

# Genome-based analyses reveal a synonymy among *Halorubrum distributum* Zvyagintseva and Tarasov 1989; Oren and Ventosa 1996, *Halorubrum terrestre* Ventosa *et al.* 2004, *Halorubrum arcis* Xu *et al.* 2007 and *Halorubrum litoreum* Cui *et al.* 2007. Emended description of *Halorubrum distributum* Zvyagintseva and Tarasov 1989; Oren and Ventosa 1996

Carmen Infante-Domínguez<sup>1†</sup>, Rafael R. de la Haba<sup>1†</sup>, Paulina Corral<sup>1</sup>, Cristina Sanchez-Porro<sup>1</sup>, David R. Arahal<sup>2</sup> and Antonio Ventosa<sup>1,\*</sup>

## Abstract

A comparative taxonomic study of *Halorubrum distributum*, *Halorubrum terrestre*, *Halorubrum arcis* and *Halorubrum litoreum* was carried out using different approaches, 16S rRNA gene sequence analysis, multilocus sequence analysis (MLSA), phylogenomic analysis based on the comparison of the core genome, orthologous average nucleotide identity (OrthoANI), Genome-to-Genome Distance Calculator (GGDC), synteny plots and polar lipid profile (PLP). The MLSA study, using the five concatenated housekeeping genes *atpB*, *EF-2*, *glnA*, *ppsA* and *rpoB*, and the phylogenomic analysis based on 1347 core translated gene sequences obtained from their genomes showed that *Halorubrum distributum* JCM 9100<sup>T</sup>, *Halorubrum terrestre* JCM 10247<sup>T</sup>, *Halorubrum arcis* JCM 13916<sup>T</sup> and *Halorubrum litoreum* JCM 13561<sup>T</sup> formed a robust cluster, clearly separated from the rest of species of the genus *Halorubrum*. The OrthoANI and digital DDH values, calculated by the GGDC, showed percentages among *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> that ranged from 98.1 to 97.5%, and 84.0 to 78.0%, respectively, while these values among those strains and the type strains of their most related species of *Halorubrum* were equal or lower than 90.8 and 41.2 %, respectively. Moreover, degree of synteny across the four genomes was very high, especially between the genomes of *Halorubrum litoreum* JCM 13561<sup>T</sup> and *Halorubrum arcis* JCM 13916<sup>T</sup>. In addition, the PLP is quite similar among the four strains studied, showing a common pattern typical of the neutrophilic species of the genus *Halorubrum*. Overall, these data show that *Hrr. distributum*, *Hrr. terrestre*, *Hrr. arcis* and *Hrr. litoreum* constitute a single species. Thus, the latter three should be considered as later, heterotypic synonyms of *Hrr. distributum* based on the rules for priority of names. We propose an emended description of *Hrr. distributum*, including the features of *Hrr. terrestre*, *Hrr. arcis* and *Hrr. litoreum*.

The genus *Halorubrum* was proposed in 1995 by McGenity and Grant [1] in order to reclassify several species previously included in the genus *Halobacterium*. Two additional *Halorubrum* species, reclassified as members of the genus *Halorubrobacterium* [2], were lately transferred into this genus [3]. This genus was emended in 2009 by Oren *et al.*

[4]. Currently, the genus *Halorubrum* is classified within the family *Halorubraceae*, order *Haloferacales*, class *Halobacteria* [5, 6], with *Halorubrum saccharovororum* as the type species. As of June 2019, the genus *Halorubrum*, with 39 species with validly published names, includes the largest number of species among all the genera within the haloarchaea [7, 8].

**Author affiliations:** <sup>1</sup>Department of Microbiology and Parasitology, Faculty of Pharmacy, University of Sevilla, 41012 Sevilla, Spain; <sup>2</sup>Department of Microbiology and Ecology, and Spanish Type Culture Collection (CECT), University of Valencia, 46980 Paterna (Valencia), Spain.

**\*Correspondence:** Antonio Ventosa, ventosa@us.es

**Keywords:** *Halorubrum*; taxonomy; synonymy; emended description; *Halorubrum distributum*.

**Abbreviations:** ANI, average nucleotide identity; DDH, DNA–DNA hybridization; GGDC, genome-to-genome distance calculator; MLSA, multilocus sequence analysis; OG, orthologous genes; OrthoANI, orthologous average nucleotide identity; PLP, polar lipid profile.

The GenBank/EMBL/DDBJ accession numbers of the 16S rRNA gene sequence and complete genome sequence of the type strain JCM 9100<sup>T</sup> are D63572 and AOJM00000000, respectively.

†These authors contributed equally to this work

Four supplementary tables and three supplementary figures are available with the online version of this article.

