‘Cellulomonas timonensis’ sp. nov., taxonogenomics description of a new bacterial species isolated from human gut

S. Ndongo¹, F. Bittar², M. Beye¹, C. Robert¹, F. Di Pinto¹, P.-E. Fournier², D. Raoult¹ and J.-C. Lagier¹
¹) Aix-Marseille Université, IRD, AP-HM, IHU-Méditerranée Infection, UMR Microbes Evolution Phylogénie (MEPHI), Marseille, France and 2) Aix Marseille Univ, IRD (Dakar, Marseille, Papeete), AP-HM, IHU-Méditerranée Infection, UMR Vecteurs - Infections Tropicales et Méditerranéennes (VITROME), Marseille, France

Abstract

‘Cellulomonas timonensis’ sp. nov. strain sn7T is a new species within the Cellulomonas genus. We present the main phenotypic characteristics and provide a complete annotation of its genome sequence. This facultative anaerobic bacterium, isolated from the stool of 38-year-old obese Frenchman, is Gram-positive, has motile rods and is sporulating. The genome is 4 057 828 bp long with 72.42% G + C content. Of the 3732 predicted genes, 3667 were protein-coding genes and 65 were RNAs.

Keywords: Cellulomonas timonensis, culturomics, genome, human gut microbiota, taxonogenomics

Original Submission: 21 December 2017; Accepted: 26 January 2018

Article published online: 8 February 2018

Introduction

The development of metagenomics has enabled a better exploration of the gut microbiota, thus bypassing the problem of noncultivable bacteria and providing an understanding of the relationship between altered gut microbiota and several pathologies such as obesity, inflammatory bowel disease and irritable bowel syndrome [1]. Nevertheless, bacterial culture remains essential in order to have a better representation of the viable population and would, in addition, allow for an extension of the known gut bacterial repertoire. Recently our laboratory has developed a new concept known as microbial culturomics. This makes it possible to explore, as comprehensively as possible, the viable population of prokaryotes associated with the human gastrointestinal tract by varying culture media and physicochemical parameters [2]. As a result of this concept, several new bacteria, including new genera and species, have been reported in the human gut microbiota. As a result, culturomics has doubled the number of species isolated at least once from the human gut [3].

Current methods of defining a new bacterial species, which are based on genetic, phenotypic and chemotaxonomic criteria, are not reproducible and cannot be applied to all bacterial genera [4,5]. Furthermore, the availability of genomic data for many bacterial species [6] has recently led to a new concept of bacterial description being proposed, including a proteomic description obtained y matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) [7,8] alongside biochemical and genomic analyses of the new species [9].

The genus Cellulomonas was first described by Bergey et al. [10] and later amended by Clark and Stackebrandt et al. [11,12]. The members of this genus are Gram-positive irregular rods with cellulolytic activity. They were cultivated in aerobic conditions, and most strains are also capable of anaerobic growth. They have L-ornithine in their peptidoglycan, contain menaquinone MK-9(H4) as the predominant respiratory quinone, have anteiso-C15:0 and C16:0 as the major fatty acids [11] and have a high genomic G + C content of 71 to 76 mol% [13]. To date, this genus consists of 28 species [9,14]. Most of these...
species were originally isolated from environmental samples, and occasionally from rumen and activated sludge.

Here we present a summary classification and a set of features for the type strain \textit{C. timonensis} sp. nov. strain sn7 (= CSUR P2058 = DSM 100699), a new bacterial species isolated by culturomics from the stool sample of an obese Frenchman, together with the description of the complete genomic sequence and its annotation.

Materials and methods

Organism information
A stool sample was collected from a 38-year-old Frenchman living in France who was included in a research protocol. The stool sample was frozen at ~80°C after sampling at the La Timone hospital in Marseille. The patient provided written informed consent. Both this study and the consent procedure were approved by the ethics committee of the Federative Research Institute IFR48, Faculty of Medicine, Marseille, France (agreement 09-022).

Strain identification by MALDI-TOF MS and 16S rRNA sequencing
The stool sample was cultured on 5% sheep’s blood–enriched Columbia agar (bioMérieux, Marcy l’Etoile, France) at 37°C in microaerophilic atmosphere generated by CampyGen (Oxoid, Dardilly, France). After 48 hours’ incubation, the isolated colonies were deposited in duplicate on a MALDI-TOF MS MSP96 target plate (Bruker Daltonics, Leipzig, Germany), then covered with 1.5 μL of a matrix solution (saturated solution of α-cyano-4-hydroxy-cinnamic acid diluted in 50% acetonitrile, 2.5% trifluoroacetic acid, completed with high-performance liquid chromatography water). Proteomic analysis of our strain was carried out with MALDI-TOF MS as previously described [9] using a MicroFlex spectrometer (Bruker). Twelve distinct deposits were made for strain sn7T from 12 isolated colonies. Twelve spectra were thus obtained, imported into MALDI BioTyper software (version 2.0; Bruker) and analysed by standard pattern matching (with default parameter settings) against the main spectra of 7567 bacteria (Bruker database completed standard pattern matching (with default parameter settings) against the main spectra of 7567 bacteria (Bruker database completed against the main spectra of 7567 bacteria (Bruker database completed by the BioTyper software). The comparison with the BioTyper database spectra enabled the identification and discrimination of the analysed species from those in the database as a result of the obtained score: a score of >2 with a validated species enabled identification at the species level, and a score of <1.7 did not enable any identification. If the colony was not identified, despite a clean spectrum, a sequencing of 16S rDNA was performed as previously described [15] to define taxonomic criteria. BLASTn searches were performed at the National Center for Biotechnology Information (NCBI) website (http://blast.ncbi.nlm.nih.gov/gate1.inist.fr/Blast.cgi) to compare and identify the 16S rDNA sequence of our strain. A threshold of 98.7% similarity was determined to define a new species without performing DNA-DNA hybridization (DDH) [4].

