MICROBIAL DIVERSITY OF CHINESE LAKES

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

By

Eulyn Pagaling BSc (Leicester)

Department of Infection, Immunity and Inflammation

University of Leicester

June 2007

THESIS CONTAINS CD/DVD

ABSTRACT

The microbial ecology of hot springs in Rehai geothermal field in Yunnan Province, China and salt lakes in Inner Mongolia, China were investigated. 16S ribosomal RNA (rRNA) genes were amplified from these environments and were cloned and sequenced. This molecular analysis showed that the 72.1°C hot spring contained *Aquificae*-dominated white streamers, while the hot water discharge (60 – 65°C) contained a laminated mat. This analysis also showed that the salt lakes were dominated by *Halorubrum* (34% of clone sequences). The salt lakes' bacterial community consisted of *Gammaproteobacteria* (20%), *Firmicutes* (20%) and *Bacteroidetes* (20%), though the bacterium *Salinibacter ruber* appeared to be the predominant bacterium in the saltern at Lake Ejinnor. Moreover, several *Gammaproteobacteria* (67% of bacterial isolates) and *Firmicutes* (23%) and a few *Archaea* (*Halorubrum* and *Haloarcula* sp.) were cultivated from the salt lakes; the majority of isolates retrieved were not detected by molecular analysis. Statistical analysis showed that temperature, Na and Mg ions, and pH were the factors that drive microbial community composition in the salt lakes, and that geographic distance is not a factor, hence the hypothesis that 'the milieu selects' can be applicable. Furthermore, the archaeal community consists of closely related lineages, whereas the bacterial community is highly diverse.

Two novel archaeal viruses that infect a haloarchaeal strain that is 98% identical to *Halorubrum saccharovorum* 16S rRNA sequence were isolated from salt lake water of Lake Bagaejinnor. Both are lytic head/tailed viruses; virus BJ1 is assigned to the *Siphoviridae*, while virus BJ2 is assigned to the *Myoviridae*, both of the order *Caudovirales*. Virus BJ1 has a 42271 base pair (bp) double stranded (ds) DNA genome, with a G+C content of 64.8 mol% that is terminally redundant and may be circularly permuted. Almost all Open Reading Frames (ORFs) are in the forward strand in the same direction consistent with a rolling circle mechanism of DNA replication. A fraction of the virus population contained extra DNA that is probably host derived, and may suggest viral integration. Virus BJ2 genome is linear dsDNA. It is incomplete, but 97602 bp (approximately 90% of the genome) has been sequenced, which has a G+C content of 51 mol% in 42 contigs. No genomic identity between the viruses has been observed. Virus BJ2 has a narrow host range; it is unable to infect the type strain of *Halorubrum saccharovorum* or closely related field isolates. Virus BJ2 is stable from 4°C to 50°C, but titres drop ten fold at 60°C. Maximal titres were observed at 2 to 3 M NaCl and titres were stable at pH 6 – 10, but reduced 10000 fold at pH 4.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisors, Prof Bill Grant and Dr Shaun Heaphy for all their help, advice and encouragement during my PhD. I'd like to thank them for the opportunities that they have given me, not just academically, but for allowing me to go on the expeditions; they were truly excellent experiences. I would like to thank Dr Richard Haigh for all his help with sequencing the haloviruses. Thanks also go to Stephan Hyman for his time in showing me the processes in TEM and for producing the images of the haloviruses. I am grateful to the members of the EU-China Consortium for feedback on my work, especially Prof Antonio Ventosa for providing archaeal strains for the virus host experiments and for characterising one of my archaeal strains.

My special thanks go to my family for their love and encouragement, for supporting me all the way, especially during the more difficult periods of my PhD. I cannot begin to list all the things I am grateful to them for, so thanks for...well, everything!

I'd like to thank everyone in the lab, past and present, for their role in making my PhD experience. Thanks to Andrew Wallace for all his help and guidance in the lab, and thanks again to him, Pam and Ahmed for making coffee breaks more interesting. I am also grateful to other people in the Department, past and present: thanks to Sarah G (for being a great friend, housemate, fellow cocktail and vodka consumer and fellow *Lost* fan, and for letting me steal - I mean borrow - chemicals from the lab), to Sarah F and Victoria B (both for being great company while you were in Leicester), to Rob F (thanks for your help with the cloning work and constantly fixing our gel scanner), Claire, Wayne, James and Rana, Adam W and Mafalda. Thanks also to Simon (for also being a great friend and housemate - we miss you loads; we won't forget you. 'Did you know Billy Kennedy is in *House*?'), Yasmin (thanks for the long phone conversations, pick-me-up shopping sessions, dinners at Stones and trips to the Hagen Daas cafe – mmmm, icecream!), Rob H and Yannis. Thanks for all the parties (especially the fancy dress ones) and the numerous trips to the pub - you have all been brilliant and have made my time in Leicester much more enjoyable. Thanks to my current housemates, Vic and Jim for sharing PhD blues. Last, but by no means least, thanks to Joe for all his help and support, especially during my writing up period. Thanks for putting up with my moods and thanks for all the hugs when I needed them.

This work was funded by European Commission research programme 'Quality of life and management of living resources', project Multigenome Access Technology for Industrial Catalysts. The TEM was done at the Advanced Microscopy Centre and the chemical analysis was done at the Department of Geology, both at the University of Leicester.

TABLE OF CONTENTS

Contents

Contents

Contents

A data DVD ROM has been enclosed with this thesis containing three folders of sequencing data and footage of the expeditions to Yunnan Province, China in March 2003, and to Inner Mongolia, China in September 2003. Details are found in Appendix 6, p 363.

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

aa	amino acid
ABV	*Acidianus* bottle-shaped virus
ALW	artificial lake water
ATP	adenosine triphosphate
ATV	*Acidianus* two tailed virus
BChl	bacteriochlorophyll
BJ	Lake Bagaejinnor
BLAST	Basic local alignment search tool
bp	base pair
CF	clone frequency
CFB	*Cytophage-Flavobacterium-Bacteroidetes* phylum
CFU	colony forming unit
CG	Lake Chagannor
CHM	classic halophile medium
D	direction of translation
DGGE	denaturing gradient gel electrophoresis
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DOC	dissolved organic carbon
dH_2O	distilled water
dsDNA	double stranded DNA
DTT	dithiothreitol
E	evenness
EBI	European bioinformatics institute
E. coli	*Escherichia coli*
EJ	Lake Ejinnor
EMBL	European molecular biology laboratories
EN	Lake Erliannor
FAPs	filamentous anoxygenic phototrophs
FASTA	Homology search program
FISH	Fluorescence *in situ* hybridisation
Fst	coefficient of differentiation
HS	hot spring
ICP-ES	Inductively coupled plasma optical emission spectrometry
ID	sequence identity
IPTG	isopropyl β-D thiogalactopyranoside
LB	Luria Bertani
LP	Lang Pu
LUCA	last common ancestor
mCHM	modified classic halophile medium
MEGA	Molecular evolutionary genetics analysis
Mr	molecular mass
Myr	Million year
NCBI	National centre for biotechnology information
nH_2O	nanopure water

OD	optical density
ORF	open reading frame
OTU	operational taxonomic unit
p	significance value
PBS	phosphate buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PFGE	pulse field gel electrophoresis
PFU	plaque forming units
PHB	poly-β-hydroxybutyrate
pI	calculated isoelectric point
ppm	parts per million
P test	parsimony test
r	correlation value
RBS	ribosomal binding site
RDP	ribosomal database project
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNase	Ribonuclease
rRNA	ribosomal RNA
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SH	Lake Shangmatala
SHOW	square haloarchaea of Walsby
SL	salt lake
SNDE	standard normal deviate equivalents
SOB	sulphur oxidising bacteria
SRB	sulphate reducing bacteria
SNDV	*Sulfolobus* neozealandicus droplet-shaped virus
TAE	tris-acetate EDTA
TBE	tris-borate EDTA
TC	Tengchong
TE	Tris-EDTA
TEM	transmission electron microscopy
TEMED	N,N,N',N'-tetramethylethylenediamine
T_m	melting temperature
T_{max}	maximum temperature for growth
T_{min}	minimum temperature for growth
T_{opt}	optimum temperature for growth
tRNA	transfer RNA
VFAs	volatile fatty acids
X-Gal	5-bromo-4-chloro-3-indolyl-β-d glactopyronoside
XH	unnamed salt lake

INTRODUCTION

The Three Domains of Life: Discovery of the Domain *Archaea*

For years, biologists grouped living organisms into five kingdoms, based on structural similarities. It was believed that organisms were split into the eukaryote/prokaryote dichotomy until work pioneered by Carl Woese in the 1970s revealed that there were three different classes of ribosomal RNAs (rRNA) and proposed that organisms were split into three domains: the *Eukarya*, *Bacteria* and *Archaea* (Woese and Fox 1977; Woese *et al.* 1990). These three domains are believed to have evolved from the Last Universal Common Ancestor (LUCA) (Penny and Poole 1999), which is the root of the Universal Tree of Life (Figure 1). Besides discovering that not all prokaryotes were phylogenetically related, it was shown that species of the *Archaea* were more closely related to *Eukarya* than the *Bacteria*. Thus evolution from LUCA went in two directions; the *Bacteria* line and the other line, with the latter branching to the *Archaea* and *Eukarya*. Table 1 shows some morphological and genetic characteristics shared by the three domains.

To date, the *Archaea* are split into three kingdoms: the *Crenarchaeota* (e.g. thermophiles), the as yet uncultivated *Korarchaeota* (to date are exclusively hyperthermophiles) and the *Euryarchaeota* (e.g. methanogens, halophiles and thermophiles). Recently, the nanoarchaea have been identified from a submarine hot vent (Huber *et al.* 2002), though it has been argued that they belong to the *Euryarchaeota* (Brochier *et al.* 2005). *Archaea* are important inhabitants of extreme environments, though some members of the domains *Bacteria* and *Eukarya* can also inhabit extreme environments. Such organisms will be the focus of this study. However, *Archaea* are not just extremophiles as they have been detected in oceans, freshwater lake sediments, and a variety of soils (forest, grassland, ruderal and permafrost), and the human mouth (Schleper *et al.* 1997; Ochsenreiter *et al.* 2003; Lepp *et al.* 2004; DeLong 2005).

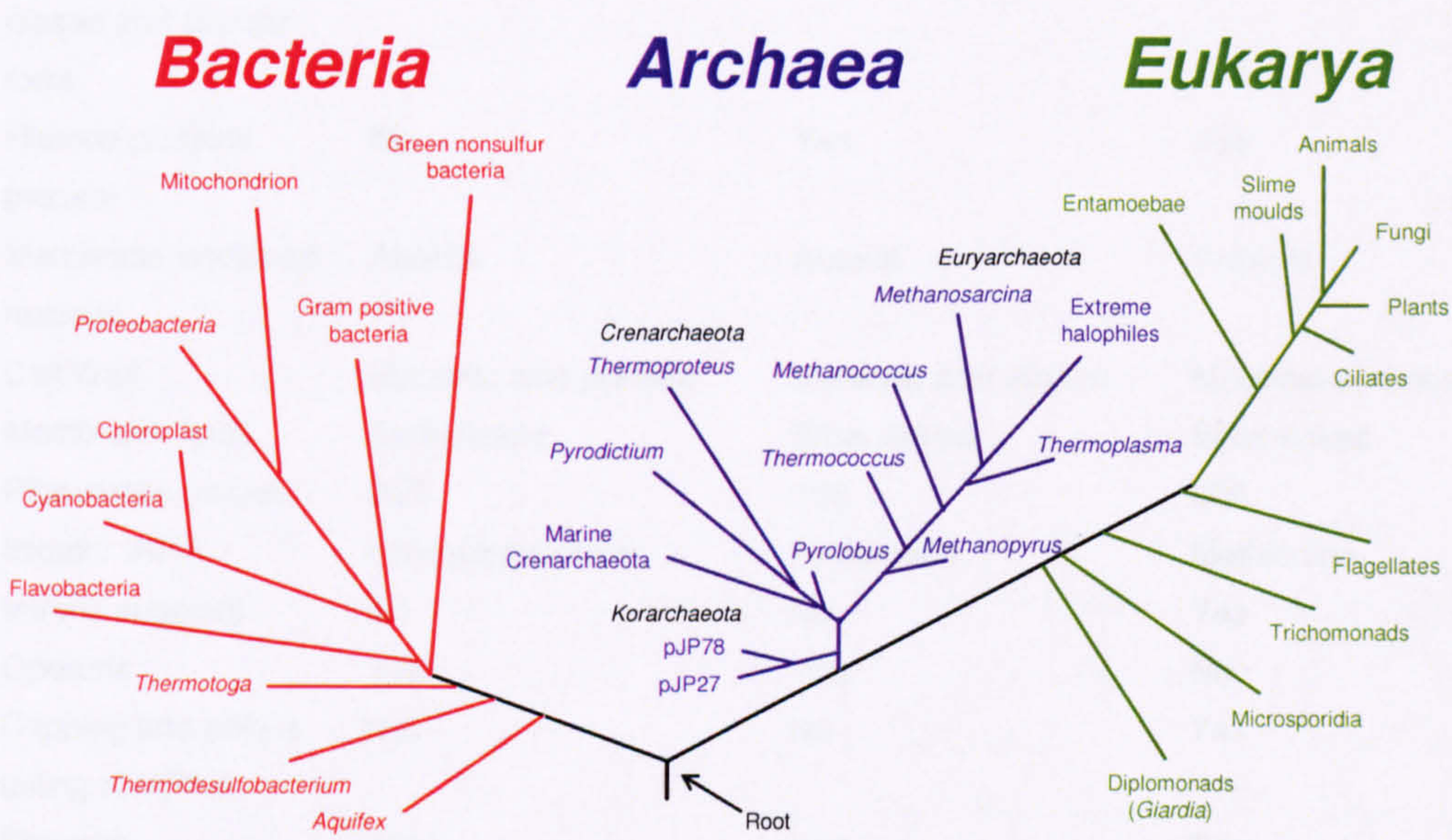


Figure 1 Universal Tree of Life.

The universal phylogenetic tree was determined from 16S ribosomal RNA (rRNA) sequences, which showed that living organisms belonged to one of the three domains: the *Bacteria*, *Archaea* and *Eukarya*. The *Bacteria* and *Archaea* contain only prokaryotic organisms. The hypothetical root or last universal common ancestor (LUCA) is indicated. Redrawn and adapted from Madigan *et al.* (2003).

Table 1 Morphological and Genetic Characteristics of the Three Domains.

The similarities between the *Bacteria*, *Archaea* and *Eukarya* are listed. Taken from Madigan *et al.* (2003).

Characteristic	***Bacteria***	***Archaea***	***Eukarya***
Prokaryotic cell structure	Yes	Yes	No
DNA in covalently closed and circular form	Yes	Yes	No
Histone proteins present	No	Yes	Yes
Membrane enclosed nucleus	Absent	Absent	Present
Cell Wall	Muramic acid present	Muramic acid absent	Muramic acid absent
Membrane lipids	Ester-linked	Ether-linked	Ester-linked
Ribosomes (mass)	70S	70S	80S
Initiator tRNA	Formylmethionine	Methionine	Methionine
Introns in genes	No	No	Yes
Operons	Yes	Yes	No
Capping and poly-A tailing of mRNA	No	No	Yes
Plasmids	Yes	Yes	Rare
Ribosome sensitivity to diphtheria toxin	No	Yes	Yes
RNA polymerases	One (4 subunits)	Several (8 – 12 subunits each)	Three (12 – 14 subunits each)
Transcription factors required	No	Yes	Yes
Promoter structure	Pribnow box	TATA box	TATA box
Sensitivity to chloramphenicol, streptomycin and kanamycin	Yes	No	No

Formation of Geothermal Features

A number of geothermal features are associated with tectonically active zones where crustal movements of the earth occur (Brock 1986). Such geothermal habitats are worldwide in distribution (Table 2). Terrestrial manifestations include hot springs and vapour discharges whose temperatures and chemical compositions are a reflection of the deep hydrothermal fluid and groundwater interactions, although the rates of outflow and chemical compositions of individual springs can undergo transient changes due to seismic activity and rapid burial and erosion in active volcanic settings. Fluid flow within these hydrothermal systems occurs through highly permeable fracture networks and lower permeable, thermally altered host rocks (Henley 1996). Hot springs are derived from deep system fluids, whereas warm springs are derived from the mixing of deep fluid with shallow groundwater (Henley 1996). Due to the localised nature to geothermal sources, geothermal features tend to occur in small areas called thermal basins (Brock 1986).

In non-volcanic areas, the formation of hot springs occurs when rainwater percolates through porous sedimentary rock, which gets heated up by geothermal heat, but does not boil because of lithostatic pressure. The temperature of the rock increases with increasing depth, which is known as a geothermal gradient. Eventually, the water encounters a fault and the heated water is forced to the surface and emerges as a hot spring. The chemical composition of such springs reflects the surrounding geology as the hot spring water contains dissolved ions that were picked up as it percolated through the rocks (Brock 1986).

In volcanic regions, water is heated by contact with magma. Magma that degasses near the surface causes the development vigorously boiling acidic hot springs (Henley 1996). If the water supply to a thermal source is scarce, steam rather than water may rise to the surface. Such geothermal features are called fumaroles (Brock 1986). Acidic hot springs tend to be sulphur rich and iron rich and are acidic due to the presence of sulphuric acid (H_2SO_4). Acidic springs are short-lived, lasting years to tens of years. Mineralisation in such springs tends to be low since the water of such springs is primarily superficial groundwater from the local region of the hills surrounding the basin, and has not undergone deep circulation. The water moves slowly down the hills into the basin and gets heated by steam from fumaroles. The steam contains hydrogen sulphide (H_2S), which gets oxidised to sulphuric acid, hence its acidic nature. In addition, H_2S can form elemental sulphur upon contact with atmospheric oxygen, which is deposited in the soil. Evaporation of the spring water means that acidity of

the fumarolic soils is high. Rain water can often leach the acidic soil further down the slope to collect in depressions and form acid springs. Such a spring has little or no water flow, and its temperature depends on whether it has contact with steam sources (Brock 1986).

Where the magma degasses much deeper, the acidic gaseous components (HCl and SO_2) are neutralised by reactions with rocks. This produces neutral to alkaline hot springs dominated by recycled surface and groundwater (Henley 1996) and tend to be flowing hot springs. The waters are highly mineralised; the dominant anions are generally bicarbonate and chloride (Brock 1986). Such hot springs are long lived, lasting hundreds of years. They are associated with geysers, which are hot springs that periodically erupt, throwing water and steam into the air. The eruption is the result of super-heated water that gets trapped in channels or vents that lead to the surface. The underground water is hottest deep within the channel, but cannot move due to the weight of water above. Steam is pushed upwards until eventually, the bubbles of steam lift the water to the surface causing an eruption. Eventually, the heat and pressure decrease and the eruption stops. Such eruptions occur at intervals as there must be sufficient water and steam. Vents that host geysers tend to be made of impermeable rock such as rhyolite.

In summary, geysers and flowing hot springs that occur at the bottom of a thermal basin are neutral to alkaline, and are highly mineralised springs of constant temperature. Further upslope, the acidic non-flowing or weakly flowing springs occur at the flanks of the hill; finally the sulphur and iron rich hot springs and furmaroles occur at the highest ground (Brock 1986).

Table 2 World Wide Distribution of Hot Springs and Geothermal Areas.

Region	Hot springs/ geothermal areas	References
Africa	Solfataric springs at Lac Abbé (Djibouti), shores of Lake Bogoria (Kenya), Ihimba hot springs, Kitagata hot springs, Nyamasizi hot springs, Semuliki hot springs, Buranga hot springs (Uganda)	(Windberger *et al.* 1989; Cioni *et al.* 1992; Bahati 2003)
Arctic	Kap Tobin, Nørrefjord, Rømerfjord (Greenland)	(Roeselers *et al.* 2007)
Asia	Domas, Cibuni, Cimanggu (Indonesia), Nakanoyu hot spring, Ganiba hot spring, Nakabusa hot spring, Beppu hot springs (Japan), Bor Khlueng Hot Spring (Thailand)	(Yamamoto *et al.* 1998; Baker *et al.* 2001; Nakagawa and Fukui 2002; Suzuki *et al.* 2002; Kanokratana *et al.* 2004)
Australia	Great Artesian Basin	(Andrews and Patel 1996)
Canada	Meager Creek hot springs	(Koch *et al.* 1999)
Europe	South-west Bulgaria, Loutra Langadas, Loutra of Nea Apollonia (Greece), Hengill geothermal area, Haegindi hot spring, Fluidir hot spring, Hverdstrokagerdi, Fludstrokir (Iceland), Termi di Agnano (Italy), Troll thermal springs (Norway)	(Emanuilova *et al.* 1993; Sonne-Hansen *et al.* 1993; Ferreira *et al.* 1997; Skirnisdottir *et al.* 2000; Papastefanou *et al.* 2001; Takacs *et al.* 2001; Hammer *et al.* 2005)
New Zealand	Kuirau Lake, Jaycee's fountain, Puara pool, Warbrick Terrace, Ohaaki Pool, Wairakei, Rotokawa, Champagne Pool, Waikite, Ngatamariki, Tokaanu, Orakei Korako, Waiotapu geothermal field	(Patel *et al.* 1985; Giggenbach *et al.* 1994; Jones *et al.* 1998a; Mountain *et al.* 2003)
Russia	Kotelnikovsky Hot Springs (Siberia)	(Bel'kova *et al.* 2005)
North America	Little Hot Creek, Casa Diablo, Shepherd's Pool, Imperial Spa (California), Valles Caldera (New Mexico), Yellowstone National Park: Obsidian Pool, Calcite Springs, Octopus Spring, Ravine Spring, Angel Terrace, Mammoth Hot Spring (Wyoming), Hunter's Hot Springs (Oregon), shores of Pyramid Lake (Nevada)	(Doemel and Brock 1977; Wannamaker 1997; Hugenholtz *et al.* 1998; Arp *et al.* 1999; Fouke *et al.* 2000; Miller and Castenholz 2000; Reysenbach *et al.* 2000b; Breitbart *et al.* 2004)
South America	Calabozos caldera, El Tatio (Chile), Las Trincheras, Mariara (Venezuela)	(Grunder *et al.* 1987; Jones and Renaut 1997; Horváth *et al.* 2000)

Tengchong and Rehai Geothermal Fields

The Province of Yunnan is located in the south west of China. Its capital city is Kunming, and east of Kunming is Tengchong (Figure 2; inset map). This area is geologically unique in that it is the only part of the Himalayan Geothermal Belt (that runs from south west Tibet to the western most part of Yunnan along the Indo-Eurasian suture zone) that is affected by Quaternary volcanism (Bai *et al.* 2001). Much geothermal exploration was conducted in the 1980s in Tengchong, which revealed that the stored heat and hot groundwater reserve were enormous, and has become the subject of many surveys as a potential source of renewable energy (Shiguang and Yong; Zhijie 1995).

The west Yunnan tectonic domain consists of many terranes that evolved between the Gondwanaland and Eurasian plates. (A terrane is a crustal block that has a distinctive geologic history that is different from the surrounding areas and that is usually bounded by faults). Data suggests that Tengchong was part of the Gondowanaland plate. However, at the beginning of the Triassic, the Tengchong terrane began to separate from the southern continent and subsequently adhered to the Eurasian plate. During the Cretaceous, the Gondowanaland plate completely separated and the Indian plate collided with the Eurasian plate, forming the Indo-Myanmar range. Tengchong is located in a high temperature zone behind this range (Zhijie 1995).

Volcanism started at Tengchong during the Pliocene or Miocene and continued throughout the Pleistocene, reaching a peak in the Early Pleistocene (Bai *et al.* 2001). The volcanoes in Tengchong are different from the volcanic rocks resulting from the continent to continent collision. They are divided into four types: the Pliocene alkaline basalt and dolerite, the Early Pleistocene calc-alkali andesite and dacite, the Middle Pleistocene alkalibasalt and the Late Pleistocene calc-alkali andesite-basalt (Zhijie 1995). The basement rocks are granite and gneiss (Bai *et al.* 2001).

The thickness of the earth's crust at Tengchong is about 40 – 50 km and is dominated by north-south faults (Bai *et al.* 2001) (see Figure 2). However, this is superimposed by many circular structures due to intrusions from the Jurassic to the Tertiary (Zhijie 1995). Geothermal fields are located in these circular structures, including the Rehai (Hot Sea) geothermal field, which is 13 km south west of Tengchong (Figure 2) (Zhijie 1995; Bai *et al.* 2001).

The Rehai (Hot Sea) field covers an area of ca. 8.5 km^2 (Zhijie 1995). It is controlled by the main faults running NNE to SSW and cross-cutting faults running NW to SE (Figure 3). The reservoir rock in the north is 69 Myr granite and in the south is Precambrian gneiss, over which lies Miocene sandstone and conglomerates. In the northern and western parts, Lower Pleistocene andesite cover sequence occurs, and in the east, Middle Pleistocene basalt is found (Bai *et al.* 2001).

The Rehai field can be divided into four main thermal areas from north to south: Luihuangtang (Sulphur Pool), Zaotanghe (Bathing River), Huangguaqing (Cucumber Gully) and Songmuqing (Pine Gully). All hot springs studied in this thesis are from the Zaotanghe area. These areas can be further divided into 20 subareas, each with different surface manifestations of geothermal activity (i.e. hot springs). These are detailed in Table 3 and Table 4. It is clear from the chemical analysis of the hot spring waters that a variety of hot springs occur in this field, with different temperatures (23 – 96.6°C) and pHs (5.5 – 8). The areas Lion's Head, Lion Pool and Frog's Mouth are featured in this thesis. The heat flow from this area has been estimated to be 147 MW, of which 81.35 MW of the heat energy is discharged by steam (Zhijie 1995).

Rehai field contains many faults, such as the Huangguaqing-Liuhuangtang fault, that are channels for the discharge of thermal fluids. Analysis of thermal waters has shown that they can be divided into four groups depending on the major dissolved ions: hot waters from springs are generally chloride-bicarbonate waters, diluted bicarbonate-chloride waters, steam-heated sulphate water, and non-thermal very dilute bicarbonate water. The reservoir temperature of the Rehai geothermal field was estimated at 230°C, though values will vary according to different geothermometers used (Zhijie 1995).

The Lang Pu Hot Pool is 6 km south west of the Rehai. Surveys were carried out to determine a possible connection between the Rehai and the Lang Pu Hot Pool, thereby identifying a common heat source. However, two separate investigations led to conflicting results, and to my knowledge, this has not been resolved (Zhijie 1995).

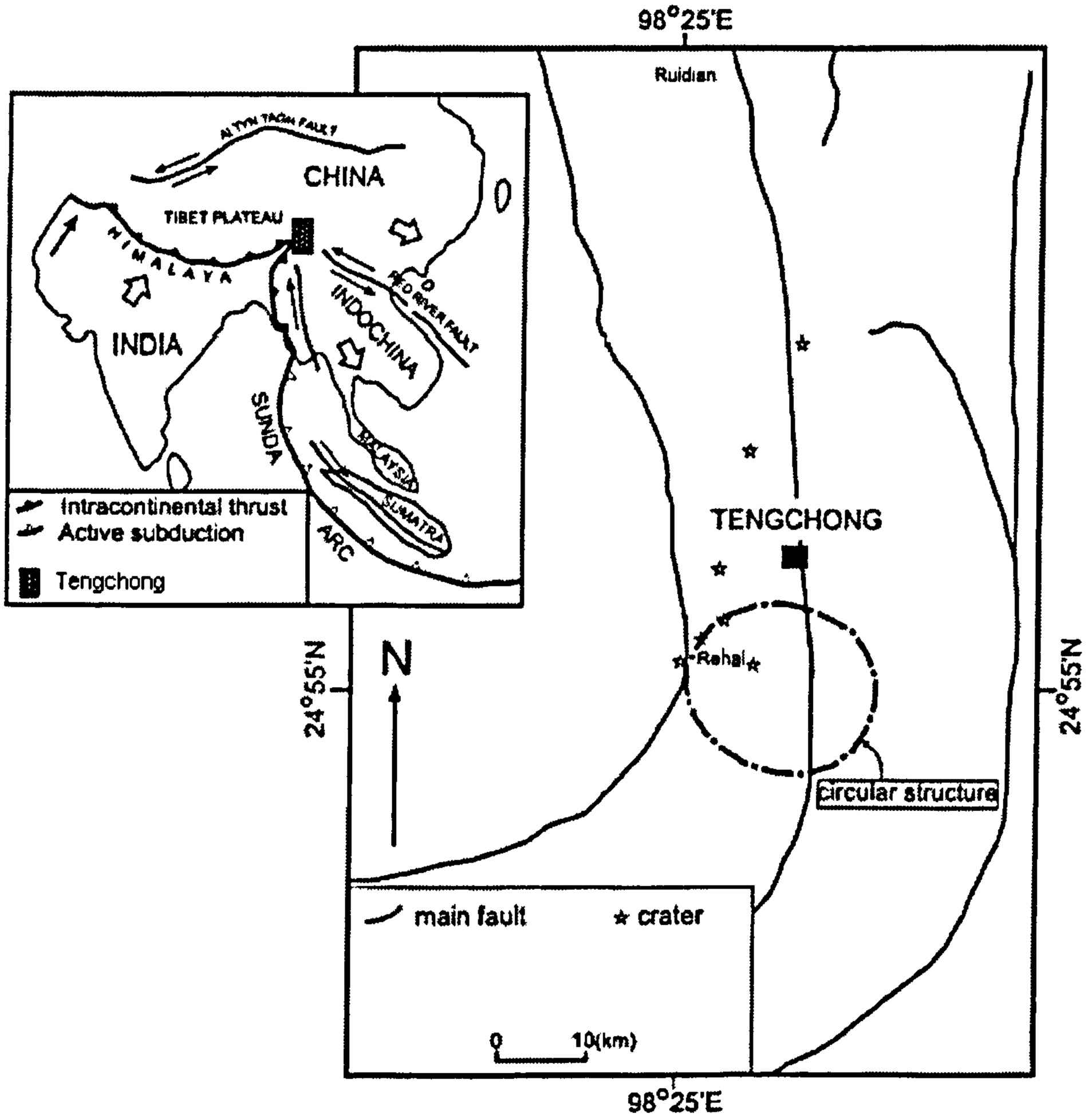


Figure 2 General Features of Tengchong and Rehai Geothermal Fields

Tengchong is located in Yunnan Province in south China, east from the capital city of Kunming. Tengchong is part of the Himalayan Geothermal Belt (inset map) and lies on a major fault line. Rehai is located 13 km south west of Tengchong in a circular structure. Taken from Bai *et al.* (2001).

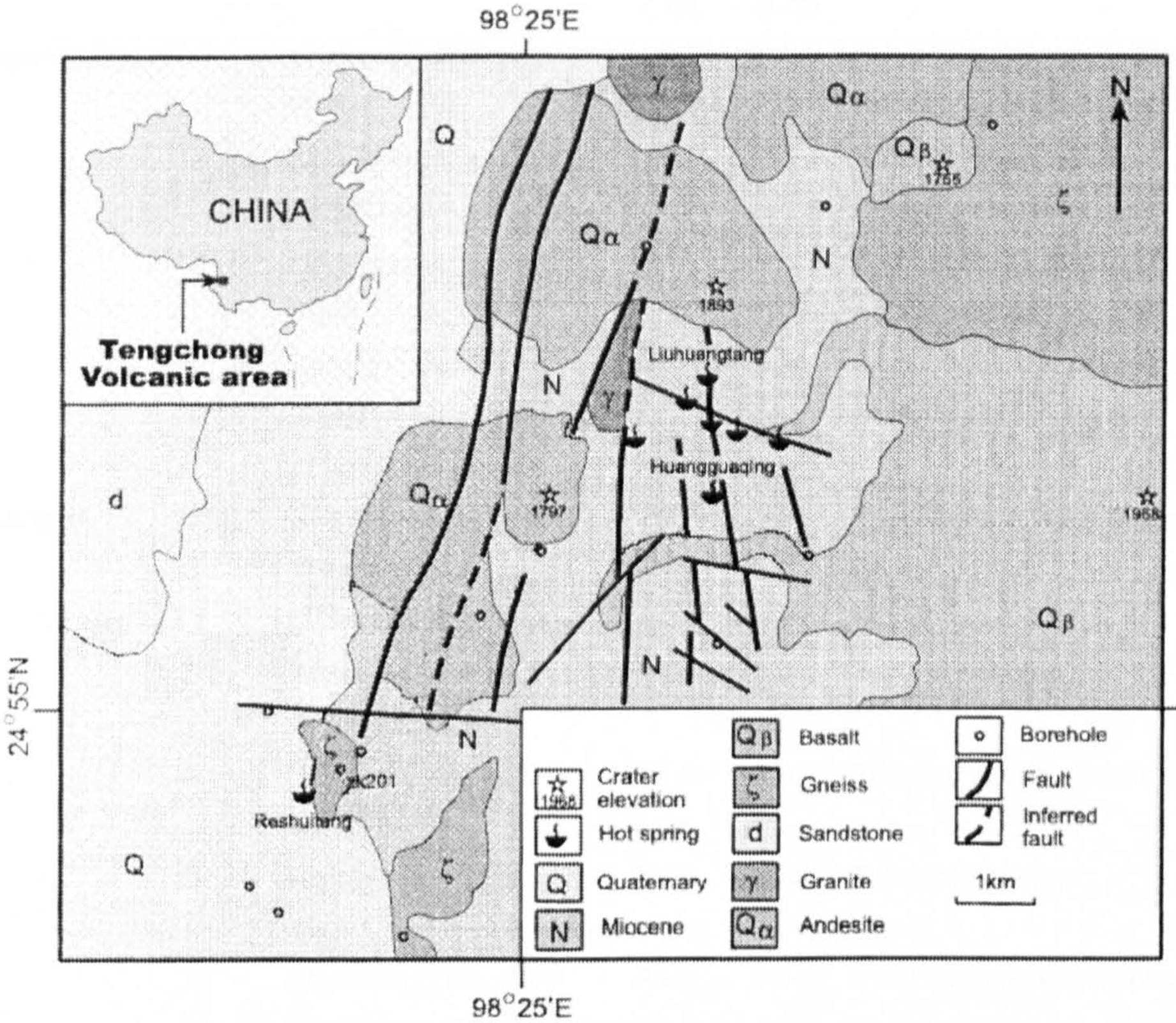


Figure 3 Geological details of Rehai Geothermal Field

The hot springs in this thesis are located in the Zaotanghe area (north of Huangguaqing). Some hot springs are indicated. There are many faults roughly running NNE to SSW and a cross-cutting fault running NW to SE. The rocks here are from the Miocene. Taken from Bai *et al.* (2001).

Table 3 Features of the Rehai (Hot Sea) Geothermal Field.

The field can be divided into four areas (Liuhuangtang, Zaotanghe, Huangguaqing and Songmuqing), which can be further subdivided into 20 subareas, each with their own geothermal characteristics. Hot springs from the Zaotanghe are featured in this thesis. Taken from (Zhijie 1995).

Area	No.	Location	Temp (℃)	Flow (L/s)	Manifestation
Liuhuangtang	1	Laogunguo (Dead Boiling pool)	92.7	0	Steaming ground ((30 m x 15 m), acid alteration
	2	Quhuangdong (Mining sulphur cave)	-	-	Steaming ground (50 m x 25 m), acid alteration
	3	Dagunguo (Big boiling Pool)	96.6	0.9	Boiling pool 3 m deep, sinter, acid alteration
	4	Zhonghuangdi (Sulphur-bearing ground)	95.3	0.21	Fumaroles, steaming ground (80 m x 40 m)
Zaotanghe	5	Shizitou (Lion's Head)	96.0	1.5	Travertine, boiling spring
	6	Yangingxiang (Spectacled springs)	94.0	5.0	Boiling springs, silica sinter (10 m x 40 m)
	7	Xianren Zaotang (Immortal baths)	43.7	1.8	3 warm springs
	8	Dabaiyan (Big white cliff)	61.8	0.91	Thermal springs, pyrite veins
	9	Bajiaoyuan (Banana Pla.)	54.0	0.38	5 thermal springs
	10	Shizitang (Lion pool)	84.5	0.5	3 hot springs, alterations
	11	Xiaobiayan (Small white cliff)	88.5	1.9	10 hot springs
	12	Zaotanghe (Bathing river)	96.0	13.0	Hydrothermal explosion, fumaroles; boiling springs, quartz veins
	13	Hamazui (Frog's mouth)	95.5	1.0	Boiling spring, travertine
	14	Liuhuangchang (Pyritic deposit)	-	-	Pyrite-binarite veins, altered host rocks
	15	Dadijiao	95.8	1.0	5 hot springs, 1 boiling spring
	16	Xiaozaotang (Minibath)	65.3	0.3	1 hot spring
	17	Zhongxiaosi (Loyal and filial temple)	44.9	0.1	Warm spring

Table 3 Continued.

Area	No.	Location	Temp (℃)	Flow (L/s)	Manifestation
Huangguaqing	18	Huangguaqing (Cucumber gully)	96.0	-	Fumaroles, acid alteration (1.2 km x 100 m)
	19	Yangjiapo	23.0	0.6	Tepid spring
Songmuqing	20	Songmuqing (Pine gully)	-	-	Alteration area of 4 km2, sulphurite, pyrite, warm ground

Table 4 Chemical Composition of Hot Spring Waters.

The chemical composition of thermal waters was determined in ppm. The locations of the springs (3, 5, 6, 7, 8, 10, 11, 12, 13, 15 and 19) are given in Table 3. Taken from Zhijie (1995).

No.	3	5	6	7	8	10	11	12	13	15	19
Temp (℃)	88	91	95	42	48.5	83	69.5	92	95.8	95.8	23
pH (field)	8	5.5	8	6	7.5	7.5	6.5	7	7.5	7.5	-
pH (lab)	8	3.1	9.4	7	7	7.5	7.3	7.8	8.55	8.2	6.4
Na	840	58	680	130	180	420	380	440	340	450	5
K	120	23	110	25	30	55	50	63	61	42	4
Ca	0.1	9.43	0.29	2.86	3.4	4.72	2.86	2.72	10.7	0.86	6.43
Mg	0.02	3.21	0.02	0.18	0.21	0.09	0.11	0.22	1.79	0.17	1.83
LI	8.31	0.60	6.20	6.20	2.10	3.90	3.80	3.80	3.40	4.00	0.04
CO_3	0	0	261	0	0	0	0	0	37	22	0
HCO_3	1160	0	700	250	353	637	606	615	512	522	44
Cl	698	16	559	57	93	256	217	295	269	372	1
SO_4	31	303	11	18	56	38	26	47	24	27	16
F	20	1.2	14	2.2	3.5	8.3	7	7.8	7	9.7	-
HBO_2	52	3	42	8	11	17	16	19	18	14	-
SiO_2	450	195	340	105	108	175	163	180	170	118	51

Definition of Thermophilic Microorganisms

Temperature is one of the most important variables in the environment; if it is too hot or too cold, then the organisms cannot grow. Classification of organisms based on temperature produces four main groups: the psychrophiles, mesophiles, thermophiles and hyperthermophiles.

The minimum and maximum temperatures that allow growth of organisms can vary greatly between organisms, and is a reflection of the temperature range of their habitats. Growth starts at T_{min} and will increase almost linearly for about 15°C to 25°C up to the optimum growth temperature T_{opt}, after which it falls abruptly for the next 5°C to 10°C to T_{max}, above which no growth occurs. These are called the cardinal temperatures, and each specific organism has a characteristic temperature range. However, attempts to define groups according to these cardinal temperatures have proved difficult since there are overlapping exceptions (Madigan *et al.* 2003).

The current definitions of these groups of organisms are as follows: a psychrophile has a T_{opt} of 15°C or lower, with a T_{max} below 20°C and a T_{min} at 0°C or lower, and a mesophile has a temperature range of 20°C and 45°C (Madigan *et al.* 2003).

A thermophile is generally defined as an organism capable of growth at a high temperature. It has been defined as having a T_{min} and T_{max} at 45°C and 80°C respectively, with a T_{opt} at 55°C to 60°C (Madigan *et al.* 2003). These boundaries have some ecological and evolutionary basis. Temperatures lower than 50°C are widespread on earth, whereas temperatures higher than 55°C to 61°C are much rarer associated exclusively to geothermal habitats. In addition, 61°C is the upper temperature limit for most eukaryotic life (plants and animals); hence only prokaryotes are found above this temperature (Brock 1986). Thermophiles can be further classified as hyperthermophiles, which have an T_{opt} of 80°C or higher (Madigan *et al.* 2003).

Phylogenetic Diversity of Thermophiles

Eukarya

Few *Eukarya* (algae, fungi and protozoa) have been isolated and studied from hot springs. *Cyanidium caldaria* is a thermophilic alga that can also grow in extremely acidic environments. The thermophilic fungus *Paecilomyces varioti* has been exploited commercially for the production of protein for animal feed (Friedman 1992).

Archaea

Branches of the *Archaea* that are observed in thermal environments are the *Crenarchaeota*, *Korarchaeota*, *Nanoarchaeota* and some members of the *Euryarchaeota* (*Thermoplasmata*, *Thermococci*, *Archaeglobi*, *Methanopyri* and other methanogens).

The *Crenarchaeota* consist of a single class, *Thermoprotei*, which is subdivided into three orders: the *Thermoproteales*, the *Desulfurococcales* and the *Sulfolobales*. They are obligately thermophilic, with growth occurring from 70°C to 113°C. Most members of the *Crenarchaeota* metabolise sulphur (Garrity and Holt 2001b). The order *Sulfolobales* have acidophilic members, while the orders *Thermoproteales* and *Desulfurococcales* are moderately acidophilic (Reysenbach 2001a).

Several classes of the *Euryarchaeota* contain thermophilic members. These are *Thermoplasmata*, *Thermococci*, *Archaeglobi* and *Methanopyri*. The *Thermoplasmata* grow optimally around 60°C and includes the only wall-less representative of the *Archaea*, *Thermoplasma*, and the extremely acidophilic thermophile, *Picrophilus* (Reysenbach 2001b). The *Thermococci* grow optimally between 75°C and 100°C. Most members of this class require sulphur for growth (Zillig and Reysenbach 2001). The *Archaeglobi* grow optimally around 80°C. They are strictly anaerobic and neutrophilic, and can utilise sulphate, sulphite, thiosulphite and/or nitrate as electron acceptors, but elemental sulphur inhibits growth (Huber and Stetter 2001a). The *Methanopyri* grow above 80°C. They are strictly anaerobic and can produce methane from CO_2 and H_2 (Garrity and Holt 2001a). Other moderately thermophilic to hyperthermophilic methanogens within the *Euryarchaeota* include *Methanobacterium*, *Methanothermobacter*, *Methanothermus*, *Methanothermococcus*, *Methanocaldococcus* and *Methanotorris* (Garrity and Holt 2001d).

The *Korarchaeota* are hyperthermophiles that form a separate deeply branching group within the domain *Archaea*. They are defined solely by their 16S rRNA gene sequence. Members of this group have been detected in Yellowstone National Park in several boiling source areas with temperatures up to 93°C (Barns *et al.* 1994). To date, there are no pure cultures of *Korarchaeota*, so their physiological characteristics are as yet unknown. However, they currently exist in a mixed enrichment culture (Brunk and Eis 1998). For this reason, their role within the thermal habitat has not been elucidated.

The first cultivated nanoarchaeon was discovered by Huber *et al* in 2002 from a submarine hot vent. This archaeon, called *Nanoarchaeum equitans*, is 400 nm in diameter and grows attached to a specific archaeal host of the genus *Ignicoccus*. Analysis of its 16S rRNA gene sequence placed it in a separate branch to the *Crenarchaeota* and the *Korarchaeota* (Huber *et al.* 2002). However, recent studies have suggested that it is a branch within the *Euryarchaeota* (Brochier *et al.* 2005). Since this discovery, nanoarchaeal lineages have been detected in an active deep-sea hydrothermal vent (McCliment *et al.* 2006). To date, nanoarchaea have not been detected in terrestrial thermal environments.

Bacteria

Hyperthermophilic *Bacteria* belong to the phyla *Aquificae*, *Thermotogae*, *Thermodesulfobacteria*, *Deinococcus-Thermus*, *Chloroflexi*, *Thermomicrobia* and *Deferribacteres*. These bacterial phyla are deeply branching. A single genus within the *Nitrospira* is also thermophilic. In addition, members of the cyanobacteria (which are oxygenic phototrophs) are found in neutral to alkaline hot springs (see below).

The *Aquificae* are moderately thermophilic to hyperthermophilic and are microaerophilic chemolithotrophic hydrogen oxidisers. Most grow between pH 6.0 and 8.0, with the exception of *Hydrogenobacter acidophilum* (Reysenbach 2001c). The *Thermotogae* are extremely thermophilic anaerobic fermenters, utilising a range of organic substrates. They grow optimally at pH 6.5 to 7.5. The cells posses a characteristic outer sheath or 'toga' (Reysenbach 2001d). The *Thermodesulfobacteria* is represented by a single genus, *Thermodesulfobacterium*. They are thermophilic, strictly anaerobic sulphate reducers that utilise lactate and pyruvate as electron donors, and sulphate or thiosulphate as electron acceptors (Garrity and Holt 2001c). The order *Thermales* within the *Deinococcus-Thermus* phylum contains thermophilic members with optimum growth temperatures between 50°C and 75°C. They are aerobic with a respiratory type of metabolism, though some grow anaerobically with nitrate and nitrite. Some members can oxidise sulphur compounds. They are often found in neutral to alkaline thermal environments (Rainey and Da Costa 2001). The *Chloroflexi* are filamentous bacteria that contain two orders: *Chloroflexales* and *Herpetosiphonales*. Members of the former order contain bacteriochlorophyll (BChl) and so are obligate or facultative phototrophs. Four genera within the *Chloroflexales* are capable of anoxygenic photosynthesis (Castenholz 2001a). The *Thermomicrobia* contains a single

representative, *Thermomicrobium roseum*. It grows optimally between 70°C and 75°C and pH 8.2 to 8.5. It is obligately aerobic with a respiratory type of metabolism, utilising oxygen as the terminal electron acceptor (Garrity and Holt 2001e). A single genus within the *Nitrospirae* is thermophilic; *Thermodesulfovibrio* is a sulphate reducing bacterium (SRB) that grows optimally at 65°C, and is obligately acidophilic and anaerobic (Maki 2001). The *Deferribacteres* contains slightly thermophilic to thermophilic members. They respire anaerobically using a range of terminal electron acceptors, including Fe(III), Mn(IV), S, Co(III) and nitrate (Garrity and Holt 2001d).

Ecology of Hot Springs: Microbial Mats

Due to a lack of predation from eukarya, cooperative populations of prokaryotes develop in hot springs to form microbial mats (Doemel and Brock 1977; Varnam 2000). The type of community that develops is primarily influenced by temperature, although pH and concentration of dissolved ions are also factors. Microbial mats only develop in neutral to alkaline hot springs, and not acidic hot springs (Varnam 2000).

Lower Temperature Mats

Lower temperature here refers to mats that develop at around 55°C or higher. Laminated mats composed of distinct layers of bacteria tend to develop in these habitats. Many examples are found in the hot springs in Yellowstone National Park in Wyoming, USA, particularly in Octopus Spring (Doemel and Brock 1977).

The top layer of laminated mats is made up of filamentous cyanobacteria; hence they are often called cyanobacterial mats. These oxygenic phototrophs are responsible for the primary production in the laminated mat and make up part of the photic zone. They have been the focus of many studies, including species adaptation to different microenvironments, UVR and disturbance (Doemel and Brock 1977; Ferris *et al.* 1997; Ward *et al.* 1998; Norris *et al.* 2002). Unicellular cyanobacteria, *Synechococcus*, are often found associated with this layer. Early selective enrichment led to the belief that cyanobacterial mats consisted solely of *Synechococcus* (Ward *et al.* 1998). These have also been the subject of many studies including vertical structure and discrete genetic distribution of species within the photic zone (Ramsing *et al.* 1997; Ramsing *et al.* 2000).

The underlying orange-brown or red layer consists of Filamentous Anoxygenic Phototrophs (FAPs) and make up the rest of the photic zone. They are part of the green non-sulphur bacteria, but appear red due to the presence of BChl a, rather than green (Castenholz 2001a; Hanada 2003). FAPs found in hot springs are taxonomically classified into three genera: *Chloroflexus*, *Heliothrix* and *Roseiflexus*. *Chloroflexus* tend to be associated with cyanobacteria particularly *Synechococcus* (Varnam 2000), usually making up the brown underlying layer (Hanada 2003). In addition to anoxygenic photosynthesis, it is believed that *Chloroflexus* provides physical cohesiveness to the mat (Ward *et al.* 1998). *Roseiflexus* tend to make up the red layer under this brown layer and have been found in a number of mats in Yellowstone National Park (Boomer *et al.* 2002; Hanada 2003).

Studies of the pigments contained in these layers have shown that all three populations absorb different parts of the light spectrum. Chlorophyll a contained in cyanobacteria has an absorption peak at 670 nm; BChl c in *Chloroflexus* has a peak at 740 nm and BChl a in *Roseiflexus* has peaks at 800 and 868 nm. Infrared light passes through cyanobacterial mats that are 5 mm thick, thereby allowing *Roseiflexus* to grow phototrophically, even under a thick cyanobacterial layer. The different absorption ranges means that there is no competition between population and allows them to coexist (Hanada 2003).

The bottom layers of the mat consist of anaerobic sediments containing different bacteria including anaerobic fermenters and SRB. Other members of the Octopus Spring cyanobacterial mat have been listed, and include *Planctomycetes*, *Thermus-Deinococcus*, *Proteobacteria* (*Alpha*, *Beta*, *Gamma* and *Delta* subdivisions), *Spirochetes*, Gram Positive bacteria (e.g. *Thermoanaerobacter*), *Thermodesulfobacterium*, *Leptospirillum* and *Nitrospira* (Ward *et al.* 1998).

Studies on the variations in carbon metabolism in Mushroom Spring in Yellowstone National Park have shown that these populations (cyanobacteria, FAPs and anaerobic bacteria) cooperate in the cycling of carbon compounds according to different light intensities (van der Meer *et al.* 2005). Presumably, other nutrient cycles occur in this mat allowing this organised community to thrive at high temperatures.

Although cyanobacterial mats predominantly contain bacteria, methanogenesis has long been recognised in laminated mats (Ward 1978) and so the only *Archaea* that have previously been

found associated with cyanobacterial mats are methanogens such as *Methanobacterium thermoautotrophicum* (Ward *et al.* 1998).

High Temperature Mats

Higher temperature here refers to mats that develop at around 70°C or higher. Streamers are formed in this habitat by the aggregation of bacteria and, in some cases, archaea. Though temperature is the most important factor in the development of streamers, it appears that significant levels of sulphide are also required (Skirnisdottir *et al.* 2000). Moreover, streamers only appear to develop in flowing streams and effluents.

Nakanoyu and Ganiba hot springs in Japan contained white streamers in slightly alkaline shallow streams. Further analysis of these streamers revealed that they consisted of sulphur particulates associated with long 'sausage-shaped' bacteria, which gave them a white 'ruffled fur-like' or 'turf-like' appearance, hence they are also called 'sulfur-turf' mats. The elongated bacteria were found to be related to the phylum of the *Aquifex-Hydrogenobacter* complex, though they form their own distinct lineage within the *Aquificae* (Yamamoto *et al.* 1998). It was shown that *Aquificae* are important in the biological oxidation of dissolved sulphide, which is subsequently oxidised to elemental sulphur by molecular oxygen (Nakagawa and Fukui 2002); hence the appearance of sulphur particulates in the streamers. It was recently shown that the carbohydrate fraction of streamers is cellulose, indicating that the bacteria in these filaments are producers of cellulose as extracellular polysaccharide, which may aid in mat building (Ogawa and Maki 2003).

Streamers can occur in a range of colours: white, blue, pink, grey and black (Reysenbach *et al.* 1994; Yamamoto *et al.* 1998; Reysenbach *et al.* 2000b; Skirnisdottir *et al.* 2000; Takacs *et al.* 2001; Nakagawa and Fukui 2002) Further studies have shown that a range of other bacteria and archaea are associated with the *Aquificae* streamers, though no streamer community is identical. The white streamers described above also contained members of the *Firmicutes* and *Proteobacteria* (Yamamoto *et al.* 1998). In contrast, blue streamers found in Haegindi and Fluidir hot springs in Iceland were a monoculture of bacteria related to the *Aquificae*; no other bacteria or archaea were detected (Takacs *et al.* 2001). Pink *Aquificae* filaments found in Octopus Spring in Yellowstone National Park also contained bacteria related to the *Thermotogales* (Reysenbach *et al.* 1994). A black filamentous community found in Calcite Springs, again from Yellowstone National Park, contained members of the *Crenarchaeota*,

related to *Thermophilum pendens*, *Pyrobaculum islandicum* and *Desulfurococcus mobilis*, though *Aquificae* still comprised > 95% of the community. Moreover, *Korarchaoeta* were detected in this community, which highlighted the fact that *Aquificae* and *Korarchaeota* can coexist in certain communities (Reysenbach *et al.* 2000b). White streamers found in the Hengill geothermal areas of Iceland was also dominated by *Aquificae*, but also contained representatives of the *Thermodesulfobacterium* group, *Thermus-Deinococcus* group, *Thermotogales* and *Nitrospira*. In addition, 77% of the archaeal sequences were closely related to *Korarchaeota*, and *Crenarchaeota*, related to *Thermofilum pendens* and *Desulfurococcus* (Skirnisdottir *et al.* 2000).

The differences observed between *Aquificae* dominated streamer communities have been explained by the geochemical differences and therefore physiological activity, in the environments that they inhabit. For example, the Calcite Springs streamers that showed high archaeal diversity contain a lot of iron, suggesting that iron may be important in the growth of certain organisms. In addition, microbial composition could be attributed to stream flow, UV and pH. However, whether these other bacteria and archaea are a functioning part of the streamer community remains to be seen. It has been postulated that conditions are favourable for aerobic sulphur and hydrogen oxidising bacteria at the surface and anaerobic sulphur and sulphate reducers inhabit the darker undermass (Skirnisdottir *et al.* 2000).

Saline Lakes

Salinity is defined as the sum total of all ion concentrations (Williams and Sherwood 1994). The conventional value of > 3 g/L is widely accepted as environments considered saline (Williams 1996). This is because it is near the calcite branch point, the salinity at which humans can taste salt and it is the salinity below which biota of higher salinities are not found (see below) and above which freshwater biota disappear (Williams 1996).

There has been increased interest in salt lakes partly because of the recognition that they represent a significant fraction of global water compartments. Salt lakes make up 0.008% of the global waters, just slightly below freshwater lakes and rivers, which contributes 0.009% (Williams 1996). However, salt lakes are important for a number of other reasons. They are an important source of minerals, especially halite (NaCl), soda (Na_2CO_3), uranium, zeolites, lithium and borax. They are a source of biochemical products such as β-carotene from *Dunaliella* used as an antioxidant and food colouring, and protein from *Spirulina* as a human food source. They also provide foodstuff for aquaculture (*Artemia* cysts). They are important in biotechnology, where many extracellular enzymes such as proteases, amylases and cellulases are used in laundry detergents. In addition, some are important habitats for flamingos and fish, while others are attractive environments that attract tourism (Williams 1996; Oren 2002a; Rees *et al.* 2003). Full reviews of the biotechnological potential of salt lakes have been reviewed elsewhere (Horikoshi 1999; Oren 2002a).

Salt lakes are common and occur on all continents. Table 5 is a compilation of some of the salt lakes around the world.

Table 5 Salt Lakes of the World.

This is a list of salt lakes around the world. Data compiled from (Williams 1996; Bowman *et al.* 2000a; Grant and Jones 2000).

Region	Lakes
North and Central America	Great Salt Lake, Salton Sea, Pyramid Lake, Big Quill, Mono Lake, Searles Lake, Walker Lake, Soap Lake, Big Soda Lake, Owens Lake, Borax Lake, Harney Lake, Summer Lake, Pyramid Lake, Union Pacific Lakes (Green River), Ragtown Soda Lakes, Lake Texacoco
Africa	Lake Fezzun, Wadi Natrun Lakes, Lake Aranguadi, Lake Kilotes, Lake Abiata, Lake Shala, Lake Chilu, Lake Hertale, Lake Metahara, Lake Natron, Lake Magadi, Lake Nakuru, Lake Bogoria, Lake Elmenteita, Lake Simbi, Lake Sonachi (Crater), Lake Oloidien, Lake Embagi, Lake Magad, Lake Manyara, Lake Balangida, Basotu Crater Lake, Lake Kusare, Lake Tulusia, Momela Lakes, Lake Lekandiro, Lake Reshitani, Lake Lgarya, Lake Ndutu, Lake Rukwa North, Lake Katwe, Lake Mahenga, Lake Kikoronga, Lake Nyamunuka, Lake Munyanyange, Lake Murumuli, Lake Nunyampaka, Lake Bodu, Lake Rombou, Lake Dijikare, Lake Monboio, Lake Yoan
Russia and Central Asia	Caspian Sea, Aral, Balkhash, Issy-Kul, Chany, Alakul, Tengiz
China	Qinghai, Dabuxun (Qinghai), Lop Nor (Sinkiang), Namu, Selin, Zarinanmu (Tibet), Dalai, various 'nurs' or 'nors' (Inner Mongolia)
Iranian Plateau and Middle East	Lake Urmia, Lake Niriz (Iran), Van Golu Lake, Lake Van, Lake Salda (Turkey), Dead Sea (Israel/Jordan)
Australia	Lake Eyre North, Lake Corangamite, Lake Bullen Merri, Lake Gnotuk, Red Rock Lake, Lake Werowrap, Lake Chidnup
Antarctica	Deep Lake, Organic Lake, Ekho Lake
Europe	Lake Feher (Hungary)
Siberia	Tanatar Lakes, Lake Baikal, Lake Khatyn

Formation of Saline Environments

There are two types of saline environments: thalassohaline, which are derived from marine waters, and so part of its salt composition is similar to seawater, and athalassohaline, which are formed by the weathering and leaching of minerals from rocks into drainage basins. A standing body of water results when there is little outflow from the basin. Since the input flow contains salts and the evaporating water contains none, salt accumulates in the lake and so are a reflection of the surrounding geology, geography and topography (Williams 1996; Grant *et al.* 1998). Hence, the chemical composition of thalassohaline brines is relatively uniform, whereas athalassohaline brines may vary widely in their chemical composition.

Saline lakes are widely distributed and the geological processes that lead to their development involves pH as the major factor (see Figure 4) (Grant *et al.* 1998). A solution of carbon dioxide forms carbonic acid, which is a weak acid that undergoes ion exchange with the surrounding rock; this causes minerals to be leached out. The final pH is influenced by the concentration of calcium ions (and to a lesser extent, magnesium ions) in the surrounding rocks. Calcium ions form insoluble calcite ($CaCO_3$), which removes alkaline carbonate ions from solution; hence brines that have a high concentration of calcium ions have a neutral pH. The Great Salt Lake in Utah is one such environment. Where the geology is deficient in calcium and magnesium ions (where rocks of volcanic or metamorphic origin are present and those of sedimentary origin are absent), carbonate becomes the dominant ion. An increase in carbonate ions results in the precipitation of calcium, then magnesium removing divalent cations from solution and allowing more soluble carbonates of sodium and potassium to accumulate. This results in soda lakes such as Lake Magadi in Kenya. In contrast to salt lakes, soda lakes are geologically very recent. In other environments, the precipitation of minerals can contribute to the release of H^+ ions, therefore reducing the pH. For example, the Dead Sea is slightly acidic because the precipitation of magnesium minerals such as sepiolite (Grant *et al.* 1998). Furthermore, a high calcium ion alkaline environment can develop from low temperature weathering of calcium and magnesium containing silicate minerals, olivine and pyroxene, which releases calcium ions and hydroxide ions. Magnesium is precipitated and carbonate is removed as calcite, but since there is an excess of calcium ions, this results in a $Ca(OH)_2$-dominated brine with a pH of around 11 (Jones *et al.* 1994). Clearly, the type of brine that develops influences the microbiota.

Salterns are used for the harvesting of salt and/or soda. Here, the saline environments are developed into evaporation ponds to allow the precipitation of salts. This occurs in the order calcite ($CaCO_3$), gypsum ($CaSO_4.2H_2O$), halite (NaCl), sylvite (KCl) and finally carnallite (KCl, $MgCl_2.6H_2O$) (Grant *et al.* 1998). Multi-pond salterns provide a range of environments with different salinities from near seawater to approaching saturation (crystalliser ponds). They are a series of connected ponds; as water evaporates and salinity increases, water is pumped or fed by gravity to the next pond, hence the salinity in each pond is essentially kept constant (Pedrós-Alió *et al.* 2000).

Salt lakes often contain significant amounts of dissolved organic carbon (DOC). Death and decay of organic matter from algae and plants growing nearby may contribute to the carbon input (Ollivier *et al.* 1994). The major sources of organic compounds come from the organisms themselves. For example, the decomposition of brine shrimp leads to the deposition of chitin from their exoskeletons, which is a substrate for halophilic bacteria (Liaw and Mah 1992). Furthermore, the decomposition of bacterial cell walls releases sugars, proteins and lipids (Ollivier *et al.* 1994). A further possible source of organic matter are compatible solutes (osmolytes) contained in the organisms. These are organic molecules that help with coping with a high salt environment. Release of such solutes is a further source of organic substrates (Ollivier *et al.* 1994; Grant *et al.* 1998).

Saline waters are often contain very little dissolved oxygen; it has been calculated that oxygen solubility at 25°C is 2 mg/L at halite precipitation (Grant *et al.* 1998). Hence salt lakes are high salt, low oxygen environments with a range of organic substrates available for growth.

Other saline environments include wall paintings, desert plants, oil-field brines and salted food (Grant *et al.* 1998). However, the focus of this study is on saline and hypersaline lakes, soda lakes and salterns.

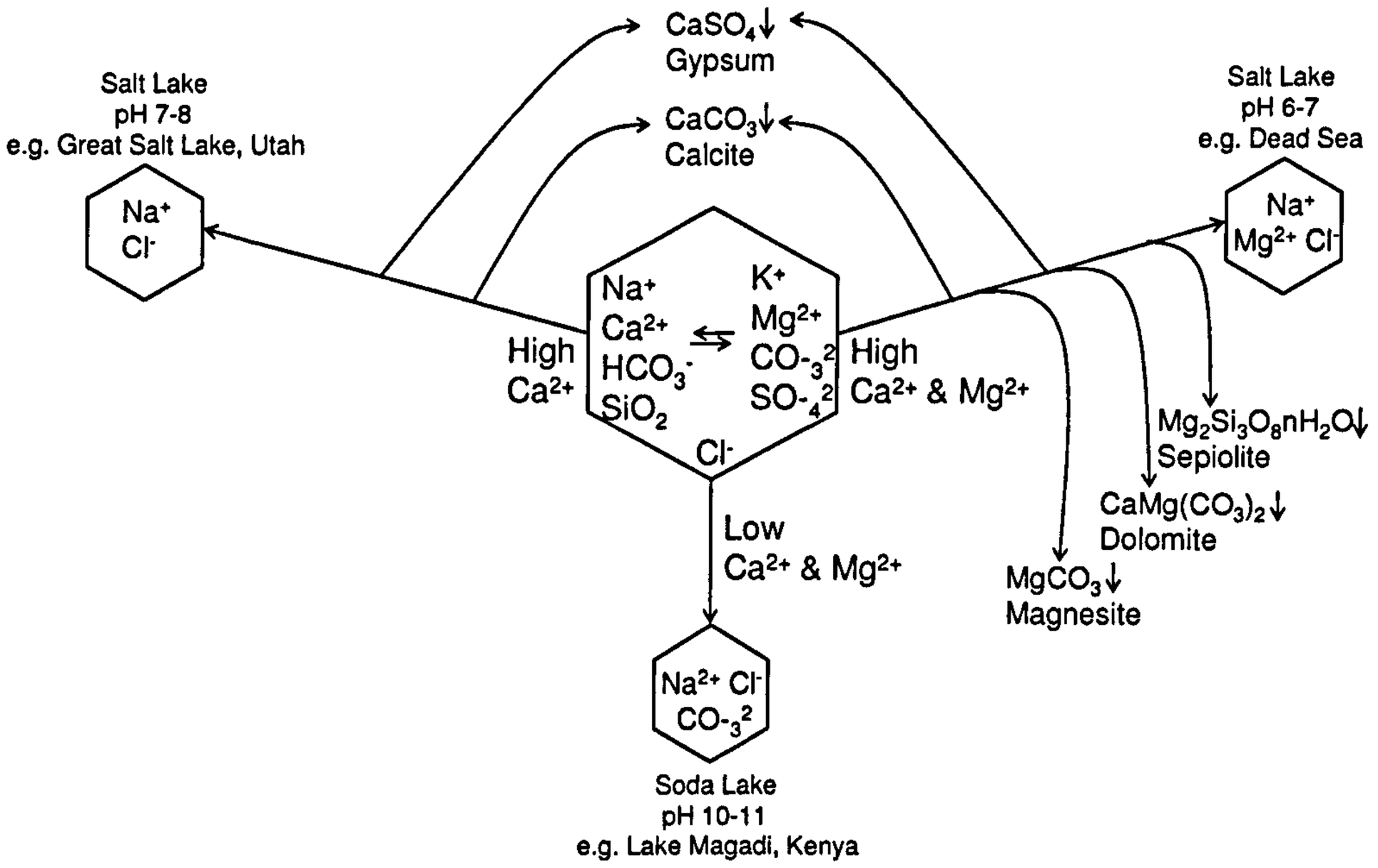


Figure 4 Schematic Representation of the Genesis of Saline and Alkaline Lakes.

Redrawn from (Grant *et al.* 1998).

Inner Mongolian Salt Lakes

More than half of China's lakes are saline (salinity > 3 g/L) (Williams 1991). Most of these occur in the west and north east, though coastal pans are located in the east (Figure 5). These lakes can be split into four main regions: Qinghai-Tibet Plateau (I), North western (II), North-central (III) and Eastern (IV) (see Figure 6). These regions reflect the geological structures, geomorphology, climate and other major factors. The Qinghai-Tibet Plateau is at an elevation of > 4000 m and is adjoined by the Himalayas in the south, the Kunlun and Qilian Mountains in the north and the Yanbajain-Qinghai Lake in the east. The North western region lies north of the Qinghai-Tibet Plateau, where salt lakes appear in basins at lower elevations (2000 - 5000 m). It adjoins Russia in the north and west and Helan Mountains in the east, and includes the Tarim Basin in Sinkiang. The North-central region adjoins Mongolia in the north, the Taihang Mountains-Greater Khingan Mountains in the east, and the Baiyu Mountain to the south, the Inner Mongolian Plateau, and Hulum Buir Basin. Finally, the Eastern region consists of a closed basin in Eastern China where there are some small saline lakes and saline groundwater (Zheng *et al.* 1993).

Tectonic movements have been the greatest influence in the development of salt lakes in China, but aeolian activity and deflation have also been a key factor (Williams 1991). Tertiary collision between the Indian and Laurasian tectonic plates caused a major uplifting of the Qinghai-Tibet Plateau, which consequently blocked rivers that were formerly discharging into the Indian Ocean. This caused changes in the climate such as a reduction in precipitation resulting in a change from warm and humid to cold and arid. Hence many salt lakes in China are found in fault lines and tectonic depressions where the arid climate caused the development of many endorheic drainage basins (closed basins) (Williams 1991).

The lakes in this study are situated on the Inner Mongolian Plateau, north west of Beijing in the north central region, and include Lakes Bagaejinnor, Chagannor, Chahannor, Ejinnor, Erliannor and Shangmatala. All salt lakes located in Inner Mongolia are athalassohaline. The north central region contains 116 salt lakes that predominantly contain trona ($Na_2CO_3.NaHCO_3.2H_2O$), mirabilite ($Na_2SO_4.10H_2O$), halite and other common salts (Zheng *et al.* 1993). Unlike most Chinese salt lakes, aeolian activity and deflation of recent origin and not of tectonic movement caused the development of large depressions in which many salt lakes could develop (Williams 1991). Two fault lines of tectonic origin running south west to north east in this area form depressions in which salt lakes can develop. These are the Erenhot

Fault and the Xilinhot Fault, which controls much of this basin (Yu *et al.* 2001). However, this region is tectonically stable (Zheng *et al.* 1993).

From the Quaternary arid centre of the Qaidam Basin-Tarim Basin (Figure 5) outwards, zones of specific rock types occur: a chloride sulphate subtype zone, a chloride-bearing magnesium sulphate subtype zone or sodium sulphate zone, and carbonate type and sodium sulphate subtype or carbonate type zone (Zheng *et al.* 1993). The salt lakes in this study lie in zone V (see Figure 7), which is the zone of sodium sulphate subtype and carbonate type. The exceptions reflect the influence of local geological conditions, for example, zone I_2 (Figure 7) contains carbonate type, which reflects recent volcanism and hydrothermal activity (Zheng *et al.* 1993). Therefore, it is expected that the lakes in this study are high in sodium sulphate and carbonate.

Quaternary saline lakes such as these often contain repetitive depositional sequences that involve alternate layers of freshened phases and saline phases. Core samples taken from the middle of saline Lake Qaganlimennuo'er showed alternate layers sand, clay and mirabilite grains at the surface, then alternate layers of mud and trona between 7 m and 20 m deep. Similarly, core samples of Lake Bayannuo'er showed alternate layers of mirabilite, mirabilite grains and mud (Zheng *et al.* 1993). Both these lakes are located north of the salt lakes in this study. A core sample taken from the middle of Lake Chagannor (approximately 23 m deep) showed alternate layers of silt, silty clay, mud and trona. The layers of trona were correlated with a shallow lake, while detrital deposition was correlated with a deeper lake (Yu *et al.* 2001). In addition, the type of chemical layers is often a reflection of the climate at that time. It has been demonstrated that different salts precipitate at different temperatures; for example, below 20°C, the dominant precipitates are halite, natron ($Na_2CO_3.10H_2O$) and mirabilite in alkali lakes of the $NaCl$-Na_2SO_4-$NaHCO_3$-Na_2CO_3 brine system, but at more than 20 – 30°C, thenardite is present in addition to trona, nahcolite and thermonatrite (Zheng *et al.* 2000). Hence, precipitates have been classed into cold, warm and eurythermal phase saline mineral assemblages according to the climatic conditions for their formation. Cold phase occur in subarid or arid climatic zone of cold or temperate zones, and include natron, mirabilite and borax. Warm phase are formed in subarid or arid areas of subtropical to tropical zones, and include half-water gypsum and fluorite. Eurythermal phase occur in warm temperature to moderate temperate zones, and include calcite, dolomite, magnesite, trona, gypsum, halite and sylvite. For a full list of minerals, see (Zheng *et al.* 2000).

Table 6 shows a compilation the major ions in some Inner Mongolian salt lakes. It was noted in one report that Lake Erliannor had 300 mg/L of Br and that Lake Chagannor was a trona lake (Zheng *et al.* 1993). However, such salt lakes are subject to large fluctuations in salinity due to climatic instability. A further report found that Lake Chagannor has an area of 21 km^2, consisting of a smaller north eastern sub-basin (East Chagannor) and a larger south western sub-basin (West Chagannor), which were joined by a narrow channel. The lake is fed by two intermittent rivers in the north east and south east, though it is also fed by precipitation. It is the largest of 16 salt lakes that are all remnants of a formerly large extensive lake (Palaeolake Chagannor). This palaeolake is believed to have occupied an area of ca 2640 km^2 (Yu *et al.* 2001).

Salt lakes are the product of an arid climate (Zheng *et al.* 2000); therefore as expected, the climate on the Inner Mongolian Plateau is generally a dry continental one. Solar radiation is often high. For example, near Dalai Hu (north of Inner Mongolia), the lowest winter air temperature is -32.4°C and the highest summer temperature is 32.6°C; annual precipitation is 303 mm and annual evotranspiration is 1.29 m (Williams 1991). The general arid climate means that dust storms are frequent. The grasslands of Xilin Gol (near Xilin Hot) are often hit with sand storms that originate in Central Western Inner Mongolia (Tao *et al.* 2003). It is also known that dust from the Gobi Desert can be blown into Beijing.

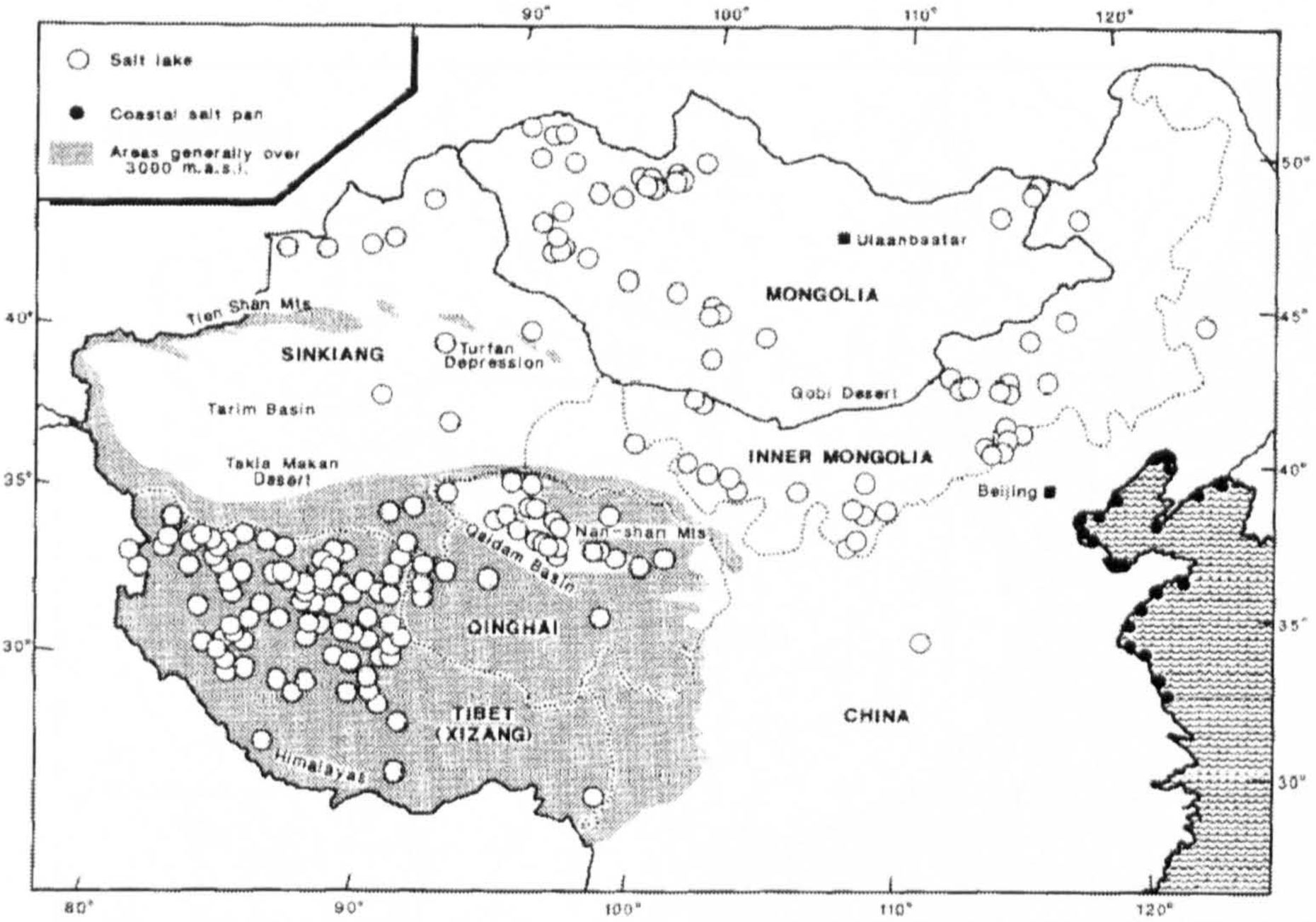


Figure 5 Distribution of Saline Lakes in China and Mongolia.

Large salt lakes are shown as open circles and coastal salt pans are also shown as black circles. m.a.s.l. is metres above sea level. Modified from Williams (1991).

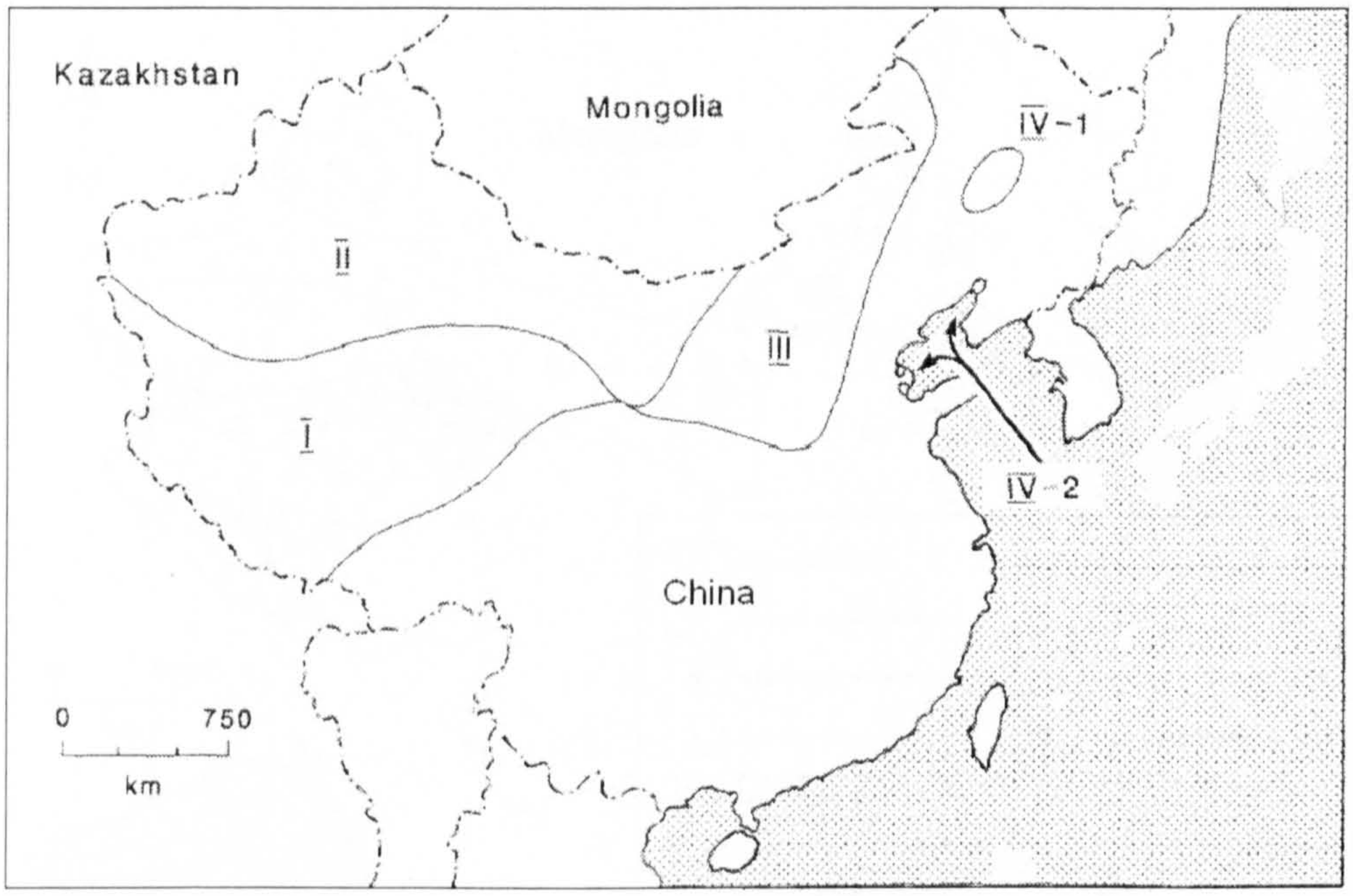


Figure 6 Four Regions of Saline Lakes in China.

The salt lakes in China can be split into four regions, which reflect the geology, climate and other factors. The salt lakes in this study are located in region III in this diagram. Modified from Zheng *et al.* (1993).

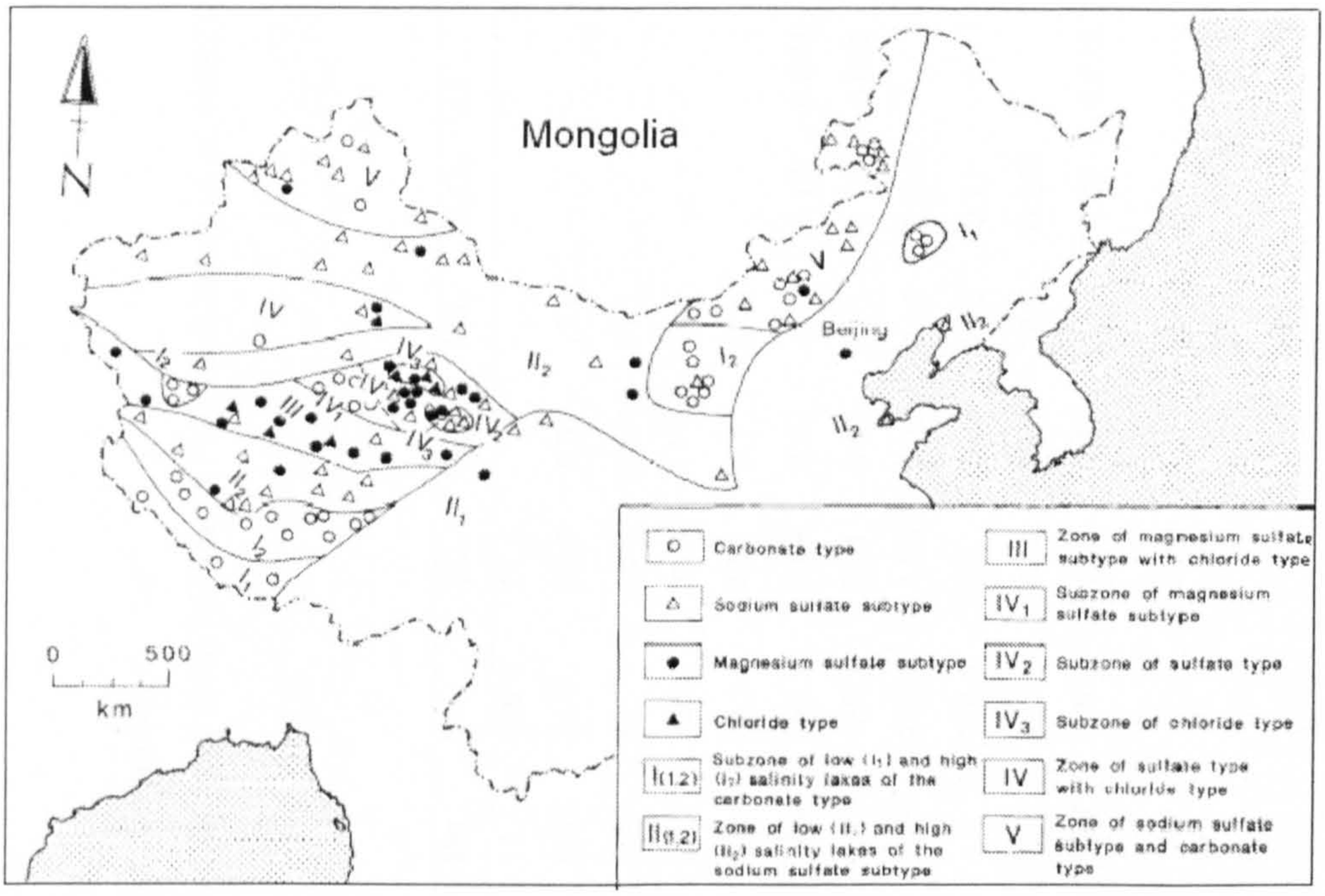


Figure 7 Distribution of Rock Subtypes in China.

The salt lakes in this study are in zone V in this diagram, which is of sodium sulphate subtype and carbonate type. Modified from Zheng *et al.* (1993).

Table 6 Chemical Composition of Salt Lakes in Inner Mongolia.

The most recent readings available are recorded below. Compiled from Williams (1991) and Zheng *et al.* (1992). '-' in the table denotes data that are not determined.

Lake	Salinity (g/L)	pH	Major ions (g/L)											Reference
			Na	K	Ca	Mg	Cl	SO_4	HCO_3	CO_3	B_2O_3	Li	PO_4	
Jartal	310.0	-	111.0	-	0.65	11.1	189.8	17.4	-	-	-	-	-	(Zheng *et al.* 1993)
Dalatulu	112.7	-	42.0	0.45	-	0.03	26.5	25.7	5.8	12.9	-	-	-	(Williams 1991)
Hadat	222.3	-	85.3	3.21	-	0.13	49.7	19.5	14.4	51.1	-	-	-	(Williams 1991)
Wudu	331.3	-	127.9	3.48	-	0.13	78.0	36.0	17.6	70.1	-	-	-	(Williams 1991)
Aledante	457.1	-	142.1	-	-	0.01	73.3	47.7	3.1	87.6	-	-	-	(Williams 1991)
Ejinnor	333.4	7.5	99.9	1.22	0.2	16.9	183.3	30.9	0.42	-	0.12	0.004	0.0008	(Zheng *et al.* 1992)
Bagaejinnor	294.7	7.7	97.8	0.77	0.17	6.81	128.1	58.1	0.59	-	0.12	0.002	0.004	(Zheng *et al.* 1992)
Shangmatala	17.29	10.2	5.2	0.49	0.02	0.25	2.52	4.26	5.22	0.14	0.039	-	-	(Zheng *et al.* 1992)
Chahannor	36.12	10.2	36.1	3.26	-	-	119.9	6.0	22.8	14.0	0.11	-	-	(Zheng *et al.* 1992)
Chagannor	173.2	9.98	67.6	0.49	-	86.3	34.2	22.0	10.9	38.0	0.225	-	0.12	(Zheng *et al.* 1992)
Erliannor	250.2	7.34	114.8	0.75	0.12	12.7	176.9	47.3	0.45	-	0.23	0.015	0.0016	(Zheng *et al.* 1992)

Definition of Halophilic Microorganisms

The term halophilic generally refers to organisms that require salt for growth, which may help maintain structural integrity. Extreme halophiles grow in salt concentrations typically greater than 1.0 – 1.5 M (Grant 2004). Moderate halophiles may be less clearly defined, but generally can be categorised as slight, moderate and borderline extreme. Halotolerant refers to organisms that can grow in both the absence and presence of high salt concentrations, but their growth rate optima are in the absence of salt. Some organisms grow in salinities from zero to saturation, but their growth rate optima is in the presence of salt, and can be termed haloversatile (Grant *et al.* 1998). The categories of halophilic organisms are shown in Table 7. However, more detailed discussions on the use of these terms are found in (Kushner 1978) and (Vreeland 1987).

Phylogenetic Diversity of Halophiles

Eukarya

Few *Eukarya* are able to live in saline environments. The upper limit for the tropical fish, *Tilapia* sp. is about 10%. Only invertebrates such as brine shrimp (*Artemia* sp.), brine flies (*Ephydra*) and eukaryotic algae (*Dunaliella*) can survive salt concentrations that are much higher (Ollivier *et al.* 1994).

Archaea

Almost all archaeal halophiles (haloarchaea) belong to the order *Halobacteriales*, family *Halobacteriaceae*. This order contains 23 validly published genera. Monophyletic groups are *Haloarcula*, *Halobioforma*, *Halococcus*, *Haloferax*, *Halorubrum*, *Halosimplex*, *Natrialba*, *Natrinema*, *Natronococcus* and *Natronorubrum*. The genera represented by a single strain are *Halobacterium*, *Halobaculum*, *Halogeometricum*, *Halomicrobium*, *Halorhabdus*, *Halosimplex*, *Natronobacterium* and *Natronomonas*. The genus *Haloterrigena* is paraphyletic (Wright 2006). Recently described genera include *Halalkalicoccus*, *Halostagnicola*, *Halovivax*, *Haloquadratum* and *Natronolimnobius* (Itoh *et al.* 2005; Xue *et al.* 2005; Castillo *et al.* 2006b; Castillo *et al.* 2006a; Burns *et al.* 2007). Of these genera *Halorubrum*, *Natrialba*, *Natronobacterium*, *Natronococcus*, *Natronomonas*, *Natronorubrum*, *Natronolimnobius* and *Halalkalicoccus* contain alkaliphilic members (Grant *et al.* 2001; Itoh *et al.* 2005; Xue *et al.* 2005). The *Halobacteriales* tend to make up the main component of the biomass in hypersaline environments (Oren 2002a).

Table 7 Definitions of Halophilic Organisms.

Taken from Grant *et al.* (1998).

Category	Salt Concentration (M)	
	Range	Optimum
Nonhalophile	0 – 1.0	< 0.2
Slight halophile	0.2 – 2.0	0.2 – 0.5
Moderate halophile	0.4 – 3.5	0.5 – 2.0
Borderline extreme halophile	1.4 – 4.0	2.0 – 3.0
Extreme halophile	2.0 – 5.2	> 3.0
Halotolerant	0 - > 1.0	< 0.2
Haloversatile	0 - > 3.0	0.2 – 0.5

They are aerobic or facultatively anaerobic and require NaCl for growth. Haloarchaea can grow in NaCl concentrations as high as 5.2 M (Grant *et al.* 2001), though some can grow in as little as 0.43 M NaCl [2.5% (w/v) NaCl] (Purdy *et al.* 2004), while others have been detected in environments where water salinity was between 0.7 – 1% (Elshahed *et al.* 2004). They are capable of growth at temperatures between 35°C and 50°C (Grant *et al.* 2001). Salt lakes often appear red due to the C-50 carotenoid pigments (bacterioruberins) found in the membranes of most of the members of this family. They also have bacteriorhodopsin, which is a light energy driven proton pump. Haloarchaea are chemoorganotrophs that use amino acids or carbohydrates as carbon sources (Grant *et al.* 2001). *Halobacterium salinarium*, *Haloferax mediterranei* and *Halorubrum vacuolatum* are gas-vacuolated, which reduces their net specific gravity enabling them to rise towards the water-air interface where conditions are less anaerobic. Such haloarchaea are most often observed in shallow pools (Oren 1994). In addition, some are capable of light driven anaerobic growth, though it is most likely that aerobic chemoorganotrophic mode of nutrition is used (Oren 1994).

Few *Euryarchaeota* of the methanogenic branch appear to contain halophilic members, and methanogenesis is known to occur at salt concentrations approaching saturation (Oren 2002a). Such methanogens include *Methanocalculus halotolerans*, *Methanohalobium evestigatum*, *Methanohalophilus mahii*, *Methanohalophilus halophilus*, *Methanohalophilus portucalensis*, *Methanosalsum zhilinae* and *Methanococcus doii* (Grant 2004).

Bacteria

The bacterial diversity in salt lakes is rather more complicated. Several branches of the domain *Bacteria* have halophilic and halotolerant members, which include cyanobacteria, *Firmicutes*, *Actinobacteria*, *Proteobacteria*, *Bacteroidetes* and *Spirochaeta* (Grant 2004).

Generally speaking, cyanobacteria are photosynthetic bacteria, responsible for primary production and sometimes nitrogen fixation. The *Firmicutes* are responsible for the degradation of organic compounds and include the anaerobic *Halanaerobiales* and *Clostridium* sp. and aerobic *Bacillus* sp. The *Proteobacteria* are well represented in salt lakes. The *Gammaproteobacteria* includes aerobic organotrophs such as *Halomonas*, *Chromohalobacter*, *Alteromonas*, *Pseudomonas* and *Vibrio*, while anoxygenic phototrophs of the *Ectothiorhodospiraceae* are present. Sulphate reducers of the *Deltaproteobacteria* division are also found in saline environments. *Bacteroidetes* are thought to degrade complex

organic substrates, and include *Salinibacter ruber* (see below). Few representatives of the *Spirochaeta* have been found in soda lakes.

Ecology of Salt lakes

Species diversity tends to decrease as salinity increases. Brine shrimp (*Artemia* sp.) and the larvae of brine flies are abundant in salinities between 10% and 30% and are therefore the only macroscopic organisms in salt lakes. *Artemia salina* tends to occur in neutral brines (Grant 2004) and graze on the alga *Dunaliella* sp and so can reach high populations, which are fed on by birds or are harvested for commercial purposes.

Dunaliella can occur at high salt concentrations (up to 1.5 M NaCl) (Grant 2004), though they are halotolerant rather than halophilic (Oren 2002a). *Dunaliella salina* is a phototrophic alga known for its β-carotene pigments, which contributes to the red coloration of salt lakes. This alga is probably responsible for all, if not most of the primary production in salt lakes (Oren 2002a). Several cyanobacteria have been found in salt lakes that contribute to primary production. *Aphanothece halophytica* and *Dactylococcopsis salina* are unicellular gas-vacuolated cyanobacteria that were originally isolated from Solar Lake in Israel (Cohen *et al.* 1977; Walsby *et al.* 1983). However, large populations of cyanobacteria are not often seen in neutral salt lakes as they are for soda lakes (see later). Anoxygenic phototrophic bacteria may also contribute to primary production such as *Ectothiorhodospira marismortui*, isolated from the Dead Sea, and *Halorhodospira halophila* (Grant 2004).

A vast range of heterotrophic organotrophs are supported by primary productivity and deposition of organic matter as they utilise various polymers and monomers.

The neutrophilic archaea are prevalent in salt lakes, hence many lakes appear red due their carotenoid pigments. They are aerobic heterotrophs that derive energy from the organic substrates formed by primary producers. Large amounts of glycerol is released from *Dunaliella* cells when they die, which is a suitable carbon source for almost all of the haloarchaea (Oren 1994). Furthermore, haloarchaea have been shown to degrade aromatic hydrocarbons, though it is not known to what extent they utilise these substrates in their habitat (Oren 1994). *Halorubrum* and the SHOW group (see later) are particularly dominant in near neutral salt lakes (Grant 2004). In contrast, it was found that the main constituent of the archaeal community in the Dead Sea (where waters are slightly acidic) was *Haloferax*,

suggested by the presence of significant amounts of their glycolipids (Oren and Gurevich 1993).

Methanogenic archaea may also be present in some salt lakes, depending on salinity. They are able to utilise methanol or methylamine as substrates for the production of methane. Examples include *Methanohalophilus halophilus* that was isolated from a salt lake in Australia, and *Methanohalophilus* sp. that was isolated from the Great Salt Lake in Utah (Grant 2004).

Methane is consumed by methane oxidising bacteria. Evidence for methane consumption was detected at hypersaline environments in the Crimea basins. Environments exhibiting 8% - 33% (w/v) total salinity showed methane consumption at 1200 ρmol cm^{-3} d^{-1}. Subsequently, several enrichment cultures of moderately halophilic methanotrophs were obtained (Sokolov and Trotsenko 1995). Recently, *Methylohalobius crimeensis* was isolated from hypersaline lakes in the Crimean Peninsula. This bacterium can tolerate NaCl concentrations from 0.2 M up to 2.5 M (1.2 – 15%), but grows optimally at 1 – 1.5 M [5.8 – 8.7%, (w/v)]; this optimum and upper limit is the highest for any methanotrophic bacterium known to date. It grows on methane or methanol and utilises the ribulose monophosphate pathway for carbon assimilation (Heyer *et al.* 2005).

Many moderate halophiles have been assigned to the genus *Halomonas* of the *Gammaproteobacteria*. They are Gram negative bacteria that grow optimally between 0.5 M and 2.5 M salt (Ventosa *et al.* 1998a). Vast numbers of halomonads have been isolated and characterised from salt lakes. Examples include *Halomonas halmophila* and *Chromohalobacter marismortui*, both isolated from the Dead Sea, and *Halomonas variabilis* isolated from the Great Salt Lake (Grant 2004). Halomonads are able to utilise a range of sugars and amino acids as carbon sources (Mata *et al.* 2002). However, *Halomonas organivorans* is unusual in that it can utilise a variety of aromatic organic compounds (García *et al.* 2004). Additional members from the *Gammaproteobacteria* found in salt lakes include halovibrios (*Halovibrio denitrificans* isolated from a hypersaline lake in north east Mongolia), alteromonads (*Marinobacter* sp.) and '*Pseudomonas halophila*' (isolated from the Great Salt Lake)(Grant 2004; Sorokin *et al.* 2006). Another important and especially abundant group of aerobic heterotrophic bacteria are the *Bacillus* of the *Firmicutes* (Ventosa *et al.* 1998a). Examples include *Halobacillus litoralis* and *Halobacillus trueperi*, which were isolated from

the Great Salt Lake (Spring *et al.* 1996). *Halobacillus dabanensis*, *Halobacillus aidingensis* and *Lentibacillus lacisalsi* were isolated from salt lakes in China (Lim *et al.* 2005a; Liu *et al.* 2005).

Fermentative anaerobic bacteria degrade a range of organic substrates. Many belong to the *Haloanaerobiales* and *Halobacteroidaceae* of the *Firmicutes*. They are known to thrive in salt concentrations of 25% (w/v) or higher (Oren 2002a). *Halanaerobium lacusrosei* was isolated from the sediments of Retba Lake near Dakar in Senegal. It was shown that yeast extract was required for growth and the fermentation products from glucose were ethanol, acetate, H_2, and CO_2, though it could utilise a range of other sugars as substrates (Cayol *et al.* 1995). *Haloanaerobium praevalens*, isolated from the Great Salt Lake, ferments pectin to form acetate, proprionate, butyrate, CO_2 and H_2. *Halanaerobacter lacunarum*, isolated from lagoonic hypersaline Lake Chokrak, ferments starch to form ethanol, acetate CO_2 and H_2 (Zeikus *et al.* 1983; Zhilina *et al.* 1991; Ollivier *et al.* 1994). *Halobacteroides halobius* isolated from the Dead Sea ferments a range of substrates including pyruvate (Oren *et al.* 1984). Clostridia are also found in hypersaline environments. *Clostridium halophilium* utilises several carbohydrates; it also uses betaine with hydrogen and several amino acids as electron donors (Fendrich *et al.* 1990).

Homoacetogenic bacteria have been isolated from salt lakes; they generate acetate from various organic compounds. *Halanaerobium saccharolyticum*, isolated from Lake Sivash (Crimea), prefers disaccharides such as sucrose to produce acetate, H_2 and CO_2 (Zhilina *et al.* 1992). Others use products from anaerobic digestion, for example, *Acetohalobium arabaticum* grows on betaine and trimethylamine and can reduce CO_2 to acetate (Ollivier *et al.* 1994).

To my knowledge, few sulphur oxidising bacteria (SOB) have been reported from salt lakes. (More have been reported from soda lakes; see below). Anoxygenic phototrophic bacteria may contribute to sulphur oxidation, for example, *Ectothiorhodospira marismortui* can grow either photoautotrophically with sulphide as electron donor, which is oxidised via extracellular sulphur to sulphate (Oren *et al.* 1989). *Halothiobacillus halophilus*, isolated from a hypersaline lake in Western Australia, grows chemolithoautotrophically on thiosulphate, tetrathionate and sulphur, oxidising them to sulphate, hence also contributing to fixation of carbon dioxide (Wood and Kelly 1991; Kelly and Wood 2000).

Sulphate reducing bacteria (SRB) utilise sulphate as the main electron acceptor during anaerobic metabolism, though thiosulphate, sulphite and sulphur may also be used. The final product is hydrogen sulphide (Ollivier *et al.* 1994). Examples include *Desulfovibrio halophilus* isolated from Solar Lake (Sinai), *Desulfohalobium retbaense* isolated from sediments of a hypersaline lake in Senegal and *Desulfocella halophila* isolated from surface sediments of the Great Salt Lake in Utah (Caumette *et al.* 1991; Ollivier *et al.* 1991; Brandt *et al.* 1999).

Ecology of Salterns

There is limited microbial diversity in ponds where salinity is approaching saturation. In a study of the Bras del Port salterns in Santa Pola, Spain, the community of the most concentrated pond was found to consist mostly of archaea, with a smaller proportion of bacteria. Although the abundance of heterotrophic prokaryotes was high, all the cell activity was found to be due to the archaea; activity was generally low in the bacteria (Pedrós-Alió *et al.* 2000). Moreover, if chlorophyll a was present, the algae responsible (*Dunaliella*) was in a non-functional state (Pedrós-Alió *et al.* 2000).

This is due to the fact that moderate halophilic bacteria are outcompeted by haloarchaea that require high salinities for the maintenance of their cell components (Grant *et al.* 2001). Indeed, the abundance of haloarchaea is what imparts the red coloration observed at many salterns (Oren and Dubinsky 1994). It is known that the reddening of salterns promotes salt precipitation since the carotenoids in the haloarchaea trap solar radiation, increasing temperature and therefore rates of evaporation (Grant *et al.* 2001). Salterns are typically around neutral pH, hence the haloarchaea found in salterns are those of the genera with neutrohalophilic members within the *Halobacteriales*, particularly *Halorubrum* and *Haloarcula* as found, for example, in the Sečovlji solar saltern in Slovenia and in the solar saltern in La Palma, Spain (Ochsenreiter *et al.* 2002; Pašić *et al.* 2005). Examples of haloarchaea in salterns are *Haloferax mediterranei*, *Haloferax gibbonsii*, *Haloferax denitrificans*, *Halogeometricum borinquense*, *Halococcus saccharolyticus*, *Haloterrigena thermotolerans*, *Halorubrum saccharovorum*, *Halorubrum coriense*, *Haloarcula hispanica* and *Haloarcula japonica* (Oren 2002b). Methanogens are not detected in such high salinities.

The dominant fraction of the haloarchaeal community found in crystalliser ponds where salinities are between 23% and 30% are the Square Haloarchaea of Walsby (SHOW). They

were first identified by Walsby in 1980 using microscopy from a natural salt pond on the Gulf of Eilat near the Rea Sea (Walsby 1980). They have since been described as the dominant fraction in many crystalliser ponds including the Cheetham Salt Works in Victoria, Australia (Burns *et al.* 2004a), the Bras del Port salterns in Santa Pola, Spain (Benlloch *et al.* 2002) and the Maras salterns in the Peruvian Andes, Peru (Maturrano *et al.* 2006). For years, this square haloarchaeon could not be cultivated despite being the most abundant prokaryote in such environments. In 2004, they were grown in co-culture in media containing low concentrations of yeast extract and with at least 2 M $MgCl_2$ and 3.3 M NaCl (Bolhuis *et al.* 2004). Shortly after, a pure culture was achieved using extinction culture techniques (Burns *et al.* 2004b). They also recognised that high salt, low nutrient media was required for the cultivation of SHOW to prevent faster growing haloarchaea from outcompeting them. The cells have a square morphology, resembling postage stamps, with gas vesicles and contain poly-β-hydroxybutyrate (PHB) granules (Burns *et al.* 2004b). They are now represented by *Haloquadratum walsbyi* (Burns *et al.* 2007).

A proportion of the prokaryotic community in salterns can be attributed to bacteria. Work by Anton *et al* (2000) first found large numbers of an extremely halophilic bacterium in the Bras del Port salterns in Alicante, Spain, exhibiting salinities between 22.4 - 37% (w/v) using fluoresence *in situ* hybridisation (FISH), which constituted from 5% to 25% of the total prokaryotic community (Antón *et al.* 2000). This study was supported by denaturing gradient gel electrophoresis (DGGE) analysis (Benlloch *et al.* 2002). This bacterium was isolated and characterised and assigned as *Salinibacter ruber* (Anton *et al.* 2002), and is affiliated with the *Bacteroidetes* phylum. *Salinibacter* is extremely halophilic requiring at least 15% NaCl, but grows optimally at 20 – 30% (w/v) (Anton *et al.* 2002). Unlike other halophilic bacteria, *Salinibacter ruber* is able to co-exist with haloarchaea, often occurring at the point of halite precipitation. Therefore, *Salinibacter* are among the most halophilic organisms within the domain *Bacteria* (Anton *et al.* 2002). Like members of the haloarchaea, *Salinibacter ruber* is an aerobic heterotroph that exploits organic nutrients from primary producers e.g. *Dunaliella*, since it has been found that it utilises glycerol early on in its growth phases (Mongodin *et al.* 2005).

Crystalliser ponds contain a predominantly prokaryotic community (Pedrós-Alió *et al.* 2000) and the lack of predation often allows microbial mats to develop. The microbial mats in the hypersaline ponds of Salins-de-Giraud in Camargue, France are fully developed in the warm

seasons in the gypsum crusts. This mat was found to be composed layers of phototrophic organisms. The top brown layer consisted of the cyanobacterium *Aphanothece*, the middle green layer contained the cyanobacterium *Phormidium* and the bottom red layer had purple sulphur oxidising bacteria, such as *Chromatium salexigens* and *Thiocapsa halophila* (Caumette *et al.* 1994). The anoxic sediments directly below these mats contained many bacteria, including *Firmicutes* (*Orenia salinaria*), *Cytophage-Flavobacterium-Bacteriodetes* group and *Alphaproteobacteria* and *Deltaproteobacteria* (Mouné *et al.* 2000; Mouné *et al.* 2003). A similar microbial mat was found in the salterns of the Israel Salt Industries in Eilat, which also developed in gypsum. *Gammaproteobacteria* (related to *Thioalkalivibrio*) dominated the white layer. The intermediate green layer was composed of cyanobacteria of the *Halothece*, *Spirulina* and *Phomidium* types. Anaerobic fermentative bacteria of the *Firmicutes* (*Haloanaerobiales*), spirochaetes (relating to *Spirochaeta halophila*) and *Planctomycetes* were detected in the deep purple/ olive green layer. *Alphaproteobacteria* also dominated the purple layer, affiliated with the *Rhodospiriales*, *Rhodobacteriales* or the *Sphingomonadales*. Archaea seemed to play a minor role in the microbial mat (Sørensen *et al.* 2005). A microbial mat of the Exportadora de Sal SA, a solar salt works located at Guerrero Negro in Baja California Sur, Mexico showed surprising complexity in its bacterial community. However, the biomass was dominated by *Chloroflexi* in all layers of the mat (Ley *et al.* 2006).

Ecology of Soda Lakes

The microorganisms that are found in soda lakes not only have to be able to cope in a high salt environment, but they also have to be adapted to extremely high pH. Hence the microbiota of soda lakes is significantly different to those of salt lakes and salterns as they have to cope to two extremes. Well documented soda lakes are those of the African Rift Valley in Kenya (such as Lake Magadi, Lake Natron and Lake Bogoria) and those of North America (such as Mono Lake and Searles Lake). Much of the recent isolation and characterisation of bacteria from soda lakes has been carried out by Sorokin and Kuenen, which has allowed the fates of reduced inorganic compounds (such as methane, hydrogen, sulphide and ammonia) to be elucidated.

Few eukarya have been observed in soda lakes. *Dunaliella salina* is unable to grow at high pH, though the brine shrimp *Artemia monica* has been observed at Mono Lake (Humayoun *et al.* 2003).

The alkaliphilic haloarchaea, as previously mentioned, are members of the genera *Halorubrum*, *Natrialba*, *Natronobacterium*, *Natronococcus*, *Natronomonas*, *Natronorubrum* (Grant *et al.* 2001), *Halalkalicoccus* (Xue *et al.* 2005) and *Natronolimnobius* (Itoh *et al.* 2005). Many alkaliphilic haloarchaea have been isolated from the soda lakes of the Kenyan Rift Valley, including *Halorubrum vacuolatum* (Mwatha and Grant 1993; Kamekura *et al.* 1997), *Natrialba magadii* (Tindall *et al.* 1984), *Natronobacterium gregoryi*, *Natronococcus occultus* (Tindall *et al.* 1984; McGenity and Grant 1993) and *Natronococcus amylolyticus* (Kanai *et al.* 1995).

Methanogenic archaea have also been identified in soda lakes. *Methanohalophilus zhilinae* is one example, which was isolated from Bosa Lake of the Wadi el Natrun in Egypt (Mathrani *et al.* 1988). Other examples include *Methanosalsus zhilinaeae* and *Methanolobus oregonense* (Ollivier *et al.* 1994).

