Systematic and Applied Microbiology 46 (2023) 126416

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.com/locate/syapm

Description of two cultivated and two uncultivated new *Salinibacter* species, one named following the rules of the bacteriological code: *Salinibacter grassmerensis* sp. nov.; and three named following the rules of the SeqCode: *Salinibacter pepae* sp. nov., *Salinibacter abyssi* sp. nov., and *Salinibacter pampae* sp. nov.

Tomeu Viver [a,b,*], Roth E. Conrad [c,d], Marianna Lucio [e], Mourad Harir [e,f], Mercedes Urdiain [a], Juan F. Gago [a], Ana Suárez-Suárez [a], Esteban Bustos-Caparros [a], Rodrigo Sanchez-Martinez [g], Eva Mayol [g], Federico Fassetta [h], Jinfeng Pang [i], Ionuț Mădălin Gridan [j], Stephanus Venter [k], Fernando Santos [g], Bonnie Baxter [l], María E. Llames [h], Adorján Cristea [m], Horia L. Banciu [n,o], Brian P. Hedlund [i], Matthew B. Stott [p], Peter Kämpfer [q], Rudolf Amann [b], Philippe Schmitt-Kopplin [e,f], Konstantinos T. Konstantinidis [c,d], Ramon Rossello-Mora [a,*]

[a] *Marine Microbiology Group, Department of Animal and Microbial Biodiversity, Mediterranean Institute for Advanced Studies (IMEDEA, CSIC-UIB), Esporles, Spain*
[b] *Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, Bremen, Germany*
[c] *Ocean Science & Engineering, School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA*
[d] *School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA*
[e] *Research Unit Analytical BioGeoChemistry, Helmholtz Munich, 85764 Neuherberg, Germany*
[f] *Chair of Analytical Food Chemistry, Technical University Munich, Maximus-von-Imhof-Forum 2, 85354 Freising, Germany*
[g] *Department of Physiology, Genetics and Microbiology, University of Alicante, 03690, San Vicent del Raspeig, Alicante, Spain*
[h] *Laboratorio de Ecología Acuática, Instituto Tecnológico Chascomús (INTECH)-CONICET-UNSAM, Escuela de Bio y Nanotecnologías -UNSAM, Buenos Aires, Argentina*
[i] *School of Life Sciences, University of Nevada, Las Vegas, NV 89154-4004, USA*
[j] *Doctoral School of Integrative Biology, Faculty of Biology and Geology, Babeș-Bolyai University, Cluj-Napoca, Romania*
[k] *Department of Biochemistry, Genetics and Microbiology, and Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa*
[l] *Great Salt Lake Institute, Westminster College, Salt Lake City, UT, 84105, USA*
[m] *Department of Taxonomy and Ecology, Faculty of Biology and Geology, Babeș-Bolyai University, Cluj Napoca, Romania*
[n] *Department of Molecular Biology and Biotechnology, Faculty of Biology and Geology, Babeș-Bolyai University, Cluj Napoca, Romania*
[o] *Emil G. Racoviță Institute, Babeș-Bolyai University, Cluj Napoca, Romania*
[p] *School of Biological Sciences, University of Canterbury, Christchurch, New Zealand*
[q] *Institute of Applied Microbiology (IFZ), Justus Liebig Universität Giessen, Giessen, Germany*

ARTICLE INFO

Keywords:
Salinibacter
Phylogeny
Genomics
Taxonomy
SeqCode
ICNP

ABSTRACT

Current -omics methods allow the collection of a large amount of information that helps in describing the microbial diversity in nature. Here, and as a result of a culturomic approach that rendered the collection of thousands of isolates from 5 different hypersaline sites (in Spain, USA and New Zealand), we obtained 21 strains that represent two new *Salinibacter* species. For these species we propose the names *Salinibacter pepae* sp. nov. and *Salinibacter grassmerensis* sp. nov. (showing average nucleotide identity (ANI) values < 95.09% and 87.08% with *Sal. ruber* M31^T, respectively). Metabolomics revealed species-specific discriminative profiles. *Sal. ruber* strains were distinguished by a higher percentage of polyunsaturated fatty acids and specific N-functionalized fatty acids; and *Sal. altiplanensis* was distinguished by an increased number of glycosylated molecules. Based on sequence characteristics and inferred phenotype of metagenome-assembled genomes (MAGs), we describe two new members of the genus *Salinibacter*. These species dominated in different sites and always coexisted with *Sal. ruber* and *Sal. pepae*. Based on the MAGs from three Argentinian lakes in the Pampa region of Argentina and the MAG of the Romanian lake Fără Fund, we describe the species *Salinibacter pampae* sp. nov. and *Salinibacter abyssi* sp. nov. respectively (showing ANI values 90.94% and 91.48% with *Sal. ruber* M31^T,

* Corresponding authors at: Marine Microbiology Group, Department of Animal and Microbial Biodiversity, Mediterranean Institute for Advanced Studies (IMEDEA, CSIC-UIB), E-07190, Esporles, Spain.
E-mail addresses: tviver@imedea.uib-csic.es (T. Viver), ramon@imedea.uib-csic.es (R. Rossello-Mora).

https://doi.org/10.1016/j.syapm.2023.126416
Received 23 January 2023; Revised 7 March 2023; Accepted 8 March 2023

respectively). *Sal. grassmerensis* sp. nov. name was formed according to the rules of the International Code for Nomenclature of Prokaryotes (ICNP), and *Sal. pepae*, *Sal. pampae* sp. nov. and *Sal. abyssi* sp. nov. are proposed following the rules of the newly published Code of Nomenclature of Prokaryotes Described from Sequence Data (SeqCode). This work constitutes an example on how classification under ICNP and SeqCode can coexist, and how the official naming a cultivated organism for which the deposit in public repositories is difficult finds an intermediate solution.

Introduction

Hypersaline environments such as manmade solar salterns constitute an excellent model system to study species composition and diversity, and the value of genetic diversity for ecosystem functioning (Conrad et al., 2022; Konstantinidis et al., 2022). The microbial communities thriving in hypersaline brines are generally dominated by *Archaea*, and mostly by the genus *Haloquadratum*, but *Bacteria* always coexist in lower abundance and are mostly represented by species of the genus *Salinibacter* (Antón et al., 2000; Ghai et al., 2011; Viver et al., 2019; Viver et al., 2021). This genus was the first member of the bacterial domain demonstrated to grow actively and in high abundance in hypersaline brines (Antón et al., 2000). *Salinibacter* also constituted an excellent example of a *Candidatus* taxon originally described using culture-independent techniques that could be brought to pure culture and formally named *Sal. ruber* (Antón et al., 2002) following the rules of the International Code for Nomenclature of Prokaryotes (ICNP; Oren et al., 2022), which requires the deposit of pure cultures as the designated material for the nomenclatural type. The isolation and classification of the first representatives of *Salinibacter* was followed by additional efforts to obtain bacterial cultures from other hypersaline systems that led to the description of two new species isolated from the Iranian Lake Aran-Bidgol, originally classified *Sal. iranicus* and *Sal. luteus* (Makhdoumi-Kakhki, et al., 2012). These species were later reclassified as a new genus *Salinivenus* (*Slv. iranica* and *lutea*; Munoz et al., 2016). More recently, the use of the high-throughput cultivation approach followed by MALDI-TOF (Matrix-Assisted Laser Desorption/Ionization Time-of-Flight) mass spectrometry in tandem with 16S rRNA gene sequencing allowed the further isolation and identification of thousands of new extremely halophilic strains, and the description of the species *Sal. altiplanensis* (Viver et al., 2015; Viver et al., 2018). *Salinibacter* is the type genus of the family *Salinibacteraceae* that, together with the genus *Salinivenus*, has been classified within the order *Rhodothermales*, class *Rhodothermia* and phylum *Rhodothermota* (Munoz et al., 2016). Biogeographic studies have demonstrated that the family *Salinibacteraceae* constitutes one of the most abundant and widely distributed bacterial lineages in hypersaline environments, with relatively high intra-family species richness (Viver et al., 2015; Mora-Ruiz et al., 2018; Conrad et al., 2022).

We recently reported a 220-isolate collection of *Salinibacter* from the solar salterns of Es Trenc in Mallorca (Conrad et al., 2022). A subset of 110 isolates were genome sequenced and the comparative analyses revealed that most of them (100 isolates) were members of the species *Sal. ruber* with a genome-aggregated average nucleotide identity (ANI) > 97% to the type strain M31^T genome. The remaining nine strains seemed to belong to a yet undescribed and co-occurring *Salinibacter* species showing ANI > 97% among themselves and < 95% with the abovementioned type strain (Conrad et al., 2022). To expand our knowledge on the global distribution of the genus *Salinibacter*, an international consortium of researchers, called Halophile Sequencing Project (HSP), was formed in 2019. The HSP consortium collected 27 hypersaline samples from 11 countries around the world with the aim to isolate and genome-sequence new strains and compare the resulting genomes with the metagenomic data of the same samples in order to assess biogeographical patterns.

In this study, we present the characterization and classification of two newly cultivated species of *Salinibacter*, and two additional yet-uncultivated species represented by the recovered metagenome assembled genomes (MAGs) from companion HSP metagenomes. As the information that can be currently retrieved from MAGs is comparable, if not even better, to what is published in many taxonomic papers of isolates (Konstantinidis et al., 2017), we opted to formally name both yet-uncultivated new species and one of the cultivated following the rules of newly published Code of Nomenclature of Prokaryotes Described from Sequence Data (SeqCode; Hedlund et al., 2022; Whitman et al., 2022). Complementing the study, we present the evaluation of metabolomics to discriminate among these species for taxonomic purposes. We propose the names *Salinibacter grassmerensis* sp. nov. which will be named following the rules of the ICNP (Oren et al., 2022), and *Salinibacter pepae* sp. nov., *Salinibacter pampae* sp. nov. and *Salinibacter abyssi* sp. nov. following the rules of the SeqCode.

Material and methods

Ionic composition of the samples in the study

To measure the ionic composition and concentration, brines samples were filtered through 0.22 μm hydrophilic PTFE filters, and the concentrations of fluoride (F^-), chloride (Cl^-), bromide (Br^-), nitrate (NO_3^-), sulfate (SO_4^{2-}), phosphate (PO_4^{3-}), sodium (Na^+), lithium (Li^+), potassium (K^+), ammonium (NH_4^+), magnesium (Mg^{2+}) and calcium (Ca^{2+}) were quantified by ion chromatography using a Metrohm, 850 ProfIC AnCat — MCS system, by Technical Research Services of Alicante University (Spain) (Table 1). The salinities of the samples were measured using a Sper Scientific Salt Refractometer (model number 300006, Arizona).

Sampling, isolation, strain dereplication and culture tests

All *Salinibacter* pure cultures included in this study were isolated using Sea Water medium (SW) at 25% salt concentration and supplemented with 0.05% (w/v) yeast extract (Rodriguez-Valera et al., 1985). The high-throughput cultivation approach was followed by the pure culture identification using the MALDI-TOF MS in tandem with 16S rRNA gene analysis (Viver et al., 2015). New isolates were stored in cryoprotectant solution containing 13% (w/v) glycerol at −80 °C. Physiological and biochemical tests were performed as detailed by Viver et al. (2018). Physiological tests were performed with SW broth or agar media. Liquid cultures were incubated at a shaking speed of 250 rpm and a temperature of 30 °C. Briefly, catalase, oxidase, anaerobic growth in the presence of arginine and DMSO; hydrolysis of Tween 20, Tween 80 and DNA, casein, gelatin and starch; arginine dihydrolase, lysine decarboxylase, tryptophanase and ornithine decarboxylase activities; methyl-red and Voges-Proskauer reactions; H_2S production, gas formation with nitrate were performed as detailed by Cowan and Steel's Manual for the Identification of Medical Bacteria (Barrow and Feltham, 1993). Temperature and pH growth range were measured in SW 25% supplemented with 0.1% yeast extract at 250 rpm. Salinity growth range was performed with SW media ranging between 10 and 35% salinities (using 5% salinity intervals). We tested the pH (5.5 to 8.5 with increments of 1 unit) and temperature (20 to 60 °C with increment of 10 °C) to assess the optimal growth conditions in a medium SW 25%. The pH of the medium was adjusted using the following buffering systems: pH 5.0–8.0,

Table 1
Sampling location, date and ionic composition (g/L) of the hypersaline sites in study. Salinity (%) given as measured using a Sper Scientific Salt Refractometer.

Location	Country	Sampling date	Coordinates	Salinity (%)	Na^+	Cl^-	Mg^{2+}	SO_4^{2-}	Ca^{2+}	Li^+	NH_4^+	NO_3^-	F^-	Br^-	PO_4^{3-}	K^+
S'Avall – Mallorca	Spain	September 2018	39°19′28″N – 2°59′21″E	36.2	62.73	180.48	36.28	41.87	0.59	0.003	0.0015	0.0678	0.0212	1.74	0	10.52
Es Trenc – Mallorca	Spain	September 2018	39°21′03″N – 3°00′45″E	34	68.01	176.71	27.65	20.72	0.36	0.002	0	0.0609	0.0142	1.36	0	8.31
Santa Pola – Alicante	Spain	May 2019	38°11′05″N – 00°35′08″W	37	76.47	217.06	40.01	60.11	0.04	0.005	0.02	0.024	0.0363	2.08	0.0075	12.23
Great Salt Lake – Utah	USA	July 2019	40°41′13″N – 112°22′33″W	37	106.10	188.56	17.78	51.4	0.23	0.086	0.0063	0.0597	0.0059	0.21	0.0006	10.68
Lake Grassmere - Marlborough	New Zealand	January 2019	41°44′33″S – 174°08′35″E	37	59.08	190.77	46.93	62.02	0.38	0.005	0.0081	0.0674	0.0251	2.14	0.0063	12.80
Colorada Chica – La Pampa	Argentina	August 2019	38°22′51″S – 63°25′41″O	33	113.28	189.27	11.14	27.49	0.49	0	0.0101	0.0562	0.0028	0.25	0.009	1.44
Colorada Grande – La Pampa	Argentina	August 2019	38°15′41″S – 63°45′02″O	32	119.49	196.05	6.49	13.9	0.55	0	0	0.0548	0.0027	0.14	0	1.43
Fără Fund	Romania	April 2019	45°52′34″N – 24°04′03″E	29	114.05	178.64	0.45	1.53	0.66	0	0.0026	0.0491	0.0022	0.01	0.0004	0.04

0.1 M citric acid/0.2 M Na_2HPO_4; pH 8.0–9.0, 0.1 M Tris/0.1 M HCl; pH 9.5–11.0, 0.05 M $NaHCO_3$/0.1 M NaOH (Gago et al., 2021). Acid production from carbohydrates (galactose, glucose and maltose) and carbon source utilization (tyrosine, tryptophan, asparagine, alanine, aspartate, proline, methionine, glucose, ribose, mannose, yeast extract, fructose, raffinose, galactose, cysteine, lactose, sucrose, maltose and lysine) were tested (Viver et al., 2018). pH was measured using a Mettler Toledo S20 SevenEasy™ tool (Schwerzenbach, Switzerland). Pigments were extracted with methanol/acetone (1:1, v/v) as an extraction solution. The extracts were cleared by centrifugation to remove turbidity. The absorption spectra were obtained using a HITACHI U-2900/2910 spectrophotometer (Tokyo, Japan). Cell morphology was examined using differential interference contrast (DIC, Nomarsky) microscopy (Zeiss Axio Imager optical microscope; Carl Zeiss, Jena, Germany). Cell motility was confirmed from an exponential phase SW 25% culture. Gram stains were prepared following the method of Dussault (1995).

Fatty acid profiles

Fatty acids were extracted from total cell lysates grown on SW medium at 25% salt concentration after one-month incubation at 30 °C, and fatty acid analysis was performed according to Kämpfer and Kroppenstedt (1996) by the separation of fatty acid methyl esters using gas chromatography (5898A, Hewlett Packard). Peaks were automatically integrated, and fatty acid names and percentages were determined with the Sherlock MIDI version 2.1 (TSBA version 4.1).

