Genome sequence and description of *Mobilicoccus massiliensis* sp. nov. isolated from the stool of a Nigerian boy with kwashiorkor

N. Mathlouthi, S. I. Traore, T. Cimmino, S. Khelaifia, T. T. Nguyen, F. Cadoret, C. Couderc, D. Raoult and J.-M. Rolain

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63 CNRS 7278 IRD 198 INSERM U1905, IHU Méditerranée Infection, Facultés de Médecine et de Pharmacie, Marseille, France

Abstract

Mobilicoccus massiliensis strain SIT2 (= CSUR P1306 = DSM 29065) is a new type strain of *Mobilicoccus* sp. nov. isolated from the stool of a 2-year-old Nigerian boy with kwashiorkor. *M. massiliensis* is Gram positive, facultatively anaerobic, nonsporulating and motile. The 3 842 438 bp long genome contains 3362 protein-coding and 49 RNA genes, including one 5S rRNA gene, one 16S rRNA gene, one 23S rRNA gene and 46 tRNA genes.

Keywords: Culturomics, genome, kwashiorkor, *Mobilicoccus massiliensis*, taxonogenomics

Original Submission: 16 August 2016; **Revised Submission:** 30 August 2017; **Accepted:** 31 August 2017

Article published online: 6 September 2017

Corresponding author: J.-M. Rolain, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63 CNRS 7278 IRD 198 INSERM U1905, IHU Méditerranée Infection, Facultés de Médecine et de Pharmacie, Marseille, France.

E-mail: jean-marc.rolain@univ-amu.fr

Mobilicoccus massiliensis strain SIT2 (= CSUR P1306 = DSM 29065) is the type strain of *Mobilicoccus* sp. nov. This bacterium was isolated from the stool of a 2-year-old Nigerian boy with a severe form of acute malnutrition known as kwashiorkor and is part of the culturomics effort, which seeks to cultivate all bacterial species from the human gut [1,2]. It is Gram positive, aerobic or facultatively anaerobic, motile and nonsporulating. The family *Dermatophilaceae* was first proposed by Austwick (1958) and was later emended by Stackebrandt et al. (1997), Stackebrandt and Schumann (2000) and Zhi et al. (2009). This family currently contains two genera: *Dermatophilus* and *Kineosphaera*. The genus *Dermatophilus* was proposed by Gordon (1954) as organisms that form branching mycelia with several transverse and longitudinal divisions, which leads to the formation of packets or clusters of cuboid cells or coccoids. Species of the genus *Dermatophilus* are bacteria isolated from the causative organism of a skin disease [3] and was reported to affect a wide variety of mammalian species. The ruling taxonomic classification of prokaryotes is based on a combination of phenotypic and genotypic criteria [4,5]. However, the three essential criteria that are used, comprising 16S rRNA gene-based phylogeny [4], G+C content and DNA-DNA hybridization [5] have several drawbacks. We recently proposed a new method, taxonogenomics, which uses genomic data in a polyphasic approach to describe new bacterial species [6]. This strategy combines phenotypic characteristics including matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) and genomic analyses [7–9].

Here we report for the first time the isolation and characterization of a novel species, *Mobilicoccus massiliensis* sp. nov., with a description of phylogenetic characteristics as well as complete genomic sequencing and annotation to distinguish this species from other species.

The study was approved by the local ethics committee of the Institut Fédératif de Recherche IFR48, Faculty of Medicine, Marseille, France, under agreement 09-022. Strain SIT2 was isolated in March 2014 by cultivation on chocolate agar Poly-ViteX (bioMérieux, Marcy l'Etoile, France) in anaerobic and aerobic condition using GasPak EZ Anaerobe Container System Sachets (Becton Dickinson (BD), San Diego, CA, USA) at 37°C. This strain exhibited a 98% 16S rRNA gene similarity with

http://dx.doi.org/10.1016/j.nmni.2017.08.007

Mobilicoccus pelagius (NZ-BAFF00000000.1), a phylogenetically valid neighbouring *Dermatophilus* species type strain (Fig. 1).

