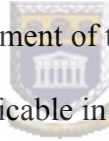


**EXPLORING DIVERSITY AND ECOLOGY OF  
NANOARCHAEA IN HYDROTHERMAL BIOTOPES**

By

**Ncebakazi Galada**

A thesis submitted in fulfilment of the requirements for the degree of  
Magister Scientiae, as applicable in the Department of Biotechnology



University of the Western Cape.

Supervisor: **Prof. D.A. Cowan**

**ACADEMIC DISSERTATION IN BIOTECHNOLOGY**

November 2005

## DECLARATION

I declare that *Exploring diversity and ecology of Nanoarchaea in hydrothermal biotopes* is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete references.

Ncebakazi Galada



November 2005

Signed: .....

# LIST OF SUPERVISORS

---

Supervisor

**Professor Don A. Cowan**

Department of Biotechnology

University of the Western Cape

Co-supervisor I

**Dr. Gillian Baker**

Brighton and Sussex Medical School

University of Sussex

Co-supervisor II

**Dr. Ana Casanueva**

Department of Biotechnology

University of the Western Cape



**University of the Western Cape**

**2005**

## **ACKNOWLEDGEMENTS**

---


---

This work was carried out at the Department of Biotechnology, at the Advanced Research Centre for Applied Microbiology (ARCAM), University of the Western Cape. Scanning electron microscopy was conducted at the University of Cape Town. The culturing experiments were conducted at Bioscience Foundation laboratory, Santiago, Chile (during a month visit). Financial assistance was provided by National research Foundation (NRF).

With great pleasure I would humbly like to express my sincere gratitude to:



- ❖ The Almighty God for giving me a chance and strength to take up this project. Through it all, I have learned to trust in Him and depend upon His Word.
- ❖ Professor Chris Gehring, who opened doors of learning for me at UWC.
- ❖ Supervisor and project leader - Professor Donald Cowan for his warm welcome into his multicultural and diverse research group, ARCAM. I am sincerely grateful to him for believing in me, his patience, and support and for grooming and allowing me to learn to the level I am today.
- ❖ First co-supervisor, Dr. Gillian Baker, who initiated this project and thought me the molecular techniques, her patience, availability and encouragement are greatly appreciated.

- ❖ Second co-supervisor, Dr. Ana Casanueva, who always made me see the light at the end of the tunnel. I appreciate her willingness to help and efforts in making things easier for me. Needless to count, she has done so much for me; I would not trade her for anything in this world.
  
  - ❖ The research team at ARCAM, for being the good people they are, it would have not been the same without them. Through them I have learned so much, from an academic point of view to a personal level. Knowing them was a lifetime experience, and as long as we have memories, yesterday remains.
- 
- ❖ My friends and colleagues, Joseph Lako, for being there always when I needed him most. Konanani Rashamuse, for making me laugh even at times when I felt the world was crushing down on me, and for his assistance and lectures. No one could ever replace him; he is one of a kind. To Tsepo Tsekoa, for advice and encouragement.
  
  - ❖ A special thanks to my family and friends, Nozolile Galada, Enoch Galada, Noxolo, Xolani, Xolile, Bomikazi, Babalwa, Lubabalo, Lulama, Sinethemba, Mbali, Nokubonga, Tabita Nkopo, Kamohelo Direko, Simphiwe Hlongwane, and Gcobisa Dudeni. Thank you for being the pillar of strength and support always, may God richly and abundantly bless you.



*To my family, amaNtshilibe amahle  
ooBhanqo, ooMdumane, ooBobese*

# **PUBLICATION**

---

The work presented on this thesis has been submitted for publication on *Microbial Ecology*

Title: New Nanoarchaeal-Specific Primers and their Use in the Identification of Novel Nanoarchaeotes in Geographically Diverse Hydrothermal Sediments

Date of submission: 31 October 2005

## **The author's contribution**

---

N. Galada conducted the experiments based on Chinese and New Zealand samples, analysed and interpreted the results under supervision of Prof. D.A. Cowan, the project leader, co-supervised by Dr. G. Baker and Dr. Ana Casanueva. Dr. J. Blamey supervised the culturing experiments during a month visit in Santiago, Chile.

Dr. G. Baker initiated this work and left at the end of her contract with University of the Western Cape.

Dr. Ana Casanueva took over the co-supervisory position after Dr. G. Baker. She conducted the experiments on Chilean samples, analyzed, interpreted the results and wrote the paper.

---




---

## CONTENTS

---



---

<b>TITLE</b> .....	<b>i</b>
<b>DECLARATION</b> .....	<b>ii</b>
<b>LIST OF SUPERVISORS</b> .....	<b>iii</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>iv</b>
<b>DEDICATION</b> .....	<b>vi</b>
<b>PUBLICATION</b> .....	<b>vii</b>
<b>ABSTRACT</b> .....	<b>xi</b>
<b>LIST OF ABBREVIATIONS</b> .....	<b>xiii</b>
<b>LIST OF FIGURES</b> .....	<b>xvi</b>
<b>LIST OF TABLES</b> .....	<b>xviii</b>
	
<b>1. INTRODUCTION</b> .....	<b>1</b>
1.1 Ecology of Archaea.....	1
1.2 Diversity of Archaea .....	2
1.3 Metabolism of Archaea .....	4
1.4 Biotechnological applicability of hyperthermophiles .....	7
1.5 The symbiotic or parasitic relationship between <i>Nanoarchaeum equitans</i> and <i>Ignicoccus</i> sp.....	8
1.6 Genomics of <i>Nanoarchaeum equitans</i> .....	9
1.7 Identification methods by molecular analysis.....	11
1.7.1 Polymerase Chain Reaction (PCR).....	12
1.7.2 Amplified Ribosomal DNA Restriction Analysis (ARDRA) .....	16
1.7.3 Fluorescence <i>In Situ</i> Hybridization (FISH).....	18



<b>2. MATERIALS AND METHODS .....</b>	<b>23</b>
2.1 Sampling and DNA extraction.....	23
2.1.1 Chinese samples.....	23
2.1.2. New Zealand samples .....	25
2.1.3. DNA extraction.....	26
2.2. Polymerase Chain Reaction and Cloning.....	27
2.2.1 Amplification of archaeal 16S rRNA gene.....	27
2.2.2 Preparation of “Ultra-Competent” Cells.....	28
2.2.3 Transformation.....	29
2.2.4 Colony PCR .....	30
2.3 Amplified Ribosomal DNA Restriction Analysis (ARDRA).....	30
2.4 Phylogenic tree construction.....	31
2.5 Culturing conditions.....	32
2.6 Microscopy.....	34
2.6.1 Light Microscopy.....	34
2.6.2 Fluorescence <i>In Situ</i> Hybridization (FISH).....	35
2.6.3 Scanning Electron Microscopy (SEM) .....	36
<b>3. RESULTS AND DISCUSSION .....</b>	<b>38</b>
3.1 DNA extraction from Chinese and New Zealand hydrothermal sediments.....	38
3.2 Amplification of 16S rRNA gene .....	42
3.3 Construction of 16S rRNA amplicon libraries from sample TC-9, TC-11, TOK-3 and TOK-17. ....	43
3.4 Amplified Ribosomal DNA Restriction Analysis (ARDRA).....	44
3.5 Amplification of nanoarchaeotes and phylogenetic analysis.....	58
3.6 Culturing and Microscopy .....	73
3.7 General Discussion .....	80
3.7.1 PCR.....	80
3.7.2 Amplified rDNA restriction analysis (ARDRA).....	80
3.7.3 Phylogenetic analysis.....	81

3.7.4 Microscopy.....	81
<b>REFERENCES.....</b>	<b>84</b>
<b>APPENDIX.....</b>	<b>90</b>



## ABSTRACT

The Nanoarchaeota were proposed as the fourth archaeal sub-division in 2002, and the only fully characterized nanoarchaeon was found to exist in a symbiotic association with the crenarchaeote, *Ignicoccus* sp. This nanoarchaeote, named *Nanoarchaeum equitans* could not be detected with “universal” archaeal 16S PCR primers and could only be amplified using specifically designed primers. In order to identify and access a wide diversity of archaeal phylotypes a new set of “universal” archaeal primers A571F (5'-GCY TAA AGS RIC CGT AGC-3') and UA1204R (5'-TTM GGG GCA TRC IKA CCT-3') was designed, that could amplify the 16S rRNA genes of all four archaeal sub-divisions. Using these primers community DNA was amplified from Chinese and New Zealand hydrothermal systems. Upon sequencing of amplicons it was discovered that Chinese and New Zealand samples contained novel nanoarchaeal phylotypes. The preliminary nanoarchaeal phylotypes were used to design nanoarchaeal-specific primer N989R (5'-GGT TTC CGG TGT CAG TTC-3'), which was coupled with A571F and used in screening of nanoarchaeotes. The nanoarchaeal phylotypes identified with these primers were further screened by amplified ribosomal DNA restriction analysis (ARDRA), which was used to explore the diversity of these phylotypes. The novel nanoarchaeotes cluster into 9 closely related clades which may represent separate species. Three of the New Zealand phylotypes form one separate clade which is closely related to the published nanoarchaeotes. The following nanoarchaeal sequences were submitted to the GenBank, TC9F (AY572420),

TC11-5 (AY571283), TC11-B6 (AY727890), TC11-B7 (AY727887), TC11-C4 (AY727886), TC11-C6 (AY727889), TC11-C8 (AY727888), AND TC11-D4 (AY727891).

Fluorescence *in situ* hybridization was also used to simultaneously visualize, identify and localize nanoarchaeotes.

***Keywords:* Archaea, Nanoarchaeota, *Nanoarchaeum equitans*, Diversity, Ecology, PCR, ARDRA, FISH**



## LIST OF ABBREVIATIONS

---

---

**ARDRA:** Amplified rDNA restriction analysis

**BLAST:** Basic Local Alignment Search Tool

**bp:** Base pairs

**BSA:** Bovine serum albumin

**CDS:** Coding DNA sequence

**CTAB:** Cetyltrimethylammonium bromide

**DAPI:** 4',6-diamidine-2-phenylindole dihydrochloride

**DMSO:** Dimethylsulfoxide

**DNA:** Deoxynucleic acid

**dNTPs:** Deoxynucleotide tri-phosphate



***E. coli:*** *Escherichia coli*

**EDTA:** Ethylenediaminetetra-acetate

**FISH:** Fluorescence *In situ* Hybridization

**g:** Gram

**× g:** Gravitational force

**GenBank:** Nucleotide sequence database

**IPTG:** Isopropyl-β-D-thiogalactoside

**IPA:** Iso-propyl alcohol

**k:** Kilo (10<sup>3</sup>)

**kb:** Kilobase

**L:** Liter

**LB:** Luria Bertani

**LBA:** Long Branch Attraction

**LGT:** Lateral Gene Transfer

**m:** Milli ( $10^{-3}$ )

**M:** Molar

**MM:** Master media

**min:** Minute

**n:** Nano ( $10^{-9}$ )

**nm:** Nanometer

**NCBI:** National center for Biotechnology Information

**O.D:** Optical Density

**PBS:** Phosphate buffered saline



**PCR:** Polymerase chain reaction

**PIPES:** Piperazine-1,2-*bis*[2-ethanesulfonic acid]

**PVPP:** Polyvinylpolypyrrolidone

**rDNA:** Ribosomal deoxynucleic acid

**rpm:** Revolution per minute

**rRNA:** Ribosomal ribonucleic acid

**s:** Second

**SDS:** Sodium dodecyl sulfate

**SEM:** Scanning electron microscope

**TBE:** Tris-borate-EDTA electrophoresis buffer

**T<sub>d</sub>:** Dissociation temperature

**TE:** Tris-EDTA buffer

**TMA:** Tetramethylammonium

**U:** Unit

**UV:** Ultra violet

**V:** Volts

**v/v:** Volume per volume

**v/w:** Volume per weight

**w/w:** Weight per weight

**w/v:** Weight per volume

**X-gal:** 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

**$\alpha$ :** Alpha

**$\beta$ :** Beta

**$\gamma$ :** Gamma

**$\lambda$ :** Lambda

**$\mu$ :** Micro ( $10^{-6}$ )

**$^{\circ}\text{C}$ :** Degrees Celsius



---



---

**LIST OF FIGURES**


---



---

<b>FIGURE 2.1:</b> Map of China .....	<b>24</b>
<b>FIGURE 2.2:</b> The Taupo Volcanic zone (North Island, New Zealand). .....	<b>25</b>
<b>FIGURE 2.3:</b> Anaerobic manifold system.....	<b>33</b>
<b>FIGURE 3.1:</b> Agarose gel electrophoresis of extracted genomic DNA from TC-9 and TC-11 .....	<b>39</b>
<b>FIGURE 3.2:</b> Genomic DNA from New Zealand samples .....	<b>40</b>
<b>FIGURE 3.3:</b> Genomic DNA from TOK 3 and TOK 17.....	<b>41</b>
<b>FIGURE 3.4:</b> 1 % agarose gel electrophoresis of 633 bp amplicons from TOK 3 and TOK 17 amplified with “universal” archaeal primers A571F/U1204R .....	<b>42</b>
<b>FIGURE 3.5:</b> Representatives of agarose gel electrophoresis of amplified 16S rRNA gene.....	<b>44</b>
<b>FIGURE 3.6:</b> Randomly selected clones digested with <i>RsaI</i> and electrophoresed in 2.5 % agarose gels in 0.5 × TBE .....	<b>46</b>
<b>FIGURE 3.7:</b> Randomly selected clones digested with <i>AluI</i> and electrophoresed in 2.5 % agarose gels in 0.5 × TBE.....	<b>47</b>
<b>FIGURE 3.8:</b> Randomly selected clones digested with <i>HaeIII</i> and electrophoresed in 2.5 % agarose gels in 0.5 × TBE.....	<b>48</b>
<b>FIGURE 3.9:</b> Randomly selected clones digested with <i>MboI</i> and electrophoresed in 2.5 % agarose gels in 0.5 × TBE .....	<b>48</b>
<b>FIGURE 3.10:</b> Rarefaction plot of accumulative ARDRA patterns of Chinese clone library.....	<b>50</b>
<b>FIGURE 3.11</b> Rarefaction plot of accumulative ARDRA patterns from the New Zealand clone library. ....	<b>51</b>
<b>FIGURE 3.12:</b> Representation of ARDRA patterns generated from <i>in silico</i> digestion with <i>RsaI</i> using DNAClub .....	<b>53</b>
<b>FIGURE 3.13:</b> Representation of ARDRA patterns generated from <i>in silico</i> digestion with <i>HaeIII</i> using DNAClub. ....	<b>54</b>



**FIGURE 3.14** Representation of ARDRA patterns generated from *in silico* digestion with *AluI* using DNAClub. .... 55

**FIGURE 3.15:** Representation of ARDRA patterns generated from *in silico* digestion with *MboI* using DNAClub. .... 56

**FIGURE 3.16:** Amplicons of nanoarchaeal rDNA with nanoarchaeal primers A571F and N989R on 1 % agarose gel electrophoresis. .... 59

**FIGURE 3.17:** Alignment of 633 bp nanoarchaeal sequences using ClustalW. .... 65

**FIGURE 3.18:** The Taupo Volcanic zone (North Island, New Zealand). .... 66

**FIGURE 3.19:** Nanoarchaeal neighbor-joining phylogenetic tree ..... 67

**FIGURE 3.20:** Light microscope photomicrographs of New Zealand and Chinese environmental samples ..... 76

**FIGURE 3.21:** Cryofixation of New Zealand/Chinese environmental samples ..... 77

**FIGURE 3.22** FISH photomicrographs of New Zealand/Chinese environmental samples ..... 78



## **LIST OF TABLES**

---

---

<b>TABLE 2.1:</b> Environmental conditions of Chinese and New Zealand samples .....	<b>26</b>
<b>TABLE 3.1:</b> Commonly observed crenarchaeal phylotypes .....	<b>70</b>
<b>TABLE 3.2:</b> Commonly observed euryarchaeal phylotypes .....	<b>71</b>
<b>TABLE 3.3:</b> Commonly observed Bacteria phylotypes .....	<b>72</b>
<b>TABLE 3.4:</b> Nanoarchaeal phylotypes from Chinese and New Zealand samples .....	<b>72</b>
<b>TABLE 3.5:</b> Growth analysis of thermophiles from TC9, TC11, TOK3 and TOK17 samples .....	<b>74</b>



# 1. Introduction

## 1.1 Ecology of Archaea

All life forms are divided into three different domains, the *Eucarya*, *Bacteria* and *Archaea* (Woese and Fox, 1977). The archaeal domain harbors organisms living in “extreme” environments according to the context of humankind, but normal to the organisms surviving at these environments. Extremophiles are classified to many categories such as acidophiles, alkaliphiles, halophiles, piezophiles, psychrophiles and thermophiles, and for the purposes of this review only thermophiles will be defined.



Hyperthermophiles are classified as those which are adapted to grow optimally at temperatures ranging from 70-121 °C (Kashefi and Lovley, 2003). Their natural biotopes include water-containing volcanic areas such as terrestrial solfataric fields and hot springs, submarine hydrothermal systems, sea mounts and abyssal hot vents (Huber and Stetter, 1991). Other thermal environments are man-made, and include smoldering coal refuse piles, geothermal power plants and household hot water geysers (Marsh and Norris, 1985; Fuchs *et al.*, 1995). Other environmental factors are also important in maintaining favorable growth conditions, including pH and salinity. For example, most terrestrial hyperthermophiles require low salinity, while those of

marine biotopes are adapted to the higher salt concentration of sea water. In some biotopes such as solfataric fields and hot springs, there may be a combination of low salinity and a wide range of pH values, ranging from pH 0.5-8.5. Marine biotopes contain high salinity but a relatively narrow pH range (pH 5-8.5).

## 1.2 Diversity of Archaea

The Archaea consist of three major phyla, the Crenarchaeota which contain six described orders (*Sulfolobales*, *Thermoproteales*, *Desulfurococcales*, *Caldisphaerales*, *Cenarchaeales* and *Nitrosopumilales*); the Euryarchaeota which is made up of 9 orders (*Archaeoglobales*, *Halobacteriales*, *Methanobacteriales*, *Methanococcales*, *Methanomicrobiales*, *Methanosarcinales*, *Methanopyrales*, *Thermococcales* and *Thermoplasmatales*) and more than 50 genera and the Korarchaeota which only contain environmental DNA sequences (Huber and Stetter, 1998). A fourth phylum, Nanoarchaeota was proposed due to its unique features such as small cell size (only 400 nm in diameter) and very small genome size, which is 490 kb (Huber *et al.*, 2002).

The Crenarchaeota are less diverse than the Euryarchaeota, with one group made up of thermophiles and hyperthermophiles that are phylogenetically distinct from

physiologically similar Euryarchaeotes. *Ignicoccus* species, the first known nanoarchaeal host, belongs to the Crenarchaeota. *Ignicoccus islandicus*, *Ignicoccus pacificus* and other environmental *Ignicoccus* species are crenarchaeotes belonging to the Desulfurococcaceae family. *I. islandicus* and *I. pacificus* were the first representatives, isolated from submarine hydrothermal systems in the Kolbeinsey Ridge north of Iceland. *Ignicoccus* species are Gram-negative, motile, coccoid cells with cell diameters between 1.2 and 3  $\mu\text{m}$ . The other six groups of Crenarchaeota are known by DNA sequences retrieved from low temperature environments (Kvist *et al.*, 2005; Dawson *et al.*, 2001). The phylum Korarchaeota was assigned based on non-culturable species that do not cluster with either Crenarchaeota or Euryarchaeota, and their properties are not known (Barns *et al.*, 1996). The Nanoarchaeota are only represented by one culturable species, *Nanoarchaeum equitans*. *N. equitans*, the first archaeal parasite known to date (Waters *et al.*, 2003) is a hyperthermophilic, obligate symbiont that only grows in co-culture with a “normal-sized” archaeon, *Ignicoccus* species. Two other nanoarchaeal phylotypes have been reported, CU-1 and OP-9, based on environmental rDNA sequences (Hohn *et al.*, 2002). Other nanoarchaeal phylotypes have been detected around the globe from different hyperthermal biotopes and their 16S rDNA sequences are readily available on GenBank (<http://www.ncbi.nlm.nih.gov>).

Hyperthermophilic archaea also exhibit different cell morphologies and sizes. This variation in morphology ranges from irregular cocci, lobed or wedge-shaped irregular discs with ultraflat areas, and regular rods or some with spheres protruding at their

ends (“golf clubs”). Some cells may be irregular with variations in shape and size. For example *Thermoplasma* species (Stetter, 1996) lacks a true cell wall. *N. equitans* cells are known to be the smallest identified to date (Huber *et al.*, 2000). Based on observations on universal phylogenetic tree, hyperthermophiles appear to be the most primitive organisms, representing phylogenetically distant groups (Huber and Stetter, 1998).

### 1.3 Metabolism of Archaea

Hyperthermophilic Bacteria and Archaea consist of anaerobic and aerobic chemolithoautotrophs and heterotrophs, the latter being able to utilize various polymeric substrates such as starch, hemicellulose, proteins and peptides (Leuschner and Antranikian, 1995). Sulfolobales are known to be the only Archaea that have an aerobic metabolism. Generally, most hyperthermophiles are anaerobic due to the low oxygen solubility at high temperatures and the presence of reducing gases such as H<sub>2</sub>S. The mode of nutrition among hyperthermophiles is also diverse. Examples are the chemolithoautotrophs, which utilize inorganic materials such as H<sub>2</sub>, ferrous irons and reduced sulfur compounds as energy sources. These organisms use CO<sub>2</sub> as the sole carbon source, and oxidize sulfur compounds. Nitrate, ferric iron, CO<sub>2</sub> and O<sub>2</sub> may serve as electron acceptors. Some hyperthermophiles are facultative or obligate heterotrophs using organic compounds as energy and carbon sources.

Within the Crenarchaeota, acidophilic extreme thermophiles are made up of strict and facultative aerobes. They belong to the genera *Sulfolobus*, *Metallosphaera*, *Acidianus* and *Desulfurolobus* (Stetter, 1996). The *Sulfolobus* are strict aerobes that can grow chemolithotrophically by oxidizing  $H_2S$  or  $S^0$  to  $H_2SO_4$  and using  $O_2$  as the terminal electron acceptor. Some members of *Thermoproteales* are facultative anaerobes that use elemental sulfur under both aerobic and anaerobic conditions (Seegerer *et al.*, 1986). Some members of *Pyrobaculum* are able to grow autotrophically by anaerobic reduction of  $S^0$  with  $H_2$  as electron donor (Nissen *et al.*, 1992) but also able to grow heterotrophically by sulfur respiration (Stetter, 1996). Some strains of *Thermophilum* and *Pyrobaculum* are obligate heterotrophs; they grow by sulfur respiration using different organic substrates.



Hyperthermophilic Euryarchaeota are more diverse than Crenarchaeota. They range from facultative aerobic chemoorganotrophs to strictly anaerobic heterotrophs (Holt, *et al.*, 1994). *Archaeoglobus* is the only true sulfate reducer. These organisms couple the oxidation of glucose, lactate, pyruvate and  $H_2$  to the reduction of sulfate. *Ferroglobus* is an iron-oxidizing chemolithotroph, and couples the oxidation of  $Fe^{2+}$  to  $Fe^{3+}$  to the reduction of  $NO_3^-$  to  $NO_2^-$ . Thermococcales are anaerobic heterotrophs that grow at a temperature range of 70 to 105 °C, depending on the species they belong to. *Pyrococcus* and *Thermococcus* (Huber and Stetter, 2001; Itoh, 2003) and *Palaeococcus* (Takai *et al.*, 2000) members are metabolically similar, and are obligate heterotrophs growing on organic substrates, usually in the presence of elemental sulfur which is reduced to hydrogen sulfide. Depending on the species, they utilize peptides, yeast extract, carbohydrates and some organic acids as carbon and energy source.

As previously stated, the phylum Nanoarchaeota is represented by only a single culturable isolate, *N. equitans*. Understanding the metabolism of *N. equitans* is important in order to gain knowledge on its symbiotic relationship with *Ignicoccus*, which has the ability to grow independent of *N. equitans*, while *N. equitans* depends on propagating *Ignicoccus* cells for growth (Huber *et al.*, 2003). Both *N. equitans* and *Ignicoccus* are hyperthermophilic obligate anaerobes. *Ignicoccus* is known to grow chemolithoautotrophically by sulfur reduction using molecular hydrogen as an electron donor (Huber *et al.*, 2000). H<sub>2</sub>S is produced as a metabolic end product of the co-culture and of the pure *Ignicoccus*. This means that *Ignicoccus* is responsible for assimilation of nutrients and metabolite production. In contrast to the host, *N. equitans* lacks genes that support chemolithoautotrophy (Waters *et al.*, 2003). Though *N. equitans* lacks most of its biosynthetic capacity, few genes have been identified that code for enzymes responsible for metabolism, such as a multifunctional prephenate dehydrogenase/chorismate mutase/ prephenate dehydrogenase and amidotransferase used in aromatic amino acid biosynthesis. In addition, *N. equitans* contains two enzymes that code for amino acid oxidative deamination: aminotransferase and glutamate dehydrogenase. Furthermore, there are also a limited number of enzymes that could be responsible for electron transfer reactions such as ATP synthase. It is, however, still not known whether *N. equitans* can produce adenosine triphosphate (ATP) by electron transport phosphorylation or whether it derives energy from its host, as it was discovered that *N. equitans* lacks most if not all of carbon assimilation genes, such as genes for glycolysis/gluconeogenesis, pentose phosphate pathway and tricarboxylic acid cycle. This supports the possibility that energy is derived from its host.



## 1.4 Biotechnological applicability of hyperthermophiles

Hyperthermophiles are major sources of thermal stable enzymes that have found industrial applicability as biocatalysts for applications such as biotransformation. These thermostable enzymes include both intracellular and extracellular enzymes (Koch *et al.*, 1991). Exploring and discovering the diversity of hyperthermophiles will enhance detection of novel enzymes with novel activities and applications. Furthermore, these enzymes can be studied in order to understand their stability and can be engineered to create novel biocatalysts with enhanced stability and altered specificity (Hough and Danson, 1999). Due to their unique characteristics, thermostable enzymes are used in biotechnological processes, such as the use of hemicellulases in paper pulp bleaching (Hongpattarakere, 2002). Amylases, pullulanases and glucosidases can be exploited in the starch processing industry, for the production of valuable products such as dextrin, glucose, fructose and trehalose (Schiraldi *et al.*, 2002). Thermostable DNA polymerase isolated from *Thermus aquaticus* (Saiki *et al.*, 1988) have been used extensively in PCR, a technique that have advanced molecular biology in amplifying DNA. In biohydrometallurgy, hyperthermophiles such as *Sulfolobus*, *Acidianus* and *Metallosphaera* can be used in leaching of ores (Huber and Stetter, 1998). As they gain industrial applicability, one of the major advantages would be the ability to retain their activity even after several hours at high temperatures. For future purposes, continuous detection of novel hyperthermophiles can aid in discovering novel thermostable enzymes, which can be used to overcome limitations caused by thermally unstable enzymes in industrial processes.

## 1.5 The symbiotic or parasitic relationship between *Nanoarchaeum equitans* and *Ignicoccus* sp.

In any given ecosystem, organisms interact with each other and with their environment in order to survive. There is a series of different kinds of interactions such as mutualism, symbiosis, commensalism, competition and parasitism. Symbiosis is an interaction between two organisms living together in more or less intimate association or even the merging of two dissimilar organisms. The various forms of symbiosis include parasitism, mutualism and commensalism. For the purpose of this review, only the symbiotic or parasitic relationship between Nanoarchaeota and *Ignicoccus* species will be considered.



In a parasitic interaction, the association is disadvantageous or destructive to one of the organisms (Margulis & Sagan, 1986). Although *Ignicoccus* survives well in the presence of *N. equitans*, a high density of nanoarchaeal cells inhibits host growth (Waters *et al.*, 2003). *Ignicoccus* is regarded as the host because of its large size compared to its symbiont *N. equitans* which is attached to it by means of an appendage (Huber *et al.*, 2003), suggesting ectosymbiosis. Furthermore, studies of the *N. equitans* genome suggest a parasitic lifestyle, due to its inability to grow on its own (Waters *et al.*, 2003). Its minimal genome suggests a limited biosynthetic and catabolic capacity, for *N. equitans* lacks genes for lipid, cofactor, amino acid or nucleotide biosynthesis. Therefore, *N. equitans* is the only archaeal parasite known to date (Waters, *et al.*, 2003). In contrast to typical genomes from parasitic/symbiotic

microbes, the genome of *N. equitans* does not show any evidence of decaying genes and contains a full complement of genes encoding informational proteins (Brochier *et al.*, 2005). In conclusion, though there is no evidence of decaying genes from *N. equitans* genome, based on the growth conditions of *N. equitans* and *Ignicoccus* and the biosynthetic information from the *N. equitans* genome, the relationship between the two microorganisms clearly indicates a parasitic-symbiotic association.

