

**Halococcoides cellulosivorans gen. nov., sp. nov., an extremely halophilic cellulose-utilizing haloarchaeon from hypersaline lakes**

Sorokin, Dmitry Y.; Khijniak, Tatiana V.; Elcheninov, Alexander G.; Toshchakov, Stepan V.; Kostrikina, Nadezhda A.; Bale, Nicole J.; Sinninghe Damsté, Jaap S.; Kublanov, Ilya V.

**DOI**

[10.1099/ijsem.0.003312](https://doi.org/10.1099/ijsem.0.003312)

**Publication date**

2019

**Document Version**

Accepted author manuscript

**Published in**

International Journal of Systematic and Evolutionary Microbiology

**Citation (APA)**

Sorokin, D. Y., Khijniak, T. V., Elcheninov, A. G., Toshchakov, S. V., Kostrikina, N. A., Bale, N. J., Sinninghe Damsté, J. S., & Kublanov, I. V. (2019). Halococcoides cellulosivorans gen. nov., sp. nov., an extremely halophilic cellulose-utilizing haloarchaeon from hypersaline lakes. *International Journal of Systematic and Evolutionary Microbiology*, 69(5), 1327-1335. Article 003312. <https://doi.org/10.1099/ijsem.0.003312>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

**International Journal of Systematic and Evolutionary Microbiology**  
**Halococcoides cellulovorans gen. nov., sp. nov., an extremely halophilic cellulose-  
utilizing haloarchaeon from hypersaline lakes**  
--Manuscript Draft--

<b>Manuscript Number:</b>	IJSEM-D-18-00298R2
<b>Full Title:</b>	Halococcoides cellulovorans gen. nov., sp. nov., an extremely halophilic cellulose-utilizing haloarchaeon from hypersaline lakes
<b>Article Type:</b>	Taxonomic Description
<b>Section/Category:</b>	New taxa - Archaea
<b>Keywords:</b>	hypersaline lakes haloarchaea cellulose cellulotrophic Halorhabdus Haloarculaceae
<b>Corresponding Author:</b>	Dimitry Y Sorokin, Ph.D., Dr.Sci. Winogradsky Institute of Microbiology, Research Centre of Biotechnology RAS Moscow, NA RUSSIAN FEDERATION
<b>First Author:</b>	Dimitry Y Sorokin, Ph.D., Dr.Sci.
<b>Order of Authors:</b>	Dimitry Y Sorokin, Ph.D., Dr.Sci. Tatiana V Khijniak, PhD Nadezhda A Kostrikina, PhD Alexander G Elcheninov Stepan V Toshchakov, PhD Nicole J Bale, PhD Jaap S Sinninghe Damsté, PhD Ilya V Kublanov, PhD
<b>Manuscript Region of Origin:</b>	RUSSIAN FEDERATION
<b>Abstract:</b>	An extremely halophilic euryarchaeon, strain HArce11T, was enriched and isolated in pure culture from the surface brines and sediments of hypersaline athalassic lakes in the Kulunda Steppe (Altai region, Russia) using amorphous cellulose as the growth substrate. The colonies of HArce11T are pale-orange, and form large zones of cellulose hydrolysis around them. The cells are nonmotile cocci of variable size with a thin monolayer cell wall. The isolate is an obligate aerobic heterotroph capable of growth with only 3 substrates: various forms of insoluble cellulose, xylan and cellobiose. HArce11T is an extremely halophilic neutrophile, growing within the salinity range from 2.5 to 5 M NaCl (optimum at 3.5-4 M). The core archaeal lipids are dominated by C20-C20 and C25-C20 dialkyl glycerol ethers (DGE), in approximately 6:1 proportion. The 16S rRNA and rpoB' gene analysis indicated that HArce11T forms a separate lineage within the family Haloarculaceae, order Halobacteriales, with the genera Halorhabdus and Halopricus as closest relatives. On the basis of the unique phenotypic properties and distinct phylogeny of the 16S-rRNA and rpoB' genes, it is suggested that strain HArce11T is classified into a new genus and species Halococcoides cellulovorans gen. nov., sp. nov. (JCM 31941T=UNIQEM U975T).
<b>Author Comments:</b>	
<b>Opposed Reviewers:</b>	



Click here to access/download  
**Response to Reviewer**  
comments-answers.doc



2 ***Halococcoides cellulosivorans* gen. nov., sp. nov., an extremely halophilic**  
3 **cellulose-utilizing haloarchaeon from hypersaline lakes**

4  
5  
6 Dimitry Y. Sorokin<sup>a,b\*</sup>, Tatiana V. Khijniak<sup>a</sup>, Alexander G. Elcheninov<sup>a</sup>, Stepan V. Toshchakov<sup>a</sup>,  
7 Nadezhda A. Kostrikina<sup>a</sup>, Nicole J. Bale<sup>c</sup>, Jaap S. Sinninghe Damsté<sup>c,d</sup>, Ilya V. Kublanov<sup>a</sup>

8  
9 <sup>a</sup>*Winogradsky Institute of Microbiology, Research Centre of Biotechnology, Russian Academy of Sciences, Moscow,*  
10 *Russia*

11 <sup>b</sup>*Department of Biotechnology, TU Delft, The Netherlands*

12 <sup>c</sup>*Department of Marine Microbiology and Biogeochemistry, NIOZ Netherlands Institute for Sea Research, and Utrecht*  
13 *University, The Netherlands*

14 <sup>d</sup>*Department of Earth Sciences – Geochemistry, Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands*

15  
16 \*Author for correspondence:

17 D.Y. Sorokin; Tel: (7-495)1350109, Fax: (7-495)1356530; e-mail: soroc@inmi.ru; d.sorokin@tudelft.nl

18  
19  
20 Running title: *Halococcoides cellulosivorans* gen. nov., sp. nov.

21  
22  
23 The GenBank accession number of the whole genome sequences of strain HArce11<sup>T</sup> is CP028858

24  
25  
26  
27

28 **An extremely halophilic euryarchaeon, strain HArce11<sup>T</sup>, was enriched and isolated in pure**  
29 **culture from the surface brines and sediments of hypersaline athalassic lakes in the Kulunda**  
30 **Steppe (Altai region, Russia) using amorphous cellulose as the growth substrate. The colonies**  
31 **of HArce11<sup>T</sup> are pale-orange, and form large zones of cellulose hydrolysis around them. The**  
32 **cells are nonmotile cocci of variable size with a thin monolayer cell wall. The isolate is an**  
33 **obligate aerobic heterotroph capable of growth with only 3 substrates: various forms of**  
34 **insoluble cellulose, xylan and cellobiose. Strain HArce11<sup>T</sup> is an extremely halophilic**  
35 **neutrophile, growing within the salinity range from 2.5 to 5 M NaCl (optimum at 3.5-4 M).**  
36 **The core archaeal lipids are dominated by C<sub>20</sub>-C<sub>20</sub> and C<sub>25</sub>-C<sub>20</sub> dialkyl glycerol ethers (DGE),**  
37 **in approximately 6:1 proportion. The phylogenetic analysis based on 16S rRNA gene, *rpoB*'**  
38 **gene and the ribosomal proteins indicated that strain HArce11<sup>T</sup> forms a separate genus-level**  
39 **lineage within the family *Haloarculaceae*, order *Halobacteriales*, with the genera *Halorhabdus***  
40 **and *Halopricus* as closest relatives. This is also in line with the ANI and DDH values being far**  
41 **below the intragenus level. On the basis of the unique phenotypic properties and distinct**  
42 **phylogeny based on multiple conservative markers, it is suggested that strain HArce11<sup>T</sup> is**  
43 **classified into a new genus and species, *Halococcoides cellulovorans* gen. nov., sp. nov. (JCM**  
44 **31941<sup>T</sup>=UNIQEM U975<sup>T</sup>).**

45  
46

#### 47 Abbreviations

48 DGE, Dialkyl glycerol ether  
49 MGE, monalkyl glycerol ether  
50 PG, phosphatidyl glycerol  
51 PGS, phosphatidyl glycerol sulfate  
52 PGP-Me, Phosphatidylglycerophosphate methylester  
53 DG, diglycosyl diether  
54 TGD, triglycosyl diether

55

56 Extremely halophilic euryarchaea of the class *Halobacteria* form dense blooms in inland salt lakes  
57 and sea solar salterns with salt concentrations close to saturation. Most of the cultured species are  
58 aerobic heterotrophs, utilizing simple soluble organic monomers, such as sugars and organic acids,  
59 or complex rich amino acid-based substrates, such as various peptons and yeast extract [1-6].

