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.
© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

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’.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
http://dx.doi.org/10.1016/j.nmni.2016.05.002

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 ≥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

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **13**, 3–12
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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

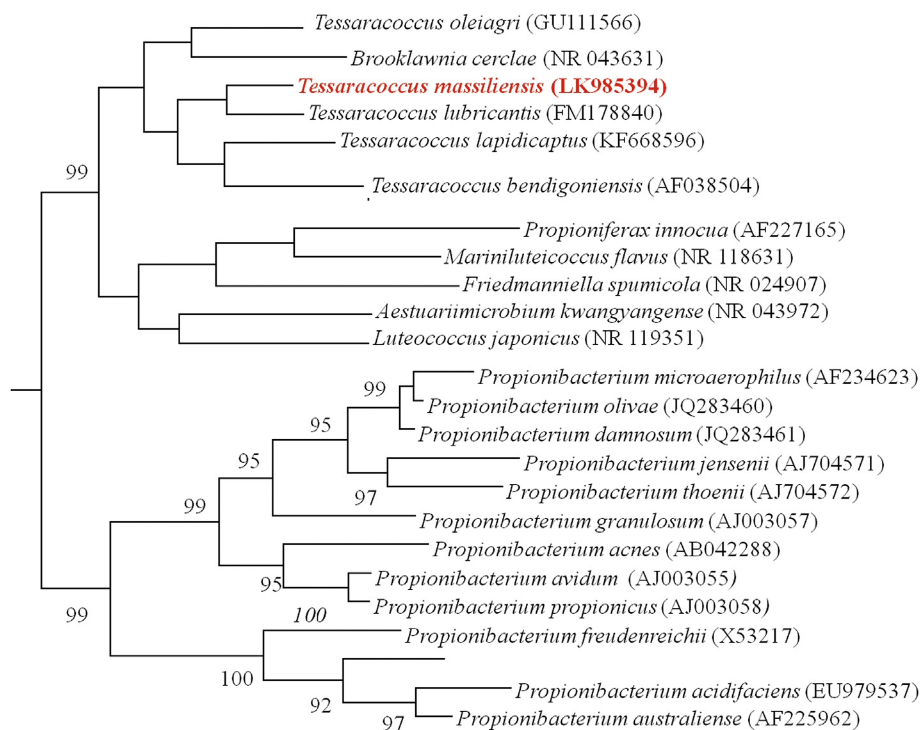

FIG. 1. Phylogenetic tree highlighting the position of relative to other close species. Phylogenetic tree showing the position of *Tessaracoccus massiliensis* strain SIT-7T (underlined) relative to other phylogenetically close members of the family *Porphyromonadaceae*. GenBank Accession numbers are 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

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **13**, 3–12
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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

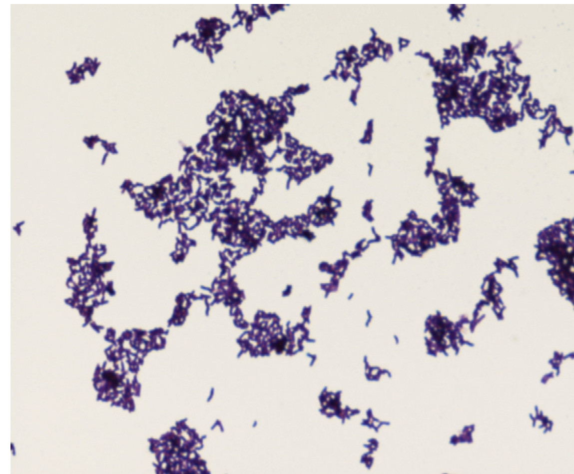

FIG. 3. Gram staining of *Tessaracoccus massiliensis* strain SIT-7^T.

