UNIVERSITÉ D'AIX-MARSEILLE

FACULTÉ DE MÉDECINE DE MARSEILLE

ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

THÈSE

Présentée et publiquement soutenue devant

LA FACULTÉ DE MÉDECINE DE MARSEILLE

Le 23 Novembre 2018

Par Khoudia DIOP

Caractérisation du Microbiote des Flores Vaginales Normale et de Vaginose Bactérienne

Pour obtenir le grade de Docteur de l'Université d'AIX-MARSEILLE

Pathologie Humaine, Spécialité Maladies Infectieuses et Microbiologie

Membres du Jury de la Thèse :

Mr le Professeur Pierre-Edouard FOURNIER	Président du jury
Mr le Professeur Max MAURIN	Rapporteur
Mme le Docteur Patricia RENESTO	Rapporteur
Mme le Professeur Florence FENOLLAR	Directrice de thèse

Institut Hospitalo-Universitaire Méditerranée Infection,

Unité de Recherche Vecteurs-Infections Tropicales et Méditerranéennes, IRD, SSA, AP-HM

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

Remerciements

Même si cette thèse est un travail personnel, je souhaite ici rendre hommage et exprimer ma profonde gratitude à tous ceux qui, de près ou de loin, ont contribué à sa réalisation et à son aboutissement.

Bien évidemment, c'est à ma **directrice de thèse**, le professeur **Florence FENOLLAR,** que vont mes premiers remerciements. Elle a été pour moi, tout au long de cette thèse, un mentor remarquable. Malgré une vie professionnelle chargée, elle a su m'apporter un soutien infaillible, une disponibilité, une écoute, une confiance et des conseils précieux et avisés. Mais au-delà de ses indéniables qualités de médecin microbiologiste et de chercheur, c'est par ses qualités humaines qu'elle a su rendre ces années agréables et enrichissantes. Elle a toujours été très attentive à mon égard, et pas seulement à la microbiologie qui nous liaient formellement. Nos échanges ont été des leçons de vie qui me guideront tout au long de ma carrière.

Mes remerciements s'adressent aussi au Professeur **Didier Raoult** et à travers lui**,** à la **Fondation Méditerranée Infectiopôle Sud,** pour m'avoir accueilli au sein de l'Institut Hospitalo-Universitaire Méditerranée Infection, et pour avoir financé mes travaux de recherche.

Mes remerciements s'adressent également à Monsieur **Pierre-Édouard Fournier**, Professeur émérite à l'Université Aix-Marseille, qui nous a fait l'honneur et le plaisir d'accepter de présider ce jury malgré ses multiples charges. Votre compétence, votre disponibilité et vos qualités humaines forcent le respect chez toutes les personnes qui ont eu l'honneur et le privilège de vous côtoyer.

Je tiens également à témoigner toute ma reconnaissance aux Professeur **Max Maurin** et Docteur **Patricia Renesto** pour avoir accepté de rapporter cette thèse et de siéger à ce jury.

Par ailleurs, j'aimerais remercier toute ma « famillede la **Culturomics** », en particulier Jean-Christophe Lagier, Saber Khelaifia et Gregory Dubourg ; ainsi que la CSUR, notamment Fréderic Cadoret et Claudia Andreu, pour m'avoir adopté et pour l'ambiance bon-enfant qui y règne et nous a fait oublier les difficultés qu'il nous arrivait de traverser.

J'exprime ma gratitude aux Docteurs **Mbayang Niang** et **Gora Diop** de l'institut Pasteur de Dakar ainsi qu'aux Docteurs **Cheikh Sokhna, Oleg Mediannikov** et **El Hadji Amadou Niang** de l'IRD pour leur sympathie, leurs conseils, la richesse de nos échanges et pour leur soutien sans faille.

Ce cheminement aurait été difficile sans mes amis et proches qui ont toujours été là pour m'encourager, et qui ont su me faire me redonner le courage pendant les moments difficiles. Je tiens particulièrement à remercier du fond du cœur mes **camarades de thèse** : Awa Diop, Issa Isaac, Maryam Tidjani, El Hadji Seck, Niokhor Dione, Camille Valles, Amadou Togo, Lamine Tall, Descartes Maxime, Fatima, Sokhna Ndongo, Pamela Afouda, Safietou Fall, Amy Diakité, Sara Bellali, ainsi que tous mes collègues, les techniciens et le personnel de l'IHU qui ont tous d'une manière ou d'une autre participé à la réussite de ma thèse.

Je n'oublierais pas de remercier tous mes amis de l'UCAD, de Marseille, de Lyon, de Paris et d'ailleurs, qui m'ont soutenu de près ou de loin et pour leur amitié pendant ces années de thèse.

Enfin, **je tiens à remercier toute ma famille**, mes parents (Alioune Badara Diop & Aby Ndiaye), ma grand-mère (Adja Khoudia Diop), mon frère et mes sœurs, qui m'ont supporté, écouté et soutenu, et qui ont partagé les moments d'exaltation mais aussi d'incertitude inhérents à la thèse. Je remercie aussi ma famille parisienne (tontons Khalil, Ousmane, Yamar, leurs épouses et enfants) pour leur soutien et encouragements.

Un **Grand Merci** à toutes et tous celles ou ceux que j'ai pu oublier de nommer.

Sommaire

RESUME/ABSTRACT 7/9

INTRODUCTION 11

CHAPITRE I : Synthèse des Données de la Littérature sur le Microbiote Vaginal 15

Article 1: Bacterial Vaginosis: what do we currently know (Revue). 16

Article 2: Exhaustive Repertoire of Human Vaginal Microbiota (Revue). 45

CHAPITRE II : Caractérisation des Flores Vaginales Normale et de Vaginose Bactérienne par Culturomique et Métagénomique 78

Article 3: Multi-Omics Strategy to Characterize Vaginal Microbiota associated with Bacterial Vaginosis: Culturomics and Metagenomics approaches. 82

Chapitre III : Taxonogénomique : Incorporation de l'Information Génomique dans la Description des Nouvelles Espèces Bactériennes 118

- **Description des nouvelles espèces isolées dans le microbiote de vaginose bactérienne** .. 124

Article 4: *Olegusella massiliensis* gen. nov. sp. nov., strain KHD7^T, a new bacterial genus isolated from the female genital tract of a patient with bacterial vaginosis. 125

Article 5: Description of *Collinsella vaginalis* strain Marseille-P2666, a new member of *Collinsella* genus isolated from the genital tract of a patient suffering from bacterial vaginosis. 135

Article 6: *Corynebacterium fournierii* sp. nov., isolated from the female genital tract of a patient with bacterial vaginosis. 172

Article 7: Isolation of the first *Janibacter* species derived from human specimen: Description of *Janibacter massiliensis* sp. nov. cultured from the vaginal discharge of a patient with bacterial vaginosis. 183

Article 8: Microbial Culturomics Broadens Human Vaginal Flora Diversity: Genome sequence and Description of *Prevotella lascolaii* sp. nov. Isolated from a Patient with bacterial Vaginosis. 217

Article 9: Description of three new *Peptoniphilus* species cultured in the vaginal fluid of a woman diagnosed with bacterial vaginosis: *Peptoniphilus pacaensis* sp. nov., *Peptoniphilus raoultii* sp. nov., and *Peptoniphilus vaginalis* sp. nov. 231

Article 10: Characterization of a novel Gram-stain-positive anaerobic coccus isolated from the female genital tract: Genome sequence and description of *Murdochiella vaginalis* sp. nov. 248

- **Caractérisation des nouvelles espèces isolées dans la flore vaginale normale** 260

Article 11: *Vaginimicrobium propionicum* gen. nov. sp. nov., a novel propionic acid bacterium derived from human vaginal discharge. 261

Article 12: Taxonogenomics and description of *Vaginella massiliensis* gen. nov. sp. nov., strain Marseille-P2517^T, a new bacterial genus isolated from the human vagina. 297

Article 13: Characterization of a New *Ezakiella* Isolated from the Human Vagina: Genome Sequence and Description of *Ezakiella massiliensis* sp. nov. 308

CHAPITRE IV : Annexes & Travaux Collaboratifs 317

Article 14: "*Anaerococcus mediterraneensis*" sp. nov., a new species isolated from human female genital tract. 320

Article 15: “*Arcanobacterium urinimassiliense”* sp. nov., a new bacterium isolated from the urogenital tract. 323

Article 16: ‘‘*Collinsella vaginalis*” sp. nov., a new bacterial species cultivated from human female genital tract. 327

- Description du génome bactérien d’*Ezakiella peruensis* **330**

Article 17: Draft Genome Sequence of *Ezakiella peruensis* Strain M6X2^{T}, a human fecal Gram-stain positive anaerobic coccus. 331

- Culturomics, un envol de l’exploration du microbiote humain **334**

Article 18: Culture of previously uncultured members of the human gut microbiota by culturomics. 335

CONCLUSION ET PERSPECTIVES.. 344

RÉFERENCES……………………………………………………………………..347

RESUME

Grâce aux avancées technologiques incluant des techniques moléculaires beaucoup plus performantes et de nouvelles stratégies OMICS, de nombreuses études se sont intéressées au microbiote vaginal ces dernières années. Elles ont révélé l'impact de ce dernier sur la santé de la femme. En effet, un déséquilibre de la communauté bactérienne vaginale la rend plus vulnérable, la prédisposant à la vaginose bactérienne ainsi qu'à des complications obstétricales et gynécologiques sévères notamment naissance prématurée et maladies sexuellement transmissibles. La prévalence de la vaginose dépend de la population étudiée. Elle a été rapportée chez 10 à 30% des femmes ayant des rapports sexuels avec des hommes et chez 25 à 50% chez celles ayant des rapports sexuels avec d'autres femmes. Elle peut être > 50% en Afrique orientale et australe. La pathogénèse de la vaginose demeure encore méconnue. Les rechutes sont très fréquentes. Le traitement classique par antibiothérapie échoue dans plus de 50% des cas. Les données sur la flore vaginale normale et anormale se sont étoffées ces dernières années. Si les techniques de culture ont permis d'isoler et de décrire de nombreuses bactéries, les méthodes moléculaires ont mis en évidence les limites de la culture en montrant que le vagin est un biotope complexe contenant une large gamme de bactéries non cultivées ou difficiles à identifier. Dans cette thèse, nous avons analysé 50 prélèvements vaginaux provenant de patientes atteintes de vaginose bactérienne et de femmes saines vivant à Marseille en France et dans une zone rurale au Sénégal. Deux approches ont été utilisées afin de cartographier exhaustivement la flore vaginale : une moléculaire, la métagénomique et une par différentes méthodes de culture, la culturomique. Nous avons pu constater une plus grande diversité bactérienne chez les patientes par rapport aux témoins avec l'augmentation d'espèces telles que *Gardnerella vaginalis*, *Atopobium vaginae* ainsi que les procaryotes sensibles à l'oxygène, y compris les Cocci anaérobies à Gram-positif et les *Prevotella*. Les femmes saines renfermaient plus d'espèces de *Lactobacillaceae* et de *Proteobacteria* dans leurs flores. De plus, nous avons

également réussi à isoler pour la première fois grâce à la culturomique un nombre important de nouvelles espèces dans la flore vaginale. Le taux de recouvrement des données obtenues par culturomique et métagénomique s'est révélé faible. En effet, sur les 581 bactéries détectées dans le microbiote vaginal, seules 285 espèces (49%) étaient identifiées par culture, 459 (79%) par métagénomique et 163 en utilisant à la fois ces 2 techniques. Ces résultats soulignent bien la complémentarité de ces 2 approches. Enfin, la combinaison de la métagénomique et la culturomique a permis l'identification d'un complexe de 11 espèces/genres bactériens associés à la vaginose : *G. vaginalis, A. vaginae, Aeroccocus christensenii, Prevotella*, *Peptoniphilus*, *Clostridium*, *Snethia amnii*, *Mycoplasma hominis, Porphyromonas*, *Facklamia languida* et *Gemella asaccharolytica*. L'utilisation de la culturomique a permis d'accroître le répertoire des bactéries humaines avec l'isolement de 27 nouvelles espèces. Parmi elles, 3 (*Peptoniphilus vaginalis, Megasphaera vaginalis* et *Atopobium massiliense*) sont étroitement apparentées, respectivement à *Peptoniphilus* sp DNF00840, *Megasphaera* sp BV3C16-1 et *Atopobium* sp S4-5, trois bactéries auparavant détectées comme associées à la vaginose en utilisant les outils moléculaires. En plus d'enrichir les connaissances sur le microbiote, ce travail souligne la diversité et la richesse du microbiote vaginal. Il a permis aussi de mieux caractériser la dysbiose de la flore vaginale lors de la vaginose bactérienne. Le faible taux de recouvrement entre les données de métagénomique et celles de culturomique montre la nécessité de persévérer dans l'isolement des bactéries par culturomique, afin de confirmer les données moléculaires et la viabilité des bactéries détectées.

Mots-clés : Vaginose bactérienne, Microbiote vaginal, Bactéries anaérobies, Culturomique et Métagénomique.

ABSTRACT

Over the last decades, thanks to advances in technology including much more efficient molecular techniques and new OMICS strategies, many studies have focused on the vaginal microbiota. They have revealed the impact of this one on women's health. Indeed, the disruption of the vaginal bacterial community makes it prone, predisposing her to bacterial vaginosis and severe obstetrical and gynecological conditions, including preterm birth, pelvic inflammatory disease, and also sexually transmitted diseases. The prevalence of bacterial vaginosis depends on the studied population. It has been reported in 10 to 30% of women who have sex with men and in 25 to 50% of women who have sex with women in developed countries. It can be greater than 50% in eastern and southern Africa. The pathogenesis of bacterial vaginosis is still unknown. Relapses are very frequent. Conventional treatment with antibiotic therapy fails in more than 50% of cases. While culture techniques have made it possible to isolate and describe many bacterial species, molecular methods have highlighted the limits of culture by showing that the vaginal tract is a complex ecosystem containing a wide range of uncultivated or difficult-to-identify bacteria. In this thesis, we analyzed 50 vaginal samples from bacterial vaginosis patients and healthy women living in Marseille-France and rural Senegal. Two approaches were used in order to map exhaustively the vaginal flora: one molecularly, the metagenomics and another with different cultures conditions, the culturomics. We found a higher bacterial diversity in patients compared to controls with the increase of species such as *Gardnerella vaginalis*, *Atopobium vaginae* as well as oxygen-sensitive prokaryotes including Gram-positive anaerobic cocci, and *Prevotella* spp. Healthy women contained more *Lactobacillaceae* species and *Proteobacteria* in their microbiota. In addition, we have also managed to isolate for the first time, thanks to culturomics, a large number of new bacterial species in the vaginal flora. The range of overlap between metagenomic and culturomics data was very low. Indeed, of the 581 species of bacteria detected in the vaginal

microbiota, only 285 species (49%) were identified by culture methods, 459 (79%) by metagenomics and 163 were identified using both these 2 techniques. These results highlight the complementarity of these two approaches. Finally, the combination of metagenomics and culturomics has allowed the identification of a complex of 11 bacterial species or genus associated with bacterial vaginosis: *Gardnerella vaginalis*, *Atopobium vaginae*, *Aeroccocus christensenii, Prevotella, Peptoniphilus, Clostridium, Snethia amnii*, *Mycoplasma hominis, Porphyromonas, Facklamia languida,* and *Gemella asaccharolytica*. The use of culturomics has extended the repertoire of human-associated bacteria with the isolation of 27 new bacterial species. Among them, three (*'Peptoniphilus vaginalis'*, *'Megasphaera vaginalis'* and *'Atopobium massiliense')* are closely related to *Peptoniphilus* sp. DNF00840, *Megasphaera* sp. BV3C16-1 and *Atopobium* sp. S4-5, three bacteria detected as associated with bacterial vaginosis using molecular tools. In addition to expanding the knowledge about the human microbiota, this work highlights the diversity and richness of the vaginal microbiota. It has also made it possible to better characterize the dysbiosis of the vaginal flora during bacterial vaginosis. The low range overlap between metagenomic and culturomics data indicates the need to persevere in the isolation of bacteria by culturomics, in order to confirm the molecular data and the viability of the bacteria detected.

Keywords: Bacterial vaginosis; Vaginal Microbiota; Anaerobic bacteria; Culturomics and Metagenomics.

INTRODUCTION

Le microbiote, normalement associé à l'homme, a une influence capitale sur le corps humain : l'immunité, la nutrition ainsi que la physiologie [1, 2]. On estime que le nombre de microorganismes présents dans le microbiome humain est 10 fois supérieur au nombre de cellules de l'organisme [3]. Les membres de ce microbiote entretiennent une relation de mutualisme avec leur hôte créant ainsi un écosystème stable et équilibré prêt à faire face à toutes perturbations [4]. Cet état d'eubiose constitue la première ligne de défense contre l'infection ou l'envahissement de microbes opportunistes [4, 5]. Malgré son importance, il reste encore beaucoup de zones d'ombres sur comment les différents acteurs du microbiote interagissent entre eux et avec leur hôte, mais aussi pourquoi ce microbiote diffère-t-il du point de vue composition, structure et fonction d'une part entre les individus sains et d'autre part entre les personnes saines et les malades. Des études ont été menées afin de comprendre le rôle de la population microbienne dans la santé et les maladies [6].

L'exploration du microbiote était d'abord principalement effectuée avec des méthodes basées sur la culture. Ces techniques n'ont permis d'identifier seulement 20% du microbiote [7]. Cependant depuis une décennie, on assiste à l'aire des nouvelles technologies avec l'avènement de l'hybridation d'ADN *in situ*, de la PCR en temps réel et du séquençage de nouvelle génération [8, 9]. Depuis 2007, les instituts nationaux de santé ont lancé le projet sur le microbiote humain (PMH) et divers échantillons de peau, de nez, de bouche, du tractus gastro-intestinal et du vagin issus d'individus sains ont été examinés afin de caractériser le microbiote « normal » [6, 10]. De là, de nombreux projets se sont intéressés au microbiote humain dont le microbiote vaginal.

Le microbiote vaginal est colonisé dès les premières heures ou lors de la naissance d'une fille par les flores cutané, vaginal ou intestinal de la mère [11]. C'est un écosystème complexe

et dynamique hébergeant plusieurs types de microorganismes, majoritairement des bactéries [12–15]. Les femmes pubères produisent environ 1 à 4 ml de liquide vaginal contenant de 10^8 à 10^9 bactéries par ml de sécrétions [16]. La flore vaginale a été décrite pour la première fois en 1892 par le gynécologue allemand Albert Döderlein. Il avait remarqué que les femmes en bonne santé présentaient une flore vaginale homogène constituée de bacilles Gram-positif (bacilles de Döderlein) et identifiés plus tard, en 1901 par Beijerink, comme appartenant au genre *Lactobacillus* [4]. Dans des conditions normales chez les femmes pré-ménopausées en bonne santé, 70 à 90% des bactéries vaginales sont des lactobacilles [17]. Parmi plus de 200 espèces de *Lactobacillus* connues dans la nomenclature, un peu plus de 20 espèces ont été retrouvées dans la flore vaginale [18]. Cependant, la flore vaginale est dominée par seulement une à deux espèces de *Lactobacillus,* les plus fréquentes étant *L. crispatus, L. jensenii, L. gasseri* et *L. iners* [19]. Cette flore vaginale normale maintient l'homéostasie et joue un rôle crucial dans la santé des femmes [20].

En effet, les lactobacilles protègent l'écosystème vaginal et maintiennent son équilibre grâce à la production de molécules antimicrobiennes, tel que le peroxyde d'hydrogène, l'acide lactique et les bactériocines [4, 21]. De plus, la production d'acide lactique provoque une acidification de l'environnement vaginal et par conséquent une diminution du pH [22]. Ce faible pH vaginal inhibe la croissance de certains agents pathogènes et espèces commensales autres que les lactobacilles. Toutefois, la composition de cette flore bactérienne n'est pas statique et varie de jour en jour en fonction des facteurs intrinsèques et extrinsèques tels que l'âge, les taux d'hormonaux œstrogènes, les pratiques sexuelles, l'environnement et les prises de médicaments comme les antibiotiques [23, 24]. Une modification de la quantité des espèces présentes dans le tractus vaginal peut provoquer une dysbiose telle que la vaginose bactérienne.

La vaginose bactérienne est l'affection vaginale la plus répandue chez les femmes en âge de procréer [25]. Sa prévalence dépend de la population étudiée. Dans les pays développés,

elle a été rapportée chez 10 à 30% des femmes ayant des rapports sexuels avec des hommes et chez 25 à 50% de celles ayant des rapports sexuels avec d'autres femmes [26, 27]. Par contre en Afrique orientale et australe, sa prévalence peut être supérieure à 50% [28]. La vaginose bactérienne peut engendrer des complications graves telles que le risque accru de développer une grossesse anormale, des infections urogénitales, des maladies inflammatoires pelviennes et l'acquisition ou la transmission de plusieurs infections sexuellement transmissibles (*Neisseria gonorrheae*, *Chlamydia trachomatis*, herpès simplex de type 2 et virus de l'immunodéficience humaine) [29–31]. La vaginose bactérienne est très souvent traitée avec des antibiotiques, principalement du métronidazole et de la clindamycine, mais le traitement échoue fréquemment et le taux de rechute est estimé à 50% après six mois [32, 33]. L'étiologie de cette affection gynécologique est méconnue et elle reste l'une des grandes énigmes médicales de la femme. Récemment, l'explosion des techniques moléculaires a augmenté nos connaissances du microbiote vaginal. Dès lors, la vaginose bactérienne apparait comme un changement taxonomique de la flore, occasionné par une forte surcroissance de bactéries anaérobies strictes ou facultatives auparavant minoritaires dans le vagin (100 à 1 000 fois plus que dans la flore normale) et d'espèces jusque-là non cultivées [34–37] à la suite d'un déficit inexpliqué des lactobacilles. Pour une meilleure appréhension de ce problème de santé publique, le retour de la culture est nécessaire afin d'isoler et d'étudier ces bactéries et potentielles nouvelles espèces associées à la vaginose observées uniquement lors de l'utilisation des méthodes moléculaires. Cette renaissance de la culture s'observe avec l'arrivée de la culturomique microbienne, une approche tendant à reproduire l'environnement originel des bactéries par multiplication des conditions de culture et variation des paramètres physico-chimiques couplées à une identification rapide des bactéries par la spectrométrie de masse MALDI-TOF [38, 39].

C'est d'ailleurs dans cette logique que s'inscrit ce travail de thèse ayant comme objectif principal de caractériser la flore vaginale humaine de façon la plus exhaustive possible pour

mieux comprendre la vaginose bactérienne. Plus précisément, il s'agit dans un premier temps de caractériser la flore vaginale normale, puis celle de vaginose bactérienne et enfin comparer ces deux types de flores afin de mieux comprendre cette dysbiose en vue de proposer des traitements beaucoup plus appropriés.

Pour mieux aborder cette thèse de doctorat, dès l'entame de nos travaux, nous avons effectué une synthèse bibliographique (**Chapitre I**) qui a fait l'objet de 2 revues de la littérature. La première revue, se concentre sur la flore vaginale saine et la vaginose bactérienne. Elle fait le point sur l'état des connaissances actuelles et les limites dans la prévention et la gestion de la vaginose. Dans la deuxième revue en revanche, en utilisant un programme informatique pour parcourir la littérature scientifique consacrée au microbiote vaginal, nous avons essayé de dresser le répertoire exhaustif de toutes bactéries trouvées dans la flore vaginale humaine. Ensuite, en utilisant à la fois les techniques de culturomique et de métagénomique, nous avons caractérisé la flore vaginale saine et celle de vaginose bactérienne en vue d'identifier les bactéries impliquées dans ce déséquilibre (**Chapitre II**). La combinaison de ces deux méthodes complémentaires, a permis la détection d'une large gamme d'espèces bactériennes précédemment non connues du microbiote vaginal dont 27 nouvelles. Douze des nouvelles espèces isolées au cours de ce travail ont été décrites par taxonogénomique (**Chapitre III**), une approche qui combine les caractéristiques phénotypiques avec les informations protéomiques et la description du génome entier annoté [40, 41]. D'autres (trois) sont décrites sous forme de «new species announcement» (**Chapitre IV**), un format qui ne relate que quelques caractéristiques principales de la bactérie [42]. Dans cette dernière partie, nous y rapportons aussi l'analyse génomique d'*Ezakiella peruensis*, la seule bactérie officiellement reconnue de son genre. Comme nous avons isolé la deuxième espèce du genre *Ezakiella (E. massilensis)* et que le génome de *E. peruensis* n'était pas disponible, nous avons donc séquencé le génome des deux espèces afin d'effectuer des comparaisons génomiques.

CHAPITRE I :

Synthèse des Données de la Littérature sur le Microbiote Vaginal

Article 1:

Bacterial Vaginosis: what do we currently know (Revue).

Khoudia Diop and Florence Fenollar

To be submitted soon in Clinical Infectious Diseases journal

Avant-propos

Le tractus vaginal est un biotope très riche en nutriments pour les microbes. Caractérisé par des techniques de culture et de biologie moléculaire, son microbiote constitue un écosystème complexe, dynamique, et capable d'influencer la santé humaine. Vu le rôle protecteur de cette flore vaginale, un changement de sa composition peut engendrer la vaginose bactérienne, le seul état gynécologique directement associé au déséquilibre des communautés bactériennes vaginales. La vaginose bactérienne est le trouble vaginal le plus répandu chez les femmes en âge de procréer, provoquant le plus souvent des leucorrhées dégageant une forte odeur et parfois une irritation, poussant les femmes à consulter [43]. En fonction de la population étudiée, la prévalence de la vaginose bactérienne varie de 4% à 75% entre les femmes asymptomatiques et celles souffrant de maladies sexuellement transmissibles, respectivement [44, 45]. Plusieurs facteurs de risque ont été rapportés : race noire, partenaires sexuels multiples, absence d'utilisation de préservatifs, pratiques sexuels non vaginales (orales ou anales suivies de vaginales) [46, 47], rapports sexuels avec d'autres femmes [48], utilisation de dispositif intra-utérin comme méthode de contraception [49], utilisation de gel de toilette intime [50], et tabagisme [51].

Le diagnostic de vaginose bactérienne dépend du praticien et des moyens du centre médical. Pendant longtemps, il reposait sur les critères d'Amsel [52] se basant sur la présence de trois de ces quatre signes : des sécrétions vaginales grisâtres homogènes et adhérentes à la paroi vaginale, une forte odeur de poisson avarié suite à l'ajout d'hydroxyde de potassium, un pH vaginal > 4,5 et la présence de « clue-cells » (Fig. 1) observée lors de l'examen microscopique du frottis vaginal après coloration de Gram. Par la suite, le score de Nugent a été développé en 1991 [53]. Ce dernier est réalisé à partir de l'examen microscopique d'un prélèvement vaginal après coloration de Gram. Son objectif est de quantifier les 3 morphotypes bactériens suivants : grands bacilles à Gram positif (*Lactobacillus* spp), petits bacilles à Gram

négatif ou variable (*Bacteroides* spp ou *Gardnerella vaginalis*), et bacilles à Gram négatif incurvées (*Mobiluncus* spp). La présence de chaque morphotype est comptée et leur proportion évaluée avec une échelle de 0 à 10, permettant de conclure à une flore vaginale normale (score ≤ 3), une flore vaginale intermédiaire (score ≤ 6) ou une flore de vaginose bactérienne (score ≥ 7). Dans l'ensemble, le score de Nugent ne fournit pas d'informations sur la composition taxonomique du microbiote vaginal. En plus des morphotypes difficiles à classer dans l'un de ces trois groupes, le score Nugent ne tient pas compte de certaines espèces fortement associées à la vaginose bactérienne telles qu'*Atopobium vaginae* et *Peptostreptococcus* [32]. En plus d'être fastidieuse, cette technique est aussi opérateur-dépendant. Afin de rationaliser le diagnostic de la vaginose bactérienne, des études récentes ont exploré l'utilisation d'approches moléculaires telle que la PCR quantitative en temps réel (qPCR) ciblant des microorganismes comme *Atopobium vaginae* et *Gardnerella vaginalis* [54–56]. Avec une sensibilité de 95% et une spécificité de 99%, cette approche moléculaire apparait comme un outil de diagnostic prometteur pour la vaginose bactérienne, mais malheureusement elles ne sont pas accessibles dans tous les laboratoires.

La vaginose bactérienne peut être asymptomatique, et ce jusqu'à 50% des cas selon les auteurs. le centre de contrôle et de prévention des maladies recommande de traiter toutes les patientes présentant des symptômes avec : 500 mg de métronidazole par voie orale deux fois par jour pendant 7 jours ou 0,75% de gel de métronidazole, un applicateur par voie intravaginale tous les soirs pour 5 nuits ou bien crème vaginale à 2% de clindamycine, un applicateur intravaginal chaque nuit pendant 7 nuits [57, 58]. Le traitement peut engendrer des effets secondaires [59]. En cas d'effets indésirables, le tinidazole a été approuvé et proposé comme thérapie alternative (soit 2 g par jour pendant 2 jours ou 1 g pendant 5 jours, par voie orale) [57, 58]. Pris localement ou oralement, ces agents antimicrobiens ont une efficacité presque similaire avec des taux de guérison d'environ 58 à 92% après un mois [60]. Non seulement

l’antibiothérapie peut avoir un impact négatif sur le microbiote vaginal [63] mais aussi, quel que soit le traitement utilisé les effets n'étaient pas durables favorisant des récidives supérieures à 50% dans les 6 à 12 mois suivant l’arrêt du traitement [61, 62]. C’est dans ce contexte que l’utilisation des probiotiques a été proposé. Divers ovules de *Lactobacillus* sont actuellement disponibles sur le marché. Les souches de *Lactobacillus reuteri* RC-14 et *Lactobacillus rhamnonus* GR-1 sembleraient les plus efficaces. Certes, ces ovules aident à restaurer la flore vaginale, mais des rechutes apparaissent le plus souvent quelques semaines voire mois après l’arrêt du traitement. Résoudre l’énigme de la pathogenèse de cette dysbiose vaginale est fondamental pour le contrôle et la prévention de ce trouble gynéco-obstétrique important.

1 **Bacterial vaginosis: What do we currently know?**

2

3 Khoudia DIOP[1] and Florence FENOLLAR[1]*

4

5 [1]Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection, Marseille,

6 France.

7

8

9 ***Corresponding author** :

10 Pr. Florence Fenollar
11 VITROME, Institut Hospitalo-Universitaire Méditerranée-infection
12 19-21 Boulevard Jean Moulin
13 13005 Marseille, France
14 Phone: + 33 (0) 4 13 73 24 01
15 Fax: + 33 (0) 4 13 73 24 02
16 E-mail address: florence.fenollar@univ-amu.fr
17

18 **Abstract count**: 96

19 **Word count**: 3,665

20 **Figures**: 2

21 **References**: 90

22 Contents

23 **Abstract** 3
24 **1. Introduction** 4
25 **2. Normal Healthy Vaginal Flora** 4
26 **2.1. Composition of normal vaginal flora** 4
27 **2.2. Vaginal microbiota during the different stages of a woman's life** 5
28 **2.3. Variability of vaginal flora according to ethnicity** 6
29 **2.4. Role of vaginal microbiota in woman health** 6
30 **3. Bacterial Vaginosis** 7
31 **3.1. Etiology and pathophysiology (Figure 1)** 7
32 **3.2. Diagnosis** 8
33 *3.2.1. Amsel's criteria* 9
34 *3.2.2. Nugent's score* 9
35 *3.2.3. Other diagnosis tools* 10
36 **3.3. Epidemiology and risk factors** 11
37 *3.3.1. Sexual practices* 11
38 *3.3.2. Other bacterial vaginosis risk factors* 12
39 **3.4. Bacterial vaginosis complications on women's health** 13
40 **3.5. Treatment and management of bacterial vaginosis** 13
41 *3.5.1. Antibiotics therapies* 14
42 *3.5.2. Probiotics therapies* 15
43 **4. Conclusions and Perspectives** 15
44 **Acknowledgments** 16
45 **References** 17

46
47

Abstract

Bacterial vaginosis is a disruption of the vaginal bacterial flora with disappearance of lactobacilli and overgrowth of resident anaerobic vaginal bacteria. Little progress has been made to identify the causal factors, although the pathophysiology of this disorder is understood. The symptoms are recognizable, nevertheless some patients do not exhibit symptoms and the number of associated obstetric and gynecological complications continues to increase. Diagnostic problems continue to dominate clinical practice, although new tools have been introduced. Therapeutic options have also increased, however recurrences remain common and the management of this public health disorder is a major challenge.

Keywords: Vaginal microbiota; *Lactobacillus*; Dysbiosis; Bacterial vaginosis; Sexually transmitted infection, Bacterial vaginosis-associated bacteria.

1. Introduction

The vaginal microbial community is complex and dynamic, consisting of a set of bacteria, typically characterized by abundant lactobacilli, that evolve during the life of the woman depending on age, hormonal estrogen levels, sexual practices and the environment [1, 2]. The vaginal microbiota plays a crucial role in women's health (infection, reproductive and that of their fetuses) [3].

Bacterial vaginosis is a dysbiosis of the vaginal microbiota characterized by a shift from lactobacilli dominance to those of a mixture of various anaerobic bacteria [4, 5]. It is the most common vaginal worldwide disorder in women of childbearing age. Bacterial vaginosis is associated with significant adverse healthcare outcomes, including increased susceptibility to sexually transmitted infections, urogenital infections, pelvic inflammatory disease, and increased risk of abnormal pregnancy [6]. The etiology of bacterial vaginosis is still unknown. Standard antibiotic therapy often fails with an estimated relapse rate of 50% at six months of follow-up [7, 8].

2. Normal Healthy Vaginal Flora

The vaginal ecosystem is colonized from the very first hours of the birth of a girl and remains throughout her life until her death [9]. Women of childbearing age produce about 1 to 4 ml of vaginal fluid containing 10^8 to 10^9 bacterial cells per ml [10].

2.1. Composition of normal vaginal flora

The vaginal flora was first described by German gynecologist Albert Döderlein in 1892, who reported a homogeneous vaginal flora of Gram-positive bacilli in healthy women [11]. They were named "Döderlein's bacilli" and were later identified as members of the *Lactobacillus* genus by Beijerink in 1901 [11]. Under normal condition, 70-90% of vaginal bacterial species in healthy premenopausal women are lactobacilli [12]. As molecular techniques have advanced, our understanding of the diversity and complexity of vaginal

bacterial community has broadened [13]. Among over than 200 *Lactobacillus* species with standing in nomenclature, over 20 species have been found in vaginal flora [14]. The flora is dominated by *Lactobacillus crispatus, Lactobacillus jensenii, Lactobacillus gasseri*, or *Lactobacillus iners* [15].

Lactobacilli produce lactic acid, hydrogen peroxide, and other substances that keep the vagina at about pH 4 and inhibit the growth of other microorganisms. Thus, many other bacteria are present at lower concentrations in healthy vaginal flora such as *Peptostreptococcus, Bacteroides, Corynebacterium, Streptococcus,* and *Peptococcus* [1].

2.2. Vaginal microbiota during the different stages of a woman's life

The colonization of the newborn's vagina begins at birth by contact with the vaginal and intestinal microbiota of the mother during the vaginal delivery [15]. For newborns delivered by cesarean section, the vagina is first populated by bacteria from the cutaneous flora of their mother [16]. At delivery, maternal estrogen contacts the vaginal walls of the baby and during the first month, the baby's vagina is under the influence of these maternal estrogens. During childhood, the girl's vaginal epithelium becomes thinner and glycogen content decrease, the pH is high and the flora is composed of intestinal and cutaneous commensal bacteria [14], with a predominance of anaerobic species [17].

Various physical and hormonal changes occur in the vagina biotope at puberty and menarche. In healthy girls, fluctuations in estrogen hormones lead to thickening of vaginal epithelial cells, glycogen production and a vaginal microbiota containing a large number of lactobacilli which produce lactic acid, thus acidifying the vaginal environment (pH <4.5) and hindering the development of anaerobic bacteria [14, 18].

During pregnancy, the vaginal environment becomes very stable with a decrease of vaginal diversity and enrichment in lactobacilli [9, 19]. This matches with increased

109 production of vaginal secretions occasioned a decrease in the vaginal pH and anaerobic
110 bacteria.

111 In postmenopausal women changes with decreasing levels of circulating estrogen leads
112 to epithelial changes and lactobacilli decrease, then hydrogen peroxide is not produced thus
113 causing an increase of vaginal pH and facilitating an overgrowth of anaerobic bacteria [20].
114 Disturbance of the vaginal microbiota is correlated with the vaginal symptoms of menopause.
115 In addition, hormone replacement therapy can restore the lactobacilli dominance in the
116 vaginal microbiota and can solve these vaginal symptoms [20].

117 **2.3. Variability of vaginal flora according to ethnicity**

118 Vaginal bacterial communities of women of childbearing age may vary between women
119 from different regions, but also between women of different ethnicities and living in the same
120 geographical area [16]. In 2011, a study of Ravel *et al.,* characterizing the vaginal microbiota
121 of asymptomatic North American women with pyrosequencing, showed that vaginal flora of
122 Asian and white American women was dominated by lactobacilli unlike Hispanic and black
123 women, of whom only 60% had a *Lactobacillus*-dominated vaginal flora [21]. In addition,
124 Caucasian and Asian women tend to have high levels of *L. crispatus* and lower *L. iners*
125 compared to African [22]. In another study using 16S rRNA gene sequencing, Fettweis *et al.*,
126 exhibited that vaginal microbiota of European ancestry women was dominated by lactobacilli
127 counter to African American women that present a mixted vaginal community containing
128 among others *Mycoplasma hominis, Aerococcus* and *L. iners* and numerous strictly anaerobes,
129 including Gram-positive anaerobic cocci, bacterial vaginosis associated-bacteria, *Sneathia,*
130 *Prevotella amnii, Megasphaera, Atopobium*, and *Gardnerella vaginalis* [23]. The vaginal pH
131 differs also between racial groups. Black American and Hispanic women had a vaginal pH
132 (4.7 and 5.0, respectively) above the norm (< 4.5) [16].

133 **2.4. Role of vaginal microbiota in woman health**

134 The stability of the vaginal flora prevents the proliferation of commensal
135 microorganisms and the colonization of pathogens, thereby preventing infections [11, 24].
136 Indeed, bacteria form biofilm in the vaginal mucosa and produce antimicrobial compounds,
137 that maintains this health equilibrium, such as hydrogen peroxide (antimicrobial product
138 protecting against deleterious microorganisms), lactic acid (which maintains the normal
139 vaginal pH between 3.5 to 4.5), and bacteriocins (antibiotics which inhibit the growth of
140 harmful microorganisms within the vagina) [11, 24].

141 During the menstrual cycle, the activation of estrogen causes the production of glycogen
142 in the epithelial cells. The fermentation of this glycogen by the lactobacilli and the epithelial
143 cells themselves leads to the production of lactic acid causing an acidification of the vaginal
144 environment which leads in turn to a decrease of the pH [25]. This low vaginal pH impairs the
145 growth of certain vaginal pathogens, but also commensal species other than lactobacilli.

146 *L. crispatus* and *L. jensenii* may produce hydrogen peroxide, an oxidizing agent, toxic
147 for catalase-negative bacteria and also susceptible *in vitro* to inhibit HIV-1 and herpes
148 simplex virus type 2 [26, 27]. The vaginal acids produced can in the presence of viral RNA
149 stimulate the maturation of dendritic cells, the activation of 17 subclasses of T helper
150 lymphocytes, and the production of protective inflammatory cytokines and interferon-γ [28].

151 **3. Bacterial Vaginosis**

152 **3.1. Etiology and pathophysiology (Figure 1)**

153 Formerly known as non-specific vaginitis [29], bacterial vaginosis is characterized by a
154 switch of the vaginal flora composition with a dramatic depletion of lactobacilli from a high
155 overgrowth of obligate or facultative anaerobes previously minority in the vagina [4, 6] such
156 as *Gardnerella vaginalis, Atopobium vaginae, Ureaplasma urealyticum, Mycoplasma*
157 *hominis, Prevotella, Peptoniphilus, Megasphaera, Mobiluncus,* and several fastidious and
158 uncultured bacteria including bacterial vaginosis-associated bacteria (BVAB-1 to 3) [9, 30,

159 31]. The factor triggering this overgrowth of anaerobic bacteria is unknown. It is linked to an
160 alkaline vaginal ecosystem dint of an increase of vaginal pH following the loss of protective
161 effects of lactobacilli.

162 The vaginal flora diversity of patients with bacterial vaginosis was described in 1921 by
163 Schröder [1]. Then, in 1955, Gardner and Dukes have asserted that the etiological agent of
164 bacterial vaginosis was *Haemophilus vaginalis* [32], a Gram-negative rod later renamed
165 *Gardnerella vaginalis* [15, 22]. Some years after, *G. vaginalis* was found in 40% of healthy
166 women and their averment was disputed [33]. In addition to *G. vaginalis,* some members of
167 anaerobic bacteria were highly associated with bacterial vaginosis, this allows to conclude
168 that bacterial vaginosis is a polymicrobial syndrome which does not follow Koch's postulates
169 [1].

170 Bacteria present in the microbiota of bacterial vaginosis form a biofilm in the vaginal
171 epithelium and that secrete a cytotoxin capable of killing the epithelial cells [14]. Besides,
172 anaerobic bacteria produce proteolytic enzymes able to degrade proteins and decarboxylases
173 that convert amino acids. Not degraded, the amines compounds become malodorous (fishy
174 odor: "Whiff test) thank an increase of the pH [34]. Then, the cytotoxicity emanating from the
175 association of organics acids present in the vagina during bacterial vaginosis and bacterial
176 polyamines lead to the production of vaginal discharge owing to the exfoliation of vaginal
177 epithelial cells [35]. In addition, bacteria particularly *G. vaginalis,* cover vaginal epithelial
178 cells causing the formation of "clue-cells", a specific characteristic of bacterial vaginosis [36].

179 **3.2. Diagnosis**

180 Bacterial vaginosis ranges from no symptoms to an increased vaginal discharge with or
181 without fish odor [32]. Its diagnosis is problematical and challenging because of its intricate
182 polymicrobial feature and a wide range of clinical features. The collection of material for
183 diagnosis can be performed during a pelvic exam using a speculum. When there is no reason

184 for a pelvic exam as part of the clinical evaluation, a self-collected vaginal swab may be also
185 provided [37]. The swab may be placed in classical bacterial transport medium or may be
186 spread on a slide and air-dried for subsequent Gram staining [38]. The transport for one or the
187 other can be carried out at ambient temperature or at 4°C. Vaginal culture is inadequate for
188 the diagnosis of bacterial vaginosis and can be misleading as it cannot identify proportions of
189 bacterial species in the vaginal specimen and "uncultivable" bacteria identified with
190 molecular tools [13, 38–40].

191 Two main basic categories of diagnostic strategies for bacterial vaginosis exists the
192 "bedside" method mainly based on real-time clinical criteria (Amsel's criteria) and
193 laboratory-based testing based on the evaluation of morphotypes on Gram staining (Nugent's
194 score). Amsel's criteria and Nugent's score are the most common diagnostic methods used for
195 bacterial vaginosis. Furthermore, Nugent's score is currently considered as the gold standard.

196 *3.2.1. Amsel's criteria*

197 The most common clinical criteria are those of Amsel *et al.* introduced in 1983 [29].
198 The Amsel's criteria require for diagnostic the presence of any three of the four following
199 conditions: (1) an increased homogeneous grey vaginal discharge adhering to the vaginal
200 walls; (2) a pH of vaginal secretions greater than 4.5; (3) a release of an amine odor (fishy
201 smell) after addition of a drop of 10% KOH (potassium hydroxide) solution on a drop of
202 vaginal secretions ("whiff test"); (4) a presence of clue cells (vaginal squamous epithelial
203 cells coated with Gram variable Coccobacilli) demonstrated by microscopic observation of
204 vaginal wet mount preparation **(Figure 2)**. One of the pitfalls of the Amsel's criteria is a
205 failure to diagnose women without any symptoms.

206 *3.2.2. Nugent's score*

207 The Nugent score, developed by Nugent *et al*. in 1991, is based on Gram staining
208 scoring [41]. The presence of the following bacterial morphotypes is evaluated: large Gram-

209 positive rods (*Lactobacillus* morphotypes), small Gram-variable rods (*G. vaginalis*
210 morphotypes), small Gram-negative rods (*Bacteroides* morphotypes), and curved Gram-
211 negative rods (*Mobiluncus* morphotypes). Gram-positive cocci are not part of the scoring
212 system, but their increased presence is not part of normal flora [38].

213 The presence of each morphotype is counted and their proportion evaluated. Thus, data
214 are scored with a scale from 0 to 10. Scores of 0 to 3 are considered to be normal and those of
215 7 to 10 are defined to bacterial vaginosis [41]. Scores comprising between 4 and 6 are
216 assigned to an intermediate flora. Intermediate vaginal flora is mentioned to physicians for
217 patient management based on clinical context. Among the patients with an intermediate flora,
218 some will correspond to bacterial vaginosis and others to a normal flora. An intermediate flora
219 score should be considered by many authors as abnormal given the high risk of transition to
220 bacterial vaginosis. Overall, Nugent score does not provide information on taxonomic
221 composition of the vaginal microbiota. In addition to morphotypes that are difficult to classify
222 in one of these three groups, Nugent Score, overlooks certain species strongly associated with
223 bacterial vaginosis such as *A. vaginae* and *Peptostreptococcus* [7, 42]. Finally, Nugent's score
224 is tedious and technician-dependent.

225 *3.2.3. Other diagnosis tools*

226 In order to overcome the diagnostic problems of BV, recent studies have explored the
227 use of molecular approaches such as clonage, fluorescence *in situ* hybridization and
228 quantitative real-time PCR (qPCR) assay for BVAB [40, 43, 44]. These molecular approaches
229 appear to be a promising diagnostic tool for bacterial vaginosis with a high specificity and
230 sensibility but unluckily they are not accessible in all laboratories.

231 Finally, alternative diagnostic strategies have been attempted such as enzymatic and
232 chromatography techniques that analyze metabolic activity including produced enzymes and
233 fatty acids [12]. Indeed, vaginal secretions from bacterial vaginosis are composed mainly of

234 succinic and acetic acids whereas those in healthy vaginal flora are mostly composed of lactic
235 acid. None are currently available for a diagnosis purpose.

236 **3.3. Epidemiology and risk factors**

237 Bacterial vaginosis may appear at any age but is more prevalent in women of
238 childbearing age. Its prevalence rates vary considerably between geographic regions of the
239 world, within the same country, and even within the same population according to ethnic
240 origin and socioeconomic status. Bacterial vaginosis occurs between 4-75% according to the
241 population studied [42, 45]. Intermediate in the USA, the prevalence of bacterial vaginosis
242 was evaluated to be low in Europe with a maximum (> 20%) in Poland, and Norway [46]. In
243 Africa, the estimated prevalence tended to be high. However, bacterial vaginosis prevalence
244 was lowest in West Africa (6-8% in Burkina Faso and 14.2% in Nigeria) than Southern and
245 eastern Africa with 32.5% in Zimbabwe, 37% Kenya, 38% Botswana, and 68.3% in
246 Mozambique [45–47].

247 *3.3.1. Sexual practices*

248 Although bacterial vaginosis is not a sexually transmitted disease, it is strongly
249 associated with sexual activities and has some characteristics of a sexually transmitted disease
250 [48]. Withal, bacterial vaginosis is diagnosed in post-pubertal women who had never sex but
251 they had a lower prevalence of bacterial vaginosis than those who had sexual experiences
252 [49]. The prevalence varies with the number of sex partners. It was evaluated at 18.8% for
253 sexually inexperienced women, 22.4% for women with one partner during their life and 43.4,
254 and 58% respectively, for women having 2-3 lifetime sex partners and those having ≥ 4-
255 lifetime sex partners [50].

256 In this dynamic, non-commercial sex worker had a lower bacterial vaginal diversity but
257 much richer in *Lactobacillus* species than commercial sex workers [51]. Compared with male
258 partners of healthy women, bacterial vaginosis related bacteria can be found in the penile

259 skin, urethra [30], spermatozoa, and prostatic fluid microbiota [52, 53] of male partners of
260 women with bacterial vaginosis. Furthermore, biofilm fragments have been found in their
261 urine and sperm [54, 55] suggesting that males partners are a reservoir, but also a heterosexual
262 transmission may occur. Nevertheless, there is not a corresponding illness in male partners
263 and use of condoms by males partners prevent acquisition and recrudescence of bacterial
264 vaginosis [56]. Furthermore, since the preputial area of some men hosts bacterial vaginosis -
265 associated microorganisms, therefore, male circumcision may reduce the risk of bacterial
266 vaginosis condition [31].

267 Prevalence rates also depend on the nature of the couple and their sexual practices. In
268 fact, bacterial vaginosis prevalence varies between 10-30% in heterosexual women, on the
269 other hand, it is more important, about 25-50% among women who have sex with women [2,
270 57]. The reasons for this difference in prevalence are not clear, however sexual activities
271 involving the transmission of vaginal fluid increase the risk of bacterial vaginosis acquisition
272 [6]. Several studies have advocated that certain sexual behaviors including non-coital sexual
273 practices like digital and penile penetration, anal and oral intercourses followed by vaginal
274 penetration enhance bacterial vaginosis risk [58]. Whilst, in lesbians, symptomatic female
275 sexual partner, receptive oral sex, and the use and sharing of unwashed sex toys constituted a
276 risk factor of bacterial vaginosis [49]. These observations have led some to think that bacterial
277 vaginosis is not an infection but rather a taxonomic change in vaginal microbiota resulting
278 from translocation of oral [12] or fecal [59] microbiota during non-coital sexual practices.

279 *3.3.2. Other bacterial vaginosis risk factors*

280 Additionally, genital hygiene can also promote the disequilibrium in the vaginal
281 microbiota. A study shows that patients who didn't bathe their vaginal region were more
282 susceptibility to bacterial vaginosis than those who bathed often the vagina, 53.9% and 40.2%
283 of prevalence respectively. Similarly, the prevalence of bacterial vaginosis is higher in

284 patients who did not change their pants frequently than among those who changed it more
285 frequently (57.6% versus 36.9%) [45]. Besides, other sexual sanitary and habits including
286 vaginal douching and washing [60], cigarette smoking [61], some contraceptives methods like
287 dispositive intra-uterine devices [62] and stress [6] may also enhance the risk of developing
288 bacterial vaginosis.

289 **3.4. Bacterial vaginosis complications on women's health**

290 Women suffering from bacterial vaginosis were vulnerable and the presence of BV-
291 related bacteria and/or sexually transmissible microorganisms in bacterial vaginosis
292 microbiota can lead to opportunists' infections. During this imbalance, 10-30% of pregnant
293 women with bacterial vaginosis give birth prematurely, a preterm delivery often accompanied
294 by severe problems of up to 70% worldwide perinatal mortality [47, 63]. At pregnancy,
295 bacterial vaginosis increases in these women the risk of preterm labor, late miscarriage,
296 intrauterine fetal demise, preterm rupture of membranes, amniotic fluid infections,
297 chorioamnionitis, post-abortion and postpartum infections [64–67].

298 In non-pregnant women, firstly, bacteria implicated in bacterial vaginosis can cause
299 cervicitis, endometritis, salpingitis, urinary tract infections [68]. After damage of the cervix,
300 bacteria can migrate from lower to upper genital tract to reach the uterus and fallopian tubes
301 causing illness such as pelvic inflammatory disease [64, 69], post-hysterectomy infections [6],
302 and even cervical cancer or tubal infertility [70, 71]. Likewise, bacterial vaginosis is
303 associated with high increased rates of acquiring herpes simplex virus [72], human
304 immunodeficiency virus [73], papillomavirus [74] and transmission of pathogens such as
305 syphilis, chancroid, gonorrhea, trichomoniasis, and *Chlamydia* [45, 75].

306 **3.5. Treatment and management of bacterial vaginosis**

307 Considering that clinical cure corresponds to the disappearance of all symptoms, the
308 treatment of bacterial vaginosis is currently focused on stopping the proliferation of BV-

309 associated microorganisms and restoring normal vaginal flora [6]. Classically, clinical
310 therapies include the use of antibiotics having a broad activity against anaerobic microbes and
311 protozoa: clindamycin and nitroimidazoles (metronidazole and tinidazole) and/or use of
312 probiotics [1, 76, 77].

313 *3.5.1. Antibiotics therapies*

314 The first line of therapy recommended by world health organization (WHO) is 500 mg
315 oral metronidazole twice a day for a week [76, 78]. However, treatment with metronidazole
316 may cause side effects like gastrointestinal pains, nausea, and vomiting [33]. Other proposed
317 therapeutic regimens include 300 mg oral clindamycin twice a day for a week, 100 mg of
318 intravaginal clindamycin ovule daily during 5 days and an application of 0.75% intravaginal
319 metronidazole gel during 5 days or 2% of intravaginal clindamycin cream at bedtime for a
320 week [78, 79]. Yet, it should be noted that local application of clindamycin may damage
321 latex-based products such as condoms and may also beget a pseudomembranous colitis [33].
322 Recently the use of tinidazole, a drug close to metronidazole, have been approved and
323 proposed an alternative therapy in oral regimen (either 2 g per day for 2 days or 1g for 5 days)
324 once metronidazole and clindamycin are not supported [76, 80].

325 Taken locally or orally, these antimicrobials agents have almost similar efficacy with
326 cure rates around 58 to 92% after 1 month of cure [79]. Nevertheless, this effects was not
327 durable thus fostering a recurrence or re-infection great than 50% within 6-12 months of
328 therapy [77, 81]. The reasons for this high relapse rate still not clear. However, it appears that
329 thanks to the formation of bacterial biofilms, these recommended therapies eradicated
330 temporarily BV-associated microorganisms or that these bacteria are reintroduced in the
331 vagina by their sex partners [31, 82, 83]. Further, the presence of some BV-associated
332 bacteria such as *Peptoniphilus lacrimalis, Megasphaera* type 2 and BVAB-1 to 3 at the

333 beginning of treatment is strongly related to bacterial vaginosis recurrence, thus causing
334 antibiotic failure [84].

335 To disrupt BV-associated biofilms and prevent adverse health effects, some researchers
336 have sought to evaluate the efficacy of other antimicrobial agents such as azithromycin,
337 secnidazole or ornidazole [85, 86]. Only secnidazole have shown an activity similar to that of
338 recommended nitroimidazoles and had also spared lactobacilli, a beneficial characteristic in
339 bacterial vaginosis's treatment [87].

340 *3.5.2. Probiotics therapies*

341 As antibiotic treatments can have a negative impact on the vaginal flora stability,
342 *Lactobacillus* probiotics, alternative and complementary therapy to the antibiotics cure, has
343 been developed to help restore and maintain the healthy vaginal flora [88]. Probiotic is a
344 living microorganism that confer a health benefit on the host when they are administered in
345 appropriate quantity [26]. Actually, only strains *L. reuteri* RC-14 and *L. rhamnonus* GR-1
346 have positive clinical effects [1, 89]. Administrated orally (twice daily) or vaginally (once a
347 week), these probiotics may restore human *Lactobacillus*-dominated microbiota and reduce
348 bacterial vaginosis-recurrence [90]. Nevertheless, they had a slight success in African woman
349 [31]. Thus, the management of bacterial vaginosis urgently requires the implementation of
350 new therapeutic strategies.

351 **4. Conclusions and Perspectives**

352 Taxonomic composition and bacterial proportion of vaginal microbiota are under
353 influence of intrinsic and external factors during the female lifespan. Over the last decades,
354 understanding of the bacterial diversity of this ecosystem was increased by molecular
355 methods. Dominated by lactobacilli that protect against infection, vaginal flora of healthy
356 women is less complex than those of patients afflicted with bacterial vaginosis presenting a
357 diverse microbiota containing numerous obligate anaerobic and uncultivable species. This

358 polymicrobial condition is associated with clinical symptoms relatively uncomplicated that do
359 not occur in all affected women thus complicating the determination of its etiology. The
360 treatment is usually unsuccessful with a high rate of relapse. Future studies that will
361 thoroughly examine the vaginal bacterial community will be needed to cultivate the bacteria
362 associated with bacterial vaginosis and the failure of its treatment, in order to study their
363 antibiotic resistance and to establish more effective alternative therapeutic strategies that
364 reduce bacterial vaginosis symptoms as well as its associated complications. Overall,
365 unlocking the enigma of bacterial vaginosis pathogenesis is key for the prevention and
366 management of this public health condition.
367
368
369
370
371 **Conflict of interest**
372 The authors declare that they have no conflict of interest.
373
374 **Acknowledgments**
375 The authors are grateful to the foundation Mediterranean Infection and the National Research
376 Agency under the program "Investissements d'avenir", reference ANR-10-IAHU-03, for
377 funding this study. The funding source has no other input in this article.

378 **References**

379 1. **Pal K, Roy S, Behera B, Kumar N, Sagiri S, *et al.*** Bacterial vaginosis: Etiology and
380 modalities of treatment-A brief note. *J Pharm Bioallied Sci* 2011;3:496.

381 2. **Bilardi J, Walker S, Mooney-Somers J, Temple-Smith M, McNair R, *et al.***
382 Women's Views and Experiences of the Triggers for Onset of Bacterial Vaginosis and
383 Exacerbating Factors Associated with Recurrence. *PLOS ONE* 2016;11:e0150272.

384 3. **Li J, McCormick J, Bocking A, Reid G**. Importance of Vaginal Microbes in
385 Reproductive Health. *Reprod Sci* 2012;19:235–242.

386 4. **Mårdh PA**. [The definition and epidemiology of bacterial vaginosis]. *Rev Fr Gynecol*
387 *Obstet* 1993;88:195–197.

388 5. **Hay PE**. Bacterial Vaginosis as a Mixed Infection. In: *Polymicrobial Diseases*.
389 Washington (DC): ASM Press. https://www.ncbi.nlm.nih.gov/books/NBK2495/ (2002,
390 accessed 1 March 2018).

391 6. **Marrazzo JM, Hillier SL**. Bacterial Vaginosis. In: *Sexually Transmitted Diseases*.
392 Elsevier. pp. 463–498.

393 7. **Bradshaw CS, Tabrizi SN, Fairley CK, Morton AN, Rudland E, *et al.*** The
394 association of Atopobium vaginae and Gardnerella vaginalis with bacterial vaginosis and
395 recurrence after oral metronidazole therapy. *J Infect Dis* 2006;194:828–836.

396 8. **Bretelle F, Rozenberg P, Pascal A, Favre R, Bohec C, *et al.*** High Atopobium vaginae
397 and Gardnerella vaginalis Vaginal Loads Are Associated With Preterm Birth. *Clin Infect*
398 *Dis* 2015;60:860–867.

399 9. **Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, *et al.*** The composition and
400 stability of the vaginal microbiota of normal pregnant women is different from that of
401 non-pregnant women. *Microbiome* 2014;2:4.

402 10. **Danielsson D, Teigen PK, Moi H**. The genital econiche: focus on microbiota and
403 bacterial vaginosis: Genital econiche. *Ann N Y Acad Sci* 2011;1230:48–58.

404 11. **Lepargneur JP, Rousseau V**. Protective role of the Doderleïn flora. *J Gynecol Obstet*
405 *Biol Reprod* 2002;31:485–494.

406 12. **Africa C, Nel J, Stemmet M**. Anaerobes and Bacterial Vaginosis in Pregnancy:
407 Virulence Factors Contributing to Vaginal Colonisation. *Int J Environ Res Public Health*
408 2014;11:6979–7000.

409 13. **Fredricks DN, Fiedler TL, Marrazzo JM**. Molecular identification of bacteria
410 associated with bacterial vaginosis. *N Engl J Med* 2005;353:1899–1911.

411 14. **Huang B, Fettweis JM, Brooks JP, Jefferson KK, Buck GA**. The Changing
412 Landscape of the Vaginal Microbiome. *Clin Lab Med* 2014;34:747–761.

413 15. **Mendling W**. Vaginal Microbiota. In: Schwiertz A (editor). *Microbiota of the Human*
414 *Body*. Cham: Springer International Publishing. pp. 83–93.

415 16. **Hickey RJ, Zhou X, Pierson JD, Ravel J, Forney LJ**. Understanding vaginal
416 microbiome complexity from an ecological perspective. *Transl Res* 2012;160:267–282.

417 17. **Hill GB, Claire KS, Gutman LT**. Anaerobes predominate among the vaginal
418 microflora of prepubertal girls. *Clin Infect Dis* 1995;20:S269–S270.

419 18. **Hickey RJ, Zhou X, Settles ML, Erb J, Malone K, *et al.*** Vaginal Microbiota of
420 Adolescent Girls Prior to the Onset of Menarche Resemble Those of Reproductive-Age
421 Women. *mBio* 2015;6:e00097-15.

422 19. **Nuriel-Ohayon M, Neuman H, Koren O**. Microbial Changes during Pregnancy, Birth,
423 and Infancy. *Front Microbiol*;7. Epub ahead of print 14 July 2016. DOI:
424 10.3389/fmicb.2016.01031.

425 20. **Muhleisen AL, Herbst-Kralovetz MM**. Menopause and the vaginal microbiome.
426 *Maturitas* 2016;91:42–50.

427 21. **Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, *et al.*** Vaginal microbiome of
428 reproductive-age women. *Proc Natl Acad Sci* 2011;108:4680–4687.

429 22. **Green KA, Zarek SM, Catherino WH**. Gynecologic health and disease in relation to
430 the microbiome of the female reproductive tract. *Fertil Steril* 2015;104:1351–1357.

431 23. **Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, *et al.*** Differences in
432 vaginal microbiome in African American women versus women of European ancestry.
433 *Microbiology* 2014;160:2272–2282.

434 24. **Tachedjian G, O'Hanlon DE, Ravel J**. The implausible "in vivo" role of hydrogen
435 peroxide as an antimicrobial factor produced by vaginal microbiota. *Microbiome*;6.
436 Epub ahead of print 6 February 2018. DOI: 10.1186/s40168-018-0418-3.

437 25. **Nelson TM, Borgogna J-LC, Brotman RM, Ravel J, Walk ST, *et al.*** Vaginal
438 biogenic amines: biomarkers of bacterial vaginosis or precursors to vaginal dysbiosis?
439 *Front Physiol*;6. Epub ahead of print 29 September 2015. DOI:
440 10.3389/fphys.2015.00253.

441 26. **Borges S, Silva J, Teixeira P**. The role of lactobacilli and probiotics in maintaining
442 vaginal health. *Arch Gynecol Obstet* 2014;289:479–489.

443 27. **Aldunate M, Tyssen D, Johnson A, Zakir T, Sonza S, *et al.*** Vaginal concentrations of
444 lactic acid potently inactivate HIV. *J Antimicrob Chemother* 2013;68:2015–2025.

445 28. **Witkin S**. The vaginal microbiome, vaginal anti-microbial defence mechanisms and the
446 clinical challenge of reducing infection-related preterm birth. *BJOG Int J Obstet*
447 *Gynaecol* 2015;122:213–218.

448 29. **Amsel R, Totten PA, Spiegel CA, Chen KCS, Eschenbach D, *et al.*** Nonspecific
449 vaginitis. *Am J Med* 1983;74:14–22.

450 30. **Zozaya M, Ferris MJ, Siren JD, Lillis R, Myers L, *et al.*** Bacterial communities in
451 penile skin, male urethra, and vaginas of heterosexual couples with and without bacterial

452 vaginosis. *Microbiome*;4. Epub ahead of print December 2016. DOI: 10.1186/s40168-
453 016-0161-6.

454 31. **Margolis E, Fredricks DN**. Bacterial Vaginosis-Associated Bacteria. In: *Molecular*
455 *Medical Microbiology*. Elsevier. pp. 1487–1496.

456 32. **Gardner HL, Dukes CD**. Haemophilus vaginalis vaginitis. *Am J Obstet Gynecol*
457 1955;69:962–976.

458 33. **Machado D, Castro J, Palmeira-de-Oliveira A, Martinez-de-Oliveira J, Cerca N**.
459 Bacterial Vaginosis Biofilms: Challenges to Current Therapies and Emerging Solutions.
460 *Front Microbiol*;6. Epub ahead of print 20 January 2016. DOI:
461 10.3389/fmicb.2015.01528.

462 34. **González Pedraza Avilés A, Ortíz Zaragoza MC, Irigoyen Coria A**. Bacterial
463 vaginosis a 'broad overview'. *Rev Latinoam Microbiol* 1999;41:25–34.

464 35. **Sobel JD**. 53 - Vaginitis, Vulvitis, Cervicitis and Cutaneous Vulval Lesions. In: Cohen
465 J, Powderly WG, Opal SM (editors). *Infectious Diseases (Fourth Edition)*. Elsevier. pp.
466 483-491.e1.

467 36. **Khan K, Shah R, Gautam M, Patil S**. Clue cells. *Indian J Sex Transm Dis*
468 2007;28:108–108.

469 37. **Menard J-P, Fenollar F, Raoult D, Boubli L, Bretelle F**. Self-collected vaginal swabs
470 for the quantitative real-time polymerase chain reaction assay of Atopobium vaginae and
471 Gardnerella vaginalis and the diagnosis of bacterial vaginosis. *Eur J Clin Microbiol*
472 *Infect Dis Off Publ Eur Soc Clin Microbiol* 2012;31:513–518.

473 38. **Money D**. The laboratory diagnosis of bacterial vaginosis. *Can J Infect Dis Med*
474 *Microbiol* 2005;16:77–79.

475 39. **Srinivasan S, Fredricks DN**. The Human Vaginal Bacterial Biota and Bacterial
476 Vaginosis. *Interdiscip Perspect Infect Dis* 2008;2008:1–22.

477 40. **Shipitsyna E, Roos A, Datcu R, Hallén A, Fredlund H, *et al.*** Composition of the
478 Vaginal Microbiota in Women of Reproductive Age – Sensitive and Specific Molecular
479 Diagnosis of Bacterial Vaginosis Is Possible? *PLoS ONE* 2013;8:e60670.

480 41. **Nugent RP, Krohn MA, Hillier SL**. Reliability of diagnosing bacterial vaginosis is
481 improved by a standardized method of gram stain interpretation. *J Clin Microbiol*
482 1991;29:297–301.

483 42. **Onderdonk AB, Delaney ML, Fichorova RN**. The Human Microbiome during
484 Bacterial Vaginosis. *Clin Microbiol Rev* 2016;29:223–238.

485 43. **Menard J-P, Fenollar F, Henry M, Bretelle F, Raoult D**. Molecular Quantification of
486 Gardnerella vaginalis and Atopobium vaginae Loads to Predict Bacterial Vaginosis. *Clin*
487 *Infect Dis* 2008;47:33–43.

488 44. **Datcu R, Gesink D, Mulvad G, Montgomery-Andersen R, Rink E, *et al.*** Vaginal
489 microbiome in women from Greenland assessed by microscopy and quantitative PCR.
490 *BMC Infect Dis* 2013;13:480.

491 45. **Bitew A, Abebaw Y, Bekele D, Mihret A**. Prevalence of Bacterial Vaginosis and
492 Associated Risk Factors among Women Complaining of Genital Tract Infection. *Int J*
493 *Microbiol* 2017;2017:1–8.

494 46. **Kenyon C, Colebunders R, Crucitti T**. The global epidemiology of bacterial vaginosis:
495 a systematic review. *Am J Obstet Gynecol* 2013;209:505–523.

496 47. **Afolabi BB, Moses OE, Oduyebo OO**. Bacterial Vaginosis and Pregnancy Outcome in
497 Lagos, Nigeria. *Open Forum Infect Dis* 2016;3:ofw030.

498 48. **Sobel JD**. Bacterial Vaginosis. *Annu Rev Med*;10.

499 49. **Cherpes TL, Hillier SL, Meyn LA, Busch JL, Krohn MA**. A Delicate Balance: Risk
500 Factors for Acquisition of Bacterial Vaginosis Include Sexual Activity, Absence of
501 Hydrogen Peroxide-Producing Lactobacilli, Black Race, and Positive Herpes Simplex
502 Virus Type 2 Serology: *Sex Transm Dis* 2008;35:78–83.

503 50. **Koumans EH, Sternberg M, Bruce C, McQuillan G, Kendrick J, *et al.*** The
504 Prevalence of Bacterial Vaginosis in the United States, 2001-2004; Associations With
505 Symptoms, Sexual Behaviors, and Reproductive Health: *Sex Transm Dis* 2007;34:864–
506 869.

507 51. **Wessels JM, Lajoie J, Vitali D, Omollo K, Kimani J, *et al.*** Association of high-risk
508 sexual behaviour with diversity of the vaginal microbiota and abundance of
509 Lactobacillus. *PLOS ONE* 2017;12:e0187612.

510 52. **Hou D, Zhou X, Zhong X, Settles ML, Herring J, *et al.*** Microbiota of the seminal
511 fluid from healthy and infertile men. *Fertil Steril* 2013;100:1261–1269.

512 53. **Gallo MF, Warner L, King CC, Sobel JD, Klein RS, *et al.*** Association between
513 semen exposure and incident bacterial vaginosis. *Infect Dis Obstet Gynecol*
514 2011;2011:842652.

515 54. **Swidsinski A, Doerffel Y, Loening-Baucke V, Swidsinski S, Verstraelen H, *et al.***
516 Gardnerella biofilm involves females and males and is transmitted sexually. *Gynecol*
517 *Obstet Invest* 2010;70:256–263.

518 55. **Swidsinski A, Dörffel Y, Loening-Baucke V, Mendling W, Verstraelen H, *et al.***
519 Desquamated epithelial cells covered with a polymicrobial biofilm typical for bacterial
520 vaginosis are present in randomly selected cryopreserved donor semen. *FEMS Immunol*
521 *Med Microbiol* 2010;59:399–404.

522 56. **Verstraelen H, Verhelst R, Vaneechoutte M, Temmerman M**. The epidemiology of
523 bacterial vaginosis in relation to sexual behaviour. *BMC Infect Dis* 2010;10:81.

524 57. **Forcey DS, Vodstrcil LA, Hocking JS, Fairley CK, Law M, *et al.*** Factors associated
525 with bacterial vaginosis among women who have sex with women: a systematic review.
526 *PloS One* 2015;10:e0141905.

527 58. **Kenyon CR, Osbak K**. Recent progress in understanding the epidemiology of bacterial
528 vaginosis: *Curr Opin Obstet Gynecol* 2014;26:448–454.

529 59. **Fenollar F, Raoult D**. Does Bacterial Vaginosis Result From Fecal Transplantation? *J*
530 *Infect Dis* 2016;214:1784–1784.

531 60. **Aslan E, Bechelaghem N**. To 'douche' or not to 'douche': hygiene habits may have
532 detrimental effects on vaginal microbiota. *J Obstet Gynaecol* 2018;1–4.

533 61. **Nelson TM, Borgogna JC, Michalek RD, Roberts DW, Rath JM, *et al.*** Cigarette
534 smoking is associated with an altered vaginal tract metabolomic profile. *Sci Rep*;8. Epub
535 ahead of print December 2018. DOI: 10.1038/s41598-017-14943-3.

536 62. **Achilles SL, Austin MN, Meyn LA, Mhlanga F, Chirenje ZM, *et al.*** Impact of
537 contraceptive initiation on vaginal microbiota. *Am J Obstet Gynecol*. Epub ahead of print
538 2 March 2018. DOI: 10.1016/j.ajog.2018.02.017.

539 63. **Svare J, Schmidt H, Hansen B, Lose G**. Bacterial vaginosis in a cohort of Danish
540 pregnant women: prevalence and relationship with preterm delivery, low birthweight and
541 perinatal infections. *BJOG Int J Obstet Gynaecol* 2006;113:1419–1425.

542 64. **Wilson BA, Thomas SM, Ho M**. The Human Vaginal Microbiome. In: *Metagenomics*
543 *of the Human Body*. Springer, New York, NY. pp. 91–115.

544 65. **Nelson DB, Hanlon AL, Wu G, Liu C, Fredricks DN**. First Trimester Levels of BV-
545 Associated Bacteria and Risk of Miscarriage Among Women Early in Pregnancy.
546 *Matern Child Health J* 2015;19:2682–2687.

547 66. **Kairys N, Garg M**. Vaginosis, Bacterial. In: *StatPearls*. Treasure Island (FL):
548 StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK459216/ (2017, accessed
549 18 January 2018).

550 67. **Brown RG, Marchesi JR, Lee YS, Smith A, Lehne B, *et al.*** Vaginal dysbiosis
551 increases risk of preterm fetal membrane rupture, neonatal sepsis and is exacerbated by
552 erythromycin. *BMC Med* 2018;16:9.

553 68. **Georgijević A, Cjukić-Ivancević S, Bujko M**. [Bacterial vaginosis. Epidemiology and
554 risk factors]. *Srp Arh Celok Lek* 2000;128:29–33.

555 69. **Sharma H, Tal R, Clark NA, Segars JH**. Microbiota and Pelvic Inflammatory Disease.
556 *Semin Reprod Med* 2014;32:43–49.

557 70. **van Oostrum N, De Sutter P, Meys J, Verstraelen H**. Risks associated with bacterial
558 vaginosis in infertility patients: a systematic review and meta-analysis. *Hum Reprod*
559 2013;28:1809–1815.

560 71. **Babu G, Singaravelu BG, Srikumar R, Reddy SV, Kokan A**. Comparative Study on
561 the Vaginal Flora and Incidence of Asymptomatic Vaginosis among Healthy Women
562 and in Women with Infertility Problems of Reproductive Age. *J Clin Diagn Res JCDR*
563 2017;11:DC18–DC22.

564 72. **Nardis C, Mosca L, Mastromarino P**. Vaginal microbiota and viral sexually
565 transmitted diseases. *Ann Ig* 2013;25:443–56.

566 73. **McClelland RS, Lingappa JR, Srinivasan S, Kinuthia J, John-Stewart GC, *et al.***
567 Evaluation of the association between the concentrations of key vaginal bacteria and the
568 increased risk of HIV acquisition in African women from five cohorts: a nested case-
569 control study. *Lancet Infect Dis*. Epub ahead of print 25 January 2018. DOI:
570 10.1016/S1473-3099(18)30058-6.

571 74. **Kero K, Rautava J, Syrjänen K, Grenman S, Syrjänen S**. Association of
572 asymptomatic bacterial vaginosis with persistence of female genital human
573 papillomavirus infection. *Eur J Clin Microbiol Infect Dis* 2017;36:2215–2219.

574 75. **Brotman RM**. Vaginal microbiome and sexually transmitted infections: an
575 epidemiologic perspective. *J Clin Invest* 2011;121:4610–4617.

576 76. Bacterial Vaginosis - 2015 STD Treatment Guidelines.
577 https://www.cdc.gov/std/tg2015/bv.htm (accessed 15 April 2018).

578 77. **Bradshaw CS, Sobel JD**. Current Treatment of Bacterial Vaginosis—Limitations and
579 Need for Innovation. *J Infect Dis* 2016;214:S14–S20.

580 78. **World Health Organization (ed)**. *Guidelines for the management of sexually*
581 *transmitted infections*. Geneva, Switzerland: World Health Organization; 2003.

582 79. **Donders GG, Zodzika J, Rezeberga D**. Treatment of bacterial vaginosis: what we have
583 and what we miss. *Expert Opin Pharmacother* 2014;15:645–657.

584 80. **Dickey LJ, Nailor MD, Sobel JD**. Guidelines for the treatment of bacterial vaginosis:
585 focus on tinidazole. *Ther Clin Risk Manag* 2009;5:485–489.

586 81. **Bilardi J, Walker S, McNair R, Mooney-Somers J, Temple-Smith M, *et al.***
587 Women's Management of Recurrent Bacterial Vaginosis and Experiences of Clinical
588 Care: A Qualitative Study. *PLOS ONE* 2016;11:e0151794.

589 82. **Marrazzo JM, Fiedler TL, Srinivasan S, Thomas KK, Liu C, *et al.*** Extravaginal
590 Reservoirs of Vaginal Bacteria as Risk Factors for Incident Bacterial Vaginosis. *J Infect*
591 *Dis* 2012;205:1580–1588.

592 83. **Bradshaw CS, Brotman RM**. Making inroads into improving treatment of bacterial
593 vaginosis – striving for long-term cure. *BMC Infect Dis*;15. Epub ahead of print
594 December 2015. DOI: 10.1186/s12879-015-1027-4.

595 84. **Marrazzo JM, Thomas KK, Fiedler TL, Ringwood K, Fredricks DN**. Relationship of
596 Specific Vaginal Bacteria and Bacterial Vaginosis Treatment Failure in Women Who
597 Have Sex with Women: A Cohort Study. *Ann Intern Med* 2008;149:20–28.

598 85. **Schwebke JR, Desmond RA**. A randomized trial of the duration of therapy with
599 metronidazole plus or minus azithromycin for treatment of symptomatic bacterial
600 vaginosis. *Clin Infect Dis Off Publ Infect Dis Soc Am* 2007;44:213–219.

601 86. **Thulkar J, Kriplani A, Agarwal N**. A comparative study of oral single dose of
602 metronidazole, tinidazole, secnidazole and ornidazole in bacterial vaginosis. *Indian J*
603 *Pharmacol* 2012;44:243–245.

604 87. **Petrina MAB, Cosentino LA, Rabe LK, Hillier SL**. Susceptibility of bacterial
605 vaginosis (BV)-associated bacteria to secnidazole compared to metronidazole, tinidazole
606 and clindamycin. *Anaerobe* 2017;47:115–119.

607 88. **Bradshaw CS, Pirotta M, De Guingand D, Hocking JS, Morton AN, *et al.*** Efficacy
608 of Oral Metronidazole with Vaginal Clindamycin or Vaginal Probiotic for Bacterial
609 Vaginosis: Randomised Placebo-Controlled Double-Blind Trial. *PLoS ONE*
610 2012;7:e34540.

611 89. **Ouarabi L, Chait YA, Seddik HA, Drider D, Bendali F**. Newly Isolated Lactobacilli
612 strains from Algerian Human Vaginal Microbiota: Lactobacillus fermentum Strains
613 Relevant Probiotic's Candidates. *Probiotics Antimicrob Proteins*. Epub ahead of print 28
614 November 2017. DOI: 10.1007/s12602-017-9360-0.

615 90. **Daliri EB-M, Lee BH**. New perspectives on probiotics in health and disease. *Food Sci*
616 *Hum Wellness* 2015;4:56–65.

618

619

620

621

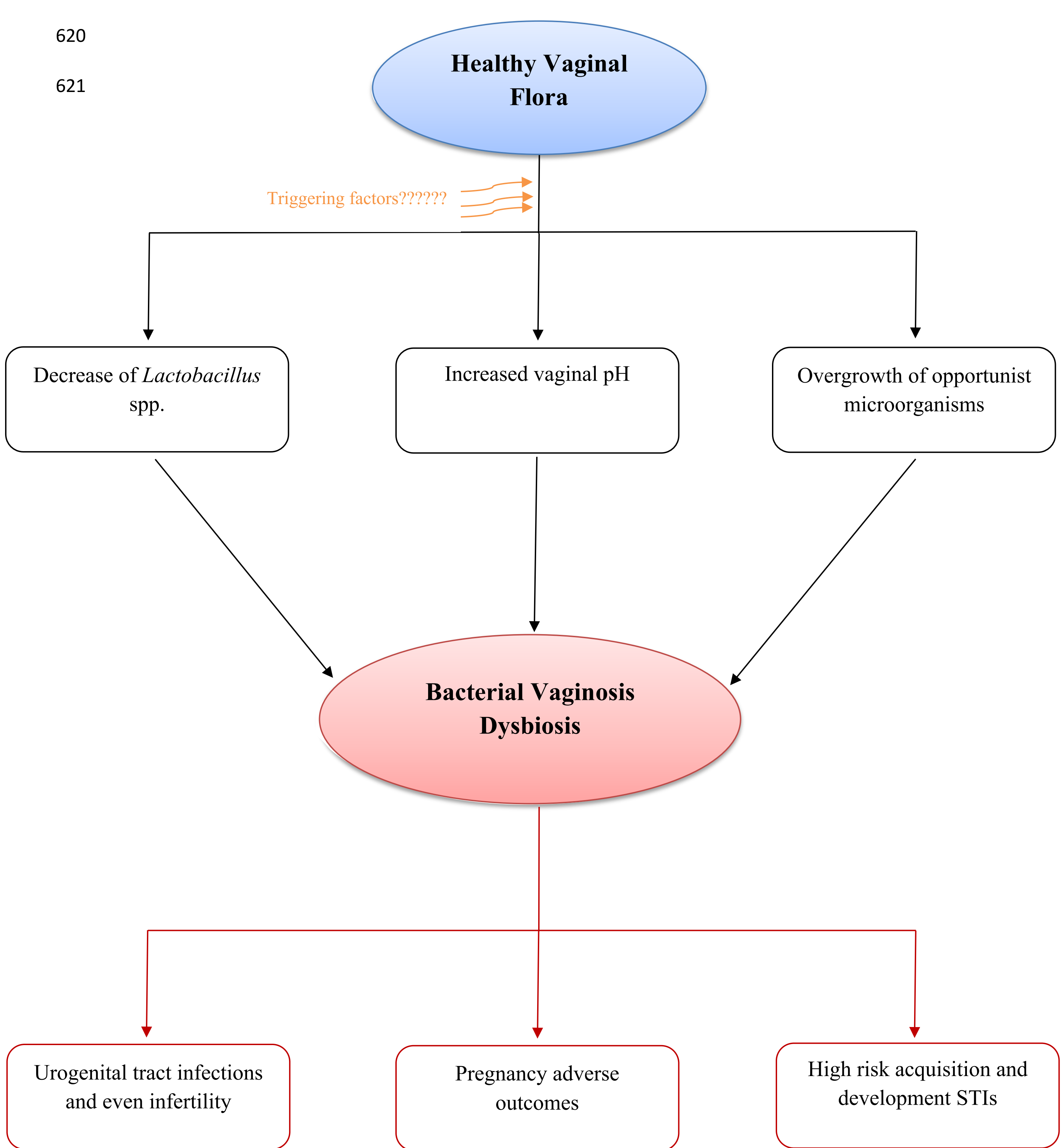
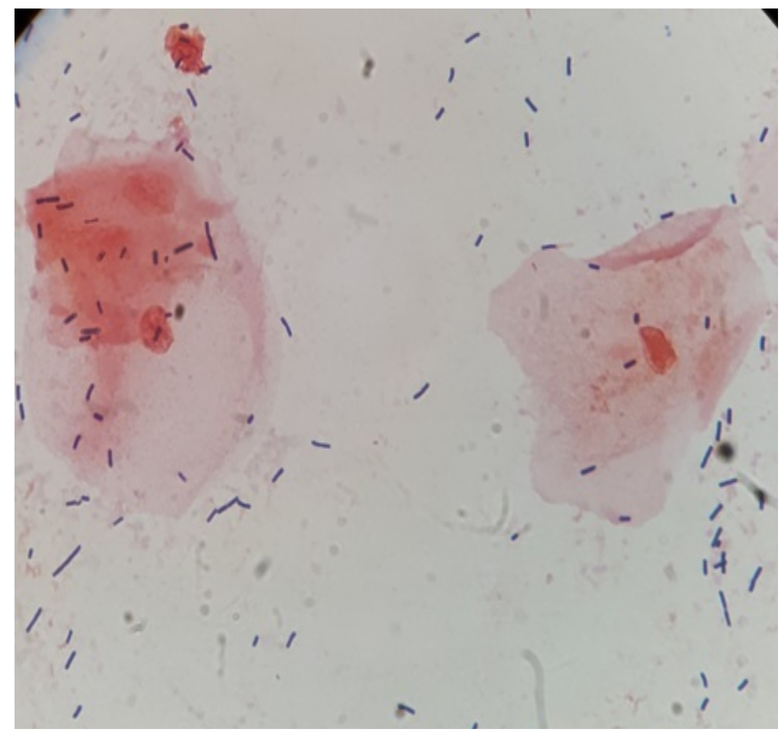
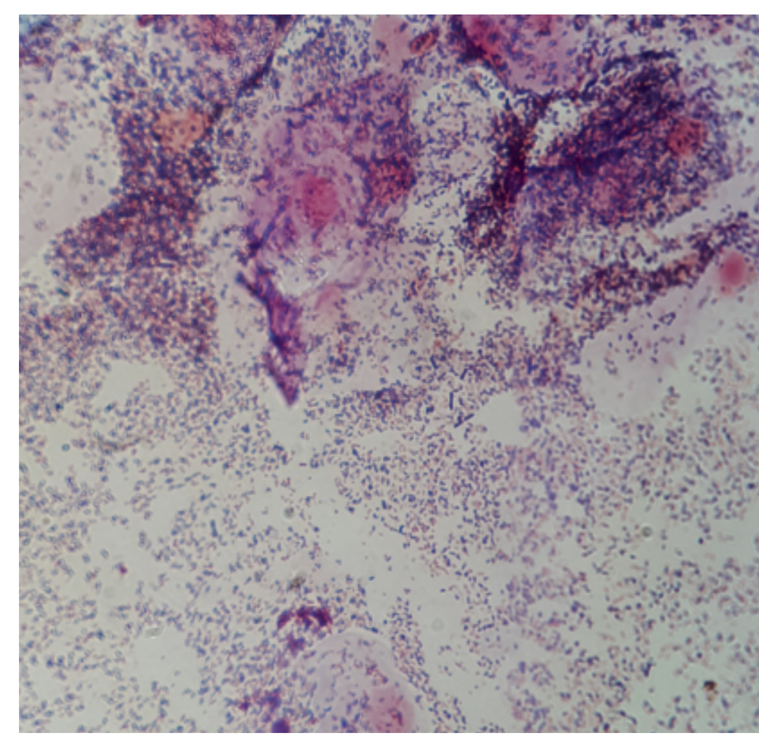


Figure 1. Bacterial vaginosis: Risks factors and impacts

622





623

624 **A- Normal Flora** **B- Bacterial vaginosis Flora and « Clue-cells »**

625

626 **Figure 2 : Gram-stain vaginal smear**

Article 2:

Exhaustive Repertoire of Human Vaginal Microbiota (Revue).

Khoudia Diop, Jean-Charles Dufour, Anthony Levasseur and Florence Fenollar

Under review in Human Microbiome Journal

Avant-propos

Les données sur le microbiote vaginal se sont accumulées ces dernières années. Les flores vaginales de sujets sains ainsi que celles de patientes présentant une vaginose bactérienne ont été étudiées. Les techniques de culture ont permis d'isoler et de décrire de nombreuses espèces bactériennes, tandis que les méthodes moléculaires ont mis en évidence les limites de la culture en montrant que le tractus vaginal était un écosystème complexe contenant de nombreuses bactéries non cultivées jusqu'à présent ou difficiles à identifier. Afin de dresser le répertoire exhaustif des bactéries de la flore vaginale, nous avons effectué une revue de la littérature scientifique en élaborant trois requêtes avec les termes MeSH utilisées pour indexer les publications s'intéressant à ladite flore. Avec ces requêtes, une recherche interrogeant la base de données PubMed a été effectuée pour chacune des 2.776 bactéries connues du microbiote humain en utilisant un programme informatique. Les articles trouvés avec le programme ont été traités et analysés manuellement afin de confirmer la présence ou non de chacune de ces bactéries dans le tractus vaginal. Lorsque cela était nécessaire, les données supplémentaires et la liste des références des articles retrouvés étaient également examinées.

Ainsi quelle que soit la technique de caractérisation utilisée, culture et/ ou moléculaire, nous avons recensé 581 bactéries dont la présence a été rapportée dans la flore vaginale humaine. Ces 581 bactéries sont réparties dans 207 genres, 96 familles et 10 phyla dont 227 *Firmicutes* (39,1%), 150 Protéobactéries (25,8%), 101 Actinobactéries (17,4%) et 74 *Bacteroidetes* (12,7%). Classé par genre, les *Lactobacillus* étaient les plus représentés avec 36 espèces différentes, suivi de *Corynebacterium,* de *Prevotella* (30 bactéries chacun) et de *Streptococcus* (n=28). Du point de vue de leur métabolisme, seulement 181 espèces sont strictement anaérobies (31%). Près de la moitié de ces bactéries sont des anaérobies (47,5%) et appartiennent au phylum *Firmicutes* (86/181) et 33,1% au phylum *Bacteroidetes* (60/181).

Nous nous sommes aussi intéressés aux outils utilisés pour caractériser la flore vaginale et à l'implication des bactéries dans la vaginose bactérienne, condition pathologique correspondant à une dysbiose vaginale. Sur les 581 bactéries détectées dans le microbiote vaginal, seules 285 espèces (49%) étaient identifiées à l'aide des méthodes de culture, 459 (79%) étaient détectés par des techniques moléculaires et 163 ont été identifiées en utilisant à la fois ces deux techniques. Ce répertoire est essentiel et représente le point de départ d'un projet visant à cartographier le microbiote vaginal humain. Une caractérisation complète du microbiote vaginal nécessite donc à la fois l'utilisation des techniques moléculaires et des méthodes de culture.

Manuscript Details

Manuscript number	HUMIC_2018_27
Title	Exhaustive Repertoire of Human Vaginal Microbiota
Article type	Review article

Abstract

Bacteria that colonize the vaginal microbiota of women play an important role in health and homeostasis. Disruption of the proportion of bacteria predisposes to dysbiosis like bacterial vaginosis or severe gynecological conditions such as preterm birth, pelvic inflammatory disease and also sexually transmitted diseases. Knowledge about normal and abnormal vaginal microbiota has become a little clearer in recent years. Culture techniques have made it possible to isolate and describe many bacterial species, whereas molecular methods have highlighted the limits of culture by showing that the vagina was a complex ecosystem containing a wide range of non-cultured or difficult-to-identify bacteria. Based on an exhaustive review of the scientific literature, we built the repertoire of all the bacteria found using culture-based and/or independent methods on the human vagina. So, whether they are valid or not, we inventoried 581 bacteria identified in the human vagina distributed into 10 taxa, mainly in the phyla of Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria with 207 distinct genera classified in 96 different families. This repertoire is essential for microbiologists and clinicians and represents the starting point for a Vaginal Microbiome Project. Such a project aimed to map the human vaginal microbiota to better understand the dysbioses or infections caused by its imbalance in order to offer more appropriate treatments.

Keywords	Bacterial vaginosis; Culture-based methods; Dysbiosis; Molecular techniques; Repertoire; Vaginal microbiota.
Manuscript category	Describing microbes from the microbiota
Manuscript region of origin	Europe

Corresponding Author	Florence Fenollar
Corresponding Author's Institution	URMITE
Order of Authors	Khoudia DIOP, Jean Charles DUFOUR, Anthony LEVASSEUR, Florence Fenollar
Suggested reviewers	Gilbert Greub, Max Maurin, Raymond Ruimy

Submission Files Included in this PDF

File Name [File Type]

cover letter.docx [Cover Letter]

Vaginal Microbiome- Manuscript.docx [Manuscript File]

Figure 1.pdf [Figure]

Figure 2.pdf [Figure]

Figure 3.pdf [Figure]

Table 1.docx [Table]

coi_disclosure.pdf [Conflict of Interest]

Submission Files Not Included in this PDF

File Name [File Type]

Supplementary Table S1.xlsx [Table]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Exhaustive Repertoire of Human Vaginal Microbiota

Khoudia DIOP[1], Jean-Charles DUFOUR[2], Anthony LEVASSEUR[3], Florence FENOLLAR[1]*

[1]Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection, Marseille, France.

[2]Aix Marseille Univ, IRD, APHM, INSERM, SESSTIM, Hop Timone, BioSTIC, Marseille, France.

[3]Aix Marseille Univ, IRD, AP-HM, MEPHI, IHU-Méditerranée Infection, Marseille, France.

***Corresponding author** :
Pr. Florence Fenollar
VITROME, Institut hospitalo-universitaire Méditerranée-infection
19-21 Boulevard Jean Moulin, 13005 Marseille, France
Phone: + 33 (0) 4 13 73 24 01
Fax: + 33 (0) 4 13 73 24 02
E-mail address: florence.fenollar@univ-amu.fr

Abstract count: 211

Word count: 2,937

21 Contents

22

23 **Abstract**.....3
24 **1. Introduction**.....4
25 **2. Methodology**.....5
26 **3. Characterization tools of vaginal flora**.....6
27 **4. Repertoire of bacterial species detected in the vagina**.....7
28 **5. Vaginal bacterial diversity: culture-based versus molecular techniques**.....9
29 **6. Vaginal bacterial microbiota in normal and abnormal conditions**.....11
30 *6.1. Normal vaginal bacterial flora*.....11
31 *6.2. Abnormal vaginal bacterial microbiota: the case of bacterial vaginosis*.....12
32 **7. Conclusions**.....14
33 **Authors' contributions**.....15
34 **Conflict of interest**.....15
35 **Acknowledgments**.....15
36 **References**.....16

37

38

39 **Abstract**

40 Bacteria that colonize the vaginal microbiota of women play an important role in health and
41 homeostasis. Disruption of the proportion of bacteria predisposes to dysbiosis like bacterial
42 vaginosis or severe gynecological conditions such as preterm birth, pelvic inflammatory disease
43 and also sexually transmitted diseases. Knowledge about normal and abnormal vaginal
44 microbiota has become a little clearer in recent years. Culture techniques have made it possible to
45 isolate and describe many bacterial species, whereas molecular methods have highlighted the
46 limits of culture by showing that the vagina was a complex ecosystem containing a wide range of
47 non-cultured or difficult-to-identify bacteria. Based on an exhaustive review of the scientific
48 literature, we built the repertoire of all the bacteria found using culture-based and/or independent
49 methods on the human vagina. So, whether they are valid or not, we inventoried 581 bacteria
50 identified in the human vagina distributed into 10 taxa, mainly in the phyla of *Actinobacteria,*
51 *Bacteroidetes, Firmicutes*, and *Proteobacteria* with 206 distinct genera classified in 96 different
52 families. This repertoire is essential for microbiologists and clinicians and represents the starting
53 point for a Vaginal Microbiome Project such a project aimed to map the human vaginal
54 microbiota, to better understand the dysbioses or infections caused by its imbalance in order to
55 offer more appropriate treatments.

56

57 **Keywords:** Bacterial vaginosis; Culture-based methods; Dysbiosis; Molecular techniques;
58 Repertoire; Vaginal microbiota.

59 **1. Introduction**

60 Microbiota associated with the human body (skin, mucosal membranes of the respiratory
61 airways, oral cavity, gastrointestinal, urinary, and genital tracts) has a considerable influence on
62 human development, physiology, and immunity [1, 2]. It is estimated that the number of
63 microorganisms in the human microbiome are teen times higher than nucleated cells [3].
64 Members of the microbial communities associated with humans interact between them and their
65 host to form a stable ecosystem that responds to disturbances [4]. This mutualistic relationship
66 constitutes the first line of defense by inhibiting and preventing the growth of pathogens [4].
67 Thus, to characterize the normal human microbiota, various body samples including skin, nose,
68 mouth, gastrointestinal tract, and vagina from healthy individuals were analyzed [3].

69 The vaginal microbiome harbors diverse communities of microorganisms, known as
70 vaginal flora which has an important impact on women's health as well as that of their newborns
71 [5]. Bacteria dominate largely vaginal microbiome. A woman in childbearing age produces
72 approximately 1 to 4 ml of vaginal fluid that contains 10^6 to 10^8 bacterial cells per ml [6]. These
73 last decades, studies on the exploration of the vaginal microbiota have increased and the advances
74 in technology, including molecular techniques as well as new OMICS strategies, have
75 demonstrated its involvement in reproductive health [1, 7–12]. The composition of the vaginal
76 microbiota depends on age, menstruations, hormonal fluctuations, sexual behaviors, and also the
77 use of drugs such as probiotics and antibiotics causing its imbalance [13–16]. As part of the
78 human microbiome project, the study of the vaginal microbiome has shown a relationship
79 between bacteria present in the vagina and diseases. The imbalance in the composition of the
80 vaginal microbiota can lead to dysbiosis such as bacterial vaginosis [16, 17]. Thus, the
81 knowledge of vaginal microbiota composition is required to better understand this vaginal
82 condition but also the host-microbiota interactions.

83 In addition to the microbiota constituents, fungal communities (mycobiome) [18] and viral
84 populations (virome) [19] are also an important part of the vaginal microbiome and have
85 relationships with vaginal bacterial components. These underestimated microbiomes play a role
86 in health and diseases such as candidiasis due to an overgrowth of *Candida albicans* [18] and
87 preterm birth caused by a higher viral vaginal diversity [19].

88 This review focuses on and exhaustively inventories bacteria present in the human vaginal
89 microbiome. It is the starting point for a Vaginal Microbiome Project aiming to characterize as
90 fully possible the human vaginal microbiota of normal and bacterial vaginosis floras, to better
91 understand this dysbiosis and better manage this public health problem.

92 ## 2. Methodology

93 To establish the repertoire of bacteria in the vaginal flora, a search in the PubMed/Medline
94 database were conducted using MeSH terms and text-words used to index articles on the subject
95 of vaginal flora. Three query patterns (QP1, QP2, and QP3) were elaborated with these MeSH
96 terms and text-words (Table 1) and using these query patterns, a computer program querying
97 PubMed/Medline for each of the 2,776 bacteria isolated from the human microbiome [20] was
98 developed. This program is a simple java-based application using freely available E-utilities APIs
99 (https://www.ncbi.nlm.nih.gov/home/develop/api/). It takes as input a species names list and
100 programmatically querying NCBI taxonomy and PubMed/Medline for each name of the list in
101 order to record in a comprehensible csv file, easy to consult, bibliographic records corresponding
102 to query patterns. The literature search was performed on May 3, 2018. The bibliographic records
103 retrieved by the program were examined and analyzed manually to confirm the identification of
104 the bacteria in the vagina. All available titles and abstracts were reviewed. If applicable, and if
105 the bacterium was not found there, the entire text was recovered as well as supplementary data.
106 The reference lists of retrieved articles were also examined for similar documents that might be

107 relevant. Data on bacteria detected, detection method, and their involvement in pathogenic
108 conditions were collected and summarized. All human studies on the vaginal microbiota were
109 reviewed, but not animal studies. Listed species were ranked at phylum, family, and genus levels
110 using the website www.ncbi.nlm.nih.gov/taxonomy.

111 ## 3. Characterization tools of vaginal flora

112 Historically, the vaginal flora was first studied by light microscopy and culture-based-
113 methods [9, 21]. Nevertheless, recently, the limitations of these culture methods have been
114 pointed out by molecular methods that identified numbers of strictly anaerobic and uncultured
115 bacteria [22–24]. These conventional methods of cultivation only managed to determine 20% of
116 the bacteria present in the vaginal ecosystem because some species are difficult to cultivate or to
117 identify [7]. Currently, most vaginal microflora studies are based on the sequencing of the 16S
118 rRNA sequence, a preserved region that lies in all bacteria with either universal or specific
119 primers and metagenomics [7, 22, 25, 26]. However, other studies have used other molecular
120 methods such as quantitative PCR (qPCR) [27, 28] and fluorescence in situ hybridization (FISH)
121 [29] that allow, in addition, to identify and quantify the number of bacterial cells. Thus, these
122 molecular methods have broadened our awareness about the complexity of the vaginal
123 microbiota. However, as observed for culture, they present a certain number of disadvantages.
124 They tend to miss the minority species and since the diversity of the vaginal flora is poorly
125 represented in the databases compared to other ecosystems such as the human gut, some species
126 are not identified [7]. Most of the metagenomics studies target the V3-V4 region of the 16S
127 rRNA gene [30]. The amplification of different variable regions of the 16S rRNA gene and
128 estimation of bacterial diversity hamper the comparison between metagenomics studies of
129 bacterial microbiota [30, 31]. Moreover, due to its intragenomic heterogenicity located mainly at
130 positions V1 and V6, when applying technics based on the 16S rRNA gene an under- or

131 overestimation of bacterial diversity can occur by grouping similar ribotypes or multiple signals
132 for a single organism, respectively [31, 32]. To reduce or compensate these estimation biases
133 (over as well as under) in taxonomic diversity of the bacterial microbiota, it is very important to
134 use the bioinformatics tools properly.

135 Despite their limitations, culture methods promoted the isolation of microorganisms present
136 in low concentrations, allowing then to study their characteristics, sequence their full genome,
137 and analyze their pathogenesis and their virulence. Therefore, for an exhaustive exploration of the
138 vaginal microbiota diversity, the return of culture methods has required a novel strategy highly
139 complementary to metagenomics called microbial culturomics. This strategy employs high-
140 throughput culture conditions with a rapid bacterial identification by matrix-assisted laser
141 desorption/ionization–time of flight (MALDI–TOF) mass spectrometry [33–35]. In addition to
142 innovating culture-based methods and bacterial identification, a new species description concept
143 called taxonogenomics has been developed to better characterize and describe bacterial species
144 [36, 37]. Taxonogenomics combines classic bacterial description and genotypic characteristics
145 such as DNA-DNA hybridization with the proteomic information obtained by MALDI-TOF mass
146 spectrometry and the description of the full genome. This concept of species description reduces
147 the number of unassigned Operational Taxonomic Units (OTUs) to known species during
148 metagenomics [35].

149 **4. Repertoire of bacterial species detected in the vagina**

150 Since the 1800s, physicians and researchers have investigated the human vaginal microbial
151 community and its relation to diseases. Thus, of the 2,776 bacterial species isolated in human
152 beings [20], a total of 581 bacterial species were found in the vagina using culture and/or
153 molecular-based techniques. These results are summarized in supplementary Table S1 and
154 referenced using a PMID number. Overall, the 581 bacteria identified in the human vaginal flora

155 are members of 10 phyla (Fig. 1), with a predominance of *Firmicutes* (227 bacterial species,
156 39.1%) followed by *Proteobacteria* (150, 25.8%) then *Actinobacteria* (101, 17.4%), and
157 *Bacteroidetes* (74, 12.7%) (Fig. 2a). They are also distributed in 206 genera (Fig. 2b) belonging
158 to 96 families (Supplementary Table S1).
159 The bacterial species from the *Firmicutes* phylum are included into 24 families, mainly the
160 *Lactobacillaceae* (39/227, 17.2%), *Streptococcaceae* (29/227, 12.8%), and *Peptoniphilaceae*
161 families (25/227, 11%). The bacterial species from the *Proteobacteria* phylum are grouped into
162 38 families. The two most frequent families are *Enterobacteriaceae* (12%, 18/150) and
163 *Pseudomonadaceae* (12%, 18/150). The bacterial species from the *Actinobacteria* phylum are
164 grouped into 16 families. Most of them belong to the *Corynebacteriaceae* (30.7%, 31/101) and
165 *Actinomycetaceae* (18.8%, 19/101) families.
166 The bacterial species from the phylum *Bacteroidetes* are included in 12 families. More than
167 half of these bacteria are members of the *Prevotellaceae* and *Bacteroidaceae* families (41.9%
168 [31/74] and 21.6% [16/74], respectively). Finally, the remaining taxa detected in human vaginal
169 flora (5%) included in order of predominance 11 species belonging to the *Tenericutes* phylum, 8
170 to the *Fusobacteria* phylum, 4 to the *Spirochaetes* phylum, 3 to the *Chlamydiae* phylum, 2 to the
171 *Synergistetes* phylum, and 1 to the *Deinococcus-Thermus* phylum (Fig. 1, Supplementary Table
172 S1). At the genus level, with 36 different species, *Lactobacillus* was the most represented
173 followed by 30 *Corynebacterium,* 30 *Prevotella*, and 28 *Streptococcus* (Fig. 2b).
174 Based on their metabolism, bacteria have been classified using the list of prokaryotes
175 according to their aerotolerant or obligate anaerobic metabolism available on the website of the
176 University Hospital Institute Méditerranée Infection (http://www.mediterranee-
177 infection.com/article.php?laref=374&titre=list-of-prokaryotes-according-to-their-aerotolerant-or-
178 obligate-anaerobic-metabolism). Strictly anaerobic bacteria are defined as unable grow in the

179 presence of oxygen while aerotolerant are those necessitating or supporting its presence [36].
180 Thus, of the 581 bacteria detected in the vaginal microbiota, 181 species are strictly anaerobic
181 (31%). Nearly half of these species (47.5%) belong to the *Firmicutes* phylum (86/181) and 33.1%
182 to the *Bacteroidetes* phylum (60/181). These anaerobic bacteria present a higher diversity at the
183 genus level, they were divided into 71 genera (Supplementary Table S1) with a predominance of
184 *Prevotella, Bacteroidetes*, and numerous Gram-positive anaerobic cocci species (*Peptoniphilus*
185 spp., *Anaerococcus* spp., ...), mostly reported in human infections [39].

186 **5. Vaginal bacterial diversity: culture-based versus molecular techniques**

187 Of the 581 species of bacteria detected in the vaginal microbiota, 122 species classified into
188 21 genera had been identified using only culture-based methods. In contrast, 296 bacteria
189 covering 111 genera (51%) had been detected using only molecular techniques. Thus, 163 (28%)
190 were identified using both culture and molecular techniques (Fig. 3).

191 Bacterial species of the *Spirochaetes* phylum (*Treponema denticola, Treponema*
192 *maltophilum, Treponema parvum,* and *Treponema vincentii*) and *Deinococcus-Thermus* phylum
193 (*Deinococcus radiophilus*) have been identified using only molecular techniques. In addition,
194 more than half of *Proteobacteria* and *Bacteroidetes* species have been detected only by molecular
195 tools.

196 Our knowledge about the vaginal microbiota has increased with technological advancement
197 but also thanks to the decreasing cost of sequencing. Numerous vaginal bacteria have been
198 detected thanks to the sequencing of their 16S rRNA gene. These data show the disparity
199 between bacteria detected in the vagina using molecular tools and those that were cultivated.
200 Indeed, several molecular studies have attested the presence in the vaginal microbiota of
201 "uncultivated" or fastidious bacterial species such as putative new species highly associated to
202 bacterial vaginosis and designated as bacterial vaginosis-associated bacteria type 1 (BVAB1),

203 BVAB2, and BVAB3, *Megasphaera* sp type 1 and 2, *Dialister* sp type 1 to 3, and *Eggerthella* sp
204 type 1 [10, 25, 40]. Therefore, new culture strategies are required to grow these "uncultured"
205 bacteria detected in the human vagina.
206 Recently, very hopeful cultivation strategies have been elaborated. This rebirth of bacterial
207 culture allowed the isolation of many fastidious and "uncultured" bacterial species in the human
208 vagina. In 2015, *Mageeibacillus indolicus*, formerly named BVAB3 when first detected in
209 vaginal samples using molecular analyses, was cultured from endometrial biopsies of women
210 with pelvic inflammatory disease [41]. In 2016, 27 vaginal samples from 15 women were
211 cultivated with both selective and non-selective culture media. Then, isolated bacteria were
212 identified using 16S rRNA sequencing (16S rRNA sequences ≥ 98% with validly published
213 species). This study allowed the isolation of 101 bacterial species including 11 previously
214 uncultured bacteria of which were 3 highly associated with BV: *Dialister* sp type 2, *Eggerthella*
215 sp type 1, and *Megasphaera* sp type 1 [10]. This study was the first to report the cultivation of
216 *Dialister* sp type 2 and *Eggerthella* sp type 1.
217 Since 2016, the application of microbial culturomics, combining a large spectrum of culture
218 conditions and rapid bacterial identification by MALDI-TOF mass spectrometry, has broadened
219 our awareness of the human vaginal bacterial diversity, allowing the isolation of several new
220 fastidious bacterial species. Among them, several new species have been characterized using
221 taxonogenomics, which includes their whole genome sequencing: "*Olegusella massiliensis*" gen.
222 nov. sp. nov. [42], "*Vaginella massiliensis*" gen. nov. sp. nov. [43], "*Dakarella massiliensis*"
223 gen. nov. sp. nov. [44], "*Ezakiella massiliensis*" sp. nov. [45], "*Massilibacteroïdes vaginae*" gen.
224 nov. sp. nov. [46], "*Prevotella lascolaii*" sp. nov. [47], "*Corynebacterium fournierii*" sp. nov.
225 [48], and "*Murdochiella vaginalis*" sp. nov. [49]. Others have been reported only as a new
226 species announcement, which includes only a few main characteristics of the bacterium:

227 “*Peptoniphilus vaginalis”* sp. nov. [50], “*Peptoniphilus raoultii”* sp. nov. [51], “*Peptoniphilus*
228 *pacaensis”* sp. nov. [52], “*Khoudiadiopia massiliensis”* gen. nov. sp. nov. [53], “*Collinsella*
229 *vaginalis”* sp. nov. [54], “*Anaerococcus mediterraneensis”* sp. nov. [55], and “*Lactobacillus*
230 *raoultii”* sp. nov. [56].

231 Even if molecular tools have shown that the human vagina includes a high proportion of
232 unculturable or fastidious bacteria, molecular data complement but do not replace those provided
233 by culture techniques. Therefore, a mapping of the diversity of the vaginal bacterial community
234 requires both synthesis and application of various approaches and techniques comprising culture-
235 based methods.

236 **6. Vaginal bacterial microbiota in normal and abnormal conditions**

237 Knowledge about the vaginal microbiome, normal as well as abnormal, has expanded in
238 recent years. Advances in technology such as culture-independent methods have shown that the
239 vagina is a dynamic and complex ecosystem, principally dominated by *Lactobacilli* [16]. Normal
240 and abnormal vaginal microbiota harbor more than 250 species of bacteria [16] and are under the
241 influence of intrinsic and extrinsic factors [13, 15].

242 *6.1. Normal vaginal bacterial flora*

243 The term “normal” vaginal bacterial flora is used to describe all bacterial species that are
244 commonly found in the vaginal biotope of healthy women. The constituents of the flora and their
245 quantities are influenced by many factors such as age, hormonal fluctuations, menstruation,
246 douching, hygiene, pregnancy, and sexual practices [57, 58].

247 Overall, normal vaginal flora is dominated by various species of *Lactobacillus*. Thus, a
248 predominance of typical Gram-positive rods, known as “Döderlein's bacilli”, is observed using
249 microscopy in normal vaginal flora [22]. Over twenty *Lactobacillus* spp. have been found in the
250 vagina of a premenopausal woman [7] whereas a woman of childbearing age is generally

251 dominated by one or two species of *Lactobacillus*; the most common are *L. crispatus, L. jensenii,*
252 *L. gasseri*, and *L. iners* [16]. *Lactobacillus* spp. protect the vaginal ecosystem through the
253 production of antimicrobial molecules that exclude and inhibit the growth and expansion of other
254 microorganisms [4]. Lactic acid, that maintains vaginal pH between 3.5 to 4.5, and hydrogen
255 peroxide, an antimicrobial product that protects against harmful microbes and which is produced
256 by *L. acidophilus* and *L. casei*, are also able to inhibit the growth of bacteria usually associated
257 with a dysbiotic state such as *Gardnerella vaginalis*, *Bacteroides* spp., *Mobiluncus* spp., and
258 anaerobic cocci [59]. However, a decrease in the production of hydrogen peroxide and
259 bacteriocins by *Lactobacilli* stimulates the growth of *G. vaginalis, Prevotella bivia, Mobiluncus*
260 spp*, Peptococcus* spp, and *Peptostreptococcus anaerobius* [60].

261 *6.2. Abnormal vaginal bacterial microbiota: the case of bacterial vaginosis*

262 The vagina is a very versatile organ that can affect the health of women and their newborns.
263 One of the main reasons why women seek gynecological care is vaginal complaints. Only one
264 gynecologic condition is directly associated with imbalance of vaginal bacterial communities:
265 bacterial vaginosis. Overall, bacterial vaginosis is a common vaginal condition [60]. However,
266 prevalence of BV depends on the studied populations. Overall, in developed nations, bacterial
267 vaginosis occurs around 10-30% among women who have sex with men (WSM) and between 25-
268 50% among women who have sex with women (WSW) [61, 62]. In contrast, the estimated
269 prevalence is greater than 50% in East/Southern Africa [63]. Besides, this common condition is
270 still of unknown etiology and remains one of the great enigmas in women's health.

271 For physicians, bacterial vaginosis is considered as a gynecologic "infection" marked by
272 the presence of three of four criteria, known as Amsel's criteria: an elevation of the vaginal pH
273 (from 3.8 - 4.2 up to 7.0), a milky creamy vaginal discharge, a malodorous odor of vaginal
274 secretions (positive with potassium test, "sniff test"), and a presence of clue cells [64]. However,

275 for microbiologists, bacterial vaginosis is not an infection but represents a dysbiosis of vaginal
276 flora. This vaginal dysbiosis is associated with adverse health outcomes [65] including increased
277 risks of abnormal pregnancy outcomes such as miscarriage, preterm birth, chorioamnionitis,
278 pelvic inflammatory disease [66], and also acquisition of sexual transmitted infections like
279 gonorrhea, *Chlamydia*, herpes simplex type 2 [22], and human immunodeficiency virus (HIV)
280 infections [67].

281 Bacterial vaginosis is characterized by a dramatic switch of vaginal bacterial flora from
282 normal predominant *Lactobacilli* to a polymicrobial flora that regroups various facultative and
283 anaerobic bacteria [60]. Indeed, vaginal flora in bacterial vaginosis contains a broad spectrum of
284 bacteria including *Gardnerella vaginalis, Atopobium vaginae, Mycoplasma hominis, Ureaplasma*
285 *urealyticum,* members of genera *Prevotella, Bacteroides, Dialister, Megasphaera,*
286 *Peptostreptococcus, Mobiluncus, Sneathia, Leptotrichia*, and putative new species belonging to
287 the *Clostridiales* order named BVAB1, BVAB2 and *Mageeibacillus indolicus* formerly known as
288 BVAB3 [24, 41, 68]. The observation of a vaginal specimen containing high levels of BVABs on
289 transmission electron microscopy reveals that BVAB1 seems to be a thin curved rod, BVAB2 a
290 short and fat rod and BVAB3 (*Mageeibacillus indolicus)* appears as a long and lancet rod-shaped
291 bacteria [22]. Besides, *Lactobacillus iners* is also considered as a marker of vaginal disorder as
292 its presence has been associated to that of other bacterial vaginosis-related bacteria such as
293 *Megasphaera*, *Leptotrichia*, and *Eggerthella* and bacterial vaginosis [69].

294 Bacterial vaginosis is also highly correlated with hormonal fluctuations and sexual
295 behaviors and practices [70, 71]. The strong presence in bacterial vaginosis microbiota of oral or
296 gut bacterial populations such as *Leptotrichia amnionii, Sneathia sanguinegens, Tannerella*
297 *forsythia*, *Treponema denticola*, *Prevotella intermedia* [59] and species of *Anaerococcus,*

Clostridium, Peptoniphilus, and *Prevotella* genera, respectively [72], suggest that BV may be linked to fecal or oral transplantation.

7. Conclusions

Overall, the composition of the vaginal flora varies depending on many factors. Compared to healthy women, those suffering from bacterial vaginosis present a complex and dynamic vaginal bacterial microbiota populated by diverse anaerobic and fastidious bacteria. Although the alteration of vaginal microbiome predisposes to healthy adverse outcomes, molecular tools suggest that some variations of vaginal flora are normal and do not necessarily cause diseases.

In recent decades, using cultivation-independent molecular methods has produced an explosion in the understanding of human bacterial microbiota in various ecosystems. Applied to the human vaginal ecosystem, these molecular methods have shown the presence of a very important proportion of uncultivable bacteria in the vaginal econiche. However, recently, the rebirth of culture techniques allowed, in addition to bacteria previously detected using only molecular tools, the isolation of fastidious and new bacterial species in the vagina. The comparison of bacteria of the vaginal microbiome repertoire found using molecular tools with those isolated by culture-based methods showed that these two different techniques enable the detection of a wide and somewhat overlapping range of vaginal bacterial species.

This repertoire is important and represents the starting point to better understand the dysbioses or infections caused by the unsteadiness of the vaginal microbiota. Thus, characterization of the vaginal microbiota by combining both cultivation and cultivation-independent methods may certainly show the role of bacteria in healthy and pathologic conditions.

