Whole genome-based taxonomy of Shewanella and Parashewanella

Alberto J. Martín-Rodríguez, Jan P. Meier-Kolthoff

Angaben zur Veröffentlichung / Publication details:

Martín-Rodríguez, Alberto J., and Jan P. Meier-Kolthoff. 2022. "Whole genome-based taxonomy of Shewanella and Parashewanella." *International Journal of Systematic and Evolutionary Microbiology* 72 (7): 005438. https://doi.org/10.1099/ijsem.0.005438.

INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

TAXONOMIC NOTE

Martín-Rodríguez and Meier-Kolthoff, *Int. J. Syst. Evol. Microbiol.* 2022;72:005438

DOI 10.1099/ijsem.0.005438

MICROBIOLOGY SOCIETY

OPEN ACCESS

Whole genome-based taxonomy of *Shewanella* and *Parashewanella*

Alberto J. Martín-Rodríguez[1,*] and Jan P. Meier-Kolthoff[2]

Abstract

The family *Shewanellaceae* currently comprises three genera, *Shewanella*, *Parashewanella* and *Psychrobium*, the latter represented by a single species. From the second half of the 1990s, the number of novel species in the *Shewanellaceae* has steadily increased, suggesting that the true diversity of this family has only begun to emerge. In recent years, efforts to provide a genus-wide, whole genome-based taxonomy for *Shewanella* have been limited by the lack of numerous type strain genome sequences. To shed light on this question, we sequenced all *Shewanella* type strains that lacked a publicly available whole-genome sequence. Using state-of-the-art phylogenomic methods, here we provide a genus-wide taxonomy of *Shewanella* and *Parashewanella* that resulted in the identification of 48 novel species represented by 73 sequenced isolates, and we propose the correction of 43 misidentified non-type-strain isolates. Our work sets a reference for family-wide comparative genomic studies addressing genetic or ecophysiological aspects of *Shewanellaceae*, as well as subsequent species descriptions.

DATA SUMMARY

Supplementary Material can be found on figshare at the following link: https://doi.org/10.6084/m9.figshare.19633374 [1].

Draft genome sequences of the strains listed in Table S1 are deposited in GenBank under BioSample accessions SAMN24537863–SAMN24537898.

INTRODUCTION

The genus *Shewanella*, named after the Scottish fisheries microbiologist James M. Shewan [2], comprises more than 70 species of *Gammaproteobacteria* with a validly published name under the International Code of Nomenclature of Prokaryotes (ICNP), excluding synonyms, according to the List of Prokaryotic names with Standing in Nomenclature (LPSN) [3, 4]. Members of this genus inhabit a diversity of aquatic and sedimentary ecosystems worldwide. *Shewanella* species are also part of the microbiota of aquatic animals. Most species are not pathogenic to humans, although certain species, predominantly *Shewanella algae*, can occasionally cause disease [5, 6].

Shewanella was the only genus in the family *Shewanellaceae* until 2014, when the genus *Psychrobium* was proposed to accommodate the psychrophilic, low G+C content (40.5 mol%) species *Psychrobium conchae* [7], which is, to date, the only member of this genus. In 2019, the genus *Parashewanella* ('beside *Shewanella*') was proposed and accepted to accommodate the species *Parashewanella curva* sp. nov. and *Parashewanella spongiae* comb. nov. [8] (formerly *Shewanella spongiae* [9]) exhibiting, among other features, a distinct phylogeny, lower G+C content than *Shewanella* spp., and the inability to produce respiratory menaquinones. Despite its name, the genus *Alishewanella* [10], in the family Alteromonadaceae, is taxonomically distant. The name *Alishewanella* ('the other *Shewanella*') was originally given because of initial misidentification of the type species *Alishewanella fetalis* as *Shewanella putrefaciens* by the conventional biochemical methods routinely employed at the time [10].

Author affiliations: [1]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden; [2]Department of Bioinformatics and Databases, Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Brunswick, Germany.
***Correspondence:** Alberto J. Martín-Rodríguez, jonatan.martin.rodriguez@ki.se
Keywords: phylogenomics; *Shewanella*; *Parashewanella*; dDDH; whole proteome; taxonomy.
Abbreviations: CTMR, Center for Translational Microbiome Research; dDDH, digital DNA–DNA hybridization; GBDP, genome BLAST distance phylogeny; ICNP, International Code of Nomenclature of Prokaryotes; LPSN, List of Prokaryotic names with Standing in Nomenclature; MiGS, Microbial Genome Sequencing Center; MIS, misassigned; ML, maximum-likelihood; MP, maximum-parsimony; TYGS, Type (Strain) Genome Server.
Draft genome sequences of the strains listed in Table S1 are deposited in GenBank under BioSample accessions SAMN24537863–SAMN24537898.
A supplementary table is available with the online version of this article.

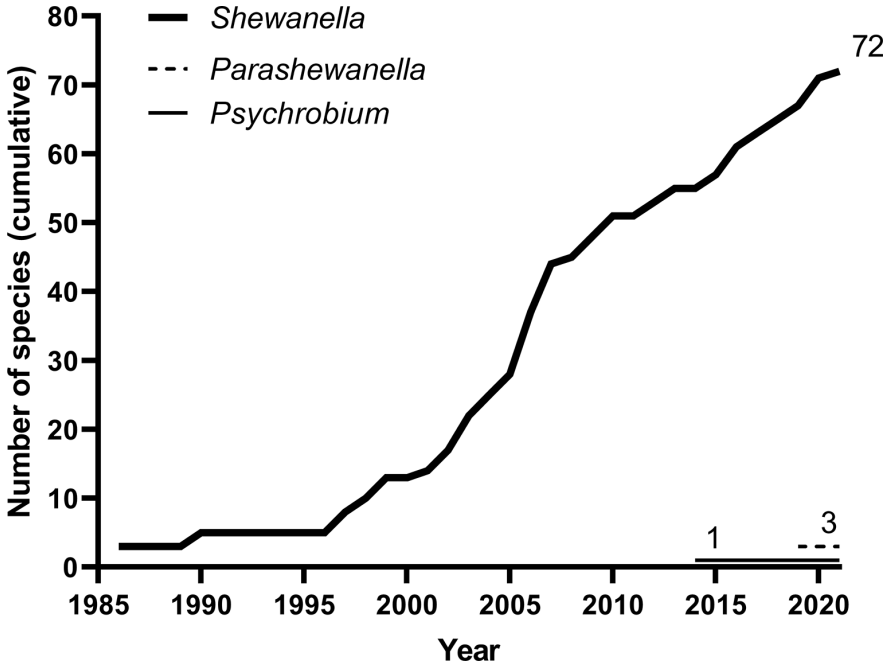

Fig. 1. Cumulative number of *Shewanella*, *Parashewanella* and *Psychrobium* species over time. Data based on LPSN as of 22 December 2021 with the amendments described in the text.

Today, next-generation sequencing technologies provide access to bacterial whole genome sequences at affordable costs and are widely used in research and diagnostics. Genomics has revolutionized bacterial taxonomy by providing reliable and reproducible means of inferring evolutionary relationships through bioinformatic analysis. Minimal standards for the use of genome data for species circumscriptions are now universally accepted [11].