## 1.6 Genomics of *Nanoarchaeum equitans*

Phylogenetic analysis of 16S rRNA sequences suggests that *N. equitans* forms a basal archaeal lineage. Due to its information processing system (replication, transcription and translation) it does not qualify as primitive archaea. In order to be eligible as a primitive archaeon it is required to possess the central metabolism capacity, primary biosynthesis and bioenergetics apparatus, and such properties are absent in *Nanoarchaeum* genome (Waters *et al.*, 2003). Currently the position of *N. equitans* is under scrutiny. It has been suggested that *N. equitans* may more probably be the representative of a fast evolving euryarchaeal lineage possibly related to Thermococcales, rather than the representative of a novel and early diverging archaeal phylum (Brochier *et al.*, 2005). The early emergence of *N. equitans* may be due to a Long Branch Attraction (LBA) artifact caused by the fast evolutionary rate. It is believed that the adaptation of *N. equitans* to symbiotic/parasitic life style might have accelerated its evolutionary rate. Lateral Gene Transfers (LGTs) may also have

contributed. Therefore the hypothesis that *N. equitans* is a novel phylum is opposed by the hypothesis that *N. equitans* is likely to be a very divergent euryarchaeon, possibly a sister of Thermococcales (Brochier *et al.*, 2005). Only 18.3 % of the coding DNA sequence (CDS) has homologs in other organisms. The remaining 81.7 % is unique to *N. equitans*. 552 coding DNA sequence with an average length of 827 bp were identified. These CDS comprise of single copies of 5S, 16S, 23S rRNA and 38 tRNA's. In addition to the metabolism related enzymes mentioned previously (section 3.1), 11 proteases and peptidases, DNA repair enzymes and recombination enzymes, endonucleases, helicase and diphosphatase have been detected (Waters *et al.*, 2003). *N. equitans* possesses information processing enzymes such as DNA-dependent RNA polymerase and RNA-modifying enzymes. In contrast to other archaeal genomes, *N. equitans* possesses a high number of split genes whose gene product is encoded by two unlinked coding DNA sequences. In total, *N. equitans* contains 553 genes. It is known to be a common feature of obligate parasitic and symbiotic bacteria to have a small genome size. This has developed by the elimination of unwanted genes in combination with restricted uptake of new genes by lateral gene transfer (Ochman *et al.*, 2001). It is suggested that *N. equitans* genome may be a derived one, through adaptation to its host, but also a genetically stable parasite that diverged anciently from the archaeal lineage. Due to the fact that it lacks most of its genes, it is not surprising that *N. equitans* possesses the smallest genome known to date.

## 1.7 Identification methods by molecular analysis

Reliable classification of microbial isolates into species and strains is required in order to analyze and monitor microbial diversity. The emerging use of molecular biology in microbial ecology has revealed that less than 1% of the microorganisms present in the environments are readily cultivable (Amann *et al.*, 1995). Molecular techniques allow less biased access to a greater portion of uncultivable microorganisms than conventional cultivation methods that give a bias representation of the microbial diversity (Robe *et al.*, 2003). Current molecular approaches rely essentially on the detection and analysis of the small subunit ribosomal RNA genes. These molecular identification methods can be grouped into two categories: molecular probes (e.g. FISH) and DNA fragment analysis (e.g. ARDRA). The former group requires prior knowledge of the targeted microbial population, whereas the latter does not necessitate prerequisite knowledge of microbial community structure before its study, but requires DNA extraction from the sample and PCR amplification with universal primers. Whatever approach is chosen, either molecular probes or DNA fingerprinting, the techniques always describe the dominant species of microbial communities. They cannot detect a population which represents less than 1/100 of the ecosystem. Subdominant species, which are probably also important for ecosystems adaptability, remain undetectable.

### 1.7.1 Polymerase Chain Reaction (PCR)

PCR is a reaction where a specific DNA sequence is amplified in a series of repeated cycles. It has been used extensively to amplify gene fragments from mixed microbial communities. Gene products derived from amplification products should contain defined fragments that can be cloned and sequenced from known priming sites. In order for the PCR to be successful, the targeted gene or sequence must be complementary to the primer sequence, therefore prior knowledge of targeted sequence is required to design the primer set to be used. Primers are small synthetic oligonucleotides that hybridize to the complementary DNA sequence. The forward and reverse primer are coupled together and used to amplify a specific DNA sequence in a PCR reaction. It is important for these primers to anneal exclusively to the target DNA in order to enhance specificity (Lexa *et al.*, 2001).

#### Primer design

The initial step when designing primers would be to retrieve sequences from the database that are homologous to the gene of interest, followed by sequence alignment using an appropriate sequence alignment program such as BioEdit. The next step would be locating at least two conserved regions that are at least 200 bp apart, such that the target sequence is contained between them. In order to achieve this, a variety of parameters must be considered when designing primers to improve primer quality. These include melting temperature, string-based alignment scores for

complimentarity, primer length, G-C content, repeat rich regions and target length (Haas *et al.*, 1998). Both forward and reverse primers should share similar thermodynamic properties, such as melting temperature, to ensure uniform hybridization and performance. Too high an annealing temperature will result in insufficient primer-template hybridization and consequently low amplicon yield. Too low an annealing temperature will result in non-specific binding that is caused by a higher number of base pair mismatches (Rychlik *et al.*, 1990). The G-C content should be between 40 and 60 % (Lowe *et al.*, 1990) to ensure an optimal balance. A G-C content lower than 20 % and higher than 80 % should be avoided. This will simultaneously enable stable specific binding and efficient melting (Li *et al.*, 1997). The GC content of the amplified DNA can also modulate the *Taq* polymerase activity and the exponential amplification of the mixture of DNA fragments may result in ratio discrepancies between the amplified DNA fragments and the original mixture. This results in semi-quantitative data (Dabert *et al.*, 2002). A pairwise comparison of primers should be conducted in order to avoid self-complimentarity of primers that will result in the formation of secondary structures or primer dimers (Kampke *et al.*, 2001). If the designed primers meet the required conditions, a primer pair is used to search the target DNA sequence. The 3' end of one primer must not bind to any site of the other primer (Proutski and Holmes, 1996). Lastly, the designed primer pair should achieve perfect complementarity matching the query sequence and it should minimize cross-hybridization with non-target sequences within the considered genome (Schretter and Milinkovitch, 2005). The selected primer pair can either be specific to the target DNA sequence or they can be “universal”. Universal primers are

complementary to nucleotide sequences which are conserved in a particular set of DNA molecules, and are able to bind to a wide variety of DNA templates. These are crucial because they significantly enlarge the template diversity, but selective priming may discriminate against certain sequences (Wintzingerode *et al.*, 1997). The selected “universal” or specific primers can be either stringent or degenerate. Degenerate primers are primers which have a number of options at several positions so as to allow annealing to and amplification of a variety of related sequences (Kwok *et al.*, 1994). Degenerate primers can be used to screen multiple sequences that share conserved regions (Wei *et al.*, 2003). Degeneracies obviously reduce the specificity of the primer(s), thus increasing mismatch opportunities and decrease primer concentration. The region between the primers should not be too far apart in order to increase efficiency of amplification. It should be noted that with degenerate primers, a limited number of primer molecules are complementary to the template.

### **PCR advantages and drawbacks**

PCR is one of a range of sensitive molecular identification methods, and its sensitivity can be advantageous and simultaneously disadvantageous. Its sensitivity allows even very small amounts of DNA to be amplifiable. Old and degraded DNA can often serve as a template. The disadvantage of this sensitivity is that small amounts of contaminating DNA can also be amplified. PCR amplification can also be biased (Farrelly *et al.*, 1995; Suzuki and Giovanonni 1996). The so called “universal



primers” cannot be totally universal (Zheng *et al.*, 1996). As more novel genotypes are identified, the so called “universal” primers become biased and fail to amplify some of the novel sequences. This indicates that the amplifiable sequences are not a true representation of what exist in the ecosystem, and more microbial representatives have been overlooked due to primer bias. Therefore, it is important to continuously update or edit the existing “universal” primers in order to incorporate the novel sequences. This will result in less amplification bias and improved representation of the studied diversity. An example of this is shown by the recently discovered *N. equitans*, which failed to amplify with the archaeal “universal” primers (Huber *et al.*, 2002). Consequently, new archaeal primers were designed that could amplify all four archaeal phyla: Crenarchaeota, Euryarchaeota, Korarchaeota and Nanoarchaeota (Baker *et al.*, 2003). There are still more undiscovered phylotypes that need to be screened in order to understand archaeal evolution, and this requires continuous assessment of primer selection used. One of the biases during PCR is due to selective priming; certain sequences may be discriminated (Amann *et al.*, 1995). Other PCR problems encountered may be due to primer design, formation of chimeric PCR products and contamination. The design of primers requires prior knowledge of the microorganism’s genome and depends on the availability of database sequences that represent the variants of the microorganisms. Therefore, the choice of primers used is very crucial in order to reduce bias. Other factors that may contribute to PCR bias include, annealing temperature and number of cycles, inhibition of polymerase by humic compounds and formation of chimeric PCR products. Primer annealing efficiency and specificity influence the PCR amplification. Non-specific binding

reduce amplification efficiency, therefore it is required of the universal primers to have uniform hybridization efficiency to guarantee the amplification of all target 16S rDNAs (Brunk *et al.*, 1996). To avoid rapid purification steps, additives such as bovine serum albumin (BSA) (Kreader, 1996), tetramethylammonium (TMA) (Kovárová and Dráber, 2000) and betaine can be added.

Regardless of the bias from PCR, it remains a crucial method for rapid amplification of DNA fragments. Continuous assessment of primers used and proper optimization of PCR conditions must be applied in order to obtain adequate results (Dieffenbach and Dveksler, 1995).



### **1.7.2 Amplified Ribosomal DNA Restriction Analysis (ARDRA)**

ARDRA is a molecular identification tool based on restriction endonuclease digestion of amplified 16S rDNA. It allows microbial characterization of the community structure in different environments (Gich *et al.*, 2000). ARDRA gives little or no information about the type of microorganisms present in the sample, but can be used to assess genotypic changes within the microbial community. This technique is based on the repetitive units of the nuclear ribosomal DNA (rDNA), which are amplified by PCR. The PCR amplicons are digested by restriction endonucleases, and the generated fragments separated according to their size using gel electrophoresis

(Schramm and Amann, 1999). Depending on the position of the restriction sites, bands of different numbers and sizes appear and form patterns, which can be used to rapidly screen large numbers of isolates. These patterns can be compared to those obtained from the database, or the DNA of the respective patterns sequenced in order to assign isolates to species or clusters (Stackebrandt and Lang, 1998). Although the 16S rDNA sequences of species may be almost identical, ARDRA is able to emphasize the few differences present without the need for extensive 16S rDNA sequence analysis (Stakenborg *et al.*, 2005). It should be noted though, that few base differences may not alter the ARDRA pattern, unless the mutation alters the restriction site. For the effectiveness of this technique, the choice of restriction enzyme is crucial in order to obtain optimum resolution. To avoid limitations, preliminary tests for the enzyme choice must be performed to ensure the highest resolution in detecting changes in communities. This can be done by preceding screening of restriction enzymes by simulating restriction sites in the available 16S rDNA sequences of interest. This will make it possible to select an appropriate enzyme combination for the analysis of a group of related organisms (Heyndrickx *et al.*, 1996). In addition to improving resolution, the typical analysis of restriction digest of isolates or clones is performed on relatively low resolution agarose gels. Thus, for community analysis the potentially large number of fragments can be resolved by using polyacrylamide gels to improve resolution and to produce a community-specific pattern.

ARDRA is PCR dependent and therefore requires cell disruption for nucleic acid extraction. This makes it impossible to associate the amplification results to a specific cell or measure the percentage of cells that contain the target sequence (Komminoth and Long, 1995). Nevertheless this technique is rendered effective for identifying phylogenetic groups in a high diverse community. It can simultaneously process numerous isolates in a very short time and yield valuable information on the similarity of isolates. Furthermore, appropriate isolates can be selected for more physiological and phylogenetic investigations. In summary, ARDRA can be used to study genetic similarities, genetic structure of a particular phylogenetic group or functional group and the impact of environmental disturbance on particular biotype (Ranjard *et al.*, 2000).



### **1.7.3 Fluorescence In situ Hybridization (FISH)**

Fluorescence *in situ* hybridization (FISH) allows the detection of specific DNA sequences in intact cells and chromosomes. The detection is by a fluorescently labeled probe that hybridizes specifically to its complementary target sequences within the intact cell. The principal steps of FISH involve alignment of rRNA sequences, identification of conserved regions to design the probes, synthesis and labeling of complimentary nucleic acid probes and finally experimental evaluation and optimization of the probe specificities and assay sensitivities. The procedure includes the following steps: (i) fixation of the specimen; (ii) preparation of the

sample, possibly including specimen pretreatment steps; (iii) hybridization with respective probes for detection of the target sequences; (iv) washing steps to remove unbound probes; (v) mounting and visualization (Moter and Göbel, 2000). FISH allows simultaneous visualization, identification, enumeration and localization of individual microbial cells. Culture based methods do not reflect the exact composition of mixed bacterial communities or microbial diversity (Wagner *et al.*, 1993), but FISH has the advantage of detecting even the unculturable microorganisms and can therefore help in understanding complex microbial communities. rRNA-targeted oligonucleotide probes (chemically synthesized single stranded, short [usually 15-25 nucleotides in length] DNA molecules) can be designed in a directed way. These probes may be targeted to signature sites of the rRNA molecules characteristic for defined taxonomic entities such as species, genera, families, orders or even domains. Sets of probes, therefore, allow for a rapid assignment of cells or rRNA of interest to major groups (Amann and Kuhl, 1998). The choice of probe in FISH must consider specificity, sensitivity and ease of cell penetration. The probes are fluorescently labeled to specifically stain different members of microbial communities. The standard probes currently used are rRNA-targeted oligonucleotides with a single fluorescent dye molecule attached to the 5' end. Commonly used dyes are fluorescein-derivatives (excitation wavelength, 490 nm, emission wavelength, 520 nm), such as Fluorescein-Isothiocyanate (FITC), 5-(6-) carboxyfluorescein-*N*-hydroxysuccinimide-ester (FluoX); rhodamine-derivatives (excitation wavelength, 550 nm, emission wavelength, 575 nm), (Tetramethyl-Rhodamine-Isothiocyanate (TRITC), and Texas Red (excitation wavelength, 578 nm, emission wavelength, 600

nm). Recently cyanine dyes like Cy3 and Cy5 have been reported to be superior to the classical dyes because they provide significant brighter staining and are very stable to photobleaching (Moter and Göbel, 2000). Aromatic diamidines like 4',6-diamidine-2-phenylindole dihydrochloride (DAPI) binds non-intercalatively to DNA with great affinity and can be used for Blue fluorescent counterstaining (Zimmer and Wähnert, 1986).

### **Advantages and disadvantages of FISH**

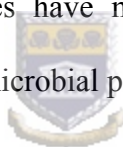
FISH has been widely used for the identification of microorganisms in microbial communities. It makes use of 16S rRNA, and permits probe design for different specificity levels based on known rRNA sequences. These probes can be either domain specific, or can target a certain genera, species, subspecies or even strains (Stahl and Amann, 1991). A key advantage of FISH is its ability to target the sequence of interest. Another advantage is that FISH can be used to identify the “non-culturable” microorganisms, since it does not require prior isolation and enrichment of microorganisms. Furthermore, FISH permits rapid, simple and accurate detection of microorganisms. It does not require extraction of nucleic acids; therefore culturable and “non-culturable” microorganisms can be easily identified and located by FISH from their original environments. The pitfalls associated with FISH include permeabilization problems. Penetration of the probe into the cell is a vital step of FISH, but cell permeability does not guarantee hybridization of the target sequence

(O'Donnell and Head, 1997). The sensitivity of FISH increases with an increased number of active metabolizing cells and a low signal is obtained if low cell concentrations or starving cells are used (Amann *et al.*, 1995). Poor accessibility of the target sequence is possible if the target is in contact with ribosomal proteins or other secondary structures of rRNA. Some of the problems encountered during FISH are non-complementarity of probe and target, ineffective probe labeling, or non-optimal hybridization conditions (e.g. hybridization at a temperature above dissociation temperature  $[T_d]$  of the probe target hybrid). One of the ways to circumvent the limitations of *in situ* identification of bacteria is to use molecules other than rRNA for phylogenetic identification of bacteria, for which nucleotide sequences would be sufficiently divergent to design species specific probes, and which would be more accessible to oligonucleotide probes (Schönhuber *et al.*, 2001). Although application of FISH is quick and easy to carry out, thorough controls are crucial to ensure good quality results. Automated Fish processes are underway, and would reduce manual work and allow improved reproducibility (Moter and Göbel, 2000).

## **Applications of FISH**

FISH has been used extensively in the field of environmental diversity research without any selective purification and/or amplification steps. It has been used to investigate microbial communities in natural environments such as aquatic habitats, soils and on root surfaces (Moter and Göbel, 2000). Huber *et al.* (2002) used FISH to

demonstrate the symbiotic relationship between *Ignicoccus* species as host and *Nanoarchaeum equitans* as symbiont, using oligonucleotide probes based on the *N. equitans* 16S rRNA sequence. It has also found application in analyzing environments like bioreactors and treatment plants (Amann *et al.*, 1996). Applicability of FISH can be improved with combination of other techniques that give additional information on the functional state of the organism, like gene expression and metabolism. An example of this is the use of FISH with microsensor measurements. Such a combination has permitted the study of changes in population structure and growth of biofilms over time (Schramm *et al.*, 1999). FISH combined with microautoradiography also allows visualization of metabolic activity (Lee *et al.*, 1999). Such combined techniques have made it possible to attribute a specific metabolic activity to a particular microbial population *in situ*.



It is widely accepted that molecular detection methods can overcome some of the limitations of conventional culturing techniques. However, any individual method is subject to some degree of bias. In order to reduce bias and to improve experimental accuracy it is preferable to use a combination of detection methods. In this study a combination of PCR, ARDRA, and FISH methods have been used to explore the diversity of Nanoarchaea in hydrothermal biotopes.



## 2. Materials and Methods

### 2.1 Sampling and DNA extraction

#### 2.1.1 Chinese samples

Chinese hydrothermal sediments were sampled from Tengchong, Rehai thermal area, located in the Yunnan Province (Grant *et al.*, 2003) with co-ordinates of 24°51'-24°58'N and 98°23'-98°28'E and an elevation of 1.098-1.912 m (Vaupen, 1999). The region consists of a number of geothermal sites, such as boiling springs, steam jets and steaming ground surfaces with alkaline hydrothermal fluids. Underground temperatures are estimated to be more than 150 °C and less than 270 °C for the heated zone (Grant *et al.*, 2003).



Samples were obtained under the auspices of the joint Eu-SA-China MGAtch Project (Grant *et al.*, 2003). Samples TC9 and TC11 used in this study were collected from hydrothermal sites at pH 8 and 82-84 °C, and pH 8 at 78°C, respectively (Table 2.1). Temperatures were measured using a Solomat 520C temperature monitor or a Hanna K-type thermocouple (Jencons, Leighton Buzzard, UK), and pH was determined with narrow range pH papers (Merk).

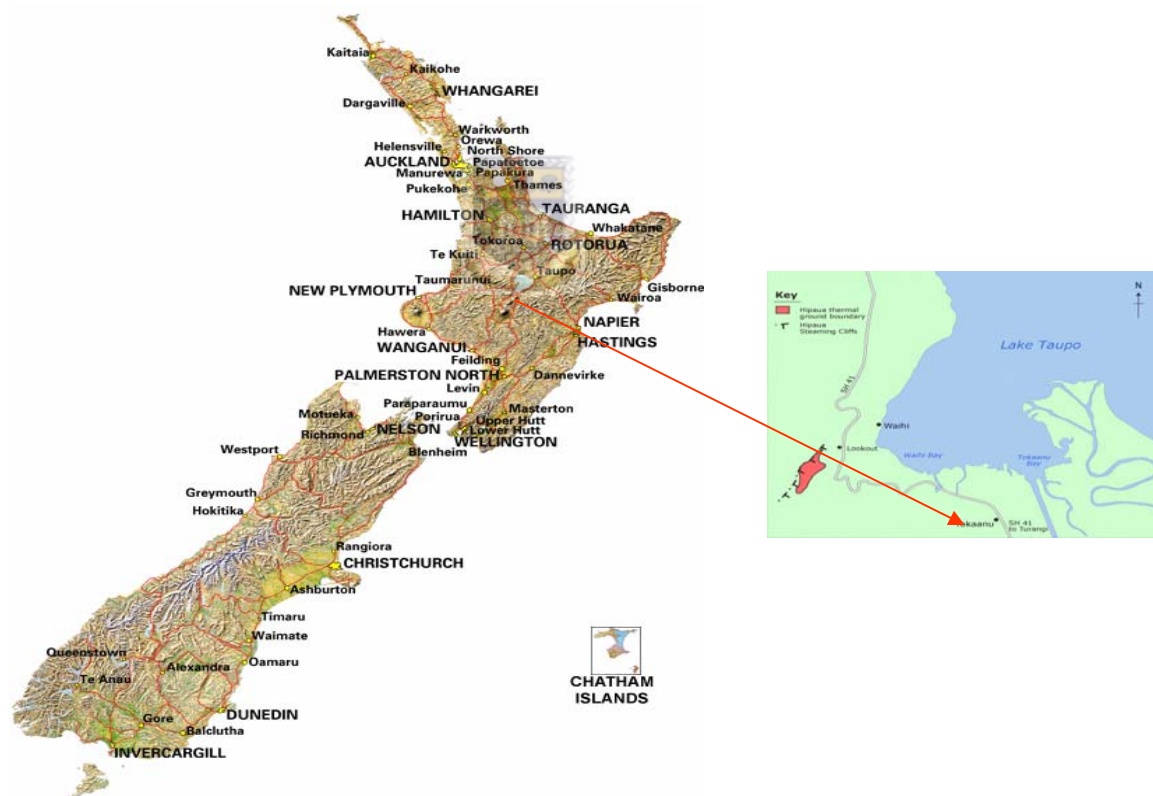


**Fig 2.1:** Map of China, showing the sampling area in Tengchong (TC), Yunnan province. Rehai is located 10.5 kilometers south of Tengchong.

Rehai is one of the unique geothermal areas with mostly alkaline hydrothermal fluids, a characteristic which is rare compared to world wide geothermal locations that are usually acidic (Grant *et al.*, 2003).

### 2.1.2. New Zealand samples

Samples were collected from Tokaanu geothermal site, North Island, New Zealand. Tokaanu region forms part of Taupo volcanic zone (Fig. 2.2, Cooper, 2002), consisting of a number of geysers, sinters deposits, hot springs and pools, steaming cliffs, fumaroles, steam vents and seepages (Jones *et al.*, 2004). Samples TOK-3 and TOK-17 were used in this study. Table 2.1 shows the environmental conditions of the sampling sites.



**Fig. 2.2:** The Taupo Volcanic zone (North Island, New Zealand). The red arrow indicates the Tokaanu geothermal area near Lake Taupo.

**Table 2.1:** Environmental conditions of Chinese and New Zealand sampling sites in terms of temperature and pH.

Sample	°C	pH
TC-9	82-84	8
TC-11	78	8
Tok 3	86.5	6.5
Tok 17	96	6.0

### 2.1.3. DNA extraction

Genomic DNA from a 0.5-1.0 g of soil or sediment was extracted using a scaled-down modified Zhou method (Stach *et al.*, 2001). In order to avoid heat shock, soil samples stored at -80 °C were transferred to -20 °C and kept for 24-48 h then thawed on ice. Using a sterile spatula, 0.5 - 1 g of sample was weighed aseptically into a sterile 2 ml eppendorf. 0.675 ml freshly prepared extraction buffer (0.1 % Cetyltrimethylammonium bromide (CTAB); 0.45 ml sterile water; 100 mM Tris-Cl pH 8; 100 mM NaH<sub>2</sub>PO<sub>4</sub>; 100 mM Ethylenediaminetetraacetic (EDTA); 1.5 M NaCl and 2.5 µl Protease K (20 mg/ml)) was added. The tubes were shaken horizontally at 37 °C for 30 minutes at 225 rpm. 0.075 ml of 20 % Sodium Dodecyl Sulphate (SDS) was added to each tube and the tubes were incubated at 65 °C for 2 hours with gentle inversion every 20 minutes. The tubes were then centrifuged at 3 000 × g for 10 minutes and the supernatant transferred to sterile eppendorfs. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1, v/v/v) was added to each tube and mixed gently. The tubes were centrifuged at 16 000 × g for 10 minutes. The

supernatants were transferred to new sterile tubes, an equal volume of chloroform added and the tubes re-centrifuged at  $16\ 000 \times g$  for 10 minutes. The chloroform wash was repeated when the supernatant was deeply colored. The supernatant was transferred to a sterile eppendorf and 0.6 volumes of iso-propyl alcohol (IPA) added. Tubes were left overnight at room temperature and centrifuged at  $10\ 000 \times g$  for 10 minutes. The DNA pellet was washed in 70 % ethanol and re-centrifuged at  $10\ 000 \times g$  for 10 minutes, the supernatant discarded and the pellet air-dried in the laminar flow cabinet. The dry pellet was resuspended in 140  $\mu$ l sterile H<sub>2</sub>O and 200  $\mu$ l of Binding Buffer (Roche) and the DNA was further purified using a High Pure PCR Template Preparation Kit (Roche).



## 2.2. Polymerase Chain Reaction and Cloning

### 2.2.1 Amplification of archaeal 16S rRNA gene

“Universal” archaeal primers A571F (5’-GCY TAA AGS RIC CGT AGC-3’) and UA1204R (5’-TTM GGG GCA TRC IKA CCT-3’) (Baker *et al.*, 2003), were used. 1  $\mu$ l of template DNA was amplified using 0.5  $\mu$ M primers A571F/U1204R with 200  $\mu$ M deoxynucleoside triphosphates (dNTPs), 1 mg/ml Bovine Serum Albumin (BSA), 3.5 mM MgCl<sub>2</sub> and 1 U of *Taq* DNA polymerase (Fermentas™) in a total reaction volume of 50  $\mu$ l. Amplifications were performed using a hotstart protocol, which consisted of an initial denaturation at 95 °C for 5 minutes, followed by

denaturation at 95 °C for 1 minute, annealing at 55 °C for 1 minute, and extension at 72 °C for 1 minute. Each reaction consisted of 30 cycles with the last cycle followed by final extension at 72 °C for 10 minutes. Identification of nanoarchaeal sequences using archaeal “universal” primers A571F/U1204R led to the design of a nanoarchaeal-specific primer set A571F and N989R (N989R: 5'-GGT TTC CGG TGT CAG TTC-3'). Preliminary nanoarchaeal sequences obtained and nanoarchaeal sequences from the database were aligned with ClustalW. Conserved regions were identified, and primer N989R was designed, which could amplify a 418 bp region. Primer N989R was coupled with archaeal “universal” primer A571F and used to amplify nanoarchaeal rDNA. The PCR conditions were as previously stated. Primer set A571F/U1204R and A571F/989R produced amplicons of 633 and 418 bp in length, respectively. PCR products were electrophoresed on 1 % agarose gels stained with 0.5 µg/ml ethidium bromide using a *Pst*I digested λ DNA marker. Gels were electrophoresed in 0.5 × TBE buffer at 100 V.

### 2.2.2 Preparation of “Ultra-Competent” Cells

Competent *E. coli* cells were prepared according to Inoue *et al.* (1990). A single, fresh bacterial colony of *E. coli* DH5<sub>α</sub> that has been incubated for 16 hours at 37 °C on a Luria-Bertani (LB) agar plate was picked and transferred into 25 ml LB broth in a 250 ml flask. The culture was incubated for 6-8 hours at 37 °C with vigorous shaking at 250 rpm. This starter culture was used to inoculate three 1 L flasks, each

containing 250 ml of SOB medium (per liter, 20 g tryptone; 5 g yeast extract; 0.5 g NaCl; 2 M MgCl<sub>2</sub>). The three flasks were inoculated with 10 ml, 4ml and 2 ml, respectively. All flasks were incubated overnight at 18 °C with moderate shaking. At an optical density at 600 nm of 0.55, culture vessel was chilled on ice for 10 minutes. Cells were harvested by centrifugation at 2500 × g for 10 minutes at 4 °C. The medium was discarded and the cells resuspended in 80 ml of ice-cold Inoue transformation buffer (55 mM MnCl<sub>2</sub>·4H<sub>2</sub>O; 15 mM CaCl<sub>2</sub>·2H<sub>2</sub>O; 250 mM KCl; 10 mM Piperazine-1,2-bis[2-ethanesulfonic acid] (PIPES, 0.5 M, pH 6.7)), and harvested by centrifugation at 2500 × g for 10 minutes at 4 °C. The medium was decanted and cells resuspended in 20 ml of ice-cold Inoue transformation buffer. 1.5 ml of dimethylsulfoxide (DMSO) was added, the bacterial suspension mixed by swirling and stored on ice for 10 minutes. Aliquots of the suspension were dispensed into pre-chilled microfuge tubes and immediately snap-frozen by immersion in a bath of liquid nitrogen. The tubes were stored at -80 °C until required.

### 2.2.3 Transformation

Amplicons from PCR reactions (Section 2.2.1) were pooled and purified using a GFX<sup>™</sup> PCR purification kit (Amersham). 30-60 ng of purified DNA amplified with A571F and UA1204R was ligated into 55 ng of InsT/Aclone<sup>™</sup> pTZ57R/T vector (Fermentas) and used to transform chemical competent *E. coli* cells. Transformed cells were plated onto Luria-Bertani (LB) agar plates supplemented with ampicillin (100 µg/ml), IPTG (20 µg/ml), and X-Gal (20 µg/ml). Using blue/white colony

selection, white colonies were randomly selected and streaked into fresh LB/ampicillin plates.

