Marine Genomics 44 (2019) 57–60

Contents lists available at ScienceDirect

Marine Genomics

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

Complete genome sequence of *Salinigranum rubrum* GX10^T, an extremely halophilic archaeon isolated from a marine solar saltern

Shuai-Bo Han[a], Xin-Jun Hou[a], Chen Wu[b], Zhe Zhao[a], Zhao Ju[a], Ran Zhang[a], Heng-Lin Cui[c], Lawrence Jordan Keen[d], Lin Xu[d,*], Min Wu[a,*]

[a] College of Life Sciences, Zhejiang University, Hangzhou, PR China
[b] Zhejiang University of Water Resources and Electric Power, Hangzhou, PR China
[c] School of Food and Biological Engineering, Jiangsu University, Zhenjiang, PR China
[d] College of Life Sciences, Zhejiang Sci-Tech University, Hangzhou, PR China

ARTICLE INFO

Keywords:
Extremely halophilic archaea
Salinigranum rubrum
Genome sequence
Marine solar saltern

ABSTRACT

Since the first genome of a halophilic archaeon was sequenced in 2000, microbes inhabiting hypersaline environments have been investigated largely based on genomic characteristics. *Salinigranum rubrum* GX10^T, the type species of the genus *Salinigranum* belonging to the euryarchaeal family *Haloferacaceae*, was isolated from the brine of Gangxi marine solar saltern near Weihai, China. Similar with most members of the class *Halobacteria*, *S. rubrum* GX10^T is an extreme halophile requiring at least 1.5 M NaCl for growth and 3.1 M NaCl for optimum growth. We sequenced and annotated the complete genome of *S. rubrum* GX10^T, which was found to be 4,973,118 bp and comprise one chromosome and five plasmids. A total of 4966 protein coding genes, 47 tRNA genes and 6 rRNA genes were obtained. The isoelectric point distribution for the predict proteins was observed with an acidic peak, which reflected the adaption of *S. rubrum* GX10^T to the halophilic environment. Genes related to potassium uptake, sodium efflux as well as compatible-solute biosynthesis and transport were identified, which were responsible for the resistance to osmotic stress. Genes related to heavy metal resistance, CRISPR-Cas system and light transform system were also detected. This study reports the first genome in the genus *Salinigranum* and provides a basis for understanding resistance strategies to harsh environment at the genomic level.

1. Introduction

Hypersaline ecosystems are widely distributed habitats and play important roles in the global ecosystem (Naghoni et al., 2017). Many hypersaline environments, such as solar salterns and lagoons, originated by evaporation of seawater (Nazareth and Gonsalves, 2014). Their salt composition is similar to that of seawater, where sodium and chloride are dominating ions, and has a pH ranging from neutral to slightly alkaline (Oren, 2002). As the dominant organism group in hypersaline environments, halophilic archaea drive the biogeochemical cycles of hypersaline ecosystems and are significant indicators for the protection and utilization of saline environments (Cui, 2016).

The genus *Salinigranum* consists of two species, *S. rubrum* GX10^T (Cui and Zhang, 2014) and *S. salinum* YJ-50-S2^T (Wang et al., 2016). Although these two strains have been well taxonomically characterized, information regarding their genomic features and stress-tolerant strategy is still lacking. *S. rubrum* GX10^T was isolated from marine solar saltern brine, where the pH was 7.1 and the salinity was 283 g/l. The genome of *S. rubrum* GX10^T was sequenced and assembled into one circular chromosome and five plasmids. The genomic sequence provides us with several aspects to understand the adaptive mechanisms of halophilic archaea in harsh environments.

