Saezia sanguinis gen. nov., sp. nov., a Betaproteobacteria member of order Burkholderiales, isolated from human blood

María J Medina-Pascual, Sara Monzón, Pilar Villalón, Isabel Cuesta, Fernando González-Romo, Sylvia Valdezate

Int J Syst Evol Microbiol. 2020 Mar;70(3):2016-2025.

https://doi.org/10.1099/ijsem.0.004010

TITLE PAGE

Title: ***Saezia sanguinis* gen. nov., sp. nov., a *Betaproteobacteria* member of order *Burkholderiales*, isolated from human blood**.

Author names: María J Medina-Pascual[1], Sara Monzón[2], Pilar Villalón[1], Isabel Cuesta[2], Fernando González-Romo[3] and Sylvia Valdezate[1]

Affiliation: [1] Laboratorio de Referencia e Investigación en Taxonomía. Centro Nacional de Microbiología. Instituto de Salud Carlos III (Majadahonda, Madrid, Spain).

[2] Unidad de Bioinformática (BU-ISCIII). Instituto de Salud Carlos III. (Majadahonda, Madrid, Spain).

[3]Hospital Clínico San Carlos (Madrid, Spain).

Corresponding autor: mjmedina@isciii.es

Keywords: *Burkholderiales*; *Comamonadaceae*; *Saezia*; *sanguinis*

Repositories:

GenBank accession for 16S rRNA gene sequence: KY039174

GenBank accession for genome data: PQSP00000000

DSMZ strain number: DSM 104959

CECT strain number: CECT 9208

Saezia sanguinis gen. nov., sp. nov., a *Betaproteobacteria* member of the order *Burkholderiales*, isolated from human blood

María J Medina-Pascual[1], Sara Monzón[2], Pilar Villalón[1], Isabel Cuesta[2], Fernando González-Romo[3] and Sylvia Valdezate[1]

[1] Laboratorio de Referencia e Investigación en Taxonomía. Centro Nacional de Microbiología. Instituto de Salud Carlos III (Majadahonda, Madrid, Spain).

[2] Unidad de Bioinformática (BU-ISCIII). Instituto de Salud Carlos III. (Majadahonda, Madrid, Spain).

[3]Hospital Clínico San Carlos (Madrid, Spain).

ABSTRACT

The taxonomic position of an unknown bacterial strain designated CNM695-12, isolated from the blood of an immunocompromised subject, was investigated via phenotypic, chemotaxonomic, genotypic and genomic analyses. Bacterial cells were determined to be Gram-negative bacilli, aerobic, non-motile and non-sporeforming. The strain showed catalase activity but no oxidase activity. Optimal growth occurred at 37^{o}C, pH 7 and with 0-1% NaCl. C16:0, summed feature 8 (comprising C18:1 ω7c), and C18:1 ω9c were the most abundant fatty acids, and UQ8 the major respiratory quinone. The polar lipids present included phosphatidylglycerol, phosphatidylethanolamine and other aminophospholipids. The 16S rRNA gene sequence showed ≈93.5% similarity to different species with validly published names within the order *Burkholderiales* (e.g., *Leptothrix mobilis* Feox-1^{T}, *Aquabacterium commune* B8^{T}, *Aquabacterium citratiphilum* B4^{T} and *Schlegelella thermodepolymerans* K14^{T}). Phylogenetic analyses based on 16S rRNA gene sequences and concatenated alignments including the sequences for 107 essential proteins, revealed the strain to form a new lineage close to family *Comamonadaceae*. The highest average nucleotide identity and average amino acid identity values were obtained with *Schlegelella thermodepolymerans* K14^{T} (69.6 and 55.7% respectively). The genome, with a size of 3.35 Mb, had a G+C content of 52.4 mol% and encoded 3056 predicted genes, 3 rRNA, 1 tmRNA and 51 tRNA. Strain CNM695-12 thus represents a new species belonging to a novel genus within the order *Burkholderiales*, for which the name *Saezia sanguinis* gen. nov., sp. nov. is proposed. The type strain is CNM695-12^{T} (=DSM 104959^{T} = CECT 9208^{T}).

Keywords: *Burkholderiales*; *Comamonadaceae*; *Saezia*; *sanguinis*

Bacteria belonging to the order *Burkholderiales* show wide phenotypic, metabolic and ecological diversity; strictly-aerobic and facultatively-anaerobic chemoorganotrophs, obligate and facultative chemolithotrophs, and nitrogen-fixing organisms are all represented (1). The type genus of this order is *Burkholderia*. The families *Alcaligenaceae, Burkholderiaceae, Comamonadaceae, Oxalobacteraceae* and *Sutterellaceae* fall within this same order, although *Burkholderiales* also comprises genera that cannot be accommodated within these families (https://www.namesforlife.com/). The order comprises Gram-negative microorganisms, mostly isolated from soil, water and vegetation, but which can infect plants and animals (1–4). Species such as *Stenotrophomonas maltophilia*, *Pandoraea apista*, *Ralstonia picketii*, *Achromobacter xylosoxidans*, and members of the *Burkholderia cepacia* complex (Bcc) can behave as opportunistic human pathogens that mainly affect immunocompromised people (5). They have caused nosocomial outbreaks (5–7). Many of the above species show intrinsic resistance to many antimicrobial agents, making it difficult, for example, to treat patients with cystic fibrosis infected with *Stenotrophomonas maltophilia* or *Burkholderia cepacia* complex species (5,6).

The bacterium described in this work was isolated from an 84 year-old man with myelodysplastic syndrome admitted to a hospital in Madrid (Spain) with a respiratory infection. After treatment with ceftriaxone, levofloxacin and amphotericin B he progressed favourably, and was discharged. A week later, however, the patient's condition worsened, and he was readmitted to hospital with fever, malaise, dyspnoea with crackles in the lungs, and oedema in the lower limbs. A haemogram revealed him to have underlying acute myeloid leukaemia. *Escherichia coli* and an unidentified Gram-negative bacillus were isolated from blood cultures. Meropenem was administered but the patient died two days later. The unidentified strain was submitted for characterization to our laboratory (Reference and Research Laboratory for Taxonomy at the National Centre for Microbiology, Madrid, Spain) where it was codified as CNM695-12.

The strain was plated on 5% sheep blood Columbia agar (BCA) and incubated at 37^{o}C for 48 h under aerobic conditions for its later identification and conservation. Gram-staining revealed small, Gram-negative bacilli. Two replicates of the strain were

examined using the VITEK MS automated mass spectrometry system (bioMérieux) - which uses matrix-assisted laser desorption ionization time-of-flight technology - following the manufacturer's protocol and making use of the VITEK MS v.3.2.0 database. However, no ID profile was obtained (Fig. S1 shows the mass spectrum obtained). Genomic DNA was extracted using the QIAamp DNA Minikit (QIAGEN) following the manufacturer's instructions, and strain identification performed by 16S rRNA gene sequence analysis, using primers fD1 and rP2 (8) for amplification, and E781 and U1115 for sequencing (9). Contigs were assembled using SeqMan Pro v.12.3.1 software (DNAStar package) and a 1467 bp sequence was obtained (GenBank accession no. KY039174). This sequence was compared to those available in different databases (http://blast.ncbi.nlm.nih.gov/Blast.cgi, https://www.ezbiocloud.net/identify and http://umr5558-bibiserv.univ-lyon1.fr/lebibi/lebibi.cgi). Percentage similarities with sequences of bacterial species with validly published names were <94%. According to the cut-offs used to classify bacterial isolates (10), strain CNM695-12 could represent a new species of a novel genus within the order *Burkholderiales*. The strongest 16S rRNA gene sequence similarity (99.8%) was obtained with strain Da-11 isolated in Australia from the hindgut of *Dermolepida albohirtum* larvae (11), but neither that strain nor its genomic DNA was available for further comparison.

Having failed to identify strain CNM695-12 by 16S rRNA sequencing and VITEK MS, the aim of the present work was to determine its taxonomic position via phenotypic, chemotaxonomic, genotypic and genomic analyses.

