Taxono-genomics and description of *Haloimpatiens massiliensis* sp. nov., a new bacterium isolated from the gut of a healthy infant

H. Anani1,2, M. Tidjani Alou2,3, A. Fontanini2,3, D. Raoult2,3,4, J.-C. Lagier2,3 and P.-E. Fournier1,2

1) UMR VITROME, SSA, Aix-Marseille Université, IRD, AP-HM, IHU Méditerranée Infection, Marseille, France, 2) IHU Méditerranée Infection, 3) UMR MEPHI, Aix Marseille Université, IRD, AP-HM, IHU Méditerranée Infection and 4) Special Infectious Agents Unit, King Fahd Medical Research Centre, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract

A polyphasic taxono-genomic strategy was used to describe a new bacterium, strain Marseille-P1935; isolated from the gut of a healthy infant. 16S rRNA sequencing showed that the isolate belongs to the genus *Haloimpatiens* in the family *Clostridiaceae*. Phenotypic analysis and whole-genome sequence analyses confirm the status of the new species. We propose the creation of the new species *Haloimpatiens massiliensis* strain Marseille-P1935T (= CSURP1935T; = DSM100591T).

Corresponding author. P.E. Fournier, UMR VITROME, SSA, Aix-Marseille Université, IRD, AP-HM, IHU Méditerranée Infection, 19–21 Boulevard Jean Moulin, 13385, Marseille cedex 05. France. E-mail: pierre-edouard.fournier@univ-amu.fr

Introduction

Culturomics is a concept of developing different culture conditions in order to enlarge our knowledge of the human microbiota through the discovery of previously uncultured bacteria [1–4]. Once isolated, bacterial strains are characterized using the taxono-genomic approach that combines matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), phylogenetic analysis, main phenotypic description and genome sequencing [5,6].

Isolation and growth conditions

In 2014, we isolated an unidentified bacterial strain from a stool sample collected from a 7-month-old healthy girl from Niger with a weight-for-height z score of 0.15. The child’s parents gave an informed consent to participating in the study. Screening was made by MALDI-TOF MS on a Microflex LT spectrometer (Bruker Daltonics, Bremen, Germany) as previously described [7]. The obtained spectrum (Fig. 1) was imported into the MALDI Biotyper 3.0 software (Bruker Daltonics) and analysed against the reference spectra of bacteria included in two databases (Bruker and the constantly updated URMS databases (https://www.mediterranee-infection.com/urms-data-base/)). The study was validated by the ethics committee of the Institut Fédératif de Recherche 48 under number 09-022. The initial growth was obtained after 72 h of culture in 5% sheep-blood-enriched Columbia agar (bioMérieux, Marcy l’Étoile, France) in strict anaerobic conditions at 37°C medium.

Phenotypic characteristics

Colonies were irregular, large and greyish with a mean diameter of 8 mm. Bacterial cells were Gram-positive, rod-shaped, with a mean length of 2.99 μm and a mean width of 0.67 μm (Fig. 2). Strain Marseille-P1935 showed catalase-negative and oxidase-negative activities. API 50CH and ZYM were performed under strict anaerobic conditions at 37°C (Table 1).
Strain identification

The 16S rRNA gene was sequenced to classify this bacterium. Amplification was done using the primer pair fD1 and rP2 (Eurogentec, Angers, France) and sequencing used the Big Dye® Terminator v1.1 Cycle Sequencing Kit and ABI Prism 3130xl Genetic Analyzer capillary sequencer (Thermofisher, Saint-Aubin, France), as previously described [8]. The 16S rRNA nucleotide sequence was assembled and corrected using the CODEONCODE ALIGNER software (http://www.codoncode.com). In 2017, strain Marseille-P1935 was considered as a representative of a new bacterial species within a new genus for which the name Khelaifella massiliensis was proposed [9]. However, this name was not validly published and, in the present study, we obtained a 99.12% 16S rRNA nucleotide identity between strain Marseille-P1935 and Haloimpatiens lingqiaonensis strain JCM-19210T (GenBank accession number KC869664), the phylogenetically closest species with standing in nomenclature (Fig. 3) [10]. We consequently deduced that strain Marseille-P1935 was most likely a member of the Haloimpatiens genus. However, in a previous work, we had demonstrated that 16S rRNA-based nucleotide similarity is often unable to classify adequately bacterial isolates at the species level [11]. Therefore, genomic analyses were performed to verify the status of strain Marseille-P1935 within the genus Haloimpatiens.

