TAXONOGENOMICS: GENOME OF A NEW ORGANISM

Genome analysis of *Lachnoclostridium phocaeense* isolated from a patient after kidney transplantation in Marseille

I. Dandachi1,3, H. Anani2, L. Hadjadj1, S. Brahimi1, J.-C. Lagier1, Z. Daoud4 and J.-M. Rolain1

1) Aix Marseille Université, IRD, APHM, MEPHI, 2) Aix Marseille Université, IRD, APHM, VITROME, IHU-Méditerranée Infection, Marseille, France, 3) Faculty of Medicine and Medical Sciences, Clinical Microbiology Laboratory, University of Balamand, Amioun, Beirut, Lebanon and 4) Clinical Microbiology and Infection Prevention, Michigan Health Clinics, Central Michigan University, Mount Pleasant, MI, USA

Abstract

*Lachnoclostridium phocaeense* is a new species in the genus *Lachnoclostridium*. *Lachnoclostridium phocaeense* is a Gram-positive anaerobic rod. This strain, Marseille-P3177T (CSUR = P3177) with the below described genome was isolated from the urine sample of a women after kidney transplantation. The strain genome is 3 500 754 bp long with 50.62% G + C content and consists of a single contig (GenBank accession number NZ_LT635479.1).

© 2021 The Authors. Published by Elsevier Ltd.

Keywords: Bacteria, culturomics, *Lachnoclostridium*, new species, taxonogenomics

Original Submission: 28 September 2020; Revised Submission: 22 February 2021; Accepted: 4 March 2021

Article published online: 16 March 2021

Corresponding author: J.-M. Rolain, Aix-Marseille Université, IRD, APHM, MEPHI, IHU Méditerranée Infection, Faculté de Pharmacie, 19–21 boulevard Jean Moulin, 13385, Marseille CEDEX 05, France. E-mail: jean-marc.rolain@univ-amu.fr

Introduction

*Lachnoclostridium* is a genus of Gram-positive, obligate anaerobic, spore-forming, motile bacteria. Organisms in this genus can grow in moderate ‘mesophilic’ as well as in extremely high ‘thermophilic’ temperatures, ranging from 20°C to 45°C and from 203°C to 633°C, respectively [1].

The *Lachnoclostridium* genus includes organisms from the Lachnospiraceae family and from several clostridial clusters such as Clostridium XIVa [1]. Clostridial cluster XIVa is known to make up a significant part of the human gut microflora [2]; it can exert anti-inflammatory effects and plays a role in homeostasis. In addition, via its components and metabolites, notably butyrate, clostridial cluster XIVa maintains intestinal health [3].

The human gut microbiota is a complex ecosystem that contains a variety of organisms including bacteria, fungi and viruses [4]. To explore this niche, bacterial cultures were used [5]; however, provided information on only the cultivable part of the human gut with a considerable fraction being uncultured. This is despite the advancement of molecular techniques such as metagenomics and 16S rRNA sequencing [6]. Recently, a new approach combining bacterial culturing under different conditions, matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF-MS) and 16s rRNA sequencing, named culturomics, was implemented. Compared with metagenomics, this approach allows the cultivation of species corresponding to previously unassigned sequences [7].

Using a previously described taxonogenomic approach [8,9] combined with culturomics, we present here the phenotypic and genomic characteristics of a *Lachnoclostridium* novel species isolated from a patient admitted to the hospital in Marseille. This is part of the culturomics project, which aims to detect and isolate new bacterial species. The new species was deposited in the *Collection de Souches de l’Unité des Rickettsies* (CSUR, WDCM 875) under the number P3177 [10].
Strain identification

*Lachnoclostridium* species, named *phocaeense* strain Marseille-P3177 had a unique spectrum upon identification with MALDI-TOF-MS on a Microflex LT spectrometer (Bruker Daltonics, Bremen, Germany). The reference spectrum obtained (Fig. 1) was imported into our database (http://www.mediterranee-infection.com/article.php?larub=280&titre=urms-database). The L. *phocaeense* 16S rRNA gene exhibited 94.6% similarity with *Lachnoclostridium contortum* strain ATCC 25540 [11], a phylogenetically close species (Fig. 2). The 94.6% value is lower than the gene sequence threshold of 98.7% 16S rRNA recommended by Stackebrandt and Ebers [12] to characterize an isolated strain as a new bacterial species without DNA–DNA hybridization.