In 2019, we provided a whole genome-based taxonomy for *Shewanella* [12]. In our study, *Shewanella upenei*, *Shewanella arctica* and *Shewanella pacifica* were defined as later heterotypic synonyms of *Shewanella algae*, *Shewanella frigidimarina* and *Shewanella japonica*, respectively. In addition, we showed that roughly half of the sequenced *Shewanella* isolates were in need of taxonomic revision. A major limitation of our study was the lack of numerous type strain sequences, which impaired a genus-wide taxonomic resolution. In the current study we sequenced all *Shewanella* type strains lacking a publicly available whole genome sequence and we took a bioinformatic approach to delineate the taxonomic relationships of *Shewanella* and *Parashewanella* strains based on genomic data. To this end, we used state-of-the-art methods for genome-based taxonomic classification as provided by the Type (Strain) Genome Server (TYGS), not only including pairwise digital DNA–DNA hybridization (dDDH) and 16S rRNA gene sequence-based phylogenetic analysis, but especially whole proteome sequence-based phylogenomic reconstruction [4, 13]. These analyses unveiled the existence of 48 novel species within *Shewanellaceae*: 46 novel *Shewanella* species and 2 novel *Parashewanella* species, most of which await formal description, substantially increasing the diversity of this family. This work supports the reclassification of 43 additional strains. Our study resolves the taxonomy of *Shewanella* and *Parashewanella* and sets a reference towards subsequent circumscriptions within these genera.

METHODS

DNA isolation and sequencing

DNA from the strains listed in Table S1 (available in the online version of this article) was extracted using the GenElute Bacterial Genomic DNA kit (Sigma). Sequencing was performed at the Center for Translational Microbiome Research (CTMR; Karolinska Institutet, Sweden) and the Microbial Genome Sequencing Center (MiGS; Pittsburgh, PA, USA).

Library preparation at CTMR was carried out on 50 ng of genomic DNA with the MGI FS library prep set according to the manufacturer's instructions. Library quality was evaluated with the TapeStation D1000 kit (Agilent). Libraries were quantified using the Quant-iT High Sensitivity dsDNA assay (ThermoFisher) using a Tecan Spark instrument. Circularized DNA of equimolarly pooled libraries was prepared using the MGI Easy Circularization kit (MGI Tech). DNBseq 2×100 bp paired-end sequencing was performed using a DNBSEQ G400 sequencing instrument (MGI) according to the manufacturer's instructions.

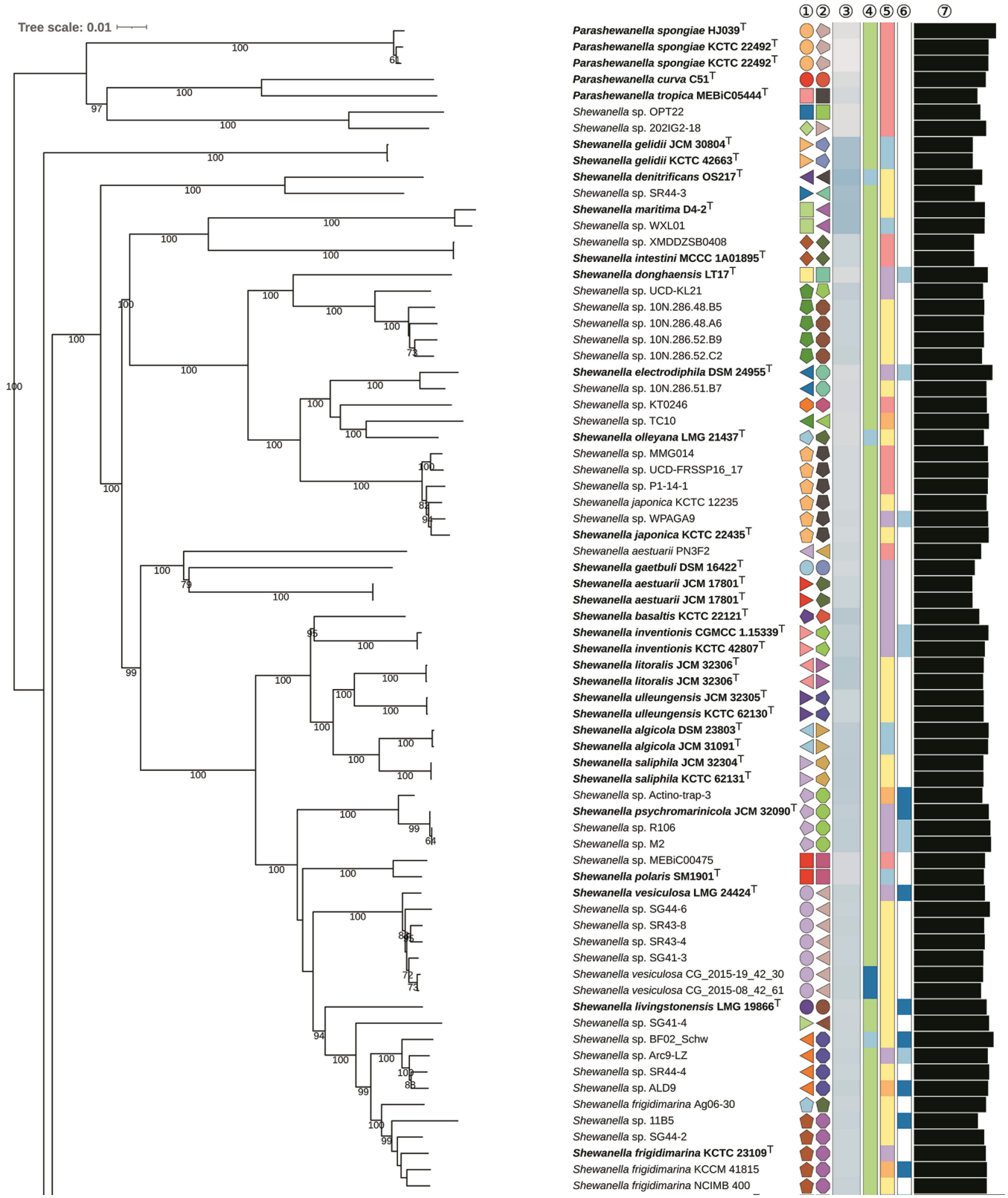

Fig. 2. Phylogenomic GBDP tree inferred with FastME 2.1.6.1 from whole proteomes (part 1). The branch lengths are scaled via GBDP distance formula d_5. Branch values are GBDP pseudo-bootstrap support values >60% from 100 replications, with an average branch support of 70.2%. The tree was midpoint-rooted. Symbols and numbers in circles are explained in the last part of this figure (Fig. 6). Non-monophyletic subspecies clusters are due to a lack of support at the subspecific level, i.e. the exact phylogenetic placement within the species clusters is not always sufficiently supported.