003956 © 2020 The Authors



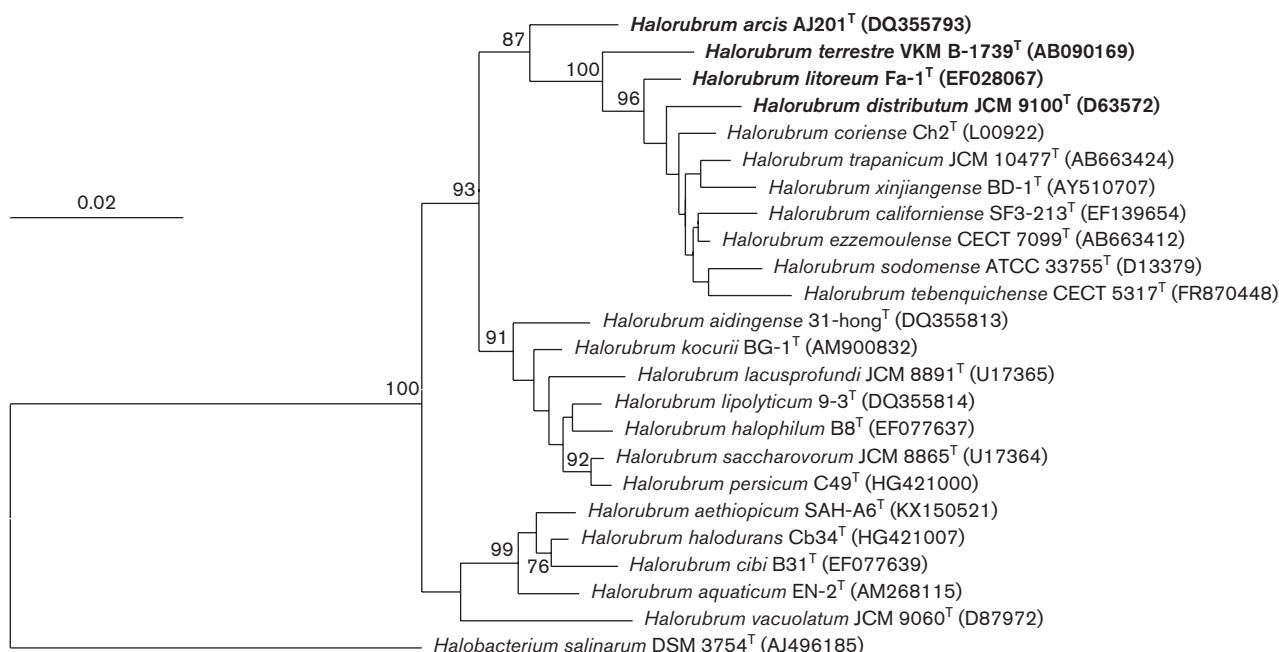
This is an open-access article distributed under the terms of the Creative Commons Attribution License.

These species were isolated from diverse hypersaline habitats, such as saline and soda lakes, salterns, coastal sabkhas or saline soils, as well as from rock salt and salted foods [4, 8]. The genus *Halorubrum* includes rods or pleomorphic cells, Gram-variable, most species are motile and their colonies are red- to orange-pigmented. They are oxidase- and catalase-positive, extremely halophilic archaea, neutrophilic and in some cases alkaliphilic (optimum pH 9–10), such as *Halorubrum alkaliphilum*, *Halorubrum gandharaense*, *Halorubrum luteum*, *Halorubrum tibetense* and *Halorubrum vacuolatum*. Their major polar lipids are phosphatidylglycerol, phosphatidylglycerol phosphate methyl ester, phosphatidylglycerol sulfate and a sulfated mannosyl glucosyl diether. However, alkaliphilic species lack phosphatidylglycerol sulfate and glycolipids. Their genomic DNA G+C content range is 60.2–71.2 mol% [7, 8]. In the past, there has been a tendency to characterize and delineate new species of haloarchaea on the basis of their 16S rRNA gene sequence comparison, and at a later stage DNA–DNA hybridization (DDH) and phenotypic features, including the polar lipid analysis with their 16S rRNA-phylogenetically most closely related species. However, several studies pointed out that the use of the 16S rRNA gene was not adequate as a phylogenetic marker for this archaeal group due to the extensive horizontal gene transfer events, the presence of several rRNA operons (in some cases divergent) in their genomes and the low evolutionary rate of this gene [9–12]. Alternative methodologies such as the use of the *rpoB'* gene or a multilocus sequence analysis (MLSA) based on the comparison of several single-copy housekeeping genes have been used [13, 14]. Recently, we have proposed the use of the five housekeeping genes *atpB*, *EF-2*, *glnA*, *ppsA* and *rpoB'* for such phylogenetic analyses as an alternative to the 16S rRNA gene [15]. In this study, we determined that some well-established species of *Halorubrum* were not separate species, such as in the case of *Halorubrum ezzemoulense* and *Halorubrum chaoviator* [15, 16] and other similar situations might happen for other species of *Halorubrum*.

Here, we have studied in detail four species of *Halorubrum*, *Hrr. distributum* and *Hrr. terrestre*, isolated from saline soils [17, 18], *Hrr. arcis*, isolated from a saline lake [19] and *Hrr. litoreum*, isolated from a saltern [20], for which the draft genome sequences are available. The results show that they are members of the same species, *Hrr. distributum* being the species name with priority according to the International Code of Nomenclature of Prokaryotes [21], and thus it should be considered as an earlier synonym of the other three species of *Halorubrum*.

The 16S rRNA gene and MLSA phylogenetic analyses were carried out as described previously [14, 15]. Since some members of the class *Halobacteria* possess highly divergent 16S rRNA genes, we compared the reference 16S rRNA gene sequence (see in Table S1, available in the online version of this article) of each of the analysed *Halorubrum* species with the 16S rRNA gene sequences extracted from the corresponding genome (partial 16S rRNA gene sequences <500 bp or low quality sequences were excluded). All the 16S rRNA gene sequences obtained from the genomes showed a similarity between 98.4–100% with the

