Systematic and Applied Microbiology xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

Genomic comparison between members of the *Salinibacteraceae* family, and description of a new species of *Salinibacter* (*Salinibacter altiplanensis* sp. nov.) isolated from high altitude hypersaline environments of the Argentinian Altiplano

Tomeu Viver[a,*], Luis Orellana[b], Pedro González-Torres[c], Sara Díaz[a], Mercedes Urdiain[a], María Eugenia Farías[d], Vladimir Benes[e], Peter Kaempfer[f], Azadeh Shahinpei[g], Mohammad Ali Amoozegar[g], Rudolf Amann[h], Josefa Antón[c], Konstantinos T. Konstantinidis[b], Ramon Rosselló-Móra[a]

[a] *Marine Microbiology Group, Department of Ecology and Marine Resources, Mediterranean Institute for Advanced Studies (IMEDEA, CSIC-UIB), E-07190 Esporles, Spain*
[b] *School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA*
[c] *Department of Physiology, Genetics and Microbiology, University of Alicante, Alicante, Spain*
[d] *Laboratorio de Investigaciones Microbiológicas de Lagunas Andinas (LIMLA), Planta Piloto de Procesos Industriales Microbiológicos (PROIMI), CCT, CONICET, Tucumán, Argentina*
[e] *European Molecular Biology Laboratory (EMBL), Genomics Core Facility, Meyerhofstr. 1, D-69117 Heidelberg, Germany*
[f] *Institut für Angewandte Mikrobiologie, Justus-Liebig-Universität Giessen, IFZ-Heinrich-Buff-Ring, Giessen, Germany*
[g] *Extremophiles Laboratory, Department of Microbiology, Faculty of Biology and Center of Excellence in Phylogeny of Living Organisms, Collage of Science, University of Tehran, Tehran, Iran*
[h] *Department of Molecular Ecology, Max-Planck-Institut für Marine Mikrobiologie, Bremen D-28359, Germany*

ARTICLE INFO

Article history:
Received 22 November 2017
Received in revised form
22 December 2017
Accepted 23 December 2017

Keywords:
Salinibacter ruber
Halophiles
MALDI-TOF MS
Salterns
Salt lake
Genomic

ABSTRACT

The application of tandem MALDI-TOF MS screening with 16S rRNA gene sequencing of selected isolates has been demonstrated to be an excellent approach for retrieving novelty from large-scale culturing. The application of such methodologies in different hypersaline samples allowed the isolation of the culture-recalcitrant *Salinibacter ruber* second phylotype (EHB-2) for the first time, as well as a new species recently isolated from the Argentinian Altiplano hypersaline lakes. In this study, the genome sequences of the different species of the phylum *Rhodothermaeota* were compared and the genetic repertoire along the evolutionary gradient was analyzed together with each intraspecific variability. Altogether, the results indicated an open pan-genome for the family *Salinibacteraceae*, as well as the codification of relevant traits such as diverse rhodopsin genes, CRISPR-Cas systems and spacers, and one T6SS secretion system that could give ecological advantages to an EHB-2 isolate. For the new *Salinibacter* species, we propose the name *Salinibacter altiplanensis* sp. nov. (the designated type strain is AN15^T = CECT 9105^T = IBRC-M 11031^T).

Introduction

Salinibacter ruber has been discovered to be the first extreme halophilic member of the domain Bacteria with cell abundances equivalent to most haloarchaea [4]. *Sal. ruber* (for the rationale of the abbreviations used see Materials and methods) was first observed using culture-independent approaches, which led to the description of *Candidatus* "Salinibacter" that was subsequently formally classified as a new species after isolation in pure culture [3]. *Sal. ruber* was the first species described from a new lineage of extreme halophilic microorganisms, monophyletic with another extremophile (*Rhodothermus marinus*), and which were both loosely affiliated to the *Chlorobium* and *Bacteroidetes* phyla [65]. Extensive research in hypersaline environments led to the isolation and classification of two additional members of the genus *Salinibacter* (*Sal. iranicus* and *Sal. luteus*) from the Iranian salt lake

* Corresponding author.
E-mail address: tviver@imedea.uib-csic.es (T. Viver).

https://doi.org/10.1016/j.syapm.2017.12.004
0723-2020/

Aran-Bidgol [36], and one new genus *Salisaeta* (*Sat. longa* [68]) that originated from a mixture of water from the Dead Sea and the Red Sea. Recently, the lineage comprising the three genera has been classified as a single phylum *Rhodothermaeota*, and the Iranian species were reclassified as the new genus *Salinivenus* [42] due to their phylogenetic distance from the type species of *Salinibacter*.

Since its discovery, numerous members of the species *Sal. ruber* have been isolated from diverse and distant environments (e.g. Refs. [2,63]). This species was shown to be formed by two closely related phylotypes (EHB-1 and EHB-2) coexisting in the same environment [4]. However, in all cases, the isolates corresponded to the most abundant phylotype EHB-1, and during almost 20 years of research the members of the second phylotype EHB-2 escaped isolation. Representatives of this species have been isolated from several hypersaline spots around the world even from distant locations and different altitudes, such as Mediterranean coastal salterns and Peruvian Altiplano salterns [63]. Furthermore, sequences related to the phylotypes have been detected in several culture-independent surveys from different locations, such as Tuz Lake in Turkey [43], lakes in the Tibetan plateau [74], the Argentinian Pampa [18], or Lake Tyrrell in Australia [54].