2. Data description

General features of *S. rubrum* GX10^T were summarized in Table 1. The cells of *S. rubrum* GX10^T were cultivated in neutral oligotrophic haloarchaeal medium (Cui and Zhang, 2014) for 7 days at 37 °C. Genomic DNA was extracted from the cell pellets using the ChargeSwitch® gDNA Mini Bacteria Kit (Life Technologies) according to the manufacturer's instructions. The genome of *S. rubrum* GX10^T was sequenced using Illumina Hiseq2000 platform and PacBio RS II sequencer at Majorbio Bio-Pharm Technology Co. Ltd. (Shanghai, China). For Illumina sequencing, a paired-end library with insert length of 300 bp

* Corresponding authors.
E-mail addresses: linxu@zstu.edu.cn (L. Xu), wumin@zju.edu.cn (M. Wu).

https://doi.org/10.1016/j.margen.2018.09.004
Received 25 May 2018; Received in revised form 12 August 2018; Accepted 19 September 2018
Available online 17 October 2018
1874-7787/

Table 1
General features of *Salinigranum rubrum* GX10^T and MIGS mandatory information.

Items	Descriptions
General features	
Gram strain	Negative
Cell shape	Pleomorphic-shaped
Color of colonies	Red
Temperature	20 to 50 °C
Optimal NaCl concentration	3.1 M
Optimal pH	7.0
MIGS data	
Investigation type	Bacteria_archaea
Project name	*Salinigranum rubrum* Genome sequencing and assembly
Latitude and longitude	37.38° N 122.46° E
Depth	Not reported
Geographical location	China: Weihai
Collection date	Not reported
Environment (biome)	Hypersaline water, [ENVO:00002012]
Environment (feature)	Saline evaporation pond, [ENVO:00000055]
Environment (material)	Water, [ENVO:00002007]
Number of replicons	6
Estimated size	4,973,118 bp
Source material identifiers	GX10 = CGMCC 1.10385 = JCM 17116
Reference of biomaterial	PMID: 24651304
Trophic level	Chemoorganotroph
Relationship to oxygen	Aerobic
Isolation and growth conditions	PMID: 24651304
Library reads sequenced	186, 033
Sequencing methods	PacBio RS II and Illumina Hiseq 2000
Assembly data	
Assembly method	Hybrid *de novo* assembly solution
Assembly name	Srub_v1.0
Genome coverage	~292×
Finishing strategy	Sequencing and assembly

was constructed. For Pacific Biosciences sequencing, 10 kb insert whole genome shotgun libraries were generated and sequenced on a Pacific Biosciences RS instrument following standard methods. The Illumina data was used to evaluate the complexity of the genome. The complete genome sequence was assembled using both the PacBio reads and Illumina reads. The assembly was produced using a hybrid *de novo* assembly solution, in which a *de-Bruijn* based assembly algorithm and a CLR reads correction algorithm were integrated in "PacBioToCA with Celera Assembler" pipeline (Koren et al., 2012). The final assembly generated six circular sequences with no gap existed.

Glimmer version 3.02 (Delcher et al., 2007) was used to predict open reading frames (ORFs) according to the manufacturers' instruction. Following this, ORFs were annotated through NCBI NR, SwissProt (Apweiler et al., 2004), KEGG (Kanehisa et al., 2004), and COG (Tatusov et al., 2000) database. RAST (Rapid Annotation using Subsystem Technology) server online (Overbeek et al., 2014) was also used for the prediction and annotation of ORFs for cross-validation. The tRNA and rRNA genes were predicted using tRNAscan-SE (Lowe and Eddy, 1997) and RNAmmer (Lagesen et al., 2007), respectively. CRISPR repeats were predicted by CRISPRfinder (Grissa et al., 2007). Calculations of the predicted pI of individual proteins were made with Protein isoelectric point calculator (Kozlowski, 2016). Circular figures of *S. rubrum* GX10^T genome were visualized by Circos v0.64 (Krzywinski et al., 2009).

The genome of *S. rubrum* GX10^T consists of one circular chromosome (4,250,563 bp, 66.9% G + C), and five circular plasmids, named as pSRU01 (237,003 bp, 57.6% G + C), pSRU02 (198,037 bp, 59.2% G + C), pSRU03 (118,877 bp, 58.9% G + C), pSRU04 (129,442 bp, 57.9% G + C) and pSRU05 (39,196 bp, 63.3% G + C), respectively. A map of the chromosome and plasmids is shown in Fig. 1. A total of 4966 protein-coding genes and 53 RNA genes were predicted. The genomic features of strain *S. rubrum* GX10^T were summarized in Table 2. The COG functional categories of *S. rubrum* GX10^T were shown in Fig. 1. Among the 4966 ORFs, only 1650 ORFs were classified into COG categories. The major categories were amino acid transport and metabolism (11.4%), energy production and conversion (9.8%), ribosomal structure and biogenesis (8.1%), inorganic ion transport and metabolism (7.8%), transcription (7.0%) and replication, recombination and repair (6.3%).

