Austin Journal of Proteomics, Bioinformatics & Genomics

Austin Publishing Group

Open Access Full Text Article

Research Article

Comparative Genomic Analysis of Halophiles Reveals New Clues to Their Adaptation Strategies in Hypersaline Environments

Shaoxing Chen[1,2]*, Jian Yang[1], Yanhong Liu[1], Chuangming Wang[1] and Zhu L Yang[2]

[1]University Key Laboratory of Crop High Quality and High Effective Cultivation and Safety Control in Yunnan Province, Honghe University, Mengzi 661100, China

[2]Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China

***Corresponding author:** Shaoxing Chen, Key Laboratory of Biodiversity and Biogeography, Kunming Institute of Botany, Chinese Academy of Sciences, 132# Lanhei Road, Heilongtan, Kunming 650201, Yunnan, China

Received: July 30, 2014; **Accepted:** September 27, 2014; **Published:** September 29, 2014

Abstract

Background: Halophiles, which have many potential applications in the biomaterial, bioremediation, and nanotechnology arenas, are microorganisms that live in environments with high salt concentrations. To elucidate the adaptive strategies that allow them to live in such hypersaline environments, the genome sequences of 66 strains of halophiles and non-halophiles (including 27 strains of halophilic bacteria, 24 strains of haloarchaea, and 15 strains of non-halophilic bacteria) were subjected to comparative genomic analysis.

Results: The G+C content of the genomic DNA sequence and acidic amino acid composition of the gene product of the haloarchaea were higher than those of both the halophilic and non-halophilic bacteria. In addition, the probability of occurrence and proportion of extra chromosomal genetic elements in the haloarchaea outweighed those of the halophilic and non-halophilic bacteria. Further, proteasome, the mRNA surveillance pathway, and basal transcription factors were present in the haloarchaea but absent in the other two groups of microorganisms. Carotenoid, sesqui-terpenoid, and tri-terpenoid were common in the haloarchaea, but occurred with a relatively low degree of frequency in the halophilic and non-halophilic bacteria. In contrast, some D-amino acids (i.e., D-glutamine, D-glutamate, D-arginine, D-ornithine, and D-alanine) and lipopolysaccharide, fluorobenzoate, limonene, and pinene were widely distributed in both types of bacteria, but absent in the family Halobacteriaceae.

Conclusion: Large-scale comparative genomic analysis of the genomes of haloarchaea, halophilic bacteria, and non-halophilic bacteria provided a novel perspective on the strategies that microorganisms adopt to adapt to hypersaline environments. Although both haloarchaea and halophilic bacteria require a high concentration of sodium chloride for growth, they employ different mechanisms of adaptation. Haloarchaea, which contain a significantly high G+C content and proportion of acidic amino acids to with stand their harsh environment, use sun light as an energy resource to balance intracellular and extracellular osmotic pressure, thus allowing them to live in hypersaline environments the same way that non-halophilic bacteria live in more common environments.

Keywords: Haloarchaea; Halobacteriaceae; Halophilic bacteria; Hypersaline environment; Carotenoids

Introduction

Halophiles, including haloarchaea and halophilic bacteria, are commonly found in salt lakes, salt mines, saline soils, artificial salterns, heavily salted hides, meats, fishes, and sauces with a high concentration of sodium chloride (NaCl) [1-3]. They have a number of useful applications in biotechnological and biomedical research [4]. Most halophiles use organic solutes to provide an osmotic balance between their cytoplasm and the surrounding medium [5].

The first genome-sequenced organism of haloarchaea is *Halobacterium* sp. NRC-1, which gave researchers an opportunity to probe the mechanisms of adaptation to hypersaline brine [6,7]. A surprising finding was that the overwhelming majority of predicted proteins were highly acidic, with a pI mode of 4.2, and very few neutral or basic proteins [8,9]. In contrast, the predicted proteins of most other non-haloarchaeal and bacterial organisms had equal fractions of acidic and basic components. The implication is that an increase in protein acidity and GC-bias in the genome is an important factor in tolerance to extreme salinity. The negatively charged residues in the haloarchaeal proteins were predominantly found at the protein surface and predicted to function as enhancers of solubility and stability in environments with high salt concentrations [10-12].

An additional characteristic observed in most haloarchaeal genomes is the presence of large megaplasmids or minichromosomes that often harbor important or even essential genes [13]. Analyses of the gene content of these large extra chromosomal elements have resulted in the discovery of expanded gene families for replication and transcription initiation [14], a variety of genes involved in cell survival, e.g., an aminoacyl transfer RNA (tRNA) synthetase [7], arsenic resistance [15], and the production of buoyant gas vesicles [7].

Citation: Chen S, Yang J, Liu Y, Wang C and Yang ZL. Comparative Genomic Analysis of Halophiles Reveals New Clues to Their Adaptation Strategies in Hypersaline Environments. Austin J Proteomics Bioinform & Genomics. 2014;1(1): 11.

Table S1: General genomic characteristics of strains.