Optimal growth occurred at 37°C after 24 hours of inoculation. Growth was observed under aerobic and anaerobic conditions after 24 hours. Colonies were 0.2–0.5 mm in diameter in gross appearance on blood-enriched Columbia agar. Cells are coccus shaped, Gram positive and non-sporulating (Fig. 2), and the motility test was positive. SIT2 showed catalase activity but was negative for oxidase.

Commercially available API ZYM and API 50CH strips (bioMérieux) were used to characterize the biochemical properties of the strain according to the manufacturer's instructions. Using an API 50CH strip, *Mobilicoccus massiliensis* SIT2 presented positive reactions for glycerol, erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl-αD-mannopyranoside, D-galactose, D-glucose, D-fructose, D-mannose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl-αD-glucopyranoside, *N*-acetylglucosamine, amygdaline, arbutin, salicin, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, D-melezitose, D-raffinose, amidone, glycogen, xylitol, gentiobiose, D-turanose, D-lyxose, D-tagatose, D-fucose, L-fucose, D-arabitol and potassium gluconate. Negative reactions were observed for L-sorbose, methyl-αD-mannopyranoside, esculin, inulin, L-arabitol, potassium 2-ketogluconate and potassium 5-ketogluconate. For API ZYM, *Mobilicoccus massiliensis* SIT2 presented positive reaction only for α-galactosidase (Table 1).

Antibiotic susceptibility of our isolates was assessed using the disk diffusion method on Mueller-Hinton agar plates supplemented with 5% blood (BD). The tested antibiotics were ceftriaxone, imipenem, vancomycin, rifampicin, gentamicin, ciprofloxacin, amoxicillin, doxycycline, ciprofloxacin, gentamicin, rifampicin, colistin, meropenem, trimethoprim/sulfonamide, amoxicillin/clavulanic acid, fosfomycin and metronidazole (Sirscan Oxoid, Montpellier, France) (Table 2).

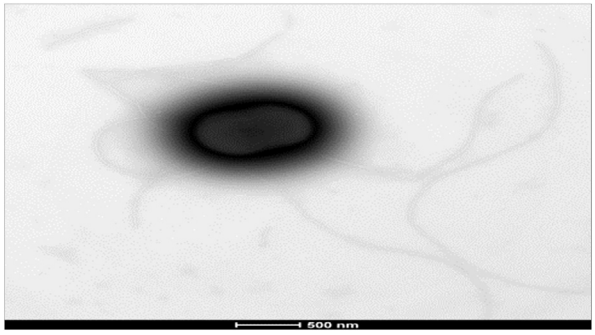

FIG. 2. Transmission electron microscopy of *Mobilicoccus massiliensis* strain SIT2 using Morgani 268D device.

MALDI-TOF MS protein analysis was carried out as previously described [2] using a Microflex spectrometer (Bruker Daltonics, Leipzig, Germany). The resulting score enabled the identification (or not) of the tested species: a score of ≥2 with a validly published species enabled identification at the species level, a score of ≥1.7 but <2 enabled identification at the genus level and a score of <1.7 did not enable any identification. No significant MALDI-TOF MS score was obtained for strain SIT2 against the Bruker database, suggesting that our isolate was not a member of a known species. Consequently, we added the spectrum from strain SIT2 to our database, and the organism was selected for sequencing on the basis of its phylogenetic position and 16S rRNA similarity to members of the genus *Dermatophilus* [2].

The phylogenetic subtree highlighted the phylogenetic position of this bacteria relative to other species. Sequences were recovered by a nucleotide BLAST (Basic Local Alignment