60 The polymer mineralizing function at hypersaline conditions is usually attributed to  
61 halophilic bacteria [3-4]. There are only few published examples of the utilization of polymeric  
62 substances, such as starch, proteins or olive oil, as growth substrates among the haloarchaeal  
63 species [7-11]. In particular, nearly nothing is known about the ability of haloarchaea to hydrolyze  
64 and utilize insoluble recalcitrant polysaccharides, such as cellulose or chitin, for growth. The  
65 glycosidase genes encoding putative cellulases (GH family 3, 5 and 9) are present in many  
66 haloarchaeal genomes (*Haloarcula*, *Halobacterium*, *Halalkalicoccus*, *Haloferax*, *Halorhabdus*,  
67 *Halovivax*, *Halostagnicola*, *Haloterrigena-Natrinema* group, *Natronococcus*), while the presence of  
68 functional beta-1,4 endoglucanases has been, to date, demonstrated only in two genera of  
69 neutrophilic haloarchaea, i.e. *Haloarcula* and *Halorhabdus* [12-14]. However, it remains to be  
70 investigated whether these haloarchaea are actually capable of using native forms of cellulose as  
71 carbon and energy source.

72 So far, only two studies have focused on the functional aspect of cellulose degradation by  
73 haloarchaea [15-16]. In those works we were able, for the first time, to enrich and isolate in pure  
74 culture a number of haloarchaeal strains utilizing various forms of native insoluble cellulose as  
75 carbon and energy source both in neutral and alkaline saturated salt brines. The cellulotrophic  
76 natronoarchaea from hypersaline alkaline lakes included 2 subgroups: two strains with relative  
77 weak cellulase activity, belonging to a known species *Natronolimnobius baerhaense* (for which the  
78 capacity for cellulose hydrolysis had not previously been demonstrated) [15] and six strains with  
79 high cellulose-degrading capacity described recently as *Natronobiforma cellulositropha* gen. nov.,  
80 sp. nov. [16]. The group of neutrophilic cellulotrophic haloarchaeal isolated from various

81 hypersaline chloride-sulfate lakes, included *Halomicrobium* sp. strain HAre13, *Halosimplex* sp.  
82 strain HArce12 and a novel lineage, strain HArce11<sup>T</sup> [15]. In this paper we describe the phenotypic  
83 and phylogenetic properties of strain HArce11<sup>T</sup> and suggest its assignment into a novel genus and  
84 species *Halococcoides cellulosivorans*.

85  
86 Surface sediments and near-bottom brines from 3 hypersaline lakes in Kulunda Steppe (Altai  
87 region, Russia) with salt concentration of 280-350 g l<sup>-1</sup> and pH from 7.5-8.1 were used to enrich for  
88 cellulotrophic haloarchaea [15]. The brine-sediment slurries from three lakes were mixed,  
89 homogenized by vortexing and the resulting mix was briefly centrifuged at low speed to remove the  
90 coarse sediment fraction, while the remaining colloidal fraction was used as an inoculum.

91 The basic mineral medium used for the enrichment and cultivation of haloarchaea contained  
92 (in g l<sup>-1</sup>): 240 NaCl, 5 KCl, 0.25 NH<sub>4</sub>Cl and 3 of K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub>, pH 6.8. After sterilization, the  
93 base was supplemented with vitamin and trace metal mix [17], 1 mM MgSO<sub>4</sub>, 20 mg l<sup>-1</sup> yeast  
94 extract and 10 mM filter-sterilized NaHCO<sub>3</sub>. Various forms of insoluble cellulose obtained from  
95 Sigma or synthesized as described previously (amorphous cellulose, [15]) were used as the only  
96 carbon and energy source at a final concentration of 1 g l<sup>-1</sup>. For the enrichment, 1 ml of colloidal  
97 sediments was used to inoculate 20 ml medium containing 1 g l<sup>-1</sup> of amorphous cellulose in 100 ml  
98 closed serum bottles placed on a rotary shaker at 37°C and at 120 rpm. The development of cells  
99 was monitored by the visual extent of cellulose degradation, the appearance of pink-orange color  
100 and by microscopy. After visible cellulose degradation and cell growth (30-40 days), the culture  
101 was serially diluted in the same medium and the maximal positive dilutions were plated onto a solid  
102 medium prepared by mixing the liquid medium (with additional solid NaCl addition to compensate  
103 for dilution with agar) and 5% extensively washed agar 3:2 at 55°C. The plates were incubated at  
104 37°C in closed plastic bags for 40-60 days. The appearance of colored colonies with large clearance  
105 zones was used as an indicator of growth of cellulolytic haloarchaea. It needs to be stressed here,

106 that such colonies were never dominating on the plates, even obtained from final positive serial  
107 dilutions, indicating a presence of high proportion of satellites probably feeding on the cellulose  
108 hydrolysis products. The cellulolytic colonies (**Fig. 1a**) were transferred to the liquid medium with  
109 amorphous cellulose and the positive cultures were further purified by several rounds of plating-  
110 liquid culture cultivation with amorphous cellulose. This yielded 3 pure cultures of cellulotrophic  
111 haloarchaea with identical 16S-rRNA gene sequence, of which strain HArce11<sup>T</sup> was chosen for  
112 further characterization.

113  
114 The phase contrast microscopy was done using the Zeiss Axioplan Imaging 2 microscope  
115 (Göttingen, Germany). For the electron microscopy of thin sections, the cells of strain HArce11<sup>T</sup>  
116 grown with amorphous cellulose were fixed in 1% (w/v) OsO<sub>4</sub> containing 3.0 M NaCl for 1 week at  
117 4°C, washed and resuspended in 3 M NaCl, stained overnight with 1% (w/v) uranyl acetate,  
118 dehydrated in ethanol series, and embedded in Epon resin. After thin sectioning, the preparations  
119 were post-stained with 1% (w/v) lead citrate and examined using the JEOL-100 model of TEM  
120 (Japan),

121 Cells of HArce11<sup>T</sup> were non-motile cocci of variable size from 0.8 to 3 µm (**Fig. 1b**). During  
122 the first stage of growth on insoluble celluloses most of the cells aggregated with cellulose  
123 particles/fibres (**Fig. 1c**), while free cells appeared only after massive cellulose hydrolysis. Electron  
124 microscopy revealed the presence of a large nucleoid and a thin, single layer cell wall, typical for  
125 many haloarchaeal species (**Fig. 1d**). The cells lysed after resuspension in solutions containing less  
126 than 10% NaCl.

127  
128 Genomic DNA was isolated by ISOLATE II Genomic DNA Kit (Bioline Reagents, UK) according  
129 to manufacturer's instructions. Fragment genomic libraries were prepared from 1 µg of genomic  
130 DNA with NEBNext Ultra DNA library preparation kit (New England Biolabs, Ipswich, MA, USA)



131 according to manufacturer's instructions to obtain mean library size of 600 - 700 bp. The library  
132 was sequenced with MiSeq™ Illumina Inc. (Illumina Inc., San Diego, CA, USA) using paired-end  
133 250-bp reads. After sequencing all reads were subjected to stringent quality filtering and trimming  
134 with CLC Genomics Workbench 10.0 (Qiagen, Germany). Sequencing adapters were trimmed with  
135 SeqPrep tool (<https://github.com/jstjohn/SeqPrep>). Finally, 925,497 read pairs were used for *de*  
136 *novo* assembly. Reads were assembled with SPADES 3.10.0 [18]. Initial assembly consisted of 166  
137 scaffolds of total length 2,793,855 nt and N50 of 2,525,738 nt. In parallel, reads were assembled  
138 with MIRA 4.0.2 genome assembler [19], resulting in assembly of total length 2,726,789 nt and  
139 N50 43612 nt. After manual curation and comparison of two assemblies using CLC Genomics  
140 Workbench 10.0 software (Qiagen, Germany) circular ungapped chromosome of strain HArce11<sup>T</sup>  
141 was obtained. Total length of the strain HArce11<sup>T</sup> chromosome is 2,723,120 bp, GC-content is  
142 65.74%. Validation of an assembly was performed by analysis of mapping of all obtained reads  
143 back to chromosome sequence performed with CLC Genomics Workbench (Qiagen, Germany).  
144 99.76% of reads were mapped resulting in final genome coverage of  $88.3 \pm 22.6$  x. Additionally,  
145 integrity of the assembly was checked by the analysis of unaligned read ends with InDel analysis  
146 tool of CLC Genomics Workbench (Qiagen, Germany). No regions, significantly enriched by  
147 partially aligned reads were found. Due to these results our genomic assembly can be considered as  
148 finalized complete genome sequence. Annotation with IMG/ER server pipeline [20] resulted in  
149 prediction of 2,641 protein-coding genes, 60 tRNA genes and one complete rRNA operon.  
150 Genomic assembly and related metadata have been deposited in NCBI database under accession  
151 numbers XCP028858, PRJNA449302, SAMN08826612 for the genomic assembly, Bioproject and  
152 Biosample, respectively.

