AIX-MARSEILLE UNIVERSITE
FACULTE DE MEDECINE DE MARSEILLE
ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

THESE DE DOCTORAT

Présenté par

BILEN Melhem

En vue de l'obtention du grade de **DOCTEUR de l'Université d'Aix-Marseille**
Spécialité : Pathologies humaines-Maladies infectieuses.

Sur le sujet suivant :

Description of the human gut microbiota by culturomics

Soutenue le 05 Juillet 2018

Membres du Jury

Professeur **Jean Christophe Lagier**	Président du jury
Professeur **Antoine Andremont**	Rapporteur
Professeur **Raymond Ruimy**	Rapporteur
Professeur **Didier Raoult**	Directeur de thèse
Professeur **Ziad Daoud**	Co-Directeur de thèse
Professeur **Dolla Sarkis**	Examinateur

Aix Marseille Univ, UMR MEPHI, Aix-Marseille Université,IRD, APHM, IHU Méditerranée-Infection, Marseille, France

Table of content/ Table des matières

Avant-propos

Le format de présentation de cette thèse correspond à une recommandation de la spécialité Maladies Infectieuses et Microbiologie, à l'intérieur du Master des Sciences de la Vie et de la Santé, qui dépend de l'Ecole Doctorale des Sciences de la Vie de Marseille.

Le candidat est amené à respecter des règles qui lui sont imposées et qui comportent un format de thèse utilisé dans le Nord de l'Europe et qui permet un meilleur rangement que les thèses traditionnelles. Par ailleurs, la partie introduction et bibliographie est remplacée par une revue envoyée dans un journal afin de permettre une évaluation extérieure de la qualité de la revue et de permettre à l'étudiant de commencer le plus tôt possible une bibliographie exhaustive sur le domaine de cette thèse. Par ailleurs, la thèse est présentée sur article publié, accepté ou soumis associé d'un bref commentaire donnant le sens général du travail. Cette forme de présentation a paru plus en adéquation avec les exigences de la compétition internationale et permet de se concentrer sur des travaux qui bénéficieront d'une diffusion internationale.

Professeur Didier RAOULT

Dedications

I dedicate my work to

- ✓ My Father Emile Bilen;
- ✓ My Uncle Nakhle Bilen;
- ✓ My family;
- ✓ My fiancée Christel Dagher;
- ✓ Lebanon.

Dédications

Je dédie mon travail à

- ✓ Mon père Emile Bilen;
- ✓ Mon oncle Nakhle Bilen;
- ✓ Ma famille;
- ✓ Ma fiancée Christel Dagher;
- ✓ Liban.

Acknowledgments

I would like to begin by expressing my sincere gratitude to my thesis director **Prof. Didier Raoult** who is a true pioneer and an inspiration to researchers in this field. His support, motivation and immense knowledge guided me all time through research and writing this thesis. It truly was an honour and privilege to have had him as my mentor.

My profound appreciation also extends to my thesis co-director **Prof. Ziad Daoud** who trusted me when he had no reason to. His patience, guidance and academic support gave me this wonderful opportunity and I hope that he now feels his trust was well placed.

I thank **Professor Jean-Christophe Lagier** for accepting to be the jury president and for all his academic support during my thesis.

I thank **Professor Antoine Andremont**, **Professeur Raymond Ruimy** for accepting to be members of the jury and reporters as well.

I thank **Professor Dolla Sarkis** for accepting to be member of the jury.

I thank my seniors **Dr. Grégory Dubourg** and **Dr. Fréderic Cadoret** for supporting and advising me during my thesis to reach the maximum possible outcome.

This work was made possible with the help of all the **Culturomics team** and my colleagues who supported me during my work.

I thank **Prof. Pierre-Edouard Fournier** for his kindness and continuous support in my work.

I thank the **IHU Méditerranée infection**, **AIX Marseille University** and **Balamand University** for hosting me during my thesis period, for their academic excellence and for the wonderful experience that i had the change to live during my stay.

I thank Adèle and Sami el Khoury for their continuous support during my stay in Marseille.

Remerciements

Je voudrais commencer par exprimer ma sincère gratitude à mon directeur de thèse, le **Prof. Didier Raoult**, qui est un véritable pionnier et une source d'inspiration pour les chercheurs dans ce domaine. Son soutien, sa motivation et son immense connaissance m'ont guidé à travers la recherche et l'écriture de cette thèse. C'était vraiment un honneur et un privilège de l'avoir comme mentor.

Ma profonde gratitude s'étend également à mon co-directeur de thèse, le **Prof. Ziad Daoud,** qui m'a fait confiance quand il n'y a aucune raison de faire. Sa patience, son soutien académique m'a donné cette merveilleuse opportunité, et j'espère qu'il sent maintenant que sa confiance était bien placée.

Je remercie le **Prof. Jean-Christophe Lagier** d'avoir accepté d'être le président du jury et pour tout son soutien académique durant ma thèse.

Je remercie le **Professeur Antoine Andremont**, le **Professeur Raymond Ruimy** d'avoir accepté d'être membres du jury et mes rapporteurs.

Je remercie le **Professeur Dolla Sarkis** d'avoir accepté d'être membre du jury.

Je remercie mes encadrants, le **Dr. Grégory Dubourg** et le **Dr. Fréderic Cadoret**, de m'avoir soutenu et conseillé pendant ma thèse pour atteindre le maximum de résultats possibles.

Ce travail a été rendu possible avec l'aide de toute l'équipe de **Culturomics** et de mes collègues qui m'ont soutenu durant mon travail.

Je remercie le **Prof. Pierre-Edouard Fournier** pour sa gentillesse et son soutien continu dans mon travail.

Je remercie **l'IHU Méditerranée**, **l'Université AIX Marseille** et **l'Université de Balamand** de m'avoir accueilli pendant ma période de thèse, pour leur excellence académique et pour l'expérience merveilleuse que j'ai eu la chance de vivre pendant mon séjour.

Je remercie Adèle et Sami el Khoury pour leur soutien continu pendant mon séjour à Marseille.

Summary

The human gut microbiota has been strongly correlated with human health and diseases and showed a potential in therapeutic developments. Since then, the scientific community has made a lot of efforts in describing its content by the means of molecular methods and specifically metagenomics. With this approach, scientists thought that culture would no longer be needed, disregarding the multiple drawbacks that can be faced such as depth bias, operational taxonomic units, incomplete genomic database, failure to distinguish dead from living organisms and inability to provide material for further experiments. Thus, culturomics was introduced with the aim of re-introducing culture in the field of microbiota description, relying on sophisticated culture methods and MALDI-TOF MS as a first line of identification. Culturomics proved its efficiency by its ability to report a significant number of new bacterial species in the gut and correlating its 16S rRNA sequences to operational taxonomic units. This emphasizes the importance of re-introducing culture techniques in describing the human gut microbiota, especially when only around 2,776 species have been isolated to date while it is approximated that up to 10^{12} bacteria may be present in 1g of stool. As the gut microbiota content depends on ethnical groups, the Pygmy population took our interest. Pygmy people are isolated ethnical groups, with particular lifestyle, diet and their digestive system represents an unknown bacterial reservoir. Thus, in the

present work, we aimed to describe the gut microbiota of Pygmy people by applying culturomics on the collected stool samples. 100,000 colonies have been isolated and tested by MALDI-TOF MS. Any unidentified organism has been tested by 16S rRNA gene sequencing for phylogenetic analysis. We succeeded in reporting 38 new species out of which, 25 (9 new genera and 16 new species) were validated with available 16S rRNA sequence on NCBI nucleotide database and 13 to be confirmed by whole genome sequencing analysis. Nevertheless, 20 of the 38 new species have been described in the taxono-genomics approach. Taxono-genomics is a recently adapted approach in the description of new bacterial species, where the phenotypic, proteomic, biochemical and genomic profiles are targeted to describe and confirm the identity of the organism of interest. While working on multiple stool samples, 195 different bacterial species were isolated and identified. Additionally, metagenomics was applied on the stool samples taken into consideration in this project and revealed that a significant number of the species isolated by culturomics were not recovered by metagenomics and that 59% of the identified OTUs were correlated to sequences corresponding to the isolated new species by culturomics (in the present study or other work). Thus, culturomics showed its efficiency in describing the human microbiota by reporting new bacterial species and correlating its sequences to previously considered un-classified sequences. Accordingly, culture remains a primary element in the human

gut microbiota description as well as a complementary technique to metagenomics. Extensive efforts should be done in order to unveil the dark matter of the human microbiota and its correlation with health and diseases, keeping in mind that most pathogenic bacteria were previously commensal. In addition, we worked on updating the present human prokaryotic repertoire. In 2015, only 2,172 different prokaryotic species were reported to have been isolated at least once from the human body as pathogens or commensals. We updated the previous database by reporting different species isolated from the human body to date, increasing it by 24% to reach a total of 2,694 species associated with human beings. Culturomics contributed up to 68.6% towards updating this repertoire by reporting 358 species, of which 242 were novel.

Key words: Culturomics, gut microbiota, Pygmy and taxono-genomics

Résumé

Le microbiote intestinal humain a été fortement corrélé avec la santé humaine et les maladies et a montré un potentiel dans les développements thérapeutiques. Depuis, la communauté scientifique a fait beaucoup d'efforts pour en décrire le contenu par les méthodes moléculaires et plus particulièrement la métagénomique. Avec cette approche, les scientifiques pensaient que la culture ne serait plus nécessaire, sans tenir compte des inconvénients multiples tels que le biais de profondeur, les unités taxonomiques opérationnelles, la base de données génomique incomplète, etc. La culturomics a été introduite dans le but de réintroduire la culture dans le domaine de la description du microbiote, en s'appuyant sur des méthodes de culture sophistiquées et sur la MALDI-TOF MS comme première ligne d'identification des organismes isolés. Cette technique a prouvée son efficacité par sa capacité à rapporter un nombre significatif de nouvelles espèces bactériennes dans l'intestin et à corréler leurs séquences à des unités taxonomiques opérationnelles (OTU). Cela souligne l'importance de réintroduire des techniques de culture dans la description du microbiote humain. Sachant que le contenu du microbiote intestinal dépend de groupes ethniques, la population Pygmée a pris notre intérêt. Les Pygmées sont des groupes ethniques isolés, avec un style de vie particulière, un régime alimentaire particulier et un système digestif qui représentent un réservoir bactérien inconnu. Dans le présent travail, nous avons cherché à décrire le

microbiote intestinal des Pygmées en appliquant la culturomics sur les échantillons de selles recueillies. 100,000 colonies ont été isolées et testées par MALDI-TOF MS. Tout organisme non identifié a été testé par séquençage du gène de l'ARNr 16S pour l'analyse phylogénétique. Nous avons réussi à isoler 38 nouvelles espèces dont 25 ont été validés avec la séquence d'ARNr 16S disponible sur la base de données de nucléotides du NCBI et 13 sont à confirmer par une analyse de séquençage du génome entier. 20 nouvelles espèces ont été décrites dans l'approche taxono-génomique. La taxono-génomique est une approche adaptée dans la description des nouvelles espèces bactériennes, où les profils phénotypiques, protéomiques, biochimiques et génomiques sont ciblés pour décrire et confirmer l'identité de l'organisme étudié. En travaillant sur plusieurs échantillons de selles, 179 espèces bactériennes différentes ont été isolées et identifiées. De plus, la métagénomique a été appliquée sur les échantillons de selles pris en compte dans ce projet et a révélé qu'un nombre significatif des espèces isolées par la culturomics n'ont pas été récupérées par métagénomique et que 59% des OTU identifiées étaient corrélées aux séquences correspondant à de nouvelle espèce isolée par Culturomics (dans la présente étude ou autre travail). Nos résultats renforcent le fait que culturomics est une technique dans la description du microbiote humain, que la culture reste un élément primordial dans la description du microbiote intestinal humain et une technique

complémentaire à la métagénomique. Des efforts considérables doivent être faits pour dévoiler la matière noire du microbiote humain, et sa corrélation avec la santé et les maladies, sachant que la plupart des bactéries pathogènes sont auparavant commensales. En plus de ce travail, nous avons travaillé sur la mise à jour du répertoire procaryote humain actuel. En 2015, seules 2 172 espèces procaryotes différentes ont été isolées au moins une fois dans le corps humain comme agents pathogènes ou commensaux. Nous avons mis à jour la base de données précédente en faisant état de différentes espèces isolées du corps humain à ce jour, l'augmentant de 24% pour atteindre un total de 2 694 espèces associées à des êtres humains. La Culturomics a contribué jusqu'à 68,6% à la mise à jour de ce répertoire en rapportant 358 espèces, dont 242 nouvelles.

Mot clés : Culturomics, microbiote digestif, Pygmée et taxono-genomics

Introduction

With up to 10^{12} bacteria estimated in 1g of stool, describing the human gut microbiota remains challenging since only around 2,776 species have been reported to be isolated from different body sites(1). Nevertheless, the human gut microbiota has been strongly correlated with health and diseases(2) and its composition was associated with obesity(3), malnutrition(4), non-alcoholic fatty liver diseases(5), Crohn's disease(6), colorectal cancer(7), gastric adenocarcinoma(8), etc.… Additionally, it has been used in therapeutic approaches development such as probiotics design(6,9,10), cancer immunotherapy (CTLA4 immunotherapy)(11) and faecal graft. Playing that essential role, describing the gut microbiota became mandatory. However, with the advances of sequencing techniques, particularly metagenomics, culture was thought be dispensable(12). Soon enough, multiple drawbacks have been faced such as the presence of operational taxonomic units (OTUs), incomplete genomics database, inability to provide material for further analyses, failure to detect certain pathogenic species and thus putting the scientific community in front of a "Dark matter" to be explored(12). Culturomics was designed with the vision of culturing what was previously called "un-cultivable bacteria" in order to complement metagenomics by deciphering the "Dark matter" of the human gut microbiota(13). This technique relies on applying sophisticated culture conditions on stool samples, in order to efficiently

isolate and describe the bacterial content of the gut microbiota(14). Isolated microorganisms are identified by MALDI-TOF MS (Matrix assisted laser desorption ionization time of flight mass spectrometry), which is a fast, efficient and cost effective technique. Unidentified organisms are subjected to 16S rRNA sequencing for phylogenetic analyses(13). In case of the isolation of a previously unknown bacterial species, taxono-genomics is used to describe and confirm its uniqueness(15). The latter is polyphasic approach, which aims to confirm the novelty of the isolated species based on its draft genome and proteomic analyses along with describing it (phenotypically and biochemically) via the standard laboratory methods(15). New species are reported in a New Species Announcement format in order to better communicate findings with the scientific community, followed by a complete descriptive paper(16). Being said, we aim in this project to study the gut microbial content of the Pygmy people by the means of culturomics and metagenomics. Knowing that the bacterial content and diversity are strongly correlated with the different ethnic groups, the Pygmy population took our interest as they have their own life style, diet and thus represent an unknown bacterial reservoir(17). In 2015, fecal samples were collected in Congo form healthy pygmy people of different ages according to Nagoya protocol from Republic of Congo. All the donors signed an informed consent prior of collection and an approval for the culturomics study was obtained (09-022) from the ethics committee

of the Institut Fédératif de Recherche 48. Samples were shipped from Congo to France in the specific protecting medium C-Top Ae-Ana (Culture Top, Marseille, France) and stored at -80°C at the laboratory (Marseille, France) for further analysis.

My objectives during my thesis in the research unit of Prof. Raoult was to learn the techniques of culturomics, which are currently part of the evolution of clinical microbiology, and to be able to import this approach to my country of origin, Lebanon. Prof. Raoult entrusted me with the analysis of the gut of Pygmies, a population living between the forest and the village, which might contain atypical bacteria. My efforts in this project allowed me to isolate 179 different bacterial species, including 38 new species with 27 confirmed. These species, according to Prof. Raoult, need to be reported individually within the framework of what has been defined as taxono-genomics, along with analyzing each time its genomic content. The latter lead to the generation of many scientific papers. Moreover, in the perspective of giving me a complete vision of the microbiota and its current state, Prof. Raoult asked me to carry out an exhaustive review of the bacterial repertoire of microbes isolated in humans (Microbiome article). In addition, I participated in a review requested to Prof. Raoult about the elements of culturomics, which I was one of the current co-authors that he considered the most apt to help in putting this work in shape. It is true that this abundance of publications, linked to the very rapid development of an

entirely new field, brings a plethora of publications. However, it should be noted that these publications are descriptive and correspond to a newly developed approach of a completely new strategy that led to the identification of 28% of the bacterial species currently known in the repertoire of the human bacteria. This lead to an unusual form that I apologize for. Moreover, I did not know how to present all the work that I had the chance to realize, so I separated it between the work where I contributed as a first author and others where I contributed with several people being involved (Those who sequenced the genomes, performed fatty acid analyses etc.).

In the present work, we are going to start by a literature review, summarizing all the efforts made in describing the human microbiota, role of culturomics in updating the human prokaryotic repertoire, the use of metagnomics and culturomics in describing the gut microbiota and importance of adapting culturomics for this matter.

In the second part, we report the results obtained by culturomics and metagenomics while working on the description of the gut microbiota of the Pygmy people along with performing a comprehensive analyses and taxono-genomics description of the isolated new bacterial species.

In the third part, we sum up with a conclusion and perspectives.

Introduction

Malgré une quantité estimée de 10^{12} bactéries par gramme de selles, la description du microbiote intestinal humain reste difficile avec environ seulement 2, 776 espèces isolées (1). Néanmoins, le microbiote intestinal humain a été fréquemment décrit comme associé à la santé ou à différentes pathologies chez l'homme (2). Sa composition a notamment été décrite comme impliquée dans l'obésité (3), la malnutrition (4), les stéatoses hépatiques non alcooliques (5), la maladie de Crohn (6), le cancer colorectal (7), les adénocarcinomes gastriques (8), etc .... En outre, le microbiote intestinal a été étudié pour le développement d'approches thérapeutiques telles que la conception de probiotiques (6,9,10), l'immunothérapie contre le cancer (CTLA4 immunothérapie) (11) et la greffe fécale. Avec ce rôle essentiel, décrire le microbiote intestinal est devenu obligatoire. Cependant, avec les progrès des techniques de séquençage, en particulier la métagénomique, on pensait que la culture était obsolète (12). Mais, récemment il a été montré que la métagénomique présentait de multiples inconvénients tels que la présence d'unités taxonomiques opérationnelles (OTU), des bases de données génomique incomplètes, l'impossibilité de fournir des analyses complémentaires, l'absence de détection de certaines espèces pathogènes et la mise en question de la "Dark matter" (12). La culturomics a été conçue dans le but de cultiver ce qui était auparavant considéré comme des «bactéries non

cultivables» afin de compléter les approches en métagénomique et en éclairant la «Dark matter» du microbiote intestinal humain (13). Cette technique repose sur l'application de conditions de culture sophistiquées et spécifiques à des échantillons de selles, afin d'isoler et de décrire efficacement le contenu microbien du microbiote intestinal (14). Les microorganismes isolés ont été identifiés par MALDI-TOF MS (Matrix assisted laser desorption ionization time of flight mass spectrometry), qui est une technique rapide, efficace et rentable. Les organismes non identifiés par cette technique ont été analysés par séquençage du gène de l'ARNr 16S permettant des analyses phylogénétiques (13). Dans le cas de l'isolement d'une espèce bactérienne jusqu'alors inconnue, l'approche taxono-génomique a été utilisée pour décrire et confirmer son statut de nouvelle espèce (15). La taxono-génomique est une approche polyphasique, qui vise à caractériser les nouvelles espèces isolées par analyse génomique, protéomique et phénotypique (15). Les nouvelles espèces ont donc été décrites selon un nouveau format de description (New species announcement) afin de palier à l'essor sans précédent de la découverte de nouvelles espèces. Ces publications sont ensuite suivies d'un document descriptif complet répondant aux standards classiques de description des nouvelles espèces auxquelles nous avons inclus une analyse génomique (16). Cela étant dit, nous visons dans ce projet à étudier le contenu microbien intestinal du peuple pygmée par le biais de la culturomics et de la

métagénomique. Sachant que le contenu bactérien et la diversité sont fortement corrélés avec les différents groupes ethniques, la population Pygmée a succité notre intérêt de par leur style de vie propre et leur régime alimentaire qui peuvent représenter un réservoir bactérien inconnu (17). En 2015, des échantillons de selles ont été collectés au Congo à partir de pygmées sains, d'âges différents en accord avec le protocole de Nagoya. Tous les donateurs ont signé un consentement éclairé avant la collecte et une approbation pour l'étude culturomics a été obtenue (09-022) du comité d'éthique de l'Institut Fédératif de Recherche 48. Des échantillons ont été expédiés du Congo vers la France dans le milieu de protection spécifique C-Top Ae-Ana (Culture Top, Marseille, France) et conservé à -80 ° C au laboratoire (Marseille, France) pour une analyse ultérieure.

Mon objectif en passant la thèse dans l'unité de recherche du Pr Raoult était d'apprendre les techniques de culturomics qui actuellement sont un élément d'évolution de la microbiologie clinique afin de pouvoir les importer dans mon pays d'origine, le Liban, dans lequel je suis chercheur biologiste. Mr Raoult m'a confié l'analyse des selles de pygmées qui représentent une population intermédiaire entre la forêt et le village et qui pouvaient contenir des bactéries atypiques. Mes efforts dans ce sens m'ont permis d'isoler 179 espèces bactériennes différentes dont 38 nouvelles espèces (27 confirmées). Ces espèces, d'après le Pr Raoult, ont toutes besoin d'être rapportées individuellement dans le cadre de ce qui a été défini comme la taxono-

génomique, avec la réalisation à chaque fois du génome de ces nouvelles espèces. Ceci a amené à écrire de très nombreuses publications. Par ailleurs, Mr Raoult, dans la perspective de me donner une vision complète, que je pourrais ensuite transférer au Liban, du microbiote et de son état actuel, m'a demandé de réaliser une revue exhaustive du répertoire bactérien des microbes connus chez l'homme, ce que j'ai réalisé pour la revue Microbiome, et de participer à une revue article qui lui avait été demandée sur les éléments de la culturomics dont j'étais un des membres actuels qu'il considérait comme étant le plus apte à aider à mettre ceci en forme. Il est vrai que cette abondance de publications, liée à l'émergence très rapide d'un champ entièrement nouveau, amène une pléthore de publications. Cependant, il faut noter que ces publications sont des publications descriptives qui correspondent à l'état d'émergence d'une stratégie totalement nouvelle et qui a amené à l'identification de 28% des espèces bactériennes actuellement connues dans le répertoire des bactéries humaines. Ceci entraine une forme inhabituelle dont je m'excuse. Par ailleurs, je ne savais pas comment présenter l'ensemble des travaux que j'ai eu la chance de réaliser, ainsi je les ai séparés entre les travaux dont je suis l'auteur principal et les travaux annexes auxquels j'ai participé, y compris la description de nouvelles espèces où souvent plusieurs personnes sont impliquées en fonction des manipulations réalisées.

Dans le présent travail, nous allons commencer par une revue de littérature,

résumant tous les efforts déployés pour décrire le microbiote humain, le rôle de la culturomics dans la mise à jour du répertoire procaryote humain, l'utilisation de la métagenomique et la contribution de la culturomics pour la description du microbiote intestinal.

Dans la deuxième partie, nous présentons les résultats obtenus par la culturomics et la métagénomique tout en travaillant sur la description du microbiote intestinal des pygmées et en réalisant une analyse complète et une description taxono-génomique des nouvelles espèces bactériennes isolées.

Dans la troisième partie, nous résumons ce travail avec une conclusion et des perspectives.

1. Hugon P, Dufour J-C, Colson P, Fournier P-E, Sallah K, Raoult D. A comprehensive repertoire of prokaryotic species identified in human beings. Lancet Infect Dis. 2015 Oct;15(10):1211–9.

2. Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. Curr Opin Gastroenterol. 2015 Jan;31(1):69–75.

3. DiBaise JK, Zhang H, Crowell MD, Krajmalnik-Brown R, Decker GA, Rittmann BE. Gut microbiota and its possible relationship with obesity. Mayo Clin Proc. 2008 Apr;83(4):460–9.

4. Million M, Tidjani Alou M, Khelaifia S, Bachar D, Lagier J-C, Dione N, et al. Increased Gut Redox and Depletion of Anaerobic and Methanogenic Prokaryotes in Severe Acute Malnutrition. Sci Rep [Internet]. 2016 Sep [cited 2017 Jun 11];6(1). Available from: http://www.nature.com/articles/srep26051

5. Abu-Shanab A, Quigley EMM. The role of the gut microbiota in nonalcoholic fatty liver disease. Nat Rev Gastroenterol Hepatol. 2010;7(12):691–701.

6. Lee BJ, Bak Y-T. Irritable Bowel Syndrome, Gut Microbiota and Probiotics. J Neurogastroenterol Motil. 2011 Jul 31;17(3):252–66.

7. Rajagopala SV, Vashee S, Oldfield LM, Suzuki Y, Venter JC, Telenti A, et al. The Human Microbiome and Cancer. Cancer Prev Res Phila Pa. 2017 Apr;10(4):226–34.

8. Nardone G, Compare D. The human gastric microbiota: Is it time to rethink the pathogenesis of stomach diseases? United Eur Gastroenterol J. 2015 Jun;3(3):255–60.

9. Kobyliak N, Conte C, Cammarota G, Haley AP, Styriak I, Gaspar L, et al. Probiotics in prevention and treatment of obesity: a critical view. Nutr Metab [Internet]. 2016 Dec [cited 2017 Jun 11];13(1). Available from: http://www.nutritionandmetabolism.com/content/13/1/14

10. Tidjani Alou M, Million M, Traore SI, Mouelhi D, Khelaifia S, Bachar D, et al. Gut Bacteria Missing in Severe Acute Malnutrition, Can We Identify Potential Probiotics by Culturomics? Front Microbiol [Internet]. 2017 [cited 2017 Jun 13];8. Available from: http://journal.frontiersin.org/article/10.3389/fmicb.2017.00899/full

11. Vétizou M, Pitt JM, Daillère R, Lepage P, Waldschmitt N, Flament C, et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. Science. 2015 Nov 27;350(6264):1079–84.

12. Greub G. Culturomics: a new approach to study the human microbiome. Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis. 2012 Dec;18(12):1157–9.

13. Lagier J-C, 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 Off Publ Eur Soc Clin Microbiol Infect Dis. 2012 Dec;18(12):1185–93.

14. Lagier J-C, Hugon P, Khelaifia S, Fournier P-E, 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 Jan;28(1):237–64.

15. Fournier P-E, Drancourt M. New Microbes New Infections promotes modern prokaryotic taxonomy: a new section “TaxonoGenomics: new genomes of microorganisms in humans.” New Microbes New Infect. 2015 Sep;7:48–9.

16. Fournier P-E, Raoult D, Drancourt M. Republication of «New

Species Announcement», a new format to prompt the description of new human microbial species. Hum Microbiome J. 2016 Sep 1;1:A1–2.

17. Gupta VK, Paul S, Dutta C. Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity. Front Microbiol [Internet]. 2017 Jun 23;8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5481955/

Part I: The human prokaryotic repertoire and the role of culturomics in deciphering its content.

Partie I : Le répertoire procaryotique humain et le rôle de la culturomics dans la description de son contenu.

Review: The role of culturomics in updating the repertoire of the isolated human bacterial species.

Published in Microbiome Journal

Briefing summary

After the creation of the first human prokaryotic repertoire in 2015, we report in this part an update, using the same methods, of the recent discoveries. We define the major techniques used in describing the human microbiota, and the importance of pursuing efforts to better understand its correlation with health and diseases. We provide a critical analysis of the data obtained and the distribution of the different species relative to multiple factors. Additionally, we enlighten the contribution and role played by culturomics in updating and expanding the previous repertoire. This part emphasizes on the importance of culturomics in describing the human microbiota by its ability to report the highest number of new bacterial species.

Résumé

Après la création du premier répertoire des procaryotes isolés chez l'homme en 2015, nous rapportons dans cette partie une mise à jour des découvertes récentes en utilisant les mêmes méthodes. Nous définissons les principales techniques utilisées pour décrire le microbiote humain, et l'importance de poursuivre ces efforts pour mieux comprendre son rôle. Nous fournissons une analyse critique des données obtenues et de la distribution des différentes espèces par rapport à plusieurs facteurs. De plus, nous discutons la contribution et le rôle joué par la culturomics dans la mise à jour et l'expansion de ce répertoire. Cette partie souligne l'importance de

la culturomics dans la description du microbiote humain par sa capacité à rapporter un grand nombre de nouvelles espèces bactériennes.

Bilen et al. Microbiome (2018) 6:94
https://doi.org/10.1186/s40168-018-0485-5

Microbiome

REVIEW Open Access

The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species

Melhem Bilen[1,2], Jean-Charles Dufour[3], Jean-Christophe Lagier[1], Fréderic Cadoret[4], Ziad Daoud[2], Grégory Dubourg[1] and Didier Raoult[1,5,6*]

Abstract

After a decade of research and metagenomic analyses, our knowledge of the human microbiota appears to have reached a plateau despite promising results. In many studies, culture has proven to be essential in describing new prokaryotic species and filling metagenomic gaps. In 2015, only 2172 different prokaryotic species were reported to have been isolated at least once from the human body as pathogens or commensals. In this review, we update the previous repertoire by reporting the different species isolated from the human body to date, increasing it by 28% to reach a total of 2776 species associated with human beings. They have been classified into 11 different phyla, mostly the *Firmicutes*, *Proteobacteria*, *Bacteroidetes*, and *Actinobacteria*. Finally, culturomics contributed up to 66.2% towards updating this repertoire by reporting 400 species, of which 288 were novel. This demonstrates the need to continue the culturing work, which seems essential in order to decipher the hidden human microbial content.

Keywords: Human microbiota, Culturomics, Gut, Isolation, Prokaryotes, New species

Background

This review reports the different bacterial species isolated at least once from the human being and emphasizes on the contribution of culturomics in unveiling and describing the human microbiota.

Introduction

Prokaryotes are known for their abundance, diversity, and presence in almost all types of niches, as well as their ability to live in different environmental conditions [1–3]. Despite more than a century of research in microbiology, estimating prokaryotic diversity remains a challenge. For instance, scientists estimate a range of between 10^7 and 10^{12} species per 1 g of stool, but few have been isolated by culture [4, 5]. However, the human microbiota composition has been strongly correlated with health and diseases where the balance of residing bacterial population was shown to be associated with an array of pathologies [6–13]. This highlighted the fact that the human microbiota can be used in therapeutic approaches, such as probiotic design or commensal replenishment [14–17] and the need to decipher its content for a better global understanding. Nevertheless, several approaches are currently being used in order to identify and describe the human microbiota. Culture is the oldest technique used for growing and isolating bacterial colonies. With the advances in sequencing techniques, particularly metagenomics, which targets the genetic material that can be recovered from any sample, the scientific community started believing that culture would no longer be needed [18, 19]. Yet, it soon appeared that these new methods have multiple drawbacks, such as depth bias and inability to detect, in some cases, the causative pathogenic bacteria such as the investigation into a shigatoxigenic *Escherichia coli* (STEC) O104:H4 outbreak [20] or *Campylobacter* and *Salmonella* in diarrhea cases [21]. In addition, incomplete genomic databases make it difficult to assign a precise taxonomic rank to a substantial number of sequences. Most importantly, these methods do not provide a pure culture of

* Correspondence: didier.raoult@gmail.com
[1]IRD, APHM, MEPHI, IHU Méditerranée Infection, Aix-Marseille Université, Marseille, France
[5]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia
Full list of author information is available at the end of the article

microorganisms, which is crucial for further strain characterization, in vitro models, or host-interaction studies [22]. Culturomics was introduced in order to optimize culture conditions and to show that the term "uncultivable organism" is misleading, since all microorganisms are cultivable using the right conditions and tools [18]. Besides the different culture conditions used, culturomics is coupled with an efficient, fast, and cost-effective identification methodology, based on MALDI-TOF MS and 16S rRNA gene sequencing, in the event of MALDI-TOF MS identification failure [18]. The complementarity between culture-independent and culture-dependent studies was well established, since only 15% of the detected species were concurrent for these two techniques [19, 23–25]. Culturomics has played an important role in the description of the human microbiota, particularly to fill the gaps of metagenomics by attributing several OTUs (operational taxonomic units) from other studies to newly isolated bacterial species [26].

A repertoire of all prokaryotes isolated by culture, at least once from the human body, was created in 2015 [27]. This review aims to update the repertoire of prokaryotes that have been isolated from different sites on the human body and complements the work published by Hugon et al. in 2015, who reported 2172 different species [27]. Additionally, it aims to highlight the contribution of culturomics in describing the human microbiota.

Culturomics: a strategy complementing metagenomics for describing the human gut microbiota

Metagenomic studies have reintroduced the study of the human gut microbiota through large amounts of generated data. Few of these studies have proved however to be significant [28–30] due to multiple factors, such as the inappropriate data analyses (i.e., bioinformatics, statistical analysis) or study design (i.e., inappropriate controls, small sampling size). Several studies have shown the ability to isolate bacterial species by culture, those very species that were not detected by high-throughput sequencing methods. For example, in 2013, Dubourg et al. showed that the number of bacterial species isolated by culture from stool samples may outrange the results of pyrosequencing [23]. Moreover, in some studies, the proportion of gram-positive to gram-negative bacteria, when comparing culture to sequencing results, revealed some differences [31–33]. Hayashi et al., in 2002, studied the digestive microbiota by the means of cloning/sequencing and anaerobic culture and revealed discrepancies between these approaches since the results were not equivalent [34]. Similar research comparing these two techniques when studying the human colonic microbiota also revealed that the reported species by culture or metagenomics were discordant with a significant number of bacterial species isolated by culture not identified by metagenomics [35]. In addition, Lagier et al., in 2012, have studied the bacterial content of three stool samples and highlighted an inconsistency between culturomics and pyrosequencing of 16S rRNA gene amplicons, where only 15% of the reported species were mutual between the two techniques [36]. These facts brought to interest the culture of microbiota with Lagier et al. in 2015. Using culturomics, 1057 prokaryotic species were identified, subsequently adding 531 species to the human gut repertoire, including 146 species which were previously isolated in humans but not in the gut, one archaea, 187 species which were not previously identified in humans, and 247 new species [26]. Additionally, taxonomic assignment, at the species level, was made possible for a significant number of unassigned sequences generated by previous metagenomic studies when comparing it to the genome sequences of the new species isolated by culturomics [26]. Thus, these two techniques complement one another and should be used together in order to more efficiently describe the human microbiota and reveal its dark matter.

Expanding the human bacterial repertoire

Methods

In 2015, our laboratory has previously established a repertoire of listing all bacteria isolated at least once from the human body between 1980 and February 16, 2015 (http://hpr.mediterranee-infection.com/arkotheque/client/ihu_bacteries/recherche/index.php) [27]. Thus, we have used the same methodology which included the three most sensitive queries selected, tested previously by a gold standard [27], to build a customized software using NCBI E-utilities to programmatically query the MeSH database, PubMed/Medline, and NCBI Taxonomy. With the means of this software, advanced searches on PubMed/Medline were made possible using a combination of NCBI taxonomy bacterial synonyms, MeSH keywords, abstract or title, and phylum information. Therefore, when using a specific bacterial species, the query pattern will allow us to retrieve the articles with their PMID that corresponds to its isolation from human [27]. In this review, we update the content of this repertoire, taking into consideration all the bacteria isolated by culture at least once from the human body and which have been reported by the List of Prokaryotic names with Standing in Nomenclature (LPSN) or NCBI taxonomy. Articles were collected from 2014 to April 17, 2018 and manually verified by title, abstract, and full text if needed, in order to confirm the results of the queries and that the article reports isolation of a bacterial species from humans. As for the pathogenicity pattern evaluation of the reported species in this repertoire, we investigated each species in the Risk Group Database

(https://my.absa.org/Riskgroups) and reported any missing species as "unclassified." Finally, in order to evaluate the oxygen requirement of the reported species, we used the database in http://www.mediterranee-infection.com/article.php?laref=374, and the taxonomic lineage was performed with the help of NCBI taxonomy (https://www.ncbi.nlm.nih.gov/taxonomy).

The human bacterial repertoire: where do we stand?

We added 604 different species to the previously reported human microbiota repertoire [27] (Additional files 1 and 2). This represents an increase of 28% in the number of species isolated from the human body [27] (Additional file 3). By adding our data to those reported in 2015 by Hugon et al., we reach a total of 2776 species isolated at least once from the human body (Fig. 1 and Additional file 3). Most of the added species belonged to the *Firmicutes*, *Actinobacteria*, and *Proteobacteria* phylum with 289 (47.76%), 137 (22.64%), and 99 (16.36%) species, respectively [27] (Table 1). The number of species isolated in *Chlamydiae*, *Deinococcus-Thermus*, *Lentisphaerae*, and *Verrucomicrobia* remained the same (Table 1). Interestingly, *Leptolyngbya ramose* and *Deferribacter desulfuricans* became the first members of the *Cyanobacteria* and *Deferribacteres* phylum, respectively, to be isolated from human beings in clinical cases [27] (Fig. 2, Additional files 1, 2, and 3, Table 1).

Nevertheless, when looking at the genus level, it is clear that *Mycobacterium* is the most common of all 602 unique genera of the 2776 species, with 151 different species belonging to it. *Mycobacterium* belongs to the *Actinobacteria* genus and is also known to cause several serious diseases in humans, including tuberculosis and leprosy [37, 38]. Recently, a novel *Mycobacterium* species (*Mycobacterium saopaulense*) was isolated from a clinical specimen from a patient who underwent LASIK surgery [39]. This shows the importance of further work on culture in order to complement our understanding of the different bacterial species that may be pathogenic but have not yet been isolated or reported. Our laboratory contributed to the identification of 13 new species of *Mycobacterium* [40–48] that have been isolated from the respiratory tract, blood, gut, femur bone, and skin. As for the most frequently occurring species epithet in this repertoire, these were *massiliensis* (162) and *timonensis* (40) (Fig. 3), both of which have been reported by our laboratory and assigned to the novel bacterial species isolated by culturomics.