16S rRNA gene reference sequence. A more detailed analysis of the pairwise alignments unveiled that lower similarities (i.e. 98.4%) were due to genome misassembly and, therefore, only microheterogeneities among 16S rRNA gene copies but not highly divergent copies were detected in the species of the genus *Halorubrum*. The nucleotide sequence of the 16S rRNA gene from the strains was aligned with ARB 6.0.5 software package [22]. The sequence similarity analysis was carried out by comparing the 16S rRNA gene sequence of *Halorubrum distributum* JCM 9100<sup>T</sup>, *Halorubrum terrestre* VKM B-1739<sup>T</sup>, *Halorubrum arcis* AJ201<sup>T</sup> and *Halorubrum litoreum* Fa-1<sup>T</sup> with the known sequences of the *Halorubrum* species showed in Table S1, using the EzBioCloud tool ([www.ezbiocloud.net/eztaxon](http://www.ezbiocloud.net/eztaxon); [23]) and ARB 6.0.5 software. The analysis based on the complete 16S rRNA gene sequences showed the percentages of similarity included in Table S2. The 16S rRNA gene sequence of the type strain of *Hrr. distributum* JCM 9100<sup>T</sup> with respect to those of *Hrr. terrestre* VKM B-1739<sup>T</sup>, *Hrr. arcis* AJ201<sup>T</sup> and *Hrr. litoreum* Fa-1<sup>T</sup> showed a similarity of 97.3, 95.8 and 98.6%, respectively, which might be interpreted as those three strains belonging to different species than *Hrr. distributum*, according to the current accepted 16S rRNA gene sequence similarity threshold for species differentiation (98.65% as a general limit, 98.2% specifically for the phylum *Euryarchaeota*) [24, 25]. However, as indicated above, the 16S rRNA gene is not considered a reliable phylogenetic marker for this archaeal group and, therefore, these results must be regarded with caution and carefully checked. The phylogenetic study based on the 16S rRNA gene sequence comparison was performed by reconstructing phylogenetic trees using the following algorithms: neighbour-joining, maximum-parsimony and maximum-likelihood. Neighbour-joining distances were corrected according to the Jukes–Cantor matrix [26] and phylogenetic neighbour-joining and maximum parsimony trees were calculated with the ARB program package version 6.0.5 [22]. Maximum-likelihood analysis was performed with PhyML version 3.3.20180214 [27] using the General Time Reversible (GTR+I+G) model of nucleotide substitution [28]. Base frequency per alignment position and positional variability by parsimony analysis of archaea taxa filters were applied in the sequence comparison and phylogenetic reconstruction analyses and the effects on the results were evaluated. To evaluate the robustness of the phylogenetic tree, a bootstrap analysis (1000 replications) was performed [29]. The phylogenetic tree based on the 16S rRNA gene reconstructed by neighbour-joining showed that *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* VKM B-1739<sup>T</sup>, *Hrr. arcis* AJ201<sup>T</sup> and *Hrr. litoreum* Fa-1<sup>T</sup> were phylogenetically related but they did not constitute a single cluster (Fig. 1). The topologies of phylogenetic trees inferred using the maximum-likelihood (Fig. S1) and maximum-parsimony algorithms were very similar to that of the tree reconstructed by neighbour-joining. As previously indicated, the comparison of the 16S rRNA gene sequences do not permit to determine in depth the phylogenetic relationships within the genus *Halorubrum* and thus an MLSA approach based on the comparison of partial sequences of the *atpB*, *EF-2*, *glnA*, *ppsA* and *rpoB'* housekeeping genes (Table S1) has been recently recommended for this genus [15]. Fig. 2 shows the phylogenetic tree obtained by the concatenation of these five



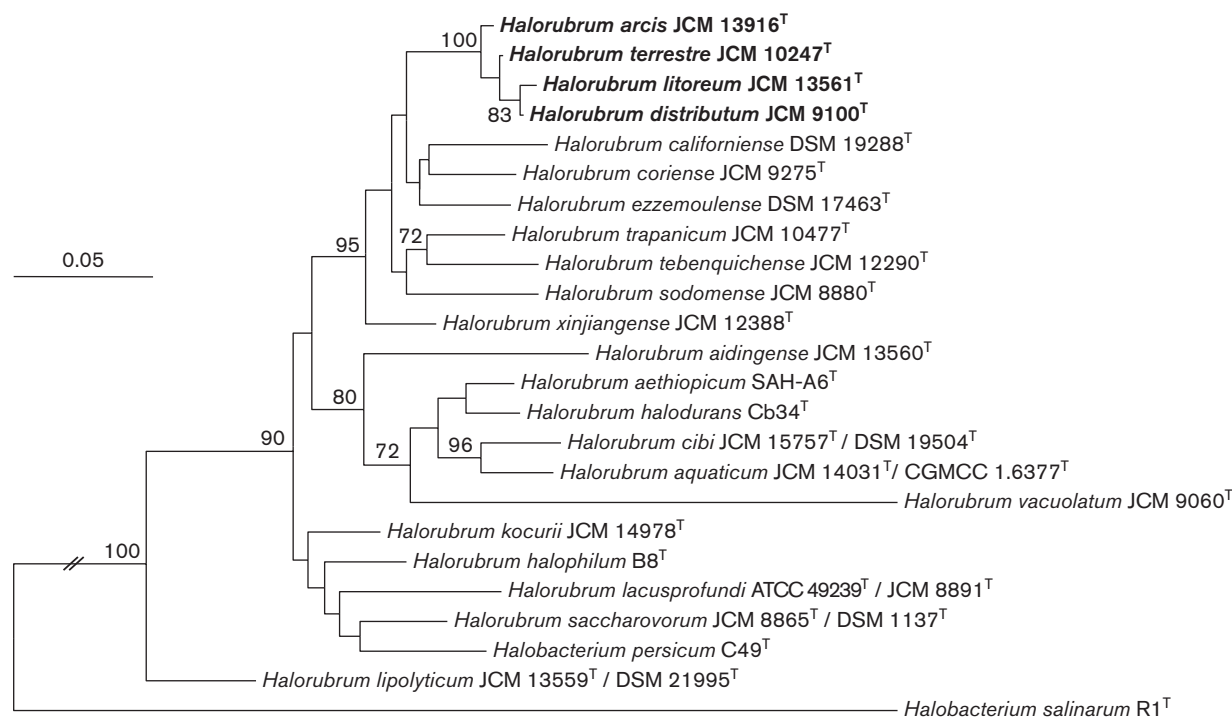
**Fig. 1.** Neighbour-joining phylogenetic tree based on 16S rRNA gene sequences comparison showing the relationships between *Halorubrum distributum*, *Halorubrum terrestre*, *Halorubrum arcis* and *Halorubrum litoreum* and other related species of the genus *Halorubrum*. The accession numbers of the sequences used are shown in parentheses after the strain designation. Bootstrap values (%) based on 1000 replicates are shown for branches with more than 70% bootstrap support. *Halobacterium salinarum* DSM 3754<sup>T</sup> was used as an outgroup. Bar, 0.02 substitutions per nucleotide position.

housekeeping genes, reconstructed by the maximum-likelihood algorithm following the same methodology as aforementioned. This tree shows a better phylogenetic separation of the species of *Halorubrum* and on the other hand, in this case *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> constituted a well-supported cluster, separated from other species of the genus *Halorubrum*. The percentage of similarity of the five concatenated genes of *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> ranged between 98.3–99.3% and those of these four species with respect to the other type strains of *Halorubrum* were always below 95.6% (Table S3), below the 96% cut-off value established for species delineation within the genus *Halorubrum* [15].