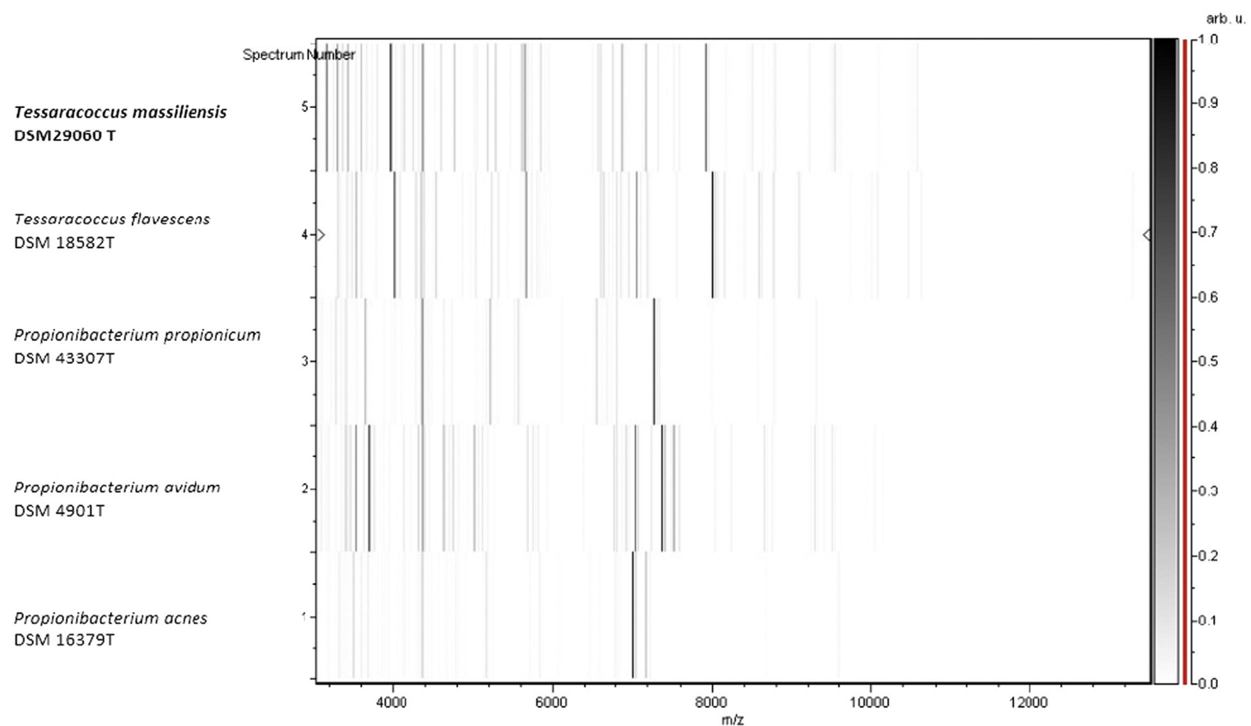

FIG. 2. Gel view comparing *Tessaracoccus massiliensis* strain SIT-7^T to the members of the genera *Tessaracoccus* and *Propionibacterium*. The gel view displays the raw spectra of all loaded spectrum files arranged in a pseudo-gel like look. The *x*-axis records the m/z value. The left *y*-axis displays 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.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **13**, 3–12
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

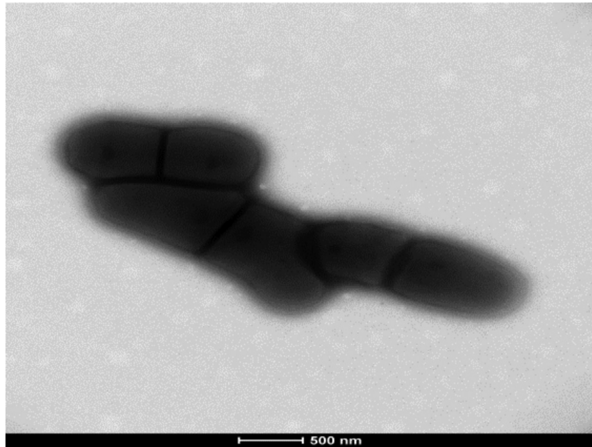

FIG. 4. Transmission electron microscopy of *Tessaracoccus massiliensis* strain SIT-7^{T}. Cells are observed on a Tecnai G20 transmission electron microscope operated at 200 keV. The scale bar represents 500 nm.

