Taxonogenomics of Culturomica massiliensis gen. nov., sp. nov., and Emergencia timonensis gen. nov., sp. nov.: two new bacterial species isolated from human stool microbiota

Afaf HAMAME
Aix Marseille Univ, IRD, AP-HM, MEPHI
Reham MAGDY WASFY
Aix Marseille Univ, IRD, AP-HM, MEPHI
Cheikh Ibrahima LO
IHU-Méditerranée Infection
Florence FENOLLAR
IHU Méditerranée-Infection, VITROME, Aix-Marseille Université
Didier RAOULT
IHU-Méditerranée Infection
Pierre-Edouard FOURNIER
IHU Méditerranée-Infection, VITROME, Aix-Marseille Université
Linda HOUHAMDI (linda.houhamdi@ap-hm.fr)
IHU Méditerranée-Infection, VITROME, Aix-Marseille Université

Article

Keywords: Culturomica massiliensis sp. nov., Emergencia timonensis sp. nov., Marseille-P2698, Marseille-P2260, New bacterial species, Human microbiota, Taxonogenomics

Posted Date: October 26th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2162345/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on May 25th, 2023. See the published version at https://doi.org/10.1038/s41598-023-35443-7.

Abstract

Two new bacterial strains, Marseille-P2698 (= CSUR P2698 = DSM 103121) and Marseille-P2260 (CSUR P2260 = DSM 101844 = SN18), were isolated from human stools by the culturomic method. We used the taxonogenomic approach, to fully describe these two new bacterial strains. The Marseille-P2698 strain was a Gram-negative, motile, non-spore-forming, rod-shaped bacterium. The Marseille-P2260 strain was a Gram-positive, motile, spore-forming rod-shaped bacterium. Major fatty acids found in Marseille-P2698 were $C_{15:0\ iso}$ (63%), $C_{15:0\ anteiso}$ (11%), and $C_{17:0\ 3\text{-}OH\ iso}$ (8%). Those found in Marseille-P2260 strain were $C_{16:00}$ (39%), $C_{18:1n9}$ (16%) and $C_{18:1n7}$ (14%). Strains Marseille-P2698 and Marseille-P2260 had 16S rRNA gene sequence similarities of 91.50% with *Odoribacter laneus*, and of 90.98% and 95.07% with *Odoribacter splanchnicus* and *Eubacterium sulci*, respectively. The exhibited digital DNA-DNA Hybridization values lower than 20.7%, and Orthologous Average Nucleotide Identity values lower than 73% compared to their closest related bacterial species *O. splanchnicus* and *E. sulci* respectively. Phenotypic, biochemical, phylogenetic, and genomic results obtained by comparative analyses provided sufficient evidence that both of the two studied strains Marseille-P2698 and Marseille-P2260 are new bacterial species for which the names *Culturomica massiliensis* gen. nov., sp. nov., and *Emergencia timonensis* gen. nov., sp. nov. were proposed, respectively.

Introduction

The human gut microbiota is considered currently as one of the most active research fields in microbiology (Sankar et al., 2015). In fact, this micflora harbours a huge biodiversity of bacteria of which a large part is still unknow (Fontaine et al., 2012). Researchers used a variety of strategies to speed up and simplify the description of new bacterial species by optimizing their *in vitro* growth conditions (Lagier et al., 2012; Lagier et al., 2015). Culturomics is one among these strategies, which relies on a diversification of culture conditions that allowed the identification of several new bacterial species isolated from the human gastrointestinal tract (Lagier et al., 2012; Lagier et al., 2015). Since 2012, it allowed the isolation of over 1,000 different human-associated bacterial species, including several hundreds of new species (Lagier et al., 2012; Lagier et al., 2015; Abdallah et al., 2017). This method highlighted the need to adopt taxonomic approaches to clinical microbiology by including the use of modern and reproducible tools, such as high throughput genomic and proteomic analyses.

In 2015, two putative new bacterial species, *Culturomica massiliensis*, and *Emergencia timonensis*, were isolated from patient's stools, and partially described (Bessis et al., 2016; Ndongo et al., 2016). The genomic sequencing of new described bacterial species constitutes currently a necessary step for performing their comparative taxogenomic descriptions with their closest related known species. In fact, several recent publications have used genomic descriptions to characterize the new species by comparison to their closest relatives strains (Ndongo et al., 2018; Ndongo et al., 2021; Zgheib et al., 2022). The aim of our current study was to complete the phenotypic, taxonomic and genomic characterization of *Culturomica massiliensis* strain Marseille-P2698, and *Emergencia timonensis* strain Marseille-P2260 and formally expose the creation of both species.

Materials And Methods

Sample collection and ethics approval

In November 30, 2015, stool samples were collected from hospitalized patients in the Timone Hospital (Marseille, France) as a part of a study of human microbiota diversity. Patients provided signed informed consent (Bessis et al., 2016; Ndongo et al., 2016). The study protocol was approved by the ethics committee of the Institut de Recherche Fédératif 48, under agreement number 09–022. All the methods were carried out in accordance with relevant Guidelines and regulations. Moreover, the patients gave an informed consent for this study. Each sample was then cultured according to the Culturomics method previously established in our laboratory (Lagier et al., 2015). Various types of bacterial colonies were isolated on 5% of sheep blood-enriched Columbia agar (bioMérieux®, Marcy l'Etoile, France). Bacterial colonies were then screened for identification by Matrix-Assisted Laser Desorption Ionization-Time Of Flight Mass Spectrometry (MALDI-TOF MS) instrument (Bruker Daltonics®, Bremen, Germany) as previously reported (Seng et al., 2009). Both of the two strains studied herein had a MALDI-TOF score lower than 2.0, which did not allow their correct identification. Their spectra were then added to the local MALDI-TOF MS database (https://www.mediterranee-infection.com/urms-data-base).

16s Rrna Gene Sequencing And Identification

The 16S rRNA sequences from both strains were directly extracted from their whole genomes sequences and then, compared by Basic Local Alignment Search Tool nucleotide (BLASTn) to the nr database (non-redundant databases) (BLAST, 2022). The obtained sequence similarity percentages allowed identification of the closest species to each strain, and to predict if it was new species (< 98.65% of similarity). Then, a phylogenetic tree was constructed based on these 16S rRNA gene sequences in comparison to the closest related species of each studied strain. Designated species sequences were downloaded from nr (Parte et al., 2020), and aligned with ClustalW. Phylogenetic trees were constructed using MEGA 11 version 11.0.10 with the maximum likelihood method and 1000 bootstrap replications (Kumar et al., 2018).

Phenotypic And Biochemical Characterizations

Optimal culture conditions were determined by testing various incubation temperatures (25, 28, 37, 42, and 50°C), atmospheres (aerobic, anaerobic and microaerophilic), NaCl concentrations (5, 5.5, 7.5, 10, 15, and 20% of NaCl) and pH levels (5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5). The morphology and motility were observed using a new generation scanning electron microscope (Hitachi High-71 Technologies Corporation, Tokyo, Japan).

Furthermore, three semi-quantitative standardized micromethods of Analytical Profile Index (API®, bioMérieux®) tests: API® 20A, API® 50 CH, and API® ZYM were used, according to the manufacturer's instructions (bioMérieux®, 2022), in order to study carbohydrate metabolism and enzymatic activities.

Fatty acid methyl ester (FAME) analysis was explored by Gas Chromatography/Mass Spectrometry, as previously reported (Sasser, 2006; Dione et al., 2016). FAMEs were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500 - SQ 8 S, Perkin Elmer®, Courtaboeuf, France). Obtained spectra were compared with those contained in the repertory databases using MS Search 2.0 operated with the Standard Reference Database 1A (National Institute of Standards and Technology-NIST, Gaithersburg, USA), and FAMEs mass spectral database (Wiley, Chichester, UK).