#### **2.2.4 Colony PCR**

Transformants were screened by colony PCR using M13F (5'-GTA AAC GAC GGC CAG T-3') and M13R (5'-CAC ACA GGA AAC AGC TAT GAC CAT-3') primers. White colonies were picked randomly and re-streaked in LB/ampicillin (100 µg/ml) media, and incubated at 37 °C overnight. Single colonies were picked using sterile tooth picks and suspended in 40 µl of the 1 × TE buffer. The colonies were boiled at 99 °C for 3 minutes and centrifuged at 16 000 × g for 2 minutes to pellet cell debris. Each supernatant was transferred to new 0.6 ml sterile eppendorf tubes, and 10 µl of supernatant was used on the PCR reaction. PCR conditions were as previously stated (Section 2.2.1), but with annealing temperature of 50 °C. PCR products containing archaeal 16S rDNA inserts were further screened by amplified ribosomal DNA restriction analysis (ARDRA).

### **2.3 Amplified Ribosomal DNA Restriction Analysis (ARDRA)**

Separate restriction digestion of M13 amplicons was performed overnight at 37 °C using *RsaI* (Fermentas), *HaeIII* (Fermentas), *MboI* (Fermentas) and *AluI* (Fermentas).



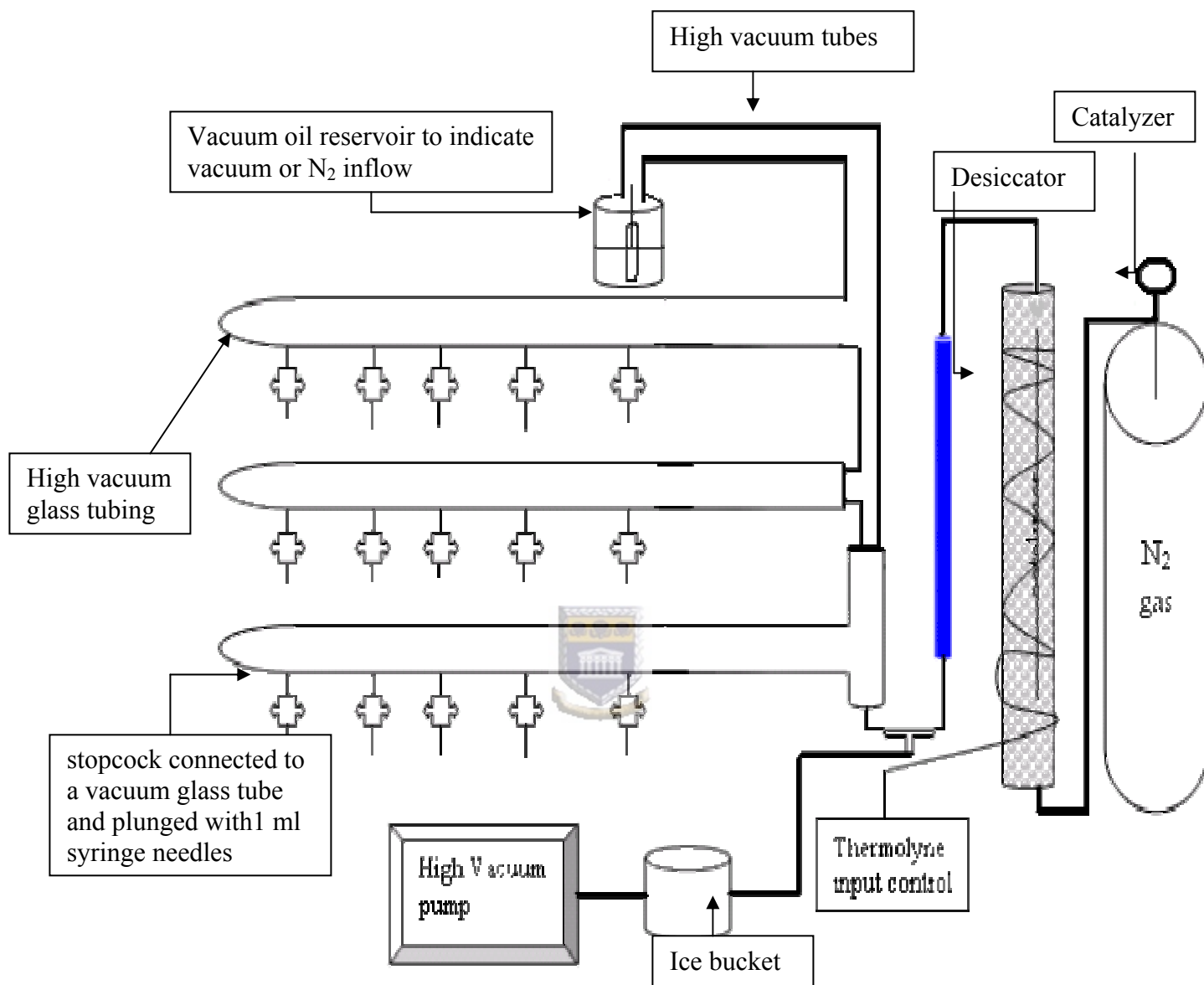
2 U of the restriction enzyme was added with 1.5  $\mu$ l of the corresponding enzyme buffer (10  $\times$  concentrated, final concentration 1  $\times$ ), and 5  $\mu$ l of the amplicon preparation. Each restriction digestion mixture was adjusted to 15  $\mu$ l with sterile distilled water and incubated overnight at 37  $^{\circ}$ C. The DNA restriction fragments were electrophoresed in a 2.5 % agarose electrophoresis gel containing 0.5  $\mu$ g/ml ethidium bromide, against a  $\lambda$  *Pst*I marker. Electrophoresis was performed in 0.5  $\times$  TBE buffer at 100 V, and visualized under ultraviolet (UV) light using a transilluminator imaging system (AlphaImager 2000).

## 2.4 Phylogenic tree construction

Clones revealing unique ARDRA patterns were grown on Luria-Bertani (LB) agar plates containing 100  $\mu$ g/ml of ampicillin and the recombinant plasmids purified using a GFX<sup>TM</sup> Plasmid Miniprep Kit (Amersham). Plasmids were sequenced (using commercial UCT and Inqaba sequencing facilities) with M13 specific primers, and were compared with the sequences in the GenBank using BLASTn (<http://www.ncbi.nlm.nih.gov>). 16S rDNA sequences were uploaded to the Bioedit CLUSTALW multiple alignment program and were aligned to closely related representatives from their taxonomic sub-division. A neighbour-joining tree based on Jukes-Cantor distances (Jukes and Cantor, 1969) with 100 bootstrap replicates was constructed using programs available in the phylogeny inference package TREECON (version 1.3b: Yves Van de Peer, University of Antwerp (UIA), 1994, 1998. [[http://bio-  
www.uia.ac.be/u/yvdp/treeconw.html](http://bio-www.uia.ac.be/u/yvdp/treeconw.html)], Van de Peer, 1997).

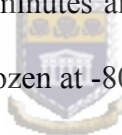
## 2.5 Culturing conditions

In order to obtain host-nanoarchaeum enrichment co-cultures, samples from hydrothermal sites were cultivated anaerobically at high temperatures, using the facilities provided by Dr. J. Blamey, Bioscience Foundation, Santiago, Chile. Chinese (TC9 & TC11) and New Zealand (TOK3 & TOK17) samples were inoculated in two different media: C7.4 (5g/L tryptone, 5g/L yeast extract, 1.25 g/L maltose, 1 g/L  $K_2HPO_4$  and 0.125 g/L cysteine HCl) and MM (contained per liter: 5 g maltose; 5 g tryptone; 5 g yeast extract; 1.2 g  $NH_4Cl$ ; 14 g NaCl; 2 g  $MgSO_4$ ; 3.5 g  $MgCl_2$ ; 0.35 g KCl ; 0.3 g  $CaCl_2$ ; 0.3 g  $KH_2PO_4$ ; 0.006 g NaBr ; 0.015 g  $H_3BO_3$  ; 0.02 g  $SrCl_2$  ; 0.006 g Sodium citrate ; 0.02 g KI; 10 ml Trazas solution (0.22 g  $FeCl_3$ ; 0.1 g  $MnSO_4$ ; 0.06 g  $Na_2WO_4 \cdot 2H_2O$  (Sodium Wolframate); 0.04 g  $NiCl_2$ ; 0.02 g  $CoSO_4$ ; 0.02 g  $ZnSO_4$ ; 0.02 g  $CuSO_4$ ; 0.02 g  $Na_2MoO_4$ ). C7.4 was adjusted to pH 6.0, pH 6.5 or pH 7.5 with NaOH or HCl as appropriate. Cultures were grown in 120 ml serum bottles containing 20 ml of media, and incubated anaerobically at 80 and 85 °C. Anaerobic conditions were maintained using an anaerobic gassing manifold system (Fig. 2.3).



**Fig. 2.3:** Schematic representation of an anaerobic manifold system, connected to a high vacuum pump and a N<sub>2</sub> cylinder.

The manifold was connected to a pressurized N<sub>2</sub> gas cylinder and a high vacuum pump. An in-line catalyser was used to trap traces of unwanted O<sub>2</sub> and the desiccator was used to absorb moisture. With all the stopcocks closed, the manifold was operated by first switching on the vacuum to remove air from the manifold tubes. 1 ml syringe needles were connected to the stopcocks for insertion into air tight serum bottles. The serum bottles were degassed while still hot from the autoclave by alternative N<sub>2</sub> gassing and evacuation, for 3 min each, with the last phase being N<sub>2</sub> gassing. After degassing the serum bottles were inoculated using 1 ml syringe needles and incubated at 80 and 85 °C. Samples were taken daily over a period of 2 days, and were observed by microscopy to monitor growth. To pellet the cells, cultures were centrifuged at 16 000 × g for 10 minutes and supernatants discarded. For long term storage, pelleted cells were kept frozen at -80 °C.



## **2.6 Microscopy**

### **2.6.1 Light Microscopy**

Approximately 10 µl of a culture was placed on a clean sterile microscope slide and covered with a clean coverslip. The cells were observed by light microscope using a 100 × oil-immersion objective (Axioplan 2, Zeiss).

### 2.6.2 Fluorescence *In Situ* Hybridization (FISH)

FISH was applied directly to environmental samples, using archaeal- and nanoarchaeal-specific probes. The archaeal-specific probe, U1115 (5'-YAA CGA GCG AGA YCC-3') and the antisense probe U1115R (5'-CGR TCT CGC TCG TTR-3') were used as positive and negative controls, respectively. A nanoarchaeal-specific probe, N1113 (5'-GGA AAC GAG CGC GAC CCG-3'), labeled with rhodamine, was used to visualize nanoarchaeal cells. *Sulfolobus sulfataricus* was also used as a positive control for archaea and *E. coli* was used as a negative control.

Cells were heat fixed using a Bunsen burner and washed with Phosphate Buffered Saline PBS (137 mM NaCl; 2.7 mM KCl; 10 mM Na<sub>2</sub>HPO<sub>4</sub>; 2 mM KH<sub>2</sub>PO<sub>4</sub>; pH 7.4). 200 µl of hybridization solution containing 50 ng/µl of each probe was added on the slides and incubated at room temperature for 1 hour. Slides were washed twice with PBS and mounted with vectashield (Sigma). Slides were viewed by fluorescent microscope using a 100 × oil-immersion objective (Axioplan 2, Zeiss) with built-in filters for rhodamine, fluorescein and DAPI.

A FISH protocol, modified according to Burggraf *et al.*, (1994), was also used. Approximately 0.1 g sample was fixed overnight in 1 ml 4 % (w/v) PBS buffered paraformaldehyde at 4 °C. After fixation, cells were spotted on a pre-cleaned gelatin

coated (0.1 % gelatin) microscope slides covered with  $\gamma$ -methacryloxypropyl-trimethoxysilane (Sigma) and dried at 46°C for 30 minutes. This was followed by successive dehydration in 50 %, 75 % and 95 % (v/v) ethanol, for 2 minutes each. The slides were air dried and 8  $\mu$ l hybridization solution (10 mM EDTA; 0.1 M NaCl; 0.1 M Tris-HCl (pH 7.2); 0.1% SDS, 50 ng probe) was added in each well of a ten-welled Teflon coated microscope slide and incubated overnight at 45 °C in a moisture chamber. Slides were washed with washing solution 1 $\times$  SET (150 mM NaCl; 20 mM Tris-HCl (pH 8); 1 mM EDTA and 0.01 1% SDS) at room temperature. To reduce stringency, a series of washing solutions was used: 5  $\times$  SET; 3  $\times$  SET and 1  $\times$  SET. The cells were stained with 10  $\mu$ g/ml 4',6-diamidino-2-phenylindole (DAPI) for 5 minutes at room temperature. The cells were immersed in washing solution for 10 minutes at room temperature, rinsed in sterile water and air dried. One drop of Gel Mount™ aqueous mounting medium (Sigma) was added to preserve the probes. Cells were covered with clean coverslips and sealed with Clarion™ (Sigma).

### 2.6.3 Scanning Electron Microscopy (SEM)

#### Cryofixation

Cryofixation was conducted according to Goldstein *et al.* (1992). An aliquot of frozen cultured cells was placed on a cryoplunge and immersed in liquid N<sub>2</sub>. The frozen cells

were then transferred to the cryopreparation chamber and cooled with liquid N<sub>2</sub>. The top surface of the cells which was covered with ice was sliced out to expose the cells to the electron beam and then transferred to the SEM chamber. The cells were sublimated to remove ice and water and were visualized using a Leica S440 scanning electron microscope (UCT Electron Microscope Unit).



### 3. Results and Discussion

In order to explore the genetic diversity of Nanoarchaea in hydrothermal samples from China and New Zealand, genomic DNA was extracted and 16S rDNA amplified and cloned to construct clone libraries. These 16S rDNA libraries were screened by amplified ribosomal DNA restriction analysis (ARDRA) for the presence of nanoarchaeotes and other existing phlotypes. The presence of nanoarchaeal phlotypes was confirmed by sequencing and sequence comparison through BLASTn. Further phylogenetic analysis clustered the Chinese and New Zealand nanoarchaeal sequences with the published nanoarchaeotes from the database.

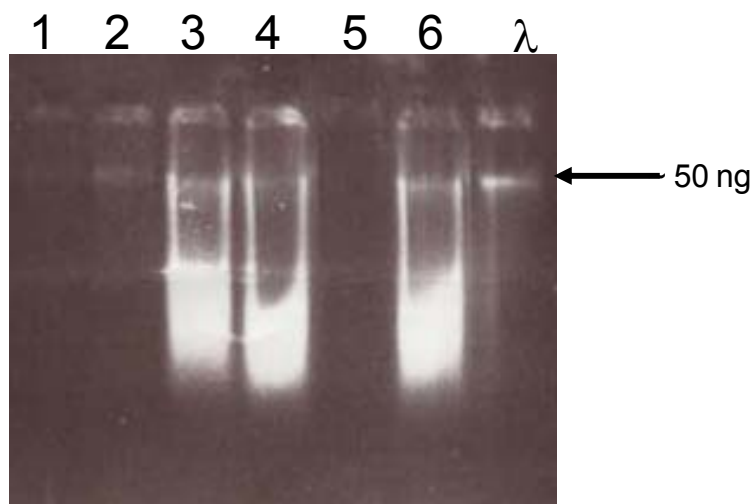


#### 3.1 DNA extraction from Chinese and New Zealand hydrothermal sediments

Metagenomic DNA was extracted from two geological different locations, China (TC-9 and TC-11) and New Zealand (TOK-3 and TOK-17). About 50 ng of genomic DNA was extracted from 0.5 g replicates of TC-9 and TC-11 sediments (Fig. 3.1). The extractions were repeated in order to obtain enough amplifiable DNA for further analysis. Despite evidence of substantial DNA shearing, some high molecular weight DNA was visible (Fig. 3.1). Yields of high molecular weight DNA was quantified by comparisons with sample of uncut lambda DNA, used as a marker. Not all extractions yielded observable DNA on the agarose gels; therefore, the DNA from each sample

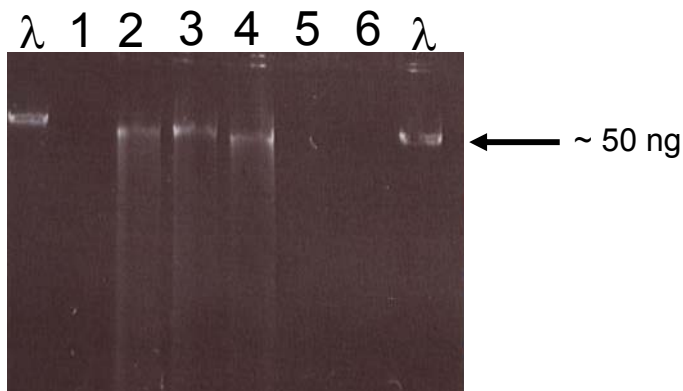


was pooled together and used as a template for polymerase chain reaction (PCR) amplification.



**Fig. 3.1:** Agarose gel electrophoresis of extracted genomic DNA from TC-9 and TC-11. Lane 1 & 2: TC-9, Lane 3 - 6: TC-11 and Lane λ: 50 ng uncut λ DNA.

DNA extraction from very high temperature New Zealand samples was difficult. DNA was extracted within a range of 48 °C to 96 °C, and out of 6 samples extracted only three yielded observable and amplifiable DNA. The lower temperature of these samples was 48, 54.3 and 61.5 °C. From the samples with temperatures of 75, 86.5 (TOK-3) and 96 °C (TOK-17) DNA extraction was unsuccessful (Fig. 3.2). This could be attributed to the soil type and the chemical content of these samples. These were black fine soils with some crystal-like material compared to other samples with sandy soil type.

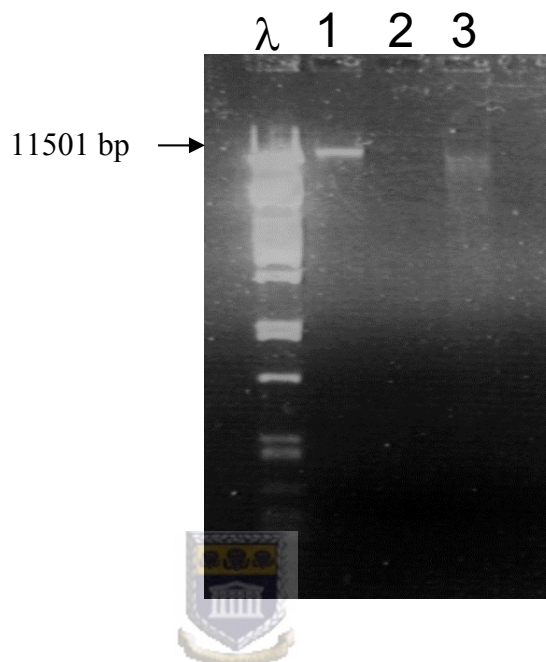


**Fig. 3.2:** 1 % agarose gel electrophoresis of genomic DNA of New Zealand samples from Tokaanu geothermal sites. Lane  $\lambda$ : uncut lambda DNA, Lane 1-6: TOK-3, TOK-7, TOK-9, TOK-10, TOK-13 and the negative control, respectively.

Clay soils are fine-textured soils, and have a high cation exchange capacity because they are negatively charged and can attract, retain and exchange cations. High cation exchange capacity increases the soils buffering capacity; therefore, this effect may affect the pH of the soil by resisting changes in pH and consequently DNA extraction. DNA is negatively charged and the presence of cations in the soil may retain the nucleic acids and result to poor or no DNA yield.

After several attempts of DNA extraction, a small amount of high molecular weight DNA was extracted from TOK-3 (Fig. 3.3, lane 1) and TOK-17 (Fig. 3.3, lane 3), using the Zhou method (Stach *et al.*, 2001). The method was modified by increased incubation and precipitation time, and decreased elution volume. TOK-3 DNA yield was 48.7 ng/ $\mu$ l and TOK-17 yielded a lower concentration of 10 ng/ $\mu$ l (Fig. 3.3). Due to the contaminants such as humic acids, the DNA pellet showed a brown color. The

crude DNA was purified with polyvinylpolypyrrolidone (PVPP), and was ready for further analysis.

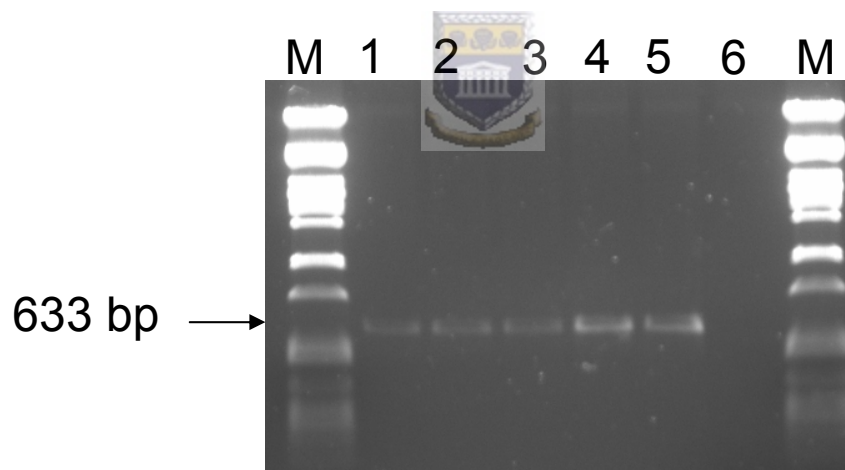


**Fig. 3.3:** 1 % agarose gel electrophoresis of genomic DNA from TOK 3 and TOK 17 using a large scale modified Zhou method (Stach *et al.*, 2001). Lane λ: *Pst*I digested lambda DNA, Lane 1-3: TOK-3, TOK-13 and TOK-17, respectively.

A small amount of DNA was lost during the purification step; the purified DNA was free of contaminants and could be amplified. The PVPP purification method is simple, easy to use and less costly compared to other purification methods used.

### 3.2 Amplification of 16S rRNA gene

Polymerase Chain Reaction (PCR) amplification of 16S rDNA sequences was used to assess environmental microbial diversity of the two samples from Chinese and New Zealand hydrothermal sediments. Initially, Nanoarchaea were detected using PCR amplification of total genomic DNA. The archaeal domain-specific primer set A571F and U1204R was used to generate a 633 bp fragment amplified from position 571 to 1204 according to the *E. coli* numbering. All samples yielded positive results in PCR assays using primers that targeted archaeal 16S rRNA genes (Fig. 3.4).

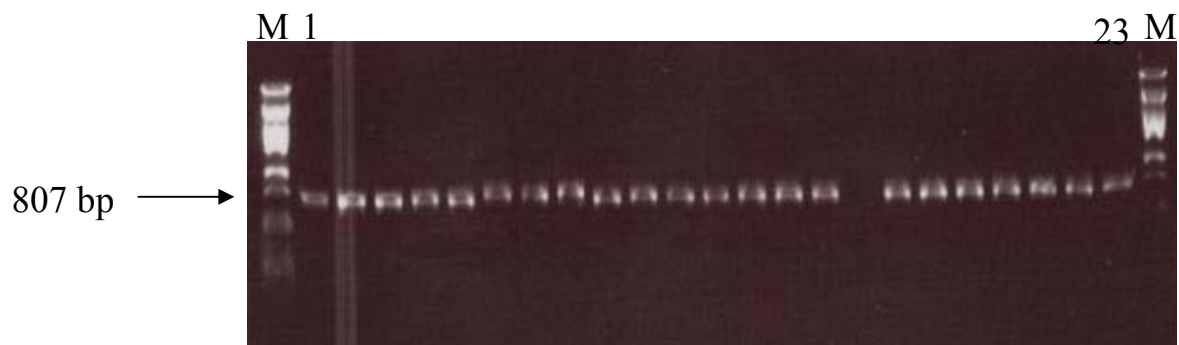


**Fig. 3.4:** 1 % agarose gel electrophoresis of 633 bp amplicons from TOK 3 and TOK 17 amplified with “universal” archaeal primers A571F/U1204R. Lane  $\lambda$ : *Pst*I digested lambda DNA, lane 1-3: TOK-3, lane 4-5: TOK-17 and lane 6: negative control.

Amplification of genomic 16S rDNA from Chinese and New Zealand samples was successful. The use of archaeal “universal” primers was advantageous, for they could amplify all four archaeal taxa. Thus, amplicons would be expected to contain Nanoarchaea and any other Archaea that would be potential nanoarchaeum hosts. Consequently enhance determination of the microbial diversity of these samples.

### **3.3 Construction of 16S rRNA amplicon libraries from sample TC-9, TC-11, TOK-3 and TOK-17.**

The 16S rDNA amplicons were purified using a gel band purification kit and cloned into the pTZ57R/T vector to construct 16S rRNA clone libraries. Chemically competent DH5 $\alpha$  *E. coli* cells, with transformation efficiency of 10<sup>8</sup>, generated 83 and 76 transformants from TC-9 and TC-11, respectively. TOK-3 and TOK-17 samples yielded 235 and 96 recombinant clones, respectively. M13 PCR primers were used to test if the plasmids contained the correct insert DNA. The constructed libraries were of good quality, with 80 % of the recombinants having the correct insert (Fig. 3.5).



**Fig. 3.5:** Representatives of agarose gel electrophoresis of amplified 16S rRNA gene. Lane M: *Pst*I digested  $\lambda$  DNA, Lane 1-23: 807 bp amplicons generated from the constructed mini libraries.

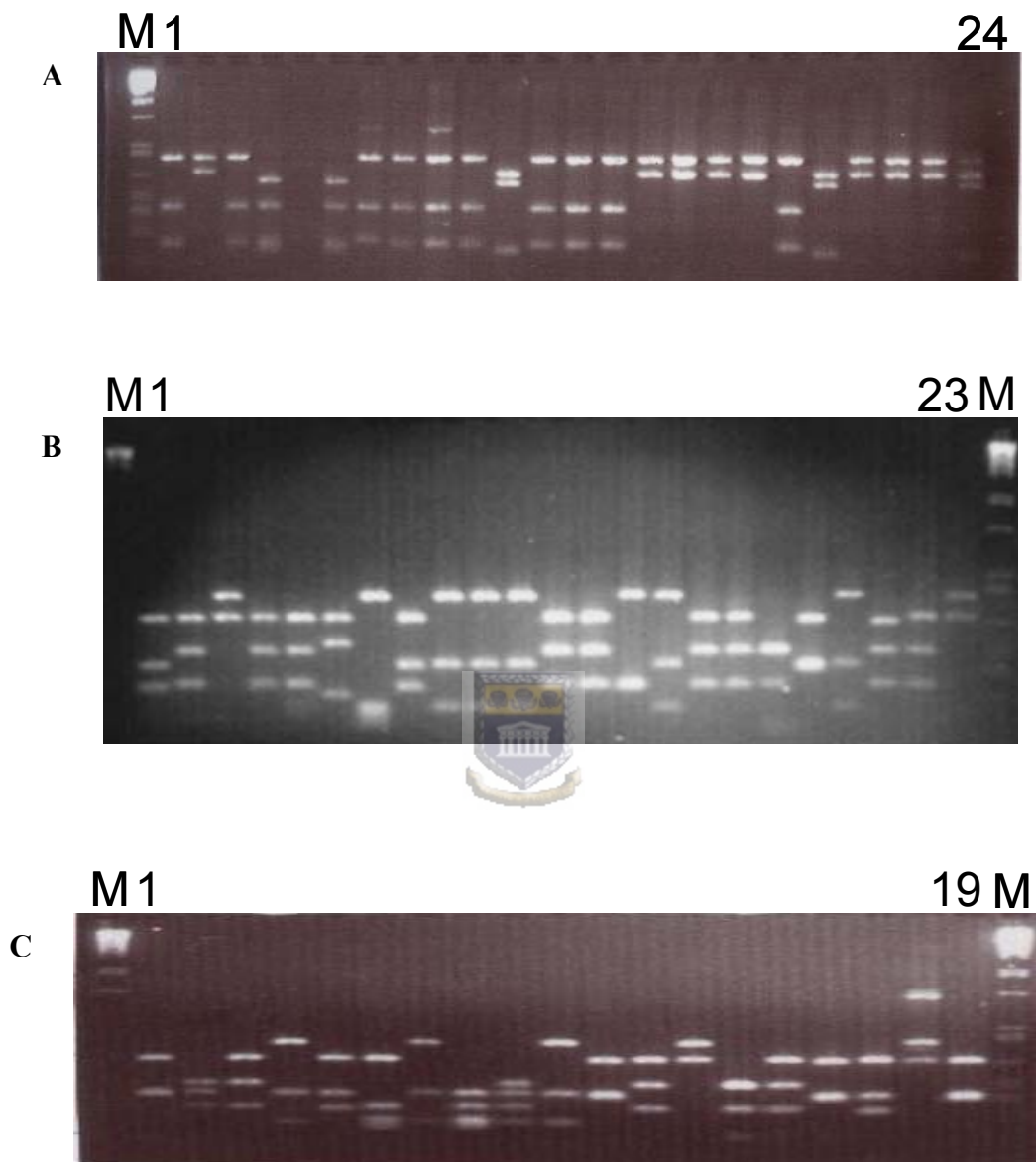
Figure 3.5 shows a representative selection of the amplified 16S rRNA gene cloned . The cloned fragments were amplified with M13 primers, generating an 807 bp fragment (Fig. 3.5). The 807 bp fragment was made up of the 633 bp fragment amplified with “universal” archaeal primers, 73 bp vector sequence to forward and 101 bp vector sequence to reverse. Colony PCR products were subject to ARDRA for further analysis.

