing *Massilia* sp. WF1 [NZ_LELH00000000], (3) *Alcaligenes faecalis* AFK2 [AB024945], (4) *Delftia* sp. Cs1–4 695 [NC_015563], and (5) *Burkholderia* sp. Ch1–1 [NZ_ADNR00000000]. Previously unannotated genes for AFK2 were determined using Geneious (v8.1.7). Putative protein-encoding genes are shown in black and mobile genetic elements in grey. Numbers represent nucleotide positions in the associated GenBank or JGI entries. Gene designations were according to the naming convention established in the *Alcaligenes faecalis* AFK2 Genbank entry and Hickey et al. [24]. Abbreviations: tetR – TetR family transcriptional repressor; gst – glutathione S-transferase; marR – MarR family transcriptional regulator; DDE – DDE superfamily endonuclease; trans. – putative transposase. For more details on gene products and predicted pathway information refer to Hickey et al. [24].

Table 1.

Characteristics of *Acidovorax* strains from this study.

	NA2	NA3[T]	P3	P4
Genome characteristics[a]				
Total size (bp)	4058023	4122625	4565632	4704737
Chromosome size (bp)	3986972	4122625	4077847	4077222
Plasmid size(s) (bp)	71051	-	258084, 156922, 72779	269150, 163888, 114839, 79638
Mol%GC	64.14	64.24	63.71	63.56
No. of protein coding genes	3803	3972	4472	4483
No. of rRNA operons	4	4	4	4
No. of tRNA genes	53	54	55	55
No. of other RNA genes	15	12	17	17
No. of genes assigned to KEGG	2108	2037	2111	2269
No. of genes assigned to COGs	2850	2763	2842	3055
Biolog GN2 carbon utilization[b]				
Tween 40	++	++	±	++
Tween 60	++	++	++	++
Pyruvic acid methyl ester	++	++	++	++
Succinic acid mono-methyl ester	++	++	++	++
D-galacturonic acid	–	–	++	++
D-glucuronic acid	–	–	++	++
β-hydroxybutyric acid	++	++	++	++
α-ketobutryic acid	++	++	++	++
α-ketoglutaric acid	++	++	++	++
α-ketovaleric acid	++	++	++	++
DL-lactic acid	++	++	++	++
Succinamic acid	++	++	++	++
L-proline	++	–	++	++
2,3-butanediol	++	++	±	–
Glycerol	++	++	++	++
Fatty acids (%)				
$C_{8:0}$ 3-OH	0.3	0.4	0.3	0.3
$C_{10:0}$ 3-OH	3.1	3.5	3.0	3.0
C12:0	2.6	2.9	2.6	2.6
C14:0	2.0	2.5	2.1	2.1
iso-$C_{15:0}$	0.2	–	0.1	0.1
$C_{15:1}$ω6c	–	0.3	–	0.1
$C_{16:1}$ω7c/ω6c	47.3	48.2	47.8	47.2

	NA2	NA3^T	P3	P4
$C_{16:1}\omega6c$	–	–	–	0.1
C16:0	27.2	28.8	26.3	26.8
$C_{17:1}\omega6c$	–	0.1	–	–
iso-$C_{17:1}\omega9c$	0.2	–	0.2	0.3
iso-$C_{17:0}$	0.1	–	–	0.2
$C_{17:0}$ cyclo	0.3	0.2	0.5	0.7
C17:0	0.3	0.6	0.3	0.3
$C_{18:1}\omega7c$	16.0	12.1	16.4	16.0
C18:0	0.2	–	0.2	0.2
$C_{18:1}\omega7c$ 11-methyl	–	0.2	0.1	0.1
$C_{19:0}$ cyclo $\omega8c$	0.2	–	–	–

[a] Based on JGI/IMG annotations

[b] Symbols: ++, positive results in both replicates; ±, positive in one of two replicates; –, negative in both replicates

Table 2.

Average nucleotide identity (ANI) between *Acidovorax* strains from this study and characterized *Acidovorax* strains with publicly available genomic data. Additional genomic comparisons to uncharacterized *Acidovorax* strains are available in Appendix A.