Two strains of methanotrophic bacteria *Methylobacter alcaliphilus* were isolated from soda lakes in Tuva in Central Asia. They grew fastest at pH 9 – 9.5 and required $NaHCO_3$ or NaCl (Khmelenina *et al.* 1997; Horikoshi 1999). Other methane oxidising bacteria were isolated from sediments of south eastern Transbaikal soda lakes. They were aerobic bacteria that were alkalitolerant or facultatively alkaliphilic, able to grow at pH 10.5 - 11.0 (optimally at pH 8.5 - 9.5). It was found that they assimilated methane and methanol, and possessed methane monooxygenase. Subsequently, they were designated as *Methylomicrobium buryatense* (Kaluzhnaya *et al.* 2001).

In an environment where carbonate is in excess, CO_2 is freely available. In combination with high solar radiation in tropical and sub-tropical regions, cyanobacterial blooms are often observed in soda lakes, which are usually predominated by filamentous species *Spirulina platensis*, though *Cyanospira*, *Chroococcus*, *Synechococcus* and *Synechocystis* are occasionally the principal species in some lakes (Jones *et al.* 1994; Grant and Jones 2000). These cyanobacterial blooms impart a green hue and are often observed at the alkaline soda lakes in Kenya, which supports vast populations of the lesser flamingo (*Phoeniconaias minor*) (Grant and Jones 2000). *Spirulina* and *Cyanospira* have been found in African soda lakes, while the filamentous green alga *Ctenocladus aircinnatus* and *Oscillatoria* have been found to fix nitrogen at Mono Lake (Oremland 1990; Grant and Jones 2000). Blooms of anoxygenic

phototrophs can also produce red blooms, which are often formed by bacteria of the genus *Halorhodospira* and *Ectothiorhodospira* (Grant and Jones 2000).

Aerobic organotrophs are supported by the primary productivity as they utilise various polymers and monomers. It is thought that daytime oxygen production by *Spirulina* sp. (that rise to the surface of the water column by gas vacuoles) is mixed in the water column by strong winds, which allows growth of such organisms (Jones *et al.* 1994). Aerobic organotrophs include various alkaliphilic halomonads. For example, *Halomonas campisalis* grows at pH 8 – 11 and is able to grow on various sugars, and can reduce both nitrate and nitrite (Mata *et al.* 2002). *Halomonas magadii*, isolated from Lake Magadi grows at pH 9 to 10, but is unable to reduce nitrite (Duckworth *et al.* 2000).

Other commonly isolated types include *Bacillus* sp., which probably play an important role in the breakdown of polymers since they have been shown to produce many extracellular enzymes including proteases, amylases, cellulases, lipases, xylanases, pectinases and chitinases, which have been exploited for biotechnological purposes (Horikoshi 1999). In 1977 Horikoshi isolated an alkaliphilic *Bacillus* strain C-125 that produced β-galactosidase and xylanase (Ikura and Horikoshi 1979; Honda *et al.* 1985). It has since been the most characterised strain physiologically, biochemically and genetically and has been re-identified as *Bacillus halodurans* (Takami *et al.* 2000). Several *Bacillus* strains were isolated from Lake Magadi and Lake Natron; strains related to *Bacillus alcalophilus* were more prevalent in shoreline mud and dry foreshore soda soils, while strains related to *Bacillus agaradherens* and *Bacillus clarkii* were more prevalent in lake waters and sediments (Jones *et al.* 1998b). Recently, a bacterium from Lake Bogoria was isolated, designated *Bacillus bogoriensis*, and found to grow optimally at pH 10 and utilised a number of sugars as carbon sources (Vargas *et al.* 2005). Other bacteria related to *Aeromonas*, *Vibrio* and *Pseudomonas* have been isolated from soda lakes (Duckworth *et al.* 1996).

Fermentative anaerobic bacteria have been isolated from various soda lakes and have been shown to utilise a range of substrates to produce volatile fatty acids (VFAs). *Alkalibacter saccharofermentans* and *Alkaliflexus imshenetskii* are both examples of saccharolytic bacteria. *Alkalibacter saccharofermentans* was isolated from alkaline Lake Nizhnee Beloe (Transbaikal region, Russia), and is a halotolerant bacterium that ferments ribose, xylose, glucose, mannose, fructose, sucrose, mannitol and peptone, though the products of glucose

fermentation are acetate, ethanol, formate, hydrogen and CO_2 (Garnova *et al.* 2004). *Alkaliflexus imshenetskii* was isolated from alkaline Lake Verkhneye Beloye in Central Asia. It does not degrade cellulose, but utilises xylose, maltose, xylan, starch and pectin to produce propionate, acetate and succinate (Zhilina *et al.* 2004). *Spirochaeta alkalica* and *Spirochaeta africana*, isolated from Lake Magadi and *Spirochaeta asiatica* isolated from Lake Khatyn, Central Asia were shown to ferment glucose to acetate, lactate, ethanol and hydrogen (Zhilina *et al.* 1996b). *Tindallia magadii*, isolated from soda deposits in Lake Magadi, is an ammonifying bacterium that utilises amino acids, preferentially arginine and ornithine and produces acetate, propionate and ammonia (Kevbrin *et al.* 1998). Moreover, various anaerobic haloalkaliphilic strains isolated from Lake Magadi were shown to ferment sugars and amino acids to isovaleric acid as an end product, with smaller amounts of isobutyric acid and acetic acid. Less halophilic strains from Lakes Elmenteita and Bogoria were shown to produce acetate and proprionate or butyrate as fermentation products, and were clustered with *Clostridium* (Jones *et al.* 1998b). Recently, one anaerobic alkaliphilic *Clostridium* was isolated from the Verkhnee Beloe soda lake in Buryatiya, Russia and was designated *Clostridium alkalicellum*. This bacterium degraded cellulose, cellobiose and xylan, and produced lactate, ethanol, acetate, hydrogen, and traces of formate during cellulose or cellobiose fermentation (Zhilina *et al.* 2005).

Acetogenic bacteria that consume products from anaerobic digestion to produce acetate have been identified from soda lakes. *Natroniella acetigena* grows optimally at pH 9.7 to 10.0 and has a requirement for sodium carbonate and chloride ions. It was isolated from Lake Magadi and has been shown to use lactate, ethanol, pyruvate, glutamate and propanol, with acetate as the main end product (Zhilina *et al.* 1996a). A further isolate from Lake Magadi, designated *Natronoinocola histidinovorans* utilises only two amino acids, histidine and glutamate, with formation of acetate and ammonium as main end products (Zhilina *et al.* 1998). *Tindallia californiensis* isolated from Mono Lake, grows on peptone, bacto-tryptone, casamino acid, yeast extract, l-serine, l-lysine, l-histidine, l-arginine and pyruvate to produce mainly acetate (Pikuta *et al.* 2003a).

SOB are divided into three genera within in the *Gammaproteobacteria*. These are *Thioalkalimicrobium*, *Thioalkalivibrio* and *Thioalkalispira*, which are members of the *Ectothiorhodospiraceae*, which derive energy by using H_2, organic compounds or hydrogen sulphide as electron donors to generate sulphate (Ollivier *et al.* 1994). Generally

Thioalkalimicrobium sp. dominate the low salt soda lakes, while *Thioalkalivibrio* sp. can tolerate saltier conditions. *Thioalkalispira microaerophila* was isolated under microoxic conditions from a denitrifying culture (Sorokin and Kuenen 2005).

Sulphur oxidation can be coupled to cycling of other elements. Enrichment cultures from soda lake sediments using thiosulphate as the electron donor and nitrate as the electron acceptor led to the discovery of *Thioalkalivibrio nitratireducens*. This bacterium could reduce nitrate to nitrite, with copious formation of sulphur globules (Sorokin *et al.* 2003). *Thioalkalivibrio denitrificans* is also a denitrifier that grows well with nitrous oxide (N_2O) and polysulphide, hence these two isolates demonstrate sulphide oxidation to elemental sulphur under denitrifying conditions (Sorokin and Kuenen 2005). Moreover, *Thioalkalivibrio paradoxus* can oxidise thiosulphate and carbon disulphide to intracellular sulphur, while *Thioalkalivibrio thiocyanoxidans* grows with thiocyanate or thiosulphate and nitrate or nitrite as the electron acceptor, representing the first complete denitrifier among the SOB (Sorokin *et al.* 2002; Sorokin and Kuenen 2005). An example of a SOB that can also oxidise hydrogen is strain AHO 1, isolated from Kenya that can oxidize reduced sulphur compounds to sulphate while utilising a range of organic compounds as energy sources and grows mixotrophically with hydrogen and acetate (Sorokin *et al.* 2000).

SRB have been identified from soda lakes. For example, *Desulfonatronum thiodismutans* and *Desulfonatronovibrio hydrogenovorans* (the former isolated from Mono Lake and the latter isolated from Lake Magadi), can both use sulphate, sulphite and thiosulphate as electron acceptors, while utilising hydrogen, formate (and ethanol for *Desulfonatronum thiodismutans*) as electron donors, forming hydrogen sulphide as the end product (Zhilina *et al.* 1997; Pikuta *et al.* 2003b).

Nitrite oxidising bacteria have been isolated from soda lakes that are closely related to *Nitrobacter*, though are not truly alkaliphilic since they grow very slowly at neutral pH; hence they have been classified as facultative alkaliphiles. These strains (AN1 - AN5) were different to neutrophilic *Nitrobacter* as they formed thick globular S-layer and could grow and oxidise nitrite in highly alkaline conditions. They were designated *Nitrobacter alkalicus* (Sorokin *et al.* 1998).

Alkaliphilic ammonia oxidising bacteria have been enriched from sediment samples from north eastern Mongolian soda lakes with a pH optimum around 9. These five strains (ANs1 – Ans5) were genetically related to the known marine species *Nitrosomonas halophila.* In a pH-controlled, ammonia-limited continuous culture, isolate ANs5 grew up to pH 11.3, which is the highest pH limit known for ammonia oxidising bacteria so far (Sorokin *et al.* 2001).

Many soda lakes in north western America contain significant amounts of arsenic (in the forms arsenate and arsenite), which can be utilised by various bacteria. Arsenate is preferentially used over sulphate or carbonate as oxidants in order to yield more energy. A haloalkaliphilic bacterium isolate SLAS-1 was isolated from the sediments of Searles Lake and was found to respire via arsenate, using either lactate or sulphide as its electron donor (Oremland *et al.* 2005). *Bacillus arsenicoselenatis* and *Bacillus selenitireducens* were isolated from Mono Lake and are typical of water column arsenate reducers. They utilise arsenate for the oxidation of lactate (Blum *et al.* 1998). Furthermore, strain MLHE-1 is capable of CO_2 fixation with arsenite and is a member of the *Ectothiorhodospiraceae* (Oremland *et al.* 2002).

Ecology in Inner Mongolian Salt Lakes

Previous reports have shown *Artemia* to be present in lakes that are in zones of sodium sulphate subtype and the magnesium sulphate and carbonate subtype and where salinity is 26 – 341 g/L, which include Lake Ejinnor and Lake Jartal (Zheng *et al.* 1993).

Natronobacterium sp. have been isolated from Lakes Chagannor and Chahannor (Jones *et al.* 1994). For example, haloalkaliphilic archaeal strains C231 and C42 were isolated from Lake Chahannor and were shown reduce nitrate to nitrite and reduce sulphur and thiosulphate to sulphide. These strains were subsequently designated *Natronobacterium nitratireducens* (Xin *et al.* 2001). Two *Natrialba* sp. were also isolated from Inner Mongolia. *Natrialba hulunbeirensis* (isolated from the Hulunbeir prefecture) and *Natrialba chahannaoensis* (isolated from Lake Chahannor) are strictly aerobic haloalkaliphiles that can utilise a range of sugars and amino acids, can reduce nitrate and can produce hydrogen sulphide from cystein (Xu *et al.* 2001). Other archaea isolated from Inner Mongolia include the haloalkaliphiles *Natronolimnobius baerhuensis* and *Natronolimnobius innermongolicus* (Itoh *et al.* 2005) and *Halobacterium jilantaiense* (isolated from Jilantai salt lake) a chemoorganotrophic, aerobic archaeon that requires 2.7 – 5.2 M NaCl and 0.05 – 0.3 M Mg^{2+} with a pH range of 5.5 – 8.5 (Yang *et al.* 2006).

Several bacteria have also been isolated from salt lakes in Inner Mongolia. *Alkalimonas amylolytica* (isolated from Lake Chahannor) and *Marinospirillum alkaliphilum* (isolated from Haoji soda lake) are alkaliphilic, slightly halophilic Gram negative bacteria affiliated with the *Gammaproteobacteria* (Zhang *et al.* 2002a; Ma *et al.* 2004a). *Salinicoccus alkaliphilus* was isolated from Baer Soda Lake. It was the first known Gram positive coccus to be both moderately halophilic and alkaliphilic; the seven previous moderately halophilic Gram positive cocci all grow at neutral pH (Zhang *et al.* 2002b). In addition, molecular analysis on sediment samples at Baer Soda Lake has been done and demonstrated an abundance of *Proteobacteria* (Ma *et al.* 2004b).

Using Molecular Techniques to Characterise Microbial Diversity

Molecular approaches used to characterise the microbial community in the environment involves extraction of community genomic DNA and then amplification of the 16S rRNA gene by Polymerase Chain Reaction (PCR) using universal primers designed against conserved regions near the ends of the gene. This mixture of PCR products is then subject to a range of molecular approaches.

Electrophoretic analysis provides a genetic fingerprint of the whole community by differential migration of the mixed PCR products by electroporation through agarose or polyacrylamide. This can be dependent on size [amplified ribosomal DNA restriction analysis (ARDRA), terminal restriction fragment length polymorphism (t-RFLP), ribosomal intergenic spacer analysis (RISA), random amplified polymorphic DNA (RAPD)] or sequence [denaturing gradient gel electrophoresis (DGGE), temperature gradient gel electrophoresis (TGGE)]. The methodology of these techniques has been reviewed elsewhere (Ranjard *et al.* 2000). However, these techniques are generally only useful in looking at changes in microbial community over a period of time by taking several fingerprint profiles at certain time intervals. In order to characterise the microbial community, bands must still be extracted and sequenced. The limitations of some of these these techniques have been discovered by others (Rees 2002; Free 2005).

Therefore, the method used in this thesis is to construct 16S rRNA gene clone libraries and to sequence clones once inserts have been screened to eliminate duplicate sequences. Although a laborious and time consuming method, it provides almost full length 16S rRNA genes available for analysis, revealing structure and genetic diversity of the community and taxonomic identification of populations.

With the advent of these molecular techniques and the ease of sequencing reactions, microbial communities of many environments have been characterised, including soil (McCraig *et al.* 1999), water (Horner-Devine *et al.* 2003), the human mouth (Kroes *et al.* 1999) and the human gut (Suau *et al.* 1999). It has allowed detection of uncultivated organisms such as the *Korarchaeota* (Barns *et al.* 1994), and has given insights into community structure and function (Eichler *et al.* 2006).

However, there are drawbacks to PCR-based methods. Templates with high % G + C content may be discriminated against during PCR due to the low efficiency of strand separation. The fidelity of DNA polymerases used during PCR may also vary, which results in mis-incorporation of nucleotides and may lead to assumptions of novel taxa present in the sample (Head *et al.* 1998). In addition, studies have shown that a proportion of amplified sequences from a mixed community form chimeras during PCR (Kopczynski *et al.* 1994). They are formed when DNA synthesis begins at one end of a sequence, but is interrupted and continues on another template sharing some degree of localised homology to the original. This results in two or more fragments from different genes joining, which is introduced into the PCR as a full length template to be amplified. It has become increasingly important that chimeras are recognised and removed from diversity studies as novel lineages could be erroneously assigned and give misleading impressions of biodiversity. These sequences could be deposited in public databases reducing the quality of such repositories. Furthermore, primers are not completely universal. In fact, a previous study showed that alternative bacterial primers designed against a different section of the 16S rRNA gene sequence was able to amplify a proportion of the microbial community in sputum samples that was not obtained with the universal 16S rRNA primer, 27Fb (Free 2005). Therefore, the widespread use of these 'universal' 16S rRNA primers can lead to a significant part of the microbial communty to be overlooked.

Differences in operon copy number causes bias towards microbes with higher copy number, thereby overestimating diversity by counting multiple signals from one organism. There is now a database providing information on operon copy number in different organisms (Klappenbach *et al.* 2001). Heterogeneity of 16S rRNA sequences can also overestimate diversity (Crosby and Criddle 2003). For example, *Haloarcula marismortui* has an inter-operon difference up to 5% between two 16S rRNA gene sequences (Mylvaganam and Dennis 1992). Hence, slightly different genes could originate from one strain, but may be interpreted as genes from two closely related organisms.

There are other limitations with molecular analysis that is not related to the problems inherent with PCR. The recovery of nucleic acids from environmental samples may not be complete. Spores are more resistant to cell lysis than vegetative cells, as are Gram positive cells compared to Gran negative cells. In addition, recovery may be reduced by degradation or adsorption of nucleic acids to matrix material (Zhou *et al.* 1996). Furthermore, it has been

noted that small cells (0.3 to 1.2 μm) are more resistant to cell lysis than larger cells and endospores (Moré *et al.* 1994).

The Need for Cultivation Methods

Only 0.1% to 10% of the microbial community in a soil sample is culturable (Head *et al.* 1998). This is a great drawback when trying to characterise the microbial communty. Despite this, cultivation methods are still required to isolate and characterise microorganisms, which provide an insight into evolution and microbial ecology (community structure and function) as well as physiology, metabolic pathways and genetics of particular organisms. For example, the recently cultivating *Haloquadratum walsbyi* was an important step in understanding the microbial ecology in hypersaline environments (Burns *et al.* 2004b). The discovery of *Salinibacter ruber* provided a better understanding of bacterial ecology in hypersaline environments and refuted previous assumptions that bacteria were outcompeted at high salinities (Anton *et al.* 2002). The completed genome of this bacterium provided insights into the coevolution of this bacterium with the haloarchaea (Mongodin *et al.* 2005). In addition, the cultivation of the dominant species in the environment may allow the isolation and characterisation of novel haloviruses also present in the environment (see later).

Therefore, in addition to molecular methodology described above, cultivation methods will be employed in this thesis to cultivate organisms from the salt lakes. Few prokaryotes have been isolated and characterised from the Inner Mongolian salt lakes (see above) and so there is potentially many novel *Bacteria* and *Archaea* that remains to be discovered. Traditional cultivation methods use conditions that are different to the natural environment that they are isolated from, but this may contribute to failure to cultivate most microorganisms (Keller and Zengler 2004). Successful approaches are those that mimic the natural environment. For example, novel halophiles were isolated from the Red Sea by using media close to the composition of the Red Sea brine (Eder *et al.* 2001). A low nutrient approach was also successful for the cultivation of *Haloquadratum walsbyi* (Burns *et al.* 2004b). Direct plating with extended incubation periods has also been shown to be successful in cultivating the dominant microbial groups in soil (Sait *et al.* 2002).

Molecular Microbial Ecology

Definition of Operational Taxonomic Unit (OTU)

Generally, an Operational Taxonomic Unit (OTU) is a group of organisms used in a taxonomic study, and can be species or individuals. In the context of this study, an OTU is a sequence of DNA that is known in the organisms, but is sufficiently variable to allow discrimination between these organisms, i.e. the 16S rRNA gene. Sequence identity in an OTU means that they are from the same taxon or that there is insufficient variation to define distinct taxa.

Diversity Indices

Statistics have long been used in general ecology to describe the biodiversity of plants and animals. Such analyses were not possible in the past, but since new molecular technologies and rapid DNA sequencing methods have emerged, these have been used by microbiologists to measure the biodiversity of microorganisms in the environment. They measure the species richness, which is the total number of different species in a given area, and species evenness, which is the distribution of the richness among the species.

The Simpson's Index accounts for both the total species richness and the proportion of this richness assigned to each species (Simpson 1949). It measures the probability that two individuals selected at random from a sample belong to the same species. However, it has been interpreted as the weighted mean of the proportion of abundances (Peet 1974). The value of the Simpson's Index decreases as species evenness increases.

The Shannon-Weaver Index similarly takes into account the total number of species present in a given environment and the richness or abundance of each of those species (Shannon and Weaver 1963). It came from information theory and measures the disorder observed in a sample. This disorder is the number of individuals belonging to each species; the more species there are and the more even their distribution, the greater the index becomes (Pielou 1966).

Species Evenness measures how similar the abundances of different species are. When there are similar proportions of all species in a sample then evenness is one, but when the abundances are different (as some species may be rarer than others) then the value varies from one.

Species richness is the total number of species in a given environment (the alpha diversity). Richness estimators are used to predict the expected number of species in an environment, including unseen species (defined as the species likely to be present in a larger sample, but that are missing from the actual sample data i.e. rare species). This does not show the complexity of an environment: no information can be inferred as to how the community is distributed among the species found.

The need for analyses that bypass OTUs

There are various drawbacks to analyses that depend on OTUs. The first, most critical point is that there is no 'universal standard' as to what level of identity defines an OTU. Some studies quote it as 95% (Martin 2002), others 97% (McCraig *et al.* 1999) and the others still at 99% (Kroes *et al.* 1999). This makes comparison of data from different sources difficult and presumably, the statistical outcomes from different definitions of OTU would be different.

The other limitation is that all OTUs are treated equally; two OTUs that may be phylogenetically distinct will be treated in the same way as two OTUs that are closely related species or subspecies. Therefore information on phylogenetic and genetic diversity is overlooked (Martin 2002).

Estimating Genetic Diversity

Lineage-per-time Plots

Population ecologists have developed methods for comparing the topology of a phylogenetic tree, which is used to estimate genetic diversity. If a phylogenetic tree is drawn such that the branch lengths from the root to the terminal taxa are equal, the distribution of divergence times can be visualised by plotting the number of lineages present as a function of time (Martin 2002). The shape of the plot will yield insights into the genetic diversity of the community, which is illustrated in Figure 8. An exponential plot (straight line) indicates that the community is experiencing constant birth and death rates. Concave plots suggest that the community contains an overabundance of highly divergent lineages (Martin 2002), which infers that the communty is very genetically diverse (Bohannan and Hughes 2003). A more genetically diverse community would be predicted to be more phenotypically diverse, if phenotypic variation is positively correlated with genetic variation as has been shown for higher organisms (Bohannan and Hughes 2003). Conversely, a convex plot indicates a

communty with an excess of closely related lineages, and infers a less genetically diverse community. Differences in tree topology can be further used to indicate the processes important in the formation of the community. An excess of closely related lineages implies that a recent selection event occurred, which may arise, for example, after antibiotic treatments (Martin 2002); an excess of highly divergent lineages may be as a result of competitive exclusion (Bohannan and Hughes 2003).

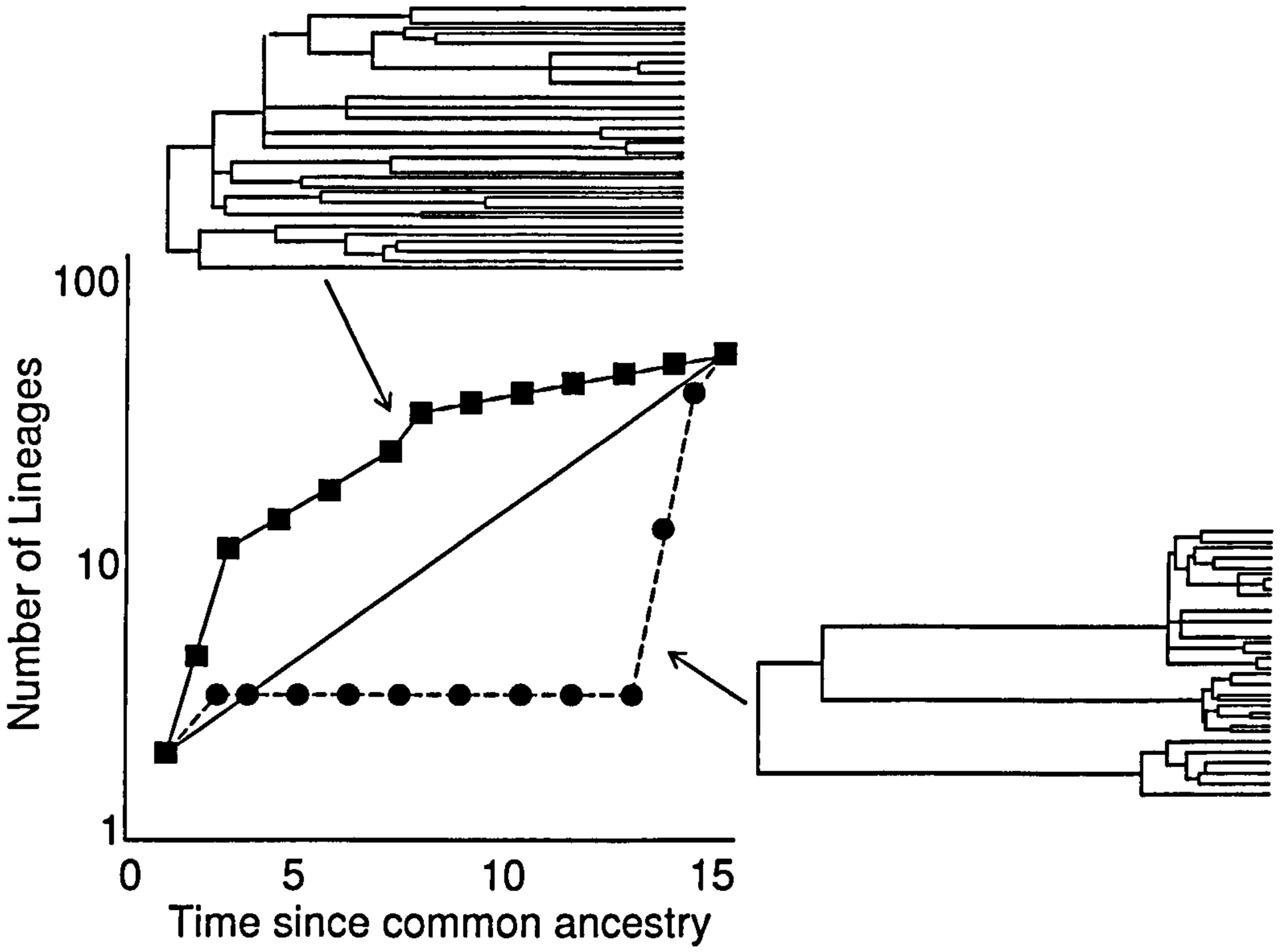


Figure 8 Lineage-per-Time Plots.

This graph illustrates the possible outcomes from a lineage-per-time plot. An exponential plot (straight line) indicates that the community is experiencing constant birth and death rates. A concave line (top) is indicative of a community with an overabundance of highly divergent lineages and high genetic diversity. The topology of the resulting phylogenetic tree is indicated along with this plot. A convex line (bottom) indicates a community with an excess of closely related lineages and less genetic diversity. The topology of this phylogenetic tree is also indicated. Redrawn from (Martin 2002).

Phylogenetic Approaches for Comparing Community Structure

The problem with the previous approaches is that different sequences are treated equally, and so the diversity for two communities may be shown to be the same by statistical comparison, even though they harbour very different sequences. For example, if two communities do not have any sequences in common, then the assumption is that diversity overlap is zero; however, if for every sequence in community A has a closely related sequence in community B, the two communities would have almost identical *phylogenetic* diversity, despite the previous observation that diversity overlap was nil (Martin 2002). Thus, phylogenetic information can be used to assess the 'degree of differentiation' between communities.

Two approaches have been used to assess the degree of differentiation. The first is the coefficient of differentiation (Fst), which compares the genetic diversity within a community to the genetic diversity of the two communities combined. For example, considering the hypothetical situation previously described, the level of diversity in each of the communities would be equal to the diversity of the two communities combined, hence Fst ~ 0. Therefore, even though the species do not overlap, the genetic diversity does (Martin 2002).

The alternative approach is the Parsimony test (P test). It is also called the Phylogenetic Test. Given a set of sequences sampled from two communities, a parsimony score can be calculated that gives a minimum number of changes to explain the observed distribution (Martin 2002).

These two approaches are then used to compare the structure of two communities, depending on whether the values for Fst and P test are statistically significant. Figure 9 illustrates the possible outcomes of these analyses. Significance for both tests implies that there is less genetic diversity in each sample than for the two combined and that the samples harbour distinct phylogenetic lineages. Insignificance for both implies that the genetic diversity in the two samples is statistically indistinguishable and that they probably came from the same community. A significant P test and insignificant Fst test implies that two communities harbour high levels of diversity (highly divergent lineages), but the phylogenetic lineages present in each sample are different. A significant Fst test and insignificant P test implies that the genetic diversity of each sample is less than for the two combined, even though the sampling is statistically indistinguishable, which occurs if each sample contains unique groups of closely related microbes that are evenly distributed across the tree (Martin 2002).

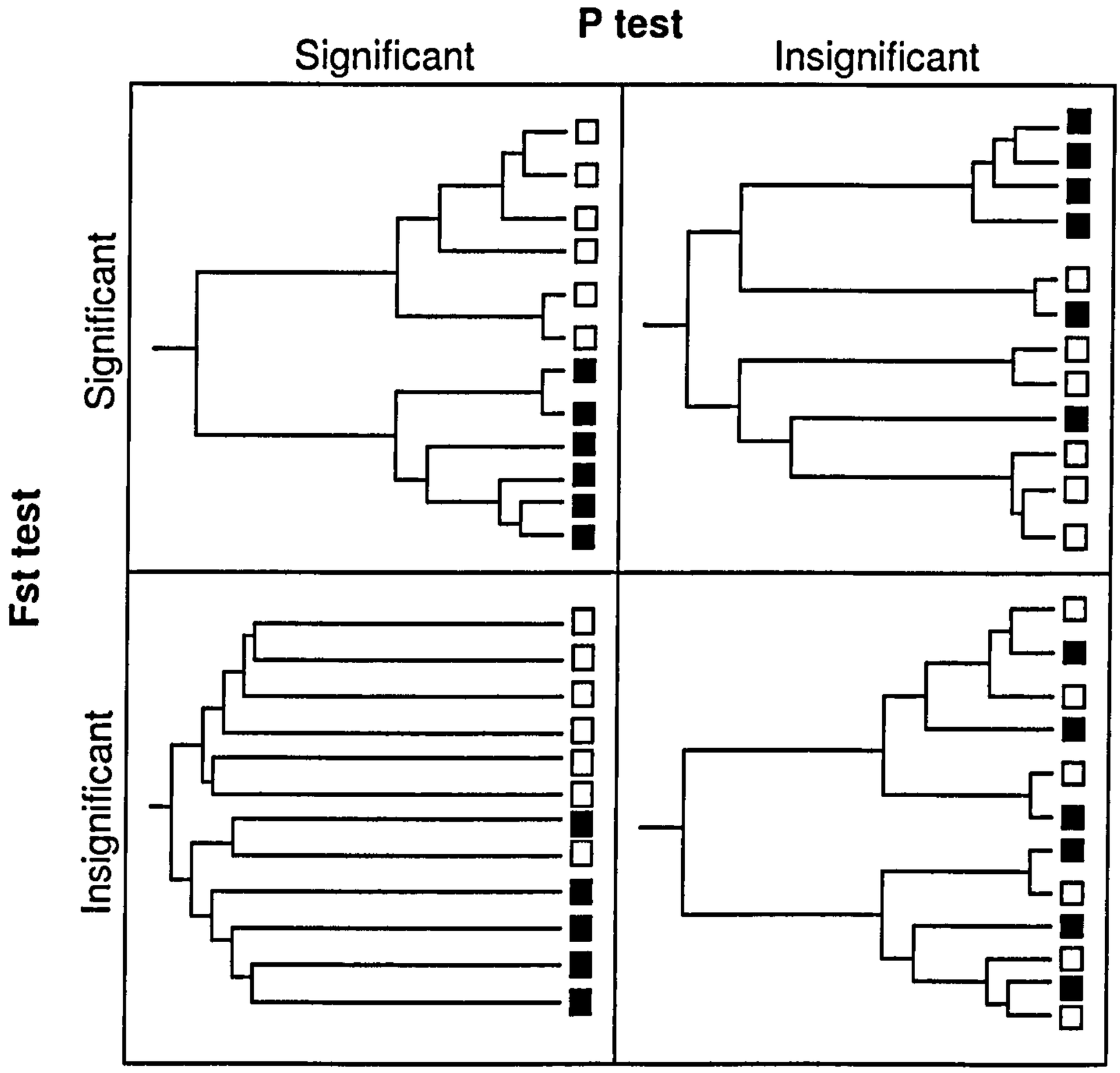


Figure 9 Outcomes of P test and Fst test.

The outcomes for the combination of the Fst and P test are illustrated and show whether two communities (white squares and black squares) are more or less genetically diverse and whether the two communities are disparate. Significance for both tests implies less genetic diversity in each sample than for the two combined and the samples harbour distinct phylogenetic lineages. Insignificance for both implies that the genetic diversity is statistically indistinguishable and that they probably came from the same community. A significant P test and insignificant Fst test implies that two communities harbour high levels of diversity, but the phylogenetic lineages present in each sample are different. A significant Fst test and insignificant P test implies each sample contains unique groups of closely related microbes that are evenly distributed across the tree. Redrawn from (Martin 2002).

Microbial Biogeography

Macro-ecologists have long been involved in the biodiversity and geographic distribution (biogeography) of plants and animals, which can be affected by speciation, extinction, dispersal and species interactions. Recently, the biogeography of microorganisms has been studied, though there are still doubts as to whether microorganisms are able to exhibit biogeography (Dolan 2006). This has stemmed from the fact that microorganisms are easily dispersed: 'everything is everywhere; the milieu selects,' has been a famously quoted hypothesis from Baas-Becking. A recent review, however, has shown what is known about the possible biogeography of microorganisms in light of macroorganisms (Martiny *et al.* 2006), simply for the reason that macroorganisms have been better studied.

The null hypothesis in biogeography is that microorganisms are distributed randomly over space. Upon rejection of this hypothesis, the second hypothesis is that the biogeographical distribution of microorganisms reflects the influence of contemporary environmental variation. This assumes that environmental factors maintain the modern microbial assemblages observed in a given sample. It also assumes that geographic distance does not affect biodiversity due to the wide dispersal of microorganisms. The third hypothesis is that the spatial variation is due to the effects of historical events, which disputes the hypothesis that 'everything is everywhere.' It implies that historical events can limit dispersal and past environmental conditions may have lead to genetic divergence between different microbial assemblages. The fourth hypothesis is that the distributions reflect the influences of both past events and contemporary environmental conditions (Martiny *et al.* 2006).

Protists that are defined by morphology are widely distributed, but are composed of genetically distinct communities; while viruses often display wide local diversity, but the ability for viruses to disperse easily between different environments means that local diversity is probably equivalent to global diversity (Dolan 2006). Bacteria, however, display different biogeographical distribution. A study on the biogeography of soil bacteria across North and South America showed that differences were largely attributed to soil pH; diversity was highest in neutral soils and lowest in acidic soils (Fierer and Jackson 2006), which differs fundamentally from the biogeography of macroorganisms. Bacteria in an estuary in Massachusetts were found to vary along with the salinity gradient and that a mixture of communities can shift to native estuarine communities provided that the community have

adequate bacterial growth rates and long residence time (Crump *et al.* 2004). Thus, with respect to bacteria, distributions can range from quite wide to limited to distinct environments.

Viruses Infecting *Archaea*

The three domain description of cellular life on earth, *Eukarya*, *Bacteria* and *Archaea* is a firmly established biological tenet (see previously) (Woese *et al.* 1990). Each domain has an associated, probably vastly diverse, virus population (Breitbart *et al.* 2002; Breitbart *et al.* 2003; Cann *et al.* 2005; Edwards and Rohwer 2005; Angly *et al.* 2006). Thousands of viruses infecting the domain *Eukarya* have been described and many of their DNA/RNA genomic sequences determined (van Regenmortel *et al.* 2000). Between 5000 and 6000 viruses infecting the domain *Bacteria* (bacteriophages or phages for short) have been described, at least morphologically, although rather fewer DNA/RNA genomic sequences have been determined (Ackermann 2007).

In contrast, we are profoundly ignorant about viruses infecting the domain *Archaea*. Just 40 or so have been described and the genomic sequences of only a dozen viruses have been determined. All archaeal viruses so far discovered have dsDNA genomes, both linear and circular (Prangishvili *et al.* 2006a; Ackermann 2007). Archaeal viruses having an RNA genome have not yet been identified and *perhaps* do not exist (Prangishvili *et al.* 2006a).

The domain *Archaea* is divided into three established kingdoms: the *Crenarchaeota*, the *Euryarchaeota* and the uncultivated *Korarchaeota* (hyperthermophiles); recently, nanoarchaea have been identified (Barns *et al.* 1996; Boone and Castenholz 2001; Huber *et al.* 2002), though this latter group has been argued to be part of the *Euryarchaeota* (see previously) (Brochier *et al.* 2005) Virus particles associated with the first two phyla have been identified (Prangishvili *et al.* 2006a). About 20 viruses of the *Crenarchaeota* have been identified, often with unusual shapes. Examples include *Sulfolobus* neozealandicus droplet-shaped virus (SNDV) that infects *Sulfolobus* (Arnold *et al.* 2000), *Acidianus* bottle-shaped virus (ABV) (Häring *et al.* 2005a) and *Acidianus* two tailed virus (ATV), which is a two-tailed virion that has an extracellular stage in its life cycle (Häring *et al.* 2005b; Prangishvili *et al.* 2006b). These morphologies have never been observed elsewhere; these viruses have no obvious relationship to phage infecting members of the domain *Bacteria* (Prangishvili *et al.* 2006a; Ackermann 2007).

Similarly, about 20 viruses infecting members of the *Euryarchaeota* have been identified (Prangishvili *et al.* 2006a) of which 15 infect haloarchaea (Dyall-Smith *et al.* 2003), which have often been termed 'haloviruses'. They are mostly head/tail viruses of the order

Caudovirales, which include the families *Myoviridae* and *Siphoviridae*, and may be distantly related to those infecting the domain *Bacteria* (Prangishvili *et al.* 2006a). Other morphotypes have also been observed, such as spindle-shaped and icosahedral (round) (Dyall-Smith *et al.* 2003); star-shaped virus-like particles were also observed in the Dead Sea (Oren *et al.* 1997a).

Haloviruses

To date, only 15 haloviruses (viruses infecting halophilic bacteria and the haloarchaea) have been reported, and only six viruses of the haloarchaea have been fully sequenced; others have been partially sequenced (Dyall-Smith *et al.* 2003). The first haloarchaeal virus, Hs1, was discovered by accident in 1974, which resembled members of the *Myoviridae* (Torsvik and Dundas 1974). Subsequently, viruses Hh1 and Hh3 were isolated from *Halobacterium salinarum* from a fermented anchovy sauce, which are both Bradley Group B siphoviruses with 37.6 kb and 29.6 kb linear dsDNA genomes respectively (Pauling 1982; Rohrmann *et al.* 1983). Virus φN is a head/tailed virus with a 56 kb linear dsDNA genome and also infects *Halobacterium salinarum* (Vogelsang-Wenke and Oesterhelt 1988). The earliest haloarchaeal virus to be intensely studied was φH, a temperate myovirus with an approximated 59 kb dsDNA genome (only 60% of the genome has been sequenced), whose host is *Halobacterium salinarum* (Schnabel *et al.* 1982).

Virus φCh1 is a temperate myovirus, with a 58.5 kb linear dsDNA genome with RNA species in mature phage particles; its host is the haloalkaliphile *Natrialba magadii* (Witte *et al.* 1997; Klein *et al.* 2002). It was isolated from a strain that originated from Africa (Tindall *et al.* 1984). This virus showed remarkable sequence similarity to the virus φH, even though their hosts live in distinct environments with respect to pH (Klein *et al.* 2002).

Other haloviruses whose genomes have been fully sequenced were isolated by the Dyall-Smith laboratory in Melbourne from hypersaline sources in Australia. Lytic myoviruses HF1 and HF2, closely related due to a large recombination event (Tang *et al.* 2004), have linear dsDNA genomes of 75.9 kb and 77.7 kb respectively, and infect the haloarchaea *Haloferax volcanii* and *Halorubrum coriense* respectively (Nuttall and Dyall-Smith 1993; Nuttall and Dyall-Smith 1995; Tang *et al.* 2002). The spindle-shaped virus His1 and the distantly related virus His2 have linear dsDNA genomes of 14.5 kb and 16 kb respectively, both with lytic and carrier status in *Haloarcula hispanica* (Bath and Dyall-Smith 1998; Bath *et al.* 2006). Finally a lytic icosahedral (round) virus SH1, having a linear dsDNA genome of 31kb infects

Haloarcula hispanica and a natural *Halorubrum* isolate (Bamford *et al.* 2005; Porter *et al.* 2005). This virus has an internal lipid layer, assigning it to the *Tectiviridae* (Bamford *et al.* 2005).

A high number of virus-like particles were detected in the Dead Sea; between 0.9 and 7.3 x 10^7 per ml of water were detected, which outnumber bacteria by a factor of 10 (Oren *et al.* 1997a) and so there is a large potential for finding novel haloviruses in the Inner Mongolian salt lakes. A limiting factor in the discovery of haloarchaeal viruses is the use of laboratory strains as hosts for isolating viruses from the environment. Using dominant ubiquitous haloarchaea that have been cultivated from the environment may allow identification of many more novel haloviruses (Dyall-Smith *et al.* 2003). During the course of this work, a number of bacterial and archaeal strains were cultivated from Inner Mongolian salt lakes. Though it is unlikely that they were the most dominant species in the salt lakes, they provided a range of hosts in which to propagate novel haloviruses.

AIMS AND OBJECTIVES

There are two parts to this thesis. The first was to investigate the microbial ecology of both the hot spring and salt lake environments by the following aims and objectives:

- Discover what microorganisms inhabit the hot springs and salt lakes using molecular methods (16S rRNA and 18S rRNA gene analysis).

- Look for the presence of SHOW, *Korarchaeota* and nanoarchaea in these environments.

- Infer structure and function of the communities; compare this to previous studies.

- Look at biodiversity within these environments by statistical analyses.

- Investigate microbial biogeography in the salt lakes, i.e. the factors involved in driving microbial community composition.

- Use cultivation methods to discover which microorganisms inhabit the salt lakes and to see how this approach compares with the data from 16S rRNA gene analyses.

The second part of this thesis was to look at the viral population in the salt lakes by the following aims and objectives:

- Isolate novel haloviruses on lawns of *Bacteria* and *Archaea* previously isolated from the salt lakes.

- Determine virus structure by transmission electron microscopy (TEM).

- Characterise viral genome (DNA or RNA, single or double stranded, circular or linear).

- Sequence viral genome and deposit the sequence in a database.

MATERIALS AND METHODS

Materials

The salts and reagents used in this study are listed in Appendix 1, p355. The antibiotics were supplied by Sigma. Two DNA ladders (1 Kb DNA ladder and the Lambda DNA *Hind* III Digest fragments) were supplied by Gibco-BRL (Invitrogen) and the Lambda DNA mixed digest ladder was supplied by Sigma. The protein standard marker Mark 12™ Unstained Standard was also supplied by Invitrogen. The restriction enzymes, DNase I, Exonuclease III and T4 DNA ligase were supplied by New England Biolabs (NEB). *Taq* polymerase was supplied by ABgene.

E. coli JM109 competent cells were supplied by Promega. *E. coli* DH5α was supplied by Invitrogen.

The pGEM T-Easy TA cloning kit and the pUC18NotI plasmid were supplied by Promega.

The GenomicPrep Cells and Tissue DNA Isolation Kit and the GenomiPhi DNA Amplification Kit were supplied by Amersham Biosciences. The UltraClean Soil DNA Isolation Kit was supplied by Mo Bio Laboratories. The Wizard® *Plus* SV Minipreps kit was supplied by Promega. The QIAEX® II Gel Extraction Kit and the QIAquick PCR Purification Kit were supplied by Qiagen.

Oligonucleotides were synthesised by VH Bio.

Sterilisation

Glassware, sampling equipment, tongue depressors and toothpicks were sterilised by autoclaving at 121°C for 15 min in an autoclave. All media were prepared using deionised distilled water (dH_2O) and sterilised by autoclaving at 121°C for 15 min. Deionised distilled water was produced by reverse osmosis using Elga water purification equipment. For particularly sensitive work such as Polymerase Chain Reaction (PCR), nanopure water (nH_2O) was used, supplied by Sigma. This water had been filter sterilised using a 0.22 µm filter and is DNase and RNase free. Heat sensitive material was also filter sterilised by passing through a 0.22 µm filter (Millipore).

Bacterial Strains, Plasmids, Primers and Growth Conditions

The bacterial strains, plasmids and primers used in this study are listed in Table 8. Diagrams of plasmids pGEM T-easy and pUC18 NotI are shown in Figure 10.

Cultures of *E. coli* strains were all grown in Luria Bertani (LB) broth or on LB agar at 37°C. The antibiotics used as required were ampicillin (100 µg/ml) and mitomycin C (10 mg/ml to 1 mg/ml). Both were made up in dH_2O, filter sterilised and stored at -20°C and 4°C respectively until use within one month.

Table 8 Bacterial Strains, Plasmids and Primers.

Bacterial Strains	Relevant Phenotype	Reference
E. coli JM109	*recA*1, *endA*1, *gyrA*96, *thi*, *hsdR*17, *supE*44, *relA*1, Δ(*lac-proAB*)/F' [*traD*36, *proAB*+, *lac1*^{q}, *lacZ*ΔM15]	Promega
DH5α	F-, φ80*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *recA*1, *endA*1, *hsdR*17(rk-, mk+) *phoA*, *supE*44, *thi*-1, *gyrA*96, *relA*1, λ-	Invitrogen

Plasmids	Relevant Phenotype	Reference
pGEM T-Easy TA cloning vector	*lacI*, *lacZ*, Amp^r	Promega
pUC18NotI	ColE1, *bla*, *lacI*, *lacZ*, Amp^r	(Yannish-Peron *et al.* 1985)

16S rRNA gene Primers[a]	Use	Sequence (5' to 3')[b]	Reference
27Fb (or FD1)	Universal bacterial forward primer	AGA GTT TGA TCC TGG CTC AG	(Weisburg *et al.* 1991)
27Fa	Universal archaeal forward primer	TCY GGT TGA TCC TGS CGG	(McGenity *et al.* 1998)
SHOWprb	SHOW forward primer	ACG GCA CAA CAG AGA CGC	(Burns *et al.* 2004b)
rp1	Universal reverse primer	ACG GHT ACC TTG TTA CGA CTT	(Weisburg *et al.* 1991)
K236F	*Korarchaeota* forward primer	GAG GCC CCA GGR TGG GAC CG	(Brunk and Eis 1998)
K1390R	*Korarchaeota* reverse primer	ACG GCC GGT GTG TRC AA	(Brunk and Eis 1998)
N3F	Nanoarchaea forward primer	TCC CGT TGA TCC TGC G	(Huber *et al.* 2002)
N1510R	Nanoarchaea reverse primer	ACG GCT ACC TTG TGT CGA CTT	(Huber *et al.* 2002)

Table 8 Continued.

18S rRNA gene Primers [a]	Use	Sequence (5' to 3') [a]	Reference
EK82F	*Eukarya* forward primer	GAA ACT GCG AAT GGC TC	(López-García *et al.* 2001)
U1492R	*Eukarya* reverse primer	GGT TAC CTT GTT ACG ACT T	(DeLong 1992)
1Af	*Eukarya* forward primer	CTG GTT GAT CCT GCC AG	(Díez *et al.* 2001)
516r	*Eukarya* reverse primer	ACC AGA CTT GCC CTC C	(Díez *et al.* 2001)

Vector Primers [a]	Use	Sequence (5' to 3')[a]	Reference
M13F	Forward vector primer	GTT TTC CCA GTC ACG AC	Promega
M13R	Reverse vector primer	CAG GAA ACA GCT ATG AC	Promega

[a] Codes for redundant bases are as follows:

D = A, G, T

H = A, C, T

K = G, T

M = A, C

N = A, C, G, T

R = A, G

S = C, G

W = A, T

Y = C, T

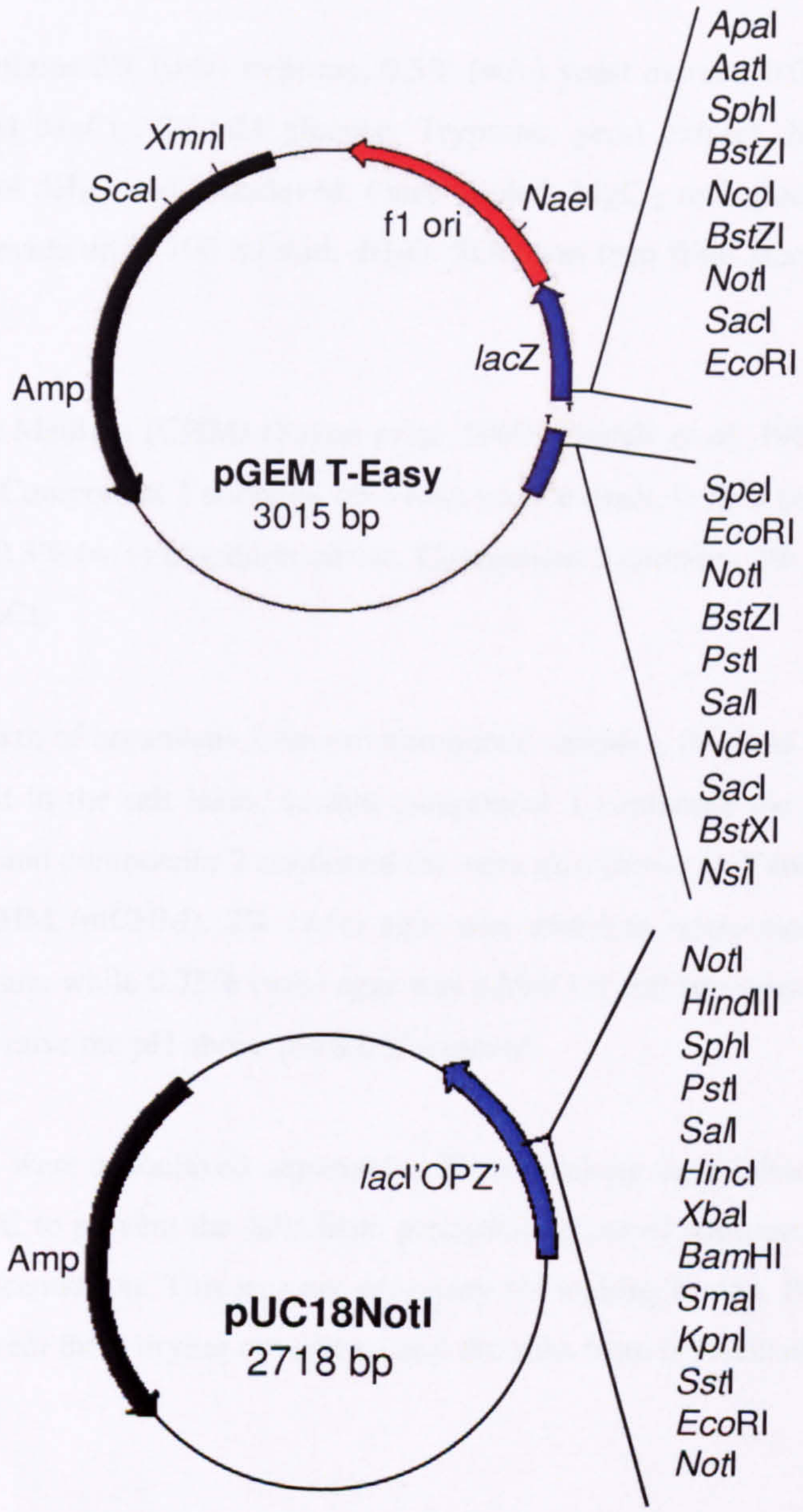


Figure 10 Plasmids.

Diagrams of plasmids pGEM T-Easy (adapted from Promega) and pUC18NotI adapted from (Yannish-Peron *et al.* 1985).

Media

LB broth contains 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl and is adjusted to pH 7.0 with 1 M NaOH. LB agar is LB broth solidified with 0.5% (w/v) agar.

SOC medium contains 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.05% (w/v) NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose. Tryptone, yeast extract, NaCl, and KCl were dissolved in 97 ml dH_2O and autoclaved. Once cooled, $MgCl_2$ and glucose were added and the final volume made up to 100 ml with dH_2O. SOC was then filter sterilised and stored at -20°C until use.

Classic Halophile Medium (CHM) (Payne *et al.* 1960; Tindall *et al.* 1980) broth is made in two components. Component 1 contains 1% (w/v) yeast extract, 0.75% (w/v) casamino acids, 0.2% (w/v) KCl, 0.3% (w/v) trisodium citrate. Component 2 contains 2% (w/v) $MgSO_4.7H_2O$ and 20% (w/v) NaCl.

However, for growth of organisms from environmental samples, this was modified to contain similar salts found in the salt lakes, so that component 1 contained the KCl concentrations shown in Table 9 and component 2 contained the salts also shown in Table 9 for each lake to make modified CHM (mCHM). 2% (w/v) agar was added to component 1 if required for mCHM bottom agars, while 0.75% (w/v) agar was added for soft top agars. NaOH was added to component 2 to raise the pH above pH 8.0 if required.

Both components were autoclaved separately. When making agar plates, the components were mixed at 60°C to prevent the salts from precipitating out of solution, since this medium has a high salt concentration. This was not necessary for making broths. Plates were stored in sealed bags to prevent them drying out and to stop the salts from crystallising.

Table 9 Concentration of salts in modified Classic Halophile Medium (mCHM).

These are the concentrations of salts in component 2 of modified Classic Halophile Medium (mCHM) for each salt lake. All values are in g/L.

Sampling Site[a]	$Na_2B_4O_7$	NaBr	$CaCl_2$	KCl	LiCl	$MgCl_2.7H_2O$	NaCl	Na_2SO_4	$NaHCO_3$	Na_2CO_3	$MgSO_4.7H_2O$
BJ1	1.62	0.83	0	2.48	0	71.16	130	91.2	0.62	0.36	-
CG1	1.45	0.13	0	1.04	0	0.21	49.1	30.56	30.2	21.76	-
EJ3	1.69	1.44	0.34	5.14	0	120	116.9	-	0.83	2.47	120
EN1	2	0.98	0	2.91	0.14	77.3	120	-	0.34	0.88	70
SH1	4.96	0.53	0.29	11.18	0	-	120	78.1	0.62	1.38	30.06
XH2	1.09	0.8	0.34	3.96	-	-	164	29.8	1.17	0.18	20.9

[a]BJ1: Lake Bagaejinnor

CG1: Lake Chagannor

EJ3: Lake Ejinnor

EN1: Lake Erliannor

SH1: Lake Shangmatala

XH2: Unnamed lake near Xilin hot

Solutions

Artificial Lake Bagaejinnor Water

Artificial lake water (ALW) for Lake Bagaejinnor was made in dH_2O containing 0.3% (w/v) KCl, 0.162% (w/v) $Na_2B_4O_7$, 0.084% (w/v) NaBr, 7.116% (w/v) $MgCl_2.7H_2O$, 13% (w/v) NaCl, 4.56% (w/v) Na_2SO_4, 0.062% (w/v) $NaHCO_3$ and 0.036% (w/v) Na_2CO_3. This was autoclaved and stored at room temperature until required.

Halovirus Diluent

Diluent was made for halovirus dilution and storage and contains 60% (v/v) artificial lake water, 29% (v/v) dH_20, 1% (v/v) 1 M Tris-HCl buffer, pH 7.2 and 10% (v/v) glycerol. This was autoclaved and stored at room temperature until use.

Phosphate Buffered Saline (PBS)

PBS contains 0.8% (w/v) NaCl, 0.121% (w/v) K_2HPO_4 and 0.034% (w/v) KH_2PO_4. This was dissolved in dH_2O and autoclaved. It was stored at room temperature until use.

Tris-HCl

Various amounts of Tris(hydroxymethyl) aminomethane (Tris) were dissolved in dH_2O depending on the desired concentration. Tris dissolves to give an alkaline solution; therefore the pH was adjusted with hydrochloric acid (HCl). This was autoclaved and stored at room temperature.

Tris-EDTA (TE)

Tris(hydroxymethyl) aminomethane/Hydrochloric acid, Ethylene diamine tetra acetic acid or Tris-EDTA (TE) consists of 10 mM Tris and 1 mM EDTA. The pH was adjusted using HCl.

5-bromo-4-chloro-3-indolyl-β-d glactopyronoside (X-Gal)

50 mg was dissolved in 1 ml N,N'-dimethyl-formamide in a glass bijoux. This was filter sterilised and used immediately as required in LB agar to make selective media.

Isopropyl β-D thiogalactopyranoside (IPTG)

119 mg was dissolved in 5 ml dH_2O and filter sterilised. This was stored at -20°C in a light-tight container until use within one month. IPTG was used as required in LB agar to make selective media.

Phenol chloroform

This mixture was used to remove proteins from preparations of nucleic acids. Chloroform denatures proteins and facilitates the separation of aqueous and organic phases. The mixture consists of equal parts of equilibrated phenol and chloroform. Isoamyl alcohol was also added to reduce foaming during extraction, so that the proportions of phenol: chloroform: isoamyl alcohol was 25:24:1. This was stored under 100 mM Tris-HCl, pH 8.0 in a light-tight bottle at 4°C until use for up to one month.

Measuring pH, Temperature and Salinity of the Hot Springs and Salt Lakes

The pH of the hot spring water or salt lake water was measured using pH strips (Merck).

The temperature was measured using a Solomat 520C temperature monitor or a Hanna K-Type thermocouple with SP weighted tanker probe (Jencons, Leighton Buzzard, UK) according to manufacturer's instructions. The temperature was measured at a distance by attaching the probe to the end of a 1 m pole.

The salinity was measured using a Hanna HI 8633 or HI9033 multirange conductivity meter (Jencons, Leighton Buzzard, UK), which was calibrated to 20°C with a temperature coefficient of 2% according to the manufacturer's instructions. All hot spring water readings were made at the 19.9 mS/cm range. All salt lake water readings were off the scale, hence they were serially diluted with distilled water and readings made at the 199.9 mS/cm range. Water conductivity gave an indication of total dissolved salts according to the following equation:

Total Ion Concentration (g/L) = 0.001K, where K = conductivity (mS/cm)

Titrations of Carbonate and Bicarbonate

Concentrations of carbonate (CO_3) and bicarbonate (HCO_3) were found by titration of lake water with H_2SO_4 using a Digital Titrator Model 16900 (Hach Systems for Analysis) according to the following reaction:

$$CaCO_3^{2-} + H_2SO_4 \rightarrow H_2CO_3 + CaSO_4 \rightarrow H_2O + CO_2 + CaSO_4$$

Titrations were done according to manufacturer's instructions. Briefly, a delivery tube was inserted in the titration cartridge containing H_2SO_4, which was flushed of air by turning the delivery knob until a few drops of H_2SO_4 were released from the end of the delivery tube. The counter was then reset to zero. The lake water samples were predicted to have a $CaCO_3$ concentration between 100 to 400 mg/L and so the 1.6 titration cartridge was used.

The sample to be titrated was then placed in a sterile Erlenmeyer flask and made up to an appropriate dilution with dH_2O so that the final volume of solution was 100 ml to make it compatible with the 1.6 titration cartridge. The indicator phenolphthalein was added (0.01% phenolphthalein dissolved in 50% (v/v) ethanol). This indicator is pink in alkali and colourless in acid. The titrator was placed in the sample such that the tip of the delivery tube was in the solution and the titrator was clamped to a stand. The delivery knob was turned to release H_2SO_4 into the sample until the solution turned from pink to colourless. This is the first end point, which occurs when the solution reaches pH 8.3, and so the number of digits required to reach this point was recorded. Phenolphthalein alkalinity (P Alkalinity) was determined by the following calculation:

P Alkalinity = Digits required x Digit multiplier,

where the digit multiplier is 1 for a 1.6 titration cartridge.

A second indicator, bromcresol green-methyl red (100 mg bromcresol green and 20 mg methyl red dissolved in 100 ml 95% (v/v) ethanol) was added to the solution. This indicator is blue in alkali and red in acid, but has several colour changes in between. It is greenish blue-grey at pH 5.1, light violet-grey at pH 4.8 and light pink at pH 4.5. The end point taken for lake water samples was pH 4.5 as it was assumed that the samples were a 'complex system' (according to manufacturer's instructions). The number of digits required to reach this end point was also recorded. Total alkalinity (T Alkalinity) was determined by the following calculation:

T Alkalinity = Total digits required x digit multiplier,

where the digit multiplier is 1 for a 1.6 titration cartridge.

The concentration of carbonate and bicarbonate was then determined by the 'Alkalinity Relationship Table' according to manufacturer's instructions. Briefly, where P Alkalinity was less than half the T alkalinity, carbonate alkalinity was calculated as follows:

CO_3 alkalinity (mg/L) = 2(P Alkalinity)

Bicarbonate alkalinity was calculated as follows:

H_2CO_3 alkalinity (mg/L) = T Alkalinity – 2(P Alkalinity).

Determining Chemical Composition of Salt Lake Water

Chemical analysis of the salt lake water was carried out by Inductively Coupled Plasma Optical Emission Spectrometry (ICP-ES). Samples were sent for analysis to the Geology Department at the University of Leicester. Details are found at www.le.ac.uk/gl/lugs/.

Measuring Geographic Distances

GPS coordinates recorded at each sampling point were imported into MapSource™ according to the manufacturer's instructions to measure the geographic distances between the sites.

Cell Methods

Culturing Organisms from Environmental Samples

Approximately 0.5 g of sediment and salt crust were resuspended in 10 ml of appropriate mCHM broth. Ten fold serial dilutions of this were made up to 10^{-6} and 100 µl of each dilution were plated onto the appropriate mCHM agar plates. Cell suspensions from agitated filters were also serially diluted and plated onto mCHM agar plates.

Each sample was done in triplicate so that organisms could be cultured under different conditions: one plate was grown at 37°C aerobically and a second plate was grown at 30°C aerobically as this more closely resembles the conditions at the time of sampling. Inoculated plates were inverted during incubation and sealed in a bag to avoid the plates drying out. A third plate was grown anaerobically at 37°C in anaerobic jars (Oxoid) that contained a platinum catalyst to absorb oxygen from the air within. The plates were stacked in the jar and flushed with a mix of 10% hydrogen in nitrogen to a pressure of 0.7 bar. The entire jar was then placed in a 37°C incubator.

Agar plates were incubated for two months, but were periodically checked for the appearance of new colonies, since many halophilic organisms are slow growing. Colonies were picked with a sterile loop and streaked onto fresh plates, and incubated under the appropriate conditions. Cultures were then restreaked if necessary until pure cultures were obtained. For organisms that were particularly difficult to separate, colonies were resuspended in the appropriate mCHM broth and serially diluted enough that the different organisms become well separated. 100 µl of each dilution was plated, incubated and single colonies again picked and streaked onto fresh agar plates.

Virus Methods

Plaque Assay

The plaque assay has been extensively used to study viruses for enumeration and purification. In this technique, host cells are grown and immobilised within a soft layer of agar to form a confluent lawn, in which viruses are allowed to propagate. The soft agar restricts the diffusion of the virus resulting in localised areas where the viruses have repeatedly infected, multiplied and burst from the host cells. These can be clearly seen in the lawn as transparent, circular clearings called 'plaques.' Counting the number of plaques directly determines the number of infectious particles applied to the plate.

mCHM bottom agars were made containing 2% (w/v) agar as described previously, which acts as a support for the soft top agar. A separate bottle of mCHM soft top agar was made containing 0.75% (w/v) agar, also as previously described, which was kept molten in a 55°C water bath until required. Meanwhile, 300 µl of host cell culture (Optical Density [OD] approximately 0.2 at 695 nm) was infected with 30 µl of virus stock. Virus particles were left to adsorb onto host cells for 15 min at room temperature. If titres of virus stock were too high, it was serially diluted in mCHM broth before used to infect host cells. Infected cultures were added to 3 ml of agar cooled to approximately 50°C in universal tubes and mixed by swirling the tube vertically between the palms. This was immediately poured on top of the bottom agars and left to set. The plates were carefully inverted and incubated in a sealed bag at 37°C for a week or longer until plaques could be clearly seen in a lawn of cells.

The total number of infectious particles is defined as plaque forming units per ml (PFU/ml) and is calculated by the following equation:

PFU/ml = Total no. plaques on plate / d x V,

where d = dilution factor and V = volume viruses added to the plate.

Plaque Purification

Several plaques found in a lawn of cells may not *all* be formed by a single virus type. Since lytic viruses are clonal in origin, it is assumed that one plaque arises from just one virus particle. By selecting a single plaque in a lawn and using it to infect subsequent cultures, one species of virus can be selectively enriched.

In a lawn of infected cells, a single plaque was selected for purification. It was picked with a sterile toothpick by stabbing the soft top agar through the centre of the plaque. Virus particles stuck to the end of the toothpick were resuspended in 100 µl of mCHM broth. This was then used in the plaque assay. This process of purification was repeated twice to ensure that the viruses were purified. Virus particles remained stable in mCHM broth when placed at 4°C for at least one overnight.

Discovery of Novel Haloviruses

Spontaneous Lysis from Lytic and Lysogenic Viruses

Bacterial and archaeal strains were grown in 5 ml mCHM broth in an orbital shaker at 37°C, 150 rpm to an OD of about 0.2 at 695 nm. 1 ml of culture was used to form lawns in soft top agar as previously described, but without the addition of viruses. The plates were incubated over a period of time until confluent lawns of cells had grown. The lawns were then checked for the appearance of plaques formed by spontaneous lysis from both lytic and lysogenic viruses.

The remaining 4 ml of culture was left in the orbital shaker at 37°C, 150 rpm for an indefinite period of time and periodically checked for the appearance of cell lysis that may be due to infection from viruses, such as increased viscosity of the culture and the appearance of cell debris. When such cultures occurred, the cell debris was pelleted by centrifugation at 13000 x g for 5 min and subsequently passed through a 0.22 µm filter (Millipore). The supernatants were used to infect lawns of phylogenetically related strains and used in a plaque assay. Confluent lawns were then checked for plaques.

Lysogenic Viruses

Mitomycin C inhibits DNA synthesis. It was used to stress cultures by adding at sub-bacteriostatic concentrations, thereby inducing lysogenic phage to become lytic. 5 ml of bacterial and archaeal cultures were grown as described and used to make lawns in soft top agar. However, on adding 1 ml of culture to molten agar, various concentrations of mitomycin C (10 mg/ml to 1 mg/ml) was added and mixed with the cells and agar. This was poured on the bottom agars and incubated at 37°C until confluent lawns had formed. The lawns were checked for the appearance of plaques.

The remaining 4 ml of culture was also subject to stress by the addition of 1 mg/ml final concentration mitomycin C and left in the orbital shaker at 37°C, 150 rpm also for an indefinite period of time and checked periodically for the appearance of cell lysis. As previously described, filtered supernatants of such cultures were used to infect phylogenetically related strains and used in the plaque assay.

Lytic Viruses

Many lytic viruses are found outside the host cell. To isolate them from the lake water for further study, the following method was used. Bacterial and archaeal strains were grown as described. Salt lake water was passed through a 0.45 μm filter followed by a 0.22 μm filter. 10 μl of filtered lake water was added to 1 ml of cell culture and incubated at 37°C in an orbital shaker at 150 rpm overnight. This would allow any lytic virus present to infect the cells and increase their titre. The infected culture was then plated in soft top agar and the lawns checked for the appearance of plaques.

Host Range Experiments

Various archaeal isolates were grown in 5 ml of mCHM broth to an OD of approximately 0.2 695 nm at 37°C in an orbital shaker, 150 rpm. 30 μl of cells were infected with various dilutions of virus stock and used in a plaque assay. Plates were checked for the appearance of plaques in a confluent lawn.

Testing Virus Tolerance to Temperature, pH and Salt

To find out what conditions virus particles remained infective, they were exposed to various ranges of temperature, pH and salinity for at least 15 min, to allow virus particles to be disrupted and rendered functionless.

To test the temperature range, 50 µl aliquots of virus stock (10^6 pfu/ml resuspended in ALW) were exposed to temperatures between 30°C and 80°C. To test the salt tolerance, salinity of virus stock was diluted with ALW (without NaCl) or salinity was raised using 5.5 M NaCl solution. To test pH tolerance, HCl was used to lower the pH, while NaOH was used to raise the pH of the ALW.

After exposing the virus particles to different conditions, appropriate dilutions were made and used in a plaque assay for enumeration.

Storage of Isolates

Environmental bacterial and archaeal isolates were kept on beads at -80°C for long term storage. The beads allowed subculturing of isolates without the need to thaw the entire culture, since repeated freeze thaw cycles can be detrimental to the isolate. Glycerol was also added as a cryoprotectant. The following protocol was adapted from (Jones *et al.* 1991).

2 mm embroidery beads (Peak Dale Craft Products) were washed once in detergent and tap water. The detergent was then removed by an acid wash in 0.1 M HCl. To remove the acid, the beads were repeatedly washed in tap water until the pH of the water on the beads was the same as the pH of the tap water. The beads were finally washed in dH_2O and dried at 45°C in an oven. The beads were aliquoted to 1.75 ml glass vials (FGB-Trident) and autoclaved.

Cultures of bacteria and archaea were grown in appropriate mCHM broth to an OD of 2 to 4 (at 695 nm), since thick cell suspensions are less likely to lose viability after repeated subculturing. Cells were maintained on beads by adding sterile glycerol to cell cultures to a final concentration of 30% (v/v) and stored at -80°C until required.

Cells grown in microtitre plates were maintained by adding sterile glycerol to the wells to a final concentration of 10% (v/v). Adhesive films (ABGene) were secured on top of the plates,

which were stored at -80°C until required. A small scraping was taken from each well and grown on LB agar when required for further analysis.

Haloviruses were also stored in glycerol, which prevents them losing their infectivity. Viruses grown in soft top agar were harvested by adding 0.5 ml of halovirus diluent to each plate and scraping off the layer of soft top agar with a sterile glass spreader into a 250 ml sorvall tube. The agar was homogenised in halovirus diluent by vortexing the tube for 30 s. The agar and cell debris was pelleted by centrifugation at 16000 x g for 20 min. The supernatant was passed through a 0.45 μm filter and then a 0.22 μm filter to further remove cell debris and agar. The supernatant was stored at 4°C. The haloviruses remained stable for over a year.

Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

This technique separates proteins in an electric field through a discontinuous polyacrylamide gel as a matrix in the presence of SDS to denature the proteins. It was used in this study to examine whole cell proteins from cultured halophilic organisms. The Laemmli method was used in this study and is outlined briefly as follows.

The 'mini-slab' gel system was used to cast the cassettes. One cassette consists of two glass plates (10 cm x 10 cm) with two 0.5 mm thick spacers in between at the two ends, into which the polyacrylamide gel is poured and set. Twenty such cassettes were carefully placed next to each other in a casting stand. The cover for the casting stand was placed over the cassettes to hold them in place. Vaseline was used to seal the cover to prevent leaks. A comb for the wells was placed into a cassette and a mark was made 1 cm below this with a marker pen to indicate how high the lower layer of the gel should be poured. This comb was then removed.

This lower layer of the gel (resolving gel) was prepared in a sterile falcon tube. This layer is responsible for separating proteins by size. 12% gels were used to resolve polypeptides less than 100 kDa, containing 40% (v/v) of 30% acrylamide mix, 26% (v/v) 1.5 M Tris-HCl, pH 8.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate, 0.1% (v/v) N,N,N',N'-tetramethylethylenediamine (TEMED) and 32% (v/v) dH_2O. Polymerisation is a free-radical catalysed reaction. Ammonium persulfate and TEMED are the catalysts and so were added last. The solution was quickly mixed and poured into the casting stand. The top of the stand was gently tapped to redistribute the gel to all the cassettes to an equal height, as previously marked. Oxygen is a free-radical scavenger and can inhibit the polymerisation reaction. To

prevent this occurring to the top of the gels and to form a smooth horizontal surface to the gels (thereby forming straight bands later), they were overlaid with water saturated isobutanol. The gels were allowed to set for 30 min or longer. Any remaining polyacrylamide gel mix in the falcon tube was used and an indicator for when the gels had set.

Once set, the water saturated isobutanol was poured off and the gels allowed to air dry. The combs to form the wells were then placed in the cassettes. The upper layer of the gel (stacking gel) was prepared at 5% consisting of 17% (v/v) of 30 % acrylamide mix, 13% (v/v) 1.0 M Tris-HCl, pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate, 0.1% (v/v) TEMED and 68% (v/v) dH_2O. Again, the catalysts were added last. The solution was mixed and immediately poured into the cassettes. The top of the stand was tapped and the combs adjusted with sterile tweezers to let out any bubbles trapped under the teeth of the combs. The stacking gel is used to compress the proteins migrating from the sample wells before they reach the resolving gel.

Once set, the cassettes were carefully removed with a sterile scalpel and wrapped in moist tissue and Clingfilm. They were stored in a Tupperware box at 4°C until use within one month.

The halophilic organisms to be examined were grown in 5 ml of mCHM broth until the culture was turbid. 1.5 ml of culture was pelleted by centrifugation at 15000 x g for 10 min in a screw-capped tube. The supernatant was discarded and the pellet used for whole cell protein analysis. Previous experience showed that this method of preparing samples for analysis gave cleaner bands (W. D. Grant, personal communication).

Prior to use, a few milligrams of dithiothreitol (DTT) were added to an aliquot of treatment buffer, pH 6.8 (25% (v/v) 0.125 M Tris-HCl, 4% (w/v) SDS, 20% (v/v) glycerol, 0.02% (w/v) bromophenol blue). Pellets were resuspended in 100 µl of treatment buffer by vortexing. The samples were then boiled in a heating block (Grant) for 15 min. Cell debris was pelleted by centrifugation at 13000 x g for 10 min. The supernatant was transferred to a fresh tube and could be stored at -20°C until use.

The EDTA in the loading buffer reduces proteolytic activity, while SDS, DTT and heat are responsible for denaturing the proteins. SDS adds a negative charge to amino acids thereby

disrupting the proteins' secondary and tertiary structures since like charges repel. Quaternary structures (interaction with other molecules by hydrophobic bonds) are disrupted by heating since this weakens the overall structure and allows SDS to bind to hydrophobic regions. Disulphide bonds formed between amino acids with a sulfhydryl group are not broken by SDS. Therefore, the reducing agent DTT is added, specifically to attack these bonds.

Once samples were prepared, a single cassette was extracted and clamped into the top chamber of a vertical electrophoresis running stand and the whole apparatus was assembled. The edges around the top of the gel were sealed in 2% (w/v) agarose made up in running buffer (3.03 g/L Tris base, 14.4 g/L glycine, 1.0 g/L SDS; dissolved in dH_2O). Once set, the top and bottom chambers of the running stand were filled with running buffer until the top and bottom edges of the gel were submerged to a depth of at least 1 mm. A fan was attached to the running stand to prevent overheating, which can cause peculiar patterns in the bands. The samples were run initially at 120 mV while they were in the stacking gel and then at 200 mV when the samples were being separated in the resolving gel. Approximate running time for these size gels was 40 min.

To visualise the bands, the gel was stained in Coomassie stain (1.25 g Coomassie brilliant blue R250, 90 ml (v/v) 50% methanol, and 10 ml (v/v) acetic acid) for 3 min with shaking. The dye penetrates the entire gel, but only permanently sticks to the proteins. Therefore, the excess dye was removed with destain (10% (v/v) methanol, 5% (v/v) acetic acid, 85% (v/v) dH_2O). This usually takes several hours, during which destain in periodically renewed.

The gel could be stored by submerging in gel drier (20% (v/v) methanol, 10% (v/v) glycerol, and 70% (v/v) dH_2O) for one overnight. The gel could then be mounted in a frame between two pieces of cellophane and left to dry.

Determining Sizes of Protein Bands

The standard marker Mark 12™ Unstained Standard was run alongside the proteins to determine the sizes of the bands and to ensure that the bands had run equally along the gel. This standard consists of the following fragments (kDa): 200, 116.3, 97.4, 66.3, 55.4, 36.5, 31.0, 21.5, 14.4, 6.0, 3.5 and 2.5.

Nucleic Acid Isolation

DNA Extraction Procedures

GenomicPrep Cells and Tissue DNA Isolation Kit (Amersham Biosciences)

Total genomic DNA was isolated from biomass collected at the sampling site by the GenomicPrep Cells and Tissue DNA Isolation Kit. The biomass (either algal mat or cells collected from water filtration) was pelleted in a 1.5 ml microfuge tubes by centrifugation at 15000 x g for 5 min and the DNA extracted according to manufacturer's protocols. Briefly, the pellet was resuspended in cell lysis solution and incubated at 80°C for 5 min to lyse the cells. Cell lysis solution contains an anionic detergent to solubilise cellular components and a DNA preservative to limit the activity of endogenous and exogenous DNases. These samples were then stable for at least 18 months at room temperature and were stored as such until further processing in the laboratory.

RNA was removed by adding 3 µl of RNase A solution and incubating the sample at 37°C for 30 min. Cytoplasmic and nuclear proteins were precipitated by protein precipitation solution and vortexing for 20 s. The protein was pelleted by centrifugation at 15000 x g for 3 min and discarded. The DNA was then ethanol precipitated (see later). DNA was resuspended in 100 µl of DNA hydration solution and allowed to rehydrate overnight at room temperature. This was then stored at 4°C until further use.

UltraClean Soil DNA Isolation Kit (Mo Bio Laboratories)

Total genomic DNA was isolated from algal mats using the UltraClean Soil DNA Isolation Kit according to the manufacturer's instructions. This method uses mechanical forces by vortexing the sample with small beads to extract DNA from tougher material such as endospores. It also removes inhibitors that may affect further processes such as PCR.

Briefly, 2.5 g of algal mat was added to Bead solution (guanidine isothiocyanate and inert particles) and gently vortexed. Solution S1 (Tris-HCl and Sodium Dodocyl Sulphate) and Inhibitor Removal Solution were also added and vortexed at maximum speed for 10 min. The beads, soil and debris were pelleted by centrifugation at 10000 x g for 30 s and discarded. Solution S2 (aqueous solution of acetate) was added to the supernatant and vortexed for 5 s. The sample was incubated at 4°C for 5 min and then was centrifuged at 10000 x g for one min.

The supernatant was transferred to a clean tube. Solution S3 (guanidine-HCl) was added to the sample and vortexed for 5 s. This was transferred to a spin filter and centrifuged at 10000 x g for 1 min. The flow through was discarded and the DNA cleaned by the addition of Solution S4 (Tris-HCl, EDTA, NaCl and 50% (v/v) ethanol) and centrifugation at 10000 x g for 30 s. The spin filter was centrifuged for 1 min to get rid of excess Solution S4. The spin filter was transferred to a clean tube and the DNA eluted by the addition of 50 µl Solution S5 (Tris-HCl, pH 8.0) and centrifugation for 30 s. The DNA was stored at -20°C until further use.

Viral Nucleic Acid Extraction

Attempts to purify virus nucleic acid from infected liquid cultures were unsuccessful. Accordingly, viruses were grown in soft top agar as previously described. 0.5 ml of halovirus diluent was added to each agar plate and the viruses harvested by scraping off the layer of soft top agar into a 250 ml sorvall tube and homogenising by vortexing for 30 s and left for the haloviruses to soak out for approximately 30 min. The agar and cell debris was pelleted by centrifugation at 16000 x g for 20 min. The supernatant was transferred to a fresh clean tube. To increase the yield of virus particles, the pellet was resuspended in 2 ml of halovirus diluent and the previous steps of homogenisation and centrifugation were repeated.

The supernatant was passed through a 0.45 µm filter and then a 0.22 µm filter to further remove agar and cell debris. To remove exogenous nucleic acids, DNase I and RNase A were added, each to a final concentration of 1 µg/ml and left to degrade nucleic acids at room temperature for 30 min. The viral genomes were protected inside the virus capsid.

Virus particles were precipitated by the addition of 1/8 volume polyethylene glycol (PEG) 6000 solution (2.5 M NaCl, 20% (w/v) PEG 6000) and left to incubate for 15 min at room temperature. The viruses were pelleted by centrifugation at 15000 x g for 5 min. The supernatant was carefully removed and the pellet dissolved in 100 µl of PBS.

To remove the viral protein coats, the pellet was mixed with an equal volume of phenol chloroform and centrifuged for 30 s. The top aqueous layer containing the viral nucleic acid was removed and transferred to a fresh tube. To remove excess phenol chloroform, 1 ml of 100% diethyl ether was added and the tube placed at -20°C. Once the sample had solidified, the ether/phenol mix was carefully removed with a sterile Pasteur pipette. Excess diethyl ether

was removed by leaving the lid off the tube and placing it in a 37°C water bath for 2 min, allowing it to evaporate from the sample.

The nucleic acid was ethanol precipitated (see below). Nucleic acid was dissolved by adding 20 µl of TE, pH 8.0 and leaving it to rehydrate at 4°C for one overnight.

Plasmid purification

10 ml of overnight cultures were grown in LB broth containing the appropriate antibiotics. The plasmids were extracted from bacterial cells using the Wizard® *Plus* SV Minipreps kit (Promega) according to manufacturer's instructions. Briefly, 1 to 10 ml of cells were pelleted by centrifugation at 15000 x g for 5 min. The pellet was resuspended in cell resuspension solution (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 100 µg/ml RNase A) and lysed with cell lysis solution (0.2 M NaOH and 1% (w/v) SDS).

Proteins were inactivated by incubation with alkaline phosphatase and then precipitated using neutralisation solution (4.09 M guanidine-HCl, 0.759 M potassium acetate, 2.12 M glacial acetic acid). Proteins were pelleted by centrifugation at 13000 x g for 10 min and the supernatant transferred to a spin column. The plasmid DNA was bound to the column by centrifugation at 13000 x g for 1 min. The DNA was washed with column wash solution (60 mM potassium acetate, 8.3 mM Tris-HCl, pH 7.5, 0.04 mM EDTA, pH 8.0, 60% (v/v) ethanol) and eluted in 100 µl of nuclease free water. Plasmids were stored at -20°C until further use.

Manipulation of Nucleic Acid

Agarose Gel Electrophoresis

This is the most widely used method of separating, identifying and purifying DNA fragments. This technique involves separating DNA in an agarose matrix run in an electric field, in the presence of a fluorescent dye, which stains the DNA to enable it to be visualised under an ultraviolet light. More detailed protocols are available from Sambrook, Fritsch and Maniatis (Sambrook *et al.* 1989) but are briefly described as follows.

There is a linear relationship between the mobility of DNA and gel concentration, therefore large DNA fragments are run on a low concentration of agarose. Generally, high molecular weight DNA (3 kb to 12 kb) was run on 0.8% (w/v) agarose gels, fragments that were 3 kb to

1 kb were run on 1.2% (w/v) agarose gels and low molecular weight fragments (1 kb to 0.1 kb) were run on 2% (w/v) agarose gels.

Electrophoresis is carried out in the presence of a buffer. In this study I used Tris-acetate EDTA (TAE) buffer. It consists of 24.2% (w/v) Tris-base, 5.71% (v/v) acetic acid and 10% (v/v) 0.5 M EDTA, pH 8, 84.29% dH_2O. Agarose was dissolved in TAE buffer and heated until molten. This was then poured into a plastic tray sealed at both ends with autoclave tape, containing a suitable comb for the wells, and left to set. Once solid, the horizontal slab gels were set on a platform inside the electrophoresis tank. The remaining TAE buffer was then used to submerge the gel to a depth of at least 1 mm.

DNA samples were mixed with loading buffer (0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF and 15% (w/v) Ficoll in dH_2O), which were loaded into the wells.

DNA is negatively charged at neutral pH, and so when an electric field is applied across the gel, it migrates towards the anode. A voltage of 1 to 5 V/cm (distance between the electrodes) was applied to the gels using a power pack (Bio-Rad) until it appeared that the loading buffer had moved significantly far down the gel. DNA molecules migrate through the gels at a rate that is inversely proportional to the log10 of the number of base pairs (Helling *et al.* 1974). Hence DNA fragments become separated since the larger molecules migrate more slowly through the gel due to greater frictional drag.

DNA was visualised using fluorescent dyes, ethidium bromide and SYBR Green I, which intercalate between the bases. They allow visualisation of DNA under ultraviolet light by absorbing the radiation and re-emitting it in a different part of the light spectrum (red-orange for ethidium bromide and green for SYBR Green I).

Gels were stained according to manufacturer's instructions. Samples containing more than 50 ng DNA were stained with ethidium bromide by addition of the stain to precast gels at a final concentration of 0.5 μg/ml. Samples containing as little as 2 ng DNA can be visualised when stained with SYBR Green I. These samples were stained once the gel had been run by soaking the gel in TAE containing SYBR Green I for at least 40 min.

Pulse Field Gel Electrophoresis (PFGE)

Pulse Field Gel Electrophoresis (PFGE) is a sensitive size-dependent separation procedure for linear DNA molecules up to about 10000 kb that do not enter the gel under conventional agarose gel electrophoresis (Carle and Olsen 1984; Schwartz and Cantor 1984; Carle *et al.* 1986). This is done by the reorientation of the electrical current; the DNA molecule is forced to change direction and therefore migrates more slowly through the gel. This also allows separation of different sized fragments as the smaller fragments will begin moving in one direction faster than the larger ones.

Viral nucleic acids were ran on a 1% PFGE agarose (BioRad) in 0.5x Tris-Borate EDTA (TBE) buffer (0.089 M Tris/borate, 0.089 M boric acid, 2 μM EDTA) at 14°C in a CHEF DR-II apparatus (BioRad). The run time was 22 h with a voltage gradient of 6 V/cm and a linearly ramped pulse time of 50 to 90 s at an angle of 120 °.

Determining Sizes of DNA Fragments

The sizes of DNA fragments were determined by running a suitable marker alongside the samples. For fragments between 100 bp to 12 kb, the 1 Kb DNA Ladder was used, which consists of the following size fragments (in bp): 12216, 11198, 10180, 9162, 8144, 7126, 6108, 5090, 4072, 3054, 2036, 1636, 1018, 517, 506, 396, 344, 298, 220, 201, 154, 134 and 75. For large fragments up to 23 kb, the Lambda DNA *Hind* III Digest marker was used. This consists of the following fragments (in bp): 23130, 9416, 6557, 4361 2322, 2027, 564 and 125. For larger fragments up to 48 kb, the Lambda DNA mixed digest ladder was used, which consists of the following size fragments (in bp): 48502, 38416, 33498, 29946, 24508, 23994, 17053, 15004, 10086, 1503.

Gel Extraction Method

DNA fragments of the desired size were extracted from agarose gels using the QIAEX® II Gel Extraction Kit (Qiagen) according to manufacturer's instructions. Briefly, nucleic acid fragments were separated by agarose gel electrophoresis as described. The gel was viewed under ultraviolet light and the desired band was cut out using a sterile scalpel. DNA was extracted from the agarose by incubation with QIAEX II beads (silica particles) and QX1 solution at 50°C for 10 min. The beads were kept resuspended by vortexing every 2 min. As the agarose melted, the DNA was released and bound to the beads. The beads were pelleted by centrifugation at 13000 x g for 30 s and the supernatant discarded. The beads were washed

in solution QX1 to remove residual agarose and then washed twice in ethanol containing buffer. The beads were air dried for at least 15 min at room temperature. To elute DNA fragments that were less than 4 kb, the beads were incubated with 20 µl of nuclease free water or Tris-Cl, pH 8.5 for 5 min at room temperature. Beads were pelleted by centrifugation at 13000 x g for 30 s. To increase yield by 10 to 15% this was repeated. DNA was stored at -20°C until further use.

Quantifying DNA concentration

DNA that was visualised by agarose gel electrophoresis was quantified by comparison to the 1636 bp band in the 1 Kb DNA Ladder. This band contains 0.1% (w/v) of the total amount of DNA loaded in one lane. The ladder was diluted by 75% (v/v) with TE to a final concentration of 25 µg/µl. To quantify DNA, 2 µg ladder was loaded, so that there was 200 ng DNA in the 1636 bp band. The concentrations of DNA in the samples were estimated by comparison of their band intensities with the 1636 bp band.

DNA concentration was also measured using ultraviolet spectrophotometry. DNA samples were diluted in nH_2O. A quartz cuvette (Sarstedt) was placed in the spectrophotometer (Pharmacia) and washed *in situ* with nH_2O with a pipette. A blank solution of nH_2O was placed in the cuvette and the counter set to zero. The blank was removed *in situ* with a pipette. DNA samples were placed in the cuvette and read at 320 nm, 280 nm and 260 nm. The cuvette was washed with nH_2O between each sample. DNA concentration was determined by the following calculation:

$$[DNA]\ (\mu g) = (A_{260} - A_{320}) \times 50dv,$$

where A_{260} is the absorbance read at 260 nm, A_{320} is the absorbance read at 320 nm, d is the dilution factor and v is the final volume. 50 is the DNA extinction (absorption) coefficient where for duplex DNA; A_{260} is approximately 1 for 50 µg/ml.

Purity of DNA samples were determined by the A_{260}/A_{280} ratio. Measurements between 1.7 and 2.0 indicated pure DNA.

Ethanol Precipitation of DNA

To purify DNA from solution, it was precipitated by the addition of 1/10 of the volume of 3 M sodium acetate and 2.5 times the volume of 95% ethanol, followed by centrifugation at 10000 x g for 10 min. The DNA pellet was washed twice in 70% ethanol and air dried at room temperature for at least 15 min or dried in a vacuum desiccator for 15 min.

A-Tailing of DNA Fragments

DNA fragments that were ligated into a TA cloning vector had to be A-tailed as this vector contains T-overhangs at the cloning site. This procedure fills in any gaps at the ends of fragments as well as adding an A-overhang. 100 ng of DNA was A-tailed in a mixture of *Taq* polymerase reaction buffer, 3 mM $MgCl_2$, 0.2 mM deoxynucleotide triphosphate mix (dNTPs), 5 U *Taq* polymerase and nH_2O to a final volume to 50 µl. This was incubated at 70°C for 30 min. The fragments were cleaned from salts, *Taq* and unused dNTPs using the PCR purification kit according to manufacturer's instructions and stored at -20°C until required.

Ligation Reaction of A-Tailed DNA Fragments

A-tailed DNA fragments were ligated into pGEM T-Easy TA cloning vector according to manufacturer's instructions. Briefly, the ligation reaction consisted of ligation buffer (60 mM Tris-HCl, pH 7.8, 20 mM $MgCl_2$, 20 mM DTT, 2 mM adenosine triphosphate, 10% (v/v) PEG), 4 U of T4 DNA ligase, 50 ng of cloning vector, 1 to 3 µl of DNA fragments and dH_2O to a final volume of 10 µl. For a maximum number of transformants, this was incubated at 4°C for one overnight.

DNA was ligated into pGEM T-Easy at a proportion of 3:1 DNA insert: vector where possible, but ligation reactions were still possible at proportions up to 1:8 DNA insert: vector. The amount of DNA used in a ligation reaction was determined by the following calculation:

Amount of insert (ng) = [amount of vector (ng) x size of insert (kb) / size of vector (kb)] x insert: vector molar ratio

Ligation Reaction of Sticky Ended DNA Fragments

Fragments of DNA that were digested by various restriction endonucleases were ligated into pUC18NotI vector digested with the same enzyme. DNA fragments were ligated in a reaction

consisting of ligation buffer (66 mM Tris-HCl, pH 7.6, 150 mM NaCl, 10% PEG 6000, 6.6 mM $MgCl_2$, 10 mM DTT, 0.1 mM ATP, 0.1 mM Spermidine), an additional 0.1 mM ATP, 10 ng of pUC18NotI, 1 U of T4 DNA ligase and dH_2O to a final volume of 10 µl. This was incubated at room temperature for one overnight.

Digestion of Nucleic Acids

Restriction endonucleases were used to digest double stranded DNA (dsDNA) according to manufacturer's instructions. Typically, 5 U of restriction endonuclease was used to digest 1 µg of DNA in a reaction containing reaction buffer and dH_20 to a final volume of 10 µl. This was incubated at 37°C for 2 h.

DNase I was used to digest dsDNA and single stranded DNA (ssDNA) molecules according to manufacturer's instructions. Briefly, 1 U of DNase I was used to digest up to 1 µg of DNA in a reaction containing reaction buffer and dH_2O to a final volume of 50 µl. This was incubated at 37°C for 10 min.

RNase A was used to digest dsRNA and ssRNA molecules according to manufacturer's instructions. Briefly, 1 U was used to digest RNA in dH_2O to a final volume of 20 µl. This was incubated at 37°C for 30 to 60 min.

To determine whether DNA molecules were linear or circular, samples were digested with Exonuclease III according to manufacturer's instructions. Exonuclease III catalyses the removal of mononucleotides from the 3'-termini of duplex DNA and therefore does not digest circular DNA molecules. The enzyme also does not digest ssDNA and so protruding 3'-termini are resistant to cleavage. 10 U of exonuclease III was used to degrade DNA in a reaction containing reaction buffer and dH_2O to a final volume of 50 µl. This was incubated at 37°C for 30 min.

All nucleic acid fragments were visualised by agarose gel electrophoresis.

Polymerase Chain Reaction (PCR)

Specific regions of DNA were amplified using polymerase chain reaction (PCR). Approximately 100 ng of template DNA was used in a reaction containing reaction buffer (75 mM Tris-HCl, pH 8.8, 20 mM $(NH_4)_2SO_4$, 0.01% (v/v) Tween 20), 0.2 mM dNTPs, 3 mM

$MgCl_2$, 20 ρmol of forward primer, 20 ρmol of reverse primer, 2.5 U of *Taq* polymerase and nH_2O to a final volume of 50 µl.

The reaction was placed initially at 95°C to denature the strands of template DNA. The temperature was then reduced to allow primers to anneal to the top and bottom strands of DNA template. The annealing temperature varied according to primer length and G+C content. The temperature was then ramped up to 72°C to allow *Taq* polymerase to attach to the primers and make new strands of DNA by incorporating complementary dNTPs to the ends of primers. 72°C is the optimal temperature for *Taq* polymerase to work. The time allowed for primer extension was varied according to the length of DNA sequenced to be amplified. Generally, 1 min is required to synthesise 1 kb DNA. This cycle of denaturing, annealing and primer extension usually occurred for thirty cycles so that specific regions between the primers were amplified. Since both strands are copied during PCR, there is an exponential increase in the copies of DNA. A final incubation at 72°C for 10 min then occurred to allow *Taq* polymerase to fully extend fragments that may have begun, but would otherwise not have gone to completion. PCR products were then cooled to 4°C and stored at -20°C until further use. PCRs were carried out in thermal cyclers Techne Progene and Biometra Trio-Thermoblock for PCRs in tubes, or for PCRs in 96 well PCR plates (ABgene). PCR products were visualised by agarose gel electrophoresis.

PCR of Ribosomal RNA (rRNA) genes

Ribosomal RNA (rRNA) genes were amplified by PCR from environmental DNA to allow identification of organisms present. The small ribosomal subunit, (the 16S rRNA gene) was amplified from bacteria, archaea, korarchaeotes and nanoarchaea. The small ribosomal subunit (the 18S rRNA gene) was amplified from eukarya. Template DNA was used in PCRs as described in the previous section. Specific running conditions were used for each of these genes and are shown in Table 10. Amplification of SHOW group 16S rRNA gene required 100 ng of DNA in a reaction containing reaction buffer, 1.75 mM $MgCl_2$, 0.2 mM dNTP, 37.5 of ρmol SHOWprb, 60 of ρmol rp1, 2 U of *Taq* polymerase and nH_2O in a final volume of 50 µl.

Table 10 Conditions for PCR.

This table shows specific running conditions for PCRs to amplify 16S and 18S rRNA genes.

rRNA gene	Primers used	Running Conditions	Reference
Bacterial 16S rRNA	27Fb and rp1	95/2' (95/30", 55/40", 72/2') x 30, 72/10', 4°C soak	(McGenity *et al.* 1998)
Archaeal 16S rRNA	27Fa and rp1	As for bacterial 16S	(McGenity *et al.* 1998)
Korarchaeota 16S rRNA	K236F and K1390R	96/1.5' (96/30", 60/30", 72/1') x 10, 94/20", 60/30", 72/1' increments of 2 s per cycle) x 25, 72/10', 4°C soak	(Brunk and Eis 1998)
Nanoarchaea 16S rRNA	N3F and N1510R	As for bacterial 16S	(Huber *et al.* 2002)
SHOW 16S rRNA	SHOWprb and rp1	95/1.5', (95/1', 46/30", 72/1.5') x 30, 72/10', 4°C soak	(Burns *et al.* 2004b)
Eukaryotic 18S rRNA	EK82F and U1492R	95/2', (95/1.5', 55/1.5', 72/1.5') x 30, 72/2', 4°C soak	(DeLong 1992)
Eukaryotic 18S rRNA	1Af and 516r	94/130", (94/30", 56/45", 72/130") x 35, 72/10', 4°C soak	(Díez *et al.* 2001)

Colony PCR

PCRs were carried out using chromosomal DNA obtained from lysed bacterial or archaeal cells, without using the full nucleic acid extraction protocols outlined previously. This method saved time and reagents and allowed screening of a large number of clones at one time.

Halophilic bacteria and haloarchaea that were grown on solid or liquid media were lysed by placing either a loopful of cells from a plate or a pellet of cells from a broth in 100 µl of nH_2O. Cells were resuspended by vortexing for 30 s. Placing cells in water was often sufficient to lyse the cells, since they can only survive in high salt solutions. However, to fully lyse them and to denature cellular proteins, they were boiled in a heating block for 20 min. The cell debris was pelleted by centrifugation at 13000 x g for 10 min. 1 µl of cell lysate provided sufficient DNA template to amplify 16S rRNA genes.

To screen transformed *E. coli* cells for the size of inserts in recombinant plasmids, 2 to 3 µl of cells (grown in appropriate broth) were lysed in 20 µl of 0.25% (v/v) Tween 20. Cellular proteins were denatured by boiling in a thermal cycler (Techne) for 20 min. Cell debris was pelleted by centrifugation at 10000 x g for 10 min. 15 µl of cell lysate was used in PCR using conditions as for bacterial 16S rRNA gene (see Table 10).

PCR Purification

PCR products were cleaned from *Taq* polymerase, salts, unused dNTP and primer dimers less than 40 bp using the QIAquick PCR Purification Kit (Qiagen). Briefly, DNA was mixed with DNA binding buffer and passed through a column by centrifugation at 13000 x g for 1 min. The DNA became bound to the column and was washed with ethanol containing buffer. DNA was then eluted in 30 to 50 µl of nuclease free water or Tris-HCl, pH 8.

This kit was also used to clean DNA samples other than PCR products if required for further work. It was also used as a means of concentrating the DNA yield by eluting in a small volume.

Isothermal Amplification of Viral Nucleic Acids

To increase yield of viral nucleic acids, the GenomiPhi DNA Amplification Kit (Amersham Biosciences) was used according to manufacturer's instructions. It exponentially amplifies ssDNA or dsDNA using random hexamer primers and a Phi29 DNA polymerase. Random

primers anneal to multiple sites on the DNA template and are extended by the Phi29 DNA polymerase. During isothermal amplification (30°C), the strands being synthesised displace each other. More random primers anneal to the newly synthesised strands and this process repeats, which allows for a large yield. Using this method can produce 1 μg of DNA from as little as 1 ng of DNA template.

Briefly, DNA template was denatured in sample buffer at 95°C for 3 min. This was cooled and added to a cooled mixture of enzyme mix (Phi29 DNA polymerase and random primers) and reaction buffer (containing salts and dNTPs). Isothermal amplification occurred at 30°C for 16 to 18 h. To stop the reaction, the enzyme was heat inactivated by placing the reaction at 65°C for 10 min. The reaction was cooled to 4°C and stored at -20°C until further use. Amplified products were visualised by agarose gel electrophoresis.

A positive control for this reaction was run with lambda DNA. A negative control was run with nH_2O. Since this is an extremely sensitive technique, this procedure was carried out under sterile conditions. However, amplified products are still expected from the negative control, but it is guaranteed in the protocol that this DNA is not expected to perform in downstream processes.

Transformation by Heat Shock

Recombinant plasmids made from pGEM T-Easy and insert DNA were transformed into *E. coli* JM109 competent cells according to manufacturer's instructions. Briefly, 2 μl of ligation reaction was placed in cold sterile cloning tubes (Simport) on ice. Tubes of JM109 cells were thawed on ice for 5 min and 50 μl was added to the ligation reaction. This was gently mixed and cooled on ice for 20 min. The cells were heat shocked by placing at 42°C for 45 to 50 s *without* shaking. This was immediately returned on ice for 2 min. 950 μl of room-temperature SOC medium was immediately added to the cells and incubated at 37°C in an orbital shaker at 150 rpm for 1.5 h. This allowed the cells to recover and to allow expression of the ampicillin resistance gene. 100 μl of cells were then plated out on selective media containing ampicillin, 0.5 mM IPTG and 80 μg/ml of X-Gal and incubated at 37°C for one overnight. Only cells containing plasmids were allowed to grow since the ampicillin resistance gene is only present on this vector. Plates were then stored at 4°C until use.

Transformation by Electroporation

50 ml of overnight cultures of bacterial cells were diluted by a 1/50 in LB broth and grown to exponential phase by incubation at 37°C with shaking for 2 h. Pellets were obtained by centrifugation for 10 min at 1500 x g. To remove salts from the cells, they were washed 3 times in 2 ml of ice-cold MOPS/glycerol (1 mM MOPS, 20% glycerol) with centrifugation for 1 min at 13000 x g. The cells were then washed a further five times by addition of 1 ml of ice cold MOPS/glycerol and centrifugation for 1 min. The cells were then resuspended in 100 µl of MOPS/glycerol and stored on ice. A voltage was passed through a 50 µl solution consisting of bacterial cells and of plasmid DNA in an electroporation cuvette at 1.5 kV, 1000 Ω, 25 µF using a Bio-Rad Gene Pulser. 1 ml of SOC was immediately added and the cells transferred to an microfuge tube, which were incubated in a 37°C water bath for 1 h to allow expression of the ampicillin resistance gene. 100 µl of cells were then plated out on selective media containing ampicillin, 0.5 mM IPTG and 80 µg/ml of X-Gal and incubated overnight at 37°C to allow cells containing plasmids to grow. Plates were then stored at 4°C until use.

Blue White Screening of Transformed cells

The pGEM T-Easy vector and pUC18NotI vector uses blue white screening of colonies to select for cells containing recombinant plasmid. pGEM T-Easy is prepared by cleavage in the *lac*Z gene and the addition of terminal T. Likewise, the multi cloning site in pUC18NotI is within the *lac*Z gene. Therefore, when DNA fragments are ligated into these vectors, it disrupts the *lac*Z gene and prevents the expression of β-galactosidase.

When cells are plated on selective media containing IPTG and X-Gal, IPTG promotes expression of β-galactosidase. If the *lac*Z gene is not disrupted by the insertion of a DNA fragment, β-galactosidase cleaves the substrate X-Gal, and turns the cells blue. If the plasmid contains an insert at the *lac*Z gene, β-galactosidase is not expressed and X-Gal remains intact, leaving the cells white. The white cells are then selected in favour for recombinant plasmid DNA.

DNA Sequencing

DNA sequencing reactions were done by Lark Technologies (Essex, UK) or AGOWA (Berlin, Germany). Samples were prepared according to their specifications.

Transmission Electron Microscopy

Preparations of virus particles were visualised by transmission electron microscopy (TEM) at the Advanced Microscopy Centre at the University of Leicester. 5 µl of virus stock containing 10^6 pfu/ml was adsorbed on fresh carbon coated glow discharged pioloform grids and fixed in glutaraldehyde vapour for 3 min. Excess sample was blotted from the grid using filter paper. Salts were removed by washing with dH_2O. The sample was visualised by negative staining by addition of 1% (v/v) uranyl acetate and viewed on JEOL 1220 transmission electron microscope fitted with a SIS Megaview III digital camera system. Captured images were viewed and analysed using the Image J program available at rsb.info.nih.gov/ij/download.html.

Computer Analyses

Viewing Gels

Polyacrylamide gels and agarose gels were visualised on a White/Ultraviolet Transilluminator (Ultra Violet Products) using white light and the images captured on a Kodak EDAS 290 camera. Images were then edited on Kodak 1D version 3.5 (Scientific Imaging Systems).

Viewing DNA Sequence Data

Chromas version 2 (Technelysium Pty Ltd) was used to visualise and edit chromatograms of DNA sequences.

Homology Searches

The DNA sequences were analysed using the BLAST program (Basic Local Alignment Search Tool). BLAST (Altschul *et al.* 1990) is available at the National Centre for Biotechnology Information (NCBI), which can be accessed at www.ncbi.nlm.nih. Some were analysed using FASTA 3 search program available at the European Bioinformatics Institute (EBI), accessed at http://www.ebi.ac.uk/fasta33/. BLASTN was used to look for homologies to known gene sequences deposited in GenBank. ORF finder, BLASTP, TBLASTX and PSI-BLAST were used to search for possible homologies to known proteins, or proteins predicted by translation of the unannotated DNA sequence in GenBank.

Assembling DNA Sequences

Genomic viral DNA was assembled using the Lasergene SeqMan version 7.0 program (DNAStar).

Annotating Viral Genomes

Potential ORFs were assigned using the programs FGENESB (www.softberry.com) and GeneMark.hmm v2.5a (http://opal.biology.gatech.edu/GeneMark/index.html). tRNA sequences were identified using the tRNAscan-SE program (http://lowelab.ucsc.edu/tRNAscan-SE/) . Translations of potential ORF sequences to amino acids were made with the SeqBuilder program (DNAStar). Statistics for each of the ORFs were calculated using the program ProtParam (http://www.expasy.ch/tools/protparam.html). GC skew was calculated using the online base composition tools http://atmolbiol-tools.ca/Jie_Zheng. Inverted repeats in the DNA sequence were identified using Einverted (http://bioweb.pastuer.fr/docs/EMBOSS/einverted.html) and PALINDROME (http://bioweb.pastuer.fr/seqanal/interfaces/palindrome.html); direct repeats were located using Palim (http://tp12.pzr.unirostock.de/~moeller/palim/index.html).

Chimera Check

The entire sequencing data set was screened for chimeras using the CHIMERA_CHECK program available from the Ribosomal Database Project II (RDP) accessed at http://rdp8.cme.msu.edu/cgis/news.cgi. The output file given by this program is a histogram showing the plot of similarities of the query sequence and with those in the database for every 10^{th} position in the sequence. A chimera can be identified when the histogram produces a bell shaped curve, with the break point occurring at the peak of the curve. In addition, it lists the closest matches in the database to the two hypothetical fragments. However, results from this program do not fully determine which sequences are chimeras, but it can highlight *suspected* chimeras.

A second program called Pintail (Ashelford *et al.* 2005) was used to verify results from the CHIMER_CHECK program. It is available to download at http://www.cf.ac.uk/biosi/research/biosoft/Pintail/pintail.html. It calculates the expected percentage differences and the 95% confidence limits between the query sequence and the sequence of its nearest neighbour found through the BLASTN database. It also calculates the observed percentage differences. Depending on how similar the observed and the expected percentage differences are, the program can conclude whether there is strong evidence of a sequence anomaly and therefore if the sequence is a chimera.

Statistical Analyses

Rarefaction Curves

Rarefaction curves are used to identify when enough sampling is done on a community to determine species diversity with some level of confidence. It plots the observed number of species against the number of individuals sampled. As more clones are sampled, less unique species are likely to be observed. Once sampling reaches saturation, the graph plateaus as more clones would have to be sampled to find more unique species.

Rarefaction curves were calculated using Analytical Rarefaction version 1.3 (Holland 2003) available to download at http://www.uga.edu/strata/software/. This program calculates the expected number of species (E (S_n)) using the equation given by Hurlbert (1971):

$$E(S_n) = \sum_i \left[1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}} \right]$$

Where S is the number of species,

n is the number of individuals selected at random from a collection,

N is the total number of individuals in that sample,

N_i is the number of individuals in the *i*th species.

Designating Operational Taxonomic Units (OTU)

16S rRNA gene sequences were aligned using the program Molecular Evolutionary Genetics Analysis (MEGA) version 3.1 (Kumar *et al.* 2004) available to download at http://www.megasoftware.net/. The Jukes and Cantor substitution model for sequence alignment was used. This alignment was then used to construct a distance matrix using the same model, which was used as the input file for the program DOTUR (Schloss and Handelsman 2005) available to download at http://www.plantpath.wisc.edu/fac/joh/dotur.html. This program assigns sequences to operational taxonomic units (OTUs) for every distance level that can be used. In this study, three of the commonly used OTU definitions were employed (95%, 97% and 99%), which is equivalent to comparing different taxonomic resolutions: at the genus, species and sub-species level (Martin 2002; Horner-Devine *et al.* 2003). This was done using the furthest neighbour clustering algorithm (default setting).

