Genome sequence and description of *Nesterenkonia massiliensis* sp. nov. strain NP1<T>

Sophie Edouard<sup>1</sup>*, Senthil Sankar<sup>1</sup>*, Nicole Prisca Makaya Dangui<sup>1</sup>, Jean-Christophe Lagier<sup>1</sup>, Caroline Michelle<sup>1</sup>, Didier Raoult<sup>1,2</sup> and Pierre-Edouard Fournier<sup>1</sup>††

<sup>1</sup>Aix-Marseille Université, Faculté de médecine, Marseille, France.
<sup>2</sup>Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

††These 2 authors contributed equally

*Correspondence: pierre-edouard.fournier@univ-amu.fr

Keywords: *Nesterenkonia massiliensis*, genome, culturomics, taxono-genomics

*Nesterenkonia massiliensis* sp. nov., strain NP1<T>, is the type strain of *Nesterenkonia massiliensis* sp. nov., a new species within the genus *Nesterenkonia*. This strain, whose genome is described here, was isolated from the feces of a 32-year-old French woman suffering from AIDS and living in Marseille. *Nesterenkonia massiliensis* is a Gram-positive aerobic coccus. Here, we describe the features of this bacterium, together with the complete genome sequencing and annotation. The 2,726,371 bp long genome (one chromosome but no plasmid) contains 2,663 protein-coding and 51 RNA genes, including 1 rRNA operon.

Introduction

*Nesterenkonia massiliensis* strain NP1<T> (= CSUR P244 = DSM 26221) is the type strain of *N. massiliensis* sp. nov. This bacterium is a Gram-positive, non-spore-forming, aerobic and motile coccus that was isolated from the fecal flora of AIDS-infected French female living in Marseille, France, as part of a “culturomics” study aiming at cultivating as many of the bacterial species in human feces as possible [1,2].

Taking advantage of the availability of more than 12,000 bacterial genome sequences [3], we recently proposed to use genomic properties in combination with phenotypic characteristics for the taxonomic classification of *Bacteria* [4-34].

Herein, we present a summary classification and a set features for *Nesterenkonia massiliensis* sp. nov., strain NP1<T> (CSUR= P244 = DSM 26221), including the description of its complete genome and annotation. These characteristics support the circumscription of the species *Nesterenkonia massiliensis*.

The genus *Nesterenkonia* was first described by Stackebrandt *et al.* in 1995 [35]. This genus belongs to the family *Micrococcaceae* within the phylum *Actinobacteria*, and is most closely related to the genera *Micrococcus*, *Arthrobacter* and *Kocuria* [35]. The *Nesterenkonia* genus includes Gram-positive, non spore-forming aerobic, mesophilic bacteria that may be halotolerant or halophilic. Currently, the genus *Nesterenkonia* includes 12 species with validly published names [36]. Members of the genus *Nesterenkonia* are ubiquitous bacteria, which have been isolated from various environments including hypersaline soil and lakes, soda lakes, sea food, and paper and cotton pulp mills [36-40]. However, prior to our study, *Nesterenkonia* species had not been reported in humans, with the exception of DNA sequences from *Nesterenkonia* sp. detected in the gut microbiota of patients with chronic kidney diseases [41].

Classification and features

A stool sample was collected from a Caucasian, AIDS-infected, 32-year-old French woman living in Marseille, France. The patient gave an informed and signed consent. This study and the assent procedure were approved by the Ethics Committee of the Institut Fédératif de Recherche IFR48, Faculty of Medicine, Marseille, France under agreement number 09-022. The fecal sample was preserved...
at -80°C after collection. Strain NP1 (Table 1) was first isolated in March 2012 by cultivation on Columbia agar (BioMerieux, Marcy l’Etoile, France) under aerobic conditions after 14 days of preincubation of the stool sample with addition of 5ml of sheep rumen in blood bottle culture. The strain exhibited a 96.7% 16S rRNA nucleotide sequence identity with N. alba [51], the phylogenetically most closely related Nesterenkonia species with standing in nomenclature (Figure 1). This value was lower than the 98.7% 16S rRNA gene sequence threshold recommended by Stackebrandt and Ebers to delineate a new species without carrying out DNA-DNA hybridization [52].

| MIGS ID | Property | Term | Evidence code* |
|---------|----------|------|----------------|
| MIGS-6.3 | Salinity | Halophilic | IDA |
| MIGS-22 | Oxygen requirement | Aerobic | IDA |
| MIGS-6 | Habitat | Human gut | IDA |
| MIGS-15 | Biotic relationship | Free living | IDA |
| MIGS-14 | Pathogenicity | Unknown | IDA |
| MIGS-4 | Geographic location | Marseille, France | IDA |
| MIGS-5 | time | March 2012 | IDA |
| MIGS-4.1 | Longitude | 5.36978 | IDA |
| MIGS-4.3 | Depth | Surface | IDA |
| MIGS-4.4 | Altitude | 0 m above sea level | IDA |

*Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene
Onontology project [50]. If the evidence is IDA, then the property was directly observed for a live isolate by one of the authors or an expert mentioned in the acknowledgements.

