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 2017

Par El Hadji SECK

Étude de la diversité des procaryotes halophiles du tube digestif par approche de culture

Pour obtenir le grade de DOCTORAT d'AIX-MARSEILLE UNIVERSITÉ

Spécialité : Pathologie Humaine

Membres du Jury de la Thèse :

Mr le Professeur Jean-Christophe Lagier	Président du jury
Mr le Professeur Antoine Andremont	Rapporteur
Mr le Professeur Raymond Ruimy	Rapporteur
Mr le Professeur Didier Raoult	Directeur de thèse

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes,

UMR 7278

Directeur : Pr. Didier Raoult

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

Sommaire

Résumé

L'alimentation est le principal déterminant du microbiote intestinal. Ces dernières années beaucoup d'études se sont penchées sur son rôle dans la variation du microbiote intestinal.

Cependant, la plupart des études qui lient le régime alimentaire et le microbiote intestinal sont axées sur les macronutriments comme les glucides, les graisses et les protéines, mais le rôle des micronutriments et des minéraux tel que le NaCl a été négligé à ce jour. Le sel est un minéral essentiel à la vie humaine et l'un des assaisonnements alimentaires les plus anciens et les plus omniprésents. Le salage est la méthode la plus ancienne de conservation des aliments. Cependant, une consommation élevée de sel a été associée à l'obésité, au syndrome métabolique, à l'hypertension artérielle, aux maladies cardiovasculaires, au cancer gastrique, aux maladies auto-immunes et à la mortalité toutes causes confondues. Plusieurs mécanismes sous-jacents, y compris le stress oxydatif, ont été étudiés. Mais la salinité dans l'intestin et l'altération possiblement associée de son microbiote, récemment identifiées comme un symbiote critique de la santé et de la maladie, n'ont pas encore été explorées chez l'homme. Ici, en testant 1334 prélèvements de selles, nous avons montré qu'une salinité élevée était associée à une diminution de la diversité globale et à l'émergence de populations microbiennes halophiles dans l'intestin. La salinité fécale était associée au régime alimentaire salé et à l'obésité, conformément aux données épidémiologiques. Aucun procaryote halophile n'a été cultivé en dessous d'un seuil de salinité fécale de 1,5 %. Au-delà de ce seuil, nous avons découvert une diversité inattendue de microbiote halophile humain dont la richesse était corrélée avec les concentrations de sel; 64 espèces différentes ont été isolées, dont 21 nouvelles espèces et 43 espèces connues dans l'environnement mais non chez les humains. Trois procaryotes extrêmement halophiles ont été isolés, dont deux Archaea appartenant au genre *Haloferax*, avec une nouvelle espèce *Haloferax massiliensis*, et un nouveau genre bactérien, *Halophilibacterium massiliense*. La découverte d'une telle diversité de procaryotes halophiles chez l'homme est assez surprenante. En effet, les procaryotes halophiles sont décrits comme des micro-organismes vivants dans des environnements hypersalés tels que l'eau de mer. La composition microbienne de l'eau de mer est principalement dominée par des bactéries halophiles. En effet plus de la moitié d'entre elles sont halophiles, qui ne sont qu'occasionnellement pathogènes pour l'homme. C'est l'une des raisons pour lesquelles, dans la littérature, il existe peu de preuves que les infections peuvent être contractées directement par l'eau de mer. Néanmoins, cela contraste avec le nombre de mesures de la qualité de l'eau organisées pour prévenir le risque d'infection humaine. Il nous a paru intéressant de vérifier au laboratoire si les bactéries pathogènes humaines étaient susceptibles de se développer dans l'eau de mer. En conclusion, nous avons démontré que la plupart des bactéries citées en microbiologie clinique n'ont pas survécu longtemps dans les eaux marines en raison de

nombreux facteurs, y compris la salinité. Les maladies marines sont dominées par les infections à *Vibrio* et *Shewenalla.* Nous avons estimé que les dépenses considérables consacrées à la prévention des infections marines ne sont pas fondées sur des données probantes. D'autres études sont nécessaires pour réévaluer le risque infectieux associé à l'eau de mer afin de réadapter les politiques de surveillance des eaux marines.

Mots-clés : microbiote digestif, halophile, sel, NaCl, eau de mer, infection, pathogène.

Abstract

Diet is the main determinant of gut microbiota. In recent years many studies have looked at its role in the variation of the gut microbiota. However, most studies linking diet and gut microbiota focus on macronutrients such as carbohydrates, fats and proteins, but the role of micronutrients and minerals such as NaCl have been neglected to date. Salt is an essential mineral to human life and one of the oldest and the most ubiquitous dietary seasoning, and salting is an important and oldest method of preserving food. However, elevated salt consumption has been associated with obesity, metabolic syndrome, high blood pressure, cardiovascular disease, gastric cancer, autoimmune disease and all-cause mortality. Several underlying mechanisms, including oxidative stress, have been investigated, but salinity in human gut and the possible associated alteration of its microbiota recently identified as a critical symbiote of health and disease, have not yet been investigated in humans. Here, by testing 1,334 stools, we have shown that high salinity is associated with a decrease in overall diversity but the emergence of halophilic microbial populations in the intestine. Fecal salinity was associated with saline diet and obesity, according to epidemiological data. No halophilic prokaryote can be grown below a fecal salinity threshold of 1.5%. Beyond this threshold, we discovered an unexpected diversity of human cultured halophilic microbiota whose richness was correlated with salt concentrations; 64 different species were isolated, including 21 new unknown species and 43 known species in the environment but not in humans. Three extremely halophilic prokaryotes were isolated, including two Archaea belonging to the genus *Haloferax*, with a new species *Haloferax massiliensis*, and a new bacterial genus, *Halophilibacteriums massiliense.* The discovery of such a diversity of halophilic bacteria in humans is quite surprising. Halophilic prokaryotes are described as micro-organisms living in hypersaline environments such as seawater. The microbial composition of seawater is dominated mainly by halophilic bacteria.

More than half of them are halophiles, which are only occasionally pathogenic to humans. This is one of the reasons why there is little evidence of infections that can be directly contracted by seawater in the literature. However, this contrast with the number of water quality measurements organized to prevent the extremely rare risk of human infections. We found very interesting to investigate whether human pathogenic bacteria were likely to develop in seawater. In conclusion, we have shown that most bacteria cited in clinical microbiology have not survived long in marine waters due to many factors, including salinity. Marine diseases are dominated by *Vibrio* and *Shewenalla* infections. We found that significant expenditures for preventing marine infections are not evidence-based. Further studies are needed to re-evaluate infectious risks associated with seawater in order to rehabilitate marine water monitoring policies.

Keywords: gut microbiota, halophilic, salt, NaCl, seawater, infection, pathogen.

Introduction

L'intestin humain est colonisé par une communauté complexe de cellules microbiennes. Les bactéries représentent la partie la plus étudiée du microbiote humain avec plus de 10^{14} cellules bactériennes (1). Les procaryotes sont connus pour jouer des rôles physiologiques importants (2) qui influencent les processus métaboliques, tels que la digestion des glucides complexes et le stockage des graisses (3,4). A cause de la variation des paramètres physico-chimiques comme le pH, la teneur en oxygène, la disponibilité en nutriments ou le stress oxydatif (5), l'écosystème microbien diffère tout le long du tube digestif. La composition de la flore intestinale au sein d'un individu est intrinsèquement associée au génotype, et à son âge. Elle présente en outre des changements dynamiques conduits par des facteurs externes tels que, l'environnement, les antibiotiques, et l'alimentation. Le régime alimentaire est le principal facteur qui influence le microbiote digestif (6). Parmi les aliments qui interviennent dans les variations du microbiote digestif, on peut citer le sel. Le chlorure de sodium (NaCl), également appelé sel de table, est un minéral essentiel pour la vie humaine, et le salé est l'un des goûts humains fondamentaux. Le sel est aussi l'un des aliments les plus anciens et le plus ubiquitaire. Le salage est une méthode importante et ancienne de conservation des aliments. En effet, le sel empêche le développement des micro-organismes qui est l'un des obstacles les plus importants dans la préservation des aliments. Néanmoins, les halophiles constituent un groupe de micro-organismes qui sont connus pour prospérer dans des environnements hypersalins (7,8). Ces derniers ont des stratégies leur permettant non seulement de résister au stress osmotique (8,9), mais aussi de mieux se multiplier en présence de sel (7). La présence de procaryotes halophiles dans l'intestin humain a déjà été postulée, sur la base de notre exposition généralisée à des aliments riches en sel et la détection récente de viables organismes dans les sels de qualité alimentaire raffinés (10,11).

Cependant, malgré des études approfondies de la diversité microbienne du tractus gastro-intestinal (12–14) et la prise de conscience que les haloarchaea peuvent être détectées dans les fèces (15,16), les bactéries halophiles ont largement échappé à la détection avant ces dernières années. Des études plus récentes ont prouvé la présence d'Archaea halophiles dans le tube digestif par des méthodes de biologie moléculaire indépendantes de la culture et un grand intérêt est accordé à leur rôle potentiel (10,17). C'est pour cette raison que dès l'entame de nos travaux, nous avons tenté de dresser un répertoire exhaustif des procaryotes halophiles déjà isolés chez l'homme. Nous nous sommes servis d'un programme informatique pour parcourir la littérature en ce qui concerne les publications relatives aux procaryotes halophiles **(Partie I)**. Ensuite nous avons essayé d'explorer le microbiote halophile du tube digestif, qui pourrait avoir été négligé dans les enquêtes précédentes et en même temps vérifier les liens putatifs qui pourraient exister entre la salinité fécale et la diversité halophile du tube digestif **(Partie II)**. Ici, nous avons choisi la culture comme moyen d'exploration, contrairement à la plupart des études qui caractérisent le microbiote halophile humain. Malgré toutes les avancées technologiques récentes de la biologie moléculaire, la culture pure reste le seul moyen pour une description des propriétés physiologiques des procaryotes et pour l'évaluation de leur potentiel de virulence (18). En outre, elle est complémentaire de la métagénomique qui produit des millions de séquences d'ADN non assignées. Chaque nouvelle espèce découverte grâce à la culture permet d'assigner de telles séquences sans signification, réduisant ainsi la "matière noire" métagénomique (référence). Enfin, la culture met en évidence des microbes vivants, tandis que la métagénomique ne détecte que des séquences d'ADN qui peuvent ne pas correspondre à aucun organisme vivant (référence). Depuis 2009, nous avons développé au sein de notre laboratoire un nouveau concept « microbial culturomics » qui, repose sur la variation des paramètres physico-chimiques des conditions de culture, de manière à exprimer au maximum la diversité microbienne. Il s'appuie sur une méthode d'identification rapide par spectrométrie de masse

MALDI-TOF complétée par le séquençage de l'ARNr 16S lorsque cela est nécessaire (13). Dans ce travail, nous avons fabriqué des milieux spécifiques contenant de fortes concentrations de sel pour cultiver des halophiles à partir d'échantillons de selles. La concentration de sel était modulable selon la catégorie halophile qu'on voulait isoler.

La majorité des halophiles habitent le milieu marin où la concentration en sel est voisine de 3,5% (19). Elles sont rarement pathogènes. La capacité de certaines bactéries à croitre dans le milieu marin a toujours suscité un intérêt en microbiologie clinique et en maladies infectieuses (20). Le nombre de bactéries présentes dans un échantillon d'eau de mer résulte d'un équilibre entre la multiplicité des interactions qui influencent leur croissance, leur reproduction, contrôlent leur survie et causent leur mort (21). Des études ont signalé que les eaux marines présentent une activité antibactérienne marquée vis-à-vis des micro-organismes non marins. Les eaux naturelles ont une forte capacité inhérente à se purifier (22). Ce processus implique des facteurs physiques, biologiques et des antagonistes chimiques (23). Malgré ces considérations, l'eau de mer constitue pour certains scientifiques un facteur de risque infectieux pour l'homme. Dans de nombreux pays, les eaux de baignade contenant plus de 35 CFU/100 mL−1 entérocoques and 126 CFU/100 mL−1 *Escherichia coli* sont considérées comme dangereuses et vecteur d'infections de diverses formes (24). L'intérêt de la détection de ces coliformes, à titre d'organismes indicateurs, réside dans le fait qu'on considère que leur survie dans l'environnement est généralement équivalente à celle des bactéries pathogènes (25). Les microbiologistes cliniques à ce jour ont été préoccupés à étudier la survie de ces coliformes fécaux, et celle des autres pathogènes semble être oubliée (26). Jusqu'à présent, peu d'études sur la survie des bactéries pathogènes dans l'eau ont été entreprises. Pourtant, de nombreuses controverses autour de la pertinence de l'utilisation des indicateurs actuels dans la prédiction du risque infectieux lié à l'eau de mer sont nées (27). En effet, la grande majorité des microbiologistes cliniques sont en accord sur le fait qu'aucun des indicateurs bactériens utilisés

ne réponde aux critères idéaux pour statuer sur la qualité sanitaire de l'eau (28). L'évaluation de la survie dans l'eau de mer des pathogènes les plus cités en microbiologie clinique pourrait permettre de mieux cerner la situation et de trouver des indicateurs alternatifs, appropriés pour la surveillance des eaux marins **(Partie III).**

Dans les deux dernières parties de ce travail **(Partie IV et Partie V)**, nous avons décrit les nouvelles espèces isolées en utilisant une nouvelle méthode de description de nouvelles espèces appelées taxonogenomics (29). Ce concept se base sur les données génomiques de la bactérie associées à la description phénotypique et protéomique (30).

CHAPITRE I

Répertoire des prokaryotes halophiles isolés chez l'homme (Revue).

Avant-propos

Pour dresser un répertoire exhaustif des bactéries halophiles isolées chez l'homme, nous nous sommes servis d'un programme informatique pour interroger pubmed pour chacune des 2172 bactéries associées à l'homme, avec trois groupes de requêtes construites à partir de termes Mesh reliés à des notices bibliographiques de publications sur les halophiles. Pour chaque bactérie, le programme permet d'obtenir un nombre « n » de publications retrouvées par requête. Les publications retrouvées ont été analysées « manuellement » afin de confirmer qu'ils étaient bien question d'un signalement ou d'une identification de la bactérie considérée comme halophile. Lors de cette analyse, nous procédions par étape en commençant par considérer d'abord les notices retrouvées par requête la plus spécifique (QP3). Si les articles obtenus par cette requête ne permettaient pas de statuer sur le caractère halophile ou halotolérante de la bactérie, nous considérions les notices des requêtes les plus sensibles (QP2, puis QP1). Lorsque cela était nécessaire, nous avons consulté l'article de description mentionné sur le site bacterio.net pour classer définitivement la bactérie comme halophile ou non. Nous avons ainsi identifié parmi les 2172 bacteries isolées chez l'homme, 258 (11,87 %) procaryotes halotolérants et 51 (2,34 %) procaryotes halophiles, dont deux Archaea. Parmi les 51 procaryotes halophiles, 31 (60,78 %) étaient des espèces modérément halophiles, 17 (33,33 %) étaient légèrement halophiles et 3 (5,89 %) extrêmement halophiles. Nous avons identifié 2 (0,7 %) espèces classées dans le groupe à risque 2 des agents pathogènes humains (*Vibrio cholerae* et *Vibrio parahaemolyticus*).

Les halophiles et les halotolérants constituent la population naturelle de l'eau de mer. Elles sont occasionnellement pathogènes. Dans ce travail, nous sommes également revenus sur les risques infectieux associés à l'eau de mer ainsi que les facteurs affectant la survie des bactéries non marines dans la mer.

Article 1: Halophilic and Halotolerant Prokaryotes in Human

El Hadji SECK[1], Jean-Charles DUFOUR[2, 3], Didier RAOULT[1, 4*] and Jean-Christophe LAGIER[1]

Soumis dans Future microbiology

HALOPHILIC AND HALOTOLERANT PROKARYOTES IN HUMAN

El Hadji SECK[1], Jean-Charles DUFOUR[2, 3], Didier RAOULT[1, 4*] and Jean-Christophe LAGIER[1]

[1]URMITE, UMR63 CNRS 7278, IRD 198, Inserm 1095, Aix Marseille Université, IHU-Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13005 Marseille.

[2] SESSTIM (UMR912), Sciences Economiques et Sociales de la Santé et Traitement de l'Information Médicale, Aix Marseille Univ, INSERM, IRD, Marseille, France.

[3] Service Biostatistique et Technologies de l'Information et de la Communication (BIOSTIC), Assistance Publique Hôpitaux de Marseille, Hôpital de la Timone, Marseille, France.

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

*Corresponding author: Prof. Jean-Christophe LAGIER

URMITE, IHU-Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13005 Marseille

Email: jclagier@yahoo.fr

Phone: 04 13 73 24 01

Fax: 04 13 73 24 02

Keywords: Bacteria, Halophilic, Halotolerant, Prokaryote, Infection, Human, Seawater

Abstract

Halophilic prokaryotes are described as microorganisms living in hypersaline environments such as marine water. However, they have been found in the human intestinal mucosa by molecular biological techniques. In addition, recently, thanks to culturomics, various halophilic prokaryotes were s isolated from human gut. Nevertheless, the human halophilic repertoire remains widely unknown. To date, it is necessary to compile a complete repertoire of the haloterant bacteria isolated in humans. By searching in the scientific literature, we identified among the 2, 172 bacteria isolated in humans, 258 (11.87 %) halotolerant prokaryotes and 51 (2.34 %) halophilic prokaryotesincluding archaea. Among 51 halophilic prokaryotes, 31 (60.78 %) were moderately halophilic species, 17 (33.33%) were slightly halophilic and 3 (5.89 %) extremely halophilic prokaryotes. At the phylum level, we identified 5 phyla among halophilic prokaryotes associated with humans: 28 (54.90%) belong to the *Firmicutes*, 15 (29.41%) to the *Proteobacteria*, 4 (7.84%) to the *Actinobacteria*, 3 (5.88%) to the *Euryarchaeaota* and 1 (1.97 %) belong to the *Bacteroidetes.* The infections acquired by exposure to halophilic and halotolerant prokaryotes are not frequent, 304 (98.38%) of the 309 halophilic and halotolerant species identified in humans were considered as Risk Group 1 and among these 51 halophilic prokaryotes were identified 2 (3.92%) species classified in risk group 2, (*Vibrio cholerae*, *Vibrio parahaemolyticus*) and 1 (1.96%) specie in the risk group 3 (*Bacillus anthracis*). This comprehensive repertoire of halophilic prokaryotes isolated in humans and their pathogenicity can indirectly improve our understanding of seawater-borne infections that today sheds much controversy.

Introduction

Most bacterial species grow at temperatures ranging from 20 °C to 40 °C, with a neutral pH and a salinity ranging from 0% to 3% (Lagier et al., 2016). Species that happen to grow, live and multiply in extraordinary conditions of pressure, heat, or salinity are called extremophiles (Grégoire et al., 2009; Rampelotto, 2016). Among the extremophiles, alcalophiles, acidophiles, barophiles, hyperthermophiles and halophiles have already been described (Oren, 2013b). Halophile are microorganisms living mainly in hypersaline environments and often require a high salt concentration for growth (Adamiak et al., 2016).

The number of microorganism present in humans is extremely larger, reaching to $10^{12} \sim 10^{14}$ CFU, particularly in human gut (Gill et al., 2006). To this day, we estimated that the number of bacterial and archaeal species associated with human is 2,172. Only 130 (5.98) of the known bacteria is pathogenic (Hugon et al., 2015). Haloarchaea DNA has been found in human intestinal mucosa by molecular biological techniques (Oxley et al., 2010). However, the authors did not specify whether it was from gut commensal prokaryote or transient microorganisms. Recently, thanks to culturomics, it was isolated from human gut various halophilic prokaryotes (Oxley et al., 2010). However the human halophilic repertoire remains widely unknown. In this work, we will comprehensively investigate the literature in order to establish a comprehensive repertoire of halophilic bacteria isolated in human. Finally, only a small number of these bacteria can be responsible for human and animal disease (DasSarma and DasSarma, 2012). We investigated the infectious risk associated with the marine environment that seems to be overestimated.

1. Halophilic prokaryotes

1. Classification

Halophilic prokaryotes are a physiologically heterogeneous group of microorganisms (Oren, 2008). Several categories of halophiles are distinguished according to the salt concentration that they require (Grégoire et al., 2009): "slightly halophilic" species develop optimally between 1–5% of NaCl (Kim et al., 2010); "moderately halophilic" species develop optimally between 5–20% of NaCl (Koh et al., 2017) and "extreme halophiles" species between 20–30% of NaCl (Moreno et al., 2010). Halotolerant species do not have an absolute requirement for salt for their growth, but often grow well at very high salt concentrations (Roberts, 2005).

2. Taxonomy and phylogeny

The halophilic microorganisms are found in the three domains of life (Oren, 2008) . Microorganisms living in hypersaline environments is dominated by prokaryotes (archaea and bacteria), with a few microbial eukaryotes, such as the photosynthetic and heterotrophic protists, fungi and the crustacean *Artemia salina* (Gunde-Cimerman et al., 2009; Oren, 2002). Most halophilic prokaryotes are moderately halophilic and belong to the phyla of *Cyanobacteria, Proteobacteria, Firmicutes, Actinobacteria, Spirochetes, and Bacteroidetes* (Kim et al., 2012; Oren, 2002). Methanogenic euryarchaeota include several genera including halophilic species. These genera are *Methanohalophilus, Methanohalobium, Methanosporirillum, Methanococci* (Arahal et al., 1996).

Halophilic and halotolerant bacteria present an important phylogenetic diversity. However there are a few homogeneous groups that consisting only of halotolerant species. There are several examples such as the order *Halobacteriales*, fermentative anaerobic bacteria of the order *Halanaerobiales* (Kivistö and Karp, 2011; Oren, 2008; Ventosa et al., 2015), the family *Halobacteriacae* and The family *Halomonadaceae* (Romano et al., 2006).

3. Adaptation to the hypersaline environment

By definition, halophilic organisms (Greek alos: salt and philein: to love) are living beings that tolerate well or need salt for their growth (DasSarma and DasSarma, 2012). Halophiles are microorganisms that live in hypersaline environments and often require a high salt concentration for their growth (Ma et al., 2010; Minegishi et al., 2008).

In order to survive in high salt concentrations, halophilic microorganisms must maintain an osmotic balance with their external environment (Mojica et al., 1997). Stress due to high salt concentration induces in most halophilic and halotolerant bacteria an adaptation of their surface and membrane structure to their highly ionic environment (Siglioccolo et al., 2011). The outer surfaces of the cytoplasmic membrane and the extracytoplasmic structures are in contact with high concentrations of inorganic ions (Roberts, 2005). In terms of their intracellular compartment, two different strategies have been developed: the accumulation of solutes compatible with their environment, mainly KCl (Oren, 2013a) and the synthesis of sugars, amino acids, glycine betaine, trehalose and ectoin (Kennedy et al., 2001). In addition to high salt concentrations (NaCl, KCl), proteins are generally destabilized due to increased hydrophobic interactions (Madigan and Oren, 1999). Halophilic proteins have therefore developed specific mechanisms that allow them to be both stable and soluble in the presence of high KCl cytoplasmic concentration (Muller and Oren, 2003). Indeed, all halophilic proteins are strongly negatively charged with hydrated carboxyl groups which maintain the solubility of proteins at high salt concentrations (Joo and Kim, 2005).

4. Usual ecosystems

a) Environment

Halophilic and halotolerant prokaryotes generally live in environments where the salt composition is dominated by NaCl (Ma et al., 2010). Sodium and chloride ions are major ions in environement derived from seawater and in thalassohalin media (Culkin and Cox, 1976a; Oren, 2015). In some athalassohaline media such as the Dead Sea, hypersaline soda lakes, saline soils, ancient halites, sodium is the main component of the sum of the anions and Cl is present in quantity. In this type of environment where salinity can exceed 10% NaCl moderately halophilic bacteria and extremely predominantly halophilic bacteria (33–37). Most halophiles inhabit the marine environment, where the salt concentration is close to 3.5% (Ma et al., 2010).

b) Food

Halophiles are involved in centuries-old processes, such as salt production and the production of fermented foods, consumed by humans (Fukushima et al., 2007; Justé et al., 2008; Roh et al., 2010a). Indeed, foods with a high salt concentration are favorable environments for the growth of halophilic microorganisms (Lee, 2013). Today, with the emergence of new biological technologies, these organisms have been described from many traditional foods such as salt (Diop et al., 2016), cheese, fermented shrimps, fermented fish but also from a range of foods manufactured products such as salted fish, skins, pork, sausages and fish sauces, soy sauce, flour where large quantities of salt are used in manufacturing and preservation processes (Lee, 2013; Roh et al., 2010a). In **Table 1** we have summarized the repertoire of halophilic bacteria isolated in foods.

2. Halophilic prokaryote in human

1. Repertoire of halophilic bacteria in human

1.1.Unusual habitat: Human

Usually, halophilic prokaryotes are described as living beings that generally flourish with extreme environmental conditions (Averhoff and Müller, 2010). However, their recent discovery in habitats with relatively low salinity has changed these considerations (Purdy et al., 2004). Studies demonstrate to day their ability (or adaptation strategy) to survive in low-salinity environments (0.5%) close to normal physiological state (Fukushima et al., 2007; Roh et al., 2007b). The intestinal content of a healthy individual has a mean salinity similar to that of plasma (135 to 145 mM sodium) (43). Over the last year, metagenomics and molecular analyses have broadened the spectrum of known human-associated Archaea to include organisms such as Halobacteria. Today we have evidence that haloarchaea colonized the human intestine (Oxley et al., 2010). In fact, *Halorubrum koreense, Halococcus morrhuae, Halorubrum saccharovorum, Halorubrum alimentarium* have been detected in human gut by studies using cultivation-independent methods (Dridi et al., 2011; Nam et al., 2008b). These results are in accord with those reported by Oxley et al. (2010). In the intestine, haloarchaea DNA sequences, and particularly *Halobacteriaceae*, have been identified in up to 20% of human colonic mucosal biopsies (Oxley et al., 2010).

1.2.Repertoire of halophilic and halotolerant prokaryote

Despite the recent discoveries on halophilic prokaryotes isolated from humans, we have a very limited knowledge about to this commonality (Oxley et al., 2010). To date, it is necessary to compile à complete repertory of the haloterant bacteria isolated in humans. For that purpose, using 50 well-known halobacteria, we conducted a literature search in order to complete/confirm our knowledge about MeSH Terms and text-words used to index article focused on such species. The 50 halobacteria belonging to the following genera: *Actibacter,*

Aeromicrobium, Aidingimonas, Agrobacterium, Halobacillus, oceanobacillus, Pausisalibacillus, Anaerosalibacter, Chromohalobacter, Aestuariicola, Actinopolyspora, Aequorivita, Bacillus, Aestuariibaculum, Aestuariibacter, Aestuariicola, Afifella, Algoriphagus, Salinisphaera, Halomonas, Paucisalibacillus, Namhaeicola, Roseovarius, amorphus, Selenihalanaerobacter, Acanthopleuribacter, Agrococcus, Agromyces, Agarivorans, Aeromonas. The literature search included articles indexed in PubMed / MEDLINE that mentioned the scientific name of the bacteria in the title or abstract. We extracted and analyzed all MeSH terms of the bibliography; only those directly or indirectly referring to halophiles were selected. Additionally, certain articles and bibliographic records were read to identify text-words that were evocative of the halophilic notion. MeSH terms and text-words were used to establish three query patterns (QP1, QP2, and QP3) **(Table 2, Figure 1a).** We have developed a computer program capable of querying PubMed / MEDLINE for each of the 2,172 bacteria isolated in humans with these three query patterns. This program uses "Entrez Programming Utilities" provided by NCBI (http://www.ncbi.nlm.nih.gov/books/NBK25501/). Previously, a similar approach has been carried out in our team by Hugon et al to create a repertoire of isolated prokaryotes in humans (Hugon et al., 2015). For each bacterium, the program provides numbers (n1, n2 and n3) of bibliographic records respectively retrieved by query patterns QP1, QP2 and QP3 (due to the syntactical construction: $n1 \geq n2 \geq n3$). The bibliographic records found were analyzed "manually" to confirm that it was indeed an identification of the bacteria in salt. For this analysis, we proceeded in steps, beginning with the records retrieved by QP3 (most specific query). If the notion of 'halophilic' was not found, we view the QP2 records and then the QP1 records (most sensitive query) **(Figure 1b).** When necessary, we read the entire text. As a last resort, when no item was found using our query patterns or items did not enable us to conclude that the bacteria

was halophilic, we consulted the article description mentioned on the site bacterio.net to definitively classify the bacteria as halophilic or not.

The results are given in the form of a PMID number. We have thus identified among the 2,172 bacteria isolated in humans, 258 (11.87 %) halotolerant prokaryotes **(suppl.Table S1, Figure 2a)** and 51 (2.34 %) halophilic prokaryote including 2 archaea **(Table 3).** Among 51 halophilic prokaryotes, 31 (60.78 %) were moderately halophilic species, 17 (33.33 %) were slightly halophilic and 3 (5.89%) extremely halophilic prokaryote (Table 2). At the phylum level, we identified 5 phyla among halophilic prokaryotes associated with humans: 28 (54.90 %) belong to the *Firmicutes*, 15 (29.41%) to the *Proteobacteria*, 4 (7.84%) to the *Actinobacteria*, 3 (5.88%) to the *Euryarchaeaota* and 1 (1.97 %) belong to the *Bacteroidetes* **(Figure 2b).** When focusing on bacterial species isolated at least once in the human gut, 168 (7.73 %) halotolerant prokaryotes and 37 (1.70 %) halophilic prokaryotes were identified including 3 archaea (*Haloferax alexandrinus, Haloferax massiliensis, Methanomassiliicoccus luminyensis*) **(Table 3).** Most human intestinal halophilic and halotolerant prokaryotes belong to the *Firmicutes, Proteobacteria* and *actinobacteria* phyla. In fact a half of the identified salt tolerate (halophilic and halotelerant) prokaryote isolated in human belongs to Firmicutes phylum **(Table 3).** The phylum of *Actinobacteria* is third in the community of human halobacteria. However, it consists mainly of halotolerant bacteria 76 (95 %), only *Nesterenkonia halobia, Nesterenkonia lacusekhoensis, Sciscionella marina,* are halophilic species. Among the halotolerant *Actinobacteria*, mycobacteria, which are of particular interest in clinical microbiology, are found (Schulze-Röbbecke, 1993)**.** In **Fig. 3 c**, we have represented the genera in which at least one halophilic representative is found. Among them, the genus *Bacillus* and *Paenibacillus* are the most representative with respectively 30 (22.22%) and 29 (21.48%) species. In particular, the genus *Chromohalobacter, Flavobacterium, Haloferax, Halobacillus, Halomonas, Methanomassiliicoccus, Nesterenkonia, Oceanobacillus, Paucisalibacillus, Planococcus,*

Salinisphaera, Terribacillus, Virgibacillus, Sciscionella are constituted only by halophilic prokaryotes **(Fig. 3c).**

Among the 51 halophilic bacteria isolated in humans, 43 (84.31 %) were isolated by culturomics (**Table 3**). Culturomics is a new approach that is based on varying physical and chemical conditions of culture to explore deeply bacterial ecosystems (Lagier et al., 2012). Culturomics has significantly increased the known microbial community of the human intestine (Lagier et al., 2015). Indeed in recent years 531 additional bacteria have been isolated in humans thanks to this technique (Lagier et al., 2016). With regard to halophilic prokaryotes, we use media containing a high salt to cultivate halophilic bacteria from human stool (Seck et al., 2015).

2. Halophilic bacteria causing human and animal diseases

Infectious agents pathogenicity involves, apart from factors related to the host, its enzymatic equipment, and growth capacity in space and time (Isenberg, 1988). Prokaryote growth is strongly influenced by the physicochemical nature of the environment (Lagier et al., 2012). Indeed, the bacterial requirements include food, energy and water sources, but also an adequate temperature, as well as suitable pH, oxygen concentration, mineral salt such as NaCl. The constitutive or energetic nutrients required for a microorganism for its development must be provided under certain environmental conditions so that it can develop well (Vartoukian et al., 2010).

The infections acquired by exposure to halopholic prokaryote are not frequent. Most infections are contracted after drowning (Robert et al., 2017). It generally concerns halotolerant microorganisms capable to grow outside salty environments but can resist in a salty environment. When considering human-associated halophilic and halotolerant bacteria, 304 (98.38%) of the 309 halophilic and halotolerant species identified in humans, were considered as Risk Group 1 and among these 51 halophilic prokaryotes were identified. The microorganisms of Risk Group1 are considered to be biological agents unlikely to cause human

or animal disease according to the American Biological Safety Association classification (American Biological Safety Association., 2017). We identified 2 (3.92%) species classified in risk group 2, characterized by European Union as a moderate individual risk and a low risk for the community (*Vibrio cholerae*, *Vibrio parahaemolyticus*) (American Biological Safety Association., 2017). They are biological agents that can cause diseases for humans but are unlikely to reach a wider community. One species, *Bacillus anthracis*, belonged to risk group 3 which includes biological agents that can cause a serious disease in humans and then spread to the community (American Biological Safety Association., 2017).

3. Sea water and infection

Sea water has a high concentration of dissolved salts (Culkin and Cox, 1976b). Its salinity varies on the surface of the terrestrial globe, depending on the latitude, the opening of the seas to the oceans, their dimensions, land inputs, currents (Boyko, 2013; Hypersaline Brines and Evaporitic Environments: Proceedings of the Bat Sheva Seminar on Saline Lakes and Natural Brines, 2011; JANSSEN, F and SCHRUM, 1999).

The saline concentration of marine waters ranges from 3-8 g / l for the Baltic Sea (Voipio, 1981) to 192.2-260 g / l for the Dead Sea (Niemi et al., 1997). The major characteristic of seawater is that it contains salt. The average salinity is about 3.5% (Millero et al., 2008) (compared to 0.9% for human organism). Life in marine environments requires special skills for microorganisms and many of these bacterial species are halophiles (Kim et al., 2012; Ma et al., 2010; Oren, 2015). Unpolluted seawater does not contain many bacteria (Shuval, 2003); the one that enters the oceans due to human activities is rapidly dying (Carlucci and Pramer, 1959). The number of bacteria in seawater is limited by factors that influence growth and reproduction (Borrego et al., 1983).

3.1.Factors affecting survival of bacteria in seawater.

The disappearance of pathogenic bacteria in seawater results from the combined action of multiple elements. The rapid death of bacteria entering the oceans can be explained by physical, chemical or biological factors (Cloutier et al., 2015).

a) Physical factors

The first element that can be cited is sedimentation, the adsorption of microbes (Gerba and McLEOD, 1976). These are inert or living particles suspended in seawater that absorb contaminants and migrate them to the bottom (Faust et al., 1975). The destructive action of sunlight is also a physical factor affecting the survival of bacteria. Indeed, light acts through its ultraviolet radiation on the bacteria that are either killed or inactivated (Alkan et al., 1995; Davies-Colley et al., 1994). Culture temperature is another factor that can influence the growth of marine bacteria. Temperature plays a limiting role on the development of bacteria in the sea, slowing down the biochemical processes and the reproduction of the germs accustomed to living around 37 ° C (Biosca et al., 1996; Maslin et al., 1995). Seasonal variations in temperate regions negatively influence bacterial life in marine waters (Van Donsel et al., 1967). Thermal sterilization is also a limiting effect on the survival of bacteria in seawater (Carlucci and Pramer, 1959). It has been shown that the germ *Escherichia coli* persist only for between 4-9 days subjected to sterilization comparable to that undergone under natural conditions (Rozen and Belkin, 2001). Pressure is only harmful at high levels (Carlucci and Pramer, 1959).

b) Chemical Factors

The salinity of surface seawater is estimated to approximately 3.5% of inorganic salt (Maslin et al., 1995). NaCl is the most representative ionic chemical compound (Culkin and Cox, 1976b). Based on their concentration, inorganic salts are the most potentially toxic substances present in the sea (Carlucci and Pramer, 1960a). Sea salts can adversely affect the survival of bacteria by a general osmotic effect. Osmotic shock caused by the transfer of cells from a relatively low salinity environment to a higher salinity environment becomes fatal to

most bacteria (Mojica et al., 1997). The survival of *E. coli* in Great Salt Lake waters (salinity of about 32%) was studied by ZoBell which observed a death of more than 95% in 1 minute (Zobell et al., 1937). The toxic effect of seawater on bacteria is also the result of specific ionic toxicity such as heavy metals (perchlorate, iodate) (Carlucci and Pramer, 1959, 1960b; Chen et al., 2007).

C) Biological factors

The survival and growth of non-marine microorganisms in seawater are mainly affected by the simple competition for nutrients (Baker et al., 1983; Carlucci and Pramer, 1960a). Seawater can be considered as a diluted culture medium in which the concentration of dissolved organic matter does not exceed 10 ppm (Carlucci and Pramer, 1959). The organic matter content of seawater is low except in the vicinity of the earth and in some lower deposits. In addition, important nutrients such as iron, nitrogen and phosphorus are also present in very small quantities (LIM and Flint, 1989). Besides the competitive effects, there are also antagonistic effects. Giaxa demonstrated that some marine bacteria were antagonistic to *Vibrio cholerae*. Studies have found, on 58 species of marine bacteria tested, 9 active antibiotic products against various test organisms, with toxic effects (Carlucci and Pramer, 1960b). Their number varied directly with the degree of pollution in wastewater. The most active antagonistic marine bacteria were species of the genera *Bacillus* and *Micrococcus* (Rosenfeld and ZoBell, 1947).

Bacteriophages were an important factor in the rapid death of bacteria in seawater (Carlucci and Pramer, 1960c).

3.2. Infectious risks of seawater

The mean salinity of the sea water is 3.5% (JANSSEN, F and SCHRUM, 1999). The species it harbor are halophilic or halotolerant (Kriss, 1967). The proportion of halobacteria among

human pathogens is relatively low, accounting for 0.2% (Supplement Table S2). They are responsible for certain pathologies reported as associated with marine waters (Hollis et al., 1976; Howard and Bennett, 1993; Sims et al., 1983). The most frequently halophilic and halotolerant species mentioned are *Pseudomonas aeruginosa, Shewanella putrefaciens, Mycobacteria spp, Micrococcus spp, Staphylococcus aureus, Streptococci, Vibrio cholerae, Vibrio parahaemolyticus, Vibrio vulnificus, Enteroccocus faecalis, Enteroccocus faecium.*.(Carlucci et al., 1959; Fiorillo et al., 2001; Cantet et al., 2013; Fleisher et al., 1993; Gonçalves et al., 1992; Marras et al., 2005; Rosenthal et al., 1975; Thorsteinsson et al., 1974; Sims et al., 1983 ; Tong et al., 2015 ; Vezzulli et al., 2013).

The infections contracted from seawater lead to clinical cases, at the extreme, that of cholera or most often that of gastroenteritis and diverse purulent skin infections, Ear, nose and throat infections (E.N.T) and ocular infections **(Table 4)** (Berger et al., 2007; Cantet et al., 2013; Chan et al., 2014; Fiorillo et al., 2001; Fleisher et al., 1993; Gonçalves et al., 1992; González-Serrano et al., 2002; Holt et al., 1997; Horseman and Surani, 2011; Lippi et al., 2016; Marras et al., 2005; Masoumi-Asl et al., 2016; Ong et al., 1991; Otsuka et al., 2007; Rosenthal et al., 1975; Thorsteinsson et al., 1974; Vezzulli et al., 2013; Vignier et al., 2013).

According to WHO estimation, contaminated marine waters are the cause of more than 120 million cases of gastrointestinal disease worldwide and more than 50 million cases of more severe respiratory disease, year (Shuval, 2003). However, these statistics seem to be largely overestimated (Fewtrell and Kay, 2015). Most of the previously published epidemiological studies on the effects of marine waters on health contain four main deficiencies (Fleisher et al., 1993). The first problem arises in the indicator bacteria studied Davies-Colley et al., 2004). According to current regulations, bathing waters containing more than 35 CFU / 100 mL^{-1} enterococci and 126 CFU 100 mL^{-1} *Escherichia coli*, are considered dangerous and may be carriers of various forms of infections (Soller et al., 2010a). Contrary to this ideal, studies have

shown that the populations of enterococci can be environmental origin are not necessarily likely to cause pathologies (Furet et al., 2009). The second bias is based on the inability to control the many temporal and spatial variations in the densities of indicator organisms (Cheung et al., 1990). In fact, it is proven that *E. coli* and other faecal coliforms only occur for a few hours in the marine environment (Rozen and Belkin, 2001; Whitman et al., 2004). The assumption that they die quickly is widely documented (Carlucci and Pramer, 1960b, 1960a, 1960c; Davies-Colley et al., 1994; Fujioka et al., 1981; Gerba et al., 1979; Gerba and McLEOD, 1976; Rozen and Belkin, 2001).

The inability to directly link the density of the indicator organism to a single person is an additional shortcoming. The presence of these organisms does not necessarily indicate a recent contamination event and in some cases may lead to an overestimation of the associated risk to public health (Halliday and Gast, 2011).

The last weakness noted is the impossibility of rigorously controlling non-water-related risk factors on previously reported associations between bathing in marine waters and disease. Bathers are also in contact with many elements, whose sanitary quality is not included in beach monitoring programs (Halliday and Gast, 2011). In fact, swimmers are often in contact with sand, shells and sediments in which fecal bacteria persist (Halliday and Gast, 2011; Phillips et al., 2011; Yamahara et al., 2012). Seafood and other staple foods raised or harvested close to shorelines and estuaries act as vectors of transmission of gastroenteritis (McLaughlin et al., 2005). Without denying the importance and especially the seriousness of bacterial contaminations from marine origin, it should nevertheless be observed that the main vector of contamination remains the factors not linked to the direct consumption of sea water (Fleisher et al., 1993; Whitman et al., 2014).

Conclusion

Given our long history with salt and our increasing consumption of salty foods, it is not surprising to detect halophilic prokaryotes in human. These results prove the survival of halophilic bacteria in humans, in this case in the digestive system and suggest that they can be regular or irregular intestinal flora members. Since human organism is not known to be a salty environment, other studies should specify the role of salt or salty products in the development of halophilic prokaryotes in humans. Indeed, they constitute with the halotolerants the natural population of the sea water. They are rarely pathogenic.

In contrast to halophile prokarytes, most human pathogens do not survive long in seawater. The latter is highly diluted and exerts antibacterial activity against non-marine microorganisms. This implies that infections from seawater remain rare. In the current literature, there is little evidence of infections after contact with sea water directly. This contrasts with the many restrictions and significant amounts allocated to beach monitoring by settlement agencies and public authorities. The French Research Institute for the Exploitation of the Sea estimated in 2014, almost 3.3 billion Euros the annual costs for marine waters monitoring in metropolitan France and the Gulf of Biscay maritime sub-region. Yet the majority of clinical microbiologists agree that none of the bacterial indicators used to date meet the ideal criteria for deciding on the sanitary quality of water. Another study would help to clarify the risk to human health associated with sea water and to find appropriate indicators for better decision-making regarding the monitoring of marine waters.

References

Adamiak, J., Otlewska, A., Gutarowska, B., and Pietrzak, A. (2016). Halophilic microorganisms in deteriorated historic buildings: insights into their characteristics. *Acta Biochim. Pol.* 63, 335–341.

Alkan, U., Elliott, D. J., and Evison, L. M. (1995). Survival of enteric bacteria in relation to simulated solar radiation and other environmental factors in marine waters. *Water Res.* 29, 2071–2080.

American Biological Safety Association. (2017). Risk group classification for infectious agents. Available at: http://www.absa.org/riskgroups/.

Amoozegar, M. A., Malekzadeh, F., Malik, K. A., Schumann, P., and Spröer, C. (2003). Halobacillus karajensis sp. nov., a novel moderate halophile. *Int. J. Syst. Evol. Microbiol.* 53, 1059–1063.

An, S.-Y., Asahara, M., Goto, K., Kasai, H., and Yokota, A. (2007). Terribacillus saccharophilus gen. nov., sp. nov. and Terribacillus halophilus sp. nov., spore-forming bacteria isolated from field soil in Japan. *Int. J. Syst. Evol. Microbiol.* 57, 51–55.

Arahal, D. R., Dewhirst, F. E., Paster, B. J., Volcani, B. E., and Ventosa, A. (1996). Phylogenetic analyses of some extremely halophilic archaea isolated from Dead Sea water, determined on the basis of their 16S rRNA sequences. *Appl. Environ. Microbiol.* 62, 3779–3786.

Asker, D., and Ohta, Y. (2002). Haloferax alexandrinus sp. nov., an extremely halophilic canthaxanthin-producing archaeon from a solar saltern in Alexandria (Egypt). *Int. J. Syst. Evol. Microbiol.* 52, 729–738.

Averhoff, B., and Müller, V. (2010). Exploring research Future microbiology: recent advances in halophilic and thermophilic extremophiles. *Res. Microbiol.* 161, 506–514.

Baker, R. M., Singleton, F. L., and Hood, M. A. (1983). Effects of nutrient deprivation on Vibrio cholerae. *Appl. Environ. Microbiol.* 46, 930–940.

Berger, P., Barguellil, F., Raoult, D., and Drancourt, M. (2007). An outbreak of Halomonas phocaeensis sp. nov. bacteraemia in a neonatal intensive care unit. *J. Hosp. Infect.* 67, 79–85.

Biosca, E. G., Amaro, C., Marco-Noales, E., and Oliver, J. D. (1996). Effect of low temperature on starvation-survival of the eel pathogen Vibrio vulnificus biotype 2. *Appl. Environ. Microbiol.* 62, 450–455.

Borrego, J. J., Arrabal, F., De Vicente, A., Gomez, L. F., and Romero, P. (1983). Study of microbial inactivation in the marine environment. *J. Water Pollut. Control Fed.*, 297–302.

Boyko, H. (2013). *Salinity and Aridity: New Approaches to Old Problems*. Springer.

Brayton, P. R., Bode, R. B., Colwell, R. R., MacDonell, M. T., Hall, H. L., Grimes, D. J., et al. (1986). Vibrio cincinnatiensis sp. nov., a new human pathogen. *J. Clin. Microbiol.* 23, 104–108.

Brenner, D. J., Hickman-Brenner, F. W., Lee, J. V., Steigerwalt, A. G., Fanning, G. R., Hollis, D. G., et al. (1983). Vibrio furnissii (formerly aerogenic biogroup of Vibrio fluvialis), a new species isolated from human feces and the environment. *J. Clin. Microbiol.* 18, 816–824.

Cabrera R, L. E., Castro E, G., Ramírez A, M. M., Llop H, A., Llanes C, R., Castañeda E, N., et al. (2007). [Isolation and identification of species from the genera Aeromonas, Vibrio, and Plesiomonas from extraintestinal samples in Cuba]. *Rev. Chil. Infectologia Organo Of. Soc. Chil. Infectologia* 24, 204–208.

Cantet, F., Hervio-Heath, D., Caro, A., Le Mennec, C., Monteil, C., Quéméré, C., et al. (2013). Quantification of Vibrio parahaemolyticus, Vibrio vulnificus and Vibrio cholerae in French Mediterranean coastal lagoons. *Res. Microbiol.* 164, 867–874.

Carlucci, A. F., and Pramer, D. (1959). Factors affecting the survival of bacteria in sea water. *Appl. Microbiol.* 7, 388–392.

Carlucci, A. F., and Pramer, D. (1960a). An evaluation of factors affecting the survival of Escherichia coli in sea water II. Salinity, pH, and nutrients. *Appl. Microbiol.* 8, 247–250.

Carlucci, A. F., and Pramer, D. (1960b). An Evaluation of Factors Affecting the Survival of Escherichia coli in Sea Water: III. Antibiotics. *Appl. Microbiol.* 8, 251.

Carlucci, A. F., and Pramer, D. (1960c). An Evaluation of Factors Affecting the Survival of Escherichia coli in Sea Water IV. Bacteriophages. *Appl. Microbiol.* 8, 254–256.

Chan, K.-S., Cheng, K.-C., Lee, M.-F., and Yu, W.-L. (2014). A fish-stunning wound infection with acute cardiac injury. *Am. J. Emerg. Med.* 32, 289.e1-289.e2.

Chen, Z., Megharaj, M., and Naidu, R. (2007). Speciation of iodate and iodide in seawater by non-suppressed ion chromatography with inductively coupled plasma mass spectrometry. *Talanta* 72, 1842–1846.

Cheung, W. H. S., Chang, K. C. K., Hung, R. P. S., and Kleevens, J. W. L. (1990). Health effects of beach water pollution in Hong Kong. *Epidemiol. Infect.* 105, 139–162.

Choi, E. J., Lee, S. H., Jung, J. Y., and Jeon, C. O. (2013). Brevibacterium jeotgali sp. nov., isolated from jeotgal, a traditional Korean fermented seafood. *Int. J. Syst. Evol. Microbiol.* 63, 3430–3436.

Cloutier, D. D., Alm, E. W., and McLellan, S. L. (2015). Influence of land use, nutrients, and geography on microbial communities and fecal indicator abundance at Lake Michigan beaches. *Appl. Environ. Microbiol.* 81, 4904–4913.

Collins, M. D., Lawson, P. A., Labrenz, M., Tindall, B. J., Weiss, N., and Hirsch, P. (2002). Nesterenkonia lacusekhoensis sp. nov., isolated from hypersaline Ekho Lake, East Antarctica, and emended description of the genus Nesterenkonia. *Int. J. Syst. Evol. Microbiol.* 52, 1145–1150.

Culkin, F., and Cox, R. A. (1976a). Sodium, potassium, magnesium, calcium and strontium in sea water. *Deep Sea Res. Oceanogr. Abstr.* 13, 789–804.

Culkin, F., and Cox, R. A. (1976b). Sodium, potassium, magnesium, calcium and strontium in sea water. in *Deep Sea Research and Oceanographic abstract*. 13, 789–804.

DasSarma, S., and DasSarma, P. (2012). "Halophiles," in *eLS*, ed.John Wiley & Sons, Ltd (Chichester, UK: John Wiley & Sons, Ltd).

Davies-Colley, R. J., Bell, R. G., and Donnison, A. M. (1994). Sunlight inactivation of enterococci and fecal coliforms in sewage effluent diluted in seawater. *Appl. Environ. Microbiol.* 60, 2049–2058.

DePaola, A., Ulaszek, J., Kaysner, C. A., Tenge, B. J., Nordstrom, J. L., Wells, J., et al. (2003). Molecular, serological, and virulence characteristics of Vibrio parahaemolyticus isolated from environmental, food, and clinical sources in North America and Asia. *Appl. Environ. Microbiol.* 69, 3999–4005.

Diop, A., Khelaifia, S., Armstrong, N., Labas, N., Fournier, P.-E., Raoult, D., et al. (2016). Microbial culturomics unravels the halophilic microbiota repertoire of table salt: description of Gracilibacillus massiliensis sp. nov. *Microb. Ecol. Health Dis.* 27, 32049.

Dridi, B., Fardeau, M.-L., Ollivier, B., Raoult, D., and Drancourt, M. (2012). Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* 62, 1902–1907.

Dridi, B., Raoult, D., and Drancourt, M. (2011). Archaea as emerging organisms in complex human microbiomes. *Anaerobe* 17, 56–63.

Esterabadi, A. H., Entessar, F., and Khan, M. A. (1973). Isolation and Identification of Aeromonas hydrophila from an Outbreak of Haemorrhagic Septicemia in Snakes. *Can. J. Comp. Med.* 37, 418–420.

Faust, M. A., Aotaky, A. E., and Hargadon, M. T. (1975). Effect of physical parameters on the in situ survival of Escherichia coli MC-6 in an estuarine environment. *Appl. Microbiol.* 30, 800–806.

Fendrich, C. (1988). Halovibrio variabilis gen. nov. sp. nov., Pseudomonas halophila sp. nov. and a New Halophilic Aerobic Coccoid Eubacterium from Great Salt Lake, Utah, USA. *Syst. Appl. Microbiol.* 11, 36–43.

Fewtrell, L., and Kay, D. (2015). Recreational Water and Infection: A Review of Recent Findings. *Curr. Environ. Health Rep.* 2, 85–94.

Fiorillo, L., Zucker, M., Sawyer, D., and Lin, A. N. (2001). The pseudomonas hot-foot syndrome. *N. Engl. J. Med.* 345, 335–338.

FLEISHER, J. M., Jones, F., Kay, D., Stanwell-Smith, R., Wyer, M. D., and Morano, R. (1993). Water and non-water-related risk factors for gastroenteritis among bathers exposed to sewage-contaminated marine waters. *Int. J. Epidemiol.* 22, 698–708.

Fujioka, R. S., Hashimoto, H. H., Siwak, E. B., and Young, R. H. (1981). Effect of sunlight on survival of indicator bacteria in seawater. *Appl. Environ. Microbiol.* 41, 690–696.

Fukushima, T., Usami, R., and Kamekura, M. (2007). A traditional Japanese-style salt field is a niche for haloarchaeal strains that can survive in 0.5% salt solution. *Saline Syst.* 3, 2.

Furet, J.-P., Firmesse, O., Gourmelon, M., Bridonneau, C., Tap, J., Mondot, S., et al. (2009). Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. *FEMS Microbiol. Ecol.* 68, 351–362.

Gerba, C. P., Goyal, S. M., LaBelle, R. L., Cech, I., and Bodgan, G. F. (1979). Failure of indicator bacteria to reflect the occurrence of enteroviruses in marine waters. *Am. J. Public Health* 69, 1116–1119.

Gerba, C. P., and McLEOD, J. S. (1976). Effect of sediments on the survival of Escherichia coli in marine waters. *Appl. Environ. Microbiol.* 32, 114–120.

Gill, S. R., Pop, M., DeBoy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., et al. (2006). Metagenomic Analysis of the Human Distal Gut Microbiome. *Science* 312, 1355–1359.

Gonçalves, J. R., Brum, G., Fernandes, A., Biscaia, I., Correia, M. J., and Bastardo, J. (1992). Aeromonas hydrophila fulminant pneumonia in a fit young man. *Thorax* 47, 482–483.

González-Serrano, C. j., Santos, J. a., García-López, M. l., and Otero, A. (2002). Virulence markers in Aeromonas hydrophila and Aeromonas veronii biovar sobria isolates from freshwater fish and from a diarrhoea case. *J. Appl. Microbiol.* 93, 414–419.

Grégoire, P., Fardeau, M. L., Guasco, S., Bouanane, A., Michotey, V., Bonin, P., et al. (2009). LES MICRO-ORGANISMES DE L'EXTRÊME. *Presse Therm. Clim.* 146, 49–61.

Gunde-Cimerman, N., Ramos, J., and Plemenitaš, A. (2009). Halotolerant and halophilic fungi. *Mycol. Res.* 113, 1231–1241.

Hakansson, A., and Molin, G. (2011). Gut Microbiota and Inflammation. *Nutrients* 3, 637–682.

Halliday, E., and Gast, R. J. (2011). Bacteria in beach sands: an emerging challenge in protecting coastal water quality and bather health. *Environ. Sci. Technol.* 45, 370–379.

Heyndrickx, M., Lebbe, L., Kersters, K., Hoste, B., De Wachter, R., De Vos, P., et al. (1999). Proposal of Virgibacillus proomii sp. nov. and emended description of Virgibacillus pantothenticus (Proom and Knight 1950) Heyndrickx et al. 1998. *Int. J. Syst. Bacteriol.* 49 Pt 3, 1083–1090.

Heyrman, J. (2003). Virgibacillus carmonensis sp. nov., Virgibacillus necropolis sp. nov. and Virgibacillus picturae sp. nov., three novel species isolated from deteriorated mural paintings, transfer of the species of the genus Salibacillus to Virgibacillus, as Virgibacillus marismortui comb. nov. and Virgibacillus salexigens comb. nov., and emended description of the genus Virgibacillus. *Int. J. Syst. Evol. Microbiol.* 53, 501–511.

Hollis, D. G., Weaver, R. E., Baker, C. N., and Thornsberry, C. (1976). Halophilic Vibrio species isolated from blood cultures. *J. Clin. Microbiol.* 3, 425–431.

Holt, H. M., Søgaard, P., and Gahrn-Hansen, B. (1997). Ear infections with Shewanella alga: a bacteriologic, clinical and epidemiologic study of 67 cases. *Clin. Microbiol. Infect.* 3, 329–333.

Hong, S. W., Kwon, S.-W., Kim, S.-J., Kim, S. Y., Kim, J. J., Lee, J. S., et al. (2014). Bacillus oryzaecorticis sp. nov., a moderately halophilic bacterium isolated from rice husks. *Int. J. Syst. Evol. Microbiol.* 64, 2786–2791.

Horseman, M. A., and Surani, S. (2011). A comprehensive review of Vibrio vulnificus: an important cause of severe sepsis and skin and soft-tissue infection. *Int. J. Infect. Dis.* 15, e157–e166.

Howard, R. J., and Bennett, N. T. (1993). Infections caused by halophilic marine Vibrio bacteria. *Ann. Surg.* 217, 525.

Huang, J., Sun, B., and Zhang, X. (2010). Shewanella xiamenensis sp. nov., isolated from coastal sea sediment. *Int. J. Syst. Evol. Microbiol.* 60, 1585–1589.

Hugon, P., Dufour, J.-C., Colson, P., Fournier, P.-E., Sallah, K., and Raoult, D. (2015). A comprehensive repertoire of prokaryotic species identified in human beings. *Lancet Infect. Dis.* 15, 1211–1219.

Hypersaline Brines and Evaporitic Environments: Proceedings of the Bat Sheva Seminar on Saline Lakes and Natural Brines (2011). Elsevier.

Isenberg, H. D. (1988). Pathogenicity and virulence: another view. *Clin. Microbiol. Rev.* 1, 40–53.

JANSSEN, F, and SCHRUM, C. (1999). A Climatological Data Set of Temperature and Salinity for the Baltic Sea and the North Sea. *Dtsch. Hydrogr. Z. Ger. J. Hydrogr.*

Jit, S., Dadhwal, M., Prakash, O., and Lal, R. (2008). Flavobacterium lindanitolerans sp. nov., isolated from hexachlorocyclohexane-contaminated soil. *Int. J. Syst. Evol. Microbiol.* 58, 1665–1669.

Joo, W.-A., and Kim, C.-W. (2005). Proteomics of Halophilic archaea. *J. Chromatogr. B* 815, 237–250.

Joseph, S. W., Daily, O. P., Hunt, W. S., Seidler, R. J., Allen, D. A., and Colwell, R. R. (1979). Aeromonas primary wound infection of a diver in polluted waters. *J. Clin. Microbiol.* 10, 46–49.

Justé, A., Lievens, B., Klingeberg, M., Michiels, C. W., Marsh, T. L., and Willems, K. A. (2008). Predominance of Tetragenococcus halophilus as the cause of sugar thick juice degradation. *Food Microbiol.* 25, 413–421.

Kennedy, S. P., Ng, W. V., Salzberg, S. L., Hood, L., and DasSarma, S. (2001). Understanding the adaptation of Halobacterium species NRC-1 to its extreme environment through computational analysis of its genome sequence. *Genome Res.* 11, 1641–1650.

Khelaifia, S., Croce, O., Lagier, J.-C., Robert, C., Couderc, C., Di Pinto, F., et al. (2015). Noncontiguous finished genome sequence and description of Virgibacillus massiliensis sp. nov., a moderately halophilic bacterium isolated from human gut. *New Microbes New Infect.* 8, 78–88.

Khelaifia, S., and Raoult, D. (2016). Haloferax massiliensis sp. nov., the first human-associated halophilic archaea. *New Microbes New Infect.* 12, 96–98. doi:10.1016/j.nmni.2016.05.007.

Kim, J.-S., Makama, M., Petito, J., Park, N.-H., Cohan, F. M., and Dungan, R. S. (2012). Diversity of Bacteria and Archaea in hypersaline sediment from Death Valley National Park, California: Diversity of Bacteria and Archaea in Hypersaline Sediment. *MicrobiologyOpen* 1, 135–148.

Kim, K. K., Lee, K. C., Oh, H.-M., and Lee, J.-S. (2010). Halomonas stevensii sp. nov., Halomonas hamiltonii sp. nov. and Halomonas johnsoniae sp. nov., isolated from a renal care centre. *Int. J. Syst. Evol. Microbiol.* 60, 369–377.

Kim, S. B., Nedashkovskaya, O. I., Mikhailov, V. V., Han, S. K., Kim, K.-O., Rhee, M.-S., et al. (2004). Kocuria marina sp. nov., a novel actinobacterium isolated from marine sediment. *Int. J. Syst. Evol. Microbiol.* 54, 1617–1620.

Kivistö, A. T., and Karp, M. T. (2011). Halophilic anaerobic fermentative bacteria. *J. Biotechnol.* 152, 114–124.

Kobayashi, T., Kimura, B., and Fujii, T. (2000). Strictly anaerobic halophiles isolated from canned Swedish fermented herrings (Surströmming). *Int. J. Food Microbiol.* 54, 81–89.

Koh, H.-W., Rani, S., Kim, S.-J., Moon, E., Nam, S. W., Rhee, S.-K., et al. (2017). Halomonas aestuarii sp. nov., a moderately halophilic bacterium isolated from a tidal flat. *Int. J. Syst. Evol. Microbiol.*

Kriss, A. E. (1967). Microbial population of oceans and seas. Available at: http://agris.fao.org/agris-search/search.do?recordID=US201300591431 [Accessed March 7, 2017].

Kuda, T., Izawa, Y., Yoshida, S., Koyanagi, T., Takahashi, H., and Kimura, B. (2014). Rapid identification of Tetragenococcus halophilus and Tetragenococcus muriaticus, important species in the production of salted and fermented foods, by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS). *Food Control* 35, 419–425.

Lagier, J.-C., 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, J.-C., Hugon, P., Khelaifia, S., Fournier, P.-E., La Scola, B., and 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, J.-C., Khelaifia, S., Alou, M. T., 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.

Lee, H.-S. (2013). Diversity of halophilic archaea in fermented foods and human intestines and their application. *J. Microbiol. Biotechnol.* 23, 1645–1653.

Lee, J.-S., Lim, J.-M., Lee, K. C., Lee, J.-C., Park, Y.-H., and Kim, C.-J. (2006). Virgibacillus koreensis sp. nov., a novel bacterium from a salt field, and transfer of Virgibacillus picturae to the genus Oceanobacillus as Oceanobacillus picturae comb. nov. with emended descriptions. *Int. J. Syst. Evol. Microbiol.* 56, 251–257.

Leveque, N., and Laurent, A. (2008). *Editorial Commentary:* A Novel Mode of Transmission for Human Enterovirus Infection Is Swimming in Contaminated Seawater: Implications in Public Health and in Epidemiological Surveillance. *Clin. Infect. Dis.* 47, 624–626.

LIM, C.-H., and Flint, K. P. (1989). The effects of nutrients on the survival of Escherichia coli in lake water. *J. Appl. Bacteriol.* 66, 559–569.

Lippi, D., Gotuzzo, E., and Caini, S. (2016). Cholera. *Microbiol. Spectr.* 4. doi:10.1128/microbiolspec.PoH-0012-2015.

Liu, W. Y., Zeng, J., Wang, L., Dou, Y. T., and Yang, S. S. (2005). Halobacillus dabanensis sp. nov. and Halobacillus aidingensis sp. nov., isolated from salt lakes in Xinjiang, China. *Int. J. Syst. Evol. Microbiol.* 55, 1991–1996.

Lo, N., Lee, S. H., Jin, H. M., Jung, J. Y., Schumann, P., and Jeon, C. O. (2015). Garicola koreensis gen. nov., sp. nov., isolated from saeu-jeot, traditional Korean fermented shrimp. *Int. J. Syst. Evol. Microbiol.* 65, 1015–1021.

Lu, J., Nogi, Y., and Takami, H. (2001). Oceanobacillus iheyensis gen. nov., sp. nov., a deep-sea extremely halotolerant and alkaliphilic species isolated from a depth of 1050 m on the Iheya Ridge. *FEMS Microbiol. Lett.* 205, 291–297.

Ma, Y., Galinski, E. A., Grant, W. D., Oren, A., and Ventosa, A. (2010). Halophiles 2010: Life in Saline Environments. *Appl. Environ. Microbiol.* 76, 6971–6981.

Madigan, M. T., and Oren, A. (1999). Thermophilic and halophilic extremophiles. *Curr. Opin. Microbiol.* 2, 265–269.

Marras, T. K., Wallace, R. J., Koth, L. L., Stulbarg, M. S., Cowl, C. T., and Daley, C. L. (2005). Hypersensitivity pneumonitis reaction to Mycobacterium avium in household water. *Chest* 127, 664–671.

Maslin, M. A., Shackleton, N. J., and Pflaumann, U. (1995). Surface water temperature, salinity, and density changes in the northeast Atlantic during the last 45,000 years: Heinrich events, deep water formation, and climatic rebounds. *Paleoceanography* 10, 527–544.

Masoumi-Asl, H., Gouya, M. M., Rahbar, M., and Sabourian, R. (2016). The epidemiology and antimicrobial resistance of cholera cases in Iran during 2013. *Iran. J. Microbiol.* 8, 232–237.

McLaughlin, J. B., DePaola, A., Bopp, C. A., Martinek, K. A., Napolilli, N. P., Allison, C. G., et al. (2005). Outbreak of Vibrio parahaemolyticus Gastroenteritis Associated with Alaskan Oysters. *N. Engl. J. Med.* 353, 1463–1470.

Millero, F. J., Feistel, R., Wright, D. G., and McDougall, T. J. (2008). The composition of Standard Seawater and the definition of the Reference-Composition Salinity Scale. *Deep Sea Res. Part Oceanogr. Res. Pap.* 55, 50–72.

Minegishi, H., Echigo, A., Nagaoka, S., Kamekura, M., and Usami, R. (2010). Halarchaeum acidiphilum gen. nov., sp. nov., a moderately acidophilic haloarchaeon isolated from commercial solar salt. *Int. J. Syst. Evol. Microbiol.* 60, 2513–2516.

Minegishi, H., Mizuki, T., Echigo, A., Fukushima, T., Kamekura, M., and Usami, R. (2008). Acidophilic haloarchaeal strains are isolated from various solar salts. *Saline Syst.* 4, 16.

Mojica, F. J., Cisneros, E., Ferrer, C., Rodriguez-Valera, F., and Juez, G. (1997). Osmotically induced response in representatives of halophilic prokaryotes: the bacterium Halomonas elongata and the archaeon Haloferax volcanii. *J. Bacteriol.* 179, 5471–5481.

Moreno, M. de L., García, M. T., Ventosa, A., Iglesias-Guerra, F., and Mellado, E. (2010). The extremely halophilic bacterium Salicola marasensis IC10 accumulates the compatible solute betaine. *Syst. Appl. Microbiol.* 33, 308–310.

Muller, V., and Oren, A. (2003). Metabolism of chloride in halophilic prokaryotes. *Extremophiles* 7, 261–266.

Nagaoka, S., Minegishi, H., Echigo, A., Shimane, Y., Kamekura, M., and Usami, R. (2011). Halostagnicola alkaliphila sp. nov., an alkaliphilic haloarchaeon from commercial rock salt. *Int. J. Syst. Evol. Microbiol.* 61, 1149–1152.

Nam, J.-H., Bae, W., and Lee, D.-H. (2008a). Oceanobacillus caeni sp. nov., isolated from a Bacillus-dominated wastewater treatment system in Korea. *Int. J. Syst. Evol. Microbiol.* 58, 1109–1113.

Nam, Y.-D., Chang, H.-W., Kim, K.-H., Roh, S. W., Kim, M.-S., Jung, M.-J., et al. (2008b). Bacterial, archaeal, and eukaryal diversity in the intestines of Korean people. *J. Microbiol.* 46, 491–501.

Namwong, S., Tanasupawat, S., Kudo, T., and Itoh, T. (2011). Haloarcula salaria sp. nov. and Haloarcula tradensis sp. nov., isolated from salt in Thai fish sauce. *Int. J. Syst. Evol. Microbiol.* 61, 231–236.

Namwong, S., Tanasupawat, S., Visessanguan, W., Kudo, T., and Itoh, T. (2007). Halococcus thailandensis sp. nov., from fish sauce in Thailand. *Int. J. Syst. Evol. Microbiol.* 57, 2199–2203.

Niemi, T. M., Ben-Avraham, Z., and Gat, J. (1997). *The Dead Sea: The Lake and Its Setting*. Oxford University Press.

Noguchi, H., Uchino, M., Shida, O., Takano, K., Nakamura, L. K., and Komagata, K. (2004). Bacillus vietnamensis sp. nov., a moderately halotolerant, aerobic, endospore-forming bacterium isolated from Vietnamese fish sauce. *Int. J. Syst. Evol. Microbiol.* 54, 2117–2120.

Nunes, I., Tiago, I., Pires, A. L., da Costa, M. S., and Veríssimo, A. (2006). Paucisalibacillus globulus gen. nov., sp. nov., a Gram-positive bacterium isolated from potting soil. *Int. J. Syst. Evol. Microbiol.* 56, 1841–1845.

Odeyemi, O. A. (2016). Incidence and prevalence of Vibrio parahaemolyticus in seafood: a systematic review and meta-analysis. *SpringerPlus* 5.

Oliver, J. D. (2005). Wound infections caused by Vibrio vulnificus and other marine bacteria. *Epidemiol. Infect.* 133, 383–391.

Ong, K. R., Sordillo, E., and Frankel, E. (1991). Unusual case of Aeromonas hydrophila endocarditis. *J. Clin. Microbiol.* 29, 1056–1057.

Oren, A. (2002). Molecular ecology of extremely halophilic Archaea and Bacteria. *FEMS Microbiol. Ecol.* 39, 1–7.

Oren, A. (2008). Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Syst.* 4, 2.

Oren, A. (2013a). Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. *Front. Microbiol.* 4, 315.

Oren, A. (2013b). *Salinibacter* : an extremely halophilic bacterium with archaeal properties. *FEMS Microbiol. Lett.* 342, 1–9.

Oren, A. (2015). Halophilic microbial communities and their environments. *Curr. Opin. Biotechnol.* 33, 119–124.

Otsuka, T., Noda, T., Noguchi, A., Nakamura, H., Ibaraki, K., and Yamaoka, K. (2007). Shewanella infection in decompensated liver disease: a septic case. *J. Gastroenterol.* 42, 87–90.

Oxley, A. P. A., Lanfranconi, M. P., Würdemann, D., Ott, S., Schreiber, S., McGenity, T. J., et al. (2010). Halophilic archaea in the human intestinal mucosa. *Environ. Microbiol.* 12, 2398–2410.

Phillips, M. C., Solo-Gabriele, H. M., Piggot, A. M., Klaus, J. S., and Zhang, Y. (2011). Relationships between sand and water quality at recreational beaches. *Water Res.* 45, 6763–6769..

Purdy, K. J., Cresswell-Maynard, T. D., Nedwell, D. B., McGenity, T. J., Grant, W. D., Timmis, K. N., et al. (2004). Isolation of haloarchaea that grow at low salinities. *Environ. Microbiol.* 6, 591–595.

Rampelotto, P. H. (2016). *Biotechnology of Extremophiles:: Advances and Challenges*. Springer.

Robert, A., Danin, P.-É., Quintard, H., Degand, N., Martis, N., Doyen, D., et al. (2017). Seawater drowning-associated pneumonia: a 10-year descriptive cohort in intensive care unit. *Ann. Intensive Care* 7, 45.

Roberts, M. F. (2005). Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Syst.* 1, 5.

Roh, S. W., and Bae, J.-W. (2009). Halorubrum cibi sp. nov., an extremely halophilic archaeon from salt-fermented seafood. *J. Microbiol. Seoul Korea* 47, 162–166.

Roh, S. W., Kim, K.-H., Nam, Y.-D., Chang, H.-W., Park, E.-J., and Bae, J.-W. (2010a). Investigation of archaeal and bacterial diversity in fermented seafood using barcoded pyrosequencing. *ISME J.* 4, 1–16.

Roh, S. W., Lee, M.-L., and Bae, J.-W. (2010b). Haladaptatus cibarius sp. nov., an extremely halophilic archaeon from seafood, and emended description of the genus Haladaptatus. *Int. J. Syst. Evol. Microbiol.* 60, 1187–1190.

Roh, S. W., Nam, Y.-D., Chang, H.-W., Sung, Y., Kim, K.-H., Lee, H.-J., et al. (2007a). Natronococcus jeotgali sp. nov., a halophilic archaeon isolated from shrimp jeotgal, a traditional fermented seafood from Korea. *Int. J. Syst. Evol. Microbiol.* 57, 2129–2131.

Roh, S. W., Nam, Y.-D., Chang, H.-W., Sung, Y., Kim, K.-H., Oh, H.-M., et al. (2007b). Halalkalicoccus jeotgali sp. nov., a halophilic archaeon from shrimp jeotgal, a traditional Korean fermented seafood. *Int. J. Syst. Evol. Microbiol.* 57, 2296–2298.

Romano, I., Giordano, A., Lama, L., Nicolaus, B., and Gambacorta, A. (2003). Planococcus rifietensis sp. nov, Isolated from Algal Mat Collected from a Sulfurous Spring in Campania (Italy). *Syst. Appl. Microbiol.* 26, 357–366.

Romano, I., Lama, L., Nicolaus, B., Poli, A., Gambacorta, A., and Giordano, A. (2006). Halomonas alkaliphila sp. nov., a novel halotolerant alkaliphilic bacterium isolated from a salt pool in Campania (Italy). *J. Gen. Appl. Microbiol.* 52, 339–348.

Rosenfeld, W. D., and ZoBell, C. E. (1947). Antibiotic production by marine microorganisms. *J. Bacteriol.* 54, 393.

Rosenthal, S. L., Zuger, J. H., and Apollo, E. (1975). Respiratory colonization with Pseudomonas putrefaciens after near-drowning in salt water. *Am. J. Clin. Pathol.* 64, 382–384.

Rozen, Y., and Belkin, S. (2001). Survival of enteric bacteria in seawater. *FEMS Microbiol. Rev.* 25, 513–529.

