Draft Genome Sequence of an Extreme Haloarchaeon 3A1-DGR Isolated from a Saltern Crystallizer of the Little Rann of Kutch, India

Kamal K. Pal, Rinku Dey, Manesh Thomas, Sucheta Ghorai, Dharmesh Sherathia, Sejal Vanpariya, Rupal Rupapara, Priya Rawal, et al.

Indian Journal of Microbiology
The Official Publication of the Association of Microbiologists of India

ISSN 0046-8991

Indian J Microbiol
DOI 10.1007/s12088-014-0483-7

Springer

Indian J Microbiol
DOI 10.1007/s12088-014-0483-7

SHORT COMMUNICATION

Draft Genome Sequence of an Extreme Haloarchaeon 3A1-DGR Isolated from a Saltern Crystallizer of the Little Rann of Kutch, India

Kamal K. Pal · Rinku Dey · Manesh Thomas · Sucheta Ghorai · Dharmesh Sherathia · Sejal Vanpariya · Rupal Rupapara · Priya Rawal · Mona Mandaliya · Bhoomika Sukhadiya · Anil K. Saxena

Received: 20 February 2014 / Accepted: 21 June 2014

Abstract Haloarchaea are predominant in the salt crystallizers of the Rann of Kutch when the concentration of salts approaches saturation levels. The obligate and extreme halophilic archaeon 3A1-DGR, isolated from a salt crystallizer pond of the Little Rann of Kutch, India, needs minimum of 10 % NaCl in the growth medium. To understand the mechanism(s) of osmotolerance and adaptation at extreme osmolarity, and to mine relevant gene(s), the genome of this haloarchaeon, 3A1-DGR, was sequenced. We report here, the 2.88 Mb draft genome sequence of the haloarchaeon 3A1-DGR, with G+C content of 68 % and the possible involvement of 43 genes in stress tolerance. Further studies of the genome of this haloarchaeon would be required to identify gene(s) that might be responsible for imparting extreme osmotolerance.

Keywords Haloarchaea · Osmotolerance · Draft whole genome sequencing · Extreme halophile

Archaea are believed to have evolved 4 billion years ago [1], and can survive in the extremities of pH, temperature, pressure, salts, etc., besides playing significant roles in the bio-geochemical cycles [2]. The salt marshes of the Great and the Little Rann of Kutch of Gujarat, India, harbour extreme halophilic archaea, besides few genera of bacilli like *Bacillus*, *Salinibacillus*, *Thalassobacillus*, *Sediminibacillus*, etc. The concentration of salts in these hypersaline regions and that of the salt crystallizers gradually approaches the point of saturation when salt starts crystallizing, allowing specific group of organisms to thrive in such conditions. The genomes of a number of bacilli of the genera *Bacillus*, *Salinibacillus*, *Thalassobacillus*, and *Sediminibacillus*, isolated from the salt crystallizers of the Little and Great Rann of Kutch, have been sequenced recently to understand the mechanism(s) of halophilism(s) and to isolate relevant gene(s) [3–7].

The present haloarchaeon 3A1-DGR was isolated from a sample (W3A; total salts concentration of 418.7 g/l; pH: 6.7) collected from a salt crystallizer pond (N 23°44.30′ E 71°10.81′) of the Little Rann of Kutch, Gujarat, India. This extreme halophilic archaeon was isolated by spread plating of the diluted samples into different known standard halophilic media including complex medium (CM), minimum growth medium (MGM), modified growth medium for haloarchaea, halophilic medium (HM), standard growth medium (SGM), DSMZ medium 1184, DSMZ medium 97, etc. containing different concentrations of NaCl (18–23 %) [8–11]. The Petri dishes were incubated at different temperatures (28, 37 and 42 °C) for 15–20 days.

The haloarchaeon 3A1-DGR is an isolate of a lineage of isolates from the salt crystallizers of the Little Rann of Kutch, India and has been described recently [12]; which lies between *Natronomonas* and *Halopenitus* on the basis of partial 16S rRNA sequences. The lineage consists of three haloarchaeal isolates, namely 3A1-DGR (16S rRNA GenBank acccession no. JF802160), H9-DGR, and 2ANA-DGR.

Electronic supplementary material The online version of this article (doi:10.1007/s12088-014-0483-7) contains supplementary material, which is available to authorized users.