For its phenotypic characterization, the growth of the bacterium at 20, 25, 30, 37, 40, 42 and 45^{o}C was tested under aerobic conditions on BCA. Plates were also incubated in a 5% CO_2 atmosphere and under anaerobic conditions at 37^{o}C. The pH range for growth was examined from pH 4.0 to 9.0, with intervals of 1 pH unit, at 37^{o}C in trypcase soya broth (TSB) adjusted by the addition of 1 M NaOH or HCl (pH verified after autoclaving). Tolerance to NaCl was investigated at 37^{o}C in TSB with 1, 2, 3, 4, 5 and 10% (w/v) NaCl. Growth on trypcase soya agar (TSA), chocolate-polivitex agar (CPA) and McConkey agar (MA) (all from bioMérieux) was also tested. Cells and ultrathin cell sections were observed by transmission electron microscopy (TEM) using a Tecnai 12 device (Philips) at 120kV, after negative staining with 2% phosphotungstic acid and using cells in the exponential phase of growth. Biochemical and enzyme profiles were determined at 37^{o}C using the API 20NE, API 20E, API 50 CH, API ZYM, ID Color

Catalase, Oxidase Reagent (all from bioMérieux) and GN2 MicroPlate (Biolog) kits following the manufacturers' instructions. Cell motility was studied using motility-indole-ornithine medium (bioMérieux) by stab-inoculation from a bacterial suspension in TSB, and from colonies grown on BCA for 48 h. The minimum inhibitory concentrations (MICs) of ampicillin, piperacillin, amoxicillin/clavulate, piperacillin-tazobactam, cefepime, ceftazidime, ceftriaxone, aztreonam, imipenem, meropenem, amikacin, gentamicin, tobramycin, minocycline, tetracycline, ciprofloxacin, levofloxacin, moxifloxacin, chloramphenicol and cotrimoxazole were determined on Muller-Hinton medium at 37^{o}C after 48 h of incubation, using the Etest (bioMérieux) and following the manufacturer's instructions. *Pseudomonas aeruginosa* ATCC 27853 and *Escherichia coli* ATCC 25922 were used as control strains. The results were interpreted according to the criteria of the Clinical and Laboratory Standards Institute (CLSI) supplement M-100 for other non-*Enterobacteriaceae* (12).

Cellular fatty acids analysis was performed at the Spanish Type Culture Collection (*Colección Española de Cultivos Tipo*) at the University of Valencia, Spain. Cells were grown on R2A medium for 48 h at 30^{o}C; extractions and determinations were performed according to the standard protocol of the MIDI Microbial Identification System (13) using an Agilent 6850 gas chromatograph (Agilent Technologies) following the TSBA6 method (14). Respiratory quinones were identified by the German Collection of Microorganisms and Cell Cultures (DSMZ) Identification Service (Braunschweig, Germany) by thin layer chromatography on silica gel followed by HPLC analysis using a reverse phase column with methanol:heptane as the eluent (15). Polar lipids were determined by the same service via two dimensional silica gel thin layer chromatography. Total lipid material was detected using molybdatophosphoric acid. Functional groups were revealed using specific spray reagents (16).

Pairwise similarity values for the CNM695-12 16S rRNA gene sequence and the closest type strains of the order *Burkholderiales* were calculated using the web-based tool available at http://www.ezbiocloud.net/tools/pairAlign (17). In addition, sequences were analysed using Bioedit software (18) after multiple alignments of the data using MAFFT v.7 software (available at https://mafft.cbrc.jp/alignment/server/) (19);. *Aquaspirillum serpens* DSM 68^{T} was used as an outgroup. A phylogenetic study was performed by constructing a neighbour-joining (NJ) tree using the Tamura 3-parameter substitution model (20) and the rate of variation among sites modelled using a gamma

distribution (shape parameter=0.5). Maximum parsimony (MP) (21) and maximum likelihood (ML) (22) trees were generated to check the robustness of the analysis. All these studies were conducted using MEGA7 software (23), and the reliability of the phylogenetic trees assessed by Bootstrap testing with 1000 replicates. A NJ tree was also constructed with the type strains of the type genera belonging to each family and to each unclassified genus within the order *Burkholderiales*.

Genomic DNA was sequenced using the Illumina NextSeq500 sequencing system (http://www.illumina.com) and the paired-end library generated by the Nextera indexing system (Illumina 1.9). A quality control analysis involving fastQC v0.11.3 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was performed and any adapter sequences removed using Trimmomatic v.0.36 software (24). Sequence reads were *de novo* assembled using SPAdes v.3.8.0 software, applying kmers 21, 33, 55 and 77 (25). An assembly quality control check was performed using QUAST software (26). Finally, Prokka v.1.12-beta software was used for automatic structural and functional annotation (27).

In-silico genome-to-genome comparisons between the newly proposed species and the closest type and non-type strains with genomes accessible in the NCBI database (https://www.ncbi.nlm.nih.gov/genome/), were performed using the GGDC 2.1 update of the web service http://ggdc.dsmz.de and employing the suggested BLAST+ software. The recommended formula of identities/HSP length, which is independent of the length of the genome sequences, was used for comparisons (28). The average nucleotide identity (ANI) and average amino acid identity (AAI) were calculated using tools available at https://www.ezbiocloud.net/tools/ani (29) and http://enve-omics.ce.gatech.edu/ (30). Phylogenetic relationships were studied as previously described (31), involving the 19 related strains with the highest ANI values and whose proteomes were available at the NCBI website. Analysis was performed using the bcgTree software tool (https://github.com/iimog/bcgTree) (32), creating a full concatenated alignment including 107 essential proteins. Genes encoding 107 proteins are shown in Table S1. A maximum likelihood tree was built using RAxML software (https://cme.h-its.org/exelixis/web/software/raxml/) (33). GTRGamma was selected as the model (-m GTRGAMMA) for RAxML, using the rapidbootstrap option (-f a) with 1000 bootstrap replicates. *Aquaspirillum serpens* DSM 68^{T} was used as an outgroup.

Additional bioinformatic analyses were performed to examine the strain's metabolic pathways (using KEGG software, http://www.genome.jp/kegg/kaas) (34), the presence of secondary metabolite biosynthesis gene clusters (using antiSMASH v.4.0 software, https://antismash.secondarymetabolites.org/) (35), antimicrobial resistance genes (using ResFinder v.3.1 software, https://cge.cbs.dtu.dk/services/ResFinder/)(36) and virulence factors (using VirulenceFinder v.2.0 software, https://cge.cbs.dtu.dk/services/VirulenceFinder/) (37); and to estimate pathogenicity (using PathogenFinder v.1.1 software, https://cge.cbs.dtu.dk/services/PathogenFinder/) (38) (the last three softwares are the property of the Center for Genomic Epidemiology).

Cells were confirmed as Gram-negative, non-sporeforming, non-motile and rod-shaped (0.5 to 0.8 µm wide and 1.4 to 2.2 µm long). Pili-like structures 9-29 nm in diameter were observed under TEM (Fig. S2). After 48 h of aerobic growth on BCA, colonies reached 1-2 mm in diameter and were punctiform, bright, opaque and convex with entire margins. Growth occurred after 2 days at 30°C and 37°C (optimum 37°C); in a 5% CO_2 atmosphere and under aerobic conditions; at pH 5-9 (optimum pH 7); and with 0-4% NaCl (optimum 0-1%). Colonies were observed after 48 h at 37°C on CPA and after 72 h on TSA and MA (without lactose fermentation). Catalase activity was detected, but unlike for most species belonging to the order *Burkholderiales*, no oxidase activity was observed. API 20NE tests were positive for nitrate reduction and the assimilation of adipic acid and malate (profile 1000060: but identifying the strain as *Comamonas testosteroni/Pseudomonas alcaligenes* with just 57.8% probability). Acetoin production was positive with the API 20E test. All tests were negative for API 50 CH. Using the API ZYM kit, alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, cystine arylamidase, acid phosphatase and naphthol-AS-BI-phosphohydrolase activities were detected. In the GN2 MicroPlate test, pyruvic acid methylester, acetic acid, sebacic acid, α-ketobutyric acid, α-ketoglutaric acid, succinic acid, bromosuccinic acid, succinamic acid, L-aspartic acid, L-glutamic acid and L-proline were seen to be utilized as carbon sources. Table 1 shows the differential characteristics of strain CNM695-12 and the type strains of the phylogenetically related species within the order *Burkholderiales*.

The major cellular fatty acids in strain CNM695-12 were C16:0 (29.8%), summed feature 8 (comprising C18:1 ω7c, 19.7%), and C18:1 ω9c (10.2%). Percentages close to 30% of C16:0 were also encountered in related species such as *Comamonas*

serinivorans, Aquabacterium commune, Rubrivivax gelatinosus and *Leptothrix mobilis*, and above 40% in C*aldimonas manganoxidans* and *Schlegelella thermodepolymerans* (39–44). In addition, C18:1 ω7c, detected in CNM695-12 as feature 8, has been reported for *Comamonas humi* and *Rhizobacter gummiphilus* (2,3). The main fatty acid in *Comamonas humi*, *Leptothrix mobilis* and *Aquabacterium commune* - summed feature 3 (comprising C16:1 ω7c and/or C16:1 ω6c) with percentages above 40% - barely reached 10% in strain CNM695-12. Although the total hydroxyl fatty acid fraction (9.5%) was similar to that seen in related species, the distribution of these acids differed, with C18:1 2OH the major hydroxyl fatty acid in CNM695-12 (4.3%) and C10:1 3OH the main one in species belonging to the genera *Comamonas*, *Schlegelella* and *Aquabacterium*, (2,40,45). Finally, cyclopropane fatty acids - typical for many Gram-negative bacteria that use the anaerobic pathway of fatty acid biosynthesis (46) - were not found in CNM695-12. As reported in the literature (47), quinones were represented only by ubiquinones (UQ) in the strain described. As occurs within the order *Burkholderiales*, UQ8 was the main respiratory quinone (94%) but UQ7, previously detected in *Comamonas terrigena* ATCC 8461^T (47), was also found in CNM695-12 (6%). The polar lipids identified were phosphatidylglycerol (PG), phosphatidylethanolamine (PE) and other aminophospholipids (PNL) (Fig. S3) whereas diphosphatidylglycerol (DPG), which is present in related species such as *Comamonas humi*, *Comamonas serinivorans* or *Rhizobacter gummiphilus* (2,48), was not observed in strain CNM695-12.