Acidovorax sp.	strain	NA2	NA3T	P3	P4
phenanthenivorans	NA2	100	96.9	97.1	97.1
phenanthenivorans	NA3T	96.9	100	97.0	97.1
phenanthenivorans	P3	97.1	97.0	100	100
phenanthenivorans	P4	97.1	97.1	100	100
acenae	T10_61	80.6	80.8	80.7	80.7
avenae	ATCC 19860T	80.8	80.7	80.8	80.8
acenae	RS-1	80.7	80.7	80.7	80.8
caeni	R-24608T	80.5	80.4	80.5	80.4
cattleyae	DSM 17101	80.7	80.5	80.7	80.7
citrulli	AAC00–1	80.9	80.9	80.9	80.9
citrulli	DSM 17060	80.7	80.6	80.8	80.7
citrulli	M6	80.8	80.8	80.9	80.8
citrulli	pslb65	80.8	80.7	80.8	80.9
citrulli	tw6	80.8	80.7	80.9	80.8
citrulli	ZJU1106	81.7	81.7	81.8	81.8
delafieldii	2AN	85.6	85.6	85.7	85.6
delafieldii	CCH3-G3	83.9	83.9	84.2	84.2
delafieldii	CCH4-A2	n.d.a	n.d.	n.d.	n.d.
ebreus	TPSY	81.1	81.2	81.0	81.1
konjaci	DSM 7481T	80.3	80.5	80.4	80.4
oryzae	ATCC 19882T	80.6	80.8	80.7	80.8
radicis	N35T	84.0	83.8	83.9	83.8
radicis	N35v	83.9	83.8	83.8	83.8
soli	DSM 25157	87.4	87.3	87.6	87.6
temperans	KY4	82.7	82.7	82.8	82.7
valerianellae	DSM 16619	81.3	81.4	81.2	81.4
wautersii	DSM 27981	81.6	81.7	81.7	81.8

[a]Insufficient hits between genomes to estimate 2-way ANI.

Table 3.

Differentiation of *Acidovorax carolinensis* from species without available genomic data[a]. Strains: 1, *A. carolinensis* spp. NA2, NA3T, P3, and P4; 2, *A. defluvii* DSM 12644T; 3, *A. facilis* DSM 649T.

	1	2	3
Biolog GN2 microplate carbon utilization			
D-galactose	-[b]	+	+
D-mannitol	-	+	+
L-arabinose	-	+	+
L-serine	-	+	+
Monomethyl succinate	-	+	+
DL-carnitine	-	-	+
D-mannose	-	w	+
D-psicose	-	-	+
D-sorbitol	-	-	+
Glycyl L-aspartic acid	-	-	+
L-alaninamide	-	+	-
L-alanyl glycine	-	-	+
L-glutamic acid	-	+	-
L-ornithine	-	+	-
L-phenylalanine	-	-	+
N-acetyl-D-glucosamine	-	-	+
Succinamic acid	-	+	+
Succinate	+	-	w
Sucrose	-	+	-
Turanose	-	+	-
α-D-glucose	-	-	+
γ-aminobutyric acid	-	-	+
γ-hydroxybutyric acid	-	+	+
Other			
Catalase reaction	-	+	+

[a]Data from Li et al. [35], Schulze et al. [46], and Willems et al. [61].

[b]Symbols: +, positive; -, negative; w, weakly positive

Table 4.

Oxygenase and dioxygenase genes putatively involved in aromatic metabolism in *A. carolinensis* strains.

Protein family (AromaDeg Cluster)	IMG Gene ID; GenBank Protein Accession No.				Blastp results		
	NA2	NA3[T]	P3	P4	Top hit(s) (Accession No.)	Perc. Iden. (Simil.)[a]	Gene Design.[a]
Benzoate	2619698514; ART47637	2619702319; ART51195	2619707999; ART55684	2619710954; ART58452	*Cupriavidus necator* (2Fe–2S)-binding protein (WP_051974707)	77–78 (87)	
Biphenyl (XXIV)	2619698485; ART50028	2619702290; ART53508	2619708030; ART56870	2619710924; ART60761	*Delftia* sp. Cs1–4, aromatic ring-hydroxylating dioxygenase subunit alpha (WP_013801305); *Burkholderia* sp. Ch1-1 aromatic ring-hydroxylating dioxygenase subunit alpha (WP_007179244)	98–99 (99)	*phnAc*
EXDO bicyclic substrates (XII, XIII)	2619698494; ART47619	2619702299; ART51178	2619708021; ART55699	2619710933; ART58435	*Alcaligenes faecalis* 3,4-dihydroxyphenanthrene dioxygenase (BAA76330)	99–100 (100)	*phnC*
EXDO monocyclic substrates (XX)	2619697639; ART46896	2619701445; ART50531	2619709018; ART57655	2619709989; ART57655	*Comamonas testosteroni* catechol 2,3-dioxygenase (WP_034382469)	97–98 (98–99)	
Gentisate (XXIII)	2619698491; ART50030	2619702296; ART53510	2619708024; ART56868	2619710930; ART60763	*Delftia* sp. Cs1–4, cupin (WP_013801312); *Alcaligenes faecalis* 1-hydroxy-2-naphthoate dioxygenase (BAA76328)	100 (100); 99 (99)	*phnG*
LigB protocatechuate (I)	2619698471; n.a.	2619702276; n.a.	2619708046; n.a.	2619710910; n.a.	*Limnohabitans* sp. 103DPR2 protocatechuate 3,4- dioxygenase (WP_055360450)	86 (93)	*pmdB*
	–	–	2619709554; ART527279	2619713890; ART61320	*Acidovorax* sp. CCH12-A10 protocatechuate 3,4-dioxyganse (WP_066789040)	100 (100)	*pmdB*[c]
LigB protocatechuate (XX)	2619698512; ART47635	2619702317; ART51193	2619708001; ART55686	2619710952; ART58450	*Cupriavidus necator* extradiol ring-cleavage dioxygenase (WP_042877061)	80–81 (87)	
LigB protocatechuate (XXXVI)[b]	–	2619703300, 2619703301; n.a.	2619707065; ART55085	2619711819; ART59161	*Noviherbaspirillum* sp. Root189, 3-(2,3-dihydroxyphenyl)propionate dioxyganase (WP_057291057)	72–74 (81–83)	