Genome sequencing

Genomic DNA was extracted using the EZ1 biorobot (Qiagen, Courtaboeuf, France) with the EZ1 DNA tissue kit
and then sequenced on the MiSeq technology (Illumina, San Diego, CA, USA) with the Nextera Mate Pair sample prep kit (Illumina), as previously described [12]. The assembly was performed with a pipeline incorporating different softwares (VELVET [13], SPADES [14] and SOAP DENOVO [15]), on trimmed data (MiSeq and Trimomatic [16–19] softwares) or untrimmed data (only MiSeq software). GAPCloser was used to reduce assembly gaps. Scaffolds <800 bp and scaffolds with a depth value < 25% of the mean depth were removed. The best assembly was selected by using different criteria (number of scaffolds, N50, number of N). The genome of strain Marseille-P1935 was 4 214 685 bp long with a 29.8 mol% G + C content and 28 contigs. The degree of genomically similarity of strain Marseille-P1935 with closely related species was estimated using the ORTHOANI software [20–22]. Strain Marseille-P1935 exhibited an 87.18% ORTHOANI value with H. lingqiaonensis (Fig. 4). An ANI value < 95% between two strains suggests that they belong to distinct species. Using the Genome-to-Genome Distance Calculation software [23–25], the digital DNA-DNA hybridization (dDDH) of strain Marseille-P1935 with H. lingqiaonensis was 34.9% (Table 2). This value was lower than the 70% threshold used for delineating prokaryotic species. Finally, the phylogenetic tree based on the genome-to-genome distance showed that strain Marseille-P1935 and H. lingqiaonensis belonged to distinct species (Fig. 5). Therefore, these various analyses suggested that strain Marseille-P1935 could be classified as a member of a new species within the genus Haloimpatiens, in the family Clostridiaceae and the phylum Firmicutes.

**Conclusion**

Strain Marseille-P1935\(^1\) exhibiting ORTHOANI and dDDH values lower than 95% and 70%, respectively, with its phylogenetically closest species with standing in nomenclature, is consequently proposed as the type strain of the new species *Haloimpatiens massiliensis* sp. nov.

**Description of Haloimpatiens massiliensis sp. nov**

*Haloimpatiens massiliensis* (mas.si.li.en'sis; L. masc. adj. massiliensis for Massilia, the Roman name of Marseille, where strain Marseille-P1935\(^1\) was first isolated). Bacterial cells are Gram-positive, rod-shaped, with a mean length of 2.99 μm and a mean width of 0.67 μm. Colonies grown on 5% sheep-blood-enriched Columbia agar (bioMérieux) are irregular, large and greyish after 72 hours of incubation in a strict anaerobic atmosphere, with a mean diameter of 8 mm. Growth occurs at 37°C. Cells grow anaerobically only.