Metabolomics

To evaluate the metabolomic composition of the members of the genus, all new isolates, together with a subset of the 12 *Sal. ruber* strains M31^T and M8 (Anton et al., 2002), UD05, DZ71, SM10, CZ25, UM08, UM07, UD05, SM14, CZ31 and CM12 (Conrad et al., 2022) and the three isolates of *Sal. altiplanensis*, were grown under the same conditions using SW 25% amended with yeast extract 0.05%, in liquid culture and shaking at 30 °C for 21 days. All cultures were sampled at the exponential phase. The metabolomes of all isolates were generated following the standard published procedure (Rossello-Mora et al., 2008). In brief, cell biomass of 3 ml of cells growing at exponential phase was collected by centrifugation and the pellet was resuspended in 1 ml of sterile double-distilled water. The pellets were lysed using sonication and the lysate was acidified with 50 µl of 98–100% formic acid. The solution formed a precipitate that was separated from the soluble fraction by centrifugation. 500 µl of methanol were added to the pelleted precipitate, and the soluble acidified supernatant was chromatographed using Bond Elut C18 columns (Varian Inc., Lake Forest, CA, USA), as previously indicated (Rossello-Mora et al., 2008). The supernatant was considered the water-soluble fraction of the cell biomass, while the pellet was considered the water-insoluble and methanol soluble fraction.

For the analysis of the chemical composition, an FT-ICR-MS analysis was carried out using a solariX FT-ICR mass spectrometer (Bruker Daltonik GmbH, Germany) equipped with a 12 T superconducting magnet (Magnex Scientific Inc., GB) and an APOLLO II electrospray ionization source (Bruker Daltonik GmbH) in the negative ionization mode. Methanolic extracts of water-soluble and insoluble cellular fraction were injected using a microliter pump at a liquid flow rate of 120 µl h^{-1}. Both sheath gas and curtain gas consisted of nitrogen. A source heater temperature of 200 °C was maintained to ensure rapid solvent evaporation of the ionized droplets. Spectra were acquired with a time domain of four megawords and 400 scans were accumulated for each mass spectrum over an m/z mass range of 92.1 to 1500. The spectra were externally calibrated based on arginine clusters and systematically internally calibrated with appropriate reference mass list, reaching accuracy values lower than 100 ppb in

routine day-to-day measurements. Data acquisition and handling were performed using Data Analysis Software (v4.1; Bruker Daltonics). The assigned elemental composition was calculated based on the exact mass differences using a software tool written in-house (NETCALC) (Tziotis et al., 2011). Final formulas were generated and categorized into groups containing carbon, hydrogen and oxygen atoms (CHO, CHNO, CHOS or CHNOS), to reconstruct the group-selective mass spectra. Final formulas were generated and categorized into groups containing carbon, hydrogen, nitrogen, oxygen and sulfur atoms (CHO, CHNO, CHOS or CHNOS), to reconstruct the group-selective mass spectra. The Van Krevelen diagrams (Van Krevelen, 1950) were used to visualize the elemental ratios of unambiguously assigned molecular formulas.

The datasets were united-variance (UV) scaled. A classification model was applied, orthogonal partial least square discriminant analysis (OPLS-DA). The three different classes were considered in the models: *Sal. ruber, Sal. altiplanensis* and *Sal. pepae*. The total amount of variance absorbed by the first two components of the models were 83% for supernatant and 80% for the pellet. The index for the goodness for the prevision was 14% for pellet and 24% for supernatant. For both datasets, i.e., pellet and supernatant, the scores of the first principal component of the model were used, and the resulting data were clustered via hierarchical cluster analysis (HCA). For Pellet we used the single linkage to calculate the distances, instead for Supernatant the Ward method. The visualization could present how the samples could cluster together. The variables that characterized the different classes (that are more perturbed in the class) were selected ranking the loadings value. The highest values were considered and plotted in the van Krevelen visualization to define their chemical space. The masses that did not contribute to the characterization of the different classes were defined as core.

Genome analyses, phylogenetic reconstruction and probe design

The new strains isolated from Spain, USA and New Zealand were sequenced using Illumina NovaSeq (2 × 150 bp, paired-end reads) at Macrogen Company, South Korea. Raw reads were filtered using of bbduk v38.82 tool (https://bbtools.jgi.doe.gov) applying the following parameters: ktrim = r, k = 28, mink = 12, hdist = 1, tbo = t, tpe = t, qtrim = rl, trimq = 20 and minlength = 100. Trimmed reads were assembled using SPAdes v. 3.10 (Bankevich et al., 2012), followed by gene prediction from contigs > 500 bp using Prodigal v2.6.3 (Hyatt et al., 2010). A hybrid assembly methodology was used for the genome of the type strain ESAV49^{Ts}, sequenced by NovaSeq and MinION Nanopore, using Unicycler v0.4.8 tool (Wick et al., 2017).

Through barrnap v0.9 tool (https://github.com/tseemann/barrnap), 16S rRNA genes sequences were extracted from genome assembled contigs. The sequences were imported into the latest updated LTP_01_2022 database that contains all sequences of the type strains classified until January 2022 (Ludwig et al 2021) available at https://imedea.uib-csic.es/mmg/ltp/. Sequences were aligned using the SINA v1.3.1 aligner (Pruesse et al., 2007) implemented in the ARB program package (Ludwig et al., 2004), and manually checked in order to improve the alignment and to finally perform the phylogenetic analyses using ARB software package v6.0.6 (Ludwig et al., 2004). Phylogenetic analysis of the almost complete 16S rRNA gene sequences were performed to reconstruct *de novo* trees based on the neighbor-joining (Saitou and Nei, 1987) and RaxML (Stamatakis 2006) algorithms using different corrections (Jukes-Cantor and GTRGAMMA, respectively) as implemented in the ARB software package (Ludwig et al., 2004). For the probe design to identify the new species, we used the currently available SILVA SSU Ref99 138.1 database, curating it to the retain the best sequences of almost full length and high SILVA quality scores (Pruesse et al., 2007), and searching the best nucleotide sequences in the 16S rRNA gene that serve as phylogenetic probes using the Probe Search and Probe Match of the ARB program package (Ludwig et al., 2004).

Pangenome analysis at intra-species level was performed as detailed in Conrad et al., 2022. To compare the pangenome statistics between strains of *Sal. pepae* and *Sal. ruber*, we randomly selected 19 genomes from the 102 *Sal. ruber* collection (Conrad et al., 2022), and this subsampling was repeated 10 times for statistical reasons.

The predicted protein sequences of the genomes included in the study were used for phylogenetic reconstruction based on single-copy core genes. Predicted proteins were compared using an all-versus-all BLAST (v2.2.31; Camacho et al., 2009) and shared reciprocal best matches in all pairwise genome comparisons were identified using *rbm.rb* script (Rodriguez-R and Konstantinidis, 2016) using a 40% sequence identity cut-off and over 50% or more of the query sequence length. All single-copy core-genes were aligned individually using MUSCLE v3.8.31 (Edgar 2004). Aligned proteins were concatenated using the script *Aln.cat.rb* (Rodriguez-R and Konstantinidis, 2016) and the phylogenetic analyses were carried out using the RaxML and neighbor-joining algorithms implemented in the ARB v6.0.6 software (Ludwig et al., 2004). Both algorithms were used with the corrections PROTGAMMA (Lanave et al., 1984) and Kimura (Kimura, 1980), respectively. The average nucleotide identity (ANI) and average amino acid identity (AAI) between all the genomes based on the BLAST tool was calculated using *ani.rb* and *aai.rb* script (Rodriguez-R and Konstantinidis, 2016), respectively.

Protein-coding genes were annotated against SwissProt and TrEMBL databases (The UniProt Consortium, 2021) using Diamond v0.9.31 tool with default settings (Buchfink et al., 2021). Moreover, proteins were annotated using Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa et al., 2016). CRISPR-Cas genes were identified using CRISPRCasFinder tool v4.2.20 (Couvin et al., 2018) and the CRISPR spacers using the online tool CRISPRfinder (Grissa et al., 2007).

Metagenome analyses

DNA extraction from environmental samples was performed as detailed in Urdiain et al. (2008). The samples from 'Es Trenc' mesocosms study were previously reported in Conrad et al. 2022. Metagenomes from Spain, Argentina, New Zealand and Romania were sequenced using Illumina NextSeq500 (2 × 150 bp, paired end reads). Bbduk of the Bbtools package (https://bbtools.jgi.doe.gov/) was used to trim the paired-end reads with the following options: ktrim = r, k = 28, mink = 12, hdist = 1, tbo = t, tpe = t, qtrim = rl, trimq = 20 and minlength = 100. Trimmed reads were assembled with SPAdes v3.13.1 (Nurk et al., 2017) assembler, using metaSPAdes mode and MEGAHIT v1.2.9 (Li et al., 2016) with default parameters. Sequences of 16S rRNA genes were extracted from the assembled metagenome contigs using barrnap v0.9 tool (https://github.com/tseemann/barrnap). Contigs > 2,000 bp and 5,000 bp were binned using MetaBAT v2 software (Kang et al., 2019). MAGs characteristics, including completeness and contamination, were calculated using MiGA tool (Microbial Genome Atlas; Rodriguez-R et al., 2018). For MAG phylogenetic identification and protein-coding genes annotation, we followed the same methodology specified in genome analysis.

Trimmed metagenomic reads were mapped to the type strain genomes and MAGs using BLASTn tool v2.2.31 (Camacho et al., 2009). A competitive read recruitment analysis was applied to calculate the difference in sequencing depth and thus, abundance between different species. For a confident read assignment, mapped reads with a tied best-hit blast match were removed from the analysis, emphasizing the unique sequence components present in the metagenomes using the script *01_MagicBlast_ShortRead-Filter.py* (Conrad et al., 2022). Recruitment were plotted using the "enveomics.R" package v1.5 (Rodriguez-R and Konstantinidis, 2016).

Genome and metagenome sequence submission

The genome sequences of the strains CM04, CZ16, CZ17, CZ26, CZ33, DZ03, DZ04, UM13 and UZ07 and the metagenomes from the 'Es Trenc' mesocosms experiment were deposited under the European Nucleotide Archive (ENA) accession code PRJEB27680 (Conrad et al., 2022). The genomes of the strains ESAV49^{Ts}, ESAV87, ESCM25, ESSP12, ESSP73, ESSP84, USCM25, USCM46, USCM88, USCM187 and NZ140^{T}; the MAGs ARCCH^{Ts}, ARCG and ROFF^{Ts}; and the metagenomes from Spain, Argentina, New Zealand and Romania, were deposited under the accession code PRJEB45291. The strains, genomes and MAGs used in this study with their accession numbers are listed in Table 2 and Sup. Spreadsheet S1.

Results and discussion

Isolation, genome binning and 16S rRNA gene phylogenetic reconstruction

From the 27 HSP samples collected from 23 different sites, we isolated 1,274 pure cultures of extreme halophiles using SW 25% and 0.05% yeast extract as substrate. From this collection, 794 isolates were identified as members of the genus *Salinibacter* using of MALDI-TOF MS and 16S rRNA gene sequencing tandem approach. The *Salinibacter* isolates were obtained from 11 samples, and from eight of them (Table 1), we obtained representatives of the new species reported here (Table 2). From the remaining 12 sites, none of the 480 pure cultures isolated could be identified as *Salinibacter* (unpublished data). The eight hypersaline sites used in the study here represented six single locations around the world (the Mallorca salterns and the Pampean lakes are geographically clustered), and the longest distance was between Santa Pola (Spain) and Lake Grassmere (New Zealand) with ~ 19,515 Km away, while the shortest distance was between Santa Pola and Mallorca, with ~ 350 Km (Sup. Figure S1).

Among the 794 *Salinibacter* isolates, 10 formed a unique lineage in the 16S rRNA gene phylogenetic reconstruction that was distinct from the *Sal. ruber* and *Sal. altiplanensis* (Fig. 1A). The 10 isolates, ESAV49^{Ts} and ESAV87 (from S'Avall, Mallorca, Spain), ESCM71 (from Es Trenc, Mallorca, Spain), ESSP12, ESSP73 and ESSP87 (from Santa Pola, Alicante, Spain), and USCM187, USCM25, USCM46 and USCM88 (from Great Salt Lake; Utah, USA), were closely related to nine isolates obtained in 2012 from the Es Trenc solar salterns in Mallorca, CM04, CZ16, CZ17, CZ26, CZ33, DZ03, DZ04, UM13 and UZ07 (Conrad et al., 2022). Altogether, the 19 isolates (Table 2) formed a distinct monophyletic group with a 16S rRNA gene identity > 99.7% that represents a new species within the genus. This new species had 16S rRNA gene identities that ranged from < 98.6% and < 97.2% (Sup. Spreadsheet S2) with the type strains of *Sal. ruber* (M31^{T}) and *Sal. altiplanensis* (AN15^{T}), respectively; these values are below the threshold that generally requires whole genome comparisons to classify new species (Stackebrandt and Ebers, 2006). In addition, the single *Salinibacter* isolate, NZ140^{T}, from Lake Grassmere (New Zealand), represented a distinct lineage within the genus (Fig. 1A), with 16S rRNA gene identities of 98.7% and < 96.7% with the type strains of *Sal. ruber* (M31^{T}) and *Sal. altiplanensis* (AN15^{T}), respectively.

The selection of the representative MAGs recovered from each metagenome using different bioinformatic tools is detailed in Sup. Text S1 and Sup. Table S1. From the *Salinibacter* MAGs, only one MAG from each of the hypersaline lakes "Laguna Colorada Chica" (ARCCH^{Ts}) and "Laguna Colorada Grande" (ARCG), both located in La Pampa province of Argentina, and Fără Fund Lake (ROFF^{Ts}) of the Transylvania region of Romania, and also formed independent lineages with 16S rRNA gene identities < 98.3% with any of the cultivated species (Fig. 1).

Table 2
Main features of the new strains and MAGs affiliated with *Salinibacter*. Bold type letter corresponds to type strains or type MAGs of the species.

Location	Genome	Accession number	Completeness (%)	Contamination (%)	Contigs	Size (bp)	N50 (bp)	%GC	CDS	rRNA	tRNA
S'Avall – Mallorca – Spain	***Salinibacter pepae* ESAV49^{Ts}**	**GCA_947077775^{Ts}**	**99**	**2**	**2**	**3,625,015**	**3,585,205**	**66.43**	**3,040**	**23S/5S/16S**	**58**
	Salinibacter pepae ESAV87	GCA_947077725	99	2.9	60	3,759,744	125,197	66.17	3,179	23S/5S/16S	57
Es Trenc – Mallorca – Spain	*Salinibacter pepae* ESCM71	GCA_947077735	99	2.9	39	3,628,964	145,479	66.39	3,041	23S/5S/16S	55
Santa Pola – Alicante – Spain	*Salinibacter pepae* ESSP12	GCA_947077745	99	2	45	3,628,964	166,466	66.41	3,064	23S/5S/16S	56
	Salinibacter pepae ESSP73	GCA_947077695	99	2	61	3,539,819	121,276	66.65	2,974	23S/5S/16S	56
	Salinibacter pepae ESSP87	GCA_947077805	99	2.9	67	3,525,093	121,901	66.68	2,980	23S/5S/16S	55
Great Salt Lake – Utah – USA	*Salinibacter pepae* USCM187	GCA_947077785	99	2	41	3,740,553	151,371	66.12	3,147	23S/5S/16S	56
	Salinibacter pepae USCM25	GCA_947077825	99	2	63	3,591,426	90,953	66.37	3,084	23S/5S/16S	55
	Salinibacter pepae USCM46	GCA_947077795	99	2	40	3,628,149	164,188	66.44	3,049	23S/5S/16S	56
	Salinibacter pepae USCM88	GCA_947077755	99	2	44	3,635,305	157,762	66.41	3,062	23S/5S/16S	56
Es Trenc– Mallorca – Spain	*Salinibacter pepae* CM04	GCA_905335865	94.1	2	598	3,615,056	8,948	66.17	3,371	23S/5S/16S	53
	Salinibacter pepae CZ16	GCA_905335995	93.1	2	666	3,474,286	7,090	66.01	3,348	23S/5S/16S	55
	Salinibacter pepae CZ17	GCA_905336005	95.1	2	553	3,474,973	8,797	66.40	3,254	23S/5S/16S	53
	Salinibacter pepae CZ26	GCA_905336055	95.1	2	513	3,630,762	10,322	65.88	3,415	23S/5S/16S	52
	Salinibacter pepae CZ33	GCA_905336145	94.1	2	620	3,626,645	8,665	65.77	3,463	23S/5S/16S	52
	Salinibacter pepae DZ03	GCA_905336415	95.1	2	527	3,509,153	9,586	66.27	3,278	23S/5S/16S	55
	Salinibacter pepae DZ04	GCA_905336425	97.1	2	469	3,462,861	11,103	66.44	3,196	23S/5S/16S	54
	Salinibacter pepae UM13	GCA_905336735	96.1	2	566	3,683,449	9,981	65.78	3,492	23S/5S/16S	53
	Salinibacter pepae UZ07	GCA_905336925	96.1	2	562	3,614,727	9,352	66.08	3,391	23S/5S/16S	53
Lake Grassmere - New Zealand	***Salinibacter grassmerensis* NZ140^{T}**	**GCA_947077765**	**99**	**3.9**	**19**	**3,781,529**	**659,197**	**64.09**	**3,176**	**23S/5S/16S**	**56**
Fără Fund – Romania	***Salinibacter abyssi* ROFF^{Ts}**	**GCA_947077815^{Ts}**	**90.2**	**2.9**	**279**	**3,018,448**	**16,630**	**66.1**	**2,705**	**23S/5S/16S**	**47**
La Pampa – Argentina	***Salinibacter pampae* ARCCH^{Ts}**	**GCA_947077715^{Ts}**	**98**	**2.9**	**225**	**3,092,045**	**20,253**	**65.53**	**2,837**	**23S/5S/16S**	**54**
	Salinibacter pampae ARCG	GCA_947077705	96	3	278	3,147,671	16,602	65.7	2,903	16S	51