S. rubrum GX10^T grew at 3.1 M NaCl for optimal growth, and the cells lysed in distilled water. Halophilic mechanisms of *S. rubrum* GX10^T were analyzed according to the complete genome. The Trk system potassium uptake protein was found, which is responsible for K^+ uptake and transport (Kraegeloh et al., 2005), including nine copies of *TrkA* genes and three copies of *TrkH* genes. Two copies of potassium voltage-gated channel subfamily proteins (KQT), two copies of KefA proteins, four copies of Kef-type K^+ transport proteins, four copies of potassium channel proteins and seven copies of pH adaptation potassium efflux system proteins were detected. All proteins mentioned above are related to potassium efflux. Presence of those potassium-related genes indicates that *S. rubrum* GX10^T accumulates K^+ and monitors potassium homeostasis in the cytoplasm to balance the hypersaline environments outside the cell. Moreover, four copies of Na^+/H^+ antiporter proteins related to sodium efflux were found. The Na^+/H^+ antiporter was able to extrude sodium out of cells in exchange for hydrogen ions to keep cytoplasm isoosmotic with the environment and to avoid intoxication of living cells (Yang et al., 2006). Furthermore, the genome contains nine genes related to the synthesis and transport of the compatible-solute glycine betaine for resistance to osmotic stress (Roberts, 2005), which included five choline-sulfatases, three glucose-methanol-choline oxidoreductases and one high-affinity choline uptake protein (BetT). These proteins are related to the metabolic pathway converting choline sulfate to glycine betaine. As well as this, one trehalose synthase gene is also found in the genome of *S. rubrum* GX10^T and trehalose is widely used as a compatible-solute (Cardoso et al., 2007). Therefore, *S. rubrum* GX10^T utilizes "the salt-in strategy" (Singh et al., 2013) as well as "the salt-out strategy" (Roberts, 2005) to resist the osmotic stress outside the cell. These genes may be a key feature of *S. rubrum* GX10^T that allows it to adapt to extremely high salt environments.

The majority of proteins in halophilic archaea have an acidic exterior that increases the hydration shell of the folded protein and prevents "salting out", thus enabling normal enzymatic function in a highly cationic environment (Hartman et al., 2010). According to genomic analysis, most proteins in *S. rubrum* GX10^T have a relatively low isoelectric point (pI). The whole proteome plot against pI shows that *S. rubrum* GX10^T possesses an extremely acidic complement of proteins in its proteome with culminating at 4.2.

Additionally, according to genomic analysis, strain *S. rubrum* GX10^T possesses fourteen heavy metal translocating P-type ATPases, which may be involved in the transport of Pb^{2+}, Cd^{2+}, Zn^{2+} and Hg^{2+} across the cytoplasmic membrane (Choudhury and Srivastava, 2001). Meanwhile, five CRISPR repeat regions were identified on the chromosome, one on plasmid pSRU01 and two on plasmid pSRU02, which indicated that *S. rubrum* GX10^T may have been exposed to potential phage infestations. A halorhodopsin-coding gene and two bacteriorhodopsin-coding genes were uncovered from the genome. The two kinds of proteins were extremely important for salty environment tolerance (Blanck and Oesterhelt, 1987). This is the first report on the presence of the *bop* and *hop* genes in the genus *Salinigranum*, extending the number of genera known to harbor *bop* and *hop* genes and the diversity of bacterioopsins.