Whole Genomic Sequencing And Bioinformatic Analyses

First, bacterial DNA was extracted using the EZ1 DNeasy Blood Tissue Kit (Qiagen® GmbH, Hilden, Germany) in line with the manufacturer's protocol (Anani et al., 2019). Whole-genome sequencing was performed using an Illumina® MiSeq sequencer (Illumina®, San Diego, CA, USA) (Caporaso et al., 2012). Then, sequenced genomes were assembled using SPAdes 3.5.0 software (Prjibelski et al., 2020), which reduces short indels and the huge number of mismatches. Raw reads in contigs less than 700-bp-long were removed. Finally, the quality of the sequenced genome was checked using BLAST against the nr/nt database. This method allowed to better explore the relationship between a submitted assembly of our new species to the International Nucleotide Sequence Database Collaboration (INSDC), i.e. DDBJ, ENA, or GenBank, and the assembly represented in the NCBI Reference Sequence (RefSeq) project. The global statistics section reported general statistics information including Gaps between scaffolds, number of scaffolds, number of contigs, total sequence length, and total ungapped length. Furthermore, taxonomic data were checked according to the best-matching-type strain with the declared new species repertory in NCBI (NCBI, 2022).

During annotation, genomic parameters were evaluated including transfer-messenger RNAs (tmRNAs) and transfer RNAs (tRNAs) using ARAGORN version 1.2 and ribosomal RNAs (rRNAs) using Barrnap version 0.9 (Laslett *et al.*, 2004; Hyatt *et al.*, 2020). Generated file (*.faa*) was used for BLAST-P analyses against the Clusters of Orthologous Genes (COGs) database, and used for CRISPR-Cas identification (Tatusov et al., 2000). Resistance genes were screened using ResFinder (Bortolaia et al., 2020). Other bioinformatic tools were also used such as AntiSMASH to search polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS) (Seemann, 2014). Circular maps of the two genomes were generated using CGView (Circular Genome Viewer) software. This Java application converts XML or tab-delimited input into a Vector Graphics format (Stothard et al., 2018).

Besides, phylogenetic trees of interest were generated with the FastME 2.1.6.1 software to highlight the position of each new bacterial strain among its closest relatives (Meier-Kolthoff *et al.*, 2019). Digital DNA-DNA Hybridization (dDDH) values were calculated to check the difference between the genomes using the following website (https://ggdc.dsmz.de). Critical limit was set at 70% below which a prokaryotic species may be considered as new ((Meier-Kolthoff et al., 2013). Orthologous average nucleotide identity (OrthoANI) version 0.93.1 was also used to calculate genomic similarities between studied species and their related taxa.

Results

Strain identification and phylogenetic analyses (Fig.1)

The species names *Culturomica massiliensis* gen. nov., sp. nov., and *Emergencia timonensis* gen. nov., sp. nov., had been previously proposed for the new species mainly as representative strains Marseille-P2698 and Marseille-P2260, respectively (Bessis *et al.,* 2016; Ndongo *et al.,* 2016). As these previous descriptions were not exhaustive, we revisited the work by including phylogenetic, morphological, and genomic data.

Strain Marseille-P2698 had a 16S rRNA gene sequence similarity and a query coverage of 91.5% and 99% respectively with *Odoribacter laneus* strain YIT 12061^{T}. Strain Marseille-P2260 exhibited 16S rRNA gene similarity and a query coverage of 92.72%, and 100% respectively, with *Eubacterium sulci* strain ATCC 35585^{T}.

Phenotypic and biochemical characterizations

Growth of strains Marseille-P2698 and Marseille-P2260 occurred on 5% sheep blood-enriched Columbia agar (bioMérieux®), after 48 hours of incubation at 37°C in a strict anaerobic atmosphere. Optimal growth was obtained at pH 7. However, strain Marseille-P2260 did not tolerate NaCl whereas strain Marseille-P2698 could grow with a NaCl concentration of 0.5%.

Strain Marseille-P2698 is a Gram-negative rod, strictly anaerobic, motile, and non-spore-forming with a size of 1.5–3 µm in length and 0.3 to 0.4 µm in diameter. It exhibits a positive catalase, but no oxidase activity. The colonies are circular, beige and from 0.7 to 1.2 mm in diameter.

Strain Marseille-P2260 is a Gram-positive, strictly anaerobic rod, ranging in length from 1 µm to 1.5 µm, and in diameter from 0.5 to 1 µm. It has no catalase or oxidase activity. Colonies of strain Marseille-P2260 are translucent with a diameter of 0.5 to 1 mm. The remaining cell characteristics of both strains compared to their closest relatives are summarized in **Table.1.** Using API® 50CH strips (bioMérieux®), positive reactions were obtained for both studied strains for glycerol, D-ribose, D-galactose, D-glucose, D-fructose, D-mannose, D-mannitol, D-sorbitol, N-acetylglucosamine, amygdalin, arbutin, esculin ferric citrate, salicin, D-cellobiose, D-maltose, D-lactose, D-saccharose, D-trehalose, D-melezitose, gentiobiose, D-tagatose. Using API® ZYM strips (bioMérieux®), positive activities were observed for esterase lipase (C8), leucine arylamidase, phosphatase acid, and naphthol-as-bi-phosphohydrolase for both studied strains. In contrast, phosphatase alkaline, esterase (C4), α-chymotrypsin, β-galactosidase, α-glucosidase, β-glucosidase, and N-acetyl-β-glucosaminidase were only positive for strain Marseille-P2260. Using API® 20A strips, positive results for the two strains were obtained for D-glucose, D-mannitol, D-lactose, D-saccharose, D-maltose, salicin, esculin ferric citrate, glycerol, D-cellobiose, D-mannose, D-melezitose, D-sorbitol, L-rhamnose, and D-trehalose. However, strain Marseille-P2698 was positive to Gelatin, unlike strain Marseille-P2260 which was negative.

The most abundant fatty acids from strain Marseille-P2698 were 13-methyl-tetradecanoic acid (63%), 12-methyl-tetradecanoic acid (11%), and 3-hydroxy-15-methyl-hexadecanoic acid (8%). Several other branched structures, mainly iso, were also detected. Specific 3-hydroxy structures were detected, mainly branched as well. For strain Marseille-P2260, the major fatty acid was hexadecanoic acid (39%) (**Table.2**).

Table.1 Phenotypic characteristics of strains Marseille-P2698 and Marseille-P2260 compared with closely related species.