A total of 7,345,408 paired-end reads with a length of 35-151 bp were obtained by genomic DNA sequencing, deriving in 6,014,312 high-quality reads (81.9%) between 50-151 bp after error correction. Assembly resulted in 245 contigs (maximum 909,344 bp; N50 = 275,526). The final assembly contained 3,349,558 bp, with a fitted mean coverage of 195x. The G+C content was 52.4 mol% (data computed by QUAST software). Annotation using Prokka v.1.12-beta software predicted 3056 genes (3001 coding sequences, 3 rRNA genes, 1 transfer-messenger-RNA and 51 tRNA genes). To ensure the quality and authenticity of the genome data, 16S rRNA sequences obtained by conventional Sanger analysis and from genome assembly were compared; both were identical. The draft genome sequence was deposited in the DDBJ/ENA/GenBank under accession number PQSP00000000. The version described in this paper is version PQSP01000000. The genome size and the G+C content of strain CNM695-12 were

similar to those of *Limnohabitans curvus* MWH-C5^T (3.03 Mb, 55.9%), *Herminiimonas fonticola* S-94^T (3.12 Mb, 51.6%) and *Advenella kashmirensis* UBA 4006 (3.24 Mb, 55.5%). Although most species belonging to the order *Burkholderiales* possess genomes of >4 Mb, sizes from 0.11 Mb (*Comamonadaceae* bacterium JGI 0001003-E14) to 11.66 Mb (*Caballeronia udeis* ES_PA-B7) have been recorded (data from https://www.ncbi.nlm.nih.gov/genome/browse). It is important to highlight that within the order *Burkholderiales*, species with smaller genome sizes are host-restricted microbial symbionts, whereas species that behave as plant or animal pathogens possess larger genomes (49). The G+C content within the order varies widely, ranging from 37.3% in *Taylorella equigenitalis* ATCC 35865^T to 72.9% in *Ideonella sakaiensis* 201-F6^T.

Pairwise nucleotide sequence similarity values based on 16S rRNA gene sequences ranged from 90.8 to 93.7%. The type strains with the greatest sequence similarity to CNM695-12 were *Leptothrix mobilis* Feox-1^T, *Aquabacterium commune* B8^T, *Aquabacterium citratiphilum* B4^T, *Schlegelella thermodepolymerans* K14^T and *Rubrivivax benzoatilyticus* JA2^T. The NJ tree based on 16S rRNA gene sequences revealed that the closest neighbour strains were *Ramlibacter tataouinensis* TTB310^T, *Comamonas humi* GAU11^T and *Comamonas serinivorans* SP-35^T, although position of strain CNM695-12 within family *Comamonadaceae* was uncertain due to the low bootstrap value (<70). Similar results were obtained using ML and MP methods (Fig. 1). In the same way, phylogenetic analysis using the type strains of type genera belonging to the order *Burkholderiales*, placed the novel strain on a branch next to *Comamonas terrigena* Hugh 247^T (Fig. S4). The ANI values based on genome data, and calculated using the EzBiocloud ANI calculator, varied between 67.6 and 69.6%, with the highest figures seen for *Schlegelella thermodepolymerans* K14^T, *Azohydromonas lata* DSM 1122^T, *Piscinibacter defluvii* SH-1^T, *Sphaerotilus natans* subsp *sulfidivorans* D-507 and *Roseateles depolymerans* KCTC 42856. In addition, the GGDC tool used to estimate DNA-DNA hybridization percentages (DDHe) returned the highest values for *Rubrivivax gelatinosus* IL144, *Sphaerotilus natans* subsp *sulfidivorans* D-507 and *Rhizobacter gummiphilus* NS21^T (≈25.5%). The smallest difference in G+C content calculated at the GGDC site, was with *Comamonas testosteroni* TK102 (9.5%). However, in all cases, the above values are below the threshold for species demarcation (95% for ANI, 70% for DDHe and >1% variation in G+C content). When AAIs (more

appropriate for comparisons between strains belonging to different taxa) were calculated, values ≈55% were obtained for *Schlegelella thermodepolymerans* K14^T, *Ramlibacter tataouinensis* TTB310^T, *Comamonas terrigena* Hugh 247^T and *Comamonas serinivorans* SP-35^T. According to a previous study (32), these percentages indicate strains to belong to different genera. The RAxML tree (Fig. 2) constructed from 107 proteins concatenated sequences also positioned strain CNM695-12 next to species of the genera *Comamonas* and *Ramlibacter* but on a clearly separate branch, forming a new lineage. The clade comprising unclassified *Burkholderiales* showed different groupings in NJ (Fig. 1) and RAxML (Fig. 2) trees. Discrepancies between phylogenetic trees inferred from a single gene and those inferred from concatenated proteins sequences have been previously described (50); events such as variations in evolution rate, convergent evolution or horizontal gene transfer have been proposed to explain these discordances (51). In order to estimate distances between proteins sequences involving in RAxML tree generation, MEGA7 software was used. These divergence values - expressed as number of amino acid substitutions per site from averaging over all sequence pairs - were calculated for each protein, for the set of ribosomal proteins and for that of non-ribosomal proteins. Non-ribosomal proteins sequences showed greater distance (mean: 0.32) than sequences belonging to ribosomal proteins (mean: 0.19). Individually, the sequences proteins with the highest divergence values (between 0.51 and 1.07) were those encoded by *sec*E, *dna*G, *rnc*, *coa*E, *his*S and *til*S genes, whereas ribosomal proteins S10, S12, L14, S11 and L19 presented the lowest distances (between 0.04 and 0.09). Table 2 shows the different phenotypic, genomic and genetic characteristics of strain CNM695-12 plus those of the type strains of related species.

Of the 3001 coding sequences predicted by Prokka v.1.12-beta software analysis, KEGG analysis associated 1569 (52.2%) to molecular functions defined in the KEGG Orthology. Despite the relatively small estimated genome size of the novel strain, the percentage of genes assigned to each pathway (based on the total number of protein-coding sequences) was very similar to those of related species (Table S2), except for genes involved in the biodegradation of compounds such as benzoate, toluene or naphthalene. The absence of genes associated with motility was also noted. However, these results indicate that the strain CNM695-12 has retained the essential genes of each metabolic pathway. Three biosynthetic gene clusters coding for secondary metabolites

were predicted by analysis with antiSMASH v.4.0 software: for bacteriocin (from nucleotide 1 to 8863), for terpene (from 195,283 to 216,131), and for nonribosomal peptide synthetase (NRPS) (from 350,234 to 396,302). No virulence factor was found and the organism was predicted to be a non-human pathogen by the Center for Genomic Epidemiology's software. ResFinder v.3.1 found no acquired resistance genes, but, in agreement with the phenotypic results and KEGG assignments, Prokka analysis predicted several genes related to bacterial resistance.

Carbapenems, third and fourth generation cephalosporins and quinolones are commonly used in the treatment of infections caused by non-fermenting Gram-negative bacilli such as *Acinetobacter* spp., *Pseudomonas aeruginosa* and *Burkholderia cepacia* complex, as well as by members of the family *Enterobacteriaceae*. However, ceftriaxone and levofloxacin, the first treatment received by the present patient, showed no *in vitro* activity against strain CNM 695-12. MIC values obtained by the Etest for beta-lactam antibiotics were: >256 µg/mL for ampicillin, piperacillin, piperacillin-tazobactam, cefepime, ceftriaxone, and aztreonam; 16 µg/mL for ceftazidime; 1 µg/mL for amoxicillin/clavulate; and 0.19 and 0.25 µg/mL for imipenem and meropenem respectively. This high resistance to beta-lactam antimicrobials (except for carbapenems and amoxicillin/clavulate) may be explained by the possession of a class A beta-lactamase showing 57.3% and 58% identity to beta-lactamases belonging to *Burkholderia dolosa* and *Klebsiella oxytoca* (as revealed by UniProt [https://www.uniprot.org/] and the Card database [https://card.mcmaster.ca/] searches respectively). Given the high MIC values obtained for ciprofloxacin, levofloxacin and moxifloxacin (>32 µg/mL), checks were made for possible replacements in the quinolone resistance-determining regions (QRDRs) of topoisomerases II (GyrA and GyrB) and IV (ParC and ParE). GyrA, GyrB, ParC and ParE sequences belonging to strain CNM695-12 aligned (as determined using the MAFFT server) with amino acid sequences of the related species *Comamonas terrigena* Hugh 247^{T} and *Ramlibacter tataouinensis* TTB310^{T}; with those of the type strain of the type genus in the order *Burkholderiales*, *Burkholderia cepacia* Ballard 717^{T}; and with those of *Escherichia coli* K12 (used as reference) (all sequences were obtained from https://www.ncbi.nlm.nih.gov/protein/). Table S3 shows the differences found among the QRDRs of GyrA, GyrB, ParC and ParE. One of the main changes - associated with increased resistance to quinolones (52) - was found in GyrA sequence: Ser83-Val

(encoded by GTT). Other replacements shared by the four aligned sequences were identified in GyrB, ParC and ParE, some of which are also associated with changes in the resistance profile to quinolones. These shared variations might be related to the taxonomic position of strains within the order *Burkholderiales* (Table 2 shows the percentage identities for the 16S rRNA, *gyr*A, *gyr*B, *par*C and *par*E genes of CNM695-12, *Comamonas terrigena* and *Ramlibacter tataouinensis*). Possession of the efflux pumps MexAB-OprM, MexXY-OprM and AcrAD-TolC (identified in the strain's genome) might also contribute to the resistance shown to several of the antimicrobials tested. Finally, strain CNM695-12 was susceptible to all other antibiotics analysed except for amikacin (MIC 96 µg/mL).