**Figure 1**: Consensus phylogenetic tree highlighting the position of *Nesterenkonia massiliensis* strain NP1 relative to other type strains within the *Nesterenkonia* genus. Genbank accession numbers are indicated for each species. Sequences were aligned using CLUSTALW and phylogenetic inferences obtained using the maximum-likelihood method in the MEGA software package. Numbers at the nodes are percentages of bootstrap values from 500 replicates that support the node. *Micrococcus luteus* was used as the outgroup. The scale bar represents a 5% nucleotide sequence divergence.
Four different growth temperatures (25, 30, 37, 45°C) were tested. Growth was observed between 25 and 45°C on blood-enriched Columbia agar (BioMerieux), with optimal growth occurring at
37°C after 24 hours of incubation. Colonies were dark yellow and 1 mm in diameter. Growth of the strain was tested under anaerobic and microaerophilic conditions using GENbag anaer and GENbag microaer systems, respectively (BioMerieux), and under aerobic conditions, with or without 5% CO₂. Optimal growth was obtained under aerobic conditions, but weak growth occurred in a microaerophilic atmosphere. No growth was observed under anaerobic conditions. Bacterial cells were Gram-positive (Figure 2), non-endospore-forming, and motile cocci. Cells grown on agar had a mean diameter and length of 0.67 μm and 1.4 μm, respectively (Figure 3).

Strain NP1ᵀ exhibited catalase but no oxidase activity. Using an API 20NE strip (BioMerieux), negative reactions were obtained for nitrate reduction, urease, indole production, glucose fermentation, arginine dihydrolase, β-galactosidase, glucose, arabinose, mannose, mannitol, N-acetyl-glucosamine, maltose, gluconate, caprate, adipate, malate, citrate, phenyl-acetate assimilation and cytochrome oxidase. Substrate oxidation and assimilation were examined with an API 50CH carbohydrate fermentation strip (BioMerieux) at 37°C. Positive reactions were observed for D-glucose, D-fructose, D-saccharose, ribose, mannose, mannitol, D-trehalose and L-rhamnose. No reaction was observed for esculin, salicin, D-cellobiose and gentiobiose. N. massiliensis is susceptible to amoxicillin, imipenem, rifampin, ciprofloxacin, gentamicin, doxycycline and vancomycin but resistant to trimethoprim/sulfamethoxazole and metronidazole. When compared with representative species from the genus Nesterenkonia, N. massiliensis strain NP1ᵀ differed in cell shape, colony color, motility, optimal growth temperature, and mannitol fermentation (Table 2).

| N. massiliensis | N. alba | N. flava | N. lacusekhoensis |
|----------------|--------|---------|------------------|
| Cell diameter (μm) | 0.7x1.4 | 0.4-0.6 x 0.8-1.2 | 0.6-0.8 x 1.2 | 0.8-0.9 x 1-1.3 |
| Oxygen requirement | Aerobic | Aerobic | Aerobic | Aerobic |
| Pigment production | Dark Yellow | White | Yellow | Bright yellow |
| Gram stain | cocci | Short rods | Short rods | Short rods |
| Optimal temperature (°C) | 37°C | 42 | 40-42 | 27-33.5 |
| Salt requirement NaCl tolerance (%) | 0-3 | 0-6 | 0-10 | 0-15 |
| Motility | + | - | - | - |
| Endospore formation | - | - | - | - |
| H2S production | - | - | - | W |
| Indole production | - | - | - | - |
| ONPG test* | NA | + | - | NA |
| Citrate test | NA | - | - | W |
| Voges–Proskauer reaction | NA | - | - | - |
| Production of Catalase | + | + | + | + |
| Oxidase | - | - | - | - |
| Nitrate reductase | - | - | - | - |
| Urease | - | - | - | NA |
| β-Galactosidase | - | NA | - | NA |
| Acid from L-arabinose | - | + | + | NA |
| Mannose | - | - | + | + |
| Mannitol | + | NA | - | - |
| Sucrose | - | - | + | + |
| D-glucose | + | + | + | + |
| D-fructose | + | - | + | + |
| D-maltose | + | - | + | + |
| D-lactose | - | - | - | - |
| D-Galactose | - | - | - | - |
| Trehalose | - | W | - | + |
D-Xylose: -
Hydrolysis of:
Starch: -
Gelatin: -
G+C content (mol%): 62.47, 60.2, 65.5, 66.1

| Habitat | Human gut, France | Black liquor treatment system of a cotton pulp mill, China | Paper mill effluent, China | Hypersaline lake, East Antarctica |
|---------|------------------|-------------------------------------------------|-------------------------|----------------------------------|

+, Positive; −, negative; W, weak reaction; NA, not available.