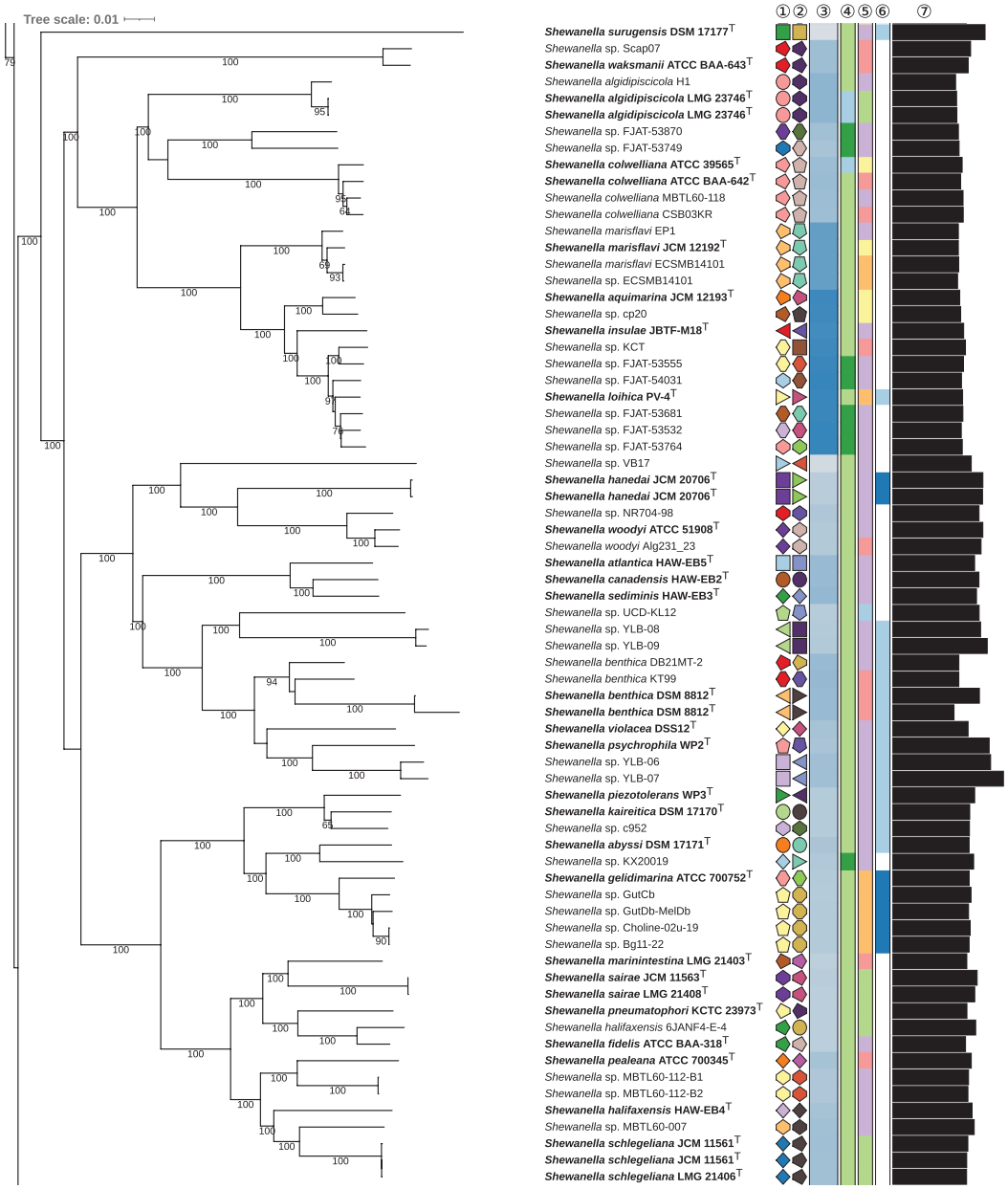

Fig. 3. Phylogenomic GBDP tree inferred with FastME 2.1.6.1 from whole proteomes (part 2). The tree parameters are listed in caption of Fig. 2.

For strains A49 and JC5 (Table S1), library preparation was performed using the TruSeq Nano DNA library preparation kit (Illumina). Libraries were sequenced on a MiSeq platform, 2×300 bp paired end reads.

Library preparation at MiGS was performed using the Illumina DNA Prep kit following the manufacturer's instructions and sequenced on an Illumina NextSeq2000 instrument (2×151 bp). Quality control and adapter trimming was performed with Illumina bcl2fastq version 2.20.0.422.

Assembly of raw sequencing reads from either platform was performed with BACTpipe version 2.6.0 (strains A49 and JC5) or 3.1 (all other strains), available at https://github.com/ctmrbio/BACTpipe.

Genome sequence dataset

Genome assemblies of *Shewanella* and *Parashewanella* strains were retrieved from NCBI GenBank on 31 August 2021. Assemblies were individually scrutinized and those excluded from RefSeq were removed.

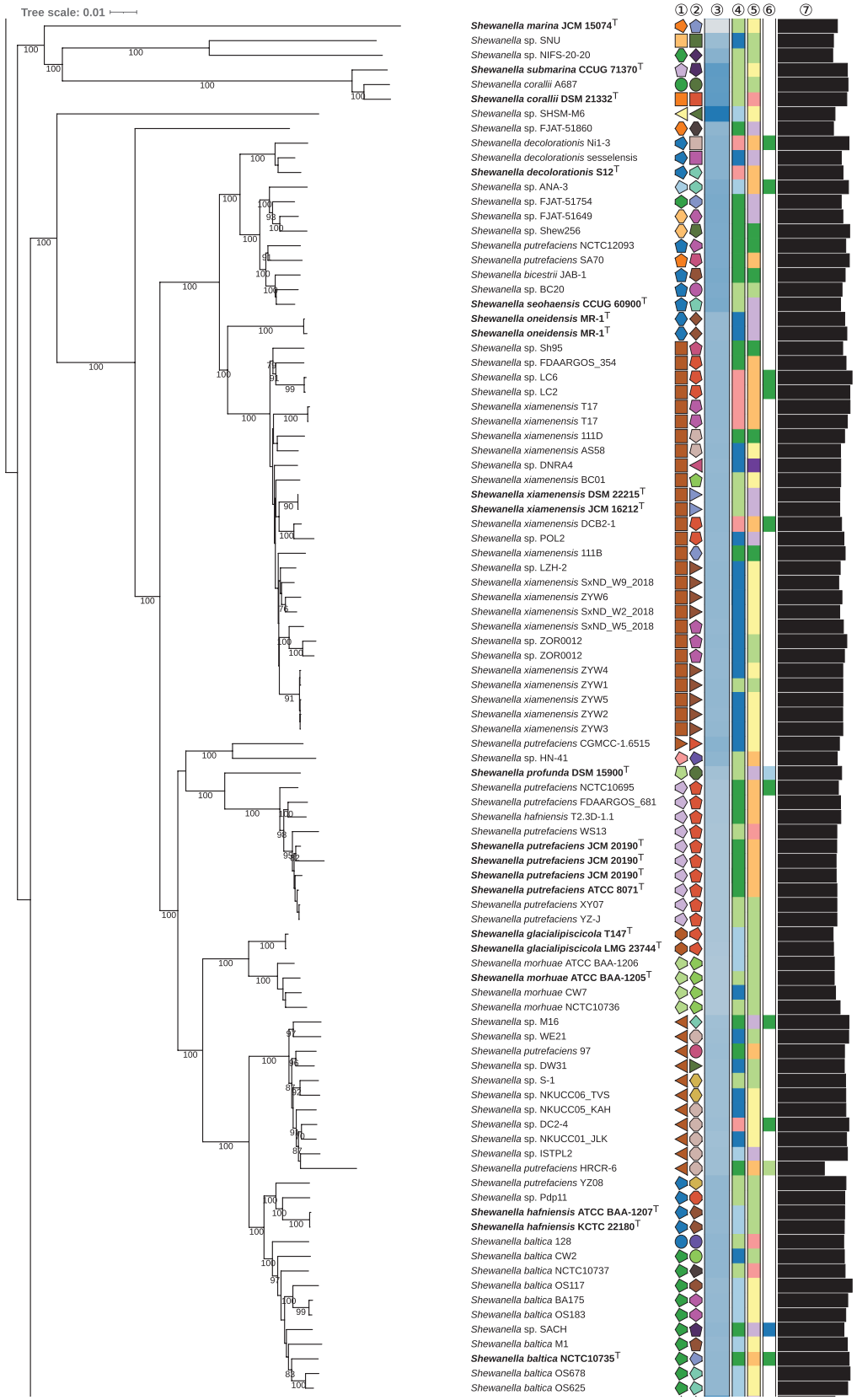

Fig. 4. Phylogenomic GBDP tree inferred with FastME 2.1.6.1 from whole proteomes (part 3). The tree parameters are listed in caption of Fig. 2.