320 **Authors' contributions**

321 KD investigated the literature, collected, analyzed and interpreted the data, and also drafted the
322 manuscript. JCD performed the search methodologies and contributed to the acquisition and
323 analysis of the data. AL participated to the data analysis and helped to write the manuscript. FF
324 designed and coordinated the study, investigated the literature, interpreted the data and drafted
325 the manuscript. All authors read and approved the final manuscript.

326

327 **Conflict of interest**

328 The authors declare no conflict of interest.

329

330 **Acknowledgments**

331 This study was funded by the "Fondation Méditerranée Infection" and the French Government
332 under the "Investissements d'avenir" program managed by the National Agency for Research
333 (reference Méditerranée Infection 10-IAHU-03).

334 **References**

335 1. **Ma B, Forney LJ, Ravel J**. Vaginal Microbiome: Rethinking Health and Disease. *Annu Rev*
336 *Microbiol* 2012;66:371–389.
337 2. **Franasiak JM, Scott RT**. Reproductive tract microbiome in assisted reproductive
338 technologies. *Fertil Steril* 2015;104:1364–1371.
339 3. **Green KA, Zarek SM, Catherino WH**. Gynecologic health and disease in relation to the
340 microbiome of the female reproductive tract. *Fertil Steril* 2015;104:1351–1357.
341 4. **Lepargneur JP, Rousseau V**. Protective role of the Doderleïn flora. *J Gynecol Obstet Biol*
342 *Reprod* 2002;31:485–494.
343 5. **Li J, McCormick J, Bocking A, Reid G**. Importance of Vaginal Microbes in Reproductive
344 Health. *Reprod Sci* 2012;19:235–242.
345 6. **Danielsson D, Teigen PK, Moi H**. The genital econiche: focus on microbiota and bacterial
346 vaginosis: Genital econiche. *Ann N Y Acad Sci* 2011;1230:48–58.
347 7. **Lamont R, Sobel J, Akins R, Hassan S, Chaiworapongsa T, *et al.*** The vaginal
348 microbiome: new information about genital tract flora using molecular based techniques:
349 Vaginal microbiome using molecular tools. *BJOG Int J Obstet Gynaecol* 2011;118:533–549.
350 8. **Braundmeier AG, Lenz KM, Inman KS, Chia N, Jeraldo P, *et al.*** Individualized medicine
351 and the microbiome in reproductive tract. *Front Physiol*;6. Epub ahead of print 1 April 2015.
352 DOI: 10.3389/fphys.2015.00097.
353 9. **Huang B, Fettweis JM, Brooks JP, Jefferson KK, Buck GA**. The Changing Landscape of
354 the Vaginal Microbiome. *Clin Lab Med* 2014;34:747–761.
355 10. **Srinivasan S, Munch MM, Sizova MV, Fiedler TL, Kohler CM, *et al.*** More Easily
356 Cultivated Than Identified Classical Isolation with molecular identification of vaginal
357 bacteria.pdf. *J Infect Dis* 2016;214:S21-8.

358 11. **Verhelst R, Verstraelen H, Claeys G, Verschraegen G, Van Simaey L, *et al.*** Comparison
359 between Gram stain and culture for the characterization of vaginal microflora: definition of a
360 distinct grade that resembles grade I microflora and revised categorization of grade I
361 microflora. *BMC Microbiol* 2005;5:61.
362 12. **Pandya S, Ravi K, Srinivas V, Jadhav S, Khan A, *et al.*** Comparison of culture-dependent
363 and culture-independent molecular methods for characterization of vaginal microflora. *J Med*
364 *Microbiol* 2017;66:149–153.
365 13. **Zapata HJ, Quagliarello VJ**. The Microbiota and Microbiome in Aging: Potential
366 Implications in Health and Age-Related Diseases. *J Am Geriatr Soc* 2015;63:776–781.
367 14. **Herbst-Kralovetz MM, Pyles RB, Ratner AJ, Sycuro LK, Mitchell C**. New Systems for
368 Studying Intercellular Interactions in Bacterial Vaginosis. *J Infect Dis* 2016;214:S6–S13.
369 15. **Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, *et al.*** Differences in vaginal
370 microbiome in African American women versus women of European ancestry. *Microbiology*
371 2014;160:2272–2282.
372 16. **Mendling W**. Vaginal Microbiota. In: Schwiertz A (editor). *Microbiota of the Human Body*.
373 Cham: Springer International Publishing. pp. 83–93.
374 17. **Reid G**. Is bacterial vaginosis a disease? *Appl Microbiol Biotechnol* 2018;102:553–558.
375 18 **Bradford LL, Ravel J**. The vaginal mycobiome: A contemporary perspective on fungi in
376 women's health and diseases. *Virulence* 2017;8:342–351.
377 19. **Wylie KM, Wylie TN, Cahill AG, Macones GA, Tuuli MG, *et al.*** The vaginal eukaryotic
378 DNA virome and preterm birth. *Am J Obstet Gynecol* 2018;219:189.e1-189.e12.
379

380 20. **Bilen M, Dufour JC, Lagier JC, Cadoret F, Daoud Z, *et al.*** The contribution of
381 culturomics to the repertoire of isolated human bacterial and archaeal species. *Microbiome*;6.
382 Epub ahead of print December 2018. DOI: 10.1186/s40168-018-0485-5.
383 21. **Martin DH**. The Microbiota of the Vagina and Its Influence on Women's Health and
384 Disease. *Am J Med Sci* 2012;343:2–9.
385 22. **Srinivasan S, Fredricks DN**. The Human Vaginal Bacterial Biota and Bacterial Vaginosis.
386 *Interdiscip Perspect Infect Dis* 2008;2008:1–22.
387 23. **Shipitsyna E, Roos A, Datcu R, Hallén A, Fredlund H, *et al.*** Composition of the Vaginal
388 Microbiota in Women of Reproductive Age – Sensitive and Specific Molecular Diagnosis of
389 Bacterial Vaginosis Is Possible? *PLoS ONE* 2013;8:e60670.
390 24. **Onderdonk AB, Delaney ML, Fichorova RN**. The Human Microbiome during Bacterial
391 Vaginosis. *Clin Microbiol Rev* 2016;29:223–238.
392 25. **Fredricks DN, Fiedler TL, Marrazzo JM**. Molecular identification of bacteria associated
393 with bacterial vaginosis. *N Engl J Med* 2005;353:1899–1911.
394 26. **Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, *et al.*** A Metagenomic Approach to
395 Characterization of the Vaginal Microbiome Signature in Pregnancy. *PLoS ONE*
396 2012;7:e36466.
397 27. **Menard JP, Mazouni C, Fenollar F, Raoult D, Boubli L, *et al.*** Diagnostic accuracy of
398 quantitative real-time PCR assay versus clinical and Gram stain identification of bacterial
399 vaginosis. *Eur J Clin Microbiol Infect Dis* 2010;29:1547–1552.
400 28. **Lopes dos Santos Santiago G, Tency I, Verstraelen H, Verhelst R, Trog M, *et al.***
401 Longitudinal qPCR Study of the Dynamics of *L. crispatus, L. iners, A. vaginae,* (Sialidase
402 Positive) *G. vaginalis*, and *P. bivia* in the Vagina. *PLoS ONE* 2012;7:e45281.

403 29. **Machado A, Almeida C, Salgueiro D, Henriques A, Vaneechoutte M, *et al.*** Fluorescence
404 in situ Hybridization method using Peptide Nucleic Acid probes for rapid detection of
405 *Lactobacillus* and *Gardnerella* spp. *BMC Microbiol* 2013;13:82.
406 30. **Bomar L, Brugger SD, Lemon KP**. Bacterial microbiota of the nasal passages across the
407 span of human life. *Curr Opin Microbiol* 2018;41:8–14.
408 31. **Sun D-L, Jiang X, Wu QL, Zhou N-Y**. Intragenomic Heterogeneity of 16S rRNA Genes
409 Causes Overestimation of Prokaryotic Diversity. *Appl Environ Microbiol* 2013;79:5962–
410 5969.
411 32. **Crosby LD, Criddle CS**. Understanding Bias in Microbial Community Analysis Techniques
412 due to *rrn* Operon Copy Number Heterogeneity. *BioTechniques* 2003;34:790–802
413 33. **Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, *et al.*** Microbial culturomics:
414 paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* 2012;18:1185–
415 1193.
416 34. **Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, *et al.*** The Rebirth of Culture
417 in Microbiology through the Example of Culturomics To Study Human Gut Microbiota. *Clin*
418 *Microbiol Rev* 2015;28:237–264.
419 35. **Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, *et al.*** Culture of previously
420 uncultured members of the human gut microbiota by culturomics. *Nat Microbiol*
421 2016;1:16203.
422 36. **Fournier PE, Drancourt M**. New Microbes New Infections promotes modern prokaryotic
423 taxonomy: a new section "TaxonoGenomics: new genomes of microorganisms in humans".
424 *New Microbes New Infect* 2015;7:48–49.

425 37. **Fournier PE, Lagier JC, Dubourg G, Raoult D**. From culturomics to taxonomogenomics:
426 A need to change the taxonomy of prokaryotes in clinical microbiology. *Anaerobe*
427 2015;36:73–78.
428 38. **Hentges DJ**. Anaerobes: General Characteristics. In: Baron S (editor). *Medical*
429 *Microbiology*. Galveston (TX): University of Texas Medical Branch at Galveston.
430 http://www.ncbi.nlm.nih.gov/books/NBK7638/ (1996, accessed 28 December 2017).
431 39. **Murphy EC, Frick I-M**. Gram-positive anaerobic cocci – commensals and opportunistic
432 pathogens. *FEMS Microbiol Rev* 2013;37:520–553.
433 40. **Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW,** ***et al.*** The composition and
434 stability of the vaginal microbiota of normal pregnant women is different from that of non-
435 pregnant women. *Microbiome* 2014;2:4.
436 41. **Austin MN, Rabe LK, Srinivasan S, Fredricks DN, Wiesenfeld HC,** ***et al.*** *Mageeibacillus*
437 *indolicus* gen. nov., sp. nov: A novel bacterium isolated from the female genital tract.
438 *Anaerobe* 2015;32:37–42.
439 42. **Diop K, Diop A, Bretelle F, Cadoret F, Michelle C,** ***et al.*** *Olegusella massiliensis* gen.
440 nov., sp. nov., strain KHD7^T, a new bacterial genus isolated from the female genital tract of a
441 patient with bacterial vaginosis. *Anaerobe* 2017;44:87–95.
442 43. **Diop K, Bretelle F, Michelle C, Richez M, Rathored J,** ***et al.*** Taxonogenomics and
443 description of *Vaginella massiliensis* gen. nov., sp. nov., strain Marseille P2517^T, a new
444 bacterial genus isolated from the human vagina. *New Microbes New Infect* 2017;15:94–103.
445 44. **Dione N, Rathored J, Tomei E, Lagier JC, Khelaifia S,** ***et al.*** *Dakarella massiliensis* gen.
446 nov., sp. nov., strain ND3^T: a new bacterial genus isolated from the female genital tract. *New*
447 *Microbes New Infect* 2017;18:38–46.

448 45. **Diop K, Andrieu C, Michelle C, Armstrong N, Bittar F, *et al.*** Characterization of a New
449 *Ezakiella* Isolated from the Human Vagina: Genome Sequence and Description of *Ezakiella*
450 *massiliensis* sp. nov. *Curr Microbiol*. 2018;75:456–463.
451 46. **Jneid J, Bou Khalil JY, Aherfi S, Blanc-Tailleur C, Raoult D, *et al.*** '*Massilibacteroïdes*
452 *vaginae*' gen. nov., sp. nov., a new genus within the family of *Bacteroïdaceae* isolated from a
453 patient with vaginosis. *New Microbes New Infect* 2017;21:1–8.
454 47. **Diop K, Diop A, Levasseur A, Mediannikov O, Robert C, *et al.*** Microbial Culturomics
455 Broadens Human Vaginal Flora Diversity: Genome Sequence and Description of *Prevotella*
456 *lascolaii* sp. nov. Isolated from a Patient with Bacterial Vaginosis. *OMICS J Integr Biol*.
457 2018;22:210–222.
458 48. **Diop K, Nguyen TT, Delerce J, Armstrong N, Raoult D, *et al.*** *Corynebacterium fournierii*
459 sp. nov., isolated from the female genital tract of a patient with bacterial vaginosis. *Antonie*
460 *Van Leeuwenhoek* 2018;1–10.
461 49. **Diop K, Diop A, Khelaifia S, Robert C, Pinto FD, *et al.*** Characterization of a novel Gram-
462 stain-positive anaerobic coccus isolated from the female genital tract: Genome sequence and
463 description of *Murdochiella vaginalis* sp. nov. *MicrobiologyOpen* 2018;e00570.
464 50. **Diop K, Mediannikov O, Raoult D, Bretelle F, Fenollar F**. '*Peptoniphilus vaginalis'* sp.
465 nov., a new species isolated from human female genital tract. *New Microbes New Infect*
466 2016;13:65–66.
467 51. **Diop K, Bretelle F, Mediannikov O, Fournier PE, Fenollar F**. '*Peptoniphilus raoultii'* sp.
468 nov., a new species isolated from human female genital tract. *New Microbes New Infect*
469 2016;13:29–31.

470 52. **Diop K, Bretelle F, Mediannikov O, Fournier PE, Fenollar F**. '*Peptoniphilus pacaensis'* sp. nov., a new species isolated from human female genital tract. *New Microbes New Infect* 2016;13:37–39.

473 53. **Diop A, Raoult D, Fenollar F, Fournier P-E**. '*Khoudiadiopia massiliensis'* gen. nov., sp. nov., strain Marseille-P2746^{T}, a new bacterial genus isolated from the female genital tract. *New Microbes New Infect* 2017;19:60–61.

476 54. **Diop K, Bretelle F, Fournier PE, Raoult D, Fenollar F**. *"Collinsella vaginalis"* sp. nov., a new bacterial species cultivated from human female genital tract. *Hum Microbiome J* 2016;2:19–20.

479 55. **Diop K, Bretelle F, Fournier PE, Fenollar F**. *'Anaerococcus mediterraneensis*' sp. nov., a new species isolated from human female genital tract. *New Microbes New Infect* 2017;17:75–76.

482 56. **Nicaise B, Bilen M, Cadoret F, Bretelle F, Fenollar F**. *'Lactobacillus raoultii'* sp. nov., a new bacterium isolated from the vaginal flora of a woman with bacterial vaginosis. *New Microbes New Infect* 2017;21:20–22.

485 57. **Georgijević A, Cjukić-Ivancević S, Bujko M**. Bacterial vaginosis. Epidemiology and risk factors. *Srp Arh Celok Lek* 2000;128:29–33.

487 58. **Bautista CT, Wurapa E, Sateren WB, Morris S, Hollingsworth B, *et al.*** Bacterial vaginosis: a synthesis of the literature on etiology, prevalence, risk factors, and relationship with chlamydia and gonorrhea infections. *Mil Med Res*;3. Epub ahead of print December 2016. DOI: 10.1186/s40779-016-0074-5.

491 59. **Africa C, Nel J, Stemmet M**. Anaerobes and Bacterial Vaginosis in Pregnancy: Virulence Factors Contributing to Vaginal Colonisation. *Int J Environ Res Public Health* 2014;11:6979–7000.

494 60. **Muzny CA, Schwebke JR**. Pathogenesis of Bacterial Vaginosis: Discussion of Current
495 Hypotheses. *J Infect Dis* 2016;214:S1–S5.
496 61. **Bilardi J, Walker S, Mooney-Somers J, Temple-Smith M, McNair R, *et al.*** Women's
497 Views and Experiences of the Triggers for Onset of Bacterial Vaginosis and Exacerbating
498 Factors Associated with Recurrence. *PLOS ONE* 2016;11:e0150272.
499 62. **Bradshaw CS, Brotman RM**. Making inroads into improving treatment of bacterial
500 vaginosis – striving for long-term cure. *BMC Infect Dis*;15. Epub ahead of print December
501 2015. DOI: 10.1186/s12879-015-1027-4.
502 63. **Chico RM, Mayaud P, Ariti C, Mabey D, Ronsmans C, *et al.*** Prevalence of malaria and
503 sexually transmitted and reproductive tract infections in pregnancy in sub-Saharan Africa: a
504 systematic review. *JAMA* 2012;307:2079–2086.
505 64. **Nugent RP, Krohn MA, Hillier SL**. Reliability of diagnosing bacterial vaginosis is
506 improved by a standardized method of gram stain interpretation. *J Clin Microbiol*
507 1991;29:297–301.
508 65. **Martin DH, Marrazzo JM**. The Vaginal Microbiome: Current Understanding and Future
509 Directions. *J Infect Dis* 2016;214:S36–S41.
510 66. **Afolabi BB, Moses OE, Oduyebo OO**. Bacterial Vaginosis and Pregnancy Outcome in
511 Lagos, Nigeria. *Open Forum Infect Dis* 2016;3:ofw030.
512 67. **McClelland RS, Lingappa JR, Srinivasan S, Kinuthia J, John-Stewart GC, *et al.***
513 Evaluation of the association between the concentrations of key vaginal bacteria and the
514 increased risk of HIV acquisition in African women from five cohorts: a nested case-control
515 study. *Lancet Infect Dis*. Epub ahead of print 25 January 2018. DOI: 10.1016/S1473-
516 3099(18)30058-6.

517 68. **Zozaya M, Ferris MJ, Siren JD, Lillis R, Myers L, *et al.*** Bacterial communities in penile
518 skin, male urethra, and vaginas of heterosexual couples with and without bacterial vaginosis.
519 *Microbiome*;4. Epub ahead of print December 2016. DOI: 10.1186/s40168-016-0161-6.
520 69. **Tamrakar R, Yamada T, Furuta I, Cho K, Morikawa M, *et al.*** Association between
521 *Lactobacillus* species and bacterial vaginosis-related bacteria, and bacterial vaginosis scores
522 in pregnant Japanese women. *BMC Infect Dis* 2007;7:128.
523 70. **Muhleisen AL, Herbst-Kralovetz MM**. Menopause and the vaginal microbiome. *Maturitas*
524 2016;91:42–50.
525 71. **Forcey DS, Vodstrcil LA, Hocking JS, Fairley CK, Law M, *et al.*** Factors associated with
526 bacterial vaginosis among women who have sex with women: a systematic review. *PloS One*
527 2015;10:e0141905.
528 72. **Fenollar F, Raoult D**. Does Bacterial Vaginosis Result From Fecal Transplantation? *J Infect*
529 *Dis* 2016;214:1784–1784.

530 **Figure legends.**

531 **Figure 1.** Phylogenetic tree of the 581 bacteria detected in the human vagina. Bacterial species
532 from *Firmicutes* are highlighted in red, *Actinobacteria* (green), *Proteobacteria* (blue),
533 *Bacteroidetes* (purple), *Fusobacteria* (dark purple), *Tenericutes* (dark blue), *Chlamydiae* (dark
534 red), *Synergistetes* (dark green), *Spirochaetes* (light blue), and *Deinococcus-Thermus* (grey),
535 respectively. The evolutionary history was inferred by using the Maximum Likelihood method
536 based on the Jukes-Cantor model (Jukes *et al.,* 1969). All positions with less than 90% site
537 coverage were eliminated. Finally, a total of 1286 positions was included in the final dataset.
538 Evolutionary analyses were conducted in MEGA7 (Kumar *et al.,* 2016). Bootstrap analyses were
539 performed with 1000 replicates.

540 **Figure 2.** Distribution of bacterial species isolated in the human vagina. a. Classification by
541 phylum of bacterial species. b. Classification of bacterial according to genus.

542 **Figure 3.** Venn Diagram of vaginal bacterial diversity: culture-based versus molecular
543 techniques.

544 **Table 1.** List of elaborated query patterns used for PubMed search.

545 **Table S1.** Vaginal bacterial microbiota. A: List of species reported in this repertoire. B, C, D:
546 Classification of the bacterial species at phylum, family, and genus levels according to NCBI
547 taxonomy (www.ncbi.nlm.nih.gov/taxonomy). E: Classification of the bacterial species based on
548 their oxygen tolerance (1, strictly anaerobic and 0, aerotolerant). F, G: The PMID of the
549 referenced article (ND, Not available Data).

550

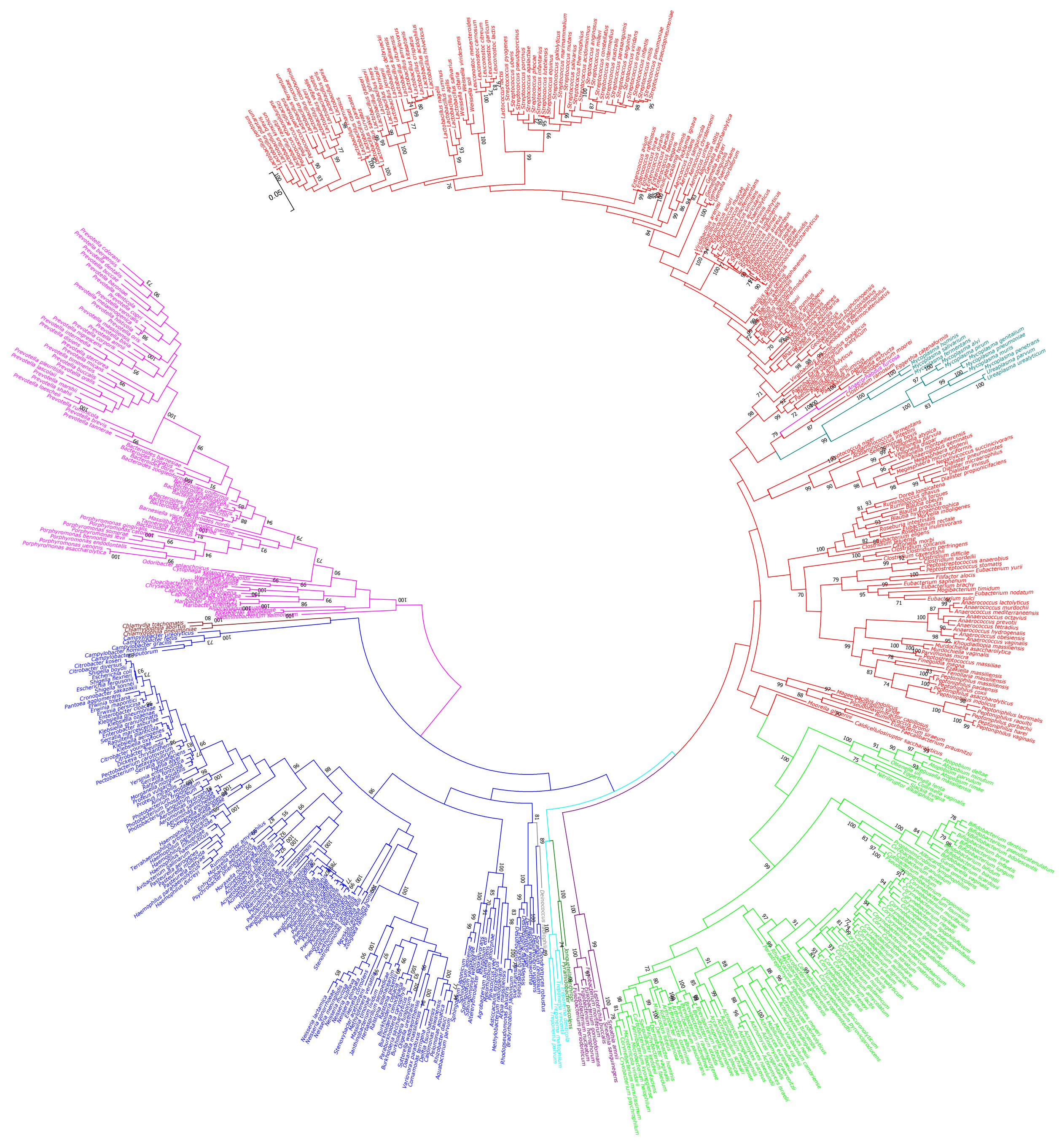
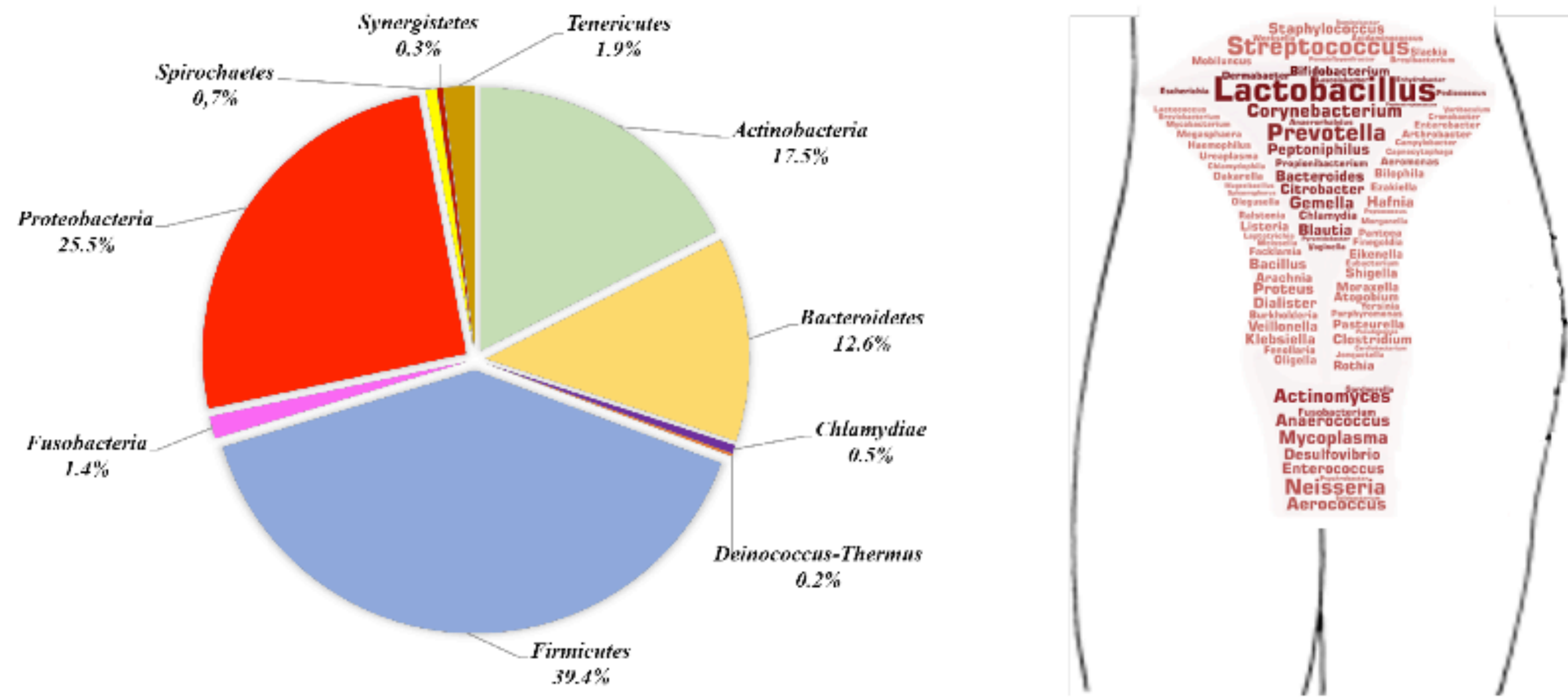


Figure 1. Phylogenetic tree of the 581 bacteria detected in the human vagina. Bacterial species from *Firmicutes* are highlighted in red, *Actinobacteria* (green), *Proteobacteria* (blue), *Bacteroidetes* (purple), *Fusobacteria* (dark purple), *Tenericutes* (dark blue), *Chlamydiae* (dark red), *Synergistetes* (dark green), *Spirochaetes* (light blue), and *Deinococcus-Thermus* (grey), respectively.

557

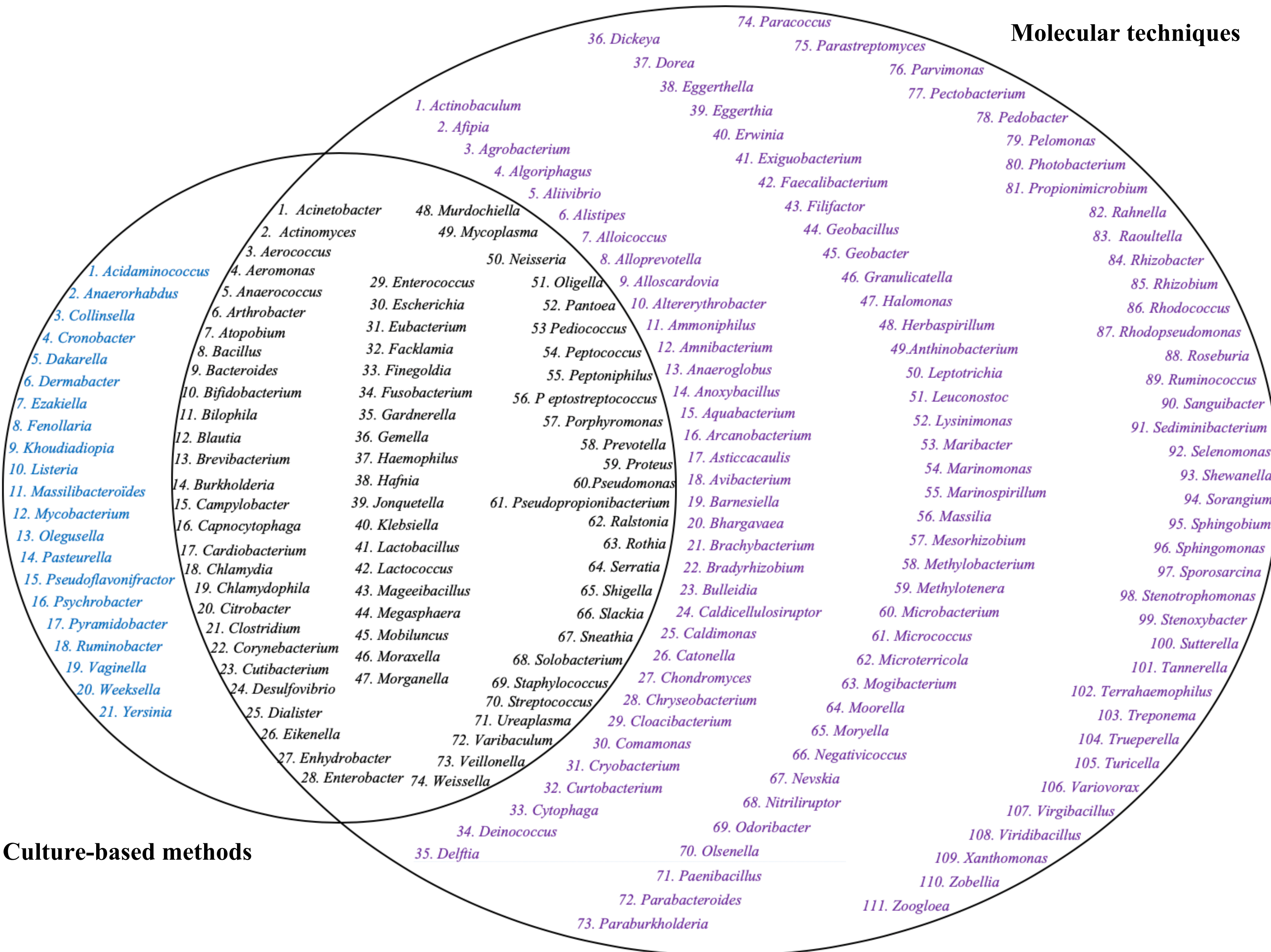


558

559

560 **Figure 2.** Distribution of bacterial species isolated in the human vagina. a. Classification by phylum of bacterial species. b. Classification of

561 bacterial according to genus.



562

563 **Figure 3.** Venn Diagram of vaginal bacterial diversity: culture-based versus molecular techniques

564 **Table 1. List of elaborated query patterns used for PubMed search.**

Query Pattern	Syntax
QP1	#3[tiab] AND (humans[Mesh] OR human[tiab] OR human[ot] OR humans[tiab] OR humans[ot] OR woman[tiab] OR woman[ot] OR women[tiab] OR women[ot] OR girl[tiab] OR girl[ot] OR girls[tiab] OR girls[ot]) AND (Vagina[MH] OR Vagina[tiab] OR Vagina[ot] OR Vaginal Diseases[MH] OR Vaginal Diseases[tiab] OR Vaginal Diseases[ot] OR Vaginal Smears[MH] OR Vaginal Smears[tiab] OR Vaginal Smears[ot] OR Vaginal Microbiota[tiab] OR Vaginal Microbiota[ot] OR Vaginal Dysbiosis [tiab] OR Vaginal Dysbiosis [ot] OR Bacterial Vaginosis[tiab] OR Bacterial Vaginosis [ot])
QP2	#3[tiab] AND (humans[Mesh] OR human[tiab] OR human[ot] OR humans[tiab] OR humans[ot] OR woman[tiab] OR woman[ot] OR women[tiab] OR women[ot] OR girl[tiab] OR girl[ot] OR girls[tiab] OR girls[ot]) AND (Vagina[MH] OR Vagina[tiab] OR Vagina[ot] OR Vaginal Diseases[MH] OR Vaginal Diseases[tiab] OR Vaginal Diseases[ot] OR Vaginal Smears[MH] OR Vaginal Smears[tiab] OR Vaginal Smears[ot] OR Vaginal Microbiota[tiab] OR Vaginal Microbiota[ot] OR Vaginal Dysbiosis [tiab] OR Vaginal Dysbiosis [ot] OR Bacterial Vaginosis[tiab] OR Bacterial Vaginosis [ot]) AND (DNA, Bacterial [MH] OR High-Throughput Nucleotide Sequencing [MH] OR isolation and purification [SH] OR RNA, Ribosomal, 16S [MH] OR Sequence Analysis, DNA [MH] OR Polymerase Chain Reaction [MH] OR Metagenome [MH] OR Metagenomics [MH] OR OR Bacteriological Techniques [MH] OR Culture Media [MH])
QP3	#3[tiab] AND (humans[Mesh] OR human[tiab] OR human[ot] OR humans[tiab] OR humans[ot] OR woman[tiab] OR woman[ot] OR women[tiab] OR women[ot] OR girl[tiab] OR girl[ot] OR girls[tiab] OR girls[ot]) AND (Vagina[MH] OR Vagina[tiab] OR Vagina[ot] OR Vaginal Diseases[MH] OR Vaginal Diseases[tiab] OR Vaginal Diseases[ot] OR Vaginal Smears[MH] OR Vaginal Smears[tiab] OR Vaginal Smears[ot] OR Vaginal Microbiota[tiab] OR Vaginal Microbiota[ot] OR Vaginal Dysbiosis [tiab] OR Vaginal Dysbiosis [ot] OR Bacterial Vaginosis[tiab] OR Bacterial Vaginosis [ot]) AND (African Continental Ancestry Group[MH] OR African Continental Ancestry Group[tiab] OR African Continental Ancestry Group[ot] OR Ethnic Groups[MH] OR Ethnic Groups[tiab] OR Ethnic Groups[ot])

565 For each species, “#3” label was automatically replaced by the program by “OR” boolean

566 operators combining the species scientific name and its synonyms listed in NCBI taxonomy.

567

CHAPITRE II :

Caractérisation des Flores Vaginales Normale et de Vaginose Bactérienne par Culturomique et Métagénomique

Avant-propos

Plusieurs études ont récemment examiné le microbiote vaginal des patientes souffrant de vaginose bactérienne, le plus souvent par des méthodes moléculaires reposant sur la métagénomique. Ces dernières ont montré que ces femmes possèdent un microbiote complexe dominé par des bactéries anaérobies. De plus, il n'existait pas d'isolat pour 80% de la communauté bactérienne vaginale identifiés par les outils moléculaires [7], dont plus d'une dizaine de bactéries détectées comme étant fortement associées à la vaginose bactérienne [37, 64, 65]. Etant donné que la métagénomique ne détecte que les séquences d'ADN des microorganismes (vivants aussi bien que morts), il est donc difficile de savoir si cet écart est dû à l'incapacité de maintenir ces bactéries vivantes une fois le prélèvement vaginal effectué ou si ces bactéries étaient déjà mortes avant que le prélèvement ne soit effectué. Ainsi, dans le but d'identifier et d'isoler les bactéries associées à la vaginose bactérienne, nous avons analysé 50 échantillons vaginaux, de femmes vivant en France (Marseille, n = 24) et au Sénégal (zone rurale, n = 10), dont 22 sujets sains et 12 patientes atteintes de vaginose bactérienne à l'aide de deux approches : la culturomique et la métagénomique. Le diagnostic de vaginose bactérienne a été établi notamment par un outil moléculaire quantitatif ciblant la quantification de 2 bactéries : *Atopobium vaginae* et *Gardnerella vaginalis*. Cet outil moléculaire est utilisé pour le diagnostic de routine de la vaginose bactérienne à l'IHU.

Concernant la culturomique, dans un premier temps, les échantillons ont été ensemencés directement sur 5 milieux de culture solides : gélose anabasal, gélose ANC (acide nalidixique, colistine), gélose Schaedler à la vitamine K1, gélose au sang et gélose Trypticase soja au sang de cheval. Les boîtes de Pétri ont été incubées à 37°C sous atmosphère anaérobie pendant 4 à 7 jours. Puis, brièvement, 2 ml du liquide vaginal restant ont été pré-incubés dans un milieu de culture liquide anaérobie et aérobie contenant du sang et du rumen filtré [38]. A différentes périodes de pré-incubation (1, 3, 7, 10, 15, 21 et 30 jours) à 37°C, 100 µl du bouillon anaérobie

ont été étalés sur les 5 milieux cités ci-dessus puis les géloses ont été incubées pendant 7 jours dans les mêmes conditions, tandis que celui aérobie a été ensemencé sur des géloses Chocolat puis incubé en aérobiose pendant 3 jours à 37°C. Les bactéries cultivées ont été isolées en culture pure puis identifiées, tout d'abord, par spectrométrie de masse MALDI-TOF (Matrix Assisted Laser Desorption Ionisation - Time of Flight), puis en cas d'échec par séquençage du gène codant pour l'ARNr 16S [66, 67]. Pour la métagénomique, l'ADN des échantillons vaginaux a été extrait par traitement mécanique utilisant un appareil FastPrep. Ensuite, ils ont été traités par lyse [68] avant d'être amplifiés puis séquencés pour le gène codant l'ARNr 16S.

En comparant les données obtenues par métagénomique à celles de la culturomique, nous n'avons pas observé, d'une part, qu'une des deux approches soit nettement plus efficace que l'autre. En revanche, lorsqu'elles sont utilisées en parallèle, les deux techniques permettent de détecter des microbiotes vaginaux diversifiés, en grande partie non chevauchants. En effet, sur les 554 bactéries détectées dans le microbiote vaginal, en utilisant ces 2 techniques en parallèle à partir de 34 échantillons, 194 espèces (35%) n'étaient identifiées que par culture, 214 (39%) que par métagénomique et 146 (26%) par les deux techniques. Ces résultats soulignent bien la complémentarité de ces deux approches. D'autre part, nous avons observé que la composition du microbiote vaginal change considérablement durant la vaginose bactérienne. Il y a en effet une augmentation significative de la diversité bactérienne connue comme inconnue, une diminution des espèces de la famille de *Lactobacillaceae* et des *Proteobacteria* appartenant aux familles *des Sutterellaceae, Enterobacteriaceae* et *Pseudomonadaceae* ainsi qu'une forte présence des espèces de *Leptotrichiaceae* (phylum *Fusobacteria*) et *Bacteroidales* (phylum *Bacteroidetes*). De plus, nous avons remarqué que la flore de vaginose bactérienne est enrichie en microorganismes anaérobies par rapport aux sujets sains. Enfin, la combinaison de la métagénomique et la culturomique a permis l'identification d'un complexe de 11 espèces/genres bactériens associés à la vaginose : *G. vaginalis, A. vaginae,*

Aeroccocus christensenii, Prevotella, *Peptoniphilus*, *Clostridium*, *Snethia amnii*, *Mycoplasma hominis, Porphyromonas*, *Facklamia languida* et *Gemella asaccharolytica*.

En utilisant la culturomique, nous avons isolé à partir de cette cohorte de 34 femmes, 340 espèces bactériennes distinctes renfermant des espèces connues du microbiote vaginal, des bactéries connues chez l'homme mais auparavant isolées dans d'autres sites que le tractus vaginal et 27 nouvelles espèces bactériennes ayant une similarité de la séquence codant pour l'ARNr 16S inférieure à 98,7% avec les espèces valides les plus proches dont 18 cultivées à partir du microbiote de vaginose bactérienne et 9 de flore vaginale normale. Des 18 nouvelles espèces isolées pour la première fois chez des patientes souffrant de vaginose bactérienne, trois : '*Peptoniphilus vaginalis'* (Similarité de séquences de 99,7% avec *Peptoniphilus* sp. DNF00840)*, 'Megasphaera vaginalis'* (99,73% avec *Megasphaera* sp. BV3C16-1) et '*Atopobium massiliense'* (99,7% et 99,8% avec respectivement *Atopobium* sp. S3MV26 et *Atopobium* sp. S4-5) sont étroitement apparentées à des bactéries détectées auparavant comme associées à la vaginose bactérienne en utilisant les outils moléculaires. La culturomique a donc permis d'isoler pour la première fois des espèces bactériennes qui jusque-là n'avaient été identifiées que par des techniques de biologie moléculaire.

La persévérance dans la culturomique, en augmentant le nombre d'échantillons vaginaux analysés et en diversifiant les zones d'étude géographique, mais aussi en utilisant de nouveaux milieux de culture imitant l'environnement vaginal, est nécessaire afin de pouvoir isoler cette part majeure du microbiote vaginal identifiée jusqu'à présent exclusivement par biologie moléculaire. L'obtention d'isolats permettra d'explorer *in vitro* les compétitions entre les bactéries du microbiote vaginal mais pourra servir également de matière première pour développer un traitement par bactériothérapie en proposant un cocktail de bactéries clés pour prévenir ou traiter la vaginose bactérienne.

Article 3:

Multi-Omics Strategy to Characterize Vaginal Microbiota associated with Bacterial Vaginosis: Culturomics and Metagenomics approaches.

Khoudia Diop, Ndeye Safietou Fall, Anthony Levasseur, Nafissatou Diagne, Dipankar Bachar, Florence Bretelle, Cheikh Sokhna, Jean-Christophe Lagier, Didier Raoult, Florence Fenollar

To be submitted soon in Microbiome Journal

Multi-Omics Strategy to Characterize Vaginal Microbiota associated with Bacterial Vaginosis: Culturomics and Metagenomics approaches

Khoudia DIOP[1], Ndeye Safietou FALL[2], Anthony LEVASSEUR[3], Nafissatou DIAGNE[2], Dipankar BACHAR[3], Florence BRETELLE[3,4], Cheikh SOKHNA[2], Jean-Christophe LAGIER[3], Didier RAOULT[3,5], and Florence FENOLLAR[1*]

[1]Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection, Marseille, France

[2]VITROME, Campus international UCAD-IRD Dakar, Sénégal

[3]Aix Marseille Univ, IRD, AP-HM, MEPHI, IHU-Méditerranée Infection, Marseille, France

[4]Department of Gynecology and Obstetrics, Gynépole, Hôpital Nord, AP-HM, Marseille, France

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

***Corresponding author**:
Prof. Florence Fenollar
Institut Hospitalo-Universitaire Méditerranée-Infection
19-21 Boulevard Jean Moulin
13005 Marseille cedex 05
France
Tel: + 33 413 732 401
Fax: + 33 413 732 402
Email: florence.fenollar@univ-amu.fr

Abstract count: 343

Word count: 4,216

28 **ABSTRACT**

29 ***Background*****:** Over the last decades, thanks to advances in technology including much more
30 efficient molecular techniques and new OMICS strategies, many studies have focused on the
31 vaginal microbiota, mainly bacterial vaginosis flora. They were complementarity to classical
32 microbiology and improve significantly knowledge on the vaginal ecosystem. Bacterial
33 vaginosis is a vaginal dysbiosis predisposing women to sexual transmitted infections. The
34 root cause of this imbalance is unknown and most often, antibiotic treatment fails. For these
35 reasons, the objective of this present study is to map exhaustively the bacterial community
36 present in bacterial vaginosis flora in order to better understand and manage this dysbiosis.
37 ***Methods*****:** The vaginal bacterial diversity of the patients with bacterial vaginosis (n =15) was
38 investigated and compared to those of healthy vaginal microbiota (n = 35). Samples (n = 50)
39 were collected from French (n = 24) and, rural-Senegal (n = 10). Microbiome profiles were
40 characterized using two OMICS strategies: one molecularly, the metagenomics and another
41 with different cultures conditions, the culturomics.
42 ***Results*****:** Combination of culturomics and metagenomics evinces the richness and diversity of
43 vaginal microbiota. Our data demonstrated that compared to healthy women, those with
44 bacterial vaginosis patients show a shift of vaginal flora. Despite interpersonal variations, the
45 vaginal microbiota of the bacterial vaginosis group can be easily distinguished to those of
46 healthy group by an increased bacterial diversity, abundance of *Bacteroidales* and
47 *Leptotrichiaeceae* and, depletion of *Proteobacteria* and *Lactobacillaceae* species. In addition,
48 a complex of 10 genera was associated with bacterial vaginosis: *Gardnerella, Atopobium,*
49 *Snethia, Urinancoccus, Aerococcus, Prevotella, Gemella, Facklamia, Porphyromonas* and
50 *Mycoplasma.* Comparing the data obtained by metagenomics with those of culturomics, the
51 two techniques make it possible to find diverse vaginal microbiota largely non-overlapping,

52 with only 146 common species. Moreover, culturomics extends the repertoire of human-
53 associated bacteria with the isolation of 27 new species.
54 ***Conclusions*:** This study provides the most representative topology of vaginal microbiota
55 structure. With culturomics, we isolated several sets of clinically significant anaerobic
56 bacteria and new species derived from human vagina, confirming that some previously
57 uncultivated species can be cultivated using an adequate strategy.
58
59 **Keywords**: Bacteria; Bacterial vaginosis; Culturomics; Metagenomics; Vaginal microbiota

BACKGROUND

Bacterial vaginosis is the most prevalent vaginal disorder in pubertal women [1]. It most often causes leucorrhoea with a foul odor and sometimes irritations that lead women to seek gynecological care [2]. It is associated with the risk of preterm birth, miscarriage, low birth weight among pregnant women [3], as well as those of pelvic inflammatory diseases [4]. Bacterial vaginosis predisposes women with serious health problems including the acquisition and transmission of various pathogens such as herpes simplex virus-2 [5], human immunodeficiency virus-1 [6], papillomavirus [7], *Neisseria gonorrhoeae*, *Trichomonas vaginalis*, and *Chlamydia trachomatis* [8, 9]. According to the studied population, its prevalence ranged from 4% to 75% between asymptomatic women and those with sexually transmitted diseases, respectively [10, 11]. To manage bacterial vaginosis condition and relieve symptoms, the Centers for Disease Control and Prevention recommend treating all patients with clinical manifestations [12] (homogeneous vaginal discharge adhering to the vaginal walls; unpleasant odor; vaginal pH > 4.5 and presence of clue cells [13]). However, 50% of affected women seem to have no symptoms and antibiotic therapy often fails with a 50% relapse rate after six months of treatment [14, 15].

A disturbance of the native bacterial flora present in the healthy vaginal tract is mainly observed in bacterial vaginosis. Under normal conditions, the healthy vagina of premenopausal women is quantitatively dominated by lactobacilli, approximately 70% to 90% of all bacteria [16]. Lactobacilli are lactic acid bacteria and some species can produce hydrogen peroxide to protect the vaginal tract and prevent the proliferation of other vaginal microorganisms [17, 18]. Previous studies have shown that during bacterial vaginosis, the vaginal tract flora undergoes significant changes ranging from predominantly *Lactobacillus* flora to a colonization without *Lactobacillus* [19]. The etiology of this gynecological condition is still unknown and, it remains one of the great enigmas on women health. To

85 understand whether modification of the vaginal microbiota can be linked with health, the
86 Human Microbiome Project (HMP) indicates that it is needful to examine the bacterial
87 communities of the healthy vaginal tract and those of the bacterial vaginosis [20].
88 In the past, the vaginal bacterial communities have been identified using culture-based
89 methods. With the inability to isolated most of the bacterial species present in the vagina, the
90 taxonomic composition of this ecosystem was not well understood. Over the last decades,
91 many studies have focused on the vaginal microbiota thanks to the advances in technology
92 such as more efficient molecular techniques (based mostly on 16S rRNA gene sequencing).
93 They revealed that vaginal microbiota is more complex than previously thinking and bacterial
94 vaginosis is a polymicrobial syndrome. Indeed, bacterial vaginosis is typified by the loss of
95 lactobacilli and overgrowth or introduction of anaerobic bacteria and Bacterial Vaginosis
96 (BV)-associated bacteria previously uncultivated using culture methods [21]. For a better
97 understanding of this public health problem, the return of culture is necessary to isolate and
98 study these BV-associated bacteria observed only with molecular methods. This renaissance
99 of culture is observed with the arrival of microbial culturomics, an approach multiplying the
100 conditions of culture and variation of physicochemical parameters coupled with a rapid
101 identification of bacteria by MALDI-TOF (Matrix Assisted Laser Desorption Ionisation-Time
102 of Flight) mass spectrometry [22, 23].
103 In this study, using metagenomics targeting 16S rRNA gene and culturomics, we
104 investigated the vaginal bacterial diversity of normal and bacterial vaginosis floras in French
105 and, rural-Senegal women in order to understand the dysbiosis of vaginal microbiota during
106 bacterial vaginosis. Ours results highlight in the one hand the complementarity of these two
107 approaches and, the other hand the diversity and richness of the vaginal microbiota. The use
108 of culturomics has extended the repertoire of human-associated bacteria with the isolation of
109 many new bacterial species in the vaginal ecosystem.

110 # RESULTS

111 **Characteristics of the studied population and profile of the vaginal microflora**

112 The present study includes 34 women aged 20 to 50-year-old (28.53±5.74), for whom
113 the vaginal flora was characterized (Table 1). Ten vaginal specimens from 5 patients with
114 bacterial vaginosis and from 5 healthy women were sampled in Senegal. Seven vaginal
115 samples from 5 patients with bacterial vaginosis and seventeen from healthy women were
116 collected in France. The culture of these 34 vaginal samples allowed the isolation and
117 identification of 340 bacteria covering 7 phyla, 51 families, and 128 genera (Table S1). On
118 the other hand, sixteen additional samples (3 diagnosed as bacterial vaginosis and 13 as
119 normal flora) were also collected from 5 French women. All the 50 specimens (34 samples +
120 16 additional) were analyzed by metagenomics barcoded the "V3-V4 region" of the gene 16S
121 rRNA. A total of 1,215,586 and 2,946,743 reads were generated from the 15 samples with
122 bacterial vaginosis and from the 35 normal vaginal flora, respectively. These sequences
123 correspond to 1,246 OTUs of which 1,229 bacteria assigned to 169 genera, 70 families, and 9
124 phyla (Table S2). Only 383 of these OTUs (31%) were classified at the species level. In the
125 bacterial vaginosis group, 46% of these OTUs corresponded to *Gardnerella vaginalis* (26%),
126 *Atopobium vaginae* (12%), and *Lactobacillus iners* (8%). Whereas *Lactobacillus crispatus*
127 (15%), *Bacillus simplex* (11%), *Escherichia coli* (10%), *L. iners*, and *G. vaginalis* (7%, of
128 each) represented the major OTUs identified in healthy women (Table S3). Overall, only
129 *Bifidobacterium scardovii* and *Facklamia ignava* were detected using both metagenomics and
130 culturomics only among patients with bacterial vaginosis. In addition, a panel of 15 species
131 (*Acidaminococcus intestini, Acinetobacter baumannii, Alistipes putidinis, Bacteroides*
132 *cellulosilyticus, B. fragilis, B. salyersiae, Enterococcus pallens, Lactobacillus mucosae,*
133 *Macrococcus caseolyticus, Morganella morganii, Phascolarctobacterium faecium,*
134 *Pseudoramibacter alactolyticus, Streptococcus australis, Streptococcus urinalis*, and

135 *Trueperella bernardiae)* was only found among normal vaginal flora using these 2 techniques
136 (Table 1).

137 **High vaginal microbiota diversity in flora with bacterial vaginosis**

138 A total of 1,194,818 and 2,484,424 reads were generated from the 12 vaginal flora with
139 bacterial vaginosis and the 22 normal, respectively. Estimation of α-diversity showed that
140 vaginal microbiomes of women with bacterial vaginosis were richer (ACE 34.8±1.7, Chao-1
141 29.5±16.5, Table 1) and more diverse (Shannon index 1.9 ± 0.7, Simpson index 0.3 ± 0.2,
142 Figure 1) than those of healthy women (28.7±1.5, 22.5±15.5, 1.3±0.8, and 0.5±0.3,
143 respectively). The bacterial communities were also more abundant during bacterial vaginosis
144 than in a normal state (evenness 0.14 in bacterial vaginosis vs 0.09 in normal vaginal flora,
145 p=0.009, unpaired t-test). Moreover, the hitherto unknown diversity (unclassified OTUs) was
146 significantly increased in patients with bacterial vaginosis (p=0.03, unpaired t-test, Figure 1).

147 According to culturomics analysis, β-diversity was higher in the bacterial vaginosis
148 group [115/241 (48%)] than in the normal vaginal flora group [110/261 (42%)]. Among the
149 patients with bacterial vaginosis, 43% of isolated species (103/241) were not previously
150 known to be present in the human vagina, including 14 new bacterial species and 4 new
151 genera (Table 2 [24–35]). Among healthy women with normal vaginal flora, 46% (119/261)
152 species were not known from the human vagina including 7 new species and 2 new genera.
153 The hitherto unknown diversity (new species) detected by culturomics increased considerably
154 during bacterial vaginosis (p=0.03, Mann Whitney test, Figure 1).

155 **Dynamic modification of vaginal microbiota during vaginosis**

156 Given the strong evidence that the structure of the vaginal microbiome differs strikingly
157 between healthy women and those with bacterial vaginosis, we sought to identify differences
158 that occur during this dysbiosis. Metagenomics analysis of samples from women with
159 bacterial vaginosis and those from healthy women generated 1,118,379 and 2,466,547 reads

160 assigned at the species level, respectively. The reads matched with a total of 360 bacterial
161 species (208 species for bacterial vaginosis and 307 for normal vaginal flora), which were
162 classified into 9 phyla: *Actinobacteria, Bacteroidetes, Epsilonbacteraeota, Firmicutes,*
163 *Fusobacteria, Proteobacteria, Synergistetes* (only in the healthy vaginal flora)*, Tenericutes,*
164 and *Verrucomicrobia* (only in bacterial vaginosis). The vaginal microbiota from bacterial
165 vaginosis includes significantly more *Actinobacteria* (67/208 versus 68/307, p=0.01, exact
166 Fischer test) and fewer *Proteobacteria* (16/208 versus 64/307, p=3.10^{-4}) than those of healthy
167 flora.

168 Interestingly, LEfSe analysis revealed that the significant abundance of *Actinobacteria*
169 in women with bacterial vaginosis was due to the increased number of bacteria belonging to
170 *Bifidobacteriaceae*. In addition, the relative abundance of *Leptotrichiaceae* (phylum
171 *Fusobacteria*) and *Bacteroidales* (phylum *Bacteroidetes*) microorganisms were also increased
172 during bacterial vaginosis. While in healthy women, three clades were significantly
173 overrepresented: two belonging to *Proteobacteria* (*Enterobacteriaceae*/*Pseudomonadaceae*
174 and *Sutterellaceae* families) and one clade to *Lactobacillaceae* (Figure 2A). An LDA score
175 for species-level abundance showed that 15 species were found to change during bacterial
176 vaginosis (Figure 2B), 11 of which increased (*G. vaginalis, A. vaginae, Snethia amnii,*
177 *Urinancoccus timonensis, Aerococcus christensenii, Prevotella amnii, Gemella*
178 *asaccharolytica, Facklamia languida, Porphyromonas asaccharolytica, P. somerae,* and
179 *Mycoplasma hominis)* and 4 decreased (*Staphylococcus haemolyticus, Escherichia coli,*
180 *Lactobacillus crispatus*, and *Bacillus simplex*). The remark is that all the diminished taxa were
181 aerotolerant species while all the augmented taxa were strictly or facultative anaerobes.

182 Microbial culturomics of healthy vaginal samples allowed the isolation of 261 bacteria
183 distributed into 7 phyla with predomination of *Firmicutes* (139), followed by 65
184 *Actinobacteria*, 40 *Bacteroidetes*, 14 *Proteobacteria*, and 1 species from each of the following

185 phyla *Epsilonbacteraeota*, *Fusobacteria,* and *Synergistetes*. Isolated bacteria were classified
186 into 108 genera including *Lactobacillus* (15), *Staphylococcus* (13), *Anaerococcus* (11),
187 *Peptoniphilus* (11), *Streptococcus* (11), *Actinomyces* (10), *Bacteroides* (10), and *Prevotella*
188 (10). On the other hand, species isolated in women with bacterial vaginosis were classified
189 into 7 phyla: 133 *Firmicutes*, 59 *Actinobacteria*, 31 *Bacteroidetes*, 9 *Proteobacteria*, 2
190 *Fusobacteria*, 2 *Synergistetes*, and 1 *Epsilonbacteraeota,* for a total of 94 genera. The major
191 genera were *Streptococcus* (13)*, Peptoniphilus* (12), *Anaerococcus* (11), *Staphylococcus* (11)*,*
192 *Clostridium* (10)*,* and *Prevotella* (10). Bacteria cultured from the microbiota of women with
193 bacterial vaginosis belonged mostly to the *Firmicutes* phylum (p=0.007, unpaired t-test) with
194 a dominance of *Clostridium* and *Peptoniphilus* spp ($p \leq 0.02$, Mann Whitney test) (Table 3).
195 Thus, completing metagenomics results, in addition to a decrease of lactobacilli and
196 *Proteobacteria,* women with bacterial vaginosis showed also an increase of *Clostridium* and
197 *Peptoniphilus* spp.
198 **Variation of vaginal microbiome composition according to area**
199 To investigate if there is an association between bacterial communities and ethnicity,
200 vaginal microbiomes of French women were compared to those of Senegalese women (Figure
201 3).
202 According to metagenomics results, rural Senegalese women exhibited a richer
203 microbiota ($p < 0.05$, two groups parametric ANOVA for Chao-1, and ACE) than that of
204 Caucasian women, but there is no significant difference in diversity ($p > 0.05$ for Shannon and
205 Simpson indices) (Table 4). Strikingly, OTUs belonging to *Bacillus*, *Escherichia*,
206 *Staphylococcus,* and *Corynebacterium* were abundant in Caucasian healthy women whereas,
207 rural Senegalese healthy women harbored diverse anaerobic species such as *Prevotella*,
208 *Ureaplasma,* and *Peptoniphilus* (Figure 4).

209 Culturomics results exhibited that vaginal microbiota was significantly more diverse
210 and richer in Caucasian women. The number of isolated species in women with bacterial
211 vaginosis (63.29±11.22 versus 30.6±11.48, $p<0.0001$) and healthy vagina (43±10.33 versus
212 24.80±12.48, $p=0.002$) increased considerably in Caucasian with many anaerobic bacteria
213 (Table 5).

214 **DISCUSSION**

215 Recently, OMICS strategies have provided an overview of bacterial communities'
216 composition, their function, and interaction with host cells. Despite being revolutionary, they
217 do not replace knowledge gained with the isolation of bacterial microorganisms. Knowledge
218 about the composition and diversity of the human vaginal microbiota in health and bacterial
219 vaginosis conditions may be useful for better understanding this dysbiosis and for preventing
220 or controlling it. Our study of 15 vaginal flora with bacterial vaginosis and 35 normal flora,
221 aimed to map exhaustively the vaginal microbiota and to understand the putative link between
222 the bacterial vaginosis condition and the microbial composition and diversity. In this goal, we
223 performed in parallel an already applied strategy to vaginal flora, the metagenomics [36, 37]
224 and a new one, the culturomics [22]. One of the major findings is that the composition of
225 vaginal microbiota changes considerably during bacterial vaginosis with a significant increase
226 in bacterial diversity. Indeed, some studies have previously shown that bacterial vaginosis
227 was obviously associated with more richer and diverse vaginal bacterial community [1, 38,
228 39], as corroborated in our study.

229 Of the 9 phyla found in the vaginal samples using both culturomics and metagenomics,
230 *Firmicutes* was largely the major phylum detected in both microbiota (healthy and bacterial
231 vaginosis). Our results have revealed a significant decrease of *Lactobacillaceae* and
232 *Proteobacteria* species in the meantime, bacterial taxa such as *Gardnerella vaginalis*,
233 *Atopobium vaginae*, *Aeroccocus christensenii, Prevotella, Peptoniphilus, Clostridium,*

234 *Snethia amnii*, *Mycoplasma hominis, Porphyromonas, Facklamia languida,* and *Gemella*
235 *asaccharolytica* increased. It is not a surprise that the presence of bacterial vaginosis is
236 marked by a depletion of protective *Lactobacillus* species and aerotolerant bacteria [17, 40–
237 42] causing a transition of the vaginal microbiota from an eubiosis status to dysbiosis [43].
238 Moreover, most of the non-*Lactobacillus* species found here in high prevalence among
239 patients with bacterial vaginosis have been already reported to be associated with bacterial
240 vaginosis [1, 38, 44–47]. The first detection of *Facklamia languida* extends the number of
241 BV-associated bacteria reported in the literature up to now [44–48]. Interestingly, *Facklamia*
242 species have been recently associated with the etiology of numerous clinical infections [49].
243 The microbial population detected in bacterial vaginosis flora was greatly similar to bacterial
244 communities found in human feces, with the presence of bacteria belonging to the genera
245 *Prevotella, Peptoniphilus*, and *Clostridium* [50]. Our data provided some insights to some
246 scientists interrogations [51, 52], with the hypothesis that the presence of gastrointestinal flora
247 associated to bacterial vaginosis may be due to their inoculation during certain sexual
248 practices. Thus, bacterial vaginosis is a dysbiosis that may result from the introduction into
249 the vagina of bacteria from another source, causing a change in pH which leads to a loss of
250 vaginal equilibrium. Then, lactobacilli are depleted, and the vaginal ecosystem becomes
251 favorable to the proliferation of many resident anaerobic species such as *A. vaginae* and *G.*
252 *vaginalis*.
253 Using metagenomics, we noticed that the vaginal microbiota of Senegalese women was
254 richer than that of Caucasian, with an abundance of anaerobic species. These results concord
255 with previous literature reports showing differential bacterial diversity between Caucasian and
256 African women [53–55]. The culturomics results on the same samples revealed the opposite
257 but this difference between the data from metagenomics and culturomics can be explained by
258 the lag time between the samples collection time and their culture. Indeed, French vaginal

259 samples were immediately inoculated upon collection, allowing thus the isolation of a wide
260 range of bacteria which were mainly anaerobic, while those from rural Senegalese women
261 were stored at -80°C few months before culturing. These data underline that storage and
262 transport constitute a pitfall of culture strategies, with the loss and death of bacterial species
263 especially those anaerobic. By testing the viability of vaginal microorganisms in 2 and 3
264 commercial transport media, respectively at different temperatures and time points, Stoner et
265 *al.*, [56] and DeMarco et *al.,* [57] noted that the microorganisms grown dependent on the
266 transport media used and also temperature and time elapsed before analysis. They indicated
267 that to prevent proliferation during transport and maintain vaginal anaerobic bacteria, culture
268 must be processed within 24 hours of 4°C storage. Despite the limits of culture, it is important
269 to underline that 7 new species were obtained from the 10 frozen Senegalese samples (Table
270 2). Overall, it would be better to associate inoculation of the sample collected at the 'patient's
271 bed' with its preincubation in a culture bottle in order to isolate extremely sensitive bacteria.
272 Several studies have tried to characterize the vaginal microbiota using different
273 strategies such as culture as well as molecular techniques [45, 58–60]. With metagenomic
274 investigations, the number of phylotypes (previously uncultivated or new undiscovered) in the
275 vagina microbiota was higher in women with bacterial vaginosis than in healthy women and
276 were similar to previous molecular studies of vaginal microbiota that detected numerous
277 uncultivated BV-associated bacteria such as BVAB1, BVAB2, *Megasphaera* sp type 1 and 2,
278 *Dialister* sp type 1 to 3, and *Eggerthella* sp type 1 [45, 46, 60, 61]. Our study is the first
279 attempt to characterize the vaginal microbiota using the culturomics strategy. Our findings
280 revealed the complementary of metagenomics and culturomics approaches allowing the
281 detection of a total of 554 bacteria (with 360 identified by metagenomics and 340 isolated by
282 culturomics, and only 146 common species detected using both techniques).

283 Culturomics extended the repertoire of vaginal flora with the isolation of 142 bacterial
284 species already detected in human but never in the vaginal flora. Overall, 27 new bacterial
285 species of which 3: *'Peptoniphilus vaginalis'*, *'Megasphaera vaginalis'* and *'Atopobium*
286 *massiliense'* closely related to *Peptoniphilus* sp. DNF00840, *Megasphaera* sp. BV3C16-1 and
287 *Atopobium* sp. S4-5, respectively that correspond to BVABs and detected previously using
288 only molecular tools. Contrary classical culture method [59], culturomics has shown this
289 prowess in bacterial isolation by almost doubling the number of cultivated species in the
290 human gut [62, 63].

291 **CONCLSION**

292 During imbalance of microbiota, it is important to know the impact of sampling and,
293 exploration techniques on the microbial community. In spite of its limitations, this study is the
294 first to investigate the vaginal microbiota of women, with and without bacterial vaginosis, in
295 France and Senegal, and also the first to use the culturomics. The metagenomics strategy
296 targeting 16S rRNA gene paired with the culturomics strategy highlights the richness and
297 diversity of the vaginal microbiota. Although culturomics do not cover all taxa in this
298 microbial econiche, it has been successfully applied to isolate several sets of bacteria
299 including 3 keys members of bacterial vaginosis flora: *Peptoniphilus* sp. DNF00840,
300 *Megasphaera* sp. BV3C16-1, and *Atopobium* sp. S4-5, which were only detected by
301 molecular tools. In the future, it would be interesting to persevere in culturomics, increasing
302 the number of analyzed vaginal samples and diversifying the geographical study areas, but
303 also using new media mimicking the vaginal environment. Obtaining isolates will allow
304 exploring *in vitro* the competitions between bacteria from vaginal microbiota but will also
305 serve as the raw material to develop a treatment by bacteriotherapy by proposing a cocktail of
306 key bacteria to prevent or treat bacterial vaginosis.

307 **MATERIALS AND METHODS**

308 **Study design**

309 This study focused on investing the vaginal microbiota of healthy women and those
310 suffering from bacterial vaginosis. The project included 34 women living in two geographical
311 areas: 24 from France (Public University Hospitals, Marseille) and 10 from rural Senegal
312 (villages of Dielmo and Ndiop, Sine-Saloum area). Only non-pregnant, HIV-negative, 18- to
313 50-year-old pre-menopausal women who received no antibiotic treatment in the 2 months
314 preceding the study were eligible to participate.

315 The ethics committees of the Institut Fédératif de Recherche IFR48 (Marseille, France)
316 and that of the Senegalese CNERS in accordance with the SEN protocol 16/04 validated this
317 study under agreement numbers 09-022 and 00039, respectively. All participants were
318 volunteers and gave informed written and signed consent.

319 **Sample collection and study process**

320 Women collected themselves their own vaginal discharges [64, 65] using Sigma
321 Transwab (Medical Wire, Corsham, United Kingdom). For French samples, a fresh culture
322 was made immediately within minutes of collecting. For swabs sampled in rural Senegal, they
323 were stored and transported to the laboratory in Dakar, the capital, in a portable freezer at -
324 20°C. As soon as they arrived, they were stored at -80°C until they were sent to Marseille in
325 dry ice. Once in Marseille, they were stored at -80°C until further analysis. Besides, we also
326 collected from 5 of the French women 16 additional vaginal specimens at follow-up visits (5,
327 5, 3, 2 and, 1 respectively). Among the 50 specimens included in the study, all were analyzed
328 using metagenomics targeting 16S rRNA gene, but only 34 were analyzed using microbial
329 culturomics. The diagnosis of bacterial vaginosis was assessed by molecular quantification of
330 the microorganisms *Atopobium vaginae* and *Gardnerella vaginalis*, as previously described,
331 [66] in parallel to Nugent score [67].

332 **High-throughput culture-based technique: Culturomics**

333 ***Culture conditions and isolation of bacteria***. Vaginal bacterial communities were isolated
334 using the culturomics concept [22, 68]. Firstly, the samples were vortexed in 3 mL of
335 Transwab, and 100 µL of the resulting broth was diluted in 900 µL of Dulbecco's phosphate-
336 buffered saline (DPBS). Ten-fold cascade dilutions were performed and, aliquots of 50 µL of
337 each dilution were directly seeded onto five solid culture media: anaerobe basal agar (Oxoid,
338 Dardilly, France), Colistin Nalidixic Acid agar, Columbia sheep blood agar, Schaedler agar
339 enriched with sheep blood and vitamin K1, and Trypticase soy agar with horse blood (all four,
340 BD Diagnostics, Le Pont-de-Claix, France). The Petri dishes were incubated at 37°C under
341 anaerobic atmosphere for 4 to7 days. Then, briefly, 2 mL of the vaginal fluid were pre-
342 incubated in both anaerobic and aerobic liquid culture medium (BD Diagnostics)
343 supplemented with sterile blood (3 mL) and filtered rumen (4 mL), both from sheep. At
344 different pre-incubation periods (1, 3, 7, 10, 15, 21, and 30 days) at 37°C, 100 µL of the broth
345 was sampled using the format plating described below. The anaerobic bottle mixture was
346 inoculated on the 5 media cited above then incubated for 7 days under the same condition
347 whereas aerobic broth was plated on Chocolate agar PolyViteX (BD Diagnostics) and
348 incubated aerobically for 3 days at 37°C. Isolated bacteria were purified and then identified
349 using MALDI–TOF mass spectrometry with a Microflex spectrometer (Bruker, Leipzig,
350 Germany) [69, 70].
351 ***Rapid bacterial identification using MALDI-TOF* mass spectrometry**. Each purified colony
352 was spotted in duplicate on a 96 MALDI-TOF target plate and covered with 2 µL of matrix
353 solution, as previously reported [69, 70]. The bacterial identification was performed using the
354 Microflex spectrometer which compares the obtained protein spectra with those of present in
355 the library (Bruker and URMS databases). Isolates with an unambiguous score ≥ 2.0 were
356 considered correctly identified at species level. For unidentified bacteria using MALDI-TOF

357 mass spectrometry (score ≤ 2.0), 16S rRNA gene sequencing was performed to determine its
358 taxonomic position [71].
359 ***Bacterial identification based on 16S rRNA gene sequencing***. Bacterial DNA was extracted
360 using the EZ1 DNA Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's
361 instructions. The DNA was subjected to an amplification using primers FD1 and RP2
362 (Eurogentec, Angers, France) targeting a conserved bacteria region and an annealing
363 temperature of 52°C. The amplified product was verified by electrophoresis gel migration
364 prior to its purification and re-amplification using the BigDye Terminator sequencing kit
365 (Qiagen, Courtaboeuf, France) with a system of 8 primers (357F, 358R 536F, 536R, 800F,
366 800R, 1050 F, and 1050R). The amplicons were purified and sequenced using an ABI PRISM
367 3130-XL capillary sequencer (Applied Biosystems, Bedford, MA, USA). The obtained
368 sequences were corrected and compared to those available in the GenBank database.
369 Sequences with a nucleotide similarity percentage below 98.7 and 95% were considered as
370 new species or genus, respectively [72, 73].
371 **High-throughput molecular method: Metagenomics**
372 ***Extraction and metagenomics sequencing***. DNA from vaginal samples was extracted by a
373 mechanical treatment performed with powder glass beads acid washed (Sigma, Lyon, France)
374 and 0.5 mm glass beads cell disruption media (Scientific Industries, Bohemia, NY, USA)
375 using a FastPrep BIO 101 instrument (Qbiogene, Strasbourg, France) at maximum speed (6.5
376 m/sec) for 90 sec. Then, the specimens were treated through two kinds of lyses methods: a
377 method with classical lysis and protease step following by purification on NucleoSpin Tissue
378 kit (Macherey Nagel, Hoerdt, France) and another using a deglycosylation step and
379 purification on the EZ1 Advanced XL device (Qiagen) [34]. Samples were first amplified on
380 these 2 extractions, pooled and barcoded, then sequenced for 16S rRNA sequencing on MiSeq
381 technology (Illumina, San Diego, CA, USA) with paired-end strategy, constructed according

382 to the 16S metagenomic sequencing library preparation (Illumina). For each protocol
383 extraction, metagenomic DNA was sequenced for the 16S rRNA gene's "V3-V4" regions as
384 previously described [74]. All reads from these two methods were grouped and clustered with
385 a threshold of 98% identity to obtain operational taxonomic units (OTUs). The paired reads
386 were filtered according to the read qualities.
387 ***Taxonomic assignments of OTUs***. The paired-end sequences were assembled into longer
388 sequences by Pandaseq [75]. The resulting fastq files of longer sequences ≥ 400nts were then
389 demultiplexed in the QIIME2 pipeline (Quantitative Insights Into Microbial Ecology) version
390 2018.2 [76]. The sequence quality was controlled and feature/OTU table was constructed in
391 DADA2 [77] of QIIME2. The OTUs/Feature sequences were blasted [78] against the
392 reference database of SSURef of Silva [79] and taxonomy was assigned by applying majority
393 voting [74, 80], considering species level ≥ 98% identity and 100% coverage. The unassigned
394 OTUs were then blasted against the IHU (Institut Hospitalo-Universitaire) database
395 containing all species isolated by culturomics. OTUs that remain not assigned to any species
396 were classed "unclassified".
397 **Data and statistical analyses**
398 For metagenomics data, α-diversity (ACE, Chao-1, Shannon, and Simpson indices) was
399 calculated using MicrobiomeAnalyst [81]. Based on their differential abundance, microbial
400 markers were determined by using the Ward clustering method based on the Euclidean
401 distance [81]. The bacterial abundance profile among our two groups (healthy women and
402 women with bacterial vaginosis) were identified using linear discriminant analysis (LDA)
403 effect size (LEfSe) methods [82]. Then, the statistically different features were computed as a
404 cladogram using GraPhlAn [83]. β-diversity (comparison of the number of taxa unique for
405 each ecosystem) was also calculated for culture data. It constituted the ratio of the
406 unique/totality of the microbiota of every group. Depending on the Gaussian distribution, *t*-

407 test or Mann Whitney test was used to compare quantitative data, and exact Fischer or Chi-
408 squared tests for proportions. Statistical analyses were performed using GraphPad Prism
409 version 5.03 and values were presented as mean and standard deviation [84].
410
411
412 **Acknowledgments**
413 The authors express their sincere gratitude to P. Perrier and B. Nicaise for their help in the
414 collection of samples and clinical data. The authors are also grateful to PE. Fournier, F. Bittar,
415 E. Tomei, and F. Cadoret.
416
417 **Authors' contributions**
418 KD, JCL, DR, and FF conceived and designed the experiments. AL, NSF, ND, FB, CS, and
419 FF contributed to the materials and analysis tools. KD and NSF perform the culturomics. KD,
420 NSF, DB, AL, JCL, and FF analyzed and interpreted the data. KD and FF wrote the paper. All
421 authors read and approved the final manuscript.
422
423 **Funding information**
424 This study was funded by the Institut Hospitalo-Universitaire (IHU) Méditerranée Infection,
425 the National Research Agency under the program "Investissements d'avenir", reference ANR-
426 10-IAHU-03, the Region Provence Alpes Côte d'Azur and European funding FEDER PRIMI.
427
428 **Conflict of interest**
429 The authors have no conflicts of interest to declare. Funding sources had no role in the design
430 and conduct of the study; collection, management, analysis, and interpretation of the data; and
431 preparation, review, or approval of the manuscript.

432 **REFERENCES**

433 1. **Ling Z, Kong J, Liu F, Zhu H, Chen X, *et al.*** Molecular analysis of the diversity of vaginal
434 microbiota associated with bacterial vaginosis. *BMC Genomics* 2010;11:488.

435 2. **Bagnall P, Rizzolo D**. Bacterial vaginosis: A practical review. *Journal of the American*
436 *Academy of Physician Assistants* 2017;30:15–21.

437 3. **Afolabi BB, Moses OE, Oduyebo OO**. Bacterial Vaginosis and Pregnancy Outcome in
438 Lagos, Nigeria. *Open Forum Infectious Diseases* 2016;3:ofw030.

439 4. **Babu G, Singaravelu BG, Srikumar R, Reddy SV, Kokan A**. Comparative Study on the
440 Vaginal Flora and Incidence of Asymptomatic Vaginosis among Healthy Women and in Women with
441 Infertility Problems of Reproductive Age. *J Clin Diagn Res* 2017;11:DC18–DC22.

442 5. **Nardis C, Mosca L, Mastromarino P**. Vaginal microbiota and viral sexually transmitted
443 diseases. *Ann Ig* 2013;25:443–56.

444 6. **McClelland RS, Lingappa JR, Srinivasan S, Kinuthia J, John-Stewart GC, *et al.***
445 Evaluation of the association between the concentrations of key vaginal bacteria and the increased risk
446 of HIV acquisition in African women from five cohorts: a nested case-control study. *Lancet Infect Dis*.
447 Epub ahead of print 25 January 2018. DOI: 10.1016/S1473-3099(18)30058-6.

448 7. **Kero K, Rautava J, Syrjänen K, Grenman S, Syrjänen S**. Association of asymptomatic
449 bacterial vaginosis with persistence of female genital human papillomavirus infection. *European*
450 *Journal of Clinical Microbiology & Infectious Diseases* 2017;36:2215–2219.

451 8. **Brotman RM**. Vaginal microbiome and sexually transmitted infections: an epidemiologic
452 perspective. *Journal of Clinical Investigation* 2011;121:4610–4617.

453 9. **Bitew A, Abebaw Y, Bekele D, Mihret A**. Prevalence of Bacterial Vaginosis and Associated
454 Risk Factors among Women Complaining of Genital Tract Infection. *International Journal of*
455 *Microbiology* 2017;2017:1–8.

456 10. **Onderdonk AB, Delaney ML, Fichorova RN**. The Human Microbiome during Bacterial
457 Vaginosis. *Clinical Microbiology Reviews* 2016;29:223–238.

458 11. **Bitew A, Abebaw Y, Bekele D, Mihret A**. Prevalence of Bacterial Vaginosis and Associated
459 Risk Factors among Women Complaining of Genital Tract Infection. *International Journal of*
460 *Microbiology* 2017;2017:1–8.

461 12. Bacterial Vaginosis - 2015 STD Treatment Guidelines.
462 https://www.cdc.gov/std/tg2015/bv.htm (accessed 15 April 2018).

463 13. **Amsel R, Totten PA, Spiegel CA, Chen KCS, Eschenbach D, *et al.*** Nonspecific vaginitis.
464 *The American Journal of Medicine* 1983;74:14–22.

465 14. **Bradshaw CS, Tabrizi SN, Fairley CK, Morton AN, Rudland E, *et al.*** The association of
466 Atopobium vaginae and Gardnerella vaginalis with bacterial vaginosis and recurrence after oral
467 metronidazole therapy. *Journal of Infectious Diseases* 2006;194:828–836.

468 15. **Bretelle F, Rozenberg P, Pascal A, Favre R, Bohec C, *et al.*** High Atopobium vaginae and

469 Gardnerella vaginalis Vaginal Loads Are Associated With Preterm Birth. *Clinical Infectious Diseases*
470 2015;60:860–867.

471 16. **Africa C, Nel J, Stemmet M**. Anaerobes and Bacterial Vaginosis in Pregnancy: Virulence
472 Factors Contributing to Vaginal Colonisation. *International Journal of Environmental Research and*
473 *Public Health* 2014;11:6979–7000.

474 17. **Lepargneur JP, Rousseau V**. [Protective role of the Doderleïn flora]. *J Gynecol Obstet Biol*
475 *Reprod (Paris)* 2002;31:485–494.

476 18. **Marrazzo JM, Hillier SL**. Bacterial Vaginosis. In: *Sexually Transmitted Diseases*. Elsevier.
477 pp. 463–498.

478 19. **Hay PE**. Bacterial Vaginosis as a Mixed Infection. In: *Polymicrobial Diseases*. Washington
479 (DC): ASM Press. https://www.ncbi.nlm.nih.gov/books/NBK2495/ (2002, accessed 1 March 2018).

480 20. **Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett C, Knight R, *et al.*** The human
481 microbiome project: exploring the microbial part of ourselves in a changing world. *Nature*
482 2007;449:804–810.

483 21. **van de Wijgert JHHM, Borgdorff H, Verhelst R, Crucitti T, Francis S, *et al.*** The Vaginal
484 Microbiota: What Have We Learned after a Decade of Molecular Characterization? *PLoS ONE*
485 2014;9:e105998.

486 22. **Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, *et al.*** The Rebirth of Culture in
487 Microbiology through the Example of Culturomics To Study Human Gut Microbiota. *Clinical*
488 *Microbiology Reviews* 2015;28:237–264.

489 23. **Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, *et al.*** Culture of previously
490 uncultured members of the human gut microbiota by culturomics. *Nature Microbiology* 2016;1:16203.

491 24. **Diop K, Diop A, Bretelle F, Cadoret F, Michelle C, *et al.*** Olegusella massiliensis gen. nov.,
492 sp. nov., strain KHD7 T , a new bacterial genus isolated from the female genital tract of a patient with
493 bacterial vaginosis. *Anaerobe* 2017;44:87–95.

494 25. **Diop K, Bretelle F, Michelle C, Richez M, Rathored J, *et al.*** Taxonogenomics and
495 description of Vaginella massiliensis gen. nov., sp. nov., strain Marseille P2517 T , a new bacterial
496 genus isolated from the human vagina. *New Microbes and New Infections* 2017;15:94–103.

497 26. **Diop K, Diop A, Khelaifia S, Robert C, Pinto FD, *et al.*** Characterization of a novel Gram-
498 stain-positive anaerobic coccus isolated from the female genital tract: Genome sequence and
499 description of *Murdochiella vaginalis* sp. nov. *MicrobiologyOpen* 2018;e00570.

500 27. **Diop K, Nguyen TT, Delerce J, Armstrong N, Raoult D, *et al.*** Corynebacterium fournierii
501 sp. nov., isolated from the female genital tract of a patient with bacterial vaginosis. *Antonie van*
502 *Leeuwenhoek* 2018;1–10.

503 28. **Diop K, Andrieu C, Michelle C, Armstrong N, Bittar F, *et al.*** Characterization of a New
504 Ezakiella Isolated from the Human Vagina: Genome Sequence and Description of Ezakiella
505 massiliensis sp. nov. *Current Microbiology*. Epub ahead of print 29 November 2017. DOI:
506 10.1007/s00284-017-1402-z.

507 29. **Diop K, Diop A, Levasseur A, Mediannikov O, Robert C, *et al.*** Microbial Culturomics
508 Broadens Human Vaginal Flora Diversity: Genome Sequence and Description of *Prevotella lascolaii*
509 sp. nov. Isolated from a Patient with Bacterial Vaginosis. *OMICS: A Journal of Integrative Biology*.
510 Epub ahead of print 15 February 2018. DOI: 10.1089/omi.2017.0151.

511 30. **Diop K, Diop A, Michelle C, Richez M, Rathored J, *et al.*** Description of three new
512 Peptoniphilus species cultured in the vaginal fluid of a woman diagnosed with bacterial vaginosis:
513 Peptoniphilus pacaensis sp. nov., Peptoniphilus raoultii sp. nov., and Peptoniphilus vaginalis sp. nov.
514 *MicrobiologyOpen*;0:e00661.

515 31. **Diop K, Bretelle F, Fournier PE, Raoult D, Fenollar F**. "Collinsella vaginalis" sp. nov., a
516 new bacterial species cultivated from human female genital tract. *Human Microbiome Journal*
517 2016;2:19–20.

518 32. **Diop A, Raoult D, Fenollar F, Fournier PE**. 'Khoudiadiopia massiliensis' gen. nov., sp.
519 nov., strain Marseille-P2746TT, a new bacterial genus isolated from the female genital tract. *New*
520 *Microbes New Infect* 2017;19:60–61.

521 33. **Diop K, Bretelle F, Fournier P, Fenollar F**. ' Anaerococcus mediterraneensis ' sp. nov., a
522 new species isolated from human female genital tract. *New Microbes and New Infections* 2017;17:75–
523 76.

524 34. **Nicaise B, Bilen M, Cadoret F, Bretelle F, Fenollar F**. 'Lactobacillus raoultii' sp. nov., a
525 new bacterium isolated from the vaginal flora of a woman with bacterial vaginosis. *New Microbes*
526 *New Infect* 2018;21:20–22.

527 35. **Fall NS, Raoult D, Sokhna C, Lagier JC**. 'Helcococcus massiliensis' sp. nov., a new
528 bacterial species isolated from the vaginal sample of a woman with bacterial vaginosis living in
529 Dielmo, Senegal. *New Microbes New Infect* 2018;25:27–29.

530 36. **Angelakis E, Bachar D, Henrissat B, Armougom F, Audoly G, *et al.*** Glycans affect DNA
531 extraction and induce substantial differences in gut metagenomic studies. *Scientific Reports*;6. Epub
532 ahead of print September 2016. DOI: 10.1038/srep26276.

533 37. **Dridi B, Henry M, El Khéchine A, Raoult D, Drancourt M**. High Prevalence of
534 Methanobrevibacter smithii and Methanosphaera stadtmanae Detected in the Human Gut Using an
535 Improved DNA Detection Protocol. *PLoS One*;4. Epub ahead of print 17 September 2009. DOI:
536 10.1371/journal.pone.0007063.

537 38. **Liu Y, Hao W, Duan C, Tan J, Xu W, *et al.*** Comparative Analysis of Vaginal Bacterial
538 Diversity in Northern-Chinese Women Associated With or Without Bacterial Vaginosis. *Journal of*
539 *Medical Microbiology & Diagnosis*;0. Epub ahead of print 30 May 2012. DOI: 10.4172/2161-
540 0703.S5-001.

541 39. **Xia Q, Cheng L, Zhang H, Sun S, Liu F, *et al.*** Identification of vaginal bacteria diversity
542 and it's association with clinically diagnosed bacterial vaginosis by denaturing gradient gel
543 electrophoresis and correspondence analysis. *Infection, Genetics and Evolution* 2016;44:479–486.

544 40. **Huang B, Fettweis JM, Brooks JP, Jefferson KK, Buck GA**. The Changing Landscape of
545 the Vaginal Microbiome. *Clinics in Laboratory Medicine* 2014;34:747–761.

546 41. **Mendling W**. Vaginal Microbiota. In: Schwiertz A (editor). *Microbiota of the Human Body*.

547 Cham: Springer International Publishing. pp. 83–93.

548 42. **Li J, McCormick J, Bocking A, Reid G**. Importance of Vaginal Microbes in Reproductive
549 Health. *Reproductive Sciences* 2012;19:235–242.

550 43. **Iebba V, Totino V, Gagliardi A, Santangelo F, Cacciotti F, *et al.*** Eubiosis and dysbiosis:
551 the two sides of the microbiota. 12.

552 44. **Margolis E, Fredricks DN**. Bacterial Vaginosis-Associated Bacteria. In: *Molecular Medical*
553 *Microbiology*. Elsevier. pp. 1487–1496.

554 45. **Fredricks DN, Fiedler TL, Marrazzo JM**. Molecular identification of bacteria associated
555 with bacterial vaginosis. *New England Journal of Medicine* 2005;353:1899–1911.

556 46. **Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, *et al.*** The composition and
557 stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant
558 women. *Microbiome* 2014;2:4.

559 47. **Zozaya M, Ferris MJ, Siren JD, Lillis R, Myers L, *et al.*** Bacterial communities in penile
560 skin, male urethra, and vaginas of heterosexual couples with and without bacterial vaginosis.
561 *Microbiome*;4. Epub ahead of print December 2016. DOI: 10.1186/s40168-016-0161-6.

562 48. **Srinivasan S, Morgan MT, Fiedler TL, Djukovic D, Hoffman NG, *et al.*** Metabolic
563 Signatures of Bacterial Vaginosis. *mBio* 2015;6:e00204-15.

564 49. **Rahmati E, Martin V, Wong D, Sattler F, Petterson J, *et al.*** Facklamia Species as an
565 Underrecognized Pathogen. *Open Forum Infect Dis*;4. Epub ahead of print 9 January 2017. DOI:
566 10.1093/ofid/ofw272.

567 50. **Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, *et al.*** Diversity of the
568 Human Intestinal Microbial Flora. *Science* 2005;308:1635–1638.

569 51. **Fenollar F, Raoult D**. Does Bacterial Vaginosis Result From Fecal Transplantation? *The*
570 *Journal of Infectious Diseases* 2016;214:1784–1784.

571 52. **Reid G**. Is bacterial vaginosis a disease? *Applied Microbiology and Biotechnology*
572 2018;102:553–558.

573 53. **Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, *et al.*** Vaginal microbiome of
574 reproductive-age women. *Proc Natl Acad Sci U S A* 2011;108:4680–4687.

575 54. **Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, *et al.*** Differences in vaginal
576 microbiome in African American women versus women of European ancestry. *Microbiology*
577 2014;160:2272–2282.

578 55. **Forney LJ, Foster JA, Ledger W**. The vaginal flora of healthy women is not always
579 dominated by Lactobacillus species. *J Infect Dis* 2006;194:1468–1469; author reply 1469-1470.

580 56. **Stoner KA, Rabe LK, Austin MN, Meyn LA, Hillier SL**. Quantitative Survival of Aerobic
581 and Anaerobic Microorganisms in Port-A-Cul and Copan Transport Systems. *J Clin Microbiol*
582 2008;46:2739–2744.

583 57. **DeMarco AL, Rabe LK, Austin MN, Stoner KA, Avolia HA, *et al.*** Survival of Vaginal

584 Microorganisms in Three Commercially Available Transport Systems. *Anaerobe* 2017;45:44–49.

585 58. **Datcu R, Gesink D, Mulvad G, Montgomery-Andersen R, Rink E, *et al.*** Vaginal
586 microbiome in women from Greenland assessed by microscopy and quantitative PCR. *BMC Infectious*
587 *Diseases* 2013;13:480.

588 59. **Pandya S, Ravi K, Srinivas V, Jadhav S, Khan A, *et al.*** Comparison of culture-dependent
589 and culture-independent molecular methods for characterization of vaginal microflora. *Journal of*
590 *Medical Microbiology* 2017;66:149–153.

591 60. **Srinivasan S, Munch MM, Sizova MV, Fiedler TL, Kohler CM, *et al.*** More Easily
592 Cultivated Than Identified: Classical Isolation With Molecular Identification of Vaginal Bacteria. *J*
593 *Infect Dis* 2016;214 Suppl 1:S21-28.

594 61. **Zozaya-Hinchliffe M, Martin DH, Ferris MJ**. Prevalence and Abundance of Uncultivated
595 Megasphaera-Like Bacteria in the Human Vaginal Environment. *Applied and Environmental*
596 *Microbiology* 2008;74:1656–1659.

597 62. **Bilen M, Dufour J-C, Lagier J-C, Cadoret F, Daoud Z, *et al.*** The contribution of
598 culturomics to the repertoire of isolated human bacterial and archaeal species. *Microbiome*;6. Epub
599 ahead of print December 2018. DOI: 10.1186/s40168-018-0485-5.

600 63. **Lagier J-C, Dubourg G, Million M, Cadoret F, Bilen M, *et al.*** Culturing the human
601 microbiota and culturomics. *Nature Reviews Microbiology* 2018;16:540–550.

602 64. **Lagier J-C, Diagne N, Fenollar F, Tamalet C, Sokhna C, *et al.*** Vaginal self-sampling as a
603 diagnosis tool in low-income countries and potential applications for exploring the infectious causes of
604 miscarriage. *Future Microbiology* 2017;12:609–620.

605 65. **Lunny C, Taylor D, Hoang L, Wong T, Gilbert M, *et al.*** Self-Collected versus Clinician-
606 Collected Sampling for Chlamydia and Gonorrhea Screening: A Systemic Review and Meta-Analysis.
607 *PLoS ONE* 2015;10:e0132776.

608 66. **Menard J-P, Fenollar F, Henry M, Bretelle F, Raoult D**. Molecular Quantification of
609 Gardnerella vaginalis and Atopobium vaginae Loads to Predict Bacterial Vaginosis. *Clinical*
610 *Infectious Diseases* 2008;47:33–43.

611 67. **Nugent RP, Krohn MA, Hillier SL**. Reliability of diagnosing bacterial vaginosis is improved
612 by a standardized method of gram stain interpretation. *Journal of clinical microbiology* 1991;29:297–
613 301.

614 68. **Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, *et al.*** Microbial culturomics:
615 paradigm shift in the human gut microbiome study. *Clinical Microbiology and Infection*
616 2012;18:1185–1193.

617 69. **Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P, *et al.*** Ongoing Revolution in
618 Bacteriology: Routine Identification of Bacteria by Matrix-Assisted Laser Desorption Ionization
619 Time-of-Flight Mass Spectrometry. *Clinical Infectious Diseases* 2009;49:543–551.

620 70. **Seng P, Abat C, Rolain JM, Colson P, Lagier J-C, *et al.*** Identification of Rare Pathogenic
621 Bacteria in a Clinical Microbiology Laboratory: Impact of Matrix-Assisted Laser Desorption
622 Ionization-Time of Flight Mass Spectrometry. *Journal of Clinical Microbiology* 2013;51:2182–2194.

623 71. **Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral J-P, *et al.*** 16S ribosomal DNA
624 sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates.
625 *Journal of clinical microbiology* 2000;38:3623–3630.

626 72. **Kim M, Oh H-S, Park S-C, Chun J**. Towards a taxonomic coherence between average
627 nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes.
628 *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*
629 2014;64:346–351.

630 73. **Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, *et al.*** Uniting the classification of
631 cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews*
632 *Microbiology* 2014;12:635–645.

633 74. **Million M, Tidjani Alou M, Khelaifia S, Bachar D, Lagier J-C, *et al.*** Increased Gut Redox
634 and Depletion of Anaerobic and Methanogenic Prokaryotes in Severe Acute Malnutrition. *Scientific*
635 *Reports*;6. Epub ahead of print September 2016. DOI: 10.1038/srep26051.

636 75. **Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD**. PANDAseq: paired-
637 end assembler for illumina sequences. *BMC Bioinformatics* 2012;13:31.

638 76. **Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, *et al.*** QIIME allows
639 analysis of high-throughput community sequencing data. *Nature Methods* 2010;7:335–336.

640 77. **Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, *et al.*** DADA2: High-
641 resolution sample inference from Illumina amplicon data. *Nature Methods* 2016;13:581–583.

642 78. **AltschuP SF, Gish W, Miller W, Myers EW**. Basic Local Alignment Search Tool. 8.

643 79. **Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, *et al.*** The SILVA ribosomal RNA
644 gene database project: improved data processing and web-based tools. *Nucleic Acids Research*
645 2012;41:D590–D596.

646 80. **Angelakis E, Yasir M, Bachar D, Azhar EI, Lagier J-C, *et al.*** Gut microbiome and dietary
647 patterns in different Saudi populations and monkeys. *Scientific Reports*;6. Epub ahead of print October
648 2016. DOI: 10.1038/srep32191.

649 81. **Dhariwal A, Chong J, Habib S, King IL, Agellon LB, *et al.*** MicrobiomeAnalyst: a web-
650 based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids*
651 *Res* 2017;45:W180–W188.

652 82. **Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, *et al.*** Metagenomic biomarker
653 discovery and explanation. *Genome Biol* 2011;12:R60.

654 83. **Asnicar F, Weingart G, Tickle TL, Huttenhower C, Segata N**. Compact graphical
655 representation of phylogenetic data and metadata with GraPhlAn. *PeerJ*;3. Epub ahead of print 18
656 June 2015. DOI: 10.7717/peerj.1029.

657 84. **Jaykaran**. “Mean ± SEM” or “Mean (SD)”? *Indian J Pharmacol* 2010;42:329.

658

Table 1. Socio-demographic characteristic and vaginal microbiota profile of 34 women.

	Bacterial vaginosis (n=12)	Normal flora (n=22)	*p*-value
Age	29.39 ± 10.80	25.82 ± 8.82	NS
Ethnicity			NS
Caucasian	7 (58.3%)	17 (77.3%)	
Rural Senegalese	5 (41.7%)	5 (22.7%)	
Bacterial load			
Bacterial load log_{10} of 16S rRNA gene/sample	4.8 ± 0.4	4.9 ± 0.3	NS
Richness and diversity index			
Reads	99,568 ± 74,868	112,928 ± 66,466	NS
ACE[1]	34.8 ± 1.7	28.7 ± 1.5	NS
Chao-1	29.5 ± 16.5	22.5 ± 15.5	0.05
Shannon[2]	1.9 ± 0.7	1.3 ± 0.8	0.02
Simpson	0.3 ± 0.2	0.5 ± 0.3	0.02
Evenness[3]	0.14 ± 0.05	0.09 ± 0.05	0.009
Taxa exclusively found using both techniques			
	Bifidobacterium scardovii *Facklamia ignava*	*Acidaminococcus intestini* *Acinetobacter baumannii* *Alistipes putredinis* *Bacteroides cellulosilyticus* *Bacteroides fragilis* *Bacteroides salyersiae* *Enterococcus pallens* *Lactobacillus mucosae* *Macrococcus caseolyticus* *Morganella morganii* *Phascolarctobacterium faecium* *Pseudoramibacter alactolyticus* *Streptococcus australis* *Streptococcus urinalis* *Trueperella bernardiae*	

[1]Richness (ACE and Chao-1) and [2]diversity (Shannon and Simpson) indices were evaluated using the MicrobiomeAnalyst pipeline. [3]Shannon evenness was calculated using the formula: E = H/ln(S) with H = Shannon index and S = total number of sequences in that cohort. *P*-value is shown only for a significant relationship. NS: not significant ($p>0.05$).

Table 2. Growth conditions, taxonomic and, the source of putative new species isolated in the human vagina.

Phylogenetic affiliation	Species name	Strain	First culture condition	Sample origin	Clinical status	16S accession number	Relatedness to known species
Actinobacteria							
Actinomycetaceae	*Varibaculum vaginae*	Marseille-P5644	Direct plating, 5% sheep blood trypticase soy agar, anaerobe, 37°C	Dielmo, Senegal	Healthy	In progress	98% *Varibaculum cambriense*
	Arcanobacterium ihumii	Marseille-P5647	Incubation in a blood culture for 20 days, 5% sheep blood CNA agar, anaerobe, 37°C	Dielmo, Senegal	Healthy	LT993248	96.64% *Arcanobacterium phocae* LT629804
Atopobiaceae	*Atopobium massiliense*	Marseille-P4126	Direct plating, 5% sheep blood CNA and trypticase soy agar, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LT986001	98.19% *Atopobium vaginae* AF325325
	Olegusella massiliensis[*]	KHD7	Incubation in a blood culture for 10 days, 5% sheep blood CNA agar, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LN998058	93.5% *Olsenella uli* NR_074414
Coriobacteriaceae	*Collinsella vaginalis*	Marseille-P2666	Incubation in a blood culture for 15 days, 5% sheep blood CNA agar, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LT598547	96.08% *Collinsella intestinalis* NR_113165
Corynebacteriaceae	*Corynebacterium feminarum*	Marseille-P4858	Direct plating, 5% sheep blood trypticase soy agar, anaerobe 37°C	Marseille, France	Healthy	In progress	98.2% *Corynebacterium similans*
	Corynebacterium fournierii	Marseille-P2948	Incubation in a blood culture for 30 days, Chocolate agar PVX, aerobe, 37°C	Marseille, France	Bacterial vaginosis	LT576414	98.7% *Corynebacterium ureicelerivorans* NR_042558
Eggerthellaceae	*Vaginimassilia timonensis**	Marseille-P4307	Direct plating, 5% sheep blood CNA agar, anaerobe, 37°C	Dielmo, Senegal	Bacterial vaginosis	LT996087	93.4% *Gordonibacter urolithinfaciens* LT900217

Intrasporangiaceae	*Janibacter massiliensis*	Marseille-P4121	Incubation in a blood culture for 10 days, Schaedler agar with 5% sheep blood and vitamin K, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LT969384	98.01% *Janibacter melonis* JN644568
Propionibacteriaceae	*Cutibacterieum timonense*	Marseille-P5998	Incubation in a blood culture for 15 days, 5% sheep blood trypticase soy agar, anaerobe, 37°C	Dielmo, Senegal	Healthy	LT996136	97.80% *Cutibacterium acnes* CP023676
	Tessaracoccus timonensis	Marseille-P5995	Direct plating, 5% sheep blood CNA agar, anaerobe 37°C	Dielmo, Senegal	Bacterial vaginosis	LT996088	97.30% *Tessaracoccus oleiagri* GU111567
	*Vaginimicrobium propionicum**	Marseille-P3275	Incubation in a blood culture for 10 days, 5% sheep blood CNA agar, Schaedler and Trypticase soy agar, anaerobe, 37°C	Marseille, France	Healthy	LT598595	92.92% *Propionimicrobium lymphophilum* LT223675
Bacteroidetes							
Flavobacteriaceae	*Vaginella massiliensis**	Marseille-P2517	Incubation in a blood culture for 7 days, Chocolate agar PVX, aerobe, 37°C	Marseille, France	Healthy	LT223570	93.03% *Weeksella virosa* NR_074495
Prevotellaceae	*Prevotella lascolaii*	KhD1	Incubation in a blood culture for 24 hours, Schaedler agar with 5% sheep blood and vitamin K, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LN998055	90% *Prevotella loescheii* FJ717335
Firmicutes							
Aerococcaceae	*Vaginisenegalia massiliensis**	Marseille-P5643	Direct plating, 5% sheep blood CNA agar, anaerobe 37°C	Dielmo, Senegal	Bacterial vaginosis	LT971014	93.77% *Facklamia hominis* NR_026393
Lactobacillaceae	*Lactobacillus raoultii*	Marseille-P4006	Incubation in a blood culture for 3 days, 5% sheep blood agar, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LT854294	98.1% *Lactobacillus farraginis* AB690214

Peptoniphilaceae	*Anaerococcus mediterraneensis*	Marseille-P2765	Incubation in a blood culture for 21 days, Schaedler agar with 5% sheep blood and vitamin K, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LT598544	97.2% *Anaerococcus lactolyticus* NR_113565
	Anaerococcus genitaliorum	Marseille-P3625	Incubation in a blood culture for 10 days, 5% sheep blood CNA agar and Schaedler, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LT900366	93.37% *Anaerococcus tetradius* GQ422749
	Anaerococcus mediannikovii	Marseille-P3915	Incubation in a blood culture for 10 days, 5% sheep blood agar, anaerobe, 37°C	Marseille, France	Healthy	LT966066	96,73% *Anaerococcus lactolyticus* NR_113565
	Ezakiella massiliensis	Marseille-P2951	Incubation in a blood culture for 21 days, 5% sheep blood CNA agar, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LT576398	98.5% *Ezakiella peruensis* KJ469554
	Helcococcus massiliensis	Marseille-P4590	Incubation in a blood culture for 15 days, 5% sheep blood agar, anaerobe, 37°C	Dielmo, Senegal	Bacterial vaginosis	LT934442	95.5% *Helcococcus seattlensis* NR_118641
	*Khoudiadiopia massiliensis**	Marseille-P2746	Incubation in a blood culture for 21 days, Schaedler agar with 5% sheep blood and vitamin K, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LT598561	89.28% *Murdochiella asaccharolytica* EU483153
	Murdochiella vaginalis	Marseille-P2341	Incubation in a blood culture for 15 days, Schaedler agar with 5% sheep blood and vitamin K, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LT576397	97.3% *Murdochiella asaccharolytica* NR_116331
	Peptoniphilus pacaensis	Kh-D5	Incubation in a blood culture for 15 days, 5% sheep blood CNA and trypticase soy agar, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LN998072	97.2% *Peptoniphilus coxii* NR_117556

	Peptoniphilus raoultii	KHD4	Incubation in a blood culture for 24 hours, 5% sheep blood CNA and trypticase soy agar, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LN998068	96% *Peptoniphilus lacrimalis* NR_041938
	Peptoniphilus vaginalis	KhD-2	Incubation in a blood culture for 24 hours, 5% sheep blood CNA and trypticase soy agar, anaerobe, 37°C	Marseille, France	Bacterial vaginosis	LN907856	98.2% *Peptoniphilus harei* NR_026358
Veillonellaceae	*Megasphaera vaginalis*	Marseille-P4857	Direct plating, 5% sheep blood CNA, anaerobe 37°C	Marseille, France	Bacterial vaginosis	LT960586	95.23% *Megasphaera micronuciformis* GU470904

*New genus.

Table 3. Comparison of the microbial culturomics results of vaginal bacterial diversity between patients with bacterial vaginosis and healthy women.

Vaginal flora	Bacterial vaginosis (n = 12)	Normal flora (n = 22)	p-value
Global Diversity			
Phyla	4.7 ± 1.6	4.2 ± 1.2	0.45[a]
Genera	30 ± 10	25 ± 9	0.26[b]
Species	50 ± 20	39 ± 13	0.06[b]
New species	1.5 ± 2	0.4 ± 0.6	0.03[a]
Diversity by Phylum			
Firmicutes	28 ± 9.8	20 ± 6.2	0.007[b]
Actinobacteria	13 ± 5.2	11 ± 4.8	0.25[b]
Bacteroidetes	5.7 ± 5.1	4.8 ± 4.2	0.70[a]
Proteobacteria	1.7 ± 1.3	1.5 ± 1.4	0.64[b]
Fusobacteria	0.2 ± 0.4	0.2 ± 0.4	0.70[a]
Synergistetes	0.2 ± 0.6	0.04 ± 0.2	0.66[a]
Epsilonbacteraeota	0.5 ± 0.5	0.4 ± 0.5	0.63[a]
Diversity by Genus			
Peptoniphilus	4.5 ± 2.8	2.2 ± 1.9	0.009[a]
Anaerococcus	2.5 ± 1.7	1.8 ± 1.8	0.20[b]
Prevotella	2.0 ± 1.9	1.1 ± 1.5	0.11[a]
Clostridium	1.3 ± 1.1	0.5 ± 0.7	0.02[a]
Lactobacillus	1.6 ± 1.6	2.2 ± 1.4	0.28[a]
Bifidobacterium	1.2 ± 1.2	0.5 ± 0.8	0.07[b]
Atopobium	1.2 ± 1.2	0.9 ± 0.8	0.48[b]
Gardnerella	0.5 ± 0.5	0.4 ± 0.5	0.81[b]

[a] Mann Whitney test. [b] unpaired *t*-test.

Table 4. Metagenomics comparison of vaginal microbiota between French and rural Senegalese women, with and without bacterial vaginosis (BV).

	Reads	ACE	Chao-1	Shannon	Simpson
BV-France (n=7)	88,001	17 ± 1.30	22.9 ± 9.10	1.9 ± 0.20	0.28 ± 0.07
BV-Senegal (n=5)	115,762	25.4 ± 1.60	39.6 ± 20.60	2.02 ± 0.40	0.29 ± 0.10
Healthy-France (n=17)	106,792	48 ± 1.98	18.06 ± 3.20	1.2 ± 0.20	0.46 ± 0.28
Healthy-Senegal (n=5)	133,792	37.4 ± 1.60	35.08 ± 5.70	1.4 ± 0.40	0.47 ± 0.29
p-value (ANOVA)	NS	0.003	0.007	NS	NS

NS: not significant ($p>0.05$).

Table 5. Comparison of vaginal species isolated from French and rural Senegalese women, with and without bacterial vaginosis (BV).

Global diversity	French women	Senegalese women	*p*-value
Bacterial vaginosis (n=12)	7	5	
Phyla	5.3 ± 0.9	3.8 ± 1.9	0.09[a]
Total species	63.3 ± 11.2	30.6 ± 11.5	0.0006[b]
Strictly anaerobic species	38 ± 11.1	15.2 ± 11.9	0.005[b]
Aerotolerant species	25.4 ± 5.7	15.6 ± 2.7	0.005[b]
Healthy flora (n=22)	17	5	
Phyla	4.8 ± 0.8	2.4 ± 0.5	0.0009[b]
Species	43 ± 10.3	24.8 ± 12.5	0.002[b]
Strictly anaerobic species	24.2 ± 9.5	8.8 ± 9	0.004[b]
Aerotolerant species	18.8 ± 3.8	15.4 ± 4.4	0.10[b]

[a] Mann Whitney test. [b] unpaired *t*-test.

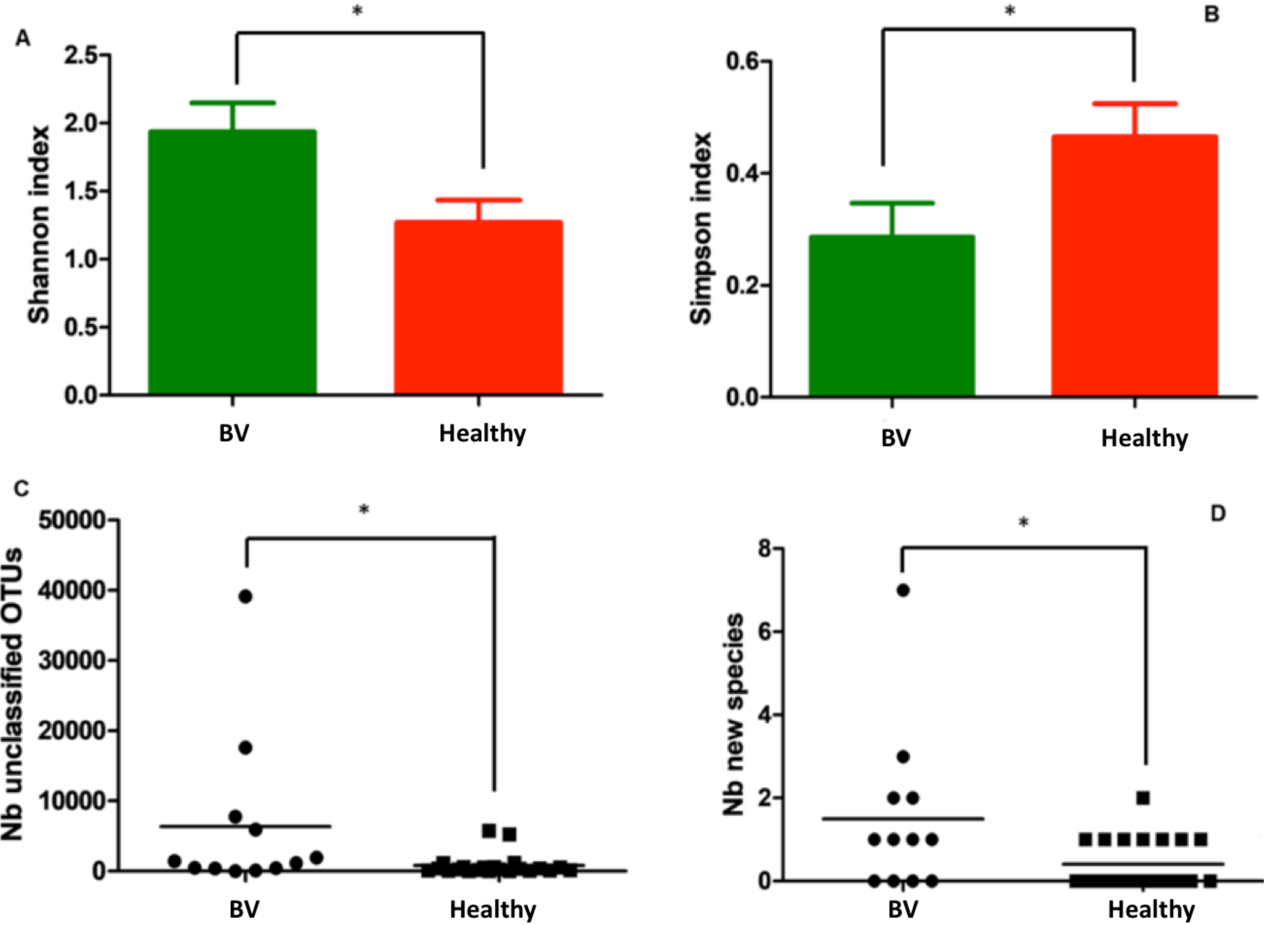


Figure 1. Comparison of vaginal microbiota diversity between flora with bacterial vaginosis (BV) and healthy. Species diversity estimated by Shannon (**A**) and Simpson indices (**B**). The hitherto unknown diversity detected using metagenomics (**C**) and culturomics (**D**) approaches.

Women with bacterial vaginosis showed a significantly increased (known as well as previously unknown) diversity. *$p < 0.05$.

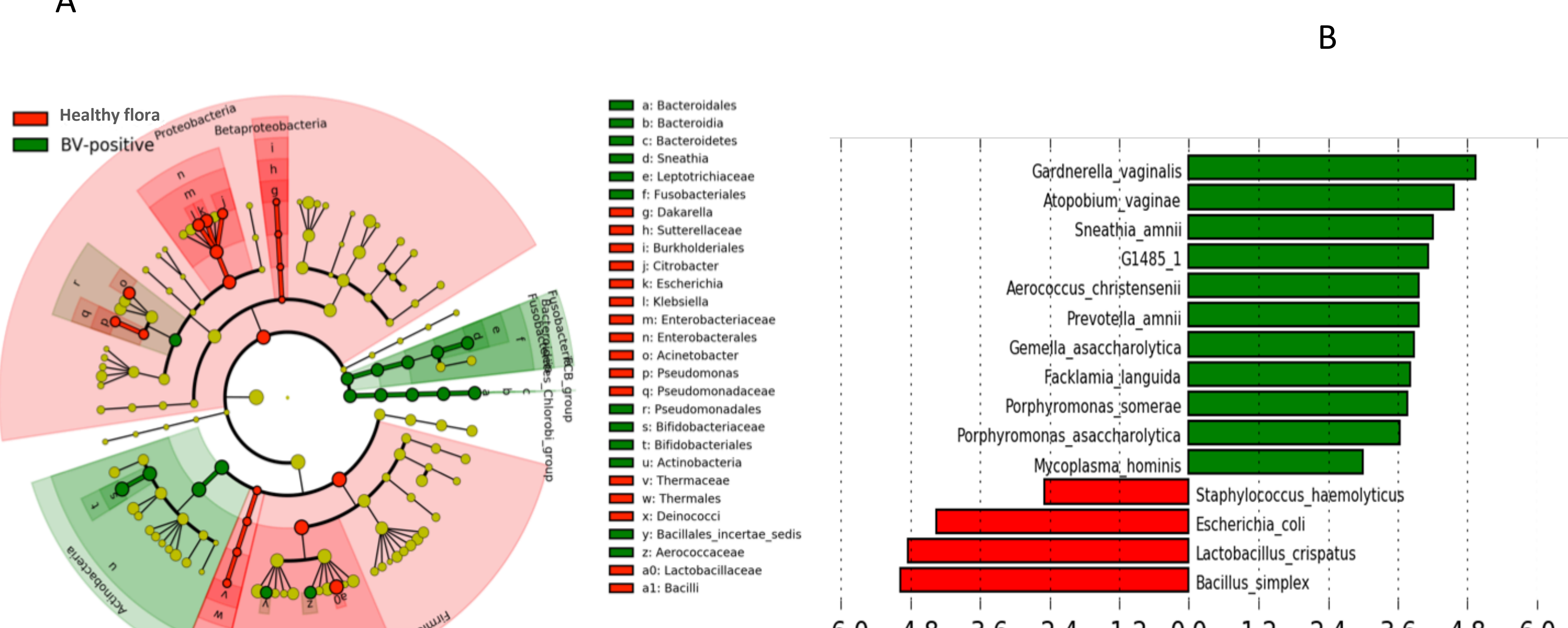


Figure 2. Identification of biomarker taxa between bacterial vaginosis (BV)-positive and healthy vaginal microbiota. **A**. Cladogram representation of taxa with different relative abundance in accordance with BV status. **B**. LDA score (linear discriminant analysis) showing abundant species as biomarkers in BV-positive and healthy vaginal microbiota.