Sánchez-Porro, C., Tokunaga, H., Tokunaga, M., and Ventosa, A. (2007). Chromohalobacter japonicus sp. nov., a moderately halophilic bacterium isolated from a Japanese salty food. *Int. J. Syst. Evol. Microbiol.* 57, 2262–2266.

Schulze-Röbbecke, R. (1993). [Mycobacteria in the environment]. *Immun. Infekt.* 21, 126–131.

Seck, E. H., Fournier, P.-E., Raoult, D., and Khelaifia, S. (2016). “Halomonas massiliensis” sp. nov., a new halotolerant bacterium isolated from the human gut. *New Microbes New Infect.* 14, 19–20.

Seck, E., Rathored, J., Khelaifia, S., Croce, O., Robert, C., Couderc, C., et al. (2015). Virgibacillus senegalensis sp. nov., a new moderately halophilic bacterium isolated from human gut. *New Microbes New Infect.* 8, 116–126.

Senghor, B., Seck, E. H., Khelaifia, S., Bassène, H., Sokhna, C., Fournier, P.-E., et al. (2017). Description of “Bacillus dakarensis” sp. nov., “Bacillus sinesaloumensis” sp. nov., “Gracilibacillus timonensis” sp. nov., “Halobacillus massiliensis” sp. nov., “Lentibacillus massiliensis” sp. nov., “Oceanobacillus senegalensis” sp. nov., “Oceanobacillus timonensis” sp. nov., “Virgibacillus dakarensis” sp. nov. and “Virgibacillus marseillensis” sp. nov., nine halophilic new species isolated from human stool. *New Microbes New Infect.* 17, 45–51.

Shimane, Y., Hatada, Y., Minegishi, H., Echigo, A., Nagaoka, S., Miyazaki, M., et al. (2011). Salarchaeum japonicum gen. nov., sp. nov., an aerobic, extremely halophilic member of the Archaea isolated from commercial salt. *Int. J. Syst. Evol. Microbiol.* 61, 2266–2270.

Shimane, Y., Hatada, Y., Minegishi, H., Mizuki, T., Echigo, A., Miyazaki, M., et al. (2010). Natronoarchaeum mannanilyticum gen. nov., sp. nov., an aerobic, extremely halophilic archaeon isolated from commercial salt. *Int. J. Syst. Evol. Microbiol.* 60, 2529–2534.

Shimoshige, H., Yamada, T., Minegishi, H., Echigo, A., Shimane, Y., Kamekura, M., et al. (2013). Halobaculum magnesiiphilum sp. nov., a magnesium-dependent haloarchaeon isolated from commercial salt. *Int. J. Syst. Evol. Microbiol.* 63, 861–866.

Shuval, H. (2003). Estimating the global burden of thalassogenic diseases: human infectious diseases caused by wastewater pollution of the marine environment. *J. Water Health* 1, 53–64.

Siglioccolo, A., Paiardini, A., Piscitelli, M., and Pascarella, S. (2011). Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. *BMC Struct. Biol.* 11, 50.

Sims, J. K., Enomoto, P. I., Frankel, R. I., and Wong, L. M. (1983). Marine bacteria complicating seawater near-drowning and marine wounds: a hypothesis. *Ann. Emerg. Med.* 12, 212–216.

Slifka, K. M. J., Newton, A. E., and Mahon, B. E. (2017). Vibrio alginolyticus infections in the USA, 1988–2012. *Epidemiol. Amp Infect.*, 1–9.

Soller, J. A., Schoen, M. E., Bartrand, T., Ravenscroft, J. E., and Ashbolt, N. J. (2010a). Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. *Water Res.* 44, 4674–4691.

Soller, J. A., Schoen, M. E., Bartrand, T., Ravenscroft, J. E., and Ashbolt, N. J. (2010b). Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. *Water Res.* 44, 4674–4691.

SPRING, S., LUDWIG, W., MARQUEZ, M. C., VENTOSA, A., and SCHLEIFER, K.-H. (1996). Halobacillus gen. nov., with Descriptions of Halobacillus litoralis sp. nov. and Halobacillus trueperi sp. nov., and Transfer of Sporosarcina halophila to Halobacillus halophilus comb. nov. *Int. J. Syst. Evol. Microbiol.* 46, 492–496.

STACKEBRANDT, E., KOCH, C., GVOZDIAK, O., and SCHUMANN, P. (1995). Taxonomic Dissection of the Genus Micrococcus: Kocuria gen. nov., Nesterenkonia gen. nov., Kytococcus gen. nov., Dermacoccus gen. nov., and Micrococcus Cohn 1872 gen. emend. *Int. J. Syst. Evol. Microbiol.* 45, 682–692.

Tamang, J. P., Watanabe, K., and Holzapfel, W. H. (2016). Review: Diversity of Microorganisms in Global Fermented Foods and Beverages. *Front. Microbiol.* 7..

Tapingkae, W., Tanasupawat, S., Itoh, T., Parkin, K. L., Benjakul, S., Visessanguan, W., et al. (2008). Natrinema gari sp. nov., a halophilic archaeon isolated from fish sauce in Thailand. *Int. J. Syst. Evol. Microbiol.* 58, 2378–2383.

Tcherpakov, M., Ben-Jacob, E., and Gutnick, D. L. (1999). Paenibacillus dendritiformis sp. nov., proposal for a new pattern-forming species and its localization within a phylogenetic cluster. *Int. J. Syst. Bacteriol.* 49 Pt 1, 239–246.

Ten, L. N., Baek, S.-H., Im, W.-T., Larina, L. L., Lee, J.-S., Oh, H.-M., et al. (2007). Bacillus pocheonensis sp. nov., a moderately halotolerant, aerobic bacterium isolated from soil of a ginseng field. *Int. J. Syst. Evol. Microbiol.* 57, 2532–2537.

Thorsteinsson, S. B., Minuth, J. N., and Musher, D. M. (1974). Clinical manifestations of halophilic non-cholera Vibrio infections. *Lancet Lond. Engl.* 2, 1283–1284.

Tian, X.-P., Zhi, X.-Y., Qiu, Y.-Q., Zhang, Y.-Q., Tang, S.-K., Xu, L.-H., et al. (2009). Sciscionella marina gen. nov., sp. nov., a marine actinomycete isolated from a sediment in the northern South China Sea. *Int. J. Syst. Evol. Microbiol.* 59, 222–228.

Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L., and Fowler, V. G. (2015). Staphylococcus aureus Infections: Epidemiology, Pathophysiology, Clinical Manifestations, and Management. *Clin. Microbiol. Rev.* 28, 603–661.

Van Donsel, D. J., Geldreich, E. E., and Clarke, N. A. (1967). Seasonal variations in survival of indicator bacteria in soil and their contribution to storm-water pollution. *Appl. Microbiol.* 15, 1362–1370.

Vartoukian, S. R., Palmer, R. M., and Wade, W. G. (2010). Strategies for culture of "unculturable" bacteria. *FEMS Microbiol. Lett.* 309, 1–7.

Ventosa, A., de la Haba, R. R., Sánchez-Porro, C., and Papke, R. T. (2015). Microbial diversity of hypersaline environments: a metagenomic approach. *Curr. Opin. Microbiol.* 25, 80–87.

Vezzulli, L., Colwell, R. R., and Pruzzo, C. (2013). Ocean Warming and Spread of Pathogenic Vibrios in the Aquatic Environment. *Microb. Ecol.* 65, 817–825.

Vignier, N., Barreau, M., Olive, C., Baubion, E., Théodose, R., Hochedez, P., et al. (2013). Human Infection with Shewanella putrefaciens and S. algae: Report of 16 Cases in Martinique and Review of the Literature. *Am. J. Trop. Med. Hyg.* 89, 151–156.

Villar, M., de Ruiz Holgado, A. P., Sanchez, J. J., Trucco, R. E., and Oliver, G. (1985). Isolation and characterization of Pediococcus halophilus from salted anchovies (Engraulis anchoita). *Appl. Environ. Microbiol.* 49, 664–666.

Voipio, A. (1981). *The Baltic Sea*. Elsevier.

von Graevenitz, A., Bowman, J., Del Notaro, C., and Ritzler, M. (2000). Human infection with Halomonas venusta following fish bite. *J. Clin. Microbiol.* 38, 3123–3124.

Wang, L., Liu, W.-Y., Gu, Z.-J., Chen, S.-F., and Yang, S.-S. (2010). Oceanobacillus manasiensis sp. nov., a moderately halophilic bacterium isolated from the salt lakes of Xinjiang, China. *J. Microbiol. Seoul Korea* 48, 312–317.

Whitman, R. L., Harwood, V. J., Edge, T. A., Nevers, M. B., Byappanahalli, M., Vijayavel, K., et al. (2014). Microbes in beach sands: integrating environment, ecology and public health. *Rev. Environ. Sci. Biotechnol.* 13, 329–368.

Whitman, R. L., Nevers, M. B., Korinek, G. C., and Byappanahalli, M. N. (2004). Solar and Temporal Effects on Escherichia coli Concentration at a Lake Michigan Swimming Beach. *Appl. Environ. Microbiol.* 70, 4276–4285.

Wu, Y.-H., Xu, X.-W., Huo, Y.-Y., Zhou, P., Zhu, X.-F., Zhang, H.-B., et al. (2008). Halomonas caseinilytica sp. nov., a halophilic bacterium isolated from a saline lake on the Qinghai-Tibet Plateau, China. *Int. J. Syst. Evol. Microbiol.* 58, 1259–1262. doi:10.1099/ijs.0.65381-0.

Yamahara, K. M., Sassoubre, L. M., Goodwin, K. D., and Boehm, A. B. (2012). Occurrence and Persistence of Bacterial Pathogens and Indicator Organisms in Beach Sand along the California Coast. *Appl. Environ. Microbiol.* 78, 1733–1745.

Yoon, J.-H., and Oh, T.-K. (2005). Bacillus litoralis sp. nov., isolated from a tidal flat of the Yellow Sea in Korea. *Int. J. Syst. Evol. Microbiol.* 55, 1945–1948.

Yoon, J.-H., Oh, T.-K., and Park, Y.-H. (2004). Transfer of Bacillus halodenitrificans Denariaz et al. 1989 to the genus Virgibacillus as Virgibacillus halodenitrificans comb. nov. *Int. J. Syst. Evol. Microbiol.* 54, 2163–2167.

Yumoto, I. (2005). Oceanobacillus oncorhynchi sp. nov., a halotolerant obligate alkaliphile isolated from the skin of a rainbow trout (Oncorhynchus mykiss), and emended description of the genus Oceanobacillus. *Int. J. Syst. Evol. Microbiol.* 55, 1521–1524.

Yumoto, I., Hirota, K., Goto, T., Nodasaka, Y., and Nakajima, K. (2005). Bacillus oshimensis sp. nov., a moderately halophilic, non-motile alkaliphile. *Int. J. Syst. Evol. Microbiol.* 55, 907–911.

Zhang, Y.-J., Tang, S.-K., Shi, R., Klenk, H.-P., Chen, C., Yang, L.-L., et al. (2012). Salinisphaera halophila sp. nov., a moderately halophilic bacterium isolated from brine of a salt well. *Int. J. Syst. Evol. Microbiol.* 62, 2174–2179.

Zobell, C. E., Anderson, D. Q., and Smith, W. W. (1937). The bacteriostatic and bactericidal action of Great Salt Lake water. *J. Bacteriol.* 33, 253.

Table 1: Halophilic prokaryotes isolated from food

Strain name	Source food	Country	Optimal NaCl	Domains	Refrences
Bacillus oryzaecorticis	Rice husks	korea	6-7%	Bacteria	(Hong et al., 2014)
Brevibacterium jeotgali	Seafood	Korea	5 %	Bacteria	(Choi et al., 2013)
Garicola koreensis	Fermented shrimp	Korea	6%	Bacteria	(Lo et al., 2015)
Gracilibacillus dipsosauri	Salt table	France	7,5-10%	Bacteria	(Diop et al., 2016)
Gracilibacillus massiliensis sp.nov	Salt table	France	7,5-10%	Bacteria	(Diop et al., 2016)
Haladaptatus cibarius	Fermented shellfish	Korea	10–30%	archea	(Roh et al., 2010b)
Halalkalicoccus jeotgali	Shrimp jeotgal	Korea	10%	archea	(Roh et al., 2007b)
Halarchaeaum acidiphilum	Commercial salt	Japan	21-24%	archea	(Minegishi et al., 2010)
Haloacular tradensis	Thai fish sauce	Thailand	20-30%	archea	(Lee, 2013)
Haloanaerobium praevalens	Surströmming	Sweden	1.7–2.5M	Bacteria	(Kobayashi et al., 2000)
Haloarcula salaria	Thai fish sauce	Thailand	20-25%	archea	(Namwong et al., 2011)
Halobacillus trueperi	Salt table	France	7,5-10 %	Bacteria	(Diop et al., 2016)
Halobacterium salinarum	Fermented fish	Thailand	3.4-4.3M	archea	(Namwong et al., 2011)
Halobaculum magnesiiphilum	Commercial salt	Japan	9-12%	archea	(Shimoshige et al., 2013)
Halococcus thailandensis	Thai fish sauce	Thailand	20-30%	archea	(Namwong et al., 2007)
Halogranum salarium	Sea salt	Korea	2.6-3.4M	archea	(Lee, 2013)
Halorubrum cibi	Shrimp jeotgal	Korea	23-25%	archea	(Roh and Bae, 2009)
Halorubrum orientalis	Olive	USA, Spain	3,4M	archea	(Tamang et al., 2016)
Halosarcina pallida	Olive	USA, Spain	3,4 M	archea	(Tamang et al., 2016)
Halostagnicola alkaliphila	Commercial rock salt	Japan	0,25M	archea	(Nagaoka et al., 2011)
Haloterrigena jeotgali	Shrimp jeotgal	Korea	15-20%	archea	(Nagaoka et al., 2011)
Micrococcus halobius	Miso	Japan	1-4%	Bacteria	(Tamang et al., 2016)
Natrinema gari	Fish sauce (nam-pla)	Thailand	2.6-3.4M	archea	(Tapingkae et al., 2008)
Natronoarchaeaum mannanilyticum	Commercial salt	Japan	4.0-4.5 M	archea	(Shimane et al., 2010)
Natronococcus jeotgali	Shrimp jeotgal	Korea	23-25%	archea	(Roh et al., 2007a)
Oceanobacillus picturae	Salt table	France	7,5-10 %	Bacteria	(Diop et al., 2016)
Paraliobacillus quinghaiensis	Salt table	France	7,5-10 %	Bacteria	(Diop et al., 2016)
Pediococcus halophilus	Salted anchovies	Philippines, Indonesia	6,5-10 %	Bacteria	(Villar et al., 1985)
Salarchaeaum japonicum	Commercial salt	Japan	2.5-3.0 M	archea	(Shimane et al., 2011)
Tetragenococcus halophilus	Fish	India	5-10 %	Bacteria	(Kuda et al., 2014)
Thalassobacillus devorans	Salt table	France	7,5-10 %	Bacteria	(Diop et al., 2016)
Thalassomonas agarivorans	Olives	Perou	0,03	Bacteria	(Tamang et al., 2016)
Vibrio parahaemolyticus	seafood	USA	10%	Bacteria	(DePaola et al., 2003)
Virgibacillus picturae	Salt table	France	7,5-10 %	Bacteria	(Diop et al., 2016)

Table 2: Query patterns syntax

Query Patterns	Syntax
QP1	#3[tiab] AND ((seawater[mesh] OR seawater*[tiab] OR sea water*[tiab] OR saline lake[tiab]) OR (halophilic[tiab] OR halophile*[tiab]) OR (Sodium Chloride[mesh] OR Saline Solution[tiab] OR Sodium Chloride[tiab] OR NaCl[tiab]) OR (soil microbiology[mesh]))
QP2	#3[tiab] AND bacteria[Mesh] AND ((seawater[mesh] OR seawater*[tiab] OR sea water*[tiab] OR saline lake[tiab]) OR (halophilic[tiab] OR halophile*[tiab]) OR (Sodium Chloride[mesh] OR Saline Solution[tiab] OR Sodium Chloride[tiab] OR NaCl[tiab]) OR (soil microbiology[mesh]))
QP3	#3[tiab] AND #2/isolation and purification[mesh] AND ((halophilic[tiab] OR halophile*[tiab]) OR (Sodium Chloride[mesh] OR Saline Solution[tiab] OR Sodium Chloride[tiab] OR NaCl[tiab]))

3 and # 2 were replaced dynamically (during program execution) for each species:

#3 = bacterium species name OR NCBI Taxonomy synonym(s) if any

#2 = MeSH term corresponding to the species name. If no MeSH terms exist for the species, the MeSH term corresponding to nearest phylum taxon from NCBI Taxonomy is chosen.

Table 3: Halophilic prokaryotes cultured at least once in human

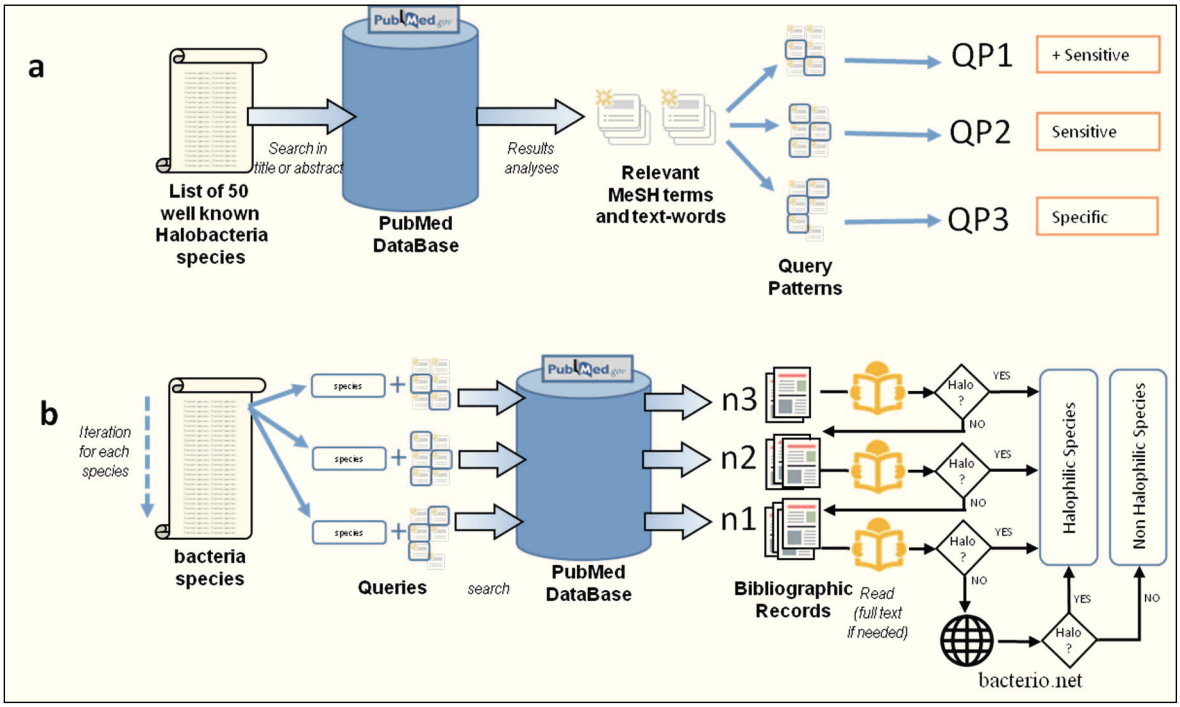
Species	Phylum	Localisation	Sodium requirement (%)	Category	PMIDList	Reference
Bacillus aquimaris	*Proteobacteria*	H(Gut)	2-5	Slight halophilic	12791884	(Choi et al., 2013)
Bacillus jeddahtimonense	*Proteobacteria*	H(Gut)	5-10	Moderately halophilic	27819657	(Lagier et al., 2016)
Bacillus jeotgali	*Proteobacteria*	H(Gut)	2-13	Moderately halophilic	11411677	(Yoon and Oh, 2005)
Bacillus litoralis	*Actinobacteria*	H(Gut)	2-3	Slight halophilic	Bacterio.net	(Yoon and Oh, 2005)
Bacillus marisflavi	*Actinobacteria*	H(Gut)	2-3	Slight halophilic	13130010	(Diop et al., 2016)
Bacillus oshimensis	*Actinobacteria*	H(Gut)	7	Moderately halophilic	15774684	(Yumoto et al., 2005)
Bacillus vietnamensis	*Actinobacteria*	H(Gut)	1	Slight halophilic	15879267	(Noguchi et al., 2004)
Chromohalobacter japonicus	*Proteobacteria*	H(Gut)	7,5-12,5	Moderately halophilic	17911293	(Sánchez-Porro et al., 2007)
Flavobacterium lindanitolerans	*Bacteroidetes*	H(Gut)	1-2	Slight halophilic	18599713	(Jit et al., 2008)
Gracilibacillus timonensis	*Firmicutes*	H(Gut)	5-10	Moderately halophilic	5333509	(Senghor et al., 2017)
Halobacillus dabanensis	*Firmicutes*	H(Gut)	0,5-25	Moderately halophilic	16166700	(Liu et al., 2005)
Halobacillus karajensis	*Firmicutes*	H(Gut)	1-24	Moderately halophilic	12892126	(Amoozegar et al., 2003)
Halobacillus trueperi	*Firmicutes*	H(Gut)	0,5-30	Moderately halophilic	27819657	(Spiring et al., 1996)
Haloferax alexandrinus	*Euryarchaeaota*	H(Gut)	25-30	extremely halophilic	12054232	(Asker and Ohta, 2002)
Haloferax massiliensis	*Euryarchaeaota*	H(Gut)	25-30	extremely halophilic	27408734	(Khelaifia and Raoult, 2016)
Halomonas hamiltonii	*Proteobacteria*	H	5-7,5	Moderately halophilic	19651714	(Kim et al., 2010)
Halomonas massiliensis	*Proteobacteria*	H(Gut)	5-10	Moderately halophilic	27621824	(Seck et al., 2016)
Halomonas salina	*Proteobacteria*	H(Gut)	1,7-12	Moderately halophilic	27819657	(Lagier et al., 2016)
Halomonas stevensii	*Proteobacteria*	H	3-7,5	Moderately halophilic	19651714	(Kim et al., 2010)
Halomonas venusta	*Proteobacteria*	H	1-12	Moderately halophilic	Bacterio.net	(von Graevenitz et al., 2000)
Kocuria marina	*Actinobacteria*	H(Gut)	15	Moderately halophilic	15388718	(Kim et al., 2004)
Methanomassiliicoccus luminyensis	*Euryarchaeaota*	H (gut)	0.5-1	Slight halophilic	22859731	(Dridi et al., 2012)
Nesterenkonia halobia	*Actinobacteria*	H	7,5	Moderately halophilic	Bacterio.net	(STACKEBRANDT et al., 1995)

Nesterenkonia lacusekhoensis	*Actinobacteria*	H	6-8	Moderately halophilic	12148619	(Collins et al., 2002)
Oceanobacillus caeni	*Firmicutes*	H(Gut)	2-5	Slight halophilic	18450698	(Nam et al., 2008a)
Oceanobacillus iheyensis	*Firmicutes*	H(Gut)	0,5-21	Moderately halophilic	11750818	(Lu et al., 2001)
Oceanobacillus manasiensis	*Firmicutes*	H(Gut)	5-10	Moderately halophilic	20571948	(Wang et al., 2010)
Oceanobacillus oncorhynchi	*Firmicutes*	H(Gut)	5--20	Moderately halophilic	16014475	(Yumoto, 2005)
Oceanobacillus picturae	*Firmicutes*	H(Gut)	5-10	Moderately halophilic	16403894	(Heyrman, 2003)
Oceanobacillus sojae	*Firmicutes*	H(Gut)	0,5-15	Moderately halophilic	27819657	(Lagier et al., 2016)
Paenibacillus dendritiformis	*Firmicutes*	H(Gut)	1-5	Slight halophilic	10028268	(Tcherpakov et al., 1999)
Paucisalibacillus globulus	*Firmicutes*	H(Gut)	1	Slight halophilic	16902018	(Nunes et al., 2006)
Planococcus rifietoensis	*Firmicutes*	H(Gut)	10	Moderately halophilic	27819657	(Romano et al., 2003)
Planomicrobium okeanokoites	*Firmicutes*	H(Gut)	1-4	Slight halophilic	11491353	(Lee et al., 2006)
Pseudomonas halophila	*Proteobacteria*	H(Gut)	1-25	Moderately halophilic	27819657	(Fendrich, 1988)
Salinisphaera halophila	*Proteobacteria*	H(Gut)	6-9	Moderately halophilic	22058321	(Zhang et al., 2012)
Sciscionella marina	*Actinobacteria*	H	3-5	Slight halophilic	19196758	(Tian et al., 2009)
Shewanella xiamenensis	*Proteobacteria*	H	1-2	Slight halophilic	19684319	(Huang et al., 2010)
Sporosarcina globispora	*Firmicutes*	H	2,4-13	Moderately halophilic	11411676	(Spring et al., 1996)
Terribacillus halophilus	*Firmicutes*	H(Gut)	5-10	Moderately halophilic	17220440	(An et al., 2007)
Terribacillus saccharophilus	*Firmicutes*	H	1-5	Slight halophilic	17220440	(An et al., 2007)
Thalassobacillus massiliensis	*Firmicutes*	H(Gut)	5-10	Moderately halophilic	27819657	(Lagier et al., 2016)
Vibrio cincinnatiensis	*Proteobacteria*	H(Gut)	0,5-6	Slight halophilic	2422196	(Brayton et al., 1986)
Vibrio fluvialis	*Proteobacteria*	H(Gut)	1-3	Slight halophilic	6971864	(Brenner et al., 1983)
Vibrio furnissii	*Proteobacteria*	H(Gut)	1-6	Slight halophilic	6630464	(Brenner et al., 1983)
Virgibacillus halodenitrificans	*Firmicutes*	H(Gut)	5-10	Moderately halophilic	15545452	(Yoon et al., 2004)
Virgibacillus massiliensis	*Firmicutes*	H(Gut)	5-10	Moderately halophilic	26649181	(Khelaifia et al., 2015)
Virgibacillus pantothenticus	*Firmicutes*	H(Gut)	4-10	Moderately halophilic	10425765	(Heyndrickx et al., 1999)
Virgibacillus promii	*Firmicutes*	H(Gut)	1-3	Slight halophilic	10425765	(Heyndrickx et al., 1999)
Virgibacillus senegalensis	*Firmicutes*	H(Gut)	5-10	Moderately halophilic	26693281	(Seck et al., 2015)

Table 4: Most frequently halophilic and halotolerant prokaryote implicated in human diseases contracted by sea water (Berger et al., 2007, 2007; Cabrera R et al., 2007; Chan et al., 2014; Esterabadi et al., 1973; Fewtrell and Kay, 2015; Fleisher et al., 1993; Gonçalves et al., 1992; Hakansson and Molin, 2011; Hollis et al., 1976; Holt et al., 1997; Horseman and Surani, 2011; Howard and Bennett, 1993; Hugon et al., 2015; Isenberg, 1988; Joseph et al., 1979; Leveque and Laurent, 2008; Lippi et al., 2016; Marras et al., 2005; Odeyemi, 2016; Oliver, 2005; Shuval, 2003; Slifka et al., 2017; Soller et al., 2010b; Thorsteinsson et al., 1974; Vignier et al., 2013, 2013; von Graevenitz et al., 2000; Yamahara et al., 2012)

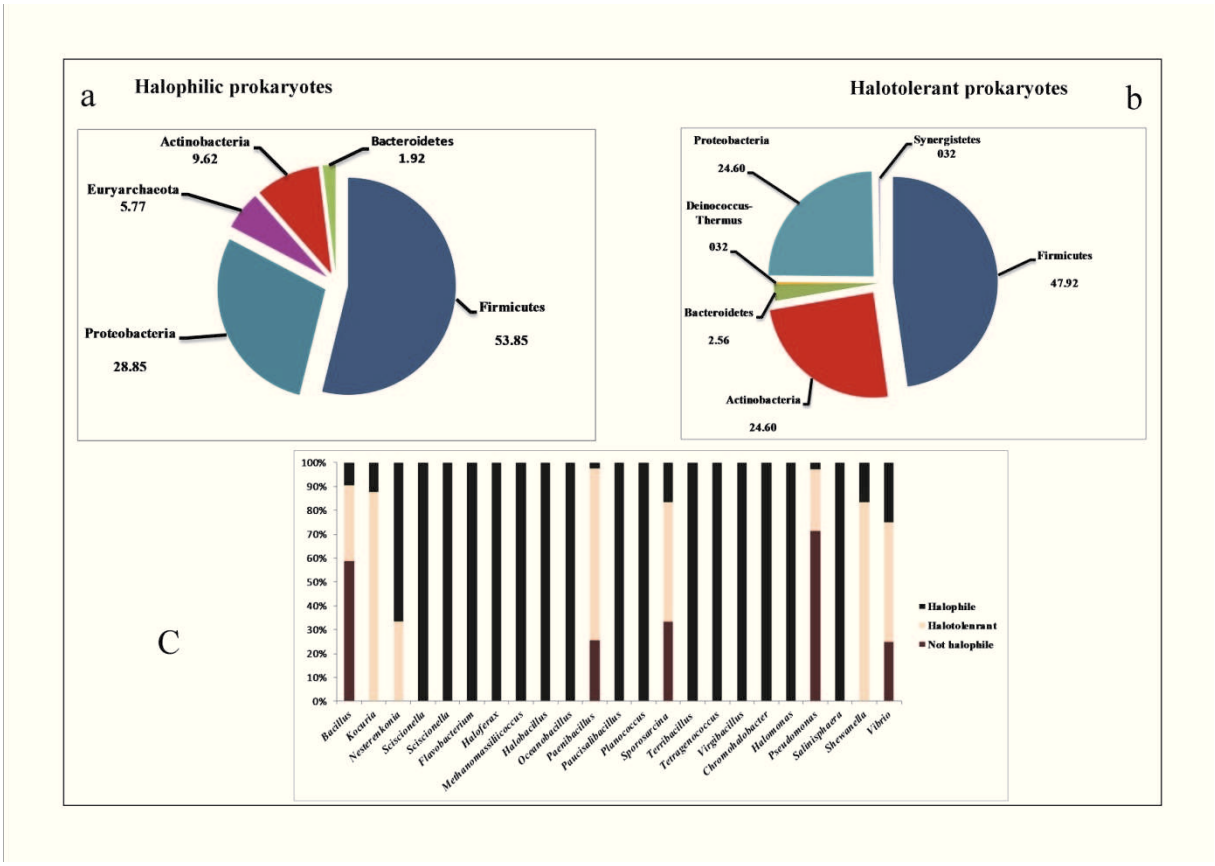
Implicated species	**Infection risk**	**Category**
Aeromonas hydrophila	endocarditis, pneumonia, conjunctivitis, urinary tract infections, wound infections	Halotolerant
Enteroccocus feacium	gastroenteritis, diarrhea, abdominal pain, nausea, vomiting	Halotolerant
Enteroccocus feacalis	gastroenteritis, diarrhea, abdominal pain, nausea, vomiting	Halotolerant
Mycobacterium avium	granulomatous pneumonia of swimming	Halotolerant
Mycobacterium marinum	bleeding bullous skin lesions, granuloma of swimming pools and aquariums	Halotolerant
Pseudomonas aeruginosa	dermatitis of spas, burning feet syndrome or Pseudomonas hot-foot	Halotolerant
Shewanella putrefaciens	bleeding bullous skin lesions	Halotolerant
Staphylococcus aureus	opportunistic pathogen, bleeding bullous skin lesions	Halotolerant
Staphylococcus spp	bleeding bullous skin lesions	Halotolerant
Staphylococcus spp	bleeding bullous skin lesions	Halotolerant
Vibrio alginolyticus	wound infection and ear infections	Halotolerant
Vibrio cholerae	cholera epidemic	Halotolerant
Vibrio parahaemolyticus	gastroenteritis, diarrhea, abdominal pain, nausea, vomiting, headaches and a mild fever	Halotolerant
Vibrio vulnificus	gastroenteritis, septicemia, cutaneous, bleeding bullous skin lesions and mucosal infection	Halotolerant

Figure 1: Extraction of publication associated with human halophilic prokaryotes in literature



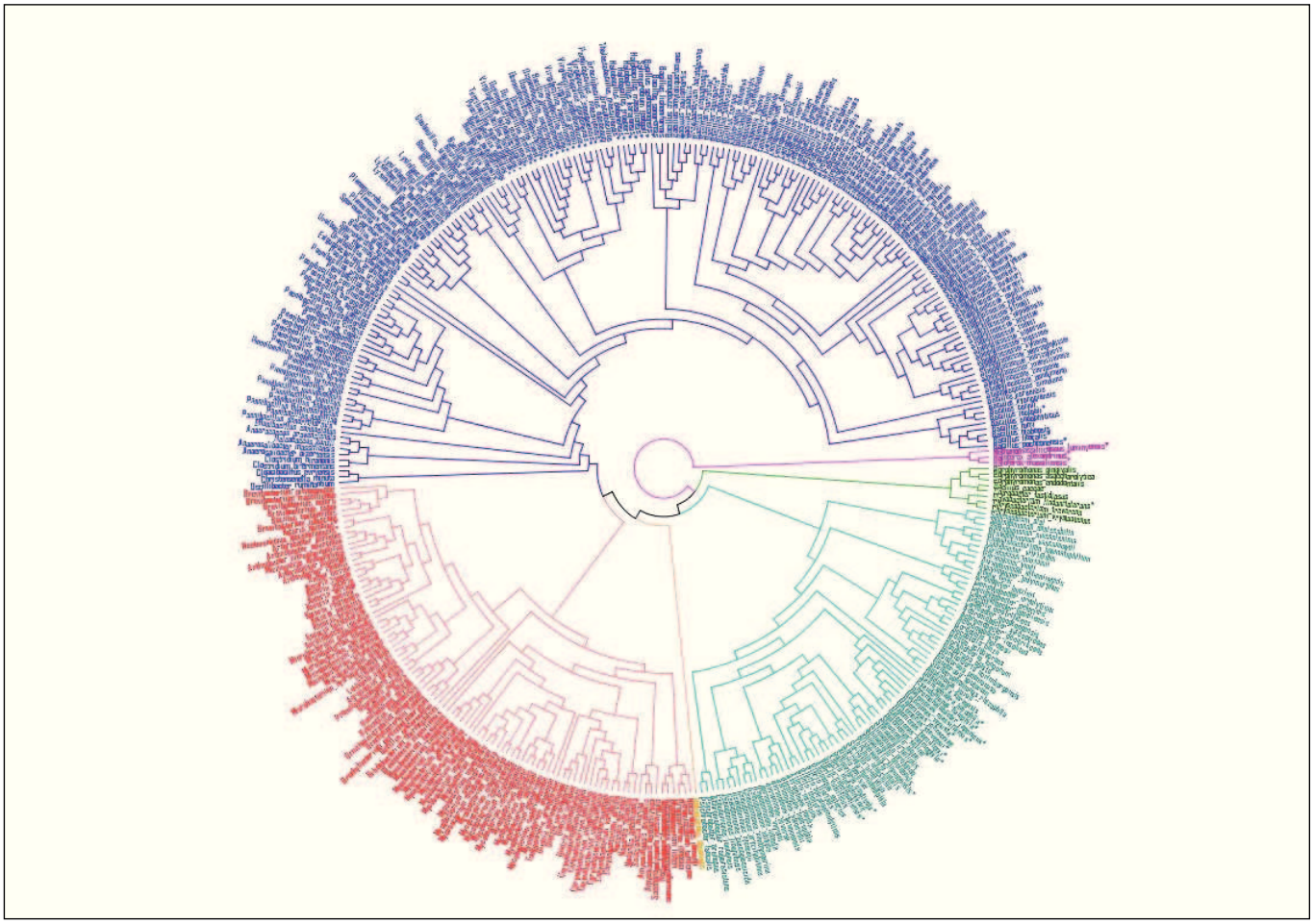
a, query parterns elaboration**, b** halophilic prokaryotic determination. The literature search included articles indexed in PubMed / MEDLINE that mentioned the scientific name of the bacteria in the title or abstract. The results are given in the form of a PMID number.

Figure 3: Distribution within the phylogenetic tree of halophilic and halotolerant bacteria isolated in human



a et **b;** Bacterial species from Firmicutes are highlighted in Blue, Proteobacteria (light blue), Actinobacteria (red), Bacteroidetes (green), Deinococcus-thermus (orange), Euryarchaeaota (purple). **c; r**epresents genus including at least one halophilic prokaryotes present in humans.

Figure 3: Distribution within the phylogenetic tree of halophilic and halotolerant bacteria isolated in human.



Bacterial species from Firmicutes are highlighted in Blue, Proteobacteria (light blue), Actinobacteria (red), Bacteroidetes (green), Deinococcus-thermus (orange), Euryarchaeaota (purple),halophilic prokaryotes (*). Phylogenetic inferences obtained using the Kimura two-parameter model. The scale bar represents 0.005% nucleotide sequence divergence.

CHAPITRE II

Etude de la diversité des prokaryotes halophiles du tube digestif par culturomics et effet du sel sur la variation du microbiome halophile.

Anvant-propos

Récemment, des études basées sur des techniques de biologie moléculaire ont détecté des micro-organismes halophiles dans le tube digestif humain. En effet, des séquences d'ADN d'haloarchaea, appartenant principalement à la famille des *Halobacteriaceae*, ont été identifiées dans 20 % des biopsies de muqueuse de côlon humain. Cependant, le répertoire halophile cultivé et non cultivé de l'intestin humain reste largement inexploré. Dans cette partie de nos travaux de thèse, notre objectif était d'explorer la diversité des halophiles du tube digestif humain à partir d'échantillons de selles en se basant sur le concept de culture « Microbial Culturomics ».

« Microbial Culturomics » est un concept innovant, développé depuis 2009, et basé sur une variation des caractères physico-chimiques des conditions de culture afin d'isoler des micro-organismes jusqu'à lors incultivés. Pour cultiver des halophiles à partir d'échantillons de selles humaines, nous avons élaboré un milieu de culture spécifique dans lequel a été ajouté d'importantes quantités de sel (NaCl). La salinité de nos milieux était variable selon la catégorie d'halophiles recherchée. Ainsi en mettant en culture 576 échantillons, 64 procaryotes ont été isolés parmi lesquels se distinguent les deux premiers Archaea halophiles isolées dans le tube digestif humain et un nouveau genre extrême halophile. Nous avons également voulu testé les liens entre la diversité halophile et la salinité fécale. Pour cela, 1.334 échantillons récoltés auprès de 1.334 participants ont été analysés. Les participants ont été recrutés sur différents continents dont l'Afrique (Bamako, Mali ; dans une région sud rurale du Sénégal, et dans le désert algérien chez les touareg), Europe (Marseille, France), Amérique (Guyane française, et la forêt Amazonienne) et en Océanie (Iles du Vent, Polynésie française). La salinité des prélèvements était comprise entre 0 et 6%. Elle était significativement plus élevée dans les régions rurales du Sénégal et de l'Amazonie comparée à l'Arabie saoudite, et la France. La diversité halophile était également significativement plus élevée au Sénégal (91%) comparée à

la Polynésie française (20%), l'Arabie saoudite (17%) et la France (0,4%). En utilisant l'analyse ROC (« Receiver operating characteristic ») sur les 576 échantillons analysés à la fois pour la salinité et en culture, nous avons trouvé le seuil 1,5% qui représente la salinité minimale requise pour la détection d'au moins une espèce procaryote halophile à partir des selles. Une corrélation très significative a été trouvée entre la salinité et la diversité halophilique. Lorsque nous avons observé le lien entre salinité fécale et obésité, nous constatons que la salinité fécale est significativement augmentée chez les obèses dans les 3 pays qui sont la France, l'Arabie Saoudite et l'Amazonie. Dans un modèle de régression logistique incluant l'obésité comme variable dépendante et âge, sexe, pays et salinité fécale comme variables indépendantes, seule la salinité fécale était un prédicateur indépendant d'obésité (OR 13,26 ; 95 % CI 2,46 – 71.36, p = 0,003). Enfin, nous avons voulu comparer les résultats obtenus par culturomics et par métagénomique. La métagénomique n'est pas le meilleur outil pour détecter les halophiles, néanmoins elle permet de relier une salinité fécale élevée à une diminution de la diversité microbienne globale du tube digestif et une déplétion de trois bactéries protectrices de l'obésité qui sont *Akkermansia muciniphila*, *Bifidobacterium longum,* et *Bifidobacterium adolescentis*.

Article 2: Salt of Stools and the Neglected Human Gut Halophilic Microbiome

El Hadji SECK[1,*], Saber KHELAIFIA[1,*], Bruno SENGHOR[1,*], Dipankar BACHAR[1], Frédéric CADORET[1], Catherine ROBERT[1], Esam Ibraheem AZHAR[2], Muhammad YASIR[2], Fehmida BIBI[2], Asif Ahmad Jiman-Fatani[3], Dabo Salimata KONATE[4], Didier MUSSO[5], Ogobara DOUMBO[4], Cheikh SOKHNA[6], Anthony LEVASSEUR[1], Jean-Christophe LAGIER[1], Matthieu MILLION[1], Didier RAOULT[1,2,6]

Soumis dans Nature

Salt in Stools, Obesity and Gut Microbiota Alteration

El Hadji SECK[1,*], Saber KHELAIFIA[1,*], Bruno SENGHOR[1,*], Dipankar BACHAR[1], Frédéric CADORET[1], Catherine ROBERT[1], Esam Ibraheem AZHAR[2], Muhammad YASIR[2], Fehmida BIBI[2], Asif Ahmad Jiman-Fatani[3], Dabo Salimata KONATE[4], Didier MUSSO[5], Ogobara DOUMBO[4], Cheikh SOKHNA[6], Anthony LEVASSEUR[1], Jean-Christophe LAGIER[1], Matthieu MILLION[1], Didier RAOULT[1,2,6]

[1]URMITE, Aix Marseille Université, UM63, CNRS 7278, IRD 198, INSERM 1095, IHU - Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13005 Marseille

[2]Special Infectious Agents Unit, King Fahd Medical Research Center and Medical Laboratory Technology Department, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

[3]Department of Medical Microbiology and Parasitology, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

[4]Malaria Research and Training Center, Department of Epidemiology of Parasitic Diseases, Faculty of Medicine, and Dentistry, UMI-CNRS 3189, GIRAFE-IRD-FMERIEUX, University of Science, Technique and Technology of Bamako, BP 1805, Bamako, Mali; Inserm U1065, Centre Méditerranéen de Médecine Moléculaire, C3M, Université de Nice-Sophia Antipolis, 151, route St Antoine de Ginestière, BP 2 3194, 06204 Nice Cedex, France

[5]Institut Louis Malardé, Papeete, Tahiti, French Polynesia

[6]Aix-Marseille Université, URMITE, IHU Méditerranée-Infection, UM63, CNRS 7278, IRD 198, Inserm U1095, Campus International UCAD-IRD, BP 1386, CP 18524, Dakar, Senegal

* These authors contributed equally to this work

Corresponding author: Prof. Didier RAOULT

Tel: 00 33 4 13 73 24 01

Fax: 00 33 4 12 73 23 07

E-mail: didier.raoult@gmail.com

Keywords: salinity, salt, gastrointestinal microbiome, halophilic, obesity

Abstract

High salt intake has been linked with several diseases and an increased risk of death[1-6] but the ability of salt to alter the gut microbiota[7], recently identified as an instrumental factor for health and disease, remains unexplored[8]. Analyzing 1,326 human fecal samples, we show that fecal salinity varies according to geographical origin, age, and obesity. By microbial culturomics[9], we deciphered the unexpected diversity of the human living halophilic microbiota, including the first 2 human-associated halophilic archaeal members[10]; 64 distinct species were isolated, including 21 new species and 41 known in the environment but not in humans. No halophiles grow in less than 1.5% salinity. Above this threshold, the richness of the halophilic microbiota correlated with fecal salinity ($r = 0.58$, $p < .0001$). Metagenomics missed the human living halophilic microbiota, but linked high fecal salinity to decreased diversity (linear regression, $p < .035$) and a depletion in anti-obesity *Akkermansia muciniphila* and *Bifidobacterium*[11-13]. Strikingly, we found that Bedouins, a traditional non-obese population of the Arabian Desert with a salty diet[14,15], have very low fecal salinity (<0.5%), and several halophiles by metagenomics (11-28 species) but none by culturomics, suggesting sodium redirection for cooling mechanisms[16,17]. This is indicative of the major bias of metagenomics, with a confusion between dead and live prokaryotes, explaining why only culturomics deciphered the link between fecal halophilic diversity and salinity, and obesity. Further studies should specify the factors driving gut salinity, and clarify if the gut microbiota alterations associated with high salt levels could be causally related to human disease such as stomach cancer and obesity.

Salt is a mineral essential for human life and one of the oldest and most ubiquitous of food seasonings. Salting is an important and very ancient method of food preservation[1]. Salt has been used for millennia to inhibit microbial proliferation through the decrease of water availability (a_w), which is among the most important hurdles used in food preservation[7]. In developed countries, salt consumption decreased with the emergence of new conservation methods and refrigerators, making preservation by salt no longer necessary[4]. The decreasing average salt intake in the last decades[18] is considered progress in public health, since high salt intake is a risk factor for several diseases, particularly obesity, independently of energy intake[3]. The gut has consistently been recently identified as a critical organ, in which salt and glucose cotransport may determine a risk for obesity through sodium-dependent glucose intestinal uptake[19]. However, so far, the salt and the microbiota in the gut have been neglected in this context.

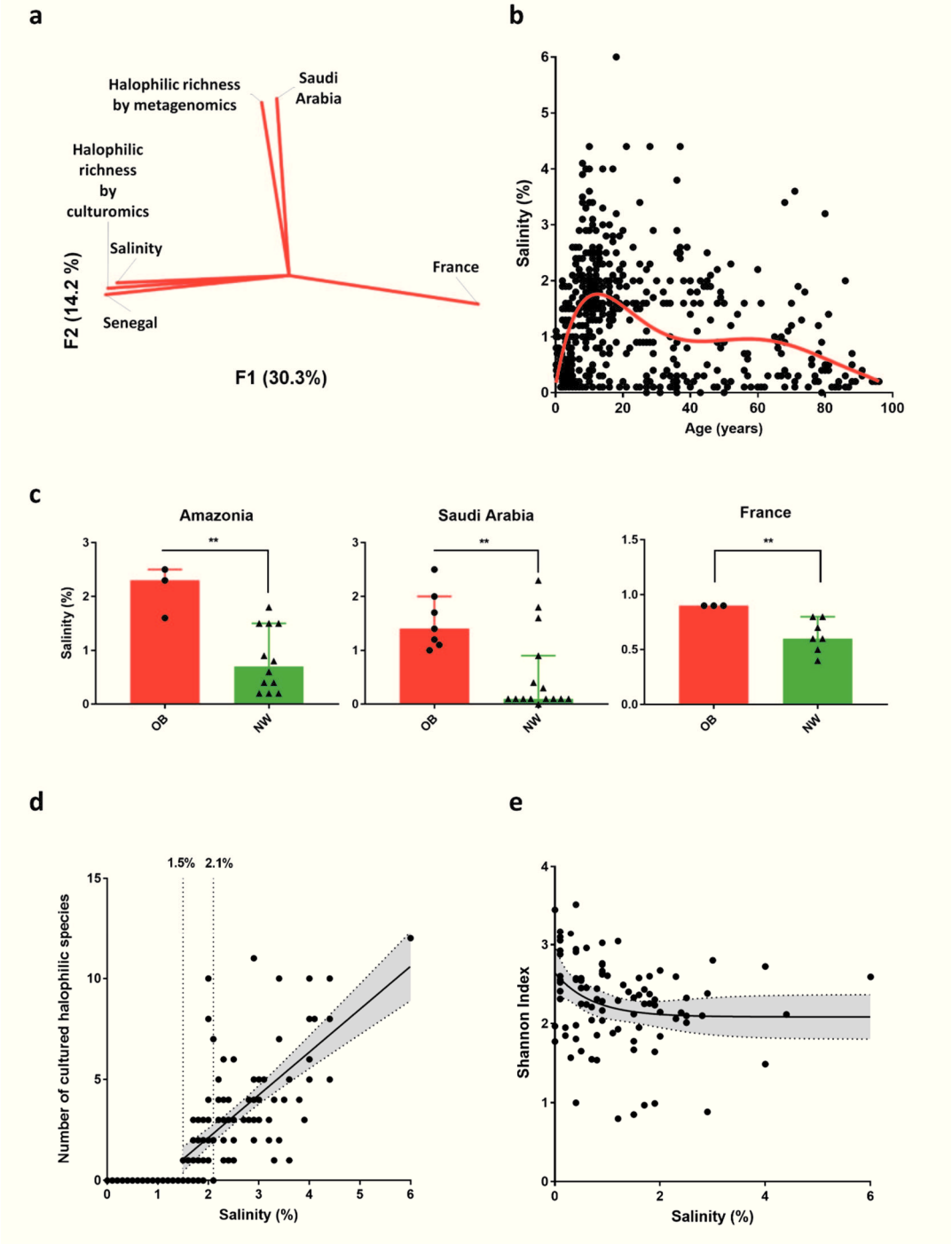
As a first step in exploring the relationship between gut salinity, geography, age, gender and obesity, we analyzed 1,326 human fecal samples from metropolitan France (n=906; 70% of samples), Senegal (n=303; 22%), Mali (n=62; 5%), Saudi Arabia (n=25; 2%), Amazonian French Guiana (n=18; 0.5%), French Polynesia (n=7; 0.4%) and the Tuareg from Algeria (n=5; 0.4%; Supplementary Table 1 & 2). Among the 25 samples from Saudi Arabia, 14 samples were collected from urban Saudis, while 11 were collected from Bedouins, a traditional nomadic population living in the desert. Among the 303 from Senegal, 9 (3%) were healthy children from Dakar, 113 (37%) were children and adults from Dielmo and 181 (60%) were children and adults from N'Diop. Dielmo and N'diop are two villages in rural southern Senegal[20]. Some samples (245/1326; 18.5%) showed significant salinity (≥1.5%, the threshold associated with halophilic growth, see below). The salinity can reach very high levels (6% (w/w); e.g., twice the ocean's average salinity) and high salinity samples (≥ 1.5%) were significantly more common in rural Southern Senegal (Dielmo & N'Diop, 217/294; 74%) and

Amazonia (10/18; 55%) compared with metropolitan France (9/906; 1%; bilateral chi-square, $p < .0001$ for both comparisons, Figure 1a & Supplementary Figure 1). This was consistent with a higher salt intake in rural people, as reported in the SALTURK study[6].

In a linear regression model including fecal salinity as the dependent variable, and age, gender and country as the independent variables, age ($p = .030$) and country ($p < .001$) but not gender ($p = .14$) were independent determinants of fecal salinity. However, age had a biphasic association with fecal salinity, with a sharp increase in the first years of life until 12 years of age, before a gradual decrease (Figure 1b). Selecting individuals with age ≤ 12 years, age increased fecal salinity independently of gender and country (linear coef. 0.08, $p < .001$), while in individuals aged ≥ 12 years, only country remained an independent predictor, suggesting that fecal salinity is an indicator of gut environment maturation in children until 12 years before relative stability. Overall, a traditional rural lifestyle and young age, but not gender, (Supplementary Figure 2) were associated with stool salinity.

Stratifying by country (Supplementary Table 3), fecal salinity was significantly and consistently higher among obese people (Figure 1c) and was a predictor of obesity independent of age, sex, and country (logistic regression, OR 13.26, 95%CI 2.46 - 71.36, $p =. 003$ – Supplementary Table 4). Taking all individuals for whom obese status and fecal salinity were available (17 obese and 39 lean), we found a dose-dependent relationship (AUC = 0.80, 95%CI 0.68-0.91, $P <. 001$ – Supplementary Figure 3). Taking all individuals for whom BMI and fecal salinity were available (n = 55), we found a significant correlation (Spearman correlation test, $P =. 016$ – Supplementary Figure 4). Altogether, these results suggest a link between salt levels in the gut and the obese phenotype. We therefore sought to characterize the gut microbiota alterations associated with high salt levels by two complementary recent techniques: metagenomics and culturomics.

Figure 1. Fecal salinity is associated with geography, age, obesity, halophilic and global microbiota diversity



(a) Principal component analysis (only variables with sum (cos^2) > 0.5 are shown. Salinity and cultured halophilic diversity depend on geography. (b) Fecal salinity according to age. A sharp increase till 12 years of age preceded a gradual decrease. (c) Fecal salinity is increased in obese people. (d) No halophiles grow below 1.5% fecal salinity. Above this threshold, cultured halophilic diversity correlates with fecal salinity (r=0.58; p < .0001). All samples with a fecal salinity > 2.1% harbor at least one live halophile. (e) A higher fecal salinity is associated with a lower overall α-diversity (linear regression, p < .035).

High salinity is an extreme environment to which relatively few organisms are adapted. Halophiles thrive in high salt concentrations by specific mechanisms to resist desiccation (osmoprotectants or selective influx of potassium ions into the cytoplasm). Strict halophiles require elevated levels of salt to grow, and have long been considered restricted to extreme environments such as salt lakes and the Dead Sea. Only recently, DNA sequences of haloarchaea have been identified in human colonic mucosal biopsies and table salt or salted shrimp, suggesting that salty food may be a source of halophilic prokaryotic DNA sequences[21]. Using culturomics with salted media to specifically study live gut halophilic prokaryotes, we analyzed 574 samples, yielding 85,326 colonies. Only 548 colonies (0.6%) needed 16S rRNA gene sequencing for their identification; 84,778 (99.4%) were correctly identified by MALDI-TOF MS. Focusing on halophilic microbiota (halotolerant bacterial species were excluded), 94 samples yielded at least one halophilic species (17%). Sixty-four different strict halophilic prokaryotic species were identified and their genomes were sequenced (Supplementary Table 5). Among these, 21 were new bacterial species[22], including one new genus, *Halophilibacterium massiliensis* (described for the first time in the Supplementary Information). Forty-one were previously known halophilic bacteria but isolated for the first time in humans, and only 2 halophilic bacteria were previously known in the human gut. Halophilic archaeal species were isolated for the first time in humans; *Haloferax alexandrines,* previously known but not in humans, and a new halophilic archaeal species, *Haloferax massiliensis* (Supplementary Figure 5)[10]. Regarding the metabolism of the 64 halo-prokaryotes isolated here, 10 were slightly halophilic bacteria, 51 were moderately halophilic bacteria species, and 3 were extremely halophilic. The 3 extremely halophilic prokaryotes included the two archaeal *Haloferax* species mentioned above and the new bacterial genus discovered in the present study (*Halophilibacterium massiliensis*). To our knowledge, this is the first time that live extreme halophiles have been isolated from humans.

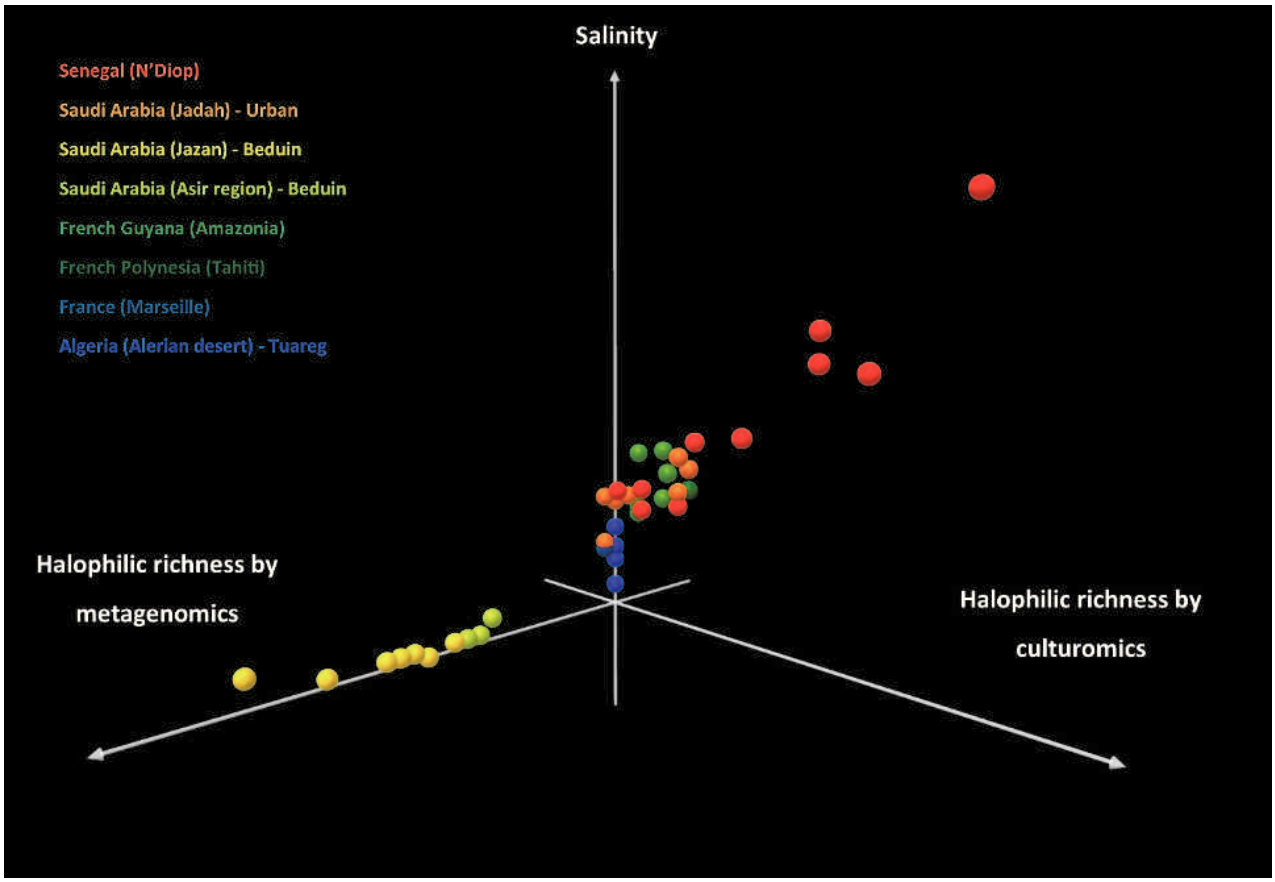
We therefore studied the determinants driving the halophilic microbiota in humans. As for salinity, age and country were independent predictors of live halophilic diversity assessed by culturomics (age, $p = .007$ - Supplementary Figure 6; gender, $p = .07$; country, $p < .001$). We found a dose-dependent relationship between fecal salinity and culture of at least one halophilic prokaryote (area under curve (AUC) = 0.996, 95%CI 0.993-0.999, $p < 1.10^{-30}$, Figure 1d & Supplementary Figure 7). The best threshold (sensitivity + specificity) was ≥ 1.5%, with a perfect negative predictive value (none of the 463 fecal samples with a salinity < 1.5% harbored any halophilic prokaryotes) and an accuracy of 97.2%. Similarly, the ≥ 2.2% threshold had a perfect positive predictive value (all the 61 fecal samples with a salinity ≥ 2.2% harbored at least one halophilic prokaryote). A very significant correlation was found between salinity and halophilic diversity among the 97 positive samples ($r = 0.58$, $p < .0001$), also confirmed with adjusted linear regression ($R^2 = 0.33$, $p < .0001$). In Saudi Arabia, the only country in whom obese and lean people were included and halophilic microbes isolated in culture, 2/3 obese versus 2/15 lean individuals had at least one halophilic species isolated in culture (Barnard's test, $p = .047$). Although preliminary, these results suggest that halophilic diversity in the gut is associated with obesity.

To decipher the gut microbiota alteration associated with high fecal salinity, we used metagenomics, which is better at assessing variations in relative abundance than culturomics. Eighty-nine fecal samples were sequenced (v3v4 large scale sequencing, Supplementary Table 6). We found that overall microbial biodiversity (Shannon index) was inversely correlated with salinity (linear regression, $p < .035$, Figure 1e), consistent with the fact that halophilic species are minority populations in the gut and that salt has the same inhibitory effect in the human gut as in food preservation. At the species level, 41 halophilic prokaryotes were identified, including 21 *Halomonas*, 6 *Kocuria*, 8 *Marinobacter* and 2 halophilic *Bacillus* species.

Analyzing the halophilic diversity by metagenomics, we serendipitously identified an outlier group which included the 11 Bedouins from Saudi Arabia from two different areas, Azir and Jazan. We recently found that this population has specific dietary habits and gut microbiota profiles[14]. Among these 11 individuals, halophilic richness ranged from 11 to 28 halophilic species, while all other individuals had a maximum of 4 halophilic species by metagenomics (Supplementary Figure 8). Moreover, this group was also different in terms of the combination of salinity and halophilic richness by metagenomics and culturomics (low fecal salinity, several halophiles by metagenomics, but none by culturomics, Figure 2). Only after exclusion of this group, the metagenomics halophilic richness correlated with salinity, and metagenomics results became consistent with culturomics results (Supplementary Figure 9). As metagenomics only reports DNA sequences, we suspected that halophiles whose sequences were found in samples with low salinity had been ingested in food[23], but did not survive in the gut because salt levels are much lower than in salty foods. A possible explanation is that Bedouins, a traditional population of the Arabian Desert with a salty diet[14, 15], have sodium redirection for cooling mechanisms[16, 17]. This is also indicative of a major bias of DNA-based metagenomics, which is a confusion between dead and live prokaryotes.

In addition, it suggests that unabsorbed salt and the distal digestive tract (ileum and colon) are critical for the link between salt, gut microbiota alteration and obesity. Moreover, the ileal and colonic microbiota itself could influence colonic NaCl absorption by short chain fatty acid (butyrate) production[24]. This also point out the predominant role of the colon for water and salt reabsorption as total colectomy is associated with dehydration and hyponatremia, as well as a greater frequency of defecation and a more liquid stool consistency[25].

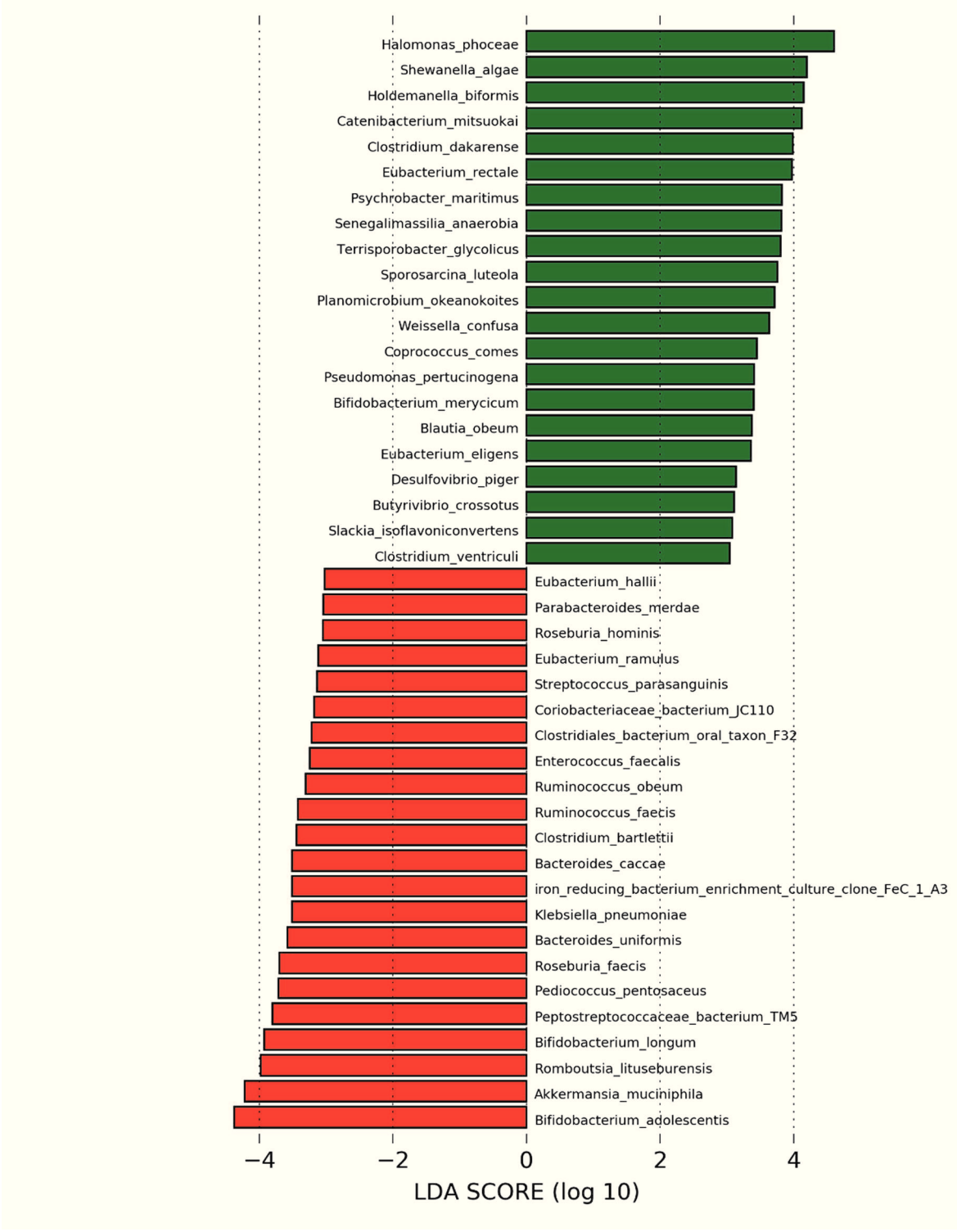
Figure 2. Metagenomics confuses dead and live prokaryotes



Halophilic prokaryotes don't thrive in absence of salt and in our study no halophiles can be cultured under 1.5% fecal salinity. Here, we found that all the fecal samples of Bedouins from Saudi Arabia had several halophiles detected by metagenomics but their fecal salinity was low (<1%) and no halophiles were cultured from their samples. This clearly evidences the major bias of DNA metagenomics: confusion between dead and live prokaryotes.

Differentiating salty (≥1.5%) and non-salty samples and using linear discriminant analysis with the LEfSe pipeline[26], we found that fecal samples with high salt levels were enriched in halophiles (*Halomonas phoceae, Shewanella algae*), but depleted in *Akkermansia muciniphila*, *Bifidobacterium adolescentis* and *Bifidobacterium longum*, which have anti-obesity properties and identified as key human symbionts for metabolic and immune homeostasis[11-13] (Figure 3 & Supplementary Table 7). This suggests a critical link between dietary salt control, preservation of a healthy gut microbiome and obesity prevention.

Figure 3. Species enriched or depleted in samples with high levels of salt



Linear discriminant analysis performed with LefSE[24]. Species enriched (green) or depleted (red) in salty stool samples (fecal salinity ≥ 1.5%) with an LDA score >3.0 are shown. Salty samples were associated with depletion in *Akkermansia muciniphila*, *Bifidobacterium adolescentis* and *Bifidobacterium longum* previously associated with leanness and immune homeostastis[11-13]. Note that the 2 most enriched species in salty samples are halophilic (*Halomonas phoceae, Shewanella algae*).

To compare culturomics and metagenomics, we selected the 51 samples analyzed by both techniques. Eighty halophilic bacteria were detected by at least one of the two techniques. Only two halophiles (2.5%) were detected by both techniques. Strikingly, 39 (48.75%) halophilic prokaryotes were obtained only by culturomics and 39 (48.75%) only by metagenomics (Supplementary Table 8, Supplementary Figure 10). As previously reported[20], this illustrates the divergence between culturomics and metagenomics, and this suggests that metagenomics is the forensic science of the gut that identifies dead or DNA relics of dietary prokaryotes[23].

Analyzing the taxonomic discrepancies between the two techniques (Supplementary Figure 10), a genus-level difference clearly appeared. All species of the following genera were found by culturomics: 9 *Bacillus,* 8 *Oceanobacillus,* 5 *Virgibacillus,* 3 *Halobacillus,* 2 *Thalassobacillus,* 1 *Gracilibacillus,* and 1 *Halomonas*, but only 2 of these species (*Bacillus* genus) were also detected by metagenomics. Conversely, the 21 *Halomonas* species, the 8 *Marinobacter* species and the 6 *Kocuria* species were mostly detected by metagenomics (among these species, only 1 *Halomonas* was identified by culturomics).

Overall, salt (sodium chloride), a micronutrient indispensable for human life, healthy growth (Type II nutrient according to Golden[27]) and prevention of dehydration, was recognized here as a major determinant of the gut microbiome, which, in turn, could have a critical instrumental role in human diseases associated with excessive salt intake, such as diabetes[2], stomach cancer[4], or obesity[3]. As the global burden of obesity rapidly and steadily increases and threatens 107 million children and 304 million adults (as of 2015[28]), the discovery of the human gut microbiome alterations associated with high salt levels may impact and provide future preventive, diagnostic and therapeutic options and allow a better understanding and control of this unprecedented and increasing worldwide pandemic. The discovery of the link between excessive salt intake and gut microbiota alteration is a new argument for young children, the

target population for obesity prevention, and a population particularly vulnerable to gut microbiota perturbation, to control their dietary salt intake to care for their lifelong microbial friends such as *A. muciniphila, B. adolescentis* and *B. longum*[11-13]. However, dietary salt is not the only factor driving fecal salinity[24] so that further studies should specify the factors driving gut salinity, and clarify if the gut microbiota alterations associated with high salt levels could be causally related to human disease such as stomach cancer and obesity.

Methods

Participants and ethics

This study included 1,326 samples prospectively collected from unrelated participants with variable geographical origin, diet, age and gender (Supplementary Table 1 & 2). One fecal sample was analyzed for each participant. Participants were recruited in different continents, countries and recruitment centers, including Africa: (Bamako and surroundings villages, Mali; Dielmo and N'Diop, rural Southern Senegal; Tuareg people of the Algerian desert, Algeria), Europe (France, Marseille), South America (French Guiana, Amazonian Forest), Saudi Arabia (Urban Jeddah and nomadic population from south western Saudi Arabia) and Oceania (French Polynesia, Iles du Vent). The stool specimens were collected after defecation in sterile plastic containers, and then aliquoted and stored at -80°C until use. Only verbal consent was necessary from participants for this study. This relied on the French bioethics decree 2007-1220 published in the official journal of the French Republic. This study was approved by the Ethics Committee of the foundation Méditerranée infection under Number 09-022.

Salinity assessment

The total mineral salt content of all 1,326 stool specimens was measured by a salinity refractometer (Thermo Scientific, Villebon sur Yvette, France). One gram of each stool specimen was diluted in 10 mL of distilled water (Thermo Scientific, Villebon sur Yvette, France) and centrifuged for 10 minutes at 5000 rpm. Then 100 µL of supernatant was deposited in the Refractometer and the result was directly displayed on the screen.

Culturing halophilic prokaryotes from stool specimens

Among the 1,326 stool samples, 576 were randomly selected and analyzed by halophilic culturomics. The culture enrichment and isolation procedures for the culture of halophilic prokaryotes were performed in a Columbia Broth medium (Sigma-Aldrich, Saint-Quentin

Fallavier, France), modified by adding (per liter): $MgCl_2$ $6H_2O$, 5 g; $MgSO_4$ $7H_2O$, 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 (Saint-Quentin-Fallavier, France).

One gram of each stool specimen was inoculated aerobically or anaerobically into 100 mL of liquid medium in flasks incubated at 37°C. The aerobic flasks were stirred at 150 rpm. Subcultures were inoculated after 3, 10, 15 and 30 incubation days for each culture condition. Then, ten-fold serial dilutions were performed from the liquid culture and then plated on agar medium containing the nutrients mentioned above. Negative controls (no inoculation of the liquid culture) were included for each culture condition. All colonies were picked and re-streaked several times to obtain pure cultures. The bacteria thus obtained were sub-cultured on a solid medium consisting of Colombia Agar medium (Sigma-Aldrich, Saint-Quentin-Fallavier, France) supplemented with 50 g/L (5%), 75 g/L (7.5%), 100 g/L (10%), 150 g/L (15%), 200 g/L (20%) or 250 g/L (25%) of NaCl to determine the halophile category to which the bacteria belonged. Slightly halophilic bacteria were defined as growing at 50 g/L of NaCl, moderately halophilic between 75-150 g/L and extremely halophilic ≥ 200 g/L.

Mass spectrometry (MALDI-TOF-MS) identification

The growing colonies were first identified using our systematic matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF-MS) screening[29] on a MicroFlex LT system spectrometer (Bruker Daltonics, Bremen, Germany). 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). 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 bacteria previously identified[30]. Protein profiles are regularly updated based on the results of clinical

diagnoses and on new species, providing new spectra.

PCR-based identification

If after three attempts the bacterial colony could not be accurately identified by MALDI-TOF or the bacterium had not previously been isolated from the human gut, 16S rRNA gene-based identification was done by amplifying and sequencing the *archaeal or bacterial* 16S rRNA gene extracted from the pure culture using the previously described DNA extraction method[31] (Dridi et al.). The archaeal 16S rRNA gene was PCR amplified using all archaeal primers SDArch0333aS15-5'-TCCAGGCCCTACGGG-3′ and SDArch0958aA19-5′YCCGGCGTTGAMTCCAATT-3′. The bacterial 16S rRNA gene was PCR amplified using primers FD1-5'- AGAGTTTGATCCTGGCTCAG-3′ and RP2-5'-ACGGCTACCTTGTTACGACTT-3′. Each 50 µl PCR consisted of 1× buffer (Qiagen), 200 µM of each dNTP, 0.2 µM of each primer, 2.5 U HotStarTaq DNA polymerase (Qiagen), and 5µl DNA. Archaeal 16S rRNA genes were amplified under the following cycle conditions: 15 min at 95°, 40 cycles at 95°C (30 s), 58°C (45 s), and 72°C (90 s), followed by a 5 min extension at 72°C. Bacterial 16S rRNA genes were amplified under the following cycle conditions: 15 min at 95°, 40 cycles at 95°C (30 s), 52°C (45 s), and 72°C (180 s), followed by a 5 min extension at 72°C. The DNA amplicon was assessed via electrophoresis on 1.5% agarose gel. Once validated, the PCR product was purified and sequenced using the BigDye Terminator Sequencing kit v.1.1 (Perkin-Elmer, Courtaboeuf, France) with the following internal primers: 536F, 536R, 800F, 800R, 1050F, and 1050R[32].