K. K. Pal (✉) · R. Dey · M. Thomas · S. Ghorai · D. Sherathia · S. Vanpariya · R. Rupapara · P. Rawal · M. Mandaliya · B. Sukhadiya
Microbiology Section, Directorate of Groundnut Research (ICAR), Ivnagar Road, PB No. 5, Junagadh, Gujarat 362001, India
e-mail: pal@nrcg.res.in; kkpal9426476749@gmail.com

A. K. Saxena
Division of Microbiology, Indian Agricultural Research Institute, New Delhi 110012, India

Published online: 10 July 2014

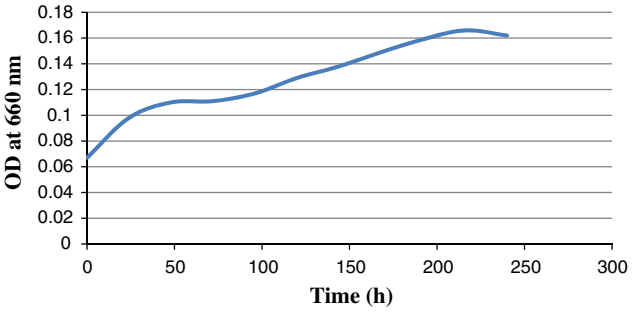

Fig. 1 Growth curve of extreme haloarchaeon 3A1-DGR at 20 % NaCl in CM medium

Comparison of 3A1-DGR with *Halopenitus* revealed that whereas *Halopenitus persicus* DC30(T) is pleomorphic (rod to triangular or disc shaped), forms pale pink-pigmented colonies, needs at least 0.02 M $MgCl_2$ for growth, optimum being 0.1 M [13], the cells of 3A1-DGR are irregularly coccoid (outer diameter of 1.4–1.5 μm) and formed hard and tough colonies which produce creamy pigmentation when incubated in dark, and light red pigmentation in illuminated conditions [12], and does not require any $MgCl_2$ for growth. Moreover, G + C content of 3A1-DGR is 68 % and contains phosphotidylglycerol phosphate (PGP), phosphotidyl glycerol phosphate methyl ester (PGPME) and diglycosyl diether (DGD) as major polar lipids against G+C content of 66 % and major polar lipids of PGP and PGPME in *Halopenitus persicus* DC30(T). It also differs from *Halopenitus malekzadehii* CC65 (T) by the fact that CC65(T) needs at least 0.02 M $MgCl_2$ for initiation of growth, 0.4 M being optimum; produces light yellow colonies and has G+C content of 63.8 % [14]. Studies revealed that the obligate haloarchaeon 3A1-DGR needs a minimum of 10 % NaCl in growth media for initiation of growth and can continue to grow up to 35 % NaCl and 42 °C, with optimum growth requirement of around 20 % NaCl, 40 °C and pH 7.5 at which it can reach a log phase of growth in around 5 days (Fig. 1) in complex medium (CM). The present isolate does not survive lyophilisation because of obligate halophilic nature and hence preserved in glycerol stock at −80 °C; in stabs and in slants at 4 °C, and in liquid broth at room temperature keeping minimum of 10 % of NaCl in the growth media and sub-cultured bi-monthly for maintenance and further use.

To understand the mechanisms of obligate halophilism, the genome of 3A1-DGR was sequenced by Roche 454 genome analyser (GS FLX). Both shotgun and 3 KB mate-paired library sequencing were performed. In shotgun sequencing, 842518 reads with average read length of 468 bases were obtained. However, in mate-paired (3 KB) library sequencing, 153960 and 106333 reads, respectively, with average read lengths of 455 and 423 bases, respectively, were obtained.

The genome was assembled de novo by GS de novo assembler v2.9 [15] with genome coverage of about 173X. The draft assembly of 3A1-DGR resulted into 3 scaffolds of 2880902 bases with the largest scaffold of 1817982 bases. The N_{50} scaffold length of 1817982 bases with an average size of 960300 bases (minimum 5278 bases; maximum 1817982 bases) was achieved. The assembly consists of 15 contigs (average size 192033 bases) with the N_{50} contig of 425867 bases and largest contigs of 628022 bases. All the assembly data were deposited in the DDBJ/EMBL/GenBank nucleotide sequence database.