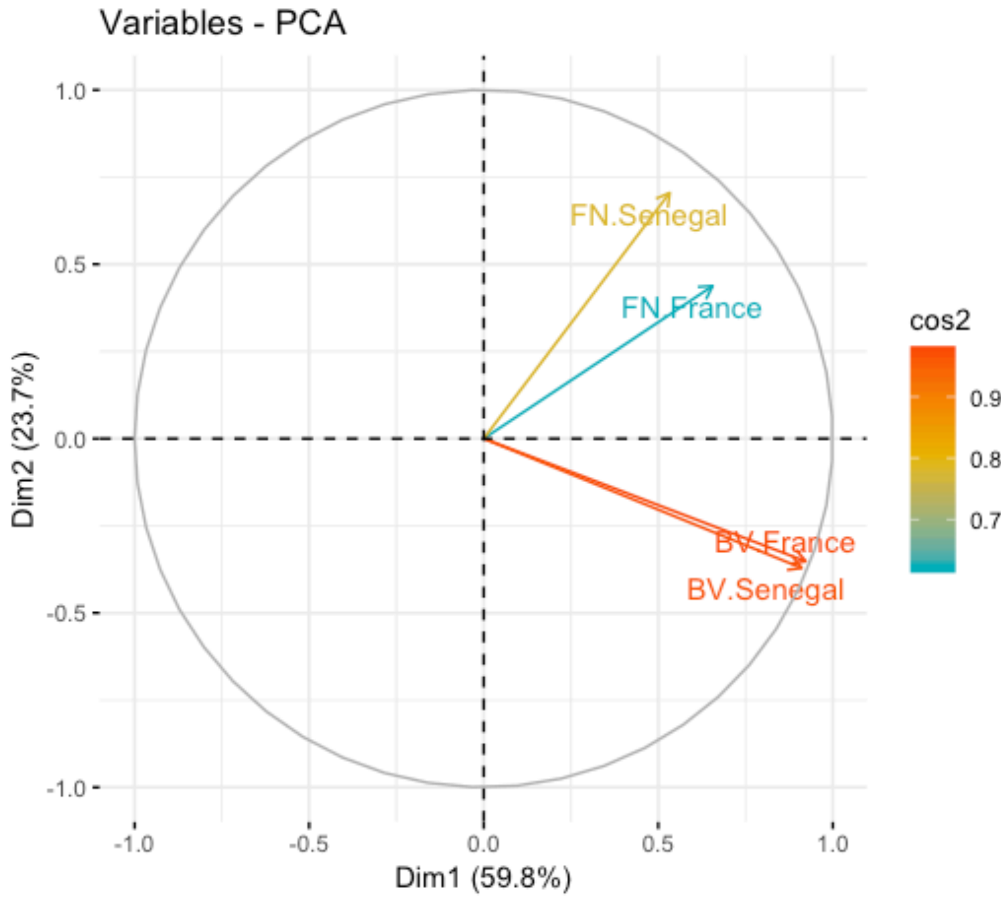


Figure 4: Principal Components Analysis (PCA) graph showing the correlation of the vaginal microbiota status (Bacterial vaginosis or healthy) and the geographical repartition (France or Senegal) for 34 vaginal samples

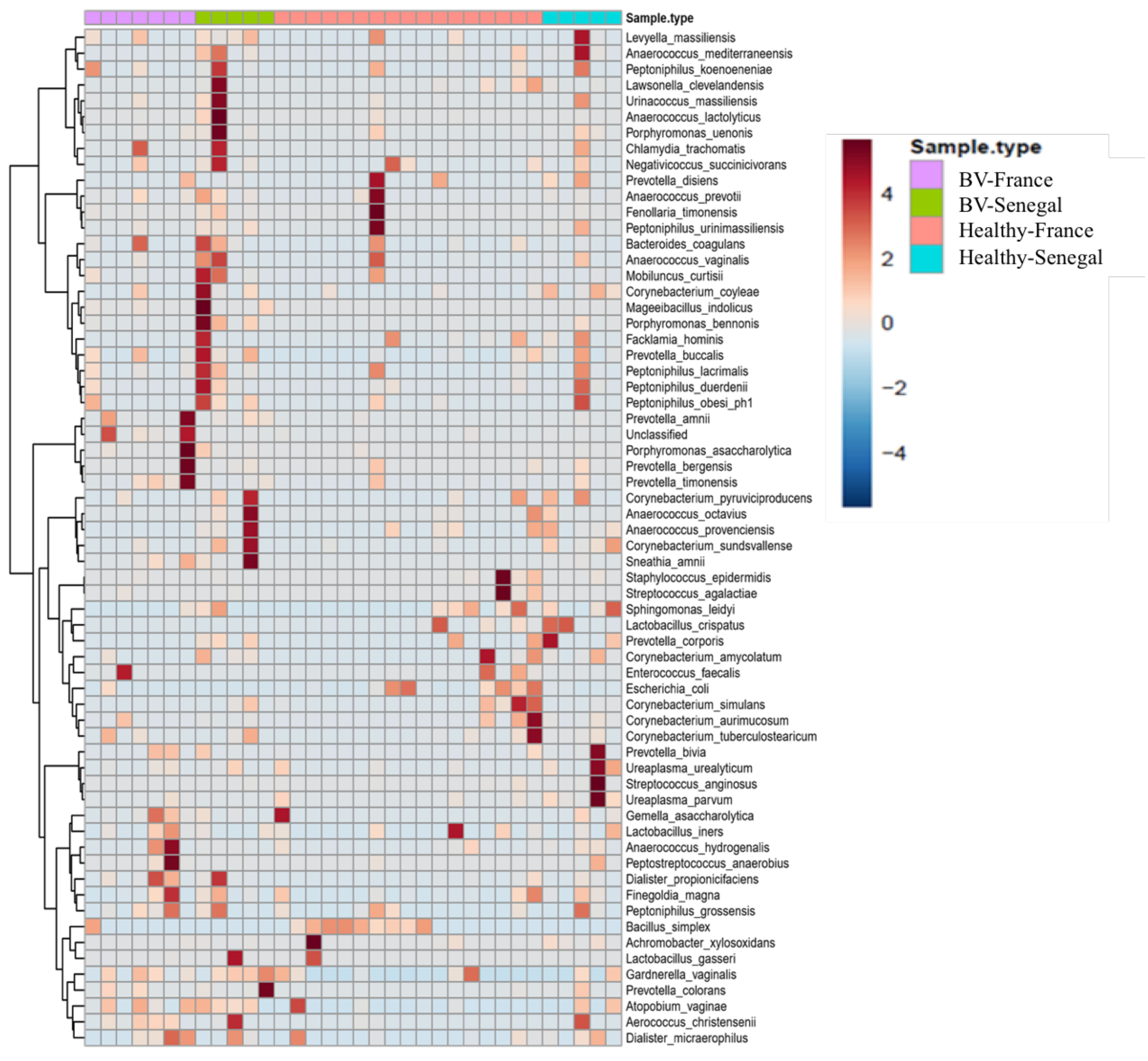


Figure 5. Heatmap by using the Ward clustering method based on the Euclidean distance.

Chapitre III :

Taxonogénomique : Incorporation de l'Information Génomique dans la Description des Nouvelles Espèces Bactériennes

Avant-propos

L'exploration du microbiote vaginal à l'aide d'outils moléculaires a révélé la forte présence dans ce biotope d'espèces bactériennes non-cultivables ou putatives nouvelles espèces fortement associées à la vaginose bactérienne telles que les bactéries associées à la vaginose de type 1 (BAVB-1), BAVB-2, BAVB-3, *Megasphaera* sp type 1 et 2, *Dialister* sp type 1 à 3 et *Eggerthella* sp type 1 [37, 64, 65]. Récemment, les nouvelles techniques de culture développées ont permis l'isolement de certaines de ces nouvelles espèces exclusivement détecter par méthodes moléculaires telles que BABV-3 nommée *Mageeibacillus indolicus* [69], *Dialister* sp type 2, *Eggerthella* sp type 1, et *Megasphaera* sp type 1 [65].

Appliquer pour la première fois dans l'exploration du microbiote vaginal, la technique de culturomique microbienne m'a permis d'isoler 20 nouvelles espèces lors de l'analyse de la flore vaginale de femmes saines et celles souffrant de vaginose bactérienne. Actuellement, un seuil de similarité de la séquence codant pour l'ARNr 16S de 98,7% a été fixé pour déterminer si deux isolats bactériens appartiennent à des espèces distinctes [70, 71]. Habituellement pour définir et caractériser les nouvelles espèces, une approche polyphasique renfermant les caractéristiques phénotypiques (critères morphologiques et biochimiques) renforcées avec la similarité des séquences du gène codant pour l'ARNr 16S est utilisée [72]. Certes, cette approche a permis de reclasser de nombreuses espèces bactériennes mais, certaines études ont montré que ce seuil de 98,7% n'était pas applicable à tous genres bactériens et que donc le séquençage de l'ADNr 16S n'était pas fiable pour distinguer certaines communautés microbiennes. C'est ainsi qu'avec l'avancé des technologies du séquençage à haut débit et la baisse de leur coût, les données génomiques deviennent de plus en plus accessibles et leur utilisation en taxonomie devient davantage souhaitée. Dès lors, pour décrire et caractériser les

nouvelles espèces bactériennes isolées par culturomique, notre laboratoire a introduit en 2014 un nouveau concept appelé taxonogénomique [40, 41].

En effet, la taxonogénomique est une approche polyphasique qui complète les méthodes de description classique par l'ajout des informations protéomiques obtenues après analyse par spectrométrie de masse MALDI-TOF, la teneur en acides gras de la paroi bactérienne et des données génomiques issues des analyses fonctionnelle et comparative de similarité des séquences génomiques. En plus de sa reproductibilité, la taxonogénomique par le séquençage du génome bactérien permet de renforcer les banques de données et réduire par conséquent le nombre de séquences non assignées à aucune espèce durant les analyses métagénomiques [73].

La description taxonogénomique d'un nouvel isolat inclue, notamment : ses conditions de culture et de croissance (pH, salinité, température et atmosphère), son aspect morphologique (taille, forme, Gram et aspect des colonies) [74], ses principales caractéristiques biochimiques et sa sensibilité aux antibiotiques [75, 76], la composition en acide gras de sa paroi [77], ses caractéristiques phylogénétiques (séquence de l'ADNr 16S et arbre phylogénétique synthétisant sa phylogénie) et ses propriétés génomiques (taille du génome, teneur en G+C, nombre total de gènes, distribution des gènes dans les catégories COG, nombre de gènes d'ARN, et l'identité génomique moyenne des séquences de gènes (AGIOS) par rapport à celles des bactéries les plus proches) [78, 79].

Dans ce chapitre, nous rapportons la description taxonogénomique de 12 de nos nouvelles espèces, le reste étant en cours d'analyse.

Neuf ont été isolées de la flore de vaginose bactérienne et appartiennent aux :

- Phylum des *Actinobacteria : Olegusella massiliensis* (famille des *Atopobiaceae*), *Collinsella vaginalis* (famille des *Coriobacteriaceae*)*, Corynebacterim fournierii* (familles des *Corynebacteriaceae*) et *Janibacter massiliensis* (familles des *Intrasporangiaceae*).
- Phylum des *Bacteroidetes* : *Prevotella lascolaii* (famille des *Prevotellaceae*)*,*
- Phylum des *Firmicutes* : *Peptoniphilus pacaensis, Peptoniphilus raoultii, Peptoniphilus vaginalis,* et *Murdochiella vaginalis* (famille des *Peptoniphilaceae*)*,*

Les 3 autres ont été isolées dans le tractus vaginal de femmes saines et appartiennent aux :

- Phylum *Actinobacteria* : *Vaginimicrobium massiliensis* (famille *Proponibacteriaceae*)
- Phylum des *Bacteroidetes* : *Vaginella massiliensis* (famille des *Flavobacteriaceae*),
- Phylum des *Firmicutes* : *Ezakiella massiliensis* (famille des *Peptoniphilaceae*).

Tableau 1 : Tableau récapitulatif des 12 nouvelles espèces décrites

Phylum	Famille	Bactéries	Isolement	Flore	Caractéristiques phénotypiques	Caractéristiques génomiques	Publications
Actinobacteria	***Atopobiaceae***	*Olegusella massiliensis*	10 jours pré-incubation, culture sur gélose ANC en anaérobie	Vaginose bactérienne	Bacille Gram+ Anaérobie stricte Catalase négative Oxydase négative	1,81 Mb 49,24% G+C 1.549 génes 51 ARNs.	Publié dans Anaerobe
	Coriobacteriaceae	*Collinsella vaginalis*	15 jours pré-incubation, culture sur gélose ANC en anaérobie	Vaginose bactérienne	Bacilles Gram+ Anaérobie stricte Catalase négative Oxydase négative	2.16 Mb 64.57% G+C 1.774 génes 50 ARNs	Révision dans IJSEM
	Corynebacteriaceae	*Corynebacterium fournierii*	30 jours pré-incubation, culture sur gélose Chocolat en aérobie	Vaginose bactérienne	Bacille Gram+ Aérobie Catalase positive Oxydase négative	2,38 Mb 65,03% G+C 2.147 génes 63 ARNs	Publié dans Antonie van Leeuwenhoek
	Intrasporangiaceae	*Janibacter massiliensis*	10 jours pré-incubation, culture sur gélose Schaedler en anaérobie	Vaginose bactérienne	Coccus Gram+ Aérobie Catalase positive Oxydase négative	2,45Mb 72.5% G+C 2.351 génes 49 ARNs	Soumis dans Antonie van Leeuwenhoek
	Propionibacteriaceae	*Vaginimicrobium propionicum*	10 jours pré-incubation, culture sur géloses Schaedler et TSA en anaérobie	Flore normale	Bacilles Gram+ Anaérobie stricte Catalase négative Oxydase négative	2,01 Mb 50,64 % G+C 1.869 génes 52 ARNs	Soumis dans IJSEM
Bacteroidetes	***Flavobacteriaceae***	*Vaginella massiliensis*	30 jours pré-incubation, culture sur gélose Chocolat en aérobie	Flore normale	Bacille Gram- Aérobie Catalase négative Oxydase positive	2,43 Mb 38,16% G+C 2.324 génes 71 ARNs	Publié dans NMNI
	Prevotellaceae	*Prevotella lascolaii*	24 h pré-incubation, culture sur gélose Schaedler en anaérobie	Vaginose bactérienne	Bacille Gram- Anaérobie stricte Catalase négative Oxydase positive	3,76 Mb 48,7% G+ C 3194 54 ARNs genes	Publié dans OMICS Integrative biology

Firmicutes	***Peptoniphilaceae***	*Ezakiella massiliensis*	15 jours pré-incubation, culture sur gélose ANC en anaérobie	Flore normale	Coccus Gram+ Anaérobie stricte Catalase positive Oxydase négative	1,74 Mb 36.69% G+C 1.606 génes 51 ARNs	Publié dans Current Microbiology
		Murdochiella vaginalis	15 jours pré-incubation, culture sur gélose Schaedler en anaérobie	Vaginose bactérienne	Coccus Gram+ Anaérobie stricte Catalase négative Oxydase positive	1,67 Mb 49.48% G+C 1.446 génes 55 ARNs	Publié dans Microbiology Open
		Peptoniphilus pacaensis	15 jours pré-incubation, culture sur géloses ANC et TSA en anaérobie	Vaginose bactérienne	Coccus Gram+ Anaérobie stricte Catalase négative Oxydase négative	1,85 Mb 49.4% G+C 1.802 génes 54 ARNs	Publié dans Microbiology Open
		Peptoniphilus raoultii	24 h pré-incubation, culture sur géloses ANC et TSA en anaérobie	Vaginose bactérienne	Coccus Gram+ Anaérobie stricte Catalase négative Oxydase négative	1,62 Mb 31,9% G+C 1.631 génes 42 ARNs	Publié dans Microbiology Open
		Peptoniphilus vaginalis	24 h pré-incubation, culture sur géloses ANC et TSA en anaérobie	Vaginose bactérienne	Coccus Gram+ Anaérobie stricte Catalase négative Oxydase négative	1,88 Mb 34,2% G+C 1.791 génes 40 ARNs	Publié dans Microbiology Open

- **Description des nouvelles espèces isolées dans le microbiote de vaginose bactérienne**

Article 4:

Olegusella massiliensis **gen. nov. sp. nov., strain KHD7^T, a new bacterial genus isolated from the female genital tract of a patient with bacterial vaginosis.**

Khoudia Diop, Awa Diop, Florence Bretelle, Frederic Cadoret, Caroline Michelle, Magali Richez, Jean-François Cocallemen, Didier Raoult, Pierre-Edouard Fournier, Florence Fenollar

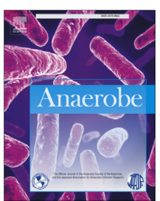
Published in Anaerobe

Anaerobe 44 (2017) 87–95

Contents lists available at ScienceDirect

Anaerobe

journal homepage: www.elsevier.com/locate/anaerobe



Research Paper
Anaerobes in the microbiome

Olegusella massiliensis gen. nov., sp. nov., strain KHD7^T, a new bacterial genus isolated from the female genital tract of a patient with bacterial vaginosis

Khoudia Diop [a], Awa Diop [a], Florence Bretelle [a, b], Frédéric Cadoret [a], Caroline Michelle [a], Magali Richez [a], Jean-François Cocallemen [b], Didier Raoult [a, c], Pierre-Edouard Fournier [a], Florence Fenollar [a, *]

[a] Aix Marseille Univ, Institut Hospitalo-Universitaire Méditerranée-Infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine, 27 Boulevard Jean Moulin, 13385 Marseille Cedex 05, France
[b] Department of Gynecology and Obstetrics, Gynépole, Marseille, Pr Boubli et D'Ercole, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, AMU, Aix-Marseille Université, France
[c] Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 18 August 2016
Received in revised form
2 February 2017
Accepted 15 February 2017

Handling Editor: Elisabeth Nagy

Keywords:
Olegusella massiliensis
Vaginal flora
Bacterial vaginosis
Culturomics
Taxono-genomics
Genome

ABSTRACT

Strain KHD7^T, a Gram-stain-positive rod-shaped, non-sporulating, strictly anaerobic bacterium, was isolated from the vaginal swab of a woman with bacterial vaginosis. We studied its phenotypic characteristics and sequenced its complete genome. The major fatty acids were C16:0 (44%), C18:2n6 (22%), and C18:1n9 (14%). The 1,806,744 bp long genome exhibited 49.24% G+C content; 1549 protein-coding and 51 RNA genes. Strain KHD7^T exhibited a 93.5% 16S rRNA similarity with *Olsenella uli*, the phylogenetically closest species in the family *Coriobacteriaceae*. Therefore, strain KHD7^T is sufficiently distinct to represent a new genus, for which we propose the name *Olegusella massiliensis* gen. nov., sp. nov. The type strain is KHD7^T.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

The female genital tract is a complex ecosystem colonized by several types of microorganisms. Its composition was described for the first time in 1892 by Doderleïn and in 1901 by Beijerink, revealing that four species of *Lactobacillus* are predominant in healthy vaginal flora: *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus jensenii*, and *Lactobacillus iners* [1,2]. The other bacteria include some anaerobic species such as *Bacteroides*, *Peptostreptococcus*, *Peptococcus*, *Corynebacterium*, and *Eubacterium* [3]. This mutualistic association maintains the stability of the vaginal environment, preventing infection by inhibiting the growth and expansion of pathogens through the production of antimicrobial molecules such as hydrogen peroxide, lactic acid, and bacteriocins [4,5].

This mutualism is disturbed in bacterial vaginosis (BV). The most common cause of vaginal discharge affecting women of child-bearing age, BV is concurrently characterized by reduced *Lactobacillus* species and increased anaerobic bacteria including *Atopobium vaginae*, *Bacteroides* spp., *Mobiluncus* spp., *Prevotella* spp., *Peptoniphilus* spp., and *Anaerococcus* spp. [6–9]. The vaginal microbiota was first studied by conventional culture methods. These methods are limited because 80% of the bacterial microbiota is considered to be fastidious or not cultivable [10]. Advances in molecular techniques, with sequencing and phylogenetic analysis of the 16S rRNA gene, enhanced understanding of the human vaginal microbiota.

* Corresponding author.
E-mail address: florence.fenollar@univ-amu.fr (F. Fenollar).

http://dx.doi.org/10.1016/j.anaerobe.2017.02.012
1075-9964/© 2017 Elsevier Ltd. All rights reserved.

Abbreviations

AGIOS	Average of Genomic Identity of Orthologous gene Sequences
bp:	base pairs
COG	Clusters of Orthologous Groups
CSUR	Collection de souches de l'Unité des Rickettsies
DDH	DNA-DNA Hybridization
DSM	Deutsche Sammlung von Mikroorganismen
FAME	Fatty Acid Methyl Ester
GC/MS	Gas Chromatography/Mass Spectrometry
kb	kilobases
MALDI-TOF	Matrix-assisted laser-desorption/ionization time-of-flight
ORF	Open Reading Frame
TE buffer	Tris-EDTA buffer
URMITE	Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes

These molecular methods allowed the detection of fastidious and uncultured bacteria, such as bacterial vaginosis-associated bacteria type 1 (BVAB1), BVAB2, and BVAB3 [11].

As part of a study on the diversity of the vaginal microbiota of patients with bacterial vaginosis using the culturomics approach, based on multiplication of culture conditions (variation of media, temperature, and atmosphere) with more rapid bacterial identification by MALDI-TOF mass spectrometry [12], we isolated a new member of the *Coriobacteriaceae* family. This family, created in 1997 by Stackebrandt, contains 35 species grouped in 13 validated genera [13,14].

Various parameters, including phenotypic and genotypic characteristics such as DNA-DNA hybridization, have been used to define a new species but they present certain limitations [15,16], so we introduced "taxono-genomics", a new approach that includes genomic analysis and proteomic information obtained by MALDI-TOF mass spectrometry analysis [17,18].

Here, we describe *Olegusella massiliensis* strain KHD7^T (= CSUR P2268 = DSM 101849), with its complete annotated genome, a new member of the *Coriobacteriaceae* family isolated in the vaginal flora of a patient with bacterial vaginosis.

2. Materials and methods

2.1. Sample collection

In October 2015, the vaginal sample of a French 33 year-old woman was collected at Hôpital Nord in Marseille (France). The patient was suffering from bacterial vaginosis, which was diagnosed as previously reported [19]. At the time of sample collection, she was not being treated with any antibiotics. She gave her written consent. This study was authorized by the local IFR48 ethics committee (Marseille, France) under agreement number 09-022. The sample was collected and transported using a Sigma Transwab (Medical Wire, Corsham, United Kingdom).

2.2. Strain identification by MALDI-TOF MS

After collection, the sample was first inoculated in a blood culture bottle (BD Diagnostics, Le Pont-de-Claix, France) supplemented with 4 mL of rumen that was filter-sterilized through a 0.2 μm pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France), and 3 mL of sheep blood (bioMérieux, Marcy l'Etoile, France). The supernatant was then inoculated on 5% sheep blood-enriched CNA agar (BD Diagnostics) under anaerobic conditions at 37 °C. Isolated colonies were deposited in duplicate on a MTP 96 MALDI-TOF target plate (Bruker Daltonics, Leipzig, Germany) for identification with a microflex spectrometer (Bruker) [20]. Briefly, 1.5 μL of matrix solution, containing solution of α-cyano-4-hydroxycinnamic acid diluted in 500 μL acetonitrile, 250 μL 10% trifluoroacetic acid and 250 μL HPLC water was deposited on each spot for ionization and crystallization. All protein spectra obtained were compared with those in the MALDI-TOF database. If the score was greater than or equal to 1.9, the strain was considered identified. Otherwise, the identification failed.

2.3. Strain identification by 16S rRNA sequencing

For unidentified strains using MALDI-TOF MS, 16S rRNA sequencing was used to achieve identification [21]. As Stackebrandt and Ebers suggested, if the 16S rRNA sequence similarity value was lower than 98.7% or 95%, the strain was defined as a new species or genus, respectively [22–24].

2.4. Morphologic observation and growth conditions

Optimal strain growth was also tested at different temperatures (25, 28, 37, 45, and 56 °C) in an aerobic atmosphere with or without 5% CO_2, and in anaerobic and microaerophilic atmospheres using GENbag Anaer and GENbag miroaer systems (bioMérieux).

For electron microscopy, detection formvar-coated grids were dropped onto a 40 μL bacterial suspension before incubation at 37 °C for 30 min. Then, the grids were incubated on 1% ammonium molybdate for 10 s, dried on blotting paper and finally observed using a Tecnai G20 transmission electron microscope (FEI, Limeil-Brevannes, France) at an operating voltage of 60 Kv. Standard procedures were used to perform Gram-staining, motility, sporulation as well as oxidase and catalase tests [25].

2.5. Biochemical analysis and antibiotic susceptibility tests

Cellular fatty acid methyl ester (FAME) analysis was performed by GC/MS. Strain KHD7^T was grown on Columbia agar enriched with 5% sheep blood (bioMérieux). Then, two samples were prepared with approximately 30 mg of bacterial biomass per tube harvested from several culture plates. Fatty acid methyl esters were prepared as described by Sasser [26]. GC/MS analyses were realized by using a Clarus 500 gas chromatograph equipped with a SQ8S MS detector (Perkin Elmer, Courtaboeuf, France). 2 μL of FAME extracts were volatilized at 250 °C (split 20 mL/min) in a Focus liner with wool and separated on an Elite-5MS column (30 m, 0.25 mm i.d., 0.25 mm film thickness) using a linear temperature gradient (70–290 °C at 6 °C/min), allowing the detection of C4 to C24 fatty acid methyl esters. Helium flowing at 1.2 mL/min was used as carrier gas. The MS inlet line was set at 250 °C and EI source at 200 °C. Full scan monitoring was performed from 45 to 500 m/z. All data were collected and processed using Turbomass 6.1 (Perkin Elmer). FAMEs were identified by a spectral database search using MS Search 2.0 operated with the Standard Reference Database 1A (National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK). Retention time correlations with estimated nonpolar retention indexes from the NIST database were obtained using a 37-component FAME mix (Supelco; Sigma-Aldrich, Saint-Quentin Fallavier, France); FAME identifications were confirmed using this index).

API ZYM, API 20A, and API 50CH strips (bioMérieux) were used

to perform the biochemical test according to the manufacturer's instructions. The strips were incubated in anaerobic conditions and respectively for 4, 24, and 48 h. Antibiotic susceptibility was tested using the E-test gradient strip method (BioMerieux) to determine the minimal inhibitory concentration (MIC) of each tested antibiotic. Strain KHD7^T was grown on blood Colombia agar (BioMerieux) and a bacterial inoculum of turbidity 0.5 McFarland was prepared by suspending the culture in sterile saline (0.85% NaCl). Using cotton swabs, the inoculum was plated on 5% horse blood enriched Mueller Hinton Agar (BioMerieux) according to EUCAST recommendations [27,28]. E-test strips (amoxicillin, benzylpenicillin, imipenem, and vancomycin) were then deposited and the plates were incubated under anaerobic conditions for 48 h. Around the strip, Elliptic zones of inhibition were formed and the intersection with the strip indicates the MIC [28]. MICs were interpreted according to the EUCAST recommendations [29]. *Escherichia coli* strain DSM 1103 was used as a quality control strain.

2.6. Genomic DNA preparation

Strain KHD7^T was grown in anaerobic conditions at 37 °C using Columbia agar enriched with 5% sheep blood (bioMérieux) after 48 h on four Petri dishes. Bacteria were resuspended in 500 μL of TE buffer; 150 μL of this suspension was diluted in 350 μL 10× TE buffer, 25 μL proteinase K, and 50 μL sodium dodecyl sulfate for lysis treatment. This preparation was incubated overnight at 56 °C. DNA was purified using phenol/chloroform/isoamylalcohol successively for extraction and followed by ethanol precipitation at −20 °C of at least 2 h each. Following centrifugation, the DNA was suspended in 65 μL EB buffer. Genomic DNA concentration was measured at 46.06 ng/μL using the Qubit assay with the high-sensitivity kit (Life technologies, Carlsbad, CA, USA).

2.7. Genome sequencing and assembly

Genomic DNA of strain KHD7^T was sequenced on the MiSeq Technology (Illumina Inc., San Diego, CA, USA) with the mate pair strategy. The gDNA was barcoded with the Nextera Mate Pair sample prep kit (Illumina) in order to be mixed with 11 other projects.

gDNA was quantified by a Qubit assay with the high sensitivity kit (Life technologies, Carlsbad, CA, USA) to 26 ng/μL. The mate pair library was prepared with 1.5 μg of genomic DNA using the Nextera mate pair Illumina guide. The genomic DNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of the fragmentation was validated on 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 6.228 kb. No size selection was performed and 556 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal at 1275 bp on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc, Santa Clara, CA, USA) and the final concentration library was measured at 37.47 nmol/L.

The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 39-h run in a 2 × 151-bp. Total information of 6.5 Gb was obtained from 696 K/mm^2 cluster density with cluster passing quality control filters of 95.6% (12,863,000 passing filter paired reads). Within this run, the index representation for strain KHD7^T was determined at 6.26%. The 805,042 paired reads were trimmed then assembled in two scaffolds.

2.8. Genome annotation and analysis

Prodigal was used for Open Reading Frames (ORFs) prediction [30] with default parameters. We excluded predicted ORFs spanning a sequencing gap region (containing N). The bacterial proteome was predicted using BLASTP (E-value 1e^{-03} coverage 0.7 and identity percent 30) against the Clusters of Orthologous Groups (COG) database. If no hit was found we searched against the NR database [31] using BLASTP with E-value of 1e^{-03} coverage 0.7 and an identity percent of 30. An E-value of 1e^{-05} was used if sequence lengths were smaller than 80 amino acids. PFam conserved domains (PFAM-A an PFAM-B domains) were searched on each protein with the hhmscan tools analysis. RNAmmer [32] and tRNAScanSE tool [33] were used to find ribosomal RNAs genes and tRNA genes respectively. ORFans were identified if all the BLASTP performed had negative results (E-value smaller than 1e^{-03} for ORFs with sequence size above 80 aa or E-value smaller than 1e^{-05} for ORFs with sequence length below 80 aa). For data management and visualization of genomic features, Artemis [34] and DNA Plotter [35] were used, respectively. We used the MAGI in-house software to analyze the mean level of nucleotide sequence similarity at the genome level. It calculated the average genomic identity of gene sequences (AGIOS) among compared genomes [36]. This software combines the Proteinortho software [37] for detecting orthologous proteins in pairwise genomic comparisons. Then the corresponding genes were retrieved and the mean percentage of nucleotide sequence identity among orthologous ORFs was determined using the Needleman-Wunsch global alignment algorithm. Genomes from the genus *Atopobium*, *Olsenella*, and *Collinsella* were used for the calculation of AGIOS values. The genome of strain KHD7^T (FLLS00000000) was compared with that of *Olsenella uli* DSM 7084 (NC_014363); *Olsenella profusa* F0195 (AWEZ00000000); *Atopobium fossor* DSM 15642 (AXXR00000000); *Atopobium parvulum* DSM 20469 (NC_013203); *Atopobium rimae* ATCC 49626 (ACFE00000000); *Collinsella tanakaei* YIT 12063 (ADLS00000000). The Multi-Agent software system DAGOBAH [38] was used to perform annotation and comparison processes, which include Figenix [39] libraries that provide pipeline analysis. We also performed GGDC analysis using the GGDC web server as previously reported [40].

3. Results

3.1. Strain characterization

3.1.1. Strain identification by MALDI-TOF

Strain KHD7^T was first isolated in November 2015 after 10 days of pre-incubation in a blood culture bottle enriched with rumen and sheep blood under anaerobic conditions and sub-cultured on CNA agar with 5% sheep blood at 37 °C, also under anaerobic conditions. MALDI-TOF MS analysis of strain KHD7^T gave a low score (1.2), suggesting that our isolate was not in the database and could be a previously unknown species.

3.1.2. Strain identification by 16S rRNA sequencing gene

The 16S rRNA gene was then sequenced and the sequence obtained (accession number LN998058) shows 93.5% similarity with *Olsenella uli*, the phylogenetically closest bacterial species with a

validly published name (Fig. 1). As this value is lower than 95% threshold defined by Stackebrandt and Ebers for defining a new genus, we classified strain KHD7^T as the type strain of a new genus named *Olegusella* (Table 1). The reference spectrum was then added to our database (See Supplementary Table S1) and compared with those of the closest species (See Supplementary Table S2).

3.1.3. Phenotypic characteristics

Strain KHD7^T grew only in anaerobic conditions. Growth was observed at temperatures ranging from 25 to 42 °C, with optimal growth at 37 °C under anaerobic conditions after 48 h of incubation. The bacterium needed NaCl concentration below 0.5% and the pH for growth ranges from 6.5 to 7.0. On blood-enriched Columbia agar, colonies were pale white and translucent with a diameter of 1–1.2 mm. Gram-staining showed a rod-shaped Gram-positive bacterium (Fig. 2). On electron microscopy, individual cells appear with a mean diameter of 0.35 µm and a mean length of 0.42 µm (Fig. 3). Strain KHD7^T is non-motile and non-sporeforming.

The major fatty acid found for this strain was C16:0 acid (44%). Several unsaturated fatty acids were described including two abundant species: C18:2n6 (22%) and C18:1n9 (14%). Fatty acids with shorter aliphatic chains were also detected such as C8:0, C10:0, and C12:0 (Table 2).

Strain KHD7^T exhibited neither catalase nor oxidase activities.

Table 1
Classification and general features of *Olegusella massiliensis* strain khD7^T.

Properties	Terms
Taxonomy	**Kingdom**: *Bacteria* **Phylum**: *Acinetobacteria* **Class**: *Coriobacteriia* **Order**: *Coriobacteriales* **Family**: *Coriobacteriaceae* **Genus**: *Olegusella* **Species**: *Olegusella massiliensis*
Type strain	KhD7
Isolation site	Human vagina
Isolation country	France
Gram stain	Negative
Cell shape	Bacilli
Motility	No
Oxygen requirements	Anaerobic
Optimal temperature	37 °C
Temperature range	Mesophilic
Habitat	Host Associated
Biotic relationship	Free living
Host name	*Homo sapiens*
Sporulation	Nonsporulating
Metabolism	NA
Energy source	Chemoorganotrophic
Pathogenicity	Unknown
Biosafety level	2

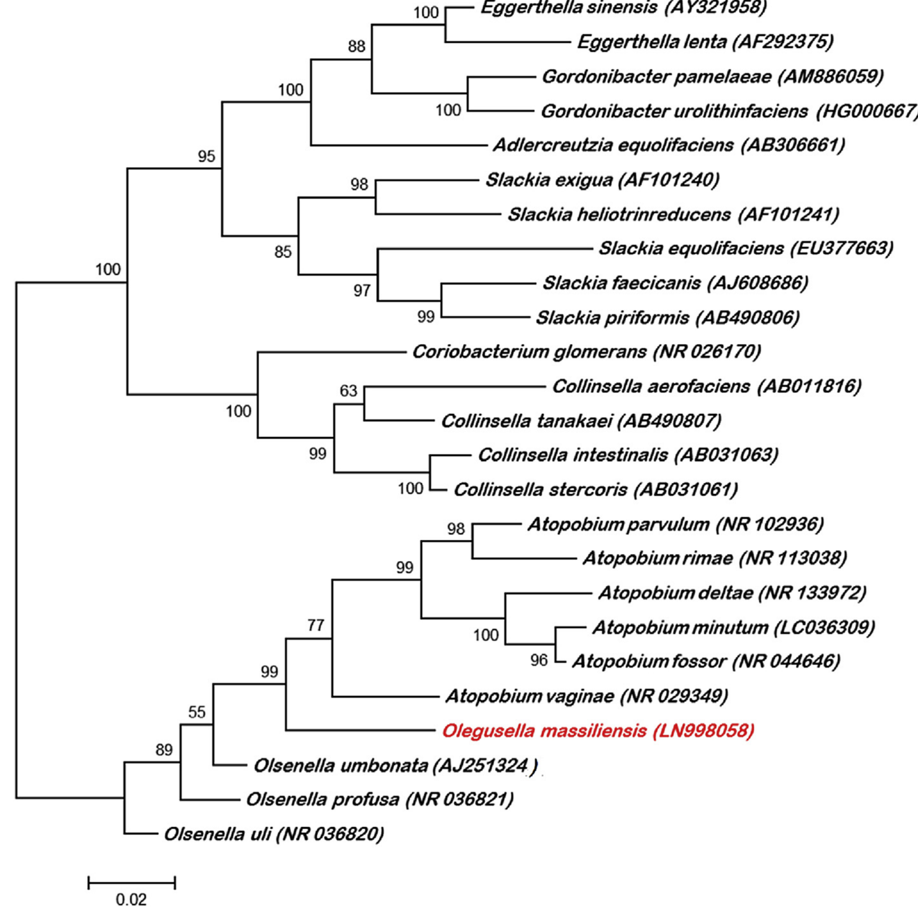


Fig. 1. Phylogenetic tree highlighting the position of *Olegusella massiliensis* strain KHD7^T relative to other close strains. GenBank accession numbers of each 16S rRNA is noted just after the name. Sequences were aligned using CLUSTALW, with default parameters and phylogenetic inferences were obtained using neighbor-joining method with 500 bootstrap replicates, within MEGA6 software. The scale bar represents a 2% nucleotide sequence divergence.

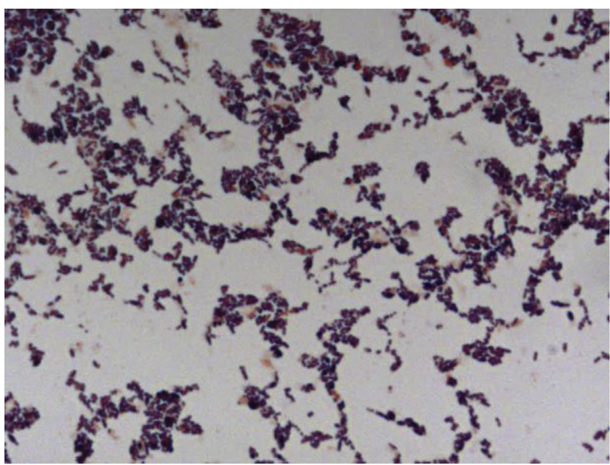


Fig. 2. Gram-staining of *Olegusella massiliensis* strain KHD7^T.

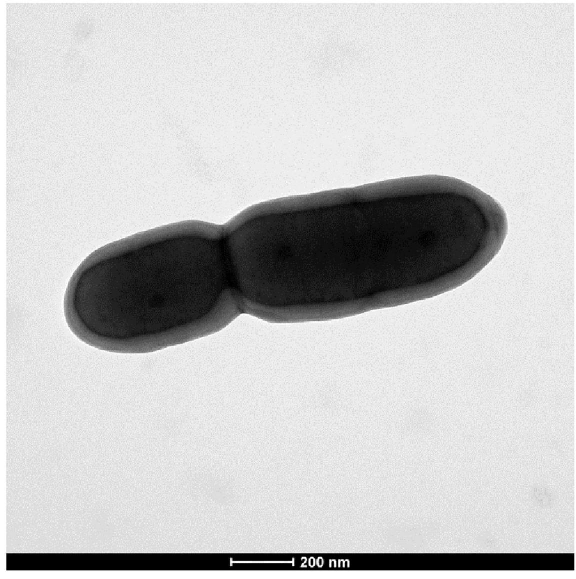


Fig. 3. Transmission electron microscopy of *Olegusella massiliensis* strain KHD7^T using a Tecnai G20 transmission electron microscope (FEI Company). The scale bar represents 200 nm.

Table 2
Cellular fatty acid composition (%).

Fatty acids	Name	Mean relative %[a]
16:0	Hexadecanoic acid	43.5 ± 0.7
18:2n6	9,12-Octadecadienoic acid	22.1 ± 0.4
18:1n9	9-Octadecenoic acid	13.8 ± 0.3
18:0	Octadecanoic acid	8.3 ± 0.1
14:0	Tetradecanoic acid	6.1 ± 0.4
10:0	Decanoic acid	1.6 ± 0.2
18:1n7	11-Octadecenoic acid	1.0 ± 0.1
18:1n6	12-Octadecenoic acid	TR
12:0	Dodecanoic acid	TR
15:0	Pentadecanoic acid	TR
16:1n7	9-Hexadecenoic acid	TR
15:0 anteiso	12-methyl-tetradecanoic acid	TR
15:0 iso	13-methyl-tetradecanoic acid	TR
8:0	Octanoic acid	TR

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

Using API ZYM strip, positive reactions were detected for leucine arylamidase, acid phosphatase, naphthol phosphohydrolase, and *N*-acetyl-beta-glucosaminidase but no reaction was observed for alkaline phosphatase, lipases (C4, C8 and C14), valine and cysteine arylamidase, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, α-mannosidase, and α-fucosidase. An API 50 CH strip revealed that strain KHD7^T metabolized D-glucose, D-mannose, *N*-acetylglucosamine, D-saccharose, and potassium 5-cetogluconate. This same strip show negative reactions for glycerol, erythritol, D-arabinose, arabinose (D and L), D-ribose, xylose, D-adonitol, methyl-βD-xylopyranoside, D-galactose, D-fructose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-αD-mannopyranoside, methyl- αD-glucopyranoside, amygdaline, arbutine, esculin ferric citrate, salicine, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-trehalose, inuline, D-melezitose, D-raffinose, starch, glycogene, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, fucose, arabitol, potassium gluconate, and potassium 2-cetogluconate. Based on API 20A strip, nitrate was not reduced, indole formation was negative. API 20A revealed also that esculin ferric citrate was hydrolyzed unlike gelatin.

Strain KHD7^T was susceptible to amoxicillin (MIC 0.38 μg/mL), benzylpenicillin (MIC 0.50 μg/mL), imipenem (MIC 1.25 μg/mL), and vancomycin (MIC 1 μg/mL). Phenotypic characteristics of strain KHD7^T compared with those of closely related species are shown in Table 3.

3.2. Genome properties

The final assembly identified two scaffolds (2 contigs) generating a genome size of 1,806,744 bp (1 chromosome, but no plasmid). The genome sequence was deposited in GenBank under accession number FLLS00000000. The G+C content was 49.24% (Table 4 and Fig. 4). Of the 1600 predicted genes, 1549 were protein-coding genes, and 51 were RNAs (two 5S rRNA, two 16S rRNA, two 23S rRNA, and 45 tRNA genes). A total of 1349 genes (87.08%) were assigned a putative function (by cogs or by NR blast): 54 genes were identified as ORFans (3.49%). The remaining genes were annotated as hypothetical proteins (224 genes, 14.46%). Genome statistics are summarized in Table 4. Genes are distributed according to COG functional categories in Table 5.

3.3. Genomic comparison

Compared to the genomes of other closed related species, the genome of strain KHD7^T (1.80 Mbp) is larger than those of *Atopobium fossor*, *Atopobium parvulum*, and *Atopobium rimae* (1.66; 1.54 and 1.63 Mbp respectively) but it is smaller than those of *Olsenella profusa*, *Olsenella uli*, and *Collinsella tanakaei* (2.72; 2.05; and 2.49 Mbp respectively). The G+C content of strain KHD7^T (49.24%) is smaller than those of *Olsenella uli*, *Olsenella profusa*, *Collinsella tanakaei*, and *Atopobium rimae* (64.70; 64.1; 60.2 and 49.30%, respectively) but larger than those of *Atopobium fossor* and *Atopobium parvulum* (45.4% and 45.70%, respectively). The gene content of strain KHD7^T (1,600) is smaller than those of *Olsenella uli*, *Olsenella profusa*, and *Collinsella tanakaei* (1,793, 2,474, and 2,150, respectively) but larger than those of *Atopobium fossor*, *Atopobium parvulum*, and *Atopobium rimae* (1,505, 1,406, and 1,511, respectively). However, the distribution of genes into COG categories was similar among all compared genomes (Fig. 5). In addition, strain KHD7^T shared on the one hand between 822 and 862 orthologous genes and on the other hand between 752 and 779 orthologous genes with the most closely related species belonging to the *Olsenella* and *Atopobium* genera (*O. uli*, *O. profusa* and *A. fossor*, *A. parvulum*, and *A. rimae*, respectively). Finally, it shared 745

Table 3
Differential characteristics of *Olegusella massiliensis* strain KHD7^T, *Olsenella uli* strain DSM 7084^T, *Olsenella umbonata* strain DSM 22620^T, *Olsenella profusa* strain DSM 13989^T, *Atopobium parvulum* strain ATCC 33793^T, *Atopobium rimae* strain ATCC 49626^T, *Atopobium fossor* strain NCTC 11919^T, *Atopobium deltae* strain CCUG 65171^T, and *Collinsella tanakaei* strain DSM 22478^T [40–46].

Properties	*Olegusella massiliensis*	*Olsenella uli*	*Olsenella umbonata*	*Olsenella profusa*	*Atopobium parvulum*	*Atopobium rimae*	*Atopobium fossor*	*Atopobium deltae*	*Collinsella tanakaei*
Cell diameter (μm)	0.3–0.4	na	0.3–0.6	0.6–0.8	0.3–0.6	na	0.5–0.9	1–1.2	0.5–1
Major fatty acid	$C_{16:0}$ (43.5%)	$C_{18:0}$ (31.7%)	$C_{18:0}$ (51%)	$C_{14:0}$-antesio (68.7%)	$C_{18:1}$ *cis*-9 FAME (38.2%)	$C_{18:1}$ *cis*-9 FAME (32.5%)	$C_{16:0}$ (33.3%)	C16:0 (33.3%)	C18:1 *cis*-9 FAME (44.91%)
DNA G+C content (mol%)	49.24	64.70	63	64.1	45.7	49.30	45.4	50.3	60.2
Production of									
Alkaline phosphatase	−	−	−	+	na	na	na	−	+
β-galactosidase		−	−	+	+	−	na	−	−
N-acetyl-glucosamine	+	−	−	+	na	na	na	−	−
Acid from									
Ribose	−	−	na	na	−	+	−	na	na
Mannitol	−	−	−	+	−	−	−	−	−
Sucrose	−	+	+	+	+	+	−	+	+
D-fructose	−	+	+	+	+	+	−	na	na
D-maltose	−	+	+	+	+	+	−	na	+
D-lactose	−	−	−	+	+	−	−	+	+
Habitat	Human vagina	Human gingival crevices	Sheep rumen	Human subgingival	Human gingival crevices	Human gingival crevices	Horse oropharyngeal	Human blood	Human faeces

+: positive reaction; −: negative reaction; na: not available data. Data are from literature except DNA G+C content which was calculated by EMBOSS software online (http://www.bioinformatics.nl/emboss-explorer/.

Table 4
Nucleotide content and gene count levels of the genome.

Attribute	Value	of total[a]
Size (bp)	1,806,744	100
G+C content (bp)	889,672	49.24
Coding region (bp)	1,610,188	89.12
Total genes	1600	100
RNA genes	51	3.18
Protein-coding genes	1549	96.81
Genes with function prediction	1349	87.08
Genes assigned to COGs	1219	78.69
Genes with peptide signals	125	8.06
Genes with transmembrane helices	371	23.95

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

orthologous genes with the most distant species belonging to the *Collinsella* genus (*C. tanakaei)* (Table 6). The same trend was observed when we analyzed the average percentage of nucleotide sequence identity, which ranged from 64.76% to 66.04% between *O. uli*, *O. profusa*, *A. parvulum*, *A. rimae*, and *A. fossor* species, but was 62.98% between strain KHD7^T and *C. tanakaei*. We obtained similar results for the analysis of the digital DNA-DNA hybridization (dDDH) using Genome-to-Genome Distance Calculator (GGDC) software (Table 7).

4. Discussion

Strain KHD7^T was isolated as part of a “culturomics” study of the vaginal flora aiming to isolate all bacterial species within the vagina. Strain KHD7^T was considered as a new genus on the basis of its unique MALDI-TOF MS spectrum, the genome comparison and its low 16S rRNA similarity level. The latter value was 93.5% with *O. uli*, which was lower than the recommended 95% threshold to define a new genus [22]. Strain KHD7^T is a member of the family *Coriobacteriaceae* belonging to the phylum *Actinobacteria*. This family comprises 35 species divided into 13 validated genera [13,14]. Most members of the *Coriobacteriaceae* are Gram-positive, non-motile, and non-sporulating bacteria. All these criteria are observed for *Olegusella massiliensis* strain KHD7^T. Bacterial species of the *Coriobacteriaceae* family have been detected in diverse habitats such as the intestinal tracts of humans and rodents, horse oropharynxes, human blood, and sheep rumen [41–46]. Furthermore, *Olsenella uli* was first isolated in the human gingival crevice; this bacterium is also associated with tissue destruction and periodontal inflammation [47].

A polyphasic taxono-genomics strategy [17,18], based on the combination of phenotypic and genomic analyses was used to characterize strain KHD7^T and the new genus from which it is the type strain. Phenotypically, strain KHD7^T exhibited a specific MALDI-TOF MS spectrum and differed from the other closed studied bacterial species in their fermentation of carbohydrate. Most often, the species of the *Coriobacteriaceae* family ferment glucose and mannose as observed for *Olegusella massiliensis*. Their differences lie on the fermentation of other carbohydrates such as ribose, mannitol, fructose, sucrose, lactose, and maltose. Unlike *O. uli*, *O. umbonata*, *O. profusa*, and *A. parvulum*, strain KHD7^T does not ferment sucrose, fructose, or maltose.

The G+C content of strain KHD7^T and its phylogenetically closest species varies from 45.4 to 64.70%. The genomic similarity of strain KHD7^T with species of *Coriobacteriaceae* family was evaluated by 2 parameters: DDH and AGIOS. The values found in DDH and AGIOS of *O. massiliensis* are in the range of those observed in the other genera of this family.

5. Conclusion

Based on the phenotypic analysis, phylogenetic and genomic results, strain KHD7^T may be a member of a new genus named *Olegusella* with *Olegusella massiliensis* as the type strain. It was isolated among the vaginal flora of a 33 year-old French woman suffering from bacterial vaginosis.

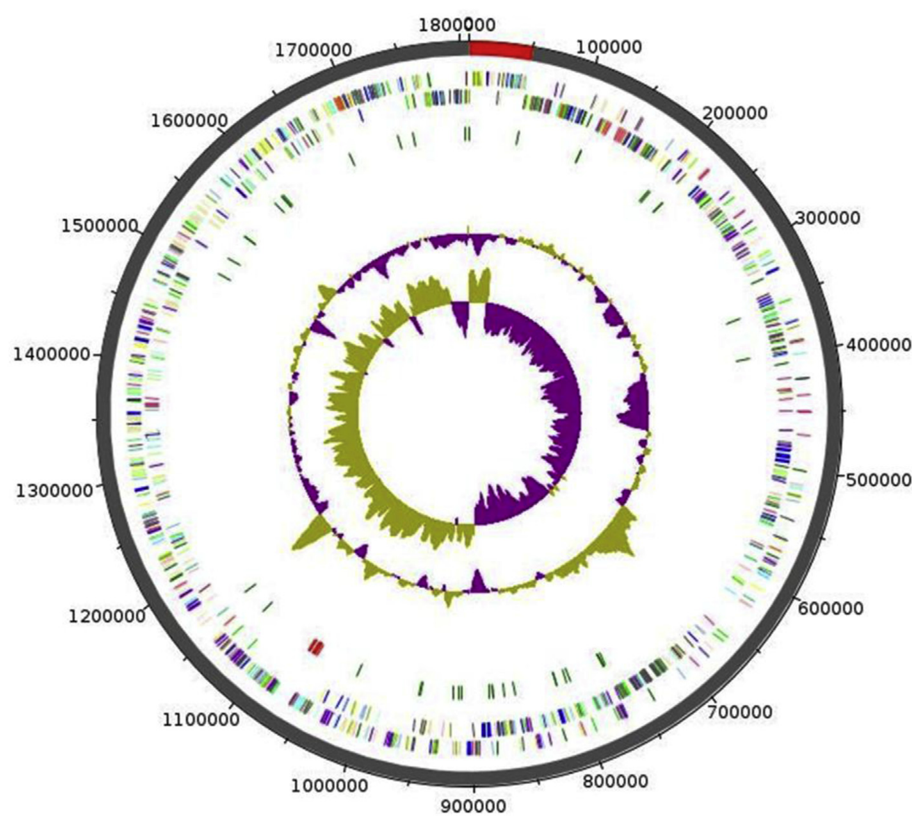


Fig. 4. Graphical circular map of the chromosome. From outside to the center: Genes on the forward strand colored by Clusters of Orthologous Groups of proteins (COG) categories (only genes assigned to COG). Genes on the reverse strand colored by COG categories (only gene assigned to COG). RNA genes (tRNAs green. rRNAs red). GC content and GC skew. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Number of genes associated with the 25 general COG functional categories.

Code	Value	% value	Description
J	125	10.25	Translation
A	0	0	RNA processing and modification
K	85	6.97	Transcription
L	74	6.07	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	17	1.39	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	54	4.43	Defense mechanisms
T	39	3.20	Signal transduction mechanisms
M	85	6.97	Cell wall/membrane biogenesis
N	2	0.16	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	15	1.23	Intracellular trafficking and secretion
O	45	3.69	Posttranslational modification, protein turnover, chaperones
X	6	0.49	Mobilome: prophages, transposons
C	53	4.35	Energy production and conversion
G	111	9.11	Carbohydrate transport and metabolism
E	113	9.27	Amino acid transport and metabolism
F	51	4.18	Nucleotide transport and metabolism
H	34	2.79	Coenzyme transport and metabolism
I	26	2.13	Lipid transport and metabolism
P	49	4.02	Inorganic ion transport and metabolism
Q	9	0.74	Secondary metabolites biosynthesis, transport and catabolism
R	121	9.93	General function prediction only
S	105	8.61	Function unknown
_	330	21.30	Not in COGs

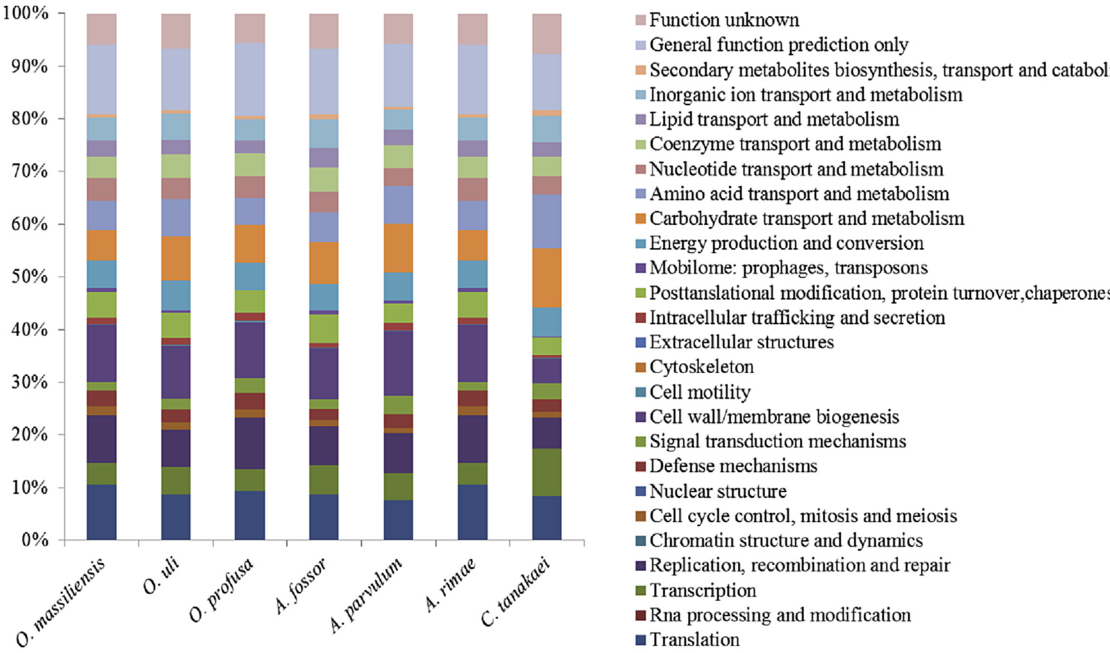


Fig. 5. Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins of *Olegusella massiliensis* strain KHD7^T among other species.

Table 6
Numbers of orthologous proteins shared between genomes (upper right) and AGIOS values obtained (lower left). The numbers of proteins per genome are indicated in bold.

	OM	OU	OP	AF	AP	AR	CT
OM	**1550**	862	822	779	755	752	745
OU	64.76%	**1775**	928	836	816	837	814
OP	64.81%	75.26%	**2593**	790	817	821	811
AF	66.04%	62.79%	62.74%	**1487**	758	753	743
AP	65.77%	63.02%	62.91%	66.67%	**1363**	899	716
AR	65.37%	64.62%	64.56%	65.65%	72.13%	**1478**	**718**
CT	62.98%	62.98%	67.42%	62.46%	62.59%	63.35%	**2194**

OM *Olegusella massiliensis* KHD7^T; OU *Olsenella uli* DSM 7084; OP *Olsenella profusa* F0195; AF *Atopobium fossor* DSM 15642; AP *Atopobium parvulum* DSM 20469; AR *Atopobium rimae* ATCC 49626; CT *Collinsella tanakaei* YIT 12063.

5.1. Taxonomic and nomenclatural proposals

5.1.1. Description of Olegusella gen. nov.

Olegusella (O.le.gu.sel'la. M.L. dim. suffix usel'la; M.L. fem. n.) was chosen to honor Dr. Oleg Mediannikov for his contribution to medical microbiology. Gram-stain-positive rods. Strictly anaerobic. Mesophilic. Non-motile. Does not exhibit catalase, oxidase nor nitrate reduction. Positive for D-glucose, D-mannose, *N*-acetylglucosamine, D-saccharose, potassium 5-cetogluconate, leucine arylamidase, acid phosphatase, naphthol phosphohydrolase, and *N*-acetyl-beta-glucosaminidase. Habitat: human vaginal flora. Type species: *Olegusella massiliensis*.

5.1.2. Description of Olegusella massiliensis gen. nov., sp. nov.

Olegusella massiliensis (mas.il'ien'sis, L. gen. fem. n. massiliensis, of Massilia, the Latin name of Marseille where the Type strain was first isolated).

Gram-stain-positive rods. Strictly anaerobic. Mesophilic. Optimal growth at 37 °C. Non-motile and non-sporulating. Colonies are pale white and translucent with 1–1.2 mm diameter on blood-enriched Colombia agar. Cells are rod-shaped with diameter approximately 0.35 μm and length approximately 0.42 μm. Strain KHD7^T exhibited neither catalase nor oxidase activities. Nitrate reduction is absent. Positive reactions were observed for D-glucose, D-mannose, *N*-acetylglucosamine, D-saccharose, potassium 5-cetogluconate, leucine arylamidase, acid phosphatase, naphthol phosphohydrolase, and *N*-acetyl-beta-glucosaminidase. The major fatty acids are C16:0 acid (44%), C18:2n6 (22%) and C18:1n9 (14%). Strain KHD7^T is susceptible to penicillin, oxacillin, ceftriaxone, imipenem, ciprofloxacin, clindamycin, erythromycin, gentamicin, metronidazole, rifampicin, teicoplanin, and vancomycin but it is resistant to colistin, doxycycline, fosfomycin and trimethoprim-sulfamethoxazole.

The 16S rRNA and genome sequences are deposited in GenBank

Table 7
dDDH values obtained by comparison of all studied genomes.

	OM	OU	OP	AF	AP	AR	CT
OM	100	25.10 ± 2.4	22.00 ± 2.35	22.00 ± 2.35	23.00 ± 2.35	20.80 ± 2.35	22.50 ± 2.4
OU		100	22.30 ± 2.35	21.70 ± 2.35	25.00 ± 2.4	24.90 ± 2.4	19.50 ± 2.3
OP			100	19.80 ± 2.3	24.00 ± 2.4	21.60 ± 2.35	20.00 ± 2.35
AF				100	20.30 ± 2.35	21.00 ± 2.3	23.60 ± 2.4
AP					100	23.90 ± 2.4	20.80 ± 2.35
AR						100	22.00 ± 2.35
CT							100

dDDH: Digital DNA-DNA hybridization. OM *Olegusella massiliensis* KHD7^T; OU *Olsenella uli* DSM 7084; OP *Olsenella profusa* F0195; AF *Atopobium fossor* DSM 15642; AP *Atopobium parvulum* DSM 20469; AR *Atopobium rimae* ATCC 49626; CT *Collinsella tanakaei* YIT 12063.

under accession numbers LN998058 and FLLS00000000 respectively. The genome is 1,806,744 bp long with a G+C content of 49.24%. The type strain KHD7^T (= CSUR P2268^T = DSM 101849^T) was isolated from the vaginal flora of a patient with bacterial vaginosis.

Conflict of interest

The authors declare no conflict of interest.

Funding

This study was supported by the Institut hospitalo-universitaire Mediterranée-Infection.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.anaerobe.2017.02.012.

References

[1] J.P. Lepargneur, V. Rousseau, Protective role of the Doderleïn flora, J. Gynecol. Obstet. Biol. Reprod. 31 (2002) 485–494.
[2] B. Chaban, M.G. Links, T.P. Jayaprakash, et al., Characterization of the vaginal microbiota of healthy Canadian women through the menstrual cycle, Microbiome 2 (2014) 23.
[3] K. Pal, S. Roy, B. Behera, et al., Bacterial vaginosis: etiology and modalities of treatment-A brief note, J. Pharm. Bioallied Sci. 3 (2011) 496.
[4] J.P. Lepargneur, V. Rousseau, Protective role of Doderleïn flora, J. Gynecol. Obstet. Biol. Reprod. 31 (2002) 485–494.
[5] S. Borges, J. Silva, P. Teixeira, The role of lactobacilli and probiotics in maintaining vaginal health, Arch. Gynecol. Obstet. 289 (2014) 479–489.
[6] J. Pépin, S. Deslandes, G. Giroux, et al., The complex vaginal flora of West African women with bacterial vaginosis, PLoS One 6 (2011) e25082.
[7] E. Shipitsyna, A. Roos, R. Datcu, et al., Composition of the vaginal microbiota in women of reproductive age – sensitive and specific molecular diagnosis of bacterial vaginosis is possible? PLoS One 8 (2013) e60670.
[8] C.S. Bradshaw, S.N. Tabrizi, C.K. Fairley, et al., The association of *Atopobium vaginae* and *Gardnerella vaginalis* with bacterial vaginosis and recurrence after oral metronidazole therapy, J. Infect. Dis. 194 (2006) 828–836.
[9] A.B. Onderdonk, M.L. Delaney, R.N. Fichorova, The human microbiome during bacterial vaginosis, Clin. Microbiol. Rev. 29 (2016) 223–238.
[10] R.F. Lamont, J.D. Sobel, R.A. Akins, et al., The vaginal microbiome: new information about genital tract flora using molecular based techniques, BJOG 118 (2011) 533–549.
[11] D.N. Fredricks, T.L. Fiedler, J.M. Marrazzo, Molecular identification of bacteria associated with bacterial vaginosis, N. Engl. J. Med. 353 (2005) 1899–1911.
[12] J.C. Lagier, F. Armougom, M. Million, et al., Microbial culturomics: paradigm shift in the human gut microbiome study, Clin. Microbiol. Infect. 18 (2012) 1185–1193.
[13] T. Clavel, C. Charrier, M. Wenning, et al., *Parvibacter caecicola* gen. nov., sp. nov., a bacterium of the family *Coriobacteriaceae* isolated from the caecum of a mouse, Int. J. Syst. Evol. Microbiol. 63 (2013) 2642–2648.
[14] Classification of bacteria, http://www.bacterio.net/-classifphyla.html#Actinobacteria, (Accessed: 06 March 2016).
[15] J.Z. Chan, M.R. Halachev, N.J. Loman, et al., Defining bacterial species in the genomic era: insights from the genus *Acinetobacter*, BMC Microbiol. 12 (2012) 302.
[16] P. Vandamme, B. Pot, M. Gillis, et al., Polyphasic taxonomy, a consensus approach to bacterial systematics, Microbiol. Rev. 60 (1996) 407–438.
[17] P.E. Fournier, M. Drancourt, New Microbes New Infections promotes modern prokaryotic taxonomy: a new section TaxonoGenomics: new genomes of microorganisms in humans, New Microbes New Infect. 7 (2015) 48–49.
[18] P.E. Fournier, J.C. Lagier, G. Dubourg, D. Raoult, From culturomics to taxonomogenomics: a need to change the taxonomy of prokaryotes in clinical microbiology, Anaerobe 36 (2015) 73–78.
[19] J.P. Menard, F. Fenollar, M. Henry, et al., Molecular quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis, Clin. Infect. Dis. 47 (2008) 33–43.
[20] P. Seng, M. Drancourt, F. Gouriet, et al., Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry, Clin. Infect. Dis. 49 (2009) 543–551.
[21] M. Drancourt, C. Bollet, A. Carlioz, et al., 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates, J. Clin. Microbiol. 38 (2000) 3623–3630.
[22] E. Stackebrandt, J. Ebers, Taxonomic parameters revisited: tarnished gold standards, Microbiol. Today 33 (2006) 152–155.
[23] M. Kim, H.-S. Oh, S.-C. Park, J. Chun, Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes, Int. J. Syst. Evol. Microbiol. 64 (2014) 346–351.
[24] P. Yarza, P. Yilmaz, E. Pruesse, et al., Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences, Nat. Rev. Micro 12 (2014) 635–645.
[25] These standards outline the test procedures for identifying the organisms, https://www.gov.uk/government/collections/standards-for-microbiology-investigations-smi#test-procedures, (Accessed 06 March 2016).
[26] M. Sasser, Bacterial Identification by Gas Chromatographic Analysis of Fatty Acids Methyl Ester (GC-FAME), 1990.
[27] E. Matuschek, D.F.J. Brown, G. Kahlmeter, Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories, Clin. Microbiol. Infect. 20 (2014) 255–266.
[28] D.M. Citron, M.I. Ostovari, A. Karlsson, E.J. Goldstein, Evaluation of the E test for susceptibility testing of anaerobic bacteria, J. Clin. Microbiol. 29 (1991) 2197–2203.
[29] The European Committee on Antimicrobial Susceptibility Testing, Breakpoint Tables for Interpretation of MICs and Zone Diameters. http://www.eucast.org (Accessed 22 November 2016).
[30] D. Hyatt, G.L. Chen, P.F. Locascio, et al., Prodigal: prokaryotic gene recognition and translation initiation site identification, BMC Bioinforma. 11 (2010) 119.
[31] K. Clark, I. Karsch-Mizrachi, D.J. Lipman, et al., GenBank, Nucleic Acids Res. 44 (2016) D67–D72.
[32] K. Lagesen, P. Hallin, E.A. Rødland, et al., RNAmmer: consistent and rapid annotation of ribosomal RNA genes, Nucleic Acids Res. 35 (2007) 3100–3108.
[33] T.M. Lowe, S.R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, Nucleic Acids Res. 25 (1997) 955–964.
[34] T. Carver, S.R. Harris, M. Berriman, et al., Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data, Bioinformatics 28 (2012) 464–469.
[35] T. Carver, N. Thomson, A. Bleasby, et al., DNAPlotter: circular and linear interactive genome visualization, Bioinformatics 25 (2009) 119–120.
[36] D. Ramasamy, A.K. Mishra, J.C. Lagier, et al., A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species, Int. J. Syst. Evol. Microbiol. 64 (2014) 384–391.
[37] M. Lechner, S. Findeiss, L. Steiner, et al., Proteinortho: detection of (Co-)orthologs in large-scale analysis, BMC Bioinforma. 12 (2011) 124.
[38] P. Gouret, J. Paganini, J. Dainat, et al., 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, Springer Berlin Heidelberg, 2011, pp. 71–87.
[39] P. Gouret, V. Vitiello, N. Balandraud, et al., FIGENIX: intelligent automation of genomic annotation: expertise integration in a new software platform, BMC Bioinforma. 6 (2005) 198.
[40] J.P. Meier-Kolthoff, A.F. Auch, H.P. Klenk, et al., Genome sequence-based species delimitation with confidence intervals and improved distance functions, BMC Bioinforma. 14 (2013) 60.
[41] P. Cools, M. Oyaert, M. Vaneechoutte, et al., *Atopobium deltae* sp. nov., isolated from the blood of a patient with Fournier's gangrene, Int. J. Syst. Evol. Microbiol. 64 (2014) 3140–3145.
[42] F.E. Dewhirst, B.J. Paster, N. Tzellas, 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. 51 (2001) 1797–1804.
[43] G.D. Bailey, D.N. Love, *Eubacterium fossor* sp. nov., an agar-corroding organism from normal pharynx and oral and respiratory tract lesions of horses, Int. J. Syst. Evol. Microbiol. 36 (1986) 383–387.
[44] I. Olsen, J.L. Johnson, L.V.H. Moore, et al., *Lactobacillus uli* sp. nov. and *Lactobacillus rimae* sp. nov. from the human gingival crevice and emended descriptions of *Lactobacillus minutus* and *Streptococcus parvulus*, Int. J. Syst. Evol. Microbiol. 41 (1991) 261–266.
[45] M. Kraatz, R.J. Wallace, L. Svensson, *Olsenella umbonata* sp. nov., a microaerotolerant anaerobic lactic acid bacterium from the sheep rumen and pig jejunum, and emended descriptions of *Olsenella*, *Olsenella uli* and *Olsenella profusa*, Int. J. Syst. Evol. Microbiol. 61 (2011) 795–803.
[46] F. Nagai, Y. Watanabe, M. Morotomi, *Slackia piriformis* sp. nov. and *Collinsella tanakaei* sp. nov., new members of the family *Coriobacteriaceae*, isolated from human faeces, Int. J. Syst. Evol. Microbiol. 60 (2010) 2639–2646.
[47] A.P. Vieira Colombo, C.B. Magalhães, F.A. Hartenbach, et al., Periodontal-disease-associated biofilm: a reservoir for pathogens of medical importance, Microb. Pathog. 94 (2015) 27–34.

<u>Article 5</u>:

Description of *Collinsella vaginalis* strain Marseille-P2666, a new member of *Collinsella* genus isolated from the genital tract of a patient suffering from bacterial vaginosis.

Awa Diop, **Khoudia Diop**, Enora Tomei, Nicholas Armstrong, Florence Bretelle, Didier Raoult, Florence Fenollar, Pierre-Edouard Fournier

In revision for International Journal of Systematic and Evolutionary Microbiology

1 ***Collinsella vaginalis* sp. nov. strain Marseille-P2666^{T}, a new member of the *Collinsella***

2 **genus isolated from genital tract of a patient suffering from bacterial vaginosis**

3

4 Awa Diop[1], Khoudia Diop[1], Enora Tomei[1], Nicholas Armstrong[1], Florence Bretelle[1, 3],

5 Didier Raoult[2,4], Florence Fenollar[1], Pierre-Edouard Fournier[1*]

6

7 [1]UMR VITROME, Aix-Marseille Université, IRD, Service de Santé des Armées,

8 Assistance Publique-Hôpitaux de Marseille, Institut hospitalo-universitaire Méditerranée-

9 infection, 19-21 Boulevard Jean Moulin 13005 Marseille, France Tel: +33 413 732 401, Fax:

10 +33 413 732 402

11 [2]UMR MEPHI, Aix-Marseille University, IRD, Assistance Publique-Hôpitaux de

12 Marseille, Institut Hospitalo-Uuniversitaire Méditerranée Infection, Marseille, France

13 [3]Department of Gynecology and Obstetrics, Gynépole, Marseille, Hôpital Nord,

14 Assistance Publique-Hôpitaux de Marseille

15 [4]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz

16 University, Jeddah, Saudi Arabia

17

18

19 *Corresponding author: Pr Pierre-Edouard Fournier

20 [1]UMR VITROME, Aix-Marseille Université, IRD, Service de Santé des Armées,

21 Assistance Publique-Hôpitaux de Marseille, Institut hospitalo-universitaire Méditerranée-

22 infection, 19-21 Boulevard Jean Moulin 13005 Marseille, France Tel: +33 413 732 401, Fax:

23 +33 413 732 402

24 **E-mail**: pierre-edouard.fournier@univ-amu.fr

25 **Running title**: *Collinsella vaginalis* sp. nov.

26 **Abstract words count**: 137

27 **Text word count**: 3684

28 **Keywords**: *Collinsella vaginalis*; bacterial vaginosis; microbial culturomics; taxono-

29 genomics; anaerobic bacteria; new species

30 **ABSTRACT**

31 A strictly anaerobic, Gram-stain-positive, non motile and non-spore-forming rod-shaped
32 bacterium, strain Marseille-P2666, was isolated from a vaginal sample of a French patient
33 suffering from bacterial vaginosis using the culturomics approach. Cells were saccharolytic
34 and were negative for catalase, oxidase, urease, nitrate reduction, indole production,
35 hydrolysis of aesculin and gelatin. Strain Marseille-P2666^T exhibited 97.04% 16S rRNA
36 sequence similarity with *Collinsella tanakaei* type strain YIT 12063^T, the phylogenetically
37 closest species with standing in nomenclature. The major fatty acids were $C_{18:1\omega9}$ (38%), $C_{16:0}$
38 (24%) and $C_{18:0}$ (19%). The G+C content of the genome sequence of strain Marseille-P2666 is
39 64.6 mol%. On the basis of its phenotypic, phylogenetic and genomic features, strain
40 Marseille-P2666^T (= CSUR 2666^T = DSM103342^T) was classified as type strain of a novel
41 species within the genus *Collinsella* for which the name *Collinsella vaginalis* sp. nov. is
42 proposed.

43 Investigating the microbial diversity of bacterial vaginosis is part of the ongoing
44 "Microbial Culturomics" project in our institute [1, 2], which consists in optimizing culture
45 conditions to explore in depth the human microbiota. In 2015, we isolated a strictly anaerobic
46 bacterial strain, strain Marseille-P2666^T, from a vaginal sample of a French woman patient
47 suffering with Bacterial vaginosis (BV). Strain Marseille-P2666^T was classified as belonging
48 to the genus *Collinsella*.

49 The genus C*ollinsella*, belonging to the family *Coriobacteriaceae* in the phylum
50 Actinobacteria [3], was first described by Kageyama *et al.* in 1999 [4]. On the basis of 16S
51 rRNA gene sequence and cell wall peptidoglycan divergence with other members of the genus
52 *Eubacterium*, these authors reclassified *Eubacterium aerofaciens* into a the new genus
53 *Collinsella*, with *Collinsella aerofaciens* being the type species [4]. Currently, five
54 C*ollinsella* species have standing in nomenclature (www.bacterio.net), namely *C. aerofaciens*
55 [4], *C. stercoris* [5], *C. intestinalis* [5], *C. tanakaei* [6] and *C. massiliensis* [7], all of which
56 had been isolated from the gastro-intestinal tract of healthy humans. All five species are non
57 spore-forming, non motile, rod-shaped cocci and contain an A4P-type peptidoglycan [4].

58 Thanks to the availability of genomic data from many bacterial species, we proposed
59 since 2012 to include the complete genome sequence analysis in a polyphasic approach for
60 the classification and description of new bacterial taxa, that we named named taxono-
61 genomics [8]. On the basis of the analysis of phenotypic and phylogenetic characteristics,
62 proteomic informations obtained by MALDI-TOF MS and genomics properties [8–10], we
63 describe here a new *Collinsella* species for which we propose the name *Collinsella vaginalis*
64 sp. nov.. Strain Marseille-P2666^T (= CSUR 2666^T = DSM103342^T) is the type strain of *C.*
65 *vaginalis* sp. nov.
66

67 Strain Marseille-P2666 was isolated in May 2015 from a vaginal sample of a 26 year-
68 old French woman diagnosed with bacterial vaginosis at the Nord hospital in Marseille,
69 France. The sample was collected using a Sigma Transwab (Medical Wire, Corsham, United
70 Kingdom) and then transported immediately to the microbiology laboratory of the Timone
71 Hospital in Marseille. The patient was not treated with any antibiotic at the time of sampling.
72 She gave an informed and signed consent and the study was validated by the ethics committee
73 of the IFR48 (Marseille, France) under agreement 09-022. For strain isolation, the vaginal
74 sample was first inoculated in an anaerobic blood culture bottle (Bactec Lytic/10 Anaerobic/F
75 Culture Vials, Becton-Dickinson, Le Pont de Claix, Isère, France) supplemented with 4 mL
76 filter-sterilized rumen fluid through a 0.2 μm pore filter (Thermo Fisher Scientific, Villebon-
77 sur-Yvette, France) and 3 mL of sheep blood (bioMérieux, Marcy l'Etoile, France) and
78 incubated at 37°C. After 15 days of incubation, 50 μL of the supernatant was inoculated on
79 5% sheep blood-enriched CNA agar (Colistin and Naladixic Acid) (Becton-Dickinson) and
80 incubated for 7 days in anaerobic atmosphere (0% O2, 100% CO2 and 100% N2) at 37°C.
81 Isolated colonies were subcultured individually using the same conditions and each
82 colony was deposited on a MTP 96 MALDI-TOF target plate (Bruker Daltonics, Leipzig,
83 Germany) in duplicate for identification with a Microflex MALDI-TOF MS spectrometer
84 (Bruker Daltonics, Leipzig, Germany), as described by Seng *et al*. [11]. The obtained protein
85 spectra were compared with those of 8687 reference spectra in the Bruker database constantly
86 enriched with our own database [12]. If the MALDI-TOF MS score was greater than 1.9 and
87 2.3, the bacterium was identified at the genus and species levels respectively. Conversely, if
88 the score was lower than these threshold, the identification was not considered as reliable and
89 the 16S rRNA gene was amplified and sequenced using the GeneAmp PCR System 2720
90 thermal cycler (Applied Bio systems, Bedford, MA, USA) and an ABI Prism 3130-XL
91 capillary sequencer (Applied Biosciences, Saint Aubin, France), respectively, as previously

92 described [13]. The obtained sequence was corrected using the Chromas Pro 1.34 software
93 (Technelysium Pty. Ltd., Tewantin, Australia) and then compared to the NCBI database using
94 the BLASTn algorithm (https://blast.ncbi.nlm.nih.gov/) for taxonomic assignment. The 16S
95 rRNA sequences of type strains from the species with a validly published name
96 (http://www.bacterio.net/) exhibiting the closest phylogenetic relationship with strain
97 Marseille-P2666 were downloaded from NCBI (ftp://ftp.ncbi.nih.gov/Genome/). Sequences
98 were aligned using MUSCLE [14]. Then, the degree of pairwise 16S rRNA sequence
99 similarity between strain Marseille-P2666 and other closely related species were calculated
100 using the GGDC web server [15] available at (http://ggdc.dsmz.de/) using the method
101 proposed by Meier-Kolthoff [16]. Phylogenetic trees were inferred in the GGDC web server
102 [15] using the DSMZ phylogenomics pipeline [17] adapted to single genes. Maximum
103 likelihood (ML) and maximum parsimony (MP)-based trees were inferred from the alignment
104 with RAxML [18] and TNT [19], respectively. For ML, rapid bootstrapping in conjunction
105 with the autoMRE bootstopping criterion [20] and subsequent search for the best tree was
106 used. The ML tree was inferred under the GTR+GAMMA model. For MP tree analysis, all
107 sites with gaps were removed and 1000 bootstrapping replicates were used in conjunction
108 with tree-bisection-and-reconnection branch swapping and ten random sequence addition
109 replicates. The sequences were checked for a compositional bias using the X^2 test as
110 implemented in PAUP* [21]. A supplementary phylogenetic tree using the Neighbor-joining
111 method is presented in supplementary data. If the 16S rRNA sequence similarity value was
112 lower than 95% or 98.65% with the most closely related species with standing in
113 nomenclature, as proposed by Stackebrandt and Ebers [22], the strain was proposed to belong
114 to a new genus or species, respectively [23].
115 In order to evaluate its ideal growth conditions, strain Marseille-P2666 was cultivated
116 on 5% sheep blood-enriched Columbia agar (bioMérieux) at various temperatures (25, 28, 37,

117 45, 56°C) under aerobic conditions with or without 5% CO_2, and in anaerobic (0% O2, 100%
118 CO2 and 100% N2) and microaerophilic atmospheres (5% O 2, 10% CO 2 and 85% N 2)
119 using GENbag Anaer and GENbag microaer systems (bioMérieux) respectively. The
120 tolerance to various NaCl concentrations (5 – 100 g/l NaCl) and pH values (pH 5, 6, 6.5, 7,
121 8.5) conditions was also tested. To observe the cell morphology, cells were fixed with 2.5%
122 glutaraldehyde in a 0.1M cacodylate buffer at 4°C for at least an hour. One drop of cell
123 suspension was deposited for approximately five minutes on glow-discharged formvar carbon
124 film on 400 mesh nickel grids (FCF400-Ni, EMS). The grids were dried on blotting paper and
125 the cells were negatively stained for 10 seconds with 1% ammonium molybdate solution in
126 filtered water at RT. Electron micrographs were acquired using a Tecnai G20 Cryo (FEI
127 company, Limeil-Brevannes, France) transmission electron microscope operated at 200 keV.
128 Gram-stain, motility and sporulation were performed as previously described [24].
129 The biochemical properties of strain Marseille-P2666 were evaluated using API ZYM,
130 API 20A, and API rapid ID 32A strips (bioMérieux) according to the manufacturer's
131 instructions. The strips were incubated in anaerobic conditions (0% O2, 100% CO2 and 100%
132 N2) at 37°C for 4, 24, and 4 hours respectively. Oxidase activity was tested using an oxidase
133 reagent (Becton-Dickenson, Le Pont de Claix, and France) and catalase activity was assessed
134 in 3% hydrogen peroxide solution (bioMérieux).
135 Amoxicillin (0.016-256 µg/mL), benzylpenicillin (0.002-32 µg/mL), ceftriaxone (0.016-
136 256 µg/mL), vancomycin (0.016-256 µg/mL), metronidazole (0.016-256 µg/mL), rifampicin
137 (0.002-32 µg/mL) and imipenem (0.002-32 µg/mL) were used to test the antibiotic
138 susceptibility of strain Marseille-P2666. The minimal inhibitory concentrations (MICs) were
139 then determined using E-test gradient strips (bioMérieux) according to the EUCAST
140 recommendations [25, 26].

141 Cellular fatty acid methyl ester (FAME) analysis was performed using Gas
142 Chromatography/Mass Spectrometry (GC/MS). Strain Marseille-P2666 was grown on 5%
143 sheep blood-enriched Columbia agar (bioMérieux). Two samples were then prepared with
144 approximately 16 mg of bacterial biomass per tube harvested from several culture plates.
145 Fatty acid methyl esters were prepared as described by Sasser [27]. GC/MS analyses were
146 carried out as described before [28]. Briefly, fatty acid methyl esters were separated using an
147 Elite 5-MS column and monitored by mass spectrometry (Clarus 500 - SQ 8 S, Perkin Elmer,
148 Courtaboeuf, France). Spectral database search was performed using MS Search 2.0 operated
149 with the Standard Reference Database 1A (NIST, Gaithersburg, USA) and the FAMEs mass
150 spectral database (Wiley, Chichester, UK). In addition, glucose fermentation products
151 measurements were performed. For this, strain Marseille-P2666 was inoculated on freshly
152 prepared Wilkins-Chalgren Anaerobe (WCA) culture broth, containing glucose [29] and
153 hemoculture flasks were incubated during 8 days at 37 °. Hydrogen, formic acid and ethanol
154 were measured from 3 independent blank and sample culture flasks using a Clarus 580 gas
155 chromatography system (Perkin Elmer, Villebon-sur-Yvette, France), a Clarus 500
156 chromatography system connected to a SQ8s mass spectrometer (Perkin Elmer) and a
157 Turbomatrix 40 Headspace sampler connected to a Clarus 500 chromatography system
158 equipped with a SQ8s mass spectrometer (Perkin Elmer), respectively. Quantities in samples
159 were given after subtraction the quantities measured in the blank flasks. Calibration curves
160 were calculated from chromatogram peak areas. Coefficients of determination were above
161 0.999, and back calculated standards all showed good accuracy with deviations below 15 %.
162 Formic acid and 2-ethylbutyric acid were high purity standards; water, ethanol, 2-propanol
163 and methanol were HPLC-grade solvents (Sigma Aldrich, Lyon, France). A more detailed
164 description of the glucose fermentation products measurements is presented in the
165 supplementary data section.

166 The genomic DNA (gDNA) of the strain Marseille-P2666^T was sequenced using a
167 MiSeq sequencer (Illumina Inc, San Diego, CA, USA) with the Mate Pair strategy. The
168 gDNA was quantified by a Qubit assay with the high sensitivity kit (Life technologies,
169 Carlsbad, CA, USA) to 68.1 ng/µl and a total of sequencing output of 5.1 Gb was obtained
170 from a 542K/mm^2 cluster density with a cluster passing quality control filters of 95.7%
171 (10,171,000 clusters). The 801,260 reads obtained by sequencing were trimmed, then
172 assembled using the Spades assembler program [30]. A more detailed description of the
173 sequencing methodology as well as the complete annotation of the genome is presented in the
174 supplementary data section.

175 A MALDI-TOF-MS score of 1.3 was obtained for strain Marseille-P2666 against our
176 database, suggesting that this isolate was not identified in the genus and species levels. The
177 MALDI-TOF MS spectrum from strain Marseille-P2666 was added to our database to
178 improve its content.

179 Using the Smith–Waterman algorithm [16], the 16S rDNA-based comparison of strain
180 Marseille-P2666 *(*EMBL-EBI accession number LT598547) against GenBank yielded a
181 highest nucleotide sequence similarity of 97.04% with *C. tanakaei* strain YIT 12063^T
182 (GenBank accession number AB490807), the phylogenetically-closest species with a validly
183 published name. As this value was lower than the 98.65% 16S rRNA sequence identity
184 threshold proposed to delineate a new species [22, 31], strain Marseille-P2666 was considered
185 as a potential new species within the genus *Collinsella* in the family *Coriobacteriaceae*. The
186 resulting combined ML/MP tree and the Neighbor-joining tree highlighting the position of
187 *Collinsella vaginalis* strain Marseille-P2666 relative to other close strains with a validly
188 published name is shown in Figure 1 and Figure 2, respectively.

189 For the phylogenetic inferences, the input nucleotide matrix comprised 21 operational
190 taxonomic units and 1,572 characters, 500 of which were variable and 351 of which were

191 parsimony-informative. The base-frequency check indicated a compositional bias (p = 0.00, α
192 = 0.05). ML analysis under the GTR+GAMMA model yielded a highest log likelihood of -
193 8308.08, whereas the estimated alpha parameter was 0.20. The ML bootstrapping did not
194 converge, hence 1,000 replicates were performed; the average support was 72.67%. MP
195 analysis yielded a best score of 1315 (consistency index 0.57, retention index 0.66) and 6 best
196 trees. The MP bootstrapping average support was 77.17%.

197 Colonies from strain Marseille-P2666 on CNA agar (Becton-Dickinson, Le pont de
198 Claix, France) under anaerobic atmosphere are grey, opaque and circular with a diameter of
199 0.5-1.2 mm after 48 hours of growth at 37°C. The growth was obtained at temperatures
200 ranging from 28 to 45 with optimal growth observed at 37°C in anaerobic atmosphere. No
201 growth was obtained in neither aerobic nor microaerophilic atmospheres. Strain Marseille-
202 P2666 needed a NaCl concentration below 5g/L and a pH ranging from 6.5 to 7.0 for its
203 growth. Bacterial cells are rod-shaped Gram-stain-positive, non-motile and non spore-forming
204 with a mean diameter of 0.4 µm and mean length of 1.8 µm and occur as single cells or in
205 short chains (Figure 3). No oxidase or catalase activity was observed.

206 Using an API ZYM strip (bioMérieux), positive results were obtained for esterase
207 (C4), esterase lipase (C8), alkaline phosphatase, leucine arylamidase, valine arylamidase,
208 cystine arylamidase, acid phosphatase, naphtol-AS-BI-phosphohydrolase and N-acetyl-β-
209 glucosaminidase but no reaction was observed for lipase (14), trypsin, α-chymotrypsin, α-
210 galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, α-mannosidase
211 and α-fucosidase. Using a Rapid ID32A strip (bioMérieux), positive reactions were obtained
212 for N-Acetyl-ß-glucosaminidase, mannose fermentation, raffinose fermentation, alkaline
213 phosphatase, arginine arylamidase, proline arylamidase, leucyl glycine arylamidase, leucine
214 arylamidase, glycine arylamidase, histidine arylamidase and serine arylamidase. Cells showed
215 no urease, arginine dihydrolase, α-galactosidase, β-galactosidase, 6-phospho-β-galactosidase,

216 α-glucosidase, β-glucosidase, α-arabinosidase, β-glucuronidase, glutamic acid decarboxylase,
217 α-fucosidase, reduction of nitrates, indole production, phenylalanine arylamidase,
218 pyroglutamic acid arylamidase, tyrosine arylamidase and glutamyl-glutamic acid arylamidase
219 activity. Using an API 20A strip (bioMérieux), strain Marseille-P2666 produced acid from D-
220 glucose, D-lactose, D-saccharose, D-maltose, salicin, D-cellobiose, D-mannose and D-
221 trehalose but not from D-mannitol, D-xylose, L-arabinose, gelatin, glycerol, D-melezitose, D-
222 raffinose, sorbitol and D-rhamnose. Esculin ferric citrate was not hydrolyzed. Indole
223 formation and urease activity were negative. Strain Marseille-P2666 differed from other
224 members of the *Collinsella* genus [4–7] in esterase, esterase lipase and cystine arylamidase
225 activities (Table 1). The most abundant cellular fatty acid found for strain Marseille-P2666
226 was the unsaturated acid $C_{18:1\omega9}$ (38%), followed by the saturated acids $C_{16:0}$ and $C_{18:0}$ (24 and
227 19%, respectively) (Table 2). The gas chromatography analysis of 8 days hemocultures
228 revealed a production of 114 ± 4 ppm of hydrogen (n=3), 8.8 ± 2.2 mM of formic acid (n=3)
229 and 4.8 ± 0.3 mM of ethanol (n=3). In addition, 6 genes encoding lactate dehydrogenase
230 associated with production of lactate were identified in the annotated genome with COG
231 databases. These results confirm the ability of this strain to perform the fermentation of
232 glucose to ethanol, formate, hydrogen and lactate as described in the Bergey's Manual of
233 Systematics of Archaea and Bacteria [32]. Cells are susceptible to benzylpenicillin (MIC 0.38
234 µg/mL), amoxicillin (MIC 0.064 µg/mL), metronidazole (MIC 0.75 µg/mL), rifampicin (MIC
235 0.008 µg/mL), vancomycin (MIC 4 µg/mL) but resistant to ceftriaxone (MIC > 256 µg/mL)
236 and imipenem (MIC > 32 µg/mL).
237 The draft genome of strain Marseille-P2666 is 2,162,909-bp long and has a G+C
238 content of 64.6 mol% (Table S1, Figure S1). It is composed of 23 scaffolds composed of 63
239 contigs. Of the 1,907 predicted genes, 1,696 were protein-coding genes and 53 were RNAs (1
240 complete rRNA operon, 47 tRNA genes and 3 ncRNA genes). A total of 1,303 genes (76.8%)

241 were assigned a putative function (by BLAST against the COGs or NR databases). A total of
242 121 genes were identified as ORFans (7.1%). The remaining 272 genes were annotated as
243 hypothetical proteins (16.0%). Strain Marseille-P2666 has many genes related to virulence,
244 including 13 bacteriocin-encoding genes (0.8%) and 50 toxin/ antitoxin modules (2.9%). By
245 using PHAST and RAST, 691 genes (40.7%) were associated with mobile genetic elements.
246 In addition, the genome of *collinsella vaginalis* exhibited 6 genes (*murE*; *MraY*; *murF*; *murC*;
247 *murD* and *murQ*) of the 20 genes found in the genome of *C. aerofaciens*, involving in
248 peptidoglycan type A4 biosynthesis and 1 gene (murJ) that encodes enzymes involved in
249 degradation of peptidoglycan (murein). Genome statistics are summarized in Table S1 and the
250 gene distribution into COGs functional categories is presented in Table S2.

251 The draft genome sequence structure of strain Marseille-P2666 is summarized in Figure
252 S1. It is smaller than those of *C. aerofaciens, Collinsella tanakei* and *C. stercoris* (2.2, 2.4,
253 2.5 and 2.5 Mb, respectively), but larger than those of *C. intestinalis* (1.8 Mb). The G+C
254 content of strain Marseille-P2666 (64.6 %) is greater than those of all compared *Collinsella*
255 species (Table S3). The gene content of strain Marseille-P2666 (1,907) is smaller than those
256 of *C. stercoris, Collinsella tanakei* and *C. aerofaciens* (2,119, 2,253 and 2437, respectively)
257 but larger than those of *C. intestinalis* (1,630) (Table S3). The gene distribution into COG
258 categories was similar among all compared genomes (Figure S2). However, *C. vaginalis*
259 possessed fewer predicted genes of the “Mobilome: prophages, transposons” category than
260 other compared *Collinsella* species (Figure S2). In addition, strain Marseille-P2666 exhibited
261 digital DNA–DNA hybridization (dDDH) values of 22.4% with *C. aerofaciens* to 23.2% with
262 *C. stercoris* (Table S4). Moreover, we observed AAI values of 64.7 to 66.9% between strain
263 Marseille-P2666 and *C. aerofaciens* and *C.intestinalis* or *C. stercoris*, respectively, these
264 values obtained confirm the affiliation of the genus but also supported the status of new
265 species of strain Marseille-P2666 (Table S5).

266 The obtained dDDH and AAI values were lower than the 70% and 95-96% threshold
267 values for species demarcation, respectively [15, 33, 34]. Finally, strain Marseille-P2666
268 exhibited the genomic G+C content differences ranging from -1.3% when compared with *C.*
269 *massiliensis* to +4.3% with *C. tanakaei*. As previously demonstrated, that the G + C content
270 deviation within species does not exceed 1% [35].
271 By taking into consideration its phenotypic (Table 1), phylogenetic (Figure 1) and
272 genomic characteristics (Supplementary data) when compared to *Collinsella* species with
273 standing in nomenclature, strain Marseille-P2666 was considered as belonging to a new
274 species within this genus, for which we propose the name *Collinsella vaginalis* sp. nov.
275

276 **Description of *Collinsella vaginalis* sp. nov.**

277 *Collinsella vaginalis* (va.gi.na'lis. L. n. fem. *vagina*, sheath, vagina; L. fem. gen. suff. –
278 *alis*, suffix denoting pertaining to; N.L. fem. adj. *vaginalis*, pertaining to the vagina).
279 Strictly anaerobic, bacterial cells are rod-shaped, Gram-stain-positive, non-motile, non-
280 sporforming, mesothermophilic, oxidase and catalase negative, with a mean diameter and
281 length of 0.4 µm and 1.8 µm, respectively. Cells occur as single rods or in short chains. After
282 two days of incubation at 37°C under anaerobic conditions, colonies on 5% sheep blood-
283 enriched Columbia agar (BioMérieux), appear grey, opaque and circular with a diameter of
284 0.5-1.2 mm. Nitrate is not reduced; esculin ferric citrate, indole formation, gelatin hydrolysis
285 and urease activities are not detected. Using an API 20A strip (BioMérieux), acid is produced
286 from D-glucose, D-lactose, D-saccharose, D-maltose, salicin, D-cellobiose, D-mannose and
287 D-trehalose but not from D-mannitol, D-xylose, L-arabinose, glycerol, D-melezitose, D-
288 raffinose, sorbitol, D-rhamnose. By using API Rapid ID32A and API ZYM strips
289 (BioMérieux), fermented reactions are observed for mannose and raffinose, N-acetyl-ß-
290 glucosaminidase, alkaline phosphatase, arginine arylamidase, proline arylamidase, leucyl-

291 glycine arylamidase, leucine arylamidase, glycine arylamidase, histidine arylamidase, serine
292 arylamidase, esterase (4), esterase lipase (8), leucine arylamidase, valine arylamidase, cystine
293 arylamidase, acid phosphatase and naphtol-AS-BI-phosphohydrolase. Arginine dihydrolase,
294 α-galactosidase, β-galactosidase, 6-phospho-β-galactosidase, α-glucosidase, β-glucosidase, α-
295 arabinosidase, β-glucuronidase, glutamic acid decarboxylase, α-fucosidase, phenylalanine
296 arylamidase, pyroglutamic acid arylamidase, tyrosine arylamidase, glutamyl glutamic acid
297 arylamidase, lipase (14), trypsin, α-chymotrypsin and α-mannosidase activities were not
298 detected. The most abundant fatty acids are 9-Octadecenoic acid ($C_{18:1\omega9}$) and Hexadecanoic
299 acid ($C_{16:0}$). *C. vaginalis* was susceptible to benzylpenicillin, amoxicillin, metronidazole,
300 rifampicin, and vancomycin and resistant to ceftriaxone and imipenem.

301 The type strain Marseille-P2666^T (= CSUR 2666 = DSM103342) was isolated from the
302 vaginal sample of a French woman suffering from bacterial vaginosis. The genome of the type
303 strain is 2,162,909-bp long and exhibits a G+C content of 64.6 mol%. The 16S rRNA and
304 genome sequences are deposited in EMBL-EBI under accession numbers LT598547 and
305 FWYK00000000, respectively.
306

307 **FUNDING INFEORMATION**

308 This study was funded by the Méditerranée-Infection foundation and the French Agence
309 Nationale de la Recherche under reference Investissements d'Avenir Méditerranée Infection
310 10-IAHU-03.

311 **CONFLICT OF INTEREST**

312 The authors declare no competing interest in relation to this research.
313

314 **ACKNOWLEDGEMENTS**

315 Genome assembly was performed by the Xegen company.

316 **References**

317 1. **Lagier J-C, Armougom F, Million M, Hugon P, Pagnier I, *et al.*** Microbial
318 culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect*
319 2012;18:1185–1193.

320 2. **Lagier J-C, Khelaifia S, Alou MT, Ndongo S, Dione N, *et al.*** Culture of previously
321 uncultured members of the human gut microbiota by culturomics. *Nat Microbiol*
322 2016;1:16203.

323 3. **Stackebrandt E, Rainey FA, Ward-Rainey NL**. Proposal for a new hierarchic
324 classification system, Actinobacteria classis nov. *Int J Syst Evol Microbiol* 1997;47:479–
325 491.

326 4. **Kageyama A, Benno Y, Nakase T**. Phylogenetic and phenotypic evidence for the
327 transfer of *Eubacterium aerofaciens* to the genus *Collinsella* as *Collinsella aerofaciens*
328 gen. nov., comb. nov. *Int J Syst Evol Microbiol* 1999;49:557–565.

329 5. **Kageyama A, Benno Y**. Emendation of genus *Collinsella* and proposal of *Collinsella*
330 *stercoris* sp. nov. and *Collinsella intestinalis* sp. nov. *Int J Syst Evol Microbiol*
331 2000;50:1767–1774.

332 6. **Nagai F, Watanabe Y, Morotomi M**. *Slackia piriformis* sp. nov. and *Collinsella*
333 *tanakaei* sp. nov., new members of the family *Coriobacteriaceae*, isolated from human
334 faeces. *Int J Syst Evol Microbiol* 2010;60:2639–2646.

335 7. **Padmanabhan R, Dubourg G, Lagier J-C, Nguyen T-T, Couderc C, *et al.*** Non-
336 contiguous finished genome sequence and description of *Collinsella massiliensis* sp. nov.
337 *Stand Genomic Sci* 2014;9:1144–1158.

338 8. **Ramasamy D, Mishra AK, Lagier J-C, Padhmanabhan R, Rossi M, *et al.*** A
339 polyphasic strategy incorporating genomic data for the taxonomic description of novel
340 bacterial species. *Int J Syst Evol Microbiol* 2014;64:384–391.

341 9. **Pagani I, Liolios K, Jansson J, Chen I-MA, Smirnova T, *et al.*** The Genomes OnLine
342 Database (GOLD) v.4: status of genomic and metagenomic projects and their associated
343 metadata. *Nucleic Acids Res* 2012;40:D571–D579.

344 10. **Sentausa E, Fournier P-E**. Advantages and limitations of genomics in prokaryotic
345 taxonomy. *Clin Microbiol Infect* 2013;19:790–795.

346 11. **Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P, *et al.*** Ongoing Revolution in
347 Bacteriology: Routine Identification of Bacteria by Matrix?Assisted Laser Desorption
348 Ionization Time?of?Flight Mass Spectrometry. *Clin Infect Dis* 2009;49:543–551.

349 12. **Lagier J-C, Hugon P, Khelaifia S, Fournier P-E, La Scola B, *et al.*** The Rebirth of
350 Culture in Microbiology through the Example of Culturomics To Study Human Gut
351 Microbiota. *Clin Microbiol Rev* 2015;28:237–264.

352 13. **Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral J-P, *et al.*** 16S ribosomal DNA
353 sequence analysis of a large collection of environmental and clinical unidentifiable
354 bacterial isolates. *J Clin Microbiol* 2000;38:3623–3630.

355 14. **Edgar RC**. MUSCLE: multiple sequence alignment with high accuracy and high
356 throughput. *Nucleic Acids Res* 2004;32:1792–1797.

357 15. **Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M**. Genome sequence-based species
358 delimitation with confidence intervals and improved distance functions. *BMC*
359 *Bioinformatics* 2013;14:60.

360 16. **Meier-Kolthoff JP, Göker M, Spröer C, Klenk H-P**. When should a DDH experiment
361 be mandatory in microbial taxonomy? *Arch Microbiol* 2013;195:413–418.

362 17. **Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, *et al.*** Complete
363 genome sequence of DSM 30083 T, the type strain (U5/41 T) of *Escherichia coli*, and a
364 proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci* 2014;9:2.

365 18. **Stamatakis A**. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
366 large phylogenies. *Bioinformatics* 2014;30:1312–1313.

367 19. **Goloboff P, Farris J, C. Nixon K**. *TNT, a free program for phylogenetic analysis*. 2008.

368 20. **Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A**. How
369 Many Bootstrap Replicates Are Necessary? In: Batzoglou S (editor). *Research in*
370 *Computational Molecular Biology*. Berlin, Heidelberg: Springer Berlin Heidelberg. pp.
371 184–200.

372 21. **L. Swofford D**. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods).*
373 *Version 4.0b10*. Sinauer Associates. Sunderland; 2002.

374 22. **Stackebrandt E, Ebers J**. *Taxonomic parameters revisited: Tarnished gold standards*.
375 2006.

376 23. **Konstantinidis KT, Ramette A, Tiedje JM**. The bacterial species definition in the
377 genomic era. *Philos Trans R Soc B Biol Sci* 2006;361:1929–1940.

378 24. **Diop A, Khelaifia S, Armstrong N, Labas N, Fournier P-E, *et al.*** Microbial
379 culturomics unravels the halophilic microbiota repertoire of table salt: description of
380 *Gracilibacillus massiliensis* sp. nov. *Microb Ecol Health Dis*;27. Epub ahead of print 18
381 October 2016. DOI: 10.3402/mehd.v27.32049.

382 25. **Citron DM, Ostovari MI, Karlsson A, Goldstein EJ**. Evaluation of the E test for
383 susceptibility testing of anaerobic bacteria. *J Clin Microbiol* 1991;29:2197–2203.

384 26. **Matuschek E, Brown DFJ, Kahlmeter G**. Development of the EUCAST disk diffusion
385 antimicrobial susceptibility testing method and its implementation in routine
386 microbiology laboratories. *Clin Microbiol Infect* 2014;20:O255–O266.

387 27. **Sasser M**. Identification of bacteria by gas chromatography of cellular fatty acids.
388 http://natasha.eng.usf.edu/gilbert/courses/Biotransport%20Phenomena/pdf/bacteria_gc_1.
389 pdf (1990, accessed 24 March 2016).

390 28. **Dione N, Sankar SA, Lagier J-C, Khelaifia S, Michele C, *et al.*** Genome sequence and
391 description of *Anaerosalibacter massiliensis* sp. nov. *New Microbes New Infect*
392 2016;10:66–76.

393 29. **Wilkins TD, Chalgren S**. Medium for Use in Antibiotic Susceptibility Testing of
394 Anaerobic Bacteria. *Antimicrob Agents Chemother* 1976;10:926–928.

395 30. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al.*** SPAdes: A New
396 Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput*
397 *Biol* 2012;19:455–477.

398 31. **Kim M, Oh H-S, Park S-C, Chun J**. Towards a taxonomic coherence between average
399 nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of
400 prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–351.

401 32. **Kageyama A, Benno Y**. *Collinsella*. In: *Bergey's Manual of Systematics of Archaea and*
402 *Bacteria*. American Cancer Society. pp. 1–6.

403 33. **Konstantinidis KT, Tiedje JM**. Towards a Genome-Based Taxonomy for Prokaryotes. *J*
404 *Bacteriol* 2005;187:6258–6264.

405 34. **Rodriguez-R LM, Konstantinidis KT**. Bypassing cultivation to identify bacterial
406 species. *Microbe* 2014;9:111–8.

407 35. **Klenk H-P, Meier-Kolthoff JP, G?ker M**. Taxonomic use of DNA G+C content and
408 DNA?DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 2014;64:352–
409 356.

410

411 **Table 1:** Compared characteristics of *Collinsella vaginalis* strain Marseille-P2666^T and other members of the genus *Collinsella*: *Collinsella*
412 *tanakaei* strain YIT 12063^T [6]; *C. stercoris* strain DSM 13279^T [5]; *C. intestinalis* strain DSM 13280^T [5]; *C. aerofaciens* strain ATCC
413 25986^T [4]; *C. massiliensis* strain GD3^T [7]. +: positive reaction; -: negative reaction; na: no available data.

Properties	*Collinsella vaginalis*	*Collinsella tanakei*	*Collinsella stercoris*	*Collinsella intestinalis*	*Collinsella aerofaciens*	*Collinsella massiliensis*
Cell diameter (µm)	0.3-0.5	0.5-1.0	0.3-0.5	0.3-0.5	0.3-0.7	0.57
Oxygen requirement	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic
Gram stain	+	+	+	+	+	+
DNA G+C content (mol %)	64.6	60.2	63.2	62.5	60.6	65.8
Spore-forming	-	-	-	-	-	-
Motility	-	-	-	-	-	-
Production of						
Alkaline phosphatase	+	+	+	+	-	+
Acid phosphatase	+	+	+	+	-	+
α-galactosidase	-	-	-	-	+	+
β-galactosidase	-	-	+	-	+	+
α-glucosidase	-	-	-	-	+	+
Esterase lipase	+	-	-	-	-	-
N-acetyl-β-glucosaminidase	+	-	+	+	-	-
Cystine arylamidase	+	-	-	-	-	-
Acid form						
Mannose	+	+	+	+	+	-
Glucose	+	+	+	+	+	-

Salicin	+	+	+	-	+	-
Trehalose	+	+	+	-	-	-
Maltose	+	+	+	-	+	-
Lactose	+	+	+	-	+	-
Rhamnose	-	-	-	-	-	-
L-arabinose	-	-	-	-	-	-
Habitat	Human vagina	Human gut	Human gut	Human gut	Human gut	Human gut

414 **Table 2:** Cellular fatty acid composition (%).

Fatty acids	Name	Mean relative % (a)
18:1ω9	9-Octadecenoic acid	37.5 ± 1.0
16:00	Hexadecanoic acid	23.5 ± 0.5
18:00	Octadecanoic acid	18.5 ± 0.4
18:2ω6	9,12-Octadecadienoic acid	11.3 ± 0.3
14:00	Tetradecanoic acid	3.5 ± 0.3
18:1ω5	13-Octadecenoic acid	2.2 ± 0.3
10:00	Decanoic acid	TR
18:1ω7	11-Octadecenoic acid	TR
20:4ω6	5,8,11,14-Eicosatetraenoic acid	TR
17:00	Heptadecanoic acid	TR
17:0 anteiso	14-methyl-Hexadecanoic acid	TR
15:00	Pentadecanoic acid	TR
12:00	Dodecanoic acid	TR
15:0 anteiso	12-methyl-tetradecanoic acid	TR
17:0 iso	15-methyl-Hexadecanoic acid	TR

415 [a] Mean peak area percentage; TR = trace amounts

416

417 **Figure legends**

418 **Figure 1.** Maximum likelihood phylogenetic tree inferred under the GTR+GAMMA model and

419 rooted by midpoint-rooting.

420 The branches are scaled in terms of the expected numbers of substitutions per site. The

421 numbers above the branches are support values when larger than 60% from ML (left) and MP

422 (right) bootstrapping.

423 **Figure 2.** Phylogenetic tree based on the 16S rRNA gene highlighting the position of

424 *Collinsella vaginalis* strain Marseille-P2666^T relative to other close.

425 GenBank accession numbers of each 16S rRNA are noted in parenthesis. Sequences were

426 aligned using CLUSTALW 2.0 software with default parameters and phylogenetic inferences

427 were obtained using the neighbor-joining method with 500 bootstrap replicates, within MEGA6

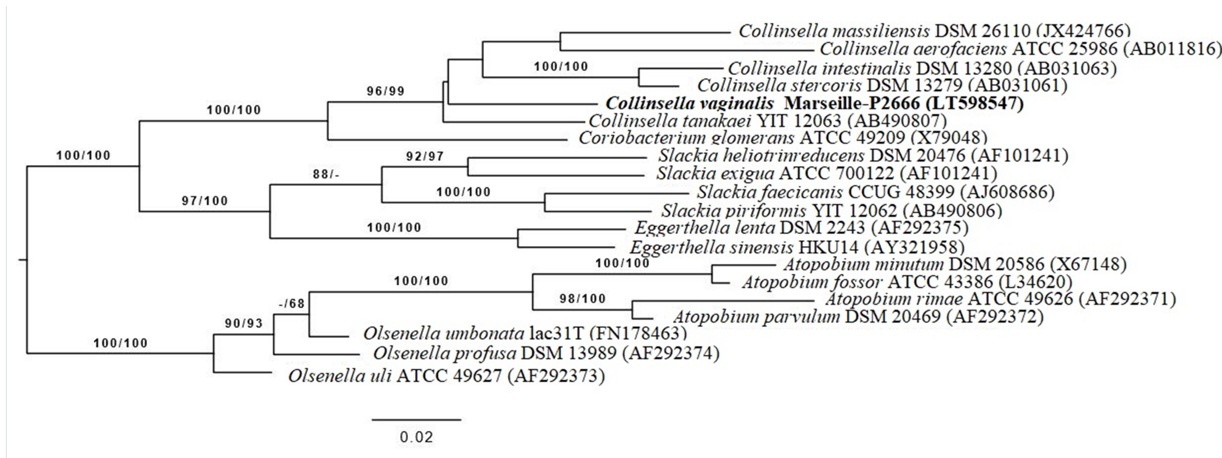
428 software. The evolutionary distances were computed using the Kimura 2-parameter method and

429 are in the units of the number of base substitutions per site. The scale bar represents a 2%

430 nucleotide sequence divergence.

431

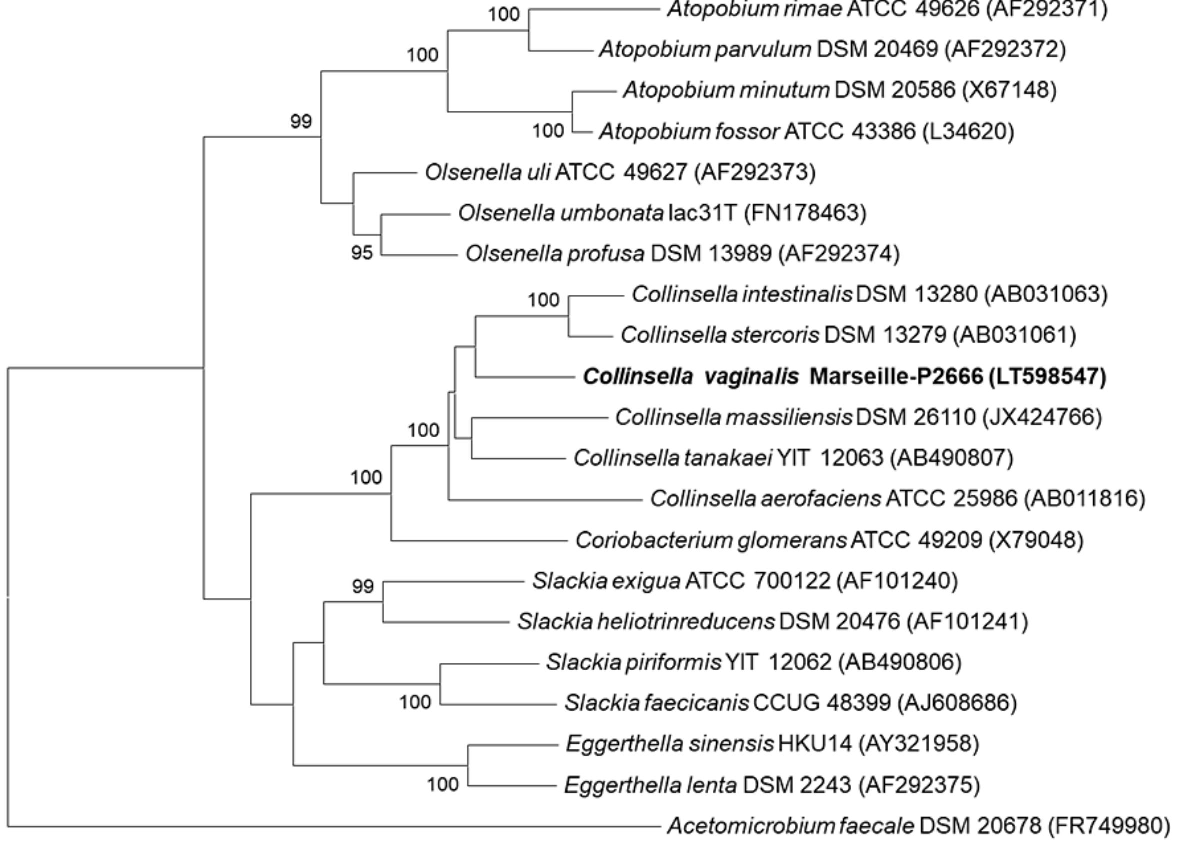
432



433 **Figure 1.** Maximum likelihood phylogenetic tree inferred under the GTR+GAMMA model and

434 rooted by midpoint-rooting.

435



436

437 **Figure 2.** Phylogenetic tree based on the 16S rRNA gene highlighting the position of

438 *Collinsella vaginalis* strain Marseille-P2666^{T} relative to other close.

439

440 **SUPPLEMENTARY DATA**

441 **Supplementary materials and methods**

442 **16S phylogenetic analysis using Neighbor-joining method.**

443 The 16S sequences of the type strains of the closest species to our new strain in the
444 BLAST search were downloaded from the NCBI ftp server (ftp://ftp.ncbi.nih.gov/Genome/).
445 Sequences were aligned using CLUSTALW 2.0 software [1], with default parameters and
446 phylogenetic inferences were obtained using the neighbor-joining method within the MEGA
447 software, version 6 [2]. The evolutionary distances were computed using the Kimura 2-
448 parameter method [3] and the partial deletion option (95%) was used. The bootstraping analysis
449 was performed with 500 replications.

450 **Glucose fermentation products measurements**

451 Wilkins-Chalgren Anaerobe (WCA) culture broth, containing glucose, was freshly
452 prepared [4]. *Collinsella vaginalis* was inoculated and hemoculture flasks were incubated
453 during 8 days at 37 °C. Hydrogen, formic acid and ethanol were measured from 3 independent
454 blank and sample culture flasks. Quantities in samples were given after subtraction the
455 quantities measured in the blank flasks. Calibration curves were calculated from chromatogram
456 peak areas. Coefficients of determination were above 0.999, and back calculated standards all
457 showed good accuracy with deviations below 15 %. Formic acid and 2-ethylbutyric acid were
458 high purity standards; water, ethanol, 2-propanol and methanol were HPLC-grade solvents
459 (Sigma Aldrich, Lyon, France).