Phenotypic and biochemical characterization

Strain Marseille-P3177 appears as translucent and whitish circular colonies with a diameter of 0.7–1 mm on a 5% sheep blood Columbia agar medium (BioMérieux, Marcy-l’Étoile, France). This species developed under anaerobic conditions at 37°C and for a period of 5 days of incubation [10].

Electron microscopy using GD6 and TechnaiG2 Cryo (FEI Company, Limeil-Brevannes, France) showed that *L. phocaeense* strain Marseille-P3177 is a Gram-positive bacillus (Fig. 3).

Biochemical characteristics of the isolated strain were determined using API ZYM and API 50CH (BioMérieux). Catalase assays (BioMérieux) and oxidase assays (Becton Dickinson, Le Pont de Claix, France) showed that this strain is oxidase and catalase negative. API ZYM revealed positive reactions for acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-galactosidase, β-galactosidase, α-glucosidase and β-glucosidase. On the other hand, using API 50CH, acid production was observed in the presence of starch (Table 1).

Antibiotic susceptibility testing was done using E-test (BioMérieux) performed on Mueller–Hinton agar supplemented with 5% blood (BioMérieux). Interpretation of the results was done according to the European Committee on Antimicrobial susceptibility testing 2018 (EUCAST). The strain was susceptible to amoxicillin, cefotaxime, ertapenem, imipenem,
FIG. 2. Phylogenetic tree analysis based on partial 16S ribosomal RNA sequences. Genbank accession numbers of partial 16S rRNA gene sequence are indicated in parenthesis. Sequences were aligned using CLUSTALW and the phylogenetic tree was obtained using the maximum likelihood bootstrap method and MEGA 7 software [20]. Numbers shown at the nodes are bootstrap percentages values obtained by 1000 times repetition analysis.
meropenem, vancomycin, teicoplanin, metronidazole, trimethoprim-sulfamethoxazole, rifampicin and gentamicin; but resistant to ciprofloxacin, fosfomycin, colistin, doxycycline, oxofloxacin and erythromycin.

**Genome sequencing**

Extracted genomic DNA of *L. phocaeense* P3177 was sequenced using MiSeq (Illumina, San Diego, CA, USA) with the mate-pair strategy. Assembly and annotation were performed with a pipeline of different softwares (SPADES [13], VELVET [11], SOAPDENOVO [14], trimmed (TRIMMOMATIC), MISEQ [15] software or untrimmed data (only MISEQ software) and XEGEN (http://www.xegen.fr/). To reduce assembly gaps, GAPCLOSER was used. Scaffolds with depth value < 25% of the mean depth and <800 bp were removed. Using different criteria (number of N, number of scaffolds and N50), the best assembly was selected. Genome coverage was 125×. The predicted bacterial protein sequences for *L. phocaeense* in addition to the five complete genomes of *Lachnoclostridium* available on NCBI were searched against the Clusters of Orthologous Groups (COG) database and NR database using BLASTP [16].

The degree of genomic similarity of Marseille-P3177 with closely related species was estimated using the ORTHOANI software [17]. Values among closely related species (Fig. 4) ranged from 67.07% between *Lachnoclostridium pacaense* and *Lachnoclostridium hylemonae* to 76.60% between *Lachnoclostridium bolteae* and *Lachnoclostridium pacaense*. When strain Marseille-P3177 was compared with these closely related species, values ranged from 67.66% with *L. saccharolyticum* to 72.53% with *L. scindens*.