Growth conditions
Different growth temperatures (28, 37, 45, 55°C) were tested under anaerobic and microaerophilic conditions using AnaeroGen and CampyGen respectively (Thermo Fisher Scientific, Courtaboeuf, France). The strain growth was also tested under aerobic conditions, in the presence of air and with or without 5% CO₂. The tolerance of this strain sn7T to salt (0–5, 50–75 and 100 g/L NaCl) and pH (6, 7 and 8.5) was calculated.

Morphologic, biochemical and antibiotic susceptibility tests
Gram staining and motility were observed from fresh colonies between blades and slats using a DM1000 photonic microscope (Leica Microsystems, Nanterre, France) with a 40 × objective lens [16]. Spore formation was determined by thermal shock (80°C for 20 minutes) and observed under a microscope. Negative staining was carried out with detection formvar-coated grids placed on a drop of 40 μL of bacterial suspension and incubated at 37°C for 30 minutes, followed by a 10-second incubation in 1% ammonium molybdate. The grids were dried on blotting paper and then observed with a Tecnai G20 transmission electron microscope (FEI Company, Liméil-Brévannes, France). We studied the biochemical characteristics of this strain using API 20NE, API ZYM and API 50CH strips according to the manufacturer’s instructions (bioMérieux). Oxidase and catalase reactions were determined using a BBL DrySlide (Becton, Le Pont de Claix, France) according to the manufacturer’s instructions. The antimicrobial activity test was performed using the disc diffusion method (i2a, Montpellier, France) [17] on Mueller-Hinton agar in a petri dish (bioMérieux).

Fatty acid methyl ester analysis by gas chromatography/mass spectrometry
Cellular fatty acid methyl ester (FAME) analysis was performed by gas chromatography/mass spectrometry (GC/MS). Two samples were prepared with approximately 60 mg of bacterial biomass per tube collected from several culture plates. FAMEs were prepared as described by Sasser [18]. GC/MS analyses were carried out as previously described [19]. Briefly, FAMEs were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500–SQ 8 S; Perkin Elmer, Courtaboeuf, France). A spectral database search was performed using MS Search 2.0 operated with the Standard Reference
Database 1A (National Institute of Standards and Technology, Gaithersburg, MD, USA) and the FAME mass spectral database (Wiley, Chichester, UK).

Genome sequencing and assembly
Genomic DNA of C. timonensis was extracted as previously described [9] and sequenced on the MiSeq sequencer (Illumina, San Diego, CA, USA) using the mate-pair strategy. The genomic DNA was barcoded in order to be mixed with 11 other projects using the Nextera Mate Pair sample prep kit (Illumina). All genome sequencing steps were performed following the process previously reported by Lagier et al. [9]. Genome assembly was performed using a pipeline that enabled the creation of an assembly with different pieces of software (Velvet [20], Spades [21] and Soap Denovo [22]) on trimmed (MiSeq and Trimomatic) [23] or untrimmed (only MiSeq software) data. For each of the six assemblies performed, GapCloser [22] was used to reduce gaps. Contamination with Phage Phix was then identified (BLASTn against Phage Phix174 DNA sequence) and eliminated. Finally, scaffolds <800 bp in size were removed. Scaffolds with a depth value of lower than 25% of the mean depth were also removed (identified as possible contaminants). The best assembly was selected by using different criteria (number of scaffolds, N50, number of N). For the studied strain, 250 bp gave the best assembly, with a depth coverage of 2 X 250 bp.

Genome annotation
Open reading frames (ORFs) were predicted using Prodigal [24] with default parameters, but the predicted ORFs were excluded if they spanned a sequencing gap region (containing N). The predicted bacterial protein sequences were searched against the Clusters of Orthologous Groups (COGs) database using BLASTP (E value of 1e-03, coverage 0.7 and identity 30%). If no hit was found, sequences were searched against the NR database using BLASTP (E value of 1e-03, coverage 0.7 and identity 30%). If the length of the sequence was <80 aa, we used an E value of 1e-05. The tRNAscanSE [25] tool was used to find transfer RNA genes, whereas ribosomal RNA genes were found using RNAmmer [26]. Lipoprotein signal peptides and the number of transmembrane helices were predicted using Phobius [27]. ORFans were identified if all the BLASTP runs performed did not give positive results (E value smaller than 1e-03 for ORFs with sequence size >80 aa or E value smaller than 1e-05 for ORFs with sequence length <80 aa). Such parameter thresholds have been used in previous studies to define ORFans. The annotation process was performed using DAGOBAH [28] including the Figenix [29] libraries, which provided the pipeline analysis.

