Front. Biol. China 2008, 3(4): 392–396
DOI 10.1007/s11515-008-0067-1

RESEARCH ARTICLE

Genetic analysis of the *br* gene in halophilic archaea isolated from Xinjiang region, China

Xiaohong XU[1], Min WU (✉)[2], Huibin ZHANG[3], Zhihu LIU[3]

1 Ningbo Institute of Technology, Zhejiang University, Ningbo 315100, China
2 College of Life Sciecne, Zhejiang University, Hangzhou 310027, China
3 The Altun Mountain National Nature Reserve, Kuerle 841000, China

Abstract Some novel members of extremely halophilic archaea, strains AJ11, AJ12 and AJ13, were isolated from the Aularz Lake located in the Altun Mountain National Nature Reserve of Xinjiang, Uygur Autonomous Region in China. Partial DNA fragments encoding a bacteriorhodopsin (BR), as well as for 16S rRNA of isolated strains, were amplified by PCR and their DNA sequences were determined subsequently. On the basis of homology and phylogenetic analysis of the 16S rDNA, we thought that the isolated strains forming a microbiological population are the members of the genus *Natrinema*. The results of genetic analysis, such as GC content, transition/transversion (Ti/Tv) rate ratios and synonymous substitution rates (*K*s) indicate that the *br* fragments, with a high level of genetic divergence, are faced with both purifying selection and bias mutation pressure. The study provides the basis for use of species and BR proteins resources.

Keywords halophilic archaea, 16S rDNA, phylogeny, bacteriorhodopsin, genetic analysis

1 Introduction

Extremely halophilic archaea belonging to the order *Halobacteriales* have been isolated from various hypersaline environments such as the Dead Sea, the Great Salt Lake and natural or artificial salterns (Radax et al., 2001). In the western part of China, there are many natural salt lakes, which could provide abundant extremely halophilic archaea resources. However, most of these salt lakes, especially in the Altun Mountain, have not been exploited because of their dreadful environment and inaccessible position. Bacteriorhodopsin (BR), which has been isolated from some extremely halophilic archaea, is a retinal-binding membrane protein with seven membrane-spanning segments and functions as a light-driven proton pump (Oesterhelt and Stoeckenius, 1971; Luecke et al., 1999; Xu and Wu, 2004). Despite an important function gene, few molecular genetic studies of the *br* gene have been reported (Ihara et al., 1999; Xu et al., 2004).

In this study, we isolated some novel strains of extremely halophilic archaea from the Altun Mountain, and amplified the partial DNA fragments encoding BR proteins from helix C to helix G as well as for 16S rRNA. The phylogenetic tree was constructed based on the 16S rDNA sequence and genetic analysis was performed using the partial *br* fragments.

2 Materials and methods

2.1 Archaeal strains and growth media

Strains AJ11, AJ12 and AJ13 were isolated from water samples, which were collected from the Aularz Lake located in the Altun Mountain National Nature Reserve in western China. Reference strain *Halobacterium salinarium* (*halobium*) R1M1 was a gift from QingGuo LI (Shanghai, China). The strains were routinely grown aerobically at 37°C in Rich medium containing (per liter) 250 g NaCl, 20 g $MgSO_4 \cdot 7H_2O$, 3 g trisodium citrate dihydrate, 2 g KCl, 0.26 g $CaCl_2$, 10 g peptone (Oxiod L37). The pH was adjusted to 7.0 with NaOH and slants and plates were prepared by adding 20 g agar per liter. (Oesterhelt and Stoeckenius, 1974).

2.2 DNA extraction

The extraction procedure was according to a previous report by Li et al. (2001) with some modifications. The

---

Translated from *Hereditas* (*Beijing*), 2007, 29(3): 376–380 [译自: 遗传]

E-mail: wumin@cls.zju.edu.cn

cells were suspended by adding 250 μL of buffer I (20 mM Tris-HCl, 100 mM EDTA) before treatments with freeze-thaw cycles and proteinase K. DNA extraction was carried out by adding 1 volume phenol and 1 volume chloroform. After extraction, DNA was precipitated with 3 M NaAc (1/10 of the final volume) and cold ethanol (twice the volume), centrifuged, dried, and finally dissolved for spectrophotometric quantification.