**Genome description**

*Lachnoclostridium phocaeense* strain Marseille P3177 genome (GenBank accession number NZ_LT635479.1) is 3,500,754 bp long with 50.62% G + C content (Table 2). The genome

---

**TABLE 1. API 50CH and API ZYM biochemical tests of *Lachnoclostridium phocaeense* P3177**

| Test Variable | Result |
|---------------|--------|
| Control       | +      |
| Glycerol      | —      |
| Erythritol    | —      |
| d-Arabinose   | —      |
| L-Arabinose   | —      |
| D-Ribose      | —      |
| D-Xylose      | —      |
| D-Xylose      | —      |
| d-Adonitol    | —      |
| d-Galactose   | —      |
| d-Glucose     | —      |
| d-Fructose    | —      |
| d-Mannose     | —      |
| L-Sorbose     | —      |
| L-Rhamnose    | —      |
| Dulcitol      | —      |
| Inositol      | —      |
| d-Mannitol    | —      |
| d-Sorbitol    | —      |
| Methyl α-d-mannopyranoside | — |
| Methyl α-d-glucopyranoside | — |
| N-Acetyl-glucosamine | — |
| Amygdalin     | —      |
| Arbutin       | —      |
| Esculin       | —      |
| Salicin       | —      |
| d-Cellobiose  | —      |
| d-Maltose     | —      |
| d-Lactose     | —      |
| d-Melitose    | —      |
| d-Saccharose  | —      |
| d-Trehalose   | —      |
| Inulin        | —      |
| d-Melezitose  | —      |
| d-Raffinose   | —      |
| Sarch         | —      |
| Glycogen      | —      |
| Xyitol        | —      |
| Gentobiose    | —      |
| d-Turanose    | —      |
| d-Lysose      | —      |
| d-Tagatose    | —      |
| d-Fucose      | —      |
| d-Arabinol    | —      |
| l-Arabinol    | —      |
| Potassium gluconate | — |
| Potassium 2-ketogluconate | — |
| Potassium 5-ketogluconate | — |
| Alkaline phosphatase | — |
| Esterase (C4) | —      |
| Esterase lipase (C8) | — |
| Lipase (C14)  | —      |
| Leucine arylamidase | — |
| Valine arylamidase | — |
| Cystine arylamidase | — |
| Trypsin       | —      |
| d-Chymotrypsin| —      |
| Acid Phosphatase | + |
| Naphtholo-AS-BI-phosphohydrolase | + |
| d-Galactosidase | + |
| β-Galactosidase | + |
| β-Glucuronidase | — |
| d-Glucosidase  | —      |
| β-Glucosidase  | +      |
| N-Acetyl-glucosaminidase | — |
| d-Mannosidase  | —      |
| d-Fucosidase  | —      |
coverage was 125×. Of the 3382 predicted genes, 3315 were protein-coding genes and 67 were RNAs (four genes were 5S rRNA, four genes were 16S rRNA, four genes were 23S rRNA, 55 genes were tRNA genes). A total of 2328 genes (70.23%) were assigned as putative function (by COGs or by NR blast). A total of 170 genes were identified as ORFans (5.13%). The remaining genes were annotated as hypothetical proteins (719 genes, 21.69%). Gene distribution into COG functional categories of *L. phocaeense* are presented in Table 3. The distribution of genes in COG categories was similar in all six species of *Lachnoclostridium* (Fig. 5).

**TABLE 2.** Genes and nucleotides content of the *Lachnoclostridium phocaeense* genome

| Variant          | Number | % of the total |
|------------------|--------|----------------|
| Size (bp)        | 3,500,754 | 100.0         |
| G + C content (bp) | 1,772,172 | 50.6          |
| Total of genes   | 3382   | 100.0         |
| RNA genes        | 67     | 2.0           |
| Coding sequence size (bp) | 3,152,738 | 90.1        |
| Protein coding genes | 33,15 | 98.0          |
| Protein associated to COGs | 1,905 | 57.5          |
| Protein with peptide signal | 300 | 9.0           |
| Protein with transmembrane helices | 733 | 22.1          |
| Genes associated to mobilome | 1259 | 38.0          |
| Genes associated to virulence | 531 | 16.0          |

Abbreviations: COGs, clusters of orthologous groups.