### Table 1. Biochemical tests of *Haloimpatiens massiliensis* sp. nov. (API 50 CH (a) and API ZYM (b))

| Test                         | Results (+/-) |
|------------------------------|---------------|
| **API 50 CH**                |               |
| (a) Bacterium: *Haloimpatiens massiliensis* sp. nov |               |
| Control                      | -             |
| Glycerol                    | -             |
| Erythrol                    | -             |
| α-arabinose                 | -             |
| α-arabinoose                | -             |
| α-ribose                    | -             |
| α-xylene                    | -             |
| α-xylose                    | -             |
| α-adonitol                  | -             |
| Methyl-β-c-xylopyranoside   | -             |
| α-galactose                 | +             |
| α-glucose                   | +             |
| α-fructose                  | +             |
| α-mannose                   | +             |
| α-sorbose                   | +             |
| α-rhamnose                  | +             |
| Dulcitol                    | +             |
| Linoctol                    | +             |
| α-mannitol                  | +             |
| α-sorbitol                  | +             |
| Methyl-α-d-mannopyranoside  | +             |
| Methyl-β-d-glucopyranoside  | +             |
| N-acetyl-glucosamine        | +             |
| Amygdaline                  | +             |
| Arbutine                    | +             |
| Esculine                    | +             |
| Salicine                    | +             |
| α-cellubiose                | +             |
| α-maltose                   | +             |
| α-lactose                   | +             |
| α-melibiose                 | +             |
| α-saccharose                | +             |
| α-trehalose                 | +             |
| Inuline                     | +             |
| α-inulose                   | +             |
| α-raffinose                 | +             |
| Amidon                      | +             |
| Glycogene                   | +             |
| Xylose                      | +             |
| Gentiobiose                 | +             |
| α-tartrate                  | +             |
| α-lysine                    | +             |
| α-tartarose                 | +             |
| α-fructose                  | +             |
| L-arabinose                 | +             |
| L-arabinose                 | +             |
| Potassium gluconate         | +             |
| Potassium 2-ketogluconate   | +             |
| Potassium 5-ketogluconate   | +             |
| **API ZYM**                 |               |
| (b) Bacterium: *Haloimpatiens massiliensis* sp. nov |               |
| Control                      | -             |
| Alkaline phosphatase        | +             |
| Esterase (C 4)              | +             |
| Esterase lipase (C 8)       | +             |
| Lipase (C 4)                | +             |
| Leucine arylamidase         | +             |
| Valine arylamidase          | +             |
| Cystine arylamidase         | +             |
| Tryptase                    | +             |
| α-chymotrypsine             | +             |
| Acid phosphatase            | +             |
| Naphthalo-AS-Bi-phosphohydrolase | +   |
| α-galactosidase             | -             |
| β-galactosidase             | -             |
| β-glucuronidase             | -             |
| α-glucosidase               | -             |
| β-glucosidase               | -             |
| N-acetyl-β-glucosaminidase  | -             |
| α-mannosidase               | -             |
| α-fucosidase                | -             |
Strain Marseille-P1935T exhibits catalase-negative and oxidase-negative activities. Using an API ZYM strip, positive reactions are observed for esterase, leucine arylamidase, naphthol-AS-BI-phosphohydrolase, alkaline and acid phosphatases, α-fucosidase and α-chymotrypsin but negative reactions are obtained for esterase lipase, lipase, valine arylamidase, cystine arylamidase, trypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, N-acetyl-β-glucosaminidase and α-mannosidase activities. Using an API 50 CH strip, strain Marseille-P1935 is able to metabolize D-galactose, D-maltose, D-tagatose, D-glucose, D-fructose, D-mannose, methyl-α-D-glucopyranoside, N-acetylglucosamine, D-lactose, D-saccharose, D-trehalose, D-turanose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-D-mannopyranoside, amygdalin, arbutin, esculin, salicin, D-cellobiose, D-melibiose, inulin, D-melezitose, D-raffinose, starch, glycogen, xylitol, gentiobiose, D-lyxose, D-fucose, L-fucose, D-arabitol, L-arabinose, potassium gluconate, potassium 5-keto-γ-gluconate and potassium 2-keto-glucuronate. However, negative reactions are obtained for glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, Kd-xylose, L-xylose, D-adonitol and methyl-β-D-xylopyranoside. The genome is 4 214 685-bp long and its G + C content is 29.8%.

The type strain, Marseille-P1935T, isolated from a stool sample collected from a 7-month-old healthy girl from Niger with a weight-for-height z score of 0.15, was deposited in the CSUR and DSMZ collections under accession numbers CSUR P1935 and DSM 100591, respectively.
FIG. 4. Heatmap generated with OrthoANI values calculated using the OAT software between Haloimpatiens massiliensis sp. nov. and other closely related species with standing in nomenclature.