No	Full name	T number	Original DB	Accession numbers	Number of nucleotides (bp)	Number of coding genes	Number of amino acids (aa)
1	*Nitrosococcus halophilus*	T01198	JGI	NC_013960	4,145,260	3,817	1,163,230
2	*Halorhodospira halophila*	T00462	JGI	NC_008789	2,678,452	2,407	815,991
3	*Halothiobacillus neapolitanus*	T01104	JGI	NC_013422	2,582,886	2,357	761,013
4	*Chromohalobacter salexigens*	T00347	JGI	NC_007963	3,696,649	3,298	1,099,708
5	*Halomonas elongata*	T01311	Max-Planck	NC_014532	4,061,296	3,474	1,171,559
6	*Desulfohalobium retbaense*	T00989	JGI	NC_013223	2,909,567	2,526	811,068
7	*Halobacillus halophilus*	T02031	Max-Planck	NC_017668	4,170,008	4,126	1,135,727
8	*Pelagibacterium halotolerans*	T01645	Zhejiang U	NC_016078	3,948,887	3,881	1,167,725
9	*Bacillus halodurans*	T00039	JAMSTEC	NC_002570	4,202,352	4,065	1,189,286
10	*Tetragenococcus halophilus*	T01640	NITE	NC_016052	2,562,720	2,555	734,179
11	*Desulfitobacterium dehalogenans*	T02146	JGI	NC_018017	4,321,753	4,011	1,211,838
12	*Dehalobacter* sp. DCA	T02278	U Toronto	NC_018866	3,069,953	2,978	871,612
13	*Dehalobacter* sp. CF	T02279	U Toronto	NC_018867	3,092,048	2,980	878,404
14	*Halothermothrix orenii*	T00838	JGI	NC_011899	2,578,146	2,342	742,032
15	*Acetohalobium arabaticum*	T01290	JGI	NC_014378	2,469,596	2,282	704,792
16	*Halobacteroides halobius*	T02401	JGI	NC_019978	2,649,255	2,468	768,933
17	*Corynebacterium halotolerans*	T02472	Bielefeld U	NC_020302	3,222,008	2,865	905,617
18	*Halothece* sp. PCC 7418	T02379	JGI	NC_019779	4,179,170	3,708	1,181,058
19	*Dehalococcoides ethenogenes*	T00223	JCVI-CMR	NC_002936	1,469,720	1,580	438,843
20	*Dehalococcoides* sp. CBDB1	T00273	Max-Planck	NC_007356	1,395,502	1,458	417,476
21	*Dehalococcoides* sp. BAV1	T00518	JGI	NC_009455	1,341,892	1,371	400,087
22	*Dehalococcoides* sp. VS	T01136	JGI	NC_013552	1,413,462	1,439	424,784
23	*Dehalococcoides mccartyi* BTF08	T02475	UFZ	NC_020387	1,452,335	1,529	436,722
24	*Dehalococcoides mccartyi* DCMB5	T02476	UFZ	NC_020386	1,431,902	1,477	427,122
25	*Dehalogenimonas lykanthroporepellens*	T01264	JGI	NC_014314	1,686,510	1,659	489,618
26	*Methanohalophilus mahii*	T01211	JGI	NC_014002	2,012,424	1,987	589,941
27	*Methanohalobium evestigatum*	T01261	JGI	NC_014253	2,406,232	2,254	626,180
28	*Halobacterium* sp. NRC-1	T00038	U.Maryland	NC_002607	2,571,010	2,622	594,376
29	*Halobacterium salinarum* R1	T00662	Max-Planck	NC_010364	2,668,776	2,749	603,761
30	*Haloarcula marismortui*	T00211	ISB	NC_006396	4,274,642	4,243	899,535
31	*Haloarcula hispanica*	T01597	CAS	NC_015948	3,890,005	3,859	872,570
32	*Haloquadratum walsbyi* DSM 16790	T00375	Max-Planck	NC_008212	3,179,361	2,646	777,797
33	*Haloquadratum walsbyi* C23	T01942	Max-Planck	NC_017459	3,260,476	2,743	798,092
34	*Natronomonas pharaonis*	T00279	Max-Planck	NC_007426	2,749,696	2,820	781,219
35	*Natronomonas moolapensis*	T02478	Max-Planck	NC_020388	2,912,573	2,749	816,750
36	*Halorubrum lacusprofundi*	T00856	JGI	NC_012029	3,692,576	3,560	787,569
37	*Halorhabdus utahensis*	T00971	JGI	NC_013158	3,116,795	2,998	911,740
38	*Halomicrobium mukohataei*	T00982	JGI	NC_013202	3,332,349	3,349	898,173
39	*Haloterrigena turkmenica*	T01154	JGI	NC_013743	5,440,782	5,113	1,087,273
40	*Natrialba magadii*	T01176	JGI	NC_013922	4,443,643	4,212	1,046,649
41	*Haloferax volcanii*	T01200	JCVI	NC_013967	4,012,900	4,015	817,712
42	*Haloferax mediterranei*	T02129	CAS	NC_017941	3,904,707	3,863	837,224
43	*Halalkalicoccus jeotgali*	T01271	Kyung Hee U	NC_014297	3,698,650	3,873	827,912
44	*Halogeometricum borinquense*	T01359	JGI	NC_014729	3,944,467	3,898	820,148
45	*Halopiger xanaduensis*	T01532	JGI	NC_015666	4,355,268	4,221	1,048,160
46	*Natrinema* sp. J7-2	T02184	Nankai China	NC_018224	3,793,615	4,302	1,053,462
47	*Natrinema pellirubrum*	T02425	JGI	NC_019962	4,354,100	4,199	1,055,529
48	*Natronobacterium gregoryi*	T02383	JGI	NC_019792	3,788,356	3,656	1,037,158
49	*Halovivax ruber*	T02402	JGI	NC_019964	3,223,876	3,099	914,088
50	*Natronococcus occultus*	T02426	JGI	NC_019974	4,314,118	4,154	1,121,232
51	Halophilic archaeon	T01603	JGI	NC_015954	3,643,158	3,476	809,071
52	*Escherichia coli* K-12 MG1655	T00007	Wisconsin, Pasteur, Regulon DB, EcoGene, ECOCYC	NC_000913	4,639,675	4,145	1,319,938
53	*Salmonella enterica* subsp. enterica serovar Typhimurium U288	T02639	U Nottingham	NC_021151	5,017,059	4,798	1,405,169
54	*Yersinia pestis* D106004	T01832	China CDC	NC_017154	4,812,922	3,781	1,241,676
55	*Pseudomonas aeruginosa* PAO1	T00035	PathoGenesis	NC_002516	6,264,404	5,571	1,860,283
56	*Shewanella baltica* BA175	T02014	JGI	NC_017571	5,199,401	4,344	1,406,846
57	*Francisella tularensis* TIGB03	T01762	Virginia Tech	NC_016933	1,968,651	1,624	502,086
58	*Neisseria meningitidis* WUE 2594 (serogroup A)	T01931	Bielefeld U	NC_017512	2,227,255	1,941	572,960
59	*Burkholderia pseudomallei* 1026b	T02049	Washington U	NC_017831	7,231,415	6,070	1,134,954
60	*Helicobacter pylori* OK113	T02510	U Tokyo	NC_020508	161,66,17	1,520	485,985
61	*Bacillus subtilis* subsp. subtilis 6051-HGW	T02503	Goettingen	NC_020507	4,215,610	4,188	1,226,616
62	*Staphylococcus aureus* M1	T02618	Hvidovre Hospital	NC_021059	2,891,564	2,727	794,042
63	*Listeria monocytogenes* N53-1	T02513	DTU	NC_020558	2,776,847	3,150	756,718
64	*Streptococcus pneumoniae* SPN034156	T02628	Sanger	NC_021006	2,024,476	1,799	538,540
65	*Lactobacillus brevis* KB290	T02528	KAGOME	NC_020819	2,587,877	2,582	692,723
66	*Clostridium pasteurianum*	T02645	JGI	NC_021182	4,990,707	4,469	1,288,879

Table S2: Statistical items of comparative genomic analysis.

No	Full name	G+C%	Optimum NaCl for growth (%)	Proportion of acidic amino acids (%)	Number of RNAs	Gene coding density (%)	Average gene length (bp/gene)
1	*Nitrosococcus halophilus*	51.58	4	12	55	85.5	1086
2	*Halorhodospira halophila*	67.98	9	14	55	91.4	1113
3	*Halothiobacillus neapolitanus*	54.71	4	11	52	88.4	1096
4	*Chromohalobacter salexigens*	63.91	3	12	90	89.2	1121
5	*Halomonas elongata*	63.61	9	13	81	86.5	1169
6	*Desulfohalobium retbaense*	57.54	8	12	60	84.6	1152
7	*Halobacillus halophilus*	41.82	3	13	91	81.9	1011
8	*Pelagibacterium halotolerans*	61.37	3	12	53	88.8	1017
9	*Bacillus halodurans*	43.69	6	13	105	84.9	1034
10	*Tetragenococcus halophilus*	36.04	5	13	79	85.9	1003
11	*Desulfitobacterium dehalogenans*	44.97	5.8	12	99	84.1	1077
12	*Dehalobacter* sp. DCA	44.61	14	12	60	85.2	1031
13	*Dehalobacter* sp. CF	44.31	0.1	12	60	85.2	1038
14	*Halothermothrix orenii*	37.88	10	13	70	86.3	1101
15	*Acetohalobium arabaticum*	36.63	15	15	85	85.6	1082
16	*Halobacteroides halobius*	32.46	8.4	13	90	87.1	1073
17	*Corynebacterium halotolerans*	68.44	10 (KCl)	13	65	86.4	1125
18	*Halothece* sp. PCC 7418	42.92	6	12	58	84.8	1127
19	*Dehalococcoides ethenogenes*	48.85	3	12	51	89.6	930
20	*Dehalococcoides* sp. CBDB1	47.03	2	11	52	89.7	957
21	*Dehalococcoides* sp. BAV1	47.17	2	12	51	89.4	979
22	*Dehalococcoides* sp. VS	47.27	2	11	51	90.2	982
23	*Dehalococcoides mccartyi* BTF08	47.28	2.5	12	49	90.2	950
24	*Dehalococcoides mccartyi* DCMB5	47.07	2.55	12	49	89.5	969
25	*Dehalogenimonas lykanthroporepellens*	55.04	2	12	52	87.1	1017
26	*Methanohalophilus mahii*	42.62	2.3	14	64	87.9	1013
27	*Methanohalobium evestigatum*	36.63	10	14	59	83.2	1068
28	*Halobacterium* sp. NRC-1	67.91	25.2	16	52	86.0	981
29	*Halobacterium salinarum* R1	68.01	22.8	16	52	87.1	971
30	*Haloarcula marismortui*	62.36	21	17	61	84.5	1007
31	*Haloarcula hispanica*	63.69	23.4	17	60	86.0	1008
32	*Haloquadratum walsbyi* DSM 16790	47.86	18	15	52	74.3	1202
33	*Haloquadratum walsbyi* C23	47.78	18	15	56	75.8	1189
34	*Natronomonas pharaonis*	63.44	20.5	18	51	86.8	975
35	*Natronomonas moolapensis*	64.53	18	17	50	84.1	1060
36	*Halorubrum lacusprofundi*	66.72	20	17	61	83.8	1037
37	*Halorhabdus utahensis*	62.90	27	17	51	87.8	1040
38	*Halomicrobium mukohataei*	65.63	20.6	17	60	86.3	995
39	*Haloterrigena turkmenica*	65.83	20	18	64	81.3	1064
40	*Natrialba magadii*	61.42	20	18	60	83.1	1055
41	*Haloferax volcanii*	66.64	15	17	58	85.1	999
42	*Haloferax mediterranei*	61.12	20	17	64	84.7	1011
43	*Halalkalicoccus jeotgali*	64.96	20	17	52	83.8	955
44	*Halogeometricum borinquense*	61.06	20	17	57	86.0	1012
45	*Halopiger xanaduensis*	65.98	25	18	60	84.9	1032
46	*Natrinema* sp. J7-2	64.25	17.6	17	58	85.3	882
47	*Natrinema pellirubrum*	64.93	15.2	18	58	82.8	1037
48	*Natronobacterium gregoryi*	62.24	20.5	18	59	82.1	1036
49	*Halovivax ruber*	64.34	20	17	52	85.1	1040
50	*Natronococcus occultus*	64.94	23	18	68	83.2	1039
51	Halophilic archaeon	63.90	20	17	51	82.6	1048
52	*Escherichia coli* K-12 MG1655	50.79	NR	11	176	85.3	1119
53	*Salmonella enterica* subsp. enterica serovar Typhimurium U288	52.18	NR	11	107	86.7	1046
54	*Yersinia pestis* D106004	47.63	NR	11	89	80.1	1273
55	*Pseudomonas aeruginosa* PAO1	66.56	NR	11	106	89.1	1124
56	*Shewanella baltica* BA175	46.19	NR	11	141	83.2	1197
57	*Francisella tularensis* TIGB03	32.30	NR	11	49	76.5	1212
58	*Neisseria meningitidis* WUE 2594 (serogroup A)	51.84	NR	12	67	77.2	1147
59	*Burkholderia pseudomallei* 1026b	67.85	NR	11	71	82.8	1191
60	*Helicobacter pylori* OK113	38.73	NR	12	42	90.2	1064
61	*Bacillus subtilis* subsp. subtilis 6051-HGW	43.51	NR	12	123	87.3	1007
62	*Staphylococcus aureus* M1	32.42	NR	12	61	83.0	1060
63	*Listeria monocytogenes* N53-1	38.10	NR	13	85	81.8	882
64	*Streptococcus pneumoniae* SPN034156	39.85	NR	13	70	79.8	1125
65	*Lactobacillus brevis* KB290	46.12	NR	10	ND	86.0	1002
66	*Clostridium pasteurianum*	30.58	NR	12	109	77.5	1117