### **3.4 Amplified Ribosomal DNA Restriction Analysis (ARDRA)**

Amplified ribosomal DNA restriction analysis (ARDRA) was used to assess the diversity of the rDNA clone libraries. Due to the large number of screened clones, not all the individual clones could be sequenced. ARDRA gives information on the

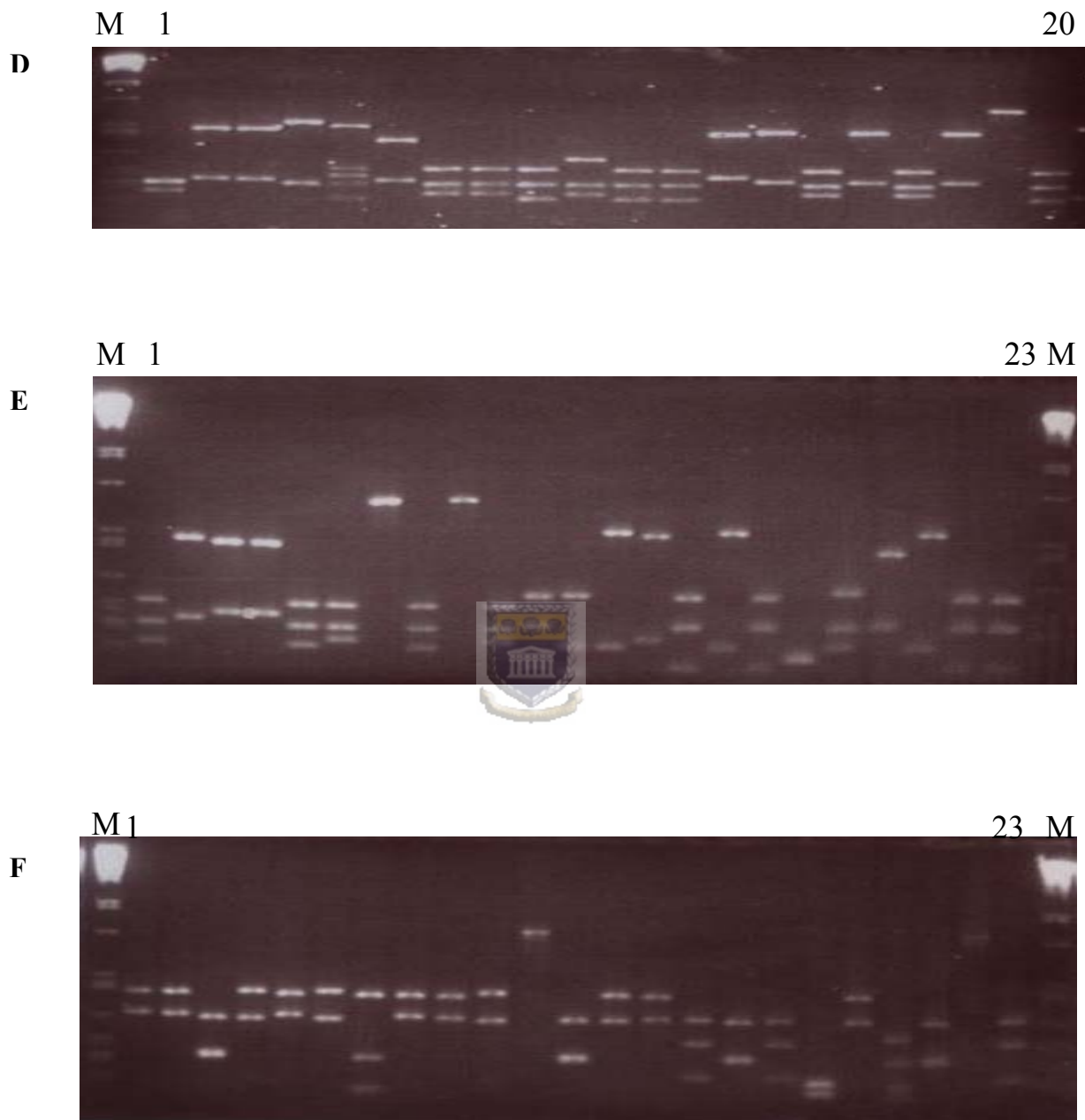
genotypic differences of the sequences but not on the type of microorganisms present. Therefore, ARDRA makes it easy to differentiate the clones according to the patterns they generate. The clones that share the same pattern can be clustered together, and one or two of the representatives sequenced in order to assign the clones to a taxonomic group. Figures 3.6 to 3.9 show the ARDRA patterns generated from *RsaI*, *AluI*, *HaeIII* and *MboI* digestions.



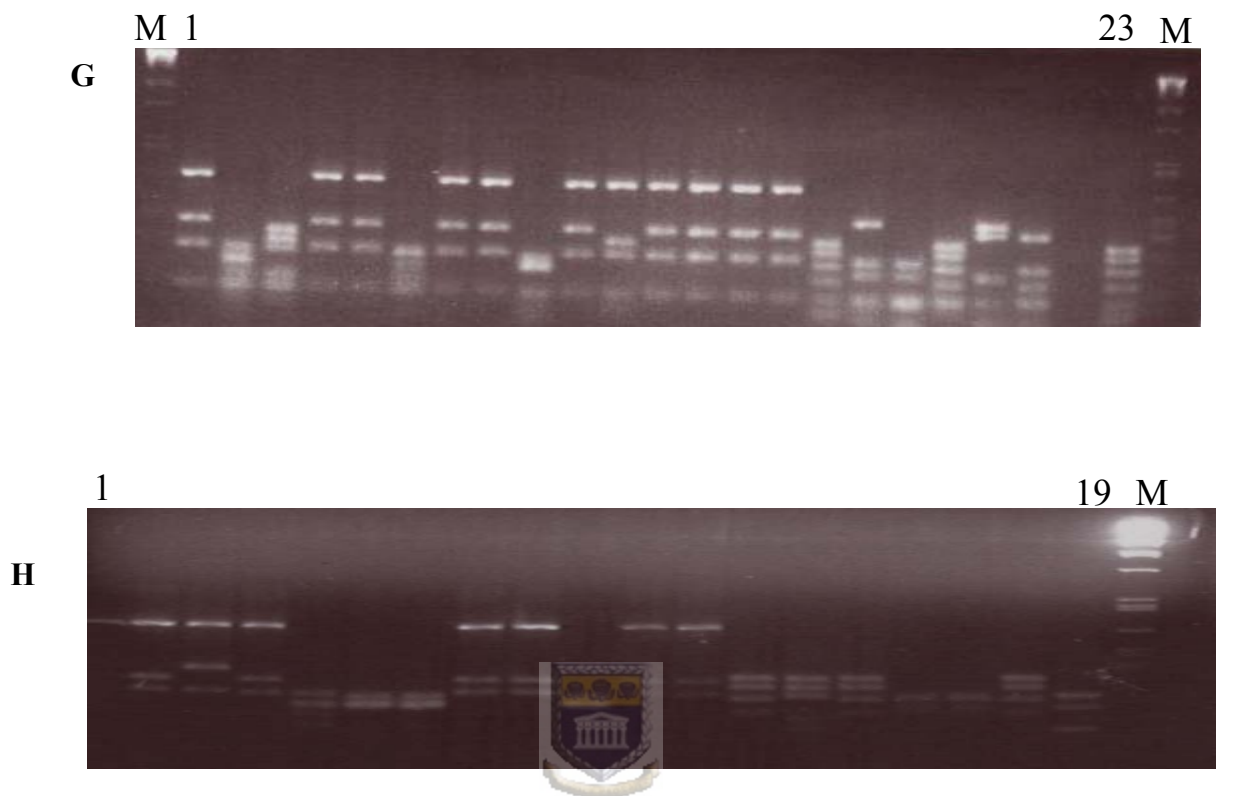


**Fig. 3.6:** Randomly selected clones digested with *Rsa*I and electrophoresed in 2.5 % agarose gels in  $0.5 \times$  TBE. Lane M: *Pst*I digested  $\lambda$  DNA.





**Fig. 3.7:** Randomly selected clones digested with *AluI* and electrophoresed in 2.5 % agarose gels in  $0.5 \times$  TBE. Lane M: *PstI* digested  $\lambda$  DNA.



**Fig. 3.8:** Randomly selected clones digested with *Hae*III and electrophoresed in 2.5 % agarose gels in  $0.5 \times$  TBE. Lane M: *Pst*I digested  $\lambda$  DNA.

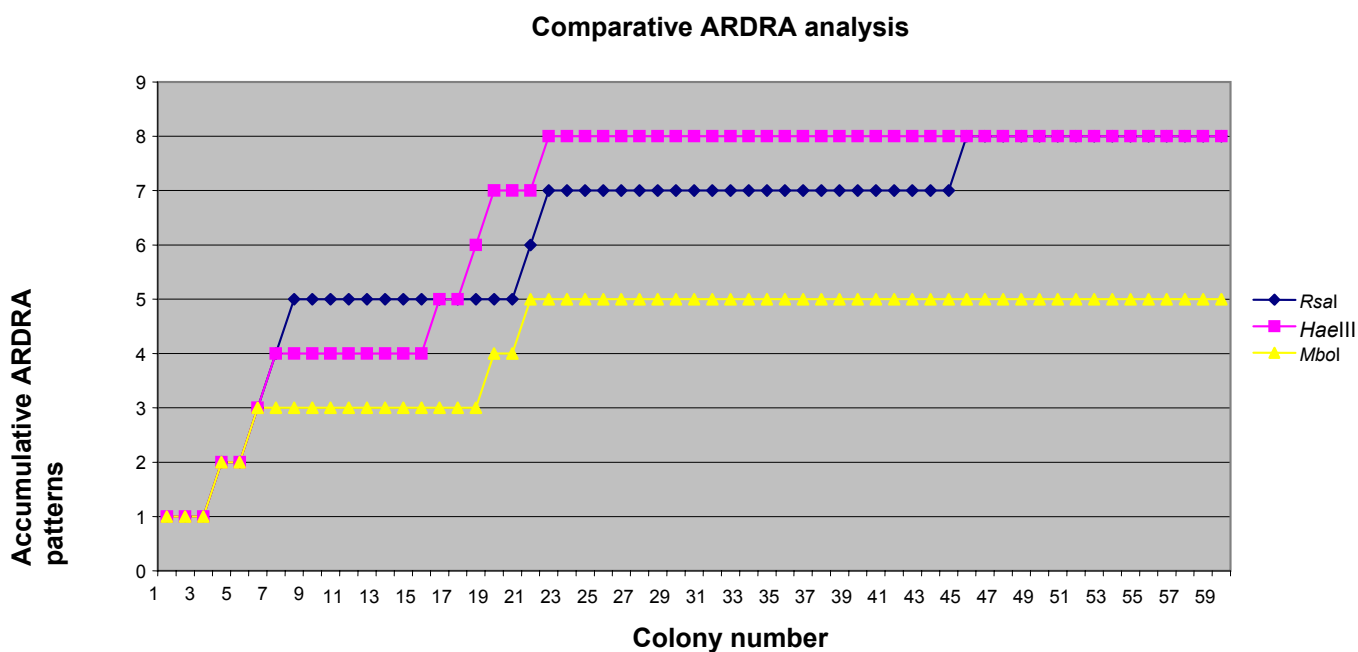


**Fig. 3.9:** Randomly selected clones digested with *Mbo*I and electrophoresed in 2.5 % agarose gels in  $0.5 \times$  TBE.

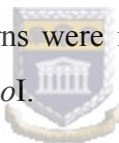
71 TC11 and 36 TC9 clones were screened by ARDRA, using *RsaI*, *HaeIII* and *MboI* endonuclease enzymes. Digestion with *MboI* was discontinued since it gave a single cleavage and thus provided little or no differentiation of the patterns (Fig. 3.9). Digestion with *RsaI* and *HaeIII* generated 8 unique patterns each, while digestion with *MboI* generated 5 patterns. 281 clones from New Zealand samples were screened using the same enzymes as described previously, except that *MboI* was substituted with *AluI*. *RsaI* generated 15 unique patterns; *HaeIII* generated 10 and *AluI* 9. Some of the patterns generated were similar for different enzymes, and the overall number of unique patterns from New Zealand samples was 15 with the combination of all three enzymes (Fig. 3.11).



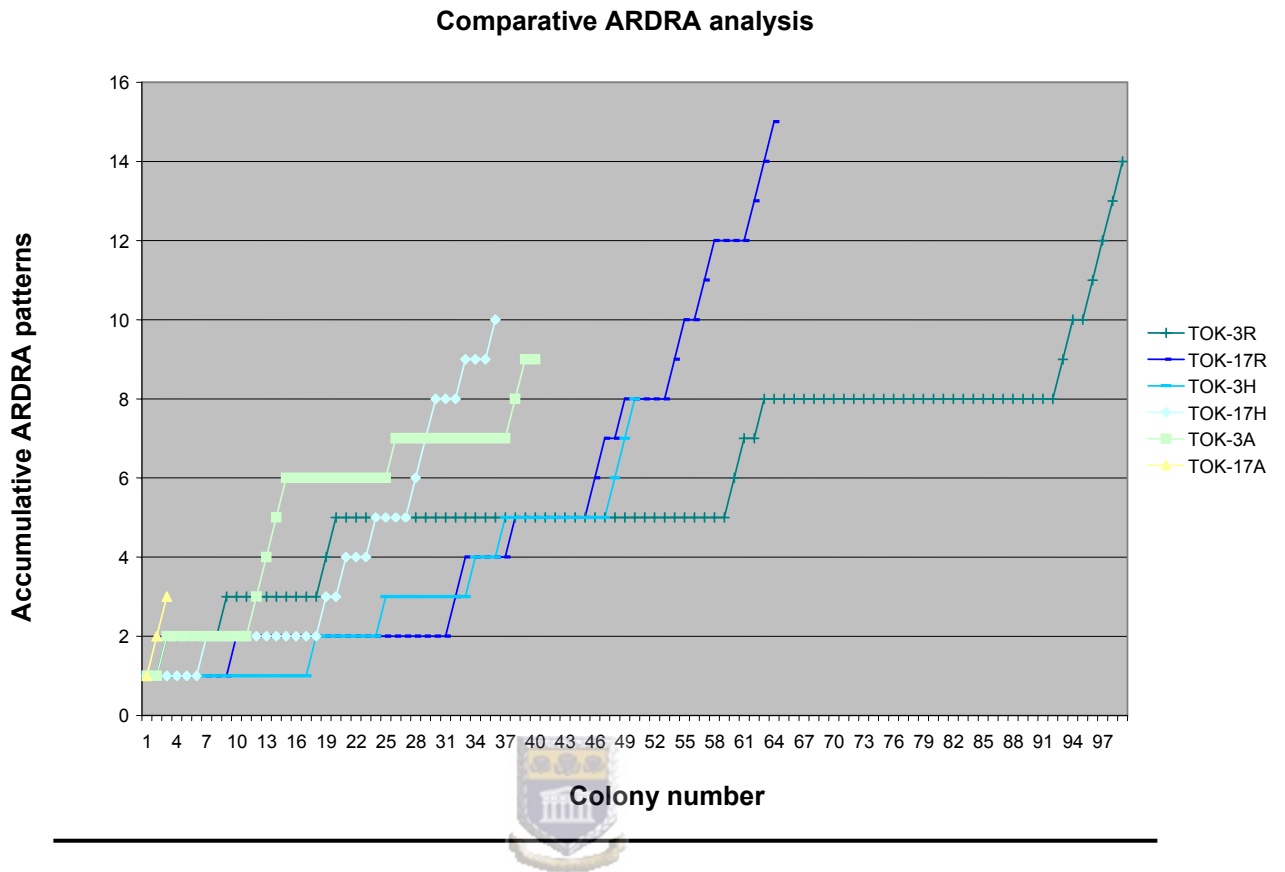
For analysis and monitoring of microbial diversity, unique groups identified by ARDRA were expressed in the form of a rarefaction plot (Fig. 3.10), and the different patterns were assembled into 8 groups. The screening of a large number of clones has several advantages, including better quantitation of major groups and a broader assessment of diversity. The rarefaction plot the number of unique patterns identified from the total number of screened clones. The point at which a plateau is reached is a simple and visual indication of the degree of genotypic diversity.



**Fig. 3.10:** Rarefaction plot of accumulative ARDRA patterns of Chinese clone library. A total of 8 unique patterns were identified after digestion with restriction endonuclease *Rsa*I, *Hae*III and *Mbo*I.



The screening of New Zealand clone library was completed before the collection curve could reach the plateau (Fig. 3.10). Comparative analysis of the Chinese and New Zealand libraries indicated that there was more archaeal diversity in the New Zealand library compared to the Chinese library. 9.3 % of Chinese library is made up of nanoarchaeum phylotypes, compared to 1.4 % from the New Zealand library. Figure 3.11 shows the collection curve of accumulative ARDRA patterns from the New Zealand library.

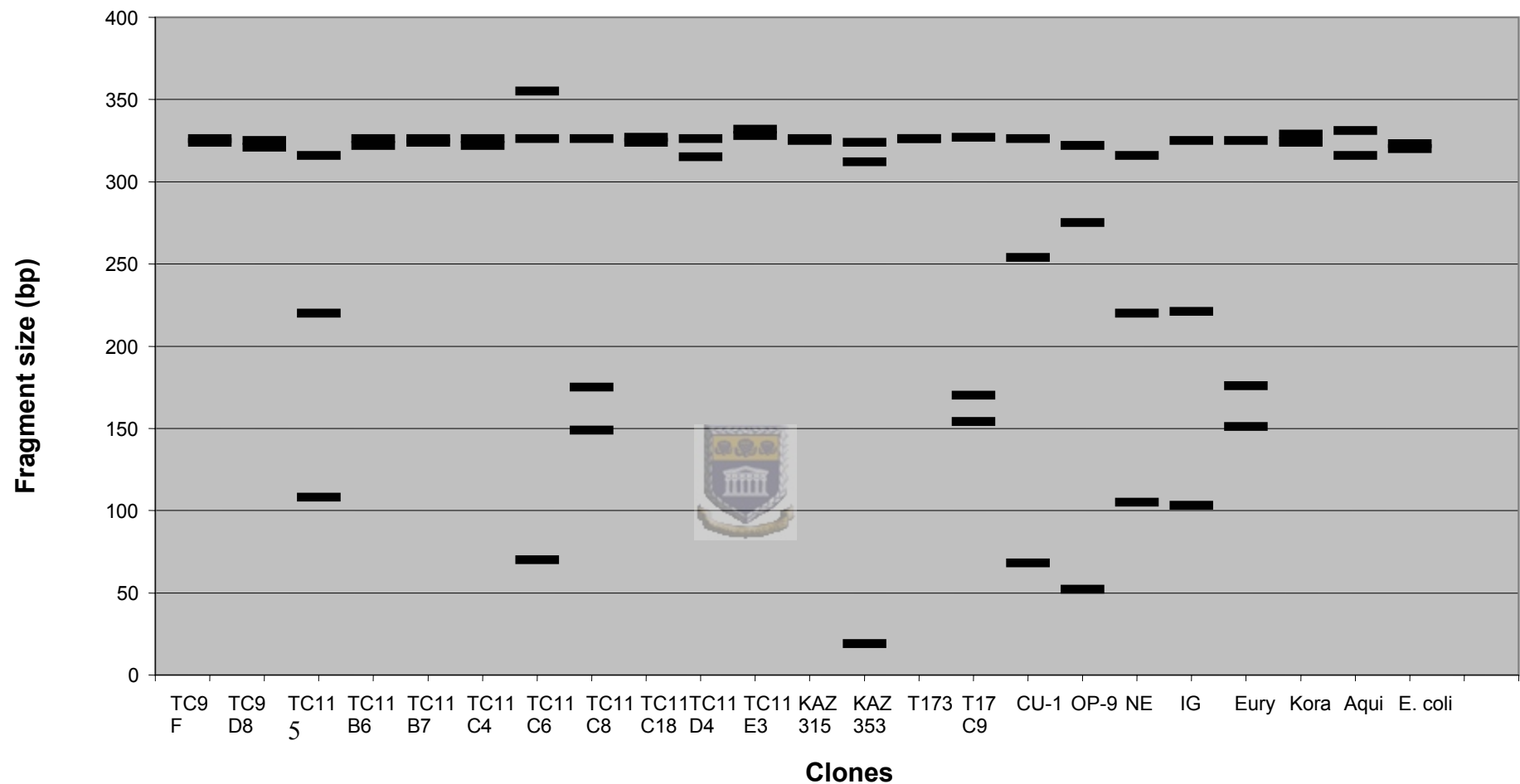


**Fig. 3.11:** Rarefaction plot of accumulative ARDRA patterns from the New Zealand clone library. Key: TOK-3 and TOK-17 are the sample names, R denotes *RsaI*, H, *HaeIII* and A, *AluI*.

Some caution is necessary in deducing phylogenetic diversity from ARDRA patterns. For example, occurrence of restriction sites in the M13 primer sequence would modify the restriction pattern of the 16S rDNA and consequently generate inaccurate estimate of diversity. Secondly, small fragments of about 100 bp or less limit the efficiency of ARDRA, small fragments being difficult to analyze manually, and can sometimes be invisible on agarose gels. To overcome these limitations, the digestion of nanoarchaeal sequences and representatives of Bacteria, Crenarchaeota,

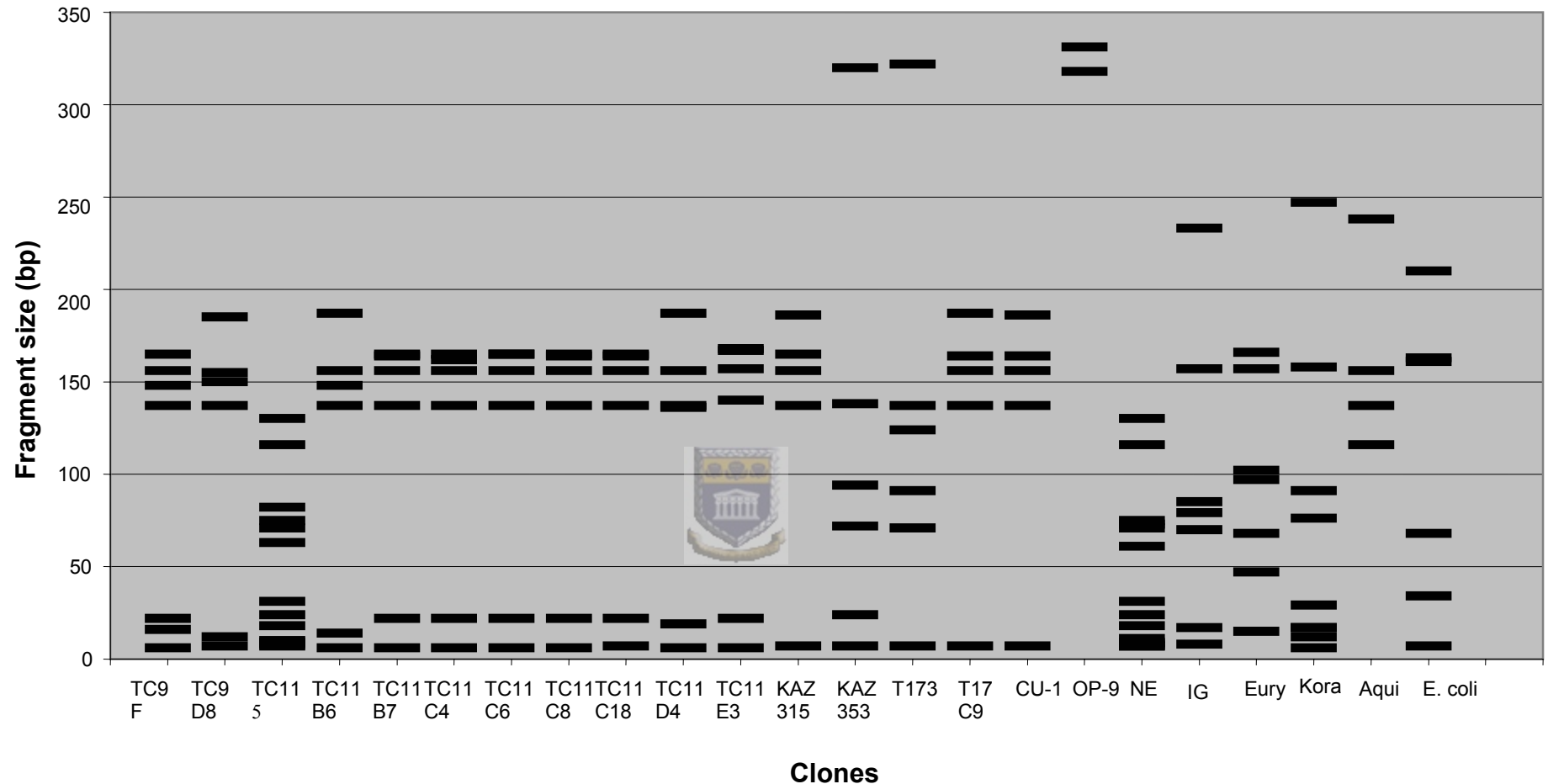
Euryarchaeota and Korarchaeota was modeled *in silico* with the same endonuclease enzymes as shown above using DNAClub software (Xiongfong Chen, [www.imtech.res.in/pub/nsa/dnaclub/dos/](http://www.imtech.res.in/pub/nsa/dnaclub/dos/)). The sequences were initially aligned with ClustalW (Thompson *et al.*, 1994), and 633 bp sequences were loaded on DNAClub for digestion. The generated fragment sizes were plotted on Excel to construct a series of restriction maps (Fig. 3.12-3.15).



ARDRA pattern generated using *RsaI*

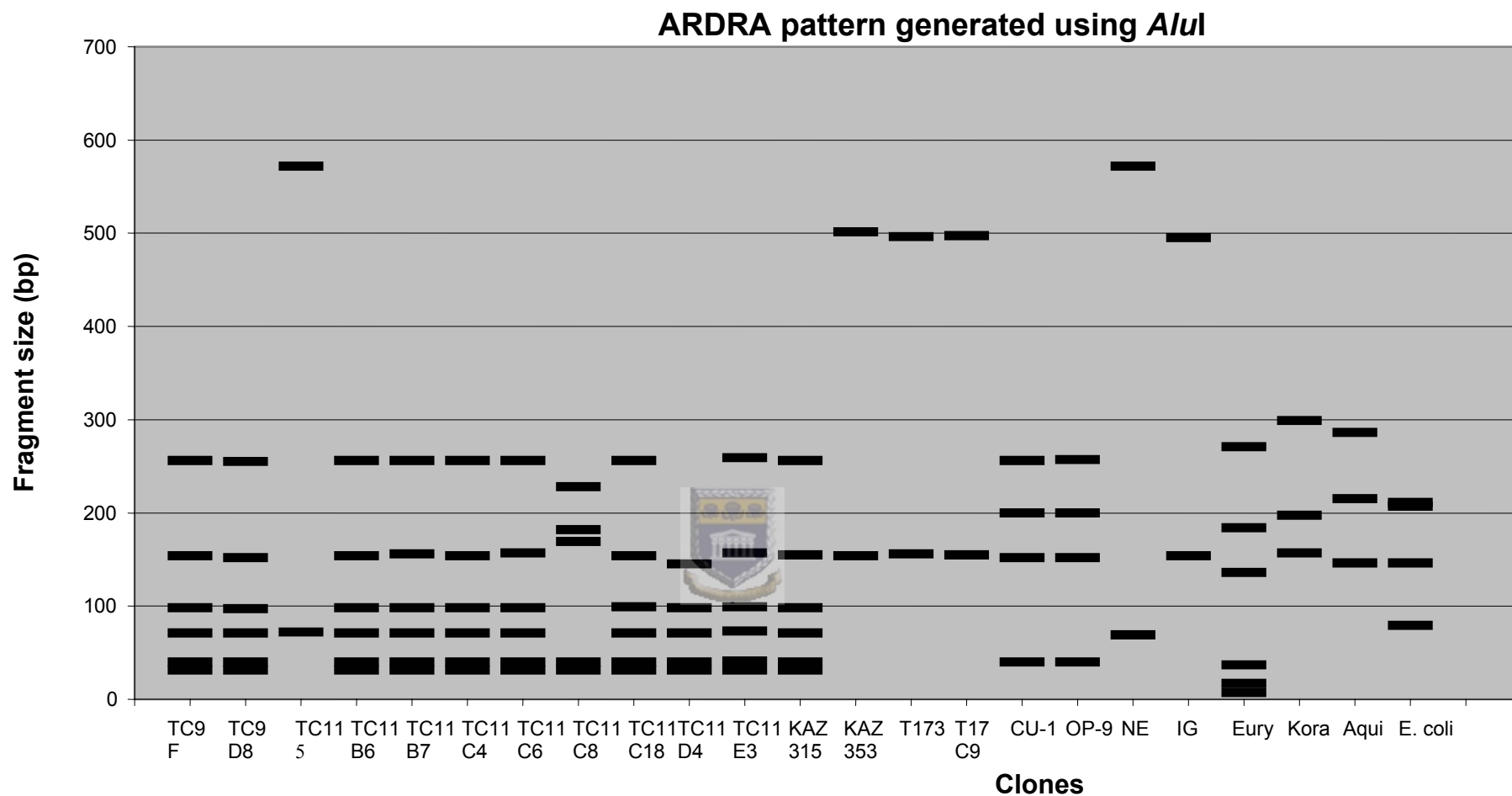
**Fig. 3.12:** Representation of ARDRA patterns generated from *in silico* digestion with *RsaI* using DNAClub. 633 bp fragments including, the priming sites of the “universal” primers A571F/U1204R, were digested. TC clones are Chinese nanoarchaeal phylotypes, KAZ and T clones are New Zealand nanoarchaeal phylotypes, CU-1 and OP-9 are uncultured published nanoarchaeal phylotypes. NE denotes *Nanoarchaeum equitans*, IG = *Ignicoccus* sp., Eury = Euryarchaeote, Kora = Korarchaeote and Aqui = *Aquifex* sp.