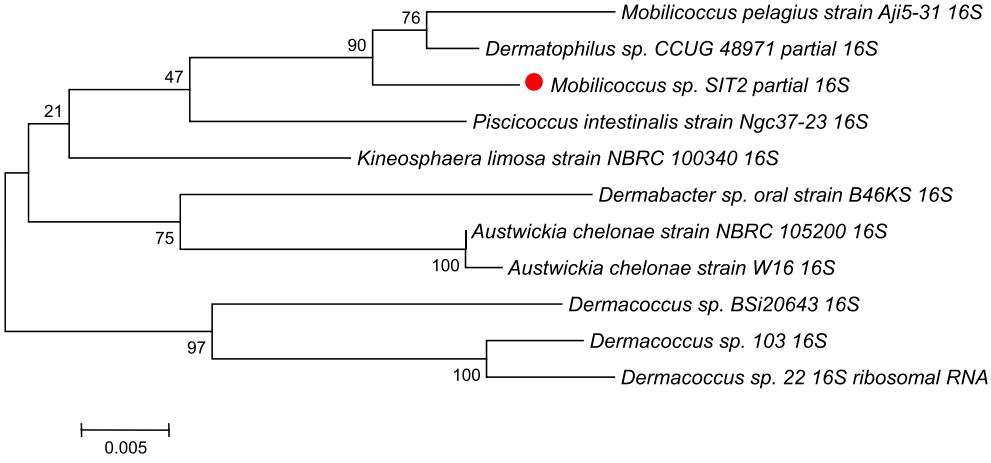

FIG. 1. Phylogenetic tree highlighting position of *Mobilicoccus massiliensis* sp. nov. strain SIT2 (= CSUR P1162 = DSM 29078) relative to other type strains within *Dermatophilus* genus.

TABLE 1. Differential phenotypic characteristics between *Mobilicoccus massiliensis* sp. nov. strain SIT2 and phylogenetically close members of other *Dermatophilaceae* species.

Property	*Mobilicoccus massiliensis*	*Mobilicoccus pelagius*	*Piscicoccus intesttinalis*
Cell diameter	0.2–0.5 mm	0.7–1.2 µm	0.7–1 µm
Oxygen requirement	F/anaerobic	F/anaerobic	F/anaerobic
Gram stain	+	+	+
Motility	+	+	+
G + C content (%)	70.5	71.6	71.5
Production of:			
Alkaline phosphatase	−	+	+
Acid phosphatase	−	−	+
Catalase	+	+	+
Oxidase	−	−	−
α-Glucosidase	−	+	+
β-Glucosidase	−	−	+
α-Galactosidase	+	−	+
β-Galactosidase	−	−	+
Leucine arylamidase	−	+	+
Pyrazinamide	−	+	+
Utilization of:			
Glycerol	+	−	−
Erythritol	+	−	−
D-Arabinose	+	−	−
L-Arabinose	+	−	−
D-Ribose	+	−	−
D-Xylose	+	+	−
L-Xylose	+	+	−
D-Adonitol	+	+	−
Methyl-αD-mannopyranoside	+	+	−
D-Galactose	+	−	+
D-Glucose	+	−	+
D-Fructose	+	−	+
D-Mannose	+	−	+
L-Rhamnose	+	+	−
Dulcitol	+	+	−
Inositol	+	−	−
D-Mannitol	+	+	−
D-Sorbitol	+	+	−
Methyl-αD-glucopyranoside	+	+	−
N-Acetylglucosamine	+	−	−
Amygdalin	+	+	−
Arbutin	+	+	+
Salicin	+	+	−
D-Cellobiose	+	+	−
D-Maltose	+	−	+
D-Lactose	+	+	−
D-Melibiose	+	−	−
D-Saccharose	+	−	−
D-Trehalose	+	−	+
D-Melezitose	+	−	−
D-Raffinose	+	+	+
Amidon	+	−	−
Glycogen	+	−	−
Xylitol	+	+	−
Gentiobiose	+	+	−
D-Turanose	+	−	−
D-Lyxose	+	+	−
D-Tagatose	+	−	−
D-Fucose	+	+	−
L-Fucose	+	+	−
D-Arabitol	+	+	−
Potassium gluconate	+	+	−
Habitat	Stool of human boy	Intestinal tract of fish	Intestinal tract of fish

+, positive result; −, negative result.

Search Tool) against the National Center for Biotechnology Information (NCBI) 16S rRNA Targeted Loci Project database. The bacterium was identified by sequence analysis of the 16S rRNA. Its phylogenetic relationships with closely related species were determined by MEGA v6. The evolutionary history was inferred by using the maximum likelihood method based on the JTT matrix-based model. Strain SIT2 exhibited a 98% 16S rRNA sequence identity with *Mobilicoccus pelagius* (NZ-BAFF00000000.1), the phylogenetically closest bacterial species with standing in nomenclature (Fig. 1).

