Genome sequence and description of Desnuesiella massiliensis gen. nov., sp. nov. a new member of family Clostridiaceae

L. Hadjadj1, M. Tidjani Alou1, C. Sokhna2, J.-C. Lagier1, D. Raoult1,3 and J.-M. Rolain1

1) Unité de recherche sur les maladies infectieuses et tropicales émergentes (URMITE), UMR CNRS, IHU Méditerranée Infection, Faculté de Médecine et de Pharmacie, Aix-Marseille-Université, Marseille, France, 2) Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM63, CNRS7278, IRD198, InsermU1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Aix-Marseille Université, Marseille, France and Dakar, Senegal and 3) Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract

Desnuesiella massiliensis, strain MT10T gen. nov., sp. nov. is a newly proposed genus within the family Clostridiaceae, isolated from the digestive microbiota of a child suffering from kwashiorkor. Desnuesiella massiliensis is a facultatively anaerobic, Gram-positive rod. Here we describe the features of this organism, together with the complete genome sequence and annotation. The 5 503 196-bp long genome (one chromosome but no plasmid) contains 5227 protein-coding and 81 RNA genes, including 14 rRNA genes.

New Microbes and New Infections © 2016 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

Keywords: Culturomics, Desnuesiella massiliensis, genome, kwashiorkor, taxono-genomics

Original Submission: 4 February 2016; Revised Submission: 12 March 2016; Accepted: 14 March 2016

Article published online: 19 March 2016

Introduction

Strain MT10T (= CSUR P1918 = DSM 101500) is the type strain of Desnuesiella massiliensis gen. nov., sp. nov. This bacterium, which is proposed to belong to the family Clostridiaceae, is a Gram-positive, flagellated, facultative anaerobic bacillus. It was isolated from the stool sample of a 1-year-old boy with kwashiorkor living in Senegal, through a culturomics study of the bacterial diversity of the faeces of children with kwashiorkor disease [1].

The newly proposed strategy of applying high throughput genome sequencing, matrix-assisted laser-desorption/ionization time-of-flight (MALDI-TOF) spectral analysis of cellular proteins, coupled with more traditional methods of phenotypic characterization has been demonstrated as a useful approach for the description of new bacterial taxa [2–5].

The family Clostridiaceae [6] belongs to the phylum Firmicutes and includes 40 genera. Members belonging to this family were isolated mainly from the environment and from commensal digestive microbiota of mammals. Some are major human pathogens, including Clostridium botulinum, Clostridium difficile, Clostridium tetani and Clostridium perfringens [7].

Here we present a summary classification and a set of features for D. massiliensis gen. nov., sp. nov. strain MT10T together with the description of the complete genome sequence and annotation. These characteristics support the circumscription of a novel genus, Desnuesiella gen. nov. within the family Clostridiaceae, with Desnuesiella massiliensis gen. nov., sp. nov. as the type species.

Organism Information

In March 2015, a faecal sample was collected from a 1-year-old boy living in Dakar, Senegal who had kwashiorkor, which is a form of severe malnutrition. Consent was obtained from the
child’s parents. The study was approved by the Institut Fédératif de Recherche 48 under agreement 09-022. The boy had not received antibiotics at the time of sample collection. The faecal specimen was preserved at –80°C after collection and sent to Marseille. Strain MT10T was isolated in liquid Columbia broth after 21 days of aerobic incubation at 37°C using the ‘cul tuo mics’ concept [1]. After a failed identification using MALDI-TOF mass spectrometry, the 16S rRNA was sequenced.

When BLAST was performed with to NCBI database, the 16S rRNA gene sequence of D. massiliensis strain MT10T (GenBank Accession number LN846906) exhibited an identity of 94.60% with Clostridium amylolyticum (Fig. 1). This value was the highest similarity observed, but was lower than the 95% 16s rRNA gene sequence threshold recommended by Stackebrandt and Ebers to delineate a new genus without carrying out DNA–DNA hybridization [8] and by Tindall et al. [9].