Strain deposition, 16S rRNA and genome sequencing accession numbers

All the strains isolated in this study were sequenced and deposited in the Collection de Souches de l'Unité des Rickettsies (CSUR, WDCM 875) and are easily available at (http://www.mediterranee-infection.com/article.php?laref=14&titre=collection-de-souches&PHPSESSID=cncregk417fl97gheb8k7u7t07) (Supplementary Table 5). All the new

prokaryote species isolated here were deposited into 2 international collections: the CSUR and the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) (Supplementary Table 5). The 16S rRNA accession numbers and the draft genomes are deposited with an available GenBank accession number (Supplementary Table 5). The main characteristics of each new species are reported in recently published papers[9,10,22,33-38] or at the end of this manuscript. All new prokaryote species have been or will be comprehensively described by taxonogenomics, including their metabolic properties, MALDI-TOF spectra and genome sequencing[39].

Comparison of culturomics and metagenomics

Seventy samples from this study were analyzed by v3v4 16S rRNA metagenomics sequencing as previously described[40]; 51 of these were analyzed both by culturomics and 16S rRNA metagenomics, allowing for a comparison of these techniques.

Statistical Analyses

An ROC analysis was used to determine if fecal salinity was a determinant for the presence of at least one halophilic bacterium. As no halophilic bacterium was found in fecal samples with a salinity < 1.5%, we defined 'salty' samples (salinity ≥ 1.5%) as a categorical variable. Fecal samples yielding at least one halophilic bacterium were considered 'halophilic' as another categorical variable.

The Kruskal-Wallis test with Dunn's test for multiple comparisons (corrected KW test) was used to compare quantitative variables: age, salinity (%), and halophilic diversity (number of different halophilic species by sample) among countries. A linear regression model, including age, gender and country of recruitment, was used to confirm that geographical origin was an independent predictor of salinity and halophilic diversity. The z-test for the comparison of

proportions of columns adjusted for multiple comparisons by the Bonferroni method was used to compare proportions among different geographical locations (gender, proportion of salty samples, and proportion of halophilic samples). All tests were two-sided and performed using SPSS v20.0 (IBM, Paris, France). A p value < .05 was considered significant.

Principal component analysis was performed using XLSTAT v19.3.01 (Addinsoft, Paris, France). Only variables with sum (Cos^2) > 0.5 are shown in the figures.

References

1. Kotchen, T. A., Cowley, Jr., A. W., Frohlich, E. D., Salt in health and disease – a delicate balance. *N. Engl. J. Med*. **368**, 1229-1237 (2013).
2. Hoffmann, S. & Cubeddu, L. X. Salt and the metabolic syndrome. *Nutr. Metab. Cardiovasc. Dis*. **19**, 123-128 (2009).
3. Ma, Y., He, F. J., MacGregor, G. A. High salt intake: independent risk factor for obesity? *Hypertension*. **66**, 843-849 (2015).
4. Joossens, J. V. et al., Dietary salt, nitrate and stomach cancer mortality in 24 countries. European Cancer Prevention (ECP) and the INTERSALT Cooperative Research Group. *Int. J. Epidemiol.* **25**, 494-504 (1996).
5. Yang, Q. et al., Sodium and potassium intake and mortality among US adults: prospective data from the Third National Health and Nutrition Examination Survey. *Arch. Intern. Med*. **171**, 1183-1191 (2011).
6. Erdem, Y. et al., The relationship between hypertension and salt intake in Turkish population: SALTURK study. *Blood Press*. **19**, 313-318 (2010).
7. Brown, A.D., Microbial water stress. *Bacteriol Rev*. **40**, 803–846 (1976).
8. Wilck, N., et al. A high-salt diet alters the composition of intestinal microbiota in mice. *Hypertension*. **64**, A321 (2014).
9. Lagier, J. C. et al., Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol*. **1**, 16203 (2016).
10. Khelaifia, S. & Raoult, D. *Haloferax massiliensis* sp. nov., the first human-associated halophilic archaea. *New Microbes New Infect*. **12**, 96-98 (2016).
11. Everard A, Belzer C, Geurts L, Ouwerkerk JP, Druart C, Bindels LB, et al. Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. Proc Natl Acad Sci U S A. 2013 May 28;110(22):9066-71.
12. Tan TG, Sefik E, Geva-Zatorsky N, Kua L, Naskar D, Teng F, et al. Identifying species of symbiont bacteria from the human gut that, alone, can induce intestinal Th17 cells in mice. Proc Natl Acad Sci USA. 2016;113(50):E8141-E50.
13. Schell MA, Karmirantzou M, Snel B, Vilanova D, Berger B, Pessi G, et al. The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract. Proc Natl Acad Sci U S A. 2002 Oct 29;99(22):14422-7.
14. Angelakis, E. et al., Gut microbiome and dietary patterns in different Saudi populations and monkeys. *Sci Rep*. **6**, 32191 (2016).
15. Leshem, M., Saadi, A., Alem, N., Hendi, K. Enhanced salt appetite, diet and drinking in traditional Bedouin women in the Negev. *Appetite*. **50**, 71-82 (2008).
16. Khan, M. S., Sasidharan, T. O., Ghosh, P. K. Water regulation in Barmer goat of the Rajasthan desert. *Experientia*. **35**, 1185-1186 (1979).
17. Shields, R., Mulholland, A. T., Elmslie, R. G. Action of aldosterone upon the intestinal transport of potassium, sodium, and water. *Gut*. **7**, 686-696 (1966).
18. Joossens, J. V. and Kesteloot, H. Trends in systolic blood pressure, 24-hour sodium excretion, and stroke mortality in the elderly in Belgium. *Am. J. Med*. **90**, 5S-11S (1991).
19. Baud, G. et al., Bile Diversion in roux-en-Y gastric bypass modulates sodium-dependent glucose intestinal uptake. *Cell. Metab*. **23**, 547-553 (2016).
20. Lagier, J. C., et al. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin. Microbiol. Infect*. **18**, 1185-1193 (2012).
21. Oxley, A. P. et al., Halophilic archaea in the human intestinal mucosa. *Environ. Microbiol*. **12**, 2398-2410 (2010).

22. Senghor, B. et al. Description of *'Bacillus dakarensis'* sp. nov., *'Bacillus sinesaloumensis'* sp. nov., *'Gracilibacillus timonensis'* sp. nov., *'Halobacillus massiliensis'* sp. nov., *'Lentibacillus massiliensis'* sp. nov., *'Oceanobacillus senegalensis'* sp. nov., *'Oceanobacillus timonensis'* sp. nov., *'Virgibacillus dakarensis'* sp. nov. and *'Virgibacillus marseillensis'* sp. nov., nine halophilic new species isolated from human stool. *New Microbes New Infect*. **17**, 45-51 (2017).
23. Lang, J. M., Eisen, J. A., Zivkovic, A. M. The microbes we eat: abundance and taxonomy of microbes consumed in a day's worth of meals for three diet types. *PeerJ*. **2**, e659 (2014).
24. Sandle, G.I. Salt and water absorption in the human colon: a modern appraisal. *Gut*. **43**, 294-9 (1998).
25. Warner, B. W. Short- and long-term complications of colectomy. *J Pediatr Gastroenterol Nutr*. **48 Suppl 2,** S72-4 (2009).
26. Segata, N. et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. **12**, R60 (2011).
27. Golden, M.H., Oedematous malnutrition. *Br Med Bull*. **54**, 433-44 (1998).
28. The GBD 2015 Obesity Collaborators. Health effects of overweight and obesity in 195 countries over 25 years. *N. Engl. J. Med*. **377**, 13-27 (2017).
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**, 543-51 (2009).
30. 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**, 2182-94 (2013).
31. Dridi, B. et al., High prevalence of *Methanobrevibacter smithii* and *Methanosphaera stadtmanae* detected in the human gut using an improved DNA detection protocol. *PLoS. One*. **4**, e7063 (2009).
32. Drancourt, M. et al., 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates, *J. Clin. Microbiol.* **38**, 3623-30 (2000).
33. Seck, E.H. et al., '*Halomonas massiliensis'* sp. nov., a new halotolerant bacterium isolated from the human gut. *New Microbes New Infect*. **14**, 19-20 (2016).
34. Roux, V. et al., Non-contiguous finished genome sequence and description of *Oceanobacillus massiliensis* sp. nov., *Stand. Genomic. Sci*. **9**, 370-84 (2013).
35. Khelaifia, S. et al., Microbial culturomics to map halophilic bacterium in human gut: Genome sequence and description of *Oceanobacillus jeddahense* sp. nov., OMICS. **20**, 248-58 (2016).
36. Seck, E.H. et al., Noncontiguous finished genome sequence and description of *Planococcus massiliensis* sp. nov., a moderately halophilic bacterium isolated from the human gut. *New Microbes. New Infect.* **10**, 36-46 (2016).
37. Khelaifia, S. et al., Noncontiguous finished genome sequence and description of *Virgibacillus massiliensis* sp. nov., a moderately halophilic bacterium isolated from human gut, *New Microbes New Infect*. **8**, 78-88 (2015).
38. Seck, E.H. et al., *Virgibacillus senegalensis* sp. nov., a new moderately halophilic bacterium isolated from human gut, *New Microbes New Infect*. **8**, 116-26 (2015).
39. Ramasamy, D. et al., A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. *Int J Syst Evol Microbiol*. **64**, 384-91 (2014).
40. Million, et al., Increased gut redox and depletion of anaerobic and methanogenic prokaryotes in severe acute malnutrition, *Sci. Rep*. **6**, 26051 (2016).

Supplementary Information

Description of new species isolated in this study and not previously reported (*Halophilibacterium massiliensis* gen. nov., sp. nov., *Sediminibacillus massiliensis* sp. nov., *Bacillus dionei* sp. nov., *Bacillus halophilus* sp. nov.)

Acknowledgements

We thank all the culturomics team for excellent growing work: You are the salt of the institute. But if the salt loses its saltiness, how can it be made salty again?

Author contributions

E.H.S., S.K., B.S., F.B. analyzed salinity, isolated the strains and described the new halophilic species, D.B., C.R., A.L. performed the metagenomics sequencing and bio-informatic analyses, E.I.A., S.K., D.M., O.D., C.S. collected samples and information from included individuals, M.M., E.I.A and D.R. wrote the M.S., M.M. and E.H.S. analyzed the data, J.C.L. supervised the study, D.R. conceived and supervised the study.

Author information

All genomes of halophilic bacteria isolated in this study have been deposited on GenBank (see Supplementary Table 3). Supplementary information is available online. Reprints and permissions information are available online atwww.nature.com/reprints. Correspondence and requests for materials should be addressed to D.R.

Competing interest declaration

No competing interest to declare.

Article 3: 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*

Publié dans 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

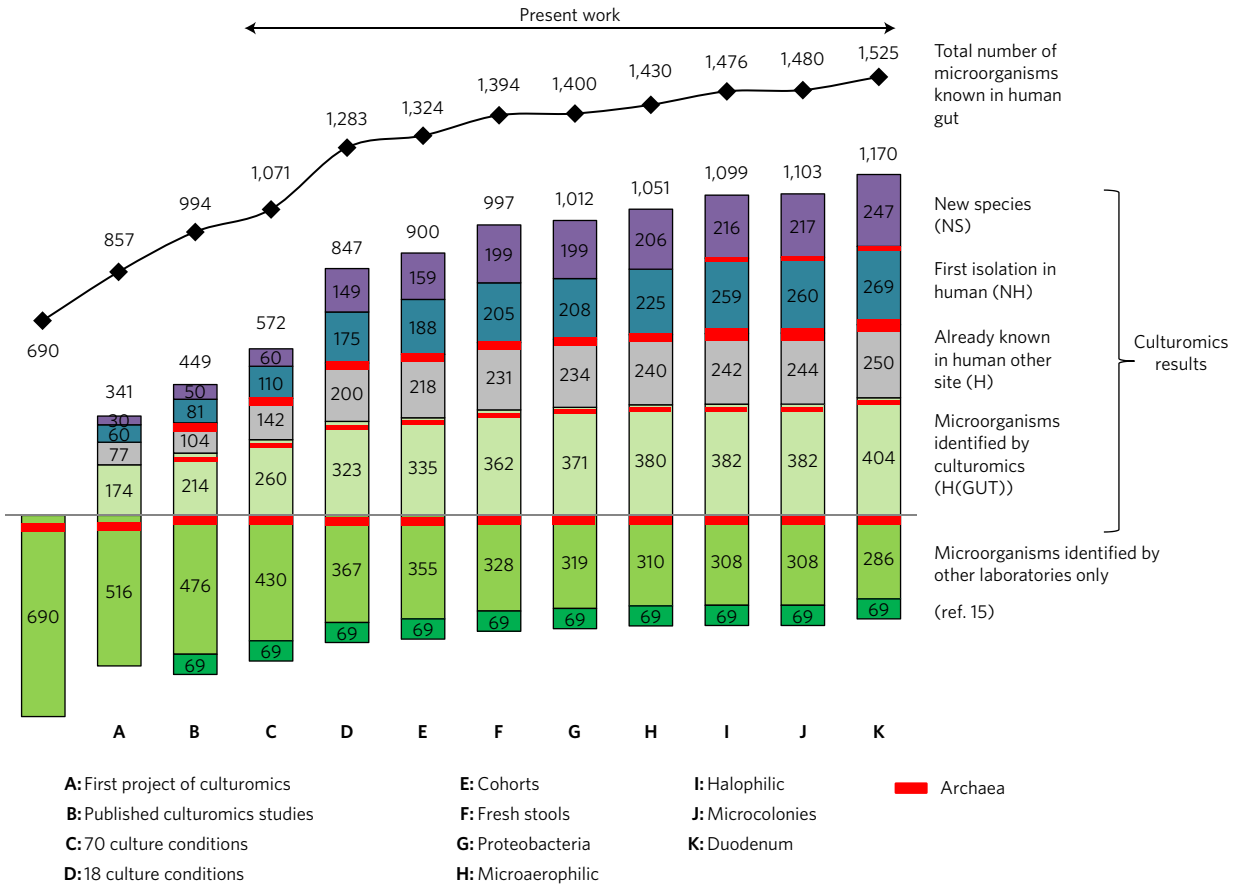


Figure 1 | Number of different bacteria and archaea isolated during the culturomics studies. Columns A and B represent the results from previously published studies, and columns C to K the different projects described herein. The bacterial species are represented in five categories: NS, new species; 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

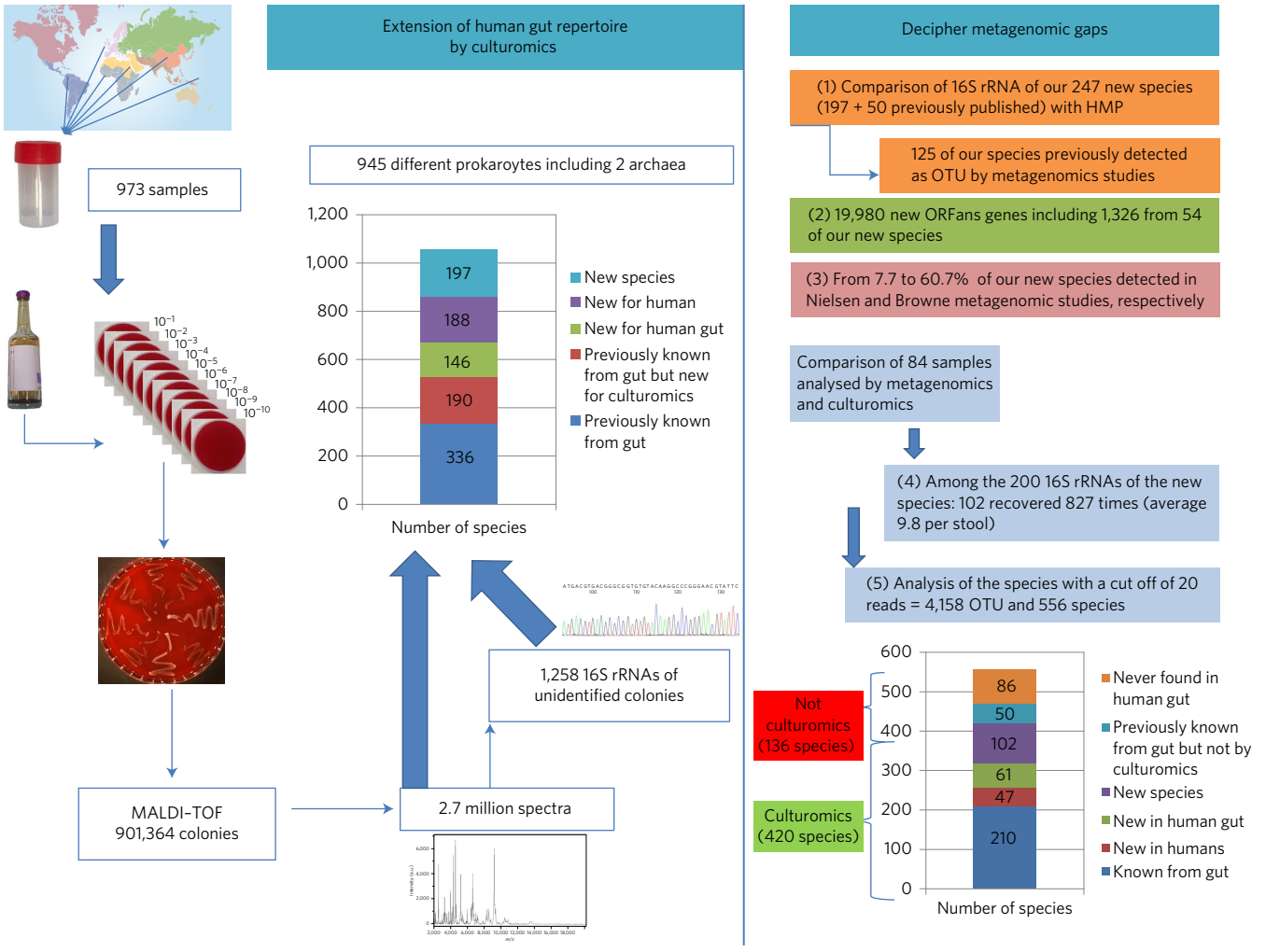


Figure 2 | Summary of the culturomics work that has extended the gut repertoire and filled some of the gaps in metagenomics.

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.

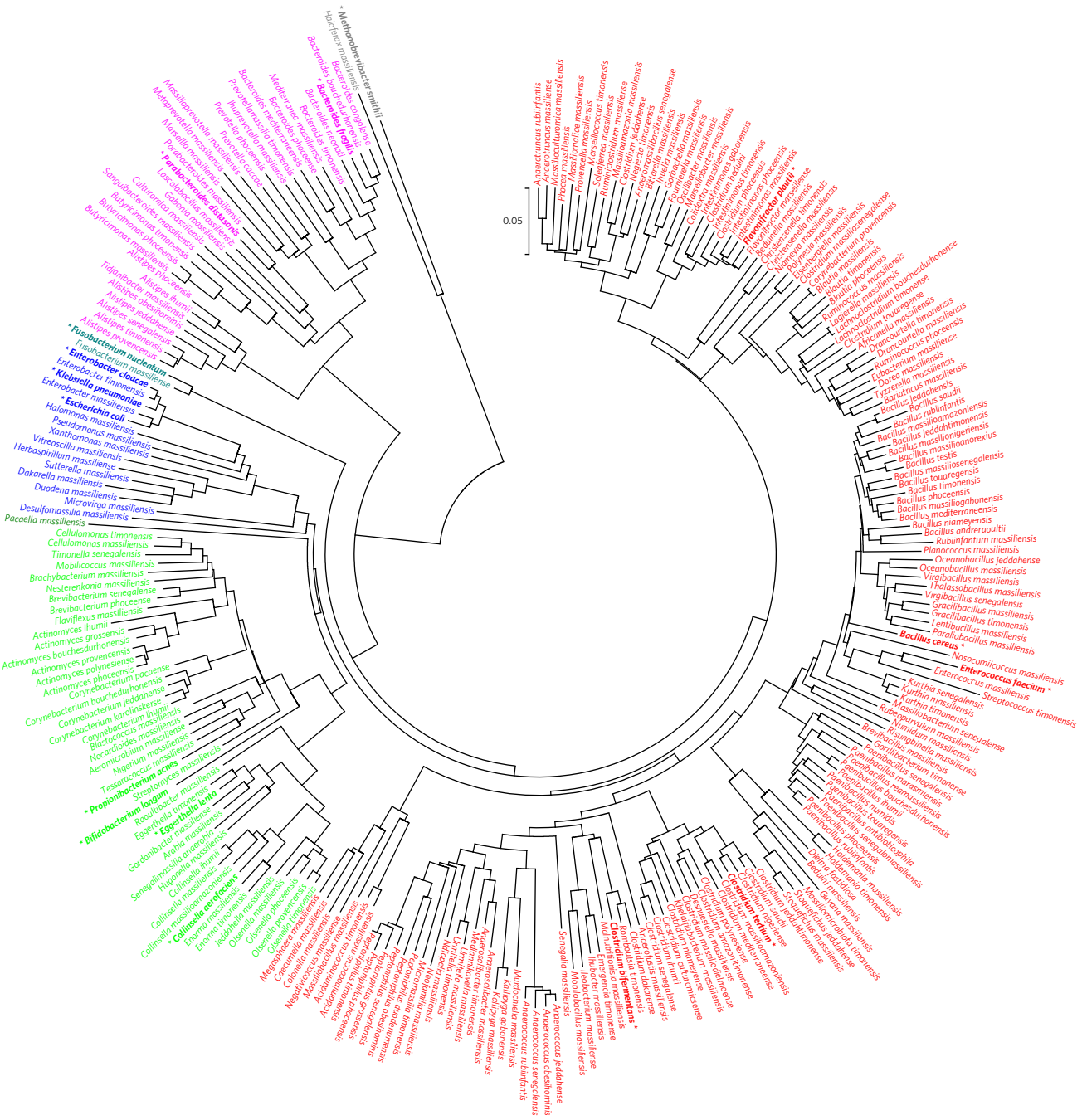


Figure 3 | Phylogenetic tree of the 247 new prokaryote species isolated by culturomics. Bacterial species from Firmicutes are highlighted in red, Actinobacteria (light green), Proteobacteria (blue), Bacteroidetes (purple), Synergistetes (green), Fusobacteria (dark green) and Archaea (grey), respectively. 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).

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 13,000g; 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.

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

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.

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.

CHAPITRE III

Evaluation de la survie des pathogènes de l'homme dans l'eau de mer

Avant-propos

L'eau de mer a une composition extrêmement homogène qui comprend principalement des bactéries halophiles. Elles sont occasionnellement pathogènes pour l'homme. Dans la littérature, il y'a peu évidence que les infections pouvent être contractées directement à partir de l'eau de mer. Néanmoins, cela contraste avec le nombre de mesures de la qualité des eaux marines mises en place et les coûts qu'elles engendrent pour prévenir le risque d'infection humaine. Par exemple, en 2014, en France et dans le golfe de Gascogne, le coût total des mesures de contrôle de la qualité des eaux marines s'élevait à environ 3,3 milliards d'euros. Aux Etats-Unis près 9 000 plages ont été fermées entre 1988-1994 pour des raisons de contamination bactérienne.

Nous avons étudié au laboratoire la survie des pathogènes les plus impliqués en microbiologie clinique dans le milieu marin, en reproduisant les conditions permettant la culture des bactéries à partir d'échantillons d'eau de mer. Notre hypothèse est la suivante : si les bactéries testées se développent dans des géloses à base d'eau de mer, en tenant compte de l'inoculum, nous pourrons considérer qu'il existe un risque de prolifération des bactéries pathogènes dans l'eau de mer. Par opposition, si nous n'observons pas de multiplication, en tenant compte de la dilution qui est considérable liée au volume de la mer, il est probable que les bactéries restent vivantes mais ne se multiplient pas dans l'eau de mer et ainsi ne représentent pas une menace pour l'homme. Parmi les 25 souches testées, 9 (36%) se multipliaient à 37 °C dans le milieu à base d'eau de mer et seulement 7 (28%) y compris *Acinetobacter baumannii*, *Escherichia coli, Klebsiella oxytoca*, *Klebsiella pneumoniae*, *Shewanella algues*, *Vibrio cholerae* et *Vibrio vulnificus* se multiplient à 25° C. Néanmoins, à 25 °C avec un inoculum de 10^4 cellules/ml, seules *Shewanella alguae*, *Vibrio cholerae* et *Vibrio vulnificus* se multipliaient. Au final, nous avons observé que la plupart des bactéries pathogènes pour l'homme ne se développent pas dans l'eau de mer dans des conditions naturelles, ce qui est expliqué par de nombreux facteurs y

compris la salinité et la température. Les températures les plus élevées sont enregistrées en été et dépassent rarement 30°C. Les maladies associées à la mer sont dominées par des infections à *Vibrio* et *Shewenalla.* Ces genres incluent des représentants halophiles et halotolérantes qui constituent la population naturelle de l'eau de mer. Contrairement à certaines idées fausses, répandues, il est clair que les infections causées par l'ingestion directe d'eau de mer semblent relativement exagérées. Les dépenses considérables consacrées à la prévention de ces infections ne sont pas fondées sur des données probantes.

Article 4: Very low risk of bacterial infection in seawater

El Hadji SECK[1], Saber KHELAIFIA[1], Jean-Christophe LAGIER[1], Didier RAOULT[1, 2*]

Soumis dans Lencet Infectious Diseases

Very low risk of bacterial infection in seawater

El Hadji SECK[1], Saber KHELAIFIA[1], Jean-Christophe LAGIER[1], Didier RAOULT[1, 2*]

[1]Aix Marseille Univ, URMITE, UMR63 CNRS 7278, IRD 198, Inserm 1095, IHU-Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13005 Marseille

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

*Corresponding author: Prof. Didier RAOULT

URMITE, IHU-Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13005 Marseille

Email: didier.raoult@gmail.com

Phone: 04 13 73 24 01

Fax: 04 13 73 24 02

Abstract

Background: The microbial composition of the sea is extremely homogeneous and mainly composed of halophilic bacteria which are rarely pathogenic. Nevertheless, the potential of certain bacteria to grow in seawater presents an interesting problem in terms of infectious diseases. However, the paucity of literature on the risk of human infections arising from salt water contrasts with the number of marine water quality measurements in place to protect humans from the risk of infection. In this study, we analyzed whether human pathogenic bacteria were likely to grow in seawater.

Methods: We evaluated the growth of 25 pathogenic strains included in the Collection de Souches de l'Unité des Rickettsies (CSUR) on natural seawater supplemented with 1.5% of bacteriological agar at 25°C and 37°C. For each strain, different inocula (10^6 to 10 cells/mL) were performed.

Findings: Of the 25 strains tested, nine (36%) grew at 37°C and seven grew (28%) at 25°C including, *Acinetobacter baumannii, Escherichia coli, Klebsiella oxytoca, Klebsiella pneumoniae, Vibrio vulnificus, Shewanella algae, and Vibrio cholerae*. Depending on the inocula, at 25°C, no growth was detected when the inocula $<10^4$ cells/ml; *Shewanella algae, Vibrio cholerae,* and *Vibrio vulnificus* grew from an inoculum of 10^4 cells/ml.

Interpretation: Our results indicated that most human pathogenic bacteria don't survive in seawater due to many factors, including salinity. It's clear that infections caused by the direct ingestion of seawater are exaggerated. Further research should be undertaken to reassess the actual infectious risks associated with seawater in order to rehabilitate the numerous measurements implemented and the huge economic resources 3·3 billion euro/year in France mobilized to monitor marine waters.

Funding: Fondation Méditerranée Infection.

Introduction

The microbial composition of seawater is extremely homogeneous.[1] More than half comprises halophilic bacteria,[2,3] which are the most common bacteria on the surface of the Earth.[2] *Pelagibacter ubiquis* and *Pelagibacter bermudensis,* two slight halophilic bacteria living in seawater and belonging to the SAR 11 clade, represent as much as 30% of the euphotic microbial community.[4] Halophilic bacterial species are occasionally commensal in humans, but are rarely human pathogens. In the literature, there is little evidence to demonstrate that infections can be contracted directly from seawater. Rare cases have been noted in near-drowning situations, where significant quantities of water can be inhaled and can cause pneumonia.[5] On the other hand, infections have been reported in relation to the consumption of seafood.[6]

Seafood filters significant volumes of water to meet their nutritional and respiratory requirements, thus concentrating pathogenic microorganisms present in the environment.[7] In addition, other staple foods raised or harvested close to shorelines and estuaries where fecal bacteria persist can be reservoirs for pathogens and vectors of gastroenteritis.[8] Surprisingly, one of the major characteristics of seawater is that it contains salt[9] which has long been used as a preservative. Indeed, brackish water, like brine, is a method of preserving food, meaning that it can avoid toxigenic bacteria and pathogenic bacteria proliferation in general.[10] The lack of literature on the risk of human infections caused by salt water contrasts with the extensive water quality measurements in place to prevent the rarely-seen risk of human infection.[11] In the United States nearly 12,000 beaches were closed between 1988 and 1994 and more than 75% of them were for reasons of microbial contamination.[12] In Europe, considerable economic resources are mobilized for beach monitoring. For example, in 2014, in France and the Gulf of Biscay maritime sub-region, the total cost of measures for monitoring the quality of marine waters was approximately 3.3 billion euros.[13] . In our laboratory, under conditions enabling us to cultivate

seawater bacteria, we investigated whether the main human pathogenic bacteria could grow in seawater agar.[14] If they do not multiply, depending on the dilution and volume of water, it is likely that the bacteria which are found to live but not multiply in seawater do not represent a threat to humans.

Materials and methods

Bacterial strains

A total of 25 bacterial strains including five strict aero-intolerant bacteria and 20 aero-tolerant or facultative aero-intolerant bacteria were tested **(Table 1).** These included the 17 most frequent pathogenic bacteria isolated in the clinical microbiology laboratory at the Timone hospital[15] (*Acinetobacter baumanii* (CSURP202), *Bacteroides fragilis* (CSURP2573), *Clostridium difficile* (CSURP346), *Enterobacter cloacae* (CSURP2284), *Enterococcus faecium* (CSURP745), *Escherichia coli* (CSURP397), *Finegoldia magna* (CSURP2342), *Klebsiella oxytoca* (CSURP2863), *Klebsiella pneumoniae* (CSURP2714), *Morganella morganii* (CSURP752), *Parvimonas micra* (CSURP2136), *Propionibacterium acnes* (CSURP742), *Pseudomonas aeruginosa* (CSURP1003), *Staphylococcus aureus* (CSURP2557), *Streptococcus agalactiae* (CSURP2760), *Streptococcus pneumoniae* (CSURP3392), and *Streptococcus pyogenes* (CSURP576)); the four major human pathogenic bacteria (*Mycobacterium tuberculosis* (CSUR2739), (CSURP782), *Salmonella Typhi* (CSURP403), and *Yersinia enterocolitica* (CSURP840), and four pathogenic bacteria associated with infections contracted in the marine environment: *Shewanella algae* (CSURP601), *Shewanella putrefaciens* (CSURP2224), *Vibrio cholerae* (CSURP2057) and *Vibrio vulnificus* (CSURP4422).[16, 17] All microorganisms studied were cultivable on 5% sheep blood agar medium (BioMerieux, Marcy l'Étoile, France) with the exception of *Neisseria meningitidis,* for which we used Polyvitex agar chocolate (BioMérieux) and incubated in an appropriate atmosphere at 37° for 24-48 hours.

Culture medium and bacterial growth

To test the viability of these strains in seawater, we used a solid culture medium consisting of natural seawater supplemented with 1·5% of bacteriological agar (BioMerieux). The seawater was collected from the Mediterranean coast in Marseille, France (Latitude: 43·25, Longitude: 5·37) and then transported to the laboratory under suitable conditions.

The salt concentration of the sample was determined using a digital refractometer (Fisher Scientific, Illkirch, France) and the pH was measured using a pH meter (Cyberscan 510PH, Eutech Instruments, Singapore). The collected seawater was then filtered (diameter 0·8 mm) to remove large debris and 15 g/L of bacteriological agar (BioMerieux) was added to obtain a solid culture medium to use for the various tests. The culture medium prepared was then sterilized in an autoclave at 120°C for 20 minutes. After autoclaving, a volume of 25 mL of medium was poured into 90 mm sterile petri dishes. The media were stored at 4°C until use.

Strain authentication using MALDI-TOF mass spectrometry

After 24–48 hours, a fresh culture colony was identified for each strain using MALDI-TOF mass spectrometry (Brüker Daltonics, Leipzig, Germany)[15] to verify authenticity before conducting the experiments. One colony was deposited in duplicate on a MALDI-TOF MSP96 target, and then covered with 1·5 µl of a matrix solution (saturated solution of α-cyano-4-hydroxycinnamic acid in 50% acetonitrile, 2·5% trifluoroacetic acid). The resulting score dictates whether a tested species can be identified: a score ≥ 2 with a validly published species enables identification at the species level, a score ≥ 1·7 but < 2 enables identification at the genus level, and a score < 1·7 does not enable any identification.

Viability tests

A bacterial suspension of 0.5 McFarland (10^8 cells/mL) was prepared from fresh colonies and a decimal dilution in cascade was then conducted by transferring 100 µL of this suspension to

an Ependorf tube containing 900 µl of PBS (Phosphate-buffered saline; BioMerieux).[1]. For each strain, six successive dilutions were conducted, giving an initial inoculum of: 10^8, 10^7, 10^6, 10^5, 10^4, and 10^3 cells/mL for each dilution, receptively. 10 µl of each suspension was then deposited on the test culture medium to obtain a final inoculum concentration of: 10^6, 10^5, 10^4, 10^3, 10^2, and 10 cells/mL. The last step was to evenly spread the inoculum , in a single stroke using a sterile spreader until no visible liquid remained on the surface of the medium (**Figure 1**). Control cultures inoculated on the reference culture media of each bacterium were simultaneously inoculated with the same methods (**Figure 1)**. The seeded petri dishes were placed in stacks of a maximum of six dishes and then incubated both at 37°C and at ambient temperature. All the experiments reported here were carried out in triplicate for the 25 clinical strains. Authenticity was then checked using MALDI-TOF before any conclusions were drawn.

Results

We were able to evaluate the growth of 25 pathogenic strains in solid medium at laboratory temperature and at 37°C. Bacterial growth was observed after 24 hours of incubation for aerotolerant bacteria and after 48-72 hours for aero-intolerant bacteria. The salt concentration of the seawater was 40 g / L and its pH was 7.8. Of the 25 strains tested, nine were able to grow in seawater at 37 °C (*Acinetobacter baumannii, Enterobacter cloacae, Enterococcus faecium, Escherichia coli, Klebsiella oxytoca, Klebsiella pneumoniae, Shewanella algae, Vibrio cholerae,* and *Vibrio vulnificus*). Growth was not observed for *Bacteroides fragilis, Clostridium difficile, Finegoldia magna, Morganella morganii, Mycobacterium tuberculosis, Neisseria meningitides, Parvimonas micra, Propionibacterium acnes, Pseudomonas aeruginosa, Salmonella Typhi, Shewanella putrefaciens, Staphylococcus aureus, Streptococcus agalactiae, Streptococcus pyogenes, Streptococcus pneumoniae,* and *Yersinia enterocolitica*), regardless of dilution (**Table 2, Fig. 2**). At laboratory temperature (25°C), only *Acinetobacter baumannii, Escherichia coli, Klebsiella oxytoca, Klebsiella pneumoniae, Vibrio vulnificus, Shewanella*

algae, and *Vibrio cholerae* grew. Of those tested, no aero-intolerant bacteria could grow in seawater (**Table 2, Fig. 2**).

Evaluation of growth according to concentration of the inoculum

Depending on the density of the bacteria cultured, results varied. At 37°C, of the nine strains that grew in seawater, only three were able to grow from an inoculum of 10^2 cells / ml (*Enterobacter cloacae, Enterococcus faecium,* and *Vibrio cholerae*) (**Fig. 3, Table 3**). Below this, no growth was observed. From an inoculum of 10^4 cells/ml, *Vibrio vulnificus* and *Shewanella algae* grew. *Escherichia coli* and *Klebsiella oxytoca,* grew only when the initial inoculum was 10^6 cells / ml (**Fig. 3, Table 3**). No colonies were observed from inoculum at 10 cells/ml for all strains tested.

At room temperature, no growth was detectable as soon as the initial population of the inoculum was lower than 10^4 cells / ml (Table 3). Of the 25 strains tested in seawater at laboratory temperature, *Shewanella algae, Vibrio cholerae* and *Vibrio vulnificus* represented those capable of growing from an inoculum of 10^4 cells/ml. *Acinetobacter baumanii, Escherichia coli* and *Klebsiella oxytoca* grew in seawater, although only above an inoculum of 10^6 cells / ml.

Discussion

Survival of main human pathogens in the marine environment has been the subject of relatively little study. The rare studies carried out on the survival of pathogenic bacteria in seawater generally concern fecal coliforms used today as indicators of the sanitary quality of seawater.[12] The value of detecting these coliforms as indicator organisms lies in the fact that their survival in the environment is considered to be generally equivalent to that of pathogenic bacteria.[19] However, the majority of clinical microbiologists agree that none of the bacterial indicators used meet the ideal criteria for deciding on the sanitary quality of water.[20] None of these studies take into account inoculum density, which is also an important parameter in bacterial culture.[21]

Here, we have demonstrated that most bacteria that are pathogenic to humans do not grow in seawater under natural conditions. The methods used to assess bacterial growth are the same as those used to grow and isolate marine bacteria from seawater samples.[22] Indeed, in many marine microbiology studies, natural seawater-based environments are used.[23] To ensure that our conclusions were accurate and reproducible, each operation was repeated three times.

The rapid death of bacteria entering the oceans can be explained by physical, chemical or biological factors.[9] There is strong evidence that salinity, sedimentation and sunlight, affect the survival of non-marine bacteria in the sea.[24] The major factor controlling the survival of bacteria in oceans is that seawater is a diluted culture medium in which the concentration of dissolved organic material does not exceed 10 ppm.[24] In addition, important nutrients such as iron, nitrogen and phosphorus are present in only small quantities[24] It is known that the population size of an organism in a chemostat culture is directly related to the quantity of growth-limiting substrate in the culture medium.[25] The temperature of seawater is another important factor that may influence the results of marine bacteria survival.[9] Although some bacteria can be cultured at 37°C in seawater in experimental conditions,[26] it is unlikely that marine waters temperature ever reach 37°C.[27] Considering the influence of seasonal variations, the average temperature of the oceans is about 3.5°C.[28] The highest temperatures are recorded during the summer and rarely exceed 30°C[28] Our data indicate that at 25°C, of the species tested, the species of the genus *Shewanella* and *Vibrio* (10^4 cell/ml) were the only ones capable of growing in seawater. Similar findings have been reported by Gram *et al.*.[26] Both genera contain halotolerant species which, together with halophiles, constitute the natural microbial population of seawater.[29] The genus *Shewanella* and *Vibrio*, are part of a very small number of human pathogens mainly or exclusively associated with seawater. There are more than eight species of *Shewanella* involved in human infections. The most widely cited are *Shewanella putrefaciens* and *Shewanella algae.*[17] These are responsible for many diseases, including gastrointestinal infections.[17] In

contrast, the vast majority of diseases associated with the marine environment are contracted in a context unrelated to salt water exposure but are usually caused by handling fish (*S. aureus, Mycobacterium spp*),[30] contact with beach sand (*K. pneumoniae, E. coli, P aeruginosa, S. aureus, Streptococcus* spp.)[8] and above all, the consumption of marine products such as seafood (*S. Typhi, E. coli, Enterococcus* spp.).[6]

Conclusion

The survival of pathogenic bacteria in seawater is an area of scientific controversy. Here, we have shown that majority of bacteria cited in clinical microbiology do not survive for any length of time in marine waters as the result of many factors, including salinity. Marine diseases are dominated by *Vibrio* and *Shewenalla* infections.[17] Contrary to common misconceptions, it would appear that infections caused by the direct ingestion of seawater are exaggerated. Pathogens that are inactivated when entering a saline environment should not be considered as etiological agents of marine infection. Further studies should be conducted to reassess the real infectious risks associated with seawater in order to review the policies currently in place for monitoring marine waters.

Bibliography

1. Droop MR, Wood EJF. *Advances in microbiology of the sea. 1 1*. Academic Press: London, 1968.

2 . Morris RM, Rappé MS, Connon SA, Vergin KL, others. SAR11 clade dominates ocean surface bacterioplankton communities. *Nature.* 2002; **420**: 806.

3 . Ventosa A, de la Haba RR, Sánchez-Porro C, Papke RT. Microbial diversity of hypersaline environments: a metagenomic approach. *Curr Opin Microbiol* . 2015; **25**: 80–87.

4. Brown SN, Giovannoni S, Cho J-C. Polyphasic taxonomy of marine bacteria from the SAR11 group Ia: Pelagibacter ubiquis (strain HTCC1062) & Pelagibacter bermudensis (strain HTCC7211). 2012.https://ir.library.oregonstate.edu/xmlui/handle/1957/29266.

5. Sims JK, Enomoto PI, Frankel RI, Wong LM. Marine bacteria complicating seawater near-drowning and marine wounds: a hypothesis. *Ann Emerg Med.* 1983; **12**: 212–216.

6 . Iwamoto M, Ayers T, Mahon BE, Swerdlow DL. Epidemiology of Seafood-Associated Infections in the United States. *Clin Microbiol Rev*. 2010; **23**: 399–411.

7. Vaillant V, Jourdan-Da Silva N, Quilici M-L, Couturier E, Le Guyader S, Delmas G *et al.* Surveillance des risques biologiques liés à la consommation de coquillages en France. *Bull Epidemiol Hebd.* 2012; **HS**: 34–37.

8. Solo-Gabriele HM, Harwood VJ, Kay D, Fujioka RS, Sadowsky MJ, Whitman RL *et al.* Beach sand and the potential for infectious disease transmission: observations and recommendations. *J Mar Biol Assoc U K.* 2016; **96**: 101–120.

9. Carlucci AF, Pramer D. Factors affecting the survival of bacteria in sea water. *Appl Microbiol*. 1959; **7**: 388–392.

10. Leistner L. Basic aspects of food preservation by hurdle technology. *Int J Food Microbiol* . 2000; **55**: 181–186.

11. Wade TJ, Pai N, Eisenberg JNS, Colford JM. Do U.S. Environmental Protection Agency water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. *Environ Health Perspect* . 2003; **111**: 1102–1109.

12. Henrickson SE, Wong T, Allen P, Ford T, Epstein PR. Marine swimming-related illness: implications for monitoring and environmental policy. *Environ Health Perspect* . 2001; **109**: 645–650.

13. Mongruel R, Jacob C. Coûts liés aux organismes pathogènes microbiens/SRM GDG. https://www.researchgate.net/profile/Celine_Jacob3/publication/268335846_Couts_lies_aux_organismes_pathogenes_microbiens_SRM_GDG/links/5478392a0cf205d1687cb69d.pdf.

14. McFeters GA, Bissonnette GK, Jezeski JJ, Thomson CA, Stuart DG. Comparative Survival of Indicator Bacteria and Enteric Pathogens in Well Water. *Appl Microbiol*. 1974; **27**: 823–829.

15. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P-E, 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**: 543–551.

16. Howard RJ, Bennett NT. Infections caused by halophilic marine Vibrio bacteria. *Ann Surg.* 1993; **217**: 525.

17. Vignier N, Barreau M, Olive C, Baubion E, Théodose R, Hochedez P *et al.* Human Infection with Shewanella putrefaciens and S. algae: Report of 16 Cases in Martinique and Review of the Literature. *Am J Trop Med Hyg.* 2013; **89**: 151–156.

18. Magniez M. Méthode de dénombrement des micro-organismes en milieu liquide (Méthode dite du nombre le plus probable). biotechnologie. (accessed 19 Jul2017). http://www.technobio.fr/2014/11/methode-de-denombrement-des-micro-organismes-en-milieu-liquide-methode-dite-du-nombre-le-plus-probable.html . Date: 2017

19 . Alkan U, Elliott DJ, Evison LM. Survival of enteric bacteria in relation to simulated solar radiation and other environmental factors in marine waters. *Water Res.* 1995; **29**: 2071–2080.

20. Gerba CP, Goyal SM, LaBelle RL, Cech I, Bodgan GF. Failure of indicator bacteria to reflect the occurrence of enteroviruses in marine waters. *Am J Public Health.* 1979; **69**: 1116–1119.

21. Rozen Y, Belkin S. Survival of enteric bacteria in seawater. *FEMS Microbiol Rev.* 2001; **25**: 513–529.

22. Rappé MS, Connon SA, Vergin KL, Giovannoni SJ. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature.* 2002; **418**: 630–633.

23. Schut F, de Vries EJ, Gottschal JC, Robertson BR, Harder W, Prins RA *et al.* Isolation of Typical Marine Bacteria by Dilution Culture: Growth, Maintenance, and Characteristics of Isolates under Laboratory Conditions. *Appl Environ Microbiol* . 1993; **59**: 2150–2160.

24. Carlucci AF, Pramer D. An evaluation of factors affecting the survival of Escherichia coli in sea water II. Salinity, pH, and nutrients. *Appl Microbiol.* 1960; **8**: 247–250.

25. Crabbendam PM, Neijssel OM, Tempest DW. Metabolic and energetic aspects of the growth of Clostridium butyricum on glucose in chemostat culture. *Arch Microbiol.* 1985; **142**: 375–382.

26. Gram L, Bundvad A, Melchiorsen J, Johansen C, Fonnesbech Vogel B. Occurrence of Shewanella algae in Danish Coastal Water and Effects of Water Temperature and Culture Conditions on Its Survival. *Appl Environ Microbiol* . 1999; **65**: 3896–3900.

27. Lea DW, Pak DK, Spero HJ. Climate impact of late Quaternary equatorial Pacific sea surface temperature variations. *Science* . 2000; **289**: 1719–1724.

28. Maslin MA, Shackleton NJ, Pflaumann U. Surface water temperature, salinity, and density changes in the northeast Atlantic during the last 45,000 years: Heinrich events, deep water formation, and climatic rebounds. *Paleoceanography* . 1995; **10**: 527–544.

29. Oren A. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Syst* 2008; **4**: 2.

30. Slany M, Jezek P, Bodnarova M. Fish Tank Granuloma Caused by Mycobacterium marinum in Two Aquarists: Two Case Reports. *BioMed Res Int* . 2013; **2013**: 161329.

Table 1: List of tested pathogenic bacteria

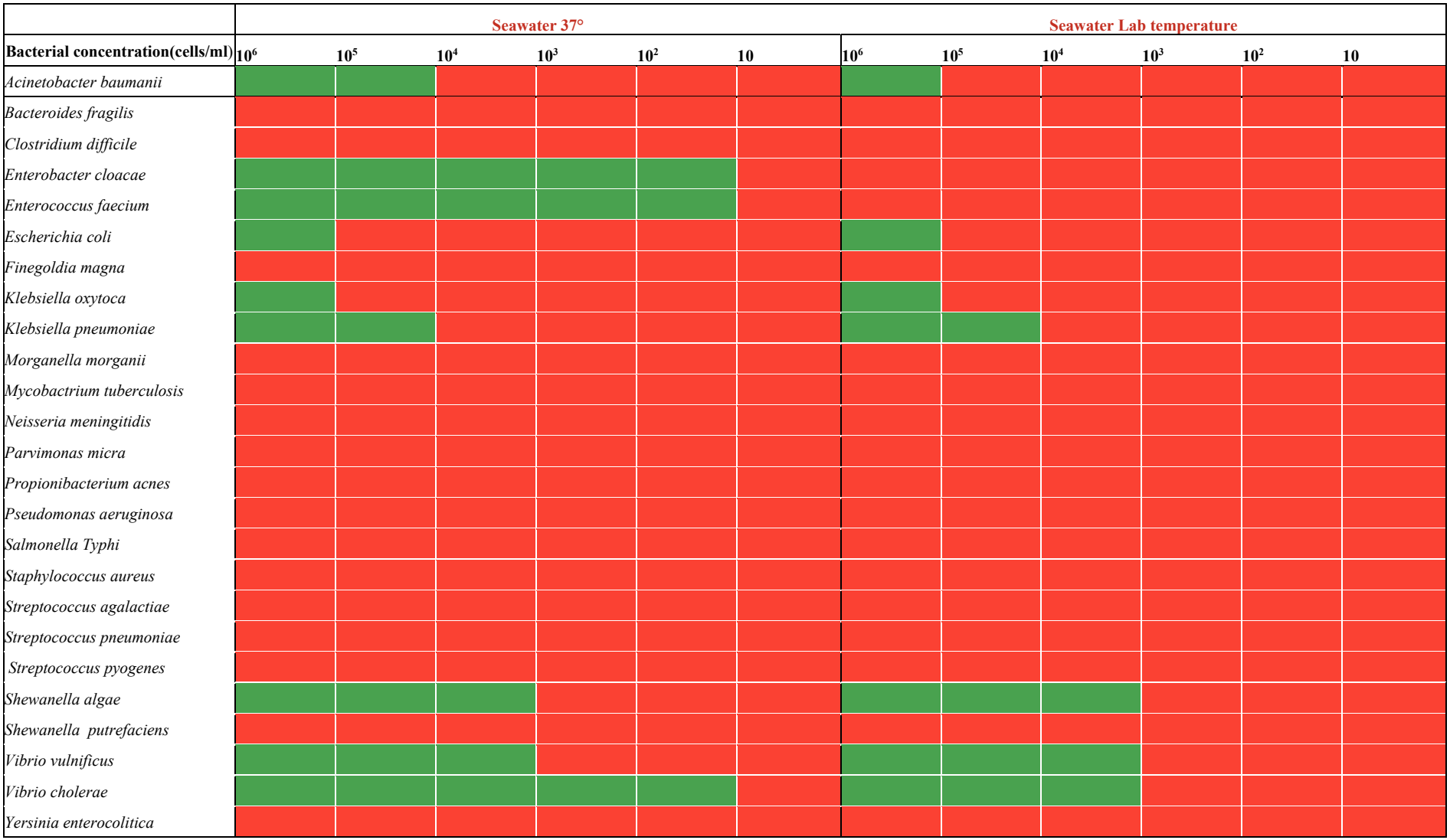
Species	CSUR	Aero-tolerant	Aero-intolerant
Acinetobacter baumanii	CSURP202	X	
Bacteroides fragilis	CSURP2573		X
Clostridium difficile	CSURP346		X
Enterobacter cloacae	CSURP2284	X	
Enterococcus faecium	CSURP745	X	
Escherichia coli	CSURP397	X	
Finegoldia magna	CSURP2342		X
Klebsiella oxytoca	CSURP2863	X	
Klebsiella pneumoniae	CSURP2714	X	
Morganella morganii	CSURP752	X	
Mycobacterium tuberculosis	2739	X	
Neisseria meningitidis	CSURP782	X	
Parvimonas micra	CSURP2136		X
Propionibacterium acnes	CSURP742		X
Pseudomonas aeruginosa	CSURP1003	X	
Salmonella Typhi	CSURP403	X	
Shewanella algae	CSURP601	X	
Shewanella putrefaciens	CSURP2224	X	
Staphylococcus aureus	CSURP2557	X	
Streptococcus agalactiae	CSURP2760	X	
Streptococcus pneumoniae	CSURP3392	X	
Streptococcus pyogenes	CSURP576	X	
Vibrio cholerae	CSURP2057	X	
Vibrio vulnificus	CSURP4422	X	
Yersinia enterocolitica	CSURP840	X	

Table 2: Evaluation of growth in sea water

Bacteria tested	**Sea water 40g /l**	
	Lab temperature	37°
Acinetobacter baumanii	+	+
Bacteroides fragilis	-	-
Clostridium difficile	-	-
Enterobacter cloacae	-	+
Enterococcus faecium	-	+
Escherichia coli	+	+
Finegoldia magna	-	-
Klebsiella oxytoca	+	+
Klebsiella pneumoniae	+	+
Morganella morganii	-	-
Mycobactrium tuberculosis	-	-
Neisseria meningitidis	-	-
Parvimonas micra	-	-
Propionibacterium acnes	-	-
Pseudomonas aeruginosa	-	-
Salmonella Typhi	-	-
Staphylococcus aureus	-	-
Streptococcus agalactiae	-	-
Streptococcus pneumoniae	-	-
Streptococcus pyogenes	-	-
Shewanella algae	+	+
Shewanella putrefaciens	-	-
Vibrio cholerae	+	+
Vibrio vulnificus	+	+
Yersinia enterocolitica	-	-

+: Growth; -: No growth

Table 3: Evaluation of growth in seawater according to inoculum density



Green: Growth

Red: No growth

Figure 1: **Growth evaluation test of strains on solid medium**

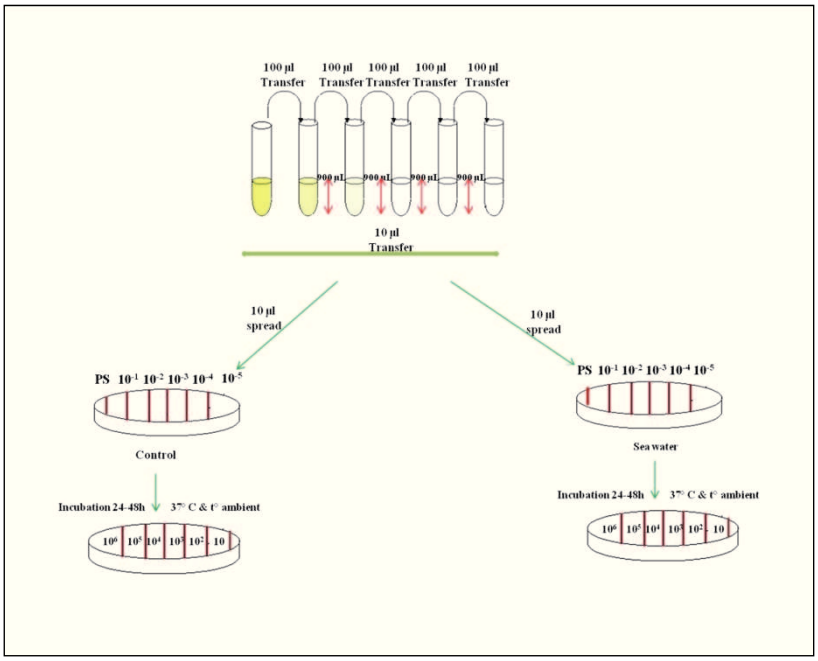
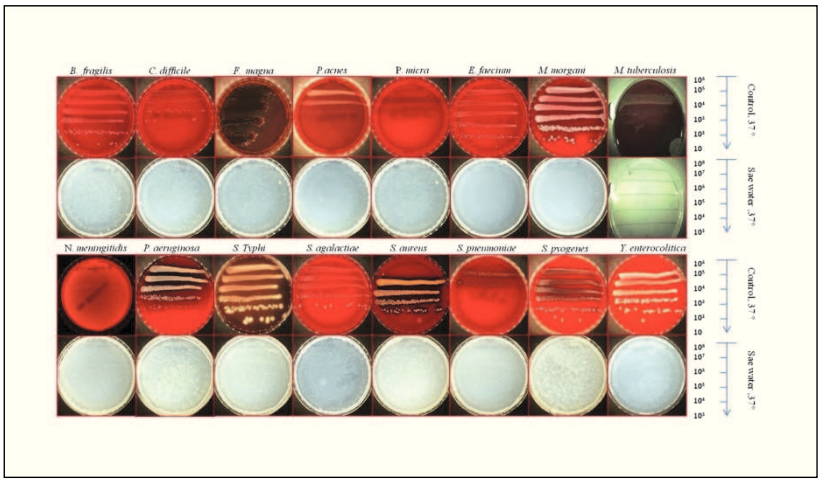
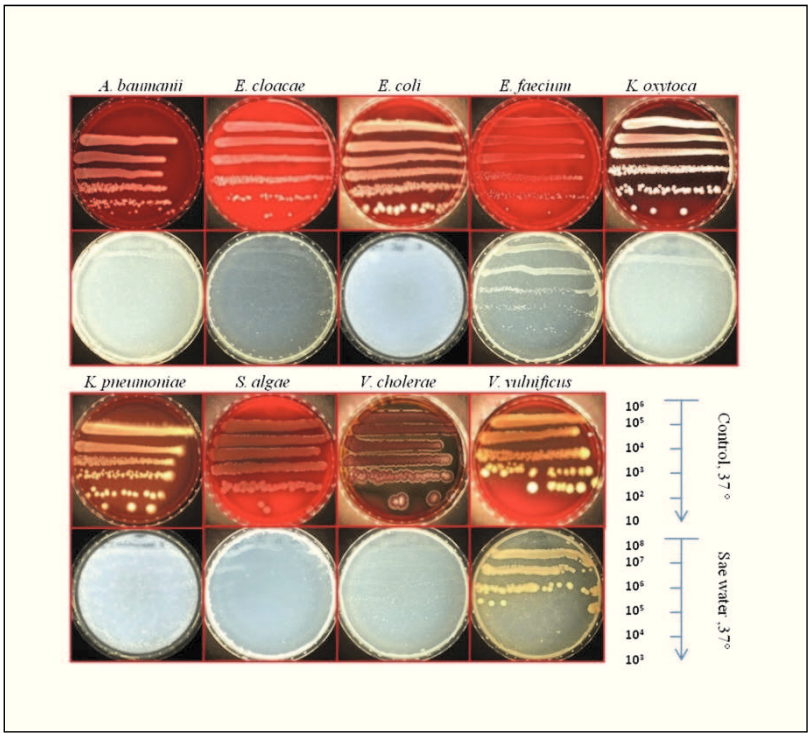


Figure 2: Summary of non-growth pathogens in seawater at 37°C.



The bacteria were arranged alphabetically and in aero-tolerant/aero-intolerant categories. The first five bacteria represented aero-intolerant bacteria and the rest was aero-tolerant bacteria.

Figure 3: Summary of pathogens growing in seawater at 37°C



CHAPITRE IV

Description des nouvelles espèces halophiles isolées par taxonogenomics

Avant-propos :

L'exploration du microbiote halophile par culturomics a permis l'isolation de 21 nouvelles espèces halophiles. Dans ce chapitre, nous présentons les descriptions d'une partie de ces nouvelles espèces halophiles isolées. Nous avons développé une méthode innovante appelée taxonogénomique. En effet, en raison de la disponibilité des données génomiques grâce au développement des nouveaux outils de séquençage de l'ADN, il a été introduit dans la description des nouvelles espèces bactériennes des informations génomiques, et protéomiques obtenues par analyse MALDI-TOF-MS. A cela s'ajoute les caractéristiques phénotypiques et la description du contenu en acides gras. Ce concept diffère des méthodes habituelles pour définir une nouvelle espèce, initialement basées sur des paramètres tels que le séquençage de l'ARNr 16S, la phylogénie, la teneur en G + C et l'hybridation ADN-ADN (DDH). Ces méthodes sont chronophages, fastidieuses et comportent de nombreuses limites.

En raison des délais observés dans la publication et l'officialisation des nouvelles espèces et tous les problèmes qu'elles engendrent, nous avons parallèlement mis en place un format interédiaire nommé « new species announcement ». Il comporte la description phénotypique et les informations phylogénétiques de l'espèce. Ce format de description intermédiaire permet de mettre rapidement à la disposition du monde scientifique les caractéristiques de base de la potentielle nouvelle espèce. Dans cette partie, nous présenterons la description de 4 espèces halophiles nouvellement isolées et 2 « new species announcement ».

Article 5: Non contiguous-finished genome sequence and description of *Virgibacillus senegalensis* sp. nov. a moderately halophilic bacterium isolated from the human gut.

E. Seck[1], J. Rathored[1], S. Khelaifia[1], O. Croce[1], C. Robert[1], C. Couderc[1], F. Di Pinto[1], C. Sokhna[2,3], D. Raoult[1,4] and J.-C. Lagier[1]

Publié dans New Microbes New Infections

Article 6: Non contiguous-finished genome sequence and description of *Planococcus massiliensis* sp. nov. a moderately halophilic bacterium isolated from the human gut.

E. H. Seck[1], S. A. Sankar[1], S. Khelaifia[1], O. Croce[1], C. Robert[1], C. Couderc[1], F. Di Pinto[1], C. Sokhna[2,3], P.-E. Fournier[1], D. Raoult[1,4] and J.-C. Lagier[1]

Publié dans New Microbes New Infections

Article 7: *Halomonas massiliensis'* sp. nov., a new halotolerant bacterium isolated from the human gut.

E. h. Seck[1], P.-E. Fournier[1], D. Raoult[1,2] and S. Khelaifia[1]

Publié dans New Microbes New Infections

Article 8: Description of "*Bacillus dakarensis* sp. nov.", "*Bacillus sinesaloumensis* sp. nov.", "*Gracilibacillus timonensis* sp .nov"."*Halobacillus massiliensis* sp. nov.", "*Lentibacillus massiliensis* sp. nov.", "*Oceanobacillus* senegalensis sp. nov.", "*Oceanobacillus timonensis* sp. nov.", "*Virgibacillus dakarensis* sp. nov.", "*Virgibacillus marseillensis* sp. nov.", nine halophilic new species isolated from human stool.

B. Senghor[1], E. H. Seck[1], S. Khelaifia[1], H. Bassène[2], C. Sokhna[2], P.-E. Fournier[1], D. Raoult[1] and J.-C. Lagier[1]

Publié dans New Microbes New Infections

Article 9: Non contiguous-finished genome sequence and description of *Gracilibacillus timonensis* sp. nov., a moderate halophile isolate from gut.

Publié dans New Microbes New Infections

Article 10: Non contiguous-finished genome sequence and description of *Bacillus salis* sp. nov. a moderately halophilic bacterium isolated from salt

El hadji Seck[1], Awa Diop[1], Jean Christophe Lagier, Frédéric Cadoret[1], Nicholas

Armstrong[1],Jérémy Delerce[1], Pierre-Edouard Fournier[1], Didier Raoult[1,2], Saber Khelaifia[1,*]

Publié dans New Microbes New Infections

- **Article 11: '*Gracilibacillus phocaeensis*' sp. nov., '*Sediminibacillus massiliensis*' sp. nov. and '*Virgibacillus ndiopensis*' sp. nov., three halophilic species isolated from salty human stools by culturomics**

B. Senghor[1], S. Khelaifia[1], H. Bassène[2], E. H. Seck[1], P.-E. Fournier[1], C. Sokhna[2], D. Raoult[1] and J.-C. Lagier[1]

Accepté dans New microbes New Infections

Virgibacillus senegalensis sp. nov., a new moderately halophilic bacterium isolated from human gut

E. Seck[1], J. Rathored[1], S. Khelaifia[1], O. Croce[1], C. Robert[1], C. Couderc[1], F. Di Pinto[1], C. Sokhna[2,3], D. Raoult[1,4] and J.-C. Lagier[1]

1) *Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM 63, CNRS 7278, IRD 198, Inserm 1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine,* 2) *Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes IRD 198, CNRS 7278, Aix-Marseille Université, Marseille, France,* 3) *Campus Commun UCAD-IRD of Hann, Dakar, Senegal and* 4) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

Virgibacillus senegalensis SK-1^T (= CSUR P1101 = DSM 28585) is the type strain of *V. senegalensis* sp. nov. It is an aerobic, Gram positive, moderately halophilic, motile bipolar flagellum isolated from a healthy Senegalese man. Here we describe the genomic and phenotypic characteristics of this isolate. The 3 755 098 bp long genome (one chromosome, no plasmid) exhibits a G + C content of 42.9% and contains 3738 protein-coding and 95 RNA genes.

Keywords: Culturomics, genome, human gut, moderately halophilic bacteria, taxonogenomics, *Virgibacillus senegalensis*
Original Submission: 12 August 2015; **Revised Submission:** 23 September 2015; **Accepted:** 24 September 2015
Article published online: 3 October 2015

Corresponding author: J.-C. Lagier, URMITE, UMR CNRS 7278, L'Institut de Recherche pour le Développement 198, INSERM U1095, Faculté de Médecine, Aix-Marseille Université, 27 Boulevard Jean Moulin, 13385 Marseille Cedex 5, France
E-mail: jclagier@yahoo.fr

Introduction

The concept of microbial culturomics is based on the variation of physicochemical parameters of the culture conditions so as to express the maximum of microbial diversity. It is based on rapid methods for identification, such as matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) and 16S rRNA amplification and sequencing for unidentified colonies. This concept considerably enriches the gut microbiota repertoire, including new species not previously isolated from humans [1,2].

This isolation was part of a culturomics study we undertook using high-salt-containing culture conditions to grow halophilic bacteria from human stool [1].

The typical parameters used to define bacterial species comprise 16S rRNA sequencing and phylogeny, G + C content genomic diversity and DNA-DNA hybridization (DDH). However, some limitations have been noted [3–6]. By using the availability of data in genomics through the development of new tools for sequencing DNA, we introduced a new taxonomic method for the description of new bacterial species. This concept, which we named taxonogenomics, includes their genomic features [7] and proteomic information obtained by MALDI-TOF analysis [8–17].

The genus *Virgibacillus* was first proposed by Heyndrickx in 1998 with the transfer of *Bacillus pantothenticus* to *Virgibacillus pantothenticus* [18]. To date, there are more than 25 recognized species [19]. These bacteria are positive, Gram-variable rods which are ellipsoidal to oval endospores and have DNA G + C content ranging from 36% to 43% [20]. These species were isolated from sediments of a salt lake [20–23], fermented seafood in traditional salt [24], a permafrost core collected from the Canadian high Arctic [25], a navy solar salt marsh [26,27], soil [28], seawater [29], field soil, a dairy product [30], residual wash water produced during processing wastewater, Spanish-style green table olives [31], saline sample of mud, salt

http://dx.doi.org/10.1016/j.nmni.2015.09.014

crust [32] and Thai fermented fish [33]. Here we present a brief classification and a set of features for strain SK-1T (= CSUR P1101 = DSM 28585), with a description of the complete genome sequence and annotation. We named this new isolate *Virgibacillus senegalensis*.

Materials and Methods

Sample and culture condition

The stool sample was collected from a healthy male Senegalese volunteer patient living in N'diop, a rural village in the Guinean–Sudanian zone in Senegal. After the patient provided signed informed consent, the sample was collected in a sterile pot and transported to our laboratory. The study and the assent procedure were approved by the National Ethics Committee of Senegal and by the ethics committees of the Institut Fédératif de Recherche 48, Faculty of Medicine, Marseille, France (agreement 09-022). The salt concentration of the stool specimen was determined by a digital refractometer (Fisher Scientific, Illkirch, France) and the pH with a pH meter (Cyberscan 510PH; Eutech Instruments, Singapore).

Strain SK-1T was isolated in February 2014 by aerobic culture on a homemade culture medium consisting of a Columbia agar culture medium (Sigma-Aldrich, Saint-Quentin Fallavier, France) modified by adding (per liter): $MgCl_2$ $6H_2O$, 5 g; $MgSO_4$ $7H_2O$, 5 g; KCl, 2 g; $CaCl_2$ $2H_2O$, 1 g; NaBr, 0.5 g; $NaHCO_3$, 0.5 g; glucose, 2 g; and 100 g/L of NaCl. The pH was adjusted to 7.5 with 10 M NaOH before autoclaving.

MALDI-TOF identification

An isolated colony was deposited in duplicate on a MALDI-TOF target to be analysed. A matrix of 1.5 µL (saturated solution of α-cyano-4-hydroxycinnamic acid diluted in 500 µL acetonitrile, 250 µL of acid tri-fluoro-acetic to 10%, and 250 µL of HPLC water) was used on each spot. This solution enables ionization and desorption of the homogeneous biological sample with which it crystallizes. The analysis was performed by a Microflex (Bruker Daltonics, Leipzig, Germany) device, and protein spectra were compared with those of the hospital database. A score was assigned indicating the reliability of the identification of the bacteria; above 1.9 was considered proper identification. Conversely, if the bacterium was not referenced in the database, sequencing the 16S rRNA was used to achieve the correct identification [34].

Identification by sequencing of 16S rRNA

Colonies not identified by the MALDI-TOF after three tests were suspended in 200 µL of distilled water for DNA extraction by EZ1 DNA Tissue Kit (Qiagen, Venlo, The Netherlands). The amplification of the 16S rRNA was performed by standard PCR in a thermocycler using the universal primer pair FD1 and rp2 according to the following amplification program: activation of the polymerase (95°C for 5 minutes), followed by 40 cycles (95°C 30 seconds, 52°C 45 seconds, 72°C 2 minutes), followed by 5 minutes at 72°C. The DNA amplified by this reaction was revealed by electrophoresis on 1.5% agarose gel. Once validated, the PCR product was purified and sequenced using the Big Dye Terminator Sequencing Kit using the internal primers 536F, 536R, 800F, 800R, 1050F and 1050R, as previously described [2].

Phylogenetic analysis

Phylogenetic analysis based on 16S rRNA of our isolates was performed to identify its phylogenetic affiliations with other near isolates, including other members of the genus *Virgibacillus*. MEGA 6 software (http://www.megasoftware.net/mega.php) allowed us to construct a phylogenetic tree. Sequence alignment of the different species was performed using Clustal W (http://www.clustal.org/clustal2/), and the evolutionary distance was calculated with the Kimura two-parameter model [35].

Biochemical, atmospheric and antimicrobial susceptibility tests

Biochemical tests were performed using the commercially available Api ZYM (bioMérieux, Marcy l'Étoile, France), API 50CH (bioMérieux) and 20 NE (bioMérieux) strips. The incubation time was 48 hours for the API 50CH and 20 NE, and 4 hours for Api ZYM. Growth of strain SK-1T was tested in aerobic atmosphere, in the presence of 5% CO_2 and also in anaerobic and microaerophilic atmospheres, created using AnaeroGen (Atmosphere Generation Systems, Dardily, France). Antibiotic susceptibility was determined by Müller-Hinton agar in a petri dish (bioMérieux). The following antibiotics were tested: doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole, imipenem and metronidazole.

Genome Sequencing Information

Genomic DNA preparation

We cultured our strain in the homemade culture. After 48 hours, bacteria grown on four petri dishes were resuspended in sterile water and centrifuged at 4°C at 2000 × *g* for 20 minutes. Cell pellets were resuspended in 1 mL Tris/EDTA/NaCl (10 mM Tris/HCl (pH7.0), 10 mM EDTA (pH8.0) and 300 mM

(NaCl)) and recentrifuged under the same conditions. The pellets were then resuspended in 200 µL Tris-EDTA buffer and proteinase K and kept overnight at 37°C for cell lysis. DNA was purified with phenol/chloroform/isoamyl alcohol (25:24:1), followed by an overnight precipitation with ethanol at −20°C. The DNA was resuspended in 205 µL Tris-EDTA buffer. DNA concentration was 155 ng/µL as measured by a Qubit fluorometer using the high-sensitivity kit (Life Technologies, Carlsbad, CA, USA).

Genome sequencing and assembly

Genomic DNA of *Virgibacillus senegalensis* was sequenced on the MiSeq Technology (Illumina, San Diego, CA, USA) with the mate pair strategy. The gDNA was bar coded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). gDNA was quantified by a Qubit assay with the high sensitivity kit (Life Technologies) to 155 ng/µL. The mate pair library was prepared with 1 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, Santa Clara, CA, USA) with a DNA 7500 lab chip. The DNA fragments ranged in size from 1 to 11 kb, with an optimal size at 4.008 kb. No size selection was performed, and 388.3 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal at 634 bp on the Covaris device S2 in microtubes (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 35.59 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 in a 2 × 251 bp. Total information of 10.6 Gb was obtained from a 1326K/mm^2 cluster density with a cluster passing quality control filters of 99.1% (24 492 260 clusters). Within this run, the index representation for *Virgibacillus senegalensis* was determined to be 7.06%. The 1 481 197 paired reads were filtered according to the read qualities. These reads were trimmed, then assembled using CLC genomicsWB4 software.

Genome annotation and comparison

Open reading frames (ORFs) were predicted using Prodigal [36] with default parameters, but the predicted ORFs were excluded if they spanned a sequencing gap region. The predicted bacterial protein sequences were searched against the GenBank database [37] and the Clusters of Orthologous Groups (COGs) database using BLASTP. The tRNAScanSE tool [38] was used to find tRNA genes, whereas ribosomal RNAs were found using RNAmmer [39] and BLASTn against the GenBank database. Lipoprotein signal peptides and the number of transmembrane helices were predicted using SignalP [40] and TMHMM [41], respectively. ORFans were identified if their BLASTP E value was lower than 1e^{-03} for alignment length greater than 80 aa. If alignment lengths were smaller than 80 aa, we used an E-value of 1e^{-05}. Such parameter thresholds have already been used in previous works to define ORFans. Artemis [42] was used for data management and DNA Plotter [43] for visualization of genomic features. Mauve alignment tool (version 2.3.1) was used for multiple genomic sequence alignment [44]. To estimate the mean level of nucleotide sequence similarity at the genome level, we used an in-lab pipeline software named Marseille Average Genomic Identity (MAGi) to calculate the average genomic identity of gene sequences (AGIOS) among compared genomes [45]. Briefly, this software combines the Proteinortho software [45] for detecting orthologous proteins in pairwise genomic comparisons, then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman-Wunsch global alignment algorithm. Genomes from the genus *Virgibacillus* and closely related genera were used for the calculation of AGIOS values. Here we compared the genome sequences of *Virgibacillus senegalensis* strain SK-1^T (GenBank accession number PRJEB1962) with those of *Virgibacillus kekensis* strain YIM kkny9 (NR_042744.1), *Virgibacillus albus* strain YIM 93624 (NR_109613.1), *Aquibacillus salifodinae* WSY08-1 (AB859945.1), *Virgibacillus halodenitrificans* DSM 10037 (AY543169), *Thalassobacillus devorans* MSP14 (JX518269.1), *Halobacillus dabanensis* HD 02 (HG931924.2), *Halobacillus kuroshimensis* DSM 18393 (AB195680.1), *Thalassobacillus devorans strain* XJSL7-8 (GQ903447.1), *Bacillaceae bacterium* EFN-4 (EU817569.1), *Virgibacillus marismortui* strain M3-23 (GQ282501.1), *Halobacillus salinus* strain GSP59 (AY505517.1), *Virgibacillus alimentarius* J18 (GU202420), *Pseudomonas aeruginosa* PAO1 (NR_074828.1) and *Virgibacillus massiliensis* (CCDP010000001).