Genomic DNA of *Mobilicoccus massiliensis* was sequenced via MiSeq Technology (Illumina, San Diego, CA, USA) with the two applications: paired end and mate pair. The reads of both applications were trimmed, and the optimal assembly was obtained through SPAdes (St Petersburg genome assembler) software with 245 contigs of coverage in eight scaffolds, which generated a genome size of 3.28 Mb. The GC% was estimated at 29% (Tables 3 and 4).

The genome was annotated by the Rapid Annotation using Subsystem Technology (RAST) bioserver [10]. The resistome

TABLE 2. Resistance gene associated with antibiotic resistance in *Mobilicoccus massiliensis* SIT2.

Characteristic	Value
ORF	1719
Gene name	*MFS*
GC%	70.6
Size (aa)	513
Function	MFS transporter
Best BLAST hit in GenBank	*Mobilicoccus pelagius*
% aa coverage	97
% aa identity	74

BLAST, Basic Local Alignment Search Tool; MFS, major facilitator superfamily; ORF, open reading frame.

TABLE 3. Nucleotide content and gene count levels of genome.

Attribute	Value	% of total[a]
Genome size (bp)	3 842 438	100
Coding region (bp)	3 415 931	88.90009
G+C content (bp)	2 707 407	70.4696
Total genes	3411	100
RNA genes	49	1.436529
Protein-coding genes	3362	100
Protein associated to function prediction	2359	70.16657
Protein associated to COGs	2099	62.43307
Protein with peptide signals	402	11.95717
Protein with transmembrane helices	738	21.95122

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.

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

Code	Value	Percentage[a]	Description
J	154	4.580607	Translation
A	1	0.029744199	RNA processing and modification
K	173	5.145746	Transcription
L	153	4.550863	Replication, recombination and repair
B	1	0.029744199	Chromatin structure and dynamics
D	23	0.6841166	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	42	1.2492564	Defense mechanisms
T	85	2.528257	Signal transduction mechanisms
M	94	2.7959547	Cell wall/membrane biogenesis
N	25	0.743605	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	30	0.89232594	Intracellular trafficking and secretion
O	79	2.3497918	Posttranslational modification, protein turnover, chaperones
C	147	4.3723974	Energy production and conversion
G	191	5.6811423	Carbohydrate transport and metabolism
E	275	8.179655	Amino acid transport and metabolism
F	60	1.7846519	Nucleotide transport and metabolism
H	114	3.3908389	Coenzyme transport and metabolism
I	106	3.1528852	Lipid transport and metabolism
P	159	4.729328	Inorganic ion transport and metabolism
Q	68	2.0226057	Secondary metabolites biosynthesis, transport and catabolism
R	302	8.982749	General function prediction only
S	155	4.610351	Function unknown
—	1263	37.566925	Not in COGs

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

was analysed with the Antibiotic Resistance Gene-ANNOTation (ARG-ANNOT) database and BLASTp in GenBank [11]. The functional annotation of protein sequences was performed using BLASTp against the GenBank and Clusters of Orthologous Groups (COGs) databases [11]. The exhaustive bacteriocin database available in our laboratories (Bacteriocins of the Unité des Maladies Infectieuses et Tropicales Emergentes (URMITE) database; http://drissifatima.wix.com/bacteriocins) was performed by collecting all currently available sequences from the databases and from NCBI. Protein sequences from this database allowed putative bacteriocins from human gut microbiota to be identified using BLASTp methodology [11]. The genome of *Mobilicoccus massiliensis* SIT2 has been deposited in GenBank with accession number CDGT01000000 and 16S rRNA accession number LK985391. The genome is 3 842 438 bp long with 70.47% G+C content. It is composed of 21 scaffolds (composed of 24 contigs). Of the 3681 predicted genes, 3362 were protein-coding genes and 49 were RNAs (one 5S rRNA gene, one 16S rRNA gene, one 23S rRNA gene and 46 tRNA genes). A total of 2437 genes were assigned as putative function (by COGs or by NR BLAST). The remaining genes were annotated as hypothetical proteins (683 genes, 20.32%) (Fig. 3).