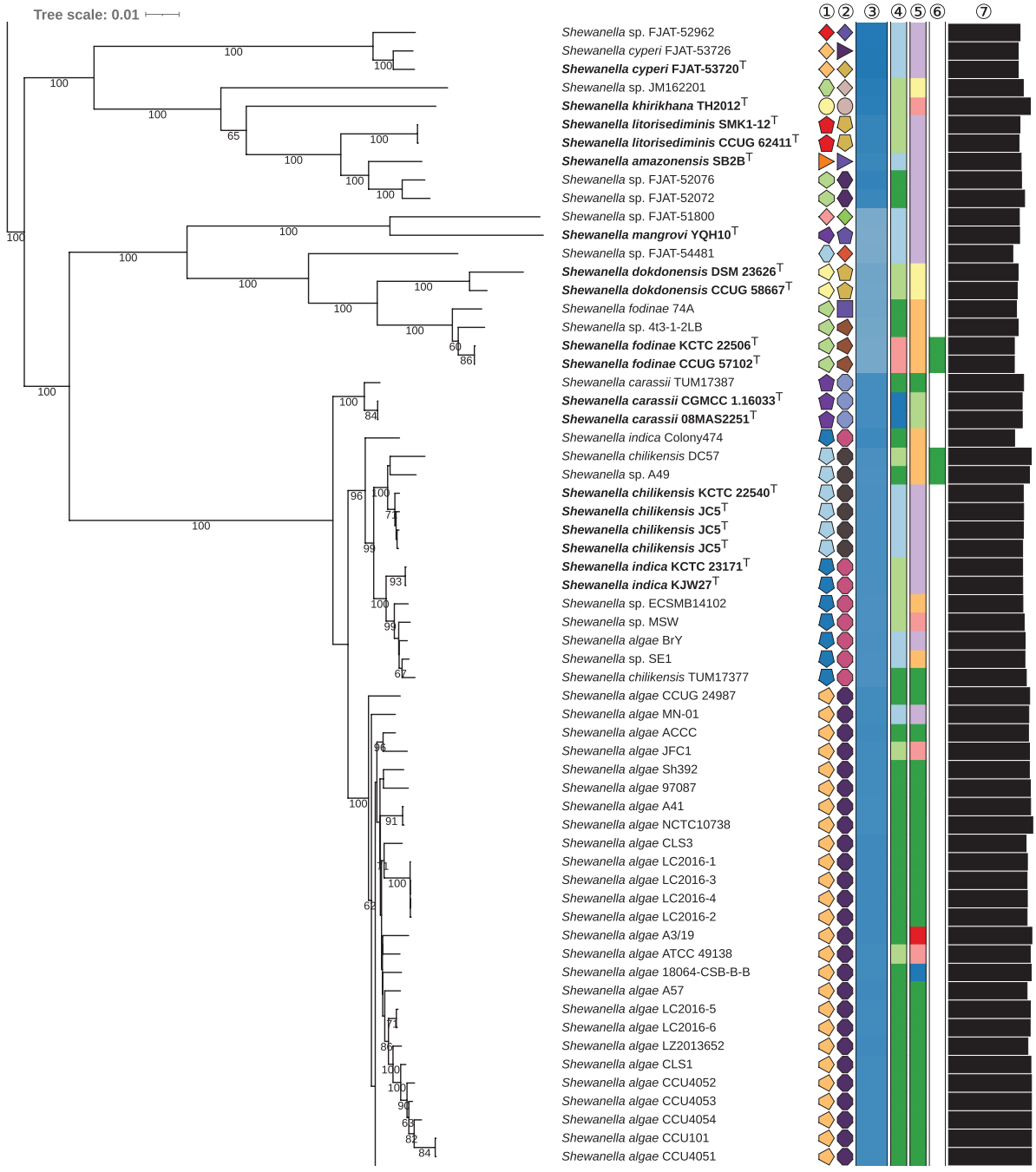

Fig. 5. Phylogenomic GBDP tree inferred with FastME 2.1.6.1 from whole proteomes (part 4). The tree parameters are listed in caption of Fig. 2.

Phylogenomic analyses

Genome sequence data were uploaded to the TYGS, available at https://tygs.dsmz.de, for a whole genome-based taxonomic analysis [13] incorporating recently introduced methodological updates and features [4]. The TYGS analysis was subdivided into the following steps:

Pairwise comparison of genome sequences

For the phylogenomic inference, all pairwise comparisons among the set of genomes were conducted using genome BLAST distance phylogeny (GBDP) and accurate intergenomic distances inferred under the algorithm 'trimming' and distance formula d_5 [14]. One hundred distance replicates were calculated each. A second GBDP phylogenomic analysis was inferred

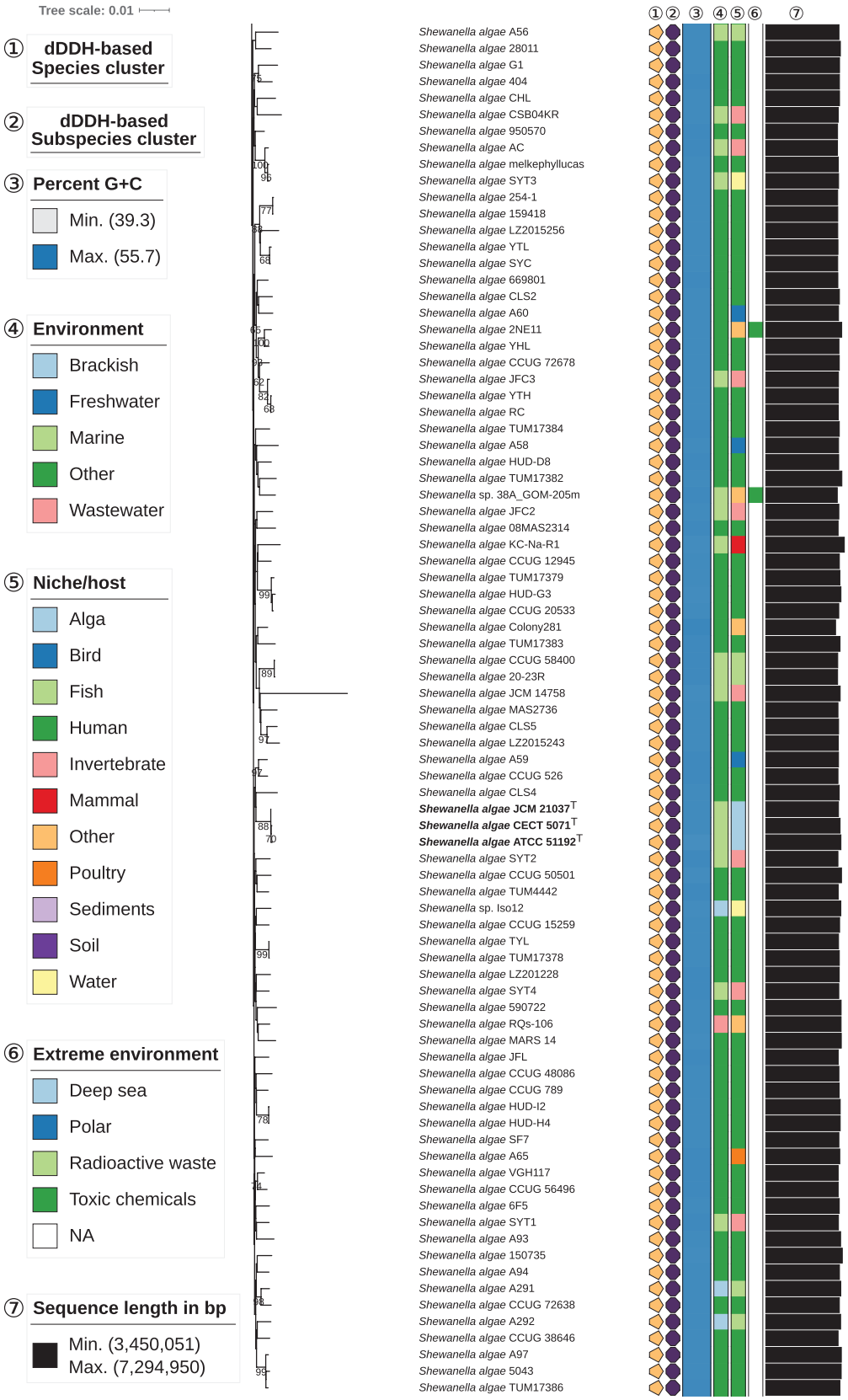

Fig. 6. Phylogenomic GBDP tree inferred with FastME 2.1.6.1 from whole proteomes (part 5). The tree parameters are listed in caption of Fig. 2.