The gut microbiota: richness and diversification

The largest human microbiota resides in the gut, and only the human colon is known to have a bacterial density of more than 10^{10} cells/g [49]. Previous studies have measured around 10 million different genes in the human gut microbiome [50] and more than 1000 species belonging to several phyla, with the most common being *Bacteroidetes*, *Firmicutes*, and *Actinobacteria*, while *Fusobacteria*, *Cyanobacteria*, *Proteobacteria*, and *Verrucomicrobia* are less frequently encountered [51, 52]. Of the 604 species reported in this study, 372 species were isolated in the gut,

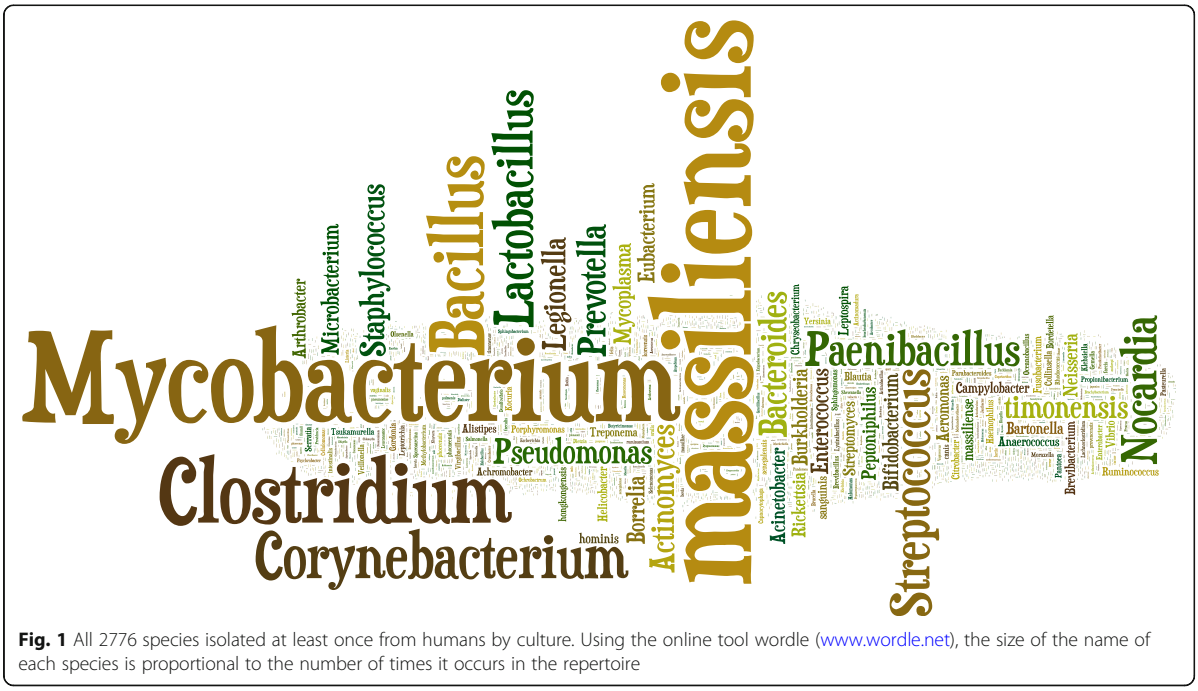


Fig. 1 All 2776 species isolated at least once from humans by culture. Using the online tool wordle (www.wordle.net), the size of the name of each species is proportional to the number of times it occurs in the repertoire

Table 1 Summary of all the results obtained in combination between Hugon et al.'s previous work and the present repertoire. Pathogenicity is represented as previously mentioned according to its risk group. Risk group 1 category refers to no or low individual and community risk, risk group 2 refers to species with moderate individual risk and a low community risk, risk group 3 represents species with high individual risk and a low community risk, risk group 4 being of a high individual and community risk, and unclassified refers to species that were not found in the risk group database

	Hugon et al., 2015 [27]	Present work
Risk group 1	2042	2042
Risk group 2	103	112
Risk group 3	27	28
Risk group 4	0	0
Unclassified	0	594
Strictly anaerobic total	386	662
Strictly anaerobic gut	244	459
Actinobacteria	558	696
Bacteroidetes	155	216
Chlamydiae	8	9
Cyanobacteria	0	1
Deferribacteres	0	1
Deinococcus-Thermus	1	1
Euryarchaeota	9	10
Firmicutes	676	962
Fusobacteria	25	29
Lentisphaerae	1	1
Proteobacteria	641	740
Spirochaetes	60	66
Synergistetes	4	7
Tenericutes	31	34
Verrucomicrobia	1	1
Unclassified	2	2
Total	2172	2776

of which 92.5% were isolated by culturomics, including 232 new species (Additional files 1 and 2 and Fig. 4).

The significant contribution of anaerobes

Because they are unable to tolerate or grow in the presence of oxygen, the distribution of anaerobic bacteria in the human body will vary by sites. However, it is expected and has been demonstrated that these organisms are more highly concentrated in the digestive tract compared to the skin, vagina, or oral cavities [53, 54]. Culturing and isolating these species is challenging and requires special conditions. There is also a slow replication rate. In recent years, various anaerobic methods have been improved, such as the Hungate roll-tube technique, anaerobic chambers, anaerobic jars, and the use of antioxidant agents [55, 56]. This study increases the number of anaerobic species isolated at least once from humans [27] to 662 (23.85%) by adding 276 (45.7%) anaerobic species, representing around half of the currently added species (Additional file 1). Therefore, when looking at the distribution of the anaerobic bacteria across the entire repertoire (2776), there are a total of 459 species in the gut, of which 215 are reported in this study (Table 1).

The revival of culture has paved the way for the exploration of the bacterial repertoire associated with human beings

Approaches to discover new bacteria

Our current knowledge of the repertoire of human-associated bacteria remains incomplete and includes only 2776 cultured pathogenic and commensal bacteria [27]. Nevertheless, culture allowed us to isolate 94 bacterial species from clinical samples, of which 43 were novel and 51 were known species that had not been reported to be isolated from humans [57, 58]. Moreover, our laboratory succeeded in using culture to isolate 25 environmental species that have also recently been considered to be associated with humans. Of these, 23 were new [44]. We also reported the correlation of 18 *Rickettsia* sp. to human infections, following its detection in ticks and fleas. Additionally, the ability to culture fastidious microorganism led to a significant evolution in infectious diseases following the Koch postulate, such as the cases of *Tropheryma whipplei* [59–61] or *Bartonella* species [62–67], where its culture enhanced diagnosis as well as our knowledge in terms of its host interactions. Likewise, 11 *Mycobacterial* species were isolated, including five pathogenic strains (*M. barassiae*, *M. bolletii*, *M. conceptionense*, *M. massiliense*, and *M. tahitimassiliense*) [44]. The ability to isolate, rather than simply obtain sequences of bacterial species, will allow researchers to further analyze their biological significance, features, and therapeutic potentials.

Following the study by Hugon et al., in 2015 [27], our laboratory was able to isolate and report 288 new species from different human body sites using enriched culture-based methods. Two hundred thirty-two new species were isolated from the human gut, of which 163 were anaerobic and 69 belonged to genera that can tolerate oxygen. Likewise, of the 14 new species isolated from the urinary tract, three were aerobic and 11 anaerobic. As for vaginal microbiota, a total of 13 new species were isolated, of which 11 were anaerobic. Finally, 10 aerobic and three anaerobic new species were isolated from the respiratory system, two aerobic and one anaerobic species from the skin, two aerobic species from a human abscess, one anaerobic species and three aerobic

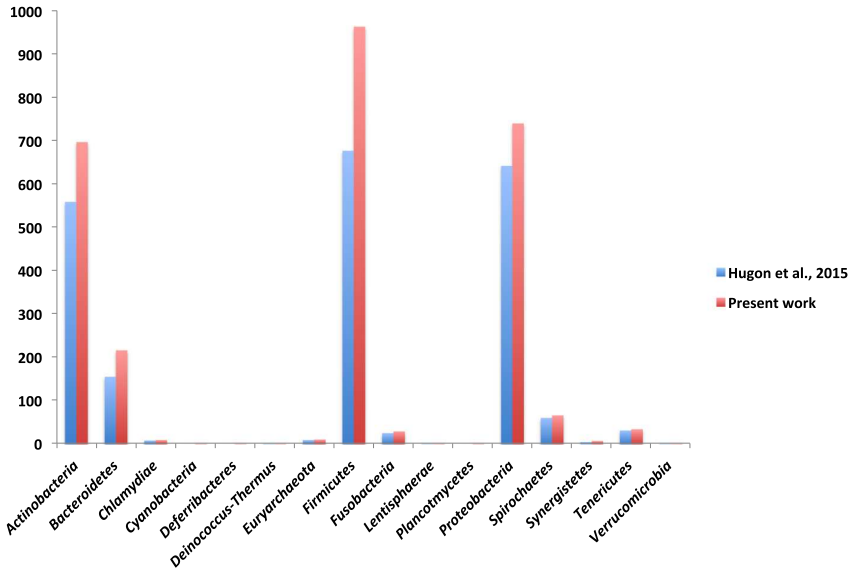


Fig. 2 Distribution of bacterial species reported in this repertoire according to their phylum, based on the NCBI taxonomy classification with the highest category being clustered in the *Firmicutes* phylum. *Cyanobacteria*, *Deferribacteres*, *Spirochaetes*, *Synergistetes*, and *Tenericutes* represent minority phyla in this repertoire, with 1, 1, 1, 3, 3, and 3 different species, respectively

species from the blood, two aerobic species from the cerebrospinal fluid, one aerobic species from a wound, one aerobic species from the peritoneal fluid, one aerobic species from a scalp pustule, two anaerobic species from the maternal colostrum, and one aerobic species from the foot of a patient with osteomyelitis.

The spectrum of these new species has now been added to the MALDI-TOF MS database at URMS database (http://www.mediterranee-infection.com/article.php?larub=280&titre=urms-database), being therefore available for clinical and microbiology laboratories using MS MALDI-TOF technology.

Taxonogenomics to fasten the reporting of the new isolated species

As previously mentioned, culturomics has succeeded in reporting and isolating a significant number of new species. It is, therefore, mandatory to describe, at the different levels, the main characteristics of these species. To do so, a new taxonogenomic approach has recently been adopted by our laboratory [68]. Taxonogenomics targets the proteomic, phenotypic, and genotypic traits of the bacterial species. Firstly, a bacterial species is considered to be novel when its MALDI-TOF score is less than 2 and its 16S rRNA gene sequence similarity to the closest phylogenetic species is less than 98.7%. When this happens, its unique proteomic spectrum, generated by MALDI-TOF MS, is added to Bruker's database and its 16S gene rRNA sequence is submitted in the NCBI nucleotide database for a faster identification in the future. The genome of the new bacterial species is then sequenced for annotation and comparison with close species in terms of DNA G+C content, size, gene content, percent of coding sequences, gene distribution in COG categories [69], number of genes coding for RNA, mobile genetic elements, transmembrane helices, signal peptides, and others if needed. The extent of genetic similarity between the compared bacterial isolates is also

Fig. 3 All 2776 species isolated at least once from humans using culture. Using the online tool wordle (www.wordle.net), the size of the name of each species is proportional to the number of times it occurs in the database. In this image, only the species name was taken into consideration, excluding the genus

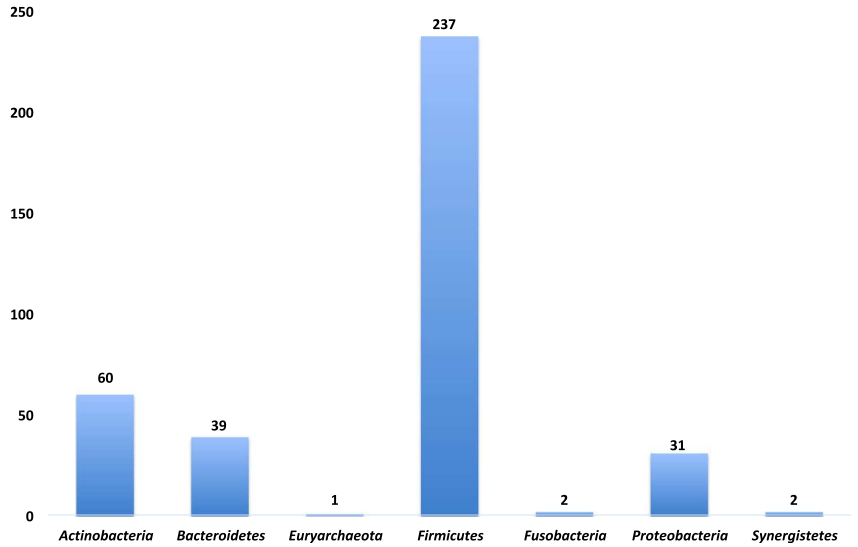


Fig. 4 Distribution of bacterial species isolated from the gut using culturomics, according to the phylum to which they belong, based on the NCBI taxonomy classification, with the highest clustering in the *Firmicutes* phylum

assessed by the Average of Genomic Identity of Orthologous Gene Sequences (AGIOS) and the establishment of the digital DDH with the help of the GGDC online software (http://ggdc.dsmz.de/distcalc2.php). New bacterial species are also tested for their ability to grow at different temperatures and environments, their resistance towards several common antibiotics, their sporulation ability, and their different biochemical characteristics.

This approach has enabled the description of over 146 new species, with an average of three species per month since April 2018. We now report new isolated species using a new format, known as the new species announcement (NSPA) [70]. The rate of identification shows no signs of plateau, and the NSPA has been proved to be efficient in terms of reporting when compared to the typical descriptive format. For instance, waiting for the complete description and validation of the name of the newly isolated species is considered to be far too long to be accessible to the scientific community when compared to NSPA, which takes around 3 months after submission to be published online. The NSPA reveals the isolation site of the new species, its growth conditions [70], the main phenotypic characteristics (gram, size, oxidase, and catalase), phylogenetic positioning among closely related species with standing in nomenclature, the MALDI-TOF spectrum, and the 16S rRNA gene sequences. Since the launch of the NSPA format in 2016, our laboratory has successfully reported over 100 novel species prior to their complete descriptions, with the average number of species reported per month being approximately three times higher than those in the standard descriptive papers. This enables faster communication of information within the scientific community, especially when, as previously mentioned, a significant number of novel species are reported by culturomics (Fig. 5).

Prokaryotes in humans and their pathogenicity

Pathogenicity is a complex mechanism that depends on the ability of a microorganism to cause diseases due to several factors such as virulence, inoculum, and host factors. Defining a microorganism as "pathogenic" appears to be very difficult in several cases, since a commensal species might become pathogenic when the opportunity arises [71]. *Escherichia coli* is the perfect example; this bacterium is usually considered commensal when found in the human intestine but can become pathogenic if it reaches extra-colonic sites [72]. Similarly, polymicrobial and nosocomial infections are both agent- and host-dependent and should also be taken into account. *Staphylococcus aureus* is frequently detected during polymicrobial infections and can also play a commensal or even a protective role during infections [41, 42, 73]. In order to assess the pathogenicity of the reported bacteria in this repertoire, we adapted the same method as Hugon et al., grouping species based on their biological risk of infection [27]. Only 10 of the 604 species, added in this current study, were reported in the risk group database. Nine belonged to risk group 2, defined by a low community risk and a relative individual infection

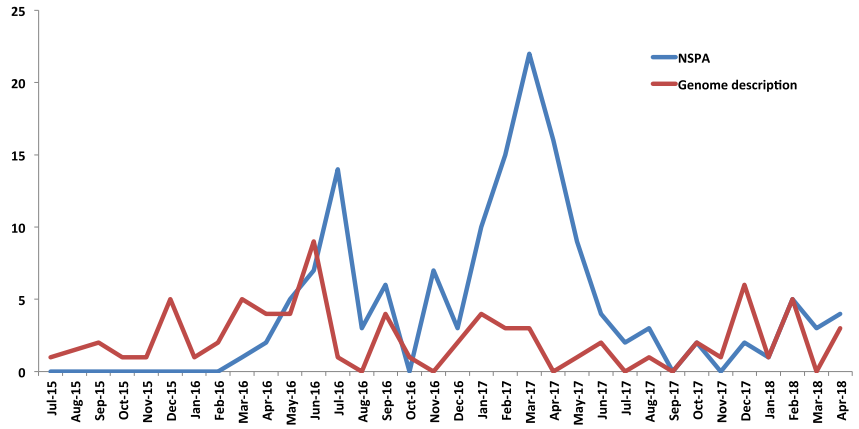


Fig. 5 Number of reported species isolated for the first time in humans by culturomics using the new species announcement format compared to the full description format

risk, and one belonged to risk group 3, defined by a low community risk and a high individual risk.

The lack of information towards the pathogenicity of the remaining species could be explained by the fact that most of them were new species that had not previously been isolated or had been isolated for the first time recently in humans. This fact remains of concern and highlights the need to create a database of all pathogenic bacteria and clinical cases that may be reported in the literature but are not yet easily accessible. Finally, in order to get an idea about the clinical occurrence of the species reported in our repertoire, which might represent some level of awareness in terms of virulence and pathogenicity, we investigated the PubMed database for articles reporting its isolation in clinical samples or clinical cases using an advanced search for articles containing the name of the bacteria (with its synonyms) with clinical, patient, or disease. Excluding new species isolated by culturomics, 148 species were reported to have been isolated in clinical specimens or clinical cases, of which eight were the cause of bacteremia (Additional files 1 and 2). After examining the isolated clinical species, we found that 58 (39%) were novel and had not been previously reported nor isolated from humans or from any other site. In addition, in 2017, Lagier et al. reported the isolation and identification of 12 bacterial species through culturomics, under different pathogenic conditions [44]. This prompted us to continue describing the human microbiota using culturomics, thus demonstrating its effective contribution towards clinical sample evaluations [26]. The fact that such a significant number of new species were detected in a clinical context highlights the need to move beyond the notion that "one bacteria causes one disease" [74]. As a perspective, high-throughput sequencing and extensive culture methods applied on clinical specimen could substantially expand the spectrum of prokaryotes involved in infectious diseases.

Metagenomics and culturomics: complementarity rather than contradiction

Human microbiota communities have been largely studied by metagenomics, which succeeded in correlating its content and diversity to health and diseases. For example, acne has been linked to alterations in the cutaneous microbiome [75] in addition to the dominance of *Firmicutes* relative to *Actinobacteria*, in psoriatic lesions [6]. Furthermore, the detection of *Helicobacter pylori* in the gastric microbiota has been correlated with age-related diseases [7] and has been found to increase the risk of developing several health issues, including gastric adenocarcinoma and peptic ulcers [8]. Colonic microbiota are currently being studied to correlate changes with colorectal cancer [9], and the gut microbiome has been suggested to play an important role in inflammatory bowel diseases such as Crohn's disease, where a decrease in obligate anaerobes belonging to the *Firmicutes* phylum was observed, along with an increase in facultative anaerobes belonging to the *Enterobacteriaceae* family [10]. Moreover, gut microbiota composition may be involved in hepatologic conditions such as non-alcoholic fatty liver disease (NAFLD) [11], besides having a certain role in the obesity phenotype [12, 13]. All these examples have led to the mass analyses of the bacterial content of the microbiota during diseases and have helped to create hypotheses for researchers on the possibility of developing new bacteriotherapy approaches, which can only be achieved whenever the different species taken into consideration are cultured and available for further experimentations [74].

Indeed, intensive culture methods have led the discovery of human commensals, which were thereafter

found to be microbes of medical importance in high-throughput sequencing-based analyses, making both methods complementary. For instance, the depletion of *Faecalibacterium prausnitzii* was correlated with inflammatory bowel diseases (IBD), including colitis and Crohn's disease [76]. In addition, *Akkermansia muciniphila*, cultivated only in 2004 from the human gut [77], is now considered a probiotic [78] as many metagenomic studies have demonstrated its association with metabolic disorders such as obesity or diabetes [14, 79]. Moreover, using a simplified culture workflow, the role of *Clostridium butyricum* in necrotizing enterocolitis was demonstrated by a case study using culturomics [80]. This points out that culture and metagenomic approaches should be grasped in a circular pattern, even though culture-based microbiota studies are difficult to apply to an entire cohort and time consuming.

This complementarity has also been demonstrated in an elegant study dedicated to show the influence of the gut microbiota composition on the efficiency of PD-1-based immunotherapy against several tumors [81]. While shotgun sequencing of fecal microbiota evidenced overrepresentation of *Akkermansia muciniphila* in responders when compared to non-responders, the culturomics approach highlighted an association between the presence of *Enterococcus hirae* and the response to PD-1 treatment. Of note, a previous strain of *Enterococcus hirae*, currently considered an oncobiotic, was cultivated from mouse spleen nodes [16]. Finally, the proof of concept of this work was definitely established in a murine model, in which an association of *E. hirae and A. muciniphila* bacterial strains were administered. As a result, efficacy of PD-1 treatment was reordered in mice previously receiving antibiotics.

Challenging operational taxonomic units

As previously mentioned, when it comes to the molecular profiling of the human microbiome, OTUs are one of the major challenges. OTUs are simply the results of the higher diversity obtained through the description of the human microbiome by metagenomics [33, 34]. However, the higher microbial diversity in molecular methods which results in a significant number of OTUs does not pose any difficulties, particularly as its reference bacteria can be cultured when the proper conditions are met [18]. In fact, significant efforts have been recently made to reveal the identity of the OTUs and to decipher the dark matter of the human microbiota using culturomics. This has succeeded in demonstrating the ability of culture to report a significant proportion of the human gut microbiota, while maintaining its overall community structure [26, 82–84]. Similarly, in a study using the concept of the culture-enriched molecular profiling of the human gut, the ability of culture to recover the majority of OTUs when after-culture sequencing was implemented has been shown, thus confirming that culture is not only complementary to molecular approaches but is also essential to the description of the human microbiome [85].

Conclusion

The study of the human microbiota is more than ever a major challenge as its implication in health and diseases has exceeded our expectations. For instance, the concept that human microbes can modulate response to anticancer therapy [16, 81] or be involved in HIV transmission [86] was difficult to imagine few years ago. These substantial advances were made possible thanks to the conjoint advances in culture, sequencing methods, and bioinformatic analysis. Herein, we highlight the considerable efforts recently made for culturing new microbes. Thus, a total of 2776 species isolated from the human body at different sites has been reached, of which 604 are reported in the present work. This represents a substantial increase of 28% within only 3 years. Culturomics contributed up to 66.2% in updating the previous repertoire and demonstrated the fundamental role of commensal culturing in describing and unveiling the hidden part of the human microbiota [26]. Beyond the impact of culturomics, this substantial increase is part of the current rebirth of culture in the field of microbiology that we are witnessing. As a matter of fact, Browne et al. recently cultured 68 new taxa, including new families and genera, while cultivating fecal specimen from six healthy individuals [87]. To the best of our knowledge, these taxa were not yet characterized. Interestingly, 37 of these new taxa belong, with different degrees of priority, to the "most wanted taxa" list. The latter was established as a part of the Human Microbiome Project (HMP) following characterization of bacterial communities at several body sites among 200 healthy volunteers and corresponded to the organisms of which its genome sequences are missing [88]. These microbes are underrepresented in culture collections, and we believe that high-throughput culture methods should, with time, enable their recovery. This points out the need of an updated repertoire of the bacteria cultured from the human beings but also raises several issues. Indeed, the pace of the human new species isolation in microbiota descriptive studies has recently accelerated, as shown in this study, but the delay of its genome's integration into the available genomic databases is incompatible with the intense research dedicated in this field. Even if we have accelerated the publication of these new isolated taxa through the new species announcement format, the time for official validation of these new species slows its integration despite having custom databases that can be built to capture the optimal diversity among assigned

sequences [26]. This also evidences the need to include genome sequencing as a part of the description of new species, at least if they were human isolated.

Taken together, these elements should encourage keeping a repertoire of prokaryotes associated with human beings real-time updated. If the vast majority of the bacteria newly isolated are in fact commensals, some can be further recovered as pathogens, as exemplified by *Akkermansia muciniphila* which has been recently isolated from blood cultures. Only time will allow adjudicating whether prokaryotes included in such a repertoire should be considered as commensals, pathogens, or bacteria passing through the human being and thus allowing a better interpretation of human microbiome studies [89].

Additional files

Additional file 1: General characteristics of the reported species in this repertoire along with its classification based on oxygen tolerance (0 being able to tolerate oxygen and 1 strictly anaerobic), risk group, potential pathogenicity, and the PMID of the article. A: New species isolated by culturomics approach but not in the gut. B: Species isolated without culturomics approach and not isolated in the gut. C: New species isolated not using culturomics approach. D: Species isolated by routine culture in our laboratory. E: Species isolated in the human gut by culturomics. F: New species isolated in the human gut by culturomics. G: Species isolated in the human gut and without using culturomics. (XLSX 40 kb)

Additional file 2: This table represents, in sheet 1, the taxonomic classification of the bacterial species according to NCBI taxonomy (https://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/) and, in sheet 2, all the species classification according to their phylum and that have not been added yet in NCBI taxonomy database. (XLSX 514 kb)

Additional file 3: This table represents all the bacterial species isolated at least once from the human being, classified according to its phylum and combining the present work and the previous species reported by Hugon et al., in 2015. (XLSX 73 kb)

Funding
This study was funded by the Mediterranée-Infection Foundation.

Availability of data and materials
All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Authors' contributions
MB investigated the literature, collected data, analyzed data, and wrote the review. J-CD designed the search methodologies and contributed to the acquisition and analyses of mass data. J-CL contributed to the critical revision of the manuscript. FC was responsible for the acquisition and interpretation of data. ZD contributed in the critical revision of the manuscript. GD contributed to the data analysis, review design, and critical revision. DR investigated the literature, collected data, analyzed data, and designed and wrote the review. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Author details
[1]IRD, APHM, MEPHI, IHU Méditerranée Infection, Aix-Marseille Université, Marseille, France. [2]Clinical Microbiology Laboratory, Faculty of Medicine and Medical Sciences, University of Balamand, PO Box: 33, Amioun, Lebanon. [3]INSERM, IRD, SESSTIM UMR912, Aix-Marseille Université, Marseille, France. [4]Assistance Publique Hôpitaux de Marseille, BIOSTIC Service Biostatistique et Technologies de l'Information et de la Communication, Hôpital de la Timone, Marseille, France. [5]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia. [6]Microbes, Evolution, Phylogeny and Infections (MEPHI), AMU, IRD, Institut Hospitalo-Universitaire Méditerranée-Infection, Aix-Marseille Université, 19-21 Boulevard Jean Moulin, 13385 Marseille CEDEX 5, France.

Received: 20 July 2017 Accepted: 18 May 2018
Published online: 24 May 2018

References

1. Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: the unseen majority. Proc Natl Acad Sci U S A. 1998;95:6578–83.
2. Antranikian G, Vorgias CE, Bertoldo C. Extreme environments as a resource for microorganisms and novel biocatalysts. Adv Biochem Eng Biotechnol. 2005;96:219–62.
3. Horner-Devine MC, Carney KM, Bohannan BJM. An ecological perspective on bacterial biodiversity. Proc Biol Sci. 2004;271:113–22.
4. Curtis TP, Sloan WT, Scannell JW. Estimating prokaryotic diversity and its limits. Proc Natl Acad Sci U S A. 2002;99:10494–9.
5. Locey KJ, Lennon JT. Scaling laws predict global microbial diversity. Proc Natl Acad Sci U S A. 2016;113:5970–5.
6. Patel RV, Lebwohl M. In the clinic. Psoriasis. Ann Intern Med. 2011;155:ITC2-1–ICT2-15. quiz ITC2-16
7. Gao Z, Tseng C, Strober BE, Pei Z, Blaser MJ. Substantial alterations of the cutaneous bacterial biota in psoriatic lesions. PLoS One. 2008;3:e2719.
8. Atherton JC, Blaser MJ. Coadaptation of *Helicobacter pylori* and humans: ancient history, modern implications. J Clin Invest. 2009;119:2475–87.
9. McColl KEL. Clinical practice. *Helicobacter pylori* infection. N Engl J Med. 2010;362:1597–604.
10. Plottel CS, Blaser MJ. Microbiome and malignancy. Cell Host Microbe. 2011; 10:324–35.
11. Rigottier-Gois L. Dysbiosis in inflammatory bowel diseases: the oxygen hypothesis. ISME J. 2013;7:1256–61.
12. Abu-Shanab A, Quigley EMM. The role of the gut microbiota in nonalcoholic fatty liver disease. Nat Rev Gastroenterol Hepatol. 2010;7:691–701.
13. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006;444:1027–31.
14. Kobyliak N, Conte C, Cammarota G, Haley AP, Styriak I, Gaspar L, et al. Probiotics in prevention and treatment of obesity: a critical view. Nutr Metab [Internet]. 2016;13 Available from: http://www.nutritionandmetabolism.com/content/13/1/14. Cited 11 Jun 2017
15. Vétizou M, Pitt JM, Daillère R, Lepage P, Waldschmitt N, Flament C, et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. Science. 2015;350:1079–84.
16. Daillère R, Vétizou M, Waldschmitt N, Yamazaki T, Isnard C, Poirier-Colame V, et al. *Enterococcus hirae* and *Barnesiella intestinihominis* facilitate cyclophosphamide-induced therapeutic immunomodulatory effects. Immunity. 2016;45:931–43.
17. Million M, Tidjani Alou M, Khelaifia S, Bachar D, Lagier J-C, Dione N, et al. Increased gut redox and depletion of anaerobic and methanogenic prokaryotes in severe acute malnutrition. Sci Rep [Internet]. 2016;6 Available from: http://www.nature.com/articles/srep26051. Cited 11 Jun 2017
18. Lagier J-C, Hugon P, Khelaifia S, Fournier P-E, 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.
19. Lagier J-C, Million M, Hugon P, Armougom F, Raoult D. Human gut microbiota: repertoire and variations. Front Cell Infect Microbiol. 2012;2:136.
20. Loman NJ, Constantinidou C, Christner M, Rohde H, Chan JZ-M, Quick J, et al. A culture-independent sequence-based metagenomics approach to the investigation of an outbreak of Shiga-toxigenic *Escherichia coli* O104:H4. JAMA. 2013;309:1502–10.

21. Singh P, Teal TK, Marsh TL, Tiedje JM, Mosci R, Jernigan K, et al. Intestinal microbial communities associated with acute enteric infections and disease recovery. Microbiome. 2015;3:45.
22. Greub G. Culturomics: a new approach to study the human microbiome. Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis. 2012;18:1157–9.
23. Dubourg G, Lagier JC, Armougom F, Robert C, Hamad I, Brouqui P, et al. The gut microbiota of a patient with resistant tuberculosis is more comprehensively studied by culturomics than by metagenomics. Eur J Clin Microbiol Infect Dis. 2013;32:637–45.
24. Pfleiderer A, Lagier J-C, Armougom F, Robert C, Vialettes B, Raoult D. Culturomics identified 11 new bacterial species from a single anorexia nervosa stool sample. Eur J Clin Microbiol Infect Dis. 2013;32:1471–81.
25. Dubourg G, Lagier JC, Robert C, Armougom F, Hugon P, Metidji S, et al. Culturomics and pyrosequencing evidence of the reduction in gut microbiota diversity in patients with broad-spectrum antibiotics. Int J Antimicrob Agents. 2014;44:117–24.
26. Lagier J-C, Khelaifia 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.
27. Hugon P, Dufour J-C, Colson P, Fournier P-E, Sallah K, Raoult D. A comprehensive repertoire of prokaryotic species identified in human beings. Lancet Infect Dis. 2015;15:1211–9.
28. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. Nature. 2006;444:1022–3.
29. Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters gut microbial ecology. Proc Natl Acad Sci U S A. 2005;102:11070–5.
30. Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, et al. Metagenomic analysis of the human distal gut microbiome. Science. 2006; 312:1355–9.
31. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 2007;449:804–10.
32. Gossling J, Slack JM. Predominant gram-positive bacteria in human feces: numbers, variety, and persistence. Infect Immun. 1974;9:719–29.
33. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, et al. Diversity of the human intestinal microbial flora. Science. 2005;308:1635–8.
34. Hayashi H, Sakamoto M, Benno Y. Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods. Microbiol Immunol. 2002;46:535–48.
35. Wilson KH, Blitchington RB. Human colonic biota studied by ribosomal DNA sequence analysis. Appl Environ Microbiol. 1996;62:2273–8.
36. Lagier J-C, 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 Off Publ Eur Soc Clin Microbiol Infect Dis. 2012;18:1185–93.
37. Cook GM, Berney M, Gebhard S, Heinemann M, Cox RA, Danilchanka O, et al. Physiology of *Mycobacteria*. Adv Microb Physiol [Internet]. 2009:81–319. Elsevier. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0065291109055027. Cited 11 Jun 2017
38. Pinheiro RO, de Souza Salles J, Sarno EN, Sampaio EP. *Mycobacterium leprae*? Host-cell interactions and genetic determinants in leprosy: an overview. Future Microbiol. 2011;6:217–30.
39. Whipps CM, Tortoli E, Matsumoto CK, Martin A, Droz S, Leão SC, et al. *Mycobacterium saopaulense* sp. nov., a rapidly growing mycobacterium closely related to members of the *Mycobacterium chelonae*? *Mycobacterium abscessus* group. Int J Syst Evol Microbiol. 2015;65:4403–9.
40. Adékambi T, Berger P, Raoult D, Drancourt M. rpoB gene sequence-based characterization of emerging non-tuberculous mycobacteria with descriptions of *Mycobacterium bolletii* sp. nov., *Mycobacterium phocaicum* sp. nov. and *Mycobacterium aubagnense* sp. nov. Int J Syst Evol Microbiol. 2006;56:133–43.
41. Ben Salah I, Cayrou C, Raoult D, Drancourt M. *Mycobacterium marseillense* sp. nov., *Mycobacterium timonense* sp. nov. and *Mycobacterium bouchedurhonense* sp. nov., members of the *Mycobacterium avium* complex. Int J Syst Evol Microbiol. 2009;59:2803–8.
42. Adékambi T, Stein A, Carvajal J, Raoult D, Drancourt M. Description of *Mycobacterium conceptionense* sp. nov., a *Mycobacterium fortuitum* group organism isolated from a posttraumatic osteitis inflammation. J Clin Microbiol. 2006;44:1268–73.
43. Adékambi T, Reynaud-Gaubert M, Greub G, Gevaudan M-J, La Scola B, Raoult D, et al. Amoebal coculture of "*Mycobacterium massiliense*" sp. nov. from the sputum of a patient with hemoptoic pneumonia. J Clin Microbiol. 2004;42:5493–501.
44. Lagier JC, Drancourt M, Charrel R, Bittar F, Ranque S, Raoult D. Many more microbes in humans: enlarging the microbiome repertoire. Clin Infect Dis Press. 2017;65:S20–9.
45. Nouioui I, Carro L, Teramoto K, Igual JM, Jando M, Del Carmen Montero-Calasanz M, et al. *Mycobacterium eburneum* sp. nov., a non-chromogenic, fast-growing strain isolated from sputum. Int J Syst Evol Microbiol. 2017;67:3174–81.
46. Paniz-Mondolfi AE, Greninger AL, Ladutko L, Brown-Elliott BA, Vasireddy R, Jakubiec W, et al. *Mycobacterium grossiae* sp. nov., a rapidly growing, scotochromogenic species isolated from human clinical respiratory and blood culture specimens. Int J Syst Evol Microbiol. 2017;67:4345–51.
47. Shahraki AH, Trovato A, Mirsaeidi M, Borroni E, Heidarieh P, Hashemzadeh M, et al. *Mycobacterium persicum* sp. nov., a novel species closely related to *Mycobacterium kansasii* and *Mycobacterium gastri*. Int J Syst Evol Microbiol. 2017;67:1766–70.
48. Davidson RM, DeGroote MA, Marola JL, Buss S, Jones V, McNeil MR, et al. *Mycobacterium talmoniae* sp. nov., a slowly growing mycobacterium isolated from human respiratory samples. Int J Syst Evol Microbiol. 2017;67:2640–5.
49. Hugon P, Lagier J-C, Robert C, Lepolard C, Papazian L, Musso D, et al. Molecular studies neglect apparently gram-negative populations in the human gut microbiota. J Clin Microbiol. 2013;51:3286–93.
50. Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, et al. An integrated catalog of reference genes in the human gut microbiome. Nat Biotechnol. 2014;32:834–41.
51. Human Microbiome Project Consortium. A framework for human microbiome research. Nature. 2012;486:215–21.
52. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486:207–14.
53. Stokes EJ. Anaerobic bacteria: role in disease. J Clin Pathol. 1975;28:424–5.
54. DiBaise JK, Zhang H, Crowell MD, Krajmalnik-Brown R, Decker GA, Rittmann BE. Gut microbiota and its possible relationship with obesity. Mayo Clin Proc. 2008;83:460–9.
55. La Scola B, Khelaifia S, Lagier J-C, Raoult D. Aerobic culture of anaerobic bacteria using antioxidants: a preliminary report. Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol. 2014;33:1781–3.
56. Summanen PH, McTeague M, Väisänen ML, Strong CA, Finegold SM. Comparison of recovery of anaerobic bacteria using the Anoxomat, anaerobic chamber, and GasPak jar systems. Anaerobe. 1999;5:5–9.
57. Seng P, Abat C, Rolain JM, Colson P, Lagier J-C, 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.
58. Drancourt M, Berger P, Raoult D. Systematic 16S rRNA gene sequencing of atypical clinical isolates identified 27 new bacterial species associated with humans. J Clin Microbiol. 2004;42:2197–202.
59. Fenollar F, Lagier J-C, Raoult D. *Tropheryma whipplei* and Whipple's disease. J Inf Secur. 2014;69:103–12.
60. Raoult D, Birg ML, Scola BL, Fournier PE, Enea M, Lepidi H, et al. Cultivation of the Bacillus of Whipple's disease. N Engl J Med. 2000;342:620–5.
61. Rolain J-M, Fenollar F, Raoult D. False positive PCR detection of *Tropheryma whipplei* in the saliva of healthy people. BMC Microbiol. 2007;7:48.
62. Okaro U, Addisu A, Casanas B, Anderson B. Bartonella species, an emerging cause of blood-culture-negative endocarditis. Clin Microbiol Rev. 2017;30:709–46.
63. Slater LN, Welch DF, Hensel D, Coody DW. A newly recognized fastidious gram-negative pathogen as a cause of fever and bacteremia. N Engl J Med. 1990;323:1587–93.
64. Perkocha LA, Geaghan SM, Yen TS, Nishimura SL, Chan SP, Garcia-Kennedy R, et al. Clinical and pathological features of bacillary peliosis hepatis in association with human immunodeficiency virus infection. N Engl J Med. 1990;323:1581–6.
65. Relman DA, Loutit JS, Schmidt TM, Falkow S, Tompkins LS. The agent of bacillary angiomatosis. An approach to the identification of uncultured pathogens. N Engl J Med. 1990;323:1573–80.
66. Welch DF, Pickett DA, Slater LN, Steigerwalt AG, Brenner DJ. *Rochalimaea henselae* sp. nov., a cause of septicemia, bacillary angiomatosis, and parenchymal bacillary peliosis. J Clin Microbiol. 1992;30:275–80.
67. Brenner DJ, O'Connor SP, Winkler HH, Steigerwalt AG. Proposals to unify the genera *Bartonella* and *Rochalimaea*, with descriptions of *Bartonella quintana* comb. nov., *Bartonella vinsonii* comb. nov., *Bartonella henselae* comb. nov., and *Bartonella elizabethae* comb. nov., and to remove the family *Bartonellaceae* from the order Rickettsiales. Int J Syst Bacteriol. 1993;43:777–86.

