Here the data on evolutionary relationships of persistent bacteria from water tanks and their close relatives are shown. Curated sequences of the hypervariable region of ribosomal ribonucleic acid (rRNA) obtained from a strain of *Aeromonas hydrophila* and two strains of *Serratia proteamaculans* after searches in the GenBank® data base were analyzed. The analysis which included 104 other bacteria strains, was carried out using molecular evolutionary genetic analysis (MEGA 7.0) software.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Value of the Data

- Data shows phylogeny of water tank bacteria and other species from several sources.
- Selection of strains for comparative whole-genome analysis can be facilitated by the data.
- Data is useful for further investigations of weak or strong biofilm producers during fouling of water tanks.

1. Data

The phylogeny data presented here have not been published in an initial study [1]. The 16S sequences from *A. hydrophila* (Fig. 1) and *S. proteamaculans* (Fig. 2) were compared with other isolates in order to gain more understanding of how they evolved. The clades formed after construction of phylogenetic trees show the evolutionary path of the sequences.

2. Experimental design, materials and methods

2.1. Molecular evolutionary genetics analysis (MEGA)

Sequences deposited in the Genbank® under accession numbers HG328351 (*A. hydrophila*), HG328350 and HG328352 (*S. proteamaculans*) from previous work [1] were analyzed.

Updated searches were carried out after which the top hits showing sequences from closely related culturable strains were selected for each genus and then subjected to phylogenetic analysis with MEGA software, version 7 [2]. After sequence alignment with ClustalW [3], the maximum likelihood statistical method based on the Tamura-Nei model [4] was used to generate phylogenetic trees. A total of 50 strains of closely related sequences from culturable strains were selected at random for each phylogenetic tree. Previously characterized 16S rRNA sequences that were used as outgroups included sequences from *Listeria monocytogenes* [5], *Pseudomonas aeruginosa* [6], *Pseudomonas fluorescens* [7] and a *Pseudomonas* species [8].
Fig. 1. The evolutionary history of *A. hydrophila* (Accession No. HG328351) was determined by using the maximum likelihood statistical method to compare with 50 other sequences from close relatives in MEGA7 software. Species of *Listeria monocytogenes* and three species of *Pseudomonas* were also investigated. The tree with the highest log likelihood was selected and the percentages of trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 55 nucleotide sequences. All positions containing gaps and missing data were eliminated.
Fig. 2. The evolutionary history of two strains of Serratia proteamaculans (HG328350 and HG3283502) was determined by using the maximum likelihood statistical method to compare with 50 other sequences from close relatives. The same outgroups and statistics shown in Fig. 1 were used. The analysis involved 56 nucleotide sequences.
Acknowledgement

The Microbiology Society United Kingdom supported the study with Grant no 2012/12/3. The sponsors had no input in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2017.10.073.

References

[1] O. Nwaiwu, M.I. Nwachukwu, Detection and molecular identification of persistent water vessel colonizing bacteria in a table water factory in Nigeria, Br. Microbiol. Res. J. 13 (2016) 1–12.
[2] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, Mol. Biol. Evol. 33 (2016) 1870–1874.
[3] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[4] K. Tamura, M. Nei M, Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees, Mol. Biol. Evol. 10 (1993) 512–526.
[5] O. Nwaiwu, Molecular serotype and evolutionary lineage of Listeria monocytogenes isolated from different Nigerian food items, Afr. J. Biotechnol. 15 (2016) 696–705.
[6] H.Y. Toschka, P. Hopfl, W. Ludwig, K.H. Schleifer, N. Ulbrich, V.A. Erdmann, Complete nucleotide sequence of a 16S ribosomal RNA gene from Pseudomonas aeruginosa, Nucleic Acids Res. 16 (1988) 2348.
[7] E. Hilario, T.R. Buckley, J.M. Young, Improved resolution on the phylogenetic relationships among Pseudomonas by the combined analysis of atp D, car A, rec A and 16S rDNA, Antonie Van Leeuwenhoek 86 (2004) 51–64.
[8] M. Eriksson, E. Sodersten, Z. Yu, G. Dalhammar, W.W. Mohn, Degradation of polycyclic aromatic hydrocarbons at low temperature under aerobic and nitrate-reducing conditions in enrichment cultures from northern soils, Appl. Environ. Microbiol. 69 (2003) 275–284.