460 **Hydrogen**

461 Hydrogen was analyzed using a Clarus 580 gas chromatography system (Perkin Elmer,
462 Villebon-sur-Yvette, France). 100 µL of headspace gas was sampled from flasks with a gastight
463 syringe, then directly injected (1 mm i.d. glass liner; split 10 mL/min) into a Shincarbon ST
464 80/100 micropacked column (2 m x 0.53 mm; Restek, Lisses, France). Injector and oven were

465 maintained at 110 and 70 °C respectively. Argon was set at 57 psi as the carrier gas. Hydrogen
466 was detected by a Thermal Conductivity Detector set at 100 °C and 40 mA in the negative
467 polarity. Data recording and processing was performed using Totalchrom 6.3.2 (Perkin Elmer).
468 A linear calibration curve was calculated by injecting known volumes of a standard mixture
469 containing 25 % (volume) of hydrogen, nitrogen, methane and carbon dioxide (Air Products,
470 Aubervilliers, France). Hydrogen amounts were expressed as number of hydrogen molecules
471 for every million molecules of air (parts per million - ppm).

472 **Formic acid**

473 Formic acid was measured with a Clarus 500 chromatography system connected to a
474 SQ8s mass spectrometer (Perkin Elmer). Formic acid calibration standards were prepared in
475 acidified water (pH 2-3 using HCl 37 %): 0,5 ; 1 ; 2,5 ; 5 ; 10 mmol/L. 1 mL of culture medium
476 was collected then centrifuged 5 minutes at 16000 x g to remove bacteria and debris, and the
477 clear supernatant was adjusted to pH 2-3. Standards and samples were then spiked with 2-
478 ethylbutyric acid as the internal standard (IS; 1 mmol/L). The samples were once again
479 centrifuged before injection. 0.5 µL of standards and samples was directly injected into a
480 splitless liner (1 mm i.d.) heated at 200 °C. Injection carry-over was decreased with 10 syringe
481 washes in methanol:water (50:50 v/v). Compounds were separated through an Elite-FFAP
482 column (30 m, 0.25 mm i.d., 0.25 mm film thickness) using a linear temperature gradient from
483 100 to 200 °C at 8 °C/min. Helium flowing at 1 mL/min was used as the carrier gas. MS inlet
484 line and Electron Ionization source were set at 200 °C. Selected Ion Recording (SIR) was
485 performed after a 4.5 min solvent delay with the following base peak masses: 29 m/z (formic
486 acid) and 88 m/z (2-ethylbutyric acid, IS). All data was collected and processed using
487 Turbomass 6.1 (Perkin Elmer, Courtaboeuf, France). A quadratic internal calibration was
488 calculated from the SIR chromatograms.

489 **Ethanol**

490 Ethanol was quantified using a Turbomatrix 40 Headspace sampler connected to a
491 Clarus 500 chromatography system equipped with a SQ8s mass spectrometer (Perkin Elmer).
492 Ethanol calibration standards were prepared in water: 0,5 ; 1 ; 2,5 ; 5 ; 10 mmol/L. 1 mL of
493 culture medium was collected into a 20 mL headspace glass vial. Standards and samples were
494 spiked with 2-propanol as the internal standard (IS; 5 mmol/L). Each vial was processed as
495 follow in the headspace sampler: 10 minutes heating at 60 °C, 1 minute pressurization at 20 psi,
496 0.03 minutes injection with needle and transfer line set at 70 and 80 °C. Volatile compounds
497 were heated at 150 °C in a glass liner (1 mm i.d.; split 10:1) before eluting through an Elite-
498 BAC2 column (30 m, 0.32 mm i.d., 1.2 mm film thickness) maintained at 40 °C. Helium was
499 flowing at 15 psi as the carrier gas. MS inlet line and Electron Ionization source were set at 150
500 °C. Selected Ion Recording (SIR) was performed after a 3 min solvent delay with the following
501 base peak masses: 31 m/z (ethanol) and 45 m/z (2-propanol, IS). All data was collected and
502 processed with Turbomass 6.1 (Perkin Elmer). Linear internal calibration was calculated from
503 the SIR chromatograms.

504 **DNA Extraction and genome sequencing**

505 After a pretreatment step by lysozyme incubation at 37°C for 2 hours, the Genomic DNA
506 (gDNA) of strain Marseille-P2666^T was extracted on the EZ1 biorobot (Qiagen, Hilden,
507 Germany) using the EZ1 DNA tissues kit. The elution volume was 50µL. gDNA was
508 quantified by a Qubit assay with the high sensitivity kit (Life technologies, Carlsbad, CA,
509 USA) to 68.1 ng/µl.

510 The gDNA was sequenced on the MiSeq sequencer (Illumina Inc, San Diego, CA, USA)
511 with the mate pair strategy. The gDNA was barcoded in order to be mixed with 11 other
512 projects using the Nextera Mate Pair sample prep kit (Illumina). The mate pair library was
513 prepared with 1.5 µg of gDNA using the Nextera mate pair Illumina guide. The genomic DNA
514 sample was simultaneously fragmented and tagged with a mate pair junction adapter. The

515 pattern of the fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent
516 Technologies Inc, Santa Clara, CA, USA) with a DNA 7500 labchip. The DNA fragments
517 ranged in size from 1.5 kb up to 11 kb with an optimal size at 9.088 kb. No size selection was
518 performed and 600ng of tagmented fragments were circularized. The circularized DNA was
519 mechanically sheared to small fragments with an optimal at 1325 bp on the Covaris device S2
520 in microtubes (Covaris, Woburn, MA, USA).The library profile was visualized on a High
521 Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc, Santa Clara, CA, USA) and the
522 final concentration library was measured at 11.99 nmol/l. The libraries were normalized at 2nM
523 and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded.
524 Automated cluster generation and sequencing run were performed in a single 39-hours run in a
525 2x151-bp.

526 A total of sequencing output of 5.1 Gb was obtained from a 542K/mm^2 cluster density
527 with a cluster passing quality control filters of 95.7% (10,171,000 clusters). Within this run, the
528 index representation for *Collinsella vaginalis* was determined to 7.88%. The 801,260 paired
529 end reads were trimmed then assembled.

530 **Genome annotation and comparison**

531 Prodigal was used for Open Reading Frame (ORF) prediction [5] with default
532 parameters. . Predicted ORFs spanning a sequencing gap region were excluded. Bacterial
533 proteome was predicted using BLASTP (E-value 1e^{-03} coverage 0.7 and identity percent 30%)
534 against the Clusters of Orthologous Groups (COG) database. If no hit was found, a search
535 against the nr database [6] was performed using BLASTP with E-value of 1e^{-03}, a coverage of
536 0.7 and an identity percent of 30 %. If sequence lengths were smaller than 80 amino acids, we
537 used an E-value of 1e^{-05}. Pfam conserved domains (PFAM-A an PFAM-B domains) were
538 searched on each protein with the hhmscan tools analysis [7]. RNAmmer [8] and tRNAScanSE
539 [9] were used to identify ribosomal RNAs and tRNAs, respectively. We predicted lipoprotein

540 signal peptides and the number of transmembrane helices using Phobius [10]. ORFans were
541 identified if the BLASTP search was negative (E-value smaller than $1e^{-03}$ for ORFs with a
542 sequence size larger than 80 aas or E-value smaller than $1e^{-05}$ for ORFs with sequence length
543 smaller than 80 aas). Artemis [11] and DNA Plotter [12] were used for data management and
544 for visualization of genomic features, respectively. Annotation and comparison processes were
545 performed using the multi-agent software system DAGOBAH [13], which include Figenix [14]
546 libraries that provide pipeline analysis. Genomes from members of the *Coriobacteriaceae*
547 family and closely related to our strain were used for the comparative genomics study.
548 Genomic informations from strain Marseille-P2666 and comparatively closest related species
549 are presented in Table 6. Finally, the average amino acid identity (AAI) was calculated, based
550 on the overall similarity between datasets of proteins of genome pairs belonging to the same
551 genus of *Collinsella* [15] available at (http://enve-omics.ce.gatech.edu/aai/index). We also
552 performed GGDC analysis using the GGDC web server, as previously reported [16].

553 **SUPPLEMENTARY TABLES**

554 **Table S1.** Nucleotide content and gene count levels of the genome of strain Marseille-P2666^T

Attribute	Value	% of total[a]
Size (bp)	2,162,909	100
G+C content (bp)	1,383,290	64.6
Coding region (bp)	1,624,759	75.1
Total genes	1,774	100
RNA genes	50	2.8
Protein-coding genes	1,724	100
Genes with function prediction	1,303	75.6
Genes assigned to COGs	1,191	69.1
Genes with peptide signals	141	8.2
Genes with transmembrane helices	389	22.6

555 a The total is based on either the size of the genome in base pairs or the total number of protein

556 coding genes in the annotated genome.

557 **Table S2:** Number of genes associated with the 25 general COG functional categories of strain

558 Marseille-P2666^T

Code	Value	% of total	Description
[J]	137	8.0	Translation
[A]	0	0	RNA processing and modification
[K]	98	5.7	Transcription
[L]	49	2.8	Replication, recombination and repair
[B]	1	0.1	Chromatin structure and dynamics
[D]	15	0.9	Cell cycle control, mitosis and meiosis
[Y]	0	0	Nuclear structure
[V]	40	2.3	Defense mechanisms
[T]	51	3.0	Signal transduction mechanisms
[M]	65	3.8	Cell wall/membrane biogenesis
[N]	5	0.3	Cell motility
[Z]	0	0	Cytoskeleton
[W]	4	0.2	Extracellular structures
[U]	19	1.1	Intracellular trafficking and secretion
[O]	50	2.9	Post-translational modification, protein turnover, chaperones
[X]	6	0.3	Mobilome: prophages, transposons
[C]	77	4.5	Energy production and conversion
[G]	182	10.6	Carbohydrate transport and metabolism
[E]	115	6.7	Amino acid transport and metabolism
[F]	52	3.0	Nucleotide transport and metabolism
[H]	63	3.7	Coenzyme transport and metabolism
[I]	33	1.9	Lipid transport and metabolism
[P]	68	3.9	Inorganic ion transport and metabolism
[Q]	15	0.9	Secondary metabolites biosynthesis, transport and catabolism
[R]	104	6.0	General function prediction only
[S]	70	4.1	Function unknown
_	533	30.9	Not in COGs

559

560 **Table S3:** Genome comparison of closely related species to *Collinsella vaginalis* strain

561 Marseille P2666^T

Species	INSDC identifier[a]	Size (Mb)	G+C (mol %)	Gene Content
Collinsella vaginalis strain Marseille-P2666^T	FWYK00000000.1	2.2	64.6	1,907
Collinsella intestinalis DSM 13280	ABXH00000000.2	1.8	62.5	1,630
Collinsella aerofaciens ATCC 25986	AAVN00000000.2	2.4	60.5	2,437
Collinsella stercoris DSM 13279	ABXJ00000000.1	2.5	63.2	2,119
Collinsella tanakei YIT 12063	ADLS00000000.1	2.5	60.2	2,253
Coriobacterium glomerans ATCC 49209	CP002628.1	2.1	60.4	1,856
Olsenella profusa DSM 13989	AWEZ00000000.1	2.7	64.2	2,707
Olsenella uli ATCC 49627	CP002106.1	2.1	64.7	1,812

562 [a] INSDC: International Nucleotide Sequence Database Collaboration.

563 **Table S4:** dDDH values (%) obtained by comparison of all studied genomes

	CT	CS	CI	CA	CG	OP	OU
CV	22.6 ± 2.4	23.2 ± 2.4	23.0 ± 2.4	22.4 ± 2.4	20.4 ± 2.3	19.1 ± 2.8	19.7 ± 2.4
CT		25.0 ± 2.4	24.7 ± 2.4	22.5 ± 2.4	21.6 ± 2.4	20.0 ± 2.4	19.5 ± 2.3
CS			28.2 ± 2.5	23.9 ± 2.4	21.3 ± 2.3	19.1 ± 2.3	20.3 ± 2.3
CI				23.6 ± 2.4	21.2 ± 2.4	19.5 ± 2.3	20.4 ± 2.3
CA					21.0 ± 2.3	19.6 ± 2.3	20.0 ± 2.3
CG						20.0 ± 2.3	20.0 ± 2.3
OP							22.3 ± 2.4

564 dDDH: Digital DNA-DNA hybridization. CV: *Collinsella vaginalis* Marseille-P2666^T; CT :
565 *Collinsella tanakaei* YIT 12063^T*;* CS : *Collinsella stercoris* DSM 13279^T; CI : *Collinsella*
566 *intestinalis* DSM 13280^T; CA : *Collinsella aerofaciens* ATCC 25986^T; CG : *Coriobacterium*
567 *glomerans* ATCC 49209^T*;* OP : *Olsenella profusa* DSM 13989^T; OU : *Olsenella uli* ATCC
568 49627^T
569

570 **Table S5:** Average amino acid identity (AAI) values (%) between *Collinsella vaginalis* strain

571 Marseille P2666^T and other closely related *Collinsella* species.

	CT	CS	CI	CA
CV	66.9	66.1	66.9	64.7
CT		68.9	69.7	65.5
CS			79.5	66.3
CI				66.4

572 CV : *Collinsella vaginalis* Marseille-P2666^T; CT : *Collinsella tanakaei* YIT 12063^T*;* CS :

573 *Collinsella stercoris* DSM 13279^T; CI : *Collinsella intestinalis* DSM 13280^T; CA : *Collinsella*

574 *aerofaciens* ATCC 25986^T.

SUPPLEMENTARY FIGURE LEGENDS

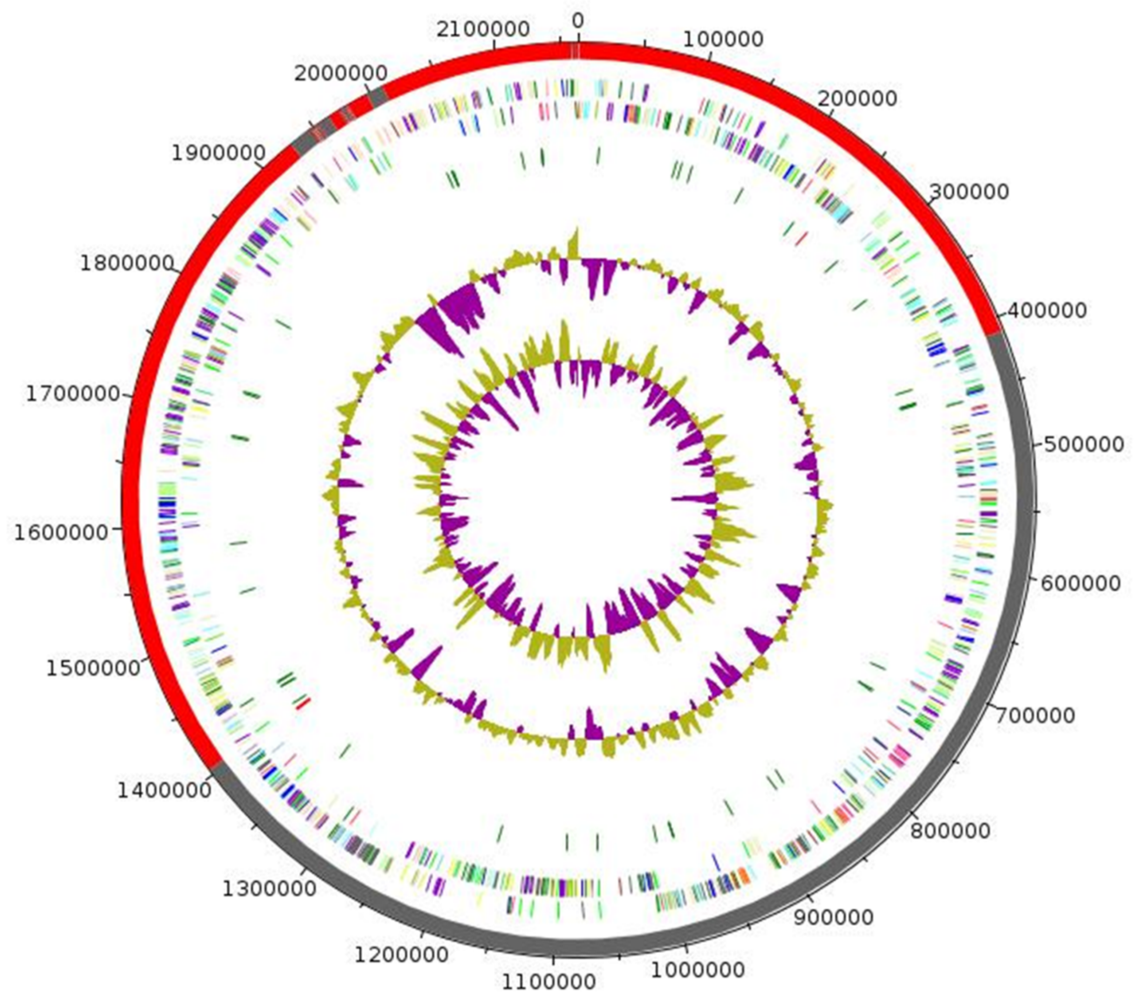
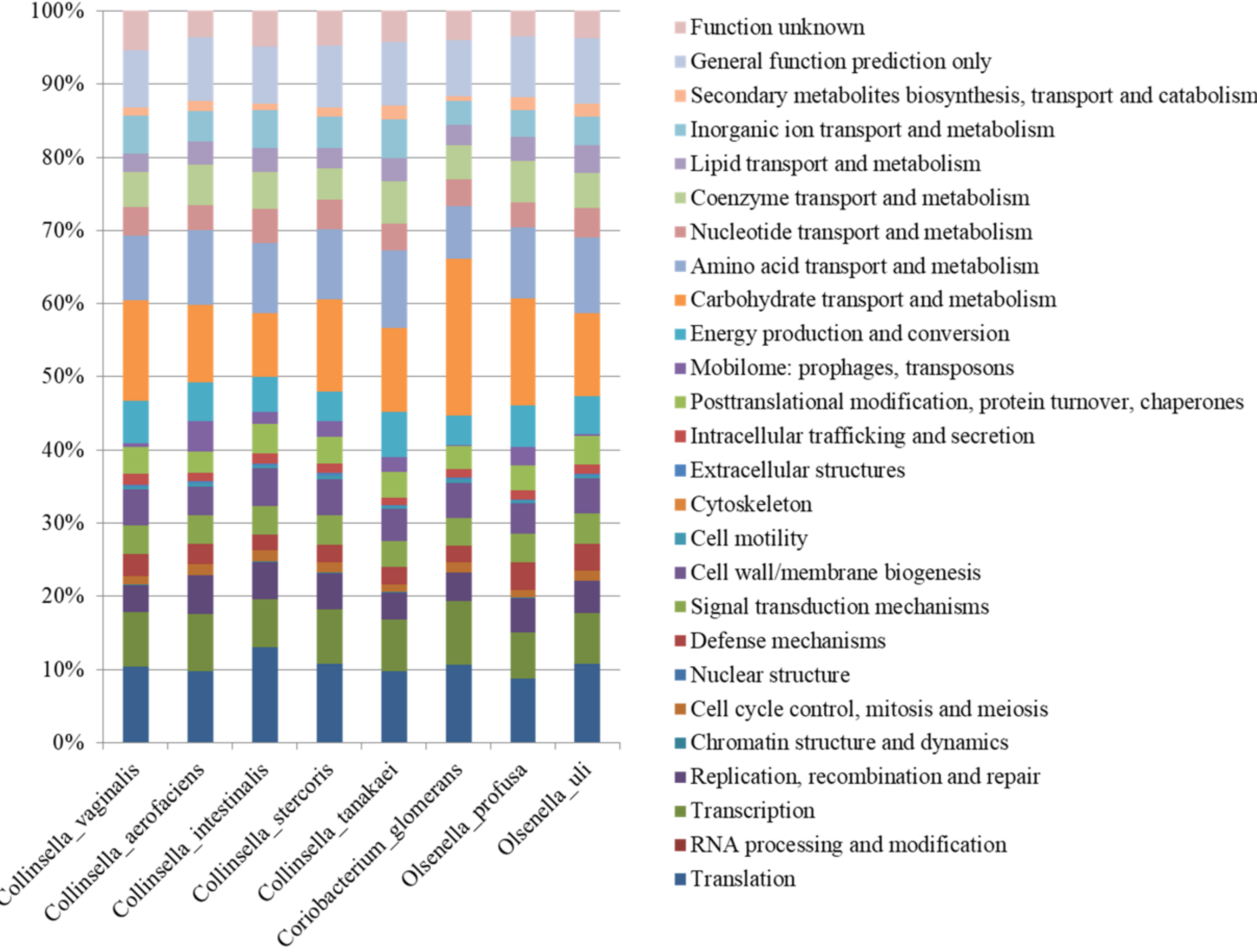


Figure S1. Graphical circular map of the genome. From the outside in: contigs (red/gray), COG category of genes on the forward strand (three circles), genes on the forward strand (blue circle), genes on the reverse strand (red circle), COG category on the reverse strand (three circles), G+C content.

581 **Figure S3.** Distribution of functional classes of predicted genes according to the clusters of
582 orthologous groups of proteins of *Collinsella vaginalis* strain Marseille-P2666^T among other
583 species.



584
585

586 **References**

587 1. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, *et al.*** Clustal W
588 and Clustal X version 2.0. *Bioinformatics* 2007;23:2947–2948.

589 2. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S**. MEGA6: Molecular
590 Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* 2013;30:2725–2729.

591 3. **Kimura M**. A simple method for estimating evolutionary rates of base substitutions
592 through comparative studies of nucleotide sequences. *J Mol Evol* 1980;16:111–120.

593 4. **Wilkins TD, Chalgren S**. Medium for Use in Antibiotic Susceptibility Testing of
594 Anaerobic Bacteria. *Antimicrob Agents Chemother* 1976;10:926–928.

595 5. **Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, *et al.*** Prodigal: prokaryotic
596 gene recognition and translation initiation site identification. *BMC Bioinformatics*
597 2010;11:1.

598 6. **Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, *et al.*** GenBank. *Nucleic*
599 *Acids Res* 2015;43:D30–D35.

600 7. **Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, *et al.*** HMMER web server:
601 2015 update. *Nucleic Acids Res* 2015;43:W30–W38.

602 8. **Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, *et al.*** RNAmmer:
603 consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007;35:3100–
604 3108.

605 9. **Lowe TM, Eddy SR**. tRNAscan-SE: a program for improved detection of transfer RNA
606 genes in genomic sequence. *Nucleic Acids Res* 1997;25:955–964.

607 10. **Käll L, Krogh A, Sonnhammer EL.** A Combined Transmembrane Topology and Signal
608 Peptide Prediction Method. *J Mol Biol* 2004;338:1027–1036.

609 11. **Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA**. Artemis: an integrated
610 platform for visualization and analysis of high-throughput sequence-based experimental
611 data. *Bioinformatics* 2012;28:464–469.

612 12. **Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J**. DNAPlotter: circular and
613 linear interactive genome visualization. *Bioinformatics* 2009;25:119–120.

614 13. **Gouret P, Paganini J, Dainat J, Louati D, Darbo E, *et al.*** *Integration of Evolutionary*
615 *Biology Concepts for Functional Annotation and Automation of Complex Research in*
616 *Evolution: The Multi-Agent Software System DAGOBAH*. Evolutionary biology – concepts,
617 biodiversity, macroevolution and genome evolution. Springer Berlin Heidelberg, Berlin,
618 Heidelberg: Pontarotti, Pierre; 2011.

619 14. **Gouret P, Vitiello V, Balandraud N, Gilles A, Pontarotti P, *et al.*** FIGENIX: intelligent
620 automation of genomic annotation: expertise integration in a new software platform. *BMC*
621 *Bioinformatics* 2005;6:198.

622 15. **Rodriguez-R LM, Konstantinidis KT**. Bypassing cultivation to identify bacterial species.
623 *Microbe* 2014;9:111–8.

624 16. **Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M**. Genome sequence-based species
625 delimitation with confidence intervals and improved distance functions. *BMC*
626 *Bioinformatics* 2013;14:1.

627

628

Article 6:

Corynebacterium fournierii **sp. nov., isolated from the female genital tract of a patient with bacterial vaginosis.**

Khoudia Diop, Thi Tien Nguyen, Jérémy Delerce, Nicholas Armstrong, Didier Raoult, Florence Bretelle, Florence Fenollar

Published in Antonie van Leeuwenhoek

Antonie van Leeuwenhoek
https://doi.org/10.1007/s10482-018-1022-z

ORIGINAL PAPER

Corynebacterium fournierii sp. nov., isolated from the female genital tract of a patient with bacterial vaginosis

Khoudia Diop · Thi Tien Nguyen · Jérémy Delerce · Nicholas Armstrong · Didier Raoult · Florence Bretelle · Florence Fenollar

Received: 20 October 2017 / Accepted: 23 January 2018
© Springer International Publishing AG, part of Springer Nature 2018

Abstract Strain Marseille-P2948^T, a novel Gram-positive, catalase-positive bacterium was isolated from a vaginal sample of a patient with bacterial vaginosis. It was characterised using the taxonogenomic approach. Phylogenetic analysis revealed that the 16S rRNA and the *rpoB* genes exhibit 98.7 and 93.4% similarity, respectively, with those of *Corynebacterium ureicelerivorans* strain IMMIB RIV-301^T. Biochemical tests of strain Marseille-P2948^T gave results that were similar to those of other validly named *Corynebacterium* species, whereas chemotaxonomic tests showed the presence of $C_{16:0}$, $C_{18:1n9}$, $C_{18:0}$, and $C_{18:2n6}$ in the fatty acid profile. The draft genome of strain Marseille-P2948^T is 2,383,644 bp long in size with a G+C content of 65.03%. Of the 2210 predicted genes, 2147 are protein-coding genes and 63 are RNAs. Based on phenotypic, phylogenic and genomic results, it was concluded that the isolate represents a new species within the genus *Corynebacterium.* The name *Corynebacterium fournierii* sp. nov. is proposed and the type strain is Marseille-P2948^T (= CSUR P2948 = DSM 103271).

Keywords *Corynebacterium fournierii* · Bacterial vaginosis · Culturomics · Taxogenomics · Genome

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10482-018-1022-z) contains supplementary material, which is available to authorized users.

K. Diop · F. Fenollar (✉)
Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13005 Marseille, France
e-mail: florence.fenollar@univ-amu.fr

T. T. Nguyen · J. Delerce · N. Armstrong · D. Raoult · F. Bretelle
Aix-Marseille Univ, IRD, AP-HM, MEPHI, IHU-Méditerranée Infection, Marseille, France

D. Raoult
Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

F. Bretelle
Department of Gynecology and Obstetrics, Gynépole, Hôpital Nord, AP-HM, Marseille, France

Abbreviations

CSUR	Collection de souches de l'Unité des Rickettsies
DSM	Deutsche Sammlung von Mikroorganismen
MALDI-TOF	Matrix-assisted laser-desorption/ ionization time-of-flight
TE buffer	Tris-EDTA buffer
URMITE	Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes

Published online: 30 January 2018

Introduction

The human vagina is a complex ecosystem colonised by diverse types of microorganisms, including archaea, protists, fungi, viruses and, mostly, bacteria (Belay et al. 1990; Fichorova et al. 2017; Pal et al. 2011). Colonisation of the vagina starts at birth and continues throughout life (Romero et al. 2014). The composition of the vaginal microbiota varies from day-to-day depending upon intrinsic and extrinsic factors (Zapata and Quagliarello 2015; Herbst-Kralovetz et al. 2016; Fettweis et al. 2014; Mendling 2016). However, symbiotic associations between the female genital tract and the vaginal flora maintains homeostasis and plays a crucial role in women's health and that of their neonates. An imbalance in the vagina microbiota can lead to dysbiosis, such as bacterial vaginosis (BV) (Narayankhedkar et al. 2015). During our investigations into the vaginal flora with a view to understanding the mechanism of bacterial vaginosis through "microbial culturomics", an approach based on using high-throughput culture conditions and matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) for bacterial identification (Lagier et al. 2012; Dubourg et al. 2013; Lagier et al. 2016), a new member of genus *Corynebacterium* was isolated, designated strain Marseille-P2948^{T}.

The genus *Corynebacterium* was created in 1896 by Lehmann and Neumann. The type species is *Corynebacterium diphtheriae*. Currently, this genus contains of 110 valid species (http://www.bacterio.net/corynebacterium.html). Bacteria within the genus *Corynebacterium* are Gram-positive, facultatively anaerobic and catalase-positive. Species have been isolated from various animal sources (Braun et al. 2016; Hoyles et al. 2013), vegetables (Fudou et al. 2002), soil (Chen et al. 2004; Negi et al. 2016) and human clinical specimens, including from the urogenital tract (Funke et al. 1997; Shukla et al. 2003; Devriese et al. 2000).

Strain Marseille-P2948^{T} was characterised and described using taxonogenomics (Fournier et al. 2015), a new polyphasic approach that combines the classical methods for bacterial description (phenotypic and genotypic characteristics, including DNA–DNA hybridization; %G+C) and peptide mass fingerprints obtained using MALDI-TOF mass spectrometry. In this paper, we describe strain Marseille-P2948^{T} (= CSUR P2948 = DSM 103271) as representative of a new bacterial species, and its draft shotgun annotated genome sequence.

Materials and methods

Ethics and sample collection

A vaginal specimen was taken from a 30-year-old woman diagnosed with bacterial vaginosis, as previously reported (Menard et al. 2008), using a Sigma Transwab (Medical Wire, Corsham, United Kingdom). The study was authorised by the local Institut Fédératif de Recherche 48 ethics committee (Faculty of Medicine, Marseille, France) under agreement number 09-022. The patient also gave her written consent. She was not receiving any antibiotic treatment at the time of the sample collection.

Bacterial strain isolation and identification

After sampling, the vaginal specimen was cultured using 1 of the 18 culturomic conditions previously described (Lagier et al. 2015). The sample was pre-incubated in an aerobic culture bottle (BD Diagnostics, Le Pont-de-Claix, France) supplemented with 4 ml sheep rumen fluid filtered at 0.2 µm using a pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France) and 3 ml sheep blood (bioMérieux, Marcy l'Etoile, France) for 1, 5, 10, 14, 21, 26, and 30 days before inoculating the broth onto Chocolat PolyViteX (PVX) agar (BD Diagnostics). After 2 days' incubation at 37 °C in aerobic conditions, purified colonies were deposited in duplicate on a MSP 96 MALDI-TOF target plate (Bruker Daltonics, Leipzig, Germany) and, as previously described, 1.5 µl of matrix solution was added to each spot. Identification was carried out using a microflex spectrometer (Bruker) (Seng et al. 2009) which compares the protein spectra found to those in the Bruker database (constantly updated with spectra of new species discovered in our laboratory). Isolates with an unambiguous score $\geq$ 2.0 were considered correctly identified at the genus and species level. In contrast, if no spectra matched with the database, and for unidentified bacteria with a clear spectrum, 16S rRNA and *rpoB* gene sequencing was performed using the universal primer pair fD1 and rp2 (Eurogentec, Angers, France) and the primer pair cory2700f and cory3130r (Khamis et al. 2004),

respectively. The obtained sequences were corrected using Chromas Pro 1.34 software (Technelysium Pty. Ltd., Tewantin, Australia) and were matched against the NCBI database using the BLAST algorithm (Drancourt et al. 2000). The 16S rRNA gene similarity level thresholds which we use to define a new species and new genus without performing DNA–DNA hybridization were < 98.7 and $< 95\%$, respectively (Stackebrandt and Ebers 2006; Kim et al. 2014; Yarza et al. 2014). This strategy lead to the designation of an apparently novel isolate as strain Marseille-P2948^T.

Phylogenetic analysis

To construct phylogenetic trees showing the position of the novel isolate relative to other species, sequences were recovered using a nucleotide BLAST against the 16S RNA Database of "The All-Species Living Tree" Project of Silva (LTPs119). First, a filter to eliminate sequences under 1450 bp in size was applied. According to the BLASTn similarity percent results (https://blast.ncbi.nlm.nih.gov/Blast.cgi), we retrieved the 16S rRNA gene sequences of the type strains of closely related species with validly published names in the List of Prokaryotic Names with Standing in Nomenclature website (http://www.bacterio.net/index.html). Then, sequences were aligned using Muscle (Edgar 2004) and phylogenetic inferences obtained using the approximate maximum-likelihood method within the FastTree software (Price et al. 2009). Local support values for the tree nodes were computed using the Shimodaira-Hasegawa test. A filter using PhyloPattern (Gouret et al. 2009) was applied to the tree to remove duplicate species in the tree or inappropriate taxonomic reference species.

Phenotypic characteristics

The growth characteristics of strain Marseille-P2948^T were tested by culturing the bacterium on Columbia agar with 5% sheep blood (bioMérieux) and incubating it at different temperatures (ambient, 28, 37, 42, and 56 °C) in dissimilar atmospheres (aerobic, anaerobic, and microaerophilic), with different pH conditions (5, 6, 6.5, 7, and 8.5) and different concentrations of NaCl (0, 5, 15, and 45%) (Mishra et al. 2013). Gram-stain, motility, sporulation, catalase and oxidase tests were performed as previously reported (Murray et al. 2007).

To view the bacterial morphology, cells were fixed with 2.5% glutaraldehyde in a 0.1 M cacodylate buffer for at least one hour at + 4 °C. One drop of cell suspension was deposited for approximately five minutes on glow-discharged formvar carbon film on 400 mesh nickel grids (FCF400-Ni, EMS). The grids were dried on blotting paper and the cells were negatively stained for 10 s with 1% ammonium molybdate solution in filtered water at RT. Electron micrographs were acquired using a Tecnai G20 Cryo (FEI) transmission electron microscope operated at 200 keV.

Biochemical tests were performed using API ZYM, API Coryne, and API 50 CH strips (bioMérieux) according to the manufacturer's instructions. Amoxicillin, benzylpenicillin, ceftriaxone, ertapenem, imipenem, erythromycin, metronidazole, ofloxacin, rifampicin, amikacin and vancomycin were used to test antibiotic susceptibility. The minimal inhibitory concentrations (MICs) were then determined according to EUCAST recommendations using E-test gradient strips (bioMérieux) (Matuschek et al. 2014).

Cellular fatty acid methyl ester (FAME) analysis was performed by Gas Chromatography/Mass Spectrometry (GC/MS). Two tubes were prepared with approximately 15 mg of bacterial biomass per tube, harvested from several culture plates. FAME were prepared for analysis as described by Sasser (Sasser 2006). GC/MS analyses were carried out as previously described (Dione et al. 2016). Briefly, FAME were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500 - SQ 8 S, Perkin Elmer, Courtaboeuf, France). A spectral database search was performed using MS Search 2.0 operated with the Standard Reference Database 1A (NIST, Gaithersburg, USA) and the FAME mass spectral database (Wiley, Chichester, UK).

Genome sequencing and assembly

Genomic DNA (gDNA) of strain Marseille-P2948^T was sequenced using the MiSeq Technology (Illumina Inc, San Diego, CA, USA) and the mate pair strategy. The gDNA was barcoded to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). gDNA was quantified by a Qubit assay with a high sensitivity kit (Life technologies, Carlsbad, CA, USA) to 105.7 ng/μl. The mate pair library was prepared with 1.5 μg of gDNA using the Nextera

mate pair Illumina guide. The gDNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of fragmentation was validated 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 to 11 kb with an optimal size at 5.203 kb. No size selection was performed and 440 ng of tagmented fragments were circularised. The circularised DNA was mechanically sheared to small fragments with an optimal size of 985 bp on a Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualised using a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc, Santa Clara, CA, USA) and the final concentration of the library was determined to be 4.17 nmol/l. The libraries were normalised at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing were performed in a single 39-h run in a 2 × 151-bp. A total of 8.8 Gb of information was obtained from a 971 K/mm^2 cluster density with a cluster passing quality control filters of 93.1% (17,376,000 passing filter paired reads). Within this run, the index representation for strain Marseille-P2948^T was determined to 7.17%. The 1,246,384 paired reads were trimmed and assembled into 24 scaffolds.

Genome annotation and analysis

Annotation and comparison processes were performed using the Multi-Agent software system DAGOBAH (Gouret et al. 2011) that includes Figenix (Gouret et al. 2005). Open Reading Frames (ORFs) were predicted using Prodigal software (Hyatt et al. 2010) with default parameters. Predicted ORFs spanning a sequencing gap region (containing N) were excluded. The protein functions were predicted using BLASTP (cut offs: E-value 1 × 10^{-03}, coverage 0.7 and 30% identity) against the Clusters of Orthologous Groups (COG) database. A search against the NR database (Clark et al. 2016) was performed if no hit was found, using BLASTP with an E-value of 1 × 10^{-03}, coverage 0.7 and 30% identity as thresholds. An E-value of 1 × 10^{-05} was used with sequence lengths smaller than 80 amino acids. The hhmscan tool analyses were used to search for Pfam conserved domains (PFAM-A and PFAM-B domains) in each protein. We used RNAmmer (Lagesen et al. 2007) and tRNAScanSE tools (Lowe and Eddy 1997) to detect ribosomal RNAs genes and tRNA genes, respectively. Visualisation and data management of genomic features was performed, using Artemis (Carver et al. 2012) and DNA Plotter (Carver et al. 2009), respectively. For the mean level of nucleotide sequence similarity analysis at the genome level, we used the MAGI home-made software. This software enables calculation of the average genomic identity of orthologous gene sequences (AGIOS) values by comparing in a pairwise manner the sequences of the strain of interest to those of members of the genus to which the strain is most closely related (Ramasamy et al. 2014). Briefly, the MAGI software, combined with the Proteinortho software (Lechner et al. 2011), detects orthologous proteins between genomes compared two by two then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFans using the Needleman–Wunsch global alignment algorithm. Digital DNA–DNA hybridization (dDDH) analysis was performed using the GGDC web server, as previously reported (Meier-Kolthoff et al. 2013).

Results and discussion

Strain identification and phylogenetic analysis

A vaginal sample was enriched in a blood culture bottle with sheep blood and rumen fluid under aerobic conditions. At different time points (24 h to 30 days), the inoculum was plated on PVX agar and incubated for 2 days at 37 °C under aerobic conditions. The first isolation of strain Marseille-P2948^T occurred after 30 days of pre-incubation. The bacterium was also found to grow on sheep blood agar. As its MALDI-TOF spectrum did not match those in our database, identification failed, suggesting that the strain is not a previously cultured bacterium or at least that the bacterium is not present in our database. The 16S RNA and *rpoB* sequences of strain Marseille-P2948^T (GenBank accession No. LT576414 and LT965931, respectively) showed 98.7 and 93.4% similarity respectively with those of *Corynebacterium ureicelerivorans* strain IMMIB RIV-301^T (GenBank

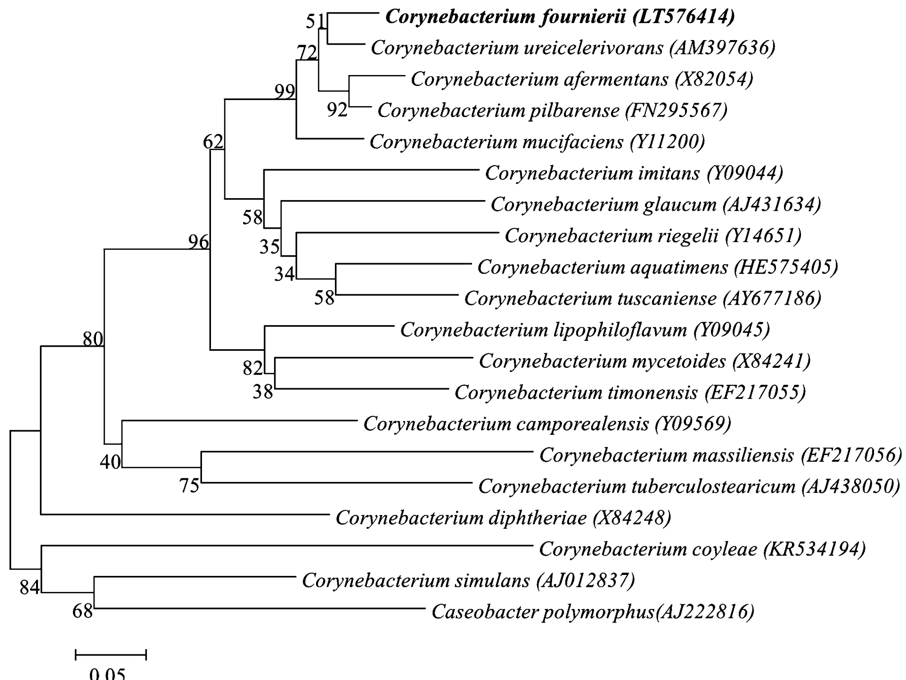


Fig. 1 Phylogenetic tree highlighting the position of *Corynebacterium fournierii* strain Marseille-P2948^T relative to other closely related strains. GenBank accession numbers for each 16S rRNA are noted in parenthesis. Sequences were aligned using Muscle v3.8.31 with default parameters and phylogenetic inferences were obtained using the neighbor-joining method with 500 bootstrap replicates, using MEGA6 software. The scale bar represents a 5% nucleotide sequence divergence

accession No. NR_042558.1), which is currently the phylogenetically closest bacterium with a validly published name (Fig. 1). Based on the thresholds recommended for delineating a new species without carrying out DNA–DNA hybridization (Stackebrandt and Ebers 2006; Kim et al. 2014; Yarza et al. 2014), strain Marseille-P2948^T can be classified as a new species within the genus *Corynebacterium*. The reference MALDI-TOF spectrum of strain Marseille-P2948^T has been added to our database (http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database).

Phenotypic features

Strain Marseille-P2948^T was found to be a facultatively anaerobic bacterium. Growth was observed between 25 and 40 °C in anaerobic and aerobic conditions on Columbia agar with 5% sheep blood. Optimal growth was observed after 48 h at 37 °C in aerobic conditions. Growth was found to occur from pH 6.0 up to 8.0 and 0–5% NaCl. Surface colonies on blood agar (bioMérieux) were observed to be circular, glistening and slightly grayish with a mean diameter of 1 mm. Cells were found to be catalase positive and asporogenous forming; motility and oxidase tests were negative. Gram-staining showed Gram-positive rods. The ultrastructure of cells of strain Marseille-P2948^T and their size was determined using transmission electron microscopy (Supplementary Figure S1). The major fatty acids were identified as $C_{16:0}$ (41%), $C_{18:1n9}$ (26%), $C_{18:0}$ (13%) and $C_{18:2n6}$ (13%). Minor amounts of other unsaturated, branched and saturated fatty acids were also detected (Supplementary Table S1). These fatty acids are consistent with those described for the members of the genus *Corynebacterium* (Bernard and Funke 2015).

The API Coryne code obtained for strain Marseille-P2948^T was 4001165, which corresponds with 99.4% confidence to *Corynebacterium macginleyi*. Using API Coryne and API ZYM tests, positive enzymatic reactions were observed for urease, esterase, esterase lipase, lipase, leucine arylamidase, cysteine arylamidase, naphthol-AS-BI-phosphohydrolase, and alkaline

Table 1 Phenotypic characteristic of strain Marseille-P2948^T and the type strains of the phylogenetically closely related species

	1	2	3	4	5	6	7	8	9
Gram stain	+	+	+	+	+	+	+	+	+
Endospore formation	−	−	na	na	−	−	−	−	−
Mobility	−	−	−	−	−	−	−	−	−
Production of									
Alkaline phosphatase	+	+	+	+	+	+	v	+	+
β-Galactosidase	−	−	−	−	−	−	−	−	−
Catalase	+	+	+	+	+	+	+	+	+
Nitrate reductase	−	−	−	−	+	−	v	−	−
Urease	+	−	−	−	−	−	−	−	+
Production of									
Xylose	−	−	−	−	−	−	−	−	±
Mannitol	−	−	−	−	−	−	−	−	−
Glucose	+	+	+	−	+	+	+	+	+
Maltose	+	−	+	−	−	−	−	+	−
Lactose	+	−	+	na	−	−	−	−	−
G+C content (mol%)	65.03	59.42	64.26	64.84	59.01	66.84	60.01	59.57	65
Habitat	Vaginal swab	Sheep milk	Throat swabs	Vaginal swab	Axillar lymph node	Human blood	Lepromatous leprosy	Human blood	Human blood

Strains: 1, *Corynebacterium fournierii* strain Marseille-P2948^T; 2, *C. camporealensis* strain CRS-51^T; 3, *C. imitans* strain NCTC^T; 4, *C. lipophiloflavum* strain DMMZ 1944^T; 5, *C. simulans* Co 553^T; 6, *C. timonense* strain 5401744^T; 7, *Corynebacterium tuberculostearicum* strain CIP 107291^T; 8, *C. tuscaniae* strain ISS-5309^T; 9, *C. ureicelerivorans* strain IMMIB RIV-2301^T. The reference for the species data comes from descriptions of the original species. + Positive, − negative, *v* variable, *na* not available data

Table 2 Nucleotide content and gene count levels of the genome

Genome statistics	Value	% of total[a]
Size (bp)	2,383,644	100
G+C content (bp)	1,536,213	65.03
Coding region (bp)	2,009,136	84.29
Number total of genes	2210	100
RNA genes	63	2.85
Protein-coding genes	2147	97.15
Genes with function prediction	1496	69.68
Genes assigned to COGs	1311	61.06
Genes with peptide signals	326	15.18
Genes with transmembrane helices	522	24.31

[a]The total is based on either the size of the genome in base pairs or the total number of protein coding genes in the annotated genome

phosphatase. Nitrate reductase, esculin, galactosidase (α and β), β-glucuronidase, glucosidase (α and β), *N*-acetyl-β-glucosaminidase, α-mannosidase, α-fucosidase, and gelatin tests were found to be negative. API Coryne strips revealed that strain Marseille-P2948^T can produce acid from glucose, maltose, lactose and saccharose. In contrast, API 50 CH strips showed that salicin, cellobiose, maltose, and saccharose can be fermented. The phenotypic differences between strain Marseille-P2948^T and its phylogenetic neighbours are summarised in Table 1.

Strain Marseille-P2948^T was found to be susceptible to benzylpenicillin (MIC 0.016 µg/ml), amoxicillin (MIC 0.125 µg/ml), imipenem (MIC 0.016 µg/ml), rifampicin (MIC 0.032 µg/ml), and vancomycin (MIC 0.125 µg/ml) but resistant to metronidazole and ceftriaxone.

Genome properties

The draft genome of strain Marseille-P2948^T (accession number FWYQ00000000) is 2,383,644 bp long with a G+C content of 65.03% (Table 2; Supplementary Figure S2). It consists of 24 scaffolds (composed of 76 contigs). Of the 2210 predicted genes, 2147 are protein-coding genes and 63 encode RNAs (3 5S rRNA genes, 4 16S rRNA genes, 1 23S rRNA gene and 55 tRNA genes). A total of 1496 genes (70%) were assigned a putative function (by COGS or by NR blast). 58 genes were identified as ORFans (2.7%). The remaining genes were annotated as hypothetical proteins (524 genes, 24%). Genome statistics are summarised in Table 2. According to COG functional categories (Supplementary Table S2), translation (160, 7.5%) and amino acid transport and metabolism (141, 6.9%) were the most predominant gene categories.

Genome comparison

The genomic comparison of strain Marseille-P2948^T's properties with those of closely related *Corynebacterium* species is detailed in supplementary Table S3. The genome size, % G+C and gene contents of strain Marseille-P2948^T (2.383 Mb, 65.03% and 2147, respectively) are in the range of those of the other species compared (Supplementary Table S3). The distribution of genes into COG categories is similar between all compared genomes (Supplementary Figure S3). However, there are a high number of genes of strain Marseille-P2948^T present in COG category M (cell wall/membrane biogenesis) than in the other species compared. Strain Marseille-P2948^T presents also a higher number of genes in COG categories X (mobilome: prophages, transposons) and H (coenzyme transport and metabolism) than *Corynebacterium lipophiloflavum, Corynebacterium simulans* and *Corynebacterium timonense* and fewer genes in COG categories P (inorganic ion transport and metabolism) and E (amino acid transport and metabolism) than *Corynebacterium tuscaniense* and *C. ureicelerivorans.*

We used two parameters to evaluate genomic similarity among *Corynebacterium* species: AGIOS

and dDDH. The AGIOS analysis revealed that strain Marseille-P2948^T shares between 963 and 1294 orthologous genes with closely related species (with *Corynebacterium camporealensis* and *Corynebacterium timonensis*, respectively). The average nucleotide sequence identity among *Corynebacterium* species with standing in nomenclature ranges from 53% between *C. diphtheriae* and *C. simulans* to 76% between *Corynebacterium imitans* and *C. ureicelerivorans*. When compared to strain Marseille-P2948^T, the values ranged from 58% with *C. diphtheriae* to 68% with *C. ureicelerivorans* (Supplementary Table S4). In addition, the dDDH values ranged from 13.5% between *C. simulans* and *C. timonense* to 32.3% between *C. simulans* and *C. diphtheriae*. dDDH values between strain Marseille-P2948^T and the compared genomes varied between 20.5 and 27.8% with C. *lipophiloflavum* and *C. ureicelerivorans*, respectively (Supplementary Table S5). These low dDDH values support the conclusion that our isolate is a distinct species different from other members of the genus *Corynebacterium*.

Conclusion

Though the 16S rRNA sequence similarity (98.7%) is close to the threshold to distinguish a new species, data from phenotypic and phylogenetic analyses, the difference in the G+C content, AGIOS values and low dDDH values (< 70%) confirm that strain Marseille-P2948^T may be classified as the representative of a new species (Klenk et al. 2014), for which the name *Corynebacterium fournierii* sp. nov. is proposed. The Digital Protologue database (Rosselló-Móra et al. 2017) TaxoNumber for strain Marseille-P2948^T is TA00393.

Description of *Corynebacterium fournierii* sp. nov.

Corynebacterium fournierii sp. nov. (four.nier'i.i. N.L. gen. n. *fournierii*, in honour of the French scientist Pierre-Edouard Fournier for his outstanding contributions to medical microbiology and taxonomy).

Gram-positive rods with a mean breadth of 0.7 µm and a mean length of 1.4 µm. Facultatively anaerobic. Catalase and urease positive. Oxidase negative. Asporogenous and non-motile. Mesophilic; optimum growth occurs in aerobic conditions at 37 °C. On blood agar, colonies appear circular and slightly grayish with a mean diameter of 1 mm. The major fatty acids are $C_{16:0}$, $C_{18:1n9}$, $C_{18:0}$, and $C_{18:2n6}$.

The type strain, Marseille-P2948^T (= CSUR P2948 = DSM 103271) was isolated from a vaginal specimen of a French patient suffering from bacterial vaginosis in Marseille, France. The 16S rRNA, *rpoB*, and whole genome shotgun sequences have been deposited in EMBL-EBI under accession number LT576414, LT965931 and FWYQ00000000, respectively.

Acknowledgements This study was supported by the Fondation Méditerranée Infection and the French National Research Agency under the "Investissements d'avenir" program, reference ANR-10-IAHU-03. We also thank TradOnline for English reviewing and Frederic Cadoret for administrative assistance.

Conflict of interest The authors declare that they have no conflict of interest.

References

Belay N, Mukhopadhyay B, De Macario EC, Galask R, Daniels L (1990) Methanogenic bacteria in human vaginal samples. J Clin Microbiol 28:1666–1668

Bernard KA, Funke G (2015) *Corynebacterium*. In: Whitman WB, Rainey F, Kämpfer P, Trujillo M, Chun J, DeVos P et al (eds) Bergey's manual of systematics of archaea and bacteria. Wiley, Chichester, pp 1–70

Braun MS, Zimmermann S, Danner M, Rashid HO, Wink M (2016) *Corynebacterium uropygiale* sp. nov., isolated from the preen gland of Turkeys (*Meleagris gallopavo*). Syst Appl Microbiol 39:88–92

Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J (2009) DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 25:119–120

Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA (2012) Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics 28:464–469

Chen HH, Li WJ, Tang SK, Kroppenstedt RM, Stackebrandt E, Xu LH et al (2004) *Corynebacterium halotolerans* sp. nov., isolated from saline soil in the west of China. Int J Syst Evol Microbiol 54:779–782

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2016) GenBank. Nucleic Acids Res 44:D67–D72

Devriese LA, Riegel P, Hommez J, Vaneechoutte M, De Baere T, Haesebrouck F (2000) Identification of *Corynebacterium glucuronolyticum* strains from the urogenital tract of humans and pigs. J Clin Microbiol 38:4657–4659

Dione N, Sankar SA, Lagier JC, Khelaifia S, Michele C, Armstrong N et al (2016) Genome sequence and description of *Anaerosalibacter massiliensis* sp. nov. New Microbes New Infect 10:66–76

Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, Raoult D (2000) 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 38:3623–3630

Dubourg G, Lagier JC, Armougom F, Robert C, Hamad I, Brouqui P et al (2013) The gut microbiota of a patient with resistant tuberculosis is more comprehensively studied by culturomics than by metagenomics. Eur J Clin Microbiol Infect Dis 32:637–645

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Fettweis JM, Brooks JP, Serrano MG, Sheth NU, Girerd PH, Edwards DJ et al (2014) Differences in vaginal microbiome in African American women versus women of European ancestry. Microbiology 160:2272–2282

Fichorova R, Fraga J, Rappelli P, Fiori PL (2017) *Trichomonas vaginalis* infection in symbiosis with *Trichomonasvirus* and *Mycoplasma*. Res Microbiol. https://doi.org/10.1016/j.resmic.2017.03.005

Fournier PE, Lagier JC, Dubourg G, Raoult D (2015) From culturomics to taxonomogenomics: a need to change the taxonomy of prokaryotes in clinical microbiology. Anaerobe 36:73–78

Fudou R, Jojima Y, Seto A, Yamada K, Kimura E, Nakamatsu T et al (2002) *Corynebacterium efficiens* sp. nov., a glutamic-acid-producing species from soil and vegetables. Int J Syst Evol Microbiol 52:1127–1131

Funke G, Efstratiou A, Kuklinska D, Hutson RA, Zoysa AD, Engler KH et al (1997) *Corynebacterium imitans* sp. nov. isolated from patients with suspected diphtheria. J Clin Microbiol 35:1978–1983

Gouret P, Vitiello V, Balandraud N, Gilles A, Pontarotti P, Danchin EG (2005) FIGENIX: intelligent automation of genomic annotation: expertise integration in a new software platform. BMC Bioinform 6:1

Gouret P, Thompson JD, Pontarotti P (2009) PhyloPattern: regular expressions to identify complex patterns in phylogenetic trees. BMC Bioinform 10:298

Gouret P, Paganini J, Dainat J, Louati D, Darbo E, Pontarotti P et al (2011) Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: the multi-agent software system DAGOBAH. In: Pontarotti P (ed) Evolutionary biology—concepts, biodiversity, macroevolution and genome evolution. Springer, Berlin, pp 71–87

Herbst-Kralovetz MM, Pyles RB, Ratner AJ, Sycuro LK, Mitchell C (2016) New systems for studying intercellular interactions in bacterial vaginosis. J Infect Dis 214:S6–S13

Hoyles L, Ortman K, Cardew S, Foster G, Rogerson F, Falsen E (2013) *Corynebacterium uterequi* sp. nov., a non-lipophilic bacterium isolated from urogenital samples from horses. Vet Microbiol 165:469–474

Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform 11:1

Khamis A, Raoult D, La Scola B (2004) rpoB gene sequencing for identification of corynebacterium species. J Clin Microbiol 42:3925–3931

Kim M, Oh HS, Park SC, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351

Klenk HP, Meier-Kolthoff JP, Göker M (2014) Taxonomic use of DNA G+C content and DNA–DNA hybridization in the genomic age. Int J Syst Evol Microbiol 64:352–356

Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108

Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C et al (2012) Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 18:1185–1193

Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, Raoult D (2015) The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 28:237–264

Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P et al (2016) Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 1:16203

Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ (2011) Proteinortho: detection of (Co-) orthologs in large-scale analysis. BMC Bioinform 12:1

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

Matuschek E, Brown DFJ, Kahlmeter G (2014) Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. Clin Microbiol Infect 20:O255–O266

Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinform 14:1

Menard JP, Fenollar F, Henry M, Bretelle F, Raoult D (2008) Molecular Quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. Clin Infect Dis 47:33–43

Mendling W (2016) Vaginal microbiota. In: Schwiertz A (ed) Microbiota of the human body, vol 902. Springer International Publishing, Cham, pp 83–93

Mishra AK, Lagier JC, Nguyen TT, Raoult D, Fournier PE (2013) Non contiguous-finished genome sequence and description of *Peptoniphilus senegalensis* sp. nov. Stand Genomic Sci 7:370–381

Murray PR, Baron EJ, Jorgensen JH, Landry ML, Pfaller MA (2007) Manual of clinical microbiology, 9th edn. ASM Press, Washington, DC

Narayankhedkar A, Hodiwala A, Mane A (2015) Clinicoetiological characterization of infectious vaginitis amongst women of reproductive age group from Navi Mumbai, India. J Sex Transm Dis 2015:1–5

Negi V, Singh Y, Schumann P, Lal R (2016) *Corynebacterium pollutisoli* sp. nov., isolated from hexachlorocyclohexane-contaminated soil. Int J Syst Evol Microbiol 66:3531–3537

Pal K, Roy S, Behera B, Kumar N, Sagiri S, Ray S (2011) Bacterial vaginosis: etiology and modalities of treatment—A brief note. J Pharm Bioallied Sci 3:496

Price MN, Dehal PS, Arkin AP (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 26:1641–1650

Ramasamy D, Mishra AK, Lagier J-C, Padhmanabhan R, Rossi M, Sentausa E et al (2014) A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 64:384–391

Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L et al (2014) The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. Microbiome 2:4

Rosselló-Móra R, Trujillo ME, Sutcliffe IC (2017) Introducing a digital protologue: a timely move towards a database-driven systematics of archaea and bacteria. Antonie Van Leeuwenhoek 110:455–456

Sasser M (2006) Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). MIDI, Technical Note

Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P, Rolain JM et al (2009) Ongoing revolution in bacteriology: routine Identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 49:543–551

Shukla SK, Bernard KA, Harney M, Frank DN, Reed KD (2003) *Corynebacterium nigricans* sp. nov.: proposed name for a black-pigmented *Corynebacterium* species recovered from the human female urogenital tract. J Clin Microbiol 41:4353–4358

Stackebrandt E, Ebers J (2006) Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 33:152

Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH et al (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol 12:635–645

Zapata HJ, Quagliarello VJ (2015) The microbiota and microbiome in aging: potential implications in health and age-related diseases. J Am Geriatr Soc 63:776–781

Article 7:

Description of *Janibacter massiliensis* sp. nov. cultured from the vaginal discharge of a patient with bacterial vaginosis.

Mossaab Maaloum*, **Khoudia Diop***, Awa Diop, Enora Tomei, Nicholas Armstrong, Jaishriram Rathored, Didier Raoult, Florence Fenollar, Pierre-Edouard Fournier

*Both authors contributed equally to the work

Submitted in Antonie van Leeuwenhoek

Antonie van Leeuwenhoek

Description of Janibacter massiliensis sp. nov. cultured from the vaginal discharge of a patient with bacterial vaginosis

--Manuscript Draft--

Manuscript Number:	
Full Title:	Description of Janibacter massiliensis sp. nov. cultured from the vaginal discharge of a patient with bacterial vaginosis
Article Type:	Original Article
Keywords:	Bacterial vaginosis; Culturomics; Janibacter massiliensis; Taxonogenomics; Vaginal microbiota.
Corresponding Author:	Pierre-Edouard Fournier Unite de recherche sur les maladies infectieuses et tropicales emergentes FRANCE
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Unite de recherche sur les maladies infectieuses et tropicales emergentes
Corresponding Author's Secondary Institution:	
First Author:	Mossaab Maaloum
First Author Secondary Information:	
Order of Authors:	Mossaab Maaloum
	Khoudia Diop
	Awa Diop
	Enora Tomei
	Magali Richez
	Jaishriram Rathored
	Florence Bretelle
	Didier Raoult
	Florence Fenollar
	Pierre-Edouard Fournier
Order of Authors Secondary Information:	
Funding Information:	Agence Nationale de la Recherche (ANR-10-IAHU-03) — Not applicable Fondation Méditerranée Infection — Dr Khoudia Diop
Abstract:	Strain Marseille-P4121 was isolated from a vaginal sample of a 45-year-old French woman with bacterial vaginosis. It is a Gram-positive, none-spore forming, non-motile and aerobic bacterium. Strain Marseille-P4121 exhibits a 98.2% 16S rRNA sequence similarity with Janibacter alkaliphilus strain SCSIO 10480T, the phylogenetically related species standing in nomenclature. Its major fatty acids are C18:1ω9 (34.4%), C16:0 (30.1%), and C18:0 (19%). The genome size of strain Marseille-P4121 is 2,452,608 bp long with a 72.5% G+C content and contains 2,351 protein-coding genes and 49 RNA genes including 3 rRNA genes. We propose that strain Marseille-P4121T (= CECT 9671 = CSUR P4121) is the type strain of the new species Janibacter massiliensis sp. nov. It was cultured from a vaginal sample of a French patient suffering from bacterial vaginosis.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Manuscript
Click here to access/download;Manuscript;MS Janibacter.docx
Click here to view linked References

1 **Description of *Janibacter massiliensis* sp. nov. cultured from the vaginal discharge of a**
2 **patient with bacterial vaginosis**

3

4 Mossaab Maaloum[1,2*], Khoudia Diop[1*], Awa Diop[1], Enora Tomei[3], Magali Richez[3],
5 Jaishriram Rathored[3], Florence Bretelle[3,4], Didier Raoult[3,5], Florence Fenollar[1] and Pierre-
6 Edouard Fournier[1†]

7

8 [1]Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection, Marseille,
9 France
10 [2]Faculty of Sciences Ben M'sik, Laboratory of Biology and Health, Hassan II University, 10
11 Casablanca, Morocco
12 [3]Aix-Marseille Univ, IRD, AP-HM, MEPHI, IHU-Méditerranée Infection, Marseille, France
13 [4]Department of Gynecology and Obstetrics, Gynépole, Hôpital Nord, AP-HM, Marseille,
14 France
15 [5]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz
16 University, Jeddah, Saudi Arabia
17 [*]Both authors contributed equally to the work
18 [†]**Corresponding author:** Prof. Pierre-Edouard FOURNIER
19 *Institut Hospitalo-Universitaire Méditerranée-Infection*
20 *19-21 Boulevard Jean Moulin*
21 *13005 Marseille Cedex 05*
22 *France*
23 *Tel : +33 413 732 401. Fax: +33 413 732 402*
24 Email*:* pierre-edouard.fournier@univ-amu.fr

25 **Running title:** *Janibacter massiliensis* sp. nov.

26 **Abstract count:** 126

27 **Word count:** 2,748

28 **ABSTRACT**

29 Strain Marseille-P4121 was isolated from a vaginal sample of a 45-year-old French woman
30 with bacterial vaginosis. It is a Gram-positive, none-spore forming, non-motile and aerobic
31 bacterium. Strain Marseille-P4121 exhibits a 98.2% 16S rRNA sequence similarity with
32 *Janibacter alkaliphilus* strain SCSIO 10480^T, the phylogenetically related species standing in
33 nomenclature. Its major fatty acids are $C_{18:1\omega9}$ (34.4%), $C_{16:0}$ (30.1%), and $C_{18:0}$ (19%). The
34 genome size of strain Marseille-P4121 is 2,452,608 bp long with a 72.5% G+C content and
35 contains 2,351 protein-coding genes and 49 RNA genes including 3 rRNA genes. We propose
36 that strain Marseille-P4121^T (= CECT 9671 = CSUR P4121) is the type strain of the new
37 species *Janibacter massiliensis* sp. nov. It was cultured from a vaginal sample of a French
38 patient suffering from bacterial vaginosis.
39
40
41 **Keywords**: Bacterial vaginosis ; Culturomics ; *Janibacter massiliensis* ; Taxonogenomics;
42 Vaginal microbiota.

43 **Abbreviations**

44 **CSUR:** Collection de souches de l'Unité des Rickettsies

45 **CECT:** Colección española de cultivos tipo

46 **MALDI-TOF:** Matrix-assisted laser-desorption/ionization time-of-flight

47 **TE buffer:** Tris-EDTA buffer

48 **INTRODUCTION**

49 The genus *Janibacter* was first described in 1997 after Martin and Groth isolated a new
50 bacterium from sludge from a wastewater treatment plant (Martin and Groth 1997). The genus
51 name derives from the two-faced Roman god Janus, as the cells of the original strain had a
52 rod-coccus cycle during growth. The genus *Janibacter* includes aerobic, Gram-positive
53 bacteria that are non-motile, non-spore forming, catalase-positive, and oxidase variable.
54 Colonies grown on solid agar are usually smooth, circular and convex. They also vary in color
55 from white to yellow (Fernández-Natal et al. 2015).

56 To date, nine *Janibacter* species have been cultured from different environments
57 (http://www.bacterio.net/janibacter.html). *Janibacter alkaliphilus* and *J. corallicola* have been
58 isolated from coral, *J. limosus* and *J. terrae* from wastewater-contaminated soils, *J. indicus*
59 from hydrothermal sediment of the Indian Ocean, *J. hoylei* from air, *J. anophelis* from insects
60 and *J. melonis*from plants (Martin and Groth 1997; Imamura et al. 2000; Lang et al. 2003;
61 Yoon 2004; Kampfer 2006; Kageyama et al. 2007; Shivaji et al. 2009; Li et al. 2012; Hamada
62 et al. 2013; Zhang et al. 2014; Fernández-Natal et al. 2015).

63 In the literature, we found nine cases of *Janibacter* infections in humans, including five
64 cases of *J. terrae* bacteremia including four in febrile patients with several underlying
65 conditions (Fernández-Natal et al. 2015) and one in a male Caucasian road worker with
66 bilateral psoas abscess (Wan et al. 2017), two cases of *J. melonis* bacteremia in a patient with
67 low-grade fever and right-sided facial swelling (Elsayed and Zhang 2005) and a second
68 patient with coeliac disease (Chander et al. 2018), one case of *J. hoylei* bacteremia in a 8-
69 week-old febrile infant (Lim et al. 2017), and a case of *Janibacter* sp. bacteremia in a man
70 treated for myeloid leukemia (Loubinoux et al. 2005) (Supplementary Table S1).

71 In this study, we describe a new *Janibacter* species that was isolated from a vaginal
72 sample of a 45-year-old French woman with bacterial vaginosis. The new species, for which

73 we propose the name *Janibacter massiliensis* sp. nov., was characterized using a combination
74 of phenotypic and genotypic characteristics using the previously described taxono-genomics
75 strategy (Ramasamy et al. 2014; Fournier et al. 2015).

76 **MATERIALS AND METHODS**

77 ***Sample collection***

78 This study was validated by the local IFR48 ethics committee (Marseille, France) under
79 agreement number 09-022. In October 2017, a vaginal sample from a 45-year-old French
80 woman suffering from bacterial vaginosis was collected at the Timone hospital in Marseille
81 (France). At the time of sampling, she was not treated with any antibiotic. She gave an
82 informed and written consent to be included in the study. The sample was collected and
83 transported to our laboratory using a Sigma Transwab (Medical Wire, Corsham, United
84 Kingdom).

85 ***Strain isolation and identification***

86 Within 2 hours after sampling, the specimen was pre-incubated in an anaerobic blood
87 culture vial (BD Diagnostics, Le Pont-de-Claix, France) enriched with 4 mL of rumen and 3
88 mL of blood, both from sheep (bioMérieux, Marcy l'Etoile, France). After ten days of pre-
89 incubation at 37°C, the supernatant was inoculated on 5% sheep blood- and vitamin K-
90 enriched Schaedler agar (BD Diagnostics). After 5 days of incubation in anaerobic
91 atmosphere at 37°C, isolated colonies were deposited individually and in duplicate on a MTP
92 96 MALDI-TOF MS target plate (Bruker Daltonics, Leipzig, Germany) for identification with
93 a Microflex spectrometer (Bruker) (Seng et al. 2009). All obtained MS spectra were compared
94 using the MALDI Biotyper software to those in the MALDI-TOF MS database (Bruker). If
95 the score was > 2.0, the strain was considered as identified. Otherwise, the identification
96 failed. For unidentified isolates, the 16S rRNA gene was amplified and sequenced to achieve
97 identification (Drancourt et al. 2000). As suggested by Stackebrandt and Ebers, if the 16S

98 rRNA sequence similarity value was lower than 98.7% or 95%, the strain was considered as a
99 putative new species or genus, respectively (Stackebrandt and Ebers 2006; Kim et al. 2014).

100 ***Bacterial morphology and growth conditions***

101 Cells were fixed with 2.5% glutaraldehyde in a 0.1 M cacodylate buffer for at least one
102 hour at +4°C. One drop of cell suspension was deposited for approximately five minutes on
103 glow-discharged formvar-carbon film on 400 mesh nickel grids (FCF400-Ni, EMS). The grids
104 were dried on blotting paper and the cells were negatively stained for 10 seconds with 1%
105 ammonium molybdate solution in filtered water at room temperature. Electron micrographs
106 were acquired using a Tecnai G20 Cryo (FEI) transmission electron microscope operated at
107 200 keV. Gram-stain, motility, sporulation, catalase and oxidase tests were performed as
108 previously reported (Murray et al. 2007).

109 The growth of strain Marseille-P4121 was tested on 5% sheep blood-enriched Columbia
110 agar (bioMérieux) at various temperatures (ambient, 28, 37, 42 and 56°C), atmospheres
111 (aerobic, anaerobic and microaerophilic), pH conditions (5, 6, 6.5, 7, and 8.5) and NaCl
112 concentrations (0, 5, 15, and 45%).

113 ***Biochemical and chemotypic features***

114 Biochemical properties of strain Marseille-P4121 were tested using API ZYM, API 20A
115 and API 32A strips (bioMérieux) according to the manufacturer's instructions. Minimal
116 inhibitory concentrations (MICs) of amoxicillin, benzylpenicillin, cefotaxime, imipenem,
117 colistin, erythromycin, metronidazole, rifampin and vancomycin were determined according
118 to EUCAST recommendations using E-test gradient strips (bioMérieux) (Citron et al. 1991;
119 Matuschek et al. 2014).

120 Cellular fatty acid methyl ester (FAME) analysis was performed by Gas
121 Chromatography/Mass Spectrometry (GC/MS). Two tubes were prepared with approximately
122 15 mg of bacterial biomass per tube, harvested from several culture plates. FAME was

123 prepared for analysis as described by Sasser (Sasser 2006). GC/MS analyses were carried out
124 as previously described (Dione et al. 2016).

125 ***Genome extraction and sequencing***

126 For genomic DNA (gDNA) of strain Marseille-P4121, a mechanical shearing was first
127 performed by acid-washed glass beads (G4649-500g Sigma) using a FastPrep BIO 101
128 instrument (Qbiogene, Strasbourg, France) at maximum speed (6.5) for 90s. Then after a 2-
129 hour lysozyme incubation at 37°C, DNA was extracted using an EZ1 biorobot (Qiagen) and
130 the EZ1 DNA Tissue kit. The gDNA was sequenced on a MiSeq sequencer (Illumina Inc, San
131 Diego, CA, USA) with the Paired-End and Mate-Pair strategies. The gDNA was barcoded in
132 order to be mixed with 16 other projects for the Nextera XT DNA sample prep kit (Illumina),
133 and with 11 other projects for the Nextera Mate-Pair sample prep kit (Illumina).

134 To prepare the Paired-End library, gDNAs were diluted to obtained 1ng of each genome
135 as input. The tagmentation step fragmented and tagged the DNA. Then, limited cycle PCR
136 amplification (12 cycles) completed the tag adapters and introduced dual-index barcodes. The
137 library profile was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies Inc,
138 Santa Clara, CA, USA) with a DNA High sensitivity labchip and the fragment size was
139 estimated to be 1.5 kb. After purification on AMPure XP beads (Beckman Coulter Inc,
140 Fullerton, CA, USA), libraries were then normalized on specific beads according to the
141 Nextera XT protocol (Illumina). Normalized libraries were pooled for sequencing on a MiSeq
142 sequencer. Automated cluster generation and Paired-End sequencing with dual index reads
143 were performed in a single 39-hour run in 2x250-bp. A total information of 0.5 Gb was
144 obtained from a 1160 k/mm 2 cluster density with a cluster passing quality control filters of
145 91.7 % (20,276,000 passed filtered clusters). Within this run, the index representation for
146 strain Marseille-P4121 was determined to be 3.48%. The 704,704 Paired-End reads were
147 trimmed and filtered according to the read qualities.

148 The Mate-Pair library was prepared with 1.5 µg of genomic DNA using the Nextera
149 Mate-Pair Illumina guide. The gDNA samples were simultaneously fragmented and tagged
150 with a Mate-Pair junction adapter. The fragmentation pattern was validated on an Agilent
151 2100 BioAnalyzer (Agilent) with a DNA 7500 labchip. The DNA fragments ranged in size
152 from 1.5 kb up to 11kb with an optimal size of 8.855 kb. No size selection was performed
153 and, for each genome, 600 ng of tagmented fragments were circularized. The circularized
154 DNA was mechanically sheared to small fragments with an optimal at 456 bp on the Covaris
155 device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a
156 High Sensitivity Bioanalyzer LabChip (Agilent) and the final concentration library was
157 measured at 6.83 nmol/l. The libraries were normalized at 2nM and pooled. After a
158 denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent
159 cartridge and then onto the instrument along with the flow cell. Automated cluster generation
160 and sequencing run were performed in a single 39-hour run in a 2x151-bp. A total information
161 of 2.9 Gb was obtained from a 290K/mm2 cluster density with a cluster passing quality
162 control filters of 98.57 % (5,767,000 passing filters paired reads). Within this run, the index
163 representation for Marseille-P4121 was determined to be 7.57%. The 436,510 paired reads
164 were trimmed and then assembled with the Paired-End reads.

165 ***Genome annotation and comparison***

166 We used Prodigal (Hyatt et al. 2010) with default parameters for predicting Open
167 Reading Frames (ORFs). The predicted ORFs were excluded if they spanned a sequencing
168 gap region (contained Ns), The predicted bacterial protein sequences were searched against
169 the Clusters of Orthologous Groups (COG) database and NR database using BLASTP. The
170 tRNAs and rRNAs were predicted using the tRNAScan-SE (Avni et al. 1997) and RNAmmer
171 tools (Lagesen et al. 2007), respectively, Signal peptides and numbers of transmembrane
172 helices were predicted using SignalP (Dyrløv Bendtsen et al. 2004) and TMHMM (Krogh et

173 al. 2001), respectively. The PHAge Search Tool (PHAST) was used to detect prophage
174 sequences (Zhou et al. 2011). Artemis (Rutherford et al. 2000) and DNA Plotter (Carver et al.
175 2009) were used for data management and visualization of genomic features, respectively.
176 Rast (Rapid Annotations using Subsystems Technology) was used for identification of
177 resistance-coding genes (Aziz et al. 2008).

178 The mean level of nucleotide sequence similarity at the genome level between strain
179 Marseille-P4121 and 7 closely related bacteria was estimated using the Average Genomic
180 Identity of orthologous gene Sequences (AGIOS) (Ramasamy et al. 2014) and digital DNA-
181 DNA hybridization (dDDH) (Meier-Kolthoff et al. 2013) parameters. The genome from strain
182 Marseille-P4121 was compared to those of *Janibacter anophelis* strain NBRC 107843^{T}
183 (BCSQ00000000), *J. corallicola* strain NBRC 107790^{T} (BCSR00000000), *J. hoylei* strain
184 PVAS-1^{T} (ALWX00000000), *J. melonis* strain CD 11-4 (LQZG00000000), *J. terrae* strain
185 NBRC 107853^{T} (BCUV00000000), *Knoellia flava* strain TL1^{T} (AVPI00000000) and
186 *Tetrasphaera duodecadis* strain DSM 12806^{T} (PJNE01000000)

187 **RESULTS and DISCUSSION**

188 ***Strain identification and phylogenetic analysis***

189 MALDI-TOF MS identification of our isolate provided a score < 1.466 with *Kytococcus*
190 *sedentarius* suggesting that the bacterium's spectrum was not in our database. Its 16S rRNA
191 gene sequence exhibited a 98.2% similarity with *J. alkaliphilus* strain SCSIO 10480^{T}
192 (accession number JN160681), the phylogenetically closest species with standing in
193 nomenclature (Figure 1). This value was lower than the 98.65% set to differentiate species
194 (Kim et al. 2014). Subsequently, strain Marseille-P4121 was placed in a new species within
195 the genus *Janibacter* for which the name *Janibacter massiliensis* sp. nov., is proposed. The
196 reference spectrum was incremented in our URMS database.

197 ***Phenotypic and chemotaxonomic* features**

198 Strain Marseille-P4121 was aero-anaerobic. After 48 hours of growth on sheep blood-
199 enriched Columbia agar (BD diagnostic), the colonies were beige, circular and convex, with a
200 diameter of approximately 2 mm. Cells were not motile and did not form spores. They stained
201 Gram-positive, were coccoid and had a diameter of 0.7-1 µm. Catalase activity, but not
202 oxidase, urease, indole and nitrate, was positive. Strain Marseille-P4121 grew in aerobic
203 atmosphere at temperatures ranging from 28 - 37 °C, pH between 6.5 and 8.5, and NaCl
204 concentration < 5 g/L. However, weak growth was also observed at 37°C in anaerobic
205 atmosphere.

206 Using API ZYM and API rapid ID 32A strips (bioMérieux), alkaline phosphatase and
207 α-glucosidase activities were positive whereas α-fucosidase, galactosidase (α and β), β-
208 glucuronidase, β-glucosidase, N-acetyl-β-glucosaminidase were negative. In addition, using
209 API ZYM, esterase, esterase lipase, lipase, leucine arylamidase, acid phosphatase and
210 Naphthol-AS-BI-phosphohydrolase enzymatic activities were positive, while, α-
211 chymotrypsin, cystine arylamidase, α-mannosidase, trypsin and valine arylamidase activities
212 were negative. Using an API 20A strip, acid was produced from D-glucose and D-lactose, but
213 not from glycerol, L-arabinose, D-cellobiose, D-maltose, D-mannitol, D-mannose, D-
214 melezitose, D-raffinose, L-rhamnose, D-saccharose, and D-xylose. Using an API 32Astrip, D-
215 mannose and D-raffinose were not assimilated. In contrast enzymatic activities were detected
216 for arginine arylamidase, proline arylamidase, phenylalanine arylamidase, tyrosine
217 arylamidase, α-arabinosidase, alanine arylamidase, glycine arylamidase, histidine arylamidase
218 and serine arylamidase. Chemotypic analysis revealed that major cellular fatty acids were
219 $C_{18:1\omega 9}$ (34.4%), $C_{16:0}$ (30.1%) and $C_{18:0}$ (19%). Several branched structures were also detected
220 in low amounts. By comparison with closely related species, strain Marseille-P4121 differed
221 in a combination of glucose acidification and absence of gelatin hydrolysis (Table 2).

222 Strain Marseille-P4121 was susceptible to benzylpenicillin (MIC 0.19 µg/mL),
223 amoxicillin (MIC 0.38 µg/mL), cefotaxime (MIC 0.064 µg/mL), imipenem (MIC 0.003
224 µg/mL), rifampicin (MIC < 0.002 µg/mL), vancomycin (MIC 0.19 µg/mL) and colistin (MIC
225 2 µg/mL), but resistant to erythromycin and metronidazole (both exhibiting MICs >256
226 µg/mL).

227 ***Genome characteristic***

228 The genome size of strain Marseille-P4121 was 2,452,608-bp long with a 72.5% G+C
229 content (Table 3, Figure 3). Of the 2,400 predicted chromosomal genes, 2,351 where protein-
230 coding genes, and 49 were RNAs (one 16S rRNA, one 23S rRNA, one 5S rRNAs and 46
231 tRNAs genes). A total of 1,175 genes (48.85%) were assigned a putative function, 447 genes
232 (18,62%) were identified as ORFans and 658 genes (27.41%) were annotated as hypothetical
233 proteins**.** The properties and statistics of the genome are summarized in Table 3 and the
234 distribution of genes into COGs functional categories is presented in Table 4.

235 ***Genome comparison***

236 The comparison of the genomes from strain Marseille-P4121 and those of its
237 phylogenetically closest neighbors is presented in Table 5. The genome size (Mb) and gene
238 count of strain Marseille-P4121 (2.4 and 2,351, respectively) was smaller than those of all
239 compared genome. In contrast, its G+C content (72.5 %) was larger than those of *J. terrae*
240 (69%), *J. anophelis* (71.3%), *J. corallicola* (71%), and *Knoellia flava* (70.9%) but smaller
241 than those of *J. hoylei* (72.8%), *J. melonis* (73.0%), and *Tetrasphaera duodecadis* (73%). In
242 addition, the distribution of genes into COG categories was similar in all compared genomes
243 (Figure 4).

244 Among *Janibacter* species, dDDH values varied from 20 +/- 2.3% between *J.*
245 *corralicola* and *J. melonis* to 28.1 +/- 2.4% between *J. anophelis* and *J. hoylei*. When strain
246 Marseille-P4121 was compared to *Janibacter* species, dDDH values varied from 19.5 +/-

247 2.25% with *J. corralicola* to 20.2 +/- 2.3% with *J. melonis* and *J. terrae* (Table 6). Moreover,
248 among *Janibacter* species, AGIOS values varied from 76.76% between *J. corralicola* and *J.*
249 *melonis* to 85.24% between *J. anophelis* and *J. terrae*. When strain Marseille-P4121 was
250 compared to *Janibacter* species, AGIOS values varied from 74.88% with *J. corralicola* to
251 76.34% with *J. melonis* (Table 7). Thus, strain Marseille-P4121 exhibited genomic similarity
252 values with *Janibacter* species in the range of those observed among *Janibacter* species
253 themselves.

254 Thus, the combination of phylogenetic, phenotypic, and genomic characteristics
255 supports the classification of strain Marseille-P4121 within the genus *Janibacter*. However, it
256 can be distinguished from other *Janibacter* species based of its phenotypic characteristics,
257 notably its fatty acid composition (Table 2), and its genomic properties, notably dDDH values
258 below the 70% threshold that delineates bacterial species (-Kolthoff et al. 2014) (Table 7).
259 Moreover, the 16S rRNA nucleotide sequence identity value (98.02%) with the closest
260 phylogenetic neighbor being lower than the 98.65% threshold fixed to separate two species
261 (Kim et al. 2014; Yarza et al. 2014) confirms the classification of strain Marseille-P4121
262 within a new species.

263 **Conclusion**

264 On the basis of phenotypic, phylogenetic and genomic data, strain Marseille-P4121
265 differed sufficiently from its closest phylogenetic neighbors to be classified into a novel
266 species belonging to the genus *Janibacter*. The name *Janibacter massiliensis* sp. nov. is
267 proposed for this new taxon.

268 **Description of *Janibacter massiliensis* sp. nov.**

269 *Janibacter massiliensis* (mas.si.li.en'sis, L. masc. adj. *massiliensis*, of Massilia, the
270 Roman name of Marseille where strain Marseille-P4121^{T} was first isolated).

271 Non-motile and non-sporulating. Gram-stain positive cocco-baccilary bacteria that
272 occur singly, in pairs or short chains, with a width of 0.7 µm and a length of 1 µm. Catalase is
273 positive. Oxidase, urease, indole, and nitrate activities negative. Gelatin and aesculin are not
274 hydrolyzed. On 5% sheep blood-enriched Columbia agar, after 48 hours of aerobic culture at
275 37°C, colonies are circular with a mean diameter of 2 mm, convex and beige to yellow. The
276 major fatty acids are $C_{18:1\omega9}$, $C_{16:0}$, and $C_{18:0}$.

277 The type strain, Marseille-P4121^{T} (= CECT 9671 = CSUR P4121), was cultured from
278 the vaginal discharge of a 45-year-old French woman with bacterial vaginosis. Its Digital
279 Protologue database (Rosselló-Móra et al. 2017) TaxoNumber is TA00784. The 16S rRNA
280 gene and whole genome shotgun sequences were deposited in GenBank under accession
281 numbers LT969384.1 and OKQN00000000, respectively.

282 **Conflict of Interest**

283 The authors declare no conflit of interest.

284

285 **Acknowledgments**

286 This study was supported by the Institut Hospitalo-Universitaire (IHU) Méditerranée
287 Infection, the National Research Agency under the program "Investissements d'avenir",
288 reference ANR-10-IAHU-03, the Région Provence Alpes Côte d'Azur and the European
289 funding FEDER PRIMI.

290 **References**

291 Avni D, Biberman Y, Meyuhas O (1997) The 5′ Terminal Oligopyrimidine Tract Confers
292 Translational Control on Top Mrnas in a Cell Type-and Sequence Context-Dependent
293 Manner. Nucleic Acids Res 25:995–1001

294 Aziz RK, Bartels D, Best AA, et al (2008) The RAST Server: rapid annotations using
295 subsystems technology. BMC Genomics 9:75

296 Carver T, Thomson N, Bleasby A, et al (2009) DNAPlotter: circular and linear interactive
297 genome visualization. Bioinformatics 25:119–120

298 Chander AM, Kochhar R, Dhawan DK, et al (2018) Genome sequence and comparative
299 genomic analysis of a clinically important strain CD11-4 of *Janibacter melonis* isolated
300 from celiac disease patient. Gut Pathog 10

301 Citron DM, Ostovari MI, Karlsson A, Goldstein EJ (1991) Evaluation of the E test for
302 susceptibility testing of anaerobic bacteria. J Clin Microbiol 29:2197–2203

303 Dione N, Sankar SA, Lagier JC, et al (2016) Genome sequence and description of
304 *Anaerosalibacter massiliensis* sp. nov. New Microbes New Infect 10:66–76

305 Drancourt M, Bollet C, Carlioz A, et al (2000) 16S ribosomal DNA sequence analysis of a
306 large collection of environmental and clinical unidentifiable bacterial isolates. J Clin
307 Microbiol 38:3623–3630

308 Dyrløv Bendtsen J, Nielsen H, von Heijne G, Brunak S (2004) Improved Prediction of Signal
309 Peptides: SignalP 3.0. J Mol Biol 340:783–795

310 Elsayed S, Zhang K (2005) Bacteremia Caused by *Janibacter melonis*. J Clin Microbiol
311 43:3537–3539

312 Fernández-Natal MI, Sáez-Nieto JA, Medina-Pascual MJ, et al (2015) First report of
313 bacteremia by *Janibacter terrae* in humans. Infection 43:103–106

314 Fournier P-E, Lagier JC, Dubourg G, Raoult D (2015) From culturomics to

315 taxonomogenomics: A need to change the taxonomy of prokaryotes in clinical
316 microbiology. Anaerobe 36:73–78
317 Hamada M, Shibata C, Tamura T, et al (2013) *Janibacter cremeus* sp. nov., an
318 *Actinobacterium* isolated from sea sediment. Int J Syst Evol Microbiol 63:3687–3690
319 Hyatt D, Chen GL, LoCascio PF, et al (2010) Prodigal: prokaryotic gene recognition and
320 translation initiation site identification. BMC Bioinformatics 11:119
321 Imamura Y, Ikeda M, Yoshida S, Kuraishi H (2000) *Janibacter brevis* sp. nov., a new
322 trichloroethylene-degrading bacterium isolated from polluted environments. Int J Syst
323 Evol Microbiol 50:1899–1903
324 Ishikawa T, Yokota A (2006) Reclassification of *Arthrobacter duodecadis* Lochhead 1958 as
325 *Tetrasphaera duodecadis* comb. nov. and emended description of the genus
326 *Tetrasphaera*. Int J Syst Evol Microbiol 56:1369–1373
327 Kageyama A, Takahashi Y, Yasumoto-Hirose M, et al (2007) *Janibacter corallicola* sp. nov.,
328 isolated from coral in Palau. J Gen Appl Microbiol 53:185–189
329 Kampfer P (2006) *Janibacter anophelis* sp. nov., isolated from the midgut of Anopheles
330 arabiensis. Int J Syst Evol Microbiol 56:389–392
331 Kim M, Oh HS, Park SC, Chun J (2014) Towards a taxonomic coherence between average
332 nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of
333 prokaryotes. Int J Syst Evol Microbiol 64:346–351
334 Krogh A, Larsson B, von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane
335 protein topology with a hidden markov model: application to complete
336 genomes11Edited by F. Cohen. J Mol Biol 305:567–580
337 Lagesen K, Hallin P, Rødland EA, et al (2007) RNAmmer: consistent and rapid annotation of
338 ribosomal RNA genes. Nucleic Acids Res 35:3100–3108
339 Lang E, Kroppenstedt, RM, Swidersk J, et al (2003) Emended description of *Janibacter*

340 *terrae*, including ten dibenzofuran-degrading strains and *Janibacter brevis* as its later
341 heterotypic synonym. Int J Syst Evol Microbiol 53:1999–2005
342 Li J, Long LJ, Yang L, et al (2012) *Janibacter alkaliphilus* sp. nov., isolated from coral
343 *Anthogorgia* sp. Antonie Van Leeuwenhoek 102:157–162
344 Lim YK, Kweon OJ, Kim HR, et al (2017) First case of bacteremia caused by J*anibacter*
345 *hoylei*. APMIS 125:665–668
346 Loubinoux J, Rio B, Mihaila L, et al (2005) Bacteremia Caused by an Undescribed Species of
347 *Janibacter*. J Clin Microbiol 43:3564–3566
348 Martin K, Groth' I (1997) *Janibacter limosus* gen. nov., sp. nov., a New *Actinomycete* with
349 meso-Uiaminopimelic Acid in the Cell Wall. Int J Syst Bacteriol 47:529–534
350 Matuschek E, Brown DFJ, Kahlmeter G (2014) Development of the EUCAST disk diffusion
351 antimicrobial susceptibility testing method and its implementation in routine
352 microbiology laboratories. Clin Microbiol Infect 20:O255–O266
353 Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M (2013) Genome sequence-based species
354 delimitation with confidence intervals and improved distance functions. BMC
355 Bioinformatics 14:60
356 Meier-Kolthoff JP, Klenk H-P, Göker M (2014) Taxonomic use of DNA G+C content and
357 DNA–DNA hybridization in the genomic age. Int J Syst Evol Microbiol 64:352–356
358 Murray PR, Baron EJ, Jorgensen JH, et al (2007) Manual of clinical microbiology., 9th edn.
359 Washington, D.C: ASM Press
360 Ramasamy D, Mishra AK, Lagier JC, et al (2014) A polyphasic strategy incorporating
361 genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol
362 Microbiol 64:384–391
363 Rosselló-Móra R, Trujillo ME, Sutcliffe IC (2017) Introducing a digital protologue: a timely
364 move towards a database-driven systematics of archaea and bacteria. Antonie Van

365 Leeuwenhoek 110:455–456

366 Rutherford K, Parkhill J, Crook J, et al (2000) Artemis: sequence visualization and

367 annotation. Bioinformatics 16:944–945

368 Sasser M (2006) Bacterial Identification by Gas Chromatographic Analysis of Fatty acids

369 Methyl Esters (GC-FAME). MIDI, Technical Note

370 Seng P, Drancourt M, Gouriet F, et al (2009) Ongoing Revolution in Bacteriology: Routine

371 Identification of Bacteria by Matrix Assisted Laser Desorption Ionization

372 Time- of- Flight Mass Spectrometry. Clin Infect Dis 49:543–551

373 Shivaji S, Chaturvedi P, Begum Z, et al (2009) *Janibacter hoylei* sp. nov., *Bacillus isronensis*

374 sp. nov. and *Bacillus aryabhattai* sp. nov., isolated from cryotubes used for collecting

375 air from the upper atmosphere. Int J Syst Evol Microbiol 59:2977–2986

376 Stackebrandt E and Ebers J (2006) Taxonomic parameters revisited: tarnished gold standards.

377 Microbiol Today 33:152

378 Wan WY, Mughal A, Ward K (2017) Bilateral psoas abscess caused by *Janibacter terrae*, an

379 unusual condition and organism resulting in laboratory and management conundrums.

380 Acta Clin Belg 72:336–339

381 Yarza P, Yilmaz P, Pruesse E, et al (2014) Uniting the classification of cultured and

382 uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol

383 12:635–645

384 Yoon J-H (2004) *Janibacter melonis* sp. nov., isolated from abnormally spoiled oriental

385 melon in Korea. Int J Syst Evol Microbiol 54:1975–1980.

386 Yu X, Du Y, Wang G (2012) *Knoellia flava* sp. nov., isolated from pig manure. Int J Syst

387 Evol Microbiol 62:384–389.

388 Zhang G, Ren H, Wang S, et al (2014) *Janibacter indicus* sp. nov., isolated from

389 hydrothermal sediment of the Indian Ocean. Int J Syst Evol Microbiol 64:2353–2357.

390 Zhou Y, Liang Y, Lynch KH, et al (2011) PHAST: A Fast Phage Search Tool. Nucleic Acids
391 Res 39:W347–W352.
392

393 **Figures legend**

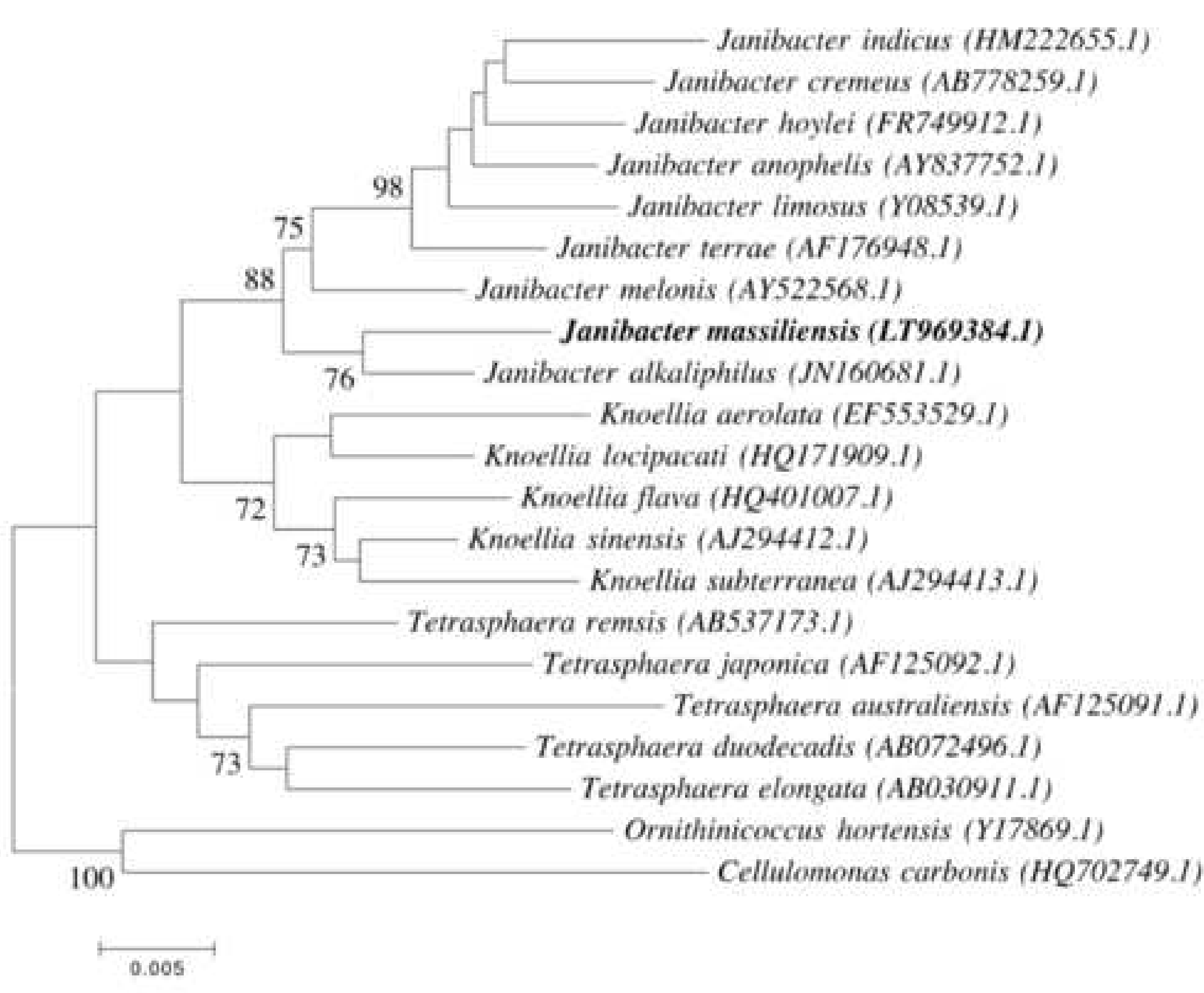
394 **Figure 1.** Phylogenetic analysis based on the comparison of 16S RNA gene sequences
395 highlighting the position of *Janibacter massiliensis* strain Marseille-P4121 among closely
396 related species. GenBank accession numbers are indicated in parentheses. Sequences were
397 aligned using ClustalW with default parameters and phylogenetic inferences were obtained
398 using the neighbor-joining method with the MEGA software version 7. The scale bar
399 represents a 0.5% nucleotide sequence divergence**.**

400 **Figure 2.** Transmission electron microscopy of *Janibacter massiliensis* strain Marseille-
401 P4121 using a Tecnai G20 transmission electron microscope (FEI Company). The scale bars
402 represent 500 (left) and 200 (right) nm, respectively.

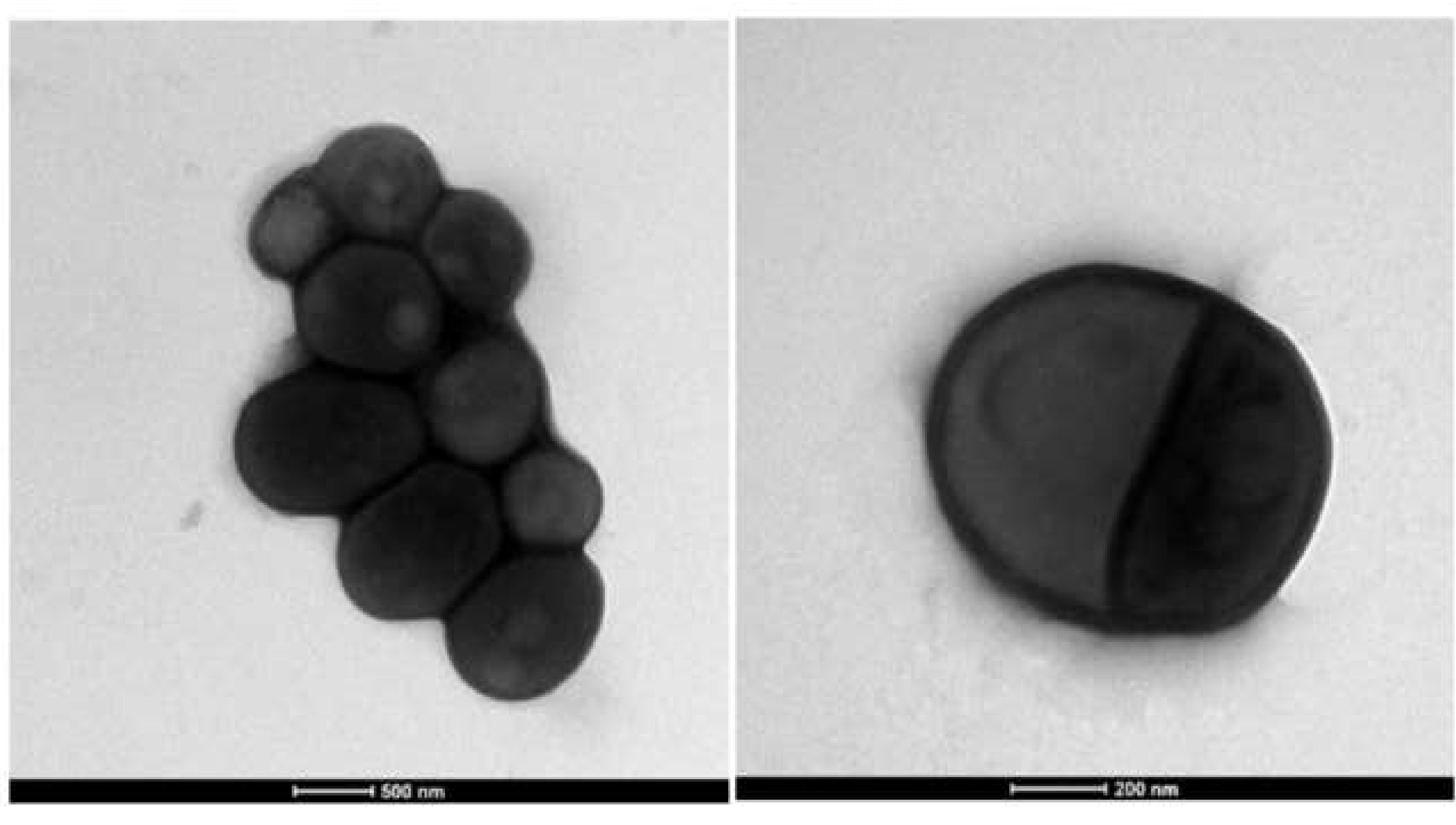
403 **Figure 3**. Graphical circular map of the chromosome from strain Marseille-P4121.
404 From the outside in: open reading frames oriented in the forward (colored by COG categories)
405 direction, open reading frames oriented in the reverse (colored by COG categories) direction,
406 RNA operon (red), and tRNAs (green), GC content plot, and GC skew (purple: negative
407 values, olive: positive values).

408 **Figure 4**: Distribution of predicted genes of *Janibacter massiliensis* and 7 closely related
409 species into COG categories.
410

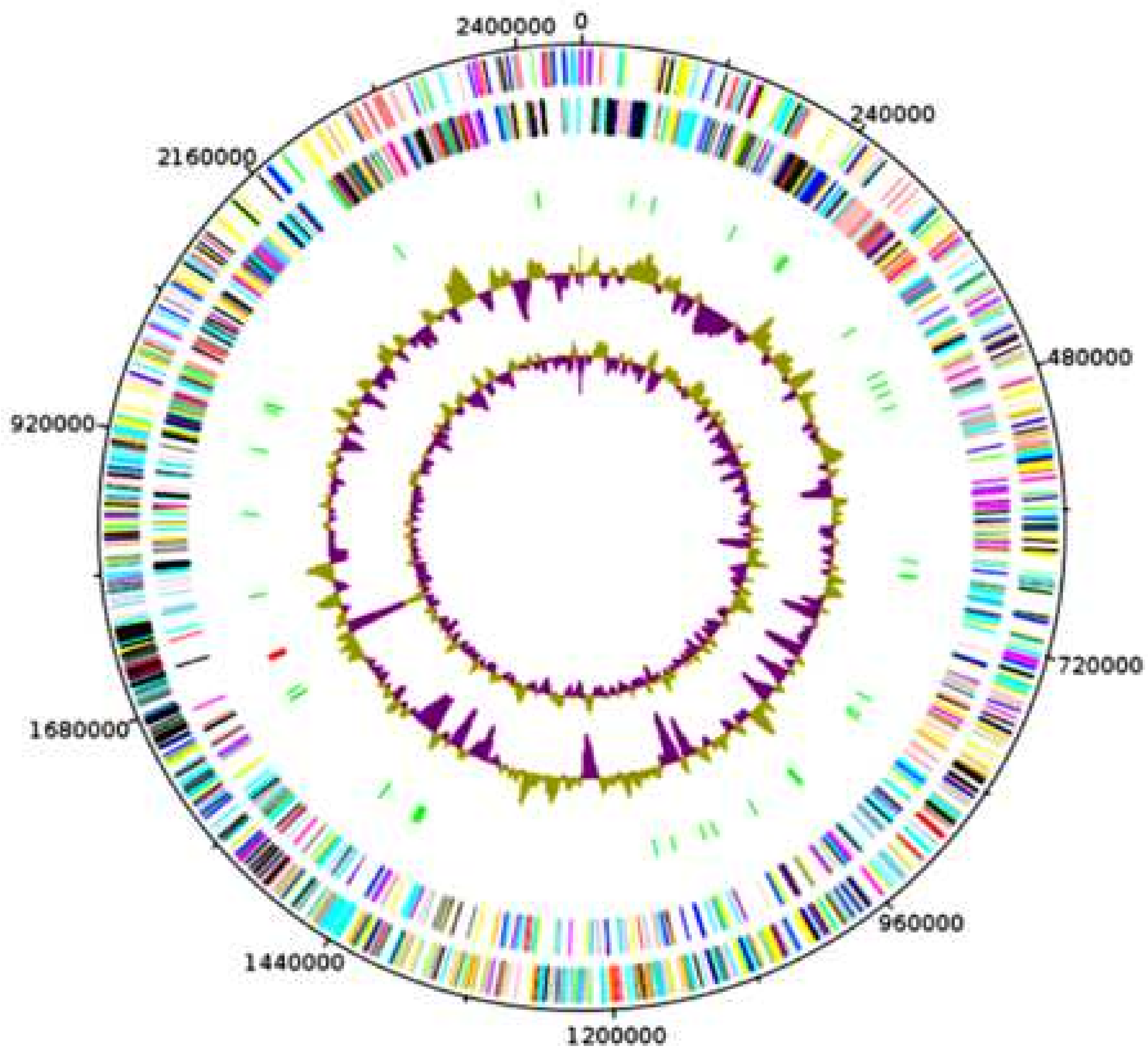
Click here to access/download;Figure;Figure 1.tif



Click here to access/download;Figure;Figure 2.tif



Click here to access/download;Figure;Figure 3.tif



Click here to access/download;Figure;figure 4.tif

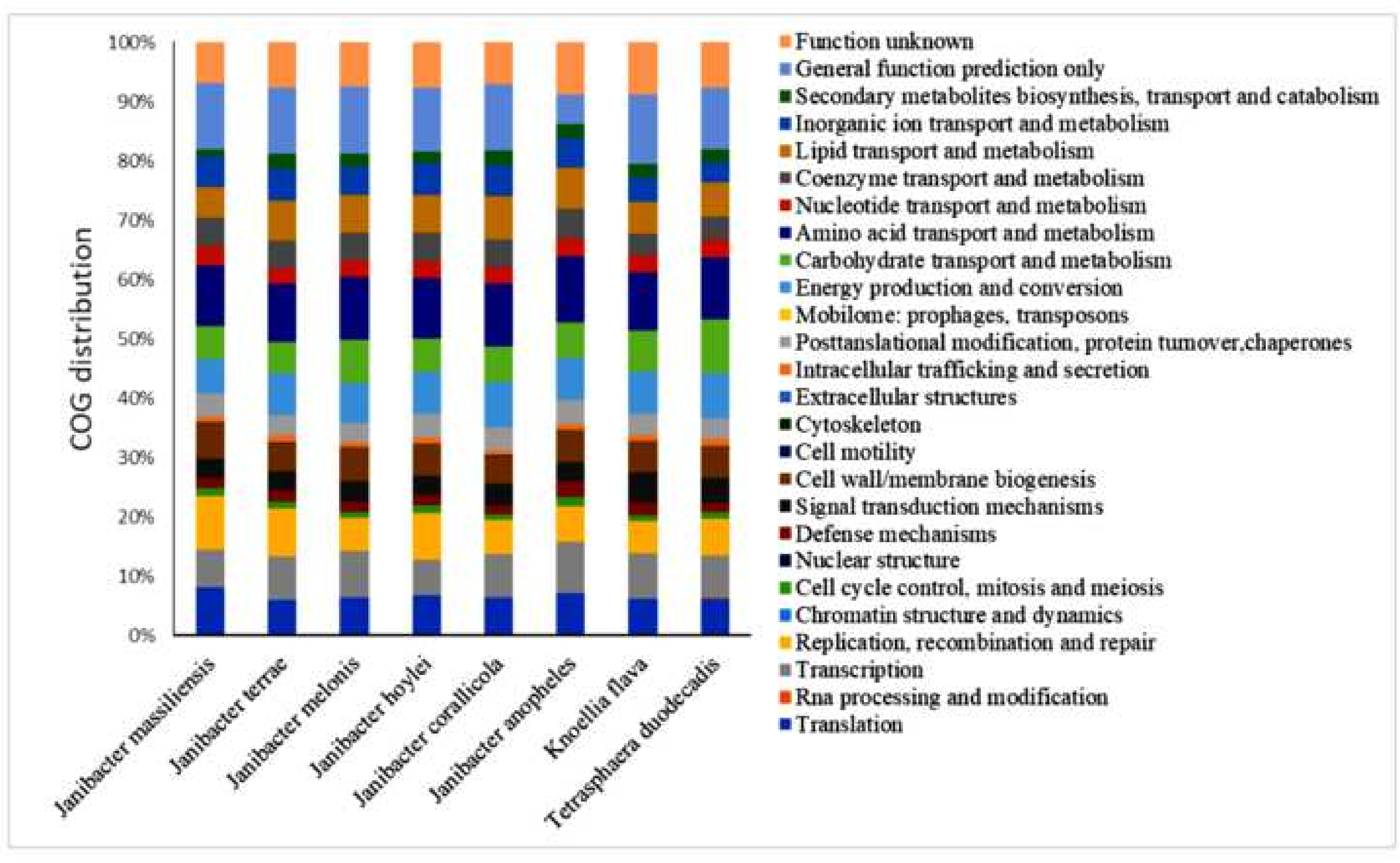


table Click here to access/download;table;Tables Janibacter.docx

1 **Table 1**. Fatty acid profiles of strain Marseille-P4121.

Fatty acids	**Name**	**Mean relative %** [a]
$C_{18\omega19}$	9-Octadecenoic acid	34.4 ± 1.0
$C_{16:0}$	Hexadecanoic acid	30.1 ± 1.0
$C_{18:0}$	Octadecanoic acid	19.0 ± 2.3
$C_{18:2\omega6}$	9,12-Octadecadienoic acid	11.8 ± 1.9
iso-$C_{17:0}$	15-methyl-Hexadecanoic acid	1.0 ± 1.5
$C_{14:0}$	Tetradecanoic acid	TR
$C_{17:0}$	Heptadecanoic acid	TR
anteiso-$C_{17:0}$	14-methyl-Hexadecanoic acid	TR
$C_{17:1\omega7}$	10-Heptadecenoic acid	TR
$C_{20:4\omega6}$	5,8,11,14-Eicosatetraenoic acid	TR
$C_{16:1\omega7}$	9-Hexadecenoic acid	TR
$C_{15:0}$	Pentadecanoic acid	TR
iso-$C_{16:0}$	14-methyl-Pentadecanoic acid	TR
anteiso-$C_{15:0}$	12-methyl-tetradecanoic acid	TR
iso-$C_{15:0}$	13-methyl-tetradecanoic acid	TR

2 TR: traces amount <1%.

Table 2. Differential Phenotypic characteristics of strain Marseille-P4121 and closely related species

Character	*Janibacter massiliensis*	*Janibacter alkaliphilus*	*Janibacter anopheles*	*Janibacter corallicola*	*Janibacter hoylei*	*Janibacter melonis*	*Janibacter terrae*	*Knoellia flava*	*Tetrasphaera duodecadis*
Cell diameter	0.7-1	na	1.0– 1.5	0.6-1.1	0.4–0.7	0.8–1.0	0.6–1.1	0.5–1.1	0.4–0.6
Indole production	-	na	na	na	-	-	-	-	-
Nitrate reduction	-	+	-	na	+	+	+	+	+
Hydrolysis of									
Aesculin	-	na	-	-	-	+	-	+	+
Gelatin	-	-	+	+	+	+	D	+	+
Enzymes activities									
Oxidase	-	-	+	na	+	-	-	-	-
Catalase	+	+	+	+	+	+	+	+	+
β-galactosidase	-	na	-	-	+	na	-	+	+
Acid from									
D-glucose	+	-	+	+	-	-	+	+	+
Lactose	+	na	-	na	-	-	-	na	na
Maltose	-	-	+	+	-	-	D	+	-
D-Mannitol	-	-	-	-	-	-	-	+	+
D-Mannose	-	-	+	+	-	-	+	+	-
Habitat	Human vagina	Coral	Midgut	Coral	Cryotubes	Oriental melon	Wastewater	Pig manure	Arable soil
Fatty acids	$C_{18:1\omega9}$ $C_{16:0}$ $C_{18:0}$	iso-$C_{17:0}$ iso-$C_{15:0}$ iso-$C_{16:0}$	$C_{16:0}$ $C_{17:1\omega8c}$ $C_{17:0}$	iso-$C_{16:0}$ $C_{17:1\omega8c}$ $C_{17:0}$	iso-C16:0 10-methyl-$C_{17:0}$ $C_{18:1\omega9c}$	iso-$C_{16:0}$ $C_{17:1\omega8c}$ $C_{18:1\omega9c}$	iso-$C_{16:0}$ $C_{18:1\omega9c}$ anteiso-$C_{17:0}$	iso-$C_{16:0}$ iso-$C_{15:0}$ $C_{17:1\omega8c}$	10-methyl-$C_{17:0}$ iso-$C_{16:0}$ iso-$C_{15:0}$
Source	Data are from this paper	(Li et al. 2012)	(Kampfer 2006)	(Kageyama et al. 2007)	(Shivaji et al. 2009)	(Yoon 2004)	(Lang et al. 2003)	(Yu et al. 2012)	(Ishikawa and Yokota 2006)

na: not available data.

6 **Table 3**. Nucleotide content and gene count levels of the genome from strain Marseille-P4121

Attribute	Value	% of total
Size (bp)	2,452,608	100%
G+C content (bp)	1,778,140	72,50%
Coding region (bp)	2,267,623	85,57%
Total genes	2,400	100%
RNA genes	49	2,04%
Protein-coding genes	2,351	97,95%
Genes with function prediction	1175	48,85%
Genes assigned to COGs	1,685	75,06%
Genes with peptide signals	142	7,18%
Genes with transmembrane helices	536	28,31%
ORFANS	447	18,62%

7

8 **Table 4**. Number of genes from strain Marseille-P4121 associated with the 25 general COG

9 functional categories

Code	Value	% value[a]	Description
J	136	5,7	Translation
A	1	_	Rna processing and modification
K	105	4,4	Transcription
L	154	6,4	Replication, recombination and repair
B	1	_	Chromatin structure and dynamics
D	21	0,9	Cell cycle control, mitosis and meiosis
Y	0	_	Nuclear structure
V	29	1,2	Defense mechanisms
T	56	2,3	Signal transduction mechanisms
M	102	4,3	Cell wall/membrane biogenesis
N	2	0,1	Cell motility
Z	0	_	Cytoskeleton
W	0	_	Extracellular structures
U	13	0,5	Intracellular trafficking and secretion
O	68	2,8	Posttanslational modification, protein turnover,chaperones
X	0	_	Mobilome: prophages, transposons
C	96	4	Energy production and conversion
G	94	3,9	Carbohydrate transport and metabolism
E	171	7,1	Amino acid transport and metabolism
F	55	2,3	Nucleotide transport and metabolism
H	81	3,4	Coenzyme transport and metabolism
I	90	3,8	Lipid transport and metabolism
P	87	3,6	Inorganic ion transport and metabolism
Q	20	0,8	Secondary metabolites biosynthesis, transport and catabolism
R	187	7,8	General function prediction only
S	116	4,8	Function unknown
_	715	29,8	Not in COGs

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

11 **Table 5**. Genomic comparison of *Janibacter massiliensis* with closely related species.

Species	Strain	INSDC identifier[a]	Size (Mb)	G+C (%mol)	Gene Content
Janibacter massiliensis	**Marseille-P4121**	**OPYK00000000**	**2.4**	**72.5**	**2,351**
Janibacter anophelis	NBRC 107843	BCSQ00000000	3.31	71.3	3,253
Janibacter corallicola	NBRC 107790	BCSR00000000	3.11	71	3,025
Janibacter hoylei	PVAS-1	ALWX00000000	3.14	72.8	3,071
Janibacter melonis	CD11-4	LQZG00000000	3.2	73.0	3,015
Janibacter terrae	NBRC 107853	BCUV00000000	3.58	69	3,494
Knoellia flava	TL1	AVPI00000000	3.56	70.9	3,388
Tetrasphaera duodecadis	DSM 12806	PJNE01000000	3.50	73	3,178

12 [a] INSDC: International Nucleotide Sequence Database Collaboration.