68. Ramasamy D, Mishra AK, Lagier J-C, Padhmanabhan 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.
69. Abdallah RA, Beye M, Diop A, Bakour S, Raoult D, Fournier P-E. The impact of culturomics on taxonomy in clinical microbiology. Antonie Van Leeuwenhoek [Internet]. 2017; Available from: http://link.springer.com/10.1007/s10482-017-0871-1. Cited 11 Jun 2017
70. Fournier P-E, Raoult D, Drancourt M. Republication of «new species announcement», a new format to prompt the description of new human microbial species. Hum Microbiome J. 2016;1:A1–2.
71. Isenberg HD. Pathogenicity and virulence: another view. Clin Microbiol Rev. 1988;1:40–53.
72. Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. Nat Rev Microbiol. 2004;2:123–40.
73. Sotto A, Richard J-L, Messad N, Molinari N, Jourdan N, Schuldiner S, et al. Distinguishing colonization from infection with *Staphylococcus aureus* in diabetic foot ulcers with miniaturized oligonucleotide arrays. Diabetes Care. 2012;35:617–23.
74. Lagier J-C, Dubourg G, Amrane S, Raoult D. Koch postulate: why should we grow bacteria? Arch Med Res. 2017;48:774–9.
75. Fournier P-E, Drancourt M. New microbes new infections promotes modern prokaryotic taxonomy: a new section "TaxonoGenomics: new genomes of microorganisms in humans". New Microbes New Infect. 2015;7:48–9.
76. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux J-J, et al. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc Natl Acad Sci U S A. 2008;105:16731–6.
77. Derrien M, Vaughan EE, Plugge CM, de Vos WM. *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. Int J Syst Evol Microbiol. 2004;54:1469–76.
78. Gómez-Gallego C, Pohl S, Salminen S, De Vos WM, Kneifel W. *Akkermansia muciniphila*: a novel functional microbe with probiotic properties. Benef Microbes. 2016;7:571–84.
79. Dao MC, Everard A, Aron-Wisnewsky J, Sokolovska N, Prifti E, Verger EO, et al. *Akkermansia muciniphila* and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. Gut. 2016;65:426–36.
80. Cassir N, Benamar S, Khalil JB, Croce O, Saint-Faust M, Jacquot A, et al. *Clostridium butyricum* strains and dysbiosis linked to necrotizing enterocolitis in preterm neonates. Clin Infect Dis Off Publ Infect Dis Soc Am. 2015;61:1107–15.
81. Routy B, Le Chatelier E, Derosa L, Duong CPM, Alou MT, Daillère R, et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. Science. 2018;359:91–7.
82. Sommer MOA. Advancing gut microbiome research using cultivation. Curr Opin Microbiol. 2015;27:127–32.
83. Goodman AL, Kallstrom G, Faith JJ, Reyes A, Moore A, Dantas G, et al. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. Proc Natl Acad Sci U S A. 2011;108:6252–7.
84. Rettedal EA, Gumpert H, Sommer MOA. Cultivation-based multiplex phenotyping of human gut microbiota allows targeted recovery of previously uncultured bacteria. Nat Commun. 2014;5:4714.
85. Lau JT, Whelan FJ, Herath I, Lee CH, Collins SM, Bercik P, et al. Capturing the diversity of the human gut microbiota through culture-enriched molecular profiling. Genome Med [Internet]. 2016;8 Available from: http://genomemedicine.biomedcentral.com/articles/10.1186/s13073-016-0327-7. Cited 11 Jun 2017
86. McClelland RS, Lingappa JR, Srinivasan S, Kinuthia J, John-Stewart GC, Jaoko W, et al. Evaluation of the association between the concentrations of key vaginal bacteria and the increased risk of HIV acquisition in African women from five cohorts: a nested case-control study. Lancet Infect Dis. 2018;18:554–64
87. Browne HP, Forster SC, Anonye BO, Kumar N, Neville BA, Stares MD, et al. Culturing of "unculturable" human microbiota reveals novel taxa and extensive sporulation. Nature. 2016;533:543–6.
88. Fodor AA, DeSantis TZ, Wylie KM, Badger JH, Ye Y, Hepburn T, et al. The "most wanted" taxa from the human microbiome for whole genome sequencing. PLoS ONE [Internet]. 2012;7 Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3406062/
89. Lagier JC, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, et al. Culturing human microbiota and culturomics. Nat Rev Microbiol. 4 In Press

Part II

Description of the gut microbiota of Pygmy people by culturomics and metagenomics along with the taxono-genomics description of the isolated new bacterial species.

Partie II

Description du microbiote intestinal des Pygmée par la culturomics et la métagénomique ainsi que la description taxonomique des nouvelles espèces bactérienne isolées.

II. i) Sample collection and culturomics:

We applied cuturomics on 15 samples out of which 5 have been subjected to the 18 culture conditions and 10 were cultured in random conditions (Table 1 and 2).

Samples underwent the process of culturomics and metagenomics (V3-V4) as previously described in previously published work (1). Bacterial growth assessment was done over 30 days of follow up by direct culturing from the blood culture bottles on COS medium at day 0, 1, 3, 5, 7, 10, 15, 21 and 30 and incubated according to the essential culture condition chosen. Isolated colonies were identified by the means of MALDI-TOF MS as previously described and any un-identified bacterial colony underwent 16S rRNA gene sequencing. Whenever identified by 16S rRNA gene sequencing, the typical mass spectrum of the understudied organism was added to the current database for a better identification in a next scene.II. ii) Results

II. ii) a) Listing of the different bacterial species identified in the current study by either culutromics or meteagenomics and its classification.

We represent in table 3 and table 4 the results obtained by culturomics and metagenomics, respectively.

II. ii) b) Listing of the new bacterial species isolated by culturomics along with its main characteristics.

We succeeded in isolating 38 new bacterial species, out of which 27 were

validated with available 16S rRNA sequence on NCBI nucleotide database (Table 5) and 13 are being confirmed by whole genome sequencing analysis.

Most of the isolated new species have been described in the taxono-genomic approach. We summarize the main characteristics of these species in table 5.

In addition, we provide in the next pages 3 examples of published taxono-genomic descriptive articles of 3 of these new species along with 1 model of the new species announcement format used. For the remaining species, we provide links to access its articles online but those that are submitted or in current revision were not included (Annex 1).

Interestingly, 3 of these new species had its 16S rRNA gene sequence correlated with previously considered operational taxonomic units. This work is further developed in the article published in 2016 entitled: Culture of previously uncultured members of the human gut microbiota by culturomics (1). This emphasizes again on the fact that we are still far from fully understanding the human gut microbial gut content and the part of the commensal bacterial species, which might have important roles that we are neglecting.

II. i) Collection d'échantillons et culturomics

Nous avons appliqué la cuturomics sur 15 échantillons dont 5 ont été soumis aux 18 conditions de culture et 10 ont été cultivés dans des conditions aléatoires (tableaux 1 et 2).

Les échantillons ont subi le protocole de la culturomics et de la métagénomique (V3-V4) comme décrit précédemment dans des travaux publiés (1). L'évaluation de la croissance bactérienne a été réalisée sur 30 jours de suivi par culture directe à partir des flacons d'hémoculture sur milieu COS aux jours 0, 1, 3, 5, 7, 10, 15, 21 et 30. Les colonies isolées ont été identifiées par MALDI-TOF MS comme décrit précédemment et toute colonie bactérienne non identifiée a subi un séquençage du gène de l'ARNr 16S. Chaque fois identifié par le séquençage du gène de l'ARNr 16S, le spectre de masse typique de l'organisme étudié a été rajouté à la base de données actuelle pour une meilleure future identification.

II. ii) Résultats

II. ii) a) Liste des différentes espèces bactériennes identifiées dans la présente étude par la méthode de culturomics ou de métagénomique avec leurs classification.

Nous présentons dans les tableaux 3 et 4 les résultats obtenus respectivement par la culturomics et la métagénomique.

II. ii) b) Liste des nouvelles espèces bactériennes isolées par culturomics avec leurs principales caractéristiques.

Nous avons réussi à isoler 38 nouvelles espèces bactériennes, dont 27 ont été validées avec la séquence d'ARNr 16S disponible sur la base de données de nucléotides NCBI (Tableau 5) et 13 sont à confirmer par l'analyse de séquençage du génome entier. La plupart des nouvelles espèces isolées ont été décrites dans l'approche taxono-génomique. Les principales caractéristiques de ces espèces sont résumées dans le tableau 5. De plus, nous fournissons dans les pages suivantes 3 exemples d'articles descriptifs taxonomiques et génomiques de 3 de ces nouvelles espèces ainsi qu'un « new species announcement format ». Pour le reste des espèces, nous fournissons des liens pour accéder aux articles en ligne (Annexe). Les articles soumis ou dont la révision est en cours n'ont pas été inclus.

Il est intéressant de noter que 3 de ces nouvelles espèces présentaient une corrélation entre la séquence du gène de l'ARNr 16S et des unités taxonomiques opérationnelles précédemment considérées. Ce travail est plus développé plus loin dans l'article publié en 2016 intitulé: Culture des membres non cultivés du microbiote intestinal humain par la culturomics (1). Ceci souligne le fait que nous sommes encore loin de comprendre complètement le contenu microbien de l'intestin humain et la population commensale.

Table 1: Information about the understudied samples in the present work.

Sample	Sex	Age	Culture conditions
F61	Female	35	18
F15	Male	39	18
F86	Female	8	18
F80	Female	50	18
F60	Male	49	18
F64	Female	12	R+RB+B (ANA 37°C)
F79	Female	47	RB ANA 37°C
F93	Female	11	Direct culture
F98	Male	8	Direct culture + RB 37°C ANA
F71	Female	45	Direct culture
FA03	Female	12	Direct culture + RB 37°C ANA
F50	Male	32	RB 37°C ANA
F62	Male	39	Direct culture
F96	Male	4	Direct culture
F70	Female	55	Direct culture

** R: rumen; B: Blood RB: Rumen+ blood; ANA: anaerobic conditions; Direct culture: Diluting stool samples with phosphate buffer saline and direct culturing on Colombia agar media supplement with 5% sheep blood (bioMérieux, Marcy l'Etoile, France) at 37 under anaerobic conditions.

Table 2: The 18 different conditions adapted in culturing samples by the

means of the culturomics approach.

Culturomics standardization by 18 culture conditions

1. Preincubation in aerobic blood culture bottle with stool filtered at 5µm, then 5% sheep blood agar, Aerobic condition, 37°C
2. Preincubation in anaerobic blood culture bottle after thermic shock at 80°C during 20min, then 5% sheep blood agar, Anaerobic condition, 37°C
3. Preincubation in anaerobic condition in marine broth, then 5% sheep blood agar, Anaerobic condition, 37°C
4. Preincubation in aerobic blood culture bottle with 5ml sheep blood, then 5% sheep blood agar, Aerobic condition, 37°C
5. Preincubation in aerobic marine broth, then 5% sheep blood agar, Aerobic condition 37°C
6. Preincubation in aerobic blood culture bottle with rumen fluid and sheep blood, then 5% sheep blood agar, Aerobic condition, 37°C
7. Preincubation in aerobic condition in trypticase soy broth, then 5% sheep blood agar, Aerobic condition, 37°C
8. Preincubation in aerobic condition in 5% sheep blood broth, then 5% sheep blood agar, Aerobic condition, 37°C
9. Preincubation in anaerobic blood culture bottle with stool filtered at 5µm, then 5% sheep blood agar, Anaerobic condition, 37°C
10. Preincubation in aerobic condition in Brain Heart Infusion broth with 5% sheep blood, then 5% sheep blood agar, Aerobic condition, 37°C
11. Preincubation in anaerobic condition in 5% sheep blood broth, then 5% sheep blood

agar , Anaerobic condition, 28°C

12. Preincubation in aerobic blood culture bottle with rumen fluid, then 5% sheep blood agar , Aerobic condition, 37°C
13. Preincubation in aerobic condition in 5% sheep blood broth, then 5% sheep blood agar, Aerobic condition, 28°C
14. Preincubation in anaerobic blood culture bottle with 5ml sheep blood, then 5% sheep blood agar , Anaerobic condition, 37°C
15. Preincubation in anaerobic blood culture bottle, then 5% sheep blood agar , Anaerobic condition, 37°C
16. Preincubation in anaerobic blood culture bottle with rumen fluid and sheep blood, then 5% sheep blood agar, Anaerobic condition, 37°C
17. Preincubation in anaerobic condition in 5% sheep blood broth, then 5% sheep blood agar, Anaerobic condition, 37°C
18. Preincubation in anaerobic blood culture bottle with rumen fluid, then 5% sheep blood agar , Anaerobic condition, 37°C

Table3: Isolated organisms after performing culturomics on the selected stool samples detailed previously in table 1. A: New species isolated by culturomics; B: New species isolated by culturomics in previous studies; C: Previously known species. *Human (H)/Non Human (NH)/ Human gut (HG). 0: Able to tolerate oxygen and 1: anaerobic.

Species	Classification	Phylum	Oxygen requirement	Source of isolation*
Actinomyces minihominis	A	Actinobacteria	0	HG
Actinomyces phoceensis	B	Actinobacteria	0	HG
Actinomyces polynesiensis	B	Actinobacteria	0	HG

Alistipes finegoldii	C	Bacteroidetes	1	HG
Alistipes indistinctus	C	Bacteroidetes	1	HG
Alistipes jeddahmassiliense	B	Bacteroidetes	1	HG
Alistipes obesi	B	Bacteroidetes	1	HG
Alistipes putredinis	C	Bacteroidetes	1	HG
Alistipes senegalensis	B	Bacteroidetes	1	HG
Alistipes shahii	C	Bacteroidetes	1	HG
Allisonella histaminiformans	C	Firmicutes	1	HG
Anaerofustis stercorihominis	C	Firmicutes	1	HG
Anaerosalibacter bizertensis	C	Firmicutes	1	HG
Anaerosphaera massiliensis	A	Firmicutes	1	HG
Anaerotroncus colihominis	C	Firmicutes	1	HG
Anaerotruncus rubiinfantis	B	Firmicutes	1	HG
Arabia massiliensis	B	Actinobacteria	1	HG
Arthrobacter creatinolyticus	C	Actinobacteria	0	H
Arthrobacter protophormiae	C	Actinobacteria	0	NH
Bacillus vallismortis	C	Firmicutes	0	HG
Bacteroides caccae	C	Bacteroidetes	1	HG
Bacteroides congonensis	A	Bacteroidetes	1	HG
Bacteroides fragilis	C	Bacteroidetes	1	HG
Bacteroides intestinalis	C	Bacteroidetes	1	HG
Bacteroides nordii	C	Bacteroidetes	1	HG
Bacteroides oleiciplenus	C	Bacteroidetes	1	HG
Bacteroides ovatus	C	Bacteroidetes	1	HG
Bacteroides thetaioteomicron	C	Bacteroidetes	1	HG
Bacteroides timonense	B	Bacteroidetes	1	HG
Bacteroides uniformis	C	Bacteroidetes	1	HG
Bacteroides vulgatus	C	Bacteroidetes	1	HG
Barnesiella intestinihominis	C	Bacteroidetes	1	HG
Bifidobacterium adolescentis	C	Actinobacteria	1	HG
Bifidobacterium pseudocatenulatum	C	Actinobacteria	1	HG
Bilophila wadsworthia	C	Proteobacteria	1	HG
Bittarella massiliensis	B	Firmicutes	1	HG
Butyricimonas phoceensis	B	Bacteroidetes	1	HG
Candida krusei	C		0	H
Christensella minuta	C	Firmicutes	1	HG
Clostridium aminovalericum	C	Firmicutes	1	HG
Clostridium bifermentans	C	Firmicutes	1	HG

Clostridium bouchedurhonensis	B	Firmicutes	1	HG
Clostridium butyricum	C	Firmicutes	1	HG
Clostridium clostridioforme	C	Firmicutes	1	HG
Clostridium culturomicsense	B	Firmicutes	1	HG
Clostridium glycolycum	C	Firmicutes	1	HG
Clostridium lavalense	C	Firmicutes	1	HG
Clostridium lituseburense	C	Firmicutes	1	NH
Clostridium mangenotii	C	Firmicutes	1	HG
Clostridium minihomine	A	Firmicutes	1	HG
Clostridium paraperfringens	C	Firmicutes	1	HG
Clostridium paraputrificum	C	Firmicutes	1	HG
Clostridium perfringens	C	Firmicutes	1	HG
Clostridium polynesiense	B	Firmicutes	1	HG
Clostridium ramosum	C	Firmicutes	1	HG
Clostridium sartagoforme	C	Firmicutes	1	HG
Clostridium sordellii	C	Firmicutes	1	HG
Clostridium sporogenes	C	Firmicutes	1	HG
Clostridium symbiosum	C	Firmicutes	1	HG
Clostridium tertium	C	Firmicutes	1	HG
Collinsella aerofaciens	C	Actinobacteria	1	HG
Collinsella bouchedurhonensis	A	Actinobacteria	1	HG
Congobacterium massiliense	A	Firmicutes	1	HG
Coprobacillus massiliensis	B	Firmicutes	1	HG
Corynebacterium afermentans	C	Actinobacteria	0	HG
Corynebacterium provencense	B	Actinobacteria	0	HG
Corynebacterium tuberculostearicum	C	Actinobacteria	0	H
Dialister succinatiphilus	C	Firmicutes	1	HG
Dorea phocaeensis	A	Firmicutes	1	HG
Dysgonomonas massiliensis	A	Bacteroidetes	0	HG
Dysgonomonas phocaensis	A	Firmicutes	0	HG
Dysgonomonas timonensis	A	Firmicutes	0	HG
Eggerthella lenta	C	Actinobacteria	1	HG
Eggerthella timonensis	A	Actinobacteria	1	HG
Enterobacter gallinarum	C	Proteobacteria	0	HG
Enterobacter thailandicus	C	Proteobacteria	0	HG
Enterococcus asini	C	Firmicutes	0	HG
Enterococcus avium	C	Firmicutes	0	HG
Enterococcus casseliflavus	C	Firmicutes	0	HG

Enterococcus devriesei	C	Firmicutes	0	HG
Enterococcus diestrammenae	C	Firmicutes	0	HG
Enterococcus dispar	C	Firmicutes	0	HG
Enterococcus durans	C	Firmicutes	0	HG
Enterococcus faecalis	C	Firmicutes	0	HG
Enterococcus faecium	C	Firmicutes	0	HG
Enterococcus galinarum	C	Firmicutes	0	HG
Enterococcus hirae	C	Firmicutes	0	HG
Enterococcus malodoratus	C	Firmicutes	0	HG
Enterococcus mediteraneensis	A	Firmicutes	0	HG
Enterococcus phoeniculicola	C	Firmicutes	0	HG
Enterococcus pseudoavium	C	Firmicutes	0	HG
Enterococcus raffinosus	C	Firmicutes	0	HG
Enterococcus thailandicus	C	Firmicutes	0	HG
Escherichia coli	C	Proteobacteria	0	HG
Flavonifractor plautii	C	Firmicutes	1	HG
Fusobacterium varium	C	Fusobacteria	1	HG
Gabonibacter massiliensis	B	Bacteroidetes	1	HG
Gabonibacter timonensis	A	Bacteroidetes	1	HG
Gardnerella vaginalis	C	Actinobacteria	0	HG
Gordonibacter pamelaeae	C	Actinobacteria	1	HG
Ihuella massiliensis	B	Firmicutes	1	HG
Intestinibacillus timonensis	A	Firmicutes	1	HG
Intestinimonas butyriciproducens	C	Firmicutes	1	HG
Intestinimonas massiliensis	B	Firmicutes	1	HG
Lachnoclostridium phocaeense	B	Firmicutes	1	H
Lactobacillus brevis	C	Firmicutes	0	HG
Lactobacillus camelliae	C	Firmicutes	0	NH
Lactobacillus coryniformis	C	Firmicutes	0	HG
Lactobacillus desctrincus	C	Firmicutes	0	H
Lactobacillus dextrinicus	C	Firmicutes	0	HG
Lactobacillus harbinensis	C	Firmicutes	0	HG
Lactobacillus kimchii	C	Firmicutes	0	NH
Lactobacillus mucosae	C	Firmicutes	0	HG
Lactobacillus paracasei	C	Firmicutes	0	HG
Lactobacillus pentosus	C	Firmicutes	0	HG
Lactobacillus perolens	C	Firmicutes	0	HG
Lactobacillus plantarum	C	Firmicutes	0	HG

Lactobacillus reuteri	C	Firmicutes	0	HG
Lactobacillus rhamnosus	C	Firmicutes	0	HG
Lactobacillus vaccinostercus	C	Firmicutes	0	HG
Lactococcus garvieae	C	Firmicutes	0	HG
Lactococcus lactis	C	Firmicutes	0	HG
Libanicoccus massiliensis	A	Actinobacteria	1	HG
Massilimaliae massiliensis	B	Proteobacteria	1	HG
Miniphocibacter massiliensis	A	Firmicutes	1	HG
Mobilibacterium timonense	A	Firmicutes	1	HG
Mogibacterium neglectum	C	Firmicutes	1	H
Morganella morganii	C	Proteobacteria	0	HG
Negativicoccus massiliensis	B	Firmicutes	1	HG
Neglecta bouchesdurhonensis	A	Firmicutes	1	HG
Neobittarella massiliensis	A	Firmicutes	1	HG
Neocorynebacterium massiliensis	A	Actinobacteria	1	HG
Niameyia massiliensis	B	Firmicutes	1	HG
Odoribacter splanchnicus	C	Bacteroidetes	1	HG
Olegusella massiliensis	B	Firmicutes	1	H
Olsenella congonensis	A	Actinobacteria	1	HG
Olsenella timonensis	B	Actinobacteria	1	HG
Paenibacillus timonensis	B	Firmicutes	0	HG
Parabacteroides distasonis	C	Bacteroidetes	1	HG
Parabacteroides goldsteinii	C	Bacteroidetes	1	HG
Parabacteroides gordonii	C	Bacteroidetes	1	HG
Parabacteroides massiliense	B	Bacteroidetes	1	HG
Parabacteroides phocaeensis	B	Bacteroidetes	1	H
Parabacteroides timonensis	A	Bacteroidetes	1	HG
Pediococcus pentosaceus	C	Firmicutes	1	HG
Peptoniphilus grossensis	B	Firmicutes	1	HG
Peptoniphilus phoceensis	B	Firmicutes	1	HG
Peptostreptococcus russellii	C	Firmicutes	1	HG
Peptostreptococcus syus	A	Firmicutes	1	HG
Phocea massiliensis	B	Firmicutes	1	HG
Phoenicibacter massiliensis	A	Actinobacteria	1	HG
Prevotella caldimerdae	C	Bacteroidetes	1	HG
Proteus mirabilis	C	Proteobacteria	0	HG
Proteus vulgaris	C	Proteobacteria	0	HG
Provencella massiliensis	B	Firmicutes	1	HG

Provencibacter massiliensis	B	Firmicutes	1	HG
Providencia alcalifaciens	C	Proteobacteria	0	HG
Providencia rettgeri	C	Proteobacteria	0	HG
Psychrobacter massiliensis	B	Proteobacteria	0	HG
Pygmaiobacter massiliensis	A	Firmicutes	1	HG
Pyramidobacter piscolens	C	Synergistetes	1	HG
Raoultibacter massiliensis	B	Actinobacteria	0	HG
Raoultibacter timonensis	A	Actinobacteria	0	HG
Robinsoniella peoriensis	C	Firmicutes	1	HG
Ruminococcus merdae	B	Firmicutes	1	HG
Ruthenibacterium lactatiformans	C	Firmicutes	1	HG
Sanguibacter massiliensis	A	Firmicutes	1	HG
Senegalemassilia anaerobia	B	Actinobacteria	1	HG
Slackia exigua	C	Actinobacteria	1	HG
Slackia isoflavoniconvertens	C	Actinobacteria	1	HG
Staphylococcus epidermidis	C	Firmicutes	0	HG
Staphylococcus hominis	C	Firmicutes	0	HG
Streptococcus lutetiensis	C	Firmicutes	0	HG
Streptococcus pyogenes	C	Firmicutes	0	HG
Streptococcus thermophilus	C	Firmicutes	0	HG
Vagococcus fluvialis	C	Firmicutes	0	HG
Vagococcus lutrae	C	Firmicutes	0	HG
Virgibacillus ndiopensis	B	Firmicutes	0	HG
Weissela paramesenteroides	C	Firmicutes	0	H

Table 4: Metagenomics (V3V4) results obtained after performing it on the selected samples in this study, detailed in table 1. *Human (H)/Non Human(NH)/ Human gut(HG). 0: Able to tolerate oxygen and 1: anaerobic.

Species	**Phylum**	**Oxygen requirement**	***Source of isolation**
Abiotrophia para-adiacens	Firmicutes	0	H
Acinetobacter calcoaceticus	Proteobacteria	0	HG
Acinetobacter guillouiae	Proteobacteria	0	H

Acinetobacter haemolyticus	Proteobacteria	0	HG
Acinetobacter johnsonii	Proteobacteria	0	H
Acinetobacter junii	Proteobacteria	0	HG
Acinetobacter kookii	Proteobacteria	0	NH
Acinetobacter schindleri	Proteobacteria	0	HG
Actinobacillus rossii	Proteobacteria	0	NH
Actinomyces gerencseriae	Actinobacteria	0	H
Actinomyces naeslundii	Actinobacteria	0	HG
Actinomyces odontolyticus	Actinobacteria	0	HG
Actinomyces timonensis	Actinobacteria	0	H
Aggregatibacter aphrophilus	Proteobacteria	0	HG
Aggregatibacter segnis	Proteobacteria	0	HG
Akkermansia muciniphila	Verrucomicrobia	0	HG
Alcaligenes faecalis	Proteobacteria	0	HG
Alistipes finegoldii	Bacteroidetes	1	HG
Alistipes indistinctus	Bacteroidetes	1	HG
Alistipes onderdonkii	Bacteroidetes	1	HG
Alistipes putredinis	Bacteroidetes	1	HG
Alistipes senegalensis	Bacteroidetes	1	HG
Alistipes shahii	Bacteroidetes	1	HG
Alistipes timonensis	Bacteroidetes	1	HG
Anaerococcus vaginalis	Firmicutes	1	HG
Anaerofustis stercorihominis	Firmicutes	1	HG
Anaerostipes hadrus	Firmicutes	1	HG
Arthrobacter citreus	Actinobacteria	0	NH
Atopobium parvulum	Actinobacteria	0	HG
Bacillus amyloliquefaciens	Firmicutes	0	HG
Bacillus flexus	Firmicutes	0	HG
Bacillus megaterium	Firmicutes	0	HG
Bacillus pumilus	Firmicutes	0	HG

Bacillus sporothermodurans	Firmicutes	0	HG
Bacillus subtilis	Firmicutes	0	HG
Bacteroides acidifaciens	Bacteroidetes	0	NH
Bacteroides caccae	Bacteroidetes	1	HG
Bacteroides clarus	Bacteroidetes	1	HG
Bacteroides dorei	Bacteroidetes	1	HG
Bacteroides eggerthii	Bacteroidetes	1	HG
Bacteroides finegoldii	Bacteroidetes	1	HG
Bacteroides fragilis	Bacteroidetes	1	HG
Bacteroides galacturonicus	Bacteroidetes	1	HG
Bacteroides nordii	Bacteroidetes	1	HG
Bacteroides ovatus	Bacteroidetes	1	HG
Bacteroides pectinophilus	Bacteroidetes	1	HG
Bacteroides plebeius	Bacteroidetes	1	HG
Bacteroides thetaiotaomicron	Bacteroidetes	1	HG
Bacteroides uniformis	Bacteroidetes	1	HG
Barnesiella intestinihominis	Bacteroidetes	1	HG
Bifidobacterium adolescentis	Actinobacteria	1	HG
Bifidobacterium kashiwanohense	Actinobacteria	1	HG
Bifidobacterium longum	Actinobacteria	1	HG
Bilophila wadsworthia	Proteobacteria	1	HG
Blastococcus aggregatus	Actinobacteria	0	HG
Blautia faecis	Firmicutes	1	HG
Blautia luti	Firmicutes	1	HG
Brachybacterium paraconglomeratum	Actinobacteria	0	HG
Brachyspira pilosicoli	Spirochaetes	0	HG
Brochothrix thermosphacta	Firmicutes	0	NH
Butyricimonas virosa	Bacteroidetes	1	HG
Butyrivibrio crossotus	Firmicutes	0	HG
Campylobacter hominis	Proteobacteria	0	HG
Campylobacter jejuni	Proteobacteria	0	HG

Campylobacter mucosalis	Proteobacteria	0	HG
Campylobacter troglodytis	Proteobacteria	0	HG
Catabacter hongkongensis	Proteobacteria	1	HG
Catenibacterium mitsuokai	Firmicutes	1	HG
Cellulosimicrobium cellulans	Actinobacteria	0	HG
Cetobacterium somerae	Fusobacteria	0	HG
Citrobacter koseri	Proteobacteria	0	HG
Cloacibacillus porcorum	Synergistetes	0	HG
Clostridium aldenense	Firmicutes	1	HG
Clostridium beijerinckii	Firmicutes	1	HG
Clostridium butyricum	Firmicutes	1	HG
Clostridium celatum	Firmicutes	1	HG
Clostridium chartatabidum	Firmicutes	1	NH
Clostridium clostridioforme	Firmicutes	1	HG
Clostridium hathewayi	Firmicutes	1	HG
Clostridium paraputrificum	Firmicutes	1	HG
Clostridium perfringens	Firmicutes	1	HG
Clostridium septicum	Firmicutes	1	HG
Clostridium sordellii	Firmicutes	1	HG
Collinsella aerofaciens	Actinobacteria	1	HG
Comamonas kerstersii	Proteobacteria	0	HG
Comamonas testosteroni	Proteobacteria	0	H
Coprobacillus cateniformis	Firmicutes	1	HG
Coprococcus comes	Firmicutes	1	HG
Coprococcus eutactus	Firmicutes	1	HG
Corynebacterium durum	Actinobacteria	0	HG
Corynebacterium flavescens	Actinobacteria	0	HG
Corynebacterium glutamicum	Actinobacteria	0	H
Corynebacterium kroppenstedtii	Actinobacteria	0	HG
Corynebacterium propinquum	Actinobacteria	0	HG
Corynebacterium tuberculostearicum	Actinobacteria	0	HG

Corynebacterium ureicelerivorans	Actinobacteria	0	HG
Corynebacterium variabile	Actinobacteria	0	HG
Corynebacterium xerosis	Actinobacteria	0	HG
Cronobacter dublinensis	Proteobacteria	0	H
Cronobacter sakazakii	Proteobacteria	0	HG
Desulfovibrio desulfuricans	Proteobacteria	0	HG
Dialister invisus	Firmicutes	1	HG
Dialister pneumosintes	Firmicutes	1	HG
Dialister propionicifaciens	Firmicutes	1	H
Dielma fastidiosa	Firmicutes	0	HG
Dietzia cinnamea	Actinobacteria	0	HG
Dorea formicigenerans	Firmicutes	1	HG
Dorea longicatena	Firmicutes	1	HG
Edwardsiella tarda	Proteobacteria	0	HG
Eggerthella lenta	Actinobacteria	1	HG
Eikenella corrodens	Proteobacteria	0	HG
Enterobacter aerogenes	Proteobacteria	0	HG
Enterobacter asburiae	Proteobacteria	0	HG
Enterobacter cloacae	Proteobacteria	0	HG
Enterococcus casseliflavus	Firmicutes	0	HG
Enterococcus faecalis	Firmicutes	0	HG
Enterococcus faecium	Firmicutes	0	HG
Erysipelatoclostridium ramosum	Firmicutes	0	H
Escherichia coli	Proteobacteria	0	HG
Eubacterium eligens	Firmicutes	1	HG
Eubacterium hallii	Firmicutes	1	HG
Eubacterium infirmum	Firmicutes	1	H
Eubacterium ramulus	Firmicutes	1	HG
Eubacterium rectale	Firmicutes	1	HG
Eubacterium siraeum	Firmicutes	1	HG
Faecalibacterium prausnitzii	Firmicutes	0	HG

Fusicatenibacter saccharivorans	Firmicutes	0	HG
Fusobacterium equinum	Fusobacteria	1	NH
Fusobacterium nucleatum	Fusobacteria	1	HG
Fusobacterium periodonticum	Fusobacteria	1	HG
Fusobacterium varium	Fusobacteria	1	HG
Gemella haemolysans	Firmicutes	0	HG
Gemella sanguinis	Firmicutes	0	HG
Gemmiger formicilis	Firmicutes	0	HG
Gordonibacter pamelaeae	Actinobacteria	1	HG
Haemophilus haemolyticus	Proteobacteria	0	HG
Haemophilus influenzae	Proteobacteria	0	HG
Haemophilus parahaemolyticus	Proteobacteria	0	H
Haemophilus parainfluenzae	Proteobacteria	0	HG
Halomonas phoceae	Proteobacteria	0	NH
Helicobacter fennelliae	Proteobacteria	0	HG
Herbaspirillum seropedicae	Proteobacteria	0	H
Holdemanella biformis	Firmicutes	1	HG
Holdemania filiformis	Firmicutes	1	HG
Ignatzschineria larvae	Proteobacteria	0	H
Intestinibacter bartlettii	Firmicutes	1	HG
Intestinimonas butyriciproducens	Firmicutes	1	HG
Klebsiella pneumoniae	Proteobacteria	0	HG
Kurthia zopfii	Firmicutes	0	NH
Lachnoclostridium phytofermentans	Firmicutes	1	NH
Lactobacillus brevis	Firmicutes	0	HG
Lactobacillus concavus	Firmicutes	0	NH
Lactobacillus crustorum	Firmicutes	0	NH
Lactobacillus delbrueckii	Firmicutes	0	HG
Lactobacillus mucosae	Firmicutes	0	HG
Lactobacillus pantheris	Firmicutes	0	NH
Lactobacillus plantarum	Firmicutes	0	HG

Lactobacillus ruminis	Firmicutes	0	HG
Lactobacillus vaccinostercus	Firmicutes	0	HG
Lactococcus garvieae	Firmicutes	0	HG
Lactococcus lactis	Firmicutes	0	HG
Leptotrichia amnionii	Fusobacteria	0	H
Leuconostoc citreum	Firmicutes	0	H
Lysinibacillus fusiformis	Firmicutes	0	HG
Massilia niabensis	Proteobacteria	1	NH
Massilia plicata	Proteobacteria	1	NH
Megasphaera elsdenii	Firmicutes	0	HG
Methylobacterium extorquens	Proteobacteria	0	H
Micrococcus luteus	Actinobacteria	0	HG
Moraxella catarrhalis	Proteobacteria	0	H
Morganella morganii	Proteobacteria	0	HG
Neisseria flavescens	Proteobacteria	1	HG
Odoribacter splanchnicus	Bacteroidetes	1	HG
Oribacterium parvum	Firmicutes	1	H
Oxalobacter formigenes	Proteobacteria	1	HG
Parabacteroides distasonis	Bacteroidetes	1	HG
Parabacteroides goldsteinii	Bacteroidetes	1	HG
Parabacteroides merdae	Bacteroidetes	1	HG
Paracoccus aminovorans	Proteobacteria	0	NH
Paracoccus marinus	Proteobacteria	0	NH
Peptoniphilus asaccharolyticus	Firmicutes	1	HG
Peptostreptococcus anaerobius	Firmicutes	1	HG
Peptostreptococcus russellii	Firmicutes	1	HG
Peptostreptococcus stomatis	Firmicutes	1	H
Phascolarctobacterium succinatutens	Firmicutes	1	H
Porphyromonas bennonis	Bacteroidetes	1	HG
Prevotella buccalis	Bacteroidetes	1	HG
Prevotella copri	Bacteroidetes	1	HG

Prevotella corporis	Bacteroidetes	1	HG
Prevotella disiens	Bacteroidetes	1	HG
Prevotella nanceiensis	Bacteroidetes	1	HG
Prevotella stercorea	Bacteroidetes	1	HG
Prevotella timonensis	Bacteroidetes	1	H
Propionibacterium acnes	Actinobacteria	1	HG
Proteus hauseri	Proteobacteria	0	H
Proteus mirabilis	Proteobacteria	0	HG
Providencia rettgeri	Proteobacteria	0	HG
Pseudomonas fulva	Proteobacteria	0	HG
Pseudomonas putida	Proteobacteria	0	HG
Psychrobacter faecalis	Proteobacteria	0	HG
Pyramidobacter piscolens	Synergistetes	1	HG
Ralstonia pickettii	Proteobacteria	0	H
Romboutsia lituseburensis	Firmicutes	1	HG
Roseburia inulinivorans	Firmicutes	1	HG
Rothia dentocariosa	Actinobacteria	0	HG
Rothia mucilaginosa	Actinobacteria	0	HG
Ruminococcus bicirculans	Firmicutes	1	HG
Ruminococcus bromii	Firmicutes	1	HG
Ruminococcus callidus	Firmicutes	1	HG
Ruminococcus champanellensis	Firmicutes	1	HG
Ruminococcus faecis	Firmicutes	1	HG
Ruminococcus flavefaciens	Firmicutes	1	H
Ruminococcus gnavus	Firmicutes	1	HG
Ruminococcus lactaris	Firmicutes	1	HG
Ruminococcus obeum	Firmicutes	1	HG
Salmonella enterica	Proteobacteria	0	HG
Sarcina ventriculi	Firmicutes	1	HG
Senegalimassilia anaerobia	Actinobacteria	1	HG
Slackia exigua	Actinobacteria	1	HG

Slackia isoflavoniconvertens	Actinobacteria	1	HG
Solobacterium moorei	Firmicutes	1	HG
Sphingobacterium siyangense	Bacteroidetes	0	NH
Staphylococcus aureus	Firmicutes	0	HG
Staphylococcus epidermidis	Firmicutes	0	HG
Stenotrophomonas acidaminiphila	Proteobacteria	0	HG
Stenotrophomonas maltophilia	Proteobacteria	0	HG
Streptococcus anginosus	Firmicutes	0	HG
Streptococcus gallolyticus	Firmicutes	0	HG
Streptococcus gordonii	Firmicutes	0	HG
Streptococcus mitis	Firmicutes	0	HG
Streptococcus pneumoniae	Firmicutes	0	HG
Streptococcus salivarius	Firmicutes	0	HG
Streptococcus salivarius subsp. thermophilus	Firmicutes	0	HG
Streptococcus sobrinus	Firmicutes	0	H
Succinatimonas hippei	Proteobacteria	0	H
Sutterella wadsworthensis	Proteobacteria	1	HG
Treponema succinifaciens	Spirochaetes	1	NH
Turicibacter sanguinis	Firmicutes	1	HG
Vagococcus fluvialis	Firmicutes	0	HG
Veillonella atypica	Firmicutes	1	HG
Veillonella dispar	Firmicutes	1	HG
Veillonella parvula	Firmicutes	1	HG
Victivallis vadensis	Lentisphaerae	1	HG
Weissella confusa	Firmicutes	0	HG
Weissella hellenica	Firmicutes	0	NH
Wohlfahrtiimonas chitiniclastica	Proteobacteria	1	H

Table 5: Differential characteristics of the new species isolated in the present study by culturomics.