153 16S rRNA and *rpoB'* gene sequences were obtained from the draft genome assemblies of  
154 strain HArce11<sup>T</sup>. The phylogenetic analysis was performed in Mega 7 package [21]. The 16S rRNA  
155 gene sequences of all species of the *Halobacteriales* order with validly described names obtained

156 from the Genbank were aligned together with the complete sequence of strain HArce11<sup>T</sup> using G-  
157 INS-i method in MAFFT server v7 [22]. The phylogenetic analysis was performed using Maximum  
158 Likelihood algorithm and the General Time Reversible (GTR) model (*G+I*, 4 categories) [23]. The  
159 *rpoB*'-based phylogenetic analysis, was performed the same way as for 16S rRNA gene. For  
160 ribosomal proteins phylogenetic analysis of 17 single-copy conserved ribosomal protein sequences  
161 (S2, S3, S11, S12, S17, S19, L3, L4, L5, L10, L11, L13, L14, L15, L23, L24, L29) were obtained  
162 from 39 available in IMG/M-ER [20] genomes of *Halobacteriales* representatives with  
163 *Natronomonas* as an outgroup. The protein sequences were aligned in MAFFT v7 [22] using L-  
164 INS-i algorithm and then concatenated using FaBox joiner alignment [24]. Phylogenetic tree based  
165 on concatenated alignment of the proteins was constructed using Maximum Likelihood method and  
166 the LG model (*G + I*, 4 categories) [25].

167 BLAST of strain HArce11<sup>T</sup> 16S rRNA gene against nucleotide sequences from cultured  
168 haloarchaeal species revealed *Halorhabdus* species and *Halapricum salinum* being the closest  
169 relatives with 94.0-92.9 and 92.5 % sequence identity, respectively. This level of relation indicates  
170 a separate genus status. Further phylogenetic analysis based of the 16S rRNA gene comparison  
171 demonstrated that strain HArce11<sup>T</sup> forms a separate lineage within the family *Haloarcelaceae* [26]  
172 with the genera *Halorhabdus* and *Halapricus* as the closest relatives (**Fig. 2 a**). Since the  
173 divergence point of “strain HArce11-*Halorhabdus*” and *Halapricum* clusters was not supported by  
174 bootstrap test, the additional markers (*rpoB* gene and ribosomal proteins) were used to infer  
175 phylogenetic position of strain HArce11<sup>T</sup> (**Fig 2 b, c**). The results support a separation of strain  
176 HArce11<sup>T</sup>, *Halorhabdus* and *Halapricum* in a distinct cluster, whereby strain HArce11<sup>T</sup> forms a  
177 longest branch suggesting its novel genus level.

178 Pairwise ANI comparison was performed using IMG built-in tool [27]. The calculated ANI  
179 values were 74.1 % between strain HArce11<sup>T</sup> and *Halapricum salinum*; 74.8 % between strain  
180 HArce11<sup>T</sup> and *Halorhabdus utahensis*; 75.1 % between strain HArce11<sup>T</sup> and *Halorhabdus tiamatea*

181 (Table 1). For digital DDH we used the Genome-to-Genome Distance Calculator 2.1 (GGDC) [28].  
182 BLAST+ was selected as local alignment tool and three formula were used: 1 – length of all HSPs  
183 divided by total genome length, 2 – sum of all identities found in HSPs divided by overall HSP  
184 length (recommended) and 3 - sum of all identities found in HSPs divided by total genome length  
185 The average *in silico* DDH values calculated from the 3 formulas between strain HArce11<sup>T</sup> and  
186 *Halapricum salinum*, *Halorhabdus utahensis* and *Halorhabdus tiamatea* were 15.7, 16.4 and 16.6  
187 %, respectively (Table 1). Thus the calculated values of both ANI and DDH were significantly  
188 below the recognized species separation (96% and 70%, respectively), [29].

189 Taken together, the phylogenetic analysis and genome-based comparison demonstrated a  
190 separate genus-level status of strain HArce11<sup>T</sup> within the *Haloarculaceae* family.

191  
192 The core membrane lipids were obtained by acid hydrolysis (5% HCl in methanol by reflux  
193 for 3 h) of the freeze-dried cells and subsequent analysis by HPLC-MS for GDGTs and archaeol  
194 derivatives according to [30]. Intact polar lipids were obtained by Bligh Dyer extraction of freeze-  
195 dried cells and subsequent HPLC-MS analysis as described in [31].

196 The core membrane lipids were dominated by archaeol [C<sub>20</sub>-C<sub>20</sub> dialkyl glycerol ether  
197 (DGE), 81% of the total] with lesser amounts of extended archaeol (C<sub>20</sub>-C<sub>25</sub> DGE, 13% of the total).  
198 Traces of the monoglycerol ether (MGE) lipids (1-C<sub>20</sub> MGE, 2-C<sub>20</sub> MGE, and 2-C<sub>25</sub> MGE) were  
199 also detected. The intact polar lipid profile (identified using multistage mass spectrometry) was  
200 quite complex, including (in order of abundance) phosphatidylglycerophosphate methylester (PGP-  
201 Me), phosphatidylglycerol (PG), a sulfophospholipid with an unknown sulfur-containing  
202 headgroup, a diglycosyl (2GL), phosphatidylglycerophosphate (PGP) and  
203 phosphatidylglycerosulfate (PGS) (Supplementary Fig. S1). When compared with the two closest  
204 phylogenetic neighbours (Table 2), only first two most abundant lipids were present in all 3  
205 species: phosphatidylglycerophosphate methylester (PGP-Me) and phosphatidylglycerol (PG).  
206 These phospholipids are most common in the members of *Halobacteria* and, in particular, the

207 domination of the PGP-Me is considered to be related to extreme salt tolerance [32]. The less  
208 abundant lipids in strain HArce11<sup>T</sup> included a glycolipid phosphatidyldiglycoside (2GL) and 2  
209 sulfolipids. Lipids belonging to the glycolipid and sulfolipid classes are also present in the two  
210 closest relatives of HArce11<sup>T</sup>. For example, the closest relative, *Halorhabdus tiamatea*, contains a  
211 three glycosyl (3GL) glycolipid and a monosulfated diglycosyl diether (S1-DGD) sulfolipid. It is  
212 probable that the structurally homologous different glyco- and sulfolipids play a similar function in  
213 maintaining membrane homeostasis at extreme salinity [33-34] (Kates 1992; Oger 2013).  
214 Sulfolipids are also commonly found in neutrophilic haloarchaea, and in particular in the members  
215 of the family *Haloarculaceae* [26].

216  
217 Strain HArce11<sup>T</sup> is an obligately aerobic saccharolytic haloarchaeon. Anaerobic growth with  
218 cellobiose as substrate was tested in 10 ml liquid cultures placed into 23 ml serum bottles, closed  
219 with butyl rubber stoppers and made anoxic by sterile evacuation-flushing with argon. The results  
220 were negative either for fermentation, or with elemental sulfur, thiosulfate, DMSO, TMA and  
221 nitrate as *e*-acceptors. During aerobic growth, strain HArce11<sup>T</sup> utilized only three substrates as  
222 their carbon and energy source: insoluble celluloses with different degree of crystallinity, including  
223 an amorphous form, Sigma celluloses, filter paper; xylan (from birch wood) and cellobiose. Weak  
224 and irregular growth was noticed with lichenan (beta-1,4/-1,3 glycan). No growth was detected with  
225 the following polysaccharides: CMC, beta 1,3/1,6 and alpha glucans, beta-mannan, beta-galactan,  
226 chitin, chitosan, pectin; heteropolysaccharides, such as beta gluco- and galacto- mannans, alginate.  
227 The soluble sugar compounds tested negative included glucose, fructose, galactose, mannose,  
228 arabinose, rhamnose, N-acetylglucosamine, glucosamine, glucuronic and galacturonic acids,  
229 maltose, lactose, trehalose, melibiose, melizitose, xylose, ribose, sorbitol, mannitol and glycerol.  
230 Likewise, no growth was observed with organic acids (C<sub>2</sub>-C<sub>10</sub> fatty acids, lactate, pyruvate, malate,  
231 succinate, fumarate) and complex organic amino acid substrates, such as various peptons and yeast

232 extract. The extremely narrow specialization on cellulose polymers of the neutrophilic haloarchaeon  
233 HArce11<sup>T</sup> is only a second example among known species of haloarchaeae, resembling its recently  
234 described alkaliphilic counterpart *Natronobiforma cellulositropha* found in various hypersaline  
235 soda lakes [16].

236 Recommended enzymatic activity tests [35] included plate assays for amylase (soluble  
237 starch), protease (casein, gelatin), esterase (tributyrin) and lipase (emulsified olive oil) using a low  
238 background of cellobiose (1 mM). Amylase activity was detected by flooding the plate with Lugol  
239 solution, for protease activity the plate was flooded with 10% TCA to denature undegraded protein,  
240 while esterase and lipase activities are evident from the visual clearance of turbid background  
241 around the colonies. All of these activities were negative. Strain HArce11<sup>T</sup> was strongly catalase  
242 positive (colony test with 3% H<sub>2</sub>O<sub>2</sub>), but only weak-positive in the oxidase activity (colony test with  
243 1% tetramethylphenyldiamine hydrochloride on filter paper). Sulfide formation from thiosulfate or  
244 sulfur during aerobic growth with cellobiose (lead acetate paper test) and indole formation from  
245 tryptophan (Kovac's reagent test, [36]) were all negative. While growing with cellobiose, strain  
246 HArce11<sup>T</sup> used only ammonium salts as the N-source (urea, nitrate, nitrite were negative).