Different growth temperatures (25, 30, 37, 45°C) were tested. Growth occurred between 25°C and 37°C, but optimal growth was observed in anaerobic conditions at 37°C, 24 h after inoculation. Colonies were smooth, opaque and approximately 1 mm in diameter on 5% sheep blood-enriched agar (BioMérieux, Marcy l’Étoile, France). Growth of the strain was tested in anaerobic and microaerophilic atmospheres using GasPak EZ Anaerobe Pouch (Becton Dickinson Co., Franklin Lakes, NJ, USA) and CampyGen Compact (Oxoid, Basingstoke, UK) systems, respectively, and in aerobic atmosphere, with or without 5% CO₂. Growth was observed under aerobic (with and without CO₂), microaerophilic and anaerobic conditions. Gram staining showed Gram-positive rods unable to form spores (Fig. 2a). A motility test produced a positive result. Cells grown on agar did not sporulate and the rods exhibited monotrichous flagella. The size of cells were determined by negative staining transmission electron microscopy on a Technai G20 Cryo (FEI, Hillsboro, OR, USA) at an operating voltage of 200 kV, the rods have a length ranging from 2.2 to 2.6 μm (mean 2.4 μm) and a width ranging from 0.4 to 0.5 μm (mean 0.43 μm) (Fig. 2b).

Desnuesiella massiliensis is catalase-, oxidase-, urease- and indole-negative but alkaline-phosphatase-positive. Fermentation of sucrose, D-ribose, D-glucose, D-lactose, D-mannose, D-maltose, D-fructose were positive but not that of L-arabinose, D-sorbitol and D-xylose. Differential phenotypic characteristics using API 50CH and API Zym system (BioMérieux) between D. massiliensis gen. nov., sp. nov. strain MT10T and others species from the family Clostridiaceae [10–14] are detailed in Table 1.

Susceptibility testing was performed by E-test method (BioMérieux) and MIC was expressed in mg/L. Desnuesiella massiliensis was susceptible to amoxicillin (0.032), ceftriaxone (1), imipenem (0.016), erythromycin (0.25), doxycycline (0.032)
and rifampicin (0.125) but resistant to colistin (>256), gentamicin (>256) and metronidazole (>256).

Extended Features Descriptions

MALDI-TOF mass spectrometry protein analysis was performed as previously described[15] using a Microflex spectrometer (Bruker Daltonics, Leipzig, Germany). Twelve distinct deposits were made for strain MT10T from 12 isolated colonies. The 12 MT10T spectra were imported into the MALDI BioTyper software (version 2.0, Bruker) and analysed by standard pattern matching (with default parameter settings) against the main spectra of 4108 bacteria, including 120 spectra from the family Clostridiaceae, used as reference data, in the BioTyper database. A score enabled the identification, or not, from the tested species: a score >2 with a validated species enabled identification at species level, a score >1.7 but <2 enabled identification at genus level; and a score <1.7 did not enable any identification. No significant

| Characteristic | 1 | 2 | 3 | 4 | 5 | 6 |
|---------------|---|---|---|---|---|---|
| Cell diameter (μm) | 2.4 | 2.75 | 3.4 | 2.5 | 1.7 | 0.5 |
| Oxygen requirement | AAF | AS | AS | na | AS | AS |
| Gram stain | + | + | – | + | v | + |
| Motility | + | – | + | + | + | + |
| Endospore formation | – | – | na | + | + | + |
| Production of: | | | | | | |
| Alkaline phosphatase | + | na | na | na | na | na |
| Catalase | – | – | na | na | na | na |
| Oxidase | – | – | na | na | na | na |
| Urase | – | na | na | – | – | – |
| β-D-glucuronidase | w | na | na | na | na | na |
| Indole | – | – | – | + | na | na |
| Leucine arylamidase | w | na | na | + | na | na |
| Cystine arylamidase | – | na | na | na | na | na |
| Valine arylamidase | – | na | na | na | na | + |
| Utilization of: | | | | | | |
| Mannitol | – | + | + | – | + | – |
| Threitol | + | – | – | na | na | na |
| Sucrose | + | + | + | – | + | + |
| L-arabinose | – | – | – | – | – | + |
| D-sorbitol | – | – | – | na | na | na |
| D-xylose | – | – | + | + | + | + |
| D-ribose | + | + | + | + | + | + |
| D-glucose | + | + | + | + | + | + |
| D-fucose | + | + | + | + | + | + |
| D-mannose | + | + | + | + | + | + |
| D-maltose | + | + | + | + | + | + |
| D-fructose | + | + | + | + | + | + |
| Glycerol | + | – | – | – | – | – |
| N-Acetylglucosamine | + | na | na | na | na | na |
| G+C content (mol%) | 32.1 | 33.1 | 32 | 28 | 32.1 | 32.1 |
| Habitat | Human gut | Anaerobic reactor | Sediment | Pork meat | Human gut | Sediment |

Strains: (1) Desnuesiella massiliensis; (2) Clostridium amylopticum; (3) Clostridium drakei; (4) Clostridium algidicarnis; (5) Clostridium beijerinckii; (6) Clostridium sulfidogenes.

