



**HAL**  
open science

## **New hybrid sequencing data of *Vreelandella sulfidaeris* DSM 15722T and *Vreelandella aquamarina* DSM 15723 provide circular and complete genomes**

Gwendoline Marquet, Margareth Renault, Françoise Irlinger, Vincent Hervé

### ► **To cite this version:**

Gwendoline Marquet, Margareth Renault, Françoise Irlinger, Vincent Hervé. New hybrid sequencing data of *Vreelandella sulfidaeris* DSM 15722T and *Vreelandella aquamarina* DSM 15723 provide circular and complete genomes. *Data in Brief*, 2026, 66, pp.112707. <10.1016/j.dib.2026.112707>. <hal-05567643>

**HAL Id: hal-05567643**

**<https://hal.inrae.fr/hal-05567643v1>**

Submitted on 26 Mar 2026

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 CC BY 4.0 - Attribution - International License



## Data Article

# New hybrid sequencing data of *Vreelandella sulfidaeris* DSM 15722<sup>T</sup> and *Vreelandella aquamarina* DSM 15723 provide circular and complete genomes



Gwendoline Marquet, Margareth Renault, Françoise Irlinger, Vincent Hervé\*

Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, 91120 Palaiseau, France

## ARTICLE INFO

**Article history:**

Received 20 January 2026

Revised 12 March 2026

Accepted 16 March 2026

Available online 19 March 2026

Dataset link: [Genome of \*Vreelandella sulfidaeris\* DSM 15722 \(Original data\)](#)

Dataset link: [Genome of \*Vreelandella aquamarina\* DSM 15723 \(Original data\)](#)

**Keywords:**

*Halomonadaceae*

*Halomonas*

Halophilic bacteria

Resequencing

PacBio

Illumina

Genomics

## ABSTRACT

*Vreelandella* is a genus of halophilic and Gram-negative bacteria belonging to the family *Halomonadaceae* (*Pseudomonadota*). Various genomes of this genus are currently available but their quality is uneven. Hybrid sequencing is now a common practice for generating complete and high-quality circular genomes. Combining long-read (PacBio) and short-read (Illumina) sequencing data, we report here the circular and complete genomes of the type strain *Vreelandella sulfidaeris* DSM 15722<sup>T</sup> (4.5 Mbp with 53.7% G + C content) and of the strain *Vreelandella aquamarina* DSM 15723 (3.6 Mbp with 56.8% G + C content). The resequencing of these genomes provides an updated and high-quality dataset that can be used to improve our understanding of the metabolism of the *Vreelandella* genus. These genomes could be easily reused to address broader ecological and evolutionary questions related to this genus.

© 2026 The Author(s). Published by Elsevier Inc.  
This is an open access article under the CC BY license  
(<http://creativecommons.org/licenses/by/4.0/>)

\* Corresponding author.

E-mail address: [vincent.herve@inrae.fr](mailto:vincent.herve@inrae.fr) (V. Hervé).

## Specifications Table

Subject	Biology
Specific subject area	Genomics, bacteriology
Type of data	Table
Data collection	Genomic data, Raw, Processed Genomic DNA extraction: NucleoMag HMW DNA kit (Macherey-Nagel) Genome sequencing platforms: Illumina Nova Seq X PE150 platform and PacBio Revo platform
Data source location	The bacterial strains were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). The data were collected and stored at the SayFood Unit, Laboratory of INRAE, Palaiseau, France.
Data accessibility	-Repository name: NCBI ( <a href="https://www.ncbi.nlm.nih.gov/">https://www.ncbi.nlm.nih.gov/</a> ) Data identification number: Bioprojects PRJNA1377305 and PRJNA1381297 Direct URL to data: DSM 15722 <sup>T</sup> : <a href="https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1377305">https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1377305</a> <a href="https://www.ncbi.nlm.nih.gov/biosample/SAMN53768409">https://www.ncbi.nlm.nih.gov/biosample/SAMN53768409</a> <a href="https://www.ncbi.nlm.nih.gov/nucleotide/JBSV5F000000000.1">https://www.ncbi.nlm.nih.gov/nucleotide/JBSV5F000000000.1</a> DSM 15723: <a href="https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1381297">https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1381297</a> <a href="https://www.ncbi.nlm.nih.gov/biosample/SAMN54104122">https://www.ncbi.nlm.nih.gov/biosample/SAMN54104122</a> <a href="https://www.ncbi.nlm.nih.gov/nucleotide/JBSXYF000000000.1">https://www.ncbi.nlm.nih.gov/nucleotide/JBSXYF000000000.1</a> Instructions for accessing these data: The complete genome sequences of <i>Vreelandella sulfidaeris</i> DSM 15722 <sup>T</sup> and <i>Vreelandella aquamarina</i> DSM 15723 are available in the National Center for Biotechnology Information (NCBI) database. -Repository name: Recherche Data Gouv ( <a href="https://recherche.data.gouv.fr/en">https://recherche.data.gouv.fr/en</a> ) Direct URL to data: <a href="https://entrepot.recherche.data.gouv.fr/dataverse/vreelandella_genome_resequencing">https://entrepot.recherche.data.gouv.fr/dataverse/vreelandella_genome_resequencing</a>  Instructions for accessing these data: All the annotations described in this paper are freely available from this repository.
Related research article	None