Properties	*Culturomica massiliensis Marseille-P2698^T*	*Odoribacter laneus YIT 12061^T*	*Odoribacter splanchnicus JCM 15291^T*	*Butyricimonas synergistica MT01^T*	*Emergencia timonensis Marseille-P2260^T*	*Eubacterium sulci ATCC 35585^T*	*Eubacterium infirmum NCTC 12940^T*
Gram stain	-	-	-	-	+	-	+
Cell shape	Rod	Rod	Fusiform	Rod	Rod Bacilli	Rod	Rod
Motility	+	NA	-	-	+	-	-
Cell diameter(µm)	0.7-1.2	NA	NA	1.0	0.5-1.0	1.0	1.0
Lenghth (µm)	0.3–0.4 × 1.5–3	9.2 x 1.15	0.7 × 1.0-5.0	1.5	1 to 1.5	0.5 x 1-2	0.5 x 1-2
Endospore	-	-	-	-	-	-	-
Temperature	37°C	37°C	37°C	37°C	37°C	37°C	37°C
Oxygen tolerance	Strict anaerobic	Anaerobic	Anaerobic	Anaerobic	Strict anaerobic	Anaerobic	Strict anaerobic
Incubation	4 days	4 days	2-3 days	2 days	3 days	3 days	7 days
Salt tolreance	0.5 %	NA	NA	NA	-	NA	NA
Optimum pH	7	NA	NA	6-7.5	7	NA	NA
Acid phosphatase	+	+	NA	NA	+	NA	NA
Catalase	+	NA	-	-	-	NA	-
Oxidase	-	NA	NA	-	-	NA	NA
Indole	-	+	+	+	-	-	-
Urease	-	-	-	-	-	-	-
βGalactosidase	-	+	-	-	+	-	NA
Ribose	+	NA	NA	NA	+	-	-
Mannose	+	-	+	-	+	-	NA
Mannitol	+	-	NA	-	+	-	-
Sucrose	+	-	-	-	+	-	-
Glucose	+	-	+	+	+	NA	-
Fructose	+	NA	+	NA	+	NA	NA
Maltose	+	-	NA	-	+	-	-
Sorbitol	+	-	NA	-	+	-	NA
Starch	+	NA	NA	NA	-	NA	NA
Lactose	+	-	-	-	+	-	-
Source	Human faeces in diabetic patient	Human faeces	Human Abdominal abscess	Mouse feces	Human faeces	Human gingival sulcus	human periodontal pockets.

+: positive result. −: negative result. **NA:** not available data.

Table.2 Cellular fatty acids composition of strains Marseille-P2698 and Marseille-P2260.

	Fatty acids	Name	Mean relative % (a)
	15:0 iso	13-methyl-tetradecanoic acid	63.4 ± 1.8
	15:0 anteiso	12-methyl-tetradecanoic acid	10.9 ± 0.2
	17:0 3-OH iso	3-hydroxy-15-methyl-Hexadecanoic acid	8.0 ± 0.7
	16:00	Hexadecanoic acid	7.8 ± 0.3
	16:0 3-OH	3-hydroxy-Hexadecanoic acid	3.1 ± 0.3
	5:0 iso	3-methyl-Butanoic acid	1.2 ± 0.0
Culturomica	18:2n6	9,12-Octadecadienoic acid	1.1 ± 0.1
massiliensis	18:1n9	9-Octadecenoic acid	1.0 ± 0.1
Strain	15:00	Pentadecanoic acid	TR
Marseille	14:00	Tetradecanoic acid	TR
P2698	18:00	Octadecanoic acid	TR
	15:0 3-OH iso	3-hydroxy-13-methyl-Tetradecanoic acid	TR
	16:0 iso	14-methyl-Pentadecanoic acid	TR
	18:1n7	11-Octadecenoic acid	TR
	17:0 iso	15-methyl-Hexadecanoic acid	TR
	14:0 iso	12-methyl-Tridecanoic acid	TR
	16:0 3-OH iso	3-hydroxy-14-methyl-Pentadecanoic acid	TR
	17:00	Heptadecanoic acid	TR
	15:00	3-OH anteiso 3-hydroxy-12-methyl-Tetradecanoic acid	TR
	16:00	Hexadecanoic acid	39.2 ± 2.4
	18:1n9	9-Octadecenoic acid	16.5 ± 1.3
	18:1n7	11-Octadecenoic acid	14.5 ± 4.8
Emergencia	14:00	Tetradecanoic acid	11.9 ± 0.6
timonensis	18:00	Octadecanoic acid	9.5 ± 1.4
Strain	18:1n3	15-Octadecenoic acid	5.6 ± 0.9
Marseille	18:2n6	9,12-Octadecadienoic acid	1.1 ± 0.3
P2260	15:00	Pentadecanoic acid	TR
	15:0 anteiso	12-methyl-tetradecanoic acid	TR
	12:00	Dodecanoic acid	TR

a: Mean peak area percentage. **TR:** Trace amount.

Genomic properties and analyses (**Fig.2**)

The whole genome of strain Marseille-P2698 was composed of 14 contigs, for a total size of 4,410,591 bp, with a G+C content of 43 mol%. This genome contained 3,679 genes, of which 3,487 were protein-coding genes. In addition, 59 RNA sequences were also identified and distributed as follows: 8 rRNAs (three 16S, two 23S, and three 5S), 51 tRNAs, and 1 tmRNA.

Genome from strain Marseille-P2260 was composed of 9 contigs, with a size of 4,661,482 bp, and a 45.8 mol% G+C content. Genome annotation identified 4,380 genes, of which 4,288 were protein-coding genes. There were 56 RNA sequences including 5 rRNAs (one 16S, one 23S, and three 5S), 51 tRNAs, and 1 tmRNA.

Odoribacter laneus, O. splanchnicus, Butyricimonas synergistica, B. faecihominis and B. virosa exhibited genome sizes ranging from 4.39 to 4.81 Mbp. The closest relatives of strains Marseille-P2698 and Marseille-P2260 are presented in **Table.3**. Eubacterium sulci, E. infirmum, A. terrae, E. minutum, and Aminipila odorimutans exhibited genome sizes ranging from 1.91 to 4.66 Mbp.

Strain Marseille-P2698 shared dDDH values of 21.6% with Odoribacter laneus, 20.6% with O. splanchnicus, 30.6% with Butyricimonas faecihominis, 21.8% with B. virosa, and 18.9% with B. synergistica. Strain Marseille-P2260 exhibited dDDH values of 20.1% with Eubacterium sulci, 21.1% with E. infirmum, 23.1% with A. terrae, 24.4% with E. minutum, and 21.1% with Aminipila odorimutans (**Table.4**).

The phylogenetic relationships of strains Marseille-P2698 and Marseille-P2260 with relative strains, based on whole-genome sequencing, is represented in **Fig.3**.

Table.3 Summary of comparative genomes and characteristics for strains Marseille-P2698 and Marseille-P2260.

	Organisms	INSDC/RefSeq	Size (Mbp)	G+C %	Protein-coding genes	rRNA	tRNA	tmRNA	genes	Genes assigned to COGs	Contigs	ba
	Strain Marseille-P2698	LT594611.1	4.41	43	3,487	8	50	1	3,679	2,113	11	4,
	Odoribacter laneus YIT 12061 ^{T}	NZ_JH594596	3.77	40.6	2,993	10	53	3	3,100	2,995	42	3,
Culturomica	*Odoribacter splanchnicus* JCM 15291^{T}	CP086000.1	4.39	43.4	3,545	12	59	1	3,757	2,106	1	4,
massiliensis strain Marseille-P2698	*Butyricimonas synergistica* MT01^{T}	NR_041690.1	4.77	43.8	3,741	8	54	1	3,922	2,113	16	4,
	Butyricimonas faecihominis 180-3^{T}	NR_126194.1	4.79	42.9	3,832	4	57	1	4,017	2,156	29	4,
	Butyricimonas virosa CCUG 56611^{T}	CP069450.1	4.81	42.4	3,837	15	57	1	4,032	2,169	1	4,
	Strain Marseille-P2260	FLKM01000000.1	4.66	45.8	4,288	5	50	1	4,380	2,895	9	4,
Emergencia	*Eubacterium sulci* ATCC 35585^{T}	CP012068.1	1.73	39.9	3,062	6	40	1	1,646	1,235	1	1,
timonensis strain Marseille-P2260	*Eubacterium infirmum* NCTC 12940^{T}	NZ_JH594435.1	1.91	39.9	3,510	4	43	1	1,781	1,305	13	1,
	Aminipila terrae CBA3637^{T}	CP047591.1	3.51	37	3,102	18	65	1	3,581	2,346	1	3,
	Eubacterium minutum ATCC 700079 ^{T}	CP016203	1.9	45.79	1,454	7	43	1	1,620	1,035	124	1,
	Anaerovorax odorimutans DSM_5092^{T}	NZ_AUFC01000001.1	3.26	31.5	2,894	4	42	1	2,944	1,976	54	3,

Table.4 Comparative digital DNA-DNA Hybridization (dDDH) values (%) between studied bacterial genomes.