## ARDRA patterns generated using *HaeIII*

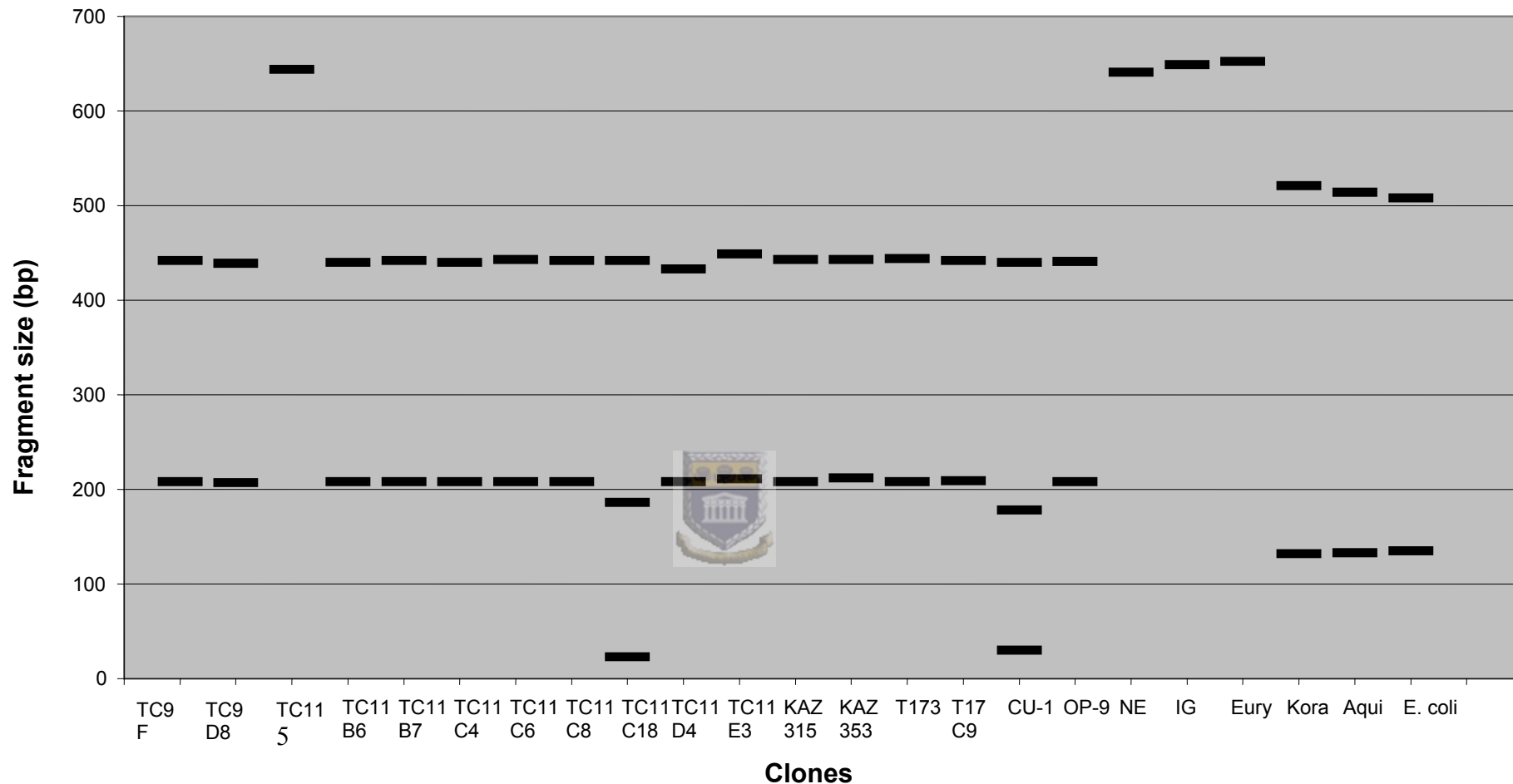


**Fig. 3.13:** Representation of ARDRA patterns generated from *in silico* digestion with *HaeIII* using DNAClub. 633 bp fragments including, the priming sites of the “universal” primers A571F/U1204R, were digested. TC clones are Chinese nanoarchaeal phylotypes, KAZ and T clones are New Zealand nanoarchaeal phylotypes, CU-1 and OP-9 are published nanoarchaeal phylotypes. NE denotes *Nanoarchaeum equitans*, IG = *Ignicoccus* sp., Eury = Euryarchaeote, Kora = Korarchaeote and Aqui = *Aquifex* sp.





**Fig. 3.14:** Representation of ARDRA patterns generated from *in silico* digestion with *AluI* using DNAClub. 633 bp fragments including, the priming sites of the “universal” primers A571F/U1204R, were digested. TC clones are Chinese nanoarchaeal phylotypes, KAZ and T clones are New Zealand nanoarchaeal phylotypes, CU-1 and OP-9 are published nanoarchaeal phylotypes. NE denotes *Nanoarchaeum equitans*, IG = *Ignicoccus* sp., Eury = Euryarchaeote, Kora = Korarchaeote and Aqui = *Aquifex* sp.

ARDRA pattern generated using *MboI*

**Fig. 3.15:** Representation of ARDRA patterns generated from *in silico* digestion with *MboI* using DNAClub. 633 bp fragments including the priming sites of the “universal” primers A571F/U1204R, were digested. TC-11 clones are Chinese nanoarchaeal phylotypes, KAZ and T clones are New Zealand nanoarchaeal phylotypes, CU-1 and OP-9 are published nanoarchaeal phylotypes. NE denotes *Nanoarchaeum equitans*, IG = *Ignicoccus* sp., Eury = Euryarchaeote, Kora = Korarchaeote and Aqoi = *Aquifex* sp.

Figures 3.12 to 3.15 are the representatives of *in silico* modeled ARDRA patterns using DNAClub software. When digested with *RsaI*, the Chinese nanoarchaeotes share similar patterns with one base pair difference at position 324,325, 326 and 327. TC11-5 also shares a similar pattern with *N. equitans* and differs by 3 base pairs. With *MboI*, all the nanoarchaeotes share similar patterns with 1-4 base pair differences. The exceptions are TC11-18 and CU-1 that share the same pattern and *N. equitans* and TC11-5, which are undigested. Digestion with *AluI* clusters the Chinese nanoarchaeotes with KAZ315 with 1-8 base pair differences. CU-1 and OP-9 share restriction sites at position 200, 240 and 496. *HaeIII* generated 1 to 14 fragmented patterns, with most of the patterns less than 100 bp long. The advantage is that these small fragments are observable *in silico* and can be easily analyzed compared to experimental data from agarose gels. The analysis of the novel nanoarchaeal phylotypes detected from this study reveals 3 distinct patterns by *MboI*, 5 with *AluI*, 6 with *RsaI* and 9 with *HaeIII*. The *in silico* modeled patterns correspond with the experimental generated pattern.

It should be noted that mutations, insertions, deletions and experimental errors may be introduced into sequences, particularly at the PCR stage, and thereby affect the restriction sites and consequently the patterns. The use of more than one enzyme is to reduce bias and ensure accuracy, and the ARDRA profile enhances comparative analysis of the patterns. All of the enzymes used were frequent cutters (four base pair cutters, *RsaI* (GT↓AC), *HaeIII* (GG↓CC) *AluI* (AG↓CT) and *MboI* (↓GATC↓). It is sensible to theoretically check the restriction patterns prior to conducting ARDRA. This assists in choosing the relevant enzymes and avoiding using enzymes that have restriction sites on

vector sequences. This does not only save time, but also minimizes the laboratory costs and the experimental bias. It is also helpful to choose enzymes that digest two or three times, thus generating three or four fragments which are easy to analyze. Enzymes generating fragments of less than 300 bp should be avoided, because they may not be visible on agarose gels. Using ARDRA for this study has helped in exploring diversity of nanoarchaeotes, and it showed the genotypic differences of the novel nanoarchaeotes detected from this study and their relatives, the published nanoarchaeotes. The unique representatives were sequenced and analyzed to construct the phylogenetic tree.

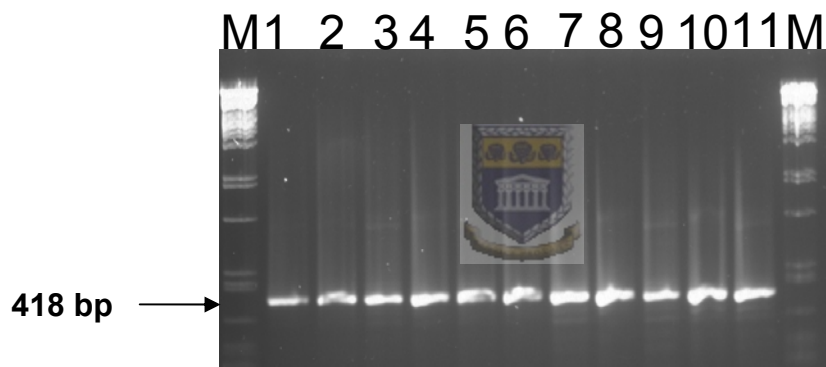
### 3.5 Amplification of nanoarchaeotes and phylogenetic analysis



Gene libraries derived from mixed amplification products contained defined fragments that were sequenced from M13 priming sites. The preliminary results of this study retrieved sequences from Chinese samples, which were submitted to the GenBank. These include nanoarchaeotes TC9F (AY572420), TC11-5 (AY571283), TC11-B6 (AY727890), TC11-B7 (AY727887), TC11-C4 (AY727886), TC11-C6 (AY727889), TC11-C8 (AY727888), and TC11-D4 (AY727891). In addition to nanoarchaeal sequences, one euryarchaeote TC9 (AY727895) and two crenarchaeote sequences (TC11-2 (AY727894); TC11-B2 (AY727893)) were also submitted. The nanoarchaeal sequences recovered after sequencing of 16S rDNA amplicons were aligned using ClustalW. Figure 3.17 shows the alignment of Chinese and New Zealand nanoarchaeal sequences with the published nanoarchaeotes from the database. The positions demarked in red indicate the

priming sites for “universal” archaeal primers A571F at position 1-16 and UA1204R at position 692-710 according to Bioedit alignment ruler.

The preliminary nanoarchaeal sequences obtained were used to design a nanoarchaeal primer, N989R. Position 421-439 was identified as a conserved region for nanoarchaeotes and was used to design primer N989R, which could be coupled with the universal archaeal primer A571F. The PCR amplification using this primer set generated 418 bp nanoarchaeal sequences (Fig. 3.16).



**Fig. 3.16:** Amplicons of nanoarchaeal rDNA with nanoarchaeal primers A571F and N989R on 1 % agarose gel electrophoresis. The primers generated 418 bp amplicons from the constructed mini libraries. Lane M: *Pst*I digested  $\lambda$  DNA, Lane 1-11: TC9-D8, TC11-B6, TC11-B7, TC11-C4, TC11-C6, TC11-C8, TC11-D4, KAZ315, KAZ353, T173 and T17C9, respectively.

The N989R primer coupled with A571F amplified nanoarchaeal 16S rDNA sequences from both New Zealand and Chinese samples. However, according to the sequence alignment shown in Figure 3.17, primer N989R could anneal to non-nanoarchaeal

sequences. It is therefore no longer specific for nanoarchaeotes. For future studies, the detection of more novel nanoarchaeal phylotypes may be facilitated by the design of new universal nanoarchaeal primers or reassessment of the currently used nanoarchaeal primers. Such primers must be based on all nanoarchaeal sequences available in the database. This needs careful selection since nanoarchaeote 16S genes share most of their conserved regions with other archaeal domains, thus minimizing the possible sequences that could be used as signature priming sites for nanoarchaeotes. The ideal nanoarchaeal-specific primer pair should amplify nanoarchaeal sequences only and exclude all other archaeal sequences. Therefore, every amplicon should be derived from a nanoarchaeote template, accelerating the screening process and minimizing experimental costs.



```
          10          20          30          40          50          60
TC9F AY572 TAAGGGGCCCC TAGCC GGACCGAT GTGGCCCC --GTG AAATCCACGGGCTCA --
9D8        TAAAGGGGCCG TAGCC --GGACCGAT GTGGTCCCC --GTG AAATCCACGGGCATA --
TC11-5 AY5 TCAACCCGTGGGCGTGCAGGGGCTACT ATCGGTCTA --GTG AAATCCCTCGGGCTCA --
TC11B6 AY7 TAAAGGGGCCG TAGCC GGACCGAT GTGGCTCCC --GTG AAATCCACGGGCATA --
TC11-B7 AY TAAAGGGGCCG TAGCC --GGACCGAT GTGGCCCC --GTG AAATCCACGGGCTCA --
TC11-C4 AY TAAAGGGGCCG TAGCC GGACCGAT GTGGCTCCC --GTG AAATCCACGGGCATA --
TC11-C6 AY TAAAGGGGCCG TAGCC GGACCGAT GTGGCTCCC --GTG AAATCCACGGGCATA --
TC11-C8 AY TAAAGGGGCCG TAGCC GGACCGAT GTGGCTCCC --GTG AAATCCACGGGCATA --
TC11-C18   TAAAGGGGCCG TAGCC GACTCGATCGTGGCCCC --GTG AAATCCACGGGCTCA --
TC11-D4 AY TAAAGGGGCCG TAGCC GGACCGAT GTGGCTCCC --GTG AAATCCACGGGCATA --
TC11-E3    TAAAGGGGCCG TAGCC GGACCGAT GTGGCCCC --GTG AAATCCACGGGCTCA --
KAZ315     TAAAGGGGCCG TAGCC TGGACCGAT GTGGCTCCT --GTG AAATCCACGGGCTCA --
KAZ353     TAAAGGGGCCG TAGCC GTACCCGT GTGACCCCT --GTG CAAATCCACGGGCTCA --
T173       TAAAGGGGCCG TAGCC TGGACCGAT GTGACCCC --GTG AAATCCACGGGCATA --
T17C9      TAAAGGGGCCG TAGCC --GGACTGGT GTGACCCCAT --GTG AAATCCACGGGCTCAC --
Nanoarchae TAAAGGGGCCG TAGCC --GGCTAGT GTGACCCCT --GTG AAATCCACGGTAATAAT
Nanoarchae TAAAGGGGTCG TAGCC --GGCTGGT ATGACCCCT --GTG AAATCCGTGGTAATAGT
N. equitan TAAAGGGGCCG TAGCC --GGCCGTG GTGGCTCCG --GTG AAATCCCTCGGGCTCA --
Ignicoccus TAAAGCGCCCG TAGCC --GGCTGGT --GGCCCCCT --CCT --AAAGCCCCGGGCTCA --
Archaeoglo TAAAGCGTCCG TAGCC --GGCTGGT AAGTCCCC --GGG AAATCTGGCGGCTTA --
KORARCHAEO TAAAGCGTCCG TAGCC --GGCCGTG AAGTCTGCA --GTT AAATCCGCTG --TAG --
Aquifex py TAAAGCGTCCG CAGCC --GGTCGGGT AAGCGGGAT --GTC AAAGCCCCAGGCTCA --
E. coli     TAAAGCGCACG CAGCC --GGTTTGT --AAGTCAGAT --GTG AAATCCCCGGGCTCA --

          70          80          90          100         110         120
TC9F AY572 ACCCGTGGGCGCGCAGGGGCTACTATCGGTCTAGGGGACGGGAGAAGCTGGGG --GTATTC
9D8        ACCCGTGGGCGCGCGGGGGCTACTATCGGTCTTGGGGACAGGAGAAGCT --AGGAGTATTC
TC11-5 AY5 ACCCGAGGGCGCGCCGGAGCTACTACC GGCTAGGGACGGGAGGGGCC --GACCGTACTC
TC11B6 AY7 ACCCGTGGGCGCGCGGGGGCTACTATCGGTCTTGGGGACAGGAGAAGCT --AGGAGTATTC
TC11-B7 AY ACCCGTGGGCGTGCAGGGGCTACTATCGGTCTAGGGGACGGGAGAAGCTGGGG --GTATTC
TC11-C4 AY ACCCGTGGGCGCGCGGGGGCTACTATCGGTCTTGGGGACAGGAGAAGCT --AGGAGTATTC
TC11-C6 AY ACCCGTGGGCGCGCGGGGGCTACTATCGGTCTTGGGGACAGGAGAAGCT --AGGAGTATTC
TC11-C8 AY ACCCGTGGGCGCGCGGGGGCTACTATCGGTCTTGGGGACAGGAGAAGTT --AGGAGTATTC
TC11-C18   ACCCGTGGGCGTGCAGGGGCTACTATCGGTCTAGGGGACGGGAGAAGCTGGGG --GTATTC
TC11-D4 AY ACCCGTGGGCGCGCGGGGGCTACTATCGGTCTTGGGGACAGGAGAAGCT --AGGAGTATTC
TC11-E3    ACCCGTGGGCGTGCAGGGGCTACTATCGGTCTAGGGGACGGGAGAAGCTTGGGGG --TATCC
KAZ315     ACCCGTGGGCGTGCAGGGGCTACTATCGGTCTAGGGGACGGGAGAAGCT --GGGGTATTC
KAZ353     ACCCGTGGGCGCGCGGGGGTTACTACC GGCTTGGGGACGGGAGAGGCC --TGGGGTATTC
T173       ACCCGTGGGCGTGCAGGGGTTACTACC GGCTAGGGGACGGGGGAGGCC --GGGGTATTC
T17C9      ACCCGTGGGCGTGCAGGGGTTACTACCAGTCTAGGGGACGGGAAGGAGCC --GAGGGTATTC
Nanoarchae TCCCGTGGGCTTGCAGGGGTTACTACTAGCCTAGGGGACGGGAAGAGCC --GAGGGTATTC
Nanoarchae ACCCACGGGCGCGCAAGGGTTACTACCAGCCTAGGGGATAGGGGAAGCC --GAGGGTATTC
N. equitan ACCCGAGGGCGCGCCGGAGCTACTACC GGCTAGGGACGGGAGGGGCC --GACCGTACTC
Ignicoccus ACCCGGGGACTGG --AGGGGTAGCGCCAGGCTAGGGGGCGGGAGAGGCC --GAGGGTACTC
Archaeoglo ACCGTCAGAAATCCGGGGGATACTGCCAGCCTAGGGACGGGAGAGGCC --GGGGGAATTC
KORARCHAEO ACCAGGC GGTTGCTGCAAA TACTGCAAGGCTAGGGAGCGGGAGGAGCC --GGGGGTATTC
Aquifex py ACCGTGGGAATGGCATCCC GAAGTCCCGACTTGAGG --CACCGGGGCGAGCGGAATTC
E. coli     ACCGTGGGAATGCATCTGATACTGGCAAGCTTGAGT --CTCGTAGAGGGGGGTAGAATTC

          130         140         150         160         170         180
TC9F AY572 CGGGGGGAGCGGTGAAATGCTGTAATCCCCGGAGGACCACCCGAGGCG --AAGGCGCC --CA
9D8        CGGGGGGAGCGGTGAAATGCTGTAATCCCCGGAGGACTACCCGCGGCG --AAAGCGCC --TA
TC11-5 AY5 CCGGGGGAGCGGTGAAATGCTGTAATCCCCGGGAGGACGACCCGTGGCG --AAAGCGGT --CG
TC11B6 AY7 CGGGGGGAGCGGTGAAATGCTGTAATCCCCGGAGGACTACCCGCGGCG --AAAGCGCC --TA
TC11-B7 AY CGGGGGGAGCGGTGAAATGCTGTAATCCCCGAAGGACCAACCCGAGGCG --AAGGCGCC --CA
TC11-C4 AY CGGGGGGAGCGGTGAAATGCTGTAATCCCCGGAGGACTACCCGCGGCG --AAAGCGCC --CA
TC11-C6 AY CGGGGGGAGCGGTGAAATGCTGTAATCCCCGGAGGACTACCCGCGGCG --AAAGCGCC --TA
TC11-C8 AY CGGGGGGAGCGGTGAAATGCTGTAATCCCCGGAGGACTACCCGCGGCG --AAAGCGCC --TA
```





```

Ignicoccus      CGGGTAGTCCCGGCTGTAAACGATGC-GGGCTAGG-TGTTGGGCGGG-CTTCGA--GCCC
Archaeoglo     CGGGTAGTCCCTGGCTGTAAACGATGC-GGGCTAGG-TGTCGGGCCAG-CTACGA--GCTG
KORARCHAEO     CGGGTAGTCCAGCCGTAAACAATGC-CGGCTAGG-TGCTGGCCAAGGTTTCGG--CCTT
Aquifex py     CGGGTAGTCCCGGCCGTAAACCATGG-GCGCTAGG-----GCTTGTCCCTTTGG-GGCAG
E.coli         CTGGTAGTCCACGCCGTAAACGATGTCGACTTGG-----GGTTGTCCCTTGA-GGCGT
    
```

```

          310          320          330          340          350          360
TC9F AY572    TTTTCG-GTGG-GTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
9D8           TTTTCG-GTGG-AGTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGC
TC11-5 AY5    GCCCA-GTGG-GTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
TC11B6 AY7    TTTTCG-GTGG-GTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
TC11-B7 AY    TTTTCG-GTGG-GTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
TC11-C4 AY    TTTTCG-GTGG-GTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
TC11-C6 AY    TTTTCG-GTGG-GTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
TC11-C8 AY    TTTTCG-GTGG-GTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
TC11-C18     TTTTCG-GTGG-AGTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
TC11-D4 AY    TTTTCG-GTGG-GTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
TC11-E3      TTTTCG-GTGG-AGTAGGG-AAGCCGTTAAGCCC-TCCGCCTGGGGAGTACGGCCGCAAGG
KAZ315      TTTTCG-GTGG-AGTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
KAZ353      TCTCG-GTGG-GACAGGGTAAAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
T173       TCTCG-GTGG-GACAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
T17C9      TCTCG-GTGG-GACAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
Nanoarchae   ACTCA-GTGG-GTAGGG-AAGCCGTTAAGCCC-ACCGCCTGGGGAGTACGGCCGCAAGG
Nanoarchae   ACCCG-GTGG-GAAAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
N. equitan   N. equitan  GCCCA-GTGG-GTAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
Ignicoccus   GCCCAAGTGC-CGCAGGG-AAGCCGTTAAGCCC-GCCGCCTNNGGGAGTACGGCCGCAAGG
Archaeoglo   GCCCG-GTGC-CGGAGGG-AAGCCGTTAAGCCC-GCCGCCTGGGGAGTACGGCCGCAAGG
KORARCHAEO   GGCCA-GTGT-CGAAGCG-AAGGCCTTAAATCCG-GCCGCCTGAGGAGTACGGCCGCAAGG
Aquifex py   G-----CTC--GCAGCT-AACGCGTTAAGCGC-CCCGCCTGGGGAGTACGGCCGCAAGC
E.coli       GGCT--TCC--GGAGCT-AACGCGTTAAGTCG-ACCGCCTGGGGAGTACGGCCGCAAGG
    
```

```

          370          380          390          400          410          420
TC9F AY572    CTGAAACTTAAAGGAAATAGGC GGGGGGGCAT-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
9D8           CTGAAACTTAAAGGAAATAGGC GGGGGGGG-CAT-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
TC11-5 AY5    CTGAAACTTAAAGGAAATAGGC GGGGGGGAGCAC-ACAAGAGGTGGG-GTGC GCGGTTTAAAT
TC11B6 AY7    CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAT-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
TC11-B7 AY    CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAT-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
TC11-C4 AY    CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAT-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
TC11-C6 AY    CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAT-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
TC11-C8 AY    CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAT-GCAAGGGGTGGG-GTGTGCGGTTTAAATG
TC11-C18     CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAT-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
TC11-D4 AY    CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAT-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
TC11-E3      CTAAACTTAAATGAATATGCGGGGGGGGCAT-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
KAZ315      CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAT-ACAAGGGGTGGG-GTGC GCGGTTTAAATG
KAZ353      CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAC-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
T173       CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAC-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
T17C9      CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAC-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
Nanoarchae   CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAC-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
Nanoarchae   CTGAAACTTAAAGGAAATAGGC GGGGGGGGCAC-ACAAGGGGTGGG-GTGTGCGGTTTAAATG
N. equitan   N. equitan  CTGAAACTTAAAGGAAATAGGC GGGGGGGAGCAC-ACAAGAGGTGGG-GTGC GCGGTTTAAAT
Ignicoccus   CTGAAACTTAAAGGAAATAGGC GGGGGGGAGCACCAAGGGGTGGA-GCCTGCGGTTTAAAT
Archaeoglo   CTGAAACTTAAAGGAAATAGGC GGGGGGGAGCACCAACGGGTGGA-GCCTGCGGTTTAAAT
KORARCHAEO   CTGAAACTTAAAGGAAATAGGC GGGGGGGAGCACCAAGGGGTGAATGCCTGCGGTTTAAAT
Aquifex py   CTGAAACTCAAAGGAAATAGGC GGGGGGGCCGC-ACAACCGGTGGA-GCCTCTGGTTCAAT
E.coli       TTAAGACTCAAATGAATGACGGGGGCCGC-ACAAGCGGTGGA-GCATGTGGTTTAAAT
    
```

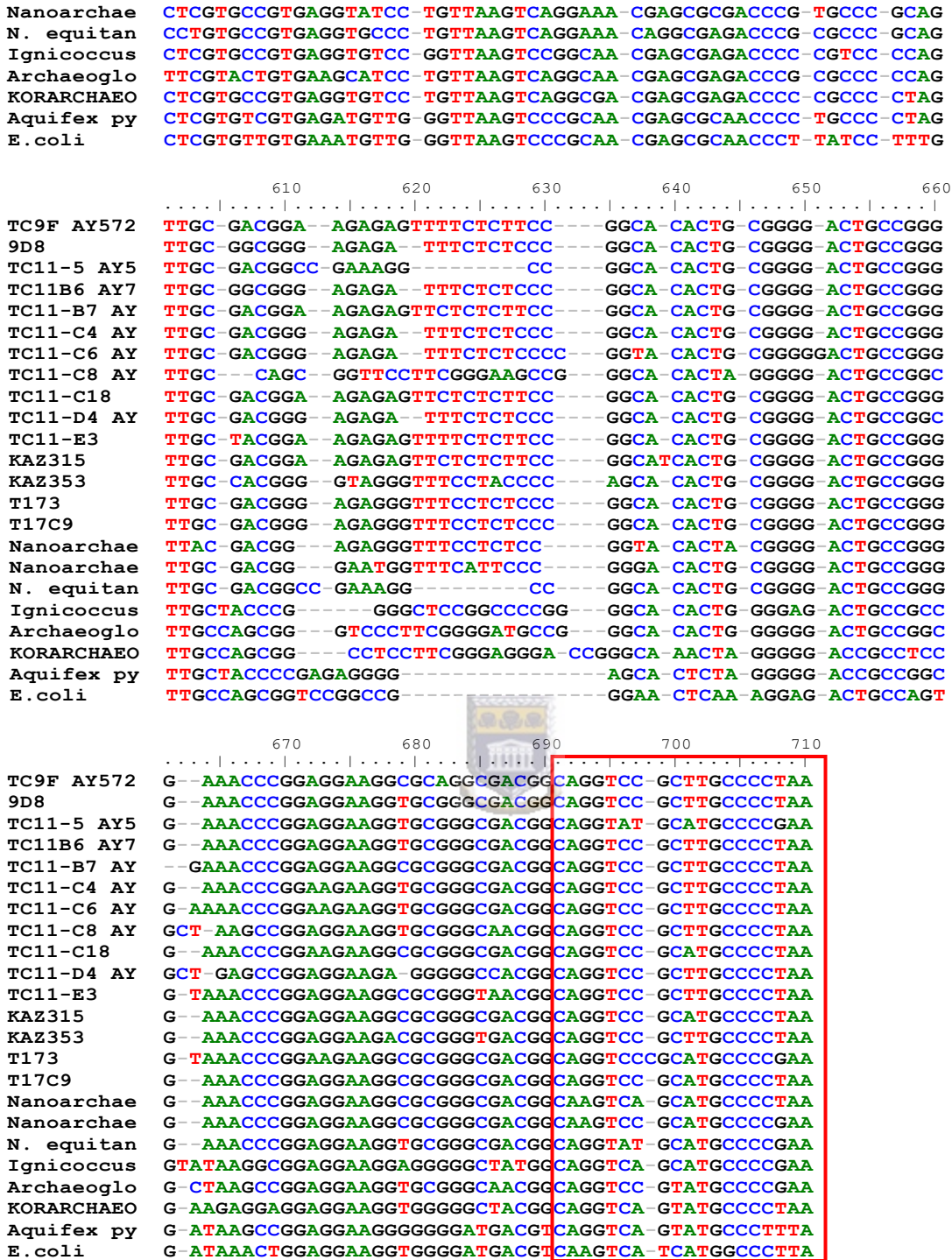
```

          430          440          450          460          470          480
TC9F AY572    GGAACTGACACCCGAAACCTTACC GGGGGG-CGACAGCA--GGATGAAGGTCA-GGC-----
9D8           GGAACTGACACCCGAAACCTTACC GGGGGG-CGACAGCA--GGATGAAGGTCA-GGC-----
    
```