+, positive result; −, negative result; v, variable; w, weak positive result; na, data not available; AS, strictly anaerobic; AAF, facultatively anaerobic.
MALDI-TOF score was obtained for strain MT10 against the Bruker database, suggesting that our isolate was a new species. We incremented our database with the spectrum from strain MT10 (Fig. 3).

**Genome Sequencing Information**

The organism was selected for sequencing on the basis of its phylogenetic position and 16S rRNA similarity to other members of the family *Clostridiaceae*. It was the first genome of *Desnuesiella massiliensis* gen. nov., sp. nov. (CYSK00000000).

After DNA extraction by the phenol–chloroform method, genomic DNA of *Desnuesiella massiliensis* was sequenced using MiSeq Technology (Illumina Inc., San Diego, CA, USA) with the mate pair strategy.

For genome annotation, open reading frames (ORFs) were predicted using PRODIGAL (http://prodigal.ornl.gov) with default parameters but the predicted ORFs were excluded if they spanned a sequencing gap. The predicted bacterial protein sequences were searched against the Clusters of Orthologous Groups (COG) databases and the GenBank database [16] using BLASTP. The tRNAscanSE tool [17] was used to find tRNA genes whereas ribosomal RNAs were found by using Rnammer [18] and BLASTn against the GenBank database. Transmembrane helices and lipoprotein signal peptides were predicted using the Phobius web server [19]. ORFans were identified if their BLASTP E-value was lower than 1e-03 for alignment length >80 amino acids. If alignment lengths were <80 amino acids, we used an E-value of 1e-05.

The genome is 5 503 196 bp long with 32.09% GC content (Fig. 4 and Table 2). It is composed of 14 scaffolds (composed of 16 contigs). Of the 5308 predicted genes, 5227 were protein-coding genes, and 81 were RNAs (eight genes are 5S rRNA, five genes are 16S rRNA, one gene is 23S rRNA, 67 genes are tRNA genes). A total of 3890 genes (74.42%) were assigned as putative function (by COGs or by NR blast). In all, 276 genes were identified as ORFans (5.28%). The remaining genes were annotated as...
hypothetical proteins (843 genes, 16.13%). The distribution of genes into COGs functional categories is presented in Table 3.

The draft genome sequence of \textit{D. massiliensis} is smaller than those of \textit{Clostridium drakei} and \textit{Clostridium beijerinckii} (5.50, 5.64 and 6.00 Mb, respectively), but larger than those of \textit{Clostridium algidicarnis} and \textit{Clostridium sulfidigenes} (3.06 and 3.72 Mb, respectively).

**FIG. 4.** Graphical circular map of the chromosome. From outside to the centre: genes on forward strand (coloured by COG categories), genes on reverse strand (coloured by COG categories), RNA genes (tRNAs green, rRNAs red), GC content, GC skew.

**TABLE 2.** Nucleotide content and gene count levels of the genome

| Attribute                  | Value | % of total |
|----------------------------|-------|------------|
| Genome size (bp)           | 5 503 196 | 100 |
| DNA coding (bp)            | 4 723 383 | 85.8 |
| DNA G+C (bp)               | 1 765 744 | 32.1 |
| DNA scaffolds              | 14     | —          |
| Total genes                | 5308   | 100 |
| Protein-coding genes       | 5227   | 98.5 |
| RNA genes                  | 81     | 1.5 |
| Pseudo genes               | 13     | —          |
| Genes in internal clusters | 2212   | —          |
| Genes with function prediction | 3890   | 74.4 |
| Genes assigned to COGs     | 3873   | 74.1 |
| Genes with Pfam domains    | 4814   | 90 |
| Genes with signal peptides | 489    | 9.3 |
| Genes with transmembrane helices | 1309 | 25 |

*The total is based on either the size of the genome in base pairs or the total number of protein-coding genes in the annotated genome.