NR: not required; ND: no data;

Table S3: Ratio of plasmid containing strains in each groups and plasmid properties.

No	Full name	Plasmid number	Plasmid name (Accession number)	Nucleotides (bp)	Percentage of each plasmid (%)	Percentage of plasmid out of its total genetic elements (%)
1	*Nitrosococcus halophilus*	1	pNHAL01 (NC_013958)	65,833	1.59	1.59
2	*Halorhodospira halophila*	0	—	0	0	0
3	*Halothiobacillus neapolitanus*	0	—	0	0	0
4	*Chromohalobacter salexigens*	0	—	0	0	0
5	*Halomonas elongata*	0	—	0	0	0
6	*Desulfohalobium retbaense*	1	pDRET01 (NC_013224)	45,263	1.56	1.56
7	*Halobacillus halophilus*	2	PL16 (NC_017669)	16,047	0.38	0.46
			PL3 (NC_017670)	3,329	0.08	
8	*Pelagibacterium halotolerans*	1	pPHB2 (NC_016079)	4,050	0.10	0.10
9	*Bacillus halodurans*	0	—	0	0	0
10	*Tetragenococcus halophilus*	0	—	0	0	0
11	*Desulfitobacterium dehalogenans*	0	—	0	0	0
12	*Dehalobacter* sp. DCA	0	—	0	0	0
13	*Dehalobacter* sp. CF	0	—	0	0	0
14	*Halothermothrix orenii*	0	—	0	0	0
15	*Acetohalobium arabaticum*	0	—	0	0	0
16	*Halobacteroides halobius*	0	—	0	0	0
17	*Corynebacterium halotolerans*	1	pCha1 (NC_020303)	86,256	2.68	2.6
18	*Halothece* sp. PCC 7418	0	—	0	0	0
19	*Dehalococcoides ethenogenes*	0	—	0	0	0
20	*Dehalococcoides* sp. CBDB1	0	—	0	0	0
21	*Dehalococcoides* sp. BAV1	0	—	0	0	0
22	*Dehalococcoides* sp. VS	0	—	0	0	0
23	*Dehalococcoides mccartyi* BTF08	0	—	0	0	0
24	*Dehalococcoides mccartyi* DCMB5	0	—	0	0	0
25	*Dehalogenimonas lykanthroporepellens*	0	—	0	0	0
26	*Methanohalophilus mahii*	0	—	0	0	0
27	*Methanohalobium evestigatum*	1	pMETEV01 (NC_014254)	16,3915	6.81	6.81
28	*Halobacterium* sp. NRC-1	2	pNRC200 (NC_002608)	36,5425	14.21	21.65
			pNRC100 (NC_001869)	19,1346	7.44	
29	*Halobacterium salinarum* R1	4	PHS1 (NC_010366)	14,7625	5.53	25.02
			PHS2 (NC_010369)	19,4963	7.31	
			PHS3 (NC_010368)	28,4332	10.65	
			PHS4 (NC_010367)	40,894	1.53	
30	*Haloarcula marismortui*	8	Chromosome II (NC_006397)	28,8050	6.74	17.12
			pNG700 (NC_006395)	410,554	9.60	
			pNG600 (NC_006394)	155,300	3.63	
			pNG500 (NC_006393)	132,678	3.10	
			pNG400 (NC_006392)	50,060	1.17	
			pNG300 (NC_006391)	39,521	0.92	
			pNG200 (NC_006390)	33,452	0.78	
			pNG100 (NC_006389)	33,303	0.78	
31	*Haloarcula hispanica*	2	Chromosome II (NC_015943)	488,918	12.57	23
			pHH400 (NC_015944)	405,816	10.43	
32	*Haloquadratum walsbyi* DSM 16790	1	PL47 (NC_008213)	46,867	1.47	1.47
33	*Haloquadratum walsbyi* C23	3	PL100 (NC_017457)	100,258	3.07	3.45
			PL6A (NC_017460)	6,129	0.19	
			PL6B (NC_017458)	6,056	0.19	
34	*Natronomonas pharaonis*	2	PL131 (NC_007427)	130,989	4.76	5.52
			PL23 (NC_007428)	23,486	0.85	
35	*Natronomonas moolapensis*	0	—	0	0	0
36	*Halorubrum lacusprofundi*	2	Chromosome II (NC_012028)	52,5943	14.24	25.92
			pHLAC01 (NC_012030)	431,338	11.68	
37	*Halorhabdus utahensis*	0	—	0	0	0
38	*Halomicrobium mukohataei*	1	pHmuk01 (NC_013201)	221,862	6.66	6.66
39	*Haloterrigena turkmenica*	6	pHTUR01 (NC_013744)	698,495	12.84	28.52
			pHTUR02 (NC_013745)	413,648	7.60	
			pHTUR03 (NC_013746)	180,781	3.32	
			pHTUR04 (NC_013747)	171,943	3.16	
			pHTUR05 (NC_013748)	71,062	1.31	
			pHTUR06 (NC_013749)	15,815	0.29	
40	*Natrialba magadii*	3	pNMAG01 (NC_013923)	378,348	8.51	15.57
			pNMAG02 (NC_013924)	254,950	5.74	
			pNMAG03 (NC_013925)	58,487	1.32	
41	*Haloferax volcanii*	4	pHV1 (NC_013968)	85,092	2.12	29.03
			pHV2 (NC_013965)	6,359	0.16	
			pHV3 (NC_013964)	437,906	10.91	
			pHV4 (NC_013966)	635,786	15.84	
42	*Haloferax mediterranei*	3	pHM100 (NC_017942)	129,210	3.31	24.48
			pHM300 (NC_017943)	321,908	8.24	
			pHM500 (NC_017944)	504,705	12.93	