247 The salt profile for growth in strain HArce11<sup>T</sup> culture was investigated using cellobiose as  
248 the substrate in medium buffered at pH 7 with potassium phosphate buff in liquid culture incubated  
249 at 37°C. Growth was observed within NaCl range from 2.5 to 5 M with an optimum at 3.5-4 M. The  
250 pH for growth with cellobiose at 4 M NaCl was investigated within the range from 5 to 9 using a  
251 combination of HEPES (4 g l<sup>-1</sup>) and potassium phosphates (5 g l<sup>-1</sup> in total) as buffers for the pH  
252 range from 5 to 8 and a combination of potassium phosphates and 0.5 M Na<sub>2</sub>CO<sub>3</sub> for the pH 8.5-9.  
253 The pH during growth was also maintained either by adding CO<sub>2</sub> into the gas phase (to decrease the  
254 actual pH) or 1 M filter-sterilized NaHCO<sub>3</sub> (to increase the actual pH). Strain HArce11<sup>T</sup> was able to  
255 grow within the pH range of 6.5-8.0 with an optimum at 7.0-7.2. Based on the data, the isolate can  
256 be classified as an extremely halophilic neutrophile. At pH 7 and 4 M NaCl, the strain grew equally

257 well at Mg concentrations from 1 to 20 mM, thus belonging to a low Mg-requiring type. The  
258 temperature profiling during growth on cellobiose at pH 7 and 4 M NaCl was done starting from 20  
259 and up to 60°C with an increment of 5°C. The growth was possible from 25 to 50°C with an  
260 optimum between 40 and 45°C.

261 Antibiotic resistance of strain HArce11<sup>T</sup> was tested at optimal growth conditions in liquid  
262 culture using cellobiose as substrate. The following antibiotics (100 mg l<sup>-1</sup>) did not inhibit growth:  
263 penicillin G, ampicillin, kanamycin, streptomycin, erythromycin, gentamicine and vancomycin. No  
264 growth was observed in presence of chloramphenicol and rifampicin at concentrations above 50 and  
265 30 mg l<sup>-1</sup>, respectively.

266 A phenotypic comparison of strain HArce11<sup>T</sup> with the closest haloarchaeal relatives from  
267 *Haloarcelaceae* is shown in **Table 2**. Interestingly, the closest relatives of HArce11<sup>T</sup>, the  
268 *Halorhabdus* species, are apparent polysaccharide degraders, according to the presence of multiple  
269 GH genes in the genome and activity tests in *H. tiamatea* [14, 37] and the proven ability of *H.*  
270 *utahensis* to grow with xylan [38]). Our tests with the type strain of *H. tiamatea* JCM 14471<sup>T</sup> and  
271 also with our own isolates closely related to this species demonstrated that these haloarchaea are,  
272 indeed, potent polysaccharide degraders capable of growth with a range of glycans as sole source of  
273 carbon and energy (**Table 2**). Especially interesting is the ability (albeit weak with never a complete  
274 utilization) of *H. tiamatea* to grow with beta-1,4 mannan. So far, only two such cases have been  
275 found among the extremely halophilic euryarchaea - in *Natronoarchaeum mannanilyticum* and  
276 recently described cellulose-utilizing *Natronobiforma cellulositropha* [16, 39]. However, the major  
277 difference between the *Halorhabdus* species and strain HArce11<sup>T</sup> is the ability of the latter to use  
278 cellulose as growth substrate : none of the tested forms of insoluble celluloses with different degree  
279 of crystallinity, including amorphous, four types of Sigma celluloses, filter paper and Avicell,  
280 supported growth of *H. tiamatea*. On the other hand, tests on CMC plates showed a presence of  
281 beta-1,4 endoglucanase activity in colonies of *H. tiamatea*. This is another demonstration, that what

282 is often claimed on the basis of test with soluble artificial analogue of cellulose (CMC) as the ability  
283 to grow with cellulose should not be considered as valid. Since the genome of another closest  
284 relative of strain HArce11<sup>T</sup>, *Halapricum salinum* [40], completely lacks genes encoding the GH-  
285 family glycosidases, it might be concluded, that it differs significantly in its key physiological  
286 specialization, most probably being an ordinary saccharolytic utilizing products of polymer  
287 hydrolysis. Taking into account that three other members of the family *Haloarculaceae* - the genera  
288 *Haloarcula*, *Halomicrobium* and *Halosimplex* do have species with confirmed ability to degrade  
289 glycans, including cellulose [12-13, 15] and chitin (*Halomicrobium*) [15], it might be speculated  
290 that such potential has already been acquired in the common ancestor of this radiation of  
291 *Halobacteria* but lost later on in some members, such as *Halapricum*, and proliferated in the others,  
292 of which strain HArce11<sup>T</sup> seems to be the most narrowly specialized. Further phylogenomic  
293 reconstructions might be able to substantiate this interesting question.

294  
295 In conclusion, strain HArce11<sup>T</sup> is the first example of an extremely halophilic euryarchaeon directly  
296 enriched and isolated from hypersaline lakes using insoluble celluloses as the growth substrate.  
297 Taking into account its unique phenotypic properties and distant phylogenetic position, as inferred  
298 from the robust phylogenetic reconstruction based on 19 conservative markers, and ANI and *in*  
299 *silico* DDH values far below the recognized intragenus levels, we propose to classify strain  
300 HArce11<sup>T</sup> in a novel genus and species *Halococcoides cellulovorans*.

301  
302

### 303 **Description of *Halococcoides* gen. nov.**

304 Ha.lo.coc.co'i.des. [Gr. n. *hals*, halos salt of the sea; N.L. masc. n. *coccus* (from Gr. masc. n.  
305 *kokkos*, grain, seed), coccus; L. suff. *-oides* (from Gr. suff. *-eides*, from Gr. n. *eidos*, that which is  
306 seen, form, shape, figure), resembling, similar; L. suff. *-oides*, resembling, similar; N.L. neutral. n.  
307 *Halococcoides*, coccus-shaped holophile].

308 Extremely halophilic euryarchaeon, a member of the family *Haloarculacea*, order *Halobacteriales*,  
309 class *Halobacteria*, found in hypersaline athalassic lakes. Specialized in utilization of cellulose as  
310 growth substrate. The type species is *Halococcoides cellulosivorans*. The recommended three-letter  
311 abbreviation for this genus is Hcd.

312

313 **Description of *Halococcoides cellulosivorans* sp. nov.**

314 *Halococcoides cellulosivorans* (cel.lu.lo.si.vo'rans N.L. neutral n. *cellulosum*, cellulose; L. pres.  
315 part. *vorans*, devouring; N.L. part. adj. *cellulosivorans*, cellulose devouring)

316

317 Cells are non-motile cocci, 0.8-3  $\mu\text{m}$ , with a thin monolayer cell wall. The colonies on amorphous  
318 cellulose agar are flat, up to 1 mm, soft and slightly orange. It is a strictly aerobic (catalase/oxidase  
319 positive) saccharolytic specialized on utilization of native forms of insoluble cellulose and xylan.  
320 Cellobiose is the only soluble sugar utilized for growth. The nitrogen source is ammonium. Nitrate  
321 and urea are not utilized. Does not grow anaerobically either by fermentation or anaerobic  
322 respiration. Does not utilize organic acids or organic nitrogen compounds as carbon and energy  
323 source. High Mg is not required for growth. Proteolytic and lipolytic activity are absent. Strain  
324 HArCell<sup>T</sup> is an extremely halophilic neutrophile, with the NaCl range for growth between 3 and 5  
325 M (optimum at 3.5-4 M) and the pH range from 6.5 to 8.0 (optimum at pH 7.0-7.2). The maximum  
326 growth temperature at 4 M NaCl with cellobiose as substrate is 50°C (optimum at 40-42°C). The  
327 core membrane lipids are dominated by C<sub>20</sub>-C<sub>20</sub> and C<sub>25</sub>-C<sub>20</sub> DGE with 1-C<sub>25</sub> MGE and 2-C<sub>20</sub> MGE  
328 as minor components. The identified intact membrane polar lipids include  
329 phosphatidylglycerophosphate methylester (PGP-Me) and phosphatidylglycerol (PG) as dominant  
330 and diglycosyl diether glycolipid (2GL) and phosphatidylglycerol sulfate (PGS) sulfolipid as minor  
331 components. The G + C content of the genomic DNA in the type strain is 65.74 mol% (genome).



332 The habitat is hypersaline lakes with near-neutral pH. The type strain (HArce11<sup>T</sup>=JCM  
333 31939<sup>T</sup>=UNIQEM U972<sup>T</sup>). The full genome accession number in the GenBank is CP028858.

334

#### 335 **Funding information**

336 This work was supported by the Russian Science Foundation (grant 16-14-00121). JSD and NB  
337 received funding from the European Research Council (ERC) under the European Union's Horizon  
338 2020 research and innovation programme (grant agreement No 694569).