In conclusion, the complete genome of *S. rubrum* GX10^T is the first genome sequence reported for the genus *Salinigranum*. This genomic analysis sheds light on understanding adaptive mechanisms of *S. rubrum* GX10^T to hypersaline environments through the "salt-in strategy" or the "salt-out strategy" to maintain osmotic balance. Furthermore, the acidic proteome enable the normal function of the cells in a hypersaline environment. Genes related to heavy metal resistance, CRISPR-Cas system

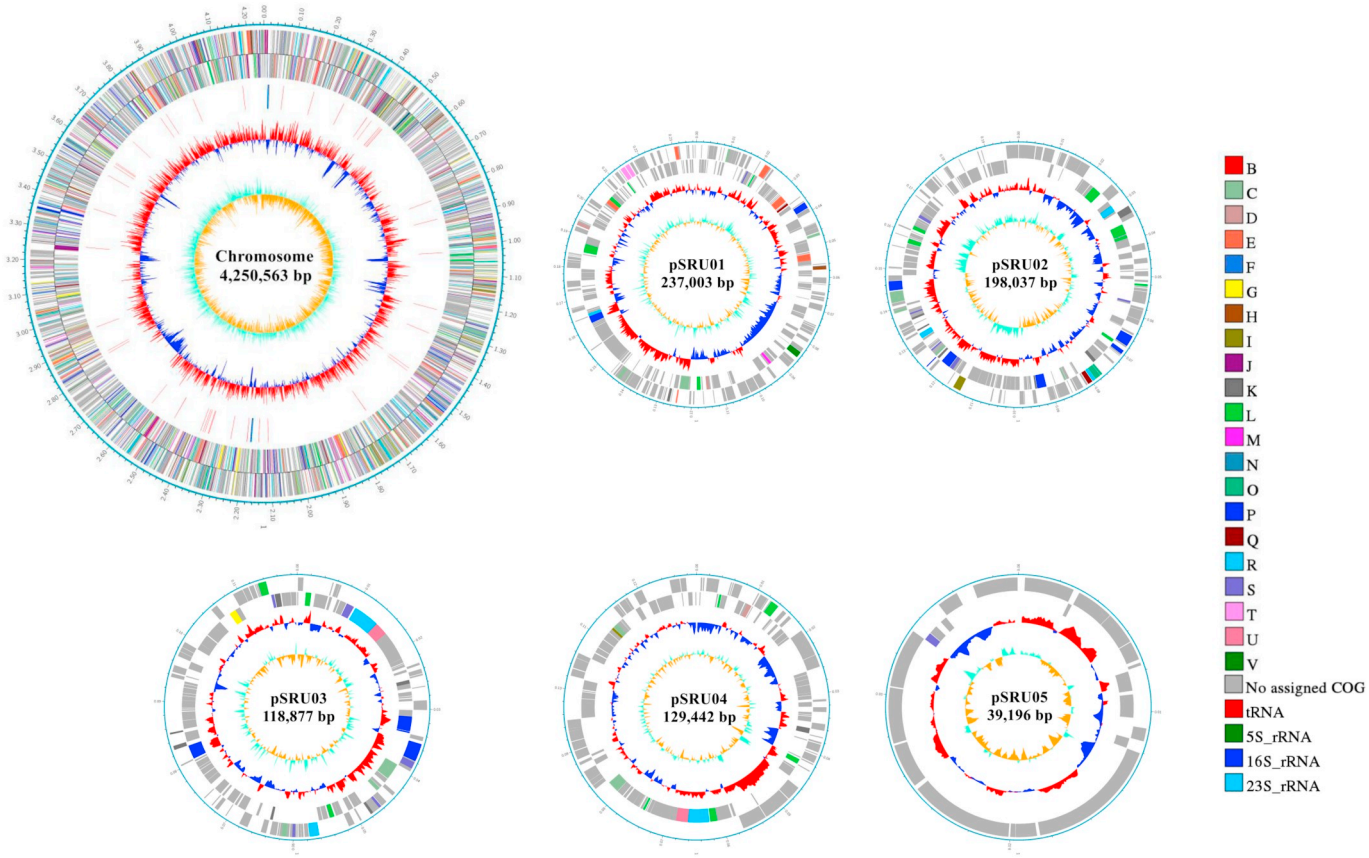

Fig. 1. Circular map of one chromosome and five plasmids in the genome of *Salinigranum rubrum* GX10^T. The tracks from the outside to the center: CDSs on forward strand, CDSs on reverse strand, rRNA genes, tRNA genes, G + C content and GC skew. The COG categories are B: chromatin structure and dynamics; C: energy production and conversion; D: cell cycle control, mitosis and meiosis; E: amino acid transport and metabolism; F: nucleotide transport and metabolism; G: carbohydrate transport and metabolism; H: coenzyme transport and metabolism; I: lipid transport and metabolism; J: translation, ribosomal structure and biogenesis; K: transcription; L: replication, recombination and repair; M: cell wall/membrane/envelope biogenesis; N: cell motility; O: post translational modification, protein turnover and chaperones; P: inorganic ion transport and metabolism; Q: secondary metabolites biosynthesis, transport and catabolism; R: general function prediction only; S: function unknown and T: signal transduction mechanisms; U: intracellular trafficking, secretion and vesicular transport; V: defense mechanisms.

Table 2
Genomic features of *S. rubrum* GX10^T.

Attribute	Chromosome	Plasmid	Plasmid	Plasmid	Plasmid	Plasmid
		pSRU01	pSRU02	pSRU03	pSRU04	pSRU05
Size (bp)	4,250,563	237,003	198,037	118,877	129,442	39,196
G + C content (%)	66.9	57.6	59.2	58.9	57.9	63.3
Total Genes	4379	289	225	138	143	45
Protein coding genes	4327	289	224	138	143	45
tRNA genes	46	0	1	0	0	0
rRNA genes	6	0	0	0	0	0
Genes assigned to COGs	1543	31	37	25	13	1
Genes assigned to KEGG	1761	52	38	31	36	6
CRISPR repeats	5	1	2	0	0	0