TABLE 2. dDDH values obtained by comparison of all studied genomes

|       | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   |
|-------|------|------|------|------|------|------|------|------|------|------|
| C. argentinense | 100  | 28.2 | 22.2 | 27   | 27   | 19.4 | 24.2 | 21.6 | 20.8 | 21.5 |
| C. botulinum    | 100  | 24.1 | 77.7 | 47.3 | 21.7 | 35.4 | 20.7 | 20.9 | 22.4 |      |
| C. cellulovorans| 100  | 22.9 | 26.5 | 22.8 | 25.7 | 20.8 | 21.6 | 24.2 |      |      |
| C. combesi      | 100  | 46.5 | 19.8 | 35.3 | 20.3 | 20.4 | 21.4 |      |      |      |
| C. sporogenes   | 100  | 21.8 | 36   | 21.7 | 21.2 | 21.8 |      |      |      |      |
| C. tepidiprofundi| 100  | 21   | 19.8 | 19.3 | 21.2 | 21.8 |      |      |      |      |
| C. tepidum      |      | 100  | 19.5 | 21.5 | 21.2 |      |      |      |      |      |
| C. tetani       |      | 100  | 19.3 | 21.2 |      |      |      |      |      |      |
| H. lingqiaonensis|  100 |  100 |  100 |  100 |  100 |  100 |  100 |  100 |  100 |  100 |
| H. massiliensis |      |      |      |      |      |      |      |      |      |      |

1, Clostridium argentinense; 2, Clostridium botulinum; 3, Clostridium cellulovorans; 4, Clostridium combesi; 5, Clostridium sporogenes; 6, Clostridium tepidiprofundi; 7, Clostridium tepidum; 8, Clostridium tetani; 9, Haloimpatiens lingqiaonensis; 10, Haloimpatiens massiliensis.
Nucleotide sequence accession number
The 16S rRNA gene and genome sequences were deposited in GenBank under accession number LN850733 and FXWX00000001.1, respectively.

Deposit in culture collections
Strain Marseille-P1935^T was deposited in two different strain collections under number (= CSURP1935; = DSM100591).

Funding sources
This work was funded by the “Fondation Méditerranée Infection”, the National Research Agency under the programme Investissements d’avenir, reference ANR-10-IAHU-03 and by Région Provence Alpes Côte d’Azur and European funding FEDER PRIMI and FEDER BIOTEK.

Conflict of interest
None to declare.

Acknowledgements
The authors thank Aurelia Caputo for submitting the genomic sequences to GenBank. The authors also thank the Hitachi Company for providing a TM4000 electron microscope.
References

[1] Lagier JC, Armougom 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] Lagier JC, Hugon P, Khelai 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, Khelai S, Alou MT, Ndongo 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] Lagier JC, Edouard S, Pagnier I, Medianiokov O, Drancourt M, Raoult D. Current and past strategies for bacterial culture in clinical microbiology. Clin Microbiol Rev 2015;28:208–36.

[5] Fournier PE, Lagier JC, Dubourg G, Raoult D. From culturomics to taxonomogenomics: a need to change the taxonomy of prokaryotes in clinical microbiology. Anaerobe 2015;36:73–8.

[6] Ramasamy D, Mishra AK, Lagier JC, Padmanabhan R, Rossi M, Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, Morel AS, Dubourg G, Prudent E, Edouard S, Gouriet F, Casalta JP, Tidjani Alou M, Cadoret F, Brah S, Diop A, Sokha N, Mehrej V, 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.

[7] Morei AS, Dubourg G, Prudent E, Edouard S, Gouriet F, Casalta JP, et al. Complementarity between targeted real-time specific PCR and conventional broad-range 16S rRNA PCR in the syndrome-driven diagnosis of infectious diseases. Eur J Clin Microbiol Infect Dis 2015;34:561–70.