To increase the resolution, we carried out a phylogenomic analysis based on the 1347 core translated gene sequences obtained from the available genomes of *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> and the type strains of other related *Halorubrum* species (Tables S1 and S4). In brief, the translated gene sequences of all the predicted coding sequences of the genomes under study were compared by BLAST search in an all-versus-all mode to identify clusters of orthologous genes (OGs), that is, genes with a common ancestor present in two or more species. Those OGs shared among all taxa constituted the core genome of the studied genomes, formed by 1415 genes. From those, only the ones present in single copy per genome (a total of 1347 genes) were selected to reconstruct

the phylogenomic tree. Enveomics collection toolbox [30] was used to carry out the aforementioned OG identification and core-genome selection. The 1347 translated gene sequences for each species were aligned with MUSCLE version 3.8.31 [31] and subsequently concatenated. An approximate maximum-likelihood tree was reconstructed using FastTree version 2.1.9 [32] with the JTT replacement matrix [33] under the CAT approximation (single rate for each site) with 20 rate categories. Local support values were estimated with the Shimodaira–Hasegawa test [34]. As shown in Fig. 3, the overall topology of the phylogenetic tree was in agreement with the MLSA tree. *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> formed a maximally supported subtree, clearly separated from the rest of species of *Halorubrum* (Fig. 3). A total of 2571 common OGs are shared among the four strains, while pairwise comparisons between the four species studied yielded 2665–2769 shared OGs (Fig. 4).

The OrthoANI percentages, determined on the basis of USEARCH comparison of the genome sequences of *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> with respect to the type strains of the related *Halorubrum* species, as implemented in OrthoANIu tool [35], indicate that the cluster formed by these strains showed a range of 97.5–98.1% among these strains, while the percentage with respect to the related species of *Halorubrum* was equal or lower than 90.8% (Table 1). The threshold of 95–96% defined for species



**Fig. 2.** Maximum-likelihood phylogenetic tree based on the five housekeeping gene (*atpB*, *EF-2*, *glnA*, *ppsA* and *rpoB*) concatenated sequences showing the relationship between *Halorubrum distributum*, *Halorubrum terrestre*, *Halorubrum arcis* and *Halorubrum litoreum* and other related species of the genus *Halorubrum*. The accession numbers of the sequences used are shown in Table S1. Bootstrap values (%) based on 1000 replicates are indicated for branches above 70%. *Halobacterium salinarum* DSM 3754<sup>T</sup> was used as an outgroup. Bar, 0.05 substitutions per nucleotide position.

delineation [36–38] clearly supports the placement of these four strains within a single species.

We also calculated the digital DDH values, determined online (<http://ggdc.dsmz.de/distcalc2.php>) using the Genome-to-Genome Distance Calculator (GGDC) version 2.0 as described by Meier-Kolthoff *et al.* [39]. The estimated digital DDH values were calculated using formula two at the GGDC website, originally described by Auch *et al.* [40] and updated by Meier-Kolthoff *et al.* [39]. The digital DDH values among *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> ranged from 84.0 to 78.0%, but the values among these strains and the type strains of species of *Halorubrum* was equal or lower than 41.2% (Table 1), showing unequivocally that the four species under study constitute a single species of *Halorubrum*, clearly separated from the rest of species of this genus. The cut-off for this parameter is well established as 70% for species delineation [38, 39].

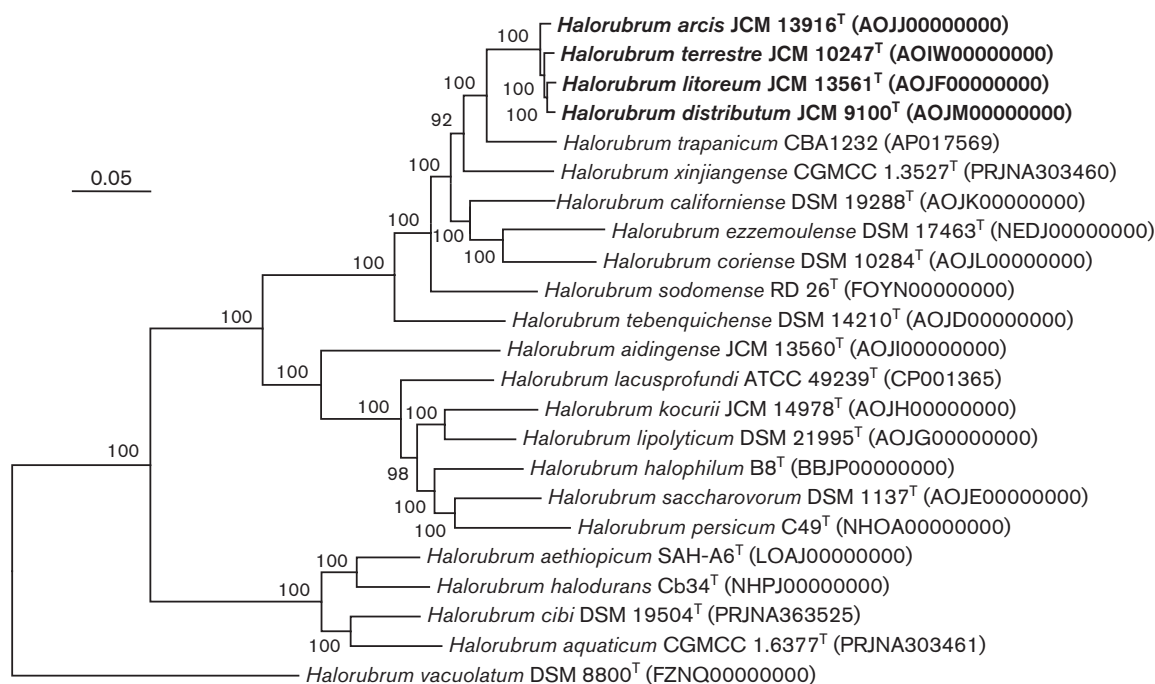
Given the high values of orthoANI and digital DDH among the four species *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup>, we inspected the conservation of blocks of order across those genomes by means of synteny plots constructed using the EDGAR 2.0 software platform [41]. As shown in Fig. S2, gene co-localization is highly conserved in those four

genomes, especially remarkable between *Hrr. litoreum* JCM 13561<sup>T</sup> and *Hrr. arcis* JCM 13916<sup>T</sup> genomes, also supporting the existence of a synonymy among the four species.