In conclusion, the results of the 16S rRNA gene and genome analyses reveal strain CNM695-12 to be a novel species within a new genus within the order *Burkholderiales* and close to family *Comamonadaceae*. The strain shares phenotypic characteristics with species belonging to this family such as their inability to produce acids from sugars but can be clearly differentiated from them by its lack of oxidase activity and via other biochemical tests. Further, the strain has a significantly smaller genome size and G+C content than most species in this family. The name *Saezia sanguinis* is proposed for the novel taxon.

Description of *Saezia* gen. nov.

Saezia (Sa.ez'i.a. N.L. fem. n. *Saezia* a tribute to the Spanish bacteriologist Juan A. Sáez-Nieto).

Cells are Gram-negative bacilli, aerobic, non-sporeforming and non-motile. Optimal growth occurs at 37^{o}C, pH 7 and with 0-1% NaCl. Catalase is produced but not oxidase. The most abundant fatty acids are C16:0, feature 8 (comprising C18:1 ω7c) and C18:1 ω9c. The major quinone is UQ8. The polar lipids present are phosphatidylglycerol, phosphatidylethanolamine and other aminophospholipids. The type species is *Saezia sanguinis*.

Description of *Saezia sanguinis* sp. nov.

Saezia sanguinis (san'gui.nis L. gen. n. *sanguinis* of blood)

Saezia sanguinis exhibits all the characteristics of genus *Saezia*. Growth occurs on BCA, MA, TSA and CPA after 48-72 h of incubation at 37^{o}C. Cells are approximately

382 0.5-0.8 µm in wide and 1.4-2.2 µm in length after 48 h of incubation on BCA at 37^{o}C.
383 Colonies grown on BCA under aerobic conditions reach 1-2 mm in diameter and are
384 punctiform, bright, opaque and convex with entire margins. Positive for catalase,
385 reduction of nitrate, assimilation of adipic acid and malate, and acetoin production.
386 Pyruvic acid methylester, acetic acid, sebacic acid, α-ketobutyric acid, α-ketoglutaric
387 acid, succinic acid, bromosuccinic acid, succinamic acid, L-aspartic acid, L-glutamic
388 acid and L-proline can be utilized as carbon sources, but not α-cyclodextrin, dextrin,
389 glycogen, tween 40, tween 80, N-acetyl-D-glucosamine, adonitol, L-arabinose, D-
390 arabitol, D-cellobiose, i-erythritol, D-fructose, L-fucose, D-galactose, gentibiose, α-D-
391 glucose, m-inositol, α-D-lactose, lactulose, maltose, D-mannitol, D-mannose, D-
392 melibiose, β-methyl-D-glucoside, D-psicose, D-raffinose, L-rhamnose, D-sorbitol,
393 sucrose, D-trehalose, turanose, xylitol, succinic acid mono-methyl ester, cis-aconitic
394 acid, citric acid, formic acid, D-galactonic acid lactone, D-galacturonic acid, D-gluconic
395 acid, glucosaminic acid, D-glucuronic acid, α-hydroxybutyric acid, β-hydroxybutyric
396 acid, ɣ-hydroxybutyric acid, p-hydroxyphenlyacetic acid, itaconic acid, α-ketovaleric
397 acid, D,L-lactic acid, malonic acid, propionic acid, quinic acid, D-saccharic acid,
398 glucuronamide, L-alaninamide, D-alanine, L-alanine, L-alanyl-glycine, L-asparagine,
399 glycyl-L-aspartic acid, glycyl-L-glutamic acid, L-histidine, hydroxy-L-proline, L-
400 leucine, L-ornithine, L-phenylalanine, L-pyroglutamic acid, D-serine, L-serine, L-
401 threonine, D,L-carnitine, ɣ-aminobutyric acid, urocanic acid, inosine, uridine,
402 thymidine, phenylethylamine, putrescine, 2-aminoethanol, 2,3-butanediol, glycerol, D-
403 L-α-glycerolphosphate, α-D-glucose-1-phosphate, D-glucose-6-phosphate. Alkaline
404 phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine
405 arylamidase, cystine arylamidase, acid phosphatase and naphthol-AS-BI-
406 phosphohydrolase activities were detected. Negative for oxidase, lipase (C14), trysin, α-
407 chymotripsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-
408 glucosidase, N-acetyl-β-glucosaminidase, α-mannosidase, or α-fucosidase.

409 The estimated genome size is 3.35 Mb and the G+C DNA content 52.4 mol%.

410 The type strain, isolated from human blood, is CNM695-12^{T} (=DSM 104959^{T} = CECT
411 9208^{T}).

412 **Supplementary material**

This article includes the following supplementary material: Table S1, Table S2, Table S3, Fig. S1, Fig. S2, Fig. S3 and Fig. S4.

Acknowledgments

The authors thank the Microscopy and Genomic Services of the *Instituto de Salud Carlos III* for help with TEM imaging and genome sequencing.

Conflicts of interest

None to report.

This work received no specific grant from any funding agency.

Ethical statement

The bacterial strain was sent to a public national reference laboratory for its identification. This study focused on the bacterial strain and no identifiable human data was used, therefore ethical approval was exempted. This work did not involve human or animal experimentation.

References

1. Garrity GM, Bell JA, Lilburn T. Order I. *Burkholderiales* ord. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (editors). *Bergey's Manual of Systematic Bacteriology* Vol 2: The *Proteobacteria*. 2th ed. New York: Springer Science & Business Media; 2006. pp. 575-759.

2. Hatayama K. *Comamonas humi* sp. nov., isolated from soil. *Int J Syst Evol Microbiol* 2014 ;64: 3976–3982. doi: 10.1099/ijs.0.067439-0.

3. Imai S, Yoshida R, Endo Y, Fukunaga Y, Yamazoe A, *et al*. *Rhizobacter gummiphilus* sp. nov., a rubber-degrading bacterium isolated from the soil of a botanical garden in Japan. *J Gen Appl Microbiol* 2013;59: 199–205. doi: 10.2323/jgam.59.199.

4. Gomila M, Bowien B, Falsen E, Moore ERB, Lalucat J. Description of *Roseateles aquatilis* sp. nov. and *Roseateles terrae* sp. nov., in the class *Betaproteobacteria*, and emended description of the genus *Roseateles*. *Int J Syst Evol Microbiol* 2008 Jan 1;58: 6–11. doi: 10.1099/ijs.0.65169-0.

5. Gilligan PH, Lum G, Vandamme AR, Whittier S. *Burkholderia*, *Stenotrophomonas*, *Ralstonia*, *Brevundimonas*, *Comamonas*, *Delftia*, *Pandoraea*, and *Acidovorax*. In: Murray PR, Baron EJ, Jorgensen JH, Pfaller MA, Yolken RH (editors). *Manual of Clinical Microbiology* vol 1. 8th ed. Washington, DC: ASM Press; 2003. p. 729–748.

6. Medina-Pascual MJ, Valdezate S, Villalón P, Garrido N, Rubio V, Saéz-Nieto JA. Identification, molecular characterisation and antimicrobial susceptibility of genomovars of the *Burkholderia cepacia* complex in Spain. *Eur J Clin Microbiol Infect Dis* 2012;31: 3385–96. doi: 10.1007/s10096-012-1707-6.

7. Tena D, Carranza R, Barberá JR, Valdezate S, Garrancho JM, *et al*. Outbreak of long-term intravascular catheter-related bacteremia due to *Achromobacter xylosoxidans* subspecies *xylosoxidans* in a hemodialysis unit. *Eur J Clin Microbiol Infect Dis* 2005; 24: 727–32. doi: 10.1007/s10096-005-0028-4.

8. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 1991;173: 697–703.

9. Baker GC, Smith JJ, Cowan DA. Review and re-analysis of domain-specific 16S primers. *J Microbiol Methods* 2003;55: 541–55. doi: 10.1016/j.mimet.2003.08.009.