Richness Estimators

The richness estimator Chao1 (Chao 1984) was calculated using DOTUR for every distance level that was used as an OTU. It is an abundance-based estimator and so estimates of unseen species are based on the number of rare species. These are classified as species with a total abundance of one (singletons) or two (doubletons). The equation is therefore given by:

$$S_{Chao1} = S_{obs} + \frac{n_1(n_1-1)}{2(n_2+1)}$$ When $n_1>0$ and $n_2\geq 0$ and when $n_1=0$ and $n_2=0$

$$S_{Chao1} = S_{obs} + \frac{n_1^2}{2n_2}$$ When $n_1=0$ and $n_2\geq 0$

Where S_{obs} is the observed number of species,

n_1 is the number of OTUs with only one sequence (singletons),

n_2 is the number of OTUs with only two sequences (doubletons).

Measuring Biodiversity

The biodiversity was measured by several diversity indices. The Simpson's Index was calculated using DOTUR for every distance level used by the equation:

$$H_{Simpson} = \frac{\sum_{i=1}^{S_{obs}} S_i(S_i-1)}{N(N-1)}$$

Where N is the total number of individuals sampled,

S_i is the number of species in the *i*th OTU,

S_{obs} is the observed number of species,

The Shannon-Weaver Index was calculated using DOTUR for every distance level that was used by the equation:

$$H_{Shannon} = -\sum_{i=1}^{S_{obs}} \frac{S_i}{N} \ln \frac{S_i}{N}$$

Species Evenness

The species evenness can be calculated from values for $H_{Shannon}$ by the equation:

$$E = \frac{H_{Shannon}}{\log N}$$

Standard Normal Deviate Equivalents (SNDE)

Raw environmental data were standardised to make the different environmental factors comparable. This was done by the following equation:

SNDE = (x – mean of the raw data) / standard deviation of the raw data

Where x is the raw data for one sampling site.

Biotic Similarity Index

Biotic similarity was calculated between communities using the Jaccard Index (Jaccard 1901), which gives a measure of incidence. It shows whether a species is present or absent from any given environment, therefore giving a measure of the beta diversity. It is dependent on three incidence counts: the number of species shared by the two samples and the number of species unique to each sample, hence, the Jaccard Index is given by the equation:

$$J_{clas} = \frac{A}{A+B+C}$$

Where A, B and C refer to counts between two matrices of biotic data, as in the table below:

	Matrix 1	
Matrix 2	Present	Absent
Present	A	B
Absent	C	-

This index was calculated using the program EstimateS (Colwell 2005) available to download at http://purl.oclc.org/estimates.

Mantel Test

To determine which factors are significantly important in driving microbial diversity, the Mantel test was used (Mantel 1967). It is a test where the variables are dissimilarity matrices summarising pairwise similarities among samples. The simple Mantel test considers two matrices; it tests if the distances in matrix A correlate with the distances in matrix B. The null hypothesis states that distances in matrix A are independent of the distances in matrix B. In the context of this study, the simple Mantel test will show if biotic similarity correlates with environmental factors such as pH, or to geographic distance, or to both.

In the simple Mantel test, r is the correlation value; positive or negative values reflect the type of relationship between the two matrices. p is the probability associated with r. Since the variables of a distance matrix are not independent, the Mantel test is calculated via permutation procedures, where the values in the matrices are randomly rearranged. Values for Mantel were recomputed for these permuted matrices. Many iterations were made to generate a distribution curve. 5040 iterations were made since a small dataset was used. At 5000 iterations, values of p are significant if it is less than 0.05; values greater than 0.05 indicates that the null hypothesis applies.

The simple Mantel test was carried out using the program zt (Bonnet and Van der Peer 2002) available to download at http://www.psb.ugent.be/~erbon/mantel/. It is given by the equation:

$$r = \frac{1}{N-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\frac{(A_{ij} - \bar{A})}{S_A} \right] \left[\frac{B_{ij} - \bar{B}}{S_B} \right]$$

Where N is the number of elements in the lower or upper triangular part of the matrix,

$\bar{A}$ is the mean for A variables measured and locations i and j,

S_A is the standard deviation of A variables

Construction of Phylogenetic Trees

Sequences were aligned using MEGA version 3.1. Phylogenetic analysis was also carried out using the same program using the Jukes and Cantor nucleotide substitution model for sequence alignment and the Neighbour-Joining method of tree inference. The support for each

node was determined by assembling a consensus tree of 500 bootstrap replicates using the same phylogenetic settings.

Principal Component Analysis (PCA)

Newick formatted trees (phylogenetic trees in a computer readable format) were constructed using MEGA 3.1. These trees were used as input files for the program UniFrac (Lozupone and Knight 2005) available at www.bmf.colorado.edu/unifrac/index.psp, which was used for Principal Component Analysis (PCA). This program uses a distance matrix to plot samples in space. The first principal component (PC1) describes the most variation possible, and the subsequent principal components (PC2, PC3, etc) explain the remaining variation not explained by the previous component or components. If the data is influenced by many independent factors, such as environmental factors, then only the first three components are considered. Plots of these components often illustrate the clustering of samples in space. The closer the samples are, the more similar they are in terms of biotic composition. The axes in the resulting plot are not related to any environmental factors since the input file are sequences only. Therefore, the factors influencing these similarities and differences must be later predicted.

Nucleotide Diversity

The nucleotide diversity was calculated using MEGA 3.1. 16S rRNA gene sequences were aligned and grouped according to the sampling site they came from. The equation used to calculate the mean nucleotide diversity for the entire community is:

$$\pi_T = \frac{q}{q-1}\sum_{i=1}^{q}\sum_{j=1}^{q}\bar{x}_i\bar{x}_j d_{ij}$$

where x_i is the estimate of average frequency of the i^{th} allele in the entire community,
q is the number of different sequences in the entire sample.

Lineage-per-time Plots

These were calculated using DOTUR assuming a molecular clock. It determines the number of lineages observed over evolutionary distance from a common ancestor (Martin 2002). Depending on the curve of the plot, conclusions were made on the composition of the microbial community. Constant birth and death rates result in a straight exponential plot. A

concave plot indicates an overabundance of highly divergent lineages, while a convex plot indicates an overabundance of closely related lineages (Martin 2002).

Coefficient of Differentiation (Fst)

16S rRNA gene sequences were aligned and grouped according to the sampling site that they came from. Pairs of environments were then selected and the Fst value for those two communities was calculated using MEGA 3.1. This was done for every possible combination of pairs of groups. It is given by the equation:

$$F_{ST} = \delta_{ST} / \pi_T$$

where δ_{ST} is the estimate of interpopulation diversity,

π_T is the mean diversity for the entire population

Fst compares the total diversity within each community to the total diversity of the communities combined. When the level of diversity in each community is equal to the level of diversity of the two communities combined, then the Fst value is zero. This shows that the genetic diversity overlaps, even though the species diversity does not (Martin 2002). A value of one indicates that the degree of differentiation for the two environments is different. The statistical significance of the coefficient of differentiation was calculated by randomly assigning sequences to communites and calculating the Fst for 1000 replicates. Values < 0.05 were considered significant (Martin 2002).

Parsimony Test (P test)

This is an alternative approach to using Fst (Martin 2002), which also uses phylogeny to determine diversity between two communities. It uses parsimony to estimate the minimum number of changes (parsimony score) to explain the observed phylogeny for two environments given the treatments for each of the 16S rRNA gene sequences present. ('Treatments' here means whether the sequence is from community A or community B). The significance of the P test value is tested by randomising the relationships between the sequences, while keeping the treatment of the sequences fixed. The number of changes (parsimony scores) for these randomly joined trees can be plotted with the frequency of trees that had that score, which results in a normal distribution curve. Significance is established if the observed number of transitions for the actual tree falls in the normal distribution where

probability is < 0.05. This indicates that the two communities contain distinct diversities (Martin 2002).

Two programs were used to determine P test values for the environments. UniFrac is an online tool available, while TreeClimber (Schloss and Handelsman 2006) is available to download at www.plantpath.wisc.edu/fac/joh/treeclimber.html. The program TreeClimber constructs 1000 randomly joining trees, while the program UniFrac is only capable of 100. 16S rRNA gene sequences were aligned in MEGA 3.1 and assembled into phylogenetic trees using the settings as previously described. This tree was then used in UniFrac, which determined p values and its significance for all combinations of pairs of environments. To use TreeClimber, sequences from pairs of environments were aligned and subsequent trees established, also using MEGA 3.1. Trees were then used in TreeClimber, which uses Fitch's algorithm to determine parsimony scores (Fitch 1971). Frequency curves, cumulative random probability and the score for each tree were determined.

Used in conjunction with values for Fst, conclusions can be made about the differentiation of diversity between environments (Martin 2002).

RESULTS: EXPEDITIONS

Descriptions of Sampling Sites and Naming of Hot Spring Samples

All hot spring environmental samples were collected from the Yunnan Province in south China. The expedition was conducted in March 2003. Hot spring water, sediment and microbial mats were collected at these sites. A summary of the samples collected are found in Table 11.

Rehai sampling area, whose coordinates are 98° 26' E, 24° 57' N, is located 10.5 km south of Tengchong and has an elevation on 1520 m. All samples from this area were named with the prefix 'TC.' The area had been developed into a national scenic geological park for tourists to visit. It is surrounded by hills, with the Zaotang River running through and covers an area of 9 km^2. Many of the thermal sites have been managed, with paved paths leading to each of them and have been given names such as 'Big Boiler' and 'Drumbeat Spring.' The local people use some of the springs to cook peanuts and eggs to sell to tourists, while many tourists use the hot spring water for hydrotherapy as it is believed that the water contains minerals and trace elements to aid circulation and metabolism, and to cure various skin diseases.

Sampling sites TC1 and TC2 were located in Drumbeat Spring, which consisted of a main source, with a flow of 1.2 – 1.5 L/s. The water from this source cascaded over rocks into a shallow rock lined pool. Sampling site TC1 was located in the lower pool, while TC2 was located in a drainage channel containing a dark green microbial mat.

Sampling site TC3 was located in Pearl Spring, which was an enclosed pool with a flow discharge of 0.2 L/s. It has been formed by the condensation of numerous fumaroles, which gave off a slight sulphurous odour.

Sampling sites TC4 to TC8 were located in a 'Hydrothermal Outbreak Pool, Scientific Research area.' This area is at the base of a cliff face above the river. It was given its name because it was reported that in 1974 and 1978, ground water overheated and sent a mixture of water vapour and soil spurting out to a height of 10 m. A white muddy sheet was then formed over the slopes. The main source for this site was a 50 cm pool with a discharge of less than 0.11 L/s located on an upper slope. Water drained from this main source into several secondary pools on lower slopes in the area. They eventually led to a large drainage channel,

which drained into a large wet area at the base of the region, which was covered in a green and yellow microbial mat. Sampling sites TC4 and TC6 were located in different points in this area that were covered in green algal mats. TC7 was located in a runoff stream from TC5. Sampling site TC8 was located in a small hot stream running down a slope where a sample of white filamentous material (streamers) was collected (Figure 11). These streamers were firmly attached to the rocks and sediment.

Figure 11 Sampling Site TC8.

White streamers (indicated by the arrow) were collected at sampling site TC8 in Rehai geothermal area.

Sampling site TC9 was located at Lion's Head, where the river flowed through a gorge. A pool of grey sediment opposite Lion's Head was sampled.

Further along the river was a 7 m high waterfall. At the base were three fumaroles spurting hot water and steam, which were called Frog's Mouth. These were covered in green microbial mass which was collected down to the rock face, providing sampling site TC10.

TC11 was located in a shallow pool located above the waterfall, which was covered in a microbial mat. Unfortunately, the spring water was diverted for domestic use and the area had a lot of debris.

TC12 to TC14 were located in an area below a solfatara cliff, which was 20 m north east from the other sampling sites. The rock here was white and yellowish, which is indicative of high acidity. Three small acidic pools that produced a gas were sampled.

Lang Pu is a village approximately 20 km south west of Tengchong in a broad valley, whose coordinates are 98° 23' 20" E, 24° 54' 30" N and has an elevation of 1119 m. Samples taken from here were named with the prefix 'LP.' It is clear that the area was being developed for human use as a thermal spa complex, probably built around 70 to 80 years ago. However, the project was not completed, leaving the site overgrown and covered with derelict buildings.

Sampling sites LP1 and LP2 were located in a small pool and associated drainage channel found on the lower slopes of the complex. Both were covered in a dark green microbial mat. Sampling sites LP3 and LP4 were located further up the complex, where a large deep basin of steaming and bubbling water was found amongst the buildings. It measured 10 m x 8 m and was several metres deep. The complex is in fact built around this pool, which appeared to be the source for the entire area. LP3 was located in the middle of this basin. Three sides of the basin were surrounded by a 3 m high wall. In one of these walls, a single metal pipe seeped hot water. This was where sampling site LP4 was located. This site contained a 3 cm thick laminated mat with clear green, brown and red striations (Figure 12).

Figure 12 Laminated mat at Sampling Site LP4.

Green, brown and red striations can be clearly seen in this sample collected at Lang Pu.

Table 11 Catalogue of Hot Spring Samples.

This is a catalogue of samples taken from each site by the Leicester team. The samples in bold were used for further analysis in molecular work.

Sampling Site	Samples Collected*
Rehai	
TC1	Biomass from filtered water
TC2	Dark green, slimy microbial mat
TC3	Biomass from filtered water
TC4	Microbial mat
TC5	Biomass from filtered water
TC6	Dark green microbial mat
TC7	-
TC8	**White streamers**
TC9	Biomass from filtered water
TC10	Microbial mat
TC11	White streamers
TC12	-
TC13	-
TC14	-
Lang Pu	
LP1	-
LP2	Dark green microbial mat
LP3	Biomass from filtered water
LP4	**Dark green laminated mat with red and brown striations**

* '-' in the table means that samples were taken at that sampling site by other groups on the expedition.

Temperature, pH and Conductivity of the Hot Springs

The graphs in Figure 13 show the conditions at various sites at the time of sampling, which may differ to previous reports, but this may be due to seismic activity and rapid burial and erosion in active volcano settings. The graph in the top panel shows temperatures of the hot springs ranging between 52°C (TC2) and 94°C (TC3), which is near the boiling point of water (only 96°C at this altitude). The middle panel shows that six of the sites (TC1, 2, 4, 9, 10, 11 and LP1 to 4) are alkaline, with TC1 being the most alkaline with a pH of 9. TC6 is the only neutral site sampled. Four of the sites (TC3, 5, 7 and 8) are slightly acidic, while TC12, 13 and 14 are very acidic with a pH of 2. The latter three sites were at the solfatara cliff. The conductivities of the sites were very low, ranging only between 0.7 and 4.8 µS/cm, which indicates low salinity.

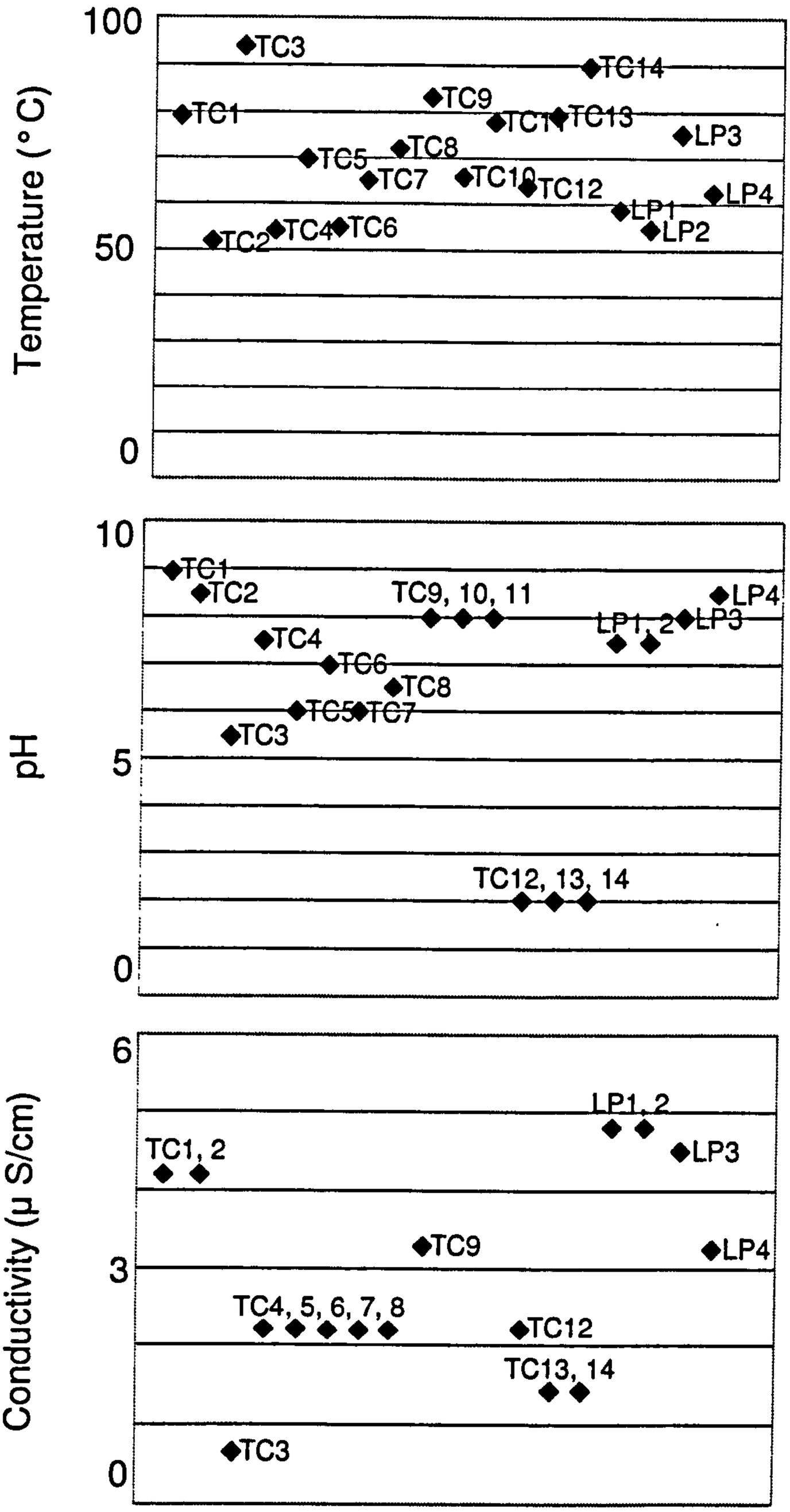


Figure 13 Hot Spring Conditions.

These graphs show the temperature (top panel), pH (middle panel) and conductivity (bottom panel) in the hot springs.

Descriptions of Sampling Sites and Naming of Salt Lake Samples

All salt lake environmental samples were collected from Inner Mongolia in north west China. Figure 14 shows the locations of seven lakes that were sampled. The nearby towns are also indicated. This expedition was conducted in September 2003, which is the end of summer in Inner Mongolia. The conditions at the time were mild, but with extreme winds coming from the Gobi Desert. Lakes Ejinnor, Bagaejinnor, the unnamed lake and Shangmatala were located in lush grassland. The terrain around lakes Chahannor, Chagannor and Erliannor was dry desert scrub covered in sedges and heathers. A summary of the GPS coordinates of these lakes are found in Figure 14 (top panel). Lake water, sediment, salt crystals and microbial mass from filtered lake water were collected from various points at these sites. A summary of the samples collected are found in Table 12.

Lake Ejinnor, whose GPS coordinates are N45 14 4.52 E116 32 4.77, was located north of Qog Ul. The word 'Ejinnor' means 'The Mother Lake.' It was a large shallow lake with evaporating lagoons on the eastern side of the main body of water. It was noted that there was a smell emitted from the lake, probably due to microbial activity. To the side were some large salterns that contained red water and gravel like salt crystals that appeared orange. All sampling sites from this lake were named with the prefix 'EJ.' A sample of biomass was taken from EJ3 in one of the red salterns (Figure 15), while EJ4 was located in a nearby drainage channel.

Sampling Site	GSP Coordinates
1. Lake Ejinnor	N45 14 4.52 E116 32 4.77
2. Lake Bagaejinnor	N45 08 527 E116 36 1.67
3. Unnamed Lake (West of Xilin Hot)	N47 55 355 E115 36 757
4. Lake Shangmatala	N43 22 751 E114 01 361
6. Lake Chagannor	N43 16 131 E112 55 636
7. Lake Erliannor	N43 44 426 E112 02 081

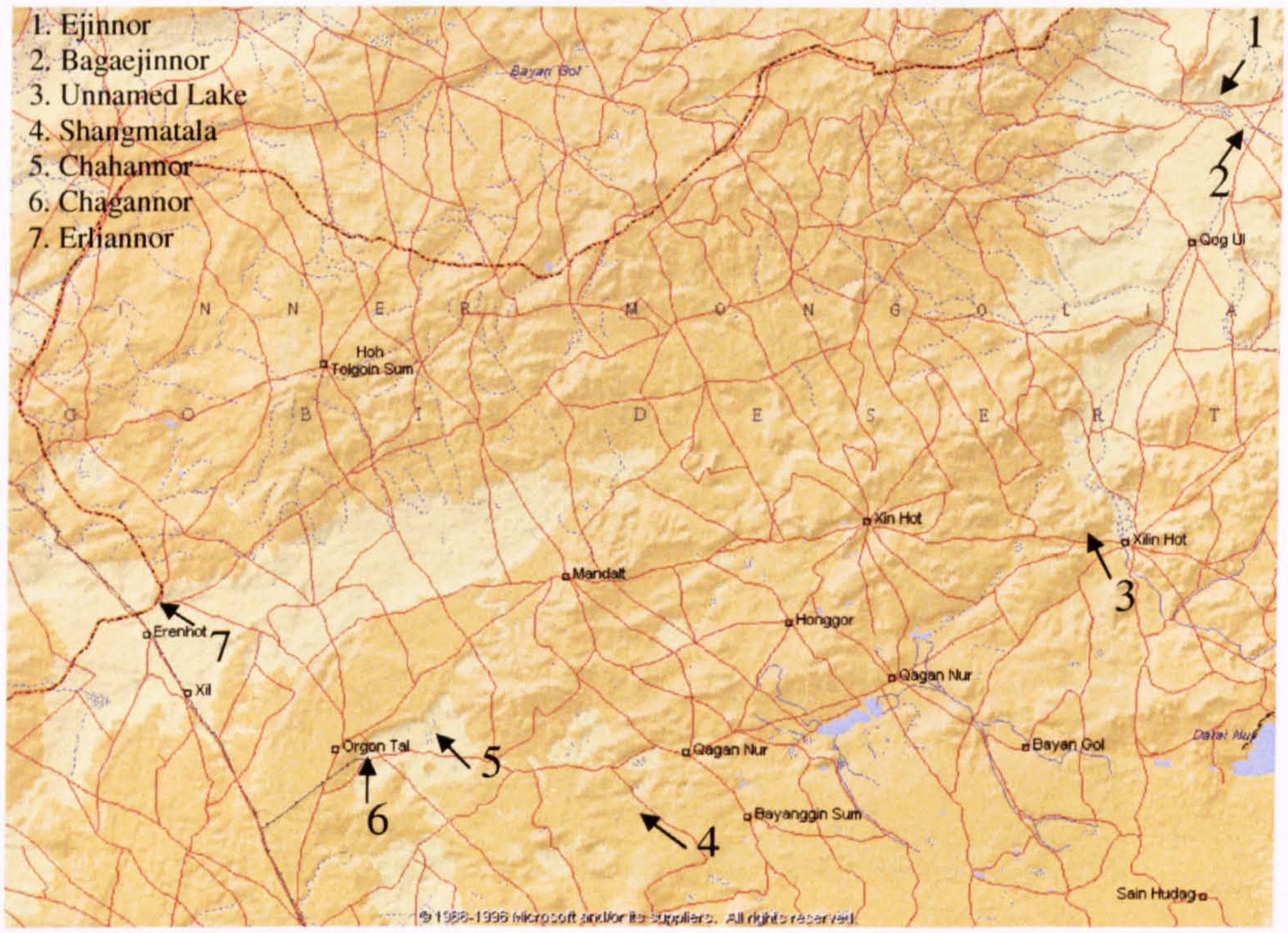


Figure 14 Locations of the Salt Lakes, Inner Mongolia.

Summary of the GPS coordinates for the salt lakes are indicated in the top panel.

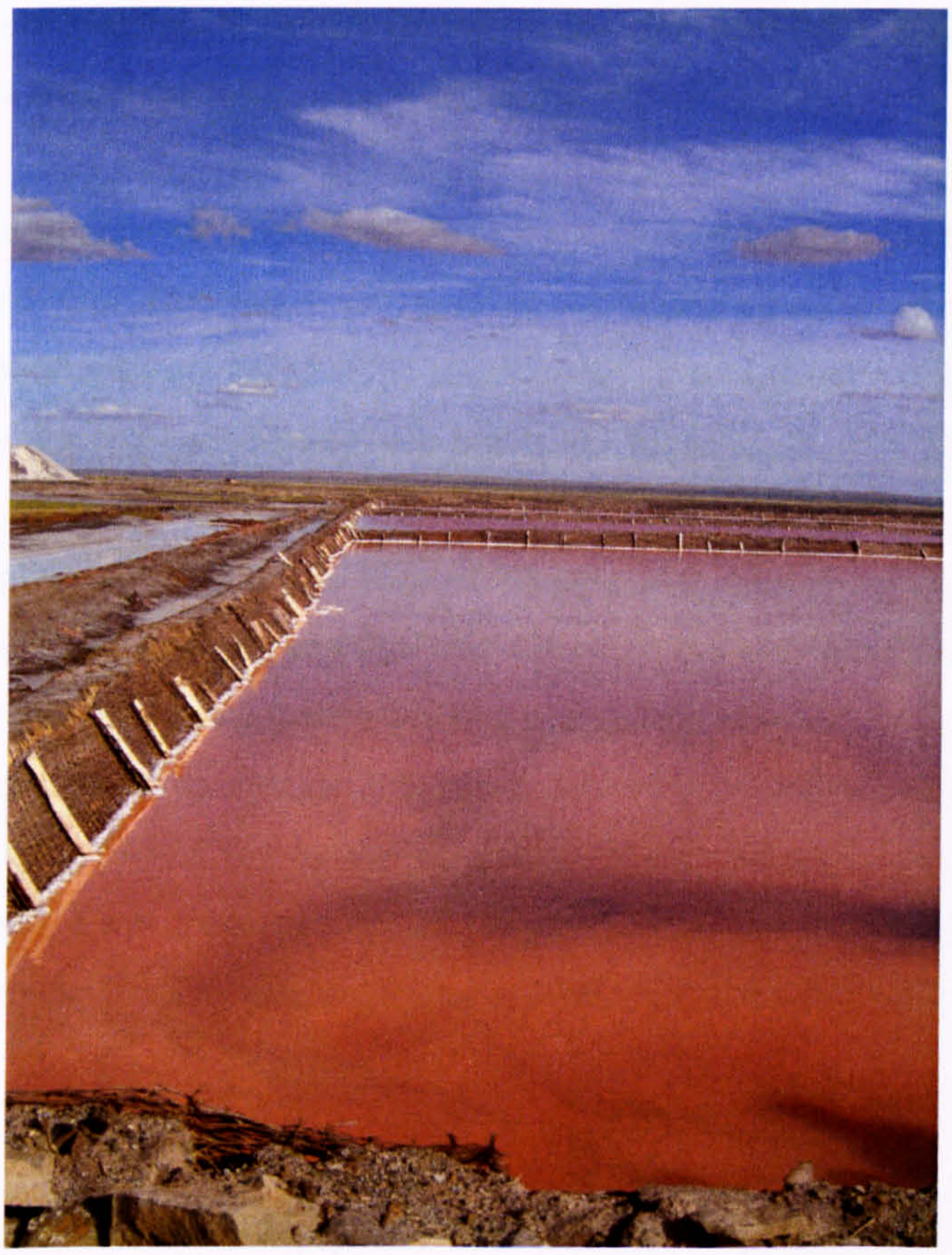


Figure 15 Sampling Site EJ3.

Samples were collected from this red saltern at Lake Ejinnor.

Lake Bagaejinnor, also north of the town Qog Ul was 25 to 30 miles south of Lake Ejinnor. The lake was located in the middle of grassland. However, the lake itself had evaporated over the summer, exposing large expanses of mud flats. The lake had therefore been reduced to a number of small pools and lagoons surrounded by mud that was very soft and dangerous to walk on. The GPS coordinates for this lake are N45 08 527 E116 36 1.67. The water in the pools were clear and colourless, but appeared pink due to the coloration of the surrounding salt crystals (Figure 16). Sampling sites from this lake were named with the prefix 'BJ.' It was apparent that there was much animal contamination due to nearby farmers allowing their livestock to roam around the lake.

An unnamed lake (it does not appear on any maps) was sampled. It is located north west of a major town called Xilin Hot, hence all sampling sites from this lake were given the prefix 'XH.' Its GPS coordinates are N47 55 355 E115 36 757. It was situated close to an abandoned soda works, with a scarp on the south side and a hill to the north. The lake was divided by several causeways. There was also a shallow lagoon cut off from the rest of the lake, which was where sampling site XH2 was located (Figure 17). The water here was clear and colourless and had many brine shrimp that were collected and later placed in buffer for further analysis. The lagoon itself had a soda crust that could support the weight of an individual. It also had thick, course surface crystals. Just below the surface was a thin aerobic layer, only a few millimetres thick, which lay on top of a thick, black and spongy anaerobic sediment. Leading from the lake was a drainage channel that connected into a 15 cm deep pool of green water (147 m from the lake), where sampling site XH3 was located (Figure 17).

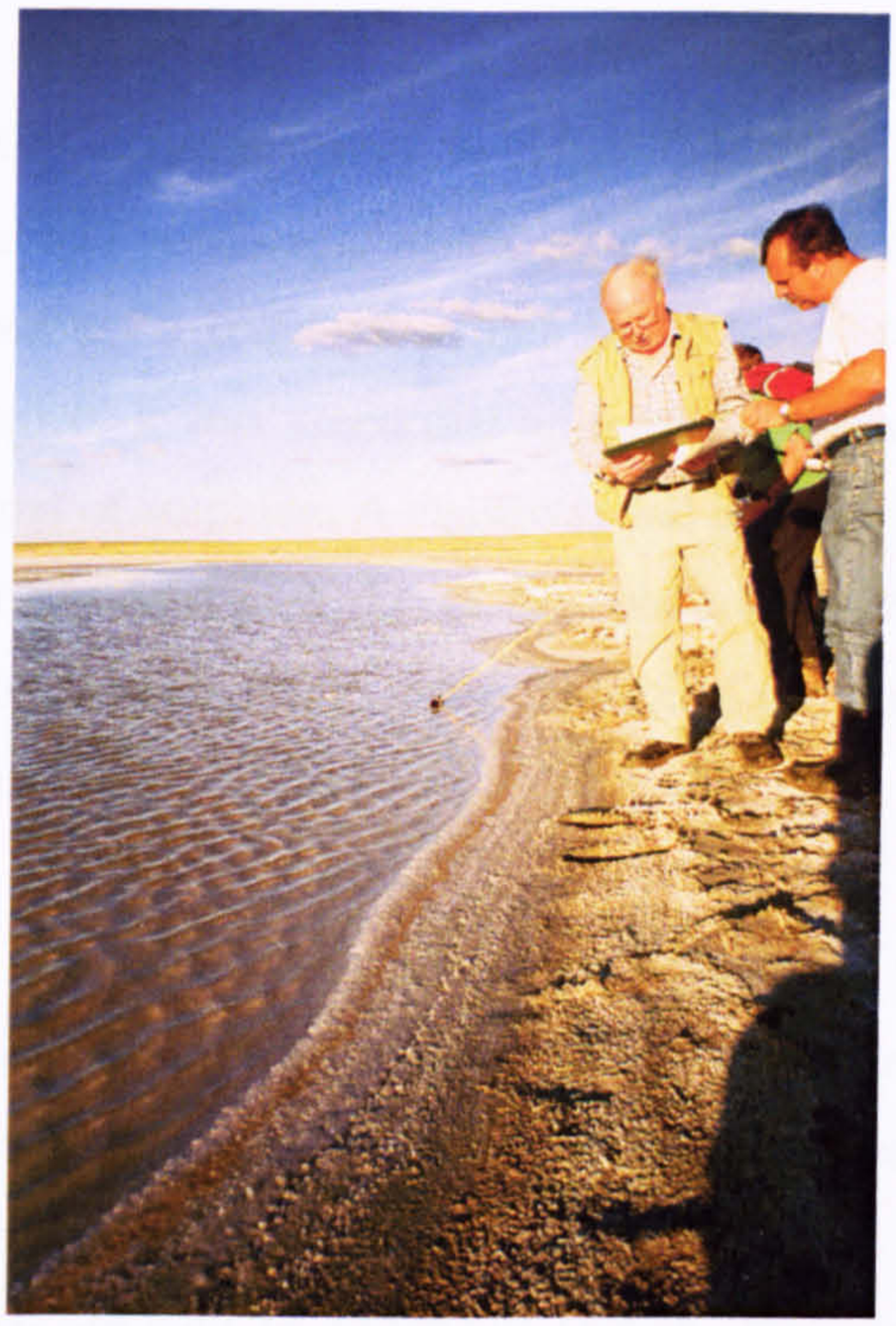
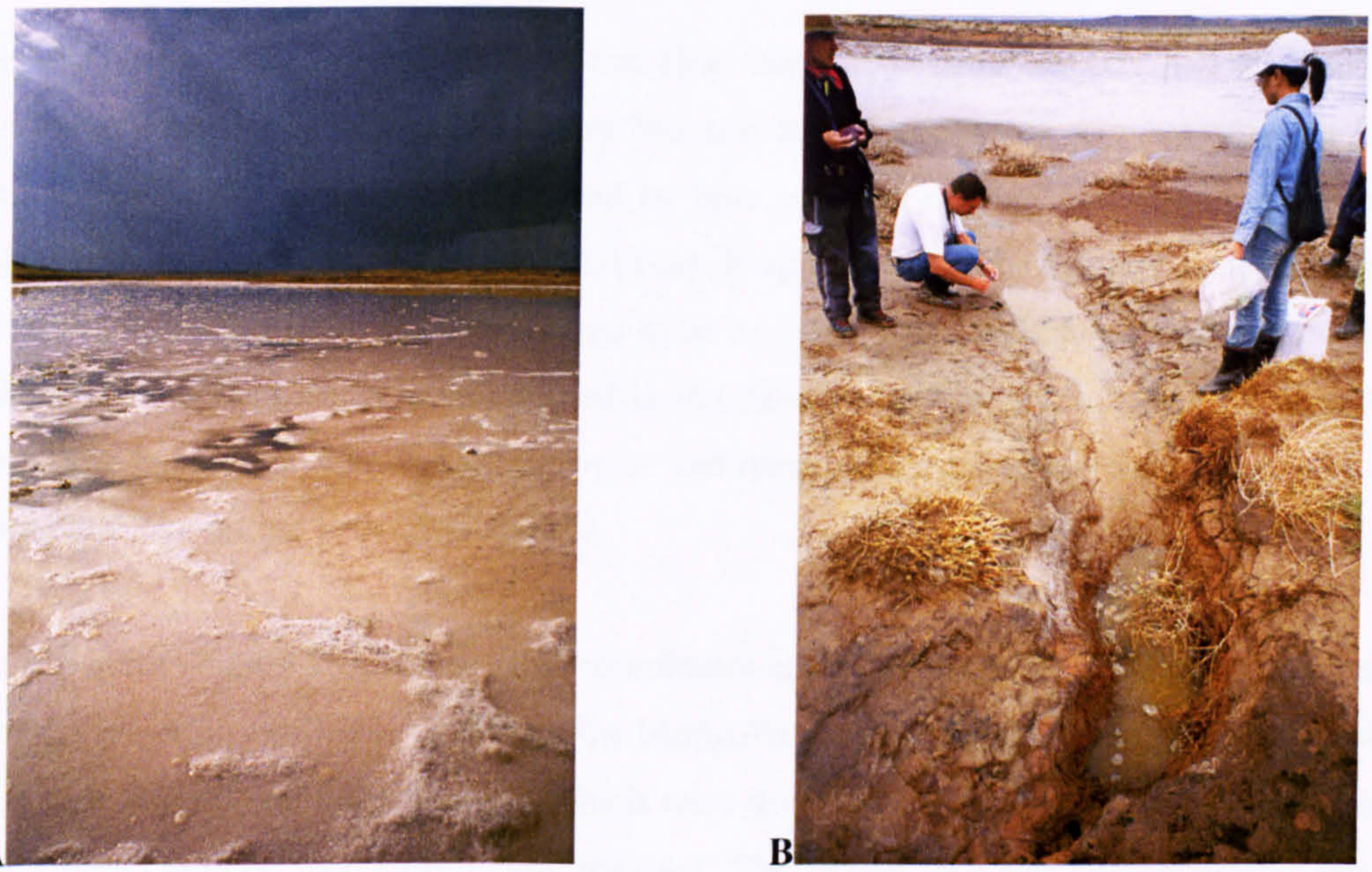


Figure 16 Sampling Site BJ1.

Sampling site BJ1 was located here in the main body of water from Lake Bagaejinnor. The surrounding salt crust had pink striations.



A B

Figure 17 Sampling Sites XH2 and XH3.

Sampling site XH2 was located in the lagoon cut off from the main body of water at the unnamed lake near Xilin Hot (A). Sampling site XH3 was located in a small pool connected to the lake by a drainage channel (B).

Lake Shangmatala (Figure 18) was located in a shallow basin surrounded by hills at an elevation of 987 m and GPS coordinates of N43 22 751 E114 01 361. The lake was surrounded by lush grassland and vegetation, which grew almost up to the water's edge. The soil nearest the lake appeared to be soda soil, which had a layer of lichen growing on the surface. This layer was sampled. It was noted that an unpleasant smelling gas was emitted, which was probably some kind of sulphide. A causeway led directly into the lake. The water in the lake was shallow and appeared pink due to the coloration of the salt crystals. On the east side, rod shaped crystals were observed on the surface of the mud. Sampling sites from this lake were given the prefix 'SH.'

Lake Chagannor (Figure 19) was situated near the largest soda works in China, 120 km south of the major town, Mandulatu. It has its own coal fired kilns and railway to transport the soda. Piles of soda were visible; hence it was clear that this was a soda lake. It was surrounded by desert scrub, but sampling took place on the south side where it was muddy with a layer of fine salt. The GPS coordinates are N43 16 131 E112 55 636. The lake water appeared to be green and the mud was grey and viscous. The mud was also very soft, which made sampling conditions unsafe. Sampling sites here were named with the prefix 'CG.'

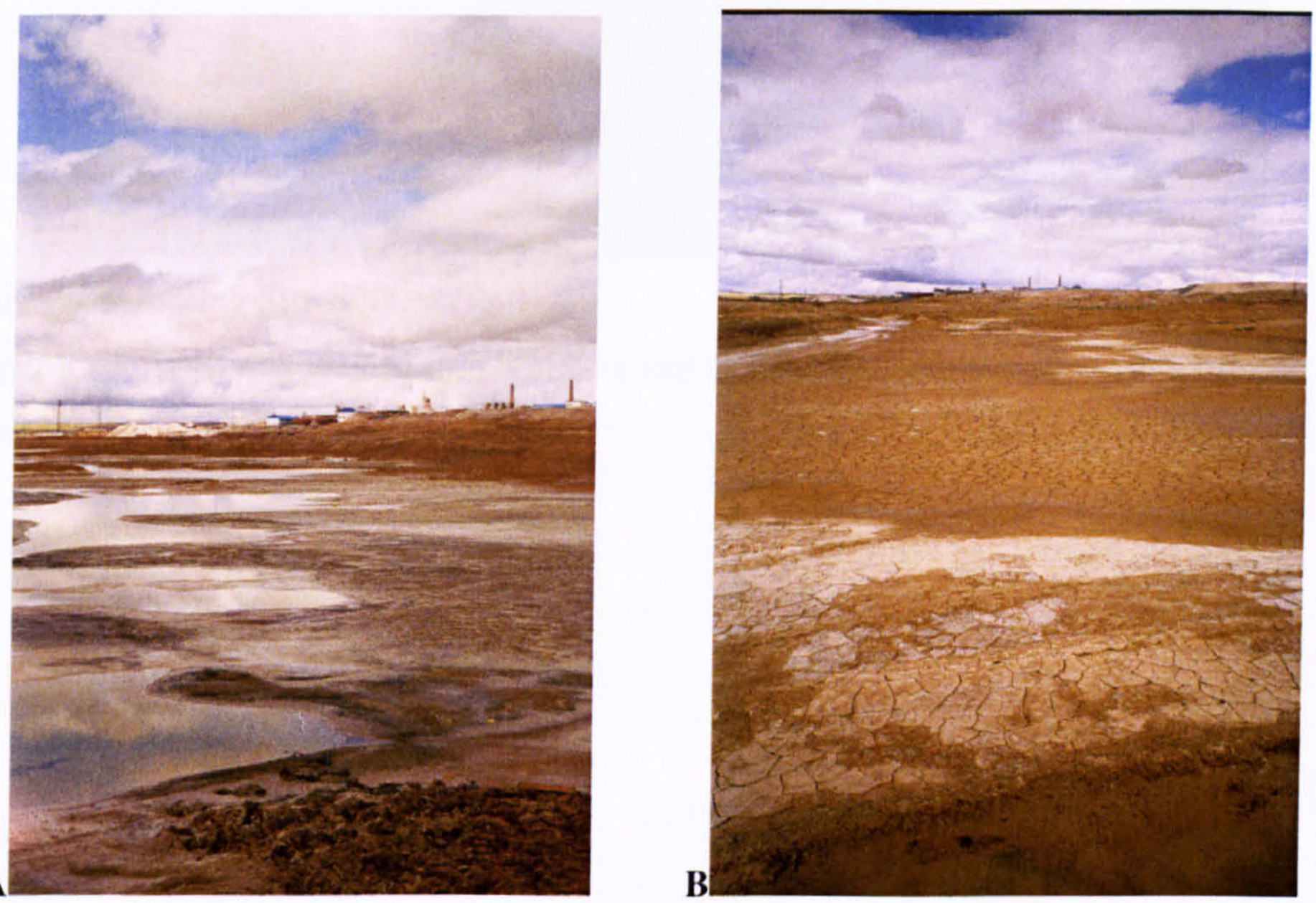
Lake Chahannor was not sampled as it was clear that the original salt lake had disappeared, since the conductivity readings were very low and the sediment was atypical of a salt lake, consisting only of clay. It was surrounded by hills and was situated in a large basin, where scrub terrain gradually converted into mud flats. It appeared that the original salt lake, which was last sampled ten years ago and known to be a concentrated soda lake, had been covered by thick sediment, probably consisting of desert sand and silt blown across from the Gobi Desert. The basin then filled with rain water and runoff to form a fresh water lake, with the original salt lake buried below the surface.

Lake Erliannor (Figure 20), whose GSP coordinates are N43 44 426 E112 02 081 was located north of a major city called Eren Hot on the Mongolian border and the Trans-Siberian railway. It was spilt into several large salterns, which were used to harvest salt and soda; hence many piles of soda could be seen surrounding the lake. The natural lake was unrecognisable due to the extensive development of these salterns. There is approximately 40 km^2 of underground soda, which is harvested in the colder months. In the summer, microbial blooms turn the salterns red and salt is harvested. At the time of sampling, all the salterns were white and the

water was clear and colourless. A few of the younger salterns had a thin brown biofilm swept to the edges by the wind, which was collected for analysis. Sampling sites here were given the prefix 'EN.'

Figure 18 Sampling Site SH1.

Sampling site SH1 was located in the main body of water at Lake Shangmatala. It was the only lake observed where vegetation grew up to the water's edge.



A B

Figure 19 Sampling Site CG1.

Lake Chagannor is situated near China's largest soda factory (A). Sampling site CG1 was located in the muddy waters in this lake. Characteristic 'cracking' in the surrounding sediment was observed (B).

Figure 20 Sampling Site EN1.

Sampling site EN1 was located in a white saltern at Lake Erliannor, where salt and soda are harvested.

Table 12 Catalogue of Salt Lake Samples.

This is a catalogue of samples taken from each site by the Leicester team. The samples in bold were used in for both molecular work and culturing work, samples in italics were used in culturing work only.

Sampling Site	Samples Collected*
Lake Ejinnor	
EJ1	-
EJ2	-
EJ3	**Biomass from filtered water**; *salt crystals; black sediment and brine; black brine*
EJ4	Microbial mat
Lake Bagaejinnor	
BJ1	**Biomass from filtered water**; *salt crystals; sediment and water; orange/ brown surface material*
Unnamed lake	
XH1	Salt crust; sediment and water; marginal surface salt with possible green algae
XH2	**Biomass from filtered water**; *salt; brine*; brine shrimp
XH3	Biomass from filtered water
Lake Shangmatala	
SH1	**Biomass from filtered water**; *salt; brine; orange salt*; soda soil; soda efflorescence
Lake Chagannor	
CG1	**Biomass from filtered water**; *salt crust; brine*
Lake Erliannor	
EN1	**Biomass from filtered water**; *salt crust; brine; surface salt and mud*
EN2	Brown biofilm

*All molecular characterisation of microbial populations was carried out on biomass collected from filtered water in order to compare the microbial communities in the water column of the different lakes. These samples were also used for cultivation methods, as well as the additional samples (e.g. salt crust, sediment) indicated in italics. Conductivity, pH and temperature readings, and chemical analysis was also carried out on the salt lake water (see later). All biogeography analysis was carried out on sequences obtained from molecular analysis of the salt lake water (see later).

'-' in the table means that samples were taken at that sampling site by other groups on the expedition.

Temperature, pH and Conductivity of the Salt Lakes

The graphs in Figure 21 show the conditions at various sites at the time of sampling. They clearly differ to previous reports (Zheng *et al.* 1992) but this may be due to seasonal differences: hot weather can encourage microbial blooms or heavy rainfall dilute the lakes to alter the pH and salinity; or it may reflect the gradual evolution of each environment as witnessed for lake Chahannor.

The temperature of the lakes ranged between 17.1°C (CG1) and 32°C (EJ1) with the majority in the 20 s. All of the lakes were slightly alkaline, however, three samples from Lake Ejinnor (EJ2, 3 and 4) and one sample from Lake Erliannor (EN2) was only slightly above neutral. The highly alkaline samples were from the small pool by the unnamed lake (XH3) and from Lake Chagannor (CG1), which had the highest pH of 10.5. Conductivities for the lakes were very high, suggesting that these are all saline environments. To obtain an accurate reading, serial dilutions of the lake water had to be carried out first. Sample EJ1 had the lowest conductivity of 166 mS/cm, while XH1 had the highest at 704 mS/cm. The majority of the lakes had conductivities between 400 and 500 mS/cm.

These results suggest that Lake Chagannor is the only less hypersaline soda lake. The small pool at the unnamed lake (XH3) is also less hypersaline and alkaline. Lakes Bagaejinnor, Shangmatala and the unnamed lake are near neutral hypersaline lakes. The salterns cut off from the main body of water at Lake Ejinnor and the salterns at Lake Erliannor are also near neutral and hypersaline.

Total ionic concentrations can be calculated from conductivity measurements (see Methods). The total ionic concentrations for the sampling sites are as follows: BJ1 contains 474 g/L, CG1 contains 202 g/L, EJ3 contains 464 g/L, EN1 contains 482 g/L, SH1 contains 487 g/L, XH2 contains 463 g/L and XH3 contains 287 g/L.

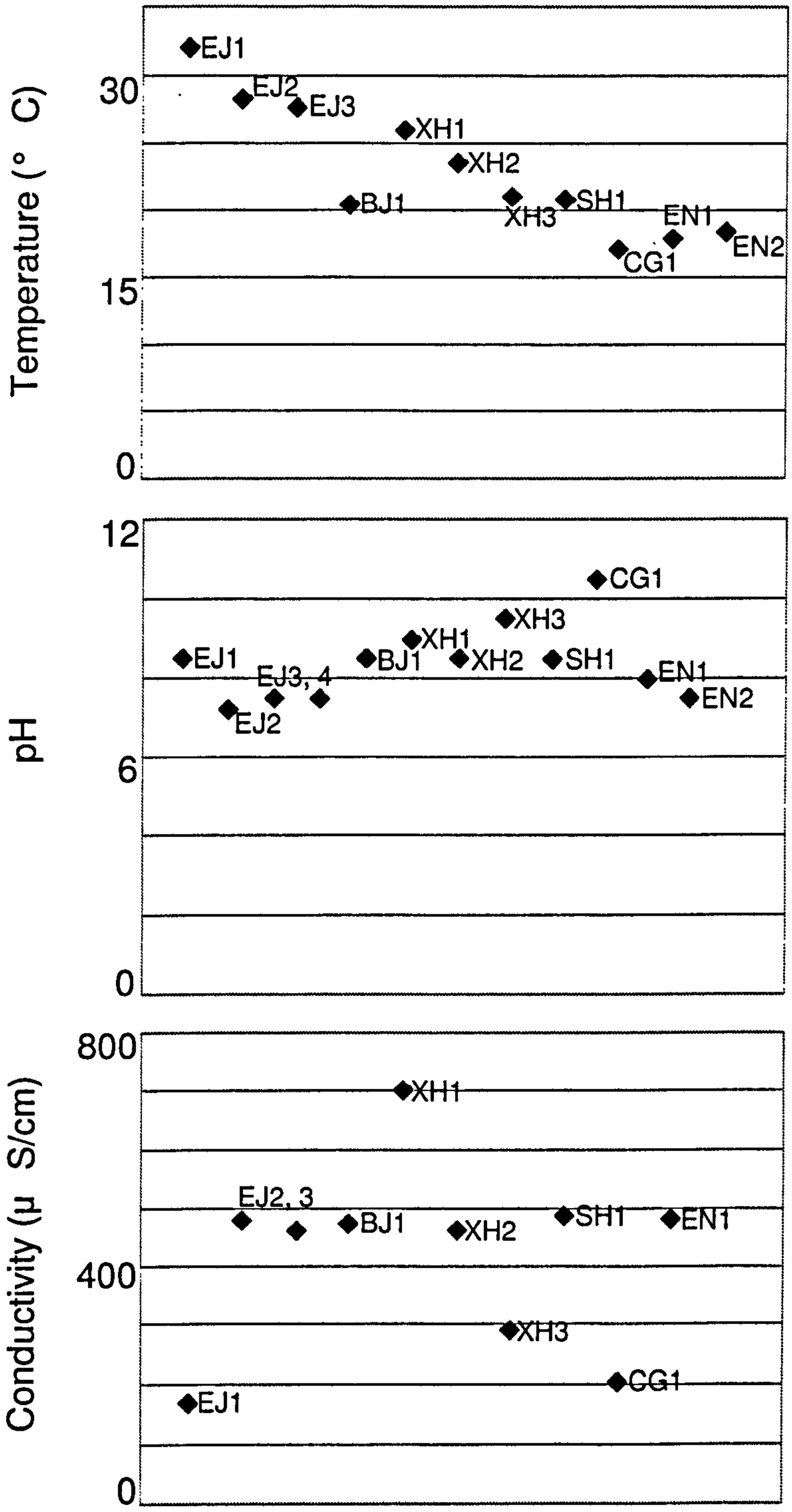


Figure 21 Salt Lake Conditions.

These graphs show the temperature (top panel), pH (middle panel) and conductivity (bottom panel) of the salt lakes.

Sampling

Water filtration

Water was filtered to collect biomass from the water column. Brine or hot spring water was sampled at a distance in 250 ml stainless steel beakers suspended on the end of a 1 m pole. Water was filtered on site through sterile 0.45 μm membrane filters (Millipore) in a 250 ml capacity polycarbonate filter unit (Sartorius) using a Nalgene hand pump, which produced a vacuum of 40-50 cm Hg under field conditions. Water was processed in this way until flow stopped, which suggested that sufficient biomass was captured on the filter. Membrane filters were removed from the apparatus using sterile tweezers and placed immediately in cold sterile stabilisation buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 2 M NaCl) and agitated to resuspend the cells. This was immediately placed on ice until further processing.

Collection of Environmental Samples

Sediment, salt crusts and algal mats were collected using sterile tongue depressors or sterile Pasteur pipettes (International Scientific Supplies, Ltd) and placed into sterile plastic bags (Whirlpacks) or sterile falcon tubes (Corning Incorporated). Filtered waters (see previous section) were also placed in sterile falcon tubes. All samples were placed immediately on ice until they could be stored at -20°C, usually within 6 hours of collection.

RESULTS: CHARACTERISING MICROBIAL POPULATIONS BY MOLECULAR METHODS

Introduction

The aim of this section was to characterise the archaeal and bacterial populations in the microbial mats in the hot springs and the water column in the salt lakes by analysing the 16S rRNA gene sequences present in community DNA. This analysis was carried out for samples at sites TC8 (white streamers) and LP4 (laminated mat) at the hot springs, and BJ1, CG1, EJ3, EN1, SH1 and XH2 at the salt lakes.

Molecular characterisation involved cloning PCR amplified 16S rRNA gene sequences from community DNA to construct 16S rRNA clone libraries. The clones were screened for inserts and unique inserts were selected by the patterns they produce upon digestion with a restriction endonuclease. These were sent for sequencing.

In addition, molecular techniques were used to look for evidence of *Eukarya*, *Korarchaeota*, nanoarchaea and the Square Haloarchaeon of Walsby (SHOW) group in these environments using specific primers designed against the 18S rRNA gene sequence (for *Eukarya*) or the 16S rRNA gene sequence for nanoarchaea, *Korarchaeota* and the SHOW group. Finally, new bacterial primers were designed to see if additional and different phylotypes could be identified in the community to those already amplified by the 'universal' primers.

Assessment of Genomic DNA Extraction Methods

Two methods were employed to extract total genomic DNA; the Amersham kit and the Mo Bio kit (see Methods). The method chosen depended on the nature of the samples. DNA from hot spring microbial mats was extracted using the Mo Bio kit since this employs mechanical shearing to break open hard material such as fungal hyphae and endospores, which aids characterisation of the eukaryal community. The results of this extraction procedure can be seen in Figure 22. The DNA had clearly been sheared as there is a smear present between 1 kb and 12 kb. However, there was sufficient concentrations of high molecular weight DNA present to be used in further analyses.

The Amersham kit was used to extract DNA from all the salt lake samples and a hot spring sample from TC8. All samples from the salt lakes were biomass captured on filters from

filtering the brine. Mechanical shearing was not required here since there was no obvious evidence of *Eukarya*. The results of this DNA extraction procedure can also be seen in Figure 22, which clearly shows intact high molecular weight material for all samples.

On comparison to the 1636 bp band in the DNA Marker (left hand lanes in Figure 22), it is clear that the Amersham kit produces a higher yield of DNA (see Methods). Results from spectrometry revealed that this kit also produced purer DNA giving an A_{260}/A_{280} ratio of an average of 1.31.

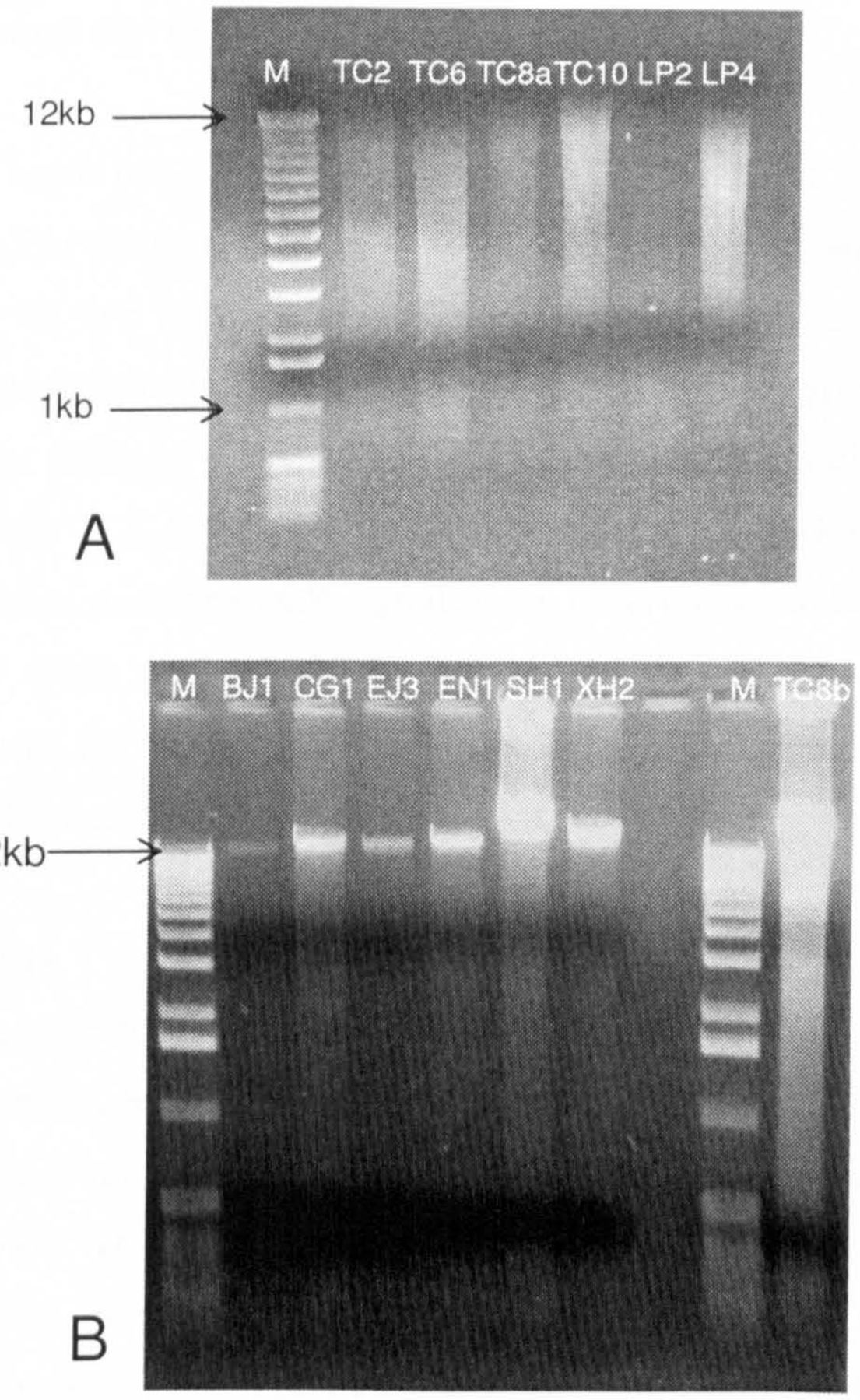


Figure 22 Results of Total Genomic DNA Extraction.

These 0.8% TAE gels show total genomic DNA extractions for environmental samples using the Mo Bio Kit (A), employing mechanical shearing to break tough material, and the Amersham kit (B), which uses only chemicals to lyse cells. Samples BJ1, CG1, EJ3, EN1, SH1 and XH2 are from salt lakes; samples TC2, TC6, TC8a, TC8b, TC10, LP2 and LP4 are from hot springs.

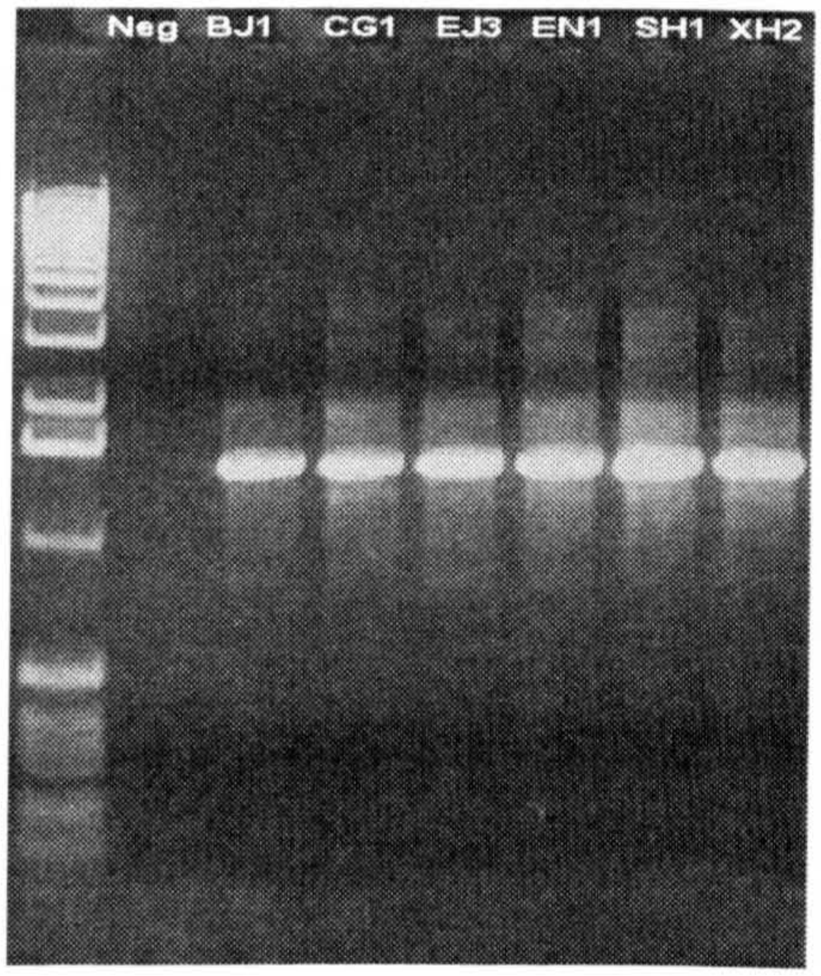
Amplification of 16S rRNA Gene by PCR

Universal bacterial primers 27Fb and rp1, and universal archaeal primers 27Fa and rp1 were used to amplify the 16S rRNA genes in bacteria and archaea respectively using PCR as described in the Methods section. Results of these amplification products are shown in Figure 23. These primers are expected to give approximately 1.5 kb products.

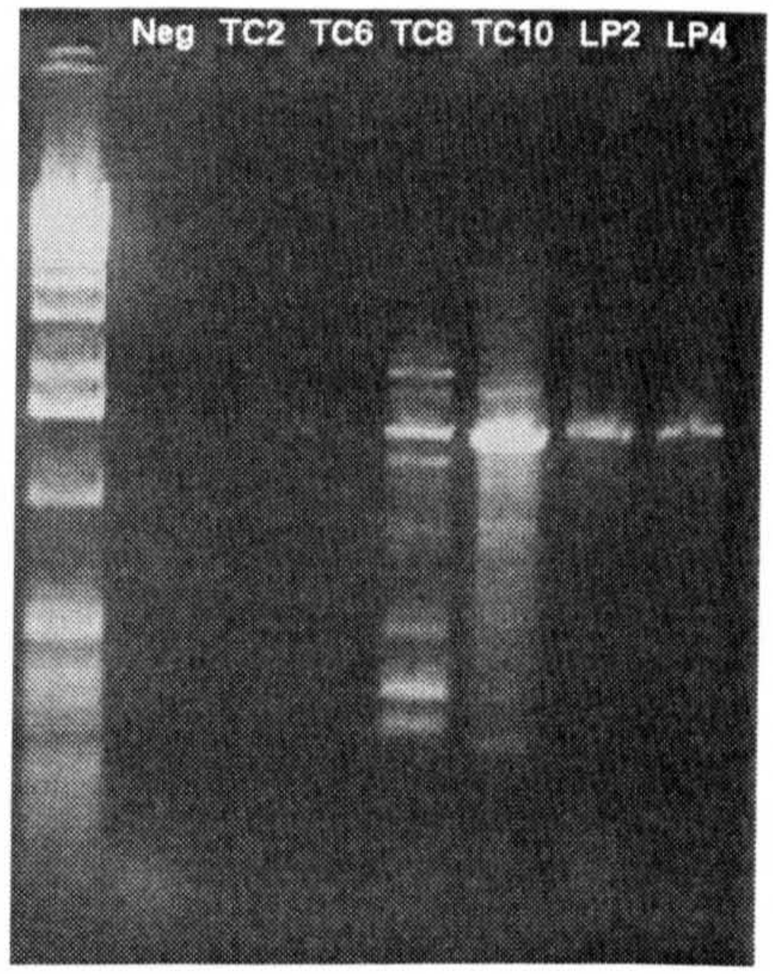
The intensities of the bands can give a rough indication of the relative numbers of organisms in each environment. For example, strong PCR signals were given when archaeal primers were used to amplify the 16S rRNA genes in salt lake samples, suggesting that archaea are abundant in all these environments. Conversely, only four of the hot springs gave PCR signals using the same primers. Sample TC6 showed a very faint band suggesting very few archaea are present here. Sample TC2 failed to give a PCR product, suggesting no archaea (or few archaea) in this environment. Strong PCR signals were observed when bacterial primers were used to amplify 16S rRNA genes from the hot spring samples. However, various band intensities were produced for the salt lakes; samples CG1, EN1 and SH1 giving particularly stronger signals than BJ1 and XH2.

Cleaning PCR Products

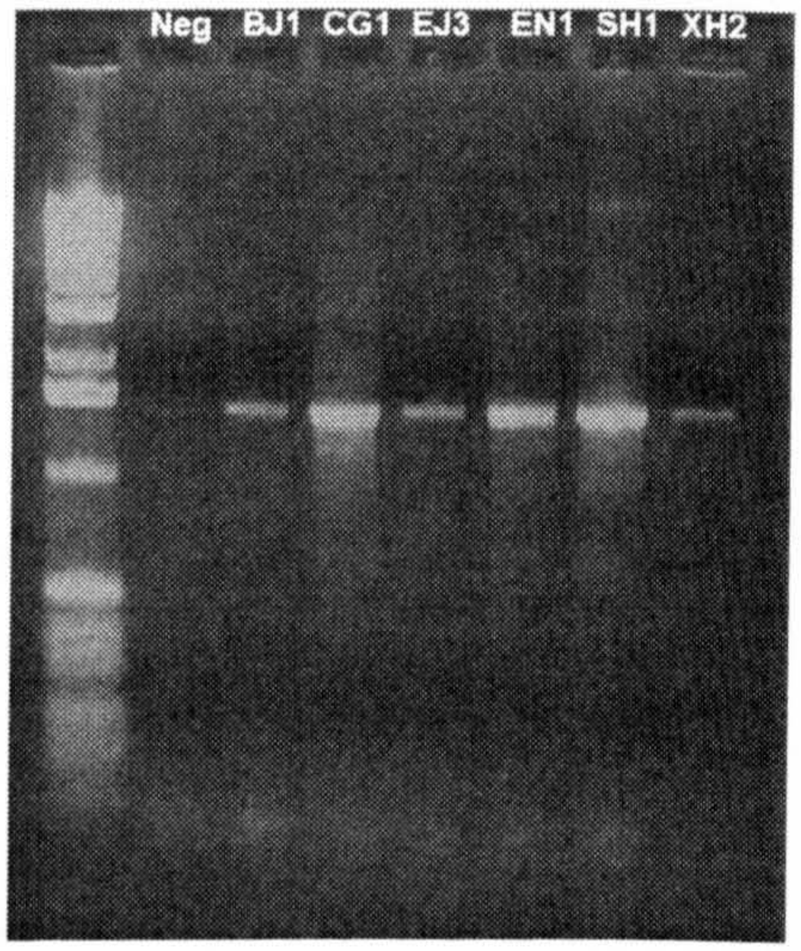
Some of the PCR amplifications produced smears and spurious bands at lower molecular weights than the 16S rRNA band, which is a problem during cloning. Experiments showed that these shorter artefacts are cloneable, and do so more efficiently than the desired longer products (data not shown). To avoid this problem, PCR reactions were gel extracted to select for the 16S rRNA band (see Methods). This process also removes primer dimers, *Taq* polymerase and contaminating salts. The PCR purification kit from Qiagen was sufficient for cleaning PCR products that did not contain these artefacts (see Methods).



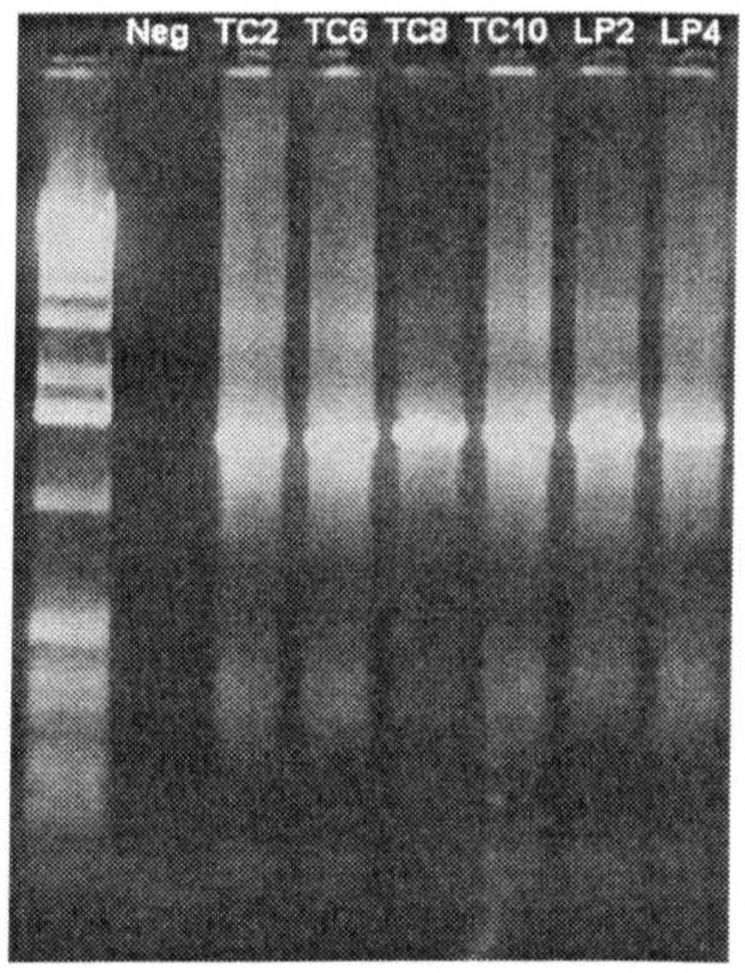
A Archaeal 16S rRNA:salt lakes



B Archaeal 16S rRNA: hot springs



C Bacterial 16S rRNA: salt lakes



D Bacterial 16S rRNA: hot springs

Figure 23 Results of 16S rRNA Gene PCR.

These 1.2% TAE gels show the results of PCR amplification of the 16S rRNA genes in both archaea and bacteria in hot springs and salt lake samples. Most samples show the expected ~1.5 kb amplicon. Gels A and B show the results for amplifying the archaeal 16S rRNA genes in salt lake (A) and hot spring (B) samples. Gels C and D show the results for amplifying the bacterial 16S rRNA genes in salt lake (C) and hot spring (D) samples.

Construction of 16S rRNA Gene Libraries

Consequent characterisation of archaeal and bacterial populations by cloning and sequencing was carried out on microbial mat samples at TC8 and LP4 and for biomass from filtered salt lake waters from BJ1, CG1, EJ3, EN1, SH1 and XH2.

Appropriate concentrations of PCR products were ligated into pGEM T-Easy cloning vector as described in the Methods section. Generally, the positive control gave 100 colonies when 100 µl cells were plated, giving a transformation efficiency of 1 x 10^5 CFU/µg vector. Approximately 90% of these clones were white. Similar numbers were generated when the 16S rRNA PCR products were cloned, showing transformation efficiencies between 10^4 to 10^5 CFU/µg vector. Between 80% and 90% of these colonies were white. The background control usually produced approximately 10 blue colonies.

Screening Clones from 16S rRNA Gene Libraries

A high throughput method of screening clones from the 16S rRNA gene libraries in 96 well plates was taken from Free (2005); this was found to be the most effective screening method. White colonies were picked with a toothpick, resuspended in 100 µl broth in a 96 well microtitre plate and incubated overnight. Clones were then screened for inserts by colony PCR using flanking vector primers M13F and M13R. Figure 24 shows the amplification products. Since these primers are set about 100 bp away from the ends of the insert, the PCR products are 200 bp larger than the expected 1.5 kb band. PCRs that did not give any bands may be because a blue colony was mistaken for a white.

Restriction Fragment Length Polymorphisms (RFLPs)

Clones were screened for the presence of identical sequences by looking at their Restriction Fragment Length Polymorphisms (RFLPs), greatly reducing the number of clones that required sequencing. This method has been used in many other 16S rRNA based studies of other environments (Head *et al.* 1998).

PCR products were digested with *Hae*III to generate RFLPs. This enzyme was chosen since it is a four base cutter and therefore cuts more frequently than five or six base cutters. It has been shown that the expected numbers of different RFLPs produced by a variety of four base cutters are not significantly different when tested *in silico* with a range of existing 16S rRNA

gene sequences in the Ribosomal Database Project (Free 2005). *Hae*III was among the enzymes tested and was therefore used in this study.

Figure 24 shows the results of a typical RFLP screen. It was not necessary to clean the PCRs prior to digestion. Different banding patterns represent different RFLPs, and it is assumed that each RFLP represents a unique 16S rRNA gene, and consequently a unique species. The patterns were analysed by eye since all RFLPs generated from a single clone library could all be run on the same gel. RFLPs were grouped accordingly and a representative clone for each group was selected for sequencing. RFLPs were compared within each clone library and not between clone libraries so that shared sequences (between clone libraries) could be identified. This also eliminated any errors made by comparisons between many gels.

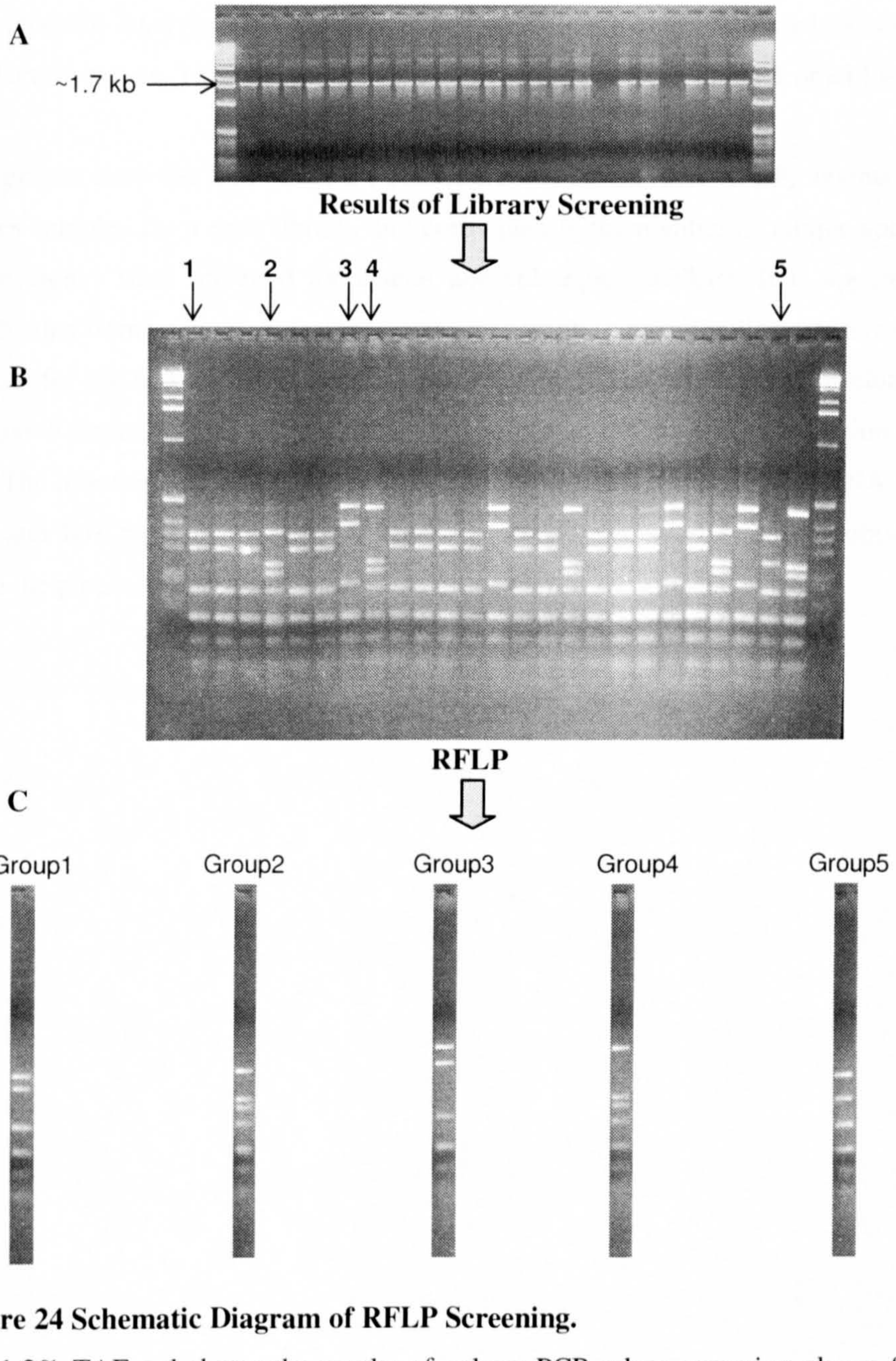


Figure 24 Schematic Diagram of RFLP Screening.

The 1.2% TAE gel shows the results of colony PCR when screening clones for the size of inserts using flanking vector primers M13F and M13R. Most samples show the expected 1.7 kb amplicon (A). The 2% TAE gel shows the results of restriction digests of various inserts amplified from clones in a 16S rRNA clone library (B). The digests were carried out using *Hae*III to produce different RFLPs. These were analysed by eye and grouped (C). A representative clone for each group was then sent for sequencing.

Rarefaction Curves

The program Analytical Rarefact version 1.3 (Holland 2003) was used to calculate data for rarefaction curves. The data was calculated from the observed number of RFLPs.

The graphs show the number of unique inserts (based on their RFLP) against the number of clones sampled from each library, and consequently the number of unique species. Up to 48 white clones were screened for inserts and subsequent RFLPs. It is apparent from these graphs that in most cases, there was sufficient sampling from the libraries to reach saturation. In fact, for the EJ3 and TC8 bacterial 16S rRNA libraries, the number of clones picked was excessive since these curves clearly plateau before any of the other rarefaction curves (Figure 25). The libraries that require more sampling are the TC8 archaeal 16S rRNA library and the LP4 and BJ1 bacterial 16S rRNA libraries to fully characterise the microbial communities since the plots are clearly still within the linear part of the graph.

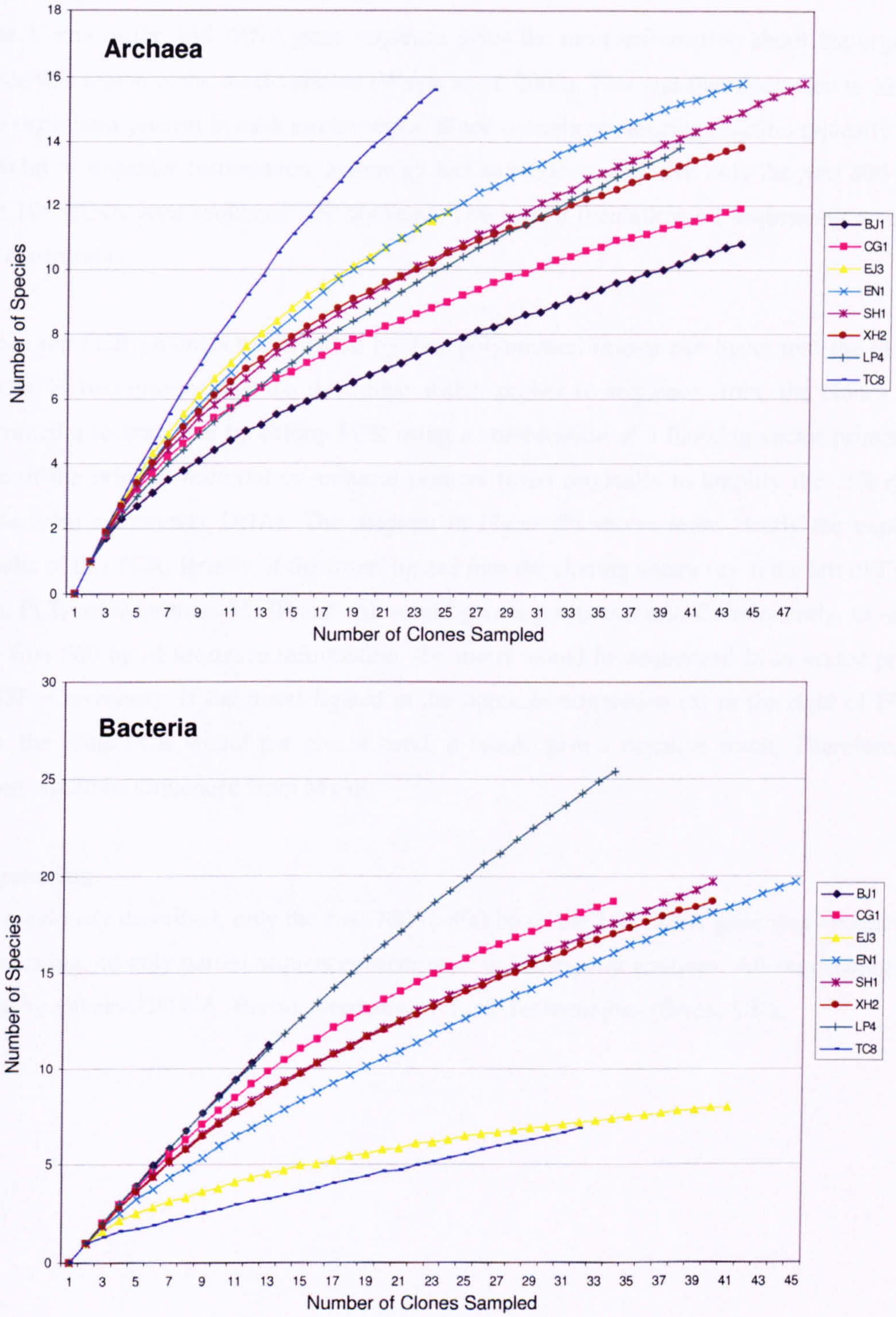


Figure 25 Rarefaction Curves.

These graphs show the rarefaction curves for all 16S rRNA libraries for archaea (top panel) and bacteria (bottom panel).

Screening the Orientation of Inserts

The 5' end of the 16S rRNA gene sequence gives the most information about the organism since this region is the most variable (Wuyts *et al.* 2002). This was therefore used to identify the organisms present in each environment. Since a single sequencing reaction typically gives 800 bp of sequence information, a strategy had to be devised so that only the *first* 800 bp of the 16S rRNA gene sequence was obtained. This would then allow all sequencing results to be comparable.

Since the PCR products are A-tailed by *Taq* polymerase, inserts can ligate into the cloning vector in two orientations. To determine which primer to sequence from, the clones were screened a second time by colony PCR using a combination of a flanking vector primer and one of the original bacterial or archaeal primers (used originally to amplify the 16S rRNA gene from community DNA). The diagram in Figure 26 shows more clearly the expected results of this PCR. Briefly, if the insert ligated into the cloning vector (as in the left of Figure 26), PCR using primers M13F and rp1 would give a positive result. Consequently, to obtain the first 800 bp of sequence information, the insert would be sequenced from vector primer M13F. Conversely, if the insert ligated in the opposite orientation (as in the right of Figure 26), the same PCR would not give a band; it would give a negative result. Therefore, the insert would be sequenced from M13R.

Sequencing

As previously described, only the first 700 to 800 bp of the 16S rRNA gene was obtained by sequencing, so only partial sequences were used in subsequent analyses. All sequencing was done by either AGOWA (Berlin, Germany) or Lark Technologies (Essex, UK).

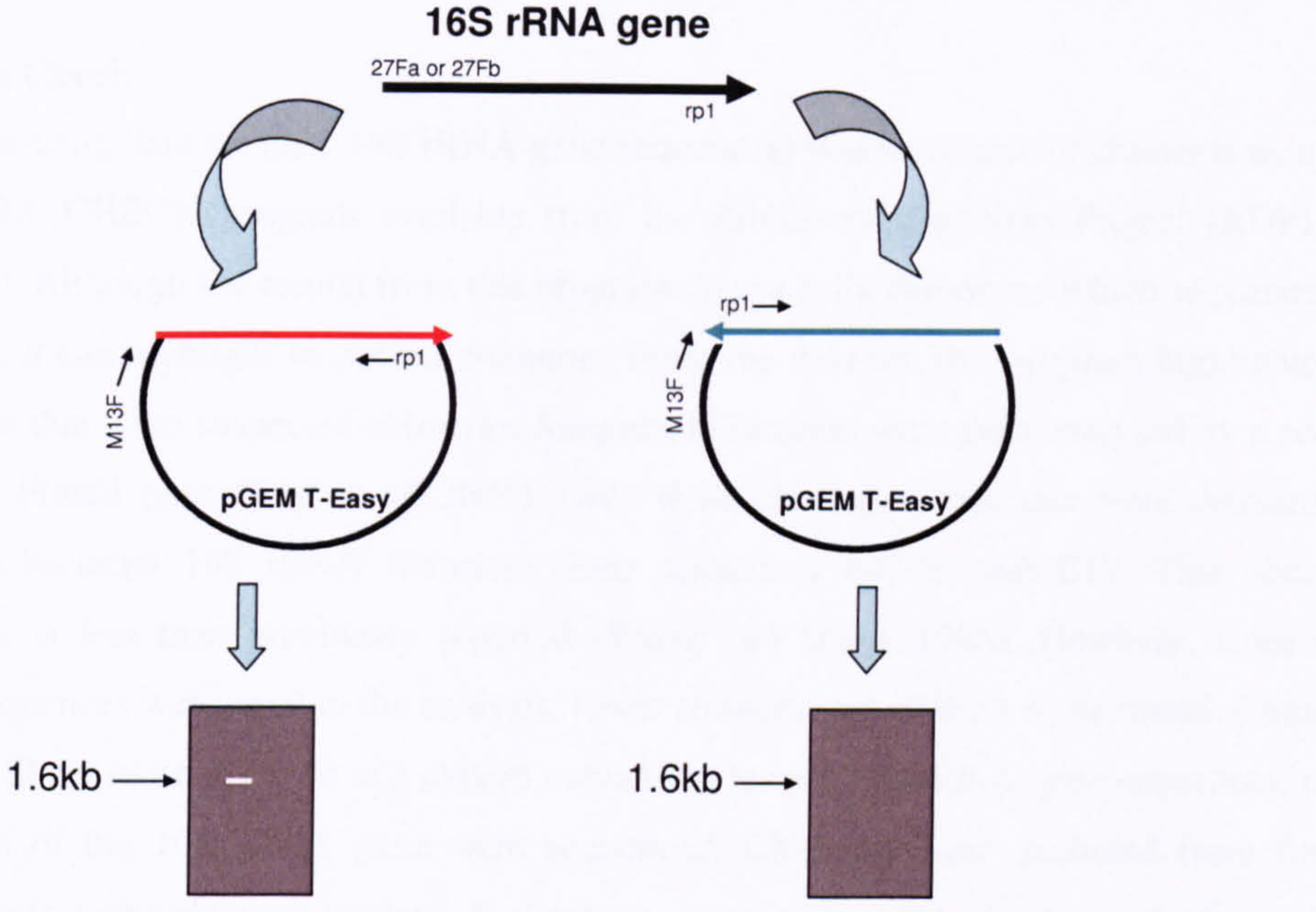


Figure 26 Screening the Orientation of Inserts.

Schematic diagram showing how clones were screened for the orientation of inserts by PCR, subsequently showing which primer the inserts should be sequenced from. A positive result (left) would mean that the insert should be sequenced from M13F; a negative result (right) would mean that the insert should be sequenced from M13R.

Sequence Analysis Using BLAST

All sequences were compared to those in the BLASTN database (Altschul *et al.* 1990). Table 13, Table 14, Table 15, Table 16 and Table 17 are the results of this analysis, which was last updated in February 2006. The tables give the nearest neighbour and the percentage identity to that neighbour. With a constantly expanding database, and more environments being characterised, these sequences may have a different nearest neighbour if analysed in the future.

Chimera Check

The sequencing data set (223 16S rRNA gene sequences) was screened for chimeras using the CHIMERA_CHECK program available from the Ribosomal Database Project (RDP) (see Methods). Although the results from this program do not fully determine which sequences are chimeras, it can highlight *suspected* chimeras. From the data set, this program highlighted 57 sequences that were suspected chimeras. Suspected chimeras were then analysed by a second program, Pintail (Ashelford *et al.* 2005). Only three chimeric sequences were detected, all from the bacterial 16S rRNA libraries: clone sequences 68, 69 and B12. This observed frequency is less than previously reported (Wang and Wang 1996). However, since only partial sequences were used in the analysis, fewer chimeras were likely to be found. Chimeras are more likely to be detected in a dataset containing longer 16S rRNA gene sequences, i.e. if both ends of the 16S rRNA gene were sequenced. Chimeras were excluded from further phylogenetic and statistical studies. A database containing a list of all chimeras currently known in the public databases is available at www.//foo.maths.uq.edu.au/~huber/doc/chimerasMar03.nds. None of the sequences reported on this website was present in this data set.

Archaeal Diversity in the Hot Springs

88% of clone sequences from the 16S rRNA gene libraries for the two hot springs (at sampling points LP4 and TC8) are related to uncultivated phylotypes that were detected in a variety of environments, including hot springs in Yellowstone National Park and Tibet, Sea of Okhotsk in Russia and in gold mines (Table 13). The remaining 12% of clone sequences are related to *Korarchaeota*, *Crenarchaeota* and known methanogens.

Table 13 Sequencing Results for the *Archaea* in Hot Springs.

Nearest neighbours from the BLASTN database to clone sequences for the archaeal 16S rRNA gene libraries for the hot springs are shown. The clone numbers have been colour coded according to the environment they are from (**Green** = LP4, **Purple** = TC8).

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
A1	Uncultured archaeon clone: OHKA5.23 (AB094546)	94	Coastal subseafloor sediments, Sea of Okhotsk (Inagaki *et al.* 2003)	18/ 40
A2	Crenarchaeotal sp. clone pJP 41 (L25301)	91	Hot spring, Yellowstone National Park, Wyoming, USA (Barns *et al.* 1994)	1/ 40
A3	Uncultured archaeon clone Oul-24 (AJ556504)	99	Anoxic rice paddy soils (Wu *et al.* 2006)	4/ 40
A4	Uncultured archaeon clone: OHKA5.23 (AB094546)	96	Coastal subseafloor sediments, Sea of Okhotsk (Inagaki *et al.* 2003)	1/ 40
A5	Uncultured euryarchaeote clone: KuA16 (AB077226)	93	Crude oil storage cavity, Kuji, Iwate, Japan (Watanabe *et al.* 2002)	3/ 40
A6	Uncultured archaeon SAGMA-T (AB050225)	90	Gold mines, South Africa (Takai *et al.* 2001)	1/ 40
A7	*Methanobacterium thermoautotrophicus* (Z37156)	99	Municipal waste-treatment facility, Illinois, USA (Zeikus and Wolfe 1972)	4/ 40
A8	Uncultured archaeon clone Oul-24 (AJ556504)	99	Anoxic rice paddy soils (Wu *et al.* 2006)	1/ 40
A9	Uncultured archaeon clone Oul-24 (AJ556504)	99	Anoxic rice paddy soils (Wu *et al.* 2006)	2/ 40
A10	Uncultured crenarchaeote Fosmid clone No. 45-H-12 (AB201308)	98	Hishikari gold mine, Japan (Nunoura *et al.* 2005)	1/ 40
A11	Uncultured archaeon clone Oul-24 (AJ556504)	98	Anoxic rice paddy soils (Wu *et al.* 2006)	1/ 40
A12	Uncultured *Crenarchaeota* clone F4 (AJ937877)	92	? (Casamayor unpublished-b)	1/ 40
A13	Uncultured archaeon SAGMA-T (AB050225)	89	Gold mines, South Africa (Takai *et al.* 2001)	2/ 40

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 13 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
C1	Uncultured crenarchaeote clone pSL4 (U63341)	98	Hot spring, Yellowstone National Park, Wyoming, USA (Barns *et al.* 1996)	1/ 24
C2	Uncultured archaeon clone archaeap14	98	Hot pool, Kuirau Park, Rotorua, New Zealand (Sunna and Bergquist 2003)	1/ 24
C3	*Methanothermobacter thermoautotrophicus* (Z37156)	99	Municipal waste-treatment facility, Illinois, USA (Zeikus and Wolfe 1972)	3/ 24
C4	Uncultured crenarchaeote clone YNP_ObP_A5 (DQ243757)	98	Hot spring, Yellowstone National Park, Wyoming, USA (Meyer-Dombard *et al.* 2005)	1/ 24
C6	Uncultured korarchaeote pBA5 (AF176347)	95	Hot spring, Yellowstone National Park, Wyoming, USA (Reysenbach *et al.* 2000b)	1/ 24
C7	Uncultured *Desulfurococcales* archaeon clone YNP_SSp_A51 (DQ243775)	96	Hot spring, Yellowstone National Park, Wyoming, USA (Meyer-Dombard *et al.* 2005)	2/ 24
C8	Uncultured archaeon clone Ou2O-50 (AJ556495)	98	Anoxic rice paddy soils (Wu *et al.* 2006)	2/ 24
C9	Uncultured archaeon isolate DGGE gel band TS5-a5-83	96	Hot spring, Daggyai Tso, Tibet (Yim *et al.* 2006)	1/ 24

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Phylogenetics of *Archaea* in the Hot Springs

Phylogenetic trees were constructed using the settings outline in the Methods. The root chosen for the archaea was *Escherichia coli*. All clone sequences are colour coded according to the hot spring that they were detected in. Only bootstrap values above 50% are shown. Phylogenetic analysis shows the distribution of the clone sequences into three monophyletic groups: *Korarchaeota*, *Crenarchaeota* and *Euryarchaeota*. Within the *Euryarchaeota*, the orders *Methanobacteriales* and *Thermoplasmatales* are represented, with clone sequences affiliated with the former. These groups are well supported, showing high bootstrap values (74% - 100%). In addition, two clone sequences form their own well supported deeply branching groups (both with bootstrap values of 100%), designated Hot Spring (HS) Clusters 1 and 2. These groups are illustrated in Figure 27.

Korarchaeota

Only one clone sequence from TC8 (C6) is affiliated with the *Korarchaeota*. It is related to uncultured korarchaeote pBA5 (95%), which was detected in a hot spring from Yellowstone National Park in Wyoming, USA. This particular clone was detected in a black filamentous mat in the Calcite Springs, where it was 83 °C and pH 7.6. The black streamers were observed along a stream where the temperature was between 60°C and 84°C (Reysenbach *et al.* 2000b).

Crenarchaeota

Several clone sequences from LP4 are affiliated with the *Crenarchaeota*. Clone sequence A10 is closely related to an uncultured crenarchaeote 45-H-12 (98%), which was detected in filamentous microbial mat growing in a discharging point of subsurface hot water stream in the Hishikari gold mine in Japan, 320 m below the land surface (Nunoura *et al.* 2005). The stream began at a subsurface aquifer through cracks of the basement rocks, ran along the mine tunnel and emptied into a waste pool. The temperature at this discharge point was 69°C, with a pH of 5.1 and a total salinity of 0.1% (w/v). In addition, the water had 0.3 mg/L of dissolved oxygen, 116 μM ammonium and 620 μM sulphate (Hirayama *et al.* 2005).

Clone sequences A1 and A4 are related to an archaeon clone: OHKA5.23 (94%), which was detected in coastal subseafloor sediments in the south western part of the Sea of Okhotsk, off the Shiretoko Peninsula at the eastern margin of Hokkaido. The 58.1 m sediment core was taken at a depth of 1225 m and was composed of pelagic clay with layers of volcanic ash as a result of several eruptions; the deepest section of the core was around 100000 years old.

Phylogenetic analysis in that study also placed this clone with the *Crenarchaeota* (Inagaki *et al.* 2003).

Clone sequence A12 is related to an uncultured *Crenarchaeota* clone F4 (92%). Unfortunately, the origin of this clone is unknown as this sequence was submitted directly to the database (Casamayor unpublished-b).

Clone sequence A2 is distantly related to a crenarchaeote sp. clone pJP 41 (91%), which was detected in a hot spring at Yellowstone National Park, USA. The pool in that study was 'Jim's Black Pool,' (now called Obsidian Pool) located in the Mud Volcano area. It measured 3 x 9 m in size and consisted of several boiling source areas with a temperature of 93°C. The water and sediment appeared black due to obsidian sand and iron sulphide. The iron content measured 15600 mg/kg. The sediment at the sampling site was 74°C and pH 7.6, while the water was slightly acidic with a pH of 6.7. Phylogenetic analysis in that study also placed this clone with the *Crenarchaeota*, with *Thermophilum pendens* as its closest neighbour (Barns *et al.* 1994). Furthermore, one clone sequence from TC8 (C1) is closely related to an uncultured crenarchaeote clone pSL4 (98%), which was also detected in Obsidian Pool as previously described. Phylogenetic analysis in that study placed this clone with the deep branching *Crenarchaeota* (Barns *et al.* 1996).

Two other clone sequences from TC8 are affiliated with the *Crenarchaeota*. Clone sequence C4 is closely related to an uncultured crenarchaeote clone YNP_ObP_A5 (98%), which was also detected in a hot spring at Yellowstone National Park. Moreover, clone sequence C7 is related to an uncultured *Desulfurococcales* archaeon clone YNP_SSp_A51 (97%), which is from the same study (Meyer-Dombard *et al.* 2005).

Euryarchaeota

The *Euryarchaeota* appear to be split into two separate clades. The clade containing the order *Thermoplasmatales* consists mostly of environmental clones and is well supported with a bootstrap value of 74%. The other clade consists of members of the order *Methanobacteriales*.

Two clone sequences from LP4 (A6 and A13) are related to an uncultured archaeon SAGMA-T (90% and 89% identity respectively). That study was based in the gold mines of South Africa. This particular clone was detected in Beatrix Mine in the extreme southern

Witwatersrand Basin. This mine contained Karoo sediments (200 to 300 Ma) that overlaid the Witwatersrand Supergroup (2.9 Ga). Compartmentalisation occurred due to Karoo basaltic dykes. Groundwater rose to the surface from springs and pans. However, water samples for that study were collected from exploratory boreholes. Clone SAGMA-T was detected in fissure water at a depth of 0.87 km. It had a temperature of 35°C and was pH 9.5. It did not contain any dissolved oxygen, but was high in chloride ions (1295 ppm). It also contained nitrates, sulphate and particulate organic carbon (all approximately 1 ppm). Phylogenetic analysis in that study placed this clone with the 'Terrestrial Miscellaneous Euryarchaeotic Group,' with the *Thermoplasmatales* as its nearest related group (Takai *et al.* 2001).

Clone sequence A5 is related to an uncultured euryarchaeote clone: KuA16 (93%), which was detected in an underground crude oil storage cavity in Kuji in Iwate, Japan. This cavity was constructed in rocky strata where the pressure of the groundwater confined the stored oil. Consequently, groundwater collected at the bottom, which was discharged to maintain storage capacity. However, this flow of groundwater established a continuous culture of microorganisms in a habitat where there was an excess of hydrocarbons for use as electron donors (Watanabe *et al.* 2002). The groundwater was anaerobic with a temperature between 14.4°C and 16°C and a pH between 7.6 and 8.8 (Watanabe *et al.* 2000). Phylogenetic analysis in that study placed this clone with the *Euryarchaeota*, but not with any known orders. Therefore, it was placed in a group arbitrarily named 'Candidate Division II.' The group's nearest related order was the *Thermoplasmatales* (Watanabe *et al.* 2002).

One clone sequence from TC8 also clusters with this group of *Euryarchaeota*. Clone sequence C8 is closely related to an uncultured archaeon clone Ou2O-50 (98%), which was detected in temperate anoxic rice paddy soils (Wu *et al.* 2006).

Euryarchaeota*: *Methanobacteriales

Methanobacteriales are strict anaerobes, and all grow by oxidising hydrogen. It is a highly specialised group that is only able to catabolise methanol, some secondary alcohols, formate or CO. They are widely distributed in nature, found in anaerobic habitats such as sediment, soil, sewage digesters and the gastrointestinal tract of animals (Boone *et al.* 2001). Known 16S rRNA gene sequences of the order *Methanobacteriales* were used in the phylogenetic tree to find out which clone sequences would branch in the same clade.

One clone sequence from LP4 (A7) and one sequence from TC8 (C3) are closely related to *Methanobacterium thermoautotrophicus* (both at 99% identity). This methanogen was first isolated from a municipal waste-treatment facility in Illinois, USA (Zeikus and Wolfe 1972). It grows at temperatures between 40°C and 70°C, but grows optimally at 65°C. Other clone sequences from LP4 (A3, A8, A9 and A11) are closely related to an uncultured archaeon clone OuI-24 (98% - 99%), which was detected in temperate anoxic rice paddy soils (Wu *et al.* 2006).

Novel Lineages

Two clone sequences from TC8 form two novel lineages that are related to, but are distinct from all known *Euryarchaeota*, which have been named Hot Spring (HS) Clusters 1 and 2.

HS Cluster 1 contains clone sequence C2, which is related to an uncultured archaeon clone archaeap14 (98%), which was detected from a hot pool in Kuirau Park in Rotorua, New Zealand (Sunna and Bergquist 2003). This pool, containing several cubic metres of water, had a temperature of 78°C and a pH of 7.5 (Saul *et al.* 1999).

HS Cluster 2 consists of clone sequence C9, which is related to an uncultured archaeon isolate DGGE gel band TS5-a5-83, which was detected in a remote hot spring in the Daggyai Tso geothermal region in southern central Tibet at an altitude of 5050 m. The water contained negligible amounts of sulphide, but high levels of dissolved silica, boron and lithium, with lower levels of calcium and magnesium. The water at the sources had a temperature of 200°C to 220°C; the temperature in the water channel was at 83°C and had a pH of 8.1 to 8.4. Phylogenetic analysis in that study placed this clone with the *Euryarchaeota*, but not within any of the known orders (Yim *et al.* 2006).

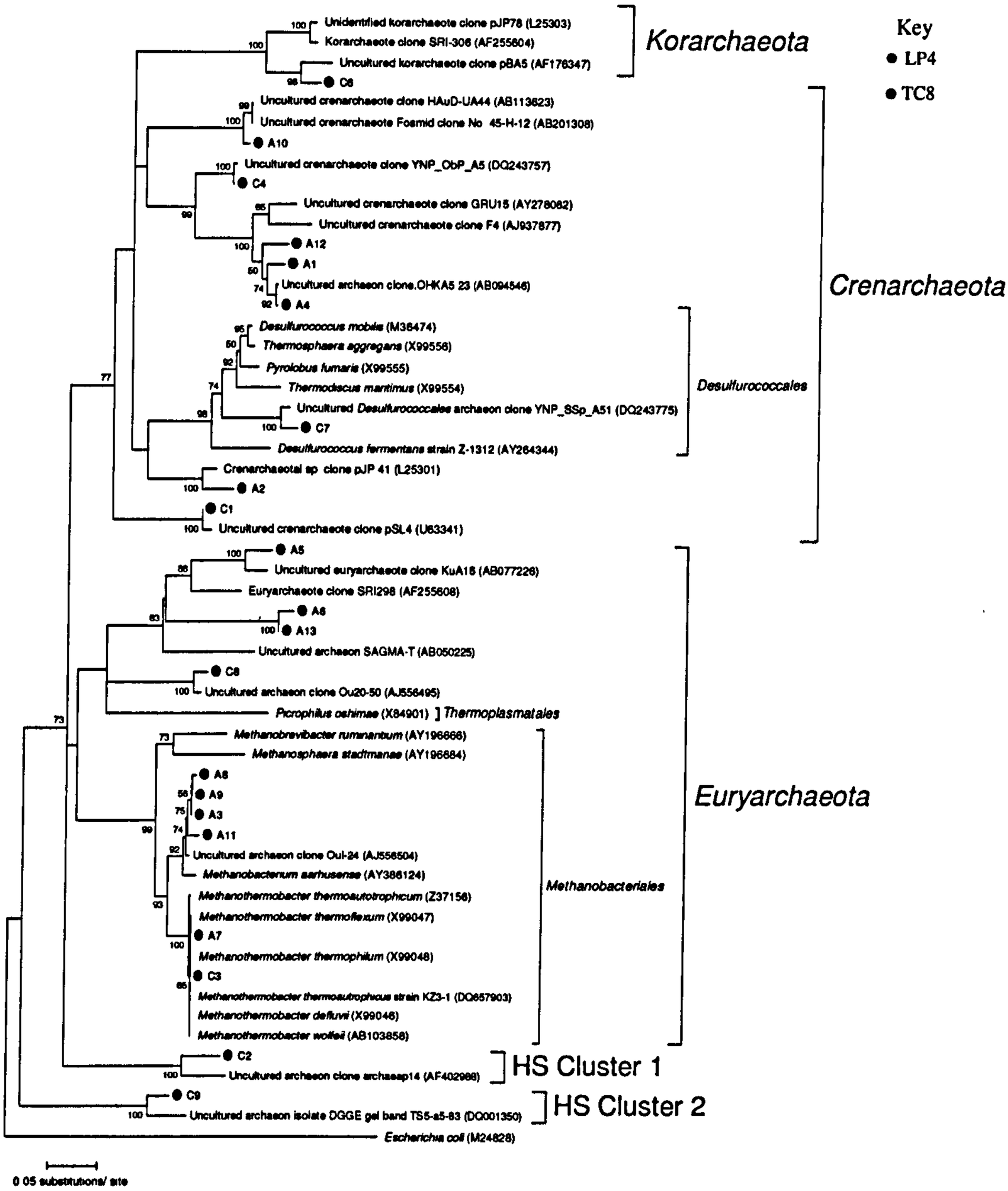


Figure 27 Phylogenetic Tree of the *Archaea* in the Hot Springs.

This phylogenetic tree shows the relationships between the clone sequences with existing 16S rRNA gene sequences for *Korarchaeota*, *Crenarchaeota* and *Euryarchaeota* from the hot springs.

Bacterial Diversity in the Hot Springs

All but four of the clone sequences are related to uncultivated phylotypes detected from a range of environments including other hot springs, marine hydrothermal systems, volcanic deposits, rice roots, mines, activated sludge and landfills (Table 14). The other clone sequences are related to known species: *Caldotoga Fontana* (*Thermotogae*), *Friedmanniella spumicola* (*Actinobacteria*) and *Thermus rehai* (*Deinococcus-Thermus*).

Table 14 Sequencing Results for the *Bacteria* in Hot Springs.

Nearest neighbours from the BLASTN database to clone sequences for the bacterial 16S rRNA gene libraries for the hot springs are shown. The clone numbers have been colour coded according to the environment they are from (**Green** = LP4, **Purple** = TC8).

