Whole-Genome Sequences of a Cluster of 14 Unidentified Related Veillonella sp. Strains from Human Clinical Samples and Type Strains of 3 Veillonella Validated Species

Fabien Aujoulat, Michel Popoff, Laure Diancourt, Alexis Criscuolo, Hélène Jean-Pierre, Estelle Jumas-Bilak, Hélène Marchandin

To cite this version:

HAL Id: pasteur-02448615
https://hal-pasteur.archives-ouvertes.fr/pasteur-02448615
Submitted on 22 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Whole-Genome Sequences of a Cluster of 14 Unidentified Related Veillonella sp. Strains from Human Clinical Samples and Type Strains of 3 Veillonella Validated Species

Fabien Aujoulat,a Michel R. Popoff,b Laure Diancourt,b Alexis Criscuolo,c Hélène Jean-Pierre,d Estelle Jumas-Bilak,e Hélène Marchandinf

aHydroSciences Montpellier, CNRS, IRD, Université de Montpellier, Montpellier, France
bCentre National de Référence des bactéries anaérobies et du botulisme, Institut Pasteur, Paris, France
cInstitut Pasteur–Bioinformatics and Biotistics Hub–C3BL, USR 3756 IP CNR, Paris, France
dCHU Montpellier, Laboratoire de Bactériologie, Montpellier, France
eHydroSciences Montpellier, CNRS, IRD, Université de Montpellier, CHU Montpellier, Laboratoire d’Ecologie Microbienne Hospitalière, Montpellier, France
fHydroSciences Montpellier, CNRS, IRD, Université de Montpellier, CHU Nîmes, Nîmes, France

ABSTRACT We report 17 draft genomes for 14 unidentified Veillonella sp. strains closely related in 16S rRNA gene-based phylogeny and type strains of 3 Veillonella species with the aims of deciphering relationships between related species, evaluating the accuracy of current thresholds for species delineation, and robustly describing new species in the genus.

Veillonella spp. are anaerobic Gram-negative cocci and important representatives of the microbiota of humans and animals. Currently, 14 species in the genus Veillonella are validly described. In the absence of discriminative phenotypic characteristics, their identification requires molecular-based methods. However, all species are not discriminated by 16S rRNA gene (rrs) analysis because several pairs of species are closely related (>99% of identical rrs nucleotides), such as V. denticariosi and V. rodentium, V. ratti and V. criceti, V. ratti and V. seminalis, and V. dispar and V. parvula (1–4). In addition, intrachromosomal heterogeneity between rrs copies (up to 1.43%) and intraspecific rrs variability that may surpass interspecific variability have been demonstrated in this genus (3, 5). This impairs the 16S rRNA gene-based identification of closely related species and suggests that applying the proposed revised threshold for a new species description, which is less than 98.7% of 16S rRNA gene identity, may not be adapted to the genus Veillonella (6). Therefore, molecular-based identification methods based on housekeeping genes such as dnaK, rpoB, and gltA were successively developed and used for the description of novel species in addition to or without associated DNA-DNA hybridization (DDH), which is the reference method for novel species description (2, 4).

We present draft genome sequences of 14 human clinical isolates and type strains of 3 Veillonella species that were not available at the time of our study. Isolates were recovered from human clinical samples from 14 patients attending the University Hospital in Montpellier, France. The study was approved by the institutional review board of the Nîmes University Hospital under the approval number 19.01.07. All strains were cultured on Columbia sheep blood agar (bioMérieux) at 37°C in an anaerobic jar with the AnaeroGen system (Oxoid Unipath) for 2 to 5 days (1, 4).

DNA was extracted with the MasterPure DNA purification kit (Epicentre Biotechnologies). Libraries were constructed with the Nextera XT DNA library preparation kit (Illumina, San Diego, CA) and sequenced on a NextSeq 500 instrument with a 2 × 150-bp paired-end protocol (on average, 1,372,528 read pairs per sample and