## 1. Value of the Data

- This study presents a complete and circular genome of *Vreelandella aquamarina* DSM 15723 and *Vreelandella sulfidaeris* DSM 15722<sup>T</sup>.
- This dataset provides higher-quality sequencing data compared to the currently available genomes of these strains.
- This dataset is a valuable resource for comparative genomics.
- This dataset may facilitate research into the metabolism, ecology and evolution of the *Vreelandella* genus.

## 2. Background

*Vreelandella* is a genus of Gram-negative bacteria belonging to the halophilic family *Halomonadaceae*. The overall taxonomy of this large and diverse family has been recently revised based on genomic data [1]. Indeed, bacterial genomes are now routinely used in taxonomy and systematics to infer phylogenetic trees [2] and for comparative genomics [3]. Therefore, high-quality and complete genomes are needed for accurate and relevant genomic analyses.

At the time of writing, 41 species of *Vreelandella* have been validly published under the International Code of Nomenclature of Prokaryotes (ICNP). Currently many *Vreelandella* genomes are available but their quality is sometimes uneven [1]. Among them, the draft genomes of *Vreelandella sulfidaeris* DSM 15722<sup>T</sup> (type strain) and *Vreelandella aquamarina* (synonym *Halomonas axialensis*) DSM 15723, two strains isolated from deep-sea hydrothermal-vent environments [4],

**Table 1**

Genome assembly and annotation characteristics of the two *Vreelandella* strains, obtained via ONT and hybrid sequencing methods.

Sequencing technology	<i>Vreelandella sulfidaeris</i> DSM 15722 <sup>T</sup>		<i>Vreelandella aquamarina</i> DSM 15723	
	ONT	PacBio/Illumina	ONT	PacBio/Illumina
Genome size (bp)	4,480,770	4,504,718	3,619,799	3,629,439
Coverage (x)	98	72	100	70
Nb of contigs	2	2	1	2
GC %	53.7	53.7	56.8	56.8
CDS	7,620	4,212	5,603	3,365
Coding density (%)	82	89	85.8	90.2
Pseudogenes	119	17	1,019	8
Hypothetical genes	1,804	334	1,735	154
rRNA	18	18	18	18
tRNA	62	62	62	62
tmRNA	1	1	1	1
ncRNA	5	6	11	12
CheckM2 completeness (%)	82.8	100	84.76	99.99
CheckM2 contamination (%)	6.56	1.08	4.99	0.42
GenBank genome assembly	GCA_007182875.1	GCA_054072535.1	GCA_007163885.2	GCA_054110395.1

have been published [5,6]. Here we report the circular and complete genomes of *Vreelandella sulfidaeris* DSM 15722<sup>T</sup> (type strain) and *Vreelandella aquamarina* DSM 15723. The resequencing of these genomes provides an updated dataset that can be used to better understand the metabolism, ecology and evolution of the *Vreelandella* genus.

### 3. Data Description

We present the complete and circular genomes of *Vreelandella aquamarina* DSM 15723 and *Vreelandella sulfidaeris* DSM 15722<sup>T</sup>, obtained through hybrid sequencing that combines Illumina NovaSeq X PE150 and PacBio Revio data. We compared our data with previously published genomes generated by Oxford Nanopore Technology (ONT) (Table 1).

Concerning the strain *V. sulfidaeris* DSM 15722<sup>T</sup>, the two different sequencing resulted in two contigs (1 chromosome and 1 plasmid), with a similar total genome size. However, for the strain *V. aquamarina* DSM 15723, while one contig was detected using Nanopore technology, the analysis of hybrid sequencing revealed the presence of two contigs (1 chromosome and 1 plasmid).

Regarding the coding density, we found higher values than for ONT sequenced genomes. Conversely, the number of pseudogenes (17 vs 119 for DSM 15722<sup>T</sup> and 8 vs 1,019 for DSM 15723), hypothetical genes (1,804 vs 334 for DSM 15722<sup>T</sup> and 1,735 vs 154 for DSM 15723) and coding DNA sequence (CDS) (7,620 vs 4,212 for DSM 15722<sup>T</sup> and 5,603 vs 3,365 for DSM 15723) are lower in our current dataset. Overall, our dataset offers genomes of higher quality with significantly higher completeness and lower contamination as determined by CheckM2 criteria (Table 1).