	CM	OL	OS	BS	BF	BV	ET	EM	AT	AO	ES	EI
CM	100%	21.6% (±4.7%)	20.6 (±4.7%)	18.9 (±4.5%)	30.6 (±4.9%)	21.8 (±4.7%)	17.8 (±4.4%)	16 (±4.4%)	17.1 (±4.4%)	21.5 (±4.7%)	18.3 (±4.5%)	18.2 (±4.5%)
OL		100%	13.3 (±6.5%)	18.6 (±4.66%)	18.5 (±4.5%)	19.9 (±4.6%)	18.1 (±4.5%)	17.7 (±4.5%)	18.1 (±4.5%)	21 (±4.7%)	18.2(±4.6%)	18.1(±4%)
OS			100%	23.7 (±4.8%)	19.2 (±4.6%)	23 (±4.8%)	19.3 (±4.5%)	18.5 (±4.5%)	19.1 (±4.6%)	21.5 (±4.7%)	19.1 (±4.6%)	17.3 (±4.5%)
BS				100%	21 (±4.7%)	20.9 (±4.7%)	19.6 (±4.6%)	20.9(±4.7%)	20.6 (±4.6%)	22.9 (±4.8%)	20.9 (±4.7%)	18.1 (±4.5%)
BF					100%	42.2 (±5.1%)	43.4 (±5.1%)	3.7 (±2%)	3.7 (±2%)	30.4 (±4.9%)	3.7 (±2%)	3.7 (±2%)
BV						100%	18 (±4.5%)	17.7 (±4.5%)	17.9 (±4.5%)	21 (±4.7%)	18.3 (±4.5%)	18.3 (±4.6%)
ET							100%	24.4 (±4.8%)	23.1 (±4.8%)	21.1 (±4.7%)	20.1 (±4.6%)	21.1 (±4.6%)
EM								100%	29.5 (±4.9%)	31.2 (±4.9%)	25.4 (±4.8%)	24.9 (±4.8%)
AT									100%	23 (±4.8%)	28.8 (±4.9%)	24.4 (±4.8%)
AO										100%	19.8 (±4.6%)	23.4 (±4.8%)
ES											100%	28.9 (±4.9%)
EI												100%

CM. Culturomica massiliensis Marseille-P2698. **OL.** Odoribacter laneus. **OS.** Odoribacter splanchnicus. **BS.** Butyricimonas synergistica. **BF.** Butyricimonas faecihominis. **BV.** Butyricimonas virosa. **ET.** Emergencia timonensis Marseille-P2260. **EM.** Eubacterium minutum. **AT.** Aminipila terrae. **AO.** Anaerovorax odorimutans. **ES.** Eubacterium sulci. **EI.** Eubacterium infirmum.

The obtained dDDH values were lower than the 70% threshold used for delineating prokaryotic species (Lagier et al., 2012). In addition, strain Marseille-P2698 exhibited OrthoANI values of 73.99% with O. laneus, 72.22% with O. splanchnicus, and 69.24% with B. synergistica. Strain Marseille-P2260 had an OrthoANI value of 67.65% with E. sulci and 67.36% with E. infirmum. These values were lower than 95 %, also suggesting that strains Marseille-P2698 and Marseille-P2260 belonged to distinct species (**Fig.4**).

Genes encoding proteins are divided into several categories according to their functions and others have unknown functions. COGs of strains Marseille-P2698 and Marseille-P2260 have a different functional distributions from their closest species (**Fig.5**).

Antibiotic resistance genes and defense mechanisms

Using the ResFinder software, antibiotic resistance genes (ARG) *erm* (F) and *tet* (Q) were detected within the genome of strain Marseille-P2698, with identity percentages of 100% and 99.8%, respectively. For strain Marseille-P2260, ARG included *erm* (B), *tet* (M), and *tet* (O), with identities ranging from 99.7% to 100% (**Table.5**).

The genomes of strains Marseille-P2698 and Marseille-P2260 contained CRISPR-Cas modules (**Table.6**). Strain Marseille-P2698 exhibited one defense mechanism composed of polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS) enzymes. NRPS-PKS was previously demonstrated to have a role in the biosynthesis of pharmaceutically-important natural products (Akey *et al.,* 2012). Using ANTI SMASH software, NRPS-PKS-like genes cluster had been detected in strain Marseille-P2698 genome, and β-lactone; containing potent medicinal properties (Robinson *et al.,* 2020; Wang *et al.,* 2021), had been detected in strain Marseille-P2698 genome. In addition, according to our results in ToxFinder-1.0 software detecting toxins (ResFinder, 2022), our two strains did not possess toxin genes.

Table.5 Antibiotic resistance genes detected in the genomes strains Marseille-P2698 and Marseille-P2260 using ResFinder software.

	Resistance gene	Identity	Alignment Length/Gene Length	Position in contig	Phenotype
	Macrolide				
	erm(F)	100	801/801	1826565..1827365	erythromycin,lincomycin,clindamycin,quinupristin,pristinamycin ia,virginiamycin s
	Lincosamide				
Culturomica	erm(F)	100	801/801	1826565..1827365	erythromycin,lincomycin,clindamycin,quinupristin,pristinamycin ia,virginiamycin s
massiliensis	**Streptogramin b**				
Marseille-P2698	erm(F)	100	801/801	1826565..1827365	erythromycin,lincomycin,clindamycin,quinupristin,pristinamycin ia,virginiamycin s
	Tetracycline				
	tet(Q)	99.8	1926/1926	1055038..1056963	doxycycline,tetracycline,minocycline
	Macrolide				
	erm(B)	100	738/738	1359368..1360105	erythromycin,lincomycin,clindamycin,quinupristin,pristinamycin ia,virginiamycin s
Emergencia	**Lincosamide**				
timonensis	erm(B)	99.86	738/738	1359368..1360105	erythromycin,lincomycin,clindamycin,quinupristin,pristinamycin ia,virginiamycin s
Marseille-P2260	**Streptogramin b**				
	erm(B)	99.9	738/738	1359368..1360105	erythromycin,lincomycin,clindamycin,quinupristin,pristinamycin ia,virginiamycin s
	Tetracycline				
	tet(M)	99.84	1920/1920	1222665..1224584	doxycycline,tetracycline,minocycline
	tet(O)	99.79	1920/1920	2008154..2010073	doxycycline,tetracycline,minocycline

Table.6 Identification of CRISPR-Cas type and subtype genes in user-submitted sequence for strains Marseille-P2698 and Marseille-P2260.

Conclusion

Culturomics combined with Taxogenomics allowed the isolation and full characterization of two new bacterial species from the human intestine. The results of phenotypic, biochemical, phylogenetic, and genomic analyses obtained for both studied strains proved that they belong to new species. Thus, we propose the creation of two new bacterial genera, *Culturomica massiliensis* gen. nov., sp. nov., and *Emergencia timonensis* gen. nov., sp. nov..

Description of *Culturomica* gen. nov.

Culturomica (*Cul.tu.ro.mi'ca* . N.L. fem, *culturomica*, referring to a new method of diversified bacterial culture). Cells are anaerobic, Gram-negative, motile and rod-shaped. The type species *is Culturomica massiliensis* sp. nov. ^{T}. It was isolated from human feces.