* N. massiliensis strain NP1T, N. alba strain DSM 19423, N. flava strain JCM 14814T, N. lacusekhoensis strain DSM 12544

Matrix-assisted laser-desorption/ionization time-of-flight (MALDI-TOF) MS protein analysis was carried out as previously described [53] using a Microflex spectrometer (Brüker Daltonics, Leipzig, Germany). Twelve individual NP1T colonies were deposited on a MTP 384 MALDI-TOF target plate (Brüker). The twelve spectra were imported into the MALDI BioTyper software (version 2.0, Brüker) and analyzed by standard pattern matching (with default parameter settings) against the main spectra of 4,706 bacteria, including 1 spectrum from N. lacusekhoensis, the only validly named Nesterenkonia species for which a spectrum was available in the BioTyper database. A score enabled the presumptive identification and discrimination of the tested species from those in a database: a score > 2 with a validated species enabled the identification at the species level; and a score < 1.7 did not enable any identification. For strain NP1T, no significant score was obtained, suggesting that our isolate was not a member of any known species (Figures 4 and 5).

Figure 4. Reference mass spectrum from Nesterenkonia massiliensis strain NP1T. Spectra from 12 individual colonies were compared and a reference spectrum was generated.
Figure 5. Gel view comparing *N. massiliensis* strain NP1\(^T\), *N. lacusekhoensis* and other species belonging to the family *Micrococcaceae*. The gel view displays the raw spectra of loaded spectrum files arranged as a pseudo-electrophoretic gel. 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. The grey scale on the right y-axis indicates the relative peak intensity in arbitrary units. Species names are on the left.
**Figure 6.** Graphical circular map of the chromosome. From the outside in, the outer two circles show open reading frames oriented in the forward (colored by COG categories) and reverse (colored by COG categories) directions, respectively. The third circle marks the rRNA gene operon (red) and tRNA genes (green). The fourth circle shows the G+C% content plot. The inner-most circle shows GC skew, purple indicating negative values and olive, positive values.

**Genome sequencing information**

**Genome project history**
The organism was selected for sequencing on the basis of its phylogenetic position, 16S rDNA similarity and phenotypic differences with other members of the genus *Nesterenkonia*, and is part of a “culturomics” study aiming at isolating individually all bacterial species within the human gut flora [1]. It was the third genome of a *Nesterenkonia* species and the first genome of *N. massiliensis* sp. nov. A summary of the project information is shown in Table 3. The Genbank accession number is CBLL000000000 and consists of 141 contigs. Table 3 shows the project information and its association with MIGS version 2.0 compliance [42].
Growth conditions and DNA isolation

*N. massiliensis* sp. nov strain NP1, (= CSUR P244 = DSM 26221) was grown aerobically on sheep blood-enriched Columbia agar at 37°C. Two petri dishes were spread and the colonies resuspended in 6x100 μl of G2 buffer (EZI DNA Tissue Kit, Qiagen). A first mechanical lysis was performed with glass powder on the Fastprep-24 device (Sample Preparation system, MP Biomedicals, USA) using 2x20 second bursts. DNA was incubated with 2.5 μg/μL of lysozyme for 30 minutes at 37°C and extracted using the BioRobot EZ 1 Advanced XL (Qiagen). The DNA was then concentrated and purified on a Qiamp kit (Qiagen). The DNA concentration measured by the Quantit Picogreen kit (Invitrogen) on the Genios Tecan fluorometer was 94 ng/μl.

Genome sequencing and assembly

DNA (5 μg) was mechanically fragmented with a Hydroshear device (Digilab, Holliston, MA, USA) with an enrichment size at 3-4 kb. The DNA fragmentation was visualized through the Agilent 2100 BioAnalyzer on a DNA labchip 7500 with an average size of 3.5 kb. The library was constructed using a 454 GS-FLX Titanium paired-end rapid library protocol. Circularization and nebulization generated a pattern with an average size of 390 bp. After 20 cycles of PCR amplification, the double stranded paired-end library was quantified using the Quant-it Ribogreen kit (Invitrogen) on the Genios Tecan fluorometer at 820 pg/μL. The library concentration equivalence was calculated as 3.86 x 10⁹ molecules/μL. The library was stored at -20°C until further use.

The library was clonally amplified with 1 cpb in two emPCR reactions with the GS Titanium SV emPCR Kit (Lib-L) v2. The yield of the emPCR was 13.67%, within the range of 5 to 20% recommended for the Roche procedure. Approximately 656,601 beads were loaded for a ¼ region on the GS Titanium PicoTiterPlate (PTP Kit 70x75, Roche) and sequenced with the GS Titanium Sequencing Kit XLR70. The run was performed overnight and analyzed on the cluster through the gsRunBrowser and Newbler assemblers (Roche). A total of 176,833 passed filter wells were obtained and generated 60.9 Mb with an average of length of 345 bp. The passed filter sequences were assembled on Newbler with 90% identity and 40 bp as overlap. The final assembly identified 18 scaffolds and 141 large contigs (>1,500 bp), and generated a genome size of 2.69Mb which corresponds to a coverage of 22.63 genome equivalents.