2.3 16S rDNA sequence PCR amplification

The 16S rDNA sequence primers, designed according to Gupta et al. (1983), were rDNA22F (5′-ATTCCGGTTGATCCTGC-3′, positions 6–22) and rDNA1521R (5′-AGGAGGTGATCCAGCCGCAG-3′, positions 1540–1521). The *E. coli* 16S rDNA sequence was used as the position reference. The PCR thermal cycling conditions were as follows: 25 cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 2 min.

2.4 Cloning and sequencing of PCR products

The desired PCR products were purified using the UNIQ-10 column DNA gel extraction kit (Sangon) according to the instructions of the manufacturer. The purified fragments were treated with T4 DNA ligase (MBI), then ligated to pUCm-T (Sangon) and transformed into *Escherichia coli* DH5α. White colonies on LB agar plates (containing 100 μg/mL ampicillin, IPTG and X-Gal) were isolated and the plasmids were extracted using the alkaline lysis method. PCR protocols, as described above, were performed using approximately 50 ng plasmid DNA as the template to verify the presence of a correct insert. Sequences were determined using automated dideoxynucleotide methods with the ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction kit (Perkin Elmer) on an ABI Prism 377XL DNA sequencer. The M13 universal sequencing primers (M13F: 5′-GTAAAACGACGGCCAGT-3′, and M13R: 5′-GGAAACAGCTATGACCATG-3′) were used for bidirectional sequencing.

2.5 Phylogenetic tree construction

The 16S rDNA sequences of strains AJ11, AJ12 and AJ13 were fitted in the halophilic archaeal alignment with Clustal X 1.8 and a few minor corrections were made manually (Xu et al., 2004). Analyses were carried out with PHYLIP (version 3.6a3). The DNADIST program was used to construct a Kimura 2-parameter evolutionary distance matrix and then a phylogenetic tree was constructed using the neighbor-joining method with randomized input order. Bootstrap analysis (1000 replications) was performed using the additional programs SEQBOOT and CONSENSE.

2.6 Identification and genetic analysis of *br* gene

The *br* PCR primers, designed on the basis of the conserved protein sequences PLLLLDL and KVGFGFI (Otomo et al., 1992), were br331F: 5′-CCGCTG(CT)TG(CT)TG(CT)T(AC)GACCTCG-3′, positions 310–331 and br686R: 5′-AGGATGA(GA)(CG)CCGAA(CG)CCGACCTT-3′, positions 707–686. The *Hbt.salinarium* *br* gene sequence was used as the position reference. The PCR amplification was performed in a 100 μL reaction mixture for 25 cycles of denaturation at 94°C for 1 min, primer annealing at 60°C for 1 min and extension at 72°C for 1 min. PCR products were examined by gel electrophoresis in 4% agarose gel using the GeneRuler TM 100 bp DNA ladder plus markers (Sangon) for size comparisons of the PCR products. Cloning and sequencing of PCR products protocols, as described above, was used. The GC content and transition/transversion (Ti/Tv) rate ratios were estimated in the region of *br* gene by Mega (version 2.1) software and manual calculation. Synonymous substitution rates (*K*s) and nonsynonymous substitution rates (*K*a) were estimated using K-estimated software (version 6.0) (Comeron, 1999).

3 Results

3.1 Phylogenetic analysis based on 16S rDNA sequence

The 16S rDNA sequences of strains AJ11, AJ12 and AJ13, which contain 1474, 1473, 1473 residues respectively, were determined. The GenBank accession numbers are AY570916, AY570917 and AY570918, respectively. Phylogenetic tree based on 16S rDNA sequence was constructed using the neighbor-joining method with bootstrap values calculated from 1000 trees subsequently (Fig. 1). On the phylogenetic tree the isolated strains clustered with the three described *Natrinema* species (McGenity et al., 1998; Xin et al., 2000). Similarity percentages of 16S rDNA sequences among the isolated strains and the type strains of *Natrinema* are listed in Table 1.