Species name	Gram	Catalase	Oxidase	Oxygen environment	Genome Size (MB)	GC %
Libanicoccus massiliensis	-	-	-	Anaerobic	2.01	65.4
Phoenicibacter massiliensis	-	-	-	Anaerobic	1.45	43.4
Mobilibacterium timonense	-	-	-	Anaerobic	2.09	53
Raoultibacter timonensis	+	+	-	Facultative anaerobic	3.94	59.9
Congobacterium massiliense	+	-	-	Anaerobic	2.2	49.6
Pygmaiobacter massiliensis	+	-	-	Anaerobic	2.8	49.4
Neobittarella massiliensis	+	-	-	Anaerobic	32.3	58.55
Miniphocibacter massiliensis	+	+	-	Anaerobic	2.08	28.2
Neocorynebacterium massiliensis	+	+	-	Facultative anaerobic	On going	On going
Sanguibacter massiliensis	+	+	-	Anaerobic	3.07	71.6
Anaerosphaera massiliensis	+	-	-	Anaerobic	On going	On going
Dysgonomonas massiliensis	-	-	-	Facultative anaerobic	3.4	37.3
Gabonibacter timonensis	+	-	-	Anaerobic	3.4	42.1
Dorea phocaeensis	+	-	-	Anaerobic	2.47	43.18
Peptostreptococcus syus	+	-	-	Anaerobic	On going	On going
Olsenella congonensis	+	-	-	Anaerobic	On going	On going
Collinsella bouchesdurhonensis	+	-	-	Anaerobic	1.88	58
Bacteroides congonensis	-	+	-	Anaerobic	6.37	43
Parabacteroides timonensis	-	+	-	Anaerobic	6.48	43.4
Eggerthella timonensis	+	+	-	Anaerobic	3.92	65.2
Actinomyces minihominis	+	-	-	Facultative anaerobic	2.04	59.7
Neglecta bouchesdurhonensis	+	-	-	Anaerobic	2.7	39.8
Clostridium minihomine	+	-	-	Anaerobic	2.7	46.1
Dysgonomonas timonensis	-	-	-	Facultative anaerobic	3.3	36.6
Dysgonomonas phocaensis	-	-	-	Facultative anaerobic	On going	On going
Enterococcus mediteraneensis	+	-	-	Facultative anaerobic	On going	On going
Intestinibacillus timonensis	-	-	-	Anaerobic	On going	On going

II. iii) Discussion:

Overall, we managed to isolate 179 different bacterial species by culturomics from the selected stool samples of the Pygmy people out of which 9 were new genera and 18 were new species. The isolated species belonged mainly to *Frimicutes* (60%), *Bacteroidetes* (17%), *Actinobacteria* (15%), *Proteobacteria* (6%), 1 *Fusobacteria* species and 1 *Synergistetes* species (Figure 1).

9 (5%) of these species were previously isolated from human but not from the gut, 4 were previously not isolated from human, 27 (15%) were new species isolated in the current study and 115 (64%) were previously know or new species isolated by culturomics in other studies. Out of the 179 reported species, 110 (61%) were anaerobic and 69 (39%) were species that are able to tolerate oxygen. As for the metagenomics results, 246 different species were recorded in total along with a significant number of un-classified sequences. Species belonged majorly to *Firmicutes* (43%), *Proteobacteria* (24%), *Bacteroidetes* (14%), *Actinobacteria* (13%), *Fusobacteria* (2%), 2 *Lentisphaerae/Verrucomicrobia* species, 2 Spirochaetes and 2 *Synergistetes* species (Figure 2).

Of these 246, 128(52%) were species that can tolerate oxygen and 118(48%) were anaerobic. Nevertheless, 28(11%) were species previously isolated from human and 20(8%) were previously not known to be isolated from human. Metagenomics recovered only 26% of the species isolated by

culturomics. Additionally, when looking into OTUs in the un-classified sequences as previously described (1), we were able to identify 211 species out of which 71 were previously identified in table 2. Out of the 140 not identified in table 2, 13 were new species isolated in the current project and 69 were new species isolated by previous culturomics studies (Table 6). Thus the number of the species identified after OTUs targeting was increased to 387. Additionally, of the 211 identified OTUs, 23 were previously isolated in human but not in the gut and 8 were previously not previously isolated from the human body. 77 of the identified OTUs were aero tolerant and 137 were strictly anaerobic. In addition to the 211 OTUs identified, 102 OTUs were detected and corresponds to bacterial species that were isolated by culturomics but no yet identified and are currently in the process of genomic analyses, with 11 being isolated in the current study. The fact that only 26% of the isolated species by culturomics were recovered by metagenomics, and that a significant number of the operational taxonomic units sequences corresponded to new species isolated by culturomics, render culturomics not only a complementary but an essential approach in describing the human gut microbiota.

To be able to describe the gut microbiota of an entire population efficiently, more samples should be studied for a better representation. In addition, unhealthy individuals should be included in order to be able to create a comparative model. However, the currently provided data reveals the

dominance of the *Firmicutes,* which is typical, in the gut of the pygmy population. The significant number of new species isolated in the small number of studied samples reveals the importance of continuing efforts to describe the gut bacterial content of this population, which represent a reservoir of unknown species. The ability to keep on culturing new bacterial species is essential for enriching the currently available genomic databases and studying the role of the commensal microbiota in health and diseases.

1) Lagier J-C, Khelaifia S, Alou MT, Ndongo S, Dione N, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 2016;1:16203.

II. iii) Discussion:

Au total, nous avons réussi à isoler 179 espèces bactériennes différentes par culturomics à partir des échantillons de selles sélectionnés chez les Pygmées. Parmi ces espèces, 9 étaient de nouveaux genres et 18 étaient de nouvelles espèces. Les espèces isolées appartenaient principalement aux *Frimicutes* (60%), aux *Bacteroidetes* (17%), aux *Actinobactéries* (15%), aux *Protéobactéries* (6%), 1 espèce au *Fusobacteria* et 1 espèce au *Synergistetes* (Figure 1).

Neuf (5%) de ces espèces étaient déjà isolées chez l'homme mais pas au niveau de l'intestin, 4 n'étaient jamais isolées chez l'homme, 27 (15%) étaient de nouvelles espèces isolées dans la présente étude et 115 (64%) étaient auparavant connues ou étaient de nouvelles espèces isolées par la culturomics dans d'autres études. Sur les 179 espèces, 110 (61%) étaient anaérobies et 69 (39%) étaient des espèces capables de tolérer l'oxygène. En ce qui concerne les résultats de la métagénomique, 246 espèces différentes ont été enregistrées au total ainsi qu'un nombre significatif de séquences non classifiées. Les espèces appartenaient majoritairement aux *Firmicutes* (43%), aux *Protéobactéries* (24%), aux *Bacteroidetes* (14%), aux *Actinobactéries* (13%), aux *Fusobactéries* (2%), 2 espèces aux *Lentisphaerae* / *Verrucomicrobia,* 2 espèces aux Spirochètes et 2 espèces aux *Synergistetes* (Figure 2).

Parmi ces 246 espèces, 128 (52%) étaient capables de tolérer l'oxygène et

118 (48%) étaient des anaérobies. De plus, 28 (11%) étaient des espèces précédemment isolées chez l'homme et 20 (8%) n'ont jamais été isolées chez l'homme. La métagénomique n'a détecté que 26% des espèces isolées par la culturomics.

De plus, en examinant les OTUs dans les séquences non classifiées comme décrit précédemment (1), nous avons pu identifier 211 espèces dont 71 ont été précédemment identifiées dans le tableau 2. Sur les 140 non identifiées dans le tableau 2, 13 étaient de nouvelles espèces isolées dans le projet actuel et 69 étaient de nouvelles espèces isolées par des études culturomics antérieures (tableau 6). Ainsi, le nombre d'espèces identifiées après le ciblage des OTUs a augmenté à 387. En outre, sur les 211 UTO identifiées, 23 étaient auparavant isolées chez l'homme mais pas au niveau de l'intestin et 8 n'ont jamais été isolées chez l'homme. 77 des OTUs identifiés étaient tolérants à l'oxygène et 137 étaient strictement anaérobies. En plus des 211 OTUs identifiées, 102 OTUs ont été détectées et correspondent à des espèces bactériennes isolées par la culturomics mais non encore identifiées, qui sont actuellement en cours d'analyse génomique.

Le fait que seulement 26% des espèces isolées par la culturomics ont été détectée par métagénomique, et qu'un nombre significatif des séquences taxonomiques opérationnelles correspondent à de nouvelles espèces isolées par la culturomics, font de cette méthode non seulement une approche complémentaire mais essentielle pour décrire le microbiote intestinal chez

l’homme.

Pour décrire efficacement le microbiote intestinal d'une population entière, il faut étudier un plus grand nombre d'échantillons pour une meilleure représentation. En outre, des patients devraient être inclus afin de pouvoir créer un modèle comparatif. Cependant, les données actuellement fournies révèlent la prédominance des *Firmicutes*, ce qui est typique dans l'intestin de la population Pygmée. Le nombre important de nouvelles espèces isolées par rapport au petit nombre d'échantillons étudiés révèle l'importance de poursuivre les efforts pour décrire le contenu bactérien de l'intestin de cette population, qui représente un réservoir d'espèces inconnues. La capacité de continuer la culture de nouvelles espèces bactériennes est essentielle pour enrichir les bases de données génomiques actuellement disponibles et étudier le rôle du microbiote commensal chez l’homme.

1) Lagier J-C, Khelaifia S, Alou MT, Ndongo S, Dione N, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 2016;1:16203.

Table 6: List of the 211 OTUs identified in this study with their main features. A: new bacterial species isolated by culturomics; B: previously detected in table 2; C: not identified in table 2; D: new species isolated in the present work; *H: Human; NH: Non human and HG: Human gut

	Classification *	**Phylum**	**Oxygen requirement**	***Source of isolation**
Acidaminococcus massiliensis	A	Firmicutes	1	HG
Acinetobacter indicus	C	Proteobacteria	0	NH
Acinetobacter lwoffii	C	Proteobacteria	0	HG
Acinetobacter septicus	C	Proteobacteria	0	HG
Acinetobacter variabilis	C	Proteobacteria	0	HG
Actinomyces bouchesdurhonensis	A	Actinobacteria	0	HG
Actinomyces grossensis	A	Actinobacteria	0	HG
Actinomyces marseillensis	A	Actinobacteria	0	H
Actinomyces minihominis	AD	Actinobacteria	0	HG
Actinomyces odontolyticus	B	Actinobacteria	0	HG
Actinomyces Turicensis	A	Actinobacteria	0	HG
Akkermansia muciniphila	B	Verrucomicrobia	0	HG
Alistipes jeddahensis	A	Bacteroidetes	1	HG
Alistipes massiliensis	A	Bacteroidetes	1	HG
Alistipes obesi	AB	Bacteroidetes	1	HG
Alistipes phoceensis	A	Bacteroidetes	1	HG
Alistipes provencensis	A	Bacteroidetes	1	HG
Alistipes putredinis	B	Bacteroidetes	1	HG
Alistipes senegalensis	AB	Bacteroidetes	1	HG
Alistipes shahii	B	Bacteroidetes	1	HG
Allisonella histaminiformans	B	Firmicutes	1	HG
Anaerocolumna aminovalericum	C	Firmicutes	1	NH
Anaerolactis massiliensis	A	Firmicutes	1	H
Anaerostipes Caccae	A	Firmicutes	1	HG
Atopobium parvulum	B	Actinobacteria	0	HG

Arcobacter butzeleri	C	Proteobacteria	0	HG
Bacillus altudinis	C	Firmicutes	0	HG
Bacillus aquimaris	C	Firmicutes	0	HG
Bacillus cereus	C	Firmicutes	0	HG
Bacillus dakarensis	A	Firmicutes	0	HG
Bacillus flexus	B	Firmicutes	0	HG
Bacillus mediterraneensis	A	Firmicutes	0	HG
Bacillus mojavensis	C	Firmicutes	0	HG
Bacillus subtilis	B	Firmicutes	0	HG
Bacillus vallimortis	C	Firmicutes	0	HG
Bacteroides bouchesdurhonense	A	Bacteroidetes	1	HG
Bacteroides caccae	AB	Bacteroidetes	1	HG
Bacteroides congonensis	AD	Bacteroidetes	1	HG
Bacteroides fragilis	B	Bacteroidetes	1	HG
Bacteroides intestinalis	AB	Bacteroidetes	1	HG
Bacteroides ndongoniae	C	Bacteroidetes	1	H
Bacteroides nordii	B	Bacteroidetes	1	HG
Bacteroides oleiciplenus	B	Bacteroidetes	1	HG
Bacteroides Ovatus	B	Bacteroidetes	1	HG
Bacteroides uniformis	B	Bacteroidetes	1	HG
Bacteroides xylanolyticus	C	Bacteroidetes	1	HG
Beduinella massiliensis	A	Firmicutes	1	HG
Bifidobacterium adolescentis	B	Actinobacteria	1	HG
Bifidobacterium longum	B	Actinobacteria	1	HG
Bifidobacterium pseudocatenulatum	B	Actinobacteria	1	HG
Bilophila wadsworthia	B	Proteobacteria	1	HG
Blautia maliae	A	Firmicutes	1	H
Blautia massiliensis	A	Firmicutes	1	HG
Blautia mediterraneensis	A	Firmicutes	1	H
Blautia phocaeensis	A	Firmicutes	1	HG
Blautia timonensis	A	Firmicutes	1	HG
Butyricimonas phoceensis	AB	Bacteroidetes	1	HG
Casaltella massiliensis	A	Firmicutes	1	HG

Catenibacterium mitsuokai	B	Firmicutes	1	HG
Citrobacter freundii	C	Proteobacteria	0	HG
Clostridium beduini	A	Firmicutes	1	HG
Clostridium bifermentans	B	Firmicutes	1	HG
Clostridium bouchesdurhonense	A	Firmicutes	1	HG
Clostridium butyricum	B	Firmicutes	1	HG
Clostridium caccae	C	Firmicutes	1	HG
Clostridium citroniae	C	Firmicutes	1	HG
Clostridium marseillense	A	Firmicutes	1	HG
Clostridium massiliosenegalense	A	Firmicutes	1	HG
Clostridium moniliforme	C	Firmicutes	1	H
Clostridium paraputrificum	B	Firmicutes	1	HG
Clostridium phoceense	A	Firmicutes	1	HG
Clostridium polynesiense	AB	Firmicutes	1	HG
Clostridium sartagoforme	B	Firmicutes	1	HG
Clostridium sordellii	B	Firmicutes	1	HG
Clostridium tertium	B	Firmicutes	1	HG
Clostridium tyrobutyricum	C	Firmicutes	1	HG
Colidextra massiliensis	A	Firmicutes	0	HG
Collinsella bouchedurhonensis	AD	Actinobacteria	1	HG
Collinsella massilioamazoniensis	A	Actinobacteria	1	HG
Congobacterium massiliense	AD	Firmicutes	1	HG
Corynebacterium accolens	C	Actinobacteria	0	HG
Cronobacter helveticus	C	Proteobacteria	0	NH
Culturomica massiliensis	A	Bacteroidetes	1	HG
Delftia tsuruhatensis	C	Proteobacteria	0	HG
Desulfomassilia massiliensis	A	Proteobacteria	1	HG
Desulfovibrio piger	C	Proteobacteria	1	HG
Dialister invisus	B	Firmicutes	1	HG
Dialister propionicifaciens	B	Firmicutes	1	H
Dorea longicatena	B	Firmicutes	1	HG
Dorea pacaensis	A	Firmicutes	1	HG
Dorea phocaensis	AD	Firmicutes	1	HG
Duodena massiliensis	A	Proteobacteria	1	HG

Dysgonomonas massiliensis	AD	Bacteroidetes	0	HG
Eisenbergiella massiliensis	A	Firmicutes	0	HG
Enterobacter Cloacae	AB	Proteobacteria	0	HG
Enterobacter gallinorum	C	Proteobacteria	0	HG
Enterobacter gergoriae	C	Proteobacteria	0	NH
Enterobacter kobei	C	Proteobacteria	0	HG
Enterococcus faecalis	AB	Firmicutes	0	HG
Enterococcus pseudoavium	B	Firmicutes	0	HG
Escherichia coli	B	Proteobacteria	0	HG
Eubacterium infirmum	B	Firmicutes	1	H
Ezakiella massiliensis	A	Firmicutes	1	H
Fusicatenibacter saccharivorans	B	Firmicutes	1	HG
Fusobacterium nucleatum	B	Fusobacteria	1	HG
Fusobacterium timonense	A	Fusobacteria	1	H
Gemella morbillorum	C	Firmicutes	0	HG
Gorbachella massiliensis	A	Firmicutes	1	HG
Gordonibacter massiliense	A	Actinobacteria	1	HG
Granulicatella adiacens	C	Firmicutes	0	HG
Halomonas meridiana	C	Proteobacteria	0	NH
Holdemanella biformis	B	Firmicutes	1	NH
Holdemania timonensis	A	Firmicutes	1	HG
Intestinibacillus timonensis	AD	Firmicutes	1	HG
Intestinibacter bartlettii	B	Firmicutes	1	HG
Intestinimonas gabonensis	A	Firmicutes	1	HG
Intestinimonas Massiliensis	AB	Firmicutes	1	HG
Jeddahella massiliensis	A	Actinobacteria	0	HG
Klebsiella oxytoca	C	Proteobacteria	0	HG
Klebsiella pneumoniae	B	Proteobacteria	0	HG
Kurthia timonensis	A	Firmicutes	0	HG
Lachnoclostridium edouardii	A	Firmicutes	1	HG
Lachnoclostridium lavalense	A	Firmicutes	1	H
Lachnoclostridium urinomassiliense	A	Firmicutes	1	H
Lactobacillus brevis	B	Firmicutes	0	HG

Lactobacillus crustorum	B	Firmicutes	0	NH
Lactobacillus fermentum	C	Firmicutes	0	HG
Lactobacillus mucosae	B	Firmicutes	0	HG
Lactobacillus plantarum	B	Firmicutes	0	HG
Lactobacillus ruminis	AB	Firmicutes	0	HG
Lactobacillus sakei	C	Firmicutes	0	HG
Lactobacillus vaccinostercus	B	Firmicutes	0	HG
Lactococcus Garvieae	B	Firmicutes	0	HG
Lactococcus lactis	B	Firmicutes	0	HG
Lactonifactor massiliensis	A	Firmicutes	1	H
Leuconostoc lactis	C	Firmicutes	0	HG
Leuconostoc pseudomesenteroides	C	Firmicutes	0	HG
Libanioccus massiliensis	AD	Actinobacteria	1	HG
Mailhella massiliensis	A	Proteobacteria	1	HG
Marseillobacter massiliensis	A	Firmicutes	1	HG
Massilioprevotella massiliensis	A	Firmicutes	1	HG
Megasphaera indica	C	Firmicutes	1	HG
Methanobrevibacter smithii	A	Euryarchaeota	1	HG
Micrococcus lylae	C	Actinobacteria	0	HG
Micromassilia vaginalis	A	Firmicutes	1	H
Mitsuokella jalaludinii	C	Firmicutes	1	HG
Moellerella wisconsensis	C	Proteobacteria	0	HG
Moryella massiliensis	A	Firmicutes	1	H
Negativibacillus massiliensis	A	Firmicutes	1	HG
Neisseria flavescens	B	Proteobacteria	1	HG
Niameyia massiliensis	B	Firmicutes	1	HG
Odoribacter splanchnicus	B	Bacteroidetes	1	HG
Olsenella congonensis	AD	Actinobacteria	1	HG
Olsenella profusa	C	Actinobacteria	1	H
Olsenella scatoligenes	C	Actinobacteria	1	NH
Oscilibacter timonense	A	Firmicutes	1	HG
Oscillibacter massiliensis	A	Firmicutes	1	HG
Pantoea Dispersa	C	Proteobacteria	0	H
Pantoea Eucrina	C	Proteobacteria	0	H

Parabacteroides bouchesdurhonensis	A	Bacteroidetes	1	HG
Parabacteroides faecis	A	Bacteroidetes	1	HG
Parabacteroides gordonii	B	Bacteroidetes	1	HG
Parabacteroides massiliensis	A	Bacteroidetes	1	HG
Parasutterella excrementihominis	C	Proteobacteria	1	HG
Peptoniphilus koenoeneniae	C	Firmicutes	1	HG
Phascolarctobacterium faecium	C	Firmicutes	1	HG
Phascolarctobacterium succinatutens	B	Firmicutes	1	H
Phoenicibacter massiliensis	AD	Actinobacteria	1	HG
Prevotella caccae	A	Bacteroidetes	1	HG
Prevotella caldimerdae	AB	Bacteroidetes	1	HG
Prevotella massiliensis	A	Bacteroidetes	1	H
Prevotella provencensis	A	Bacteroidetes	1	HG
Prevotella salivae	C	Bacteroidetes	1	H
Prevotellamassilia phoceensis	A	Bacteroidetes	1	HG
Prevotellamassilia timonensis	A	Bacteroidetes	1	HG
Providencia heimbachae	C	Proteobacteria	0	HG
Providencia stuartii	C	Proteobacteria	0	HG
Pseudomonas fluorescens	C	Proteobacteria	0	HG
Pseudomonas lundensis	C	Proteobacteria	0	HG
Pseudoruminococcus massiliensis	A	Firmicutes	0	HG
Pygmaiobacter massiliensis	AD	Firmicutes	1	HG
Pyramidobacter piscolens	B	Synergistetes	1	HG
Raoultella ornithinolytica	C	Proteobacteria	0	HG
Raoultibacter massiliensis	AD	Actinobacteria	0	HG
Raoultibacter timonensis	AB	Actinobacteria	0	HG
Robinsoniella peoriensis	B	Firmicutes	1	HG
Romboutsia lituseburensis	B	Firmicutes	1	HG
Romboutsia massiliensis	A	Firmicutes	1	HG
Romboutsia timonensis	A	Firmicutes	1	HG
Ruminococcus bicirculans	B	Firmicutes	1	HG
Ruminococcus gauvreauii	C	Firmicutes	1	HG

Ruminococcus phoceensis	A	Firmicutes	1	HG
Rummeliibacillus suwonensis	C	Firmicutes	0	HG
Ruthenibacterium lactatiformans	B	Firmicutes	1	HG
Salmonella enterica	B	Proteobacteria	0	HG
Sanguibacter massiliense	AD	Firmicutes	1	HG
Senegalimassilia anaerobia	AB	Actinobacteria	1	HG
Slackia isoflavoniconvertens	B	Actinobacteria	1	HG
Stenotrophomonas acidaminiphila	B	Proteobacteria	0	HG
Streptoccoccus pneumoniae	A	Firmicutes	0	H
Streptococcus alactolyticus	C	Firmicutes	0	HG
Streptococcus constellatus	C	Firmicutes	0	HG
Streptococcus dentisani	C	Firmicutes	0	H
Streptococcus Thermophilus	C	Firmicutes	0	HG
Sutterella wadsworthensis	B	Proteobacteria	1	HG
Tidjanibacter massiliensis	A	Bacteroidetes	1	HG
Veillonella parvula	B	Firmicutes	1	HG
Victivallis vadensis	B	Lentisphaerae	1	HG
Vitreoscilla massiliensis	A	Proteobacteria	1	HG
Weissella cibaria	C	Firmicutes	0	HG
Weissella viridescens	C	Firmicutes	0	HG

Figure 1: Distribution of the species isolated by culturomics according to their phyla.

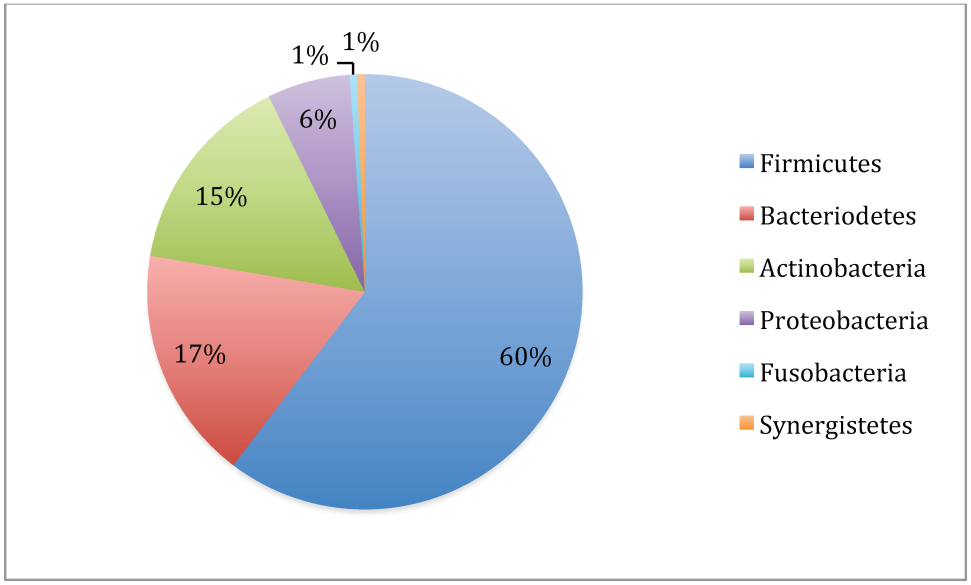
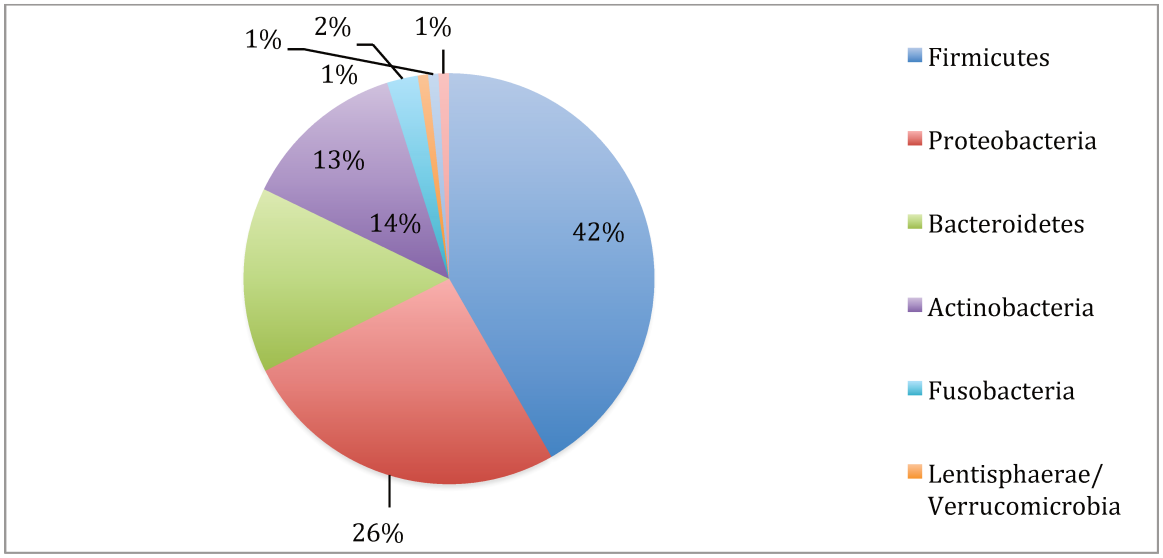


Figure 2: Distribution of the species recovered by metagenomics according to its phyla.



II. iv) Associated Articles
II. iv) Articles associés

Article 1

Sanguibacter massiliensis sp. nov ., *Actinomyces minihominis* sp. nov ., *Clostridium minihomine* sp. nov., *Neobittarella massiliensis* gen. nov., and *Miniphocibacter massiliensis* gen. nov., new bacterial species isolated by culturomics from human stool samples

Published in NMNI journal

Sanguibacter massiliensis sp. nov., *Actinomyces minihominis* sp. nov., *Clostridium minihomine* sp. nov., *Neobittarella massiliensis* gen. nov. and *Miniphocibacter massiliensis* gen. nov., new bacterial species isolated by culturomics from human stool samples

M. Bilen[1,2], M. D. Mbogning Founkou[1], F. Cadoret[4], G. Dubourg[1], Z. Daoud[2] and D. Raoult[1,3]

1) *Aix Marseille Univ, IRD, MEPHI, IHU-Méditerranée Infection, Marseille, France,* 2) *Clinical Microbiology Department, Faculty of Medicine and Medical Sciences, University of Balamand, Amioun, Lebanon,* 3) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia and* 4) *Assistance Publique Hôpitaux de Marseille (AP-HM), IHU Méditerranée Infection, Marseille, France*

Abstract

We present the description of *Sanguibacter massiliensis* sp. nov., *Actinomyces minihominis* sp. nov., *Clostridium minihomine* sp. nov., *Neobittarella massiliensis* gen. nov. and *Miniphocibacter massiliensis* gen. nov., new bacterial species isolated by culturomics from human stool samples.

Keywords: *Actinomyces minihominis* sp. nov., *Clostridium minihomine* sp. nov., *Miniphocibacter massiliensis* gen. nov., *Neobittarella massiliensis* gen. nov., *Sanguibacter massiliensis* sp. nov

Original Submission: 9 February 2018; **Revised Submission:** 5 March 2018; **Accepted:** 9 March 2018

Article published online: 16 March 2018

Corresponding author. D. Raoult, Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), CNRS 7278, IRD 198, INSERM 1095, UM63, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, 27 Boulevard Jean Moulin, 13385, Marseille cedex 5, France.
E-mail: didier.raoult@gmail.com

As part of the effort aiming to describe the human gut microbiota by the means of culturomics [1], we report in the present study the isolation of four new bacterial species and two new bacterial genera. Before initiation of the study, approval from the Institut Federatif de Recherche IFR48 (Marseille, France) was obtained under number 09-022, along with signed informed consent provided by the donors.

All stool samples were independently diluted with phosphate-buffered saline and incubated in a blood culture bottle supplemented with 5% sheep's blood and 5% filtered rumen. A 30-day follow-up was done in order to assess the bacterial growth, and matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) and 16S rRNA gene sequencing were used for strain identification, as previously described [2,3]. Because the organisms' spectra were missing from the MALDI-TOF MS database, we were unable to identify them. Thus, 16S rRNA gene sequencing was applied. A similarity threshold of less than 98.65% between the isolated strains and the phylogenetically closest species with standing in nomenclature was adapted to delimitate a new species, and a divergence of more than 5% was adapted to delimitate a new genus [4].

Strain Marseille-P3815 was isolated on Columbia medium supplemented with 5% sheep's blood (COS; bioMérieux, Marcy l'Étoile, France) from the stool samples of a 12-year-old healthy Pygmy girl after 2 days' incubation under anaerobic conditions at 37°C. This strain is a Gram-positive rod, catalase positive but oxidase negative. Its colonies have a diameter ranging between 0.2 to 0.9 mm, with its cells' average size being 1.0 × 0.4 µm. It has a smooth and grey appearance after 48 hours of growth on COS medium. Strain Marseille-P3815 exhibited a 96.96% sequence similarity with *Sanguibacter inulinus* strain ST50 (NR_029277.1). It was thus classified as a new species, *Sanguibacter massiliensis* (ma.ssi.lien′sis, L. adj. neut., *massiliensis,* from 'Massilia,' the antic name of Marseille, France, where the

https://doi.org/10.1016/j.nmni.2018.03.002

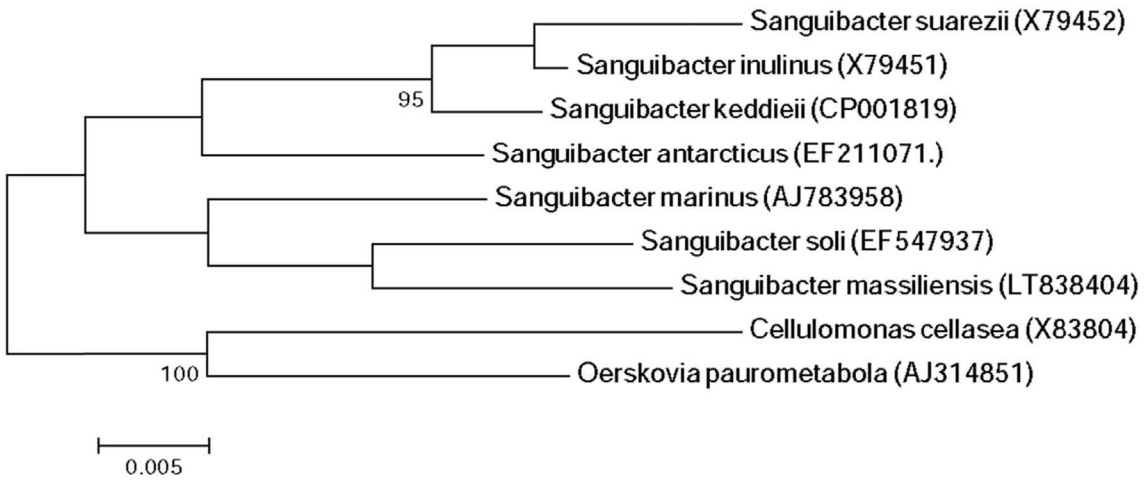


FIG. 1. Phylogenetic tree showing position of *Sanguibacter massiliensis* strain Marseille-P3815 towards its closest species. CLUSTALW was used for alignment and MEGA software for phylogenetic inferences generation with maximum likelihood method. After 500 repeats, bootstrap values are shown on nodes, with only values above 90% kept.

strain was isolated) (Fig. 1). The strain Marseille-P3815^T is the type strain of the species *Sanguibacter massiliensis*.

Strain Marseille-P3850 was isolated on COS (bioMérieux) from the stool samples of a 12-year-old healthy Pygmy girl after 5 days' incubation under anaerobic conditions at 37°C. This strain is a Gram-positive rod, catalase and oxidase negative. Its colonies have a diameter ranging between 0.3 to 1 mm, with its cells' average size being 1.0 × 0.5 µm. It has a smooth and grey appearance after 48 hours of growth on COS medium. Strain Marseille-P3850 exhibited a 93.36% 16S rRNA gene sequence similarity with *Actinomyces marimammalium* strain CCUG 41710 (NR_025395.1). Thus it is classified as a new species, *Actinomyces minihominis* (mini.ho.min′is, L. adj. neut., *minihominis,* referring to the Pygmy people from whom the strain was isolated and who are characterized by their small size) (Fig. 2). The strain Marseille-P3850^T is the type strain of the species *Actinomyces minihominis*.

Strain Marseille-P4642 was isolated on COS (bioMérieux) from the stool samples of a 39-year-old healthy Pygmy man after 10 days' incubation under anaerobic conditions at 37°C. This strain is a Gram-positive bacilli, catalase and oxidase negative. Its colonies have a diameter ranging between 0.05 to 1 mm, with its cells' average size being 2.2 × 0.43 µm. It has a smooth and grey appearance after 48 hours of growth on COS medium. Strain Marseille-P4642 exhibited a 98.13% 16S rRNA gene sequence similarity with *Clostridium jeddahense* strain JCD (NR_144697.1); thus, it is classified as a new species, *Clostridium minihomine* (mini.ho.min′e, L. adj. neut., *minihomine,* to refer to the Pygmy people from whom the strain was isolated and who are characterized by their small size) (Fig. 3). The strain Marseille-P4642^T is the type strain of the species *Clostridium minihomine*.

Strain Marseille-P4047 was isolated on COS medium (bioMérieux) from the stool samples of a healthy Senegalese male subject after 5 days' incubation under anaerobic conditions at 37°C. This strain is a Gram-positive coccobacilli, catalase and oxidase negative. Its colonies have a diameter ranging between 0.04 to 0.9 mm, with its cells' average size being 1.6 × 0.6 µm. It has a shiny grey appearance after 48 hours of growth on COS medium. Strain Marseille-P4047 exhibited an 89.22% 16S rRNA gene sequence similarity with *Acetanaerobacterium elongatum* strain Z7 (NR_042930.1); thus, it is classified as a new genus, *Neobitarella* (neo.bita.rel′la, L. adj. fem., in honor of microbiologist Fadi Bittar). *Neobitarella massiliensis* is the type species of the new genus *Neobittarella* (ma.ssi.lien′sis, L. adj. fem., *massiliensis,* from 'Massilia,' the antic name of Marseille, France, where the strain was isolated) (Fig. 4). The strain Marseille-P4047^T is the type strain of the species *Neobittarella massiliensis*.