and light transform system were also identified. Overall these attributes and specific genes contribute to the robustness of the strain, allowing it to survive well in its extreme habitat. This study contributes to the expansion of our knowledge into genetic mechanisms that enable halophilic archaea to endure hypersaline environments and may facilitate biotechnological applications of the halophilic archaeon.

3. Nucleotide sequence accession number

The complete genome sequence of *S. rubrum* GX10^T has been deposited in GenBank under the accession numbers CP026309 (Chromosome), CP026310 (plasmid pSRU01), CP026311 (plasmid pSRU02), CP026312 (plasmid pSRU03), CP026313 (plasmid pSRU04) and CP026314 (plasmid pSRU05).

Acknowledgement

This work was supported by grants from the Science & Technology Basic Resources Investigation Program of China (2017FY100300) and the Scientific Research Foundation of Zhejiang Sci-Tech University (17042187-Y).

References

Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Yeh, L.S.L., 2004. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 32, D115–D119. https://doi.org/10.1093/nar/gkh131.

Blanck, A., Oesterhelt, D., 1987. The halo-opsin gene. II. Sequence, primary structure of halorhodopsin and comparison with bacteriorhodopsin. EMBO J. 6 (1), 265–273.

Cardoso, F.S., Castro, R.F., Borges, N., Santos, H., 2007. Biochemical and genetic characterization of the pathways for trehalose metabolism in *Propionibacterium freudenreichii*, and their role in stress response. Microbiol. Sgm 153, 270–280. https://doi.org/10.1099/mic.0.29262-0.

Choudhury, R., Srivastava, S., 2001. Zinc resistance mechanisms in bacteria. Curr. Sci. 81 (7), 768–775.

Cui, H.L., 2016. Taxonomy of halophilic archaea: a review. Microbiol. China 43 (5), 1113–1122.

Cui, H.L., Zhang, W.J., 2014. *Salinigranum rubrum* gen. nov., sp. nov., a member of the family *Halobacteriaceae* isolated from a marine solar saltern. Int. J. Syst. Evol. Microbiol. 64, 2029–2033. https://doi.org/10.1099/ijs.0.061606-0. Pt 6.

Delcher, A.L., Bratke, K.A., Powers, E.C., Salzberg, S.L., 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23 (6), 673–679. https://doi.org/10.1093/bioinformatics/btm009.

Grissa, I., Vergnaud, G., Pourcel, C., 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 35 (Web Server issue), W52–W57. https://doi.org/10.1093/nar/gkm360.

