Massilibacterium senegalense gen. nov., sp. nov., a new bacterial genus isolated from the human gut

M. Tidjani Alou[1], J. Rathored[1], J.-C. Lagier[1], S. Khelaifia[1], N. Labas[1], C. Sokhna[1,2], A. Diallo[1,2], D. Raoult[1,3] and G. Dubourg[1]

1) *Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes IRD 198, CNRS 7278, Aix-Marseille Université, Marseille, France,* 2) *Campus Commun UCAD-IRD of Hann, Dakar, Senegal and* 3) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

Massilibacterium senegalense gen. nov., sp. nov., strain mt8^T, is the type strain of *Massilibacterium* gen. nov., a new genus within the *Bacillaceae* family. This Gram-negative facultative anaerobic rod was isolated from the gut microbiota of a severely malnourished boy. Its phenotypic description is hereby presented with a complete annotation of its genome sequence. This genome is 5 697 950 bp long and contains 5615 protein-coding genes and 178 RNA genes, among which are 40 rRNA genes.

Keywords: Culturomics, genome, gut microbiota, *Massilibacterium senegalense*, taxonogenomics
Original Submission: 1 December 2015; **Revised Submission:** 11 January 2016; **Accepted:** 14 January 2016
Article published online: 22 January 2016

Corresponding author: G. Dubourg, Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, INSERM 1095, Faculté de médecine, 27 Boulevard jean Moulin, 13385 Marseille cedex 05, France
E-mail: greg.dubourg@gmail.com
The first two authors contributed equally to this article, and both should be considered first author.

Introduction

The human gut microbiota is a complex and vast ecosystem harbouring eukaryotes, viruses, archaea and bacteria, these being by far the most abundant [1]. Its cell count is estimated to approximately 10^{14}, representing ten times the human somatic cell count, and its collective bacterial genome size is 150 times the size of the human genome [1–4]. The development of metagenomics has allowed a better exploration of gut microbiota by bypassing the noncultivable bacteria problem and unveiling links between altered gut microbiota and several diseases such as obesity, inflammatory bowel disease and irritable bowel syndrome [2]. It has also been demonstrated that the microbiota plays key roles in digestion and in immunologic and metabolic functions [2–4]. Nevertheless, a cultivation approach would be a complementary way to explore the gut microbiome in order to have a better representation of the viable population. In addition, it would allow further knowledge about the gut bacterial repertoire.

A new approach was developed in our laboratory in order to explore as exhaustively as possible the human gut microbiota by multiplying culture conditions with different atmospheres, media and temperatures [5]. This approach, known as culturomics, allowed us to isolate a new member of the *Bacillaceae* family. This family was created by Cohn in 1872 and consists of 52 validated genera (http://www.bacterio.net/). *Bacillus* is the type genus of this family, containing genera that are mostly aerobic or facultative anaerobic, rod-shaped, spore-forming, Gram-positive bacteria. These ubiquitous species are found in many ecosystems—mainly soil but also other environmental and clinical samples. Most *Bacillaceae* species are harmless, but some can be opportunistic pathogens, and *Bacillus anthracis,* the agent of anthrax, is well known to be pathogenic for humans [6].

Bacterial classification is currently based on phylogenetic relationships built on the 16S ribosomal RNA gene, phenotypic and genotypic characteristics including G+C content and DNA-DNA hybridization [7–9]. However, a great breakthrough has been

http://dx.doi.org/10.1016/j.nmni.2016.01.010

made in the last years in the area of genome sequencing, partly due to its decreasing cost. In fact, to this day, almost 70 000 genomes have been sequenced (https://gold.jgi.doe.gov/). With the development of this innovation, we proposed a new concept of bacterial description, including a proteomic description with a matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) profile [10,11] alongside a biochemical and genomic description of the new species [12–17].

We describe here a new member of this family, the genus *Massilibacterium,* isolated in the faeces of a patient with kwashiorkor. *Massilibacterium senegalense* is the type species (= CSUR P1510 = DSM 100455) of this new genus.