Strain Marseille-P4678 was isolated on COS (bioMérieux) from the stool samples of a 39-year-old healthy Pygmy man after 10 days' incubation under anaerobic conditions at 37 °C. This strain is a Gram-positive cocci, catalase positive and oxidase negative. Its colonies have a diameter ranging between 0.02 to 0.08 mm, with its cells' average diameter being 0.6 µm. It has a smooth and grey appearance after 48 hours of growth on COS medium. Strain Marseille-P4047 exhibited an 89.69% 16S rRNA gene sequence similarity with *Parvimonas micra* strain 3119B (NR_036934.1); it is thus classified as a new genus, *Miniphocibacter* (mini.phoci.bacter, L. adj. masc., *miniphocibacter* composed of *mini,* referring to the small size of the Pygmy

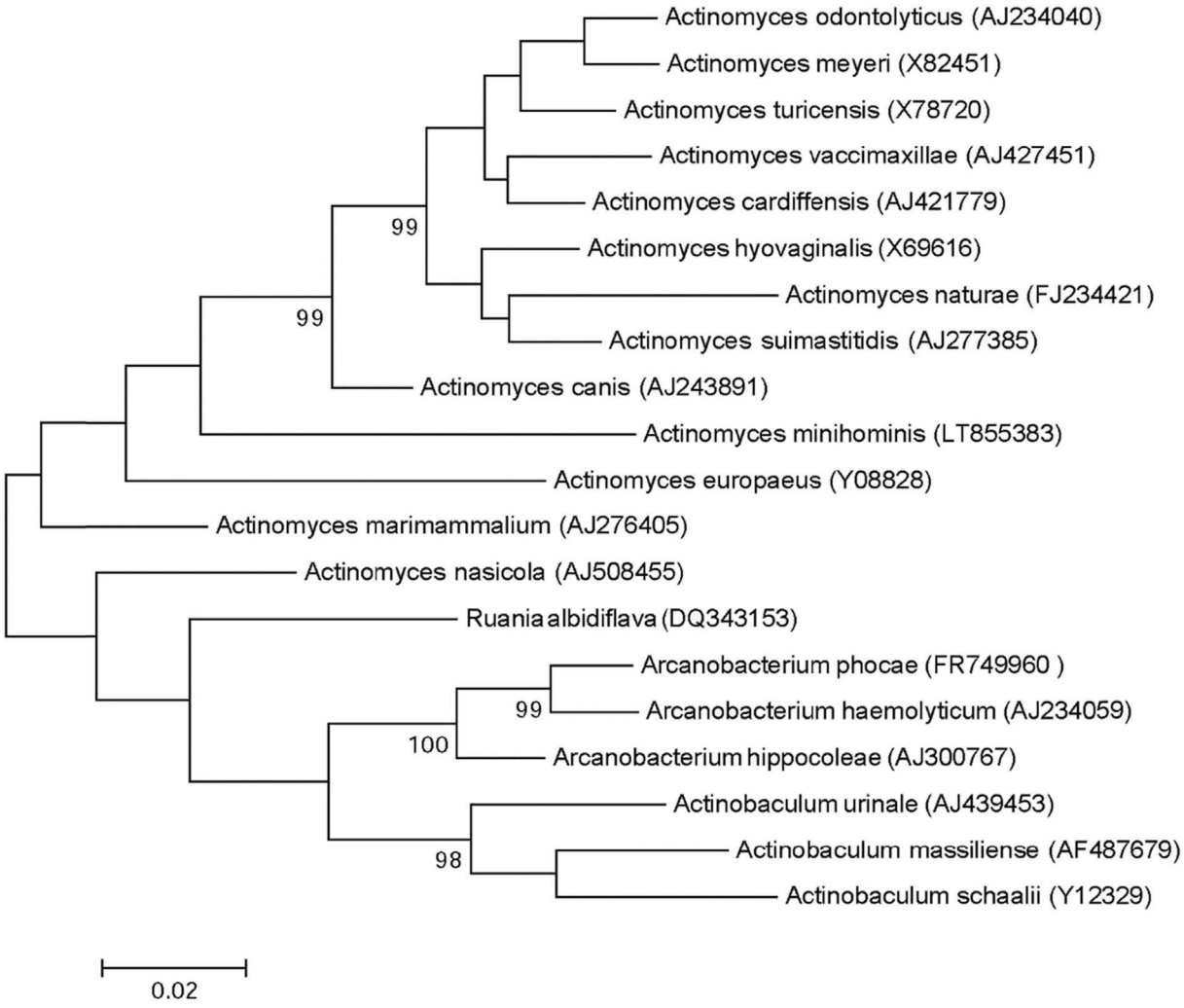


FIG. 2. Phylogenetic tree showing position of *Actinomyces minihominis* strain Marseille-P3850 towards its closest species. CLUSTALW was used for alignment and MEGA software for phylogenetic inferences generation with maximum likelihood method. After 500 repeats, bootstrap values are shown on nodes, with only values above 90% kept.

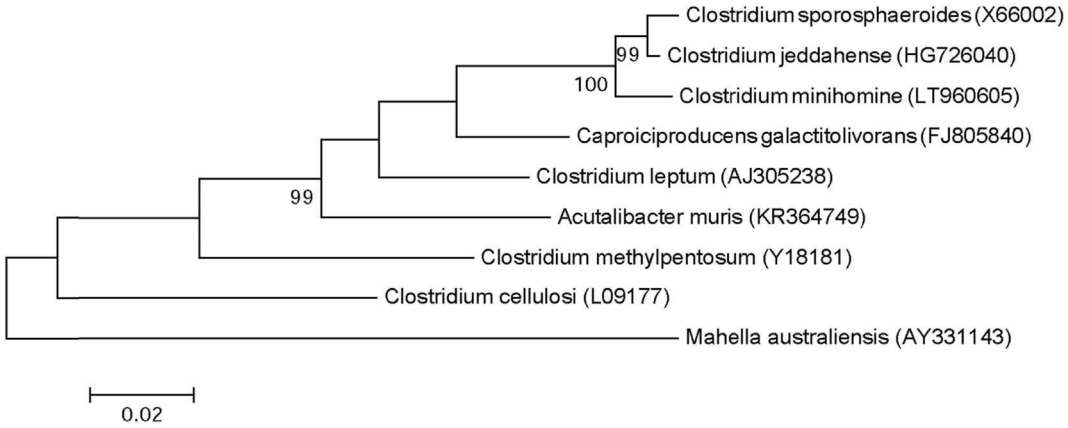


FIG. 3. Phylogenetic tree showing position of *Clostridium minihomine* strain Marseille-P4042 towards its closest species. CLUSTALW was used for alignment and MEGA software for phylogenetic inferences generation with maximum likelihood method. After 500 repeats, bootstrap values are shown on nodes, with only values above 95% kept.

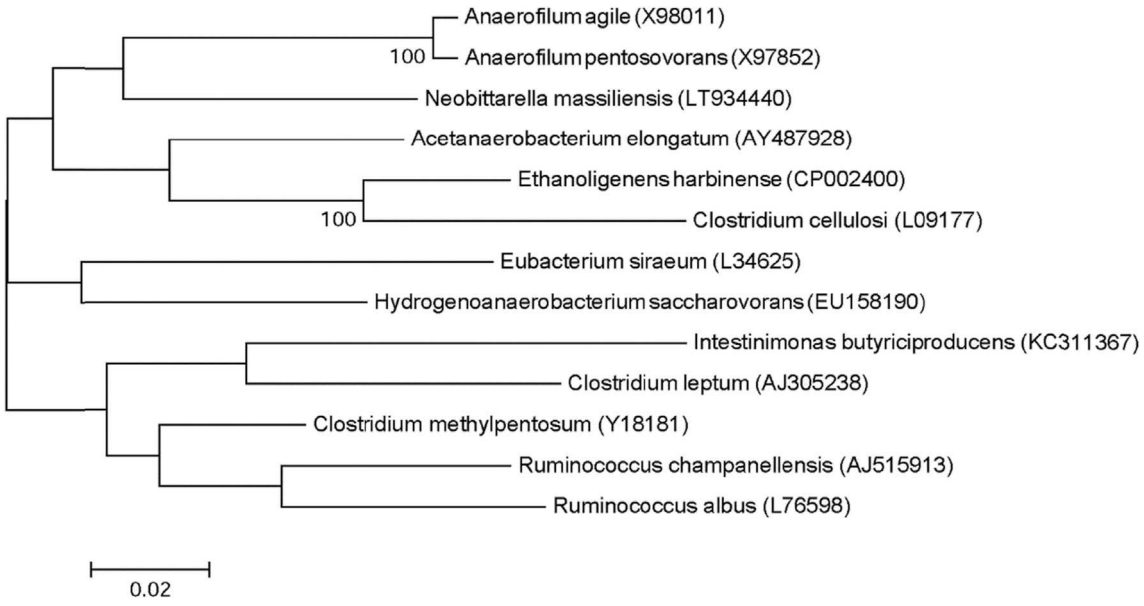


FIG. 4. Phylogenetic tree showing position of *Neobittarella massiliensis* strain Marseille-P4047 towards its closest species. CLUSTALW was used for alignment and MEGA software for phylogenetic inferences generation with maximum likelihood method. After 500 repeats, bootstrap values are shown on nodes, with only values above 95% kept.

people from whom the strain was isolated, and *phoci,* referring to Phocae, the Latin name of the city that the founders of Marseille came from). *Miniphocibacter massiliensis* is the type species of the genus *Miniphocibacter* (ma.ssi.lien′sis, L. adj. masc., *massiliensis,* from 'Massilia,' the antic name of Marseille, France, where the strain was isolated) (Fig. 5). The strain Marseille-P4678^T is the type strain of the species *Miniphocibacter massiliensis.*

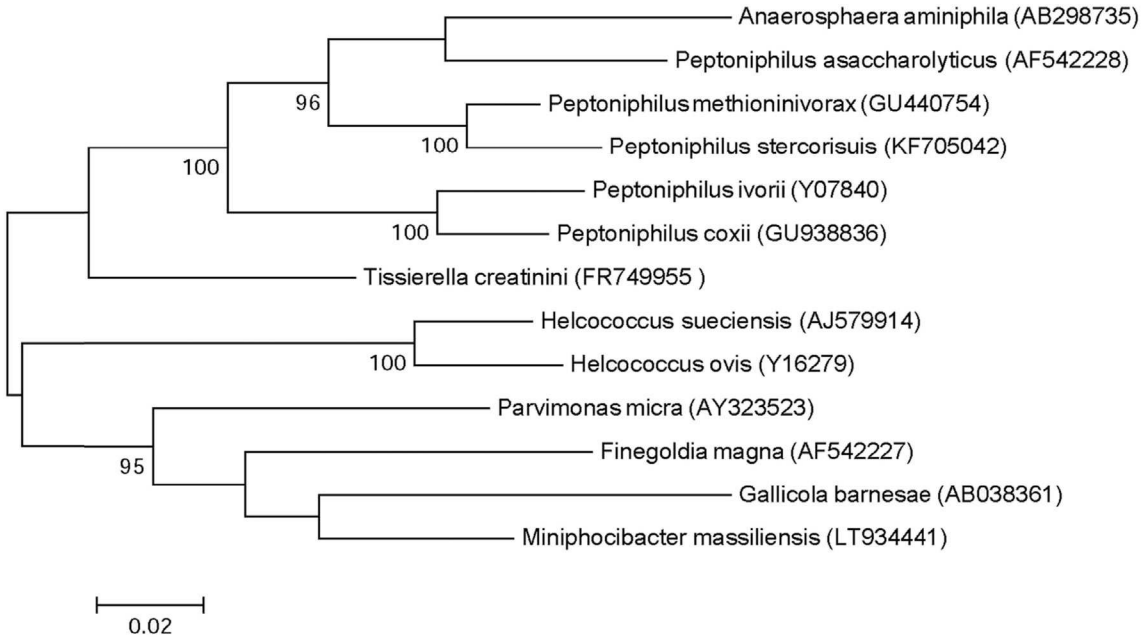


FIG. 5. Phylogenetic tree showing position of *Miniphocibacter massiliensis* strain Marseille-P4678 towards its closest species. CLUSTALW was used for alignment and MEGA software for phylogenetic inferences generation with maximum likelihood method. After 500 repeats, bootstrap values are shown on nodes, with only values above 95% kept.

MALDI-TOF MS spectrum

The MALDI-TOF MS spectrum of the reported organisms are available online (http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database).

Nucleotide sequence accession number

The 16S rRNA gene sequences of *Sanguibacter massiliensis* sp. nov., *Actinomyces minihominis* sp. nov., *Clostridium minihomine* sp. nov., *Neobittarella massiliensis* gen. nov. and *Miniphocibacter massiliensis* gen. nov. were deposited in GenBank under the following accession numbers respectively: LT838404, LT855383, LT960605, LT934440 and LT934441.

Deposit in a culture collection

Strain Marseille-P3815, strain Marseille-P3850, strain Marseille-P4642, strain Marseille-P4047 and strain Marseille-P4678 were deposited in the Collection de Souches de l'Unité des Rickettsies (CSUR, WDCM 875) under number the following numbers respectively: P3815, P3850, P4642, P4047 and P4678.

Acknowledgement

This study was funded by the Fondation Méditerranée Infection.

Conflict of interest

None declared.

References

[1] Lagier JC, Hugon P, Khelaifia 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.
[2] 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.
[3] Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, Raoult D. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 2000;38:3623–30.
[4] Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 2014;64(Pt 2):346–51.

Article 2

Taxonogenomics description of *Libanicoccus massiliensis* gen. nov., sp. nov.

Published in NMNI journal

ORIGINAL ARTICLE

Libanicoccus massiliensis gen. nov., sp. nov., a new bacterium isolated from human stool

M. Bilen[1,2], F. Cadoret[1], M. Richez[1], E. Tomei[1], Z. Daoud[2], D. Raoult[1,3] and P.-E. Fournier[1]

1) *Aix-Marseille Université, Institut hospitalo-universitaire Méditerranée-infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Marseille, France,* 2) *Clinical Microbiology Department, Faculty of Medicine and Medical Sciences, University of Balamand, Amioun, Lebanon and* 3) *Special Infectious Agents Unit, King Fahd Medical Research Centre, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

Strain Marseille-P3237 was isolated from a stool sample of a healthy 35-year-old Congolese pygmy female. This anaerobic, Gram-negative, non-spore-forming and non-motile coccus-shaped bacterium is a member of the order *Coriobacteriales*. It exhibits a 2 009 306-bp genome with a 65.46 mol% G+C content and is closely related to, but distinct from, members of the *Olsenella* genus. We propose the creation of the new genus *Libanicoccus* gen. nov. and of the new species *Libanicoccus massiliensis* sp. nov.

Keywords: Culturomics, Genome, *Libanicoccus massiliensis*, Pygmy, Taxonogenomics
Original Submission: 31 October 2017; **Revised Submission:** 6 November 2017; **Accepted:** 6 November 2017
Article published online: 10 November 2017

Corresponding author: P.E. Fournier, Aix-Marseille Université, Institut hospitalo-universitaire Méditerranée-infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, 19–21 Boulevard Jean Moulin, 13005 Marseille, France.
E-mail: pierre-edouard.fournier@univ-amu.fr

Introduction

The human gut contains 10^{11} to 10^{12} bacteria per gram of stool. This complex microflora is known for its microbial diversity and role in health and diseases [1]. Deciphering the gut microbiota has become a challenge in the twenty-first century [2] and has been attempted using different tools yielding increasingly complex results [3]. To date, more than 2000 different bacterial species belonging to the human gut microbiota have been reported [4]. In our laboratory, we have developed a new technique named culturomics to isolate previously uncultured human gut bacteria [5,6]. Basically, stool samples are cultured under various conditions and all isolated colonies are identified using matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS). Among bacterial isolates that fail MALDI-TOF MS identification, those that show sufficient 16S rRNA gene sequence divergence with species with standing in nomenclature are further characterized using the taxonogenomics strategy, which combines phenotypic assays and genome sequencing and analysis [3,7]. In the present study, using the taxonogenomics approach, we describe the new genus *Libanicoccus* gen. nov. within the family *Coriobacteriaceae*. Strain Marseille-P3237^T (= CSUR P3237 = CCUG 71182) is the type strain of the new species *Libanicoccus massiliensis* gen. nov., sp. nov.

Material and methods

Ethics and sample collection

A stool sample from a healthy 35-year-old pygmy woman was collected in Congo and preserved at −80°C for further analysis at the URMITE Laboratory (Marseille, France). The sample donor gave a signed and informed consent. The study was approved by the ethics committee of the Institut Fédératif de Recherche IFR48 (Marseille, France) under number 09-022.

https://doi.org/10.1016/j.nmni.2017.11.001

Strain isolation

The stool sample was diluted in PBS (Life Technologies, Carlsbad, CA, USA) and pre-incubated for 3 days in a blood culture vial (BD BACTEC®, Plus Anaerobic/F Media, Le Pont de Claix, France) supplemented with 5 mL of sheep blood and 5 mL of filter-sterilized rumen at 37°C. Then, the culture suspension was inoculated on 5% sheep blood-enriched Columbia agar (bioMérieux, Marcy l'Etoile, France) and incubated at 37°C in anaerobic atmosphere.

MALDI-TOF MS and 16S rRNA gene sequencing identification

The individual identification of isolated colonies was first attempted using MALDI-TOF MS, as previously described [5,6]. The reference spectrum obtained for each colony was compared with the Bruker database using the MALDI BIOTYPER software version 3.0 (Bruker Daltonics, Bremen, Germany). Any score <1.9 was considered unreliable. In this case, colonies were subjected to 16S rRNA gene amplification and sequencing, using a GeneAmp 2720 thermal cycler (Applied Biosystems, Foster City, CA, USA) and an ABI Prism 3130xl Genetic Analyzer capillary sequencer (Applied Biosystems) as previously described [8]. Each 16S rRNA nucleotide sequence was compared with the nr database of the National Center for Biotechnology Information using the BLAST software (https://blast.ncbi.nlm.nih.gov/). We used the 16S rRNA sequence similarity thresholds of 95% and 98.65% proposed by Kim et al. to consider bacterial isolates as putatively belonging to a new genus or a new species without performing DNA–DNA hybridization [9]. Finally, to determine the phylogenetic position of strain Marseille-P3237 with regard to species with standing in nomenclature, its 16S rRNA gene sequence was compared with the 16S rRNA database of the 'All-Species Living Tree' project of Silva (LTPs121) [10]. Sequence alignment was obtained using Muscle [11] and phylogenetic relationships were inferred with the maximum-likelihood method within the FastTree software [12].

Growth conditions

Culture of strain Marseille-P3237 was attempted using several conditions to determine its optimal growth requirements. First, strain Marseille P3237 was inoculated on 5% sheep blood-enriched Columbia agar (bioMérieux) and incubated in aerobic, micro-aerophilic and anaerobic conditions at 28, 37, 45 and 55°C. The GENbag anaer and GENbag microaer systems (bioMérieux) were used to evaluate the bacterial growth in anaerobic and microaerophilic atmospheres, respectively. In addition, optimal halophily and pH were estimated using 0, 5, 15 and 45% NaCl concentrations and pH values of 6, 6.5, 7 and 7.5, respectively.

Morphological and biochemical assays

The API 20A, API ZYM and API 50CH strips (bioMérieux) were used to biochemically characterize strain Marseille-P3237. Sporulation ability was tested after exposing a bacterial suspension to a thermic shock at 80°C for 10 min. Motility was evaluated using a DM1000 photonic microscope (Leica Microsystems, Nanterre, France) with a 1000× magnification. Cell morphology was observed using electron microscopy and the following protocol. Bacteria were fixed with 2.5% glutaraldehyde in 0.1 M cacodylate buffer for at least 1 h at 4°C. Then, a drop of cell suspension was deposited for approximately 5 min on glow-discharged formvar carbon film on 400-mesh nickel grids (FCF400-Ni, EMS). The grids were dried on blotting paper and cells were negatively stained for 10 seconds with a solution of 1% ammonium molybdate in filtered water at room temperature. Electron micrographs were acquired with a Tecnai G20 Cryo (FEI) transmission electron microscope operated at 200 keV.

Fatty acid methyl ester analysis

Cellular fatty acid methyl ester analysis was performed by gas chromatography/mass spectrometry. Two samples were prepared with approximately 21 mg of bacterial biomass per tube, harvested from several culture plates. FAME analysis was carried out as previously described [5].

Antibiotic susceptibility testing

The antibiotic susceptibility of strain Marseille-P3237 was assessed using the E-test method for the following molecules: benzylpenicillin, amoxicillin, cefotaxime, ceftriaxone, imipenem, amikacin, erythromycin, daptomycin, rifampicin, minocycline, teicoplanin, vancomycin, colistin and metronidazole (bioMérieux).

Genomic DNA extraction and genome sequencing

Genomic DNA (gDNA) of strain Marseille-P3237 was extracted as previously described [5]. A final concentration of 67.8 ng µL was measured with the Qubit assay and the high sensitivity kit (Life Technologies, Carlsbad, CA, USA). Afterwards, gDNA was sequenced on a MiSeq sequencer (Illumina, San Diego, CA, USA). Briefly, 1.5 µg of gDNA was used for mate-pair library preparation using the Nextera mate pair Illumina guide (Illumina). After tagmentation and fragmentation of the gDNA with a mate-pair junction adapter, the fragmentation pattern was confirmed using an Agilent 2100 BioAnalyzer (Agilent Technologies Inc, Santa Clara, CA, USA) with a DNA 7500 labchip. The DNA fragments ranged in size from 1.5 kb up to 11 kb with an optimal size at 5.282 kb. No size selection was performed and 191.8 ng of tagmented fragments were circularized. Mechanical shearing of the circularized DNA was

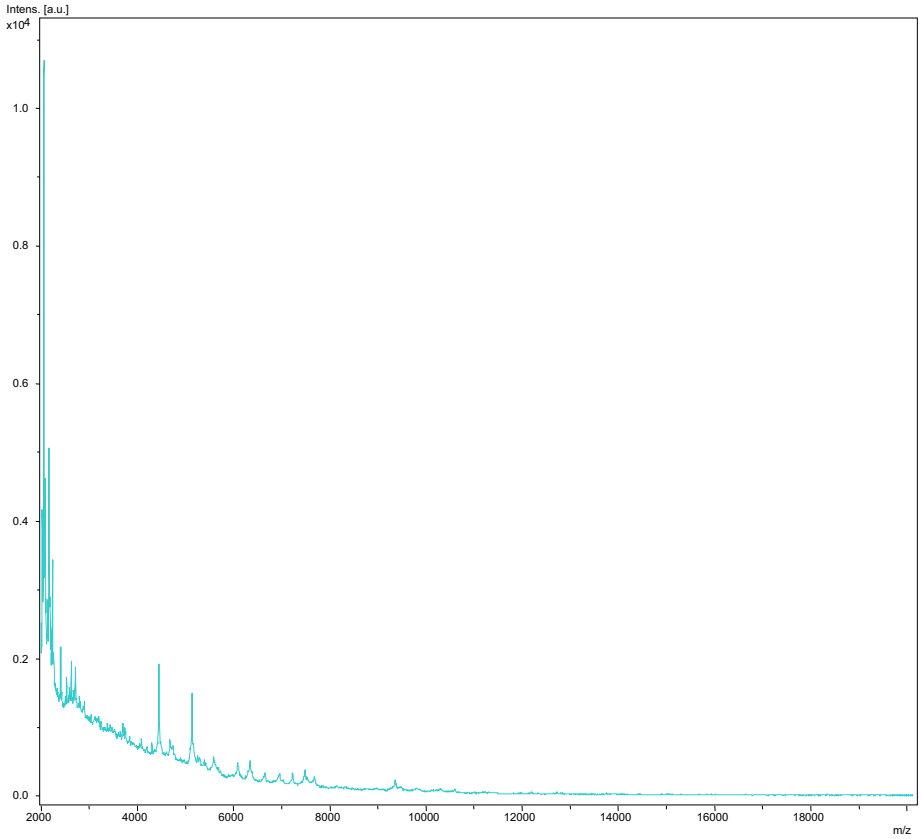


FIG. 1. MALDI-TOF MS-generated mass spectrum of *Libanicoccus massiliensis* gen. nov., sp. nov., strain Marseille-P3237^T.

performed with an optimal size of 1261 bp on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). Library profile visualization was performed on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc) and the final concentration library was detected as 0.7626 nmol/L. Libraries were normalized to 2 nM, followed by a denaturation step and dilution to reach a 15 pM concentration. An automated cluster generation and sequencing run were performed in a single 39-h run in a 2 × 251-bp. Total information of 11.1 Gb was obtained from a 1332 K/mm^2 cluster density with a cluster passing quality control filters of 87.9% (21 937 000 passing filter paired reads). Within this run, the index representation for strain Marseille-P3237 was determined to be 3.05%. The 668 978 paired-reads were trimmed and then assembled.

The genome of strain Marseille-P3237 was assembled, annotated and compared with other closely related species as previously described [5]. The genomic comparison included *Olsenella profusa* (GenBank accession number NZ_AWEZ00000000.1), *Olsenella uli* (NZ_JQCO00000000.1), *Slackia exigua* (ACUX00000000), *Atopobium vaginae* (NZ_ACGK00000000.2), *Atopobium parvulum* (NC_013203.1), *Atopobium rimae* (NZ_ACFE00000000.1), *Atopobium minutum* (NZ_AGXC00000000.1), *Collinsella tanakaei* (NZ_ADLS00000000.1) and *Collinsella intestinalis* (NZ_ABXH00000000.2).

Results and discussion

Strain identification and phylogenetic analysis

The identification of strain Marseille-P3237 using MALDI-TOF MS failed. The generated reference mass spectrum (Fig. 1) was added to the URMS database (http://www.mediterranee-infection.com/article.php?larub=280&titre=urms-database).

Strain Marseille-P3237 exhibited a 93.05% sequence identity with *O. uli* (GenBank accession number AF292373), the phylogenetically closest species with a validly published name (Fig. 2). This 16S rRNA gene sequence divergence being greater than 5%, we investigated whether strain Marseille-P3237 could be the representative strain of a new species within a new genus (Table 1) [13]. The 16S rRNA gene sequence from strain

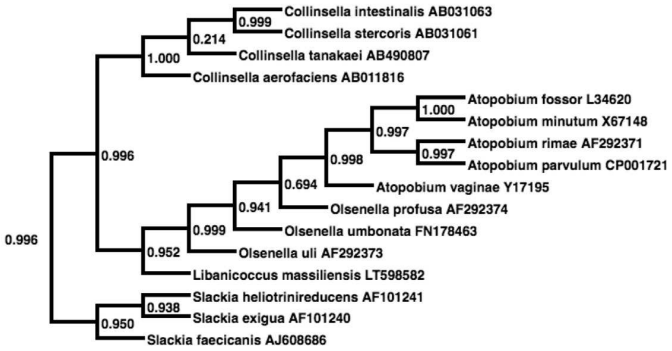


FIG. 2. Phylogenetic position of *Libanicoccus massiliensis* strain Marseille-P3237 relative to other closely related species. Phylogenetic relationships were inferred using the maximum-likelihood method. Numbers at the nodes are bootstrap values obtained from 1000 replicates with Shimodaira–Hasegawa test. The scale bar represents a 1% nucleotide sequence divergence.

TABLE 1. General features of *Libanicoccus massiliensis* gen. nov., sp. nov. strain Marseille-P3237^T

Properties	Term
Current classification	Domain: *Bacteria* Phylum: *Actinobacteria* Class: *Coriobacteria* Order: *Coriobacteriales* Family: *Coriobacteriaceae* Genus: *Libanicoccus* Species: *Libanicoccus massiliensis* Type Strain: Marseille-P3237^T
Gram stain	Negative
Cell shape	Coccus
Motility	Non-motile
Sporulation	Negative
Growth temperature range	30°C–42°C
Optimal growth temperature	37°C

Marseille-P3237 was submitted to EMBL-EBI under accession number LT598582. The MALDI-TOF MS spectrum of strain Marseille-P3237 was compared with those of other members of the *Coriobacteriaceae* family (Fig. 3).

Phenotypic and biochemical characterization

Strain Marseille-P3237 grew at an optimal temperature of 37°C under anaerobic conditions (Fig. 4). No growth was observed in aerobic or microaerophilic atmospheres. It also tolerated a pH range between 6 and 8.5 but could not sustain NaCl concentrations >5 g/L. In optimal culture conditions, bacterial colonies were dark white and rough, with a diameter of 0.8–1.2 mm. Cells of strain Marseille-P3237 were not motile, unable to sporulate, Gram-negative cocci. They were also catalase- and oxidase-negative. Using electron microscopy, bacterial cells had a mean diameter of 1.06 µm (Fig. 5).

Using an API 20A strip, strain Marseille-P3237 could hydrolyse esculin but could neither produce indole nor hydrolyse gelatine. In addition, it was urease negative and could not acidify D-glucose, D-saccharose, D-lactose, D-melezitose, D-xylose, D-mannose, D-trehalose, L-rhamnose, L-arabinose, D-raffinose, D-mannitol, D-cellobiose, glycerol, salicin, D-maltose and D-sorbitol.

Using an API ZYM strip, strain Marseille-P3237 exhibited α-fucosidase, alkaline phosphatase, acid phosphatase, esterase lipase (C8), Valine arylamidase, Leucine arylamidase and naphthol-AS-BI phosphohydrolase activities but was negative for α-galactosidase, β-galactosidase, α-chymotrypsin, β-glucuronidase, α-glucosidase, β-glucosidase, esterase (C4), *N*-acetyl-β-glucosaminidase, lipase (C14), Cystine arylamidase, trypsin and α-mannosidase activities.

Using an API 50CH strip, the following sugars were fermented: D-ribose, L-arabinose, D-galactose, D-xylose, glycerol, D-fructose, D-glucose, D-mannose, D-sorbitol, methyl-αD-mannopyranoside, *N*-acetylglucosamine, amygdaline, D-mannitol, salicin, arbutine, D-cellobiose, esculin, D-lactose, D-saccharose, D-melibiose, D-maltose, D-melezitose, D-trehalose, starch, xylitol, D-tagatose, D-raffinose, potassium gluconate and gentobiose. In contrast, strain Marseille-P3237 exhibited negative reactions for inositol, erythritol, D-lyxose, L-fucose, LL-sorbose, D-arabitol, L-arabitol, L-rhamnose, inulin, methyl-αD-glucosamin, potassium 5-ketogluconate, D-turanose, D-adonitol, glycogen, methyl-β-D-xylopyranoside, L-xylose, D-fucose, D-ulcitol and D-arabinose. By comparison with members of the *Olsenella* genus, strain Marseille-P3237 differed in acid production from D-mannose and D-glucose (Table 2).

The major fatty acid found for this strain was 9-octadecenoic acid (18:1n9, 38%). The other most abundant fatty acids were the saturated structures hexadecanoic acid (16:0, 28 %) and octadecanoic acid (18:0, 11 %) (Table 3).

Strain Marseille-P3237 MICs (mg/L) of 0.38, 0.75, 1, 1, 0.047, >256, 0.016, 1.5, 0.19, 4, 0.5, 32, >256, 0.125 and for benzylpenicillin, amoxicillin, cefotaxime, ceftriaxone, imipenem, amikacin, erythromycin, daptomycin, rifampicin, minocycline,

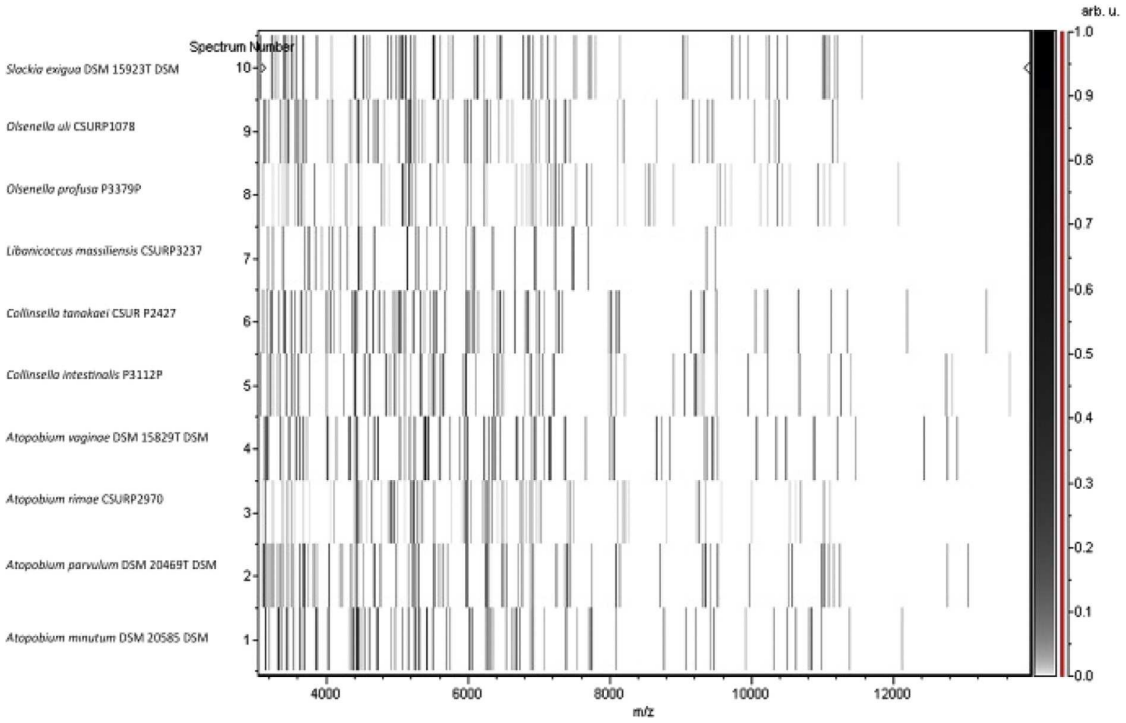


FIG. 3. Gel view comparing *Libanicoccus massiliensis* gen. nov., sp. nov. strain Marseille-P3237^T with other closely related species. The gel view displays the raw spectra of loaded spectrum files arranged in a pseudo-gel like look. M/z values are shown on the *x* axis and the left *y*-axis represents the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a Grey scale scheme code. The right *y*-axis indicates the relation between the colour of a peak and its intensity, in arbitrary units.

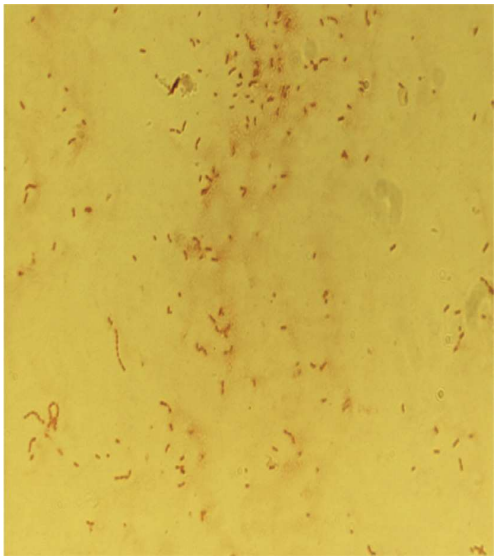


FIG. 4. Gram staining of *Libanicoccus massiliensis* gen. nov., sp. nov. strain Marseille-P3237^T.

teicoplanin, vancomycin, colistin and metronidazole, respectively.

Genome properties

The genome from strain Marseille-P3237 is 2 009 306-bp long with a 65.46 mol% G+C content (Table 4). It is composed of one scaffold (two contigs). Of the 1805 predicted genes, 1747 are protein-coding genes and 58 are RNAs (two complete rRNA operons and an additional 5S rRNA and 51 tRNA genes). A total of 1375 genes (78.71%) are assigned a putative function (by BLAST against COGs or nr). A total of 108 genes are identified as ORFans (6.18%). The remaining 216 genes are annotated as hypothetical proteins (12.36%, Table 4). A graphical representation of the genome is depicted in Fig. 6. The distribution of genes into COG functional categories is presented in Table 5.

Genomic comparison

The compared distribution of functional classes of predicted genes from strain Marseille-P3237 and closely related species, according to the COGs database, is represented in Fig. 7. The

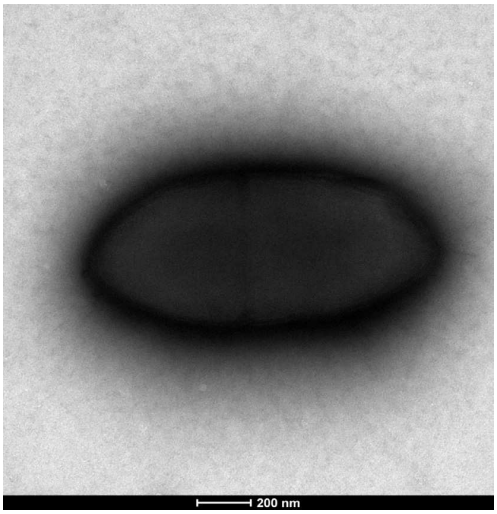


FIG. 5. Transmission electron microscopy of *Libanicoccus massiliensis* gen. nov., sp. nov. strain Marseille-P3237^T.