16S rRNA phylogenetic tree
Sequences were recovered after a nucleotide BLAST against the 16S rRNA database of the All-Species Living Tree SILVA project (LTPs119). First, a filter to eliminate sequences smaller than 1450 bp was applied. (Sometimes this filter is decreased to retrieve more sequences.) Pass filter sequences were aligned using Muscle [30], and phylogenetic inferences were obtained using the approximate maximum-likelihood method within FastTree software [31]. Numbers at the nodes corresponding to local values were computed by the Shimodaira-Hasegawa test. A filter using PhyloPattern [32] was applied to the tree to remove duplicate species and bad taxonomic reference species. This pipeline was performed in DAGOBAH [28], which include the Figenix [29] libraries.

Genome comparison analysis
Species to be compared were automatically retrieved from the 16s RNA tree using PhyloPattern. For each selected species, the complete genome sequence, proteome sequence and ORFeome sequence were retrieved from the NCBI FTP site. If one specific strain did not have a complete and available genome, a complete genome of the same species was used. If ORFeomes and proteomes were not predicted, Prodigal was used with default parameters to predict them. All proteomes were analysed using proteinOrtho [33]; then, for each couple of genomes, a similarity score was computed. This score is the mean value of nucleotide similarity between all couples of orthologous genes between the two genomes studied (average genomic identity of orthologous gene sequences (AGIOS) tool). An annotation of all proteomes was also performed to define the distribution of functional classes of predicted genes.

---

**FIG. 1.** Reference matrix-assisted desorption ionization—time of flight mass spectrometry analysis of *Cellulomonas timonensis* strain sn7T.
according to COGs proteins (using the same method as that for genome annotation). The comparison process was performed by DAGOBAD [28], which includes the Figenix [29] libraries, which provided pipeline analysis, and PhyloPattern [32] for tree manipulation. To evaluate the genomic similarity among the Cellulomonas strains studied, we set two parameters: digital DDH, which exhibits a high correlation with DDH [34,35], and AGIOS, which was designed to be independent of DDH.

Results

Strain identification and phylogenetic analyses

Strain sn7T was isolated in May 2015 by cultivation on 5% sheep’s blood–enriched agar under microaerophilic conditions. This strain was not identified by MALDI-TOF MS, and its spectrum was added to our database (Fig. 1). Sequencing of the 16S rRNA gene demonstrated that this strain sn7T exhibited a nucleotide sequence similarity of 98.4% with Cellulomonas cellulosa (GenBank accession no. KR922256), the phylogenetically closest bacterial species with a validly published name (Fig. 2). Its 16S rRNA sequence was deposited in the European Molecular Biology Laboratory–European Bioinformatics Institute (EMBL-EBI) database under accession number LN870311. This value was lower than the 98.7% 16S rRNA gene sequence threshold recommended by Stackebrandt and Ebers [4] to delineate a new species without carrying out DDH. Strain sn7T is thus a new species which has been named Cellulomonas timonensis (Table 1). The other closest species were C. chitinilytica (97.94%), C. biazotea (97.7%), C. fimii (97.5%) and C. xylanilytica (96.9%). The species C. timonensis, C. chitinilytica and C. fimii shared a single cluster, whereas C. xylanilytica is present in a distant

![Phylogenetic tree highlighting position of Cellulomonas timonensis strain sn7T relative to other type strains within genus Cellulomonas. GenBank accession numbers are indicated at right of species name. Sequences were aligned using CLUSTAL W, and phylogenetic inferences were obtained using maximum-likelihood method within MEGA software. Numbers at nodes are bootstrap values obtained by repeating analysis 500 times to generate majority consensus tree. Scale bar indicates 1% nucleotide sequence divergence.](http://example.com/phylogenetic-tree.png)
clade in the phylogenetic tree. A comparison between the spectral differences of \textit{C. timonensis} and other closest species is represented in a gel view (Fig. 3).

**Phenotypic description**

Growth of strain sn7\textsuperscript{T} was observed between 28 and 37°C on 5% sheep’s blood–enriched Colombia agar, and optimal growth was achieved at 37°C after 24 hours’ incubation in aerobic conditions and 48 hours’ incubation in anaerobic and microaerobic atmospheres. Cells were motile and sporulating.

Colonies were irregular, with a diameter of 1.5 to 2 mm on blood-enriched Colombia agar after 48 hours. The results of pH testing showed that strain sn7\textsuperscript{T} can survive under pH conditions ranging between 7 and 8.5, but has optimal growth at pH 7. It grows in salinity concentrations ranging from 0 to 5 g/L NaCl. Gram staining (Fig. 4) showed Gram-positive rods. Using electron microscopy, the rods had a mean diameter of 0.3 μm and a length of 1.3 μm (Fig. 5).