10. *Interpretative criteria for identification of bacteria and funghi by DNA target sequencing; approved guide*. CLSI document MM18-A. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.

11. Pittman GW, Brumbley SM, Allsopp PG, O'Neill SL. Assessment of gut bacteria for a paratransgenic approach to control *Dermolepida albohirtum* larvae. *Appl Environ Microbiol* 2008;74: 4036–4043. doi: 10.1128/AEM.02609-07.

12. CLSI *Performance standards for antimicrobial susceptibility testing*. CLSI supplement M100. 29th ed. Wayne, PA: Clinical and Laboratory standard Institute; 2019.

13. Sasser M. Identification of bacteria by gas chromatography of cellular fatty acids. In: *MIDI technical note 101*. Newark, DE: MIDI Inc; 1990.

14. *Sherlock Microbial Identification version 6.1*. MIS system operating manual. Newark, DE: MIDI Inc; 2008.

15. Tindall BJ. A comparative study of lipid composition of *Halobacterium saccharovorum* from various sources. *Sys Appl Microbiol* 1990;13: 128–130.

16. Tindall BC, Sikorski J, Smibert R, Kreig N. Phenotypic characterization and the principles of comparative systematics. In: Reddy CA, Beveridge T, Brenak J, Marzluf G, Schmidt T, Snyder L (editors). *Methods for General and Molecular Microbiology*. 3rd ed. Washington, DC: ASM Press; 2007 pp. 330–393. doi: 10.1128/9781555817497.ch15.

17. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y *et al*. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67: 1613–1617. doi: 0.1099/ijsem.0.001755.

18. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser*. 1999;41: 95–98.

19. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 2017; 1-7. doi/10.1093/bib/bbx108.

20. Tamura K. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Mol Biol Evol* 1992;9: 678-687.

21. Fitch WM. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool.* 1971;20: 406–416. doi: 10.2307/2412116.

22. Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 1981;17: 368–376.

23. Kumar S, Stecher G, Tamura K. Mega7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33: 1870–1874. doi: 10.1093/molbev/msw054.

24. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30: 2114–2120. doi: 10.1093/bioinformatics/btu170.

25. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M *et al*. Spades: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19: 455-477. doi: 10.1089/cmb.2012.0021.

26. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;29: 1072–1075. doi: 10.1093/bioinformatics/btt086.

27. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30: 2068–2069. doi: 10.1093/bioinformatics/btu153.

28. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14: 60. doi: 10.1186/1471-2105-14-60.

29. Yoon S-H, Ha S, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;110: 1281–1286. doi: 10.1007/s10482-017-0844-4.

30. Rodriguez-R LM, Konstantinidis KT. Bypassing cultivation to identify bacterial species. *Microbe* 2014;9: 111–118. doi: 10.1128/microbe.9.111.1.

31. Dupont CL, Rusch DB, Yooseph S, Lombardo M-J, Alexander Richter R *et al*. Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J* 2012;6: 1186–1199. doi: 10.1038/ismej.2011.189.

32. Ankenbrand MJ, Keller A. bcgTree: automatized phylogenetic tree building from bacterial core genomes. *Genome* 2016;59: 783–791. doi: 10.1139/gen-2015-0175.

33. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30: 1312–1313. doi: 10.1093/bioinformatics/btu033.

34. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 2007;35: W182–W185. doi: 10.1093/nar/gkm321.

35. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ *et al*. antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* 2017;45: W36–W41. doi: 10.1093/nar/gkx319.

36. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S *et al*. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 2012;67: 2640–2644. doi: 10.1093/jac/dks261.

37. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS *et al*. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *J Clin Microbiol* 2014;52: 1501–1510. doi: 10.1128/JCM.03617-13.

38. Cosentino S, Voldby Larsen M, Møller Aarestrup F, Lund O. PathogenFinder - distinguishing friend from foe using bacterial whole genome sequence data. *PLoS ONE* 2013;8: e77302. doi:10.1371/journal.pone.0077302.

39. Zhu D, Xie C, Huang Y, Sun J, Zhang W. Description of *Comamonas serinivorans* sp. nov., isolated from wheat straw compost. *Int J Syst Evol Microbiol* 2014;64: 4141–4146. doi: 10.1099/ijs.0.066688-0.

40. Chen W-M, Cho N-T, Yang S-H, Arun AB, Young C-C *et al*. *Aquabacterium limnoticum* sp. nov., isolated from a freshwater spring. *Int J Syst Evol Microbiol* 2012;62: 698–704. doi: 10.1099/ijs.0.030635-0.

41. Hiraishi A, Hoshino Y, Satoh T. *Rhodoferax fermentans* gen. nov., sp. nov., a phototrophic purple nonsulfur bacterium previously referred to as the "*Rhodocyclus gelatinosus*-like" group. *Arch Microbiol* 1991;155: 330-336.

42. Spring S, Kampfer P, Ludwig W, Schleifer K-H. Polyphasic characterization of the genus *Leptothrix*: new descriptions of *Leptothrix mobilis* sp. nov. and *Leptothrix discophora* sp. nov. nom. rev. and emended description of *Leptothrix cholodnii* emend. *Syst Appl Microbiol* 1996;19: 634–643.

43. Takeda M, Kamagata Y, Ghiorse WC, Hanada S, Koizumi J. *Caldimonas manganoxidans* gen. nov., sp. nov., a poly(3-hydroxybutyrate)-degrading, manganese-oxidizing thermophile. *Int J Syst Evol Microbiol* 2002;52: 895–900. doi: 10.1099/ijs.0.02027-0.

44. Elbanna K, Lütke-Eversloh T, Van Trappen S, Mergaert J, Swings J *et al Schlegelella thermodepolymerans* gen. nov., sp. nov., a novel thermophilic bacterium that degrades poly(3-hydroxybutyrate-co-3-mercaptopropionate). *Int J Syst Evol Microbiol* 2003;53: 1165–1168. doi: 10.1099/ijs.0.02562-0.

45. Chou Y-J, Sheu S-Y, Sheu D-S, Wang J-T, Chen W-M. *Schlegelella aquatica* sp. nov., a novel thermophilic bacterium isolated from a hot spring. *Int J Syst Evol Microbiol* 2006;56: 2793–2797. doi: 10.1099/ijs.0.64446-0.

46. Hartig C, Loffhagen N, Harms H. Formation of *trans* fatty acids is not involved in growth-linked membrane adaptation of *Pseudomonas putida*. *Appl Environ Microbiol* 2005;71:1915–1922. doi: 10.1128/AEM.71.4.1915-1922.2005.

47. Collins MD, Jones D. Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implications. *Microbiol Rev* 1981;45: 316-354.

48. Stackebrandt E, Verbarg S, Fruhling A, Busse H-J, Tindall BJ. Dissection of the genus *Methylibium*: reclassification of *Methylibium fulvum* as *Rhizobacter fulvus* comb. nov., *Methylibium aquaticum* as *Piscinibacter aquaticus* gen. nov., comb. nov. and *Methylibium subsaxonicum* as *Rivibacter subsaxonicus* gen. nov., comb. nov. and emended descriptions of the genera *Rhizobacter* and *Methylibium*. *Int J Syst Evol Microbiol* 2009;59:2552–2560. doi: 10.1099/ijs.0.008383-0.

562 49. Voronina OL, Kunda MS, Ryzhova NN, Aksenova EI, Semenov AN *et al*. The variability of the
563 order *Burkholderiales* representatives in the healthcare units. *BioMed Res Int* 2015: 1–9.
564 doi: 10.1155/2015/680210.

565 50. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A *et al*. A standardized bacterial
566 taxonomy based on genome phylogeny substantially revises the tree of life. *Nat*
567 *Biotechnol*.2018;36: 996–1004. doi: 10.1038/nbt.4229.

568 51. Wu D, Jospin G, Eisen JA. Systematic identification of gene families for use as '"markers"'
569 for phylogenetic and phylogeny-driven ecological studies of Bacteria and Archaea and
570 their major subgroups. *PLOS ONE* 2013;8: e77033. doi:10.1371/journal.pone.0077033.

571 52. Cullen ME, Wyke AW, Kuroda R, Fisher LM. Cloning and characterization of a DNA gyrase A
572 gene from *Escherichia coli* that confers clinical resistance to 4-quinolones. *Antimicrob*
573 *Agents Chemother* 1989;33: 886–894. doi: 10.1128/AAC.33.6.886.

574 53. Heulin T, Barakat M, Christen R, Lesourd M, Sutra L *et al*. *Ramlibacter tataouinensis* gen.
575 nov., sp. nov., and *Ramlibacter henchirensis* sp. nov., cyst-producing bacteria isolated from
576 subdesert soil in Tunisia. *Int J Syst Evol Microbiol* 2003;53: 589–594. doi:
577 10.1099/ijs.0.02482-0.

578 54. Lee HJ, Lee SH, Lee S-S, Lee JS, Kim Y *et al*. *Ramlibacter solisilvae* sp. nov., isolated from
579 forest soil, and emended description of the genus *Ramlibacter*. *Int J Syst Evol Microbiol*
580 2014;64: 1317–1322. doi: 10.1099/ijs.0.058396-0.