We also examined the predicted secondary metabolism using antiSMASH and found more regions encoding biosynthetic gene clusters in our new dataset. For *V. sulfidaeris* DSM 15722<sup>T</sup>, 9 secondary metabolite regions were detected in the ONT sequenced genome. In our newly sequenced genomes, we also detected these nine regions, as well as an extra region related to homoserine lactone production. Similarly, for the strain DSM 15723, five secondary metabolite regions were detected in the ONT sequenced genome. In our complete and circular genome, these five regions were also detected, but we also found two more secondary metabolite regions: one related to the synthesis of a type I polyketide synthase and another related to a cluster of ribosomally synthesized and posttranslationally modified peptide (RiPPs) recognition element (RRE).

## 4. Experimental Design, Materials and Methods

### 4.1. Bacterial culture

The bacterial strains *Vreelandella sulfidaeris* DSM 15722<sup>T</sup> (= Esulfide1 = ATCC BAA-803 = CECT 5817 = NZp) and *Vreelandella aquamarina* DSM 15723 (= 1723a = Althf1 = ATCC BAA-802 = CECT 5812 = KCTC 22,194) were grown in Marine broth without agitation for 3 days at 28°C. The cell pellets were collected by centrifugation of 1 mL of culture at 5000 x g for 10 min at 4°C.

### 4.2. DNA extraction and library preparation

We extracted total DNA using the NucleoMag HMW DNA kit (Macherey-Nagel, Hoerd, France) following the manufacturer's protocol with some modifications: we added lysozyme (25 µL, 100 mg/mL) to the bacterial pellet in HM1 buffer, and incubated samples for 1 h at 37°C before the addition of proteinase K (included in the kit). After RNase treatment, samples were mixed with beads. After several steps of washing, the beads were air-dried for 15 min at room temperature. We eluted DNA in 100 µL of 5 mM Tris/HCl, pH 8.5 buffer. DNA quantity and quality were assessed using the Nanodrop One (ThermoFisher Scientific, Waltham, USA) and the Agilent 4150 TapeStation with the Genomic DNA ScreenTape System (Agilent, Santa Clara, USA). Library preparation and sequencing were performed by BMKGENE (Münster, Germany) using two platforms: Illumina NovaSeq X PE150 and PacBio Revio.

### 4.3. Genome assembly and annotation

Previously published genomes of *V. sulfidaeris* DSM 15722<sup>T</sup> [5] and *V. aquamarina* (previously known as *V. axialensis*) DSM 15723 [6] sequenced using Oxford Nanopore Technology (ONT) were downloaded from NCBI. To compare these two genomes with our newly sequenced ones, we analysed them with the same software with default parameters.

Our genome assembly was performed by Hifiasm v0.12 [7]. Circlator v1.5.5 was applied to circularize the assembled sequences and resolve high-quality ends [8]. The sequence was further corrected using Illumina reads with Pilon v1.22 [9] in order to generate a highly accurate assembly for downstream analysis. Genome annotation was done with Bakta version 1.11.0 [10]. Genome completeness and contamination were assessed by CheckM2 version 1.1.0 [11]. Analysis of secondary metabolism was performed using antiSMASH version 8.0.4 with default parameters [12].

## Limitations

None.

## Ethics Statement

The authors have read and followed the ethical requirements for publication in *Data in Brief*. The current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

## Data Availability

[Genome of \*Vreelandella sulfidaeris\* DSM 15722 \(Original data\)](#) ((NCBI),)

[Genome of \*Vreelandella aquamarina\* DSM 15723 \(Original data\)](#) ((NCBI).)

## CRedit Author Statement

**Gwendoline Marquet:** Methodology, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization; **Margareth Renault:** Investigation, Resources, Writing – review & editing; **Françoise Irlinger:** Resources, Writing – review & editing; **Vincent Hervé:** Conceptualization, Validation, Data curation, Writing – original draft, Project administration.

## Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We are grateful to the INRAE MIGALE bioinformatics facility (MIGALE, INRAE, 2020. Migale bioinformatics Facility, doi: [10.15454/1.5572390655343293E12](https://doi.org/10.15454/1.5572390655343293E12)) for providing help, computing and storage resources.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- [1] R.R. de la Haba, D.R. Arahala, C. Sánchez-Porro, M. Chuvochina, S. Wittouck, P. Hugenoltz, A. Ventosa, A long-awaited taxogenomic investigation of the family *Halomonadaceae*, *Front. Microbiol.* 14 (2023) 1293707, doi:[10.3389/fmicb.2023.1293707](https://doi.org/10.3389/fmicb.2023.1293707).
- [2] D.H. Parks, M. Chuvochina, D.W. Waite, C. Rinke, A. Skarshewski, P.-A. Chaumeil, P. Hugenoltz, A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life, *Nat. Biotechnol.* 36 (2018) 996–1004, doi:[10.1038/nbt.4229](https://doi.org/10.1038/nbt.4229).
- [3] H. Tettelin, D. Riley, C. Cattuto, D. Medini, Comparative genomics: the bacterial pan-genome, *Curr. Opin. Microbiol.* 11 (2008) 472–477, doi:[10.1016/j.mib.2008.09.006](https://doi.org/10.1016/j.mib.2008.09.006).
- [4] J.Z. Kaye, M.C. Márquez, A. Ventosa, J.A. Baross, *Halomonas neptunia* sp. nov., *Halomonas sulfidaeris* sp. nov., *Halomonas axialensis* sp. nov. and *Halomonas hydrothermalis* sp. nov.: halophilic bacteria isolated from deep-sea hydrothermal-vent environments, *Int. J. Syst. Evol. Microbiol.* 54 (2004) 499–511, doi:[10.1099/ijs.0.02799-0](https://doi.org/10.1099/ijs.0.02799-0).
- [5] M. Saito, A. Nishigata, J. Galipon, K. Arakawa, Complete genome sequence of *Halomonas sulfidaeris* strain Esulfide1 isolated from a metal sulfide rock at a depth of 2200 meters, obtained using Nanopore sequencing, *Microbiol. Resour. Announc.* 8 (2019) e00327–19, doi:[10.1128/mra.00327-19](https://doi.org/10.1128/mra.00327-19).
- [6] M. Tsurumaki, S. Deno, J. Galipon, K. Arakawa, Complete genome sequence of halophilic deep-sea bacterium *Halomonas axialensis* strain Althf1, *Microbiol. Resour. Announc.* 8 (2019) e00839–19, doi:[10.1128/mra.00839-19](https://doi.org/10.1128/mra.00839-19).
- [7] H. Cheng, G.T. Concepcion, X. Feng, H. Zhang, H. Li, Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm, *Nat. Methods* 18 (2021) 170–175, doi:[10.1038/s41592-020-01056-5](https://doi.org/10.1038/s41592-020-01056-5).
- [8] M. Hunt, N.D. Silva, T.D. Otto, J. Parkhill, J.A. Keane, S.R. Harris, Circlator: automated circularization of genome assemblies using long sequencing reads, *Genome Biol.* 16 (2015) 294, doi:[10.1186/s13059-015-0849-0](https://doi.org/10.1186/s13059-015-0849-0).
- [9] B.J. Walker, T. Abeel, T. Shea, M. Priest, A. Bouuelliel, S. Sakthikumar, C.A. Cuomo, Q. Zeng, J. Wortman, S.K. Young, A.M. Earl, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement, *PLoS One* 9 (2014) e112963, doi:[10.1371/journal.pone.0112963](https://doi.org/10.1371/journal.pone.0112963).
- [10] S. Beyvers, L. Jelonek, A. Goesmann, O. Schwengers, Bakta Web – rapid and standardized genome annotation on scalable infrastructures, *Nucleic Acids Res.* 53 (2025) W51–W56, doi:[10.1093/nar/gkaf335](https://doi.org/10.1093/nar/gkaf335).
- [11] A. Chklovski, D.H. Parks, B.J. Woodcroft, G.W. Tyson, CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning, *Nat. Methods* 20 (2023) 1203–1212, doi:[10.1038/s41592-023-01940-w](https://doi.org/10.1038/s41592-023-01940-w).
- [12] K. Blin, S. Shaw, L. Vader, J. Szenei, Z.L. Reitz, H.E. Augustijn, J.D.D. Cediell-Becerra, V. de Crécy-Lagard, R.A. Koetsier, S.E. Williams, P. Cruz-Morales, S. Wongwas, A.E. Segurado Luchsinger, F. Biermann, A. Korenskaia, M.M. Zdouc, D. Meijer, B.R. Terlouw, J.J.J. van der Hoof, N. Ziemert, E.J.N. Helfrich, J. Masschelein, C. Corre, M.G. Chevette, G.P. van Wezel, M.H. Medema, T. Weber, antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation, *Nucleic Acids Res.* 53 (2025) W32–W38, doi:[10.1093/nar/gkaf334](https://doi.org/10.1093/nar/gkaf334).