3.2 Detection of *br* gene by PCR

The nucleotide sequences of the partial *br* gene from strains AJ11, AJ12 and AJ13 (GenBank accession No. AY570919, AY570920 and AY570921), with 401 bp in total, were sequenced. Despite the difference between the *br* gene fragments from the isolated strains, the deduced amino acid residues are completely identical. Compared to BR protein sequences from different strains, about 97% of amino acid residues are identical between strain AJ11 and arg-4, while less than 60% homology between AJ11 and other reported strains. The result confirmed that the

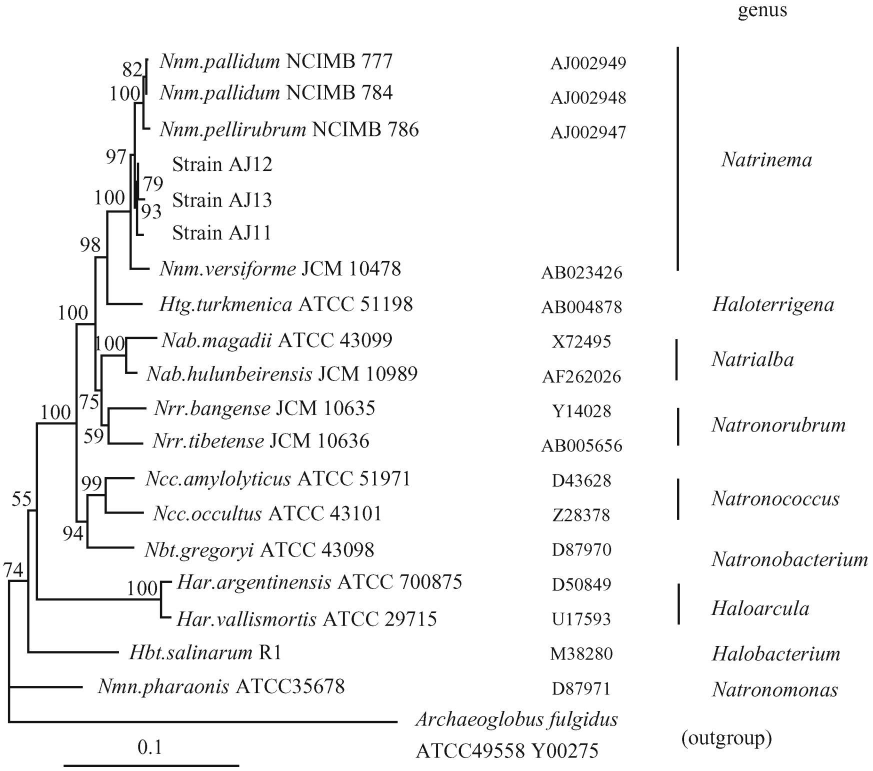

Fig. 1 Phylogenetic tree based on the 16S rDNA sequences from halobacterial strains. Note: The 16S rDNA sequence from *Archaeoglobus flugidus* was used as the outgroup. The tree was constructed using the neighbor-joining method with bootstrap values calculated from 1000 trees. The number at each branch point represents the percentage bootstrap support. The bar at the bottom represents 0.1 *K*nu.

Table 1 Similarity of l6S rDNA sequence of extremely halophilic archaea strains (%)

	1	2	3	4	5	6	7	8
1. *Nnm.pallidum* NCIMB 777		–	–	–	–	–	–	–
2. *Nnm.pallidum* NCIMB 784	98.7		–	–	–	–	–	–
3. *Nnm.pellirubrum*	98.0	97.7		–	–	–	–	–
4. *Nnm.versiforme*	96.8	96.8	96.4		–	–	–	–
5. Strain AJ11	97.7	97.8	97.3	98.1		–	–	–
6. Strain AJ12	97.9	98.0	97.6	98.5	99.7		–	–
7. Strain AJ13	97.3	97.7	97.3	98.2	99.5	99.7		–
8. *Htg.turkmenia*	94.5	94.7	94.3	95.6	96.1	96.4	96.0	
9. *Hbt.salinarum*	87.7	87.8	87.5	89.3	89.3	89.3	89.1	88.1

nucleotide fragments from PCR reaction are just the part of *br* gene that encodes the BR protein from helix C to helix G.

3.3 Genetic analysis of the partial *br* fragments

According to all reported *br* fragments from helix C to helix G, the GC contents at different codon positions were calculated and curved (Fig. 2). All the *br* fragments obtained from 16 different halophilic archaea were rich in G+C. The GC contents varied from 60.1% to 67.9%, with an average of 64.6%. The average of GC contents at third codon position was about 90%, and the marked increasing trend among three GC contents at different codon positions in all *br* genes is as follows: the second, first, and third position. The results indicate that the *br* gene is faced with strong evolution force.