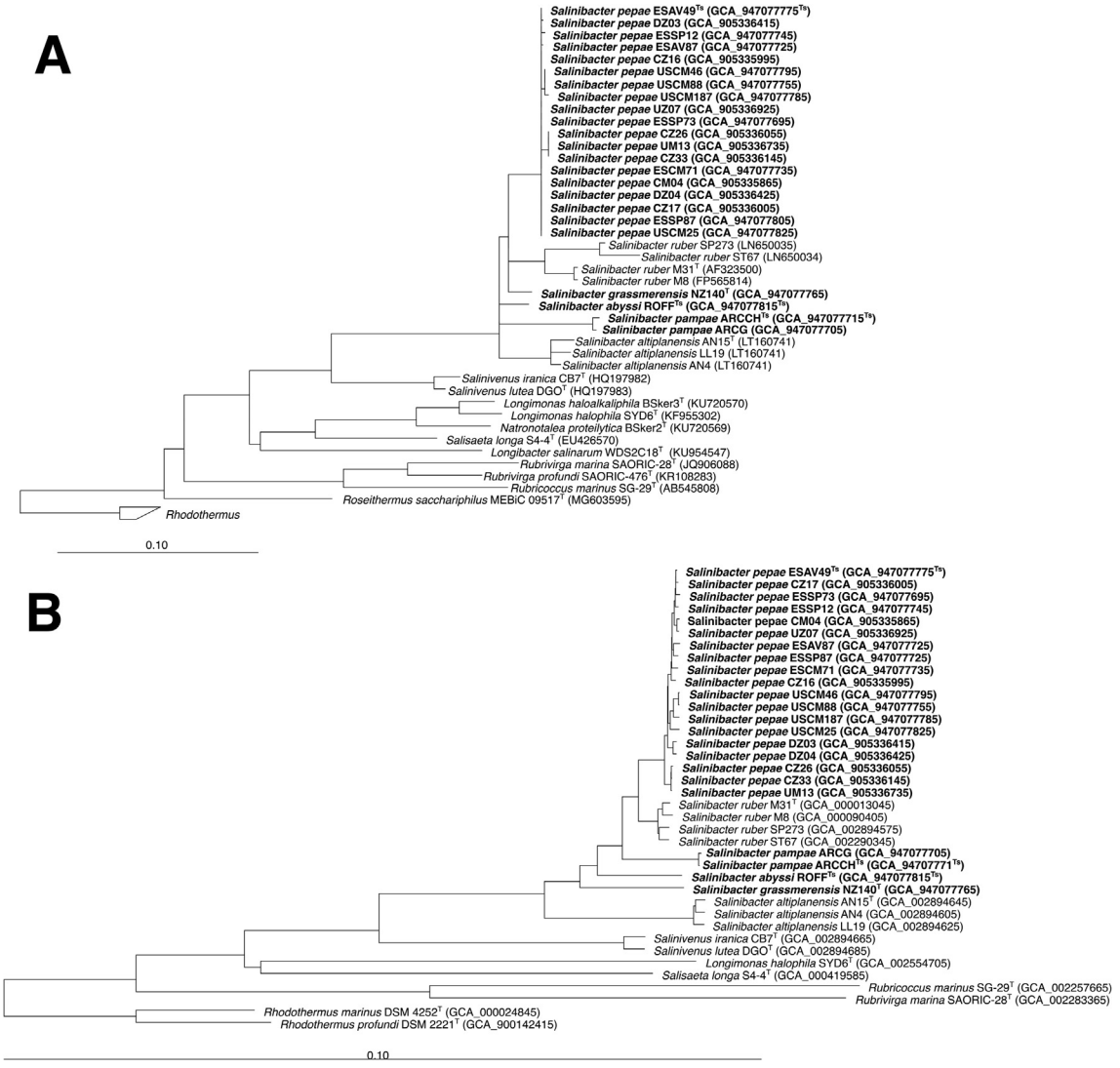

Fig. 1. (A) Phylogenetic reconstruction based on the 16S rRNA gene sequence of all *Salinibacter* species available in the LTP_01_2022 and the MAGs recovered from metagenomes. The tree was reconstructed using the maximum likelihood algorithm and is the result of the consensus of different approaches using distinct filters and datasets. The multifurcations indicate a branching order that could not be resolved. Bar indicates 10% sequence divergence. In brackets the accession number of each sequence is given. (B) Phylogenetic reconstruction based on concatenated 592 core orthologous gene sequences shared between all genomes and MAGs included in the analysis. The tree was reconstructed using the maximum likelihood algorithm. Bar indicates 10% sequence divergence. In brackets the accession number of each sequence is given.

Proposal of new Salinibacter species names

The isolate genomes and MAGs (Table 2 and Table 3), together with the phenotypic characterization of the isolates allowed the description of four new species (see below). Since naming new taxa under the ICNP is only possible for pure cultures deposited in culture collections without distribution restrictions (Oren et al., 2022), we named the one of the newly cultivated species using this bacteriological code (ICNP). In addition, based on the newly published SeqCode, that allows designating genomes such as MAGs as nomenclatural types for novel species (Hedlund et al., 2022), we name two new species based on MAGs, and one cultivated for which the deposit in a second collection was hindered by the fastidious nature of the isolate. The strain ESAV49^{Ts} was deposited at the Spanish culture collection (CECT) with success, but the attempts to obtain a deposit number from a second collection was delayed for at least one year, and in order not to delay the publication of the new descriptions we decided to name the new species for which this strain is designated type under the SeqCode. Thus, for one cultivated species we propose the name *Salinibacter grassmerensis* sp. nov. (named after the Lake Grassmere location of its isolation, with the designated type strain NZ140^{T} = CECT 30523^{T} = ICMP 24464^{T}). For the second cultivated, but yet unsuccessful deposit in a second culture collection we propose *Salinibacter pepae* sp. nov. (named after the discoverer of the first *Salinibacter* organism Professor Pepa Antón, with the designated type strain ESAV49^{Ts} = CECT 30522^{Ts}), and for the two MAGs, we propose *Salinibacter pampae* sp. nov. (named after the grassland plain in South America, referring to the Pampa region in Argentina, with the sequence accession number GCA_947077715^{Ts} = ARCCH^{Ts}) and *Salinibacter abyssi* sp. nov. (named after the belief that the lake Fără Fund was bottomless, with the type sequence accession number GCA_947077815^{Ts} = ROFF^{Ts}). The formal protologues are given in Table 4. We expect that in the course of the months after this contribution has been published, the deposit of the strain ESAV49^{Ts} is finally successful. After the second deposit we will append a corrigendum to this manuscript to definitively validly publish the name also under the ICNP.

Table 3
Pairwise genomic comparison of the genomes at the nucleotide level (ANI, lower triangle) and protein translated genes (AAI, upper triangle) between members of the *Salinibacteraceae* family. In the genome comparisons that the number of genes shared was lower than 20% of the genomes the ANI value was not calculated (ND = Not Determined). *Salinibacter pepae* ESAV49^{Ts}, ESAV87, ESCM71, ESSP12, ESSP73, ESSP87, CM04, CZ16, CZ17, CZ26, CZ33, DZ03, DZ04, UM13, UZ07, USCM187, USCM25, USCM46 and USCM88; *Salinibacter ruber* M31^{T}; *Salinibacter pampae* ARCCH^{Ts} and ARCG; *Salinibacter abyssi* ROFF^{Ts}; *Salinibacter grassmerensis* NZ140^{T}; *Salinibacter altiplanensis* AN15^{T}; *Salinivenus iranica* CB7^{T}; *Salinivenus lutea* DGO^{T}.

Specific name	Strain	ESAV49^{Ts}	ESAV87	ESCM71	ESSP12	ESSP73	ESSP87	CM04	CZ16	CZ17	CZ26	CZ33	DZ03	DZ04	UM13	UZ07	USCM187	USCM25	USCM46	USCM88	M31^{T}	ARCCH^{Ts}	ARCG	ROFF^{Ts}	NZ140^{T}	AN15^{T}	CB7^{T}	DGO^{T}
Sal. pepae	ESAV49^{Ts}		98.08	98.15	98.71	99.24	98.30	98.58	97.83	98.68	97.21	97.33	97.62	97.58	97.30	98.53	97.36	97.38	97.6	97.65	94.49	91.99	89.69	89.17	87.86	84.36	70.10	69.89
	ESAV87	99.01		97.99	97.96	98.10	98.21	97.87	97.48	97.85	96.93	97.00	97.80	97.75	96.99	97.90	97.06	97.50	97.48	97.42	94.46	92.00	89.32	88.79	87.65	84.31	69.86	69.78
	ESCM71	99.07	99.02		98.22	98.33	98.12	97.73	97.95	98.17	97.28	97.33	97.65	97.67	97.29	97.54	97.12	97.31	97.40	97.35	94.24	92.06	89.74	89.00	87.76	84.33	69.91	70.11
	ESSP12	99.50	98.97	99.05		98.69	98.24	98.32	97.57	98.63	97.04	97.17	97.53	97.62	97.13	98.53	97.14	97.68	97.54	97.51	94.48	92.06	89.66	88.75	87.67	84.30	70.18	69.98
	ESSP73	99.70	98.90	98.98	99.37		98.62	98.84	97.92	98.88	97.73	97.89	97.79	97.87	97.78	98.54	97.50	97.87	97.61	97.62	94.61	92.21	89.90	89.18	88.08	84.56	70.29	70.11
	ESSP87	99.09	98.97	99.15	99.10	99.02		98.57	98.20	98.32	97.37	97.44	97.73	97.93	97.41	98.24	97.55	97.75	97.53	97.52	94.63	92.13	89.98	89.07	88.16	84.51	70.27	70.05
	CM04	99.39	98.83	98.99	99.37	99.33	99.17		97.88	98.56	97.14	97.18	97.60	97.66	97.11	98.85	97.04	97.38	97.31	97.29	94.41	92.33	89.44	88.60	87.85	84.33	70.5	69.99
	CZ16	99.07	99.07	99.06	99.01	98.98	99.09	99.02		98.12	97.59	97.54	97.61	97.65	97.58	97.48	96.93	97.14	97.49	97.44	94.03	92.03	89.46	88.27	88.17	84.35	70.24	70.23
	CZ17	99.41	98.86	99.02	99.37	99.34	99.13	99.30	99.04		97.41	97.42	97.80	98.07	97.40	98.33	97.35	97.59	97.77	97.70	94.40	92.06	89.67	89.02	87.99	84.53	70.30	70.11
	CZ26	98.43	98.39	98.50	98.40	98.45	98.42	98.40	98.48	98.45		99.87	97.88	97.77	99.86	96.96	97.21	97.33	97.41	97.40	94.54	92.08	89.65	88.77	80.00	84.30	70.14	70.25
	CZ33	98.47	98.38	98.49	98.39	98.49	98.44	98.43	98.50	98.44	99.99		97.93	97.83	99.91	97.07	97.24	97.35	97.50	97.42	94.54	92.30	89.43	88.64	87.92	84.36	70.11	70.23
	DZ03	98.55	98.56	98.59	98.58	98.56	98.64	98.55	98.66	98.58	98.68	98.69		99.29	97.88	97.37	97.63	97.85	97.97	97.93	94.08	92.25	89.79	89.03	87.90	84.58	80.03	70.11
	DZ04	98.60	98.54	98.64	98.62	98.59	98.71	98.60	98.68	98.65	98.63	98.64	99.59		97.83	97.51	97.71	97.78	97.92	97.88	94.17	92.02	90.07	89.07	88.03	84.74	70.13	70.04
	UM13	98.48	98.40	98.48	98.43	98.50	98.44	98.41	98.50	98.47	99.99	99.99	98.67	98.68		97.04	97.28	97.39	97.46	97.39	94.40	92.11	89.41	88.50	87.88	84.28	70.21	70.24
	UZ07	99.36	98.86	98.87	99.31	99.21	99.04	99.65	98.91	99.21	98.30	98.29	98.45	98.50	98.32		96.81	97.01	97.10	97.10	94.40	92.12	89.26	88.35	87.63	84.06	70.33	70.14
	USCM187	98.40	98.45	98.42	98.35	98.31	98.40	98.26	98.44	98.34	98.32	98.33	98.58	98.54	98.33	98.15		97.64	98.42	98.44	93.90	91.98	89.43	88.85	87.65	84.25	69.97	69.61
	USCM25	98.48	98.53	98.56	98.44	98.45	98.50	98.41	98.53	98.46	98.51	98.55	98.70	98.67	98.55	98.32	98.69		97.97	97.98	94.24	92.19	89.44	88.64	87.81	84.22	70.25	69.93
	USCM46	98.42	98.46	98.42	98.37	98.32	98.44	98.26	98.46	98.29	98.34	98.33	98.59	98.54	98.32	98.23	99.03	98.73		99.78	94.42	91.99	90.02	88.74	88.10	84.76	70.06	69.83
	USCM88	98.43	98.43	98.42	98.38	98.33	98.45	98.28	98.45	98.31	98.36	98.36	98.58	98.55	98.40	98.20	99.06	98.70	99.84		94.39	92.20	89.94	88.56	88.04	84.70	70.01	69.84
Sal. ruber	M31^{T}	94.96	94.90	94.90	94.98	94.97	94.93	95.06	94.84	95.03	95.09	95.09	94.61	94.87	95.08	95.11	94.72	94.82	94.73	94.74		92.27	89.77	90.00	88.41	84.58	69.90	58.65
Sal. pampae	ARCCH^{Ts}	90.83	90.83	90.86	90.00	90.89	90.86	90.98	91.03	90.89	91.05	91.06	90.98	90.91	91.01	90.88	90.84	90.89	90.91	90.87	90.94		87.89	88.16	88.50	84.99	70.39	70.18
	ARCG	90.72	90.68	90.71	90.67	90.76	90.76	90.78	90.75	90.72	90.84	90.83	90.8	90.79	90.83	90.72	90.65	90.74	90.76	90.72	90.91	99.39		88.16	88.5	84.99	70.39	70.18
Sal. abyssi	ROFF^{Ts}	90.18	90.14	90.20	90.20	90.28	90.12	90.39	90.25	90.23	90.17	90.30	90.16	90.29	90.39	90.34	90.10	90.24	90.17	90.13	91.24	88.11	87.87		86.27	83.04	69.46	69.27
Sal. grassmerensis	NZ140^{T}	86.68	86.63	86.62	86.66	86.70	86.65	86.78	86.80	86.78	86.77	86.77	86.76	86.74	86.86	84.77	86.58	84.64	86.59	86.64	87.08	86.23	86.11	86.45		83.93	69.53	69.39
Sal. altiplanensis	AN15^{T}	84.01	84.09	84.02	84.03	84.09	84.09	84.27	84.35	84.23	84.20	84.26	84.18	84.22	84.3	84.25	84.01	83.97	84.00	84.01	84.17	83.80	83.71	83.88	83.41		69.58	69.33
Slv. iranica	CB7^{T}	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D		95.56
Slv. lutea	DGO^{T}	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	N D	

Genotype of the cultivated species

The genomes of all isolates were sequenced and deposited in public repositories under the accession numbers given in Table 2. The tree reconstruction based on the 592 orthologous genes that formed the core genome of the families *Salinibacteraceae* and *Salisaetaceae*, and the genomes of *Rhodothermus marinus* DSM 4252^{T} and *Rhodothermus profundis* DSM 2221^{T} (Fig. 1B) showed high agreement with that based on the 16S rRNA gene (Fig. 1A), with only a minor discrepancy on the position of *Sal. grassmerensis* that showed the basal position within the genus. The hierarchical clustering based on the presence or absence of orthologous groups and that observed on AAI clustering showed a similar cluster distribution (Sup. Figure S2 and S3) that mirrored the reconstructed phylogenies (Fig. 1).