Description of *Culturomica massiliensis* sp. nov.

Culturomica massiliensis (*mas.si.li.en'sis*. L. fem. adj.*massiliensis*, form "*Massilia*", the Latin name of Marseille, France, where the type strain was iso-lated). Bacterial cells are Gram-negative and rod-shaped. They are motile and non-spore-forming, with lengths ranging from 1.5 to 3.0 µm and diameters from 0.3 to 0.4 µm. The optimal growth conditions are 37°C, a strict anaerobic at-mosphere, and neutral pH. On 5% sheep blood-enriched Columbia agar, colonies appear small, circular, beige to white, and measure between 0.7 to 1.2 mm in diameter. They ferment D-galactose, D-glucose, D-fructose, D-mannose, inosi-tol, D-mannitol, D-sorbitol, N-acetylglucosamine, amygdalin, arbutin, esculin, salicin, D-cellobiose, D-maltose, D-glucose, glycerol, D-lactose, D-saccharose, D-trehalose, D-melezitose, Amidon, gentiobiose, D-tagatose, potassium glu-conate and potassium 5-ketogluconate. Enzymatic activities for phosphatase al-kaline, esterase, esterase lipase, leucine arylamidase, α-chymotrypsin, phospha-tase acid, naphthol-as-bi-phosphohydrolase, β-galactosidase, α-glucosidase, β-glucosidase, and N-acetyl-β-glucosaminidase are present. The major cell wall fatty acids are $C_{15:0\ iso}$ (63 %), $C_{15:0\ anteiso}$ (11 %), and $C_{17:0\ 3\text{-}OH\ iso}$ (8%). The genome size from strain Marseille-P2698 is 4.41Mbp-long with a G+C content of 43 mol%.

Bacteria	Element	CRISPR Id /gene name	Start	End	Spacer /Gene	Repeat consensus /cas genes
	CRISPR		112	199	1	ATGGGACTCTTTTTTTGTAAATA
	Cas cluster	CAS	71544	73337	1	Cas3_0_I
	Cas cluster	CAS-TypeIB	613607	622650	7	Cas6_0_I-III, Cas7_3_IB, Cas5_1_IB, Cas3_0_I,
		Cas6_0_I-III				Cas4_0_I-II, Cas1_0_I-II-III, Cas2_0_I-II-III-V
Culturomica	CRISPR	Cas7_3_IB	622870	625330	37	CTTTTAATTGAACTAAGGTAGAATTGAAAC
massiliensis	CRISPR	Cas5_1_IB	992376	993764	21	GTTTCAATACTACTTAGTTCTATTAAAAG
Marseille-P2698	CRISPR	Cas3_0_I	2048339	2049405	15	GTCTCAATGCCGTGATTACAGCTATTCTCTAAATC
	Cas cluster	Cas4_0_I-II	2049852	2057279	6	Cas2_0_I-II-III-V, Cas1_0_I-II-III, Cmr6_0_IIIB,
						Cmr4_0_IIIB,Cmr3_0_IIIB, Cas10_1_IIIB
	CRISPR	Cas1_0_I-II-III	3461867	3462226	5	ATTTCAATTCTACTCCAGTTCTATTAAAAT
	CRISPR	Cas2_0_I-II-III-V	3750732	3750812	1	AATTCTCTTTAGCGTTGGTTTGTG
	CRISPR		882	1086	2	GGCATATATTATCAATACAGTATTTCCCTATTTT
	CRISPR		300401	300525	1	TTGCGGCAACGCCGTGAGATTTTGGAATCATTATA
	CRISPR		522369	522503	1	TTGCAGCAACGCTGTCAATATCGTTGTGTTCATTATACCATGA
	CRISPR		1380564	1381183	9	GTCGCATCCTCATGGATGCGTGGATTGAAAT
	CRISPR	CAS-TypeIC	1429920	1430223	4	GTCTTGCTCCGCATGGAGCAAGTGGATTGAAAT
	Cas cluster		1453194	1461082	7	Cas3_0_I, Cas5_0_IC, Cas8c_0_IC, Cas7_0_IC, Cas4_0_I-II, Cas1_0_IC, Cas2_0_I-II-III-V
Emergencia	CRISPR		1461248	1463328	30	GTCGCACTCCGCATGGAGTGCGTGGATTGAAAT
timonensis	CRISPR		1608002	1608146	1	AAAGCATAAGGACTTTCATCAATTTGTTTCACAAATTGTGACAAGTCCATTA
Marseille-P2260	CRISPR		1944682	1944814	1	GGTATAATGAATCCATAACCAAACGACGTTGTCGCAAGGTAG
	CRISPR		225209	225286	1	ATCCAAGCCAATTTCTTATCGAGGCT
	CRISPR		530585	530712	1	TATTTTCAGCGGAATGCACATGGTATAATAAACGCA
	CRISPR		853819	853957	1	TGGTATAATGGTTCCACGATCTTACGGCGTTGCCGAAGTGTAGGAACCA
	CRISPR		1673013	1673146	1	CATTATACGACAAAACGCAGCCAAATCAGAGATTTGGGCGTT
	CRISPR		1798696	1798841	1	TCCTTATGCTTTAAACGCCTAAATCTGTGATTTAGCTGCGTTTTGTCGTATAA
	CRISPR		1910685	1910831	1	AATGAACGCTTCCTTGTCAGCAATCTTGATTAAGATTGTGCGTTTATTATACCA
	CRISPR		1927335	1927459	1	CCGTTCATCTTTCGGCGTTGCCGCAAGGTAGGAA
	CRISPR		2025969	2026111	1	TGTTCTGCTTAACTGACTCCATATCCAGCCCTCTGTGTTACGCAGATTT

The type strain Marseille-P2698 (= CSUR P2698 = DSM 103121) was iso-lated from a stool specimen of a 66-year-old man with diabetes mellitus.

Description of *Emergencia timonensis* gen. nov .

Emergencia timonensis (*e.mer.gen'cia* N.L. fem. n., *Emergencia,* for emergence, in reference to the discovery of emerging human bacteria).

Description of *Emergencia timonensis* sp. nov.

Emergencia timonensis (ti.mo.nen'sis. L. fem. adj., timonensis, from Timone, the name of a university hospital in Marseille, France, where the type strain was isolated).

Bacterial cells are Gram-positive rod-shaped and bacilli. They are motile, non-spore-forming, with lengths ranging from 1.0 x 1.5 µm and a diameter of 0.5 to 1.0 µm. The optimal growth conditions are 37°C, a strict anaerobic atmosphere, and pH.7. On 5% sheep, blood-enriched Columbia agar, colonies of strain

Marseille-P2698 appear small, translucent, and measure from 0.7 to 1.2 mm in diameter. Cells ferment D-glucose, D-mannitol, D-lactose, D-saccharose, D-maltose, salicin, esculin ferric citrate, glycerol, D-cellobiose, D-mannose, D-melezitose, D-sorbitol, L-rhamnose, D-trehalose, glycerol, D-arabinose, L-arabinose, D-ribose, D-xylose, D-galactose, D-glucose, D-fructose, D-mannose, L-rhamnose, dulcitol, D-mannitol, D-sorbitol, Methyl-αD-glucopyranoside, N-acetylglucosamine, amygdalin, arbutin, esculin, salicin, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, D-melezitose, D-raffinose, gentiobiose, D-turanose, D-tagalose, L-fucose, and potassium gluconate. In addition, enzymatic activities such as esterase lipase, leucine arylamidase, acid phosphatase, and naphthol phosphohydrolase are present. The major cell wall fatty acids are $C_{16:00}$ (39 %), $C_{18:1n9}$ (16 %), and $C_{18:1n7}$ (14%).