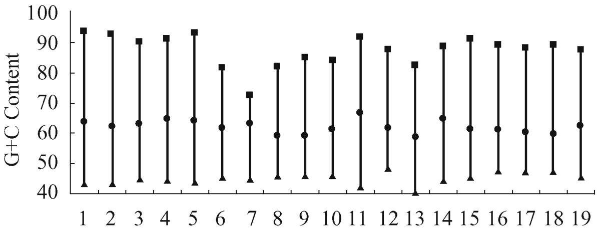

Fig. 2 GC content of different codons in *br* gene segments. Note: •: 1st Codon; ▲: 2nd Codon; ■: 3 rd Codon. Strains and references: 1: *Hrr*.sp.aus-1 (Mukohata et al., 1988); 2, 12: *Hrr. sodomense*, *Htg*.sp.arg-4 (Ihara et al., 1999); 3: *Hrr*. sp.aus-2 (Uegaki et al., 1991); 4: XZ515 (Wang et al., 2000); 5, 7, 9:mex, shark, port (Otomo et al., 1992); 6: *Har. vallismor-tis* (Kitajima et al., 1996); 8: *Har.argentinensis* (Tateno et al., 1994); 10: *Har. japonica* (Yatsunami et al., 2000); 11: *Har. muko-hataei* (Sugiyama et al., 1994); 13: *Hbt.salinarum* (Dunn et al., 1981); 14: *Har*. sp. AJ4 (Xu et al., 2004); 15: *Nnm*. sp. AB3 (Yu et al., 2005); 16–18: AJ11, AJ12, AJ13; 19: The average.

The average of transition/transversion (Ti/Tv) rate ratios, which were estimated using 16 partial *br* fragments, was about 0.57. The results indicate that there was low bias towards either transitional or transversional substitution. Among different genera of *Halobacteriaceae*, four have been found to possess light-driven proton pump protein. According to the origin these bR proteins have been subdivided into four kinds: archaerhodopsin (aR) from *Halorubrum*, bR from *Halobacterium*, cruxrhodopsin (cR) from *Haloarcula* and deltarhodopsin (dR) from *Haloterrigena* (Xu et al., 2004; Mukohata et al., 1999). Synonymous substitution rates (*K*s) and nonsynonymous substitution rates (*K*a) in the coding regions of the four different kinds of *br* genes were subsequently calculated according to the method of Kimura 2-parameter (Table 2). The result indicates that *br* gene bears a high level of genetic divergence.

Table 2 Nonsynonymous substitution rates (*Ka*), synonymous substitution rates (*Ks*) and the ratio between them in different kinds of *br* gene segments

	AJ11 *br* vs. *ar*	AJ11 *br* vs. *br*	AJ11 *br* vs. *cr*	AJ11 *br* vs. *dr*
*K*a	0.387	0.408	0.459	0.018
*K*s	1.148	0.854	0.830	0.193
*K*a/*K*s	0.337	0.478	0.553	0.093

4 Discussion

The strains AJ11, AJ12 and AJ13 clustered with each other in the phylogenetic tree based on 16S rDNA sequence, and the similarity between their 16S rDNA sequences is more than 99.5%, suggesting the strains isolated from the Aularz Lake form a microbiological population. The results that the isolated strains clustered with the three described *Natrinema* species in the phylogenetic tree and the cluster was supported by a high bootstrap value (97%) indicating that the isolated strains are the member of *Natrinema*. There are abundant *Natrinema* species resources in the salt lakes of Xinjiang region, China, which indicates the regional distribution of this type of strains (Yu et al., 2005; Xu et al., 2005).

The substitution occurring at the second position is an entirely nonsynonymous substitution and *br* is a highly expressed function gene. Therefore, its GC content is mainly affected by purifying selection and appears to be low. However the substitution occurring at the third position is an almost synonymous substitution. Therefore, the GC content is mostly determined by biased mutation pressure and is high (90%). The substitution occurring at the first position is a slightly synonymous substitution, thus the GC content at the first position is between that of the second and third position. Obviously the GC content is determined by both purifying selection and bias mutation pressure. Ng et al. (2000) reported that the whole genome GC content in the strain *Halobacterium* NRC-1 is 67.9%, similar to the GC contents in *br* fragments (Stackebrandt and Goebel, 1994; Ng et al., 2000). The result that there are similar GC contents in *br* genes implies halophilic archaea, apart from *Halobacterium* NRC-1, possess the GC bias.