The draft genome sequence of *Mobilicoccus massiliensis* is larger than those of *Mobilicoccus pelagius* Aji5-31, *Dermatophilus congolensis* DSM 44180, *Dermacoccus nishinomiyaensis, Arsenicicoccus* spp. and *Austwickia chelonae* NBRC 105200 (3.54, 2.62, 3.03, 3.53 and 3.54 MB respectively) but smaller than those of *Kineosphaera limosa* NBRC 100340 (4.5 MB) (Table 5).

The G+C content of *Mobilicoccus massiliensis* is smaller than those of *Mobilicoccus pelagius* Aji5-31 and *Arsenicicoccus* spp. (71.9 and 72.7% respectively), but larger than those of *Dermatophilus congolensis* DSM 44180, *Kineosphaera limosa* NBRC 100340, *Dermacoccus nishinomiyaensis* and *Austwickia chelonae* NBRC 105200 (59.4, 70.4, 69.1 and 66.1% respectively). The gene content of *Mobilicoccus massiliensis* is larger than those of *Mobilicoccus pelagius* Aji5-31, *Dermatophilus congolensis* DSM 4418, *Kineosphaera limosa* NBRC 100340, *Dermacoccus nishinomiyaensis, Arsenicicoccus* spp. and *Austwickia chelonae* NBRC 105200 (3090, 2340, 4375, 2745, 3271 and 3046 respectively) (Table 5).

The comparison of amino acid sequence homology of the predicted genes, as shown in Fig. 4, by bidirectional BLAST hits taken from the RAST annotation [10] is a useful way to evaluate the protein similarity using BLAST between NBRC 104925 and the fully sequenced SIT2. Fig. 5 provides the distribution of functional classes of predicted genes of *M. massiliensis* and *M. pelagius*.

Antimicrobial susceptibility testing demonstrate that the strain *M. massiliensis* SIT2 was susceptible to ceftriaxone,

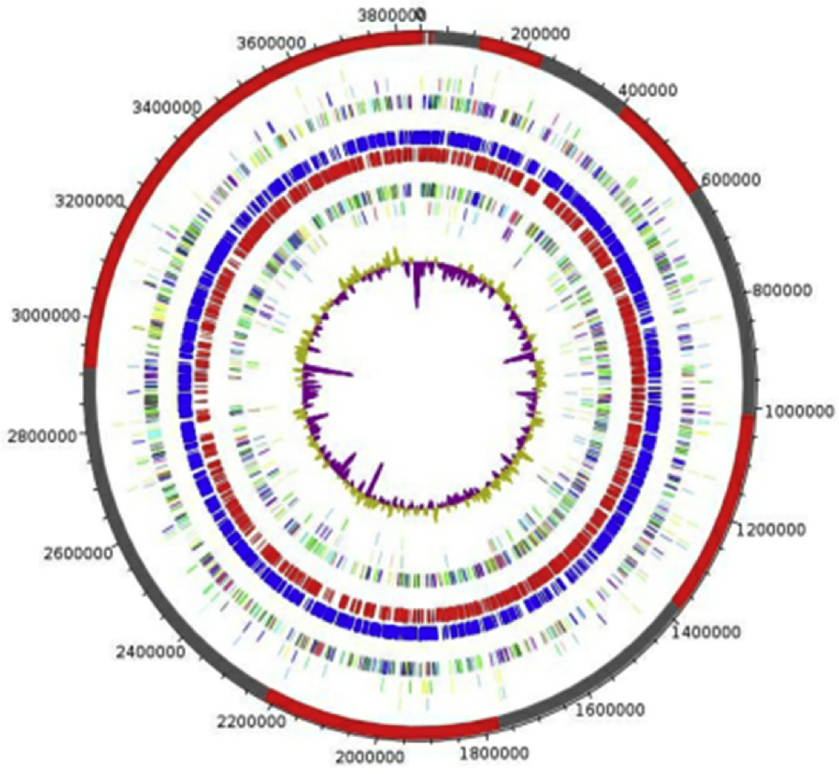

FIG. 3. Graphical circular map of genome.

imipenem, vancomycin, rifampicin, gentamicin, ciprofloxacin, amoxicillin, doxycycline, ciprofloxacin, gentamicin, rifampicin and colistin but resistant to trimethoprim/sulfamethoxazole, fosfomycin and metronidazole. *In silico* analysis of resistome revealed the presence of resistance genes (Table 2).