DNA G+C content was calculated from the genomic sequences (Table S1) and compared to the DNA G+C content provided in the species descriptions. For *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> the *in silico* values were 68.1, 68.0, 67.3 and 68.9 mol%, respectively, whereas the wet values were 63.9 ( $T_m$ ) [17], 64.4 ( $T_m$ ) [18], 65.7 ( $T_m$ ) [19] and 64.9 mol% ( $T_m$ ) [20]. Therefore, significant discrepancies exist between *in silico* and conventional G+C values probably due to the inaccuracies of the applied conventional method and, accordingly, emendation of the species descriptions seems necessary. Although *in silico* G+C values among the strains belonging to the same species should not vary more than 1% [42], in this case, all the strains are within this 1% difference with respect to *Hrr. distributum* JCM 9100<sup>T</sup>, the type strain of the species name with priority.

For the polar lipid profile (PLP) analyses, the cell biomass of the strains was obtained after 10 days of aerobic incubation in modified SW20 liquid medium under optimal conditions: 20% (w/v) NaCl, 37 °C and pH 7.5. The extraction of membrane polar lipids of halophilic archaea and their profile detection were performed following the methods described by Corral *et al.*





**Fig. 3.** Approximately maximum-likelihood core protein phylogenomic tree including the genomes of *Halorubrum distributum*, *Halorubrum terrestre*, *Halorubrum arcis* and *Halorubrum litoreum* and other related species of the genus *Halorubrum*. This tree was based on the JTT substitution matrix under the CAT approximation with 20 rate categories from the alignment of 1342 shared orthologous single-copy translated genes of these genomes. All genomes were retrieved from GenBank (Table S1). Local support values based on the Shimodaira–Hasegawa test over 70% are shown above the branch. Bar, 0.05 substitutions per nucleotide position.

[43, 44]. The lipids pattern was obtained by one-dimensional high-performance thin-layer chromatography (HPTLC) and then revealed by universal and specific stains for the identification of chemical nature of the polar lipid bands present in the HPTLC plate [43–46]. The lipid extracts of *Halobacterium salinarum* DSM 3754<sup>T</sup>, *Halorubrum saccharovorum* DSM 1137<sup>T</sup> and *Halorubrum tibetense* JCM 11889<sup>T</sup>, as representative of the alkaliphilic species, were used as reference in this study.

The thin-layer chromatography results of the polar lipids (Fig. S3) revealed that *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> possess a similar pattern of polar lipid profile, showing in common the major lipids: phosphatidylglycerol, phosphatidylglycerol phosphate methyl ester, phosphatidylglycerol sulfate and the main polar lipid component of this genus: sulfated mannosyl glucosyl diether. Biphosphatidylglycerol was also found and minor phospholipids were also detected. The polar lipid profile of all these strains possessed all major lipids described for neutrophilic species of the genus *Halorubrum* [1, 4].

In summary, the genome-based study shows that *Hrr. distributum* JCM 9100<sup>T</sup>, *Hrr. terrestre* JCM 10247<sup>T</sup>, *Hrr. arcis* JCM 13916<sup>T</sup> and *Hrr. litoreum* JCM 13561<sup>T</sup> constitute a single species. In accordance to the International Code of Nomenclature of Prokaryotes [21], the name *Hrr. distributum* has priority and thus, *Hrr. terrestre*, *Hrr. arcis* and *Hrr. litoreum*

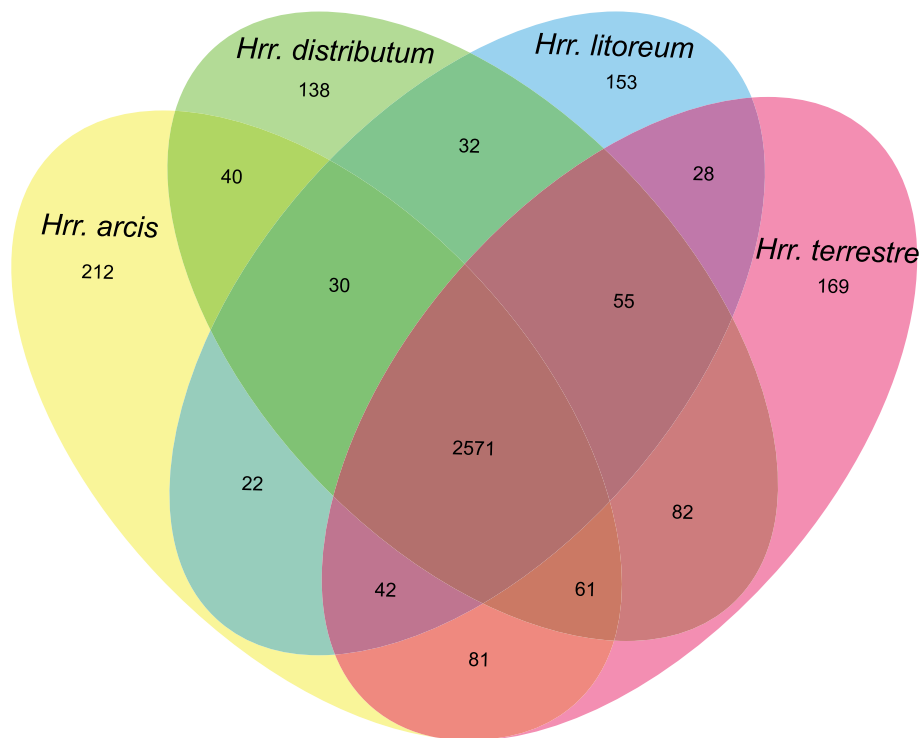
should be considered as later heterotypic synonyms of *Hrr. distributum*. Based on these corroborations we propose the emended description of the species *Hrr. distributum*, including the features of *Hrr. terrestre*, *Hrr. arcis* and *Hrr. litoreum*.