TABLE 2. Differential characteristics of *Libanicoccus massiliensis* gen. nov., sp. nov. strain Marseille-P3237^T (present study), *Olsenella scatoligenes* strain SK9K4^T [14], *Olsenella uli* strain VPI D76D-27C^T (15) and *Olsenella profusa* strain D315A-29^T [15]

Properties	*L. massiliensis*	*O. scatoligenes*	*O. uli*	*O. profusa*
Cell length (μm)	1.6	1–2	NA	0.8–2
Oxygen requirement	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic
Gram stain	Negative	Positive	Positive	Positive
Salt requirement	—	NA	+	+
Motility	—	—	—	—
Endospore formation	—	—	NA	—
Indole production	—	—	—	—
Production of:				
Alkaline phosphatase	+	+	NA	NA
Catalase	—	—	—	—
Oxidase	—	NA	NA	NA
Urease	—	—	NA	NA
β-galactosidase	—	+	NA	NA
N-acetyl-glucosamine	+	NA	NA	NA
Acid production from:				
L-arabinose	—	—	—	+
D-mannose	—	+	+	+
D-mannitol	—	+	—	+
D-glucose	—	+	+	+
D-maltose	—	+	—	+
D-lactose	—	+	—	+
G+C content (%)	65.46%	62.1	64	64
Habitat	Human gut	Pig gut	Human gingival crevices	Human subgingival plaque

NA, data not available.

genome from strain Marseille-P3237 was smaller than those of *O. profusa*, *O. uli*, *S. exigua* and *C. tanakaei* (2009, 2725, 2057, 2096 and 2492 Mb, respectively), but larger than those of *A. vaginae*, *A. parvulum*, *A. rimae*, *A. minutum* and *C. intestinalis* (1419, 1544, 1626, 1707 and 1809 Mb, respectively). The G+C content of strain Marseille-P3237 was larger than those of

TABLE 3. Cellular fatty acid composition of *Libanicoccus massiliensis* gen. nov., sp. nov. strain Marseille-P3237^T

Fatty acids	Name	Mean relative % [a]
18:1n9	9-Octadecenoic acid	37.8 ± 1.8
16:0	Hexadecanoic acid	28.4 ± 1.4
18:0	Octadecanoic acid	10.6 ± 0.5
14:0	Tetradecanoic acid	10.4 ± 0.5
18:1n5	13-Octadecenoic acid	7.2 ± 1.0
10:0	Decanoic acid	2.4 ± 0.4
18:2n6	9,12-Octadecadienoic acid	1.8 ± 0.7
15:0	Pentadecanoic acid	TR
12:0	Dodecanoic acid	TR
15:0 anteiso	12-methyl-tetradecanoic acid	TR

[a]Mean peak area percentage; TR = trace amounts <1%.

TABLE 4. Nucleotide content of strain Marseille-P3237^T and gene count levels of the genome

	Number	%
Size (bp)	2 009 306	100
Number of G+C nucleotides	1 313 861	65.46
Total number of genes	1805	100
Number of protein-coding genes	1747	96.79
Total number of RNA genes	58	3.21
Number of tRNA genes	51	2.82
Number of rRNA (5S, 16S, 23S) genes	7	0.39
Coding sequence size	1 782 220	88.7
Protein-coding gene sequence size	1 768 986	88.04
tRNA gene sequence size	3993	0.2
rRNA gene (5S, 16S, 23S) sequence size	9241	0.46
Number of proteins associated to COGs	1223	70.00
Number of proteins classified as orfans	108	6.18
Number of proteins with signal peptides	152	8.70
Number of genes associated with antibiotic resistance	1	0.06
Number of genes associated with PKS or NRPS	1	0.06
Number of genes associated with virulence	339	19.40

O. profusa, *O. uli*, *A. vaginae*, *A. parvulum*, *A. rimae*, *A. minutum*, *C. intestinalis*, *C. tanakaei* and *S. exigua* (65.46, 64.17, 64.69, 42.66, 45.69, 49.26, 48.94, 62.47, 60.24 and 62.15 mol% respectively). The gene content of strain Marseille-P3237 is smaller than those of *O. profusa*, *C. intestinalis*, *O. uli*, *C. tanakaei* and *S. exigua* (1747, 2650, 1784, 1770, 2212 and 2029, respectively), but larger than those of *A. vaginae*, *A. parvulum*, *A. rimae* and *A. minutum* (1179, 1353, 1548 and 1539, respectively). Moreover, strain Marseille-P3237 exhibited the highest Average Genomic Identity of Orthologous gene Sequences values (Table 6) with *O. profusa* and *O. uli* (61.51 and 62.78%, respectively) and dDDH values (Table 7) with these species (20.2% (18–22.6) and 19.9% (17.7–22.3), respectively) that were lower than that obtained within the *Olsenella* genus (22.4%; 20.1–24.8).

Conclusion

On the basis of phenotypic, biochemical, genomic and phylogenetic results, we formally propose the creation of the new genus *Libanicoccus* gen. nov., with *Libanicoccus massiliensis*

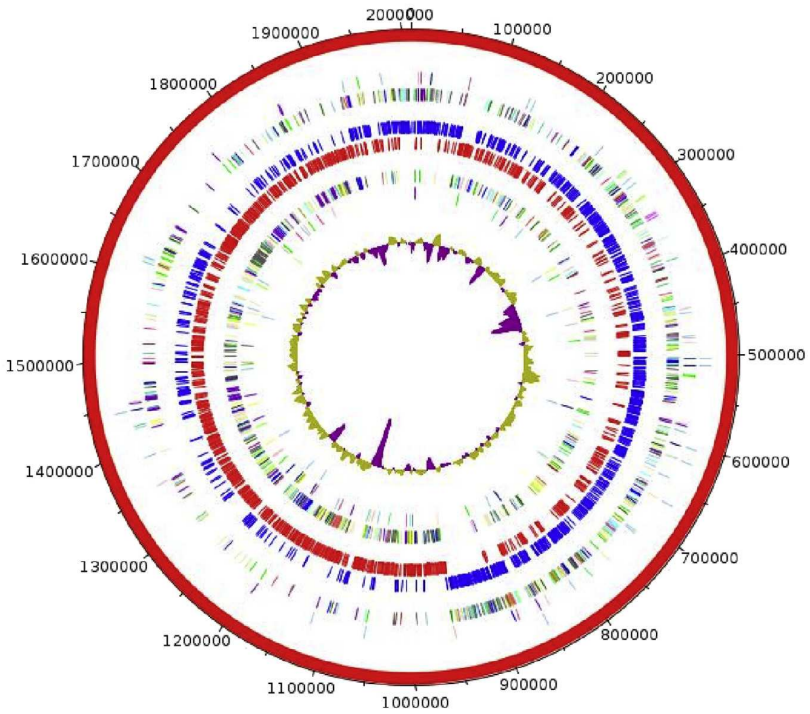


FIG. 6. Graphical circular map of the genome of *Libanicoccus massiliensis* strain Marseille-P3237^T. From outside in: contigs (red/grey), COG categories of genes on the forward strand (three circles), genes on the forward strand (blue circle), genes on the reverse strand (red circle), COG categories on the reverse strand (three circles), G+C content.

TABLE 5. **Numbers of genes associated with the 25 general COG functional categories**

Code	Value	% of total[a]	Description
[J]	160	9.16	Translation
[A]	0	0	RNA processing and modification
[K]	85	4.86	Transcription
[L]	71	4.06	Replication, recombination and repair
[B]	0	0	Chromatin structure and dynamics
[D]	20	1.14	Cell cycle control, mitosis and meiosis
[Y]	0	0	Nuclear structure
[V]	37	2.12	Defence mechanisms
[T]	51	2.92	Signal transduction mechanisms
[M]	65	3.72	Cell wall/membrane biogenesis
[N]	7	0.40	Cell motility
[Z]	0	0	Cytoskeleton
[W]	4	0.23	Extracellular structures
[U]	18	1.03	Intracellular trafficking and secretion
[O]	43	2.46	Post-translational modification, protein turnover,chaperones
[X]	19	1.09	Mobilome: prophages, transposons
[C]	69	3.95	Energy production and conversion
[G]	142	8.13	Carbohydrate transport and metabolism
[E]	159	9.10	Amino acid transport and metabolism
[F]	56	3.20	Nucleotide transport and metabolism
[H]	63	3.61	Coenzyme transport and metabolism
[I]	48	2.75	Lipid transport and metabolism
[P]	64	3.66	Inorganic ion transport and metabolism
[Q]	26	1.49	Secondary metabolites biosynthesis, transport and catabolism
[R]	119	6.81	General function prediction only
[S]	48	2.75	Function unknown
–	524	29.99	Not in COGs

[a]The total is based on the total number of protein-coding genes in the annotated genome.

sp. nov. being the type species. Strain Marseille-P3237^T, isolated from the gut microbiota of a healthy 35-year-old Congolese pigmy female, is the type strain of *L. massiliensis* gen. nov., sp. nov.

Description of *Libanicoccus* gen. nov.

Libanicoccus gen. nov. (li. ba.ni.coc'cus N.L. masc. n, *Libanicoccus*, composed of *Liban*, the country of origin of the microbiologist who first cultivated strain Marseille-P3237^T, and *coccus* for the shape of bacterial cells).

Gram-negative cocci. Strictly anaerobic. Mesophilic. Not motile. Unable to sporulate. Absent catalase, oxidase and indole productions. Hydrolyses esculin. Positive for α-fucosidase, alkaline phosphatase, acid phosphatase, esterase lipase (C8), valine arylamidase, leucine arylamidase and naphthol-AS-BI phosphohydrolase activities. Ferments D-ribose, L-arabinose, D-galactose, D-xylose, glycerol, D-fructose, D-glucose, D-mannose, D-sorbitol, methyl-αD-mannopyranoside, *N*-acetyl-glucosamine, amygdaline, D-mannitol, salicin, arbutine, D-cellobiose, esculin, D-lactose, D-saccharose, D-melibiose, D-maltose, D-melezitose, D-trehalose, starch, xylitol, D-tagatose, D-raffinose, potassium gluconate and gentobiose. Habitat: human digestive tract. Type species: *Libanicoccus massiliensis*. Type strain *Libanicoccus massiliensis* strain Marseille-P3237^T (= CSUR P3237 = CCUG 71182).

Description of *Libanicoccus massiliensis* sp. nov.

Libanicoccus massiliensis, sp. nov. (ma.ssi.li.en'sis. L. masc. adj. *massiliensis* pertaining to Massilia, the Roman name of Marseille, where strain Marseille-P3237^T was first isolated).

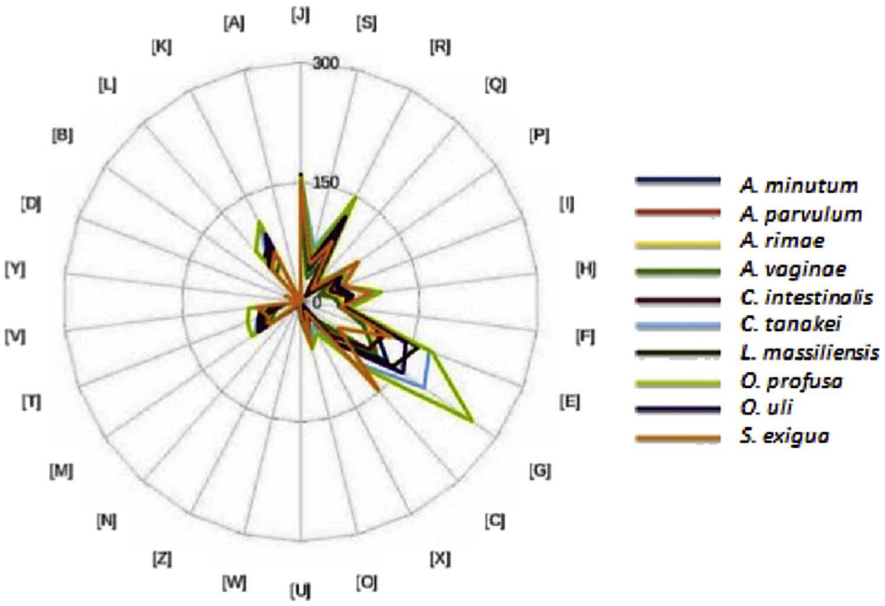


FIG. 7. Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins.

TABLE 6. Numbers of orthologous proteins between strain Marseille-P3237^T and other closely related species (upper right), numbers of proteins per genome (bold numbers), and Average Genomic Identity of Orthologous gene Sequences values (%, lower left)

	AV	OP	AR	SE	OU	LM	AM	CI	CT	AP
AV	**1179**	676	645	482	695	629	680	601	613	658
OP	51.87	**2650**	808	624	920	865	799	759	875	818
AR	54.15	59.38	**1548**	560	818	734	743	669	714	904
SE	49.53	60.25	57.01	**2029**	624	591	544	577	625	560
OU	53.06	62.94	56.53	60.81	**1770**	814	820	747	826	809
LM	58.27	61.51	53.92	56.9	62.78	**1747**	738	796	870	751
AM	57.22	55.35	57.54	54.14	58.02	56.91	**1539**	704	768	767
CI	51.9	61.42	56.84	60.42	62.31	60.82	55.67	**1784**	928	691
CT	53.82	59.71	56.27	58.88	59.6	60.92	55.29	63.12	**2212**	738
AP	56.19	57.91	63.66	54.48	54.38	54.37	57.78	54.92	55.34	**1353**

Abbreviations: AV, *Atopobium vaginae* strain DSM 15829; OP, *Olsenella profusa* strain DSM 13989; AR, *Atopobium rimae* strain ATCC 49626; SE, *Slackia exigua* strain ATCC 700122; OU, *Olsenella uli* strain DSM 7084; LM, *Libanicoccus massiliensis* strain Marseille-P3237^T; AM, *Atopobium minutum* strain NCFB 2751; CI, *Collinsella intestinalis* strain DSM 13280; CT, *Collinsella tanakaei* strain YIT 12063; AP, *Atopobium parvulum* strain DSM 20469.

TABLE 7. Digital DNA–DNA hybridization values (%) obtained by pairwise genomic comparison of strain Marseille-P3237^T with other closely related species using the GGDC formula 2 software. The inherent uncertainty in approximating dDDH values from intergenomic distances established on models derived from empirical test data sets are represented in confidence intervals.

	LM	AV	OP	AR	SE	OU	AM	CI	CT	AP
LM	100	17.2 [15.1–19.5]	20.2 [18–22.6]	20.3 [18.1–22.7]	19.8 [17.6–22.2]	19.9 [17.7–22.3]	23.4 [21.1–25.8]	21.4 [19.2–23.9]	21.2 [19–23.7]	21.9 [19.6–24.3]
AV		100	19 [16.8–21.4]	23.3 [21.1–25.8]	24.5 [22.1–26.9]	20.3 [18.1–22.7]	20.7 [18.5–23.1]	24.5 [22.2–27]	26.9 [24.5–29.4]	20.3 [18.1–22.8]
OP			100	21.6 [19.3–24]	20.4 [18.2–22.8]	22.4 [20.1–24.8]	20.7 [18.5–23.2]	19.5 [17.3–21.9]	20 [17.8–22.5]	24 [21.7–26.5]
AR				100	21.8 [19.5–24.2]	25.2 [22.9–27.7]	25.5 [23.2–28]	24.7 [22.4–27.2]	22 [19.7–24.4]	23.9 [21.6–26.4]
SE					100	32.6 [30.2–35.1]	20.6 [18.3–23]	18.9 [16.7–21.2]	20.5 [18.3–22.9]	22.1 [19.8–24.5]
OU						100	21.4 [19.1–23.8]	20.4 [18.2–22.8]	19.5 [17.3–21.9]	25.3 [23–27.8]
AM							100	22.5 [20.2–25]	22.2 [20–24.7]	20.2 [18–22.6]
CI								100	24.7 [22.3–27.1]	22 [19.7–24.4]
CT									100	20.8 [18.6–23.3]
AP										100

Abbreviations: AV, *Atopobium vaginae* strain DSM 15829; OP, *Olsenella profusa* strain DSM 13989; AR, *Atopobium rimae* strain ATCC 49626; SE, *Slackia exigua* strain ATCC 700122; OU, *Olsenella uli* strain DSM 7084; LM, *Libanicoccus massiliensis* strain Marseille-P3237^T; AM, *Atopobium minutum* strain NCFB 2751; CI, *Collinsella intestinalis* strain DSM 13280; CT, *Collinsella tanakaei* strain YIT 12063, AP, *Atopobium parvulum* strain DSM 20469.

Colonies are dark white and rough with a diameter of 0.8–1.2 mm on blood-enriched Columbia agar. Strictly anaerobic. Mesophilic. Can grow at a pH range of 6.0 to 8.5 but cannot grow at NaCl concentrations >5 g/L. Cells are Gram-negative cocci with a mean diameter of 1.06 µm. Not motile and non-sporulating. Catalase, oxidase and indole negative.

Using an API 20A strip, strain Marseille-P3237 could hydrolyse esculin but not gelatin. In addition, it was urease negative and could not acidify D-glucose, D-saccharose, D-lactose, D-melezitose, D-xylose, D-mannose, D-trehalose, L-rhamnose, L-arabinose, D-raffinose, D-mannitol, D-cellobiose, glycerol, salicin, D-maltose and D-sorbitol. Using an API ZYM strip, strain Marseille-P3237 exhibited α-fucosidase, alkaline phosphatase, acid phosphatase, esterase lipase (C8), valine arylamidase, leucine arylamidase and naphthol-AS-BI phosphohydrolase activities but was negative for α-galactosidase, β-galactosidase, α-chymotrypsin, β-glucuronidase, α-glucosidase, β-glucosidase, esterase (C4), *N*-acetyl-β-glucosaminidase, lipase (C14), cystine arylamidase, trypsin, and α-mannosidase activities. Using an API 50CH strip, D-ribose, L-arabinose, D-galactose, D-xylose, glycerol, D-fructose, D-glucose, D-mannose, D-sorbitol, methyl-αD-mannopyranoside, *N*-acetyl-glucosamine, amygdaline, D-mannitol, salicin, arbutine, D-cellobiose, esculin, D-lactose, D-saccharose, D-melibiose, D-maltose, D-melezitose, D-trehalose, starch, xylitol, D-tagatose, D-raffinose, potassium gluconate and gentiobiose were fermented. Inositol, erythritol, D-lyxose, L-fucose, L-sorbose, D-arabitol, L-arabitol, L-rhamnose, inulin, methyl-αD-glucosamin, potassium 5-ketogluconate, D-turanose, D-adonitol, glycogen, methyl-β-D-xylopyranoside, L-xylose, D-fucose, D-ulcitol and D-arabinose were not fermented.

The major fatty acid is 9-octadecenoic acid (18:1n9, 38 %), followed by hexadecanoic acid (16:0, 28 %) and octadecanoic acid (18:0, 11 %).

The 16S rRNA gene and genome sequences are deposited in the EBI/EMBL database under accession numbers LT598582 and LT671675, respectively. The G+C content of the genome is 65.46%. Habitat: human digestive tract. Type strain *Libanicoccus massiliensis* strain Marseille-P3237^T (= CSUR P3237 = CCUG 71182).

Acknowledgments

The authors thank the Xegen Company (http://www.xegen.fr/) for automating the genomic annotation process. This study was funded by the Méditerranée Infection foundation.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. Curr Opin Gastroenterol 2015;31:69–75.
[2] Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature 2007;449:804–10.
[3] Lagier J-C, 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.
[4] Raoult D, Henrissat B. Are stool samples suitable for studying the link between gut microbiota and obesity? Eur J Epidemiol 2014;29:307–9.
[5] Elsawi Z, Togo AH, Beye M, Dubourg G, Andrieu C, Armsrtong N, et al. *Hugonella massiliensis* gen. nov., sp. nov., genome sequence, and description of a new strictly anaerobic bacterium isolated from the human gut. MicrobiologyOpen 2017;6(4).
[6] Lagier J-C, Elkarkouri K, Rivet R, Couderc C, Raoult D, Fournier P-E. Non contiguous-finished genome sequence and description of *Senegalemassilia anaerobia* gen. nov., sp. nov. Stand Genomic Sci 2013;7:343–56.
[7] Fournier P-E, Drancourt M. New Microbes New Infections promotes modern prokaryotic taxonomy: a new section "TaxonoGenomics: new genomes of microorganisms in humans." New Microbe New Infect 2015;7:48–9.
[8] Morel A-S, Dubourg G, Prudent E, Edouard S, Gouriet F, Casalta J-P, et al. Complementarity between targeted real-time specific PCR and conventional broad-range 16S rDNA PCR in the syndrome-driven diagnosis of infectious diseases. Eur J Clin Microbiol Infect Dis 2015;34:561–70.
[9] Kim M, Oh H-S, Park S-C, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 2014;64:346–51.
[10] The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks [Internet]. [cited 2017 Jul 10]. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3965112/.
[11] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004;32:1792–7.
[12] Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 2009;26:1641–50.
[13] Bilen M, Cadoret F, Fournier P-E, Daoud Z, Raoult D. "*Libanicoccus massiliensis*" gen. nov., sp. nov., a new bacterium isolated from a stool sample from a pygmy woman. New Microbe New Infect 2016;15:40–1.
[14] Li X, Jensen RL, Højberg O, Canibe N, Jensen BB. *Olsenella scatoligenes* sp. nov., a 3-methylindole- (skatole) and 4-methylphenol- (p-cresol) producing bacterium isolated from pig faeces. Int J Syst Evol Microbiol 2015;65:1227–33.
[15] Dewhirst FE, Paster BJ, Tzellas N, Coleman B, Downes J, Spratt DA, et al. Characterization of novel human oral isolates and cloned 16S rDNA sequences that fall in the family *Coriobacteriaceae*: description of *Olsenella* gen. nov., reclassification of *Lactobacillus* uli as *Olsenella uli* comb. nov. and description of *Olsenella profusa* sp. nov. Int J Syst Evol Microbiol 2001;51:1797–804.

Article 3

Genomic and phenotypic description of a newly isolated human species *Collinsella bouchesdurhonensis* sp. nov.

Published in MicrobiologyOpen journal

Received: 8 September 2017 | Revised: 15 November 2017 | Accepted: 21 November 2017

DOI: 10.1002/mbo3.580

ORIGINAL RESEARCH

WILEY MicrobiologyOpen Open Access

Genomic and phenotypic description of the newly isolated human species *Collinsella bouchesdurhonensis* sp. nov.

Melhem Bilen[1,2] | Mamadou Beye[1] | Maxime Descartes Mbogning Fonkou[1] | Saber Khelaifia[1] | Frédéric Cadoret[1] | Nicholas Armstrong[1] | Thi Tien Nguyen[1] | Jérémy Delerce[1] | Ziad Daoud[2] | Didier Raoult[1,3] | Pierre-Edouard Fournier[1]

[1]Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, INSERM 1095, Assistance Publique-Hôpitaux de Marseille, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, Marseille, France

[2]Clinical Microbiology Department, Faculty of Medicine and Medical sciences, University of Balamand, Amioun, Lebanon

[3]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

Correspondence
Pierre-Edouard Fournier, Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, INSERM 1095, Assistance Publique-Hôpitaux de Marseille, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, Marseille, France.
Email: pierre-edouard.fournier@univ-amu.fr

Funding information
Mediterranee Infection Foundation; Agence Nationale de la Recherche; Investissements d'Avenir Méditerranée Infectio

Abstract

Using culturomics, a recently developed strategy based on diversified culture conditions for the isolation of previously uncultured bacteria, we isolated strain Marseille-P3296^T from a fecal sample of a healthy pygmy female. A multiphasic approach, taxono-genomics, was used to describe the major characteristics of this anaerobic and gram-positive bacillus that is unable to sporulate and is not motile. The genome of this bacterium is 1,878,572 bp-long with a 57.94 mol% G + C content. On the basis of these characteristics and after comparison with its closest phylogenetic neighbors, we are confident that strain Marseille-P3296^T (=CCUG 70328 = CSUR P3296) is the type strain of a novel species for which we propose the name *Collinsella bouchesdurhonensis* sp. nov.

KEYWORDS
Collinsella bouchesdurhonensis, culturomics, genome, pygmy, taxonomy

1 | INTRODUCTION

Over the past decade, metagenomics has been extensively adopted to enhance the description of the human gut microbiota (Gill et al., 2006; Ley, Turnbaugh, Klein, & Gordon, 2006; Ley et al., 2005). However, many detected DNA sequences cannot be assigned to as-yet cultured microorganisms, suggesting that a significant part of the human gut microbiota content remains to be isolated and described (Rinke et al., 2013). Thus, the ability to culture and obtain pure bacterial colonies is mandatory for a better description, analysis, and correlation with health and diseases (Lagier et al., 2015). Stool samples are recognized as a representative model of the gut microbiota (Raoult & Henrissat, 2014). To improve bacterial isolation from these specimens, a strategy named culturomics was recently developed (Lagier, Armougom, Million et al., 2012; Lagier, Armougom, Mishra, et al., 2012; Lagier, El Karkouri, et al., 2012 Mishra, Lagier, Robert, Raoult, & Fournier, 2012; Seng et al., 2009). This method relies on culturing samples using diversified combinations of culture media, temperatures, atmospheres, and incubation times. All isolated colonies from a specimen are identified using matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS). In case of identification's failure by MALDI-TOF MS, 16S rRNA amplification, and sequencing is systematically performed for further phylogenetic analysis with closely related bacterial species (Lagier, Armougom, Mishra, et al. 2012; Lagier,

El Karkouri,et al 2012; Mishra et al., 2012; Seng et al., 2009). Prior to the development of culturomics, only 688 bacteria were reported to be isolated from the human gut (Lagier et al., 2016). To date, culturomics has permitted the isolation of more than 1,000 distinct human gut bacterial species, including a significant number of novel species (Lagier et al., 2016). Using culturomics, we report the isolation of strain Marseille-P3296^T from a human stool sample, which we believe to be the representative strain of a new *Collinsella* species and propose the name *Collinsella bouchesdurhonensis* (*C. bouchesdurhonensis*) sp. nov. The *Collinsella* genus was first described by Kageyama, Benno, & Nakase, (1999), after the reclassification of *Eubacterium aerofaciens* into a new genus on the basis of a high 16S rRNA gene sequence divergence with other members of the *Eubacterium* genus. In addition to the type species *C. aerofaciens* (Kageyama et al., 1999), the *Collinsella* genus currently contains *C. intestinalis* (Kageyama & Benno, 2000),*C. massiliensis*(Padmanabhan et al., 2014), *C. stercoris* (Kageyama & Benno, 2000), and *C. tanakaei* (Nagai, Watanabe, & Morotomi, 2010). Members of the *Collinsella* genus are gram-positive anaerobic bacilli from the human gut microbiota that were suggested to play a role in health and diseases such as rheumatoid arthritis (Chen et al., 2016) and irritable bowel syndrome (Lee & Bak, 2011). They were also isolated from patients with Crohn's disease, ulcerative colitis, and colon cancer (Whitman et al., 2012).

Herein, we describe *C. bouchesdurhonensis* sp. nov. strain Marseille-P3296^T (=CCUG 70328 = CSUR P3296) using the taxono-genomics approach.

2 | MATERIAL AND METHODS

2.1 | Ethics and sample collection

The feces donor is a healthy 50-year-old pygmy female from Congo. She gave an informed and signed consent. Samples were stored at the URMITE laboratory (Marseille, France) at −80°C for further analysis. The culturomics study was approved by the ethics committee of the Institut Fédératif de Recherche 48 under number 09–022.

2.2 | Strain isolation

After being diluted with phosphate-buffered saline (Life Technologies, Carlsbad, CA), stool samples were incubated in an anaerobic blood culture bottle (BD BACTEC®, Plus Anaerobic/F Media, Le Pont de Claix, France) supplemented with 5% sheep blood and 5% sterile-filtered rumen at 37°C. Following initial growth in the blood culture vial, the bacterium was subcultured on 5% sheep blood–enriched Columbia agar (bioMérieux, Marcy l'Etoile, France).

2.3 | Identification of colonies identification

Colony identification was performed by the mean of MALDI-TOF MS as previously described (Elsawi et al., 2017; Lagier, El Karkouri, Nguyen, Armougom, Raoult, and Fournier, 2012). For strains that were not identified, 16S rRNA gene amplification and sequencing was performed as formerly done (Morel et al., 2015). Then, sequences were assembled and modified using the CodonCode Aligner software (http://www.codoncode.com) and a comparative analysis with the sequences of closely related species was performed using the BLAST software (http://blast.ncbi.nlm.nih.gov.gate1.inist.fr/Blast.cgi). A new species was considered when the similarity threshold between the 16S rRNA gene sequence of the understudied strain and its closest phylogenetic species with stranding in nomenclature was below 98.65% or considered as a new genus in case the similarity threshold was below 95% (Kim, Oh, Park, & Chun, 2014). The generated mass spectrum and 16S rRNA gene sequence were added to the UR-MS and the EMBL-EBI databases, respectively.

2.4 | Growth conditions

To identify the optimal growth conditions, strain Marseille-P3296^T was cultured under several temperature, atmosphere, pH, and salinity conditions. First, strain Marseille-P3296^T was cultured and incubated under aerobic, anaerobic (GENbag anaer, bioMérieux) and microaerophilic (GENbag Microaer, bioMérieux) conditions on 5% sheep blood-enriched Columbia agar (bioMérieux) at the following temperatures: 28, 37, 45, and 55°C. Furthermore, the halotolerance and acidity tolerance were estimated using 0, 5, 10, 50, 75, and 100 g/L NaCl concentrations and pH values of 6, 6.5, 7, and 8.5, respectively.

2.5 | Morphological and biochemical assays

The biochemical characteristics of strain Marseille-P3296^T were described using multiple API strips (ZYM, 20A and 50CH, bioMérieux), using bacteria that had been cultivated on 5% sheep blood-enriched Columbia agar, in anaerobic atmosphere at 37°C for 24 hr. In addition, the sporulation ability was tested by exposing a bacterial suspension to a thermic shock (80°C) for 20 min. Gram staining and cell motility were observed under a 100X magnification using a DM1000 photonic microscope (Leica Microsystems, Nanterre, France). Finally, cell morphology was observed by means of a Tecnai G20 Cryo (FEI) transmission electron microscope as previously described (Elsawi et al., 2017).

2.6 | FAME analysis

GC/MS was used for cellular fatty acid methyl ester (FAME) analysis after preparing two cell samples containing around 15 mg of bacterial biomass per tube and analysis was performed as previously described (Dione et al., 2016). Moreover, Elite 5-MS column was used for FAME separation and checked by mass spectrometry (Clarus 500-SQ 8 S, Perkin Elmer, Courtaboeuf, France). Then, the FAME mass spectral database (Wiley, Chichester, UK) and MS Search 2.0 operated with the Standard Reference Database 1A (NIST, Gaithersburg, USA) were used for FAME identification.

2.7 | Antibiotic resistance profile

The antibiotic resistance profile of the strain was evaluated using the *E*-test method and the following molecules: benzylpenecilin, amoxicillin, cefotaxime, ceftriaxone, imipenem, rifampicin, minocycline, tigecycline, amikacin, teicoplanin, vancomycin, colistin, daptomycin, and metronidazole (Biomerieux, France).

2.8 | DNA extraction and genome sequencing

Strain Marseille-P3296^T genomic DNA (gDNA) was extracted after subjecting it first to a mechanical treatment with the FastPrep BIO 101 instrument (Qbiogene, Strasbourg, France) using acid washed glass beads (Sigma) at a speed of 6.5 m/s for 90 s. Then, an incubation of 3 hr with lysozyme was done at 37°C for a DNA extraction assay using the EZ1 biorobot (Qiagen). The elution volume was set to 50 µl and the final gDNA concentration (25.8 ng/µl) was determined using a Qubit assay with the high sensitivity kit (Life technologies, Carlsbad, CA, USA).

The gDNA from strain Marseille-P3296^T was sequenced using a MiSeq sequencer (Illumina Inc, San Diego, CA, USA) and the Mate Pair strategy as formerly described (Elsawi et al., 2017). The size of the DNA fragments obtained ranged from 1.5 kb to up to 11 kb with an optimal size at 7.933 kb. Additionally, a circularization of 600 ng of tagmented DNA was done with no size selection. Mechanical shearing of the circularized DNA was performed using the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA) in order to obtain small DNA fragments with an optimal size of 981 bp.

High Sensitivity Bioanalyzer LabChip (Agilent Technologies) was used for libraries profile visualizations and a 19.28 nmol/L final concentration was measured. Normalization of the libraries was done at 2 nmol/L and pooled. Then, following denaturation, libraries were diluted to reach a concentration of 15 pmol/L. Libraries were loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Cluster generation was done automatically and a single 2 x 251-bp run was performed for sequencing. A total of 9.5 Gb information was acquired from a cluster density of 1,050 K/mm^2 with a quality threshold of 92.5% (18,644,000 passing filter paired reads). Strain Marseille-P3296^T index representation in this run was 7.78%. The 1,451,051 paired reads were trimmed and then assembled.

2.9 | Genome assembly

The genome assembly was performed using an in-house pipeline combining various softwares (Velvet (Zerbino & Birney, 2008), Spades (Bankevich et al., 2012), and Soap Denovo (Luo et al., 2012)), on trimmed (MiSeq and Trimmomatic softwares (Bolger, Lohse, & Usadel, 2014)) or untrimmed data (only MiSeq software). Gap reduction was performed using GapCloser tool for each assembly (Luo et al., 2012). Then, phage Phix was adapted for contamination detection and elimination. Finally, any scaffolds with a size lower than 800 bp or with a depth value less than 25% of the mean depth, were considered as potential contaminants and were eliminated. The best assembly was chosen based on multiple characteristics such as number of Ns, N50 and number of scaffolds. For the studied strain, the Velvet software gave the best assembly, with a mean depth coverage of 386.

2.10 | Genome annotation

The Prodigal software (Hyatt et al., 2010) was used with default parameters for Open Reading Frames (ORFs) detection and any ORFs that are spanning a sequencing gap were excluded.

The Clusters of Orthologous Groups (COG) database was used for bacterial protein sequences detection using BLASTP as previously described (Elsawi et al., 2017). Also, transfer RNA genes were searched using the tRNAScanSE (Lowe & Eddy, 1997) software and RNAmmer tool was used for ribosomal RNA genes detection (Lagesen et al., 2007). Moreover, Phobius (Käll, Krogh, & Sonnhammer, 2004) was used for lipoprotein signal peptides and transmembrane helices detection. ORFans were identified based on the BLASTP results as previously determined (Elsawi et al., 2017). Finally, annotation was performed with the DAGOBAH software (Gouret et al., 2011) that contains the Figenix Librairies (Gouret et al., 2005).

2.11 | 16S rRNA phylogenetic analysis

Sequences of the strains to be considered in the phylogenetic analyses were obtained after performing a BLASTn search against the 16S rRNA database of "The All-Species Living Tree" Project of Silva (Yilmaz et al., 2014). Sequences were aligned with CLUSTALW (Thompson, Higgins, Gibson, & Gibson, 1994) and MEGA software (Kumar, Tamura, & Nei, 1994) was used for phylogenetic inferences generation with the maximum-likelihood method.

2.12 | Genome comparison analysis

GenBank was used for retrieving the ORFeome, proteome, and complete sequences of the species being used for the comparative analyses. Nevertheless, the species were automatically recovered using Phylopattern (Gouret, Thompson, & Pontarotti, 2009) from the 16S rRNA tree. If no genome was available for a specific strain, a complete genome from another strain of the same species was used. If ORFeome and proteome were not predicted, Prodigal was used with default parameters to predict them from the genome sequence. ProteinOrtho was used for proteome analyses (Lechner et al., 2011). Then, AGIOS similarity scores were obtained with MAGI software for each couple of genomes (Ramasamy et al., 2014). These scores represent the mean value of nucleotide similarity between all couples of orthologous genes of the selected genomes. Proteome annotation was also performed for functional classes of predicted genes determination based on the clusters of orthologous groups of proteins (as was performed for genome annotation).

The comparison protocol was done with DAGOBAH (Gouret et al., 2011) which provided pipeline analysis, and using Phylopattern (Gouret et al., 2009) for tree manipulation.

3 | RESULTS AND DISCUSSION

3.1 | Strain identification

MALDI-TOF MS could not identify strain Marseille-P3296^T. Hence, its type spectrum (Figure 1a) was added to the URMS database and compared to spectra from other closely related species (Figure 1b). The 16S rRNA gene sequence exhibited a 96.19% sequence similarity with *Collinsella aerofaciens* strain JCM10188^T (GenBank accession number AB011816), the phylogenetically closest species with a validly published name (Figure 2). Strain Marseille-P3296^T exhibiting a 16S rRNA divergence greater than 1.35% with its closest phylogenetic neighbor (Kim et al., 2014), we investigated whether it possessed sufficient phenotypic and genomic differences with *C. aerofaciens* and other *Collinsella* species to be considered as a new species. The 16S rRNA sequence from strain Marseille-P3296^T was submitted to GenBank with the number LT623900.

3.2 | General biochemical and phenotypic characteristics of strain Marseille-P3296^T

Strain Marseille-P3296^T is a strictly anaerobic bacterium, able to grow at temperatures between 37°C and 42°C but optimally at 37°C). This strain can sustain a pH range of 6–8.5 and NaCl concentrations up to 5 g/L. After 24 hr of culture on agar, colonies are white and smooth with a diameter of 0.1–0.5 mm. Cells are rods, gram-negative with a mean width of 0.5 μm and a mean length of 2.6 μm (Table 1, Figures S1& S2), are not motile, are unable to sporulate and exhibit no catalase and oxidase activities.

Strain Marseille-P3296^T was susceptible to benzylpenecilin, amoxicillin, cefotaxime, ceftriaxone, imipenem, rifampicin, tigecycline, teicoplanin, vancomycin, and metronidazole (minimal inhibitory concentrations [μg/ml] of 0.006, 0.64, 0.125, 0.56, 0.12, 0.004, 0.19, 0.016, 1, and 0.064, respectively), but resistant to minocycline, amikacin, colistin, and daptomycin (4, >256, >256, and 2, respectively). Similar antibiotic susceptibility profiles were previously reported for *C. massiliensis* (Padmanabhan et al., 2014), *Collinsella aerofaciens*, *C. intestinalis*, and *C. tanakei* (Whitman et al., 2012)

Using API ZYM, strain Marseille-P3296^T exhibited alkaline phosphatase, valine arylamidase, esterase lipase (C8), α-fucosidase, esterase (C4), αchymotrypsin, leucine arylamidase, and acid phosphatase activities but was negative for β-glucosidase, lipase (C14), trypsin, naphtol-AS-BI-phosphohydrolase, N-acetyl-β-glucosaminidase, cystine arylamidase, α- mannosidase, β- galactosidase, α-galactosidase, β-glucuronidase, andα-glucosidase.