Genome annotation

Open Reading Frames (ORFs) were predicted using Prodigal [54] 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 [55] and the Clusters of Orthol-
ogous Groups (COG) databases using BLASTP. The tRNAscanSE tool [56] was used to find tRNA genes, whereas ribosomal RNAs were found by using RNAmmer [57] and BLASTn against the GenBank database. Lipoprotein signal peptides and the number of transmembrane helices were predicted using SignalP [58] and TMHMM [59] respectively. ORFans were identified if their BLASTP E-value was lower than 1e⁻³ for alignment length greater than 80 amino acids. If alignment lengths were smaller than 80 amino acids, we used an E-value of 1e⁻⁵. Such parameter thresholds have already been used in previous works to define ORFans. Artemis [60] and DNA Plotter [61] were used for data management and visualization of genomic features, respectively. The Mauve alignment tool (version 2.3.1) was used for multiple genomic sequence alignment [62]. To estimate the mean level of nucleotide sequence similarity at the genome level, we used the Average Genomic Identity Of gene Sequences (AGIOS) home-made software [34]. Briefly, this software combines the Proteinortho software [63] for detecting orthologous proteins in pairwise comparisons of genomes, then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman-Wunsch global alignment algorithm. As only one genome was available for the genus *Nesterenkonia*, we used genomes from closely related genera for the calculation of AGIOS values. *N. massiliensis strain* NT1ᵀ was compared to *Nesterenkonia alba* strain DSM 19423 (GenBank accession number ATXP00000000), *Micrococcus luteus* strain NCTC2665 (CP001628), *Kocuria rhizophila* strain DSM 2048 (AP009152) and *Arthrobacter arilaitensis* strain RE117 (FQ311875).

### Genome properties

The genome of *N. massiliensis strain* NP1ᵀ is 2,726,371bp long (1 chromosome, but no plasmid) with a 62.47% G+C content (Table 4, Figure 6, Figure 7). Of the 2,714 predicted genes, 2,663 were protein-coding genes, and 51 were RNAs. Three rRNA genes (one 16S rRNA, one 23S rRNA and one 5S rRNA) and 48 predicted tRNA genes were identified in the genome. A total of 1,962 genes (73.68%) were assigned a putative function. One hundred and ninety-nine genes were identified as ORFans (7.47%). The remaining genes were annotated as hypothetical proteins. The properties and the statistics of the genome are summarized in Table 4. The distribution of genes into COGs functional categories is presented in Table 5 and a comparison is presented in Table 6.

| Attribute                              | Value    | % of total |
|----------------------------------------|----------|------------|
| Genome size (bp)                       | 2,726,371| 100        |
| DNA coding region (bp)                 | 2,473,018| 90.40      |
| DNA G+C content (bp)                   | 1,703,271| 62.47      |
| Total genes                            | 2,714    | 100        |
| RNA genes                              | 51       | 1.88       |
| Protein-coding genes                   | 2,663    | 98.12      |
| Genes with function prediction (Cog + NR) | 1,962     | 73.68      |
| Genes assigned to COGs                 | 1,950    | 73.23      |
| Genes with peptide signals             | 241      | 8.90       |
| Genes with transmembrane helices       | 599      | 22.49      |

*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.
Nesterenkonia massiliensis

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

| Code | Value | %age | Description                                      |
|------|-------|------|--------------------------------------------------|
| J    | 149   | 5,60 | Translation                                      |
| A    | 1     | 0,04 | RNA processing and modification                  |
| K    | 162   | 6,08 | Transcription                                   |
| L    | 188   | 7,06 | Replication, recombination and repair            |
| B    | 1     | 0,04 | Chromatin structure and dynamics                 |
| D    | 20    | 0,75 | Cell cycle control, mitosis and meiosis          |
| Y    | 0     | 0    | Nuclear structure                                |
| V    | 36    | 1,35 | Defense mechanisms                               |
| T    | 78    | 2,93 | Signal transduction mechanisms                   |
| M    | 112   | 4,21 | Cell wall/membrane biogenesis                    |
| N    | 3     | 0,11 | Cell motility                                   |
| Z    | 0     | 0    | Cytoskeleton                                    |
| W    | 0     | 0    | Extracellular structures                         |
| U    | 29    | 1,09 | Intracellular trafficking and secretion          |
| O    | 71    | 2,67 | Posttranslational modification, protein turnover, chaperones |
| C    | 129   | 4,84 | Energy production and conversion                 |
| G    | 148   | 5,56 | Carbohydrate transport and metabolism            |
| E    | 239   | 8,97 | Amino acid transport and metabolism              |
| F    | 67    | 2,52 | Nucleotide transport and metabolism              |
| H    | 88    | 3,30 | Coenzyme transport and metabolism                |
| I    | 73    | 2,74 | Lipid transport and metabolism                   |
| P    | 160   | 6,01 | Inorganic ion transport and metabolism           |
| Q    | 48    | 1,80 | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 293   | 11,00| General function prediction only                 |
| S    | 142   | 5,33 | Function unknown                                 |
| _    | 713   | 26,77| Not in COGs                                     |