Hartman, A.L., Norais, C., Badger, J.H., Delmas, S., Haldenby, S., Madupu, R., Eisen, J.A., 2010. The complete genome sequence of Haloferax volcanii DS2, a Model Archaeon. PLoS One 5 (3) (doi: ARTN e9605).

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., Hattori, M., 2004. The KEGG resource for deciphering the genome. Nucleic Acids Res. 32, D277–D280. https://doi.org/10.1093/nar/gkh063.

Koren, S., Schatz, M.C., Walenz, B.P., Martin, J., Howard, J.T., Ganapathy, G., Adam, M. Phillippy, 2012. Hybrid error correction and *de novo* assembly of single-molecule sequencing reads. Nat. Biotechnol. 30 (7), 693–700. https://doi.org/10.1038/nbt.2280.

Kozlowski, L.P., 2016. IPC - isoelectric point calculator. Biol. Direct 11. https://doi.org/10.1186/s13062-016-0159-9.

Kraegeloh, A., Amendt, B., Kunte, H.J., 2005. Potassium transport in a halophilic member of the *Bacteria* domain: identification and characterization of the K^+ uptake systems TrkH and TrkI from *Halomonas elongata* DSM 2581^T. J. Bacteriol. 187 (3), 1036–1043. https://doi.org/10.1128/Jb.187.3.1036-1043.2005.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Marra, M.A., 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19 (9), 1639–1645. https://doi.org/10.1101/gr.092759.109.

Lagesen, K., Hallin, P., Rodland, E.A., Staerfeldt, H.H., Rognes, T., Ussery, D.W., 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35 (9), 3100–3108. https://doi.org/10.1093/nar/gkm160.

Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25 (5), 955–964. https://doi.org/10.1093/Nar/25.5.955.

Naghoni, A., Emtiazi, G., Amoozegar, M.A., Cretoiu, M.S., Stal, L.J., Etemadifar, Z., Bolhuis, H., 2017. Microbial diversity in the hypersaline Lake Meyghan, Iran. Sci. Rep. 7 (1), 11522. https://doi.org/10.1038/s41598-017-11585-3.

Nazareth, S.W., Gonsalves, V., 2014. Halophilic *Aspergillus penicillioides* from athalassohaline, thalassohaline, and polyhaline environments. Front. Microbiol. 5, 412. https://doi.org/10.3389/fmicb.2014.00412.

Oren, A., 2002. Diversity of halophilic microorganisms: Environments, phylogeny, physiology, and applications. J. Ind. Microbiol. Biotechnol. 28 (1), 56–63. https://doi.org/10.1038/sj/jim/7000176.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Stevens, R., 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 42 (Database issue), D206–D214. https://doi.org/10.1093/nar/gkt1226.

Roberts, M.F., 2005. Organic compatible solutes of halotolerant and halophilic microorganisms. Saline Syst. 1, 5. https://doi.org/10.1186/1746-1448-1-5.

Singh, S.P., Raval, V., Purohit, M.K., 2013. Strategies for the salt tolerance in bacteria and archeae and its implications in developing crops for adverse conditions. Plant Acclimation to Environ. Stress 85–99. https://doi.org/10.1007/978-1-4614-5001-6_4.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., Koonin, E.V., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28 (1), 33–36. https://doi.org/10.1093/Nar/28.1.33.

Wang, Z., Xu, J.Q., Xu, W.M., Li, Y., Zhou, Y., Lu, Z.Z., Cui, H.L., 2016. *Salinigranum salinum* sp. nov., isolated from a marine solar saltern. Int. J. Syst. Evol. Microbiol. 66 (8), 3017–3021. https://doi.org/10.1099/ijsem.0.001138.

Yang, L.F., Jiang, J.Q., Zhao, B.S., Zhang, B., Feng, D.Q., Lu, W.D., Yang, S.S., 2006. A Na^+/H^+ antiporter gene of the moderately halophilic bacterium *Halobacillus dabanensis* D-8^T: cloning and molecular characterization. FEMS Microbiol. Lett. 255 (1), 89–95. https://doi.org/10.1111/j.1574-6968.2005.00055.x.