43	*Halalkalicoccus jeotgali*	6	p1 (NC_014298)	406,285	10.98	24.05
			p2 (NC_014299)	363,534	9.83	
			p3 (NC_014300)	44,576	1.21	
			p4 (NC_014301)	44,459	1.20	
			p5 (NC_014302)	23,727	0.64	
			p6 (NC_014303)	6,951	0.19	
44	*Halogeometricum borinquense*	5	pHBOR01 (NC_014735)	362,194	9.18	28.49
			pHBOR02 (NC_014731)	339,010	8.59	
			pHBOR03 (NC_014736)	210,350	5.33	
			pHBOR04 (NC_014732)	194,834	4.94	
			pHBOR05 (NC_014737)	17,535	0.44	
45	*Halopiger xanaduensis*	3	pHALXA01 (NC_015658)	436,718	10.03	15.78
			pHALXA02 (NC_015667)	181,778	4.17	
			pHALXA03 (NC_015659)	68,763	1.58	
46	*Natrinema* sp. J7-2	1	pJ7-1 (NC_018225)	95,989	2.53	2.53
47	*Natrinema pellirubrum*	2	pNATPE01 (NC_019967)	28,7800	6.61	12.94
			pNATPE02 (NC_019963)	27,5821	6.33	
48	*Natronobacterium gregoryi*	0	—	0	0	0
49	*Halovivax ruber*	0	—	0	0	0
50	*Natronococcus occultus*	2	p1 (NC_019975)	12,939	0.30	6.97
			p2 (NC_019976)	287,963	6.67	
51	Halophilic archaeon	2	phalar01 (NC_015955)	705,810	19.37	20.02
			phalar02 (NC_015959)	23,659	0.65	
52	*Escherichia coli* K-12 MG1655	0	—	0	0	0
53	*Salmonella enterica* subsp. enterica serovar Typhimurium U288	3	pSTU288-1 (NC_021155)	148,711	2.96	3.18
			pSTU288-2 (NC_021156)	11,067	0.22	
			pSTU288-3 (NC_021157)	4,675	0.09	0.09
54	*Yersinia pestis* D106004	3	pPCY1 (NC_017156)	9,611	0.20	3.58
			pCD1 (NC_017153)	68,342	1.42	
			pMT1 (NC_017155)	94,249	1.96	
55	*Pseudomonas aeruginosa* PAO1	0	—	0	0	0
56	*Shewanella baltica* BA175	2	pSBAL17501 (NC_017570)	72,392	1.39	2.56
			pSBAL17502 (NC_017572)	60,958	1.17	
57	*Francisella tularensis* TIGB03	0	—	0	0	0
58	*Neisseria meningitidis* WUE 2594 (serogroup A)	0	—	0	0	0
59	*Burkholderia pseudomallei* 1026b	1	Chromosome II (NC_017832)	3,138,747	43.40	43.40
60	*Helicobacter pylori* OK113	0	—	0	0	0
61	*Bacillus subtilis* subsp. subtilis 6051-HGW	0	—	0	0	0
62	*Staphylococcus aureus* M1	1	pSK67-M1 (NC_021060)	27,439	0.95	0.95
63	*Listeria monocytogenes* N53-1	0	—	0	0	0
64	*Streptococcus pneumoniae* SPN034156	0	—	0	0	0
65	*Lactobacillus brevis* KB290	9	pKB290-1 (NC_020820)	42,449	1.64	6.09
			pKB290-2 (NC_020821)	35,388	1.37	
			pKB290-3 (NC_020826)	35,340	1.37	
			pKB290-4 (NC_020822)	25,335	0.98	
			pKB290-5 (NC_020823)	17,882	0.69	
			pKB290-6 (NC_020827)	11,627	0.45	
			pKB290-7 (NC_020824)	10,300	0.40	
			pKB290-8 (NC_020828)	8,556	0.33	
			pKB290-9 (NC_020825)	5,866	0.23	
66	*Clostridium pasteurianum*	0	—	0	0	0

"—": do not have any plasmids.

However, no study to date has investigated the metabolic pathways of and other differences between the halophiles (haloarchaea and halophilic bacteria) and non-halophilic bacteria. In this paper, we present the results of comprehensive analysis of the genomes of these three groups of organisms, which was carried out to obtain an in-depth understanding of the genomic characteristics that allow for survival in harsh natural environments.

Methods

Group information

The complete genomes of the haloarchaea, halophilic bacteria, and non-halophilic bacteria used for statistical analysis in this study were downloaded from the public database of the Kyoto Encyclopedia of Genes and Genomes (KEGG) (www.genome.jp/kegg/) [16] and GenBank (http://www.ncbi.nlm.nih.gov/genbank/). The genomes of the sequenced halophilic bacteria were grouped as Group I, which contained 27 strains; those of the haloarchaea as Group II, comprising 24 strains; and those of the non-halophilic bacteria or normal bacteria, in which NaCl is not required for regular growth, as Group III, which included 15 strains (see Supplementary Data Table S1).

Statistical items

General genomic information, including species name, genomic accession number, T number in KEGG [16], total number of nucleotides and amino acids, original database, and number of coding genes, is provided in table S1 of the Supplementary Data. The optimum NaCl concentration for growth, G+C content (G: Guanine; C: Cytosine), acidic amino acids, tRNA, gene density, and average gene length can be found in table S2 of the Supplementary Data, and plasmid information (i.e., name, accession number, nucleotides, percentage of plasmids in total genetic elements, and megaplasmid or minichromosome) in Table S3. Finally, comparisons of the KEGG

pathways [16] of Groups I, II, and III were performed (data not shown). The pathways shared by all strains were omitted from the study.

Statistical methods

The main statistical analyses were carried out using Sigma Plot 12.2 (scatter plot or box plot) (http://www.sigmaplot.com/products/sigmaplot/sigmaplot-details.php) or Origin 7.5, whereas the proportion of acidic amino acids was calculated online (http://www.bio-soft.net/sms/index.html). The probability values (*p*-values) were obtained via a *t*-test performed using Statistical Product and Service Solutions (SPSS) statistical software.

Results

NaCl requirements

The prokaryotes, including Archaea and Bacteria, can be divided into two main groups, a NaCl-dependent group and a NaCl-independent group, based on their requirements for NaCl for growth. Most of the strains in Group I needed a relatively low NaCl concentration (~5% in W/V) for optimum growth (Figure 1), whereas the 27 strains in Group III were NaCl-independent (Supplementary Data Table S1). However, the NaCl concentration needed for the optimum growth of the Group II strains was 20%, significantly higher than that for the other two groups (Figure 1).

G+C content

It is well known that haloarchaea (the Group II strains) possess a high G+C content (> 60%), except for the strains from genus *Haloquadratum* (see Figure 2 and Supplementary Data Table S2). In this study, the G+C content of the haloarchaea was markedly higher than that of the halophilic bacteria (~45%) and non-halophilic bacteria (~45%) ($p < 0.0001$), although there was no significant difference in average G+C content between Groups I and III ($p = 0.3762$) (Figure 2). Moreover, several strains in Group I (Nos. 2, 4, 5, 8, and 17) and Group III (Nos. 55 and 59) had a similar G+C content to the Group II strains (Figure 2 and Table S2).

Proportion of acidic amino acids

The proportion of the 20 natural amino acids was analyzed (data not shown). Acidic amino acids (or negatively charged amino acids), which mainly include asparagine (Asp, D) and glutamate (Glu, E), play a crucial role in osmotic regulation, particularly in haloarchaea and halophilic bacteria. The proportion of acidic amino acids in Group II (average = 17.0%) was found to be significantly higher than that in Group I (average = 12.5%) ($p < 0.0001$) or Group III (average = 11.5%) ($p = 0.0029$) (see Table S2). The proportion of negatively charged amino acids generally rose with an increase in the

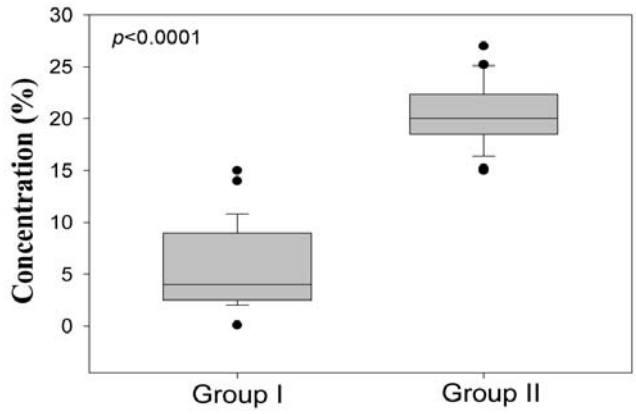

Figure 1: NaCl concentration for optimum growth. Some strains of conventional bacteria (Group III) have an NaCl tolerance, but NaCl is not needed for their optimum growth. These strains are not presented here.

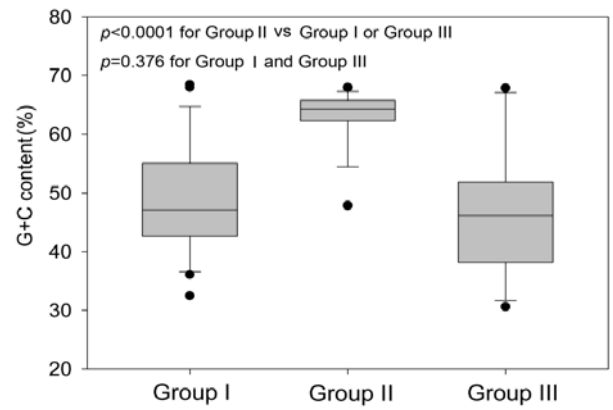

Figure 2: Comparison of G+C content. The genomic sequences were downloaded from KEGG, and calculated by DNASTAR.

