16S rDNA primers and the unbiased assessment of thermophile diversity

G.C. Baker and D.A. Cowan

Institute for Microbial Biotechnology and Metagenomics, University of the Western Cape, Private Bag X17, Bellville 7535, Cape Town, South Africa

Abstract

Our understanding of thermophile diversity is based predominantly on PCR studies of community DNA. 'Universal' and domain-specific rRNA gene PCR primers have historically been used for the assessment of microbial diversity without adequate regard to the degree of specificity of primer pairs to different prokaryotic groups. In a reassessment of the published primers commonly used for 'universal' and archaeal 16 S rDNA sequence amplification we note that substantial variations in specificity exist. An unconsidered choice of primers may therefore lead to significant bias in determination of microbial community composition. In particular, Archaea-specific primer sequences typically lack specificity for the Korarchaeota and Nanoarchaea and are often biased towards certain clades. New primer pairs specifically designed for 'universal' archaeal 16 S rDNA sequence amplification, with homology to all four archaeal groups, have been designed. Here we present the application of these new primers for preparation of 16 S libraries from thermophile communities.

Keywords: Bacteria, Archaea, Korarchaeota, Nanoarchaeota, primer, rRNA, 16 S.

Introduction

Before the advent of PCR-based technologies our knowledge of microbial community diversity was restricted to those taxa that were culturable *in vitro*. PCR of 16 S rDNA from environmental samples and subsequent sequence analysis has facilitated a greater understanding of microbial diversity [1], but many important taxa will continue to be missed in our inventories if care is not taken to amplify DNA with adequately universal primers. In the 1980s, sets of primers were designed on the basis of nucleotides conserved in the organisms that had been sequenced at that time (e.g. [2]). Since the 1980s, thousands of additional sequences have been added to the databases, and with this increase in sequence information, new taxonomic groups have been discovered [36]. In the domain Archaea, two new sub-divisions have been proposed: the Korarchaeota [3] and the Nanoarchaeota [4]. Both of these taxa comprise hyperthermophilic organisms, which through their study may provide new information on the nature of thermophily and a greater understanding of early evolution [7]. The rDNA sequences of these taxa differ from Crenarchaeota and Euryarchaeota, and in the case of Nanoarchaeota there is mismatch at important priming sites. Specific primers have been designed for Nanoarchaeota [4] and Korarchaeota [8] that can be used when expressly searching for new members of these taxa. However, when conducting routine phylogenetic analysis of samples from thermal environments, use of a broad-based primer pair that shows equal complementarity to all archaeal taxa would be advantageous. In the course of selecting a set of primers that would amplify all archaeal groups, without bias towards particular taxa, we reviewed a range of published primers and examined

their specificities [9]. Here we present a summary of our analysis of archaeal 16 S rDNA primers and provide a graphic representation of nucleotide conservation in archaeal 16 SrDNAs.

Experimental

In silico

A 1300 bp ClustalW alignment was made of 16 S rRNA gene sequences that were representive of the major subdivisions within the archaeal domain. These included four Euryarchaeote (X05567, D50849, M59126, U20163), four Crenarchaeote (X03235, M35966, M36474, U51469), one Nanoarchaeote (AJ318041) and two Korarchaeote (AF176347, L25852) sequences from the NCBI database. The alignment was tabulated and annotated with published primer sequences reported to be complementary to 16 S rRNA genes [9]. Bases that were conserved in all 11 sequences were highlighted on a secondary structure model of the Euryarchaeote *Methanococcus jannaschii,* downloaded from the European Ribosomal RNA database (http://oberon.fvms.ugent.be:8080/rRNA/index.html). The model was annotated with archaeal priming sites (primer sequences are shown in Table 1). Four new Archaea-specific primers were designed from the alignment, with 100% complementarity to all archaeal sub-divisions, but with mismatch to bacterial and eukaryote sequences. A selection of 47 published archaeal primers and the four new primers were submitted to the Probe Match facility of the Ribosomal RNA Database Project (http://rdp.cme.msu.edu). DNA sequences in the RDP database with 100% complementarity to the primers were noted and tabulated taxonomically. For each primer a score was given for the number of 100% matches in each of the eight categories of Archaea, in the Bacteria, and the eukaryotes [9].

Table 1 16 S rRNA primer sequences

All primers have been re-named using the following convention. First letters, specificity as judged by RDP Probe Match Analysis and manual analysis of the archaeal alignment (A, Archaea; UA, universal Archaea; E, Eubacteria; EK, Euryarchaeote; TC, thermophilic Crenarchaeote; M, Methanogen; Mb, Methanogen-biased; N, Nanoarchaeote; b, biased); number, position (*E. coli* numbering) of 5_ end of primer in gene; F/R, forward or reverse.