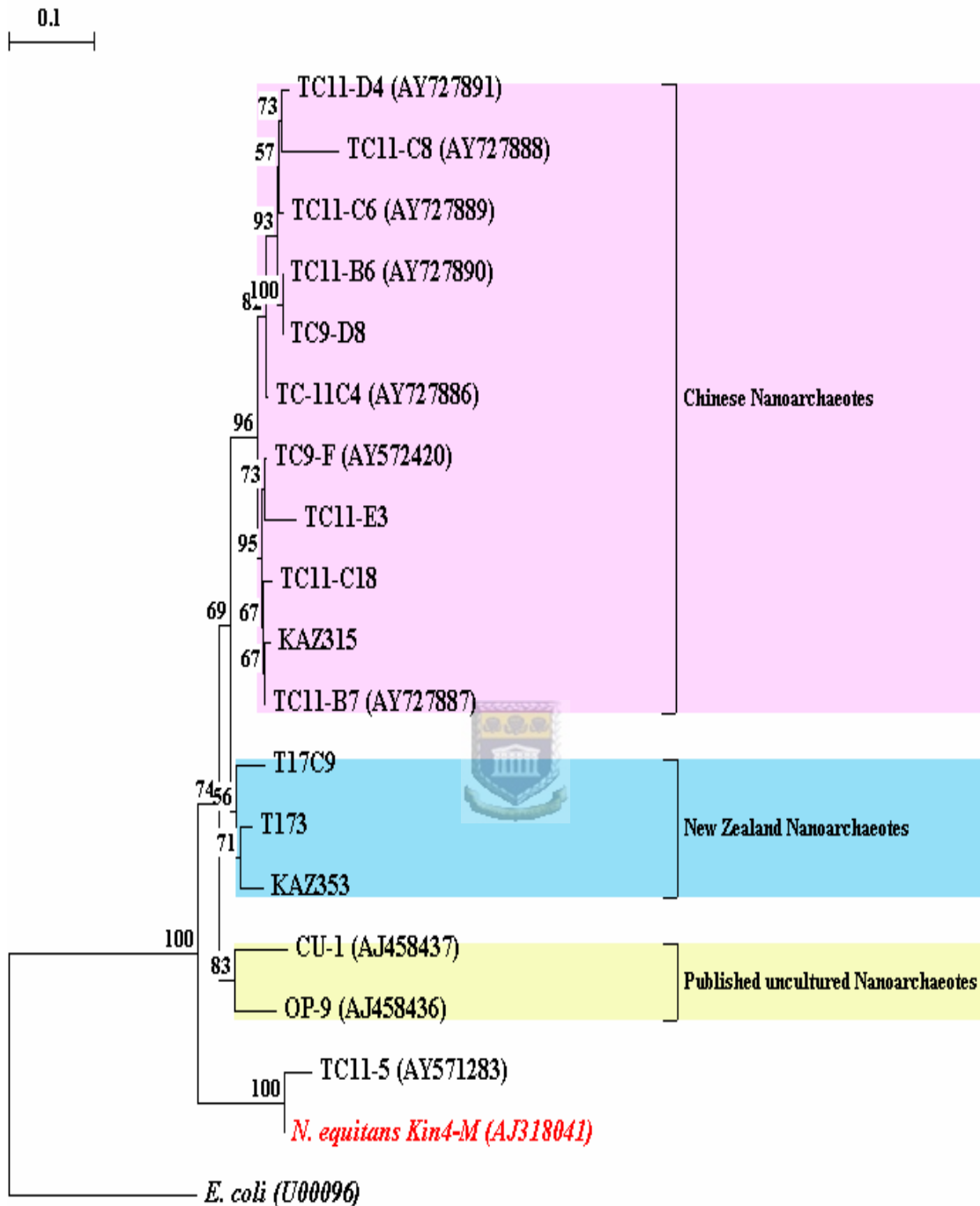
TC11-5 AY5	GGATTCGACGCCGGAACTCACC	CGGGGCGTGCAGCA	CAATGATGGTC	GGCC
TC11B6 AY7	GGAACTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGAAGGTCA	GGC
TC11-B7 AY	GGAACTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGAAGGTCA	GGC
TC11-C4 AY	GGAACTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGAAGGTCA	GGC
TC11-C6 AY	GGAACTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGAAGGTCA	GGC
TC11-C8 AY	GGAACTGACACCCGAAACCTTACC	CGGGGCGACAGCG	GGATGAAGGTCA	GGC
TC11-C18	GGAACTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGAAGGTCA	GGC
TC11-D4 AY	GGAACTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGAAGGTCA	GGC
TC11-E3	GGAACTGACTCCGGAAACCTTACC	CGGGGCGACTGCTG	GATGAAGGTCA	TGC
KAZ315	GGAACTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGAATGTCA	TGC
KAZ353	GGACCTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGATGGTCA	GGC
T173	GGAACTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGATGGTCA	GGC
T17C9	GGACCTGACACCCGAAACCTTACC	CGGGGCGACAGCA	GGATGATGGTCA	GGC
Nanoarchae	GGACCTAACACCCGAAACCTTACC	CGGGGAGCGACAGCA	GGATGATGGTCA	GGC
Nanoarchae	GGACCTGACACCCGAAATCTTACC	CGGGGCGACAGCA	GGATGATGGTCA	GGC
N. equitan	GGATTGACACCCGAAACCTTACC	CGGGGCGTGCAGCA	CAATGATGGTC	GGCC
Ignicoccus	GGAGTCAACGCCGGAAACCTTACC	CGGGGCGACAGCA	GGATGAAGGTCA	GGC
Archaeoglo	GGATTCAACGCCGGAAAGCTTACC	CGGGGAGACAGCT	CGATGAAGGTCA	GGC
KORARCHAEO	GGACTCAACGCCGGAAATCTTACC	CGGGGCGACAGCA	GGATGAAGGTCA	GGC
Aquifex py	GGATTGTAACCGAAGAACCTTACC	CGGGGCGTTGACATGCC	GGGGAGACTCCGAAAGC	
E.coli	CGATGCAACCGAAGAACCTTACC	TGGTCTTGACATCCA	CGGAAGTTTTCA	GAGA

	490	500	510	520	530	540
TC9F AY572	TGAA	GAC	CTTACCTGACGCGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
9D8	TGAA	GAC	CTTACCTGACACGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
TC11-5 AY5	TGAA	GGG	CTTACCGGAGGCGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
TC11B6 AY7	TGAA	GAC	CTTACCTGACACGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
TC11-B7 AY	TGAA	GAC	CTTACCTGACGCGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
TC11-C4 AY	TGAA	GAC	CTTACCTGACACGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
TC11-C6 AY	TGAA	GAC	CTTACCTGACACGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
TC11-C8 AY	TGAA	GAC	TCTACCAGACAGCT	GAGAGGTGGTGCATGGCCGCCGTCAG		
TC11-C18	TGAA	GAC	CTTACCTGACACGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
TC11-D4 AY	TGAA	GAC	CTTACCTGACACGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
TC11-E3	TAAAT	GAC	CTTACCTGACACGCT	GAGAGGAGGTGCATGGCCGCCATCAG		
KAZ315	TGAA	GAC	CTTACCTGACGCGCT	GAGAGGAGGTGCGTGGCCGCCGTCAG		
KAZ353	TGAA	GAC	CTTACCGGACGCGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
T173	TGAA	GAC	CTTACCGGACGCGCT	GAGAGGAGGTGCATGGTCGCCGTCAG		
T17C9	TGAA	GAC	CTTACCGGACGCGCT	GAGAGGAGGTACATGGCCGCCGTCAG		
Nanoarchae	TGAA	GAC	CTTACCGGACAGCT	GAGAGGAGGTGCATGGCTGCCGTCAG		
Nanoarchae	TAAA	GAC	CTTACCGGACAGCT	GAGAGGAGGTGCATGGCTGCCGTCAG		
N. equitan	TGAA	GGG	CTTACCGGAGGCGCT	GAGAGGAGGTGCATGGCCGCCGTCAG		
Ignicoccus	TGAA	GAC	CTTACCTGACGCGCT	GAGAGGAGGTGCATGGCCGTCGCCAG		
Archaeoglo	TGAA	GAC	CTTACCAGAGTAGCT	GAGAGGTGGTGCATGGCCGCCGTCAG		
KORARCHAEO	TGAA	GAC	CTTACCAGACGCGCT	GAGGGGTGGTGCATGGCCGTCGCCAG		
Aquifex py	GGAGT	TGTGGAAGTCTCTGAC	CC	CCCCGGCACAGGTGGTGCATGGCCGTCGTCAG		
E.coli	TGAGAATGTG	CTTTCGGGAAACCGT		GAGACAGGTGCTGCATGGCTGTCGTCAG		

	550	560	570	580	590	600
TC9F AY572	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	GCAG	
9D8	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	GCCG	
TC11-5 AY5	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CAGGCGGACCCG	CGCCC	GCAG	
TC11B6 AY7	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	GCCG	
TC11-B7 AY	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	GCAG	
TC11-C4 AY	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	GCCG	
TC11-C6 AY	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	GCCG	
TC11-C8 AY	CTCGTACTGTGAAGCATCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	CTAG	
TC11-C18	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	GCAG	
TC11-D4 AY	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	GCCG	
TC11-E3	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCATGAAA	CGAGCGTACCCT	CGCCCTGCTG		
KAZ315	CTCGTGCCGTGAGGTGTCC	TGTTAAGTCAGGAAA	CGAGCGGACCCG	CGCCC	GCAG	



**Fig. 3.17:** Alignment of 633 bp nanoarchaeal sequences using ClustalW. The red demarcated positions indicate the priming regions. Position 1-16, 421-4439, and 692-710 are for A571F, N989R and U1204R, respectively.



**Fig. 3.18:** Neighbor-joining phylogenetic tree based on 100 bootstrap replicates constructed using Jukes & Cantor distances. The tree demonstrates phylogenetic relationships among the Chinese, New Zealand and the published nanoarchaeal phylotypes as inferred from the 633 bp rDNA data set.




According to the phylogenetic tree based on Jukes and Cantor (1969) distance estimations (Fig. 3.18), all novel nanoarchaeal phylotypes detected in this study form a paraphyletic clade with the published nanoarchaeotes, including *N. equitans* with a 100 % bootstrap support. The Chinese nanoarchaeotes are related to each other by 57 to 100 % bootstrap support, and New Zealand nanoarchaeotes supported by 56 to 74 % bootstrap values. The New Zealand nanoarchaeotes are related to the Chinese by 67 to 74 % bootstrap support. TC11-5 shares a monophyletic clade with *N. equitans*, with 100 % bootstrap support. This is a good indication that culturable nanoarchaeotes can be grown from Chinese samples for further analysis. When compared to other phyla (Fig. 3.19), the relationship of New Zealand nanoarchaeotes with Chinese nanoarchaeotes is supported by a weak bootstrap support of 38 % and 48 % to uncultured published nanoarchaeotes CU-1 and OP-9. This weak bootstrap support strongly confirms the difference of these three minor clades within the nanoarchaeotes. Each of the sequences contain particular small fragment sequences or individual bases that are variable, thus distinguishing them from their nanoarchaeal relatives. It is concluded that the novel sequences represent distinct nanoarchaeal phylotypes. It is also observed from Figure 3.18 & 3.19 that nanoarchaeotes cluster according to their geographical origin. Since this study has analysed only a limited range of samples from the world-wide range of hydrothermal biotopes, it is suggested that the true diversity of nanoarchaeotes is probably much wider. This would be supported by the very different environmental conditions that exist in different geothermal areas, which would be expected to greatly influence adaptation and evolutionary divergence.

The data generated from this study contribute to the increasing database of Nanoarchaeota as novel phylotypes are continuously discovered. If the novel nanoarchaeotes are symbiotic, the potential hosts may be some of the commonly observed phylotypes in the microbial consortium of the screened clone libraries. A broad taxonomic range of sequences from community DNA was identified. These include Bacteria, Crenarchaeota, Euryarchaeota and Nanoarchaeota representatives (Table 3.1-3.4).



**Table 3.1:** Representatives of the commonly observed crenarchaeal phylotypes from Chinese and New Zealand samples.

Clone representative	Blast match	Accession number	Similarity (%)	E-Value
T17A4	<i>Ignicoccus pacificus</i>	AJ271794	95	0.0
T178	<i>Pyrobaculum calidifontis</i>	AB078332	96	0.0
T17B5	<i>Pyrobaculum neutrophilum</i>	X81886	97	0.0
KAZ343	<i>Pyrodictium occultum</i>	M21087	85	0.0
TC11-2	 <i>Desulfurococcus mobilis</i>	M36474	97	0.0
T3C14	<i>Sulfolobus shibatae</i>	M32504	93	0.0
T3C20	<i>Sulfolobus islandicus</i>	AY247899	93	0.0
TK17C5	<i>Staphylothermus marinus</i>	X99560	94	0.0
T3B18	<i>Acidianus convivator</i>	AJ634763	94	0.0
T178	<i>Pyrobaculum organotrophum</i>	AB063647	96	0.0



**Table 3.2:** Representatives of the commonly observed euryarchaeal phylotypes from Chinese and New Zealand samples.

Clone	Blast match	Accession number	Similarity (%)	E-Value
KAZ174	<i>Archaeoglobus fulgidus</i>	AE000965	96	0.0
T172	<i>Archaeoglobus veneficus</i>	Y10011	96	0.0
T17A6	<i>Thermococcus waimanguensis</i>	AF098975	99	E-160
TC9-2	<i>Archaeoglobus profundus</i>	AF297529	96	0.0
T3C8	<i>Thermococcus waiotapuensis</i>	AY099187	99	0.0
KAZ1724	<i>Thermococcus peptonophilus</i>	AJ298871	95	0.0

**Table 3.3:** Representatives of the commonly observed Bacteria phylotypes from Chinese and New Zealand samples.

Clone	Blast match	Accession number	Similarity (%)	E-Value
KAZ344	<i>Aquifex aeolicus</i>	AJ309733	94	E-125
KAZ313	<i>Hydrogenivirga caldilitoris</i>	AB120294	95	0.0
TC11-12	<i>Hydrogenobacter subterranea</i>	AB026268	98	0.0
KAZ311	<i>Aquifex pyrophilus</i>	M83548	95	0.0

**Table 3.4:** Nanoarchaeal phylotypes from Chinese and New Zealand samples,

Clone	Blast match	Accession number	Similarity (%)	E-Value
KAZ315	<i>N. equitans</i>	AJ318041	85	E-124
T173	<i>N. equitans</i>	AJ318041	90	E-104
TC11D4	<i>N. equitans</i>	AJ318041	90	7E-86
TC11E3	<i>N. equitans</i>	AJ318041	87	7E-64
TC11-6	<i>N. equitans</i>	AJ318041	90	5E-86
TC11-18	<i>N. equitans</i>	AJ318041	89	1E-83
T17C9	<i>N. equitans</i>	AJ318041	90	E-118
TC11-4	<i>N. equitans</i>	AJ318041	89	1E-90
KAZ353	<i>N. equitans</i>	AJ318041	89	E-127

Table 3.1 to 3.3 shows some of the phylotypes that were commonly observed in Chinese and New Zealand samples. The presence of unique phylotypes provides an estimate of the microbial diversity on these biotopes, but is not a true reflection of the existing microbial diversity. This limitation may be due to a number of factors, such as the amplification bias from primer selection and PCR conditions. For culturable microorganisms, unfavorable culturing conditions can be another factor, since some environmental conditions can be difficult to mimic in the laboratory, thereby limiting a number of microorganisms from being detected.

### 3.6 Culturing and Microscopy



In order to visualize host-nanoarchaeon co-cultures, samples were cultured anaerobically on two different media, C7.4 and Master media (MM) at 80 and 85 °C. Light microscope was used to monitor the growth of cells. There was no apparent difference on the growth of cell using these two media. Table 3.5 demonstrates growth analysis observed by light microscope.

**Table 3.5:** Growth analysis of thermophiles from TC9, TC11, TOK3 and TOK17 samples. The numbers on brackets indicates the incubation time.

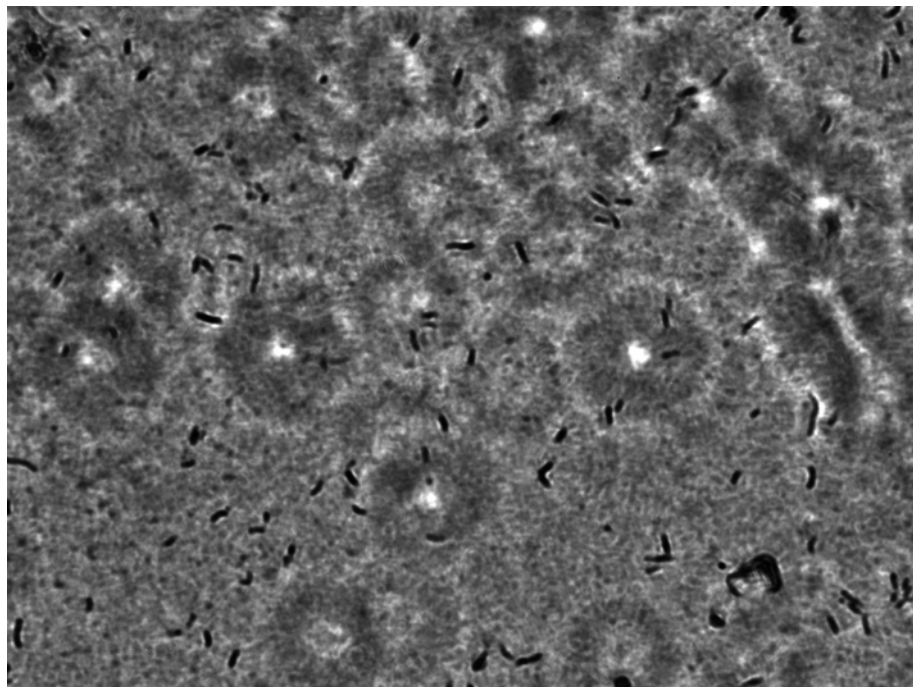
80 °C				
pH	TC9 (50 h)	TC11(50 h)	TOK3 (51 h)	TOK17(51 h)
6.0	Rods and cocci >10 cells/view	Filaments and small rods >20 cells/view	N/A	N/A
6.5	Rods and cocci >10 cells/ view	Few rods >10 cells/view 20 cocci/view few filaments	N/A	30 cocci 3-4 rods
7.5	Rods and cocci 9 cells/ view	Filaments, <10 cocci and 4 rods/view	>20/view cocci	100 cocci/view and 3-4 rods
MM	5 cocci per view	Filaments, <10 cocci and 4 rods/view	>20/view cocci	100 cocci/view
85 °C				
pH	TC9 (35 h)	TC11 (39 h)	TOK3 (37 h)	TOK17 (38 h)
6.0	Diplococcus rods, and cocci , >10 cells/view	cocci and rods with flagelli, 20 cells/view	rods and cocci, >20 cells/view	
6.5	Few cocci	Few rods, >10/view	rods and cocci, >10 cells/view	N/A
7.5	Few rods, >3 cells/view, 7-8 cocci/view	Filaments and cocci, >20 cells/view	Rods and cocci, some with flagelli, >20 cells/view	cocci and diplococcus rods, >20 cells/view
MM	rods and cocci, >10 cells/view	Filaments, numerous cocci and rods >20 cells/view	N/A	cocci and diplococcus rods, >10 cells/view

Both media and temperature conditions generated active growth, as observed by cell mobility (Fig. 3.20). The ability of other hyperthermophiles to grow in these media was a positive indication of the potential for nanoarchaeotes to grow in co-culture, provided that their thermophilic hosts were present. Due to small size of

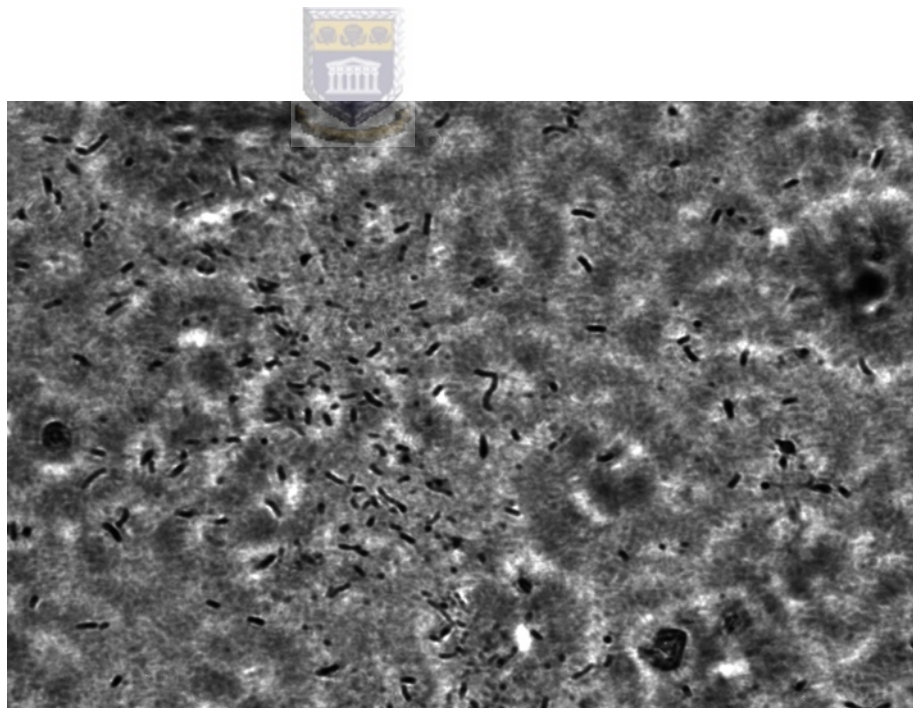
nanoarchaeotes, it was impossible to visualize them under light microscopy. Alternative microscopic methods, suitable for visualizing nanoarchaeotes include fluorescence *in situ* hybridization (FISH) and Scanning Electron Microscopy (SEM). Light microscopy revealed the basic morphology of the cells present, including both rods and cocci (Fig. 3.20). SEM experiments were performed on cultured cells, and preliminary results showed coccoid cells of about 2  $\mu\text{m}$  and 400 nm in diameter (Fig. 3.21A). These sizes correspond to general dimensions of Thermococcales and nanoarchaeote cells, respectively. The nano-sized cells are possibly nanoarchaeotes, and the bigger cells may be their potential hosts. No nano-sized cells were observed attached to any coccoid cell. However, this is not surprising as nanoarchaeotes are known to detach from the host during the late exponential growth phase (Huber *et al.* 2002). Another factor that may result in the nano-sized cells being detached from their host cells could be due to centrifugation. The cells were pelleted after culturing and the supernatant discarded. It is possible that the nanoarchaeotes, if present, could have detached from their host and were largely lost in the supernatant.

Initial FISH experiments were tested on environmental samples and positive signal was observed, although substantial interference was noted from non-specific binding and soil debris (Fig. 3.22). Fluorescence *in situ* hybridization on cultured cells was unsuccessful. This was probably due to cell damage as a result of freezing and thawing. For future experiments, the cultured cells need to be stored in glycerol, or be fixed immediately with glutaraldehyde for FISH experiments.

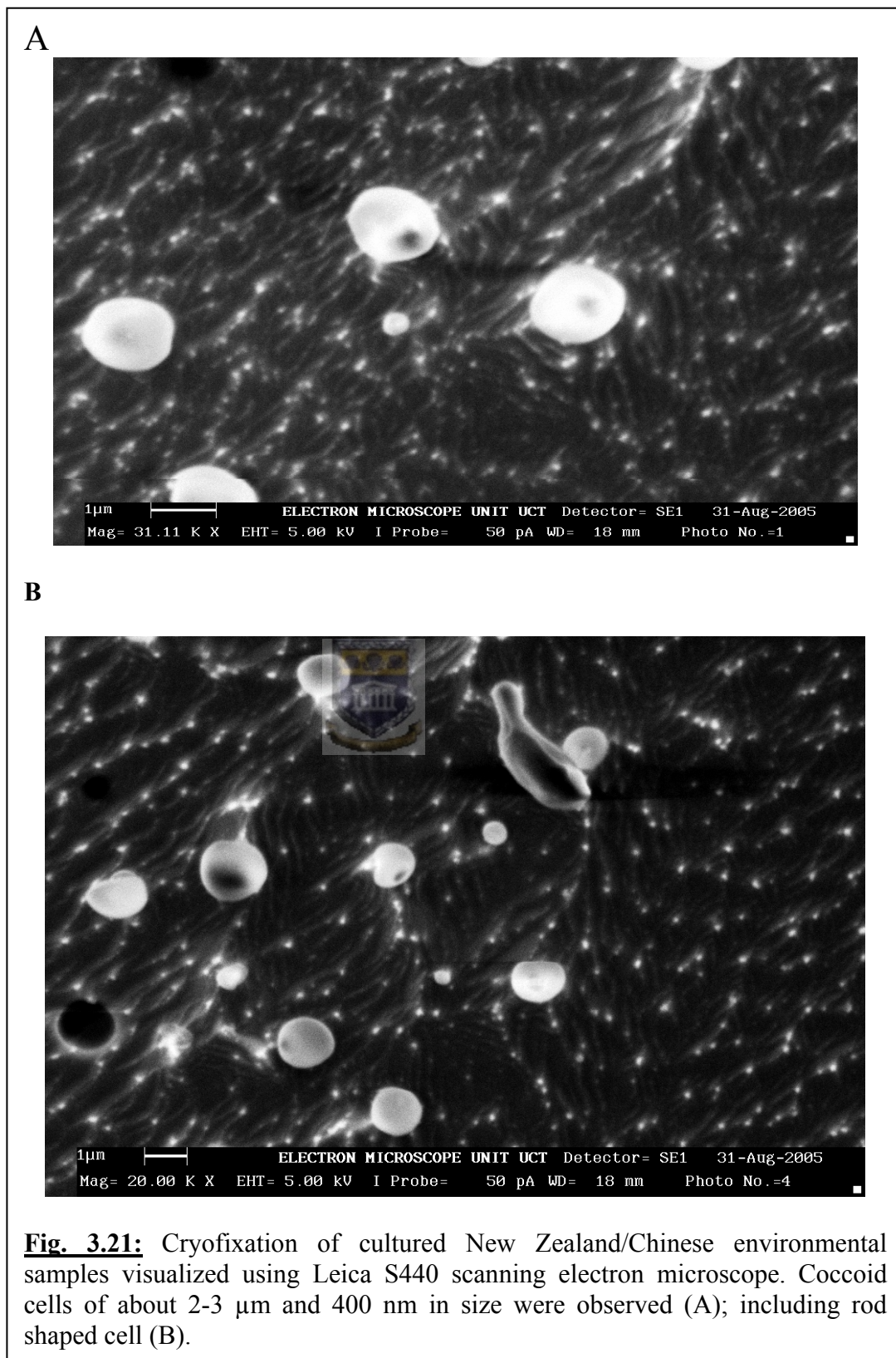
A

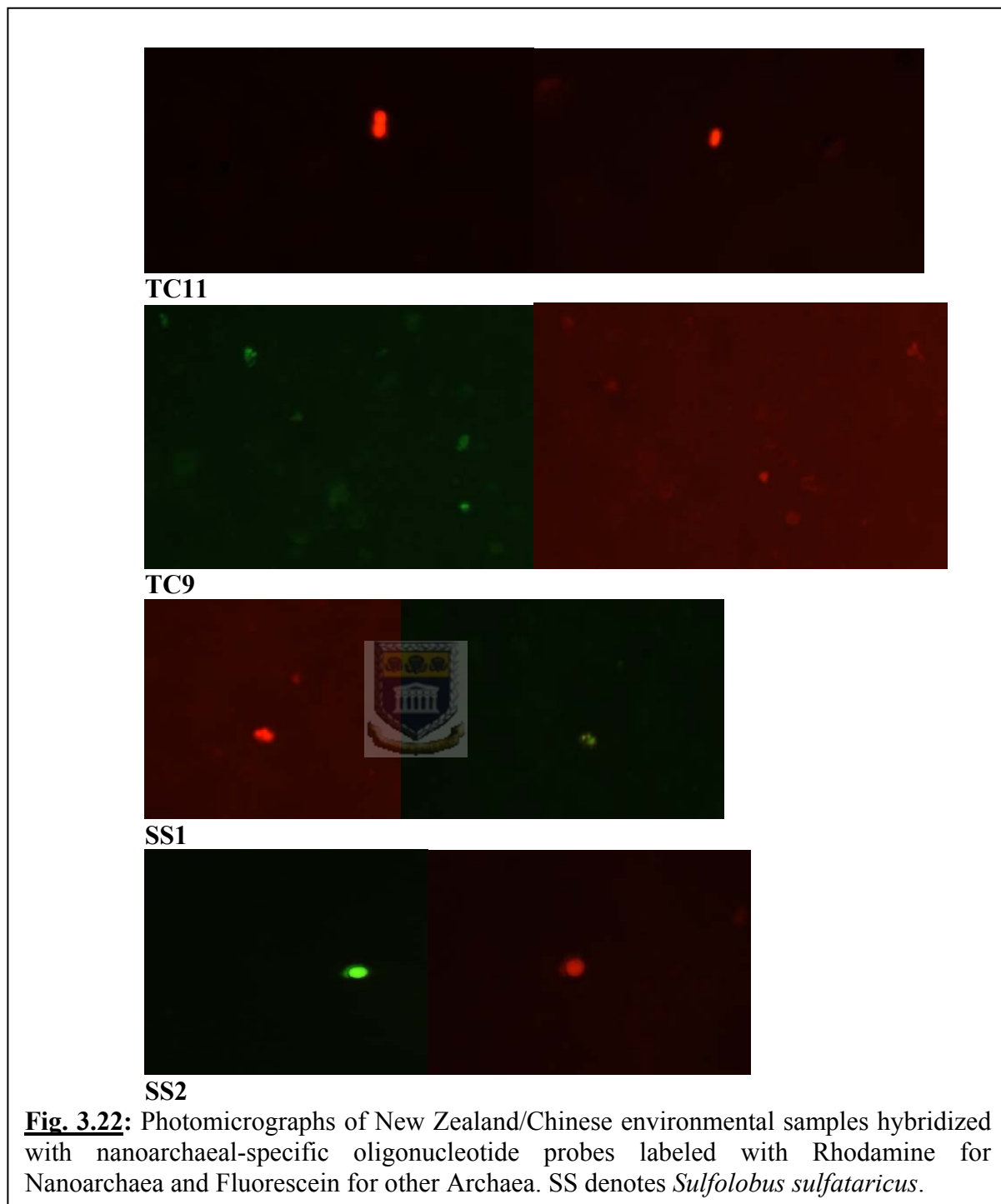


B



**Fig. 3.20** : Photomicrographs of New Zealand and Chinese environmental samples cultured at 80 and 85 °C. Cocci and rod shaped cells were observed under light microscope.