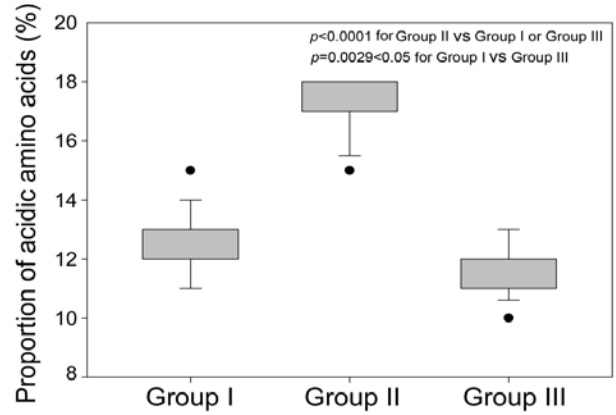

Figure 3: Proportion of acidic amino acids. The gene products were downloaded from KEGG, and the proportion of amino acids was analyzed online (http://www.bio-soft.net/sms/index.html).

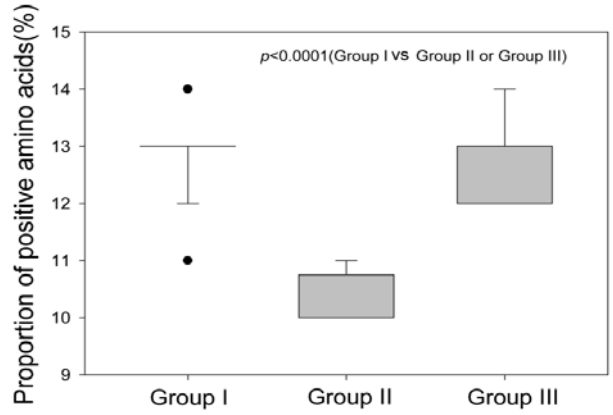

Figure S1: Proportion of positive amino acids.

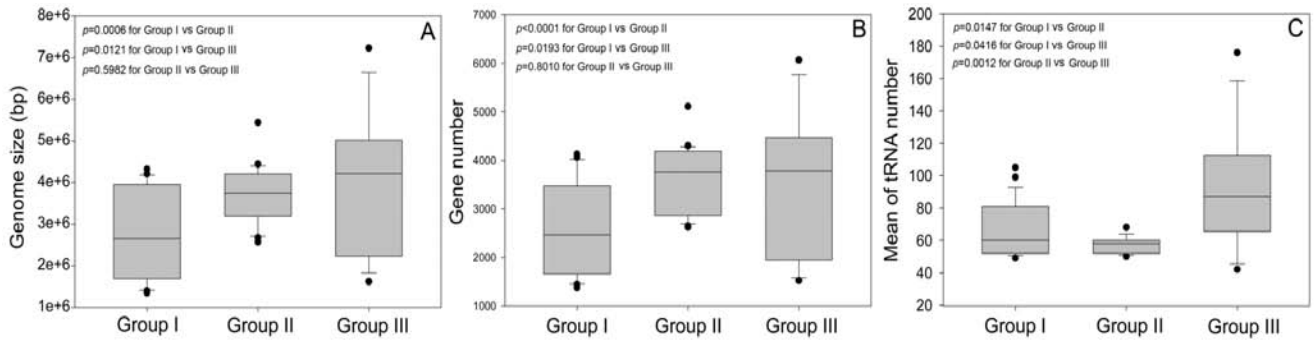

Figure 4: Genome size, gene number, and tRNA number. A. Genome size; B. Gene number; C. tRNA number.

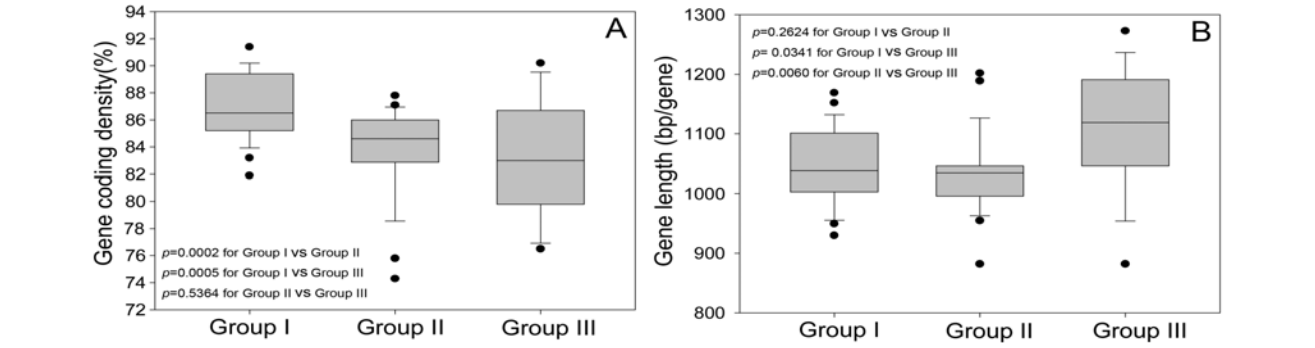

Figure 5: Gene length and density. Genes from the extra-chromosome and the main genome were combined for gene length and density analysis. A. Gene coding density; B. Average gene length.

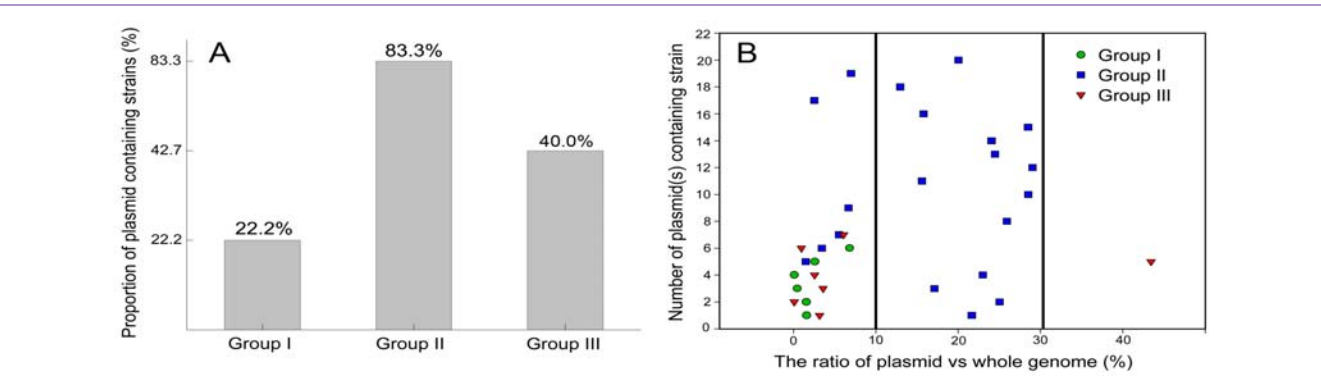

Figure 6: Proportion of plasmid-containing strains and general properties of plasmids. A. Ratio of plasmid-containing strains; B. Number of plasmids in plasmid-containing strains.

NaCl concentration for growth (Figure 3), whereas the proportion of positive amino acids decreased with such an increase (Supplementary Data Figure S1).

Genome size, gene number, and tRNA number

The average genome size in Group II was larger than that in Group I (Figure 4A), as was the number of genes (Figure 4B). See Table S1 in addition to the two figures. The average genome size and number of genes in Group III varied more widely (Figure 4), which rendered these strains unsuitable for direct comparison. Interestingly, the range of variation in tRNA number in the Group II was much narrower than that in either Group I or Group III (Figure 4C and Table S2). There are generally 61 tRNAs in all microorganisms (the exception being the several tRNAs in rare codons), but the tRNA numbers presented here are those given by KEGG (www.genome.jp/kegg/), which are predicted by the software based on analysis of the genomic sequence. Hence, the tRNA number does not depend on the tRNA type.

Gene length and density

The gene coding density in Group II was very similar to that in Group III ($p > 0.5$), and much lower than that in Group I ($p < 0.05$) (Figure 5A). The halophilic bacteria (Group I) featured the highest gene coding density of the three groups. However, the average gene length in Group I was quite similar to that in Group II ($p = 0.2624 > 0.05$), albeit much lower than that in Group III ($p < 0.05$) (Figure 5B). Hence, we concluded haloarchaea and halophilic bacteria share a similar average gene length, which differs considerably from that of

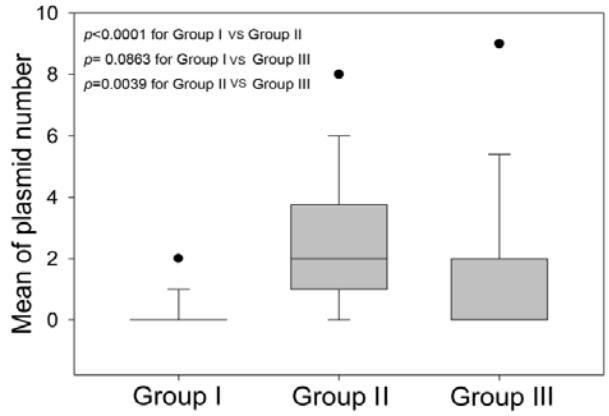

Figure S2: The average number of plasmid.

non-halophilic bacteria.