The isolates of *Sal. pepae* showed ANI values > 97% among themselves. In addition, *Sal. ruber* M31^{T} showed ANI values < 95.09% and 87.08% and a shared genomic fraction < 72.3% and 53.7% with *Sal. pepae* and *Sal. grassmerensis,* respectively. Moreover, *Sal. altiplanensis* AN15^{T} showed ANI values of < 84.3% and 86.8% and a shared genomic fraction < 40.7% and 38% with the same species, respectively (Table 3, Fig. 2, Sup. Figure S2 and Sup. Spreadsheet S3 and S4). These results confirmed that, genomically, the classification of both new species was justified (Rosselló-Móra and Amann, 2015). The genomic characteristics observed for the genomes of the new species were in line with those of the remaining *Salinibacter* species (Viver et al., 2018), with genome sizes ranging between 3.46 Mbp (for DZ04) and 3.78 Mbp (for NZ140^{T}), and with an average genome size of 3.59 Mbp (SD = 83,767 bp). The genome of *Sal. pepae* ESAV49^{Ts} was fully sequenced and consisted of a circular chromosome of 3,585,205 bp and a small plasmid of 39,810 bp. *Sal. grassmerensis* NZ140^{T} showed the lowest G + C mol% value (64.09%) among all members of the family *Salinibacteraceae* (Table 2). The number of predicted coding Sequences (CDS) was similar among strains, with an average of 3,200 CDS (SD = 162.1), while the average number of annotated tRNAs was 54.7 (SD = 1.65) (Table 2).

The pangenome of *Sal. ruber* was considered open and highly diverse, probably conferring ecological advantages to this species as evidenced, for instance, by the fact that the pangenome based on isolates recovered from a single location was of a similar size to that of a similar number of *Escherichia coli* genome collected worldwide (Conrad et al., 2022). Species with an open pangenome encode a larger number of dispensable genes, increasing their adaptability to the environment (or changing environmental conditions) due to the availability of an increased number of biochemical pathways (Medini et al., 2005). To compare *Sal. ruber* pangenome size with that of *Sal. pepae*, we randomly selected 19 genomes from the 102 *Sal. ruber* collection (Conrad et al., 2022), and this subsampling was repeated 10 times for statistical reasons (Sup. Table S2). Both pangenomes were open, but that of *Sal. ruber* was consistently larger ($\gamma > 0.35$, average of 0.38 and SD = 0.014, the γ parameter that reflects the slope of the curve that represents the total nonredundant genes; Tettelin et al., 2005; Tettelin et al., 2008) than for *Sal. pepae* ($\gamma = 0.30$) (Fig. 3). Consequently, the pangenome of the 19 *Sal. ruber* genomes consisted of 7,509 genes on average (SD = 262) and that of *Sal. pepae* of 6,052 genes. Members of *Sal. ruber* encoded a larger number of strain-specific genes (3,316 on average and a mean addition of 279 new genes per genome included) than members of *Sal. pepae* (2,133 genes and a mean addition of 114 new genes per genome included), but the number of core genes was similar for both species (1,943 on average in *Sal. ruber* and 1,979 in members of *Sal. pepae*). As observed in our results (see below), both species coexisted, but *Sal. ruber* always occurred in abundances three times higher than *Sal. pepae*. One reason for this dominance could reside in the larger *Sal. ruber* pangenome, that would confer ecological advantages and/or adaptations to a wider number of niches (Tettelin et al., 2008). The search for autapomorphic characters (unique to the taxon) among the 19 members of *Sal. pepae* rendered 81 species-specific genes. From them, 9 could not be annotated (no match identified using TREMBL, UniProt and KEGG databases) and 13 corresponded to uncharacterized proteins. Among the annotated proteins with a known function, we identified glycosylhydrolases and transferases, Na^{+} solute symporters, ABC transporters, transcriptional regulators and integrases (Sup. Spreadsheet S5). However, none of these functions could be associated with an explicit phenotype that could serve as diagnostic for the species.

Single-strain species descriptions (SSSDs) preclude the identification of intraspecific diversity. Here, unfortunately among the 189 *Salinibacter* isolates from Lake Grassmere and 1,274 *Salinibacter* isolates worldwide, *Sal. grassmerensis* was just represented by a single strain, NZ140^{T}. *Sal. grassmerensis* NZ140^{T} encoded for 725 species-specific genes, 165 of which could not be annotated (no match identified using TREMBL, UniProt and KEGG databases) and 225 of which were annotated as uncharacterized proteins. As relevant features, we identified 25 genes annotated as glycosyltransferases and 16 of them were found in a cluster (genomic island) of 36 consecutive genes (between gene ID 164 to 200, Sup. Spreadsheet S6). In addition, we identified genes annotated as methyl- and sulfotransferases, ABC

Table 4
Protologues of the new species descriptions.

Species name	*Salinibacter pepae*	*Salinibacter grassmerensis*	*Salinibacter pampae*	*Salinibacter abyssi*
Specific epithet	*pepae*	*grassmerensis*	*pampae*	*abyssi*
Species status	sp. nov.	sp. nov.	sp. nov.	sp. nov.
Species etymology	pe'pae. N.L. gen. n. adj. pepae named after the microbiologist Pepa Antón expert in bacteria and viruses in extremophilic sites and who discovered the existence of the first *Salinibacter*	grass.mer.en'sis. N.L. masc. adj. *grassmerensis,* of the lake Grassmere located in New Zealand	pam'pae. N.L. gen. n. *pampae,* of the pampa, the grassland plain in South America, referring to the Pampa region in Argentina.	a.bys'si. L. gen. n. *abyssi*, of a bottomless pit, referring to lake Fără Fund ('without bottom')
Description of the new taxon and diagnostic traits	Straight rod cells, 3.0–6.0 μm long, forming red colonies after 15 days growth on SW agar media at 25% of salts at 30 °C. Colonies are circular and convex with an entire margin and with a diameter of 0.5–1.0 mm. Cells are flagellar and motile. Cells exhibit growth in the ranges of 15–34% salt concentration, optimum temperature at 30 °C and pH 7. The organism is positive in catalase, oxidase, Tween20, Tween80 and lysine decarboxylase. The organism is negative in indole, methyl-red, Voges-Proskauer, casein, DNA, Starch and gelatin hydrolysis, H_2S and nitrate production, acid production from carbohydrates, anaerobic growth in presence of arginine and DMSO, ornithine and adenine decarboxylase. The major fatty acids were $C_{15:0}$ iso, $C_{18:1}$ ω7*c* and summed feature 3†.	Slightly curved rod cells, 2.0–4.0 μm long, forming red colonies after 15 days growth on SW agar media at 25% of salts at 30 °C. Colonies are circular and convex with an entire margin and with a diameter of 0.5–1.0 mm. Cells are flagellar and motile. Cells exhibit growth in the ranges of 15–34% salt concentration, optimum temperature at 30 °C and pH 7. The organism is positive in catalase, oxidase, Tween20, Tween80 and lysine decarboxylase and hydrolysis of Starch. The organism is negative in indole, methyl-red, Voges-Proskauer, casein, DNA, and gelatin hydrolysis, H_2S and nitrate production, acid production from carbohydrates, anaerobic growth in presence of arginine and DMSO, ornithine and adenine decarboxylase. The major fatty acids were $C_{15:0}$ iso, $C_{18:1}$ ω7*c* and summed feature 3†.	The MAG encoded for oxidase, catalase, lysine decarboxylase and hydrolysis of starch. Genes could not be detected for nitrate reductase or the genes for the complete denitrification pathway, ornithine and adenine decarboxylase, arginine dihydrolase, and therefore such traits were considered negative. The MAG encoded for genes for flagella assembly, indicating putative motility.	The MAG encoded for oxidase, catalase, lysine decarboxylase and hydrolysis of starch. Genes could not be detected for nitrate reductase or the genes for the complete denitrification pathway, ornithine and adenine decarboxylase, arginine dihydrolase, and therefore such traits were considered negative. The MAG encoded the genes for flagella assembly, indicating motility.
Country of origin	Spain	New Zealand	Argentina	Romania
Region of origin	Mallorca Island	Marlborough	La pampa	Transylvania
Date of isolation (dd/mm/yyyy)	30/11/2018	20/04/2020	26/08/2019	01/04/2019
Source of isolation	Brines	Brines	Brines	Brines
Samping date (dd/mm/yyyy)	22/08/2018	24/01/2020	22/08/2019	01/04/2019
Latitude (xx°xx′xx″ N/S)	39°19′26″N	41°44′12.3″S	38°22′47.2″S	45°52′34″N
Longitude (xx°xx′xx″ E/W)	2°59′22″E	174°10′07.5″E	63°25′42.6″O	24°04′03″E
Altitude (meters above sea level)	0	0	15	420
16S rRNA gene accession nr.	–	–	–	–
Genome accession number [RefSeq; EMBL; …]	GCA_947077775[Ts]	GCA_947077765	GCA_947077715[Ts]	GCA_947077815[Ts]
Genome status	Closed	Draft – 99% completeness	Draft – 98% completeness	Draft – 90.2% completeness
Genome size	3.46	3.78	3.09	3.02
G + C mol%	66.43	64.09	65.53	66.1
Number of strains in study	19	1	–	–
Number of MAGs in study	–	–	2	1
Source of isolation of non-type strain	Brines	Brines	Brines	Brines
Information related to the Nagoya Protocol	Permit ABSCH-IRCC-ES-241224–1	–	Permit from the Ministerio de la Producción given to M. E. Llames on Aug 13st, 2019	–
Designation of the Type Strain or genomic assembly	ESAV49[Ts]	NZ140[T]	ROFF[Ts]	ARCCH[Ts]
Strain Collection Numbers	CECT 30522[Ts]	CECT 30523[T] = ICMP 24464[T]		
Metagenomic Raw Data			SAMEA12118482	SAMEA111452366
SeqCode registry URL	seqco.de/i:24081		seqco.de/i:23671	seqco.de/i:23670

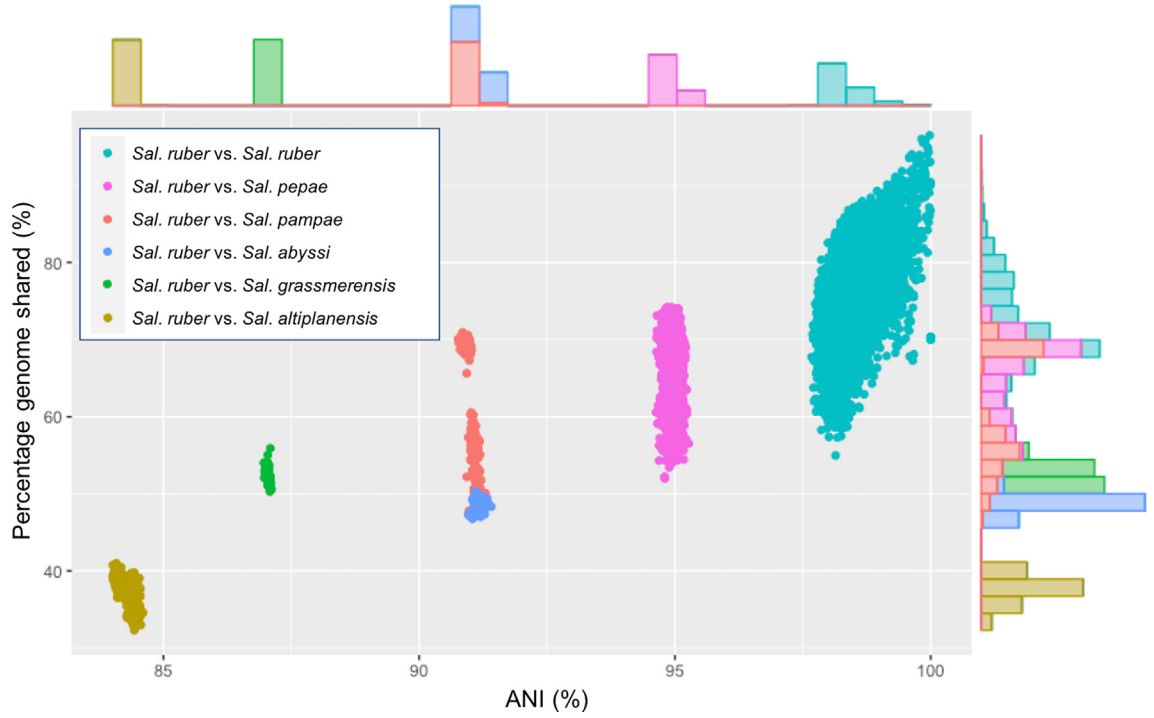

Fig. 2. Genomic diversity among *Salinibacter* species assessed by ANI and shared genome fraction. The shared genome fraction (y-axis) is plotted against ANI (x-axis) of pairwise compared genomes between strains affiliated to *Salinibacter* ruber and the other species belonging to *Salinibacter* genus. The graphs to the top and right of the panel show the kernel density estimates for each axis (distribution of the datapoints).

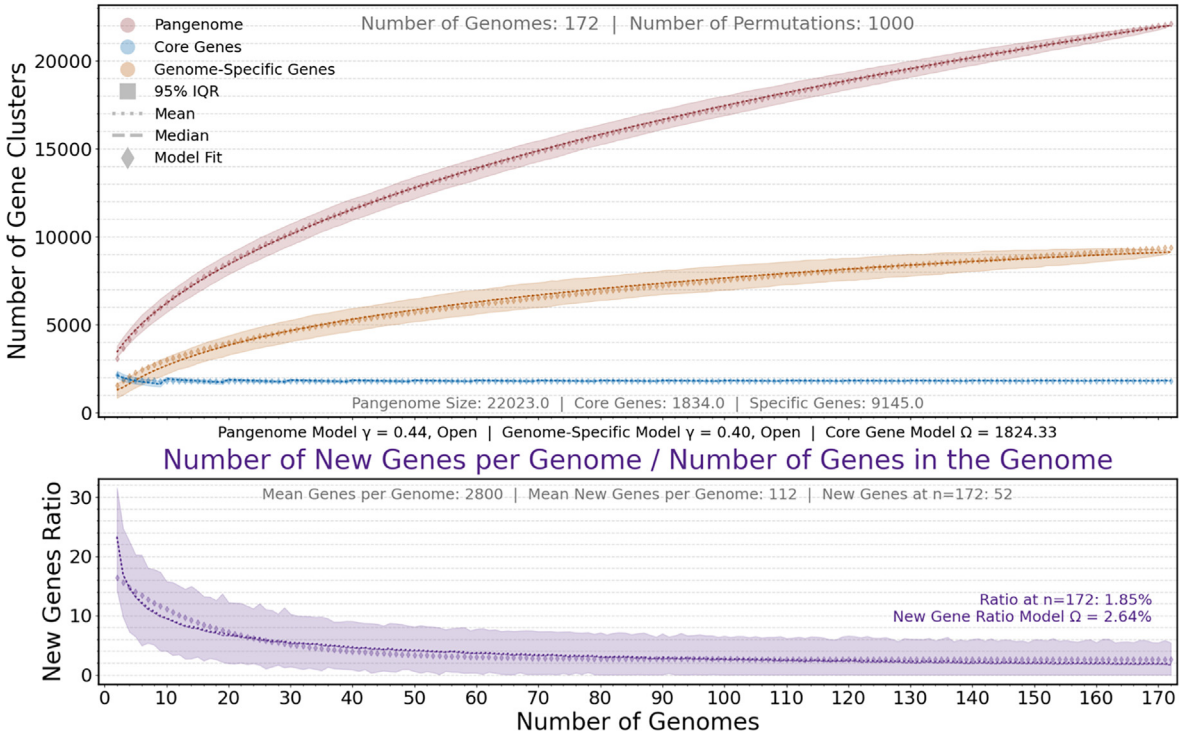

Fig. 3. Pangenome comparison of *Sal. pepae* genomes. The top panel shows the mean, median, 95% empirical confidence interval of permuted values, and the model fit for each of three curves showing the total non-redundant genes in the pangenome (in red), total number of core genes (in blue), and total number of isolate-specific genes (in orange; y-axis) plotted against the number of genomes sampled (x-axis). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

transporters, TonB-dependent receptors and phage integrase genes (Sup. Spreadsheet S6). The enrichment of glycosyltransferases and the presence of sulfotransferases resembled our observations with *Sal. ruber* M8, where genes of similar functions were linked to physiological advantages in tolerating different salinities (Pena et al., 2010). The glycosyltransferase genes could thus be associated with cell osmoregulation and their adaptability to adapt to changes in salinity (Luley-Goedl and Nidetzky, 2011), an important environmental driver in salterns.