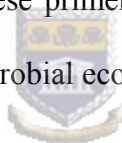


Ribosomal targeted oligonucleotide probes were hybridized to paraformaldehyde fixed environmental samples to facilitate identification of Nanoarchaea with the fluorescence microscope. Hybridization with both of the hybridization probes yielded strong signals but with no apparent specificity to nanoarchaeotes. No fluorescence was observed when the antisense probe was used. Figure 3.22 shows signals from fluorescent labeling experiments. TC11 shows a red fluorescence due to the use of a rhodamine labeled probe. TC9 gives both a Fluorescein (green) and rhodamine (red), signal as an indication of the presence of archaeal and possibly nanoarchaeal cells. SS (SS, *Sulfolobus sulfataricus*) shows both red and green signals on the same spot. This indicates that both probes might be binding non-specifically to all archaea. With the identification of a wide range of phylotypes, the challenge for future studies is to design new probes specific for nanoarchaeotes. Positive signals with such probes would assist with an understanding of the morphology and localization of nanoarchaeotes within the microbial community.

## 3.7 General Discussion

### 3.7.1 PCR

DNA from all samples was PCR amplified using primers developed for this study. All four samples produced the required amplicons, which were ligated into the pTZ57R/T vector and successfully cloned in *E. coli* DH5<sub>α</sub> strain. The A571F/U1204R primers could amplify 16S rRNA gene sequences from all three archaeal domains; the Crenarchaeota, Euryarchaeota and Nanoarchaeota. These primers were shown to amplify a broad taxonomic range of sequences from community DNA. With the identification of new taxa, the continued use of the so called “universal” primers must be assessed. Confirmation that these primers remain truly “universal” is critical for their continued use for general microbial ecology.



### 3.7.2 Amplified rDNA restriction analysis (ARDRA)

ARDRA was used extensively on this study, and with great success. ARDRA enabled the differentiation of phylotypes through digestion with different enzymes. Phylotypes that generated unique patterns were sequenced and analyzed further by phylogenetic analysis tools. ARDRA not only contributes to an exploration of the diversity of the microbial communities of the different samples analyzed, but also minimized the experimental cost. Instead of sequencing all screened phylotypes, only representative clones were sequenced. As mentioned previously, prior checking of

restriction sites by using database sequences is crucial for a good endonuclease selection.

### 3.7.3 Phylogenetic analysis

Phylogenetic analysis of nucleotide sequences derived from Chinese and New Zealand libraries has placed nanoarchaeotes into three distinct groups, based on their geological origin. These groups are Chinese, New Zealand and the published nanoarchaeotes. Other phylogenetic groups that were identified from screened samples include *Thermococcus* species, *Archaeoglobus* (Euryarchaea), *Pyrobaculum* and *Sulfolobus* species (Crenarchaea). As these genera inhabit the same environments as Nanoarchaea, they may be assumed to be potential host cells. These results raise the question of whether *Ignicoccus* is the only host of Nanoarchaea, or whether other hyperthermophiles may serve as host species. It is not known whether nanoarchaeotes are truly parasitic, or live in symbiotic association with other thermophiles. These questions may be addressed with the isolation of new host-parasite pairs.

### 3.7.4 Microscopy

The light microscope enabled visualization of live cells, and active growth was observed during culturing. SEM experiments showed coccoid cell that fall within the nanoarchaeal cell diameter. This observation is promising for the presence of nanoarchaeotes; however the nano-sized coccoid may not be a nanoarchaeote. In the

presence of host-nanoarchaeal pairs, SEM would reveal the mechanism of attachment of nanoarchaeotes to the host cells and differences in cell morphology would be identified.

FISH may be used to identify nanoarchaeotes and to visualize their physical localization with their host cells. Though FISH was not successful on cultured cells, it is still a promising tool to be used to explore the morphology and localization of nanoarchaeotes. Using FISH for future studies will allow simultaneous visualization, identification, enumeration and localization of individual cells.

All of the three microscope methods used have their own advantages and disadvantages. Combination of all three will allow effective identification. Cultured cells allow effective identification due to higher cell densities than with environmental samples. The cells present in environmental sample might be attached to soil particles or other sample debris and may give false positive results. It is also advisable not to freeze the cells prior to FISH, for this will damage the cells due to heat shock. For future studies on this project the challenge is to modify the nanoarchaeal probes, based on the novel nanoarchaeal phylotype sequences retrieved from this study, and to use these probes on properly fixed fresh cultured microbial consortium.

In conclusion, the diversity of nanoarchaeotes was explored on two mini-libraries from two continents. Eleven phylotypes were identified from a Chinese library and 4

from New Zealand library. This study has shown that there is genetic diversity between the novel nanoarchaeal phylotypes and the published sequences and that they cluster into three groups according to their geographical origin.



## REFERENCES

1. **Amann, R. and K hl, M. (1998)** *In situ* methods for assessment of microorganisms and their activities. *Current Opinion in Microbiology*, **1**: 352-358.
2. **Amann, R.I., Ludwig, W. and Schleifer, K-H. (1995)** Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiological Reviews*, **59**: 143-169.
3. **Amann, R., Snaidr, J., Wagner, M., Ludwig, W. and Schleifer, K.H. (1996)** In situ visualization of high genetic diversity in a natural microbial community. *Journal of Bacteriology*, **178**: 3496-3500.
4. **Baker, G.C., Smith, J.J., and Cowan D.A. (2003)** Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods*, **55**: 541-55.
5. **Barns, S.M., Delwiche, C.F., Palmer, J.D. and Pace. N.R. (1996)** Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proceedings of the National Academy of Sciences U.S.A*, **93**: 9188-9193.
6. **Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F. and Forterre, P. (2005)** Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biology*, **6**: R42.1-R42.10.
7. **Brunk, C.F., Avaniss-Aghajani, E. and Brunk, C.A. (1996)** A computer analysis of primer and probe hybridization potential with bacterial small-subunit rRNA sequences. *Applied and Environmental Microbiology*, **62**: 872-879.
8. **Burggraf, S., Mayer, T., Amann, R., Schadhauer, S., Woese, C.R. and Stetter, K.O. (1994)** Identifying members of the domain archaea with rRNA-targeted oligonucleotide probes. *Applied and Environmental Microbiology*, **60**: 3112-3119.
9. **Cooper, L. (2002)** Hipaua Steaming Cliffs and Little Waihi Landslide. *Tephra*, **19**: 42-47.
10. **Dabert, P., Delgen s, J-P., Moletta, R. and Godon, J-J. (2002)** Review: Contribution of molecular microbiology to the study in water pollution removal of microbial community dynamics. *Environmental Science and Biotechnology*, **1**: 39-49
11. **Dawson, S.C., DeLong E.F. and N.R. Pace (2001)** Phylogenetic and ecological perspectives on uncultured Crenarchaeota and Korarchaeota. In Dworkin *et al.* (eds.). *The Prokaryotes: An evolving electronic database for the microbiological community*, 3<sup>rd</sup> ed., release 3.7, Springer-Verlag, NY.
12. **Farrelly, V., Rainey, F.A. and Stackebrandt, E. (1995)**. Effect of genome size and *rrn* gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Applied and Environmental Microbiology*, **61**: 2798-2801.
13. **Fuchs, T., Huber, H., Teiner, K., Burggraf, S., Stetter, K.O. (1995)** *Metallosphaera prunae*, sp. nov., a novel metal-mobilizing thermoacidophilic

- Archaeum, isolated from a uranium mine in Germany. *Systematic and Applied Microbiology*, **18**: 560-566.
14. **Gich, F.B., Amer, E., Figueras, J.B., Abella, C.A., Balaguer, M.D., Poch, M. (2000)** Assessment of microbial community structure changes by amplified ribosomal DNA restriction analysis (ARDRA). *International Microbiology*, **3**: 103–106.
  15. **Goldstein, J.I., Newbury, D.E., Echlin, P., Joy, D.C., Romig, A.D., Lyman, Jr., C.E., Fiori, C. and Lifshin, E. (1992)** Scanning electron microscopy and X-ray microanalysis, 2<sup>nd</sup> ed., pp. 1-819, Plenum Press, NY.
  16. **Grant, M., Ventosa, A., Cowan, D., Yanhe, M. and Jones, B. (2003).** Identification and sustainable use of metabolic and diversity as a source of new valuable products. *MultiGenome Access Technology for Industrial Catalysts*, 1-12.
  17. **Haas, S., Vingron, M., Poustka, A. and Wiemann, S. (1998)** Primer design for large scale sequencing. *Nucleic Acids Research*, **26**: 3006-3012.
  18. **Heyndrickx, M., Vauterin, L., Vandamme, P., Kersters, K. and De Vos, P. (1996)** Applicability of combined amplified ribosomal DNA restriction analysis (ARDRA) patterns in bacterial phylogeny and taxonomy. *Journal of Microbiology Methods*, **26**: 247-259.
  19. **Hohn, M.J., Hedlund, B.P. and Huber, H. (2002)** Detection of 16S rDNA sequences representing the novel phylum “Nanoarchaeota”: Indication for a wide distribution in high temperature biotopes. *Systematic and Applied Microbiology*, **25**: 551-554.
  20. **Hongpattarakere, T. (2002)** Hyperthermostable cellulolytic and hemicellulolytic enzymes and their biotechnological applications. *Songklanakarinn Journal of Science and Technology*, **24**: 481-491.
  21. **Holt, J.G., Krieg, N.R., Sneath, P.H.A., Staly, J.T. and Williams, S.T. (1994)** *Bergey's manual of determinative bacteriology*, 9<sup>th</sup> ed., pp. 1-787, Williams & Wilkins, Baltimore.
  22. **Hough, D.W. and Danson, M. (1999)** Extremozymes. *Current Opinion in Chemical Biology*, **3**: 39-46.
  23. **Huber, G. and Stetter, K.O. (1991)** *Sulfolobus metallicus*, sp. nov., a novel strictly chemolithoautotrophic thermophilic archaeal species of metal mobilizers. *Systematic and Applied Microbiology*, **14**: 372-378.
  24. **Huber, H. and Stetter, K.O. (1998)** Review article: Hyperthermophiles and their possible potential in biotechnology. *Journal of Biotechnology*, **64**: 39-52.
  25. **Huber, H. and Stetter, K.O. (2001)** Family I. *Desulfurococcaceae*. In: Boone *et al.* (eds.), The archaea and the deeply branching and phototrophic bacteria, *Bergey's Manual of Systematic Bacteriology*, 2<sup>nd</sup> ed., vol. 1, pp. 180–181, Springer, New York.
  26. **Huber, H., Burggraf, S., Mayer, T., Wyszchony, I., Rachel, R. and Stetter, K.O. (2000)** *Ignicoccus* gen nov., a novel genus of hyperthermophilic, chemolithoautotrophic archaea, represented by two new species, *Ignicoccus islandicus* sp. nov. and *Ignicoccus pacificus* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*, **50**: 2093-2100.

27. **Huber, H., Hohn, M.J., Rachel, R., Fuchs, T., Wimmer, V.C. and Stetter, K.O. (2002)** A new phylum of archaea represented by a nanosized hyperthermophilic symbiont. *Nature*, **417**: 63-67.
28. **Huber, H., Hohn, M.J., Stetter, K.O. and Rachel, R. (2003)** The phylum Nanoarchaeota: Present knowledge and future perspectives of a unique form of life. *Research in Microbiology*, **154**: 165-171.
29. **Itoh, T. (2003)** Taxonomy of nonmethanogenic hyperthermophilic and related thermophilic archaea. *Journal of Bioscience and Bioengineering*, **96**: 203–212.
30. **Inoue, H., Nojima, H. and Okayama, H. (1990)** High efficiency transformation of *Escherichia coli* with plasmids. *Gene*, **96**: 23-28.
31. **Jones, B., Renaut, R.W., Rosen, M.R. (2004)** Taxonomic fidelity of silicified filamentous microbes from hot-spring systems in the Taupo volcanic zone, North Island, New Zealand. *Transactions: Earth Sciences*, **94**: 475-483.
32. **Jukes, T.H., and Cantor C.R. (1969)** Evolution of protein molecules. In: Munro (ed.), Mammalian protein metabolism, pp. 21–32, *Academic Press*, NY.
33. **Kampke, T., Kieninger, M. and Mecklenbug, M. (2001)** Efficient primer design algorithms. *Bioinformatics*, **17**: 214-225.
34. **Kashefi, K and Lovley, D.R. (2003)** Extending the upper temperature limit for life. *Science*, **301**: 934.
35. **Koch, R., Sprengel, A., Lemke, K., Antranikyan, G. (1991)** Purification and properties of a hyperthermoactive  $\alpha$ -amylase from the archaeobacterium *Pyrococcus woesslei*. *Archaea Microbiology*, **155**: 572-578.
36. **Komminoth, P. and Long, A.A. (1995)** *In situ* polymerase chain reaction methodology, applications and non-specific pathways. In: Mannheim (ed.), PCR application manual, 97-106, Mannheim, Germany.
37. **Kovárová, M. and Dráber, P. (2000)** New specificity and yield enhancer of polymerase chain reactions. *Nucleic Acids Research*, **28**: E70-e70.
38. **Kreder, C.A. (1996)** Relief of amplification in PCR with bovine serum albumin or T4 gene 32 protein. *Applied and Environmental Microbiology*, **62**: 1102-1107.
39. **Kvist, T., Mengewein, A., Manzei, S., Ahring, B.K. and Westermann, P. (2005)** Diversity of thermophilic and non-thermophilic crenarchaeota at 80 °C. *Federation of European Microbiological Societies Microbiology Letters*, **244**: 61–68.
40. **Kwok, S., Chang, S.Y., Sninsky, J.J. and Wang, A. (1994)** A guide to the design and use of mismatched and degenerate primers. *PCR Methods and Applications*, **3**: S39-47.
41. **Lee, N., Nielsen, P.H., Andreasen, K.H., Juretschko, S., Nielsen, J.L., Schleifer, K.-H., and Wagner, M. (1999)** Combination of fluorescent *in situ* hybridization and microautoradiography – a new tool for structure-function analyses in microbial ecology. *Applied and Environmental Microbiology*, **65**: 1289-1297.
42. **Leuschner, C. and Antranikian, G. (1995)** Heat-stable enzymes from extremely thermophilic and hyperthermophilic microorganisms. *World Journal of Microbiology and Biotechnology*, **11**: 95-114.



43. **Lexa, M., Horak, J., and Brzobohaty, B. (2001)** Virtual PCR. *Bioinformatics*, **17**: 192-193.
44. **Li, P., Kupfer, K.C., Davies, C.J., Burbee, D., Evans, G.A. and Garner, H.R. (1997)** PRIMO: A primer design program that applies base quality statistics for automated large-scale DNA sequencing. *Genomics*, **40**: 476-485.
45. **Lowe, T.M., Sharefkin, J., Yang, S.Q., Dieffenbach, C.W. (1990)** A computer program for selection of oligonucleotide primers for polymerase chain reactions. *Nucleic Acids Research*, **18**: 1757-1761.
46. **Margulis, L. and Sagan, D. (1986)** Microcosmos: Four billion years of evolution from our microbial ancestors. Summit Books, New York.
47. **Marsh, R.M and Norris, P.R. (1985)** The isolation of some thermophilic, autotrophic iron- and sulfur-oxidising bacteria. *Federation of European Microbiological Societies Microbiology Letters*, **17**: 311-315.
48. **Moter, A. and Göbel, U.B. (2000)**. Fluorescence *in situ* hybridization (FISH) for direct visualization of microorganisms. *Journal of Microbiological methods*, **41**: 85-112.
49. **Nissen, A.N., Anker, L., Munk, N., Lange, N. K. (1992)** Xylanases for the Pulp and Paper Industry. In: Visser *et al.* (eds.), Xylan and Xylanases, pp. 325-337, *Elsevier Science Publishers*, Amsterdam.
50. **Ochman, H. and Moran, N.A. (2001)** Genes lost and genes found: Evolution of bacterial pathogenesis and symbiosis. *Science*, **292**: 1096-1099.
51. **O'Donnell, T. and Head, I.M. (1997)** Rapid enumeration of specific bacterial population. In *Molecular Microbial Ecology: Applications in wastewater treatment*, 18-20, University of Newcastle upon Tyne.
52. **Proutski, V. and Holmes, E. (1996)** Primer Master: a new program for the design and analysis of PCR primers. *Computer Applications in the Biosciences*, **12**: 253-255.
53. **Ranjard, L., Poly, F. and Nazaret, S. (2000)** Monitoring complex bacterial communities using culture-independent molecular techniques: Application to soil environment. *Research Microbiology*, **151**: 167-177.
54. **Robe, P., Nalin, R., Capellano, C., Voges, T.M. and Simonet, P. (2003)** Extraction of DNA from soil. *European Journal of soil Biology*, **39**:183-190.
55. **Rychlik, W., Spencer, W.J., and Rhoads, R.E. (1990)** Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Research*, **18**: 6409-6412.
56. **Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A. (1988)** Primer directed enzymatic amplification of DNA with thermostable DNA polymerase. *Science*, **239**: 487-491.
57. **Schiraldi, C. Giuliano, M. and De Rosa, M. (2002)** Perspectives on biotechnological applications of Archaea. *Archaea*, 75- 86.
58. **Schönhuber, W., Bourhis, G-L., Tremblay, J., Amann, R. and Kulakauskas, S. (2001)** Utilization of tmRNA sequences for bacterial identification. *BioMed Central Microbiology*, **20**:1471-2180.
59. **Schramm, A. and Amann, R. (1999)** Nucleic acid-based techniques for analyzing the diversity, structure, and dynamics of microbial communities in wastewater treatment. In: Rehm *et al.* (eds.), *Biotechnology: A multi volume*

- comprehensive treatise, Environmental processes I, Vol. 11a, pp. 85–108, Wiley-VCH, Weinheim.
60. **Schretter, C., and Milinkovitch, M. C. (2005)** Oligonucleotide design by multilevel optimization. *Unit of Evolutionary Genetics*, Technical report, 1-9.
  61. **Segerer, A., Neuner, A., Kristjansson, J.K. and Stetter, K.O. (1986)** *Acidianus infernus* gen. nov., sp. nov., and *Acidianus brierleyi* comb. nov.: Facultatively aerobic extremely acidophilic thermophilic sulfur-metabolizing archaeobacteria. *International Journal of Systematic Bacteriology*, **36**: 559-564.
  62. **Stach, J.E.M., Bathe, S., Clapp, J.P., Burns, R.G. (2001)** PCR-SSCP comparison of 16S rDNA sequence diversity in soil DNA obtained using different isolation and purification methods. *Federation of European Microbiological Societies Microbiology Ecology*, **36**: 139–151.
  63. **Stackebrandt, E. and Lang, E. (1998)** Molecular detection and identification of prokaryotes. OECD workshop molecular methods for safe drinking water, 1-10.
  64. **Stahl, D.A. and Amann, R. (1991)** Development and application of nucleic acid probes in bacterial systematics. In *Nucleic Acid Techniques in Bacterial Systematics*. Stackebrandt and Goodfellow (eds.), pp. 205-248, John Wiley & Sons, Chichester.
  65. **Stakenborg, T., Vicca, J., Butaye, P., Maes, D., De Baere, T., Verhelst, R., Peeters, J., de Kruif, A., Haesebrouck, F. and Vaneechoutte, M. (2005)** Evaluation of amplified rDNA restriction analysis (ARDRA) for the identification of mycoplasma species. *BioMed Central Infectious Diseases*, **5**: 46.
  66. **Stetter, K. O. (1996)** Hyperthermophilic prokaryotes. *Federation of European Microbiological Societies Microbiology Reviews*, **18**:149-158.
  67. **Suzuki, M., and Giovannoni, S.J. (1996)** Bias caused by template annealing in the amplification mixtures of 16S rRNA genes by PCR. *Applied and Environmental Microbiology*, **62**: 625-630.
  68. **Takai, K., Sugai, A., Itoh, T. and Horikoshi, K. (2000)** *Palaeococcus ferrophilus* gen. nov., sp. nov., a barophilic, hyperthermophilic archaeon from a deep-sea hydrothermal vent chimney. *International Journal of Systematic and Evolutionary Microbiology*, **50**: 489–500.
  69. **Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994)** CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**: 4673-4680.
  70. **Van de Peer, Y., Jansen, J., De Rijk P., and De Wachter, R. (1997)** Database on the structure of small ribosomal subunit RNA. *Nucleic Acids Research*, **25**: 111-116.
  71. **Vaupen, S.B. (1999)** Renewable Energy Markets in China: An analysis of renewable energy markets in Guangdong, Jiangxi, Jilin, and Yunnan provinces with updated information from Beijing. *Subcontractor Report*, 1-105.
  72. **Wagner, M., Amann, R., Lemmer, H., Schleifer, K.H. (1993)** Probing activated sludge with proteobacteria-specific oligo-nucleotides: Inadequacy of

- culture-dependent methods for describing microbial community structure. *Applied and Environmental Microbiology*, **59**:1520-1525.
73. **Waters, E., Hohn, M.J., Ahel, I., Graham, D.E., Adams, M.D., Barnstead, M., Beeson, K.Y., Bibbs, L., Bolanos, R., Keller, M., Kretz, K., Lin, X., Mathur, E., Ni, J., Podar, M., Richardson, T., Sutton, G.G., Simon, M., Söll, D., Stetter, K.O., Short, J.M. and Noordewier, M. (2003)** The genome of *Nanoarchaeum equitans*: Insights into early archaeal evolution and derived parasitism. *Proceedings of the National Academy of Sciences*, **22**: 100:12984-12988.
74. **Wei, X., Kuhn, D.N. and Narasimhan, G. (2003)** Degenerate primer design via clustering. *Proceedings of the IEEE Computer Society Bioinformatics Conference*, 75-83.
75. **Wintzingerode, V.F., Göbel, U.B. and Stackebrandt, E. (1997)** Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *Federation of European Microbiological Societies Microbiology Reviews*, **21**: 213-229.
76. **Woese, C.R. and Fox, G.E. (1977)** Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, **74**: 5088-5090.
77. **Zheng, D., Alm, E.W., Stahl, D.A. and Raskin, L. (1996)** Characterization of universal small-subunit rRNA hybridization probes for quantitative molecular microbial ecology studies. *Applied and Environmental Microbiology*, **62**: 4504-4513.
78. **Zimmer, C. and Wähnert, U. (1986)** Nonintercalating DNA-binding ligands: specificity of the interaction and their use as tools in biophysical, biochemical and biological investigations of the genetic materials. *Progressive Biophysical Molecular Biology*, **47**: 31-112.



Draft Manuscript for Review

**New Nanoarchaeal-Specific Primers and their Use in the Identification of Novel Nanoarchaeotes in Geographically Diverse Hydrothermal Sediments**

Journal:	<i>Extremophiles</i>
Manuscript ID:	EXT-06-Jan-0006
Manuscript Type:	Original Paper
Date Submitted by the Author:	09-Jan-2006
Complete List of Authors:	Galada, Ncebakazi; University of the Western Cape, Department of Biotechnology Casanueva, Ana; University of the Western Cape, Department of Biotechnology Baker, Gillian; University of the Western Cape, Department of Biodiversity & Conservation Biology Grant, William; Univ. of Leicester, Infection, Immunity and Inflammation Heaphy, Shaun; University of Leicester, Infection, Immunity and Inflammation Jones, Brian; Genencor International B.V., - Yanhe, Ma; Chinese Academy of Sciences, Institute of Microbiology Ventosa, Antonio; University of Sevilla, Department of Microbiology and Parasitology, Faculty of Pharmacy Blamey, Jenny; Fundacion Biociencia Cowan, Don; University of the Western Cape, Department of Biotechnology
Keyword:	Nanoarchaeota, 16S rRNA, phylogeny, primer, hyperthermophilic

powered by ScholarOne  
Manuscript Central™

1 **New Nanoarchaeal-Specific Primers and their Use in the Identification of Novel**  
2 **Nanoarchaeotes in Geographically Diverse Hydrothermal Sediments**

3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13 Ncebakazi Galada<sup>1</sup>, Ana I Casanueva<sup>1</sup>, Gillian C Baker<sup>2</sup>, William D Grant<sup>3</sup>, Shaun  
14  
15 Heaphy<sup>3</sup>, Brian Jones<sup>4</sup>, Ma Yanhe<sup>5</sup>, Antonio Ventosa<sup>6</sup>, Jenny Blamey<sup>7</sup> and Don A Cowan<sup>1\*</sup>  
16  
17

18  
19  
20 Department of Biotechnology<sup>1</sup> and Department of Biodiversity & Conservation Biology<sup>2</sup>, University of the  
21  
22 Western Cape, Bellville, 7535, South Africa; Department of Microbiology and Immunology, University of  
23  
24 Leicester, LE1 7RH, UK<sup>3</sup>; Genencor International, Leiden, The Netherlands<sup>4</sup>; Institute of Microbiology,  
25  
26 Chinese Academy of Sciences, 100080, Beijing, China<sup>5</sup>; Department of Microbiology and Parasitology,  
27  
28 University of Seville, 41012, Seville, Spain<sup>6</sup>; Fundación Biociencia, José Domingo Cañas 2280, Ñuñoa,  
29  
30 Santiago, Chile<sup>7</sup>.



31  
32  
33 Correspondence to: Don A Cowan; Department of Biotechnology, University of the  
34  
35 Western Cape, Private Bag X17, Bellville, 7535, South Africa; Phone: (2721) 959 2083;  
36  
37 Fax: (2721) 959 3505; Email: [dcowan@uwc.ac.za](mailto:dcowan@uwc.ac.za)  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

18 **Abstract**

19 The Nanoarchaeota were proposed as the fourth sub-division of the Archaea in 2002.  
20 The first, and only fully characterised Nanoarchaeon was found to exist in a symbiotic  
21 association with the hyperthermophilic Crenarchaeote, *Ignicoccus*. This nanoarchaeote,  
22 named *N. equitans*, did not amplify with “standard archaeal 16S PCR primers” and could  
23 only be amplified using specifically designed primers. We have designed two new sets of  
24 primers; “universal archaeal primers,” that amplify the 16S rRNA gene of all four  
25 archaeal sub-divisions, and “nanoarchaeal-specific primers based on all known  
26 nanoarchaeal 16S rRNA gene sequences. Using these primers we have amplified  
27 community DNA from Chinese, New Zealand and Chilean hydrothermal sites and have  
28 identified novel nanoarchaeal phylotypes. The novel sequences suggest the existence of  
29 new genera. The novel phylotypes are partly separated on the basis of geography; New  
30 Zealand phylotypes constituting a unique clade, but with similar phylotypes found in both  
31 Asia and South America.

32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

33 Keywords: Nanoarchaeota, 16S rRNA, phylogeny, primer, hyperthermophilic.

## 34 Introduction

35 The Archaea are a biotechnologically important group of organisms. In the past decade  
36 two new sub-divisions have been described within the Archaea: the Korarchaeota and  
37 Nanoarchaeota (Barns *et al.*, 1996; Huber *et al.*, 2002). The sub-division Nanoarchaeota  
38 is known from a single isolate (Huber *et al.*, 2003), two environmental DNA samples  
39 (Hohn *et al.*, 2002), 19 DNA sequences (McCliment *et al.*, 2006) and 8 DNA sequences  
40 from this study. *Nanoarchaeum equitans*, the type specimen for the Nanoarchaeota, is a  
41 400 nm coccus that is found on the cell surface of the chemolithotrophic Crenarchaeote  
42 *Ignicoccus*. Studies of the *Nanoarchaeum* genome suggest a parasitic lifestyle, as it must  
43 derive the majority of its nutrients from its host due to the absence of a number of  
44 metabolic pathways (Waters *et al.*, 2003). *Ignicoccus* cells grow equally well in pure  
45 culture as in co-culture with *Nanoarchaeum*, but too high a burden of *N. equitans* cells is  
46 reported to inhibit *Ignicoccus* growth, suggesting that their relationship is not mutualistic  
47 (Waters *et al.*, 2003).