*The total is based on the total number of protein coding genes in the annotated genome*

Genomic comparison

We compared the genome of *N. massiliensis* strain NP1T to those of *N. alba* strain DSM 19423, *M. luteus* strain NCTC2665, *K. rhizophila* strain DSM 2048 and *A. arilaitensis* strain RE117. The draft genome of *N. massiliensis* had a larger size than those of *N. alba* and *M. luteus* (2.72, 2.59 and 2.7 Mb, respectively) but was smaller than that of *A. arilaitensis* (3.92 Mb). The G+C content of *N. massiliensis* was lower than those of *N. alba*, *M. luteus* and *K. rhizophila* (62.4, 63.8, 73 and 71.2%, respectively) but higher than that of *A. arilaitensis* (59.3%). The gene content of *N. massiliensis* was larger than those of *N. alba*, *M. luteus* and *K. rhizophila* (2,714, 2,403, 2,345 and 2,414 genes, respectively) but smaller than that of *A. arilaitensis* (3,771). In addition, *N. massiliensis* shared 1,142, 1,132, 1,087 and 1,208 orthologous genes with *N. alba*, *M. luteus* and *K. rhizophila*, respectively. *N. massiliensis* exhibited an AGIOS value
of 75.61% with \textit{N. alba}, and 68.57 to 70.5 with other members of the family \textit{Micrococcaceae}.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure7.png}
\caption{Distribution of functional classes of predicted genes on the chromosomes of \textit{Nesterenkonia massiliensis} (NM), \textit{Nesterenkonia alba} (NA), \textit{Micrococcus luteus} (ML), \textit{Kocuria rhizophila} (KR) and \textit{Arthrobacter arilaitensis} (AA) according to the COG category. For each genome, we indicated the percentage of each gene category.}
\end{figure}

\begin{table}
\centering
\begin{tabular}{lcccc}
\hline
 & \textit{Nesterenkonia massiliensis} & \textit{Nesterenkonia alba} & \textit{Micrococcus luteus} & \textit{Kocuria rhizophila} & \textit{Arthrobacter arilaitensis} \\
\hline
\textit{Nesterenkonia massiliensis} & \textbf{2,663} & 1,142 & 1,132 & 1,087 & 1,208 \\
\textit{Nesterenkonia alba} & 75.61 & \textbf{2,351} & 1,046 & 1,033 & 1,116 \\
\textit{Micrococcus luteus} & 70.5 & 1,046 & \textbf{2,236} & 1,152 & 1,241 \\
\textit{Kocuria rhizophila} & 69.76 & 69.96 & 74.23 & \textbf{2,357} & 1,224 \\
\textit{Arthrobacter arilaitensis} & 68.57 & 68.04 & 69.35 & 69.05 & \textbf{3,436} \\
\hline
\end{tabular}
\caption{Genomic comparison of \textit{N. massiliensis} and 4 other members of the family \textit{Micrococcaceae}.}
\end{table}

\textsuperscript{1}Upper right, numbers of orthologous proteins shared between genomes; lower left, AGIOS values; bold numbers indicate the numbers of proteins per genome.

\section*{Conclusion}

http://standardsingenomics.org
On the basis of phenotypic, phylogenetic and genomic analyses, we formally propose the creation of *Nesterenkonia massiliensis* sp. nov. that contains the strain NP1<sup>T</sup>. The strain has been isolated from the fecal flora of an AIDS-infected patient living in Marseille, France. Several other bacterial species were also cultivated from different fecal samples through diversification of culture conditions [4-34], thus suggesting that the human fecal flora of humans remains only partially known.

**Description of Nesterenkonia massiliensis sp. nov.**

*Nesterenkonia massiliensis* (mas.si.li.en’sis. L gen. fem. n. massiliensis of *Massilia*, the Roman name of Marseille, France, where the type strain was isolated).

Grows between 25 and 45°C on blood-enriched Columbia agar (BioMerieux). Optimal growth obtained at 37°C in aerobic atmosphere. Weak growth in microaerophilic atmosphere. No growth under anaerobic condition. Colonies are dark yellow and 1 mm in diameter. Cells are Gram-positive, non-endospore-forming, and motile cocci, with a mean diameter and length of 0.67 µm and 1.4 µm, respectively. Catalase positive, oxidase negative. Negative reactions obtained for nitrate reduction, urease, indole production, glucose fermentation, arginine dihydrolase, β-galactosidase, glucose, arabinose, mannose, mannitol, N-acetyl-glucosamine, maltose, gluconate, caprate, adipate, malate, citrate, phenylacetate assimilation and cytochrome oxidase. Fermentation of D-glucose, D-fructose, D-saccharose, ribose, mannose, mannitol, D-trehalose and L-rhamnose. No reaction observed for esculin, salicin, D-cellulbiose and gentiobiose. Cells are susceptible to amoxicillin, imipenem, rifampin, ciprofloxacin, gentamicin, doxycycline and vancomycin but resistant to trimethoprim/sulfamethoxazole and metronidazole.