Table 1. Strains representing novel species identified upon phylogenomic analyses

A brief description of the isolation source of each strain, inferred from the available GenBank BioSample information or published data if available, is provided. At the relevant instances, proposed names by the respective authors are indicated.

Strain	Isolation source	Proposed name
Shewanella sp. NIFS-20–20	Marine fish	–
Shewanella corallii A687	Marine fish	–
Shewanella sp. SHSM-M6	Brackish water	–
Shewanella sp. FJAT-52962	Sediments	*Shewanella sedimentimangrovi* Liu *et al.* 2021 [36]
Shewanella sp. JM162201	Seawater	–
Shewanella sp. FJAT-52076 *Shewanella* sp. FJAT-52072	Sediments Sediments	–
Shewanella sp. FJAT-51800	Sediments	*Shewanella avicenniae* Liu *et al.* 2021 [36]
Shewanella sp. cp20	Seawater	–
Shewanella sp. KCT *Shewanella* sp. FJAT-53555	Marine invertebrate Sediments	–
Shewanella sp. FJAT-54031	Sediments	–
Shewanella sp. FJAT-53764	Sediments	–
Shewanella sp. FJAT-53681	Sediments	–
Shewanella sp. FJAT-53532	Sediments	–
Shewanella sp. FJAT-54481	Sediments	*Shewanella yunxiaonensis* Liu *et al.* 2021 [36]
Shewanella sp. SNU WT4	Freshwater fish	–
Shewanella sp. VB17	Marine sediments	–
Shewanella sp. YLB-08 *Shewanella* sp. YLB-09	Deep sea sediments Deep sea sediments	*Shewanella eurypsychrophilus* Yu *et al.* 2021 [37]
Shewanella sp. YLB-06 *Shewanella* sp. YLB-07	Deep sea sediments Deep sea sediments	*Shewanella psychropiezotolerans* Yu *et al.* 2021 [37]
Shewanella benthica KT99	Deep sea invertebrates and water	–
Shewanella benthica DB21MT-2	Deep sea sediments	–
Shewanella sp. SR44-3	Seawater	–
Shewanella sp. FJAT-51860	Sediments	–
Shewanella sp. ANA-3	As-treated pier in brackish water	–
Shewanella sp. FJAT-51754	Sediments	–
Shewanella sp. FJAT-51649 *Shewanella* sp. Shew256	Sediments Human	–
Shewanella putrefaciens CGMCC-1.6515	Freshwater	–
Shewanella sp. HN-41	Rocks	–
Shewanella putrefaciens 97 *Shewanella* sp. M16 *Shewanella* sp. DW31 *Shewanella* sp. S-1 *Shewanella* sp. NKUCC06_TVS *Shewanella* sp. WE21 *Shewanella putrefaciens* HRCR-6 *Shewanella* sp. NKUCC05_KAH *Shewanella* sp. ISTPL2 *Shewanella* sp. DC2-4 *Shewanella* sp. NKUCC01_JLK	Unknown As-containing sediments Freshwater fish Marine fish Freshwater Freshwater fish Groundwater with potential radioactive waste Freshwater Lake sediments Acid mine drainage Freshwater	–
Shewanella baltica 128	Marine invertebrate	–
Shewanella sp. SG41-4	Seawater	–

Continued

Table 1. Continued

Strain	Isolation source	Proposed name
Shewanella sp. BF02_Schw	Subglacial brine	–
Shewanella sp. Arc9-LZ	Deep sea sediments	
Shewanella sp. SR44-4	Seawater	
Shewanella sp. ALD9	Sea ice floe	
Shewanella sp. FJAT-53870	Sediments	–
Shewanella sp. FJAT-53749	Sediments	–
Shewanella sp. OPT22	Marine invertebrate	- (*Parashewanella*)
Shewanella sp. 202IG2-18	Marine invertebrate	- (*Parashewanella*)
Shewanella sp. UCD-KL21	Marine sediments	–
Shewanella sp. 10 N.286.48.A6	Seawater	–
Shewanella sp. 10 N.286.48.B5	Seawater	
Shewanella sp. 10 N.286.52.B9	Seawater	
Shewanella sp. 10 N.286.52.C2	Seawater	
Shewanella sp. KT0246	Marine invertebrate	–
Shewanella sp. TC10	Marine biofilm	–
Shewanella sp. UCD-KL12	Marine alga	–
Shewanella sp. c952	Deep sea sediments	–
Shewanella sp. KX20019	Sediments	–
Shewanella sp. GutCb	Faeces, Arctic ocean	–
Shewanella sp. GutDb-MelDb	Faeces, Arctic ocean	
Shewanella sp. Choline-02u-19	Sea ice floe	
Shewanella sp. Bg11-22	Sea ice floe	
Shewanella sp. NR704-98	Marine sediments	*Shewanella nanhaiensis* Cao *et al.* 2021 [29]
Shewanella sp. MBTL60-112-B1	Marine sediments	–
Shewanella sp. MBTL60-112-B2	Marine sediments	
Shewanella sp. MBTL60-007	Marine sediments	–
Shewanella frigidimarina Ag06-30	Seawater	–
Shewanella aestuarii PN3F2	Marine invertebrate	–

using the amino acid sequences of the entire proteome as input, which was expected to provide a better resolved phylogeny in datasets of e.g. only remotely related strains. Digital DDH values and confidence intervals were calculated using the recommended settings of the GGDC 3.0 [4, 14].

Phylogenetic inference

The resulting intergenomic distances were used to infer a balanced minimum-evolution tree with branch support via FastME 2.1.6.1 including SPR postprocessing [15]. Branch support was inferred from 100 pseudo-bootstrap replicates each. The trees were rooted at the midpoint [16] and visualized with iTOL [17].

Type-based species and subspecies clustering

The type-based species clustering using a 70% dDDH radius around each of the 108 type strains (of 75 distinct species) was done as previously described [13]. Subspecies clustering was done using a 79% dDDH threshold as previously introduced [18]. All resulting clusters were annotated in the iTOL visualization.

16S rRNA gene sequence phylogeny and pairwise comparisons

The 16S (small subunit, SSU) rRNA gene sequence with the best RNAmmer score was extracted from each of the 381 assemblies via RNAmmer [19] if present in the sequence data. Phylogenies were inferred by the specialized DSMZ single-gene phylogeny pipeline [4]. That is, a multiple sequence alignment was created with MUSCLE [20]. Maximum likelihood (ML) and maximum parsimony (MP) trees were inferred from the alignment with RAxML [21] and TNT [22], respectively. For the ML tree, rapid bootstrapping in conjunction with the autoMRE bootstopping criterion [23] and subsequent search for the best tree was used; for the MP tree, 1000 bootstrapping replicates were used in conjunction with tree-bisection-and-reconnection branch swapping and

Table 2. Correction of previously misidentified strains and species assignments of isolates previously identified at the genus level

The isolation source and correct name of each strain are indicated.