The analysis of the genome did not demonstrate the presence of bacteriocin and nonreducing polyketide synthases. SIT2 is equipped with an intact flagellar system of 41 Coding DNA Sequence (CDS) encoding six cytoplasmic signal transduction proteins, the products of the *che* genes (*cheA, cheB, cheR, cheW, cheY* and *cheZ*), transmembrane proteins with receptor functions termed methyl-accepting chemotaxis proteins or MCPs, flagellar assembly proteins (*FliP, FliQ, FliR, flhA, flhB*), chemotaxis protein (*motA, motB*), flagellar motor switch protein (*FliG, FliM, FliN, FliY*), rod, hook and filament (*FlgC, FlgG, FlgK, FlgL, fliD, fliC*) and regulation (RNA polymerase sigma factor for flagellar operon *FliA*).

On the basis of phenotypic, phylogenetic and genomic analyses (taxonogenomics), we propose that strain SIT2 represents a novel species of the genus *Mobilicoccus* for which the name

TABLE 5. Genome features of *Mobilicoccus* SIT2 genome compared to other *Dermatophilaceae* species.

Strain	Accession No.	Size (Mb)	GC%	Gene	Protein
Mobilicoccus SIT2	NZ_CDGT00000000.1	3.84	70.5	3377	3182
Mobilicoccus pelagius	NZ_BAFE00000000.1	3.54	71.9	3090	2895
Dermatophilus congolensis	NZ_AUCS00000000.1	2.62	59.4	2340	2204
Kineosphaera limosa	NZ_BAHD00000000.1	4.85	70.4	4375	4033
Dermacoccus nishinomiyaensis	NZ_CP008889.1	3.03	69.1	2745	2619
Arsenicicoccus spp.	NZ_CP012070.1	3.53	72.7	3271	3052
Austwickia chelonae	NZ_BAGZ00000000.1	3.54	66.1	3046	2903

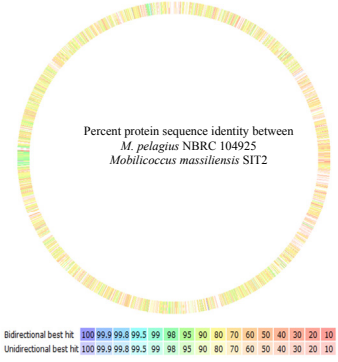

FIG. 4. Proteomic comparison and *in silico* DNA-DNA hybridization between *Mobilicoccus massiliensis* SIT2 and *Mobilicoccus pelagius* NBRC 104925.