The draft genome of 3A1-DGR was annotated further using Rapid Annotation of Subsystem Technology (RAST) server [16], Glimmer 3 [17, 18], GeneMarkS [19], tRNA-Scan-SE [20], RNAmmer [21], KEGG [22], and Signal P [23] for predicting subsystems, coding sequences (CDS), tRNA, rRNA genes, signal peptides, biochemical pathways, etc.

Using the different software tools, we identified 2894 coding sequences (CDS). The draft genome also encodes 46 RNA genes (43 tRNA and 3 rRNA genes) and 254 subsystems. There was one each of 5S, 16S and 23S rRNA in the draft genome of the haloarchaeon 3A1-DGR. Among the CDS, 1931 are not in a subsystem (nonhypothetical, 540; hypothetical, 1391), whereas 963 (nonhypothetical, 936; hypothetical, 27) are in a subsystem. This indicates that nearly 66 % of the CDS could not be placed in a subsystem and nearly 49 % of the CDS have hypothetical functions. RAST annotation also revealed the association of 43 genes in stress responses in this organism: 10 in osmotic stress (10 in choline and betaine uptake and betaine biosynthesis), 28 in oxidative stress (2 in protection from reactive oxygen species [ROS], 16 in oxidative stress, 2 in glutathione:biosynthesis and gamma-glutamyl cycle, 4 in glutathione:non-redox reactions, 3 in rubrerythrin, and 1 in glutaredoxins), 2 in detoxification and 3 in no subcategory. Further analysis identified 15 signal peptides, and three closest neighbours of haloarchaeon 3A1-DGR as *Halogeometricum borinquense* DSM 11551 (genome ID 469382.4), *Haloquadratum walsbyi* DSM 16790 (genome ID 362976.10) and *Haloarcula marismortui* ATCC 43049 (genome ID 272569.1) with <90 % similarity. In addition, 306 genes have been mapped to different pathways involved in the biosynthesis and degradation of amino acids and derivatives, including 50 in branched-chain amino acid pathways. Similarly, 167 genes have also been mapped to different pathways of central carbohydrate metabolism including 26 in the serine-glyoxylate cycle.

To unravel the mechanism(s) of osmotolerance and to identify relevant gene(s), comparative genomics will be initiated on completion of the filling of the gaps in the

sequence data. Deciphering the functions of hypothetical proteins would also lead to identification of a number of novel genes. Future studies are underway in this direction.

Nucleotide sequence accession numbers

This Whole Genome Shotgun (WGS) project has been deposited at DDBJ/EMBL/GenBank under the accession number ATCR00000000. The version described in this paper is version ATCR02000000. Bioproject registered under accession: PRJNA183189 ID:146531.

Acknowledgments The work was carried out in the subproject "Exploring bacterial and archaebacterial diversity in Kutch eco-region of Gujarat for agricultural and industrial applications" under the Application of Microorganisms in Agriculture and Allied Sectors (AMAAS) of the Indian Council of Agricultural Research (ICAR). Authors thank ICAR for funding through AMAAS. Authors thank the Directors of the Directorate of Groundnut Research, Junagadh, and NBAIM, Mau, for help during the course of this study. The authors thank Macrogen Inc., South Korea, and Sequencher Tech Pvt. Ltd., Ahmedabad, India, for the genomic services.