339

#### 340 **Conflict of interest:**

341 The authors declare that there is no conflict of interests.

342

343

#### 344 **References**

- 345 1. Ventosa A, Oren A, Ma Y. (Editors). *Halophiles and hypersaline environments. Current*  
346 *research and future trends*. Springer: Heidelberg, Dordrecht, London, New York; 2011.  
347 387 pp.
- 348 2. Horikoshi K. (Editor). Halophiles, Part III. In: *Extremophiles Handbook*, Springer: Tokyo;  
349 2011. vol. 1: 255-402.
- 350 3. **Andrei AS, Banciu HL, Oren A.** Living with salt: metabolic and phylogenetic diversity of  
351 archaea inhabiting saline ecosystems. *FEMS Microbiol Lett* 2012; 330: 1-9.
- 352 4. **Oren A.** Halophilic microbial communities and their environments. *Curr Opin Biotechnol*  
353 (2015); 33: 119-124.
- 354 5. **Oren A.** Life at high salt concentrations. In: *The Prokaryotes. Ecophysiology and*  
355 *Biochemistry*, 4<sup>th</sup> Edn. Rosenberg E *et al.* (Editors) Springer: New York; 2013. pp. 429-440..
- 356 6. **Grant BD, Jones BE.** Bacteria, archaea and viruses of soda lakes. In: *Soda lakes of East*  
357 *Africa*, Schagerl M (Editor) Springer: Switzerland; 2016. pp. 97-147.

- 358 7. **Bhatnagar T, Boutaiba S, Hacene H, Cayol J-L, Fardeau M-L et al.** Lipolytic activity  
359 from *Halobacteria*: Screening and hydrolase production. *FEMS Microbiol Lett* 2005; 248:  
360 133–140.
- 361 8. **Enache M, Kamekura M.** Hydrolytic enzymes of halophilic microorganisms and their  
362 economic values. *Rom J Biochem* 2010; 47: 47–59.
- 363 9. **Moshfegh M, Shahverdi AR, Zarrini G, Faramarzi MA.** Biochemical characterization of  
364 an extracellular polyextremophilic  $\alpha$ -amylase from the halophilic archaeon *Halorubrum*  
365 *xinjiangense*. *Extremophiles* 2013; 17: 677–687.
- 366 10. **Selim S, Hagagy N, Aziz MA, El-Meleigy E-S, Pessione E.** Thermostable alkaline  
367 halophilic-protease production by *Natronolimnobius innermongolicus* WN18. *Nat Prod Res*  
368 2014; 28: 1476–1479.
- 369 11. **Amoozegar MA, Siroosi M, Atashgahi S, Smidt H, Ventosa A.** Systematics of  
370 haloarchaea and biotechnological potential of their hydrolytic enzymes. *Microbiology* 2017;  
371 163: 623-645.
- 372 12. **Li TX, Yu H-Y.** Halostable cellulase with organic solvent tolerance from *Haloarcula* sp.  
373 LLSG7 and its application in bioethanol fermentation using agricultural wastes. *J Ind*  
374 *Microbiol Biotechnol* 2013; 13: 1357-1365.
- 375 13. **Li X, Yu H-Y.** Characterization of a halostable endoglucanase with organic solvent-tolerant  
376 property from *Haloarcula* sp. G10. *Int J Biol Macromol* 2013; 62: 101-106.
- 377 14. **Werner J, Ferrer M, Michel G, Mann AJ, Huang S, et al.** *Halorhabdus tiamatea*:  
378 proteogenomics and glycosidase activity measurements identify the first cultivated  
379 euryarchaeon from a deep-sea anoxic brine lake as potential polysaccharide degrader.  
380 *Environ Microbiol* 2014; 16: 2525-2537.

- 381 15. **Sorokin DY, Toschakov SV, Kolganova TV, Kublanov IV.** Halo(natrono)archae isolated  
382 from hypersaline lakes utilize cellulose and chitin as growth substrates. *Front Microbiol*  
383 2015; 6: article 942.
- 384 16. **Sorokin DY, Khijniak TV, Kostrikina NA, Elcheninov AG, Toshchakov SV et al.**  
385 *Natronobiforma cellulositropha* gen. nov., sp. nov., a novel haloalkaliphilic member of the  
386 family *Natrialbaceae* (class *Halobacteria*) from hypersaline alkaline lakes. *Syst Appl*  
387 *Microbiol* 2018; 41: 355-362.
- 388 17. **Pfennig N, Lippert KD.** Über das Vitamin B12-Bedürfnis phototropher Schwefelbakterien.  
389 *Arch Mikrobiol* 1966; 55: 245-256.
- 390 18. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M et al.** SPAdes: A new  
391 genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*  
392 2012; 19: 455–477.
- 393 19. **Chevreur B, Wetter T, Suhai S.** Genome sequence assembly using trace signals and  
394 additional sequence information. *Computer science and biology: Proceedings of the German*  
395 *Conference on Bioinformatics (GCB)*. 1999; 99: 45-56.
- 396 20. **Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E et al.** IMG 4 version of the  
397 integrated microbial genomes comparative analysis system. *Nucleic Acids Res* 2014;  
398 42(Database issue): D560-567.
- 399 21. **Kumar S, Stecher G, Tamura K.** MEGA7: Molecular evolutionary genetics analysis  
400 version 7.0 for bigger datasets. *Mol Biol Evol.* 2016; 33: 1870-1874.
- 401 22. **Katoh K, Rozewicki J, Yamada KD.** MAFFT online service: multiple sequence alignment,  
402 interactive sequence choice and visualization. *Brief Bioinform* 2017; bbx108.
- 403 23. **Nei M, Kumar S.** *Molecular evolution and phylogenetics.* Oxford University Press : New  
404 York; 2000. 333 pp.

- 405 24. **Villesen P.** FaBox: An online toolbox for FASTA sequences. *Mol Ecol Notes* 2007; 7: 965–  
406 968.
- 407 25. **Le SQ, Gascuel O.** An improved general amino acid replacement matrix. *Mol Biol Evol*  
408 2008; 25: 1307–1320.
- 409 26. **Oren A.** *Haloarculaceae*. In: *Bergey's Manual of Systematics of Archaea and Bacteria*.  
410 2017. John Wiley & Sons, Inc. DOI: 10.1002/9781118960608.fbm00293.
- 411 27. **Chen IMA, Markowitz VM, Chu K, Palaniappan K, Szeto E, et al.** IMG/M: Integrated  
412 genome and metagenome comparative data analysis system. *Nucleic Acids Res* 2017; 45:  
413 D507–D516.
- 414 28. **Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M.** Genome sequence-based species  
415 delimitation with confidence intervals and improved distance functions. *BMC*  
416 *Bioinformatics* 2013; 14: 60.
- 417 29. **Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, et al.** Proposed minimal  
418 standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol*  
419 *Microbiol* 2018; 68: 461–466.
- 420 30. **Weijers JWH, Panoto E, van Bleijswijk J, Schouten S, Balk M.** Constraints on the  
421 biological source(s) of the orphan branched tetraether membrane lipids. *Geomicrobiol J*  
422 2009; 26: 402-414.
- 423 31. **Sinninghe Damsté JS, Rijpstra WIC, Hopmans EC, Jung MY, Kim JG.** Intact polar and  
424 core glycerol dibiphytanyl glycerol tetraether lipids of group I. 1a and I. 1b *Thaumarchaeota*  
425 in soil. *Appl Environ Microbiol* 2012; 78: 6866-6874.
- 426 32. **Tenchov B, Vescio EM, Sprott GD, Zeidel ML, Mathai JC.** Salt tolerance of archaeal  
427 extremely halophilic lipid membranes. *J Biol Chem* 2006; 281: 10016–10023.

- 428 33. **Kates M.** Archaeobacterial lipids: structure, biosynthesis and function. In: *The*  
429 *Archaeobacteria: Biochemistry and Biotechnology*, Danson MJ, Hough DW, Lunt GG (Eds).  
430 Portland Press, London; 1992. pp 51–72.
- 431 34. **Oger PM, Cario A.** Adaptation of the membrane in Archaea. *Biophys Chem* 2013; 183: 42–  
432 56.
- 433 35. **Oren A, Ventosa A, Grant WD.** Proposed minimal standards for description of new taxa in  
434 the order *Halobacteriales*. *Int J Syst Bacteriol* 1997; 47: 233–238.
- 435 36.  **Holding AJ, Collee JG.** Routine biochemical tests. *Meth Microbiol* 1971; 6A: 1-32.
- 436 37. **Antunes A, Taborda M, Huber R, Moiss C, Fernanda M et al.** *Halorhabdus tiamatea*  
437 sp. nov., a non-pigmented, extremely halophilic archaeon from a deep-sea, hypersaline  
438 anoxic basin of the Red Sea, and emended description of the genus *Halorhabdus*. *Int J Syst*  
439 *Evol Microbiol* 2008; 58: 215-220
- 440 38. **Wainø M, Ingvorsen K.** Production of b-xylanase and b-xylosidase by the extremely  
441 halophilic archaeon *Halorhabdus utahensis*. *Extremophiles* 2003; 7: 87-93.
- 442 39. **Shimane Y, Hatada Y, Minegishi H, Mizuki T, Echigo A et al.** *Natronoarchaeum*  
443 *mannanilyticum* gen. nov., sp. nov., an aerobic, extremely halophilic archaeon isolated from  
444 commercial salt. *Int J Syst Evol Microbiol* 2010; 60: 2529–2534.
- 445 40. **Song HS, Cha I-T, Yim KJ, Lee H-W, Hyun D-W et al.** *Halapricum salinum* gen. nov.,  
446 sp. nov., an extremely halophilic archaeon isolated from non-purified solar salt. *Ant van*  
447 *Leeuwenhoek* 2014; 105: 979–986.