Another relevant feature among the *Salinibacter* species is the presence of rhodopsins. *Sal. ruber* was the first cultivated member of the bacterial domain from salterns in which the presence of these non-chlorophyllic photosynthetic genes were detected (Pena et al., 2005). All members of the *Salinibacteraceae* family carry xanthorhodopsin, a special rhodopsin lineage first detected in *Sal. ruber* (Balashov et al., 2005) and later detected in other cultivated (e.g., Vollmers et al., 2013) or yet uncultivated (Boeuf et al., 2021) bacteria. In addition, *Sal. ruber* exhibits a high variability in terms of numbers and kinds (i.e., sensory rhodopsins, halorhodopsins and xanthorhodopsins) of rhodopsins (Pena et al., 2010; Viver et al., 2018). As typically for the genus, all new isolates also encode one copy of a xanthorhodopsin gene, but that was the only rhodopsin detected in the genome (strain ESSP84 was an exception) of *Sal. pepae* that encoded one additional halorhodopsin gene (Sup. Spreadsheet S7). Therefore, the frequency of rhodopsins in the genome seems to distinguish *Sal. ruber* from the new species described here that only code for the xanthorhodopsin (*Sal. grassmerensis*, almost all *Sal. pepae*, *Slv. iranica* and *Slv. lutea*) or for an additional sensory rhodopsin (*Sal. altiplanensis*). This versatility might also be related to the ecological dominance of *Sal. ruber* over its coexisting close relatives.

Altogether, 15 CRISPR spacers with different nucleotide sequences were identified within *Sal. pepae* (Sup. Spreadsheet S8). All encoded for the Type I-E CRISPR-Cas system, but isolate USCM187 additionally encoded a Type III-B. It was remarkable that all CRISPR spacers present in the strains isolated from Great Salt Lake were also present in one or more of the strains isolated from the Spanish salterns (s'Avall, Es Trenc and Santa Pola). Also, some spacers were shared between the Santa Pola and the Mallorca isolates. This might indicate that either these strains had been infected by similar viral populations and these viruses are common between locations, or the infection occurred before dispersal of the strains or in their immediate ancestors (Villamor et al., 2018; Viver et al., 2018). *Sal. grassmerensis* NZ140^T encoded Type I-C and Type I-E CRISPR-Cas system, with unique CRISPR spacers among the *Salinibacter* genomes (Sup. Table S3).

Phenotypes of the cultivated new species

For a standardized description, we proceeded with the classical phenotyping of *Salinibacter* species as done in the past (Antón et al., 2002; Makhdoumi-Kakhki et al., 2012; Viver et al., 2018). The phenotypic properties of *Sal. pepae* and *Sal. grassmerensis* are summarized in Table 5. All strains formed red-pigmented colonies with optimum growth at 30 °C and pH 7. Interestingly, the salinity optimum (w/v) of both new lineages ranged between 15% and 20% (w/v), which was lower than those of *Sal. ruber* and *Sal. altiplanensis*, which range between 20% and 25%. Cells were motile curved rods (Fig. 4 and Sup. Table S4). All *Salinibacter* strains were positive for catalase and oxidase, lysine decarboxylase, and hydrolysis of Tween 80; and negative for anaerobic growth with DMSO, arginine dihydrolase and ornithine decarboxylase activities, indole production, Voges-Proskauer, methyl-red, production of H_2S, hydrolysis of DNA and casein and in the acid production from maltose, glucose and lactose. Variable traits between *Sal. pepae* isolates were the hydrolysis of Tween 20, starch and gelatin. *Sal. grassmerensis* NZ140^T was negative for the hydrolysis of Tween 20 and positive for starch. The strains of *Sal. pepae* CM04 and CZ16 grew with most carbon sources tested (Table 5). Interestingly, *Sal. grassmerensis* NZ140^T grew with all carbon sources tested.

The fatty acid profiles (Table 6) of the new strains were in concordance with the strains of *Sal. ruber* and *Sal. altiplanensis* species (Antón et al., 2002; Viver et al., 2018), being the summed feature 3† (measure that cannot be disaggregated and sums the two fatty acids, *iso*-$C_{15:0}$ 2-OH and/or $C_{16:1}$ ω7t), $C_{18:1}$ ω7c and *iso*-$C_{15:0}$ the major compounds. However, we identified a large number of fatty acids of *Sal. pepae* not identified in the other strains of the *Salinibacter* species, including $C_{13:0}$ 3-OH, *iso*-$C_{14:0}$, $C_{15:1}$ ω6c, $C_{16:0}$ 10-methyl, *cyclo*-$C_{17:0}$, *iso*-$C_{17:1}$ ω7c, *iso*-$C_{17:0}$, $C_{17:1}$ ω8c, $C_{17:1}$ ω6c, $C_{17:0}$, *iso*-$C_{16:0}$ 3-OH and $C_{18:0}$. The fatty acids $C_{17:0}$ 2-OH and $C_{15:0}$ were identified in the majority of *Sal. pepae* isolates and in *Sal. grassmerensis* NZ140^T. We did not identify any species-specific fatty acid profile.

Metabolomics of the cultivated species

High-throughput metabolomics has discriminated biogeographical patterns (Rossello-Mora et al., 2008), growth conditions (Brito-Echeverría et al., 2011), and helped distinguishing wild and laboratory-adapted types of *Salinibacter* (Antón et al., 2013). To our knowledge, this will be the first report on the application of ICR-FT/MS-based metabolomics to taxonomic descriptions. For statistical reasons, the analyses were only applied to the three species with at least 3 available isolates (i.e., *Sal. ruber, Sal. altiplanensis* and *Sal. pepae*). The *Sal. grassmerensis* NZ140^T metabolome was not included in the comparative analysis, as the single isolate available for this species would not provide statistically robust results. Accordingly, its metabolome was only used to compare with the discriminative metabolites identified for *Sal. ruber, Sal. altiplanensis* and *Sal. pepae*. The water-soluble and water-insoluble (but methanol soluble) cellular fractions at the exponential phase rendered a total of 25,244 signals, ~ 17,000 of which were unique of each fraction and isolate and were not considered for our study. On the other hand, 8,218 signals (4,789 for the water-soluble and 3,429 for the -insoluble fractions) were shared between strains and were considered for the statistical analyses. Among them, 3,460 signals could be annotated according to CHNOPS compositions using their masses (42% of the total; Sup. Spreadsheet S9 and S10). A core metabolome of 1,484 annotated metabolites was shared by all members of the genus (Sup. Figure S4).

The core metabolome (Fig. 5A) was formed by a total of 463 and 567 molecular compositions from the water-soluble and water-insoluble cellular fractions, respectively. The chemical compositions derived from exact masses covered two main areas in van Krevelen diagrams representative of chemical classes in the elements CHO, CHNO, CHOS and CHNOS that corresponded to diverse fatty acids and glycosylated compounds. The former correlates with the common *Salinibacter* profiles shown above, and the latter with the conspicuous occurrence of glycosyltransferases in their genomes (Pena et al., 2010; Conrad et al., 2022). Both fractions showed a similar distribution of representative molecular compositions of saturated and polyunsaturated fatty acids (Fig. 5A, left panel). A greater number of fatty acids (~ 155 compounds) were common between the three species (Sup. Figure S5, inserted Venn diagram). Moreover, about 112 and 84 different fatty acids were found specific to core metabolites of the water-soluble and water-insoluble fractions, respectively (Figure S5, inserted Venn diagram). The core fatty acid profiles in the water-soluble fractions were dominated by saturated, mono-, di- and tri-unsaturated fatty acids, while water insoluble-fractions were dominated by polyunsaturated fatty acids (mainly tetra-unsaturated fatty acids and more, Sup. Figure S5). Besides, the water-insoluble fractions showed the presence of specific oxygenated molecular compositions in van Krevelen diagram (compositional ratios $0.33 < O/C < 0.72$ and $1.12 < H/C < 2.23$), mainly annotated to glycosylated-like compounds (Fig. 5A, middle area in the right panel).

Table 5
Diagnostic phenotypic characteristics. All strains showed a red pigmentation, optimum growth temperature at 30 °C, optimum growth pH 7, showed motility and gram-negatively staining. Positive tests for all strains were: catalase, oxidase, lysine decarboxylase and hydrolysis of Tween80. Negative test for all strains were: anaerobic growth with DMSO, arginine dihydrolase and ornithine decarboxylase, indole production, Voges-Proskauer, methyl-red, production of H2S, gas formation with nitrate, hydrolysis of DNA and casein, acid production from the carbohydrates maltose, glucose and lactose.

	Sal. pepae																			*Sal. grassmerensis*	*Sal. ruber*	
Isolates	ESAV49^{Ts}	ESAV87	ESCM2571	ESSP12	ESSP73	ESSP87	CM04	CZ16	CZ17	CZ26	CZ33	DZ03	DZ04	UM13	UZ07	USCM187	USCM25	USCM46	USCM88	NZ140^{T}	M31^{T}	M8
Optimum growth salinity	15	15	20	15	15	15	20	15	15	20	15	15	20	15	15	15	15	15	15	15	20	20
Hydrolysis:																						
Tween20	+	+	+	+	+	+	+	+	+	+	–	–	–	+	+	+	+	+	+	–	+	–
Starch	–	–	–	+	+	–	+	+	–	+	+	+	+	+	+	+	+	–	–	+	+	+
Gelatin	–	–	–	–	–	+	+	–	–	+	+	+	+	+	–	–	–	–	–	ND	+	+
Carbon sources																						
Tyrosine	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Tryptophan	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Asparagine	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	+	+	–
Alanine	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Aspartate	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Proline	–	–	–	–	–	–	+/-	+	–	–	–	+	–	–	–	–	–	–	–	+	+	–
Methionine	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Glucose	–	–	–	–	–	+	+	+	–	+	–	–	–	–	+	–	–	–	–	+	–	–
Ribose	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Mannose	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Yeast	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Fructose	–	–	+	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Raffinose	–	–	+	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Galactose	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Cysteine	–	–	+	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Lactose	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Succinate	–	–	+	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Malate	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–
Lysine	–	–	–	–	–	–	+/-	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–

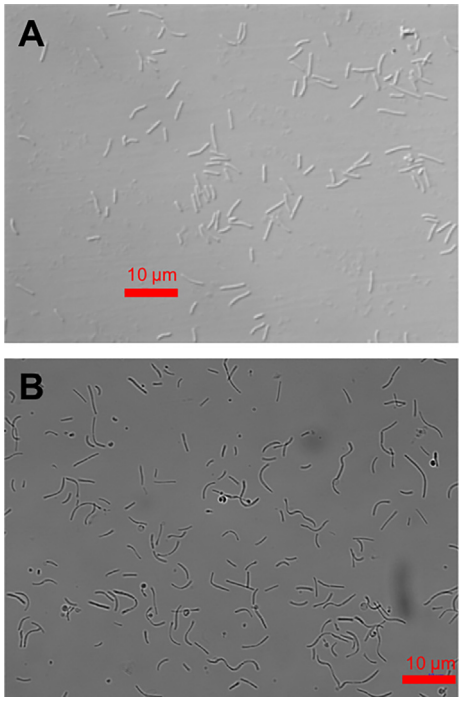

Fig. 4. Cell morphologies of *Sal. pepae* strain ESAV49^{Ts} and *Sal. grassmerensis* strain NZ140^{T}. The bar scale indicates 10 µm.

The three species *Sal. ruber, Sal. altiplanensis* and *Sal. pepae* could be distinguished by their metabolomes. The clustering based on metabolite similarity (Sup. Figure S4) mirrored all reconstructed phylogenies (Fig. 1) and the hierarchical clustering based on presence or absence of orthologous groups (OGs) and AAI (Sup. Figure S2 and S3). In all cases, *Sal. pepae* and *Sal. ruber* appeared closely related and distinct from *Sal. altiplanensis*. When specifically focusing for the discriminative metabolites (i.e., significantly enhanced in their abundance but not necessarily exclusive in comparison with the other groups) that could be annotated (Fig. 5B, 5C, 5D), we observed 468 species-specific metabolites for *Sal. ruber,* 632 for *Sal. pepae* and 704 of *Sal. altiplanensis*. In fact, the main classes of differentiating compounds in both fractions were generally related to saturated and unsaturated fatty acids, fatty acyls and glycosylated-like compounds (Fig. 5). Both fractions of *Sal. pepae*, *Sal. ruber* and *Sal. altiplanensis* were represented by a high diversity of fatty acids, including saturated, mono-, di-, tri-, tetra-, and more unsaturated fatty acids with up to 7 double bond equivalent and 4 hydroxyl moieties (Sup. Figure S6). *Sal. pepae* was mostly represented by a high number of saturated fatty acids (~ 39 compounds) followed by polyunsaturated fatty acids (~ 32 compounds) composed mainly of tetra- and more unsaturated compounds (Sup. Figure S6). Among the saturated fatty acids of *Sal. pepae,* 22 of them were non-hydroxylated and 63 carried at least one hydroxyl moiety in the fatty acyl chain. Accordingly, *Sal. altiplanensis* has the highest number of di- and tri-unsaturated fatty acids, whereas *Sal. ruber* showed the lowest number of fatty acids compared to *Sal. pepae* and *Sal. altiplanensis* (Sup. Figure S6). Besides, specific N-functionalized fatty acyls (i.e., of CHNO-molecular compositions; Fig. 5C upper-left brown shaded area in the diagram) were found only discriminant in the water-soluble fraction of *Sal. ruber*. As an exception, the water-soluble fraction of *Sal. altiplanensis* showed an increased number of oxygenated compounds (composed of CHO-molecular compositions) with the molecular ratios (i.e., $0.31 < O/C < 0.88$ and $1.14 < H/C < 2.11$) that are characteristic of glycosylated-like compounds (Fig. 5D middle blue shaded area in the diagram).

On the other hand, the water-insoluble fraction showed various discriminating compounds that were different in their chemical composition for each species covering a wide range of saturated and unsaturated compounds (Fig. 5BCD, right-panel, red-shaded upper-left area of the diagram). Specifically, we observed an important distinct pattern regarding the distribution of oxygen-containing compounds between *Sal. pepae* and *Sal. ruber* (Fig. 5B and 5C, right panel, grey shaded middle area of the diagram). In fact, *Sal. pepae and Sal. ruber* showed an increased abundance of oxygenated compounds over a wide range of molecular ratios, namely ($0.25 < O/C < 0.69$) and ($0.6 < O/C < 1$) probably specific for glycosylated-like compounds and of their derivatives. By looking at the 20 most discriminating metabolites per class in both fractions, *Sal. pepae* and *Sal. altiplanensis* strains shared eleven metabolites (i.e., unsaturated CHO-based fatty acids, sulfate-based unsaturated CHOS-based fatty acids and CHNO-containing metabolites) in water-soluble fraction, while *Sal. pepae* and *Sal. altiplanensis* pairwise strains shared only three unsaturated CHO-based fatty acids in water-insoluble fraction (Sup. Spreadsheet S9a and S10a). In contrast, *Sal. pepae* and *Sal. ruber* showed no common metabolites with the pairwise strains, namely *Sal. ruber/Sal. altiplanensis* and *Sal. pepae/Sal. altiplanensis* in water-soluble or water-insoluble fractions, respectively. Overall, the variance in the chemical compositions of the studied cellular fractions demonstrated the uniqueness of its metabolome quality in relation to the used strains.