**TABLE 1.** Classification and general features of \textit{Cellulomonas timonensis} sn7\textsuperscript{T}

| Property                  | Term                                      |
|---------------------------|-------------------------------------------|
| Current classification    | Domain: Bacteria                          |
|                           | Phylum: Actinobacteria                    |
|                           | Class: Actinobacteria                     |
|                           | Order: Micrococcaceae                     |
|                           | Family: Cellulomonadaceae                 |
|                           | Genus: Cellulomonas                       |
|                           | Species: \textit{Cellulomonas timonensis}  |
|                           | Type strain: sn7\textsuperscript{T}       |
| Gram stain                | Positive                                  |
| Cell shape                | Rod                                       |
| Motility                  | Motile                                    |
| Sporulation               | Sporulating                               |
| Temperature range         | Mesophilic                                |
| Optimum temperature       | 37°C                                      |
| Oxygen requirement        | Facultative anaerobic                     |
| Salinity                  | 0–5 g/L                                   |
| Optimum salinity          | 1 g/L                                     |
| pH                        | 7–8.5                                     |
| Optimum pH                | 7                                         |
| Pathogenicity             | Unknown                                   |
| Habitat                   | Human gut                                 |
| Isolation                 | Human faeces                              |

**FIG. 3.** Gel view comparing \textit{Cellulomonas timonensis} strain sn7\textsuperscript{T} to other close species. Gel view displays raw spectra of loaded spectrum files arranged in pseudo-gel-like look. X-axis records m/z value. Left y-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity is expressed by greyscale scheme code. Colour bar and right y-axis indicate relation between colour of peak and its intensity in arbitrary units. Displayed species are indicated at left.

**FIG. 4.** Gram staining of \textit{Cellulomonas timonensis} strain sn7\textsuperscript{T}.

Colonies were irregular, with a diameter of 1.5 to 2 mm on blood-enriched Colombia agar after 48 hours. The results of pH testing showed that strain sn7\textsuperscript{T} can survive under pH conditions ranging between 7 and 8.5, but has optimal growth at pH 7. It grows in salinity concentrations ranging from 0 to 5 g/L NaCl. Gram staining (Fig. 4) showed Gram-positive rods. Using electron microscopy, the rods had a mean diameter of 0.3 μm and a length of 1.3 μm (Fig. 5).
Catalase was positive for strain sn7T, and oxidase was negative. Using the API ZYM gallery, positive reactions were observed for alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase (C14), leucine arylamidase, valine arylamidase, α-chymotrypsin, acid phosphatase, naphthol-AS-BI-phosphohydrolase, β-glucosidase and N-acetyl-β-glucosaminidase. Cystine arylamidase, trypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, α-mannosidase and α-fucosidase were negative. Using the API 20 NE system, a positive reaction was obtained for a nitrate reduction, indole formation, glucose fermentation, esculin hydrolysis and β-galactosidase, and assimilation for glucose, arabinose and mannose were observed. All other reactions were negative, including urease and gelatin hydrolysis. An API 50CH strip showed positive reactions for glycerol, L-arabinose, D-ribose, D-xylose, D-galactose, D-glucose, D-fructose, D-mannose, arbutin, FIG. 5. Transmission electron microscopy of Cellulomonas timonensis strain sn7T using Tecnai G20 device at operating voltage of 60 kV. Scale bar = 500 nm.

| Property               | C. timonensis | C. cellasea | C. massiliensis | C. fimii | C. chitinilytica | C. soli | C. humilata | C. xylanilytica | C. terrae |
|------------------------|---------------|-------------|----------------|---------|-----------------|--------|-------------|----------------|----------|
| Oxygen requirement     | Facultative   | Facultative | Aerobic        | Facultative | Facultative | Aerobic | Facultative | Aerobic/microaerophilic | Facultative | Facultative |
| Salt requirement       | 0–5 g/L       | NA          | 5 g/L          | NA      | NA              | 0–3 g/L | <4 g/L      | NA             | NA       |
| Modality               | +             | –           | +              | –       | –               | –      | –           | –              | –        |
| Endospore formation    | +             | NA          | –              | NA      | NA              | –      | –           | –              | –        |
| Production of:         |               |             |                |         |                 |        |             |                |          |
| Alkaline phosphatase   | +             | NA          | NA             | NA      | NA              | NA     | NA          | NA             | NA       |
| Catalase               | –             | –           | –              | –       | –               | +      | –           | –              | –        |
| Oxidase                | –             | NA          | –              | –       | –               | –      | –           | –              | –        |
| Nitratreductase        | +             | –           | –              | –       | –               | –      | –           | –              | –        |
| Urease                 | –             | –           | –              | –       | –               | –      | –           | –              | –        |
| β-Galactosidase        | +             | +           | –              | +       | +               | NA     | +           | NA             | NA       |
| N-Acetyl-glucosamine   | –             | NA          | NA             | NA      | NA              | –      | NA          | NA             | NA       |
| Acid from:             |               |             |                |         |                 |        |             |                |          |
| L-Arabinose            | +             | –           | –              | +       | +               | –      | NA          | NA             | NA       |
| Ribose                 | +             | –           | NA             | +       | +               | –      | –           | –              | –        |
| Mannose                | +             | +           | +              | +       | +               | +      | +           | +              | +        |
| Mannitol               | –             | –           | –              | –       | –               | –      | –           | –              | –        |
| Sucrose                | +             | +           | –              | –       | –               | –      | +           | +              | +        |
| α-Glucose              | +             | –           | –              | –       | –               | +      | +           | +              | +        |
| α-Fructose             | +             | +           | NA             | +       | +               | +      | –           | –              | +        |
| α-Maltose              | +             | –           | NA             | NA      | NA              | –      | +           | +              | +        |
| α-Lactose              | –             | –           | NA             | –       | –               | –      | NA          | +              | +        |
| Habitat                | Human gut     | Soil        | Human gut      | Soil    | Cattle farm compost | Soil     | Soil        | Elm tree       | Soil    |