The genome size from strain Marseille-P2260 is 4.66 Mbp with a G+C content of 45.8 mol%.

The type strain Marseille-P2260 (= CSUR P2260 = DSM 101844) was isolated from a stool sample of patient with abdominal pain.

Declarations

Data Availability Statement

The genomic and 16S rRNA gene sequences of *Culturomica massiliensis* sp. nov. were deposited in the GenBank database under accession numbers FLKM00000000 and LN998061, respectively.

The genomic and 16S rRNA gene sequences of *Emergencia timonensis* sp. nov. were deposited in the GenBank database under accession numbers FLKM00000000 and LN998061, respectively.

Author Contributions

Conceptualization and methodology: D.R., P.E.F., F.F., L.H..; Software and analysis: A.H.; Validation: L.H., P.E.F., F.F., C.I.L.; Investigation: A.H., R.M.W.; Writing-original draft preparation: A.H, R.M.W, P.E.F, L.H.; Supervision, L.H., P.E.F.; Project administration, P.E.F., D.R.; Funding acquisition, D.R. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the French Government under the « Investissements d'avenir » (Investments for the Future) program managed by the Agence Nationale de la Recherche (ANR, fr: National Agency for Research), (reference: Méditerranée Infection 10-IAHU-03). This work was also supported by Région Provence Alpes Côte d'Azur and European funding FEDER PRIMI.

Transparency declaration

The authors declare that they have no competing interests. The study was approved by the ethics committee of the Institut de Recherche Fédératif 48 under Authorization number 09-022 with the consent of the patients.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Abdallah RA, Beye M, Diop A, Bakour S, Raoult D, Fournier PE. The impact of culturomics on taxonomy in clinical microbiology. *Antonie van Leeuwenhoek 2017 11010*. 2017;110(10):1327-1337. doi:10.1007/S10482-017-0871-1.
2. Akey DL, Gehret JJ, Khare D, Smith JL. Insights from the Sea: Structural Biology of Marine Polyketide Synthases. *Nat Prod Rep*. 2012;29(10):1038. doi:10.1039/C2NP20016C.
3. Anani H, Abdallah RA, Khoder M, et al. Colibacter massiliensis gen. nov. sp. nov., a novel Gram-stain-positive anaerobic diplococcal bacterium, isolated from the human left colon. *Sci Rep*. 2019;9(1). doi:10.1038/S41598-019-53791-1.
4. Bessis S, Ndongo S, Lagier JC, Fournier PE, Raoult D. "Emergencia timonensis," a new bacterium isolated from the stool of a healthy patient. *New Microbes New Infect*. 2016;12:73-75. doi:10.1016/J.NMNI.2016.05.001.
5. BioMérieux France - Diagnostic Microbiologique clinique ou industriel. Accessed February 3, 2022. https://www.biomerieux.fr/
6. BLAST: Basic Local Alignment Search Tool. Accessed August 29, 2022. https://blast.ncbi.nlm.nih.gov/Blast.cgi.
7. Bortolaia V, Kaas RS, Ruppe E, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother*. 2020;75(12):3491-3500. doi:10.1093/jac/dkaa345
8. Caporaso JG, Lauber CL, Walters WA, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*. 2012;6(8):1621-1624. doi:10.1038/ismej.2012.8.
9. Dione N, Sankar SA, Lagier JC, et al. Genome sequence and description of Anaerosalibacter massiliensis sp. nov. *New Microbes New Infect*. 2016;10:66-76. doi:10.1016/J.NMNI.2016.01.002
10. Fontaine B, Perrard A, Bouchet P. 21 years of shelf life between discovery and description of new species. *Curr Biol*. 2012; 22(22):R943-R944. doi:10.1016/J.CUB.2012.10.029.
11. Home - Assembly - NCBI. Accessed February 4, 2022. https://www.ncbi.nlm.nih.gov/assembly/

12. 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(1):1-11. doi:10.1186/1471-2105-11-119/TABLES/5
13. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*. 2018;35(6):1547. doi:10.1093/MOLBEV/MSY096
14. Lagier JC, Edouard S, Pagnier I, Mediannikov O, Drancourt M, Raoult D. Current and past strategies for bacterial culture in clinical microbiology. *Clin Microbiol Rev*. 2015; 28(1):208-236. doi:10.1128/CMR.00110-14
15. Lagier JC, Armougom F, Million M, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect*. 2012;18(12):1185-1193. doi:10.1111/1469-0691.12023
16. 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(1):237-264. doi:10.1128/CMR.00014-14/ASSET/EC9AA70E-2915-4655-BCD0-6A04B11458FC/ASSETS/GRAPHIC/ZCM0011524990005.JPEG
17. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res*. 2004;32(1):11-16. doi:10.1093/NAR/GKH152
18. 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(1):1-14. doi:10.1186/1471-2105-14-60/TABLES/2
19. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun*. 2019;10(1). doi:10.1038/S41467-019-10210-3
20. Ndongo S, Lagier J-C, Fournier P-E, Raoult D, Khelaifia S. NEW SPECIES “Culturomica massiliensis”, a new bacterium isolated from the human gut. *New Microbes New Infect*. 2016;13:60-61. doi:10.1016/j.nmni.2016.06.010
21. Ndongo S, Beye M, Dubourg G, et al. Genome analysis and description of Xanthomonas massiliensis sp. nov., a new species isolated from human faeces. *New Microbes New Infect*. 2018;26:63. doi:10.1016/J.NMNI.2018.06.005
22. Ndongo S, Maaloum M, Richez M, et al. Vitreoscilla massiliensis sp. nov., Isolated From the Stool of an Amazonian Patient. *Curr Microbiol*. 2021;78(8):3313-3320. doi:10.1007/S00284-021-02577-8
23. Parte AC, Carbasse JS, Meier-Kolthoff JP, Reimer LC, Göker M. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol*. 2020;70(11):5607-5612. doi:10.1099/IJSEM.0.004332.
24. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. *Curr Protoc Bioinforma*. 2020;70(1). doi:10.1002/CPBI.102.
25. ResFinder 4.1. Accessed May 31, 2022. https://cge.food.dtu.dk/services/ResFinder/
26. Robinson SL, Wackett LP. Rings of Power: Enzymatic Routes to β-Lactones. *Compr Nat Prod III*. Published online January 1, 2020:323-345. doi:10.1016/B978-0-12-409547-2.14634-7.
27. Sankar SA, Lagier JC, Pontarotti P, Raoult D, Fournier PE. The human gut microbiome, a taxonomic conundrum. *Syst Appl Microbiol*. 2015;38(4):276-286. doi:10.1016/J.SYAPM.2015.03.004
28. Sasser M. CULTURING OF THE BACTERIA. *Newark, NY Microb ID*. Published online 2006.
29. Seemann T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068-2069. doi:10.1093/bioinformatics/btu153
30. Seng P, Drancourt M, Gouriet F, 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(4):543-551. doi:10.1086/600885
31. Stothard P, Grant JR, Van Domselaar G. Visualizing and comparing circular genomes using the CGView family of tools. *Brief Bioinform*. 2018;20(4):1576-1582. doi:10.1093/bib/bbx081
32. Tatusov RL, Galperin MY, Natale DA, Koonin E V. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*. 2000;28(1):33-36. doi:10.1093/NAR/28.1.33
33. Wang J, Shi Y, Jiang D. β-Lactone Derivatives and Their Anticancer Activities: A Short Review. *Curr Top Med Chem*. 2021;21(18):1645-1656. doi:10.2174/1568026621666210402142150.
34. Zgheib R, Hasni I, Mbaye B, et al. Buttiauxella massiliensis sp. nov., Isolated from a Human Bone Infection. *Curr Microbiol*. 2022;79(2). doi:10.1007/S00284-021-02714-3