The G+C content of the genome is 62.47%. The 16S rRNA and genome sequences are deposited in GenBank under accession numbers JX424770 and CBL00000000, respectively. The habitat of the microorganism is the human digestive tract. The type strain NP1<sup>T</sup> (= CSUR P244 = DSM 26221) was isolated from the fecal flora of a 32-year-old French female suffering from AIDS.

**Acknowledgements**

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

**References**

1. Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, Bittar F, Fournous G, Gimenez G, Maraninchi M, *et al.* Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* 2012; 18:1185-1193. PubMed

2. Dubourg G, Lagier JC, Armougom F, Robert C, Hamad I, Brouqui P, Raoult D. The gut microbiota of a patient with resistant tuberculosis is more comprehensively studied by culturomics than by metagenomics. *Eur J Clin Microbiol Infect* 2013; 32:637-645. PubMed

3. Database GOLD. http://www.genomesonline.org/cgi-bin/GOLD/index.cgi

4. Ramasamy D, Mishra AK, Lagier JC, Padmanabhan R, Rossi-Tamisier M, Sentausa E, Raoult D, Fournier PE. A polyphasic strategy incorporating genomic data for the taxonomic description of new bacterial species. *Int J Syst Evol Microbiol* 2014; (In press). PubMed

5. Roux V, Million M, Robert C, Magne A, Raoult D. Non-contiguous finished genome sequence and description of *Oceanobacillus massiliensis* sp. nov. *Stand Genomic Sci* 2013; 9:370-384. PubMed

6. Kokcha S, Mishra AK, Lagier JC, Million M, Leroy Q, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Bacillus timonensis* sp. nov. *Stand Genomic Sci* 2012; 6:346-355. PubMed

7. 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; 6:116-125. PubMed

8. Mishra AK, Gimenez G, Lagier JC, Robert C, Raoult D, Fournier PE. Genome sequence and description of *Alistipes senegalensis* sp. nov. *Stand Genomic Sci* 2012; 5:178-187. PubMed
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; 6:315-324. PubMed http://dx.doi.org/10.4056/sigs.2956294

10. Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Clostridium senegalense* sp. nov. *Stand Genomic Sci* 2012; 6:386-395. PubMed http://dx.doi.org/10.4056/sigs.2685971

11. Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Peptoniphilus timonensis* sp. nov. *Stand Genomic Sci* 2012; 7:1-11. PubMed http://dx.doi.org/10.4056/sigs.2956294

12. Mishra AK, Lagier JC, Rivet R, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Paenibacillus senegalensis* sp. nov. *Stand Genomic Sci* 2012; 7:70-81. PubMed http://dx.doi.org/10.4056/sigs.3056450

13. Lagier JC, Gimenez G, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Herbaspirillum massiliense* sp. nov. *Stand Genomic Sci* 2012; 7:200-209. PubMed

14. Roux V, El Karkouri K, Lagier JC, Robert C, Raoult D. Non-contiguous finished genome sequence and description of *Kurthia massiliensis* sp. nov. *Stand Genomic Sci* 2012; 7:221-232. PubMed http://dx.doi.org/10.4056/sigs.3206554

15. 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; 7:233-245. PubMed http://dx.doi.org/10.4056/sigs.3256677

16. Ramasamy D, Kokcha S, Lagier JC, Nguyen TT, Raoult D, Fournier PE. Genome sequence and description of *Aeromicrobium massiliensis* sp. nov. *Stand Genomic Sci* 2012; 7:246-257. PubMed http://dx.doi.org/10.4056/sigs.3306717

17. Lagier JC, Ramasamy D, Rivet R, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Cellulomonas massiliensis* sp. nov. *Stand Genomic Sci* 2012; 7:258-270. PubMed http://dx.doi.org/10.4056/sigs.3316719

18. Lagier JC, Elkarkouri K, Rivet R, Couderc C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Senegalemassilia anaerobia* gen. nov., sp. nov. *Stand Genomic Sci* 2013; 7:343-356. PubMed http://dx.doi.org/10.4056/sigs.3246665

19. Mishra AK, Hugon P, Lagier JC, Nguyen TT, Robert C, Couderc C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Peptoniphilus obesi* sp. nov. *Stand Genomic Sci* 2013; 7:337-369. PubMed http://dx.doi.org/10.4056/sigs.32766871

20. Mishra AK, Lagier JC, Nguyen TT, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Peptoniphilus senegalenensis* sp. nov. *Stand Genomic Sci* 2013; 7:370-381. PubMed http://dx.doi.org/10.4056/sigs.3366764

21. 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. PubMed http://dx.doi.org/10.4056/sigs.3396830

22. Hugon P, Ramasamy D, Lagier JC, Rivet R, Couderc C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Alistipes obesi* sp. nov. *Stand Genomic Sci* 2013; 7:427-439. PubMed http://dx.doi.org/10.4056/sigs.3336746