Proportion of plasmid-containing strains and plasmid properties

The plasmids used for statistical analysis in this study included conventional plasmids, megaplasmids, and minichromosomes. The percentage (ratio) of plasmid-presenting strains among the haloarchaea, halophilic bacteria, and non-halophilic bacteria was 22.2% (6/27) for Group I, 40% (6/15) for Group III, and 83.3% (20/24) for Group II (Supplementary Data Table S3). Hence, the proportion of plasmid-containing strains in Group II was markedly higher than that in the other two groups. It seems that plasmids are widely distributed in haloarchaea, but limited in both halophilic and non-halophilic bacteria.

In addition, the average number of plasmids (ratio of plasmid-containing strains) in Groups I, II, and III were 0.26 (7/27), 2.54 (61/24), and 1.27 (19/15), respectively (Supplementary Data Figure S2). If the *Lactobacillus brevis* KB290 strain in Group III, which harbored the richest plasmids (9/1) of the 66 genome-sequenced strains, was excluded from plasmid calculation, then the average number (ratio) of plasmids in Group III would immediately decrease to 0.71 (10/14). The proportion of plasmids in the total genetic elements in the three groups differed significantly. Those in each of the Group I and III strains accounted for less than 10% of the total number of genetic elements, with the exception of the *Burkholderia pseudomallei* 1026b strain (Figure 6B and Supplementary Data Table S3). In contrast, most of the plasmids in Group II accounted for more than 20% of the total number of such elements (Figure 6).

Comparative analysis of metabolic pathways

The metabolic pathways that were widely distributed among the Group I, II, and III strains are not presented in this paper, leaving 61 distinct pathways for comparison (Figure 7). The pathways of fluorobenzoate degradation (row 12), D-glutamine and D-glutamate metabolism (row 20), D-arginine and D-ornithine metabolism (row 21), D-alanine metabolism (row 22), lipopolysaccharide biosynthesis (row 26), peptidoglycan biosynthesis (row 27), limonene and pinene degradation (row 46), flagellar assembly (row 55), homologous recombination (row 60), and non-homologous end-joining (row 61) were shared by the Group I and III strains, whereas that of polycyclic aromatic hydrocarbon degradation (row 35) was shared primarily by those in Groups I and II. The pathway of carbon fixation in photosynthetic organisms (row 42) was found in Group I alone, whereas those of carotenoid biosynthesis (row 47), sesquiterpenoid and triterpenoid biosynthesis (row 48), mRNA surveillance (row 57), basal transcription factors (row 58), and proteasome (row 59) were found primarily in Group II (Figure 7).

Discussion

Microorganisms that inhabit hypersaline environments are designated halophiles. Depending upon the salt concentration they require for optimum growth, they are classified as haloarchaea (Group II), which grow optimally in media with 15%-30% (2.5 M--5.2 M) NaCl, or halophilic bacteria (Group I), which grow optimally in media with 3%-15% (0.5 M-2.5 M) NaCl (Figure 1). Non-halophilic bacteria, in contrast, are microorganisms that achieve optimal growth in media with less than 1% (0.2 M) NaCl.

Halophiles have evolved two major strategies to cope with the high osmotic pressure in their hypersaline environments. Most aerobic halophilic bacteria produce compatible solutes, such as betaine, ectoine, β-carotene, or trehalose, whereas haloarchaea take advantage of the accumulation of intracellular potassium to balance that pressure. However, several species of halophilic bacteria, namely, *Salinibacter* and *Salisaeta*, have a similar mechanism to haloarchaea, coping with their hypersaline environments via a low degree of water activity [17,18]. A number of the current study's findings are worthy of particular note.

First, the G+C content of the haloarchaea far outweighed that of either the halophilic or non-halophilic bacteria, although a few exceptions were found in all three groups of microorganisms (Figure 2). Examination of the protein coding genes showed that the use of codons with G or C in the third codon position in the haloarchaea was over 90%, which may be attributable to the high G+C content of haloarchaea [19]. That content may contribute to the stability of genetic materials through replication, transcription, and gene expression. However, the reason that G or C always appears in the third codon position in haloarchaea requires elucidation.

Second, as is well known, high concentrations of salt lead to protein aggregation. Cations can capture the combined H_2O from the protein molecule, and denature it. However, the proteins in haloarchaea are unlikely to be denatured by the universal denaturing NaCl concentration, as on the contrary, they need a high NaCl concentration to perform their biological activity. The proportion of acidic amino acids in the haloarchaeal proteins in this study reached 17% or even higher, which was markedly higher than those in the proteins from the halophilic or non-halophilic bacteria (Figure 3). Acidic amino acids (glutamate or aspartate), combined with the intracellular or extracellular cations needed to avoid a configuration change in proteins, may play a critical role in the mechanism of adaptive evolution.

Third, the distribution range of tRNA number (Figure 4A), genome size (Figure 4B), gene number (Figure 4C), gene coding density (Figure 5A), and gene length (Figure 5B) among the various strains of haloarchaea was much narrower than that among the halophilic and non-halophilic bacteria strains. Unlike haloarchaea, halophilic and non-halophilic bacteria are represented by a wide

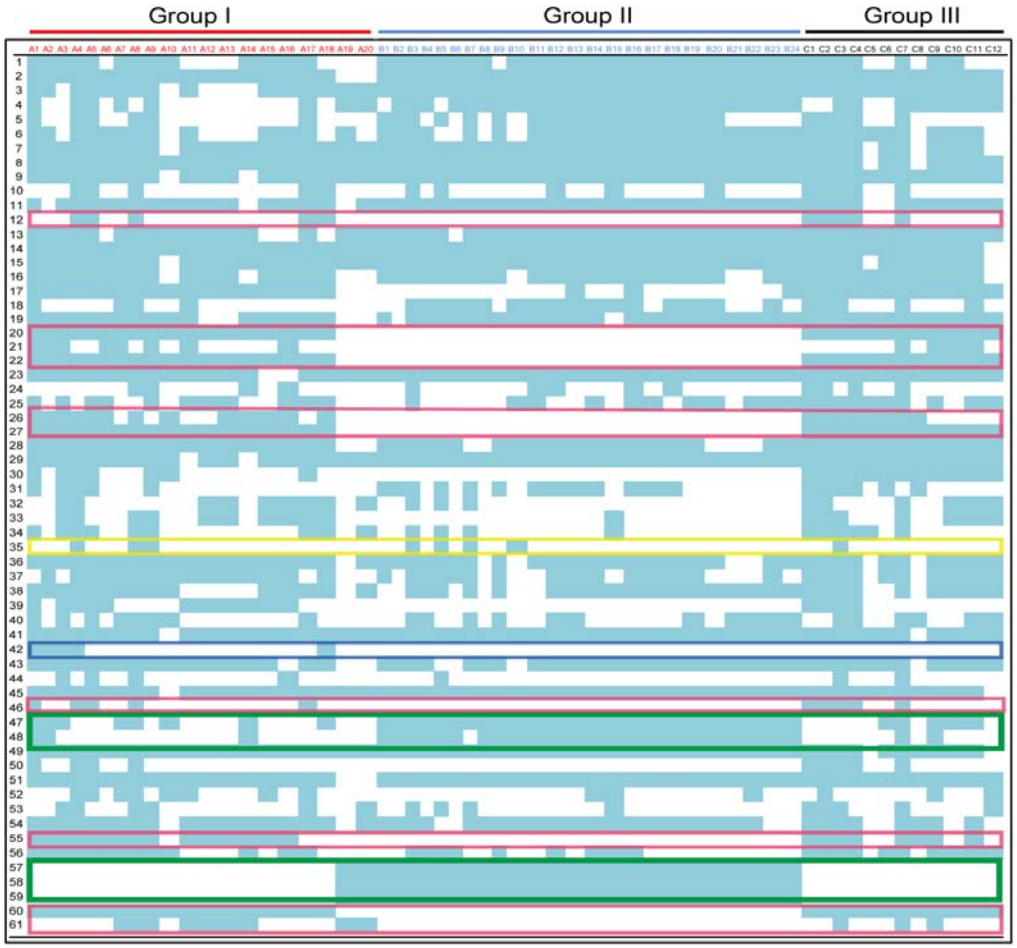

Figure 7: Comparative analysis of metabolic pathways. The pathways shared by all halophilic bacteria (Group I), haloarchaea (Group II), and non-halophilic bacteria (Group III) were omitted from the study. The pathways boxed in red (the pathways of 12, 20-22, 26-27, 46, 55, and 60-61) were common to both the halophilic (Group I) and non-halophilic bacteria (Group III). That boxed in yellow (the pathway of 35) was mainly shared by halophilic bacteria (Group I) and haloarchaea (Group II). The pathway boxed in blue (the pathway of 42) was found only in the halophilic bacteria (Group I), and those boxed in green (the pathways of 47-48 and 57-59) were found primarily in the haloarchaea (Group II).