Ihara et al. (1999) reported that the genetic divergence time of *br* gene is long, based on the use of the 16S rDNA sequences from different halophilic archaea possessing *br* gene as a chronometer. Therefore, the low Ti/Tv ratio result is consistent with the frequent reports that at high levels of genetic divergence, Ti/Tv appears to be low (Yang and Yoder, 1999). Comparison of the *br* fragments from strain AJ11 to other genus strains indicates that nucleotide substitutions at synonymous sites are almost saturated, because the GC content at the third position is about 90% so that Ks is close to 1.0. Because *K*s value is relative to the genetic divergence, the large *K*s value between *br* genes from AJ11 and other strains in genus of *Halobacterium*, *Haloarcula* and *Halorubrum* indicates AJ11 has a high genetic divergence from others. On the contrary, the small *K*s value between *br* genes from AJ11 and *Htg*. sp. arg-4 indicates there is a low genetic divergence between them. The low values of *K*a/*K*s ratio also imply a strong selective constraint against the nonsynonymous sites of the *br* gene from halophilic archaea.

At present, five genus of halophilic archae possess the *br* gene: *Halorubrum*, *Halobacterium*, *Haloarcula*, *Haloterrigena* and *Natrinma*. Although isolated strains belong to *Natrinma*, according to the genetic analysis, the *br* genes from strains AJ11, AJ12 and AJ13 should be classified into dR but not a new type. This study provides the basis for use of species and BR proteins resources.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 30370029) and the Scientific Research Start-up Foundation of Ningbo Institute of Technology, Zhejiang University (No.1140257G413).

References

Comeron J M (1999). K-estimator: calculation of the number of nucleotide substitutions per site and the confidence intervals. Bioinformatics, 15(9): 763–764

Dunn R, McCoy J, Simsek M, Majumdar A, Chang S H, Rajbhandary U L, Khorana H G (1981). The bacteriorhodopsin gene. Proc Natl Acad Sci USA, 78(11): 6744–6748

Gupta R, Lanter J M, Woese C R (1983). Sequence of the ribosomal RNA from *Halobacterium volcanii*, an archaebacte-rium. Science, 221: 656–659

Ihara K, Umemura T, Katagiri I, Kitajima-Ihara T, Sugiyama Y, Kimura Y, Mukohata Y (1999). Evolution of the archaeal rhodopsins: evolution rate changes by gene duplication and functional differentiation. J Mol Biol, 285(1): 163–174

Kitajima T, Hirayama J, Ihara K, Sugiyama Y, Kamo N, Mukohata Y (1996). Novel bacterial rhodopsins from *Haloarcula vallismortis*. Biochem Biophys Res Commun, 220(2): 341–345

Li L, Wu M, Qiao S Y, Chen K S (2001). Cloning and expression of the bacteriorhodopsin gene. Journal of Zhejiang University (Engineering Science), 35(3): 324–327 (in Chinese)

Luecke H, Schobert B, Richter H T, Cartailler J P, Lanyi J K (1999). Structural changes in bacteriorhodopsin during ion transport at 2 angstrom resolution. Science, 286(5438): 255–260

McGenity T J, Gemmell R T, Grant W D (1998). Proposal of a new halobacterial genus *Natrinema* gen. nov., with two species *Natrinema pellirubrum* nom. nov. and *Natrinema pallidum* nom. nov. Int J Syst Bacteriol, 48(4): 1187–1196

Mukohata Y, Ihara K, Tamura T, Sugiyama Y (1999). Halobacterial rhodopsin. J Biochem, 125(4): 649–657

Mukohata Y, Sugiyama Y, Ihara K, Yoshida M (1998). An Australian *Halobacteria* contains a novel proton pump retinal protein: archaerhodopsin. Biochem Biophys Res Commun, 151(3): 1339–1345

Ng W V, Kennedy S P, Mahairas G G, Berquist B, Pan M, Shukla H D, Lasky S R, Baliga N S, Thorsson V, Sbrogna J, Swartzell S, Weir D, Hall J, Dahl T A, Welti R, Goo Y A, Leithauser B, Keller K, Cruz R, Danson M J, Hough D W, Maddocks D G, Jablonski P E, Krebs M P, Angevine C M, Dale H, Isenbarger T A, Peck R F, Pohlschroder M, Spudich J L, Jung K W, Alam M, Freitas T, Hou S, Daniels C J, Dennis P P, Omer A D, Ebhardt H, Lowe T M, Liang P, Riley M, Hood L, DasSarma S (2000). Genome sequence of *Halobacterium* species NRC-1. Proc Natl Acad Sci USA, 97(22): 12176–12181