	IMG Gene ID; GenBank Protein Accession No.				Blastp results		
Protein family (AromaDeg Cluster)	NA2	NA3^T	P3	P4	Top hit(s) (Accession No.)	Perc. Iden. (Simil.)^a	Gene Design.^{a}
Phthalate (VI)	2619698529; ART47651	2619702334; ART51209	2619707984; ART55670	2619710969; ART58466	*Hydrogenophaga* sp. RAC07 MarR family transcriptional regulator (WP_069049060)	85 (92)	*ophA2*
Phthalate (VI)	–	–	2619709578; ART57294	2619713912; ART61337	*Acidovorax* sp. CCH12-A4 MarR family transcriptional regulator (WP_066786982)	99 (100)	*ophA2*^{c}

^{a}Abbreviations: Perc. Iden. – amino acid identity (percentage); Simil. – amino acid similarity; Gene Design. – gene designation; n.a. – not annotated as protein using NCBI pipeline. Only gene designations for those potentially involved in phenanthrene metabolism are shown. Coverage for all Blastp comparisons between strains from this study was 98–100%.

^{b}Gene for putative dioxygenase in strain NA3T was split in two; Gene ID numbers for both genes are provided. Values for amino acid identity/similarity are for the *in silico* recombined whole gene.

^{c}Genes are plasmid-borne.

Table 5.

Description of *Acidovorax carolinensis* sp. nov. according to Digital Protologue TA00473 assigned by the the www.imedea.uib.es/dprotologue website.

TAXONUMBER	TA00473
SPECIES NAME	*Acidovorax carolinensis*
GENUS NAME	*Acidovorax*
SPECIFIC EPITHET	*carolinensis*
SPECIES STATUS	sp. nov.
SPECIES ETYMOLOGY	ca.ro.li.nen'sis. N.L. masc. adj. *carolinensis* of or belonging to Carolina, refering to North Carolina, USA, where each of the strains were first isolated
SUBMITTER	David Singleton
E-MAIL OF THE SUBMITTER	drsingle@email.unc.edu
HAS THE TAXON BEEN SUBJECTED TO EMENDATION?	NO
DESIGNATION OF THE TYPE STRAIN	NA3
STRAIN COLLECTION NUMBERS	=LMG 30136, =DSM 105008
16S rRNA GENE ACCESSION NUMBER	EU910093.1
GENOME ACCESSION NUMBER [RefSeq]	NZ_CP021361.1
GENOME STATUS	Complete
GENOME SIZE	4122 kbp
GC mol %	64.24
COUNTRY OF ORIGIN	USA
REGION OF ORIGIN	Charlotte, North Carolina
DATE OF ISOLATION UNKNOWN (<*yyyy*)	2005
SOURCE OF ISOLATION	Contaminated soil
SAMPLING DATE	2002–01–01
GEOGRAPHIC LOCATION	Charlotte, NC former manufactured gas plant site
TEMPERATURE OF THE SAMPLE [In Celsius degrees]	20
pH OF THE SAMPLE	7.5
NUMBER OF STRAINS IN STUDY	4
SOURCE OF ISOLATION OF NON- TYPE STRAINS	Contaminated soil
GROWTH MEDIUM, INCUBATION CONDITIONS [Temperature, pH, and further information] USED FOR STANDARD CULTIVATION	Nutrient agar/R2A, 30–32.5°C, pH 7–7.5
CONDITIONS OF PRESERVATION	15% glycerol; –80°C