Code	Sequence (5'–3')	Reference	Code	Sequence (5'–3')	Reference
A1F	ATTCCGGTTGATCCTGC	[12]	Ab127R	CCACGTGTTACTSAGC	[21]
A2F	TTCCGGTTGATCCYGCCGGA	[10,13,14]	A348R	CCCCGTAGGGCCYGG	[22]
EK4F	CTGGTTGATCCTGCCAG	[15]	EK510R	CTTGCCCRGCCCTT	[21]
A109F	ACKGCTCAGTAACACGT	[16]	TC518R	ACACCAGRCTTGCCCCCCGCTT	[22]
A333F	TCCAGGCCCTACGGG	[10]	U529R	ACCGCGGCKGCTGGC	[21]
U341F	CCTACGGGRSGCAGCAG	[17]	U534R	GWATTACCGCGGCKGCTG	[21]
A344F	ACGGGGTGCAGCAGGCGCGA	[18]	M704R	TTACAGGATTTCACT	[21]
U515F	GTGCCAGCMGCCGCGGTAA	[10,19]	Ab909R	TTTCAGYCTTGCGRCCGTAC	[8]
U519F	CAGCMGCCGCGGTAATWC	[20]	Ab927R	CCCGCCAATTCCTTTAAGTTTC	[14]
UA571F	GCYTAAAGSRICCGTAGC	[9]	A976R	YCCGGCGTTGAMTCCAATT	[10]
UA751F	CCGACGGTGAGRGRYGAA	[9]	A1115R	GGGTCTCGCTCGTTG	[10]
Ab779F	GCRAASSGGATTAGATACCC	[8]	EKb1242R	CCATTGTAGCSCGCGTG	[21]
Eb787F	ATTAGATACCCTGGTA	[21]	UA1204R	TTMGGGGCATRCIKACCT	[9]
Ab787F	ATTAGATACCCGGGTA	[21]	UA1406R	ACGGGCGGTGWGTRCAA	[9]
A1040F	GAGAGGWGGTGCATGGCC	[10]	N1406R	ACGGGCGGTGAGTGCAA	[4]
A1098F	GGCAACGAGCGMGACCC	[10]	U1406R	GACGGGCGGTGTGTRCA	[10,17]
Mb1225F	ACACGCGTGCTACAAT	[21]			

In vitro

New primers (UA571F/UA1204R and UA751F/UA1406R) were tested in the laboratory in comparison with a published primer pair (A2F/U1510R), which is frequently used for amplification of archaeal small-subunit rDNA from environmental samples (e.g. [5,10,11]). The primers were used to amplify DNA from type strain Archaea and environmental samples and were tested on *Escherichia coli* to check that they did not amplify non-archaeal prokaryote DNA [9].

Results

A model of the 16 S rRNA of *M. jannaschii* has been generated (Figure 1) that indicates conserved bases between the 11 archaeal 16 S rDNA sequences that were aligned. Although almost 50% of the nucleotides are conserved between all taxa, there are few regions where the conserved bases are adjacent over sufficiently long stretches for accurate primer design. In all priming regions there is some mismatch if base degeneracies are not incorporated into the primers. This model may thus be employed as a simple tool to establish where degeneracies are needed in order to have 100% primer-template complementarity for all archaeal taxa. Figure 2 demonstrates the specificities of a range of 16 S primers to sequences from 10 different taxa. Specificity is defined as 100% complementarity between the primers and at least one sequence in the RDP database of that taxon. None of the primers examined are truly universal and there are significant differences in specificity between similar primers. For example, primers A1F, A2F and EK4F all anneal at the 5' of the 16 S rRNA gene. They differ in sequence by less than 10% (see Table 1) and in the 3' position by between 2 and 4 bases (see Figure 1). The specificities of these three primers are, however, quite different. Primer A2Fa has the broadest specificity, with 100% identity to DNA sequences from six archaeal taxa, whereas EK4F only has 100% identity to Methanomic-robacteria and Eukaryote sequences. All of the published Archaea-specific primers examined either complement non-archaeal sequences or have mismatch to particular archaeal taxa. The newly designed primers (UA571F/UA1204R and UA751F/UA1406R) utilize regions of conservation common to all four Archaea and do not have 100% identity to any bacterial or eukaryotic sequences. At positions where there is mismatch, degenerate bases and inosine residues are incorporated.

Laboratory assessment of primers

The new primer pairs effectively amplify DNA from *Sulfolobus, Thermococcus* and *Pyrococcus* type strains [9] and have been used to amplify DNA from environmental DNA samples from hotsprings in New Zealand [9] and China (results not shown). However, a library constructed from primer pair UA751F/UA1406R contained a large number of chimaeric artifacts (results not shown).