Strain	Isolation source	Correct name
Shewanella sp. A49	Human	*Shewanella chilikensis* A49
Shewanella sp. SE1	Brackish water	*Shewanella indica* SE1
Shewanella algae BrY	Sediments	*Shewanella indica* BrY
Shewanella sp. ECSMB14102	Marine biofilm	*Shewanella indica* ECSMB14102
Shewanella sp. MSW	Marine invertebrate	*Shewanella indica* MSW
Shewanella sp. 38A_GOM-205m	Oil-amended biotrap in seawater	*Shewanella algae* 38A_GOM-205m
Shewanella sp. ECSMB14101	Marine biofilm	*Shewanella marisflavi* ECSMB14101
Shewanella sp. 4t3-1-2LB	Bicycle in a canal	*Shewanella fodinae* 4t3-1-2LB
Shewanella putrefaciens NCTC12093	Human	*Shewanella seohaensis* NCTC12093
Shewanella bicestrii JAB-1	Human	*Shewanella seohaensis* JAB-1
Shewanella sp. BC20	Marine fish	*Shewanella seohaensis* BC20
Shewanella putrefaciens SA70	Hospital	*Shewanella seohaensis* SA70
Shewanella sp. Sh95	Human	*Shewanella xiamenensis* Sh95
Shewanella sp. POL2	Lake sediments	*Shewanella xiamenensis* POL2
Shewanella sp. FDAARGOS_354	Not disclosed or published	*Shewanella xiamenensis* FDAARGOS_354
Shewanella sp. LC6	Industrial wastewater	*Shewanella xiamenensis* LC6
Shewanella sp. LC2	Industrial wastewater	*Shewanella xiamenensis* LC2
Shewanella sp. DNRA4	Rice field soil	*Shewanella xiamenensis* DNRA4
Shewanella sp. LZH-2	Freshwater	*Shewanella xiamenensis* LZH-2
Shewanella sp. ZOR0012	Freshwater fish	*Shewanella xiamenensis* ZOR0012
Shewanella hafniensis T2.3D-1.1	Groundwater	*Shewanella putrefaciens* T2.3D-1.1
Shewanella putrefaciens YZ08	Marine fish	*Shewanella hafniensis* YZ08
Shewanella sp. Pdp11	Marine fish	*Shewanella hafniensis* Pdp11
Shewanella sp. SACH	Antarctic soil	*Shewanella baltica* SACH
Shewanella sp. MEBiC00475	Marine invertebrate	*Shewanella polaris* MEBiC00475
Shewanella sp. Actino-trap-3	Sea ice	*Shewanella psychromarinicola* Actino-trap-3
Shewanella sp. M2	Deep sea sediments	*Shewanella psychromarinicola* M2
Shewanella sp. R106	Deep sea sediments	*Shewanella psychromarinicola* R106
Shewanella sp. SG44-6	Seawater	*Shewanella vesiculosa* SG44-6
Shewanella sp. SR43-8	Seawater	*Shewanella vesiculosa* SG43-8
Shewanella sp. SG41-3	Seawater	*Shewanella vesiculosa* SG41-3
Shewanella sp. SR43-4	Seawater	*Shewanella vesiculosa* SR43-4
Shewanella sp. SG44-2	Seawater	*Shewanella frigidimarina* SG44-2
Shewanella sp. 11B5	Arctic seawater	*Shewanella frigidimarina* 11B5
Shewanella sp. WXL01	Alga	*Shewanella maritima* WXL01
Shewanella sp. Scap07	Marine invertebrate	*Shewanella waksmanii* Scap07
Shewanella sp. XMDDZSB0408	Marine invertebrate	*Shewanella intestini* XMDDZSB0408
Shewanella sp. 10 N.286.51.B7	Seawater	*Shewanella electrodiphila* 10 N.286.51.B7
Shewanella sp. MMG014	Marine invertebrate	*Shewanella japonica* MMG014

Continued

Table 2. Continued

Strain	Isolation source	Correct name
Shewanella sp. UCD-FRSSP16_17	Marine invertebrate	*Shewanella japonica* UCD-FRSSP16_17
Shewanella sp. P1-14-1	Marine invertebrate	*Shewanella japonica* P1-14-1
Shewanella sp. WPAGA9	Deep sea sediments	*Shewanella japonica* WPAGA9
Shewanella halifaxensis 6JANF4-E-4	Marine fish	*Shewanella fidelis* 6JANF4-E-4

ten random sequence addition replicates. The sequences were checked for a compositional bias using the X^2 test as implemented in PAUP* [24].

Calculation of pairwise SSU similarity values was done according to the recommended method described in [25] and as implemented in the DSMZ phylogeny server for single genes [4].

RESULTS AND DISCUSSION

The genus *Shewanella* is rapidly expanding and in 2019, the genus *Parashewanella* was proposed [8] to accommodate new members of the family *Shewanellaceae*. The rapid expansion of *Shewanellaceae*, in particular the genus *Shewanella*, is illustrated in Fig. 1. For this representation we considered taxa with a validly published name under the ICNP listed in the LPSN [3, 4] as of 22 December 2021, synonyms and ortographic variants excluded. Besides, *S. upenei* (later heterotypic synonym of *S. algae* [12]), *S. pacifica* (later heterotypic synonym of *S. japonica* [12]) and (*Para*)*shewanella ircinae*, for which type strain material availability issues have been reported [8, 26], were also excluded. The species *Shewanella piezotolerans* and *Shewanella psychrophila* [27] were considered, although it should be noted that the corresponding type strains could neither be obtained from the China General Microbiological Culture Collection Centre (CGMCC) nor the Japanese Collection of Microorganisms (JCM) because of quality issues with the deposited type material. Thus, 72 *Shewanella* species, 3 *Parashewanella* species and one *Psychrobium* species were considered for this figure at the moment of writing. Fig. 1 shows a pronounced increase in the number of *Shewanella* species from the second half of the 1990s. This is only partially concomitant with the development and implementation of next-generation sequencing technologies. Thus, numerous type strains had lacked a whole genome sequence, including some described in recent years (Table S1).

To provide a family-wide, whole genome-based taxonomy for *Shewanella* and *Parashewanella*, we retrieved all assemblies available for these genera from NCBI GenBank as of 31 August 2021 and filtered out those excluded from RefSeq for diverse reasons (derived from metagenomes, presence of many frameshifted proteins, highly fragmented assemblies, or genome size too small). The final dataset comprised 381 assemblies (File S1) including the strains sequenced in this study (Table S1). To delineate species and subspecies, pairwise dDDH distances were calculated by the TYGS with the GBDP method, and subjected to a type-based species clustering using the widely accepted cutoffs of 70 and 79% for species and subspecies definitions, respectively [11, 13, 18]. In addition, to provide an improved phylogenomic resolution of distant clades, a whole proteome-based GBDP phylogenomic analysis was inferred. The resulting whole proteome-based and whole genome sequence-based phylogenies are shown in Figs 2–6 and S1, respectively. The resulting species and subspecies clusters are annotated in Figs 2–6 and S1 but pairwise dDDH distances are provided for convenience in File S2 as well. Hereforth, misassigned binomials will be followed by '(MIS)'.