GRAM STAIN	Negative
CELL SHAPE	Rod
CELL SIZE (length or diameter)	1.2
MOTILITY	Motile
IF MOTILE	Flagellar
IF FLAGELLATED	Single, polar
SPORULATION (resting cells)	None
COLONY MORPHOLOGY	Cream color to pale yellow, circular, 1.0–1.5mm diameter.
TEMPERATURE RANGE	4–40
LOWEST TEMPERATURE FOR GROWTH	4, nutrient agar plates
HIGHEST TEMPERATURE FOR GROWTH	40, nutrient agar plates
TEMPERATURE OPTIMUM	30–32.5
LOWEST pH FOR GROWTH	5.5, nutrient broth
HIGHEST pH FOR GROWTH	8.0, nutrient broth
pH OPTIMUM	7.0–7.5
pH CATEGORY	Neutrophile
LOWEST NaCl CONCENTRATION FOR GROWTH	0, nutrient broth
HIGHEST NaCl CONCENTRATION FOR GROWTH	1%, nutrient broth
SALINITY CATEGORY	Nonhalophile (NaCl inhibitory at <1 % NaCl)
RELATIONSHIP TO O2	Aerobe
O2 CONDITIONS FOR STRAIN TESTING	Aerobiosis; anaerobiosis
CARBON SOURCE USED [class of compounds]	Organic acids, aromatic compounds
CARBON SOURCE USED [specific compounds]	Phenanthrene
CARBON SOURCE NOT USED [specific compounds]	Acenaphthene, anthracene, benzene, ethylbenzene, fluoranthene, fluorene, naphthalene, phthalate, pyrene, salicylate, toluene, xylene
Positive tests with BIOLOG	Tween 40, Tween 80, Methyl pyruvate, succinic acid mono-methyl ester, β-hydroxybutyric acid, α-ketobutyric acid, α-ketoglutaric acid, α-ketovaleric acid, DL-lactic acid, succinamic acid, glycerol
Negative tests with BIOLOG	α-Cyclodextrin, Dextrin, Glycogen, Tween 40, N-Acetyl-D-galactosamine, N-Acetyl-D-glucosamine, Adonitol, L-Arabinose, D-Arabitol, D-Cellobiose, i-Erythritol, D-Fructose, L-Fucose, D- Galactose, Gentiobiose, D-Glucose, m-Inositol, α-D-Lactose, Lactulose, Maltose, D-Mannitol, D-Mannose, D-Melibiose, β-Methyl-D-Glucoside, D-Psicose, D-Raffinose, L-Rhamnose, D- Sorbitol, Sucrose, D-Trehalose, Turanose, Xylitol, Acetic Acid, Cis-Aconitic Acid, Citric Acid, Formic Acid, D-Galactonic Acid Lactone, D-Glucosaminic Acid, D-Glucuronic Acid, α- Hydroxy-Butyric Acid, γ-Hydroxy Butyric Acid, p-Hydroxy Phenylacetic Acid, Itaconic Acid, α-Keto Valeric Acid, Malonic Acid, Propionic Acid, Quinic Acid, D-Saccharic Acid, Sebacic Acid, Bromo-Succinic Acid, Glucuronamide, L-Alaninamide, D-Alanine, L-Alanine, L- Alanylglycine, L-Asparagine, L-Aspartic Acid, L-Glutamic Acid, Glycyl-L-Aspartic Acid, Glycyl-L-Glutamic Acid, L-Histidine, Hydroxy-L-Proline, L-Leucine, L-Ornithine, L-Phenylalanine, L-Pyroglutamic Acid, D-Serine, L-Serine, L-Threonine, DL-Carnitine, γ-Amino Butyric Acid, Urocanic Acid, Inosine, Uridine, Thymidine, Phenyethylamine, Putrescine, 2- Aminoethanol, D,L-α-Glycerol Phosphate, Gluose-1-Phosphate, Glucose-6-Phosphate

Variable tests with BIOLOG	D-galacturonic acid, D-glucuronic acid, 2,3-butanediol, L-proline
NITROGEN SOURCE	NH_4
TERMINAL ELECTRON ACCEPTOR	Oxygen
ENERGY METABOLISM	Chemoorganotroph
OXIDASE	Positive
CATALASE	Negative
POSITIVE TESTS	Lipase, nitrate reduction
NEGATIVE TESTS	Gelatin hydrolysis, cellulase, DNase, skim milk protease, starch hydrolysis, urease
QUINONE TYPE	R8
MAJOR FATTY ACIDS	C16:1ω7c/C16:1ω6c, C16:0, C18:1ω7c
PHOSPOLIPID PATTERN OR DIAGNOSTIC PHOSPOLIPID	Phosphatidylethanolamine, diphosphatidylglycerol, phosphatidylglycerol
BIOSAFETY LEVEL	1
HABITAT	Terrestrial biome
BIOTIC RELATIONSHIP	Free-living
KNOWN PATHOGENICITY	None