23. Mishra AK, Hugon P, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Peptoniphilus grossensis* sp. nov. *Stand Genomic Sci* 2012; 7:320-330. PubMed

24. Mishra AK, Hugon P, Lagier JC, Nguyen TT, Couderc C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Enorma massiliensis* gen. nov., sp. nov., a new member of the Family *Coriobacteriaceae*. *Stand Genomic Sci* 2013; 8:290-305. PubMed http://dx.doi.org/10.4056/sigs.3426906

25. Ramasamy D, Lagier JC, Gorlas A, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Bacillus massiliosenegalensis* sp. nov. *Stand Genomic Sci* 2013; 8:264-278. PubMed http://dx.doi.org/10.4056/sigs.3496989

26. Ramasamy D, Lagier JC, Nguyen TT, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Dielma fastidiosa* gen. nov., sp. nov., a new member of the Family *Erysipelotrichaceae*. *Stand Genomic Sci* 2013;
Nesterenkonia massiliensis

8:336-351. PubMed http://dx.doi.org/10.4056/sigs.3567059

27. Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Genome sequence and description of Timonella senegalensis gen. nov., sp. nov., a new member of the suborder Micrococcineae. Stand Genomic Sci 2013; 8:318-335. PubMed http://dx.doi.org/10.4056/sigs.3476977

28. Mishra AK, Pfeiderer A, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous-finished genome sequence and description of Bacillus massilienox sp. nov. Stand Genomic Sci 2013; 8:465-479. PubMed http://dx.doi.org/10.4056/sigs.4087826

29. Hugon P, Mishra AK, Lagier JC, Nguyen TT, Couderc C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of Brevibacillus massiliensis sp. nov. Stand Genomic Sci 2013; 8:1-14. PubMed http://dx.doi.org/10.4056/sigs.3466975

30. Hugon P, Mishra AK, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of Anaerococcus vaginalis. Stand Genomic Sci 2012; 6:356-365. PubMed http://dx.doi.org/10.4056/sigs.2716452

31. Hugon P, Ramasamy D, Robert C, Couderc C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of Kallipyga massiliensis gen. nov., sp. nov., a new member of the family Clostridiales Incertae Sedis XI. Stand Genomic Sci 2013; 8:500-515. PubMed http://dx.doi.org/10.4056/sigs.4047997

32. Padmanabhan R, Lagier JC, Dangui NPM, Michelle C, Couderc C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of Megasphaera massiliensis. Stand Genomic Sci 2013; 8:525-538. PubMed http://dx.doi.org/10.4056/sigs.4077819

33. Mishra AK, Edouard S, Dangui NPM, Lagier JC, Caputo A, Blanc-Tailleur C, Ravaux I, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of Nosocomicoccus massiliensis sp. nov. Stand Genomic Sci 2013; 9:205-219. PubMed http://dx.doi.org/10.4056/sigs.4378121

34. Mishra AK, Lagier JC, Pfeiderer A, Nguyen TT, Caputo A, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of Holdemania massiliensis sp. nov. Stand Genomic Sci 2013; 9:395-409. PubMed http://dx.doi.org/10.4056/sigs.4628316

35. Stackebrandt E, Koch C, Gvozdiak O, Schumann P. Taxonomic dissection of the genus Micrococcus: Kocuria gen. nov., Nesterenkonia gen. nov., Kyotococcus gen. nov., Dermacoccus gen. nov., and Micrococcus Cohn 1872 gen. emend. Int J Syst Bacteriol 1995; 45:682-692. PubMed http://dx.doi.org/10.1099/00207713-45-4-682

36. Govender L, Naidoo L, Setati ME. Nesterenkonia suensis sp. nov., a haloalkaliphilic actinobacterium isolated from a salt pan. Int J Syst Evol Microbiol 2013; 63:41-46. PubMed http://dx.doi.org/10.1099/ijsem.0.035006-0

37. Collins MD, Lawson PA, Labrenz M, Tindall BJ, Weiss N, Hirsch P. Nesterenkonia lacusekhoensis sp. nov., isolated from hypersaline Ekho Lake, East Antarctica, and emended description of the genus Nesterenkonia. Int J Syst Evol Microbiol 2002; 52:1145-1150. PubMed http://dx.doi.org/10.1099/ijsem.0.02118-0

38. Li WJ, Zhang YQ, Schumann P, Liu HY, Yu LY, Zhang YQ, Stackebrandt E, Xu LH, Jiang CL. Nesterenkonia halophila sp. nov., a moderately halophilic, alkalitolerant actinobacterium isolated from a saline soil. Int J Syst Evol Microbiol 2008; 58:1359-1363. PubMed http://dx.doi.org/10.1099/ijsem.0.064226-0