The metabolome of *Sal. grassmerensis* NZ140^{T} was similar in chemical composition with that overall observed for *Sal. ruber*, *Sal. altiplanensis* and *Sal. pepae*. Despite *Sal. grassmerensis* was only represented by NZ140^{T}, which does not allow for statistical comparisons, we investigated here the pairwise comparison of the discriminating compounds of *Sal. ruber*, *Sal. altiplanensis* and *Sal. pepae* with the total assigned compounds of *Sal. grassmerensis* NZ140^{T} in either water-soluble or water-insoluble fractions. The count of assigned compounds (in %) shared with the discriminating compounds in each species (when comparing pairs of data) was as follow 47% (*Sal. ruber*), 38% (*Sal. pepae*) and 35% (*Sal. altiplanensis*) (Sup. Figure S7). In the water-insoluble fraction, the count of assigned compounds (in %) of *Sal. grassmerensis* NZ140^{T} that correspond to the discriminating compounds of *Sal. ruber* and *Sal. altiplanensis* were two-fold higher than those of *Sal. pepae* (Sup. Figure S7, bottom panels) covering an extensive range of saturated and unsaturated molecular ratios (i.e., $1.37 \leq H/C \leq 2.11$ and $0.09 \leq O/C \leq 0.3$). In the water-soluble fraction, the count of assigned compounds (in %) of *Sal. grassmerensis* NZ140^{T} shared with the discriminating compounds of *Sal. ruber* and *Sal. pepae* covered similar but distinct areas of saturated and unsaturated compounds (~ 28%) in van Krevelen diagrams (Sup. Figure S7, top left panel and top middle panel). In contrast, a specific class of glycosylated-like compounds (~ 14%) primarily composed of CHO-containing compounds was common only between *Sal. grassmerensis* NZ140^{T} and the discriminant compounds of *Sal. altiplanensis* (Sup. Figure S7, top right panel). The overall metabolome was typical for the genus, but the composition demonstrated to be unique compared to the other species, a fact that still would require further investigation with additional strains.

A deeper study of the annotated compounds (Sup. Spreadsheet S9 and S10) will be necessary to discern the different isomers that each mass could be related to, and probably would need some additional experimentation. However, at a first glance the approach seems promising as metabolomically we can discriminate the three cultivated species with multiple strains. Altogether, this high-throughout approach is feasible and should be further implemented in taxonomic studies to evaluate its usefulness to overcome the simplicity of the current phenotyping practices in most taxonomic studies which address biochemical properties, often based on commercial kits, with of irrelevant phenotypic and chemotaxonomic tests that add little insight into the biology of the taxon under study (Sutcliffe et al., 2021). The advantage of these high-throughput methods is that, as occurs with MALDI-TOF, they generate absolute molecular mass/abundance information that can be accumulated in databases and may serve long time as source for analysis and discrimination (Rosselló-Móra, 2012).

Table 6
Long-chain cellular fatty acid composition of the described isolates of genus *Salinibacter*.

	Sal. pepae																			*Sal. grassmerensis*	*Sal. ruber*				*Sal. altiplanensis*		
Fatty acid	CZ16	CZ17	CZ26	CZ33	CM04	UM13	DZ04	DZ03	UZ07	USCM25	USCM187	USCM46	ESSP73	ESSP87	ESSP12	ESAV49^{Ts}	ESAV87	USCM46	USCM88	NZ140^{T}	M31^{T}	M8	SP273	ST67	AN4	AN15^{T}	LL19
Unknown 13.565*	0.6	0.5	0.4	1.4	1.0	0.6	0.6	0.4	–	–	0.7	0.6	0.4	0.9	–	0.4	0.5	0.9	–	–	0.8	–	–	–	–	1.7	–
Unknown 14.263*	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.5	–	–	–	–	–	–	–	–	–
$C_{13:0}$ 3-OH	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.7	–	–	–	–	–	–	–	–	–	–	–
iso-$C_{14:0}$	–	–	–	–	–	–	–	–	–	–	0.2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
iso-$C_{15:1}$F	0.4	0.5	0.3	–	0.3	0.4	–	–	0.8	–	0.8	–	–	–	–	0.4	0.5	0.5	–	–	–	–	1.2	–	–	–	–
iso-$C_{15:0}$	18.5	20.2	26.2	24.4	22.4	21.5	19.1	22.8	24.2	21.3	20.3	21.5	23.3	24.4	20.7	21.0	20.0	22.5	23.3	22.3	23.8	27.7	23.6	27.3	21.6	21.4	23.5
antesio-$C_{15:0}$	4.0	4.8	5.1	5.3	5.1	4.9	3.6	4.2	5.2	5.1	4.6	4.6	5.3	5.9	4.0	5.1	4.6	4.3	4.5	3.5	3.6	4.6	4.3	3.4	–	4.4	5.8
$C_{15:0}$	1.1	1.2	0.9	0.8	0.9	–	1.1	1.1	0.9	0.8	0.6	1.1	1.3	0.7	1.2	0.7	0.5	0.7	1.0	0.8	–	–	–	–	–	–	–
$C_{15:1}$ ω6*c*	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.5	–	–	–	–	–	–	–	–	–	–	–
iso-$C_{16:0}$	2.9	2.5	3.0	3.9	3.5	2.7	3.4	3.3	3.0	3.3	6.3	3.3	2.6	2.6	3.4	3.3	3.1	2.1	4.0	3.0	–	–	1.8	–	2.2	–	2.3
$C_{16:1}$ ω5*c*	1.3	0.3	0.4	0.8	0.7	1.9	1.4	–	–	0.6	1.1	1.1	–	–	0.8	0.9	0.7	1.3	–	–	–	–	–	–	–	1.8	–
$C_{16:0}$	4.7	4.0	4.6	4.2	4.5	4.1	5.8	4.8	3.3	3.9	3.1	4.0	3.7	4.1	5.6	2.5	3.1	4.2	4.1	2.1	6.2	5.4	4.4	8.0	4.1	4.8	3.8
$C_{16:0}$ 10-methyl	–	–	2.0	–	–	–	–	–	2.8	–	–	–	2.2	1.5	2.4	–	1.0	–	–	–	–	–	–	–	2.0	–	–
Unknown 16.582*	–	–	0.4	0.5	0.4	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
cyclo-$C_{17:0}$	–	–	–	–	–	–	–	–	1.9	–	–	–	1.7	1.3	1.4	–	–	–	–	–	–	–	–	–	–	–	–
iso-$C_{17:1}$ ω7*c*	1.9	0.8		1.9	2.2	1.0	2.1	1.6	–	1.2	3.4	2.0	–	2.0	1.4	2.4	–	1.9	1.8	–	–	–	–	–	–	–	–
iso-$C_{17:0}$	0.5	0.8	1.0	1.0	0.8	–	0.7	0.7	0.7	0.9	0.6	0.8	–	–	–	0.5	0.4	–	0.7	–	–	–	–	–	–	2.5	–
antesio-$C_{17:0}$	1.6	2.1	2.0	2.2	2.4	0.8	1.8	1.7	1.8	3.1	2.0	1.9	1.1	–	–	1.4	1.1	1.2	1.5	1.7	–	–	–	–	1.8	1.7	–
$C_{17:1}$ ω8*c*	1.1	0.5	0.5	0.5	0.7	1.2	1.3	0.6	0.7	0.5	1.1	0.6	–	–	–	0.5	–	1.5	0.5	–	–	–	–	–	–	–	–
$C_{17:1}$ ω6*c*	2.0	1.9	1.4	1.3	1.7	1.7	1.5	1.4	–	1.4	1.6	1.7	–	–	–	1.5	1.3	2.2	1.4	–	–	–	–	–	–	–	–
$C_{17:0}$	0.7	0.8	0.7	0.6	0.8	–	0.6	0.8	–	0.5	–	0.7	–	–	–	–	–	–	0.5	–	–	–	–	–	–	–	–
iso-$C_{16:0}$ 3-OH	1.3	Tr.	–	–	–	Tr.	2.4	–	–	–	Tr.	–	–	–	–	–	–	2.0	–	–	–	–	–	–	–	–	–
$C_{18:1}$ ω7*c*	20.8	21.2	20.2	19.2	18.4	22.5	18.3	23.7	20.0	20.6	17.8	21.1	22.7	25.4	19.8	21.9	22.5	16.3	21.3	24.1	27.3	23.6	24.5	21.4	25.4	18.1	18.9
$C_{18:0}$	0.8	0.9	0.8	–	–	1.2	1.2	–	–	–	1.0	1.0	–	–	–	–	–	1.4	–	–	–	–	–	–	–	–	–
iso-$C_{17:0}$ 3-OH	2.7	2.8	4.7	3.1	3.4	2.5	3.0	3.6	2.0	3.3	3.1	3.0	3.4	2.9	3.8	3.0	3.6	3.0	3.0	3.5	3.6	3.2	4.0	3.1	4.2	3.6	3.5
$C_{17:0}$ 2-OH	1.0	1.4	1.5	1.1	0.9	1.2	0.8	1.3	0.7	1.3	1.1	1.1	1.3	1.3	1.0	1.2	1.6	1.0	1.0	1.1	–	–	–	–	–	–	–
$C_{18:0}$ 3-OH	Tr.	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Summed feature 3†	32.1	23.7	25.3	28.1	30.2	31.7	31.2	28.0	31.8	31.3	30.6	30.8	30.7	29.0	32.9	32.2	30.8	32.4	30.6	35.4	30.1	35.6	30.9	36.7	38.3	36.9	39.8

*Unknown fatty acid; numbers indicate equivalent chain length.

† Fatty acids that could not be separated by GC using the Microbial Identification System (Microbial ID) software were considered summed features. Summed feature 3 contains $C_{15:0}$ iso 2-OH and/or $C_{16:1}$ ω7*t*.

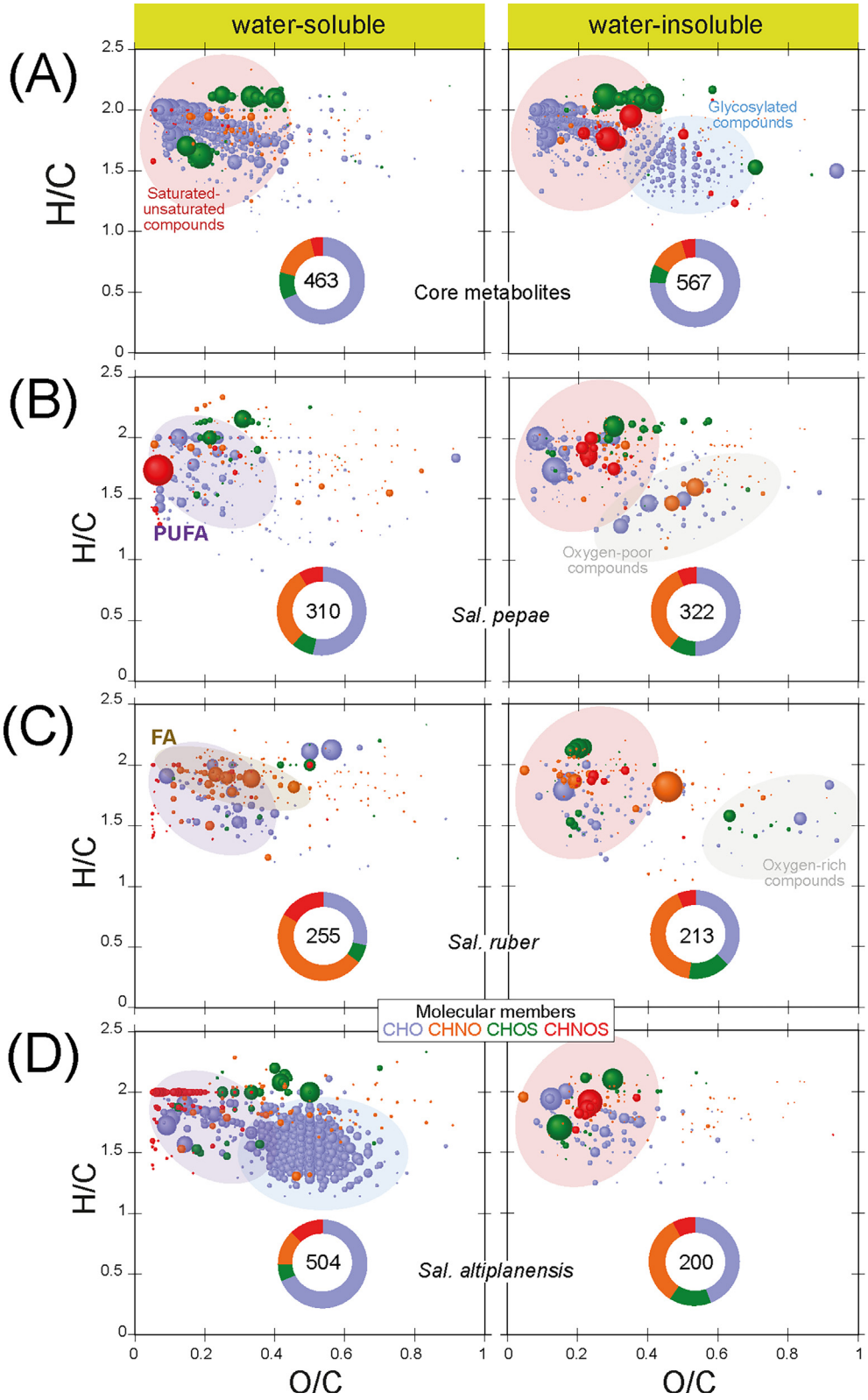

Fig. 5. Van Krevelen graphical presentation of the core metabolomes (A), and the discriminative metabolomes of *Sal. pepae* (B), *Sal. ruber* (C) and *Sal. altiplanensis* (D). The Van Krevelen diagram (H/C versus O/G) is based on all discriminating *m/z* values independent of their origin but colored only as a function of their attributed elementary composition (CHO, blue; CHON, orange; CHOS, green; CHONS, red). The metabolomes were analyzed using the water-soluble (left panel) and water-insoluble (but methanol soluble, right panel) fractions of all samples, and the species-discriminative metabolites are given in Supplementary Spreadsheet Tables S9 and S10, respectively. The shaded areas correspond to: purple - polyunsaturated fatty acids (PUFA); brown – fatty acyls (N-based molecular compositions; FA); red – Saturated and unsaturated compounds; light blue –glycosylated compounds, light grey – oxygen containing compounds.

MAGs of uncultivated Salinibacter species

The number of contigs of each assembly, assembly size and data quality of the MAGs of *Sal. abyssi* ROFF^{Ts} and *Sal. pampae* ARCCH^{Ts} and ARCG are provided in Table 2. The ANI between *Sal. pampae* ARCCH^{Ts} and ARCG was 99.39%, indicating that both MAGs belong to the same species (Richter and Rosselló-Móra, 2009). MAGs of the two species showed ANI values < 91.48% and AAI < 91.99% among themselves and with type strains of other *Salinibacter* species (Table 3). The 16S rRNA gene was binned as part of the *Sal. abyssi* ROFF^{Ts} MAG but not for *Sal. pampae* ARCCH^{Ts} or ARCG MAGs. Based on phylogenetic affiliation, sequence coverage, and the concordance with the phylogenomic reconstructions, we could confidently assign a non-binned 16S rRNA gene sequence to their corresponding MAGs (Fig. 1 and Sup. Figure S8). The procedure of the 16S rRNA gene assignment is described in Sup. Text S2.

According to KEGG metabolic reconstruction, we identified the genes for the central carbohydrate metabolism pathways, including the complete glycolysis, gluconeogenesis, pyruvate metabolism, the TCA cycle (Krebs cycle) and the pentose phosphate pathways. Most ribosomal proteins and flagella genes were also recovered as part of the MAGs (Sup. Table S5). In comparison with the phenotypic traits used to diagnose cultured *Salinibacter* species (Table 5), the MAGs contain genes for oxidase, catalase and lysine decarboxylase activities, as well as starch hydrolysis. Genes that were not detected included the nitrate reductase, other genes for denitrification, ornithine and adenine decarboxylase and arginine dihydrolase. Therefore, such traits might be considered negative, although the possibility that the genes/-functions are present but not recovered as part of the MAG cannot be excluded until the genome is closed.