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
B1	Toluene-degrading methanogenic consortium bacterium Eub 6 (AF423186)	91	Pensacola aquifer, Florida, USA (Ficker *et al.* 1999)	2/ 34
B2	*Cyanobacterium* sp. (OS type I) (L04709)	97	Octopus spring, Yellowstone National Park, Wyoming, USA (Weller *et al.* 1992)	2/ 34
B3	Uncultured bacterium clone: GAB-B04 (AB183860)	94	Geothermal aquifer, Victoria, Great Artesian Basin, Australia (Kimura *et al.* 2005)	2/ 34
B4	Uncultured *Bacteroidetes* bacterium clone HrhB26 (AM159243)	88	Rice roots and the rhizosphere (Lu *et al.* 2006)	1/ 34
B5	Uncultured bacterium clone pLW-102 (DQ066982)	90	Sediment of Lake Washington, USA (Nercessian *et al.* 2005)	1/ 34
B6	Uncultured bacterium clone JSX31 (DQ340754)	98	Hot Spring (Xiao *et al.* unpublished)	1/ 34
B7	Toluene-degrading methanogenic consortium bacterium Eub 6 (AF423186)	90	Pensacola aquifer, Florida, USA (Ficker *et al.* 1999)	2/ 34
B8	Uncultured bacterium clone:IBC2-15 (AB175561)	90	Iheya North field, Mid-Okinawa Trough, Japan (Nakagawa *et al.* 2005)	1/ 34
B9	Uncultured bacterium clone ZZ12AC2 (AY214187)	94	Benzene-contaminated groundwater (Alfreider and Vogt unpublished)	2/ 34
B10	Uncultured deltaproteobacterium clone GuBH2-AD/TzT-67 (AJ519663)	97	Uranium mining waste piles and mill tailings (Geissler *et al.* unpublished)	1/ 34
B11	Uncultured *Chloroflexaceae* bacterium clone Witch_8 (AF421757)	95	Witch Pond, Yellowstone National Park, Wyoming, USA (Boomer *et al.* 2002)	1/ 34

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 14 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
B13	Uncultured bacterium clone 1700a2-23 (AY917295)	96	Recent Hawaiian volcanic deposits (Gomez-Alvarez and Nuesslein unpublished)	1/ 34
B14	Uncultured bacterium clone HC71_17 (AY508318)	89	Environmental samples (Manning *et al.* unpublished)	1/ 34
B15	Uncultured bacterium clone AKIW776 (DQ129254)	93	Air of two cities in Texas (Andersen unpublished)	2/ 34
B16	Uncultured bacterium clone BE325FW032701CTS_hole1-69 (DQ088759)	97	Continental crust (Lin *et al.* 2006)	1/ 34
B17	Uncultured *Deinococcus* sp. clone DR546AS36a (DQ128151)	95	Gold mines, Witwatersrand Basin, South Africa (Onstott *et al.* 2003)	1/ 34
B18	Uncultured betaproteobacterium clone: HAuD-LB/2-3 (AB176697)	95	Geothermal water, Hishikari gold mine, Japan (Hirayama *et al.* 2005)	1/ 34
B19	Uncultured *Deinococcus* sp. clone DR546AS36a (DQ128151)	94	Gold mines, Witwatersrand Basin, South Africa (Onstott *et al.* 2003)	1/ 34
B20	Uncultured bacterium clone 1700a2-23 (AY917295)	97	Recent Hawaiian volcanic deposits (Gomez-Alvarez and Nuesslein unpublished)	1/ 34
B21	Uncultured bacterium clone HC71_17 (AY508318)	94	Environmental samples (Manning *et al.* unpublished)	1/ 34
B22	Uncultured bacterium clone A11 (DQ145124)	96	Wetland sediments (Jackson unpublished)	1/ 34
D1	Uncultured *Aquificales* bacterium clone pKB (AF453506)	94	Kamchatka, Russia (Takacs-Vesbach *et al.* unpublished)	18/ 33
D2	*Caldotoga fontana* (AJ237665)	94	? (Xue *et al.* unpublished)	1/ 33
D3	*Friedmanniella spumicola* (AF062535)	93	Activated sludge foam, Victoria, Australia (Maszenan *et al.* 1999)	1/ 33
D4	*Thermus rehai* (AF331969)	99	Tengchong, Rehai, Yunnan province, China (Lin *et al.* 2002)	1/ 33

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 14 Continued

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
D5	Uncultured *Aquificales* bacterium clone pKB (AF453506)	94	Kamchatka, Russia (Takacs-Vesbach *et al.* unpublished)	1/ 33
D6	Uncultured *Aquificales* bacterium clone pKB (AF453506)	94	Kamchatka, Russia (Takacs-Vesbach *et al.* unpublished)	1/ 33
D7	Uncultured deltaproteobacterium clone SM1G01 (AF445695)	93	Angel Terrace, Mammoth Hot Springs, Yellowstone National Park, Wyoming, USA (Bonheyo *et al.* unpublished)	1/ 33
D8	Uncultured *Aquificales* bacterium clone pKB (AF453506)	94	Kamchatka, Russia (Takacs-Vesbach *et al.* unpublished)	1/ 33
D9	Uncultured *Aquificales* bacterium clone pKB (AF453506)	93	Kamchatka, Russia (Takacs-Vesbach *et al.* unpublished)	1/ 33
D10	Uncultured *Aquificales* bacterium clone pKB (AF453506)	93	Kamchatka, Russia (Takacs-Vesbach *et al.* unpublished)	1/ 33
D11	Unidentified green non-sulfur bacterium OPB11 (AF027032)	96	Obsidian Pool, Yellowstone National Park, Wyoming, USA (Hugenholtz *et al.* 1998)	1/ 33
D12	Uncultured *Aquificales* bacterium clone pKB (AF453506)	93	Kamchatka, Russia (Takacs-Vesbach *et al.* unpublished)	1/ 33
D13	Uncultured *Hydrogenothermus* sp. clone OPPB005 (AY861748)	92	Obsidian Pool, Yellowstone National Park, Wyoming, USA (Spear *et al.* 2005)	1/ 33
D14	Uncultured *Thermotogales* bacterium clone WB3B071 (AY862075)	95	Washburn Spring, Yellowstone National Park, Wyoming, USA (Spear *et al.* 2005)	1/ 33
D15	Uncultured Termite group 1 bacterium clone OPPA005 (AY861728)	93	Obsidian Pool, Yellowstone National Park, Wyoming, USA (Spear *et al.* 2005)	2/ 33

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Phylogenetics of *Bacteria* in the Hot Springs

Phylogenetic trees were drawn using the clone sequences, their nearest neighbours and existing rRNA gene sequences of type species of known deeply branching and phototrophic bacteria. Sequences that did not give any information about the relatedness of the clone sequences were removed from the tree. The root chosen for the phylogenetic tree was *Methanospirillum hungatei*. The resulting tree is shown in Figure 29. Phylogenetic analysis divided the clone sequences into eleven monophyletic groups. Clone sequences are affiliated with the deeply branching groups *Aquificae* and *Thermotogae*, and with *Nitrospirae*, *Deinococci* and *Chloroflexi*. There are also clone sequences related to the phototrophic bacteria cyanobacteria and *Chlorobia*. Moreover, clone sequences are related to *Actinobacteria*, *Bacteroidetes* and *Proteobacteria* (*Delta* and *Beta* subdivisions). All groups are well supported with bootstrap values between 58% and 99%. In addition, there are lineages that do not affiliate with any of the known phyla within the domain *Bacteria*. These have been designated Hot Spring (HS) Clusters 1 to 7. All are well supported with bootstrap values between 90% and 99%.

Aquificae

All clone sequences affiliated with this group are from TC8. Clone sequences D1, D5, D6, D8, D9, D10 and D12 are related to an uncultured *Aquificales* bacterium clone pKB (93% - 94%), which was detected in a near-neutral thermal Spring in Kamchatka, Russia's largest volcanic belt flanked by the Pacific Ocean and the Sea of Okhotsk (Takacs-Vesbach *et al.* unpublished). Active volcanism occurs in the Kamchatka arc where the Pacific Plate subducts beneath the Kamchatka Peninsula (Hochstaedter *et al.* 1996). A representative *Aquificales* clone from TC8 (D1) was used in further analysis. Figure 28 shows the phylogenetic relationship between D1 and existing *Aquificales* 16S rRNA gene sequences found in other streamer communities that can be classified into different branches: J, S, H and P (Skirnisdottir *et al.* 2000). Clone D1 clearly forms its own lineage distinct from other known *Aquificae*.

Clone sequence D13 is related to an uncultured *Hydrogenothermus* sp. clone OPPB005 (92%), which was detected in a Obsidian Pool in Yellowstone National Park, Wyoming, USA, which was previously described (Spear *et al.* 2005).

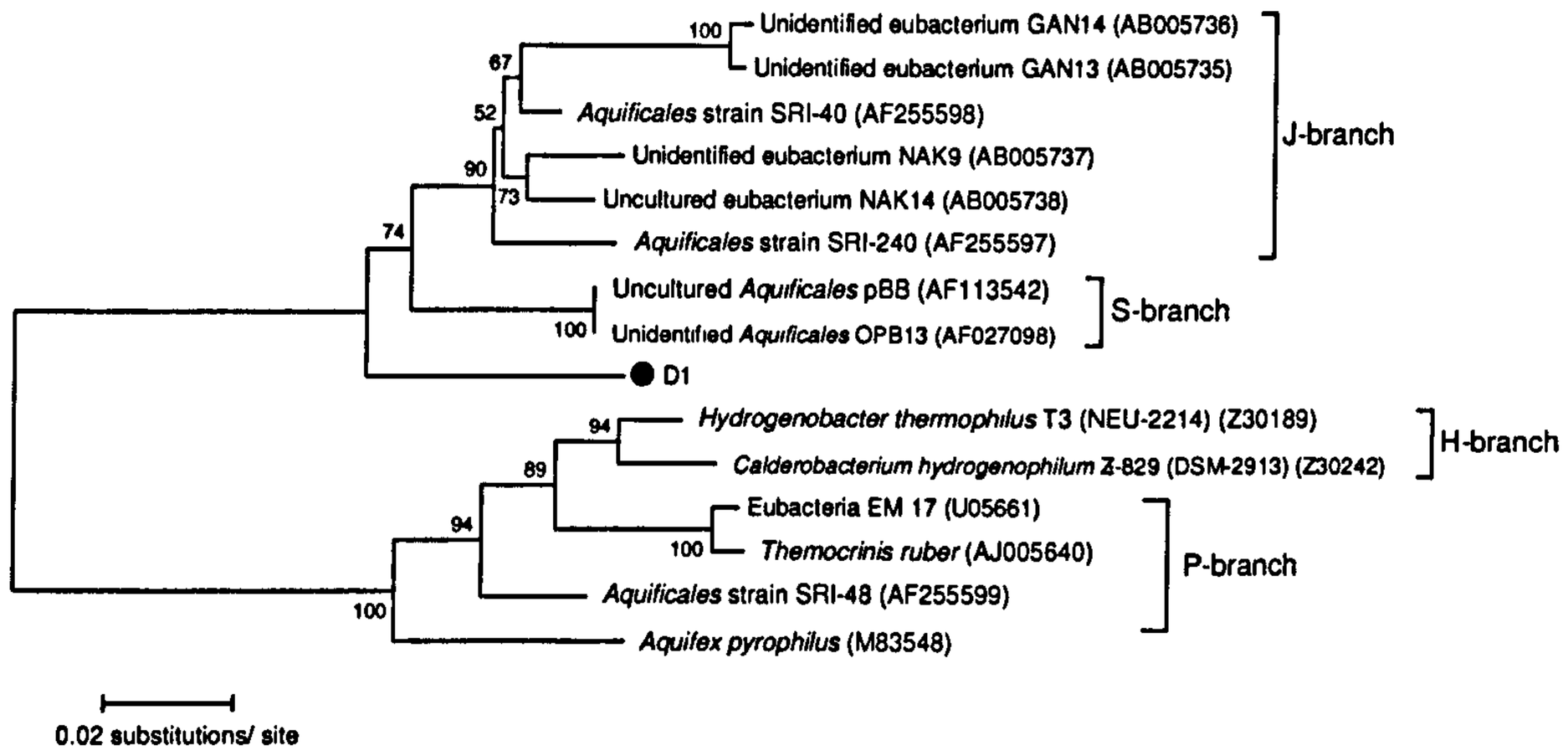


Figure 28 *Aquificae* from Streamer Communities.

Phylogenetic tree showing the relationships between clone sequence D1 from TC8 with existing 16S rRNA gene sequences of *Aquificae* from other streamer communities.

Thermotogae

Just one clone from TC8 is affiliated with the *Thermotogae*. Clone sequence D2 is related to *Caldotoga fontana* (94%). Unfortunately, the origin of this clone is unknown since the authors have removed this sequence from the database (Xue *et al.* unpublished). Clone sequence D2 is also closely related to *Fervidobacterium nodosum*, showing 91% identity.

Nitrospirae

Three clone sequences from LP4 branch within this clade, showing relatedness to the genus *Thermodesulfovibrio*, which is the only thermophilic genus in this phylum. Clone sequence B3 is related to an uncultured bacterium clone: GAB-B04 (94%), which was detected in a geothermal aquifer associated with the subsurface of the Great Artesian Basin in Australia. This basin occupies 1.7 million km^2 and covers more than a fifth of the Australian continent. It consists of layers of permeable sandstone and impermeable shale, formed between 100 and 250 million years ago. The groundwater in the aquifer is recharged by rainfall and is warmed by geothermal heat, thereby increasing water pressure. The groundwater in that study was collected from a bore hole near Blackfall in western Queensland. The flow rate does not permit the formation of microbial mats; therefore the bore water reflected the environment deep in the aquifer. The temperature of this bore water was 64.4°C and had a pH of 8.0. Phylogenetic analysis in that study also placed this clone with the *Nitrospirae*, with *Thermodesulfovibrio islandicus* as its nearest neighbour (Kimura *et al.* 2005). Clone sequences B14 and B21 are related to an uncultured bacterium clone HC71_17 (89% and 94% identity respectively), which was detected from environmental samples (Manning *et al.* unpublished).

Deinococcus-Thermus

Three clone sequences branch with the two orders of this group; the *Deinococcales* and *Thermales*. Clone sequences B17 and B19 are related to an uncultured *Deinococcus* sp. clone DR546AS36a (95% and 94% identity respectively), within the order *Deinococcales*. This was detected in ultradeep gold mines in the Witwatersrand Basin, South Africa. The mined zone was composed of quartz, sulphide, uraninite and gold bearing layer, sandwiched by quartzite and conglomerate. Rock, air and water samples were collected 3.2 km below the surface for analysis. Phylogenetic analysis in that study also placed this clone with *Deinococcus*, with *Deinococcus proteolyticus* as its nearest neighbour (Onstott *et al.* 2003).

The other clone sequence is from TC8 and is affiliated with the order *Thermales*. Clone sequence D4 is closely related to *Thermus rehai* (99%), which was isolated from a hot spring in Rehai near Tengchong in the Yunnan Province, China where the water temperature varied between 55°C and 103°C, with a pH of 6.5. It is unclear if it was isolated from the same hot springs as in this study. This particular isolate is a Gram negative, non-spore forming, non-motile rod. Growth occurs between 40°C and 80°C, but grows optimally at 65°C to 70°C. It also grows between pH 4.5 and pH 10.5, but grows optimally between pH 7.5 and 8.5. It utilises glucose, acetate and galactose and can convert nitrate to nitrite (Lin *et al.* 2002).

Chloroflexi

Clone sequence B11 from LP4 is related to an uncultured *Chloroflexaceae* bacterium clone Witch_8 (95%), which was detected in a red layer community in a hot spring at Yellowstone National Park, USA. The spring, called Witch Pond was located in the remote Witch Creek Thermal Basin near Heart Lake. The sampling site had a temperature of 47°C to 50°C, and a pH of 8.7. The microbial mat collected had a visible red layer. Phylogenetic analysis in that study placed this clone in a group that was related to, but distinct from known *Chloroflexus* species. This new 'red cluster' within the green non-sulphur bacterial lineage was designated 'YRL-B' (Boomer *et al.* 2002).

Actinobacteria

One clone sequence from TC8 (D3) is related to *Friedmanniella spumicola* (93%), which was isolated from stable foam on the surface of an aerobic reactor in an activated sludge plant treating waste from an orange juice processing plant, located in Mildura in Victoria, Australia. This particular isolate is a Gram positive, non-spore forming, strictly aerobic coccus that forms aggregates. It grows at temperatures between 15°C to 37°C, but grows optimally at 25°C. It also grows at pH 5.5 - 8.0, but grows optimally at pH 7.0 - 7.5. Phylogenetic analysis in that study showed *Friedmanniella antarctica* as its nearest relative (Maszenan *et al.* 1999).

One clone sequence from LP4 (B15) is related to an uncultured bacterium clone AKIW776 (93%), which was detected in the air of two cities in Texas (Andersen unpublished).

Chlorobia

Only two clone sequences from LP4 are affiliated with this group. Clone sequence B5 is distantly related to an uncultured bacterium clone pLW-102 (90%). This was detected in the sediment of Lake Washington, USA. This is a freshwater lake where the top layer of the

sediment contained steep gradients of methane and oxygen. It had a high population of methanotrophs and so methane oxidation occurred at high rates. Samples were collected from a study site 60 m deep. Phylogenetic analysis in that study placed this clone in the 'unclassified bacteria' group (Nercessian *et al.* 2005). However, clone B5 is distantly related to *Chlorobium ferrooxidans* showing 81% identity. Clone sequence B6 is closely related to an uncultured bacterium clone JSX31 (98%), which was detected in a hot spring (Xiao *et al.* unpublished). Clone sequence B6 also shows 85% identity to *Chlorobium ferrooxidans*.

Bacteroidetes

Just one clone sequence from LP4 branches within this clade. Clone sequence B4 is distantly related to an uncultured *Bacteroidetes* bacterium clone HrhB26 (88%), which was detected in rice roots (Lu *et al.* 2006).

Cyanobacteria

One clone sequence from LP4 (B2) is related to a *Cyanobacterium* sp. (OS type I) (97%), which was detected in a microbial mat in Octopus spring, located in Yellowstone National Park, USA. This spring had a temperature of 50°C - 55°C. Phylogenetic analysis and determination of the secondary structure of the 16S rRNA gene sequence in that study showed that the clone was related to cyanobacteria and placed this clone near *Anacystic nidulans* (Weller *et al.* 1992).

Deltaproteobacteria

One clone sequence from LP4 (B10) is related to an uncultured deltaproteobacterium clone GuBH2-AD/TzT-67 (97%), which was detected in uranium mining waste piles and mill tailings (Geissler *et al.* unpublished). Branching with the *Deltaproteobacteria*, but is distinct from known sulphate reducers is from TC8 (D7); it is related to an uncultured deltaproteobacterium clone SM1G01 (93%), which was detected in Angel Terrace at the Mammoth Hot Springs of Yellowstone National Park, USA (Bonheyo *et al.* unpublished).

Betaproteobacteria

One clone from LP4 (B18) is related to an uncultured betaproteobacterium clone: HAuD-LB/2-3 (95%), which was detected in subsurface geothermal water stream in the Hishikari gold mine in Japan, which was described previously. Phylogenetic analysis in that study also placed this clone with *Thiobacillus aquaesulis* as its nearest neighbour (Hirayama *et al.* 2005).

Novel Lineages

Several clone sequences did not affiliate with any known groups and were designated HS Clusters 1 to 7.

HS Cluster 1 has just one clone sequence from LP4 (B16), which branches near to, but is distinct from the *Chrysiogenetes*. It is related to an uncultured bacterium clone BE325FW032701CTS_hole1-69 (97%), which was detected in a continental crust (Lin *et al.* 2006).

HS Cluster 2 branches near to the *Bacteroidetes*. It consists of just two clone sequences from LP4 (B13 and B20), which are both related to an uncultured bacterium clone 1700a2-23 (96% and 97% identity respectively), which was detected in recent Hawaiian volcanic deposits (Gomez-Alvarez and Nuesslein unpublished).

HS Cluster 3 contains one sequence from LP4 (B22), which is related to an uncultured bacterium clone A11 (96%), whuch was detected in wetland sediments (Jackson unpublished). These sequences form a separate branch that does not bear much relation to any of the groups in the domain *Bacteria*.

HS Cluster 4 consists of just one clone sequence from LP4 (B9), which is related to an uncultured bacterium clone ZZ12AC2 (94%), which was detected in benzene-contaminated groundwater (Alfreider and Vogt unpublished). These sequences branch near to the *Chloroflexi*.

Branching near to the *Thermomicrobia-Dictyoglomi* branch, HS Cluster 5 consists of just one clone from LP4 (B8); it is distantly related to an uncultured bacterium clone:IBC2-15 (90%), which was detected in a deep-sea hydrothermal system of the Iheya North field in the Mid-Okinawa Trough, south of Japan. Hydrothermal activity occurred in a volcanic ridge in a continental margin. The sediment here was full of organic molecules; therefore the vent fluids became highly alkaline, with high concentrations of ammonium ions, carbon dioxide, hydrogen sulphide and methane (Nakagawa *et al.* 2005).

HS Cluster 6 consists of three clone sequences that also branch near to the *Thermomicrobia-Dictyoglomi*. Clone sequences B1 and B7 from LP4 are related to a toluene-degrading methanogenic consortium bacterium Eub 6, which was detected in a microbial consortium that was enriched from creosote-contaminated aquifer sediments and maintained on toluene as its sole carbon source for ten years (Ficker *et al.* 1999). The aquifer solids were from the Pensacola aquifer in Florida, USA, which consisted of sand deposits with discontinuous layers of silt and clay. The top 30 m was contaminated with creosote and pentachlorophenol. The sample was taken from an actively methanogenic sandy zone at a depth of 6 m (Edwards and Grbić-Galić 1994). Phylogenetic analysis in that study showed that this clone did not group closely with any known organism and placed it in a separate clade near to *Dictyoglomi* (Ficker *et al.* 1999). Also in this cluster, clone sequence D11 from TC8 is related to an unidentified green non-sulphur bacterium OPB11 (96%), which was detected in Obsidian Pool in Yellowstone National Park, USA, which was previously described. Phylogenetic analysis in that study placed this clone with the 'Green non-sulphur bacteria' group, with *Dehalococcoides ethenogenes* as its nearest neighbour (Hugenholtz *et al.* 1998).

Branching outside the known groups within the domain *Bacteria* is HS Cluster 7. It consists of just one sequence from TC8 (D14), which is related to an uncultured *Thermotogales* bacterium clone WB3B071 (95%), which was detected in Washburn Spring in Yellowstone National Park, USA. The temperature at this sampling site was 86°C and had a pH of 6.2. Analysis of the water chemistry revealed that it contained 167 µM sulphide, 44 µM sulphate, approximately 5.8 µM methane and approximately 9.75 µM carbon dioxide. In addition, there was 68000 µg/L aluminium and 340 µg/L manganese (Spear *et al.* 2005). Despite being a *Thermotogales* clone, my phylogenetic analysis has clearly placed this clone in a branch separate from the *Thermotogae*.

'HS Clone' consists of one sequence from TC8 (D15); it is related to an uncultured Termite group 1 bacterium clone OPPA005 (93%), which was detected in Obsidian Pool in Yellowstone National Park, USA (Spear *et al.* 2005). However, the clone sequence only showed high localised similarity to a short stretch of the 'Termite group 1 bacterium clone OPPA005' sequence (339 bases showing 93% identity). When used in phylogenetic analysis, 'Termite group 1 bacterium clone OPPA005' did not appear to show any relatedness to clone sequence D15 and so was removed from the tree. The second closest match was used in the analysis instead (an uncultured clone RsaP85), but this similarly did not show any relatedness

to clone sequence D15 and branched with the *Thermotogae* instead. Clone sequence D15 has clearly formed its own separate branch near to, but significantly different from the *Deinococcus-Thermus* branch (Figure 29).

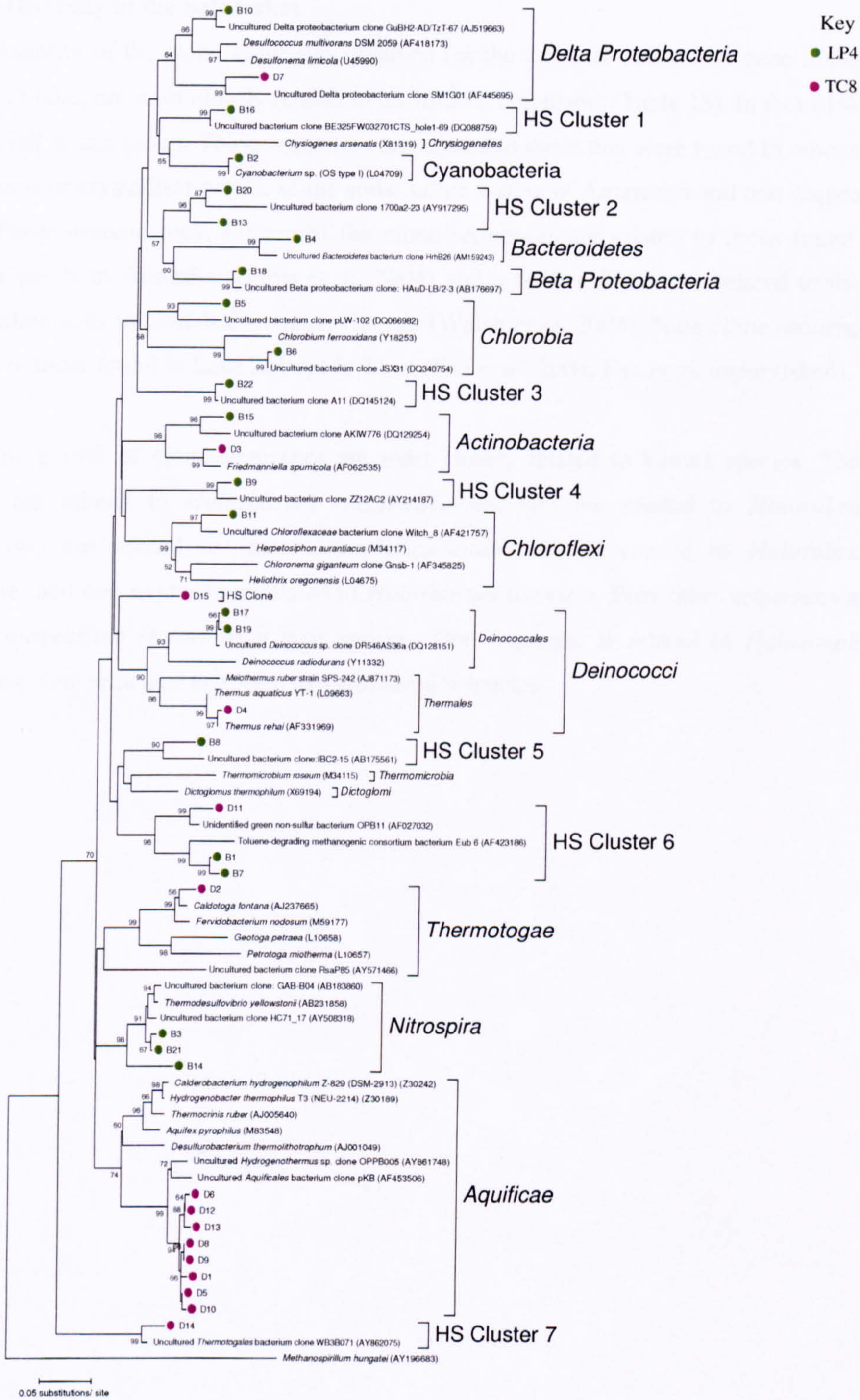


Figure 29 Phylogenetic Tree of the *Bacteria* in the Hot Springs.

Phylogenetic tree showing the relationships between the clone sequences with existing 16S rRNA gene sequences for the bacterial community from the hot springs.

Archaeal Diversity of the Salt Lakes

The vast majority of the clone sequences obtained for the archaeal 16S rRNA gene libraries from the salt lakes are most closely related to uncultured organisms (Table 15). In fact 81% of sequences fall in this group. These sequences are related to those that were found in other salt lakes, salterns or crystalliser ponds, saline soils, saline waters of Antarctica and one sequence found in Permo-triassic rock. Fifteen of the clone sequences are related to those found in crystalliser ponds in Australia (Burns *et al.* 2004) and a further fifteen are related to those found in saline soils in British Columbia, Canada (Walsh *et al.* 2005). Nine clone sequences are related to those found in Lake Zabuye in Tibet (Fan *et al.* 2004; Fan *et al.* unpublished).

The remaining 19% of clone sequences are most closely related to known species. Three sequences are related to *Halorubrum saccharovorum*; two are related to *Halorubrum tibetense*; two are related to *Halorubrum vacuolatum*; one is related to *Halorubrum xinjiangense*; and one sequence is related to *Halorubrum terrestre*. Four other sequences are related to unspecified *Halorubrum* type species. One sequence is related to *Halosimplex carlsbadense*. One sequence is related to a *Haloarcula* species.

Table 15 Sequencing Results for the *Archaea* in Salt Lakes.

Nearest neighbours from the BLASTN database to clone sequences for the archaeal 16S rRNA gene libraries for the salt lakes are shown. The clone numbers have been colour coded according to the environment they are from (**Red** = BJ1, **Blue** = CG1, **Orange** = EJ3, **Magenta** = EN1, **Cyan** = SH1, **Grey** = XH2).

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
1	*Halorubrum* sp. AUS-1 (D32081)	99	Western Australia (Mukohata *et al.* 1988)	4/ 44
2	Haloarchaeon CSW4.22.4 (AY498649)	99	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	1/ 44
3	Haloarchaeon CSW4.22.4 (AY498649)	98	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	22/ 44
4	Uncultured archaeon clone ss033c (AJ969781)	96	British Columbia, Canada (Walsh *et al.* 2005)	8/ 44
5	*Halorubrum saccharovorum* (U17364)	94	San Francisco Bay (Tomlinson and Hochstein 1976)	3/ 44
6	Haloarchaeon CSW4.22.4 (AY498649)	96	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	1/ 44
7	Uncultured haloarchaeon clone Sec16SC8 (DQ071588)	95	Adriatic solar saltern (Pašić *et al.* 2005)	1/ 44
8	Uncultured archaeon ORGANIC4_A (AF142984)	95	Vestfold Hilds, eastern Antarctica (Bowman *et al.* 2000)	1/ 44
9	Haloarchaeon CSW4.22.4 (AY498649)	94	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	1/ 44
10	Uncultured archaeon ORGANIC5_A (AF142983)	94	Vestfold Hilds, eastern Antarctica (Bowman *et al.* 2000)	12/ 44
11	Uncultured haloarchaeon clone HC10 (AY524141)	95	San Diego, USA (Bidle *et al.* 2005)	7/ 44
12	*Halorubrum tibetense* (AF435111)	96	Lake Zabuye, Tibet (Fan *et al.* 2004)	2/ 44
13	Uncultured archaeon clone ss049a (AJ969815)	97	British Columbia, Canada (Walsh *et al.* 2005)	2/ 44
14	Uncultured archaeon clone ss057a (AJ969912)	98	British Columbia, Canada (Walsh *et al.* 2005)	1/ 44
15	*Halorubrum tibetense* (AF435111)	97	Lake Zabuye, Tibet (Fan *et al.* 2004)	11/ 44

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 15 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
16	Uncultured archaeon clone ss_032a (AJ969845)	89	British Columbia, Canada (Walsh *et al.* 2005)	1/ 44
17	*Halorubrum vacuolatum* (D87972)	95	Lake Magadi, Kenya (Mwatha and Grant 1993)	3/ 44
154	Uncultured haloarchaeon clone ZB-A56 (AF505709)	95	Zabuye Lake, Tibet (Fan *et al.* unpublished)	2/ 44
155	Uncultured archaeon clone LCKW-A3 (DQ129955)	93	Lake Chaka, north western China (Jiang *et al.* 2006)	1/ 44
156	*Halorubrum vacuolatum* (D87972)	96	Lake Magadi, Kenya (Mwatha and Grant 1993)	2/ 44
19	Uncultured archaeon clone 11-UMH 8% pond (AF477906)	97	'Bras del Port', Santa Pola, Alicante, Spain (Benlloch *et al.* 2002)	4/ 25
20	Haloarchaeon CWS4.22.4 (AY498649)	96	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	2/ 25
22	Uncultured archaeon clone 32-UMH 31% pond (AF477937)	97	'Bras del Port', Santa Pola, Alicante, Spain (Benlloch *et al.* 2002)	3/ 25
23	Uncultured haloarchaeon clone ZB-A53 (AF505707)	90	Zabuye Lake, Tibet (Fan *et al.* unpublished)	2/ 25
24	Uncultured archaeon clone ss_019a (AJ969885)	92	British Columbia, Canada (Walsh *et al.* 2005)	1/ 25
25	*Halosimplex carlsbadense* (AJ586107)	90	Permian halite deposit, south eastern Mexico (Vreeland *et al.* 2002)	1/ 25
26	Haloarchaeon CSW2.24.4 (AY498650)	95	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	1/ 25
27	Uncultured archaeon clone ss007 (AJ969930)	94	British Columbia, Canada (Walsh *et al.* 2005)	4/ 47
28	Haloarchaeon CSW6.14.5 (AY498648)	96	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	13/ 47
29	Uncultured archaeon clone ss_048 (AJ969840)	93	British Columbia, Canada (Walsh *et al.* 2005)	1/ 47
30	Uncultured archaeon clone HW11 (AJ344308)	97	Permo-Triassic rock salt (Wieland *et al.* unpublished)	2/ 47

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 15 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
31	*Halobacterium* sp. NCIMB 720 (AB074563)	92	? (Kamekura and Mizuki Published only in database (2001))	2/ 47
32	*Halorubrum* AUS-1 (D32081)	98	Western Australia (Mukohata *et al.* 1988)	1/ 47
33	Uncultured archaeon clone ss007 (AJ969930)	94	British Columbia, Canada (Walsh *et al.* 2005)	4/ 47
34	Uncultured archaeon clone ss007 (AJ969930)	93	British Columbia, Canada (Walsh *et al.* 2005)	1/ 47
35	Uncultured archaeon clone ss_048 (AJ969840)	94	British Columbia, Canada (Walsh *et al.* 2005)	1/ 47
36	Uncultured archaeon MSP41 (AB012059)	87	Lake Magadi, Kenya (Grant *et al.* 1999)	1/ 47
37	Uncultured archaeon clone ss057j (AJ969823)	97	British Columbia, Canada (Walsh *et al.* 2005)	5/ 47
38	Uncultured haloarchaeon clone ZB-A56 (AF505709)	91	Zabuye Lake, Tibet (Fan *et al.* unpublished)	2/ 47
39	Uncultured archaeon clone ss035 (AJ969859)	97	British Columbia, Canada (Walsh *et al.* 2005)	3/ 47
157	Uncultured euryarchaeote clone DGGE Band 26 (EM071497)	90	Monegros Desert (Casamayor unpublished-a)	2/ 47
158	Uncultured archaeon clone ss_032a (AJ969845)	98	British Columbia, Canada (Walsh *et al.* 2005)	1/ 47
159	*Halorubrum* sp. F100 (DQ309090)	98	Salt lake and Bolluk lake, Turkey (Ozcan *et al.* unpublished)	3/ 47
160	Uncultured haloarchaeon MSP41 (AB012059)	86	Lake Magadi, Kenya (Grant *et al.* 1999)	1/ 47
40	*Halorubrum saccharovorum* (U17364)	94	San Francisco Bay (Tomlinson and Hochstein 1976)	15/ 48
41	*Halorubrum* sp. AUS-1 (D32081)	98	Western Australia (Mukohata *et al.* 1988)	1/ 48
42	Uncultured archaeon clone ss033c (AJ969781)	96	British Columbia, Canada (Walsh *et al.* 2005)	8/ 48
43	Uncultured haloarchaeon MSP41 (AB012059)	90	Lake Magadi, Kenya (Grant *et al.* 1999)	1/ 48

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 15 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
44	Uncultured archaeon clone ss057u (AJ969856)	98	British Columbia, Canada (Walsh *et al.* 2005)	3/ 48
45	Haloarchaeon CSW4.22.4 (AY498649)	99	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	5/ 48
46	Uncultured haloarchaeon clone ZB-A53 (AF505707)	91	Zabuye Lake, Tibet (Fan *et al.* unpublished)	1/ 48
47	Uncultured haloarchaeon clone ZB-A53 (AF505707)	91	Zabuye Lake, Tibet (Fan *et al.* unpublished)	2/ 48
48	Uncultured haloarchaeon clone HC10 (AY524141)	94	San Diego, USA (Bidle *et al.* 2005)	1/ 48
49	*Halorubrum saccharovorum* (U17364)	94	San Francisco Bay (Tomlinson and Hochstein 1976)	1/ 48
161	Uncultured haloarchaeon clone ZB-A53 (AF505707)	91	Zabuye Lake, Tibet (Fan *et al.* unpublished)	1/ 48
162	Uncultured haloarchaeon clone ZB-A53 (AF505707)	88	Zabuye Lake, Tibet (Fan *et al.* unpublished)	1/ 48
50	Haloarchaeon CSW2.27.5 (AY498646)	98	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	4/ 44
51	*Halorubrum xinjiangense* (AY510707)	99	Xiao-Er-Kule Lake, Xinjiang, China (Feng *et al.* 2004)	1/ 44
52	Haloarchaeon CSW2.27.5 (AY498646)	99	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	6/ 44
53	Haloarchaeon CSW2.27.5 (AY498646)	99	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	1/ 44
54	Haloarchaeon CSW4.22.4 (AY498649)	99	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	4/ 44
55	Uncultured archaeon clone ss033c (AJ969781)	97	British Columbia, Canada (Walsh *et al.* 2005)	7/ 44
56	Haloarchaeon CSW2.27.5 (AY498646)	99	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	11/ 44
57	Uncultured haloarchaeon clone ZB-A53 (AF505707)	90	Zabuye Lake, Tibet (Fan *et al.* unpublished)	1/ 44
58	*Haloarcula* sp. (AB010964)	94	? (Oren *et al.* 1999)	1/ 44

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 15 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
59	Uncultured haloarchaeon clone ZB-A53 (AF505707)	88	Zabuye Lake, Tibet (Fan *et al.* unpublished)	1/ 44
60	Uncultured archaeon clone DIE3 (AJ308007)	92	Potassium mine, Germany (Ochsenreiter *et al.* 2002)	2/ 44
61	*Halorubrum terrestre* (AB090169)	93	Saline soils (Ventosa *et al.* 2004)	3/ 44
62	Haloarchaeon CSW2.27.5 (AY498646)	97	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	1/ 44
63	Haloarchaeon CSW8.8.11 (AY498645)	97	Corio Bay, Victoria, Australia (Burns *et al.* 2004)	1/ 44

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Phylogenetics of *Archaea* in the Salt Lakes

Phylogenetic analysis shows the distribution of sequences into eight monophyletic assemblages within the order *Halobacteriales*. Many branch within the *Halorubrum*, *Halosimplex*, *Haloarcula*, *Halobacterium*, *Halobaculum*, *Halogeometricum*, *Halomicrobium* and the alkaliphilic group *Natronomonas*. However, there are additional lineages that form between these nodes, designated Salt Lake (SL) Clusters 1 to 3. This is summarised in Figure 30. All known genera within the order *Halobacteriales* are represented in the tree; however, none of the sequences in the SL Clusters affiliate closely with any of them. There appears to be good bootstrap support for the branches *Halorubrum* (82%) and *Natronomonas* (100%). All the SL Clusters are also well supported (98% - 100%), with SL Cluster 3 forming an outer group to the *Halobacteriales*.

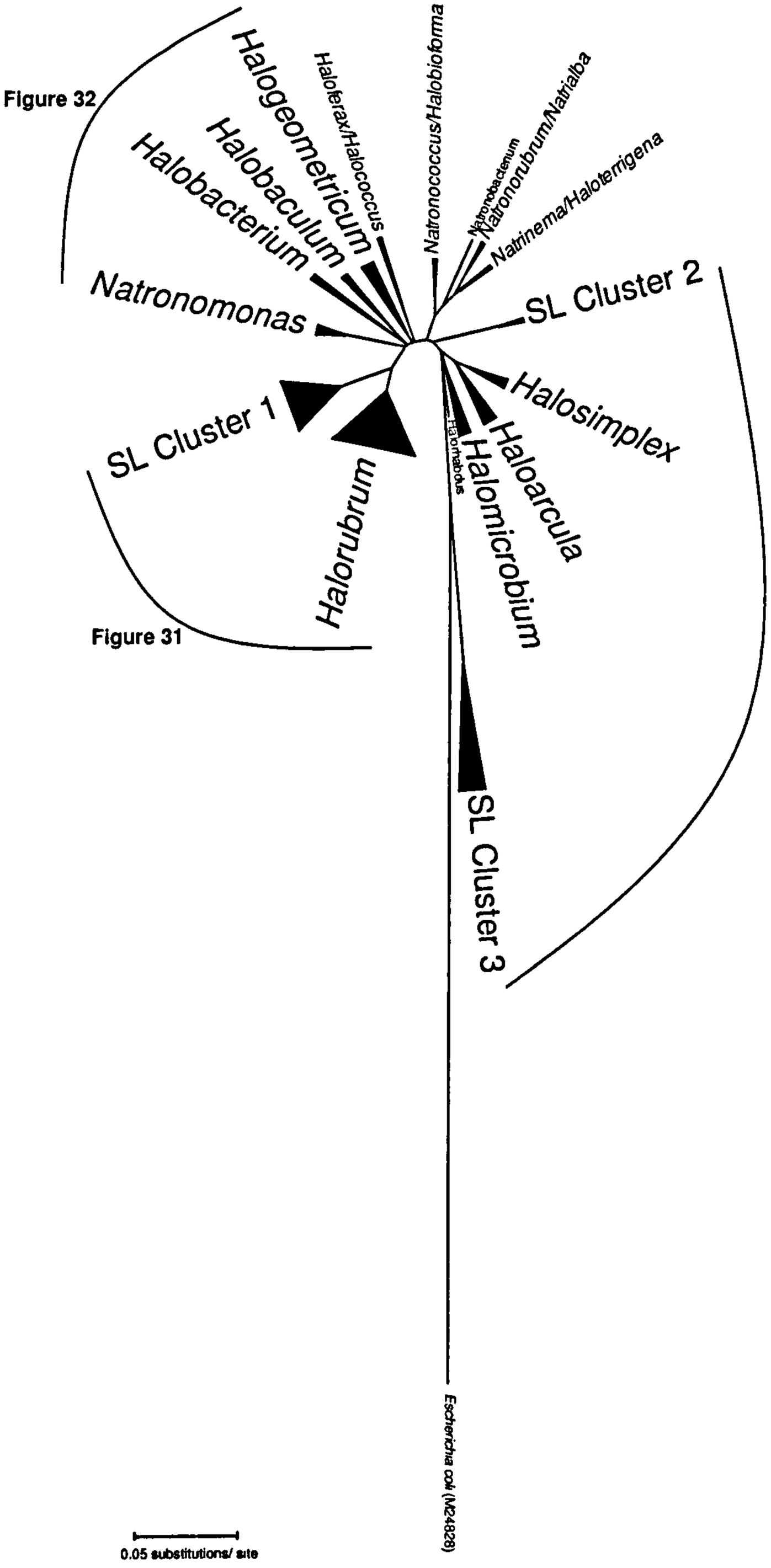


Figure 32

Figure 31

Figure 33

Figure 30 Phylogenetic Tree Summarising all *Archaea* found in the Salt Lakes.

Halorubrum

34% of clone sequences branch with the *Halorubrum*, therefore comprising the largest group detected in all six salt lakes. This group is shown in more detail in Figure 31. One clone sequence from BJ1 (5) and two from SH1 (40 and 49) are related to *Halorubrum saccharovorum* (all at 94%). The 16S rRNA gene of *Halorubrum saccharovorum* is the closest relative to the dominant clone found in the 16S rRNA library for SH1, showing a clone frequency of 15. *Halorubrum saccharovorum* was first isolated by Tomlinson and Hochstein in 1976 from a mixture of mud and brine from a saltern in the southern section of San Francisco Bay. This isolate is a strict aerobe, requiring 1.5 - 5.2 M NaCl for growth, with an optimum concentration of 3.5 - 4.5 M (Grant *et al.* 2001). Two clone sequences from CG1 (12 and 15) are closely related to *Halorubrum tibetense* (96% and 97% respectively). The gene of *Halorubrum tibetense* is the closest relative to the second most frequent clone in the library for CG1, with a clone frequency of 11. This is a haloalkaliphilic archaeon isolated from Lake Zabuye on the Tibetan Plateau, which is an alkaline chloride-sulphate salt lake with a pH of 9.4 (Fan *et al.* 2004). Two more clone sequences from CG1 (17 and 156) are related to *Halorubrum vacuolatum* (95% and 96% respectively). This is also an obligate haloalkaliphilic archaeon requiring an optimum pH of 9.5 for growth (Grant *et al.* 2001). *Halorubrum vacuolatum* was first isolated from Lake Magadi in Kenya, an alkaline soda lake with a pH of 11 (Mwatha and Grant 1993). One clone sequence from XH2 (51) is closely related to *Halorubrum xinjiangense* (99%), which was isolated from Xiao-Er-Kule Lake in Xinjiang, China and grows optimally at 3.1 - 3.4 M NaCl and pH 7.0 - 7.5 (Feng *et al.* 2004). Another clone sequence from XH2 (61) is related to *Halorubrum terrestre* (93%). Isolated from saline soils, *Halorubrum terrestre* grows optimally at 25% (w/v) NaCl and pH 7.5 (Ventosa *et al.* 2004). One clone sequence from EN1 (159) is closely related to *Halorubrum* sp. F100 (98%). This was found in saline lakes in Turkey (Ozcan *et al.* unpublished). Clone sequences from BJ1 (1), EN2 (32) and SH1 (41) are closely related to *Halorubrum* AUS-1, which was isolated from a nameless clay pan in Western Australia. This isolate grows on medium containing 25% (w/v) NaCl at pH 7.4 (Mukohata *et al.* 1988).

The remaining clone sequences are related to uncultured isolates that cluster with the *Halorubrum*. Four clone sequences from CG1, EN1 and SH1 (14, 37, 39 and 44) are related to uncultured archaeon clones ss057_a, ss057j, ss057u and ss035, which were sequences detected in saline soils in British Columbia, Canada at distances between 0 cm to 300 cm from the salt spring source exhibiting 7% (w/v) NaCl, and pH 7 to 10. The phylogenetic tree

in that study similarly placed clones ss057_a, ss057j, ss057u and ss035 with the *Halorubrum* (Walsh *et al.* 2005). Five sequences from BJ1, EJ3, SH1 and XH2 (2, 3, 20, 45 and 54) are related to haloarchaeon clone CSW4.22.4 (96% - 99%), which was an isolate from crystalliser ponds of Corio Bay in Victoria, Australia, which exhibited a total salt concentration of 33% (w/v) and a pH of 8.1 (Burns *et al.* 2004) The 16S rRNA gene of this isolate is the closest relative to the dominating clone found in the 16S rRNA library for BJ1, with a clone frequency of 22.

Natronomonas

This group of haloalkaliphiles grow at 2 - 5.2 M NaCl and pH 7 - 10 (Grant *et al.* 2001). Only two clone sequences from the libraries are found in this group. One sequence from CG1 (13) is closely related to an uncultured clone ss049a (97%), which was also detected in saline soils in British Columbia, which was previously described. Phylogenetic analysis in that study showed this clone to be distantly related to *Natronomonas pharaonis* (Walsh *et al.* 2005). One clone sequence from XH2 (63) is closely related to an uncultured clone CSW8.8.11 (97%), which was from Australian crystalliser ponds in Corio Bay, which was previously described. Although both clone sequences (13 and 63) are related to uncultured archaea, they are distributed within the same clade as *Natronomonas pharaonis* (Figure 32) showing 95% and 94% identity respectively to *Natronomonas pharaonis*.

Halobacterium

Only one clone sequence from EJ3 (24) clusters within the same clade as *Halobacterium* (Figure 32). It is related to an uncultured clone ss_019a (92%), found in saline soils in British Columbia, which was previously described. Clone sequence 24 is also distantly related to *Halobacterium salinarium* showing 87% identity.

Halobaculum

One clone sequence from CG1 (154) is related to an uncultured haloarchaeon clone ZB-A56 (95%), which is from Lake Zabuye (as previously described). These sequences are distributed within the same clade as *Halobaculum gomorrense* (Figure 32). Clone sequence 154 is distantly related to *Halobaculum gomorrense*, showing 88% identity.

Halogeometricum

This genus currently has just one recognised species, *Halogeometricum borinquense*, which is distantly related to *Haloferax* (Montalvo-Rodriguez *et al.* 1998). Three clone sequences appear in the *Halogeometricum* clade (Figure 32). Clone sequence 26 from EJ3 is related to clone CSW2.24.4, which is a sequence found in an Australian crystalliser pond in Corio Bay, which was previously described. Phylogenetic analysis in that study similarly placed this clone near to, but distantly related to *Halogeometricum borinquense* (Burns *et al.* 2004). One clone sequence from EN1 (30) is closely related to uncultured archaeon clone HW11 (97%), which is a sequence found in Permo-Triassic rock salt (Wieland *et al.* unpublished). Clone sequence 25 (from EJ3) showed similarity to *Halosimplex*, but phylogenetic analysis places this sequence near *Halogeometricum*. Clone sequences 26 and 30 form separate branches within the *Halogeometricum* clade with low bootstrap values (see Figure 32), suggesting that they are only distantly related; they only show 88 - 89% identity to *Halogeometricum borinquense*.

Haloarcula

Three clone sequences affiliate with the *Haloarcula* (Figure 33). One clone sequence from BJ1 (7) is related to an uncultured clone Sec16SC8 (95%), which was a sequence found in the crystallisers of an Adriatic solar saltern. This had a water activity of 0.759 at 30°C (the water activity for a saturated NaCl solution at 30°C is 0.769) and was at pH 8.0. The phylogenetic tree in that study also showed clone Sec16SC8 affiliated with *Haloarcula* (Pašić *et al.* 2005). One clone sequence from XH2 (58) is related to a *Haloarcula* sp. (94%). One sequence from EN2 (31) is related to *Haloarcula* sp. NCIMB 720 (92%). This is an extreme halophile whose sequence was only published in the NCBI database (Kamekura and Mizuki Published only in database (2001)).

Halosimplex

The *Halosimplex* are a deep branch distantly related to *Haloarcula* (Figure 33). They are unique in that they have three different genes encoding their 16S rRNA. The strain *Halosimplex carlsbadense* was isolated from a Permian halite deposit in south eastern Mexico. It requires a defined medium for growth containing glycerol and acetate or pyruvate (Vreeland *et al.* 2002). Three clone sequences from EN2 (27, 33, 34) branch with the *Halosimplex* clade. Although related to an uncultured archaeon ss007, which was detected in

saline soils in British Columbia (93% - 94%), they all show 93% identity to *Halosimplex carlsbadense* strain 2-9-1(T).

Halomicrobium

This genus currently contains just one species, *Halomicrobium mukohataei*, which was isolated from salt flats in Argentina and was originally somewhat related to the *Haloarcula*. However, analysis of its morphology, polar lipid composition and 16S rRNA gene sequence showed that it was significantly different from *Haloarcula* and was proposed as a new genus within the *Halobacteriales*. It is an extremely halophilic aerobe and a facultative anaerobe in the presence of nitrate (Oren *et al.* 2002). Three clone sequences appear in the same clade as *Halomicrobium mukohataei* (Figure 33). One clone sequence from CG1 (155) is related to an uncultured clone LCKW-A3 (93%), which was from Lake Chaka. This athalassohaline lake possesses a total salinity of 32.5% (w/v) with a water temperature between -8°C to 4.2°C in the winter and 6°C to 20°C in the summer. It also had a pH of 7.4 (Jiang *et al.* 2006). Two clone sequences from EN1 (29 and 35) are related to an uncultured clone ss_48 (93% and 94%), which was from saline soils in British Columbia, Canada, which was previously described. That study also placed clone ss_48 near the *Haloarcula-Halomicrobium* branch (Walsh *et al.* 2005).

Novel Lineages

These novel lineages have been designated Salt Lake (SL) Clusters 1 to 3 as shown in Figure 31, Figure 32 and Figure 33.

Branching near the *Halorubrum* group, SL Cluster 1 (Figure 31) is the second largest group comprising 27% of the clone sequences. Several clone sequences from XH2 (50, 52, 53, 56 and 62) and one from EN1 (28) are closely related to uncultured clones from Australian crystalliser ponds in Corio Bay (previously described). Clone sequences related to haloarchaeon clone 6.14.5 is the most frequently observed clone in the library for EN1, with a clone frequency of 13. Clone sequences related to haloarchaeon clone 2.27.5 are also the dominating clones in the library for XH2, with a clone frequency of 11. Clone sequences from BJ1 (4), CG1 (16), EN2 (158), SH1 (42) and XH2 (55) are related to clones from saline soils in British Columbia, Canada (previously described). Clone sequences from BJ1 (8) and CG1 (10) are related to uncultured clones ORGANIC4_A and ORGANIC5_A (95% and 94% respectively), which were found in anoxic sediments of Organic Lake in the Vestfold Hilds in

eastern Antarctica. This lake had a total salinity of 20% (w/v) and a temperature of -7°C (Bowman *et al.* 2000). Clone sequences related to ORGANIC5_A are the most frequently occurring in the 16S rRNA library for CG1, with a clone frequency of 12. One clone sequence from XH2 (60) is related to an uncultured archaeon DIE3 (92%), which was detected in a hypersaline pond from the bottom of a slag heap of a former potassium mine in Germany. This pond had a pH of 5.6 with a total salinity above 32% (w/v). Phylogenetic analysis in that study also placed this clone in an unknown cluster that formed an outer group to *Halorubrum* (Ochsenreiter *et al.* 2002). One clone sequence from CG1 (11) and one sequence from SH1 (48) are related to a haloarchaeon clone HC10 (95% and 94% respectively), which was detected in a solar saltern in San Diego, USA (Bidle *et al.* 2005). One clone sequence from EJ3 (19) is related to an uncultured clone 11-UHM 8% pond (97%), which was detected in a multipond solar saltern 'Bras del Port' located in Alicante, Spain, that had a salinity of 8% (w/v) total salinity (Benlloch *et al.* 2002)

SL Cluster 2 is a small group branching within the *Halobacteriales* (Figure 33). It consists of just one clone sequence from EJ3 (22), which is related (93%) to an uncultured archaeon from a solar saltern in Alicante, Spain, showing 31% (w/v) total salinity (Benlloch *et al.* 2002).

SL Cluster 3 appears to form its own separate lineage to the order *Halobacteriales.* Several sequences from EJ3 (23), EN2 (38), SH1 (46, 47, 161 and 162) and XH2 (57and 59) are related (88% - 91%) to sequences detected in Lake Zabuye (Fan *et al.* unpublished). Two sequences from EN2 (36 and 160) and one from SH1 (43) are related to an uncultured haloarchaeon clone MSP41, which was detected in a salt crystallising pond at Lake Magadi, Kenya. Similarly, phylogenetic analysis in that study positioned this clone in a novel lineage on the periphery of known haloarchaea, suggesting that this clone was a member of a deeply branching group of the *Euryarchaeota* (Grant *et al.* 1999). Phylogenetic analysis has shown that this novel lineage is also distinct from known methanogens, and members of the *Crenarchaeota*, *Korarchaeota* and nanoarchaea (Figure 33).

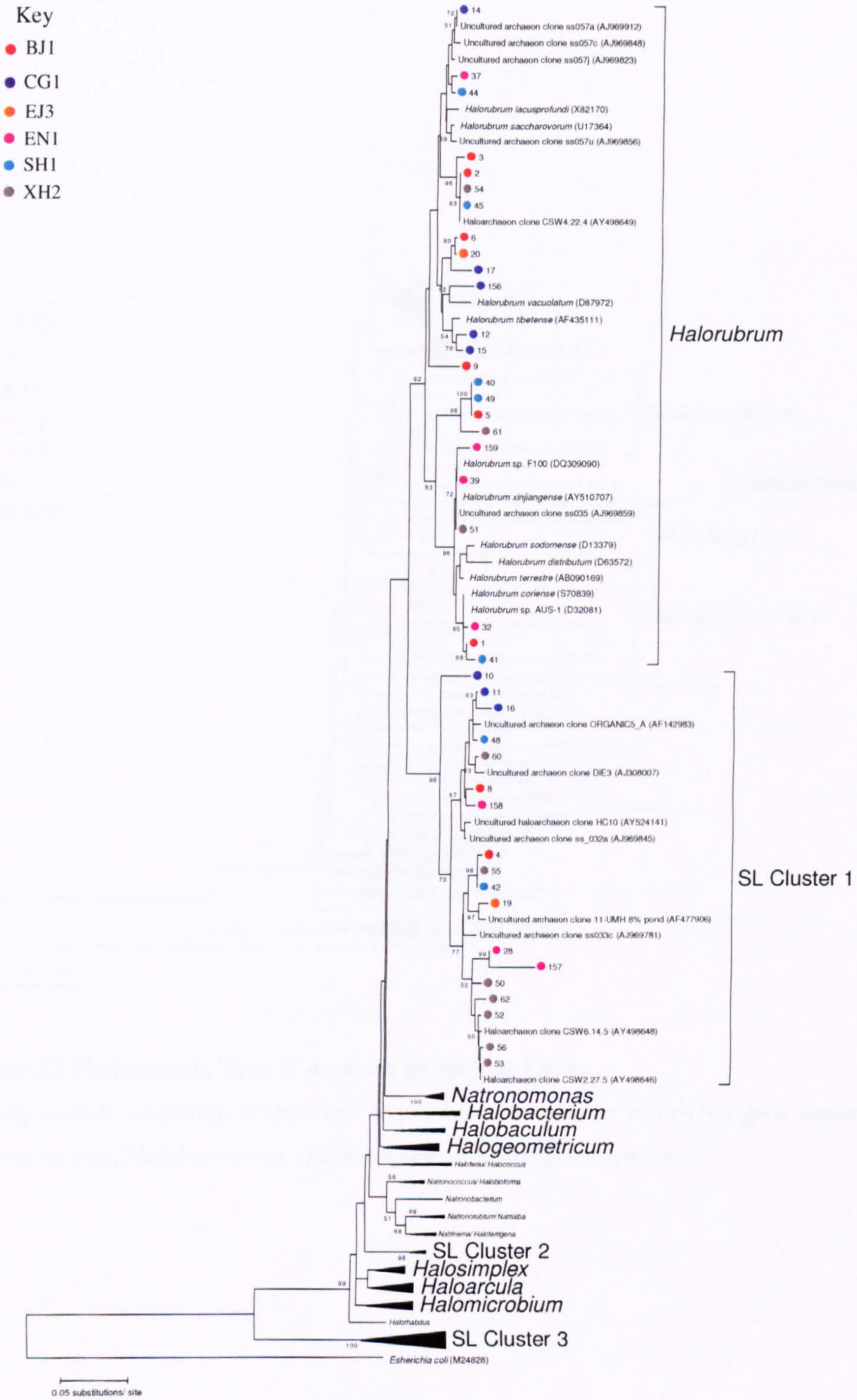


Figure 31 Phylogenetic Tree of *Archaea* in the Salt Lakes.

Phylogenetic relationships of the clone sequences with existing 16S rRNA gene sequences for *Halorubrum* and uncultured isolates.

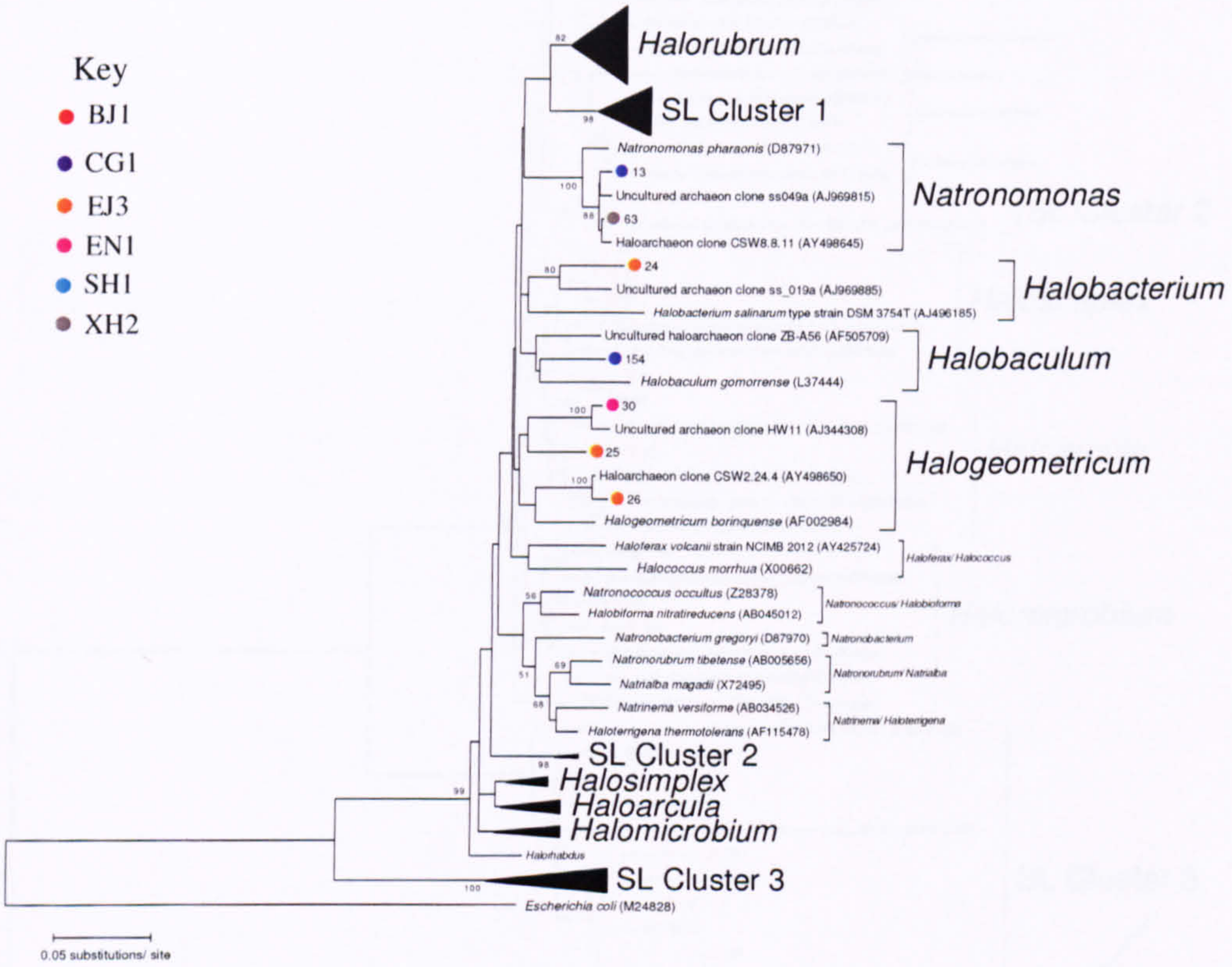


Figure 32 Phylogenetic Tree of *Archaea* in the Salt Lakes.

Phylogenetic relationships of the clone sequences with existing 16S rRNA gene sequences for *Natronomonas*, *Halobacterium*, *Halobaculum* and *Halogeometricum*.

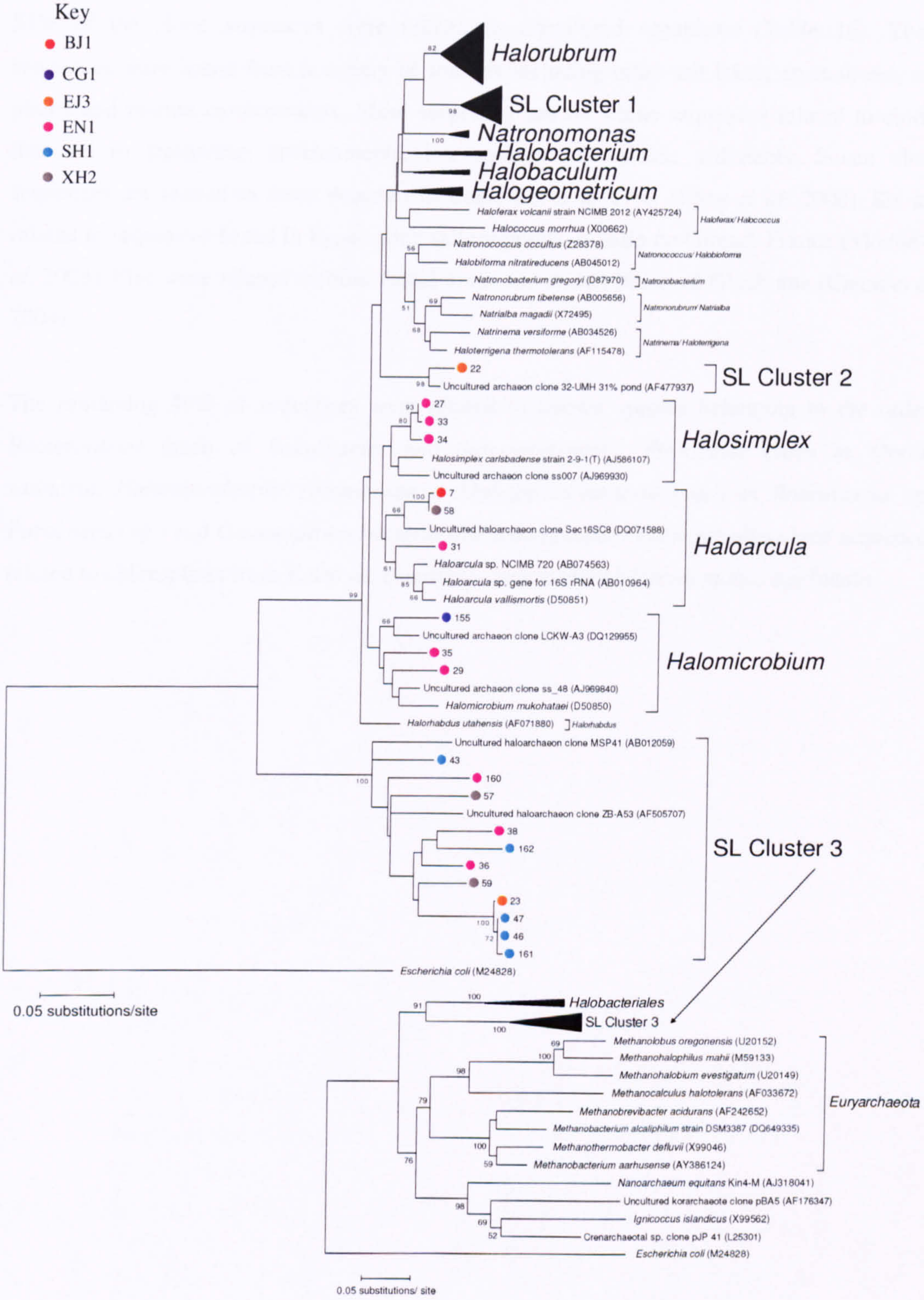


Figure 33 Phylogenetic Tree of *Archaea* in the Salt Lakes.

Phylogenetic relationships of the clone sequences with existing 16S rRNA gene sequences for *Halosimplex*, *Haloarcula*, *Halomicrobium* and other uncultured isolates.

Bacterial Diversity of the Salt Lakes

51% of the clone sequences were related to uncultured organisms (Table 16). These sequences were found from a variety of sources including other salt lakes, crystallisers, salt plains and marine environments. More surprising are the clone sequences related to clones detected in freshwater environments, hot springs and anoxic sediments. Seven clone sequences are related to those detected in Lake Chaka in Tibet (Jiang *et al.* 2006). Six are related to sequences found in hypersaline saltern ponds in Salin-de-Giraud, France (Mouné *et al.* 2003). Five were related to those found in the Great Salt Plains of Oklahoma (Caton *et al.* 2004).

The remaining 49% of sequences were related to known species belonging to the orders *Bacteroidetes* (such as *Salinibacter* and *Salegentibacter*), *Firmicutes* (such as *Orenia salinaria*, *Halanaerobacter chitinivorans*), *Alphaproteobacteria* (such as *Roseovarius* sp., *Paracoccus* sp.) and *Gammaproteobacteria* (the halomonads). There are also clone sequences related to chloroplasts from *Eukarya*; *Dunaliella salina* and *Chlamydomonas applanata*.

Table 16 Sequencing Results for the *Bacteria* in Salt Lakes.

Nearest neighbours from the BLASTN database to clone sequences for the bacterial 16S rRNA gene libraries for the salt lakes are shown. The clone numbers have been colour coded according to the environment they are from (**Red** = BJ1, **Blue** = CG1, **Orange** = EJ3, **Magenta** = EN1, **Cyan** = SH1, **Grey** = XH2).

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
64	Uncultured bacterium clone ZB118 (AY327199)	89	Mesophilic sulphide-rich spring (Elshahed *et al.* 2003)	1/ 13
65	*Halomonas variabilis* strain GSP3 (AY505527)	97	Great salt plains, Oklahoma, USA (Caton *et al.* 2004)	1/ 13
66	*Halomonas salina* strain GSP33 (AY505525)	99	Great salt plains, Oklahoma, USA (Caton *et al.* 2004)	1/ 13
67	Uncultured *Aquificales* bacterium clone pKB (AF453506)	94	Kamchatka, Russia (Takacs-Vesbach *et al.* unpublished)	1/ 13
70	*Salinibacter ruber* strain POLA 18 (AF323502)	94	Salterns, Santa Pola, Alicante, Spain (Anton *et al.* 2002)	1/ 13
72	Uncultured freshwater bacterium clone 965005D05.xl (DQ065453)	96	Artificial mesocosms (Horner-Devine *et al.* 2003)	1/ 13
73	Uncultured bacterium isolate JH10_C67 (AY568819)	90	Ganghwa island, Korea (Cho *et al.* unpublished-b)	1/ 13
74	Uncultured eubacterium clone LKB122 (AJ746517)	95	Gouzikeng Landfill, China (Huang *et al.* 2004)	1/ 13
75	Uncultured bacterium gene clone: ODP1251B5.13 (AB177341)	93	Pacific Ocean margins (Inagaki *et al.* 2006)	1/ 13
76	*Halomonas campisalis* (AF054286)	97	Alkali Lake, Washington, USA (Mormile *et al.* 1999)	1/ 34
77	*Halanaerobiaceae* bacterium SLAS-1 (AY965613)	90	Searles Lake, Mojave Desert (Oremland *et al.* 2005)	3/ 34
78	*Halanaerobium* sp. strain AN-BI5B (AM157647)	97	Bannock basin (Daffonchio *et al.* 2006)	1/ 34
79	*Paracoccus* sp. mdw-1 (AY819696)	95	? (Xu unpublished)	3/ 34
80	Gammaproteobacterium M12-20B (AY730235)	95	Mono Lake, California, USA (Bano and Hollibaugh unpublished)	3/ 34

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 16 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
81	*Natronoanaerobium halophilum* isolate G-M14CH-4 (AJ271451)	93	Soda lakes (Jones *et al.* 1998)	1/ 34
82	Phototrophic bacterium (BN 9624) (X93474)	97	? (Imhoff and Suling 1996)	1/ 34
83	Uncultured candidate division OP11 clone OPB92 (AF027030)	88	Yellowstone National Park, Wyoming, USA (Hugenholtz *et al.* 1998)	1/ 34
84	Uncultured alphaproteobacterium clone G4 (DQ003190)	94	Subsurface microbial community (Boyd *et al.* unpublished)	1/ 34
85	*Halomonas* sp. 3019 (AM110974)	97	East Pacific (Zeng unpublished)	3/ 34
86	Uncultured bacterium clone E4aG09 (DQ103651)	92	Hypersaline endoevaporitic microbial mat, Israel (Sørensen *et al.* 2005)	1/ 34
87	Uncultured low G+C Gram-positive bacterium clone mixed culture C-1 (DQ206426)	98	Mono Lake, California, USA (Hollibaugh *et al.* 2006)	3/ 34
88	*Natronoanaerobium halophilum* isolate G-M14CH-4 (AJ271451)	95	Soda lakes (Jones *et al.* 1998)	2/ 34
89	Arctic seawater bacterium Bsw20461 (DQ064620)	99	Arctic seawater (Zeng *et al.* unpublished-b)	1/ 34
90	Uncultured bacterium clone CSR-16 (AY699368)	91	Hot springs, north western Iceland (Hobel *et al.* 2005)	3/ 34
91	*Deltaproteobacterium* MLMS-1 (AY459365)	92	Mono Lake, California, USA (Hoeft *et al.* 2004)	2/ 34
92	*Natronoanaerobium halophilum* isolate G-M14CH-4 (AJ271451)	96	Soda lakes (Jones *et al.* 1998)	1/ 34
93	Uncultured bacterium clone E6aF12 (DQ103656)	95	Hypersaline endoevaporitic microbial mat, Israel (Sørensen *et al.* 2005)	1/ 34
163	*Halomonas* sp. A-07	99	Soda lakes, Tanzania (Pradhan *et al.* unpublished)	2/ 34
94	*Salinibacter* sp. 5Sm6 (AY987850)	99	Maras salterns, Peruvian Andes (Maturrano *et al.* 2006)	7/ 41
95	Uncultured bacterium clone LCKW-B4 (DQ129879)	99	Lake Chaka, China (Jiang *et al.* 2006)	3/ 41

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 16 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
96	Uncultured bacterium clone LCKW-B3 (DQ129878)	99	Lake Chaka, China (Jiang *et al.* 2006)	2/ 41
97	*Chlamydomonas applanata* 16S rRNA gene; chloroplast gene for chloroplast product (AF394204)	95	? (Visviki and Santikul 2000)	2/ 41
98	*Halovibrio denitrificans* strain HGD 3 (DQ072718)	98	Hypersaline lake, Mongolia (Sorokin *et al.* 2006)	1/ 41
99	*Salegentibacter* sp. 18III/A01/068 (AY576719)	97	Mediterranean Sea (Agogué *et al.* 2005)	1/ 41
100	Uncultured bacterium clone LCKS0-B3 (DQ129888)	98	Lake Chaka, China (Jiang *et al.* 2006)	1/ 41
164	Halophilic eubacterium EHB-2 (AJ242998)	99	'Bras del Port', Santa Pola, Alicante, Spain (Antón *et al.* 2000)	24/ 41
101	Uncultured CFB group bacterium isolate SA-20 (AJ495620)	94	Salin-de-Giraud, France (Mouné *et al.* 2003)	2/ 45
102	Glacial ice bacterium G500K-19 (AF479330)	98	Glacial environments (Christner 2002)	1/ 45
103	*Alkalilimnicola halodurans* isolate 34Alc (AJ404972)	94	Lake Natron, east Africa Rift Valley (Yakimov *et al.* 2001)	1/ 45
104	Uncultured bacterium clone E6aH12 (DQ103652)	88	Hypersaline endoevaporitic microbial mat, Israel (Sørensen *et al.* 2005)	1/ 45
105	Uncultured bacterium VC2.1 Bac16 (AF068794)	88	*In situ* growth chamber, Mid-Atlantic Ridge hydrothermal vent (Reysenbach *et al.* 2000c)	1/ 45
106	*Dunaliella salina* 16S rRNA gene; chloroplast gene for chloroplast product (AF547096)	99	Salterns, southern France (Oren 2005)	2/ 45
107	Uncultured proteobacterium clone BSB3-217m (AY360507)	88	Black Sea (Vetriani *et al.* 2003)	2/ 45
108	*Dunaliella salina* 16S rRNA gene; chloroplast gene for chloroplast product (AF547096)	99	Salterns, southern France (Oren 2005)	1/ 45

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 16 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
109	*Orenia salinaria* (Y18485)	88	Salin-de-Giraud, France (Mouné *et al.* 2000)	1/ 45
110	*Dunaliella salina* 16S rRNA gene; chloroplast gene for chloroplast product (AF547096)	97	Salterns, southern France (Oren 2005)	19/ 45
111	*Halomonas salina* strain GSP33 (AY505525)	97	Great salt plains, Oklahoma, USA (Caton *et al.* 2004)	4/ 45
112	Uncultured alphaproteobacterium clone BIxi22 (AJ318202)	90	Industrial biofilter, Germany (Friedrich *et al.* 2002)	1/ 45
113	Uncultured deltaproteobacterium isolate SB-15 (AJ495661)	95	Salin-de-Giraud, France (Mouné *et al.* 2003)	2/ 45
114	Uncultured bacterium 'KTK 32' (AJ133617)	91	Kebrit Deep, Red Sea (Eder *et al.* 1999)	1/ 45
115	*Dunaliella salina* 16S rRNA gene; chloroplast gene for chloroplast product (AF547096)	98	Salterns, southern France (Oren 2005)	1/ 45
165	Uncultured bacterium clone At12OctA1 (AY862782)	95	Salar de Atacama, northern Chile (Demergasso *et al.* unpublished)	1/ 45
166	*Salinibacter ruber* strain POLA 13 (AF323503)	93	Salterns, Santa Pola, Alicante, Spain (Anton *et al.* 2002)	1/ 45
167	*Flexibacter tractuosus* strain:IFO 15989 (AB078072)	87	Marine (Nakagawa *et al.* 2002)	1/ 45
116	Uncultured bacterium clone CH-B4 (AY280399)	84	*Paralvinella sulfincola* tube and the adjacent substratum (Page *et al.* 2004)	3/ 40
118	*Halomonas variabilis* strain GSP3 (AY505527)	98	Great salt plains, Oklahoma, USA (Caton *et al.* 2004)	1/ 40
119	Uncultured bacterium 'KTK 32' (AJ133617)	97	Kebrit Deep, Red Sea (Eder *et al.* 1999)	1/ 40
120	Uncultured bacterium clone: UTFS-OF09-d22-24 (AB200299)	92	Activated sludge (Oe *et al.* unpublished)	1/ 40

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 16 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
121	*Vibrio metschnikovii* (NCTC 11170) (X74712)	99	Aquatic environments (Wallet *et al.* 2005)	1/ 40
122	Uncultured bacterium clone LCKS0-B7 (DQ129890)	97	Lake Chaka, China (Jiang *et al.* 2006)	1/ 40
123	*Dunaliella salina* 16S rRNA gene; chloroplast gene for chloroplast product (AF547096)	98	Salterns, southern France (Oren 2005)	6/ 40
124	*Halanaerobacter lacunarum* (U32593)	96	Lake Chokrak, Kerch Peninsula (Zhilina *et al.* 1991)	5/ 40
125	*Halanaerobacter lacunarum* (U32593)	96	Lake Chokrak, Kerch Peninsula (Zhilina *et al.* 1991)	4/ 40
126	*Halanaerobium* sp. strain AN-BI5B (AM157647)	97	Bannock basin (Daffonchio *et al.* 2006)	6/ 40
127	Uncultured bacterium clone PL-11B10 (AY570581)	91	Oil reservoir, western Canadian Sedimentary Basin (Grabowski *et al.* 2005)	1/ 40
128	*Pseudoalteromonas* sp. AC-144 (AJ519790)	97	East Pacific (Zeng *et al.* unpublished-a)	1/ 40
129	*Halomonas* sp. IB-559 (AJ309560)	95	? (Gilvanova *et al.* unpublished)	1/ 40
130	Uncultured *Ectothiorhodospiraceae* bacterium clone LA7-B9 (AF513947)	95	Inland waters of remote Hawaiian islands (Donachie *et al.* unpublished)	1/ 40
131	*Halomonas variabilis* strain GSP3 (AY505527)	96	Great salt plains, Oklahoma, USA (Caton *et al.* 2004)	2/ 40
132	Uncultured bacterium clone M3_c07 (DQ015372)	94	Mice gut (Ley *et al.* 2005)	1/ 40
133	Uncultured *Flavobacteriaceae* bacterium clone LA1-B21N (AF513959)	96	Inland waters of remote Hawaiian islands (Donachie *et al.* unpublished)	1/ 40
134	*Halanaerobacter lacunarum* (U32593)	96	Lake Chokrak, Kerch Peninsula (Zhilina *et al.* 1991)	1/ 40

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 16 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
168	*Dunaliella salina* 16S rRNA gene; chloroplast gene for chloroplast product (AF547096)	99	Salterns, southern France (Oren 2005)	1/ 40
135	*Halomonas venusta* strain DSM 4743 (AJ306894)	97	Seawater (Baumann *et al.* 1983)	9/ 40
136	*Halanaerobacter chitinivorans* (X89076)	90	Saltern, Chula Vista, California, USA (Liaw and Mah 1992)	5/ 40
137	Uncultured bacterium clone LCKS0-B7 (DQ129890)	99	Lake Chaka, China (Jiang *et al.* 2006)	2/ 40
138	*Chryseobacterium* sp. CHNTR56 (DQ337588)	96	Swine effluent impacted environments (Maxwell *et al.* unpublished)	1/ 40
139	*Rhodobacteraceae* bacterium CL-SP20 (AY962295)	98	Solar saltern (Cho *et al.* unpublished-a)	1/ 40
140	*Halanaerobium* sp. strain AN-BI5B (AM157647)	98	Bannock basin (Daffonchio *et al.* 2006)	2/ 40
141	Uncultured soil bacterium clone 519 (AY493959)	91	Soil aggregates (Kim and Crowley unpublished)	3/ 40
142	Uncultured proteobacterium isolate SB-43 (AJ495688)	94	Salin-de-Giraud, France (Mouné *et al.* 2003)	1/ 40
143	*Halanaerobacter lacunarum* (U32593)	98	Lake Chokrak, Kerch Peninsula (Zhilina *et al.* 1991)	1/ 40
144	Uncultured bacterium gene clone: KNA6-NB09 (AB179665)	93	Tono uranium mine, Japan (Miyoshi *et al.* 2005)	1/ 40
145	*Halomonas* sp. LCKW-Isolate10A (DQ129871)	99	Lake Chaka, China (Jiang *et al.* 2006)	1/ 40
146	Uncultured bacterium clone At12OctA1 (AY862782)	89	Salar de Atacama, Chile (Demergasso *et al.* unpublished)	1/ 40
147	*Salinibacter* sp. 5Sm6 (AY987850)	95	Maras salterns, Peruvian Andes (Maturrano *et al.* 2006)	3/ 40
148	Uncultured CFB group bacterium isolate SB-47 (AJ495692)	93	Salin-de-Giraud, France (Mouné *et al.* 2003)	1/ 40

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Table 16 Continued.

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)	CF/ tot[b]
149	Uncultured CFB group bacterium isolate SA-20 (AJ495620)	93	Salin-de-Giraud, France (Mouné *et al.* 2003)	1/ 40
150	Uncultured forest soil bacterium clone DUNssu064 (AY913287)	92	Forest soil, Germany (Roesch *et al.* unpublished)	3/ 40
151	Uncultured *Flavobacteriaceae* bacterium clone LA1-B21N (AF513959)	96	Inland waters of remote Hawaiian islands (Donachie *et al.* unpublished)	1/ 40
152	Uncultured eubacterium clone SM1H04 (AF445703)	91	Yellowstone National Park, Wyoming, USA (Bonheyo *et al.* unpublished)	2/ 40
153	Uncultured bacterium clone LCKSO-B3 (DQ129888)	98	Lake Chaka, China (Jiang *et al.* 2006)	1/ 40

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

[b] Clone frequencies (CF)/total number of clones screened (tot); based on the RFLP data.

Phylogenetics of *Bacteria* in the Salt Lakes

Different phylogenetic trees were drawn containing clone sequences and existing 16S rRNA gene sequences from all major bacterial lines of descent known to have halophilic or halotolerant members (Grant 2004), as well as the sequences of the clones' nearest neighbours. The resulting tree distributes the clone sequences from the salt lakes into nine monophyletic assemblages: *Bacteroidetes*, cyanobacteria, *Aquificae*, *Actinobacteria*, *Firmicutes*, *Alphaproteobacteria*, *Deltaproteobacteria* and *Gammaproteobacteria*. Chloroplast products (*Chlorophyceae*) are also shown. None of the sequences affiliated closely to any of the *Spirochaeta* sequences, so these were removed from the final version of the phylogenetic tree, summarised in Figure 34. Many of these groups are well supported: Chloroplasts (97%), cyanobacteria (68%), *Aquificae* (67%) and *Gammaproteobacteria* (67%). There are also additional well supported lineages (80% to 98%) that branch between these known groups, designated Salt Lake (SL) Clusters 1 to 3. Two additional SL Clusters (4 and 5) are shown near the *Aquificae* and the *Firmicutes*, both with low bootstrap values. The largest groups are the *Proteobacteria* with 30% of the clone sequences and the *Bacteroidetes* and *Firmicutes*, both with 20% of the clone sequences.

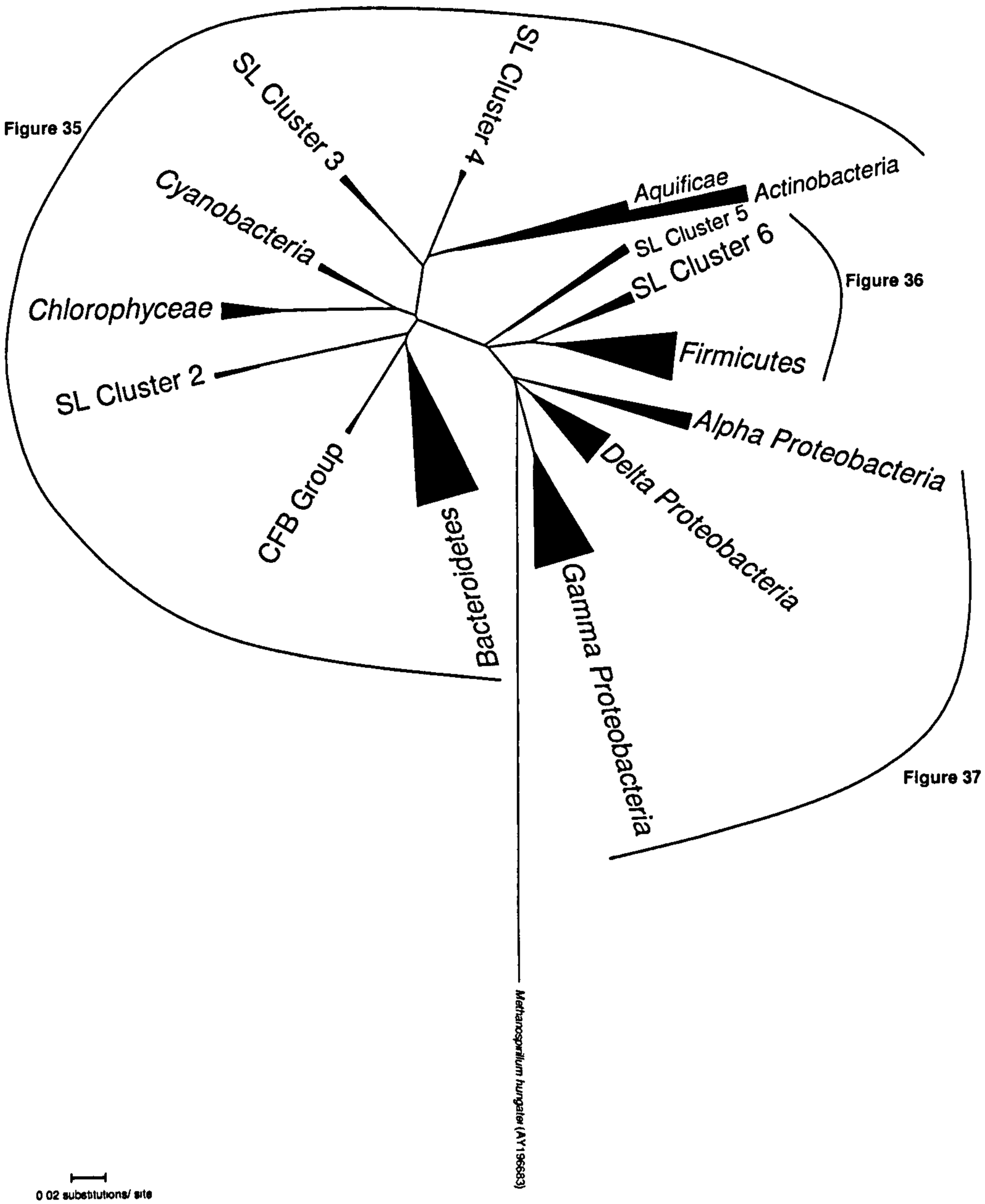


Figure 35

Figure 36

Figure 37

Figure 34 Phylogenetic Tree Summarising all *Bacteria* found in the Salt Lakes.

***Bacteroidetes* and CFB Group**

This group is shown in Figure 35, which has been magnified from the original tree and therefore does not show the root. A small group of sequences related to *Salinibacter* has formed their own well supported clade (93%) within the *Bacteroidetes*. One clone sequence from EJ3 (164) is closely related to halophilic eubacterium EHB-2 (99%). Clone sequences related to eubacterium EHB-2 were the most frequently occurring in the 16S rRNA library for EJ3, with a clone frequency of 24. It was originally detected in salterns in Alicante, Spain, which had total salinities between 22.4 - 37% (w/v), where it was recognised that there was a large number of *Salinibacter*, making up the major bacterial constituent of the microbial community (Anton *et al.* 2002). Another clone sequence from EJ3 (94) is closely related to *Salinibacter* sp. 5Sm6 (99%), which was isolated from Maras salterns in the Peruvian Andes, formed from hypersaline water emerging from a spring, said to have originated from ancient salt deposits that run underground dating to 110 million years ago and are marine in origin. This particular isolate was found in a saltern that had a total salinity of 31% (w/v) and was isolated on media that had a pH of 7.2. This isolate also grew at 37°C (Maturrano *et al.* 2006). *Salinibacter ruber* strains POLA 13 and POLA 18 were first isolated from crystalliser ponds in Santa Pola in Alicante, Spain. These isolates grew optimally at 150 - 300 g/L NaCl, at a pH 6.5 - 8.0 and at temperatures between 37°C - 47°C (Antón *et al.* 2000). These were the closest neighbours to clone sequences 70 and 166 respectively, showing 94% and 93% identity (Table 16). However, phylogenetic analysis has shown clone sequences 70 and 166 to be part of a group that branch near to, but is distantly related to *Salinibacter*.

A small group that is within the *Salinibacter* clade (showing a bootstrap value of 79%) consists of one clone sequence from EJ3 (96) and one from XH2 (147). Clone sequence 96 is closely related to an uncultured clone LCKW-B3 (99%), but sequence 147 was shown to have the highest similarity to *Salinibacter* sp. 5Sm6, with a sequence identity of 95%. However, full length alignment of clone sequence 147 has clearly placed it in this separate group. Clone LCKW-B3 is from Lake Chaka in Tibet. Phylogenetic analysis in that study similarly placed clone LCKW-B3 in a group near the *Salinibacter*. Another sequence from EJ3 (95) is closely related to an uncultured clone LCKW-B4 (99%), also from Lake Chaka. It is only distantly related to *Salinibacter* but is found with the *Bacteroidetes* (Jiang *et al.* 2006).

Other clone sequences within the *Salinibacter* clade are from CG1 (86 and 90). Sequence 86 is closely related to an uncultured clone E4aG09, which was detected in a hypersaline

endoevaporitic microbial mat. This mat was collected from a saltern from the Israel Salt Industries, which had a total salinity of 20% (w/v). The phototrophic mat had developed in the gypsum crust that covered the bottom of the pond. The mat contained distinct layers of communities; this clone was detected in the white gypsum layer. Phylogenetic analysis in that study similarly placed this clone outside, but near to the *Salinibacter* group (Sørensen *et al.* 2005). Clone sequence 90 shows 91% identity to an uncultured clone CSR-16, which was detected in an intertidal hot spring on the north western coast of Iceland. This is a changeable environment where the temperatures fluctuate between 45°C and 95°C. The system is supplied with fresh and slightly alkaline groundwater, but it mixes with the seawater at the outlets, causing severe changes in salinity, oxygen and light. Phylogenetic analysis in that study also placed this clone with the CFB group, with *Cytophaga* as its nearest neighbour (Hobel *et al.* 2005).

Apart from the *Salinibacter* clade, several other sequences are found within the *Bacteroidetes*. Clone sequences from SH1 (133) and XH2 (151) are related to *Flavobacteriaceae* clone LA1-B21N (both at 96%), detected in inland waters of remote Hawaiian islands, either from hypersaline Lake Laysan or from a brackish pond (Donachie *et al.* unpublished). Lake Laysan contains evaporated sea water and has a total salinity of 7.6% (w/v) (Donachie *et al.* 2004). One clone sequence from EJ3 (99) is closely related to *Salegentibacter* 18III/A01/068 (97%), which was isolated from the surface microlayer of the north western Mediterranean Sea (Agogué *et al.* 2005). One clone sequence from EN1 (167) is distantly related to *Flexibacter tractuosus* strain IFO 15989 (87%), which is a marine bacterium and is the type strain of the species (Nakagawa *et al.* 2002). One clone sequence from XH2 (138) is related to *Chryseobacterium* CHNTR56 (96%), which was isolated from swine effluent impacted environments (Maxwell *et al.* unpublished).

One clone sequence from XH2 (150) is distantly related to forest soil bacterium DUNssu064 (92%), detected in Germany (Roesch *et al.* unpublished). One clone sequence from BJ1 (75) is related to an uncultured bacterium clone: ODP1251B5.13 (93%), which is from deep marine sediments of the Pacific Ocean Margin (at the Peru and Cascadia Margins) that contain methane and methane hydrate. Phylogenetic analysis in that study similarly placed this clone with the *Bacteroidetes* (Inagaki *et al.* 2006). One sequence from BJ1 (72) is related to freshwater clone 965005D065.x1 which was detected in a controlled experiment where 16 mesocosms were made, each with a different input of inorganic nutrients to establish a

primary productivity gradient (Horner-Devine *et al.* 2003). One clone sequence from EN1 (101) and one from XH2 (149) are related to a bacterial isolate SA-20 (94% and 93% respectively), which was found in anoxic sediments underlying cyanobacterial mats in hypersaline saltern ponds in Salin-de-Giraud, France. Halite precipitation occurred at this pond, which had a salinity of 250 to 320 ‰. Phylogenetic analysis from that study also placed this isolate with the CFB group, with *Cytophaga fermentans* as its nearest neighbour. In addition, one clone sequence from XH2 (148) is related to an uncultured CFB group bacterium isolate SB-47 (93%), which was also detected in Salin-de-Giraud salterns. Phylogenetic analysis placed this clone in a separate branch in the middle of the CFB group (Mouné *et al.* 2003). Similarly, clone sequence 148 appears to form a separate clade, but is positioned as an outer branch near the *Bacteroidetes*.

Chlorophyceae

Illustrated in Figure 35 are several sequences are related to chloroplast 16S rRNA gene. It is possible that *Eukarya* were present in the salt lakes, and the chloroplasts in them were lysed and their DNA extracted along with the rest of the community DNA. The detection of these genes in the clone libraries has therefore given an insight into the eukaryotic diversity in the salt lakes. Several clone sequences from EN1 (106, 108, 110 and 115) and SH1 (123 and 168) are related to chloroplasts in *Dunaliella salina* (97% - 99%). Clone sequences related to chloroplasts in *Dunaliella salina* are in fact the most frequently observed in the 16S rRNA libraries for both EN1 and SH1, with a clone frequency of 19 and 6 respectively. It was first discovered in 1838 in a saltern in the south of France by Michel Felix Dunal, but was not named until 1905 (Oren 2005). It is a phototrophic alga whose carotenoids contribute to the coloration of salt lakes. One clone sequence from EJ3 (97) is related to *Chlamydomonas applanata* with 95% identity over 554 bp. However, it shows 99% identity over 367 bp to *Dunaliella salina*, thus phylogenetic analysis shows that this clone is more related to *Dunaliella salina*.

Cyanobacteria

Only one clone sequence from BJ1 (64) appears in the same clade as *Synechococcus elongatus* (see Figure 35), an oxygenic photosynthetic bacterium. Clone sequence 64 is related to clone ZB118, showing 95% identity, which was detected in a mesophilic sulphide-rich spring, called Zodletone spring. It is located north of Zodletone Mountain in the Anadarko Basin of western Oklahoma. Brine from the basin was ejected at the source along

with petroleum. The sulphide concentration in this water was between 8 to 10 mM, which maintained an anoxic environment in the water and underlying sediments. It was also slightly saline, containing 0.2 M NaCl. Phylogenetic analysis in that study placed clone ZB118 with the cyanobacteria, with *Oscillatoria terebriformis* as its nearest neighbour, a known mat-building organism (Elshahed *et al.* 2003).

Aquificae

It is surprising to find sequences related to *Aquificae*, since cultivated representatives of this group are not halophilic or halotolerant (Grant 2004), and are obligate thermophiles (Reysenbach 2001). However, they have been found in deep sea hydrothermal vent systems (Reysenbach *et al.* 2000a) so are presumably tolerant to marine conditions. Nevertheless, there appears to be one clone sequence from BJ1 (67) that is related to an uncultured *Aquificales* clone pKB (94%), which comes from a near-neutral thermal spring in Kamchatka, Russia's largest volcanic belt (Takacs-Vesbach *et al.* unpublished). This was also detected in the library for TC8 (Table 14).

Moreover, other sequences appear in the same clade as this *Aquificales* clone. A clone sequence from EN1 (105) is distantly related to an uncultured bacterium clone VC2.1 Bac16 (88%), which was from a growth chamber deployed in a Mid-Atlantic Ridge hydrothermal vent; it had a maximum temperature of 70°C that decreased to 20°C. That study placed this clone as a group on the periphery of the *Aquificales* group and in fact branched deeper within the bacterial 16S rRNA phylogeny (Reysenbach *et al.* 2000c). Another clone sequence from SH1 (116) is related to a sequence (84%) that came from a *Paralvinella sulfincola* tube and the adjacent substratum on an active deep-sea vent chimney (Page *et al.* 2004).

Actinobacteria

Actinobacteria are high%G+C Gram positives that are readily found in saline environments (Grant 2004). Two clone sequences appear in the same clade as *Microbacterium* sp. isolate D9-23 (see Figure 35), which was found in a marine aquaculture biofilter (Krieger *et al.* unpublished). One clone sequence from SH1 (127) is distantly related to clone PL- 11B10 (91%), which was detected in the waters of a low temperature biodegraded oil reservoir. That study was intended to look at the microbial diversity associated with the Pelican Lake oil field in the western Canadian Sedimentary Basin. Fluid temperature was 18°C - 20°C, containing 3 g/L NaCl and a pH of 8.4. That study placed this clone with the *Firmicutes*, with

Desulfosporosinus meridiei as its nearest neighbour (Grabowski *et al.* 2005). However, due to the low sequence identity to its nearest neighbour (only 87%), perhaps clone PL- 11B10 was wrongly positioned. Clone sequence 152 from XH2 is distantly related to eubacterium clone SM1H04 (91%), which was detected in Angel Terrace at the Mammoth Hot Springs in Yellowstone National Park, Wyoming, USA (Bonheyo *et al.* unpublished).

Firmicutes

Firmicutes are low%G+C Gram positives, also readily isolated from saline environments (Grant 2004) and was in fact on of the largest groups detected in five of the six salt lakes. This group only consists of several haloanaerobes; *Halanaerobacter*, *Halanaerobium*, *Orenia* and *Natronoanaerobium*, which can be seen in Figure 36. This has also been magnified from the original tree and therefore does not show the root.

Several clone sequences from SH1 (124, 125 and 134) and one from XH2 (143) are related to *Halanaerobacter lacunarum* (96 - 98%). This isolate was found in the silt of lagoonic hypersaline Lake Chokrak on the Kerch Peninsula, which exhibited 25% (w/v) total salinity (Zhilina *et al.* 1991). One clone sequence from XH2 (136) was related to *Halanaerobacter chitinivorans* (98%), which was found in sediment in a solar saltern in Chula Vista in southern California, which exhibited 25 - 30% (w/v) total salinity. It grew optimally at 2 - 3 M NaCl and at temperatures between 30°C and 45°C. The medium used for isolation was at pH 7.2. This isolate required chitin for growth (Liaw and Mah 1992). One sequence from CG1 (77) is distantly related to *Halanaerobiaceae* bacterium SLAS-1 (90%), which was isolated from Searles Lake; an alkaline salt-saturated brine, rich in arsenic. Searles Lake is located in the Monjave Desert about 270 km south of Mono Lake. This particular isolate required a minimum salinity of 20% (w/v). Growth was optimal at pH 9.5, with very little growth produced below pH 9.1. It was demonstrated that this isolate could use arsenate as its electron acceptor, while utilising sulphide as its electron donor. Phylogenetic analysis in that study also placed this clone with the *Firmicutes*; however, it was only distantly related to its nearest neighbour *Halothermothrix orenii*, with a sequence identity of 83.9% (Oremland *et al.* 2005).

One clone sequence from CG1 (78), six from SH1 (126) and two from XH2 (140) are closely related a *Halanaerobium* sp. AN-B15B (97 - 98%). Clone sequences related to *Halanaerobium* sp. AN-B15B were one of the most dominant in the library for SH1, with a

clone frequency of 6. It was isolated from the oxic-anoxic transition of a deep sea halocline. That study was conducted in the Bannock basin, where the total salinity varied between 4.6% - 25% (w/v) (Daffonchio *et al.* 2006).

One clone sequence from EN1 (109) is distantly related to *Orenia salinaria* (88%). This fermentative bacterium was isolated from anoxic sediment that was rich in sulphides in Salin-de-Giraud salterns, France. The salterns showed 20 - 34% (w/v) total salinity; however, this particular isolate grew optimally at 5 - 10% (w/v) NaCl, but can tolerate between 2 - 25% (w/v). It also grew optimally at temperatures between 40°C - 45°C and at a pH between 7.2 - 7.4, but tolerates between pH 5.5 - 8.5 (Mouné *et al.* 2000).

Three clone sequences from CG1 (81, 88 and 92) are related to *Natronoanaerobium halophilum* isolate G-M14CH-4 (93%, 95% and 96% respectively), which was detected in soda lakes (Jones *et al.* 1998).

The remaining clone sequences are related to uncultured isolates. One clone sequence from EJ3 (100), one from SH1 (122) and two from XH2 (137, 153) are related to two clones found in Lake Chaka, which was previously described (97- 98%). Phylogenetic analysis in that study similarly placed these two clones with the *Firmicutes*. One sequence from CG1 (93) is distantly related to uncultured clone E6aF12 (94%), which was detected in a hypersaline endoevaporitic microbial mat from a saltern in Israel (previously described). Phylogenetic analysis similarly placed this clone with the *Firmicutes*, with *Halocella cellulolytica* its nearest neighbour (Sørensen *et al.* 2005).

Another clone sequence from CG1 (84) had high localised similarity to an alphaproteobacterium (94% similarity over 369 bp) (see Table 16), but my phylogenetic analysis has placed this clone with the *Firmicutes*, distantly related to the uncultured clones of this group.

Proteobacteria

This group comprises the largest detected from all six salt lakes. Sequences from three divisions of *Proteobacteria* were found as seen in Figure 37. (4% belong to the *Alphaproteobacteria*, 6% to the *Deltaproteobacteria* and 20% to the *Gammaproteobacteria* of the total clone libraries).

Alphaproteobacteria

One clone sequence from XH2 (139) is closely related to a *Rhodobacteriaceae* clone (98%), which is from a solar saltern (Cho *et al.* unpublished-a). One clone sequence from SH1 (117) is related to *Roseovarius* sp. 2S5-2 (95%), which is a novel bacterium that was cultured from Seto Inland Sea in 1986, grown on media that was adjusted to pH 7.8 - 8.2 and was shown to produce iodine (Fuse *et al.* 2003). One clone sequence from EN2 (112) is distantly related to an uncultured alphaproteobacterium clone BIxi22 (90%), which was detected in an industrial biofilter for waste abatement in an animal-rendering plant, Germany (Friedrich *et al.* 2002). One clone sequence from CG1 (79) is related to a *Paracoccus* sp. mdw-1 (95%), which is an isolate that is known to degrade methomyl (Xu unpublished).

Deltaproteobacteria

One clone sequence from CG1 (91) is distantly related to a *Deltaproteobacteria* clone MLMS-1 (92%), which was isolated from anoxic bottom water of Mono Lake, and is known to reduce arsenate using sulphide as the electron donor. Mono Lake a hypersaline alkaline lake in California, found in a closed basin, with a salinity of 90 g/L and a pH of 9.8. That study placed this clone with the *Deltaproteobacteria*, near to but significantly distinct from known sulphate reducers (Hoeft *et al.* 2004). One sequence from EN1 (113) and a sequence from XH2 (142) are related to uncultured isolates SB-15 and SB-43 respectively (95% and 98% respectively), which were found in anoxic sediments underlying microbial mats in Salin-de Giraud salterns, France (see previously). Phylogenetic analysis in that study also placed these clones with the *Deltaproteobacteria* with *Desulfovibrio retbaense* as their nearest neighbour (Mouné *et al.* 2003). Another clone sequence from EN1 (104) is distantly related to an uncultured clone E6aH12 (88%), which was detected in a hypersaline endoevaporitic microbial mat in Israel (see previously). Phylogenetic analysis similarly placed E6aH12 with the *Deltaproteobacteria*, near *Desulfocella halophila* (Sørensen *et al.* 2005). One clone sequence from BJ1 (74) is related to eubacterium LKB122 (95%), which was detected in the leachate of Gouzikeng Landfill, a closed municipal solid waste landfill. Phylogenetic analysis in that study positioned this clone with the *Betaproteobacteria* (Huang *et al.* 2004). However, my phylogenetic analysis has positioned this clone in the middle of the *Deltaproteobacteria*, with *Desulphobulbus* as its closest relative.

Gammaproteobacteria

Clone sequences from BJ1 (65) and SH1 (118 and 131) are related to *Halomonas variabilis* strain GSP 3 (96 - 98%). In addition, one clone sequence from BJ1 (66) and EN1 (111) are related to *Halomonas salina* strain GSP33 (99% and 97% respectively). These isolates were found in the Great Salt Plains in Oklahoma, USA (Caton *et al.* 2004). This is a barren salt flat where Permian brine emerges and evaporates to leave a white salt crust. Rainfall can dissolve the salt crust to form ephemeral salt ponds. Since the plains are exposed, they are subject to severe changes in salinity (ground water salinity ranges from 4% to 37%), temperature (with summer highs of 56°C and nightly lows of 23°C), pH (average pH of 9.1, but with a recorded low of pH 7.3) and UV exposure (Miller *et al.* 2004). One clone sequence from CG1 (76) is related to *Halomonas campisalis* (97%), which was found in the sediments of the Salt Plain of Alkali Lake in Washington, USA (Mormile *et al.* 1999). This isolate grew at a total salinity range between 0.5 - 15% (w/v), but grew optimally at 5% (w/v). It is an alkaliphile, growing between pH 8 and 11. It can also grow between 4°C – 50°C (Mata *et al.* 2002). One sequence from XH2 (135) is closely related to *Halomonas venusta* strain DSM 4743 (97%). This isolate grew at a salinity range between 0 - 20% (w/v), but grew optimally at 0.5 - 7% (w/v). It grows between pH 5 and 10 and between 4°C - 45°C (Mata *et al.* 2002). Clone sequences related to *Halomonas venusta* strain DSM 4743 were the most frequently occurring in the library for XH2, with a clone frequency of 9.

One clone sequence from CG1 (163) is closely related to *Halomonas* sp. A-07 (99%), which was detected in soda lakes in Tanzania (Pradhan *et al.* unpublished). Another clone sequence from CG1 (85) is closely related to *Halomonas* sp. 3019 (97%), which was detected in deep sea sediment in the east Pacific (Zeng unpublished). One clone sequence from SH1 (129) is related to *Halomonas* sp. IB-559, showing 95% identity (Gilvanova *et al.* unpublished). One clone sequence from XH2 (145) is closely related to *Halomonas* sp. LCKW-Isolate10 (99%), which is another isolate found in Lake Chaka (previously described). Phylogenetic analysis in that study placed this clone with the *Gammaproteobacteria*, with *Halomonas ventosae* as its nearest neighbour (Jiang *et al.* 2006).

One clone sequence from EJ3 (98) is closely related to *Halovibrio denitrificans* strain HGD 3 (98%), which is an extremely halophilic denitrifying bacterium that was found in a hypersaline inland lake in north eastern Mongolia using anaerobic enrichments with acetate as an electron donor and nitrate as an electron acceptor. Growth occurred in 2 M - 5 M NaCl, but

grew optimally at 2 – 3 M and at pH 6.7 – 8.5. Anaerobic growth occurred with 2 – 4 M NaCl in the presence of nitrate and nitrite, but not with N_2O (Sorokin *et al.* 2006).

One clone sequence from SH1 (121) is closely related to *Vibrio metschnikovii* (99%). This bacterium is usually found in aquatic environments, including marine waters. They have also been isolated from clams and oysters. However, it is also a known pathogen causing diarrhoea, septicaemia and pneumonia (Wallet *et al.* 2005). Another sequence from SH1 (128) was related to *Pseudoalteromonas* sp. AC-144 (97%), which is a psychrophilic bacterium isolated from Antarctic deep sea sediment (Zeng *et al.* unpublished-a).

One clone sequence from SH1 (130) is related to an uncultured *Ectothiorhodospiraceae* bacterium clone LA7-B9 (95%), which was detected in the inland waters of remote Hawaiian islands, which was previously described (Donachie *et al.* unpublished). A clone sequence from CG1 (82) is related to a phototrophic bacterium (BN 9624) (97%), which was shown to be related to *Ectothiorhodospira* species (Imhoff and Suling 1996).