448

449

450

451

452

453 **Table 1.** Average pairwise genomic Nucleotide Identity (ANI-P) and digital DNA-DNA  
 454 hybridization analyses (% similarity) of strain HArce11<sup>T</sup> with the nearest phylogenetic relatives  
 455 from the family *Haloarculaceae*.

Compared with:	ANI-P			Digital DDH (average from 3 formulas)	
	Strain HArce11 <sup>T</sup>	<i>Halorhabdus</i> <i>tiamatea</i>	<i>Halothabodus</i> <i>utahensis</i>	<i>Halapricum</i> <i>salinum</i>	Strain HArce11 <sup>T</sup>
<i>Halorhabdus</i> <i>tiamatea</i> SARL4B <sup>T</sup>	<b>75.1</b>		85.6	75.7	16.6
<i>Halorhabdus</i> <i>utahensis</i> AX-2 <sup>T</sup>	<b>74.8</b>	85.6		75.3	16.4
<i>Halapricum</i> <i>salinum</i> CBA1105 <sup>T</sup>	<b>74.1</b>	75.7	75.2		15.7

456

457 **Table 2.** Comparative property of cellulotrophic haloarchaeon strain HArce11<sup>T</sup> with the nearest  
 458 phylogenetic relatives in *Haloarculaceae*: *Halorhabdus tiamatea* [14, 37], *Halopricum salinum*  
 459 [40].

Feature	Strain HArce11 <sup>T</sup>	<i>Halorhabdus tiamatea</i> JCM 14471 <sup>T</sup>	<i>Halopricum salinum</i> CBA1105 <sup>T</sup>
Cell morphology	Non-motile coccoids	Pleomorphic, non-motile	Pleomorphic cocci, non-motile
Pigmentation	Pale orange	-	Red
Growth substrates: <u>polymers</u>	Insoluble celluloses, xylan	pullulan <sup>#</sup> , starch, xyloglycan <sup>#</sup> , xylane <sup>#</sup> , arabinoxylane <sup>#</sup> , glycomannan <sup>#</sup> , beta-mannan (weak) <sup>#</sup>	-
<u>sugars</u>	Cellobiose	Galactose, maltose, mannose <sup>#</sup> , xylose <sup>#</sup>	Glucose, mannose, maltose, sucrose
<u>others</u>			glutamate
Number of cellulase genes (GH families) in the genome	GH5 (24); GH9 (3); GH12 (2)	GH5 (6); GH9 (1); GH12 (1)	none
Anaerobic growth	-	+ (fermentative, denitrification)	-
Esterase/lipase	- (tributyryn/ olive oil)	+ (C8)/nd	Tweens/nd
Protease activity	- (casein, gelatin)	+ (gelatin)	-
Oxidase/catalase	weak/+	-/+	+/-
Salinity range (opt.) M NaCl	2.5-5 (3.5-4.0)	1.6-5 (4.5)	2.5-6.0 (3.2)
pH range (opt.)	6.5-8.0 (7.0-7.2)	6.0-8.5 (7.0-7.5)	7.0-8.0 (7.0)
Temperature (°C)	max. 50 (opt. 43)	max. 55 (opt. 45)	max. 45 (37)
Core lipids	C <sub>20</sub> -C <sub>20</sub> , C <sub>25</sub> -C <sub>20</sub>	DGE (undefined)	nd
Intact membrane polar lipids	PGP-Me, PG, DGD, PGP, PGS; unknown sulfolipid	PG, PGP-Me, TGD, S <sub>1</sub> -DGD	PG, PGP-Me, 3 unidentified glycolipids
DNA G+C (mol%)	65.7 (genome)	61.7 (T <sub>m</sub> )	66.0 (T <sub>m</sub> )
Habitat	Hypersaline salt lakes in s-w Siberia	Deep-sea hypersaline brines (Red Sea)	Solar saltern

460 Phospholipids: (PGP-Me) phosphatidylglycerophosphate methylester, (PG) phosphatidylglycerol, (GL-PG)  
 461 phosphatidylglycose, (DGD) diglycosyl glycerol diether, (PGS) phosphatidylglycerol sulfate, (PGP)  
 462 phosphatidylglycerophosphate; glycolipids: (S<sub>1</sub>-DGD) monosulfated diglycosyl diether, TGD (triglycosyl glycerol  
 463 diether).

464 \* based on the genomic data and activity measurements but not yet validated by growth experiments

465 #determined in this work; negative results for *H. tiamatea* included amylopectin, dextrans, inulin, galactan,  
 466 galactomannan, beta-1,3 glycans, arabinan, arabinogalactan and various forms of native insoluble cellulose

467

468 **Legends to the figures**

469  
 470 **Fig. 1** Morphology of strain HArce11<sup>T</sup> growing at 4 M total NaCl and 37°C. (a) colonies on  
 471 amorphous cellulose plates forming large hydrolysis zones; (b) phase contrast microphotograph of  
 472 cells grown with amorphous cellulose in liquid culture; (c) phase contrast microphotograph of cells  
 473 forming biofilm on a cellulose fiber; (d) electron microscopy of thin sections of cells grown with  
 474 amorphous cellulose. **CW**, cell wall; **CM**, cytoplasmic membrane; **N**, nucleoid.

475  
 476 **Fig. 2.** Phylogeny of strain HArce11<sup>T</sup>.  
 477 (a) Maximum Likelihood 16S rRNA gene sequence-based phylogenetic tree showing position of  
 478 HArce11<sup>T</sup> (in bold) within the order *Halobacteriales*. Branch lengths (see scale) correspond to the  
 479 number of substitutions per site with corrections, associated with the model (GTR, G + I, 4  
 480 categories). All positions with less than 95% site coverage were eliminated. Totally 1435 positions  
 481 were used in the alignment of 119 sequences. Numbers at nodes indicate bootstrap values of 1000  
 482 repetitions, bootstrap values below 50% are not shown. *Halomarina* genus was used as an outgroup.  
 483 (b) Maximum Likelihood *rpoB'* gene sequence-based tree showing position of strain HArce11<sup>T</sup> (in  
 484 bold) within the order *Halobacteriales*. All parameters were the same as in 16S rRNA gene-based  
 485 phylogeny. Totally 1827 positions were used in the alignment of 81 sequences. *Halomarina* genus  
 486 was used as an outgroup.  
 487 (c) Maximum Likelihood tree based on 17 ribosomal proteins alignment showing position of strain  
 488 HArce11<sup>T</sup> (in bold) within the order *Halobacteriales*. Branch lengths (see scale) correspond to the  
 489 number of substitutions per site with corrections, associated with the model (LG, G + I, 4  
 490 categories). All positions with less than 95% site coverage were eliminated. Totally 2938 positions  
 491 were used in the alignment of 40 amino acid sequences. *Natronomonas* genus was used as an  
 492 outgroup