Results

Phenotypic description

Strain SK-1T was isolated in February 2014 (Table 1) by aerobic culture on a homemade culture medium at 37°C after 48 hours. No significant MALDI-TOF result for the strain SK-1^T against

TABLE 1. Classification of *Virgibacillus senegalensis* strain SK-1

Property	Term
Current classification	Domain: *Bacteria* Phylum: *Firmicutes* Class: *Bacilli* Order: *Bacillales* Family: *Bacillaceae* Genus: *Virgibacillus* Species: *Virgibacillus senegalensis* Type strain: SK-1
Gram stain	Positive
Cell shape	Rod shaped
Motility	Motile by polar flagellum
Sporulation	Endospore forming
Temperature range	Mesophile
Optimum temperature	37°C
pH	pH 5 to 9
Optimum pH	7.5
Salinity	0.5–10%
Optimum salinity	7.5%
Oxygen requirement	Aerobic

our MALDI-TOF database was obtained, suggesting that our isolate was not a previously known species. We added the specter of SK-1T strain to our database (Fig. 1). The gel view allowed us to observe the spectral differences with other members of *Virgibacillus* genus (Fig. 2). PCR-based identification of the 16S rRNA of our new isolate (GenBank accession number LK021111) yielded 96.3% 16S rRNA sequence similarity with the reference *Virgibacillus kekensis* (GenBank accession number NR042744), the phylogenetically closest validated *Virgibacillus* species (Fig. 3).

After growth for 24 hours on our homemade culture medium at 37°C, the surface colonies were circular, greyish, shiny and smooth, with a diameter of 1 to 2 mm. *V. senegalensis* is Gram positive (Fig. 4).

Growth was observed at temperatures ranging from 25 to 40°C, with an optimum at 37°C. The growth required a salinity ranging from 5 to 200 g/L of NaCl (optimum at 75 g/L). The optimum pH for growth was 7.5 (pH range 5 to 9). Growth of the strain SK-1^T was tested in an aerobic atmosphere, in the presence of 5% CO_2 and also in anaerobic and microaerophilic atmospheres created using AnaeroGen (Atmosphere Generation Systems), respectively. The strain was strictly aerobic and also grew in the presence of 5% CO_2 but did not grow in an anaerobic atmosphere. The size and ultrastructure of cells were determined by negative staining transmission electron microscopy 2 to 6 µm in length and 0.5 µm in diameter (Fig. 5). Using the commercially available Api ZYM, Api 20NE (bioMérieux), to characterize the biochemical *V. senegaiensis* strain SK-1^T, positive reactions were observed for urease, β-glucosidase,

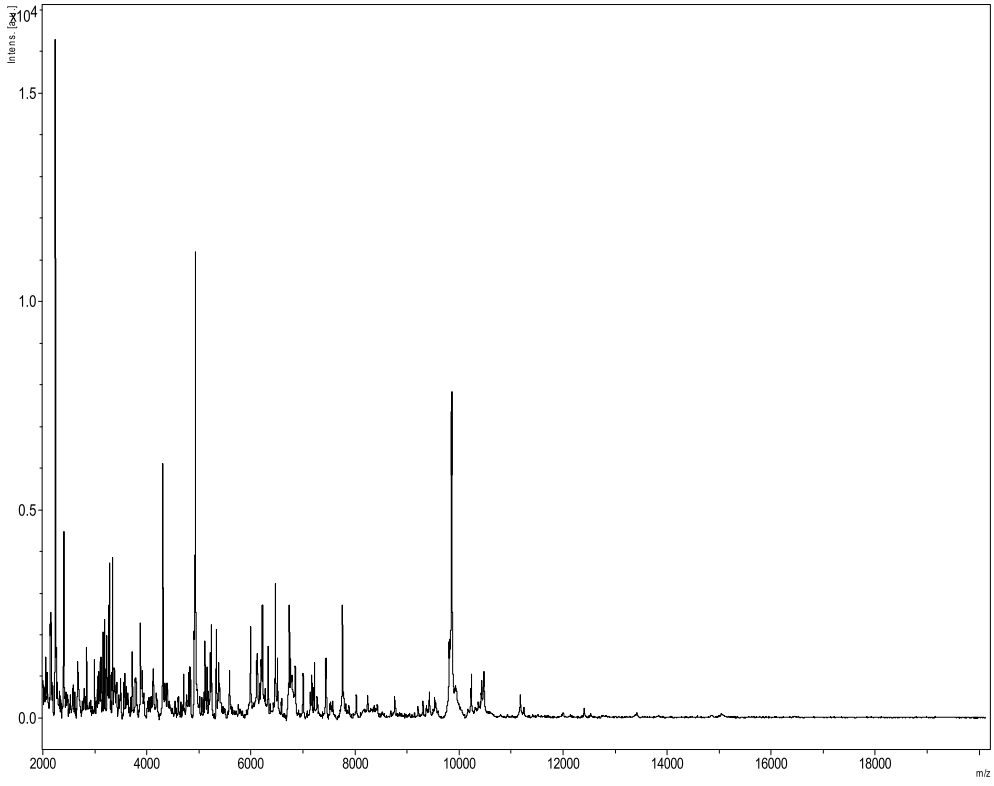


FIG. 1. Reference mass spectrum from *Virgibacillus senegalensis* sp. nov. SK-1^T. Spectra from 12 individual colonies were compared and reference spectrum generated.

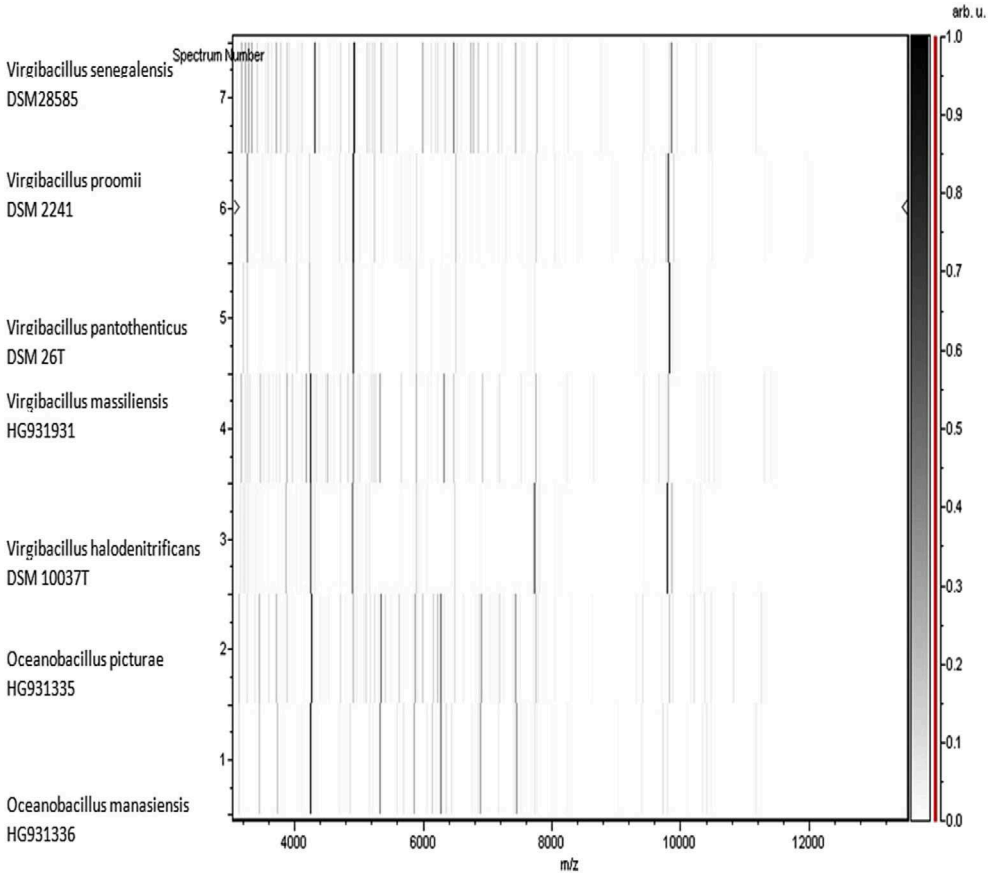


FIG. 2. Gel view comparing *Virgibacillus senegalensis* sp. nov. SK-1^{T}to members of family *Virgibacillus* and *Oceanobacillus*. Gel view displays raw spectra of all loaded spectrum files arranged in pseudo-gel-like look. X-axis records m/z value. Left y-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity is expressed by greyscale scheme code. Color bar and right y-axis indicate relation between color peak; peak intensity expressed in arbitrary units. Displayed species are indicated at left.

protease, β-galactosidase and arginine hydrolase. All other tested reactions were negative, notably nitrate reduction alkaline phosphatase and N-acetyl-β-glucosaminidase. The stain was also catalase and oxidase negative. Substrate oxidation and assimilation were examined with an API 50CH strip (bioMérieux) at 37°C. Negative reactions were obtained for D-mannose, D-lactose, L-arabinose, D-galactose, D-ribose, D-sucrose, D-fructose, D-glucose, D-mannitol and D-maltose. Phenotypic characteristics were compared to those of the most closely related species (Table 2).

Finally, antimicrobial susceptibility testing demonstrate that the strain SK-1^T was susceptible to doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole and imipenem, but resistant to metronidazole.

Genome Sequencing Information

Genome properties

The draft genome of *V. senegalensis* consists of nine scaffolds with 59 contigs.

The genome is, 755 098 bp long with a 42.9% G + C content (Table 3, Fig. 6). Of the 3833 predicted genes, 3738 (96.46%) were protein-coding genes and 95 (2.44%) were RNAs (14 genes were 5S rRNA, five genes 16S rRNA, seven genes 23S rRNA, 69 genes tRNA). A total of 2773 genes (62.82%) were assigned a putative function, 2427 genes (65.43%) were assigned to COGs and 155 genes (4.04%) contained peptide signals, whereas 980 (25.56%) genes had transmembrane helices. A total of 245 genes were identified as ORFans (6.39%). The remaining genes were annotated as hypothetical proteins. The properties and statistics of the genome are summarized in

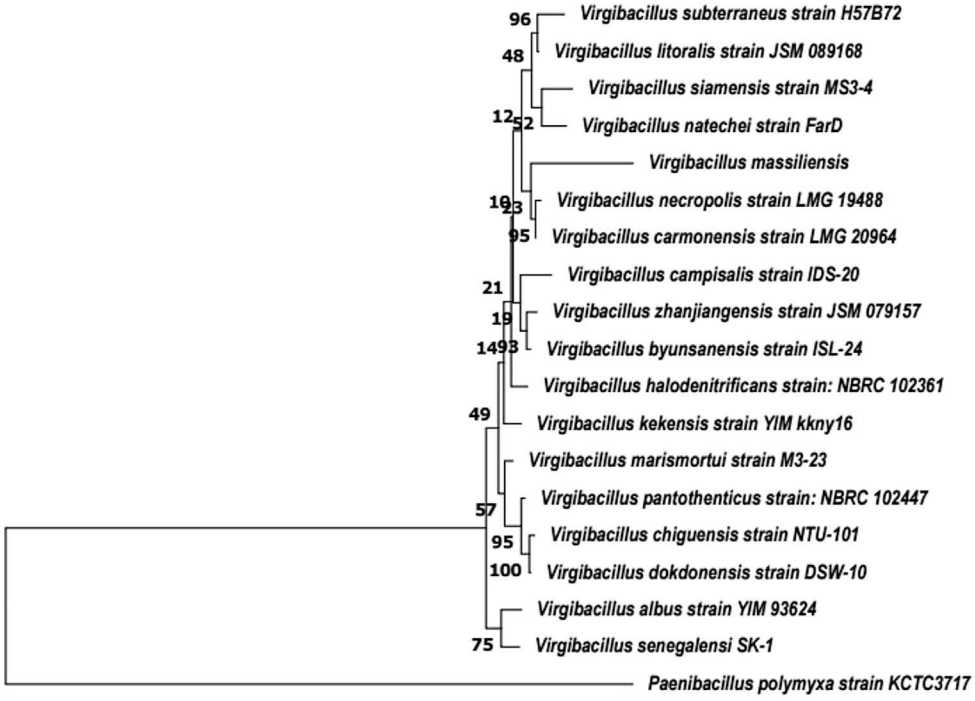


FIG. 3. Phylogenetic tree highlighting position of *Virgibacillus senegalensis* strain SK-1T (= CSUR P1101 = DSM 28585) relative to other type strains of *Virgibacillus albus* strain YIM 93624 (NR_109613.1), *Virgibacillus kekensis* strain YIM kkny16 (NR_042744.1), *Virgibacillus alimentarius* strain J18 (NR_108710.1), *Virgibacillus marismortui* strain M3-23 (GQ282501.1), *Virgibacillus necropolis* strain LMG 19488 (NR_025472.1), *Virgibacillus carmonensis* strain LMG 20964 (NR_025481.1), *Virgibacillus subterraneus* strain H57B72 (FJ746573.1), *Virgibacillus zhanjiangensis* strain JSM 079157 (FJ425904.1), *Virgibacillus litoralis* strain JSM 089168 (FJ425909.1), *Virgibacillus dokdonensis* strain DSW-10 (NR_043206.1), *Virgibacillus siamensis* strain MS3-4 (AB365482.1), *Virgibacillus salarius* strain SA-Vb1 (NR_041270.1), *Virgibacillus halophilus* strain 5B73C (NR_041358.1), *Virgibacillus natechei* strain FarD (NR_132721.1), *Virgibacillus chiguensis* strain NTU-101 (NR_044086.1), *Virgibacillus dokdonensis* strain DSW-10 (NR_043206.1), *Virgibacillus campisalis* strain IDS-20 (GU586225.1), *Virgibacillus pantothenticus* strain NBRC 102447 (AB681789.1), *Virgibacillus halodenitrificans* strain NBRC 102361 (AB681753.1), *Virgibacillus byunsanensis* strain ISL-24 (FJ357159.1), *Virgibacillus massiliensis* strain Vm-5 (HG931931.1) and *Paenibacillus polymyxa* strain KCTC3717 (AY359637.1). GenBank accession numbers are indicated in parentheses. Sequences were aligned using Clustal W (http://www.clustal.org/clustal2/), and phylogenetic inferences were obtained using maximum-likelihood method within MEGA 6 (http://www.megasoftware.net/mega.php). *Paenibacillus polymyxa* was used as outgroup. Scale bar = 0.005% nucleotide sequence divergence.

Tables 3 and 4. The distribution of genes into COGs functional categories is presented in Table 5.

Genome comparison

The draft genome of V. senegalensis SK-1^T is smaller than those of Halobacillus kuroshimensis DSM 18393, Virgibacillus halodenitrificans DSM10037, Thalassobacillus devorans XJSL7-8, Thalassobacillus devorans XJSL7-8, Halobacillus dabanensis HD 02 and Pseudomonas aeruginosa PAO1 (3.85, 3.92, 3.94, 4.1 and 6.26 Mb, respectively) but larger than that of Virgibacillus alimentarius J18 (3.05 Mb). The G + C content of V. senegalensis SK-1^T is smaller than those of Halobacillus kuroshimensis DSM 18393 and Pseudomonas aeruginosa PAO1 (47.0% and 66.60%, respectively) and larger than those of Virgibacillus alimentarius J18, Virgibacillus halodenitrificans DSM10037 and Halobacillus dabanensis HD 02 (37.1%, 37.4% and 41.5% respectively) but equal to Thalassobacillus devorans XJSL7-8 (42.9%). Protein-coding genes of V. senegalensis SK-1^T were smaller than those of Virgibacillus halodenitrificans DSM10037, Thalassobacillus devorans XJSL7-8, Halobacillus kuroshimensis DSM 18393, Halobacillus dabanensis HD 02 and Pseudomonas aeruginosa PAO1 (3748, 3752, 3832, 3835 and 5572 Mb, respectively) but larger than those of Virgibacillus alimentarius J18 (2889 Mb). Total gene content of V. senegalensis SK-1^T (3883) is smaller than those of Halobacillus kuroshimensis DSM 18393, Halobacillus dabanensis HD 02 and Pseudomonas aeruginosa PAO1 (3915, 4011, and 5697 respectively) but larger than those of Thalassobacillus devorans XJSL7-8, Virgibacillus halodenitrificans DSM10037 and Virgibacillus alimentarius J18 (3840, 3822 and 3022, respectively).

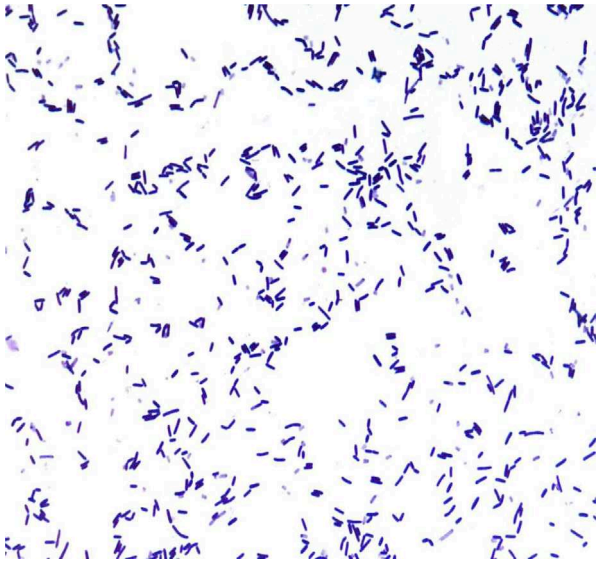


FIG. 4. Gram staining of *Virgibacillus senegalensis* sp. nov. SK-1^T.

Among species with standing in nomenclature, AGIOS values ranged from 66.41% between *V. senegalensis* SK-1^T and *Halobacillus kuroshimensis* DSM 18393 to 73.39% between *Halobacillus dabanensis* HD 02 and *Halobacillus kuroshimensis* DSM 18393. To evaluate the genomic similarity among studied strains, in addition to AGIOS [7], which was designed to be independent from DDH, we determined a digital DDH that exhibited a high correlation with DDH [46,47]. Digital DDH ranged from 18.4% to 27.2% between the different species tested (Table 6, Fig. 7).

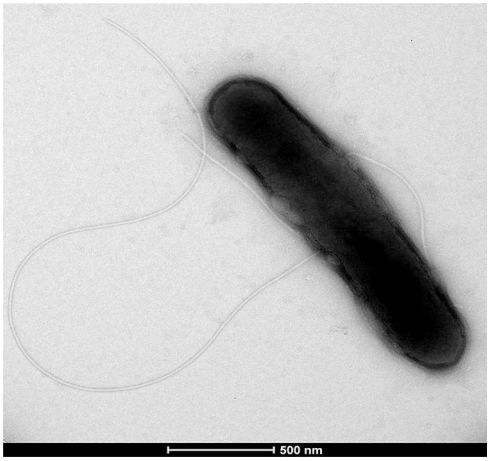


FIG. 5. Transmission electron microscopy of *Virgibacillus senegalensis* sp. nov. SK-1^T. Cells were observed on Tecnai G20 transmission electron microscope operated at 200 keV. Scale bar = 500 nm.

TABLE 2. **Differential characteristics of *Virgibacillus* species**

Property	*V. senegalensis*	*V. massiliensis*	*V. olivae*	*V. salarius*	*V. marismortui*	*V. sediminis*	*V. xinjiangensis*	*V. kekensis*	*V. halodenitrificans*	*V. proomii*	*V. dokdonensis*
Cell diameter (µm)	0.6–0.9	0.5–0.8	0.4–0.6	0.6–0.9		0.4–0.7	1.4–2.4	0.3–0.5	0.6–0.8	0.5–0.7	
Oxygen requirement	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic
Gram stain	+	+	+	+	+	+	+	+	+	+	+
Salt requirement	+	+	+	+	−	+	−	−	+	NA	+
Motility	+	+	+	+	+	+	+	+	+	+	+
Endospore formation	+	+	+	+	+	+	+	+	+	+	+
Indole	−	−	−	−	−	−	−	−	−	−	−
Production of											
Alkaline phosphatase	−	−	NA	NA		−	NA	−	NA	NA	−
Catalase	−	+	+	+	+	+	+	+	+	+	+
Oxidase	−	+	+	−	+	+	−	+	+	NA	+
Nitrate reductase	−	+	NA	−	NA	+	−	+	+	NA	−
Urease	+	−	NA	−	NA	+	−	−	+	NA	−
β-Galactosidase	+	−	NA	−	−	−	−	−	+	+	−
N-acetyl-glucosamine	−	−	NA	+	+	−	NA	−	NA	+	−
Acid from:											
L-Arabinose	−	−	NA	−	−	−	−	−	−	−	−
Ribose	−	−	NA	NA	NA	+	−	−	NA	+	+
D-Mannose	−	+	−	+	+	−	−	+	+	+	+
D-Manitol	−	+	NA	−	−	−	−	w	+	+	−
D-Sucrose	−	+	−	−	−	−	+	−	−	+	+
D-Glucose	−	+	−	+	+	+	+	+	+	+	+
D-Fructose	−	+	+	+	+	+	+	−	+	+	+
D-Maltose	−	+	−	+	−	+	−	+	+	+	−
D-Lactose	−	−	−	−	−	−	−	−	+	+	+
Habitat	Human gut	Human gut	Waste wash water	Salt lake	Mural paintings	Salt lake	Salt lake	Salt lake	Solar saltern	Soil	Soil

NA, data not available; w, weak reaction.

TABLE 3. Nucleotide content and gene count levels of genome

Attribute	Value	% of total[a]
Size (bp)	3 755 098	100
G + C content (%)	1 610 937	42.9
Coding region (bp)	3 129 675	83.34
Total genes	3883	100
RNA genes	95	2.44
Protein-coding genes	3738	98.46
Genes with function prediction	2773	62.82
Genes assigned to COGs	2421	65.43
Genes with peptide signals	155	4.04
Genes with transmembrane helices	980	25.56
CRISPRs	2	0.05
Genes with Pfam domains	2011	52.46

COGs, Clusters of Orthologous Groups database; CRISPR, clustered regularly interspaced short palindromic repeat.
[a]Total is based on either size of genome (bp) or total number of protein-coding genes in annotated genome.

TABLE 4. Number of genes associated with the 25 general COGs functional categories

Code	Value	% value[a]	Description
J	172	4.60	Translation
A	0	0	RNA processing and modification
K	262	7.01	Transcription
L	196	5.24	Replication, recombination and repair
B	1	0.03	Chromatin structure and dynamics
D	30	0.80	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	49	1.31	Defense mechanisms
T	153	4.09	Signal transduction mechanisms
M	156	4.17	Cell wall/membrane biogenesis
N	64	1.71	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	51	1.36	Intracellular trafficking and secretion
O	92	2.46	Posttranslational modification, protein turnover, chaperones
C	137	3.67	Energy production and conversion
G	275	7.36	Carbohydrate transport and metabolism
E	305	8.16	Amino acid transport and metabolism
F	82	2.19	Nucleotide transport and metabolism
H	93	2.49	Coenzyme transport and metabolism
I	102	2.73	Lipid transport and metabolism
P	202	5.40	Inorganic ion transport and metabolism
Q	74	1.98	Secondary metabolites biosynthesis, transport and catabolism
R	448	11.99	General function prediction only
S	287	7.68	Function unknown
—	352	9.06	Not in COGs

COGs, Clusters of Orthologous Groups database.
[a]Total is based on total number of protein-coding genes in annotated genome.

Conclusion

On the basis of phenotypic, genomic and phylogenetic analyses, we formally propose the creation of *Virgibacillus senegalensis* sp. nov., represented here by the SK-1T strain. The strain was isolated from a stool sample of a Senegalese healthy individual.

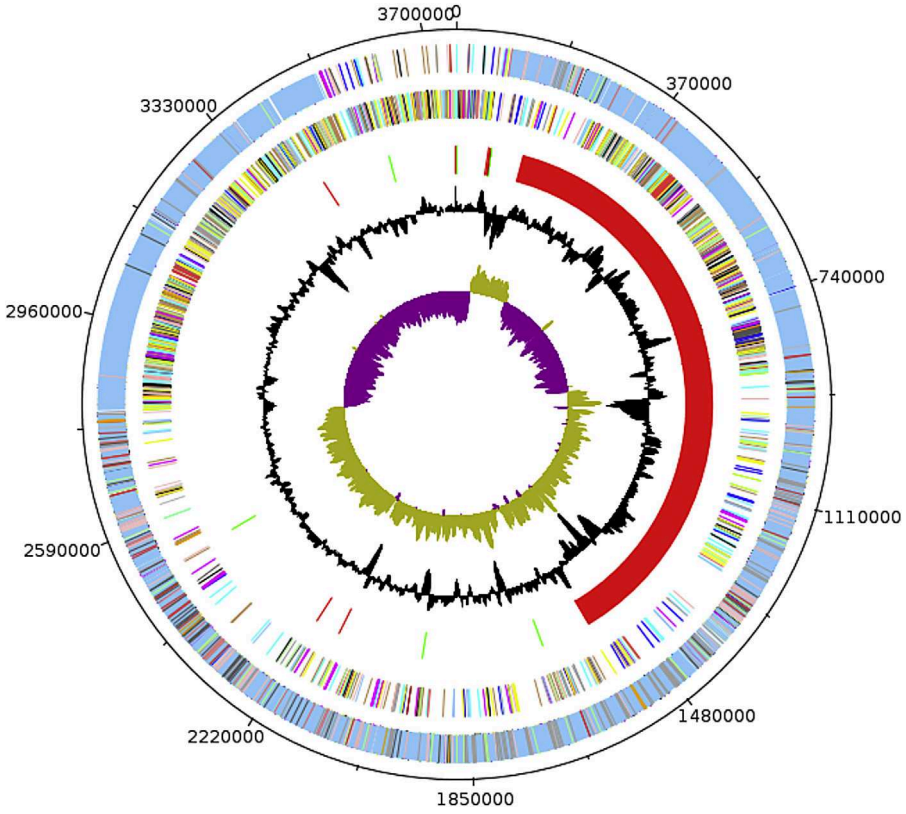


FIG. 6. Graphical circular map of *Virgibacillus senegalensis* sp. nov. SK-1^T chromosome. From outside in, outer two circles show open reading frames oriented in forward (colored by COGs categories) and reverse (colored by COGs categories) directions, respectively. Third circle marks tRNA genes (green). Fourth circle shows percentage G + C content plot. Innermost circle shows GC skew, purple indicating negative values and olive positive values. COGs, Clusters of Orthologous Groups database.

TABLE 5. Numbers of orthologous proteins shared between genomes (upper right), average percentage similarity of nucleotides corresponding to orthologous proteins shared between genomes (lower left) and numbers of proteins per genome (bold)

	Virgibacillus senegalensis	*Halobacillus dabanensis*	*Halobacillus kuroshimensis*	*Thalassobacillus devorans*	*Virgibacillus alimentarius*	*Virgibacillus halodenitrificans*	*Pseudomonas aeruginosa*	*Virgibacillus massiliensis*
V. senegalensis	**3378**	1791	1786	1776	1446	1741	601	588
H. dabanensis	66.82	**4063**	2218	2122	1581	1866	697	642
H. kuroshimensis	66.41	73.39	**3926**	2119	1554	1861	695	630
T. devorans	67.76	69.22	68.46	**3880**	1596	1923	687	618
V. alimentarius	67.16	66.25	64.45	66.55	**3119**	1663	607	604
V. halodenitrificans	67.19	66.41	64.71	66.67	71.39	**3876**	661	676
P. aeruginosa	53.94	52.73	55.86	53.96	49.86	50.02	**5681**	244
V. massiliensis	66.21	65.24	63.48	65.57	69.67	70.87	50.45	**1768**

TABLE 6. Pairwise comparison of *Virgibacillus senegalensis* with eight other species[a]

	Virgibacillus senegalensis	*Halobacillus dabanensis*	*Halobacillus kuroshimensis*	*Thalassobacillus devorans*	*Virgibacillus alimentarius*	*Virgibacillus halodenitrificans*	*Pseudomonas aeruginosa*	*Virgibacillus massiliensis*
V. senegalensis	100% ± 00	24.3% ± 2.55	26.7% ± 2.54	22.8% ± 2.54	26.4% ± 2.54	21.8% ± 2.56	20.8% ± 2.53	22.8% ± 2.53
H. dabanensis		100% ± 00	26.3% ± 2.55	24.4% ± 2.56	27.7% ± 2.55	27% ± 2.55	21.2% ± 2.57	20.0% ± 2.53
H. kuroshimensis			100% ± 00	24.6% ± 2.55	27.2% ± 2.55	24.7% ± 2.57	21.8% ± 2.54	19.5% ± 2.53
T. devorans				100% ± 00	22.1% ± 2.55	21.3% ± 2.57	18.4% ± 2.56	20.1% ± 2.53
V. alimentarius					100% ± 00	23% ± 2.56	21.1% ± 2.53	20.0% ± 2.54
V. halodenitrificans						100% ± 00	22.5% ± 2.54	21.3% ± 2.55
P. aeruginosa							100% ± 00	20.3% ± 2.52
V. massiliensis								100% ± 00

[a]Comparison made using GGDC, formula 2 (DDH estimates based on identities/HSP length). Confidence intervals indicate 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 16S rRNA (Fig. 1) and phylogenomic analyses as well as GGDC results. DDH, DNA-DNA hybridization; HSP, high-scoring pair.

Taxonomic and nomenclatural proposals

Description of *Virgibacillus senegalensis* sp. nov.

Virgibacillus senegalensis (se.ne.ga.len'sis. L. masc. adj. *senegalensis* of Senegalia, the Roman name for Senegal, where the type strain was isolated). Growth occurred between 15°C and 45°C on a homemade culture medium (described above), with optimal growth observed at 37°C in an aerobic atmosphere. Strain SK-1T required a salinity ranging from 5 to 200 g/L of NaCl (optimum at 100 g/L). The optimum pH for growth was 7.5 (range, 5 to 9). The strain SK-1T was strictly aerobic and also grew in the presence of 5% CO_2, but no growth was observed under anaerobic and microaerophilic conditions. The colonies of the strain SK-1T were circular, greyish, shiny and smooth, with a diameter of 2 to 6 mm. Cells stained Gram positive. They were motile by polar flagella, spore forming (2 to 6 µm in length and 0.5 µm in diameter) and generally occurred individually or in pairs. Strain SK-1T is catalase and oxidase negative. Using API 50 CH and API20 NE (bioMérieux), strain SK-1T was positive

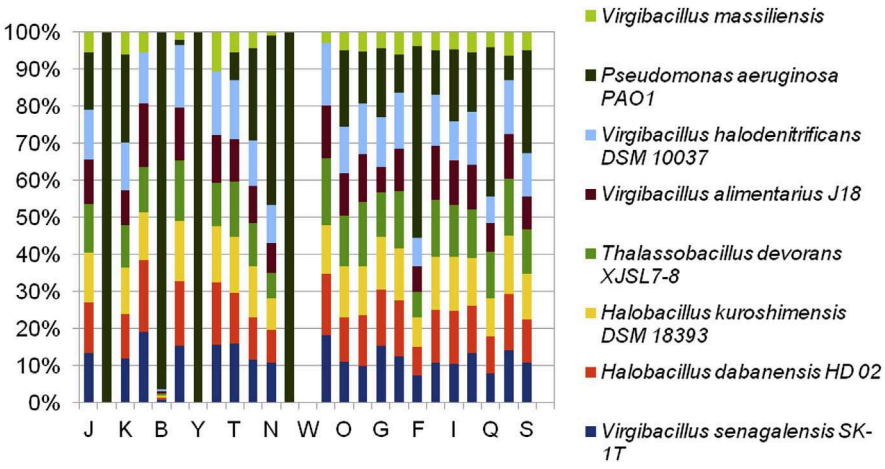


FIG. 7. Distribution of functional classes of predicted genes according to COGs proteins. COGs, Clusters of Orthologous Groups (COGs) database.

for reduction of nitrates but negative for phosphatase alkaline activity, â-galactosidase, áN-acetyl-â-glucosaminidase and urease. Strain SK-1T was negative for ribose, L-arabinose and D-lactose assimilation and positive for D-glucose, D-fructose, D-mannose, D-mannitol, D-maltose and D-sucrose. The strain SK-1T was susceptible to doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole and imipenem, but resistant to metronidazole.

The percentage of G + C content of the genome is 42.9%. The 16S rRNA and genome sequences are deposited in GenBank under accession numbers LK021111. The habitat of the microorganism is the human digestive tract. The type strain SK-1T (= CSUR P1101, = DSM 28585) was isolated from a stool specimen of a healthy Senegalese man.

Acknowledgements

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

Conflict of Interest

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;1:237–64.
[2] 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.
[3] Tindall BJ, Rosselló-Móra R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. Int J Syst Evol Microbiol 2010;60:249–66.
[4] Wayne LG, Brenner DJ, Colwell PR, Grimont PAD, Kandler O, Krichevsky, et al. Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematic. Int J Syst Bacteriol 1987;37:463–4.
[5] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[6] Rosselló-Móra R. DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In: Stackebrandt E, editor. Molecular identification, systematics, and population structure of prokaryotes. Berlin: Springer; 2006. p. 23–50.
[7] 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 new bacterial species. Int J Syst Evol Microbiol 2014;64:384–91.
[8] Welker M, Moore ER. Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 2011;34:2–11.
[9] Lagier JC, El Karkouri K, Nguyen TT, Armougom F, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Anaerococcus senegalensis* sp. nov. Stand Genomic Sci 2012;6:116–25.
[10] Lagier JC, Armougom F, Mishra AK, Nguyen TT, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Alistipes timonensis* sp. nov. Stand Genomic Sci 2012;6:315–24.
[11] Roux V, El Karkouri K, Lagier JC, Robert C, Raoult D. Non-contiguous finished genome sequence and description of *Kurthia massiliensis* sp. nov. Stand Genomic Sci 2012;7:221–32.
[12] Kokcha S, Ramasamy D, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Brevibacterium senegalense* sp. nov. Stand Genomic Sci 2012;7:233–45.
[13] Ramasamy D, Kokcha S, Lagier JC, Nguyen TT, Raoult D, Fournier PE. Genome sequence and description of *Aeromicrobium massiliense* sp. nov. Stand Genomic Sci 2012;7:246–57.
[14] Lagier JC, Elkarkouri K, Rivet R, Couderc C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Senegalemassilia anaerobia* gen. nov., sp. nov. Stand Genomic Sci 2013;7:343–56.
[15] Lagier JC, El Karkouri K, Mishra AK, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Enterobacter massiliensis* sp. nov. Stand Genomic Sci 2013;7:399–412.
[16] Hugon P, Mishra AK, Lagier JC, Nguyen TT, Couderc C, Raoult D, et al. Non-contiguous finished genome sequence and description of *Brevibacillus massiliensis* sp. nov. Stand Genomic Sci 2013;8:1–14.
[17] Lagier JC, Gimenez G, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Herbaspirillum massiliense* sp. nov. Stand Genomic Sci 2012;7:200–9.
[18] Heyndrickx ML, Kersters K, DeVos P, Forsyth G, Logan NA. Virgibacillus: a new genus to accommodate *Bacillus pantothenticus* (Proom and Knight 1950). Emended description of *Virgibacillus pantothenticus*. Int J Syst Bacteriol 1998;48:99–106.
[19] Heyrman J, Logan NA, Busse HJ, Balcaen A, Lebbe L, Rodriguez-Diaz M, et al. *Virgibacillus carmonensis* sp. nov., *Virgibacillus necropolis* sp. nov. and *Virgibacillus picturae* sp. nov., three novel species isolated from deteriorated mural paintings, transfer of the species of the genus *Salibacillus* to *Virgibacillus*, as *Virgibacillus marismortui* comb. nov. and *Virgibacillus salexigens* comb. nov., and emended description of the genus *Virgibacillus*. Int J Syst Evol Microbiol 2003;53:501–11.
[20] Carrasco IJ, Marquez MC, Ventosa A. *Virgibacillus salinus* sp. nov., a moderately halophilic bacterium from sediment of a saline lake. Int J Syst Evol Microbiol 2009;59:3068–73.
[21] Chen YG, Cui XL, Wang YX, Zhang YQ, Tang SK, Li WJ, et al. *Virgibacillus sediminis* sp. nov., a moderately halophilic bacterium isolated from a salt lake in China. Int J Syst Evol Microbiol 2009;59:2058–63.
[22] Jeon C, Kim J, Park DJ, Xu LH, Jiang CL, Kim CJ. *Virgibacillus xinjiangensis* sp. nov., isolated from a Salt Lake of Xin-jiang Province in China. J Microbiol 2009;47:705–9.
[23] Zhang YJ, Zhou Y, Ja M, Shi R, Chun-Yu WX, Yang LL, et al. *Virgibacillus albus* sp. nov., a novel moderately halophilic bacterium isolated from Lop Nur salt lake in Xinjiang Province, China. Antonie Van Leeuwenhoek 2012;102:553–60.
[24] Kim J, Jung MJ, Roh SW, Nam YD, Shin KS, Bae JW. *Virgibacillus alimentarius* sp. nov., isolated from a traditional Korean food. Int J Syst Evol Microbiol 2011;61:2851–5.
[25] Niederberger TD, Steven B, Charvet S, Barbier B, Whyte LG. *Virgibacillus arcticus* sp. nov., a moderately halophilic, endospore-forming bacterium from permafrost in the Canadian high Arctic. Int J Syst Evol Microbiol 2009;59:2219–25.

[26] Yoon JH, Kang SJ, Jung YT, Lee KC, Oh HW, Oh TK. *Virgibacillus byunsanensis* sp. nov., isolated from a marine solar saltern. Int J Syst Evol Microbiol 2010;60:291–5.
[27] Lee SY, Kang CH, Oh TK, Yoon JH. *Virgibacillus campisalis* sp. nov., from a marine solar saltern. Int J Syst Evol Microbiol 2012;62:347–51.
[28] An SY, Asahara M, Goto K, Kasai H, Yokota A. *Virgibacillus halophilus* sp. nov., spore-forming bacteria isolated from soil in Japan. Int J Syst Evol Microbiol 2007;57:1607–11.
[29] Yoon JH, Kang SJ, Lee SY, Lee MH, Oh TK. *Virgibacillus dokdonensis* sp. nov. isolated from a Korean island, Dokdo, located at the edge of the East Sea in Korea. Int J Syst Evol Microbiol 2005;55:1833–7.
[30] Seiler H, Wenning M. *Virgibacillus halotolerans* sp. nov., isolated from a dairy product. Int J Syst Evol Microbiol 2013;63:3358–63.
[31] Quesada T, Aguilera M, Morillo JA, Ramos-Cormenzana A, Monteoliva-Sanchez M. *Virgibacillus olivae* sp. nov., isolated from waste wash-water from processing of Spanish-style green olives. Int J Syst Evol Microbiol 2007;57:906–10.
[32] Chen YG, Cui XL, Fritze D, Chai LH, Schumann P, Wen ML, et al. *Virgibacillus kekensis* sp. nov., a moderately halophilic bacterium isolated from a salt lake in China. Int J Syst Evol Microbiol 2008;58:647–53.
[33] Tanasupawat S, Chamroensaksri N, Kudo T, Itoh T. Identification of moderately halophilic bacteria from Thai fermented fish and proposal of *Virgibacillus siamensis* sp. nov. J Gen Appl Microbiol 2010;56:369–79.
[34] Seng P, Abat C, Rolain M, Colson P, Lagier J, Gouriet F. Laboratory: impact of matrix-assisted laser desorption ionization–time of flight mass spectrometry. J Clin Microbiol 2013;51:2182–94.
[35] Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980;16:111–20.
[36] 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.
[37] Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res 2012;40:D48–53.
[38] 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.
[39] 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.
[40] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340:783–95.
[41] Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001;305:567–80.
[42] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–5.
[43] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 2009;25:119–20.
[44] Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004;14:1394–403.
[45] Lechner M, Findeib S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 2011;12:124.
[46] Auch AF, Von Jan M, Klenk HP, Goker M. DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2010;2:117–34.
[47] 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.

Noncontiguous finished genome sequence and description of *Planococcus massiliensis* sp. nov., a moderately halophilic bacterium isolated from the human gut

E. H. Seck[1], S. A. Sankar[1], S. Khelaifia[1], O. Croce[1], C. Robert[1], C. Couderc[1], F. Di Pinto[1], C. Sokhna[2,3], P.-E. Fournier[1], D. Raoult[1,4] and J.-C. Lagier[1]

1) *Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, UM 63, CNRS 7278, IRD 198, Inserm 1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de Médecine,* 2) *Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes IRD 198, CNRS 7278, Aix-Marseille Université, Marseille, France,* 3) *Campus Commun UCAD-IRD of Hann, Dakar, Senegal and* 4) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

We propose the main phenotypic characteristics and the complete genome sequence and annotation of *Planococcus massiliensis* strain ES2^T (= CSUR P1103 = DSM 28915), the type strain of *P. massiliensis* sp. nov., isolated from a faeces sample collected from a healthy Senegalese man. It is an aerobic, Gram-positive, moderately halophilic, motile and rod-shaped bacterium. The 3 357 017 bp long genome exhibits a G+C content of 46.0% and contains 3357 protein-coding genes and 48 RNA genes.

Keywords: Culturomics, genome, human gut microbiota, *Planococcus massiliensis*, taxonogenomics
Original Submission: 25 October 2015; **Revised Submission:** 18 December 2015; **Accepted:** 18 December 2015
Article published online: 31 December 2015

Corresponding author: J.-C. Lagier, URMITE, UMR CNRS 7278, IRD 198, INSERM U1095, Faculté de Médecine, Aix-Marseille Université, 27 Boulevard Jean Moulin, 13385 Marseille Cedex 5, France
E-mail: jean-christophe.lagier@ap-hm.fr

Introduction

Planococcus massiliensis strain ES2^T (= CSUR P1103 = DSM 28915) is the type strain of *P. massiliensis* sp. nov. This bacterium was isolated from a stool sample of a healthy Senegalese man. This isolation is part of a wider culturomics study, with an aim to maximize the culture conditions to explore the human microbiota in depth [1]. For this project, several hundred samples from healthy individuals, antibiotic-treated individuals, or people with, for example, obesity, anorexia nervosa, or malnutrition were analysed by culturomics in order to extend our knowledge of gut microbes [2]. In this case, we created media containing a high salt concentration in order to cultivate halophilic microorganisms [2]. Furthermore, the availability of genomic data for many bacterial species [3] inspired us to propose a new concept for the description of new species of bacteria by adding proteomic information obtained by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) [4] and genomic analyses [5]. This concept changes the current methods of defining a new bacterial species, which are based on the genetic, phenotypic and chemotaxonomic criteria that are not reproducible and cannot be applied to all the bacterial genus [6–8].

Here we present a summary classification and a set of features for the type strain *Planococcus massiliensis* sp. nov. strain ES2^T, together with the description of the complete genomic sequence and its annotation. These characteristics support the circumscription of the species *Planococcus massiliensis.* To our knowledge, *Planococcus massiliensis* is the first representative member of the genus of *Planococcus* isolated from a human subject. To date, 17 recognized species are representatives of *Planococcus* (*P. alkanoclasticus, P. antarcticus, P. citreus, P. columbae, P. donghaensis, P. halocryophilus, P. kocurii, P. maitriensis, P.*

http://dx.doi.org/10.1016/j.nmni.2015.12.006

maritimus, P. mcmeekini, P. okeanokoites, P. plakortidis, P. psychrophylus, P. rifietoensis, P. salinarum, P. stackebrandtii) (http://www.bacterio.net). All these species are Gram-positive, aerobic cocci that are able to grow at moderately low temperatures and high salt concentrations, and have been predominantly isolated from saline environments [9,10].

Material and Methods

Sample collection and culture conditions

Signed informed consent was obtained from each person included in the study. The study and the assent procedure were approved by the National Committee of Senegal and the ethics committee of Federative Research Institute 48 (Faculty of Medicine, Marseille, France) under agreement 09-022. The sample was obtained from a native Senegalese man living in Ndiop, a rural village in the Guinean–Sudanese zone in Senegal. The specimen was collected in sterile plastic containers, formed into aliquots and stored at −80°C until use. For each sample, pH and salinity were systematically determined with a pH meter (ThermoFisher Scientific, Saint Aubin, France) and a digital refractometer (ThermoFisher Scientific) before any analysis was performed. Then it was cultured in a liquid Colombia broth culture medium (Sigma-Aldrich, Saint-Quentin Fallavier, France) modified by adding (per litre): 5 g $MgCl_2$ $6H_2O$; 5 g $MgSO_4$ $7H_2O$; 2 g KCl; 1 g $CaCl_2$ $2H_2O$; 0.5 g NaBr; 0.5 g $NaHCO_3$; 2 g glucose; and 100 g NaCl. pH was adjusted to 7.5 with 10 M NaOH.

Strain identification by MALDI-TOF and sequencing of 16S rRNA

MALDI-TOF analysis of proteins was used for the identification of bacteria. Each colony was deposited in duplicate on a MALDI-TOF MSP96 target and covered with 1.5 µL of a matrix solution (saturated solution of α-cyano-4-hydroxycinnamic acid in 50% acetonitrile, 2.5% trifluoroacetic acid) to allow crystallisation of molecules. MALDI-TOF was performed using the LT Microflex spectrometer (Bruker Daltonics, Leipzig, Germany). All spectra were recorded in positive linear mode for the mass range from 2000 to 20 000 Da (parameters: ion source 1 (ISI), 20 kV; IS2, 18.5 kV lens, 7 kV). The generated spectra were then compared to the Bruker database, which was supplemented with the new species found through the culturomics project. The resulting score enabled the identification or not of tested species: a score of $\geq$2 with a validly published species enables identification at the species level, a score of $\geq$1.7 but <2 enables identification at the genus level and a score of <1.7 does not enable any identification.

After three assays, unidentified colonies were sequenced using 16S rRNA for formal identification. Isolated colonies were suspended in 200 µL distilled water for DNA extraction using an EZ1 DNA Tissue Kit (Qiagen, Venlo, Netherlands). The amplification of the 16S rRNA was performed by a standard PCR using the universal primer pair FD1 5′-AGAGTTTGATCCTGGCTCAG-3′ and RP2 5′-ACGGCTACCTTGTTACGACTT-3′ [11]. The PCR product was purified and sequenced using the Big Dye Terminator Sequencing kit v.1.1 (Perkin-Elmer, Courtaboeuf, France) with the following internal primers: 536F, 536R, 800F, 800R, 1050F and 1050R; 16S rRNA amplification and sequencing were carried out as previously described by Steven *et al.* [12]. The 16S rRNA nucleotide sequence was corrected using Chromas Pro 1.34 software (Technelysium, Tewantin, Australia), and the BLASTn searches were performed by National Center for Biotechnology Information (http://blast.ncbi.nlm.nih.gov.gate1.inist.fr/Blast.cgi). MEGA6 (Molecular Evolutionary Genetics Analysis) software [13] allowed us to construct a phylogenetic tree. Sequences alignment of the different species was performed by CLUSTALW, and calculation of the evolutionary distance was done with the Kimura two-parameter model [14,15].

Atmospheric, sporulation and microscopy tests

Growth of the strain $ES2^T$ was tested under aerobic atmosphere, in the presence of 5% CO_2, and also in anaerobic and microaerophilic atmospheres created using AnaeroGen and CampyGen respectively (ThermoFisher Scientific).

Spore formation was determined by thermal shock and observed under a microscope.

Gram staining and motility were observed by the use of the DM1000 photonic microscope (Leica Microsystems, Nanterre, France). Cell morphology was examined with a Tecnai G20 (FEI, Limeil-Brevannes, France) transmission electron microscope.

Biochemistry and antimicrobial susceptibility

Biochemical tests were realized by using the commercially available API ZYM, API 50CH and API 20 NE strips (bioMérieux, Marcy l'Étoile, France). Oxidase and catalase reactions were determined by using a BD BBL DrySlide (Becton Dickinson, Franklin Lakes, NJ, USA) according to the manufacturer's instructions.

Sensitivity to antibiotics was determined by using Sircan Discs (i2a, Montpellier, France) on Mueller-Hinton agar in a petri dish (bioMérieux). Doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole, imipenem and metronidazole activity were tested on our strain.

Genomic DNA preparation

After 48 hours of culture, the bacteria were resuspended in sterile water and centrifuged at 4°C at 2000 × *g* for 20 min. Cell pellets were resuspended in 1 mL Tris/EDTA/NaCl solution (10 mM Tris/HCl (pH7.0), 10 mM EDTA (pH8.0) and 300 mM NaCl) and recentrifuged under the same conditions. The pellets were then resuspended in 200 µL of Tris-EDTA (TE) buffer and proteinase K and kept overnight at 37°C for cell lysis. DNA was purified with phenol/chloroform/isoamylalcohol (25:24:1), followed by an overnight precipitation with ethanol at −20°C. It was then resuspended in 205 µL TE buffer and quantified (155 ng/µL) by a Qubit fluorometer using the high-sensitivity kit (ThermoFisher Scientific).

Genome sequencing and assembly

Genomic DNA of *Planococcus massiliensis* was sequenced on the MiSeq Technology (Illumina, San Diego, CA, USA) with the mate pair strategy. The gDNA was barcoded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). The mate pair library was prepared with 1 µg of gDNA using the Nextera mate pair Illumina guide. The gDNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 lab chip. The DNA fragments ranged in size from 1 to 11 kb, with an optimal size at 4.008 kb. No size selection was performed, and 388.3 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments, with an optimum of 634 bp, on the Covaris device S2 in microtubes (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 35.59 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 the sequencing run were performed in a single 39-hour run at a 2 × 251 bp read length. Total information of 10.6 Gb was obtained from a 1326K/mm^2 cluster density with a cluster passing quality control filters of 99.1% (24 492 260 clusters). Within this run, the index representation for *Planococcus massiliensis* was determined to be 7.06%. The 1 481 197 paired reads were filtered according to the read qualities. These reads were trimmed, then assembled using the CLC genomics WB4 software.

Genome annotation and comparison

Open reading frames (ORFs) were predicted using Prodigal [16] with default parameters, but the predicted ORFs were

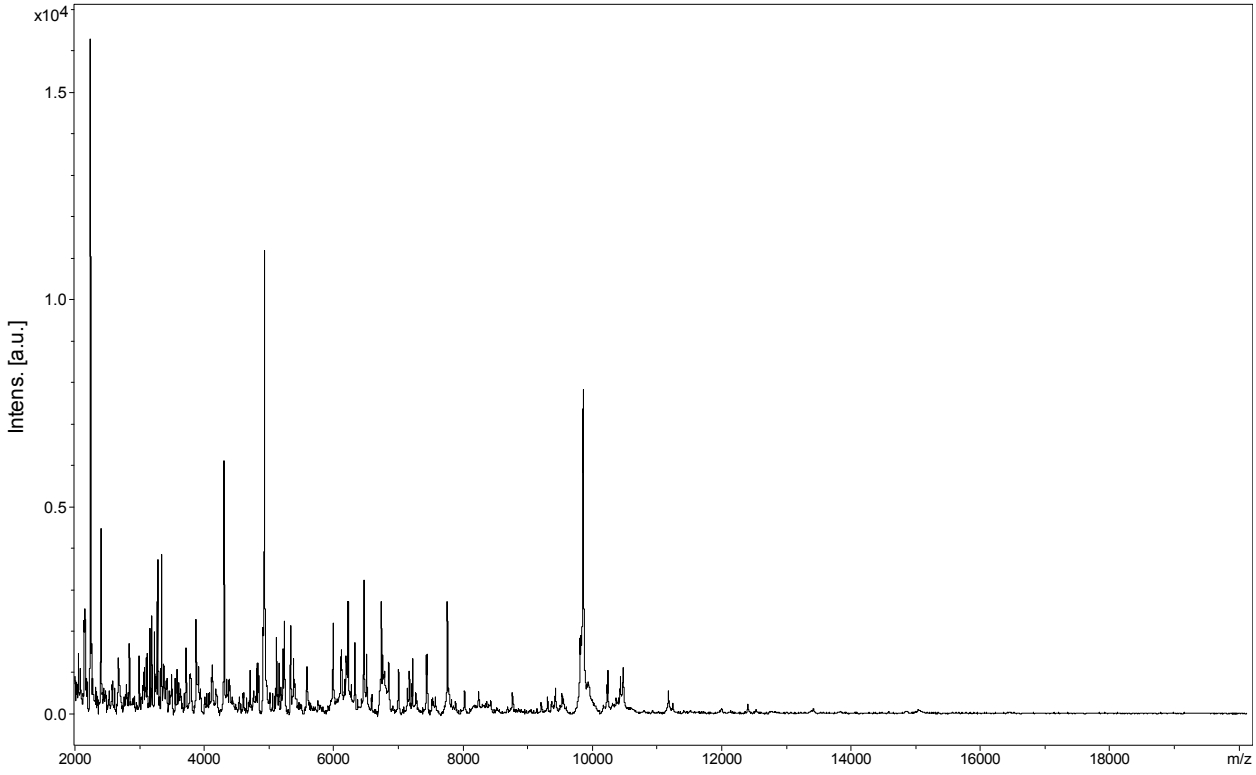


FIG. 1. Reference mass spectrum from Planococcus massiliensis strain ES2^T.

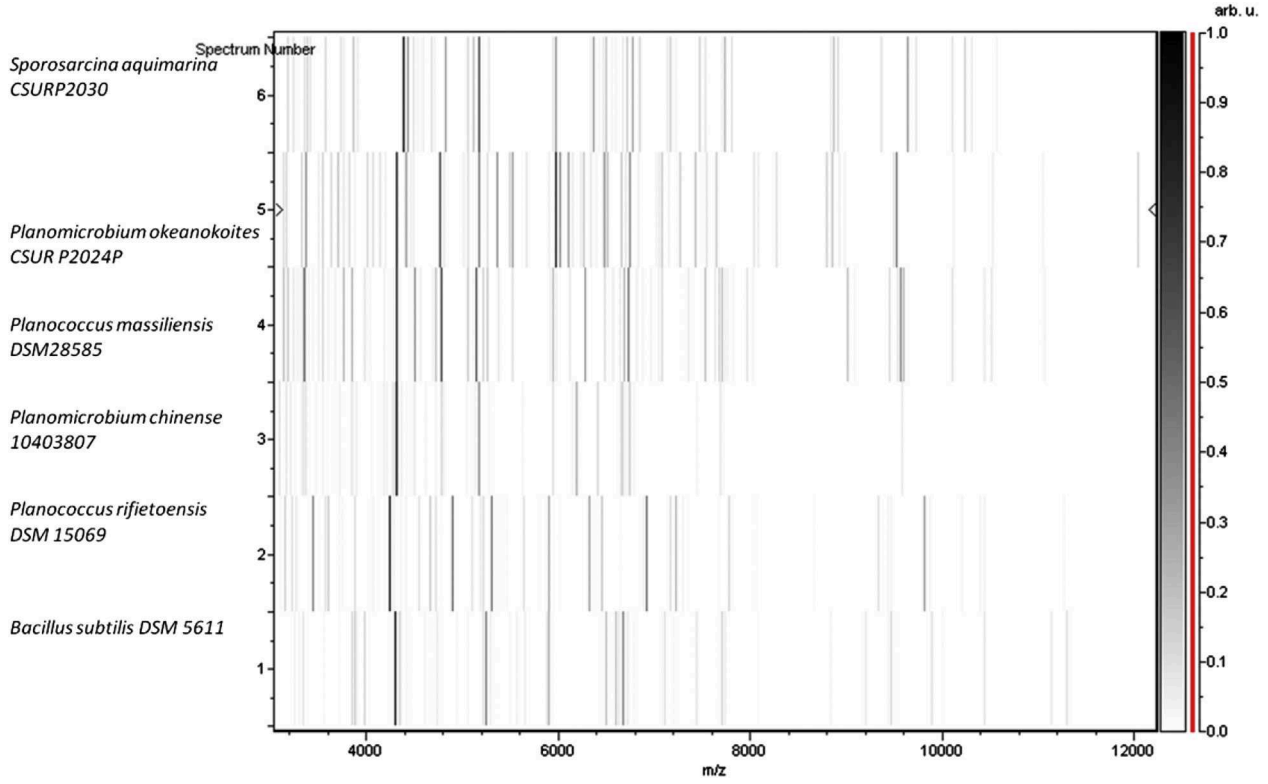


FIG. 2. Gel view comparing Planococcus massiliensis strain ES2^T to other species within genera Planomicrobium, Planococcus and Bacillus. Gel view displays raw spectra of loaded spectrum files arranged in pseudo-gel-like look. x-axis records m/z value. Left y-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity is expressed by greyscale scheme code. Color bar and right y-axis indicate relation between color peak; peak intensity in arbitrary units. Displayed species are indicated on left.

excluded if they spanned a sequencing gap region. The predicted bacterial protein sequences were searched against the GenBank database [17] and the Clusters of Orthologous Groups (COGs) database using BLASTP. The tRNAScanSE tool [18] was used to find tRNA genes, whereas ribosomal RNAs were found using RNAmmer [19] and BLASTn against the GenBank database. Lipoprotein signal peptides and the number of transmembrane helices were predicted using SignalP [20] and TMHMM [21] respectively. ORFans were identified if their BLASTP *E* value was lower than 1e-03 for alignment length greater than 80 amino acids. If alignment lengths were smaller than 80 amino acids, we used an *E* value of 1e-05. Such parameter thresholds have already been used in previous works to define ORFans. Artemis [22] was used for data management and DNA Plotter [23] for visualization of genomic features. The Mauve 2.3.1 alignment tool was used for multiple genomic sequence alignment [24]. To estimate the mean level of nucleotide sequence similarity at the genome level, we used the MAGI homemade software to calculate the average genomic identity of gene sequences (AGIOS) among compared genomes. Briefly, this software combines the Proteinortho software [25] for detecting orthologous proteins in pairwise genomic comparisons, then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman-Wunsch global alignment algorithm. Genomes from the genus *Planococcus* and closely related genera were used for the calculation of AGIOS values.

The genome of *Planococcus massiliensis* strain ES2^T (GenBank accession no. CCXS00000000) was compared to the one of *Planomicrobium glaciei* strain CHR43 (NTS) (GenBank accession no. AUYR00000000), *Planococcus halocryophilus* strain Or1 (GenBank accession no. ANBV01000000), *Planococcus donghaensis* strain MPA1U2 (NTS) (GenBank accession no. AEPB01000000) and *Bacillus subtilis* subsp. *spizizenii* strain TU-B-10 (GenBank accession no. CP002905).

To estimate the overall similarity between the genomes, the Genome-to-Genome Distance Calculator (GGDC) was used [26,27]. The system calculates the distances by comparing the genomes to obtain high-scoring segment pairs (HSP) and interfering distances from a set of formulas: 1, HSP length/total length; 2, identities/HSP length; and 3, identities/total length.

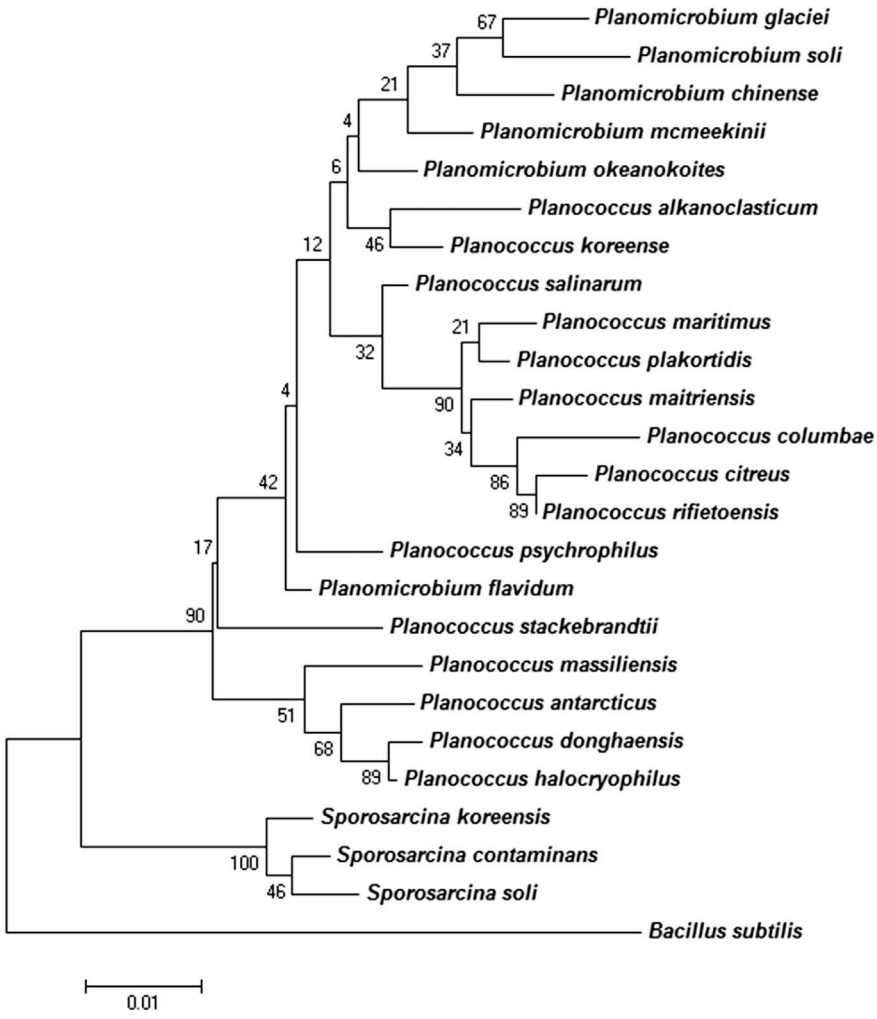


FIG. 3. Phylogenetic tree highlighting position of Planococcus massiliensis sp. nov. strain ES2^T (1536 bp) relative to other type strains within genus. Planomicrobium glaciei strain 0423 (EU036220), Planomicrobium soli strain XN13 (JQ772482), Planomicrobium chinense strain DX3-12 (AJ697862), Planomicrobium mcmeekinii strain S23F2 (AF041791), Planomicrobium okeanokoites strain ATCC 33414 (D55729), Planomicrobium alkanoclasticum strain MAE2 (AF029364), Planomicrobium koreense strain JG07 (AF144750), Planococcus salinarum strain ISL-16 (FJ765415), Planococcus maritimus strain TF-9 (AF500007), Planococcus plakortidis strain MTCC 8491 (JF775504), Planococcus maitriensis strain S1 (AJ544622), Planococcus columbae strain PgEx11 (AJ966515), Planococcus citreus strain ATCC 14404 (X62172), Planococcus rifietoensis strain M8 (AJ493659), Planomicrobium psychrophilus strain CMS 53°r (AJ314746), Planomicrobium flavidum strain ISL-41 (FJ265708), Planococcus stackebrandtii strain K22−03 (AY437845), Planococcus massiliensis (LK021122−1516 bp), Planococcus antarcticus strain CMS 26or (AJ314745), Planococcus donghaensis strain JH1 (EF079063), Planococcus halocryophilus strain Or1 (JF742665), Sporosarcina koreensis strain F73 (DQ073393), Sporosarcina contaminans strain CCUG 53915 (FN298444), Sporosarcina soli strain I80 (DQ073394). GenBank accession numbers are indicated in parentheses. Sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained using maximum likelihood method within MEGA software. Bacillus subtilis subsp. spizizenii strain TU-B-10 (AF074970) was used as outgroup. Scale bar = 0.005% nucleotide sequence divergence.

Results

Strain identification

We did not obtain a significant MALDI-TOF score for strain ES2^T against the Bruker database, suggesting that our isolate was not a known species. Its spectrum was added to our database (Figure 1). The gel view highlighted the spectral differences with other members of the genus *Planococcus* (Figure 2). PCR-based identification of the 16S rRNA of our new isolate (GenBank accession no. LK021122) revealed 1516 bp long sequences. This indicated a 97.95% 16S rRNA sequence similarity with *Planococcus halocryophilus* (GenBank accession no. AJ314745), the phylogenetically closest validated *Planococcus* species (Figure 3). The other closest species were *P. donghaensis* (97.72%), *Planococcus glaciei* (97.06%) and *B. subtilis* (91.92%). The species *P. massiliensis, P. halocryophilus* and *P. donghaensis* shared a single cluster, whereas *P. glaciei* was present in a distant clade in the phylogenetic tree (Figure 3). This value of similarity remains lower than the 98.7% 16S rRNA gene sequence threshold recommended by Stackebrandt and Ebers to delineate a new species without carrying out DNA-DNA hybridization [6]. Thus, this bacterium was considered to be a new species called *Planococcus massiliensis* strain ES2^T sp. nov.

Physiologic and biochemical characteristics

Strain ES2^T is able to grow at temperatures between 25 and 40° C (optimum 37°C) and pH 6–9 (optimum pH 7.0–8.0), and it tolerates NaCl concentrations between 5 to 200 g/L (optimum 75 g/L). We tested the *Planococcus massiliensis* growth on 5% sheep's blood–enriched Columbia agar (bioMérieux, Marcy l'Étoile, France, at 37°C) but we observed weak growth with colonies measuring about 0.3 to 0.6 mm after 48 hours of growth. It is a motile, non-spore-forming and Gram-positive bacterium (Figure 4). Atmospheric testing demonstrated that *Planococcus massiliensis* was strictly aerobic and grew in the presence of 5% CO_2 but did not grow in an anaerobic atmosphere. Colonies that grow on our homemade culture medium were orange, circular, entire, smooth and convex, and they had a diameter of 1.0 to 2.0 mm after 48 hours. Individual cells exhibited a diameter of 0.6 to 0.9 µm and had a slightly curved form with a flagellum under electron microscopy (Figure 5).

Using API galleries, we observed positive reactions for esterase, lipase, trypsin, naphthol-AS-BI-phosphohydrolase, α-glucosidase, D-glucose, D-fructose, D-mannose, D-ribose and D-arabinose. Negative reactions were observed for leucine arylamidase, valine arylamidase, β-galactosidase, alkaline phosphatase, cystine arylamidase, α-chymotrypsin, acid phosphatase, α-galactosidase, β-glucuronidase, β-glucosidase, α-mannosidase, α-fucosidase, arginine dihydrolase, N-acetyl-β-glucosaminidase, nitrate, D-galactose, D-mannitol and urease. The strain was also oxidase positive but catalase negative. Phenotypic characteristics were compared to those of the most closely related species (Table 1).

Antimicrobial susceptibility testing demonstrate that strain ES2^T was susceptible to doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin,

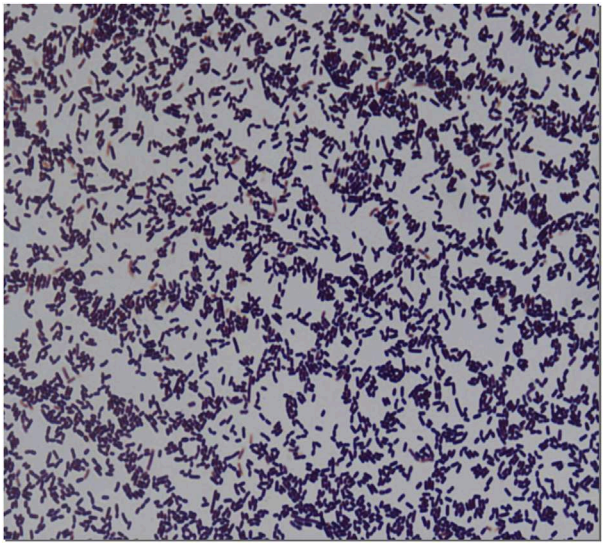


FIG. 4. Gram staining of Planococcus massiliensis strain ES2^T.

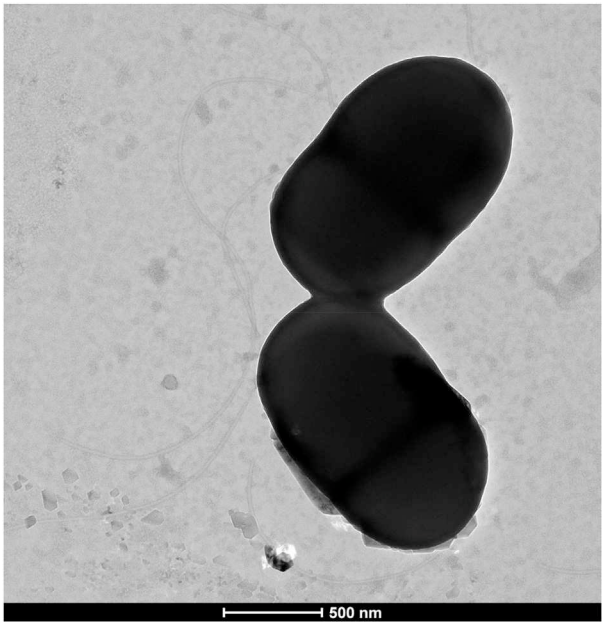


FIG. 5. Transmission electron microscopy of Planococcus massiliensis strain ES2^T. Cells are observed on Tecnai G20 transmission electron microscope operated at 200 keV. Scale bar = 500 nm.

TABLE 1. Differential characteristics of Planococcus massiliensis strain ES2^T compared to other closely related Planococcus species

Property	*P. massiliensis*	*P. okeanokoites*	*P. koreense*	*P. mcmeekinii*	*P. donghaensis*	*P. halocryophilus*	*P. glaciei*	*P. salinarum*	*P. columbae*	*P. alkanoclasticum*	*P. soli*
Cell diameter (μm)	0.6–0.9	0.4–0.8	0.4–0.8	0.6–0.9	0.8–1.2	0.8–1.2	0.4–0.8	0.4–0.8	0.8–1.0	0.4–0.8	0.8–1.0
Oxygen requirement	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic
Gram stain	+	+ to v	+ to v	+	+	+	+	+	+	+ to v	+
Salt requirement	+	+	+	+	+	+	+	–	+	+	+
Motility	+	+	+	+	–	+	+	–	+	+	+
Endospore formation	–	+	+	+	–	–	+	+	–	+	+
Indole	–	–	–	–	–	+	–	NA	–	–	–
Production of:											
Alkaline phosphatase	–	NA	NA	NA	–	NA	+	NA	NA	+	–
Catalase	–	+	+	+	NA	NA	+	NA	NA	+	+
Oxidase	+	w	–	–	+	+	–	+	–	–	–
Nitrate reductase	–	–	–	+	–	–	+	–	–	–	–
Urease	–	–	–	–	NA	–	–	NA	NA	NA	NA
Arginine dihydrolase	–	NA	NA	NA	NA	–	–	NA	NA	NA	NA
B-Galactosidase	–	NA	NA	NA	+	NA	–	–	NA	–	–
N-acetyl-β-glucosaminidase	–	NA	–	NA	+	NA	NA	NA	NA	–	–
Acid from:											
L-Arabinose	–	–	–	–	–	–	–	NA	–	–	NA
D-Ribose	+	+	–	–	+	+	+	NA	NA	–	NA
D-Mannose	+	–	–	–	–	+	NA	–	NA	–	–
D-Mannitol	–	–	–	–	–	+	NA	–	–	–	NA
D-Sucrose	–	–	–	–	+	+	NA	–	+	–	–
D-Glucose	+	–	w	+	+	+	NA	–	–	+	–
D-Fructose	+	+	–	+	–	+	NA	+	+	+	–
D-Maltose	–	–	+	w	+	+	NA	–	NA	–	–
D-Lactose	–	–	+	–	–	+	NA	–	+	–	–
Habitat	Human gut	Fermented seafood	Fermented seafood	Fermented seafood	Sea	Soil	Glacier	Coastal sediment	Coastal sediment	Coastal sediment	Soil

+, positive result; –, negative result; v, variable result; w, weakly positive result; NA, data not available.

TABLE 2. Nucleotide content and gene count levels of genome

Attribute	Value	% of total[a]
Size (bp)	3 357 017	100
G+C content (bp)	1 544 227	46.0
Coding region (bp)	2 972 253	88.53
Total genes	3405	100
RNA genes	48	1.40
Protein-coding genes	3357	98.59
Genes with function prediction	2319	68.10
Genes assigned to COGs	2405	70.63
Genes with peptide signals	188	5.52
Genes with transmembrane helices	776	22.79

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.

ceftriaxone, ciprofloxacin, gentamycin, penicillin, trimethoprim/sulfamethoxazole and imipenem, but it was resistant to metronidazole.

Genome properties

The GenBank Bioproject number is PRJEB6479 and consists of 192 large contigs. The draft genome of *P. massiliensis* ES2^T consists of six scaffolds with 32 contigs and generated a genome size of 3 357 017 bp with a 46.0% G+C content (Table 2, Figure 6). Of the 3405 predicted genes, 3357 are protein-coding genes and 48 are RNAs (eight genes are 5S rRNA, two are 16S rRNA, three are 23S rRNA and 35 are tRNA). A total of 2601 genes (66.90%) were assigned a putative function. A total of 75 genes (1.93%) were identified as ORFans. The remaining genes were annotated as hypothetical proteins. The properties and statistics of the genome are summarized in Table 2.

The distribution of genes into COGs functional categories is presented in Table 3.