Table 6. dDDH values obtained by pairwise comparison of all studied genomes

	J. massiliensis	*J. anophelis*	*J. corallicola*	*J. hoylei*	*J. melonis*	*J. terrae*	*K. flava*	*T. duodecadis*
Janibacter massiliensis	100	20.00 ±2.30	19.50 ±2.25	20.10 ±2.30	20.20 ±2.30	20.20 ±2.30	20.20 ±2.30	19.90 ±2.30
Janibacter anophelis		100	22.30 ±2.35	28.10±2.40	20.70 ±2.30	27.80 ±2.40	20.10 ±2.30	19.80 ±2.30
Janibacter corallicola			100	22.10±2.35	20.00 ±2.30	22.10 ±2.35	19.70 ±2.30	19.70 ±2.30
Janibacter hoylei				100	21.00±2.35	27.60 ±2.40	20.20 ±2.35	19.70 ±2.30
Janibacter melonis					100	20.50 ±2.30	20.10 ±2.30	19.90 ±2.30
Janibacter terrae						100	20.20 ±2.30	19.50 ±2.30
Knoellia flava							100	20.50±2.35
Tetrasphaera duodecadis								100

Table 7. Pairwise genomic comparison of *Janibacter massiliensis* with closely related species

	J. massiliensis	*J. anophelis*	*J. corallicola*	*J. hoylei*	*J. melonis*	*J. terrae*	*K. flava*	*T. duodecadis*
Janibacter massiliensis	**2356**	1341	1236	1214	1330	1371	1239	1195
Janibacter anophelis	75.50	**3218**	1888	1872	1877	2149	1608	1491
Janibacter corallicola	74.88	80.46	**3016**	1632	1696	1878	1478	1407
Janibacter hoylei	75.47	85.17	79.95	**3173**	1617	1858	1407	1307
Janibacter melonis	76.34	77.76	76.76	77.98	**3041**	1901	1631	1525
Janibacter terrae	75.71	85.24	80.34	84.54	78.05	**3529**	1621	1528
Knoellia flava	75.23	74.76	73.84	74.83	75.57	74.95	**3352**	1546
Tetrasphaera duodecadis	75.14	74.52	74.01	74.35	75.23	74.63	77.37	**3208**

Numbers of orthologous proteins shared between genomes (upper right), AGIOS values (lower left) and numbers of proteins per genome (bold numbers).

Supplementary Table S1. Human infections caused by *Janibacter* species

Species	Number of cases	Diagnosis	Clinical source	Treatments	References
Janibacter terrae	5	Acute bronchitis	Blood	Amoxicillin/clavulanic	(Fernández-Natal et al. 2015)
		Multilobar pneumonia	Blood	Ceftazidime + amikacin + metronidazole	
		Cholangitis	Blood	Amoxicillin/clavulanic + tobramycin	
		Myelodysplastic syndrome	Blood	Piperacillin/tazobactam	
		Psoas abscess	Blood	meropenem, vancomycin and fluconazole	(Wan et al. 2017)
Janibacter melonis	2	Low-grade fever	Blood	Intravenous cefazolin	(Elsayed and Zhang 2005)
		Celiac disease with gastrointestinal symptoms	Duodenal mucosa	na	(Chander et al. 2018)
Janibacter hoylei	1	Irritability, reduced appetite, and fever	Blood	Vancomycin	(Lim et al. 2017)
Janibacter sp.	1	Acute myeloid leukemia	Blood	na	(Loubinoux et al. 2005)
Janibacter massiliensis	1	Bacterial vaginosis	Vagina smears	Clindamycin	Data are from this paper

Article 8:

Microbial Culturomics Broadens Human Vaginal Flora Diversity: Genome sequence and Description of *Prevotella lascolaii* sp. nov. Isolated from a Patient with bacterial Vaginosis.

Khoudia Diop, Awa Diop, Anthony Levasseur, Oleg Mediannikov, Catherine Robert, Nicholas Armstrong, Carine Couderc, Florence Bretelle, Didier Raoult, Pierre-Edouard Fournier, Florence Fenollar

Published in OMICS Journal of Integrative Biology

OMICS A Journal of Integrative Biology
Volume 22, Number 3, 2018
© Mary Ann Liebert, Inc.
DOI: 10.1089/omi.2017.0151

Research Article

Microbial Culturomics Broadens Human Vaginal Flora Diversity: Genome Sequence and Description of *Prevotella lascolaii* sp. nov. Isolated from a Patient with Bacterial Vaginosis

Khoudia Diop,[1] Awa Diop,[1] Anthony Levasseur,[1] Oleg Mediannikov,[1] Catherine Robert,[1] Nicholas Armstrong,[1] Carine Couderc,[1] Florence Bretelle,[2] Didier Raoult,[1,3] Pierre-Edouard Fournier,[1] and Florence Fenollar[1]

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

Abstract

Microbial culturomics is a new subfield of postgenomic medicine and omics biotechnology application that has broadened our awareness on bacterial diversity of the human microbiome, including the human vaginal flora bacterial diversity. Using culturomics, a new obligate anaerobic Gram-stain-negative rod-shaped bacterium designated strain khD1^T was isolated in the vagina of a patient with bacterial vaginosis and characterized using taxonogenomics. The most abundant cellular fatty acids were $C_{15:0}$ anteiso (36%), $C_{16:0}$ (19%), and $C_{15:0}$ iso (10%). Based on an analysis of the full-length 16S rRNA gene sequences, phylogenetic analysis showed that the strain khD1^T exhibited 90% sequence similarity with *Prevotella loescheii*, the phylogenetically closest validated *Prevotella* species. With 3,763,057 bp length, the genome of strain khD1^T contained (mol%) 48.7 G + C and 3248 predicted genes, including 3194 protein-coding and 54 RNA genes. Given the phenotypical and biochemical characteristic results as well as genome sequencing, strain khD1^T is considered to represent a novel species within the genus *Prevotella*, for which the name *Prevotella lascolaii* sp. nov. is proposed. The type strain is khD1^T (= CSUR P0109, = DSM 101754). These results show that microbial culturomics greatly improves the characterization of the human microbiome repertoire by isolating potential putative new species. Further studies will certainly clarify the microbial mechanisms of pathogenesis of these new microbes and their role in health and disease. Microbial culturomics is an important new addition to the diagnostic medicine toolbox and warrants attention in future medical, global health, and integrative biology postgraduate teaching curricula.

Keywords: culturomics, taxonogenomics, *Prevotella lascolaii*, bacterial vaginosis, microbiome science

Introduction

The symbiotic relationship between humans and their associated bacteria plays a crucial role in their health. Changes in the proportion of microbial species in the vagina predispose that person to dysbioses such as bacterial vaginosis (BV) (Narayankhedkar et al., 2015). First studies using traditional culture methods identified only 20% of bacteria present in the vagina (Lamont et al., 2011). The vaginal flora diversity has been revealed further using molecular methods, sequencing, and phylogenetic analysis of the 16S rRNA gene, which show the detection of fastidious and uncultured bacteria, such as bacterial vaginosis-associated bacteria type 1 (BVAB1), BVAB2, and BVAB3 (Fredricks et al., 2005).

Recently, a new approach named ''Microbial Culturomics,'' involving high-throughput culture conditions and matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) for bacterial identification, was initiated and used to study the human microbiota (Dubourg et al., 2013; Lagier et al., 2012). Culturomics broadened our awareness about the bacterial diversity of the human microbiome by analyzing different samples (such as stool, small-bowel, and colonic samples) from healthy individuals and patients with various diseases (such as anorexia nervosa, obesity, malnutrition,

[1]Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM 63, CNRS UMR 7278, IRD 198, INSERM U1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de Médecine, Aix-Marseille University, Marseille, France.
[2]Department of Gynecology and Obstetrics, Gynépole, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, Marseille, France.
[3]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia.

and HIV) from different geographical origins (Europe, rural and urban Africa, Polynesia, India, …) (Lagier et al., 2016).

In addition to improving culture and bacterial identification, culturomics is used with a new classification and nomenclature concept called taxonogenomics to better characterize and describe bacterial species (Fournier and Drancourt, 2015; Fournier et al., 2015). Taxogenomics combines classic bacterial description and phenotypic/genotypic characteristics such as DNA-DNA hybridization with the proteomic information obtained by MALDI-TOF mass spectrometry (MS) and the description of the complete genome.

We isolated a new member of the genus *Prevotella* in a culturomics study of the vaginal flora, which aimed to map the vaginal microbiome in healthy women and patients with BV to identify bacteria involved in this dysbiosis. Amended in 2012 (Sakamoto and Ohkuma, 2012), the *Prevotella* genus was created in 1990 by reclassifying some *Bacteroides* species. It contains gram-negative rod, strict anaerobic, nonspore forming, and nonmotile bacteria with *Prevotella melaninogenica* as the type strain (Shah and Collins, 1990).

Here follows the description of *Prevotella lascolaii* strain khD1^T (=CSUR P0109, =DSM 101754) with its annotated whole genome, isolated in the vaginal flora of a patient suffering from BV.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

Materials and Methods

Ethics and sample collection

The vaginal sample of a 33-year-old French woman was collected at Nord Hospital in Marseille (France) in October 2015 using a Sigma Transwab (Medical Wire, Corsham, United Kingdom). As previously described (Menard et al., 2008), the patient was suffering from abnormal vaginal discharge and diagnosed with BV. During the sample collection, she was not treated with any antibiotics and she signed a written consent. The local ethics committee of the IFR48 (Marseille, France) had also authorized this study under agreement number 09-022.

Strain identification by MALDI-TOF MS

Initially, the vaginal sample was inoculated in a blood culture bottle (BD Diagnostics, Le Pont-de-Claix, France) enriched with sheep's blood (bioMérieux, Marcy l'Etoile, France) and rumen filtered at 0.2 μm by a pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France). Fifty microliters of the supernatant was plated onto Schaedler agar enriched with vitamin K and sheep's blood (BD Diagnostics). Then, after 4 days of incubation at 37°C in anaerobic conditions, purified colonies were deposited on an MALDI-TOF target plate (Bruker Daltonics, Leipzig, Germany) in duplicate and, as previously described, 1.5 μL of matrix (Seck et al., 2015) was added on each spot.

The identification was carried out using a Microflex spectrometer (Bruker) (Seng et al., 2009), which compares identified protein spectra to those on the MALDI-TOF database containing 7567 references (composed of the Bruker database incremented with our data). The reliability of bacteria identification was indicated by a score. If the score was greater than 1.9, the bacterium was considered identified. Conversely, if the score was less than 1.9 it was not in the database or identification failed.

Strain identification by 16S rRNA sequencing

To identify unidentified bacterium, the 16S rRNA gene was sequenced using fD1-rP2 primers (Eurogentec, Angers, France). The obtained sequence was corrected using ChromasPro 1.34 software (Technelysium Pty. Ltd., Tewantin, Australia) and matched against the NCBI database using the BLAST algorithm (Drancourt et al., 2000).

Phylogenetic tree

All species from the same genus of the new species were retrieved and 16S sequences were downloaded from NCBI. Sequences were aligned using CLUSTALW, with default parameters and phylogenetic inferences obtained using the neighbor-joining method with 500 bootstrap replicates, using MEGA6 software.

Growth conditions

To evaluate ideal growth, the strain khD1^T was cultivated on Columbia agar with 5% sheep's blood and incubated at different temperatures (25°C, 28°C, 37°C, 45°C, and 56°C) in an aerobic atmosphere with or without 5% CO_2 and also in anaerobic and microaerophilic atmospheres using GENbag anaer and GENbag miroaer systems (bioMérieux), respectively.

Morphology

To observe cell morphology, cells were fixed with 2.5% glutaraldehyde in a 0.1 M cacodylate buffer for at least an hour at 4°C. One drop of cell suspension was deposited for $\sim$5 min on glow-discharged formvar carbon film on 400 mesh nickel grids (FCF400-Ni; EMS). The grids were dried on blotting paper and the cells were negatively stained for 10 sec with 1% ammonium molybdate solution in filtered water at RT. Electron micrographs were acquired using a Tecnai G20 Cryo (FEI) transmission electron microscope operated at 200 keV. Gram staining, motility, and sporulation were performed as previously conducted (Murray et al., 2007).

Biochemical analysis

The biochemical characteristics of strain khD1^T have been determined using the API ZYM, 20A, and 50CH strips (bioMérieux) according to the manufacturer's instructions. API ZYM was performed for the research of enzymatic activities. It allows the systematic and rapid study of 19 enzymatic reactions using very small sample quantities. While API 20A (20 cupules) was used for the biochemical identification of the isolate and 50CH API (50 cupules) to study carbohydrate metabolism.

Cellular fatty acid methyl ester (FAME) analysis was performed using gas chromatography/mass spectrometry (GC/MS). Two samples were prepared with $\sim$35 mg of bacterial biomass per tube harvested from several culture plates. FAMEs were prepared as described by Sasser (Sasser, 2006). First, fatty acids were released from lipids with a saponification step at 100°C during 30 min in the presence of 1 mL NaOH 3.75 M in water/methanol (50% v:v). Then, free fatty acids were transformed to methyl esters at 80°C during 10 min after adding 2 mL of HCl 6 N/methanol (54/46% v:v). The resulting FAMEs were then extracted in 1 mL of hexane/MTBE (50% v:v). Organic extracts were finally washed with 3 mL of NaOH 0.3 M to

remove free acids. GC/MS analyses were carried out using a Clarus 500 gas chromatograph connected to a SQ8S single quadrupole MS detector (Perkin Elmer, Courtaboeuf, France).

Two microliters of both FAME extracts were volatized at 250°C (split 20 mL/min) in a Focus liner with wool. Compounds were separated on an Elite-5MS column (30 m, 0.25 mm i.d., 0.25 mm film thickness) using a linear temperature gradient (70–290°C at 6°C/min) enabling the detection of C4 to C24 FAMEs. Helium flowing at 1.2 mL/min was used as carrier gas. MS inlet line was set at 250°C and electron ionization source at 200°C. Full scan monitoring was performed from 45 to 500 m/z. All data were collected and processed using Turbomass 6.1 (Perkin Elmer).

FAMEs were identified using the identity spectrum search using the MS Search 2.0 software, operated with the Standard Reference Database 1A (NIST, Gaithersburg, USA) and the FAME mass spectral database (Wiley, Chichester, United Kingdom). A 37-component FAME mix (Supelco; Sigma-Aldrich, Saint-Quentin Fallavier, France) was used to calculate the correlation between chromatographic retention times and nonpolar retention indexes from the NIST database. MS Search identifications were therefore validated if reverse/forward search scores were above 750 and if nonpolar retention indexes were correlated to the chromatographic retention time.

Antibiotic susceptibility tests

Amoxicillin, benzylpenicillin, imipenem, metronidazole, and vancomycin were used to test antibiotic susceptibility of strain khD1^T. The minimal inhibitory concentrations (MICs) were then determined using E-test gradient strips (bioMérieux) according to the EUCAST recommendations (Citron et al., 1991; Matuschek et al., 2014).

Genomic DNA preparation

Strain khD1^T was cultured on 5% sheep's blood-enriched Columbia agar (bioMérieux) at 37°C anaerobically. Bacteria grown on three Petri dishes were resuspended in $4\times100\,\mu$L of Tris–EDTA (TE) buffer. Next, 200 μL of this suspension was diluted in 1 mL TE buffer for lysis treatment, which included a 30-min incubation with 2.5 μg/μL lysozyme at 37°C, followed by an overnight incubation with 20 μg/μL proteinase K at 37°C. Extracted DNA was then purified using

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

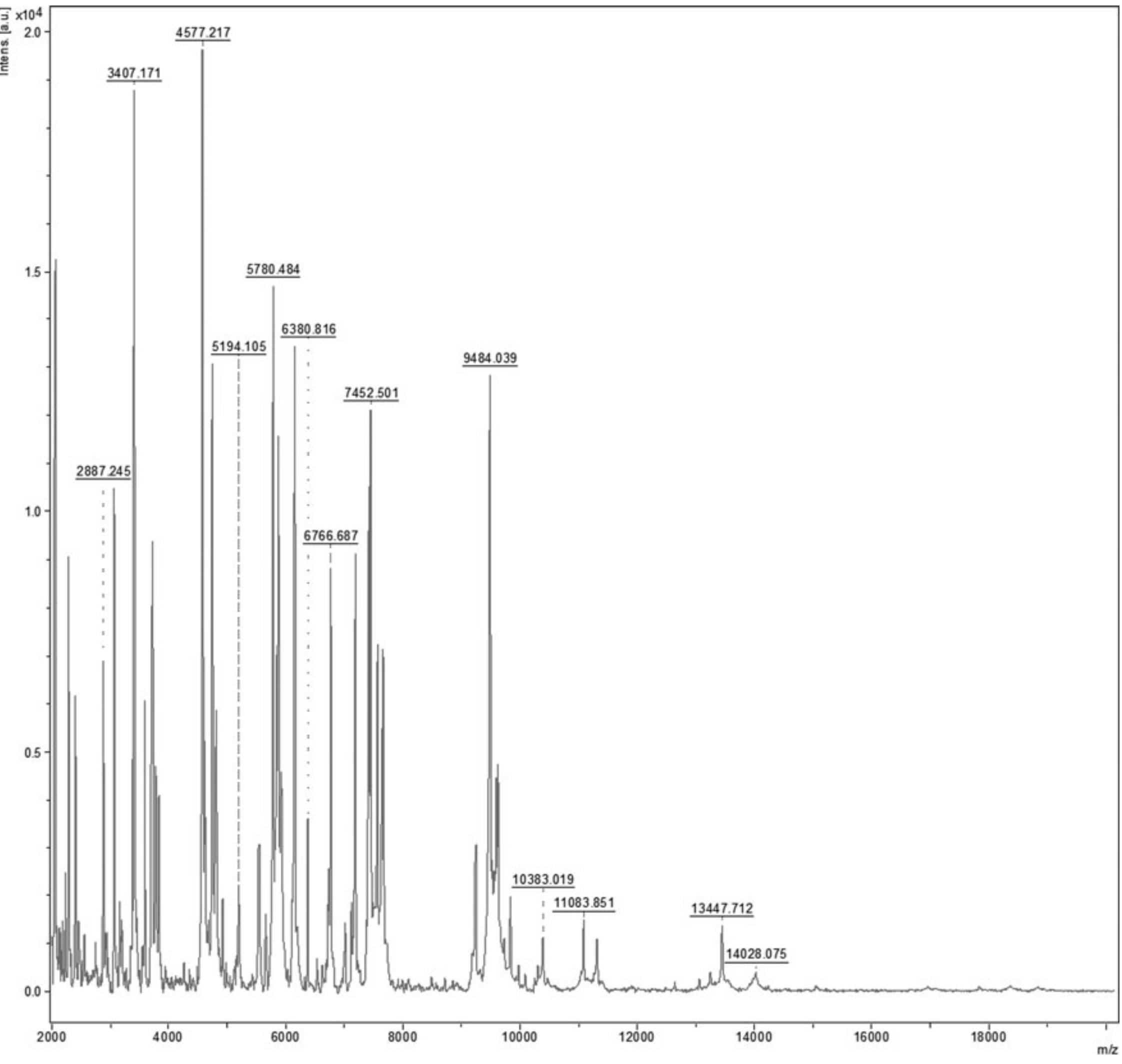


FIG. 1. Reference mass spectrum from the *Prevotella lascolaii* strain khD1^T.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

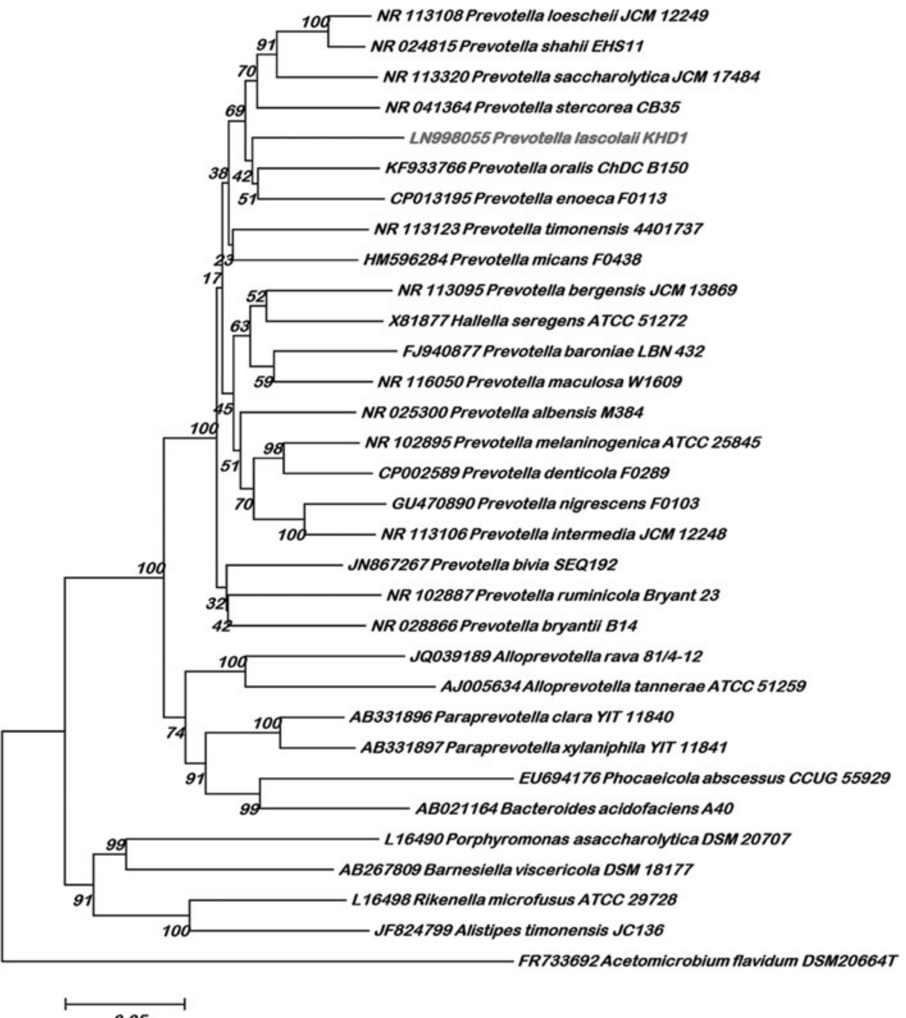


FIG. 2. Phylogenetic tree highlighting the position of *Prevotella lascolaii* strain khD1^{T} relative to other close strains. GenBank accession numbers of each 16S rRNA are noted before the name. Sequences were aligned using Muscle v3.8.31 with default parameters, and phylogenetic inferences were obtained using the neighbor-joining method with 500 bootstrap replicates, within MEGA6 software. The scale bar represents a 0.05% nucleotide sequence divergence.

three successive phenol–chloroform extractions and ethanol precipitations at –20°C overnight. After centrifugation, the DNA was resuspended in 160 μL TE buffer.

Genome sequencing and assembly

Genomic DNA (gDNA) of strain khD1^{T} was sequenced on the MiSeq Technology (Illumina, Inc., San Diego, CA, USA) with the mate-pair strategy. The gDNA was barcoded with the Nextera Mate-Pair sample prep kit (Illumina) to be mixed with 11 other projects.

gDNA was quantified by a Qubit assay with a high-sensitivity kit (Life technologies, Carlsbad, CA, USA) to 105.7 ng/μL. The mate-pair library was prepared with 1.5 μg of genomic DNA using the Nextera mate-pair Illumina guide. The genomic DNA sample was simultaneously fragmented and tagged with a mate-pair junction adapter. The pattern of fragmentation was validated on an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA) with a

Table 1. Classification and General Features of *Prevotella lascolaii* Strain khD1^{T}

	Term
Current classification	Domain: *Bacteria* Phylum: *Bacteroidetes* Class: *Bacteroidia* Order: *Bacteroidales* Family: *Prevotellaceae* Genus: *Prevotella* Species: *Prevotella lascolaii* Type strain: khD1
Gram stain	Negative
Cell shape	Rod
Motility	Nonmotile
Sporulation	Nonsporulating
Temperature range	Anaerobic
Optimum temperature	37°C

DNA 7500 LabChip. The DNA fragments ranged in size from 1.5 to 11 kb with an optimal size at 5.203 kb. No size selection was performed and 440 ng of tagmented fragments were circularized.

The circularized DNA was mechanically sheared to small fragments with an optimal size of 985 bp on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was viewed on a High-Sensitivity Bioanalyzer LabChip (Agilent Technologies, Inc., Santa Clara, CA, USA) and the final concentration library was measured at 4.17 nM.

The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 39-h run in a 2×151 bp.

Total information of 8.8 Gb was obtained from a 971 K/mm^2 cluster density with a cluster passing quality control filters of 93.1% (17,376,000 passing filter paired reads). Within this run, the index representation for strain khD1^T was determined to be 7.17%. The 1,246,384 paired reads were trimmed and then assembled in 27 scaffolds.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

Genome annotation and analysis

Open reading frames (ORFs) were predicted using Prodigal software (Hyatt et al., 2010) with default parameters. Predicted ORFs spanning a sequencing gap region (containing N) were excluded. We predicted the bacterial proteome using BLASTP (E-value 1e-03 coverage 0.7 and identity percent 30) against the Clusters of Orthologous Groups (COGs) database. A search against the NR database (Clark et al., 2016) was performed if no hit was found, using BLASTP with E-value of 1e-03 coverage 0.7 and an identity percent of 30. An E-value of 1e-05 was used with sequence lengths smaller than 80 amino acids. The hhmscan tool analyses were used for searching PFam conserved domains (PFAM-A and PFAM-B domains) on each protein.

We used RNAmmer (Lagesen et al., 2007) and tRNAScanSE tools (Lowe and Eddy, 1997) to find ribosomal RNA genes and tRNA genes, respectively. Viewing and data managing genomic features were performed using Artemis (Carver et al., 2012) and DNA Plotter (Carver et al., 2009), respectively. For the mean level of nucleotide sequence similarity analysis at the genome level, we used the MAGI home-made software. It calculated the average genomic identity of gene sequences (AGIOS) among compared genomes (Ramasamy et al., 2014). The Proteinortho (Lechner et al., 2011) software was incorporated with the MAGI home-made software for detecting orthologous proteins in pair-wise genomic comparisons. Next, the corresponding genes were retrieved and the mean percentage of nucleotide sequence identity among orthologous ORFs was determined using the Needleman–Wunsch global alignment algorithm.

The Multi-Agent Software System DAGOBAH (Gouret et al., 2011) was used to perform annotation and comparison processes, which included Figenix (Gouret et al., 2005) libraries providing pipeline analysis. GGDC analysis was performed using the GGDC web server as previously reported (Meier-Kolthoff et al., 2013).

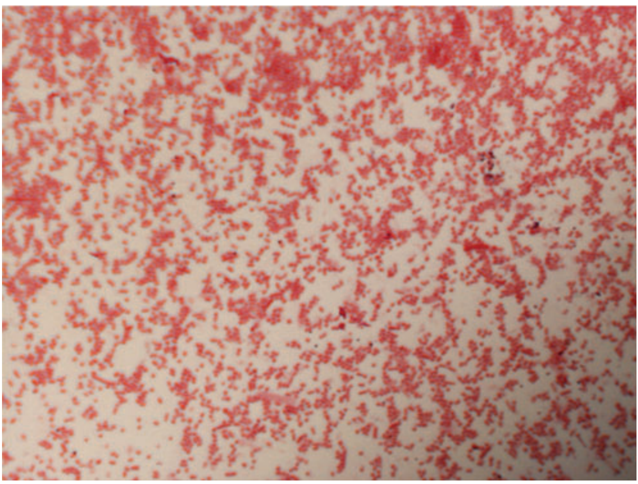


FIG. 3. Gram staining of *Prevotella lascolaii* strain khD1^T.

Results

Strain identification and phylogenetic analysis

P. lascolaii strain khD1^T was first isolated after 24 h preincubation of the vaginal sample in a blood culture bottle enriched with rumen, which was filter sterilized through a 0.2 μm pore filter (Thermo Fisher Scientific), and sheep's blood (bioMérieux) under anaerobic conditions at 37°C. Then, 50 μL of the supernatant was inoculated on Schaedler agar enriched with sheep's blood and vitamin K (BD Diagnostics) in the same conditions for 4 days. The MALDI-TOF identification gave us a score of 1.3. As the strain was not in the database, the reference spectrum (Fig. 1) was incremented in our database and the gene 16S rRNA was sequenced.

The sequence obtained (number accession LN998055) exhibited 90% similarity with *Prevotella loescheii*, the phylogenetically closest bacterial species with a validly published

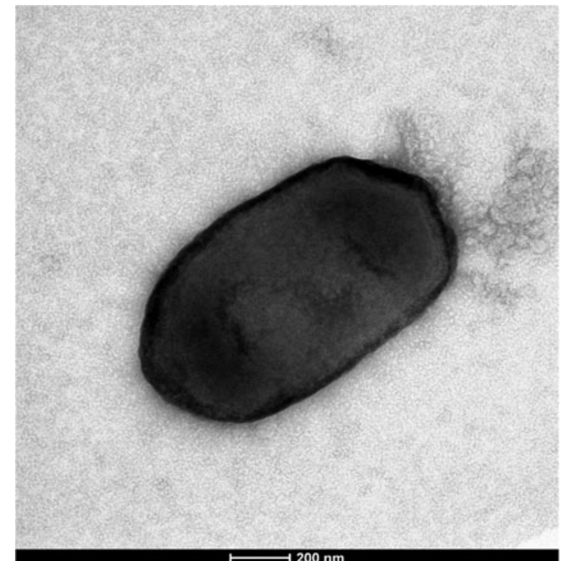


FIG. 4. Electron micrographs of *Prevotella lascolaii* strain khD1^T using a Tecnai G^{20} Cryo (FEI) transmission electron microscope operated at 200 keV. The scale bar represents 200 nm.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

TABLE 2. PHENOTYPIC CHARACTERISTICS THAT DIFFERENTIATE *PREVOTELLA LASCOLAII* STRAIN KHD1^T SP. NOV. FROM RELATED *PREVOTELLA* SPECIES, *P. LOESCHEII*, *P. SHAHII*, *P. ORALIS*, *P. STERCOREA*, *P. ENOECA*, *P. TIMONENSIS*, AND *P. MICANS*

Characteristic	Prevotella lascolaii	Prevotella loescheii	Prevotella shahii	Prevotella oralis	Prevotella stercorea	Prevotella enoeca	Prevotella timonensis	Prevotella micans
Cell diameter (μm)	0.3–0.5	0.4–0.6	0.5–0.8	0.5–1	0.25–0.42	0.5	0.8–1.4	0.7
Endospore formation	–	–	–	–	–	–	–	na
Indole	–	–	–	–	–	–	na	+
Production of								
Alkaline phosphatase	+	na	+	na	+	na	+	+
Catalase	–	–	–	–	–	–	na	–
Nitrate reductase	–	–	–	–	na	–	na	–
Urease	–	na	–	–	–	na	na	na
β-galactosidase	+	na	+	na	+	na	+	+
N-acetyl-glucosamine	+	na	+	na	+	na	+	+
Production of								
L-arabinose	+	–	–	–	–	–	+	–
Ribose	+	–	na	na	na	–	+	–
Mannose	–	+	+	+	+	+	–	+
Sucrose	–	+	+	+	+	–	+	+
D-glucose	–	+	+	+	+	+	+	+
D-fructose	–	+	na	+	na	+	na	+
D-maltose	–	+	+	+	+	+	+	+
D-lactose	–	+	+	+	+	+	+	+
Major cellular fatty acids[a]	$C_{15:0}$ anteiso, $C_{16:0}$, $C_{15:0}$ iso	$C_{15:0}$ anteiso, $C_{18:1n9c}$, $C_{15:0}$ iso	$C_{18:1n9c}$, $C_{16:0}$, $C_{16:0}$ 3-OH	$C_{16:0}$, $C_{18:1n9c}$, $C_{16:0}$ 3-OH, $C_{15:0}$ anteiso	$C_{18:1n9c}$, $C_{15:0}$ iso, $C_{15:0}$ anteiso	$C_{15:0}$ anteiso, $C_{16:0}$, $C_{16:0}$ 3-OH, $C_{15:0}$ iso	$C_{14:0}$, $C_{16:0}$, $C_{18:2\ n6,9c}$/ $C_{18:0}$	na
G+C content (mol%)	48.7	46.9	44.3	43.1	48.2	47	40.50	46
Habitat	Human vagina	Human oral cavity	Human oral cavity	Human oral cavity	Human feces	Human gingiva	Breast abscess	Human oral cavity

The reference for the species data comes from descriptions of the original species. +, –, and na data.
[a]Major cellular fatty acids listed in order of predominance.
+, positive; –, negative; na, not available.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

Table 3. Cellular Fatty Acid Composition (%) of *Prevotella lascolaii* Strain khD1^T (Data from This Study) Compared with Closest Species

Fatty acids	*Name*	Prevotella lascolaii	Prevotella loescheii	Prevotella shahii	Prevotella oralis	Prevotella stercorea	Prevotella enoeca	Prevotella timonensis
Saturated straight chain								
14:00	Tetradecanoic acid	1.5	1.1	10.9	2.1	0.8	4	**19.5**
15:00	Pentadecanoic acid	tr	3.8	1.0	tr	tr	na	na
16:00	Hexadecanoic acid	**18.8**	**12.5**	**16.9**	**19.2**	3.8	**17**	**15.3**
17:00	Heptadecanoic acid	tr	1.5	na	tr	na	na	na
18:00	Octadecanoic acid	tr	0.9	2.8	0.9	0.8	na	**16**
Unsaturated straight chain								
18:1n9	9-Octadecenoic acid	2.3	**15.0**	**18.7**	**18.6**	**14.7**	na	na
18:2n6	9,12-Octadecadienoic acid	4.0	2.0	na	na	2,2	na	**16**
20:4n6	5,8,11,14-Eicosatetraenoic acid	tr	na	na	na	na	na	na
Hydroxy acids								
16:0 3-OH	3-hydroxy-hexadecanoic acid	4.4	6.1	**16.3**	**10.4**	1	**10**	na
17:0 3-OH	3-hydroxy-heptadecanoic acid	7.7	na	na	na	na	na	na
18:0 3-OH	3-hydroxy-octadecanoic acid	tr	na	na	na	na	na	na
Saturated branched chain								
5:0 anteiso	2-methyl-butanoic acid	tr	na	na	na	na	na	na
14:0 iso	12-methyl-tridecanoic acid	1.5	2.1	4.4	3.0	2.7	3	**14**
15:0 iso	13-methyl-tetradecanoic acid	**9.9**	3.2	3.4	3.2	**23.7**	8	na
15:0 anteiso	12-methyl-tetradecanoic acid	**36.1**	**24.0**	6.8	**20.6**	**26.2**	**36**	na
16:0 iso	14-methyl-pentadecanoic acid	3.2	0.8	1.0	1.7	2.7	na	na
17:0 iso	15-methyl-hexadecanoic acid	4.8	1.1	na	tr	1.7	2	na
17:0 anteiso	14-methyl-hexadecanoic acid	4.3	1.7	na	1.5	1.3	na	na

Bold represents the majority fatty acid for this species; na, not available data; tr, trace amounts <1%. The reference for the species data comes from descriptions of the original species. *P. micans* was not listed because its complete fatty acid profile was not available.

Table 4. Nucleotide Content and Gene Count Levels of the Genome

Attribute	Value	% of total[a]
Size (bp)	3,763,057	100
G+C content (bp)	1,832,608	48.7
Coding region (bp)	3,186,418	84.67
Total genes	3248	100
RNA genes	54	1.60
Protein-coding genes	3194	98.33
Genes with function prediction	2034	63.68
Genes assigned to COGs	1691	52.9
Genes with peptide signals	643	20.13
Genes with transmembrane helices	2541	79.55

[a]The total is based on either the size of the genome in base pairs or the total number of protein-coding genes in the annotated genome.
COG, Clusters of Orthologous Group.

name (Fig. 2). Thus, as this value was under the threshold of 98.7%, established to delineate a new species (Kim et al., 2014; Stackebrandt and Ebers, 2006), strain khD1^T was classified as a new species within the *Prevotella* genus and named *P. lascolaii* (Table 1).

Phenotypic and biochemical characteristics

Cultivated on Columbia agar (bioMérieux) for 48 h in anaerobic conditions at 37°C, *P. lascolaii* strain khD1^T colonies were grayish-white, shiny, smooth, and circular with a diameter of 1.4 to 2 mm. Gram staining showed gram-negative short rod-shaped bacilli or coccobacilli (Fig. 3). Under electronic microscopy, individual cells had a mean diameter of 0.65 μm and mean length of 0.9 μm (Fig. 4). Nonmotile and nonspore-forming, *P. lascolaii* exhibited positive oxidase activity. Nevertheless, catalase activity was negative and nitrate was not reduced. Strictly anaerobic, strain khD1^T grows at temperatures between 25°C and 42°C, with optimal growth at 37°C after 48 h of incubation. Its growth also needs an NaCl concentration under 5 g/L and pH ranging from 6.5 to 8.5.

API ZYM strips revealed that strain khD1^T exhibited positive reactions for alkaline phosphatase, α-chymotrypsin, acid phosphatase, naphthol-AS-BI-phosphohydrolase, galactosidase (α and β), glucosidase (α and β), N-acetyl-β-glucosaminidase, and α-fucosidase enzymes. However, esterase, esterase lipase, lipase, leucine, cystine and valine arylamidase, trypsin, β-glucuronidase, and α-fucosidase were negative. API 50CH shows that strain khD1^T ferments arabinose, ribose, galactose, methyl-αD-mannopyranoside, β-galactosidase, melezitose, glycogen, turanose, tagose, and potassium 5-ketogluconate.

In contrast, arabinose, xylose, glucose, fructose, mannose, mannitol, cellobiose, maltose, lactose, sucrose, and starch were not metabolized. The same results were also observed using API 20A; ferric citrate esculin was hydrolyzed, but urease was not exhibited and carboxylates were not fermented. These

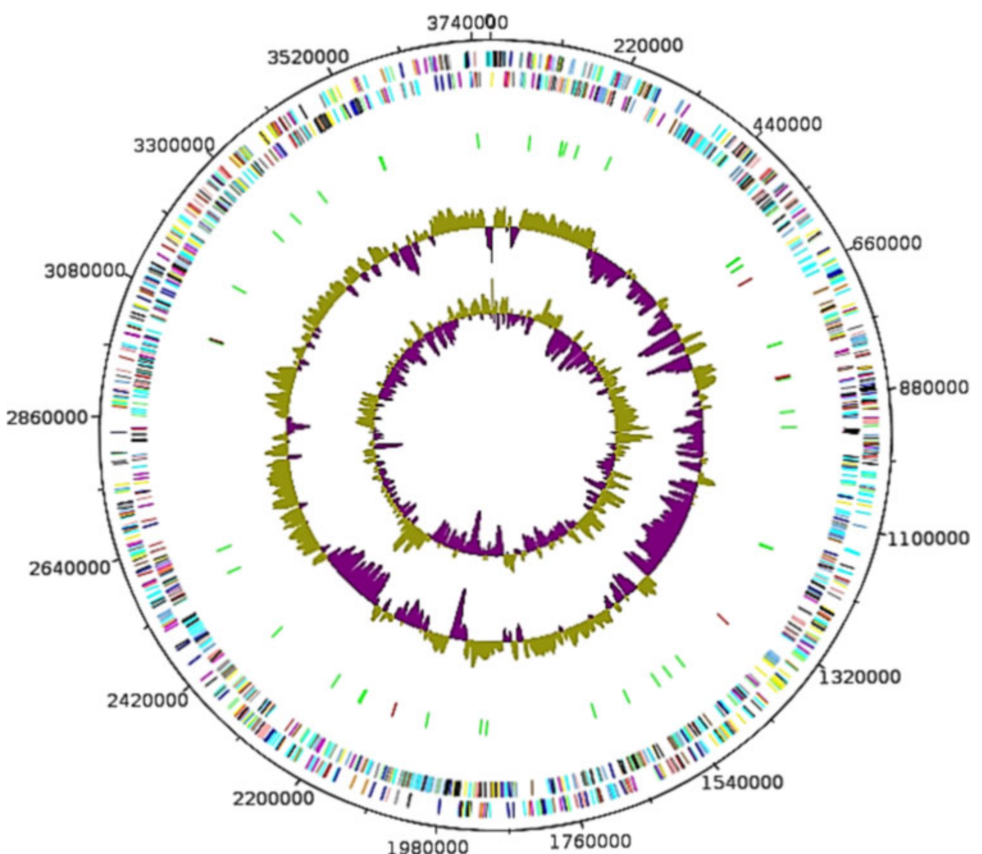


FIG. 5. Graphical circular map of the chromosome. From outside to the center: Contigs (*red/gray*), COG category of genes on the forward strand (three *circles*), genes on forward strand (*blue circle*), genes on the reverse strand (*red circle*), COG category on the reverse strand (three *circles*), GC content. COG, Clusters of Orthologous Group.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

Table 5. Number of Genes Associated with the 25 General Clusters of Orthologous Group Functional Categories

Code	Value	% value	Description
J	133	7.9	Translation
A	0	0	RNA processing and modification
K	88	5.2	Transcription
L	159	9.4	Replication, recombination, and repair
B	0	0	Chromatin structure and dynamics
D	25	1.5	Cell cycle control, mitosis, and meiosis
Y	0	0	Nuclear structure
V	53	3.1	Defense mechanisms
T	49	2.9	Signal transduction mechanisms
M	169	10.0	Cell wall/membrane biogenesis
N	4	0.2	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	31	1.8	Intracellular trafficking and secretion
O	68	4.0	Posttranslational modification, protein turnover, chaperones
X	14	0.8	Mobilome: prophages, transposons
C	83	4.9	Energy production and conversion
G	131	7.7	Carbohydrate transport and metabolism
E	114	6.7	Amino acid transport and metabolism
F	59	3.5	Nucleotide transport and metabolism
H	69	4. 1	Coenzyme transport and metabolism
I	46	2.7	Lipid transport and metabolism
P	77	4. 6	Inorganic ion transport and metabolism
Q	8	0.5	Secondary metabolite biosynthesis, transport, and catabolism
R	202	11.9	General function prediction only
S	109	6.4	Function unknown
—	1504	47.1	Not in COGs

phenotypic characteristics of *P. lascolaii* strain khD1^T are summarized in Table 2.

The major fatty acids of strain khD1T were similar to those found in members of *Prevotella* genus (Table 3) with saturated structures: 12-methyl-tetradecanoic acid (36%), hexadecanoic acid (19%), and 13-methyl-tetradecanoic acid (10%). Several branched structures and characteristic 3-hydroxy fatty acids were also described.

P. lascolaii khD1^T is sensitive to imipenem (MIC 0.47 μg/mL) and metronidazole (MIC 0.19 μg/mL) but resistant to amoxicillin (MIC >256 μg/mL), benzylpenicillin (MIC >256 μg/mL), and vancomycin (MIC 24 μg/mL).

Genome properties

The draft genome of *P. lascolaii* khD1^T (accession number FKKG00000000) is 3,763,057 bp long with 48.7% G+C content (Table 4). It contains 27 scaffolds assembled in 42 contigs (Fig. 5). Of the 3248 predicted genes, 3194 were protein-coding genes and 54 were RNAs (4 genes were 5S rRNA, 1 gene was 16S rRNA, 1 gene was 23S rRNA, and 47 genes were tRNA genes). A total of 2034 genes (63.68%) were assigned as putative functions (by cogs or NR blast). Two hundred twelve genes were identified as ORFans (6.63%). The remaining genes were annotated as hypothetical proteins (897 genes=>27.52%). Genome statistics is summarized in Table 4 and the distribution of the genes in COG functional categories is presented in Table 5.

Genomic comparison

The genome comparison of *P. lascolaii* strain khD1^T with the closest related species of *Prevotella* genus (Table 6) shows that the draft genome sequence of our strain (3.76 Mbp) was bigger than those of *Prevotella enoeca* and *Prevotella micans* (2.86 and 2.43 Mbp, respectively) but smaller than those of *P. loescheii* (7.01 Mbp). The G+C content of strain khD1^T (48.7 mol%) is larger than those of all the compared *Prevotella* species except *P. stercorea* (49 mol%). However, gene distribution in COG categories was similar among all compared genomes (Fig. 6). In addition, the AGIOS analysis revealed that strain khD1^T shares 975 orthologous genes with *P. micans* and 1285 with *Prevotella oralis*, whereas the analysis of the average percentage of nucleotide sequence identity ranged from 65.38% to 70.94% with *P. micans* and *P. stercorea*, respectively (Table 7). Similar results were also observed in the analysis of the digital DNA-DNA hybridization (dDDH) (Table 8).

Description of P. lascolaii *strain khD1^T sp. nov.*

P. lascolaii (las.co.la'ii N.L. gen. masc. n. *lascolaii* of La Scola, the family name of the French microbiologist Bernard La Scola) is strictly anaerobic and is nonmotile and nonspore forming. It has positive oxidase activity. No production of

Table 6. Genome Comparison of Closely Related Species with the *Prevotella lascolaii* Strain khD1^T

Species	INSDC identifier	Genome size (Mbp)	G+C percent	Protein-coding genes
Prevotella lascolaii strain khD1	FKKG00000000	3.76	48.7	3194
Prevotella stercorea DSM 18206	AFZZ00000000	6.19	49	2677
Prevotella oralis ATCC 33269	AEPE00000000	5.67	44.5	2353
Prevotella loescheii JCM 12249	ARJO00000000	7.01	46.6	2828
Prevotella enoeca JCM 12259	BAIX00000000	2.86	46.5	2806
Prevotella micans DSM 21469	BAKH00000000	2.43	45.5	2828
Prevotella shahii DSM 15611	BAIZ00000000	3.49	44.4	3371
Prevotella timonensis 4401737	CBQQ000000000	6.34	42.5	2685

INSDC, International Nucleotide Sequence Database Collaboration.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

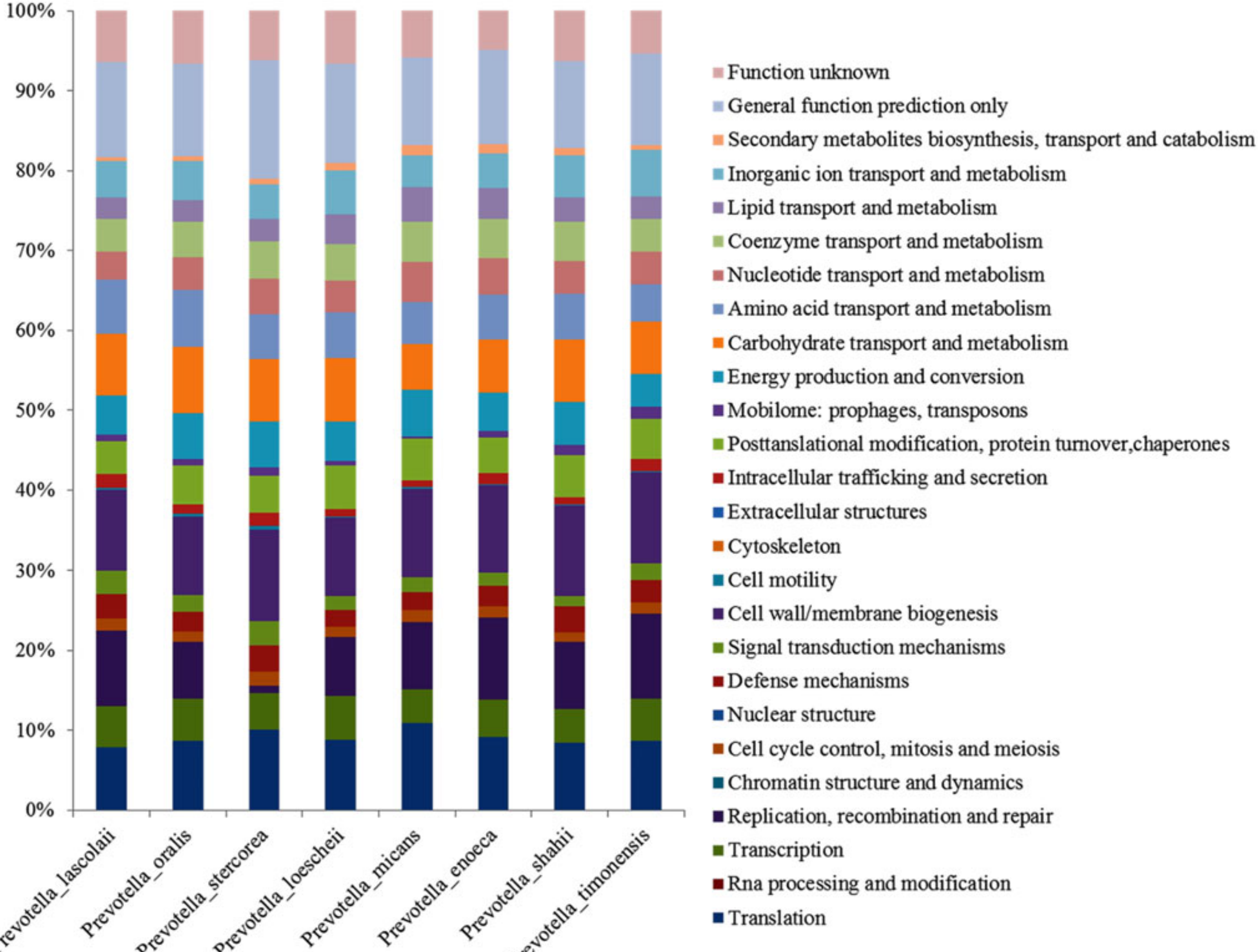


FIG. 6. Distribution of functional classes of predicted genes according to the COG of proteins of *Prevotella lascolaii* strain khD1^T among other species.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

urease or catalase was observed. Cells are mesophilic, with optimal growth at 37°C, and are gram-negative bacilli with nearly 0.65 μm of diameter and 0.9 μm of length. On Columbia agar after 2 days of incubation at 37°C under anaerobic conditions, colonies appear grayish-white, shiny, smooth, and are circular with a diameter between 1.4 and 2 mm. It is moderately saccharolytic, and arabinose, ribose, galactose, melezitose are fermented while fructose, glucose, lactose, maltose, mannose, mannitol, raffinose, rhamnose, salicin, cellobiose, sucrose, trehalose, and xylose are not fermented. Ferric citrate esculin is hydrolyzed, but gelatin and urease are not hydrolyzed. Indole and catalase are not produced and nitrate is not reduced.

P. lascolaii exhibited positive enzymic reactions for alkaline phosphatase, α-chymotrypsin, acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-galactosidase, β-galactosidase,

Table 7. Numbers of Orthologous Proteins Shared Between Genomes (Upper Right) and Average Genomic Identity of Gene Sequence Values Obtained (Lower Left)

	Prevotella lascolaii	Prevotella oralis	Prevotella stercorea	Prevotella loescheii	Prevotella micans	Prevotella enoeca	Prevotella shahii	Prevotella timonensis
P. lascolaii	**3194**	1285	1252	1255	975	1083	1099	1264
P. oralis	68.30%	**2353**	1226	1370	1038	1154	1185	1296
P. stercorea	70.94%	67.62%	**2677**	1228	968	1044	1057	1217
P. loescheii	67.21%	67.97%	67.78%	**2828**	1082	1215	1353	1311
P. micans	65.38%	66.46%	65.68%	65.70%	**2301**	935	956	983
P. enoeca	67.32%	69.23%	67.08%	68.27%	66.26%	**2806**	1055	1128
P. shahii	66.17%	67.21%	66.52%	81.03%	64.82%	67.33%	**3371**	1135
P. timonensis	66.97%	69.03%	66.71%	67.17%	65.66%	67.89%	67.02%	**2685**

The numbers of proteins per genome are indicated by bold numbers.

Table 8. Pairwise Comparison of *Prevotella lascolaii* with Other Species Using GGDC, Formula 2 (DDH Estimates Based on Identities/HSP length).[a]

	Prevotella lascolaii	Prevotella oralis	Prevotella stercorea	Prevotella loescheii	Prevotella micans	Prevotella enoeca	Prevotella shahii	Prevotella timonensis
P. lascolaii	100%	19.8% ± 2.3	31.6% ± 2.4	21.1% ± 2.3	20.0% ± 2.35	19.8% ± 2.3	22.4% ± 2.4	28.1% ± 2.4
P. oralis		100%	20.5% ± 2.3	19.7% ± 2.25	21.9% ± 2.35	20.0% ± 2.3	20.2% ± 2.35	21.0% ± 2.35
P. stercorea			100%	20.2% ± 2.3	21.4% ± 2.35	22.7% ± 2.35	21.5% ± 2.35	21.1% ± 2.4
P. loescheii				100%	24.0% ± 2.4	28.5% ± 2.45	24.9% ± 2.4	24.1% ± 2.4
P. micans					100%	29.4% ± 2.45	20.9% ± 2.3	25.2% ± 2.6
P. enoeca						100%	21.3% ± 2.3	24.0% ± 2.35
P. shahii							100%	25.7% ± 2.4
P. timonensis								100%

[a]The confidence intervals indicate the inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size). These results are in accordance with the 16S rRNA (Fig. 1) and phylogenomic analyses as well as the GGDC results.

DDH, DNA-DNA hybridization; HSP, high-scoring segment pairs.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

α-glucosidase, β-glucosidase, N-acetyl-β-glucosaminidase, and α-fucosidase. The major fatty acids are $C_{15:0}$ anteiso (36%), $C_{16:0}$ (19%), and $C_{15:0}$ iso (10%).

P. lascolaii khD1^T is sensitive to imipenem and metronidazole but resistant to amoxicillin, benzylpenicillin, and vancomycin. Its genome contains 48.7%mol G+C and measured 3,763,057 bp long. The 16S rRNA and genome sequences are both deposited in GenBank under accession numbers LN998055 and FKKG00000000, respectively. The type strain khD1^T (= DSM 101754, = CSUR P0109) was isolated in the vaginal sample of a 33-year-old French woman afflicted with BV.

Discussion

Metagenomics has enhanced our knowledge of the relationships between human vaginal microbiome, health, and diseases, and also has shown the presence of a number of unknown and uncultured microorganisms such as BVAB1, BVAB2, and BVAB3 (Fredricks et al., 2005). In the postgenomic era, new technology and omics methodologies are being intensively developed. Culturomics is one of these new approaches dynamically describing new bacteria. Based on a multiplication of culture conditions combined with a rapid identification of bacteria, it was recently introduced and applied to samples from various body sites, including the human vagina.

First application of culturomics was to study the gut microbiota. Thus, microbial culturomics has expanded the diversity of the human microbiome to 1057 species, including 197 potential new bacterial species (Lagier et al., 2016). Recently, it has also enabled the culture and description of new bacterial species found in the vagina (Diop et al., 2016; 2017a; 2017b).

In this article, we described the isolation as well as the phenotypic and genomics characteristics of a new bacterial species *P. lascolaii* isolated from a vaginal sample of a 33-year-old French woman afflicted with BV. We described the sample using a polyphasic taxono-genomic strategy (Ramasamy et al., 2014) in sequencing its genome. The phylogenetic and genomic results agreed that *P. lascolaii* is indeed distinct from its phenotypically closest species and constitutes a new species.

After sampling under strict protocols, the sample was rapidly transported to the laboratory and cultured as soon as possible in aseptic conditions. This strictly anaerobic and nonmotile bacterium was also isolated in another vaginal specimen of a patient with BV and in stool samples, thus confirming that it is not a contamination but a member of the human microbiome (unpublished data). As suggested by several authors (Fenollar and Raoult, 2016), this also leads us to believe that BV results from fecal transplantation. To prove the authenticity of our isolate, a pure culture was deposed in two different microorganism collections: the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) and the Collection de Souches de l'Unité des Rickettsies (CSUR).

This work demonstrates the ability of culturomics and its taxonogenomics approach to, respectively, explore the human microbiome and describe new bacterial species. It should also be noted that this work does not attempt to describe the medical importance of this new bacterium in BV. Instead, it expands the human vaginal flora and by sequencing the genome of new species reduces the number of sequences not assigned to a known microorganism under metagenomics. To better understand the role of these species in vaginal health and vaginal dysbiosis, further laboratory experimentation will be needed to study their pathogenesis and virulence.

Conclusions

Phenotypic and phylogenetic analyses and genomic results mean we can propose strain khD1^T as the representative of a new species named *P. lascolaii* sp. nov. The type strain khD1^T was isolated from the vaginal sample of a patient suffering from BV. Using culturomics, which uses high-throughput culture conditions with a rapid bacterial identification by MALDI-TOF, several potential new bacterial species were found in the human vagina, thus suggesting that the vagina flora is a complex and still unknown ecosystem and its diversity should be explored as fully as possible. In sum, microbial culturomics is an important new addition to the diagnostic medicine toolbox and warrants attention in future medical, global health, and integrative biology postgraduate teaching curricula.

Acknowledgments

This study was supported by Méditerranée Infection and the National Research Agency under the program ''Investissements d'avenir,'' reference ANR-10-IAHU-03.

The authors thank the Xegen Company (www.xegen.fr) for automating the genomic annotation process. They also thank TradOnline for reviewing the English.

Author Disclosure Statement

The authors declare that no conflicting financial interests exist.

References

Carver T, Harris SR, Berriman M, Parkhill J, and McQuillan JA. (2012). Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics 28, 464–469.

Carver T, Thomson N, Bleasby A, Berriman M, and Parkhill J. (2009). DNAPlotter: Circular and linear interactive genome visualization. Bioinformatics 25, 119–120.

Citron DM, Ostovari MI, Karlsson A, and Goldstein EJ. (1991). Evaluation of the E test for susceptibility testing of anaerobic bacteria. J Clin Microbiol 29, 2197–2203.

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, and Sayers EW. (2016). GenBank. Nucleic Acids Res 44, 67–72.

Diop K, Bretelle F, Michelle C, et al. (2017a). Taxonogenomics and description of *Vaginella massiliensis* gen. nov., sp. nov., strain Marseille P2517^T, a new bacterial genus isolated from the human vagina. New Microbes New Infect 15, 94–103.

Diop K, Diop A, Bretelle F, et al. (2017b). *Olegusella massiliensis* gen. nov., sp. nov., strain KHD7^T, a new bacterial genus isolated from the female genital tract of a patient with bacterial vaginosis. Anaerobe 44, 87–95.

Diop K, Raoult D, Bretelle F, and Fenollar F. (2016). ''*Murdochiella vaginalis*'' sp. nov., a new bacterial species cultivated from the vaginal flora of a woman with bacterial vaginosis. Hum Microbiome J 2, 15–16.

Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral J-P, and Raoult D. (2000). 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 38, 3623–3630.

Dubourg G, Lagier JC, Armougom F, et al. (2013). The gut microbiota of a patient with resistant tuberculosis is more comprehensively studied by culturomics than by metagenomics. Eur J Clin Microbiol Infect Dis 32, 637–645.

Fenollar F, and Raoult D. (2016). Does bacterial vaginosis result from fecal transplantation? J Infect Dis 214, 1784–1784.

Fournier PE, and Drancourt M. (2015). New Microbes New Infections promotes modern prokaryotic taxonomy: A new section ''TaxonoGenomics: New genomes of microorganisms in humans.'' New Microbes New Infect 7, 48–49.

Fournier PE, Lagier JC, Dubourg G, and Raoult D. (2015). From culturomics to taxonomogenomics: A need to change the taxonomy of prokaryotes in clinical microbiology. Anaerobe 36, 73–78.

Fredricks DN, Fiedler TL, and Marrazzo JM. (2005). Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med 353, 1899–1911.

Gouret P, Paganini J, Dainat J, et al. (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*. Pontarotti P, ed. Heidelberg, Germany: Springer Berlin, 71–87.

Gouret P, Vitiello V, Balandraud N, Gilles A, Pontarotti P, and Danchin EG. (2005). FIGENIX: Intelligent automation of genomic annotation: Expertise integration in a new software platform. BMC Bioinformatics 6, 198.

Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, and Hauser LJ. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11, 119.

Kim M, Oh HS, Park SC, and Chun J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64, 346–351.

Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, and Ussery DW. (2007). RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35, 3100–3108.

Lagier JC, Armougom F, Million M, et al. (2012). Microbial culturomics: Paradigm shift in the human gut microbiome study. Clin Microbiol Infect 18, 1185–1193.

Lagier JC, Khelaifia S, Alou MT, et al. (2016). Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 1, 16203.

Lamont R, Sobel J, Akins R, et al. (2011). The vaginal microbiome: New information about genital tract flora using molecular based techniques: Vaginal microbiome using molecular tools. BJOG Int J Obstet Gynaecol 118, 533–549.

Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, and Prohaska SJ. (2011). Proteinortho: Detection of (Co-) orthologs in large-scale analysis. BMC Bioinformatics 12, 1.

Lowe TM, and Eddy SR. (1997). tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25, 955–964.

Matuschek E, Brown DFJ, and Kahlmeter G. (2014). Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. Clin Microbiol Infect 20, O255–O266.

Meier-Kolthoff JP, Auch AF, Klenk HP, and Göker M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14, 1.

Menard JP, Fenollar F, Henry M, Bretelle F, and Raoult D. (2008). Molecular quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. Clin Infect Dis 47, 33–43.

Murray PR, Baron EJ, Jorgensen JH, Landry ML, and Pfaller MA. (2007). *Manual of Clinical Microbiology*, 9th ed. Washington, DC: ASM Press.

Narayankhedkar A, Hodiwala A, and Mane A. (2015). Clinicoetiological characterization of infectious vaginitis amongst women of reproductive age group from Navi Mumbai, India. J Sex Transm Dis 2015, 1–5.

Ramasamy D, Mishra AK, Lagier JC, et al. (2014). A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 64, 384–391.

Sakamoto M, and Ohkuma M. (2012). Reclassification of *Xylanibacter oryzae* Ueki et al. 2006 as *Prevotella oryzae* comb. nov., with an emended description of the genus *Prevotella*. Int J Syst Evol Microbiol 62, 2637–2642.

Sasser M. (2006). Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). MIDI, Technical Note #101.

Seck E, Rathored J, Khelaifia S, et al. (2015). *Virgibacillus senegalensis* sp. nov., a new moderately halophilic bacterium isolated from human gut. New Microbes New Infect 8, 116–126.

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

Seng P, Drancourt M, Gouriet F, et al. (2009). Ongoing revolution in bacteriology: Routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 49, 543–551.
Shah HN, and Collins DM. (1990). NOTES: *Prevotella*, a new genus to include *Bacteroides melaninogenicus* and related species formerly classified in the genus *Bacteroides*. Int J Syst Evol Microbiol 40, 205–208.
Stackebrandt E, and Ebers J. (2006). Taxonomic parameters revisited: Tarnished gold standards. Microbiol Today 33, 152.

Address correspondence to:
Pr. Florence Fenollar, MD, PhD
URMITE, UM 63, CNRS UMR 7278, IRD 198, INSERM U1095
Aix-Marseille University
27 Bd Jean Moulin
Marseille 13005
France

E-mail: florence.fenollar@univ-amu.fr

Abbreviations Used

AGIOS = average genomic identity of gene sequences
BV = bacterial vaginosis
BVAB = bacterial vaginosis-associated bacteria
COG = Clusters of Orthologous Groups
CSUR = Collection de souches de l'Unité des Rickettsies
DSM = Deutsche Sammlung von Mikroorganismen
FAMEs = fatty acid methyl esters
GC/MS = gas chromatography/mass spectrometry
MALDI-TOF = matrix-assisted laser desorption/ionization–time of flight
MICs = minimal inhibitory concentrations
MTBE = methyl tert-butyl ether
ORFs = open reading frames
TE buffer = Tris–EDTA buffer

Downloaded by Disc Inserm from online.liebertpub.com at 02/17/18. For personal use only.

Article 9:

Description of three new *Peptoniphilus* species cultured in the vaginal fluid of a woman diagnosed with bacterial vaginosis: *Peptoniphilus pacaensis* sp. nov., *Peptoniphilus raoultii* sp. nov., and *Peptoniphilus vaginalis* sp. nov.

Khoudia Diop, Awa Diop, Caroline Michelle, Magali Richez, Jaishriram Rathored, Florence Bretelle, Pierre-Edouard Fournier, Florence Fenollar

Published in MicrobiologyOpen

Received: 1 February 2018 | Revised: 5 April 2018 | Accepted: 30 April 2018

DOI: 10.1002/mbo3.661

ORIGINAL ARTICLE

Description of three new *Peptoniphilus* species cultured in the vaginal fluid of a woman diagnosed with bacterial vaginosis: *Peptoniphilus pacaensis* sp. nov., *Peptoniphilus raoultii* sp. nov., and *Peptoniphilus vaginalis* sp. nov.

Khoudia Diop[1] | Awa Diop[1] | Caroline Michelle[2] | Magali Richez[2] | Jaishriram Rathored[1] | Florence Bretelle[2,3] | Pierre-Edouard Fournier[1] | Florence Fenollar[1]

[1]Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection, Marseille, France

[2]Aix-Marseille Univ, IRD, AP-HM, MEPHI, IHU-Méditerranée Infection, Marseille, France

[3]Department of Gynecology and Obstetrics, Gynépole, Hôpital Nord, AP-HM, Marseille, France

Correspondence
Florence Fenollar
Institut Hospitalo-Universitaire Méditerranée-Infection, 19-21 Boulevard Jean Moulin, 13005 Marseille Cedex 05, France.
Email: florence.fenollar@univ-amu.fr

Méditerranée Infection and the National Research Agency under the program "Investissements d'avenir", reference ANR-10-IAHU-03, supported this study.

Abstract

Three previously unidentified Gram-positive anaerobic coccoid bacteria, strains KhD-2^T, KHD4^T, and Kh-D5^T, isolated from a vaginal swab, were characterized using the taxonogenomics concept. The phylogenic analysis, phenotypic characteristics, and genotypic data presented in this report attest that these three bacteria are distinct from previously known bacterial species with standing in nomenclature and represent three new *Peptoniphilus* species. Strain KhD-2^T is most closely related to *Peptoniphilus* sp. DNF00840 and *Peptoniphilus harei* (99.7% and 98.2% identity, respectively); strain KHD4^T to *Peptoniphilus lacrimalis* (96%) and strain Kh-D5^T to *Peptoniphilus coxii* (97.2%). Strains KhD-2^T, KHD4^T, and Kh-D5^T DNA G+C contents are, respectively, 34.23%, 31.87%, and 49.38%; their major fatty acid was $C_{16:0}$ (41.6%, 32.0%, and 36.4%, respectively). We propose that strains KhD-2^T (=CSUR P0125 = DSM 101742), KHD4^T (=CSUR P0110 = CECT 9308), and Kh-D5^T (=CSUR P2271 = DSM 101839) be the type strains of the new species for which the names *Peptoniphilus vaginalis* sp. nov., *Peptoniphilus raoultii* sp. nov., and *Peptoniphilu pacaensis* sp. nov., are proposed, respectively.

KEYWORDS

bacterial vaginosis, culturomics, human microbiota, *Peptoniphilus pacaensis*, *Peptoniphilus raoultii*, *Peptoniphilus vaginalis*, taxogenomics

1 | INTRODUCTION

Since the 1800s, physicians and researchers investigate the vaginal bacterial community using both cultivation and culture-independent methods (Pandya et al., 2017; Srinivasan et al., 2016). To date, many species from the vaginal microbiota have been identified. The healthy vaginal flora is associated to a biotope rich in *Lactobacilli* species (Li, McCormick, Bocking, & Reid, 2012). The vaginal microbiota has a beneficial relationship with its host and can also impact women's health, that of their partners as well as their neonates (Lepargneur & Rousseau, 2002; Srinivasan & Fredricks, 2008). A depletion of vaginal *Lactobacilli* can lead to bacterial vaginosis (BV). This disease is a dysbiosis that may be associated to sexually transmitted infections as well as miscarriage and preterm birth in pregnant women (Afolabi, Moses, & Oduyebo, 2016; Martin & Marrazzo, 2016).

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2018 The Authors. *MicrobiologyOpen* published by John Wiley & Sons Ltd.

A microbial culturomics study exploring the bacterial community of the vaginal econiche flora in healthy women and patients suffering from bacterial vaginosis enabled the isolation of three Gram-positive-staining, anaerobic, and coccoid bacteria in the vaginal discharge of a woman with bacterial vaginosis (Lagier et al., 2015, 2016). These bacteria exhibited phylogenetic and phenotypic proximity to species of the *Peptoniphilus* genus. Created after the division of *Peptostreptococcus* genus into five genera (Ezaki et al., 2001), the *Peptoniphilus* genus belonging to the Peptoniphilaceae family that regroup members of the genera *Peptoniphilus*, *Parvimonas*, *Murdochiella*, *Helcococcus*, *Gallicola*, *Finegoldia*, *Ezakiella*, *Anaerosphaera*, and *Anaerococcus* (Johnson, Whitehead, Cotta, Rhoades, & Lawson, 2014; Patel et al., 2015). The *Peptoniphilus* genus is currently made of 16 valid published species (http://www.bacterio.net/peptoniphilus.html). These bacteria employ amino acids and peptone as a major energy sources (Ezaki et al., 2001). They are mainly cultivated from diverse human samples such as sacral ulcer, vaginal discharge, as well as ovarian, peritoneal, and lacrymal gland abscesses (Ezaki et al., 2001; Li et al., 1992; Ulger-Toprak, Lawson, Summanen, O'Neal, & Finegold, 2012).

Herein, we describe the isolation and taxonogenomic characterization (Fournier, Lagier, Dubourg, & Raoult, 2015) of strains KhD-2^T, $KHD4^T$, and Kh-$D5^T$ as type strains of three new *Peptoniphilus* species for which the names *Peptoniphilus vaginalis* sp. nov. (=CSUR P0125, =DSM 101742), *Peptoniphilus raoultii* sp. nov. (=CSUR P0110, =CECT 9308), and *Peptoniphilus pacaensis* sp. nov. (=CSUR P2271, =DSM 101839), are proposed, respectively. All the three strains were cultivated from the vaginal swab of the same patient.

2 | MATERIALS AND METHODS

2.1 | Samples and ethics

The vaginal specimen from a French 33-year-old woman with bacterial vaginosis was sampled at Hospital Nord in Marseille (France) in October 2015 using a Sigma Transwab (Medical Wire, Corsham, United Kingdom). Bacterial vaginosis was diagnosed as previously described (Menard, Fenollar, Henry, Bretelle, & Raoult, 2008). The patient had not received any antibiotic for several months. The local IFR48 ethics committee in Marseille (France) authorized the study (agreement number: 09-022). In addition, the patient gave her signed informed consent.

2.2 | Bacterial strain isolation and identification

After sampling, the specimen was preincubated in a blood culture bottle (Becton-Dickinson Diagnostics, Le Pont-de-Claix, France). The blood culture bottle was enriched with 3 ml of sheep blood (bioMérieux, Marcy l'Etoile, France) and 4 ml of rumen fluid, filter-sterilized through a 0.2 μm pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France). Various preincubation periods (1, 3, 7, 10, 15, 20, and 30 days) were tested. Then, 50 μl of the supernatant were inoculated on both Colistin-nalidixic acid (CNA) used for selective enrichment of Gram-positive bacteria and trypticase soy agar plates used for cultivation of nonfastidious and fastidious microorganisms (both BD Diagnostics), and then incubated for 4 days under anaerobic conditions at 37°C. Isolated colonies were purified and subsequently identified by matrix-assisted laser-desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry with a Microflex spectrometer (Bruker, Leipzig, Germany) that compared the new spectra with those present in the library (Bruker database and URMITE database, constantly updated), as previously reported (Seng et al., 2009). If the score was >1.99, the bacterium was considered as identified at the genus level (score between 2.0 and 2.299) or species level (score from 2.3 to 3.0). When the score was <1.7, no identification was considered reliable. The 16S rRNA sequence of unidentified isolates was obtained using an ABI Prism 3130xl Genetic Analyzer capillary sequencer (Applied Biosystems, Bedford, MA, USA), as previously described (Morel et al., 2015; Seng et al., 2009). Finally, the sequences were compared to the NCBI nr database using the BLAST algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi). If the 16S rRNA sequence similarity value was lower than 98.7%, the isolate was considered as a putative new species (Kim, Oh, Park, & Chun, 2014; Stackebrandt & Ebers, 2006; Yarza et al., 2014).

2.3 | Phylogenetic analysis

The 16S rRNA sequences of isolates not identified using mass spectrometry and those of members of the family Peptoniphilaceae with standing in nomenclature (downloaded from the nr database) were aligned using CLUSTALW (Thompson, Higgins, & Gibson, 1994) with default setting. The phylogenetic inferences were performed using both the neighbor-joining and maximum-likelihood methods with the software MEGA version 6 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013).

2.4 | Phenotypic characteristics

For each new isolate, cell morphology was visualized using optical and electron microscopy. Oxidase, catalase, motility, sporulation tests, as well as Gram stain were performed as already reported (Murray, Baron, Jorgensen, Landry, & Pfaller, 2007). Cells were fixed for electron microscopy for at least 1 hour at 4°C with 2.5% glutaraldehyde in a 0.1 mol L^{-1} cacodylate buffer. One drop of cell suspension was deposited for about 5 min on a glow-discharged formvar carbon film on 400-mesh nickel grids (FCF400-Ni, EMS). The grids were dried on a blotting paper. Then, the cells were negatively stained at room temperature for 10 s with a 1% ammonium molybdate solution in filtered water. Micrographs were obtained using a Tecnai G20 Cryo (FEI) transmission electron microscope operated at 200 keV.

In order to characterize the best growth conditions of each isolate, bacteria were inoculated on 5% sheep blood-enriched Columbia agar (bioMérieux) incubated at various atmospheres (aerobic, anaerobic, and microaerophilic) and temperatures (56, 42, 37, 28, and

25°C) (Mishra, Lagier, Nguyen, Raoult, & Fournier, 2013). Several salinity (NaCl concentrations of 0%, 5%, 15%, and 45%) and pH (5, 6, 6.5, 7, and 8.5) conditions were also tested.

Biochemical analyses were realized using various strips (API ZYM, API 20A, and API 50CH) according to the manufacturer's instructions (bioMérieux) (Avguštin, Wallace, & Flint, 1997; Durand et al., 2017). The tests were performed in anaerobic chamber. The strips were incubated there for 4, 24, and 48 hr, respectively.

For the analysis of cellular fatty acid methyl ester (FAME), gas chromatography/mass spectrometry (GC/MS) was achieved. All three isolates were grown anaerobically at 37°C on 5% sheep blood-enriched Columbia agar (bioMérieux). For each isolate, after 2 days of incubation, two aliquots with roughly 25–70 mg of bacterial biomass per tube were prepared. FAME preparation and GC/MS analyses were performed as already reported (Dione et al., 2016; Sasser, 2006). FAMEs were separated with an Elite 5-MS column and monitored by MS (Clarus 500-SQ 8 S, Perkin Elmer, Courtaboeuf, France). A spectral database search was done with MS Search 2.0 operated using the standard reference database 1A (NIST, Gaithersburg, USA) as well as the FAMEs mass spectral database (Wiley, Chichester, UK).

The susceptibility of all three isolates was tested for 11 antibiotics: amoxicillin (0.16–256 μg/ml), benzylpenicillin (0.002–32 μg/ml), ceftriaxone (0.002–32 μg/ml), ertapenem (0.002–32 μg/ml), imipenem (0.002–32 μg/ml), amikacin (0.16–256 μg/ml), erythromycin (0.16–256 μg/ml), metronidazole (0.16–256 μg/ml), ofloxacin (0.002–32 μg/ml), rifampicin (0.002–32 μg/ml), and vancomycin (0.16–256 μg/ml). Minimal inhibitory concentrations (MICs) were estimated using E-test strips (bioMérieux) and according to EUCAST recommendations (Citron, Ostovari, Karlsson, & Goldstein, 1991; Matuschek, Brown, & Kahlmeter, 2014).

2.5 | Genome sequencing and analyses

After a pretreatment of 2 hr at 37°C using lysozyme, the genomic DNAs (gDNAs) of strains KhD-2^T, KHD4^T, and Kh-D5^T were extracted using the EZ1 biorobot and EZ1 DNA Tissue kit (Qiagen). An elution volume of 50 μl was obtained for each sample. The gDNAs were quantified by a Qubit assay (Life technologies, Carlsbad, CA, USA) at 74.2, 22.4, and 16.4 ng/μl, respectively. Genomic sequencing of each strain was performed with a MiSeq sequencer (Illumina Inc, San Diego, CA, USA) and the Mate Pair strategy.

The Mate Pair library was prepared with the Nextera Mate Pair guide (Illumina) using 1.5 μg of gDNA. The gDNA samples were fragmented and tagged using a Mate Pair junction adapter (Illumina). Then, the fragmentation pattern was validated using a DNA 7500 labchip on an Agilent 2100 BioAnalyzer (Agilent Technologies Inc, Santa Clara, CA, USA). No size selection was done. Thus, 537, 600, and 480.7 ng of tagmented fragments were, respectively, circularized. Circularized DNAs were mechanically cut to smaller fragments using Optima on a bimodal curve at 507 and 1,244 bp for KhD-2^T, 975 and 1,514 bp for KHD4^T, and 609 and 999 bp for Kh-D5^T on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The libraries profiles were visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc, Santa Clara, CA, USA) and the final concentrations libraries were determined. Then, the libraries were normalized at 2 nmol L^{-1}, pooled, denatured, diluted at 15 pmol L^{-1}, loaded onto the reagent cartridge, and onto the instrument. Sequencing was performed in a single 39-hr run in a 2 × 250-bp.

The genome assembly was performed with a pipeline that enabled to create an assembly with various software such as Velvet (Zerbino & Birney, 2008), Spades (Bankevich et al., 2012), and Soap Denovo (Luo et al., 2012), on trimmed data with MiSeq and Trimmomatic (Bolger, Lohse, & Usadel, 2014) software or untrimmed data with only MiSeq software. In order to reduce gaps, GapCloser was used (Luo et al., 2012). Phage contamination was searched (blastn against Phage Phix174 DNA sequence) and eliminated. Finally, scaffolds with sizes under 800 bp and scaffolds with a depth value lower than 25% of the mean depth were identified as possible contaminants and removed. The best assembly was considered by using several criteria including number of scaffolds, N50, and number of N. Spades gave the best assembly for the three studied strains with depth coverage of 518x.

Prodigal was used to predict open reading frames (ORFs) (Hyatt et al., 2010) using default parameters. However, the predicted ORFs were excluded if they spanned a sequencing gap region (containing Ns). The predicted bacterial protein sequences were analyzed as previously reported (Alou et al., 2017). tRNA genes were found using the tRNAScan-SE tool (Lowe & Eddy, 1997), while RNAmmer was used to find ribosomal RNAs (Lagesen et al., 2007). Phobius was used to predict lipoprotein signal peptides and the number of transmembrane helices (Käll, Krogh, & Sonnhammer, 2004). ORFans were identified when the BLASTP search failed to provide positive results (*E*-value smaller than 1e^{-03} for ORFs with a sequence size larger than 80 aa or an *E*-value smaller than 1e^{-05} for ORFs with a sequence length smaller than 80 aa), as previously reported (Alou et al., 2017). For genomic comparison, the closest species with validly published names in the 16S RNA phylogenetic tree were identified with the Phylopattern software (Gouret, Thompson, & Pontarotti, 2009). The complete genome, proteome, and ORFeome sequences were retrieved for each selected species in NCBI. An annotation of the entire proteome in order to define the distribution of functional classes of predicted genes according to the COG classification of their predicted protein products was performed as already reported (Alou et al., 2017). Annotation and comparison processes were done using the DAGOBAH software as previously described (Alou et al., 2017; Gouret et al., 2005, 2011). Finally, in order to evaluate the genomic similarity between the genomes, we determined two previously described parameters: average amino acid identity (AAI) based on the overall similarity between two genomic datasets of proteins available at (http://enve-omics.ce.gatech.edu/aai/index) and digital DNA–DNA hybridization (dDDH) (Auch, von Jan, Klenk, & Göker, 2010; Meier-Kolthoff, Auch, Klenk, & Göker, 2013; Alou et al., 2017; Rodriguez & Konstantinidis, 2014; Chun et al., 2018).

3 | RESULTS

3.1 | Strain identification and phylogenetic analysis

The MS identification of the three bacteria, secluded, respectively, after 24 hr (strains KhD-2^T and KHD4^T) and 15 days (Kh-D5^T) of preincubation, failed. This suggested that these isolates were not in the database and may be unknown species. Pairwise analysis of 16S rRNA sequences attested that strain KhD-2^T exhibited 92.8% and 87.4% sequence similarities with strains KHD4^T and Kh-D5^T, respectively, and strains KHD4^T and Kh-D5^T had an 88.7% identity. BLASTN sequence searches demonstrated that the three strains were related to the genus *Peptoniphilus*, suggesting that each strain represented a new species within this genus. Strain KhD-2^T exhibited a 16S rRNA similarity of 99.7% with *Peptoniphilus* sp. strain DNF00840 (GenBank KQ960236) over 1,842 bp and 98.2% with *Peptoniphilus harei* (GenBank NR_026358.1) over 1,488 bp. Strain KHD4^T exhibited a 16S rRNA similarity of 96% with *Peptoniphilus lacrimalis* (GenBank NR_041938.1) over 1,489 bp. Finally, strain Kh-D5^T exhibited a 16S rRNA similarity of 97.2% with *Peptoniphilus coxii* (GenBank NR_117556.1) over 1,491 bp (Figure 1). As these percentage similarities were under the threshold of 98.7% established to delineate new species (Kim et al., 2014; Stackebrandt & Ebers, 2006; Yarza et al., 2014), strains KhD-2^T, KHD4^T, and Kh-D5^T were considered as representative strains of putative new *Peptoniphilus* species. The names *P. vaginalis* sp. nov., *P. raoultii* sp. nov., and *P. pacaensis* sp. nov. are, respectively, proposed.

The reference MALDI-TOF MS spectra of our isolates were added in our database (http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database) and then compared to those of other *Peptoniphilus* spp. (Figure 2).

3.2 | Phenotypic features

Cells from all three novel strains (KhD-2^T, KHD4^T, and Kh-D5^T) were Gram- -positive cocci (mean diameter of 0.6–0.7 μm for each). After 4 days of incubation, colonies on blood agar were grey and circular, and all had a diameter ranging from 1 to 2 mm. For all the three strains, growth occurred only in anaerobic atmosphere. Besides, optimal growth occurred at 37°C, with a pH between 6.5 and 8.5, and a NaCl concentration lower than 5%. They exhibited no catalase, oxidase, and urease activities. Using API 20A strips, all tests including aesculin, arabinose, cellobiose, gelatin, glucose, glycerol, indole, lactose, maltose, mannitol, mannose, raffinose, rhamnose, saccharose, sorbitol, trehalose, urease, and xylose were negative for strains KHD4^T and Kh-D5^T, whereas for strain KhD-2^T, indole formation was positive, and gelatin was hydrolyzed. API ZYM strips showed that the three isolates exhibited positive reactions for acid phosphatase, esterase, and Naphthol-AS-BI-phosphohydrolase. In addition, strains KhD-2^T and KHD4^T had *N*-acetyl-β-glucosaminidase and leucine arylamidase activities. In contrast, an alkaline phosphatase activity was observed for strains KhD-2^T and Kh-D5^T. All other remaining tests including valine arylamidase, lipase, cystine arylamidase, trypsin, galactosidase, glucosidase, β-glucuronidase, α-mannosidase, and α-fucosidase were negative. Using API 50CH strips, all three isolates fermented ribose, tagatose, and potassium-5-ketogluconate. However, they did not ferment adonitol, aesculin, arabinose, arabitol, cellobiose, dulcitol, erythritol, fructose, fucose, galactose, glucose, glycerol, glycogen, inulin, lyxose, inositol, mannose, mannitol, maltose, melibiose, potassium gluconate, potassium-2-ketogluconate, salicine, saccharose, sorbitol, sorbose, trehalose, melezitose, raffinose, rhamnose, starch, turanose, xylitol, and xylose. Table 1 displayed the phenotypic differences between these bacteria and other *Peptoniphilus* spp.

The fatty acid composition of the three strains was as following: strain KhD-2^T contained saturated acid $C_{16:0}$ (41.6%) and $C_{14:0}$ (14.7%); unsaturated acids were also detected (Table 2); strains KHD4^T and Kh-D5^T contained $C_{16:0}$ (32% and 36%, respectively), $C_{18:2\omega6}$ (26% and 24%, respectively), and $C_{18:1\omega9}$ (26% and 21%, respectively) (Table 2). These fatty acid results were likened to those of related species in Table 2 (Johnson et al., 2014; Rooney, Swezey, Pukall, Schumann, & Spring, 2011). Strain KhD-2^T can be distinguished from its nearest neighbor *P. harei* by the production of $C_{14:0}$ (14.7% vs. 4.4%). Strain KHD4^T can be distinguished from its closest related species *P. lacrimalis* by the presence of fatty acids: $C_{14:0}$, $C_{17:0}$ iso 3-OH, and anteiso-$C_{17:0}$. Finally, strain Kh-D5^T showed a fairly similar profile with its neighbors *P. coxii* and *Peptoniphilus ivorii* with some differences such as the presence of antesio-$C_{5:0,}$ only in strain Kh-D5^T (4.5%), of iso-$C_{5:0}$ in *P. coxii* (5.5%), and $C_{17:0}$ iso 3-OH and antesio-$C_{17:0}$, solely in *P. ivorii* (7.7% and 3.8%, respectively). Besides, the three strains were sensitive to amoxicillin, benzylpenicillin, ceftriaxone, ertapenem, imipenem, metronidazole, rifampicin, and vancomycin, but resistant to amikacin, erythromycin, and ofloxacin (Table 3).

3.3 | Genome characteristics

Strains KhD-2^T, KHD4^T, and Kh-D5^T exhibited genomes sizes of 1,877,211, 1,623,601, and 1,851,572 bp long, respectively (Figure 3). The genome characteristics were detailed in Table 4. The repartition of genes into the 25 general COG categories was represented in Table 5 and Figure 4. When compared to other *Peptoniphilus* species, the three strains had genome sizes, G+C contents and total gene counts in the same range (Table 6, Figure 5). Although, base composition varies widely among bacterial species, the genes within a given genome are relatively similar in G+C content with the exception of recently acquired genes. As a matter of fact, DNA sequences acquired by horizontal transfer often bear unusual sequence characteristics and can be distinguished from ancestral DNA notably by a distinct G+C content (Lawrence & Ochman, 1997). The region between 100,000 and 600,000 bp of the chromosome from strain KhD-5^T showed a high variation in G+C content (Figure 3). Thus, 43 genes putatively acquired by horizontal gene transfer were identified in this region, including 25 genes specific for strain KhD-5^T and 18 genes shared with strain *Peptoniphilus urinimassiliensis*. Consequently, the presence of these genes may play a role in the

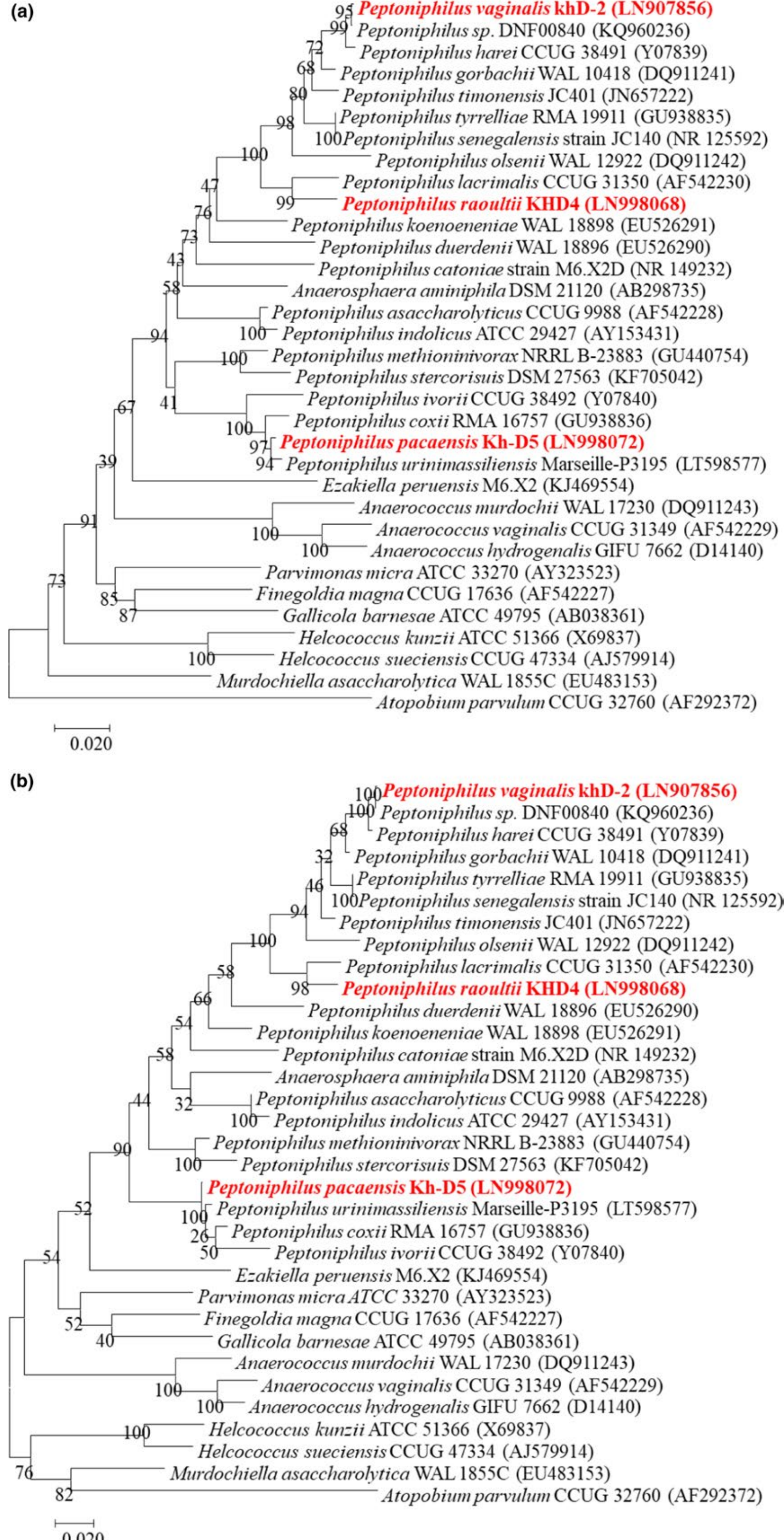


FIGURE 1 Phylogenetic analysis based on the 16S RNA gene sequence highlighting the position of *Peptoniphilus vaginalis* strain KhD-2^T, *Peptoniphilus raoultii* strain KHD4^T, and *Peptoniphilus pacaensis* strain Kh-D5^T relative to other closely related strains. GenBank accession numbers are indicated in parentheses. Sequences were aligned using Muscle v3.8.31 with default parameters and, phylogenetic inferences were performed using the neighbor-joining (a) and maximum-likelihood (b) methods with the software MEGA version 6. The scale bar represents a 2% nucleotide sequence divergence

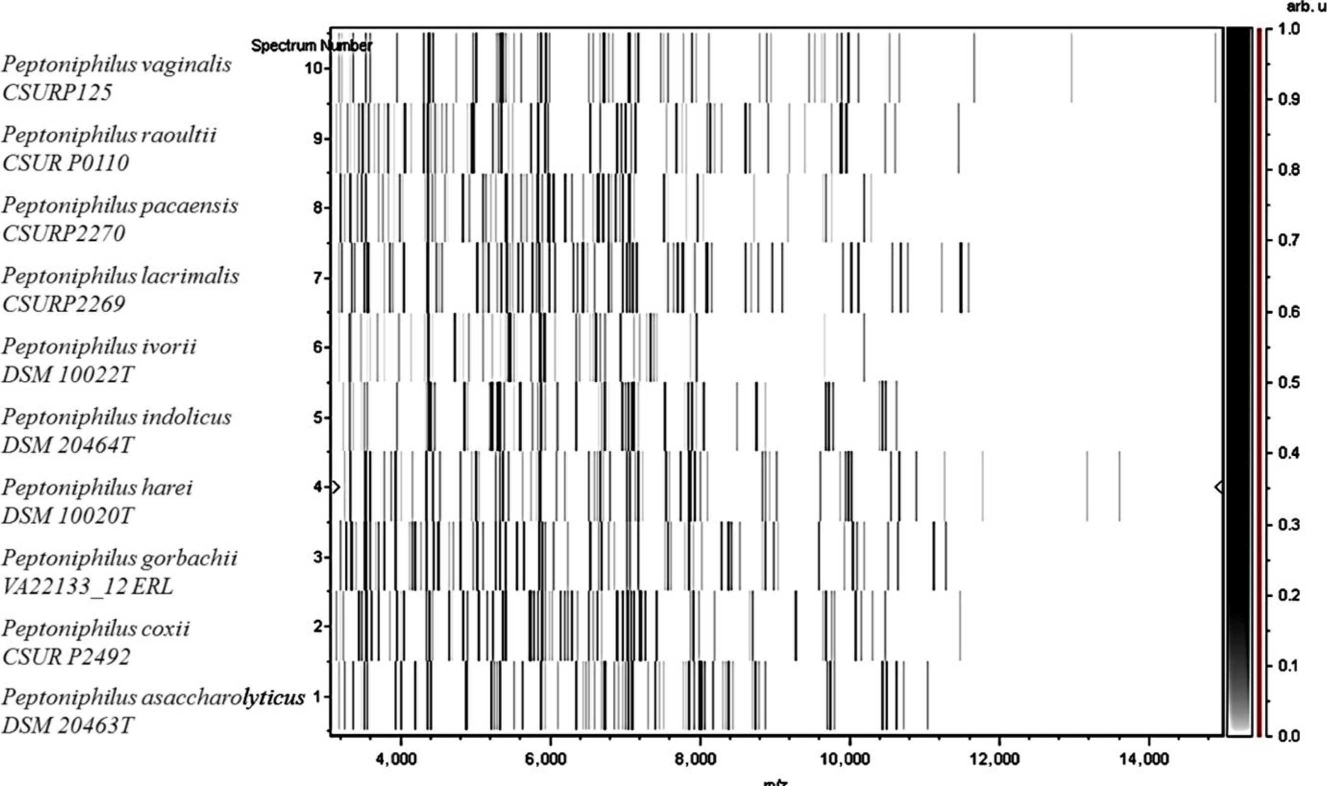


FIGURE 2 Gel view comparing strains KhD-2^T, KHD4^T, and Kh-D5^T to other species within the genus *Peptoniphilus*. The gel view displays the raw spectra of loaded spectrum files arranged in a pseudo-gel-like look. The *x*-axis records the *m/z* value. The left *y*-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a gray scale scheme code. The right *y*-axis indicates the relation between the color of a peak and its intensity, in arbitrary units. Displayed species are indicated on the left

significant difference in genomic G+C content observed between strain KhD-5^T and other compared *Peptoniphilus* species as well as the similar genomic G+C content observed between strain KhD-5^T and *P. urinimassiliensis*.

The dDDH values ranked from 20.1% ± 2.3% between *P. harei* and *P. duerdenii* to 56.4% ± 2.75% between *P. lacrimalis* and *P. urinimassiliensis* (Table 7). When comparing the three new strains to other *Peptoniphilus* species, strain KhD-2^T exhibited dDDH values ranging from 22.7% ± 2.4% with *Peptoniphilus indolicus* to 47.3% ± 2.55% with *P. coxii*; dDDH values from strain KHD4^T ranged from 19.0% ± 2.25% with *P. harei* to 44.3% ± 2.55% with *P. coxii*; and strain Kh-D5^T exhibited dDDH values ranging from 20.7% ± 2.35% with *P. coxii* to 45.0% ± 2.60% with *P. urinimassiliensis* (Table 7). Furthermore, the AAI values ranged from 51.3% between *P. coxii* and *P. indolicus* to 84.0% between *P. indolicus* and *Peptoniphilus asaccharolyticus* (Table 8). Comparing the three new isolates to their neighbors, strain KhD-2^T shared AAI values ranging from 51.5% with *P. urinimassiliensis* to 92.9% with *P. harei*, AAI values of strain KHD4^T ranging from 50.9% with *P. urinimassiliensis* to 70.6% with *P. lacrimalis*, and strain Kh-D5^T exhibited AAI values ranging from 50.2% with *P. asaccharolyticus* to 92.9% with *P. urinimassiliensis* (Table 8). According to the fact that the threshold of dDDH and AAI values for distinguishing different species are 70% and 95%–96%, respectively (Chun et al., 2018; Klappenbach et al., 2007; Meier-Kolthoff et al., 2013; Richter & Rosselló-Móra, 2009; Rodriguez-R & Konstantinidis, 2014), these data confirm the classification of strains KhD-2^T, KHD4^T, and Kh-D5^T in distinct species.

4 | DISCUSSION

The aim of this study was to investigate, using culturomics, the vaginal flora of a woman with bacterial vaginosis. Indeed, bacterial vaginosis is a gynecologic disorder marked by a perturbation of the vaginal microbiota equilibrium with a loss of commensal *Lactobacillus* spp. and their replacement with anaerobic bacteria including *Atopobium vaginae*, *Bacteroides* spp., *Mobiluncus* spp., *Prevotella* spp., and numerous Gram-positive anaerobic cocci (Bradshaw et al., 2006; Onderdonk, Delaney, & Fichorova, 2016; Shipitsyna et al., 2013). Gram-positive anaerobic cocci were associated to various infections (Murdoch, 1998). They represent about 24%–31% of anaerobic bacteria cultivated in clinical specimens (Murdoch, Mitchelmore, & Tabaqchali, 1994). In this present study, three novel Gram-positive-staining, anaerobic cocci (KhD-2^T, KHD4^T, and Kh-D5^T) were cultured in the vaginal discharge of a patient suffering from bacterial vaginosis. These bacteria exhibited sufficient MALDI-TOF MS profiles, 16S rRNA sequence,

TABLE 1 Compared phenotypic characteristics of *Peptoniphilus vaginalis* strain KhD-2^T, *Peptoniphilus raoultii* strain KHD4^T, *Peptoniphilus pacaensis* strain Kh-D5^T, and other closely related *Peptoniphilus* species. Data were obtained from the original descriptions of species

Properties	*P. vaginalis*	*P. raoultii*	*P. pacaensis*	*P. harei*	*P. lacrimalis*	*P. coxii*	*P. duerdenii*	*P. indolicus*	*P. asaccharolyticus*
Cell diameter (μm)	0.66	0.7	0.7	0.5–1.5	0.5–0.7	<0.7	≥0.7	0.7–1.6	0.5–1.6
% G+C	34.23	31.87	49.38	34.44	30.22	44.62	34.24	31.69	32.30
Major fatty acid (%)	$C_{16:00}$ (41.6)	$C_{16:00}$ (32)	$C_{16:00}$ (36.4)	$C_{16:00}$ (31.2)	$C_{16:00}$ (27.7)	$C_{16:00}$ (49.9)	$C_{16:00}$ (33)	$C_{16:00}$ (19.4)	C18:2ω6 (22.0)
Production of									
Alkaline phosphatase	+	–	+	–	–	–	–	+	+
Indole	+	–	–	+	–	–	+	+	–
Catalase	–	–	–	+	na	–	–	–	–
Urease	–	–	–	–	–	–	–	–	–
β-galactosidase	–	–	–	–	–	–	–	–	–
N-Acetyl-β-glucosaminidase	+	+	–	na	na	–	–	na	na
Acid from									
Ribose	+	+	+	–	–	–	–	–	–
D-fructose	+	–	–	–	–	–	–	–	–
Habitat	Human vagina	Human vagina	Human vagina	Human sacral ulcer	Human eyes	Human specimens	Human vagina	Summer mastitis of cattle	Human vagina

+, positive; –, negative; v, variable and na (not available) data.

TABLE 2 Cellular fatty acid profiles (%) of strains KhD-2^T, KHD4^T, and Kh-D5^T compared with other *Peptoniphilus* species

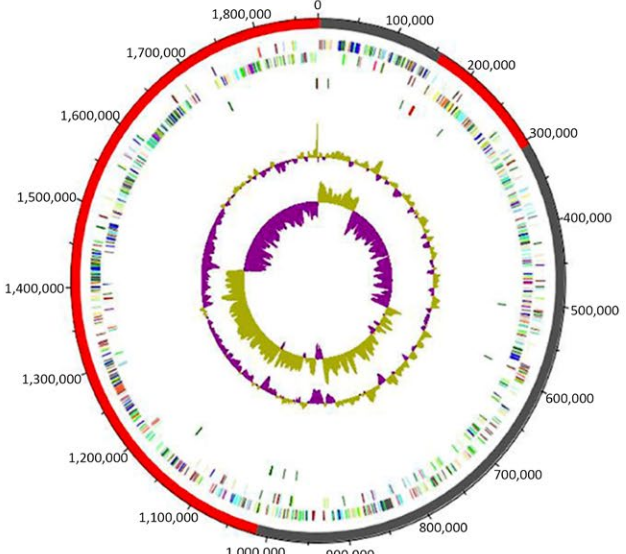
Fatty acids	Name	1	2	3	4	5	6	7	8	9	10
C4:00	Butanoic acid	TR	–	–	–	–	–	–	–	–	–
iso-C5:0	3-Methyl-butanoic acid	–	–	–	–	–	5.5	–	–	–	–
anteiso-C5:0	2-Methyl-butanoic acid	TR	–	4.5	–	–	–	–	–	–	–
C10:0	Decanoic acid	–	–	TR	TR	–	–	2.8	TR	–	–
C12:0	Dodecanoic acid	TR	–	TR	–	TR	TR	–	1.2	TR	2.3
C13:0	Tridecanoic acid	TR	–	–	–	–	–	–	–	–	–
C14:0	Tetradecanoic acid	**14.7**	TR	4.9	4.4	2.9	8.6	4.4	**12.6**	4.4	5.4
C14:1ω5	9-Tetradecenoic acid	TR	–	–	–	–	–	–	–	–	–
C15:0	Pentadecanoic acid	1.1	TR	TR	–	–	1.4	–	–	–	–
C16:0	Hexadecanoic acid	**41.6**	**32.0**	**36.4**	**32.1**	**27.7**	**49.9**	**33.0**	**19.4**	**29.5**	**14.4**
C16:0 9,10-methylene	2-Hexyl-cyclopropaneoctanoic acid	–	TR	–	–	–	–	–	–	–	–
C16:1ω5	11-Hexadecenoic acid	TR	–	–	–	–	–	–	–	–	–
C16:1ω7	9-Hexadecenoic acid	6.2	1.0	TR	1.0	3.2	–	–	–	1.0	3.9
C16:1ω9	7-Hexadecenoic acid	TR	–	–	–	–	–	–	3.6	–	–
C17:0	Heptadecanoic acid	TR	TR	TR	–	–	–	–	–	–	–
C17:0 iso 3-OH	3-Hydroxy-heptadecanoic acid	–	–	–	6.0	3.0	–	–	–	7.7	-
anteiso-C17:0	14-Methyl-hexadecanoic acid	TR	–	–	4.2	1.8	–	–	2.6	3.8	1.6
C17:1ω7	10-Heptadecenoic acid	TR	–	–	–	–	–	–	–	–	–
C18:0	Octadecanoic acid	3.9	8.8	3.6	7.2	**11.2**	**13.1**	**16.2**	2.5	4.8	9.4
C18:1ω7	11-Octadecenoic acid	4.8	3.7	2.0	1.9	3.5	–	–	3.5	2.6	–
C18:1ω9	9-Octadecenoic acid	**12.1**	**25.8**	**21.2**	**17.0**	**25.7**	**17.3**	**22.6**	6.2	**11.4**	**20.2**
C18:2ω6	9,12-Octadecadienoic acid	**12.0**	**26.4**	**24.4**	**17.0**	**13.6**	3.2	**21.1**	**13.0**	**24.0**	**22.0**

Strains: 1, *P. vaginalis* strain KhD-2^T; 2, *P. raoultii* strain KHD4^T; 3, *P. pacaensis* strain Kh-D5^T; 4, *Peptoniphilus harei* DSM 10020^T; 5, *P. lacrimalis* DSM 7455^T; 6, *P. coxii* CSUR 2492^T; 7, *P. uerdenii* WAL 18896^T; 8, *P. indolicus* DSM 20464^T, 9, *P. ivorii* CCUG 38492^T and 10, *P. asaccharolyticus* CCUG 9988^T. Strains 1, 2, 3, and 6 data are from this study and strains 4, 5, 7 to 9, data come from Rooney et al., 2011 and Johnson et al., 2014. Predominant products are shown in bold; TR, trace amounts < 1%; –, not detected.

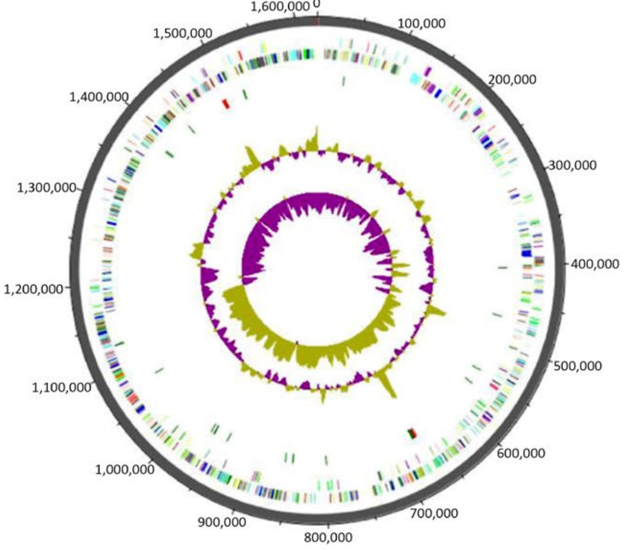
TABLE 3 Minimal inhibitory concentrations (MIC µg/µl) of antibiotics for *P. vaginalis* strain KhD-2^T, *P. raoultii* strain KHD4^T, and *P. pacaensis* strain Kh-D5^T

Antibiotics	Concentration (µg/ml)	*P. vaginalis* strain KhD-2^T	*P. raoultii* strain KHD4^T	*P. pacaensis* strain Kh-D5^T
Amoxicillin	0.016–256	0.032	0.016	0.016
Benzylpenicillin	0.002–32	0.094	0.002	0.002
Ceftriaxone	0.002–32	0.064	0.064	0.064
Ertapenem	0.002–32	0.002	0.003	0.002
Imipenem	0.002–32	0.004	0.002	0.002
Metronidazole	0.016–256	0.125	0.032	0.032
Rifampicin	0.002–32	0.002	0.002	0.002
Vancomycin	0.016–256	0.094	0.094	0.094
Amikacin	0.016–256	>256	>256	>256
Erythromycin	0.016–256	1	2	2
Ofloxacin	0.002–32	>256	>256	2

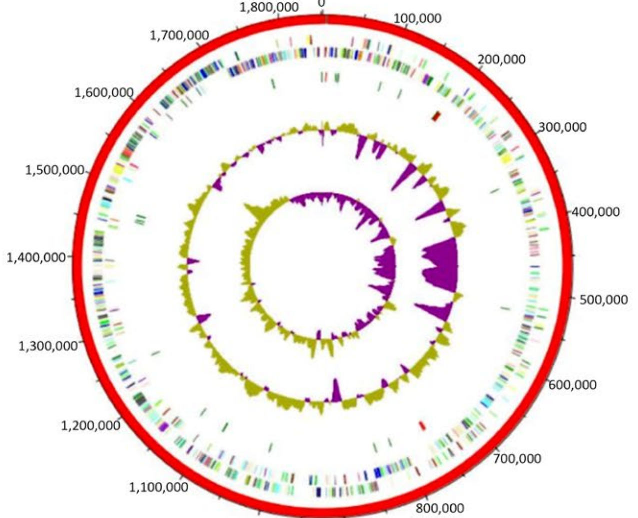
FIGURE 3 Graphical circular map of the three genomes. From outside to the center: Contigs (red/gray), COG category of genes on the forward strand (three circles), genes on forward strand (blue circle), genes on the reverse strand (red circle), COG category on the reverse strand (three circles), G+C content



Peptoniphilus vaginalis strain KhD-2^{T}



Peptoniphilus raoultii strain KHD4^{T}



Peptoniphilus pacaensis strain Kh-D5^{T}

TABLE 4 Nucleotide and gene count levels of the genomes

	P. raoultii		*P. vaginalis*		*P. vaginalis*	
Attribute	**Value**	**% of total[a]**	**Value**	**% of total[a]**	**Value**	**% of total[a]**
Size (bp)	1,623,601	100%	1,877,211	100%	1,851,572	100%
G+C content (bp)	517,506	31.87%	642,534	34.22%	914,357	49.38%
Coding region (bp)	1,467,557	90.39%	1,692,527	90.16	3,579,496	85.07%
Total genes	1,624	100%	1,780	100%	1,801	100%
RNA genes	42	2.59%	40	2.35%	54	3.00%
Protein-coding genes	1,520	93.60%	1,698	95.39%	1,699	94.34%
Genes with function prediction	1,222	75.25%	1,375	77.24%	1,323	73.45%
Genes assigned to COGs	1,048	65.53%	1,204	67.64%	1,175	65.24%
Genes with peptide signals	162	9.97%	169	9.49%	231	12.83%
Genes with transmembrane helices	349	21.49%	403	22.64%	414	22.98%

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

phenotypic, and genomic differences with *Peptoniphilus* species to be regarded as representative strains of three new species within this genus. Currently, this genus contains 16 species with validly published names. Most of them have been observed in human clinical specimens (Ezaki et al., 2001).

Data from phylogenetic analysis and genomic comparison exhibited the heterogeneity of this genus and revealed that strain KhD-2^T and *Peptoniphilus* sp. DNF00840^T share 99.79% 16S rRNA gene sequence similarity, an ANI value of 96.83% and 75.0% of dDDH. In fact, to differentiate bacterial species, thresholds lower than 98.7%, 94%, and 70% were delimited for 16S rRNA sequence identity, ANI, and dDDH values, respectively. Therefore, the obtained values suggest that the two strains (KhD-2^T and *Peptoniphilus* sp. DNF00840^T) belong to the same species. Unlike other *Peptoniphilus* spp., strains KhD-2^T, KHD4^T, and Kh-D5^T ferment ribose and tagatose. The study of their genomes revealed that strain Kh-D2^T had 75 genes associated to carbohydrate metabolism, including 4 genes (1 *rbsA* gene, 2 *rbsR* genes, and 1 *rpiB* gene) encoding proteins involved in fermentation of ribose; the genome from strain KHD4^T contained 61 genes associated to carbohydrate metabolism of which one *rpiB* gene is involved in fermentation of ribose; and strain KhD-5^T had 58 genes associated to carbohydrate metabolism with 3 genes implicated in ribose fermentation (2 *rpiB* genes and 1 *rbsK*) and 1 gene encoding a tagatose biphosphate aldolase enzyme involved in tagatose fermentation. In addition, the genomes of strains Kh-D2^T, KHD4^T, and KhD-5^T also had 25 genes (5 genes encoding proteins responsible for the degradation of histidine, 1 of lysine, 2 of threonine, 12 of methionine, and 5 of arginine), 20 genes (5 of histidine, 1 of lysine, 1 of threonine, 7 of methionine, and 6 of arginine), and 21 genes (14 which degraded methionine, 6 for arginine and 1 for lysine), associated to amino acid degradation, respectively.

Finally, we propose that strains KhD-2^T, KHD4^T, and Kh-D5^T are type strains of *P. vaginalis* sp. nov., *P. raoultii* sp. nov., and *P. pacaensis* sp. nov., respectively.

4.1 | Description of *P. vaginalis* sp. nov

Peptoniphilus vaginalis (va.gi.na'lis. L. n. fem. gen. *vaginalis* from the feminine organ vagina; vaginalis pertaining to the vagina).

Gram-stain—positive. Coccus-shaped bacterium with a mean diameter of 0.66 μm. *Peptoniphilus vaginalis* sp. nov. is a mesophilic bacterium; its optimal growth occurs at temperature 37°C, a pH ranking from 6.5 to 8.5, and a NaCl concentration lower than 5%. Colonies are circular, translucent, gray, and have a diameter of 1–1.5 mm on Columbia agar. Cells are strictly anaerobic, not motile, and non-spore-forming. Catalase, oxidase, and urease activities are negative. Nitrate reduction is also negative nevertheless indole production is positive. *P. vaginalis* shows positive enzymatic activities for acid phosphatase, alkaline phosphatase, esterase, esterase lipase, leucine arylamidase, Naphthol-AS-BI-phosphohydrolase, and *N*-acetyl-β-glucosaminidase. *P. vaginalis* ferments fructose, potassium 5-ketogluconate, ribose, and tagatose. $C_{16:0}$, $C_{14:0}$, $C_{18:1\omega9}$, and $C_{18:2\omega6}$ are its main fatty acids. Strain KhD-2^T is sensitive to amoxicillin, benzylpenicillin, ceftriaxone, imipenem, ertapenem, metronidazole, rifampicin, and vancomycin but resistant to amikacin, erythromycin, and ofloxacin. Its 1,623,601-bp genome contains 34.23% G+C. In EMBL-EBI, the 16S rRNA gene sequence is deposited under accession number LN907856 and the draft genome sequence under accession number FXLP00000000. The type strain of *Peptoniphilus vaginalis* sp. nov. is strain KhD-2^T (=CSUR P0125 = DSM 101742), which was cultured from the vaginal discharge of a woman suffering from bacterial vaginosis.

4.2 | Description of *P. raoultii* sp. nov

Peptoniphilus raoultii (ra.oul'ti.i. N. L. masc. gen. n. *raoultii* of Raoult, to honor French scientist Professor Didier Raoult for his outstanding contribution to medical microbiology).