One sequence clone from EN1 (103) is related to *Alkalilimnicola halodurans* isolate (94%), which is an isolate that was found in sediments from Lake Natron in the African Rift Valley. The site where this isolate was found had a total salinity of 2% (w/v) and a pH of 10 - 11. This particular isolate grows optimally at pH above 8.5, with a total salinity range between 3% - 8% (w/v) NaCl and at 35°C. Though only moderately halophilic, this isolate is also halotolerant, growing at salinities up to 28% (w/v) NaCl (Yakimov *et al.* 2001).

One clone sequence from CG1 (80) is related to *Gammaproteobacterium* M12-20B (95%), which was detected in Mono Lake, which was previously described (Bano and Hollibaugh unpublished). Another clone sequence from CG1 (89) is closely related to an Arctic seawater bacterium (99%) (Zeng *et al.* unpublished-b).

Novel Lineages

These novel lineages have been called Salt Lake (SL) Clusters 1 to 5 (Figure 35 and Figure 36).

SL Cluster 1 branches near to the *Bacteriodetes* (Figure 35). It consists of just one clone sequence from SH1 (120), which is distantly related to an uncultured bacterium clone: UTFS-OF09-d22-24 (92%), which was detected in activated sludge (Oe *et al.* unpublished).

SL Clusters 2 and 3 branch between the cyanobacteria and the *Aquificae* (Figure 35). SL Cluster 2 consists of one sequence from BJ1 (73), which is distantly related to an uncultured bacterium isolate JH10_C67 (90%), which was detected in the intertidal flat of Ganghwa Island (Cho *et al.* unpublished-b). This island is located in the estuary of the Han River on the west coast of South Korea. These marine sediments are periodically exposed and then flooded by seawater, causing large changes in temperature and salinity (Kim *et al.* 2005). Another sequence from XH2 (141) is distantly related to uncultured soil bacterium clone 519 (91%), which was detected in soil aggregates (Kim and Crowley unpublished).

SL Cluster 3 consists of one sequence from EN1 (114) and one from SH1 (119), which are both related to an uncultured bacterium 'KTK 32' (91% and 97% respectively). This clone is from highly saline brine sediments of Kebrit Deep in the northern Red Sea. Kebrit Deep is a round basin approximately 1 km in diameter and 1549 m deep. It is filled with anaerobic brine with a pH of 5.5, a total salinity of 26% (w/v) and a temperature of 23.3°C. The brine had a high gas content, containing large amounts of CO_2 and H_2S. Phylogenetic analysis in that study similarly placed this clone in a separate clade that did not belong to any of the known genus in the domain *Bacteria*, but was positioned between the orders *Thermotogales* and *Aquificae* (Eder *et al.* 1999).

SL Cluster 4 branches from the *Aquificae*, consisting of several clone sequences (Figure 35). One sequence from CG1 (83) is distantly related to an uncultured candidate division OP11 bacterium clone OPB92 (88%). This clone was detected in Obsidian Pool in Yellowstone National Park, Wyoming, USA. This pool had a temperature of 75°C to 95°C and was rich in reduced iron, sulphide and CO_2. Phylogenetic analysis in that study also placed this clone in a novel lineage that did not bear any relation to known bacterial lineages, only showing ≤ 80% similarity to existing 16S rRNA gene sequences. The group, designated OP11, was positioned near the *Deinococcus-Thermus* branch. That study also demonstrated the presence of this division in a range of environments including Carolina Bay sediment, Amazonian soil and Australian deep subsurface water, thereby demonstrating the ubiquitous nature of this novel lineage in the environment (Hugenholtz *et al.* 1998). One sequence from XH2 (144) is distantly related to an uncultured bacterium gene clone: KNA6-NB09 (93%), which was detected in groundwater from aquifers in sedimentary and granite rock in Tono uranium mine, Japan. The water was anaerobic with a pH of 7.8 to 9.6 and a temperature of 19°C to 20.1°C.

This particular clone was detected in groundwater that was filtered through a 0.2 µm filter, but captured on a 0.1 µm filter. Phylogenetic analysis in that study similarly placed this clone in a separate clade unrelated to known bacteria; it was placed in a separate group that was weakly related to the candidate division OP11 as previously described, designated WS6 (Miyoshi *et al.* 2005). One clone sequence from EN1 (107) is distantly related to an uncultured proteobacterium clone BSB3-217m (88%), though my phylogenetic analysis has clearly distinguished this clone from the *Alpha*, *Delta* or *Gammaproteobacteria*. Clone BSB3-217m was detected in the oxic/anoxic chemocline in the Black Sea at a depth of 217 m, where there is little oxygen but a lot of sulphide (Vetriani *et al.* 2003). One clone sequence from EN1 (165) and one from XH2 (146) are related to an uncultured bacterium clone At12OctA1 (95% and 89% respectively), which was detected in Salar de Atacama in northern Chile (Demergasso *et al.* unpublished). This saltern was found in the Atacama Desert, 2300 m above sea level. It was located on salt deposits of the Pre-Cordilleran Depression. It was also in a basin and so received leachate from the surrounding volcanic rock and underground water. The conditions varied considerably; salinity was between 0.9% and 36% (w/v), pH between 6.26 and 8.94 and temperature between 10.6°C and 23.9°C (Demergasso *et al.* 2004).

SL Cluster 5 branches near the *Firmicutes* (Figure 36). One clone sequence from EN1 (102) is closely related to a Glacial ice bacterium G500K-19, showing 98% sequence identity (Christner 2002). One clone sequence from SH1 (132) is related to an uncultured bacterium clone M3_c07 (94%), which was detected in the gut of obese mice (Ley *et al.* 2005). One clone sequence from CG1 (87) is closely related to an uncultured low G+C Gram-positive bacterium clone mixed culture C-1 (98%), which was also detected in Mono Lake (previously described). Phylogenetic analysis in that study positioned this clone as distantly related to *Bacillus selenitireducens*, which belongs to the *Gammaproteobacteria* (Hollibaugh *et al.* 2006).

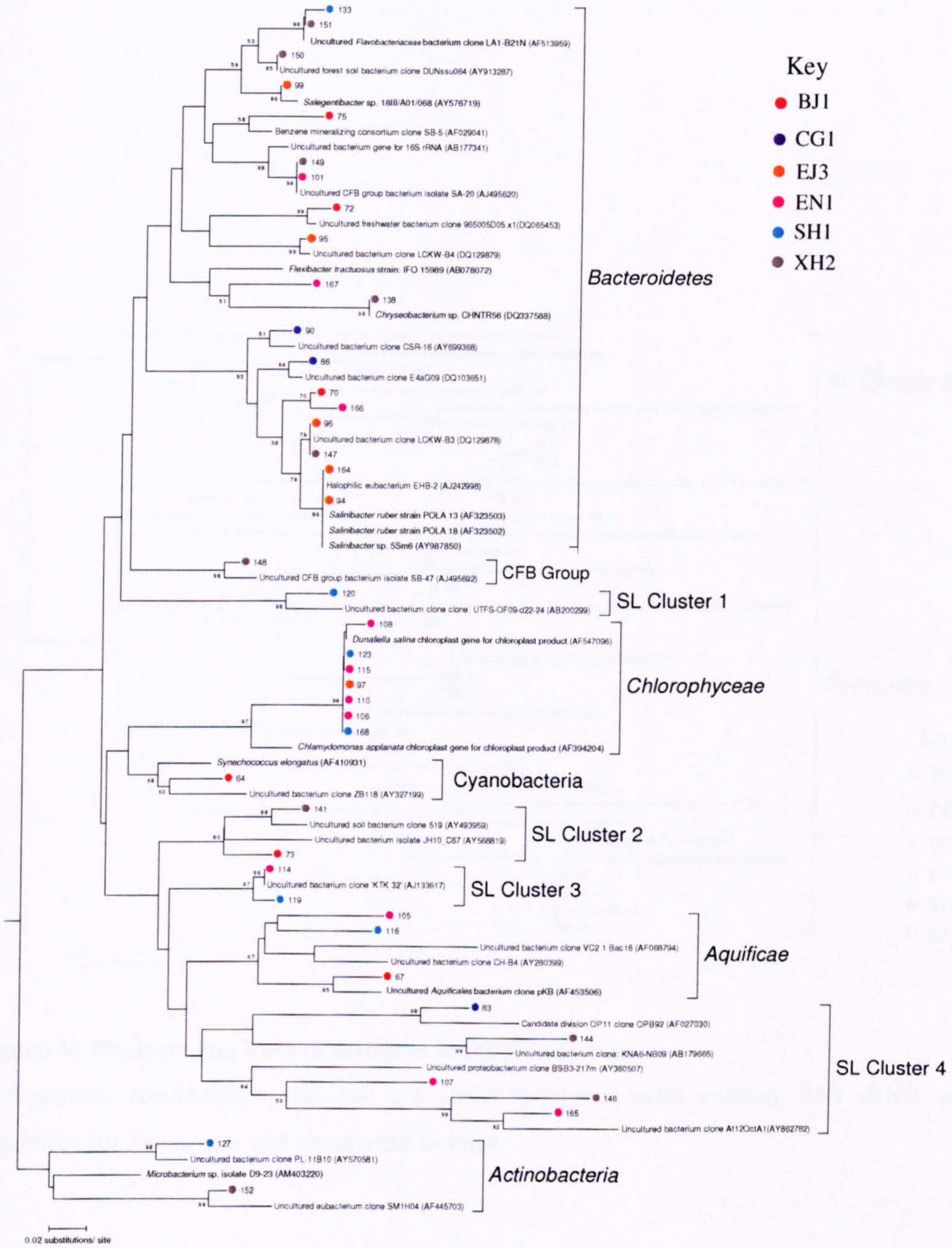


Figure 35 Phylogenetic Tree of *Bacteria* in the Salt Lakes.

Phylogenetic relationships of the clone sequences with existing 16S rRNA gene sequences for *Bacteroidetes*, CFB group, chloroplasts, cyanobacteria, *Aquificae*, *Actinobacteria* and uncultured isolates.

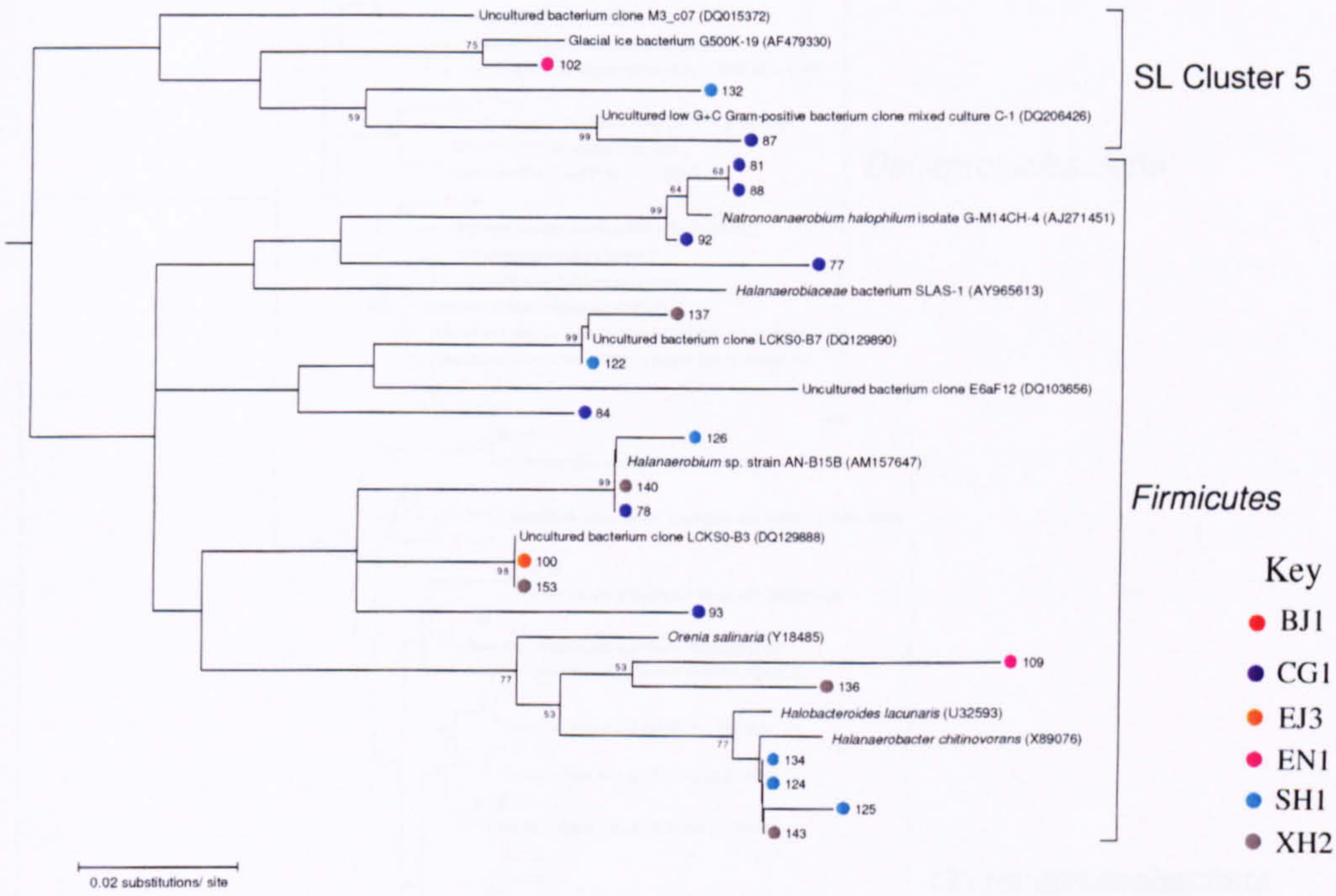


Figure 36 Phylogenetic Tree of *Bacteria* in the Salt Lakes.

Phylogenetic relationships showing the clone sequences with existing 16S rRNA gene sequences for *Firmicutes* and uncultured isolates.

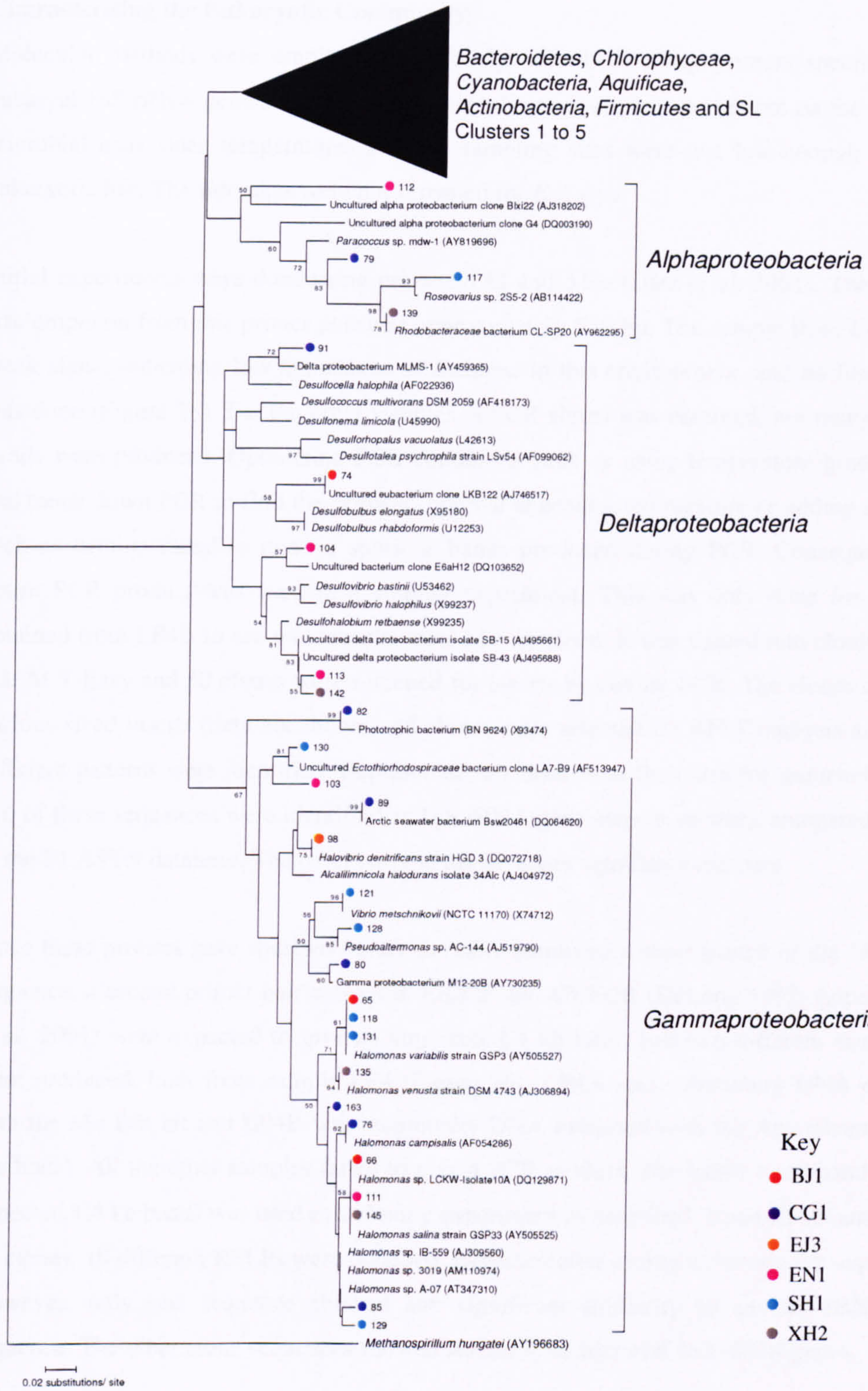


Figure 37 Phylogenetic Tree of *Bacteria* in the Salt Lakes.

Phylogenetic relationships showing the clone sequences with existing 16S rRNA gene sequences for *Proteobacteria* and uncultured isolates.

Characterising the Eukaryotic Community

Molecular methods were employed (previously described) using primers specific for the eukaryal 18S rRNA gene sequence. This molecular analysis was carried out on the hot spring microbial mats since temperatures at some sampling sites were just low enough to sustain eukaryotic life. The salt lakes were not screened for *Eukarya*.

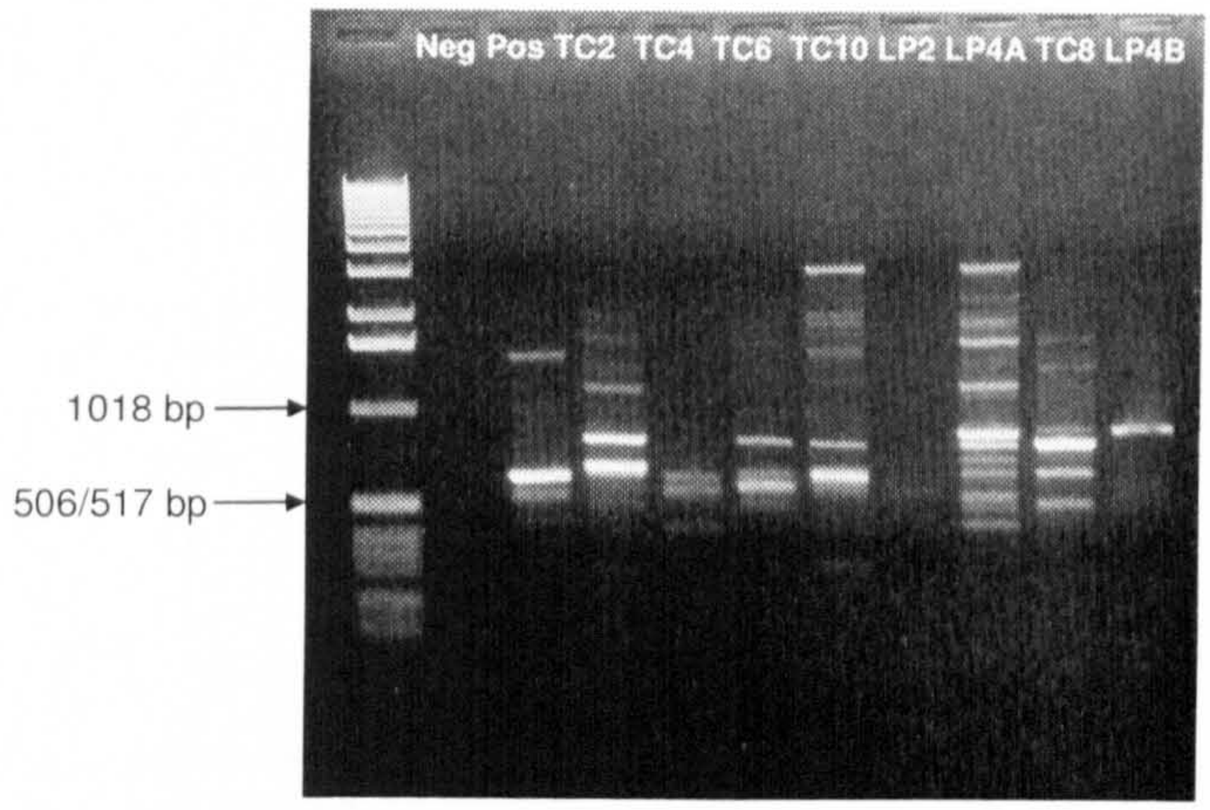
Initial experiments were done using primers 1Af and 516r (Díez *et al.* 2001). The expected size amplicon from this primer pair was approximately 550 bp. The sample from LP2 gave a weak signal indicating low abundance of *Eukarya* in this environment, and no further work was done (Figure 38). For the other samples, a PCR signal was obtained, but many spurious bands were produced. Optimising PCR conditions such as using temperature gradient PCR and touch down PCR to find the primers' optimal annealing temperature or adding enhancers such as betaine failed to remove spurious bands produced during PCR. Consequently, the entire PCR product was used in a cloning experiment. This was only done for the PCR obtained from LP4B to see what results would be obtained. It was ligated into cloning vector pGEM T-Easy and 50 clones were screened for inserts by colony PCR. The clones contained various sized inserts (data not shown). 45 clones were selected for RFLP analysis and only 4 different patterns were identified. Representative clones were then sent for sequencing. Only two of these sequences were identified as 18S rRNA gene sequences when compared to those in the BLASTN database. The other two did not give any significant matches.

Since these primers gave spurious bands and only amplified a short stretch of the 18S rRNA sequence, a second primer pair was used. EK82F and U1492R (DeLong 1992; López-García *et al.* 2001) were expected to give an amplicon 1.4 kb long. Just two different sized bands were produced, both from sample LP4 (Figure 38). LP4A was community DNA extracted with the Mo Bio kit and LP4B was community DNA extracted with the Amersham kit (see Methods). All the other samples failed to give a PCR product. The larger sized band (i.e. the expected 1.4 kb band) was used in a cloning experiment as described. Upon RFLP analysis on 30 clones, 10 different RFLPs were obtained. Representative clones were sent for sequencing. However, only one sequence showed any significant similarity to an 18S rRNA gene sequence. The other clone sequences showed similarity to bacterial 16S rRNA genes.

In addition to BLAST analysis to find the highest localised homology (homology over small strectches of the gene), the FASTA database was used to find homologies to the full length of the gene. All sequences showed low percentage identity to their nearest neighbour.

An unrooted phylogenetic tree was constructed using the clone sequences and known 18S rRNA gene sequences (Figure 39). LP4_euk1 is distantly related to an uncultured eukaryotic picoplankton clone PG5.22 (91%) according to BLAST. LP4_euk2 is distantly related to an uncultured eukaryote clone: d7-Band_L-Cont-Upper (88%) also according to BLAST. However, upon phylogenetic analysis, this latter clone (d7-Band_L-Cont-Upper) did not affiliate with any of the clades in the tree and was therefore removed from the final version. In fact, LP4_euk1 and LP4_euk2 appear to be closely related to each other and formed a well supported branch (with a bootstrap value of 99%) near to, but significantly different from the Ciliates. LP4_euk3 was found to be distantly related to *Ceratium longipes* strain ccmp1770 (89%) according to BLAST analysis. However, FASTA analysis showed it to be more related to *Naegleria clarkii*. Clearly phylogenetic analysis suggests that it is in fact somewhat related to the *Schizopyrenidae*, though it clearly forms its own separate branch.

Primers 1Af and 516r



Primers EK82F and 1492r

Figure 38 Results of 18S rRNA gene PCR.

These 1.2% TAE gels show the PCR results using eukaryotic primers 1Af/516r (top panel) and EK82F/1492r (bottom panel) on hot spring environmental samples. The positive PCR was done using acanthamoeba genomic DNA. Community DNA from LP4A was extracted using the Mo Bio kit, while community DNA from LP4B was extracted using the Amersham kit.

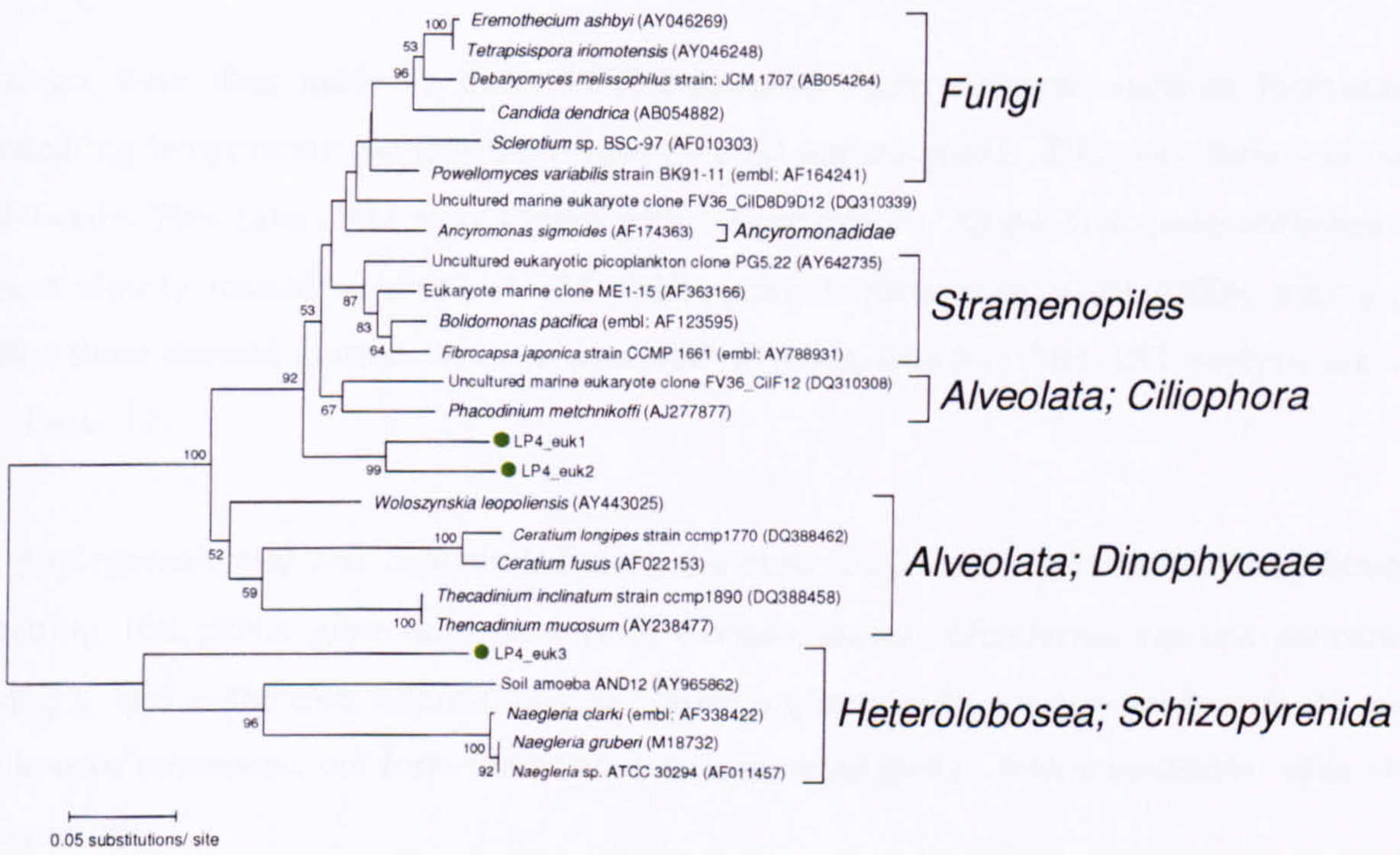


Figure 39 Phylogenetic Tree of *Eukarya* in LP4 Microbial Mat.

Phylogenetic tree showing the relationships between the clone sequences with existing 18S rRNA gene sequences for *Eukarya* from the hot spring LP4 microbial mat.

Nanoarchaea Specific Primers

Primers designed by Huber were used to amplify nanoarchaea that may be present in the environmental samples (Huber *et al.* 2002). PCR signals were given by hot spring samples LP4, TC6 and TC8 and salt lake samples EN1, SH1 and XH2. These were subsequently cloned to construct a nanoarchaeal 16S rRNA gene library. Twelve clones from each environment were screened for inserts. Various sized inserts were obtained (data not shown) and only 15 clones in total were used in subsequent RFLP analysis. Only six different patterns were obtained and representative clones were sent for sequencing. These failed to give specific nanoarchaeal sequences. Instead, the primers had amplified archaeal 16S rRNA gene sequences.

Efforts were then made to make PCR conditions more stringent such as increasing the annealing temperature, so that the primers would not misprime. This was done with sample LP4 only. This gave eight more clones with unique RFLPs. Again, five clone sequences were more closely related to archaeal 16S rRNA gene sequences upon BLASTN analysis. The other three showed matches to *Listeria monocytogenes*. Results of BLAST analysis are shown in Table 17.

A phylogenetic tree was constructed using the clone sequences, their nearest neighbours and existing 16S rRNA gene sequences from *Crenarchaeota*, *Euryarchaeota* and nanoarchaea, using *E. coli* as the root. Interestingly, the clone sequences obtained do not bear much relation to known sequences, but form their own well supported group, with a bootstrap value of 86% (Figure 40).

All clone sequences from EN1 and one clone sequence from TC6 are distantly related to an unidentified archaeal clone Rot13 (84% to 92%), which was detected in anoxic sediment in Lake Rotsee in Switzerland. This lake had a low input of organic material and a high methane production. Phylogenetic analysis in that study placed this clone in a novel lineage, distinct from known *Euryarchaeota* (Zepp Falz *et al.* 1999). One clone sequence from XH2 is distantly related to an unidentified archaeal clone ABS3 (83%), which was detected in excised, washed rice roots producing methane that were grown in flooded microcosms. Phylogenetic analysis in that study also placed this clone in a separate lineage, significantly different from other known *Euryarchaeota* (Groβkopf *et al.* 1998). One clone sequence from LP4 is distantly related to an uncultured archaeon clone HSWK30 (87%), which was detected in

microbial communities in cold sulphidic springs of Eschenlohe near Garmisch-Partenkirchen, Germany. The temperature at the sampling site was 8°C and the pH was 7.4. Sulphide concentration was 0.5 mg/L. Phylogenetic analysis in that study placed this clone in a separate assemblage, distinct from known sequences (Rudolph *et al.* 2004). Another clone sequence from LP4 is distantly related to an uncultured euryarchaeote clone LDS50 (86%), which was detected in the sediment of eutrophic Lake Dagow in north eastern Germany. Moreover, a further LP4 clone sequence is distantly related to an uncultured euryarchaeote clone PSS5 (85%), which was detected in the same environment. This lake was found to contain many methanogens, so the rate of methane production was high (Glissman *et al.* 2004). Another clone sequence from LP4 is distantly related to an unidentified archaeon clone pISA1 (82%), which was detected in a deep-sea hydrothermal vent in the Iheya Basin in the Mid-Okinawa trough, south of Japan. This clone was found in water simmering sediments that was composed of carbonate-rich hydrothermal precipitates with disseminated sulphides. The temperature was 100°C at a depth of 1398 m. Phylogenetic analysis in that study placed this clone in a separate clade that was separate from known *Euryarchaeota* (Takai and Horikoshi 1999). Another two LP4 clone sequences are distantly related to an uncultured archaeon clone St_T_94, which was detected in the sediment of Lake Stechlin, a temperate oligotrophic lake in north eastern Germany (Chan *et al.* unpublished).

Based on this experiment, it seems that nanoarchaea are not present in these environments. Despite this, these primers have found a novel portion of the microbial community that is not related to any known organisms.

Table 17 Sequencing Results for *Archaea* in Hot Springs and Salt Lakes.

Nearest neighbours from the BLASTN database to clone sequences for the 16S rRNA gene libraries are shown. The clone numbers have been colour coded according to the environment they are from (**Magenta** = EN1, **Grey** = XH2, **Green** = LP4, **Black** = TC6).

Clone No.	Nearest Neighbour (Accession Numbers)	% ID[a]	Origin (reference)
EN1_n1	Unidentified archaeal clone Rot13 (Y18089)	87	Lake Rotsee, Switzerland (Zepp Falz *et al.* 1999)
EN1_n2	Unidentified archaeal clone Rot13 (Y18089)	87	Lake Rotsee, Switzerland (Zepp Falz *et al.* 1999)
EN1_n3	Unidentified archaeal clone Rot13 (Y18089)	84	Lake Rotsee, Switzerland (Zepp Falz *et al.* 1999)
XH2_n1	Unidentified archaeal clone ABS3 (AJ227951)	83	Rice roots (Großkopf *et al.* 1998)
LP4_n1	Uncultured archaeon clone HSWK30 (AJ631246)	87	Sulphidic springs, Eschenlohe, Garmisch-Partenkirchen, Germany (Rudolph *et al.* 2004)
LP4_n2	Uncultured euryarchaeote clone LDS50 (AY133943)	86	Sediment, Lake Dagow, north eastern Germany (Glissman *et al.* 2004)
LP4_n3	Unidentified archaeon clone pISA1 (AB019751)	82	Deep-sea hydrothermal vent, Iheya Basin, Mid-Okinawa trough (Takai and Horikoshi 1999)
LP4_n4	Uncultured euryarchaeote clone PSS5 (AY133888)	85	Sediment, Lake Dagow, north eastern Germany (Glissman *et al.* 2004)
LP4_n5	Uncultured archaeon clone St_T_94 (AY531729)	93	Sediment, Lake Stechlin, north eastern Germany (Chan *et al.* unpublished)
LP4_n6	Uncultured archaeon clone St_T_94 (AY531729)	94	Sediment, Lake Stechlin, north eastern Germany (Chan *et al.* unpublished)
TC6_n1	Unidentified archaeal clone Rot13 (Y18089)	92	Lake Rotsee, Switzerland (Zepp Falz *et al.* 1999)

[a] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

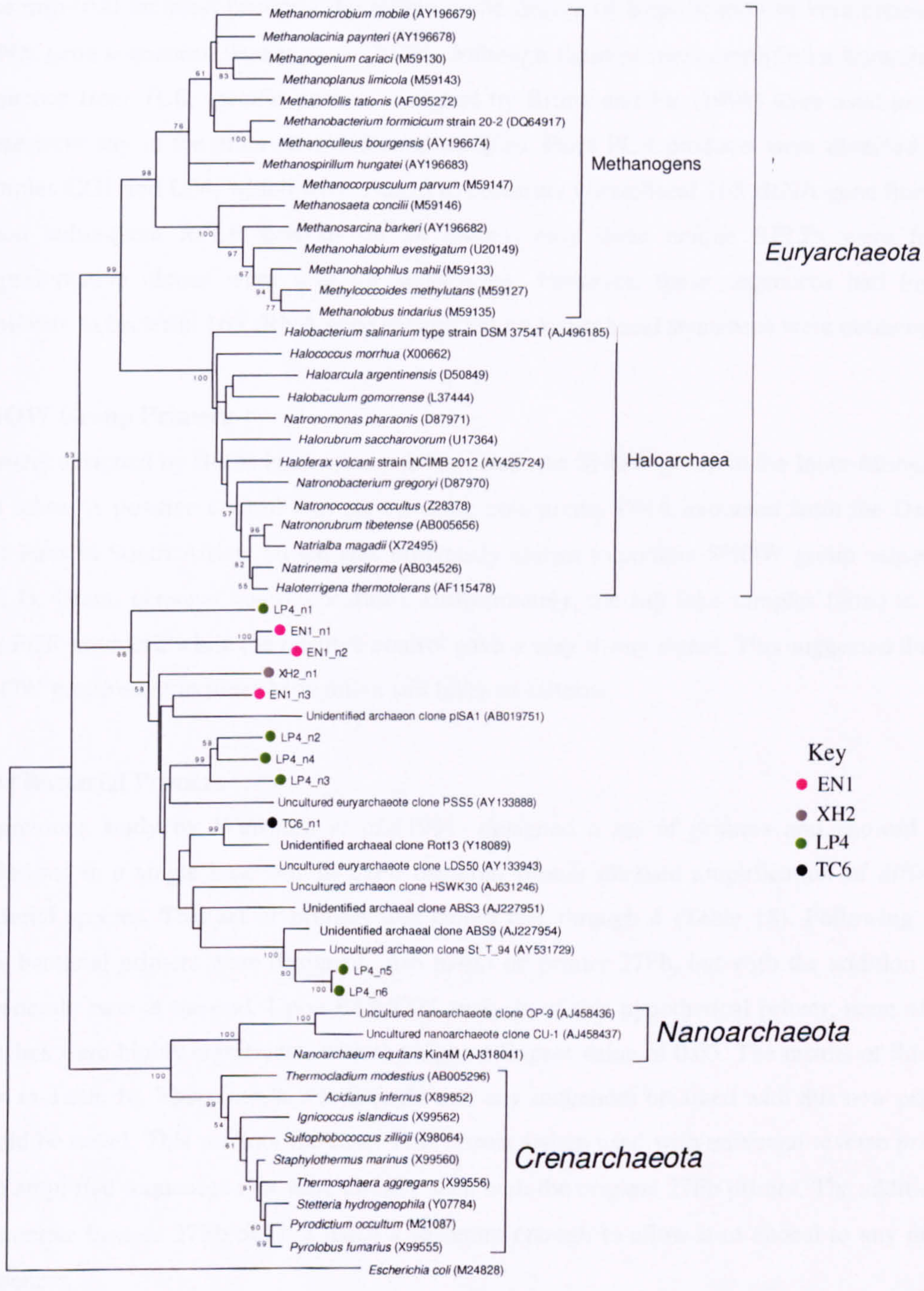


Figure 40 Phylogenetic Tree of *Archaea* in Hot Springs and Salt Lakes.

Phylogenetic tree showing the relationships between clone sequences and existing 16S rRNA gene sequences for nanoarchaea, *Crenarchaeota* and *Euryarchaeota* from the hot springs and salt lakes.

Korarchaeota Specific Primers

The universal archaeal primer 27Fa allows some degree of amplification of korarchaeal 16S rRNA gene sequences (Baker *et al.* 2003). Although these primers amplified a korarchaeote sequence from TC8, specific primers designed by Brunk and Eis (1998) were used to see if there were any in the other environmental samples. Faint PCR products were obtained from samples CG1 and LP4, which were cloned to construct korarchaeal 16S rRNA gene libraries. Upon subsequent RFLP analysis on 24 clones, only three unique RFLPs were found. Representative clones were sent for sequencing. However, these sequences had highest similarity to bacterial 16S rRNA gene sequences; no korarchaeal sequences were obtained.

SHOW Group Primers

Primers designed by Burns *et al*, were used to detect the SHOW group in the Inner Mongolian salt lakes. A positive control was set up using community DNA extracted from the Darling Salt Pans in South Africa, which was previously shown to contain SHOW group sequences (W. D. Grant, personal communication). Unfortunately, the salt lake samples failed to give any PCR products, while the positive control gave a very strong signal. This suggested that no SHOW group exist in Inner Mongolian salt lakes or salterns.

New Bacterial Primers

A previous study by Weisburg *et al* (1991) designed a set of primers and showed that variations in a single base within 27Fb bacterial primer allowed amplification of different bacterial species. This set of primers was called fD1 through 4 (Table 18). Following this, new bacterial primers were designed, also based on primer 27Fb, but with the addition of a degenerate base at the end. Upon BLASTN analysis of this hypothetical primer, none of the matches were highly significant, with the highest Expect value of 0.03. The results of this are seen in Table 18. Therefore, it was thought that any sequences obtained with this new primer would be novel. This was not the case. This primer (when used with universal reverse primer rp1) amplified sequences that were already seen with the original 27Fb primer. The additional degenerate base to 27Fb did not make it different enough to allow it to anneal to any novel sequences.

Table 18 Bacterial Primers and their Targets.

This table shows variations of bacterial forward primer, 27Fb and their target groups for amplification during PCR. The bases that are different to the original primer are highlighted in **red**.

Primer	Sequence (5' to 3')	Target[a]	Reference
fD1 (27Fb)	AGA GTT TGA TCC TGG CTC AG	**Gram-positive bacteria and relatives**: *Bacillus*, *Clostridium*, *Staphylococcus*, *Listeria*, *Lactobacillus*, *Streptococcus*, *Mycoplasma*, *Spiroplasma*, *Ureaplasma*, *Achloeplasma*, *Erysipelothrix*, *Fuseobacterium*, *Arthrobacter*, *Mycobacterium*, *Streptomyces*. **Purple bacteria and relatives (*Proteobacteria*)**: *Rochalimaea*, *Brucella*, *Rhodopseudomonas*, *Agrobacterium*, *Rhodospirillum*, *Pseudomonas*, *Neisseria*, *Caulobacter*, *Myxococcus*, *Campylobacter*, *Rickettsia*, *Ehrlichia*. **Cyanobacteria**: *Anacystis* (*Synechococcus*). ***Bacteroidetes*, *Flavobacterium***. ***Deinococcus* and relatives**: *Deinococcus*, *Thermus*. ***Spirochetes***: *Treponema*, *Spirochaeta*. ***Planctomyces* and relatives**. ***Chlorobium*- green sulphur bacteria**. ***Thermotogae***.	(Weisburg *et al.* 1991)
fD2	AGA GTT TGA TC**A** TGG CTC AG	**Enteric members of *Gammaproteobacteria***: *Escherichia*, *Shigella*, *Salmonella*, *Serratia*, *Erwinia*, *Citrobacter*, etc. *Oceanospirillum*, *Haemophilus*, *Actinobacillus*, *Vibrio*, *Pasturella*.	(Weisburg *et al.* 1991)
fD3	AGA GTT TGA TCC TGG CT**T** AG	**Spirochetes of the genus *Borrelia***	(Weisburg *et al.* 1991)

[a]The targets for fD5 to 7 were determined by BLASTN analysis. The numbers of matches are those with the highest Expect value.

Table 18 Continued.

Primer	Sequence (5' to 3')	Target[a]	Reference
fD4	AGA ATT TGA TCT TGG TTC AG	**Genus *Chlamydia***	(Weisburg *et al.* 1991)
fD5	AGA GTT TGA TCC TGG CTT AA	**(95 matches): *Deltaproteobacteria*, *Alphaproteobacteria*** (*Wolbachia*), ***Betaproteobacteria*** (*Burkholderia*), ***Gammaproteobacteria*** (*Salmonella*). ***Firmicutes*** (*Clostridium, Paenibacillus, Bacillus, Pseudomonas*). ***Actinobacteria*** (*Streptomyces*). ***Chloroflexi*** (*Thermomicrobia*). ***Bacteriodetes, Chlorobi.***	This study
fD6	AGA GTT TGA TCC TGG CTT AC	**(77 matches): *Betaproteobacteria*** (*Mitsuaria, Herbaspirillum, Naxibacter, Halobacillus*), ***Deltaproteobacteria*** (*Desulfovibrio*), ***Gammaproteobacteria*** (*Methylococcus*). ***Firmicutes*** (*Geobacillus, Mycoplasma*). ***Actinobacteria*** (*Holophaga*). **Cyanobacteria** (*Nostocales*).	This study
fD7	AGA GTT TGA TCC TGG CTT AT	**(34 matches): *Gammaproteobacteria*** (*Pseudomonas, Methylophaga*), ***Alphaproteobacteria*** (*Sphingomonas*). ***Actinobacteria*, Cyanobacteria.**	This study

[a]The targets for fD5 to 7 were determined by BLASTN analysis. The numbers of matches are those with the highest Expect value.

RESULTS: CHARACTERISING MICROBIAL POPULATIONS BY CULTIVATION METHODS

Introduction

In this chapter, the microbial community was investigated by cultivating microorganisms from salt lakes. Direct plating on nutrient rich media containing similar salts to those found in the salt lakes were used to allow maximum recovery of organisms. All isolates were sorted into groups, firstly by colony morphology, and then by analysing their whole cell protein patterns. They were later identified by their 16S rRNA gene sequence. Upon comparison between the isolates with the sequences obtained from the 16S rRNA gene libraries, it was clear that cultivation methods picked up organisms that molecular techniques did not. Assuming that PCR is random and representative, then cultivation methods detect an atypical portion of the environment. Therefore, culturing microorganisms from the environment remains a valid and crucial methodology for assessing microbial populations.

Salt Composition of Salt Lake Brine and Modification of CHM

The elements contained in the lakes were determined by Inductively Coupled Plasma Optical Emission Spectrometry (ICP-ES), calculated in parts per million (ppm), while the carbonate and bicarbonate was determined by titration with H_2SO_4, given in mg/L (see Methods). Since 1 ppm is equivalent to 1 mg/L, the molarities of the elements were determined and are shown in Table 19. A list of possible salts to incorporate all the elements found in the analysis was compiled and relevant recipes were made up according to the compounds' molecular mass. However, not all of the elements were accounted for by this method (Cl and Na; see Table 20), and so there are probably other trace compounds that are present in the lakes in the form of chlorides and other sodium salts.

Unfortunately, classic halophile medium (CHM) could not be made containing *all* of the salts calculated in this analysis since it is made in two components. The total volume of distilled water is halved and one half used to dissolve nutrients such as yeast extract, casamino acids and agar; and the other half for the salts. Hence, not all the salts could be dissolved in the latter component, especially where NaCl is approaching saturation. In addition, the high salt concentrations caused the agar to precipitate out of solution once the two components were mixed, despite being mixed at 60°C or higher. Splitting the volume of water unevenly so that more water was used for dissolving the salts meant that the nutrient component was too

viscous and would not mix with the salt solution. Therefore, the levels of some of the salts were reduced by 30% to 55% until all salts were dissolved, hence producing the recipes found in the Methods section.

Table 19 Chemical Composition of the Salt Lake Waters.

Calculated concentrations of elements found in the salt lake water samples as determined by Inductively Coupled Plasma Optical Emission Spectrometry (ICP-ES). Bicarbonate (HCO_3) and carbonate (CO_3) concentrations were found by titration with H_2SO_4.

Sample/ Element	B (mM)	Br (mM)	Ca (mM)	Cl (M)	K (mM)	Li (mM)	Mg (M)	Na (M)	S (M)	HCO_3 (mM)	CO_3 (mM)
BJ1	4.25	8.05	0.77	4.61	33.2	0.33	0.35	5.32	1.07	7.4	3.3
CG1	3.81	1.25	0.12	1.08	14	0.014	0.001	2.89	0.43	360	410
EJ3	4.43	14	3.04	4.36	68.9	1.34	2.08	2.82	0.94	9.84	23.3
EN1	5.24	9.57	0.12	5.33	39	3.3	0.86	4.2	0.48	4.1	8.3
SH1	13	5.18	2.62	4.69	150	1.93	0.26	5.38	0.81	7.4	13
XH2	2.87	7.78	3.1	5.4	53.1	0.23	0.085	5.06	0.33	13.9	1.7

Table 20 Salt Composition of the Salt Lake Waters.

Calculated concentrations of possible salts found in the lake water. The concentration of excess ions that were not incorporated into the recipes is also shown.

Salt/Sample	BJ1	CG1	EJ3	EN1	SH1	XH2
$Na_2B_4O_7.10H_2O$ (mM)	4.25	3.81	4.43	5.24	13	2.87
NaBr (mM)	8.05	1.25	14	9.57	5.18	7.78
$CaCl_2$ (mM)	0.77	0.12	3.04	0.12	2.62	3.1
KCl (mM)	33.2	14	68.9	39	150	53.1
LiCl (mM)	0.33	0.01	1.34	3.3	1.93	0.23
$MgCl_2.6H_2O$ (mM)	350	1.02	1140	380	-	-
NaCl (M)	3.15	0.84	2	4.16	4.22	4.55
Na_2SO_4 (M)	1.07	0.43	-	-	0.55	0.24
$NaHCO_3$ (mM)	7.4	360	9.84	4.1	7.4	13.9
Na_2CO_3 (mM)	3.3	410	23.3	8.3	13	1.7
$MgSO_4.7H_2O$ (mM)	-	-	940	480	260	84.9
Unaccounted Elements						
Na (M)	-	-	0.741	-	-	-
Cl (M)	0.726	0.224	-	0.369	0.454	0.790

Cultivation of Microorganisms on Solid Media

Salt crust, brine, sediment and biomass captured on filters were used to inoculate plates of mCHM (see Methods). Dilutions of these innocula were made and also plated on mCHM since previous studies have shown that this increases the viability of organisms by reducing overcrowding (Burns *et al.* 2004). These plates were then placed in different conditions; some were placed at 30°C, some at 37°C and another set at 37°C in anaerobic jars. Two different incubation temperatures were used; 30°C was chosen to mimic the conditions at the time of sampling with the aim of isolating different organisms. After inoculation, the plates were checked every day for a week, and then every week for one to two months. Numerous colonies with different morphologies were picked and made into pure cultures. The colony morphologies and the time it took for that colony to appear were recorded.

Whole Cell Protein Analysis and Identification of Isolates

Isolates were grouped according to colony morphology. These were grown in mCHM broth and the whole cell proteins of all isolates were analysed by SDS-PAGE gels. Isolates with similar colony morphologies were run on the same gel. A representative isolate was selected from those that had identical whole cell protein patterns and used for further analysis. The resulting bands from this analysis can be seen in Figure 41. The 16S rRNA genes of each isolate were amplified by colony PCR. These were then sent for sequencing to allow the identification of the isolates. The first 700 – 800 bp of the bacterial isolates' sequences and approximately 1.4 kb of the archaeal isolates' sequences were compared to those in the BLASTN database; the results are seen in Table 21.

Isolate Diversity with Incubation Time and Temperature

Some samples failed to give any microbial growth. These were salt crust from EJ3 grown in aerobic conditions at 30°C and 37°C, and salt crust and biomass from filter samples from BJ1 and EJ3 grown in anaerobic conditions.

Colonies began to appear after a week of incubation at both 30°C and 37°C on plates inoculated with samples from BJ1, EJ3, EN1 and CG1. Most of these isolates were identified as *Halomonas* and *Bacillus* species. There was also an isolate related to *Chromohalobacter* sp. IA1-Ch2 identified from BJ1, an isolate from CG1 related to an isolate from Crater Lake, Kenya and and isolate related to *Pontibacillus marinus* strain BH030004 from EJ3.

In the second week, growth had appeared on plates from the remaining samples, SH1 and XH2. The majority of the isolates were identified as *Bacillus* and *Halobacillus* species. In addition, isolates related to *Marinimicrobium agarilyticum* and two *Gammaproteobacteria* isolates were identified on plates for CG1, isolates related to *Halovibrio denitrificans* strain HGD 3 and *Alkalibacterium* sp. ARD M22 were identified from SH1 and one isolate affiliated with *Halomonas* from XH2. Moreover, growth started to appear on the plates grown in anaerobic conditions. These isolates gave off a pungent odour, characteristic of anaerobes. The anaerobic isolates were streaked onto fresh plates and grown in aerobic conditions to ensure that they were obligate anaerobes. No growth occurred on these plates, and these isolates were related to *Halanaerobium praevalens*.

By week three, new isolates were identified on plates for samples BJ1, EJ3, EN1 and SH1. These were related to *Halomonas* species such as *Halomonas ventosae*. In addition, an isolate related to *Filobacillus* sp. MO21 was isolated from SH1. Furthermore, week three saw the appearance of archaeal isolates. These showed as small, deep red colonies. They were later identified as isolates affiliated with *Halorubrum*. A single *Haloarcula* species was identified from sample EN1. No archaeal isolates were identified from samples CG1 or XH2.

New isolates continued to be seen into week four from all samples except CG1. Among these were isolates related to *Chromohalobacter* (from EJ3), *Bacillales* (from EN1), *Marinobacter*, *Salegentibacter* and *Amphibacillus* (from SH1) and *Bacillus*, *Gracilibacillus* and *Halomonas* (from XH2). More colonies of archaea were observed, but only an isolate related to *Halorubrum lacusprofundi* was identified as a new isolate. By week five, just one new isolate was observed from sample XH2, which was related to *Idiomarina loihiensis*. No new isolates were observed by week six.

Attempts to cultivate different organisms at different temperatures failed. Many of the organisms observed on the plates grown at 37°C were also observed on plates grown at 30°C, though some were not seen until later; for example, the archaea did not appear on the plates at 30°C until much after week 3. This may be because the organisms have wide temperature ranges for growth to enable them to adapt to seasonal changes in temperature. Perhaps incubations at much lower temperatures would allow cultivation of different organisms.

Overall, more *Bacteria* were isolated from these environments than *Archaea*. Most of the bacterial isolates observed were *Gammaproteobacteria*, with the dominant group being *Halomonas* species, and *Firmicutes*, with the dominant group being *Bacillus* species. All but one of the archaeal isolates were *Halorubrum*.

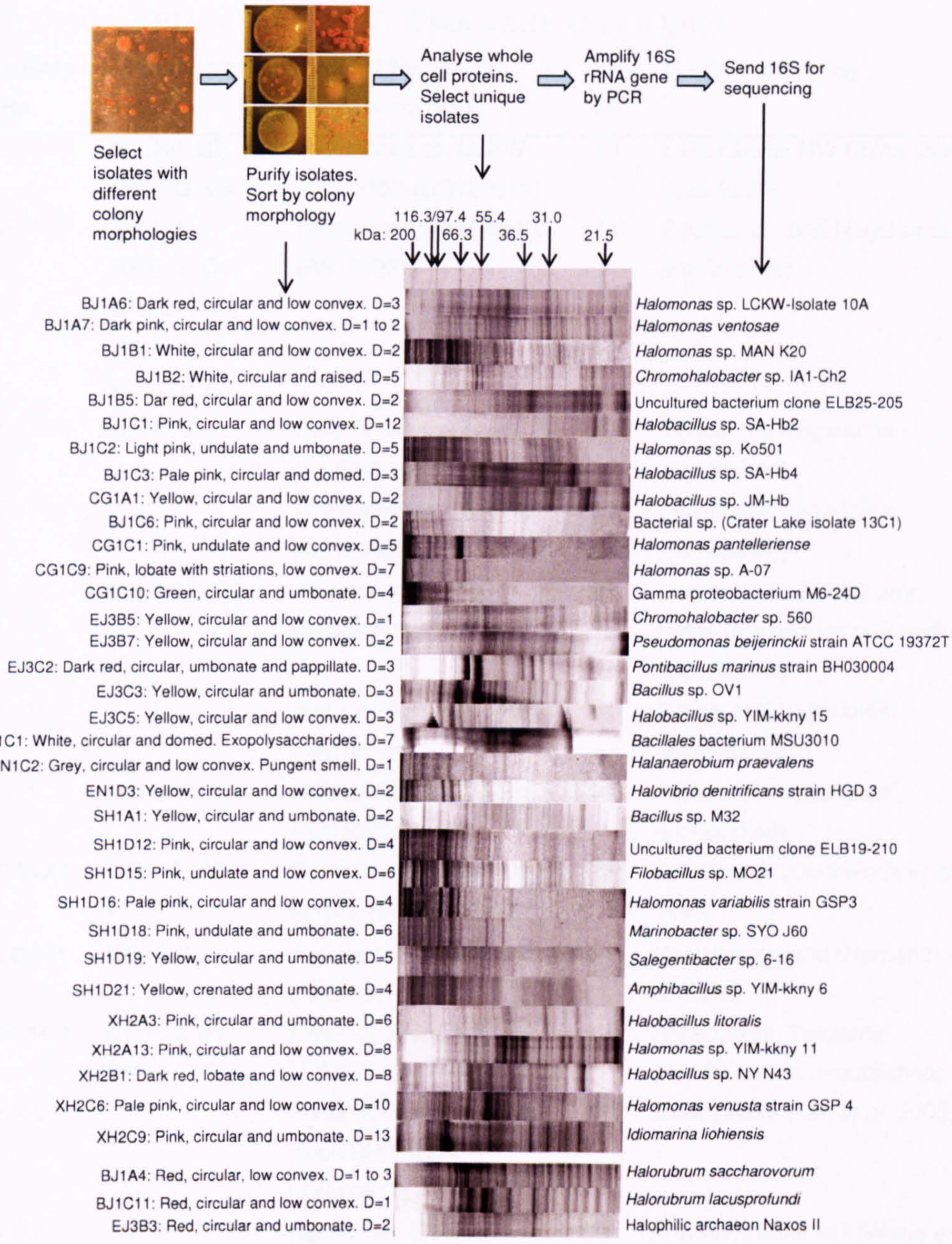


Figure 41 Whole Cell Protein Analysis.

Schematic diagram showing how isolates were purified, sorted and identified. The colony morphologies of the isolates and the whole cell protein analysis are illustrated. D is the diameter of the colonies in mm.

Table 21 Sequencing Results for the Cultivated Isolates in the Salt Lakes.

Nearest neighbours from the BLASTN database to the cultivated isolates are shown. The isolates have been colour coded according to the environment they are from (**Red** = BJ1, **Blue** = CG1, **Orange** = EJ3, **Magenta** = EN1, **Cyan** = SH1, **Grey** = XH2).

W [a]	Isolate No.	Other isolate Nos.[b]	Nearest Neighbour (Accession No.)	% ID[c]	Origin (reference)
1	**BJ1A6**	BJ1B6, B7, C5, CG1C8	*Halomonas* sp. LCKW-Isolate10A (DQ129871)	99	Lake Chaka, NW China (Jiang *et al.* 2006)
	BJ1B1	BJ1B3, XH2C1, C4, EN1D1, D5, D6, D7, D13, D14, D15	*Halomonas* sp. MAN K20 (AB166931)	98	? (Okamoto and Naganuma unpublished)
	BJ1B2	BJ1B10	*Chromohalobacter* sp. IA1-CH2 (AB189308)	99	? (Hua and Naganuma unpublished)
	BJ1C1		*Halobacillus* sp. SA-Hb2 (AB189302)	99	? (Hua and Naganuma unpublished)
	BJ1C2	SH1D11, D20	*Halomonas* sp. Ko501 (AF550585)	98	Smoker pipe, Kolbeinsey Ridge (Steinsbu and Torsvik unpublished)
	BJ1C3		*Halobacillus* sp. SA-Hb4 (AB189304)	99	? (Hua and Naganuma unpublished)
	BJ1C6	BJ1C8	*Halobacillus* sp. JM-Hb (AB189294)	99	? (Hua and Naganuma unpublished)
	CG1A1	CG1A2, A7	Bacterial sp. (Crater Lake Isolate 13C1) (X92155)	99	Soda Lakes (Duckworth *et al.* 1996)
	CG1C1	CG1B9	*Halomonas pantelleriense* (X93493)	98	Pantelleria island (Romano *et al.* 1996)
	CG1C9	CG1C3, B1, B4	*Halomonas* sp. A-07 (AY347310)	99	Soda Lakes, Tanzania (Pradhan *et al.* unpublished)
	EJ3C2		*Pontibacillus marinus* strain BH030004 (AY603977)	99	Solar saltern (Lim *et al.* 2005)
	EJ3C3		*Bacillus* sp. OV1 (AY647318)	99	Arizona pristine soil (Wang *et al.* unpublished-a)

[a] Week that these isolates first appeared on the plates.

[b] Isolates that turned out to be identical (via whole cell protein analysis or 16S rRNA gene sequencing).

[c] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

Table 21 Continued.

W [a]	Isolate No.	Other isolate Nos.[b]	Nearest Neighbour (Accession No.)	% ID[c]	Origin (reference)
2	CG1A5		*Bacillus saliphilus* (AJ493660)	96	Campania, Italy (Romano *et al.* 2005)
	CG1A6		*Halobacillus* sp. Y2 (DQ232875)	94	Submitted to database (Liu unpublished)
	CG1A8		*Marinimicrobium agarilyticum* (AY839870)	97	Tidal flat sediment, South Sea, Korea Strait (Lim *et al.* 2006)
	CG1B3	CG1C6	Gammaproteobacterium M6-24A (AY730234)	95	Mono lake, California, USA (Bano and Hollibaugh unpublished)
	CG1C10		Gammaproteobacterium M6-24D (AY730243)	95	Mono lake, California, USA (Bano and Hollibaugh unpublished)
	EJ3C5		*Halobacillus* sp. YIM-kkny15 (AY121438)	98	Lake Keke, W China (Chai *et al.* unpublished)
	EN1C2		*Halanaerobium praevalens* (AB022035)	99	Can of Swedish salted herrings (Kobayashi published only in NCBI database (1999))
	EN1D3	BJ1A9, EN1D4, SH1D6, D2, D8, XH2C7, C2	*Halovibrio denitrificans* strain HGD3 (DQ072718)	98	Hypersaline inland lakes (Sorokin *et al.* 2006)
	SH1D6B		*Alkalibacterium* sp. ARD M22 (AB167074)	99	? (Okamoto and Naganuma unpublished)
	XH2B1	XH2A1, A2	*Halobacillus* sp. NT N43 (AB166985)	99	? (Okamoto and Naganuma unpublished)
	XH2A3		*Halobacillus litoralis* (X94558)	98	Sediment, Great Salt Lake, Utah (Spring *et al.* 1996)
	XH2A13	SHD4, D7, D9, D14	*Halomonas* sp. YIM-kkny11 (AY121436)	97	Lake Keke, W China (Chai *et al.* unpublished)

[a] Week that these isolates first appeared on the plates.

[b] Isolates that turned out to be identical (via whole cell protein analysis or 16S rRNA gene sequencing).

[c] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

Table 21 Continued.

W [a]	Isolate No.	Other isolate Nos.[b]	Nearest Neighbour (Accession No.)	% ID[c]	Origin (reference)
3	BJ1A4	BJ1B11, BJ1A12B	*Halorubrum saccharovorum* (U17364)	98	San Francisco Bay (Tomlinson and Hochstein 1976)
	EJ3B3	EN1D16	Halophilic archaeon Naxos II (AJ400624)	99	Island of Naxos, Greece (Radax *et al.* unpublished)
	EN1C10 A		*Halorubrum* sp. F100 (DQ309090)	99	Salt lake, Turkey (Ozcan *et al.* unpublished)
	EN1D10		*Haloarcula* sp. AJ4 (AY208973)	98	? (Xu and Wu unpublished)
	SH1D13		*Halorubrum aidingense* strain 31-hong (DQ355813)	98	Salt lake, Xin-Jiang, China (Cui *et al.* 2006)
	BJ1A7	SH1D10	*Halomonas ventosae* (AY268080)	99	Saline soils, Jaén, Spain (Martínez-Cánovas *et al.* 2004)
	BJ1B5		Uncultured bacterium clone ELB25-205 (DQ015839)	96	Lake Bonney, Antarctica (Glatz *et al.* 2006)
	SH1D12		Uncultured bacterium clone ELB19-210 (DQ015832)	96	Lake Bonney, Antarctica (Glatz *et al.* 2006)
	SH1D15		*Filobacillus* sp. MO21 (AY553112)	93	Great Salt Plains of Oklahoma (Caton *et al.* 2004)
	SH1D16		*Halomonas variabilis* strain GSP3 (AY505527)	98	Great Salt Plains of Oklahoma (Caton *et al.* 2004)
4	BJ1C11		*Halorubrum lacusprofundi* (X82170)	98	Deep Lake, Antarctica (Franzmann *et al.* 1988)
	EJ3B5		*Chromohalobacter* sp. 560 (AB105160)	99	Saline environments (Tokunaga *et al.* unpublished)
	EJ3B7	EJ3B9	*Chromohalobacter beijerinckii* (AB021386)	99	Salted beans (Peçonek *et al.* 2006)

[a] Week that these isolates first appeared on the plates.

[b] Isolates that turned out to be identical (via whole cell protein analysis or 16S rRNA gene sequencing).

[c] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

Table 21 Continued.

W [a]	Isolate No.	Other isolate Nos.[b]	Nearest Neighbour (Accession No.)	% ID[c]	Origin (reference)
4	EN1C1		*Bacillales* bacterium MSU3010 (AY647316)	98	Saline soil, Mahasarakham Province, Thailand (Chookietwattana *et al.* unpublished)
	SH1A1		*Bacillus* sp. M32 (AY912092)	99	Wadi El-Natrun in Egypt, Africa (Ghozlan *et al.* unpublished)
	SH1D18		*Marinobacter* sp. SYO J60 (AB166912)	96	? (Okamoto and Naganuma unpublished)
	SH1D19		*Salegentibacter* sp. 6-16 (AJ783959)	98	Near-shore marine-cultural sediment (Wang *et al.* unpublished-b)
	SH1D21	SH1D22	*Amphibacillus* sp. YIM-kkny6 (AY121432)	99	Lake Keke, W China (Chai *et al.* unpublished)
	XH2A9		*Bacillus baekryungensis* strain GSP56 (AY505508)	94	Great Salt Plains of Oklahoma (Caton *et al.* 2004)
	XH2A10	XH2A11	*Gracilibacillus orientalis* (AM040718)	99	Lake Ejinnor, Inner Mongolia (Carrasco *et al.* 2006)
	XH2C6		*Halomonas venusta* strain GSP4 (AY553064)	98	Great Salt Plains of Oklahoma (Caton *et al.* 2004)
5	XH2C9		*Idiomarina loihiensis* L2TR (AF288370)	99	Hydrothermal vent, Lō'ihi' submarine volcano, Hawaii (Hou *et al.* 2004)

[a] Week that these isolates first appeared on the plates.

[b] Isolates that turned out to be identical (via whole cell protein analysis or 16S rRNA gene sequencing).

[c] Percentage identity (% ID) of the clone sequence compared to its nearest neighbour.

Phylogenetics of Cultivated *Archaea* in Salt Lakes

The 16S rRNA sequences for the archaeal isolates were sequenced from both ends of the PCR product, to give approximately 1.4 kb of sequence information for phylogenetic analysis. An unrooted tree was drawn containing the isolate sequences and their nearest neighbours, shown in Figure 42a (top panel). The isolates that were recovered from the salt lakes were all from the order *Halobacteriales*. All isolates belonged to the genus *Halorubrum*, except for one, identified as a *Haloarcula*. The two clades within the *Halorubrum* and the *Haloarcula* are well supported with bootstrap values of 100%. A second phylogenetic tree was drawn using only the first 800 bp of the 16S rRNA sequence from cultivated isolates, their nearest neighbours and related 16S rRNA sequences from the molecular analysis in the previous chapter (Figure 42b; bottom panel).

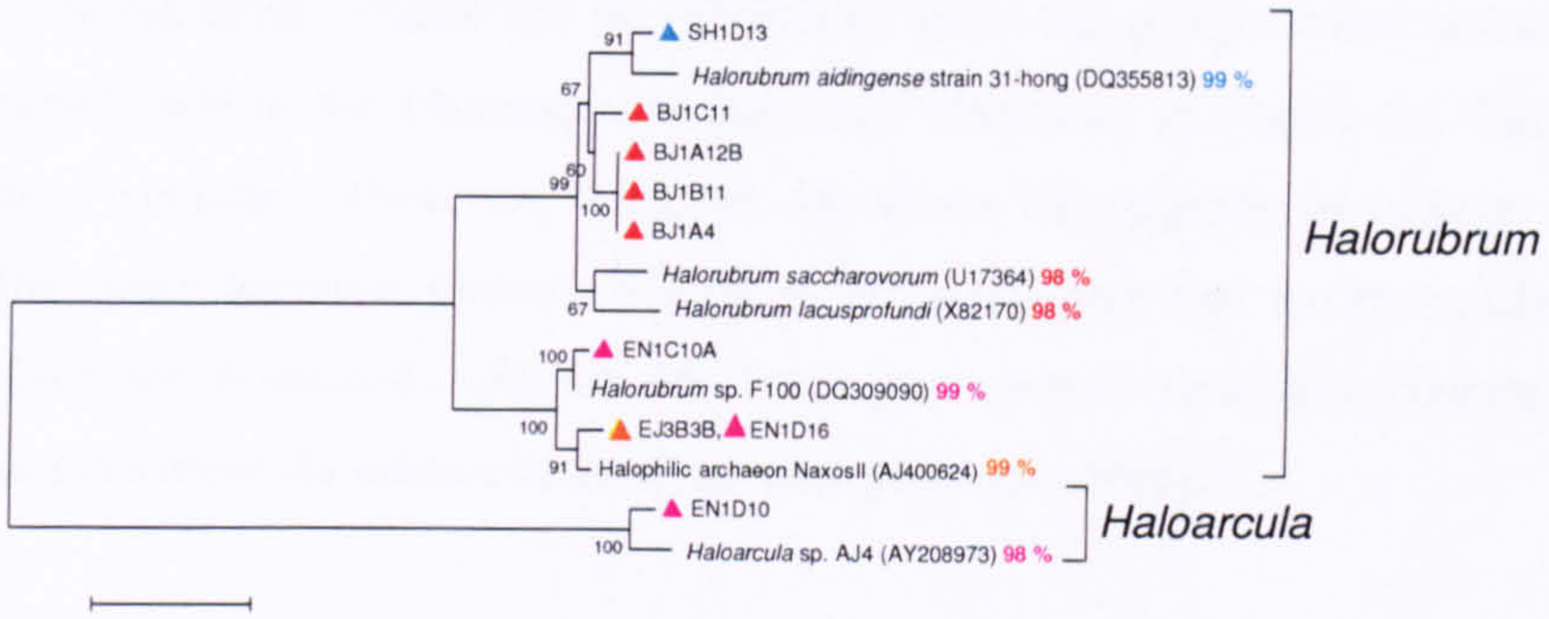
Halorubrum

Isolate SH1D13 is closely related to *Halorubrum* sp. 31-hong (99%), which has been renamed to *Halorubrum aidingense* strain 31-hong, and was cultivated from Aiding salt lake in Xin-Jiang, China. *Halorubrum aidingense* cells are motile, Gram negative rods. It is an aerobe that grows at a salinity range of 1.7 – 4.3 M NaCl, but grows optimally at 2.6 M. It can also grow between pH 7 to 9, but grows best at pH 7.5 and has a temperature range of 25°C – 52°C, but grows optimally at 40°C – 42°C (Cui *et al.* 2006). Isolate BJ1C11 is related to *Halorubrum lacusprofundi* (98%), which was first isolated from Deep Lake, Antarctica. Originally thought to be a psychrophile, further research determined that *Halorubrum lacusprofundi* had an optimum temperature for growth from 31°C – 37°C (Franzmann *et al.* 1988). Isolates BJ1A4, A12B and B11 are related to *Halorubrum saccharovorum* (98%), which was detected in the molecular analysis and has therefore been previously described. Isolates BJ1A4 and BJ1B11 have had further taxonomy analysis and appear to be a novel species (A. Ventosa, personal communication). Isolate EN1C10A is related to *Halorubrum* sp. F100 (99%), which was isolated from a salt lake in Turkey (Ozcan *et al.* unpublished). Isolate EJ3B3B is closely related to a halophilic archaeon NaxosII (99%), which was cultivated from the island of Naxos, Greece (Radax *et al.* unpublished).

Haloarcula

Isolate EN1D10 is related *Haloarcula* sp. AJ4 (98%), which is an extreme halophile, whose origin is not known (Xu and Wu unpublished).

a



b

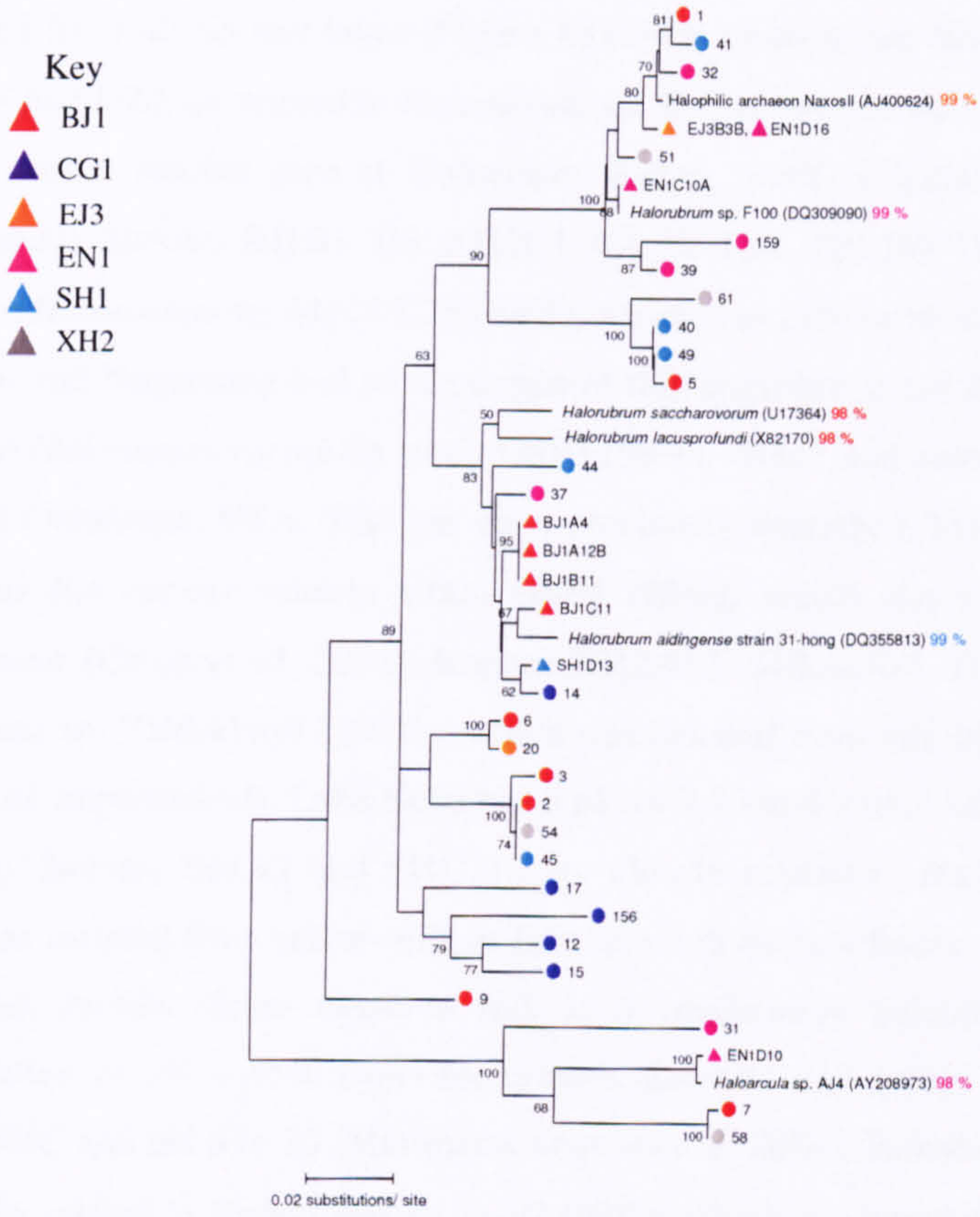


Figure 42 Phylogenetic Tree of Cultivated *Archaea*.

a; Unrooted phylogenetic tree of the relationships between the cultivated isolates with existing 16S rRNA gene sequences for *Halorubrum* and *Haloarcula*.

b; Unrooted phylogenetic tree of the relationships between the cultivated isolates with existing 16S rRNA gene sequences and related sequences from the molecular analysis (circles; see previous chapter) for *Halorubrum* and *Haloarcula*. The closest percentage similarities (identities) between the cultivated isolates and known sequences are indicated.

Phylogenetics of Bacterial Isolates

The cultivated bacterial isolates can be subdivided into four groups. Most isolates belong to the halomonads within the *Gammaproteobacteria*, illustrated in Figure 43. The other large group is the *Firmicutes*, illustrated in Figure 44, where the majority of isolates are *Bacillus* species. The other bacterial isolates belong to *Bacteroidetes* and an unclassified bacterial species, which are illustrated in Figure 45. These phylogenetic trees also contain related 16S rRNA sequences from the molecular analysis (see previous chapter).

Gammaproteobacteria

The *Gammaproteobacteria* comprises 67% of all isolates, making it the largest group cultivated from all six salt lakes (Figure 43). 76% of these are halomonads. Isolates BJ1C2, SH1D11 and D20 are related to *Halomonas* sp. Ko501 (98%), which was isolated from basalt and an extinct smoker pipe at Kolbeinsey Ridge, north of Iceland (Steinsbu and Torsvik unpublished). Isolates BJ1B1, B3, XH2C1, C4, EN1D1, D5, D6, D7, D13, D14 and D15 are related to *Halomonas* sp. MAN K20 (98%), which was cultivated in the unpublished study by Okamoto and Naganuma and so the origin of this organism is not known. Isolate SH1D16 is related to *Halomonas variabilis* strain GSP3 (98%), which was cultivated from the Great Salt Plains of Oklahoma, USA. This has been previously described. Moreover, isolate XH2C6 is related to *Halomonas venusta* strain GSP4 (98%), which was cultivated from the same environment (Caton *et al.* 2004). Isolates XH2A13, SHD4, D7, D9 and D14 are related to *Halomonas* sp. YIM-kkny11 (97%), which was isolated from salt lake Keke in western China (Chai *et al.* unpublished). Lake Keke has a pH of 7.25 and a total salinity of 185 g/L (Xiao *et al.* 2000). Isolates BJ1A7 and SH1D10 are closely related to *Halomonas ventosae* (99%), which was isolated from saline soils in Jaén in south eastern Spain. *Halomonas ventosae* is a capsulated, motile, Gram negative rod. It is moderately halophilic, requiring total salt concentration of 3% – 15% (w/v) for growth. Growth also occurs at temperatures between 15°C – 50°C and pH 6 to 10 (Martínez-Cánovas *et al.* 2004). Isolates CG1C9, C3, B1 and B4 are closely related to *Halomonas* sp. A-07 (99%), which is a novel alkaliphile from Tanzania soda lakes (Pradhan *et al.* unpublished). Isolate CG1C8 and isolates BJ1B6, B7, A6 and C5 are related to *Halomonas* sp. LCKW-Isolate10A (97% and 99% identity respectively), which was cultivated from Lake Chaka in north western China. The conditions in this lake have been previously described (Jiang *et al.* 2006). Isolates CG1C1 and B9 are related to *Halomonas pantelleriense* (98%), which is a haloalkaliphilic, strictly aerobic bacterium isolated from Pantelleria island, in the Sicily Channel (Romano *et al.* 1996). This island is the emergent

portion of a volcano that rises 1000 m above the sea floor and so is largely composed of volcanic rocks. The island contains a lot of pantellerite rock (Civetta *et al.* 1998). An isolate from BJ1 (BJ1B5) and an isolate from SH1 (SH1D12) are distantly related to uncultured bacterial clones ELB25-205 and ELB19-210, respectively (both at 96% identity), which were detected in Lake Bonney, a chemically stratified, permanently ice-covered Antarctic Lake. These particular clones were detected in the east lobe of the lake, where high concentrations of nitrate and nitrite occur due to denitrification in the west lobe. The temperature here ranges from -2°C at a depth of 40 m, to 5°C at a depth of 15 m. The salinity can vary, with conductivity readings ranging from 0 just below the ice surface, to 10 S/m at depths of 20 m to 40 m. ELB19 represents an area in the chemocline where the oxygen starts to decrease, while ammonium, N_2O, nitrate, nitrite and salinity increase; ELB25 represents an area where there is minimum oxygen, but high levels of ammonium, N_2O, nitrate and nitrite, and the salinity is ten times higher than sea water (Glatz *et al.* 2006). My phylogenetic analysis has clearly placed these clones with the halomonads. The isolates are interesting since the clones that they are related to clearly have not been previously cultivated; and since the isolates are not closely related anyway, they are probably novel bacteria that can be a potential focus for further study.

Isolates EJ3B7 and B9 are closely related to *Chromohalobacter beijerinckii* (previously named *Pseudomonas beijerinckii* strain ATCC 19372T) (99%) (Peçonek *et al.* 2006). This bacterium was isolated from salted beans and was first described by Hof in 1935. It is an aerobic, polarly flagellated, Gram negative, rod. It Isolates BJ1B2 and B10 are closely related to *Chromohalobacter* sp. IA1-Ch2 (99%), which was cultivated by Hua and Naganuma in a study that is currently unpublished, and so the origins of these organisms are not known. Isolate EJ3B5 is closely related to *Chromohalobacter* sp. 560, with an identity of 99% (Tokunaga *et al.* unpublished), which was isolated from a salted squid gut specimen (Tokunaga *et al.* 2004).

Isolates XH2C7, C2, BJ1A9, SH1D6, D2, D8 and EN1D3 and D4 are related to *Halovibrio denitrificans* strain HGD 3 showing 98% identity. This isolate has been previously described (Sorokin *et al.* 2006).

Other isolates belong to the family *Alteromonadaceae* within the *Gammaproteobacteria.* Isolate SH1D18 is distantly related to *Marinobacter* sp. SYO J60, with an identity of 96%

(Okamoto and Naganuma unpublished). Isolate CG1A8 is distantly related to *Marinobacter* strain M18 (97%), which has been renamed to *Marinimicrobium agarilyticum*, and was cultivated from tidal flat sediment of the Jeonnam Province of the South Sea of Korea (the Korea Strait). *Marinimicrobium agarilyticum* showed agar dissolving activity. It is a Gram negative, strictly aerobic, non-spore forming rod that is motile. It grew in the presence of 0% - 13% (w/v) NaCl, but grew optimally at 0 to 1% (w/v), making this bacterium moderately halotolerant rather than halophilic. It also grew at a temperature range of 12°C – 40°C (optimum 35°C) and a pH range of 6 – 10.5 (optimum pH 7 – 8) (Lim *et al.* 2006).

Within the *Idiomarinaceae*, isolate XH2C9 is closely related to *Idiomarina loihiensis* L2TR (99%), which is a deep sea bacterium, isolated from a hydrothermal vent at a depth of 1300 m at the Lõ'ihi' submarine volcano, approximately 30 km south of Hawaii. It was found to grow in partially oxygenated water at the periphery of the vent. It can grow in a range of temperatures from 4°C to 46°C and a wide range of salinities, from 0.5% to 20% (w/v) NaCl. The genome of this bacterium was recently sequenced and it was revealed that it catabolises amino acids rather than sugar for carbon and energy. It also revealed that it produces a viscous exopolysaccharide and so it was proposed that these bacteria survive in hydrothermal vents by attaching to proteinaceous particles by secreted exopolysaccharide, then degrade, uptake and metabolise the amino acids (Hou *et al.* 2004).

Isolate CG1C10 is distantly related to a gammaproteobacterium M6-24D (95%), while isolates CG1B3 and C6 are distantly related to a gammaproteobacterium M6-24A (95%); both isolates were originally found in Mono Lake in California, USA (Bano and Hollibaugh unpublished). This lake has been described previously.

Firmicutes

The second largest group to be cultivated from the salt lakes is the *Firmicutes*, which comprises 23% of all isolates (Figure 44). Isolate BJ1C3 is closely related to *Halobacillus* sp. SA-Hb4 (99%); isolate BJ1C1 is closely related to *Halobacillus* sp. SA-Hb2 (99%) and isolates BJ1C6 and C8 are closely related to *Halobacillus* sp. JM-Hb (99%), which are all from the same study (Hua and Naganuma unpublished). Isolate XH2A3 is related to *Halobacillus litoralis* (98%), which is a Gram positive, motile, spore-forming, strictly aerobic rod isolated from hypersaline sediments of the Great Salt Lake in Utah (Spring *et al.* 1996). This lake has an area of about 6000 km^2. Salinity of this lake has varied considerably due to

the rise in water levels over the years. The addition of a causeway also caused changes in salinity, ranging from 16% to 29% (w/v) in the north part of the lake, and in the south part from 6% to 28% (w/v) (Stephens 1990). Isolates XH2B1, A1 and A2 are closely related to *Halobacillus* sp. NT N43 (99%), which was isolated by Okamoto and Naganuma (unpublished). Isolate EJ3C5 is related to *Halobacillus* sp. YIM-kkny15 (98%), which was isolated from salt Lake Keke in western China, which was previously described. Moreover, isolate SH1D21 and 22 are closely related to *Amphibacillus* sp. YIM-kkny6 (99%), which is from the same lake (Chai *et al.* unpublished). Isolates XH2A10 and A11 are closely related to *Gracilibacillus* sp. EJ-15 (99%), which was cultivated from Lake Ejinnor in Inner Mongolia, and has been renamed to *Gracilibacillus orientalis*. It is a moderately halophilic, motile, Gram positive rod that forms terminal endospores. It is strictly aerobic and can grow at a NaCl range of 1% to 20% (w/v), but is optimal 10% (w/v). It can also grow at a pH range of 5 to 9 (optimal at pH 7.5), and a temperature range of 4°C to 45°C (optimal at 37°C) (Carrasco *et al.* 2006). Isolate CG1A5 is distantly related to *Bacillus* sp. 6AG (96%), which has been renamed to *Bacillus saliphilus* and was cultivated from an algal mat found in a mineral pool in Malvizza in Campania, Italy. This is a haloalkaliphilic, aerobic, Gram positive bacterium that grows optimally at 37°C, but growth is observed between 4°C and 50°C. It also grows at a pH range of 7 – 10, but grows optimally at pH 9, and is able to grow at > 0% to 25% (w/v) NaCl, but grows optimally at 16% (w/v) (Romano *et al.* 2005). Isolate EN1C1 is related to *Bacillales* bacterium MSU3010 (98%), which was cultivated from saline soil in Mahasarakham Province, Thailand (Chookietwattana *et al.* unpublished). Isolate SH1A1 is closely related to *Bacillus* sp. M32 (99%), which was cultivated from Wadi El-Natrun in Egypt, Africa (Ghozlan *et al.* unpublished). This depression contains several alkaline saline lakes, with salinities between 283 – 540 g/L and pH between 8.5 and 9.5 (Taher 1999). Isolate EJ3C3 is closely related to *Bacillus* sp. OV1 (99%), which was found in Arizona pristine soils (Wang *et al.* unpublished-a). Isolate EJ3C2 is closely related to *Pontibacillus marinus* strain BH030004 (99%), which is a moderately halophilic, motile, Gram positive, rod that was isolated from a solar saltern of the Yellow Sea in Korea. This strict aerobe grows at 15°C to 40°C, but grows optimally at 30°C, and can tolerate pH 6 to 9 (optimal at pH 7 – 7.5). It also grows in 1% to 9% (w/v) NaCl, with optimum growth at 2% – 5% (w/v) (Lim *et al.* 2005). Isolate XH2A9 is distantly related to *Bacillus baekryungensis* strain GSP56 (94%), which was found in the Great Salt Plains of Oklahoma (previously described). In addition, isolate SH1D15 is distantly related to *Filobacillus* sp. MO21 (93%), which is from the same sampling site (Caton *et al.* 2004). Isolate CG1A6 is distantly related to *Halobacillus* sp. Y2

(94%). Unfortunately the site that this isolate was found is not known as this sequence was submitted directly into the database (Liu unpublished). Isolate SH1D6B is closely related to *Alkalibacterium* sp. ARD M22 (99%), which is from the unpublished study by Okamoto and Naganuma. Isolate EN1C2 is related to *Halanaerobium praevalens* (99%), which is a strictly anaerobic halophile that was found in a can of Swedish salted herrings (Kobayashi published only in NCBI database (1999)). It has also been found in surface sediments of the Great Salt Lake in Utah, USA where salinity was more than 20% (w/v). It was later found that it grew optimally in 13% (w/v) total salt (Zeikus *et al.* 1983).

Bacteroidetes

Isolate SH1D19 is related to *Salegentibacter* sp. 6-16 (98%), which was cultivated from near-shore marine-cultural sediment (Wang *et al.* unpublished-b).

Bacterial Isolate

Isolates CG1A7, A1 and A2 are closely related to a Crater Lake isolate 13C1 (99%), which is a soda lake alkaliphile (Duckworth *et al.* 1996). This is an 'unclassified' bacterial isolate, but phylogenetic analysis has placed this clone near to a *Microbacterium* species, which is an *Actinobacteria* (Figure 45).

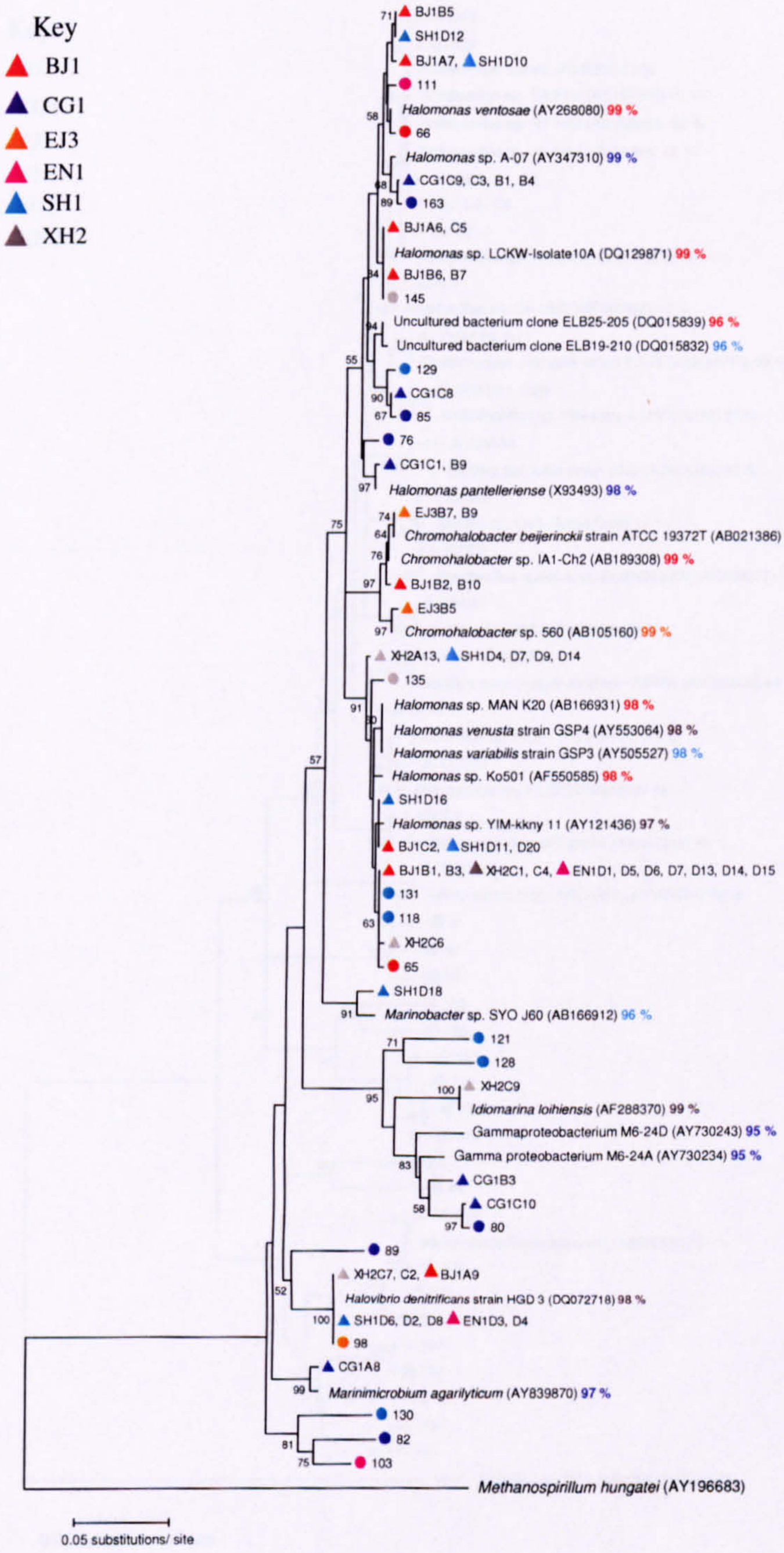


Figure 43 Phylogenetic Tree of *Gammaproteobacteria* in the Salt Lakes.

This phylogenetic tree shows the relationships between cultivated isolates with existing 16S rRNA gene sequences and related 16S rRNA sequences from the molecular analysis (circles; see previous chapter) for the *Gammaproteobacteria*. The closest percentage similarities between the cultivated isolates and known sequences are indicated.

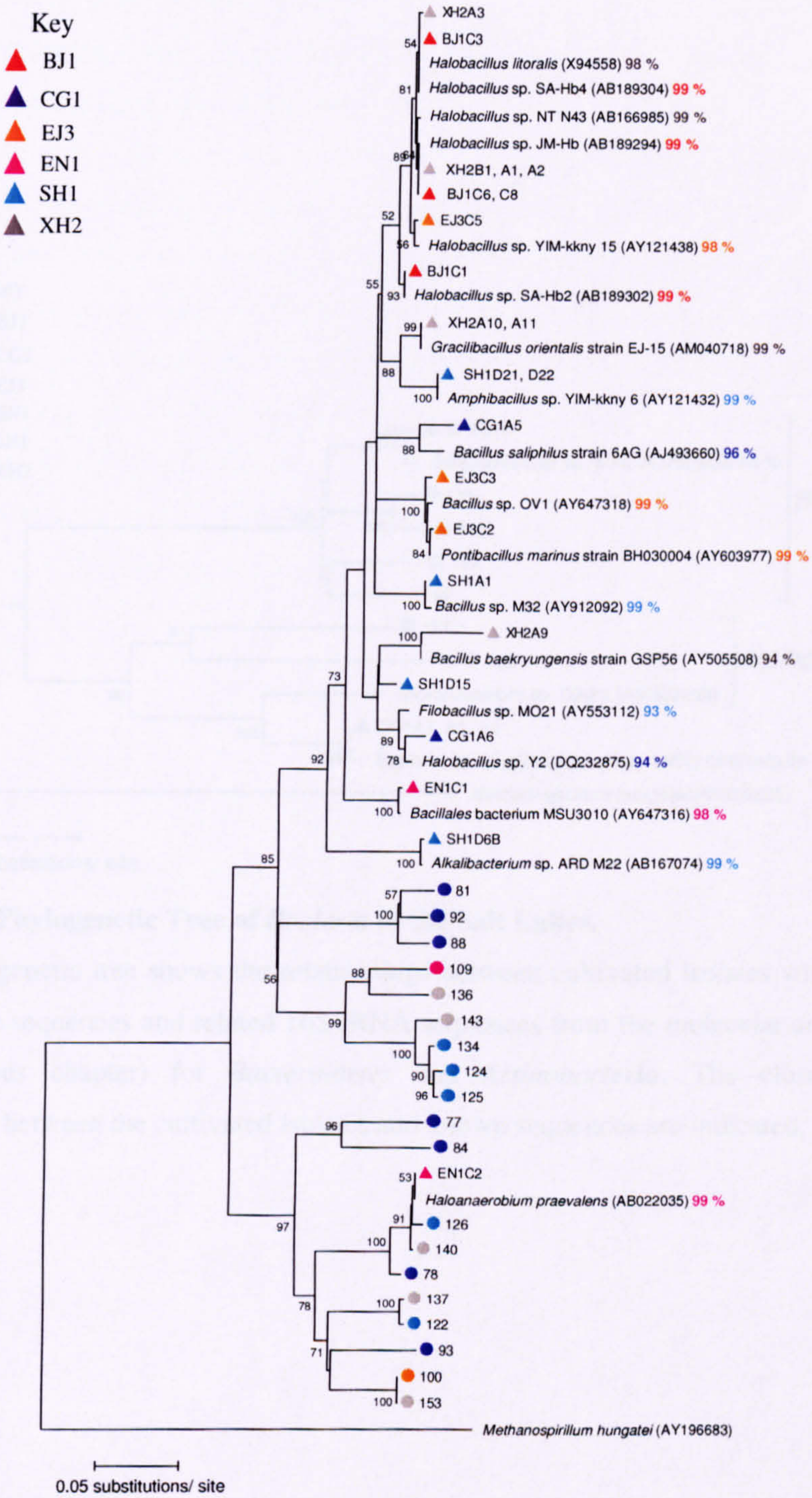


Figure 44 Phylogenetic Tree of *Firmicutes* in the Salt Lakes.

This phylogenetic tree shows the relationships between cultivated isolates with existing 16S rRNA gene sequences and related 16S rRNA sequences from the molecular analysis (circles; see previous chapter) for the *Firmicutes*. The closest percentage similarities between the cultivated isolates and known sequences are indicated.

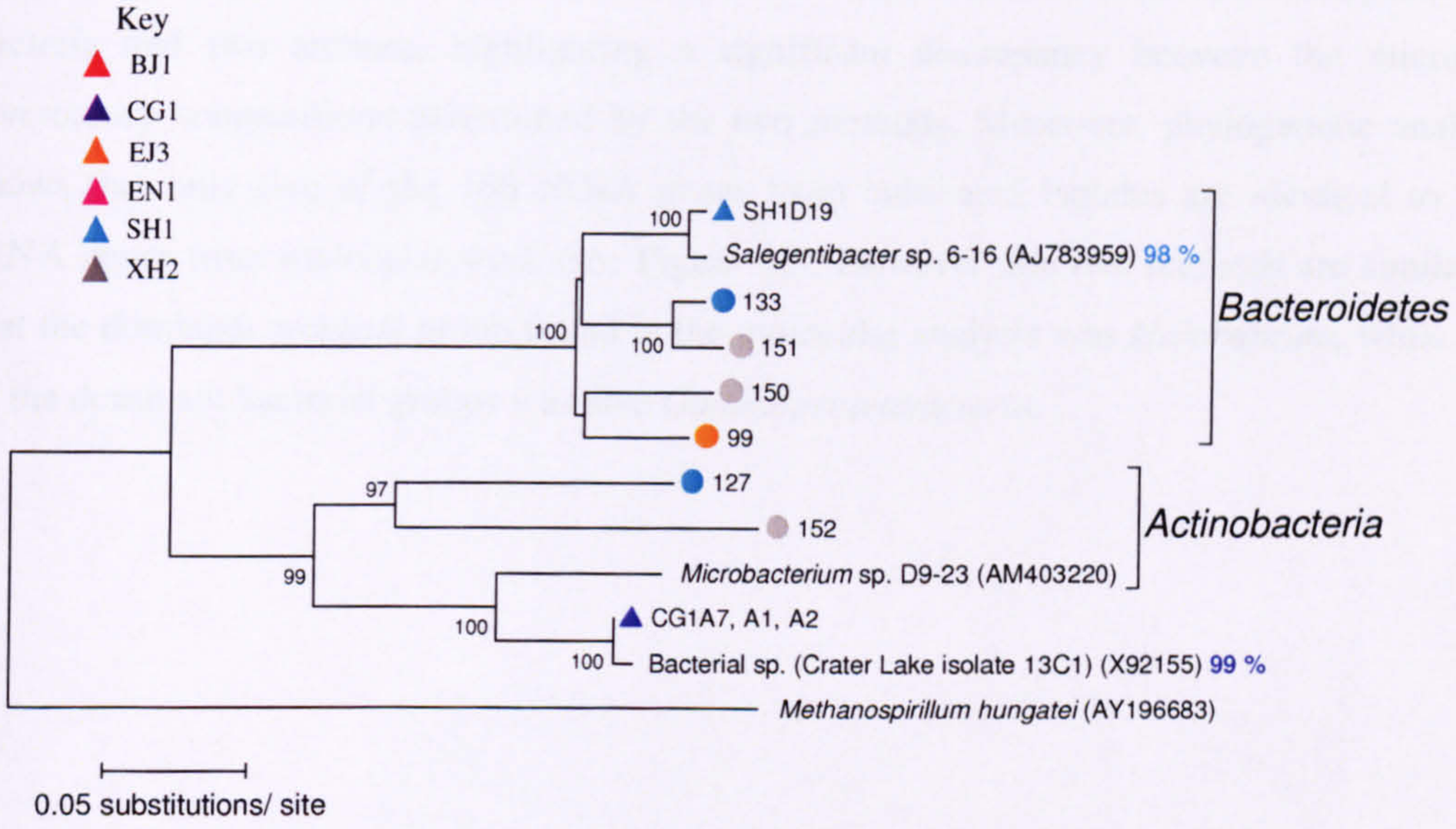


Figure 45 Phylogenetic Tree of *Bacteria* in the Salt Lakes.

This phylogenetic tree shows the relationships between cultivated isolates with existing 16S rRNA gene sequences and related 16S rRNA sequences from the molecular analysis (circles; see previous chapter) for *Bacteroidetes* and *Actinobacteria*. The closest percentage similarities between the cultivated isolates and known sequences are indicated.

Comparison with Molecular Analysis

Cultivation methods have shown that only a small diversity of isolates could be identified, with most of the bacterial isolates belonging to *Gammaproteobacteria* and archaeal isolates belonging to *Halorubrum*. This is a major difference with molecular methods, where a greater diversity of organisms was identified (see previous chapter). This is illustrated in Figure 46. Only six cultivated isolates were observed in the data from the molecular analysis; four bacteria and two archaea, highlighting a significant discrepancy between the microbial community compositions determined by the two methods. Moreover, phylogenetic analysis shows that only five of the 16S rRNA genes from cultivated isolates are identical to 16S rRNA genes from molecular work (see Figure 43). However, the two methods are similar in that the dominant archaeal group found in the molecular analysis was *Halorubrum*, while one of the dominant bacterial groups was also *Gammaproteobacteria*.

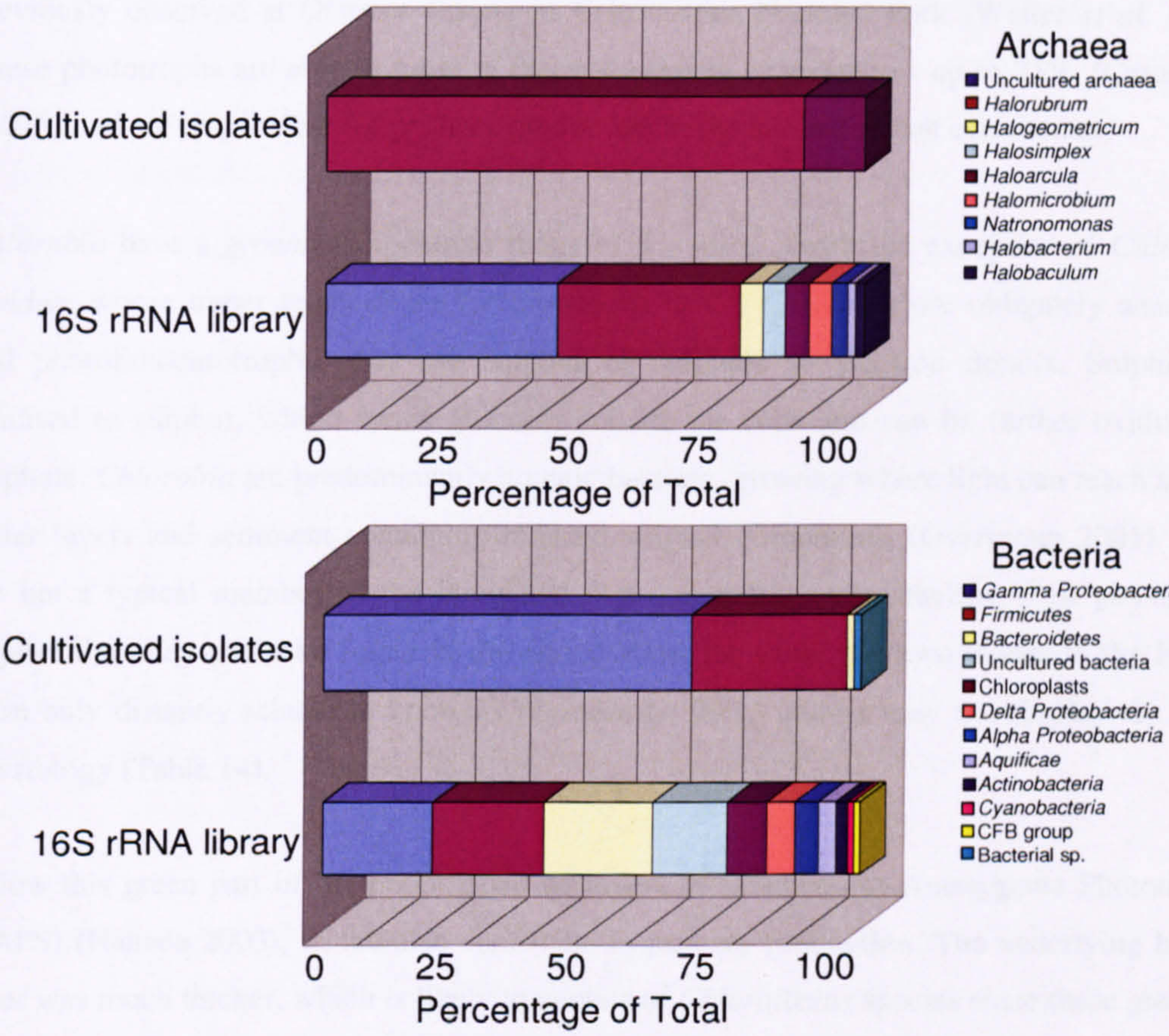


Figure 46 Diversity of Cultivated Isolates and 16S rRNA Clone Libraries.

This bar chart compares the 16S rRNA gene sequences from 163 sequenced clones and 58 sequenced isolates for the archaea (top panel) and bacteria (bottom panel) detected in the salt lakes.

DISCUSSION: MICROBIAL ECOLOGY OF THE HOT SPRINGS

Community Composition and Structure of Laminated Mat

The laminated mat at sampling site LP4 contained very clear striations. The uppermost green layer was only a few millimetres thick and is likely to be composed of cyanobacteria. One clone was related to a *Cyanobacterium* sp. (type I) (97%) (Table 14, p161), which was previously observed at Octopus Spring in Yellowstone National Park (Weller *et al.* 1992). These phototrophs are able to grow in thermal areas in temperatures up to 75°C (Castenholz 2001b) and are responsible for primary production in the laminated mat community.

Chlorobia have a growth temperature range of 5 – 40°C, (with the exception of *Chlorobia tepidum* whose upper temperature for growth is 52°C). *Chlorobia* are obligately anaerobic and photolithoautotrophic that use sulphur or sulphide as electron donors. Sulphide is oxidised to sulphur, which forms globules outside the cells and can be further oxidised to sulphate. *Chlorobia* are predominantly aquatic bacteria, growing where light can reach anoxic water layers and sediment containing reduced sulphur compounds (Overmann 2001). They are not a typical member of the laminated mat community; nevertheless, their physiology implies that they could be found in the photic zone. However, the two clones in the library were only distantly related to known *Chlorobia* (< 90%) and so may not display the same physiology (Table 14).

Below this green part of the photic zone lie layers of Filamentous Anoxygenic Phototrophs (FAPS) (Hanada 2003), which also contribute to primary production. The underlying brown layer was much thicker, which is likely to consist of *Chloroflexus* species since these greenish brown to brown filaments are known to be associated with cyanobacteria (Varnam 2000; Hanada 2003). Below this was a red layer community, which is likely to consist of *Roseiflexus*, a red filamentous bacterium (Hanada 2003). One clone from the library was related to an uncultured *Chloroflexaceae* (95%) (Table 14), which was previously detected in Witch Pond in Yellowstone National Park that was also part of a red layer community in a laminated mat, and was shown to be somewhat related to *Roseiflexus castenholzii* (Boomer *et al.* 2002).

This layered mat community lay on top of anoxic sediment. This is likely to contain other bacteria and archaea that ferment the organic substrates from the upper layers. Many bacteria

detected in the clone library such as the *Proteobacteria*, *Bacteroidetes* and *Actinobacteria* (Table 14) have growth temperature ranges that would mean that their growth would be restricted to lower layers of the mat where they are protected from the boiling water at the surface.

