

Draft genome sequence of a human-associated isolate of *Haloferax alexandrinus* strain Arc-hr, an extremely halophilic archaea

S. Khelaifia^{1,2}, A. Caputo¹, F. Djossou³ and D. Raoult^{1,2}

1) Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, CNRS (UMR 7278), IRD (198), INSERM (U1095), AMU (UM63), 2) Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, Aix-Marseille Université, Marseille, France and 3) Infectious and Tropical Diseases Department, Centre Hospitalier Andree-Rosemon, Cayenne, French Guiana

Abstract

We report the draft genome sequence of *Haloferax alexandrinus* strain Arc-hr (CSUR P798), isolated from the human gut of a 10-year-old Amazonian individual. Its 3 893 626 bp genome exhibits a 66.00% GC content. The genome of the strain Arc-hr contains 37 genes identified as ORFans, seven genes associated to halocin and 11 genes associated with polyketide synthases or nonribosomal peptide synthetases.

© 2016 The Author(s). Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

Keywords: Culturomics, genome sequence, genomics, *Haloferax alexandrinus*, halophilic archaea

Original Submission: 19 September 2016; **Revised Submission:** 26 October 2016; **Accepted:** 4 November 2016

Article published online: 18 November 2016

Corresponding author: S. Khelaifia, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, Aix-Marseille Université, 13385 Marseille Cedex 5, Marseille, France
E-mail: khelaifia_saber@yahoo.fr

Haloferax alexandrinus is an extremely halophilic archaea initially isolated from a solar saltern in Alexandria, Egypt. This Gram-negative halophilic archaea is highly pleomorphic, nonmotile and strictly aerobic, and it requires at least 10 g/L NaCl for growth [1]. Interestingly, DNA related to halophilic archaea from *Halorubrum* and *Halobacterium* genera was detected by PCR in colon mucous biopsy samples collected from patients with inflammatory bowel disease [2], but no isolate was made from any of these specimens. Recently we reported the first isolation of *H. alexandrinus* strain Arc-hr (Collection de Souches de l'Unité des Rickettsies (CSUR) P798) from a human stool specimen as part of a culturomics study of intestinal microflora of a 10-year-old Amazonian individual using media containing high salt concentrations [3]. *H. alexandrinus* strain Arc-hr is a Gram-negative archaeon, highly pleomorphic and strictly aerobic, with optimal growth observed at 37°C, pH 7.5 and 100 g/L NaCl. Sequencing the 16S ribosomal DNA confirmed the affiliation of the strain Arc-hr (GenBank accession no.

HG931927) to the genus *Haloferax* and yielded a 99.9% sequence similarity with the reference *H. alexandrinus* JCM 10717^T (GenBank accession no. NR_113438). Strain Arc-hr therefore represents the first *H. alexandrinus* isolate cultured from the human gut.

The complete genome of *H. alexandrinus* strain Arc-hr was sequenced on the MiSeq Technology (Illumina, San Diego, CA, USA) with the mate-pair strategy as previously described [4]. Open reading frames (ORFs) were predicted using Prodigal [5] with default parameters, but the predicted ORFs were excluded if they spanned a sequencing gap region (containing N) [6]. Noncoding genes and miscellaneous features were predicted using RNAmmer [7], ARAGORN [8], Rfam [9], PFAM [10] and Infernal [11]. Coding DNA sequences were predicted using Prodigal [5], and functional annotation was achieved using BLAST+ [6] and HMMER3 [12] against the UniProtKB database [13].

The genome the strain Arc-hr is 3 893 626 bp long with 66.00% GC content. It is composed of one scaffold (composed of five contigs); of the 3770 predicted genes, 3679 were protein-coding genes and 58 were RNAs (two genes are 5S rRNA, three genes are 16S rRNA, two genes are 23S rRNA and 51 genes are tRNA genes). A total of 2019 genes (65.38%) were assigned a putative function. The remaining genes were

annotated as hypothetical proteins (924 genes, 29.92%). Moreover, the *H. alexandrinus* strain Arc-hr genome contains 37 genes identified as ORFans, seven genes associated to halocin and 11 genes associated with polyketide synthases or nonribosomal peptide synthetases.

Deposit in a culture collection

Strain Arc-hr was deposited in CSUR (WDCM 875) under number CSUR P974.

Nucleotide sequence accession number

The *H. alexandrinus* strain Arc-hr genome sequence has been deposited in European Molecular Biology Laboratory under the accession number CCDK010000001:CCDK010000010.

Acknowledgements

This study was funded by the Fondation Méditerranée Infection. We thank M. Lardièrre for English-language review.

Conflict of Interest

None declared.

References

- [1] Asker D, Ohta Y. *Haloferax alexandrinus* sp. nov., an extremely halophilic canthaxanthin-producing archaeon from a solar saltern in Alexandria (Egypt). *Int J Syst Evol Microbiol* 2002;52:729–38.
- [2] Oxley AP, Lanfrancioni MP, Würdemann D, Ott S, Schreiber S, McGenety TJ, et al. Halophilic archaea in the human intestinal mucosa. *Environ Microbiol* 2010;12:2398–410.
- [3] Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol* 2016;7(1):16203.
- [4] Khelaifia S, Lagier JC, Bibi F, Azhar EI, Croce O, Padmanabhan R, et al. Microbial culturomics to map halophilic bacterium in human gut: genome sequence and description of *Oceanobacillus jeddahense* sp. nov. *Omics* 2016;20:248–58.
- [5] Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;8:119.
- [6] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics* 2009;15(10):421.
- [7] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007;35:3100–8.
- [8] Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 2004;32:11–6.
- [9] Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. Rfam: an RNA family database. *Nucleic Acids Res* 2003;31:439–41.
- [10] Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. *Nucleic Acids Res* 2012;40:D290–301.
- [11] Nawrocki EP, Kolbe DL, Eddy SR. Infernal 1.0: inference of RNA alignments. *Bioinformatics* 2009;25:1335–7.
- [12] Eddy SR. Accelerated profile HMM searches. *PLoS Comp Biol* 2011;7:e1002195.
- [13] UniProt Consortium. Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res* 2011;39:D214–9.