TABLE 5 Number of genes associated with the 25 general COG functional categories

	P. vaginalis		*P. raoultii*		*P. pacaensis*		
Code	**Value**	**% value**	**Value**	**% value**	**Value**	**% value**	**Description**
J	170	9.70	170	10.69	171	9.78	Translation
A	0	0	0	0	0	0	RNA processing and modification
K	75	4.28	63	3.96	78	4.46	Transcription
L	64	3.65	65	4.09	63	3.60	Replication, recombination, and repair
B	0	0	0	0	0	0	Chromatin structure and dynamics
D	20	1.14	18	1.13	23	1.31	Cell cycle control, mitosis, and meiosis
Y	0	0	0	0	0	0	Nuclear structure
V	61	3.48	40	2.51	60	2.97	Defense mechanisms
T	44	2.51	43	2.70	52	3.64	Signal transduction mechanisms
M	50	2.85	50	3.14	55	3.14	Cell wall/membrane biogenesis
N	7	0.39	7	0.44	8	0.45	Cell motility
Z	0	0	0	0	0	0	Cytoskeleton
W	3	0.17	3	0.18	2	0.11	Extracellular structures
U	15	0.85	16	1.00	15	0.85	Intracellular trafficking and secretion
O	58	3.31	51	3.20	54	3.08	Posttranslational modification, protein turnover, chaperones
X	68	3.88	22	1.38	44	2.51	Mobilome: prophages, transposons
C	83	4.74	66	4.15	75	4.29	Energy production and conversion
G	40	2.28	47	2.95	48	2.74	Carbohydrate transport and metabolism
E	115	6.56	105	6.60	112	6.40	Amino acid transport and metabolism
F	57	3.25	52	3.27	58	3.31	Nucleotide transport and metabolism
H	71	4.05	52	3.27	84	4.80	Coenzyme transport and metabolism
I	56	3.19	53	3.33	45	2.57	Lipid transport and metabolism
P	68	3.88	48	3.02	69	3.94	Inorganic ion transport and metabolism
Q	19	1.08	18	1.13	11	0.62	Secondary metabolites biosynthesis, transport, and catabolism
R	111	6.33	107	6.73	98	5.60	General function prediction only
S	62	3.54	51	3.20	71	4.06	Function unknown
-	547	31.23	541	34.04	573	32.78	Not in COGs

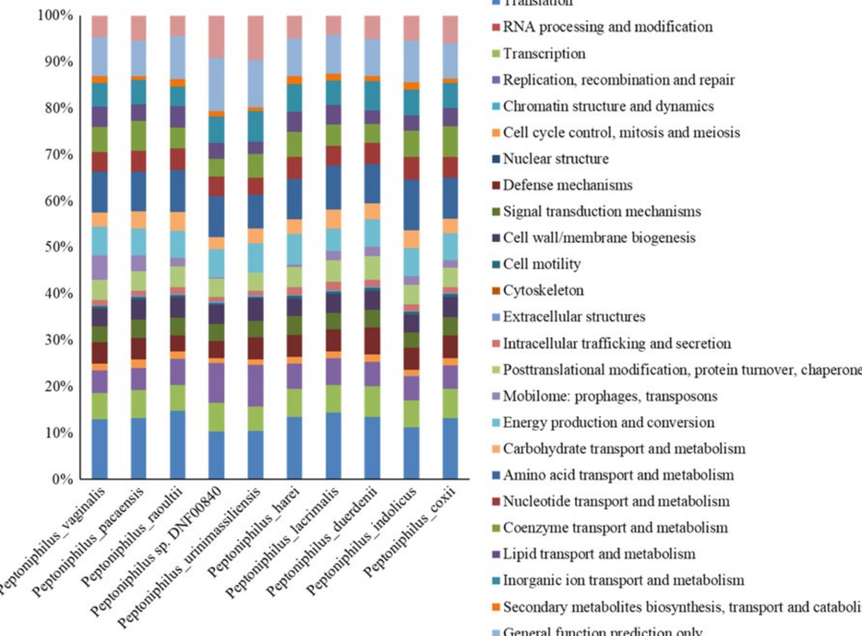


FIGURE 4 Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins of *P. vaginalis* strain KhD-2^T, *P. raoultii* strain KHD4^T, and *P. pacaensis* strain Kh-D5^T among other species

TABLE 6 Genome comparison of closely related species to *P. vaginalis* strain KhD-2^T, *P. raoultii* strain KHD4^T, and *P. pacaensis* strain Kh-D5^T

Species	INSDC identifier[a]	Size (Mbp)	G+C Percent	Gene Content	Number of contigs	N50 Value
***P. vaginalis* KhD-2^T**	**FXLP00000000**	**1.88**	**34.2**	**1,791**	**5**	**707,77**
***P. raoultii* KHD4^T**	**FMWM00000000**	**1.62**	**31.9**	**1,631**	**2**	**1,62**
***P. pacaensis* Kh-D5^T**	**FLQT00000000**	**1.85**	**49.4**	**1,802**	**3**	**1,84**
Peptoniphilus sp. DNF00840	LSDH00000000	1.88	34.3	1,671	91	50,04
Peptoniphilus urinimassiliensis Marseille-P3195	FTPC00000000	1.82	49.7	1,770	5	563,37
Peptoniphilus harei ACS-146-V-Sch2b	AENP00000000	1.84	34.4	1,749	32	111,2
Peptoniphilus lacrimalis CCUG 31350	ARKX00000000	1.85	30.2	1,785	22	190,04
Peptoniphilus duerdenii WAL 18896	AEEH00000000	2.12	34.2	1,963	61	96,77
Peptoniphilus indolicus ATCC 29427	AGBB00000000	2.24	31.7	2,145	302	11,79
Peptoniphilus coxii RMA 16757	LSDG00000000	1.84	44.6	1,783	48	103,89
Peptoniphilus asaccharolyticus DSM 20463	FWWR00000000	2.23	32.3	2,054	17	1,358,172

[a]INSDC: International Nucleotide Sequence Database Collaboration. Text and values in bold have been used to highlight new species.

FIGURE 5 Phylogenetic tree based on whole genome sequence showing the position of *P. vaginalis* strain KhD-2^T, *P. raoultii* strain KHD4^T, and *P. pacaensis* strain Kh-D5^T relative to their nearest neighbors. GenBank accession numbers are indicated in parentheses. Sequences were aligned using Mugsy software, and phylogenetic inferences were performed using the maximum likelihood method with the software FastTree. The scale bar represents a 2% nucleotide sequence divergence

TABLE 7 dDDH values obtained by comparison of all studied genomes using GGDC, Formula 2 (DDH Estimates Based on Identities/HSP length)[a]

	P. vaginalis strain KhD-2^T	*P. raoultii* strain KHD4^T	*P. pacaensis* strain Kh-D5^T	*P. urini-massiliensis*	*P. harei*	*P. lacrimalis*	*P. duerdenii*	*P. indolicus*	*P. coxii*	*P. asaccharolyticus*
P. vaginalis	100 ± 00	22.9 ± 2.35	40.0 ± 2.50	35.3 ± 2.50	45.8 ± 2.60	25.6 ± 2.40	32.0 ± 2.45	22.7 ± 2.40	47.3 ± 2.55	33.20 ± 2.45
P. raoultii		100 ± 00	29.8 ± 2.45	40.5 ± 2.50	19.0 ± 2.25	20.4 ± 2.30	36.4 ± 2.55	22.2 ± 2.35	44.3 ± 2.55	28.40 ± 2.45
P. pacaensis			100 ± 00	45.0 ± 2.60	42.0 ± 2.55	41.9 ± 2.55	38.7 ± 2.50	27.3 ± 2.45	20.7 ± 2.35	29.30 ± 2.45
P. urinimassiliensis				100 ± 00	32.9 ± 2.50	56.4 ± 2.75	42.9 ± 2.50	33.0 ± 2.45	20.1 ± 2.30	32.30 ± 2.45
P. harei					100 ± 00	34.3 ± 2.50	39.2 ± 2.50	20.1 ± 2.30	36.2 ± 2.45	33.30 ± 2.45
P. lacrimalis						100 ± 00	39.3 ± 2.50	25.1 ± 2.40	40.6 ± 2.50	31.90 ± 2.45
P. duerdenii							100 ± 00	24.3 ± 2.35	38.2 ± 2.50	32.80 ± 2.50
P. indolicus								100 ± 00	44.0 ± 2.55	26.70 ± 2.45
P. coxii									100 ± 00	35.40 ± 2.45
P. asaccharolyticus										100 ± 00

[a]The confidence intervals indicate the inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size).

TABLE 8 AAI values obtained by comparison of all studied genomes

	P. raoultii strain KHD4^T	*P. pacaensis* strain Kh-D5^T	*P. urini-massiliensis*	*P. harei*	*P. lacrimalis*	*P. duerdenii*	*P. indolicus*	*P. coxii*	*P. asaccharolyticus*
P. vaginalis	62.7	51.2	51.5	92.9	61.5	57.0	55.9	53.2	57.9
P. raoultii		50.0	50.9	61.6	70.6	56.2	55.4	52.5	56.8
P. pacaensis			92.9	51.8	51.2	51.8	50.4	74.1	50.2
P. urinimassiliensis				52.0	52.7	52.2	51.4	73.4	51.3
P. harei					64.2	58.5	56.4	51.7	58.5
P. lacrimalis						58.0	55.9	51.8	57.1
P. duerdenii							54.7	53.1	57.0
P. indolicus								51.3	84.0
P. coxii									51.2

Gram-stain—positive. Coccus-shaped bacterium with a mean diameter of 0.7 μm. *Peptoniphilus raoultii* sp. nov. is a mesophilic bacterium; its optimal growth occurs at temperature 37°C, a pH ranking from 6.5 to 8.5, and a NaCl concentration lower than 5%. Colonies are circular, translucent, gray, and have a diameter of 1–1.5 mm on Columbia agar. Cells are strictly anaerobic, not motile, and non-spore-forming. Catalase, oxidase, urease, indole, and nitrate activities are negative. *P. raoultii* exhibits positive enzymatic activities for acid phosphatase, esterase, esterase lipase, leucine arylamidase, Naphthol-AS-BI-phosphohydrolase, and *N*-acetyl-β-glucosaminidase. *P. raoultii* ferments potassium 5-ketogluconate, ribose, and tagatose. $C_{16:0}$, $C_{18:2\omega6}$, and $C_{18:1\omega9}$ are its main fatty acids. Strain KHD4^T is sensitive to amoxicillin, benzylpenicillin, ceftriaxone, imipenem, ertapenem, metronidazole, rifampicin, and vancomycin but resistant to amikacin, erythromycin, and ofloxacin. The genome is 1,877,211 bp long and contains 31.87% G+C. In EMBL-EBI, the 16S rRNA gene sequence is deposited under accession number LN998068 and the draft genome sequence under accession number FMWM00000000. Strain KHD4^T (=CSUR P0110 = CECT 9308) is the type strain of *P. raoultii* sp. nov., which was cultured from the vaginal discharge of a woman suffering from bacterial vaginosis.

4.3 | Description of *P. pacaensis* sp. nov

Peptoniphilus pacaensis (pa.ca.en'sis N. L. gen. masc. n. *pacaensis*, from the acronym PACA, of Provence-Alpes-Côte d'Azur, the region where the type strain was first cultured and characterized).

Gram-stain—positive. Coccus-shaped bacterium with a mean diameter of 0.7 μm. *Peptoniphilus pacaensis* sp. nov. is a mesophilic bacterium; its optimal growth occurs at temperature 37°C, a pH ranking from 6.5 to 8.5, and a NaCl concentration lower than 5%. Colonies are circular, translucent, gray, and have a diameter of 1–1.5 mm on Columbia agar. Cells are strictly anaerobic, not motile, and non-spore-forming. Catalase, oxidase, urease, indole, and nitrate activities are negative. *P. pacaensis* shows positive enzymatic activities for alkaline phosphatase, acid phosphatase, esterase, esterase lipase, and Naphthol-AS-BI-phosphohydrolase. *P. pacaensis* ferments potassium 5-ketogluconate, ribose, and tagatose. $C_{16:0}$, $C_{18:2\omega6}$, and $C_{18:1\omega9}$ are its main fatty acids. Strain Kh-D5^T is sensitive to amoxicillin, benzylpenicillin, ceftriaxone, imipenem, ertapenem, metronidazole, rifampicin, and vancomycin but resistant to amikacin, erythromycin, and ofloxacin. Its genome is 1,851,572 bp long with a 49.38% G+C content. In EMBL-EBI, the 16S rRNA gene sequence is deposited under accession number LN998072 and the draft genome sequence under accession number FLQT00000000. The type strain of *P. pacaensis* sp. nov. is strain Kh-D5^T (=CSUR P2270 = DSM 101839), which was cultured from the vaginal discharge of a woman suffering from bacterial vaginosis.

ACKNOWLEDGEMENTS

The authors thank Frederic Cadoret for administrative assistance and the Xegen Company (www.xegen.fr) for automating the genomic annotation process.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Khoudia Diop http://orcid.org/0000-0002-9296-563X
Pierre-Edouard Fournier http://orcid.org/0000-0001-8463-8885

REFERENCES

Afolabi, B. B., Moses, O. E., & Oduyebo, O. O. (2016). Bacterial vaginosis and pregnancy outcome in Lagos, Nigeria. *Open Forum Infectious Diseases*, *3*, ofw030. https://doi.org/10.1093/ofid/ofw030

Alou, M. T., Rathored, J., Michelle, C., Dubourg, G., Andrieu, C., Armstrong, N., … Fournier, P. E. (2017). *Inediibacterium massiliense* gen. nov., sp. nov., a new bacterial species isolated from the gut microbiota of a severely malnourished infant. *Antonie van Leeuwenhoek*, *110*, 737–750. https://doi.org/10.1007/s10482-017-0843-5

Auch, A. F., von Jan, M., Klenk, H.-P., & Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Standards in Genomic Sciences*, *2*, 117–134. https://doi.org/10.4056/sigs.531120

Avguštin, G., Wallace, R. J., & Flint, H. J. (1997). Phenotypic diversity among ruminal isolates of *Prevotella ruminicola*: Proposal of *Prevotella brevis* sp. nov., *Prevotella bryantii* sp. nov., and *Prevotella albensis* sp. nov. and redefinition of *Prevotella ruminicola*. *International Journal of Systematic and Evolutionary Microbiology*, *47*, 284–288.

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

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

Bradshaw, C. S., Tabrizi, S. N., Fairley, C. K., Morton, A. N., Rudland, E., & Garland, S. M. (2006). The association of *Atopobium vaginae* and *Gardnerella vaginalis* with bacterial vaginosis and recurrence after oral metronidazole therapy. *Journal of Infectious Diseases*, *194*, 828–836. https://doi.org/10.1086/506621

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D. R., da Costa, M. S., … Trujillo, M. E. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, *68*, 461–466. https://doi.org/10.1099/ijsem.0.002516

Citron, D. M., Ostovari, M. I., Karlsson, A., & Goldstein, E. J. (1991). Evaluation of the E test for susceptibility testing of anaerobic bacteria. *Journal of Clinical Microbiology*, *29*, 2197–2203.

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

Durand, G. A., Pham, T., Ndongo, S., Traore, S. I., Dubourg, G., Lagier, J. C., … Million, M. (2017). *Blautia massiliensis* sp. nov., isolated from a fresh human fecal sample and emended description of the genus Blautia. *Anaerobe*, *43*, 47–55. https://doi.org/10.1016/j.anaerobe.2016.12.001

Ezaki, T., Kawamura, Y., Li, N., Li, Z.-Y., Zhao, L., & Shu, S. (2001). Proposal of the genera *Anaerococcus* gen. nov., *Peptoniphilus* gen. nov. and *Gallicola* gen. nov. for members of the genus *Peptostreptococcus*. *International Journal of Systematic and Evolutionary Microbiology*, *51*, 1521–1528. https://doi.org/10.1099/00207713-51-4-1521

Fournier, P. E., Lagier, J. C., Dubourg, G., & Raoult, D. (2015). From culturomics to taxonomogenomics: A need to change the taxonomy of prokaryotes in clinical microbiology. *Anaerobe*, *36*, 73–78. https://doi.org/10.1016/j.anaerobe.2015.10.011

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 Heidelberg: Springer. https://doi.org/10.1007/978-3-642-20763-1

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. (2005). FIGENIX: Intelligent automation of genomic annotation: Expertise integration in a new software platform. *BMC Bioinformatics*, *6*, 1.

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*, 1.

Johnson, C. N., Whitehead, T. R., Cotta, M. A., Rhoades, R. E., & Lawson, P. A. (2014). *Peptoniphilus stercorisuis* sp. nov., isolated from a swine manure storage tank and description of *Peptoniphilaceae* fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 3538–3545. https://doi.org/10.1099/ijs.0.058941-0

Käll, L., Krogh, A., & Sonnhammer, E. 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

Klappenbach, J. A., Goris, J., Vandamme, P., Coenye, T., Konstantinidis, K. T., & Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*, *57*, 81–91.

Lagesen, K., Hallin, P., Rodland, 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., Hugon, P., Khelaifia, S., Fournier, P. E., La Scola, B., & Raoult, D. (2015). The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clinical Microbiology Reviews*, *28*, 237–264. https://doi.org/10.1128/CMR.00014-14

Lagier, J. C., Khelaifia, S., Alou, M. T., Ndongo, S., Dione, N., Hugon, P., … Durand, G. (2016). Culture of previously uncultured members of the human gut microbiota by culturomics. *Nature Microbiology*, *12*, 16203. https://doi.org/10.1038/nmicrobiol.2016.203

Lawrence, J. G., & Ochman, H. (1997). Amelioration of bacterial genomes: Rates of change and exchange. *Journal of Molecular Evolution*, *44*, 383–397. https://doi.org/10.1007/PL00006158

Lepargneur, J. P., & Rousseau, V. (2002). Protective role of the Doderleïn flora. *Journal de Gynecologie*, *Obstetrique et Biologie de la Reproduction*, *31*, 485–494.

Li, N., Hashimoto, Y., Adnan, S., Miura, H., Yamamoto, H., & Ezaki, T. (1992). Three new species of the genus *Peptostreptococcus* isolated from humans: *Peptostreptococcus vaginalis* sp. nov., *Peptostreptococcus lacrimalis* sp. nov., and *Peptostreptococcus lactolyticus* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, *42*, 602–605.

Li, J., McCormick, J., Bocking, A., & Reid, G. (2012). Importance of vaginal microbes in reproductive health. *Reproductive Sciences*, *19*, 235–242. https://doi.org/10.1177/1933719111418379

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. https://doi.org/10.1093/nar/25.5.0955

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., … Tang, 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

Martin, D. H., & Marrazzo, J. M. (2016). The vaginal microbiome: Current understanding and future directions. *Journal of Infectious Diseases*, *214*, S36–S41. https://doi.org/10.1093/infdis/jiw184

Matuschek, E., Brown, D. F., & Kahlmeter, G. (2014). Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. *Clinical Microbiology & Infection*, *20*, O255–O266. https://doi.org/10.1111/1469-0691.12373

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, *14*, 1.

Menard, J. P., Fenollar, F., Henry, M., Bretelle, F., & Raoult, D. (2008). Molecular quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. *Clinical Infectious Diseases*, *47*, 33–43. https://doi.org/10.1086/588661

Mishra, A. K., Lagier, J. C., Nguyen, T. T., Raoult, D., & Fournier, P.-E. (2013). Non contiguous-finished genome sequence and description of *Peptoniphilus senegalensis* sp. nov. *Standards in Genomic Sciences*, *7*, 370–381. https://doi.org/10.4056/sigs.3366764

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

Murdoch, D. A. (1998). Gram-positive anaerobic cocci. *Clinical Microbiology Reviews*, *11*, 81–120.

Murdoch, D. A., Mitchelmore, I. J., & Tabaqchali, S. (1994). The clinical importance of gram-positive anaerobic cocci isolated at St Bartholomew's Hospital, London, in 1987. *Journal of Medical Microbiology*, *41*, 36–44. https://doi.org/10.1099/00222615-41-1-36

Murray, P. R., Baron, E. J., Jorgensen, J. H., Landry, M. L., & Pfaller, M. A. (2007). *Manual of clinical microbiology*, 9th ed. Washington, D.C: ASM Press.

Onderdonk, A. B., Delaney, M. L., & Fichorova, R. N. (2016). The human microbiome during bacterial vaginosis. *Clinical Microbiology Reviews*, *29*, 223–238. https://doi.org/10.1128/CMR.00075-15

Pandya, S., Ravi, K., Srinivas, V., Jadhav, S., Khan, A., Arun, A., … Madhivanan, P. (2017). Comparison of culture-dependent and culture-independent molecular methods for characterization of vaginal microflora. *Journal of Medical Microbiology*, *66*, 149–153.

Patel, N. B., Tito, R. Y., Obregón-Tito, A. J., O'Neal, L., Trujillo-Villaroel, O., Marin-Reyes, L., … Lewis, C. M. Jr (2015). *Ezakiella peruensis* gen. nov., sp. nov. isolated from human fecal sample from a coastal traditional community in Peru. *Anaerobe*, *32*, 43–48. https://doi.org/10.1016/j.anaerobe.2014.12.002

Richter, M., & Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, *106*, 19126–19131. https://doi.org/10.1073/pnas.0906412106

Rodriguez-R, L. M., & Konstantinidis, K. T. (2014). Bypassing cultivation to identify bacterial species. *Microbe*, *9*, 111–118.

Rooney, A. P., Swezey, J. L., Pukall, R., Schumann, P., & Spring, S. (2011). *Peptoniphilus methioninivorax* sp. nov., a Gram-positive anaerobic coccus isolated from retail ground beef. *International Journal of Systematic and Evolutionary Microbiology*, *61*, 1962–1967. https://doi.org/10.1099/ijs.0.024232-0

Sasser, M. (2006). *Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME)*. New york, NY: MIDI, Technical Note.

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

Shipitsyna, E., Roos, A., Datcu, R., Hallén, A., Fredlund, H., Jensen, J. S., … Unemo, M. (2013). Composition of the vaginal microbiota in women of reproductive age–sensitive and specific molecular diagnosis of bacterial vaginosis is possible? *PLoS ONE*, *8*(4), e60670. https://doi.org/10.1371/journal.pone.0060670

Srinivasan, S., & Fredricks, D. N. (2008). The human vaginal bacterial biota and bacterial vaginosis. *Interdisciplinary Perspectives on Infectious Diseases*, *2008*, 1–22. https://doi.org/10.1155/2008/750479

Srinivasan, S., Munch, M. M., Sizova, M. V., Fiedler, T. L., Kohler, C. M., Hoffman, N. G., … Fredricks, D. N. (2016). More easily cultivated than identified: Classical isolation with molecular identification of vaginal bacteria. *Journal of Infectious Diseases*, *214*(Suppl 1), S21–S28. https://doi.org/10.1093/infdis/jiw192

Stackebrandt, E., & Ebers, J. (2006). Taxonomic parameters revisited: Tarnished gold standards. *Microbiology Today*, *33*, 152.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, *30*, 2725–2729. https://doi.org/10.1093/molbev/mst197

Thompson, J. D., Higgins, D. G., & 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. https://doi.org/10.1093/nar/22.22.4673

Ulger-Toprak, N., Lawson, P. A., Summanen, P., O'Neal, L., & Finegold, S. M. (2012). *Peptoniphilus duerdenii* sp. nov. and *Peptoniphilus koenoeneniae* sp. nov., isolated from human clinical specimens. *International Journal of Systematic and Evolutionary Microbiology*, *62*, 2336–2341. https://doi.org/10.1099/ijs.0.031997-0

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K. H., … Rosselló-Móra, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology*, *12*, 635–645. https://doi.org/10.1038/nrmicro3330

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

How to cite this article: Diop K, Diop A, Michelle C, et al. Description of three new *Peptoniphilus* species cultured in the vaginal fluid of a woman diagnosed with bacterial vaginosis: *Peptoniphilus pacaensis* sp. nov., *Peptoniphilus raoultii* sp. nov., and *Peptoniphilus vaginalis* sp. nov. *MicrobiologyOpen*. 2018;e661. https://doi.org/10.1002/mbo3.661

Article 10:

Characterization of a novel Gram-stain-positive anaerobic coccus isolated from the female genital tract: Genome sequence and description of *Murdochiella vaginalis* sp. nov.

Khoudia Diop, Awa Diop, Saber Khelaifia, Catherine Robert, Fabrizio Di Pinto, Jérémy Delerce, Didier Raoult, Pierre-Edouard Fournier, Florence Bretelle, Florence Fenollar

Published in MicrobiologyOpen

Received: 29 August 2017 | Revised: 3 November 2017 | Accepted: 7 November 2017

DOI: 10.1002/mbo3.570

ORIGINAL RESEARCH

Characterization of a novel Gram-stain-positive anaerobic coccus isolated from the female genital tract: Genome sequence and description of *Murdochiella vaginalis* sp. nov.

Khoudia Diop[1] | Awa Diop[1] | Saber Khelaifia[1] | Catherine Robert[1] | Fabrizio Di Pinto[1] | Jérémy Delerce[1] | Didier Raoult[1,2] | Pierre-Edouard Fournier[1] | Florence Bretelle[1,3] | Florence Fenollar[1]

[1]Aix-Marseille Univ, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM 63, CNRS UMR 7278, IRD 198, INSERM U1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de Médecine, Marseille, France

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

[3]Department of Gynecology and Obstetrics, Gynépole, Marseille, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, Marseille, France

Correspondence
Florence Fenollar, Aix-Marseille Univ, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM 63, CNRS UMR 7278, IRD 198, INSERM U1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de Médecine, Marseille, France.
Email: florence.fenollar@univ-amu.fr

Funding information
This study was supported by the Fondation Méditerranée Infection and the French National Research Agency under the “Investissements d’avenir” program, reference ANR-10-IAHU-03

Abstract

Strain Marseille-P2341^T, a nonmotile, nonspore-forming, Gram-stain-positive anaerobic coccus, was isolated in the vaginal specimen of a patient with bacterial vaginosis using culturomics. Its growth occurred at temperatures ranging from 25 to 42°C, with pH between 6.5 and 8.5, and at NaCl concentrations lower than 5%. The major fatty acids were $C_{18:1n9}$ (27.7%) and $C_{16:0}$ (24.4%). Its genome is 1,671,491 bp long with 49.48 mol% of G+C content. It is composed of 1,501 genes: 1,446 were protein-coding genes and 55 were RNAs. Strain Marseille-P2341^T shared 97.3% of 16S rRNA gene sequence similarity with *Murdochiella asaccharolytica*, the phylogenetically closest species. These results enabled the classification of strain Marseille-P2341^T as a new species of the genus *Murdochiella* for which we proposed the name *Murdochiella vaginalis* sp. nov. The type strain is strain Marseille-P2341^T (=DSM 102237, =CSUR P2341).

KEYWORDS
bacterial vaginosis, culturomics, genome, *Murdochiella vaginalis*, taxono-genomics, vaginal microbiota

1 | INTRODUCTION

Due to vaginal secretions and, sometimes, urine, the vagina is a humid biotope which constitutes a complex ecosystem colonized by several types of microorganisms (Pal et al., 2011). Its composition was described for the first time in 1892 by Döderlein, who revealed that the vaginal flora is homogeneous and composed of Gram-positive bacteria known as Döderlein bacilli (Lepargneur & Rousseau, 2002). Since then, many studies have been conducted, some of which suggest that this complex ecosystem is mostly dominated by the *Lactobacillus* genus (De Vos et al., 2009) with four main species: *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus jensenii*, and *Lactobacillus vaginalis*. This constitutes the first line of defense against genital infections (Bohbot & Lepargneur, 2012; Turovskiy, Sutyak

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2018 The Authors. *MicrobiologyOpen* published by John Wiley & Sons Ltd.

Noll, & Chikindas, 2011). An imbalance in this flora is observed in bacterial vaginosis.

The vaginal microflora diversity of a patient suffering from bacterial vaginosis was first described by Schröder in 1921 (Pal et al., 2011). This dysbiosis is characterized by a progressive decrease or even a lack of normal *Lactobacillus* flora accompanied by an increased pH of the vaginal lumen and an abnormal proliferation of previously underrepresented bacteria and Gram-stain-negative anaerobic bacteria (*Gardnerella vaginalis*, *Atopobium vaginae*, *Mobiluncus curtisii*, etc.) (Pépin et al., 2011; Shipitsyna et al., 2013). The mechanism of bacterial vaginosis is unknown; its empirical treatment and relapse rate is estimated at 50% at 3 months (Bretelle et al., 2015). This disturbance is associated with some complications in pregnant women such as miscarriage, chorioamnionitis, and preterm birth (Bretelle et al., 2015; Svare, Schmidt, Hansen, & Lose, 2006).

Initially studied using conventional culture methods, the understanding of the human vaginal microbiota was enhanced through the use of molecular techniques involving sequencing and phylogenetic analysis of the 16S rRNA gene (Lamont et al., 2011). These molecular methods enabled the detection of fastidious and uncultured bacteria such as bacterial vaginosis-associated bacteria (BVAB): BVAB1 BVAB2, and BVAB3 (Fredricks, Fiedler, & Marrazzo, 2005). In order to identify all bacteria (uncultured and fastidious) present in the vagina and involved in this alteration, we studied normal vaginal flora and those from bacterial vaginosis using the concept of "microbial culturomics," based on the multiplication of culture conditions with variations in temperature, media, pH, and atmospheric conditions, and rapid bacterial identification using matrix-assisted laser-desorption/ionization (MALDI) time-of-flight (TOF) mass spectrometry (MS) (Lagier et al., 2012, 2015). This microbial culturomics approach enabled us to isolate a new member of the *Murdochiella* genus that did not correspond to other species of this genus. This strain is designated as Marseille-P2341^T. The *Murdochiella* genus was created in 2010, to include strain recovered from a human abdominal wall abscess and in a sacral pilonidal cyst aspirate (Ulger-Toprak, Liu, Summanen, & Finegold, 2010). This genus has only one valid species: *Murdochiella asaccharolytica*.

The description of new bacterial species is based on phenotypic and genotypic characteristics but has some limitations (Chan, Halachev, Loman, Constantinidou, & Pallen, 2012; Vandamme et al., 1996). In this manuscript we use taxonogenomics, a new approach combining classic characteristics with the proteomic information obtained from MALDI-TOF MS and the description of the annotated whole genome (Fournier & Drancourt, 2015; Fournier, Lagier, Dubourg, & Raoult, 2015), to describe *Murdochiella vaginalis* sp. nov. (=DSM 102237 = CSUR P2341).

2 | MATERIALS AND METHODS

2.1 | Sample ethics and strain isolation

Using a Sigma Transwab (Medical Wire, Corsham, United Kingdom), the vaginal specimen of a 33-year-old French woman was collected

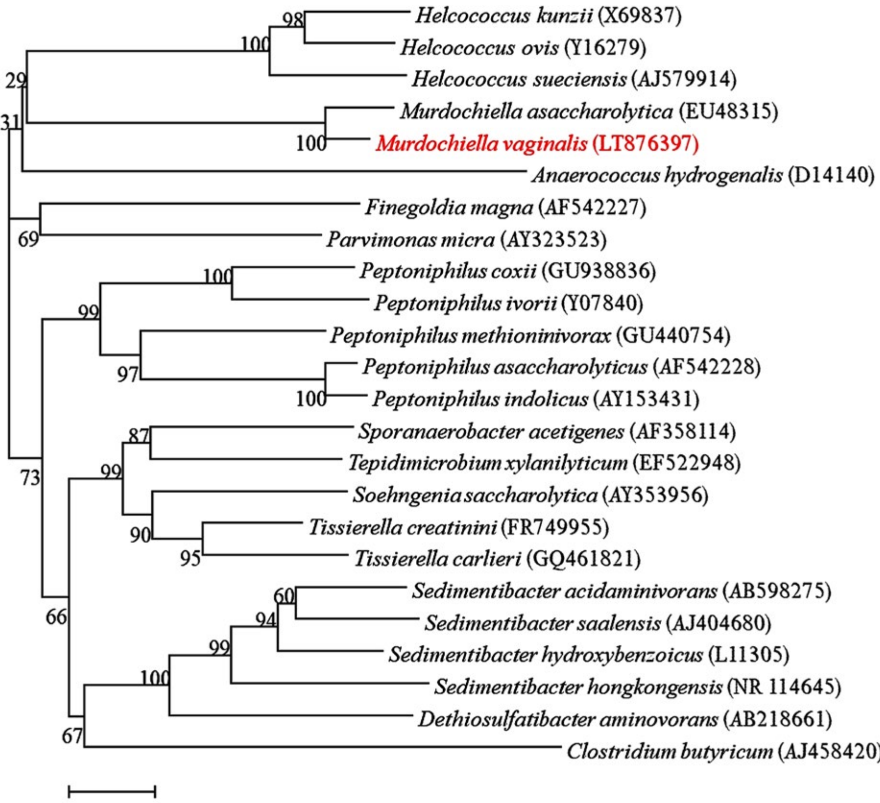


FIGURE 1 Phylogenetic tree highlighting the position of *Murdochiella vaginalis* strain Marseille-P2341^T relative to other close strains. GenBank accession numbers of each 16S rRNA are noted in parenthesis. Sequences were aligned using Muscle v3.8.31 with default parameters and phylogenetic inferences were obtained using the neighbor-joining method with 500 bootstrap replicates, within MEGA6 software. The scale bar represents a 2% nucleotide sequence divergence

TABLE 1 Classification and general features of *Murdochiella vaginalis* Marseille-P2341^T

Properties	Terms
Taxonomy	**Kingdom**: *Bacteria*
	Phylum: *Firmicutes*
	Class: *Clostridia*
	Order: *Clostridiales*
	Family: *Peptoniphiliaceae*
	Genus: *Murdochiella*
	Species: *M. vaginalis*
Type strain	Marseille-P2341^T
Isolation site	Human vagina
Isolation country	France
Gram stain	Positive
Cell shape	Coccus
Motility	No
Oxygen requirements	Anaerobic
Optimal temperature	37°C
Temperature range	Mesophilic

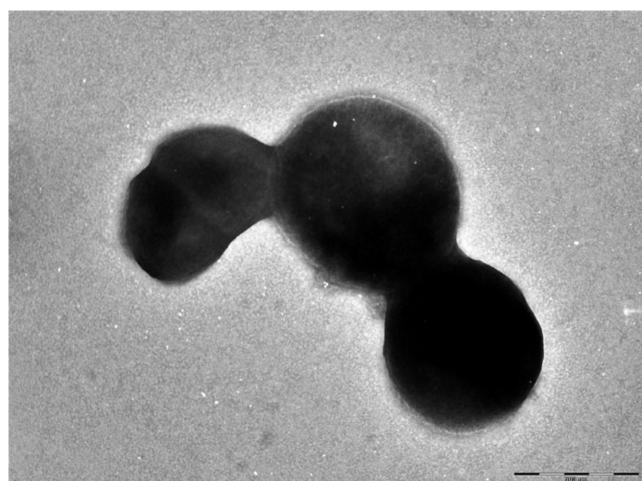
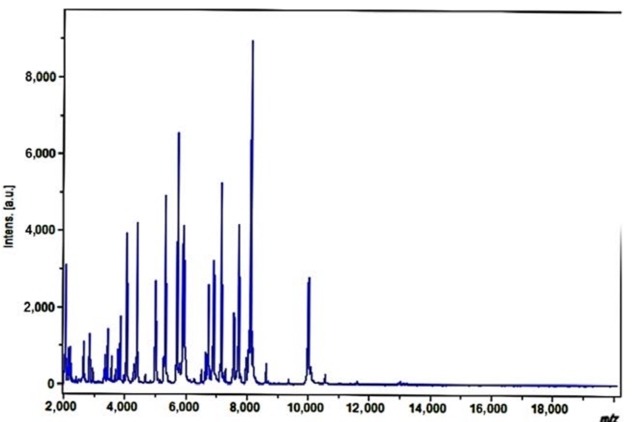


FIGURE 3 Transmission electron microscopy of *Murdochiella vaginalis* strain Marseille-P2341^T, using a Tecnai G20 transmission electron microscope (FEI Company). The scale bar represents 100 nm

and transported to the La Timone hospital in Marseille (France). Diagnosed as previously reported (Menard, Fenollar, Henry, Bretelle, & Raoult, 2008), the patient was suffering from bacterial vaginosis. At the time the sample was collected, she was not being treated with any antibiotics. The study was authorized by the local IFR48 ethics committee (Marseille, France) under agreement number 09-022 and the patient also signed written consent. After sampling, the specimen was preincubated in a blood culture bottle (BD Diagnostics, Le Pont-de-Claix, France) enriched with 4 ml of rumen that was filter-sterilized through a 0.2 μm pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France) and 3 ml of sheep's blood (bioMérieux, Marcy l'Etoile, France). After different preincubation periods (1, 3, 7, 10,15, 20, and 30 days), 50 μl of the supernatant was inoculated on Schaedler agar (BD Diagnostics) and then incubated for 7 days under anaerobic conditions at 37°C.

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

Isolated colonies were deposited in duplicate on a MTP 96 MALDI-TOF target plate (Bruker Daltonics, Leipzig, Germany) for identification with a microflex spectrometer (Bruker), as previously described (Seng et al., 2009). All obtained protein spectra were loaded into the MALDI Biotyper Software (Bruker Daltonics) and compared, as previously described (18), using the standard pattern-matching algorithm, which compared the acquired spectrum with those present



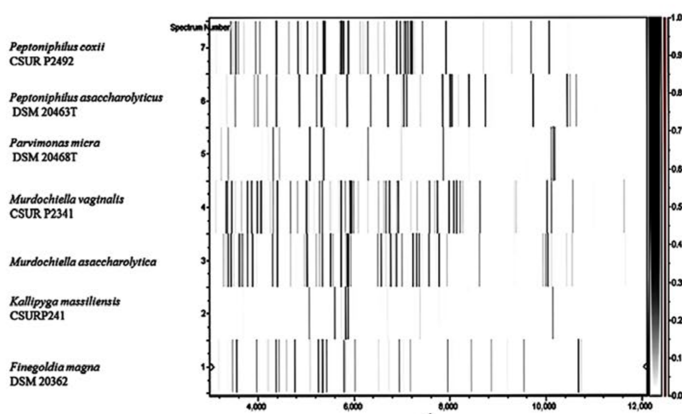


FIGURE 2 MALDI-TOF information. (**a**) Reference mass spectrum from *Murdochiella vaginalis* strain Marseille-P2341T spectra. (**b**) Gel view comparing *M. vaginalis* strain Marseille-P2341T to other species within Peptoniphilaceae family. The gel view displays the raw spectra of loaded spectrum files arranged with a pseudo-gel like appearance. The *x*-axis records the *m*/*z* value. The left *y*-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a gray scale scheme code. The right *y*-axis indicates the relation between the color of a peak and its intensity, in arbitrary units. Displayed species are indicated on the left

TABLE 2 Cellular fatty acid composition (%)

Fatty acids	Name	Mean relative % (a)
18:1n9	9-Octadecenoic acid	27.7 ± 6.6
16:0	Hexadecanoic acid	24.2 ± 4.1
18:2n6	9,12-Octadecadienoic acid	15.7 ± 4.4
18:0	Octadecanoic acid	13.4 ± 2.2
14:0	Tetradecanoic acid	5.9 ± 7.0
18:1n7	11-Octadecenoic acid	3.7 ± 0.6
15:0 iso	13-methyl-tetradecanoic acid	1.4 ± 1.7
17:0	Heptadecanoic acid	1.0 ± 0.1
14:0 3-OH	3-hydroxy-Tetradecanoic acid	TR
20:0	Eicosanoic acid	TR
18:0 9,10-methylene	2-octyl-Cyclopropaneoctanoic acid	TR
5:0 iso	3-methyl-butanoic acid	TR
20:4n6	5,8,11,14-Eicosatetraenoic acid	TR
15:0	Pentadecanoic acid	TR
16:1n5	11-Hexadecenoic acid	TR
17:0 anteiso	14-methyl-Hexadecanoic acid	TR
17:0 iso	15-methyl-Hexadecanoic acid	TR
20:1n9	11-Eicosenoic acid	TR
15:0 anteiso	12-methyl-tetradecanoic acid	TR
17:1n7	10-Heptadecenoic acid	TR
10:0	Decanoic acid	TR
20:2n6	11,14-Eicosadienoic acid	TR
12:0	Dodecanoic acid	TR
19:0	Nonadecanoic acid	TR
22:5n2	7,10,13,16,19-docosapentaenoic acid	TR
16:0 9,10-methylene	2-Hexyl-Cyclopropaneoctanoic acid	TR
13:0	Tridecanoic acid	TR
4:0	Butanoic acid	TR
22:6n3	4,7,10,13,16,19-Docosahexaenoic acid	TR

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

in the library (the Bruker database and our constantly updated database). If the score was greater than 1.9, the bacterium was considered to be identified at the species level. If not, identification failed and to achieve identification for unidentified colonies, the 16S rRNA gene was sequenced using fD1-rP2 primers (Eurogentec, Angers, France) and the obtained sequence was matched against the NCBI database using the BLAST algorithm (Drancourt et al., 2000). As suggested, if the 16S rRNA gene sequence similarity value was lower than 95% or 98.7%, the strain was defined as a new genus or species, respectively (Kim, Oh, Park, & Chun, 2014; Stackebrandt & Ebers, 2006).

2.3 | Phylogenetic analysis

All species from the same order of the new species were retrieved and 16S sequences were download from NCBI, by parsing NCBI eUtils results and the NCBI taxonomy page. Sequences were aligned using CLUSTALW, with default parameters and phylogenetic inferences obtained using the neighbor-joining method with 500 bootstrap replicates, within MEGA6 software.

2.4 | Growth conditions and morphological observation

To evaluate ideal growth, the strain Marseille-P2341^T was cultivated on Columbia agar with 5% sheep's blood (bioMérieux) and incubated at different temperatures (25, 28, 37, 45, and 56°C) in an aerobic atmosphere with or without 5% CO_2, and in anaerobic and microaerophilic atmospheres, using GENbag Anaer and GENbag microaer systems (bioMérieux). The salinity and pH conditions were also tested at different concentrations of NaCl (0%, 5%, 15%, and 45%) and different pH (5, 6, 6.5, 7, and 8.5).

Oxidase and catalase tests, Gram-stain, motility, and sporulation were performed using standard procedures (Murray, Baron, Jorgensen, Landry, & Pfaller, 2007). To observe cell morphology, they were fixed with 2.5% glutaraldehyde in 0.1 mol/L cacodylate buffer for at least 1 hr at 4°C. A drop of cell suspension was then 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 s with 1% ammonium molybdate solution in filtered water at RT. Electron micrographs were acquired using a Tecnai G20 Cryo (FEI) transmission electron microscope operated at 200 keV.

2.5 | Biochemical and antibiotic susceptibility tests

Biochemical tests were performed using API ZYM, API 20A, and API 50CH strips (bioMérieux) according to the manufacturer's instructions. The strips were incubated for 4, 24, and 48 hr respectively.

Cellular fatty acid methyl ester (FAME) analysis was performed using Gas Chromatography/Mass Spectrometry (GC/MS). Strain Marseille-P2341^T was grown on Columbia agar enriched with 5% sheep's blood (bioMérieux). Two samples were then prepared with approximately 50 mg of bacterial biomass per tube harvested from several culture plates. Fatty acid methyl esters were prepared as described by Sasser (Sasser, 2006). GC/MS analyses were carried out as previously described (Dione et al., 2016). In brief, fatty acid methyl esters were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500—SQ 8 S, Perkin Elmer, Courtaboeuf, France). A spectral database search was performed using MS Search 2.0 operated with the Standard Reference Database 1A (NIST, Gaithersburg, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK).

Antibiotic susceptibility was tested using the disc diffusion method (Le Page et al., 2015). The results were read using Scan 1200 (Interscience, Saint-Nom-la-Bretèche, France).

TABLE 3 Differential characteristics of *Murdochiella vaginalis* and the phylogenetically related species. *Murdochiella vaginalis* strain Marseille-P2341^T, *Murdochiella asaccharolytica* strain WAL 1855C^T, *Finegoldia magna* strain CCUG 17636^T, *Peptoniphilus indolicus* ATCC 29427^T, *Parvimonas micra* CCUG 46357^T, *Helcococcus sueciensis* CCUG 47334^T, and *Anaerococcus hydrogenalis* JCM 7635^T

Properties	*M. vaginalis*	*M. asaccharolytica*	*F. magna*	*P. indolicus*	*P. micra*	*H. sueciensis*	*A. hydrogenalis*
Cell diameter (µm)	0.6–0.8	0.5–0.6	0.8–1.6	0.7–1.6	0.3–0.7	na	0.7–1.8
Oxygen requirement	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Facultative anaerobic	Anaerobic
DNA G+C content (mol%)	49.5	na	na	31.69	28.65	29.5	29.64
Production of							
Alkaline phosphatase	−	−	Variable	+	+	+	−
Indole	−	−	−	+	−	−	+
Catalase	−	−	Variable	na	Variable	−	−
Nitrate reductase	−	−	−	+	−	−	−
Urease	−	−	−	−	−	−	Variable
β-galactosidase	+	−	−	−	−	−	-
N-acetyl-glucosamine	+	−	−	na	−	+	na
Acid from							
Mannose	+	−	−	−	−	−	+
Glucose	+	−	−	−	−	+	+
Lactose	−	−	−	−	−	+	+
Raffinose	−	−	−	−	−	−	+
Habitat	Vaginal discharges	Human wound	Human specimen	Summer mastitis of cattle	Human specimen	Human wound	Vaginal discharges

+, positive reaction; −, negative reaction; na, data not available.

2.6 | Genomic DNA preparation

Genomic DNA (gDNA) of strain Marseille-P2341^T was extracted in two steps: a mechanical treatment was first performed using acid-washed glass beads (G4649-500 g Sigma) and a FastPrep BIO 101 instrument (Qbiogene, Strasbourg, France) at maximum speed (6.5) for 3 × 30 s. Then after 2 hr of lysozyme incubation at 37°C, DNA was extracted on the EZ1 biorobot (Qiagen, Hilden, Germany) using the EZ1 DNA tissue kit. The elution volume was 50 µl. The gDNA was quantified by a Qubit assay using the high sensitivity kit (Life technologies, Carlsbad, CA, USA) to 103 ng/µl.

TABLE 4 Nucleotide content and gene count levels of the genome

Attribute	Value	% of total[a]
Size (bp)	1,671,491	100
G+C content (bp)	827,028	49.48
Coding region (bp)	1,511,436	90.42
Total genes	1,501	100
RNA genes	55	3.66
Protein-coding genes	1,446	100
Genes with function prediction	1,056	73.03
Genes assigned to COGs	965	66.74
Genes with peptide signals	160	11.06
Genes with transmembrane helices	369	25.52

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

2.7 | Genome sequencing and assembly

gDNA was sequenced on the MiSeq Technology (Illumina Inc, San Diego, CA, USA) using the mate pair strategy. The gDNA was barcoded using the Nextera Mate Pair sample prep kit (Illumina) to be mixed with 11 other projects. The mate pair library was prepared with 1.5 µg of genomic DNA using the Nextera mate pair Illumina guide. The genomic DNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of fragmentation was validated on 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 to 11 kb with an optimal size at 3.716 kb. No size selection was performed and 652 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with a bi-modal pattern at 644 bp and 1,613 bp on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc, Santa Clara, CA, USA) and the final concentration library was measured at 53.40 nmol/L.

The libraries were normalized at 2 nmol/L and pooled. After a denaturation step and dilution at 15 pmol/L, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 39-hr run in a 2 × 251-bp.

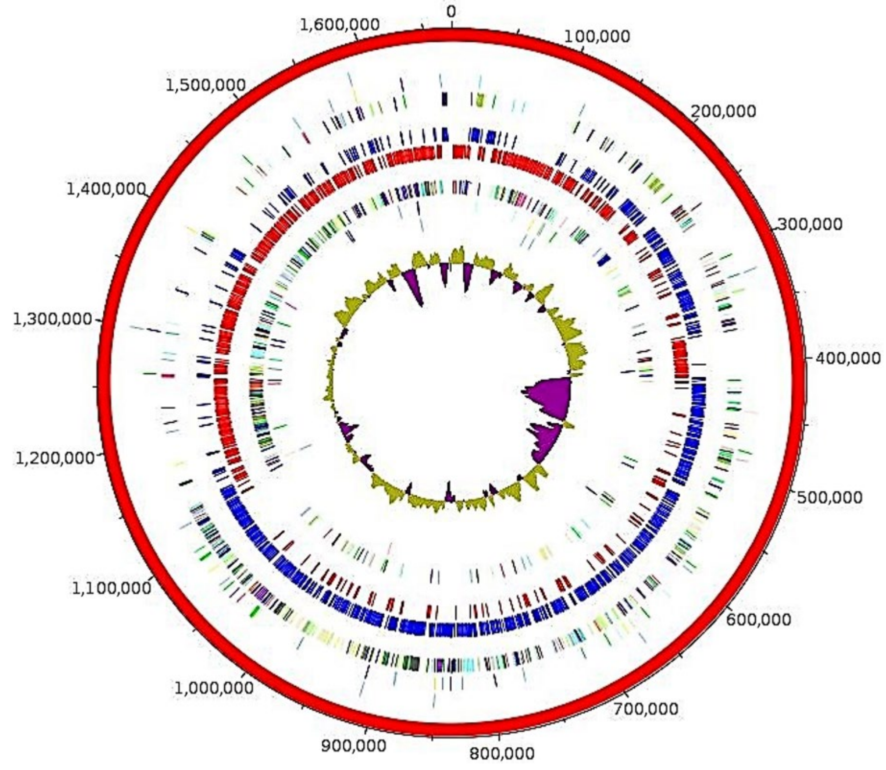


FIGURE 4 Graphical circular map of the genome. From outside to the center: Contigs (red/gray), COG category of genes on the forward strand (three circles), genes on forward strand (blue circle), genes on the reverse strand (red circle), COG category on the reverse strand (three circles), GC content

TABLE 5 Number of genes associated with the 25 general COG functional categories

Code	Value	% of total	Description
[J]	157	10.857538	Translation
[A]	0	0	RNA processing and modification
[K]	71	4.910097	Transcription
[L]	57	3.9419088	Replication, recombination and repair
[B]	0	0	Chromatin structure and dynamics
[D]	16	1.1065007	Cell cycle control, mitosis and meiosis
[Y]	0	0	Nuclear structure
[V]	45	3.1120331	Defense mechanisms
[T]	32	2.2130015	Signal transduction mechanisms
[M]	44	3.042877	Cell wall/membrane biogenesis
[N]	4	0.2766252	Cell motility
[Z]	0	0	Cytoskeleton
[W]	1	0.0691563	Extracellular structures
[U]	15	1.0373445	Intracellular trafficking and secretion
[O]	53	3.6652837	Post-translational modification, protein turnover, chaperones
[X]	8	0.5532504	Mobilome: prophages, transposons
[C]	60	4.149378	Energy production and conversion
[G]	81	5.60166	Carbohydrate transport and metabolism
[E]	80	5.5325036	Amino acid transport and metabolism
[F]	51	3.526971	Nucleotide transport and metabolism
[H]	52	3.5961275	Coenzyme transport and metabolism
[I]	34	2.351314	Lipid transport and metabolism
[P]	46	3.1811898	Inorganic ion transport and metabolism
[Q]	9	0.62240666	Secondary metabolites biosynthesis, transport and catabolism
[R]	92	6.3623796	General function prediction only
[S]	42	2.9045644	Function unknown
_	481	33.26418	Not in COGs

In total, 9.2 Gb of information was obtained from a 1042 K/mm^2 cluster density with a cluster passing quality control filters of 91.6% (18,078,000 passing filter paired reads). Within this run, the index representation for strain Marseille-P2341^T was determined to 13.14%. The 2,375,075 paired reads were trimmed then assembled in a scaffold.

2.8 | Genome annotation and analysis

Prodigal was used for open reading frame (ORF) prediction (Hyatt et al., 2010) with default parameters. We excluded predicted ORFs spanning a sequencing gap region (containing *N*). The bacterial proteome was predicted using BLASTP (E-value 1e^{-03} coverage 0.7 and identity percent 30) against the Clusters of Orthologous Groups (COG) database. If no hit was found, we searched against the NR database (Clark, Karsch-Mizrachi, Lipman, Ostell, & Sayers, 2016), using BLASTP with E-value of 1e^{-03} coverage 0.7 and an identity percent of 30. An E-value of 1e^{-05} was used if sequence lengths were shorter than 80 amino acids. PFam conserved domains (PFAM-A an PFAM-B domains) were searched on each protein with the hhmscan tools analysis. RNAmmer (Lagesen et al., 2007) and tRNAScanSE tools (Lowe & Eddy, 1997) were used to find ribosomal RNAs genes and tRNA genes, respectively. ORFans were identified if all the BLASTP performed had negative results (E-value smaller than 1e^{-03} for ORFs with sequence size above 80 aa or E-value smaller than 1e^{-05} for ORFs with sequence length below 80 aa). For data management and visualization of genomic features, Artemis (Carver, Harris, Berriman, Parkhill, & McQuillan, 2012) and DNA Plotter (Carver, Thomson, Bleasby, Berriman, & Parkhill, 2009) were used, respectively. We used the home-made MAGI software to analyze the mean level of nucleotide sequence similarity at the genome level. It calculated the average genomic identity of gene sequences (AGIOS) among compared genomes (Ramasamy et al., 2014). This software combines the Proteinrtho software (Lechner et al., 2011) for detecting orthologous proteins in pairwise genomic comparisons. The corresponding genes were then retrieved and the mean percentage of nucleotide sequence identity among orthologous ORFs was determined using the Needleman-Wunsch global alignment algorithm. The Multi-Agent software system DAGOBAH (Gouret et al., 2011) was used to perform the annotation and comparison processes, which included Figenix (Gouret et al., 2005) libraries for pipeline analysis. We also performed GGDC analysis using the GGDC web server, as previously reported (Meier-Kolthoff, Auch, Klenk, & Göker, 2013).

3 | RESULTS

3.1 | Strain identification

Strain Marseille-P2341^T was first isolated after 15 days of preincubation of a vaginal sample in a blood culture bottle supplemented with rumen and sheep's blood under anaerobic conditions and then sub-cultured on Schaedler agar. A score of 1.3 was also obtained with MALDI-TOF MS identification, suggesting that this isolate was not in the database. The 16S rRNA gene sequence (accession number LT576397) of the strain exhibited 97.3% nucleotide sequence similarity with *M. asaccharolytica*, the phylogenetically-closest species with a validly published name (Figure 1). As this value was lower than 98.7%, the threshold recommended for delineating a new species (Kim et al., 2014; Stackebrandt & Ebers, 2006), strain Marseille-P2341^T was classified as a new species named *M. vaginalis* (Table 1). The reference

TABLE 6 Genome comparison of closely related species to *Murdochiella vaginalis* strain Marseille-P2341^T

Species	INSDC identifier	Size (Mb)	G+C (mol%)	Gene Content
M. vaginalis strain Marseille-P2341^T	LT632322	1.671	49.48	1,501
Anaerococcus hydrogenalis DSM 7454	ABXA00000000.1	1.89	29.64	2,069
Helcococcus kunzii NCFB 2900	AGEI00000000.1	2.10	29.35	1,882
Peptoniphilus indolicus ATCC 29427	AGBB00000000.1	2.24	31.69	2,269
Helcococcus sueciensis CCUG 47334	AUHK00000000.1	1.57	28.40	1,445
Peptoniphilus coxii RMA 16757	LSDG00000000.1	1.84	44.62	1,86
Parvimonas micra ATCC 33270	ABEE00000000.2	1.70	28.65	1,678

INSDC, International Nucleotide Sequence Database Collaboration.

spectrum of the strain Marseille-P2341^T (Figure 2a) was then added to our database and compared to other known species of the family *Peptoniphilaceae* (Johnson, Whitehead, Cotta, Rhoades, & Lawson, 2014). Their differences are shown in the gel view which was obtained (Figure 2b).

3.2 | Phenotypic characteristics

Only grown in anaerobic conditions, strain Marseille-P2341^T grows at temperatures between 25 to 42°C, with optimal growth at 37°C after 48 hr of incubation. It needs NaCl concentrations lower than 5 g/L and a pH ranging from 6.5 to 8.5. After 2 days of incubation at 37°C under anaerobic conditions on Columbia agar (bioMérieux), colonies are circular, white, and opaque with a diameter of 2–2.5 mm. Gram-staining shows a Gram-positive coccus. Individual cells show a diameter ranging from 0.6 to 0.8 μm under an electron microscope (Figure 3). Nonmotile and nonspore-forming, strain Marseille-P2341^T exhibited positive oxidase activity. Nevertheless, catalase activity was negative and nitrate was not reduced.

Using an API ZYM strip, positive reactions were observed for leucine arylamidase, Naphtol-AS-BI-phosphohydrolase, α and β-galactosidase, glucosidase (α and β), N-acetyl-β-glucosaminidase, α-mannosidase, and α-fucosidase. Alkaline phosphatase, lipases, and other reactions were negative. On an API 20A strip, we observed an acidification of glucose and an API 50CH strip revealed that only galactose, glucose, mannose, and potassium 5-ketogluconate were metabolized. All the other reactions were negative on both API strips. The most abundant fatty acids found were 9-Octadecenoic acid and Hexadecanoic acid (28% and 24%, respectively). Interesting minor fatty acids (<1%) are also described (Table 2). Cells were susceptible to oxacillin, penicillin, ceftriaxone, ciprofloxacin, clindamycin, doxycycline, erythromycin, fosfomycin, gentamycin, trimethoprim-sulfamethoxazole, rifampicin, and vancomycin but resistant to colistin. The phenotypic characteristics of strain Marseille-P2341^T were compared to those of closely related species and are summarized in Table 3 (Collins, 2004; Ezaki et al., 2001; Ezaki, Yamamoto, Ninomiya, Suzuki, & Yabuuchi, 1983; Murdoch & Shah, 1999; Tindall & Euzeby, 2006; Ulger-Toprak et al., 2010).

3.3 | Genome properties

The genome measures 1,671,491 bp long and has 49.48 mol% of G+C content (Table 4, Figure 4). It is composed of one scaffold composed of one contig. Of the 1,501 predicted genes, 1,446 were protein-coding genes and 55 were RNAs (two genes were 5S rRNA, two genes were 16S rRNA, two genes were 23S rRNA, 49 genes were tRNA genes). A total of 1,056 genes (73.03%) were assigned a putative function (by cogs or by NR blast). 56 genes were identified as ORFans (3.87%). The remaining 292 genes were annotated as hypothetical proteins (20.19%). Genome statistics are summarized in Table 4 and the distribution of the genes in COGs functional categories is presented in Table 5.

3.4 | Genomic comparison

The comparison of the genome of our species with the closest related species (Table 6) reveals that the genome sequence of strain Marseille-P2341^T (1.67 Mbp) is larger than that of *Helcococcus sueciensis* (1.57 Mbp), but smaller than those of *Parvimonas micra*, *Peptoniphilus coxii*, *Anaerococcus hydrogenalis*, *Helcococcus kunzii*, and *Peptoniphilus indolicus* (1.70, 1.84, 1.89, 2.10, and 2.24, respectively). The G+C content of strain Marseille-P2341 ^{T} (49.48 mol%) is greater than those of all compared species. The gene content of strain Marseille-P2341^T (1,446) is almost equal to that of *H. sueciensis* but is smaller than those of other compared genomes. However, in all the compared genomes, the distribution of genes in COG categories was similar. Nevertheless, there are fewer genes of *M. vaginalis* present in the COG categories X (Mobilome: prophages, transposons) and W (Extracellular structures) than other compared species (Figure 5). Moreover, the AGIOS analysis shows that strain Marseille-P2341^T shares between 509 and 542 orthologous genes with closely related species (Table 7) and analysis of the average percentage of nucleotide sequence identity ranged from 50.8% to 56.4% with *P. micra* and *H. sueciensis*, respectively (Table 7). In addition, the digital DNA-DNA hybridization (dDDH) of strain Marseille-P2341^T and its closest species varied between 22.40% to 36% with 22.40, 23.60, 23.70, 25.50, 25.90, and 36% for *H. kunzii*, *A. hydrogenalis*, *P. micra*, *P. coxii*, *H. sueciensis*, and *P. indolicus*, respectively. Unfortunately, *M. asaccharolytica* was not included in this comparison because its genome was not sequenced.

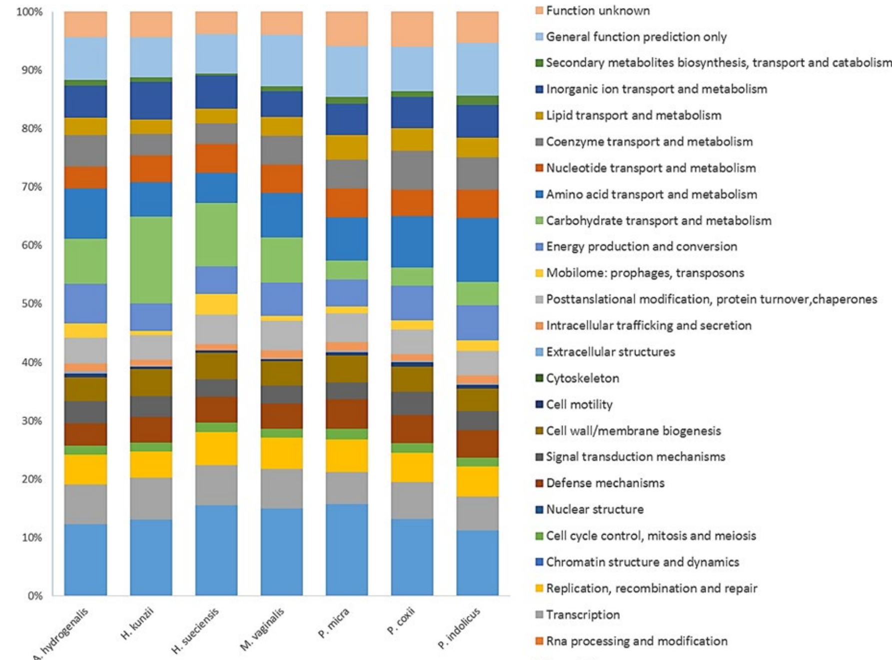


FIGURE 5 Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins of *Murdochiella vaginalis* strain Marseille-P2341T among other species

4 | DISCUSSION

During the study of vaginal microbiota using culturomics, with the aim of exploring the vaginal flora as exhaustively as possible and identifying the bacteria involved in bacterial vaginosis in order to better manage this infection, strain Marseille-P2341^T was identified in the vaginal sample of a patient suffering from bacterial vaginosis. Its phenotypic characteristics, MALDI-TOF MS, 16S rRNA gene sequencing, and genome comparison with close phylogenic relatives enabled us to classify strain Marseille-P2341^T as a new species of the genus *Murdochiella*. The 16S rRNA gene sequence similarity was 97.3% with *M. asaccharolytica*, which was lower than the 98.7% threshold recommended for defining a new species (Kim et al., 2014; Stackebrandt & Ebers, 2006). Created in 2010, the genus *Murdochiella* contains Gram-positive staining anaerobic cocci bacteria which have been detected in human clinical samples (Ulger-Toprak et al., 2010). Members of this genus are nonmotile and nonsporulating, as observed for strain Marseille-P2341^T.

A polyphasic taxono-genomic strategy based on the combination of phenotypic and genomic analyses (Fournier & Drancourt, 2015; Fournier et al., 2015) was used to describe the new species whose strain Marseille-P2341^T is the type strain. Strain Marseille-P2341^T exhibited a specific MALDI-TOF MS spectrum and differed from the other studied closed bacterial species in their fermentation of carbohydrate. Bacteria in the *Murdochiella* genus are asaccharolytic and do not ferment carbohydrates. However, the *M. vaginalis* strain Marseille-P2341^T produces acid from glucose and mannose. This observation was confirmed by the annotation of the genome with the COGs database (Figure 5), showed that 7.7% of Marseille-P2341 predicted genes' were dedicated to carbohydrate transport and metabolism functions. These genes include carbohydrate enzymes such as glucose-6-phosphate isomerase, 6-phosphogluconolactonase, 6-phosphofructokinase, fructose-bisphosphate aldolase, triose-phosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, phosphoglycerate mutase, enolase, pyruvate kinase, phosphomannomutase involved in carbohydrate metabolism, mainly in the process of glucose, fructose, and mannose metabolism.

TABLE 7 Numbers of orthologous proteins shared between genomes (upper right) and AGIOS values obtained (lower left)

	Murdochiella vaginalis	*Anaerococcus hydrogenalis*	*Helcococcus kunzii*	*Parvimonas micra*	*Helcococcus sueciensis*	*Peptoniphilus indolicus*	*Peptoniphilus coxii*
M. vaginalis	**1,446**	538	514	511	509	525	542
A. hydrogenalis	51.39	**2,069**	538	516	526	565	580
H. kunzii	51.12	57.33	**1,882**	541	653	511	534
P. micra	50.80	57.96	59.47	**1,678**	530	533	534
H. sueciensis	56.37	59.46	63.43	58.83	**1,445**	491	514
P. indolicus	52.45	58.27	56.33	58.43	59.21	**2,269**	614
P. coxii	52.67	53.15	52.95	53.78	50.25	52.93	**1,860**

The numbers of proteins per genome are indicated in bold.

The G+C content of strain Marseille-P2341^T and its phylogenetically-closest species ranges from 28.40 to 49.48 mol% and, as previously demonstrated, the difference in the G+C content is, at most, 1% in a species. Thus, overall, these values justify the strain Marseille-P2341^T being classified as a distinct species. The AGIOS and GGDC values also confirm it belongs to a new species (Klenk, Meier-Kolthoff, & Göker, 2014).

5 | TAXONOMIC AND NOMENCLATURE PROPOSAL

5.1 | Description of *Murdochiella vaginalis* sp. nov

Murdochiella vaginalis (va.gi.na'lis. L. n. *vagina*, sheath, vagina; L. fem. suff. *–alis*, suffix denoting pertaining to; N.L. fem. adj. *vaginalis*, pertaining to the vagina, of the vagina).

Obligate anaerobic *M. vaginalis* cells are Gram-stain-positive and coccus-shaped. They are nearly 0.7 μm in diameter, nonmotile, nonspore-forming, mesophilic, and occur in pairs or short chains. After 2 days of incubation on Columbia agar with 5% sheep's blood (bioMérieux) at 37°C under anaerobic conditions, colonies appear circular, white, and opaque with a diameter of 2–2.5 mm. Nitrate is not reduced; catalase and urease are also negative. Weakly saccharolytic, acid is produced only from glucose, mannose, and galactose. Positive reactions are observed for leucine arylamidase, Naphtol-AS-BI-phosphohydrolase, α- galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, N-acetyl-β-glucosaminidase, α-mannosidase, and α-fucosidase. The most abundant fatty acids are $C_{18:1n9}$ (27.7%) and $C_{16:0}$ (24.4%). The type strain is susceptible to oxacillin, penicillin, ceftriaxone, ciprofloxacin, clindamycin, doxycycline, erythromycin, fosfomycin, gentamycin, trimethoprim-sulfamethoxazole, vancomycin, and rifampicin but resistant to colistin.

Its genome contains 49.48 mol% of G+C content and measures 1,671,491 bp long. The 16S rRNA and whole-genome sequences are both deposited in EMBL-EBI under accession numbers LT576397 and LT632322 respectively. The type strain Marseille-P2341^T (=DSM 102237, =CSUR P2341) was isolated from the vaginal sample of a French woman suffering from bacterial vaginosis.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors thank the Xegen Company (www.xegen.fr) for automating the genomic annotation process. We also thank TradOnline for English reviewing and Claudia Andrieu for administrative assistance.

ORCID

Khoudia Diop http://orcid.org/0000-0002-9296-563X
Saber Khelaifia http://orcid.org/0000-0002-9303-3893
Pierre-Edouard Fournier http://orcid.org/0000-0001-8463-8885

REFERENCES

Bohbot, J. M., & Lepargneur, J. P. (2012). La vaginose en 2011: Encore beaucoup d'interrogations. *Gynécologie Obstétrique & Fertilité*, *40*, 31–36. https://doi.org/10.1016/j.gyobfe.2011.10.013

Bretelle, F., Rozenberg, P., Pascal, A., Favre, R., Bohec, C., Loundou, A., … Groupe de Recherche en Obstetrique Gynecologie (2015). High *Atopobium vaginae* and *Gardnerella vaginalis* vaginal loads are associated with preterm birth. *Clinical Infectious Diseases*, *60*, 860–867. https://doi.org/10.1093/cid/ciu966

Carver, T., Harris, S. R., Berriman, M., Parkhill, J., & McQuillan, J. A. (2012). Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics*, *28*, 464–469. https://doi.org/10.1093/bioinformatics/btr703

Carver, T., Thomson, N., Bleasby, A., Berriman, M., & Parkhill, J. (2009). DNAPlotter: Circular and linear interactive genome visualization. *Bioinformatics*, *25*, 119–120. https://doi.org/10.1093/bioinformatics/btn578

Chan, J. Z., Halachev, M. R., Loman, N. J., Constantinidou, C., & Pallen, M. J. (2012). Defining bacterial species in the genomic era: Insights from the genus *Acinetobacter*. *BMC Microbiology*, *12*, 302. https://doi.org/10.1186/1471-2180-12-302

Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2016). GenBank. *Nucleic Acids Research*, *44*, D67–D72. https://doi.org/10.1093/nar/gkv1276

Collins, M. D. (2004). *Helcococcus sueciensis* sp. nov., isolated from a human wound. *International Journal of Systematic and Evolutionary Microbiology*, *54*, 1557–1560. https://doi.org/10.1099/ijs.0.63077-0

De Vos, P., Garrity, G. M., Jones, D., Krieg, N. R., Ludwig, W., Rainey, F. A., … Whitman, W. B. (2009). The Firmicutes. In G. M. Garrity (Ed.), *Bergey's Manual of Systematic Bacteriology* (pp. 465–511). New York, NY: Springer.

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

Drancourt, M., Bollet, C., Carlioz, A., Martelin, R., Gayral, J.-P., & Raoult, D. (2000). 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. *Journal of Clinical Microbiology*, *38*, 3623–3630.

Ezaki, T., Kawamura, Y., Li, N., Li, Z. Y., Zhao, L., & Shu, S. (2001). Proposal of the genera *Anaerococcus* gen. nov., *Peptoniphilus* gen. nov. and *Gallicola* gen. nov. for members of the genus Peptostreptococcus. *International Journal of Systematic and Evolutionary Microbiology*, *51*, 1521–1528. https://doi.org/10.1099/00207713-51-4-1521

Ezaki, T., Yamamoto, N., Ninomiya, K., Suzuki, S., & Yabuuchi, E. (1983). Transfer of *Peptococcus indolicus*, *Peptococcus asaccharolyticus*, *Peptococcus prevotii*, and *Peptococcus magnus* to the Genus *Peptostreptococcus* and Proposal of *Peptostreptococcus tetradius* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, *33*, 683–698. https://doi.org/10.1099/00207713-33-4-683

Fournier, P. E., & Drancourt, M. (2015). New Microbes New Infections promotes modern prokaryotic taxonomy: A new section "TaxonoGenomics: New genomes of microorganisms in humans". *New Microbes and New Infections*, *7*, 48–49. https://doi.org/10.1016/j.nmni.2015.06.001

Fournier, P. E., Lagier, J. C., Dubourg, G., & Raoult, D. (2015). From culturomics to taxonomogenomics: A need to change the taxonomy of prokaryotes in clinical microbiology. *Anaerobe*, *36*, 73–78. https://doi.org/10.1016/j.anaerobe.2015.10.011

Fredricks, D. N., Fiedler, T. L., & Marrazzo, J. M. (2005). Molecular identification of bacteria associated with bacterial vaginosis. *New England Journal of Medicine*, *353*, 1899–1911. https://doi.org/10.1056/NEJMoa043802

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). New York, NY: Springer, Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20763-1

Gouret, P., Vitiello, V., Balandraud, N., Gilles, A., Pontarotti, P., & Danchin, E. G. (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

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

Johnson, C. N., Whitehead, T. R., Cotta, M. A., Rhoades, R. E., & Lawson, P. A. (2014). *Peptoniphilus stercorisuis* sp. nov., isolated from a swine manure storage tank and description of *Peptoniphilaceae* fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 3538–3545. https://doi.org/10.1099/ijs.0.058941-0

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

Klenk, H. P., Meier-Kolthoff, J. P., & Göker, M. (2014). Taxonomic use of DNA G+C content and DNA–DNA hybridization in the genomic age. *International Journal of Systematic and Evolutionary Microbiology*, *64*, 352–356. https://doi.org/10.1099/ijs.0.056994-0

Lagesen, K., Hallin, P., Rodland, 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., … Trape, J. F. (2012). Microbial culturomics: Paradigm shift in the human gut microbiome study. *Clinical Microbiology & Infection*, *18*, 1185–1193. https://doi.org/10.1111/1469-0691.12023

Lagier, J. C., Hugon, P., Khelaifia, S., Fournier, P. E., La Scola, B., & Raoult, D. (2015). The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clinical Microbiology Reviews*, *28*, 237–264. https://doi.org/10.1128/CMR.00014-14

Lamont, R., Sobel, J., Akins, R., Hassan, S., Chaiworapongsa, T., Kusanovic, J., & Romero, R. (2011). The vaginal microbiome: New information about genital tract flora using molecular based techniques. *BJOG: An International Journal of Obstetrics and Gynaecology*, *118*, 533–549. https://doi.org/10.1111/j.1471-0528.2010.02840.x

Le Page, S., van Belkum, A., Fulchiron, C., Huguet, R., Raoult, D., & Rolain, J.-M. (2015). Evaluation of the PREVI® Isola automated seeder system compared to reference manual inoculation for antibiotic susceptibility testing by the disk diffusion method. *European Journal of Clinical Microbiology and Infectious Diseases*, *34*, 1859–1869. https://doi.org/10.1007/s10096-015-2424-8

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

Lepargneur, J. P., & Rousseau, V. (2002). Protective role of the Doderleïn flora. *Journal de Gynecologie, Obstetrique et Biologie de la Reproduction*, *31*, 485–494.

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. https://doi.org/10.1093/nar/25.5.0955

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, *14*, 60. https://doi.org/10.1186/1471-2105-14-60

Menard, J., Fenollar, F., Henry, M., Bretelle, F., & Raoult, D. (2008). Molecular Quantification of *Gardnerella vaginalis* and *Atopobium vaginae* Loads to Predict Bacterial Vaginosis. *Clinical Infectious Diseases*, *47*, 33–43. https://doi.org/10.1086/588661

Murdoch, D., & Shah, H. N. (1999). Reclassification of *Peptostreptococcus magnus* (Prevot 1933) Holdeman and Moore 1972 as *Finegoldia magna* comb. nov. and Peptostreptococcus micros (Prevot 1933) Smith 1957 as *Micromonas micros* comb. nov. *Anaerobe*, *5*, 555–559. https://doi.org/10.1006/anae.1999.0197

Murray, P. R., Baron, E. J., Jorgensen, J. H., Landry, M. L., & Pfaller, M. A. (2007). *Manual of clinical microbiology* (9th ed.). Washington, D.C.: ASM Press.

Pal, K., Roy, S., Behera, B., Kumar, N., Sagiri, S., & Ray, S. (2011). Bacterial vaginosis: Etiology and modalities of treatment-A brief note. *Journal of Pharmacy And Bioallied Sciences*, *3*, 496. https://doi.org/10.4103/0975-7406.90102

Pépin, J., Deslandes, S., Giroux, G., Sobéla, F., Khonde, N., Diakité, S., … Frost, E. (2011). The complex vaginal flora of west african women with bacterial vaginosis. *PLoS ONE*, *6*, e25082. https://doi.org/10.1371/journal.pone.0025082

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

Sasser, M. (2006). Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). MIDI, Technical Note.

Seng, P., Drancourt, M., Gouriet, F., La Scola, B., Fournier, P., 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

Shipitsyna, E., Roos, A., Datcu, R., Hallén, A., Fredlund, H., Jensen, J. S., … Unemo, M. (2013). Composition of the vaginal microbiota in women of reproductive age – Sensitive and specific molecular diagnosis of bacterial vaginosis is possible? *PLoS ONE*, *8*, e60670. https://doi.org/10.1371/journal.pone.0060670

Stackebrandt, E., & Ebers, J. (2006). Taxonomic parameters revisited: Tarnished gold standards. *Microbiology Today*, *33*, 152.

Svare, J., Schmidt, H., Hansen, B., & Lose, G. (2006). Bacterial vaginosis in a cohort of Danish pregnant women: Prevalence and relationship with preterm delivery, low birthweight and perinatal infections. *BJOG: An International Journal of Obstetrics and Gynaecology*, *113*, 1419–1425. https://doi.org/10.1111/j.1471-0528.2006.01087.x

Tindall, B. J., & Euzeby, J. P. (2006). Proposal of *Parvimonas* gen. nov. and *Quatrionicoccus* gen. nov. as replacements for the illegitimate, prokaryotic, generic names *Micromonas* Murdoch and Shah 2000 and *Quadricoccus* Maszenan et al. 2002, respectively. *International Journal of Systematic and Evolutionary Microbiology*, *56*, 2711–2713. https://doi.org/10.1099/ijs.0.64338-0

Turovskiy, Y., Sutyak Noll, K., & Chikindas, M. L. (2011). The etiology of bacterial vaginosis. *Journal of Applied Microbiology*, *110*, 1105–1128. https://doi.org/10.1111/j.1365-2672.2011.04977.x

Ulger-Toprak, N., Liu, C., Summanen, P. H., & Finegold, S. M. (2010). *Murdochiella asaccharolytica* gen. nov., sp. nov., a Gram-stain-positive, anaerobic coccus isolated from human wound specimens. *International Journal of Systematic and Evolutionary Microbiology*, *60*, 1013–1016. https://doi.org/10.1099/ijs.0.015909-0

Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K., & Swings, J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiological Reviews*, *60*, 407–438.

How to cite this article: Diop K, Diop A, Khelaifia S, et al. Characterization of a novel Gram-stain-positive anaerobic coccus isolated from the female genital tract: Genome sequence and description of *Murdochiella vaginalis* sp. nov. *MicrobiologyOpen*. 2018;e570. https://doi.org/10.1002/mbo3.570

- **Caractérisation des nouvelles espèces isolées dans la flore vaginale normale**

Article 11:

Vaginimicrobium propionicum gen. nov. sp. nov., a novel propionic acid bacterium derived from human vaginal discharge.

Khoudia Diop, Frederic Cadoret, Thi Tien Nguyen, Jean-Pierre Baudoin, Nicholas Armstrong, Didier Raoult, Florence Bretelle, Pierre-Edouard Fournier, Florence Fenollar

Submitted in International Journal of Systematic and Evolutionary Microbiology

International Journal of Systematic and Evolutionary Microbiology

Vaginimicrobium propionicum gen. nov., sp. nov., a novel propionic acid bacterium derived from Human vaginal discharge

--Manuscript Draft--

Manuscript Number:	
Full Title:	Vaginimicrobium propionicum gen. nov., sp. nov., a novel propionic acid bacterium derived from Human vaginal discharge
Article Type:	Taxonomic Note
Section/Category:	New taxa - Actinobacteria
Keywords:	Vaginimicrobium propionicum; Bacterial vaginosis; Vaginal microbiota; Propionibacteriaceae; taxonogenomics, Anaerobic bacteria
Corresponding Author:	Florence FENOLLAR, PhD, MD IHU Mediterranee Infection Marseille, FRANCE
First Author:	Khoudia DIOP
Order of Authors:	Khoudia DIOP
	Frederic CADORET
	Thi Tien NGUYEN
	Jean-Pierre Baudoin
	Nicholas Armstrong
	Didier Raoult
	Florence Bretelle
	Pierre-Edouard FOURNIER
	Florence FENOLLAR, PhD, MD
Manuscript Region of Origin:	FRANCE
Abstract:	A Gram-stain-positive anaerobic rod-shaped bacterium, designated strain Marseille-P3275T, was isolated using culturomics from the vaginal discharge of healthy French woman. Strain Marseille-P3275T was non-motile and did not form spores. Cells had neither catalase nor oxidase activity. The major fatty acids were C16:0 (29%), C18:1ω9 (18%) and iso-C15:0 (17%). The genomic G+C content was 50.64 mol%. The phylogenic analysis based on 16S rRNA gene sequence suggested that strain Marseille-P3275T was related to members of the family Propionibacteriaceae (between 90.32 to 92.92% sequence similarity) with formation of a clade with the monospecific genus Propionimicrobium (type species Propionimicrobium lymphophilum). Based on these phylogenetic, and phenotypic distinctiveness, strain Marseille-P3275T was classed in a new genus, Vaginimicrobium, as Vaginimicrobium propionicum gen. nov., sp. nov. The type strain is Marseille-P3275T (=CSUR 3275 =CECT 9697).

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

1 **_Vaginimicrobium propionicum_ gen. nov., sp. nov., a novel propionic acid bacterium derived**

2 **from Human vaginal discharge**

3 Khoudia Diop[1], Frederic Cadoret[2], Thi Tien Nguyen[2], Jean-Pierre Baudoin[2], Nicholas

4 Armstrong[2], Didier Raoult[2,3], Florence Bretelle[2,4], Pierre-Edouard Fournier[1] and Florence

5 Fenollar[1*]

6 [1]Aix-Marseille Univ, IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection, Marseille,

7 France

8 [2]Aix-Marseille Univ, IRD, AP-HM, MEPHI, IHU-Méditerranée Infection, Marseille, France

9 [3]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz

10 University, Jeddah, Saudi Arabia

11 [4]Department of Gynecology and Obstetrics, Gynépole, Hôpital Nord, AP-HM, Marseille, France

12 [*]**Corresponding author**: Prof. Florence Fenollar, VITROME, Institut Hospitalo-Universitaire

13 Méditerranée Infection, 19-21 Boulevard Jean Moulin 13385 Marseille Cedex 05, France. Phone:

14 + 33 (0) 4 13 73 24 01. Fax: + 33 (0) 4 13 73 24 02. E-mail address: florence.fenollar@univ-

15 amu.fr

16 **Category:** New Taxa - *Actinobacteria*

17 **Keywords**: *Vaginimicrobium propionicum*; Bacterial vaginosis; Vaginal microbiota;

18 *Propionibacteriaceae*; taxonogenomics, Anaerobic bacteria.

19 **Running title:** *Vaginimicrobium propionicum* sp. nov.

20 **Abstract count:** 131

21 **Word count:** 2,959

22 **Abbreviations:** AAI, Average amino acid identity; CSUR, Collection de souches de l'Unité des

23 Rickettsies; dDDH, digital DNA-DNA hybridization; DSM, Deutsche Sammlung von

24 Mikroorganismen; MALDI-TOF, Matrix-assisted laser-desorption/ionization time-of-flight.

26 **ABSTRACT**

27 A Gram-stain-positive anaerobic rod-shaped bacterium, designated strain Marseille-P3275^T, was
28 isolated using culturomics from the vaginal discharge of healthy French woman. Strain Marseille-
29 P3275^T was non-motile and did not form spores. Cells had neither catalase nor oxidase activity.
30 The major fatty acids were $C_{16:0}$ (29 %), $C_{18:1\omega9}$ (18 %) and iso-$C_{15:0}$ (17 %). The genomic G+C
31 content was 50.64 mol%. The phylogenic analysis based on 16S rRNA gene sequence suggested
32 that strain Marseille-P3275^T was related to members of the family *Propionibacteriaceae*
33 (between 90.32 to 92.92 % sequence similarity) with formation of a clade with the monospecific
34 genus *Propionimicrobium* (type species *Propionimicrobium lymphophilum*). Based on these
35 phylogenetic, and phenotypic distinctiveness, strain Marseille-P3275^T was classed in a new
36 genus, *Vaginimicrobium*, as *Vaginimicrobium propionicum* gen. nov., sp. nov. The type strain is
37 Marseille-P3275^T (=CSUR 3275 =DSM 103714).

38 The human vagina is home to a diverse community of microbes called vaginal microbiota [1].
39 This vaginal microbiota has a significant impact on the health of women and their fetuses [2]. A
40 variation in the composition of vaginal microbiota can lead to bacterial vaginosis [3] or
41 obstetrics and gynecologic diseases such as chorioamnionitis, miscarriage, pelvic inflammatory
42 disease, as well as sexual transmitted diseases [4, 5]. In a Human vaginal microbiome project
43 aimed at exploring as fully as possible the vaginal microbiota of healthy and bacterial vaginosis
44 patients' using culturomics approach [6, 7], we isolated a novel bacterium, near to member of
45 *Propionibacteriaceae* family and designated as Marseille-P3275^T.

46 The family *Propionibacteriaceae* was described in 1957 by Delwiche then emended in 1997 and
47 2009 by Stackebrandt [8] and Zhi [9], respectively. At time of the writing, this family contains
48 20 genera including *Aestuariimicrobium, Auraticoccus, Brooklawnia, Friedmanniella,*
49 *Granulicoccus, Luteococcus, Mariniluteicoccus, Microlunatus, Micropruina, Naumannella,*
50 *Propionibacterium, Propionicicella, Propioniciclava, Propionicimonas, Propioniferax,*
51 *Propionimicrobium,* and *Tessaracoccus* (www.bacterio.net/-classifphyla.html), in addition to
52 three other genera: *Acidipropionibacterium*, *Cutibacterium* and *Pseudopropionibacterium*
53 derived recently from the subdivision of the genus *Propionibacterium* [10]. Members of this
54 family have been isolated in various biotopes, from human specimen [11–13], animal sources
55 [14–16], soils and vegetables [17–19]. Herein, follows the taxonogenomics description [20, 21]
56 of strain Marseille-P3275^T (=CSUR 3275 =DSM 103714) containing its phenotypic
57 characteristics coupled with phylogenetic and genomic characteristics.

58 Strain Marseille-P3275^T was isolated from the vaginal swab of a 23 years-old French woman,
59 without bacterial vaginosis. The vaginal discharge was sampled at Hospital "La Timone" in
60 Marseille (France) using a Sigma Transwab (Medical Wire, Corsham, United Kingdom).

61 Bacterial vaginosis was diagnosed as previously described [22]. The patient gave a written and
62 signed consent and the study was validated by the ethic committee of the Institut Federatif de
63 Recherche IFR48 under agreement number 09-022. At the time of sampling, the patient had not
64 received any antibiotics treatments.

65 For bacterial culture, 0.5 ml of the specimen was rapidly pre-incubated in a blood culture bottle
66 (Becton-Dickinson, Le Pont de Claix, France) enriched with 4 ml of rumen fluid that was filter-
67 sterilized through a 0.2 µm pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France)
68 and 3 ml of sheep's blood (bioMérieux, Marcy l'Etoile, France) and then incubated at 37°C.
69 After 10 days of pre-incubation periods, 100 µl of the broth was plated both on colistin nalidixic
70 acid (CNA) and Schaedler agar (BD Diagnostics) and then incubated for 4 days under anaerobic
71 conditions at 37°C. Isolated colonies were purified and subsequently identified by Matrix-
72 Assisted Laser-Desorption/Ionization Time-Of-Flight (MALDI-TOF) mass spectrometry with a
73 Microflex LT spectrometer (Bruker, Leipzig, Germany), as previously described [23]. The
74 obtained spectra were imported into MALDI Biotyper 3.0 software (Bruker Daltonics) and
75 analyzed by standard pattern matching (with default parameter settings) against the main spectra
76 included in the database (Bruker database constantly updated with MEPHI database). When the
77 MALDI-TOF failed to identify a bacterium, the identification and phylogenetic affiliation of the
78 isolate were performed by amplifying and sequencing its 16S rRNA gene using the universal
79 primers pair fD1 and rp2 (Eurogentec, Angers, France) and the ABI Prism 3130-XL Genetic
80 Analyzer capillary sequencer (Applied Biosystems, Bedford, MA, USA), as previously described
81 [24, 25]. The obtained sequence was corrected using Chromas Pro 1.34 software (Technelysium
82 Pty. Ltd., Tewantin, Australia). Then the identification of the bacterium and its taxonomic
83 position were determined using the BLASTn program [26] in the EzBioCloud database [27] that

84 contains the type strains of all validly published species and available at

85 https://www.ezbiocloud.net.

86 The 16S rRNA sequences of the type strains of closest species to the strain Marseille-P3275^{T}'s

87 indicated in LPSN (http://www.bacterio.net) were retrieved from the NCBI database. Pairwise

88 similarity was evaluated using Meier-Kolthoff's recommended method for the 16S rRNA gene

89 [28] implemented in GGDC web server (http://ggdc.dsmz.de/; [29]). The sequences were aligned

90 with MUSCLE [30] and phylogenies were inferred in GGDC web server [29] using the DSMZ

91 phylogenomics pipeline adapted to single genes [31]. Maximum likelihood (ML) and maximum

92 parsimony (MP) trees were constructed from this alignment with RAxML [32] and TNT [33],

93 respectively. For ML, rapid bootstrapping in conjunction with the autoMRE bootstopping

94 criterion [34] and subsequent search for the best tree was used; for MP, 1000 bootstrapping

95 replicates were used in conjunction with tree-bisection-and-reconnection branch swapping and

96 ten random sequence addition replicates. The sequences were checked for a compositional bias

97 using the X^2 test as implemented in PAUP* [35]. The downloaded sequences were also aligned

98 using CLUSTALW [36] and phylogenetic tree was constructed using Neighbor-joining methods

99 within MEGA version 6 software [37]. The stability of the groupings was estimated by

100 bootstrap's analysis (1,000 replicates).

101 The best growth condition of stain Marseille-P3275^{T} was characterized by culturing the

102 bacterium on 5% sheep blood agar (bioMérieux) and incubated it at several temperature (56, 42,

103 37, 28, and 25°C) under different atmospheres: anaerobic and microaerophilic conditions created

104 using GENbag Anaer and GENbag Microaer systems, respectively (bioMérieux, Marcy-l'Étoile,

105 France) and aerobic conditions, with or without 5% of CO_2. The tolerance of diverse salinity

106 concentrations (NaCl concentrations of 0-100 g/l) and pH (5, 6, 6.5, 7, and 8.5) conditions were
107 also tested.

108 Ultrastructure of cells was observed using a Tecnai G20 (FEI company, Limeil-Brevannes,
109 France) transmission electron microscope operated at 200 kV. Gram-stain, sporulation, mobility,
110 catalase and oxidase activity were determined as previously described [38]. Biochemical features
111 of strain Marseille-P3275^T were characterized using API ZYM, API 20A, and API 32A strips
112 (bioMérieux,) according to the manufacturer's instructions. Antibiotics susceptibility was tested
113 by estimated the Minimal inhibitory concentrations (MICs) using E-test gradient strips
114 (bioMérieux) according to EUCAST recommendations [39, 40].

115 For the analysis of cellular fatty acid methyl ester (FAME), Gas Chromatography/Mass
116 Spectrometry (GC/MS) was used. Strain Marseille-P3275^T was grown anaerobically at 37°C on
117 5% sheep blood-enriched Columbia agar (bioMérieux). After 2 days of incubation, two aliquots
118 were prepared with approximately 10 mg of bacterial biomass per tube. FAME analyses were
119 prepared as described by Sasser [41]. GC/MS analyses were carried out as previously described
120 [42]. Briefly, fatty acid methyl esters were separated using an Elite 5-MS column and monitored
121 by mass spectrometry (Clarus 500 - SQ 8 S, Perkin Elmer, Courtaboeuf, France). A spectral
122 database search was performed using MS Search 2.0 operated with the Standard Reference
123 Database 1A (NIST, Gaithersburg, USA) and the FAME mass spectral database (Wiley,
124 Chichester, UK).

125 Short chain fatty acids analysis (SCFA) were measured with a Clarus 500 chromatography
126 system connected to a SQ8s mass spectrometer (Perkin Elmer, Courtaboeuf, France) such as
127 detailed previously [43] with modifications. Acetic, propanoic, isobutanoic, butanoic,

128 isopentanoic, pentanoic, isohexanoic, hexanoic and heptanoic were purchased from Sigma
129 Aldrich (Lyon, France). A stock solution was prepared in water/methanol (50% v/v) at a final
130 concentration of 50 mmol/L and then stored at -20°C. Calibration standards were freshly
131 prepared in acidified water (pH 2-3 with HCl 37%) from the stock solution at the following
132 concentrations: 0.5; 1; 5; 10 mmol/l. SCFA were analyzed from 3 independent culture bottles
133 (both blank and samples). Culture medium was collected then centrifuged 5 minutes at 16000 x g
134 to remove bacteria and debris. The clear supernatant was adjusted to pH 2-3 and spiked with 2-
135 ethylbutyric acid as the internal standard (IS) at a final concentration of 1 mmol/l (Sigma
136 Aldrich). The solution was once again centrifuged before injection. Aqueous samples were
137 directly injected (0.5 µl) in a split less liner heated at 200°C. Injection carry-over was decreased
138 with 10 syringe washes in methanol:water (50:50 v/v). Compounds were then separated on an
139 Elite-FFAP column (30 m, 0.25 mm id., 0.25mm film thickness) using a linear temperature
140 gradient from 100 to 200°C at 8°C/min. Helium flowing at 1 ml/min was used as carrier gas. MS
141 inlet line and Electron Ionization source were set at 200°C. To insure compound selectivity,
142 Selected Ion Recording (SIR) was performed after a 4.5 min solvent delay with the following
143 masses: 43 m/z (isobutanoic acid), 60 m/z (acetic, butanoic, pentanoic, isopentanoic, hexanoic
144 and heptanoic acids) 74 m/z (isohexanoic and propanoic acid), 88 m/z (2-ethylbutyric acid, IS).
145 All data was collected and processed using Turbomass 6.1 (Perkin Elmer, Courtaboeuf, France).
146 Quadratic internal calibration was calculated for each acid using the peak areas from the
147 associated SIR chromatograms. Coefficients of determination were all above 0.999. Back
148 calculated standards and calculated quality controls (0.5 and 5 mmol/l) all showed good accuracy
149 with deviations below 15%. SCFA quantities in samples were presented after subtraction of the
150 quantities measured in the blank samples.

151 The genomic DNA (gDNA) of strain Mareille-P3275^T was quantified by a Qubit assay with the
152 high sensitivity kit (Life technologies, Carlsbad, CA, USA) to 16.8 ng/µl. gDNA of strain
153 Mareille-P3275^T was sequenced on the MiSeq Technology (Illumina Inc, San Diego, CA, USA)
154 with the mate pair strategy. The gDNA was barcoded in order to be mixed with 11 other projects
155 with the Nextera Mate Pair sample prep kit (Illumina). The library profile was visualized on a
156 High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc, Santa Clara, CA, USA) and the
157 final concentration library was measured at 8.82 nmol/l [38]. A total information of 9.5 Gb was
158 obtained from a 1050 K/mm^2 cluster density with a cluster passing quality control filters of 92.5
159 % (18,644,000 passing filter paired reads). The 1,003,034 paired reads were trimmed then
160 assembled using Spades assembler program [44].

161 The genome annotation and analysis processes were performed as previously reported [45] using
162 the Multi-Agent software system DAGOBAH [46], which include Figenix libraries [47] to
163 provide pipeline analysis. To evaluate the genetic difference between strain Marseille-P3275^T and
164 its neighbors, two parameters were determined: the digital DNA-DNA hybridization (dDDH)
165 value using the genome to genome distance calculator [29, 48–50] and the average amino acid
166 identity (AAI) based on the overall similarity between two genomic datasets of proteins and
167 available at http://enve-omics.ce.gatech.edu/aai/index.

168 No MALDI-TOF identification of the strain Marseille-P3275^T was observed. However,
169 phylogenetic analysis based on 16S rRNA gene sequences using neighbor-joining method (Fig.
170 1) shows that strain Marseille-P3275^T belong to the family *Propionibacteriaceae* and its clustered
171 with *Propionimicrobium lymphophilum* DSM 4903^T [13]. This cluster was sustained by a in
172 bootstrap value of 92.0%. The same topology was also found in a tree combining the ML/MP
173 methods (Fig. 2). The input nucleotide matrix using for phylogenetic inferences comprised 15

174 operational taxonomic units and 1,561 characters, of which 321 were variable and 217 were
175 parsimony-informative. The BLASTn analysis of strain Marseille-P3275^T exhibited a 16S rRNA
176 nucleotide sequence similarity of 92.92, 92.48, 92.37 and 91.93%, respectively with *P.*
177 *lymphophilum* DSM 4903, *Propionibacterium australiense* LCDC-98A072,
178 *Pseudopropionibacterium propionicum* NBRC 14587 and *Cutibacterium acnes* DSM 1897 and
179 of between 90.32% - 91.92% with others validly published members of *Propionibacteriaceae*
180 family. Since then, its MALDI-TOF reference spectrum was incremented in our database
181 (http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database) and
182 compared to those of closest related species (Fig. S1).

183 Strain Marseille-P3275^T colonies, on the surface of Columbia agar plate enriched with 5% sheep
184 blood (bioMérieux) after 3 days anaerobic growth, had a diameter of 0.2-0.6 mm and were
185 circular, entire, convex, smooth, shiny, and white to pale cream. The bacterium grew in anaerobic
186 and microaerophilic atmospheres at temperature between 28-42 °C but optimum growth was
187 observed at 37 °C after 2 days of incubation anaerobically. The growth of strain Marseille-
188 P3275^T occurred at pH ranging between 6-7.5 and NaCl concentration less than 5 g/l. Cells of
189 strain Marseille-P3275^T were Gram-stain-positive, non-spore-forming, non-motile, rod-shaped
190 bacterium with 0.4-0.6 µm in diameter 0.65- 0.85µm in length (Fig. 3).

191 Strain Marseille-P3275^T was catalase and oxidase negative. The main phenotypic characteristics
192 compared with those most closely related species based on 16S rRNA gene sequence
193 comparisons are presented in Table 1. The results of the other phenotypic tests of the new strain
194 are given in the description section of the species. The major fatty acids were $C_{16:0}$ (29%), $C_{18:1\omega9}$
195 (18%) and iso-$C_{15:0}$ (17%). Several branched structures, mainly iso, were also described. Cellular
196 fatty acid profile of strain Marseille-P3275^T and its closest species is summarized in Table 2.

197 Strain Marseille-P3275^T produced SCFAs after 24, 48 and 72h of culture in reduced WCA broth.
198 A production of propanoic (10.0 ± 1.3 mM) and acetic (3.0 ± 0.8 mM) acids were measured after
199 72 h of culture. Isobutyric acid was also detected (< 0.5 mM).

200 The draft genome of strain Marseille-P3275^T (accession number LT706985) is 2,010,679 bp long
201 with 50.64 mol% of G+C content (Fig. S2, Table S1). It is composed of 1 scaffolds (composed of
202 1 contigs). Of the 1,921 predicted genes, 1,869 were protein-coding genes and 52 were RNAs
203 (two 5S rRNA, two 16S rRNA, two 23S rRNA, and 46 tRNA genes). A total of 1,386 genes
204 (74.16%) were assigned as putative function by COGs or NR BLAST. 92 genes were identified
205 as ORFans (4.92%). Using PHAST and RAST, 777 genes were associated to mobilome elements
206 (41.57%). The genome analysis reveals that strain Marseille-P3275^T contains also 345 virulence-
207 related genes (18.46%) of which 16 encoding for bacteriocin (0.86%) and 65 genes for toxin /
208 antitoxin (3.48%). The remaining genes (337 genes, 18.03%) were annotated as hypothetical
209 proteins. The genome statistics are recapitulated in Table S1. According to COG functional
210 categories presented in Table S2, translation (151; 8.08%), carbohydrate transport and
211 metabolism (143; 7.65%) and amino acid transport and metabolism (125; 6.69%) were the most
212 predominant gene categories.

213 The genomic comparison of strain Marseille-P3275^T's with those of its neighbors is itemized in
214 Table S3. The genome size, G+C percent and gene contents of strain Marseille-P3275^T (2.01 Mb,
215 50.64% and 1,869, respectively) are smaller than those of others compared genomes but are very
216 close to those of *P. lymphophilum* (2.04 Mb, 56.06% and 1,839, respectively) (Table S3).
217 However, strain Marseille-P3275^T shows a genomic difference of 5.42% in G+C content
218 compared to *P. lymphophilum*. The repartition of genes into COGs categories are identical in all
219 compared genomes (Fig. S3). Nevertheless, strain Marseille-P3275^T presents few genes in COGs

220 categories, T (signal transduction mechanisms), H (coenzyme transport and metabolism), P
221 (inorganic ion transport and metabolism) and Q (secondary metabolites biosynthesis, transport
222 and catabolism) compared to its neighbors (Fig. S3). The dDDH values among compared species
223 ranged from 18.40 ± 2.30% between *P. acidifaciens* and *P. propionicum* to 27 ± 2.40% between
224 *A. thoenii* and *P. freudenreichii* (Table S4). Strain Marseille-P3275^T exhibited dDDH values
225 ranking from 21.30 ± 2.35% with *P. freudenreichii* to 29.60 ± 2.45% with *A. acidipropionici*
226 (Table S4). In addition, the AAI values ranked from 52.35% between *T. flavescens* and *P.*
227 *lymphophilum* to 85.43% between *C. acnes* to *C. avidum* (Table S5). Furthermore, strain
228 Marseille-P3275^T shared AAI values of 52.04% with P. *propionicum* to 66.08 with P.
229 *lymphophilum* (Table S5).

230 The combination of phylogenetic (Fig. 1 and 2), phenotypic (Tables 1 and 2) and genomic
231 characteristics (supplementary data) evinces that strain Marseille-P3275^T is a member of the
232 family *Propionibacteriaceae*. However, it can be easily differentiated from other members of
233 *Propionibacteriaceae* on the basis of its fatty acids composition (Table 2) and its genomic data,
234 especially the difference of DNA G+C content greater than 5% (between 5.42% compared with
235 *P. lymphophilum* to 17.66% with *T. flavescens*) [31] as well as the dDDH and AAI values lower
236 than threshold of 70% and 95-96% [29, 48–50], respectively for bacterial species demarcation
237 (supplementary data). In addition, its very low 16S rRNA gene sequence similarity value
238 (between 90.32- 92.92%) with the 95% limit set to distinguish two genera [51–53] corroborate
239 the classification of strain Marseille-P3275^T into a new genus. On the basis of phenotypic
240 features, phylogenetic inferences and genomic data, strain Marseille-P3275^T differed sufficiently
241 from its closest related species and may be placed into a novel species, belonging to a new genus

242 within *Propionibacteriaceae* family. The name*, Vaginimicrobium propionicum* gen. nov, sp. nov. is

243 proposed for this new taxon.

244 DESCRIPTION OF *VAGINIMICROBIUM* GEN. NOV.

245 *Vaginimicrobium* (Va.gi.ni.mi.cro'bi.um. N.L.n. neutr. *Vaginimicrobium,* combination of vagina,

246 the Latin name of vagina, and microbium, a microbe).

247 Cells are anaerobic, non-motile, non-spore-forming and have a Gram-positive-staining. They are

248 pleomorphic or club-shaped rods that occur in single cells, in pairs or short chains. The rods are

249 0.75µm of long and 0.5µm of wide. Catalase, and oxidase activities tests are negative. Nitrate is

250 not reduced to nitrite and indole is not produced. Urease is positive. Aesculin and gelatin are not

251 hydrolyzed. Based on its 16S rRNA gene sequence, the genus be part of *Propionibacteriaceae*

252 family. The type species is *Vaginimicrobium propionicum*. Habitat is human vagina.

253 DESCRIPTION OF *VAGINIMICROBIUM PROPIONICUM* SP. NOV.

254 *Vaginimicrobium propionicum* (pro.pio.nic'um. N.L. n. acidum propionicum, propionic acid; L.

255 neut. suff. -icum, suffix used with the sense of pertaining to; N.L. neut. adj. *propionicum*,

256 pertaining to propionic acid).

257 In addition to properties given in the genus description, cultured anaerobically on 5% sheep

258 blood-enriched Columbia agar after 48h at 37°C, colonies are circular, entire, convex, smooth,

259 shiny, white to pale cream and have diameter of 0.2-0.6 mm. Using API ZYM and 32A strips,

260 arginine arylamidase, leucine arylamidase, proline arylamidase, phenylalanine arylamidase,

261 tyrosine arylamidase, Naphthol-AS-BI-phosphohydrolase, β-glucuronidase, and glucosidase (α

262 and β) activities are detected. Whereas, alkaline phosphatase, esterase lipase, lipase, cystine

263 arylamidase, trypsin, α-chymotrypsin, acid phosphatase, galactosidase (α and β), N-acetyl-β-

264 glucosaminidase, α-arabinosidase, α-mannosidase and α-fucosidase activities are not observed.
265 However, on API ZYM strip, esterase and valine arylamidase activities are weakly detected. In
266 addition, leucyl glycine arylamidase, acid pyroglutamic arylamidase, alanine arylamidase,
267 glycine arylamidase, histidine arylamidase, and serine arylamidase activities are negative on API
268 32A strip. Using API 20A strip, D-glucose, D-saccharose, D-maltose, salicine, L-arabinose, D-
269 melezitose, D-xylose and D-trehalose are fermented. D-mannitol, D-lactose, glycerol, D-
270 cellobiose, D-mannose, D-raffinose, D-sorbitol and L-rhamnose are not assimilated.
271 Contrariwise, on API 32A strip, acid is formed from D-mannose and D-raffinose. Propionic and
272 acetic acids are the major end products of the glucose fermentation. In addition, isobutyric acid is
273 also detected. The major fatty acids are $C_{16:0}$, $C_{18:1\omega9}$ and iso-$C_{15:0}$. Our isolate is susceptible to
274 amoxicillin (MIC < 0.002), benzylpenicillin (MIC = 0.032), imipenem (MIC = 0.032) and
275 vancomycin (MIC = 0.50).

276 Its genome exhibits 50.64% of DNA G+C content. The 16S rRNA gene and genome sequences
277 were deposited in Genbank under accession number LT598595 and LT706985, respectively. The
278 type strain, Marseille-P3275^T (=CSUR 3275 =DSM 103714), was isolated from the vaginal
279 specimen of a 23 years-old French woman without bacterial vaginosis.

280 **Funding information**

281 This study was funded by the "Fondation Méditerranée Infection" and the French Government
282 under the "Investissements d'avenir" program managed by the National Agency for Research
283 (reference Méditerranée Infection 10-IAHU-03).

284 **Acknowledgements**

285 We express our gratitude to Xegen Company (www.xegen.fr) for automating the genomic
286 annotation process. We also thank TradOnline for English reviewing and Aurélia Caputo for
287 submitting the genomic sequences to GenBank.

288 **Conflict of interest**

289 The authors declare no conflicts of interest.

290 **References**

291 1. **Pal K, Roy S, Behera B, Kumar N, Sagiri S, *et al.*** Bacterial vaginosis: Etiology and
292 modalities of treatment-A brief note. *J Pharm Bioallied Sci* 2011;3:496.
293 2. **Li J, McCormick J, Bocking A, Reid G**. Importance of Vaginal Microbes in Reproductive
294 Health. *Reprod Sci* 2012;19:235–242.
295 3. **Onderdonk AB, Delaney ML, Fichorova RN**. The Human Microbiome during Bacterial
296 Vaginosis. *Clin Microbiol Rev* 2016;29:223–238.
297 4. **Svare J, Schmidt H, Hansen B, Lose G**. Bacterial vaginosis in a cohort of Danish pregnant
298 women: prevalence and relationship with preterm delivery, low birthweight and perinatal
299 infections. *BJOG Int J Obstet Gynaecol* 2006;113:1419–1425.
300 5. **Marrazzo JM, Hillier SL**. Bacterial Vaginosis. In: *Sexually Transmitted Diseases*. Elsevier.
301 pp. 463–498.
302 6. **Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, *et al.*** The Rebirth of Culture
303 in Microbiology through the Example of Culturomics To Study Human Gut Microbiota. *Clin*
304 *Microbiol Rev* 2015;28:237–264.
305 7. **Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, *et al.*** Culture of previously
306 uncultured members of the human gut microbiota by culturomics. *Nat Microbiol*
307 2016;1:16203.
308 8. **Stackebrandt E, Rainey FA, Ward-Rainey NL**. Proposal for a new hierarchic classification
309 system, *Actinobacteria* classis nov. *Int J Syst Evol Microbiol* 1997;47:479–491.
310 9. **Zhi XY, Li WJ, Stackebrandt E**. An update of the structure and 16S rRNA gene sequence-
311 based definition of higher ranks of the class *Actinobacteria*, with the proposal of two new
312 suborders and four new families and emended descriptions of the existing higher taxa. *Int J*
313 *Syst Evol Microbiol* 2009;59:589–608.

314 10. **Scholz CFP, Kilian M**. The natural history of cutaneous propionibacteria, and
315 reclassification of selected species within the genus *Propionibacterium* to the proposed novel
316 genera *Acidipropionibacterium* gen. nov., *Cutibacterium* gen. nov. and
317 *Pseudopropionibacterium* gen. nov. *Int J Syst Evol Microbiol* 2016;66:4422–4432.
318 11. **Downes J, Wade WG**. *Propionibacterium acidifaciens* sp. nov., isolated from the human
319 mouth. *Int J Syst Evol Microbiol* 2009;59:2778–2781.
320 12. **Aubin GG, Bémer P, Kambarev S, Patel NB, Lemenand O, *et al.*** *Propionibacterium*
321 *namnetense* sp. nov., isolated from a human bone infection. *Int J Syst Evol Microbiol*
322 2016;66:3393–3399.
323 13. **Stackebrandt E**. *Propionimicrobium* gen. nov., a new genus to accommodate
324 *Propionibacterium lymphophilum* (Torrey 1916) Johnson and Cummins 1972, 1057(AL) as
325 *Propionimicrobium lymphophilum* comb. nov. *Int J Syst Evol Microbiol* 2002;52:1925–1927.
326 14. **Bernard KA, Shuttleworth L, Munro C, Forbes-Faulkner JC, Pitt D, *et al.***
327 *Propionibacterium australiense* sp. nov. Derived from Granulomatous Bovine Lesions.
328 *Anaerobe* 2002;8:41–47.
329 15. **Sugawara Y, Ueki A, Abe K, Kaku N, Watanabe K, *et al.*** *Propioniciclava tarda* gen. nov.,
330 sp. nov., isolated from a methanogenic reactor treating waste from cattle farms. *Int J Syst*
331 *Evol Microbiol* 2011;61:2298–2303.
332 16. **Li GD, Chen X, Li QY, Xu FJ, Qiu SM, *et al.*** *Tessaracoccus rhinocerotis* sp. nov., isolated
333 from the faeces of Rhinoceros unicornis. *Int J Syst Evol Microbiol* 2016;66:922–927.
334 17. **Lucena-Padros H, Gonzalez JM, Caballero-Guerrero B, Ruiz-Barba JL, Maldonado-**
335 **Barragan A**. *Propionibacterium olivae* sp. nov. and *Propionibacterium damnosum* sp. nov.,
336 isolated from spoiled packaged Spanish-style green olives. *Int J Syst Evol Microbiol*
337 2014;64:2980–2985.

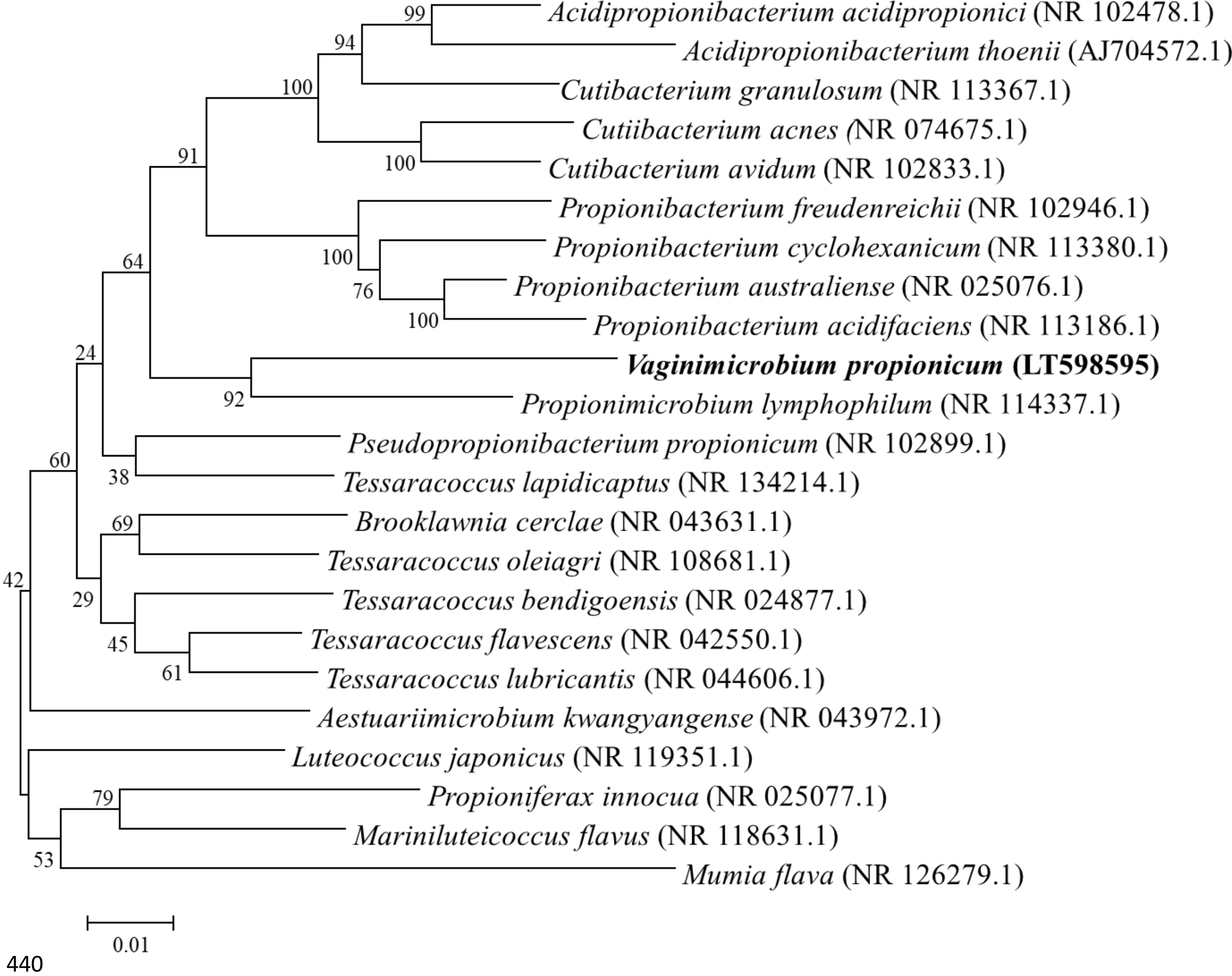
338 18. **Jung SY, Kim HS, Song JJ, Lee SG, Oh TK, *et al.*** *Aestuariimicrobium kwangyangense*
339 gen. nov., sp. nov., an ll-diaminopimelic acid-containing bacterium isolated from tidal flat
340 sediment. *Int J Syst Evol Microbiol* 2007;57:2114–2118.
341 19. **Zhang X, Zhang J, Zhang Y, Xin Y, He H**. *Friedmanniella flava* sp. nov., a soil
342 actinomycete. *Int J Syst Evol Microbiol* 2013;63:1771–1775.
343 20. **Fournier PE, Drancourt M**. New Microbes New Infections promotes modern prokaryotic
344 taxonomy: a new section "TaxonoGenomics: new genomes of microorganisms in humans".
345 *New Microbes New Infect* 2015;7:48–49.
346 21. **Fournier PE, Lagier JC, Dubourg G, Raoult D**. From culturomics to taxonomogenomics:
347 A need to change the taxonomy of prokaryotes in clinical microbiology. *Anaerobe*
348 2015;36:73–78.
349 22. **Menard JP, Fenollar F, Henry M, Bretelle F, Raoult D**. Molecular Quantification of
350 *Gardnerella vaginalis* and *Atopobium vaginae* Loads to Predict Bacterial Vaginosis. *Clin*
351 *Infect Dis* 2008;47:33–43.
352 23. **Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P, *et al.*** Ongoing Revolution in
353 Bacteriology: Routine Identification of Bacteria by Matrix-Assisted Laser Desorption
354 Ionization Time-of-Flight Mass Spectrometry. *Clin Infect Dis* 2009;49:543–551.
355 24. **Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, *et al.*** 16S ribosomal DNA
356 sequence analysis of a large collection of environmental and clinical unidentifiable bacterial
357 isolates. *J Clin Microbiol* 2000;38:3623–3630.
358 25. **Morel AS, Dubourg G, Prudent E, Edouard S, Gouriet F, *et al.*** Complementarity between
359 targeted real-time specific PCR and conventional broad-range 16S rDNA PCR in the
360 syndrome-driven diagnosis of infectious diseases. *Eur J Clin Microbiol Infect Dis*
361 2015;34:561–570.

362 26. **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, *et al.*** Gapped BLAST and
363 PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*
364 1997;25:3389–3402.
365 27. **Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, *et al.*** Introducing EzBioCloud: a taxonomically
366 united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol*
367 *Microbiol* 2017;67:1613–1617.
368 28. **Meier-Kolthoff JP, Göker M, Spröer C, Klenk HP**. When should a DDH experiment be
369 mandatory in microbial taxonomy? *Arch Microbiol* 2013;195:413–418.
370 29. **Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M**. Genome sequence-based species
371 delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*
372 2013;14:60.
373 30. **Edgar RC**. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
374 *Nucleic Acids Res* 2004;32:1792–1797.
375 31. **Meier-Kolthoff JP, Klenk H-P, Göker M**. Taxonomic use of DNA G+C content and DNA–
376 DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 2014;64:352–356.
377 32. **Stamatakis A**. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
378 phylogenies. *Bioinforma Oxf Engl* 2014;30:1312–1313.
379 33. **Goloboff PA, Farris JS, Nixon KC**. TNT, a free program for phylogenetic analysis.
380 *Cladistics* 2008;24:774–786.
381 34. **Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A**. How
382 many bootstrap replicates are necessary? *J Comput Biol J Comput Mol Cell Biol*
383 2010;17:337–354.