IUPAC, International Union of Pure and Applied Chemistry; TR = trace amounts <1%.

Mean peak area percentage ± standard deviation.

### TABLE 3. Cellular fatty acid composition (%)

| Fatty acid | IUPAC name                           | Mean relative % |
|------------|--------------------------------------|-----------------|
| 15:0 anteiso | 12-methyl-Tetradecanoic acid         | 74.0 ± 0.2      |
| 17:0 anteiso | 14-methyl-Hexadecanoic acid          | 8.4 ± 0.6       |
| 15:1n5 iso | 13-methyl-Tetradec-9-enoic acid      | 5.0 ± 0.6       |
| 16:0        | Hexadecanoic acid                    | 3.9 ± 0.1       |
| 16:0 iso    | 14-methyl-Pentadecanoic acid         | 2.0 ± 0.3       |
| 15:0        | Pentadecanoic acid                   | 1.6 ± 0.1       |
| 15:0 iso    | 13-methyl-Hexadecanoic acid          | 1.4 ± 0.1       |
| 17:1 cyclo  | 11-methyl-Cyclohexylundecanoic acid  | 1.0 ± 0.1       |
| 18:0        | Octadecanoic acid                    | TR              |
| 18:2n6      | 9,12-Octadecadienoic acid            | TR              |
| 5:0 anteiso | 2-methyl-Butanoic acid               | TR              |
| 17:0        | Heptadecanoic acid                   | TR              |
| 17:0 iso    | 15-methyl-Hexadecanoic acid          | TR              |
| 6:0 iso     | 4-methyl-Pentanoic acid              | TR              |

IUPAC, International Union of Pure and Applied Chemistry; TR = trace amounts <1%.

*Mean peak area percentage ± standard deviation.

© 2018 The Authors. Published by Elsevier Ltd, NMNI, 23, 7–16
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
esculin ferric citrate, salicin, D-cellobiose, D-maltose, D-saccharose, D-trehalose, amidon and glycogen. Negative reactions were recorded for erythritol, D-arabinose, L-xylose, D-adonitol, methyl-β-D-xylopyranoside, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-α-D-mannopyranoside, methyl-α-D-glucopyranoside, N-acetyl-glucosamine, amygdalin, D-lactose, D-melibiose, inulin, D-melezitose, D-raffinose, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate. Table 2 shows a comparison of the principal phenotypic and biochemical features between C. timonensis and other species belonging to the Cellulomonas genus. Cells were susceptible to ceftriaxone, ciprofloxacin, clindamycin, doxycycline, erythromycin, gentamicin, penicillin, rifampicin, teicoplanin

### TABLE 4. Nucleotide content and gene count levels of genome

| Attribute                  | Value       | % of total |
|----------------------------|-------------|------------|
| Size (bp)                  | 4 057 828   | 100        |
| G + C content (%)          | 2 938 504   | 72.41      |
| Coding region (bp)         | 3 726 909   | 91.84      |
| Total genes                | 3732        | 100        |
| Protein-coding genes       | 3667        | 100        |
| Genes with function prediction | 2727        | 74.36      |
| Genes assigned to COGs     | 2405        | 65.58      |
| Genes with peptide signals | 481         | 13.11      |
| Genes with transmembrane helices | 913       | 24.89      |
| ORFans                     | 176         | 4.79       |
| Genes associated with PKS or NRPS | 28          | 0.76       |
| No. of antibiotic resistance genes | 0            | 0          |

COGs, Clusters of Orthologous Groups database; PKS, polyketide synthase.