[8] Tidjani Alou M, Cadoret F, Brah S, Diao A, Soukha C, Mehrej V, et al. ‘Khelfafla massiliensis’, ‘Niameybacter massiliensis’, ‘Brachybacterium massiliens’, ‘Enterobacter amianensis’, ‘Massilibacillus massiliensis’, new bacterial species and genera isolated from the gut microbiota of healthy infants. New Microb New Infect 2017;19:1–7.

[9] Dildar Wu, Zheng N-F, Sun C, Zhang W-W, Han S-B, Pan J, et al. Halomaiptens ingoiaeni gen. nov., sp. nov., an anaerobic bacterium isolated from paper-mill wastewater. Int J Syst Evol Microbiol 2016;66:628–32.

[10] Rossi-Tamisier M, Be Amar S, Raoult D, Fournier PE. Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. Int J Syst Evol Microbiol 2015;65:1929–34.

[11] Diop A, Khelai S, Armstrong N, Labas N, Fournier PE, Raoult D, et al. Microbial culturomics unravels the halophilic microbiota repertoire of table salt: description of Gratillibacillus massiliensis sp. nov. Microb Ecol Health Dis 2016;27:23049.

[12] Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 2008;18:821–9.

[13] Bolger AM, Lohse M, Usadel B. Trimomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–20.

[14] Anani H, Hassn I, Zgheib R, Fadlane A, Raoult D, Fournier P-E. Whole-genome sequence of French clinical Peptoniphilus catoniae strain PB546. Microbiol Resour Announc 2019;8(45):e01241-19. Cuomo CA, editor.

[15] Anani H, Khodor M, Raoult D, Fournier P-E. Whole-genome sequence of French clinical Olivibacter jilunii strain PB502. Microbiol Resour Announc 2019;8(31):e00701–19. Stewart FJ, editor.

[16] Anani H, Raoult D, Fournier PE. Whole-genome sequence of Halomaiptens lngoiaeni strain PB956. Microbiol Resour Announc 2019;8(43). pii: e00699-19.

[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;66:100–3.

[18] Anani H, Guilhot E, Andreieu C, Fontanini A, Raoult D, Fournier PE. Prevotella ihumii sp. nov., a new bacterium isolated from a stool specimen of a healthy woman. New Microb New Infect 2019;32:10067.

[19] Anani H, Abou Abdallah R, Chelkha N, Fontanini A, Raoult D, Fournier PE. Prevotella ihumii strain P8956. Microbiol Resour Announc 2019;8(43). pii: e00699-19.

[20] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[21] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[22] Anani H, Khodor M, Raoult D, Fournier P-E. Whole-genome sequence of Peptoniphilus catoniae strain PB546. Microbiol Resour Announc 2019;8(45):e01241-19. Cuomo CA, editor.

[23] Anani H, Khodor M, Raoult D, Fournier P-E. Whole-genome sequence of French clinical Olivibacter jilunii strain PB502. Microbiol Resour Announc 2019;8(31):e00701–19. Stewart FJ, editor.

[24] Anani H, Raoult D, Fournier PE. Whole-genome sequence of Halomaiptens lngoiaeni strain PB956. Microbiol Resour Announc 2019;8(43). pii: e00699-19.

[25] 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 2015;65:3302–9.

[26] Anani H, Abou Abdallah R, Chelkha N, Fontanini A, Raoult D, Fournier PE. Prevotella ihumii sp. nov., a new bacterium isolated from a stool specimen of a healthy woman. New Microb New Infect 2019;32:10067.

[27] Anani H, Abou Abdallah R, Chelkha N, Fontanini A, Raoult D, Fournier PE. Prevotella ihumii strain P8956. Microbiol Resour Announc 2019;8(43). pii: e00699-19.

[28] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[29] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[30] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[31] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[32] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[33] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[34] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[35] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[36] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.

[37] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence database and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 2015;65:3302–9.