With the mean of API 50 CH, fermentation was observed in the presence of methyl-αD-glucopyranoside, D-galactose, potassium gluconate, D-mannose, D-arabitol, D-fructose, L-sorbose, gentiobiose, D-melezitose, D-glucose, D-lactose, D-trehalose, glycerol, D-raffinose, D-turanose, D-maltose, D-tagatose, D-xylose, and D-saccharose (sucrose). In contrast, no fermentation was obtained with erythritol, L-arabinose, D-ribose, D-adonitol, L-xylose, dulcitol, D-arabinose, inositol, L-rhamnose, methyl-βD-xylopyranoside, D-melibiose, inulin, methyl-D-mannopyranoside, glycogen, D-fucose, starch, xylitol, D-lyxose, potassium 5-ketogluconate, potassium 2-ketogluconate, L-fucose, and L-arabitol.

Moreover, when testing strain Marseille-P3296^T with API 20A, it could not form indole and was urease-negative. Additionally,

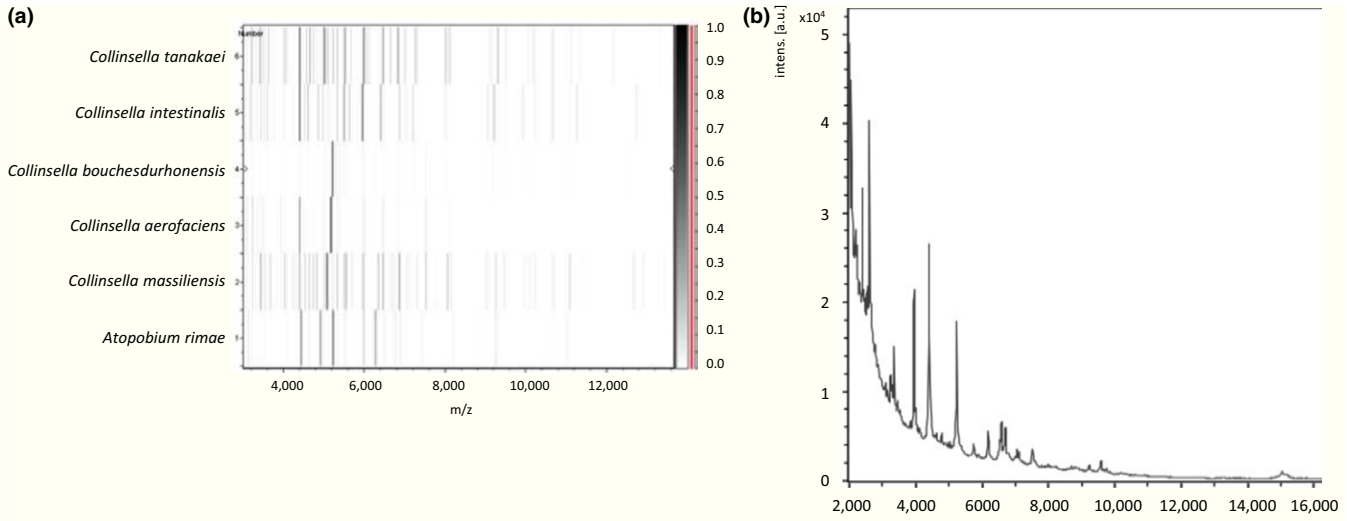


FIGURE 1 (a) Gel view comparing *Collinsella bouchedurhonensis* strain Marseille-P3296^T to other species. This figure shows in a pseudo-gel like format the raw spectra of the different strains shown on the left. The *x*-axis shows the m/z value. The left *y*-axis represents the running spectrum number obtained from subsequent spectra loading. The gray scale scheme code represents the intensity of the peaks and the *y*-axis shows the correlation between the peak color and its intensity. (b) Reference mass spectrum representing *Collinsella bouchesdurhonensis* strain Marseille-P3296^T acquired after analyzing 12 spectra

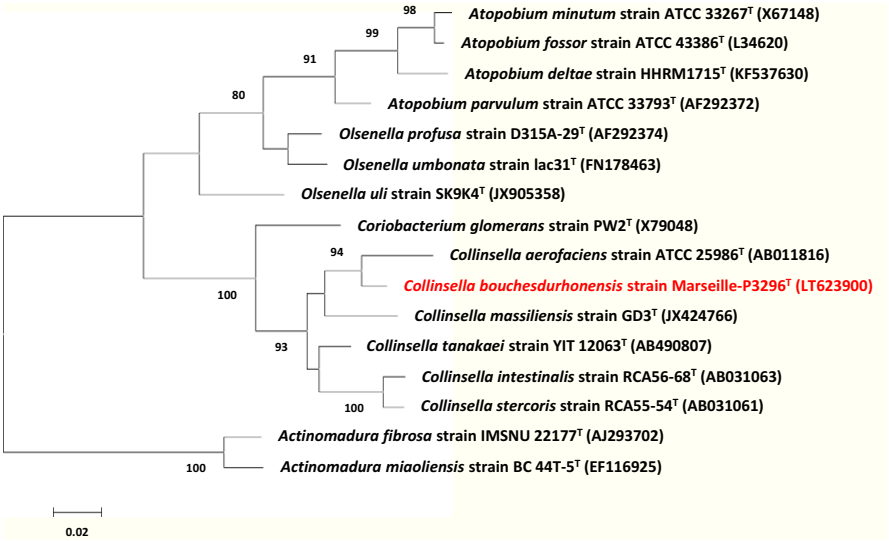


FIGURE 2 Phylogenetic tree representing the position of *Collinsella bouchesdurhonensis* strain Marseille-P3296^T relative to other closely related species. The 16S rRNA Database of "The All-Species Living Tree" Project of Silva (LTPs121) was used for sequence collection. The Muscle and FastTree softwares were used for sequence alignment and phylogenetic inferences with the maximum-likelihood method, respectively. Numbers at the nodes are percentages of bootstrap values obtained by repeating the analysis 1,000 times to generate a majority consensus tree. Only values >70% are displayed. *Actinomadura fibrosa* strain IMSNU 22177^T and *Actinomadura miaoliensis* strain BC 44T-5^T were used as outgroup

TABLE 1 Classification and general characteristics of *C. bouchesdurhonensis* strain Marseille-P3296^T

Property	Term	References
Current classification	Domain: *Bacteria*	(Woese, Kandler, & Wheelis, 1990)
	Phylum: *Actinobacteria*	(Cavalier-Smith, 2002)
	Class: *Actinobacteria*	(Stackebrandt, Rainey, & Ward-Rainey, 1997)
	Order: *Coriobacteriales*	(Gupta, Chen, Adeolu, & Chai, 2013; Stackebrandt et al., 1997)
	Family: *Coriobacteriaceae*	(Gupta et al., 2013)
	Genus: *Collinsella*	(Kageyama et al., 1999)
	Species: *Collinsella bouchesdurhonensis*	(Bilen, Cadoret, Daoud, Fournier, & Raoult, 2016)
	Type Strain: Marseille-P3296	
Gram stain	Negative	
Motility	Negative	
Cell shape	Bacillus	
Sporulation	Negative	
Optimum temperature	37°C	
Temperature range	37–42°C	

acidification of glucose, lactose, maltose, salicin, arabinose, xylose, cellobiose, mannitol, mannose, rhamnose, melezitose, saccharose, sorbitol, and trehalose was observed, in contrast to glycerol and raffinose. Cells were also gelatin-positive, but β-glucosidase-negative. A comparison of phenotypic and biochemical features between compared species is presented in Table 2.By comparison with other studied species, Strain Marseille-P3296^T differed in a combination of production of N-acetyl-β-glucosamine, absence of β-galactosidase activity, and acidification of D-mannitol. In contrast, all compared *Collinsella* species included Gram-positive bacilli that were unable to sporulate and did neither exhibit any urease activity nor could produce acid from L-arabinose. The G + C content of *C. bouchesdurhonensis*, *C. massiliensis*, *C. aerofaciens*, *C. intestinalis*, *and C*, *stercoris* is 57.94, 65.8, 61, 61.2, and 64.4 mol %, respectively (Table 2).

TABLE 2 Differential phenotypic characteristics of *C. bouchesdurhonensis* strain Marseille-P3296^T, *C. aerofaciens*, *C. massiliensis*, *C. intestinalis*, and *C. stercoris*

Properties	*Collinsella bouchesdurhonensis*	*Collinsella Massiliensis*	*Collinsella aerofaciens*	*Collinsella intestinalis*	*Collinsella stercoris*
Cell length (μm)	2,6	1,19	1.2–4.3	1.3–2.4	1.2–2.2
Oxygen requirement	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic
Gram stain	Positive	Positive	Positive	Positive	Positive
Motility	–	–	Na	–	–
Endospore formation	–	–	–	–	–
Indole	–	–	Na	Na	Na
Production of					
Alkaline phosphatase	+	+	–	+	+
Catalase	–	–	Na	Na	Na
Oxidase	–	–	Na	Na	Na
Urease	–	–	–	–	–
β-galactosidase	–	+	+	Na	Na
N-acetyl-β-glucosamine	+	–	Na	Na	Na
Acid from					
L-Arabinose	–	–	–	–	–
D-Ribose	–	–	–	–	+
D-Mannose	+	–	+	+	+
D-Mannitol	+	–	–	–	–
D-glucose	+	–	+	+	+
D-fructose	+	–	+	+	+
D-maltose	+	–	+	+	–
D-lactose	+	–	+	+	–
G + C content (mol%)	57,94	65,8	61	61,2	64,4
Habitat	Human gut	Human gut	Human gut	Human gut	Human gut

The major fatty acid of strain Marseille-P3296^T was the unsaturated 9-octadecanoic acid (C18:1n9, 35%). Several abundant saturated fatty acids were also identified, including hexadecanoic acid (C16:0, 30%), octadecanoic acid (C18:0, 7%), tetradecanoic acid (C14:0, 13%), and dodecanoic acid (C12:0, 6%) (Table S1). By comparison, *C. aerofaciens*, *C. intestinalis*, *and C. stercoris* possessed octadecanoic acid as major fatty acid (Nagai et al., 2010).

3.3 | Genome characteristics of strain marseille-P3296^T

The genome of strain Marseille P3296^T is 1,878,572-bp long with a 57.94 mol% G + C content. It is made of 15 scaffolds (for a total of 24 contigs). Of the 1,711 predicted genes, 51 are RNAs (1 16S rRNA, 1 5S rRNA, 48 tRNA genes, and 1 23S rRNA) and 1,660 are protein-coding genes. Fifty-two genes are detected as ORFans (3,17%) and 1,348 genes (82,15%) are assigned a putative function (by BLAST against nr or COGs). The 193 genes (11.76%) remaining are defined as hypothetical proteins (Table S2). A graphical representation of the genome from strain Marseille-P3296^T is presented in Figure S3. Table S4 details the distribution of genes into COG functional categories.

3.4 | Comparative analysis between the genomes of strain Marseille-P3296^T and closely related species

The draft genome sequence of strain Marseille-P3296^T was compared to those of *Collinsella aerofaciens* (*C. aerofaciens*) (AAVN00000000), *C. tanakaei* (ADLS00000000), *C. stercoris* (ABXJ00000000), *C. intestinalis* (ABXH00000000), *C. massiliensis* (NZ_CAPI01000000), *Olsenella uli* (*O. uli*, NR_075775.1), *Atopobium parvulum* (*A. parvulum*, NC_013203.1), *A. rimae* (ACFE00000000), *A. minutum* (AGXC00000000), and *Coriobacterium glomerans* (*C. glomerans*, NC_015389.1).

The draft genome of strain Marseille-P3296^T (1,88 Mb) is smaller than those of *C. aerofaciens*, *C. tanakaei*, *C. stercoris*, *O. uli*, *C. massiliensis*, and *C, glomerans* (2,44, 2,48, 2,40, 2,05, 2,33, and 2,12 Mb, respectively), but larger than those of *C. intestinalis*, *A. minutum*, *A. rimae*, and *A. parvulum* (1,8, 1,71, 1,63, and 1,54 MB, respectively). Also, the G+C content of strain Marseille-P3296^T (57.9), is smaller than those of

C. massiliensis, *C. aerofaciens*, *C. intestinalis*, *C. tanakaei*, *C. stercoris*, *O. uli*, and *C. glomerans* (65.8, 60.5, 62.5, 60.2, 63.2, 64.7, and 60.4%, respectively), but larger than those of *A. minutum*, *A. rimae*, and *A. parvulum* (48.9, 49.3, and 45.7%, respectively) (Table S3).

As for the gene content of strain Marseille-P3296^T (1,711), it is smaller than those of *C. massiliensis*, *C. aerofaciens*, *C. intestinalis*, *C. tanakaei*, *C. stercoris*, *O. uli*, *and C. glomerans* (2,046, 2,222, 1,626, 2,258, 2,106, 1,827, and 1,859, respectively), but larger than those of *A. minutum*, *A. rimae*, and *A. parvulum* (1,593, 1,523, and 1,411, respectively). The distribution of functional classes of predicted genes of strain Marseille-P3296^T according to the COGs database is presented in Figure S4. The distribution was similar in all studied genomes.

Strain Marseille-P3296^T shared the highest number of orthologous proteins with *C. tanakaei* (988). It also shared 971, 964, 890, and 860 orthologous proteins with *C. intestinalis*, *C. aerofaciens*, *C. massiliensis*, and *C. stercoris*, respectively, and lower numbers of orthologous proteins with species that belonged to other compared genera (Table S5). Moreover, in an effort to determine the degree of genetic relatedness of strain Marseille-P3296^T with, and among *Collinsella* species, we used two parameters, AGIOS and dDDH. Within the *Collinsella* genus, excluding strain Marseille-P3296^T, AGIOS values ranged from 72.77% between *C. aerofaciens* and *C. tanakaei* to 81.12% between *C. stercoris* and *C. intestinalis* (Table S5). When compared to other *Collinsella* species, strain Marseille-P3296^T exhibited AGIOS values ranging from 71.57% with *C. massiliensis* to 74.07% with *C. aerofaciens* (Table S5). In addition, within the *Collinsella* genus, excluding strain Marseille-P3296^T, dDDH values ranged from 22.1% between *C. aerofaciens* and *C. massiliensis* to 24.8% between *C. stercoris* and *C. tanakaei* (Table S6). When compared with other *Collinsella* species, strain Marseille-P3296^T exhibited dDDH values ranging from 21.9% with *C. massiliensis* to 25% with *C. aerofaciens* (Table S6). For both parameters, strain Marseille-P3296^T exhibited genomic similarity values in the range of those observed between other *Colinsella* species. In addition, the dDDH values lower than the 70% threshold that delimits bacterial species are consistent with the 16S rRNA sequence comparison and support the classification of strain Marseille-P3296^T within a new species (Tindall, Rosselló-Móra, Busse, Ludwig, & Kämpfer, 2010; Wayne, 1988).

4 | CONCLUSION

On the basis of genomic, phenotypic, and biochemical features, we suggest the creation of a new species, *Collinsella bouchesdurhonensis* sp. nov. Strain Marseille-P3296^T is the type strain of *C. bouchesdurhonensis* sp. nov.

4.1 | Description of *Collinsella bouchesdurhonensis* sp. nov

Collinsella bouchesdurhonensis (bou.che.du.rho.nen'sis. N.L. fem. adj. *bouchesdurhonensis*, referring to Bouches-du-Rhône, the name of the department around Marseille, France, where this study was performed and strain Marseille-P3296^T was isolated).

Strain Marseille-P3296^T is a strictly anaerobic bacterium, able to grow at temperatures between 37°C and 42°C but optimally at 37°C). This strain can sustain a pH range of 6–8.5 and up to 5% NaCl concentration.

The species exhibits alkaline phosphatase, valine arylamidase, esterase lipase (C8), α-fucosidase, esterase (C4), αchymotrypsin, leucine arylamidase, and acid phosphatase activities. Fermentation is observed in the presence of methyl- αD-glucopyranoside, D-galactose, potassium gluconate, D-mannose, D-arabitol, D-fructose, L-sorbose, gentiobiose, D-melezitose, D-glucose, D-lactose, D-trehalose, glycerol, D-raffinose, D-turanose, D-maltose, D-tagatose, D-xylose, and D-saccharose (sucrose). Strain Marseille-P3296^T cannot form indole and is urease-negative. Additionally, acidification of glucose, lactose, maltose, salicin, arabinose, xylose, cellobiose, mannitol, mannose, rhamnose, melezitose, saccharose, sorbitol, and trehalose is observed. The major fatty acid is 9-octadecenoic acid.

The genome size 1,878,572-bp long with a 57.94 mol% G + C content. The 16S rRNA and genome sequences of *C. bouchesdurhonensis* sp. nov. are deposited in EMBL-EBI under accession numbers LT623900 and FTLD00000000, respectively. The type strain is Marseille-P3296^T (=CSURP3296 = CCUG70328) and was isolated from the stool sample of a healthy 50-year-old pygmy woman from Congo.

ACKNOWLEDGMENTS

The authors acknowledge the Xegen Company (http://www.xegen.fr/) for automating the genomic annotation process, Claudia Andrieu for administrative assistance and Magdalen Lardière for literature assistance. This study was funded by the Mediterranee Infection Foundation and the French Agence National de la Recherche under reference Investissements d'Avenir Méditerranée Infection 10-IAHU-03.

CONFLICTS OF INTEREST

No conflicts of interest are to be declared.

ORCID

Melhem Bilen http://orcid.org/0000-0002-0178-5124

Pierre-Edouard Fournier http://orcid.org/0000-0001-8463-8885

REFERENCES

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computer Biology*, *19*, 455–477. https://doi.org/10.1089/cmb.2012.0021

Bilen, M., Cadoret, F., Daoud, Z., Fournier, P.-E., & Raoult, D. (2016). "*Collinsella bouchesdurhonensis*" sp. nov., identified in human stool sample. *New Microbes and New Infections*, *17*, 39–40. https://doi.org/10.1016/j.nmni.2016.11.004

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Cavalier-Smith, T. (2002). The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. *International Journal of Systematic and Evolutionary Microbiology*, *52*, 7–76. https://doi.org/10.1099/00207713-52-1-7

Chen, J., Wright, K., Davis, J. M., Jeraldo, P., Marietta, E. V., Murray, J., … Taneja, V. (2016). An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. *Genome Medicine*, *8*, 43. https://doi.org/10.1186/s13073-016-0299-7

Dione, N., Sankar, S. A., Lagier, J.-C., Khelaifia, S., Michele, C., Armstrong, N., … Fournier, P.-E. (2016). Genome sequence and description of *Anaerosalibacter massiliensis* sp. nov. *New Microbes and New Infections*, *10*, 66–76. https://doi.org/10.1016/j.nmni.2016.01.002

Elsawi, Z., Togo, A.H., Beye, M., Dubourg, G., Andrieu, C., Armsrtong, N., … Khelaifia, S. (2017) *Hugonella massiliensis* gen. nov., sp. nov., genome sequence, and description of a new strictly anaerobic bacterium isolated from the human gut. *Microbiologyopen*. *6*, 4.https://doi.org/10.1002/mbo3.458

Gill, S. R., Pop, M., Deboy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., … Nelson, K. E. (2006). Metagenomic analysis of the human distal gut microbiome. *Science*, *312*, 1355–1359. https://doi.org/10.1126/science.1124234

Gouret, P., Paganini, J., Dainat, J., Louati, D., Darbo, E., Pontarotti, P., & Levasseur, A. (2011). Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: The multi-agent software system DAGOBAH. In. P. Pontarotti (Ed.). *Evolutionary Biology – Concepts, biodiversity, macroevolution and genome evolution* (pp. 71–87). Berlin and Heidelberg: Springer. https://doi.org/10.1007/978-3-642-20763-1_5

Gouret, P., Thompson, J. D., & Pontarotti, P. (2009). PhyloPattern: Regular expressions to identify complex patterns in phylogenetic trees. *BMC Bioinformatics*, *10*, 298. https://doi.org/10.1186/1471-2105-10-298

Gouret, P., Vitiello, V., Balandraud, N., Gilles, A., Pontarotti, P., & Danchin, E. G. J. (2005). FIGENIX: Intelligent automation of genomic annotation: Expertise integration in a new software platform. *BMC Bioinformatics*, *6*, 198. https://doi.org/10.1186/1471-2105-6-198

Gupta, R. S., Chen, W. J., Adeolu, M., & Chai, Y. (2013). Molecular signatures for the class *Coriobacteriia* and its different clades; proposal for division of the class *Coriobacteriia* into the emended order *Coriobacteriales*, containing the emended family *Coriobacteriaceae* and *Atopobiaceae* fam. nov., and *Eggerthellales* ord. nov., containing the family *Eggerthellaceae* fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, *63*, 3379–3397. https://doi.org/10.1099/ijs.0.048371-0

Hyatt, D., Chen, G.-L., Locascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*, 119. https://doi.org/10.1186/1471-2105-11-119

Kageyama, A., & Benno, Y. (2000). Emendation of genus *Collinsella* and proposal of *Collinsella stercoris* sp. nov. and *Collinsella intestinalis* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, *50*, 1767–1774. https://doi.org/10.1099/00207713-50-5-1767

Kageyama, A., Benno, Y., & Nakase, T. (1999). Phylogenetic and phenotypic evidence for the transfer of *Eubacterium aerofaciens* to the genus *Collinsella* as *Collinsella aerofaciens* gen. nov., comb. nov. *International Journal of Systematic Bacteriology.*, *49*, 557–565. https://doi.org/10.1099/00207713-49-2-557

Käll, L., Krogh, A., & Sonnhammer, E. L. L. (2004). A combined transmembrane topology and signal peptide prediction method. *Journal of Molecular Biology*, *338*, 1027–1036. https://doi.org/10.1016/j.jmb.2004.03.016

Kim, M., Oh, H.-S., Park, S.-C., & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 346–351. https://doi.org/10.1099/ijs.0.059774-0

Kumar, S., Tamura, K., & Nei, M. (1994). MEGA: Molecular evolutionary genetics analysis software for microcomputers. *Computer and Applied Biosciences*, *10*, 189–191.

Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H.-H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, *35*, 3100–3108. https://doi.org/10.1093/nar/gkm160

Lagier, J.-C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., … Raoult, D. (2012). Microbial culturomics: Paradigm shift in the human gut microbiome study. *Clinical Microbiology and Infection*, *18*, 1185–1193. https://doi.org/10.1111/1469-0691.12023

Lagier, J.-C., Armougom, F., Mishra, A. K., Nguyen, T.-T., Raoult, D., & Fournier, P.-E. (2012). Non-contiguous finished genome sequence and description of *Alistipes timonensis* sp. nov. *Standards in Genomic Sciences*, *6*, 315–324. https://doi.org/10.4056/sigs.2685971

Lagier, J.-C., Edouard, S., Pagnier, I., Mediannikov, O., Drancourt, M., & Raoult, D. (2015). Current and past strategies for bacterial culture in clinical microbiology. *Clinical Microbiology Reviews*, *28*, 208–236. https://doi.org/10.1128/CMR.00110-14

Lagier, J.-C., El Karkouri, K., Nguyen, T.-T., Armougom, F., Raoult, D., & Fournier, P.-E. (2012). Non-contiguous finished genome sequence and description of *Anaerococcus senegalensis* sp. nov. *Standards in Genomic Sciences*, *6*, 116–125. https://doi.org/10.4056/sigs.2415480

Lagier, J. C., Khelaifia, S., Alou, M. T., Ndongo, S., Dione, N., Hugon, P., … Raoult, D. (2016). Culture of previously uncultured members of the human gut microbiota by culturomics. *Nature Microbiology*, *1*, 16203. https://doi.org/10.1038/nmicrobiol.2016.203

Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., & Prohaska, S. J. (2011). Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics*, *12*, 124. https://doi.org/10.1186/1471-2105-12-124

Lee, B. J., & Bak, Y.-T. (2011). Irritable bowel syndrome, gut microbiota and probiotics. *Neurogastroenterology and Motility*, *17*, 252–266. https://doi.org/10.5056/jnm.2011.17.3.252

Ley, R. E., Bäckhed, F., Turnbaugh, P., Lozupone, C. A., Knight, R. D., & Gordon, J. I. (2005). Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences USA*, *102*, 11070–11075. https://doi.org/10.1073/pnas.0504978102

Ley, R. E., Turnbaugh, P. J., Klein, S., & Gordon, J. I. (2006). Microbial ecology: Human gut microbes associated with obesity. *Nature*, *444*, 1022–1023. https://doi.org/10.1038/4441022a

Lowe, T. M., & Eddy, S. R. (1997). tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, *25*, 955–964.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., … Wang, J. (2012). SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *GigaScience*, *1*, 18. https://doi.org/10.1186/2047-217X-1-18

Mishra, A. K., Lagier, J.-C., Robert, C., Raoult, D., & Fournier, P.-E. (2012). Non contiguous-finished genome sequence and description of *Peptoniphilus timonensis* sp. nov. *Standards in Genomic Sciences*, *7*, 1–11. https://doi.org/10.4056/sigs.2956294

Morel, A.-S., Dubourg, G., Prudent, E., Edouard, S., Gouriet, F., Casalta, J.-P., … Raoult, D. (2015). Complementarity between targeted real-time specific PCR and conventional broad-range 16S rDNA PCR in the syndrome-driven diagnosis of infectious diseases. *European Journal of Clinical Microbiology and Infectious Diseases*, *34*, 561–570. https://doi.org/10.1007/s10096-014-2263-z

Nagai, F., Watanabe, Y., & Morotomi, M. (2010). *Slackia piriformis* sp. nov. and *Collinsella tanakaei* sp. nov., new members of the family *Coriobacteriaceae*, isolated from human faeces. *International Journal of Systematic and Evolutionary Microbiology*, *60*, 2639–2646. https://doi.org/10.1099/ijs.0.017533-0

Padmanabhan, R., Dubourg, G., Nguyen, T.-T., Couderc, C., Rossi-Tamisier, M., Caputo, A., … Fournier, P.-E. (2014). Non-contiguous finished genome sequence and description of *Collinsella massiliensis* sp. nov. *Standards in Genomic Sciences*, *9*, 1144–1158. https://doi.org/10.4056/sigs.5399696

Ramasamy, D., Mishra, A. K., Lagier, J.-C., Padhmanabhan, R., Rossi, M., Sentausa, E., … Fournier, P.-E. (2014). A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 384–391. https://doi.org/10.1099/ijs.0.057091-0

Raoult, D., & Henrissat, B. (2014). Are stool samples suitable for studying the link between gut microbiota and obesity? *European Journal of Epidemiology*, *29*, 307–309. https://doi.org/10.1007/s10654-014-9905-4

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J.-F., … Woyke, T. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, *499*, 431–437. https://doi.org/10.1038/nature12352

Seng, P., Drancourt, M., Gouriet, F., La Scola, B., Fournier, P.-E., Rolain, J. M., & Raoult, D. (2009). Ongoing revolution in bacteriology: Routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clinical Infectious Diseases*, *49*, 543–551. https://doi.org/10.1086/600885

Stackebrandt, E., Rainey, F. A., & Ward-Rainey, N. I. (1997). Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. *International Journal of Systematic Bacteriology*, *47*, 479–491.

Thompson, J. D., Higgins, D. G., Gibson, T. J., Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*, 4673–4680.

Tindall, B. J., Rosselló-Móra, R., Busse, H.-J., Ludwig, W., & Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *International Journal of Systematic and Evolutionary Microbiology*, *60*, 249–266. https://doi.org/10.1099/ijs.0.016949-0

Wayne, L. G. (1988). International committee on systematic bacteriology: Announcement of the report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Zentralblatt Bakteriologie, Mikrobiologie und Hygiene*, *268*, 433–434.

Whitman, W. B., Goodfellow, M., Kämpfer, P., Busse, H.-J., Trujillo, M.E., Ludwig, W., & Suzuki, K. (2012). *Bergey's manual of systematic bacteriology: Volume 5: The Actinobacteria*. New York, USA: Springer.

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences USA*, *87*, 4576–4579.

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., … Glöckner, F. O. (2014). The SILVA and all-species living tree project (LTP) taxonomic frameworks. *Nucleic Acids Research*, *42*, D643–D648.https://www.ncbi.nlm.nih.gov/pm.

Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, *18*, 821–829. https://doi.org/10.1101/gr.074492.107

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Bilen M, Beye M, Mbogning Fonkou MD, et al. Genomic and phenotypic description of the newly isolated human species *Collinsella bouchesdurhonensis* sp. nov.. *MicrobiologyOpen*. 2018;e580. https://doi.org/10.1002/mbo3.580

Article 4

Eggerthella timonensis sp. nov, a new species isolated from the stool sample of a pygmy female.

Published in MicrobiologyOpen journal

Received: 7 August 2017 | Revised: 15 November 2017 | Accepted: 21 November 2017

DOI: 10.1002/mbo3.575

ORIGINAL RESEARCH

WILEY MicrobiologyOpen Open Access

Eggerthella timonensis sp. nov, a new species isolated from the stool sample of a pygmy female

Melhem Bilen[1,2] | Maxime Descartes Mbogning Fonkou[1] | Enora Tomei[1] | Nicholas Armstrong[1] | Fadi Bittar[1] | Jean-Christophe Lagier[1] | Ziad Daoud[2] | Pierre-Edouard Fournier[1] | Didier Raoult[1,3] | Frédéric Cadoret[1]

[1]URMITE UM63, CNRS 7278, IRD 198, INSERM 1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, Aix-Marseille Université, Marseille Cedex, France

[2]Clinical Microbiology Department, Faculty of Medicine and Medical sciences, University of Balamand, Amioun, Lebanon

[3]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

Correspondence
Frédéric Cadoret, URMITE UM63, CNRS 7278, IRD 198, INSERM 1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, Aix-Marseille Université, Marseille Cedex, France.
Email: frederic.n.cadoret@gmail.com

Funding information
Fondation Méditerranée Infection

Abstract

Eggerthella timonensis strain Marseille-P3135 is a new bacterial species, isolated from the stool sample of a healthy 8-year-old pygmy female. This strain (LT598568) showed a 16S rRNA sequence similarity of 96.95% with its phylogenetically closest species with standing in nomenclature *Eggerthella lenta* strain DSM 2243 (AF292375). This bacterium is a nonspore forming, Gram-positive, nonmotile rod with catalase but no oxidase activity. Its genome is 3,916,897 bp long with 65.17 mol% of G + C content. Of the 3,371 predicted genes, 57 were RNAs and 3,314 were protein-coding genes. Here, we report the main phenotypic, biochemical, and genotypic characteristics of *E. timonensis* strain Marseille-P3135 (=CSUR P3135, =CCUG 70327); ti.mo.nen′sis, N.L. masc. adj., with *timonensis* referring to La Timone, which is the name of the hospital in Marseille (France) where this work was performed). Strain is a nonmotile Gram-positive rod, unable to sporulate, oxidase negative, and catalase positive. It grows under anaerobic conditions between 25°C and 42°C but optimally at 37°C.

KEYWORDS
culturomics, *Eggerthella timonensis*, genome, new species, pygmy, taxonogenomics

1 | INTRODUCTION

The human gut microbiota has drawn more attention to with the advancement and development of new sequencing techniques (Gill et al., 2006; Ley, Turnbaugh, Klein, & Gordon, 2006; Ley et al., 2005). Yet, we face several limitations when using these techniques, especially when it comes to depth bias, incomplete database, and the obtention of raw material for further analysis (Greub, 2012). However, the ability to cultivate and isolate pure colonies is mandatory to describe the human gut microbiota, thus the need to develop a technique that enhances the efficiency of these two factors (Lagier et al., 2015). When talking about the human gut, stool samples are the best representatives of its microbiome since only 1 g of human stool sample might contain up to 10^{11}–10^{12} bacteria (Raoult & Henrissat, 2014). Before the introduction of culturomics, only 688 bacteria and 2 archaea had been recognized in the human gut (Lagier et al., 2016). Culturomics was developed with the purpose of optimizing growth conditions of previously uncultured bacteria in order to fill the missing gaps in the human microbiome (Lagier et al., 2012a). In general, culturomics consists in culturing samples by using 18 different conditions along with isolating pure colonies for further identifications using the matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (matrix-assisted laser desorption/ionization time-of-flight mass spectrometry [MALDI-TOF MS]) approach and 16S rRNA gene sequencing. Any unidentified colonies are subject to 16S rRNA gene sequencing and a series of descriptive experiments targeting the phenotypic, biochemical,

and genomic characteristics at the same time (Lagier et al., 2012b; Seng et al., 2009). Using this methodology, we were able to isolate a new strain *Eggerthella timonensis*, a member of the genus *Eggerthella* (Bilen, Cadoret, Daoud, Fournier, & Raoult, 2016). *Eggerthella lenta*, formerly known as *Eubacterium lentum*, is the type strain of *Eggerthella* genus and was first reported in 1935 by Arnold Eggerth (Eggerth, 1935; Kageyama, Benno, & Nakase, 1999; Moore, Cato, & Holdeman, 1971). Species belongs to *Eggerthella* genus, *Actinobacteria* phylum in the *Coriobacteriaceae* family and known for its ability to grow under anaerobic conditions (Kageyama et al., 1999). Moreover, *Eggerthella* species have been reported to colonize the human gut microbiome and have been correlated to several health problems such as anal abscess and ulcerative colitis (Lau et al., 2004a).

In this study, we describe *E. timonensis* strain Marseille-P3135 (=CCUG 70327 [Culture Collection University Gothenburg], =CSUR P3135 [Collection de Souches de l'Unité des Rickettsies]) using a polyphasic approach by targeting multiple phenotypic, biochemical, and genotypic aspects.

2 | MATERIAL AND METHODS

2.1 | Strain isolation

Before stool sample collection in Congo in 2015, an approval was obtained from the ethic committee (09-022) of the Institut Hospitalo-Universitaire Méditerranée Infection (Marseille, France). The stool sample was collected from a healthy 8-year-old pygmy female accordingly to Nagoya protocol. Stool samples were shipped from Congo to France in the specific protecting medium C-Top Ae-Ana (Culture Top, Marseille, France) and stored at −80°C for further study and analysis.

Samples were inoculated in a blood culture bottle (BD BACTEC®, Plus Anaerobic/F Media, Le Pont de Claix, France) supplemented with 5% of rumen and 5% of sheep blood at 37°C. Bacterial growth and isolation was assessed during 30 days on 5% sheep blood–enriched Columbia agar solid medium (bioMérieux, Marcy l'Etoile, France). MALDI-TOF MS was used for colonies identification. When the latter fails to identify tested colonies, 16S rRNA gene sequencing was used (Lagier et al., 2012b; Seng et al., 2009). On average, 10,000 colonies have been tested for each stool sample.

2.2 | MALDI-TOF MS and 16S rRNA gene sequencing

Using a MSP 96 MALDI-TOF target plate, bacterial colonies were spotted and identified by the means of MALDI-TOF MS using a Microflex LT spectrometer as previously described (Seng et al., 2009). In case of MALDI-TOF's identification failure due to lack of a reference strain in the database, 16S rRNA sequencing was used for further analysis using the GeneAmp PCR System 2720 thermal cyclers (Applied Biosystems, Foster City, CA, USA) and the ABI Prism 3130xl Genetic Analyzer capillary sequencer (Applied Biosystems) (Morel et al., 2015). Sequences were assembled and modified using CodonCode Aligner software (http://www.codoncode.com) and finally blasted against the online database of National Center for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov.gate1.inist.fr/Blast.cgi). Once blasted, a sequence similarity of less than 98.65% with the closest species was used to define a new species and 95% for defining a new genus (Kim, Oh, Park, & Chun, 2014). Subsequently, the mass spectrum of the new species was added to the URMITE [Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes] database (http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database) and its 16S rRNA gene sequence was submitted to EMBL-EBI with an accession number of LT598568.

2.3 | Phylogenetic analysis

16S rRNA sequences of strain's closest species were obtained from the database of "The All-Species Living Tree" Project of Silva (LTPs121) ("The SILVA and 'All-species Living Tree Project (LTP)' taxonomic frameworks," n.d.), aligned with Muscle software and phylogenetic inferences were done using FastTree with the approximately maximum-likelihood method (Price, Dehal, & Arkin, 2009). Moreover, Shimodaira–Hasegawa test was adapted in order to compute the support local values shown on the nodes. Bad taxonomic reference strains were removed along with duplicates using phylopattern (Gouret, Thompson, & Pontarotti, 2009). This pipeline was done using the DAGOBAH software (Gouret et al., 2011), which comprises Figenix (Gouret et al., 2005) libraries.

2.4 | Growth conditions

In order to obtain the optimal growth conditions, the strain was cultured under several conditions in terms of temperature, atmosphere, pH, and salinity. First, the strain was cultured and incubated under aerobic, anaerobic, and micro-aerophilic conditions on 5% sheep blood–enriched Colombia agar (bioMérieux) at the following temperatures: 28°C, 37°C, 45°C, and 55°C. Bacterial growth under anaerobic and microaerophilic environment was tested using the GENbag anaer and GENbag microaer systems (Thermofisher Scientific, Basingstoke), respectively. Furthermore, salinity tolerance was tested by assessing growth at 37°C under anaerobic condition using different NaCl concentrations (0, 5, 10, 50, 75, and 100 g/L NaCl). As well, optimal pH for growth was evaluated by testing multiple pH: 6, 6,5, 7, and 8.5

2.5 | Morphological and biochemical assays

In order to biochemically describe strain Marseille-P3135; different API tests (ZYM, 20A and 50CH, bioMérieux) were used. Sporulation ability of this bacterium was tested by exposing a bacterial suspension for 10 min to a thermal shock at 80°C, and then cultured on COS media. Moreover, the motility of the strain was detected using a DM1000 photonic microscope (Leica Microsystems, Nanterre, France) under a 100× objective lens. Also, a bacterial suspension was fixed with a solution of 2.5% glutaraldehyde in 0.1 mol/L cacodylate buffer for more than 1 hr at 4°C for observation under the Morgagni

268D (Philips) transmission electron microscope. Finally, Gram staining results and images were obtained by DM1000 photonic microscope (Leica Microsystems) using a 100× oil-immersion objective lens.