Using the Bio-EDIT interface, a BLAST search was conducted against ARG-ANNOT, a database for acquired antibiotic

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

| Code | Value | % of total | Description                              |
|------|-------|------------|------------------------------------------|
| [J]  | 195   | 5.882353   | Translation                               |
| [A]  | 0     | 0          | RNA processing and modification          |
| [K]  | 201   | 6.0633483  | Transcription                            |
| [L]  | 107   | 3.227753   | Replication, recombination and repair    |
| [B]  | 0     | 0          | Chromatin structure and dynamics         |
| [D]  | 40    | 1.2066365  | Cell cycle control, mitosis and meiosis  |
| [Y]  | 0     | 0          | Nuclear structure                        |
| [V]  | 89    | 2.6847663  | Defence mechanisms                       |
| [T]  | 101   | 3.0467572  | Signal transduction mechanisms           |
| [M]  | 101   | 3.0467572  | Cell wall/membrane biogenesis            |
| [N]  | 12    | 0.3619909  | Cell motility                            |
| [W]  | 2     | 0.06033183 | Extracellular structures                  |
| [U]  | 28    | 0.8446456  | Intracellular trafficking and secretion  |
| [O]  | 78    | 2.3529413  | Post-translational modification, protein turnover, chaperones |
| [X]  | 48    | 1.4479638  | Mobilome: prophages, transposons         |
| [C]  | 111   | 3.3484166  | Energy production and conversion         |
| [E]  | 165   | 4.977355   | Amino acid transport and metabolism      |
| [F]  | 72    | 2.1719458  | Nucleotide transport and metabolism      |
| [H]  | 115   | 3.4690802  | Coenzyme transport and metabolism        |
| [I]  | 63    | 1.9004526  | Lipid transport and metabolism           |
| [P]  | 78    | 2.3529413  | Secondary metabolites biogenesis         |
| [Q]  | 24    | 0.7239819  | Secondary metabolites biogenesis         |
| [R]  | 179   | 5.3996983  | General function prediction only          |
| [G]  | 98    | 2.956295   | Function unknown                         |
| [S]  | 1410  | 42.533936  | Not in COGs                              |

Abbreviations: COGs, clusters of orthologous groups.
resistance genes (ARGs). The BLAST search was done under an e-value of $10^{-5}$, moderately stringent conditions for in silico ARG prediction [18]. ARG-ANNOT BLAST search revealed the presence of one resistance gene against tetracycline. This is in accordance with the antibiotic susceptibility testing performed, which showed that this strain was resistant to tetracycline. The bacteriocin database available in our research unit (Bacteriocins of the URMITE database BUR; available at http://drissfatima.wix.com/bacteriocins) was set up through the collection of all available sequences from NCBI and databases. Protein sequences from the aforementioned database allow the identification of bacteriocins from the human gut microbiota using BLASTp methodology [19]. Resistome analysis via this database showed the presence of 25 bacteriocin-associated genes.

**FIG. 5.** Functional distribution of COG categories in *L. phocaeense*, *L. hylemonae* (Genbank accession number NZ_CP036524.1), *L. pacaense* (Genbank accession number UOUF0100001.1), *L. saccharolyticum* (Genbank accession number NC_014376.1), *L. bolteae* (Genbank accession number NZ_CP022464.2) and *L. scindens* (Genbank accession number NZ_CP036170.1).