From a total of 2,705 proteins predicted in *Sal. abyssi* ROFF^{Ts}, we identified 133 not encoded in any of the other genomes of the family *Salinibacteraceae*. After annotation with TREMBL database, 30 proteins were not annotated (no match identified using TREMBL, UniProt and KEGG databases) and 32 were annotated as uncharacterized proteins (Sup. Spreadsheet S11). The remaining species-specific genes mostly encoded for poorly characterized functions, but we were able to identify several phage tail proteins, membrane transporters (as ABC transporters and sodium-independent anion transporter) and chaperonins among them. From a total of 2,837 proteins predicted in *Sal. pampae* ARCCH^{Ts} and 2,903 predicted in ARCG, in the pangenome analysis we identified 47 species-specific proteins against other genomes of the *Salinibacteraceae* family, 133 ARCCH^{Ts}-specific genes and 137 ARCG-specific genes. Based on protein annotation using TREMBL database, 80 of the 317 specific genes to *Sal. pampae* were not annotated (no match identified using TREMBL database) and 58 annotated as uncharacterized proteins. The remaining species-specific genes encoded poorly characterized functions, but we were able to identify several glycosyl- and methyltransferases associated with membrane transport (i.e., potassium transporter TrkH, TRK potassium uptake and ABC transporters) and a ParA gene among them (Sup. Spreadsheet S12 to S14). Similar to the majority of *Salinibacter* species (Pena et al., 2010; Viver et al., 2018), *Sal. abyssi and Sal. pampae* only encoded one copy of each sensory rhodopsin, xanthorhodopsin and halorhodopsin. *Sal. abyssi* ROFF^{Ts} encoded for CRISPR spacers but the CRISPR-Cas systems were not found in the recovered MAGs. The sequences of the CRISPR spacers were also identified in the genomes of *Sal. pepae* (ESSP87, USCM187 and CM04; Sup. Spreadsheet S8), indicating that closely related viruses are present in the locations from which we recovered *Sal. pepae* genomes, or the viral infection occurred before dispersal. *Sal. pampae* did not contain any CRISPR-Cas system.

Environments and biogeographic distribution of Salinibacter species

As shown in Table 1, the salinities of the samples from which the genomes were recovered from ranged from ~ 30%, in Fără Fund, to ~ 40%, in Santa Pola, and NaCl was the major type of salt in all samples. The ionic composition of the different brines was as expected for their origin (Oren, 2002; Decampo and Jones, 2014). Specifically, the three Mediterranean coastal salterns, S'Avall, Es Trenc and Santa Pola, showed a typical ionic composition generated by the concentration of seawater due to evaporation with dominance of sodium (between 62 and 76 g/L) and chloride (between 176 and 217 g/L), and nearly equivalent concentrations of magnesium (between 20 and 60 g/L). Similarly results, but higher proportion of magnesium salts, were observed in the coastal Lake Grassmere (46.9 g/L), which has also a marine origin. On the other hand, in the inland hypersaline sites such as Great Salt Lake, Laguna Colorada Grande, Laguna Colorada Chica and Fără Fund, in which salts are probably originating from the dissolution of underlying diapiric rocks, sodium (between 106 and 119 g/L) and chloride (between 178 and 196 g/L) were more abundant than magnesium (between 0.4 and 17.8 g/L) or sulphate (between 1.5 and 51 g/L), which accounted for a very minor proportion.

The ionic composition, together with the environmental conditions (climate, human manipulation of the semiartificial solar salterns, etc.,) and the geographical distances, ranging from 350 Km to 19,500 Km along our sampled sites (Sup. Figure S1), must have an influence on the microbial community structures, and therefore on the occurrence of *Salinibacter* species. To assess the impact of these factors on the biogeography of the species studied here, we performed a competitive read mapping using as references the genomes of the type strains of all previously known and newly proposed here *Salinibacter* species against the metagenomes of the eight hypersaline sites from where the new species were retrieved (Fig. 6). Conspicuously, *Sal. ruber* and *Sal. pepae* were present in all samples, and therefore, they represent the geographically most widely distributed species. Both always coexisting, *Sal. ruber* always with abundance 2 to 5 times higher than *Sal. pepae* (Fig. 6). As speculated above, the larger pangenome of *Sal. ruber* may partly account for this ecological success. Contrary, *Sal. altiplanensis* was neither recruited in any of the metagenomes used here nor in any of the large collection of samples studied previously based on amplicon sequencing, even in the samples from where it was isolated (Mora-Ruiz et al., 2018), and despite the fact that *Sal. altiplanensis* was the only member of the family cultivated from these samples (Viver et al., 2018). The three additional new species described here represented the most abundant *Salinibacter* populations in their environments of origin. *Sal. grassmerensis*, with a relative abundance of 3.7% in the Lake Grassmere (New Zealand), was additionally detected in Es Trenc (Spain) with a relative abundance of 0.021% (Fig. 6). *Sal. abyssi* with a relative abundance of 18.9% in Fără Fund, also recruited reads with high identity (> 95%) in the metagenomes from S'Avall, Es Trenc, and Great Salt Lake, with a relative abundance of 0.81%, 0.26% and 0.11%, respectively, but was not detected in the Argentinian samples. Finally, *Sal. pampae* represented the dominant *Salinibacter* with a relative abundance of 18.9% and 1.9%, in the Laguna Colorada Chica and Laguna Colorada Grande respectively, and was also detected in S'Avall, Es Trenc and Great Salt Lake samples (Fig. 6), and not detected in the Romanian sample.

We could not observe a pattern that would be related to the ionic composition of the samples. *Sal. ruber* and *Sal. pepae* always coexisted in all samples, and the new species detected that were dominant in their own original environment did not follow any ionic pattern neither. For instance, *Sal. ruber* and *Sal. pepae* occurred in all studied environments and were the dominant bacterial species in the Mediterranean salterns (high sodium content) and in the Great Salt Lake which magnesium content is much lower. Therefore, these major salts appear to not play an important role for the occurrence of at least these two species. Conspicuously, the three new species *Sal. grassmerensis, Sal. abyssi* and *Sal. pampae*, besides dominating in their own samples, were also detected in the salterns of Mallorca, but not in the other samples. One of the explanations of the co-occurrence of all species described here could be due to an efficient dispersal

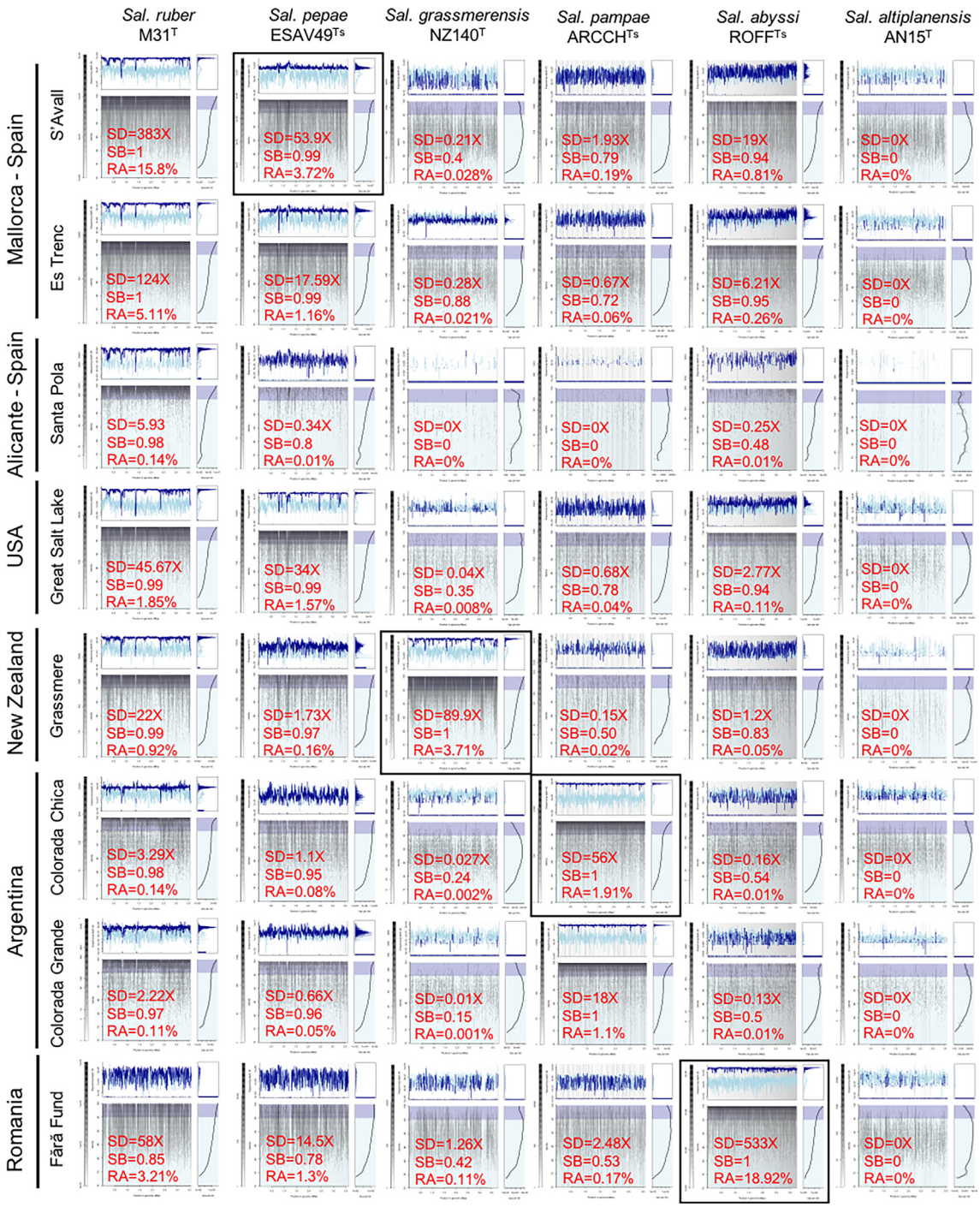

Fig. 6. Read-recruitment approach to identify and track *Salinibacter* species in metagenomic samples. Each line represents a metagenomic sample from where the isolates or MAGs used for the new *Salinibacter* species characterization were retrieved. The columns represent the different *Salinibacter* species and the type material genomes used for recruitment of metagenomic reads. Note that we considered the recruited reads representing the type material genome (reference genome) as reads mapping between 95 and 100% similarity (dark blue) and the reads representing different *Salinibacter* populations are represented by reads mapping at < 95% similarity (light blue). Each individual recruitment plot shows the sequencing depth (SD) or coverage (y-axis) of the genome (x-axis) by reads representing the genome (dark blue) and the remaining populations (light blue). The sequencing breadth (SB) estimates the proportion of genome that has been mapped by metagenomic reads. Relative abundance (RA) indicates the percentage of metagenomic reads mapping to each *Salinibacter* genome. Black squares indicate the sample from where the type strains or MAGs were recovered. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

through the oceanic waterbodies in where the *Salinibacter* strains could be transported in a dormant state through geological time scales and then concentrated in the salterns by the cyclic man-powered seawater feeding. These observations would be also supported by the fact that we found all CRISPR spacers of *Sal. pepae* from the isolates of Great Salt Lake, and from *Sal. abyssi* MAG in one or more of the strains from the Spanish salterns. However, the lack of detection of the *Sal. grassmerensis, Sal. abyssi* and *Sal. pampae* in the Santa Pola salterns, is not consistent with the abovementioned hypothesis as this location is the closest to the Mallorca salterns among all other sites studied.

Table 7
Difference alignments of probe target sites.

Organism	Probe	Probe sequence (5′-3′)	rRNA target pos. (*E. coli* numbering)	Ref.
Sal. altiplanensis	SalAl174	ACCGUACGUCGUCUGGAC	16S, 174–192	This study
Sal. abyssi	SalAb598	UCGGAGGUGAAAGUCCAU	16S, 598–616	This study
Sal. grassmerensis	SalGr152	AUCACGGGAAACUGUGGC	16S, 152–170	This study
Sal. pampae	SalPa1448	UACGUUCAGAGCGAGACG	16S, 1448–1466	This study
Sal. pepae	SalPe460	UAGCUUCAGAGCGAGACG	16S, 460–478	This study
Sal. ruber	SalRu183	GAUCCCGCAUGGGGACCA	16S, 183–201	This study
Salinibacter genus	EHB412[a]	ACACCCCUAUGGGGCGUA	16S, 414–432	Antón et al., 2000
Sal. ruber (EHB-1)	EHB586[a]	GGGCAGCAAGUCGGAUGU	16S, 588–606	Antón et al., 2000
Sal. ruber (EHB-2)	EHB1451[a]	AGCCGGAGGGAGAGCGGC	16S, 1452–1470	Antón et al., 2000
Sal. altiplanensis	EHB130[b]	AGACAACCUGCCCAAAAG	16S, 126–155	Viver et al., 2018

Therefore, one possibility is that these new species are part of the rare biosphere (Pedrós-Alió, 2007). Also, the particular flow dynamics of the waterbodies of the Western Mediterranean waters could effectively physically isolate the two coastal sites from each other (Barceló-Llull et al., 2019). The presence of the new species in coastal sites as well as inland lake in Romania or Argentina is also challenging to explain but, as we have demonstrated previously, migrating birds could effectively transport actively growing extreme halophile communities between distant sites (Brito-Echeverría et al., 2009). More samples from other hypersaline sites and additional isolates of *Salinibacter* from around the world will shed light on the dispersal patterns of these extreme halophiles in the future. It should be pointed out that the present study is just a snapshot of a single time and site of the hypersaline environments, and that the physicochemical dynamics influence on the relative occurrence and dominance of distinct species and subspecies units in such environments (Viver et al., 2021), and perhaps what has not been detected based in the current sampling, may show increased abundances under other environmental conditions.

Probe design for future studies

For further ecological studies we have explored the currently available SILVA SSU Ref99 138.1 database, curating it to retain the best sequences of almost full length and high SILVA quality scores (Pruesse et al., 2007), to provide specific probes for the new species. For each of the new species, we selected the best probe that theoretically will give enough specificity to identify and quantify them using fluorescence microscopy, southern blots or any other technique that would require specific oligonucleotide sequences. The resulting sequences are given in Table 7 and Sup. Spreadsheet S15, together with the previously designed probes for the *Salinibacter* genus (EHB), *Sal. ruber* (EHB-1), the second *Sal. ruber* phylotype (EHB-2) and *Sal. altiplanensis*. The designed probes could be used in the future to identify and study morphology in environmental samples.

Concluding remarks

In summary, combining cultivation with metagenomics allowed us to better characterize the taxonomic diversity of *Salinibacter* compared to what would have not been feasible based on cultivation alone. On the other hand, the use of high throughput genome- and metabolomics-inferred phenotypes allow a fine-tuned comparison of strains for descriptive purposes, with the advantage compared to alternative phenotypic methods (e.g., biology plates), that produced data can be obtained (and re-analyzed) from the public databases (Rosselló-Móra and Amann, 2015). Only time will reveal whether naming uncultivated prokaryotes under the SeqCode (Hedlund et al., 2022), and the use of high-throughput phenotyping methods for cultivated strains (Rosselló-Móra and Amann, 2015), will become popular. Our results presented here shows that these approaches have significant advantages over traditional methods for the same purposes.

Sampling permits

The samples from S'Avall, Es Trenc and Santa Pola were taken in accordance with the permit ESNC27, with the unique identifier ABSCH-IRCC-ES-241224–1 that has been provided by the Dirección General de Biodiversidad y Calidad Ambiental del Ministerio para la Transición Ecológica of the Spanish Government. The samples of the Pampa region were taken under the permit of the Dirección de Recursos Naturales of the Argentinian Ministerio de la Producción given to M. E. Llames under the umbrella of the project *Microplayers on macro roles: functional patterns of microbial communities in shallow lakes*. The sample of the Great Salt Lake was permitted by the state of Utah.

Data availability

All data is publicly available in research repositories.

Acknowledgements

The authors would especially like to thank the whole team at Salinas d'Es Trenc and Gusto Mundial Balearides, S.L. (Flor de Sal d´Es Trenc), and to the family Dezcallar owner of the Salterns of S'Avall for allowing access to their facilities. Aharon Oren is acknowledged for supervising the taxonomic etymologies of the new species and new combinations. This study was funded by the Spanish Ministry of Science, Innovation and Universities projects PGC2018-096956-B-C41, RTC-2017-6405-1 and PID2021-126114NB-C42, which were also supported by the European Regional Development Fund (FEDER). RRM acknowledges the financial support of the sabbatical stay at Georgia Tech and HelmholzZentrum München by the grants PRX18/00048 and PRX21/00043 respectively also from the Spanish Ministry of Science, Innovation and Universities. This research was carried out within the framework of the activities of the Spanish Government through the "Maria de Maeztu Centre of Excellence" accreditation to IMEDEA (CSIC-UIB) (CEX2021-001198). KTK's research was supported, in part, by the U.S. National Science Foundation (Award No. 1831582 and No. 2129823). IMG. AC and HLB were financially supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI – UEFISCDI, project number PN-III-P4-ID-PCE-2020-1559, within PNCDI III. HLB acknowledges Ocna Sibiului City Hall (Sibiu County, Romania) for granting the access to Fără Fund Lake and A. Baricz and D.F. Bogdan for technical support during sampling and sample preparation. MBS thanks Dominion Salt for their assistance in sample Lake Grassmere. MELL acknowledges the financial support of the Argentinian National Scientific and Technical Research Council (Grant CONICET-NSFC 2017 N° IF-2018-10102222-APN-GDCT-CONICET) and the National Geographic Society (Grant # NGS 357R-18). BPH was supported by NASA (award 80NSSC18M0027). TV acknowledges the "Margarita Salas" postdoctoral grant, funded by the Spanish Ministry of Universities, within the framework of Recovery, Transformation and Resilience Plan, and

funded by the European Union (NextGenerationEU), with the participation of the University of Balearic Islands (UIB).