**EMENDED DESCRIPTION OF HALORUBRUM DISTRIBUTUM ZVYAGINTSEVA AND TARASOV 1989; OREN AND VENTOSA 1996**

*Halorubrum distributum* (dis.tri.bu'tum. L. neut. part. adj. *distributum*, distributed [widely]).

The description is that of Zvyagintseva and Tarasov [17] and Oren and Ventosa [3] with the following modifications: cells are rod-shaped, pleomorphic, flat or disk-shaped, aerobic growth occurs at 12–30% (w/v) NaCl, pH 5.0–9.0 and 20–55 °C. Optimum NaCl concentration, pH and temperature for growth are 20–25% (w/v), pH 7.0–7.5 and 37–42 °C. The polar lipid profile includes phosphatidylglycerol, phosphatidylglycerol phosphate methyl ester, phosphatidylglycerol sulfate and one glycolipid chromatographically identical to sulfated mannosyl glucosyl diether. Biphosphatidylglycerol is also found as a minor component, and minor phospholipids are also detected.

The G+C content of the genomic DNA is 67.3–68.9 mol% (genome).



**Fig. 4.** Venn diagram showing the core orthologous and unique proteins for the type strains of *Halorubrum distributum* JCM 9100<sup>T</sup>, *Halorubrum terrestre* JCM 10247<sup>T</sup>, *Halorubrum arcis* JCM 13916<sup>T</sup> and *Halorubrum litoreum* JCM 13561<sup>T</sup>.

**Table 1.** GGDC (lower triangle) and OrthoANlu (upper triangle) values for the type strains of the species of the genus *Halorubrum* with available genomes. Cells are filled with darker colours for higher values.

GGDC/OrthoANlu	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. <i>Hrr. arcis</i>	100	97.7	97.6	97.5	90.5	89.1	88.7	87.7	87.5	88.3	86.9	81.9	82.4	83.2	83.3	83.0	82.8	81.0	81.1	80.5	81.0	77.9	
2. <i>Hrr. terrestre</i>	78.6	100	97.9	97.9	90.6	89.1	88.8	87.8	87.3	88.6	86.8	81.9	82.5	83.2	83.4	83.6	83.0	83.1	81.1	81.2	80.7	81.0	77.9
3. <i>Hrr. litoreum</i>	78.0	83.2	100	98.1	90.8	89.0	89.0	87.7	87.5	88.4	86.9	81.8	82.6	83.0	83.4	83.5	83.1	83.5	81.4	81.3	80.8	80.7	78.1
4. <i>Hrr. distributum</i>	79.3	82.7	84.0	100	90.7	89.3	88.8	88.0	87.3	88.5	86.7	81.7	82.7	83.2	83.6	83.6	83.2	82.8	81.2	81.4	80.8	80.9	78.0
5. <i>Hrr. trapanicum</i>	41.2	41.1	41.2	40.9	100	88.8	89.1	87.6	87.6	89.6	87.0	82.0	82.4	83.1	83.4	83.4	83.1	83.1	81.3	81.1	80.8	80.8	78.0
6. <i>Hrr. xinjiangense</i>	37.1	37.2	37.4	37.2	36.8	100	88.4	87.3	87.3	87.8	86.5	81.7	82.3	83.2	83.0	83.1	82.8	82.8	81.2	80.9	80.8	80.6	77.8
7. <i>Hrr. californiense</i>	36.8	36.5	36.7	36.6	36.6	35.7	100	88.1	88.1	87.5	86.5	81.7	81.9	82.8	82.8	83.0	82.4	83.0	80.8	80.7	80.8	80.6	78.0
8. <i>Hrr. ezzemoulense</i>	34.2	34.7	34.0	34.6	33.6	33.4	34.9	100	88.4	87.2	86.0	81.5	81.7	82.7	82.9	82.9	82.6	82.6	80.4	80.8	80.3	81.0	77.8
9. <i>Hrr. coriense</i>	33.8	33.9	33.5	33.6	33.4	33.4	35.5	35.8	100	86.9	86.3	81.4	82.2	82.8	83.0	82.8	82.6	82.7	81.1	81.0	80.9	80.7	77.8
10. <i>Hrr. sodomense</i>	35.6	35.7	35.7	35.6	38.0	34.2	33.8	32.7	32.5	100	86.5	81.6	82.2	82.6	83.1	82.7	82.7	82.6	81.0	80.8	80.5	80.4	77.5
11. <i>Hrr. tebenquichense</i>	32.0	32.1	32.0	32.0	31.9	31.5	31.3	30.9	31.3	31.4	100	81.7	82.0	82.6	82.9	83.1	82.9	82.7	81.0	81.0	80.7	80.9	77.5
12. <i>Hrr. aidingense</i>	25.6	25.6	25.5	25.4	25.6	25.5	25.2	25.1	25.3	25.1	25.3	100	83.2	83.3	84.0	83.4	83.2	83.3	80.7	80.6	80.3	80.5	77.7
13. <i>Hrr. lacusprofundi</i>	26.1	25.9	25.9	26.0	25.8	25.5	25.6	25.3	25.3	25.4	25.5	27.2	100	87.3	88.3	87.6	87.6	86.9	80.9	80.6	80.5	80.8	77.8
14. <i>Hrr. kocurii</i>	26.8	26.8	26.5	26.7	26.5	26.7	26.6	26.2	26.7	26.3	26.2	27.7	33.4	100	89.4	87.9	87.6	87.4	81.4	81.1	81.4	81.3	77.9
15. <i>Hrr. lipolyticum</i>	27.0	27.2	27.0	27.2	26.9	26.8	26.6	26.3	26.5	26.6	26.4	28.2	34.9	38.5	100	88.6	88.6	87.8	81.6	81.5	81.1	81.2	78.1
16. <i>Hrr. halophilum</i>	27.4	27.5	27.2	27.5	26.9	26.7	26.7	27.0	26.3	26.3	26.6	27.5	34.0	34.7	36.5	100	88.8	88.0	81.3	81.4	81.0	81.4	77.8
17. <i>Hrr. saccharovororum</i>	26.8	26.7	26.8	26.8	26.6	26.4	26.2	26.2	26.1	26.2	26.2	27.4	34.2	34.7	36.2	36.9	100	87.8	81.5	81.5	80.9	81.2	77.8
18. <i>Hrr. persicum</i>	26.7	26.6	27.3	26.6	26.4	26.6	26.3	26.1	26.1	26.1	26.1	27.3	33.1	33.8	34.7	34.7	34.5	100	81.3	81.2	81.3	81.2	78.1
19. <i>Hrr. aethiopicum</i>	25.0	25.2	25.2	25.0	25.0	24.9	24.8	24.5	24.9	24.8	25.0	24.8	24.5	25.3	25.2	25.1	25.1	25.0	100	91.7	89.1	88.4	78.5
20. <i>Hrr. halodurans</i>	25.0	25.0	25.1	25.1	24.8	24.9	24.6	24.9	24.6	24.7	24.8	24.9	24.3	24.9	25.1	25.5	25.7	25.1	45.0	100	88.9	88.2	78.4
21. <i>Hrr. cibi</i>	24.8	24.8	24.9	24.7	24.6	24.9	24.9	24.3	24.7	24.2	24.6	24.6	24.2	25.7	24.9	24.8	24.7	25.0	37.3	36.7	100	89.9	78.3
22. <i>Hrr. aquaticum</i>	25.0	25.3	24.8	25.0	24.7	24.6	24.5	25.1	24.6	24.4	24.9	24.9	24.4	25.3	25.3	25.4	25.0	25.1	35.2	34.9	39.2	100	78.4
23. <i>Hrr. vacuolatum</i>	22.8	22.8	22.8	22.8	22.4	22.9	22.8	22.7	22.6	22.3	22.7	22.6	23.3	23.1	23.0	22.8	23.0	23.1	23.4	23.2	23.3	23.3	100