**Table 4.** Nucleotide content and gene count levels of genome.

### TABLE 5. Number of genes associated with 25 general COGs functional categories

| Code | Value | % of total | Description                                      |
|------|-------|------------|--------------------------------------------------|
| J    | 184   | 5.017726   | Translation                                       |
| K    | 189   | 5.154077   | RNA processing and modification                   |
| L    | 88    | 2.399782   | Replication, recombination and repair             |
| B    | 1     | 0.027270248| Chromatin structure and dynamics                  |
| D    | 27    | 0.736296   | Cell cycle control, mitosis and meiosis          |
| Y    | 0     | 0          | Nuclear structure                                 |
| V    | 84    | 2.290701   | Defense mechanisms                                |
| T    | 99    | 2.6997545  | Signal transduction mechanisms                    |
| M    | 105   | 2.863376   | Cell wall/membrane biogenesis                     |
| N    | 52    | 1.4180529  | Cell mobility                                     |
| Z    | 0     | 0          | Cytoskeleton                                      |
| W    | 19    | 0.5181347  | Extracellular trafficking and secretion           |
| U    | 35    | 0.95445865 | Intracellular trafficking and secretion           |
| O    | 102   | 2.7815652  | Posttranslational modification, protein turnover, chaperones |
| X    | 12    | 0.32724297 | Mobilome: prophages, transposons                  |
| C    | 156   | 4.2541585  | Energy production and conversion                  |
| G    | 343   | 9.353695   | Carbohydrate transport and metabolism             |
| E    | 281   | 7.66294    | Amino acid transport and metabolism               |
| F    | 96    | 2.6179438  | Nucleotide transport and metabolism               |
| H    | 142   | 3.8723752  | Coenzyme transport and metabolism                 |
| I    | 111   | 3.026976   | Lipid transport and metabolism                    |
| P    | 170   | 4.6359425  | Inorganic ion transport and metabolism            |
| Q    | 86    | 2.345213   | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 248   | 6.7630215  | General function prediction only                  |
| S    | 121   | 3.2997     | Function unknown                                  |
| —    | 1262  | 34.415054  | Not in COGs                                      |

COGs, Clusters of Orthologous Groups database.

**Table 5.** Number of genes associated with 25 general COGs functional categories.

### FIG. 6.** Graphical circular map of chromosome of Cellulomonas timonensis strain sn7. From outside to centre: genes on forward strand coloured by COGs categories (only genes assigned to COGs), genes on reverse strand coloured by COGs categories (only gene assigned to COGs), RNA genes (tRNAs green, rRNAs red), GC content and GC skew. COGs, Clusters of Orthologous Groups database.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
and vancomycin but were resistant to colistin, fosfomycin, oxacillin and trimethoprim/sulfamethoxazole.

**FAME analysis by GC/MS**

The major fatty acid was the saturated and branched 12-methyltetradecanoic acid (74%). The detected compounds are mainly all saturated fatty acids. Very few unsaturated species are listed in Table 3.

**Genome properties**

The genome is 4,057,828 bp long with 72.42% G + C content (Table 4, Fig. 6). (Accession no. FCOT00000000). It is composed of 13 scaffolds (composed of 13 contigs). Of the 3,732 predicted genes, 3,667 were protein-coding genes and 65 were RNAs (five genes were 5S rRNA, five genes 16S rRNA, five genes 23S rRNA and 50 genes tRNA). A total of 2,727 genes (74.37%) were assigned a putative function (by COGs or by NR BLAST). A total of 176 genes (4.80%) were identified as ORFans. The remaining genes were annotated as hypothetical proteins (643 genes were >17.53%). Table 5 represents the distribution of *C. timonensis* genes into the different COGs categories.

**Genome comparison**

We compared the genome of *C. timonensis* with other close species: *Cellulomonas flavigena*, *Cellulomonas terrae*, *Cellulomonas xylanilytica*, *Cellulomonas chitinilytica*, *Cellulomonas fimii* and *Cellulomonas cellasea* (Table 6). The draft genome sequence of *C. timonensis* is smaller than those of *C. flavigena* and *C. fimii* (4.06, 4.12 and 4.27 Mbp respectively), but larger than those of *C. xylanilytica*, *C. cellasea*, *C. chitinilytica* and *C. terrae* (1.78, 3.91, 1.17 and 1.84 Mbp respectively). The G + C content of *C. timonensis* is smaller than those of *C. cellasea*, *C. flavigena* and *C. fimii* (72.42, 74.55, 74.29 and 74.72% respectively), but larger than those of *C. xylanilytica*, *C. chitinilytica* and *C. terrae* (69.55, 68.99 and 69.55% respectively). The gene content of *C. timonensis* is smaller than those of *C. flavigena* and *C. fimii* (3,667, 3,676 and 3,762 respectively), but larger than those of *C. xylanilytica*, *C. cellasea*, *C. chitinilytica* and *C. terrae* (3,510, 3,560, 2,627 and 3,606 respectively) (Table 6). We observed an identical distribution of

---

**TABLE 6. Genome comparison of closely related species to *Cellulomonas timonensis* strain sn7**

| Cellulomonas species | INSDC No. | Size (Mbp) | G + C % | Protein-coding genes |
|----------------------|-----------|-----------|---------|---------------------|
| C. timonensis strain sn7 | FCOT00000000 | 4.05 | 72.41 | 3,667 |
| C. xylanilytica strain XIL11 | BBGX00000000.1 | 1.77 | 69.55 | 3,510 |
| C. cellasea strain DSM 20118T | AXNT00000000.1 | 1.39 | 74.55 | 3,230 |
| C. flavigena strain DSM 20109 | CP001964.1 | 4.12 | 74.29 | 3,676 |
| C. chitinilytica strain X.bu-b | BBHG00000000.1 | 1.17 | 68.99 | 2,627 |
| C. terrae strain DB5 | BBGZ00000000.1 | 1.83 | 69.55 | 3,606 |
| C. fimii strain DSM 20113T | CP002666.1 | 4.26 | 74.72 | 3,725 |

INSDC, International Nucleotide Sequence Database Collaboration.
genes into COGs categories in all compared genomes (Fig. 7). An analysis of orthologous genes shared among the different genomes revealed that C. timonensis shared 106, 353, 1623, 1742, and 1487 orthologous genes with C. cellasea, C. chitinilytica, C. terrae, C. xylanilytica, and C. flavigena, respectively. Among species with standing in nomenclature, AGIOS values ranged from 64.51% to 87.97% among compared species, with the exception of C. timonensis. When compared to other species, the AGIOS values ranged from 66.61% with C. cellasea to 77.31% with C. xylanilytica (Table 7).