493



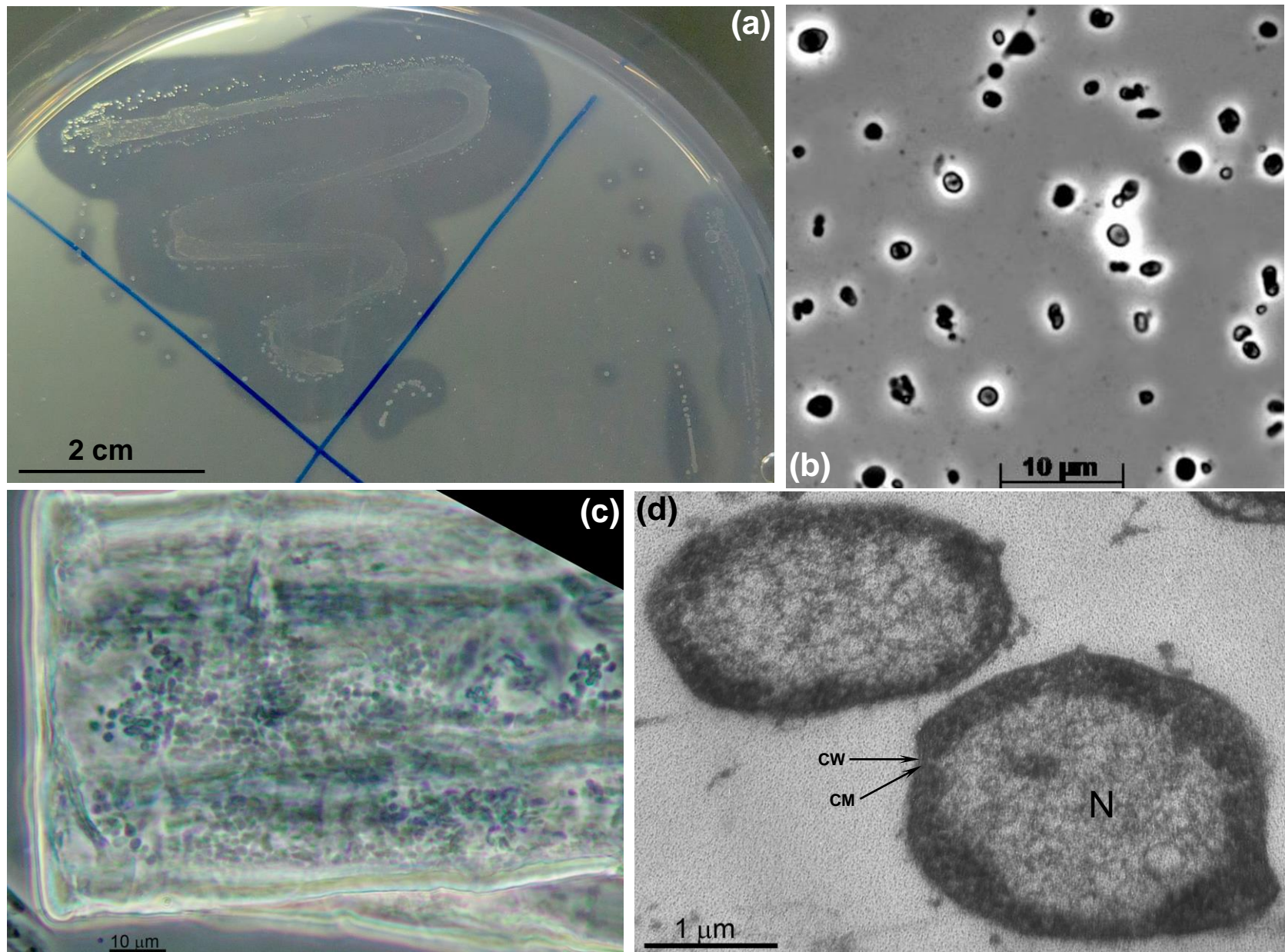


Fig.1

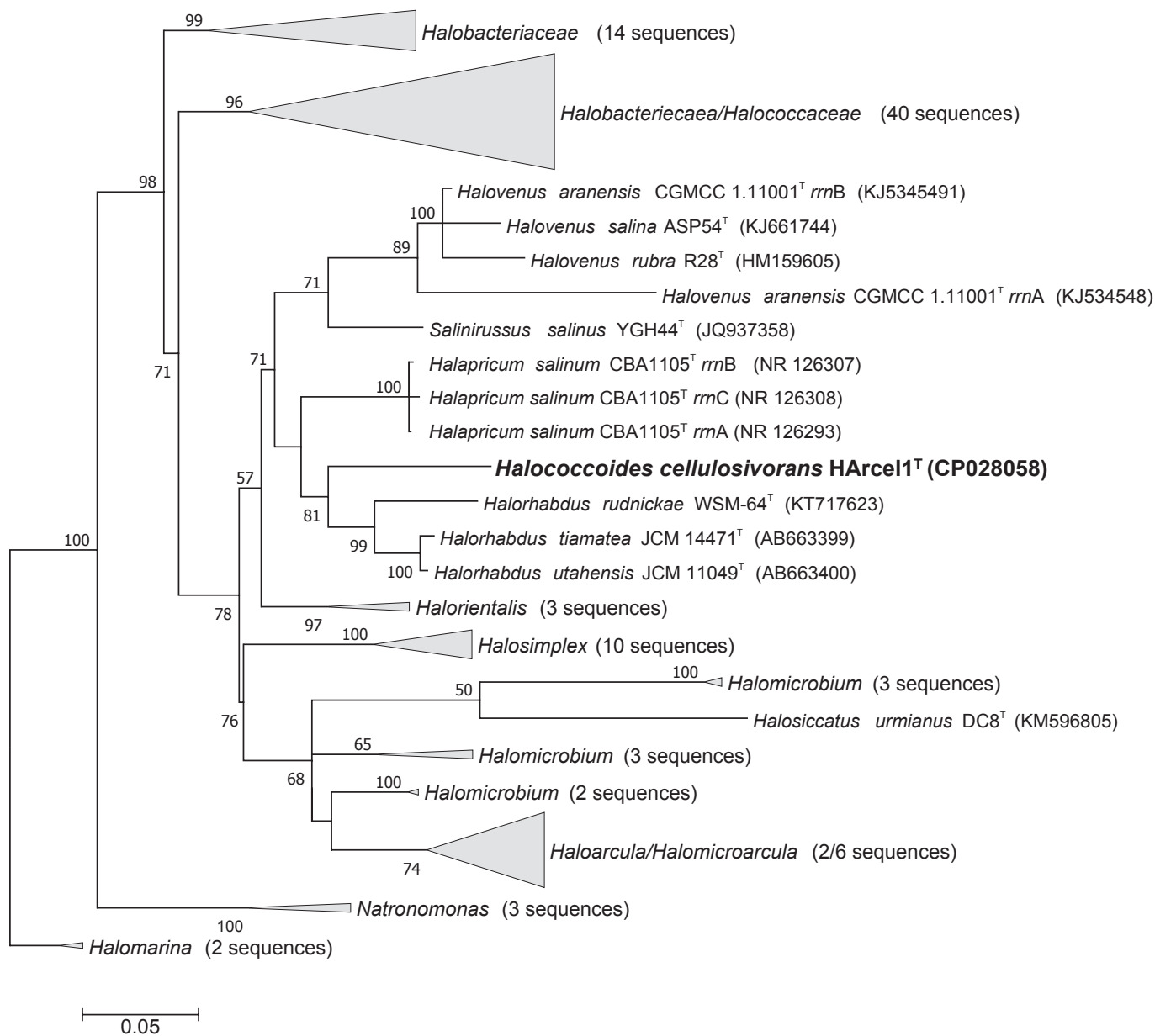
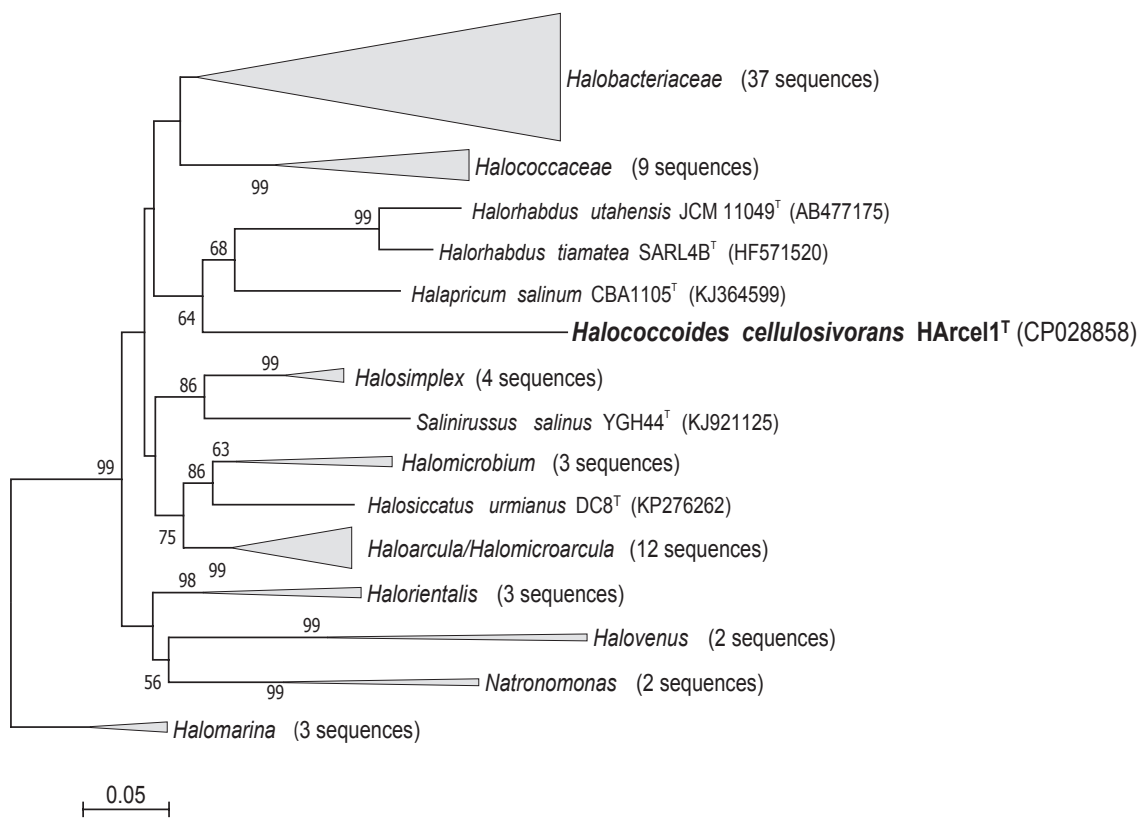
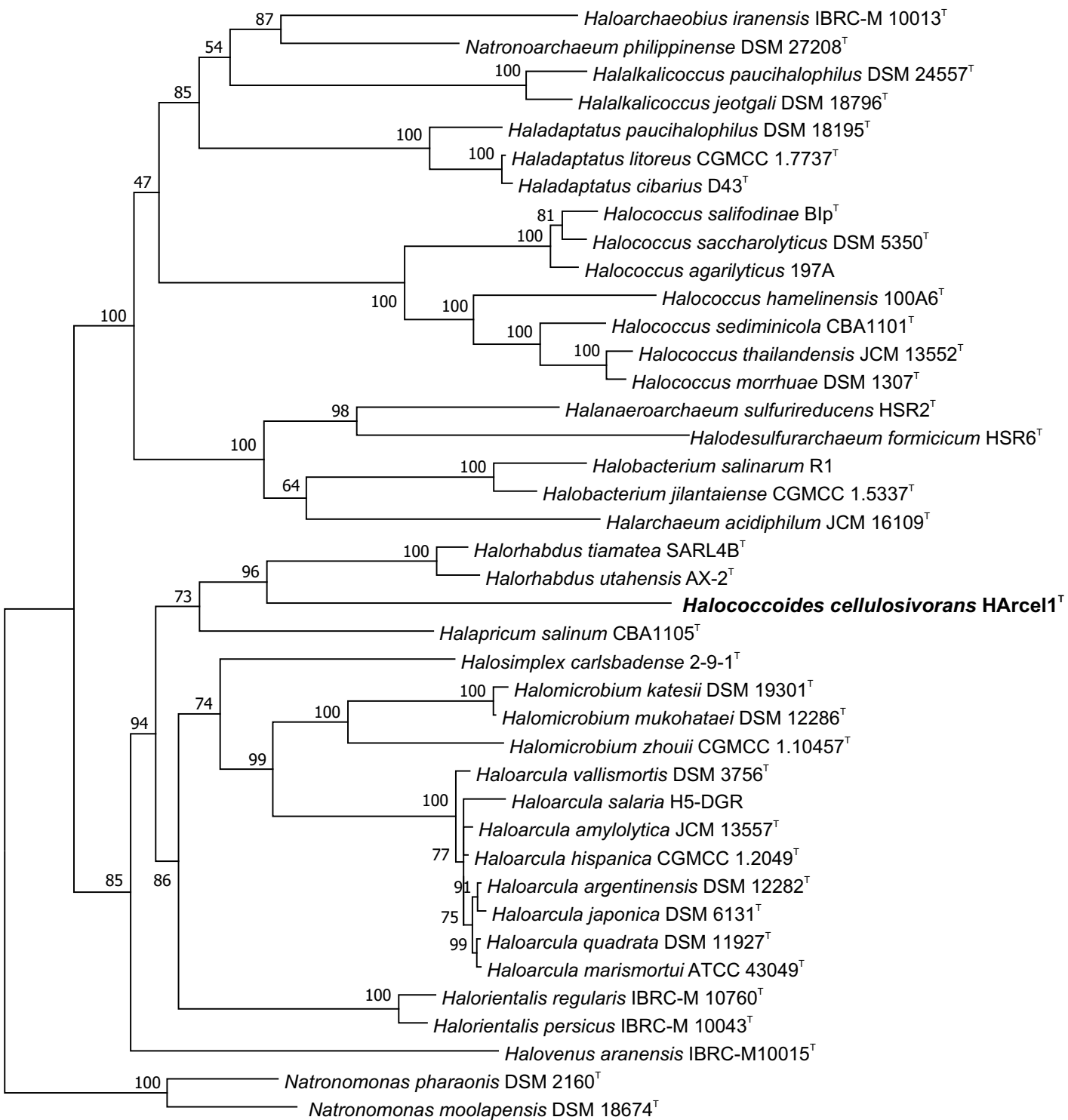