A1. *Nitrosococcus halophilus*; A2. *Halorhodospira halophila*; A3. *Halothiobacillus neapolitanus*; A4. *Chromohalobacter salexigens*; A5. *Halomonas elongata*; A6. *Desulfohalobium retbaense*; A7. *Halobacillus halophilus*; A8. *Pelagibacterium halotolerans*; A9. *Bacillus halodurans*; A10. *Tetragenococcus halophilus*; A11. *Desulfitobacterium dehalogenans*; A12. *Dehalobacter* sp. DCA; A13. *Dehalobacter* sp. CF; A14. *Halothermothrix orenii*; A15. *Acetohalobium arabaticum*; A16. *Halobacteroides halobius*; A17. *Corynebacterium halotolerans*; A18. *Halothece* sp. PCC 7418; A19. *Methanohalophilus mahii*; A20. *Methanohalobium evestigatum*; B1. *Halobacterium* sp. NRC-1; B2. *Halobacterium salinarum* R1; B3. *Haloarcula marismortui*; B4. *Haloarcula hispanica*; B5. *Haloquadratum walsbyi* DSM 16790; B6. *Haloquadratum walsbyi* C23; B7. *Natronomonas pharaonis*; B8. *Natronomonas moolapensis*; B9. *Halorubrum lacusprofundi*; B10. *Halorhabdus utahensis*; B11. *Halomicrobium mukohataei*; B12. *Haloterrigena turkmenica*; B13. *Natrialba magadii*; B14. *Haloferax volcanii*; B15. *Haloferax mediterranei*; B16. *Halalkalicoccus jeotgali*; B17. *Halogeometricum borinquense*; B18. *Halopiger xanaduensis*; B19. *Natrinema* sp. J7-2; B20. *Natrinema pellirubrum*; B21. *Natronobacterium gregoryi*; B22. *Halovivax ruber*; B23. *Natronococcus occultus*; B24. *Halophilic archaeon*; C1. *Escherichia coli* K-12 MG1655; C2. *Yersinia pestis* D106004; C3. *Pseudomonas aeruginosa* PAO1; C4. *Shewanella baltica* BA175; C5. *Francisella tularensis* TIGB03; C6. *Neisseria meningitidis* WUE 2594 (serogroup A); C7. *Burkholderia pseudomallei* 1026b; C8. *Helicobacter pylori* OK113; C9. *Bacillus subtilis* subsp. subtilis 6051-HGW; C10. *Staphylococcus aureus* M1; C11. *Listeria monocytogenes* N53-1; C12. *Lactobacillus brevis* KB290.

1. Ascorbate and aldarate metabolism; 2. Fatty acid biosynthesis; 3. Fatty acid metabolism; 4. Synthesis and degradation of ketone bodies; 5. Geraniol degradation; 6. Lysine degradation; 7. Histidine metabolism; 8. Tyrosine metabolism; 9. Phenylalanine metabolism; 10. Chlorocyclohexane and chlorobenzene degradation; 11. Benzoate degradation; 12. Fluorobenzoate degradation; 13. Tryptophan metabolism; 14. Phenylalanine, tyrosine and tryptophan biosynthesis; 15. Novobiocin biosynthesis; 16. beta-Alanine metabolism; 17. Taurine and hypotaurine metabolism; 18. Phosphonate and phosphinate metabolism; 19. Cyanoamino acid metabolism; 20. D-Glutamine and D-glutamate metabolism; 21. D-Arginine and D-ornithine metabolism; 22. D-Alanine metabolism; 23. Glutathione metabolism; 24. Other glycan degradation; 25. Polyketide sugar unit biosynthesis; 26. Lipopolysaccharide biosynthesis; 27. Peptidoglycan biosynthesis; 28. Glycerolipid metabolism; 29. Inositol phosphate metabolism; 30. Arachidonic acid metabolism; 31. alpha-Linolenic acid metabolism; 32. Dioxin degradation; 33. Xylene degradation; 34. Toluene degradation; 35. Polycyclic aromatic hydrocarbon degradation; 36. Chloroalkane and chloroalkene degradation; 37. Naphthalene degradation; 38. Aminobenzoate degradation; 39. Nitrotoluene degradation; 40. Styrene degradation; 41. C5-Branched dibasic acid metabolism; 42. Carbon fixation in photosynthetic organisms; 43. Lipoic acid metabolism; 44. Atrazine degradation; 45. Porphyrin and chlorophyll metabolism; 46. Limonene and pinene degradation; 47. Carotenoid biosynthesis; 48. Sesquiterpenoid and triterpenoid biosynthesis; 49. Sulfur metabolism; 50. Caprolactam degradation; 51. Biosynthesis of unsaturated fatty acids; 52. Nonribosomal peptide structures; 53. Degradation of aromatic compounds; 54. Bacterial chemotaxis; 55. Flagellar assembly; 56. Phosphotransferase system (PTS); 57. mRNA surveillance pathway; 58. Basal transcription factors; 59. Proteasome; 60. Homologous recombination; 61. Non-homologous end-joining.

variety of species in different phylogenetic lineages (orders), thus reflecting a broad range of major genetic information [20].

Fourth, like most bacterial genomes, the haloarchaeal genomes in this study ranged from 2.5-5.4 Mbp (Supplementary Data Table S1), with a single main circular chromosome and, on occasion, accessory plasmids or extra chromosomal elements (Supplementary Data Table S3). Some large plasmids or megaplasmids containing several important or essential genes are classified as minichromosome or chromosome II (Supplementary Data Table S3; see also [7]). Our results show that the probability of extra chromosomal elements occurring in haloarchaea is greater than 83% and that the percentage of extra chromosomal elements in their total genetic elements is between 12% and 30%. Both figures far outweigh those for halophilic or non-halophilic bacteria (Figure 6). Studies of the megaplasmids in *Halobacterium* sp. NRC-1 and similar plasmids in other *Halobacterium* strains suggest that they are highly dynamic and rapidly evolving [21]. The widely distributed and highly dynamic extra chromosomal elements in haloarchaea can be attributed to the high frequency of homologous recombination [22] and imprecise excision.

Fifth, the plasma membranes of Archaea differ from those of Bacteria, and exhibit remarkable structural and chemical diversity. The realization in the early 1970s that these cell walls do not contain peptidoglycan, a major component of bacterial cell walls, was an initial pillar on which the establishment of Archaea as a distinct phylogenetic kingdom rested [23]. The lipopolysaccharide-containing outer membranes that are a typical characteristic of gram-negative bacteria are absent in Archaea [24]. In Archaea, including haloarchaea, flagellar biosynthesis is reminiscent of bacterial type IV pilus biosynthesis [25]. Apart from the flagella, the functional roles played by the putative archaeal pili and pilus-like structures are unknown. However, in the current study, few flagellins, which are related to the flagellar assembly, were found in the haloarchaea in comparative analysis of the metabolic pathways (Figure 7), suggesting that the flagellins of haloarchaea are quite different from those of halophilic or non-halophilic bacteria.

Sixth, Soloshonok and Izawa [26] indicated that many D-amino acids or secondary metabolites can be found in the cell walls of microorganisms. In this study, we did not find the metabolic pathways of D-glutamine, D-glutamate, D-arginine, D-ornithine, or D-alanine (Figure 7). Hence, we hypothesize that these D-amino acid pathways are most likely absent in haloarchaea. In addition to D-amino acids, we found fluorobenzoate, limonene, and pinene to be widely distributed in the two types of bacteria (Figure 7), yet absent in the family Halobacteriaceae. As there are few reports on other groups of Archaea, this absence may constitute a selective hallmark of the distinction between Archaea and Bacteria.

Seventh, homologous recombination or non-homologous end-joining frequently occurs in Archaea, including haloarchaea [27]. However, our survey of the comparative metabolic pathways of haloarchaea, halophilic bacteria, and non-halophilic bacteria found no such occurrence in haloarchaea. This finding suggests that the DNA or protein sequences of the homologous recombination- or non-homologous end-joining-related enzymes or proteins of haloarchaea are distinct from those of others.

Eighth, the proteasome that occurred in the haloarchaea was absent in the halophilic and non-halophilic bacteria (Figure 7). An ATP-dependent protease in Bacteria appears to be homologous to Archaea and the eukaryotic proteasome, and that in Archaea shares the simple architecture of bacterial proteases. However, the subunits are homologous to the eukaryotic proteasome, thus suggesting the existence of a bridge between bacteria and eukaryotic organisms [28]. Moreover, the mRNA surveillance pathway and basal transcription factors in haloarchaea are similar to proteasome, but neither was found in the halophilic or non-halophilic bacteria in this study.