The type strain, 1m<sup>T</sup> (=VKMB-1733<sup>T</sup>=ATCC51197<sup>T</sup>=CGMCC 1.3491<sup>T</sup>=CIP 105238<sup>T</sup>=JCM 9100<sup>T</sup>=NCIMB 13203<sup>T</sup>), was isolated from a saline soil sample. The DNA G+C content of this strain is 68.1 mol% (genome).

*Hrr. terrestre* 4p<sup>T</sup> (=VKM B-1739<sup>T</sup>=CIP 108318<sup>T</sup>=DSM 23453<sup>T</sup>=JCM 10247<sup>T</sup>), *Hrr. arcis* AJ201<sup>T</sup> (=CGMCC 1.5343<sup>T</sup>=DSM 23454<sup>T</sup>=JCM 13916<sup>T</sup>) and *Hrr. litoreum* Fa-1<sup>T</sup> (=CGMCC 1.5336<sup>T</sup>=DSM 23497<sup>T</sup>=JCM 13561<sup>T</sup>) are additional strains of *Halorubrum distributum*.

#### Funding information

This study was supported by the Spanish Ministry of Economy and Competitiveness (MINECO) through project CGL2017-83385-P, which included European (FEDER) funds, and Junta de Andalucía (Spain).

#### Acknowledgements

C. I.-D. and R. R. de la H. were recipients of a predoctoral fellowship and a short-stay grant 'Jose Castillejo-2018', respectively, from the Spanish Ministerio de Educación, Cultura y Deporte and Spanish Ministerio de Ciencia, Innovación y Universidades, respectively. We thank Lidia Rodrigo for her advice on the comparative genomic analysis of the strains.

#### Conflicts of interest

The authors declare that there are no conflicts of interest.