**TABLE 3.** Number of genes associated with the 25 general COG functional categories

| Code | Value | % of total | Description |
|------|-------|------------|-------------|
| A    | 0     | 0          | RNA processing and modification |
| B    | 1     | 0.02       | Chromatin structure and dynamics |
| C    | 232   | 4.44       | Energy production and conversion |
| D    | 60    | 1.15       | Cell cycle control, mitosis and meiosis |
| E    | 286   | 5.47       | Amino acid transport and metabolism |
| F    | 115   | 2.2        | Nucleotide transport and metabolism |
| G    | 379   | 7.25       | Carbohydrate transport and metabolism |
| H    | 154   | 2.95       | Coenzyme transport and metabolism |
| I    | 122   | 2.33       | Lipid transport and metabolism |
| J    | 321   | 6.14       | Translation, ribosomal structure and biogenesis |
| K    | 475   | 9.1        | Transcription |
| L    | 182   | 3.5        | Replication, recombination and repair |
| M    | 243   | 4.65       | Cell wall/membrane biogenesis |
| N    | 87    | 1.66       | Cell motility |
| O    | 182   | 3.48       | Post-translational modification, protein turnover, chaperones |
| P    | 183   | 3.5        | Inorganic ion transport and metabolism |
| Q    | 54    | 1.03       | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 436   | 8.34       | General function prediction only |
| S    | 218   | 4.17       | Function unknown |
| T    | 294   | 5.62       | Signal transduction mechanisms |
| U    | 47    | 0.9        | Intracellular trafficking and secretion |
| V    | 188   | 3.6        | Defence mechanisms |
| W    | 14    | 0.27       | Extracellular structures |
| Y    | 0     | 0          | Nuclear structure |
| Z    | 2     | 0.04       | Cytoskeleton |
| ⍗    | 1354  | 25.9       | Not in COGs |

*The total is based on the total number of protein-coding genes in the annotated genome.
TABLE 4. Genomic comparison of Desnuesiella massiliensis with five others members of the family Clostridiaceae*

| C. drakei | C. algidicarnis | C. beijerinckii | C. beijerinckii | C. sulfidigenes | D. massiliensis |
|-----------|----------------|----------------|----------------|----------------|---------------|
| 5748      | 747            | 818            | 911            | 761            | 910           |
| 61.97     | 62.01          | 62.50          | 62.57          | 68.84          | 68.69         |
| 57.32     | 61.78          | 62.06          | 61.57          | 68.69          | 69.39         |
| 5227      | 747            | 61.60          | 68.69          | 69.39          | 5227          |
| 3148      | 72.16          | 68.84          | 69.39          | 5227           |               |

*Numbers of orthologous proteins shared between genomes (above diagonal), AGIOS values (below diagonal) and numbers of proteins per genome (bold numbers).

The G+C content of D. massiliensis is smaller than those of C. drakei (32.09, 35.02%, respectively), but larger than those of C. algidicarnis, C. beijerinckii and C. sulfidigenes (30.26%, 29.86% and 30.00%, respectively).

The gene content of D. massiliensis is smaller than that of C. drakei (5227, 5748, respectively), but larger than those of C. algidicarnis, C. beijerinckii and C. sulfidigenes (2787, 5020 and 3148, respectively).

Table 4 summarizes the number of orthologous genes and the average percentage of nucleotide sequence identity between the different genomes studied. The nucleotide sequence identity of orthologous genes ranges from 57.3% to 69.9% among previously published genomes.

Conclusions

On the basis of phenotypic, phylogenetic and genomic analysis (taxonogenomics), we formally propose the creation of Desnuesiella massiliensis gen. nov., sp. nov., which contains the strain MT10T. This bacterium has been isolated from the digestive flora of a child living in Dakar, Senegal suffering from kwashiorkor.

Taxonomic and Nomenclatural Proposals

Description of Desnuesiella massiliensis gen. nov., sp. nov.
Desnuesiella (Des.nue.si.ell’a. ML. dim. suffix tella; M.L. fem. n. Desnuesiella named after the French bacteriologist Christelle Desnues, Aix-Marseille University, Marseille, France), massiliensis (mas.si.li.en’sis. L. gen. fem. n. massiliensis of Massilia, the Roman name of Marseille, France, where the type strain was isolated).