Oesterhelt D, Stoeckenius W (1971). Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. Nature, 233(39): 149–152

Oesterhelt D, Stoeckenius W (1974). Isolation of the cell membrane of *Halobacterium halobium* and its fraction into red and purple membrane. Methods Enzymol, 31: 667–678

Otomo J, Urabe Y, Tomioka H, Sasabe H (1992). The primary structures of helices A to G of three new bacteriorhodopsin-like retinal proteins. J Gen Microbiol, 138(11): 2389–2396

Radax C, Gruber C, Stan-Lotter H (2001). Novel haloarchaeal 16S rRNA gene sequences from Alpine Permo-Triassic rock salt. Extremophiles, 5(4): 221–228

Stackebrandt E, Goebel B M (1994). Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Bacteriol, 44(4): 846–849

Sugiyama Y, Yamada N, Mukohata Y(1994). The light-driven proton pump, cruxrhodopsin-2 in *Haloarcula* sp.arg-2 (bR$^+$, hR$^-$), and its coupled ATP formation. Arch Biochem Biophys, 1188(3): 287–292

Tateno M, Ihara K, Mukohata Y(1994). The novel ion pump rhodopsins from *Haloarcula* form a family independent from both the bacteriorhodopsin and archaerhodopsin families/tribes. Arch Biochem Biophys, 315(1): 127–132

Uegaki K, Sugiyama Y, Mukohata Y (1991). Archaerhodopsin-2, from *Halobacterium* sp. aus-2 further reveals essential amino acid residues for light-driven proton pumps. Arch Biochem Biophys, 286(1): 107–110

Wang H, Zhan S X, SunN Q A, Xu D Q, Zhao W, Huang W D, Li Q G (2000). Primary structure of helix C to helix G of a new retinal protein in *H.* sp. xz515. Chin Sci Bull, 45(12): 1108–1112

Xin H, Itoh T, Zhou P, Suzuki K, Kamekura M, Nakase T (2000). *Natrinema versiforme* sp. nov., an extremely halophilic archaeon from Aibi salt lake, Xinjiang, China. Int J Syst Evol Microbiol, 50(3): 1297–1303

Xu X H, Wu M (2004). Spectral analysis and application of photoelectric characteristics in bacteriorhodopsin. Journal of Zhejiang University (Science Edition), 31(5): 570–574 (in Chinese)

Xu X H, Wu M, Cao Y, Wu Y H (2005). Isolation and phylogenetic analysis of halophilic archaea AJ6. Journal of Zhejiang University (Science Edition), 32(1): 83–88 (in Chinese)

Xu X W, Ren P G, Liu S J, Wu M, Zhou P J (2005). *Natrinema altunense* sp. nov., an extremely halophilic archaeon isolated from a salt lake in Altun Mountain in Xinjiang, China. Int J Syst Evol Microbiol, 55: 1311–1314

Xu X W, Wu M, Huang W D (2004). Isolation and 16S rDNA sequence analysis of extremely halophilic archaea from saline in Altun region. Journal of Zhejiang University (Engineering Science), 38(4): 508–512 (in Chinese)

Xu X W, Wu M, Ruan H, Hong J, Huang W D (2004). Partial Sequence of the gene for a new BR protein, and nucleotide sequence of 16S rRNA encoding gene from a halophilic archaeum AJ4. Chinese Journal of Biochemistry and Molecular Biology, 20(1): 55–60 (in Chinese)

Xu X W, Wu M, Ruan H, Huang W D, Dilbar Tohty (2004). Sequencing and analysis partial sequence of the gene for a novel Br protein. Hereditas (Beijing), 26(3): 343–348 (in Chinese)

Yang Z, Yoder A D (1999). Estimation of the transition/transversion rate bias and species sampling. J Mol Evol, 48(3): 274–283

Yatsunami R, Kawakami T, Ohtani H, Nakamura S (2000). A novel bacteriorhodopsin-like protein from *Haloarcula japonica* strain TR-1: gene cloning, sequencing, and transcript analysis. Extremophiles, 4(2): 109–114

Yu X W, Dilbar Tohty, Wu Y H, Wu M (2005). Bacteriorhodopsin gene and phylogenetic analysis on halophilic archaea AB3. Journal of Zhejiang University (Engineering Science), 39(3): 418–422 (in Chinese)