High-throughput culturing in tandem with MALDI-TOF MS screening and 16S rRNA identification of isolates from hypersaline systems is a robust strategy for retrieving rare taxa from environmental samples [70]. Using this strategy, the pure cultures of two strains corresponding to the *Sal. ruber* EHB-2 phylotype are described in this current study together with members of a new species of *Salinibacter* thriving in hypersaline lakes of the Argentinian Altiplano. The genomes of these strains, together with the type strains of *Salinivenus*, were sequenced and compared with the available *Rhodothermaeota* genomes to reveal new genomic features of this extremely halophilic lineage.

Materials and methods

Strains and name abbreviations

Reference strains of *Sal. ruber* M31^T and M8 were obtained from our strain collection, and the type strains of *Slv. iranica* (CB7^T) and *Slv. lutea* (DGO^T) were provided by the co-author M. Amoozegar from his collection. In order to simplify the identification of the names, a three letter abbreviation was used, which was already commonly used for *Halobacteria* in accordance with the recommendation made by the International Committee on Systematic Bacteriology, Subcommittee on Taxonomy of *Halobacteriaceae* [46] as: *Salinibacter* = *Sal.*; *Salinivenus* = *Slv.*; *Salisaeta* = *Sat.*; and *Rhodothermus* = *Rho*.

Sample processing and strains studied

Two different athalassohaline salt lakes (Ojo Rojo in Antofalla and Salar de Llullialliaco), both located in the Argentinian Altiplano at altitudes above 3600 m, were sampled in February 2011 (Table 1). The salinity of the brines was 34% and the pH was 7. Cultures were obtained using salt water medium (SW) at a salt concentration of 25% [61]. The isolated strains were screened by MALDI-TOF MS (Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry) using whole cell biomass, as previously published [70]. An in-house MALDI-TOF profile database was used to generate a dendrogram and detect new *Salinibacter* isolates. 16S rRNA gene PCR amplification of the isolates was performed as previously published [70]. The two strains ST67 and SP273 had been isolated previously in a survey from the Trinitat (Tarragona, Spain) and Santa Pola (Alicante, Spain) solar salterns, respectively [70]. The type strains of *Slv. iranica* and *Slv. lutea* (CB7^T and DGO^T, respectively) were added to the genome sequencing strategy. For the global comparisons, the genomes of the already characterized *Sal. ruber* M31^T and M8 strains, and *Salisaeta longa* S4-4^T were incorporated (Table 1).

DNA extraction and genome assembly analysis

DNA extraction was performed as detailed by Urdiain et al. [67] and different methods of sequencing were used: Illumina Miseq (PE 300 × 2), Illumina Hiseq (PE 100 × 2), Roche GS FLX and Pacific Biosciences PacBio RS (Table 2). Illumina and Roche GS FLX reads were trimmed with a PHRED score quality threshold of 20 using SolexaQA v3.1.4 [14]. Different assembly softwares were used for each sequencing platform. Genomes sequenced by Illumina MiSeq were assembled using IDBA v1.1.1 [48] and by Illumina HiSeq using Velvet v1.0.13 [73]. A hybrid assembly methodology was used for assembling strain ST67, sequenced by Roche and PacBio: trimmed sequences from Roche were assembled using SPADES v.3.1.1 [6] and then ordered by using the long read information from the PacBio backbone using SSPACE-LongReads v1.1 [9]. Gene prediction from assembled contigs was conducted by using GeneMark.hmm with default parameters [8], and functional annotation was based on protein level searches against NCBI databases with Blast2Go v3.0.10 [13]. The annotations were compared with the RAST annotation, and metabolic pathways were analyzed using KAAS-KEGG [41]. CRISPR spacers were predicted in the genomes using CRISPRfinder [25].

Tree reconstructions based on rRNA and housekeeping genes

16S rRNA gene sequences were retrieved from the genomes and the alignments, and tree reconstructions were performed using the ARB software package version 5.5 [32]. The new sequences were added to the reference dataset LTP115 [72] and aligned using the SINA v1.2.12 tool (SILVA Incremental Aligner [55]) implemented within the ARB software package. Final alignments were manually improved following the reference alignment in ARB-editor. Sequences were used to reconstruct de novo trees using the neighbor-joining algorithm. 23S rRNA gene sequences and multilocus sequence analysis (MLSA) with 29 single-copy genes were also extracted from assembled genomes. The 23S rRNA genes were added to the LSURef 115 SILVA dataset, and selected sequences were aligned using the SINA aligner implemented in ARB-editor. The MLSA genes selected in this study were the same as those used in the revised phylogeny of *Bacteroidetes* [42]. MLSA genes were aligned individually using MUSCLE v3.8.31 [21] and were concatenated posteriorly. The neighbor-joining (NJ) [64] and RaxML v8.2.0 [66] algorithms were used for phylogenetic reconstructions as implemented in ARB. Tree reconstructions with NJ were performed using the Jukes Cantor correction, and RaxML reconstructions with the GTRGAMMA correction.

Core and pan-genome analysis; phylogenetic reconstruction

Predicted protein sequences were compared using an all-versus-all BLAST v2.2.28 [1] with available reference sequences in order to identify the shared reciprocal best matches in all pairwise genome comparisons using a 50% sequence similarity cut-off and over 50% or more of the query sequence length. All proteins shared between all sequenced genomes were aligned using MUSCLE v3.8.31 [21]. The concatenated and aligned orthologous genes were used to build phylogenetic trees in RAxML v8.2.0 [66]. The variable genes were defined as those present in two or more genomes but not in all genomes. The presence or absence of variable

Please cite this article in press as: T. Viver, et al., Genomic comparison between members of the *Salinibacteraceae* family, and description of a new species of *Salinibacter* (*Salinibacter altiplanensis* sp. nov.) isolated from high altitude hypersaline environments of the Argentinian Altiplano, Syst. Appl. Microbiol. (2018), https://doi.org/10.1016/j.syapm.2017.12.004