Other bacteria were detected in the clone library that are valid members of the hot spring community. However, it is not clear if these bacteria are part of the laminated mat community, or if they were inadvertently sampled from the surrounding water or sediment. The *Nitrospirae* are a deeply branching group of bacteria consisting of metabolically diverse members including nitrifiers, sulphate reducers and magnetotactic forms. Three clone sequences branched within this clade, showing relatedness to the genus *Thermodesulfovibrio* (Table 14 and Figure 29, p172), which is the only thermophilic genus in this phylum. It is a fermentative, sulphate reducing bacterium that oxidises substrates to acetate. It is strictly anaerobic and acidophilic, with a temperature range of 40 – 70°C (Maki 2001). The clones from the library represent distant relatives of *Thermodesulfovibrio* and may not show the same physiology; this can only be found by cultivation methods. Two clones from the library were affiliated to the *Deinococcales* (Table 14). *Deinococcales* are aerobic chemoorganotrophs, and can be either mesophilic or thermophilic, with growth occurring optimally at 25°C – 35°C and 45°C – 50°C respectively. The thermophilic representatives are often found in hot springs (Battista and Rainey 2001), though to my knowledge they are not often associated with laminated mats.

Methanogenesis is known to occur in laminated mats (Ward 1978), and so only methanogenic archaea have previously been found in these environments (Ward *et al.* 1998). Indeed, five clones from the library were affiliated with the *Methanobacteriales*, possibly contributing to methanogenesis within the laminated mat community. Other archaeal clones from the library are affiliated with the *Crenarchaeota* and other *Euryarchaeota* (that do not bear much resemblance to existing species) (Figure 27, p159; Figure 40, p220), however, these are not known to be associated with laminated mats. The clones are not closely related to known species (< 90%) and so their physiology and function in the community is not known.

Eukarya are not known to be associated with thermal laminated mats (Varnam 2000), and the fact that there were few *Eukarya* detected in the 18S rRNA library supports this (Figure 39, p216). The ones that were detected are likely to be part of the hot spring water community

that got trapped in the sample. Few *Eukarya* have been isolated from hot springs and so their role within the community is yet to be elucidated.

Comparison to Other Laminated Mats

The laminated mat was found in sampling site LP4, which had a temperature of 60 - 65°C and a pH of 8.5. Laminated mats are typically found in neutral to alkaline hot springs with temperatures up to approximately 55°C, and so this laminated mat is in its upper temperature limit for development.

The laminated mats that consist of layers of *Cyanobacterium* (including *Synechococcus*), as well as *Chloroflexus*, *Roseiflexus* and anoxic sediment have been found in a number of hot springs. Many examples have been found in Yellowstone National Park (Doemel and Brock 1977; Ward *et al.* 1998; Ramsing *et al.* 2000). In addition, laminated mats in Nakabusa hot spring, Japan contain these characteristic layers, associated with non-phototrophic thermophilic bacteria such as *Thermus* species and sulphur oxidising bacteria (SOB) (Hanada 2003).

Few clones in the library were affiliated to cyanobacteria, or *Chloroflexus* and *Roseiflexus*, though this may be because sampling of the clone library did not go to completion according to the rarefaction curves (Figure 25, p149), and thus further sampling from the library may throw up more of these sequences. *Synechococcus* are often found associated with filamentous cyanobacteria (Varnam 2000). However, no *Synechococcus* sequences were detected in the library. This may consistent with studies that demonstrate that *Synechococcus* are absent from hot springs in Iceland and Alaska, and high-temperature forms of *Synechococcus* are absent in New Zealand, Italy and Japan, indicating endemism of thermophilic cyanobacteria (Castenholz 1996; Ward *et al.* 1998) or it may again be because an insufficient number of clones were sampled.

Though not typical members of laminated mats, green-sulphur bacterium-like populations have previously been detected in the photic zone (Ramsing *et al.* 2000), which may explain the presence of *Chlorobia*-like species in the clone library. Moreover, *Chlorobia* species have been observed in New Zealand unlaminated mats (Overmann 2001), and green sulphur bacterium-like sequences have been retrieved from cyanobacterial mats at Octopus Spring

(Ward *et al.* 1998), suggesting that certain *Chlorobia* species are more heat adapted than previously recognised.

Other *Bacteria* in the library were affiliated with the *Nitrospirae*, *Deinococcales*, *Proteobacteria* (*Deltaproteobacteria* and *Betaproteobacteria* subdivisions), *Bacteroidetes* and *Actinobacteria* (Figure 34, p198). *Nitrospirae*-like sequences have previously been detected in the sediment of Obsidian Pool in Yellowstone National Park (Hugenholtz *et al.* 1998), so likewise, the clones in the LP4 library may have been from the underlying sediment. Thermophilic members of the *Deinococcales* are often found in hot springs (Battista and Rainey 2001), though presumably their aerobic metabolism would limit their distribution. *Proteobacteria* have been found associated with cyanobacterial mats (Nold *et al.* 1996; Ward *et al.* 1998). *Bacteroidetes* and *Actinobacteria* have been detected in inter-tidal hot springs (Hobel *et al.* 2005). Hence these other bacteria are valid inhabitants of hot springs.

Methanogenesis has long been recognised in laminated mats (Ward 1978) and so the only archaea that have previously been found associated with cyanobacterial mats are methanogens, which was confirmed in the molecular analysis. However, the finding of several *Crenarchaeota* sequences is surprising since there has been little evidence showing their association with laminated mats. Archaeal signatures affiliated with the *Crenarchaeota* were detected in hot spring sediments at Yellowstone National Park (Barns *et al.* 1994) and so the clone sequences related to *Crenarchaeota* in the LP4 library may be from sediment underlying the laminated mat. Similarly, other *Euryarchaeota* are not known to be associated with laminated mats, however, several euryarchaeal like sequences were detected in the library that do not bear much resemblance to known species. Euryarchaeal sequences were detected in the hot spring water in Nagasaki, Japan that also did not bear any similarity to existing representatives (Takai and Sako 1999); hence the clone sequences relating the *Euryarchaeota* in the LP4 library may also be from the overlying hot water.

Eukarya are not known to be associated with laminated mats (Varnam 2000). Similarly, eukaryotic algae were shown to be in very little numbers, or completely absent from microbial mats in Arctic hot springs (Roeselers *et al.* 2007). Unfortunately, few studies have been done on *Eukarya* associated with hot springs. One study conducted on a variety of environments in Yellowstone National Park, showed that fungi could be isolated from

environments at 62°C, while eukaryotic algae were absent from environments above 56°C – 60°C (Tansey and Brock 1972).

Community Composition and Structure of Streamer Mat

The microbial community observed at TC8 is composed of macroscopic white bundles, forming streamers that were firmly stuck to the surrounding rocks. The majority of the clones were related to an *Aquificales* clone (93 – 94%) (Table 14). *Aquificae* are hyperthermophilic, microaerophilic and chemolithotrophic bacteria that grow best between pH 6 and 8 (with the exception of *Hydrogenobacter acidophilum*) (Reysenbach 2001c). It has been suggested that the *Aquificae* are the main primary producers in these streamer mats (Skirnisdottir *et al.* 2000) since *Aquifex* sp. are capable of fixing CO_2. They utilise molecular hydrogen, thiosulphate and elemental sulphur as electron donors, and use oxygen and nitrate as electron acceptors (Reysenbach 2001c).

Other *Bacteria* in the library were affiliated with the *Thermales* (*Deinococcus-Thermus*), *Thermotogae*, *Actinobacteria* and the *Deltaproteobacteria* (Figure 34). One clone was related to *Thermus rehai* (99%) (Table 14). *Thermus rehai* was cultivated from a water sample from a hot spring, and not from a microbial mat (Lin *et al.* 2002), hence it is possible that this bacterium was trapped in the streamers and is not part of the mat community. *Thermus rehai* is capable of reducing nitrate to nitrite and utilising glucose, acetate and galactose for growth (Lin *et al.* 2002), and so may have a role in breaking down organic substrates. Some strains of the order *Thermales* may also be chemolithoheterotrophic, oxidising sulphur compounds (Rainey and Da Costa 2001). One clone was somewhat related to *Caldotoga fontana* (94%), which was affiliated with the *Thermotogae* (Table 14 and Figure 29). These are anaerobic heterotrophs that ferment a range of substrates including yeast extract, glucose and xylose; hence may also have a role in breaking down organic substrates (Reysenbach 2001d). One clone was affiliated with the *Actinobacteria* (93% identical to *Friedmanniella spumicola*) and another with the *Deltaproteobacteria* (93% identical to an uncultivated clone) (Table 14), once again demonstrating that some members of these lineages may be more adapted to high temperatures that previously recognised.

Three of the clones from the library (D11, D15, D14) appear to form their own novel lineages, though only two of these show high bootstrap values (HS Clusters 6 and 7) (Figure 29).

Nevertheless, this highlights novel thermophilic bacteria residing in this hot spring environment. It is uncertain what roles these bacteria play within this community.

Archaea were associated with the white streamers. One clone from the library was affiliated with the *Korarchaeota*, which have only previously been detected in Yellowstone National Park by their 16S rRNA genes (Barns *et al.* 1996) and their biological significance has not been elucidated as they have not been cultivated. Several clones were affiliated with the *Crenarchaeota*; one clone was related to an uncultured *Desulfurococcales* (96%) (Table 13, p153); they are hyperthermophilic anaerobes, facultative anaerobes or aerobes. Like the *Aquificae*, they can grow chemolithotrophically by hydrogen oxidation, with elemental sulphur, nitrate, nitrite or thiosulphate and CO_2 as a sole carbon source. However, they can also grow organotrophically by sulphur respiration of various organic substrates or by fermentation (Huber and Stetter 2001b), thereby providing a dual role in carbon fixation and utilising organic substrates from primary production. Two other clones were affiliated with *Euryarchaeota*; one was related to a clone that was detected in rice soils (98%) and has possible methanogenic properties (Wu *et al.* 2006b) and the other is related to a known methanogen, *Methanothermobacter thermoautotrophicus* (99%) (Table 13). Therefore, it is tempting to speculate that some members of the consortium produce methane, though further studies would be needed to confirm this. Two clones formed their own well supported lineages, with bootstrap values of 100% (Figure 27), indicating the presence of two novel lineages within this consortium that are distinct from known organisms. The role that these organisms play in the community is therefore unknown.

Comparison to Other Streamer Mats

There have been a number of streamers around the world whose microbial composition has been characterised (Reysenbach *et al.* 1994; Yamamoto *et al.* 1998; Reysenbach *et al.* 2000b; Skirnisdottir *et al.* 2000; Takacs *et al.* 2001; Nakagawa and Fukui 2002). Although they appear in a range of colours they are all predominantly made up of *Aquificales*, as for the streamers observed at TC8. Streamers consist of colourless sulphur bacteria (*Aquificae*) and elemental sulphur particles that form in shallow streams of sulphide-containing hot springs, where temperatures reach 73°C, with a pH range of 6 to 9 (Yamamoto *et al.* 1998). This presumably means that the water at TC8 was high in sulphide, but no data was obtained to confirm this. The TC8 sampling site had a pH of 6.5, which is very close to the optimal pH for growth for *Aquifex* sp. (pH 6.8) (Reysenbach 2001c).

In a study by Skirnisdottir *et al* (2000) known sequences from streamer-forming *Aquificae* were constructed into a phylogenetic tree and the sequences were split into discrete branches. Bacteria in the P-branch were dominant in the high-temperature, low sulphide springs; bacteria in the J branch were dominant the high sulphide springs, while bacteria in the H branch were found in low ratios in some low sulphide springs. The same phylogenetic tree was attempted here, but with a representative sequence from TC8 incorporated. This tree showed that the *Aquificae* at TC8 was different to all those previously documented, forming its own separate lineage (Figure 28, p165); hence the streamers at TC8 represent a novel streamer community.

Other *Bacteria* and *Archaea* were observed in the TC8 libraries (Table 14). *Thermotogae*, *Deinococcus-Thermus*, *Proteobacteria*, *Crenarchaeota* and *Korarchaeota* have been observed at other streamer communities (Yamamoto *et al.* 1998; Reysenbach *et al.* 2000b; Skirnisdottir *et al.* 2000). *Actinobacteria* and *Euryarchaeota* have not been detected previously in streamer communities; hence the overall community composition of the streamers at TC8 is unique. Such differences have been explained by the geochemical differences and therefore physiological activity in the environments that they inhabit as well as stream flow, UV, pH. However, whether these other *Bacteria* and *Archaea* are a functioning part of the streamer community remains to be seen. The fact that the TC8 *Aquificae* did not affiliate closely with any known sequences suggests that the conditions at TC8 are unique; however, further investigation is required to elucidate this.

Few studies have been conducted on streamer mat communities, but it has been suggested that the *Aquificae* are the main primary producers in the mat and that aerobic sulphur and hydrogen oxidising bacteria are found at the surface and sulphate reducers inhabit the darker undermass (Skirnisdottir *et al.* 2000).

DISCUSSION: MICROBIAL ECOLOGY OF THE SALT LAKES

Geochemistry of Inner Mongolian Salt Lakes

The Inner Mongolian salt lakes are athalassohaline formed from leachate of surrounding volcanic rocks (see Introduction). The chemistry of salt lakes tend to vary with seasons (e.g. dilution of brine due to rainfall), therefore, it was not until the expedition that the chemistry of the salt lakes and salterns could be elucidated. These conditions may not be permanent, but it is important to recognise the sort of environments they were at the time of sampling.

pH measurements (Figure 21, p138) suggest that Lake Ejinnor (EJ3), Lake Bagaejinnor (BJ1) and the Unnamed lake (XH2) were near neutral lakes, where the predominant ions were Na and Cl ions (Table 19, p 226). Presumably the surrounding rocks were also high in Ca ions, which is known to remove alkaline CO_3 ions from solution to keep the brine neutral (Grant *et al.* 1998). These lakes are therefore classified as near neutral hypersaline lakes. Geographically, these three lakes were in the east of the region, separate from the others (Figure 14, p128), and consequently share a similar geochemistry. EJ3 was at a point in Lake Ejinnor that had been developed into salterns. They were still red and so it was likely that salt was harvested.

Lake Chagannor (CG1) was clearly a soda lake, since it was situated near the largest soda works in China. It also had a high pH (10.5) and chemical analysis of the brine showed high levels of CO_3 and HCO_3 ions. It had a lower salinity compared to the previous lakes, though Na and Cl ions were still the major ions (Table 19). Presumably, the surrounding geology was deficient in Ca and Mg ions, allowing CO_3 ions to remain in solution and raising the pH (Grant *et al.* 1998). It is therefore classified as a less hypersaline soda lake.

The soil around Lake Shangmatala (SH1) appeared to be soda, but pH measurements (Figure 21) and chemical analysis of the lake water (Table 19) showed that it was also predominantly a near neutral hypersaline lake. Likewise, Lake Erliannor is known to contain a lot of underground soda, but this was not apparent from the pH measurements (pH 8) or the chemical analysis of the lake water. This too is classified as a near neutral hypersaline lake.

Sampling point EN1 was also at an area in Lake Erliannor that had been developed into salterns. The warmer months promote the growth of haloarchaea and the reddening cause

precipitation of salt for harvesting. The higher air temperature during the summer months would also promote rates of evaporation. At the time of sampling, the salterns were white and so underground soda was harvested instead. The Lakes Erliannor, Chagannor Shangmatala were situated in the south west of the region, so again it is possible that they have a similar geochemistry.

Black sediment at EJ3 and XH2 suggests biological sulphate production and the odour at EJ3 and SH1 suggests microbial activity producing gases such as hydrogen sulphide (Grant *et al.* 1998).

Lake Chahannor was not sampled since it was clear that it was a freshwater lake (data not shown). However, previous studies have shown this to be a soda lake (Xin *et al.* 2001; Xu *et al.* 2001). Layers of alternating mud and trona are common in Quaternary lakes in China (Zheng *et al.* 1993) and so it is possible that this soda lake may not be lost, but is just undergoing another freshened phase.

Community Composition and Structure of the Salt Lakes

Eukarya

Lush grassland and vegetation was only observed around SH1. Plants growing near to the surface of salt lakes often contribute to organic matter (Ollivier *et al.* 1994). Brine shrimp (*Artemia* sp.) were observed in XH2, which is a common eukaryote observed in saline waters (Grant 2004).

The presence of another eukaryote was indicated by the amplification of chloroplast genes from *Dunaliella salina* at EJ3, EN1 and SH1 (Table 16, p190). The amplification of chloroplast genes from chlorophyte algae has also been observed in other hypersaline environments (Humayoun *et al.* 2003). *Dunaliella* are responsible for the primary productivity in salt lakes, which is able to support aerobic and anaerobic heterotrophic bacteria. *Dunaliella* utilise glycerol as a compatible solute to cope with the high salinity environment (Grant 2004), and such algal exudates are also a major source of organic carbon.

Although *Dunaliella salina* was found to be the most dominant clone at EN1 and SH1, this is quite possibly an artefact of multiple copies of the 16S rRNA gene in multiple chloroplasts

present in these organisms. Hence it may not indicate that these were the most common organisms in these lakes.

Archaea

The archaeal community of the salt lakes was largely composed of organisms belonging to the order *Halobacteriales*. Many members of the *Halobacteriales* are extremely halophilic (Grant *et al.* 2001), reflecting the hypersaline nature of these environments. They possess red carotenoids (Grant *et al.* 2001) that are responsible for the colouration of the salt crusts around the lakes and the water in the saltern at EJ3.

Halorubrum species were the dominant haloarchaeal constituent and were ubiquitous i.e. occurring in libraries for all six salt lakes (Table 15, p174). These are extremely halophilic and strictly aerobic organisms requiring between 1.5 and 5.2 M NaCl for growth (Grant *et al.* 2001), suggesting that the brine was at least partially oxygenated, though no data for this was acquired. The majority of the cultivated organisms were also *Halorubrum* species (Table 21, p232; Figure 42, p237). The *Halorubrum* species detected in CG1 were haloalkaliphilic (*Halorubrum tibetense* and *Halorubrum vacuolatum*), while the species in the other lakes were neutrophilic (Table 15), which is consistent with the alkaline pH of CG1 and the near neutral pHs of the other sampling sites (Figure 21).

Natronomonas were observed in CG1 and XH2. Members of this group are also haloalkaliphilic and grow at 2 - 5.2 M NaCl and pH 7 - 10 (Grant *et al.* 2001). Moreover, one clone from the CG1 library affiliated with the *Halobaculum*, and was related (95%) to a clone previously detected in Lake Zabuye, which is an alkaline salt lake with a pH of 9.4 (Fan *et al.* 2004). Therefore, all the haloarchaea detected at CG1 are consistent with the lake's high pH. However, it was surprising that no more alkaliphilic species belonging to other genera were observed in this lake.

The species present in the Salt Lake (SL) Clusters (Figure 31, Figure 32, Figure 33) have < 90% similarity to documented sequences, thus it is not possible to deduce their physiology or roles within the communities. Clones from BJ1 and CG1 libraries were related to clones previously detected in Organic Lake in the Vestfold Hills, Antarctica, which was a major component of the haloarchaeal community in CG1 as demonstrated by clone frequencies (Table 15). Moreover, one isolate that was cultivated from BJ1 showed 98% identity to

Halorubrum lacusprofundi, which was initially isolated from Deep Lake, Antarctica (Franzmann *et al.* 1988). This suggests that some of the haloarchaea in Lakes Bagaejinnor and Chagannor have a broad temperature range for growth, which may allow them to remain viable throughout the year, particularly during the winter months. Further investigation would be needed to prove this, for example, by incubating the isolate at low temperatures and looking for growth.

SL Cluster 3 (Figure 33, p188) was a group that formed its own lineage distinct from the order *Halobacteriales*, which were present at EJ3, EN1, SH1 and XH2. Several clones from the EJ3, EN1, SH1 and XH2 libraries that affiliated with SL Cluster 3 were similar to a clone previously detected in Lake Magadi, which was suggested to be a deeply branching group of the *Euryarchaeota* (Grant *et al.* 1999). The characteristics or roles of these organisms in this group are not known since their existence has only been detected by molecular methods. Another group of *Euryarchaeota* were detected in EN1 and XH2 when nanoarchaeal specific primers were used (Figure 40, p220). This lineage also appeared to be novel as it did not bear much resemblance to existing 16S rRNA sequences (< 90% identity). These sequences were associated with *Euryarchaeota* that produced methane, although it is not clear whether these organisms have the same characteristics since the clones had very low sequence similarities to known methanogens (83 – 87%).

Bacteria

As previously mentioned, the primary production from *Eukarya* supports the growth of *Bacteria*, mainly *Proteobacteria*. Indeed, this was the largest group detected amongst all six salt lakes, with the *Gammaproteobacteria* being the largest group detected by both molecular (Figure 37, p212) and cultivation methods (Figure 43, p243).

Many moderate halophiles have been assigned to the genera *Halomonas* and *Halovibrio*. One isolate cultivated from CG1 was related to *Halomonas pantelleriense* (98%); one clone from the CG1 library was related to *Halomonas campisalis* (97%), and isolates cultivated from SH1, XH2 and CG1 were related to *Halomonas variabilis* and *Halomonas venusta* (96% - 98%), and clones detected in libraries for BJ1, SH1, XH2 were also related to these bacteria (98%) (Table 16, p190, Table 21, p232; Figure 37, p212, Figure 43, p243); all are capable of growth at alkaline pH. *Halomonas pantelleriense*, *Halomonas venusta* and *Halomonas campisalis* have been shown to reduce nitrate, while *Halomonas campisalis* can reduce nitrite

(Mata *et al.* 2002). *Halomonas ventosae* (cultivated from BJ1 and SH1) is also able to reduce nitrate and nitrite (Martínez-Cánovas *et al.* 2004). Isolates from BJ1, EN1, SH1 and XH2 were related to *Halovibrio denitrificans* (98%) (Figure 43), and one clone from the EJ3 library (98%) was related to the same bacterium (Table 16). This is a facultative anaerobe that also reduces nitrate and uses short chain fatty acids as electron donors and carbon sources, producing nitrite and N_2O as intermediates (Sorokin *et al.* 2006).

A small group of clones were affiliated to the anoxygenic phototrophs *Ectothiorhodospiraceae*, detected in libraries for CG1 and SH1 (Table 16, Figure 37). They usually inhabit places that are anoxic, but are sufficiently illuminated, for example at sediment surfaces (Ollivier *et al.* 1994). *Alkalilimnicola halodurans*, detected in EN1, is within the *Ectothiorhodospiraceae*, but unlike other members of this family is not anoxygenic or phototrophic. It is an aerobe that preferentially metabolises volatile fatty acids and requires vitamins for growth (Yakimov *et al.* 2001).

Sequences of the *Deltaproteobacteria* were present in BJ1, CG1, EN1 and XH2 libraries (Figure 37). They were associated with the sulphate reducing bacteria (SRB), many of which have been previously isolated from salt lakes (see Introduction). Moreover, elemental sulphur was calculated to be approximately 0.5 M - 1 M in the lake waters, which were probably in the form of various sulphates, providing suitable substrate for such bacteria. All clones appear to be somewhat related to genera classified as slightly halophilic sulphate reducers (Ollivier *et al.* 1994). The final product of sulphate reduction is hydrogen sulphide, which is released into the environment (Ollivier *et al.* 1994), which may account for the smell observed at some of the lakes.

One clone from CG1 was related to *Deltaproteobacteria* clone MLMS-1 (92%) (Table 16), which was shown reduce arsenate to arsenite using sulphide as the electron donor (Hoeft *et al.* 2004). Another bacterial clone from CG1 was somewhat related (90%) to a bacterium of the phylum *Firmicutes* (*Halanaerobiaceae* bacterium SLAS-1) that was also shown to use arsenate as its electron acceptor, while utilising sulphide as its electron donor (Oremland *et al.* 2005). In addition, this isolate could only grow at pH 9.1 – 9.5, again demonstrating the alkaliphilic nature of the organisms residing in this lake. A source of sulphide is the SRB as discussed, thereby demonstrating interrelationships with the *Deltaproteobacteria* in the salt lakes.

A small group of *Alphaproteobacteria* were also detected in libraries for CG1, EN1, SH1 and XH2 (Table 16, Figure 37), which display a range of metabolisms. One clone from the SH1 library was related to a *Roseovarius* sp. (95%) that was shown to produce iodine (Fuse *et al.* 2003), while another clone from the CG1 library was related to a *Paracoccus* sp. (95%) that was shown to degrade methomyl (Xu unpublished). Furthermore, a clone related to *Rhodobacteriaceae* (98%) was detected in the XH2 library, which is a group that is metabolically diverse (Table 16).

Known anaerobic fermentative halophilic bacteria are members of the order *Halanaerobiales* and *Halobacteroidaceae*, within the *Firmicutes*. They were detected in all the salt lake libraries, except BJ1. These are not typical members of the water column, but usually reside in anoxic sediments (Mouné *et al.* 2000; Mouné *et al.* 2003). Their presence may be because the sediment was disturbed during sampling or that the water contained an oxic-anoxic gradient, though there is no data to support this. A clone sequence related to *Orenia salinaria* (88%) was detected in EN1 (Table 16), which ferments glucose to produce ethanol, formate, acetate, lactate, CO_2 and H_2 (Mouné *et al.* 2000). Clones from SH1 and CH2 libraries were related to *Halanaerobacter lacunarum* (96% - 98%) (Table 16), which ferments starch to form ethanol, acetate CO_2 and H_2. One isolate cultivated from EN1 was related to *Halanaerobium praevalens* (99%) (Table 16), which ferments pectin to form acetate, proprionate, butyrate, CO_2 and H_2 (Ollivier *et al.* 1994). The carcasses of brine shrimp are responsible for the deposition of chitin and other organic molecules in hypersaline environments (Liaw and Mah 1992). Since these were only observed at XH2, this may explain the presence of *Halanaerobacter chitinivorans*; a clone detected in the XH2 library showed 98% identity to this bacterium, which uses chitin for growth (Liaw and Mah 1992). Several clones from CG1 library were related to *Natronoanaerobium halophilum* (93% - 96%), which was previously isolated from Lake Magadi (Jones *et al.* 1998b). In addition, *Halanaerobium* (detected in CG1 library) have been detected in soda lake environments or have alkaliphilic members (Tsai *et al.* 1995; Jones *et al.* 1998b), further demonstrating the alkaline nature of CG1.

A number of cultivated isolates were affiliated with *Bacillus* sp., (*Halobacillus*, *Amphibacillus*, *Pontibacillus*, *Filobacillus*, and *Gracilibacillus*) (Table 21, Figure 44, p244). These aerobic members of the *Firmicutes* have been well documented in both salt lakes and

soda lakes, many of which have been isolated and characterised (Ventosa *et al.* 1998a). They can utilise a number of organic substrates; many alkaliphilic *Bacillus* sp. have been exploited for their production of extracellular enzymes (Horikoshi 1999).

Many clones affiliated with *Bacteroidetes* were also detected in the salt lakes (Figure 34). A large number of clones from the EJ3 library were related to *Salinibacter ruber* (99%) as demonstrated by its high clone frequency (Table 16). *Salinibacter ruber* is a motile rod that contains red pigment (Anton *et al.* 2002), which probably contributed to the red colouration of the water at EJ3 (Oren and Rodriguez-Valera 2001) that was not observed at any of the other sampling sites. A number of other clone sequences from BJ1, CG1, EN1 and XH2 also appeared in the same clade as *Salinibacter* that may represent novel *Salinibacter* species (Figure 35, p210).

A variety of other members of the CFB were detected in the libraries for the salt lakes (Figure 35), but their roles are not easy to elucidate. Members of the CFB are associated with the degradation of complex organic substrates (Reichenbach 1991) and so they too may be responsible for degrading organic matter from algae or submerged vegetation. An isolate that was cultivated from SH1 was related to *Salegentibacter* sp. (97%), part of the CFB group (Table 16). To date, only one other *Salegentibacter* has been isolated from a hypersaline environment, which was from a hypersaline lake in Antarctica (McCammon and Bowman 2000); others have been isolated from a sea cucumber and a sea urchin (Nedashkovskaya *et al.* 2004; Nedashkovskaya *et al.* 2005)

A few *Actinobacteria* were detected by molecular methods (Figure 34), but again their roles are not easy to elucidate due to low sequence similarity to known organisms. However, one cultivated bacterium from CG1 whose closest affiliation was with the *Actinobacteria*, closely resembled a bacterium isolated from Crater Lake (a soda lake from the African Rift Valley), making it a valid member of the soda lake community.

Aside from the major phylogenetic groups described, which were mostly specialised halophilic bacteria, the bacterial community contained members of other divisions making this community far more complex than the archaeal community.

BJ1 contained a clone related to the *Cyanobacterium*, *Synechococcus* (< 90%). Cyanobacteria are oxygenic photosynthetic bacteria, however, they play a minor role in primary production where salinity is above 20% (w/v) (Grant *et al.* 1998). Perhaps during the rainy months when the lake is diluted, these organisms become a major component of the microbial community, but this was not observed at the time of sampling.

Clones from the salt lake libraries were related to clones that were initially detected in cold climates such as glacial ice and Arctic seawater, for example, Arctic seawater bacterium Bsw20461 (Table 16). This again may indicate that some of the bacteria have a broad temperature range for growth to give them a selective advantage in the colder months.

Several clones from the BJ1, CG1, EN1 and SH1 libraries were related to marine bacteria (Table 16). Furthermore, one isolate from SH1 was related to *Salegentibacter* sp. 18III/A01/068 (98%), which is a marine bacterium (Figure 45, p245). The presence of marine bacteria in these environments is intriguing since these salt lakes were derived from leachate of dissolved volcanic rocks and are not marine in origin. Marine bacteria have been shown to be halotolerant, resisting salinities up to 27% (w/v) (Kaye and Baross 2000). Since some bacteria are susceptible to wind dispersal (Echigo *et al.* 2005), it is possible that these bacteria have been blown in from the sea, and their halotolerance has allowed them to survive in the salt lakes. An isolate was cultivated from XH2 that was closely related (99%) to *Idiomarina loihiensis* (Figure 43, p243). Originally isolated from the periphery of a deep sea hydrothermal vent, it has been shown to have a growth temperature range of 4°C to 46°C and a salinity range of 0.5% to 20% (w/v) NaCl (Hou *et al.* 2004). This would allow this bacterium to a very good survivor of both temperature and salinity fluctuations in the salt lake. It is believed that it derives its energy from amino acids rather than sugar (Hou *et al.* 2004). One possible source for proteinaceous particles is the brine shrimp and organic molecules from decomposing halophilic microorganisms as previously discussed (Ollivier *et al.* 1994).

Sequences previously identified from thermal sites were detected. For example, a clone from BJ1 was related (94%) to an *Aquificales* clone (Table 16, Figure 35), which is very peculiar since *Aquificae* are thermophilic, growing at temperatures from 70°C or above (Reysenbach 2001c). Therefore, the possibility that this was a cross contaminant or a mix up from the molecular work on the hot spring environments must be considered. However, the finding of other clones related to 16S rRNA signatures that also originated around thermal sites does not

support this theory. A further two clones from EN1 and SH1 were found to cluster with the *Aquificae*; one clone related to a sequence originally detected at a hydrothermal vent (Reysenbach *et al.* 2000c) and the other clone related to a sequence detected on an active deep-sea vent chimney (Page *et al.* 2004). Moreover, one clone from CG1 was related to a clone detected in an intertidal hot spring (Hobel *et al.* 2005), one from XH2 was related to a clone detected in Mammoth Hot Springs in Yellowstone National Park (Bonheyo *et al.* unpublished). Yet another clone that was in a cluster on the periphery of the *Aquificae* (SL Cluster 4) was related to a clone detected in Obsidian Pool in Yellowstone National Park, USA, designated 'candidate division OP11.' However, specific primers designed against this OP11 group and were later found to be present in a range of environments including Carolina Bay sediment, Amazonian soil and Australian deep subsurface water (Hugenholtz *et al.* 1998). Owing to the low sequence similarity of these sequences to existing 16S rRNA sequences (84% - 94%), it may not be appropriate to construe the growth temperature ranges for these bacterial clones; the bacteria that were detected may not be thermophilic at all. Alternatively, perhaps like OP11, the sequences are from bacteria that are more ubiquitous in nature than previously recognised. Another possibility is that they may represent a novel lineage distantly related to the *Aquificae* that are adapted to lower temperatures. Recently, there has been evidence for the existence of mesophilic members of the order *Thermotogales* (Nesbø *et al.* 2006). However, it is not clear what role these bacteria contribute to the salt lake environment.

Other sequences that were detected in the libraries were relatives of bacteria found in more temperate conditions, such as forest soil and fresh water (Table 16). Specific bacterial sequences have been correlated with laboratory contaminants, particularly when the environment being studied has low biomass (Tanner *et al.* 1998). However, the high biomass observed at the lakes and the large number of clones picked suggests that these are not contaminants. It therefore suggests that these bacteria are physiologically diverse. For example, bacteria isolated from fresh water lakes on the Tibetan Plateau showed remarkable halotolerance, demonstrating an adaptation to a width of niches, which was predicted to be important in the colonisation of diverse habitats (Wu *et al.* 2006a). This demonstrates the ubiquitous nature of some organisms.

Sequences related to bacteria that were associated with the microbial flora characterised in swine effluent, animal waste treatments and activated sludge were also detected in the libraries (Table 16). It was observed that there was much animal contamination at Lake

Shangmatala and there were pigs roaming around Lake Bagaejinnor. Moreover, birds often feed on brine shrimp, which may leave faecal pellets. It is highly likely then that these clones represent contamination from animal waste material.

Comparison to Other Hypersaline Environments

Eukarya

Only certain specialised invertebrates survive salinities exceeding 10% (w/v). Brine shrimp are widely distributed, occurring in China (Xin *et al.* 1994), Africa (Kaiser *et al.* 2006) and America (Humayoun *et al.* 2003). *Artemia* species often occur in shallow pools and lakes at neutral pH that are subject to daily fluctuations in temperature, salinity and oxygen. However, *Artemia monica* are specially adapted to the deep lake water and relatively constant conditions by producing cysts that are buoyant and resistant to dehydration (Lenz 1982) and are resistant to high pH (Bowden *et al.* 1980). It is possible then that the brine shrimp observed in Inner Mongolia are also a different species of the *Artemia* genus.

Dunaliella is a common halophilic eukaryote with a wide distribution. Various species have been detected in a range of environments including hypersaline lakes in Vestfold Hills, Antarctica (Bowman *et al.* 2000b), the Dead Sea (Oren and Shilo 1982) and coastal solar salterns (Benlloch *et al.* 2002).

Archaea

The vast majority of archaeal clones belonged to the order *Halobacteriales*, consistent with many previous studies on hypersaline environments. It is now apparent that many neutral hypersaline environments where NaCl approaches saturation harbour climax populations of almost identical haloarchaea belonging to *Haloruhrum* (Grant *et al.* 1998) and *Haloquadratum walsbyi*. However, the predominating group for the Inner Mongolian lakes was *Haloruhrum*; *Haloquadratum walsbyi* was not detected often in the clone libraries. Despite this, other saline environments have also been found to be dominated solely by *Haloruhrum* sp., for example, the archaeal community in crystallisers of an Adriatic solar saltern (Pašić *et al.* 2005) and the archaeal community in an Australian crystalliser pond in Corio Bay, Victoria (Burns *et al.* 2004a).

There are currently eight genera that include haloalkaliphilic representatives: *Haloruhrum*, *Natrialba*, *Natronobacterium*, *Natronococcus*, *Natronomonas*, *Natronoruhrum*,

Natronolimnobius and *Halalkalicoccus*. Surprisingly, only haloalkaliphilic representatives of *Halorubrum* and *Natronomonas* were detected in CG1, but none from the other genera. However, the haloarchaeal diversity of soda lakes is not restricted to these genera. Clones affiliated to *Halobaculum* were detected in CG1 that were related to clones from alkaline Lake Zabuye (Fan *et al.* unpublished). Several clones related to *Haloferax* were detected from soda lakes of the Kenyan-Tanzanian Rift Valley (Rees *et al.* 2004). *Haloarcula* were also detected in the hypersaline alkaline 'Eight Mile Flat' in Nevada, USA (Ochsenreiter *et al.* 2002). This suggests that alkaliphilic organisms are more diverse than the previously recognised genera.

A dominant fraction of the microbial community in salterns has been found to be relatives of the recently cultivated *Haloquadratum walsbyi* (see Introduction). Therefore, it was surprising that they were not detected in the Inner Mongolian salt lakes, even when using *Haloquadratum walsbyi* specific primers, particularly in the salterns at Lake Ejinnor. This may suggest that the salinity at EJ3 was too low for growth, since they require 23% - 30% total salinity (Burns *et al.* 2004a). However, it was noted that SHOW grow well under low nutrient conditions (Bolhuis *et al.* 2004) and therefore it is more likely that the brine at EJ3 was nutrient rich, though there is no data to confirm this. This would allow faster growing haloarchaea to dominate. This appeared to be the case for Adriatic salterns, where no SHOW group were detected despite the salinity approaching saturation. This was also explained by the nutrient-enriched water source as well as the milder weather conditions (Pašić *et al.* 2005).

The absence of *Crenarchaeota* and documented methanogens in the salt lakes can be attributable to the high salinity. Walsh *et al* (2005) found that as soil salinity increased, the number of crenarchaeotes and methanogens decreased; they were only present in soils between 7 – 13% (w/v) NaCl (Walsh *et al.* 2005). This pattern was also observed in the shift in salinity across the water-sediment interface at Lake Chaka, China (Jiang *et al.* 2006).

SL Cluster 3 was a group that formed its own lineage on the periphery of the order *Halobacteriales*. It was suggested that this group was a deeply branching group of the *Euryarchaeota* (Grant *et al.* 1999). My phylogenetic analysis has shown that this group also exists in Lake Zabuye, a soda lake on the Tibetan plateau (Fan *et al.* unpublished). Therefore, this group is not an anomaly, but a group existing in two hypersaline salterns and two hypersaline salt lakes in Inner Mongolia (EJ3, EN1, SH1 and XH2) (Table 16, Figure 33) as

well as two hypersaline soda lakes (Lake Zabuye, Tibet and Lake Magadi, Kenya). This work also confirms that this group is not restricted to alkaline brines. Phylogenetic analysis has shown that they are distinct from other known *Euryarchaeota* (Figure 33, p188). Another group of *Euryarchaeota* detected when nanoarchaeal specific primers were used, that also appeared to form a distinct lineage where the sequences did not resemble existing 16S rRNA sequences. This group has also been observed in saline soils in British Columbia, Canada, where NaCl was only 7% (w/v) and the group was designated 'Uncultured *Euryarchaeota* group 2 (UEG-2)' (Walsh *et al.* 2005). These findings suggest that adaptation to high salt environments is not restricted to the *Halobacteriales*, but is in fact more widespread within the domain *Archaea*.

Bacteria in Hypersaline Lakes

Most of the *Bacteria* detected in the salt lake libraries for BJ1, SH1 and XH2, are valid members of salt lakes. The dominant groups from the 16S rRNA libraries in these lakes were the *Proteobacteria* (particularly the *Gammaproteobacteria*), the *Firmicutes* and the *Bacteroidetes*. Many isolates that were cultivated were related to halomonads and *Bacillus* species (Figure 43, Figure 44).

Similar studies on neutral hypersaline inland lakes have shown similar trends. A study on 16 lakes on the Tibetan Plateau that ranged in salinity from freshwater (0.02%) to hypersaline (22.3%) revealed that the hypersaline lakes were dominated by *Gammaproteobacteria*, as demonstrated by DGGE (Wu *et al.* 2006a) though it was not shown which genera they belonged to. Similarly a significant proportion of the bacterial diversity in Lake Chaka, also a neutral athalassohaline lake in China, belonged to the *Gammaproteobacteria*, though this was not the predominating group (see later) (Jiang *et al.* 2006). Halomonads often constitute the major bacterial group in less concentrated hypersaline environments. Various halomonads have been cultivated from the Great Salt Plains of Oklahoma, an environment with varying salinity and pH, which have subsequently been characterised (Caton *et al.* 2004).

Low levels of dissolved oxygen and a number of sources of organic matter encourage the growth of anaerobic fermentative organisms, which was observed in the Inner Mongolian hypersaline lakes. *Firmicutes* relating to the *Halanaerobiales* have been found in a variety of salt lakes and salterns (Zeikus *et al.* 1983; Oren *et al.* 1984; Zhilina *et al.* 1991; Ollivier *et al.* 1994; Cayol *et al.* 1995; Grant 2004) as well as the Red Sea (Eder *et al.* 2001) and Lake Ekho

in the Vestfold Hills, Antarctica (Bowman *et al.* 2000a). However, this is in contrast to the bacterial community observed at Lake Chaka, where only 7% of clones were affiliated with the *Firmicutes* (Jiang *et al.* 2006). Likewise, no *Firmicutes* were detected in hypersaline lakes on the Tibetan Plateau (Wu *et al.* 2006a) and none was detected in the hypersaline lakes of the Atacama desert, Chile (Demergasso *et al.* 2004).

One of the most abundant groups as the *Bacteroidetes*, which is consistent with the community observed in the water column at Lake Chaka, where this was the predominant group, many relating to *Salinibacter* (Jiang *et al.* 2006). It was also observed as the most dominant group in a range of hypersaline environments in the Atacama Desert, Chile (Demergasso *et al.* 2004).

Bacteria in Salterns

The sampling points EJ3 and EN1 were from hypersaline salterns and so these two sites will be considered separately. Environments in salterns are generally considered to be the most extreme with respect to NaCl concentrations, which often approach saturation. Many studies have been carried out on salterns that contain thalassohaline waters (Caumette *et al.* 1994; Antón *et al.* 2000; Benlloch *et al.* 2002; Burns *et al.* 2004a; Pašić *et al.* 2005; Sørensen *et al.* 2005; Maturrano *et al.* 2006), but to my knowledge, few have been carried out on athalassohaline salterns.

It is widely accepted that in thalassohaline salterns, a climax population of archaea from the family *Halobacteriales* (Grant 2004) and a smaller proportion of halophilic bacteria (Antón *et al.* 2000) dominate. Despite the fact that the Inner Mongolian lakes are athalassohaline, the community at EJ3 reflects this view. A large number of clones related to *Salinibacter ruber* was detected in the saltern at EJ3 (Table 16). *Salinibacter ruber* is amongst the most halophilic organisms within the domain *Bacteria* and has been demonstrated as a predominant member in various saltern communities in Spain (Antón *et al.* 2000; Benlloch *et al.* 2002; Rosselló-Mora *et al.* 2003). They have also been detected in Lake Chaka, an athalassohaline lake in China (Jiang *et al.* 2006). The hypothesis that *Salinibacter ruber* is geographically restricted to coastal solar salterns or that it is unable to grow in athalassohaline environments (Demergasso *et al.* 2004) must therefore be rejected.

However, despite this generalised view, there are salterns that do not contain such a community. For example, the Maras salterns in the Peruvian Andes, were found to be dominated by bacteria related to '*Pseudomonas halophila*' DSM 3050; *Salinibacter* only represented a minor component (Maturrano *et al.* 2006). Likewise, the saltern at EN1 is atypical of saltern microbiota. The bacterial community was not dominated by *Salinibacter ruber*, or indeed by any particular bacterium, but contained a variety of bacteria from several genera. This confirms the prediction by Oren and Rodriguez-Valera (2001) that salterns at different geographical locations may harbour microbial communities with different structures (Oren and Rodriguez-Valera 2001). In fact, a study on two geographically distinct salterns (one from Newark, California, USA and the other from Eilat, Israel) using whole community pigments and lipids analysis were found to contain very different communities. The more diverse and complex lipid and pigment patterns implied a more complex microbial community in the USA plant, which was predicted to be the result of milder weather and a more nutrient rich water source (Litchfield *et al.* 2000). Additional studies on the same salterns comparing the metabolic activities also suggested that salterns in different geographic locations may have communities with different microbial communities despite the similarities in salt content and processing for salt production (Litchfield *et al.* 2001).

Bacteria in Soda Lakes

Lake Chagannor was the only less hypersaline soda lake that was sampled. Most *Bacteria* were valid members of the soda lake community. The majority of the bacterial community here was attributed to the haloalkaliphilic representatives of the *Proteobacteria* and the *Firmicutes* belonging to the *Halanaerobiaceae* and *Natronoanaerobiaceae*. Several isolates that were cultivated from this lake were affiliated with *Halomonas* and *Bacillus* species that show tolerance to high pH (Figure 43, Figure 44).

Cyanobacterial blooms only occur in soda lakes after rainfall allowing dilution of the surface waters (Jones *et al.* 1998b). However, no cyanobacteria were detected at CG1. The anoxygenic phototrophs of the family *Ectothiorhodospiraceae* have most commonly been isolated from alkaline brines in athalassohaline environments, such as desert lakes (Ollivier *et al.* 1994) and contribute to primary production. They have been isolated from the alkaline Lake Magadi (Jones *et al.* 1998b) and recently have been detected in the inland waters of remote Hawaiian islands (Donachie *et al.* unpublished). A relative of the

Ectothiorhodospiraceae was detected in the CG1 16S rRNA library, possibly contributing to primary production.

Two clones from the CG1 library were related to bacteria that reduce arsenate, implying that Lake Chagannor contains arsenate and allows this metabolic process to occur. This has been observed elsewhere. Several soda lakes in western USA that are also situated in arid volcanic terrain contain high concentrations of arsenic oxyanions (Hoeft *et al.* 2004). Searles Lake and Mono Lake are examples of alkaline lakes that are rich in arsenic. Isolates cultivated from these lakes were demonstrated to metabolise arsenate (Oremland *et al.* 2002; Oremland *et al.* 2005).

Mono Lake and lakes of the Kenyan-Tanzanian Rift Valley are well studied examples of hypersaline soda lakes. Molecular studies in Mono Lake showed that the majority of the sequences from the oxygenated surface waters affiliated with the *Actinobacteria*, which is in complete contrast to CG1, where no *Actinobacteria* were observed at all. In addition, the anoxic waters were dominated by *Bacillus* and *Clostridium* (Humayoun *et al.* 2003), whereas the only aerobic bacilli were cultivated and the only anoxygenic species in CG1 were *Halanaerobiaceae*, *Natronoanaerobiaceae* and *Ectothiorhodospiraceae*. Members of the *Gammaproteobacteria* belonging to the genus *Thioalkalivibrio* were detected in both Mono Lake and in the Rift Valley (Humayoun *et al.* 2003; Rees *et al.* 2004), but were not detected at CG1. *Thioalkalivibrio* are typical inhabitants of soda lakes and are responsible for sulphur oxidation in aerobic alkaline brines. They have been isolated from soda lakes in south east Siberia, Kenya and Egypt (Sorokin *et al.* 2002).

However, a study on the Inner Mongolian Baer soda lake (with a pH of 10) showed that the sediment samples also had an abundance of *Proteobacteria*, with most of the abundance with the *Gammaproteobacteria* division (Ma *et al.* 2004b), which better reflects the bacterial community at CG1. The Wadi An Natrun is a depression in the Sahara desert containing seven alkaline hypersaline lakes (Mesbah *et al.* In Press). Similarly, there was a relatively high abundance of *Gammaproteobacteria* and *Firmicutes* in the water column, which is also consistent with this study.

The differences noted in the microbial community of the soda lakes discussed may be a reflection of salinity. The microbial community at Lake Chagannor is more similar to the

microbial communities observed in Baer Soda Lake and lakes in the Wadi An Natrun, where salinity was less hypersaline, but are dissimilar to microbial communities in Mono Lake and the lakes in the Rift Valley, where conditions are hypersaline.

DISCUSSION: CHARACTERISATION OF MICROBIAL DIVERSITY BY CULTIVATION METHODS

Cultivated *Archaea*

Archaea can be slow growing, some requiring 21 days for the appearance of red colonies on a plate (Grant *et al.* 2001). Indeed, the *Archaea* did not appear on the plates until week three (Table 21, p232). This was also observed when *Archaea* were cultivated from Australian crystalliser ponds (Burns *et al.* 2004a). A study on crystalliser ponds showed that most of the haloarchaeal groups were cultivable (Burns *et al.* 2004a). The same result was obtained when investigating the microbiota of a solar saltern in La Palma, Spain (Ochsenreiter *et al.* 2002). This is in contrast to my results, where only two genera were cultivated, despite evidence from the 16S rRNA libraries indicating that many more were present (Figure 46, p247). Extended incubation time has shown to be significant in the isolation of *Natronomonas* groups (Burns *et al.* 2004a); though no more archaeal colonies were observed beyond week five, hence all the *Archaea* that could be isolated had already been done so. The lack of archaeal isolates could not be because the media did not contain enough salt since this would have meant that no *Archaea* could be isolated at all, but *Halorubrum* and *Haloarcula* were isolated, which require 1.5 M – 5.2 M and 2 M – 5.2 M NaCl for growth, respectively (Grant *et al.* 2001).

A more likely explanation is that many of the *Archaea* are not cultivable using the methods in this study. A study by Ochsenreiter *et al.* (2002) showed that when media containing the major ions from natural habitats was used to isolate microorganisms from hypersaline environments, few isolates were obtained. However, isolates were obtained when a different approach involving enrichment cultures was undertaken (Ochsenreiter *et al.* 2002). The media used in this study was similarly designed against the results from optical emission spectrometry. Therefore, other cultivation methods may have been more successful in obtaining more diverse archaeal isolates. A further study by Burns *et al.* (2004a) demonstrated that utilising a wide range of carbon and energy sources allowed cultivation of a number of microorganisms that showed similarity to those detected by molecular methods (Burns *et al.* 2004a). Therefore using a carbon-rich media may have been more successful in cultivating more *Archaea* than an amino acid-rich media as was used in this study.

Cultivated *Bacteria*

The *Bacteria* were faster growing than the *Archaea*; some bacterial species even appearing within the first week. There were many more bacterial isolates than archaeal, though the diversity of isolates was again quite low, despite evidence from the molecular characterisation of a more diverse community (Figure 46). The fact that both the cultivated *Archaea* and *Bacteria* showed low diversity suggests that the cultivation methods employed were highly biased. The majority of bacterial isolates were *Gammaproteobacteria*, particularly halomonads, and *Firmicutes*, particularly *Bacillus* sp. (Figure 43, Figure 44, Figure 45). All isolates were related to bacteria that are known to be halophilic or show tolerance to high salinities and so they are all valid members of the microbiota of saline environments.

Results from the culturing work are consistent with other cultivation studies. Isolates that were retrieved from Lake Chaka were also predominantly from the *Gammaproteobacteria* that were related to *Halomonas* species, with a few related to *Bacillus* species (Jiang *et al.* 2006). The fast-growing isolates from enrichments of samples from the Great Salt Plain of Oklahoma were predominantly from the *Gammaproteobacteria* (Caton *et al.* 2004). Similarly, another study on the same environment showed that *Halomonas* species dominated (Miller *et al.* 2004). Relatives of *Idiomarina* were also cultivated from the Great Salt Plains, demonstrating that they do not only occur in marine environments. Moreover, cultivation studies on a solar saltern in Chungnam Province, Korea showed that the *Gammaproteobacteria* and *Firmicutes* predominated, and were also shown to be moderate halophiles, some related to *Halomonadaceae*, *Alteromonadaceae*, *Idiomarinaceae* and *Bacillaceae* (Yeon *et al.* 2005).

Although no *Bacillus* species were detected in the libraries, they were one of the dominant groups that were cultivated. *Bacillus* sp. are one of the most commonly found aerobic bacteria in salt lakes. *Bacillus*, *Halobacillus* and *Marinococcus* species are moderately halophilic bacteria (Ventosa *et al.* 1998a), which have been cultivated from a number of saline environments (Spring *et al.* 1996; Jones *et al.* 1998b; Lim *et al.* 2005a; Lim *et al.* 2005b; Liu *et al.* 2005; Vargas *et al.* 2005).

The observation that a large proportion of *Gammaproteobacteria* was cultivated can be attributed to the nutrient-rich approach that was undertaken. A study was done by Pinhassi and Berman (2003) where bacteria were cultivated from samples from aquatic environments

using dilution culture and nutrient addition. It was found that *Gammaproteobacteria* prefer elevated concentrations of nutrients and organic carbon, but in the presence of low nutrient concentrations, they are selectively outcompeted by *Alphaproteobacteria* (Pinhassi and Berman 2003). Perhaps a low nutrient approach would have allowed the cultivation of other *Bacteria.*

RESULTS: MICROBIAL BIOGEOGRAPHY AND BIODIVERSITY

Introduction

In this chapter, results from the molecular analysis were used to analyse the microbial diversity in the salt lakes and hot springs. Clear differences between the samples were determined by Principal Component Analysis (PCA). Operational Taxonomic Units (OTUs) were defined in order to investigate which factors significantly influence biotic similarity. These factors include geographic distance (and therefore history) and environmental factors such as pH and ionic concentrations. Subsequently, various diversity indices used in ecology, were used to measure the richness, diversity and relative abundance of microorganisms in each environment. In addition, microbial diversity was defined at the genetic level using nucleotide diversity, lineage-per-time plots, degree of differentiation (Fst) and parsimony test (P test).

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) was conducted for both the archaeal and bacterial communities in the salt lakes (from BJ1, CG1, EJ3, EN1, XH2 and XH3) and the hot springs (from TC8 and LP4) using UniFrac (Lozupone and Knight 2005). The results for an additional sampling point, XH3 were provided by Wang (2006), which was a small pool found alongside the unnamed salt lake near Xilin Hot (Wang 2006).

Results from PCA are illustrated in Figure 47 and Figure 48. The program calculates the biotic similarity between the samples based on sequence data and then plots this similarity in space. The three axes represent the first three Principal Components (PCs), where the first PC describes the most variation possible, and the subsequent components explain the remaining variation not explained by the previous component or components. Since this data is influenced by a number of factors, only the first three PCs are considered. The closer the samples are, the more similar their biotic composition.

For the *Archaea*, approximately 61.2% of the variability is represented in the first three Principal Components (PCs). PC1 explains 32.9% of the variability, PC2 explains 14.8% and PC3 explains 13.5%. Generally, the hot spring samples are significantly separate from the salt lake samples at all three PCs. LP4 and TC8 are very similar, and are loaded heavily on PC1, representing key differences between the hot spring and salt lake samples. However, when

considering the salt lake samples only, CG1 appears to be diametrically opposite BJ1, EJ3, EN1, SH1, XH2 and XH3 at all three PCs.

For the *Bacteria*, approximately 51.5% of the variability is represented in the first three PCs, suggesting that there are more independent factors causing variations between the samples for the *Bacteria* than the *Archaea*. PC1 explains 23.5% of the variability, PC2 explains 14.6% and PC3 explains 13.4%. Like the *Archaea*, the hot springs are separate from the salt lakes on PC1, which represents key differences. The same is true at PC3, but PC2 places LP4 and TC8 diametrically opposite each other. When considering the salt lake samples only, CG1 again seems to be diametrically opposite the other salt lakes since all salt lakes load heavily on PC3, except CG1.

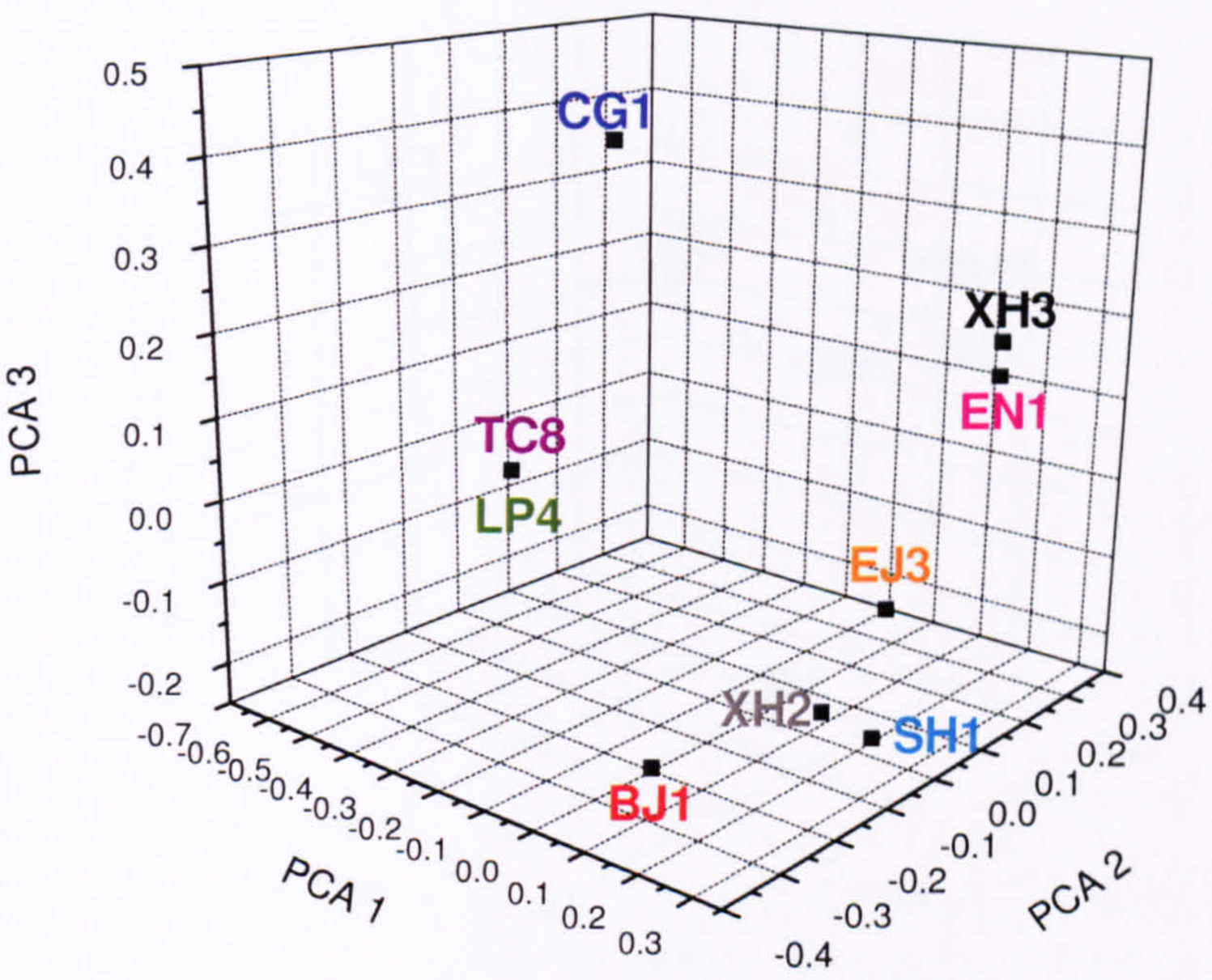


Figure 47 Principal Component Analysis for the *Archaea*.

This plot shows the separation of the samples in space according to three principal components (PCs), accounting for 61.2% of the variability. The data point for TC8 and LP4 is identical.

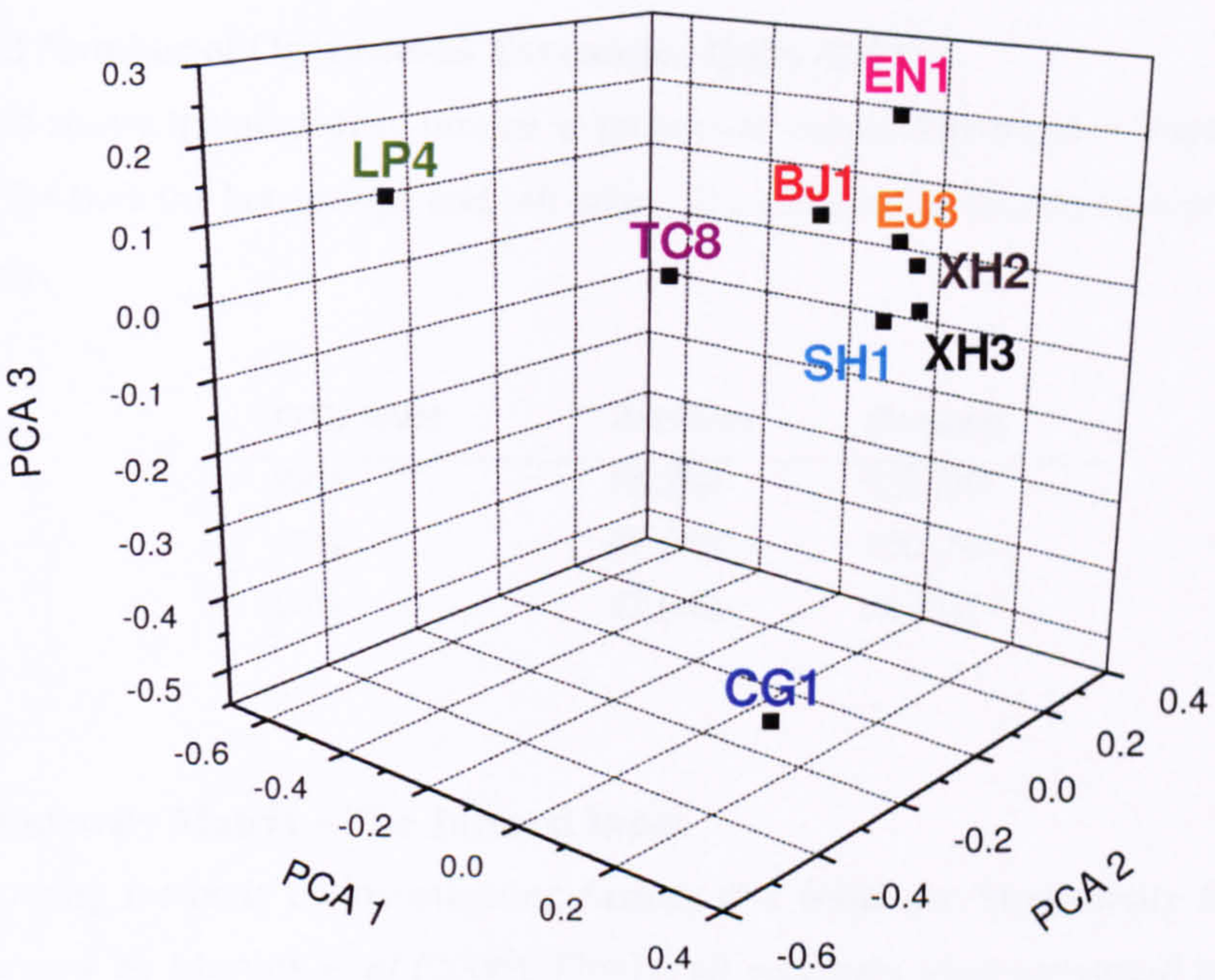


Figure 48 Principal Component Analysis for the *Bacteria*.

This plot shows the separation of the samples in space according to three principal components (PCs), accounting for 51.5% of the variability.

Operational Taxonomic Units (OTUs)

16S rRNA gene sequences of organisms were used to determine biotic similarity between the the salt lakes and the hot springs. The sequences were aligned using MEGA 3.1 (Kumar *et al.* 2004), and the output file from this was used to define Operational Taxonomic Units (OTUs) using DOTUR (Schloss and Handelsman 2005). In this study, three of the commonly used OTU definitions were used (95%, 97% and 99%), which is equivalent to comparing different taxonomic resolutions: at the genus, species and sub-species level (Horner-Devine *et al.* 2003). Table 22 shows the number of unique sequences at each of these OTU definitions.

Table 22 Number of Operational Taxonomic Units (OTUs).

This table shows the number of unique sequences at various Operational Taxonomic Units (OTUs) for both the hot springs and salt lakes. The numbers in brackets are results for the salt lakes only.

OTU level	*Archaea*	*Bacteria*
99%	88 (69)	130 (99)
97%	67 (49)	100 (76)
95%	57 (40)	86 (63)

Biotic Similarity Matrix – The Jaccard Index

The following methods of investigating factors that influence biodiversity have been taken from a review by Martiny *et al* (2006). Firstly, all variables were converted into resemblance matrices.

The Jaccard Index was calculated using the program EstimateS version 7.5 (Colwell 2005) for data generated at each OTU level. This index measures incidence data; it shows only the presence or absence (occurrence) of species in each environment. Values for the Jaccard Index were then put into a biotic similarity matrix. Table 23 and Table 24 show the different matrices constructed at each OTU for *Archaea* and *Bacteria*. A value of zero indicates that two environments do not have any organisms in common at that particular OTU and a value of 1 indicates that two environments harbour identical microbial communities. Hence, it is clear that none of the environments contain identical archaeal or bacterial communities, as a value of 1 was never obtained. The null hypothesis ('microorganisms are distributed randomly over space:' see Introduction) can therefore be rejected for this particular area.

Clearly a much larger data set is required to test this over greater geographical distances (see Discussion).

None of the archaeal sequences from the hot springs were similar to those of the salt lakes at any OTU level. However, one of the bacterial sequences from the salt lake BJ1 was 99% similar to a sequence recovered from the hot spring TC8 (data not shown).

Table 23 Biotic Similarity Matrices for the *Archaea*.

These matrices show biotic similarity determined by the Jaccard Index for the *Archaea* at three definitions of Operational Taxonomic Units (OTUs) for the salt lakes only. A value of 1 indicates identical microbial communities.

99%

	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	1						
CG1	0	1					
EJ3	0.071	0	1				
EN1	0.043	0	0	1			
SH1	0.214	0	0	0	1		
XH2	0.166	0	0	0.035	0.1	1	
XH3	0.043	0	0	0.066	0.041	0.035	1

97%

	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	1						
CG1	0	1					
EJ3	0.071	0	1				
EN1	0.095	0.043	0	1			
SH1	0.307	0.058	0.066	0.2	1		
XH2	0.285	0.055	0	0.136	0.266	1	
XH3	0.043	0.086	0	0.24	0.136	0.130	1

95%

	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	1						
CG1	0.071	1					
EJ3	0.076	0	1				
EN1	0.176	0.105	0	1			
SH1	0.454	0.133	0.142	0.222	1		
XH2	0.6	0.133	0.066	0.157	0.384	1	
XH3	0.176	0.166	0	0.368	0.222	0.222	1

Table 24 Biotic Similarity Matrices for the *Bacteria*.

These matrices show biotic similarity determined by the Jaccard Index for the *Bacteria* at three definitions of Operational Taxonomic Units (OTUs) for the salt lakes only. A value of 1 indicates identical microbial communities.

99%

	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	1						
CG1	0	1					
EJ3	0	0	1				
EN1	0.04	0	0.041	1			
SH1	0.038	0	0	0	1		
XH2	0	0	0.076	0.057	0	1	
XH3	0	0	0	0	0	0.088	1

97%

	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	1						
CG1	0.038	1					
EJ3	0	0	1				
EN1	0.086	0.030	0.043	1			
SH1	0.041	0.062	0.043	0.103	1		
XH2	0.076	0.057	0.080	0.129	0.129	1	
XH3	0.086	0.097	0	0.066	0.103	0.166	1

95%

	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	1						
CG1	0.045	1					
EJ3	0	0	1				
EN1	0.095	0.071	0.05	1			
SH1	0.045	0.071	0.05	0.111	1		
XH2	0.04	0.064	0.142	0.1	0.222	1	
XH3	0.1	0.16	0	0.115	0.208	0.230	1

Geographic-Distance Matrix

This matrix shows the distances between each of the environments in this study. This was calculated using the GSP coordinates measured during the expeditions (see Methods).

Environmental Similarity Matrices

To construct environmental similarity matrices, all raw values needed to be standardised to make them comparable since this accommodates for different units of different variables. Therefore, the Standard Normal Deviate Equivalents (SNDE) values were calculated for temperature, pH, and concentrations of Na, Mg, K, Cl, S, CO_3^{2-} and HCO_3^- (see Methods). In order to construct a similarity matrix with this new standardised data, one minus the Euclidean distance between the SNDE values of two lakes was calculated. Therefore, a value of 1 indicates that that environmental factor is identical for those two lakes. Table 27 to Table 35 show several environmental similarity matrices.

Table 25 Geographic-distance Matrix.

This matrix shows the distances between the sampling points at the salt lakes in km.

km	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	360.5	X					
EJ3	12.5	326.2	X				
EN1	395.7	88.4	395.2	X			
SH1	284.9	89.8	288.8	165.1	X		
XH2	158.6	227.2	165.4	286.6	140.2	X	
XH3	157.6	229.1	164.5	297.8	142	0.147	X

Table 26 Standard Normal Deviate Equivalents (SNDE).

This table shows the Standard Normal Deviate Equivalents (SNDE) for various environmental factors measured at each of the sampling points at the salt lakes.

	Temp (ºC)	pH	Na (M)	Mg (M)	K (M)	Cl (M)	S (M)	CO_3 (mM)	HCO_3 (mM)
BJ1	-0.2019	-0.2156	0.9916	-0.2338	-0.4376	0.3963	1.3289	-0.4101	-0.3936
CG1	-1.1629	107964	-0.7728	-0.7016	-0.8492	-1.7373	-0.4377	2.2646	2.2669
EJ3	1.8048	-1.2216	-0.8236	2.0851	0.3277	0.2452	0.9700	-0.2786	-0.3752
EN1	-0.9367	-0.7186	0.1784	0.4498	-0.3133	0.8315	-0.2997	-0.3772	-0.4185
SH1	-0.1170	-0.2156	1.0352	-0.3545	2.0662	0.4447	0.6112	-0.3463	-0.3936
XH2	0.6743	-0.2156	0.8029	-0.5890	-0.0110	0.8739	-0.7137	-0.4206	-0.3445
XH3	-0.0606	0.7904	-1.4118	-0.6560	-0.7827	-1.0543	-1.4590	-0.4318	-0.3415

Table 27 Environmental Similarity Matrix for Temperature.

Temp (ºC) [a]	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	0.03903	X					
EJ3	-1.00673	-1.96770	X				
EN1	0.26514	0.77329	-1.74159	X			
SH1	0.91521	-0.04576	-0.92278	0.18035	X		
XH2	0.12382	-0.83715	-0.13055	-0.61104	0.20861	X	
XH3	0.85868	-0.10229	-0.86541	0.12382	0.94347	0.26514	X

Table 28 Environmental Similarity Matrix for pH.

pH [a]	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	-1.01201	X					
EJ3	-0.00601	-2.01802	X				
EN1	0.49699	-1.51502	0.497	X			
SH1	1	-1.01201	-0.00601	0.49699	X		
XH2	1	-1.01201	-0.00601	0.49699	1	X	
XH3	-0.006	-0.006	-1.01201	-0.50901	-0.006	-0.006	X

Table 29 Environmental Similarity Matrix for Sodium Concentration.

Na (M) [a]	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	-0.76443	X					
EJ3	-0.81526	0.94914	X				
EN1	0.186767	0.048806	0.002024	X			
SH1	0.956343	-0.80799	-0.85882	0.143201	X		
XH2	0.811214	-0.57564	-0.62647	0.375553	0.767648	X	
XH3	-1.40322	0.36103	0.41186	-0.59016	-1.44696	-1.21461	X

[a] These environmental matrices show 1 minus the Euclidean distance between the SNDE values of the two sampling points being compared.

Table 30 Environmental Similarity Matrix for Magnesium Concentration.

Mg (M) [a]	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	0.53219	X					
EJ3	-1.31894	-1.78675	X				
EN1	0.316381	-0.15143	-0.63532	X			
SH1	0.87936	0.65283	-1.43958	0.195741	X		
XH2	0.64479	0.8874	-1.67415	-0.03883	0.76543	X	
XH3	0.57777	0.95442	-1.74117	-0.10593	0.69841	0.93298	X

Table 31 Environmental Similarity Matrix for Potassium Concentration.

K (M) [a]	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	0.58842	X					
EJ3	0.234718	-0.17686	X				
EN1	0.81567	0.46409	0.359048	X			
SH1	-1.50378	-1.91536	-0.73850	-1.37945	X		
XH2	0.57341	0.16183	0.661308	0.69774	-1.07719	X	
XH3	0.65487	0.93355	-0.11041	0.53054	-1.84891	0.22828	X

Table 32 Environmental Similarity Matrix for Chlorine Concentration.

Cl (M) [a]	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	-1.13368	X					
EJ3	0.848889	-0.98257	X				
EN1	0.564802	-1.56888	0.413691	X			
SH1	0.951642	-1.18202	0.800534	0.613157	X		
XH2	0.522491	-1.61119	0.37138	0.957689	0.570846	X	
XH3	-0.45066	0.31698	-0.29955	-0.88586	-0.49902	-0.92817	X

[a] These environmental matrices show 1 minus the Euclidean distance between the SNDE values of the two sampling points being compared.

Table 33 Environmental Similarity Matrix for Sulphur Concentration.

S (M) [a]	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	-0.76658	X					
EJ3	0.641163	-0.40774	X				
EN1	-0.62857	0.86199	-0.24973	X			
SH1	0.282325	-0.04891	0.641162	0.089104	X		
XH2	-1.04261	0.72397	-0.68377	0.58596	-0.32493	X	
XH3	-1.78789	-0.02131	-1.42905	-0.15932	-1.07021	0.25472	X

Table 34 Environmental Similarity Matrix for Carbonate Concentration.

CO_3 (mM) [a]	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	-1.67172	x					
EJ3	0.868465	-1.54321	X				
EN1	0.967116	-1.64186	0.901349	X			
SH1	0.936206	-1.61095	0.932259	0.96909	X		
XH2	0.989478	-1.68527	0.857943	0.95654	0.925684	X	
XH3	0.978297	-1.69645	0.846762	0.945413	0.914503	0.988819	X

Table 35 Environmental Similarity Matrix for Bicarbonate Concentration.

HCO_3 (mM) [a]	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3
BJ1	X						
CG1	-1.66051	X					
EJ3	0.98159	-1.64210	X				
EN1	0.9751	-1.68541	0.95669	X			
SH1	1	-1.66051	0.98159	0.975	X		
XH2	0.95095	-1.61146	0.96936	0.92605	0.95095	X	
XH3	0.94794	-1.60845	0.96635	0.92304	0.94794	0.99699	X

[a] These environmental matrices show 1 minus the Euclidean distance between the SNDE values of the two sampling points being compared.

Factors Driving Microbial Community Composition – The Mantel Test

The Mantel tests were carried out using the zt program (Bonnet and Van der Peer 2002) on the data for the salt lakes only. This program calculates the simple Mantel test, which will show if biotic similarity correlates with environmental factors, and whether this correlation is statistically significant. The value of r is the correlation value; positive or negative values reflect the type of relationship between the two matrices, while p is the probability associated with r. Values of p are significant if it is less than 0.05; values greater than 0.05 indicates that the null hypothesis applies. (The null hypothesis states that distances in matrix A are independent of the distances in matrix B).

Table 36 shows the results of the simple Mantel test. The results show that for both the *Archaea* and *Bacteria*, geographic distance is not a significant factor in influencing the microbial community composition, since p values across all OTU levels are > 0.05. Hence the effect of history also does not affect microbial community composition.

For the *Archaea*, the factors that appear to be significant in driving differences in community composition at the genus level are temperature, pH and the concentration of Mg ions. Temperature appears to have a positive correlation (r = 0.558919, p = 0.027183), which suggests that biotic similarity between environments increases if the temperature of the environments are more similar. The situation is the same for the concentration of Mg (r = 0.568478, p = 0.034127) and for pH (r = 0.612896, p = 0.011905); as these factors become more similar, the biotic composition of the lakes also become more similar. However, pH is dependent on carbonate and bicarbonate, but these factors are not statistically significant (p = 0.153373 and p = 0.221825 respectively). At the species level, pH and concentration of Na ions are a significant factors, with positive correlations to biotic similarity (r = 0652055, p = 0.008333 for pH; r = 0.349779, p = 0.048016 for Na). Surprisingly, carbonate and bicarbonate still are not statistically significant. At the subspecies level, pH and concentration of Na ions are still a significant factors, with positive correlations to biotic similarity (r = 0.611123, p = 0.005952 for pH; r = 0.369830, p = 0.038095 for Na). However, at this level, carbonate and bicarbonate are also significant (r = 0.435762, p = 0.031151 for carbonate; r = 0.429136, p = 0.053373 for bicarbonate).

For the *Bacteria*, the factors that are significant in influencing differences in community composition at the genus level are temperature (r = 0.522541, p = 0.027381) and

concentration of Mg (r = 0.512640, p = 0.027381). Again, the positive value for r indicates a positive correlation of these factors with biotic similarity. At the species level, pH and concentration of Mg are significant factors, again with positive correlations (r = 0.531519, p = 0.046429 for pH; r = 0.607582, p = 0.030159 for Mg). At the subspecies level, concentration of bicarbonate appears to be the significant factor (r = 0.375262, p = 0.020040). However, the concentration of carbonate has a p value that is near to the cut off point (p = 0.062103). Together, it would suggest that pH would also be significant, but unexpectedly, its p value is not near the cut off point (p = 0.094048).

Table 36 The Simple Mantel Test.

This shows the results of the simple Mantel test for the *Archaea* and the *Bacteria* at three definitions of OTUs for the salt lakes only. r values represent the possible correlations, while the p values represent statistical significance. Correlations are significant if $p < 0.05$, which are in **red**.

Archaea

OTU	99%		97%		95%	
Factor	r	p	r	p	r	p
Distance	-0.060155	0.391270	-0.103315	0.325595	-0.114317	0.314484
Temp	0.402505	0.067460	0.488340	0.059127	**0.558919**	**0.027183**
pH	**0.611123**	**0.005952**	**0.652055**	**0.008333**	**0.612896**	**0.011905**
Na	**0.369830**	**0.038095**	**0.349779**	**0.048016**	0.266982	0.108333
Cl	0.422015	0.069841	0.381560	0.087103	0.275187	0.182341
Mg	0.349224	0.120635	0.510477	0.068056	**0.568478**	**0.034127**
S	-0.059673	0.389087	0.105037	0.350000	-0.051399	0.424405
K	-0.174892	0.312103	-0.320044	0.139683	-0.189825	0.383532
CO_3	**0.435762**	**0.031151**	0.435839	0.159524	0.361955	0.153373
HCO_3	**0.429136**	**0.053373**	0.422993	0.213889	0.347359	0.221825

Bacteria

OTU	99%		97%		95%	
Factor	r	p	r	p	r	p
Distance	0.014208	0.465278	-0.211539	0.171627	-0.300727	0.081944
Temp	0.021471	0.494643	0.468950	0.076984	**0.522541**	**0.027381**
pH	0.327979	0.094048	**0.531519**	**0.046429**	0.358587	0.138095
Na	-0.052422	0.410913	-0.165519	0.187500	-0.277499	0.083135
Cl	0.239571	0.171627	0.154325	0.318056	-0.002551	0.539881
Mg	-0.024314	0.466667	**0.607582**	**0.030159**	**0.512640**	**0.027381**
S	0.125110	0.308532	-0.008087	0.478175	-0.047244	0.406944
K	0.250513	0.173810	-0.022989	0.549008	-0.131126	0.422421
CO_3	0.372461	0.062303	0.317234	0.153770	0.217162	0.347421
HCO_3	**0.375262**	**0.020040**	0.306015	0.211905	0.214920	0.358333

Diversity Indices

The values for expected richness (Chao 1) for *Archaea* are generally higher for the hot spring TC8 (Table 37). The values for TC8 remains at 28.5 for all levels of OTU, while the Chao 1 value for LP4 at the subspecies level is similar (33), which decreases to 11 at the genus level. Values for the *Bacteria* are generally higher for LP4 than for TC8 (41.75 – 52.9 for LP4 and 29 – 43 for TC8) (Table 38). Between the two hot springs, TC8 has the greatest archaeal diversity, with a Shannon-Weaver Index value of 2.67 for all levels of OTU, while LP4 has values between 1.57 and 1.88 (Table 37). This is supported by the Simpson Index. The evenness is also higher for TC8 (E = 0.96; 1 indicates completely even distribution). However, these values are still an underestimate for TC8, since the sampling effort for the archaeal 16S rRNA library did not reach saturation (Figure 25, p18). The opposite pattern is observed for the bacterial diversity; LP4 has the highest Shannon-Weaver Index (3.02 – 3.14), the highest Simpson Index (37.4 – 56.1) and the highest evenness (E = 0.95 – 0.96) (Table 38).

Looking at the Chao1 values for the salt lakes (see Table 37 and Table 38), SH1 has the highest archaeal richness values at all levels of OTU (26 to 55.5), while CG1 has the lowest (10.25 to 16.17). Again, SH1 has the highest values for bacterial richness at all levels of OTU (50 to 72). However, EJ3 appears to have the lowest bacterial richness (7 to 8).

The archaeal diversity generally appears to be the highest for EJ3 over all levels of OTUs, showing a Shannon-Weaver Index of 2.3 (Table 37). This is supported by the value for the Simpson Index (12.5). This diversity also appears to be the most evenly distributed amongst the OTUs detected, with an evenness value of 0.93. There is a slight discrepancy at the subspecies level, where EN1 appears to have a slightly higher diversity indicated by Shannon-Weaver Index (2.39), but this is not supported by the Simpson Index, which stills shows EJ3 to have the highest diversity. BJ1 had the lowest archaeal diversity over all levels of OTU as shown by both the Shannon-Weaver and Simpson Index (1.39 – 1.59 and 2.88 – 3.37, respectively). This diversity is also the least even amongst the OTUs detected, indicating that there is one dominating species present in that environment. However, these values may be an underestimate as sampling of the BJ1 archaeal 16S rRNA library did not go to completion (Figure 25).

The bacterial diversity is the greatest for CG1 for all levels of OTU, with a Shannon-Weaver Index of 2.48 to 2.77 (Table 38). This is confirmed by the Simpson Index values (15.5 -

25.14). The distribution of the abundances amongst the OTUs are also the most even in CG1 (E = 0.92 – 0.94). EJ3 appears to have the lowest bacterial diversity for all levels of OTU, with values between 0.87 and 1.29. This is also confirmed by the values from the Simpson Index (1. 65 to 2.58). This environment also has the lowest evenness (E = 0.45 – 0.62).

Table 37 Diversity Indices for the *Archaea*.

This shows the estimates of archaeal richness and diversity at three definitions of OTUs for the salt lakes and hot springs. Highest values for the salt lakes are in **red**; lowest values for the salt lakes are in **blue**.

99% (subspecies)

Index	BJ1	CG1	EJ3	EN1	SH1	XH2	LP4	TC8
OTUs	11	12	12	16	16	14	13	16
$H_{Shannon}$[a]	1.59	2.03	2.3	2.39	2.19	2.21	1.88	2.67
$H_{Simpson}$[a]	3.37	6.39	12.5	8.84	6.89	8.44	4.33	27.6
S_{Chao1}[b]	25	16.17	21	22.13	55.5	31	33	28.5
Evenness (E)[c]	0.66	0.82	0.93	0.86	0.79	0.84	0.73	0.96

97% (species)

Index	BJ1	CG1	EJ3	EN1	SH1	XH2	LP4	TC8
OTUs	11	11	12	15	14	11	10	16
$H_{Shannon}$[a]	1.59	1.83	2.3	2.21	2.07	1.63	1.67	2.67
$H_{Simpson}$[a]	3.37	5.31	12.5	7.34	6.21	3.45	4.04	27.6
S_{Chao1}[b]	25	14	21	18.5	34	22.5	15	28.5
Evenness (E)[c]	0.66	0.73	0.93	0.82	0.78	0.68	0.73	0.96

95% (genus)

Index	BJ1	CG1	EJ3	EN1	SH1	XH2	LP4	TC8
OTUs	9	9	12	12	12	10	9	16
$H_{Shannon}$[a]	1.39	1.58	2.3	1.94	1.91	1.53	1.57	2.67
$H_{Simpson}$[a]	2.88	4.18	12.5	5.64	5.63	3.18	3.7	27.6
S_{Chao1}[b]	14	10.25	21	13.67	26	17	11	28.5
Evenness (E)[c]	0.63	0.72	0.93	0.78	0.77	0.66	0.71	0.96

[a] Higher values indicate greater microbial diversity.

[b] Higher values indicate greater microbial richness.

[c] Values nearest to 1 indicate even distribution of abundances.

Table 38 Diversity Indices for the *Bacteria*.

This shows the estimates of bacterial richness and diversity at three definitions of OTUs for the salt lakes and hot springs. Highest values for the salt lakes are in **red**; lowest values for the salt lakes are in **blue**.

99% (subspecies)

Index	BJ1	CG1	EJ3	EN1	SH1	XH2	LP4	TC8
OTUs	12	19	8	19	18	19	26	12
$H_{Shannon}$[a]	2.20	2.77	1.29	1.85	2.52	2.59	3.14	1.6
$H_{Simpson}$[a]	12	25.14	2.58	3.44	12.78	12.78	56.1	2.96
S_{Chao1}[b]	45	31.5	8	31.66	72	34.67	45.64	43
Evenness (E)[c]	0.89	0.94	0.62	0.63	0.86	0.88	0.96	0.64

97% (species)

Index	BJ1	CG1	EJ3	EN1	SH1	XH2	LP4	TC8
OTUs	12	18	7	19	16	18	25	9
$H_{Shannon}$[a]	2.20	2.7	0.87	1.85	2.27	2.5	3.06	0.97
$H_{Simpson}$[a]	12	22	1.65	3.44	8.51	11.05	40.07	1.68
S_{Chao1}[b]	45	27.66	7	31.66	60	30.5	52.9	29
Evenness (E)[c]	0.89	0.93	0.45	0.63	0.82	0.86	0.95	0.44

95% (genus)

Index	BJ1	CG1	EJ3	EN1	SH1	XH2	LP4	TC8
OTUs	12	15	7	19	15	18	24	9
$H_{Shannon}$[a]	2.20	2.48	0.87	1.88	2.22	2.5	3.02	0.97
$H_{Simpson}$[a]	12	15.5	1.65	3.44	8.23	11.05	37.4	1.68
S_{Chao1}[b]	45	32	7	31.66	50	30.5	41.75	29
Evenness (E)[c]	0.89	0.92	0.45	0.64	0.82	0.86	0.95	0.44

[a] Higher values indicate greater microbial diversity.

[b] Higher values indicate greater microbial richness.

[c] Values nearest to 1 indicate even distribution of abundances.

Nucleotide Diversity

The nucleotide diversity of the microbial communities for the salt lakes and hot springs were assessed using MEGA 3.1 (Kumar *et al.* 2004). The results for the two hot springs show that LP4 had the higher bacterial nucleotide diversity and the lower archaeal nucleotide diversity (Table 39). This is consistent with previous diversity indices (Table 37 and Table 38). The results in Table 40 for the hot springs infer that on average, two randomly sampled archaeal sequences are 28% different, which is higher than for the bacterial community (22%). This implies that the overall nucleotide diversity is higher in the archaeal community than the bacterial community.

The results for the salt lakes show that BJ1 had the lowest archaeal nucleotide diversity, while EJ3 had the highest archaeal nucleotide diversity (Table 39). This is consistent with the previous diversity indices (Shannon-Weaver and Simpson), which show that BJ1 had the lowest archaeal diversity and EJ3 had the highest archaeal diversity (see Table 37). Moreover, EJ3 had the lowest bacterial nucleotide diversity, which is again consistent with the previous diversity indices (Shannon-Weaver and Simpson: see Table 38). However, SH1 shows the highest bacterial nucleotide diversity, which is inconsistent with the diversity indices, but is consistent with the Chao1 (richness) index. The results in Table 40 for the salt lakes infer that on average, two randomly sampled archaeal sequences are approximately 18% different. This is higher in the bacterial community, where on average the difference between two random sequences is 24% different (Table 40). This suggests a higher nucleotide diversity overall in the bacterial community than the archaeal community.

Table 39 Nucleotide Diversity.

This table shows the nucleotide diversity within each 16S rRNA library. The highest values are in **red** and the lowest are in **blue** (for the salt lakes only).

Sampling Site	*Archaea*	*Bacteria*
BJ1	**0.079 ± 0.007**	0.261 ± 0.016
CG1	0.090 ± 0.009	0.235 ± 0.016
EJ3	**0.186 ± 0.012**	**0.102 ± 0.007**
EN1	0.168 ± 0.012	0.241 ± 0.013
SH1	0.150 ± 0.011	**0.272 ± 0.015**
XH2	0.098 ± 0.008	0.264 ± 0.015
LP4	0.219 ± 0.014	0.232 ± 0.014
TC8	0.268 ± 0.014	0.090 ± 0.006

Table 40 Nucleotide Diversity.

These tables show the nucleotide diversity for entire communities of *Archaea* and *Bacteria* in the salt lakes and hot springs.

Hot springs

Nucleotide diversity	*Bacteria*	*Archaea*
Mean diversity for entire community	0.222 ± 0.016	0.277 ± 0.015

Salt lakes

Nucleotide diversity	*Bacteria*	*Archaea*
Mean diversity for entire community	0.243 ± 0.017	0.182 ± 0.014

Lineage-per-time Plots

These plots were determined using DOTUR (Schloss and Handelsman 2005). The interpretation of the different outcomes of these plots is explained in the Introduction (Figure 8, p70). The results for *Archaea* and *Bacteria* are illustrated in Figure 49. These graphs show the number of lineages present on a phylogenetic tree relative to evolutionary time.

The hot spring, LP4 shows a convex plot for the archaeal community, indicating that the *Archaea* here are less divergent and may have experienced a more recent selection event (Martin 2002). The *Bacteria* produce a concave plot, showing an excess of highly divergent lines, indicating that selection factors maintain a high diversity in the bacterial community (Martin 2002). The plots for *Archaea* for the hot spring TC8 is also both concave, but the plot for the *Bacteria* in TC8 is linear, suggesting constant birth and death rates. This means that no clear conclusions can be made about the genetic diversity and that the null expectation applies [that the time between successive divergence events decreases as the distance from the root decreases (Martin 2002)].

The salt lake samples show that BJ1, CG1 and EJ3 have constant birth and death rates for the *Archaea*. However, EN1, SH1 and XH2 give a convex plot, suggesting a recent selection event. The *Bacteria* in these salt lakes show yet a different pattern. Constant birth and death rates appear in *Bacteria* for sites EJ3, SH1 and XH2. A convex plot for *Bacteria* appears in CG1, suggesting the presence of less divergent lineages. However, BJ1 and EN1 have concave plots, suggesting that selection that maintains a high diversity.

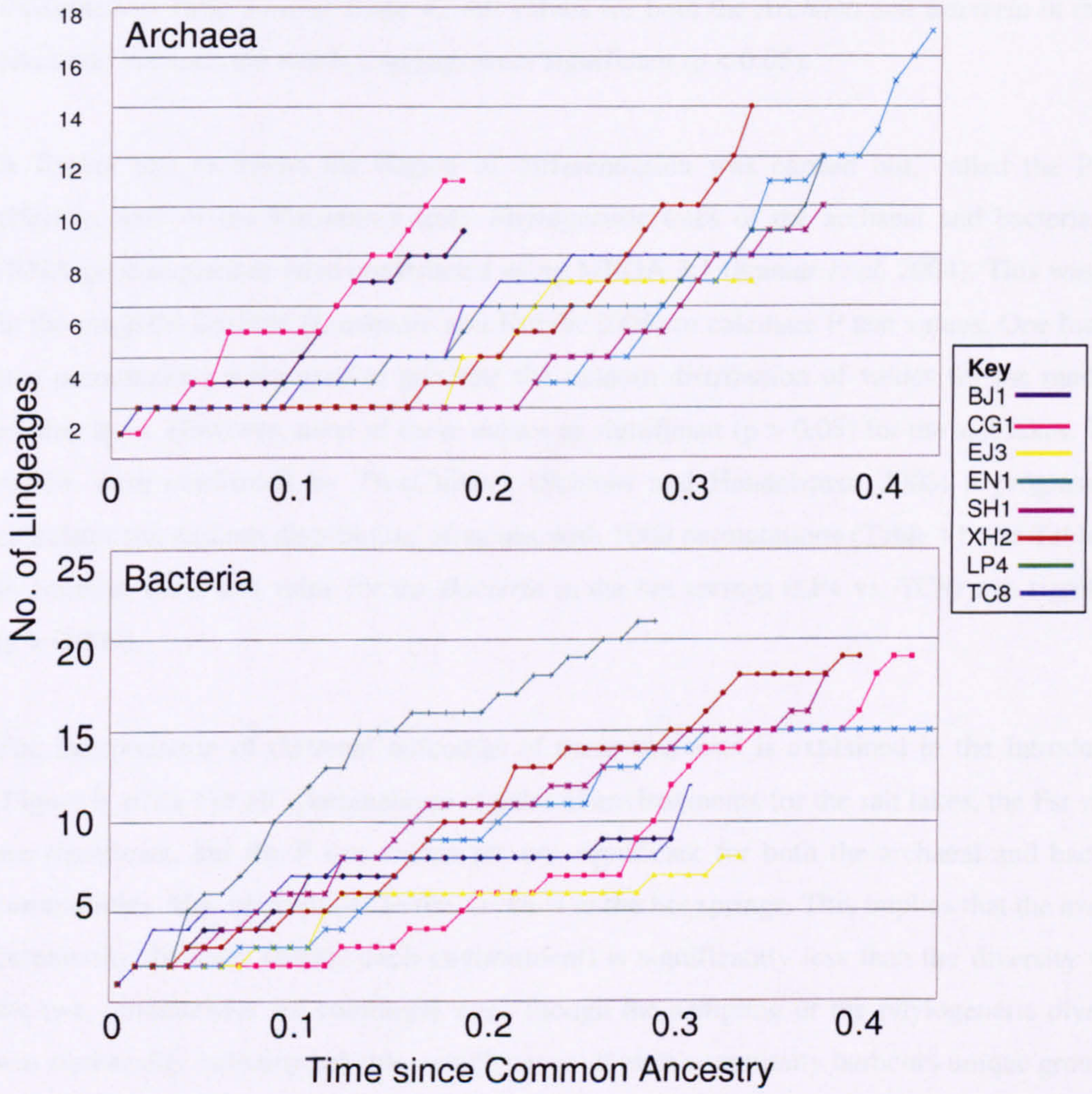


Figure 49 Lineage-per-time Plots.

These plots show the number of lineages present on a phylogenetic tree relative to time (in arbitrary units) since the common ancestor for the archaeal and bacterial communities for the salt lakes and hot springs.

Assessing the Degree of Differentiation

The program MEGA 3.1 (Kumar *et al.* 2004) was used to calculate coefficient of differentiation (Fst) values between pairs of communities. The salt lakes and hot springs were treated separately. The statistical significance of the Fst was assessed by randomly assigning sequences to communities and calculating the Fst for 1000 permutations. The values are illustrated in Table 41 and Table 42 All values for both the *Archaea* and *Bacteria* in the salt lakes and between the two hot springs were significant ($p < 0.05$).

A further test to assess the degree of differentiation was carried out, called the P Test (Phylogenetic or the Parsimony test). Phylogenetic trees of the archaeal and bacterial 16S rRNA gene sequences were constructed using MEGA 3.1 (Kumar *et al.* 2004). This was used in the program UniFrac (Lozupone and Knight 2005) to calculate P test values. One hundred tree permutations were used to generate the random distribution of values for the randomly joining trees. However, none of these values as significant ($p > 0.05$) for the salt lakes. These results were confirmed by TreeClimber (Schloss and Handelsman 2006) a program that calculates the random distribution of values with 1000 permutations (Table 41 and Table 42). In contrast, the P test value for the *Bacteria* in the hot springs (LP4 vs. TC8) was significant ($p = 0.009$).

The interpretation of different outcomes of these two tests is explained in the Introduction (Figure 9, p72). For all combinations of pairs of environments for the salt lakes, the Fst values are significant, but the P test values are not significant for both the archaeal and bacterial communities. This also applies to the *Archaea* in the hot springs. This implies that the average community diversity (within each environment) is significantly less than the diversity when the two communities are combined, even though the sampling of the phylogenetic diversity was statistically indistinguishable, which occurs if each community harbours unique groups of closely related microbes and that these groups are interspersed across the tree (Martin 2002).

However, the *Bacteria* in the hot springs had significant values for both the Fst and P test. This indicates less diversity within each community than for the two communities combined and that the two communities contain distinct phylogenetic lineages (Martin 2002). This evaluation has further proven that the diversity was disparate, with little phylogenetic overlap.

Table 41 Fst and P-test for the *Archaea*.

These tables show values calculated for Fst (using MEGA) and P-test (using TreeClimber) between the salt lakes and hot springs. Significance (p) values are in brackets. Values that are significant are in bold ($p<0.05$).

Fst	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3	LP4
CG1	**0.002**	x						
	(0.007)							
EJ3	**-0.008**	**-0.032**	x					
	(0.011)	**(0.010)**						
EN1	**0.109**	**0.092**	**-0.025**	x				
	(0.011)	**(0.009)**	**(0.008)**					
SH1	**0.078**	**0.067**	**0.012**	**-0.002**	x			
	(0.009)	**(0.008)**	**(0.006)**	**(0.007)**				
XH2	**0.087**	**0.065**	**0.014**	**0.009**	**-0.005**	x		
	(0.008)	**(0.008)**	**(0.009)**	**(0.004)**	**(0.005)**			
XH3	**0.120**	**0.085**	**-0.009**	**-0.015**	**-0.010**	**0.007**	x	
	(0.011)	**(0.009)**	**(0.008)**	**(0.002)**	**(0.005)**	**(0.004)**		
TC8	x	x	x	x	x	x	x	**-0.005**
								(0.008)

P test	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3	LP4
CG1	0.097	x						
	(0.125)							
EJ3	0.114	0.274	x					
	(0.139)	(0.371)						
EN1	0.285	0.093	0.434	x				
	(0.420)	(0.126)	(1.000)					
SH1	0.342	0.102	0.058	0.261	x			
	(0.630)	(0.126)	(0.073)	(0.466)				
XH2	0.304	0.169	0.187	0.226	0.225	x		
	(0.525)	(0.233)	(0.233)	(0.472)	(0.363)			
XH3	0.120	0.224	0.156	0.206	0.281	0.134	x	
	(0.155)	(0.362)	(0.188)	(0.905)	(0.502)	(0.226)		
TC8	x	x	x	x	x	x	x	0.292
								(0.414)

Table 42 Fst and P-test for the *Bacteria*.

These tables show values calculated for Fst (using MEGA) and P-test (using TreeClimber) between the salt lakes and hot springs. Significance (p) values are in brackets. Values that are significant are in bold ($p<0.05$).

Fst	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3	LP4
CG1	-0.011	x						
	(0.008)							
EJ3	0.008	0.037	x					
	(0.007)	(0.011)						
EN1	0.003	0.029	0.032	x				
	(0.008)	(0.004)	(0.009)					
SH1	-0.006	0.002	0.029	0 (0.002)	x			
	(0.008)	(0.002)	(0.009)					
XH2	0 (0.007)	0.010	0.037	0.006	-0.003	x		
		(0.003)	(0.008)	(0.002)	(0.002)			
XH3	-0.007	0 (0.003)	0.035	0.006	-0.010	-0.009	x	
	(0.006)		(0.008)	(0.003)	(0.002)	(0.002)		
TC8	x	x	x	x	x	x	x	0.139
								(0.009)

P test	BJ1	CG1	EJ3	EN1	SH1	XH2	XH3	LP4
CG1	0.242	x						
	(0.341)							
EJ3	0.379	0.048	x					
	(0.685)	(0.056)						
EN1	0.371	0.060	0.201	x				
	(0.760)	(0.087)	(0.264)					
SH1	0.280	0.250	0.180	0.216	x			
	(1.000)	(0.544)	(0.225)	(0.412)				
XH2	0.269	0.113	0.174	0.161	0.173	x		
	(1.000)	(0.188)	(0.229)	(0.882)	(0.304)			
XH3	0.373	0.153	0.189	0.257	0.220	0.214	x	
	(0.782)	(0.934)	(0.261)	(0.633)	(0.717)	(0.780)		
TC8	x	x	x	x	x	x	x	**0.002**
								(0.009)

Summary

A summary of findings for the hot springs are as follows:

- PCA showed that the samples have very similar archaeal communities, but different bacterial communities.

- TC8 had the highest archaeal richness and diversity, whereas LP4 had the highest bacterial richness and diversity.

- This was supported by the nucleotide diversity. Overall, the archaeal community showed the highest nucleotide diversity.

- The lineage-per-time plots indicated that LP4 had a stable and highly diverse bacterial community with an archaeal community of less diverse, closely related lineages.

- The lineage-per-time plots also showed that the archaeal community in TC8 is stable and highly diverse, while the bacterial community showed constant birth and death rates.

- Significant Fst and insignificant P test indicates that the archaeal communities are not disparate between LP4 and TC8.

- Significant Fst and P test indicates that the bacterial communities are disparate between LP4 and TC8.

A summary of findings for the salt lakes are as follows:

- PCA identified CG1 as the most different sample compared to the other salt lakes.

- The Jaccard Index showed that no two samples harboured identical communities and that microbial diversity is not randomly distributed (rejection of the null Hypothesis).

- The simple Mantel test revealed that the likely factors causing the differences in biotic composition were concentrations of Na, Mg, CO_2 and HCO_3, temperature and pH.

- SH1 had the highest archaeal and bacterial richness (Chao1). CG1 had the lowest archaeal richness, while EJ3 had the lowest bacterial richness.

- EJ3 had the highest archaeal diversity and lowest bacterial diversity (Shannon-Weaver and Simpson). CG1 had the highest bacterial diversity, while BJ1 had the lowest archaeal diversity (this is an underestimate).

- Nucleotide diversity studies generally agreed with the Shannon-Weaver and Simpson diversity indices, (with the exception for SH1 bacterial diversity). Overall, the bacterial community showed the highest nucleotide diversity.

- The majority of lineage-per-time plots indicated that the archaeal and bacterial communities are less diverse and may have undergone a recent selection event.

- Significant Fst values and insignificant P test values for both the archaeal and bacterial communities implies that the average community diversity (within each environment) is significantly less than the diversity when the two communities are combined, which occurs if each community harbours unique groups of closely related microbes that are spread across the tree (Martin 2002). The communities are not disparate.

DISCUSSION: MICROBIAL BIODIVERSITY IN THE HOT SPRINGS

Sample Coverage

This was assessed by rarefaction curves (Figure 25, p149); they showed that further sampling was required from the TC8 archaeal 16S rRNA library and the LP4 and bacterial 16S rRNA library as these plots did not plateau, and so it should be noted that analysis of diversity for the archaeal community in TC8 and the bacterial community in LP4 will be underestimated

Statistical Analysis of Microbial Biodiversity and Genetic Diversity

The Principal Component Analysis (PCA) suggested that the archaeal community composition of the two hot spring samples were very similar, since the two points clustered together in space (Figure 47, p276). However, they differed significantly in the bacterial diversity as the two points did not cluster together (Figure 48, p277). Both hot spring mats were composed of different bacterial populations, hence the appearance of different microbial mats (laminated cyanobacterial mat at LP4 and the *Aquificae* streamers at TC8).

The richness estimator (Chao1), and diversity indices (Shannon-Weaver and Simpson) were calculated for the two environments (Table 37, p292, Table 38, p293). Between the two environments, LP4 had the highest bacterial richness and diversity (Chao1 ~ 42-53, Shannon-Weaver ~ 3, Simpson ~ 37 – 56). This was supported by the analysis of nucleotide diversity, which showed that between the two environments, LP4 had higher nucleotide diversity (0.232: Table 39, p295). Furthermore, the lineage-per-time plots were concave for the bacterial community (Figure 49, p297), which is indicative of a stable community and that phylogenetically, it consists of highly divergent lineages, and therefore contains high genetic diversity. The interpretation of the various outcomes for the lineage-per-time plots is explained in the Introduction (Figure 8, p70). The value for bacterial diversity in LP4 is similar to those observed for diverse microbial environments, the human mouth and gut (Kroes *et al.* 1999; Suau *et al.* 1999; Martin 2002), thereby indicating that LP4 is also a bacterially diverse environment. (OTUs were defined at 99% identity in the mouth; and 98% identity in the gut). All these analyses support the fact that laminated cyanobacterial mats are a complex community consisting of cyanobacteria, anoxygenic phototrophs, thermoanaerobic fermenters and possibly a number of other heat adapted bacteria. The lower archaeal richness and diversity may be because few *Archaea* are known to be associated with laminated cyanobacterial mats (Ward *et al.* 1998). The lineage-per-time plots indicate that the archaeal

community contains an excess of closely related lineages and that the archaeal community may have undergone a recent selection event (Martin 2002). Perhaps the constant fluctuation in water temperature does not allow a stable community, unlike the *Bacteria* in cyanobacterial mats that have been shown to cope with daily fluctuations in temperature as well as disturbance (Ferris *et al.* 1997; Ward *et al.* 1998).

Between the two environments, TC8 had the highest archaeal richness and diversity (Chao1 ~ 29, Shannon-Weaver = 2.67, Simpson ~ 28: Table 37, p292), which was supported by the analysis of nucleotide diversity (0.268: Table 39). Moreover, the lineage-per-time plots indicate that phylogenetically, the archaeal community contains an excess of highly divergent lineages, and therefore shows higher genetic diversity (Figure 49). It follows that this community did not have a high bacterial diversity since it was dominated by *Aquificae*. The lineage-per-time plot for the *Bacteria* was linear and not concave, hence supporting the fact that phylogenetically, the bacterial community is not highly diverse (Figure 8). The higher archaeal diversity could be attributed to the high temperature of the water that the streamers were submerged in as this would support more *Archaea*, particularly *Crenarchaeota* and *Korarchaeota* that require 70°C - 113°C for growth (Barns *et al.* 1994; Garrity and Holt 2001b), which would be detrimental to most *Bacteria*. Indeed, one clone from TC8 was affiliated with the *Korarchaeota*, while no clones from LP4 were affiliated with this group. Furthermore, two clusters that were completely unrelated to existing species were detected in the TC8 library (Figure 27, p159). Unfortunately, this is not consistent with the PCA that suggested that the two environments harboured similar archaeal communities. Perhaps in terms of genetic diversity across all eight samples (including the salt lakes), the analysis became insensitive to the differences in the two archaeal communities i.e. perhaps if only LP4 and TC8 were analysed by PCA, the variation in the two communities would become more apparent. Furthermore, the diversity for the archaeal community in TC8 was an underestimate as not enough clones were sampled according to the rarefaction curves (Figure 25, p 149). Perhaps further sampling of the library would help resolve this discrepancy.

The interpretation of different outcomes of the Fst and P tests are explained in the Introduction (Figure 9, p72). The Fst and P tests showed that in terms of phylogeny, the archaeal diversity between the two environments was not disparate, and contained groups of closely related microbes that are interspersed across the tree (Martin 2002). Indeed, the phylogenetic tree for the archaeal community showed groups of clones from LP4 and TC8

affiliated with the *Korarchaeota*, *Crenarchaeota* and *Euryarchaeota* (Figure 27, p159). However, the bacterial community showed that the diversity was disparate, with little phylogenetic overlap (Martin 2002), which is consistent with previous discussions.

These statistical analyses and the results from the molecular data show that the microbial community composition that develop at higher temperatures (in this case, 72.1°C) are drastically different to the composition that develop in lower temperatures (60 - 65°C). It is likely that the major factor for this is temperature and pH, although other factors such as water flow, oxygen availability and sulphide concentrations have also been implicated (Skirnisdottir *et al.* 2000; Nakagawa and Fukui 2002).

DISCUSSION: MICROBIAL BIODIVERSITY AND BIOGEOGRAPHY IN THE SALT LAKES

Sample Coverage

This was assessed by rarefaction curves (Figure 25, p149). The majority of the plots had reached saturation point since they had come to a plateau, suggesting that the predominant phylogenetic groups had been sampled. The only rarefaction curve that did not reach saturation was for the bacterial library for BJ1. This is clearly still in the linear phase, and so it should be noted that analysis of diversity for BJ1 will be underestimated.

Statistical Analysis of Microbial Biodiversity and Genetic Diversity

The archaeal and bacterial richness was the highest in SH1 (26 – 55.5 and 50 – 72 respectively). The richer archaeal and bacterial communities may be attributed the availability of organic matter from vegetation that grew up to the waters edge, which was not observed for any of the other salt lakes. CG1 had the lowest archaeal richness (10.25 – 16.17), while EJ3 had the lowest bacterial richness (7 – 8) (Table 37, p292 and Table 38, p293).

The Chao1 values are not consistent with the values from the diversity indices, Shannon-Weaver and Simpson. Chao1 estimates richness (the total number of individuals) of an environment by counting the number of observed species and estimates the number of species that were not observed i.e. the rarer species. It does not show the complexity of an environment: no information can be inferred as to how the community is distributed among the species found. However, both the Shannon-Weaver and Simpson indices is influenced by the both the number of observed species and the abundance or frequency of each of those species; both indices are affected if the distribution amongst the species is even. Since Chao1 estimates richness, while Shannon-Weaver and Simpson measures frequency and richness, and since they are affected by different factors, possibly means that they may not necessarily agree with each other.