<table>
<thead>
<tr>
<th>Species</th>
<th>Strain</th>
<th>Accession no.</th>
<th>BioSample no.</th>
<th>No. of paired-end reads</th>
<th>Coverage (×)</th>
<th>Genome size (bp)</th>
<th>No. of contigs</th>
<th>N50 (bp)</th>
<th>G+C content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>V. ratti</em></td>
<td>ATCC 17746&lt;sup&gt;2&lt;/sup&gt;</td>
<td>RQUX0000000000</td>
<td>SAMN10465589</td>
<td>1,163,748</td>
<td>156</td>
<td>2,236,286</td>
<td>49</td>
<td>95,000</td>
<td>42.8</td>
</tr>
<tr>
<td><em>V. caviae</em></td>
<td>DSM 20738&lt;sup&gt;1&lt;/sup&gt;</td>
<td>RQUY0000000000</td>
<td>SAMN10465590</td>
<td>1,391,318</td>
<td>211</td>
<td>1,963,997</td>
<td>57</td>
<td>83,095</td>
<td>38.4</td>
</tr>
<tr>
<td><em>V. seminalis</em></td>
<td>ADV 4313.2&lt;sup&gt;1,2&lt;/sup&gt;</td>
<td>RQUZ0000000000</td>
<td>SAMN10465591</td>
<td>1,399,442</td>
<td>184</td>
<td>2,252,690</td>
<td>27</td>
<td>152,260</td>
<td>42.2</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>CHU110 = 14/09/02-B-6149</td>
<td>RQUA0000000000</td>
<td>SAMN10465592</td>
<td>973,834</td>
<td>155</td>
<td>1,881,921</td>
<td>43</td>
<td>107,364</td>
<td>39.2</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>Va140 = CHU594 = 03/11/05-B-4237</td>
<td>RQVB0000000000</td>
<td>SAMN10465593</td>
<td>1,206,819</td>
<td>199</td>
<td>1,822,704</td>
<td>23</td>
<td>147,662</td>
<td>39.3</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>Va138 = CHU732 = 06/09/06-B-3252</td>
<td>RQVC0000000000</td>
<td>SAMN10465594</td>
<td>740,997</td>
<td>112</td>
<td>1,972,307</td>
<td>42</td>
<td>104,904</td>
<td>39.5</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>Va143 = CHU740 = 09/10/06-B-1508</td>
<td>RQVD0000000000</td>
<td>SAMN10465595</td>
<td>1,518,558</td>
<td>240</td>
<td>1,893,839</td>
<td>35</td>
<td>130,014</td>
<td>39.3</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>3891 = 12/03/14-B-3204</td>
<td>RQVE0000000000</td>
<td>SAMN10465596</td>
<td>1,295,195</td>
<td>205</td>
<td>1,890,765</td>
<td>38</td>
<td>161,698</td>
<td>39.4</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>3913 = 09/03/14-B-7013</td>
<td>RQVF0000000000</td>
<td>SAMN10465597</td>
<td>698,249</td>
<td>118</td>
<td>1,770,519</td>
<td>15</td>
<td>394,883</td>
<td>39.4</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>3627 = CHU4040 = 99141203627</td>
<td>RQVG0000000000</td>
<td>SAMN10465598</td>
<td>1,308,145</td>
<td>226</td>
<td>1,738,782</td>
<td>34</td>
<td>132,446</td>
<td>39.3</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>CHU3735 = CNR 79/14 = 14/11/14-B-2378</td>
<td>RQVH0000000000</td>
<td>SAMN10465599</td>
<td>1,123,247</td>
<td>190</td>
<td>1,776,932</td>
<td>51</td>
<td>93,431</td>
<td>39.4</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>3960 = CHU4076 = 99141423960</td>
<td>RQVI0000000000</td>
<td>SAMN10465600</td>
<td>746,487</td>
<td>120</td>
<td>1,860,088</td>
<td>69</td>
<td>60,297</td>
<td>39.7</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>3310 = CHU5981 = 99162773310</td>
<td>RQVJ0000000000</td>
<td>SAMN10465601</td>
<td>1,640,876</td>
<td>258</td>
<td>1,870,378</td>
<td>16</td>
<td>311,068</td>
<td>39.6</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>Va137 = 19/07/06-3404</td>
<td>RQVK0000000000</td>
<td>SAMN10465602</td>
<td>1,966,481</td>
<td>310</td>
<td>1,902,403</td>
<td>36</td>
<td>128,606</td>
<td>39.6</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>Va139 = 10/01/06-3360</td>
<td>RQVL0000000000</td>
<td>SAMN10465603</td>
<td>2,226,458</td>
<td>303</td>
<td>2,172,276</td>
<td>76</td>
<td>67,981</td>
<td>39.5</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>Va140 = 20/06/06-2457</td>
<td>RQVM0000000000</td>
<td>SAMN10465604</td>
<td>2,034,756</td>
<td>333</td>
<td>1,806,031</td>
<td>12</td>
<td>229,161</td>
<td>39.2</td>
</tr>
<tr>
<td><em>V. sp.</em></td>
<td>Va141 = 08/06/05-3381</td>
<td>RQVN0000000000</td>
<td>SAMN10465605</td>
<td>1,898,363</td>
<td>281</td>
<td>1,969,382</td>
<td>66</td>
<td>75,004</td>
<td>39.1</td>
</tr>
</tbody>
</table>

<sup>a</sup>BioProject number, PRJNA506647.
All whole-genome sequences (WGSs) are presented in Table 1. Bacteria of trimmed Veillonella of relevant gene content, particularly virulence and antibiotic resistance genes, in besides taxonomic purposes, WGS will also allow the characterization of clinically proposing minimal standards for description of new species in the genus Veillonella. Besides taxonomic purposes, WGS will also allow the characterization of clinically relevant gene content, particularly virulence and antibiotic resistance genes, in Veillonella spp. considered opportunistic human pathogens.

Data availability. This whole-genome shotgun project was deposited at DDBJ/ENA/GenBank under the accession numbers RQUX00000000 to RQVN00000000 as listed in Table 1 (BioProject number PRJNA506647; BioSample numbers SAMN10465589 to SAMN10465605). The versions described in this paper are versions RQUX01000000 to RQVN01000000. The Sequence Read Archive accession numbers are SRX5189907 to SRX5189923.

REFERENCES