Genome comparison

The draft genome of *Planococcus massiliensis* strain ES2^T is smaller than those of *Planomicrobium glaciei, Planococcus halocryophilus* and *Bacillus subtilis* subsp. *spizizenii* (3.35, 3.9, 3.43 and 4.21 Mb respectively) but larger than those of *Planococcus donghaensis* (3.30 Mb). The G+C content of *Planococcus massiliensis* is smaller than those of *P. glaciei* (46.0 and 47.0% respectively) but larger than those of *Planococcus halocryophilus, P. donghaensis* and *Bacillus subtilis* (39.9, 39.7 and 43.8% respectively). The gene content of *P. massiliensis* is smaller than those of *P. glaciei, P. halocryophilus* and *B. subtilis* (3405, 3967, 3429 and 4307 respectively) but larger than that of *P. donghaensis* (3251). The number of rRNA genes varied from four for *P. donghaensis,* 13 for *P. massiliensis,* 30 for *B. subtilis,* 60

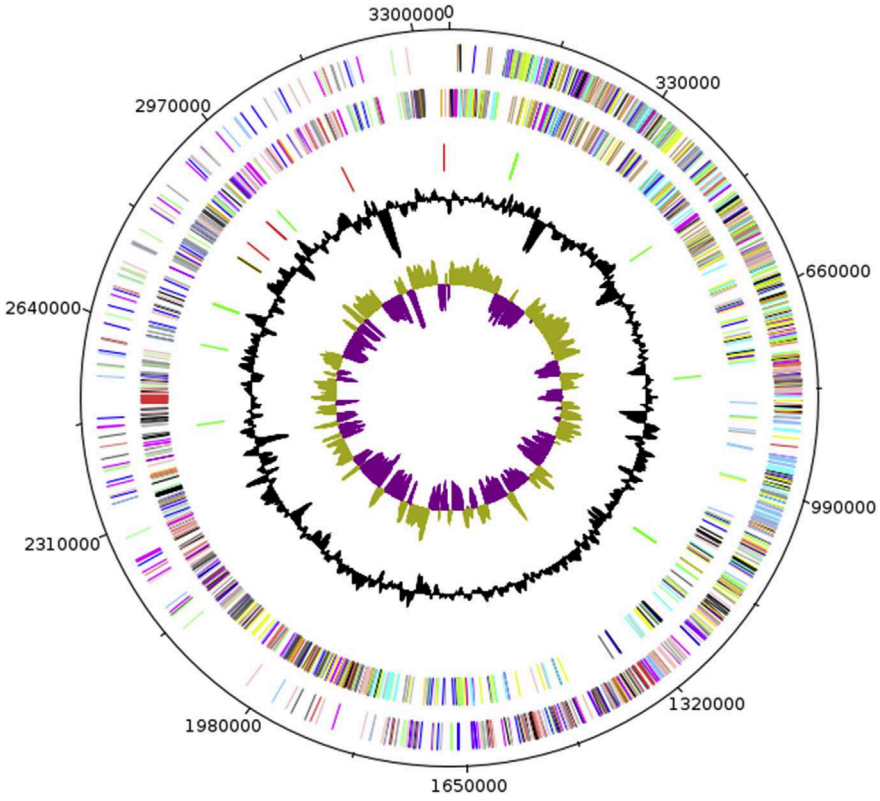


FIG. 6. Graphical circular map of genome of Planococcus massiliensis strain ES2^T. From outside to center: contigs (red/grey), COGs category of genes on forward strand (three circles), genes on forward strand (blue circle), genes on reverse strand (red circle), COGs category on reverse strand (three circles), GC content.

TABLE 3. Number of genes associated with 25 general COGs functional categories

Code	Value	% Value	Description
J	174	5.18	Translation
A	0	0	RNA processing and modification
K	248	7.38	Transcription
L	133	3.96	Replication, recombination and repair
B	1	0.02	Chromatin structure and dynamics
D	34	1.01	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	64	1.90	Defense mechanisms
T	152	4.52	Signal transduction mechanisms
M	132	3.93	Cell wall/membrane biogenesis
N	41	1.22	Cell motility
Z	1	0.02	Cytoskeleton
W	0	0	Extracellular structures
U	48	1.42	Intracellular trafficking and secretion
O	95	2.82	Posttranslational modification, protein turnover, chaperones
X	14	0.41	Phages, Prophages, Transposable elements, Plasmids
C	163	4.85	Energy production and conversion
G	216	6.84	Carbohydrate transport and metabolism
E	317	9.44	Amino acid transport and metabolism
F	80	2.38	Nucleotide transport and metabolism
H	95	2.82	Coenzyme transport and metabolism
I	131	3.90	Lipid transport and metabolism
P	169	5.03	Inorganic ion transport and metabolism
Q	85	2.53	Secondary metabolites biosynthesis, transport and catabolism
R	487	14.50	General function prediction only
S	269	8.01	Function unknown
—	1183	35.23	Not in COGs

COGs, Clusters of Orthologous Groups database.

for *P. halocryophilus* and 62 for *P. glaciei* respectively. A large number of genes assigned to COGs functional categories for amino acid transport and metabolism, transcription, carbohydrate transport and metabolism and translation were identified. Nevertheless, we observed a relative lower number of genes assigned for amino acid transport and metabolism in *P. massiliensis* compared to other species (Figure 7). The genes for RNA processing and modification, nuclear structure and extracellular structures were absent in all the genomes. Finally, the genes coding for COGs category cytoskeleton were present only in *P. massiliensis* and *P. glaciei* (Figure 7). In addition, *P. massiliensis* shared 3880, 2775, 3146 and 4099 orthologous genes with *P. glaciei, P. halocryophilus, P. donghaensis* and *B. subtilis* (Table 4). The average nucleotide sequence identity ranged from 85.84% between *P. donghaensis* and *P. halocryophilus* to 56.69% between *P. halocryophilus* and *B. subtilis* (Table 4). The genomic similarity level between strain ES2^T and closely related *Planomicrobium* and *Planococcus* species was also estimated using the GGDC (Table 5). This comparison of the genomes using

TABLE 4. Numbers of orthologous proteins shared between genomes (upper right) and AGIOS values obtained (lower left)

	PM	PD	PG	PH	BS
PM	3357[a]	2255	2275	2195	1533
PD	70.55	3146[a]	2318	2340	1544
PG	74.92	70.50	3880[a]	2256	1562
PH	67.78	85.84	67.70	2775[a]	1497
BS	59.21	58.76	59.25	56.69	4099[a]

AGIOS, average genomic identity of orthologous gene sequences; BS, *Bacillus subtilis;* PD, *Planococcus donghaensis;* PG, *Planomicrobium glaciei;* PH, *Planococcus halocryophilus;* PM, *Planococcus massiliensis.*
[a]Number of proteins per genome.

TABLE 5. Pairwise comparisons of Planomicrobium species using GGDC, formula 2 (DDH estimates based on identities/HSP length)[a]

	PM	PD	PG	PH	BS
PM	100.00%	19.1% ± 2.76	20.9% ± 2.91	18.9% ± 2.76	27.4% ± 2.54
PD		100.00%	19.2% ± 2.73	39.2% ± 3.34	27.7% ± 2.54
PG			100.00%	19.2% ± 2.73	28.6% ± 2.54
PH				100.00%	29.7% ± 2.54
BS					100.00%

BS, *Bacillus subtilis;* DDH, DNA-DNA hybridization; GGDC, Genome-to-Genome Distance Calculator; PD, *Planococcus donghaensis;* PG, *Planomicrobium glaciei;* PH, *Planococcus halocryophilus;* PM, *Planococcus massiliensis.*
[a]Confidence intervals indicate inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size) [27]. Distance formulas are explained in Auch *et al.* [26]. Formula 2 is recommended, particularly for draft genomes.

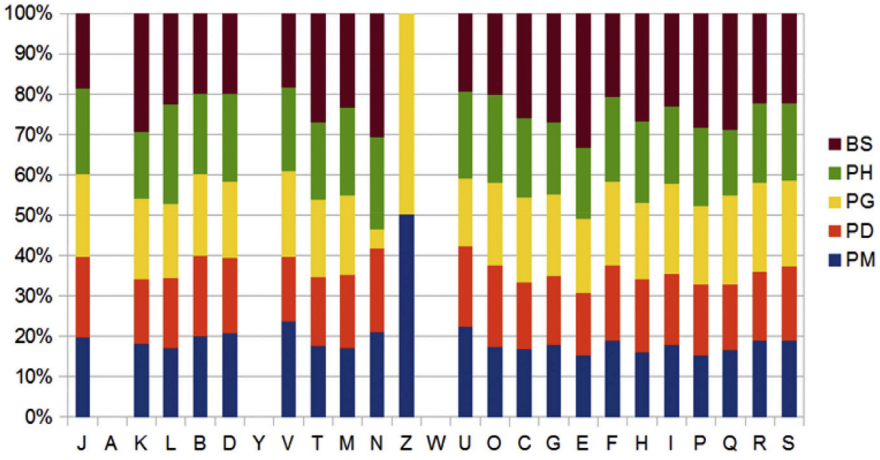


FIG. 7. Distribution of functional classes of predicted genes according to clusters of orthologous groups of proteins. BS, Bacillus subtilis; PD, Planococcus donghaensis; PG, Planomicrobium glaciei; PH, Planococcus halocryophilus; PM, Planococcus massiliensis.

GGDC revealed that *P. massiliensis* shows a slightly higher DNA-DNA hybridization (DDH) estimate with *P. glaciei* compared to those with *P. halocryophilus* and *B. subtilis* (Table 5). For *B. subtilis,* a higher DDH value was estimated with *P. halocryophilus* but did not vary much to the other genomes. These results are in accordance with the 16S rRNA (Figure 1). However, given the confidence intervals (Table 5), the DDH estimates do not show significant differences.

Conclusion

In the context of culturomics studies, several new bacterial species are isolated and then characterized. It is in this context that we studied the phenotypic and phylogenetic characteristics and conducted genomic analyses on strain ES2^T. Results allowed us to formally propose the creation of *Planococcus massiliensis* sp. nov., represented by the strain ES2^T. *P. massiliensis* represents the eighth halophilic bacterium isolated from human stool. Because the colon is not a high-salinity environment, it would be interesting to examine the role of salt or salty products as a potential source of any unusual taxon, such as halophilic.

Taxonomic and nomenclatural proposals

Description of *Planococcus massiliensis* sp. nov.

Planococcus massiliensis (mas.si.li.en'sis, L., masc. adj., *massiliensis* for Massilia, the old Roman name for Marseille, where the strain was isolated).

Strain ES2^T grows at an optimum temperature of 37°C, at pH 7.0–8.0 and at NaCl concentration of 75 g/L. Cells are Gram-positive, strictly aerobic, straight or curved rods (0.6–0.9 µm), motile, and nonendospore forming. Colonies are orange, circular, entire, smooth and convex, 1.0–2.2 mm in diameter.

P. massiliensis shows positive reactions for esterase, lipase, trypsin, naphthol-AS-BI-phosphohydrolase, α-glucosidase, D-glucose, D-fructose, D-mannose, D-ribose and D-arabinose. The strain is also oxidase positive but catalase negative.

Strain ES2^T is susceptible to doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole and imipenem.

P. massiliensis ES2^T (= CSUR P1103, = DSM 28915) was isolated from a stool sample of a healthy Senegalese man. It exhibited a genome size of 3 357 017 bp with a 46.0% G+C content. The 16S rRNA sequence was deposited in GenBank under accession number LK021122, and the whole genome shotgun sequence has been deposited in GenBank under accession number CCXS00000000.

Acknowledgements

The authors thank the Xegen Company (http://www.xegen.fr/) for automating the genomic annotation process and C. Andrieu for administrative assistance. This study was funded by the Fondation Méditerranée Infection.

Conflict of Interest

None declared.

References

[1] Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–93.
[2] Lagier JC, Hugon P, 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;1:237–64.
[3] Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, et al. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. Int J Syst Bacteriol 1987;37:463–4.
[4] Welker M, Moore ERB. Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 2011;34:2–11.
[5] 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 2011;64:384–91.
[6] Stackebrandt E. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[7] Tindall BJ, Rosselló-Móra R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. Int J Syst Evol Microbiol 2000;60:249–66.
[8] Vandamme P, Pot B, Gillis M, de Vos P, Kersters K, Swings J. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Rev 1996;60:407–38.
[9] Nadia C, Mykytczuk S, Wilhelm RC, Whyte LG. *Planococcus halocryophilus* sp. nov., an extreme sub-zero species from high Arctic permafrost. Int J Syst Evol Microbiol 2012;62:1937–44.
[10] Yoon JH, Kang SJ, Lee SY, Oh KH, Oh TK. *Planococcus salinarum* sp. nov., isolated from a marine solar saltern, and emended description of the genus *Planococcus*. Int J Syst Evol Microbiol 2010;60:754–8.
[11] Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 1991;173:697–703.
[12] Steven B, Briggs G, McKay CP, Pollard WH, Greer CW, Whyte LG. Characterization of the microbial diversity in a permafrost sample from the Canadian high Arctic using culture dependent and culture-independent methods. FEMS Microbiol Ecol 2007;59:513–23.

[13] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. c: Molecular Evolutionary Genetics Analysis, version 6.0. Mol Biol Evol 2013;30: 2725–9.
[14] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–80.
[15] Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980;16:111–20.
[16] 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.
[17] Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res 2012;40:D48–53.
[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] 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.
[20] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340:783–95.
[21] Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001;305:567–80.
[22] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–5.
[23] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 2009;25:119–20.
[24] Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004;14:1394–403.
[25] Lechner M, Findeib S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 2011;12:124.
[26] Auch AF, von Jan M, Klenk HP, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2010;2:117–34.
[27] 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.

'*Halomonas massiliensis*' sp. nov., a new halotolerant bacterium isolated from the human gut

E. h. Seck[1], **P.-E. Fournier**[1], **D. Raoult**[1,2] **and S. Khelaifia**[1]

1) *Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, INSERM 1095, Faculté de médecine, Marseille, France and* 2) *Special Infectious Agents Unit, King Fahd Medical Research Centre, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

We report here the main characteristics of '*Halomonas massiliensis*' strain Marseille-P2426^T (CSUR P2426), a new species of the *Halomonas* genus that was isolated from the stool sample of a healthy 24-year-old Senegalese man.

Keywords: Culturomics, genomics, '*Halomonas massiliensis*', halotolerant, taxonomy
Original Submission: 18 July 2016; **Accepted:** 25 July 2016
Article published online: 29 July 2016

Corresponding author: S. Khelaifia, Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, INSERM 1095, Faculté de médecine, 27 Boulevard jean Moulin, 13385 Marseille cedex 05, France
E-mail: khelaifia_saber@yahoo.fr

A new bacterium named strain Marseille-P2426, was isolated from a stool sample of a healthy Senegalese man in May 2016. This isolate is part of a wider culturomics study aiming to cultivate bacterial diversity colonizing the human gut [1]. The study was approved by the Institut Fédératif de Recherche 48 (Faculty of Medicine, Marseille, France), under agreement number 09-022 and the patient gave an informed consent. No identification was obtained for the strain Marseille-P2426 using our systematic matrix-assisted laser desorption–ionization time-of-flight mass spectrometry (MALDI-TOF-MS) screening on a MicroFlex spectrometer (Bruker Daltonics, Bremen, Germany) [2]. The strain Marseille-P2426 was cultured in a homemade liquid medium [3] incubated for 2 days in an aerobic atmosphere at 37°C. Optimal growth for the strain Marseille-P2426 was obtained at 37°C at pH 7; strain Marseille-P2426 is halotolerant and tolerates NaCl concentration up to 20%. Strain Marseille-P2426 is a Gram-negative bacterium, motile non-spore-forming and does not exhibit catalase or oxidase activities. The growing colonies on our homemade culture medium were orange, circular, entire, smooth, convex, with a diameter of 1.0–2.0 mm. Individual cells have a slightly curved form with a flagellum and exhibited 2 µm in length and 0.6 µm in diameter under electron microscopy. We sequenced the complete 16S rRNA gene using universal primers FD1 and RP2 (Eurogentec, Angers, France) as previously described [4], using a 3130-XL sequencer (Applied Biosciences, Saint Aubin, France). Strain Marseille-P2426 exhibited a 16S rRNA gene sequence identity of 97.5% with *Halomonas alkaliantarctica* strain CRSS^T (GenBank Accession number, NR_114902), the phylogenetically closest species with standing in nomenclature (Fig. 1), which classifies it as a member of a new species within the *Halomonas* genus in the family *Halomonadaceae* [5]. So far, the genus *Halomonas* includes more than 80 species with validly published names [6]. *Halomonas* species are described as halotolerant or halophilic, Gram-negative, aerobic and rod-shaped bacteria [7]. Strain Marseille-P2426 exhibited a 16S rRNA sequence divergence >1.3% with *Halomonas alkaliantarctica*, the closest related species with standing in nomenclature [8], which classifies it as a new representative of the *Halomonas* genus isolated from human gut. Based on the information reported here, we

http://dx.doi.org/10.1016/j.nmni.2016.07.013

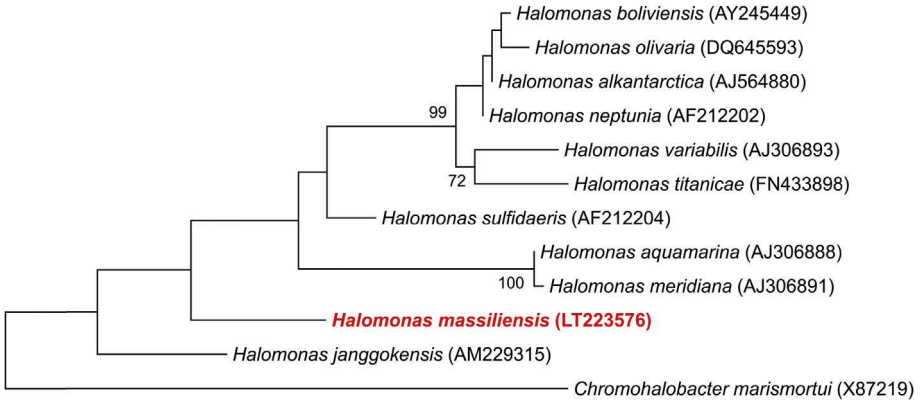


FIG. 1. Phylogenetic tree highlighting the position of '*Halomonas massiliensis*' strain Marseille-P2426^T (in red) relative to other phylogenetically closest members of the *Halomonas* genus. Numbers at the nodes are percentages of bootstrap values obtained by repeating 500 times the analysis to generate a majority consensus tree. Only values > 95% were displayed. The scale bar represents a 2% nucleotide sequence divergence.

propose the creation of '*Halomonas massiliensis*' sp. nov., (mas.si.li.en'sis. L. masc. adj. *massiliensis* of Massilia, the old Roman name for Marseille, where the strain was isolated) and the strain Marseille-P2426^T is the type strain.

MALDI-TOF-MS spectrum accession number

The MALDI-TOF-MS spectrum of "*H. massiliensis*" is available at: http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database.

Nucleotide sequence accession number

The 16S rRNA gene sequence of the strain Marseille-P2426^T was deposited in GenBank under Accession number LT223576.

Deposit in a culture collection

Strain Marseille-P2426^T was deposited in the Collection de Souches de l'Unité des Rickettsies (CSUR, WDCM 875) under number P2426.

Conflict of Interest

None declared.

Acknowledgments

This research was funded by the Mediterranee-Infection Foundation.

References

[1] Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, Raoult D. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 2015;28:237–64.
[2] Seng P, Abat C, Rolain JM, Colson P, Lagier JC, Gouriet F, et al. Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol 2013;51:2182–94.
[3] Khelaifia S, Croce O, Lagier JC, Robert C, Couderc C, Di Pinto F, et al. Noncontiguous finished genome sequence and description of *Virgibacillus massiliensis* sp. nov., a moderately halophilic bacterium isolated from human gut. New Microbes New Infect 2015;8:78–88.
[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] Polia A, Espositoa E, Orlandoc P, Lamaa L, Giordanoa A, Appoloniab F. Nicolausa B,Gambacortaa A, *Halomonas alkaliantarctica* sp. nov., isolated from saline lake Cape Russell in Antarctica, an alkalophilic moderately halophilic, exopolysaccharide-producing bacterium. Syst. Appl. Microbiol 2007;30:31–8.
[6] Euzeby J. Validation list N°143. Int J Syst Evol Microbiol 2012;62:1–4.
[7] Chaohua M, Fangfang J, Yusong W, Wei Z, Min L, Wujun J. *Halomonas huangheensis* sp. nov., a moderately halophilic bacterium isolated from a saline–alkali soil. Int J Syst Evol Microbiol 2014;64:915–20.
[8] 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.

Description of *'Bacillus dakarensis'* sp. nov., *'Bacillus sinesaloumensis'* sp. nov., *'Gracilibacillus timonensis'* sp. nov., *'Halobacillus massiliensis'* sp. nov., *'Lentibacillus massiliensis'* sp. nov., *'Oceanobacillus senegalensis'* sp. nov., *'Oceanobacillus timonensis'* sp. nov., *'Virgibacillus dakarensis'* sp. nov. and *'Virgibacillus marseillensis'* sp. nov., nine halophilic new species isolated from human stool

B. Senghor[1], E. H. Seck[1], S. Khelaifia[1], H. Bassène[2], C. Sokhna[2], P.-E. Fournier[1], D. Raoult[1] and J.-C. Lagier[1]

1) *Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, Inserm 1095, Institut Hospitalo-universitaire Méditerranée-Infection, Faculté de Médecine, Marseille, France and* 2) *Aix-Marseille Université, URMITE, UMR63, CNRS7278, IRD198, Inserm 1095, Campus commun IRD-UCAD de Hann, BP 1386 CP 18524 Dakar, Senegal*

Abstract

We report the main characteristics of *'Bacillus dakarensis'* P3515^T sp. nov., *'Bacillus sinesaloumensis'* P3516^T sp. nov., *'Gracilibacillus timonensis'* P2481^T sp. nov., *'Halobacillus massiliensis'* P3554^T sp. nov., *'Lentibacillus massiliensis'* P3089^T sp. nov., *'Oceanobacillus senegalensis'* P3587^T sp. nov., *'Oceanobacillus timonensis'* P3532^T sp. nov., *'Virgibacillus dakarensis'* P3469^T sp. nov. and *'Virgibacillus marseillensis'* P3610^T sp. nov., that were isolated in 2016 from salty stool samples (≥1.7% NaCl) from healthy Senegalese living at Dielmo and N'diop, two villages in Senegal.

Keywords: Culturomics, halophilic species, human gut microbiota, new species, taxonogenomics
Original Submission: 23 December 2016; **Revised Submission:** 19 January 2017; **Accepted:** 25 January 2017
Article published online: 2 February 2017

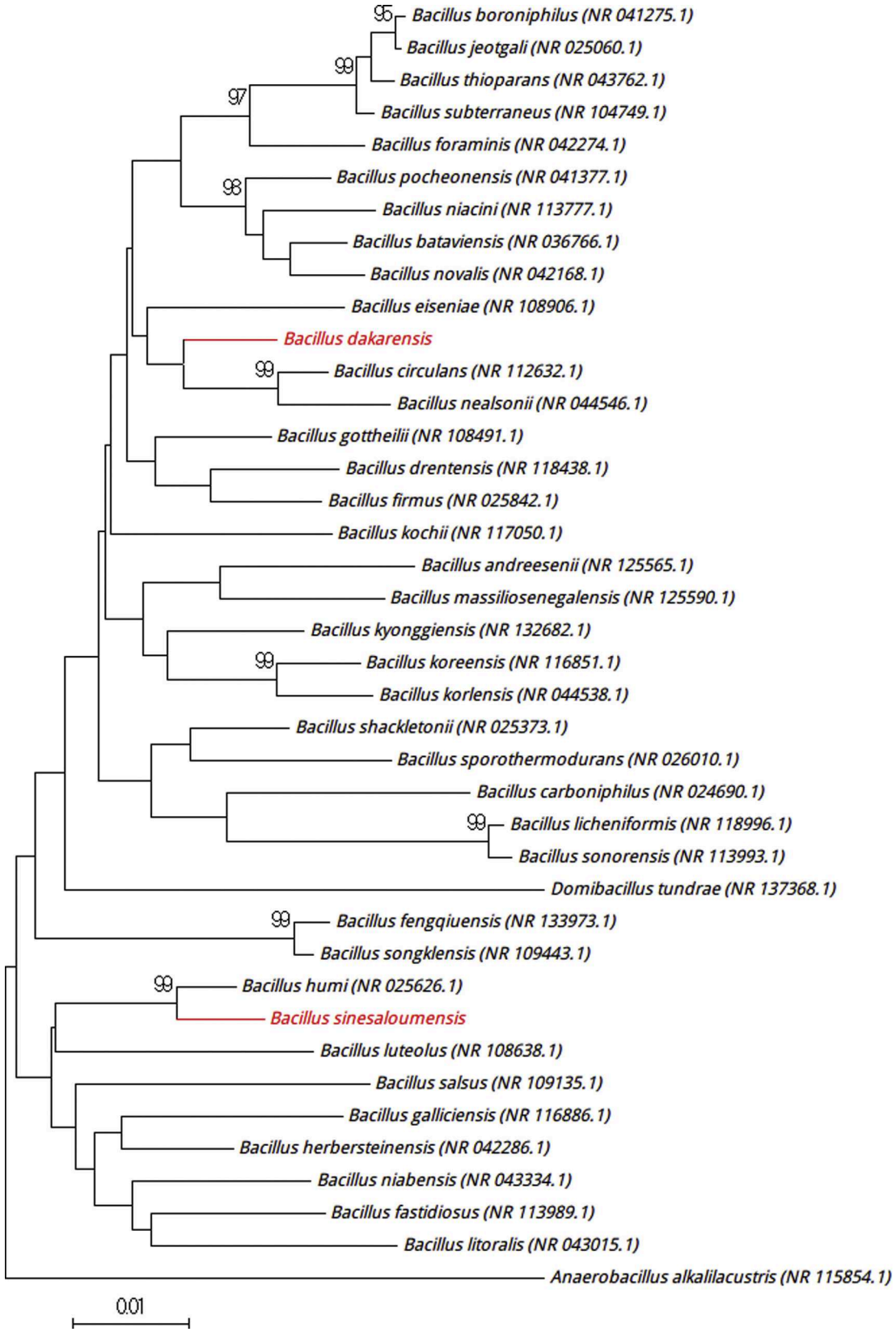
Corresponding author: J.-C. Lagier, Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, Inserm 1095, Institut Hospitalo-universitaire Méditerranée-Infection, Faculté de Médecine, 27 Boulevard Jean Moulin, 13385, Marseille Cedex 05, France
E-mail: jclagier@yahoo.fr
B. Senghor, E.H. Seck contributed equally to this article, and both should be considered first author.

Culturomics has allowed to culture 247 new bacterial species, greatly increasing our understanding of the human gut repertoire thanks to high-throughput culture conditions with a rapid identification method of grown colonies using matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) [1]. As a part of the culturomics approach, we isolated in 2016 from stool samples from healthy patients from Senegal nine bacteria that could not be identified by our MALDI-TOF MS screening on a Microflex spectrometer (Bruker Daltonics, Bremen, Germany) [2]. The study was approved by the ethics committee of the Institut Hospitalo-Universitaire Méditerranée Infection under number 2016-011, and all patients provided written informed consent.

The salinity of the stool specimens was measured by a salinity refractometer (Thermo Scientific, Villebon-sur-Yvette, France). One gram of each stool specimen was diluted in 10 mL of distilled water and centrifuged for 10 minutes at 5000*g*. Then 100 µL of supernatant was deposited in the refractometer, and the result were directly displayed on the screen (in ‰) and then reported (in % NaCl).

Stool samples were inoculated in an aerobic blood culture bottles (Becton Dickinson, Le Pont-de-Claix, France) including an halophilic medium prepared in modifying a Columbia broth medium (Sigma-Aldrich, Saint-Quentin-Fallavier, France) by adding (per litre: 1% (w/v) $MgSO_4$, 0.1% (w/v) $MgCl_2$, 0.4% (w/v) KCl, 0.1% (w/v) $CaCl_2$, 0.05% (w/v) $NaHCO_3$, 0.2% (w/v) of glucose, 0.5% (w/v) of yeast extract (Becton Dickinson), and from 10 to 15% (w/v) NaCl according to the required salinity

http://dx.doi.org/10.1016/j.nmni.2017.01.010



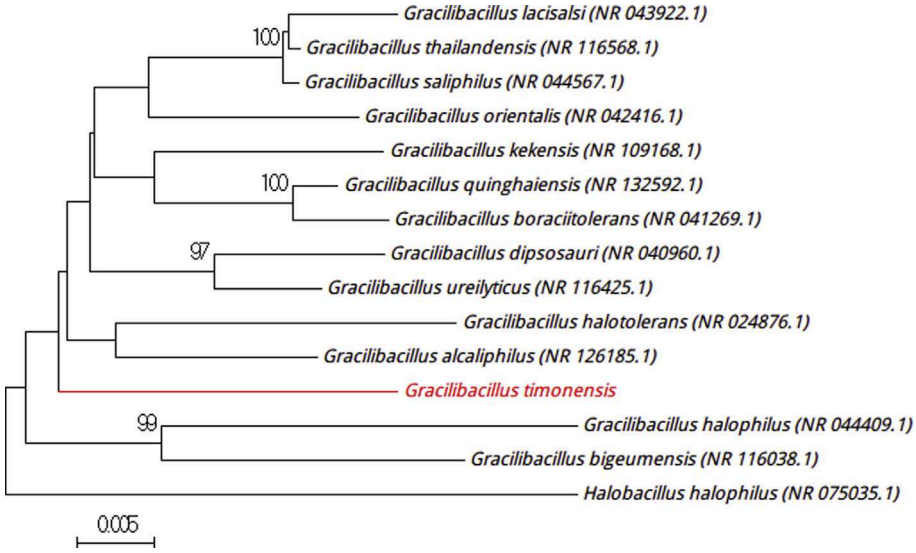


FIG. 2. Phylogenetic tree showing position of *'Gracilibacillus timonensis'* Marseille-P2481^T relative to other phylogenetically close neighbours. Sequences alignment and phylogenetic inferences were realized as explained for Fig. 1. Scale bar indicates 0.5% nucleotide sequence divergence.

with a pH adjusted to 7.5 and incubated for 3 days in aerobic atmosphere at 37°C. The amount of solutes per litre was given in weight volume (w/v) percentage concentration and calculated by the following formula: Concentration in % (w/v) = 100 × [(mass solute in g)/(volume solution in mL)].

The initial agar-grown colonies were obtained after 24 to 48 hours of incubation at 37°C. The 16S rRNA gene was sequenced using fD1-rP2 primers as previously described, using a 3130-XL sequencer (Applied Biosciences, Saint-Aubin, France). Because all strains exhibited a 16S rRNA sequence homology of <98.7% with its phylogenetically closest species, we propose the creation of nine new species according to the nomenclature [3].

Strain Marseille-P3515^T was isolated in a stool sample (4% NaCl) from a 17-year-old adolescent boy living in N'diop. The strain was first isolated in a halophilic medium with 15% (w/v) NaCl. Growth occurred in culture media with NaCl concentration ranging from 0 to 15% (w/v) NaCl (optimum, 7.5% (w/v) NaCl). Agar-grown colonies are beige, circular and shiny with a mean diameter of 1.2 mm. Bacterial cells were Gram positive, rod shaped and polymorphic. Strain Marseille-P3515^T was motile and spore forming with a positive catalase and oxidase reaction. Strain Marseille-P3515^T exhibited a 97.96% sequence identity with *Bacillus circulans* strain NBRC 13626 (GenBank accession no. NR_112632 (Fig. 1). We propose to classify strain Marseille-P3515^T as a new species of the genus *Bacillus* within the family *Bacillaceae* in the phylum *Firmicutes*. Strain Marseille-P3515^T is the type strain of *'Bacillus dakarensis'* (da.ka.ren'sis, N.L. masc. adj. *dakarensis,* the name of the capital of Senegal where the stool sample was collected).

Strain Marseille-P3516^T was isolated in a stool sample (3.8% NaCl) from a 10-year-old girl from N'diop in the same halophilic medium as for strain Marseille-P3515^T, but this strain could grow in media ranging from 0 to 10% (w/v) NaCl (optimum, 5% (w/v) NaCl). Colonies are beige, circular and shiny with a mean diameter of 1.2 mm. Bacterial cells were Gram positive, rod shaped and polymorphic. Strain Marseille-P3516^T was catalase and oxidase positive and exhibited a 98.5% sequence identity with *Bacillus humi* strain LMG 22167 (GenBank accession no. NR_025626.1) (Fig. 1) [4], which allowed us to classify it as a member of the genus *Bacillus* within the family *Bacillaceae* in the phylum *Firmicutes*. Strain Marseille-P3516^T is the type strain of the new species *'Bacillus sinesaloumensis'* (si.ne.sa.lou.men'sis, N.L. masc. adj. *sinesaloumensis,* a vast agropastoral area in Senegal where the village of N'diop, where the stool sample was obtained, is situated).

Strain Marseille-P2481^T is a Gram-positive, motile, spore-forming organism isolated from a stool sample (1.7% NaCl) from a 10-year-old boy from N'diop. Catalase activity is positive and oxidase negative. The first isolation was realized in a halophilic medium with 10% (w/v) NaCl, but the strain could tolerate up to 20% (w/v) NaCl. The growing colonies are orange and circular, with a mean diameter of 2 mm. Individual

FIG. 1. Phylogenetic tree showing position of *'Bacillus dakarensis'* Marseille-P3515^T and *'Bacillus sinesaloumensis'* Marseille-P3516^T relative to other phylogenetically close neighbours. Sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained using maximum-likelihood method within MEGA software. Numbers at nodes are percentages of bootstrap values obtained by repeating analysis 500 times to generate majority consensus tree. Only bootstrap scores of at least 75 were retained. Scale bar indicates 1% nucleotide sequence divergence.

cells exhibit a diameter of 1.9 × 0.5 µm and have a slightly curved form with a flagellum under electron microscopy. Strain P2481^T exhibited a 16S rRNA gene sequence identity of 97% with *Gracilibacillus alcaliphilus* strain SG103 ^{T} (GenBank accession NR_126185) (Fig. 2) [5]. We proposed the creation of a new species within the *Gracilibacillus* genus in the family *Bacillaceae* in the phylum *Firmicutes*. Strain Marseille-P2481^T is the type strain of the new species *'Gracilibacillus timonensis'* (ti.m-o.n.en'sis, N.L. fem. adj. *timonensis,* from the name Hôpital de la Timone, the hospital of Marseille, France).

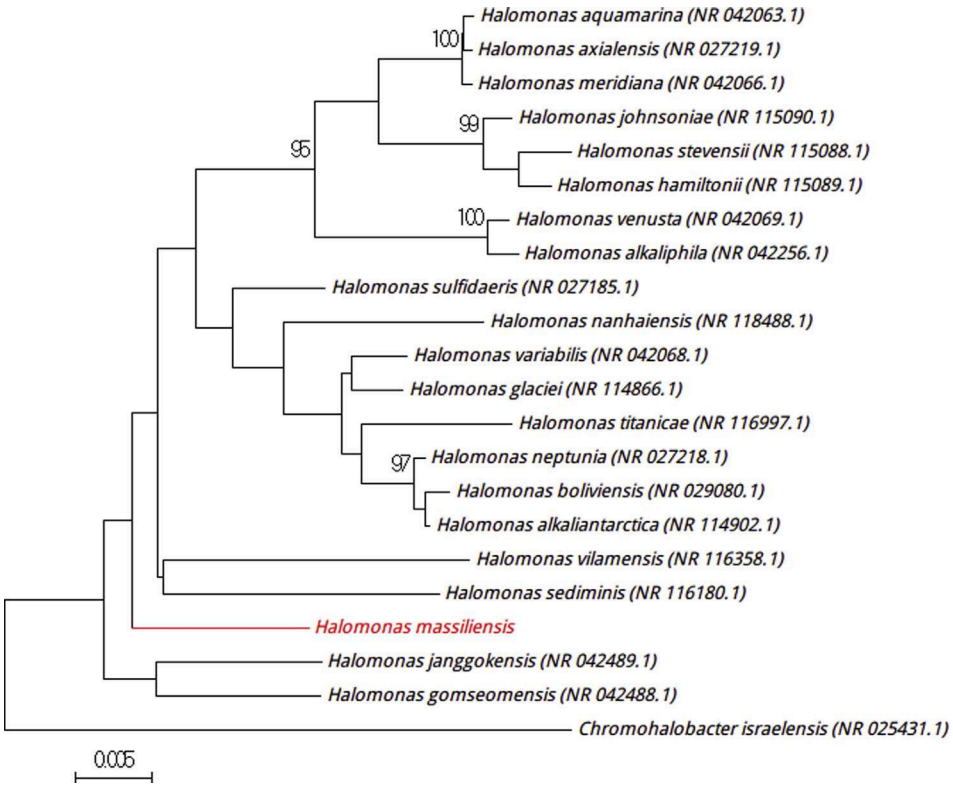


FIG. 3. Phylogenetic tree showing position *'Halobacillus massiliensis'* Marseille-P3554^T relative to other phylogenetically close neighbours. Sequences alignment and phylogenetic inferences were realized as explained for Fig. 1. Scale bar indicates 0.5% nucleotide sequence divergence.

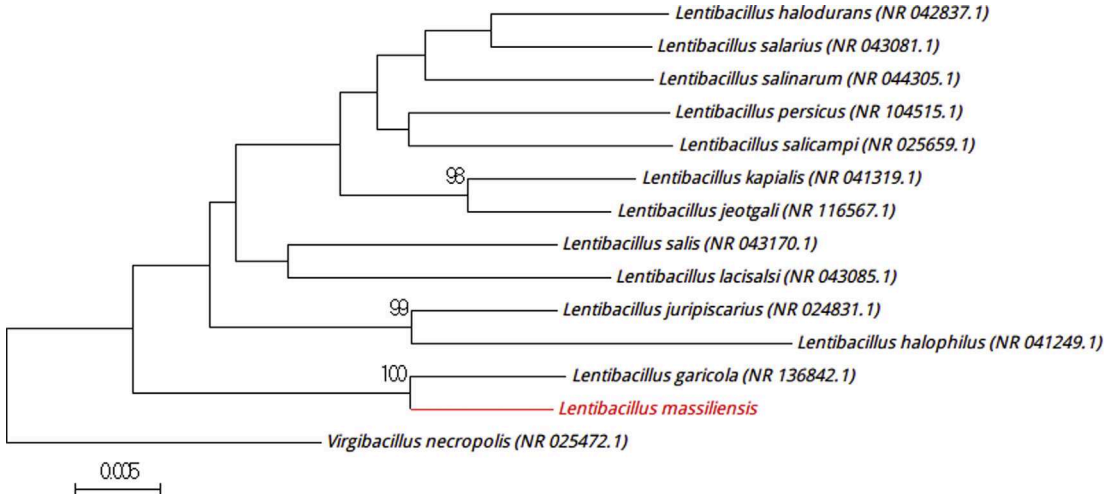


FIG. 4. Phylogenetic tree showing position *'Lentibacillus massiliensis'* Marseille-P3089^T relative to other phylogenetically close neighbours. Sequences alignment and phylogenetic inferences were realized as explained for Fig. 1. Scale bar indicates 0.5% nucleotide sequence divergence.

Strain Marseille-P3554^T was isolated in a stool sample (2.1%) from a 12-year-old boy living in N'diop. The strain Marseille-P3554^T is Gram positive, catalase negative and oxidase positive. The first isolation was realized in a halophilic medium with 10% (w/v) NaCl, but the strain was able to grow with 5 to 20% (w/v) NaCl, with optimum growth at 7.5% (w/v) NaCl. The agar colonies are grey, circular, smooth and convex, with a mean diameter of 2 mm. Individual cells exhibit a diameter of 0.3–0.6 × 2.0–4.0 µm. Strain Marseille-P3554^T exhibited a 16S rRNA sequence divergence of >1.3% with *Halobacillus yeomjeoni* strain MSS-402 (GenBank accession no. NR_043251) (Fig. 3) [6]. On the basis of this finding, we propose to classify *'Halobacillus massiliensis'* as a new representative of the *Halobacillus* genus belonging to the family *Bacillaceae* in the phylum *Firmicutes*. Strain Marseille-P3554^T is the type strain of *'Halobacillus massiliensis'* (mas.si.li.en′sis, L. masc. adj. *massiliensis,* 'of Massilia,' the old Roman name for Marseille, where the strain was isolated).

Strain Marseille-P3089^T is also a Gram-positive, motile, spore-forming, catalase-negative and oxidase-positive bacterium. This strain was isolated from a stool sample (2.1% NaCl) from a 19-year-old man from N'diop. Strain Marseille-P3089^T required salinity ranging from 0.5 to 20% (w/v) NaCl (optimum, 10% (w/v) NaCl). The colonies are yellow and circular, and entire, smooth and convex, with a 1.5 mm diameter. Individual cells exhibit a diameter of 0.3–0.6 × 2.0–4.0 µm. Strain Marseille-P3089^T exhibited a 16S rRNA sequence similarity of 98.3% with *Lentibacillus garicola* (strain SL-MJ1, GenBank accession NR_136842), its closest related species with standing in nomenclature (Fig. 4) [7]. On the basis of this difference, we propose the creation of *'Lentibacillus massiliensis,'* a member of the genus *Lentibacillus,* family *Bacillaceae* and phylum *Firmicutes.* Strain Marseille-P3089^T is the type strain of *'Lentibacillus massiliensis'* (mas.si.li.en′sis, L. masc. adj. *massiliensis,* 'of Massilia,' the old Roman name for Marseille, where the strain was isolated).

Strain Marseille-P3587^T was isolated from a stool sample (2.1% NaCl) from a 7-year-old boy living in N'diop. Growth was first obtained in 15% (w/v) NaCl but occurs at 5 to 20% (w/v) NaCl, with an optimum at 7.5% (w/v) NaCl. Colonies are smooth, circular, yellowish and shiny with a mean diameter of 2.5 mm. Bacterial cells are Gram positive, rod shaped and polymorphic, and the strain Marseille-3587^T is catalase and oxidase positive. The 16S rRNA gene showed 96.6% sequence identity with *Oceanobacillus polygoni* strain SA9 (GenBank

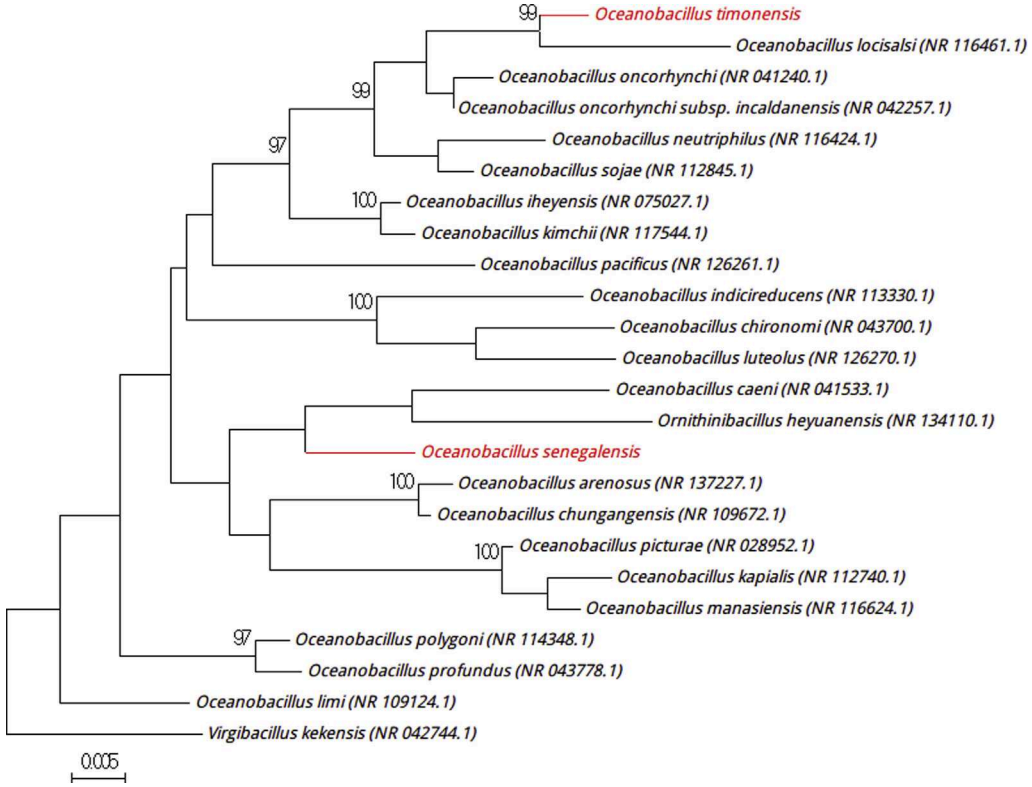


FIG. 5. Phylogenetic tree showing position of *'Oceanobacillus senegalensis'* Marseille-P3587^T and *'Oceanobacillus timonensis'* Marseille-P3532^T relative to other phylogenetically close neighbours. Sequences alignment and phylogenetic inferences were realized as explained for Fig. 1. Scale bar indicates 0.5% nucleotide sequence divergence.

accession no. NR_114348.1), the phylogenetically closest species (Fig. 5) [8]. Considering the sequence similarity of <98.7%, we proposed the creation of a new species, *'Oceanobacillus senegalensis,'* within the phylum *Firmicutes*. Strain Marseille-P3587^T is the type strain of the new species *'Oceanobacillus senegalensis'* (se.ne.ga.len′sis, L. masc. adj. *senegalensis,* 'related to Senegal,' the name of the country where the stool sample was collected).

Strain Marseille-P3532^T was isolated from the same stool sample as for strain Marseille-P3516^T. The first isolation was obtained in a halophilic medium with 15% (w/v) NaCl. However, the strain was able to grow with 0.5 to 20% (w/v) NaCl, with optimum growth at 7.5% (w/v) NaCl. Grown colonies are beige, circular and shiny with a mean diameter of 2 mm. Bacterial cells are Gram positive, rod shaped and polymorphic. Strain Marseille-3532^T is catalase and oxidase positive and exhibited a 98.2% sequence identity with *Oceanobacillus oncorhynchi* strain R-2 (GenBank accession no. NR_041240.1) (Fig. 5) [9], the phylogenetically closest species with standing in nomenclature. Therefore, we proposed to classify the species *'Oceanobacillus timonensis'* as a member of the genus *Oceanobacillus* within the phylum *Firmicutes*. Strain Marseille-P3532^T is the type strain of the new species *'Oceanobacillus timonensis'* (ti.mo.nen′sis, L. fem. adj. *timonensis,* 'related to Timone,' the name of the main university hospital in Marseille, France, where the strain was isolated).

Strain Marseille-P3469^T was isolated from a stool sample (2% NaCl) from a 15-year-old adolescent boy in the village of Dielmo. The first isolation was obtained in a halophilic medium with 10% (w/v) NaCl. However, the strain was able to grow between 0 and 20% (w/v) NaCl, with optimum growth observed at 10% (w/v) NaCl. Agar-grown colonies are cream coloured, circular and shiny, with a mean diameter of 4 mm. Bacterial cells are Gram positive, rod shaped and polymorphic. Strain Marseille-3469^T is catalase and oxidase positive. Strain Marseille 3469^T exhibited a 96.6% sequence identity with *Virgibacillus byunsanensis* strain ISL-24 (GenBank accession no. NR_116615.1) (Fig. 6) [10], the phylogenetically closest species, which led us to classify it as a member of the genus *Virgibacillus* within the family *Bacillaceae* in the phylum *Firmicutes*. Strain Marseille-P3469^T is the type strain of the proposed new species *'Virgibacillus dakarensis'* (da.ka.ren′sis, N.L. fem. adj. *dakarensis,* 'related to Dakar,' the capital of Senegal, where the stool sample was collected).

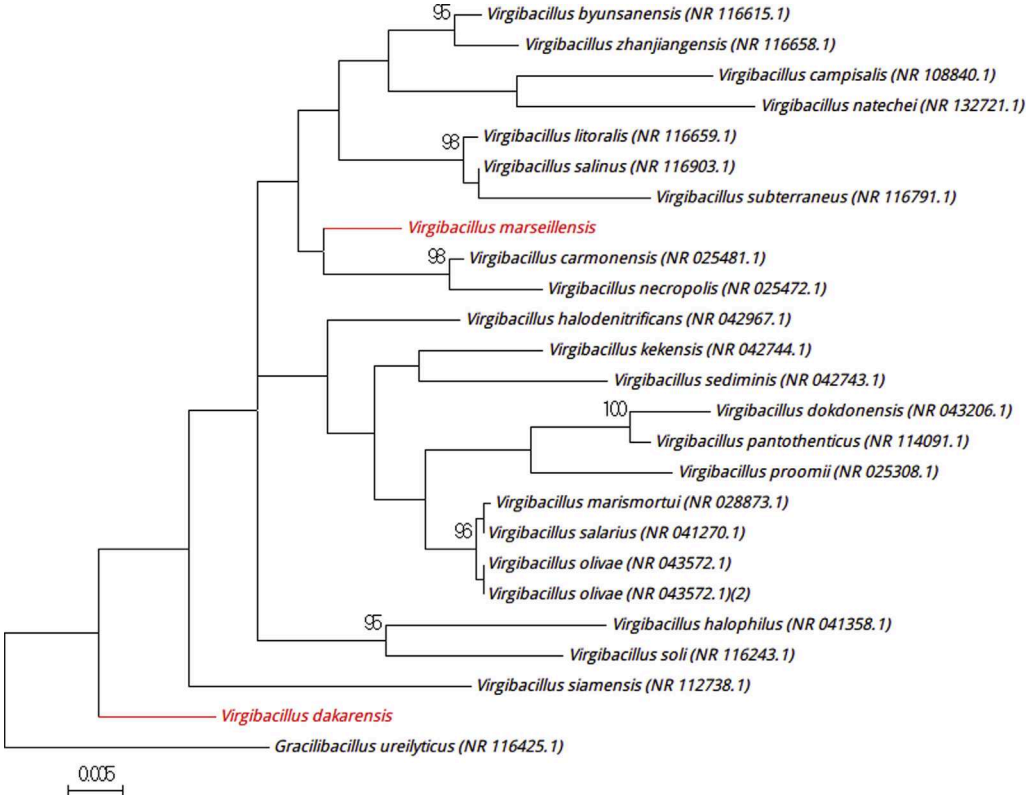


FIG. 6. Phylogenetic tree showing position of *'Virgibacillus dakarensis'* Marseille-P3469^T and *'Virgibacillus marseillensis'* Marseille-P3610T relative to other phylogenetically close neighbours. Sequences alignment and phylogenetic inferences were realized as explained for Fig. 1. Scale bar indicates 0.5% nucleotide sequence divergence.

Strain Marseille-P3610^T was isolated in a stool sample (2.6% NaCl) from a 7-year-old boy living in N'diop. The strain was first isolated in a halophilic medium with 10% (w/v) NaCl but requires 2 to 20% (w/v) NaCl, with optimum growth at 10% (w/v) NaCl. Colonies are cream, circular and shiny with a mean diameter of 4 mm. Bacterial cells are Gram positive, rod shaped and polymorphic. Strain Marseille-3610^T is catalase and oxidase positive. Strain 3610^T exhibited a sequence similarity of 98% with *Virgibacillus carmonensis* strain LMG-20964 (GenBank accession no. NR_025481), the phylogenetically closest species (Fig. 6) [4]. Therefore, we propose the classification of *'Virgibacillus marseillensis'* as a member of the family *Bacillaceae* in the phylum *Firmicutes*. Strain Marseille-P3610^T is the type strain of the new species *'Virgibacillus marseillensis'* (mar.sei.llens's, N.L. fem. adj. *marseillensis,* 'related to Marseille,' the name of the region in France where the laboratory of La Timone hospital, where the strain was isolated, is situated).

MALDI-TOF MS spectra

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

Nucleotide sequence accession numbers

The 16S r RNA gene sequences were deposited in GenBank under the following accession numbers: *'Bacillus dakarensis'* Marseille-P3515^T (LT671589), *'Bacillus sinesaloumensis'* Marseille-P3516^T (LT671591), *'Gracilibacillus timonensis'* Marseille-P2481 (LT223702), *'Halobacillus massiliensis'* Marseille-P3554^T (LT671593), *'Lentibacillus massiliensis'* Marseille-P3089^T (LT598549), *'Oceanobacillus senegalensis'* Marseille-P3587^T (LT714172) *'Oceanobacillus timonensis'* Marseille-P3532^T (LT671597), *'Virgibacillus dakarensis'* Marseille-P3469^T (LT631544) and *'Virgibacillus marseillensis'* Marseille-P3610^T (LT714170).

Deposit in a culture collection

The strains were deposited in the Collection de Souches de l'Unité des Rickettsies (CSUR, WDCM 875) under numbers *'Bacillus dakarensis'* Marseille-P3515^T (P3515), *'Bacillus sinesaloumensis'* Marseille-P3516^T (P3516), *'Gracilibacillus timonensis'* Marseille-P2481^T (P2248), *'Halobacillus massiliensis'* P3554^T (P3354), *'Lentibacillus massiliensis'* Marseille-P3089^T (P3089), *'Oceanobacillus senegalensis'* Marseille-P3587^T (P3587), *'Oceanobacillus timonensis'* Marseille-P3532^T (P3532), *'Virgibacillus dakarensis'* Marseille-P3469^T (P3469) and *'Virgibacillus marseillensis'* Marseille-P3610^T (P3610).

Acknowledgement

This study was funded by the Fondation Méditerranée Infection.

Conflict of Interest

None declared.

References

[1] Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 2016;1:16203.
[2] Lo CI, Fall B, Sambe-Ba B, Diawara S, Gueye MW, Mediannikov O, et al. MALDI-TOF mass spectrometry: a powerful tool for clinical microbiology at Hôpital Principal de Dakar, Senegal (West Africa). PLoS One 2015;10:e0145889.
[3] 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.
[4] Heyrman J. *Virgibacillus carmonensis* sp. nov., *Virgibacillus necropolis* sp. nov. and *Virgibacillus picturae* sp. nov., three novel species isolated from deteriorated mural paintings, transfer of the species of the genus *Salibacillus* to *Virgibacillus*, as *Virgibacillus marismortui* comb. nov. and *Virgibacillus salexigens* comb. nov., and emended description of the genus *Virgibacillus*. Int J Syst Evol Microbiol 2003;53:501–11.
[5] Hirota K, Hanaoka Y, Nodasaka Y, Yumoto I. *Gracilibacillus alcaliphilus* sp. nov., a facultative alkaliphile isolated from indigo fermentation liquor for dyeing. Int J Syst Evol Microbiol 2014;64:3174–80.
[6] Yoon JH, Kang SJ, Lee CH, Oh HW, Oh TK. *Halobacillus yeomjeoni* sp. nov., isolated from a marine solar saltern in Korea. Int J Syst Evol Microbiol 2005;55:2413–7.
[7] Jung WY, Lee SH, Jin HM, Jeon CO. *Lentibacillus garicola* sp. nov., isolated from myeolchi-aekjeot, a Korean fermented anchovy sauce. Antonie Van Leeuwenhoek 2015;107:1569–76.
[8] Hirota K, Hanaoka Y, Nodasaka Y, Yumoto I. *Oceanobacillus polygoni* sp. nov., a facultatively alkaliphile isolated from indigo fermentation fluid. Int J Syst Evol Microbiol 2013;63:3307–12.
[9] Yumoto I. *Oceanobacillus oncorhynchi* sp. nov., a halotolerant obligate alkaliphile isolated from the skin of a rainbow trout (*Oncorhynchus mykiss*), and emended description of the genus *Oceanobacillus*. Int J Syst Evol Microbiol 2005;55:1521–4.
[10] Yoon JH, Kang SJ, Jung YT, Lee KC, Oh HW, Oh TK. *Virgibacillus byunsanensis* sp. nov., isolated from a marine solar saltern. Int J Syst Evol Microbiol 2010;60:291–5.

Genome sequence and description of *Gracilibacillus timonensis* sp. nov. strain Marseille-P2481, a moderate halophilic bacterium first isolated from the human gut microflora: an example of the microbial culturomics revolution

Awa Diop[1], El hadji Seck[1], Gregory Dubourg[1], Nicholas Armstrong, Caroline Michelle[1], Didier Raoult[1,2], Pierre-Edouard Fournier[1]*

[1]Aix-Marseille Université, Institut hospitalo-universitaire Méditerranée-infection, URMITE, UM63, CNRS 7278, IRD 198, Inserm U1095, 19-21 Boulevard Jean Moulin, 13005 Marseille, France Tel: +33 413 732 401, Fax: +33 413 732 402.

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

*Corresponding author: Pr Pierre-Edouard Fournier

Email: pierre-edouard.fournier@univ-amu.fr

Running title: *Gracilibacillus timonensis* sp.nov, a moderate halophilic bacteria from human gut.

ABSTRACT

Microbial culturomics represents an ongoing revolution in the characterization of the human gut microbiota. By using three media containing high salt concentrations (100, 150, and 200 g/L NaCl), we attempted an exhaustive exploration of the halophilic microbial diversity of the human gut, and isolated strain Marseille-P2481 (= CSUR 2481 = DSM 103076), a new moderately halophilic bacterium. This bacterium is a Gram-positive, aerobic, spore-forming rod, that is motile by use of a flagellum and exhibits catalase, but not oxidase, activity. Strain Marseille-P2481 was cultivated in media containing up to 20% (w/v) NaCl, with optimal growth being obtained at 37°C, pH 7.0-8.0 and 7.5% (w/v) NaCl). The major fatty acids were 12-methyl-tetradecanoic acid and Hexadecanoic acid. Its draft genome is 4,548,390 bp-long, composed of 11 scaffolds, with a G+C content of 39.8%. It encodes 4,395 predicted genes (4,332 protein-coding and 63 RNA genes). Strain Marseille-P2481 showed 96.57% 16S rRNA sequence similarity with *Gracilibacillus alcaliphilus* strain SG103^T, the phylogenetically closest species with standing in nomenclature. On the basis of its specific features, strain Marseille-P2481^T was classified as type strain of a new species within the genus *Gracilibacillus* for which the name *Gracilibacillus timonensis* sp. nov. is proposed.

Keywords: *Gracilibacillus timonensis*; halophile; culturomics; taxono-genomics; human gut flora

INTRODUCTION

One of the most important methods of food preservation in history has been the use of salt (NaCl). Salt has also become an indispensable ingredient of any kitchen. Considered previously as hostile to most forms of life by limiting the growth of certain bacteria, it was demonstrated to favor the emergence and growth of others, mainly halophilic bacteria (Cantrell *et al.*, 2011). Several recent studies have reported the isolation of new halophilic species from the human gut microflora (Lagier *et al.*, 2015b; Khelaifia *et al.*, 2016). Therefore, exploring the diversity of halophilic microorganisms in the human gut flora may provide important insights into our understanding of their presence, interactions with the human digestive environment and their influence on health.

In order to explore the human gut halophilic microbiota, and as part of the ongoing Microbial Culturomics study in our laboratory (Lagier *et al.*, 2012, 2016), we used high salt-containing culture media, which enabled us to isolate a new moderately halophilic bacterial strain, Marseille-P2481, that belongs to the genus *Gracilibacillus*. First proposed by Waino *et al.* in 1999 (Wainø *et al.*, 1999), the genus *Gracilibacillus* currently includes 12 species (www.bacterio.net) with validly published names (Hirota *et al.*, 2014). These are Gram-positive, moderately halophilic or halotolerant, mobile bacteria. Most cells form endospores or filaments (Huo *et al.*, 2010a) and were isolated from diverse salty environmental samples including salty lakes, sea water (Jeon *et al.*, 2008; Gao *et al.*, 2012), soil (Chen *et al.*, 2008; Huo *et al.*, 2010b) and or food (Chamroensaksri *et al.*, 2010; Diop *et al.*, 2016).

Using the taxono-genomics approach that includes phenotypic features, proteomic information obtained by matrix-assisted laser-desorption/ionization time-of-flight mass

spectrometry (MALDI-TOF MS) and analysis of the complete genome sequence (Sentausa and Fournier, 2013), we present here the characterization of a new halophilic species for which we propose the name *Gracilibacillus timonensis* sp. nov. Strain Marseille-P2481^{T} (= CSUR 2481 = DSM 103076) is the type strain of *Gracilibacillus timonensis* sp. nov.

MATERIALS AND METHODS

Sample collection and culture conditions

A stool sample was collected from a 10 year-old healthy young Senegalese boy living in N'diop (a rural village in the Guinean-Sudanian zone of Senegal). The patient's parents gave an informed consent and the study was approved by the National Ethics Committee of Senegal, and by the local ethics committee of the IFR48 (Marseille, France) under agreement 09-022.

The salinity of the sample was measured using a digital refractometer (Fisher scientific, Illkirch, France) and its pH measured using a pH-meter (Eutech Instruments, Strasbourg, France).

Strain Marseille-P2481 was isolated in aerobic conditions, on a home-made culture medium consisting of Columbia agar enriched with 42g/L NaCl (Sigma-Aldrich, Saint-Louis, MO, USA), as previously described (Diop *et al.*, 2016).

MALDI-TOF MS strain identification

Twelve individual colonies were deposited in duplicate on a MTP 96 MALDI-TOF target plate for identification with a Microflex MALDI-TOF MS spectrometer (Bruker Daltonics, Leipzig, Germany), as previously reported (Lagier *et al.*, 2015a). The obtained protein spectra were compared with those of XX spectra in the Bruker database enriched with our own database (Lagier *et al.*, 2015a). The strain was identified at the species level if the MALDI-TOF MS score was greater than 1.9. If the score was lower than this threshold, the identification was not considered as reliable and the 16S rRNA gene was sequenced.

16S rRNA gene sequencing identification

The 16S rRNA gene was amplified using the broad-range primer pair FD1 and rp2 (Drancourt *et al.*, 2000). The primers were obtained from Eurogentec (Seraing, France). The obtained amplicon was sequenced using the Big Dye Terminator Sequencing kit and the following

internal primers: 536F, 536R, 800F, 800R, 1050F, 1050R, 357F, 357R, as previously described (Drancourt *et al.*, 2000). The sequence was then compared to the NCBI database using the BLASTn algorithm. If the 16S rRNA gene sequence similarity value was lower than 95% or 98.7% with the most closely related species with standing in nomenclature, as proposed by Stackebrandt and Ebers, the strain was proposed to belong to a new genus or species, respectively (Konstantinidis *et al.*, 2006)

Phylogenetic analysis

The 16S sequences of the type strains of the closest species to our new strain in the BLAST search were downloaded from the NCBI ftp server (ftp://ftp.ncbi.nih.gov/Genome/). Sequences were aligned using the CLUSTALW 2.0 software (Larkin *et al.*, 2007) and phylogenetic inferences were obtained using the neighbor-joining method within the MEGA software, version 6 (Tamura *et al.*, 2013). The bootstraping analysis was performed with 500 replications.

Morphological observation

To observe the cell morphology, transmission electron microscopy of the strain was performed using a Tecnai G20 Cryo (FEI company, Limeil-Brevannes, France) at an operating voltage of 60Kv after negative staining. Gram staining was performed and observed using a photonic microscope Leica DM2500 (Leica Microsystems, Nanterre, France) with a 100X oil-immersion objective. The motility of the strain was assessed by observation of a fresh colony using a DM1000 photonic microscope (Leica Microsystems) at 40x, while sporulation was tested following a thermic shock at 80°C during 20 minutes, and the viability of cells was verified by subculturing them on the same medium before heating.

Atmospheric tests, biochemical and antimicrobial susceptibility

In order to evaluate the optimal culture conditions, strain Marseille-P2481 was cultivated on Chapman agar at different temperatures (25, 28, 37, 45 and 56 °C) under aerobic conditions, and in anaerobic and microaerophilic atmospheres using GENbag Anaer and GENbag microaer systems (bioMérieux), respectively. The pH (pH 5, 6, 6.5, 7, and 8.5) and salinity (0 – 45 g/L NaCl) conditions were also tested.

Biochemical tests were performed using the API ZYM, API 50 CH and API 20 NE strips (BioMerieux, Marcy-l'Etoile, France), according to the manufacturer's instructions. The API ZYM was incubated for 4 hours and the other two strips for 48 hours.

The antibiotic susceptibility of strain Marseille-P2481 was determined using the disk diffusion method as previously described. The following antibiotics were tested: Penicillin G, amoxicillin, ceftriaxone, imipenem, rifampicin, erythromycin, gentamicin and metronidazole. The results were interpreted using the Scan 1200 automate (Interscience, Saint Nom la Bretêche, France).

Fatty acid methyl ester (FAME) analysis by GC/MS

Cellular fatty acid methyl ester (FAME) analysis was performed by Gaz Chromatography / Mass Spectrometry (GC/MS). Two samples were prepared with approximately 70 mg of bacterial biomass per tube harvested from several culture plates. FAMEs were prepared as described by Sasser (Sasser, 1990). GC/MS analyses were carried out as previously described (Dione *et al.*, 2016). Briefly, FAMEs were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500 - SQ 8 S, Perkin Elmer, Courtaboeuf, France). Spectral database search was performed using the MS Search 2.0 software operated with the Standard Reference Database 1A (NIST, Gaithersburg, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK).

Extraction and Genome sequencing

After a pretreatement by lysozyme incubation at 37°C for 2 hours, the DNA of strain Marseille-P2481 was extracted on the EZ1 biorobot (Qiagen) with EZ1 DNA Tissue kit. The elution volume was 50µL. The gDNA was quantified by a Qubit assay with the high sensitivity kit (Life technologies, Carlsbad, CA, USA) to 185 ng/µl.

A MiSeq sequencer and the mate pair strategy (Illumina Inc, San Diego, CA, USA) were used to sequence the gDNA. The gDNA was barcoded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). The mate pair library was prepared with 1.5 µg of gDNA using the Nextera mate pair 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 to 11kb with an optimal size at 5.314 kb. No size selection was performed and 600 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared with an optimal size at 939 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 8.38 nmol/l. The libraries were normalized at 2nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded. Automated cluster generation and sequencing run were performed in a single 39-hours run in a 2x251-bp.

A total sequencing output of 6.52 Gb was obtained from a 696 K/mm^2 cluster density with a cluster passing quality control filters of 95.6 % (12,863,388 passing filter paired reads). Within

this run, the index representation for strain Marseille-P2481 was determined to be 9.39 %. The 1,207,306 paired reads were trimmed and then assembled.

Genome annotation and comparison

Prodigal was used for Open Reading Frame (ORF) prediction (Hyatt *et al.*, 2010) with default parameters. Predicted ORFs spanning a sequencing gap region were excluded. Bacterial protein sequences were 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, a search against the nr database (Benson *et al.*, 2015) was performed using BLASTP with E-value of $1e^{-03}$, a coverage of 0.7 and an identity percent of 30 %. If sequence lengths were smaller than 80 amino acids, we used an E-value of $1e^{-05}$. 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 (Lowe and Eddy, 1997) were used to identify ribosomal RNAs and tRNAs, respectively. We predicted lipoprotein signal peptides and the number of transmembrane helices using Phobius (Käll *et al.*, 2004). ORFans were identified if the BLASTP search was negative (E-value smaller than $1e^{-03}$ for ORFs with a sequence size larger than 80 aas or E-value smaller than $1e^{-05}$ for ORFs with sequence length smaller than 80 aas). Artemis (Carver *et al.*, 2012) and DNA Plotter (Carver *et al.*, 2009) were used for data management and for visualization of genomic features, respectively. Genomes from members of the genus *Gracilibacillus* and closely related genera were used for the calculation of AGIOS values. The genome of strain Marseille-P2481 (EMBL-EBI accession number FLKH00000000) was compared with those of *Amphibacillus jilinensis* strain Y1 (AMWI00000000), *Gracilibacillus halophilus* strain YIM-C55.5 (APML00000000), *G. boraciitolerans* strain JCM 21714 (BAVS00000000), *Bacillus subterraneus* strain MITOT1 (JXIQ00000000), *B. clausii* strain KSM-K16 (AP006627), *B.*

shackletonii strain LMG 18435 (LJJC00000000), *B. alcalophilus* strain ATCC 27647 (ALPT00000000), *Virgibacillus soli* strain PL205 *(*LGPD00000000) and *Halalkalibacillus halophilus* strain DSM 18494 *(*AUHI00000000). Annotation and comparison processes were performed using the multi-agent software system DAGOBAH (Gouret *et al.*, 2011), which include Figenix (Gouret *et al.*, 2005) libraries that provide pipeline analysis. We estimated the degrees of genomic sequence similarity among compared genomes using the following tools. First, we used the MAGI home-made software. This software calculates the average genomic identity of orthologous gene sequences (AGIOS) among compared genomes (Ramasamy *et al.*, 2014). It combines the Proteinortho software (Lechner *et al.*, 2011) for detecting orthologous proteins in pairwise genomic comparisons, then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman-Wunsch global alignment algorithm. Second, we determined the digital DNA-DNA hybridization (dDDH) values among compared genomes using the Genome-to-Genome Distance Calculator (GGDC) web server as previously reported (Klenk *et al.*, 2014). Finally, the Average nucleotide identity by orthology analysis based on the overall similarity between two genome sequences was performed using the OrthoANI algorithm version v0.91 (Ouk Kim *et al.*, 2016).

Accession Numbers

The 16S rRNA and genome sequences were deposited in EMBL-EBI under accession numbers LT223702 and FLKH00000000, respectively.

RESULTS

Characterization of strain Marseille-P2481

Strain identification and phylogenetic analysis

A MALDI-TOF MS score of 1.4 was obtained for strain Marseille-P2481 against our database (Bruker database) suggesting that our isolate was not in the database. The MALDI-TOF MS spectrum from strain Marseille-P2481 (Figure 2) was added to our database and a gel view showed the spectral differences between our isolate and other members of the family *Bacillaceae* was performed (Figure 3). The 16S rDNA-based identification of strain Marseille-P24 *(*EMBL-EBI accession number LT223702) yielded a 96.57% 16S rRNA gene sequence identity with *Gracilibacillus alcaliphilus* strain SG103^{T} (GenBank accession number NR_126185), the phylogenetically closest species with a validly published name (Figure 1). As this value was lower than the 98.65% 16S rRNA sequence identity threshold recommended to define a new species without carrying out DNA-DNA hybridization [35], strain Marseille-P2481 was considered as a potential new species within the *Gracilibacillus* genus.

Physiological and biochemical characteristics

Isolated for the first time in our home-made halophilic medium with 10% (w/v) NaCl, strain Marseille-P2481 able to grow up to 20% (w/v) NaCl under aerobic conditions, but was also able to grow in anaerobic and microaerophilic (at 37°C) atmospheres. Colonies are creamy orange, circular with a mean diameter of 0.2 µm after 2 days of growth at 37°C. Cells are Gram-stain-positive (Figure 4), endospore-forming and motile with a peritrichous flagellum, rods, that are slightly curved, with mean diameter and length of 0.5 and 1.9 µm, respectively (Figure 5). Strain

Marseille-P2481 exhibits positive catalase but no oxidase activity. General features and classification of *Gracilibacillus* timonensis strain Marseille-P2481 are summarized in Table 1.

Using API 50CH, API ZYM, and API 20NE strips, positive results were obtained for esterase, α-glucosidase, 4-nitrophenyl-βD-galactopyranoside, alkaline phosphatase and naphtol-AS-BI-phosphohydrolase activities. Acid was produced from D-glucose, D-fructose, D-galactose and potassium-5-ketogluconate. Nitrate was not reduced and indole was not produced but esculine was hydrolyzed. All the rest was negative on both API strips. Strain Marseille-P2481 differed from other members of the genus *Gracilibacillus* in ß–galactosidase, L-arabinose and D-mannitol metabolism (Table 2). The cellular fatty acids from strain Marseille-P2481 are mainly saturated and the most abundant were 12-methyl-tetradecanoic acid, Hexadecanoic acid and 14 methyl-Hexadecanoic acid (45 %, 16% and 14% respectively). No unsaturated fatty acids were detected (Table 3). Cells are resistant to Penicillin G, amoxicillin, ceftriaxone and metronidazole, but susceptible to imipenem, rifampicin, gentamicin and erythromycin.

Genome properties

The genome is 4,548,390 bp long with 39.8 % GC content. It is composed of 11 scaffolds (composed of 12 contigs). Of the 4,395 predicted genes, 4,330 were protein-coding genes and 65 were RNAs (4 complete rRNA operons, 2 additional 5S rRNA and 51 tRNA genes). A total of 3,043 genes (70.24 %) were assigned a putative function (by COGs or by BLAST against nr). A total of 214 genes were identified as ORFans (4.94%). The remaining genes were annotated as hypothetical proteins (861 genes => 19.92%). The genome statistics are presented in Table 4 and the distribution of genes into COGs functional categories is summarized in Table 5.

Comparative genomics

The draft genome sequence structure of strain Marseille-P2481 is summarized in Figure 6. It is smaller than those of *B. shackletonii* and *V. soli* (4.548, 5.298 and 5.003 Mb, respectively), but larger than those of *A. jilinensis*, *B. subterraneus*, *G. halophilus*, *B. clausii*, *H. halophilus*, *G. boraciitolerans* and *B. alcalophilus* (3.837, 3.9, 3.034, 4.304, 2.708, 3.652 and 4.37 Mb, respectively). The G+C content of strain Marseille-P2481 is smaller than those of *B. subterraneus* and *B. clausii* (39.8, 42.13 and 44.75%, respectively), but larger than those of *A. jilinensis*, *B. shackletonii*, *V. soli*, *G. halophilus*, *H. halophilus*, *G. boraciitolerans* and *B. alcalophilus* (37.27, 36.71, 37.37, 37.92, 37.43, 35.83 and 37.39%, respectively). The gene content of strain Marseille-P2481 is smaller than those of *B. shackletonii* and *G. boraciitolerans* (4,332, 4,727 and 4,450 genes, respectively), but larger than those of *A. jilinensis*, *B. subterraneus*, *V. soli*, *G. halophilus*, *B. clausii*, *H. halophilus* and *B. alcalophilus* (3,594, 3,465, 4,253, 2,968, 4,087, 2,771 and 3,745 genes, respectively). The gene distribution into COG categories was similar among all compared genomes (Figure 7). In addition, AGIOS analysis showed that at least 37.90% of the orthologous proteins of G. timonensis were found among others compared genomes. Strain Marseille-P2481 shared between 1,642, 1,553, 1,496, 1.493, 1,458 and 1,301 orthologous proteins with *V. soli*, *G. halophilus*, *B. shackletonii*, *B. clausii*, *A. jilinensis* and *G. boraciitolerans*, respectively. (Table 6). The average percentage of nucleotide sequence identity ranged from 59.69% to 58.41% at the intra-species level between *G. timonensis* and the two *Gracilibacillus* species, but it ranged from 66.78% to 65.19% at interspecies level between strain Marseille-P2481 and *A. jilinensis and H. halophilus*. This value is lower between *G. timonensis* and other species belonging to the genus *Bacillus*. In addition, digital DNA-DNA hybridization (dDDH) values of strain Marseille-P2481 and the compared closest species varied between 17.4 to 38.9% and were 19.3, 19.9, 20.6, 22.6,

23.6, 25.9, 29.1, 30.9 and 38.9% for *G. halophilus, G. boraciitolerans, H. halophilus, B. shackletonii, B. alcalophilus, A. jilinensis, B. subterraneus, B. clausii* and *V. soli* respectively *(*Table 7). Finally, The ANI values calculated using OrthoANI between strain Marseille-P2481 and the compared genomes ranged from 65.1 to 71.6% similarity (Figure 8). These dDDH and ANI values were less than 70% and 95-96% threshold values for species demarcation respectively (Meier-Kolthoff *et al.*, 2013; Klappenbach *et al.*, 2007; Richter and Rosselló-Móra, 2009).