(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-

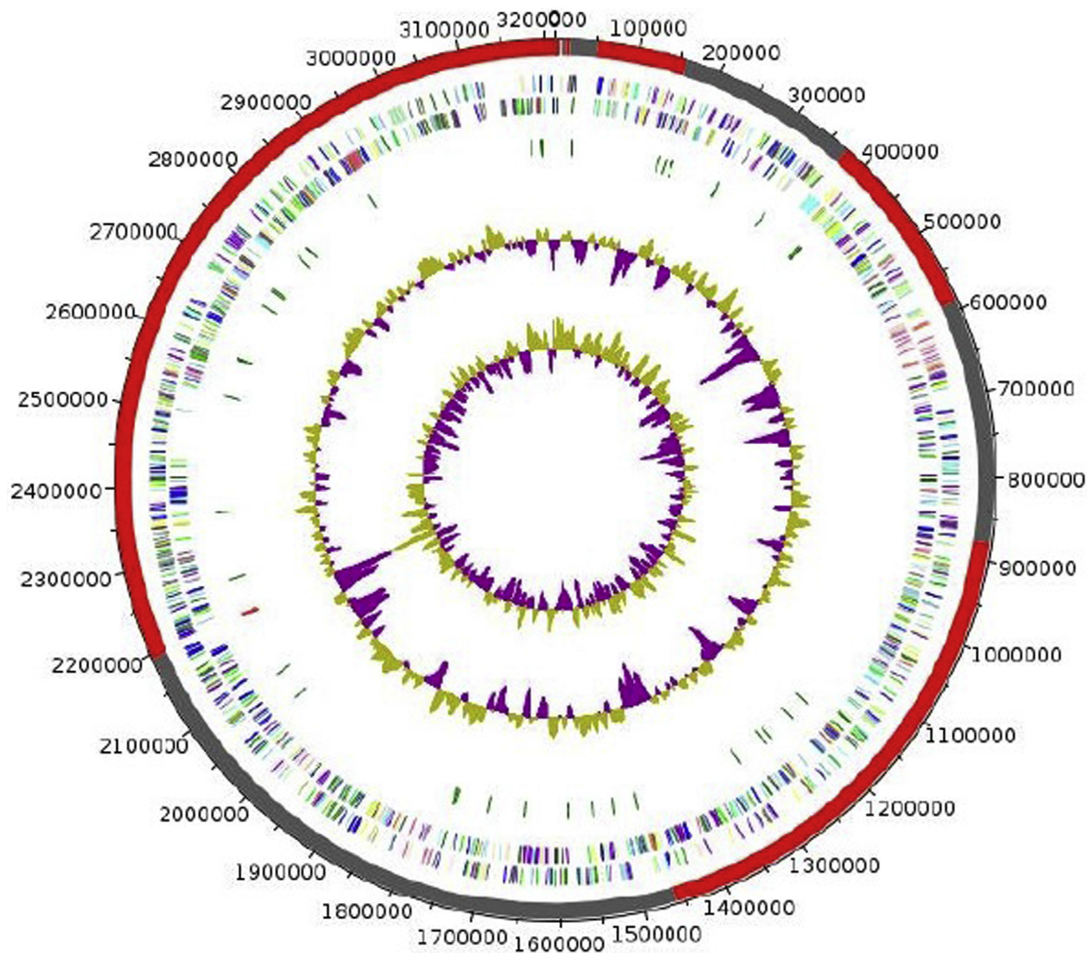

FIG. 5. Graphical circular map of the *Tessaracoccus massiliensis* strain SIT-7^{T} chromosome. From the outside to center: the outer two circles show open reading frames oriented in the forward (coloured by COG categories) and reverse (coloured by COG categories) directions, respectively. The third 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.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **13**, 3–12
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **13**, 3–12
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

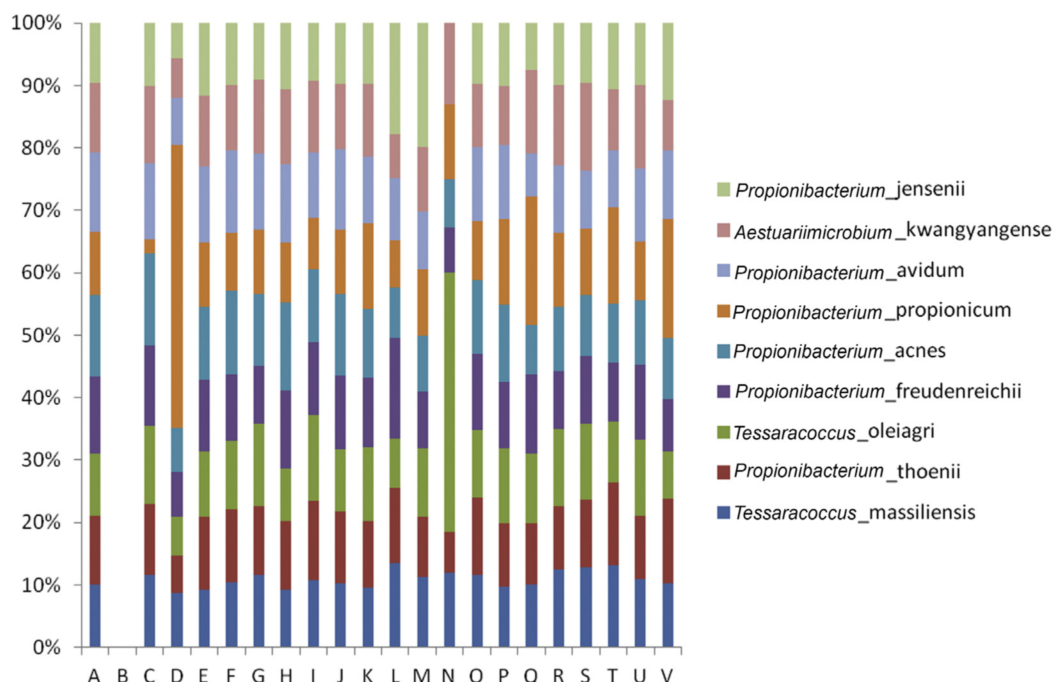

FIG. 6. Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins.

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).

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **13**, 3–12
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, NMNI, **13**, 3–12
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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 *Tessarococcus* 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.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **13**, 3–12
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

[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.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, *NMNI*, **13**, 3–12
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).