Shewanella spp. genome size ranged from 3.45 Mbp (*Shewanella putrefaciens* HRCR-6 (MIS)) to 7.29 Mbp (*Shewanella* sp. YLB-07), with a G+C content between 40.2 mol% (*Shewanella donghaensis* LT17^T) and 55.7 mol% (*Shewanella* sp. SHSM-M6). *Parashewanella* genomes ranged from 4.22 Mbp (*Parashewanella tropica* MEBiC05444^T) to 5.47 Mbp (*Parashewanella spongiae* HJ039^T). The G+C content of *Parashewanella* spp., in a narrow range from 39.3 mol% (*Parashewanella spongiae* KCTC 22492^T) to 40.8 mol% (*Parashewanella tropica* MEBiC05444^T), is distinctively lower than that of *Shewanella* spp. Of note, the genome assembly obtained for *Shewanella benthica* DSM 8812^T (SAMN24537884) was significantly larger (5.70 Mbp) than that of the same strain available in GenBank (SAMN16273954, 4.03 Mbp). The G+C content of both assemblies was the same (45.76 mol%). The dDDH value (inferred from the sequence identity-based GBDP formula d_4) between both assemblies was 99.9% (File. S2), clearly indicating that the smaller genome is part of the larger one, thus ruling out a contamination (this would have been reflected by a much lower dDDH value). The former interpretation is consistent with the inspection of both assemblies with Mauve [28] (data not shown).

The genome sequence-based and whole proteome-based reconstructions were largely consistent with each other, supporting the existence of overall 124 distinct species and 155 subspecies among *Shewanella* (121 species and 152 subspecies) and *Parashewanella* (3 species, 3 subspecies). However, the genome sequence-based phylogeny (Fig. S1) had a significantly lower branch support on average (45%) compared to the proteome-based one (70%) (Figs. 2–6), yielding an uncertain phylogenetic placement of *Parashewanella* species only in the former case. We thus based our analysis on the better resolved proteome-based reconstruction to infer the main

taxonomic consequences. Our phylogenomic reconstructions supported the identification of 46 novel *Shewanella* species and 2 novel *Parashewanella* species among the sequenced isolates (Table 1). Some of the novel species identified in our 381-genome dataset were described during 2021 but do not have a validly published name yet, except *Shewanella nanhaiensis* Cao *et al*. 2021, with the type strain designated NR704-98^T [29], validly published and included in the LPSN while writing this paper.

A highly populated clade comprising diverse strains belonging to the same novel species is that formed by *Shewanella putrefaciens* 97 (MIS), *Shewanella* sp. M16, *Shewanella* sp. DW31, *Shewanella* sp. S-1, *Shewanella* sp. NKUCC06_TVS, *Shewanella* sp. WE21, *Shewanella putrefaciens* HRCR-6 (MIS), *Shewanella* sp. NKUCC05_KAH, *Shewanella* sp. ISTPL2, *Shewanella* sp. DC2-4, and *Shewanella* sp. NKUCC01_JLK, representing distinct subspecies that inhabit environments spanning from host-associated (marine and freshwater fish) to contaminated sites, thereby showcasing a remarkable ecophysiological adaptability. Further clades comprising four representative strains each are those formed by: (1) *Shewanella* sp. BF02_Schw, *Shewanella* sp. Arc9-LZ, *Shewanella* sp. SR44-4, and *Shewanella* sp. ALD9, isolated from environments spanning from the deep sea to polar; (2) *Shewanella* sp. 10 N.286.48.A6, S*hewanella* sp. 10 N.286.48.B5, *Shewanella* sp. 10 N.286.52.B9, and *Shewanella* sp. 10 N.286.52.C2, isolated from seawater; and (3) *Shewanella* sp. GutCb, *Shewanella* sp. GutDb-MelDb, *Shewanella* sp. Choline-02u-19, and *Shewanella* sp. Bg11-22, isolated from the Arctic. Other clades with more than one representative strain included the ones formed by: (1) *Shewanella* sp. FJAT-52076 and *Shewanella* sp. FJAT-52072, isolated from sediments; (2) *Shewanella* sp. KCT and *Shewanella* sp. FJAT-53555, two distinct subspecies isolated from *Meretrix lusoria* and sediments, respectively; (3) *Shewanella* sp. FJAT-51649 and *Shewanella* sp. Shew256, two distinct subspecies isolated from sediments and a human clinical specimen, respectively, thereby representing a potential disease-causing species [30]; and (4) *Shewanella* sp. strains MBTL60-112-B1 and MBTL60-112-B2 retrieved from marine sediments. Of note, strains *Shewanella* sp. OPT2 (MIS) and *Shewanella* sp. 202IG2-18 (MIS) represent two distinct novel *Parashewanella* species. Overall, the preponderant source of isolates representing novel species were sediments (31/73 isolates from 26 distinct species). Fifteen isolates belonging to 12 distinct novel species were host-associated. In addition to novel species identification, both genome sequence-based and whole proteome-based phylogenomic reconstructions supported the species assignment or reclassification of 43 additional isolates listed in Table 2. Except for the placement of two strains (*Shewanella putrefaciens* SA70 (MIS) and *Shewanella indica* Colony474), the type-based species clustering matched monophyletic groups in the phylogenomic reconstruction throughout. These deviations are due to the circumstance that this dataset is not ultrametric [31]; however, the phylogenomic reconstruction avoids drawing wrong conclusions regarding species affiliation. Thus, *S. putrefaciens* SA70 (MIS) belongs to the species *Shewanella seohaensis,* whereas *S. indica* Colony474 is placed in the same species cluster than that of the type strain of *S. indica*.

Similarity of (partial) 16S rRNA gene sequences is still considered a primary taxonomic marker because of, among other reasons, its historical use, including the existence of comprehensive reference databases as well as the availability and common use of universal primers that amplify variable regions across the almost entire length of the gene. However, while similarity of (partial) 16S rRNA gene sequences generally provides sufficient resolution to delineate the taxonomic position of bacterial isolates at the genus level, there is evidence of its limited usefulness when it comes to species circumscriptions [12, 32–34]. Recent work indicates limitations even at delineating bacterial genera [35]. We were able to extract complete or partial 16S rRNA sequences from 311 of the 381 assemblies (File S3) and used them to infer the phylogeny of the isolates. This reconstruction did not yield sufficient branch support and the resulting topology was not interpretable (Fig. S2), suggesting that the 16S rRNA gene has insufficient phylogenetic resolution in this taxonomic group. The insufficient taxonomic resolution of 16S rRNA is reflected in the pairwise sequence similarities from type strains, with numerous validly published and distinct species exhibiting 16S rRNA gene sequence similarities higher than the accepted 98.7 [11] or 98.8% thresholds [25] for species delineation (File S3). For example, the full-length 16S rRNA gene of *Shewanella khirikhana* TH2012^T has 100% pairwise sequence similarity with respect to the full-length gene of *Shewanella amazonensis* SB2B^T and *Shewanella frigidimarina* KCTC 23109^T, respectively, and 99.28% with respect to the full-length gene of *Shewanella cyperi* FJAT-53720^T. Likewise, the pairwise sequence similarity of the full-length 16S rRNA gene of *Parashewanella curva* C51^T with respect to the full-length genes of *Shewanella colwelliana* ATCC BAA-642^T, *Shewanella decolorationis* S12^T, *Shewanella sediminis* HAW-EB3^T and *Parashewanella tropica* MEBiC05444^T, is 100%, 99.08, 99.02 and 98.95%, respectively. Diverse other cases are compiled in File S3 (see the second tab in this file for a list of the complete or partial 16S rRNA gene lengths extracted from complete or draft genomes). All in all, this highlights the limited usefulness of 16S rRNA gene sequence similarity as a taxonomic marker for *Shewanellaceae* even at the genus level.