Two parameters were used to evaluate genomic similarity among the studied strains: AGIOS (Table 7), which was designed to be independent of DDH, and digital DDH, which exhibits a high correlation with DDH [34,35] (Table 8).

**Conclusion**

On the basis of phenotypic, phylogenetic and genomic analyses, we formally propose the creation of *Cellulomonas timonensis* sp. nov., which contains the strain sn7T. This bacterial strain has been isolated from the faecal flora of a 38-year-old obese Frenchman.

**Description of Cellulomonas timonensis** sp. nov.

*Cellulomonas timonensis* (tim.o.nen’sis, L. masc. adj., timonensis, ‘of Timone,’ the name of the hospital where strain sn7T was first cultivated).

C. timonensis is a facultative anaerobic Gram-positive bacterium which is rod shaped with a mean diameter of 0.3 μm and a length of 1.3 μm. Optimal growth of strain sn7T occurs at 37°C in aerobic conditions. Colonies are white and smooth with an irregular diameter of 1.5 to 2 mm on 5% sheep’s blood–enriched Colombia agar. Cells are sporulating and motile.

C. timonensis shows negative reactions for oxidase, cystine arylamidase, trypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, α-mannosidase, and α-fucosidase. Positive reactions were observed for catalase, esterase (C4), acid phosphatase, nitrate reduction, indole formation, glucose fermentation and esculin hydrolysis. Strain sn7T showed resistance to colistin, fosfomycin, oxacillin and trimethoprim/sulfamethoxazole. The fatty acids are mainly composed of 12-methyl-tetradecanoic acid (74%).

This strain exhibited a G + C content of 72.42% and a genome length of 4 057 828 bp. The 16S rRNA sequence and the whole genome shotgun sequence have been deposited in EMBL-EBI under accession numbers LN870311 and FCOT00000000, respectively. C. timonensis strain sn7T (=CSUR P2058 = DSM 100699) was isolated from a stool sample of an obese Frenchman.

**Conflict of interest**

None declared.

---

### TABLE 7. Number of orthologous proteins shared between Cellulomonas genomes (upper right)

|                | C. chitinilytica | C. flavigena | C. flavigena | C. flavigena | C. flavigena | C. flavigena |
|----------------|-----------------|--------------|--------------|--------------|--------------|--------------|
| C. timonensis  | 106 (100–100%)  | 353 (100–100%) | 1623 (100–100%) | 1742 (100–100%) | 1487 (100–100%) | 1487 (100–100%) |
| C. cellasea    | 22.1 (19.8–24.6%) | 28.4 (26–30.9%) | 21.9 (19.6–24.3%) | 21.2 (19–23.7%) | 28.9 (26.5–31.4%) | 31.3 (30–33.1%) |
| C. xylanilytica| 30.8 (28.4–33.3%) | 30.8 (28.4–33.3%) | 30.8 (28.4–33.3%) | 30.8 (28.4–33.3%) | 30.8 (28.4–33.3%) | 30.8 (28.4–33.3%) |
| C. flavigena   | 28 (25.7–30.5%)  | 100 (100–100%) | 100 (100–100%) | 100 (100–100%) | 100 (100–100%) | 100 (100–100%) |
| C. terrae      | 28.8% (26–28.8%) | 28.8% (26–28.8%) | 28.8% (26–28.8%) | 28.8% (26–28.8%) | 28.8% (26–28.8%) | 28.8% (26–28.8%) |

**TABLE 8. Pairwise comparison of Cellulomonas timonensis sn7T with other Cellulomonas species using GGDC, formula 2 (DDH estimates based on identities/HSP length)**

|                | C. timonensis | C. cellasea | C. flavigena | C. timonensis | C. flavigena | C. xylanilytica |
|----------------|--------------|-------------|--------------|--------------|--------------|----------------|
| C. timonensis  | 26.30 (24–28.8%) | 26.3 (23.8–28.6%) | 25.6 (23.9–29.1%) | 24.3 (24.3–29.1%) | 24.3 (24.8–29.3%) | 24.3 (24.8–29.3%) |
| C. cellasea    | 26.1 (23.8–28.6%) | 26.1 (23.8–28.6%) | 26.1 (23.8–28.6%) | 26.1 (23.8–28.6%) | 26.1 (23.8–28.6%) | 26.1 (23.8–28.6%) |
| C. flavigena   | 30.6 (28.2–33.1%) | 30.6 (28.2–33.1%) | 30.6 (28.2–33.1%) | 30.6 (28.2–33.1%) | 30.6 (28.2–33.1%) | 30.6 (28.2–33.1%) |
| C. xylanilytica| 30.7 (28.3–33.2%) | 30.7 (28.3–33.2%) | 30.7 (28.3–33.2%) | 30.7 (28.3–33.2%) | 30.7 (28.3–33.2%) | 30.7 (28.3–33.2%) |