DISCUSSION

Thanks to the concept of "microbial culturomics", aiming at exploring the diversity of the human microbiota as exhaustively as possible, many new bacterial species have been discovered (Lagier *et al.*, 2016). This concept is based on the diversification of physico-chemical parameters of culture conditions (Lagier *et al.*, 2012, 2015a, 2016) to mimick the entirety of selective constraints that have shaped the natural flora for millions of years, Using hypersaline conditions, many hitherto unknown bacteria extremely and or moderately halophilic have been identified in humans, including strain Marseille-P2481. Its phenotypic, phylogenetic and genomic characteristics suggested that it represents a new species within the genus *Gracilibacillus*. Members of this genus are Gram-positive, aerobic, motile and peritrichous bacteria and are moderately halophilic, and form endospores and white colonies. However, *Gracilibacillus timonensis* sp. nov. is different from other *Gracilibacillus* species in colony color and metabolism of ß–galactosidase, L-arabinose and D-mannitol. In addition, the dDDH, the OrthoANI and AGIOS values of *G. timonensis* compared to other known species confirm strongly its new species status.

CONCLUSION

The moderately halophilic strain Marseille-P2481^T was isolated from a stool sample of a 10-year-old healthy Senegalese boy as part of a study of halophilic bacteria from the human gut. Based on its phenotypic, phylogenetic and genomic characteristics analysis, this strain is proposed to represent a novel species in the genus *Gracilibacillus*, for which the name *Gracilibacillus timonensis* sp. nov. is proposed, Marseille-P2481^T being the type strain.

Description of Gracilibacillus timonensis *sp. nov.*

Gracilibacillus timonensis (*ti.mo.nen'sis*, N. L adj. masc., *timonensis* of Timone, the name of the main hospital of Marseille, France, where the type strain was first isolated).

This bacterium is Gram-stain-positive and moderately halophilic. Cells are motile (with peritrichous flagella), spore-forming rods and are slightly curved with mean diameter and length of 0.5 and 1.9 µm, respectively. Colonies are creamy orange, circular and have a mean diameter of 0.2 µm on our home-made halophilic culture medium (7.5% (w/v) NaCl) after 48h of growth at 37°C. The bacterium is preferentially aerobic but is able to grow in anaerobic and microaerophilic (at 37°C) atmospheres. Strain Marseille-P2481^T is able to grow in media containing up to 20% NaCl but no growth occurs in the absence of NaCl. The optimal culture conditions are 37°C, pH 7.0-8.0 and 7.5% (w/v) NaCl.

Esculine, esterase, catalase, α-glucosidase, 4-nitrophenyl-βD-galactopyranoside, alkaline phosphatase and naphtol-AS-BI-phosphohydrolase activities are present but oxidase, indole and nitrate reductase activities are absent.Cellular fatty acids are mainly saturated structures, with 12-methyl-tetradecanoic acid (45%) and Hexadecanoic acid (16%) being the most abundant. No unsaturated structure was found. The genomic DNA G+C content is 39.8 mol%. The 16S rRNA

and genome sequences are deposited in EMBL-EBI under accession numbers LT223702 and FLKH00000000, respectively. The type strain of *Gracilibacillus timonensis* is strain Marseille-P2481^{T} (= CSUR P2481 = DSM 103076).

ACKNOWLEDGEMENTS

The authors thank the Xegen Company (www.xegen.fr) for automating the genomic annotation process. This study was supported by the “Fondation Méditerranée Infection”.

CONFLICT OF INTEREST

The authors declare that they have no competing interest in relation to this research.

REFERENCES

Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. (2015). GenBank. *Nucleic Acids Res* **43**: D30–D35.

Cantrell SA, Dianese JC, Fell J, Gunde-Cimerman N, Zalar P. (2011). Unusual fungal niches. *Mycologia* **103**: 1161–1174.

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.

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

Chamroensaksri N, Tanasupawat S, Akaracharanya A, Visessanguan W, Kudo T, Itoh T. (2010). Gracilibacillus thailandensis sp. nov., from fermented fish (pla-ra). *Int J Syst Evol Microbiol* **60**: 944–948.

Chen Y-G, Cui X-L, Zhang Y-Q, Li W-J, Wang Y-X, Xu L-H, *et al.* (2008). Gracilibacillus halophilus sp. nov., a moderately halophilic bacterium isolated from saline soil. *Int J Syst Evol Microbiol* **58**: 2403–2408.

Dione N, Sankar SA, Lagier J-C, 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.

Diop A, Khelaifia S, Armstrong N, Labas N, Fournier P-E, Raoult D, *et al.* (2016). Microbial culturomics unravels the halophilic microbiota repertoire of table salt: description of *Gracilibacillus massiliensis* sp. nov. *Microb Ecol Health Dis* **27**.. e-pub ahead of print, doi: 10.3402/mehd.v27.32049.

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. *J Clin Microbiol* **38**: 3623–3630.

Gao M, Liu Z-Z, Zhou Y-G, Liu H-C, Ma Y-C, Wang L, *et al.* (2012). Gracilibacillus kekensis sp. nov., a moderate halophile isolated from Keke Salt Lake. *Int J Syst Evol Microbiol* **62**: 1032–1036.

Gouret P, Paganini J, Dainat J, Louati D, Darbo E, Pontarotti P, et al. Berlin: Springer-Verlag; 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; pp. 71–87.

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 Bioinformatics* **6**: 198.

Hirota K, Hanaoka Y, Nodasaka Y, Yumoto I. (2014). Gracilibacillus alcaliphilus sp. nov., a facultative alkaliphile isolated from indigo fermentation liquor for dyeing. *Int J Syst Evol Microbiol* **64**: 3174–3180.

Huo Y-Y, Xu X-W, Cui H-L, Wu M. (2010a). Gracilibacillus ureilyticus sp. nov., a halotolerant bacterium from a saline-alkaline soil. *Int J Syst Evol Microbiol* **60**: 1383–1386.

Huo Y-Y, Xu X-W, Cui H-L, Wu M. (2010b). Gracilibacillus ureilyticus sp. nov., a halotolerant bacterium from a saline-alkaline soil. *Int J Syst Evol Microbiol* **60**: 1383–1386.

Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.

Jeon CO, Lim J-M, Jang HH, Park D-J, Xu L-H, Jiang C-L, *et al.* (2008). Gracilibacillus lacisalsi sp. nov., a halophilic Gram-positive bacterium from a salt lake in China. *Int J Syst Evol Microbiol* **58**: 2282–2286.

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

Khelaifia S, Lagier J-C, Bibi F, Azhar EI, Croce O, Padmanabhan R, *et al.* (2016). Microbial Culturomics to Map Halophilic Bacterium in Human Gut: Genome Sequence and Description of *Oceanobacillus jeddahense* sp. nov. *OMICS J Integr Biol* **20**: 248–258.

Klappenbach JA, Goris J, Vandamme P, Coenye T, Konstantinidis KT, Tiedje JM. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**: 81–91.

Klenk H-P, Meier-Kolthoff JP, G?ker 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.

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

Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* **35**: 3100–3108.

Lagier J-C, 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 J-C, Hugon P, Khelaifia S, Fournier P-E, La Scola B, Raoult D. (2015a). The Rebirth of Culture in Microbiology through the Example of Culturomics To Study Human Gut Microbiota. *Clin Microbiol Rev* **28**: 237–264.

Lagier J-C, 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.

Lagier J-C, Khelaifia S, Azhar EI, Croce O, Bibi F, Jiman-Fatani AA, *et al.* (2015b). Genome sequence of Oceanobacillus picturae strain S1, an halophilic bacterium first isolated in human gut. *Stand Genomic Sci* **10**.. e-pub ahead of print, doi: 10.1186/s40793-015-0081-2.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947–2948.

Lechner M, Findei\s s S, Steiner L, Marz M, Stadler PF, Prohaska SJ. (2011). Proteinortho: detection of (co-) orthologs in large-scale analysis. *BMC Bioinformatics* **12**: 124.

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.

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

Ouk Kim Y, Chun J, Lee I, Park S-C. (2016). OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* **66**: 1100–1103.

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

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.

Richter M, Rosselló-Móra R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci* **106**: 19126–19131.

Sasser M. (1990). Identification of bacteria by gas chromatography of cellular fatty acids. http://natasha.eng.usf.edu/gilbert/courses/Biotransport%20Phenomena/pdf/bacteria_gc_1.pdf (Accessed June 13, 2017).

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

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

Vandamme P, Pot B, Gillis M, De Vos P, Kersters K, Swings J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* **60**: 407–438.

Wainø M, Tindall BJ, Schumann P, Ingvorsen K. (1999). Gracilibacillus gen. nov., with description of Gracilibacillus halotolerans gen. nov., sp. nov.; transfer of Bacillus dipsosauri to Gracilibacillus dipsosauri comb. nov., and Bacillus salexigens to the genus Salibacillus gen. nov., as Salibacillus salexigens comb. nov. *Int J Syst Evol Microbiol* **49**: 821–831.

Table 1: Classification and general features of *Gracilibacillus timonensis* strain Marseille-P2481^T according to the MIGS recommendations [23]

MIGS ID	Property classification	Term	Evidence code[a]
		Domain: *Bacteria*	TAS [37]
		Phylum: *Firmicutes*	TAS [38]
		Class: *Bacilli*	TAS [37]
		Order: *Bacillales*	TAS [37]
		Family: *Bacillaceae*	TAS [37]
		Genus: *Gracilibacillus*	TAS [5]
		Species: *Gracilibacillus timonensis*	IDA
		Type strain: Marseille-P2481^T	IDA
	Gram stain	Positive	IDA
	Cell shape	Rods	IDA
	Motility	Motile	IDA
	Sporulation	Spore-forming	IDA
	Temperature (°C)	Mesophile (25-45)	IDA

	Optimum temperature	37°C	IDA
	pH range:	6.0-9.0	IDA
	Optimal pH	7.0-8.0	
	Carbon source	Unknown	IDA
MIGS-6	Habitat	Human gut	IDA
MIGS-6.3	NaCl range:	75-200	IDA
	Optimum NaCl	75 g/L	
MIGS-22	Oxygen requirement	Aerobic	IDA
MIGS-15	Biotic relationship	Free living	IDA
MIGS-14	Pathogenicity	Unknown	IDA

[a]Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature). These evidence codes are from the Gene Ontology project [39].

Table 2: Differential characteristics of *Gracilibacillus timonensis* strain Marseille-P2481^T and other closely related members of the genus *Gracilibacillus*

Properties	***G. timonensis***	***G. saliphilus***	***G. bigeumensis***	***G. halophilus***	***G. boraciitolerans***	***G. kekensis***	***G. halotolerans***	***G. alcaliphilus***
Cell diameter (µm)	0.5 – 0.8	0.7 – 0. 9	0.3 – 0.5	0.3 – 0.5	0.5 – 0. 9	0.2 – 1.05	0.4 – 0.6	0.5 – 0.7
Pigmentation	Creamy orange	Creamy white	Creamy	White	Dirty white	Creamy white	Creamy white	Creamy white
Oxygen requirement	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic
Gram stain	+	+	+	+	+	+	+	+
Salt requirement	+	+	+	+	+	+	+	+
Motility	+	+	+	+	+	+	+	+
Sporulation	+	+	+	+	+	+	+	+
Indole	-	-	-	-	-	-	-	-
Production of								
Alkaline phosphate	+	+	+	+	+	NA	+	-
Catalase	+	+	+	+	+	NA	+	+
Oxidase	-	+	+	+	+	-	+	-
Nitrate reductase	-	+	-	+	-	-	+	+
Urease	-	+	-	-	-	-	+	+
Arginine Dihydrolase	NA	-	-	-	-	NA	-	-
β-galactosidase	-	+	+	+	+	NA	-	NA
α -galactosidase	-	-	-	-	+	NA	+	-
N-acetyl-glucosamine	-	+	-	-	NA	NA	NA	+
L-arabinose	-	+	+	-	+	+	+	+
Ribose	-	+	-	+	+	+	+	+
D-mannose	-	+	+	-	+	+	-	-
D-mannitol	-	+	+	+	+	+	+	+
D-sucrose	+	+	+	+	NA	+	-	+
D-glucose	+	+	+	+	+	+	+	+
D-fructose	+	+	+	+	+	+	+	+
D-maltose	-	+	+	-	+	+	-	+
D-lactose	-	+	+	-	+	+	-	+
DNA G+C content (mol %)	39.8	40.1	37.9	42.3	35.8	35.8	38	41.3
Habitat	Cooking salt	Salt lake	Solar saltern soil	Salty soil	Soil	Salty lake	Saline soil	Fermentation liquor for dyeing

G. timonensis Marseille-P2481^T, *G. Gracilibacillus bigeumensis* BH097^T [40], *G. halophilus* YIM-C55.5^T [8], *G. boraciitolerans* T-16X^T [41], *G. saliphilus* YIM91119^T [42], *G. kekensis* K170^T [11], *G. halotolerans* NN^T [5], *G. alcaliphilus* SG103^T [7]. NA= no data available

Table 3: Total cellular fatty acid composition of *Gracilibacillus timonensis* strain Marseille-P2481^T

Fatty acids	IUPAC name	Mean relative %[a]
15:0 anteiso	12-methyl-tetradecanoic acid	45.4 ± 1.5
16:0	Hexadecanoic acid	15.6 ± 1.1
17:0 anteiso	14-methyl-Hexadecanoic acid	13.9 ± 0.6
15:0 iso	13-methyl-tetradecanoic acid	10.3 ± 0.6
17:0 iso	15-methyl-Hexadecanoic acid	5.8 ± 1.0
16:0 iso	13-methyl-Pentadecanoic acid	3.4 ± 0.4
18:0	Octadecanoic acid	1.2 ± 0.1
15:0	Pentadecanoic acid	1.1 ± 0.2
14:0 iso	12-methyl-Tridecanoic acid	1.1 ± 0.1
17:0	Heptadecanoic acid	1.1 ± 0.1
14:0	Tetradecanoic acid	TR
10:0	Decanoic acid	TR
12:0	Dodecanoic acid	TR
13:0 anteiso	10-methyl-Dodecanoic acid	TR
13:0 iso	11-methyl-Dodecanoic acid	TR

[a] Mean peak area percentage calculated from the analysis of FAMEs in 2 sample preparations ± standard deviation (n=3); TR= trace amounts < 1%

Table 4: Nucleotide content and gene count of the genome

Attribute	Value	% of total[a]
Size (bp)	4,548,390	100%
G+C content (bp)	1,808,751	39.8 %
Coding region (bp)	3,844,022	85.07 %
Total genes	4,395	100%
RNA genes	63	1.76 %
Protein-coding genes	4,332	98.23 %
Genes with function prediction	3,043	68.95 %
Genes assigned to COGs	2,797	63.94 %
Genes with peptide signals	474	11.20 %
Genes with transmembrane helices	1,191	27.68 %

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

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

Code	Value	% value	Description
J	212	4.89	Translation
A	0	0	RNA processing and modification
K	266	6.14	Transcription
L	103	2.37	Replication, recombination and repair
B	1	0.02	Chromatin structure and dynamics
D	52	1.20	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	98	2.26	Defense mechanisms
T	154	3.46	Signal transduction mechanisms
M	147	3.39	Cell wall/membrane biogenesis
N	49	1.13	Cell motility
Z	0	0	Cytoskeleton
W	3	0.06	Extracellular structures
U	30	0.69	Intracellular trafficking and secretion
O	107	2.46	Posttranslational modification, protein turnover, chaperones
X	57	1.31	Mobilome: prophages, transposons

C	113	2.60	Energy production and conversion
G	478	11.03	Carbohydrate transport and metabolism
E	201	4.63	Amino acid transport and metabolism
F	100	2.30	Nucleotide transport and metabolism
H	138	3.18	Coenzyme transport and metabolism
I	94	2.16	Lipid transport and metabolism
P	192	4.43	Inorganic ion transport and metabolism
Q	66	1.52	Secondary metabolites biosynthesis, transport and catabolism
R	288	6.64	General function prediction only
S	212	4.89	Function unknown
-	1,535	35.43	Not in COGs

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.

	GT	**VS**	**GH**	**BSh**	**BC**	**AJ**	**BA**	**BS**	**GB**	**HH**
GT	4,332	1,642	1,553	1,496	1,493	1,458	1,423	1,349	1,301	1,266
VS	57.63%	4,253	1,388	1,671	1,535	1,364	1,489	1,493	1,158	1,232
GH	58.41%	59.07%	2,968	1,320	1,275	1,307	1,250	1,244	1,245	1,166
BSh	58.07%	60.84%	57.71%	4,727	1,611	1,347	1,532	1,637	1,115	1,318
BC	55.64%	58.36%	56.44%	57.78%	4,087	1,360	1,591	1,426	1,073	1,244
AJ	66.78%	58.14%	57.88%	58.70%	55.80%	3,594	1,305	1,229	1,084	1,164
BA	53.00%	58.08%	58.75%	56.86%	58.58%	53.39%	3,745	1,417	1,051	1,244
BS	57.00%	58.80%	56.44%	59.75%	55.87%	57.00%	56.70%	3,465	1,086	1,235
GB	59.69%	57.95%	60.19%	59.64%	55.88%	58.23%	58.27%	57.71%	4,450	964
HH	65.19%	57.49%	56.67%	58.82%	55.10%	65.82%	53.45%	56.81%	56.67%	2,771

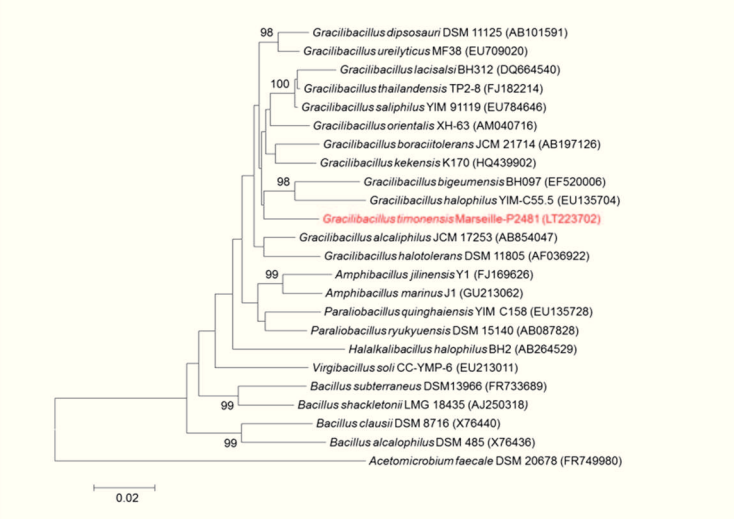
GT : *Gracilibacillus timonensis* strain Marseille-P2481^{T}; HH : *Halalkalibacillus halophilus* strain DSM 18494; AJ : *Amphibacillus jilinensis* strain Y1; VS : *Virgibacillus soli* strain PL205; BC : *Bacillus clausii* strain KSM-K16; BS : *Bacillus subterraneus* strain MITOT1; BSh : *Bacillus shackletonii* strain LMG 18435; BA : *Bacillus alcalophilus* strain ATCC 27647; GH : *Gracilibacillus halophilus* strain YIM-C55.5^{T}; GB : *Gracilibacillus boraciitolerans* strain JCM 21714.

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

	VS	GH	BSh	BC	AJ	BA	BS	GB	HH
GT	38.9% ± 2.5	19.3% ± 2.3	22.6% ± 2.35	30.9% ± 2.45	25.9% ± 2.4	23.6% ± 2.4	29.1% ± 2.45	19.9% ± 2.3	20.6% ± 2.35
VS		20.8% ± 2.35	24.7% ± 2.4	32.1% ± 2.45	24.1% ± 2.3	20% ± 2.34	23.5% ± 2.35	21.1% ± 2.35	24.2% ± 2.45
GH			28.6% ± 2.45	29.9% ± 2.45	18.6% ± 2.3	27.2% ± 2.4	18.6% ± 2.3	17.4% ± 2.2	18.6% ± 2.25
BSh				37.7% ± 2.45	32.6% ± 2.45	32.7% ± 2.45	28.3% ± 2.45	27.6% ± 2.45	18.2% ± 2.25
BC					25.5% ± 2.25	25.4% ± 2.4	25% ± 2.4	33.7% ± 2.45	25.6% ± 2.4
AJ						19.8% ± 2.3	18.2% ± 2.3	24.6% ± 2.4	18.7% ± 2.5
BA							26.4% ± 2.4	18.2% ± 2.25	21.4% ± 2.35
BS								28% ± 2.4	25.8% ± 2.4
GB									20.8% ± 2.3
HH									

dDDH: Digital DNA-DNA hybridization. GT : *Gracilibacillus timonensis* strain Marseille-P2481^{T}; HH : *Halalkalibacillus halophilus* strain DSM 18494; AJ : *Amphibacillus jilinensis* strain Y1; VS : *Virgibacillus soli* strain PL205; BC : *Bacillus clausii* strain KSM-K16; BS : *Bacillus subterraneus* strain MITOT1 ; BSh : *Bacillus shackletonii* strain LMG 18435; BA : *Bacillus alcalophilus* strain ATCC 27647; GH : *Gracilibacillus halophilus* strain YIM-C55.5^{T}; GB : *Gracilibacillus boraciitolerans* strain JCM 21714

Figure 1: Phylogenetic tree highlighting the position of *Gracilibacillus timonensis* strain Marseille-P2481^T relative to other closely related species.



GenBank accession numbers are indicated after the name. Sequences were aligned using Muscle, and phylogenetic inferences obtained using the maximum-likelihood method within the FastTree software. Numbers at the nodes are bootstrap values computed with the Shimodaira-Hasegawa test.

Figure 2: Reference mass spectrum from *Gracilibacillus timonensis* strain Marseille-P2481^T

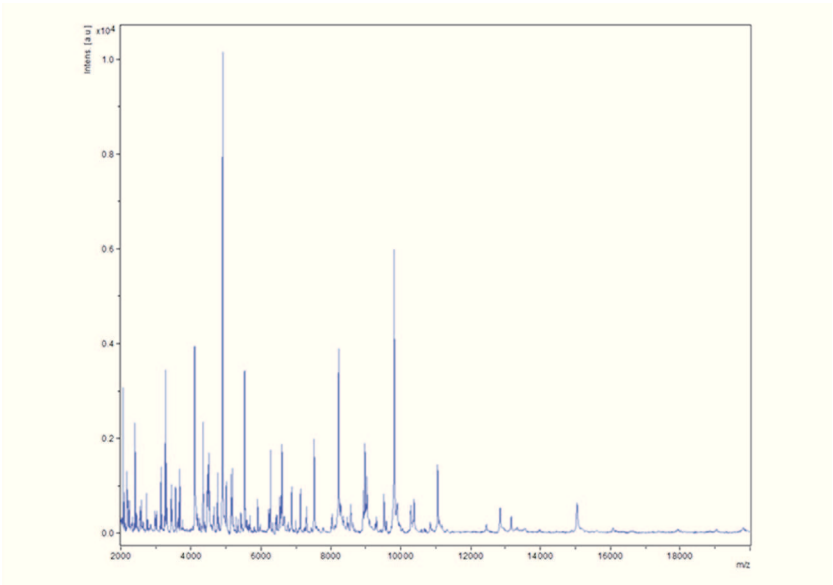


Figure 3: Gel view comparing *Gracilibacillus timonensis* strain Marseille-P2481^{T} to other species within the genera *Gracilibacillus* and *Thalassobacillus*.

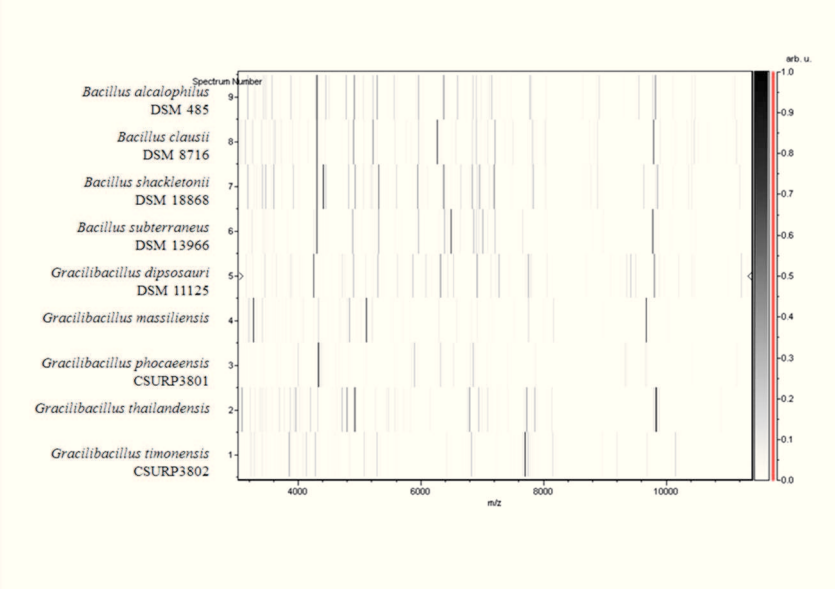


Figure 4: Gram staining of *Gracilibacillus timonensis* strain Marseille-P2481^{T}.

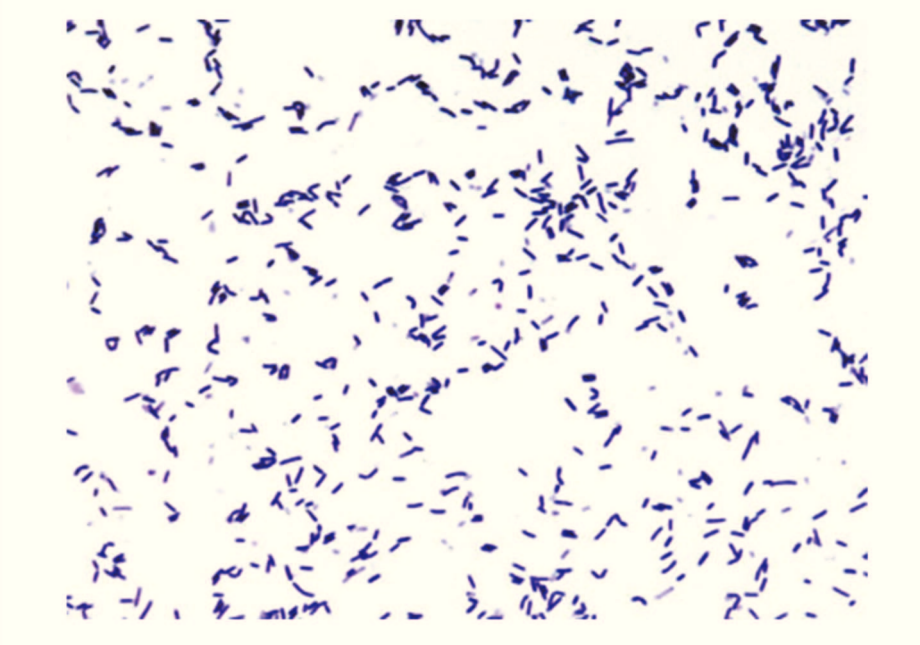


Figure 5: Transmission electron microscopy of *Gracilibacillus timonensis* strain Marseille-P2481^{T}. The scale bar represents.

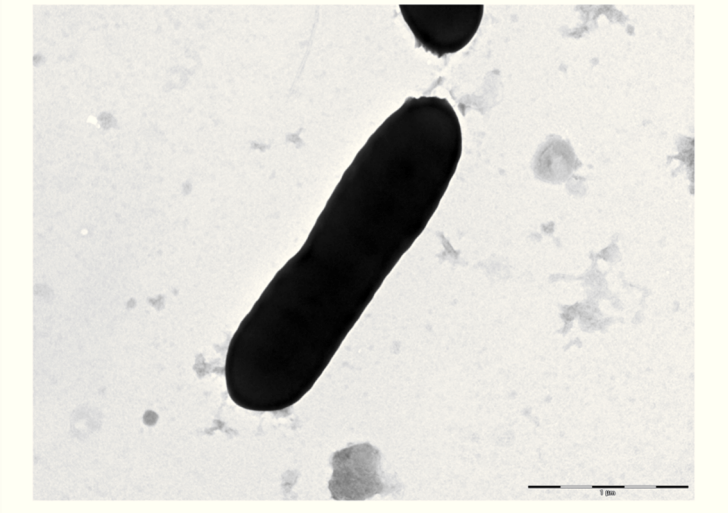
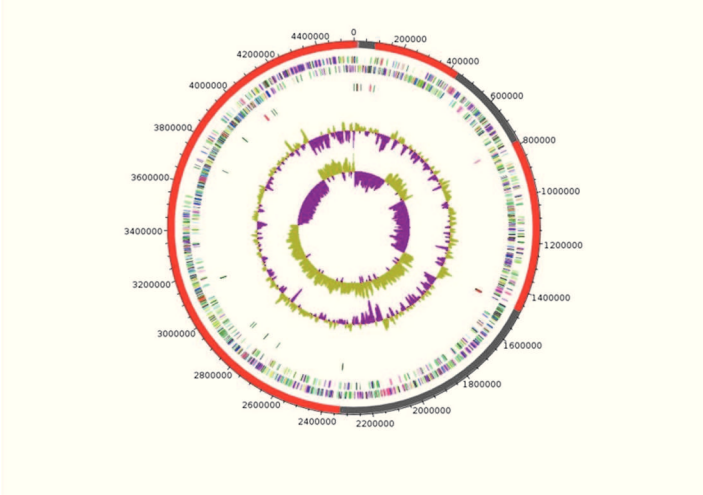


Figure 6: Graphical circular map of the chromosome.



From the 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.

Figure 7: Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins of *Gracilibacillus timonensis* strain Marseille-P2481^T among other species.

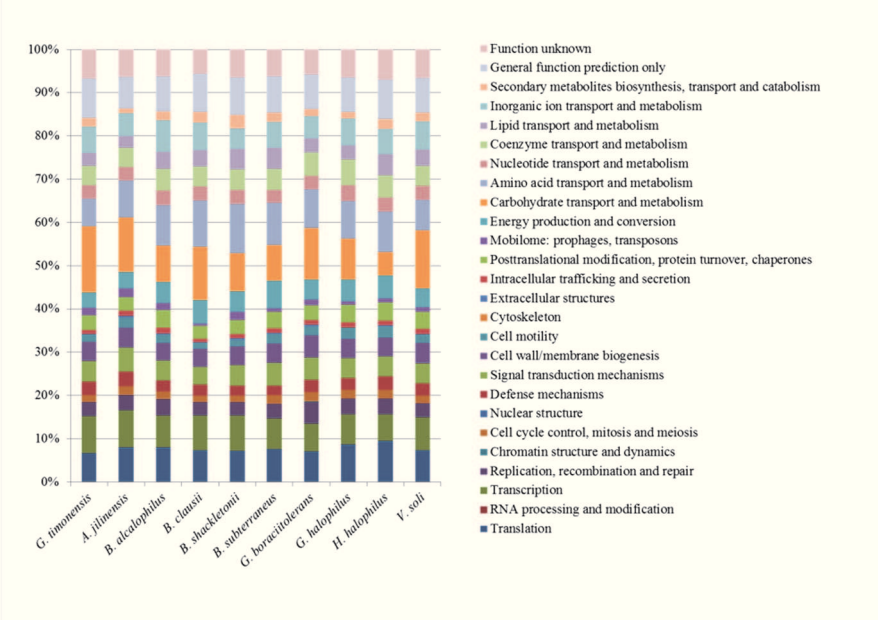
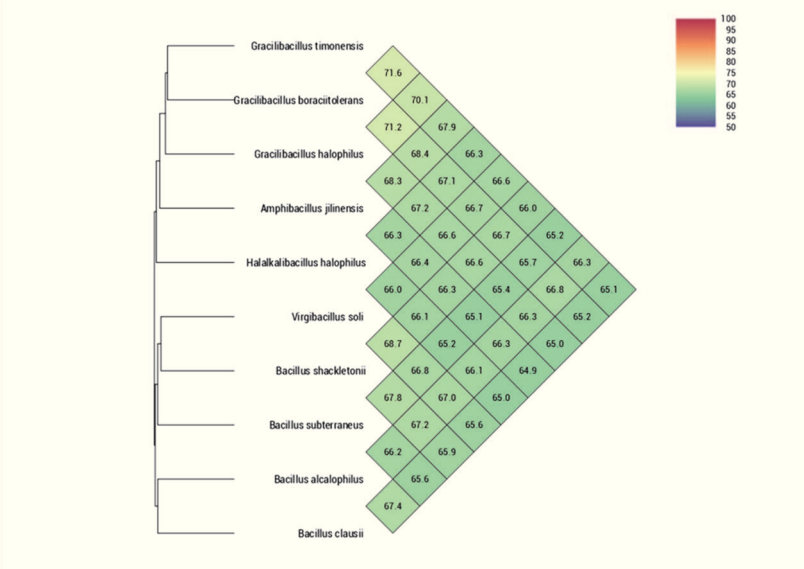


Figure 8: Heatmap generated with OrthoANI values between *Gracilibacillus timonensis* strain Marseille-P2481^T among other closely related species.



Microbial culturomics to isolate halophilic bacteria from table salt: Genome sequence and description of the moderately halophilic bacterium *Bacillus salis* sp. nov.

El hadji Seck[1], Awa Diop[1], Jean Christophe Lagier, Frédéric Cadoret[1], Nicholas Armstrong[1], Jérémy Delerce[1], Pierre-Edouard Fournier[1], Didier Raoult[1,2], Saber Khelaifia[1,*]

[1] URMITE, UM 63, CNRS 7278, IRD 198, Inserm 1095, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, Aix-Marseille Université

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

*Corresponding author: Dr. Saber Khelaifia

URMITE, UM63, UMR CNRS 7278, IRD198, INSERM U1095, Faculté de Médecine, Aix-Marseille Université, IHU-Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13385 Marseille Cedex 5, France. Tel: +33 4 13 73 82 95, Fax: +33 4 13 73 24 02. E-mail: khelaifia_saber@yahoo.fr

Keywords: *Bacillus salis,* genome, taxono-genomics, culturomics, halophiles

Abstract

Bacillus salis strain ES3^T (= CSUR P1478 = DSM 100598) is the type strain of *B. salis* sp. nov. It's an aerobic, Gram-positive, moderately halophilic, motile and spore-forming bacterium. It was isolated from commercial table salt as part of a broad "culturomics" study aiming to maximize the culture conditions for the in-depth exploration of halophilic bacteria in salty food. Here we describe the phenotypic characteristics of this isolate, its complete genome sequence and annotation, together with a comparison with close bacteria. Phylogenetic analysis based on 16S rRNA gene sequences indicated 97.5% of similarity with *Bacillus aquimaris,* the closest species. The 8,329,771 bp long genome (1 chromosome, no plasmid) exhibits a G+C content of 39.19%. It is composed of 18 scaffolds with 29 contigs. Of the 8,303 predicted genes, 8,109 were protein-coding genes, and 194 were RNAs. A total of 5,778 genes (71.25%) were assigned a putative function.

Abbreviations

AGIOS: Average of Genomic Identity of Orthologous gene Sequences

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: Gaz Chromatography/Mass Spectrometry

MALDI-TOF MS: Matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry

ORF: Open Reading Frame

URMITE: Unité des Maladies Infectieuses et Tropicales Emergentes

Introduction

Halophiles are considered as micro-organisms living in hypersaline environments and which often require a high salt concentration for growth. They are involved in centuries old processes, such as production of salt and fermented food consumed by humans (Kivistö and Karp 2011; Lee 2013). Today, with the emergence of new biological technologies, these organisms have been isolated and described from many traditional foods (Lee 2013) such as salt (Diop et al. 2016).

Despite all the recent technological advances in molecular biology, pure culture is the only way to characterize the physiological properties of bacteria and to evaluate their potential virulence (Vartoukian et al. 2010). Therefore, we tried to investigate the population of halophilic prokaryotes in the human gut and salty food by using a culture approach. This approach allowed us to isolate a new member of the *Bacillus* genus. This bacterium is Gram-negative, strictly aerobic, moderately halophilic, motile, and was isolated from commercial table salt. This isolation was part of a "culturomics" study, using high salt culture conditions in order to cultivate halophilic bacteria from human feces and environment samples (Lagier et al. 2012). This isolate is described using a new and innovative method that we have implemented (Ramasamy et al. 2014). The old methods based on 16S rRNA sequencing, phylogeny, G + C content and DNA-DNA hybridization (DDH), are fastidious and include many limitations (Auch et al. 2010; Ramasamy et al. 2014).

The emergence of new tools for DNA sequencing and technology such as MALDI-TOF-MS has allowed an increase in available genomic and proteomic data during the last years (Qin et al. 2010; Seng et al. 2013). These technological advances allowed us to develop a new way of describing bacterial species and taking into account genomic and protonic information (Bouvet et al. 2014).

Here, we present a summary classification and a set of features for *B. salis* strain ES3^T (= CSUR P1478 = DSM 100598) together with the description of the complete genomic sequence and its annotation.

1. Materials and methods

1.1. Strain isolation and identification

Culture condition

Culture has been realized in aerobic atmosphere on a home-made culture medium consisting of a Columbia agar culture medium (Sigma-Aldrich, Saint-Quentin Fallavier, France) modified by adding (per liter): $MgCl_2$ $6H_2O$, 5 g; $MgSO_4$ $7H_2O$, 5 g; KCl, 2 g; $CaCl_2$ $2H_2O$, 1g; NaBr, 0,5 g; $NaHCO_3$, 0,5 g, glucose, 2 g and 100g/L of NaCl. The pH was adjusted to 7.5 with 10M NaOH before autoclaving (Diop A et al. 2016).

MALDI-TOF MS identification

The identification of our strain was carried out by a MALDI-TOF MS analysis with a Microflex spectrometer (Bruker Daltonics, Leipzig, Germany) as previously described (Lo et al. 2015). Obtained spectra were then compared by using the MALDI Biotyper 3.0 software (Bruker) and also the URMITE's database which is constantly updated. If no identification was possible at the genus or species level (score <1.7), sequencing of the 16S rRNA gene was used in order to achieve a correct identification (Weisburg et al. 1991; Drancourt et al. 2000).

Strain identification by sequencing of the 16S rRNA gene

DNA extraction was performed using the EZ1 DNA Tissue Kit and BioRobot EZ1 Advanced XL (Qiagen, Courtaboeuf, France). The 16S rRNA gene was amplified using PCR technology and universal primers fD1 and rP2 (Weisburg et al. 1991) (Eurogentec, Angers, France). The amplifications and sequencing of the amplified products were performed as previously described (Morel et al. 2015) . Then, 16S rRNA gene sequences were assembled and corrected using Codoncode Aligner software (http://www.codoncode.com) and compared with those

available in GenBank (http://www.ncbi.nlm.nih.gov/genbank/). Identification at the species level was defined by a 16S rRNA gene sequence similarity ≥99% with the sequence of the type strain in GenBank. When the percentage of identity was lower than 98.7%, the studied strain was considered as a new species (Tindall et al. 2014).

1.2. Phylogenetic classification

Phylogenetic analysis based on 16S rRNA of our isolate was performed to identify its phylogenetic affiliations with other close isolates, including other members of the genus *Bacillus*. The MEGA 6 (Molecular Evolutionary Genetics Analysis) software allowed us to construct a phylogenetic tree (Tamura et al. 2013). Sequence alignment of the different species was performed using CLUSTAL W (Thompson et al. 1994) and Evolutionary distance matrices for the neighbour-joining method were calculated using the algorithm of the Kimura two-parameter model (Kimura 1980).

1.3. Physiologic and phenotypic characteristics

Phenotypic tests

The phenotypic characteristics of this strain were studied by testing different parameters. Regarding the temperatures, we studied the growth at 25, 30, 37, 45 and 56°C. Growth at various NaCl concentrations (0.5, 5, 7.5, 10, 15, 200 and 250 %) was investigated. The optimal pH for growth was determined by testing different pH: 5, 6, 6.5, 7, 7.5, 8, 9 and 10. Growth of strain ES3^T was tested under aerobic atmosphere, in the presence of 5% CO_2 and also under anaerobic and microaerophilic atmospheres, created using AnaeroGenTM (ThermoFisher scientific, Saint Aubin, France) and CampyGenTM (ThermoFisher scientific) respectively.

Microscopy

Gram staining and motility were observed by using a light microscope DM1000 (Leica Microsystems, Nanterre, France). Cell morphology was studied using a Tecnai G^{20} Cryo (FEI Company, Limeil-Brévannes, France) transmission electron microscope operated at 200 keV,

after negative staining of bacteria: cells were first fixed with 2.5 % glutaraldehyde in 0.1M cacodylate buffer for at least 1h at 4°C. A 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 s with 1% ammonium molybdate solution in filtered water at RT. Formation of spores was determined after thermal shock and observed under a microscope.

Biochemical test

Acid production from carbohydrates was determined by using the API 50CHB system (bioMérieux, Marcy l'Etoile, France). Other physiological tests were performed with the API 20NE system (bioMérieux) and API ZYM (bioMérieux), according to the manufacturer's instructions.

Antibiotic susceptibility test

Antibiotic susceptibility was determined on Mueller-Hinton agar in a Petri dish using the disk diffusion method according to EUCAST recommendations (BioMérieux) (Matuschek et al. 2014). The following antibiotics were tested: doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim / sulfamethoxazole, imipenem, and metronidazole.

Fatty acids analysis

Cellular fatty acid methyl ester (FAME) analysis was performed by GC/MS. Two samples were prepared with approximately 85 mg of bacterial biomass per tube harvested from several culture plates. Fatty acid methyl esters were prepared as described by Sasser (2006) (Sasser 2006). GC/MS analyses were carried out as described before (Dione et al. 2015). Briefly, 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). 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).

1.4. Genome sequencing

Genomic DNA (gDNA) of *Bacillus salis* was extracted in two steps: a mechanical treatment was first performed by acid washed glass beads (G4649-500g Sigma) using a FastPrep BIO 101 instrument (Qbiogene, Strasbourg, France) at maximum speed (6.5m/sec) for 90s. Then after a 2-hour lysozyme incubation at 37°C, DNA was extracted on the EZ1 biorobot (Qiagen) with EZ1 DNA tissues kit. The elution volume is 50µL. gDNA was quantified by a Qubit assay with the high sensitivity kit (Life technologies, Carlsbad, CA, USA) to 120 ng/µl.

GDNA was sequenced on the MiSeq Technology (Illumina Inc, San Diego, CA, USA) with the mate pair strategy. The gDNA was barcoded in order to be mixed with 11 others projects with the Nextera Mate Pair sample prep kit (Illumina). 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 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 11kb with an optimal size at 6.859 kb. No size selection was performed and 600ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal at 921 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 39.94 nmol/l. The libraries were normalized at 2nM and this library was added as two spots and all were 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 a sequencing run were performed in a single 39-hour run in a 2x251-bp. Total

information of 5.5 Gb was obtained from a 572 K/mm2 cluster density with a cluster passing quality control filters of 96.33 % (11,740,000 passing filter paired reads). Within this run, the index representation for *Bacillus salis* was determined to be of 14.60 %. The 1,662,573 paired reads were trimmed then assembled.

1.5. Genome annotation and comparison

The genome's assembly was performed with a pipeline that enabled to create an assembly with different software Velvet (Zerbino and Birney 2008), Spades (Bankevich et al. 2012) and Soap Denovo (Luo et al. 2012)), on trimmed (MiSeq and Trimmomatic (Bolger et al. 2014)softwares) or untrimmed data (only MiSeq software). For each of the six assemblies performed, GapCloser (Luo et al. 2012) was used to reduce gaps. Then, contamination with Phage Phix was identified (blastn against Phage Phix174 DNA sequence) and eliminated. Finally, scaffolds which size was under 800 bp were removed and scaffolds with a depth value lower than 25% of the mean depth were removed (identified as possible contaminants). The best assembly was selected by using different criteria (number of scaffolds, N50, number of N). For the studied strain, Spades gave the best assembly, with a depth coverage of 99x.

Open Reading Frames (ORFs) were predicted using Prodigal (Hyatt et al. 2010) with default parameters but the predicted ORFs were excluded if they were spanning a sequencing gap region (contained N). The predicted bacterial protein sequences were searched against the Clusters of Orthologous Groups (COG) database using BLASTP (E-value $1e^{-03}$, coverage 0.7 and identity percent 30%). If no hit was found, it searched against the NR database using BLASTP with a E-value of $1e^{-03}$, coverage 0.7 and identity percent 30%. If the sequence length was smaller than 80 amino acids, we used an E-value of $1e^{-05}$. The tRNAScanSE tool (Lowe and Eddy 1997) was used to find transfer RNA genes, whereas ribosomal RNAs were found using RNAmmer (Lagesen et al. 2007). Lipoprotein signal peptides and the number of transmembrane helices were predicted using Phobius (Käll et al. 2004). ORFans were identified

if the BLASTP performed didn't give positive results (E-value was lower than $1e^{-03}$ for ORFs with sequence size greater than 80 amino acids or if alignment lengths were smaller than 80 amino acids, we used an E-value of $1e^{-05}$). Such parameter thresholds have already been used in previous works to define ORFans. The annotation process was performed in the Multi-Agent software system DAGOBAH (Gouret et al. 2011), that include Figenix (Gouret et al. 2005) libraries that provided pipeline analysis.

Artemis was used for data management and DNA Plotter (Carver et al. 2009) for visualization of genomic features. The Mauve alignment tool (version 2.3.1) was used for multiple genomic sequence alignment (Darling et al. 2004). To estimate the mean level of nucleotide sequence similarity at the genome level, we used the MAGI home-made software to calculate the average genomic identity of gene sequences (AGIOS) among compared genomes. Briefly, this software is combined with the Proteinortho software (Lechner et al. 2011) for detecting orthologous proteins in pairwise genomic comparisons, then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman-Wunsch global alignment algorithm. Genomes from the genus *Bacillus* and closely related genera were used for the calculation of AGIOS values. The genomic similarity was evaluated among studied species close to the isolate by the dDDH (http://ggdc.dsmz.de/distcalc2.php).

2. Results and discussion

2.1. Strain identification and phylogenetic analyses

Strain ES3^T was first isolated in May 2014 (Table 1) after 30 days of pre-incubation in aerobic culture on our home-made culture medium at 37°C. No significant MALDI-TOF score was obtained for strain ES3^T against the Bruker and URMITE databases, suggesting that our isolate was not a member of a known species (Seng P et al. 2013). An almost complete 16S rRNA gene sequence of strain ES3^T (Accession number: LN827530) comprising 1505 nt was analyzed.

Comparative 16S rRNA gene sequences analyses showed that strain ES3^T is phylogenetically affiliated to the *Bacillus* genus (Figure 1). The phylogenetic distinctiveness (16S rRNA gene sequence similarity of < 97 %) confirms that strain ES3^T represents a distinct species from the recognized species belonging to *Bacillus* genus (Coorevits et al. 2011). In fact, strain ES3^T exhibited 97.5% nucleotide sequence similarity with *Bacillus aquimaris,* the phylogenetically closest species with a validly published name (Yoon JH at al. 2003a). The reference spectrum for strain ES3^T was thus incremented in our database (Figure 2) and then compared to other known species of the genus *Bacillus*. The differences exhibited are showed in the obtained gel view (Figure 3).

2.2. Phenotypic description

Strain ES3^T formed creamy, smooth,circular and slightly irregular colonies, of 5–8 mm in diameter after incubation at 37°C for 2 days on our halophilic medium under aerobic atmosphere. Growth occurred between 25–40°C, but not at 55°C. No growth was observed without NaCl, and the strain grew at salt concentrations ranged between 1–25 % (w/v) NaCl, with optimum growth occurring at 10% (w/v) NaCl. Growth occurred between pH 6 and pH 10 with an optimum at pH 7.5. Cells were motile and spore-forming. Gram staining (Figure 4) showed Gram-positive rods. Strain ES3^T exhibited catalase activity but no oxidase. Measured by electron microscopy, the rods had a mean diameter of 1.8 µm and a length of 5.9 µm (Figure 5).

Biochemical test

Using API 50CH strip, positive reactions was observed for D-glucose, D-fructose, D-mannose, arbutin, esculin ferric citrate, salicin, D-maltose, D-saccharose, D-trehalose, melezitose and D-raffinose, amidon; and negative reactions were recorded for glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl-ßD-xylopyranoside, D-galactose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-αD-

mannopyranoside, methyl-αD glucopyranoside, N-Acetyl-glucosamine, D-cellobiose, inulin, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate.

Using API 20NE, positive reactions were obtained for esculin ferric citrate, potassium nitrate, L-tryptophane, D-glucose (fermentation), L-arginine and urea. Glucose was assimilated. Nitrophenyl-βD-galactopyranoside, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate, phenylacetic acid were not assimilated.

When assayed with the API ZYM system, alkaline phosphatase, esterase (C4), esterase lipase (C8), acid phosphatase and naphthol-AS-BI-phosphohydrolase had an enzymatic activity, but lipase (C14), leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, N-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase had no activity. Those features have been compared with close species in Table 2.

Antibiotic susceptibility test

Cells were resistant to metronidazole but susceptible to imipenem, doxycycline, rifampicin, vancomycin, amoxicillin, ceftriaxone, gentamicin, trimethoprim/sulfamethoxazole, erythromycin, ciprofloxacin, nitrofurantoin, ampicillin, and penicillin.

Fatty acids analysis

The major fatty acids found for this strain were branched: 12-methyl-tetradecanoic acid (60 %), 14-methyl-Hexadecanoic acid (17 %) and 13-methyl-tetradecanoic acid (10 %). The most abundant fatty acids were saturated ones (99 %), as presented in Table 3.

2.3. Genome properties

The draft genome of strain ES3^T is 8,329,771bp long with 39.19% of G+C content (Table 4, Figure 6). It is composed of 18 scaffolds with 29 contigs. Of the 8,303 predicted genes, 8,109 were protein-coding genes, and 194 were RNAs (20 genes are 5S rRNA, 2 genes are 16S rRNA, 2 genes are 23S rRNA and 170 genes are tRNA genes). A total of 5,778 genes (71.25%) were assigned a putative function (by cogs or by NR blast). 180 genes (2.22%) were identified as ORFans. The remaining genes were annotated as hypothetical proteins (1,748 genes => 21.569%). In Table 4 we summarized statistics and the properties of the genome. In Table 5 is presented the distribution of genes into COGs functional.

2.4. Genome comparison

We compared the genome sequence of strain ES3^T (Accession number: FNMN00000000) with that of halophilic bacteria close to our strain: *Halobacillus halophilus* strain DSM 2266 (HE717023), *Bacillus endophyticus* Hbe603 (NZ_CP011974), *Bacillus marisflavi* JCM 11544 (LGUE00000000), *Paenibacillus sabinae* T27 (CP004078) and *Paenibacillus terrae* HPL-003 (CP003107). The draft genome of strain ES3^T (8.32Mb) was larger than that of *B. endophyticus, B. marisflavi, H. halophilus, P. sabinae* and *P. terrae* (4.86; 4.31; 4.17; 5.27; 6.08 Mb respectively). Its G+C content (39.19%) was smaller than that of *B. marisflavi, H. halophilus, P. sabinae* and *P. terrae* (48.60, 41.82, 52.6 and 46.80% respectively) but larger than that of *B. endophyticus* (36.60 %). The gene content of strain ES3^T (8,303) was larger than that of *B. endophyticus, B. marisflavi, H. halophilus, P. sabinae* and *P. terrae* (4,816; 4,319; 4,857 and 5,396 respectively). However the distribution of genes into COG categories was similar in all compared genomes (Figure 7). In addition, strain ES3^T shared more orthologous genes with species belonging to the same genus (*B. endophyticus, B. marisflavi* respectively 1,153 and 1,151*)* than with others species belonging to other genus (*H. halophilus, P. sabinae* and *P. terrae* respectively shared 997, 701 and 725 orthologous genes) (Table 6). The average percentage of nucleotide sequence identity ranged from 65.34% to 65.84% at the intraspecies

level between strain ES3^T and the two *Bacillus* species but it ranged from 57.74% to 60.05% between strain ES3^T and the two others *Paenibacillus* species. Similar results were obtained for the analysis of the DDH using GGDC software (Tableau 7).

Conclusion

Based on the phenotypic properties (Table 2), Phylogenetic tree (Figure 1), MALDI-TOF analyses (Figure 3) and genomic comparison (taxonogenomics (Table 6 and Table 7) and GGDC results, we propose the creation of *Bacillus salis* sp. nov., represented by the type strain ES3^T.

Description of *Bacillus salis* sp. nov.

Bacillus salis (sa'lis, L. gen. n. salis, from salt, in which the strain was first identified)

Colonies grew after 48h of incubation at 37°C on our home-made culture medium were creamy, smooth, circular and slightly irregular measuring 5–8 mm in diameter. Cells were Gram-positive rods and had a mean diameter of 1.8 µm and a length of 5.9 µm. The strain was able to form subterminal ellipsoidal spores and was motile with a single polar flagella. Growth occurred optimally at 37 °C, pH 7.5 and at 10% NaCl.

API 50CH strip showed positive reactions for D-glucose, D-fructose, D-mannose, arbutin, esculin ferric citrate, salicin, D-maltose, D-saccharose, D-trehalose, melezitose and D-raffinose, amidon. Negative reactions were recorded for glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl-ßD-xylopyranoside, D-galactose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-αD-mannopyranoside, methyl-αD glucopyranoside, N-Acetyl-glucosamine, D-cellobiose, inulin, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium 5-

ketogluconate, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate.

Using API 20NE, positive reactions were obtained for esculin ferric citrate, potassium nitrate, L-tryptophane, D-glucose (fermentation), L-arginine and urea. Glucose was assimilated. Nitrophenyl-βD-galactopyranoside, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate, phenylacetic acid were not assimilated.

When assayed with the API ZYM system, alkaline phosphatase, esterase (C4), esterase lipase (C8), acid phosphatase and naphthol-AS-BI-phosphohydrolase had an enzymatic activity, but lipase (C14), leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, N-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase had no activity.

The type strain was sensitive to imipenem, doxycycline, rifampicin, vancomycin, amoxicillin, ceftriaxone, gentamicin (500 µg), trimethoprim/ sulfamethoxazole, erythromycin, ciprofloxacin nitrofurantoin, ampicillin, penicillin and gentamicin (15 µg) but resistant to metronidazole (500 µg).

The major fatty acids found for this strain were branched: 12-methyl-tetradecanoic acid (60 %), 14-methyl-Hexadecanoic acid (17 %) and 13-methyl-tetradecanoic acid (10 %). The most abundant fatty acids were saturated ones (99 %). The G+C content of the genome is of 39.19%.

The 16S rRNA gene sequence and whole-genome shotgun sequence of *B. salis* strain ES3^T were deposited in GenBank under accession numbers LN827530 and FNMN00000000, respectively. The type strain of *Bacillus salis* is strain ES3^T (= CSUR P1478 = DSM 100598) and was isolated from salt.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank the Xegen Company (www.xegen.fr) for automating the genomic annotation process and Magdalen Lardière for English language correction. This study was funded by the "Fondation Méditerranée Infection".

References

Auch AF, von Jan M, Klenk HP, Göker M (2010) Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2:117–134.

Bankevich A, Nurk S, Antipov D, et al (2012) SPAdes: A new genome assembly algorithm and its applications to aingle-aell sequencing. J Comput Biol 19:455–477.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Bouvet P, Ferraris L, Dauphin B, et al (2014) 16S rRNA gene sequencing, multilocus sequence analysis, and mass spectrometry identification of the proposed new species "*Clostridium neonatale*" J Clin Microbiol 52:4129–4136.

Coorevits A, Logan NA, Dinsdale AE, et al (2011) *Bacillus thermolactis* sp. nov., isolated from dairy farms, and emended description of Bacillus thermoamylovorans. Int J Syst Evol Microbiol 61:1954–1961.

Darling AC, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403.

Dione N, Khelaifia S, Lagier J-C, Raoult D (2015) The aerobic activity of metronidazole against anaerobic bacteria. Int J Antimicrob Agents 45:537–540.

Diop A, Khelaifia S, Armstrong N, et al (2016) Microbial culturomics unravels the halophilic microbiota repertoire of table salt: description of *Gracilibacillus massiliensis* sp. nov. Microb Ecol Health Dis 27:32049.

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

Gouret P, Paganini J, Dainat J, et al (2011) Integration of svolutionary 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 Heidelberg, pp 71–87

Gouret P, Vitiello V, Balandraud N, et al (2005) FIGENIX: Intelligent automation of genomic annotation: expertise integration in a new software platform. BMC Bioinformatics 6:198.

Hyatt D, Chen G-L, LoCascio PF, et al (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:1.

Käll L, Krogh A, Sonnhammer EL (2004) A combined transmembrane topology and signal peptide prediction method. J Mol Biol 338:1027–1036.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120.

Kivistö AT, Karp MT (2011) Halophilic anaerobic fermentative bacteria. J Biotechnol 152:114–124.

Lagesen K, Hallin P, Rodland EA, et al (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108.

Lagier J-C, Armougom F, Million M, et al (2012) Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 18:1185–1193.

Lechner M, Findei\s s S, Steiner L, et al (2011) Proteinortho: Detection of (Co-) orthologs in large-scale analysis. BMC Bioinformatics 12:1.

Lee H-S (2013) Diversity of halophilic archaea in fermented foods and human intestines and their application. J Microbiol Biotechnol 23:1645–1653.

Lo CI, Fall B, Sambe-Ba B, et al (2015) MALDI-TOF Mass Spectrometry: A powerful tool for clinical microbiology at Hôpital principal de Dakar, Senegal (West Africa). PLOS ONE 10:e0145889.

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.

Luo R, Liu B, Xie Y, et al (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience 1:18.

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.

Ma Y, Xia Z, Liu X, Chen S (2007) *Paenibacillus sabinae* sp. nov., a nitrogen-fixing species isolated from the rhizosphere soils of shrubs. Int J Syst Evol Microbiol. 57(1):6-11

Morel A-S, Dubourg G, Prudent E, et al (2015) Complementarity between targeted real-time specific PCR and conventional broad-range 16S rDNA PCR in the syndrome-driven diagnosis of infectious diseases. Eur J Clin Microbiol Infect Dis 34:561–570.

Qin J, Li R, Raes J, et al (2010) A human gut microbial gene catalog established by metagenomic sequencing. Nature 464:59–65.

Ramasamy D, Mishra AK, Lagier J-C, 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.

Reva ON, Smirnov VV, Pettersson B, Priest FG (2002) *Bacillus endophyticus* sp. nov., isolated from the inner tissues of cotton plants (Gossypium sp.). Int J Syst Evol Microbiol 52:101–107.

Sasser M (2006) Bacterial identification by Gas Chromatographic Analysis of Fatty Acids Methyl Esters (GC-FAME. MIDI Tech. Note 101

Seng P, Abat C, Rolain JM, et al (2013) 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.

Spring S, Lidwing W, Marquez MC, et al (1996) *Halobacillus* gen. nov., with descriptions of *Halobacillus litoralis* sp. nov., and *Halobacillus trueperi* sp. nov., and transfer of *Sporosarcina halophila* to *Halobacillus halophilus* comb. nov. Int J Syst Evol Microbiol 46:492–496.

Tamura K, Stecher G, Peterson D, et al (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680.

Tindall BJ (2014) The designated type strain of *Pseudomonas halophila* Fendrich 1989 is DSM 3051, the designated type strain of *Halovibrio variabilis* Fendrich 1989 is DSM 3050, the new name *Halomonas utahensis* (Fendrich 1989) Sorokin and Tindall 2006 is created for the species represented by DSM 3051 when treated as a member of the genus *Halomonas*, the combination *Halomonas variabilis* (Fendrich 1989) Dobson and Franzmann 1996 is rejected, and the combination *Halovibrio denitrificans* Sorokin et al. 2006 is validly published with an emendation of the description of the genus *Halovibrio Fendrich* 1989 emend. Sorokin et al. 2006. Opinion 93. Judicial Commission of the International Committee on Systematics of Prokaryotes. Int J Syst Evol Microbiol 64:3588–3589.

Vartoukian SR, Palmer RM, Wade WG (2010) Strategies for culture of "unculturable" bacteria. FEMS Microbiol Lett 309:1–7.

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173:697–703.

Yoon JH, Kim I-G, Kang KH, et al (2003a) *Bacillus marisflavi* sp. nov. and *Bacillus aquimaris* sp. nov., isolated from sea water of a tidal flat of the Yellow Sea in Korea. Int J Syst Evol Microbiol 53:1297–1303.

Yoon J-H, Oh H-M, Yoon B-D, et al (2003b) *Paenibacillus kribbensis* sp. nov. and *Paenibacillus terrae* sp. nov., bioflocculants for efficient harvesting of algal cells. Int J Syst Evol Microbiol 53:295–301.

Zerbino DR, Birney E (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829.

Table 1: Classification and general features of *Bacillus salis* strain ES3^T

Property	Term
Current classification	Domain: Bacteria
	Phylum: Firmicutes
	Class: Bacilli
	Order: Bacillales
	Family: Bacillaceae
	Genus: Bacillus
	Species: Bacillus salis
	Type strain: ES3^T
Gram stain	Positive
Cell shape	Rod-shaped
Motility	motile
Sporulation	endospore-forming
Temperature range	Mesophile
Optimum temperature	37°C
Optimum pH	7.5
Salinity	5.0-200 g/L
Optimum salinity	100 g/L
Oxygen requirement	aerobic

Table 2: Differential characteristics of *Bacillus salis* strain ES3^T and other strains: *Bacillus marisflavi* strain TF-11^T (Yoon et al. 2003a), *Bacillus endophyticus* strain 2DT^T (Reva et al. 2002), *Halobacillus halophilus* strain SL-4^T (Spring et al. 1996), *Paenibacillus terrae* strain

AM141^T (Yoon et al. 2003b) and *Paenibacillus sabinae* strain T27^T (Ma et al. 2007). +, positive; -, negative; na, not available; v, variable.

Characteristics	***B. salis***	***B. marisflavi***	***B. endophyticus***	***H. halophilus***	***P. terrae***	***P. sabinae***
Cell diameter (µm)	1.8	0.6-0.8	0.5-1.5	0.6-0.8	0.8-1.1	0.7-3.2
Oxygen requirement	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic	Aerobic
Gram stain	+	+ to v	+ to v	+	v	+
Motility	+	+	-	+	+	+
Endospore formation	+	+	-	+	+	+
Catalase	+	+	-	+	+	+
Oxidase	-	-	+	+	-	-
Nitrate reductase	+	na	-	-	+	+
Urease	+	-	-	-	-	na
β-galactosidase	-	na	na	na	-	na
N-acetyl-β-glucosaminidase	-	na	na	na	+	na
L-Arabinose	-	-	+	na	-	-
D-Ribose	+	+	+	na	-	+
D-mannose	+	+	+	+	+	na
D-mannitol	-	+	-	-	+	na
D-sucrose	-	-	+	-	+	-
D-glucose	+	+	+	-	+	+
D-fructose	+	+	-	-	-	-
D-maltose	-	-	-	na	+	+
D-lactose	-	-	-	na	-	-
Starch	+	+	+	na	na	na
Gelatin	+	+	+	na	na	na
Habitat	Table salt	Sea water	Soil sediment	Soil	Soil	Salt lake

Table 3: Cellular fatty acid composition (%).

Fatty acids	Name	Mean relative % (a)
15:0 anteiso	12-methyl-tetradecanoic acid	59.6 ± 1.1
17:0 anteiso	14-methyl-Hexadecanoic acid	17.3 ± 1.0
15:0 iso	13-methyl-tetradecanoic acid	10.1 ± 1.6
16:0	Hexadecanoic acid	3.7 ± 0.2
14:0	Tetradecanoic acid	2.7 ± 0.4
16:0 iso	14-methyl-Pentadecanoic acid	2.1 ± 0.3
17:0 iso	15-methyl-Hexadecanoic acid	1.5 ± 0.1
16:1n9	7-Hexadecenoic acid	TR
5:0 anteiso	2-methyl-Butanoic acid	TR
14:0iso	12-methyl-Tridecanoic acid	TR
13:0 anteiso	10-methyl-Dodecanoic acid	TR
17:1 iso	15-methylHexadecenoic acid	TR
19:0 anteiso	16-methyl-Octadecanoic acid	TR
18:0	Octadecanoic acid	TR
16:1 iso	14-methylPentadecenoic acid	TR
13:0 iso	11-methyl-Dodecanoic acid	TR
12:0	Dodecanoic acid	TR

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

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

Attribute	Value	% of total[a]
Size (bp)	8,329,771	100
G+C content (bp)	3,263,777	39.18
Coding region (bp)	6, 920,184	83.07
Total genes	8, 303	100
RNA genes	194	2.33
Protein-coding genes	8,109	97.66
Genes with function prediction	5, 778	71.25
Genes assigned to COGs	5,277	65.07
Genes with peptide signals	869	10.71
Genes with transmembrane helices	2,032	25.05

[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

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

Code	Value	% value	Description
J	475	5.85	Translation
A	0	0	RNA processing and modification
K	400	4. 93	Transcription
L	215	2.65	Replication, recombination and repair
B	2	0.02	Chromatin structure and dynamics
D	102	1.25	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	130	1.60	Defense mechanisms
T	288	3.55	Signal transduction mechanisms
M	260	3.20	Cell wall/membrane biogenesis
N	118	1.45	Cell motility
Z	0	0	Cytoskeleton
W	15	0.18	Extracellular structures
U	66	0.81	Intracellular trafficking and secretion
O	234	2.88	Posttranslational modification, protein turnover, chaperones
X	56	0.69	Mobilome: prophages, transposons
C	358	4.41	Energy production and conversion
G	431	5.31	Carbohydrate transport and metabolism
E	571	7.04	Amino acid transport and metabolism
F	208	2.56	Nucleotide transport and metabolism
H	318	3. 92	Coenzyme transport and metabolism
I	333	4.10	Lipid transport and metabolism
P	323	3. 98	Inorganic ion transport and metabolism
Q	176	2.17	Secondary metabolites biosynthesis, transport and catabolism
R	560	6. 90	General function prediction only
S	403	4.96	Function unknown
-	2,832	34. 92	Not in COGs

Table 6: Number of orthologous proteins shared between genomes (upper right) and AGIOS values obtained (lower left). BS= *Bacillus salis* strain ES3^T; BE= *Bacillus endophyticus* strain Hbe603; BM= *Bacillus marisflavi* strain JCM 11544; PS= *Paenibacillus sabinae* strain T27^T; PT= *Paenibacillus terrae* strain HPL-003; HH= *Halobacillus halophilus* strain DSM 2266

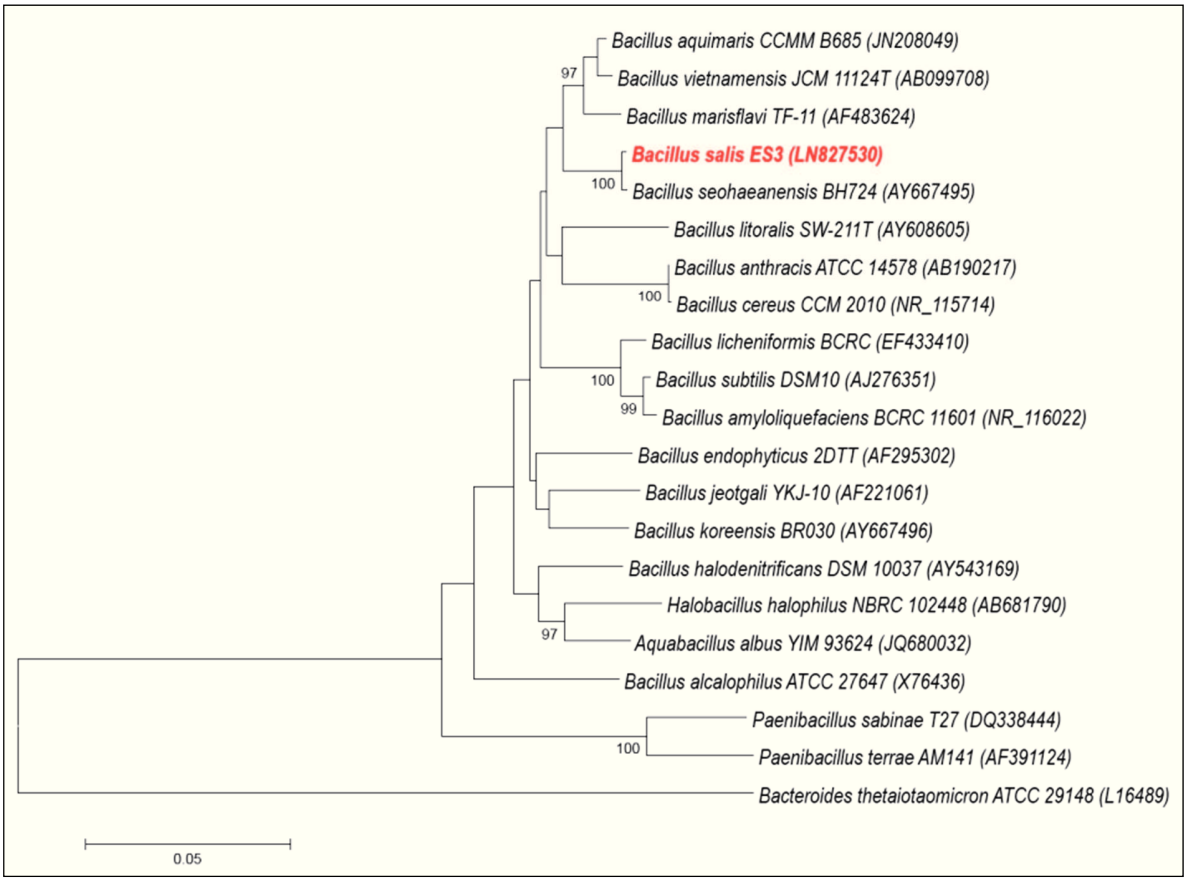
	BS	**BE**	**BM**	**PS**	**PT**	**HH**
BS	8,118	1,153	1,151	701	725	997
BE	65.34%	4,846	1,036	657	717	818
BM	65.84%	62.01%	4,356	639	678	822
PS	57.74%	57.64%	60.32%	4,866	735	518
PT	60.05%	60.41%	60.35%	67.59%	5,446	528
HH	66.03%	62.50%	61.65%	57.85%	59.29%	4,055

Table 7: Pairwise comparison of strain ES3^T with other species using GGDC, formula 2 (DDH estimates based on identities / HSP length)* upper right. BS= *Bacillus salis* strain ES3^T; BE= *Bacillus endophyticus* strain Hbe603; BM= *Bacillus marisflavi* strain JCM 11544; PS= *Paenibacillus sabinae* strain T27; PT= *Paenibacillus terrae* strain HPL-003; HH= *Halobacillus halophilus* strain DSM 2266

	BE	**BM**	**PS**	**PT**	**HH**
BS	23.20% ± 2.38	19.0% ± 2.30	30.50% ± 2.45	22.00% ± 2.39	20.40% ± 2.32
BE		26.50% ± 2.42	29.20% ± 2.44	28.50% ± 2.44	29.80% ± 2.45
BM			28. 90% ± 2.44	28.50% ± 2.44	22.70% ± 2.37
PS				26.00% ± 2.41	29.40% ± 2.44
PT					28.70% ± 2.44

*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 (Figure 1) and phylogenomic analyses as well as the GGDC results.

Figure 1. Phylogenetic tree highlighting the position of *Bacillus salis* strain ES3^T relative to other close species. GenBank accession numbers are indicated in parentheses.



Sequences were aligned using CLUSTALW, and phylogenetic inferences obtained using the Kimura two-parameter model within the MEGA 6 software. *Bacteroides thetaiotaomicron* was used as an outgroup. The scale bar represents 0.05% nucleotide sequence divergence.

Figure 2. Reference mass spectrum from *Bacillus salis* strain ES3^T. Spectra from 12 individual colonies were compared and a reference spectrum was generated.