384 35. **L. Swofford D**. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods).*
385 *Version 4.0b10*. Sinauer Associates. Sunderland; 2002. Epub ahead of print 1 January 2002.
386 DOI: 10.1111/j.0014-3820.2002.tb00191.x.
387 36. **Thompson JD, Higgins DG, Gibson TJ**. CLUSTAL W: improving the sensitivity of
388 progressive multiple sequence alignment through sequence weighting, position-specific gap
389 penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–4680.
390 37. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S**. MEGA6: Molecular
391 Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 2013;30:2725–2729.
392 38. **Togo AH, Durand G, Khelaifia S, Armstrong N, Robert C, *et al.*** *Fournierella*
393 *massiliensis* gen. nov., sp. nov., a new human-associated member of the family
394 *Ruminococcaceae*. *Int J Syst Evol Microbiol* 2017;67:1393–1399.
395 39. **Citron DM, Ostovari MI, Karlsson A, Goldstein EJ**. Evaluation of the E test for
396 susceptibility testing of anaerobic bacteria. *J Clin Microbiol* 1991;29:2197–2203.
397 40. **Matuschek E, Brown DFJ, Kahlmeter G**. Development of the EUCAST disk diffusion
398 antimicrobial susceptibility testing method and its implementation in routine microbiology
399 laboratories. *Clin Microbiol Infect* 2014;20:O255–O266.
400 41. **Sasser M**. Bacterial Identification by Gas Chromatographic Analysis of Fatty acids Methyl
401 Esters (GC-FAME). *MIDI, Technical Note*.
402 42. **Dione N, Sankar SA, Lagier JC, Khelaifia S, Michele C, *et al.*** Genome sequence and
403 description of *Anaerosalibacter massiliensis* sp. nov. *New Microbes New Infect* 2016;10:66–
404 76.
405 43. **Zhao G, Nyman M, Åke Jönsson J**. Rapid determination of short-chain fatty acids in
406 colonic contents and faeces of humans and rats by acidified water-extraction and direct-
407 injection gas chromatography. *Biomed Chromatogr* 2006;20:674–682.

408 44. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al.*** SPAdes: A New
409 Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol*
410 2012;19:455–477.
411 45. **Tidjani Alou M, Khelaifia S, Michelle C, Andrieu C, Armstrong N, *et al.*** *Anaerococcus*
412 *rubiinfantis* sp. nov., isolated from the gut microbiota of a Senegalese infant with severe
413 acute malnutrition. *Anaerobe* 2016;40:85–94.
414 46. **Gouret P, Paganini J, Dainat J, Louati D, Darbo E, *et al.*** Integration of Evolutionary
415 Biology Concepts for Functional Annotation and Automation of Complex Research in
416 Evolution: The Multi-Agent Software System DAGOBAH. In: Pontarotti P (editor).
417 *Evolutionary Biology – Concepts, Biodiversity, Macroevolution and Genome Evolution*.
418 Springer Berlin Heidelberg; 2011. pp. 71–87.
419 47. **Gouret P, Vitiello V, Balandraud N, Gilles A, Pontarotti P, *et al.*** FIGENIX: intelligent
420 automation of genomic annotation: expertise integration in a new software platform. *BMC*
421 *Bioinformatics* 2005;6:1.
422 48. **Auch AF, von Jan M, Klenk HP, Göker M**. Digital DNA-DNA hybridization for microbial
423 species delineation by means of genome-to-genome sequence comparison. *Stand Genomic*
424 *Sci* 2010;2:117–134.
425 49. **Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, *et al.*** Proposed minimal
426 standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol*
427 *Microbiol* 2018;68:461–466.
428 50. **Rodriguez-R LM, Konstantinidis KT**. Bypassing Cultivation To Identify Bacterial Species:
429 Culture-independent genomic approaches identify credibly distinct clusters, avoid cultivation
430 bias, and provide true insights into microbial species. *Microbe Mag* 2014;9:111–118.

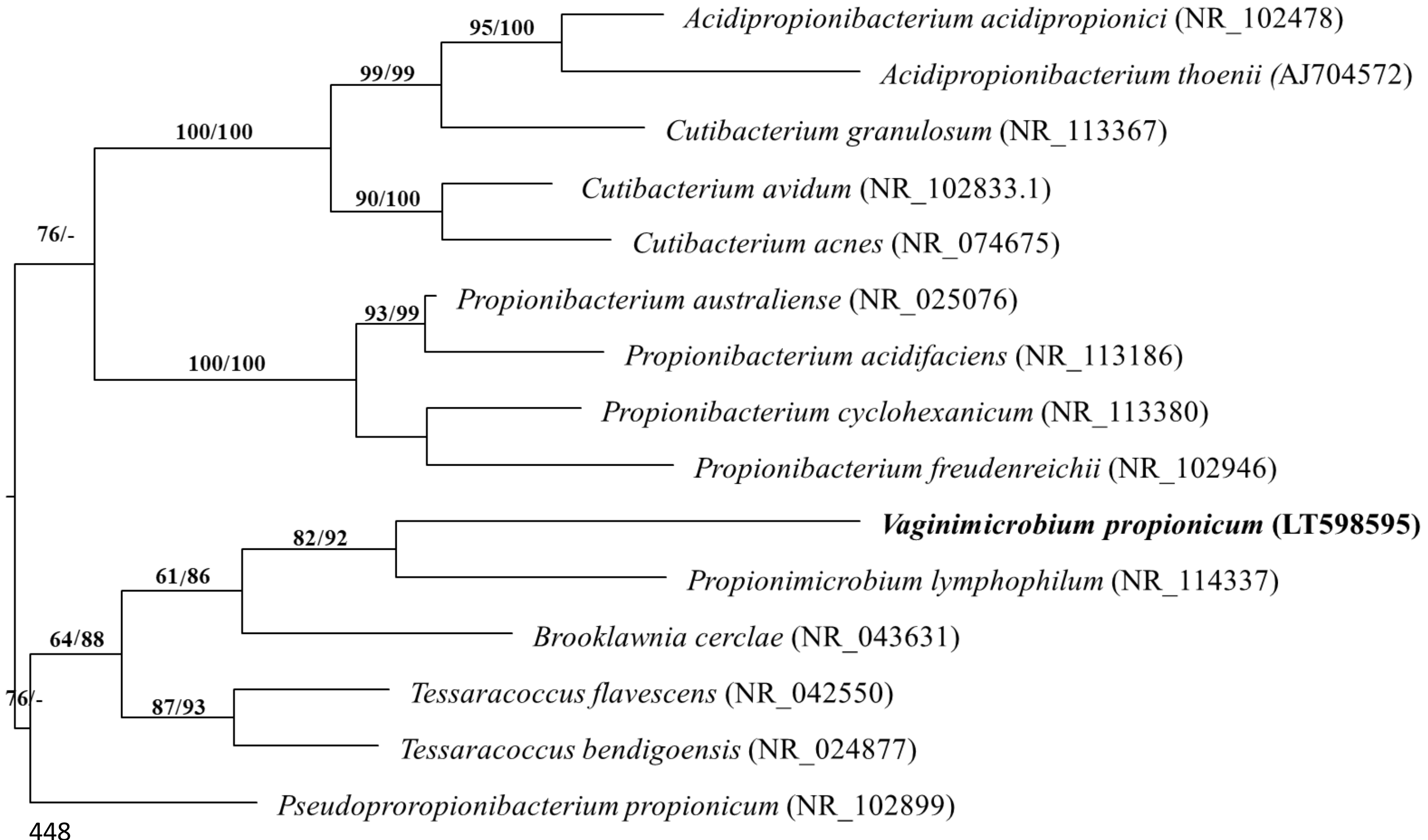
431 51. **Stackebrandt E, Ebers J.** Taxonomic parameters revisited: tarnished gold standards.
432 *Microbiol Today* 2006;33:152.
433 52. **Kim M, Oh HS, Park SC, Chun J**. Towards a taxonomic coherence between average
434 nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of
435 prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–351.
436 53. **Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, *et al.*** Uniting the classification
437 of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev*
438 *Microbiol* 2014;12:635–645.
439



440

441 **Figure 1.** Phylogenetic analysis based on the 16S rRNA gene sequences comparison highlighting
442 the position of *Vaginimicrobium propionicum* strain Marseille-P3275^{T} relative to other closely
443 related strains. GenBank accession numbers are indicated in parentheses. Sequences were aligned
444 using Muscle v3.8.31 with default parameters and, phylogenetic inferences were performed using
445 neighbor-joining method with the software MEGA version 6. The scale bar represents a 1%
446 nucleotide sequence divergence.

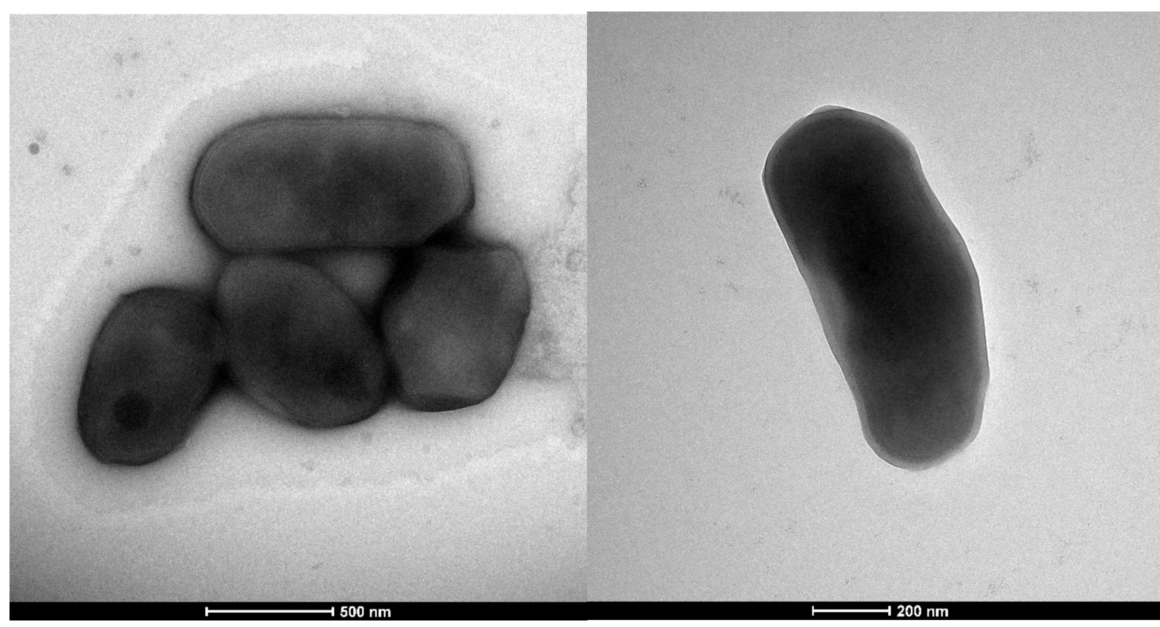
447



448

449 **Figure 2.** ML tree inferred under the GTR+GAMMA model and rooted by midpoint-rooting. The
450 branches are scaled in terms of the expected number of substitutions per site. The numbers above
451 the branches are support values when larger than 60% from ML (left) and MP (right)
452 bootstrapping.

453



454 **Figure 3.** Transmission electron microscopy of *Vaginimicrobium propionicum* strain Marseille-

455 P3275^T using a Tecnai G20 transmission electron microscope (FEI Company).

456 **Table 1.** Differential phenotypic features between strain Marseille-P3275^T and closely related species.

457 Strains: 1, *Vaginimicrobium propionicum* Marseille-P3275^T; 2, *Propionimicrobium lymphophilum* DSM 4903^T; 3,

458 *Acidipropionibacterium acidipropionici* NCFB 570^T; 4, *Acidipropionibacterium thoenii* DSM 20276^T; 5, *Cutibacterium acnes* 6609^T;

459 6, *Cutibacterium avidum* DSM 4901^T; 7, *Propionibacterium acidifaciens* DSM 21887^T; 8, *Propionibacterium freudenreichii* DSM

460 20271^T; 9, *Pseudopropionibacterium propionicum* DSM 43307^T and 10, *Tessaracoccus flavescens* DSM 18582^T.

Character	1	2	3	4	5	6	7	8	9	10
Indole production	-	-	-	-	+	-	-	-	-	-
Nitrate reduction	-	v	+	-	+	-	-	-	+	+
Hydrolysis of										
Aesculin	-	-	+	+	-	+	-	+	-	w
Gelatin	-	-	-	-	+	+	-	-	+	-
Enzyme activities										
Catalase	-	v	v	+	+	+	-	+	-	+
N-acetyl-β-glucosaminidase	-	nd	+	+	+	+	nd	nd	-	-
β-Galactosidase	-	nd	nd	nd	w	-	nd	nd	+	+
Fermentation of										
Arabinose	+	-	+	-	-	+	nd	+	-	+
Cellobiose	-	-	+	-	-	-	-	-	-	+
Glucose	+	+	+	+	+	+	+	+	+	+
Lactose	-	-	+	+	-	+	+	-	v	-
Maltose	+	+	+	+	-	+	+	-	+	+
Mannitol	-	-	+	-	-	-	+	-	+	-
Mannose	-	-	+	+	+	+	+	+	v	+
Melezitose	+	-	+	+	-	+	-	-	-	-
Rhamnose	-	-	+	-	-	-	+	-	-	-
Xylose	+	nd	+	+	-	-	nd	-	-	+

461

462 +, Positive; -, negative; v, variable; and nd, not available data. Data were obtained from the original descriptions of species.

463 **Table 2.** Cellular fatty acid composition (%) of strain Marseille-P3275^T compared to those of closest related species.

464 Strains: 1, *V. propionicum* Marseille-P3275^T (data from this study); 2, *P. lymphophilum* CCUG 27816^T; 3, *A. acidipropionici* CCUG

465 36815^T; 4, *A. thoenii* CCUG 28149^T; 5, *C. acnes* CCUG 1794^T; 6, *C. avidum* CCUG 36734^T; 7, *P. acidifaciens* CCUG 57100^T; 8, *P.*

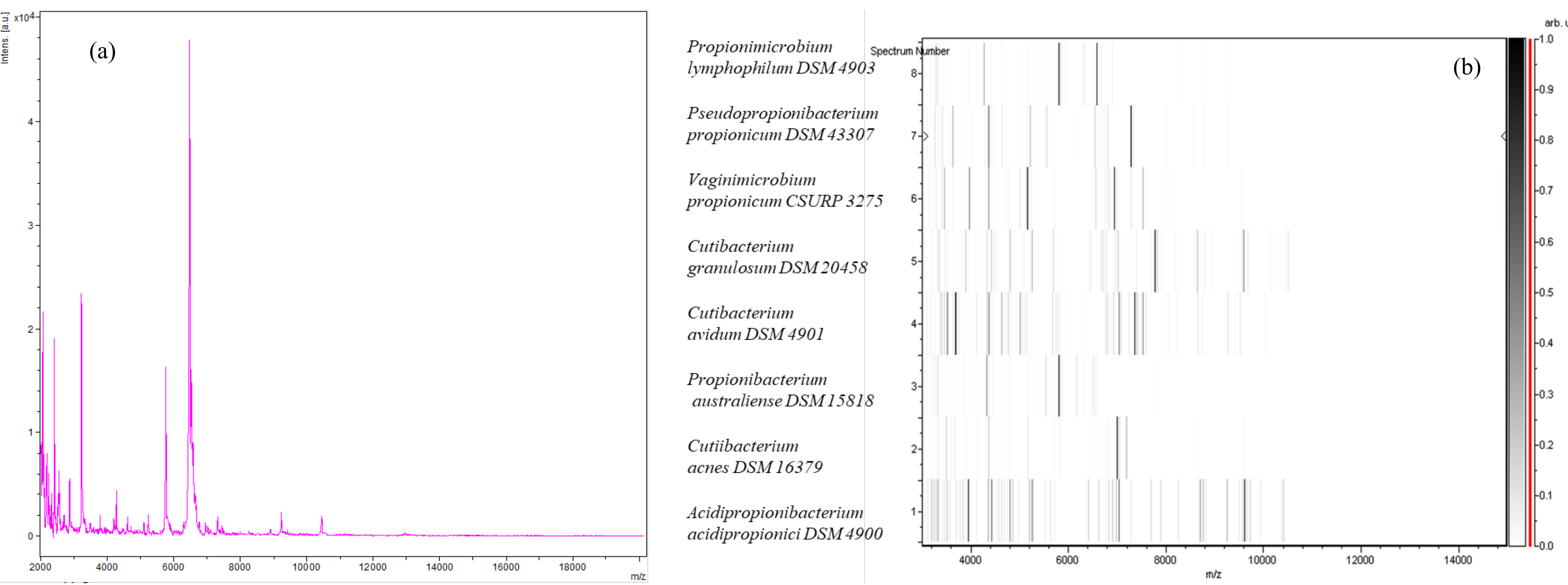
466 *freudenreichii* CCUG 18835a^T; 9, *P. propionicum* CCUG 4939^T and 10, *T. flavescens* DSM 18582^T.

Fatty acids	**1**	**2***	**3***	**4***	**5***	**6***	**7***	**8***	**9***	**10**†
$C_{12:0}$	TR	-	-	-	-	-	-	-	-	-
$C_{14:0}$	2.6	3.3	-	-	-	TR	-	TR	7.6	1.4
iso-$C_{14:0}$	2.5	-	TR	-	-	-	-	-	3.2	2.4
$C_{14:0}$ 2OH	-	-	TR	-	TR	1.5	TR	4.3	-	-
$C_{14:0}$ 3OH	-	-	3.6	2.7	-	2.7	**10.8**	-	-	-
iso-$C_{14:0}$ 3OH	-	-	3.2	1.7	6.1	7.6	2.5	1.1	-	1.2
$C_{15:0}$	4.7	1.6	9.8	9.6	-	3.0	**19.0**	TR	**13.1**	-
antesio-$C_{15:0}$	7.3	**22.7**	**19.2**	**16.2**	8.7	8.8	**20.7**	**57.4**	**33.3**	**49.6**
iso-$C_{15:0}$	**16.9**	-	**45.8**	**23.0**	**49.8**	**43.2**	**24.5**	5.2	**29.4**	3.2
$C_{16:0}$	**28.9**	**32.2**	2.4	4.4	5.2	3.5	1.6	3.7	3.9	**11.5**
iso-$C_{16:0}$	2.7	-	TR	1.1	-	-	-	TR	-	5.0
$C_{16:1\omega7}$	TR	-	-	-	-	-	-	-	-	-
$C_{17:0}$	TR	1.2	TR	3.0	-	-	1.3	TR	-	-
antesio-$C_{17:0}$	TR	3.0	TR	1.6	1.0	1.2	2.7	**10.6**	-	-
iso-$C_{17:0}$	TR	-	1.3	3.6	6.3	3.3	1.6	1.6	-	-
iso-$C_{17:1}$	-	6.9	-	1.0	1.6	3.3	2.5	-	-	-
$C_{18:0}$	7.1	1.3	TR	2.7	4.2	-	-	-	-	**17.5**
$C_{18:1\omega7}$	TR	-	-	**18.7**	TR	-	-	-	-	-
$C_{18:1\omega9}$	**18.1**	7.6	2.8	3.2	3.7	-	-	TR	-	2.2
$C_{18:2\omega6}$	6.6	**17.4**	2.6	4.5	6.5	4.0	1.2	1.4	6.1	-

467 *data are from CCUG (http://www.ccug.se). †data are from Lee and Lee, 2008.

468 Predominant products are shown in bold; TR, trace amounts < 1 %; -, not detected.

469 SUPPLEMENTARY DATA



471 **Figure S1. a.** Reference mass spectrum of strain Marseille-P3275^T. **b.** Gel view comparing strain Marseille-P3275^T to other species

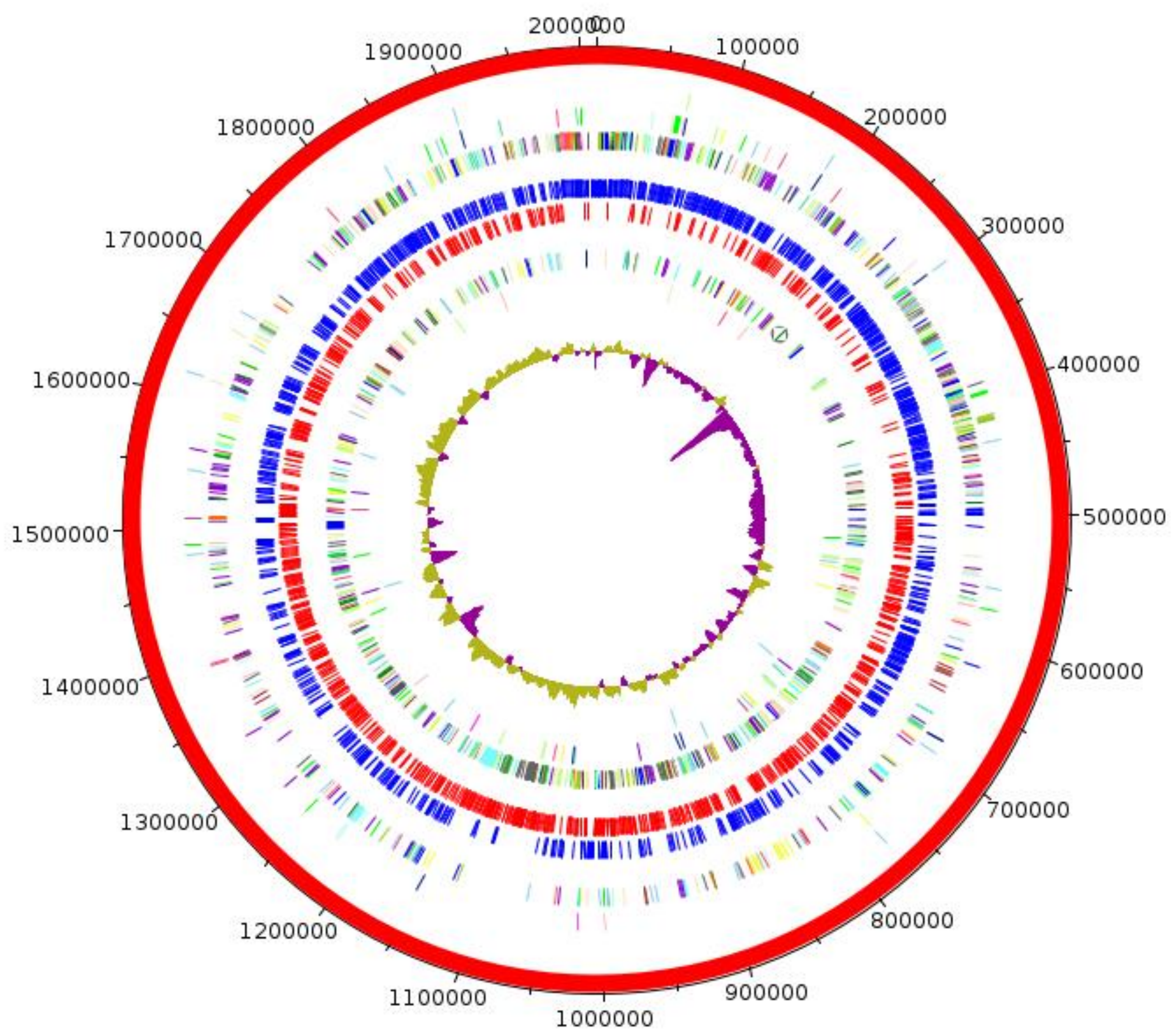
472 within the family.

473 The gel view displays the raw spectra of loaded spectrum files arranged in a pseudo-gel-like look. The x-axis records the m/z value.

474 The left y-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by

475 a gray scale scheme code. The right y-axis indicates the relation between the color of a peak and its intensity, in arbitrary units.

476 Displayed species are indicated on the left.



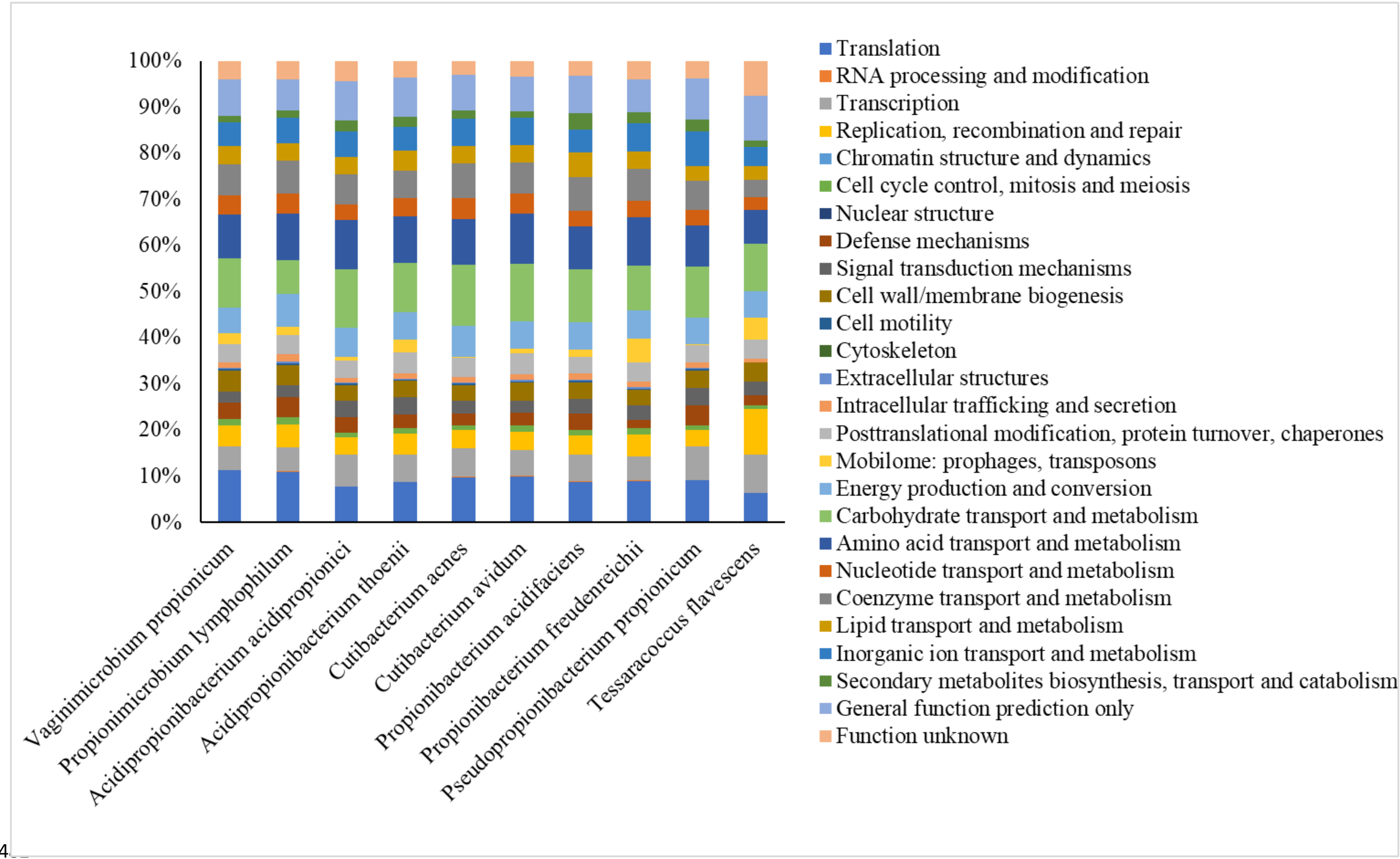
477

478 **Figure S2.** Graphical circular map of strain Marseille-P3275^{T}'s genome.

479 From outside to the center: Contigs (red/gray), COG category of genes on the forward strand (three

480 circles), genes on forward strand (blue circle), genes on the reverse strand (red circle), COG

481 category on the reverse strand (three circles), G+C content**.**



482

483 **Figure S3.** Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins of

484 *Vaginimicrobium propionicum* strain Marseille-P3275^T among other species.

485 **Table S1.** Nucleotide content and gene count levels of the genome

Attribute	Genome	
	Value	**% total[a]**
Genome size (bp)	2,010,679	100
G+C content (bp)	1,018,208	50.64
Coding region (bp)	1,858,683	92.44
Total genes	1,921	100
RNA genes	52	2.71
Protein-coding genes	1,869	97.29
Protein assigned to COGs	1,228	65.70
Protein associated to ORFans	92	4.92
Protein associated to hypothetical protein	337	18.03
Protein with function prediction	1,386	74.16
Protein with peptide signals	184	9.84
Protein with TMH	400	21.40
Gene associated to PKS or NRPS	9	0.48
Gene associated to bacteriocin genes	16	0.86
Genes associated to mobilome	777	41.57
Genes associated to virulence	345	18.46
Genes associated to toxin / antitoxin	65	3.48
Genes with paralogues (evalue:$1e^{-10}$)	239	12.79
Genes with paralogues (evalue:$1e^{-25}$)	121	6.47
Genes larger than 5000 nucleotides	2	0
Genes with Pfam-A domains	1,710	89

486 [a] The total is based on either the size of the genome in base pairs or the total number of protein

487 coding genes in the annotated genome.

488 **Table S2.** Number of genes associated with the 25 general COG functional categories

Code	Value	% of total	Description
[J]	151	8.08	Translation
[A]	1	0.05	RNA processing and modification
[K]	69	3.69	Transcription
[L]	61	3.26	Replication, recombination and repair
[B]	0	0	Chromatin structure and dynamics
[D]	20	1.07	Cell cycle control, mitosis and meiosis
[Y]	0	0	Nuclear structure
[V]	49	2.62	Defense mechanisms
[T]	31	1.66	Signal transduction mechanisms
[M]	62	3.32	Cell wall/membrane biogenesis
[N]	4	0.21	Cell motility
[Z]	0	0	Cytoskeleton
[W]	4	0.21	Extracellular structures
[U]	14	0.75	Intracellular trafficking and secretion
[O]	55	2.94	Post-translational modification, protein turnover, chaperones
[X]	31	1.66	Mobilome: prophages, transposons
[C]	74	3.96	Energy production and conversion
[G]	143	7.65	Carbohydrate transport and metabolism
[E]	125	6.69	Amino acid transport and metabolism
[F]	56	2.70	Nucleotide transport and metabolism
[H]	89	4.76	Coenzyme transport and metabolism
[I]	53	2.83	Lipid transport and metabolism
[P]	69	3.69	Inorganic ion transport and metabolism
[Q]	20	1.07	Secondary metabolites biosynthesis, transport and catabolism
[R]	106	5.67	General function prediction only
[S]	54	2.89	Function unknown
_	641	34.30	Not in COGs

489

490 **Table S3.** Genome comparison of closely related species to strain Marseille-P3275^T

Species	Strain	INSDC identifier[a]	Size(Mb)	GC Percent	Gene Content
Vaginimicrobium propionicum	**Marseille-P3275**	**LT706985.1**	**2.01**	**50.64**	**1,869**
Propionimicrobium lymphophilum	DSM 4903	AUIB00000000.1	2.04	56.06	1,839
Acidipropionibacterium acidipropionici	CGMCC 1.2230	CP013126.1	3.65	68.77	3,162
Acidipropionibacterium thoenii	DSM 20276	AUHZ00000000.1	2.94	67.97	2,645
Cutibacterium acnes	6609	CP002815.1	2.56	60.01	2,348
Cutibacterium avidum	ATCC25577	AGBA00000000.1	2.55	63.42	2,376
Propionibacterium acidifaciens	F0233	ACVN00000000.2	3.02	70.36	2,913
Propionibacterium freudenreichii	DSM 20271	CP010341.1	2.65	67.34	2,320
Pseudopropionibacterium propionicum	F0230a	CP002734.1	3.45	66.06	2,938
Tessaracoccus flavescens	SST-39T	CP019607.1	3.60	68.3	3,243

491 [a] INSDC: International Nucleotide Sequence Database Collaboration.

492 **Table S4:** dDDH values obtained by comparison of all studied genomes Using GGDC, Formula 2 (DDH Estimates Based on

493 Identities/HSP length)*

	V. propionicum	***P. lymphophilum***	***A. acidipropionici***	***A. thoenii***	***C. acnes***	***C. avidum***	***P. acidifaciens***	***P. freudenreichii***	***P. propionicum***	***T. flavescens***
Vaginimicrobium propionicum	100	25.50±2.40	29.60±2.45	26.60±2.40	26.40±2.45	24.60±2.40	24.40±2.40	21.30±2.35	26.70±2.40	29.40±2.45
Propionibacterium lymphophilum		100	21.90±2.35	21.40±2.35	20.40±2.3	21.20±2.35	20.00±2.30	18.90±2.30	19.10±2.35	21.00±2.35
Acidipropionibacterium acidipropionici			100	23.10±2.40	20.40±2.35	21.20±2.35	19.60±2.30	23.00±2.35	19.40±2.30	19.70±2.30
Acidipropionibacterium thoenii				100	20.60±2.35	20.80±2.35	19.50±2.30	27.00±2.40	20.30±2.35	20.80±2.30
Cutibacterium acnes					100	23.50±2.35	20.80±2.30	20.50±2.35	22.20±2.35	22.00±2.35
Cutibacterium avidum						100	18.80±2.30	19.40±2.30	20.70±2.30	19.20±2.30
Propionibacterium acidifaciens							100	20.50±2.30	18.40±2.30	18.60±2.25
Propionibacterium freudenreichii								100	19.40±2.30	20.90±2.35
Pseudopropionibacterium propionicum									100	19.70±2.30
Tessaracoccus flavescens										100

494 *The confidence intervals indicate the inherent uncertainty in estimating DDH values from intergenomic distances based on models

495 derived from empirical test data sets (which are always limited in size).

496 **Table S5.** Average amino acid identity (AAI) values (%) between strain Marseille-P3275^{T} and closest related species.

	V. propionicum	*P. lymphophilum*	*A. acidipropionici*	*A. thoenii*	*C. acnes*	*C. avidum*	*P. acidifaciens*	*P. freudenreichii*	*P. propionicum*
Vaginimicrobium propionicum	-								
Propionibacterium lymphophilum	66.08	-							
Acidipropionibacterium acidipropionici	52.66	52.60	-						
Acidipropionibacterium thoenii	52.55	52.52	73.38	-					
Cutibacterium acnes	52.54	52.90	65.87	65.02	-				
Cutibacterium avidum	53.14	53.53	66.64	65.90	85.43	-			
Propionibacterium acidifaciens	55.37	55.21	55.41	55.36	55.05	55.12	-		
Propionibacterium freudenreichii	55.87	55.55	56.65	56.85	55.18	55.07	65.19	-	
Pseudopropionibacterium propionicum	52.04	52.61	53.13	53.22	53.39	53.57	54.34	53.49	-
Tessaracoccus flavescens	52.44	52.35	53.10	53.40	53.04	53.55	53.36	53.61	60.74

497

Article 12:

Taxonogenomics and description of *Vaginella massiliensis* gen. nov. sp. nov., strain Marseille-P2517^T, a new bacterial genus isolated from the human vagina.

Khoudia Diop, Florence Bretelle, Caroline Michelle, Magali Richez, Jaishriram Rathored, Didier Raoult, Pierre-Edouard Fournier, Florence Fenollar

Published in New Microbes and New Infections

Taxonogenomics and description of *Vaginella massiliensis* gen. nov., sp. nov., strain Marseille P2517^T, a new bacterial genus isolated from the human vagina

K. Diop[1], **F. Bretelle**[2], **C. Michelle**[1], **M. Richez**[1], **J. Rathored**[1], **D. Raoult**[1,3], **P.-E. Fournier**[1] and **F. Fenollar**[1]

1) *Institut hospitalo-universitaire Méditerranée-infection, Aix-Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine,* 2) *Department of Gynecology and Obstetrics, Gynépole, Aix-Marseille Université, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, Marseille, France and* 3) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

An obligate aerobic, Gram-negative, nonmotile and nonsporulating rod designated Marseille P2517 was isolated from the vaginal flora. We describe its features, annotate the genome and compare it to the closest species. The 16S rRNA analysis shows 93.03% sequence similarity with *Weeksella virosa,* the phylogenetically closest species. Its genome is 2 434 475 bp long and presents 38.16% G+C. On the basis of these data, it can be considered as a new genus in the *Flavobacteriaceae* family, for which we proposed the name *Vaginella massiliensis* gen. nov., sp. nov. The type strain is Marseille P2517^T.
© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

Keywords: Bacterial vaginosis, *Flavobacteriaceae*, genome, vaginal flora, *Vaginella massiliensis*
Original Submission: 13 October 2016; **Accepted:** 9 November 2016
Article published online: 25 November 2016

Corresponding author: F. Fenollar, Aix-Marseille Université, Institut hospitalo-universitaire Méditerranée-infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine, 27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France
E-mail: florence.fenollar@univ-amu.fr

Introduction

Bacterial vaginosis is a common yet poorly understood condition affecting women of childbearing age in both industrialized and developing countries. Bacterial vaginosis is characterized simultaneously by an abnormal loss of the normal Doderleïn flora accompanied by an unexplained overgrowth of anaerobic bacteria that were previously minor in the vagina [1,2]. In pregnant women, this vaginal dysbiosis is the consequence of certain complications such as miscarriage, preterm birth or chorioamnionitis [3]. Bacterial vaginosis is mostly treated with antibiotics, mainly metronidazole and clindamycin, but treatment frequently fails; the relapse rate is estimated at 50% at 6 months [4,5].

In order to describe the vaginal flora as fully as possible and to better understand the condition in order to provide better treatment, we studied the vaginal microbiota from healthy women and patients with bacterial vaginosis using the culturomics concept. This is based on the multiplication of culture conditions (atmosphere, media, and temperature) and a rapid bacterial identification using matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) [6].

A new member of *Flavobacteriaceae* was therefore isolated. Proposed in 1985 by Jooste, it was only in 1992 that the name of this family was validated by Reichenbach. The type genus is *Flavibacterium* (http://www.bacterio.net/flavobacteriaceae.html) [7]. The family currently contains 114 genera (http://www.bacterio.net/-classifgenerafamilies.html). Some species are found in soil and the marine environment, while others are pathogens found in fish and the human urogenital tract [8].

The classical bacterial description presents some limitations. Hence, in order to describe a new bacterium, our laboratory introduced taxonogenomics, a new approach that complements classic features with the proteomic information obtained by MALDI-TOF MS and the description of the annotated whole genome [9,10].

In the following section, we describe the *Vaginella massiliensis* strain Marseille P2517^T (= DSM 102346^T = CSUR P2517), a new genus isolated from a vaginal swab taken from a healthy 22-year-old French woman without bacterial vaginosis.

Materials and Methods

Sample collection

As previously described [11], a vaginal sample was taken from a healthy 22-year-old French woman without bacterial vaginosis at

New Microbe and New Infect 2017; **15:** 94–103
© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
http://dx.doi.org/10.1016/j.nmni.2016.11.006

La Timone Hospital in Marseille (France) in January 2016 using a Sigma Transwab (Medical Wire, Corsham, United Kingdom). The study was authorized by the local IFR48 ethics committee (Marseille, France) under agreement number 09-022. The patient also provided written informed consent. When the sample was collected, she was not receiving any antibiotic treatment.

Strain identification by MALDI-TOF MS and 16S rRNA sequencing

The vaginal sample was first preincubated in a blood culture bottle (BD Diagnostics, Le Pont-de-Claix, France) supplemented with 4 mL rumen and filtered at 0.2 μm using a pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France) and 3 mL sheep's blood (bioMérieux, Marcy l'Etoile, France). After 7 days of preincubation, 50 μL of the supernatant was inoculated on Chocolat PolyViteX (PVX) agar (BD Diagnostics). After 2 days of incubation at 37°C in aerobic conditions, purified colonies were deposited in duplicate on a MSP 96 MALDI-TOF MS target plate (Bruker Daltonics, Leipzig, Germany), and, as previously described, 1.5 μL of matrix solution was added to each spot. Identification was carried out using a Microflex spectrometer (Bruker) [12], which compares the protein spectra found to those in the Bruker database (constantly updated with spectra of new species discovered in our laboratory). If the score is greater than 1.9, the bacterium is correctly identified. In contrast, if no spectra match the database, and for unidentified bacteria with a clear spectrum, 16S rRNA gene sequencing is performed [13]. As Stackebrandt and Ebers suggested [14], if the 16S rRNA sequence similarity value is lower than 95% or 98.7%, the strain is defined as a new genus or species, respectively.

Phylogenetic tree

A custom Python script was used to automatically retrieve all species from the same order of the new genus and to download 16S sequences from National Center for Biotechnology Information (NCBI) by parsing NCBI eUtils results and the NCBI taxonomy page. This only retains sequences from type strains. In the event of multiple sequences for one type strain, it selects the sequence obtaining the best identity rate from the BLASTn alignment with our sequence. The script then separates 16S sequences into two groups: one containing the sequences of strains from the same family (group A) and one containing the others (group B). Finally, it only retains the 15 closest strains from group A and the closest one from group B. If it is impossible to retrieve 15 sequences from group A, the script selects more sequences from group B to achieve at least nine strains from both groups.

Growth conditions

The ideal growth was tested by cultivating the strain Marseille P2517^T on Colombia agar with 5% sheep's blood incubated at different temperatures (25, 28, 37, 45 and 56°C) and different atmospheres (anaerobic, microaerophilic and aerobic). The anaerobic and microaerophilic atmospheres were generated using, respectively, GENbag anaer and GENbag microaer systems (bioMérieux). Salinity and pH conditions were also tested at different concentrations of NaCl (0, 5, 15 and 45%) and different pH (5, 6, 6.5, 7 and 8.5).

Morphological, biochemical and antibiotic susceptibility tests

Sporulation, motility, Gram stain, catalase and oxidase tests were performed using standard test procedures (https://www.gov.uk/government/collections/standards-for-microbiology-investigations-smi#test-procedures). In order to observe cell morphology, samples were fixed with 2.5% glutaraldehyde in 0.1 M cacodylate buffer for at least an hour at 4°C. One drop of cell suspension was deposited for approximately 5 minutes 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 1% ammonium molybdate solution in filtered water at room temperature. Electron micrographs were acquired using a Tecnai G^{20} Cryo (FEI Company, Limeil-Brevannes, France) transmission electron microscope operated at 200 keV. Biochemical characteristics were studied using API ZYM, API 20NE and API 50CH strips (bioMérieux) according to the manufacturer's instructions.

A cellular fatty acid methyl ester (FAME) analysis was performed by gas chromatography mass spectrometry (GC/MS). Two samples were prepared with approximately 90 mg of bacterial biomass per tube collected from several culture plates. FAMEs were prepared as described by Sasser (http://www.midi-inc.com/pdf/MIS_Technote_101.pdf). GC/MS analyses were carried out using a Clarus 500 gas chromatograph equipped with a SQ8S MS detector (Perkin Elmer, Courtaboeuf, France). FAME extracts (2 mL) were volatilized at 250°C (split 20 mL/min) in a Focus liner with wool and separated on an Elite-5MS column (30 m, 0.25 mm i.d., 0.25 mm film thickness) using a linear temperature gradient (70–290°C at 6°C/min), enabling the detection of C4 to C24 fatty acid methyl esters. Helium flowing at 1.2 mL/min was used as a carrier gas. The MS inlet line was set at 250°C and EI source at 200°C. Full scan monitoring was performed from 45 to 500 *m/z*. All data were collected and processed using Turbomass 6.1 (Perkin Elmer). FAMEs were identified through a spectral database search using MS Search 2.0 operated with the Standard Reference Database 1A (National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK). Retention time correlations with estimated nonpolar retention

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **15**, 94–103
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

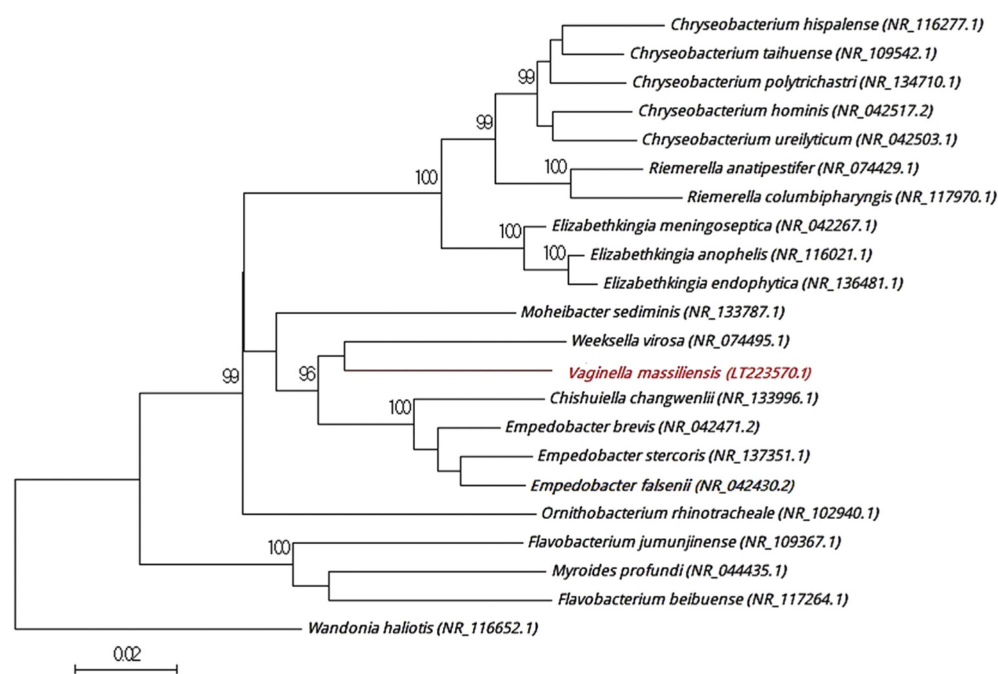


FIG. 1. Phylogenetic tree highlighting position of *Vaginella massiliensis* strain Marseille P2517^T relative to other close strains. GenBank accession numbers of each 16S rRNA are noted immediately after name. Sequences were aligned using Muscle 3.8.31 with default parameters, and phylogenetic inferences were obtained by neighbour-joining method with 500 bootstrap replicates within MEGA6 software. Only bootstraps >95% are shown. Scale bar represents 0.02% nucleotide sequence divergence.

indices from the NIST database were obtained using a 37-component FAME mix (Supelco; Sigma-Aldrich, Saint-Quentin Fallavier, France); FAME identifications were confirmed using this index.

The antibiotic susceptibility of strain Marseille P2517^T was tested using the disk diffusion method (Sirscan discs, Perols, France).

Genomic DNA preparation

Strain Marseille P2517^T was grown in aerobic conditions at 37°C using Columbia agar enriched with 5% sheep's blood (bioMérieux) after 48 hours on four petri dishes. Bacteria were resuspended in 500 μL of Tris-EDTA (TE) buffer; 150 μL of this suspension was diluted in 350 μL 10× TE buffer, 25 μL proteinase K and 50 μL sodium dodecyl sulfate for lysis treatment. This preparation was incubated overnight at 56°C. DNA was purified using phenol/chloroform/isoamylalcohol successively for extraction, followed by ethanol precipitation at −20°C for at least 2 hours each. After centrifugation, the DNA was suspended in 65 μL TE buffer.

TABLE 1. Classification and general features of *Vaginella massiliensis* strain Marseille P2517^T

Characteristic	Term
Current classification	Domain: *Bacteria* Phylum: *Bacteroidetes* Class: *Flavobacteriia* Order: *Flavobacteriales* Family: *Flavobacteriaceae* Genus: *Vaginella* Species: *Vaginella massiliensis* Type strain: Marseille P2517
Gram stain	Negative
Cell shape	Rod
Motility	Nonmotile
Sporulation	Non–spore forming
Temperature range	Aerobic
Optimum temperature	37°C

Genome sequencing and assembly

Genomic DNA (gDNA) of strain Marseille P2517^T was sequenced on the MiSeq Technology (Illumina, San Diego, CA, USA) using the mate-pair strategy. The gDNA was barcoded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). gDNA was quantified using a Qubit assay with the High Sensitivity kit (Life Technologies, Carlsbad, CA, USA) to 153 ng/μL. The mate-pair library was prepared with 1.5 μg of gDNA using the Nextera mate pair Illumina guide. The gDNA sample was simultaneously fragmented and tagged with a mate-pair junction adapter. The fragmentation pattern was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 labchip. The DNA fragments ranged in size from 1.5 to 11 kb with an optimal size of 7.455 kb. No size selection was

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **15**, 94–103
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

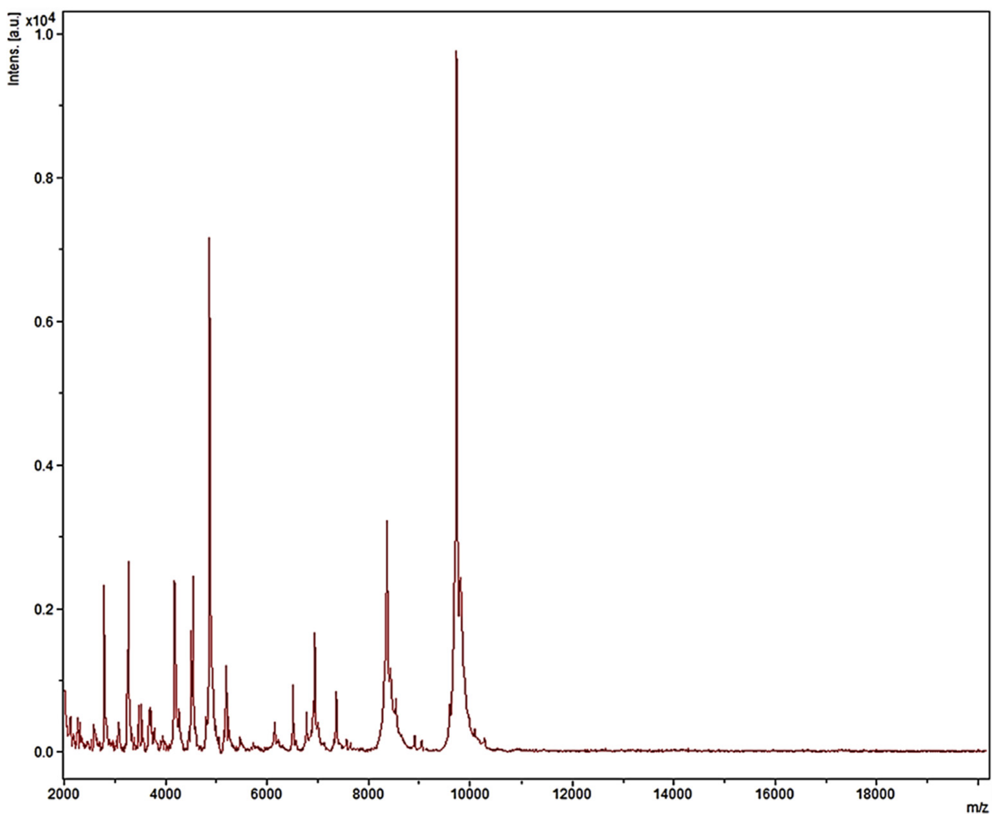


FIG. 2. Reference mass spectrum from *Vaginella massiliensis* strain Marseille P2517^T.

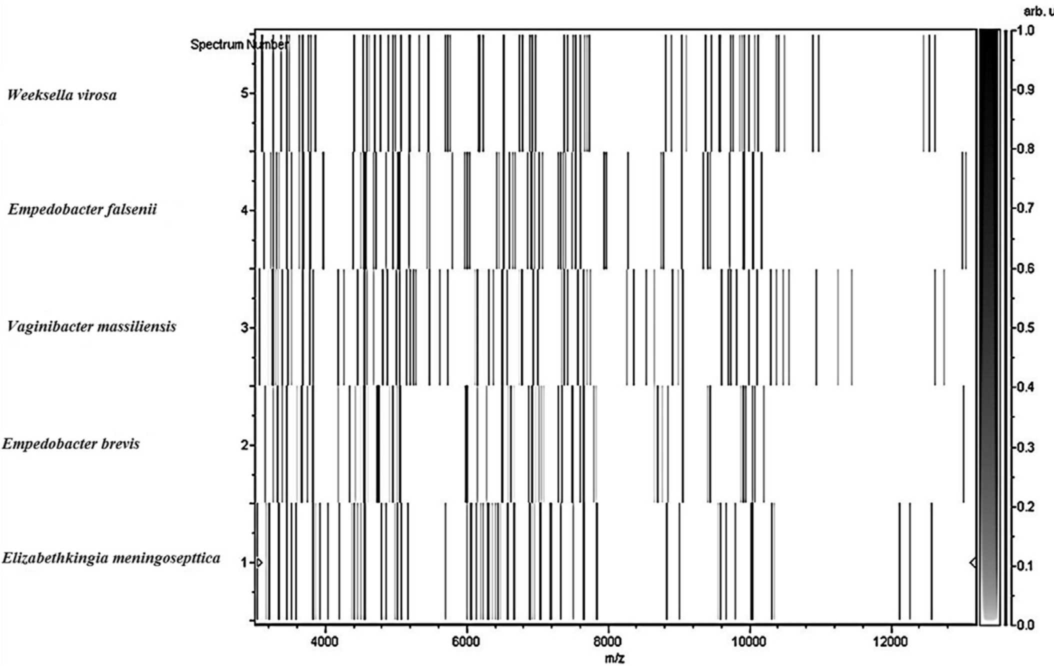


FIG. 3. Gel view comparing *Vaginella massiliensis* strain Marseille P2517^T to other species within genera *Empedobacter*, *Elizabethkingia* and *Weeksella*. Gel view displays raw spectra of loaded spectrum files arranged in pseudo-gel-like look. *x*-axis records *m/z* value. Left *y*-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity is expressed by greyscale scheme code. Right *y*-axis indicates relation between color of peak and its intensity in arbitrary units. Displayed species are indicated at left.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **15**, 94–103
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

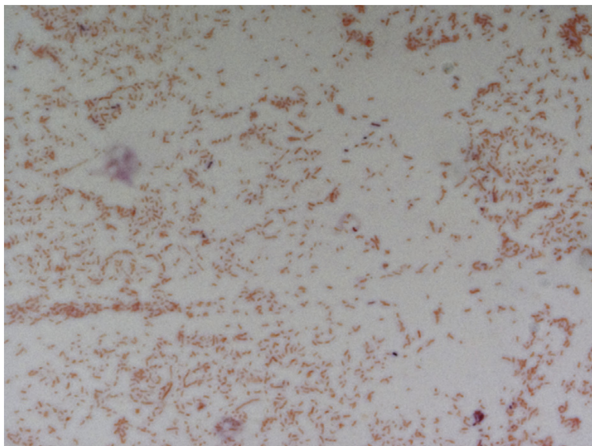


FIG. 4. Gram staining of *Vaginella massiliensis* strain Marseille P2517^T.

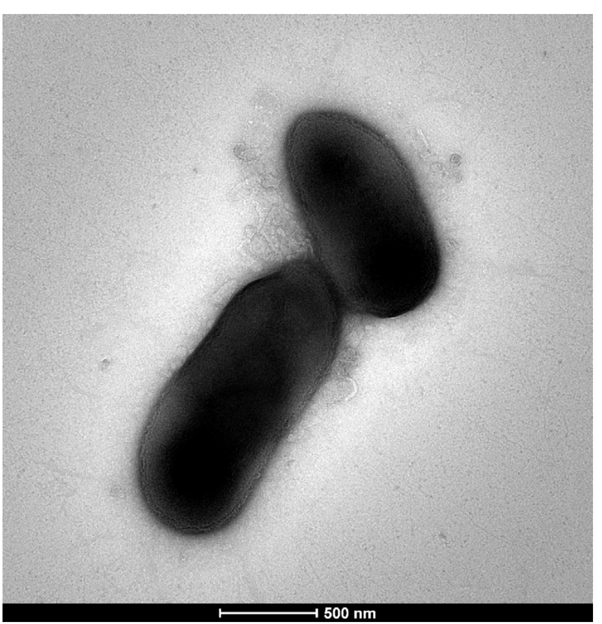


FIG. 5. Electron micrographs of *Vaginella massiliensis* strain Marseille P2517^T using Tecnai G^{20} Cryo (FEI Company) transmission electron microscope operated at 200 keV. Scale bar = 500 nm.

TABLE 2. **Cellular fatty acid composition (%)**

Fatty acid	Name	Mean relative %[a]
15:0 iso	13-Methyl-tetradecanoic acid	60.2 ± 0.6
17:1n5 anteiso	14-Methyl-11-hexadecenoic acid	8.5 ± 0.3
17:0 3-OH	3-Hydroxy-heptadecanoic acid	6.2 ± 0.4
15:0 2-OH	2-Hydroxy-pentadecanoic acid	5.1 ± 0.3
17:1n5 iso	15-Methyl-11-hexadecenoic acid	4.9 ± 0.1
17:0 iso	15-Methyl-hexadecanoic acid	3.7 ± 0.1
5:0 iso	3-Methyl-butanoic acid	3.2 ± 0.7
16:0	Hexadecanoic acid	2.2 ± 0.1
15:0 3-OH	3-Hydroxy-pentadecanoic acid	1.8 ± 0.1
15:1n5 iso	13-Methyltetradec-9-enoic acid	1.1 ± 0.1
18:2n6	9,12-Octadecadienoic acid	1.0 ± 0.1
15:0 anteiso	12-Methyl-tetradecanoic acid	TR
18:1n9	9-Octadecenoic acid	TR
18:0	Octadecanoic acid	TR
16:0 iso	14-Methyl-pentadecanoic acid	TR
16:1n6 iso	14-Methylpentadec-9-enoic acid	TR
15:0	Pentadecanoic acid	TR
17:1n7 anteiso	14-Methylhexadec-9-enoic acid	TR
14:0	Tetradecanoic acid	TR
16:0 3-OH	3-Hydroxy-hexadecanoic acid	TR

TR, trace amounts <1%
[a]Mean peak area percentage ± SD.

performed, and 410.7 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal size of 1115 bp on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies), and the final concentration library was measured at 12.49 nmol/L.

The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 39-hour run at a 2 × 251 bp read length.

A total of 9.2 Gb of information was obtained from a 1042K/mm^2 cluster density with a cluster passing quality control filters of 91.6% (18 078 000 passing filter paired reads). Within this run, the index representation for strain P2517^T was determined to 6.87%. The 1 241 784 paired reads were trimmed, then assembled in nine scaffolds.

Genome annotation and analysis

Open reading frames (ORFs) were predicted using Prodigal software [15] with default parameters. Predicted ORFs spanning a sequencing gap region (containing N) were excluded. We predicted the bacterial proteome sequences using BLASTP (*E* value 1e-03, coverage 0.7 and identity percentage 30) against the Clusters of Orthologous Groups (COGs) database. A search against the NR database [16] was performed if no hit was found using BLASTP with an *E* value of 1e-03 coverage of 0.7 and an identity percentage of 30. An *E* value of 1e-05 was used with sequence lengths smaller than 80 aa. The hmmscan analysis tools were used for searching PFam conserved domains (PFAM-A and PFAM-B domains) on each protein. We used RNAmmer [17] and the tRNAScanSE tool [18] to find ribosomal RNAs genes and tRNA genes respectively. For visualization and for data management of genomic features, we used Artemis [19] and DNA Plotter [20] respectively. For the mean level of nucleotide sequence similarity analysis at the genome level, we used MAGI homemade software. It calculated the average genomic identity of orthologous gene sequences (AGIOS) among compared genomes [21]. The Proteinortho [22] software was incorporated into the MAGI homemade software for detecting orthologous proteins in pairwise genomic comparisons. The corresponding

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **15**, 94–103
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

TABLE 3. Differential characteristics of *Vaginella massiliensis, Weeksella virosa, Empedobacter brevis, Empedobacter falsenii, Chishuiella changwenlii* and *Moheibacter sediminis* [26–29]

Property	*Vaginella massiliensis*	*Weeksella virosa*	*Empedobacter brevis*	*Empedobacter falsenii*	*Chishuiella changwenlii*	*Moheibacter sediminis*
Cell diameter (μm)	0.54–0.68	0.6	NA	NA	0.5–0.6	0.2–0.3
Oxygen requirement	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic
Gram stain	−	−	−	−	−	−
Indole	−	+	NA	+	+	NA
Major fatty acid	iso-C15:0 (60.2%) anteiso-C17:1n5 (8.5%)	iso-C15:0 (46.8%) iso-C17:0 3-OH (13.6%)	iso-C15:0 (24.5%) iso-C17:0 3-OH (17.9%)	iso-C15:0 (24.5%) iso-C17:0 3-OH (17.6%)	iso-C15:0 (19.6%) iso-C17:0 3-OH (17.8%)	iso-C15:0 (43.2%) iso-C17:0 3-OH (24.0%)
%G+C (%mol/L)	38.16	35.9	32.8	32.1	30.0	38.2
Production of:						
Alkaline phosphatase	+	+	NA	+	+	+
Catalase	−	+	NA	+	+	+
Oxidase	+	+	NA	+	+	+
Nitrate reductase	−	−	NA	−	−	−
Urease	−	−	−	+	−	−
β-Galactosidase	−	−	NA	−	−	−
N-acetyl-glucosamine	−	+	−	−	−	−
Acid from:						
L-Arabinose	−	+	−	−	−	−
Mannose	−	−	−	−	+	−
Mannitol	−	−	−	−	+	−
D-Glucose	+	−	−	−	+	−
D-Fructose	−	−	−	−	+	−
D-Maltose	−	+	+	−	+	+
Habitat	Human vagina	Human urinogenital tract	Clinical material	Surgical wound	Freshwater	Sediment

+, positive result; −, negative result; NA, data not available.

TABLE 4. Nucleotide content and gene count levels of genome

Attribute	Value	% of total[a]
Size (bp)	2 434 475	100%
G+C content (bp)	928 861	38.16%
Coding region (bp)	2 208 924	90.73%
Total genes	2395	100%
RNA genes	71	2.96%
Protein-coding genes	2324	100%
Genes with function prediction	1618	69.62%
Genes assigned to COGs	1320	56.8%
Genes with peptide signals	524	22.55%
Genes with transmembrane helices	447	19.23%

COGs, Clusters of Orthologous Groups database.
[a]Total is based on either size of genome in base pairs or total number of protein-coding genes in annotated genome.

genes were then retrieved, and the mean percentage of nucleotide sequence identity among orthologous ORFs was determined using the Needleman-Wunsch global alignment algorithm. Genomes from the genus *Weeksella, Empedobacter* and *Elizabethkingia* were used for the calculation of AGIOS values. The genome of strain Marseille P2517^T (European Molecular Biology Laboratory (EMBL)/European Bioinformatics Institute (EBI) accession number FLMR00000000) was compared to those of *Weeksella virosa* DSM 16922 (NC_015144.1), *Empedobacter brevis* ATCC 43319 (NZ_ARNT00000000.1), *Empedobacter falsenii* strain 282 (NZ_JSYQ00000000.1), *Elizabethkingia anophelis* strain B2D (JNCG00000000.1), and *Elizabethkingia meningoseptica* ATCC 13253 (BARD00000000.1); all these genomes were reannotated with Prodigal. The multiagent software system DAGOBAH [23] was used to perform the annotation and comparison process, including Figenix [24] libraries, which provide pipeline analysis. Genome-to-Genome Distance Calculator (GGDC) analysis was performed using the GGDC Web server, as previously reported [25].

Results

Strain characterization

Strain identification. Strain Marseille P2517^T was first cultivated in January 2016 after 7 days of preincubation in a blood culture bottle supplemented with sheep's blood and rumen under aerobic conditions and then inoculated on PVX agar incubated for 2 days at 37°C in aerobic conditions. Scores of 1.35 and 1.5 were obtained with the MALDI-TOF MS identification, suggesting that this isolate was not in the database and consequently was not a known species. The 16S rRNA sequence (accession no. LT223570) of our strain showed 93.03% nucleotide sequence similarity with *Weeksella virosa,* the phylogenetically closest species with a validly published name (Fig. 1). Because this 16S rRNA nucleotide sequence similarity was lower than 95%, the threshold recommended by Stackebrandt and Ebers [14] for delineating a new genus, strain Marseille P2517^T was classified as a new genus, *Vaginella,* with *Vaginella*

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **15**, 94–103
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

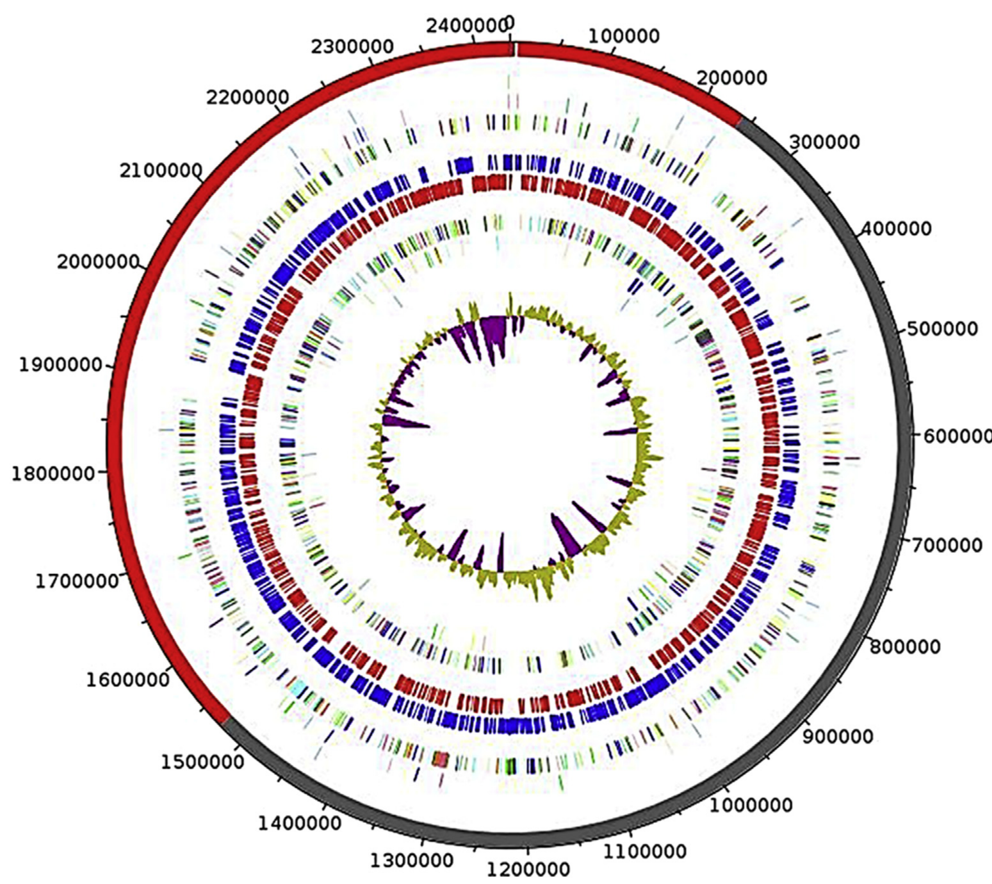


FIG. 6. Graphical circular map of chromosome. From outside to center: contigs (red/grey), COGs category of genes on forward strand (three circles), genes on forward strand (blue circle), genes on reverse strand (red circle), COGs category on reverse strand (three circles), G+C content. COGs, Clusters of Orthologous Groups database.

massiliensis as the type species (Table 1). The reference spectrum of strain Marseille P2517^T (Fig. 2) was then incremented in our database and compared to other known species of the family *Flavobacteriaceae*. Their differences are shown in a gel view in Fig. 3.

Phenotypic characteristics. Cultivated on blood agar for 2 days at 37°C under aerobic conditions, colonies of strain Marseille P2517^T are yellow, opaque, circular and smooth with a diameter of 1.7 to 2 mm. The strain grows only under aerobic conditions at 25, 28 and 37°C, but optimal growth was observed at 37°C after 48 hours of incubation. Its growth also requires a pH ranging from 6.5 to 8.5 and a NaCl concentration lower than 5 g/L. Nonmotile and non–spore forming, strain Marseille P2517^T exhibits positive oxidase activity; however, the catalase activity was negative. Under the microscope, bacterial cells are Gram negative and rod shaped (Fig. 4), and individual cells have a diameter ranging 0.54 to 0.68 μm and a length ranging 1.2 to 1.5 μm (Fig. 5, Table 1).

Using an API 20NE strip, we observed that nitrate and nitrite were not reduced, and urease and indole activities were absent. β-Glucosidase and esculin were not hydrolyzed, unlike gelatin. There was also no assimilation from D-glucose, L-arabinose, D-mannose, D-mannitol, D-maltose and *N*-acetyl-glucosamine. API 50CH showed that strain Marseille P2517^T metabolized only D-glucose; and acid was not produced from other carbohydrates: arabinose, D-ribose, D-xylose, D-adonitol, D-galactose, D-fructose, D-mannose, L-rhamnose, D-mannitol, D-sorbitol, inositol, methyl-αD-mannopyranoside, methyl- αD-glucopyranoside, amygdalin, arbutin, esculin ferric citrate, salicin, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-trehalose, inulin, D-melezitose, D-raffinose, starch, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, fucose, arabitol, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate. API ZYM revealed positive reactions for alkaline phosphatase, esterase, esterase lipase, lipase, leucine arylamidase, cystine arylamidase, valine arylamidase, trypsin, α-chymotrypsin, acid phosphatase and

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **15**, 94–103
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

TABLE 5. Number of genes associated with 25 general COGs functional categories

Code	Value	% of total	Description
J	177	7.6	Translation
A	0	0	RNA processing and modification
K	53	2.2	Transcription
L	78	3.3	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	24	1.0	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	47	2.0	Defense mechanisms
T	36	1.5	Signal transduction mechanisms
M	116	4.9	Cell wall/membrane biogenesis
N	16	0.6	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	19	0.8	Intracellular trafficking and secretion
O	81	3.4	Post-translational modification, protein turnover, chaperones
X	10	0.4	Mobilome: prophages, transposons
C	81	3.4	Energy production and conversion
G	56	2.4	Carbohydrate transport and metabolism
E	111	4.7	Amino acid transport and metabolism
F	51	2.1	Nucleotide transport and metabolism
H	94	4.0	Coenzyme transport and metabolism
I	87	3.7	Lipid transport and metabolism
P	106	4.5	Inorganic ion transport and metabolism
Q	32	1.3769363	Secondary metabolites biosynthesis, transport and catabolism
R	109	4.6901894	General function prediction only
S	55	2.3666093	Function unknown
—	1004	43.201378	Not in COGs

COGs, Clusters of Orthologous Groups database.

naphtol-AS-BI-phosphohydrolase. Reactions for other enzymes such as galactosidase (α and β) and oxidases were negative. FAME analysis demonstrated that the most abundant compound was 13-methyl-tetradecanoic acid (60%). This strain shows many saturated and unsaturated branched fatty acids (iso and anteiso). Several hydroxy fatty acids were also present (Table 2).

Resistant to gentamicin (CN 500 μg) and metronidazole (Met 4), cells were sensitive to ceftriaxone (CRO 30 μg), ciprofloxacin (CIP 5 μg), clindamycin (DA 15 μg), colistin (CT 50 μg), doxycycline (DO 30 μg), erythromycin (E 15 μg), fosfomycin (POS 50 μg), oxacillin (OX 5 μg), penicillin (P 10 μg), rifampicin (RA 30 μg), teicoplasmin (TEC 30 μg), trimethoprim/sulfamethoxazole (SXT 25 μg), vancomycin (VA 30 μg) and imipenem (MP 10).

All the phenotypic characteristics of strain Marseille P2517^T were compared to those of the closely related *Flavobacteriaceae* species [26–29] (Table 3).

Genome properties

The draft genome of strain Marseille P2517^T is 2 434 475 bp long with 38.16 mol% G+C content (Table 4, Fig. 6). It is composed of nine scaffolds (composed of ten contigs). On the 2395 predicted genes, 2324 were protein-coding genes and 71 were RNAs (five 5S rRNA, four 16S rRNA, two 23S rRNA and 60 tRNA genes). A total of 1618 genes (69.62%) were assigned a putative function (by COGs or by NR blast), and 123 genes were identified as ORFans (5.29%). The remaining genes were annotated as hypothetical proteins (528 genes, 22.72%). Table 5 provides the distribution of the genes into COGs functional categories.

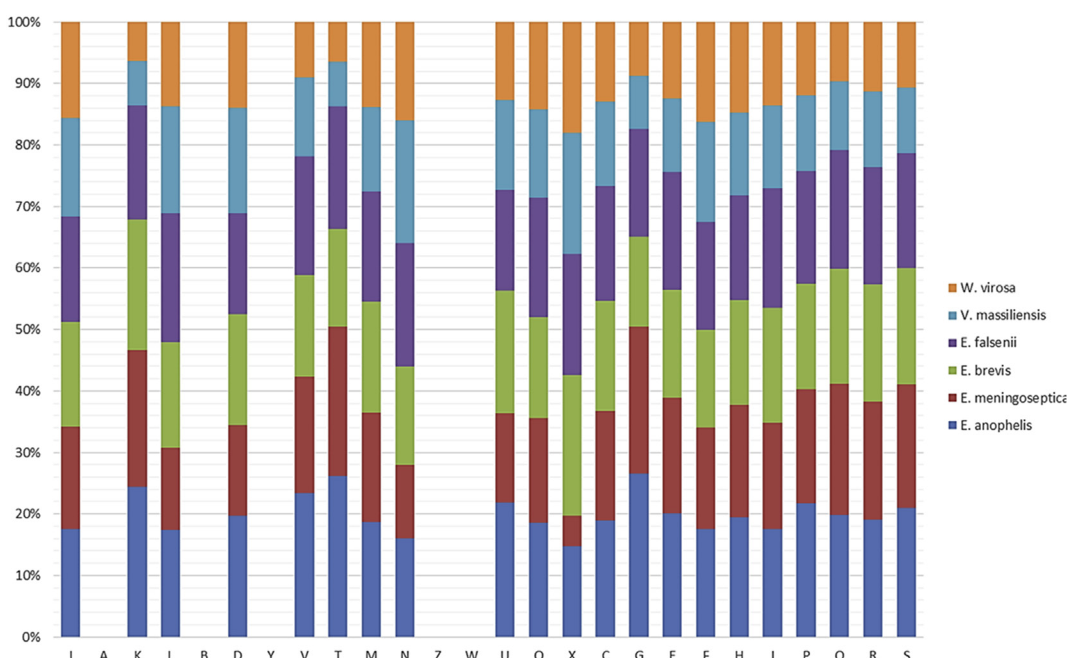


FIG. 7. Distribution of functional classes of predicted genes according to clusters of orthologous groups of proteins of *Vaginella massiliensis* strain Marseille P2517^T among other species.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **15**, 94–103
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

TABLE 6. Numbers of orthologous proteins shared between genomes (upper right) and AGIOS values obtained (lower left)

	Vm	Eb	Ef	Wv	Ea	Em
Vm	**2324**	1214	1200	1142	702	690
Eb	72.18%	**3567**	1426	1208	926	902
Ef	72.54%	86.26%	**3500**	1199	890	876
Wv	75.10%	72.37%	72.80%	**2118**	685	686
Ea	67.43%	68.71%	68.73%	67.63%	**4042**	941
Em	67.26%	68.37%	68.29%	67.38%	83.65%	**3459**

Numbers of proteins per genome are indicated in bold.
AGIOS, average genomic identity of orthologous gene sequences; Ea, *Elizabethkingia anophelis* B2D; Eb, *Empedobacter brevis* ATCC 43319; Ef, *Empedobacter falsenii* 282; Em, *Elizabethkingia meningoseptica* ATCC13253; Vm, *Vaginella massiliensis* Marseille P2517^T; Wv, *Weeksella virosa* DSM16922.

Genomic comparison

Comparison of the genome of our strain Marseille P2517^T with those of the closest species revealed that the draft genome sequence of strain Marseille P2517^T (2.34 Mb) was smaller than those of *Empedobacter falsenii* (3.71 Mb), *Empedobacter brevis* (3.79 Mb), *Elizabethkingia meningoseptica* (3.84 Mb) and *Elizabethkingia anophelis* (4.02 Mb) but larger than those of *Weeksella virosa* (2.27 Mb). The G+C content of strain Marseille P2517^T was larger than those of all the compared genomes: *E. meningoseptica* (36.4%), *W. virosa* (35.9%), *E. anophelis* (35.6%), *E. brevis* (32.8%) and *E. falsenii* (32.1%).

The gene content of strain Marseille P2517^T (2395) was smaller than those of *E. falsenii, E. brevis, E. anopheles* and *E. meningoseptica* (3610, 3633, 4108 and 3500 respectively) but larger than those of *W. virosa* (2192). Nevertheless, the distribution of genes into COGs categories was similar among all compared genomes (Fig. 7). Otherwise, the AGIOS analysis showed that strain Marseille P2517^T shares between 1214 and 690 orthologous genes with its closely related species: 1214, 1200, 1142, 702 and 690 with *E. brevis, E. falsenii, W. virosa, E. anopheles* and *E. meningoseptica* respectively (Table 6). Analysis of the average percentage of nucleotide sequence identity between strain Marseille P2517^T and other species ranged from 67.26% with *E. meningoseptica* and 75.10% with *W. virosa* (Table 6). Similar results were also observed for analysis of digital DNA-DNA hybridization (Table 7).

Conclusion

Phenotypic, phylogenetic analyses and genomic results enable us to propose that strain Marseille P2517^T may be the representative of a novel genus, *Vaginella,* with *Vaginella massiliensis* as the type strain. It was isolated from the normal vaginal flora of a 22-year-old Frenchwoman.

Taxonomic and nomenclatural proposals

Description of Vaginella *gen. nov. Vaginella* (va.gi.nel′la, L. fem. n. *vagina,* 'vagina,' part of the female genital tract; L. dim. suff. *-ella;* N.L. dim. fem. n. *Vaginella,* 'small vagina,' referring to the source of the isolation of the type strain).

The organism is an obligate aerobic, Gram-negative and rod-shaped bacilli. It is nonmotile and non–spore forming. It has negative catalase activity, and nitrate not reduced, with no urease production and positive oxidase activity. Habitat is human vagina flora. The type species is *Vaginella massiliensis* strain P2517^T.

Description of Vaginella massiliensis *strain Marseille P2517^T gen. nov., sp. nov. Vaginella massiliensis* (mas.si.li.en′sis, L. gen. adj. *massiliensis,* from Massilia, the Latin name of Marseille, France, where the organism was first grown, identified, and characterized).

The organism is obligate aerobic, nonmotile, nonsporulating and mesophilic, with optimal growth at 37°C. *Vaginella* cells are Gram negative and rod shaped, with a mean diameter of 0.61 μm and a length of 1.35 μm; it is oxidase positive and catalase negative; the major fatty acid is 13-methyl-tetradecanoic acid (60%). On Columbia agar, colonies are yellow, opaque, circular, smooth and approximately 1.85 mm in diameter. Nitrate reduction, urease and indole formation are negative. They are asaccharolytic; acid is produced only from glucose. Gelatin is hydrolyzed. Cells are susceptible to ceftriaxone, ciprofloxacin, clindamycin, colistin, doxycycline, erythromycin, fosfomycin, oxacillin, penicillin, rifampicin, teicoplasmin, trimethoprim/sulfamethoxazole, vancomycin and imipenem but are resistant to gentamicin and metronidazole.

TABLE 7. Digital DNA-DNA hybridization values obtained by comparison of all studied genomes

	Vm	Eb	Ef	Wv	Ea	Em
Vm	**100%**	21.10 ± 2.3%	19.10 ± 2.3%	21.40 ± 2.35%	26.10 ± 2.4%	25.50 ± 2.4%
Eb		**100%**	28.40 ± 2.4%	18.90 ± 2.3%	21.30 ± 2.35%	16.80 ± 2.2%
Ef			**100%**	19.90 ± 2.3%	24.90 ± 2.4%	19.30 ± 2.3%
Wv				**100%**	24.80 ± 2.4%	22.60 ± 2.35%
Ea					**100%**	23.70 ± 2.35%
Em						**100%**

Ea, *Elizabethkingia anophelis* B2D; Eb, *Empedobacter brevis* ATCC 43319; Ef, *Empedobacter falsenii* 282; Em, *Elizabethkingia meningoseptica* ATCC13253; Vm, *Vaginella massiliensis* Marseille P2517^T; Wv, *Weeksella virosa* DSM16922.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **15**, 94–103
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

The genome of *Vaginella massiliensis* is 2 434 475 bp long and exhibits 38.16% G+C content. Its 16S rRNA gene sequence and that of draft genome are both deposited in EMBL/EBI under accession numbers LT223570 and FLMR00000000, respectively. The type strain Marseille P2517^T (= DSM 102346^T = CSUR P2517) was isolated from the vaginal swab of a healthy Frenchwoman.

Acknowledgements

The authors thank the Xegen Company (http://www.xegen.fr/) for automating the genomic annotation process. This study was funded by the Fondation Méditerranée Infection. We also thank K. Griffiths for English-language review.

Conflict of Interest

None declared.

References

[1] Onderdonk AB, Delaney ML, Fichorova RN. The human microbiome during bacterial vaginosis. Clin Microbiol Rev 2016;29:223–38.
[2] Shipitsyna E, Roos A, Datcu R, Hallén A, Fredlund H, Jensen JS, et al. Composition of the vaginal microbiota in women of reproductive age—sensitive and specific molecular diagnosis of bacterial vaginosis is possible? PLoS One 2013;8:e60670.
[3] Svare J, Schmidt H, Hansen B, Lose G. Bacterial vaginosis in a cohort of Danish pregnant women: prevalence and relationship with preterm delivery, low birthweight and perinatal infections. BJOG 2006;113:1419–25.
[4] Bradshaw CS, Tabrizi SN, Fairley CK, Morton AN, Rudland E, Garland SM. The association of *Atopobium vaginae* and *Gardnerella vaginalis* with bacterial vaginosis and recurrence after oral metronidazole therapy. J Infect Dis 2006;194:828–36.
[5] Bretelle F, Rozenberg P, Pascal A, Favre R, Bohec C, Loundou A, et al. High *Atopobium vaginae* and *Gardnerella vaginalis* vaginal loads are associated with preterm birth. Clin Infect Dis 2015;60:860–7.
[6] Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–93.
[7] Bernardet JF, Nakagawa Y, Holmes B. Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended description of the family. Int J Syst Evol Microbiol 2002;52:1049–70.
[8] Bernardet JF, Segers P, Vancanneyt M, Berthe F, Kersters K, Vandamme P. Cutting a Gordian knot: emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. nov. (basonym, *Cytophaga aquatilis* Strohl and Tait 1978). Int J Syst Evol Microbiol 1996;46:128–48.
[9] Fournier PE, Lagier JC, Dubourg G, Raoult D. From culturomics to taxonomogenomics: a need to change the taxonomy of prokaryotes in clinical microbiology. Anaerobe 2015;36:73–8.
[10] Fournier PE, Drancourt M. *New Microbes New Infections* promotes modern prokaryotic taxonomy: a new section 'Taxonogenomics: new genomes of microorganisms in humans'. New Microb New Infect 2015;7:48–9.
[11] Menard JP, Fenollar F, Henry M, Bretelle F, Raoult D. Molecular Quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. Clin Infect Dis 2008;20:33–43.
[12] Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 2009;49:543–51.
[13] 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.
[14] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152.
[15] Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 2010;11:119.
[16] Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res 2016;44:D67–72.
[17] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–8.
[18] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997;25:955–64.
[19] Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics 2012;28:464–9.
[20] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 2009;25:119–20.
[21] Ramasamy D, Mishra AK, Lagier JC, 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.
[22] Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 2011;12:124.
[23] Gouret P, Paganini J, Dainat J, Louati D, Darbo E, Pontarotti P, et al. Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: the multi-agent software system DAGOBAH. In: Pontarotti P, editor. Evolutionary biology—concepts, biodiversity, macroevolution and genome evolution. Berlin: Springer; 2011. p. 71–87.
[24] Gouret P, Vitiello V, Balandraud N, Gilles A, Pontarotti P, Danchin EG. FIGENIX: intelligent automation of genomic annotation: expertise integration in a new software platform. BMC Bioinformatics 2005;6:198.
[25] Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence–based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 2013;14:60.
[26] Zhang RG, Tan X, Liang Y, Meng TY, Liang HZ, Lv J. Description of *Chishuiella changwenlii* gen. nov., sp. nov., isolated from freshwater, and transfer of *Wautersiella falsenii* to the genus *Empedobacter* as *Empedobacter falsenii* comb. nov. Int J Syst Evol Microbiol 2014;64:2723–8.
[27] Kampfer P, Avesani V, Janssenes M, Charlier J, De Baere T, Vaneechouette M. Description of *Wautersiella falsenii* gen. nov., sp. nov., to accommodate clinical isolates phenotypically resembling members of the genera *Chryseobacterium* and *Empedobacter*. Int J Syst Evol Microbiol 2006;56:2323–9.
[28] Zhang RG, Tan X, Zhao X-M, Deng J, Lv J. *Moheibacter sediminis* gen. nov., sp. nov., a member of the family *Flavobacteriaceae* isolated from sediment, and emended descriptions of *Empedobacter brevis, Wautersiella falsenii* and *Weeksella virosa*. Int J Syst Evol Microbiol 2014;64:1481–7.
[29] Holmes B, Snell JJS, Lapage SP. Revised description, from clinical strains, of *Flavobacterium breve* (Lustig) Bergey et al. 1923 and proposal of the neotype strain. Int J Syst Evol Microbiol 1978;28:201–8.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **15**, 94–103
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Article 13:

Characterization of a New *Ezakiella* Isolated from the Human Vagina: Genome Sequence and Description of *Ezakiella massiliensis* sp. nov.

Khoudia Diop, Claudia Andrieu, Caroline Michelle, Nicholas Armstrong, Fadi Bittar, Florence Bretelle, Pierre-Edouard Fournier, Didier Raoult, Florence Fenollar

Published in Current Microbiology

Current Microbiology
https://doi.org/10.1007/s00284-017-1402-z

Characterization of a New *Ezakiella* Isolated from the Human Vagina: Genome Sequence and Description of *Ezakiella massiliensis* sp. nov.

Khoudia Diop[1] · Claudia Andrieu[1] · Caroline Michelle[1] · Nicholas Armstrong[1] · Fadi Bittar[1] · Florence Bretelle[1,2] · Pierre-Edouard Fournier[1] · Didier Raoult[1,3] · Florence Fenollar[1]

Received: 23 March 2017 / Accepted: 16 November 2017
© Springer Science+Business Media, LLC, part of Springer Nature 2017

Abstract

The study of the vaginal microbiota using the "culturomics concept" allowed us to isolate, from the vaginal swab of an asymptomatic 20-year-old woman who had sexual relations with another woman with bacterial vaginosis, an unknown Gram-positive anaerobic coccus-shaped bacterium that was designated strain Marseille-P2951^T and characterized using taxono-genomics. Strain Marseille-P2951^T is non-motile and non-spore forming and exhibits catalase and oxidase activities. Its 16S rRNA gene-based identification showed 98.5% identity with *Ezakiella peruensis*, the phylogenetically closest species. The major fatty acids are C18:1n9 (58%) and C16:0 (22%). With a 1,741,785 bp length, the G+C content of the genome is 36.69%. Of a total of 1657 genes, 1606 are protein-coding genes and 51 RNAs. Also, 1123 genes are assigned a putative function and 127 are ORFans. Phenotypic, phylogenetic, and genomics analyses revealed that strain Marseille-P2951^T (=CSUR P2951 =DSM 103122) is distinct and represents a new species of the genus *Ezakiella*, for which the name *Ezakiella massiliensis* sp. nov. is proposed.

Introduction

Gram-positive anaerobic cocci are mainly represented by the *Peptostreptococci*. These bacteria are part of the commensal flora of humans and animals, and are also often associated with a variety of human infections [21]. They represent approximately 24–31% of all anaerobic bacteria isolated in clinical samples and are the most frequent species that belong to the genus *Peptostreptococcus* [20]. A taxonomic revision of this group has occasioned the division of *Peptostreptococcus* into 5 genera: *Anaerococcus, Finegoldia, Gallicola, Parvimonas*, and *Peptoniphilus* [8]. In 2015, a new genus of Gram-positive anaerobic cocci called *Ezakiella* was identified from a human fecal sample in a coastal traditional community in Peru [24, 25]. This genus has only one species: *Ezakiella peruensis* [16]. In a study characterizing the vaginal flora of women with and without bacterial vaginosis using the "culturomics concept," based on the multiplication of culture conditions (media, temperature, pH, and atmosphere) combined with rapid bacterial identification by Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) [13, 14], we identified an unknown strain of *Ezakiella* designated Marseille-P2951^T which did not correspond to the other species of this genus.

We report the phenotypic, phylogenetic, and genomic characteristics of *Ezakiella massiliensis* sp. nov., strain Marseille-P2951^T (=CSUR [Collection de souches de l'Unité des Rickettsies] P2951 = DSM [Deutsche Sammlung von Mikroorganismen] 103122), isolated from the vaginal sample of a healthy woman who had sexual relations with another woman with bacterial vaginosis.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00284-017-1402-z) contains supplementary material, which is available to authorized users.

✉ Florence Fenollar
florence.fenollar@univ-amu.fr

[1] Aix-Marseille Univ, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, AMU UM 63, CNRS UMR7278, IRD 198, INSERM U1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de Médecine, 19-21 Boulevard Jean Moulin, 13005 Marseille Cedex 05, France

[2] Department of Gynecology and Obstetrics, Gynépole, Marseille, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, AMU, Aix-Marseille Université, Marseille Cedex 05, France

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

Published online: 29 November 2017

Materials and Methods

Ethics and Sample Collection

In October 2015, a vaginal sample from an asymptomatic 20-year-old woman who had sexual relations with another woman with bacterial vaginosis was collected at the Hospital "La Timone" in Marseille (France). The diagnosis was conducted as previously reported [18]. At the time of sample collection, she was not being treated with antibiotics. She gave her written consent. This study was authorized by the local IFR48 ethics committee (Marseille, France) under agreement number 09-022. The sample was collected and transported using a Sigma Transwab (Medical Wire, Corsham, United Kingdom).

Strain Identification by MALDI-TOF MS and 16S rRNA Gene Sequencing

After sampling, the vaginal specimen was cultured using 1 of the 18 culturomics conditions previously described [14]. Isolated colonies were identified using MALDI-TOF, as previously described [4, 5, 28]. For unidentified colonies, the 16S rRNA gene was sequenced and the obtained sequence was matched against the NCBI database using the BLAST algorithm [7]. As suggested, if the 16S rRNA gene sequence similarity value was <98.7%, the strain was defined as a new species [12, 29].

Phylogenetic Tree

All species from the same order of the new species were retrieved and 16S sequences were downloaded from NCBI. Sequences were aligned using CLUSTALW, with default parameters and phylogenetic inferences obtained using the neighbor-joining method with 500 bootstrap replicates, using MEGA6 software.

Morphologic and Biochemical Characterization

To observe cell morphology, individual cells were visualized using a Tecnai G20 electron microscope [4]. Oxidase and catalase tests, Gram stain, motility, and sporulation were performed as previously conducted [23]. Biochemical tests were performed using API ZYM, API 20A, and API 50CH strips (bioMérieux, Marcy l'Etoile, France).

Growth Conditions

To determine optimal growth, the strain Marseille-P2951^{T} was cultivated on Columbia agar with 5% sheep blood (bioMérieux) and incubated at different temperatures in aerobic, anaerobic, and microaerophilic atmospheres [4, 5]. The salinity and pH conditions were also tested [4, 5].

Cellular Fatty Acid Analysis

Cellular fatty acid methyl ester (FAME) analysis was performed by Gas Chromatography/Mass Spectrometry (GC/MS). Strain Marseille-P2951^{T} was grown on Columbia agar (bioMérieux). Then, two samples were prepared with approximately 9 mg of bacterial biomass per tube harvested from several culture plates. FAME was prepared as described by Sasser [27]. GC/MS analyses were carried out as described before [3]. Briefly, fatty acid methyl esters were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500—SQ 8S, Perkin Elmer, Courtaboeuf, France). A spectral database search was performed using MS Search 2.0 operated with the Standard Reference Database 1A (NIST, Gaithersburg, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK).

Antibiotic Susceptibility Testing

Amoxicillin, benzylpenicillin, ceftriaxone, imipenem, metronidazole, and vancomycin were used to test antibiotic susceptibility of strain Marseille-P2951^{T}. The minimal inhibitory concentrations (MICs) were then determined using E-test gradient strips (bioMérieux) [1, 17].

DNA Extraction, Genome Sequencing, and Assembly

After a pre-treatment of 2 h with lysozyme incubation at 37 °C, genomic DNA (gDNA) of strain Marseille-P2951^{T} was extracted on the EZ1 biorobot (Qiagen) with the EZ1 DNA tissue kit. The elution volume was 50 µL. gDNA was quantified by a Qubit assay with the high sensitivity kit (Life technologies, Carlsbad, CA, USA) at 74.2 ng/µL.

Then gDNA was sequenced on MiSeq Technology (Illumina Inc, San Diego, CA, USA) with the mate pair strategy. The gDNA was barcoded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina) [4, 5]. Then, the genome sequencing and assembly were performed as previously described [4]. The reads were trimmed and assembled using the CLC genomics Workbench v4.7.2 software (CLC bio, Aarhus, Denmark).

Genome Annotation and Analysis

The Multi-Agent software system DAGOBAH [10] was used to perform annotation and comparison processes, which include Figenix [11] libraries that provide pipeline analysis. The genome of strain Marseille-P2951^{T} was compared to those of closest species. In order to evaluate the genomic similarity among all compared genomes, two parameters

were determined: dDDH, which exhibits a high correlation with DNA–DNA Hybridization (DDH) [19], and AGIOS (Average Genomic Identity of Orthologous gene Sequences) [26], which was designed to be independent from DDH.

Results

Strain Identification and Phylogenetic Analysis

The MALDI-TOF MS identification of our isolate failed (score <1.7). The 16S rRNA gene sequence of the strain showed 98.5% identity with *Ezakiella peruensis*, the phylogenetically closest species with a validly published name (Fig. 1). As this value was <98.7% recommended to delimited a species [12, 29], the strain Marseille-P2951^T was classified as a new species, named *Ezakiella massiliensis* sp. nov. The reference spectrum (Fig. 2a) was incremented in our database (http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database) and then compared to other closest species; their differences are shown in the gel view obtained (Fig. 2b).

Phenotypic and Biochemical Characteristics

Strain Marseille-P2951^T is a Gram-positive anaerobic coccus, 0.7 µm in diameter (Supplementary Figure S1A and B), non-motile, and non-spore forming. After 3 days of growth on blood agar (BD diagnostic), colonies were clear and gray, circular, and convex, with a diameter of approximately 0.8 mm. This strict anaerobe grows at temperatures between 28 and 37 °C, but optimal growth was observed at 37 °C after 48 h of incubation. Bacterial growth requires a NaCl concentration <5 g/L and a pH ranging from 6.5 to 8.5. Strain Marseille-P2951^T exhibited catalase and oxidase activities, whereas nitrate and nitrite were not reduced. The results of the API ZYM, 20A, and 50CH are summarized in Table 1 and were compared to those of the closest species [2, 8, 22, 25, 30]. The fatty acid profiles of strain Marseille-P2951^T and the closest species are recapped in Table 2. Strain Marseille-P2951^T was susceptible to amoxicillin (MIC 0.016 µg/mL), benzylpenicillin (MIC 0.003 µg/mL), ceftriaxone (MIC 0.016 µg/mL), imipenem (MIC 0.032 µg/mL), metronidazole (MIC 0.047 µg/mL), and vancomycin (MIC 0.19 µg/mL).

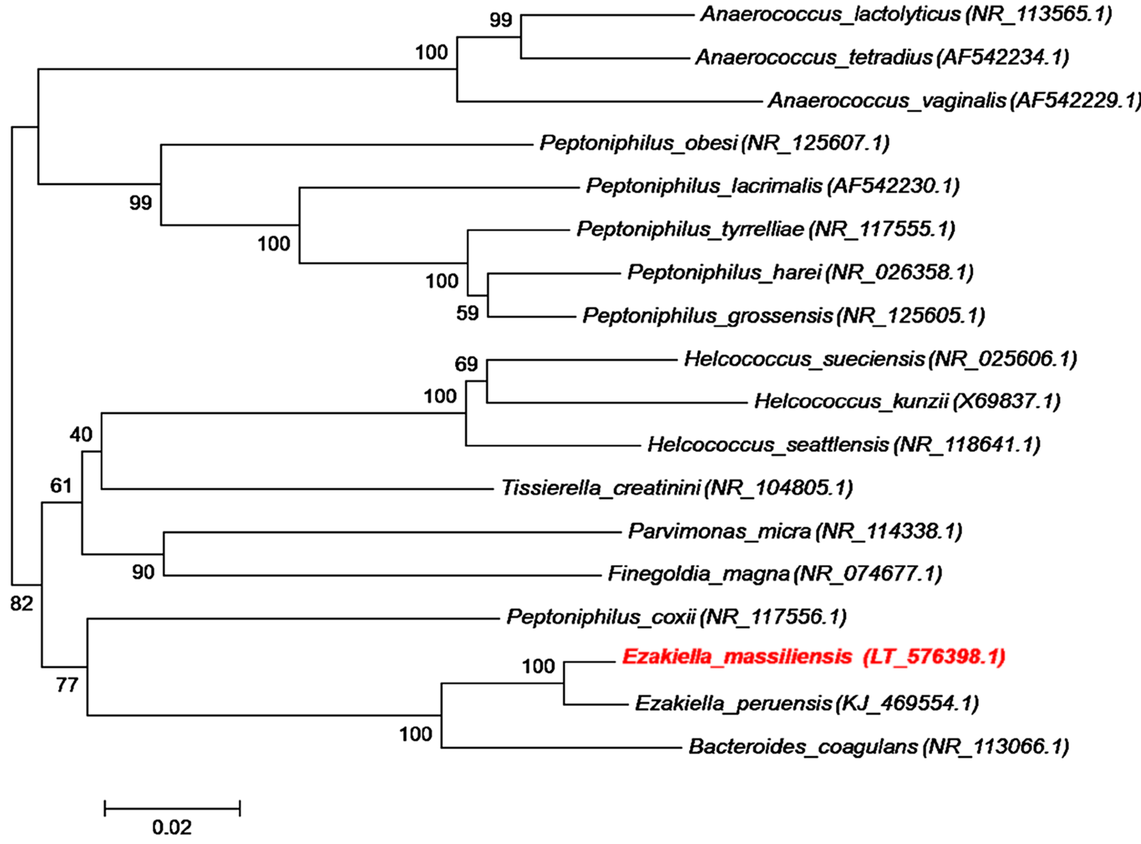
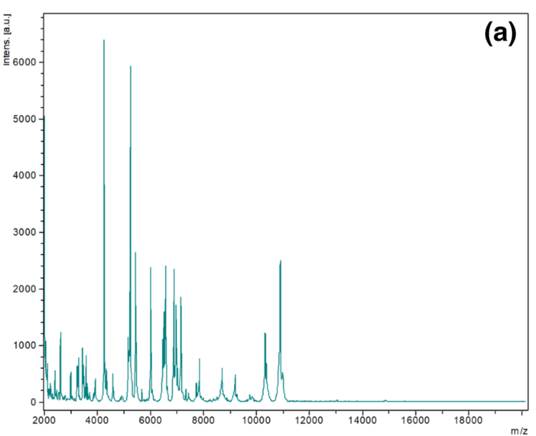


Fig. 1 Phylogenetic tree highlighting the position of *Ezakiella massiliensis* strain Marseille-P2951^T relative to other close strains. GenBank accession numbers for each 16S rRNA are noted in parenthesis. Sequences were aligned using Muscle v3.8.31 with default parameters and phylogenetic inferences were obtained using the neighbor-joining method with 500 bootstrap replicates, using MEGA6 software. The scale bar represents a 2% nucleotide sequence divergence



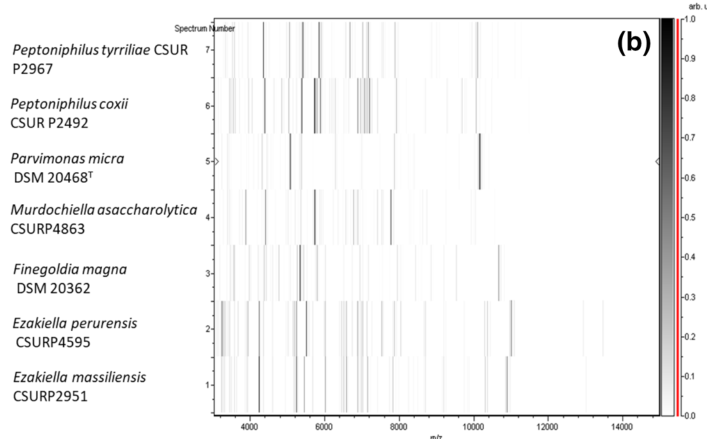


Fig. 2 **a** Reference mass spectrum from *Ezakiella massiliensis* strain Marseille-P2951^T spectra. **b** Gel view comparing the *Ezakiella massiliensis* strain Marseille-P2951^T to other species within the family. The gel view displays the raw spectra of loaded spectrum files arranged in a pseudo-gel-like look. The *x*-axis records the *m/z* value. The left *y*-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a gray scale scheme code. The right *y*-axis indicates the relation between the color of a peak and its intensity, in arbitrary units. Displayed species are indicated on the left

Genome Properties

The draft genome of strain Marseille-P2951^T is 1,741,785 bp long with 36.69% of G+C content (Table 3; Fig. 3). It is composed of 1 scaffold (composed of 1 contig). Of the 1657 predicted genes, 1606 were protein-coding genes and 51 were RNAs (6 genes are rRNA, 45 genes are tRNA). A total of 1123 genes (69.93%) were assigned a putative function and 127 genes were identified as ORFans (7.91%). Tables 3 and 4 summarize the genome statistics and the distribution of genes into COG (Clusters of Orthologous Groups) functional categories, respectively.

Genomic Comparison

The comparison of the genome of strain Marseille-P2951^T with those of *Ezakiella peruensis* and other Gram-positive anaerobic cocci is detailed in supplementary Table S1. The distribution of genes into COG categories was similar among all compared genomes (Supplementary Figure S2). However, there were fewer genes of strain Marseille-P2951^T present in the COG categories C (Energy production and conversion) and G (Carbohydrate transport and metabolism) than other compared species. In addition, the AGIOS analysis revealed that strain Marseille-P2951^T shares 541 orthologous genes with *Helcococcus sueciensis* and 874 with *Ezakiella peruensis* (Supplementary Table S2). The analysis of the average percentage of nucleotide sequence identity ranged from 60.10 to 94.20% with *Peptoniphilus coxii* and *Ezakiella peruensis*, respectively (Supplementary Table S2). Moreover, the dDDH of strain Marseille-P2951^T and its closest species varied between 23.90 and 62 mol% with *Helcococcus kunzii* and *Anaerococcus tetradius*, respectively (Supplementary Table S3).

Discussion

Only 20% of vaginal bacteria were identified by culture [15]. The understanding of vaginal flora was enhanced by the use of molecular techniques and culturomics. Molecular methods allowed the detection of uncultured and fastidious bacteria in the vagina, such as bacterial vaginosis-associated bacteria type 1 (BVAB1), BVAB2, and BVAB3 [9]. Also, culturomics concepts enabled the culture of new bacterial species, such as *Vaginella massiliensis* [4], *Olegusella massiliensis* [5], *Murdochiella vaginalis* [6], and *Ezakiella massiliensis*, reported in this paper.

Ezakiella massiliensis strain Marseille-P2951^T is a new member of the genus *Ezakiella*. The similarity of its 16S rRNA gene sequence less than the threshold delineated to distinguish a new species, the difference in the G+C content, the AGIOS values, and the GGDC values to 70% confirmed that strain Marseille-P2951^T is a distinct species than *Ezakiella perurensis*. The genus *Ezakiella* was created in 2015 [25] and contains Gram-positive anaerobic cocci that were detected in diverse human clinical samples. Bacteria of this group were also asaccharolytic. All these characteristics were observed in strain Marseille-P2951^T.

Table 1 Differential characteristics of *Ezakiella massiliensis* and phylogenetically related species: *Ezakiella peruensis* strain M6.X2[T], *Finegoldia magna* strain CCUG 17636[T], *Peptoniphilus asaccharolyticus* strain ATCC 14963, *Parvimonas micra* CCUG 46357[T], *Helcococcus kunzii* NCFB 2900[T], and *Anaerococcus prevotii* ATCC 9321[T] [2, 8, 22, 25, 30]

Properties	*Ezakiella massiliensis*	*Ezakiella peruensis*	*Finegoldia magna*	*Peptoniphilus asaccharolyticus*	*Parvimonas micra*	*Helcococcus kunzii*	*Anaerococcus prevotii*
Cell diameter (µm)	0.7	<1	0.8–1.6	0.5–0.9	0.3–0.7	na	na
Oxygen requirement	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Facultatively anaerobic	Anaerobic
Gram stain	+	+	+	+	+	+	+
Spore formation	–	–	–	–	–	–	na
Motility	–	–	–	–	–	–	–
DNA G+C content (mol%)	36.69	38,4	na	31	28.65	30	33
Production of							
Alkaline phosphatase	+	+	variable	–	+	–	–
Indole	+	+	–	variable	–	–	–
Catalase	+	–	variable	+	variable	–	+
Nitrate reductase	–	–	–	–	–	–	–
Urease	–	–	–	–	–	–	+/–
β-galactosidase	–	–	–	–	–	–	–
N-acetyl-glucosamine	–	–	–	na	–	+	na
Acid from							
Mannose	–	–	–	–	–	–	+
Glucose	–	–	–	–	–	–	+/–
Lactose	–	–	–	–	–	+	–
Raffinose	–	–	–	–	–	–	+
Habitat	Vaginal discharge	Human feces	Human specimen	Human specimen	Human specimen	Human specimen	Human specimen

na no available data

+: positive reaction

–: negative reaction

Table 2 Cellular fatty acid profiles (%) of *Ezakiella massiliensis* strain Marseille-P2951[T] (data from this study) compared with closest species [25]

Fatty acid	*Ezakiella massiliensis*	*Ezakiella peruensis*	*Finegoldia magna*	*Peptoniphilus asaccharolyticus*	*Parvimonas micra*	*Helcococcus kunzii*	*Anaerococcus prevotii*
C10:0	–	7.3	9.4	–	–	–	–
C12:0	–		4.8	12.0	–	–	–
C14:0	6.9	3.0	5.9	5.4	1.6	2.5	2.0
Iso-$C_{15:0}$	–	–	–	2.6	–	–	–
C16:0	**21.8**	**18.3**	**17.6**	**14.4**	**13.4**	**30.0**	**17.1**
C17:0		5.2		–	–	–	–
Anteiso-$C_{17:0}$	1.3	–	4.5	1.6	–	–	1.7
Iso-C17:1/C16:0 DMA	–	–	**18.2**	–	–	–	–
C18:0	5.0	–	–	9.4	6.8	**16.0**	**11.5**
C18:1 ω7c	–	4.2	–	–	–	–	–
$C_{18:1}$ ω9c	**58.1**	**39.8**	3.6	**20.2**	**15.5**	**19.3**	**19.3**
C18:2ω6,9c/C18:0 ANTE	4.9	**13.2**	5.6	**22.0**	**58.3**	**29.4**	**20.7**
$C_{18:1}$ ω9c DMA	–	–	**11.1**	6.6	–	–	–

Bold represents the majority fatty acid for this species; values <1% are not shown

Conclusion

Phenotypic, phylogenetic, and genomic results confirmed that strain Marseille-P2951^T is distinct from its phenotypically closest species and can represent a new species, within the genus *Ezakiella*, for which we proposed the name *Ezakiella massiliensis*. The type strain Marseille-P2951^T was isolated from the vaginal sample of a healthy woman who had sexual relations with another woman with bacterial vaginosis.

Table 3 Nucleotide content and gene count levels of the genome

Attribute	Genome (total)	
	Value	% of total[a]
Size (bp)	1,741,785	100
G + C content (bp)	639,071	36.69
Coding region (bp)	1,581,924	90.82
Total genes	1657	100
rRNA genes	6	0.36
Protein-coding genes	1606	100
Genes with function prediction	1123	69.93
Genes assigned to COGs	1029	64.07
Genes with peptide signals	210	13.08
Genes with transmembrane helices	408	25.40

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

Taxonomic and Nomenclature Proposal

Description of *Ezakiella massiliensis* sp. nov.

Ezakiella massiliensis (mas.si.li.en'sis, L. fem. adj. massiliensis, from *Massilia*, the Latin name of Marseille where the type strain was first isolated).

Ezakiella massiliensis is non-motile and non-sporulating. It is mesophilic, with optimal growth at 37 °C after 48 h. Colonies are clear and gray, circular, and convex, with 0.8 mm of diameter on blood agar. Obligate anaerobe, the cells are gram-positive and coccus-shaped, with a diameter of 0.7 µm, positive for indole, catalase, and oxidase, negative for nitrate reductase. Gelatin and starch were not hydrolyzed. Lipases, valine arylamidase, trypsin, α-chymotrypsin, phosphatase acid, α- and β-galactosidase, β-glucuronidase, α- and β-glucosidase, *N*-acetyl-β-glucosaminidase were negative, although alkaline phosphatase, leucine arylamidase, and naphthol-AS-BI-phosphohydrolase were positive.

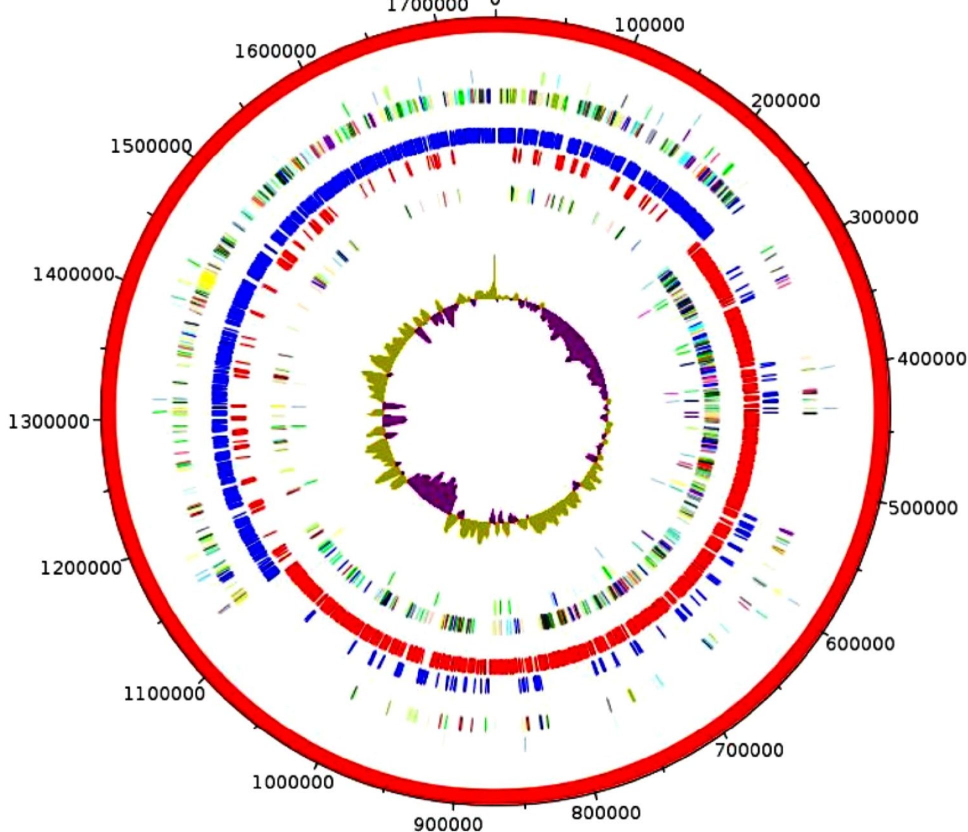


Fig. 3 Graphical circular map of the genome. From outside to the center: Contigs (red/gray), COG category of genes on the forward strand (three circles), genes on forward strand (blue circle), genes on the reverse strand (red circle), COG category on the reverse strand (three circles), G+C content. (Color figure online)

Table 4 Number of genes associated with the 25 general COG functional categories

Code	Value	% of total	Description
[J]	171	10.647572	Translation
[A]	0	0	RNA processing and modification
[K]	61	3.7982564	Transcription
[L]	59	3.6737237	Replication, recombination, and repair
[B]	0	0	Chromatin structure and dynamics
[D]	21	1.3075966	Cell cycle control, mitosis and meiosis
[Y]	0	0	Nuclear structure
[V]	59	3.6737237	Defense mechanisms
[T]	36	2.241594	Signal transduction mechanisms
[M]	47	2.9265256	Cell wall/membrane biogenesis
[N]	5	0.3113325	Cell motility
[Z]	0	0	Cytoskeleton
[W]	2	0.124533	Extracellular structures
[U]	19	1.1830635	Intracellular trafficking and secretion
[O]	51	3.1755917	Posttranslational modification, protein turnover, chaperones
[X]	21	1.3075966	Mobilome: prophages, transposons
[C]	49	3.0510585	Energy production and conversion
[G]	36	2.241594	Carbohydrate transport and metabolism
[E]	100	6.2266498	Amino acid transport and metabolism
[F]	50	3.1133249	Nucleotide transport and metabolism
[H]	65	4.0473228	Coenzyme transport and metabolism
[I]	31	1.9302616	Lipid transport and metabolism
[P]	82	5.105853	Inorganic ion transport and metabolism
[Q]	13	0.8094645	Secondary metabolites' biosynthesis, transport, and catabolism
[R]	77	4.7945204	General function prediction only
[S]	61	3.7982564	Function unknown
_	577	35.92777	Not in COGs

It is asaccharolytic; acid is not produced from sugars. The major fatty acids are C18:1n9 (58%) and C16:0 (22%). Bacterial cells were susceptible to amoxicillin, ceftriaxone, benzylpenicillin, imipenem, metronidazole, and vancomycin. The genome is 1,741,785 bp long and the DNA G+C content is 36.69 mol%. The 16S rRNA and genome sequences are both deposited in EMBL/EBI under accession numbers LT576398 and LT635475, respectively. The respective DPD (Digital Protologue Database) TaxonNumber is TA00324. The type strain, strain Marseille-P2951^T (=CSUR P2951 =DSM 103122), was isolated from the vaginal sample of a healthy woman.

Acknowledgements The authors thank the Xegen Company (http://www.xegen.fr) for automating the genomic annotation process. We also thank TradOnline for reviewing the English.