The values for the diversity indices, Shannon-Weaver and Simpson are shown in Table 37 and Table 38. EJ3 has the highest Shannon-Weaver and Simpson values for archaeal diversity at all OTU levels (2.3 and 12.5 respectively). This was supported by the analysis for nucleotide diversity, since EJ3 had the highest nucleotide diversity for the archaeal community (0.186: Table 39, p295). Conversely, EJ3 had the lowest Shannon-Weaver and

Simpson values for the bacterial community (0.87 – 1.29 and 1.65 - 2.58 respectively). Again, this was supported by the nucleotide diversity, as EJ3 scored the lowest (0.102). Furthermore, EJ3 had the lowest estimate for bacterial richness (Chao1; see above). It also had the lowest evenness values (0.45 – 0.62), suggesting that there was one dominating bacterial species. This is consistent with the findings from the molecular data that showed this saltern to contain a climax population of haloarchaea and very few bacterial species, the dominating clones being relatives of *Salinibacter* (Table 16, p190).

CG1 had the highest bacterial diversity (Shannon-Weaver = 2.48 – 2.77, and Simpson = 15.5 – 25.14), but this was not supported by nucleotide diversity. CG1 had the lowest observed salinity (Figure 21, p138), and so it would follow that this lake had the highest bacterial diversity as there would be fewer haloarchaea to compete with. Hence, it would be expected that CG1 had the lowest archaeal diversity, however, BJ1 had the lowest archaeal diversity (Shannon-Weaver = 1.39 – 1.59, Simpson = 2.88 – 3.37), which was supported by nucleotide diversity (0.079). However, these are an underestimate, since the sampling effort of the 16S rRNA library did not reach saturation (Figure 25, p149).

Since all the measures of richness and diversity do not agree, no clear conclusions can be made.

The observed values for the Shannon-Weaver index (neither the *Archaea* nor *Bacteria* exceeding 2.77) are consistent with those found for other saline environments, such as the crystallisers of an Adriatic solar saltern (Pašić *et al.* 2005). These values are less than microbially diverse environments such as the human mouth and gut (Kroes *et al.* 1999; Suau *et al.* 1999; Martin 2002), hence these saline environments are not microbially diverse.

Estimating Diversity by Phylogeny

These methods are not concerned with numbers, but looks at the distribution of lineages throughout the phylogenetic tree. The interpretation of the different outcomes for lineage-per-time plots is outline in the Introduction (Figure 8, p70). The lineage-per-time plots for the archaeal communities in EN1, SH1 and XH2, and the bacterial community in CG1 were convex (Figure 49, p297), suggesting that, in terms of phylogeny, the archaeal community contains an excess of closely related lineages, and that a recent selection event may have occurred (Martin 2002). The linear plots, indicative of constant birth and death rates were

shown in archaeal communities in BJ1, CG1 and EJ3, and the bacterial communities in EJ3, SH1 and XH2, and point towards communities containing less divergent lineages, and therefore less genetically diverse communities.

The likely cause for this is the high salinity of the environments, which means that only specialised *Archaea* are able to thrive. Indeed, molecular analysis showed that the majority of *Archaea* in the salt lakes belonged to the highly specialised group, *Halobacteriales*. A further selection factor for CG1 is the highly alkaline pH; hence all haloarchaea in this environment must be tolerant to high pH as well as high salinity. This was again supported by the molecular analysis that showed that clones were related to known haloalkaliphilic *Archaea*.

The coexistence of several closely related organisms has been termed 'microdiversity,' which has been observed elsewhere. The haloarchaeal community in the salterns in Spain (Benlloch *et al.* 2002) was shown to contain 'microdiversity.' Similarly, closely related sequences affiliated with microdiverse groups appeared in the DGGE analysis of lakes on the Tibetan Plateau (Wu *et al.* 2006a).

However, the plots for the bacterial communities in BJ1and EN1 were concave (Figure 49), suggesting that these environments maintain an excess of highly divergent lineages, and are therefore genetically diverse. It was noted that the salterns at EN1 appear red during the summer months, which aids the precipitation of salt for harvesting. However, during the colder months, the salterns lose their pigment, implying a decline in archaeal community and so soda is harvested instead. At the time of sampling, the EN1 salterns were white; hence it is possible that the decline in haloarchaea allowed more bacteria to thrive, leading to a more phylogenetically diverse environment. It was also noted that there was animal contamination at BJ1, which may account for the input of bacteria. Indeed, the molecular analysis showed some clones that were related to bacteria found in animal waste (Table 16), particularly at Lake Bagaejinnor and Lake Shangmatala. These would not be related to the halophilic bacterial species in the lake and may account for the phylogenetic diversity observed.

The interpretation for the different results for the Fst and P tests are explained in the Introduction (Figure 9, 72). The Fst and P tests are consistent with microbial communities that contain closely related lineages. Significant Fst values and insignificant P test values for both the archaeal and bacterial communities (Table 41, p299 and Table 42, p300) implies that the

average community diversity (within each environment) is significantly less than the diversity when the two communities are combined, even though the sampling of the phylogenetic diversity was statistically indistinguishable, which occurs if each community harbours unique groups of closely related microbes that are spread across the phylogenetic tree (Martin 2002). It therefore implies that the communities are not disparate.

Microbial Biogeography

When considering just the Principal Component Analysis (PCA) results for the salt lake communities, CG1 was clearly the most different for both the archaeal and bacterial communities as CG1 it did not cluster with the points for the other salt lakes (Figure 47, p276 and Figure 48, p277). Indeed, measurements taken on the expedition combined with 16S rRNA data support these findings; CG1 was found to have the lowest salinity, lowest temperature and highest pH, and the 16S rRNA data showed that CG1 contained haloalkaliphilic species of the genera that were detected in the other salt lakes.

The null hypothesis states that the microbial diversity in hypersaline environments is randomly distributed (Martiny *et al.* 2006). The figures calculated for the Jaccard Index clearly demonstrate that this is not the case, since none of the two environments being compared scored a value of 1 (Table 23, p280 Table 24, p281). Since the null hypothesis was rejected, there were clearly factors that affected the distribution of microbes in hypersaline environments. Factors that could explain these variations were investigated.

The biogeography of the *Archaea* and *Bacteria* within the salt lakes was assessed using methods outlined in the review by Martiny *et al* (2006). The results obtained suggest that only environmental factors were significant in driving biotic similarity, namely temperature, pH and ion concentrations, and it was not the influence of geographical distance and evolutionary time. However, this was not the case when samples were taken at a greater geographic distance (intercontinental distances). When the Inner Mongolian salt lakes were used in an analysis that included a salt lake from Argentina, geographic distance, and therefore evolutionary time, was highly significant in driving bacterial biotic composition (Wang 2006). The observation that biotic similarity decreased as geographic distance increased was explained by the fact that dispersal barriers were more likely, which meant that evolutionary events such as speciation and extinction gave rise to differences in the communities (Wang 2006). It is well documented that prokaryotic dispersal can occur over large distances by wind

(Echigo *et al.* 2005). The winds observed on the Inner Mongolian Steppe would allow such dispersal. However, it is clear that the geochemistry of the lakes and seasonal changes are responsible for determining whether these organisms are able to survive and colonise. Therefore, the hypothesis 'everything is everywhere; the milieu selects' appears to be applicable for the microbiota in the Inner Mongolian salt lakes from the area sampled.

For the *Archaea*, the factors driving community composition at the genus level were temperature, pH and Mg ions. The r values were all positive, indicating that as these factors became more similar, biotic similarity also increased. The pH and concentrations of Na, carbonate and bicarbonate then become the significant factors in influencing archaeal community composition at the species and subspecies level. Since the major group found within the six salt lakes were from the *Halobacteriales*, it follows that Na and Mg ions (probably in the form of NaCl and $MgCl_2$) were found to be significant, as all members of the *Halobacteriales* are extremely halophilic, requiring salinities approaching saturation (Grant *et al.* 2001). pH is also considered one of the important factors in determining the microbial community in hypersaline brine (Grant 2004). *Archaea* adapt to high salinity by accumulation of KCl, thereby excluding Na ions to maintain osmotic stability (Galinski and Trüper 1994). However, the volume of cells is small compared with the volume of brine, so even at low concentrations of K ions, archaeal cells are able to cope with high external NaCl concentrations.

For the *Bacteria*, temperature and concentration of Mg ions were significant at the genus level. Again, all r values were positive integers. Since the majority of bacterial species detected were halophilic, again it follows that Mg ions was a significant factor. Surprisingly, concentration of Na ions was not significant. At the species level, concentration of Mg ions is the only significant factor. *Bacteria* adapt to the high salinity environment by the accumulation of compatible solutes, organic compounds that have osmotic potential (Roberts 2005), and so it appears that salinity is no longer significant at the subspecies level, but bicarbonate ions (and perhaps carbonate ions) becomes the most significant, both of which affect the pH.

These findings are consistent with previous studies. Crump *et al* (2004) found that bacterial community composition correlated with salinity at a small geographic scale (5 km) (Crump *et al.* 2004). Another study conducted in salterns, also at a small geographic scale (3 km) found

that archaeal, bacterial and eukaryal community composition correlated with salinity (Casamayor *et al.* 2002). In addition, it was observed that the haloarchaeal community composition and abundance correlated with salinity over several hypersaline lakes in the Vestfold Hills, Antarctica (Bowman *et al.* 2000a). Moreover, the bacterial community composition in 16 lakes on the Tibetan Plateau showed that salinity was the major factor responsible for differences in microbial communities and that altitude (and therefore geographic distance) was not a factor (Wu *et al.* 2006a). A study of saline soil in British Columbia, Canada correlated the fluctuations in archaeal community composition with both salinity and pH; the more alkaline soils showed more of the alkaliphilic members of the *Halobacteriales* (Walsh *et al.* 2005).

Most haloarchaea have an optimal growth temperature between 35°C and 50°C (Grant *et al.* 2001). Walsh and colleagues (2005) found that a predominant ribogroup related to *Halorubrum lacusprofundi*, an archaeon that has a growth temperature range of 4°C – 40°C (Franzmann *et al.* 1988; Grant *et al.* 2001) was detected in saline soils (Walsh *et al.* 2005). Similarly, an isolate related to *Halorubrum lacusprofundi* was cultivated from BJ1 (Figure 42, p237), and several clones from BJ1 and CG1 libraries were related to 16S rRNA sequences detected in Antarctica (Table 16, p190). It has been postulated that organisms capable of growth at low temperatures would have a selective advantage over the warm-adapted relatives during the colder months (Walsh *et al.* 2005). Experiments with samples from a saltern showed that at 35°C, dense growth of archaea were observed at 35% and 40% salt, but at 25°C, very little archaeal growth was observed (Ventosa *et al.* 1998b), again indicating that those archaea capable of growth at low temperatures would do better. It is also known that temperature affects some organisms' ability tolerate high salinity (Grant *et al.* 1998). For example, *Marinococcus halophilus* grows at 0.01 M NaCl at 20°C, but it requires 0.5 M NaCl at 25°C (Ventosa *et al.* 1998b). Therefore, it follows that seasonal changes in environmental temperature is also significant in driving community composition. Temperature also plays a role in competition between halophilic bacteria and archaea. Growth of moderately halophilic bacteria and haloarchaea overlap at intermediate salinity ranges (20% – 30%). It was found that temperature was a deciding factor, with bacterial growth being favoured at lower temperatures (Ventosa *et al.* 1998b).

RESULTS: CHARACTERISATION OF TWO NOVEL HALOARCHAEAL VIRUSES

Identification of Hosts and Host Range

Haloviruses were isolated by incubating some of the salt lake water with a liquid culture of archaeal cells (see Methods) that were previously retrieved from the salt lakes (see culturing chapter). Many different sized plaques were produced in a confluent lawn of archaeal cells (data not shown). Only two different sized plaques were picked for plaque purification; subsequently two different haloarchaeal viruses were isolated, designated virus BJ1 and virus BJ2. Virus BJ1 was propagated on archaeal isolate BJ1B11 and virus BJ2 was propagated on archaeal isolate BJ1A4. These archaeal hosts were characterised by their 16S rRNA gene sequence and were affiliated with the genus *Halorubrum* within the order *Halobacteriales.* The relationship of these archaeal strains with documented *Halorubrum* species has been described previously (see culturing chapter). Alignment of the 16S rRNA gene sequences of *Halorubrum* strains BJ1B11 and BJ1A4 showed that over ~1400 bp, they were 99.57% identical (six base mismatches). They have had further taxonomy analysis and appear to be a novel species (A. Ventosa, personal communication).

The host ranges of these haloarchaeal viruses were tested using the other archaeal strains previously isolated from the salt lakes. The host range of virus BJ1 was not determined as it was difficult to see these plaques. Virus BJ2 was able to form plaques on archaeal isolate BJ1B11, but not on *Halorubrum* strains BJ1A12B, BJ1C11, SH1D13 or on the type strain *Halorubrum saccharovorum.* Furthermore, other field strains donated from the University of Seville also isolated from Inner Mongolian salt lakes (A. Ventosa, personal communication) were not hosts for virus BJ2, suggesting that it has a narrow host range.

Plaque Morphology

Plaques for both viruses BJ1 and BJ2 required one to two weeks to appear since the hosts are slow growing. Virus BJ1 produced various sized plaques, which ranged between 1 to 5 mm in diameter and were irregularly shaped. This was probably due to variations in growth conditions. It was evident that the larger plaques were turbid with an undefined hazy edge, though no attempts to culture archaea from these turbid plaques were made. Virus BJ2 produced circular plaques with a defined edge that were approximately 1 mm in diameter, and were produced consistently upon infection. The plaques are illustrated in Figure 50.

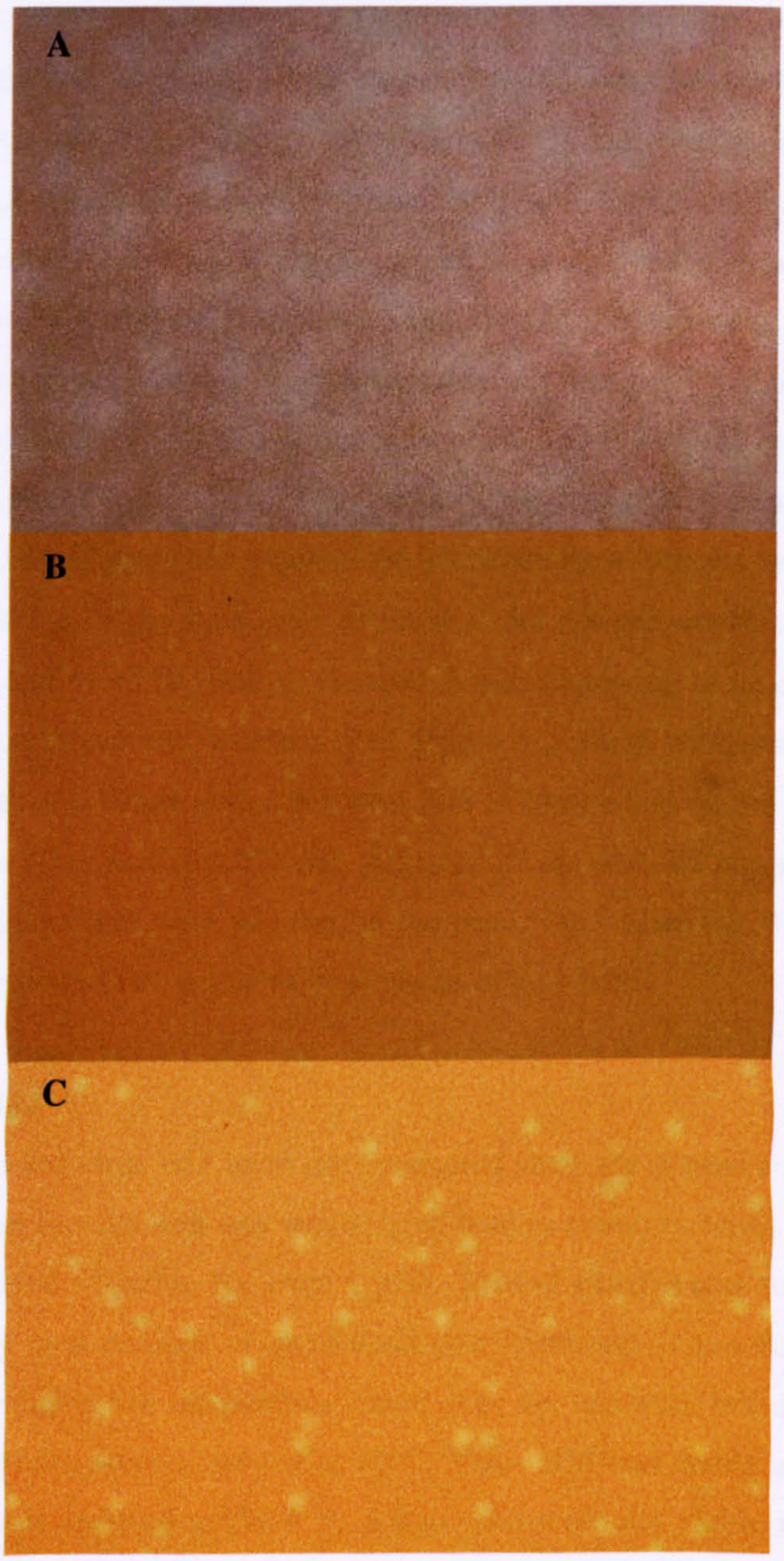


Figure 50 Plaque Morphology.

Plaques for virus BJ1 are shown in panels A and B. Sizes of these plaques varied according to growth conditions. Plaques for virus BJ2 is shown in panel C, and were produced consistently upon infection.

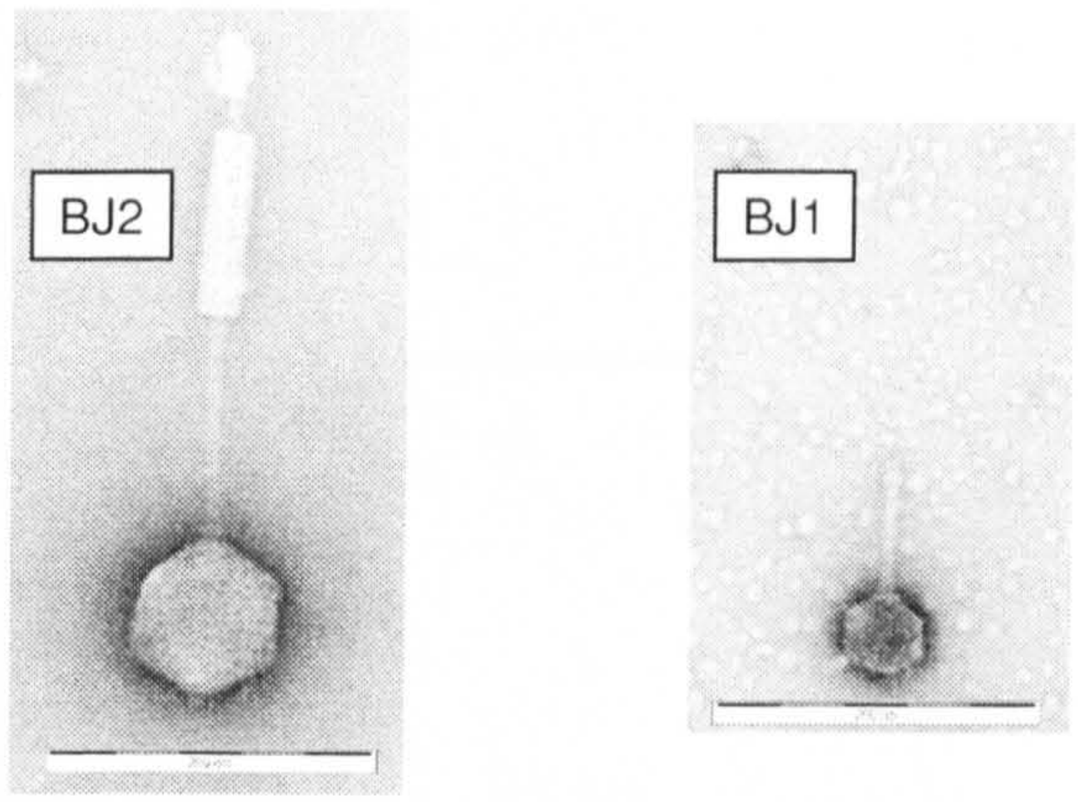
Transmission Electron Microscopy (TEM)

Suspensions of virus particles were fixed and stained with uranyl acetate, which are illustrated in Figure 51 (top and bottom panels). Both viruses have icosahedral heads, collars and tails. There are no obvious base plates or tail fibres on either of the virus particles. The vesicles attached to the end of virus BJ2 tails are probably cell debris. Additional features on virus BJ2 are a protuberance on the icosahedral head opposite the collar and a sheath surrounding the tail. This sheath appears in different positions along the tail (bottom panel). However, it may be the result of osmotic shock of the virus particles during fixing and staining procedures that caused this sheath to contract and to subsequently break away from the collar, allowing it to move along the exposed central tail. The presence of dissociated heads from tails, and collapsed heads on all preparations support the fact that these viruses were under osmotic stress. Stepwise washing of the salts prior to fixation did not prevent this (data not shown). The icosahedral head of virus BJ1 shows electron dense shadowing in its centre. Sizes of the features are shown on the central schematics of Figure 51, which were determined by image averaging. The length of a single vertex for virus BJ1 is 28 nm, and 46 nm for virus BJ2. The average length of the virus particles BJ1 and BJ2 are 127 nm and 400 nm, respectively. Virus BJ2 often formed dimers that were attached at the base plates (data not shown), though this may be due to absorption of two particles onto pieces of cell debris.

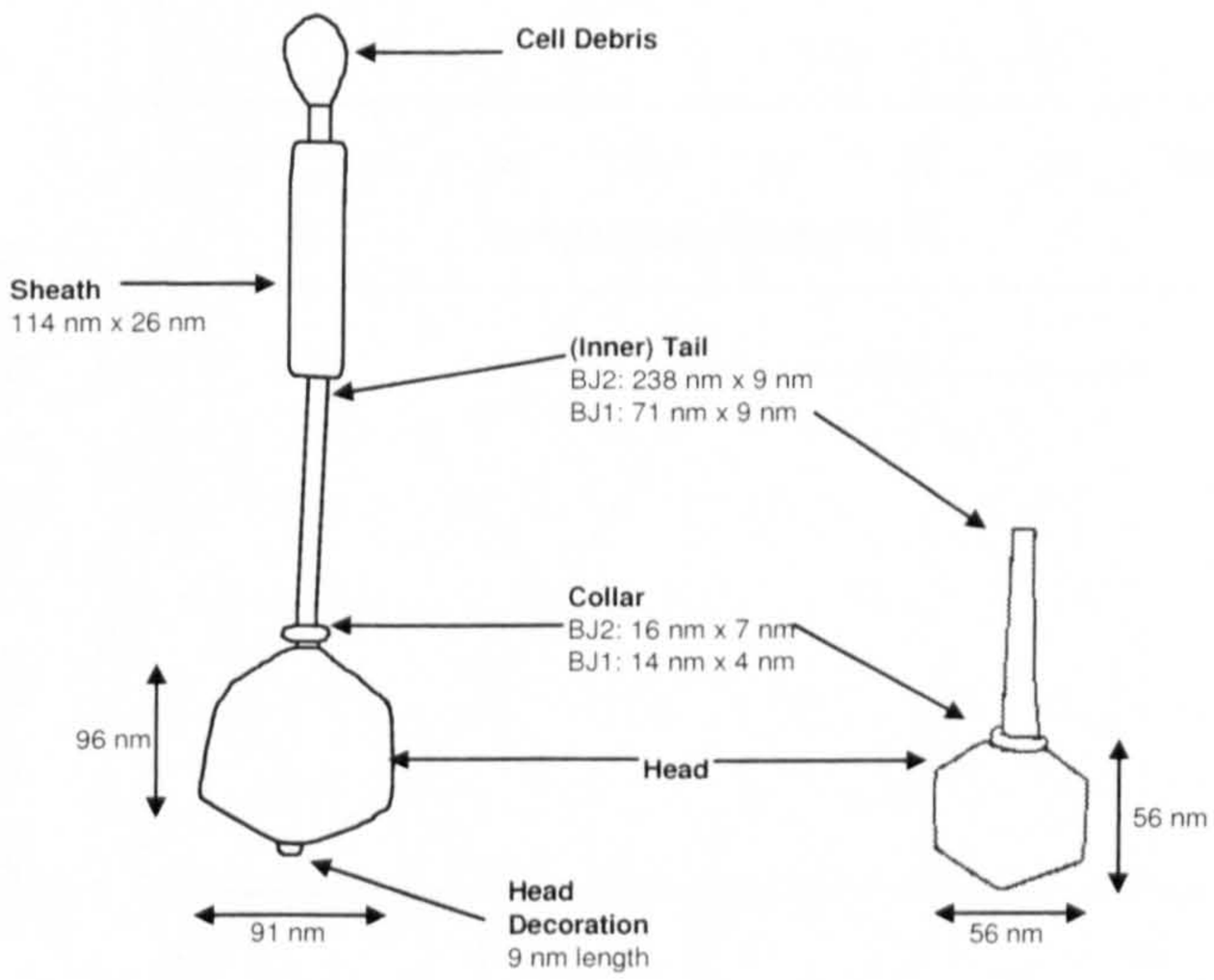
Stability of Virus BJ2

Stability experiments for virus BJ1 were not conducted since the plaque sizes produced with each experiment were inconsistent and would have lead to counting errors. Virus BJ2, when stored at 4°C, remained infective for over a year. Several experiments were carried out to define virus stability in a number of conditions (see Methods). Virus BJ2 titres remained stable up to 50°C (roughly the temperature range over the seasons), but dropped 10 fold at 60°C. No viable particles were observed at 70°C. Virus BJ2 was stable at pH 6 to 10, but titres were significantly reduced at pH 4 by 4 logs, and were not viable at pH 2. Maximal titres were observed at salt concentrations of 2 – 3 M NaCl, but there was a 10 fold loss observed at 1 M and 5 M NaCl. In the absence of NaCl, only 0.1% of the virus population remained viable. These results are illustrated in Figure 52.

A



B



C

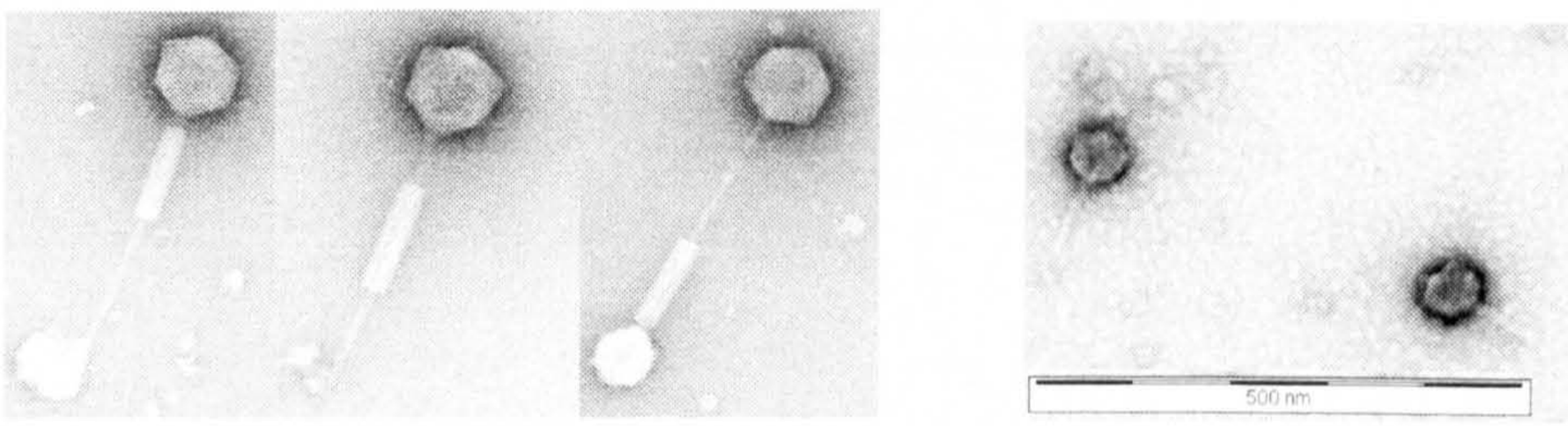


Figure 51 Electron Microscopy and Schematic Diagrams of Virus BJ1 and BJ2.

A: Electron micrograph images of viruses BJ1 and BJ2 at the same scale, the scale bar is 200 nm.

B: Schematic diagrams of viruses BJ1 and BJ2 annotated with discernable features and size of these features averaged from a number of different images and preparations.

C: Electron micrograph images of virus BJ2 showing sheath movement along the inner tail (left). Further examples of virus BJ1 are shown on the right.

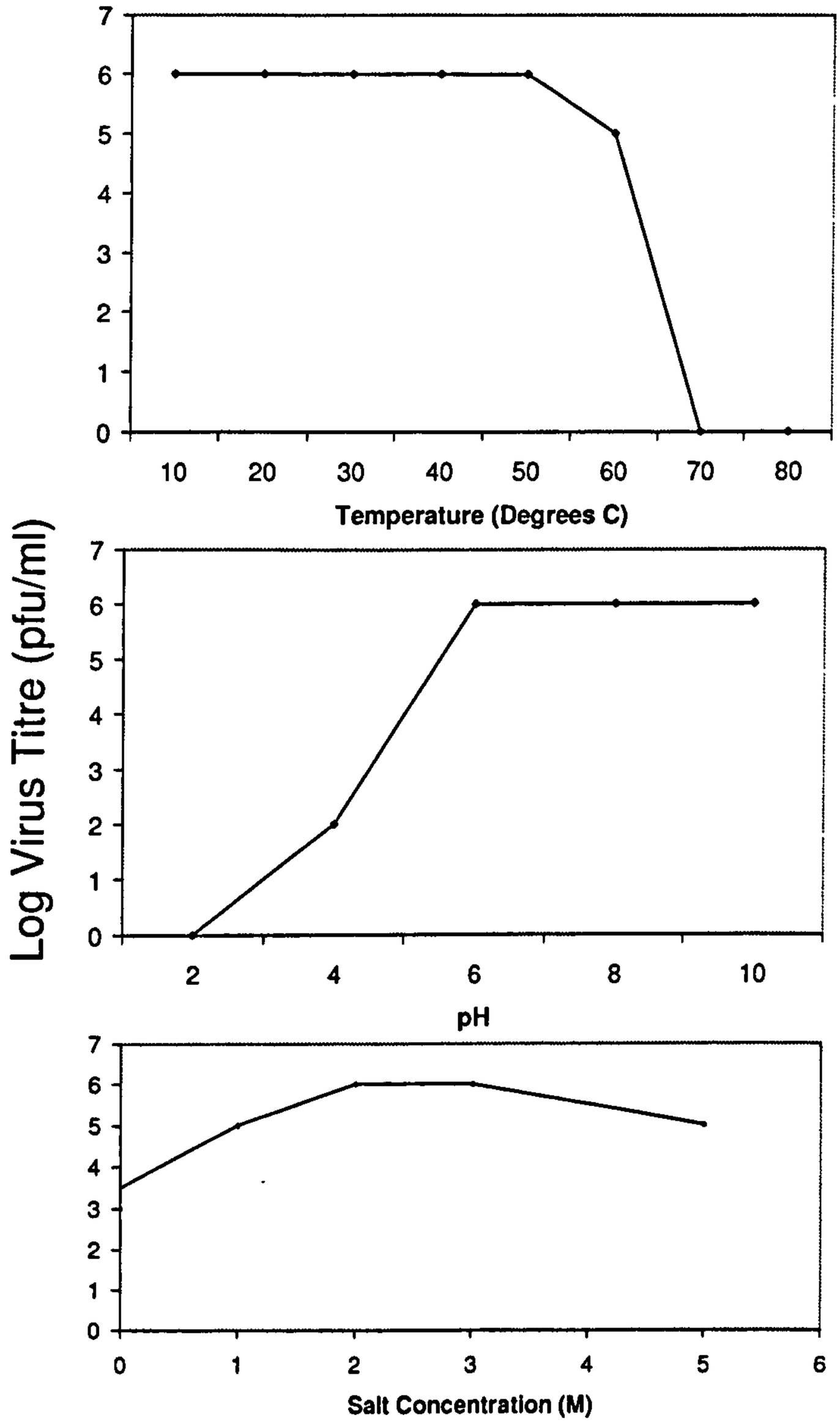


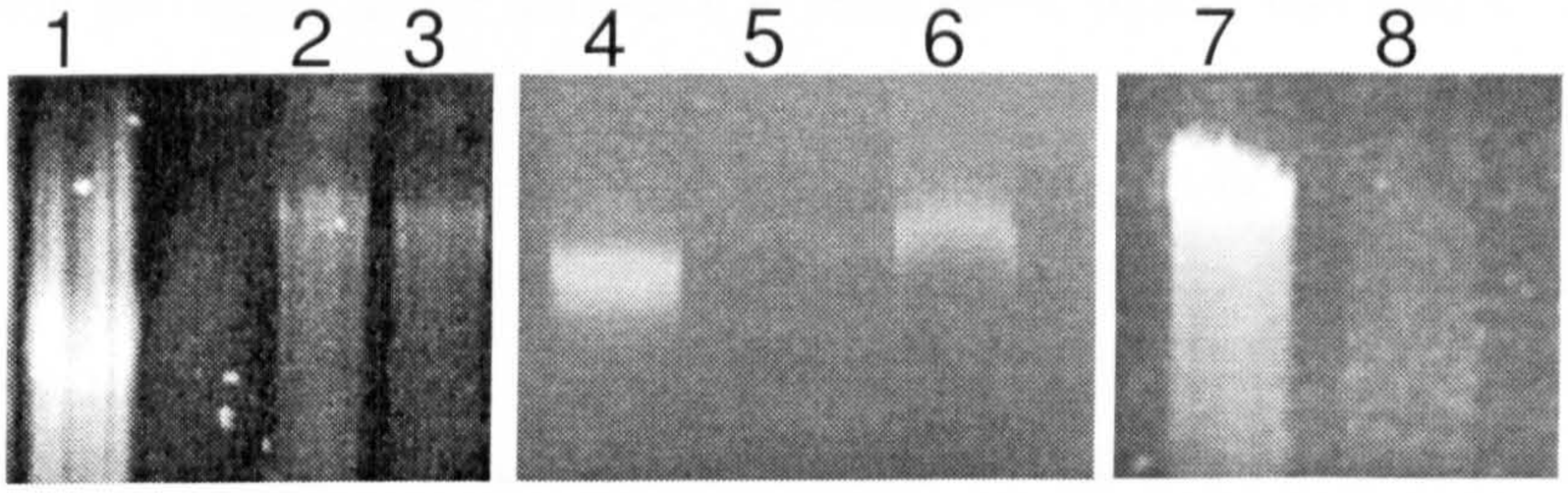
Figure 52 Stability Tests for Virus BJ2.

Virus BJ2 stability experiments with variations in temperature (top panel); pH (middle panel) and NaCl concentration (bottom panel).

Linear dsDNA Virus Genomes

Twenty plates containing plaques were sufficient for viral nucleic acid extraction. Approximately 1 – 2 µg was extracted for virus BJ1 and typically between 25 and 200 ng/µl was extracted for virus BJ2 (data not shown). The sizes of the genomic nucleic acids were difficult to determine. When both genomes were run on 1.2% TAE agarose, single bands were produced that were larger than the 23 kb DNA marker band (see Figure 53, left panel). PFGE indicated a genomic size of 23 – 48 kb for virus BJ1 (data not shown), but a size could not be determined for virus BJ2.

Various nucleases were used to characterise the genomes, which gave identical results for both viruses. Negative control experiments, where genomes were incubated with just reaction buffer and water, showed that no virus associated nucleases were responsible for the degradation observed in these experiments. The middle and right panel of Figure 53 shows the results for virus BJ1 only. Lanes 4 and 7 show undigested genome controls, lane 5 shows that the genome was sensitive to DNase digestion and lane 6 shows that the genome was insensitive to RNase. Treatment with different restriction endonucleases (*Bam*HI, *Sst*I and *Xho*I for virus BJ1 and *Hae*III, *Msp*I, *Hind*III, *Hinf*I, *Sac*I, *Sma*I for virus BJ2) gave various sized bands indicative of double stranded genomes (see later). Exonuclease III, specific for linear dsDNA substantially degraded both genomes, but failed to cut circular double stranded plasmid DNA. These results indicate that both haloarchaeal viruses harbour linear dsDNA genomes.



Key:

1. 23 kb DNA marker band
2. Virus BJ2 genome
3. Virus BJ1 genome
4. Uncut viral genome
5. DNase treated viral genome
6. RNase treated viral genome
7. Uncut viral genome
8. Exonuclease III treated viral genome

Figure 53 Characterising Viral Genomes.

Lanes1 and 2 in the 0.8% TAE agarose gel show that genomes of viruses BJ1 and BJ2 are larger than the 23 kb marker band. Lanes 4 and 7 are undigested controls of virus BJ1 genome; lane 5 shows DNase treated genome; Lane 6 shows RNase treated genome; Lane 8 shows exonuclease III treated genome. Identical results were obtained with the BJ2 genome.

Genome Sequencing – General Strategy

Since sequencing projects require large concentrations of DNA, isothermal amplification of genomic DNA was done for virus BJ2. This typically produced 1 – 2 μg DNA (data not shown). This was not required for virus BJ1 genomic DNA.

A variety of restriction endonucleases were used to cut genomic DNA to give different restriction patterns. Endonucleases that cut infrequently were used to construct a scaffold library, which provided a 'frame' on which to build up on. *Bam*HI was used for the scaffold library for virus BJ1 and *Hind*III was used for virus BJ2. Furthermore, the *Bam*HI restriction pattern allowed sizing of the genome, which was estimated at of 42.72 kb for virus BJ1 (Figure 55). Significant degradation of genomic DNA was observed when *Hae*III and *Msp*I were used on virus BJ1 genome, but not when used on virus BJ2 DNA, indicating a higher GC content in the former, as these enzymes recognise GC palindromes. This observation was later confirmed (see below). These enzymes were therefore not appropriate for constructing libraries for the virus BJ1 genome. Subsequent libraries using other endonucleases (see previously) were made to generate smaller overlapping fragments to add to the scaffold.

Clones with inserts larger than 1 kb were selected for sequencing. Sticky-ended ligation often allowed genomic fragments to ligate to other genomic fragments, producing non-contiguous inserts. These were identified in the sequence analysis and spliced into their appropriate sections before assembled into contigs. Likewise, isothermal amplification produced non-contiguous fragments by deletions, particularly at regions of high GC. Unfortunately these were only identified once a significant number of contiguous sequences were assembled and were found where there were large conflicts in the consensus sequence. This genome sequencing strategy is illustrated in Figure 54.

Primer walking was done across large *Bam*HI and *Hind*III fragments once many overlapping fragments were inserted into the assembly. Finally, gaps were closed using PCR. Primers were designed at the ends of the contigs and then different primer pairs were used in a PCR using viral DNA as a template. These primers are listed in the appendix (Appendix 3, Appendix 4, Appendix 5, p359-362). The average genome coverage for virus BJ1 was approximately 4.

Sequencing the ends of linear genomes is problematic since they cannot be cloned and many have proteins or stem-loop structures associated with them, which makes sequencing reactions difficult. However, primers designed around virus BJ1 genome allowed amplification all around the molecule and so the termini of the genome were never established (see below).

Assessment of Isothermal Amplification

Isothermal amplification is an extremely sensitive reaction that uses Phi29 DNA polymerase, and so all reactions were carried out under sterile conditions. Amplified DNA products are expected in reactions that do not have template DNA, but it is guaranteed that this product will not clone (see Methods). Despite this, some of this negative control DNA product was cloned into pGEM T-Easy after digestion with *Hae*III. Six clones were selected and the DNA fragments were sequenced. Two of the sequences had significant matches in the database to bacteria; *Serratia marcescens* and *Burkholeria thailandensis*. *Serratia* could be a laboratory contaminant. The other four sequences had insignificant matches to human DNA, thermophilic bacteria, amoeba and red flour beetle.

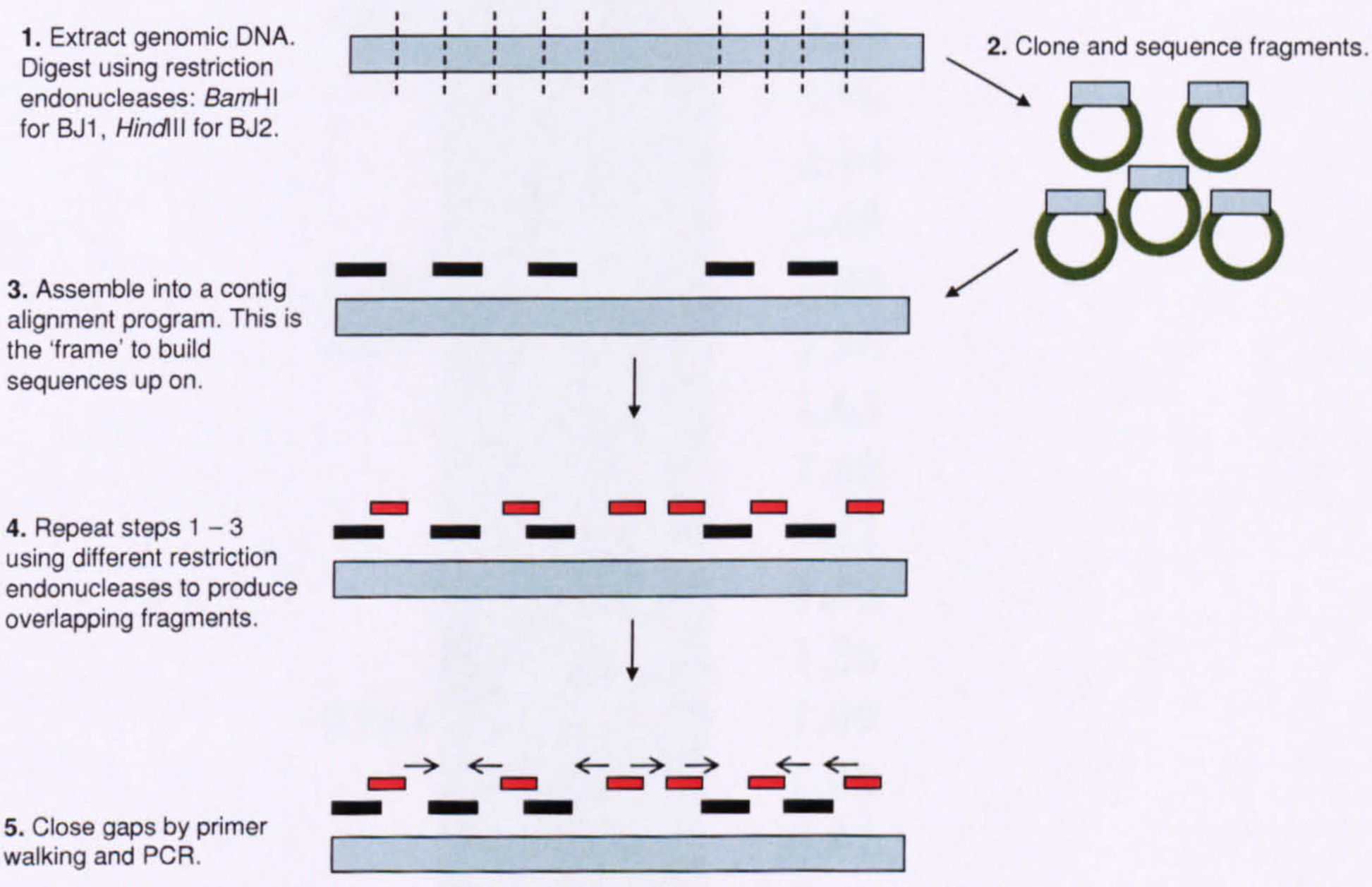
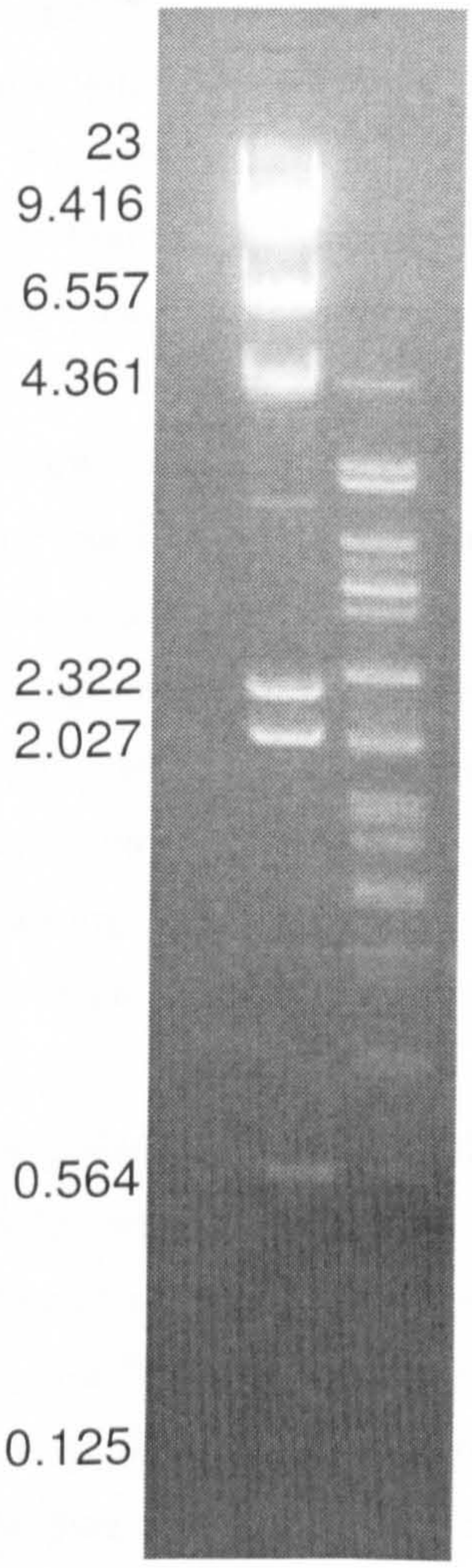


Figure 54 Strategy for Sequencing Viral Genomes.



Fragment size (kb)
4.70
4.40
3.80
3.51
3.19
2.96
2.44
2.05
2.00
1.96
1.63
1.60
1.51
1.32
1.28
1.09
1.00
0.82
0.80
0.57
0.09
Total 42.72

Figure 55 *Bam*HI Digest of Virus BJ1 Genome.

This 1.2% TAE agarose gel shows *Bam*HI restriction fragments of virus BJ1 genome. The sizes of the DNA marker bands are shown on the left in kb. The sizes of genomic fragments are listed on the right (including those that were not clear from the photograph). The total sizes of the fragments give genome BJ1 an estimated size of 42.72 kb.

Genome Sequence of BJ1

The double-stranded genomic DNA isolated from virus particles yielded a circular sequence 42271 bp long with a G+C content of 64.8 mol% Figure 56 (EMBL Accession AM419438). In conjunction with the exonuclease III susceptibility showing that the DNA is linear, this result indicates that the genome is terminally redundant (and may be circularly permuted), and so the sequence termini will not be obvious. It is unclear if the BJ1 genome ever forms a circular molecule but if it does then *cos* sites are unlikely to be involved as digests with three infrequent cutting restriction enzymes (*Hind*III, *Eco*RV and *Eco*RI) followed by melting at 80°C failed to show any change in the number of bands compared to un-melted digests (data not shown). It is unclear if the genome termini are modified in any way. In the absence of an obvious end for the genome, the cumulative GC skew of the sequence was analysed (Figure 57). Skew minima and maxima often represent initiation and termination points of DNA replication in prokaryotes and viruses with a cumulative increase in skew related to the direction of replication and transcription (Grigoriev 1999). A clear maximum was observed at about 43000 followed by a sharp change with the minima from 1-8000. This in conjunction with the Open Reading Frame (ORF) map and pattern of operons was used to designate a +1 start of the genome, (see Figure 56). The cumulative GC skew is consistent with the reading direction of most ORFs and a rolling circle pattern of DNA replication.

A single tRNA for phenylalanine (GAA anticodon recognising a UUC codon) was identified using the tRNAScan-SE program. Potential ORFs were assigned using the programs FGENESB and GeneMark.hmm v2.5a (set for prokaryotes); these analyses predicted 63 and 66 ORFs, respectively, encoding polypeptides larger than 30 amino acids. Regions upstream and downstream of these predicted ORFs were further analysed for putative ribosome binding sites (RBS), and overlapping start and stop codons, and several additional ORFs were found. BLAST searches using the amino acid sequences of all predicted ORFs were used to differentiate between possible genes e.g. ORFs 5 and 6 have significant matches (see below), so putative ORFs in the opposite strand with no BLAST matches have been discounted. By combining all of the data it was concluded that BJ1 probably contains 70 ORFs (Figure 56 and Table 43) [If ORFs greater than 60 amino acids in size are counted then the number of ORFs drops to 55] Taking the upper estimate of 70 gives an ORF density of 1.65 /kb. This is fairly close to the figure of 1.7 ORFs/kb observed for other archaeal virus genomes (Bamford *et al.* 2005). The majority of the ORFs have initiation codons of ATG (62) and the rest are GTG (8).

The Shine/Dalgarno sequence (AGGAGGTGA) from *Halorubrum saccharovorum* 16S rRNA sequence (Accession HSU17364), which is the closest phylogenetic match to the virus' host was used to search 5-15 bp upstream of each putative start site for the presence of putative ribosome binding sites. 51 of the 70 ORFs had sequences suggestive of a RBS, (see Table 43). One particular stretch of 6 predicted genes (ORF43 – ORF48) showed no obvious RBSs at all. A lack of a RBS for some genes is not surprising as archaeal transcription/translation is a mosaic of prokaryotic and eukaryotic mechanisms and the first gene of an operon, or a singly transcribed gene often lacks a RBS (Bell and Jackson 1998; Tolstrup *et al.* 2000; Satorius-Neef and Pfeifer 2004).

The majority of the ORFs (59/70) had a low calculated isoelectric point (pI < 5), which is similar to the acidic proteins of haloarchaeal organisms (Mongodin *et al.* 2005) and haloviruses (Witte *et al.* 1997; Bamford *et al.* 2005; Bath *et al.* 2006). Just three small ORFs (less than 74 amino acids) were predicted to be extremely basic (pI > 10). No ORFs larger than 100 amino acids had a pI above 6.3. Sixty three ORFs and the tRNA are coded on one strand (designated forward) and 7 are on the reverse strand. One gene, ORF 30 (13255 - 14700 bp) overlaps entirely with another, ORF31 (13270 - 14487 bp), running in the opposite direction. It seems probable that both ORFs are coding, ORF30 because it overlaps with the start and stop codons of the ORFs before and after it i.e. 29 and 32, and has a good consensus RBS, and ORF31 because it shows significant homology to integrases (see below).

BJ1 ORF analysis

BLAST (BLASTP and TBLASTN) and PSI-BLAST were used to search for possible homologies to known proteins, or proteins predicted by translation of the unannotated DNA sequence in GenBank during December 2006 and January 2007. BLASTN analysis of the whole virus genome showed significant matches (E 10^{-9} to 10^{-4}) to small segments of several haloarchaeal sequences i.e. *Natronomonas pharaonis*, *Halobacterium* sp. NRC-1 and *Haloarcula marismortui*. BLASTX analysis identified four regions of the genome having significant matches to database proteins either from haloviruses or haloarchaea, discussed below. The putative ORFs were individually analysed using BLASTX and BLASTP. InterPro was also used to search for functional domains. Using these approaches no match or function could be ascribed to 50 of the 70 ORFs. Of the 20 significant matches, most were haloarchaeal virus entries or haloarchaeal entries. Of these 20, four were matches to database

entries with no identifiable function, i.e.: ORF9 shows 59% identity (E 10^{-8}) to ORF 58 of halovirus φCh1 (Accession AAM88732); ORF10 has 54% identity (E 10^{-5}) to a protein from *Haloquadratum walsbyi* (Accession CAJ52235); this protein may contain a SWIM zinc finger capable of mediating DNA or protein interactions; ORF17 shows similarity (E 10^{-13}) to a protein from *Natronomonas pharaonis* (Accession CR936257.1). ORF55 shows similarity (E 10^{-7}) to a protein of φCh1 (Accession NP665930.1). One ORF, 24, was suggested to be a DNA binding protein by InterPro but does not have a significant match to any described protein.

The remaining 15 ORFs could have functions tentatively ascribed to them on the basis of amino acid similarity. They can be placed into three groups. (1) Those probably concerned with DNA replication, gene expression and possibly integration, i.e. ORFs 5, 6, 16, 20, 21, 31, 35, 39 and 43. ORF5 has some similarity (E 10^{-3}) to bacterial proteins with a DnaJ-like domain and may be involved in DNA replication. ORF6 has 65% similarity (E 10^{-6}) to protein Vng6287h of *Halobacterium* sp. NRC-1 (Accession AAG20925) belonging to a family of metal regulated homodimeric repressors with a 'winged helix' DNA binding domain. ORF16 is probably the DNA polymerase and has 60% identity (E 10^{-67}) to a *Haloarcula marismortui* protein (Accession YP_136906); this is part of the CDC-6 superfamily of NTPases involved in DNA replication, recombination, and repair (COG1474.2). ORF20 has 54% identity (E 10^{-13}) to the halovirus φH1 repressor (Accession AAV47198.1), a possible Mn-dependent repressor protein with a winged helix DNA binding domain (COG1321). ORF21 has 66% identity (E 10^{-17}) to the *Haloquadratum walsbyi* PadR transcriptional regulator (Accession CAJ51359.1) belonging to pfam03551. ORF31 shows significant similarity to a number of integrases from archaeal and prokaryotic species from the λ bacteriophage recombinase family, pfam00589; the highest similarity is to a *Haloarcula marismortui* phage integrase with 45% identity (E 10^{-66}) (Accession AAV47153). ORF35 is probably a DNA helicase and has 62% identity (E 10^{-128}) to *Haloarcula marismortui* protein (Accession AAV47142) that is part of the Cdc-46/Mcm family of DNA dependent ATPases (COG1241.2). ORF39 has 68% identity (E 0.05) to the ArsR-like transcriptional regulator (Accession CAJ51299) from *Haloquadratum walsbyi.* ORF43 shows 56% identity (E 10^{-37}), to a halovirus HF1 protein (Accession AAO61337.1) which may be a YonJ like, small subunit of the DNA polymerase, (COG1311).

(2) Those proteins likely to be involved in virus assembly, i.e. ORFs 48, 49, 50, 52 and 53. ORF48 shows 54% identity to the central region of *Listonella pelagia* phage phiHSIC small terminase subunit (Accession YP_224235.1). ORF49 has similarity with a number of large terminase proteins, the C-terminal region (228 - 430) shows strongest similarity, 43% identity (E 0.01) to *Streptococcus pneumoniae* bacteriophage EJ-1 large terminase (Accession CAE82121). ORF50 shows 54% identity (E 10^{-77}) to the putative portal protein (NP_665924) of *Natrialba magadii* virus φCh1, pfam04860, COG5518. ORF52 shows 49% identity (E 10^{-13}) to the capsid protein gpD (AAM88683) of halovirus φCh1. ORF53 shows 48% identity (E 10^{-29}) to hp32 (Accession CAA56442) of *Halobacterium salinarum* virus φH and 47% identity (E 10^{-24}) to the capsid protein gpE (Accession AAG32163) of halovirus φCh1.

(3) Those proteins with other identifiable functions, i.e. ORF1, which shows homology to LPS biosynthesis proteins with 51% identity (E 10^{-15}) to an *Enterococcus faecium* glycosyl transferase (Accession EAN10921), COG1442.

Nucleotide Features

Nine direct repeats were observed with greater than 13 nucleotides; the largest was 17 nucleotides, i.e. GGCGGCATCCAACTCGG repeated at positions 34076 and 34120. All of the repeats were located in putative ORFs and no significance can be inferred from them. A number of perfect and imperfect inverted repeat/stemloop structures were identified, often having loops 100s - 1000s of nucleotides in size. One perfect palindrome is located at nucleotides 14226GTCCGCTGGA/TCCAGCGGAC14247 in ORF31, the putative integrase gene. Another palindrome separated by 3 nucleotides (lower case) is 42048ACTATCCGACtggGTCGGATAGT42070; again both are present in putative ORFs and their significance is unclear although the last palindrome is located 209 nucleotides from the 3' end of the genome.

Sequence Heterogeneity

*Bam*HI digests of virus DNA gave rise to a fragment of about 3.5 kb, as judged by agarose gel electrophoresis, present in sub-stoichiometric amounts relative to the other bands. This was fully sequenced and found not to fit into the genome assembly. Primers derived from this sequence were used with virus sequence primers and virus genomic DNA as a template. Products were observed with primers derived from the 3' end of gene 32, suggesting that a minor subfraction of virus DNA did contain this *Bam*HI fragment. Sequencing showed that

the site of insertion was at nucleotide 14790 in gene 32 and showed that this part of gene 32 was rich in CGX repeats, (See Table 44). No PCR products were derived defining either the location or 5' end of the insertion/substitution. Instead primer walking was done out from the defined 3' end of the insertion. Approximately 8.7 kb of sequence has been determined having a G+C content of 72.6 mol%, notably higher than the 64 mol% determined for the rest of the virus genome and close to that reported for *Halorubrum saccharovorum* (71 mol%). Predicted ORFs have much higher homologies to known haloarchaeal proteins than the other viral genes, see Table 44. It is most likely this sequence is derived from the host genomic DNA due to an integration/excision event.

Table 43 ORFs and Proposed Genes for Halovirus BJ1.

ORF	Position[a]		D[b]	No. of residues	Mr[c]	pI[d]	RBS sequence[e] (RBS distance[f])	Functional assignment
	Start	Stop						
1	130	990	+	286	33000.0	4.58	-	glycosyl transferase
2	1146	1805	+	219	24797.9	4.94	-	
3	1980	2093	-	37	3711.5	8.50	GGAGGTG (-5)	
4	2207	2425	+	72	8099.2	6.96	-	
5	2541	3191	-	216	24075.2	4.66	GAGG (-10)	DnaJ domain protein
6	3178	3393	-	71	8155.9	4.34	-	metal regulated repressor
7 V	3547	3993	+	148	15888.1	4.07	GAG (-6)	
8	3993	4463	+	156	17665.2	4.25	AGGAGGTGA (-8)	
9	4456	4851	+	131	14528.9	4.22	AGGAGGTGA (-7)	similar to PhiCh1 ORF58
10	4844	5218	+	124	13690.3	4.72	GGAGGT (-6)	similar to *Haloarcula* protein
11	5208	5357	+	49	5171.4	3.80	GAGGTG (-8)	
12	5350	5574	+	74	8261.1	4.56	AGGAGGT (-6)	
13	5571	5744	+	57	6098.9	10.39	GGAGG (-5)	

[a]Numbers refer to the BJ1 genome sequence

[b]D = Direction of translation

[c]Mr = Molecular mass

[d]pI = Calculated isoelectric point

[e]Putative 5' upstream ribosome binding site (RBS) sequence complementary to the 3' end of *Halorubrum saccharovorum* 16S rRNA sequence (exact match AGGAGGTGA)

[f]The distance of the RBS from the initiation codon calculated from the last nucleotide of the determined RBS motif.

Table 43 Continued.

ORF	Position[a]		D[b]	No. of residues	Mr[c]	pI[d]	RBS sequence[e] (RBS distance[f])	Functional assignment
	Start	Stop						
14	5741	5986	+	81	8947.0	5.84	GGAGG (-8)	
15	5998	6417	-	139	15371.3	4.27	GAGG (-7)	
16	6637	7713	+	358	40040.8	4.96	AGGTG (-9)	DNA polymerase
17	7919	8560	-	213	23637.8	4.28	AGGA (-8)	
18	8689	8949	+	86	9328.6	4.85	-	
19	8950	9153	+	67	7903.8	5.21	GGTG (-10)	
20	9159	9446	-	95	10617.7	4.60	GGAG (-4)	similar to PhiH1 repressor
21	9660	10022	+	120	13739.9	4.58	GGA (-7)	similar to PadR-like transcriptional regulators
22	10022	10153	+	43	4602.9	4.02	GGTG (-8)	
23	10153	10890	+	245	26691.8	3.91	GGAGG (-8)	
24	10880	11806	+	308	34331.4	4.33	GGAGG (-9)	
25 V	11803	11946	+	47	5206.6	4.14	GGTGA (-7)	

[a]Numbers refer to the BJ1 genome sequence

[b]D = Direction of translation

[c]Mr = Molecular mass

[d]pI = Calculated isoelectric point

[e]Putative 5' upstream ribosome binding site (RBS) sequence complementary to the 3' end of *Halorubrum saccharovorum* 16S rRNA sequence (exact match AGGAGGTGA)

[f]The distance of the RBS from the initiation codon calculated from the last nucleotide of the determined RBS motif.

Table 43 Continued.

ORF	Position[a]		D[b]	No. of residues	Mr[c]	pI[d]	RBS sequence[e] (RBS distance[f])	Functional assignment
	Start	Stop						
26	11946	12671	+	241	27096.7	4.68	GGTGA (-7)	
27 V	12668	12760	+	30	3278.9	4.53	GGAGGTG (-6)	
28	12757	13092	+	111	12160.5	5.75	GAGGTGA (-5)	
29	13092	13262	+	56	6225.5	3.78	GGAGG (-8)	
30	13255	14700	+	481	51625.0	6.15	AGGAGG (-6)	
31	13270	14487	-	405	45959.9	4.96	-	Integrase
32	14701	14826	+	41	4326.8	4.02	GGAGGTGA (-9)	
33	14819	15307	+	162	17892.6	4.57	GAGGTGA (-7)	
34	15310	15531	+	73	8339.6	11.61	AGGAGGTG (-9)	
35	15489	17603	+	704	78350.6	4.71	(GAAAA)	Helicase
36	17606	18058	+	150	16621.2	4.44	GGAGG (-9)	
37	18055	18519	+	154	17457.1	4.30	(GGGGG)	
38 V	18512	18817	+	101	11264.4	5.04	GAGGTG (-8)	restriction alleviation protein

[a]Numbers refer to the BJ1 genome sequence

[b]D = Direction of translation

[c]Mr = Molecular mass

[d]pI = Calculated isoelectric point

[e]Putative 5' upstream ribosome binding site (RBS) sequence complementary to the 3' end of *Halorubrum saccharovorum* 16S rRNA sequence (exact match AGGAGGTGA)

[f]The distance of the RBS from the initiation codon calculated from the last nucleotide of the determined RBS motif.

Table 43 Continued.

ORF	Position[a]		D[b]	No. of residues	Mr[c]	pI[d]	RBS sequence[e] (RBS distance[f])	Functional assignment
	Start	Stop						
39 V	18814	19073	+	86	9917.0	6.05	GAGGTG (-9)	transcriptional regulator
40 V	19071	19241	+	56	5937.6	10.26	GGAGG (-8)	
41 V	19129	19806	+	225	25860.6	6.27	-	
42	19803	19982	+	59	6420.9	3.96	GAGGTG (-6)	
tRNA-phe	19973	20046	+	-	-	-	-	tRNA-phe
43	20365	21843	+	492	55057.3	4.94	-	similar to YonJ
44	21840	21998	+	52	5903.4	4.26	-	
45	22001	22111	+	36	3901.2	4.79	-	
46	22108	22416	+	102	11197.8	4.35	-	
47	22416	22577	+	53	6140.6	4.24	-	
48	22574	23083	+	169	18470.5	4.31	-	terminase small subunit
49	23080	24423	+	447	50106.2	4.92	GAGG (-8)	terminase large subunit

[a]Numbers refer to the BJ1 genome sequence

[b]D = Direction of translation

[c]Mr = Molecular mass

[d]pI = Calculated isoelectric point

[e]Putative 5' upstream ribosome binding site (RBS) sequence complementary to the 3' end of *Halorubrum saccharovorum* 16S rRNA sequence (exact match AGGAGGTGA)

[f]The distance of the RBS from the initiation codon calculated from the last nucleotide of the determined RBS motif.

Table 43 Continued.

ORF	Position[a]		D[b]	No. of residues	Mr[c]	pI[d]	RBS sequence[e] (RBS distance[f])	Functional assignment
	Start	Stop						
50	24427	26382	+	651	73227.1	4.47	-	phage portal protein
51	26461	26586	+	41	4439.6	4.40	GAG (-9)	
52	26590	27933	+	447	47210.4	3.93	AGGAGG (-9)	capsid protein
53	27949	29031	+	360	39640.3	4.16	GTGA (-8)	capsid protein
54	29040	29219	+	59	6346.8	3.82	GAGGTGA (-4)	
55	29222	29572	+	116	12437.3	3.91	-	similar to PhiCh1 ORF12
56	29576	30451	+	291	32631.7	4.56	GGAGG (-9)	
57	30444	30761	+	105	11392.4	4.06	-	
58	30758	31210	+	150	17362.2	4.75	AGG (-10)	
59	31207	31734	+	175	19569.4	4.46	GGAGGT (-5)	
60 V	31766	32680	+	304	32401.7	3.77	GAGGTGA (-7)	
61	32680	33177	+	165	17484.6	4.03	AGGAGGTGA (-8)	
62	33281	34408	+	375	37904.1	4.14	-	

[a]Numbers refer to the BJ1 genome sequence

[b]D = Direction of translation

[c]Mr = Molecular mass

[d]pI = Calculated isoelectric point

[e]Putative 5' upstream ribosome binding site (RBS) sequence complementary to the 3' end of *Halorubrum saccharovorum* 16S rRNA sequence (exact match AGGAGGTGA)

[f]The distance of the RBS from the initiation codon calculated from the last nucleotide of the determined RBS motif.

Table 43 Continued.

ORF	Position[a]		D[b]	No. of residues	Mr[c]	pI[d]	RBS sequence[e] (RBS distance[f])	Functional assignment
	Start	Stop						
63	34444	34731	+	95	10261.5	4.82	-	
64	34771	35439	+	222	23954.8	4.02	-	
65	35446	36633	+	395	42306.2	4.08	TGA (-7)	
66	36634	38226	+	530	51806.3	3.73	AGGAGGTG (-10)	glycine rich protein
67	38229	40979	+	916	100113.4	4.04	GGAGGTG (-15)	type III fibronectin domain
68	41059	41400	+	113	12162.1	3.75	GGAG (-6)	similar to Phi-H hp20
69	41403	41843	+	146	15854.3	4.57	AGGTG (-9)	
70	41840	42151	+	103	11124.1	3.94	GGTGA (-4)	

[a]Numbers refer to the BJ1 genome sequence

[b]D = Direction of translation

[c]Mr = Molecular mass

[d]pI = Calculated isoelectric point

[e]Putative 5' upstream ribosome binding site (RBS) sequence complementary to the 3' end of *Halorubrum saccharovorum* 16S rRNA sequence (exact match AGGAGGTGA)

[f]The distance of the RBS from the initiation codon calculated from the last nucleotide of the determined RBS motif.

Table 44 Predicted ORFs in *Bam*HI Fragment.

Predicted ORFs on the sequence inserted into gene 32 and their highest BLASTX matches. Nucleotide numbering is from the 5' end of the insertion sequence: nucleotide 8685 corresponds to nucleotide 14790 in the BJ1 genomic sequence. The sequence at the site of insertion was tgctcggtcgtcaa/CGACGCCGACGACGGCGA; lowercase variant, upper case virus BJ1 ORF 32. ORFs are in the forward direction with respect to the virus genome unless indicated by a '–' sign.

ORF	Position Start	Stop	Size (aa)	Homologs (% Identity)
V10*	1	277	*	67% - ornithine cyclodeaminase *Natronomonas pharaonis* DSM 2160
V9-	749	351	132	36% -hypothetical protein VNG6157H *Halobacterium* sp. NRC-1
V8-	1910	843	355	70% - cell division protein pelota *Natronomonas pharaonis* DSM 2160
V7-	3051	1936	169	28% - hypothetical protein NP4342A *Natronomonas pharaonis*
V6	3346	3753	135	38% - hypothetical protein rrnAC2062 *Haloarcula marismortui*
V5	3912	4685	257	38% - Alpha/beta hydrolase fold protein *Ralstonia eutropha* JMP134
V4	4747	5058	103	75% - hypothetical protein HQ2797A *Haloquadratum walsbyi* DSM 16790
V3-	7408	5900	502	73% - RtcB-like protein 1 *Natronomonas pharaonis* DSM 2160
V2-	7934	7503	143	61% - hypothetical protein NP3986A *Natronomonas pharaonis* DSM 2160
V1*	8326	8684	119*	64% - 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMG-CoA reductase) *Haloferax volcanii*
32*	8685	9059	*	100% - Phage BJ1 hypothetical protein

*indicates a truncated ORF because of incomplete sequencing (V10) or the insertion event itself (V1 and ORF 32).

aa indicates the number of amino acids.

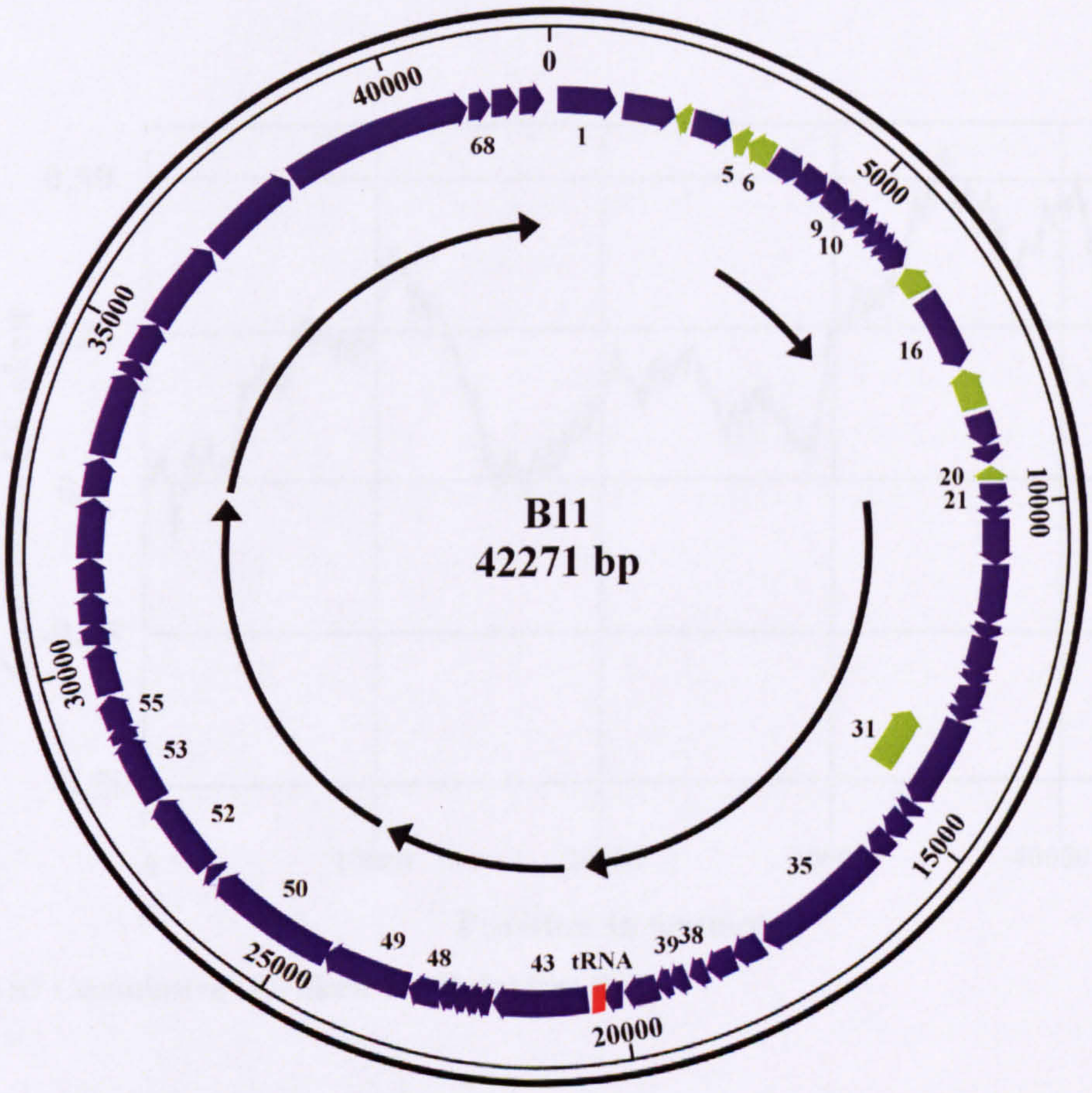


Figure 56 Genome of Halovirus BJ1.

The diagram of the virus BJ1 genome is drawn in circular form. The major features are shown including the predicted ORFs, blue arrows in the forward direction, green arrows in the reverse. The tRNA gene is in red. ORFs mentioned in the text are numbered. The outer scale bar is in base pairs. The inner curved arrows indicate putative operons.

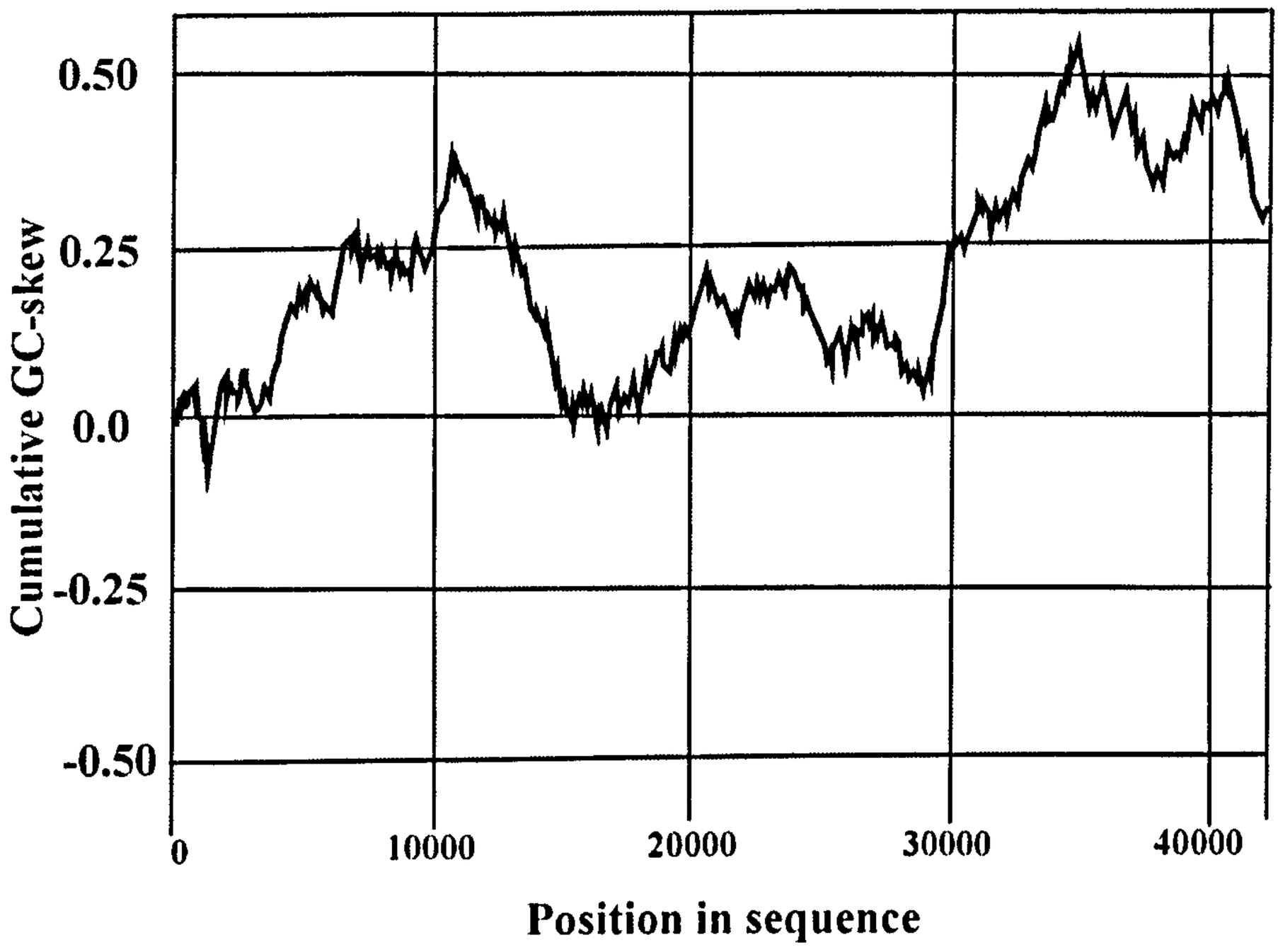


Figure 57 Cumulative GC Skew for Halovirus BJ1.

Genome Sequence of Virus BJ2

The genome sequence for virus BJ2 is incomplete. Thus far, 97602 nucleotides of the genome have been sequenced, which are split into 42 contigs, the largest of which is 10564 nucleotides long (Appendix 2, p357). Twenty eight contigs are larger than 1000 nucleotides in length and account for 86490 nucleotides of sequence (89% total). Statistical analysis suggests that these contigs account for approximately 90% of the total genome sequence (R. Haigh, personal communication).

BLASTN and BLASTX analysis showed no homology with virus BJ1 genomic sequence. A number of putative ORFs and domains, listed in Table 45 and Table 46, were identified using BLASTX, ORF finder and BLASTP. Significant matches were made to halophilic phage, including HF1, HF2, ΦCh1 and Halophage AAJ-2005, and to halophilic archaea, including *Natronomonas pharaonis*, *Haloquadratum walsbyi*, *Halobacterium* NRC-1 and *Haloarcula marismortui*. A number of significant matches were also made to other methanogenic archaea (including *Methanosphaera stadtmanae* DSM 3091, uncultured methanogenic archaeon and *Methylococcus capsulatus* str. Bath), and thermophilic archaea (including *Pyrococcus horikoshii* OT3, *Pyrococcus furiosus* DSM 3638 and *Thermococcus kodakarensis* KOD1). Moreover, a number of matches were made to bacteria including *Escherichia coli*, *Rhodobacter sphaeroides* 2.4.1, *Prochlorococcus marinus* subsp. *marinus* str. CCMP1375, *Haemophilus influenzae* 86-028NP and *Desulfotomaculum reducens* MI-1. In addition, significant matches were made to bacterial phages (*Vibrio* phage KVP40 and *Lactococcus* phage 712). Thus the genome of virus BJ2 appears to be a mosaic.

ORFs for a DNA polymerase, DNA ligase, DNA recombinase, ribonucleoside diphosphate reductase α and β chains, thymidylate synthase, thymidylate kinase, terminase, tape measure and head decoration proteins were found. Putative domains were detected for many of these ORFs, including Thy1, TMPK, CDC48_N, AAA and DNA_ligase_A_M. A number of ORFs did not have any significant matches. A full list of contigs is in Appendix 2.

Twenty four tRNA genes were predicted, including four for Leu and Asp; three for Arg; two for Gly, Val and Ser; and one for the other amino acids except Ala, Phe, Glu, His, Met, Cys, and Trp, which were absent. A G/C nucleotide was present in the anticodon wobble position in fourteen of these tRNAs, which expands decoding capability of tRNAs; as might be expected in an AT rich virus growing in a GC rich host. Two major operons were predicted,

one run of at least nine tRNA genes and another of at least ten tRNA genes is coded for in just 1994 nucleotides and 3486 nucleotides respectively.

Palindromes

The BJ1 genome contained three of each of the palindromes CTAG and GATC. However there were 36 of each of the related palindromes CGAG and GCTC. All palindromes were evenly spaced in the genome. In contrast, there were no CTAG and GATC palindromes present in 90% BJ2 genome. However, there were 13 of each of the related palindromes CGAG and GCTC present.

Table 45 Proposed Genes for Halovirus BJ2.

Homologues of ORFs detected in virus BJ2 genome were found using BLASTP.

Contig	% Identity (Expect)	Homologue (BLASTP)	Putative Domains Detected
1	42 (E 10^{-50})	Hypothetical protein VNG1325C [*Halobacterium* sp. NRC-1]	Thy1 domain
	42 (E 10^{-50})	Thymidylate synthase thyX (TS) (TSase)	
1	23(E 10^{-07})	Putative membrane protein of prophage CP-933X [*Escherichia coli*)	
	23 (E 10^{-07})	Putative tail fiber protein [*Escherichia coli* O157:H7 str. *Sakai*]	
1	38 (E 10^{-05})	Conserved hypothetical protein [*Aurantimonas* sp. SI85-9A1]	DUF1130 Domain
8	79 (0.0)	Ribonucleoside-diphosphate reductase , alpha subunit 1 [*Natronomonas pharaonis* DSM 2160]	RNR_1 Domain
8	76 (E 10^{-150})	Ribonucleoside-diphosphate reductase , beta subunit 1 [*Natronomonas pharaonis* DSM 2160]	RNRR2 Domain
8	39 (E 10^{-33})	Thymidylate kinase [*Halobacterium* sp. NRC-1]	TMPK Domain
8	50 (E 10^{-42})	Deoxycytidine triphosphate deaminase [*Methanosphaera stadtmanae* DSM 3091]	Dcd Domain
8	38 (E 10^{-05})	Conserved hypothetical protein [*Aurantimonas* sp. SI85-9A1]	DUF1130 Domain
8	49 (E 10^{-17})	Predicted MazG family pyrophosphatase [*Prochlorococcus marinus* subsp. *marinus* str. CCMP1375]	
8	33 (0.001)	Unknown [Virus PhiCh1]	
8	28 (0.37)	Ribonucleoside reductase alpha chain [*Halophage* AAJ-2005]	
14	29 (E 10^{-04})	Conserved hypothetical protein [uncultured methanogenic archaeon]	
14	43 (E 10^{-04})	Hypothetical protein HF2p086 [Halovirus HF2]	

Table 3 Continued.

Contig	% Identity (Expect)	Homologue (BLASTP)	Putative Domains Detected
61	60 (0.0)	AAA-type ATPase (transitional ATPase homolog) [*Haloquadratum walsbyi* DSM 16790]	CDC48_N Domain AAA Domain SpoVK Domain
	59 (0.0)	Cell division control protein 48 [*Haloarcula marismortui* ATCC 43049]	
61	51 (E 10^{-73})	Hypothetical protein rrnAC0170 [*Haloarcula marismortui* ATCC 43049]	Band_7_1 Domain
61	34 (E 10^{-12})	DNA ligase [*Vibrio* phage KVP40]	DNA_ligase_A_M Domain
61	48 (E 10^{-12})	146aa long hypothetical protein [*Pyrococcus horikoshii* OT3]	
61	47 (E 10^{-17})	Hypothetical protein HF2p028 [Halovirus HF2]	
73	29 (0.010)	Tail tape measure protein [*Lactococcus* phage 712]	

Table 46 Proposed Genes for Halovirus BJ2.

Homologues of ORFs detected in virus BJ2 genome were found using TBLASTX.

Contig	% Identity (expect)	Homologue (TBLASTX)
6	22 (E 10^{-05})	Terminase, ATPase subunit [*Haemophilus influenzae* 86-028NP]
7	32 (E 10^{-04})	Hypothetical protein RSP_6185 [*Rhodobacter sphaeroides* 2.4.1]
9	32 (E 10^{-10})	Hypothetical protein RB2654_04994 [*Rhodobacterales* bacterium HTCC2654]
17	46 (E 10^{-08})	Prophage MuMc02, head decoration protein, putative [*Methylococcus capsulatus* str. Bath] (late)
19	26 (E 10^{-10})	Hypothetical protein HalHV1gp084 [Halovirus HF1]
30	47 (E 10^{-16})	Hypothetical protein HF2p028 [Halovirus HF2]/ hypothetical protein HalHV1gp023 [Halovirus HF1]
45-35 inc	56 (E 10^{-22})	Hypothetical protein PF0131.2n [*Pyrococcus furiosus* DSM 3638]
48	40 (E 10^{-05})	DNA repair and recombination protein RadA [*Thermococcus kodakarensis* KOD1]
52	30 (E 10^{-101})	DNA polymerase B1 [*Halobacterium* sp. NRC-1]
58	28 (E 10^{-07})	Hypothetical protein DredDRAFT_0393 [*Desulfotomaculum reducens* MI-1]
66	30 (E 10^{-05})	Hypothetical protein HF2p084 [Halovirus HF2]
67	34 (E 10^{-04})	Unknown [Virus PhiCh1]
80	20 (0.003)	Hypothetical protein ERGA_CDS_05370 [*Ehrlichia ruminantium* str. *Gardel*]

DISCUSSION: CHARACTERISATION OF TWO NOVEL HALOARCHAEAL VIRUSES

Assessment of Isothermal Amplification

The yields of genomic DNA obtained from virus BJ2 were not sufficient for cloning and sequencing. Yields were increased by isothermal amplification, which uses Phi29 DNA polymerase. It is known that DNA products are obtained even in the absence of a template due to the sensitive nature of this procedure, which are said to be artefacts such as primer-derived multimers and should not clone (see Methods). However, despite using sterile conditions, the DNA product in the negative control produced cloneable sequences, but they (mostly) had no significant matches to the database. This therefore places a doubt over the origin of some of the sequenced genomic DNA, particularly those that do not have any significant matches to the database. However, it may be safe to assume that the concentrations of amplified contaminating DNA would be negligible compared to viral genomic DNA and so genomic DNA would be preferentially cloned. The only way to resolve this would be to complete the BJ2 genome, and any 'leftover' DNA can be assumed to be contaminant. Most sequences from amplified DNA were stripped from the analysis since there were sequences from unamplified DNA overlapping them.

Another flaw with this procedure is the formation of non-contiguous fragments. Massive deletions were detected in the amplified products, particularly at regions of high GC, which had to be edited during contig assembling.

Isothermal amplification, also called multiple displacement amplification, has been used to characterise microbial populations in contaminated sediments. It was found that certain regions of genomic DNA was over-amplified and therefore over represented in the libraries (Abulencia *et al.* 2006). Indeed, it is clear that certain contigs have a relatively high coverage, which may be due to this over-amplification (Appendix 2, p357).

It is for these reasons that isothermal amplification does not appear to be appropriate for sequencing whole genomes. Perhaps it is more appropriate for amplifying rare genes in the environment, though care must still be taken when interpreting the sequences.

Classification and Nomenclature of Viruses BJ1 and BJ2

Morphological criteria used for virus classification is outlined by the International Committee for Taxonomy of Viruses (van Regenmortel *et al.* 2000). Both virus BJ1 and BJ2 are icosahedral head/tailed viruses (Figure 51, p315) and as such are assigned to the order *Caudovirales* with examples infecting members of both the domains *Bacteria* and *Archaea.*

The techniques used to prepare the samples for TEM caused osmotic shock to the virus particles, which was apparent from the presence of dissociated head from tails, and collapsed heads on all preparations. This has been observed by Nuttall and Dyall-Smith (1993) and suggested that this was attributed to removal of the supporting salt solution and drying down of the stain. It was suggested that longer absorption times and longer fixation times with glutaraldehyde prevented this (Nuttall and Dyall-Smith 1993).

If it is assumed that both viruses BJ1 and BJ2 underwent osmotic stress, then the tails of BJ1 must not be able to contract. The lack of a contractile tail and base plate, and the presence of striations in the tail fibre would assign virus BJ1 to the Bradley classification group B and might be tentatively be assigned to the family *Siphoviridae*.

The tail of BJ2 is surrounded by a contractile outer sheath. Unfortunately, there are no pictures of the sheath in a non-contracted state. However, the features of this virus are similar to a Bradley group A virus and suggest that it belongs to the family *Myoviridae*, despite lacking a base plate and tail fibres.

The size of the BJ1 virus particles is 56 nm: 71nm (head: tail) (Figure 51) and so its length is comparable to other known haloviruses of the same classification, though its dimensions distinguishes them from those previously reported (Dyall-Smith *et al.* 2003). The size of the BJ2 virus particles is 96 nm: 238 nm (head: tail), which makes it the largest known halovirus of the same classification to date; it is over twice as long as HF1 and HF2 (Nuttall and Dyall-Smith 1993; Dyall-Smith *et al.* 2003).

Growth Conditions

Virus BJ2 produced plaques quite clearly and reliably, BJ1 did not and further refinements in culture conditions will be required to study the stability and host range of this virus Figure 50, p313). Virus BJ2 was able to grow on two related *Halorubrum* isolates from Lake

Bagaejinnor but it could not form plaques on the closely related type species *Halorubrum saccharovorum.* To date, there are only two other haloviruses known to infect *Halorubrum*: HF2 infects *Halorubrum saccharovorum* and *Halorubrum coriense* and SH1 infects a natural *Halorubrum* isolate (Nuttall and Dyall-Smith 1993; Porter *et al.* 2005). If host ranges reflect different cell adhesins carried on the virus particle (Tang *et al.* 2004), then presumably viruses BJ1, BJ2, HF2 and possibly SH1 have similar cell adhesins.

The stability experiments (Figure 52, p316), showed that virus BJ2 is stable up to 50°C; solar radiation intensity in summer in the Mongolian steppes is high and temperatures in shallow lakes and pools could easily reach such temperatures. The virus is also stable to a wide range of pH, from 6 - 10 and high NaCl concentrations, up to 5 M which is close to saturation and the conditions in which we found it. Hence, there is clearly a need for high NaCl concentration, however, the necessity for $MgCl_2$ (required by some haloviruses, such as SH1 (Porter *et al.* 2005)) was not tested. This tolerance to high salinities would be expected since all haloarchaeal species grow at high salinities (Grant *et al.* 2001). The combination of moderate stability to temperature and haloalkaline conditions may make virus BJ2 a useful model for developing biosensing activities or nanoparticles at high pH and alkalinity.

GC Content and tRNAs

The GC content of viruses BJ1 and BJ2 are quite distinct, 65 and 51mol% respectively. The GC content of *Halorubrum* sp. vary from about 63 – 71 mol% (Grant *et al.* 2001; Cui *et al.* 2006). The value for *Halorubrum aidingense* is 64 mol% and the value for *Halorubrum lacusprofundi* is 65.3 - 65.8 mol%, which are similar to virus BJ1. (The reported value for *Halorubrum saccharovorum* of 71 mol%). Other members of the *Halobacteriales* also have GC-rich DNA (Grant *et al.* 2001) hence, the GC content of BJ1 is comparable to that of other haloviruses.

Clearly, the GC content of BJ2 is distinctly different. However, it is comparable to that of viruses HF1 and HF2, which have a GC content of 55.8 mol% (Nuttall and Dyall-Smith 1993; Nuttall and Dyall-Smith 1995). Furthermore, His1 and His2 genomes are approximately 40 mol%, which is about 20% lower than that of their host *Haloarcula hispanica* (Bath *et al.* 2006). However, it was shown that His1 and His2 were poorly adapted to grow in its host and that no virus-encoded tRNA genes were discovered to compensate for the differences in codon usage between virus and host, hence it was postulated that these viruses had alternative

hosts (Bath *et al.* 2006). In contrast, virus BJ1 had a striking number of tRNA genes coded for by the virus (twenty four found so far versus one for virus BJ1). In fact, that is probably the most number of tRNA genes found encoded in a halovirus to date. This indicates the virus has not recently occupied this strain as a host, but has at least partially compensated for the GC rich codon usage of its host by substantial tRNA gene acquisition. Moreover, virus BJ2 had a narrow host range, suggesting that it has had time to adapt to this particular host.

BLASTN analysis of the first tRNA operon (9 tRNA genes in 1994 nucleotides) gives no significant matches. However the second operon, 10 genes in 3486 nucleotides has significant matches (~ E 10^{-5}) to the euryarchaeotes *Methanothermus fervidus*, *Methanosphaera stadtmanae* and *Methanothermobacter thermautotrophicus*. The first, most significant match is to a tRNA gene in a cluster of 5 tRNA genes from an organism isolated from an Icelandic hot spring. This supports the idea that BJ2 has acquired these genes from an archaeal host.

Palindromes

Both genomes appear to avoid tetrameric palindrome sequences, particularly CTAG and GATC. Only three of each of these palindromes were present in virus BJ1 genome, and so far they are completely absent from the virus BJ2 genome. Such sequences are expected to arise every 256 nt, but they appear to be selected against by many haloviruses. For example, these palindromes are absent from the genomes of HF1, HF2, His2 and SH1 (Bath *et al.* 2006). This selection pressure is thought to be due to the avoidance of restriction-modification systems in the host cells (Bickle and Krüger 1993), and there is evidence that CTAG and GATC palindromes are used by haloarchaeal systems (Holmes *et al.* 1991; Allers and Mevareeh 2005). Hence there are relatively more of the related palindromes CGAG and GCTC present in both BJ1 and BJ2 genomes.

BJ1 ORFs

Cumulative GC skews have been used to find the origin of replication in many prokaryotes; the lowest point indicates the origin of replication and the highest point indicates termination (Grigoriev 1998; Grigoriev 1999), which helped in deciding the beginning and end of the BJ1 genome (Figure 56, p335). The ORF arrangement is similar to that of other viruses of both *Bacteria* and *Archaea* in that they are divided into distinct modules: genes relating to establishing cell infection are in the early region, genes relating to DNA synthesis are in the middle region, and genes relating to virus assembly and cell lysis are located in the late region

(Hendrix *et al.* 1999; Brussow and Desiere 2001). No significant matches were made to the early genes in virus BJ1, but other haloviruses do not have close database matches either (Tang *et al.* 2002). The suggestion of operons indicated in Figure 56 is entirely speculative and based on the presence of overlapping stop and start signals; one run of ORFs from 43 - 48 has no RBS at all. Proteins with putative functions involved in DNA replication and transcription are found in ORFs 1 - 43, putative structural proteins are found after ORF48 consistent with early and late expression of operons.

Virus BJ1 genome is 42271 bp long and is therefore comparable to other halovirus genome lengths (Dyall-Smith *et al.* 2003). Of the ORFs identified in BJ1 described in the Results (Table 43, p328), all of the statistically significant matches are recorded. Six of the ORFs (9, 20, 50, 52, 53, 55) are most closely related to the haloarchaeal temperate, isometric viruses φCh1 (Klein *et al.* 2002) and the intensively studied, φH (Stolt and Zillig 1994). These two viruses are closely related to each other, the completed genome of φCh1 shows 97% homology to the genome of φH, which is about 60% complete. ORF 43 is most closely related to a gene from the haloarchaeal isometric virus HF1. There are no similarities with the ORFs from either the spindle (His1, His2) or icosahedral (SH1) shaped haloarchaeal viruses. The most significant matches were ORFs 16, 31, 35, which are almost certainly the polymerase, integrase and helicase functions respectively of the virus, having highly significant matches to full length proteins in *Haloarcula marismortui.* ORF50 was also closely related to the putative portal protein (NP_665924) of *Natrialba magadii* virus φCh1.

Speculatively, almost all ORFs are in the forward strand in the same direction consistent with a rolling circle mechanism of DNA replication. The 7 ORFs on the reverse strand including the integrase may be poorly expressed. A few proteins had GTG starts (but with good RBS sequences) and some proteins lacked RBS sequences altogether, presumably both coding features control/reduce expression levels. The fact that the putative integrase gene is coded for on the minor strand with no RBS, and that it overlaps with gene 30 on the major strand may indicate that its expression is tightly controlled; perhaps most infections are lytic with a small proportion of lysogenic events. Plaques were clearly turbid, though there is no direct experimental evidence of a lysogenic life cycle.

BJ2 ORFs

BJ2 genome length so far is 97602 bp, which makes it the largest of known halovirus genomes (Dyall-Smith *et al.* 2003; Bamford *et al.* 2005), except for Ja1, which was estimated by reassociation kinetics and not by sequencing (Wais *et al.* 1975). Although not fully sequenced, it has some interesting properties.

Many ORFs have been detected that could be assigned to early, middle and late regions (Table 45, p339 and Table 46, p341). Genes related to establishing cell infection (early region) include those involved in nucleotide metabolism. Homologues of ribonucleoside-diphosphate reductase alpha and beta subunits were found that are part of putative domains RNR_1 and RNRR2, respectively. They catalyse the conversion of ribonucleotides to deoxyribonucleotides, which provides the precursors necessary for DNA synthesis. A homologue of ribonucleotide reductase alpha chain was also found, which also provides precursors for DNA synthesis. In addition, a homologue of thymidylate kinase was found that is part of putative domain TMPK, which is involved in catalysing the phosphorylation of thymidine monophosphate (TMP) to thymidine diphosphate (TDP) utilizing ATP as its preferred phophoryl donor. In addition, a pyrophosphatase was found. All these genes are located in contig 8 (Table 45), which again suggests a modular arrangement of genes in viruses. Genes that relate to DNA synthesis are found in the middle region. Homologues of ATPase were found that is associated with a wide variety of cellular activities, including membrane fusion, proteolysis, and DNA replication. This ATPase and a DNA ligase were found in contig 61 (Table 45). Other genes involved in DNA synthesis include DNA repair and recombination protein (contig 48) and DNA polymerase group B (contig 52). Genes have also been found involved in virus assembly. Homologues of a tape measure protein, a head decoration protein (capsid protein) and a tail fibre protein have been found in contigs 73, 17 and 1 respectively, and so speculatively could be located in the late region. A homologue of a thymidylate synthase that is part of the putative domain THY1 was also located in contig1, and is involved in complementing the thymidine growth requirement of the virus.

Many of these ORFs obtained significant matches with homologues in *Archaea*, *Bacteria*, bacterial viruses as well as haloviruses (HF1, HF2, ΦCh1 and Halophage AAJ-2005). The genome of virus BJ2 thus shows surprising mosaicism. However, this has also been observed for halovirus HF2. Subsequently, it was postulated that a stepwise movement and evolution of genes between halophiles with different salt optima allowed genes from non-halophiles to

eventually be transferred to haloviruses. The existence of a clear progression from marine to moderate to extremely halophilic species and the fact that the bacterium *Salinibacter ruber* coexists with halophilic archaea (see previously) provides sufficient bridges to account for the gene transfer seen in both HF2 (Tang *et al.* 2002) and presumably virus BJ2. Moreover, it is well documented that tailed bacteriophages have access to a common genetic pool by horizontal exchange (Hendrix *et al.* 1999) and that there is extensive horizontal gene transfer between *Bacteria* and *Archaea* (Mongodin *et al.* 2005), suggesting that haloarchaeal viruses would share many genes with bacteriophages.

Sequence Heterogeneity

Although virus BJ1 stocks are clonal in origin, the genomic DNA preparation is obviously and necessarily derived from a virus pool. Genome sequence projects often therefore give rise to heterogeneous sequences. We found one substantial region of heterogeneity in gene 32 at nucleotide 14790 (Table 44, p334) involving either a large insertion or more probably a substitution event (since terminally redundant virus genomes usually package genomes in a 'head full' mechanism). To distinguish between these possibilities requires more sequencing. The variant sequence probably involves the acquisition of host derived DNA since the GC content is higher (72.6%) than that of the virus (64.8%) and close to that reported for *Halorubrum saccharovorum* (71%). Obviously this insertion/substitution has taken place about 300 nucleotides away from the putative integrase gene. The integrase gene in viruses is often the site of insertion as well. We speculate that this variant sequence in the virus population is the result of an integration/excision event (possibly aberrant) during the virus infection to prepare genomic DNA. Whether the virus population with this variant sequence is viable will require further studies. Certainly virus populations with insertions and or substantial genomic deletions can be viable or at least rescued by functional virus genomes.

The Need for New Screening Technologies

As described in the Introduction, very few viruses infecting the domain *Archaea* have been described and as yet we have little idea as to the extent of virus diversity in this domain. Sixteen head tail archaeal viruses have been reported so far, although diversity studies have shown that this morphology is relatively rare in environments where archaea dominate (Oren *et al.* 1997a; Rice *et al.* 2001). Hence, viruses described here are not likely to be common or dominant members of the virus community infecting haloarchaea in saline waters. Lytic viruses were screened for on archaeal lawns. These requirements for host culturability, good

lawn formation and plaquing are extremely restrictive and therefore extremely biased. In addition the screening conditions used standard halophile media not closely related to either the chemical or nutritional conditions found in the lake. There is a genuine need to develop other isolation and culture techniques to study both the dominant virus populations and the true extent of archaeal virus variation in samples such as these - perhaps using a combination of electron microscopy and metagenomic sequence studies.

Evolution of Archaeal Viruses

If we assume that this classification is phylogenetically justified then it could indicate that the *Caudovirales* originated before the divergence of the *Bacteria* and *Archaea* (Zillig *et al.* 1996). An alternative explanation is that the *Caudovirales* originally infected members of the domain *Bacteria* but that horizontal gene exchange from mesophilic *Bacteria* to the *Archaea* and the subsequent stabilisation of these genes in the *Archaea* allowed the *Caudovirales* to spread into the domain *Archaea* [Certainly there is a diverse bacterial community in the water of Lake Bagaejinnor] (Prangishvili *et al.* 2006a). Sequences such as those reported here should help in resolving these possibilities.

CONCLUSION

The microbial ecology of hot springs in Rehai geothermal field in Yunnan Province, China and salt lakes in Inner Mongolia, China was investigated. The hot springs were located in Tengchong and Lang Pu, and were mostly near neutral to alkaline hot springs, at temperatures ranging from 52°C to 94°C. The salt lakes located in Inner Mongolia are athalassohaline in origin, formed in closed drainage basins by the leaching of surrounding rocks. Most were alkaline and hypersaline, apart from Lake Chagannor, which was highly alkaline with a pH of 10.5 and a slightly lower salinity than the other salt lakes.

The microbial ecology was investigated by molecular analysis of the 16S rRNA sequences from community DNA. The laminated mat exposed to temperatures between 60 – 65°C was found alongside a deep basin of steaming and bubbling water and consisted of layers of cyanobacteria, filamentous anaerobic phototrophs (*Chloroflexi*), thermoanaerobic fermenters, other heat adapted bacteria (*Nitrospirae*, *Thermales*, *Chlorobia*). Previous studies on similar mats have shown that the cyanobacteria and filamentous anaerobic phototrophs are located in the top layers (the photic zone) and are responsible for primary production, and the thermoanaerobic fermenters located in the lower layers utilise organic substrates from primary producers. The laminated mat in this study is likely to have a similar structure since it contained distinct layers of microbes. The bacterial community was highly diverse, comparable to that of the human mouth and gut. It also contained a less diverse archaeal community, consisting of methanogenic archaea (*Methanobacteriales*) and *Crenarchaeota.* However, the microbial mat structure observed at a higher temperature site was clearly different. The white streamer microbial mat found at the 'Hydrothermal Outbreak Pool' was located in a shallow stream at 72.1°C, and consisted predominantly of filamentous sulphur bacteria belonging to the *Aquificae*. These were shown to be phylogenetically distinct from known *Aquificae* from other streamer communities. They have been shown to form bundles with elemental sulphur particles, hence their filamentous appearance. It also contained a highly diverse archaeal community consisting of members of the *Crenarchaeota*, *Euryarchaeota* and *Korarchaeota*. It has been suggested that the *Aquificae* are the main primary producers in the mat and that aerobic sulphur and hydrogen oxidising bacteria are found at the surface and sulphate reducers inhabit the darker undermass. Statistical analyses supported the observation that the bacterial communities between the two hot spring environments are completely different, while the archaeal communities show some similarity.

Lake Bagaejinnor, Lake Shangmatala and the unnamed salt lake near Xilin Hot are near neutral hypersaline lakes. There was evidence for the presence of the eukaryote *Dunaliella*, which is likely to be responsible for primary production in the lake. There was also evidence for the presence of the anoxygenic phototrophs belonging to the *Ectothiorhodospiraceae*, which would also contribute to primary production. Cyanobacteria were also present, but these are only significant in primary production after heavy rainfall. One macroscopic eukaryote, *Artemia* was observed in the unnamed lake; this is a source of chitin deposition. Vegetation grew up to the water's edge at Lake Shangmatala, which is also a likely source of organic matter. The archaeal community consisted predominantly of archaea belonging to the *Halobacteriales* (haloarchaea), particularly *Halorubrum* sp. The bacterial community consisted predominantly of *Gammaproteobacteria* (halomonads), *Firmicutes* (*Haloanaerobiales*) and *Bacteroidetes*.

The salterns at Lake Ejinnor and Lake Erliannor were significantly different from each other. The salterns at Lake Ejinnor were red, while those at Lake Erliannor were white, suggesting that they harbour distinct populations. The saltern at Lake Ejinnor consisted of a climax population of bacteria related to *Salinibacter ruber* and of various haloarchaea. Hence, it has been shown that *Salinibacter* are not restricted to thalassohaline environments. No SHOW group were detected, despite the fact that they are thought to be characteristically associated with saltern brines. The community at Lake Erliannor was atypical of saltern communities, containing a far more diverse bacterial community as well as various haloarchaea. These findings confirm that salterns at different geographical locations may harbour microbial communities with different structures.

Lake Chagannor was a less hypersaline soda lake. The archaeal community consisted of haloalkaliphilic archaea belonging to the *Halobacteriales*, such as *Natronomonas* sp. The bacterial community also consisted of haloalkaliphilic species, such as *Natronoanaerobium* sp. and alkaliphilic halomonads.

No *Crenarchaeota* or known *Euryarchaeota* were detected in any of the salt lakes. However, several clones were detected that formed a lineage on the periphery of the *Halobacteriales*, suggesting that adaptation to high salinities may extend to *Archaea* not of this order. Clones from the libraries were also related to bacteria found in hot springs and fresh water

environments, suggesting that some bacteria are more ubiquitous in nature than previously recognised.

It appears that the main factors driving microbial community composition were environmental and not geographical distance. Hence the hypothesis 'everything is everywhere, the milieu selects' can be applicable to this environment, apart from the theory that 'everything is everywhere.' In particular, temperature, concentrations of Na and Mg ions and pH (and therefore CO_3 and HCO_3 ions) were significant. For the archaeal community, temperature, pH and concentrations of Na and Mg ions are significant in driving community composition at the genus and species level, but pH and concentration of Na ions are responsible for driving community composition at the subspecies level. This was similar for the bacterial community, though pH only affected the community composition at the species and subspecies levels. Both results are consistent with the distinct microbial community observed at Lake Chagannor, the only soda lake in the analysis.

The statistical analysis was inconclusive. However, it can be said that microbial diversity in salt lakes is relatively low as the values obtained for various indices were less than those obtained from human mouth and gut, environments that are microbially rich and diverse. In addition, the communities generally consisted of closely related lineages, consistent with the fact that only highly specialised microbes are able to survive such environments.

Few *Archaea* were cultivated from the salt lakes; most were *Halorubrum* sp., one other isolate was a *Haloarcula* sp. These were slow growing and did not appear on plates until week three. The bacteria were faster growing, and far more isolates were cultivated. Most belonged to the *Gammaproteobacteria* (halomonads, alteromonads, *Halovibrio* and *Idiomarina*) and the *Firmicutes* (*Bacillus* and *Haloanaerobium*). One isolate was affiliated to the *Bacteroidetes* and another was distantly related to the *Actinobacteria.* The fact that both the cultivated *Archaea* and *Bacteria* showed low diversity suggests that the cultivation methods employed were highly biased. Future work would include trying other cultivation methods, such as enrichment cultures to prevent the slow-growing microbes from being outcompeted by the fast-growing ones. Most isolates were identified by the first 800 bp of their 16S rRNA gene sequences. Those that showed 98% identity to known microbes are possibly new species. Hence, future work could also include further identification experiments. The proposed

minimal standards for identifying new taxa within the *Halobacteriales* have been described (Oren *et al.* 1997b).

Two novel haloarchaeal viruses were isolated from Lake Bagaejinnor. They infect a haloarchaeal strain that shows 98% identity to *Halorubrum saccharovorum*. Both are lytic; virus BJ1 forms turbid plaques of various sizes, while virus BJ2 forms small, clear round plaques. They are both head/tailed viruses. Based on morphology, virus BJ1 is assigned to the *Siphoviridae*, while virus BJ2 is assigned to the *Myoviridae*.