Fig. 2a



**Fig. 2b**



0.05

Fig. 2c

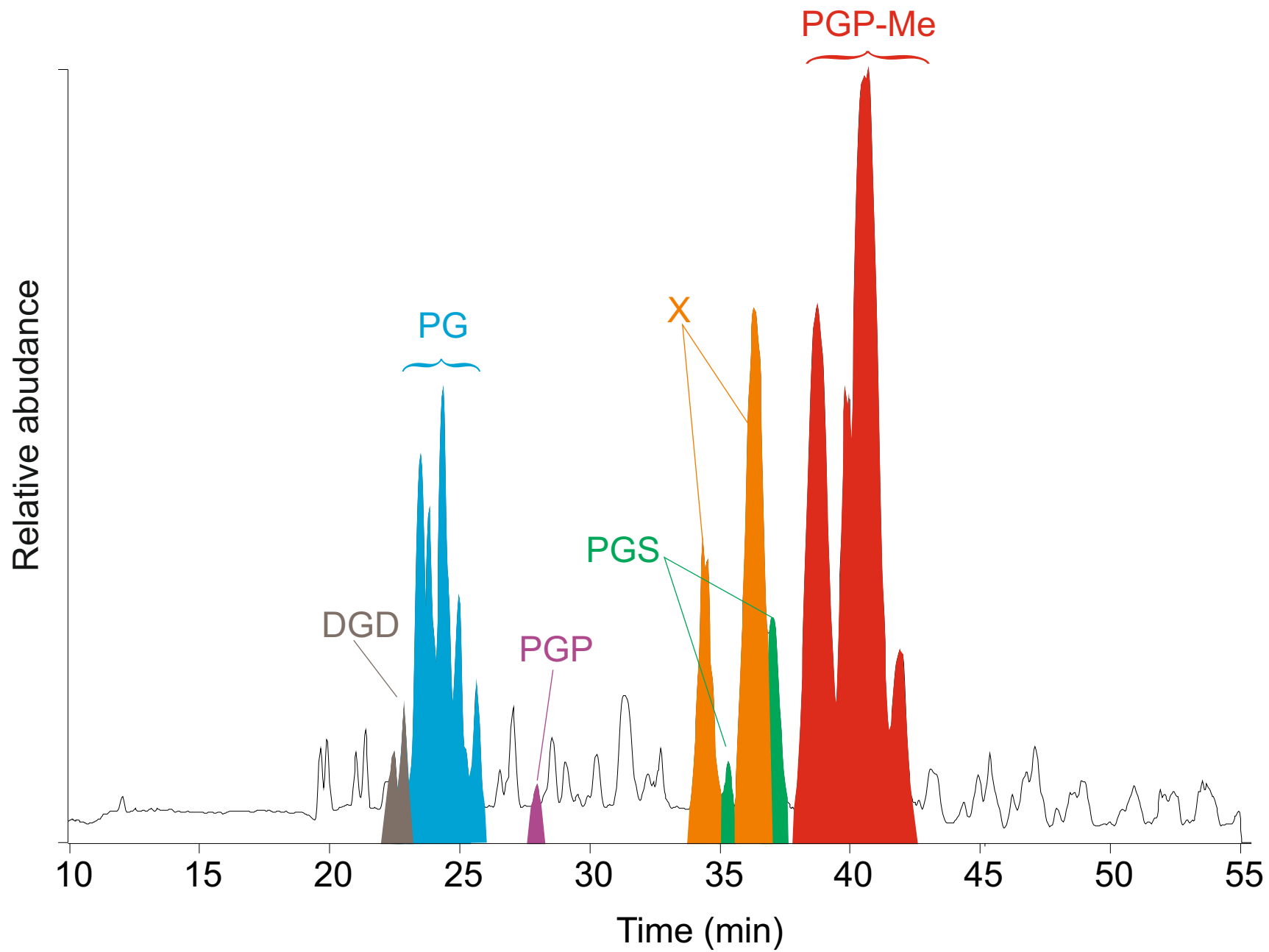
## Supplementary data file

### ***Halococcoides cellulovorans* gen. nov., sp. nov., an extremely halophilic cellulose-utilizing haloarchaeon from hypersaline lakes**

Dimitry Y. Sorokin, Tatiana V. Khijniak, Nadezhda A. Kostrikina, Alexander G. Elcheninov, Stepan V. Toshchakov, Nicole J. Bale, Jaap S. Sinninghe Damsté, Ilya V. Kublanov

### Supplementary Figure S1

Partial base peak chromatogram (Gaussian smoothed) of the HPLC-ESI/MS analysis of intact polar lipids in the cell extract of strain Harcell<sup>T</sup>. Peak labels: PGP-Me = phosphatidylglycerophosphate methylester, PG = phosphatidylglycerol, DGD = diglycosyl diether, X = unknown sulfur containing headgroup, PGP = phosphatidylglycerophosphate and PGS = phosphatidylglycerosulfate. Double or multiple peaks are due to the presence of the polar head group with both the archaeol core (C<sub>20</sub>-C<sub>20</sub> dialkyl glycerol ether) and the extended archaeol core (C<sub>20</sub>-C<sub>25</sub>) as well as their unsaturated homologs.



Supplementary Figure S1