581 55. Kalmbach S, Manz W, Wecke J, Szewzyk U. *Aquabacterium* gen. nov., with description of
582 *Aquabacterium citratiphilum* sp. nov., *Aquabacterium parvum* sp. nov. and *Aquabacterium*
583 *commune* sp. nov., three in situ dominant bacterial species from the Berlin drinking water
584 system. *Int J Syst Bacteriol* 1999;49: 769–777.

585 56. Suyama T, Shigematsu T, Takaichit S, Nodasakaf Y, Tokiwa Y *et a*l. *Roseateles*
586 *depolymerans* gen. nov., sp. nov., a new bacteriochlorophylla-containing obligate aerobe
587 belonging to the β-subclass of the *Proteobacteria*. *Int J Syst Bacteriol* 1999;49: 449-457.

588 57. Rakshak K, Ravinder K, Nupur, Srinivas TNR, Anil Kumar P. *Caldimonas meghalayensis* sp.
589 nov., a novel thermophilic *betaproteobacterium* isolated from a hot spring of Meghalaya in
590 northeast India. *Antonie Van Leeuwenhoek* 2013;104: 1217–1225. doi: 10.1007/s10482-
591 013-0043-x.

592 58. Kang W, Kim PS, Hyun D-W, Lee J-Y, Kim HS *et al*. *Comamonas piscis* sp. nov., isolated from
593 the intestine of a Korean rockfish, *Sebastes schlegelii*. *Int J Syst Evol Microbiol* 2016;66:
594 780–785. doi: 10.1099/ijsem.0.000790.

595

596

Table 1. Differential characteristics of strain CNM695-12 and the type strain of phylogenetically related species within order *Burkholderiales*.

+,Positive; -, Negative; ND, data no available. SF8: Summed feature 8 (comprising C18:1 ω7c); SF3: Summed feature 3 (comprising C16:1 ω7c and/or C16:1 ω6c). PG: Phosphatidylglycerol; PE: Phosphatidylethanolamine; DPG: diphosphatidylglycerol; PS: phosphatidylserine; GL: glycolipid; PNL: unidentified aminophospholipid; PL: unidentified phospholipids

	CNM695-12	*Comamonas humi* GAU11 ^{T} (2)	*Comamonas serinivorans* SP-35^{T} (39)	*Ramlibacter tataouinensis* TTB310^{T} (53,54)	*Schlegelella thermodepolymerans* K14^{T} (45)	*Aquabacterium commune* B8^{T} (40,55)	*Rubrivivax gelatinosus* DSM 1709^{T} (41)	*Leptothrix mobilis* Feox-1^{T} (42)	*Rhizobacter gummiphilus* NS21^{T} (3,48)	*Roseateles depolymerans* 61A^{T} (4,56)	*Caldimonas manganoxidans* HS^{T} (43,57)
Isolation source	Blood	Soil	Wheat straw compost	Meteorite fragment	Activated sludge sample	Drinking water system biofilms	Mud	Sediment of freshwater lake	Soil	River water	Hot spring
Cell shape (size, µm)	Straight rods (0.5-0.8 by 1.4-2.2)	Rods (0.5-0.9 by 0.8-4.2)	Short rods (0.6-0.8 by 0.8-2)	Rods (0.2 by 3) or cyts (0.8 by 0.9)	Rods (0.5-0.6 by 1.0-2.8)	Rods (0.5 by 2-4)	Straight-slightly curved rods (0.4-0.7 by 1-3)	Straight rods (0.6-0.8 by 1.5-12)	Straight rods (0.6-0.7 by 1.2-1.5)	Straight rods (0.5 by 2)	Rods, (0.5-0.7 by 2.2-3.5)
Gram stain	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative
Flagella	-	-	-	-	+: Polar monotrichous	+: Single polar	+: Single polar	+: Single polar	+: Single polar	+: Several polar	+: Single polar
Growth temperature ºC (optimum)	20-42 (37)	15-30 (25-30)	20-37 (30)	(30)	45-50 (50)	6-34	25-30	10-37	10-35 (15-30)	5-43 (35)	(50)
Growth pH (optimum)	5-9 (7)	5-10 (6-8)	5-9 (7)	(7.5)	6–9 (7)	6.5-9.5	5-9	6.5-8.5	6-9 (7-8)	5-8 (6.5)	(8-9)
Growth NaCl range (% w/v) (optimum)	0-4 (0-1)	0-3	0-3	ND	ND	0-0.4	<1	ND	≤0.7	ND	0-2
Catalase production	+	+	+	+	+	-	ND	ND	+	-	Weakly +
Oxidase production	-	+	+	+	+	+	ND	+	+	+	+
Relation to O_2	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic, facultatively anaerobic	Aerobic	Facultatively anaerobic	Aerobic	Aerobic
Major fatty acids (>10%)	$C_{16:0}$, SF8, $C_{18:1}$ ω9c	SF3, $C_{16:0}$, $C_{18:1}$ ω7c	$C_{16:0}$, SF3, $C_{17:0}$ cyclo,	SF3, $C_{16:0}$, $C_{12:0}$	$C_{16:0}$, $C_{17:0}$ cyclo	SF3, $C_{16:0}$	$C_{16:0}$,, $C_{16:1}$ ω7c , $C_{18:1}$	$C_{16:1}$ ω9c, $C_{16:0}$	SF3, $C_{16:0}$, $C_{17:0}$ cyclo, $C_{18:1}$ ω7c	$C_{16:1}$ ω7c, $C_{16:0}$, $C_{18:1}$ ω7c/Rt/9t	$C_{16:0}$, $C_{16:1}$, $C_{18:1}$
Major polar lipids	PG, PE, PNL	DPG, PE, PG	DPG, PE, PG	ND	ND	PE, PG, PS, DPG, PL	ND	ND	DPG, PE, PG	ND	DPG, PE, GL, PL
Major ubiquinone	UQ8	UQ8	UQ8	ND	ND	UQ8	UQ8, MK8	UQ8	UQ8	UQ8	UQ8
G+C DNA content (mol%)*	52.4	68.2	63.1	69.6	70	66	71.9	68	70.8	66.3	62.2

	CNM695-12	*Comamonas humi* GAU11 ^{T} (2)	*Comamonas serinivorans* SP-35^{T} (39)	*Ramlibacter tataouinensis* TTB310^{T} (53,54)	*Schlegelella thermodepolymerans* K14^{T} (45)	*Aquabacterium commune* B8^{T} (40,55)	*Rubrivivax gelatinosus* DSM 1709^{T} (41)	*Leptothrix mobilis* Feox-1^{T} (42)	*Rhizobacter gummiphilus* NS21^{T} (3,48)	*Roseateles depolymerans* 61A^{T} (4,56)	*Caldimonas manganoxidans* HS^{T} (43,57)
Genome size (Mb)*	3.35	ND	4.52	4.07	3.83	3.56 †	5.04 †	4.91†	6.40	5.66	3.53

Data for related strains are taken from indicated references.

* Data from www.ncbi.nlm.nih.gov/genome.

†Information not available for type strains. Data from: *Aquabacterium* sp. strain UBA666, *R. gelatinosus* strain IL144 and *L. cholodnii* strain sp-6.

Table 2. Differential phenotypic, genomic and genetic characteristics of strain CNM695-12 and type strains of related species belonging to family *Comamonadaceae*. Data for related strains are taken from indicated references. +, Positive; -, Negative; ND, data no available.

	CNM695-12	*Comamonas terrigena* Hugh 247^T (2,58)	*Ramlibacter tataouinensis* TTB310^T (54)
Phenotypic Characteristics			
Oxidase activity	-	+	+
API ZYM:			
Cystine arylamidase	+	-	w
API 20 NE:			
Indole production	-	ND	+
Hydrolisis esculin	-	ND	+
Assimilation:			
Potassium gluconate	-	+	-
Adipic acid	+	+	-
Malate	+	-	-
BIOLOG GN2:			
Tween 40	-	+	-
Succinic Acid Mono-Methyl Ester	-	+	-
α-Hydroxybutyric Acid	-	+	-
β-Hydroxybutyric Acid	-	+	+
ɣ-Hydroxybutyric Acid	-	-	+
α-Ketobutyric Acid	+	+	-
α-Ketoglutaric Acid	+	-	-
α-Ketovaleric Acid	-	+	-
D,L-Lactic Acid	-	+	+
Propionic Acid	-	+	+
Sebacic Acid	+	+	-
Succinic Acid	+	-	-
Bromosuccinic Acid	+	-	-
Succinamic Acid	+	-	-
L-Aspartic Acid	+	-	-
L-Glutamic Acid	+	+	-
L-Leucine	-	+	-
L-Proline	+	+	-
L-Pyroglutamic acid	-	+	-
L-Threonine	-	+	-
Genomic and Genetic Characteristics *			
Genome size (Mb)	3.35	4.63	4.07
G+C content (%)	52.4	65.1	69.6
Coding sequences genes	3001	4000	3843
rRNA	3	4	3
tRNA	51	61	43
ANI (%) †	100	69.2	69.2
GGDC (%DDH estimated) ‡	0 (100)	0.2337 (18.2)	0.2030 (21.6)
AAI (%) §	100	54.5	54.6
% Identity with CNM695-12:			
16S rRNA	100	91.1	92.0
***gyr*A**	100	83.3	83.5
***gyr*B**	100	69.0	70.5
***par*C**	100	72.0	73.8
***par*E**	100	77.7	80.3

*Data from https://www.ncbi.nlm.nih.gov/genome/. NCBI reference sequences: NC_015677.1 (*R. tataouinensis* TTB310^T) and NZ_BCNR00000000.1 (*Comamonas terrigena* Hugh 247^T)

† ANI values were calculated at https://www.ezbiocloud.net/tools/ani

‡ Distances between genomes (GGDC) and DDH estimated were calculated at http://ggdc.dsmz.de

§ AAI values were calculated at http://enve-omics.ce.gatech.edu/

¶ % identity calculated after multiple alignments of sequences at MAFFT v7 and using MegAlign (DNAStar package 12.3.1).