Funding This study was supported by Méditerranée Infection and the National Research Agency under the program "Investissements d'avenir," reference ANR-10-IAHU-03.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Citron DM, Ostovari MI, Karlsson A, Goldstein EJ (1991) Evaluation of the E test for susceptibility testing of anaerobic bacteria. J Clin Microbiol 29:2197–2203
2. Collins MD, Facklam RR, Rodrigues UM, Ruoff KL (1993) Phylogenetic analysis of some *Aerococcus*-like organisms from clinical sources: description of *Helcococcus kunzii* gen. nov., sp. nov. Int J Syst Evol Microbiol 43:425–429
3. Dione N, Sankar SA, Lagier JC, Khelaifia S, Michele C, Armstrong N, Richez M, Abrahão J, Raoult D, Fournier PE (2016) Genome sequence and description of *Anaerosalibacter massiliensis* sp. nov. New Microbes New Infect 10:66–76
4. Diop K, Bretelle F, Michelle C, Richez M, Rathored J, Raoult D, Fournier PE, Fenollar F (2017) Taxonogenomics and description of *Vaginella massiliensis* gen. nov., sp. nov., strain Marseille P2517^T, a new bacterial genus isolated from the human vagina. New Microbes New Infect 15:94–103
5. Diop K, Diop A, Bretelle F, Cadoret F, Michelle C, Richez M, Cocallemen JF, Raoult D, Fournier PE, Fenollar F (2017)

Olegusella massiliensis gen. nov., sp. nov., strain KHD7T, a new bacterial genus isolated from the female genital tract of a patient with bacterial vaginosis. Anaerobe 44:87–95
6. Diop K, Raoult D, Bretelle F, Fenollar F (2016) "*Murdochiella vaginalis*" sp. nov., a new bacterial species cultivated from the vaginal flora of a woman with bacterial vaginosis. Hum Microbiome J 2:15–16
7. Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, Raoult D (2000) 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 38:3623–3630
8. Ezaki T, Kawamura Y, Li N, Li ZY, Zhao L, Shu S (2001) Proposal of the genera *Anaerococcus* gen. nov., *Peptoniphilus* gen. nov. and *Gallicola* gen. nov. for members of the genus *Peptostreptococcus*. Int J Syst Evol Microbiol 51:1521–1528
9. Fredricks DN, Fiedler TL, Marrazzo JM (2005) Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med 353:1899–1911
10. 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: Pontarotti P (ed) Evolutionary biology—concepts, biodiversity, macroevolution and genome evolution. Springer, Berlin, pp 71–87
11. Gouret P, Vitiello V, Balandraud N, Gilles A, Pontarotti P, Danchin EG (2005) FIGENIX: intelligent automation of genomic annotation: expertise integration in a new software platform. BMC Bioinform 6:198
12. Kim M, Oh H-S, Park SC, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351
13. Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, Bittar F, Fournous G, Gimenez G, Maraninchi M, Trape JF, Koonin EV, La Scola B, Raoult D (2012) Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 18:1185–1193
14. Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, Raoult D (2015) The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 28:237–264
15. Lamont RF, Sobel JD, Akins RA, Hassan SS, Chaiworapongsa T, Kusanovic JP, Romero R (2011) The vaginal microbiome: new information about genital tract flora using molecular based techniques. BJOG 118:533–549
16. LPSN *Ezakiella*. In: LPSN. http://www.bacterio.net/ezakiella.html. Accessed 26 Jan 2017
17. Matuschek E, Brown DF, Kahlmeter G (2014) Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. Clin Microbiol Infect 20:O255–O266
18. Menard J, Fenollar F, Henry M, Bretelle F, Raoult D (2008) Molecular quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. Clin Infect Dis 47:33–43
19. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinform 14:60
20. Murdoch DA (1998) Gram-positive anaerobic cocci. Clin Microbiol Rev 11:81–120
21. Murdoch DA, Mitchelmore IJ, Tabaqchali S (1994) The clinical importance of gram-positive anaerobic cocci isolated at St Bartholomew's Hospital, London, in 1987. J Med Microbiol 41:36–44
22. Murdoch DA, Shah HN (1999) Reclassification of *Peptostreptococcus magnus* (Prevot 1933) Holdeman and Moore 1972 as *Finegoldia magna*. comb.nov. and *Peptostreptococcus micros* (Prevot 1933) Smith 1957 as *Micromonas micros* comb. nov.. Anaerobe 5:555–559
23. Murray PR, Baron EJ, Jorgensen JH, Landry ML, Pfaller MA (2007) Manual of clinical microbiology, 9th edn. ASM Press, Washington, D.C.
24. Oren A, Garrity GM (2015) List of new names and new combinations previously effectively, but not validly, published. Int J Syst Evol Microbiol 65:1105–1111
25. Patel NB, Tito RY, Obregón-Tito AJ, O'Neal L, Trujillo-Villaroel O, Marin-Reyes L, Troncoso-Corzo L, Guija-Poma E, Hamada M, Uchino Y, Lewis CM, Lawson PA (2015) *Ezakiella peruensis* gen. nov., sp. nov. isolated from human fecal sample from a coastal traditional community in Peru. Anaerobe 32:43–48
26. Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, Sentausa E, Raoult D, Fournier PE (2014) A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 64:384–391
27. Sasser M (2006) Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). MIDI, Technical Note
28. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P, Rolain JM, Raoult D (2009) Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 49:543–551
29. Stackebrandt E, Ebers J (2006) Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 33:152
30. Tindall BJ, Euzeby JP (2006) Proposal of *Parvimonas* gen. nov. and *Quatrionicoccus* gen. nov. as replacements for the illegitimate, prokaryotic, generic names *Micromonas* Murdoch and Shah 2000 and *Quadricoccus* Maszenan et al. 2002, respectively. Int J Syst Evol Microbiol 56:2711–2713

CHAPITRE IV :

Annexes & Travaux Collaboratifs

Avant-propos

La culturomique microbienne est une technique de culture à haut débit basée sur la multiplication des conditions de culture et l'utilisation de la spectrométrie de masse MALDI-TOF. Tous les isolats sont donc d'abord identifiés par spectrométrie de masse MALDI-TOF. Après trois tentatives, si un isolat reste toujours non identifié, la séquence du gène codant pour l'ARNr 16S est alors amplifiée puis séquencée. Après un BLAST, si la similarité avec la souche type de l'espèce la plus proche est inférieure à 98,7%, la bactérie peut être considérée comme une potentielle nouvelle espèce. Depuis sa naissance en 2009, l'approche de la culturomique microbienne a surtout été appliqué au microbiote digestif humain et a de ce fait permis l'isolement de 1.057 espèces bactériennes distinctes avec 531 espèces isolées pour la première fois dont près de 200 nouvelles espèces [80]. Vu le nombre de nouvelles espèces cultivées grâce à cette approche, il est donc quasi impossible de séquencer puis annoter tous leurs génomes et de publier leurs descriptions taxonogénomiques au même rythme que leur isolement [81]. Afin d'informer la communauté scientifique de l'isolement des nouvelles espèces et de mettre rapidement à leur disposition les informations relatives à ces espèces, un nouveau format de description rapide a été développé : le « new species announcement » [42].

Un « new species announcement » ne contient que les caractéristiques majeures (taxonomiques aussi bien phénotypiques) de la nouvelle espèce telles que l'origine de l'échantillon à partir duquel la souche a été isolée, son aspect morphologique (taille, forme, Gram, et aspect colonies), la présence d'activités de catalase et d'oxydase, le pourcentage de similarité de ARNr avec l'espèce valide la plus proche et sa position phylogénétique. De plus, il décrit aussi l'étymologie du nom de la nouvelle espèce. Dans cette dernière partie, nous présentons les « new species announcement » de 3 de nos nouveaux isolats : *Anaerococcus mediterraneensis* et *Collinsella vaginalis* cultivées à partir des échantillons de patientes atteintes de vaginose bactérienne et *Arcanobacterium urinimassiliense* comportant 2 souches,

une isolée des urines d'un bébé de 7 semaines présentant une gastro-entérite et une deuxième trouvée dans la flore vaginale d'une femme saine. Pour toutes ces espèces, les descriptions complètes selon le modèle taxonogénomique seront très prochainement disponibles.

Dans cette section, nous rapportons aussi la première séquence génomique de la souche d'*Ezakiella peruensis* M6.X2^{T}. Il s'agit d'un coque Gram-positif anaérobie isolé du microbiote intestinal d'un individu en bonne santé, représentant le seul et l'unique membre valide du genre *Ezakiella*. L'analyse de son génome révèle qu'il mesure 1.672.788 pb de long et qu'il contient 1.589 gènes codant pour des protéines, dont 26 gènes impliqués dans la résistance aux antibiotiques avec un codant pour celle à la vancomycine.

Article 14:

"*Anaerococcus mediterraneensis*" sp. nov., a new species isolated from human female genital tract.

Khoudia Diop, Florence Bretelle, Pierre-Edouard Fournier, Florence Fenollar

Published in New Microbes and New Infections

NEW SPECIES

'*Anaerococcus mediterraneensis*' sp. nov., a new species isolated from human female genital tract

K. Diop[1], **F. Bretelle**[1,2], **P.-E. Fournier**[1] **and F. Fenollar**[1]

1) *Institut hospitalo-universitaire Méditerranée-infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine and* 2) *Department of Gynecology and Obstetrics, Gynépole, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, AMU, Aix-Marseille Université, Marseille, France*

Abstract

We report the principal characteristics of '*Anaerococcus mediterraneensis*' strain Marseille P2765, a new member of the *Anaerococcus* genus. Strain Marseille P2765 was isolated in a vaginal sample of a 26-year-old patient with bacterial vaginosis.
© 2017 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

Keywords: *Anaerococcus* mediterraneensis, bacterial vaginosis, culturomics, human microbiota, vaginal flora
Original Submission: 19 October 2016; **Revised Submission:** 6 February 2017; **Accepted:** 27 February 2017
Article published online: 3 March 2017

Corresponding author: F. Fenollar, Aix-Marseille Université, Institut hospitalo-universitaire Méditerranée-infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine, 27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France
E-mail: florence.fenollar@univ-amu.fr

We are currently studying the human microbiota by culturomics in our laboratory in Marseille, France [1]. As part of this study, we isolated in the vaginal flora of a 26-year-old French woman with bacterial vaginosis [2] a bacterium which could not be identified using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) performed with a MicroFlex spectrometer (Bruker Daltonics, Leipzig, Germany) [3]. The agreement number of the National Ethics Committee of the IFR48 (Marseille, France) for this study is 09-022. The patient provided written consent.

First, the vaginal sample was preincubated at 37°C for 21 days in a blood culture bottle (BD Diagnostics, Le Pont-de-Claix, France) enriched with 4 mL rumen that was filter-sterilized through a 0.2 μm pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France) and 3 mL of sheep's blood (bioMérieux, Marcy l'Etoile, France). After 21 days of preincubation, the sample was inoculated on Schaedler agar enriched with sheep's blood and vitamin K (BD Diagnostics) and incubated for 7 days in anaerobic conditions at 37°C. On sheep's blood agar (bioMérieux), colonies were white with a mean diameter of 2 mm. Bacterial cells were Gram-positive cocci. Catalase activity was positive; oxidase activity was negative.

The 16S rRNA gene was amplified and sequenced using the universal primers (fD1 and rp2) and a 3130-XL sequencer (Applied Biosciences, Saint Aubin, France), as described elsewhere [4]. Strain Marseille P2765 exhibited a 97.2% sequence identity with *Anaerococcus lactolyticus* strain JCM 8140 (GenBank accession no. NR_113565.1), the phylogenetically closest validated species (Fig. 1). This degree of similarity was lower than the 98.7% threshold to define a new species [5], and we propose that strain Marseille P2765 be considered representative of a new species within the *Anaerococcus* genus in the phylum *Firmicutes*. The *Anaerococcus* genus was created by Ezaki *et al.* [6]. This genus is one of the three genera obtained after the subdivision of the *Peptostreptococcus* genus [6]. Bacterial species from the *Anaerococcus* genus have been already reported from diverse human clinical specimens [6]. *Anaerococcus lactolyticus*, the phylogenetically closest validated species, was first isolated from vaginal discharges [6] like strain Marseille P2765[T].

Because strain Marseille P2765 is more than 2.8% divergent in the 16S rRNA gene sequence with its closest phylogenetic neighbour [7], we propose that it may be the representative strain of a novel species named '*Anaerococcus mediterraneensis*' (me.di.ter.ra.ne.en′sis, L. masc. adj., *mediterraneensis*, 'of Mediterraneum,' the Latin name of the Mediterranean Sea by which Marseille, where strain P2765 was isolated, is located). Strain

© 2017 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
http://dx.doi.org/10.1016/j.nmni.2017.02.007

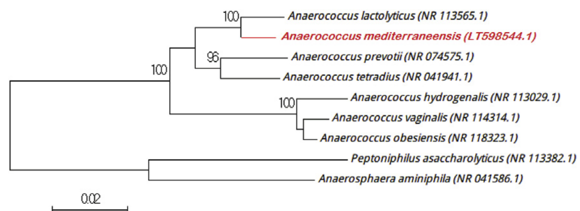


FIG. 1. Phylogenetic tree showing phylogenetic position of '*Anaerococcus mediterraneensis*' strain Marseille 2765^T relative to close species in genus *Anaerococcus*. GenBank accession numbers are indicated after name. Sequences were aligned using Muscle v3.8.31 with default parameters, and phylogenetic inferences were obtained using neighbour-joining method with 500 bootstrap replicates within MEGA6 software. Only bootstrap values >95% are shown. Scale bar represents 2% nucleotide sequence divergence.

Marseille P2765^T is the type strain of the new species '*Anaerococcus mediterraneensis*' sp. nov.

MALDI-TOF MS spectrum

The MALDI-TOF MS spectrum of '*Anaerococcus mediterraneensis*' is available online (http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database).

Nucleotide sequence accession number

The 16S rRNA gene sequence was deposited in European Molecular Biology Laboratory–European Bioinformatics Institute under accession number LN598544.1.

Deposit in a culture collection

The type isolate of '*Anaerococcus mediterraneensis*' was deposited in the collection Deutsche Sammlung von Mikroorganismen (DSM 103343) and the Collection de Souches de l'Unité des Rickettsies (CSUR, WDCM 875) under number P2765.

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] Menard JP, Fenollar F, Henry M, Bretelle F, Raoult D. Molecular quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. Clin Infect Dis 2008;20:33–43.
[3] 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.
[4] 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.
[5] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[6] Ezaki T, Kawamura Y, Li N, Li ZY, Zhao L, Shu S. Proposal of the genera *Anaerococcus* gen. nov., *Peptoniphilus* gen. nov. and *Gallicola* gen. nov. for members of the genus *Peptostreptococcus*. Int J Syst Evol Microbiol 2001;51:1521–8.
[7] Yarza P, Richter M, Peplies J, Euzeby J, Amann R, Schleifer KH, et al. The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. Syst Appl Microbiol 2008;31:241–50.

© 2017 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **17**, 75–76
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Article 15:

"*Arcanobacterium urinimassiliense*" sp. nov., a new bacterium isolated from the urogenital tract.

Khoudia Diop, Aurelie Morand, Jean-Christophe Dubus, Pierre-Edouard Fournier, Didier Raoult, Florence Fenollar

Published in New Microbes and New Infections

'*Arcanobacterium urinimassiliense*' sp. nov., a new bacterium isolated from the urogenital tract

K. Diop[1], A. Morand[1,2], J. C. Dubus[2], P.-E. Fournier[1], D. Raoult[1,3] and F. Fenollar[1]

1) *Aix-Marseille Université, Institut hospitalo-universitaire Méditerranée-infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine,* 2) *Médecine infantile—Hôpital Nord, Chemin des Bourrely, Marseille, France and* 3) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

Herein we report the main characteristics of *'Arcanobacterium urinimassiliense'* strain Marseille-P3248^T (=CSUR P3248) isolated from a urine sample of a 54-day-old girl with rotavirus gastroenteritis.

© 2017 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

Keywords: '*Arcanobacterium urinimassiliense*', culturomics, human microbiota, rotavirus gastroenteritis, taxonomy
Original Submission: 8 February 2017; **Revised Submission:** 18 March 2017; **Accepted:** 28 March 2017
Article published online: 31 March 2017

Corresponding author: F. Fenollar, Aix-Marseille Université, Institut hospitalo-universitaire Méditerranée-infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine, 19-21 Boulevard Jean Moulin, 13385 Marseille cedex 05, France.
E-mail: florence.fenollar@univ-amu.fr

In 2016, as a part of the culturomics study [1,2] of the human microbiome, a bacterial strain that could not be identified by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) using a Microflex spectrometer (Bruker Daltonics, Bremen, Germany) [3] was isolated from a urine sample of a 54-day-old girl with rotavirus gastroenteritis. The study was approved by the local ethics committee of the Institut Federatif de Recherche IFR48 (Marseille, France) under the agreement number 09-022. The parents provided written informed consent.

The sample was preincubated in an anaerobic blood culture bottle (Becton Dickinson, Le Pont-de-Claix, France) supplemented with 4 mL rumen that was filter-sterilized through a 0.2 μm pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France) and 5 mL of defibrined sheep's blood (bioMérieux, Marcy l'Etoile, France). After 30 days of preincubation, the supernatant was cultured on homemade R medium (Timone Hospital, Marseille, France) and then incubated in anaerobic atmosphere generated using the GENbag Anaer system (bioMérieux). After 3 days of incubation, strain Marseille-P3248^T was isolated. On agar, colonies were small and beige with a mean diameter of 200 μm. Bacterial cells were Gram variable and rod shaped, with length ranging from 400 to 600 nm and width ranging from 300 to 400 nm. Strain Marseille-P3248^T was nonmobile. Catalase and oxidase reactions were negative.

The 16S rRNA gene was sequenced using the fD1-rP2 primers as previously described [4], using a 3130-XL sequencer (Applied Biosciences, Saint-Aubin, France). Strain Marseille-P3248^T exhibited 94.7% sequence similarity with *Arcanobacterium phocae* strain DSM 10002 (GenBank accession no. NR-117159) [5], its phylogenetically closest species with standing in nomenclature (Fig. 1). Because this sequence was smaller than the 98.65% threshold set defined by Kim *et al.* [6] to support a new species, strain Marseille-P3248^T can be classified as a new species of *Arcanobacterium* genus belonging to the family *Actinomycetaceae* classified within the *Actinobacteria* phylum.

Because strain Marseille-P3248^T exhibited a 16S rRNA gene sequence divergence of 3.95% with its phylogenetically closest species with standing in nomenclature [6], we propose that strain Marseille-P3248 may be the representative strain of the new species called *'Arcanobacterium urinimassiliense'* (u.ri.ni.mas.sil.ien′se, N.L. u.ri.no, N.L. gen. fem. *urina*, 'urine,' from which this bacterium was first cultivated; and mas.si.li.en′sis, L. gen. adj. *massiliensis*, from 'Massilia,' the Latin name of

© 2017 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
http://dx.doi.org/10.1016/j.nmni.2017.03.003

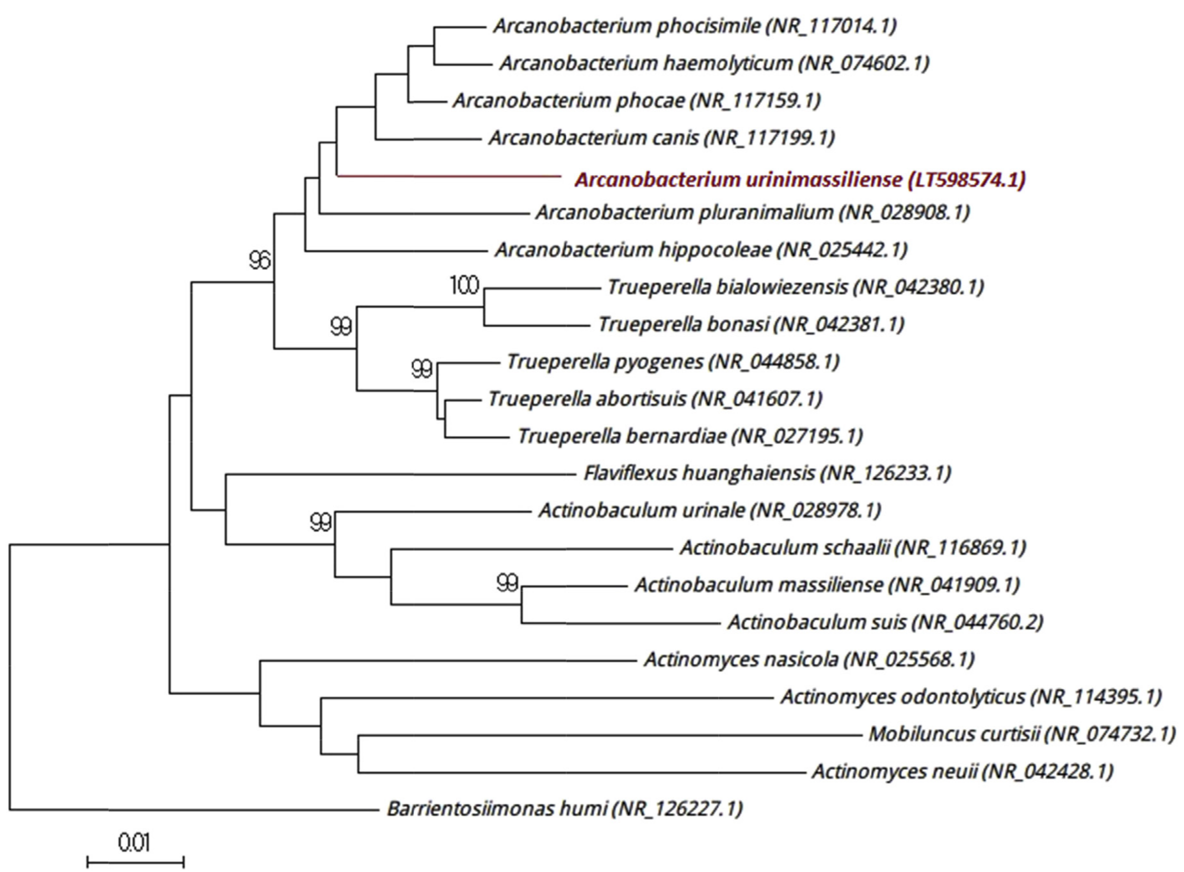


FIG. 1. Phylogenetic tree showing position of *'Arcanobacterium urinimassiliense'* strain Marseille-P3248^T relative to other phylogenetically close neighbours. Sequences were aligned by CLUSTALW, and phylogenetic inferences were obtained by maximum-likelihood method within MEGA software. Number at node is percentages of bootstrap value ($\geq$95%) obtained by repeating analysis 500 times to generate majority consensus tree. GenBank accession numbers are indicated in parentheses. Scale bar indicates 1% nucleotide sequence divergence.

Marseille, France, where the microorganism was first isolated). Strain Marseille-P3248^T is the type strain.

MALDI-TOF MS spectrum

The MALDI-TOF MS spectrum of *'Arcanobacterium urinimassiliense'* strain Marseille-P3248^T is available online (http://mediterranee-infection.com/article.php?laref=256&titre=urms-database).

Nucleotide sequence accession number

The 16S rRNA gene sequence of the strain Marseille-P3248^T was deposited in GenBank under accession number LT598574.

Deposit in a culture collection

Strain Marseille-P3248^T was deposited in the Collection de Souches de l'Unité des Rickettsies (CSUR) under number P3248.

Acknowledgement

This study was funded by the Fondation Méditerranée Infection.

Conflict of Interest

None declared.

© 2017 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **18**, 15–17
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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] Hilt EE, McKinley K, Pearce MM, Rosenfeld AB, Zilliox MJ, Mueller ER, et al. Urine is not sterile: use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. J Clin Microbiol 2014;5:871–6.

[3] 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.

[4] 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.

[5] Ramos CP, Foster G, Collins MD. Phylogenetic analysis of the genus actinomyces based on 16S rRNA gene sequences: description of *Arcanobacterium phocae* sp. nov., *Arcanobacterium bernardiae* comb. nov., and *Arcanobacterium pyogenes* comb. nov. Int J Syst Evol Microbiol 1997;47:46–53.

[6] 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;64:346–51.

© 2017 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **18**, 15–17
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Article 16:

‘‘*Collinsella vaginalis*” sp. nov., a new bacterial species cultivated from human female genital tract.

Khoudia Diop, Florence Bretelle, Pierre-Edouard Fournier, Didier Raoult, Florence Fenollar

Published in Human Microbiome Journal

Human Microbiome Journal 2 (2016) 19–20

Contents lists available at ScienceDirect

Human Microbiome Journal

journal homepage: www.elsevier.com/locate/humic

"*Collinsella vaginalis*" sp. nov., a new bacterial species cultivated from human female genital tract

Khoudia Diop [a], Florence Bretelle [a,b], Pierre-Edouard Fournier [a], Didier Raoult [a,c], Florence Fenollar [a,*]

[a] *Aix-Marseille Université, Institut hospitalo-universitaire Méditerranée-infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, Faculté de médecine, 27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France*

[b] *Department of Gynecology and Obstetrics, Gynépole, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, AMU, Aix-Marseille Université, France*

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

ARTICLE INFO

Article history:
Received 14 September 2016
Revised 19 October 2016
Accepted 8 November 2016
Available online 15 November 2016

Keywords:
"*Collinsella vaginalis*"
Culturomics
Vaginal flora
Bacterial vaginosis
Human microbiota

ABSTRACT

We present a brief description of "*Collinsella vaginalis*" strain P2666 (=CSUR P2666), a new bacterium that was cultivated from the vaginal sample of a 26-year-old woman affected from bacterial vaginosis.

© 2016 Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

As part of the study of the human microbiota thanks to the approach of microbial culturomics [1], we analyzed the vaginal flora of women with bacterial vaginosis [2]. In May 2015, we cultivated from the vaginal swab of a 26 year-old French patient a bacterial strain that could not be identified using matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (Microflex spectrometer, Bruker Daltonics, Leipzig, Germany) [3]. The study was authorized by the local ethics committee of the IFR48 (Marseille, France; agreement 09-022). The patient gave also her written consent.

The initial growth of strain Marseille P2666 was obtained at 37 °C under anaerobic conditions after 7 days of culture on CNA (Colistin and Naladixic Acid) agar (BD Diagnostics, Le Pont-de-Claix, France) after 15 days of pre-incubation in a blood culture bottle (BD Diagnostics) enriched with 4 ml of rumen that was filter-sterilized through at 0.2 µm pore filter (Thermo Fisher Scientific, Villebon-sur-Yvette, France) and 3 ml of sheep blood (bioMérieux, Marcy l'Etoile, France). Bacterial cells are rod-shaped Gram-positive, strictly anaerobic, non-motile, and non-spore-forming with a mean diameter of 0.4 µm and a mean length of 1.8 µm. After 2 days of incubation at 37 °C under anaerobic conditions on blood agar (bioMérieux), colonies are grey, circular, and opaque with a diameter of 0.5–1.2 mm. Strain Marseille P2666 exhibited neither oxidase nor catalase activity.

The 16S rDNA sequence was obtained after amplification using the universal primer pair (fD1 and rp2) and a 3130-XL sequencer (Applied Biosciences, Saint-Aubin, France), as previously reported [4]. 16S rRNA gene sequence-based identification of strain Marseille P2666 exhibited 96.08% of identity with *Collinsella intestinalis* strain JCM 10643 (GenBank accession number NR_113165), the phylogenetically closest bacterium with a validly published name (Fig. 1). As this sequence was below the 98.7% threshold to define a new species [5], strain Marseille P2666 was considered as a new species within the *Collinsella* genus in the *Coriobacteriaceae* family. Created in 1999, the genus *Collinsella* contains currently 4 species [6]; all were isolated from human faeces.

Strain Marseille P2666 presents a 16S rRNA divergence around 3.8% with its phylogenetically closest species [7], we propose that strain Marseille P2666 may be the representative of a novel species named "*Collinsella vaginalis*" sp. nov. (va.gi.na'lis. L. n. vagina sheath, vagina; L. masc. suff. -alis suffix denoting pertaining to; N.L. masc. adj. vaginalis pertaining to vagina, of the vagina). Strain Marseille P2666^T is the type strain of the new species "*Collinsella vaginalis*" sp. nov.

MALDI-TOF MS spectrum accession number. The MALDI-TOF MS spectrum of Collinsella vaginalis" is available at http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database.

* Corresponding author.
E-mail address: florence.fenollar@univ-amu.fr (F. Fenollar).

http://dx.doi.org/10.1016/j.humic.2016.11.001
2452-2317/© 2016 Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

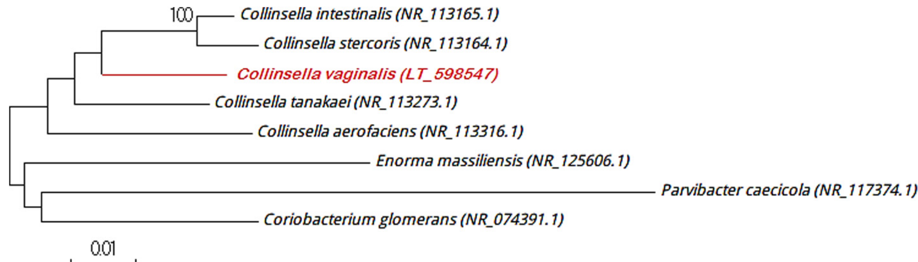


Fig. 1. Phylogenetic tree highlighting the position of "*Collinsella vaginalis*" strain Marseille P2666^T relative to other closest species. GenBank accession numbers are indicated after the name. Sequences were aligned using Muscle v3.8.31 with default parameters and phylogenetic inferences were obtained using neighbor-joining method with 500 bootstrap replicates, within MEGA6 software. Only bootstrap bigger than 95% are shown. The scale bar indicates a 1% nucleotide sequence divergence.

Nucleotide sequence accession number. The 16S rRNA gene sequence was deposited in EMBL-EBI under accession number LT598547.

Deposit in culture collection. Strain Marseille P2666 was deposited in the "Collection de Souches de l'Unité des Rickettsies" (CSUR, WDCM 875) under number P2666.

Acknowledgments

This research is funded by the Fondation Méditerranée Infection.

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] Menard JP, Fenollar F, Henry M, Bretelle F, Raoult D. Molecular quantification of *Gardnerella vaginalis* and *Atopobium vaginae* loads to predict bacterial vaginosis. Clin Infect Dis 2008;20:33–43.

[3] 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.

[4] 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.

[5] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.

[6] Collinsella. http://www.bacterio.net/collinsella.html. [Accessed: 10-Sept-2016].

[7] Yarza P, Richter M, Peplies J, Euzeby J, Amann R, Schleifer KH, et al. The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. Syst Appl Microbiol 2008;31:241–50.

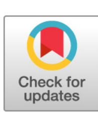
- **Description du génome bactérien d'*Ezakiella peruensis***

Article 17:

Draft Genome Sequence of *Ezakiella peruensis* Strain M6X2^T, a human fecal Gram-stain positive anaerobic coccus.

Awa Diop, **Khoudia Diop**, Enora Tomei, Didier Raoult, Florence Fenollar, Pierre-Edouard Fournier

Published in Genome Announcement

PROKARYOTES


Draft Genome Sequence of *Ezakiella peruensis* Strain M6.X2, a Human Gut Gram-Positive Anaerobic Coccus

Awa Diop,[a] Khoudia Diop,[a] Enora Tomei,[a] Didier Raoult,[a,b] Florence Fenollar,[a] Pierre-Edouard Fournier[a]

[a]Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, Aix-Marseille Université, UM 63, CNRS UMR7278, IRD 198, INSERM U1095, Assistance Publique-Hôpitaux de Marseille, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de Médecine, Marseille, France

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

Downloaded from http://genomea.asm.org/ on March 2, 2018 by guest

ABSTRACT We report here the draft genome sequence of *Ezakiella peruensis* strain M6.X2^T. The draft genome is 1,672,788 bp long and harbors 1,589 predicted protein-encoding genes, including 26 antibiotic resistance genes with 1 gene encoding vancomycin resistance. The genome also exhibits 1 clustered regularly interspaced short palindromic repeat region and 333 genes acquired by horizontal gene transfer.

Ezakiella peruensis is the type and only species of the genus *Ezakiella*, created in 2015 (1). *E. peruensis* occupies a unique position in an undefined family within the phylum *Firmicutes* (1). This microorganism is a Gram-positive anaerobic coccus. Gram-positive anaerobic cocci include many commensal species of humans and animals and also some human pathogens (2). The type strain M6.X2^T was isolated from a fecal sample of a healthy individual residing in a coastal traditional community in Peru (1). It is nonmotile and non-spore forming. Here, we present the annotated draft genome sequence of *E. peruensis* strain M6.X2^T (DSM 27367 = NBRC 109957 = CCUG 64571), which we obtained from the DSMZ collection.

Genomic DNA of *E. peruensis* strain M6.X2^T was sequenced using a MiSeq sequencer with the mate-pair strategy (Illumina, Inc., San Diego, CA, USA). DNA was quantified by a Qubit assay with a high-sensitivity kit (Life Technologies, Carlsbad, CA, USA) at 38.4 ng/μl. The 576,285 high-quality paired-end reads were trimmed and then assembled using the SPAdes assembler program (3). The draft genome sequence was annotated using Prokka software (4). Functional annotation was achieved using the BLASTp algorithm (5) against the Clusters of Orthologous Groups (COGs) database and the Rapid Annotations using Subsystems Technology (RAST) web server (6). Ribosomal RNAs (5S, 16S, and 23S rRNAs) were predicted using RNAmmer software (7).

The genome was 1,672,788-bp long, assembled in five scaffolds (seven contigs) with a G+C content of 36.9%. Overall, 1,589 protein-coding sequences were identified, including 1,165 (73.31%) protein-coding genes that had orthologs in the COGs database, 1,052 of which were assigned a putative function. A total of 46 tRNA loci and 1 rRNA operon (16S, 5S, and 23S rRNA) were identified in the genome. Strain M6.X2^T exhibited 26 genes associated with antibiotic resistance and toxic compounds, including one *vanW* gene encoding vancomycin resistance. No toxin/antitoxin module or bacteriocin-associated gene was identified. The genome of *E. peruensis* harbored 1 clustered regularly interspaced short palindromic repeat locus of 763 bp with 12 repeats (mean repeat length = 36 bp). We also detected 333 putative genes acquired by horizontal gene transfer, including 209 from bacteria within the order *Clostridiales*.

Accession number(s). The 16S rRNA and genome sequences from *Ezakiella peruensis* strain M6.X2^T are available in GenBank under accession numbers KJ469554 and OCSL00000000, respectively.

Received 28 November 2017 **Accepted** 6 February 2018 **Published** 1 March 2018

Citation Diop A, Diop K, Tomei E, Raoult D, Fenollar F, Fournier P-E. 2018. Draft genome sequence of *Ezakiella peruensis* strain M6.X2, a human gut Gram-positive anaerobic coccus. Genome Announc 6:e01487-17. https://doi.org/10.1128/genomeA.01487-17.

Copyright © 2018 Diop et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Pierre-Edouard Fournier, pierre-edouard.fournier@univ-amu.fr.

ACKNOWLEDGMENTS

This study was funded by the Méditerranée-Infection Foundation and the French Agence National de la Recherche under reference Investissements d'Avenir Méditerranée-Infection 10-IAHU-03.

REFERENCES

1. Patel NB, Tito RY, Obregón-Tito AJ, O'Neal L, Trujillo-Villaroel O, Marin-Reyes L, Troncoso-Corzo L, Guija-Poma E, Hamada M, Uchino Y, Lewis CM, Lawson PA. 2015. *Ezakiella peruensis* gen. nov., sp. nov. isolated from human fecal sample from a coastal traditional community in Peru. Anaerobe 32:43–48. https://doi.org/10.1016/j.anaerobe.2014.12.002.
2. Ulger-Toprak N, Liu C, Summanen PH, Finegold SM. 2010. *Murdochiella asaccharolytica* gen. nov., sp. nov., a Gram-stain-positive, anaerobic coccus isolated from human wound specimens. Int J Syst Evol Microbiol 60:1013–1016. https://doi.org/10.1099/ijs.0.015909-0.
3. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
4. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
5. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421.
6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.
7. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.

Downloaded from http://genomea.asm.org/ on March 2, 2018 by guest

- **Culturomics, un envol de l’exploration du microbiote humain**

Article 18:

Culture of previously uncultured members of the human gut microbiota by culturomics.

Jean-Christophe Lagier, Saber Khelaifia, Maryam Tidjani Alou, Sokhna Ndongo, Niokhor Dione, Perrine Hugon, Aurelia Caputo, Frédéric Cadoret, Sory Ibrahima Traore, El Hadji Seck, Gregory Dubourg, Guillaume Durand, Gaël Mourembou, Elodie Guilhot, Amadou Togo, Sara Bellali, Dipankar Bachar, Nadim Cassir, Fadi Bittar, Jérémy Delerce, Morgane Mailhe, Davide Ricaboni, Melhem Bilen, Nicole Prisca Makaya Dangui, Souleymane Ndeye Mery Dia Badiane, Camille Valles, Donia Mouelhi, **Khoudia Diop**, Matthieu Million, Didier Musso, Jõnatas Abrahao, Esam Ibraheem Azhar, Fehmida Bibi, Muhammad Yasir, Aldiouma Diallo, Cheikh Sokhna, Felix Djossou, Véronique Vitton, Catherine Robert, Jean Marc Rolain, Bernard La Scola, Pierre-Edouard Fournier, Anthony Levasseur and Didier Raoult

Published in Nature Microbiology

nature microbiology

LETTERS

PUBLISHED: 7 NOVEMBER 2016 | ARTICLE NUMBER: 16203 | DOI: 10.1038/NMICROBIOL.2016.203

OPEN

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

Jean-Christophe Lagier[1], Saber Khelaifia[1], Maryam Tidjani Alou[1], Sokhna Ndongo[1], Niokhor Dione[1], Perrine Hugon[1], Aurelia Caputo[1], Frédéric Cadoret[1], Sory Ibrahima Traore[1], El Hadji Seck[1], Gregory Dubourg[1], Guillaume Durand[1], Gaël Mourembou[1], Elodie Guilhot[1], Amadou Togo[1], Sara Bellali[1], Dipankar Bachar[1], Nadim Cassir[1], Fadi Bittar[1], Jérémy Delerce[1], Morgane Mailhe[1], Davide Ricaboni[1], Melhem Bilen[1], Nicole Prisca Makaya Dangui Nieko[1], Ndeye Mery Dia Badiane[1], Camille Valles[1], Donia Mouelhi[1], Khoudia Diop[1], Matthieu Million[1], Didier Musso[2], Jônatas Abrahão[3], Esam Ibraheem Azhar[4], Fehmida Bibi[4], Muhammad Yasir[4], Aldiouma Diallo[5], Cheikh Sokhna[5], Felix Djossou[6], Véronique Vitton[7], Catherine Robert[1], Jean Marc Rolain[1], Bernard La Scola[1], Pierre-Edouard Fournier[1], Anthony Levasseur[1] and Didier Raoult[1]*

Metagenomics revolutionized the understanding of the relations among the human microbiome, health and diseases, but generated a countless number of sequences that have not been assigned to a known microorganism[1]. The pure culture of prokaryotes, neglected in recent decades, remains essential to elucidating the role of these organisms[2]. We recently introduced microbial culturomics, a culturing approach that uses multiple culture conditions and matrix-assisted laser desorption/ionization–time of flight and 16S rRNA for identification[2]. Here, we have selected the best culture conditions to increase the number of studied samples and have applied new protocols (fresh-sample inoculation; detection of microcolonies and specific cultures of Proteobacteria and microaerophilic and halophilic prokaryotes) to address the weaknesses of the previous studies[3–5]. We identified 1,057 prokaryotic species, thereby adding 531 species to the human gut repertoire: 146 bacteria known in humans but not in the gut, 187 bacteria and 1 archaea not previously isolated in humans, and 197 potentially new species. Genome sequencing was performed on the new species. By comparing the results of the metagenomic and culturomic analyses, we show that the use of culturomics allows the culture of organisms corresponding to sequences previously not assigned. Altogether, culturomics doubles the number of species isolated at least once from the human gut.

The study of the human gut microbiota has been revived by metagenomic studies[6–8]. However, a growing problem is the gaps that remain in metagenomics, which correspond to unidentified sequences that may be correlated with an identified organism[9]. Moreover, the exploration of relations between the microbiota and human health require—both for an experimental model and therapeutic strategies—the growing of microorganisms in pure culture[10], as recently demonstrated in elucidations of the role of *Clostridium butyricum* in necrotizing enterocolitis and the influence of gut microbiota on cancer immunotherapy effects[11,12]. In recent years, microbial culture techniques have been neglected, which explains why the known microbial community of the human gut is extremely low[13]. Before we initiated microbial culturomics[13] of the approximately 13,410 known bacterial and archaea species, 2,152 had been identified in humans and 688 bacteria and 2 archaea had been identified in the human gut. Culturomics consists of the application of high-throughput culture conditions to the study of the human microbiota and uses matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) or 16S rRNA amplification and sequencing for the identification of growing colonies, some of which have been previously unidentified[2]. With the prospect of identifying new genes of the human gut microbiota, we extend here the number of recognized bacterial species and evaluate the role of this strategy in resolving the gaps in metagenomics, detailing our strategy step by step (see Methods). To increase the diversity, we also obtained frozen samples from healthy individuals or patients with various diseases from different geographical origins. These frozen samples were collected as fresh samples (stool, small-bowel and colonic samples; Supplementary Table 1). Furthermore, to determine appropriate culture conditions, we first reduced the number of culture conditions used (Supplementary Table 2a–c) and then focused on specific strategies for some taxa that we had previously failed to isolate (Supplementary Table 3).

First, we standardized the microbial culturomics for application to the sample testing (Supplementary Table 1). A refined analysis of our first study, which had tested 212 culture conditions[4], showed that all identified bacteria were cultured at least once using one of the 70 best culture conditions (Supplementary Table 2a). We applied these 70 culture conditions (Supplementary Table 2a) to the study of 12 stool samples (Supplementary Table 1). Thanks to the implementation of the recently published repertoire of human bacteria[13] (see Methods), we determined that the isolated bacteria included 46 bacteria known from the gut but not recovered by culturomics before this work (new for culturomics), 38 that had

[1]Aix Marseille Université URMITE, UM63, CNRS 7278, IRD 198, INSERM 1095, 27 Boulevard Jean Moulin, 13385 Marseille Cedex 5, France. [2]Institut Louis Malardé, Papeete, Tahiti, Polynésie Française. [3]Departamento de Microbiologia Laboratorio de Virus, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil. [4]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah 21589, Saudi Arabia. [5]Institut de Recherche pour le Développement, UMR 198 (URMITE), Campus International de Hann, IRD, BP 1386, CP, 18524 Dakar, Sénégal. [6]Department of Infectious and Tropical Diseases, Centre Hospitalier de Cayenne, Cayenne, French Guiana. [7]Service de Gastroentérologie, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, 13915 Marseille, France. *e-mail: didier.raoult@gmail.com

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

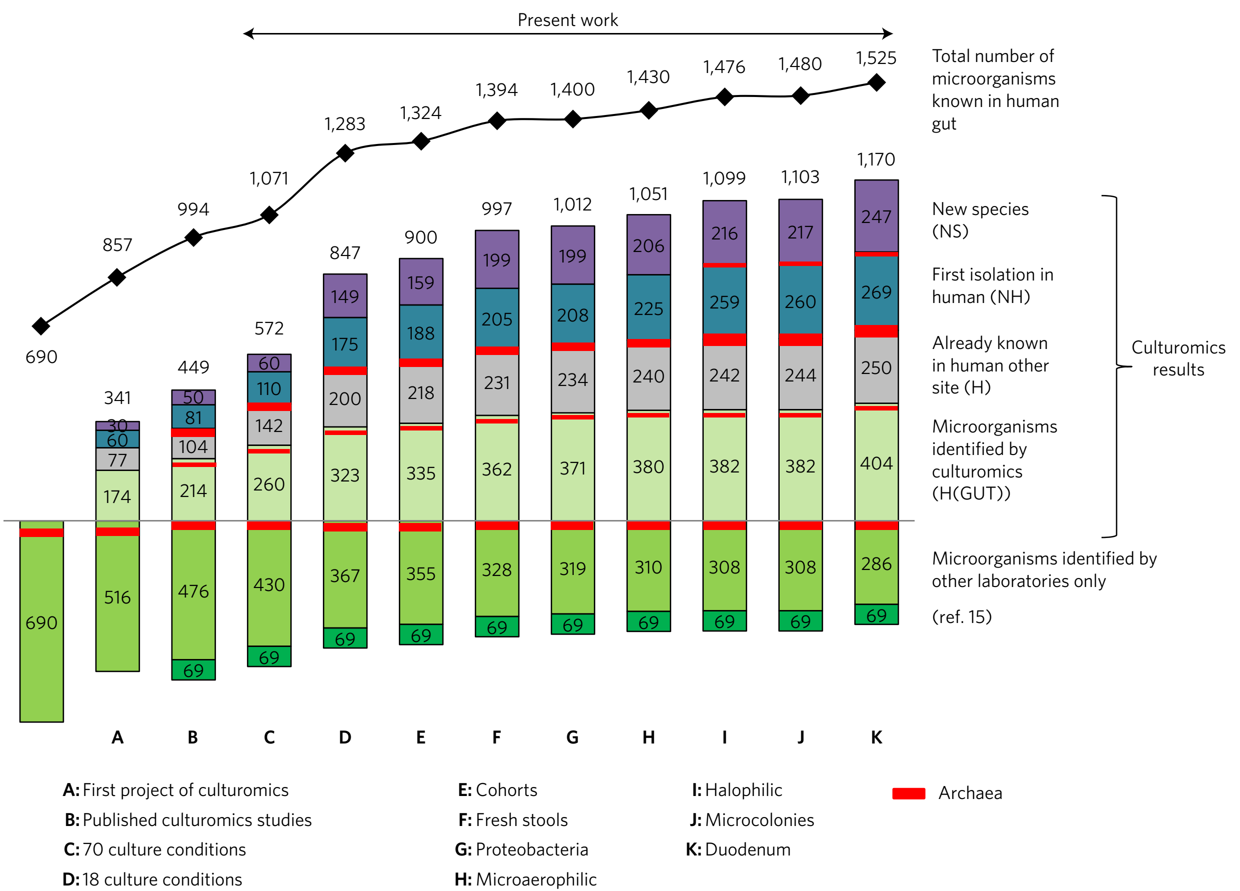


Figure 1 | Number of different bacteria and archaea isolated during the culturomics studies. Columns A and B represent the results from previously published studies, and columns C to K the different projects described herein. The bacterial species are represented in five categories: NS, new species; NH, prokaryotes first isolated in humans; H, prokaryotes already known in humans but never isolated from the human gut; H (GUT), prokaryotes known in the human gut but newly isolated by culturomics; and prokaryotes isolated by other laboratories but not by culturomics.

already been isolated in humans but not from the gut (non-gut bacteria), 29 that had been isolated in humans for the first time (non-human bacteria) and 10 that were completely new species (unknown bacteria) (Fig. 1 and Supplementary Tables 4a and 5).

Beginning in 2014, to reduce the culturomics workload and extend our stool-testing capabilities, we analysed previous studies and selected the 18 best culture conditions[2]. We performed cultures in liquid media in blood culture bottles, followed by subcultures on agar (Supplementary Table 2b). We designed these culture conditions by analysing our first studies. The results of those studies indicated that emphasizing three components was essential: pre-incubation in a blood culture bottle (56% of the new species isolated), the addition of rumen fluid (40% of the new species isolated) and the addition of sheep blood (25% of the new species isolated)[2–5]. We applied this strategy to 37 stool samples from healthy individuals with different geographic provenances and from patients with different diseases (Supplementary Table 1). This new strategy enabled the culture of 63 organisms new to culturomics, 58 non-gut bacteria, 65 non-human bacteria and 89 unknown bacteria (Fig. 1 and Supplementary Tables 4a and 5).

We also applied culturomic conditions (Supplementary Table 2c) to large cohorts of patients sampled for other purposes (premature infants with necrotizing enterocolitis, pilgrims returning from the Hajj and patients before or after bariatric surgery) (Supplementary Table 1). A total of 330 stool samples were analysed. This enabled the detection of 13 bacteria new to culturomics, 18 non-gut bacteria, 13 non-human bacteria and 10 unknown species (Fig. 1 and Supplementary Tables 4a and 5).

Among the gut species mentioned in the literature[13] and not previously recovered by culturomics, several were extremely oxygen-sensitive anaerobes, several were microaerophilic and several were Proteobacteria, and we focused on these bacteria (Supplementary Table 3). Because delay and storage may be critical with anaerobes, we inoculated 28 stools immediately upon collection. This enabled the culture of 27 new gut species for culturomics, 13 non-gut bacteria, 17 non-human bacteria and 40 unknown bacteria (Fig. 1 and Supplementary Tables 3a and 4). When we specifically tested 110 samples for Proteobacteria, we isolated 9 bacteria new to culturomics, 3 non-gut bacteria and 3 non-human bacteria (Fig. 1 and Supplementary Tables 4a and 5). By culturing 242 stool specimens exclusively under a microaerophilic atmosphere, we isolated 9 bacteria new to culturomics, 6 non-gut bacteria, 17 non-human bacteria and 7 unknown bacteria (Fig. 1 and Supplementary Tables 4a and 5). We also introduced the culture of halophilic prokaryotes from the gut and microcolony detection. The culture of halophilic bacteria was performed using culture media supplemented with salt for 215 stool samples, allowing the culture of 48 halophilic prokaryotic species, including one archaea (*Haloferax alexandrinus*), 2 new bacteria for culturomics, 2 non-gut bacteria, 34 non-human bacteria, 10 unknown bacteria and one new halophilic archaea (*Haloferax massiliensis* sp. nov.) (Fig. 1 and Supplementary Tables 4a and 5). Among these 48 halophilic prokaryotic species, 7 were slight halophiles (growing with 10–50 g l^{-1} of NaCl), 39 moderate halophiles (growing with 50–200 g l^{-1} of NaCl) and 2 extreme halophiles (growing with 200–300 g l^{-1} of NaCl).

We also introduced the detection of microcolonies that were barely visible to the naked eye (diameters ranging from 100 to

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

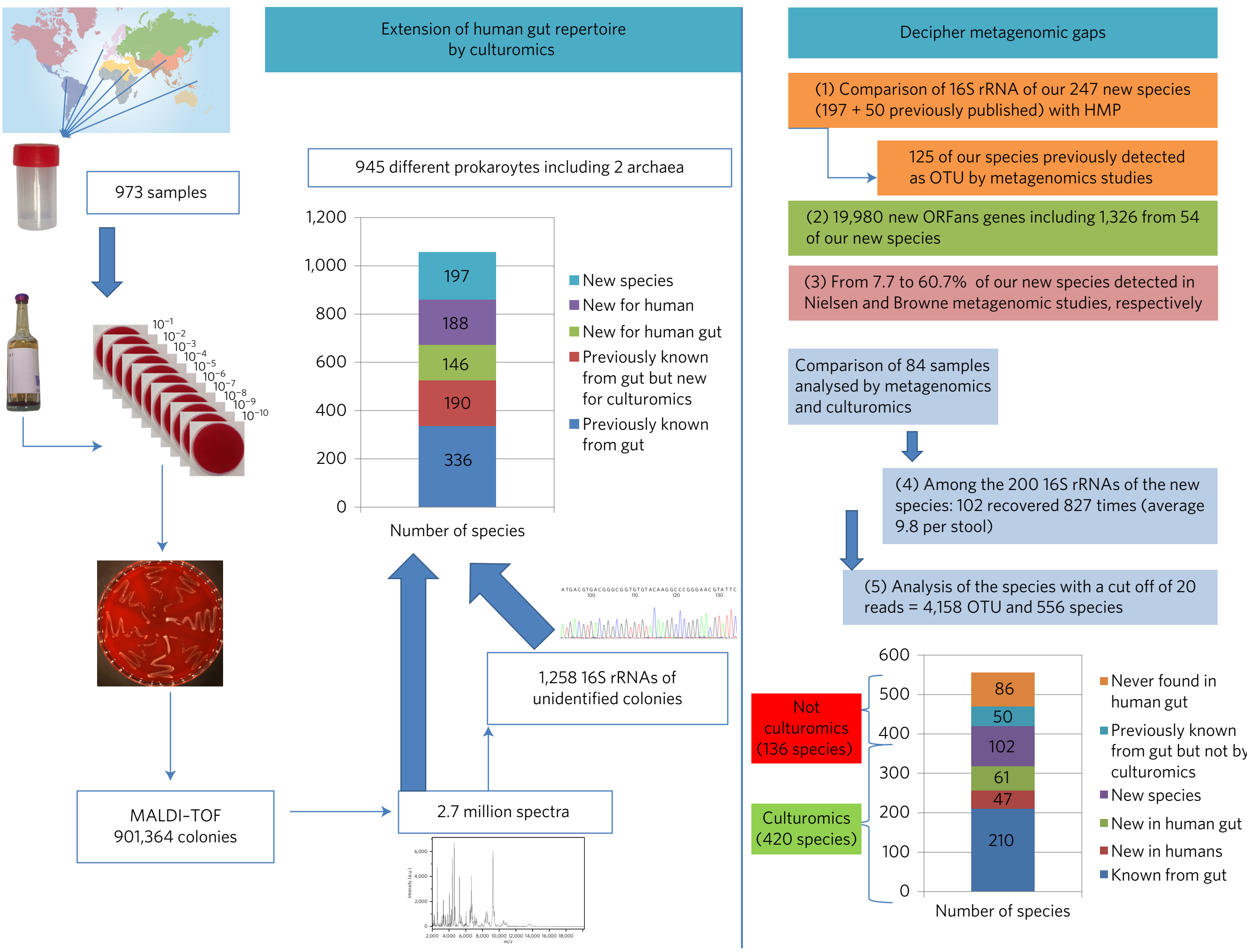


Figure 2 | Summary of the culturomics work that has extended the gut repertoire and filled some of the gaps in metagenomics.

300 μm) and could only be viewed with magnifying glasses. These colonies were transferred into a liquid culture enrichment medium for identification by MALDI–TOF mass spectrometry (MS) or 16S rRNA amplification and sequencing. By testing ten stool samples, we detected two non-gut bacteria, one non-human bacterium and one unknown bacterium that only formed micro-colonies (Fig. 1 and Supplementary Tables 4a and 5). Finally, by culturing 30 duodenal, small bowel intestine and colonic samples, we isolated 22 bacteria new to culturomics, 6 non-gut bacteria, 9 non-human bacteria and 30 unknown bacteria (Fig. 1 and Supplementary Tables 4a and 5). To continue the exploration of gut microbiota, future culturomics studies could also be applied to intestinal biopsies.

In addition, we performed five studies to evaluate the role of culturomics for deciphering the gaps in metagenomics[9]. First, we compared the 16S rRNA sequences of the 247 new species (the 197 new prokaryotic species isolated here in addition to the 50 new bacterial species isolated in previous culturomic studies[3–5]) to the 5,577,630 reads from the 16S rRNA metagenomic studies listed by the Human Microbiome Project (HMP) (http://www.hmpdacc.org/catalog). We found sequences, previously termed operational taxonomic units (OTUs), for 125 of our bacterial species (50.6%). These identified bacterial species included *Bacteroides bouchedurhonense*, which was recovered in 44,428 reads, showing that it is a common bacterium (Supplementary Table 6). Second, because the genome sequencing of 168 of these new species allowed the generation of 19,980 new genes that were previously unknown (ORFans genes) (Supplementary Table 7), we blasted these with 13,984,809 contigs/scaffolds from the assembly of whole metagenomic studies by HMP, enabling the detection of 1,326 ORFans (6.6%) from 54 of our new bacterial species (including 45 detected also from 16S) (Supplementary Table 8). Therefore, at least 102 new bacterial species were found but not identified in previous metagenomic studies from the HMP. Third, we searched for our 247 new species in the 239 human gut microbiome samples from healthy individuals described by Browne *et al.*, in which 137 bacterial species were isolated[15]. We captured 150 of our new species in these metagenomics data, representing 60.7% (Supplementary Table 9). Moreover, we also identified 19 of our species (7.7%) from 396 human stool individuals described by Nielsen *et al.*, from which 741 metagenomic species and 238 unique metagenomic genomes were identified[16] (Supplementary Table 9). Fourth, we analysed the 16S rRNA metagenomic sequences of 84 stools also tested by culturomics (Supplementary Table 10). We compared the OTUs identified by blast with a database including the 16S rRNA of all species isolated by culturomics. Among the 247 16S rRNA of the new species, 102 were recovered 827 times, with an average of 9.8 species per stool. Finally, analysis of these species using a cutoff threshold of 20 reads identified 4,158 OTUs and 556 (13.4%) species (Supplementary Table 11), among which 420 species (75.5%) were recovered by culturomics. Of these, 210 (50%) were previously found to be associated with the human gut, 47 were not previously found in humans (11.2%), 61 were found in humans but not in the gut (14.5%) and 102 (24.3%) were new species.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Figure 3 | Phylogenetic tree of the 247 new prokaryote species isolated by culturomics. Bacterial species from Firmicutes are highlighted in red, Actinobacteria (light green), Proteobacteria (blue), Bacteroidetes (purple), Synergistetes (green), Fusobacteria (dark green) and Archaea (grey), respectively. The sequences of 16 prokaryotic species belonging to six phyla previously known from the human gut and more frequently isolated by culture in human gut are highlighted in bold and by an asterisk.

Interestingly, among the 136 species not previously found by culturomics, 50 have been found in the gut and 86 have never previously been found in the human gut (Fig. 2 and Supplementary Table 11).

Overall, in this study, by testing 901,364 colonies using MALDI–TOF MS (Supplementary Table 1), we isolated 1,057 bacterial species, including 531 newly found in the human gut. Among them, 146 were non-gut bacteria, 187 were non-human bacteria, one was a non-human halophilic archaeon and 197 were unknown bacteria, including two new families (represented by *Neofamilia massiliensis* gen. nov., sp. nov. and *Beduinella massiliensis* gen. nov., sp. nov.) and one unknown halophilic archaeon (Fig. 1 and Supplementary Table 4a). Among these, 600 bacterial species belonged to Firmicutes, 181 to Actinobacteria, 173 to Proteobacteria (a phylum that we have under-cultured to date; Supplementary Table 5), 88 to Bacteroidetes, 9 to Fusobacteria, 3 to Synergistetes, 2 to Euryarchaeota, 1 to Lentisphaerae and 1 to Verrucomicrobia (Supplementary Table 4a). Among these 197 new prokaryotes species, 106 (54%) were detected in at least two stool samples, including a species that was cultured in 13 different stools (*Anaerosalibacter massiliensis*) (Supplementary Table 4a). In comparison with our contribution, a recent work using a single culture medium was able to culture 120 bacterial species, including 51 species known from the gut, 1 non-gut bacterium, 1 non-human bacterium and 67 unknown bacteria, including two new families (Supplementary Table 12).

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

To obtain these significant results we tested more than 900,000 colonies, generating 2.7 million spectra, and performed 1,258 molecular identifications of bacteria not identified through MALDI–TOF, using 16S rRNA amplification and sequencing. The new prokaryote species are available in the Collection de Souches de l'Unité des Rickettsies (CSUR) and Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) (Supplementary Tables 4a and 5). All 16S sequences of the new species and the species unidentified by MALDI–TOF, as well as the genome sequences of the new species, have been deposited in GenBank (Supplementary Tables 5 and 13). In addition, thanks in part to an innovative system using a simple culture for the archaea without an external source of hydrogen[17], among these prokaryotes we isolated eight archaeal species from the human gut, including two new ones for culturomics, one non-gut archaea, four non-human archaea and one new halophilic species.

We believe that this work is a key step in the rebirth of the use of culturing in human microbiology[2–5,16] and only the efforts of several teams around the world in identifying the gut microbiota repertoire will allow an understanding and analysis of the relations between the microbiota and human health, which could then participate in adapting Koch's postulates to include the microbiota[21]. The rebirth of culture, termed culturomics here, has enabled the culturing of 77% of the 1,525 prokaryotes now identified in the human gut (Fig. 1 and Supplementary Table 4b). In addition, 247 new species (197 cultured here plus 50 from previous studies) and their genomes are now available (Fig. 3). The relevance of the new species found by culturomics is emphasized because 12 of them were isolated in our routine microbiology laboratory from 57 diverse clinical samples (Supplementary Table 14). In 2016, 6 of the 374 (1.6%) different identifications performed in the routine laboratory were new species isolated from culturomics. As 519 of the species found by culturomics in the gut for the first time (Fig. 1) were not included in the HMP (Supplementary Table 15) and because hundreds of their genomes are not yet available, the results of this study should prompt further genome sequencing to obtain a better identification in gut metagenomic studies.

Methods

Samples. To obtain a larger diversity of gut microbiota, we analysed 943 different stool samples and 30 small intestine and colonic samples from healthy individuals living or travelling in different geographical regions (Europe, rural and urban Africa, Polynesia, India and so on) and from patients with diverse diseases (for example, anorexia nervosa, obesity, malnutrition and HIV). The main characteristics are summarized in Supplementary Table 1. Consent was obtained from each patient, and the study was approved by the local Ethics Committee of the IFR48 (Marseille, France; agreement no. 09–022). Except for the small intestine and stool samples that we directly inoculated without storage (see sections 'Fresh stool samples' and 'Duodenum and other gut samples'), the faecal samples collected in France were immediately aliquoted and frozen at −80 °C. Those collected in other countries were sent to Marseille on dry ice, then aliquoted and frozen at −80 °C for between 7 days and 12 months before analysis.

Culturomics. Culturomics is a high-throughput method that multiplies culture conditions in order to detect higher bacterial diversity. The first culturomics study concerned three stool samples, 212 culture conditions (including direct inoculation in various culture media), and pre-incubation in blood culture bottles incubated aerobically and anaerobically[4]. Overall, 352 other stool samples, including stool samples from patients with anorexia nervosa[3], patients treated with antibiotics[5], or Senegalese children, both healthy and those with diarrhoea[22], were previously studied by culturomics, and these results have been comprehensively detailed in previous publications[3–5]. In this work, we only included the genome sequences of the 50 new bacterial species isolated in these previous works to contribute to our analysis of culturomics and to fill some of the gaps left by metagenomics. In addition, these previously published data are clearly highlighted in Fig. 1, illustrating the overall contribution of culturomics in exploring the gut microbiota.

Bacterial species isolated from our new projects and described here were obtained using the strategy outlined in the following sections.

Standardization of culturomics for the extension of sample testing. A refined analysis allowed the selection of 70 culture conditions (Supplementary Table 2a) for the growth of all the bacteria[4]. We applied these culture conditions to 12 more stool samples and tested 160,265 colonies by MALDI–TOF (Supplementary Table 1). The 18 best culture conditions were selected using liquid media enrichment in a medium containing blood and rumen fluid and subculturing aerobically and anaerobically in a solid medium (Supplementary Table 2b)[2]. Subcultures were inoculated every three days on solid medium, and each medium was kept for 40 days. We applied these culture conditions to 40 stool samples, ultimately testing 565,242 colonies by MALDI–TOF (Supplementary Table 1).

Cohorts. In parallel to these main culturomics studies, we used fewer culture conditions to analyse a larger number of stool samples. We refer to these projects as cohorts. Four cohorts were analysed (pilgrims returning from the Hajj, premature infants with necrotizing enterocolitis, patients before and after bariatric surgery, and patients for acidophilic bacterial species detection). A total of 330 stool samples generated the 52,618 colonies tested by MALDI–TOF for this project (Supplementary Table 1).

Pilgrims from the Hajj. A cohort of 127 pilgrims was included and 254 rectal swabs were collected from the pilgrims: 127 samples were collected before the Hajj and 127 samples were collected after the Hajj. We inoculated 100 µl of liquid sample in an 8 ml bottle containing Trypticase Soy Broth (BD Diagnostics) and incubated the sample at 37 °C for 1 day. We inoculated 100 µl of the enriched sample into four culture media: Hektoen agar (BD Diagnostics), MacConkey agar+Cefotaxime (bioMérieux), Cepacia agar (AES Chemunex) and Columbia ANC agar (bioMérieux). The sample was diluted 10^{-3} before being plated on the MacConkey and Hektoen agars and 10^{-4} before being plated on the ANC agar. The sample was not diluted before being inoculated on the Cepacia agar. Subcultures were performed on Trypticase Soy Agar (BD Diagnostics) and 3,000 colonies were tested using MALDI–TOF.

Preterm neonates. Preterm neonates were recruited from four neonatal intensive care units (NICUs) in southern France from February 2009 to December 2012 (ref. 12). Only patients with definite or advanced necrotizing enterocolitis corresponding to Bell stages II and III were included. Fifteen controls were matched to 15 patients with necrotizing enterocolitis by sex, gestational age, birth weight, days of life, type of feeding, mode of delivery and duration of previous antibiotic therapy. The stool samples were inoculated into 54 preselected culture conditions (Supplementary Table 2c). The anaerobic cultures were performed in an anaerobic chamber (AES Chemunex). A total of 3,000 colonies were tested by MALDI–TOF for this project.

Stool analyses before and after bariatric surgery. We included 15 patients who had bariatric surgery (sleeve gastrectomy or Roux-en-Y gastric bypass) from 2009 to 2014. All stool samples were frozen before and after surgery. We used two different culture conditions for this project. Each stool sample was diluted in 2 ml of Dulbecco's phosphate-buffered saline, then pre-incubated in both anaerobic (BD Bactec Plus Lytic/10 Anaerobic) and aerobic (BD Bactec Plus Lytic/10 Aerobic) blood culture bottles, with 4 ml of sheep blood and 4 ml of sterile rumen fluid being added as previously described[4]. These cultures were subcultured on days 1, 3, 7, 10, 15, 21 and 30 in 5% sheep blood Columbia agar (bioMérieux), and 33,650 colonies were tested by MALDI–TOF.

Acidophilic bacteria. The pH of each stool sample was measured using a pH meter: 1 g of each stool specimen was diluted in 10 ml of neutral distilled water (pH 7) and centrifuged for 10 min at 13,000*g*; the pH values of the supernatants were then measured. Acidophilic bacteria were cultured after stool enrichment in a liquid medium consisting of Columbia Broth (Sigma-Aldrich) modified by the addition of (per litre) 5 g $MgSO_4$, 5 g $MgCl_2$, 2 g KCl, 2 g glucose and 1 g $CaCl_2$. The pH was adjusted to five different values: 4, 4.5, 5, 5.5 and 6, using HCl. The bacteria were then subcultured on solid medium containing the same nutritional components and pH as the culture enrichment. They were inoculated after 3, 7, 10 or 15 incubation days in liquid medium for each tested pH condition. Serial dilutions from 10^{-1} to 10^{-10} were then performed, and each dilution was plated on agar medium. Negative controls (no inoculation of the culture medium) were included for each condition.

Overall, 16 stool samples were inoculated, generating 12,968 colonies, which were tested by MALDI–TOF.

Optimization of the culturomics strategy. In parallel with this standardization period, we performed an interim analysis in order to detect gaps in our strategy. Analysing our previously published studies, we observed that 477 bacterial species previously known from the human gut were not detected. Most of these species grew in strict anaerobic (209 species, 44%) or microaerophilic (25 species, 5%) conditions, and 161 of them (33%) belonged to the phylum Proteobacteria, whereas only 46 of them (9%) belonged to the phylum Bacteroidetes (Supplementary Table 3). The classification was performed using our own database: (http://www.mediterranee-infection.com/article.php?laref=374&titre=list-of-prokaryotes-according-to-their-aerotolerant-or-obligate-anaerobic-metabolism). Focusing on these bacterial species, we designed specific strategies with the aim of cultivating these missing bacteria.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Fresh stool samples. As the human gut includes extremely oxygen-sensitive bacterial species, and because frozen storage kills some bacteria[10], we tested 28 stool samples from healthy individuals and directly cultivated these samples on collection and without storage. Each sample was directly cultivated on agar plates, enriched in blood culture bottles (BD Bactec Plus Lytic/10 Anaerobic) and followed on days 2, 5, 10 and 15. Conditions tested were anaerobic Columbia with 5% sheep blood (bioMérieux) at 37 °C with or without thermic shock (20 min/80 °C), 28 °C, anaerobic Columbia with 5% sheep blood agar (bioMérieux) and 5% rumen fluid and R-medium (ascorbic acid 1 g l^{-1}, uric acid 0.4 g l^{-1}, and glutathione 1 g l^{-1}, pH adjusted to 7.2), as previously described[23]. For this project, 59,688 colonies were tested by MALDI–TOF.

Proteobacteria. We inoculated 110 stool samples using pre-incubation in blood culture bottles (BD Bactec Plus Lytic/10 Anaerobic) supplemented with vancomycin (100 μg l^{-1}; Sigma-Aldrich). The subcultures were performed on eight different selective solid media for the growth of Proteobacteria. We inoculated onto MacConkey agar (Biokar-Diagnostics), buffered charcoal yeast extract (BD Diagnostic), eosine-methylene blue agar (Biokar-Diagnostics), Salmonella–Shigella agar (Biokar-Diagnostics), Drigalski agar (Biokar-Diagnostics), Hektoen agar (Biokar-Diagnostics), thiosulfate-citrate-bile-sucrose (BioRad) and Yersinia agar (BD Diagnostic) and incubated at 37 °C, aerobically and anaerobically. For this project, 18,036 colonies were tested by MALDI–TOF.

Microaerophilic conditions. We inoculated 198 different stool samples directly onto agar or after pre-incubation in blood culture bottles (BD Bactec Plus Lytic/10 Anaerobic bottles, BD). Fifteen different culture conditions were tested using Pylori agar (bioMérieux), Campylobacter agar (BD), Gardnerella agar (bioMérieux), 5% sheep blood agar (bioMérieux) and our own R-medium as previously described[23]. We incubated Petri dishes only in microaerophilic conditions using GENbag microaer systems (bioMérieux) or CampyGen agar (bioMérieux), except the R-medium, which was incubated aerobically at 37 °C. These culture conditions generated 41,392 colonies, which were tested by MALDI–TOF.

Halophilic bacteria. In addition, we used new culture conditions to culture halophilic prokaryotes. The culture enrichment and isolation procedures for the culture of halophilic prokaryotes were performed in a Columbia broth medium (Sigma-Aldrich), modified by adding (per litre): $MgCl_2{\cdot}6H_2O$, 5 g; $MgSO_4{\cdot}7H_2O$, 5 g; KCl, 2 g; $CaCl_2{\cdot}2H_2O$, 1 g; NaBr, 0.5 g; $NaHCO_3$, 0.5 g and 2 g of glucose. The pH was adjusted to 7.5 with 10 M NaOH before autoclaving. All additives were purchased from Sigma-Aldrich. Four concentrations of NaCl were used (100 g l^{-1}, 150 g l^{-1}, 200 g l^{-1} and 250 g l^{-1}).

A total of 215 different stool samples were tested. One gram of each stool specimen was inoculated aerobically into 100 ml of liquid medium in flasks at 37 °C while stirring at 150 r.p.m. Subcultures were inoculated after 3, 10, 15 and 30 incubation days for each culture condition. Serial dilutions from 10^{-1} to 10^{-10} were then performed in the culture medium and then plated on agar medium. Negative controls (no inoculation of the culture medium) were included for each culture condition. After three days of incubation at 37 °C, different types of colonies appeared: yellow, cream, white and clear. Red and pink colonies began to appear after the 15th day. All colonies were picked and re-streaked several times to obtain pure cultures, which were subcultured on a solid medium consisting of Colombia agar medium (Sigma-Aldrich) NaCl. The negative controls remained sterile in all culture conditions, supporting the authenticity of our data.

Detection of microcolonies. Finally, we began to focus on microcolonies detected using a magnifying glass (Leica). These microcolonies, which were not visualized with the naked eye and ranged from 100 to 300 μm, did not allow direct identification by MALDI–TOF. We subcultured these bacteria in a liquid medium (Columbia broth, Sigma-Aldrich) to allow identification by MALDI–TOF after centrifugation. Ten stool samples were inoculated and then observed using this magnifying glass for this project, generating the 9,620 colonies tested.

Duodenum and other gut samples. Most of the study was designed to explore the gut microbiota using stool samples. Nevertheless, as the small intestine microbiota are located where the nutrients are digested[24], which means there are greater difficulties in accessing samples than when using stool specimens, we analysed different levels of sampling, including duodenum samples (Supplementary Table 1). First, we tested five duodenum samples previously frozen at −80 °C. A total of 25,000 colonies were tested by MALDI–TOF. In addition, we tested samples from the different gut levels (gastric, duodenum, ileum and left and right colon) of other patients. We tested 25,048 colonies by MALDI–TOF for this project. We tested 15 culture conditions, including pre-incubation in blood culture bottles with sterile rumen fluid and sheep blood (BD Bactec Plus Lytic/10 Anaerobic), 5% sheep blood agar (bioMérieux), and incubation in both microaerophilic and anaerobic conditions, R-medium[23] and Pylori agar (bioMérieux). Overall, we tested 50,048 colonies by MALDI–TOF for this project.

Archaea. The culture of methanogenic archaea is a fastidious process, and the necessary equipment for this purpose is expensive and reserved for specialized laboratories. With this technique, we isolated seven methanogenic archaea through culturomic studies as previously described[25–27]. In addition, we propose here an affordable alternative that does not require specific equipment[17]. Indeed, a simple double culture aerobic chamber separated by a microfilter (0.2 μm) was used to grow two types of microorganism that develop in perfect symbiosis. A pure culture of *Bacteroides thetaiotaomicron* was placed in the bottom chamber to produce the hydrogen necessary for the growth of the methanogenic archaea, which was trapped in the upper chamber. A culture of *Methanobrevibacter smithii* or other hydrogenotrophic methanogenic archaea had previously been placed in the chamber. In the case presented here, the methanogenic archaea were grown aerobically on an agar medium supplemented with three antioxidants (ascorbic acid, glutathione and uric acid) and without the addition of any external gas. We subsequently cultured four other methanogenic archaeal species for the first time aerobically, and successfully isolated 13 strains of *M. smithii* and 9 strains of *Methanobrevibacter oralis* from 100 stools and 45 oral samples. This medium allows aerobic isolation and antibiotic susceptibility testing. This change allows the routine study of methanogens, which have been neglected in clinical microbiology laboratories and may be useful for biogas production. Finally, to culture halophilic archaea, we designed specific culture conditions (described in the 'Halophilic bacteria' section).

Identification methods. The colonies were identified using MALDI–TOF MS. Each deposit was covered with 2 ml of a matrix solution (saturated α-cyano acid-4-hydroxycinnamic in 50% acetonitrile and 2.5% trifluoroacetic acid). This analysis was performed using a Microflex LT system (Bruker Daltonics). For each spectrum, a maximum of 100 peaks was used and these peaks were compared with those of previous samples in the computer database of the Bruker Base and our homemade database, including the spectra of the bacterial species identified in previous works[28,29]. An isolate was labelled as correctly identified at the species level when at least one of the colonies' spectra had a score ≥1.9 and another of the colonies' spectra had a score ≥1.7 (refs 28,29).

Protein profiles are regularly updated based on the results of clinical diagnoses and on new species providing new spectra. If, after three attempts, the species could not be accurately identified by MALDI–TOF, the isolate was identified by 16S rRNA sequencing as previously described. A threshold similarity value of >98.7% was chosen for identification at the species level. Below this value, a new species was suspected, and the isolate was described using taxonogenomics[30].

Classification of the prokaryotes species cultured. We used our own online prokaryotic repertoire[13] (http://hpr.mediterranee-infection.com/arkotheque/client/ihu_bacteries/recherche/index.php) to classify all isolated prokaryotes into four categories: new prokaryote species, previously known prokaryote species in the human gut, known species from the environment but first isolated in humans, and known species from humans but first isolated in the human gut. Briefly, to complete the recent work identifying all the prokaryotes isolated in humans[13], we examined methods by conducting a literature search, which included PubMed and books on infectious diseases. We examined the Medical Subject Headings (MeSH) indexing provided by Medline for bacteria isolated from the human gut and we then established two different queries to automatically obtain all articles indexed by Medline dealing with human gut isolation sites. These queries were applied to all bacterial species previously isolated from humans as previously described, and we obtained one or more articles for each species, confirming that the bacterium had been isolated from the human gut[13].

International deposition of the strains, 16S rRNA accession numbers and genome sequencing accession number. Most of the strains isolated in this study were deposited in CSUR (WDCM 875) and are easily available at http://www.mediterranee-infection.com/article.php?laref=14&titre=collection-de-souches&PHPSESSID=cncregk417fl97gheb8k7u7t07 (Supplementary Tables 4a and b). All the new prokaryote species were deposited into two international collections: CSUR and DSMZ (Supplementary Table 5). Importantly, among the 247 new prokaryotes species (197 in the present study and 50 in previous studies), we failed to subculture 9 species that were not deposited, of which 5 were nevertheless genome sequenced. Apart from these species, all CSUR accession numbers are available in Supplementary Table 5. Among these viable new species, 189 already have a DSMZ number. For the other 49 species, the accession number is not yet assigned but the strain is deposited. The 16S rRNA accession numbers of the 247 new prokaryotes species are available in Supplementary Table 5, along with the accession number of the known species needing 16S rRNA amplification and sequencing for identification (Supplementary Table 14). Finally, the 168 draft genomes used for our analysis have already been deposited with an available GenBank accession number (Supplementary Table 5) and all other genome sequencing is still in progress, as the culturomics are still running in our laboratory.

New prokaryotes. All new prokaryote species have been or will be comprehensively described by taxonogenomics, including their metabolic properties, MALDI–TOF spectra and genome sequencing[30]. Among these 247 new prokaryote species, 95 have already been published (PMID available in Supplementary Table 5), including 70 full descriptions and 25 'new species announcements'. In addition, 20 are under

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

review and the 132 others are ongoing (Supplementary Table 5). This includes 37 bacterial species already officially recognized (as detailed in Supplementary Table 5). All were sequenced successively with a paired-end strategy for high-throughput pyrosequencing on the 454-Titanium instrument from 2011 to 2013 and using MiSeq Technology (Illumina) with the mate pair strategy since 2013.

Metagenome sequencing. Total DNA was extracted from the samples using a method modified from the Qiagen stool procedure (QIAamp DNA Stool Mini Kit). For the first 24 metagenomes, we used GS FLX Titanium (Roche Applied Science). Primers were designed to produce an amplicon length (576 bp) that was approximately equivalent to the average length of reads produced by GS FLX Titanium (Roche Applied Science), as previously described. The primer pairs commonly used for gut microbiota were assessed *in silico* for sensitivity to sequences from all phyla of bacteria in the complete Ribosomal Database Project (RDP) database. Based on this assessment, the bacterial primers 917F and 1391R were selected. The V6 region of 16S rRNA was pyrosequenced with unidirectional sequencing from the forward primer with one-half of a GS FLX Titanium PicoTiterPlate Kit 70×75 per patient with the GS Titanium Sequencing Kit XLR70 after clonal amplification with the GS FLX Titanium LV emPCR Kit (Lib-L).

Sixty other metagenomes were sequenced for 16S rRNA sequencing using MiSeq technology. PCR-amplified templates of genomic DNA were produced using the surrounding conserved regions' V3–V4 primers with overhang adapters (FwOvAd_341F TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGG NGGCWGCAG; ReOvAd_785RGTCTCGTGGGCTCGGAGATG TGTATAAGA GACAGGACTACHVGGGTATCTAATCC). Samples were amplified individually for the 16S V3–V4 regions by Phusion High Fidelity DNA Polymerase (Thermo Fisher Scientific) and visualized on the Caliper Labchip II device (Illumina) by a DNA 1K LabChip at 561 bp. Phusion High Fidelity DNA Polymerase was chosen for PCR amplifications in this biodiversity approach and deep sequencing: a thermostable DNA polymerase characterized by the greatest accuracy, robust reactions and high tolerance for inhibitors, and finally by an error rate that is approximately 50-fold lower than that of DNA polymerase and sixfold lower than that of Pfu DNA polymerase. After purification on Ampure beads (Thermo Fisher Scientific), the concentrations were measured using high-sensitivity Qbit technology (Thermo Fisher Scientific). Using a subsequent limited-cycle PCR on 1 ng of each PCR product, Illumina sequencing adapters and dual-index barcodes were added to each amplicon. After purification on Ampure beads, the libraries were then normalized according to the Nextera XT (Illumina) protocol. The 96 multiplexed samples were pooled into a single library for sequencing on the MiSeq. The pooled library containing indexed amplicons was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and paired-end sequencing with dual index reads of 2 × 250 bp were performed in a single 39-hour run. On the instrument, the global cluster density and the global passed filter per flow cell were generated. The MiSeq Reporter software (Illumina) determined the percentage indexed and the clusters passing the filter for each amplicon or library. The raw data were configured in fasta files for R1 and R2 reads.

Genome sequencing. The genomes were sequenced using, successively, two high-throughput NGS technologies: Roche 454 and MiSeq Technology (Illumina) with paired-end application. Each project on the 454 sequencing technology was loaded on a quarter region of the GS Titanium PicoTiterPlate and sequenced with the GS FLX Titanium Sequencer (Roche). For the construction of the 454 library, 5 μg DNA was mechanically fragmented on the Covaris device (KBioScience-LGC Genomics) through miniTUBE-Red 5Kb. The DNA fragmentation was visualized through the Agilent 2100 BioAnalyser on a DNA LabChip7500. Circularization and fragmentation were performed on 100 ng. The library was then quantified on Quant-it Ribogreen kit (Invitrogen) using a Genios Tecan fluorometer. The library was clonally amplified at 0.5 and 1 cpb in 2 emPCR reactions according to the conditions for the GS Titanium SV emPCR Kit (Lib-L) v2 (Roche). These two enriched clonal amplifications were loaded onto the GS Titanium PicoTiterPlates and sequenced with the GS Titanium Sequencing Kit XLR70. The run was performed overnight and then analysed on the cluster through gsRunBrowser and gsAssembler_Roche. Sequences obtained with Roche were assembled on gsAssembler with 90% identity and 40 bp of overlap. The library for Illumina was prepared using the Mate Pair technology. To improve the assembly, the second application in was sometimes performed with paired ends. The paired-end and the mate-pair strategies were barcoded in order to be mixed, respectively, with 11 other genomic projects prepared with the Nextera XT DNA sample prep kit (Illumina) and 11 others projects with the Nextera Mate Pair sample prep kit (Illumina). The DNA was quantified by a Qbit assay with high-sensitivity kit (Life Technologies). In the first approach, the mate pair library was prepared with 1.5 μg genomic DNA using the Nextera mate pair Illumina guide. The genomic DNA sample was simultaneously fragmented and tagged with a mate-pair junction adapter. The profile of the fragmentation was validated on an Agilent 2100 Bioanalyzer (Agilent Technologies) with a DNA 7500 LabChip. The DNA fragments, which ranged in size, had an optimal size of 5 kb. No size selection was performed, and 600 ng of 'tagmented' fragments measured on the Qbit assay with the high-sensitivity kit were circularized. The circularized DNA was mechanically sheared to small fragments, with optimal fragments being 700 bp, on a Covaris S2 device in microtubes. The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies). The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. To prepare the paired-end library, 1 ng of genome as input was required. DNA was fragmented and tagged during the tagmentation step, with an optimal size distribution at 1 kb. Limited-cycle PCR amplification (12 cycles) completed the tag adapters and introduced dual-index barcodes. After purification on Ampure XP beads (Beckman Coulter), the library was normalized and loaded onto the reagent cartridge and then onto the instrument along with the flow cell. For the 2 Illumina applications, automated cluster generation and paired-end sequencing with index reads of 2 × 250 bp were performed in single 39-hour runs.

ORFans identification. Open reading frames (ORFs) were predicted using Prodigal with default parameters for each of the bacterial genomes. However, the predicted ORFs were excluded if they spanned a sequencing gap region. The predicted bacterial sequences were searched against the non-redundant protein sequence (NR) database (59,642,736 sequences, available from NCBI in 2015) using BLASTP. ORFans were identified if their BLASTP E-value was lower than 1e-03 for an alignment length greater than 80 amino acids. We used an E-value of 1e-05 if the alignment length was <80 amino acids. These threshold parameters have been used in previous studies to define ORFans (refs 12–14). The 168 genomes considered in this study are listed in Supplementary Table 7. These genomes represent 615.99 Mb and contain a total of 19,980 ORFans. Some of the ORFans from 30 genomes were calculated in a previous study[4] with the non-redundant protein sequence database containing 14,124,377 sequences available from NCBI in June 2011.

Metagenomic 16S sequences. We collected 325 runs of metagenomic 16S rRNA sequences available in the HMP data sets that correspond to stool samples from healthy human subjects. All samples were submitted to Illumina deep sequencing, resulting in 761,123 Mo per sample on average, and a total of 5,970,465 high-quality sequencing reads after trimming. These trimmed data sets were filtered using CLC Genomics Workbench 7.5, and reads shorter than 100 bp were discarded. We performed an alignment of 247 16S rRNA sequences against the 5,577,630 reads remaining using BLASTN. We used a 1e-03 e-value, 100% coverage and 98.7% cutoff, corresponding to the threshold for defining a species, as previously described. Finally, we reported the total number of aligned reads for each 16S rRNA sequence (Supplementary Table 8).

We collected the sequences of the 3,871,657 gene non-redundant gene catalogue from the 396 human gut microbiome samples (https://www.cbs.dtu.dk/projects/CAG/)[15]. We performed an alignment of 247 16S rRNA sequences against the 3,871,657 gene non-redundant gene catalogue using BLASTN with a threshold of 1e-03 e-value, 100% coverage and 98.7% cutoff. The new species identified in these data are reported in Supplementary Table 9. We collected the raw data sets of 239 runs deposited at EBI (ERP012217)[16]. We used the PEAR software (PMID 24142950) for merging raw Illumina paired-end reads using default parameters. We performed an alignment of 247 16S rRNA sequences against the 265,864,518 merged reads using BLASTN. We used a 1e-03 e-value, 100% coverage and 98.7% cutoff. The list of the new species identified in these data is included in Supplementary Table 9.

Whole metagenomic shotgun sequences. We collected the contigs/scaffolds from the assembly of 148 runs available in the HMP data sets. The initial reads of these samples were assembled using SOAPdenovo v.1.04 (PMID 23587118). These assemblies correspond to stool samples from healthy human subjects and generated 13,984,809 contigs/scaffolds with a minimum length of 200 bp and a maximum length of 371,412 bp. We aligned the 19,980 ORFans found previously against these data sets using BLASTN. We used a 1e-05 e-value, 80% coverage and 80% identity cutoff. Finally, we reported the total number of unique aligned ORFans for each species (Supplementary Table 8).

Study of the gaps in metagenomics. The raw fastq files of paired-end reads from an Illumina Miseq of 84 metagenomes analysed concomitantly by culturomics were filtered and analysed in the following steps (accession no. PRJEB13171).

Data processing: filtering the reads, dereplication and clustering. The paired-end reads of the corresponding raw fastq files were assembled into contigs using Pandaseq[31]. The high-quality sequences were then selected for the next steps of analysis by considering only those sequences that contained both primers (forward and reverse). In the following filtering steps, the sequences containing N were removed. Sequences with length shorter than 200 nt were removed, and sequences longer than 500 nt were trimmed. Both forward and reverse primers were also removed from each of the sequences. An additional filtering step was applied to remove the chimaeric sequences using UCHIME (ref. 32) of USEARCH (ref. 33). The filtering steps were performed using the QIIME pipeline[34]. Strict dereplication (clustering of duplicate sequences) was performed on the filtered sequences, and they were then sorted by decreasing number of abundance[35–37]. For each metagenome, the clustering of OTUs was performed with 97% identity. Total OTUs from the 84 metagenomes (Supplementary Table 10) clustered with 93% identity.

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

Building reference databases. We downloaded the Silva SSU and LSU database1 and release 123 from the Silva website and, from this, a local database of predicted amplicon sequences was built by extracting the sequences containing both primers. Finally, we had our local reference database containing a total of 536,714 well-annotated sequences separated into two subdatabases according to their gut or non-gut origin. We created four other databases containing 16S rRNA of new species sequences and species isolated by culturomics separated into three groups (human gut, non-human gut, and human not reported in gut). The new species database contains 247 sequences, the human gut species database 374 sequences, the non-human gut species database 256 sequences and the human species not reported in gut database 237 sequences.

Taxonomic assignments. For taxonomic assignments, we applied at least 20 reads per OTU. The OTUs were then searched against each database using BLASTN (ref. 38). The best match of ≥97% identity and 100% coverage for each of the OTUs was extracted from the reference database, and taxonomy was assigned up to the species level. Finally, we counted the number of OTUs assigned to unique species.

Data availability. The GenBank accession numbers for the sequences of the16SrRNA genes of the new bacterial species as well as their accession numbers in both Collection de Souches de l'Unité des Rickettsies (CSUR, WDCM 875) and the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) are listed in Supplementary Table 5. Sequencing metagenomics data have been deposited in NCBI under Bioproject PRJEB13171.

Received 20 April 2016; accepted 14 September 2016; published 7 November 2016

References

1. Lagier, J. C., Million, M., Hugon, P., Armougom, F. & Raoult, D. Human gut microbiota: repertoire and variations. *Front. Cell Infect. Microbiol.* **2,** 136 (2012).
2. Lagier, J. C. *et al.* The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin Microbiol Rev.* **28,** 237–264 (2015).
3. Pfleiderer, A. *et al.* Culturomics identified 11 new bacterial species from a single anorexia nervosa stool sample. *Eur. J. Clin. Microbiol. Infect. Dis.* **32,** 1471–1481 (2013).
4. Lagier, J. C. *et al.* Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin. Microbiol. Infect.* **18,** 1185–1193 (2012).
5. Dubourg, G. *et al.* Culturomics and pyrosequencing evidence of the reduction in gut microbiota diversity in patients with broad-spectrum antibiotics. *Int. J. Antimicrob. Agents* **44,** 117–124 (2014).
6. Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444,** 1022–1023 (2006).
7. Ley, R. E. *et al.* Obesity alters gut microbial ecology. *Proc. Natl Acad. Sci. USA* **102,** 11070–11075 (2005).
8. Gill, S. R. *et al.* Metagenomic analysis of the human distal gut microbiome. *Science* **312,** 1355–1359 (2006).
9. Rinke, C. *et al.* Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499,** 431–437 (2013).
10. Lagier, J. C. *et al.* Current and past strategies for bacterial culture in clinical microbiology. *Clin. Microbiol. Rev.* **28,** 208–236 (2015).
11. Vetizou, M. *et al.* Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **350,** 1079–1084 (2015).
12. Cassir, N. *et al. Clostridium butyricum* strains and dysbiosis linked to necrotizing enterocolitis in preterm neonates. *Clin. Infect. Dis.* **61,** 1107–1115.
13. Hugon, P. *et al.* A comprehensive repertoire of prokaryotic species identified in human beings. *Lancet Infect. Dis.* **15,** 1211–1219 (2015).
14. The Human Microbiome Project Consortium. A framework for human microbiome research. *Nature* **486,** 215–221 (2012).
15. Browne, H. P. *et al.* Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* **533,** 543–546 (2016).
16. Nielsen, H. B. *et al.* Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* **32,** 822–828 (2014).
17. Khelaifia, S. *et al.* Aerobic culture of methanogenic archaea without an external source of hydrogen. *Eur. J. Clin. Microbiol. Infect. Dis.* **35,** 985–991 (2016).
18. Rettedal, E. A., Gumpert, H. & Sommer, M. O. Cultivation-based multiplex phenotyping of human gut microbiota allows targeted recovery of previously uncultured bacteria. *Nat. Commun.* **5,** 4714 (2014).
19. Hiergeist, A., Gläsner, J., Reischl, U. & Gessner, A. Analyses of intestinal microbiota: culture versus sequencing. *ILAR J.* **56,** 228–240 (2015).
20. Rajilic-Stojanovic, M. & de Vos, W. M. The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol. Rev.* **38,** 996–1047 (2014).
21. Byrd, A. L. & Segre, J. A. Infectious disease. Adapting Koch's postulates. *Science* **351,** 224–226 (2016).
22. Samb-Ba, B. *et al.* MALDI–TOF identification of the human gut microbiome in people with and without diarrhea in Senegal. *PLoS ONE* **9,** e87419 (2014).
23. Dione, N., Khelaifia, S., La Scola, B., Lagier, J.C. & Raoult D. A quasi-universal medium to break the aerobic/anaerobic bacterial culture dichotomy in clinical microbiology. *Clin. Microbiol. Infect.* **22,** 53–58 (2016).
24. Raoult, D. & Henrissat, B. Are stool samples suitable for studying the link between gut microbiota and obesity? *Eur. J. Epidemiol.* **29,** 307–309 (2014).
25. Khelaifia, S., Raoult, D. & Drancourt, M. A versatile medium for cultivating methanogenic archaea. *PLoS ONE* **8,** e61563 (2013).
26. Khelaifia, S. *et al.* Draft genome sequence of a human-associated isolate of methanobrevibacter arboriphilicus, the lowest-G+C-content archaeon. *Genome Announc.* **2,** e01181 (2014).
27. Dridi, B., Fardeau, M.-L., Ollivier, B., Raoult, D. & Drancourt, M. *Methanomassiliicoccus luminyensis* gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **62,** 1902–1907 (2012).
28. Seng, P. *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.* **51,** 2182–2194 (2013).
29. Seng, P. *et al.* Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin. Infect. Dis.* **49,** 543–551 (2009).
30. Ramasamy, D. *et al.* A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. *Int. J. Syst. Evol. Microbiol.* **64,** 384–391 (2014).
31. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for Illumina sequences. *BMC Bioinformatics* **13,** 31 (2012).
32. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27,** 2194–2200 (2011).
33. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26,** 2460–2461 (2010).
34. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7,** 335–336 (2010).
35. Stoeck, T. *et al.* Massively parallel tag sequencing reveals the complexity of anaerobic marine protistan communities. *BMC Biol.* **7,** 72–77 (2009).
36. Mondani, L. *et al. Microbacterium lemovicicum* sp. nov., a bacterium isolated from a natural uranium-rich soil. *Int. J. Syst. Evol. Microbiol.* **63,** 2600–2606 (2013).
37. Boissiere, A. *et al.* Midgut microbiota of the malaria mosquito vector *Anopheles gambiae* and interactions with *Plasmodium falciparum* infection. *PLoS Pathog.* **8,** e1002742 (2012).
38. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410 (1990).

Acknowledgements

The authors thank R. Valero, A.A. Jiman-Fatani, B. Ali Diallo, J.-B. Lekana-Douki, B. Senghor, A. Derand, L. Gandois, F. Tanguy, S. Strouk, C. Tamet, F. Lunet, M. Kaddouri, L. Ayoub, L. Frégère, N. Garrigou, A. Pfleiderer, A. Farina and V. Ligonnet for technical support. This work was funded by IHU Méditerranée Infection as a part of a Foundation Louis D grant and by the Deanship of Scientific Research (DSR), King Abdulaziz University, under grant no. 1–141/1433 HiCi.

Author contributions

D.R. conceived and designed the experiments. J.-C.L., S.K., M.T.A., S.N., N.D., P.H., A.C., F.C., S.I.T., E.H.S., G.Dub., G.Dur., G.M., E.G. A.T., S.B., D.B., N.C., F.B., J.D., M.Ma., D.R., M.B., N.P.M.D.N., N.M.D.B., C.V., D.M., K.D., M.Mi., C.R., J.M.R., B.L.S., P.-E.F. and A.L. performed the experiments. D.M., J.A., E.I.A., F.B., M.Y., A.D., C.S., F.D. and V.V. contributed materials/analysis tools. J.-C.L., A.C., A.L. and D.R. analysed the data. J.-C.L., A.L. and D.R. wrote the manuscript. All authors read and approved the final manuscript.

Additional information

Supplementary information is available for this paper. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.R.

Competing interests

The authors declare no competing financial interests.

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© 2016 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

CONCLUSION ET PERSPECTIVES

Dans l'ensemble, le microbiote vaginal agit sur la santé des femmes et sa composition varie en fonction de nombreux facteurs. Certaines variations de cette flore vaginale sont normales, n'ayant pas *a priori* d'impact sur la santé. La vaginose bactérienne est quant à elle une dysbiose vaginale pouvant engendrer des troubles gynéco-obstétriques. Elle se manifeste par des leucorrhées nauséabondes, particulièrement gênantes pour les patientes, mais cette dysbiose peut être aussi asymptomatique. La vaginose bactérienne est une pathologie commune de la femme. Sa prévalence dépend, toutefois, des populations étudiées ainsi des pratiques sexuelles. Au cours du temps, différents outils diagnostiques ont été développés. Son diagnostic a reposé pendant longtemps exclusivement sur les critères Amsel puis sur le score de Nugent. La biologie moléculaire a permis d'avoir une approche plus rationnelle du diagnostic ; notamment, la PCR quantitative en temps réel ciblant deux bactéries *Atopobium vaginae* et *Gardnerella vaginalis*, dont la forte augmentation dans la flore vaginale est un excellent marqueur de vaginose bactérienne. Quel que soit le traitement utilisé, antibiothérapie à base d'imidazolés ou clindamycine associée ou non à des ovules de *Lactobacillus*, la thérapie échoue dans la majorité des cas.

Ce travail sur le microbiote vaginal, fut non seulement le premier à étudier les flores vaginales des femmes françaises et sénégalaises mais aussi à utiliser la culturomique comme moyen de caractérisation de cette flore. Ainsi, comparer aux femmes pubères en bonne santé, les personnes souffrant de vaginose bactérienne présentent un microbiote vaginal très complexe et dynamique. La composition du microbiote vaginal lors de la vaginose bactérienne est donc différente de celle du microbiote normal. Au cours de nos travaux, nous avons aussi constaté que cette distinction comprend à la fois une expansion de la diversité bactérienne et une prolifération et/ou apparition de certains taxons surtout de bactéries anaérobies et parfois fastidieuses. La caractérisation de la flore en utilisant la culturomique et la métagénomique,

nous a permis d'une part de démontrer la complémentarité entre ces deux techniques avec seulement 26% de recouvrement. D'autre part avec ces deux approches, nous avons détecté une communauté bactérienne diversifiée et variée peuplant le tractus vaginal humain mais aussi montré que les patientes atteintes de la vaginose bactérienne présentent une absence de *Proteobacteria* et des lactobacilles, espèces protectrices et indispensables au maintien de l'équilibre vaginal. Chez ces patientes, nous avons aussi trouvé un taux élevé d'espèces de l'ordre des *Bacteroidales* (phylum *Bacteroidetes*) et de la famille *Leptotrichiaceae* (phylum *Fusobacteria*), dont de nombreuses espèces pathogènes.

Au cours de cette thèse de doctorat, la stratégie de la culturomique associée à la spectrométrie de masse MALDI-TOF et si nécessaire au séquençage du gène codant pour l'ARNr 16S nous a permis de cultiver et d'identifier un large spectre bactérien. Le répertoire des bactéries cultivées dans le microbiote vaginal a été augmenté de plus de 150 espèces isolées pour la première fois dans le tractus vaginal renfermant 27 nouvelles espèces dont 3 associées à la vaginose bactérienne, *Peptoniphilus* sp. DNF00840*, Megasphaera* sp. BV3C16-1 et *Atopobium* sp. S3MV26, auparavant identifiées uniquement à l'aide de méthodes moléculaires. Ces résultats prouvent que certaines espèces non-cultivables identifiées par biologie moléculaire sont en fait susceptible d'être cultivées. Sur les 27 nouvelles espèces isolées : 18 étaient cultivées dans la flore de vaginose bactérienne et 9 dans la flore normale. Ces nouvelles espèces sont majoritairement des anaérobies strictes. Douze ont été décrites parmi lesquelles deux sont déjà officiellement validées (validation liste n°184 *: Corynebacterium fournierii* et *Murdochiella vaginalis*).

La persévérance dans la culturomique, en augmentant le nombre d'échantillons vaginaux analysés et en diversifiant les zones d'étude géographique, mais aussi en utilisant de nouveaux milieux de culture imitant l'environnement vaginal, est nécessaire afin de pouvoir isoler cette part majeure du microbiote vaginal identifiée jusqu'à présent exclusivement par

biologie moléculaire. L'obtention d'isolats permettra d'explorer *in vitro* les compétitions entre les bactéries du microbiote vaginal mais pourra servir également de matière première pour développer un traitement par bactériothérapie en proposant un cocktail de bactéries clés pour prévenir ou traiter la vaginose bactérienne.

REFERENCES :

1. **Ma B, Forney LJ, Ravel J**. Vaginal Microbiome: Rethinking Health and Disease. *Annual Review of Microbiology* 2012;66:371–389.

2. **Franasiak JM, Scott RT**. Reproductive tract microbiome in assisted reproductive technologies. *Fertility and Sterility* 2015;104:1364–1371.

3. **Green KA, Zarek SM, Catherino WH**. Gynecologic health and disease in relation to the microbiome of the female reproductive tract. *Fertility and Sterility* 2015;104:1351–1357.

4. **Lepargneur JP, Rousseau V**. [Protective role of the Doderleïn flora]. *J Gynecol Obstet Biol Reprod (Paris)* 2002;31:485–494.

5. **Iebba V, Totino V, Gagliardi A, Santangelo F, Cacciotti F, *et al.*** Eubiosis and dysbiosis: the two sides of the microbiota. 12.

6. **NIH HMP Working Group, Peterson J, Garges S, Giovanni M, McInnes P, *et al.*** The NIH Human Microbiome Project. *Genome Res* 2009;19:2317–2323.

7. **Lamont R, Sobel J, Akins R, Hassan S, Chaiworapongsa T, *et al.*** The vaginal microbiome: new information about genital tract flora using molecular based techniques: Vaginal microbiome using molecular tools. *BJOG: An International Journal of Obstetrics & Gynaecology* 2011;118:533–549.

8. **Lopes dos Santos Santiago G, Tency I, Verstraelen H, Verhelst R, Trog M, *et al.*** Longitudinal qPCR Study of the Dynamics of L. crispatus, L. iners, A. vaginae, (Sialidase Positive) G. vaginalis, and P. bivia in the Vagina. *PLoS ONE* 2012;7:e45281.

9. **Machado A, Almeida C, Salgueiro D, Henriques A, Vaneechoutte M, *et al.*** Fluorescence in situ Hybridization method using Peptide Nucleic Acid probes for rapid detection of Lactobacillus and Gardnerella spp. *BMC Microbiol* 2013;13:82.

10. **Human Microbiome Project Consortium**. Structure, function and diversity of the healthy human microbiome. *Nature* 2012;486:207–214.

11. **Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, *et al.*** The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome* 2014;2:4.

12. **Belay N, Mukhopadhyay B, De Macario EC, Galask R, Daniels L**. Methanogenic bacteria in human vaginal samples. *Journal of clinical microbiology* 1990;28:1666–1668.

13. **Fichorova R, Fraga J, Rappelli P, Fiori PL**. Trichomonas vaginalis infection in symbiosis with Trichomonasvirus and Mycoplasma. *Research in Microbiology* 2017;168:882–891.

14. **Lee JE, Lee S, Lee H, Song YM, Lee K, *et al.*** Association of the Vaginal Microbiota with Human Papillomavirus Infection in a Korean Twin Cohort. *PLoS One*;8. Epub ahead of print 22 May 2013. DOI: 10.1371/journal.pone.0063514.

15. **Pal K, Roy S, Behera B, Kumar N, Sagiri S, *et al.*** Bacterial vaginosis: Etiology and modalities of treatment-A brief note. *Journal of Pharmacy and Bioallied Sciences* 2011;3:496.

16. **Danielsson D, Teigen PK, Moi H**. The genital econiche: focus on microbiota and bacterial vaginosis: Genital econiche. *Annals of the New York Academy of Sciences* 2011;1230:48–58.

17. **Africa C, Nel J, Stemmet M**. Anaerobes and Bacterial Vaginosis in Pregnancy: Virulence Factors Contributing to Vaginal Colonisation. *IJERPH* 2014;11:6979–7000.

18. **Huang B, Fettweis JM, Brooks JP, Jefferson KK, Buck GA**. The Changing Landscape of the Vaginal Microbiome. *Clinics in Laboratory Medicine* 2014;34:747–761.

19. **Mendling W**. Vaginal Microbiota. In: Schwiertz A (editor). *Microbiota of the Human Body*.

Cham: Springer International Publishing. pp. 83–93.

20. **Li J, McCormick J, Bocking A, Reid G**. Importance of Vaginal Microbes in Reproductive Health. *Reproductive Sciences* 2012;19:235–242.

21. **Tachedjian G, O'Hanlon DE, Ravel J**. The implausible "in vivo" role of hydrogen peroxide as an antimicrobial factor produced by vaginal microbiota. *Microbiome*;6. Epub ahead of print 6 February 2018. DOI: 10.1186/s40168-018-0418-3.

22. **Nelson TM, Borgogna J-LC, Brotman RM, Ravel J, Walk ST, *et al.*** Vaginal biogenic amines: biomarkers of bacterial vaginosis or precursors to vaginal dysbiosis? *Frontiers in Physiology*;6. Epub ahead of print 29 September 2015. DOI: 10.3389/fphys.2015.00253.

23. **Borges S, Silva J, Teixeira P**. The role of lactobacilli and probiotics in maintaining vaginal health. *Archives of Gynecology and Obstetrics* 2014;289:479–489.

24. **Witkin SS, Linhares IM, Giraldo P**. Bacterial flora of the female genital tract: function and immune regulation. *Best Practice & Research Clinical Obstetrics & Gynaecology* 2007;21:347–354.

25. **Muzny CA, Schwebke JR**. Pathogenesis of Bacterial Vaginosis: Discussion of Current Hypotheses. *Journal of Infectious Diseases* 2016;214:S1–S5.

26. **Bilardi J, Walker S, Mooney-Somers J, Temple-Smith M, McNair R, *et al.*** Women's Views and Experiences of the Triggers for Onset of Bacterial Vaginosis and Exacerbating Factors Associated with Recurrence. *PLOS ONE* 2016;11:e0150272.

27. **Bradshaw CS, Brotman RM**. Making inroads into improving treatment of bacterial vaginosis – striving for long-term cure. *BMC Infectious Diseases*;15. Epub ahead of print December 2015. DOI: 10.1186/s12879-015-1027-4.

28. **Chico RM, Mayaud P, Ariti C, Mabey D, Ronsmans C, *et al.*** Prevalence of malaria and sexually transmitted and reproductive tract infections in pregnancy in sub-Saharan Africa: a systematic review. *JAMA* 2012;307:2079–2086.

29. **Martin DH, Marrazzo JM**. The Vaginal Microbiome: Current Understanding and Future Directions. *Journal of Infectious Diseases* 2016;214:S36–S41.

30. **Afolabi BB, Moses OE, Oduyebo OO**. Bacterial Vaginosis and Pregnancy Outcome in Lagos, Nigeria. *Open Forum Infectious Diseases* 2016;3:ofw030.

31. **McClelland RS, Lingappa JR, Srinivasan S, Kinuthia J, John-Stewart GC, *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*. Epub ahead of print 25 January 2018. DOI: 10.1016/S1473-3099(18)30058-6.

32. **Bradshaw CS, Tabrizi SN, Fairley CK, Morton AN, Rudland E, *et al.*** The association of Atopobium vaginae and Gardnerella vaginalis with bacterial vaginosis and recurrence after oral metronidazole therapy. *The Journal of infectious diseases* 2006;194:828–836.

33. **Bretelle F, Rozenberg P, Pascal A, Favre R, Bohec C, *et al.*** High Atopobium vaginae and Gardnerella vaginalis Vaginal Loads Are Associated With Preterm Birth. *Clinical Infectious Diseases* 2015;60:860–867.

34. **Mårdh PA**. [The definition and epidemiology of bacterial vaginosis]. *Rev Fr Gynecol Obstet* 1993;88:195–197.

35. **Marrazzo JM, Hillier SL**. Bacterial Vaginosis. In: *Sexually Transmitted Diseases*. Elsevier. pp. 463–498.

36. **Donders GGG**. Definition and classification of abnormal vaginal flora. *Best Pract Res Clin Obstet Gynaecol* 2007;21:355–373.

37. **Fredricks DN, Fiedler TL, Marrazzo JM**. Molecular identification of bacteria associated with bacterial vaginosis. *New England Journal of Medicine* 2005;353:1899–1911.

38. **Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, *et al.*** Microbial culturomics: paradigm shift in the human gut microbiome study. *Clinical Microbiology and Infection* 2012;18:1185–1193.

39. **Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, *et al.*** The Rebirth of Culture in Microbiology through the Example of Culturomics To Study Human Gut Microbiota. *Clinical Microbiology Reviews* 2015;28:237–264.

40. **Fournier PE, Drancourt M**. New Microbes New Infections promotes modern prokaryotic taxonomy: a new section "TaxonoGenomics: new genomes of microorganisms in humans". *New Microbes and New Infections* 2015;7:48–49.

41. **Fournier PE, Lagier JC, Dubourg G, Raoult D**. From culturomics to taxonomogenomics: A need to change the taxonomy of prokaryotes in clinical microbiology. *Anaerobe* 2015;36:73–78.

42. **Fournier PE, Raoult D, Drancourt M**. New Species Announcement: a new format to prompt the description of new human microbial species. *New Microbes and New Infections* 2017;15:136–137.

43. **Bagnall P, Rizzolo D**. Bacterial vaginosis: A practical review. *Journal of the American Academy of Physician Assistants* 2017;30:15–21.

44. **Onderdonk AB, Delaney ML, Fichorova RN**. The Human Microbiome during Bacterial Vaginosis. *Clinical Microbiology Reviews* 2016;29:223–238.

45. **Bitew A, Abebaw Y, Bekele D, Mihret A**. Prevalence of Bacterial Vaginosis and Associated Risk Factors among Women Complaining of Genital Tract Infection. *International Journal of Microbiology* 2017;2017:1–8.

46. **Cherpes TL, Hillier SL, Meyn LA, Busch JL, Krohn MA**. A Delicate Balance: Risk Factors for Acquisition of Bacterial Vaginosis Include Sexual Activity, Absence of Hydrogen Peroxide-Producing Lactobacilli, Black Race, and Positive Herpes Simplex Virus Type 2 Serology: *Sexually Transmitted Diseases* 2008;35:78–83.

47. **Koumans EH, Sternberg M, Bruce C, McQuillan G, Kendrick J, *et al.*** The Prevalence of Bacterial Vaginosis in the United States, 2001-2004; Associations With Symptoms, Sexual Behaviors, and Reproductive Health: *Sexually Transmitted Diseases* 2007;34:864–869.

48. **Forcey DS, Vodstrcil LA, Hocking JS, Fairley CK, Law M, *et al.*** Factors associated with bacterial vaginosis among women who have sex with women: a systematic review. *PloS one* 2015;10:e0141905.

49. **Garg K, Khare A, Bansal R, Sharma S, Chaudhary N**. Effects of Different Contraceptive Methods on Cervico-Vaginal Cytology. *J Clin Diagn Res* 2017;11:EC09-EC11.

50. **Aslan E, Bechelaghem N**. To 'douche' or not to 'douche': hygiene habits may have detrimental effects on vaginal microbiota. *Journal of Obstetrics and Gynaecology* 2018;1–4.

51. **Nelson TM, Borgogna JC, Michalek RD, Roberts DW, Rath JM, *et al.*** Cigarette smoking is associated with an altered vaginal tract metabolomic profile. *Scientific Reports*;8. Epub ahead of print December 2018. DOI: 10.1038/s41598-017-14943-3.

52. **Amsel R, Totten PA, Spiegel CA, Chen KCS, Eschenbach D, *et al.*** Nonspecific vaginitis. *The American Journal of Medicine* 1983;74:14–22.

53. **Nugent RP, Krohn MA, Hillier SL**. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *Journal of clinical microbiology* 1991;29:297–301.

54. **Shipitsyna E, Roos A, Datcu R, Hallén A, Fredlund H, *et al.*** Composition of the Vaginal

Microbiota in Women of Reproductive Age – Sensitive and Specific Molecular Diagnosis of Bacterial Vaginosis Is Possible? *PLoS ONE* 2013;8:e60670.

55. **Menard JP, Fenollar F, Henry M, Bretelle F, Raoult D**. Molecular quantification of Gardnerella vaginalis and Atopobium vaginae loads to predict bacterial vaginosis. *Clin Infect Dis* 2008;47:33–43.

56. **Datcu R, Gesink D, Mulvad G, Montgomery-Andersen R, Rink E, *et al.*** Vaginal microbiome in women from Greenland assessed by microscopy and quantitative PCR. *BMC Infectious Diseases* 2013;13:480.

57. Bacterial Vaginosis - 2015 STD Treatment Guidelines. https://www.cdc.gov/std/tg2015/bv.htm (accessed 15 April 2018).

58. **Dickey LJ, Nailor MD, Sobel JD**. Guidelines for the treatment of bacterial vaginosis: focus on tinidazole. *Ther Clin Risk Manag* 2009;5:485–489.

59. **Machado D, Castro J, Palmeira-de-Oliveira A, Martinez-de-Oliveira J, Cerca N**. Bacterial Vaginosis Biofilms: Challenges to Current Therapies and Emerging Solutions. *Frontiers in Microbiology*;6. Epub ahead of print 20 January 2016. DOI: 10.3389/fmicb.2015.01528.

60. **Donders GG, Zodzika J, Rezeberga D**. Treatment of bacterial vaginosis: what we have and what we miss. *Expert Opinion on Pharmacotherapy* 2014;15:645–657.

61. **Bradshaw CS, Sobel JD**. Current Treatment of Bacterial Vaginosis—Limitations and Need for Innovation. *Journal of Infectious Diseases* 2016;214:S14–S20.

62. **Bilardi J, Walker S, McNair R, Mooney-Somers J, Temple-Smith M, *et al.*** Women's Management of Recurrent Bacterial Vaginosis and Experiences of Clinical Care: A Qualitative Study. *PLOS ONE* 2016;11:e0151794.

63. **Ling Z, Liu X, Chen W, Luo Y, Yuan L, *et al.*** The Restoration of the Vaginal Microbiota After Treatment for Bacterial Vaginosis with Metronidazole or Probiotics. *Microbial Ecology* 2013;65:773-780.

64. **Zozaya-Hinchliffe M, Martin DH, Ferris MJ**. Prevalence and Abundance of Uncultivated Megasphaera-Like Bacteria in the Human Vaginal Environment. *Applied and Environmental Microbiology* 2008;74:1656–1659.

65. **Srinivasan S, Munch MM, Sizova MV, Fiedler TL, Kohler CM, *et al.*** More Easily Cultivated Than Identified: Classical Isolation With Molecular Identification of Vaginal Bacteria. *J Infect Dis* 2016;214 Suppl 1:S21-28.

66. **Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, *et al.*** 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. *Journal of clinical microbiology* 2000;38:3623–3630.

67. **Seng P, Abat C, Rolain JM, Colson P, Lagier JC, *et al.*** Identification of Rare Pathogenic Bacteria in a Clinical Microbiology Laboratory: Impact of Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. *Journal of Clinical Microbiology* 2013;51:2182–2194.

68. **Angelakis E, Bachar D, Henrissat B, Armougom F, Audoly G, *et al.*** Glycans affect DNA extraction and induce substantial differences in gut metagenomic studies. *Scientific Reports*;6. Epub ahead of print September 2016. DOI: 10.1038/srep26276.

69. **Austin MN, Rabe LK, Srinivasan S, Fredricks DN, Wiesenfeld HC, *et al.*** Mageeibacillus indolicus gen. nov., sp. nov: A novel bacterium isolated from the female genital tract. *Anaerobe* 2015;32:37–42.

70. **Stackebrandt E and Ebers J.** Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006;33:152.

71. **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. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY* 2014;64:346–351.

72. **Vandamme P, Pot B, Gillis M, De Vos P, Kersters K, *et al.*** Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiological reviews* 1996;60:407–438.

73. **Abdallah RA, Beye M, Diop A, Bakour S, Raoult D, *et al.*** The impact of culturomics on taxonomy in clinical microbiology. *Antonie Van Leeuwenhoek* 2017;110:1327–1337.

74. **Murray PR, Baron EJ, Jorgensen JH, Landry ML, Pfaller MA.** *Manual of clinical microbiology.* 9th ed. Washington, D.C : ASM Press; 2007.

75. **Citron DM, Ostovari MI, Karlsson A, Goldstein EJ**. Evaluation of the E test for susceptibility testing of anaerobic bacteria. *J Clin Microbiol* 1991;29:2197–2203.

76. **Le Page S, van Belkum A, Fulchiron C, Huguet R, Raoult D, *et al.*** Evaluation of the PREVI® Isola automated seeder system compared to reference manual inoculation for antibiotic susceptibility testing by the disk diffusion method. *European Journal of Clinical Microbiology & Infectious Diseases* 2015;34:1859–1869.

77. **Sasser M**. Bacterial Identification by Gas Chromatographic Analysis of Fatty acids Methyl Esters (GC-FAME). *MIDI, Technical Note*.

78. **Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, *et al.*** The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res* 2001;29:22–28.

79. **Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, *et al.*** A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. *International Journal of Systematic and Evolutionary Microbiology* 2014;64:384–391.

80. **Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, *et al.*** Culture of previously uncultured members of the human gut microbiota by culturomics. *Nature Microbiology* 2016;1:16203.

81. **Lagier JC, Dubourg G, Million M, Cadoret F, Bilen M, *et al.*** Culturing the human microbiota and culturomics. *Nature Reviews Microbiology* 2018;16:540–550.