Desnuesiella massiliensis are Gram-positive rods, flagellated, motile, facultative anaerobic, mesophilic. Optimal growth is achieved at 37°C. Colonies are moderately opaque and approximately 1 mm in diameter on 5% sheep blood-enriched agar. Cells have a mean length of 2.4 μm and a mean width of 0.43 μm.

Desnuesiella massiliensis is catalase-, oxidase-, urease- and indole-negative but alkaline-phosphatase-positive. A positive reaction was obtained for the fermentation of sucrose, D-ribose, D-glucose, D-lactose, D-mannose, D-maltose, D-fructose, but not for D-arabinose, D-sorbitol and D-xylose.

Desnuesiella massiliensis was susceptible to amoxicillin, ceftriaxone, imipenem, erythromycin, doxycycline and rifampicin but resistant to colistin, gentamicin and metronidazole.

The G+C content of the genome is 32.09%. The 16S rRNA and genome sequences are deposited in GenBank under accession numbers LN846906 and CYSK00000000, respectively. The type strain MT10T (= CSUR P1918 = DSM 101500) was isolated from the stool of a child living in Dakar, Senegal suffering from kwashiorkor disease.

Acknowledgements

The authors thank Xegen company for automating the genome annotation process.

Transparency Declaration

The authors declare that they have no competing interests.

References

[1] Lagier JC, Armougou F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–93.
[2] Ramazamy D, Mishra AK, Lagier JC, Padmanabhan R, Rossi M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 2014;64:384–91.
[3] Bendjama E, Loucif L, Diene SM, Michelle C, Gacemi-Kirane D, Rolain JM. Non-contiguous finished genome sequence and description of “Bacillus massilioalgeriensis” sp. nov. Stand Genomic Sci 2014;9:1046–61.
[4] Hadjadj L, Bentorki AA, Michelle C, Amourea K, Djahoudi A, Rolain JM. Genome sequence and description of “Mannheimia massilio-gueldensis” sp. nov. New Microbes New Infect 2015;8:131–6.
Keita MB, Padmanabhan R, Caputo A, Robert C, Delaporte E, Razult D, et al. Non-contiguous finished genome sequence and description of “Gorillibacterium massiliense” gen. nov, sp. nov., a new member of the family Paenibacillaceae. Stand Genomic Sci 2014;9:807–20.

Skerman VBD, McGowan, Sneath PHA. Approved lists of bacterial names. Int J Syst Bacteriol 1980;30:225–420.

Wells CL, Wilkins TD. Clostridia: spore forming anaerobic bacilli. In: Baron S, et al., editors. Baron’s Medical Microbiology. 4th ed. University of Texas Medical Branch; 1996.

Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2015;33:152–5.

Tindall BJ, Rossello-Mora R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. Int J Syst Evol Microbiol 2010;60:249–66.

Lawson P, Dainty RH, Kristiansen N, Berg J, Collins MD. Characterization of a psychrotrophic Clostridium causing spoilage in vacuum-packed cooked pork: description of Clostridium algidicarnis sp. nov. Lett Appl Microbiol 1994;19:153–7.

Liou JS, Balkwill DL, Drake GR, Tanner RS. Clostridium carboxidivorans sp. nov., a solvent-producing clostridium isolated from an agricultural settling lagoon, and reclassification of the acetogen Clostridium scatologenes strain SLI as Clostridium drakei sp. nov. Int J Syst Evol Microbiol 2005;55:2085–91.

Sallam A, Steinbüchel A. Clostridium sulfidigenes sp. nov., a mesophilic, proteolytic, thiosulfate- and sulfur-reducing bacterium isolated from pond sediment. Int J Syst Evol Microbiol 2009;59:1661–5.

Song L, Dong X. Clostridium amylolyticum sp. nov., isolated from H2-producing UASB granules. Int J Syst Evol Microbiol 2008;58:2132–5.

Suresh K, Prakash D, Rastogi N, Jain RK. Clostridium nitrophenolicum sp. nov., a novel anaerobic p-nitrophenol-degrading bacterium, isolated from a subsurface soil sample. Int J Syst Evol Microbiol 2007;57:1886–90.

Seng P, Drancourt M, Gouriet F, La SB, 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.

Benson DA, Karlish-Mitarachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res 2012;40(Database issue):D48–53.

Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997;25:955–64.

Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–18.

Kall L, Krogh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids Res 2007;35(Web Server issue):W429–32.