Author Contributions

RRM and TV designed the experiment. TV, MU, JFG, ASS and EBC collected and processed samples from Mallorca. RSM, EM and FS collected and processed the sample from Alicante. FF and MEL collected and processed the samples from Argentina. JP, BB and BPH collected, processed and isolated the strains of the sample from Utah. IMG, AC and HLB collected and processed the sample from Romania. MBS collected and processed the sample from New Zealand. ML, MH, and PSK analyzed the metabolomic samples. PK analyzed the fatty acid profiles. REC, SV and KTK supervised the metagenomic and genomic computational analysis. RA and TV designed the specific probes. TV analyzed the data and together with RRM wrote the manuscript. All coauthors also read, commented, and corrected the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.syapm.2023.126416.

References

Antón, J., Rosselló-Mora, R., Rodríguez-Valera, F., Amann, R., 2000. Extremely halophilic bacteria in crystallizer ponds from solar salterns. Appl. Environ. Microbiol. 66, 3052–3057.

Antón, J., Oren, A., Benlloch, S., Rodríguez-Valera, F., Amann, R., Rosselló-Móra, R., 2002. *Salinibacter ruber* gen nov., sp. nov., a novel extremely halophilic member of the bacteria from saltern crystallizer ponds. Int. J. Syst. Evol. Microbiol. 52, 485–491.

Antón, J., Lucio, M., Peña, A., Cifuentes, A., Brito-Echeverría, J., Moritz, F., Tziotis, D., López, C., Urdiain, M., Schmitt-Kopplin, P., Rosselló-Móra, R., 2013. High metabolomic microdiversity within co-occurring isolates of the extremely halophilic bacterium *Salinibacter ruber*. PLoS One. 8, e64701.

Balashov, S.P., Imasheva, E.S., Boichenko, V.A., Antón, J., Wang, J.M., Lanyi, J.K., 2005. Xanthorhodopsin: a proton pump with a light-harvesting carotenoid antenna. Science. 309, 2061–2064.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.

Barceló-Llull, B., Pascual, A., Ruiz, S., Escudier, R., Torner, M., Tintoré, J., 2019. Temporal and spatial hydrodynamic variability in the Mallorca channel (western Mediterranean Sea) from 8 years of underwater glider data. J. Geophys. Res. Oceans. 124, 2769–2786.

Barrow, G., Feltham, R. (Eds.), Cowan and Steel's Manual for the Identification of Medical Bacteria, Cambridge University Press, Cambridge, 1993.

Boeuf, D., Eppley, J.M., Mende, D.R., Malmstrom, R.R., Woyke, T., DeLong, E.F., 2021. Metapangenomics reveals depth-dependent shifts in metabolic potential for the ubiquitous marine bacterial SAR324 lineage. Microbiome. 9, 1–18.

Brito-Echeverría, J., López-López, A., Yarza, P., Antón, J., Rosselló-Móra, R., 2009. Occurrence of *Halococcus* spp. in the nostrils salt glands of the seabird *Calonectris diomedea*. Extremophiles. 13, 557–565.

Brito-Echeverría, J., Lucio, M., López-López, A., Antón, J., Schmitt-Kopplin, P., Rosselló-Móra, R., 2011. Response to adverse conditions in two strains of the extremely halophilic species *Salinibacter ruber*. Extremophiles. 15, 379–389.

Buchfink, B., Reuter, K., Drost, H.G., 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat. Meth. 18, 366–368.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinform. 10, 1–9.

Conrad, R.E., Viver, T., Gago, J.F., Hatt, J.K., Venter, S.N., Rossello-Mora, R., Konstantinidis, K.T., 2022. Toward quantifying the adaptive role of bacterial pangenomes during environmental perturbations. ISME J. 16, 1222–1234.

Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Néron, B., Rocha, E.P., Vergnaud, G., Gautheret, D., Pourcel, C., 2018. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res. 46. W246-W251.

Cowan, S.T., Steel, K.J., 2003. Manual for the Identification of Medical Bacteria. Cambridge University Press, London.

Decampo, D.M., Jones, G.F., 2014. Geochemistry of Saline Lakes. Treat. Geochem. (Second Edition). 7, 437–469.

Dussault, H.P., 1995. An improved technique for staining red halophilic bacteria. J. Bacteriol. 70, 484–485.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracyand high throughput. Nucleic Acids Res. 32, 1792–1797.

Gago, J.F., Viver, T., Urdiain, M., Pastor, S., Kämpfer, P., Ferreira, E., Rossello-Mora, R. (2021) Description of three new *Alteromonas* species *Alteromonas antoniana* sp. nov., *Alteromonas lipotrueae* sp. nov. and *Alteromonas lipotrueiana* sp. nov. isolated from marine environments, and proposal for reclassification of the genus *Salinimonas* as *Alteromonas*. Syst. Appl. Microbiol. 44, 126226.

Ghai, R., Pašić, L., Fernández, A.B., Martin-Cuadrado, A.B., Mizuno, C.M., McMahon, K. D., Papke, R.T., Stepanauskas, R., Rodriguez-Brito, B., Rohwer, F., Sánchez-Porro, C., Ventosa, A., Rodriguez-Valera, F., 2011. New abundant microbial groups in aquatic hypersaline environments. Scientific Rep. 1, 1.

Grissa, I., Vergnaud, G., Pourcel, C., 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 35. W52-W57.

Hedlund, B.P., Chuvochina, M., Hugenholtz, P., Konstantinidis, K.T., Murray, A.E., et al, 2022. SeqCode: a nomenclatural code for prokaryotes described from sequence data. Nat. Microbiol. 7, 1702–1708.

Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform. 11, 119–130.

Kämpfer, P., Kroppenstedt, R.M., 1996. Numerical analysis of fatty acid patternsof coryneform bacteria and related taxa. Can. J. Microbiol. 42, 989–1005.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2016. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 44, 457–462.

Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., Wang, Z., 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7, e7359.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Konstantinidis, K.T., Rossello-Mora, R., Amann, R., 2017. Uncultivated microbes in need of their own taxonomy. ISME J. 11, 2399–2406.

Konstantinidis, K.T., Viver, T., Conrad, R.E., Venter, S.N., Rossello-Mora, R., 2022. Solar salterns as model systems to study the units of bacterial diversity that matter for ecosystem functioning. Curr. Op. Biotech. 73, 151–157.

Lanave, C., Preparata, G., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20, 86–93.

Li, D., Luo, R., Liu, C.M., Leung, C.M., Ting, H.F., Sadakane, K., Yamashita, H., Lam, T. W., 2016. MEGAHIT v1. 0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods. 102, 3–11.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S., Hermann, S., Jost, R., König, A., Liss, T., Lussmann,T., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N.,Vildig, A., Lenke, M., Ludwig, T., Bode, A., Schleifer, K.H. (2004) ARB; a software environment for sequence data. Nucleic Acids Res. 32, 1363–1371.

Ludwig, W., Viver, T., Westram, R., Gago, J.F., Bustos-Caparros, E., Knittel, K., Amann, R., Rossello-Mora, R., 2021. Release LTP_12_2020, featuring a new ARB alignment and an improved 16S rRNA tree of prokaryotic type strains. Syst. Appl. Microbiol. 44, 126218.

Luley-Goedl, C., Nidetzky, B., 2011. Glycosides as compatible solutes: biosynthesis and applications. Nat. Product Rep. 28, 875–896.

Makhdoumi-Kakhki, A., Amoozegar, M.A., Ventosa, A., 2012. *Salinibacter iranicus* sp. nov. and *Salinibacter luteus* sp. nov., isolated from a salt lake, and emended descriptions of the genus *Salinibacter* and of *Salinibacter ruber*. Int. J. Syst. Evol. Microbiol. 62, 1521–1527.

Medini, D., Donati, C., Tettelin, H., Masignani, V., Rappuoli, R., 2005. The microbial pan-genome. Current Op. Gen. Develop. 15, 589–594.

Mora-Ruiz, M.D.R., Cifuentes, A., Font-Verdera, F., Pérez-Fernández, C., Farias, M.E., González, B., Orfila, A., Rosselló-Móra, R., 2018. Biogeographical patterns of bacterial and archaeal communities from distant hypersaline environments. Syst. Appl. Microbiol. 41, 139–150.

Munoz, R., Rosselló-Móra, R., Amann, R., 2016. Revised phylogeny of Bacteroidetes and proposal of sixteen new taxa and two new combinations including *Rhodothermaota* phyl. nov. Syst. Appl. Microbiol. 39, 281–1196.

Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 27, 824–834.

Oren, A., 2002. Halophilic microorganisms and their environments. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Oren, A., Arahal, D., Goeker, M., Moore, E., Rossello-Mora, R., Sutcliffe, I., 2022. International Code of Nomenclature of Prokaryotes. Prokaryotic Code, IJSEM. in press.

Pedrós-Alió, C., 2007. Dipping into the rare biosphere. Science. 315, 192–193.

Pena, A., Valens, M., Santos, F., Buczolits, S., Anton, J., Kämpfer, P., Busse, H.J., Amann, R., Rossello-Mora, R., 2005. Intraspecific comparative analysis of the species *Salinibacter ruber*. Extremophiles. 9, 151–161.

Pena, A., Teeling, H., Huerta-Cepas, J., Santos, F., Yarza, P., Brito-Echeverría, J., Lucio, M., Schmitt-Kopplin, P., Meseguer, I., Schenowitz, C., Dossat, C., Barbe, V., Dopazo, J., Rossello-Mora, R., Schüler, M., Glöckner, F.O., Amann, R., Gabaldón, T., Antón, K., 2010. Fine-scale evolution: genomic, phenotypic and ecological differentiation in two coexisting Salinibacter ruber strains. ISME J. 4, 882–895.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., Glöckner, F.O., 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. 35, 7188–7196.

Richter, M., Rosselló-Móra, R., 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. U. S. 45, 19126–19131.

Rodriguez-R, L.M., Konstantinidis, K.T., 2016. The enveomics collection: a tool-box for specialized analyses of microbial genomes and metagenomes. PeerJ 4, e1900v1.
Rodriguez-R, L.M., Gunturu, S., Harvey, W.T., Rosselló-Mora, R., Tiedje, J.M., Cole, J.R., Konstantinidis, K.T., 2018. The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. Nucl. Ac. Res. 46. W282-W288.
Rodriguez-Valera, F., Ventosa, A., Juez, I., J.f., 1985. Variation of environmental features and microbial populations with salt concentrations in a multi-ponds saltern. Microbial Ecol. 11, 107–115.
Rosselló-Móra, R., 2012. Towards a taxonomy of Bacteria and Archaea based on interactive and cumulative data repositories. Env. Microbiol. 14, 318–334.
Rosselló-Móra, R., Amann, R., 2015. Past and future species definitions for Bacteria and Archaea. Syst. Appl. Microbiol. 38, 209–216.
Rossello-Mora, R., Lucio, M., Pena, A., Brito-Echeverría, J., Lopez-Lopez, A., Valens-Vadell, M., Frommberger, M., Anton, J., Schmitt-Kopplin, P., 2008. Metabolic evidence for biogeographic isolation of the extremophilic bacterium *Salinibacter ruber*. ISME J. 2, 242–253.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
Stackebrandt, E., Ebers, J., 2006. Taxonomic parameters revisited: tarnished gold standards. Microbiol. Today. 33, 152–155.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phyloge-netic analyses with thousands of taxa and mixed models. Bioinform. 22, 2688–2690.
Sutcliffe, I.C., Rossello-Mora, R., Trujillo, M.E., 2021. Addressing the sublime scale of the microbial world: reconciling an appreciation of microbial diversity with the need to describe species. New Microb. New Inf. 43, 100931.
Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Fraser, C.M., 2005. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial “pan-genome”. Proc. Natl. Acad. Sci. 102, 13950–13955.
Tettelin, H., Riley, D., Cattuto, C., Medini, D., 2008. Comparative genomics: the bacterial pan-genome. Cur. Op. Microbiol. 11, 472–477.
The UniProt Consortium, 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 49. D480-D489.
Tziotis, D., Hertkorn, N., Schmitt-Kopplin, P., 2011. Kendrick-analogous network visualization of ion cyclotron resonance Fourier transform mass spectra: improved options for the assignment of elemental compositions and the classification of organic molecular complexity. Europ. J. Mass Spectro. 17, 415–421.
Urdiain, M., López-López, A., Gonzalo, C., Busse, H.J., Langer, S., Kämpfer, P., Rosselló-Móra., R., 2008. Reclassification of *Rhodobium marinum* and *Rhodobium pfennigii* as *Afifella marina* gen. nov. comb. nov. and *Afifella pfennigii* comb. nov., a new genus of photoheterotrophic Alphaproteobacteria and emended descriptions of *Rhodobium*, *Rhodobium orientis* and *Rhodobium gokarnense*. Syst. Appl. Microbiol. 31 (5), 339–351.
Van Krevelen, D.W., 1950. Graphical-statistical method for the study of structure and reaction processes of coal. Fuel 29, 269–284.
Villamor, J., Ramos-Barbero, M.D., González-Torres, P., Gabaldón, T., Rosselló-Móra, R., Meseguer, I., Martínez-García, M., Santos, F., Antón, J., 2018. Characterization of ecologically diverse viruses infecting co-occurring strains of cosmopolitan hyperhalophilic *Bacteroidetes*. ISME J. 12, 424–437.
Viver, T., Cifuentes, A., Díaz, S., Rodríguez-Valdecantos, G., González, B., Antón, J., Rosselló-Móra, R., 2015. Diversity of extremely halophilic cultivable prokaryotes in Mediterranean, Atlantic and Pacific solar salterns: evidence that unexplored sites constitute sources of cultivable novelty. Syst. Appl. Microbiol. 38, 266–275.
Viver, T., Orellana, L., Gonzalez-Torres, P., Diaz, S., Urdiain, M., Farías, M.E., Benes, V., Kämpfer, P., Shahinpei, A., Amoozegar, M.A., Amann, R., Anton, J., Konstantinidis, K.T., Rosselló-Móra, R., 2018. 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. 41, 198–212.
Viver, T., Orellana, L.H., Díaz, S., Urdiain, M., Ramos-Barbero, M.D., González-Pastor, J. E., Oren, A., Hatt, J.K., Amann, R., Antón, F., Konstantinidis, K.T., Rosselló-Móra, R., 2019. Predominance of deterministic microbial community dynamics in salterns exposed to different light intensities. Environ. Microbiol. 21, 4300–4315.
Viver, T., Conrad, R.E., Orellana, L.H., Urdiain, M., González-Pastor, J.E., Hatt, J.K., Amann, R., Antón, J., Konstantinidis, K.T., Rosselló-Móra, R., 2021. Distinct ecotypes within a natural haloarchaeal population enable adaptation to changing environmental conditions without causing population sweeps. ISME J. 15, 1178–1191.
Vollmers, J., Voget, S., Dietrich, S., Gollnow, K., Smits, M., Meyer, K., Brinkhoff, T., Simon, M., Daniel, R., 2013. Poles apart: Arctic and Antarctic *Octadecabacter* strains share high genome plasticity and a new type of xanthorhodopsin. PloS one. 8, 63422.
Whitman, W.B., Chuvochina, M., Hedlund, B.P., Hugenholtz, P., Konstantinidis, K.T., Murray, A., Palmer, M., Parks, D.H., Probst, A.J., Reysenbach, A.L., Rodriguez-R, L. M., Rossello-Mora, R., Sutcliffe, I., Venter, S.N., 2022. Development of the SeqCode: a proposed nomenclatural code for uncultivated prokaryotes with DNA sequences as type. Syst. Appl. Microbiol., 126305.
Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comp. Biol. 13, e1005595.