48  
49 *Nanoarchaeum equitans* (Huber *et al.*, 2002) failed to amplify using published universal  
50 archaeal primers and all three known nanoarchaeal sequences were obtained by  
51 amplification with Nanoarchaea-specific 16S rDNA primers (Hohn *et al.*, 2002). We  
52 have recently designed a new set of primers with 100% similarity to sequences from all  
53 four existing archaeal sub-divisions with the objective of providing an unbiased  
54 assessment of archaeal diversity in environmental samples (Baker *et al.*, 2003). These  
55 primers proved to amplify archaeote type strains and targeted sequences of the latter from  
56 environmental samples that did not amplify using published universal archaeal primers

1  
2  
3  
4 57 (Baker and Cowan, 2004; Baker *et al.*, 2003). The acquisition of new nanoarchaeal  
5  
6 58 sequences has allowed us to design a new “nanoarchaeal-specific” PCR primer set. Here  
7  
8 59 we report the use of both primer sets to identify novel nanoarchaeal taxa from Chinese,  
9  
10 60 New Zealand and Chilean hydrothermal sediments.  
11  
12  
13  
14

## 15 62 **Materials and Methods**

16  
17  
18 63 **Sampling.** Hydrothermal sediment samples were taken from location TC11 and TC9  
19  
20 64 (pH 8.0, average temperature 83 °C) in the Rehai thermal region near Tenchong, SW  
21  
22 65 China; TOK17 (pH 6.0, 96 °C) and TOK3 (pH 6.5, 86.5 °C) in the Tokaanu region in the  
23  
24 66 North Island, New Zealand; TA22 (pH 6.3, 80 °C), TA101 (pH 7.0, 85 °C), TA62 (pH  
25  
26 67 8.0, 85 °C) and TA91 (pH 5.0, 85 °C) in the Géiseres del Tatio volcanic region, Chile.  
27  
28  
29  
30

31  
32 69 **DNA extraction, PCR and cloning.** DNA was extracted from duplicate 0.5g sediment  
33  
34 70 samples using a modified Zhou method (Stach *et al.*, 2001). 620 bp fragments of the 16S  
35  
36 71 rRNA gene were amplified from genomic DNA extracted from Chinese and New  
37  
38 72 Zealand samples, using universal archaeal primer pair A571F and UA1204R and a  
39  
40 73 standard protocol (Baker *et al.*, 2003). Using nanoarchaeal-specific primer pair A571F  
41  
42 74 (Baker *et al.*, 2003) and N961R (5'-cmattaaaccgcraccc-3'), 400 bp fragments of the 16S  
43  
44 75 rRNA gene were amplified from genomic DNA from Chilean samples. Amplicons were  
45  
46 76 purified using GFX™ PCR purification kit (Amersham Biosciences), ligated into  
47  
48 77 pTZ57R vector (InsT/Aclone™, Fermentas Life Sciences) and transformed into  
49  
50 78 electrocompetent DH5α *E. coli*.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3  
4 80 **Amplified Ribosomal DNA Restriction Analysis.** White colonies from the resulting  
5  
6 81 libraries were amplified by colony PCR using the original PCR primers. The amplicons  
7  
8 82 were digested subsequently with *RsaI*, *MboI*, *HaeIII* and *AluI* at 37 °C for 16 hours and  
9  
10 83 compared on 2.5% agarose gels by amplified rDNA restriction analysis (ARDRA).  
11  
12

13 84  
14  
15 85 **Phylogenetic analysis.** A single clone representing each of the ARDRA types was  
16  
17 86 sequenced using M13F and M13R primers. Sequences were checked for similarity to  
18  
19 87 known sequences using BLASTn (<http://www.ncbi.nlm.nih.gov>) and aligned to closely  
20  
21 88 related representatives from their taxonomic sub-division, using ClustalW. A neighbour-  
22  
23 89 joining tree based on Jukes-Cantor distances with 100 bootstrap replicates was derived  
24  
25 90 using TREECON (Van de Peer *et al.*, 1996).  
26  
27  
28  
29

## 30 91 31 32 92 **Results and Discussion**

### 33 34 93 **Phylogenetic Diversity in the Chinese and New Zealand Samples**

35  
36  
37 94 Inserts from 71 and 36 clones from TC11 and TC9 amplicon libraries, respectively, were  
38  
39 95 amplified using the original universal archaeal primers A571F and UA1204R. Digestion  
40  
41 96 of these using *RsaI*, *HaeIII* and *MboI* yielded 8 unique ARDRA patterns. Sequencing of  
42  
43 97 the different ARDRA patterns revealed the presence of euryarchaeal, crenarchaeal and  
44  
45 98 novel nanoarchaeal sequences. The phylogenetic positions of the novel nanoarchaeal  
46  
47 99 phlotypes are shown in Figure 1. Samples TC9 and TC11 from China contained  
50  
51 100 nanoarchaeal phlotypes which clustered into 5 closely related clades and differed from  
52  
53 101 each other by 1-12%. These new nanoarchaeotes shared 80-89% base identity with the  
54  
55 102 three previously published nanoarchaeal sequences and consistently grouped together  
56  
57  
58  
59  
60

1  
2  
3  
4 103 with a bootstrap support of 100% irrespective of the calculation algorithm applied (Fig.  
5  
6 104 1). These sequence data have been submitted to the GenBank database under accession  
7  
8 105 numbers AY727886 to AY727891, AY572420 and AY571283. The remaining sequences  
9  
10 106 in the TC11 and TC9 libraries clustered mostly with the Crenarchaeotes. TC11-C2 had  
11  
12 107 97% sequence similarity to *Sulfophobococcus zilligii*, whilst TC11-C3 and TC11-B2  
13  
14 108 formed a novel clade that clustered with uncultured Crenarchaeotes including  
15  
16 109 *Cenarchaeum symbiosum*. TC9-2F had 98% sequence identity to uncultured  
17  
18 110 Archaeoglobales.  
19  
20  
21  
22  
23

24  
25 112 The inserts from 281 clones from the New Zealand samples TOK3 and TOK17 were  
26  
27 113 amplified using the universal archaeal primers and digested with *RsaI*, *HaeIII* and *AluI*.  
28  
29 114 15 unique ARDRA patterns were identified in the New Zealand libraries, containing  
30  
31 115 representatives of Crenarchaeota, Euryarchaeota and Nanoarchaeota. The novel  
32  
33 116 nanoarchaeal sequences, TOK173, TOK17C9 and TOK315, which differ from each other  
34  
35 117 by 3-5%, were more closely related to the published nanoarchaeotes than to the Chinese  
36  
37 118 nanoarchaeotes (Fig. 1). These New Zealand phlotypes shared 85-91% identity with the  
38  
39 119 published nanoarchaeal sequences, as compared to 84-96% with the Chinese sequences.  
40  
41  
42 120 Other sequences from the New Zealand libraries had high similarities to crenarchaeotes,  
43  
44 121 such as *S. zilligii*, as well as euryarchaeotes, mostly *Thermococcus* and *Archaeoglobus*  
45  
46 122 species.  
47  
48  
49  
50

51  
52 123  
53  
54 124 *Nanoarchaeum equitans*, the only cultured representative of the Nanoarchaeota, is a  
55  
56 125 nano-sized exosymbiont or parasite (Waters *et al.*, 2003) of *Ignicoccus*, a marine  
57  
58  
59  
60

1  
2  
3 126 hyperthermophilic archaeon (Huber *et al.*, 2002). At present it is unknown whether the  
4  
5 127 uncultured members of the group are free-living or associated with a host. If the new  
6  
7  
8 128 Chinese and New Zealand Nanoarchaeotes are symbiotic or parasitic, we might infer that  
9  
10 129 their host is one of the other observed phylotypes. TC11-C2 is closely related to  
11  
12 130 *Sulfophobococcus zilligi*, a genus in the same family as *Ignicoccus*. This phylotype was  
13  
14 131 comparatively rare in the rRNA gene library. The diversity of amplified products is  
15  
16 132 however not directly related to the diversity of the original community as there is intrinsic  
17  
18 133 bias in DNA extraction and PCR techniques. In order to minimise this bias we pooled  
19  
20 134 separate DNA extractions and PCR amplicons prior to cloning. In this study we have  
21  
22 135 demonstrated concomitant amplification of DNA from genetically distinct taxa from  
23  
24 136 three sub-divisions of Archaea. This provides support that our extraction and  
25  
26 137 amplification procedures were not taxonomically biased, but it does not guarantee an  
27  
28 138 accurate reflection of phylotype abundance.  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

#### 140 **New “nanoarchaeal-specific” primer design**

141 A new “nanoarchaeal-specific” primer was designed on the basis of the archaeal  
142 alignment (Fig. 2) by searching for conserved areas within the nanoarchaeal sequences  
143 which were not conserved across the other 3 archaeal sub-divisions. The region chosen  
144 for primer N961R had 100% complementarity to the Chinese and New Zealand  
145 nanoarchaeal phylotypes, as well as to the published environmental DNA sequences,  
146 except for *N. equitans* which had 2 mismatches. Primer N961R was, therefore, designed  
147 containing degeneracies at those two positions.

148

1  
2  
3  
4 149 The new primer N961R was used in the laboratory in conjunction with universal archaeal  
5  
6 150 primer A571F, in comparison with universal archaeal primer pair A571F/UA1204R. Both  
7  
8 151 primer sets were tested against crenarchaeal (*Solfolobus solfataricus*), euryarchaeal  
9  
10 152 (*Thermococcus litoralis*), nanoarchaeal DNA (clone TC11B6) and a hydrothermal  
11  
12 153 sediment sample (TA101) from the Tatio volcanic region in Chile.

13  
14  
15 154 The archaeal universal primers amplified 16S rRNA gene from all four samples (Fig. 3).  
16  
17 155 The “nanoarchaeal-specific” primer set was used successfully to amplify DNA from  
18  
19 156 nanoarchaeal DNA and the Chilean hydrothermal sediment sample and did not amplify  
20  
21 157 crenarchaeal or euryarchaeal DNA (Fig. 3). The PCR product amplified with the  
22  
23 158 nanoarchaeal-specific primers from the Chilean sediment sample was cloned and the  
24  
25 159 resultant recombinants were sequenced using M13 primers. Sequence analysis revealed  
26  
27 160 high identity to nanoarchaeal phylotypes.  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



### 162 **Nanoarchaeal Diversity in the Chilean Hydrothermal Samples**

163 The nanoarchaeal-specific primer set was used to amplify 16S rRNA genes from Chilean  
164 hydrothermal samples. Reamplification of inserts from 22 clones from the TA22 library,  
165 43 clones from the TA101 library and 48 clones from the TA62 library was achieved  
166 using the same primer set. These inserts were digested with *RsaI*, *HaeIII* and *AluI*. A  
167 single ARDRA pattern was identified for all samples, with the exception of sample  
168 TA101 which resulted in a second pattern with *AluI*. The two nanoarchaeal sequences  
169 had high identity (95-99%) to the Chinese nanoarchaeal sequences (Fig. 4).

170

1  
2  
3 171 The exact phylogenetic position of these Chilean phylotypes is not precise due to the  
4  
5 172 short sequence used to construct the phylogenetic tree. Nanoarchaeal-specific primers  
6  
7  
8 173 should be designed to allow amplification of a larger portion of the 16S rRNA gene.  
9  
10 174 However, due to the lack of nanoarchaeal sequences before position 571 and after  
11  
12 175 position 1204, it is difficult to design primers which would amplify the full nanoarchaeal  
13  
14 176 diversity.  
15  
16  
17  
18 177

19  
20  
21 178 The placement of the Nanoarchaeota within the Archaea is less clear, its position  
22  
23 179 depending on the algorithm and outgroups chosen (Brochier *et al.*, 2005). It was hoped  
24  
25 180 that an accurate branching position could be determined with the discovery of further  
26  
27 181 members of the clade (Huber *et al.*, 2002). Unfortunately however, despite this study  
28  
29 182 increasing the number of known nanoarchaeal phylotypes, it has not provided full  
30  
31 183 resolution of the Nanoarchaeota branching position.  
32  
33  
34  
35  
36 184

37  
38 185 Previous studies (Hohn *et al.*, 2002; Huber *et al.*, 2002) have shown that nanoarchaeotes  
39  
40 186 are found in marine and continental hydrothermal systems of pH 5.5 – 6.0. In this study  
41  
42 187 the sediments from which nanoarchaeote sequences were amplified had a pH range of 5.0  
43  
44 188 to 8.0, with temperatures between 80 and 96 °C.  
45  
46  
47  
48 189

49  
50 190 As previously noted (Hohn *et al.*, 2002), there is a wide geographical distribution of  
51  
52 191 members of the Nanoarchaeota, with new members being found in Asia, Oceania and  
53  
54 192 South America. The lack of a complete sequence for these novel phylotypes makes it  
55  
56 193 impossible to deduce whether they can be classified as different species, and possibly  
57  
58  
59  
60

1  
2  
3 194 genera, due to their divergence within the available sequence. It is possible to speculate  
4  
5 195 that if a similar divergence was found in the rest of the 16S rRNA gene, these phylotypes  
6  
7  
8 196 would definitely be classified as different taxa. It is interesting to note that the Chinese  
9  
10 197 and Chilean samples formed a phylogenetic cluster, separate from the New Zealand  
11  
12 198 phylotypes, and all these in turn are related but different to the published sequences.  
13  
14  
15  
16

199

17  
18 200 The design of universal archaeal primers A571F and UA1204R, as well as nanoarchaeal-  
19  
20 201 specific primer N961R, has been useful in identifying novel phyla and significantly  
21  
22 202 increase our knowledge of the sub-division Nanoarchaeota with the identification of  
23  
24 203 novel nanoarchaeal phylotypes that could possibly constitute several new species.  
25  
26  
27  
28

204

### 205 **Acknowledgements**



29  
30  
31  
32 206 Chinese sediment samples were obtained under the auspices of the joint Chinese-EU-SA  
33  
34 207 MGATech project (QLRT-2001-01972). Chilean sediment samples were obtained by Dr  
35  
36 208 J. Blamey. GCB was supported by the Royal Society/NRF program. NG and AC were  
37  
38 209 supported by the NRF HBU-RDP program. DNA sequencing was conducted by Di  
39  
40 210 James at the University of Cape Town Sequencing Facility and Inqaba Biotech, South  
41  
42 211 Africa. We gratefully acknowledge the assistance of Joseph Lako and William Stafford.  
43  
44  
45  
46  
47

212

### 213 **References**

- 50  
51  
52 214 1. Baker GC, Cowan DA (2004) 16S rDNA primers and the unbiased assessment of  
53  
54 215 thermophile diversity. *Biochem Soc Trans* 32:218-221.  
55  
56  
57  
58  
59  
60

- 1  
2  
3 216 2. Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-  
4  
5 217 specific 16S primers. *J Microbiol Methods* 55:541-555.  
6  
7  
8  
9 218 3. Barns SM, Delwiche CF, Palmer JD, Pace NR (1996) Perspectives on archaeal  
10  
11 219 diversity, thermophily and monophyly from environmental rRNA sequences. *Proc*  
12  
13 220 *Natl Acad Sci USA* 93[17], 9188-9193.  
14  
15  
16  
17 221 4. Brochier C, Gribaldo S, Zivanovic Y, Confalonieri F, Forterre P (2005)  
18  
19 222 Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving  
20  
21 223 euryarchaeal lineage related to Thermococcales? *Genome Biol* 6: R42.  
22  
23  
24  
25 224 5. Hohn MJ, Hedlund BP, Huber H (2002) Detection of 16S rDNA sequences  
26  
27 225 representing the novel phylum "Nanoarchaeota": Indication for a wide distribution  
28  
29 226 in high temperature biotopes. *Syst Appl Microbiol* 25:551-554.  
30  
31  
32  
33 227 6. Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO (2002) A new  
34  
35 228 phylum of Archaea represented by a nano-sized hyperthermophilic symbiont.  
36  
37 229 *Nature* 417: 63-67.  
38  
39  
40  
41 230 7. Huber H, Hohn MJ, Stetter KO, Rachel R (2003) The phylum Nanoarchaeota:  
42  
43 231 Present knowledge and future perspectives of a unique form of life. *Res Microbiol*  
44  
45 232 154:165-171.  
46  
47  
48  
49 233 8. McCliment EA, Voglesonger KM, O'Day Pa, Dunn EE, Holloway JR, Cary SC  
50  
51 234 (2006) Colonization of nascent, deep-sea hydrothermal vents by a novel Archaeal  
52  
53 235 and Nanoarchaeal assemblage. *Environ Microbiol* 8: 114-125.  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 236 9. Stach JEM, Bathe S, Clapp JP, Burns RG (2001) PCR-SSCP comparison of 16S  
4  
5  
6 237 rDNA sequence diversity in soil obtained using different isolation and purification  
7  
8 238 methods. FEMS Microbiol Ecol 36: 139-151.
- 9  
10  
11 239 10. Van de Peer Y, Chapelle S, De Wachter R (1996) A quantitative map of  
12  
13 240 nucleotide substitution rates in bacterial rRNA. Nucleic Acids Res 24: 3381-3391.
- 14  
15  
16  
17 241 11. Waters E, Hohn MJ, Ahel I, Graham DF, Adams MD, Barnstead M, *et al.* (2003)  
18  
19 242 The genome of Nanoarchaeum equitans: Insights into early archaeal evolution and  
20  
21 243 derived parasitism. Proc Natl Acad Sci USA 100:12984-12988.
- 22  
23  
24  
25  
26 244



Peer Review

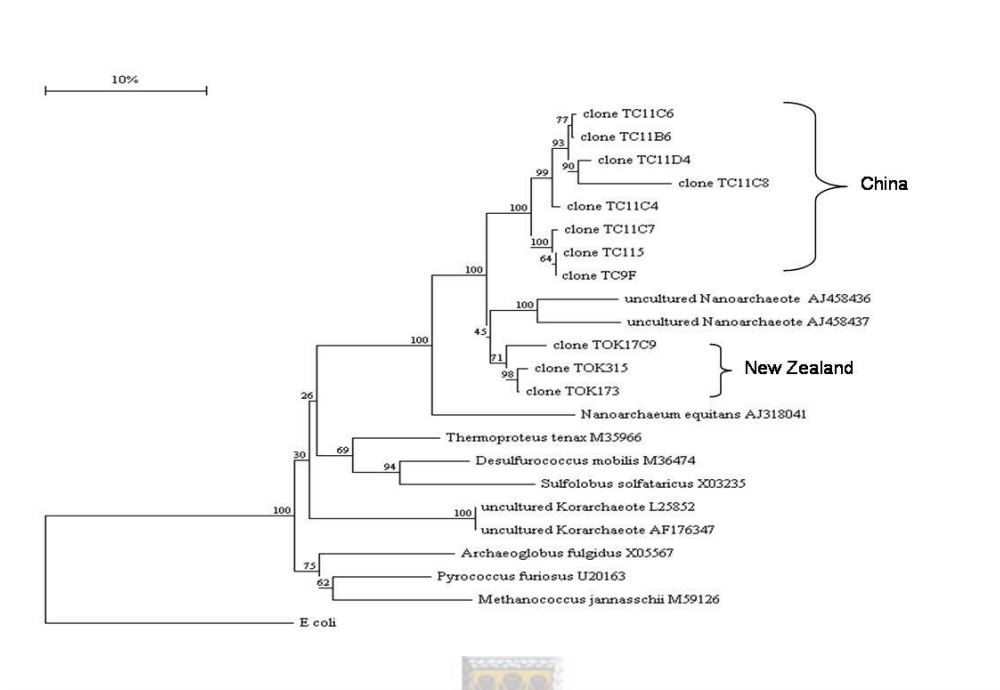


1  
2  
3  
4 244 **Figure 1.** Neighbour-joining tree based on Jukes-Cantor distances of a 640 bp alignment  
5  
6 245 of 16S rRNA gene sequences. The Chinese and New Zealand clones were amplified with  
7  
8 246 primers A571F and UA1204R.  
9

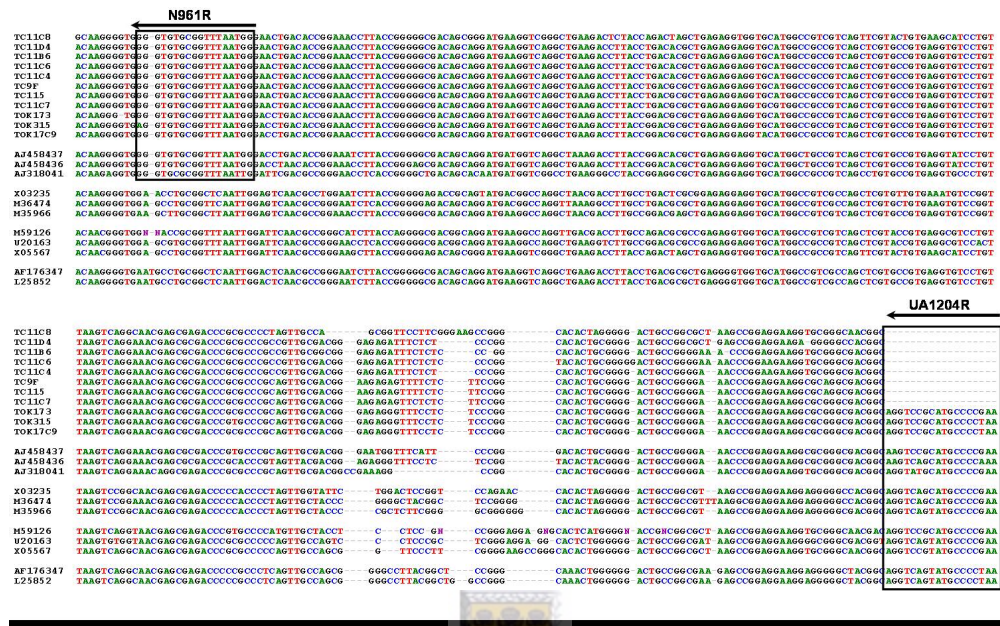
10  
11 247  
12  
13  
14  
15 248 **Figure 2.** Alignment of representative members from the four archaeal sub-divisions.  
16  
17 249 Clustal alignment from base 933 to 1204 of archaeal 16S rRNA gene sequences. Row 1:  
18  
19 250 Chinese and New Zealand Nanoarchaeotes; Row 2: published Nanoarchaeotes; Row 3:  
20  
21 251 Crenarchaeotes; Row 4: Euryarchaeotes; Row 5: Korarchaeotes. The positions of the  
22  
23 252 universal archaeal primer UA1204R<sup>2</sup> and the newly designed Nanoarchaeal-specific  
24  
25 253 primer N961R are marked by a black arrow.  
26  
27  
28

29  
30 254  
31  
32 255 **Figure 3.** TBE 1% agarose gel containing PCR products using primers A571F/UA1204R  
33  
34 256 (Lanes 1-4) and A571F/N961R (Lanes 5-8) from crenarchaeal (Lanes 1 & 5),  
35  
36 257 euryarchaeal (Lanes 2 & 6), nanoarchaeal (Lanes 3 & 7) and hydrothermal sediment  
37  
38 258 (Lanes 4 & 8) DNA templates. *Pst*I-digested  $\lambda$  DNA was used as marker.  
39  
40  
41

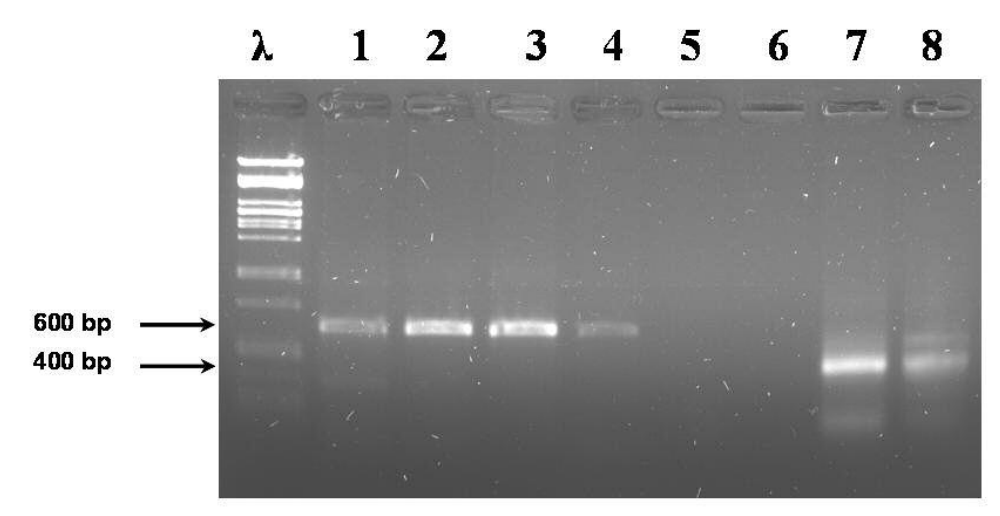
42 259  
43  
44  
45 260 **Figure 4.** Neighbour-joining tree based on Jukes-Cantor distances of a 400 bp alignment  
46  
47 261 of 16S rRNA gene sequences. Chilean clones TA62NA4 and TA101Na3 were amplified  
48  
49 262 with primers A571F and N961R, while the other Chinese (TC11 and TC9) and New  
50  
51 263 Zealand (TOK) clones were amplified with universal archaeal primers A571F and  
52  
53 264 UA1204R. Chinese and Chilean Nanoarchaeotes form a cluster separate from the New  
54  
55 265 Zealand sequences and the Yellowstone and Kamchatka phylotypes.  
56  
57  
58  
59  
60



**Figure 1. Neighbour-joining tree based on Jukes-Cantor distances of a 640 bp alignment of 16S rRNA gene sequences. The Chinese and New Zealand clones were amplified with primers A571F and UA1204R.**



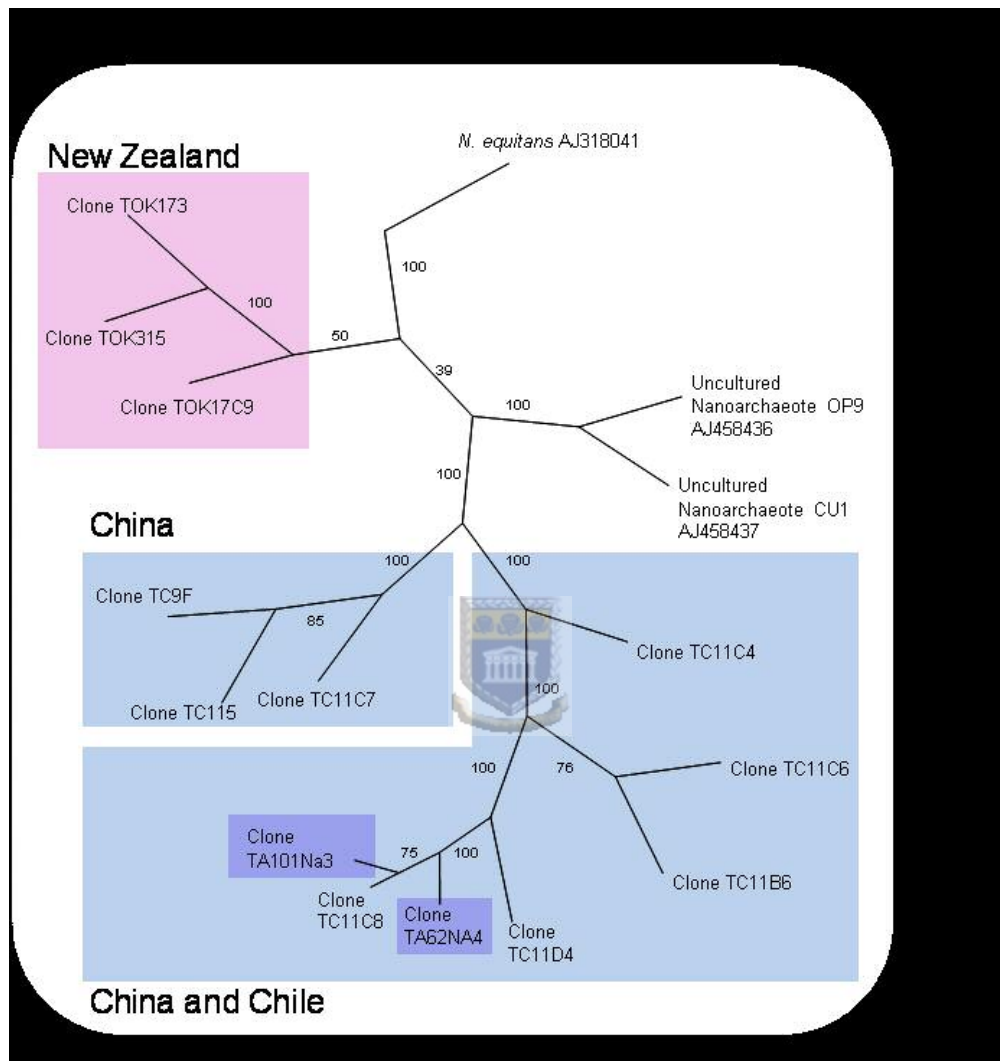
**Figure 2. Alignment of representative members from the four archaeal sub-divisions. Clustal alignment from base 933 to 1204 of archaeal 16S rRNA gene sequences. Row 1: Chinese and New Zealand Nanoarchaeotes; Row 2: published Nanoarchaeotes; Row 3: Crenarchaeotes; Row 4: Euryarchaeotes; Row 5: Korarchaeotes. The positions of the universal archaeal primer UA1204R2 and the newly designed Nanoarchaeal-specific primer N961R are marked by a black arrow.**



27  
28 **Figure 3.** TBE 1% agarose gel containing PCR products using primers A571F/UA1204R  
29 (Lanes 1-4) and A571F/N961R (Lanes 5-8) from crenarchaeal (Lanes 1 & 5),  
30 euryarchaeal (Lanes 2 & 6), nanoarchaeal (Lanes 3 & 7) and hydrothermal sediment  
31 (Lanes 4 & 8) DNA templates. PstI-digested  $\lambda$  DNA was used as marker.

32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Review



**Figure 4. Neighbour-joining tree based on Jukes-Cantor distances of a 400 bp alignment of 16S rRNA gene sequences. Chilean clones TA62NA4 and TA101Na3 were amplified with primers A571F and N961R, while the other Chinese (TC11 and TC9) and New Zealand (TOK) clones were amplified with universal archaeal primers A571F and UA1204R. Chinese and Chilean Nanoarchaeotes form a cluster separate from the New Zealand sequences and the Yellowstone and Kamchatka phylotypes.**