Figures

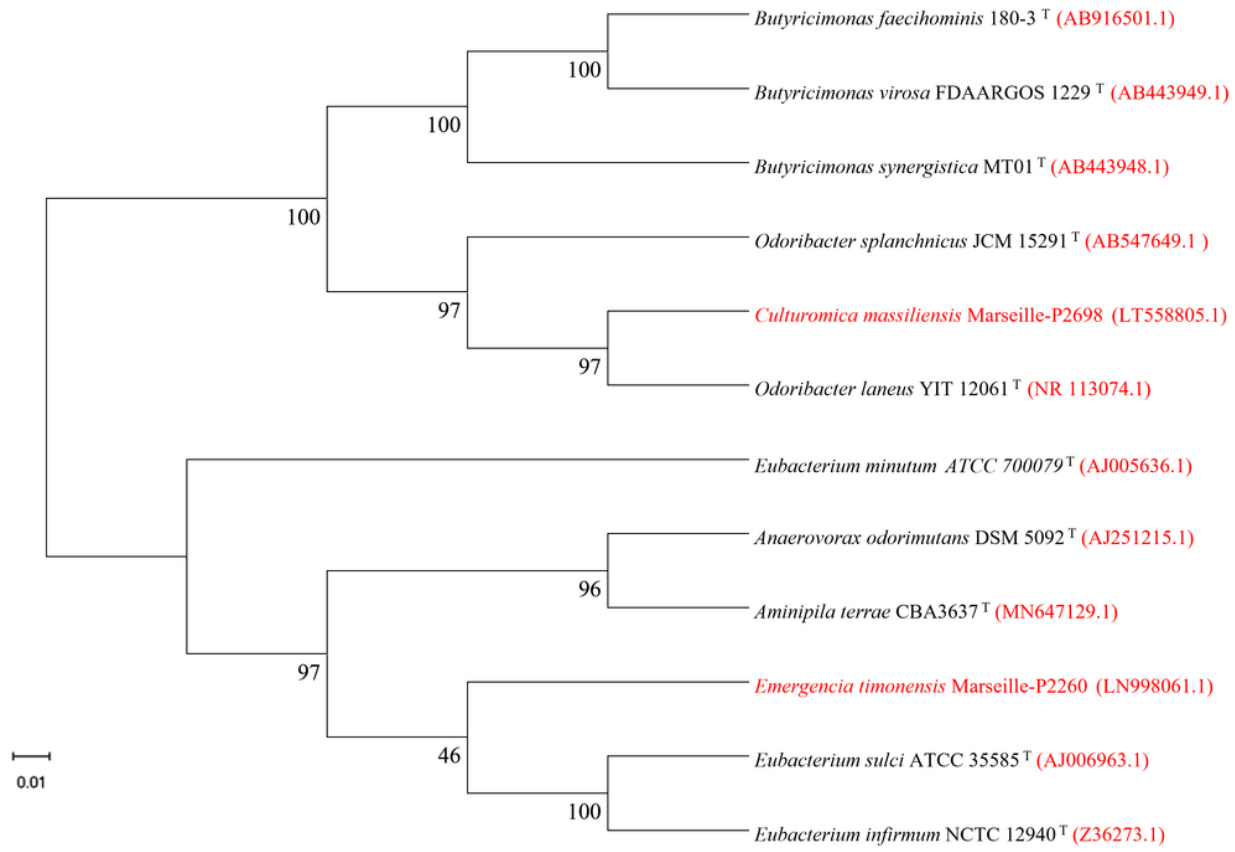

Figure 1

Phylogenetic tree with the position of new species (strains Marseille-P2698 and Marseille-P2260) among closely related species. The following phylogenetic tree was performed from the comparison of 16S rRNA sequences. The accession numbers of 16S rRNA gene are mentioned in parentheses. Bootstrap appears at the nodes. MUSCLE software was used to align sequences. The tree was designed with the MEGA-X software. The used methodology is the Maximum Likelihood method and Kimura 2-parameter model.

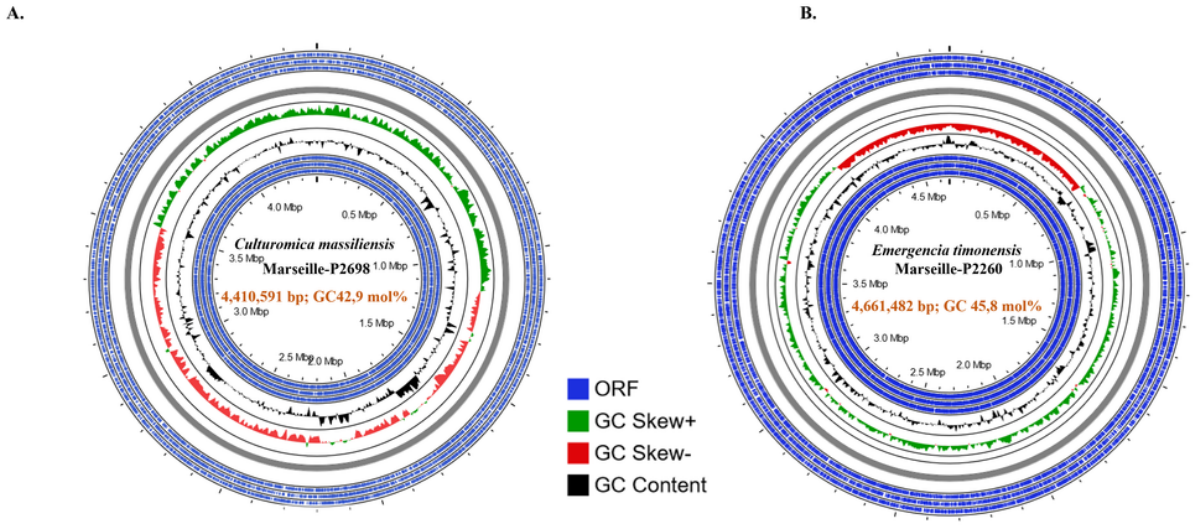

Figure 2

Circular genome map of strains Marseille-P2698 (left) and Marseille-P2260 (right) generated by the CGView software. From outside to the center: blue rings demonstrate the ORFs (Open Reading Frames) on both forward and reverse strands, green and red rings represent both positive and negative GC skew respectively, and black ring represents the GC content plot.