Finally, carotenoid is a major component of halorhodopsin (light-driven chloride pump) [29] and bacteriorhodopsin (light-driven proton pump) [30]. In addition to carotenoid, sesquiterpenoid and triterpenoid were widely distributed in the haloarchaea, but barely present in either the halophilic or non-halophilic bacteria (Figure 7). Because haloarchaea can utilize sunlight to perform photosynthesis, whereas halophilic bacteria cannot, these bacteria need to take advantage of the synthesis of a compatible solute to cope with the hypersaline environment and achieve growth. From this perspective, it is obvious that the strategy evolved by haloarchaea is considerably more beneficial than that evolved by halophilic bacteria. In the condition of ever-present solar irradiance, the haloarchaea living in hypersaline environments thrive equally well to non-halophilic bacteria living in less hostile environments.

In sum, the genome composition of haloarchaea, including their high G+C and acidic amino acid content, reveal apparent traits of adaptive evolution when these species live in a hypersaline environment for long periods. The higher G+C content in haloarchaea leads to greater sequence similarity, which means that haloarchaea have a higher probability of homologous recombination than do halophilic or non-halophilic bacteria. As a result, the plasmid-containing ratio and plasmid proportion in haloarchaea are higher than those in the other two groups. In a harsh environment lacking in nutrition and full of salt, haloarchaea use their purple membrane structure to cope. Haloarchaea are one of the very few microorganisms that lack chloroplast but are able to draw upon the sun for energy synthesis.

Acknowledgements

We would like to thank Dr. Yan-Jia Hao from the Kunming Institute of Botany at the Chinese Academy of Sciences for assistance with data analysis. We are also grateful for the support we received for the study from the CAS/SAFEA International Partnership Program for Creative Research Teams and the Hundred Talents Program of the Chinese Academy of Sciences, and the National Natural Science Foundation of China (31460003).

References

1. Grant WD, Kamekura M, McGenity TJ, Ventosa A. Class III. Halobacteria class. nov. In: Boone D, Castenholz R (eds), Bergey's manual of systematic bacteriology. New York: Springer. 2001; 2: 294-334.

2. DasSarma S, DasSarma P. Halophiles. In: Encyclopedia of Life Sciences. New York: John Wiley & Sons. 2012; 175-186.

3. DasSarma P, Coker J, Huse V, DasSarma S. Halophiles, biotechnology. In: Flickinger M (ed), Encyclopedia of industrial biotechnology, bioprocess, bioseparation, and cell technology. New York: John Wiley & Sons. 2010; 2769-2777.

4. Samuel BS, Gordon JI. A humanized gnotobiotic mouse model of host Archaeal-Bacterial mutualism. Proc Natl Acad Sci. 2006; 103: 10011-10016.

5. Oren A. *Salinibacter*: An extremely halophilic bacterium with Archaeal properties. FEMS Microbiol Lett. 2013; 342: 1-9.

6. Ng WV, Kennedy SP, Mahairas GG, Berquist B, Pan M, et al. Genome sequence of *Halobacterium* species NRC-1. Proc Natl Acad Sci USA. 2000; 97: 12176-12181.

7. DasSarma S. Genome sequence of an extremely halophilic archaeon. In: Fraser CM, Read T, Nelson KE (eds), Microbial genomes. Totowa, NJ: Humana. 2004; pp. 383-399.

8. Kennedy SP, Ng WV, Salzberg SL, Hood L, DasSarma S. Understanding the adaptation of *Halobacterium* species NRC-1 to its extreme environment through computational analysis of its genome sequence. Genome Res. 2001; 11: 1641-1650.

9. Paul S, Bag SK, Das S, Harvill ET, Dutta C. Molecular signature of hypersaline adaptation: Insights from genome and proteome composition of halophilic prokaryotes. Genome Biol. 2008; 9: R70.

10. Capes MD, DasSarma P, DasSarma S. The core and unique proteins of haloarchaea. BMC Genomics. 2012; 13: 39.

11. Arakawa T, Tokunaga M. Electrostatic and hydrophobic interactions play a major role in the stability and refolding of halophilic proteins. Protein Pept Lett. 2004; 11: 125-132.

12. Britton KL, Baker PJ, Fisher M, Ruzheinikov S, Gilmour DJ, et al. Analysis of protein solvent interactions in glucose dehydrogenase from the extreme halophile *Haloferax mediterranei*. Proc Natl Acad Sci USA. 2006; 103: 4846-4851.

13. DasSarma S, Capes M, DasSarma P. Haloarchaeal megaplasmids. In: Schwartz E (ed), Megaplasmids. Berlin: Springer-Verlag Berlin and Heidelberg GmbH. 2008; 3-30.

14. Capes MD, Coker JA, Gessler R, Grinblat-Huse V, DasSarma SL, et al. The information transfer system of halophilic Archaea. Plasmid. 2011; 65: 77-101.

15. Wang G, Kennedy SP, Fasiludeen S, Rensing C, DasSarma S. Arsenic resistance in *Halobacterium* sp. strain NRC-1 examined by using an improved gene knockout system. J Bacteriol. 2004; 186: 3187-3194.

16. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res. 2010; 38: D355-D360.

17. Vaisman N, Oren A. *Salisaeta longa* gen. nov., sp. nov., a red, halophilic member of the Bacteroidetes. Int J Syst Evol Microbiol. 2009; 59: 2571-2574.

18. Oren A, Elevi Bardavid R, Kandel N, Aizenshtat Z, Jehlička J. Glycine betaine is the main organic osmotic solute in a stratified microbial community in a hypersaline evaporitic gypsum crust. Extremophiles. 2013; 17: 445-451.

19. Oren A. Halophilic microorganisms and their environments. In: Joseph S (ed), Cellular origin, life in extreme habitats and astrobiology. Dordrecht: Kluwer Academic Publishers. 2002; 323-355.

20. Vreeland RH. Advances in understanding the biology of halophilic microorganisms. In: Ventosa A, Márquez MC, Sánchez-Porro C, de la Haba RR (eds), Taxonomy of halophilic Archaea and Bacteria. Dordrecht: Springer Science+Business Media. 2012; 59-80.

21. Pfeifer F, Ghahraman P. Plasmid pHH1 of *Halobacterium salinarium*: Characterization of the replicon region, the gas vesicle gene cluster and insertion elements. Mol Gen Genet. 1993; 238: 193-200.

22. Williams D, Gogarten JP, Papke RT. Quantifying homologous replacement of loci between haloarchaeal species. Genome Biol E. 2012; 4: 1223-1244.

23. Wolfe RS. The Archaea: A personal overview of the formative years. In: Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E (eds), The prokaryotes, Vol. 3: Archaea. Bacteria: Firmicutes, actinomycetes. Berlin: Springer. 2006; 3-9.

24. Kandler and König. Cell envelopes of Archaea: Structure and chemistry. In: Kates M, Kushner DJ, Matheson AT (eds), The biochemistry of Archaea (Archaebacteria). Amsterdam: Elsevier Science Publishers B.V. 1993; 223-259.

25. Jarrell KF, McBride MJ. The surprisingly diverse ways that prokaryotes move. Nat Rev Microbiol. 2008; 6: 466-476.

26. Soloshonok V, Izawa K. Asymmetric synthesis and application of alpha-amino acids. Washington, DC: American Chemical Society. 2009.

27. White MF. Homologous recombination in the Archaea: The means justify the ends. Biochem Soc Trans. 2011; 39: 15-19.

28. Zwickl P, Ng D, Woo KM, Klenk HP, Goldberg AL. An archaebacterial ATPase, homologous to ATPases in the eukaryotic 26S proteasome, activates protein breakdown by 20S proteasomes. J Biol Chem. 1999; 274: 26008-26014.

29. Schobert B, Lanyi JK. Halorhodopsin is a light-driven chloride pump. J Biol Chem. 1982; 257: 10306-10313.

30. Oesterhelt D, StoeckeniusW. Functions of a new photoreceptor membrane. Proc Natl Acad Sci USA. 1973; 70 : 2853-2857.

Austin J Proteomics Bioinform & Genomics - Volume 1 Issue 1 - 2014
ISSN : 2471-0423 | www.austinpublishinggroup.com

Citation: Chen S, Yang J, Liu Y, Wang C and Yang ZL. Comparative Genomic Analysis of Halophiles Reveals New Clues to Their Adaptation Strategies in Hypersaline Environments. Austin J Proteomics Bioinform & Genomics. 2014;1(1): 11.