#### References

- McGenity TJ, Grant WD. Transfer of *Halobacterium saccharovororum*, *Halobacterium sodomense*, *Halobacterium trapanicum* NRC 34021 and *Halobacterium lacusprofundi* to the genus *Halorubrum* gen. nov., as *Halorubrum saccharovororum* comb. nov., *Halorubrum sodomense* comb. nov., *Halorubrum trapanicum* comb. nov., and *Halorubrum lacusprofundi* comb. nov. *Syst Appl Microbiol* 1995;18:237–243.
- Kamekura M, Dyall-Smith ML. Taxonomy of the family *Halobacteriaceae* and the description of two new genera *Halorubrobacterium* and *Natrialba*. *J Gen Appl Microbiol* 1995;41:333–350.
- Oren A, Ventosa A. A proposal for the transfer of *Halorubrobacterium distributum* and *Halorubrobacterium coriense* to the genus *Halorubrum* as *Halorubrum distributum* comb. nov. and *Halorubrum coriense* comb. nov., respectively. *Int J Syst Bacteriol* 1996;46:1180.
- Oren A, Arahal DR, Ventosa A. Emended descriptions of genera of the family *Halobacteriaceae*. *Int J Syst Evol Microbiol* 2009;59:637–642.
- Gupta RS, Naushad S, Baker S. Phylogenomic analyses and molecular signatures for the class *Halobacteria* and its two major clades: a proposal for division of the class *Halobacteria* into an emended order *Halobacteriales* and two new orders, *Haloferacales* ord. nov. and *Natrialbales* ord. nov., containing the novel families *Haloferacaceae* fam. nov. and *Natrialbaceae* fam. nov. *Int J Syst Evol Microbiol* 2015;65:1050–1069.
- Gupta RS, Naushad S, Fabros R, Adeolu M. A phylogenomic reappraisal of family-level divisions within the class *Halobacteria*: proposal to divide the order *Halobacteriales* into the families *Halobacteriaceae*, *Haloarculaceae* fam. nov., and *Halococcaceae* fam. nov., and the order *Haloferacales* into the families, *Haloferacaceae* and *Halorubraceae* fam. nov. *Antonie van Leeuwenhoek* 2016;109:565–587.
- Parte AC. LPSN - List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. *Int J Syst Evol Microbiol* 2018;68:1825–1829.
- Amoozegar MA, Siroosi M, Atashgahi S, Smidt H, Ventosa A. Systematics of haloarchaea and biotechnological potential of their hydrolytic enzymes. *Microbiology* 2017;163:623–645.
- Boucher Y, Douady CJ, Sharma AK, Kamekura M, Doolittle WF. Intragenomic heterogeneity and intergenomic recombination among haloarchaeal rRNA genes. *J Bacteriol* 2004;186:3980–3990.
- López-López A, Benlloch S, Bonfá M, Rodríguez-Valera F, Mira A. Intragenomic 16S rDNA divergence in *Haloarcula marismortui* is an adaptation to different temperatures. *J Mol Evol* 2007;65:687–696.
- Sun D-L, Jiang X, Wu QL, Zhou N-Y. Intra-genomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. *Appl Environ Microbiol* 2013;79:5962–5969.
- Papke RT, Corral P, Ram-Mohan N, de la Haba RR, Sánchez-Porro C et al. Horizontal gene transfer, dispersal and haloarchaeal speciation. *Life* 2015;5:1405–1426.
- Minegishi H, Kamekura M, Itoh T, Echigo A, Usami R et al. Further refinement of the phylogeny of the *Halobacteriaceae* based on the full-length RNA polymerase subunit B' (*rpoB'*) gene. *Int J Syst Evol Microbiol* 2010;60:2398–2408.
- Papke RT, White E, Reddy P, Kamekura M et al. A multilocus sequence analysis (MLSA) approach to *Halobacteriales* phylogeny and taxonomy. *Int J Syst Evol Microbiol* 2011;61:2984–2995.
- de la Haba RR, Corral P, Sánchez-Porro C, Infante-Domínguez C, Makkay AM et al. Genotypic and lipid analyses of strains from the archaeal genus *Halorubrum* reveal insights into their taxonomy, divergence, and population structure. *Front Microbiol* 2018;9:512.
- Corral P, de la Haba RR, Infante-Domínguez C, Sánchez-Porro C, Amoozegar MA et al. *Halorubrum chaoviator* Mancinelli et al. 2009 is a later, heterotypic synonym of *Halorubrum ezzemoulense* Kharroub et al. 2006. Emended description of *Halorubrum ezzemoulense* Kharroub et al. 2006. *Int J Syst Evol Microbiol* 2018;68:3657–3665.
- Zvyagintseva IS, Tarasov AL. Extreme halophilic bacteria from saline soils. *Microbiology* 1987;56:664–669.
- Ventosa A, Gutiérrez MC, Kamekura M, Zvyagintseva IS, Oren A. Taxonomic study of *Halorubrum distributum* and proposal of *Halorubrum terrestre* sp. nov. *Int J Syst Evol Microbiol* 2004;54:389–392.
- Xu X-W, Wu Y-H, Zhang H-B, Wu M. *Halorubrum arcis* sp. nov., an extremely halophilic archaeon isolated from a saline lake on the Qinghai-Tibet Plateau, China. *Int J Syst Evol Microbiol* 2007;57:1069–1072.
- Cui H-L, Lin Z-Y, Dong Y, Zhou P-J, Liu S-J et al. *Halorubrum litoreum* sp. nov., an extremely halophilic archaeon from a solar saltern. *Int J Syst Evol Microbiol* 2007;57:2204–2206.
- Parker CT, Tindall BJ, Garrity GM. International Code of Nomenclature of Prokaryotes. *Int J Syst Evol Microbiol* 2019;69:S1–S111.
- Ludwig W, Strunk O, Westram R, Richter L, Meier H et al. ARB: a software environment for sequence data. *Nucleic Acids Res* 2004;32:1363–1371.
- Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y et al. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.
- Kim M, Oh H-S, Park S-C, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–351.
- 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.
- Jukes TH, Cantor CR. Evolution of protein molecules. In: Munro HN (editor). *Mammalian Protein Metabolism*, 3. New York: Academic Press; 1969. pp. 21–132.
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010;59:307–321.
- Tavaré S. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures Math Life Sci* 1986;17:57–86.
- Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 1985;39:783–791.
- Rodríguez-R LM, Konstantinidis KT. The envomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints*;2016:e1900v1.

31. Edgar RC. Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2002;94:1792–1797.
32. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;5:e9490.
33. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 1992;8:275–282.
34. Shimodaira H, Hasegawa M. Multiple comparisons of Log-Likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 1999;16:1114–1116.
35. Yoon S-H, Ha S-min, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* 2017;110:1281–1286.
36. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P *et al.* DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007;57:81–91.
37. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 2009;106:19126–19131.
38. Chun J, Oren A, Ventosa A, Christensen H, Arahall 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.
39. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
40. Auch AF, von Jan M, Klenk H-P, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2010;2:117–134.
41. Blom J, Kreis J, Spänig S, Juhre T, Bertelli C *et al.* EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Res* 2016;44:W22–W28.
42. Meier-Kolthoff JP, Klenk H-P, 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.
43. Corral P, de la Haba RR, Sánchez-Porro C, Amoozegar MA, Papke RT *et al.* *Halorubrum persicum* sp. nov., an extremely halophilic archaeon isolated from sediment of a hypersaline lake. *Int J Syst Evol Microbiol* 2015;65:1770–1778.
44. Corral P, de la Haba RR, Sánchez-Porro C, Amoozegar MA, Thane Papke R *et al.* *Halorubrum halodurans* sp. nov., an extremely halophilic archaeon isolated from a hypersaline lake. *Int J Syst Evol Microbiol* 2016;66:435–444.
45. Angelini R, Corral P, Lopalco P, Ventosa A, Corcelli A. Novel ether lipid cardiolipins in archaeal membranes of extreme haloalkaliphiles. *Biochim Biophys Acta* 1818;2012:1365–1373.
46. Corcelli A, Lobasso S. Characterization of lipids of halophilic archaea. *Methods Microbiol* 2006;35:585–613.

### Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at [microbiologyresearch.org](http://microbiologyresearch.org).