References

1. Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria and Eucarya. Proc Natl Acad Sci (USA) 87:4576–4579
2. Oren A (2008) Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline Syst 4:2. doi:10.1186/1746-1448-4-2
3. Pal KK, Dey R, Sherathia D, Dalsania T, Savsani K, Patel I, Thomas M, Ghorai S, Vanpariya S, Rupapara R, Acharya N, Rawal P, Joshi P, Sukhadiya B, Mandaliya M, Saxena AK (2013) Draft genome sequence of *Salinibacillus aidingensis* strain MSP4, an obligate halophilic bacterium isolated from a salt crystallizer of the Rann of Kutch, India. Genome Announc 1(4):e00253–13. doi:10.1128/genomeA.00253-13
4. Pal KK, Dey R, Thomas M, Sherathia D, Dalsania T, Patel I, Savsani K, Ghorai S, Vanpariya S, Sukhadiya B, Mandaliya M, Rupapara R, Rawal P, Saxena AK (2013) Draft genome sequence of *Bacillus* sp. strain SB47, an obligate extreme halophilic bacterium isolated from a salt pan of the Little Rann of Kutch, India. Genome Announc 1(5):e00816–13. doi:10.1128/genomeA.00816-13
5. Pal KK, Dey R, Thomas M, Sherathia D, Dalsania T, Patel I, Savsani K, Ghorai S, Vanpariya S, Sukhadiya B, Mandaliya M, Rupapara R, Rawal P (2013) Draft genome sequence of the extremely halophilic *Bacillus* sp. strain SB49, isolated from a salt crystallizer pond of the Little Rann of Kutch, India. Genome Announc 1(5):e00869–13. doi:10.1128/genomeA.00869-13
6. Pal KK, Dey R, Sherathia D, Sukhadiya B, Dalsania T, Patel I, Savsani K, Thomas M, Vanpariya S, Mandaliya M, Rupapara R, Rawal P, Ghorai S, Bhayani S, Shah A, Saxena AK (2013) Draft genome sequence of an obligate and moderately halophilic bacterium, *Thalassobacillus devorans* strain MSP14, the first draft genome of the genus *Thalassobacillus*. Genome Announc 1(6):e01103–e01113. doi:10.1128/genomeA.01103-13
7. Dey R, Pal, KK, Sherathia D, Sukhadiya B, Dalsania T, Patel I, Savsani K, Thomas M, Vanpariya S, Mandaliya M, Rupapara R, Rawal P, Ghorai S, Bhayani S, Shah A, Saxena AK (2014) Insight into the first draft genome sequence of the genus *Sediminibacillus*, *Sediminibacillus halophilus* strain NSP9.3. Genome Announc 2(1):e01133–13. doi:10.1128/genomeA.01133-13
8. Asker D, Ohta Y (2002) *Haloferax alexandrinus* sp. nov., an extremely halophilic canthaxanthin–producing archaeon from a solar saltern in Alexandria (Egypt). Int J Syst Evol Microbiol 52:729–738
9. Nuttall SD, Dyall-Smith ML (1993) HF1 and HF2: novel bacteriophages of halophilic archaea. Virology 197:678–684
10. Oren A, Lau PP, Fox GE (1988) The taxonomic status of '*Halobacterium marismortui*' from the Dead Sea: a comparison with *Halobacterium vallismortis*. Syst Appl Microbiol 10:251–258
11. Porter K, Kukkaro P, Bamford JK, Bath C, Kivelä HM, Dyall-Smith ML, Bamford DH (2005) SH1: a novel, spherical halovirus isolated from an Australian hypersaline lake. Virology 335:22–33
12. Thomas M, Pal KK, Dey R, Saxena AK, Dave SR (2012) A novel haloarchaeal lineage widely distributed in the hypersaline marshy environment of Little and Great Rann of Kutch in India. Cur Sci 103:1078–1084
13. Amoozegar MA, Makhdoumi-Kakhki A, Mehrshad M, Fazeli SA, Azarbaijani R, Ventosa A (2012) *Halopenitus persicus* gen. nov., sp. nov., an archaeon from an island salt lake. Int J Syst Evol Microbiol 62(Pt 8):1932–1936. doi:10.1099/ijs.0.036921-0
14. Amoozegar MA, Makhdoumi-Kakhki A, Mehrshad M, Fazeli SA, Ventosa A (2013) *Halopenitus malekzadehii* sp. nov., an extremely halophilic archaeon isolated from a salt lake. Int J Syst Evol Microbiol 63(Pt 9):3232–3236. doi:10.1099/ijs.0.050500-0
15. Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J (2010) De novo assembly of human genomes with massively parallel short read sequencing. Genome Res 20:265–272
16. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008) The RAST Server: rapid annotations using subsystems technology. BMC Genom 9:75. doi:10.1186/1471-2164-9-75
17. Salzberg SL, Delcher AL, Kasif S, White O (1998) Microbial gene identification using interpolated Markov models. Nucleic Acids Res 26:544–548
18. Delcher AL, Bratke KA, Powers EC, Salzberg SL (2007) Identifying bacterial genes and endosymbiont DNA with glimmer. Bioinformatics 23:673–679
19. Besemer J, Lomsadze A, Borodovsky M (2001) GeneMarkS: a self-training method for prediction of gene starts in microbial genome. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 29:2607–2618
20. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964
21. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108
22. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004) The KEGG resource for deciphering the genome. Nucleic Acids Res 32:277–280
23. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786