Materials and Methods

Organism information

As part of a culturomics study of the gut microbiota of children with severe acute malnutrition, a stool sample was collected from a 2-month-old Senegalese boy with kwashiorkor (body mass index, 14 kg/m^2) in April 2014. The patient was not treated with antibiotics at the time of sample collection; the sample was stored at −80°C. This study was authorized by the child's parents and was approved by the Institut Fédératif de Recherche 48 (Faculty of Medicine, Marseille, France) under agreement 09-022.

Strain identification by MALDI-TOF and 16S rRNA sequencing

Using the 18 culture conditions of the culturomics concept, the fecal sample was cultivated, and the obtained colonies were identified by MALDI-TOF as described below [5]. Proteomic analysis of our strain was carried out with MALDI-TOF as previously described [10,11]. A Microflex spectrometer (Bruker Daltonics, Leipzig, Germany) was used with a MTP 96 MALDI-TOF target plate (Bruker) on which 12 individual colonies were deposited. Twelve spectra were thus obtained, imported into MALDI BioTyper 2.0 software (Bruker) and analysed by standard pattern matching (with default parameter settings) against the main spectra of 7567 bacteria. Comparison with the BioTyper database spectra enabled the identification and discrimination of the analysed species from those in the database in accordance with the obtained score: a score >2 with a validated species enabled the identification at the species level, and a score <1.7 did not enable any identification. After a failed identification of the colony with a clean spectrum, it was identified by sequencing the 16S ribosomal RNA as previously described [18]. A threshold of 98.7% similarity level was determined to define a new species without performing DNA-DNA hybridization [19].

Growth conditions

In order to determine the ideal growth condition of *M. senegalense,* different growth temperatures (28, 30, 37, 45 and 56°C) were tested under anaerobic and microaerophilic atmospheres using GENbag anaer and GENbag miroaer systems respectively (bioMérieux, Marcy l'Étoile, France). The strain growth was also tested aerobically with or without 5% CO_2 supplementation.

Morphologic, biochemical and antibiotics susceptibility tests

The phenotypic characteristics (Gram staining, sporulation, motility, catalase, oxidase) were analysed as previously described [20]. Antibiotic susceptibility testing was performed using the disk diffusion method according to EUCAST 2015 recommendations (http://www.eucast.org/). Using API 20NE, API ZYM and API 50CH strips, we investigated the biochemical characteristics of the strain according to the manufacturer's instructions (bioMérieux). Electronic microscopy was performed with detection Formvar-coated grids which were deposited on a 40 µL bacterial suspension drop and incubated at 37°C for 30 minutes. Then followed a 10-second incubation on ammonium molybdate 1%. The grids were dried on blotting paper and finally observed with a Tecnai G20 transmission electron microscope (FEI Company, Limeil-Brevannes, France).

Genomic DNA preparation

M. senegalense strain mt8^T was cultured on 5% sheep's blood–enriched Columbia agar (bioMérieux) at 37°C aerobically. Bacteria grown on three petri dishes were resuspended in 4 × 100 µL of Tris-EDTA (TE) buffer. Then 200 µL of this suspension was diluted in 1 mL TE buffer for lysis treatment that included a 30-minute incubation with 2.5 µg/µL lysozyme at 37°C, followed by an overnight incubation with 20 µg/µL proteinase K at 37°C. Extracted DNA was then purified using three successive phenol–chloroform extractions and ethanol precipitations at −20°C overnight. After centrifugation, the DNA was resuspended in 160 µL TE buffer.

Genome sequencing and assembly

Genomic DNA (gDNA) of *M. senegalense* was sequenced on the MiSeq Technology (Illumina, San Diego, CA, USA) with the mate pair strategy. The gDNA was barcoded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). gDNA was quantified by a Qubit assay with a high sensitivity kit (Thermo Fisher Scientific, Waltham, MA, USA) to 66.2 ng/µL. The mate pair library was prepared with 1 µg of gDNA using the Nextera mate pair Illumina guide. The gDNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The