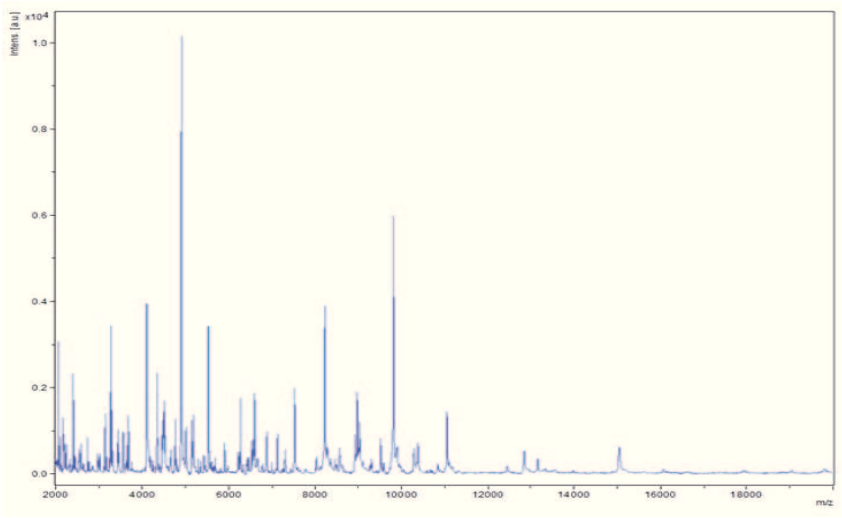
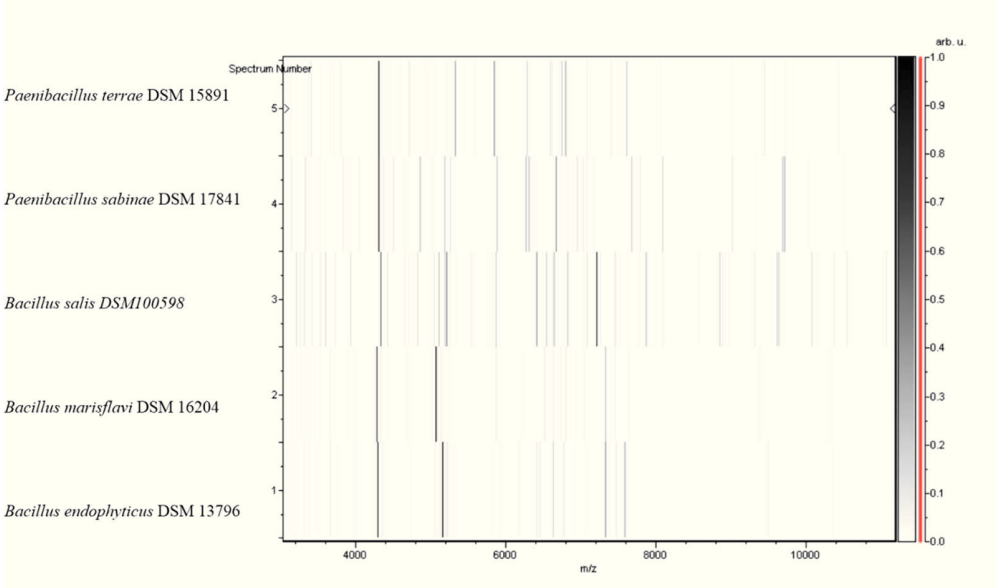


Figure 3. Gel view comparing *Bacillus salis* strain ES3^T to the members of the genera *Bacillus, Paenibacillus.*



The gel view displays the raw spectra of all 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 color bar and the right y-axis indicate the relation between the color a peak is displayed with and the peak intensity in arbitrary units. Displayed species are indicated on the left.
Figure 4. Gram staining of *Bacillus salis* strains ES3^T

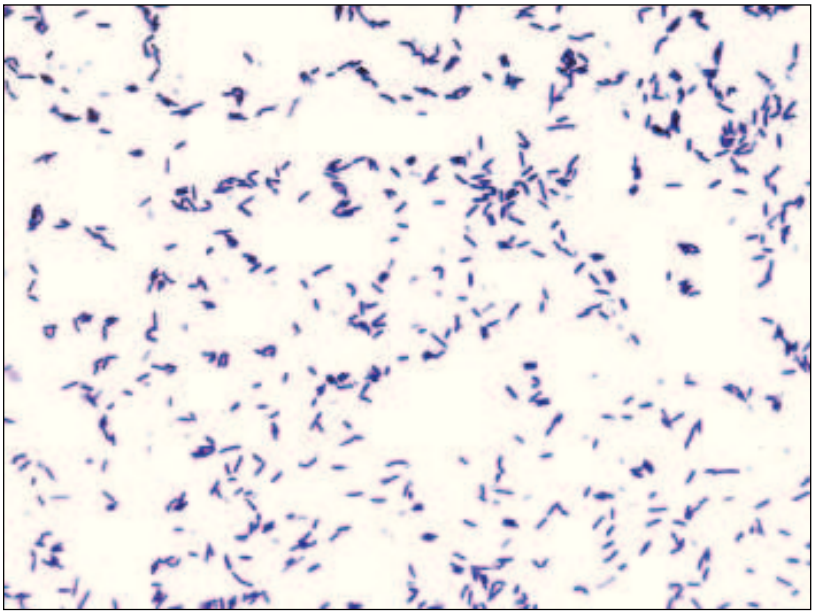


Figure 5. Transmission electron microscopy of *Bacillus salis* strain ES3^T. Cells are observed on a Tecnai G20 transmission electron microscope operated at 200 keV. The scale bar represents 500 nm.

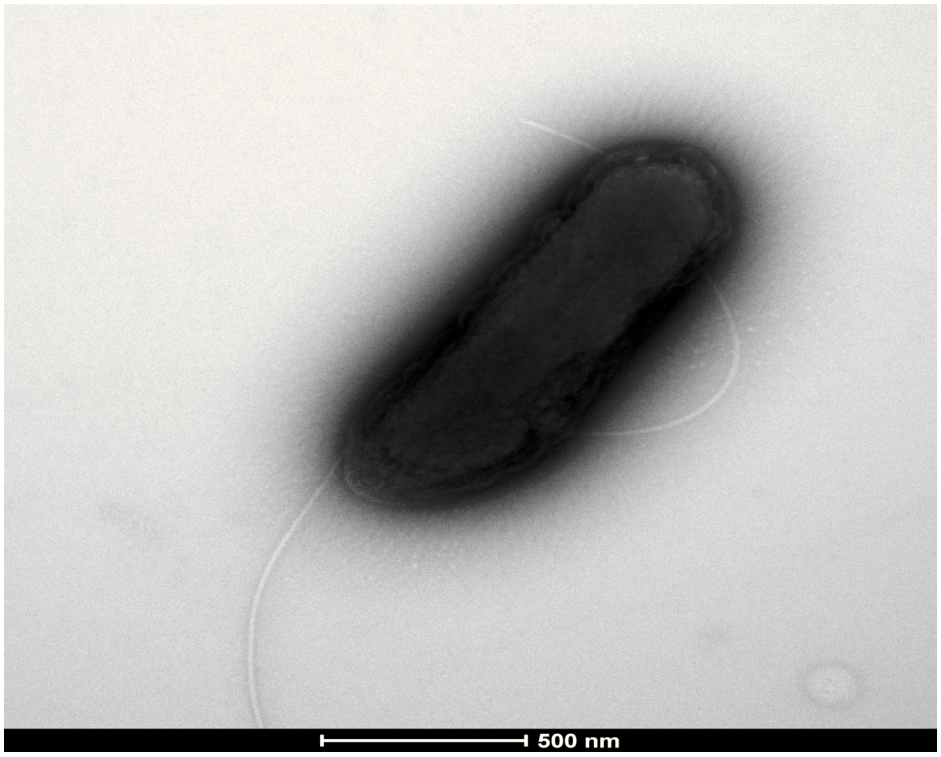


Figure 6: Graphical circular map of the *Bacillus salis* strain ES3^T chromosome. From the outside to center: the outer two circles show open reading frames oriented in the forward (colored by COG

categories) and reverse (colored by COG categories) directions, respectively. The third circle marks the tRNA genes (green). The fourth circle shows the G+C% content plot. The inner-most circle shows GC skew, purple indicating negative values whereas olive is for positive values.

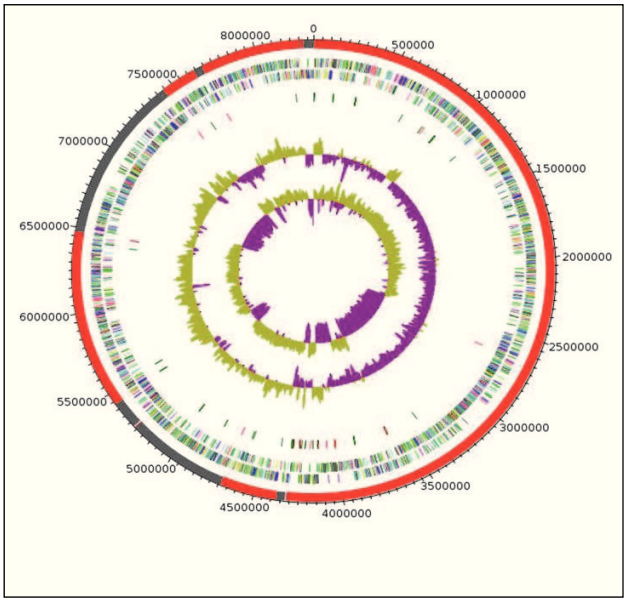
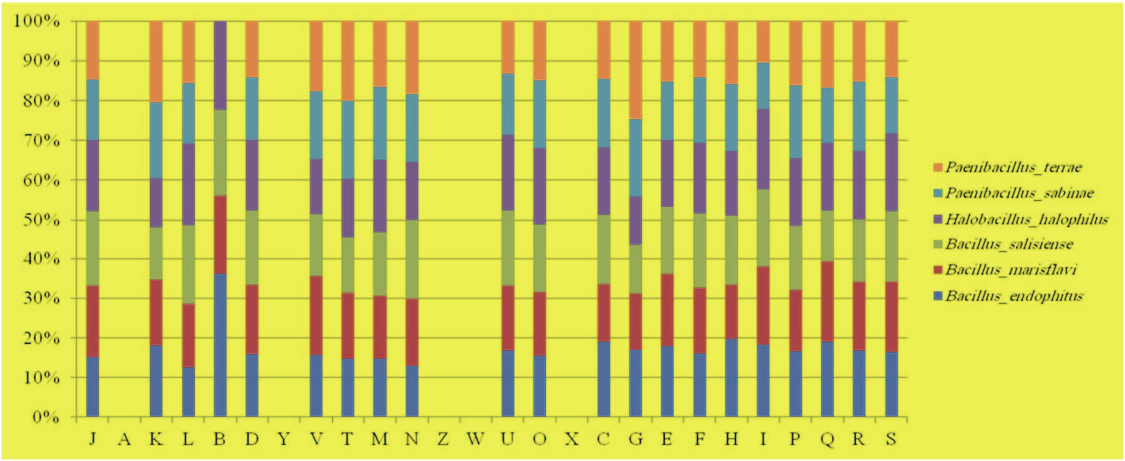


Figure 7. Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins.



NEW SPECIES

'*Gracilibacillus phocaeensis*' sp. nov., '*Sediminibacillus massiliensis*' sp. nov. and '*Virgibacillus ndiopensis*' sp. nov., three halophilic species isolated from salty human stools by culturomics

Q10 B. Senghor[1], S. Khelaifia[1], H. Bassène[2], E. H. Seck[1], P.-E. Fournier[1], C. Sokhna[2], D. Raoult[1] and J.-C. Lagier[1]

1) *URMITE, Aix Marseille Université, UM63, CNRS 7278, IRD 198, INSERM 1095, IHU—Méditerranée Infection, Marseille, France and* 2) *Aix-Marseille* Q1 *Université, URMITE, UMR63, CNRS 7278, IRD 198, INSERM 1095, Dakar, Senegal*

Abstract

Q3 We report the isolation of three bacterial strains that could not be identified by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry screening. '*Gracilibacillus phocaeensis*' sp. nov., '*Sediminibacillus massiliensis*' sp. nov. and '*Virgibacillus ndiopensis*' sp. nov. are Q4 halophilic species isolated from salty human stools by culturomics.

Original Submission: 19 June 2017; **Revised Submission:** 28 August 2017; **Accepted:** 31 August 2017
Article published online: XXX

Corresponding author: J.-C. Lagier, URMITE, Aix Marseille Université, UM63, CNRS 7278, IRD 198, INSERM 1095, IHU—Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13005 Q2 Marseille, France
E-mail: jclagier@yahoo.fr

Culturomics is a new approach using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF MS) for bacteria identification and aiming to cultivate individually all bacterial species from the human gut and also from other human mucosa microbiota. Thus, this approach has allowed a considerable increase in the gut microbiota repertoire, with the description of more than 247 new species in the last few years [1]. Here we report the isolation of three bacterial strains that could not be identified by our MALDI-TOF MS screening on a Microflex spectrometer (Bruker Daltonics, Bremen, Germany) [2,3]. These strains were isolated in 2017 from the salty stools (>1.7% NaCl) of healthy Senegalese individuals. The study was approved by the ethics committee of the Institut Hospitalo-Universitaire Méditerranée Infection under number 2016-011, and all patients provided signed informed consent.

The percentage of NaCl in the stool specimens was determined using a salinity refractometer (Thermo Scientific, Villebon-sur-Yvette, France) by diluting 1 g in 10 mL of distilled water and centrifuging it for 10 minutes at 5000*g*. Then 100 µL of supernatant was deposited in the refractometer; the result was in a straight line, displayed on the screen in *per mille* and then reported in percentage of NaCl.

To cultivate the bacteria from stool samples, we used an aerobic blood culture bottle (Becton Dickinson, Le Pont-de-Claix, France) containing a halophilic medium prepared by modifying a Columbia broth medium (Sigma-Aldrich, Saint-Quentin-Fallavier, France), as detailed in our previous study [4]. The amount of solute per liter was determined by the following formula: concentration (in %, w/v) = 100 × [(mass solute in g)/(volume solution in mL)].

All strains were first isolated in a halophilic culture medium with 15% (w/v) NaCl.

The initial agar-grown colonies were obtained after 24 hours of incubation at 37°C in aerobic conditions. The 16S rRNA genes were sequenced using the universal primer pair fD1–rP2 as previously described [5] using a 3130-XL sequencer (Applied Biosciences, Saint-Aubin, France). Because all the strains exhibited a 16S rRNA sequence homology of <98.7% with their phylogenetically closest species, we thus propose the creation of these three new species according to the nomenclature [6].

Strain Marseille-P3801^T was isolated from stool samples (2% NaCl) of a 20-year-old man from N'Diop. Strain Marseille-P3801^T can grow in media ranging from 2 to 20% (w/v) NaCl (optimum at 7.5 (w/v) NaCl). The growing colonies are yellow

http://dx.doi.org/10.1016/j.nmni.2017.08.006

and circular with a mean diameter of 2 mm. Bacterial cells were motile by using peritrichous flagella under electron microscopy, and were Gram positive, rod shaped and polymorphic, and catalase and oxidase positive. Strain Marseille-P3801^T exhibited a 98.45% sequence identity with *Gracilibacillus thailandensis* strain TP2-8 (GenBank accession no. NR_116568.1) (Fig. 1) [7], which allowed us to classify it as a member of the genus *Gracilibacillus* within the family *Bacillaceae* in the phylum *Firmicutes*. Strain Marseille-P3801^T is the type strain of the new species '*Gracilibacillus phocaeensis*' (pho.ca.een'sis, N.L. masc. adj., from *phocaeensis,* related to the Phocaeans, the founders of Marseille).

Strain Marseille-P3518^T was isolated from stool samples (2% NaCl) of a 15-year-old boy from Dielmo. Agar-grown colonies were beige, circular and shiny with a mean diameter of 2 mm. Bacterial cells were Gram positive, rod shaped and polymorphic, and had positive catalase and oxidase reaction. Strain Marseille-P3518^T exhibited a 97.4% sequence identity with *Sediminibacillus albus* strain NHBX5 (GenBank accession no. NR_044031.1) [8], the phylogenetically closest species with standing in nomenclature (Fig. 2), which putatively classifies it as a member of the genus *Sediminibacillus* within the family *Bacillaceae* in the phylum *Firmicutes*. Strain Marseille-P3518^T is the type strain of the new species *Sediminibacillus massiliensis* (ma.si.lien'sis, L. masc. adj., from *massiliensis,* related to the university hospital in Marseille, France, where the strain was isolated).

Strain Marseille-P3835^T was isolated from in stool samples (3.7% NaCl) of a 11-year-old boy from Dielmo. Strain Marseille-P3835^T is Gram positive, and catalase and oxidase positive. The strain was able to grow in 0.5 to 15% (w/v) NaCl, with an optimum growth at 5% (w/v) NaCl. The agar colonies are pink and circular, with a mean diameter of 2 mm. Strain Marseille-P3835^T exhibited a 16S rRNA sequence similarity of 98.6% with *Virgibacillus zhanjiangensis* strain JSM 079157 (GenBank accession no. NR_116658.1) (Fig. 3) [9]. On the basis of this result, we propose to classify '*Virgibacillus ndiopensis*' as a new representative of the *Virgibacillus* genus belonging to the family *Bacillaceae,* of the phylum *Firmicutes*. Strain Marseille-P3835^T is the type strain of '*Virgibacillus ndiopensis*' (ndiop.en'sis,

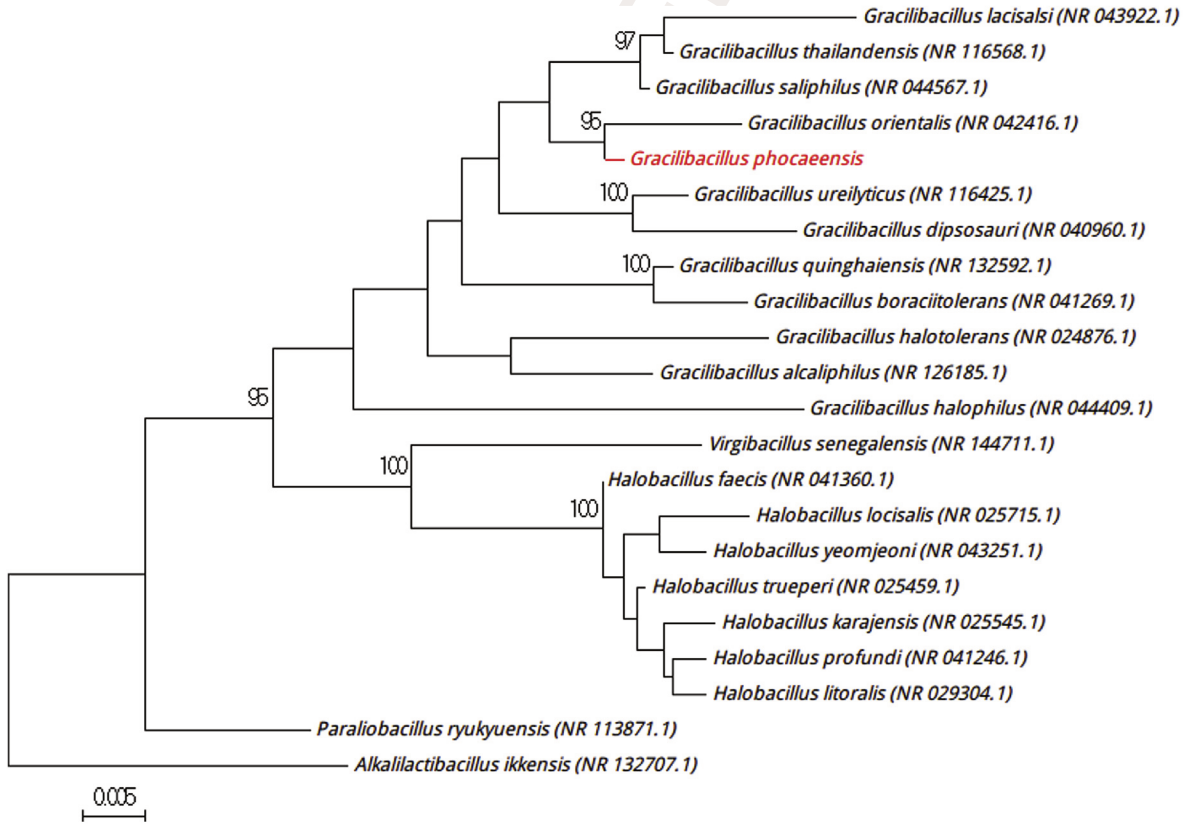


FIG. 1. Phylogenetic tree showing position of '*Gracilibacillus phocaeensis*' Marseille-P3801^T relative to other phylogenetically close neighbours. 16S rRNA gene sequences were aligned using ClustalW, and phylogenetic inferences were obtained using maximum-likelihood method within MEGA software. Numbers at nodes are percentages of bootstrap values obtained by repeating analysis 500 times to generate majority consensus tree. Only bootstrap scores of ≥75 were retained. Scale bar indicates 0.005 nucleotide sequence divergence.

Please cite this article in press as: Senghor B, et al., '*Gracilibacillus phocaeensis*' sp. nov., '*Sediminibacillus massiliensis*' sp. nov. and '*Virgibacillus ndiopensis*' sp. nov., three halophilic species isolated from salty human stools by culturomics, New Microbes and New Infections (2017), http://dx.doi.org/10.1016/j.nmni.2017.08.006

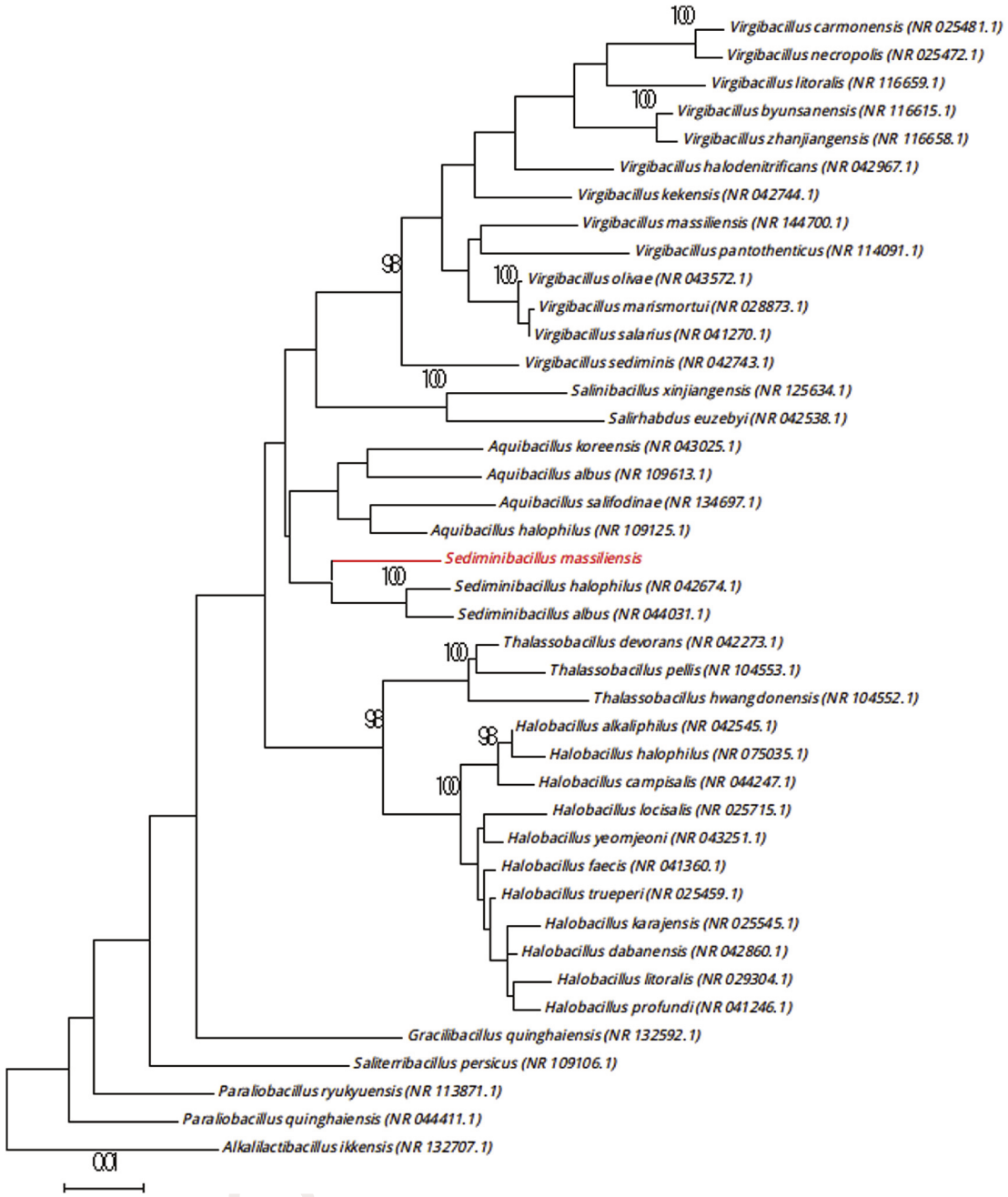


FIG. 2. Phylogenetic tree showing position of '*Sediminibacillus massiliensis*' Marseille-P3518^T relative to other phylogenetically close neighbours. Sequences alignment and phylogenetic inferences were realized as explained in Fig. 1. Scale bar represents 0.01 nucleotide sequence divergence.

L. masc. adj., from *ndiopensis*, related to N'Diop, a Senegalese village from which stool samples were collected).

MALDI-TOF MS spectrum

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

Nucleotide sequence accession number

The 16S rRNA gene sequences were deposited in GenBank under accession numbers '*Gracilibacillus phocaeensis*' Marseille-P3801^T (in progress), '*Sediminibacillus massiliensis*' Marseille-P3518^T (LT671588) and '*Virgibacillus ndiopensis*' Marseille-P3835^T (LT883149).

Please cite this article in press as: Senghor B, et al., '*Gracilibacillus phocaeensis*' sp. nov., '*Sediminibacillus massiliensis*' sp. nov. and '*Virgibacillus ndiopensis*' sp. nov., three halophilic species isolated from salty human stools by culturomics, New Microbes and New Infections (2017), http://dx.doi.org/10.1016/j.nmni.2017.08.006

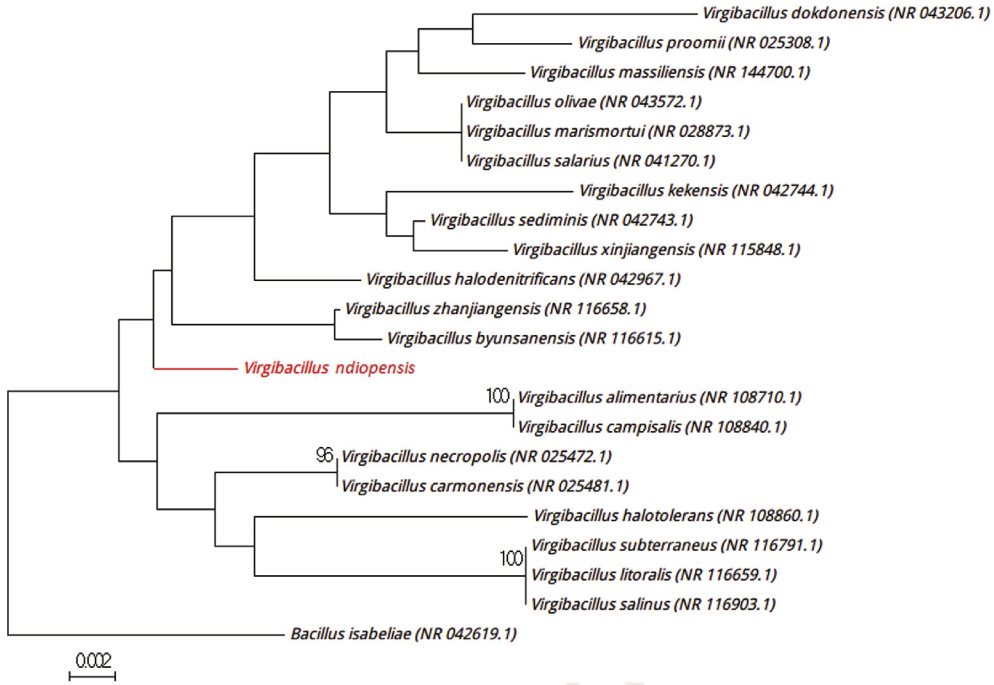


FIG. 3. Phylogenetic tree showing position of '*Virgibacillus ndiopensis*' Marseille-P3835^T relative to other phylogenetically close neighbours. Sequences alignment and phylogenetic inferences were realized as explained for Fig. 1. Scale bar represents 0.002 nucleotide sequence divergence.

Deposit in a culture collection

The strains were deposited in the Collection de Souches de l'Unité des Rickettsies (CSUR, WDCM 875) under the following accession numbers: '*Gracilibacillus phocaeensis*' Marseille-P3801^T (P3801), '*Sediminibacillus massiliensis*' Marseille-P3518^T (3518) and '*Virgibacillus ndiopensis*' Marseille-P3835^T (P3835).

Conflict of interest

None declared.

Acknowledgement

Q8 This study was funded by the Fondation Méditerranée Infection.

References

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

[2] Seng P, Abat C, Rolain JM, Colson P, Lagier JC, Gouriet F, et al. Identification of rare pathogenic bacteria in a clinical microbiology michel laboratory: impact of matrix-assisted laser desorption ionization–time of flight mass spectrometry. J Clin Microbiol 2013;51:2182–94.

[3] .

[4] Senghor B, Seck EH, Khelaifia S, Bassène H, Sokhna C, Threenier PE, et al. Description of *'Bacillus dakarensis'* sp. nov., *'Bacillus sinesaloumensis'* sp. nov., *'Gracilibacillus timonensis'* sp. nov., *'Halobacillus massiliensis'* sp nov., *'Lentibacillus massiliensis'* sp. nov., *'Oceanobacillus senegalensis'* sp. nov., *'Oceanobacillus timonensis'* sp. nov., *'Virgibacillus dakarensis'* sp. nov. and *'Virgibacillus marseillensis'* sp. nov., nine halophilic new species isolated from human stool. New Microbes New Infect 2017;17:45–51.

[5] Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. Bacteriology 1991;173:697–703. Q9

[6] .

[7] Chamroensaksri N, Tanasupawat S, Akaracharanya A, Visessanguan W, Kudo T, et al. *Gracilibacillus thailandensis* sp. nov., from fermented fish (pla-ra). Int J Syst Evol Microbiol 2010;60:944–8.

[8] Wang X, Xue Y, Ma Y. *Sediminibacillus albus* sp. nov., a moderately halophilic, Gram-positive bacterium isolated from a hypersaline lake, and emended description of the genus *Sediminibacillus* Carrasco et al. 2008. Int J Syst Evol Microbiol 2009;59(pt. 7):1640–4.

[9] Peng QZ, Chen J, Zhang YQ, Chen QH, Peng DJ, Cui XL, et al. *Virgibacillus zhanjiangensis* sp. nov., a marine bacterium isolated from sea water. Antonie Van Leeuwenhoek 2009;96:645–52.

Please cite this article in press as: Senghor B, et al., '*Gracilibacillus phocaeensis*' sp. nov., '*Sediminibacillus massiliensis*' sp. nov. and '*Virgibacillus ndiopensis*' sp. nov., three halophilic species isolated from salty human stools by culturomics, New Microbes and New Infections (2017), http://dx.doi.org/10.1016/j.nmni.2017.08.006

CHAPITRE V (ANNEXE)

Travaux collaboratifs sur la description de nouvelles espèces isolées par culturomics.

Avant-propos

La culture microbienne a connu un renouveau ces dernières années grâce à l'émergence de nouvelles techniques d'identification rapide et peu coûteuse de bactéries telles que la spectrométrie de masse (MALDI-TOF). Dans notre unité, nous avons mis en place depuis 2009 un nouveau concept de culture nommé « Microbial culturomics ». C'est une technique de culture à haut débit basée sur la multiplication des conditions de culture en faisant varier des paramètres tels que la température et l'atmosphère d'incubation, le pH du milieu, couplées à une identification rapide des bactéries par MALDI-TOF. La connaissance du microbiote digestif humain a connu un essor exponentiel ces cinq dernières années grâce à culturomics. En effet 1170 bactéries différentes ont pu être isolées à partir de l'intestin humain, dont 247 nouvelles espèces. Ainsi, la culturomics a permis de doubler le nombre de bactéries cultivées dans le microbiote digestif humain.

Cette dernière partie de mon travail a été consacrée aux travaux collaboratifs sur la description de nouvelles espèces isolées par culturomics. Il s'agit de la description par taxonogénomique de ces espèces bactériennes isolées lors de l'exploration du microbiote des enfants sévèrement malnutris atteints de kwashiorkor.

Article 12: *Bacillus kwashiokori* gen. nov., sp. nov., a new bacterial genus isolated from a malnourished child using culturomics.

El hadji Seck[1] | Mamadou Beye[1] | Sory Ibrahima Traore[1] | Saber Khelaifia[1] | Caroline Michelle[1] | Carine Couderc[1] | Souleymane Brah[2] | Pierre-Edouard Fournier[1] |Didier Raoult[1, 3] | Fadi Bittar[1]

Publié dans MicrobiologyOpen

Article 13: *Tessaracoccus massiliensis* sp. nov., a new bacterial species isolated from human gut.

E. Seck[1], S. I. Traore[1], S. Khelaifia[1], M. Beye[1], C. Michelle[1], C. Couderc[1], S. Brah[2], P.-E. Fournier[1], D. Raoult[1, 3] and G. Dubourg[1]

Publié dans New Microbes New Infections

Received: 28 April 2017 | Accepted: 15 August 2017

DOI: 10.1002/mbo3.535

ORIGINAL RESEARCH

WILEY MicrobiologyOpen Open Access

Bacillus kwashiorkori sp. nov., a new bacterial species isolated from a malnourished child using culturomics

El hadji Seck[1] | Mamadou Beye[1] | Sory Ibrahima Traore[1] | Saber Khelaifia[1] | Caroline Michelle[1] | Carine Couderc[1] | Souleymane Brah[2] | Pierre-Edouard Fournier[1] | Didier Raoult[1,3] | Fadi Bittar[1]

[1]Aix-Marseille Univ, CNRS 7278, AP-HM, IRD 198, INSERM 1095, URMITE, IHU Méditerranée Infection, Marseille, France

[2]Hôpital National de Niamey, Niamey, Niger

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

Correspondence
Fadi Bittar, URMITE, IHU Méditerranée Infection, Marseille, France.
Email: fadi.bittar@univ-amu.fr

Funding information
Fondation Méditerranée Infection

Abstract

Strain SIT6^T was isolated from the fecal flora of a severely malnourished child as part of a broad "culturomics" study aiming to maximize the culture conditions for the in-depth exploration of the human microbiota. An analysis of the 16S rRNA gene sequence showed that strain SIT6^T shared 94.1% 16S rRNA gene sequence similarity with *Bacillus thermoamylovorans* DKP^T (NR_029151), the phylogenetically closest type species. Colonies are creamy white, circular, 4–5 mm in diameter after cultivation at 37°C for 24 hr on 5% sheep blood-enriched Colombia agar. Growth occurs at temperatures in the range of 25–56°C (optimally at 37°C). Strain SIT6^T is a gram-positive, facultative anaerobic rod and motile by means of peritrichous flagella and sporulating; it is catalase and oxidase positive. The 2,784,637-bp-long genome, composed of 16 contigs, has a G+C content of 35.19%. Of the 2,646 predicted genes, 2,572 were protein-coding genes and 74 were RNAs. The major fatty acids are saturated species (15:0 iso, 16:0 and 17:0 anteiso). Of the 14 detected fatty acids, 11 are saturated, either linear or branched (iso and anteiso). Digital DNA–DNA hybridization (dDDH) estimation and average genomic identity of orthologous gene sequences (AGIOS) of the strain SIT6^T against genomes of the type strains of related species ranged between 18.6% and 38.3% and between 54.77% and 65.50%, respectively. According to our taxonogenomics results, we propose the creation of *Bacillus kwashiorkori* sp. nov. that contains the type strain SIT6^T (=CSUR P2452^T, =DSM 29059^T).

KEYWORDS

Bacillus kwashiorkori, culturomics, genome, taxonogenomics

1 | INTRODUCTION

Although the human intestinal flora is intrinsically associated with the host genotype and age, many external factors can affect and modify this microbiota, such as antibiotics, probiotics, and diet (Angelakis, Armougom, Million, & Raoult, 2012; Chen, He, & Huang, 2014; Moreno-Indias, Cardona, Tinahones, & Queipo-Ortuño, 2014). Recently, genomic and metagenomic advances have widely participated in describing the human microbiota, but culture isolation remains the only means and the first step to characterize the physiological and genomic properties of a given bacterium and to describe a potential new species (Vartoukian, Palmer, & Wade,

2010). For this reason, in our laboratory we have developed a new strategy called culturomics, which is based on the application of various culture conditions followed by rapid identification using matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) to explore the bacterial composition (Lagier et al., 2012). This new concept has allowed us to significantly increase the bacterial species associated with the human digestive tract and to find many new species (Lagier et al., 2016). Using this strategy (i.e., culturomics), we were able to isolate a new species belonging to the genus *Bacillus*.

This new isolate was described according to the new method that we have implemented (taxonogenomics) (Kokcha et al., 2012; Lagier, Elkarkouri, Rivet, Couderc, & Raoult, 2013; Seck et al., 2016). In brief, it involves using proteomic, fatty acid, and genomic features (Ramasamy et al., 2014; Welker & Moore, 2011; Seng et al., 2013), along with phenotype and some conventional methods, such as 16S rRNA phylogeny and the G+C content.

In this article, we describe the strain SIT6^T (=CSUR P2452^T, =DSM 29059^T) isolated from the stool sample of a kwashiorkor patient.

2 | MATERIALS AND METHODS

2.1 | Organism information

The study and consent procedure were approved by the National Ethics Committee of Nigeria and the Ethics Committee of the Federative Research Institute 48 (Faculty of Medicine, Marseille, France) under the agreement number 09-022. The stool sample was obtained from a 4-month-old Nigerian child suffering from acute malnutrition (kwashiorkor). The patient was not being treated with antibiotics at the time of the sample collection and the sample was stored at −80°C. The stool sample was cultured in blood culture bottles supplemented with sheep blood (BioMérieux, Marcy l'Etoile, France). During a 30-day preincubation period at 37°C in aerobic atmosphere, the liquid culture is then spread on Columbia agar with 5% sheep blood COS medium (BioMérieux, Marcy l'Etoile, France) and the isolated colonies are subsequently identified.

2.2 | Strain identification by MALDI-TOF MS and 16S rRNA sequencing

MALDI-TOF MS analysis of proteins was used to identify the bacteria. Each colony was deposited in duplicate on a MALDI-TOF MSP 96 target and then covered with 1.5 µl of a matrix solution (saturated solution of α-cyano-4-hydroxycinnamic acid in 50% acetonitrile, 2.5% trifluoroacetic acid) to allow the crystallization of molecules. MALDI-TOF MS was performed using the LT Microflex spectrometer (Bruker Daltonics, Leipzig, Germany). All spectra were recorded in positive linear mode for the mass range from 2,000 to 20,000 Da (parameters: ion source 1 [ISI], 20 kV; IS2, 18.5 kV lens, 7 kV). The generated spectra were then compared to the Bruker database, with the addition of new species found through the "culturomics" project. The resulting score dictates whether a tested species can be identified: a score ≥2 with a validly published species enables identification at the species level, a score ≥1.7 but <2 enables identification at the genus level, and a score <1.7 does not enable any identification.

Following three assays, unidentified colonies were identified using 16S rRNA gene sequencing as described previously (Bittar et al., 2014). The isolated colony was suspended in 200 µl distilled water for DNA extraction using an EZ1 DNA Tissue Kit with a BioRobot EZ1 Advanced XL (Qiagen, Courtaboeuf, France). The amplification of the 16S rRNA gene was performed using the universal primer pair fD1 and rP2 (Eurogentec, Angers, France) (Weisburg, Barns, Pelletier, & Lane, 1991). The PCR product was purified and sequenced using the BigDye Terminator v1.1 Cycle Sequencing Kit (PerkinElmer, Courtaboeuf, France) with the following internal primers: 536F, 536R, 800F, 800R, 1050F, and 1050R, and ABI Prism 3130xl Genetic Analyzer capillary sequencer (Applied Biosystems). 16S rRNA amplification and sequencing were carried out as described previously by Morel et al. (2015). The 16S rRNA nucleotide sequences were assembled and corrected using CodonCode Aligner software (http://www.codoncode.com). Then, the BLASTn searches against the GenBank NCBI database (http://blast.ncbi.nlm.nih.gov.gate1.inist.fr/Blast.cgi) and EzBioCloud's Identify Service (http://www.ezbiocloud.net/identify) (Yoon et al., 2017) were performed to determine the percentage of similarity with the closest bacteria.

The MEGA 7 (Molecular Evolutionary Genetics Analysis) software (Kumar, Stecher, & Tamura, 2016) allowed us to construct a phylogenetic tree. Sequence alignment of the different species was performed using CLUSTALW and the calculation of the evolutionary distance was done with the Kimura two-parameter model (Kimura, 1980; Thompson, Higgins, & Gibson, 1994).

2.3 | Growth conditions

In order to determine the ideal growth conditions for strain SIT6^T, different growth temperatures (25°C, 28°C, 30°C, 37°C, 45°C, and 56°C) were tested under anaerobic and microaerophilic atmospheres using GENbag Anaer and GENbag microaer systems, respectively (BioMérieux, Marcy l'Etoile, France). The strain growth was also tested aerobically with and without 5% CO_2. The growth of strain SIT6^T was tested under different pH using a Columbia agar with 5% sheep blood COS medium (BioMérieux, Marcy l'Etoile, France) with NaCl, $MgCl_2$, $MgSO_4$, KCl, $CaCl_2$, and glucose. The pH was modified by adding HCl to the medium and measured with a pH meter. The optimal pH for growth was determined by testing at different pH 5, 6, 6.5, 7, 7.5, 8, and 8.5. Growth at various NaCl concentrations (0.5%, 5%, 7.5%, 10%, 15%, and 200%) was investigated.

2.4 | Morphologic, biochemical, and antibiotic susceptibility tests

Gram staining was performed and observed using a Leica DM 2500 photonic microscope (Leica Microsystems, Nanterre, France) with a 100× oil immersion lens. A thermal shock (80°C during 20 min) was applied on fresh colonies in order to test sporulation. The motility of

the strain was tested by observing fresh colonies using a DM1000 photonic microscope (Leica Microsystems) with a 40× objective lens. Catalase (BioMérieux) activity was determined in 3% hydrogen peroxide solution and oxidase activity was assessed using an oxidase reagent (Becton-Dickinson, Le Pont-de-Claix, France).

Antibiotic susceptibility testing was performed using SIRscan Discs (i2a, Montpellier, France) on Mueller-Hinton agar according to EUCAST 2015 recommendations (Matuschek, Brown, & Kahlmeter, 2014). The following antibiotics were tested: doxycycline (30 µg), rifampicin (30 µg), vancomycin (30 µg), erythromycin (15 µg), ampicillin (10 µg), ceftriaxone (30 µg), ciprofloxacin (5 µg), gentamicin (500 µg), penicillin (10 µg), trimethoprim/sulfamethoxazole (25 + 23.75 µg), imipenem (10 µg), metronidazole (4 µg), clindamycin (15 µg), colistin (50 µg), and oxacillin (5 µg).

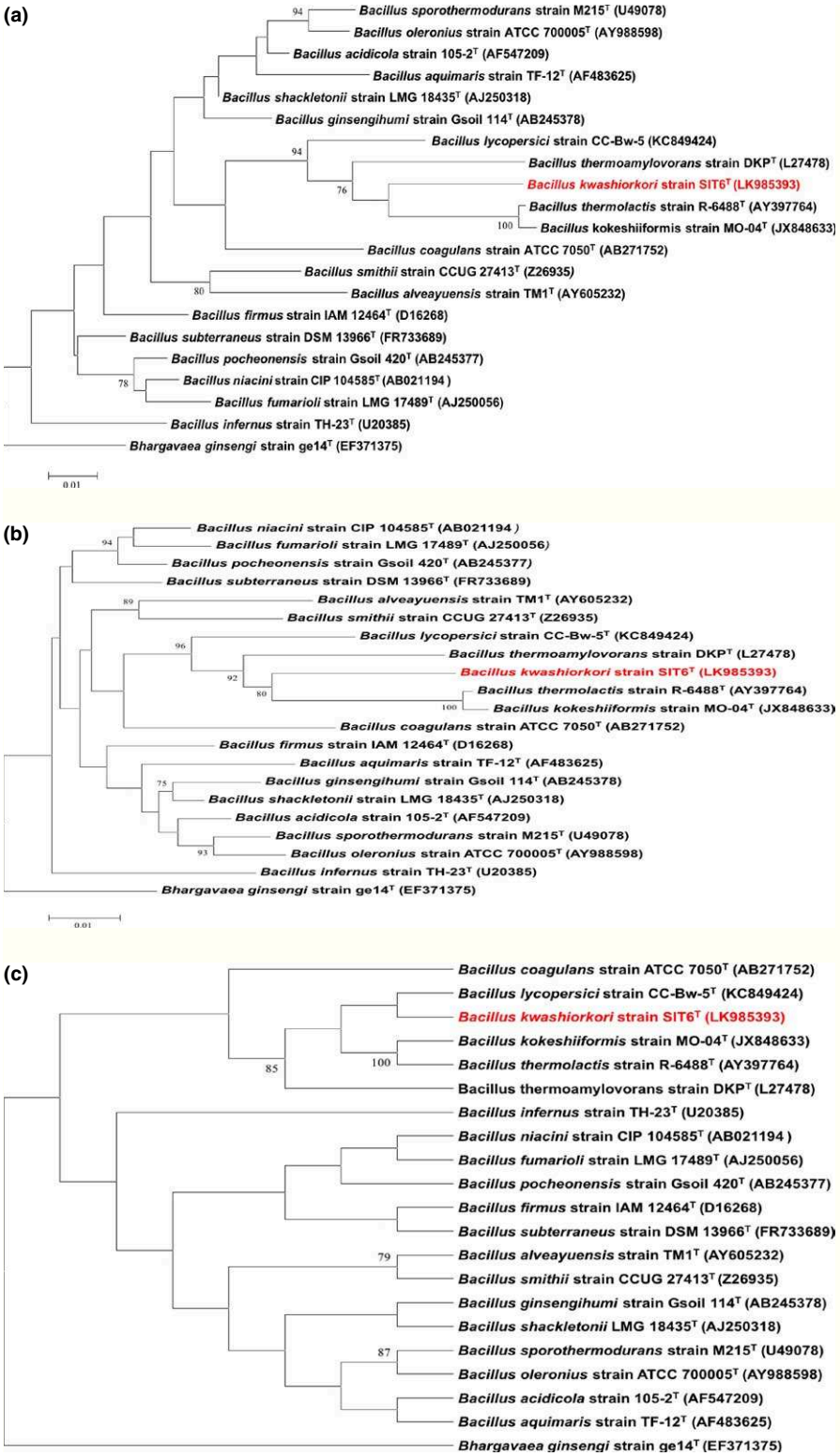


FIGURE 1 Phylogenetic tree showing the position of *Bacillus kwashiorkori* SIT6^T (red) relative to other phylogenetically close members of the family *Bacillaceae*. GenBank accession numbers are indicated in parentheses. Sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained using (a) the maximum-likelihood method, (b) the neighbor-joining method and (c) the maximum parsimony method within the MEGA software. Numbers at the nodes are percentages of bootstrap values obtained by repeating the analysis 1,000 times to generate a majority consensus tree. Only values >70% were displayed. *Bhargavaea ginsengi* ge14^T (EF371375) was used as out-group

Using the commercially available biochemical API 20NE, API ZYM, and API 50CH strips, we investigated the biochemical characteristics of our strain according to the manufacturer's instructions (BioMérieux).

Negative staining was done in order to visualize the cell morphology. Cells 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 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 the cells were negatively stained for 10 s with 1% ammonium molybdate solution in filtered water at room temperature. Electron micrographs were acquired with a Tecnai G20 Cryo (FEI) transmission electron microscope operated at 200 keV.

2.5 | FAME analysis by gas chromatography/mass spectrometry

Cellular fatty acid methyl ester (FAME) analysis was performed by GC/MS. Two samples were prepared with approximately 2 mg of bacterial biomass each, harvested from five culture plates. Fatty acid methyl esters were prepared as described by Sasser (2006). GC/MS analyses were carried out as described previously (Dione et al., 2016). In brief, fatty acid methyl esters were separated using an Elite 5-MS column and monitored by mass spectrometry (MS) (Clarus 500-SQ 8 S, PerkinElmer, Courtaboeuf, France). A spectral database search was performed using MS Search 2.0 operated with the Standard Reference Database 1A (NIST, Gaithersburg, MD, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK).

2.6 | Genomic DNA preparation

After pretreatment by a lysozyme (incubation at 37°C for 2 hr), the DNA of strain SIT6^T was extracted on the EZ1 BioRobot (Qiagen) with the EZ1 DNA tissues kit. The elution volume was 50 µl. Genomic DNA (gDNA) was quantified by a Qubit assay with the high sensitivity kit (Life Technologies, Carlsbad, CA) to 55.8 ng/µl.

2.7 | Genome sequencing and assembly

Genomic DNA (gDNA) of *B. kwashiorkori* was sequenced on MiSeq Technology (Illumina Inc., San Diego, CA) with the mate pair strategy. The gDNA was barcoded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). gDNA was quantified by a Qubit assay with the high sensitivity kit (Thermo Fisher Scientific, Waltham, MA) to 66.2 ng/µl. The mate pair library was prepared with 1 µg of gDNA using the Nextera mate pair Illumina guide. The gDNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of the fragmentation was validated on an Agilent 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA) with a DNA 7500 LabChip. The DNA fragments ranged in size from 1 kb to 11 kb, with an optimal size at 3.927 kb. No size selection was performed and 505 ng of tagged fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal size of 597 bp on the Covaris device S2 in microtubes (Covaris, Woburn, MA). The library profile was viewed on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc.) and the final concentration library was measured at 59.2 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. An automated cluster generation and sequencing run was performed in a single 39-hr run in a 2 × 251 bp.

TABLE 1 Classification and general features of *Bacillus kwashiorkori* strain SIT6^T

Property	Term
Current classification	Domain: Bacteria
	Phylum: Firmicutes
	Class: Bacilli
	Order: Bacillales
	Family: Bacillaceae
	Genus: *Bacillus*
	Species: *kwashiorkori*
	Type strain: SIT6^T
Gram stain	Positive
Cell shape	Rod shaped
Motility	Motile
Sporulation	Endospore forming
Temperature range	Mesophile
Optimum temperature	37°C
Optimum pH	7.5
Salinity	0.0–5.0 g/L
Optimum salinity	0 g/L
Oxygen requirement	Facultative aerobic

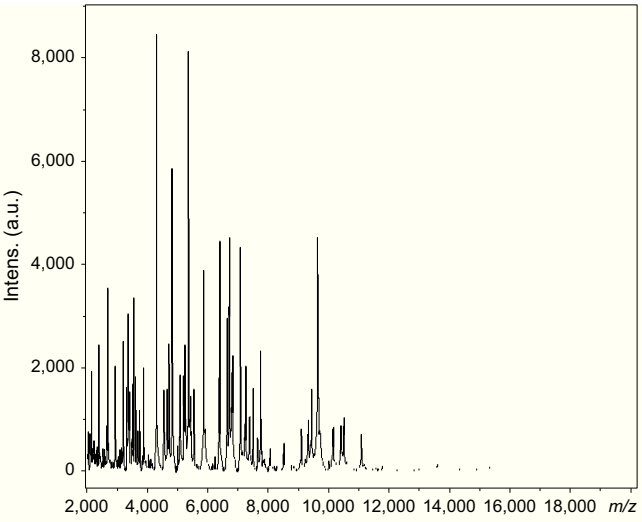


FIGURE 2 Reference mass spectrum from *Bacillus kwashiorkori* SIT6^T strain. Spectra from 12 individual colonies were compared and a reference spectrum was generated

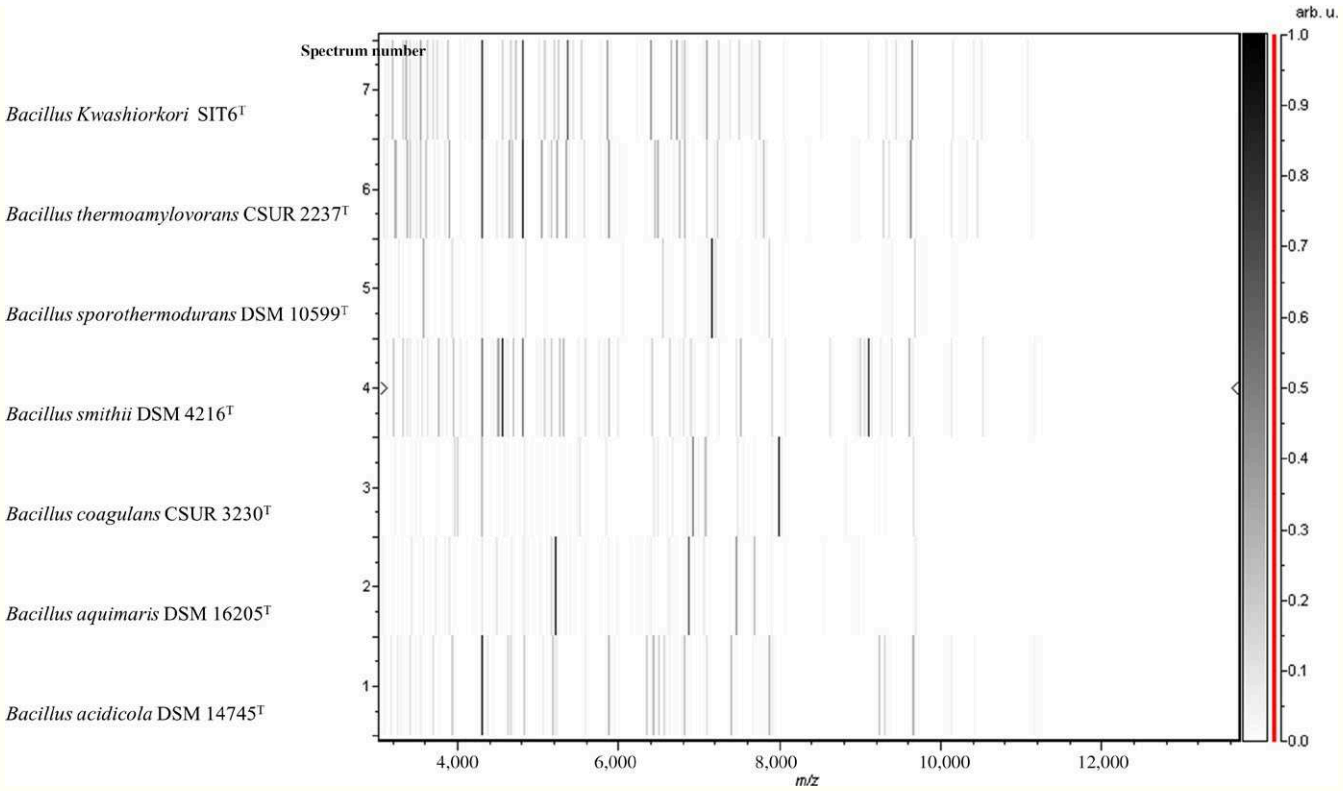


FIGURE 3 Gel view comparing *Bacillus kwashiorkori* SIT6^T spectra with other members of the genus *Bacillus*. The gel view displays the raw spectra of all loaded spectrum files arranged in a pseudo-gel-like look. The *x*-axis records the m/z value. The left *y*-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a gray scale scheme code. The color bar and the right *y*-axis indicate the relation between the color a peak is displayed with and the peak intensity in arbitrary units. Displayed species are indicated on the right

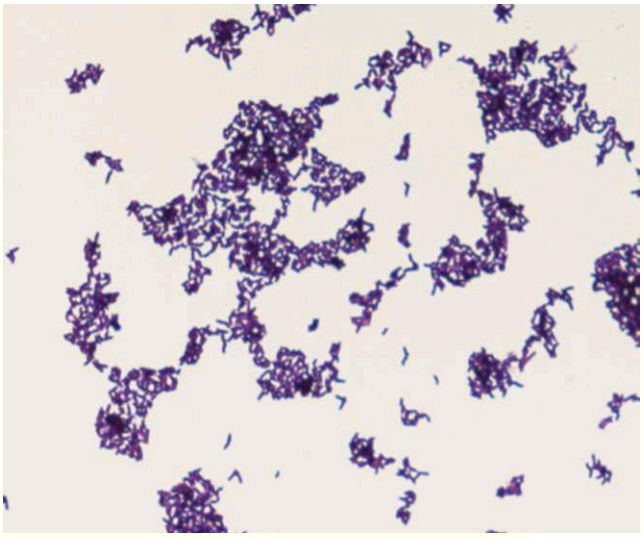


FIGURE 4 Gram staining of *Bacillus kwashiorkori* SIT6^T

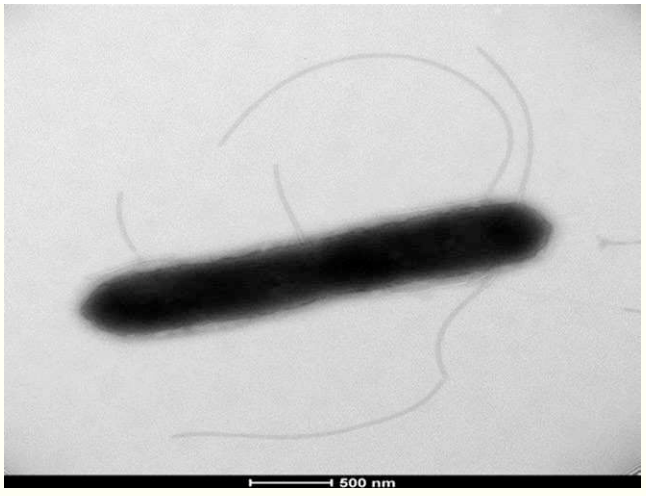


FIGURE 5 Transmission electron microscopy of *Bacillus kwashiorkori* SIT6^T using a Morgani 268D (Philips) at an operating voltage of 60 kV. The scale bar represents 500 nm

2.8 | Genome annotation and analysis

Open reading frames (ORFs) were predicted using Prodigal (http://prodigal.ornl.gov/) with default parameters. However, the predicted ORFs were excluded if they spanned a sequencing gap region. The predicted bacterial protein sequences were searched against GenBank and Clusters of Orthologous Group (COG) databases using BLASTP. The tRNAs and rRNAs were predicted using the tRNAScan-SE and RNAmmer tools, respectively. Signal peptides and numbers of transmembrane helices were predicted using SignalP (Nielsen,

Engelbrecht, Brunak, & von Heijne, 1997) and TMHMM (Krogh, Larsson, von Heijne, & Sonnhammer, 2001), respectively. Mobile genetic elements were predicted using PHAST (Zhou, Liang, Lynch, Dennis, & Wishart, 2011) and RAST (Aziz et al., 2008). ORFans were identified if their BLASTP E-value was lower than 1e-03 for alignment length >80 amino acids. If alignment lengths were <80 amino acids, we used an E-value of 1e-05. Such parameter thresholds have already been used in previous work to define ORFans. Artemis (Carver, Harris, Berriman, Parkhill, & McQuillan, 2012) and DNA plotter (Carver, Thomson, Bleasby, Berriman, & Parkhill, 2009) were used for data management and visualization of genomic features, respectively. The mauve alignment tool (version 2.3.1) was used for multiple genomic sequence alignment (Darling, Mau, Blattner, & Perna, 2004). The mean level of nucleotide sequence similarity at the genome level between *B. kwashiorkori* and other *Bacillus* species was estimated using the Average Genomic Identity of Orthologous Gene Sequences (AGIOS) in-house software (Ramasamy et al., 2014). This software combines the functionality of other software programs: Proteinortho (Lechner et al., 2011) (detects orthologous proteins between genomes compared two by two, then retrieves the corresponding genes) and the Needleman–Wunsch global alignment algorithm (determines the mean percentage of nucleotide sequence identity among orthologous ORFs). Finally, the Genome-to-Genome Distance Calculator (GGDC) web server (http://ggdc.dsmz.de) was used to estimate the similarity between the compared genomes (Auch, Jan, Klenk, & Göker, 2010; Meier-Kolthoff, Auch, Klenk, & Göker, 2013).

Average nucleotide identity at the genome level between *B. kwashiorkori* (CTDX00000000) and the other species *B. firmus* (BCUY00000000), *B. shackletonii* (LJJC00000000), *B. smithii* (BCVY00000000), *B. aquimaris* (LQXM00000000), *B. thermoamylovorans* (CCRF00000000), *B. coagulans* (CP003056), *B. alveayuensis* (JYCE00000000), *B. sporothermodurans* (LQYN00000000), *B. acidicola* (LWJG00000000), and *B. ginsengihumi* (JAGM00000000) was estimated using BLASTN and a home made software, following the algorithm described by Ouk, Chun, Lee, and Park (2016).

3 | RESULTS AND DISCUSSION

3.1 | Strain identification and phylogenetic analyses

Strain SIT6^T was first isolated in May 2014 after a 30-day preincubation in a blood culture bottle with sheep blood and cultivation on 5% sheep blood-enriched Colombia agar in an aerobic atmosphere at 37°C. No significant MALDI-TOF score was obtained for strain SIT6^T against the Bruker and URMITE databases, suggesting that the isolate was not a member of a known species. Strain SIT6^T shared 94.1% 16S rRNA gene sequence similarity with *B. thermoamylovorans* DKP^T (NR_029151) using GenBank NCBI database (reference RNA sequences). Although the 16S rRNA gene sequence of strain SIT6^T showed 94.58% similarity with *Bacillus kokeshiiformis* MO-04^T and 94.57% similarity with *Bacillus thermolactis* R-6488^T by EzBioCloud's identify server. Figure 1a, b, and c present the phylogenetic trees of strain SIT6^T relative to other closest type species with a validly published name using maximum-likelihood, neighbor-joining, and maximum parsimony methods, respectively. Consequently, as this 16S rRNA nucleotide sequence similarity was lower than the threshold of 98% recommended by Tindall, Rosselló-Mora, Busse, Ludwig, and Kämpfer (2010) to delineate a new species; it was classified as a new species called *Bacillus kwashiorkori* SIT6^T (Table 1). Furthermore, this percentage of similarity comprised in the range of percentage similarity of *Bacillus* species (82.7–100%), confirming the new species status (Rossi-Tamisier, Benamar, Raoult, & Fournier, 2015). The reference spectrum for strain SIT6^T was thus incremented in the URMITE database (http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database) (Figure 2) and then compared to other known species of the genus *Bacillus*. The differences exhibited are shown in the obtained gel view (Figure 3).

TABLE 2 Cellular fatty acid composition

Fatty acids	IUPAC name	Mean relative %[a]
15:0 iso	13-Methyl-tetradecanoic acid	19.6 ± 1.2
16:0	Hexadecanoic acid	19.5 ± 0.4
17:0 anteiso	14-Methyl-hexadecanoic acid	16.5 ± 1.3
18:1n12	6-Octadecenoic acid	12.7 ± 1.9
18:0	Octadecanoic acid	9.3 ± 0.4
17:0 iso	15-Methyl-hexadecanoic acid	6.9 ± 1.7
15:0 anteiso	12-Methyl-tetradecanoic acid	4.5 ± 0.1
18:2n6	9,12-Octadecadienoic acid	4.0 ± 0.2
16:0 iso	14-Methyl-pentadecanoic acid	3.9 ± 0.5
18:1n5	13-Octadecenoic acid	1.5 ± 0.1
14:0	Tetradecanoic acid	TR
17:0	Heptadecanoic acid	TR
15:0	Pentadecanoic acid	TR
14:0 iso	12-Methyl-tridecanoic acid	TR

[a]Mean peak area percentage ± standard deviation (n = 3); TR, trace amounts <1%.

3.2 | Phenotypic description

Growth of strain SIT6^T was observed between 25°C and 56°C on 5% sheep blood Colombia agar and optimal growth was achieved at 37°C after 24 hr incubation in aerobic conditions (37°C was the temperature at which this strain grows most rapidly). Poor growth occurred under microaerophilic and anaerobic conditions. Cells were motile and sporulating. Colonies were circular, white with a mean diameter of 5 mm on blood-enriched Colombia agar. Gram staining (Figure 4) showed gram-positive rods. Using electron microscopy, the rods had a mean diameter of 1.8 µm and a length of 5.9 µm (Figure 5). Catalase and oxidase activities were positive for strain SIT6^T.

The major fatty acids are saturated species (15:0 iso, 16:0 and 17:0 anteiso). Of the 14 detected fatty acids, 11 are saturated, either linear

TABLE 3 Phenotypic characteristics of *Bacillus kwashiorkori* SIT6[T] with other phylogenetically close *Bacillus* strains[a]

Properties	*B. kwashiorkori*	*B. thermoamylovorans*	*B. thermolactis*	*B. smithii*	*B. aquimaris*	*B. acidicola*	*B. sporothermodurans*	*B. alveayuensis*	*B. coagulans*
Cell diameter (μm)	0.2–3.0	0.45–4.0	0.6–10	0.5–1.1	0.45–1	3.1–5.9	NA	0.7–4.5	0.6–1
Oxygen requirement	Facultative Anaerobic	Facultative Anaerobic	Facultative Anaerobic	Facultative Anaerobic	Aerobic	Facultative Anaerobic	Aerobic	Facultative Anaerobic	Facultative Anaerobic
Gram stain	+	+	+	+	v	+	+	+	+
Motility	+	–	–	+	+	NA	+	+	+
Endospore formation	+	+	+	+	+	+	+	+	+
Indole	–	–	–	–	NA	–	–	–	–
Production of									
Catalase	+	+	+	+	+	NA	+	NA	NA
Oxidase	+	–	+	+	–	–	+	–	NA
Nitrate reductase	–	+	+	–	+	NA	–	–	+
Urease	–	–	–	–	–	NA	–	–	–
D-Galactosidase	+	+	–	–	+	NA	NA	–	+
N-acetyl-glucosamine	–	–	–	–	–	NA	NA	–	+
N-acetyl-β-glucosaminidase	–	NA	NA	+	NA	NA	NA	NA	NA
Alpha-mannosidase	+	NA	NA	NA	NA	NA	NA	NA	+
Acid from									
L-Arabinose	–	–	+	+	NA	–	NA	–	+
Glycerol	–	–	–	+	–	+	NA	+	NA
Erythritol	–	NA	NA	+	NA	NA	NA	–	NA
D-Ribose	–	+	+	+	NA	+	NA	–	–
D-Xylose	–	–	+	+	–	+	–	–	–
D-Adonitol	–	NA	NA	–	NA	NA	NA	–	NA
Methyl-β-D-xylopyranose	–	NA	NA	–	NA	NA	NA	–	NA
L-Sorbose	–	+	NA	–	NA	–	NA	NA	+
L-Rhamnose	–	–	NA	+	NA	–	NA	NA	–
Dulcitol	–	NA	NA	–	NA	NA	NA	NA	+
Inositol	–	NA	NA	+	NA	NA	NA	NA	NA
D-Sorbitol	–	NA	NA	+	NA	–	NA	NA	–
Esculin	+	NA	NA	+	NA	NA	NA	NA	NA
D-Raffinose	+	–	–	NA	–	NA	–	–	NA

(Continues)

TABLE 3 (Continued)

Properties	*B. kwashiorkori*	*B. thermoamylovorans*	*B. thermolactis*	*B. smithii*	*B. aquimaris*	*B. acidicola*	*B. sporothermodurans*	*B. alveayuensis*	*B. coagulans*
D-Lyxose	–	NA	NA	–	NA	NA	NA	NA	NA
D-Fucose	–	NA	NA	–	NA	NA	NA	NA	–
L-Arabitol	–	NA	NA	NA	NA	–	NA	NA	–
D-Mannose	+	–	–	+	–	+	–	+	+
D-mannitol	+	–	v	+	–	+	–	–	–
D-Glucose	+	+	–	+	+	+	+	+	+
D-Fructose	+	+	–	+	+	+	–	+	+
D-Maltose	+	–	+	+	+	+	+	–	+
D-Lactose	–	–	–	+	–	v	–	–	+
Galactose	–	–	–	+	–	+	–	+	+
Habitat	Human gut	Oil	Milk	Milk	Sea	Acidic sphagnum	Water, milk	Sea	Milk

NA, not available; v, variable.

[a]*Bacillus thermoamylovorans* DKP[T] (Combet-Blanc et al., 1995), *Bacillus thermolactis* R-6488[T] (Coorevits et al., 2011), *Bacillus smithii* NRRL NRS-173[T] (Bae, Lee, & Kim, 2005), *Bacillus aquimaris* TF-12[T] (Yoon, Kim, Kang, Oh, & Park, 2003), *Bacillus sporothermodurans* M215[T] (Heyndrickx et al., 2012), *Bacillus acidicola* 105-2[T] (Albert, 2005), *Bacillus alveayuensis* TM1[T] (Bae et al., 2005), and *Bacillus coagulans* 2-6[T] (De Clerck et al., 2004).

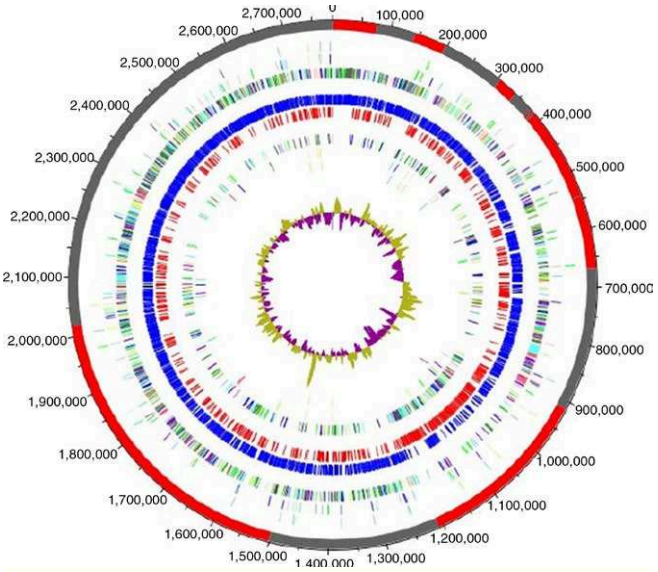


FIGURE 6 Graphical circular map of the chromosome. From outside to the center: Genes on the forward strand colored by clusters of orthologous groups (COG) categories (only gene assigned to COG), genes on the reverse strand colored by COG categories (only gene assigned to COG), RNA genes (tRNAs green, rRNAs red), G+C content and G+C skew

TABLE 4 Nucleotide content and gene count levels of the genome

Attribute	Value	% of total[a]
Genome size (bp)	2,784,637	100
DNA coding region (bp)	2,319,082	83.28
DNA G+C content (bp)	980,134	35.19
Total genes	2,646	100
RNA genes	74	2.79
tRNA genes	63	2.38
Protein-coding genes	2,572	97.20
Genes with function prediction	1,749	68.00
Genes assigned to COGs	1,715	66.68
Protein associated with hypothetical protein	487	18.93
Protein associated with ORFan	156	6.06
Genes with peptide signals	250	9.72
Genes with transmembrane helices	720	27.99
Genes associated with PKS or NRPS	6	0.2
Genes associated with mobilome	1,402	54.51
Genes associated with toxin/antitoxin	84	3.26

[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.

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

Code	Value	%[a]	Description
J	130	5.05	Translation
A	0	0	RNA processing and modification
K	141	5.48	Transcription
L	140	5.44	Replication, recombination, and repair
B	1	0.04	Chromatin structure and dynamics
D	18	0.70	Cell cycle control, mitosis, and meiosis
Y	0	0	Nuclear structure
V	43	1.67	Defense mechanisms
T	88	3.42	Signal transduction mechanisms
M	81	3.15	Cell wall/membrane biogenesis
N	26	1.01	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	26	1.01	Intracellular trafficking and secretion
O	87	3.38	Posttranslational modification, protein turnover, chaperones
C	120	4.67	Energy production and conversion
G	137	5.33	Carbohydrate transport and metabolism
E	138	5.37	Amino acid transport and metabolism
F	39	1.52	Nucleotide transport and metabolism
H	52	2.02	Coenzyme transport and metabolism
I	67	2.61	Lipid transport and metabolism
P	147	5.72	Inorganic ion transport and metabolism
Q	34	1.32	Secondary metabolites biosynthesis, transport, and catabolism
R	241	9.37	General function prediction only
S	180	6.99	Function unknown
–	857	33.32	Not in COGs

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

TABLE 6 Genomic comparison of *Bacillus kwashiorkori* with other *Bacillus* spp.

Species	Strain	Genome accession number	Genome size (Mb)	GC (%)	Gene content
B. kwashiorkori	SIT6^T	CTDX00000000	2.78	35.19	2,572
B. alveayuensis	TM1^T	JYCE00000000	6.70	38.13	6,689
B. shackletonii	LMG 18435^T	LJJC00000000	5.29	36.70	4,727
B. coagulans	2-6^T	CP003056	3.07	47.29	2,971
B. ginsengihumi	Gsoil 114^T	JAGM00000000	3.92	35.85	3,832
B. firmus	IAM 12464^T	BCUY00000000	4.97	41.45	4,922
B. aquimaris	TF-12^T	LQXM00000000	4.42	44.57	4,432
B. sporothermodurans	M215^T	LQYN00000000	4.04	35.65	4,211
B. smithii	NRRL NRS-173^T	BCVY00000000	3.38	40.75	3,619
B. acidicola	105-2^T	LWJG00000000	5.13	39.39	4,876
B. thermoamylovorans	DKP^T	CCRF00000000	3.70	37.27	3,441

or branched (iso and anteiso). The fatty acid composition of strain SIT6^T is detailed in Table 2.

Table 3 shows the biochemical features of *B. kwashiorkori* SIT6^T and the most closely related species.

Bacterial cells were resistant to metronidazole, but susceptible to imipenem, doxycycline, rifampicin, vancomycin, amoxicillin, ceftriaxone, gentamicin, trimethoprim/sulfamethoxazole, erythromycin, ciprofloxacin, and gentamicin.

3.3 | Genome properties

The genome is 2,784,637 bp long with 35.19% G+C content (Figure 6). It is composed of 16 scaffolds, composed of 16 contigs. Of the 2,646 predicted genes, 2,572 were protein-coding genes, and 74 were RNAs (7 genes are 5S rRNA, 2 genes are 16S rRNA, 2 genes are 23S rRNA, and 63 genes are tRNA). A total of 1,749 (68%) were assigned as putative function (by COGs of NR blast). A total of 156 genes were identified as ORFans (6.07%). The remaining genes were annotated as hypothetical proteins (487 genes [18.93%]). Genome content is detailed in Table 4, while Table 5 presents the distribution of the genes into COG functional categories.

The genome sequence has been deposited in GenBank under accession number CTDX00000000.