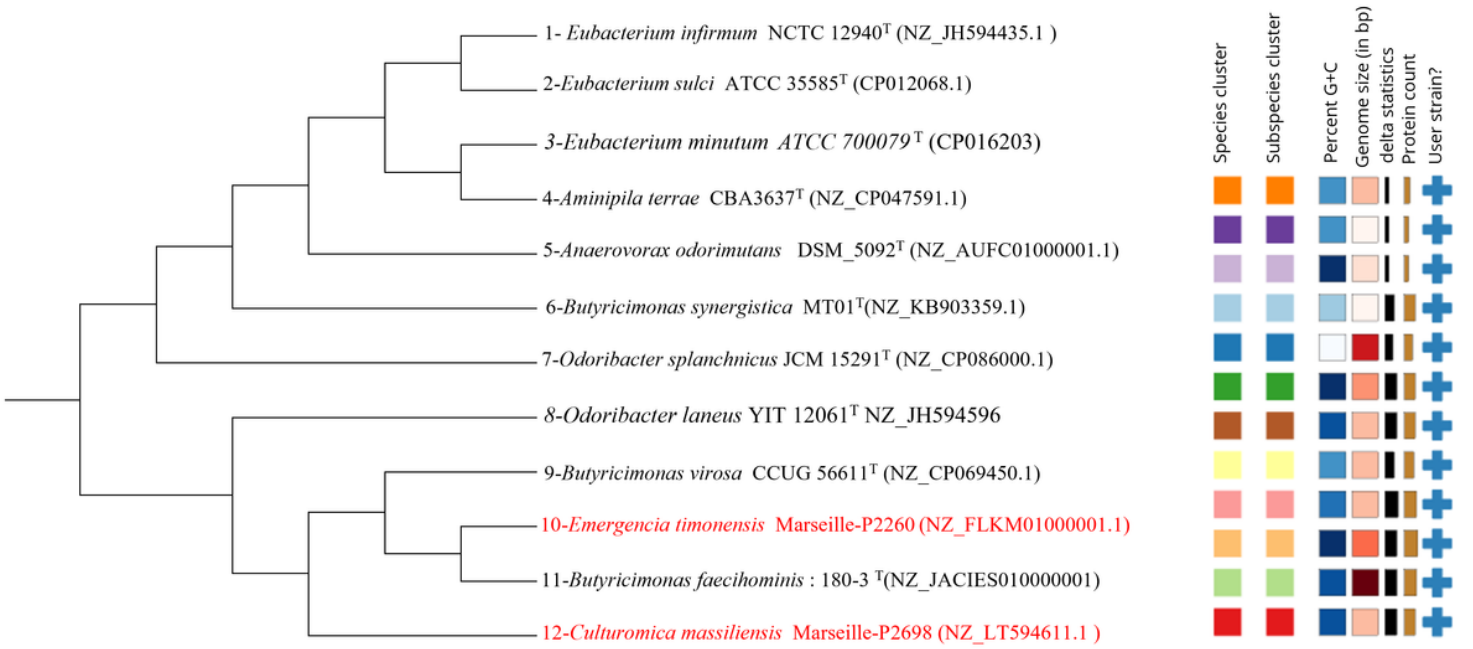

Figure 3

Whole-Genome sequence-based phylogenetic tree of strains Marseille-P2698 and Marseille-P2260 using TYGS design tree with FastME 2.1.4 software based on Genome BLAST Distance Phylogeny (GBDP) parameters. Distances were calculated from the genome sequences, and branch lengths were calculated by GBDP distance formula d5.

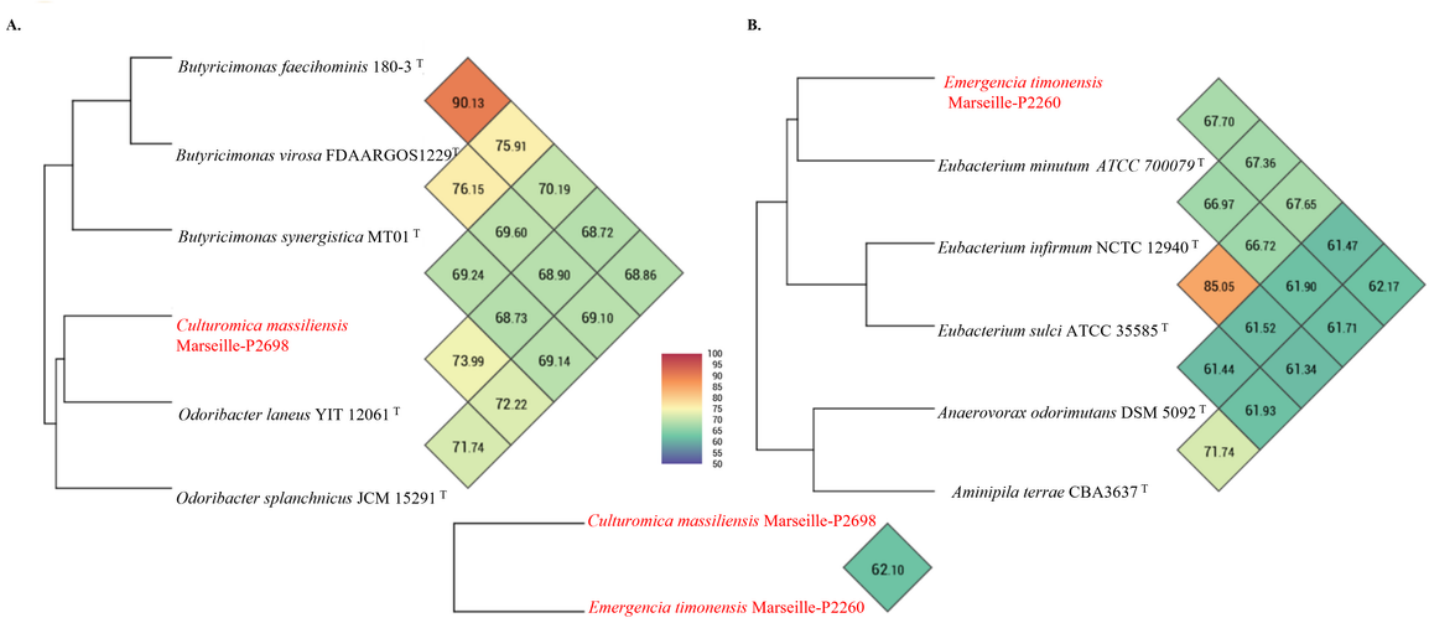

Figure 4

Heatmap showing the average nucleotide identity by orthology (OrthoANI) between Mar-seille-P2698 (A) and Marseille-P2260 (B) compared to their closest related bacterial species.

Genes encoding proteins are divided into several categories according to their functions and others have unknown functions. COGs of strains Marseille-P2698 and Marseille-P2260 have a different functional distributions from their closest species (**Fig.5**).

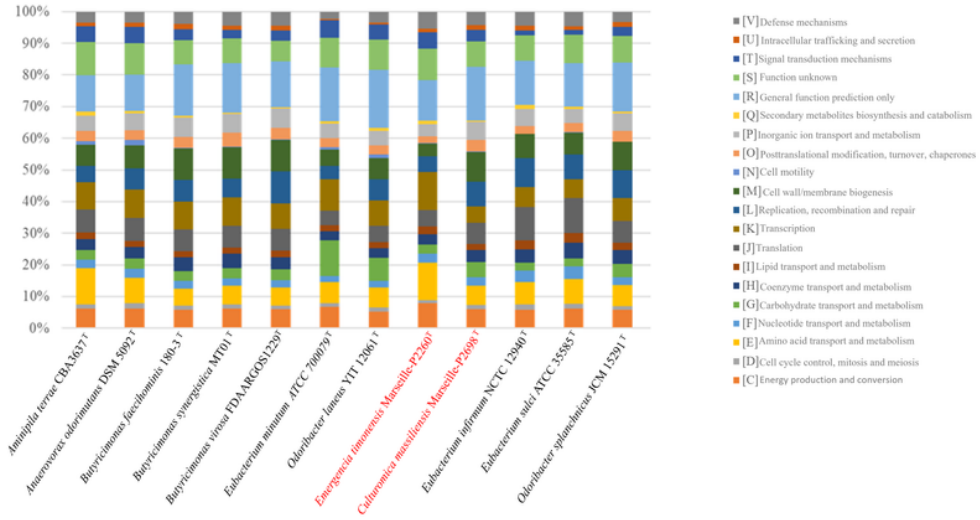

Figure 5

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