Table S1. Genes used to generate the concatenated amino acid sequences for phylogenetic inference by RAxML of the strain CNM695-12 and related strains within the order *Burkholderiales* (31).

Gene symbol	Description	Accession/domain in TIGRFAM/ Pfam databases	Position in the concatenated amino acid sequence
***ala*S**	Alanyl-tRNA synthetase	TIGR00344	7571-8432
***arg*S**	Arginyl- tRNA synthetase	PF00750.15	2037-2566
***asp*S**	Aspartyl- tRNA synthetase	TIGR00459	14893-15483
***cgt*A**	Obg family GTPase CgtA	TIGR02729	33435-33771
***coa*E**	Dephospho-CoA kinase	TIGR00152	5963-6138
***cys*S**	Cysteinyl-tRNA synthetase	TIGR00435	13761-14200
***dna*A**	Chromosomal replication initiator protein DnaA	TIGR00362	8433-8833
***dna*G**	DNA primase	TIGR01391	26732-27202
***dna*K**	Chaperone protein DnaK	TIGR02350	30658-31280
***dna*N**	DNA polymerase III, beta subunit	TIGR00663	19208-19571
***dna*X**	DNA polymerase III, gamma and tau subunits	TIGR02397	32666-33125
***eng*A**	Ribosome-associated GTPase EngA	TIGR03594	33977-34408
era	GTP-binding protein Era	TIGR00436	14201-14482
ffh	Signal recognition particle protein	TIGR00959	20262-20707
fmt	Methionyl-tRNA formyltransferase	TIGR00460	15484-15772
frr	Ribosome recycling factor	TIGR00496	17721-17906
***fts*Y**	Signal recognition particle-docking protein FtsY	TIGR00064	4308-4613
***gly*S**	Glycyl-tRNA synthetase	TIGR00388	8834-9130
gmk	Guanylate kinase	TIGR03263	33772-33976
***grp*E**	Co-chaperone GrpE	PF01025.15	2567-2709
***gyr*A**	DNA gyrase, subunit A	TIGR01063	24734-25579
***gyr*B**	DNA gyrase, subunit B	TIGR01059	23936-24733
***his*S**	Histidyl-tRNA synthetase	TIGR00442	14483-14892
***ile*S**	Isoleucyl-tRNA synthetase	TIGR00392	9131-10045
***inf*B**	Translation initiation factor IF-2	TIGR00487	16895-17720
***inf*C**	Translation initiation factor IF-3	TIGR00168	6489-6643
***ksg*A**	Dimethyladenosine transferase	TIGR00755	19572-19802
***lep*A**	GTP-binding protein LepA	TIGR01393	27203-27804
***leu*S**	leucyl-tRNA synthetase	TIGR00396	10046-10881
***lig*A**	DNA ligase, NAD-dependent	TIGR00575	17907-18539
***mnm*A**	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	TIGR00420	12508-12845
***mra*W**	MraW methylase family	PF01795.15	2710-3007
***nus*A**	Transcription termination factor NusA	TIGR01953	27948-28435
***nus*G**	Transcription termination/ antitermination factor NusG	TIGR00922	20008-20180
pgk	Phosphoglycerate kinase	PF00162.15	1-397
***phe*S**	Phenylalanyl-tRNA synthase α subunit	TIGR00468	15773-16120
***phe*T**	Phenylalanyl-tRNA synthase β subunit	TIGR00472	16121-16894
***prf*A**	Peptide chain release factor 1	TIGR00019	3293-3642
***pro*S**	Prolyl-tRNA synthetase	TIGR00409	10882-11458
***pyr*G**	CTP synthetase	TIGR00337	7036-7570
***rec*A**	RecA protein	TIGR02012	28436-28768
***rfb*A**	Ribosome-binding factor A	TIGR00082	4614-4728
rnc	Ribonuclease III	TIGR02191	30450-30657
***rpl*A**	Ribosomal protein L1	TIGR01169	26227-26457
***rpl*B**	Ribosomal protein L2	TIGR01171	26458-26731
***rpl*C**	Ribosomal protein L3	PF00297.18	676-889
***rpl*D**	Ribosomal protein L4	PF00573.18	1836-2036
***rpl*E**	Ribosomal protein L5	PF00281.15	498-675
***rpl*F**	Ribosomal protein L6	PF00347.19	890-1066
***rpl*I**	Ribosomal protein L9	TIGR00158	6139-6285
***rpl*J**	Ribosomal protein L10	PF00466.16	1666-1835
***rpl*K**	Ribosomal protein L11	TIGR01632	27805-27947
***rpl*L**	Ribosomal protein L7/L12	TIGR00855	19883-20007
***rpl*M**	Ribosomal protein L13	TIGR01066	25580-25721
***rpl*N**	Ribosomal protein L14	TIGR01067	25722-25843
***rpl*O**	Ribosomal protein L15	TIGR01071	25844-25986
***rpl*P**	Ribosomal protein L16	TIGR01164	26090-26226
***rpl*Q**	Ribosomal protein L17	TIGR00059	3876-4004

rplR	Ribosomal protein L18	TIGR00060	4005-4122
rplS	Ribosomal protein L19	TIGR01024	23139-23254
rplT	Ribosomal protein L20	TIGR01032	23515-23632
rplU	Ribosomal protein L21	TIGR00061	4123-4223
rplV	Ribosomal protein L22	TIGR01044	23633-23741
rplW	Ribosomal protein L23	PF00276.16	398-497
rplX	Ribosomal protein L24	TIGR01079	25987-26089
rpmA	Ribosomal protein L27	TIGR00062	4224-4307
rpmB	Ribosomal protein L28	TIGR00009	3153-3228
rpmC	Ribosomal protein L29	TIGR00012	3229-3292
rpmF	Ribosomal protein L32	TIGR01031	23455-23514
rpmH	Ribosomal protein L34	TIGR01030	23411-23454
rpmI	Ribosomal protein L35	TIGR00001	3008-3074
rpoA	DNA-directed RNA polymerase, α subunit	TIGR02027	30133-30449
rpoB	DNA-directed RNA polymerase, β subunit	TIGR02013	28769-30132
rpoC	DNA-directed RNA polymerase, β′ subunit	TIGR02386	31281-32665
rpsB	Ribosomal protein S2	TIGR01011	22514-22760
rpsC	Ribosomal protein S3	TIGR01009	22281-22513
rpsD	Ribosomal protein S4	TIGR01017	22761-22967
rpsE	Ribosomal protein S5	TIGR01021	22968-23138
rpsF	Ribosomal protein S6	TIGR00166	6372-6488
rpsG	Ribosomal protein S7	TIGR01029	23255-23410
rpsH	Ribosomal protein S8	PF00410.15	1280-1410
rpsI	Ribosomal protein S9	PF00380.15	1150-1279
rpsJ	Ribosomal protein S10	TIGR01049	23742-23844
rpsK	Ribosomal protein S11	PF00411.15	1411-1544
rpsL	Ribosomal protein S12	TIGR00981	22157-22280
rpsM	Ribosomal protein S13	PF00416.18	1545-1665
rpsO	Ribosomal protein S15	TIGR00952	20181-20261
rpsP	Ribosomal protein S16	TIGR00002	3075-3152
rpsQ	Ribosomal protein S17	PF00366.16	1067-1149
rpsR	Ribosomal protein S18	TIGR00165	6286-6371
rpsS	Ribosomal protein S19	TIGR01050	23845-23935
rpsT	Ribosomal protein S20	TIGR00029	3643-3739
secA	Preprotein translocase, SecA subunit	TIGR00963	20708-21602
secE	Preprotein translocase, SecE subunit	TIGR00964	21603-21722
secG	Preprotein translocase, SecG subunit	TIGR00810	19803-19882
secY	Preprotein translocase, SecY subunit	TIGR00967	21723-22156
serS	Seryl-tRNA synthetase	TIGR00414	11459-11880
smpB	SmpB protein	TIGR00086	4729-4873
thrS	Threonyl-tRNA synthetase	TIGR00418	11881-12507
tig	Trigger factor	TIGR00115	5238-5669
tilS	tRNA(Ile)-lysidine synthetase	TIGR02432	33126-33434
tsf	Translation elongatio factor Ts	TIGR00116	5670-5962
tyrS	Tyrosyl-tRNA synthetase	TIGR00234	6644-7035
uvrB	Excinuclease ABC, B subunit	TIGR00631	18540-19207
valS	Valyl-tRNA synthetase	TIGR00422	12846-13760
ybeY	Conserved hypothetical protein YbeY	TIGR00043	3740-3875
ychF	GTP-binding protein YchF	TIGR00092	4874-5237

Table S2. Number of genes assigned to each molecular function defined in the KEGG Orthology (KO) for strain CNM695-12 and related species *Comamonas terrigena* Hugh 247^T and *Schlegelella thermodepolymerans* K14^T . In brackets, percentage of genes assigned, based on total number of coding-protein sequences.