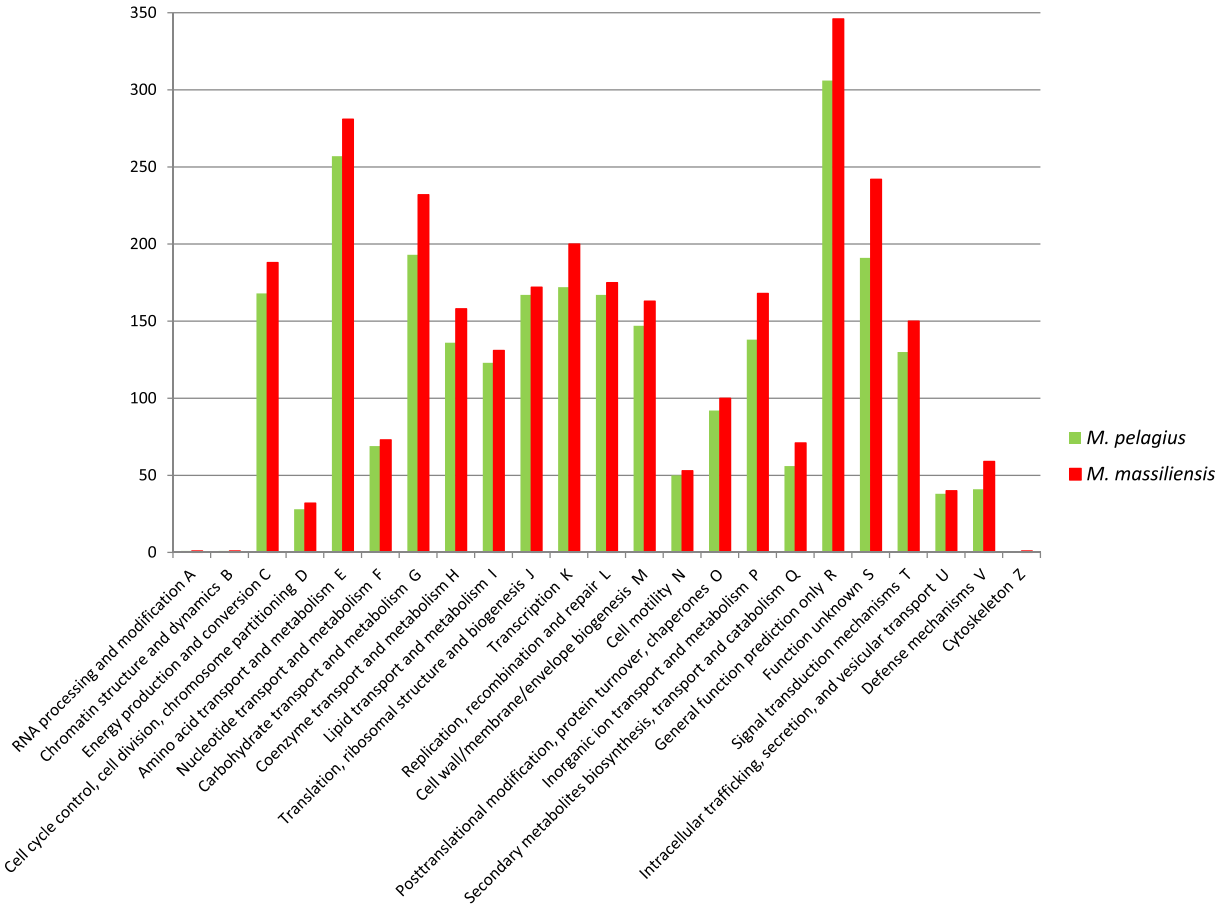

FIG. 5. Distribution of functional classes of predicted genes of *Mobilicoccus massiliensis* and *Mobilicoccus pelagius* according to clusters of orthologous groups of proteins.

Mobilicoccus massiliensis is proposed. The genome sequences are deposited in GenBank under accession numbers CDGT01000000 and 16S LK985391 respectively.

Conflict of interest

None declared.

Acknowledgements

We are grateful to L. Hadjadj (Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63 CNRS 7278 IRD 198 INSERM U1905, IHU Méditerranée Infection, Facultés de Médecine et de Pharmacie, Aix- Marseille Université, Marseille, France) for technical assistance. Supported in part by the Centre National de la Recherche Scientifique (CNRS) and Infectiopôle Sud Fondation.

References

[1] Dubourg G, Lagier JC, Robert C, Armougom F, Hugon P, Metidji S, et al. Culturomics and pyrosequencing evidence of the reduction in gut microbiota diversity in patients with broad-spectrum antibiotics. Int J Antimicrob Agents 2014;44:117–24.

[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] Hamada M, Iino T, Iwami T, Harayama S, Tamura T, Suzuki K. *Mobilicoccus pelagius* gen. nov., sp. nov. and *Piscicoccus intestinalis* gen. nov., sp. nov., two new members of the family *Dermatophilaceae*, and reclassification of *Dermatophilus chelonae* (Masters et al. 1995) as *Austwickia chelonae* gen. nov., comb. nov. J Gen Appl Microbiol 2010;56:427–36.

[4] Stackebrandt E, Frederiksen W, Garrity GM, Grimont PA, Kämpfer P, Maiden MC, et al. Report of the *ad hoc* committee for the re-evaluation of the species definition in bacteriology. Int J Syst Evol Microbiol 2002;52:1043–7.

[5] 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.

[6] 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.

[7] Kokcha S, Mishra AK, Lagier JC, Million M, Leroy Q, Raoult D. Non contiguous – finished genome sequence and description of *Bacillus timonensis* sp. nov. Stand Genom Sci 2012;6:346–55.

[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 Genom Sci 2012;6:116–25.

[9] Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Clostridium senegalense* sp. nov. Stand Genom Sci 2012;6:386–95.

[10] Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. BMC Genom 2008;9:75.

[11] Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 2014;58:212–20.