DDH, DNA-DNA hybridization; GGDC, Genome-to-Genome Distance Calculator; HSP, high-scoring segment pairs.
Acknowledgements

The authors thank the Xegen Company (www.xegen.fr) for automating the genomic annotation process. This work has benefited from the French State support, managed by the ‘Agence Nationale pour la Recherche’ including the “Programme d’Investissement d’avenir” under the reference Méditerranée Infection 10-IAHU-03.

References

[1] Salazar N, Arboleya S, Valdés L, Stanton C, Ross P, Ruiz L, et al. The human intestinal microbiome at extreme ages of life. Dietary intervention as a way to counteract alterations. Front Genet 2015:4;406.
[2] Lagier JC, Hugon P, Khelifa S, Fournier PE, La Scola B, Raoult D. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 2015;28:237–64.
[3] Lagier JC, Khelifa S, Aoufi MT, N’Dongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 2016;1:16203.
[4] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;3:152–5.
[5] Vandamme P, Pot B, Gillis M, de Vos P, Kersters K, Swings J. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Rev 1996;60:407–38.
[6] Wayne LG. International committee on systematic bacteriology: announcement of the report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Zentralbl Bakteriol Mikrobiol Hyg A 1988;268:433–4.
[7] Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 2009;49:543–51.
[8] Seng P, Abat C, Rolain JM, Colson P, Lagier JC, Gouriet F, et al. Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of matrix-assisted laser desorption ionization-time of-flight mass spectrometry. J Clin Microbiol 2013;51:2182–94.
[9] Lagier JC, Ramasamy D, Rivet R, Raoult D, Fournier PE. Non-contiguous-finished genome sequence and description of Cellulomonas massiliensis sp. nov. Stand Genomic Sci 2012;7:208–70.
[10] Bergey DH, Harrison FC, Breed RS, Hammer BW, Hunton FM, editors. Bergey’s manual of determinative bacteriology. Baltimore: Williams & Wilkins; 1923.
[11] Stackebrandt E, Schumann P, Prauser H. The family Cellulomonadaceae. In: The prokaryotes: a handbook on the biology of bacteria. New York: Springer; 2006. p. 981–1001.
[12] Stackebrandt E, Schumann P, Genus I. Cellulomonas Bergey, Harrison, Breed, Hammer and Hunton 1923, 154 emend. mut. char. Clark 1952. In: Whitman W, Goodfellow M, Kämpfer P, Busse H-J, Trujillo M, editors. Bacterial identification of rare pathogenic bacteria in a clinical microbiology environment and clinical unidentifiable bacterial isolates. J Clin Microbiol 2000;38:3623–30.
[13] Lagier JC, Edouard S, Pagnier I, Medaniakkov O, Drancourt M, Raoult D. Current and past strategies for bacterial culture in clinical microbiology. Clin Microbiol Rev 2015;28:208–36.
[14] Le Page S, van Belkum A, Fulchiron C, Huguet R, Raoult D, Rolain JM. Evaluation of the PREVIBIS isola automated seeder system compared to reference manual inoculation for antibiotic susceptibility testing by the disk diffusion method. Eur J Clin Microbiol Infect Dis 2015;34:1859–69.
[15] Sasser M. Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). Newark, NY: Microbial ID; 2006.
[16] Dione N, Sankar SA, Lagier JC, Khelifa S, Michele C, Armstrong N, et al. Genome sequence and description of Aerolysinobacter massiliensis sp. nov. New Microbes New Infect 2016;10:66–76.
[17] Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 2008;18:821–9.
[18] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–20.
[19] Käll L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 2004;338:1027–103.
[20] Gouret P, Paganini J, Dinaut J, Loutai D, Darbo E, Pantorotti P, et al. Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: the multi-agent software system DAGOBH. In: Pantorotti P, editor. Evolutionary biology: concepts, biodiversity, macroevolution and genome evolution. Berlin: Springer Verlag; 2011. p. 71–87.
[21] Kall L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 2004;338:1027–103.
[22] Gouret P, Paganini J, Dinaut J, Loutai D, Darbo E, Pantorotti P, et al. Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: the multi-agent software system DAGOBH. In: Pantorotti P, editor. Evolutionary biology: concepts, biodiversity, macroevolution and genome evolution. Berlin: Springer Verlag; 2011. p. 71–87.
[23] Käll L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 2004;338:1027–103.
[24] Gouret P, Paganini J, Dinaut J, Loutai D, Darbo E, Pantorotti P, et al. Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: the multi-agent software system DAGOBH. In: Pantorotti P, editor. Evolutionary biology: concepts, biodiversity, macroevolution and genome evolution. Berlin: Springer Verlag; 2011. p. 71–87.
[25] Käll L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 2004;338:1027–103.