Categories	Subcategories	CNM695-12	*C. terrigena*	*S. thermodepolymerans*
Metabolism	Carbohydrate metabolism	182 (6.1%)	222 (5.6%)	241 (6.6%)
	Energy metabolism	134 (4.5%)	155 (3.9%)	149 (4.2%)
	Lipid metabolism	49 (1.6%)	62 (1.6%)	60 (1.7%)
	Nucleotide metabolism	59 (2%)	75 (1.9%)	71 (2%)
	Amino acid metabolism	181 (6%)	231 (5.8%)	229 (6.5%)
	Metabolism of other aminoacid	36 (1.2%)	49 (1.2%)	50 (1.4%)
	Glycan biosynthesis and metabolism	37 (1.2%)	32 (0.8%)	36 (1.02%)
	Metabolism of cofactors and vitamins	105 (3.5%)	136 (3.4%)	128 (3.6%)
	Metabolism of terpenoids and polyketides	26 (0.9%)	16 (0.4%)	35 (1%)
	Biosynthesis of other secondary metabolites	31 (1%)	31 (0.8%)	37 (1.1%)
	Xenobiotics biodegradation and metabolism	39 (1.3%)	56 (1.4%)	119 (3.4%)
Genetic information processing	Transcription	5 (0.2%)	5 (0.1%)	4 (0.1%)
	Translation	82 (2.7%)	83 (2.1%)	83 (2.4%)
	Folding, sorting and degradation	39 (1.3%)	41 (1%)	41 (1.2%)
	Replication and repair	70 (2.3%)	69 (1.7%)	73 (2.1%)
Environmental information processing	Membrane transport	86 (2.9%)	95 (2.4%)	122 (3.5%)
	Signal transduction	71 (2.4%)	96 (2.4%)	95 (2.7%)
Cellular processes	Transport and catabolism	6 (0.2%)	7 (0.2%)	7 (0.2%)
	Cell growth and death	19 (0.6%)	24 (0.6%)	24 (0.7%)
	Cellular community eukaryotes	1 (-)	-	-
	Cellular community prokaryotes	63 (2.1%)	73 (1.8%)	106 (3%)
	Cell motility	2 (0.1%)	51 (1.3%)	48 (1.4%)
Organismal System		32 (1.1%)	53 (1.3%)	54 (1.5%)
Human diseases	Infectious diseases: bacterial	15 (0.5%)	17 (0.4%)	14(0.4%)
	Other diseases	25 (0.8%)	64 (1.6%)	63 (1.8%)
	Drug resistance: antimicrobial	25 (0.8%)	24 (0.6%)	24 (0.7%)
	Drug resistance: antineoplasic	7 (0.2%)	8 (0.2%)	8 (0.2%)
Not included in other categories and uncharacterized proteins		142 (4.7%)	385 (9.6%)	120 (3.4%)
Number of coding sequences		3001	4000	3514
Percentage number KO assigned		52.3% (1569)	54% (2160)	58.1% (2041)

Table S3. Alignments of those positions with changes from QRDRs of GyrA, GyrB, ParC and ParE from CNM695-12, *Escherichia coli* K12, *Comamonas terrigena* Hugh 247^T, *Ramlibacter tataouinensis* TTB310^T and *Burkholderia cepacia* Ballard 717^T. Numeration is according to *E. coli* sequences.

	GyrA (QRDR: from 67 to 106aa)										GyrB (QRDR: from 426 to 464aa)							
Positions	69	74	**83[1]**	84	85	89	92	98	100	103	437	440	**447[1]**	450	458	459	461	463
E. coli	V	I	**S**	A	V	I	M	L	Y	V	N	N	**K**	I	F	D	M	S
CNM695-12	I	M	**V**	*	I	V	*	*	H	I	D	F	**R**	*	Y	E	L	*
C. terrigena	I	*	*****	*	*	*	*	*	H	*	D	F	**R**	*	A	E	L	*
R. tataounensis	I	*	**Q**	S	*	*	L	M	H	*	D	F	**R**	*	Y	E	L	T
B. cepacia	I	*	**T**	*	*	*	*	*	*	I	D	F	**R**	V	Y	*	L	*

	ParC (QRDR: from 64-103aa)												
Positions	**73**	**74**	**80[1]**	**81**	**82**	**84[1]**	**86**	**88**	**89**	**92**	**94**	**95**	**100**
E. coli	K	Y	**S**	A	C	**E**	M	L	M	P	S	Y	V
CNM695-12	R	F	**Q**	*	A	**D**	L	R	*	D	*	Q	*
C. terrigena	R	F	**Q**	S	A	**D**	L	R	*	D	N	Q	*
R.tataounensis	R	F	**Q**	*	A	**D**	L	R	*	D	*	Q	I
B. cepacia	*	*	**Q**	S	A	**D**	L	R	L	D	*	L	I

	ParE (QRDR: from 420 to 458)																	
Positions	**428**	**429**	**431**	**432**	**434**	**437**	**438**	**441[1]**	**444**	**447**	**450**	**451**	**452**	**454**	**455**	**456**	**457**	**458**
E. coli	Q	A	D	R	Y	I	M	**K**	I	T	V	S	S	E	V	L	A	S
CNM695-12	M	G	N	K	*	V	L	**R**	*	A	*	E	Q	R	L	F	G	N
C. terrigena	M	G	*	K	T	V	L	**R**	V	*	*	D	R	R	L	F	*	N
R.tataounensis	M	G	*	K	C	V	L	**R**	V	*	*	E	R	R	L	F	*	N
B. cepacia	M	G	*	K	*	*	L	**R**	V	*	T	E	R	R	L	F	*	N

[1] Marked positions have been described as being involved in major changes that increase resistance to quinolones.

Figure legends

Fig. 1. Neighbour-joining phylogenetic tree based on 16S rRNA gene sequences (1393 alignment positions) showing relationships between strain CNM695-12 and closely related taxa within the order *Burkholderiales*. Filled circles indicate those nodes that were also recovered using maximum-parsimony and maximum-likelihood methods, whereas open circles indicate nodes also recovered using the maximum-likelihood algorithm. Bootstrap values ≥70% (expressed as percentages of 1000 replications) are shown above (NJ) and/or below nodes (MP). *Aquaspirillum serpens* DSM 68^T was used as an outgroup. Bar, 0.02 substitutions per nucleotide position. GenBank accession numbers or NCBI reference sequences are given in brackets.

Fig. 2. RAxML phylogenetic tree inferred from the concatenated amino acid sequences of 107 proteins (Table S1 shows encode genes) and involving CNM695-12 and 19 related strains within order *Burkholderiales*. Node labels indicate bootstrap confidence levels (values ≥70% are shown). *Aquaspirillum serpens* DSM 68^T was used as an outgroup. Bar, 0.05 substitutions per amino acid position. NCBI reference sequences are given in brackets.

Fig. S1 Mass spectrum of strain CNM695-12 obtained by using the VITEK MS automated mass spectrometry system (bioMérieux). The horizontal axis shows mass/charge ratio (Da) and the vertical axis shows the relative intensities of ions (%). Peaks >12,500 Da are not shown due to their low intensity.

Fig. S2. Transmission electron micrographs of cells of strain CNM695-12 non-stained (a) and negatively stained with 2% phosphotungstic acid (b) showing overall morphology and pili-like structures. Ultrathin sections of cells of strain CNM695-12 stained with 2% phosphotungstic acid (c).

Fig. S3. Chromatogram of polar lipids from strain CNM695-12 obtained by thin layer chromatography.

Fig. S4. Neighbour-joining phylogenetic tree based on 16S rRNA gene sequences (1397 alignment positions) showing relationships between CNM695-12 and type strains of type genera within order *Burkholderiales*. Bootstrap values ≥70% (expressed as percentages of 1000 replications) are shown. Bar, 0.01 substitutions per nucleotide

position. GenBank accession numbers or NCBI reference sequences are given in brackets.