**Description of Lachnoclostridium phocaeense sp. nov.**

*Lachnoclostridium phocaeense* (pho.ca.e.en’se, L. neut. adj. pho-caeense, referring to the town Phocaea, the Latin name of the city that was later named Marseille, in France, where the type strain was first isolated). *Lachnoclostridium phocaeense* strain Marseille-P3177 is a new species in the genus *Lachnoclostridium* that was isolated from a 51-year-old woman’s urine sample after kidney transplantation in Marseille. The species’ optimal growth conditions are 37°C for 5 days under anaerobic conditions. Colonies are 0.7–1 mm in diameter on blood-supplemented agar. *Lachnoclostridium phocaeense* is a strictly anaerobic Gram-positive rod. It is also catalase and oxidase negative.
Funding
This work was supported by the French Government under the Investissements d’avenir (Investments for the Future) programme managed by the Agence Nationale de la Recherche (ANR, fr: National Agency for Research), (reference: Méditerranée Infection 10-IAHU-03).

Conflicts of interest
There are no conflicts of interest or financial disclosures for any authors.

Acknowledgement
We thank the genomics platform of the of the Institut Hospitalo-Universitaire (IHU)-Méditerranée Infection for their technical assistance.

References
[1] Yutin N, Galperin MY. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. Environ Microbiol 2013;15:2631–41.
[2] Lopetuso LR, Scaldaferri F, Petito V, Gasbarrini A. Commensal Clostridia: leading players in the maintenance of gut homeostasis. Gut Pathog 2013;5(1): 23-4749-5-23.
[3] Guo P, Zhang K, Ma X, He P. Clostridium species as probiotics: potentials and challenges. J Anim Sci Biotechnol 2020;11:24. eCollection 2020.
[4] Fraher MH, O'Toole PW, Quigley EM. Techniques used to characterize the gut microbiota: a guide for the clinician. Nat Rev Gastroenterol Hepatol 2012;9:312–22.
[5] Lagier JC, Edouard S, Pagnier I, Mediannikov O, Drancourt M, Raoult D. Current and past strategies for bacterial culture in clinical microbiology. Clin Microbiol Rev 2015;28:208–36.
[6] Turnbaugh PJ, Ley RE, Hamady M, Fraser-Ligget CM, Knight R, Gordon JL. The human microbiome project. Nature 2007;449(7164):804–10.
[7] Lagier JC, Khelaifi S, Alou MT, Ndongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microflora by culturomics. Nat Microbiol 2016;1:16203.
[8] Ramasamy D, Mishra AK, Lagier JC, Padmanabhan R, Rossi M, Sentaux 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.
[9] Fournier PE, Lagier JC, Dubourg G, Raoult D. From culturomics to taxonomogenomics: a new challenge for microbial identification. Microbiol Today 2015;64:384–91.
[10] Brahimi S, Cadoret F, Fournier PE, Moal V, Raoult D. Lachnoclostridium unimassiliense sp. nov. and ‘Lachnoclostridium phocaense’ sp. nov., two new bacterial species isolated from human urine after kidney transplantation. New Microbe. New Infect 2017;16:73–5.
[11] Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 2008;18:821–9.
[12] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:5–152.
[13] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012;19:455–77.
[14] Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 2012;1:1-18.
[15] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexibly trimmer for Illumina sequence data. Bioinformatics 2013;30:2114–20.
[16] Anani H, Abou Abdallah R, Chellaha N, Fontanini A, Ricaboni D, Malhe M, et al. Draft genome and description of Merdibacter massiliensis gen. nov., sp. nov., a new bacterium genus isolated from the human ileum. Sci Rep 2019;9. 7931-019-44343-8.
[17] Lee I, Ouk Kim Y, Park SC, Chun J. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2016;66:1100–3.
[18] Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 2014;58:212–20.
[19] Drissi F, Buffet S, Raoult D, Merhej V. Common occurrence of antibacterial agents in human intestinal microbiota. Front Microbiol 2015;6:441.
[20] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 2016;33:1870–4.