Virus BJ1 has a 42271 bp dsDNA genome, with a G+C content of 64.8 mol%, comparable to that of other haloviruses. It is terminally redundant and may be circularly permuted, though the termini of the genome are not clearly defined; it is not known if they are modified. Almost all ORFs are in the forward strand in the same direction consistent with a rolling circle mechanism of DNA replication. ORFs are arranged into early, middle and late regions, and showed significant similarity to ORFs from other haloarchaeal viruses and archaea. A fraction of the virus population contained extra DNA that is probably host derived since the GC content matches more closely to host DNA and predicted ORFs have much higher homologies to known haloarchaeal proteins than other viral genes. It suggests that an insertion/substitution event has taken place, and it may be a remnant of this process.

Virus BJ2 genome is linear dsDNA. It is incomplete, but 97602 bp (approximately 90% of the genome) has been sequenced, which has a G+C content of 51 mol% in 42 contigs, making it the largest of known halovirus genomes. So far, 24 tRNA genes have been detected (compared to one in virus BJ1), which is consistent with an AT-rich virus infecting a GC-rich host. It suggests that this virus has not recently occupied the host and has adapted to the GC rich codon usage of its host by substantial tRNA gene acquisition. The virus genome has been shown to be a mosaic of ORFs from haloarchaeal viruses, bacterial viruses, archaea and bacteria. No genomic identity between the viruses has been observed. Virus BJ2 has a narrow host range; it is unable to infect the type strain of *Halorubrum saccharovorum* or closely related field isolates. Virus BJ2 is stable from 4°C to 50°C, but titres drop ten fold at 60°C. Maximal titres were observed at 2 to 3 M NaCl and titres were stable at pH 6 – 10, but reduced 10000 fold at pH 4.

Many interesting features remain to be discovered about these two viruses, including patterns of transcription, protein functions, lysogenic potential, viability of the variant virus and completion of the BJ2 sequence. Assignment of protein functions to unclassified ORFs is probably easier using a virus as a model than any other genome. A systematic effort on this front will reduce the number of unclassified ORFs that metagenomic and archaeal sequencing projects so often throw up.

APPENDICES

Appendix 1 List of chemicals used in this study.

Salts	Supplier
$Na_2B_4O_7$	Fisons Scientific
NaBr	Sigma
$CaCl_2$	BDH Limited
KCl	Fisher Scientific
LiCl	Sigma
$MgCl_2.7H_2O$	Fisher Scientific
NaCl	Fisher Scientific
Na_2SO_4	Acros Organics
Na_2CO_3	Fisher Scientific
$NaHCO_3$	Fisher Scientific
$MgSO_4$	Fisons Scientific
NaOH Pellets	Fisons Scientific
K_2HPO_4	Fisons Scientific
KH_2PO_4	Fisher Scientific
$MgCl_2$	Fisher Scientific
Tri-sodium citrate	Fisher Scientific
Media	
Yeast extract	Difco or Oxoid
Casamino acids	Difco
Tryptone	Oxoid
Agar	Difco
Glucose	
Powders and Reagents	
5-bromo-4-chloro-3-indolyl-β-d glactopyronoside (X-Gal)	Melford Laboratories
Isopropyl β-D thiogalactopyranoside (IPTG)	Melford Laboratories
Tris(hydroxymethyl) aminomethane (Tris)	Fisher Scientific
Ethylene diamine tetra acetic acid (EDTA)	BDH Laboratories
30 % acrylamide mix	National Diagnostics
Sodium Dodecyl Sulphur (SDS)	Sigma, Fisher Scientific
Ammonium persulfate	Sigma
N,N,N',N'-tetramethylethylenediamine (TEMED)	Sigma
Dithiothreitol (DTT)	Sigma

Appendix 1 Continued.

Powders and Reagents	**Supplier**
Ficoll	Sigma
Polyethylene glycol (PEG) 6000	Fisher Scientific
PCR Nucleotide mix (dNTPs)	Promega
ATP	Promega
Agarose	Melford Laboratories
Tween 20	BDH Laboratories
Glycine	Fisher Scientific
3-[N-Morpholino]propane sulfonic acid (MOPS)	Sigma
Solvents	
Phenol	Fisher Scientific
Chloroform	Fisher Scientific
Diethyl ether	Fisons Scientific
Ethanol	Leicester University Chemistry Dept
Methanol	Fisher Scientific
Glycerol	Fisher Scientific
N,N'-dimethyl formamide	Sigma
Isoamyl alcohol	Fisons Scientific
Glutaraldehyde	Fisher Scientific
Acids	
HCl	Fisher Scientific
Glacial acetic acid	Fisher Scientific
Stains	
bromophenol blue	BDH Laboratories
Coomassie brilliant blue R250	BDH Laboratories
xylene cyanol FF	Bio-Rad Laboratories
Indicators	
Phenolphthalein	BDH Laboratories
bromcresol green	Fisons Scientific
methyl red	BDH Laboratories
Nucleic Acid Stains	
Ethidium bromide	Sigma
SYBR Green I	Invitrogen

Appendix 2 Contigs for Virus BJ2 Genome.

Name	Contig Length	Total Seq. Length	Number Seq.	Number(Top, Bottom)	Avg. Coverage	Cumulative Contig Length
8	10572	39569	50	(22,28)	3.75	10564
14	4161	13022	16	(11,5)	3.13	14725
80	4047	12900	17	(10,7)	3.19	18772
42	3774	10330	12	(4,8)	2.74	22546
9	3100	10081	12	(4,8)	3.25	25646
49	2701	5942	8	6,2	2.20	28347
47	2394	6516	8	5,3	2.72	30741
58	2292	4816	6	3,3	2.10	33033
40	2178	2399	3	1,2	1.10	35211
30	2049	5849	8	2,6	2.85	37260
7	1853	5466	7	3,4	2.95	39113
83	1720	5873	7	5,2	3.41	40833
6	1693	5474	8	8,0	3.23	42526
19	1823	12250	16	11,5	6.72	40936
5	1674	9524	11	7,4	5.69	45887
21	1648	1801	2	2,0	1.09	47535
48	1619	8882	11	6,5	5.49	49154
75	1576	2475	3	2,1	1.57	50730
66	1560	2488	3	2,1	1.59	52290
50	1534	3253	5	3,2	2.12	53824
69	1465	6959	10	5,5	4.75	55289
67	1342	1649	3	2,1	1.23	56631
56	1215	3059	4	3,1	2.52	57846
73	1055	2979	4	2,2	2.82	59961
38	986	2746	4	3,1	2.78	60947
17	952	1753	2	1,1	1.84	61899
59	945	1494	2	1,1	1.58	62844
45	926	926	1	1,0	1.00	63770
70	925	925	1	1,0	1.00	64695
57	766	1531	2	2,0	2.00	65461

Appendix 2 Continued.

Name	Contig Length	Total Seq. Length	Number Seq.	Number(Top, Bottom)	Avg. Coverage	Cumulative Contig Length
77	755	755	1	1,0	1.00	66216
71	752	1484	2	2,0	1.97	66968
64	731	731	1	1,0	1.00	67699
53	698	6926	10	10,0	9.92	68397
84	645	645	1	1,0	1.00	69042
62	585	585	1	1,0	1.00	69627
72	514	514	1	1,0	1.00	70141
37	336	336	1	1,0	1.00	70477
79	336	336	2	1,1	1.88	70813
61	7222	16315	20	(12,8)	2.26	78295
52	10490	43835	61	(38,23)	4.18	88785
1	8817	50526	66	(31,35)	5.73	97602
totals 42	**97602**	**317402**	**415**			

Appendix 3 Virus BJ1 PCR Primers for Gap Filling.

Primer	Sequence (5' to 3')	Tm[a]
18L	GAT GTT CCG ATG GCA GAA GCA TGT GAA GCC	64
52R	CGG CGA GGA GGC CAT GAT GCA TCG C	66
23L	CGA CGC GCC CCA CAG CAC CAG C	66
23R	CGA CAG CTC CGG GCG ATA TAC AGC C	64
29L	CGA AGG CCA GGT CAA CTA GTT GTT ATC TAC GG	64
29R	TCC GCA AGC CCG AAC TCG TTC GCG	64
07L	ACG ACG ATC GAC AAC CTG CCA CAT TCG C	64
07R	CCG AAG AGA CGG GGT ACA GCC TCC	64
39L	CGA GGG AAG AAA TGG CTC GAA CAA CCG	63
39R	GAA ATC GAT GGG ATC GAC TCC ATC TGT TCG	63
24L	TCG ACG ACG AGG GCA CCG TCC TCA CC	67
24R	ATG TCA GAC GCA CCG CCG CCA CCA CC	67
57L	CGA CGT TCC AGC CGA CGT TTG AGC C	64
57R	GAT GCC GTC CTC GAT CTG CTG GCG	64
33L	GCC AGC CAT CCA CAC CTC GTC GC	64
53R	CGA GTG ACA ACG CGT GGG ATC GCC	64
25L	TCC AGA TCC ATC AAT ACC CGT GAC GAA CG	63
25R	GCG GTG GGT GAA GAA TGA TGA ACG TCG	63
21L	GAC CAG ACC GCT TGT CTG GTC ATG G	63
21R	GAC AGA ATC AGA GAA CAC ATA TAC TGG TCA GCC	63
58-1	ACG AAG TTG TGG TTG TCA GCG	54
58-2	GCC TCC CGG TCG AGC G	56
58-3	GTC TTG GAG TCG TAC CAT GCG	56
58-4	GGA CCA GTT CGA CGA TCT ACC	56
49-1	CAG TCT CCT TGT GGA CGT CG	56
49-2	GAC CTC CCG GAC GAG CTC	57
49-3	GAC AGC GCC ATC CTC CGC GTG CTC	66
49-4	TCC GCG AGT TTC CGG AGT TTC TGC GTC C	67
49-5	CGC GAG GCG CTT CAG CG	57
49-6	AGC TGA TCG TGT TTC CGA GC	54
49-7	GCA AAC GCG GAC ACG ACG	55
49-8	AAG CTT AAG CAC CGC CAC GC	56
49-9	CGG CGA GTG GTC CGA TGG	57
26-1	GAG TTC GGT GAC GTC GAC C	55

[a]melting temperature

Appendix 3 Continued.

Primer	Sequence (5' to 3')	Tm[a]
26-2	GCG GGC TGG GCG TCC	56
26-3	ACC TCG CCC TCC GCA TCC	57
INT1	GTT AGC GAG ACC CCC GAC CTC GCG	66
INT2	TTA TTC GCG GAC GAA AAC ACG GCG CCG C	66
23-1	TCC CGG ACT CAC CGA CGG TTT ATG CGC G	67
23-2	CCG CCG TCG ACC GCG GAG AAG TCG	68
23L2	CGT CGG CGA ACC GCT CCG CGA CG	68

[a]melting temperature

Appendix 4 Virus BJ1 Primers for Primer Walking.

Primer	Sequence (5' to 3')
B-02f	GTA GTC CGA CTC CGT GG
B-04f	TAC GTC GAC GCC GGT GC
B-05f	CGC CTT GAC GGC ACC AC
B-05r	CTC GAC TGC GAG ATG CG
BF-08f	GTC GGC GGA CTT CGA GC
BF-08r	GCG AGG AGA TGG ACT TCC
BF-12f	AGT CGC ACG TCT CAG CG
BF-12r	CGA CAA CAA GGA GGT GAC G
BF-16f	AGC TCT CCT TCC TGC GG
BF-16r	CTG ATC ACC CAC CAG GC
S-02f	TGA CGA CTA CAT CCA GCC C
S-28r	AAA CTA TCC GAC TGG GTC GG
S-66f	CGA GGT CGG GAA GTG GC
S-66r	CCT CAA CGA GCA GCA GCG
S-80f	GAC GCC TGC GCC TCG
S-80r	TGG CTG TCG GGG CTC G
B-01f	CCA GAC GAC GAA CTG TCC G
B-05xf	TCG CGA TGT CGA CGC CG
BF-12xr	AAG TAC GAC CTC GCC GAG C
BF-16xf	TCG GAC TGT TCG CCA TCG G
BF-16xr	CCG GCG GCT GCA GTG C

Appendix 5 Virus BJ2 PCR Primers for Gap Filling.

Primer	Sequence (5' to 3')	Tm[a]
8L	TTT CTC GCA AGA CAC CAA CC	52
8R	AAC GAA CGT GCC GGG AGA G	55
1L	GTT GCT ACT ATG CGG CGG C	55
1R	CGA ACG TGT CTT GGT CGT GG	56
14L	ACC TCC ACC TCC ACC ACC	55
14R	GTC TAT GTA AAT GAC GCG G	49
80L	CAG TAT TCT CTG TGT CGC CC	54
80R	GGG AAT GTC TGA TGT AGA CC	52
51L	TTC CCC ACT CGG AAT AGC GC	56
51R	TCT GGC GGA TAG TTC CAC GG	56
42L	TAT AGG GTG AAT TTC GGC CC	52
42R	AAT CCA GTA GTC CAT CCA CG	52
46L	ATC GCT GAA ACC ACT TCC CG	54
46R	AGC ATA CTG GGC AAT TCG CC	54
52L	GTT GAC ATC GAG GAG GTC GG	56
52R	CGT GGA CTA TCT CTT TCG GG	54
82L	TAC GAT AGG CTC ACA CGG CC	56
82R	AAC TGC CCG GCA TAC CCG	55
28L	AAT CTC AGC GAC CTC TCC GC	56
28R	GGT CAC TCC TCT TTC CAG CG	56
61L	CTC AGG CTC GTC ATC AGT CG	
61R	GCT TCC GAT GAA CCA TCC CG	
76L	AGA TGA GGG CCG CCT CTC GG	
76R	GAA GAC GCT GAC AGA ACT GG	
78L	TAC GAA CGA GGT CGT CGC CC	
78R	TTC GTC GGG GCA CTA AC CC	
55L	CCG AGG GTC TAC GTA GAT GG	
55R	GAT TAC GCT CGC TGG TCG GG	

[a]melting temperature

Appendix 6 Data DVD ROM

A data DVD ROM has been enclosed with this thesis containing three folders of sequencing data and footage of the expeditions. The folder 'characterising microbial populations by molecular methods' contains all 16S rRNA and 18S rRNA sequences from the chapter of the same name. The folder 'characterising microbial populations by cultivation methods' contains 16S rRNA sequences of all cultivated isolates from the chapter of the same name. The folder 'Virus BJ2' contains the 42 contig sequences from haloarchaeal virus BJ2. The files 'Yunnan 2003' contain footage of the expedition to the Yunnan Province in south China in March 2003. The files 'Inner Mongolia 2003' contain footage of the expedition to Inner Mongolia, China, in September 2003. Both films can be viewed on any computer with a DVD drive using appropriate software e.g. Windows Media Player.

REFERENCES

Abulencia, C. B., D. L. Wyborski, *et al.* (2006). "Environmental whole-genome amplification to access microbial populations in contaminated sediments." Appl Environ Microbiol **72**: 3291-301.

Ackermann, H. W. (2007). "5500 Phages examined on the electron microscope." Arch Virol **152**: 227-43.

Agogué, H., E. O. Casamayor, *et al.* (2005). "A survey on bacteria inhabiting the sea surface microlayer of coastal ecosystems." FEMS Microbiol Ecol **54**: 269-80.

Alfreider, A. and C. Vogt (unpublished). Analysis of bacterial populations in benzene-contaminated groundwater, UFZ Centre for Environmental Research, Leipzig, Germany.

Allers, T. and M. Mevareeh (2005). "Archaeal genetics - the third way." Nature Rev Genet **6**: 58-73.

Altschul, S. F., W. Gish, *et al.* (1990). "Basic local alignment tool." J Mol Biol **215**: 403-10.

Andersen, G. L. (unpublished). A high-density 16S rRNA gene microarray reveals a high level of bacterial diversity in the air of two cities *in* Texas, Lawrence Berkeley National Lab, California, USA.

Andrews, K. T. and B. K. C. Patel (1996). "*Fervidobacterium gondwanense* sp. nov., a new thermophilic anaerobic bacterium isolated from nonvolcanically heated geothermal waters of the Great Artesian Basin of Australia." Int J Syst Bacteriol **46**: 265–9.

Angly, F. E., B. Felts, *et al.* (2006). "The marine viromes of four oceanic regions." PLoS Biol **4**: 368.

Antón, J., R. Rosselló-Mora, *et al.* (2000). "Extremely halophilic *Bacteria* in crystallizer ponds from solar salterns." Appl Environ Microbiol **66**: 3052-7.

Anton, J., A. Oren, *et al.* (2002). "*Salinibacter ruber* gen. nov., sp. nov., a novel, extremely halophilic member of the *Bacteria* from saltern crystalliser ponds." Int J Syst Evol Microbiol **52**: 485-91.

Arahal, D. R., W. Ludwig, *et al.* (2002). "Phylogeny of the family *Halomonadaceae* based on 23S and 165 rDNA sequence analyses." Int J Syst Evol Microbiol **52**: 241-9

Arnold, H. P., U. Ziese, *et al.* (2000). "SNDV, a novel virus of the extremely thermophilic and acidophilic Archaeon *Sulfolobus*." Virology **272**: 409-16.

Arp, G., V. Thiel, *et al.* (1999). "Biofilm exopolymers control microbialite formation at thermal springs discharging into the alkaline Pyramid Lake, Nevada, USA." Sedim

Geol **126**: 159–76.

Ashelford, K. E., N. A. Chuzhanova, *et al.* (2005). "At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies." Appl Environ Microbiol **71**: 7724-36.

Bahati, G. (2003). Geothermal energy in Uganda, country update. International Geothermal Conference, Reykjavík.

Bai, D., M. A. Meju, *et al.* (2001). "Magnetotelluric images of deep crustal structure of the Rehai geothermal field near Tengchong, southern China." Geophys. J. Int. **147**: 677-87.

Baker, G. C., S. Gaffar, *et al.* (2001). "Bacterial community analysis of Indonesian hot springs." FEMS Microbiol Lett **200**: 103-9.

Baker, G. C., J. J. Smith, *et al.* (2003). "Review and re-analysis of domain-specific 16S primers." J Microbiol Meth **55**: 541-55.

Bamford, D. H., J. J. Ravantti, *et al.* (2005). "Constituents of SH1, a novel lipid-containing virus infecting the halophilic Euryarchaeon *Haloarcula hispanica*" J Virol **79**: 9097-107.

Bano, N. and J. T. Hollibaugh (unpublished). Aerobic heterotrophs isolated from Mono Lake, California, University of Georgia, USA.

Barns, S. M., R. E. Fundyga, *et al.* (1994). "Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment." Proc Natl Acad Sci U S A **91**: 1609-13.

Barns, S. M., C. F. Delwiche, *et al.* (1996). "Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA Sequences." Proc Natl Acad Sci U S A **93**: 9188-93.

Bath, C. and M. Dyall-Smith (1998). "His1, an archaeal virus of the *Fuselloviridae* family that infects *Haloarcula hispanica*." J Virol **72**: 9392-95.

Bath, C., T. Cukalac, *et al.* (2006). "His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, Salterprovirus." Virology **350**: 228-39.

Battista, J. R. and F. Rainey. "Order I. *Deinococcales*." In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001.

Baumann, L., R. D., Bowditch *et al.* (1983). "Description of *Deleya* gen. nov. created to accommodate the marine species *Alcaligenes aestus, A. pacificus, A. cupidus, A. venustus* and *Pseudomonas marina*." Int J Syst Bacteriol **33**: 793-802.

Bel'kova, N. L., V. V. Parfenova, *et al.* (2005). "Biodiversity and activity of the microbial

community in the Kotelnikovsky Hot Springs (Lake Baikal)." Biol Bulletin **32**: 549-55.

Bell, S. D. and S. P. Jackson (1998). "Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features." Trends Microbiol **6**: 222-8.

Benlloch, S., A. López-López, *et al.* (2002). "Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern." Environ Microbiol **4**: 349-60.

Bickle, T. A. and D. H. Krüger (1993). "Biology of DNA restriction" Microbiol Rev **57**: 434-50.

Bidle, K., W. Amadio, *et al.* (2005). "A phylogenetic analysis of haloarchaea found in a solar saltern." Bios **76**: 89-96

Blum, J. S., B. A. Bindi, *et al.* (1998). "*Bacillus arsenicoselenatis* sp. nov., and *Bacillus selenitireducens* sp. nov.; two haloalkaliphiles from Mono Lake, California that respire oxyanions of selenium and arsenic." Arch Microbiol **171**: 19-30.

Bohannan, B. J. M. and J. Hughes (2003). "New approaches to analysing microbial biodiversity data." Curr Opin Microbiol **6**: 282-7.

Bolhuis, H., E. M. te Poele, *et al.* (2004). "Isolation and cultivation of Walsby's square archaeon." Environ Microbiol **6**: 1287-91.

Bonheyo, G. T., B. W. Fouke, *et al.* (unpublished). Microbial 16S rRNA diversity within travertine depositional facies from 73 degrees to 25 degrees C at Angel Terrace, Mammoth Hot Springs, Yellowstone National Park, USA, University of Illinois, USA.

Bonnet, E. and Y. Van der Peer (2002). "zt: a software tool for simple and partial Mantel tests." J Stat Software **7**: 1-12.

Boomer, S. M., D. P. Lodge, *et al.* (2002). "Molecular characterization of novel red green nonsulfur bacteria from five distinct hot spring communities in Yellowstone National Park." Appl Environ Microbiol **68**: 346-55.

Boone, D. R., R. Castenholz, *et al.* (Eds). Bergey's Manual of Systematic Bacteriology. 2nd ed. New York, Springer-Verlag, 2001.

Boone, D. R., W. B. Whitman, *et al.* "Order I. Methanobacteriales." In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001.

Bowden, S. T., M. L. Davis, *et al.* (1980). Sibling species of *Artemia*. The Brine Shrimp *Artemia*. G. Persoone, P. Sorgeloos, O. Roels and E. Jasper. Belgium, Universa Press. **1**: 155-67.

Bowman, J. P., S. A. McCammon, *et al.* (2000a). "The microbial composition of three limnologically disparate hypersaline Antarctic lakes." FEMS Microbiol Lett **183**: 81-8.

Bowman, J. P., S. M. Rea, *et al.* (2000b). "Diversity and community structure within anoxic sediment from marine salinity meromictic lakes and a coastal meromictic marine basin, Vestfold Hilds, Eastern Antarctica." Environ Microbiol **2**: 227-37.

Boyd, E. S., D. E. Cummings, *et al.* (unpublished). Analysis of subsurface microbial community structure using colonized mineral surfaces and investigation of methodological biases, Montana State University, USA.

Brandt, K. K., B. K. C. Patel, *et al.* (1999). "*Desulfocella halophila* gen. nov., sp. nov., a halophilic, fatty-acid-oxidizing, sulfate reducing bacterium isolated from sediments of the Great Salt Lake." Int J Syst Bacteriol **49**: 193–200.

Breitbart, M., P. Salamon, *et al.* (2002). "Genomic analysis of uncultured marine viral communities." Proc Natl Acad Sci U S A **99**: 14250-5.

Breitbart, M., I. Hewson, *et al.* (2003). "Metagenomic analyses of an uncultured viral community from human feces." J Bacteriol **185**: 6220-3.

Breitbart, M., L. Wegley, *et al.* (2004). "Phage community dynamics in hot springs." Appl Environ Microbiol **70**: 1633–40.

Brochier, C., S. Gribaldo, *et al.* (2005). "Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to *Thermococcales*?" Genome Biol **6**: R42.

Brock, T. D. "Introduction: An Overview of the Thermophiles." In Thermophiles: General, Molecular, and Applied Microbiology. Ed: T. D. Brock. John Wiley & Sons, Inc., 1986

Brunk, C. F. and N. Eis (1998). "Quantitative measure of small-subunit rRNA gene sequences of the Kingdom *Korarchaeota*." Appl Environ Microbiol **64**: 5064-6.

Brussow, H. and F. Desiere (2001). "Comparative phage genomics and the evolution of *Siphoviridae*: insights from dairy phages." Molec Microbiol **39**: 213-22.

Burns, D. G., H. M. Camakaris, *et al.* (2004a). "Combined use of cultivation-dependent and cultivation-independent methods indicates that members of most haloarchaeal groups in an Australian crystallizer pond are cultivable." Appl Environ Microbiol **70**: 5258-65.

Burns, D. G., H. M. Camakaris, *et al.* (2004b). "Cultivation of Walsby's square haloarchaeon." FEMS Microbiol Lett **238**: 469-73.

Burns, D. G., P. H. Janssen, *et al.* (2007). "*Haloquadratum walsbyi* gen. nov., sp. nov., the square haloarchaeon of Walsby, isolated from saltern crystallizers in Australia and Spain." Int J Syst Evol Microbiol **57**: 387-92.

Cann, A. J., S. E. Fandrich, *et al.* (2005). "Analysis of the virus population present in equine

faeces indicates the presence of hundreds of uncharacterised virus genomes." Virus Genes **30**: 151-6.

Carle, G. F. and M. V. Olsen (1984). "Separation of chromosomal DNA molecules from yeast by orthogonal-field-alteration gel electrophoresis." Nucleic Acids Res **12**: 5647-64.

Carle, G. F., M. Frank, *et al.* (1986). "Electrophoretic separations of large DNA molecules by periodic inversions of the electric field." Science **232**: 65-8.

Carrasco, I. J., M. C. Marquez, *et al.* (2006). "*Gracilibacillus orientalis* sp. nov., a novel moderately halophilic bacterium isolated from a salt lake in Inner Mongolia, China." Int J Syst Evol Microbiol **56**: 599-604.

Casamayor, E. O., R. Massana, *et al.* (2002). "Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparison of genetic fingerprinting methods in a multipond solar saltern." Environ Microbiol **4**: 338-48.

Casamayor, E. O. (unpublished-a). Microbial diversity in the salt lakes of the Monegros Desert, Acc. Cala Sant Francesc, Spain.

Casamayor, E. O. (unpublished-b). Published directly in database.

Castenholz, R. (1996). "Endemism and biodiversity of thermophilic cyanobacteria." Nova Hedwigia **112**: 33–47.

Castenholz, R. "Class I. 'Chloroflexi'". In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001a.

Castenholz, R. "Phylum BX. Cyanobacteria. " In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001b.

Castillo, A. M., M. C. Gutiérrez, *et al.* (2006a). "*Halostagnicola larsenii* gen. nov., sp. nov., an extremely halophilic archaeon from a saline lake in Inner Mongolia, China." Int J Syst Evol Microbiol **56**: 1515-7.

Castillo, A. M., M. C. Gutiérrez, *et al.* (2006b). "*Halovivax asiaticus* gen. nov., sp. nov., a novel extremely halophilic archaeon isolated from Inner Mongolia, China." Int J Syst Evol Microbiol **56**: 765-70.

Caton, T. M., L. R. Witte, *et al.* (2004). "Halotolerant aerobic heterotrophic bacteria from the Great Salt Plains of Oklahoma." Microbiol Ecol **48**: 449-62.

Caumette, P., Y. Cohen, *et al.* (1991). "Isolation and characterization of *Desulfovibrio halophilus* sp. nov., a halophilic sulfate-reducing bacterium isolated from Solar Lake (Sinai)." Syst Appl Microbiol **14**: 33-8.

Caumette, P., R. Matheron, *et al.* (1994). "Microbial mats in the hypersaline ponds of

Mediterranean salterns (Salins-de-Giraud, France)." FEMS Microbiol Ecol **13**: 273-86.

Cayol, J.-L., B. Ollivier, *et al.* (1995). "*Haloanaerobium lacusroseus* sp. nov., an extremely halophilic fermentative bacterium from the sediments of a hypersaline lake." Int J Syst Bacteriol **45**: 790–7.

Chai, L. H., X. L. Cui, *et al.* (unpublished). The microbial diversity of salt Lake Keke in western China, Yunnan University, China.

Chan, O. C., P. Casper, *et al.* (unpublished). Archaeal community in Lake Stechlin sediment, Institute of Freshwater Ecology and Inland Fisheries, Neuglobsow, Germany.

Chao, A. (1984). "Non-parametric estimation of the number of classes in a population." Scandin J Stat **11**: 265-70.

Cho, B. C., D. H. Choi, *et al.* (unpublished-a). Novel bacterium affiliated with family *Rhodobacteraceae*, isolated from a solar saltern, Seoul National University, Korea.

Cho, H. Y., J. H. Hyun, *et al.* (unpublished-b). Bacterial community structure in the intertidal flat of Ganghwa island by 16S rRNA gene analysis, Korea Ocean Research & Development Institute, Kyoung Gi, Republic of Korea.

Chookietwattana, K., S. Chen, *et al.* (unpublished). Fatty acid analysis and genotypic characterization of halophilic bacteria isolated from saline soil in Mahasarakham Province, Thailand, Suranaree University of Technology, Nakhon Ratchasima, Thailand.

Christner, B. C. Recovery of bacteria from glacial and subglacial environments. Department of Microbiology, Columbus, Ohio State University. **PhD**, 2002

Cioni, R., G. Fanelli, *et al.* (1992). "Lake Bogoria hot springs (Kenya): geochemical features and geothermal implications." J Volcanol Geothermal Res **50**: 231-46

Civetta, L., M. D'antonio, *et al.* (1998). "The geochemistry of volcanic rocks from Pantelleria Island, Sicily Channel: petrogenesis and characteristics of the mantle source region." J Petrology **39**: 1453–91.

Cohen, Y., W. E. Krumbein, *et al.* (1977). "Solar Lake (Sinai). 2. Distribution of photosynthetic microorganisms and primary production." Limnol Oceanography **22**: 609-20.

Colwell, R. K. (2005). *Estimate*S: Statistical estimation of species richness and shared species from samples. Version 7.5.

Crosby, L. D. and C. S. Criddle (2003). "Understanding bias in microbial community analysis techniques due to *rrn* operon copy number heterogeneity." BioTechniques **34**: 790-802.

Crump, B. C., C. S. Hopkinson, *et al.* (2004). "Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time." Appl Environ Microbiol **70**: 1494-505.

Cui, H. L., D. Tohty, *et al.* (2006). "*Halorubrum lipolyticum* sp. nov. and *Halorubrum aidingense* sp. nov., isolated from two salt lakes in Xin-Jiang, China." Int J Syst Evol Microbiol **56**: 1631-4.

Daffonchio, D., S. Borin, *et al.* (2006). "Stratified prokaryote network in the oxic-anoxic transition of a deep-sea halocline." Nature **440**: 203-7.

DeLong, E. F. (1992). "Archaea in coastal marine environments." Proc Natl Acad Sci U S A **89**: 5685-9.

DeLong, E. F. (2005). "Microbial community genomics in the ocean." Nat Rev Microbiol **3**: 459-69.

Demergasso, C. S., E. O. Casamayor, *et al.* (2004). "Distribution of prokaryotic genetic diversity in athalassohaline lakes of the Atacama Desert, Northern Chile." FEMS Microbiol Ecol **48**: 57–69.

Demergasso, C. S., E. O. Casamayor, *et al.* (unpublished). Distribution of prokaryotic genetic diversity in Salar de Atacama, Northern Chile, Universidad Catolica del Norte, Antofagasta, Chile.

Díez, B., C. Pedrós-Alió, *et al.* (2001). "Application of denaturing gradient gel electrophoresis (DGGE) to study the diversity of marine picoeukaryotic assemblages and comparison of DGGE with other molecular methods." Appl Environ Microbiol **67**: 2942-51.

Doemel, W. N. and T. D. Brock (1977). "Structure, growth, and decomposition of laminated algal-bacterial mats in alkaline hot springs." Appl Environ Microbiol **34**: 433-52.

Dolan, J. R. (2006). "Microbial biogeography?" J Biogeogr **33**: 199-200.

Donachie, S. P., J. P. Bowman, *et al.* (2004). "*Psychroflexus tropicus* sp. nov., an obligately halophilic *Cytophaga–Flavobacterium–Bacteroides* group bacterium from an Hawaiian hypersaline lake." Int J Syst Evol Microbiol **54**: 935–40.

Donachie, S. P., S. Hou, *et al.* (unpublished). Diverse uncultured and cultured bacteria in inland waters of remote Hawaiian islands: hypersaline Lake Laysan and a brackish pond on Pearl and Hermes Atoll, University of Hawaii, Honolulu, USA.

Duckworth, A. W., W. D. Grant, *et al.* (1996). "Phylogenetic diversity of soda lake alkaliphiles." FEMS Microbiol Ecol **19**: 181-91.

Duckworth, A. W., W. D. Grant, *et al.* (2000). "*Halomonas magadii* sp. nov., a new member of the genus *Halomonas*, isolated from a soda lake of the East African Rift Valley."

Extremophiles **4**: 53-60.

Dyall-Smith, M., S.-L. Tang, *et al.* (2003). "Haloarchaeal viruses: how diverse are they?" Res Microbiol **154**: 309-13.

Echigo, A., M. Hino, *et al.* (2005). "Endospores of halophilic bacteria of the family *Bacillaceae* isolated from non-saline Japanese soil may be transported by Kosa event (Asian dust storm)." Saline Syst **1**: 1-13.

Eder, W., W. Ludwig, *et al.* (1999). "Novel 16S rRNA gene sequences retrieved from highly saline brine sediments of kebrit deep, Red Sea." Arch Microbiol **172**: 213-8.

Eder, W., L. L. Jahnke, *et al.* (2001). "Microbial diversity of the brine-seawater interface of the Kebrit Deep, Red Sea, studied via 16S rRNA gene sequences and cultivation methods." Appl Environ Microbiol **67**: 3077-85.

Edwards, E. A. and D. Grbić-Galić (1994). "Anaerobic degradation of toluene and *o*-xylene by a methanogenic consortium." Appl Environ Microbiol **60**: 313-22.

Edwards, R. A. and F. Rohwer (2005). "Viral metagenomics." Nature Rev Microbiol **3**: 504-10.

Eichler, S., R. Christen, *et al.* (2006). "Composition and dynamics of bacterial communities of a drinking water supply system as assessed by RNA- and DNA-based 16S rRNA gene fingerprinting." Appl Environ Microbiol **72**: 1858–72.

Elshahed, M. S., J. M. Senko, *et al.* (2003). "Bacterial diversity and sulfur cycling in a mesophilic sulfide-rich spring." Appl Environ Microbiol **69**: 5609-21.

Elshahed, M. S., F. Z. Najar, *et al.* (2004). "Survey of archaeal diversity reveals an abundance of halophilic archaea in a low-salt, sulfide- and sulfur-rich spring." Appl Environ Microbiol **70**: 2230-9.

Emanuilova, E., M. Kambourova, *et al.* (1993). "Thermoalkaliphilic lipase-producing *Bacillus* selected by continuous cultivation." FEMS Microbiol Lett **108**: 247-50.

Fan, H., Y. Xue, *et al.* (2004). "*Halorubrum tibetense* sp. nov., a novel haloalkaliphilic archaeon from Lake Zabuye in Tibet, China." Int J Syst Evol Microbiol **54**: 1213-6.

Fan, H., Y. Xue, *et al.* (unpublished). Uncultured haloarchaeon from Zabuye Lake, Tibet China, Chinese Academy of Sciences, Beijing, China.

Fendrich, C., H. Hippe, *et al.* (1990). "*Clostridium halophilium* sp. nov. and *C. litorale* sp. nov., an obligate halophilic and a marine species degrading betaine in the Stickland reaction." Arch Microbiol **154**: 127-32.

Feng, J., P. J. Zhou, *et al.* (2004). "*Halorubrum xingjiangense* sp. nov., a novel halophile isolated from saline lakes in China." Int J Syst Evol Microbiol **54**: 1789-91

Ferreira, A. C., M. F. Nobre, *et al.* (1997). "*Deinococcus geothermalis* sp. nov. and *Deinococcus murrayi* sp. nov., two extremely radiation-resistant and slightly thermophilic species from hot springs." Int J Syst Bacteriol **47**: 939–47.

Ferris, M. J., S. C. Nold, *et al.* (1997). "Population structure and physiological changes within a hot spring microbial mat community following disturbance." Appl Environ Microbiol **63**: 1367-74.

Ficker, M., K. Krastel, *et al.* (1999). "Molecular characterisation of a toluene-degrading methanogenic consortium." Appl Environ Microbiol **65**: 5576-85.

Fierer, N. and R. B. Jackson (2006). "The diversity and biogeography of soil bacterial communities." Proc Natl Acad Sci U S A **103**: 626-31.

Fitch, W. M. (1971). "Toward defining the course of evolution: minimum change for a specific tree topology." Syst Zool **20**: 406-16.

Fouke, B. W., J. D. Farmer, *et al.* (2000). "Depositional facies and aqueous-solid geochemistry of travertine-depositing hot springs (Angel Terrace, Mammoth Hot Springs, Yellowstone National Park, U.S.A.)." J Sedim Res **70**: 565-85.

Franzmann, P. D., E. Stackebrandt, *et al.* (1988). "*Halobacterium lacusprofundi* sp. nov., a halophilic bacterium isolate from Deep Lake, Antarctica." Syst Appl Microbiol **11**: 20-7.

Free, R. C. Molecular analysis of the lower respiratory tract microbiota. Infection, Immunity and Inflammation, University of Leicester. **PhD**, 2005.

Friedman, S. M. "Thermophilic Microorganisms." In Encyclopedia of Microbiology. Ed: J. Lederberg. California, Academic Press, Inc., 1992.

Friedrich, U., K. Prior, *et al.* (2002). "High bacterial diversity of a waste gas-degrading community in an industrial biofilter as shown by a 16S rDNA clone library." Environ Microbiol **4**: 721-34.

Fuse, H., H. Inoue, *et al.* (2003). "Production of free and organic iodine by *Roseovarius* spp." FEMS Microbiol Lett **229**: 189-94.

Galinski, E. A. and H. G. Trüper (1994). "Microbial behaviour in salt-stressed ecosystems." FEMS Microbiol Rev **15**: 95-108.

García, M. T., E. Mellado, *et al.* (2004). "*Halomonas organivorans* sp. nov., a moderate halophile able to degrade aromatic compounds." Int J Syst Evol Microbiol **54**: 1723–8.

Garnova, E. S., T. N. Zhilina, *et al.* (2004). "Anaerobic, alkaliphilic, saccharolytic bacterium *Alkalibacter saccharofermentans* gen. nov., sp. nov. from a soda lake in the Transbaikal region of Russia." Extremophiles **8**: 309-16.

Garrity, G. and J. G. Holt. "Class VII. "Methanopyri *class. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001a.

Garrity, G. and J. G. Holt. "Phylum AI. Crenarchaeota *phy. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001b.

Garrity, G. and J. G. Holt. "Phylum AII. *Euryarchaeota* phy. nov." In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001c.

Garrity, G. and J. G. Holt. "Phylum BIII. Thermodesulfobacteria *phy. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001d.

Garrity, G. and J. G. Holt. "Phylum BIX. Deferribacteres *phy. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001e.

Garrity, G. and J. G. Holt. "Phylum BVII. Thermomicrobia *phy. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001f.

Geissler, A., T. Tzvetkova, *et al.* (unpublished). Comparison of natural bacterial communities found in uranium mining waste piles and mill tailings, Institute of Radiochemistry, Dresden, Germany.

Ghozlan, H. A., A. A. Gaballa, *et al.* (unpublished). Bacterial diversity at Wadi El-Natrun, Egypt Faculty of Science, Alexandria, Egypt.

Giggenbach, W. F., D. S. Sheppard, *et al.* (1994). "Geochemical structure and position of the Waiotapu geothermal field, New Zealand." Geothermics **23**: 599-644.

Gilvanova, E. A., S. K. Paavilainen, *et al.* (unpublished). Microorganisms tolerate to high concentration of sodium nitrite, Russian Academy of Sciences.

Glatz, R. E., P. W. Lepp, *et al.* (2006). "Planktonic microbial community composition across steep physical/chemical gradients in permanently ice-covered Lake Bonney, Antarctica." Geobiology **4**: 53-67

Glissman, K., K. J. Chin, *et al.* (2004). "Methanogenic pathway and archaeal community structure in the sediment of eutrophic Lake Dagow: effect of temperature." Microbiol Ecol **48**: 389-99.

Gomez-Alvarez, V. and K. Nuesslein (unpublished). Comparative 16S rRNA gene sequence

analysis of microbial communities on recent Hawaiian volcanic deposits (1959, 1921, 1790 & 1700): colonization and succession of microbes in extreme environments - Data 2002, University of Massachusetts, USA.

Grabowski, A., O. Nercessian, *et al.* (2005). "Microbial diversity in production waters of a low-temperature biodegraded oil reservoir." FEMS Microbiol Ecol **54**: 427-43.

Grant, S., W. D. Grant, *et al.* (1999). "Novel archaeal phylotypes from an East African alkaline saltern." Extremophiles **3**: 139-45

Grant, W. D., R. T. Gemmell, *et al.* "Halophiles." In Extremophiles: Microbial Life in Extreme Environments. Eds: K. Horikoshi and W. D. Grant. Wiley-Liss, Inc., 1998a, b.

Grant, W. D. and B. E. Jones. "Alkaline Environments." In Encyclopedia of Microbiology. Ed: J. Lederberg. Academic Press, 2000.

Grant, W. D., M. Kamekura, *et al.* "Class III. *Halobacteria* class nov." In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001.

Grant, W. D. (2004). "Life at low water activity." Phil. Trans. R. Soc. Lond. B **359**: 1249-67.

Grigoriev, A. (1998). "Analyzing genomes with cumulative skew diagrams." Nucleic Acids Res **26**: 2286-90.

Grigoriev, A. (1999). "Strand-specific compositional asymmetries in doublestranded DNA viruses." Virus Res **60**: 1-19.

Groβkopf, R., S. Stubner, *et al.* (1998). "Novel euryarchaeotal lineages detected on rice roots and in the anoxic bulk soil of flooded rice microcosms." Appl Environ Microbiol **64**: 4983-9.

Grunder, A. L., J. M. Thompson, *et al.* (1987). "The hydrothermal system of the Calabozos caldera, central Chilean Andes." J Volcanol Geothermal Res **32**: 287-98

Hammer, Ø., B. Jamtveit, *et al.* (2005). "Evolution of fluid chemistry during travertine formation in the Troll thermal springs, Svalbard, Norway." Geofluids **5**: 140–50.

Hanada, S. (2003). "Filamentous anoxygenic phototrophs in hot springs." Microbes Environ **18**: 51-61.

Häring, M., R. Rachel, *et al.* (2005a). "Viral diversity in hot springs of Pozzuoli, Italy, and characterisation of a unique archaeal virus, *Acidianus* bottle-shaped virus, from a new family, the *Ampullaviridae*." J Virol **79**: 9904-11.

Häring, M., G. Vestergaard, *et al.* (2005b). "Independent virus development outside a host." Nature **436**: 1101-2.

Head, I. M., J. R. Saunders, *et al.* (1998). "Microbial evolution, diversity and ecology: a decade of ribosomal RNA analysis of uncultivated microorganisms." Microbiol Ecol **35**: 1-21.

Helling, R. B., H. M. Goodman, *et al.* (1974). "Analysis of endonuclease R-*Eco*RI fragments of DNA from lambdoid bacteriophages and other viruses by agarose-gel electrophoresis." J Virol **14**: 1235-44.

Hendrix, R. W., M. C. Smith, *et al.* (1999). "Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage." Proc Natl Acad Sci U S A **96**: 2192-7.

Henley, R. W. "Chemical and Physical Context for Life in Terrestrial Hydrothermal Systems: Chemical Reactors for the Early Development of Life and Hydrothermal Ecosystems." In Evolution of Hydrothermal Ecosystems on Earth (and Mars?). Eds: G. R. Bock and J. A. Goode. West Sussex, John Wiley & Sons, Ltd., 1996.

Herbst, D. B. (1998). "Potential salinity limitations on nitrogen fixation in sediments from Mono Lake, California." Int J Salt Lake Res **7**: 261-74.

Heyer, J., U. Berger, *et al.* (2005). "*Methylohalobius crimeensis* gen. nov., sp. nov., a moderately halophilic, methanotrophic bacterium isolated from hypersaline lakes of Crimea." Int J Syst Evol Microbiol **55**: 1817-26.

Hirayama, H., K. Takai, *et al.* (2005). "Bacterial community shift along a subsurface geothermal water stream in a Japanese gold mine." Extremophiles **9**: 169-84.

Hobel, C. F., V. T. Marteinsson, *et al.* (2005). "Investigation of the microbial ecology of intertidal hot springs by using diversity analysis of 16S rRNA and chitinase genes." Appl Environ Microbiol **71**: 2771-6.

Hochstaedter, A. G., P. Kepezhinskas, *et al.* (1996). "Insights into the volcanic arc mantle wedge from magnesian lavas from the Kamchatka arc." J Geophys Res **101**: 697-712.

Hoeft, S. E., T. R. Kulp, *et al.* (2004). "Dissimilatory arsenate reduction with sulfide as electron donor: experiments with mono lake water and Isolation of strain MLMS-1, a chemoautotrophic arsenate respirer." Appl Environ Microbiol **70**: 2741-7.

Holland, S. M. (2003). Analytic Rarefaction 1.3.

Hollibaugh, J. T., C. Budinoff, *et al.* (2006). "Sulfide oxidation coupled to arsenate reduction by a diverse microbial community in a soda lake." Appl Environ Microbiol **72**: 2043-9.

Holmes, M. L., S. D. Nuttall, *et al.* (1991). "Construction and use of halobacterial shuttle vectors and further studies on *Haloferax* DNA gyrase." J Bacteriol **173**: 3807-13.

Honda, H., T. Kudo, *et al.* (1985). "Two types of xylanases of alkaliphilic *Bacillus* sp. no. C-

125." Can J Microbiol **31**: 538–42.

Horikoshi, K. (1999). "Alkaliphiles: some applications of their products for biotechnology." Microbiol Molec Biol Rev **63**: 735-50.

Horner-Devine, M. C., M. A. Leibold, *et al.* (2003). "Bacterial diversity patterns along a gradient of primary productivity." Ecol Lett **6**: 613–22.

Horváth, Á., L. O. Bohus, *et al.* (2000). "Radon concentrations in hot spring waters in northern Venezuela." J Environ Radioactivity **47**: 127-33.

Hou, S., J. H. Saw, *et al.* (2004). "Genome sequence of the deep-sea gamma-proteobacterium *Idiomarina loihiensis* reveals amino acid fermentation as a source of carbon and energy." Proc Natl Acad Sci U S A **101**: 18036-41.

Hua, N.-P. and T. Naganuma (unpublished). Molecular characterization of halophilic bacteria isolated from diverse habitats, Hiroshima University, Japan.

Huang, L. N., H. Zhou, *et al.* (2004). "Phylogenetic diversity of Bacteria in the leachate of a full-scale recirculating landfill." FEMS Microbiol Ecol **50**: 175-83

Huber, H. and K. O. Stetter. "Class VI. Archaeglobi *class. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001a.

Huber, H. and K. O. Stetter. "Order II. *Desulfurococcales.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001b.

Huber, H., M. J. Hohn, *et al.* (2002). "A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont." Nature **417**: 63-7.

Hugenholtz, P., C. Pitulle, *et al.* (1998). "Novel division level bacterial diversity in a Yellowstone hot spring." J Bacteriol **180**: 366–76.

Hughes, J. B., J. J. Hellmann, *et al.* (2001). "Counting the uncountable: statistical approaches to estimating microbial diversity." Appl Environ Microbiol **67**: 4399-406.

Humayoun, S. B., N. Bano, *et al.* (2003). "Depth distribution of microbial diversity in Mono Lake, a meromictic soda lake in California." Appl Environ Microbiol **69**: 1030-42.

Hunter, E. M., H. J. Mill, *et al.* (2006). "Microbial community diversity associated with carbon and nitrogen cycling in permeable shelf sediments." Appl Environ Microbiol **72**: 5689–701.

Hurlbert, S. H. (1971). "The nonconcept of species diversity: a critique and alternative parameters." Ecology **52**: 577-86.

Ikura, Y. and K. Horikoshi (1979). "Isolation and some properties of b-galactosidase

producing bacteria." Agricul Biol Chem **43**: 85–8.

Imhoff, J. F. and J. Suling (1996). "The phylogenetic relationship among *Ectothiorhodospiraceae*: a reevaluation of their taxonomy on the basis of 16S rDNA analyses." Arch Microbiol **165**: 106-13.

Inagaki, F., T. Nunoura, *et al.* (2006). "Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin." Proc Natl Acad Sci U S A **103**: 2815-20.

Inagaki, F., M. Suzuki, *et al.* (2003). "Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk." Appl Environ Microbiol **69**: 7224-35.

Itoh, T., T. Yamaguchi, *et al.* (2005). "*Natronolimnobius baerhuensis* gen. nov., sp. nov. and *Natronolimnobius innermongolicus* sp. nov., novel haloalkaliphilic archaea isolated from soda lakes in Inner Mongolia, China." Extremophiles **9**: 111-6.

Jaccard, P. (1901). Bulletin del la Société Vaudoisedes Sciences Naturelles. **37: 241-72.**

Jackson, C. R. (unpublished). Phylogenetic diversity of Proteobacteria associated with different sizes of particles in wetland sediments, Southeastern Louisiana University, USA.

Jiang, H., H. Dong, *et al.* (2006). "Microbial diversity in water and sediment of Lake Chaka, an athalassohaline lake in northwestern China." Appl Environ Microbiol **72**: 3832-45.

Jones, D., P. A. Pell, *et al.* "Maintenance of Bacteria on Glass Beads at -60°C to -76°C." In Maintenance of Microorganisms and Cultured Cells: A Manual of Laboratory Methods. Eds: B. E. Kirsop and A. Doyle. Academic Press Limited, 1991.

Jones, B. E., W. D. Grant, *et al.* "Alkaliphiles: diversity and identification." Bacterial Diversity and Systematics. Ed: F. G. e. a. Priest. New York, Plenum Press, 1994.

Jones, B. and R. W. Renaut (1997). "Formation of silica oncoids around geysers and hot springs at El Tatio, northern Chile." Sedimentology **44**: 287-304.

Jones, B., R. W. Renaut, *et al.* (1998a). "Microbial biofacies in hot-spring sinters; a model based on Ohaaki Pool, North Island, New Zealand." J Sedim Res **68**: 413-34.

Jones, B. E., W. D. Grant, *et al.* (1998b). "Microbial diversity of soda lakes." Extremophiles **2**: 191-200.

Kaiser, H., A. K. Gordon, *et al.* (2006). "Review of the African distribution of the brine shrimp genus *Artemia*." Water SA **32**: 597-604.

Kaluzhnaya, M., V. Khmelenina, *et al.* (2001). "Taxonomic characterization of new alkaliphilic and alkalitolerant methanotrophs from soda lakes of the southeastern

Transbaikal region and description of *Methylomicrobium buryatense* sp. nov." Syst Appl Microbiol **24**: 166-76.

Kamekura, M., M. L. Dyall-Smith, *et al.* (1997). "Diversity of alkaliphilic halobacteria: proposals for transfer of *Natronobacterium vacuolatum*, *Natronobacterium magadii*, and *Natronobacterium pharaonis* to *Halorubrum*, *Natrialba*, and *Natronomonas* gen. nov., respectively, as *Halorubrum vacuolatum* comb. nov., *Natrialba magadii* comb. nov., and *Natronomonas pharaonis* comb. nov., respectively." Int J Syst Bacteriol **47**: 853-7.

Kamekura, M. and T. Mizuki (Published only in database (2001)). An extreme halophile in the order Halobacteriales, Noda Institute for Scientific Research, Chiba, Japan.

Kanai, H., T. Kobayashi, *et al.* (1995). "*Natronococcus amylolyticus* sp. nov., a haloalkaliphilic archaeon." Int J Syst Bacteriol **45**: 762-6.

Kanokratana, P., S. Chanapan , *et al.* (2004). "Diversity and abundance of bacteria and archaea in the Bor Khlueng Hot Spring in Thailand." J Basic Microbiol **44**: 430 - 44.

Kaye, J. Z. and J. A. Baross (2000). "High incidence of halotolerant bacteria in Pacific hydrothermal-vent and pelagic environments." FEMS Microbiol Ecol **32**: 249-60.

Keller, M. and K. Zengler (2004). "Tapping into microbial diversity." Nature Rev Microbiol **2**: 141-50.

Kelly, D. P. and A. P. Wood (2000). "Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halothiobacillus* gen. nov. and *Thermithiobacillus* gen. nov." Int J Syst Evol Microbiol **50**: 511–6.

Kevbrin, V. V., T. N. Zhilina, *et al.* (1998). "*Tindallia magadii* gen. nov., sp. nov.: an alkaliphilic anaerobic ammonifier from soda lake deposits." Curr Microbiol **37**: 94-100.

Khmelenina, V., M. Kaluzhnaya, *et al.* (1997). "Isolation and characterization of halotolerant alkaliphilic methanotrophic bacteria from Tuva soda Lakes." Curr Microbiol **35**: 257-61.

Kim, B.-S., H.-M. Oh, *et al.* (2005). "Archaeal diversity in tidal flat sediment as revealed by 16S rDNA analysis." The J Microbiol **43**: 144-51.

Kim, J.-S. and D. E. Crowley (unpublished). Fingerprinting bacterial communities in soil aggregates, University of California Riverside, USA.

Kimura, H., M. Sugihara, *et al.* (2005). "Microbial community in a geothermal aquifer associated with the subsurface of the Great Artesian Basin, Australia." Extremophiles **9**: 407-14.

Klappenbach, J. A., P. R. Saxman, *et al.* (2001). "rrndb: the ribosomal RNA operon copy number database." Nucleic Acids Res **29**: 181-4.

Klein, R., U. Baranyi, *et al.* (2002). "*Natrialba magadii* Virus φCh1: first complete nucleotide sequence and functional organization of a virus infecting a haloalkaliphilic archaeon." Molec Microbiol **45**: 851-63.

Kobayashi, T. (published only in NCBI database (1999)). Strictly anaerobic halophiles isolated from the can of Swedish salted herrings, Tokyo University of Fisheries, Japan.

Koch, I., J. Feldmannb, *et al.* (1999). "Arsenic in the Meager Creek hot springs environment, British Columbia, Canada." Sci Tot Environ **236**: 101-17.

Kopczynski, E. D., M. M. Bateson, *et al.* (1994). "Recognition of chimeric small-subunit ribosomal DNAs composed of genes from uncultivated microorganisms." Appl Environ Microbiol **60**: 746-8.

Krieger, B. U., N. Rezakhani, *et al.* (unpublished). Diversity of nitrate-reducing and denitrifying bacteria in a marine aquaculture biofilter, University of Bayreuth, Germany.

Kroes, I., P. W. Lepp, *et al.* (1999). "Bacterial diversity within the human subgingival crevice." Proc Natl Acad Sci U S A **96**: 14547–52.

Kumar, S., K. Tamura, *et al.* (2004). "MEGA 3: Integrated software for molecular evolutionary genetics analysis and sequence alignment." Brief Bioinform **5**: 150-63.

Kushner, D. J. (1978). Life in high salt and solute concentrations. Microbial life in extreme environments. D. J. Kushner. London, Academic: 318-368.

Lenz, P. H. Population studies on Artemia in Mono Lake, California. University of California, Santa Barbara. **PhD**, 1982.

Lepp, P. W., M. M. Brinig, *et al.* (2004). "Methanogenic archaea and human periodontal disease." Proc Natl Acad Sci U S A **101**: 6176-81.

Ley, R. E., F. Backhed, *et al.* (2005). "Obesity alters gut microbial ecology." Proc Natl Acad Sci U S A **102**: 11070-5.

Ley, R. E., J. K. Harris, *et al.* (2006). "Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat." Appl Environ Microbiol **72**: 3685-95.

Liaw, H. J. and R. A. Mah (1992). "Isolation and Characterisation of *Haloanaerobacter chitinovorans* gen. nov., sp. nov., a halophilic, anaerobic, chitinolytic bacterium from a solar saltern." Appl Environ Microbiol **58**: 260-6.

Lim, J.-M., C. O. Jeon, *et al.* (2005a). "*Lentibacillus lacisalsi* sp. nov., a moderately halophilic bacterium isolated from a saline lake in China." Int J Syst Evol Microbiol

55: 1805-9.

Lim, J. M., C. O. Jeon, *et al.* (2005b). "*Pontibacillus marinus* sp. nov., a moderately halophilic bacterium from a solar saltern, and emended description of the genus *Pontibacillus*." Int J Syst Evol Microbiol **55**: 1027-31.

Lim, J. M., C. O. Jeon, *et al.* (2006). "*Marinimicrobium koreense* gen. nov., sp. nov. and *Marinimicrobium agarilyticum* sp. nov., novel moderately halotolerant bacteria isolated from tidal flat sediment in Korea." Int J Syst Evol Microbiol **56**: 653–7.

Lin, L., C. Chen, *et al.* (2002). "*Thermus rehai* sp. nov., isolated from Rehai of Tengchong, Yunnan Province, China." J Basic Microbiol **5**: 337– 44.

Lin, L.-H., J. Hall, *et al.* (2006). "Planktonic microbial communities associated with fracture-derived groundwater in a deep gold mine of South Africa." Geomicrobiol J **23**: 475-97

Lippert and E. A. Galinski (1992). "Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying." Appl Microbiol Biotechnol **37**: 61-5.

Litchfield, C. D., A. Irby, *et al.* (2000). "Comparisons of the polar profiles of two solar salterns located in Newark, California, U.S.A., and Eilat, Israel." Extremophiles **4**: 259-65.

Litchfield, C. D., A. Irby, *et al.* (2001). "Comparative metabolic diversity in two solar salterns." Hydrobiologia **466**: 73-80.

Liu, W. Y., J. Zeng, *et al.* (2005). "*Halobacillus dabanensis* sp. nov. and *Halobacillus aidingensis* sp. nov., isolated from salt lakes in Xinjiang, China." Int J Syst Evol Microbiol **55**: 1991-6.

Liu, L. (unpublished). Direct Submission to NCBI, College of Life Sciences, Jiangsu Province, China.

López-García, P., F. Rodríguez-Valera, *et al.* (2001). "Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton." Nature **409**: 603-10.

Lozupone, C. and R. Knight (2005). "UniFrac: a new phylogenetic method for comparing microbial communities." Appl Environ Microbiol **71**: 8228-35.

Lu, Y., D. Rosencrantz, *et al.* (2006). "Structure and activity of bacterial community inhabiting rice roots and the rhizosphere." Environ Microbiol **8**: 1351-60.

Ma, Y., Y. Xue, *et al.* (2004a). "*Alkalimonas amylolytica* gen. nov., sp. nov., and *Alkalimonas delamerensis* gen. nov., sp. nov., novel alkaliphilic bacteria from soda lakes in China and East Africa." Extremophiles **8**: 193-200.

Ma, Y., W. Zhang, *et al.* (2004b). "Bacterial diversity of the Inner Mongolian Baer Soda Lake as revealed by 16S rRNA gene sequence analyses." Extremophiles **8**: 45-51.

Madigan, M. T., J. M. Martinko, *et al.* Brock Biology of Microorganisms (International Edition). 10th ed. Pearson Education, Inc, 2003

Maki, J. S. "Genus IV. *Thermodesulfovibrio.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001.

Manning, T. I., S. Boomer, *et al.* (unpublished). Attempts to cultivate RLMO bacteria using site-specific water chemistry data, Western Oregon University, USA.

Mantel, N. (1967). "The detection of disease clustering and generalized regression approach." Cancer Res **27**: 209-20.

Martin, A. P. (2002). "Phylogenetic approaches for describing and comparing the diversity of microbial communities." Appl Environ Microbiol **68**: 3673-82.

Martínez-Cánovas, M. J., E. Quesada, *et al.* (2004). "*Halomonas ventosae* sp. nov., a moderately halophilic, denitrifying, exopolysaccharide-producing bacterium." Int J Syst Evol Microbiol **54**: 733-7.

Martiny, J. B. H., B. J. M. Bohannan, *et al.* (2006). "Microbial biogeography: putting microorganisms on the map." Nature Rev Microbiol **4**: 102-12.

Maszenan, A. M., R. J. Seviour, *et al.* (1999). "*Friedmanniella spumicola* sp. nov. and *Friedmanniella capsulata* sp. nov. from activated sludge foam: gram-positive cocci that grow in aggregates of repeating groups of cocci." Int J Syst Bacteriol **49**:1667-80.

Mata, J. A., J. Martínez-Cánovas, *et al.* (2002). "A detailed phenotypic characterisation of the type strains of *Halomonas* species." Syst Appl Microbiol **25**: 360-75.

Mathrani, I. M., D. R. Boone, *et al.* (1988). "*Methanohalophilus zhilinae* sp. nov., an alkaliphilic, halophilic, methylotrophic methanogen." Int J Syst Bacteriol **38**: 139-42.

Maturrano, L., F. Santos, *et al.* (2006). "Microbial diversity in Maras salterns, a hypersaline environment in the Peruvian Andes." Appl Environ Microbiol **72**: 3887-95.

Maxwell, S., G. K. Sims, *et al.* (unpublished). Distribution of tetracycline and tylosin resistance genes in bacteria isolated from swine effluent impacted environments, University of Illinois, USA.

McCammon, S. A. and J. P. Bowman (2000). "Taxonomy of Antarctic *Flavobacterium* species: description of *Flavobacterium gillisiae* sp. nov., *Flavobacterium tegetincola* sp. nov. and *Flavobacterium xanthum* sp. nov., nom. rev. and reclassification of [*Flavobacterium*] *salegens* as *Salegentibacter salegens* gen. nov., comb. nov." Int J Syst Evol Microbiol **50**: 1055-63.

McCliment, E. A., K. M. Voglesonger, *et al.* (2006). "Colonization of nascent, deep-sea

hydrothermal vents by a novel archaeal and nanoarchaeal assemblage." Environ Microbiol **8**: 114–25.

McCraig, A. E., L. A. Glover, *et al.* (1999). "Molecular analysis of bacterial community structure and diversity in unimproved and improved upland grass pastures." Appl Environ Microbiol **65**: 1721-30.

McGenity, T. J. and W. D. Grant (1993). "The haloalkaliphilic archaeon (archaebacterium) *Natronococcus occultus* represents a distinct lineage within the *Halobacteriales*, most closely related to the other haloalkaliphilic lineage (*Natronobacterium*)." Syst Appl Microbiol **16**: 239-43.

McGenity, T. J., R. T. Gemmell, *et al.* (1998). "Proposal of a new halobacterial genus *Natrinema* gen. nov., with two species *Natrinema pellerubra* nom. nov. and *Natrinema pallida* nom. nov." Int J Syst Bacteriol **48**: 1187-96.

Mesbah, N., S. Abou-El-Ela, *et al.* (In Press). Novel and unexpected prokaryotic diversity in water and sediments of the alkaline, hypersaline lakes of the Wadi An Natrun, Egypt, The University of Georgia, USA.

Meyer-Dombard, D. R., E. L. Shock, *et al.* (2005). "Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA." Geobiology **3**: 211-27.

Miller, S. R. and R. W. Castenholz (2000). "Evolution of thermotolerance in hot spring cyanobacteria of the genus *Synechococcus*." Appl Environ Microbiol **66**: 4222–9.

Miller, R. V., A. E. Kirkwood, *et al.* (2004). Microbiology at the Salt Plains Microbiology Observatory. Third Annual Microbial Observatories Principle Investigators Workshop.

Miyoshi, T., T. Iwatsuki, *et al.* (2005). "Phylogenetic characterization of 16S rRNA gene clones from deep-groundwater microorganisms that pass through 0.2-micrometer-pore-size filters." Appl Environ Microbiol **71**: 1084-8.

Mongodin, E. F., K. E. Nelson, *et al.* (2005). "The genome of *Salinibacter ruber*: Convergence and gene exchange among hyperhalophilic bacteria and archaea." Proc Natl Acad Sci U S A **102**: 18147-52.

Montalvo-Rodriguez, R., R. H. Vreeland, *et al.* (1998). "*Halogeometricum borinquense* gen. nov., sp. nov., a novel halophilic archaeon from Puerto Rico." Int J Syst Bacteriol **48**: 1305-12.

Moré, M., J. B. Herrick, *et al.* (1994). "Quantitative cell lysis of indigenous microorganisms and rapid extraction of microbial DNA from sediment." Appl Environ Microbiol **60**: 1572-80.

Mormile, M. R., M. F. Romine, *et al.* (1999). "*Halomonas campisalis* sp. nov., a denitrifying, moderately haloalkaliphilic bacterium." Syst Appl Microbiol **22**: 551-8.

Mouné, S., C. Eatock, *et al.* (2000). "*Orenia salinaria* sp. nov., a fermentative bacterium isolated from anaerobic sediments of Mediterranean salterns." Int J Syst Evol Microbiol **50**: 721-9.

Mouné, S., P. Caumette, *et al.* (2003). "Molecular sequence analysis of prokaryotic diversity in the anoxic sediments underlying cyanobacterial mats of two hypersaline ponds in Mediterranean salterns." FEMS Microbiol Ecol **44**: 117-30.

Mountain, B. W., L. G. Benning, *et al.* (2003). "Experimental studies on New Zealand hot spring sinters: rates of growth and textural development." Can J Microbiol **40**: 1643-67.

Mukohata, Y., Y. Sugiyama, *et al.* (1988). "An Australian halobacterium contains a novel proton pump retinal protein: archaerhodopsin." Biochem Biophys Res Commun **151**: 1339-45.

Mwatha, W. E. and W. D. Grant (1993). "*Natronobacterium vacuolata* sp. nov., a haloalkaliphilic archaeon isolated from Lake Magadi, Kenya." Int J Syst Bacteriol **43**: 401-4.

Mylvaganam, S. and P. P. Dennis (1992). "Sequence heterogeneity between the two genes encoding 16S rRNA from the halophilic archaebacterium *Haloarcula marismortui*." Genetics **130**: 399-410.

Nakagawa, T. and M. Fukui (2002). "Phylogenetic characterization of microbial mats and streamers from a Japanese alkaline hot spring with a thermal gradient." J Gen Appl Microbiol **48**: 211-22.

Nakagawa, Y., T. Sakane, *et al.* (2002). "Phylogenetic structure of the genera *Flexibacter*, *Flexithrix*, and *Microscilla* deduced from 16S rRNA sequence analysis." J Gen Appl Microbiol **48**: 155-65.

Nakagawa, S., K. Takai, *et al.* (2005). "Variability in microbial community and venting chemistry in a sediment-hosted backarc hydrothermal system: Impacts of subseafloor phase-separation." FEMS Microbiol Ecol **54**: 141-55.

Nedashkovskaya, O. I., M. Suzuki, *et al.* (2004). "*Salegentibacter holothuriorum* sp. nov., isolated from the edible holothurian *Apostichopus japonicus*." Int J Syst Evol Microbiol **54**: 1107–10.

Nedashkovskaya, O. I., S. B. Kim, *et al.* (2005). "*Salegentibacter mishustinae* sp. nov., isolated from the sea urchin *Strongylocentrotus intermedius*." Int J Syst Evol

Microbiol **55**: 235–8.

Nercessian, O., E. Noyes, *et al.* (2005). "Bacterial populations active in metabolism of C1 compounds in the sediment of Lake Washington, a freshwater lake." Appl Environ Microbiol **71**: 6885-99.

Nesbø, C. L., M. Dlutek, *et al.* (2006). "Evidence for existence of "Mesotogas," members of the order *Thermotogales* adapted to low-temperature environments." Appl Environ Microbiol **72**: 5061-8.

Ng, W. V., S. P. Kennedy, *et al.* (2000). "Genome sequence of *Halobacterium* species NRC-1." Proc Natl Acad Sci U S A **97**: 12176-81.

Nold, S. C., E. D. Kopczynski, *et al.* (1996). "Cultivation of aerobic chemoorganotrophic Proteobacteria and Gram-positive bacteria from a hot spring microbial mat." Appl Environ Microbiol **62**: 3917–21.

Norris, T. B., T. R. McDermott, *et al.* (2002). "The long-term effects of UV exclusion on the microbial composition and photosynthetic competence of bacteria in hot-spring microbial mats." FEMS Microbiol Ecol **39**: 193-209.

Nunoura, T., H. Hirayama, *et al.* (2005). "Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments." Environ Microbiol **7**: 1967-84.

Nuttall, S. and M. Dyall-Smith (1993). "HF1 and HF2: novel bacteriophages of halophilic archaea." Virology **197**: 678-84.

Nuttall, S. and M. Dyall-Smith (1995). "Halophage HF2: genome organisation and replication strategy." J Virol **69**: 2322-7.

Ochsenreiter, T., F. Pfeifer, *et al.* (2002). "Diversity of Archaea in hypersaline environments characterized by molecular-phylogenetic and cultivation studies." Extremophiles **6**: 267-74.

Ochsenreiter, T., D. Selezi, *et al.* (2003). "Diversity and abundance of *Crenarchaeota* in terrestrial habitats studied by 16S rRNA surveys and real time PCR." Envrion Microbiol **5**: 787-97.

Oe, H., T. Fukushima, *et al.* (unpublished). Microbial community analysis of enhanced biological phosphorus removal activated sludge, The University of Tokyo, Japan.

Ogawa, K. and Y. Maki (2003). "Cellulose as extracellular polysaccharide of hot spring sulfur-turf bacterial mat." Biosci Biotech Biochem **67**: 2652-4.

Okamoto, T. and T. Naganuma (unpublished). Distribution of euryhaline halophilic microorganisms, Hiroshima University, Japan.

Ollivier, B., C. E. Hatchikian, *et al.* (1991). "*Desulfohalobium retbaense* gen. nov., sp. nov., a halophilic sulfate-reducing bacterium from sediments of a hypersaline lake in Senegal." Int J Syst Bacteriol **41**: 74-81.

Ollivier, B., P. Caumette, *et al.* (1994). "Anaerobic bacteria from hypersaline environments." Microbiol Rev **58**: 27-38.

Onstott, T. C., D. P. Moser, *et al.* (2003). "Indigenous and contaminant microbes in ultradeep mines." Environ Microbiol **5**: 1168-91.

Oremland, R. S. (1990). "Nitrogen fixation dynamics of two diazotrophic communities in Mono Lake, California." Appl Environ Microbiol **56**: 614-22.

Oremland, R. S., S. E. Hoeft, *et al.* (2002). "Anaerobic oxidation of arsenite in Mono Lake water and by a facultative, arsenite-oxidizing chemoautotroph, strain MLHE-1." Appl Environ Microbiol **68**: 4795-802.

Oremland, R. S., J. F. Stolz, *et al.* (2004). "The microbial arsenic cycle in Mono Lake, California." FEMS Microbiol Ecol **48**: 15-27.

Oremland, R. S., T. R. Kulp, *et al.* (2005). "A microbial arsenic cycle in a salt-saturated, extreme environment." Science **308**: 1305-8.

Oren, A. and M. Shilo (1982). "Population dynamics of *Dunaliella parva* in the Dead Sea." Limnol Oceanography **27**: 201-11.

Oren, A., W. G. Weisburg, *et al.* (1984). "*Halobacteroides halobius* gen. nov., sp. nov., a moderately halophilic anaerobic bacterium from the bottom sediments of the Dead Sea." Syst Appl Microbiol **5**: 58-70.

Oren, A. (1988). "Anaerobic degradation of organic compounds at high salt concentrations." Antonie van Leeuwenhoek **54**: 267-77.

Oren, A., M. Kessel, *et al.* (1989). "*Ectothiorhodospira marismortui* sp. nov., an obligately anaerobic, moderately halophilic purple sulfur bacterium from a hypersaline sulfur spring on the shore of the Dead Sea." Arch Microbiol **151**: 524-9.

Oren, A. and P. Gurevich (1993). "Characterisation of the dominant halophilic archaea in a bacterial bloom in the Dead Sea." FEMS Microbiol Ecol **12**: 249-56.

Oren, A. (1994). "The ecology of the extremely halophilic archaea." FEMS Microbiol Rev **13**: 415-40.

Oren, A. and Z. Dubinsky (1994). "On the red coloration of saltern crystallizer ponds. II. Additional evidence for the contribution of halobacterial pigments." Int J Salt Lake Res **3**: 9-13.

Oren, A., G. Bratbak, *et al.* (1997a). "Occurrence of virus-like particles in the Dead Sea."

Extremophiles **1**: 143-9.

Oren, A., A. Ventosa, *et al.* (1997b). "Proposed minimal standards for description of new taxa in the order *Halobacteriales*." Int J Syst Bacteriol **47**: 233-8.

Oren, A. and F. Rodriguez-Valera (2001). "The contribution of halophilic Bacteria to the red coloration of saltern crytallizer ponds." FEMS Microbiol Ecol **36**: 123-30.

Oren, A., A. Ventosa, *et al.* (1999). "*Haloarcula quadrata* sp. nov., a square, motile archaeon isolated from a brine pool in Sinai (Egypt)." Int J Syst Bacteriol **49**: 1149-55.

Oren, A., R. Elevi, *et al.* (2002). "*Halomicrobium mukohataei* gen. nov., comb. nov., and emended description of *Halomicrobium mukohataei*." Int J Syst Evol Microbiol **52**: 1831-5.

Oren, A. (2002a). "Diversity of halophilic microorganisms: environments, phylogeny, physiology and applications." J Ind Microbiol Biotechnol **28**: 56-63.

Oren, A. (2002b). "Molecular ecology of extremely halophilic Archaea and Bacteria." FEMS Microbiol Ecol **39**: 1-7.

Oren, A. (2005). "A hundred years of *Dunaliella* research: 1905–2005." Saline Syst **1**: 1-14.

Overmann, J. "Class I. 'Chlorobia'". In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001.

Ozcan, B., C. Cokmus, *et al.* (unpublished). Diversity of halophilic archaea in Salt Lake and Bolluk Lake (Turkey), Mustafa Kemal University, Hatay, Turkey.

Page, A., S. K. Juniper, *et al.* (2004). "Microbial diversity associated with a *Paralvinella sulfincola* tube and the adjacent substratum on an active deep-sea vent chimney." Geobiology **2**: 225-38

Papastefanou, C., M. Manolopoulou, *et al.* (2001). "Radon measurements along active faults in the Langadas Basin, northern Greece." Nat Hazards Earth Syst Sci **1**: 159–64.

Pašić, L., S. G. Bartual, *et al.* (2005). "Diversity of halophilic archaea in the crystallizers of an Adriatic solar saltern." FEMS Microbiol Ecol **54**: 491-8.

Patel, B. K. C., H. W. Morgan, *et al.* (1985). "Thermophilic anaerobic spirochetes in New Zealand hot springs." FEMS Microbiol Lett **26**: 101-6.

Pauling, C. (1982). "Bacteriophages of *Halobacterium halobium*: isolation from fermented fish sauce and primary characterisation." Can J Microbiol **28**: 916-21.

Payne, J. I., S. N. Sehgal, *et al.* (1960). "Immersion refractometry of some halophilic bacteria." Can J Microbiol **6**: 9-15.

Peçonek, J., C. Gruber, *et al.* (2006). "Reclassification of *Pseudomonas beijerinckii* Hof 1935 as *Chromohalobacter beijerinckii* comb. nov., and emended description of the

species." Int J Syst Evol Microbiol **56**: 1953-7.

Pedrós-Alió, C., J. I. Calderón-Paz, *et al.* (2000). "The microbial food web along salinity gradients." FEMS Microbiol Ecol **32**: 143-55.

Peet, R. K. (1974). "The measurement of species diversity." Ann Rev Ecol Syst **5**: 285-307.

Penny, D. and A. Poole (1999). "The nature of the last universal common ancestor." Curr Opin Genet Devel **9**: 672-7.

Pielou, E. C. (1966). "Species-diversity and pattern-diversity in the study of ecological succession." J Theoret Biol **10**: 370-83.

Pikuta, E. V., R. B. Hoover, *et al.* (2003a). "*Tindallia californiensis* sp. nov., a new anaerobic, haloalkaliphilic, spore-forming acetogen isolated from Mono Lake in California." Extremophiles **7**: 327-34.

Pikuta, E. V., R. B. Hoover, *et al.* (2003b). "*Desulfonatronum thiodismutans* sp. nov., a novel alkaliphilic, sulfate-reducing bacterium capable of lithoautotrophic growth " Int J Syst Evol Microbiol **53**: 1327-32.

Pinhassi, J. and T. Berman (2003). "Differential growth response of colony-forming α- and γ-Proteobacteria in dilution culture and nutrient addition experiments from Lake Kinneret (Israel), the Eastern Mediterranean Sea, and the Gulf of Eilat." Appl Environ Microbiol **69**: 199-211.

Porter, K., P. Kukkaro, *et al.* (2005). "SH1: A novel, spherical halovirus isolated from an Australian hypersaline lake." Virology **335**: 22-33.

Pradhan, D., D. Marsic, *et al.* (unpublished). Novel alkaliphiles from Tanzania soda lakes, University of Alabama - Huntsville, USA.

Prangishvili, D., P. Forterre, *et al.* (2006a). "Viruses of the archaea: a unifying view." Nature Rev Microbiol **4**: 837-48.

Prangishvili, D., G. Vestergaard, *et al.* (2006b). "Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an extracellular stage of the reproductive cycle." J Molec Biol **359**: 1203-16.

Purdy, K. J., T. D. Creswell-Maynard, *et al.* (2004). "Isolation of haloarchaea that grow at low salinities." Environ Microbiol **6**: 591-5.

Radax, C., E. B. M. Denner, *et al.* (unpublished). NaxosII, an extremely halophilic archaeon isolated from the island of Naxos, Greece, University of Salzburg, Austria.

Rainey, F. and M. S. Da Costa. "Order II. Thermales *ord. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001.

Ramsing, N. B., M. J. Ferris, *et al.* (1997). "Light-induced motility of thermophilic *Synechococcus* isolates from Octopus Spring, Yellowstone National Park." Appl Environ Microbiol **63**: 2347-54.

Ramsing, N. B., M. J. Ferris, *et al.* (2000). "Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter-thick photic zone of a hot spring cyanobacterial mat." Appl Environ Microbiol **66**: 1038-49.

Ranjard, L., F. Poly, *et al.* (2000). "Monitoring complex bacterial communities using culture-independent molecular techniques: application to soil environment." Res Microbiol **151**: 167-77.

Rees, H. C. Environmental Gene Screening. Department of Microbiology and Immunology, University of Leicester. **PhD**, 2002.

Rees, H. C., S. Grant, *et al.* (2003). "Detecting cellulase and esterase enzyme activity encoded by novel genes present in environmental DNA." Extremophiles **7**: 415-22.

Rees, H. C., W. D. Grant, *et al.* (2004). "Diversity of Kenyan soda lake alkaliphiles assessed by molecular methods." Extremophiles **8**: 63-71.

Reichenbach, H. "The order *Cytophagales*." In The Prokaryotes. Eds: A. Balows, H. G. Trüper, M. Dworkin, W. Harder and K. H. Schleifer. 2nd ed. New York, Springer-Verlag, 1991.

Reysenbach, A., G. S. Wickham, *et al.* (1994). "Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park." Appl Environ Microbiol **60**: 2113-9.

Reysenbach, A., A. Banta, *et al.* (2000a). "Microbial essentials at hydrothermal vents." Nature **404**: 835.

Reysenbach, A., M. Ehringer, *et al.* (2000b). "Microbial diversity at 83°C in Calcite Springs, Yellowstone National Park: another environment where the *Aquifcales* and "Korarchaeota" coexist." Extremophiles **4**: 61-7.

Reysenbach, A. L., K. Longnecker, *et al.* (2000c). "Novel bacterial and archaeal lineages from an in situ growth chamber deployed at a Mid-Atlantic Ridge hydrothermal vent." Appl Environ Microbiol **66**: 3798-806.

Reysenbach, A. "Class I. Thermoprotei *class. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001a.

Reysenbach, A. "Class IV. Thermoplasmata *class. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York,

Springer-Verlag, 2001b.

Reysenbach, A. "Phylum BI. Aquificae *phy. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001c.

Reysenbach, A. "Phylum BII. Thermotogae *phy. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001d.

Rice, G., K. Stedman, *et al.* (2001). "Viruses from extreme thermal environments." Proc Natl Acad Sci U S A **98**: 13341-5.

Roberts, M. F. (2005). "Organic compatible solutes of halotolerant and halophilic microorganisms." Saline Syst **1**.

Roesch, C., J. Bruns, *et al.* (unpublished). Bacterial diversity of a German forest soil, as assessed by *nif*H, *nos*Z, and SSU clone libraries and tRFLP, University of Cologne, Germany.

Roeselers, G., T. B. Norris, *et al.* (2007). "Diversity of phototrophic bacteria in microbial mats from Arctic hot springs (Greenland)." Environ Microbiol **9**: 26–38.

Rohrmann, G. F., R. Cheney, *et al.* (1983). "Bacteriophages of *Halobacterium halobium*: virion DNAs and proteins." Can J Microbiol **29**: 627-9.

Romano, I., B. Nicolaus, *et al.* (1996). "Characterization of haloalkaliphilic strictly aerobic bacterium, isolated from Pantelleria island." Syst Appl Microbiol **19**: 326-33

Romano, I., L. Lama, *et al.* (2005). "*Bacillus saliphilus* sp. nov., isolated from a mineral pool in Campania, Italy." Int J Syst Evol Microbiol **55**: 159-63.

Rosselló-Mora, R., N. Lee, *et al.* (2003). "Substrate uptake in extremely halophilic microbial communities revealed by microautoradiography and fluorescence in situ hydridization." Extremophiles **7**: 409-13.

Rudolph, C., C. Moissl, *et al.* (2004). "Ecology and microbial structures of archaeal/bacterial strings-of-pearls communities and archaeal relatives thriving in cold sulfidic springs." FEMS Microbiol Ecol **50**: 1-11

Sait, M., Hugenholtz, P., *et al.* (2002). "Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys." Environ Microbiol **4**: 654-666.

Sambrook, J., E. F. Fritsch, *et al.* Molecular Cloning: A Laboratory Manual. 2nd ed. New York, Cold Spring Harbour, 1989.

Satorius-Neef, S. and F. Pfeifer (2004). "*In vivo* studies on putative shine-Dalgarno sequences

of the halophilic archaeon *Halobacterium salinarium*." Molec Microbiol **51**: 579-88.

Saul, D. J., R. A. Reeves, *et al.* (1999). "*Thermus* diversity and strain loss during enrichment." FEMS Microbiol Ecol **30**: 157-62.

Schleper, C., W. Holben, *et al.* (1997). "Recovery of crenarchaeotal ribosomal DNA sequences from freshwater-lake sediments." Appl Environ Microbiol **63**: 321-3.

Schloss, P. D. and J. Handelsman (2005). "Introducing DOTUR, a computer program for defining Operational Taxonomic Units and estimating species richness." Appl Environ Microbiol **71**: 1501-6.

Schloss, P. D. and J. Handelsman (2006). "Introducing TreeClimber, a test to compare microbial community structures." Appl Environ Microbiol **72**: 2379-84.

Schnabel, I. I., W. Zilig, *et al.* (1982). "*Halobacterium halobium* phage φH." Molec Gen Genet **235**: 197-204.

Schwartz, D. C. and C. R. Cantor (1984). "Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis." Cell **37**: 67-75.

Shannon, C. E. and W. Weaver. The mathematical theory of communication. Urbana, University of Illinois Press, 1963.

Sher, J., R. Elevi, *et al.* (2004). "Glycerol metabolism in the extremely halophilic bacterium *Salinibacter ruber*." FEMS Microbiol Lett **232**: 211-5.

Shiguang, X. and X. Yong Study on Kunming low-temperature geothermal field. Kunming, Yunnan Geological Engineering Survey Institute: 2289-94.

Simpson, E. H. (1949). "Measurement of diversity." Nature **163**: 688.

Skirnisdottir, S., G. O. Hreggvidsson, *et al.* (2000). "Influence of sulfide and temperature on species composition and community structure of hot spring microbial mats." Appl Environ Microbiol **66**: 2835-41.

Sokolov, A. P. and Y. A. Trotsenko (1995). "Methane consumption in (hyper)saline habitats of Crimea (Ukraine)." FEMS Microbiol Ecol **18**: 299-304.

Sonne-Hansen, J., I. M. Mathrani, *et al.* (1993). "Xylanolytic anaerobic thermophiles from Icelandic hot-springs." Appl Microbiol Biotechnol **38**: 537-41.

Sørensen, K. B., D. E. Canfield, *et al.* (2005). "Community composition of a hypersaline endoevaporitic microbial mat." Appl Environ Microbiol **71**: 7352-65.

Sorokin, D. Y., G. Muyzer, *et al.* (1998). "Isolation and characterisation of a novel facultatively alkaliphilic *Nitrobacter* species - *Nb. alkalicus* sp. nov." Arch Microbiol **170**: 345-52.

Sorokin, D. Y., T. P. Tourova, *et al.* (2000). "A new facultatively autotrophic hydrogen- and

sulfur-oxidizing bacterium from an alkaline environment." Extremophiles 4: 237-45.

Sorokin, D. Y., T. P. Tourova, *et al.* (2001). "Isolation and properties of obligately chemolithotrophic and extremely alkalitolerant ammonia-oxidizing bacteria from Mongolian soda lakes." Arch Microbiol **176**: 170-7.

Sorokin, D. Y., T. P. Tourova, *et al.* (2002). "*Thioalkalivibrio thiocyanooxidans* sp. nov. and *Thioalkalivibrio paradoxus* sp. nov., novel alkaliphilic, obligately autotrophic, sulfur-oxidising bacteria from soda lakes able to grow with thiocyanate." Int J Syst Evol Microbiol **52**: 657-64.

Sorokin, D. Y., T. P. Tourova, *et al.* (2003). "*Thioalkavibrio nitratireducens* sp. nov., a nitrate-reducing member of an autotrophic denitrifying consortium from a soda lake." Int J Syst Evol Microbiol **53**: 1779-83.

Sorokin, D. Y. and J. G. Kuenen (2005a). "Chemolithotrophic haloalkaliphiles from soda lakes." FEMS Microbiol Ecol **52**: 287-95.

Sorokin, D. Y. and J. G. Kuenen (2005b). "Haloalkaliphilic sulfur-oxidising bacteria in soda lakes." FEMS Microbiol Rev **29**: 685-702.

Sorokin, D. Y., T. P. Tourova, *et al.* (2006). "Extremely halophilic denitrifying bacteria from hypersaline inland lakes, *Halovibrio denitrificans* sp. nov. and *Halospina denitrificans* gen. nov., sp. nov., and evidence that the genus name *Halovibrio* Fendrich 1989 with the type species *Halovibrio variabilis* should be associated with DSM 3050." Int J Syst Evol Microbiol **56**: 379-88.

Spear, J. R., J. J. Walker, *et al.* (2005). "Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem." Proc Natl Acad Sci U S A **102**: 2555-60

Spring, S., W. Ludwig, *et al.* (1996). "*Halobacillus* gen nov., with descriptions of *Halobacillus litoralis* sp. nov. and *Halobacillus trueperi* sp. nov., and transfer of *Sporosarcina halophila* to *Halobacillus halophilus* comb. nov." Int J Syst Bacteriol **46**: 492-6.

Steinsbu, B. O. and T. Torsvik (unpublished). Characterization of bacteria isolated from basalt and an extinct smoker pipe at Kolbeinsey Ridge, University of Bergen, Norway.

Stephens, D. W. (1990). "Changes in lake levels, salinity and the biological community of Great Salt Lake (Utah, USA), 1847–1987." Hydrobiologia **197**: 139-46.

Stolt, P. and W. Zillig (1994). "Transcription of the Halophage Phi H repressor gene is abolished by transcription from an inversely oriented lytic promoter." FEBS Lett **344**: 125-8.

Suau, A., R. Bonnet, *et al.* (1999). "Direct analysis of genes encoding 16S rRNA from

complex communities reveals many novel molecular species within the human gut." Appl Environ Microbiol **65**: 4799–807.

Sunna, A. and P. L. Bergquist (2003). "A gene encoding a novel extremely thermostable 1,4-beta-xylanase isolated directly from an environmental DNA sample." Extremophiles **7**: 63-70.

Suzuki, T., T. Iwasaki, *et al.* (2002). "*Sulfolobus tokodaii* sp. nov. (*Sulfolobus* sp. strain 7), a new member of the genus *Sulfolobus* isolated from Beppu Hot Springs, Japan." Extremophiles **6**: 39-44.

Taher, A. G. (1999). "Inland saline lakes of Wadi El Natrun depression, Egypt " Int J Salt Lake Res **8**: 149-70.

Takacs, C. D., M. Ehringer, *et al.* (2001). "Phylogenetic characterization of the blue filamentous bacterial community from an Icelandic geothermal spring." FEMS Microbiol Ecol **35**: 123-8.

Takacs-Vesbach, C. D., A. Banta, *et al.* (unpublished). The microbial community of a near-neutral thermal spring in Kamchatka: expansion of the geographic occurrence of a novel lineage within the *Aquificales*, Portland State University, Oregon, USA.

Takai, K. and K. Horikoshi (1999). "Genetic diversity of archaea in deep-sea hydrothermal vent environments." Genetics **152**: 1285-97.

Takai, K. and Y. Sako (1999). "A molecular view of archaeal diversity in marine and terrestrial hot water environments." FEMS Microbiol Ecol **28**: 177-88.

Takai, K., D. P. Moser, *et al.* (2001). "Archaeal diversity in waters from deep South African gold mines." Appl Environ Microbiol **67**: 5750-60.

Takami, K., K. Nakasone, *et al.* (2000). "Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*." Nucleic Acids Res **28**: 4317-31.

Tang, S.-L., S. Nuttall, *et al.* (2002). "HF2: a double-stranded DNA tailed haloarchaeal virus with a mosaic genome." Molec Microbiol **44**: 283-96.

Tang, S.-L., S. Nuttall, *et al.* (2004). "Haloviruses HF1 and HF2: evidence for a recent and large recombination event." J Bacteriol **186**: 2810-7.

Tanner, M. A., B. M. Goebel, *et al.* (1998). "Specific ribosomal DNA sequences from diverse environmental settings correlate with experimental contaminants." Appl Environ Microbiol **64**: 3110-3.

Tansey, M. R. and T. D. Brock (1972). "The upper temperature limit for eukaryotic organisms." Proc Natl Acad Sci U S A **69**: 2426-8.

Tao, G., Y. Xiao, *et al.* (2003). "Climatology and trends of the temporal and spatial distribution of sandstorms in Inner Mongolia." Water, Air and Soil Pollution: Focus **3**: 51-66.

Tindall, B. J., A. A. Mills, *et al.* (1980). "An alkaliphilic red halophilic bacterium with a low magnesium requirement from a Kenyan soda lake." J Gen Microbiol **116**: 257-60.

Tindall, B. J., H. N. M. Ross, *et al.* (1984). "*Natronobacterium*, gen. nov. and *Natronococcus*, gen. nov., two new genera of haloalkaliphilic archaebacteria." Syst Appl Microbiol **5**: 41-57.

Tokunaga, H., M. Ishibashi, *et al.* (2004). "Highly efficient renaturation of beta-lactamase isolated from moderately halophilic bacteria." FEBS Lett **558**: 7-12.

Tokunaga, H., K. Mitsuo, *et al.* (unpublished). Major outer membrane proteins in moderately halophilic eubacteria isolated from saline environments, Kagoshima University, Japan.

Tolstrup, N., C. W. Sensen, *et al.* (2000). "Two different and highly organized mechanisms of translation initiation in the archaeon *Sulfolobus solfataricus*." Extremophiles **4**: 175-9.

Tomlinson, G. A. and L. I. Hochstein (1976). "*Halobacterium saccharovorum* sp. nov., a carbohydrate-metabolising, extremely halophilic bacterium." Can J Microbiol **22**: 587-91.

Torsvik, T. and I. D. Dundas (1974). "Bacteriophage of *Halobacterium salinarium*." Nature **248**: 680-1.

Tsai, C.-R., J. L. Garcia, *et al.* (1995). "*Haloanaerobium alcaliphilum* sp. nov., an anaerobic moderate halophile from the sediments of Great Salt Lake, Utah." Int J Syst Bacteriol **45**: 301–7.

van der Meer, M. T. J., S. Schouten, *et al.* (2005). "Diel variations in carbon metabolism by green nonsulfur-like bacteria in alkaline siliceous hot spring microbial mats from Yellowstone National Park." Appl Environ Microbiol **71**: 3978-86.

van Regenmortel, M. H. V., C. M. Fauquet, *et al.* (Eds). Virus Taxonomy: Classification and Nomenclature of Viruses. 7th report. Academic Press, 2000.

Vargas, V. A., O. D. Delgado, *et al.* (2005). "*Bacillus bogoriensis* sp. nov., a novel alkaliphilic halotolerant bacterium isolated from a Kenyan soda lake." Int J Syst Evol Microbiol **55**: 899-902.

Varnam, A. H. "Extreme Environments." In Environmental Microbiology. Manson, 2000.

Ventosa, A. "Taxonomy and Phylogeny of Moderately Halophilic Bacteria." In Bacterial Diversity and Systematics. Ed: F. G. e. a. Priest. New York, Plenum Press, 1994.

Ventosa, A., M. C. Márquez, *et al.* (1998a). "Moderately halophilic Gram-positive bacterial

diversity in hypersaline environments." Extremophiles **2**: 297-304.

Ventosa, A., J. Nieto, *et al.* (1998b). "Biology of moderately halophilic aerobic bacteria." Microbiol Molec Biol Rev **62**: 540-44.

Ventosa, A., M. C. Gutierrez, *et al.* (2004). "Taxonomic study of *Halorubrum distributum* and proposal of *Halorubrum terrestre* sp. nov." Int J Syst Evol Microbiol **54**: 389-92.

Vetriani, C., H. V. Tran, *et al.* (2003). "Fingerprinting microbial assemblages from the oxic/anoxic chemocline of the Black Sea." Appl Environ Microbiol **69**: 6481-8.

Vogelsang-Wenke, H. and D. Oesterhelt (1988). "Isolation of a halobacterial phage with a fully cytosine-methylated genome." Molec Gen Genet **211**: 407-14.

Vreeland, R. H. (1987). "Mechanism of halotolerance in microorganisms." CRC Crit Rev Microbiol **14**: 311-55.

Vreeland, R. H., S. Straight, *et al.* (2002). "*Halosimplex carlsbadense* gen. nov., sp. nov., a unique halophilic archaeon, with three 16S rRNA genes, that grows only in defined medium with glycerol and acetate or pyruvate." Extremophiles **6**: 445-52.

Wais, A. C., M. Kon, *et al.* (1975). "Salt-dependent bacteriophage infecting *Halobacterium cutirubrum* and *H. halobium*." Nature **256**: 314-5.

Wallet, F., M. Tachon, *et al.* (2005). "*Vibrio metschnikovii* Pneumonia." Emerging Infectious Diseases **11**: 1641-2.

Walsby, A. E. (1980). "A square bacterium." Nature (London) **283**: 69-71.

Walsby, A. E., J. Van Rijn, *et al.* (1983). "The biology of a new gas-vacuolate Cyanobacterium, *Dactylococcopsis salina* sp. nov., in Solar Lake " Proceedings of the Royal Society of London. Series B, Biological Sciences **217**: 417-47.

Walsh, D. A., R. T. Papke, *et al.* (2005). "Archaeal diversity along a soil salinity gradient prone to disturbance." Environ Microbiol **7**: 1655-66.

Wang, G. C. and Y. Wang (1996). "The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species." Microbiology **142**: 1107-14.

Wang, H. Biogeography of Prokaryotes in Hypersaline Environments, Department of Genetics, University of Leicester. **M.Sc. Molecular Genetics**, 2006.

Wang, G., S. Y. Carroll, *et al.* (unpublished-a). The diversity of *Archaea* and *Bacteria* in two similar Arizona pristine soils, The University of Arizona, USA.

Wang, Y. N., X. Dai, *et al.* (unpublished-b). Analysis of a bacterial community in a near-shore marine-cultural sediment, Chinese Academy of Sciences, Zhongguancun, China.

Wannamaker, P. E. (1997). "Tensor CSAMT survey over the sulphur springs thermal area,

Valles Caldera, New Mexico, U.S.A., Part I: Implications for structure of the western caldera." Geophysics **62**: 451-65

Ward, D. M. (1978). "Thermophilic methanogenesis in a hot-spring algal-bacterial mat." Appl Environ Microbiol **35**: 1019-26.

Ward, D. M., M. J. Ferris, *et al.* (1998). "A natural view of microbial biodiversity within hot spring cyanobacterial mat communities." Microbiol Molecul Biol Rev **62**: 1353-70.

Watanabe, K., K. Watanabe, *et al.* (2000). "Molecular characterization of bacterial populations in petroleum-contaminated groundwater discharged from underground crude oil storage cavities." Appl Environ Microbiol **66**: 4803–9.

Watanabe, K., Y. Kodama, *et al.* (2002). "Diversity, abundance, and activity of archaeal populations in oil-contaminated groundwater accumulated at the bottom of an underground crude oil storage." Appl Environ Microbiol **68**: 3899-907.

Weisburg, W. G., S. M. Barns, *et al.* (1991). "16S ribosomal DNA amplification for phylogenetic study." J Bacteriol **173**: 697-703.

Weller, R., M. M. Bateson, *et al.* (1992). "Uncultivated cyanobacteria, *Chloroflexus*-like inhabitants, and *Spirochete*-like inhabitants of a hot spring microbial mat." Appl Environ Microbiol **58**: 3964-9.

Wheelis, M. L., O. Kandler, *et al.* (1992). "On the nature of global classification." Proc Natl Acad Sci U S A **89**: 2930-4.

Wieland, H., H. Stan-Lotter, *et al.* (unpublished). Molecular characterization of haloarchaeal populations from Permo-Triassic rock salt, University of Salzburg, Austria.

Williams, W. D. (1991). "Chinese and Mongolian saline lakes: a limnological overview." Hydrobiologia **210**: 39-66.

Williams, W. D. and J. E. Sherwood (1994). "Definition and measurement of salinity in salt lakes." Int J Salt Lake Res **3**: 53-63.

Williams, W. D. (1996). "The largest, highest and lowest lakes of the world: Saline lakes." Verh. Internat. Verein. Limnol. **26**: 61-79.

Windberger, E., R. Huber, *et al.* (1989). "*Thermotoga thermarum* sp. nov. and *Thermotoga neapolitana* occurring in African continental solfataric springs." Arch Microbiol **151**: 506-12.

Witte, A., U. Baranyi, *et al.* (1997). "Characterisation of *Natronobacterium magadii* phage φCh1, a unique archaeal phage containing DNA and RNA." Molec Microbiol **23**: 603-16.

Woese, C. R. and G. E. Fox (1977). "Phylogenetic structure of the prokaryotic domain: the

primary kingdoms." Proc Natl Acad Sci U S A **74**: 5088-90.

Woese, C. R., R. Gutell, *et al.* (1983). "Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids." Microbiol Rev **47**: 621-69.

Woese, C. R. (1987). "Bacterial evolution." Microbiol Rev **51**: 221-71.

Woese, C. R., O. Kandler, *et al.* (1990). "Towards a natural system of organisms: proposal for the domains Archaea, Bacteria and Eucarya." Proc Natl Acad Sci U S A **87**: 4576-9.

Wood, A. P. and D. P. Kelly (1991). "Isolation and characterisation of *Thiobacillus halophilus* sp. nov., a sulphur-oxidising autotrophic eubacterium from a Western Australian hypersaline lake." Arch Microbiol **156**: 277-80.

Wright, A.-D. G. (2006). "Phylogenetic relationships within the order *Halobacteriales* inferred from 16S rRNA gene sequences." Int J Syst Evol Microbiol **56**: 1223-7.

Wu, Q. L., G. Zwart, *et al.* (2006a). "Bacterioplankton community composition along a salinity gradient of sixteen high-mountain lakes located on the Tibetan plateau, China." Appl Environ Microbiol **72**: 5478-85.

Wu, X. L., M. W. Friedrich, *et al.* (2006b). "Diversity and ubiquity of thermophilic methanogenic archaea in temperate anoxic soils." Environ Microbiol **8**: 394-404.

Wuyts, J., Y. Van der Peer, *et al.* (2002). "The European database on small subunit ribosomal RNA." Nucleic Acids Res **30**: 183-5.

Xiao, Y.-K., W.-G. Liu, *et al.* (2000). "Variations in isotopic compositions of chlorine in evaporation-controlled salt lake brines of Qaidam Basin, China " Chinese J Oceanol Limnol **18**: 169-77.

Xiao, C., S. Zhou, *et al.* (unpublished). Direct submission to NCBI, Zhongshan (Sun Yat-Sen) University, Guangdong, China.

Xin, N., J. Sun, *et al.* (1994). "International study on *Artemia*. New Survey of *Artemia* resources in the People's Republic of China." Int J Salt Lake Res **3**: 105-12.

Xin, H., T. Itoh, *et al.* (2001). "*Natronobacterium nitratireducens* sp. nov., a haloalkaliphilic archaeon isolated from a soda lake in China." Int J Syst Evol Microbiol **51**: 1825-9.

Xu, Y., Z. Wang, *et al.* (2001). "*Natrialba hulunbeirensis* sp. nov. and *Natrialba chahannaoensis* sp. nov., novel haloalkaliphilic archaea from soda lakes in Inner Mongolia Autonomous region, China." Int J Syst Evol Microbiol **51**: 1693-8.

Xu, J. (unpublished). 16S ribosomal RNA of mdw-1 which can degrade methomyl, Life Science, Jiangsu, China.

Xu, X. W. and M. Wu (unpublished). An extreme halophile in the family *Halobacteriaceae*, Zhejiang University, China.

Xue, Y., H. Fan, *et al.* (2005). "*Halalkalicoccus tibetensis* gen. nov., sp. nov., representing a novel genus of haloalkaliphilic archaea." Int J Syst Evol Microbiol **55**: 2501-5.

Xue, Y., Y. Xu, *et al.* (unpublished). Taxonomical position of a thermophilic bacterium, *Caldotoga fontana* gen. nov., sp.nov, Chinese Academy of Sciences, Beijing, China.

Yakimov, M. M., L. Giuliano, *et al.* (2001). "*Alcalilimnicola halodurans* gen. nov., sp. nov., an alkaliphilic, moderately halophilic and extremely halotolerant bacterium, isolated from sediments of soda-depositing Lake Natron, East Africa Rift Valley." Int J Syst Evol Microbiol **51**: 2133-43.

Yamamoto, S. and S. Harayama (1995). "PCR amplification and direct sequencing of gyrB genes with universal primers and their application to the detection and taxonomic analysis of *Pseudomonas putida* strains." Appl Environ Microbiol **61**: 1104-9.

Yamamoto, H., A. Hiraishi, *et al.* (1998). "Phylogenetic evidence for the existence of novel thermophilic bacteria in hot spring sulfur-turf microbial mats in Japan." Appl Environ Microbiol **64**: 1680-7.

Yang, Y., H.-L. Cui, *et al.* (2006). "*Halobacterium jilantaiense* sp. nov., a halophilic archaeon isolated from a saline lake in Inner Mongolia, China." Int J Syst Evol Microbiol **56**: 2353-5.

Yannish-Peron, C., J. Vierra, *et al.* (1985). "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors." Gene **33**: 103–19.

Yeon, S.-L., W.-J. Jeong, *et al.* (2005). "The diversity of culturable organotrophic bacteria from local solar salterns." The J Microbiol **43**: 1-10.

Yim, L. C., J. Hongmei, *et al.* (2006). "Highly diverse community structure in a remote central Tibetan geothermal spring does not display monotonic variation to thermal stress." FEMS Microbiol Ecol **57**: 80-91.

Yu, G., S. P. Harrison, *et al.* (2001). Lake status records from China: Data base document.

Zeikus, J. G. and R. S. Wolfe (1972). "*Methanobacterium thermoautotrophicus* sp. n., an anaerobic, autotrophic, extreme thermophile." J Bacteriol **109**: 707-13.

Zeikus, J. G., P. W. Hegge, *et al.* (1983). "Isolation and description of *Haloanaerobium praevalens* gen. nov. and sp. nov., an obligately anaerobic halophile common to Great Salt Lake sediments." Curr Microbiol **9**: 225-33.

Zeng, R. (unpublished). Phylogenetic analysis of cultured bacteria in the deep sea sediment of the east Pacific, Third Institute of Oceanography, Xiamen, China.

Zeng, R. Y., N. W. Lin, *et al.* (unpublished-a). Study of low temperature active chitinase secreted by three psychrophilic bacteria isolated from Antarctic deep sea sediment,

Third Institute of Oceanography, Fujian, China.

Zeng, Y., H. Li, *et al.* (unpublished-b). Protease-producing bacteria isolated from Arctic seawater, Polar Research Institute of China, Shanghai, China.

Zepp Falz, K., C. Holliger, *et al.* (1999). "Vertical distribution of methanogens in the anoxic sediment of Rotsee (Switzerland)." Appl Environ Microbiol **65**: 2402-8.

Zhang, W., Y. Xue, *et al.* (2002a). "*Marinospirillum alkaliphilum* sp. nov., a new alkaliphilic helical bacterium from Haoji soda lake in Inner Mongolia Autonomous Region of China." Extremophiles **6**: 33-7.

Zhang, W., Y. Xue, *et al.* (2002b). "*Salinicoccus alkaliphilus* sp. nov., a novel alkaliphile and moderate halophile from Baer Soda Lake in Inner Mongolia Autonomous Region, China." Int J Syst Evol Microbiol **52**: 789-93.

Zheng, X., M. Zhang, *et al.* Salt lakes in Inner Mongolia of China. Science Press, Beijing, 1992.

Zheng, M., J. Tang, *et al.* (1993). "Chinese saline lakes." Hydrobiologia **267**: 23-36.

Zheng, M., Y. Zhao, *et al.* (2000). "Palaeoclimatic indicators of China's Quaternary." Acta Geologica Sinica **74**: 259-65.

Zhijie, L. "The Rehai (Hot Sea) Geothermal System in Tengchong County, Yunnan Province, China." In Terrestrial Heat Flow and Geothermal Energy in Asia. Eds: M. L. Gupta and M. Yamano. Rotterdam, 1995.

Zhilina, T. N., L. V. Miroshnikova, *et al.* (1991). "*Halobacteroides lacunaris* sp. nov. - a new saccharolytic anaerobic extremely halophilic organism from lagoonic hypersaline Lake Chokrak." Mikrobiologiya **60**: 714-24.

Zhilina, T. N., G. A. Zavarzin, *et al.* (1992). "Ecology, physiology and taxonomy studies on a new taxon of *Haloanaerobiaceae*, *Haloincola saccharolytica* gen. nov., sp. nov." Syst Appl Microbiol **15**: 275-84.

Zhilina, T. N., G. A. Zavarzin, *et al.* (1996a). "*Natroniella acetigena* gen. nov. sp. nov., an extremely haloalkaliphilic, homoacetic bacterium: a new member of *Haloanaerobiales*." Curr Microbiol **32**: 320-6.

Zhilina, T. N., G. A. Zavarzin, *et al.* (1996b). "*Spirochaeta alkalica* sp. nov., *Spirochaeta africana* sp. nov., and *Spirochaeta asiatica* sp. nov., alkaliphilic anaerobes from the continental soda lakes in central Asia and the East African Rift." Int J Syst Bacteriol **46**: 305-12.

Zhilina, T. N., G. A. Zavarzin, *et al.* (1997). "*Desulfonatronovibrio hydrogenovorans* gen. nov., sp. nov., an alkaliphilic, sulfate-reducing bacterium." Int J Syst Bacteriol **47**:

144-9.

Zhilina, T. N., E. N. Detkova, *et al.* (1998). "*Natronoincola histidinovorans* gen. nov., sp. nov., a new alkaliphilic acetogenic anaerobe." Curr Microbiol **37**: 177-85.

Zhilina, T. N., R. Appel, *et al.* (2004). "*Alkaliflexus imshenetskii* gen. nov. sp. nov., a new alkaliphilic gliding carbohydrate-fermenting bacterium with propionate formation from a soda lake." Arch Microbiol **182**: 244-53.

Zhilina, T. N., V. V. Kevbrin, *et al.* (2005). "*Clostridium alkalicellum* sp. nov., an obligately alkaliphilic cellulolytic bacterium from a soda lake in the Baikal Region." Microbiology **74**: 557-66.

Zhou, J., M. A. Bruns, *et al.* (1996). "DNA recovery of soils of diverse composition." Appl Environ Microbiol **62**: 316-.

Zillig, W., D. Prangishvili, *et al.* (1996). "Viruses, plasmids and other genetic elements of thermophilic and hyperthermophilic archaea." FEMS Microbiol Rev **18**: 225-36.

Zillig, W. and A. Reysenbach. "Class IV. Thermococci *class. nov.*" In Bergey's Manual of Systematic Bacteriology. Eds: D. R. Boone, R. Castenholz and G. Garrity. 2nd ed. New York, Springer-Verlag, 2001.