3.4 | Comparison with other *Bacillus* spp. genomes

The draft genome of *B. kwashiorkori* (2.78 Mb) is smaller in size than those of *Bacillus alveayuensis*, *Bacillus shackletonii*, *Bacillus coagulans*, *Bacillus ginsengihumi*, *Bacillus firmus*, *Bacillus aquimaris*, *Bacillus sporothermodurans*, *Bacillus smithii*, *Bacillus acidicola*, and *Bacillus thermoamylovorans* (6.70, 5.29, 3.07, 3.92, 4.97, 4.42, 4.04, 3.38, 5.13, and 3.70 Mb, respectively) (Table 6). *Bacillus kwashiorkori* has a lower G+C content (35.19%) than those of *B. alveayuensis*, *B. shackletonii*, *B. coagulans*, *B. ginsengihumi*, *B. firmus*, *B. aquimaris*, *B. sporothermodurans*, *B. smithii*, *B. acidicola*, and *B. sporothermodurans* (38.13%, 36.70%, 47.29%, 35.85%, 41.45%, 44.57%, 35.65%, 40.75%, 39.39%, and 37.27%, respectively) (Table 6). The protein content of *B. kwashiorkori* (2,572) is lower than those of *B. alveayuensis*, *B. shackletonii*, *B. coagulans*, *B. ginsengihumi*, *B. firmus*, *B. aquimaris*, *B. sporothermodurans*, *B. smithii*, *B. acidicola*, and *B. thermoamylovorans* (6,689, 4,727, 2,971, 3,832, 4,922, 4,432, 4,211, 3,619, 4,876, and 3,441, respectively) (Table 6). However, the distribution of genes into COG categories is similar in all compared genomes (Figure 7). In addition, AGIOS values ranged from 54.77% to 67.06% among the *Bacillus* species compared (Table 7). The range of AGIOS varied from 54.77% to 65.50% between *B. kwashiorkori* and other compared *Bacillus* species (Table 7). Moreover, *B. kwashiorkori* shares 455, 500, 340, 375, 541, 490, 461, 283, 451, and 476 orthologous genes with *B. alveayuensis*, *B. shackletonii*, *B. coagulans*, *B. ginsengihumi*, *B. firmus*, *B. aquimaris*, *B. sporothermodurans*, *B. smithii*, *B. acidicola*, and *B. thermoamylovorans*, respectively (Table 7). Of the species with standing in nomenclature, ANI values ranged from 66.46% between *B. coagulans* and *B. aquimaris* to 72.53% between *B. sporothermodurans* and *Bacillus shackletonii*. When comparing *B. kwashiorkori* to other species, the ANI value ranged from 66.74% between *B. kwashiorkori* and *B. coagulans* to 69.92% between *B. kwashiorkori* and *B. thermoamylovorans* (Table 8). The low ANI values confirmed it as a new species because ANI values bigger than 95 indicated that strains belong to the same species (Konstantinidis, Ramette, & Tiedje, 2006). Finally, digital DNA–DNA hybridization (dDDH) estimation of the strain SIT6^T against the compared genomes confirmed its new species status, as it ranges between 18.6 and 38.3 (below the cutoff of 70%).

4 | CONCLUSION

Based on the phenotypic properties (Table 2), phylogenetic tree (Figure 1), MALDI-TOF analyses (Figure 3), and genomic comparison (taxonogenomics [Table 6 and Table 7] and GGDC results), we propose the creation of *B. kwashiorkori* sp. nov. represented by the strain SIT6^T.

4.1 | Description of *B. kwashiorkori* sp. nov

Bacillus kwashiorkori (kwa.shi.or.ko'ri. L. adj. masc., in reference to Kwashiorkor) species are gram-positive, facultative aerobic, short rods, 1.8–5.9 μm in size, and motile by means of peritrichous flagella and sporulating. Colonies are creamy white, circular, 4–5 mm in diameter after cultivation at 37°C for 24 hr on 5% sheep blood-enriched Colombia agar. Growth occurs at temperatures in the range of 25–56°C (optimally at 37°C). It is catalase and oxidase positive. Concerning the biochemical characteristics, the API 50CH strip showed positive reactions for D-glucose, D-fructose, D-mannose, arbutin, esculin ferric citrate, salicin, D-maltose, saccharose, D-trehalose, melezitose, D-raffinose,

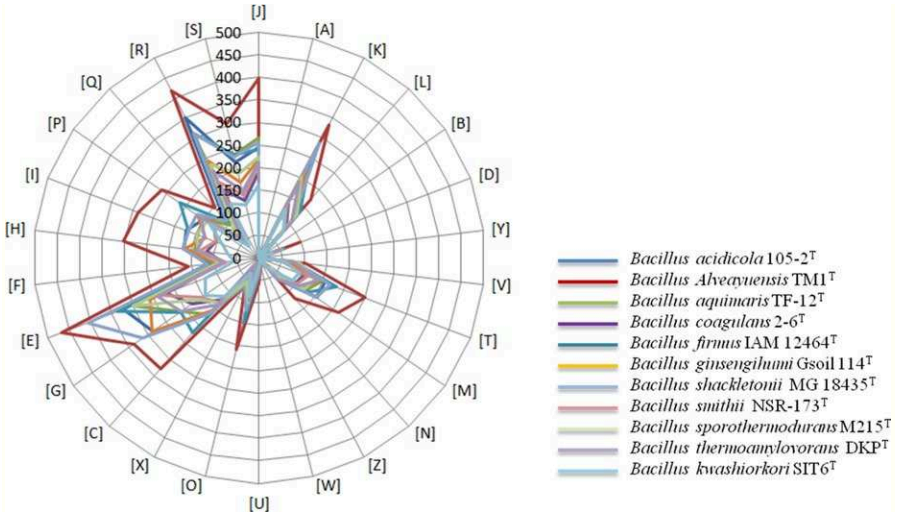


FIGURE 7 Distribution of functional classes of predicted genes of *Bacillus kwashiorkori* SIT6^T with 10 members of *Bacillus* genus

TABLE 7 Number of orthologous proteins shared between genomes (upper right triangle), average percentage similarity of nucleotides corresponding to orthologous protein shared between genomes (lower left triangle), and number of proteins per genome (bold numbers)

	B. kwashiorkori	*B. firmus*	*B. shackletonii*	*B. smithii*	*B. aquimaris*	*B. thermoamylovorans*	*B. coagulans*	*B. alveayuensis*	*B. sporothermodurans*	*B. acidicola*	*B. ginsengihumi*
B. kwashiorkori	**2,572**	541	500	283	490	476	340	455	461	451	375
B. firmus	56.84	**4,922**	928	552	984	642	529	839	837	887	700
B. shackletonii	59.33	56.57	**4,727**	592	905	656	628	736	1026	1073	838
B. smithii	55.15	59.30	55.92	**3,619**	510	457	466	525	530	603	616
B. aquimaris	62.32	58.30	55.88	56.28	**4,432**	574	508	755	779	848	640
B. thermoamylovorans	58.24	58.30	56.48	60.21	55.62	**3,441**	526	583	574	577	596
B. coagulans	54.77	58.14	54.96	59.33	56.18	58.45	**2,971**	484	553	616	676
B. alveayuensis	65.14	57.97	57.07	57.05	63.75	57.46	56.32	**6,689**	639	693	604
B. sporothermodurans	58.97	57.07	60.75	58.46	57.02	59.13	55.77	58.53	**4,211**	871	718
B. acidicola	63.97	57.72	58.77	57.15	63.91	57.09	57.12	63.93	59.39	**4876**	862
B. ginsengihumi	65.50	57.61	59.44	57.66	63.01	58.01	58.03	65.22	60.56	67.06	**3832**

TABLE 8 Average nucleotide identity (ANI) pairwise comparisons among sequenced species in the *Bacillus* genus

	B. kwashiorkori	*B. firmus*	*B. shackletonii*	*B. smithii*	*B. aquimaris*	*B. thermoamylovorans*	*B. coagulans*	*B. alveayuensis*	*B. sporothermodurans*	*B. acidicola*	*B. ginsengihumi*
B. kwashiorkori	100	68.67	68.52	67.94	67.41	69.92	66.74	68.17	68.67	67.64	68.28
B. firmus		100	68.07	67.68	68.02	66.74	67.52	67.90	68.75	68.70	67.43
B. shackletonii			100	68.67	68.69	67.68	67.95	68.24	72.53	70.61	70.06
B. smithii				100	67.95	67.95	68.52	68.91	68.41	68.41	68.50
B. aquimaris					100	66.77	66.46	67.96	68.88	68.31	68.17
B. thermoamylovorans						100	67.87	67.81	68.16	67.01	67.80
B. coagulans							100	66.84	67.69	68.46	69.39
B. alveayuensis								100	68.44	68.34	68.33
B. sporothermodurans									100	69.87	70.47
B. acidicola										100	69.06
B. ginsengihumi											100

ANI values are in percentages. Strains with ANI values over 95% are considered to belong to the same species.

and amidon. Negative reactions were recorded for glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl-β-D-xylopyranoside, D-galactose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-α-D-mannopyranoside, methyl-α-D glucopyranoside, *N*-acetyl-glucosamine, D-cellobiose, D-trehalose, inulin, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate, and potassium 5-ketogluconate.

Using API 20NE, a positive reaction was obtained for D-maltose, D-glucose, D-mannitol, and esculin ferric citrate. But potassium nitrate, L-tryptophan, L-arginine, urea, nitrophenyl β-D-galactopyranoside, L-arabinose, *N*-acetyl-glucosamine, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate, and phenylacetic acid were not assimilated.

API ZYM showed positive reactions for alkaline phosphatase, esterase (C4), cystine aminopeptidase, chymotrypsin, acid phosphatase, phosphoamidase, galactosidase, and mannosidase, but negative results for esterase lipase (C8), lipase (C14), leucine aminopeptidase, valine aminopeptidase, trypsin, α-glucuronidase, glucosaminidase, and α-fucosidase.

The major fatty acids are saturated species (15:0 iso, 16:0 and 17:0 anteiso). Strain SIT6^T was resistant to metronidazole, but susceptible to imipenem, doxycycline, rifampicin, vancomycin, amoxicillin, ceftriaxone, gentamicin, trimethoprim/sulfamethoxazole, erythromycin, ciprofloxacin, and gentamicin.

The G+C content of the genome is 35.19%. The 16S rRNA gene sequence and whole-genome shotgun sequence of *B. kwashiorkori* SIT6^T are deposited in GenBank under accession numbers LK985393 and CTDX00000000, respectively. The strain SIT6^T (=CSUR P2452^T, =DSM 29059^T) was isolated from the fecal flora of a Nigerian 4-month-old child suffering from acute malnutrition (kwashiorkor). Habitat is the human gut.

ACKNOWLEDGMENTS

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

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

ORCID

Saber Khelaifia http://orcid.org/0000-0002-9303-3893
Fadi Bittar http://orcid.org/0000-0003-4052-344X

REFERENCES

Albert, R.-A., Archambault, J., Rosselló-Mora, R., Tindall, B.-J., & Matheny, M. (2005). *Bacillus acidicola* sp. nov., a novel mesophilic, acidophilic species isolated from acidic *Sphagnum* peat bogs in Wisconsin. *International Journal of Systematic and Evolutionary Microbiology*, *55*, 2125–2130.

Angelakis, E., Armougom, F., Million, M., & Raoult, D. (2012). The relationship between gut microbiota and weight gain in humans. *Future Microbiology*, *7*, 91–109.

Auch, A. F., 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.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., … Osterman A. L. (2008). The RAST server: Rapid annotations using subsystems technology. *BMC Genomics*, *9*, 75.

Bae, S. S., Lee, J.-H., & Kim, S.-J. (2005). *Bacillus alveayuensis* sp. nov., a thermophilic bacterium isolated from deep-sea sediments of the Ayu Trough. *International Journal of Systematic and Evolutionary Microbiology*, *55*, 1211–1215.

Bittar, F., Keita, M. B., Lagier, J.-C., Peeters, M., Delaporte, E., & Raoult, D. (2014). *Gorilla gorilla gorilla* gut: A potential reservoir of pathogenic bacteria as revealed using culturomics and molecular tools. *Scientific Report*, *4*, 7174.

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* (Oxford, England), *28*, 464–469.

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

Chen, J., He, X., & Huang, J. (2014). Diet effects in gut microbiome and obesity. *Journal of food science*, *79*, 442–451.

Combet-Blanc, Y., Ollivier, B., Streicher, C., Patel, B. K. C., Dwivedi, P. P., Pot, B., … Garcia, J.-L. (1995). *Bacillus thermoamylovorans* sp. nov., a moderately thermophilic and amylolytic bacterium. *International Journal of Systematic and Evolutionary Microbiology*, *45*, 9–16.

Coorevits, A., Logan, N. A., Dinsdale, A. E., Halket, G., Scheldeman, P., Heyndrickx, M., … De Vos, P. (2011). *Bacillus thermolactis* sp. nov., isolated from dairy farms, and emended description of *Bacillus thermoamylovorans*. *International Journal of Systematic and Evolutionary Microbiology*, *61*, 1954–1961.

Darling, A.-C.-E., Mau, B., Blattner, F.-R., & Perna, N.-T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Research*, *14*, 1394–1403.

De Clerck, E., Rodriguez-Diaz, M., Forsyth, G., Lebbe, L., Logan, N. A., & DeVos, P. (2004). Polyphasic characterization of *Bacillus coagulans* strains, illustrating heterogeneity within this species, and emended description of the species. *Systematic and Applied Microbiology*, *27*, 50–60.

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.

Heyndrickx, M., Coorevits, A., Scheldeman, P., Lebbe, L., Schumann, P., Rodríguez-Diaz, M., … De Vos, P. (2012). Emended descriptions of *Bacillus sporothermodurans* and *Bacillus oleronius* with the inclusion of dairy farm isolates of both species. *International Journal of Systematic and Evolutionary Microbiology*, *62*, 307–314.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*, 111–120.

Kokcha, S., Ramasamy, D., Lagier, J.-C., Robert, C., Raoult, D., & Fournier, P.-E. (2012). Non-contiguous finished genome sequence and description of *Brevibacterium senegalense* sp. nov. *Standards in Genomic Sciences*, *7*, 233–245.

Konstantinidis, K., Ramette, A., & Tiedje, J.-M. (2006). The bacterial species definition in the genomic era. *Philosophical transactions Royal Society B Biological Sciences*, *361*, 1929–1940.

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov

model: Application to complete genomes. *Journal of Molecular Biology*, *305*, 567–580.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*, 1870–1874.

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

Lagier, J.-C., Elkarkouri, K., Rivet, R., Couderc, C., & Raoult, D. (2013). Non contiguous-finished genome sequence and description of *Senegalemassilia anaerobia* gen. nov., sp. nov. *Standards in Genomic Sciences*, *7*, 343–356.

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

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.

Matuschek, E., Brown, D.-F.-J., & Kahlmeter, G. (2014). Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. *Clinical Microbiology and Infection*, *20*, O255–O266.

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.

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

Moreno-Indias, I., Cardona, F., Tinahones, F. J., & Queipo-Ortuño, M. I. (2014). Impact of the gut microbiota on the development of obesity and type 2 diabetes mellitus. *Frontiers in microbiology*, *5*, 190.

Nielsen, H., Engelbrecht, J., Brunak, S., & von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering*, *10*, 1–6.

Ouk, K.-Y., Chun, J., Lee, I., & Park, S.-C. (2016). OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *International Journal of Systematic and Evolutionary Microbiology*, *66*, 1100–1103.

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.

Rossi-Tamisier, M., Benamar, S., Raoult, D., & Fournier, P.-E. (2015). Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. *International Journal of Systematic and Evolutionary Microbiology*, *65*, 1929–1934.

Sasser, M. (2006). *Bacterial Identification by Gas Chromatographic Analysis of Fatty Acids Methyl Esters (GC-FAME)*. Technical Note 101. Newark, DE: MIDI Inc.

Seck, E., Sankar, S. A., Khelaifia, S., Croce, O., Robert, C., Couderc, C., … Fournier, P. (2016). Non contiguous-finished genome sequence and description of *Planococcus massiliensis* sp. nov., a moderately halophilic bacterium isolated from the human gut. *New Microbes and New Infections*, *10*, 36–46.

Seng, P., Abat, C., Rolain, M., Colson, P., Lagier, J., & Gouriet, F. (2013). Laboratory: Impact of matrix-assisted laser desorption ionization – time of flight mass spectrometry. *Journal of clinical microbiology*, *51*, 2182–2194.

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.

Tindall, B.-J., Rosselló-Mora, R., Busse, H.-J., Ludwig, W., & Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *International Journal of Systematic and Evolutionary Microbiology*, *60*, 249–266.

Vartoukian, S.-R., Palmer, R.-M., & Wade, W.-G. (2010). Strategies for culture of 'unculturable' bacteria. *FEMS microbiology letters*, *309*, 1–7.

Weisburg, W.-G., Barns, S.-M., Pelletier, D.-A., & Lane, D.-J. (1991). 16S Ribosomal DNA amplification for phylogenetic study. *Journal of bacteriology*, *173*, 697–703.

Welker, M., & Moore, E.-R.-B. (2011). Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. *Systematic and applied microbiology*, *34*, 2–11.

Yoon, S.-H., Ha, S.-M., Kwon, S., Lim, J., Kim, Y., Seo, H., & Chun, J. (2017). Introducing EzBioCloud: A taxonomically united database of 16S rRNA and whole genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*, *67*, 1613–1617.

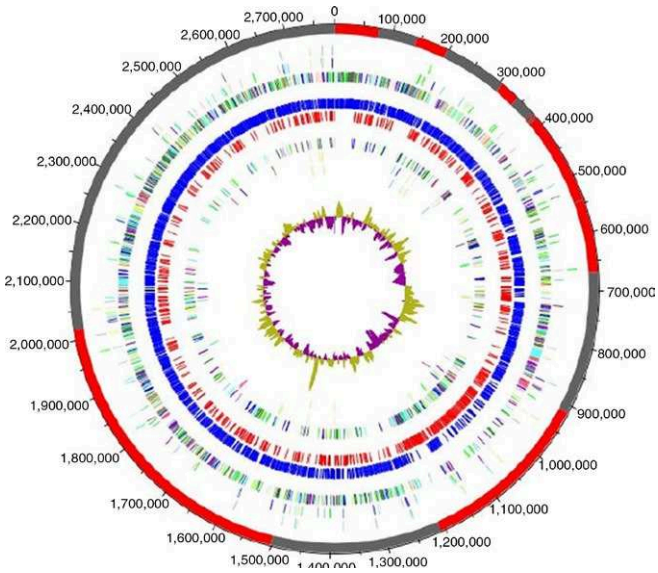
Yoon, J.-H., Kim, I.-G., Kang, K. H., Oh, T.-K., & Park, Y.-H. (2003). *Bacillus marisflavi* sp. nov. and *Bacillus aquimaris* sp. nov., isolated from sea water of a tidal flat of the Yellow Sea in Korea. *International Journal of Systematic and Evolutionary Microbiology*, *53*, 1297–1303.

Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., & Wishart, D. S. (2011). PHAST: A fast phage search tool. *Nucleic Acids Research*, *39*, W347–W352.

How to cite this article: Seck EH, Beye M, Traore SI, et al. *Bacillus kwashiorkori* sp. nov., a new bacterial species isolated from a malnourished child using culturomics. *MicrobiologyOpen*. 2017;e535. https://doi.org/10.1002/mbo3.535

Graphical Abstract

The contents of this page will be used as part of the graphical abstract of html only.
It will not be published as part of main article.



Here we describe the phenotypic characteristics, MALDI-TOF MS profile, the complete genome sequence, and genomic analyses (taxonogenomics) for the strain SIT6^{T} (the type strain for the new proposed species *Bacillus kwashiorkori*). This new species was isolated from the stool sample of a Nigerian 4-month-old child suffering from acute malnutrition (kwashiorkor) using culturomics.

Tessaracoccus massiliensis sp. nov., a new bacterial species isolated from the human gut

E. Seck[1], S. I. Traore[1], S. Khelaifia[1], M. Beye[1], C. Michelle[1], C. Couderc[1], S. Brah[2], P.-E. Fournier[1], D. Raoult[1,3] and G. Dubourg[1]

1) *Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, INSERM 1095, Faculté de médecine, Marseille, France,* 2) *Hôpital National de Niamey, Niamey, Niger and* 3) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

A new *Actinobacterium*, designated *Tessaracoccus massiliensis* type strain SIT-7^T (= CSUR P1301 = DSM 29060), have been isolated from a Nigerian child with kwashiorkor. It is a facultative aerobic, Gram positive, rod shaped, non spore-forming, and non motile bacterium. Here, we describe the genomic and phenotypic characteristics of this isolate. Its 3,212,234 bp long genome (1 chromosome, no plasmid) exhibits a G+C content of 67.81% and contains 3,058 protein-coding genes and 49 RNA genes.

Keywords: culturomics, genome, human gut, taxono-genomics, *Tessaracoccus massiliensis*
Original Submission: 23 February 2016; **Revised Submission:** 28 April 2016; **Accepted:** 3 May 2016
Article published online: 28 May 2016

Corresponding author: G. Dubourg, Aix-Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, INSERM 1095, Faculté de médecine, 27 Boulevard Jean Moulin, 13385 Marseille Cedex 05, France
E-mail: greg.dubourg@gmail.com

Introduction

Kwashiorkor is a nutritional disorder most often seen in low-income and middle-income countries [1]. It is a virulent form of severe acute malnutrition caused by a lack of protein intake. People suffering from kwashiorkor typically have an extremely emaciated appearance in all body parts except their ankles, feet and belly, which swell with fluid. Clinical manifestations include, generalized oedema, anorexia and skin ulcerations [2]. The gastrointestinal microbiota also appears to be involved. To elucidate this hypothesis we have studied by 'culturomics' the human digestive microbiota of children suffering from kwashiorkor [3,4]. In this work we isolated a wide range of bacteria, including several new species including our isolate, the strain SIT-7^T.

Each newly isolated bacterium in this work is described by using a new and innovative method that we have implemented. Indeed, due to the availability of genomic data through the development of new tools for the sequencing of DNA [5], we introduced a new way of describing the novel bacterial species [6]. This includes, among other features, their genomic [7–11] and proteomic information obtained by matrix-assisted laser desorption-ionization time-of-flight (MALDI-TOF-MS) analysis [12]. This concept differs from the usual methods to define a new species, which are based on parameters such as sequencing of the 16S rRNA, phylogeny, G+C content, and DNA–DNA hybridization. These methods are time consuming and involve numerous limitations [13,14].

Here, we present a brief classification and a set of features for strain SIT-7^T (= CSUR P1301 = DSM 29060) and we also describe the sequence of the complete genome and its annotation. We named this new isolate *Tessaracoccus massiliensis*.

The genus *Tessaracoccus*, belonging to the family *Propionibacteriaceae*, was first described by Maszenan *et al.* in 1999 [15] and contains five species with validly published names at the time of writing: *Tessaracoccus bendigoensis*, *Tessaracoccus flavescens*, *Tessaracoccus lubricantis, Tessaracoccus oleiagri* and *Tessaracoccus lapidicaptus* [16–19]. All five species are Gram-positive and non-motile bacteria able to live in various different environments [18]. *Tessaracoccus massiliensis* is the first representative of the genus *Tessaracoccus* to be isolated in humans through 'culturomics'.

http://dx.doi.org/10.1016/j.nmni.2016.05.002

Materials and Methods

Sample informations

The culture sample comes from Niamey (Niger). It was collected from a 2-year-old child with kwashiorkor, a form of acute malnutrition.

Consent was obtained from the child's parents at the National Hospital of Niamey and the study was approved by the Institut Fédératif de Recherche 48 (Faculty of Medicine, Marseille, France), under agreement number 09-022. The patient did not receive antibiotics at the time of sample collection and the stool sample was collected in sterile plastic containers, divided into aliquots and stored at −80°C. The pH of the aliquots was also tested by pH meter (ThermoFisher scientific, Saint Aubin, France).

Strain identification and phylogenetic classification

The isolate was cultured in blood culture bottles supplemented with sheep blood, (BioMérieux, Marcy l'Etoile, France) during a 7-day pre-incubation period at 37°C in aerobic atmosphere.

The identification of the strain was carried out by MALDI-TOF mass spectrometric analysis with a Microflex (Bruker Daltonics, Leipzig, Germany), as previously described [19]. Obtained spectra were then compared with the Bruker database and our database, which is continuously updated. If no identification is possible at the genus or species level, sequencing of the 16S rRNA gene is used to achieve a correct identification [20]. DNA extraction was performed using EZ1 DNA Tissue Kit (Qiagen, Courtaboeuf, France). The DNA extract was amplified using PCR technology and universal primers FD1 and RP2 [21] (Eurogentec, Angers, France). The amplifications and sequencing of the amplified products were performed as previously described [22]. Then, 16S rRNA gene sequences were compared with those which are available in GenBank (http://www.ncbi.nlm.nih.gov/genbank/). Identification at the species level was defined by a 16S rRNA gene sequence similarity $\geq$99% with the sequence of the prototype strain in GenBank. When the percentage of identity is <98.7%, the studied strain is considered as a new species [11].

Phylogenetic analysis based on 16S rRNA of our isolate was performed to identify its phylogenetic affiliations with other near isolates, including other members of the genus *Tessaracoccus*. The MEGA 6 (Molecular Evolutionary Genetics Analysis) software allowed us to construct a phylogenetic tree [23]. Sequence alignment of the different species was performed using CLUSTAL W [24,25] and the evolutionary distance was calculated with the Kimura two-parameter model [26].

Physiological and phenotypic characteristics

The phenotypic characteristics of this strain were studied by testing different temperatures and growth atmospheres. With regards to the temperature, we studied the growth at 25, 30, 37, 45 and 56°C. The test was conducted under aerobic conditions in the presence of 5% CO_2 and also in anaerobic and microaerophilic atmospheres, created by using AnaeroGen™ (ThermoFisher Scientific, Saint Aubin, France) and CampyGen™ (ThermoFisher scientific) respectively. Gram staining and motility were observed by using a light microscope DM1000 (Leica Microsystems, Nanterre, France). Cell morphology was examined after negative staining of bacteria, with a G20 Tecnai transmission electron microscope (FEI Company, Limeil-Brévannes, France). Formation of spores was determined after thermal shock and observed under a microscope.

Biochemical characteristics of our strain were studied using the API 20NE, ZYM and 50 CH strips (BioMérieux) according to manufacturer's instructions. Antibiotic susceptibility was determined on Mueller–Hinton agar (BioMérieux) using a disc diffusion method (i2A, Montpellier, France). The following antibiotics were tested: doxycycline (30 µg), rifampicin (30 µg), vancomycin (30 µg), erythromycin (15 µg), ampicillin (10 µg), ceftriaxone (30 µg), ciprofloxacin (5 µg), gentamicin (500 µg), penicillin (10 µg), trimethoprim/sulfamethoxazole (1, ;25 + 23.75 µg), imipenem (10 µg), and metronidazole (4 µg), clindamycin (15 µg), colistin (50 µg) and oxacillin (5 µg).

Genomic sequencing

Genomic DNA (gDNA) of *T. massiliensis* was sequenced using MiSeq Technology (Illumina Inc, San Diego, CA, USA) with the mate pair strategy. The gDNA was barcoded to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). The mate pair library was prepared with 1 µg of gDNA using the Nextera mate pair Illumina guide. The gDNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The 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 kb up to 11 kb with an optimal size of 4.008 kb. No size selection was performed and 388.3 ng of tagged fragments were circularized. The circularized DNA was mechanically sheared to small fragments, with an optimal size of 634 bp, on the Covaris device S2 in microtubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc.) and the final concentration library was measured at 35.59 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 × 251-bp. Total information of 10.6 Gb was obtained from a 1326 K/mm^2 cluster density with a cluster passing quality control filters of 99.1% (24 492 260 clusters). Within this run, the index representation for *T. massiliensis* was determined to be 7.06%. The 1 481 197 paired reads were filtered according to the read qualities. These reads were trimmed, and then assembled using the CLC genomics WB4 software.

Genome annotation and comparison

Open reading frames (ORFs) were predicted using PRODIGAL [27] with default parameters but the predicted ORFs were excluded if they spanned a sequencing gap region. The predicted bacterial protein sequences were searched against the GenBank database [28] and the Clusters of Orthologous Groups (COG) database using BLASTP (E-value 1e^{-03}, coverage 0.7 and identity percent 30%). If no hit was found, a search against the N R database using BLASTP was performed, with an E-value of 1e^{-03}, coverage 0.7 and identity 30%. If the sequence length was <80 amino acids, we used an E-value of 1e^{-05}. The tRNASCANSE tool [29] was used to find tRNA genes, whereas ribosomal RNAs were found using RNAMMER [30] and BLASTN against the GenBank database. Lipoprotein signal peptides and the number of transmembrane helices were predicted using PHOBIUS [31]. ORFans were identified if their BLASTP E-value was lower than 1e^{-03} for alignment length >80 amino acids. If alignment lengths were <80 amino acids, we used an E-value of 1e^{-05}. Such parameter thresholds have already been used in previous works to define ORFans. ARTEMIS [32] was used for data management and DNA PLOTTER [33] for visualization of genomic features. The MAUVE alignment tool (version 2.3.1) was used for multiple genomic sequence alignment [34]. To estimate the mean level of nucleotide sequence similarity at the genome level, we used the MAGI home-made software to calculate the average genomic identity of gene sequences among compared genomes. Briefly, this software combines the PROTEINORTHO software [35] for detecting orthologous proteins in pairwise genomic comparisons, then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman–Wunsch global alignment algorithm. Annotation and comparison processes were performed in the MULTI-AGENT software system DAGOBAH [36], that include FIGENIX [37] libraries that provide pipeline analysis. Genomes

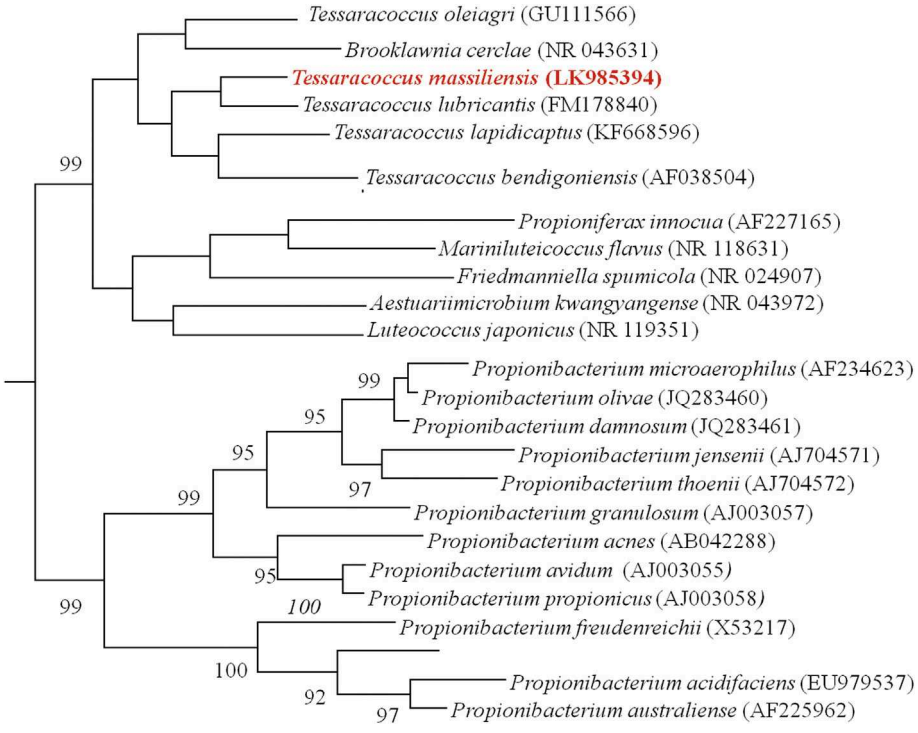


FIG. 1. Phylogenetic tree highlighting the position of relative to other close species. Phylogenetic tree showing the position of *Tessaracoccus massiliensis* strain SIT-7T (underlined) relative to other phylogenetically close members of the family *Porphyromonadaceae*. GenBank Accession numbers are indicated in parentheses. Sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained using the maximum-likelihood method within the MEGA software. Numbers at the nodes are percentages of bootstrap values obtained by repeating the analysis 500 times to generate a majority consensus tree. Only values >90% were displayed. The scale bar indicates a 2% nucleotide sequence divergence

from the genus *Tessaracoccus* and closely related genera were used for the calculation of average genomic identity of gene sequences values. We compared the genome sequence of *T. massiliensis* strain SIT-7^T with those of *Propionibacterium propionicum* F0230a (CP002734), *Aestuariimicrobium kwangyangense* DSM 21549 (ATXE01000000), *Propionibacterium jensenii* DSM 20535 (AUDD00000000), *Propionibacterium freudenreichii* subsp. *shermanii* CIRM-BIA1 (FN806773), *Propionibacterium avidum* 44067 (CP005287), *Propionibacterium acnes* KPA171202 (AE017283) and *Propionibacterium thoenii DSM 20276*

TABLE 1. Classification and general features of *Tessaracoccus massiliensis* strain SIT-7^T

Property	Term
Current classification	Domain: *Bacteria* Phylum: *Actinobacteria* Class: *Actinobacteria* Order: *Propionibacteriales* Family: *Propionibacteriaceae* Genus: *Tessaracoccus* Species: *Tessaracoccus massiliensis* Type strain: SIT-7^T
Gram stain	Positive
Cell shape	Rod-shaped
Motility	Non-motile
Sporulation	Non-endospore forming
Temperature range	Mesophile
Optimum temperature	37°C
pH	pH 6 to 9
Optimum pH	7.5
Salinity	0.0–5.0 g/L
Optimum salinity	0
Oxygen requirement	Facultative aerobic

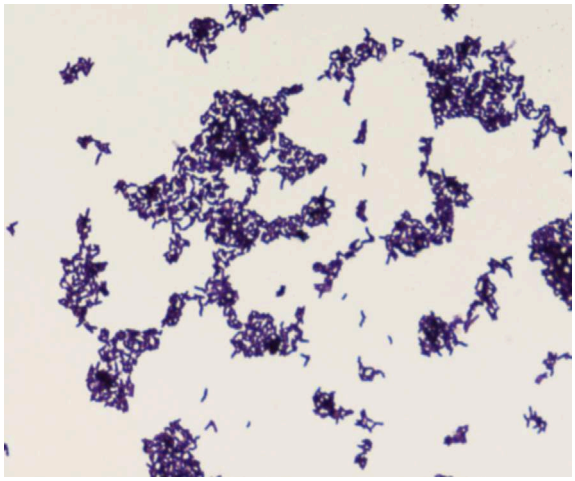


FIG. 3. Gram staining of *Tessaracoccus massiliensis* strain SIT-7^T.

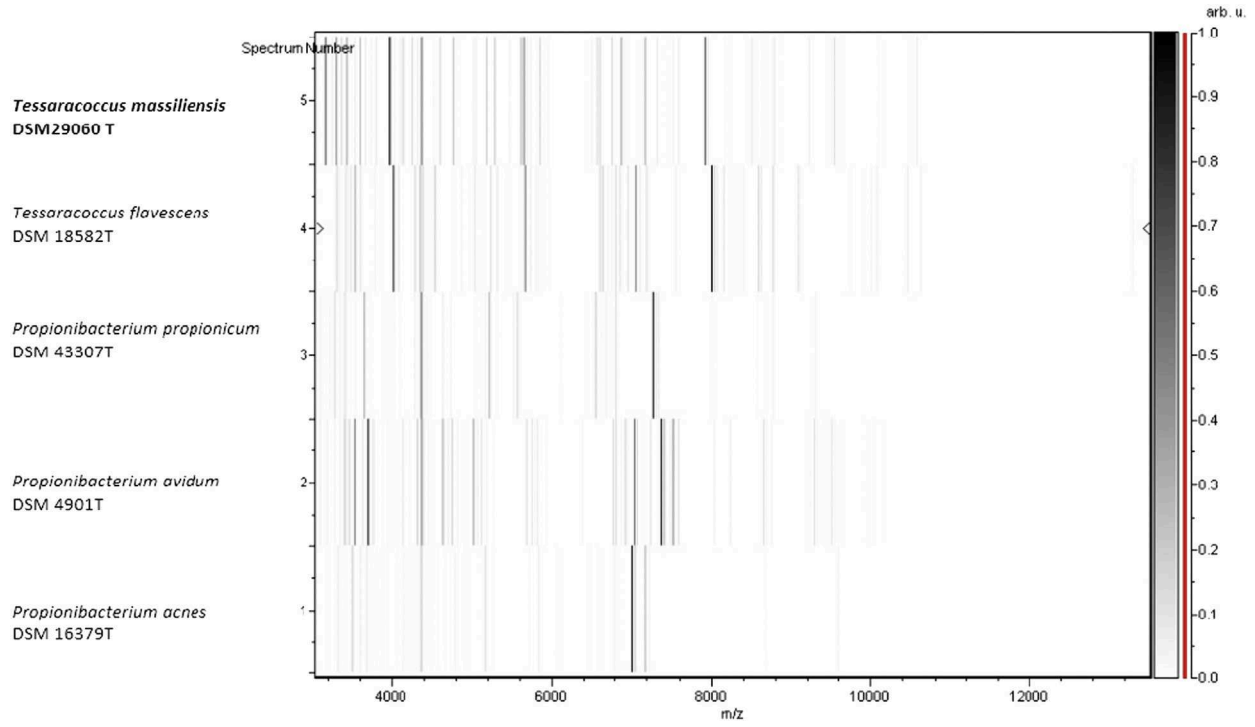


FIG. 2. Gel view comparing *Tessaracoccus massiliensis* strain SIT-7^T to the members of the genera *Tessaracoccus* and *Propionibacterium*. The gel view displays the raw spectra of all loaded spectrum files arranged in a pseudo-gel like look. The *x*-axis records the m/z value. The left *y*-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a Grey-scale scheme code. The colour bar and the right *y*-axis indicate the relation between the colour a peak is displayed with and the peak intensity in arbitrary units. Displayed species are indicated on the left.

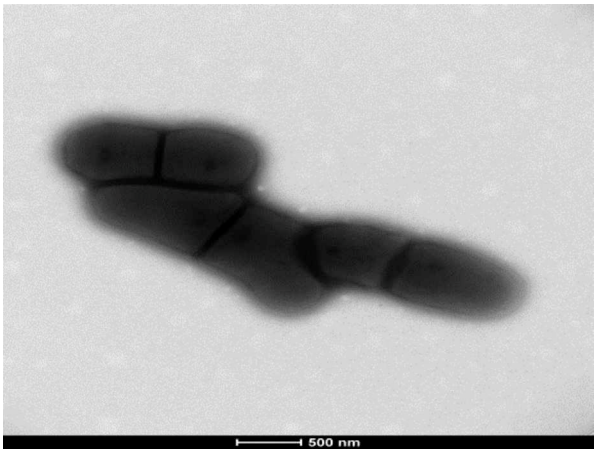


FIG. 4. Transmission electron microscopy of *Tessaracoccus massiliensis* strain SIT-7^T. Cells are observed on a Tecnai G20 transmission electron microscope operated at 200 keV. The scale bar represents 500 nm.

(KE384018) *Tessaracoccus oleiagri* CGMCC 1.9159 (PRJNA303532) . The genomic similarity was evaluated among closely studied species from *T. massiliensis* by the DNA–DNA hybridization. *Tessaracoccus massiliensis* was locally aligned 2-by-2 using BLAT algorithm [38,39] against each selected genome previously cited and DNA–DNA hybridization values were estimated from a generalized model [40].

Results

Phylogenic classification

Strain SIT-7^T was first isolated in April 2014 by incubation in a blood-culture bottle with sheep blood and cultivation on 5% sheep blood-enriched Colombia agar in aerobic conditions after a 7-day pre-incubation period at 37°C. MALDI-TOF applied on colonies did not give a reliable identification for the strain SIT-

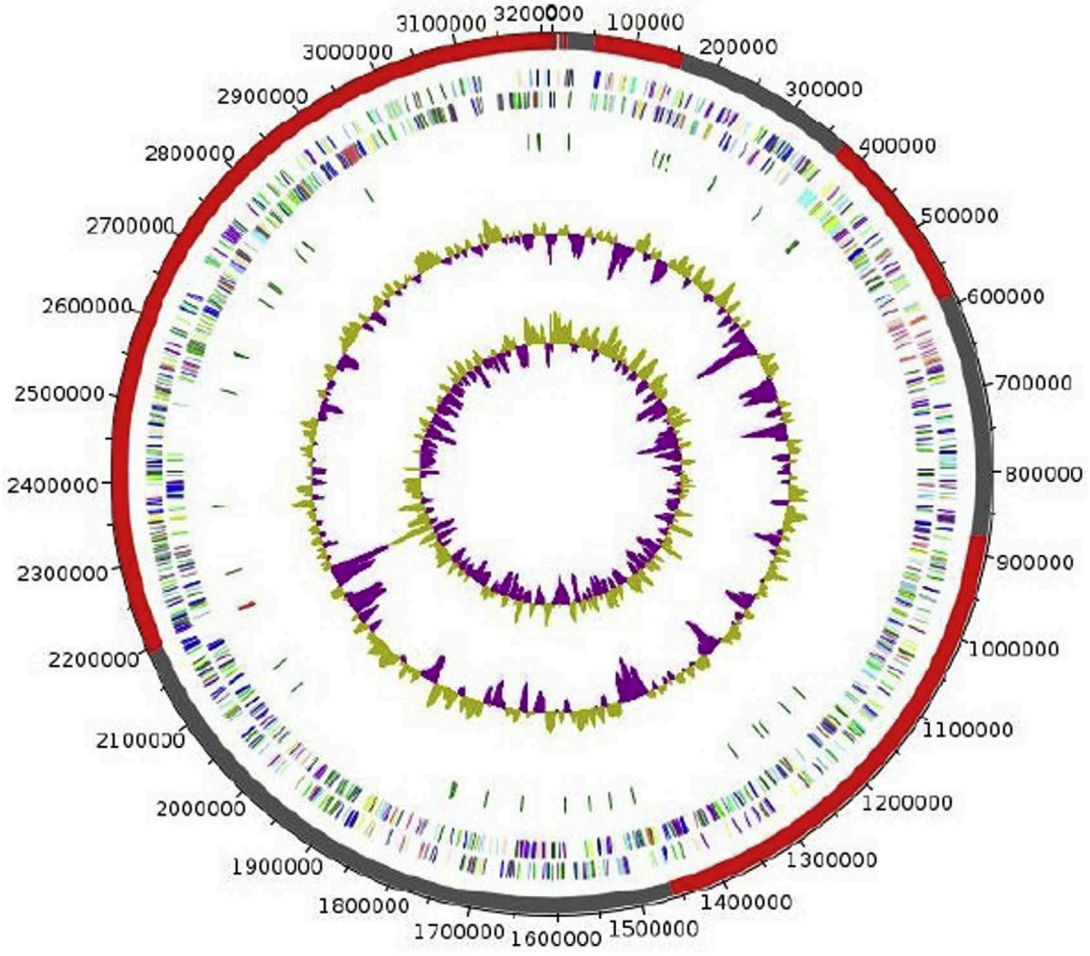


FIG. 5. Graphical circular map of the *Tessaracoccus massiliensis* strain SIT-7^T chromosome. From the outside to center: the outer two circles show open reading frames oriented in the forward (coloured by COG categories) and reverse (coloured by COG categories) directions, respectively. The third circle marks the tRNA genes (green). The fourth circle shows the G+C% content plot. The inner-most circle shows GC skew, purple indicating negative values whereas olive is for positive values.

TABLE 2. Differential characteristics of *Tessaracoccus massiliensis* strain SIT-7^T and other strains

Properties	*T. massiliensis*	*T. lubricantis*	*T. bendigoensis*	*T. oleiagri*	*T. flavescens*	*T. lapidicaptus*	*P. avidum*	*P. propionicum*	*P. acnes*
Cell diameter (µm)	0.6–0.9	0.5–2	0.5–1.1	0.48–1	0.6–1.2	0.45–1	0.25–0.75	0.2–0.8	0.4–0.5
Indole (tryptophanase)	–	–	–	–	–	na	–	+	+
Production of									
Catalase	+	+	+	+	+	+	na	–	+
Oxidase	–	–	–	–	–	–	na	na	na
Nitrate reductase	+	+	+	+	+	+	–	–	na
Urease	–	–	–	–	–	–	na	na	na
β-galactosidase	+	–	+	+	+	+	–	na	na
N-acetyl-glucosamine	+	na	–	na	+	–	+	na	na
Acid from									
L-Arabinose	+	–	+	+	+	–	na	na	na
D-mannose	+	+	–	+	–	+	+	+	+
D-mannitol	–	–	+	–	–	+	+	+	+
D-glucose	+	+	+	+	–	+	+	+	+
D-fructose	+	+	+	na	+	+	+	+	+
D-maltose	–	+	–	na	+	–	+	+	+
D-lactose	–	+	–	+	+	+	+	–	–
Habitat	human gut	metal-working fluid	activated sludge biomass	saline soil	saline soil	Iberian pyrite belt	chlorosolvent	human mouth	skin

Differential characteristics compared with *Tessaracoccus bendigoensis* strain Ben 106T, *Tessaracoccus flavescens* strain SST-39T, *Tessaracoccus lapidicaptus* strain IPBSL-7(T), *Tessaracoccus lubricantis* strain KSS-17Se(T), *Tessaracoccus oleiagri* strain SL014B-20A1(T), *Propionibacterium avidum strain* ATCC 25577, *Propionibacterium propionicum* strain ATCC 14157, *Propionibacterium acnes* strain ATCC 6919.
+, positive; –, negative; na, not available.

7^T. Phylogenetic analysis based on 16S rRNA showed that this strain has a 97.5% nucleotide sequence similarity with *T. lubricantis*, the phylogenetically closest *Tessaracoccus* species with a valid published name (Fig. 1). However, this percentage remains lower than the threshold to delineate a new species, fixed at 98.7% by Stackebrandt and Ebers [13]. This allows us to report the strain as a new species within the genus of *Tessaracoccus* (Table 1). The 16S rRNA sequence of *T. massiliensis* was deposited in EMBL-EBI under accession number LK985394. We consequently added the spectrum of strain SIT-7^T to our MALDI-TOF database, allowing further correct identification. We also performed a gel view to compare spectra between available species, which highlights spectral differences with *T. flavescens* (Fig. 2).

Phenotypic description

After 48 h of growth on 5% sheep blood-enriched Colombia agar at 37°C at pH 7.5 in aerobic conditions, the surface colonies were circular, greyish, shiny and smooth with a diameter of 0.7–1.1 mm. *Tessaracoccus massiliensis* is Gram positive (Fig. 3), non-motile and non-spore-forming. Growth is observed at temperatures ranging from 25 to 45°C, with an optimum at 37°C. Strain SIT-7^T grew in an aerobic atmosphere, in the presence of 5% CO_2, and also in anaerobic and microaerophilic atmospheres. Growth was also achieved after 48 h on trypticase soy agar. Under the microscope, cells are rod-shaped and measure 4–5 µm in length and 0.6 µm in diameter (Fig. 4). The strain is catalase-positive and oxidase-negative.

Using the commercially available API 20NE strip, *T. massiliensis* strain SIT-7^T demonstrates positive reactions for potassium nitrate, D-mannose, L-arabinose, D-glucose, but negative for L-arginine, urea, D-maltose, malic acid and L-tryptophan. Using an API ZYM strip, positive reactions were observed for esterase(C4), leucine arylamidase, α-chymotrypsin, alkaline phosphatase, cystine arylamidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase, α-mannosidase and β-galactosidase.

Substrate oxidation and assimilation were examined with an API 50CH strip at 37°C. Positive reactions were obtained for D-mannose, D-arabinose, D-galactose, D-ribose, D-sucrose, D-fructose and D-glucose. We compared some of these phenotypic characteristics with the most closely related species (Table 2). By comparison with *T. lubricantis*, its phylogenetically closest neighbour, *T. massiliensis* differed in β-galactosidase, acid formation from L-arabinose, D-maltose and D-lactose.

Our strain SIT-7^T was susceptible to doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole and imipenem, but resistant to metronidazole.

TABLE 3. Nucleotide content and gene count levels of the genome

Attribute	Value	% of total[a]
Size (bp)	3 212 234	100
G+C content (bp)	2 177 721	67.81
Coding region (bp)	3 129 675	83.34
Total genes	3107	100
RNA genes	49	1.57
Protein-coding genes	3058	98.46
Genes with function prediction	384	12.55
Genes assigned to clusters of orthologous groups	1770	57.88
Genes with peptide signals	409	13.57
Gene associated to resistance genes	4	0.13
Gene associated to bacteriocin genes	13	0.42
Protein without homologue (orfan)	405	13.24

[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.

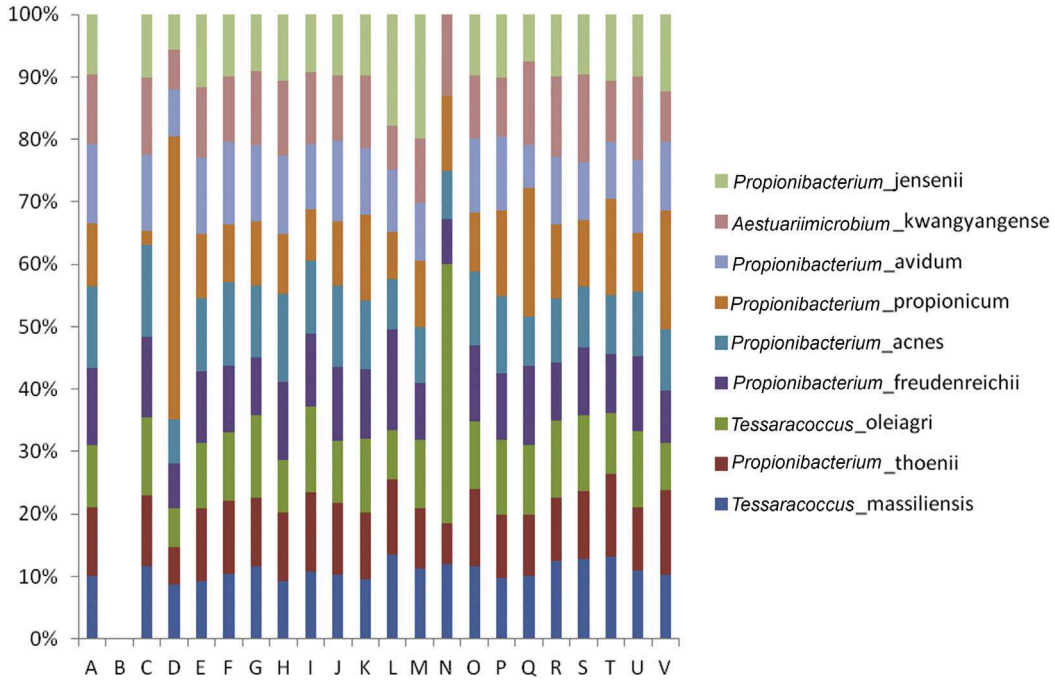


FIG. 6. Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins.

Genome properties

The genome is 3 212 234 bp long with 67.81% G+C content (Table 3 and Fig. 5). It is composed of 17 scaffolds (composed of 20 contigs). On the 3107 predicted genes, 3058 were protein-coding genes, and 49 were RNAs (one gene is 5S rRNA, one gene is 16S rRNA, one gene is 23S rRNA, 46 genes are tRNA genes). A total of 2042 genes (66.78%) were assigned as putative function (by COGs or by NR BLAST). In all, 405 genes were identified as ORFans (13.24%). The remaining genes (499 genes; 16.32%) were annotated as hypothetical proteins. The details of statistics of the genome are presented in Table 3. Table 4 and Fig. 6 shows the proportion of genes divided in the different COGs functional categories. The genome sequence has been deposited in EMBL-EBI under accession number CCYJ00000000.

TABLE 4. Number of genes associated with the 25 general clusters of orthologous group functional categories

Code	Value	% value[a]	Description
J	144	4.7	Translation
A	1	0.03	RNA processing and modification
K	126	4.12	Transcription
L	184	6.01	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	17	0.55	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	46	1.50	Defence mechanisms
T	62	2.02	Signal transduction mechanisms
M	81	2.64	Cell wall/membrane biogenesis
N	0	0	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	22	0.71	Intracellular trafficking and secretion
O	69	2.25	Post-translational modification, protein turnover,chaperones
C	133	4.34	Energy production and conversion
G	205	6.70	Carbohydrate transport and metabolism
E	169	5.52	Amino acid transport and metabolism
F	65	2.12	Nucleotide transport and metabolism
H	87	2.84	Coenzyme transport and metabolism
I	61	1.99	Lipid transport and metabolism
P	113	3.695	Inorganic ion transport and metabolism
Q	37	1.20	Secondary metabolites biosynthesis, transport and catabolism
R	238	7.78	General function prediction only
S	112	3.66	Function unknown
_	1288	42.11	Not in clusters of orthologous groups

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

Genome comparison

The draft genome sequence of *T. massiliensis* is smaller than those of *P. propionicum* F0230a (CP002734), (3.21 and 3.45 Mb respectively), but larger than *P. avidum* 44067 (CP005287), *P. acnes* KPA171202 (AE017283), *P. freudenreichii* subsp. *shermanii* CIRM-BIA1 (FN806773), *P. thoenii* DSM 20276 (KE384018), *A. kwangyangense* DSM 21549 (ATXE01000000), *P. jensenii* DSM 20535 (AUDD00000000) and *T. oleiagri* CGMCC 1.9159 (PRJNA303532) (2.52; 2.56; 2.61; 2.93, 2.99; 3.02; 3.15 MB, respectively).

The G+C content of *T. massiliensis* is smaller than those *of P. thoenii*, *A. kwangyangense*, *P. jensenii*, *T. oleiagri* (68%, 68.5%, 68.7%, 69.7%, respectively), but larger than those of *P. freudenreichii* subsp. *shermanii*, *P. avidum*, *P. propionicum* and *P. acnes* (67.27%, 63.48%, 66.06% and 60.01%, respectively).

TABLE 5. **The numbers of orthologous proteins shared between genomes (upper right)**[a]

	Aestuariimicrobium kwangyangense	*Tessaracoccus massiliensis*	*Propionibacterium freudenreichii*	*Propionibacterium thoenii*	*Propionibacterium acnes*	*Propionibacterium propionicum*	*Propionibacterium avidum*	*Tessaracoccus oleiagri*	*Propionibacterium jensenii*
A. kwangyangense	**2633**	1231	1007	1196	1123	1082	1155	1252	1197
T. massiliensis	69.97	**3052**	1030	1180	1111	1249	1133	1382	1170
P. freudenreichii	69.129	68.65	**2318**	1075	1027	1013	1041	1003	1090
P. thoenii	69.47	68.54	69.4	**2629**	1280	1113	1326	1163	1497
P. acnes	66.09	65.63	65.96	70.18	**2373**	1084	1411	1092	1305
P. propionicum	69.4	72.3	67.93	68.2	65.35	**3064**	1087	1193	1122
P. avidum	68.19	67.68	67.77	72.49	81.27	67.06	**2302**	1108	1348
T. oleiagri	70.13	74.56	68.73	68.62	65.57	72.01	67.63	**2916**	1155
P. jensenii	69.94	68.96	69.53	83.34	70.4	68.57	72.92	69.21	**2700**

[a]Average percentage similarity of nucleotides corresponding to orthologous proteins shared between genomes (lower left) and numbers of proteins per genome (bold).

TABLE 6. **Pairwise comparison of *Tessaracoccus massiliensis* with other species using GGDC, formula 2 (DNA–DNA hybridization estimates based on identities / HSP length)**[a]

	Tessaracoccus massiliensis	*Propionibacterium avidum*	*Tessaracoccus oleiagri*	*Propionibacterium jensenii*	*Propionibacterium thoenii*	*Propionibacterium freudenreichii*	*Propionibacterium propionicum*	*Propionibacterium acnes*	*Aestuariimicrobium kwangyangense*
Tessaracoccus massiliensis	100.00% [100%–100%]	13.50% [11.2%–16.3%]	19.30% [17.1%–21.7%]	19.20% [17%–21.6%]	20.20% [18%–22.6%]	19.10% [16.9%–21.5%]	19.40% [17.2%–21.8%]	20.30% [18.1%–22.7%]	14.00% [11.2%–17.3%]
Propionibacterium avidum		100.00% [100%–100%]	19.90% [17.7%–22.3%]	21.40% [19.1%–23.8%]	15.50% [12.6%–18.9%]	19.70% [17.5%–22.1%]	13.00% [10.3%–16.3%]	23.50% [21.2%–25.9%]	19.80% [17.6%–22.2%]
Tessaracoccus oleiagri			100.00% [100%–100%]	13.60% [10.8%–16.9%]	19.40% [17.2%–21.8%]	19.30% [17.1%–21.7%]	19.50% [17.3%–21.9%]	20.40% [18.2%–22.8%]	19.60% [17.4%–22%]
Propionibacterium jensenii				100.00% [100%–100%]	27.80% [25.4%–30.3%]	14.10% [11.3%–17.4%]	21.40% [19.1%–23.8%]	20.10% [17.9%–22.5%]	18.90% [16.8%–21.3%]
Propionibacterium thoenii					100.00% [100%–100%]	14.10% [11.3%–17.4%]	20.30% [18.1%–22.8%]	20.60% [18.4%–23%]	18.80% [16.7%–21.2%]
Propionibacterium freudenreichii						100.00% [100%–100%]	19.30% [17.1%–21.7%]	20.90% [18.7%–23.4%]	19.70% [17.5%–22.1%]
Propionibacterium propionicum							100.00% [100%–100%]	12.80% [10.1%–16%]	19.60% [17.4%–22%]
Propionibacterium acnes								100.00% [100%–100%]	12.90% [10.2%–16.2%]
Aestuariimicrobium kwangyangense									100.00% [100%–100%]

[a]The confidence intervals indicate the inherent uncertainty in estimating DNA–DNA hybridization 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. 2) and phylogenomic analyses as well as the GGDC results.

The gene content of *T. massiliensis* is smaller than those *P. propionicum*, but larger than those of *P. freudenreichii* subsp. *shermanii*, *P. avidum*, *P. acnes*, *A. kwangyangense*, *P. thoenii*, *P. jensenii* and *T. oleiagri* (2439, 2242, 2938, 2966, 2684, 2687 and 2754, respectively).

Tessaracoccus massiliensis also shared 69.97, 1030, 1111, 1133, 1170, 1180, 1249, 1382, orthologous genes with, *A. kwangyangense*, *P. freudenreichii*, *P. acnes*, *P. avidum*, *P. jensenii*, *P. thoenii*, *P. propionicum* and *T. oleiagri*, respectively (Table 5).

The DNA–DNA hybridization was of 13.50% (11.2%–16.3%) with *P. avidum*, 14.00% (11.2%–17.3%) with *A. kwangyangense*, 19.10% (16.9%–21.5%) with *P. freudenreichii*, 19.20% (17%–21.6%) with *P. jensenii*, 19.30% (17.1%–21.7%) with *T. oleiagri*, 19.40% (17.2%–21.8%) with *P. propionicum*, 20.20% (18%–22.6%) with *P. jensenii* and 20.30% (18.1%–22.7%) with *P. acnes* (Table 6). These data confirm *T. massiliensis* as a unique species.

Finally, we observed that all compared genomes have nearly the same number of genes in each of the COG categories (Table 4, Fig. 6).

Conclusions

Based on phenotypic, genomic and phylogenetic analyses, we formally propose the creation of *Tessaracoccus massiliensis* sp. nov, represented here by the strain SIT-7^T. The strain was isolated from a stool sample of a Nigerian child suffering from kwashiorkor, and represents the first *Tessaracoccus* species isolated in human, as in culturomics studies.

Description of *Tessaracoccus massiliensis* sp. nov.

Tessaracoccus massiliensis (mas.si.li.en'sis. L. masc. adj. *massiliensis* of Massilia, the old Roman name for Marseille, where the strain was isolated).

Facultative anaerobic, Gram-positive, oxidase-negative, catalase-positive, non-endospore-forming, non-motile rods. Colonies are circular, greyish, shiny and smooth with a 0.7–1.1 mm diameter on Columbia agar + 5% sheep blood. Good growth occurs at 25–45°C (optimum 37°C) and with a pH between 6 and 9 (optimum 7.5). Cells measure 4–5 µm in length and 0.6 µm in diameter.

Using API 50CH and ZYM strip, positive reactions were observed for D-mannose, D-arabinose, D-galactose, D-ribose, D-sucrose, D-fructose and D-glucose, esterase(C4), leucine arylamidase, α-chymotrypsin, alkaline phosphatase, cystine arylamidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase, α-mannosidase and β-galactosidase. API 20NE strip showed that *T. massiliensis* assimilates D-mannose, D-glucose and L-arabinose. Potassium nitrate is reduced. SIT-7^T was susceptible to doxycycline, rifampicin, vancomycin, nitrofurantoin, amoxicillin, erythromycin, ampicillin, ceftriaxone, ciprofloxacin, gentamicin, penicillin, trimethoprim/sulfamethoxazole and imipenem.

The G+C content of the genome is 67.81%. Accession numbers of the sequences of 16S rRNA and genome deposited in EMBL-EBI are LK985394 and CCYJ00000000, respectively. The habitat of the microorganism is the human digestive tract. The type strain SIT-7^T (= CSUR P1301, = DSM29060) was isolated from a stool specimen of a Nigerian child suffering from kwashiorkor.

Conflict of Interest

None declared.

Acknowledgements

The authors thank the Xegen Company (www.xegen.fr) for automating the genomic annotation process, Karolina Griffiths for reviewing the English and Claudia Andrieu for administrative assistance. This study was funded by the *Fondation Méditerranée Infection*.

References

[1] Manary MJ, Heikens GT, Golden M. Kwashiorkor: more hypothesis testing is needed to understand the aetiology of oedema. Malawi Med J 2009;21:106–7.
[2] Williams CD. Kwashiorkor: a nutritional disease of children associated with a maize diet. Bull WHO 2003;81:912–4.
[3] 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.
[4] 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.
[5] 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.
[6] Hugon P, Mishra AK, Lagier JC, Nguyen TT, Couderc C, Raoult D, et al. Non-contiguous finished genome sequence and description of *Brevibacillus massiliensis* sp. nov. Stand Genomic Sci 2013;8:1–14.
[7] Kokcha S, Ramasamy D, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Brevibacterium senegalense* sp. nov. Stand Genomic Sci 2012;7:233–45.

[8] Lagier JC, El Karkouri K, Nguyen TT, Armougom F, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Anaerococcus senegalensis* sp. nov. Stand Genomic Sci 2012;6:116–25.
[9] Lagier JC, Armougom F, Mishra AK, Nguyen TT, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Alistipes timonensis* sp. nov. Stand Genomic Sci 2012;6:315–24.
[10] Lagier JC, El karkouri K, Rivet R, Couderc C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Senegalemassilia anaerobia* gen. nov., sp. nov. Stand Genomic Sci 2013;7:343–56.
[11] Lagier JC, El Karkouri K, Mishra AK, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Enterobacter massiliensis* sp. nov. Stand Genomic Sci 2013;7:399–412.
[12] Welker M, Moore ER. Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 2011;34:2–11.
[13] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[14] Vandamme P, Pot B, Gillis M, de Vos P, Kersters K, Swings J. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Rev 1996;60:407–38.
[15] Maszenan AM, Seviour RJ, Patel BK, Schumann P, Rees GN. *Tessaracoccus bendigoensis* gen. nov., sp. nov., a gram-positive coccus occurring in regular packages or tetrads, isolated from activated sludge biomass. Int J Syst Bacteriol 1999;49:459–68.
[16] Lee DW, Lee SD. *Tessaracoccus flavescens* sp. nov., isolated from marine sediment. Int J Syst Evol Microbiol 2008;58:785–9.
[17] Kämpfer P, Lodders N, Warfolomeow I, Busse HJ. *Tessaracoccus lubricantis* sp. nov., isolated from a metalworking fluid. Int J Syst Evol Microbiol 2009;59:1545–9.
[18] Cai M, Wang L, Cai H, Li Y, Wang YN, Tang YQ, et al. *Salinarimonas ramus* sp. nov. and *Tessaracoccus oleiagri* sp. nov., isolated from a crude oil-contaminated saline soil. Int J Syst Evol Microbiol 2011;61:1767–75.
[19] Puente-Sanchez F, Sanchez-Roman M, Amils R, Parro V. *Tessaracoccus lapidicaptus* sp. nov., an actinobacterium isolated from the deep subsurface of the Iberian pyrite belt. Int J Syst Evol Microbiol 2014;64:3546–52.
[20] Dubourg G, Lagier JC, Armougom F, Robert C, Hamad I, Brouqui P, et al. The gut microbiota of a patient with resistant tuberculosis is more comprehensively studied by culturomics than by metagenomics. Eur J Clin Microbiol Infect Dis 2013;32:637–45.
[21] Ash C, Farrow JA, Dorsch M, Stackebrandt E, Collins MD. Comparative analysis of *Bacillus anthracis, Bacillus cereus*, and related species based on the basis of reverse transcriptase sequencing of 16S rRNA. Int J Syst Bacteriol 1991;41:343–6.
[22] Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol 1991;173:697–703.
[23] 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.
[24] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 2013;30:2725–9.
[25] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–80.
[26] Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980;16:111–20.
[27] 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.
[28] Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res 2012;40:D48–53.
[29] 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.
[30] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucl Acids Res 2007;35:3100–8.
[31] Käll L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 2004;338:1027–36.
[32] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–5.
[33] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 2009;25:119–20.
[34] Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004;14:1394–403.
[35] 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.
[36] 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, Heidelberg: Springer-Verlag; 2011. p. 71–87.
[37] 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.
[38] Kent WJ. BLAT–the BLAST-like alignment tool. Genome Res 2002;12:656–64.
[39] Auch AF, von Jan M, Klenk HP, Göker M. Digital DNA–DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2010;2:117–34.
[40] 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.

Conclusion

Le microbiome humain est le principal facteur instrumental à l'interface entre l'environnement et l'hôte humain. L'alimentation est le principal déterminant du microbiote intestinal. Les procaryotes vivant uniquement en milieu salé sont appelés halophiles, contrairement aux organismes halotolérants qui n'ont pas besoin de sel mais peuvent croître dans des conditions salines. Des phylotypes identiques ont été détectés dans l'intestin humain, le sel de table et dans des crevettes salées, ce qui suggère que ces aliments peuvent être considérés comme des probiotiques halophiles. Nous avons récemment confirmé que le sel de table est une source potentielle de bactéries halophiles par la description de son répertoire procaryotique. Compte tenu de notre longue histoire avec le sel et de notre consommation croissante d'aliments salés, il n'est pas surprenant de détecter des procaryotes halophiles chez l'humain. Dans ce travail, nous avons prouvé définitivement, par culturomics et par métagénomique, que les microorganismes halophiles colonisent l'intestin de l'homme et suggèrent qu'elles peuvent être parmi les commensaux de la flore intestinale. En effet nous avons découvert une diversité inattendue de procaryotes halophiles dont la richesse était corrélée avec les concentrations de sel fécales. Ainsi, le sel devrait être ajouté aux composants alimentaires capables d'altérer globalement et spécifiquement le microbiote intestinal humain. Il pourrait jouer un rôle critique dans les maladies humaines associées à une consommation excessive de sel, telles que le diabète et l'obésité. D'autres études devraient préciser tous les facteurs qui déterminent la salinité intestinale et le rôle des procaryotes halophiles dans l'intestin.

Par ailleurs, nous avons montré dans ce travail, que la plupart des pathogènes cités en microbiologie clinique ne se développent pas dans l'eau de mer. Cela s'explique par la salinité et d'autres facteurs tel que la dilution, la prédation, la température et les rayons solaires. Les bactéries non marines peuvent ainsi rester en vie dans la mer sous forme inactivée, perdant leur

susceptibilité à transmettre des maladies. La prévalence des infections associées à l'eau de mer semble exagérée. Les études de référence sur l'évaluation des risques infectieuses liées aux eaux marines comportent plusieurs biais non négligeables. Dans la littérature actuelle, il existe peu de données suggérant qu'on peut contracter des infections par ingestion direct de l'eau de mer. La plupart des infections attribuées à l'eau de mer sont dues à la consommation de fruits mer et autres produits marins. Chaque année quelques 800 millions de repas potentiellement contaminés, récoltés dans des eaux de mer polluées, sont consommés. Les infections bactériennes associées au milieu marin peuvent également être acquises par l'introduction de pathogènes marins dans les tissus traumatisés. Cette situation contraste avec la multiplicité des mesures de la qualité des eaux et les moyens démesurés mobilisés pour la surveillance des eaux marines. Nous suggérons des études supplémentaires prenant en compte tous les facteurs adjacents pour réévaluer le risque infectieux lié à la mer afin de revoir les programmes de surveillance des plages.

Références

1. Qin J, Li R, Raes J, Arumugam M, Burgdorf S, Manichanh C, et al. Europe PMC Funders Group Europe PMC Funders Author Manuscripts A human gut microbial gene catalog established by metagenomic sequencing. 2013;464(7285):59‑65.

2. Palmer C, Bik EM, DiGiulio DB, Relman D a, Brown PO. Development of the human infant intestinal microbiota. PLoS Biol. juill 2007;5(7):e177.

3. Turnbaugh PJ, Ley RE, Mahowald M a, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 21 déc 2006;444(7122):1027‑31.

4. Bäckhed F, Ding H, Wang T, Hooper L V, Koh GY, Nagy A, et al. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci U S A. 2 nov 2004;101(44):15718‑23.

5. Wang M, Ahrné S, Jeppsson B, Molin G. Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes. FEMS Microbiol Ecol. 1 oct 2005;54(2):219‑31.

6. Flint HJ. The impact of nutrition on the human microbiome. Nutr Rev. août 2012;70 Suppl 1:S10-3.

7. Oren A. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline Syst. janv 2008;4:2.

8. Oren A. Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. Front Microbiol. janv 2013;4(November):315.

9. Kennedy SP, Ng WV, Salzberg SL, Hood L, Dassarma S. Understanding the Adaptation of Halobacterium Species NRC-1 to Its Extreme Environment through Computational Analysis of Its Genome Sequence. 2001;(Eisenberg 1995):1641‑50.

10. Application T. Diversity of Halophilic Archaea in Fermented Foods and Human Intestines and Their Application. 2013;23(12):1645‑53.

11. Kivistö AT, Karp MT. Halophilic anaerobic fermentative bacteria. J Biotechnol. 10 avr 2011;152(4):114‑24.

12. Lagier J-C, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis. déc 2012;18(12):1185‑93.

13. Lagier J-C, Million M, Hugon P, Armougom F, Raoult D. Human gut microbiota: repertoire and variations. Front Cell Infect Microbiol. janv 2012;2(November):136.

14. Hugon P, Lagier J-C, Robert C, Lepolard C, Papazian L, Musso D, et al. Molecular studies neglect apparently gram-negative populations in the human gut microbiota. J Clin Microbiol. oct 2013;51(10):3286‑93.

15. Gill SR, Pop M, Deboy RT, Eckburg PB, Peter J, Samuel BS, et al. NIH Public Access. 2011;312(5778):1355‑9.

16. Nam Y-D, Chang H-W, Kim K-H, Roh SW, Kim M-S, Jung M-J, et al. Bacterial, archaeal, and eukaryal diversity in the intestines of Korean people. J Microbiol Seoul Korea. oct 2008;46(5):491‑501.

17. Oxley AP a, Lanfranconi MP, Würdemann D, Ott S, Schreiber S, McGenity TJ, et al. Halophilic archaea in the human intestinal mucosa. Environ Microbiol. sept 2010;12(9):2398‑410.

18. Vartoukian SR, Palmer RM, Wade WG. Strategies for culture of « unculturable » bacteria. FEMS Microbiol Lett. 1 août 2010;309(1):1‑7.

19. Ma Y, Galinski E a, Grant WD, Oren A, Ventosa A. Halophiles 2010: life in saline environments. Appl Environ Microbiol. déc 2010;76(21):6971‑81.

20. Sims JK, Enomoto PI, Frankel RI, Wong LM. Marine bacteria complicating seawater near-drowning and marine wounds: a hypothesis. Ann Emerg Med. 1983;12(4):212–216.

21. Carlucci AF, Pramer D. Factors affecting the survival of bacteria in sea water. Appl Microbiol. 1959;7(6):388–392.

22. Carlucci AF, Pramer D. An Evaluation of Factors Affecting the Survival of Escherichia coli in Sea Water: III. Antibiotics. Appl Microbiol. 1960;8(4):251.

23. Carlucci AF, Pramer D. An evaluation of factors affecting the survival of Escherichia coli in sea water II. Salinity, pH, and nutrients. Appl Microbiol. 1960;8(4):247–250.

24. Soller JA, Schoen ME, Bartrand T, Ravenscroft JE, Ashbolt NJ. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water Res. sept 2010;44(16):4674‑91.

25. Wade TJ, Sams E, Brenner KP, Haugland R, Chern E, Beach M, et al. Rapidly measured indicators of recreational water quality and swimming-associated illness at marine beaches: a prospective cohort study. Environ Health. 2010;9(1):66.

26. McFeters GA, Bissonnette GK, Jezeski JJ, Thomson CA, Stuart DG. Comparative Survival of Indicator Bacteria and Enteric Pathogens in Well Water. Appl Microbiol. 1 mai 1974;27(5):823‑9.

27. Gerba CP, Goyal SM, LaBelle RL, Cech I, Bodgan GF. Failure of indicator bacteria to reflect the occurrence of enteroviruses in marine waters. Am J Public Health. 1979;69(11):1116–1119.

28. Savichtcheva O, Okabe S. Alternative indicators of fecal pollution: Relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. Water Res. 1 juill 2006;40(13):2463‑76.

29. Ramasamy D, Mishra AK, Lagier J-C, Padhmanabhan R, Rossi M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol. 1 févr 2014;64(Pt 2):384‑91.

30. Elsawi Z, Togo AH, Beye M, Dubourg G, Andrieu C, Armsrtong N, et al. Hugonella massiliensis gen. nov., sp. nov., genome sequence, and description of a new strictly anaerobic bacterium isolated from the human gut. MicrobiologyOpen. août 2017;6(4).