In conclusion, through whole-genome sequencing of type strains and subsequent phylogenomic analysis, we have resolved the taxonomy of *Shewanella* and *Parashewanella*, including the identification of 48 novel species, most of which await formal description by the respective authors, thereby substantially increasing the taxonomic diversity of these genera. Our work constitutes a reference for future taxonomic studies as well as comparative genomic analyses aimed at unravelling different aspects of the complex ecophysiology of *Shewanella* and *Parashewanella* species.

Funding information
This study was funded by the Anna och Gunnar Vidfelts fond för biologisk forskning, grant number: 2019-051-Vidfelts fond/SOJOH (A.J.M.-R.). Ferring Phamaceuticals funded the Center for Translational Microbiome Research where most sequencing efforts were carried out.

Author contributions

Conceptualization: A.J.M.-R.; methodology: A.J.M.-R., J.P.M.-K.; software: J.P.M-K.; validation: A.J.M.-R., J.P.M.-K.; formal analysis: J.P.M.-K.; investigation: A.J.M.-R., J.P.M.-K.; resources: A.J.M.-R., J.P.M.-K.; data curation: A.J.M.-R., J.P.M.-K.; writing – original draft preparation: A.J.M.-R.; writing – review and editing: A.J.M.-R., J.P.M.-K.; visualization: A.J.M.-R., J.P.M.-K.; supervision: A.J.M.-R.; project administration: A.J.M.-R.; funding: A.J.M.-R.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. Martín-Rodríguez AJ, Meier-Kolthoff JP. Whole genome-based taxonomy of *Shewanella* and *Parashewanella*. *Figshare*. 2022. DOI: 10.6084/m9.figshare.19633374.
2. MacDonell MT, Colwell RR. Phylogeny of the vibrionaceae, and recommendation for two new genera, *Listonella* and *Shewanella*. *Syst Appl Microbiol* 1985;6:171–182.
3. Parte AC, Sardà Carbasse J, 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;70:5607–5612.
4. Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 2022;50:D801–D807.
5. Martín-Rodríguez AJ, Martín-Pujol O, Artiles-Campelo F, Bolaños-Rivero M, Römling U. *Shewanella* spp. infections in Gran Canaria, Spain: retrospective analysis of 31 cases and a literature review. *JMM Case Rep* 2017;4:e005131.
6. Janda JM, Abbott SL. The genus *Shewanella*: from the briny depths below to human pathogen. *Crit Rev Microbiol* 2012;7828:1–21.
7. Nogi Y, Abe M, Kawagucci S, Hirayama H. *Psychrobium conchae* gen. nov., sp. nov., a psychrophilic marine bacterium isolated from the Iheya North hydrothermal field. *Int J Syst Evol Microbiol* 2014;64:3668–3675.
8. Xu S, Yu K, Su H, Chen B, Huang W, *et al*. Proposal of *Parashewanella* gen. nov. to accommodate *Parashewanella curva* sp. nov. and *Parashewanella spongiae* comb. nov. in the *Shewanellaceae*. *Int J Syst Evol Microbiol* 2019;69:1259–1264.
9. Yang SH, Kwon KK, Lee HS, Kim SJ. *Shewanella spongiae* sp. nov., isolated from a marine sponge. *Int J Syst Evol Microbiol* 2006;56:2879–2882.
10. Vogel BF, Venkateswaran K, Christensen H, Falsen E, Christiansen G, *et al*. Polyphasic taxonomic approach in the description of *Alishewanella fetalis* gen. nov., sp. nov., isolated from a human foetus. *Int J Syst Evol Microbiol* 2000;50 Pt 3:1133–1142.
11. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, *et al*. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461–466.
12. Thorell K, Meier-Kolthoff JP, Sjöling Å, Martín-Rodríguez AJ. Whole-genome sequencing redefines *Shewanella* taxonomy. *Front Microbiol* 2019;10:1861.
13. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 2019;10:2182.
14. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:1–14.
15. Lefort V, Desper R, Gascuel O. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol* 2015;32:2798–2800.
16. Farris JS. Estimating phylogenetic trees from distance matrices. *Am Nat* 1972;106:645–668.
17. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 2019;47:W256–W259.
18. Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, *et al*. Complete genome sequence of DSM 30083(T), the type strain (U5/41(T)) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci* 2014;9:2.
19. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, *et al*. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007;35:3100–3108.
20. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–1797.
21. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312–1313.
22. Goloboff PA, Farris JS, Nixon KC. TNT, a free program for phylogenetic analysis. *Cladistics* 2008;24:774–786.
23. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. How many bootstrap replicates are necessary? *J Comput Biol* 2010;17:337–354.
24. Swofford DL. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0 b10. Sinauer Assoc Sunderl,
25. Meier-Kolthoff JP, Göker M, Spröer C, Klenk H-P. When should a DDH experiment be mandatory in microbial taxonomy? *Arch Microbiol* 2013;195:413–418.
26. Rameshkumar N. The status of the species *Shewanella irciniae* Lee *et al.* 2006. Request for an Opinion. *Int J Syst Evol Microbiol* 2015;65:2774.
27. Xiao X, Wang P, Zeng X, Bartlett DH, Wang F. *Shewanella psychrophila* sp. nov. and *Shewanella piezotolerans* sp. nov., isolated from west Pacific deep-sea sediment. *Int J Syst Evol Microbiol* 2007;57:60–65.
28. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, *et al*. Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics* 2009;25:2071–2073.
29. Cao W-R, Li X, Sun Y-Y, Jiang M-Y, Xu X-D, *et al*. *Shewanella nanhaiensis* sp. nov., a marine bacterium isolated from sediment of south china sea, and emended descriptions of *Shewanella woodyi*, *Shewanella hanedai* and *Shewanella canadensis*. *Int J Syst Evol Microbiol* 2021;71:005152.
30. Almuzara M, Montaña S, Lazzaro T, Uong S, Parmeciano Di Noto G, *et al*. Genetic analysis of a PER-2-producing *Shewanella* sp. strain harbouring a variety of mobile genetic elements and antibiotic resistance determinants. *J Glob Antimicrob Resist* 2017;11:81–86.
31. Meier-Kolthoff JP, Klenk HP, Göker M. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 2014;64:352–356.
32. Rossi-Tamisier M, Benamar S, Raoult D, Fournier P-E. Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. *Int J Syst Evol Microbiol* 2015;65:1929–1934.
33. Janda JM, Abbott SL. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* 2007;45:2761–2764.
34. Martín-Rodríguez AJ, Suárez-Mesa A, Artiles-Campelo F, Römling U, Hernández M. Multilocus sequence typing of *Shewanella algae* isolates identifies disease-causing *Shewanella chilikensis* strain 6I4. *FEMS Microbiol Ecol* 2019;95.
35. Barco RA, Garrity GM, Scott JJ, Amend JP, Nealson KH, *et al*. A genus definition for *Bacteria* and *Archaea* based on standard genome relatedness index. *MBio* 2020;11:e02475-19.
36. Liu G-H, Zhang Q, Narsing Rao MP, Yang S, Tang R, *et al*. Stress response mechanisms and description of three novel species *Shewanella avicenniae* sp. nov., *Shewanella sedimenti-mangrovi* sp. nov. and *Shewanella yunxiaonensis* sp. nov., isolated from mangrove ecosystem. *Antonie van Leeuwenhoek* 2021;114:2123–2131.

37. Yu L, Jian H, Gai Y, Yi Z, Feng Y, *et al*. Characterization of two novel psychrophilic and piezotolerant strains, *Shewanella psychropiezotolerans* sp. nov. and *Shewanella eurypsychrophilus* sp. nov, adapted to an extreme deep-sea environment. *Syst Appl Microbiol* 2021;44:126266.