2.6 | Fatty acid methyl ester (FAME) composition of strain Marseille-P3135

Using gas chromatography/mass spectrometry (GC/MS), Cellular FAME analysis was performed. Harvested from several culture plates, two samples were made with <1 mg of bacterial biomass per tube, and then FAME and GC/MS were done as previously described (Dione et al., 2016).

2.7 | Antibiotic susceptibility

An antibiotic resistance profile was developed by the use of the E test method. The following antibiotics strips were used: vancomycin, rifampicin, benzylpenicillin, amoxycillin, imipenem, tigecycline, amikacin, erythromycin, minocycline, teicoplanin, colistin, daptomycin, metronidazole, and ceftriaxone.

2.8 | DNA extraction and genome sequencing

To extract the genomic DNA (gDNA) of strain Marseille-P3135, FastPrep BIO 101 (Qbiogene, Strasbourg, France) was used for a mechanical treatment with acid-washed beads (G4649-500 g Sigma). Then, samples were incubated with lysozyme after 2 hr and a half at 37°C and EZ1 biorobot (Qiagen) was used for DNA extraction according the to manufacture guidelines. Qubit was used for DNA quantification (69.3 ng/µl).

As for genome sequencing, MiSeq Technology (Illumina Inc, San Diego, CA, USA) was used with mate-pair and paired-end methods. Also, Nextera XT kit (Illumina) and Nextera Mater pair kit (Illumina) were used for samples barcoding. The DNA of the strain was mixed with 11 paired-end projects and 11 mate-pair projects. Pair-end libraries were prepared by using 1 ng of gDNA, which was fragmented and tagged. Twelve PCR amplification cycles accomplished the tag adapters and added dual-index barcodes. Subsequently, purification was done using AMPure XP bead (Beckman Coulter Inc, Fullerton, CA, USA), and libraries' normalization was done as described in Nextera XT protocol (Illumina) for pooling and sequencing on MiSeq. A single run of 39 hr in 2 × 250-bp was done for paired-end sequencing and clusters generation. This library was loaded on two flowcells. Total information of 6.5 and 4.3 Gb was obtained from a 685 and 446 k/mm^2 cluster density with a cluster quality 95.1% and 94.8% (12,615,000 and 8,234,000 passed filtered clusters). Index representation for strain Marseille-P3135 was determined to be of 4.57% and 3.83%. The 576,647 and 315,481 paired-end reads were filtered based on the quality of the reads.

As for mate-pair libraries preparation, 1.5 µg of gDNA were used with Nextera mate-pair Illumina protocol. Mate-pair junction adapter were used to tag fragmented gDNA, and Agilent 2100 BioAnalyzer (Agilent Technologies Inc, Santa Clara, CA, USA) was used to validate the profile of the fragmentation on a DNA 7500 labchip. The average size of the fragmented DNA ranged between 1 kb up to 11 kb with an ideal size at 9.1 kb. Six hundred nanograms of tagmented DNA (with no size selection) was circularized and sheared mechanically on Covaris device S2 in microtubes (Covaris, Woburn, MA, USA) to reach a size of 956 bp. Then, libraries concentration was set to 8.46 nmol/L, and High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc, Santa Clara, CA, USA) was used for libraries profile visualization. This library was loaded on two flowcells. Before pooling the libraries, it was normalized at 2 nmol/L. Samples were denatured and diluted at 15 pmol/L before loading it on a the reagent cartridge and then onto the instrument along with the flow cell. A single run of 39 hr was done for cluster generation and sequencing in a 2 × 151-bp. 5.1 and 2.8 Gb 2.8 were obtained from 542 and 274 K/mm^2 cluster density with a cluster threshold of 95.7% and 97.6% (10,171,000 and 5,537,000 passing filtered paired reads). Index representation of the studied strain was determined to be of 8.53% and 9.24%. The 867,401 and 511,563 paired reads were trimmed and then assembled.

Genome assembly, annotation, and comparison were made with the same pipeline as previously discussed in our previous work (Elsawi et al., 2017).

3 | RESULTS AND DISCUSSION

3.1 | Strain Marseille-P3135 identification

After comparing the 16S rRNA gene sequence of the present strain with other organisms, it was found that it exhibited a sequence similarity of 96.95% with *E. lenta* (DSM 2243; AF292375), its phylogenetically closest species with standing in nomenclature (Figure 1). The phylogenetic analysis clearly supports that the studied strain is a member of the *Eggerthella* genus. Having more than 1.3% sequence divergence with its closest species, we can suggest that the isolate represents a new species named *E. timonensis* (Bilen et al., 2016). Typical spectrum obtained by MALDI-TOF MS of the strain can be found in supplementary Figure 1.

3.2 | Phenotypical and biochemical analysis of strain Marseille-P3135

The strain is a nonmotile Gram-positive rod, unable to sporulate, oxidase negative, and catalase positive. It grows under anaerobic conditions between 25°C and 42°C but optimally at 37°C. As for acidity tolerance, this strain was able to survive in media with pH ranging between 6 and 8.5 and could sustain only a 5 g/L NaCl concentration. Colonies have a smooth appearance with a mean diameter of 0.5 mm. Moreover, cells had a length of 0.7–1.6 µm when seen under electron microscope and an average diameter of 0.4 µm (Figure 2).

Gel view comparing the mass spectrum of the strain and other closely related species are represented in Figure 3. We can observe that *E. lenta* and *E. timonensis* clearly presented common peaks notably around 4,550 and 6,640 m/z. *Gordonibacter pamelaeae* and *Adlercreutzia equolifaciens* which are closed of *E. timonensis* but not members of *Eggerthella* genus have a clear different spectrum profile.

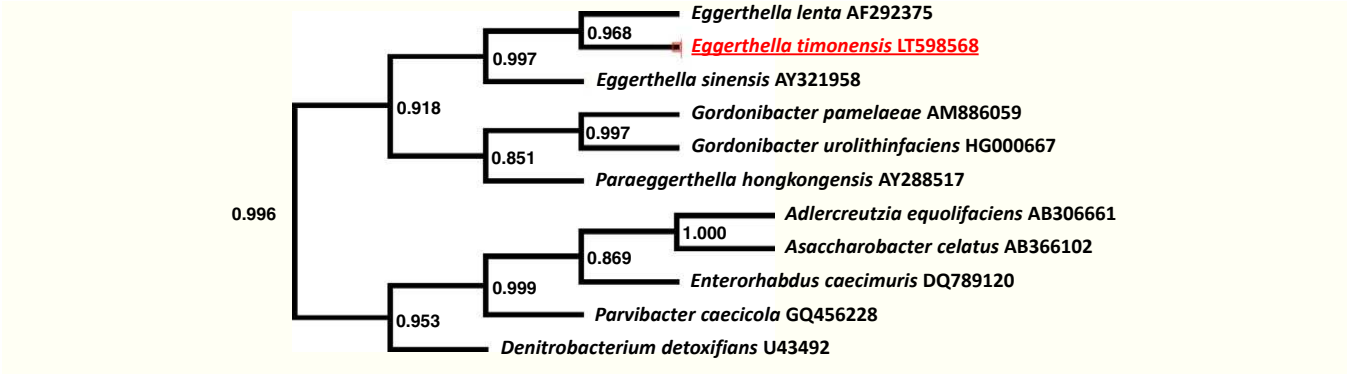


FIGURE 1 Phylogenetic tree showing the placement of *Eggerthella timonensis* strain Marseille-P3135 relative to other close species. 16S RNA sequences are recovered after a nucleotide blast with the database of Silva project "The All-Species Living Tree" (LTPs121). Muscle was used for sequence alignment and Fast tree for approximately maximum likelihood sequence tree generation. Local values obtained with the Shimodaira-Hasegawa test are shown on the nodes. Duplicated and bad taxonomic reference species were removed

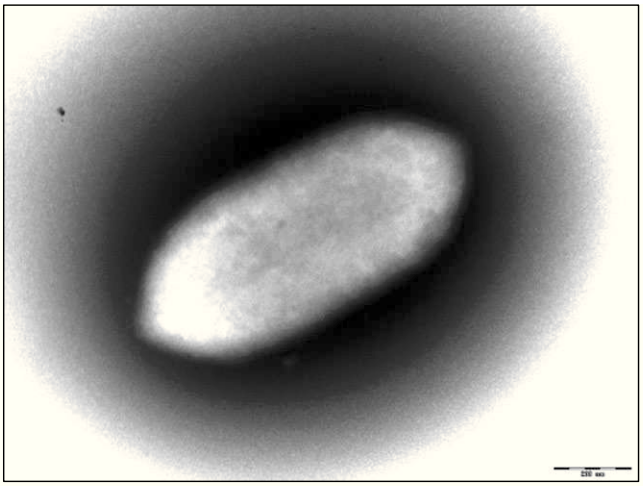


FIGURE 2 Electron micrographs of *Eggerthella timonensis* strain Marseille-P3135 generate wit Morgagni 268D (Philips) transmission electron microscope operated at 80 keV. Scale bar = 200 nm

Examined traits using API20A, API50CH, and APIZYM are detailed in supplementary Table 1. A comparison of some biochemical features was done in Table 1 between the studied strain and the literature data of closely related species (Lau et al., 2004b; Würdemann et al., 2009).

The major fatty acids content of this bacterium were 9-Octadecenoic acid (37%), Octadecanoic acid (28%), and Hexadecanoic acid (28%) (Supplementary Table 2). As for its closely related species, *E. lenta* and *G. pamelaeae* had 12-methyl tetradecanoic acid as its major component (Kageyama et al., 1999; Würdemann et al., 2009) (Table 1). These data allow us to suggest that the major fatty acid could be not representative to a genus.

Strain Marseille-P3135 had a minimal inhibitory concentration (μg/ml) for vancomycin of 1.5, 0.047 for rifampicin, 1.5 for benzylpenicillin, 0.75 for amoxycillin, 0.94 for imipenem, 0.25 for tigecycline, 0.016 for erythromycin, 0.38 for minocycline, 0.125 for Teicoplanin, 64 for daptomycin, 0.19 for metronidazole, and more than 256 for ceftriaxone, colistin, and amikacin.

Composed of two scaffolds, the genome of strain Marseille-P3135 is 3,916,897 bp long with 65.17 mol% G+C content. When analyzing the detected 3,371 predicted genes, 57 were RNAs (2 genes are 23S rRNA, 2 genes are 5S rRNA, 2 genes are 16S rRNA, and 51 genes are tRNA genes) and 3,314 were protein-coding genes. Moreover, 2,524

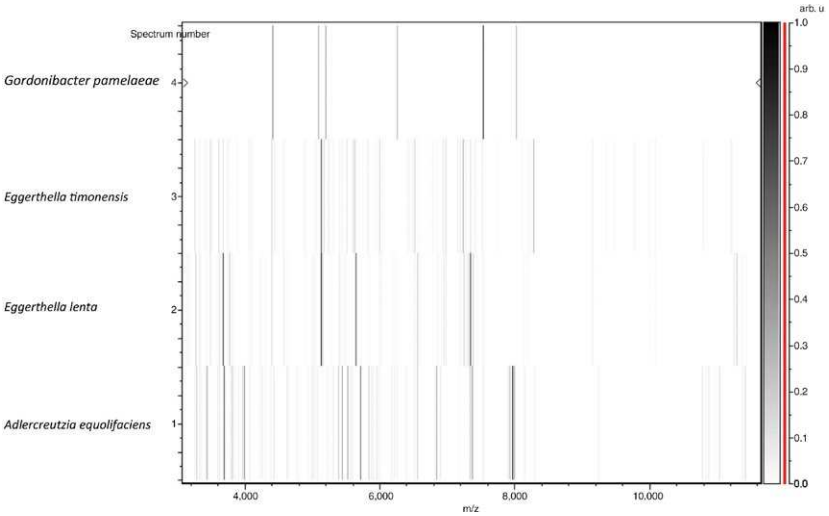


FIGURE 3 Gel view comparing mass spectra of *Eggerthella timonensis* strain Marseille-P3135 to other species by displaying the raw spectra of different species in a pseudo-gel like arrangement. The *x*-axis represents the *m*/*z* value and the left *y*-axis correspond to the running spectrum number deriving from subsequent spectra loading. Intensities of the peaks are represented by with a gray scale. Also, the correlation between the peak color and its intensity is represented in the right *y*-axis with arbitrary units. Species shown for this analysis are noted on the left

TABLE 1 Differential characteristics of *Eggerthella timonensis* strain Marseille-P3135 with *Eggerthella lenta* strain NCTC 11813 (Kageyama et al., 1999), *Eggerthella sinensis* (Lau et al., 2004b), and *Gordonibacter pamelaeae* strain 7-10-1-b (T) (Würdemann et al., 2009)

Properties	*E. timonensis*	*E. lenta*	*E. sinensis*	*G. pamelaeae*
Cell length (μm)	0.7–1.6	1.0–6.0	0.79	1.01 ± 0.21
Gram stain	Positive	Positive	Positive	Positive
Indole	–	–	–	Na
Endospore formation	–	–	–	–
Oxygen requirement	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic
Major fatty acid	9-Octadecenoic acid	12-methyl tetradecanoic acid	Na	12-methyl tetradecanoic acid
Motility	–	–	–	–
Acid from:				
D-Ribose	+	Na	Na	Na
L-Arabinose	+	–	–	Na
D-glucose	+	Na	–	+
D-Mannose	+	–	–	–
D-Mannitol	+	–	Na	Na
Production of:				
β-galactosidase	–	–	–	Na
Catalase	+	Na	+	+
Urease	–	–	–	Na
Alkaline phosphatase	–	–	–	–
Habitat	Human gut	Human gut	Blood culture	Human colon
G+C content (mol%)	65.17	64.2	Na	64

genes (76.16%) were assigned by cogs or by NR blast a putative function. ORFans were found in 190 genes (5.73%) and 495 genes (14.94%) were annotated as hypothetical proteins (Table 2). Circular visualization of the species genome can be seen in supplementary Figure 2. The distribution of the genes into clusters of orthologous groups (COG) functional categories is represented in Table 3.

3.3 | Comparison of genome properties

The draft genome sequence of the present new species was compared to with *G. pamelaeae* (FP929047) which is close but outside the *Eggerthella* genus and *E. lenta* (ABTT00000000) as the closest species and alone member of the genus for which the genome is available. The draft genome sequence of our strain was larger than that of *G. pamelaeae* and *E. lenta* (3,608 and 3,632 Mb, respectively). The G+C content was larger too (64% and 64.2%, respectively). The gene content was larger than that of *G. pamelaeae* and *E. lenta* (2,027 and 3,070, respectively). The functional classes' distribution of predicted genes of the present genome according to the COGs of proteins is shown in Figure S3. The latter showed an identical profile for the three compared strains.

E. timonensis shared higher number of proteins with *E. lenta* (80.76%) (Table 4).

Subsequently, DNA–DNA hybridization values between *E. timonensis* and other species with standing in nomenclature was of 43.6 with *E. lenta*, 21.2 with *G. pamelaeae* (Table 5). Interestingly, these data show that the genome of the strain was closer than *E. lenta* one and further of *G. pamelaeae* supporting the hypothesis that strain Marseille-P3135 as a unique species which is close to species of the *Eggerthella* genus (Kim et al., 2014; Tindall, Rosselló-Móra, Busse, Ludwig, & Kämpfer, 2010; Wayne et al., 1987).

4 | CONCLUSION

In conclusion, culturomics helped us in the isolation of a new species previously uncultured from the human gut normal flora and its description using a taxonogenomics approach. Given its 16S rRNA gene sequence divergence higher than 1.3% with its phylogenetically closest species with standing in nomenclature, we propose a new species *E. timonensis*, type strain Marseille-P3135 (=CSUR P3135, =CCUG 70327).

4.1 | *E. timonensis* sp. nov. description

E. timonensis (ti.mo.nen′sis, N.L. fem. adj., with *timonensis* referring to La Timone, which is the name of the hospital in Marseille (France) where this work was performed).

The strain Marseille-P3135^T (=CSUR P3135, =CCUG 70327) is the type strain of the species *E. timonensis*.

It is a nonmotile Gram-positive rod, unable to sporulate, oxidase negative, and catalase positive. It grows under anaerobic conditions optimally at 37°C. Colonies have a smooth appearance with a mean

TABLE 2 Nucleotide content and gene count levels of the genome

	Number	Percent[a]
Size (bp)	3,916,897	100
Number of G + C	2,552,694	65.17
Number total of genes	3,371	100
Number total of protein genes	3,314	98.31
Number total of RNA genes	57	1.69
Number total of tRNA genes	51	1.51
Number total of RNA (5S, 16S, 23S) genes	6	0.18
Coding sequence size	3,459,399	88.32
Coding sequence gene protein size	3,446,256	87.98
Coding sequence tRNA gene size	3,987	0.10
Coding sequence (5S, 16S, 23S) gene size	9,156	0.23
Number of protein-coding gene	3,314	100
Number of protein associated to COGs	2,046	61.74
Number of protein associated to orfan	190	5.73
Number of protein with peptide signal	492	14.85
Number of gene associated to resistance genes	0	0
Number of gene associated to PKS or NRPS	14	0.42
Number of genes associated to virulence	555	16.75
Number of protein with TMH	967	29.18

COG, clusters of orthologous groups.
[a]The total is considered either the genome size or the number of protein-coding genes found after annotating the genome.

TABLE 3 Number and percentage of genes correlated with some COG functional categories

Code	Value	% of total	Description
[J]	167	5.04	Translation
[A]	17	0.51	Cell motility
[K]	190	5.73	Amino acid transport and metabolism
[L]	98	2.96	Cell wall/membrane biogenesis
[B]	70	2.11	Defense mechanisms
[D]	0	0	Nuclear structure
[Y]	191	5.76	Transcription
[V]	0	0	RNA processing and modification
[T]	0	0	Chromatin structure and dynamics
[M]	370	11.16	Energy production and conversion
[N]	114	3.44	Posttranslational modification, protein turnover, chaperones
[Z]	99	2.99	Signal transduction mechanisms
[W]	25	0.75	Cell cycle control, mitosis, and meiosis
[U]	34	1.03	Intracellular trafficking and secretion
[O]	79	2.38	Replication, recombination, and repair
[X]	20	0.60	Extracellular structures
[C]	0	0	Cytoskeleton
[G]	97	2.93	Coenzyme transport and metabolism
[E]	94	2.84	Carbohydrate transport and metabolism
[F]	105	3.17	Lipid transport and metabolism
[H]	10	0.30	Mobilome: prophages, transposons
[I]	65	1.96	Nucleotide transport and metabolism
[P]	175	5.28	General function prediction only
[Q]	45	1.36	Secondary metabolites biosynthesis, transport, and catabolism
[R]	105	3.17	Function unknown
[S]	125	3.77	Inorganic ion transport and metabolism
_	1.268	38.26	Not in COGs

COG, clusters of orthologous groups.

diameter of 0.5 mm. Moreover, cells had a length of 0.7–1.6 μm when seen under electron microscope and an average diameter of 0.4 μm.

It is able to produce esterase C4, esterase lipase C8, acid phosphatase, and naphtol-AS-BI-phosphohydrolase. As well, it can ferment D-galactose, glycerol, D-ribose, L-arabinose, D-xylose, D-glucose, N-acetylglucosamine, D-fructose, D-mannitol, D-mannose, L-rhamnose, D-sorbitol, D-cellobiose, amygdalin, methyl-αD-mannopyranoside, arbutin, escutin ferric citrate, salicin, D-maltose, D-lactose, D-melibiose, D-trehalose, D-saccharose/sucrose, D-raffinose, D-melezitose, amidon, gentiobiose, D-tagatose, D-arabitol, L-fucose, potassium 5-Ketogluconate, and potassium 2-Ketogluconate. Finally, it can ferment glucose, manitol, saccharose, lactose, maltose, xylose, salicin, arabinose, mannose, cellobiose, trehalose, and rhamnose.

TABLE 4 This table represents in its upper right the total number of shared orthologous proteins along with the percent of similarity of the nucleotides corresponding to it in the lower left

	ET	GP	EL	DD	EC	AE
ET	**3,314**	902	1,543	937	1,164	1,204
GP	65.13	**2,027**	901	486	613	619
EL	70.56	65.79	**3,07**	942	1,146	1,19
DD	62.43	61.34	62.56	**1,762**	864	890
EC	63.23	63.83	64.36	62.04	**2,455**	1,15
AE	62.78	63.68	64.32	62.0	80.76	**2,281**

ET, *Eggerthella timonensis*; GP, *Gordonibacter pamelaeae*; EL, *Eggerthella lenta*; DD, *Denitrobacterium detoxificans*; EC, *Enterorhabdus caecimuris*; AE, *Adlercreutzia equolifaciens*.
Values with bold font represent the numbers of proteins per genome.

TABLE 5 Pairwise comparison of strain Marseille-P3135 with other species. GGDC formula 2 was used: DDH estimates based on identities/HSP length)[a]

	ET	EL	GP	AE	EC	DD
ET	100%	43.6 [40.3%–47.1%]	21.2 [18%–24.9%]	15.1 [12.2%–18.6%]	15.1 [12.2%–18.6%]	13.2 [10.5%–16.5%]
EL		100%	22.3 [19%–25.9%]	15.3 [12.4%–18.7%]	15 [12.1%–18.4%]	13.4 [10.7%–16.7%]
GP			100%	15 [12.1%–18.4%]	15 [12.2%–18.5%]	13.3 [10.6%–16.6%]
AE				100%	26.1 [22.8%–29.7%]	13.4 [10.6%–16.7%]
EC					100%	13.4 [10.7%–16.8%]
DD						100%

ET, *Eggerthella timonensis*; GP, *Gordonibacter pamelaeae*; EL, *Eggerthella lenta*; DD, *Denitrobacterium detoxificans*; EC, *Enterorhabdus caecimuris*; AE, *Adlercreutzia equolifaciens*.
[a]Inherent uncertainty in DDH values estimation are represented in the confidence intervals.

The draft genome of the type strain is 3,916,897 bp long with a G + C content of 65.17%. The 16S rRNA gene and genome sequences were deposited in EMBL-EBI under accession number LT598568 and FXXA00000000, respectively. *E. timonensis* strain Marseille-P3135 was isolated from the stool samples of a healthy 8-year-old pygmy female.

ACKNOWLEDGMENTS

The authors acknowledge Xegen Company (http://www.xegen.fr/) for performing the genomic analysis, Claudia Andrieu for her administrative assistance, and Magdalen Lardière for her linguistic assistance. This study was funded by the "Fondation Méditerranée Infection."

CONFLICT OF INTEREST

None to be declared.

ORCID

Fadi Bittar http://orcid.org/0000-0003-4052-344X
Jean-Christophe Lagier http://orcid.org/0000-0003-0771-8245
Pierre-Edouard Fournier http://orcid.org/0000-0001-8463-8885
Frédéric Cadoret http://orcid.org/0000-0002-9553-8959

REFERENCES

Bilen, M., Cadoret, F., Daoud, Z., Fournier, P.-E., & Raoult, D. (2016). "*Eggerthella timonensis*" gen. nov. isolated from a human stool sample. *New Microbes and New Infections*, *17*, 41–42. https://doi.org/10.1016/j.nmni.2016.11.001

Dione, N., Sankar, S. A., Lagier, J.-C., Khelaifia, S., Michele, C., Armstrong, N., … Fournier, P.-E. (2016). Genome sequence and description of Anaerosalibacter massiliensis sp. nov. *New Microbes and New Infections*, *10*, 66–76. https://doi.org/10.1016/j.nmni.2016.01.002

Eggerth, A. H. (1935). The Gram-positive Non-spore-bearing Anaerobic Bacilli of Human Feces. *Journal of Bacteriology*, *30*, 277–299.

Elsawi, Z., Togo, A. H., Beye, M., Dubourg, G., Andrieu, C., Armsrtong, N., … Khelaifia, S. (2017). Hugonella massiliensis gen. nov., sp. nov., genome sequence, and description of a new strictly anaerobic bacterium isolated from the human gut. *MicrobiologyOpen*, *6*, e00458. https://doi.org/10.1002/mbo3.458

Gill, S. R., Pop, M., DeBoy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., … Nelson, K. E. (2006). Metagenomic Analysis of the Human Distal Gut Microbiome. *Science*, *312*, 1355–1359. https://doi.org/10.1126/science.1124234

Gouret, P., Paganini, J., Dainat, J., Louati, D., Darbo, E., Pontarotti, P., & Levasseur, A. (2011). Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: The multi-agent software system DAGOBAH. In *Evolutionary biology – Concepts, biodiversity, macroevolution and genome evolution* (pp. 71–87). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-20763-1_5

Gouret, P., Thompson, J. D., & Pontarotti, P. (2009). PhyloPattern: Regular expressions to identify complex patterns in phylogenetic trees. *BMC Bioinformatics*, *10*, 298. https://doi.org/10.1186/1471-2105-10-298

Gouret, P., Vitiello, V., Balandraud, N., Gilles, A., Pontarotti, P., & Danchin, E. G. J. (2005). FIGENIX: Intelligent automation of genomic annotation: Expertise integration in a new software platform. *BMC Bioinformatics*, *6*, 198. https://doi.org/10.1186/1471-2105-6-198

Greub, G. (2012). Culturomics: A new approach to study the human microbiome. *Clinical Microbiology and Infection*, *18*, 1157–1159. https://doi.org/10.1111/1469-0691.12032

Kageyama, A., Benno, Y., & Nakase, T. (1999). Phylogenetic evidence for the transfer of Eubacterium lentum to the genus *Eggerthella* as *Eggerthella lenta* gen. nov., comb. nov. *International Journal of Systematic and Evolutionary Microbiology*, *49*, 1725–1732. https://doi.org/10.1099/00207713-49-4-1725

Kim, M., Oh, H.-S., Park, S.-C., & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 346–351. https://doi.org/10.1099/ijs.0.059774-0

Lagier, J.-C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., … Raoult, D. (2012a). Microbial culturomics: Paradigm shift in the human gut microbiome study. *Clinical Microbiology and Infection*, *18*, 1185–1193. https://doi.org/10.1111/1469-0691.12023

Lagier, J.-C., Edouard, S., Pagnier, I., Mediannikov, O., Drancourt, M., & Raoult, D. (2015). Current and past strategies for bacterial culture in clinical microbiology. *Clinical Microbiology Reviews*, *28*, 208–236. https://doi.org/10.1128/CMR.00110-14

Lagier, J.-C., El Karkouri, K., Nguyen, T.-T., Armougom, F., Raoult, D., & Fournier, P.-E. (2012b). Non-contiguous finished genome sequence and description of *Anaerococcus senegalensis* sp. nov. *Standards in Genomic Sciences*, *6*, 116–125. https://doi.org/10.4056/sigs.2415480

Lagier, J.-C., Khelaifia, S., Alou, M. T., Ndongo, S., Dione, N., Hugon, P., … Raoult, D. (2016). Culture of previously uncultured members of the human gut microbiota by culturomics. *Nature Microbiology*, *1*, 16203. https://doi.org/10.1038/nmicrobiol.2016.203

Lau, S. K. P., Woo, P. C. Y., Fung, A. M. Y., Chan, K., Woo, G. K. S., & Yuen, K. (2004a). Anaerobic, non-sporulating, Gram-positive bacilli bacteraemia characterized by 16S rRNA gene sequencing. *Journal of Medical Microbiology*, *53*, 1247–1253. https://doi.org/10.1099/jmm.0.45803-0

Lau, S. K. P., Woo, P. C. Y., Woo, G. K. S., Fung, A. M. Y., Wong, M. K. M., Chan, K., … Yuen, K. (2004b). Eggerthella hongkongensis sp. nov. and eggerthella sinensis sp. nov., two novel Eggerthella species, account for half of the cases of Eggerthella bacteremia. *Diagnostic Microbiology and Infectious Disease*, *49*, 255–263. https://doi.org/10.1016/j.diagmicrobio.2004.04.012

Ley, R. E., Bäckhed, F., Turnbaugh, P., Lozupone, C. A., Knight, R. D., & Gordon, J. I. (2005). Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 11070–11075. https://doi.org/10.1073/pnas.0504978102

Ley, R. E., Turnbaugh, P. J., Klein, S., & Gordon, J. I. (2006). Microbial ecology: Human gut microbes associated with obesity. *Nature*, *444*, 1022–1023. https://doi.org/10.1038/4441022a

Moore, W. E. C., Cato, E. P., & Holdeman, L. V. (1971). Eubacterium lentum (Eggerth) Prévot 1938: Emendation of Description and Designation of the Neotype Strain. *International Journal of Systematic and Evolutionary Microbiology*, *21*, 299–303. https://doi.org/10.1099/00207713-21-4-299

Morel, A.-S., Dubourg, G., Prudent, E., Edouard, S., Gouriet, F., Casalta, J.-P., … Raoult, D. (2015). Complementarity between targeted real-time specific PCR and conventional broad-range 16S rDNA PCR in the syndrome-driven diagnosis of infectious diseases. *European Journal of Clinical Microbiology & Infectious Diseases*, *34*, 561–570. https://doi.org/10.1007/s10096-014-2263-z

Price, M. N., Dehal, P. S., & Arkin, A. P. (2009). FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution*, *26*, 1641–1650. https://doi.org/10.1093/molbev/msp077

Raoult, D., & Henrissat, B. (2014). Are stool samples suitable for studying the link between gut microbiota and obesity? *European Journal of Epidemiology*, *29*, 307–309. https://doi.org/10.1007/s10654-014-9905-4

Seng, P., Drancourt, M., Gouriet, F., La Scola, B., Fournier, P.-E., Rolain, J. M., & Raoult, D. (2009). Ongoing revolution in bacteriology: Routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clinical Infectious Diseases*, *49*, 543–551. https://doi.org/10.1086/600885

The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks [WWW Document] (n.d.) Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3965112/ (accessed 7.10.17).

Tindall, B. J., Rosselló-Móra, R., Busse, H.-J., Ludwig, W., & Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *International Journal of Systematic and Evolutionary Microbiology*, *60*, 249–266. https://doi.org/10.1099/ijs.0.016949-0

Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., … Truper, H. G. (1987). Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *International Journal of Systematic and Evolutionary Microbiology*, *37*, 463–464. https://doi.org/10.1099/00207713-37-4-463

Würdemann, D., Tindall, B. J., Pukall, R., Lünsdorf, H., Strömpl, C., Namuth, T., … Oxley, A. P. A. (2009). *Gordonibacter pamelaeae* gen. nov., sp. nov., a new member of the Coriobacteriaceae isolated from a patient with Crohn's disease, and reclassification of *Eggerthella hongkongensis* Lau et al. 2006 as *Paraeggerthella hongkongensis* gen. nov., comb. nov. *International Journal of Systematic and Evolutionary Microbiology*, *59*, 1405–1415. https://doi.org/10.1099/ijs.0.005900-0

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Bilen M, Fonkou MDM, Tomei E, et al. *Eggerthella timonensis* sp. nov, a new species isolated from the stool sample of a pygmy female. *MicrobiologyOpen*. 2017;e575. https://doi.org/10.1002/mbo3.575

Part III

Conclusion and perspectives

Partie III

Conclusion et perspectives

Conclusion and Perspectives

The advances in sequencing, bioinformatics and culture techniques have made the study of the human microbiome very promising. Given the crucial role that plays in maintaining the human health, describing the human microbiome became mandatory. Yet we are still far from describing its entire content, in comparison to the large number of prokaryotes that exist. Resources should be assigned to perform clinical research in the aim of describing the healthy human microbiome and to attempt correlating its content to diseased individuals microbiome. The ability to isolate new bacterial organisms and describing their role on human health is fundamental if we are to more precisely assess the community structure and for drawing new therapeutic strategies for treatment and monitoring, such as *Akkermansia muciniphila* in inflammations and metabolic disorders. Conclusively, our work has expanded our knowledge in terms of the human gut microbiota by its ability to isolate a significant number of non-human bacterial species and new bacterial species along with correlating its sequences to operational taxonomic units. This emphasizes on the fundamental role of culture in describing and unveiling the hidden part of the human microbiome and its commensals that can be used in the future in bacterio-therpay approaches. As well, having a human prokaryotic repertoire is a key step towards unveiling the human microbiome, aiding the scientific community to communicate the recent discoveries and providing useful resources for investigators and clinicians.

Conclusion et Perspectives

Les progrès du séquençage, de la bioinformatique et des techniques de culture ont rendu l'étude du microbiome humain très prometteuse. Étant associé à la santé ou à différentes pathologies chez l'homme, la description du microbiome humain est devenue obligatoire. Pourtant, nous sommes encore loin de décrire l'ensemble de son contenu en comparaison avec le grand nombre de procaryotes qui existent. Des ressources devraient être affectées à la recherche clinique dans le but de décrire le microbiome humain sain et d'essayer de corréler son contenu avec le microbiome des individus malades. La capacité d'isoler de nouveaux organismes bactériens et de décrire leur rôle dans la santé humaine est fondamentale. Plus précisément évaluer le contenu bactérien du microbiote humain permettra d'élaborer de nouvelles stratégies thérapeutiques pour le traitement et la surveillance, comme *Akkermansia muciniphila* dans les inflammations et les troubles métaboliques. En conclusion, notre travail a élargi nos connaissances en termes de microbiote intestinal humain par sa capacité à isoler un nombre important d'espèces bactériennes non humaines et de nouvelles espèces bactériennes dont leurs séquences ont été corrélé aux unités taxonomiques opérationnelles. Ceci souligne le rôle fondamental de la culture dans la description du microbiome humain et de ses commensaux qui peuvent être utilisés dans l'avenir dans des approches bactério-thérapeutique. De plus, avoir un répertoire de procaryotes humains permettra à la communauté scientifique de communiquer les découvertes récentes.

Annex/Annexe

Online access to the currently published new species isolated this project and where I contribute as a first author. L'accès en ligne aux articles des nouvelles espèces isolé dans ce projet et où je contribue en tant que premier auteur :

"*Phoenicibacter massiliensis*" gen. nov., sp. nov. , a new bacterium isolated from the human gut of a pygmy woman
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5496386/
Bacteroides congonensis sp. nov., a new bacterial species from the human gut microbiota
https://www.sciencedirect.com/science/article/pii/S2452231716300112
Gabonibacter timonensis sp., nov., a new bacterium isolated from the human gut of a pygmy female
https://www.ncbi.nlm.nih.gov/pubmed/28179984
Collinsella bouchesdurhonensis sp. nov., identified in human stool sample
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5328908/
Eggerthella timonensis gen. nov., isolated from a human stool sample
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5328910/
Parabacteroides timonensis sp. nov., identified in human stool
https://www.sciencedirect.com/science/article/pii/S2452231716300100
Pygmaiobacter massiliensis sp., nov., a new bacterium isolated from the human gut of a pygmy female
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5284491/
Raoultibacter timonensis gen. nov., sp. nov., a new bacterium isolated from the human gut of a pygmy female
https://www.ncbi.nlm.nih.gov/pubmed/28203374
'*Congobacterium massiliense*' gen. nov. sp. nov., a new bacterium isolated from the gut of a pygmy (Baka) woman
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5192236/
"*Libanicoccus massiliensis*" gen. nov., sp. nov., a new bacterium isolated from a stool sample from a pygmy woman
https://www.sciencedirect.com/science/article/pii/S2052297516301123
"*Mobilobacterium massiliense*" gen.nov. and "*Mobilobacterium timonense*" sp.nov., identified in human stool specimen
DOI: 10.1016/j.nmni.2016.11.008
Olsenella congonensis sp.nov., identified in human stool sample
https://www.sciencedirect.com/science/article/pii/S2052297517300495
Non-contiguous finished genome sequences and description of *Bacteroides*

mediterraneensis sp. nov., *Bacteroides ihuae* sp. nov., *Bacteroides togonis* sp. nov., *Bacteroides ndongoniae* sp. nov., *Bacteroides ilei* sp. nov., *Bacteroides congonensis* sp. nov., identified by culturomics.
https://www.sciencedirect.com/science/article/pii/S2052297518300520

Contribution to other work with data from the current study/ Contribution des données de la présente étude à d'autres travaux:

In the present study, I co-author with my colleagues an article published in Nature microbiology. This manuscript was made to introduce the advancements done by culturomics on the description of the human gut microbiota. This represents a global effort done on different stool samples by the culturomics team. Part of the data included in this manuscript was taken from the results obtained while analyzing the Pygmy stool. Accordingly, new species isolated by this work were correlated to operational taxonomic units generated by other labs. As well, bacterial species previously considered as non-human were added to the human gut prokaryotic database and thus have led to the enlargement of the latter.

Dans ce papier, je participe avec mes collègues à un article publié dans Nature microbioloy. Ce manuscrit présente les progrès réalisés par la culturomics sur la description du microbiote intestinal humain. Ceci correspond à un effort global de culturomics appliqué sur différents échantillons de selles par l'équipe de culturomics. Une partie des données incluses dans ce manuscrit correspond à des résultats obtenus lors de

l'analyse des selles des Pygmées. En conséquence, de nouvelles espèces isolées par ce travail ont été corrélées à des unités taxonomiques opérationnelles générées par d'autres laboratoires. De plus, des espèces bactériennes considérées auparavant comme non-humaines ont été ajoutées à la base de données procaryotes de l'intestin humain.

“Culture of previously uncultured members of the human gut microbiota by culturomics”

https://www.nature.com/articles/nmicrobiol2016203

In the present study, I co-author with the correspondent authors a review published in Nature reviews microbiology. This manuscript reports the achievements done after applying culture and culturomics on the human microbiota description. After performing my literature review: “The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species” (Part I of the thesis manuscript), I was able to summarize the advancements done by culture to describe the human microbiota through creating a database of the human isolated bacterial species. According, I took part of this manuscript by analyzing these advancements and by showing the contribution to date of culturomics overtime in this field and its efficiency.

Dans la présente étude, je participe à une revue publiée dans Nature reviews microbiology. Ce manuscrit traite les avancements accomplis après l'application de la culture et de la culturomics à la description du microbiote

humain. Après avoir effectué ma revue de littérature " The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species " (partie I du manuscrit de thèse), j'ai pu résumer la contribution de la culture pour décrire le microbiote humain en créant une base de données des espèces bactériennes isolées chez l'homme. Alors, j'ai participé à ce manuscrit en analysant les avancements fait part la culturomics et la culture dans ce domaine.

"Culturing the human microbiota and culturomics"
https://www.nature.com/articles/s41579-018-0041-0