39. Li WJ, Chen HH, Kim CJ, Zhang YQ, Park DJ, Lee JC, Xu LH, Jiang CL. Nesterenkonia sandarakina sp. nov. and Nesterenkonia lutea sp. nov., novel actinobacteria, and emended description of the genus Nesterenkonia. Int J Syst Evol Microbiol 2005; 55:463-466. PubMed http://dx.doi.org/10.1099/ijsem.0.063281-0

40. Li WJ, Chen HH, Zhang YQ, Schumann P, Stackebrandt E, Xu LH, Jiang CL. Nesterenkonia halotolerans sp. nov. and Nesterenkonia xinjiangensis sp. nov., actinobacteria from saline soils in the west of China. Int J Syst Evol Microbiol 2004; 54:837-841. PubMed http://dx.doi.org/10.1099/ijsem.0.02935-0

41. Vaziri ND, Wong J, Pahl M, Piceno YM, Yuan J, DeSantis TZ, Ni Z, Nguyen TH, Andersen GL. Chronic kidney disease alters intestinal microbial flora. Kidney Int 2013; 83:308-315. PubMed http://dx.doi.org/10.1038/ki.2012.345

42. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. Nat Biotechnol 2008; 26:541-547. PubMed http://dx.doi.org/10.1038/nbt1360
43. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaebacteria*, *Bacteria*, and *Eucarya*. Proc Natl Acad Sci USA 1990; **87**:4576-4579. PubMed [http://dx.doi.org/10.1073/pnas.87.12.4576](http://dx.doi.org/10.1073/pnas.87.12.4576)

44. Garrity GM, Holt JG. The Road Map to the Manual. In: Garrity GM, Boone DR, Castenholz RW (eds), Bergey’s Manual of Systematic Bacteriology, Second Edition, Volume 1, Springer, New York, 2001, p. 119-169.

45. Stackebrandt E, Rainey FA, Ward-Rainey NL. Proposal for a New Hierarchic Classification System, *Actinobacteria* classis nov. *Int J Syst Bacteriol* 1997; **47**:479-491. PubMed [http://dx.doi.org/10.1099/00207713-47-2-479](http://dx.doi.org/10.1099/00207713-47-2-479)

46. Zhi XY, Li WJ, Stackebrandt E. An update of the structure and 16S rRNA gene sequence-based definition of higher ranks of the class *Actinobacteria*, with the proposal of two new suborders and four new families and emended descriptions of the existing higher taxa. *Int J Syst Evol Microbiol* 2009; **59**:589-608. PubMed [http://dx.doi.org/10.1099/ijs.0.65780-0](http://dx.doi.org/10.1099/ijs.0.65780-0)

47. Skerman VBD, McGowan V, Sneath PHA. Approved Lists of Bacterial Names. *Int J Syst Bacteriol* 1980; **30**:225-420. PubMed [http://dx.doi.org/10.1099/00207713-30-1-225](http://dx.doi.org/10.1099/00207713-30-1-225)

48. Buchanan RE. Studies in the Nomenclature and Classification of the Bacteria: II. The Primary Subdivisions of the *Schizomycetes*. *J Bacteriol* 1917; **2**:155-164. PubMed

49. Pribram E. A Contribution to the classification of microorganisms. *J Bacteriol* 1929; **18**:361-394. PubMed

50. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**:25-29. PubMed [http://dx.doi.org/10.1038/75556](http://dx.doi.org/10.1038/75556)

51. Luo HY, Wang YR, Miao LH, Yang PL, Shi PJ, Fang CX, Yao B, Fan YL. *Nesterenkonia alba* sp. nov., an alkaliphilic actinobacterium isolated from the black liquor treatment system of a cotton pulp mill. *Int J Syst Evol Microbiol* 2009; **59**:863-868. PubMed [http://dx.doi.org/10.1099/ijs.0.003376-0](http://dx.doi.org/10.1099/ijs.0.003376-0)

52. Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006; **33**:152-155.

53. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, Raoult D. Ongoing revolu-
54. Prodigal. http://prodigal.sourceforge.net

55. Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res 2012; 40:D48-D53. PubMed http://dx.doi.org/10.1093/nar/gkr1202

56. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997; 25:955-964. PubMed http://dx.doi.org/10.1093/nar/25.5.0955

57. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007; 35:3100-3108. PubMed http://dx.doi.org/10.1093/nar/gkm160

58. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004; 340:783-795. PubMed http://dx.doi.org/10.1016/j.jmb.2004.05.028

59. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001; 305:567-580. PubMed http://dx.doi.org/10.1006/jmbi.2000.4315

60. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. Artemis: sequence visualization and annotation. Bioinformatics 2000; 16:944-945. PubMed http://dx.doi.org/10.1093/bioinformatics/16.10.944

61. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNA Plotter: circular and linear interactive genome visualization. Bioinformatics 2009; 25:119-120. PubMed http://dx.doi.org/10.1093/bioinformatics/btn578

62. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004; 14:1394-1403. PubMed http://dx.doi.org/10.1101/gr.2289704

63. 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. PubMed http://dx.doi.org/10.1186/1471-2105-12-124