Discussion

Analysis of the model shown in Figure 1 clearly demonstrates variability of nucleotides between archaeal taxa at published priming sites. Primer design is a compromise between primer-template complementarity and other primer attributes, such as melting temperature, G:C ratio and secondary structure.

Figure 1 Secondary structure map of *M.jannaschii* (http://oberon.fvms.ugent.be:8080/rRNA/index.html) annotated with Archaea-specific priming sites

Nucleotides (105-1406; *E. coli* numbering) conserved between representatives of the Euryarchaeota, Crenarchaeota, Korarchaeota and Nanoarchaeota are highlighted.

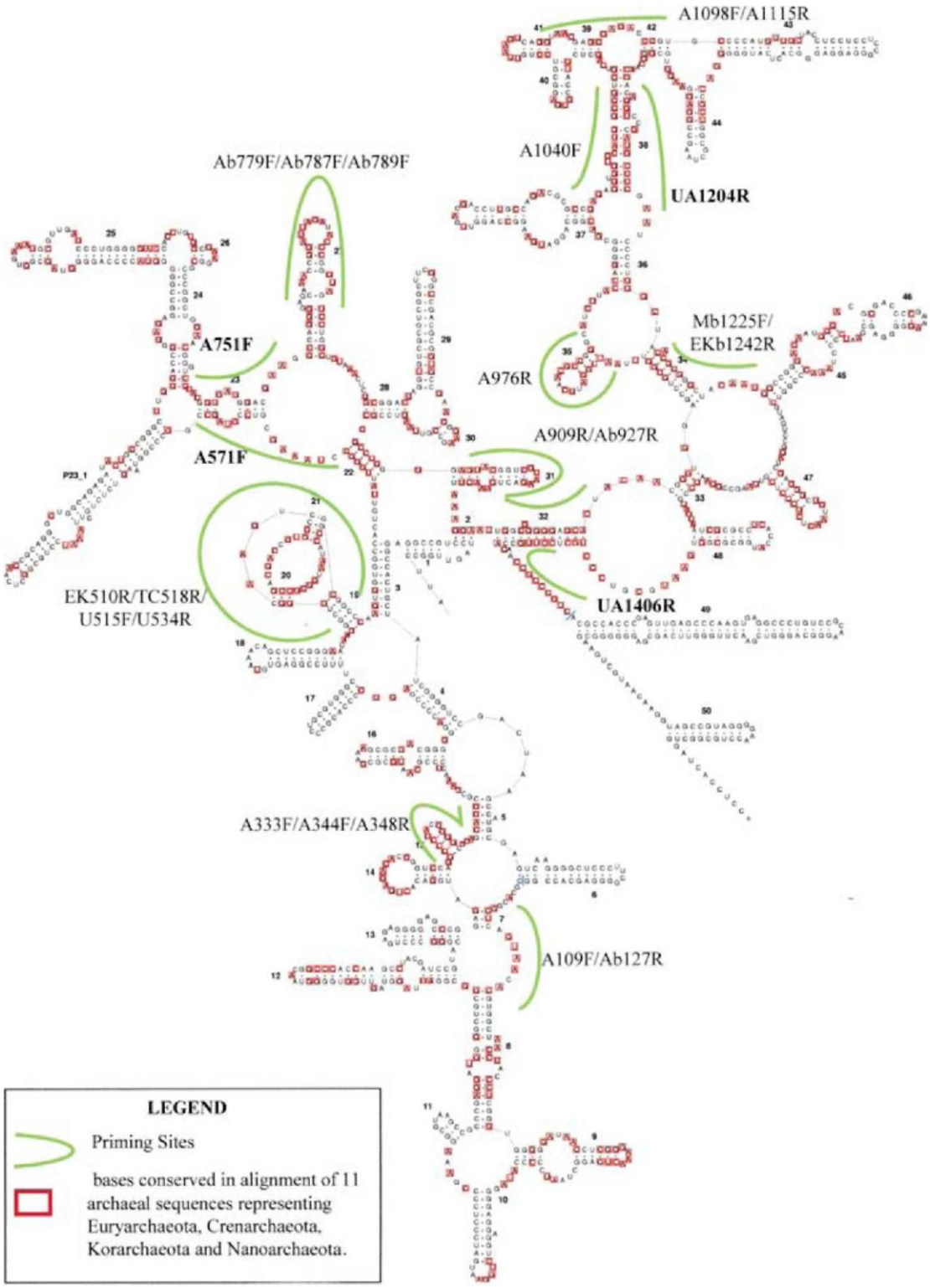
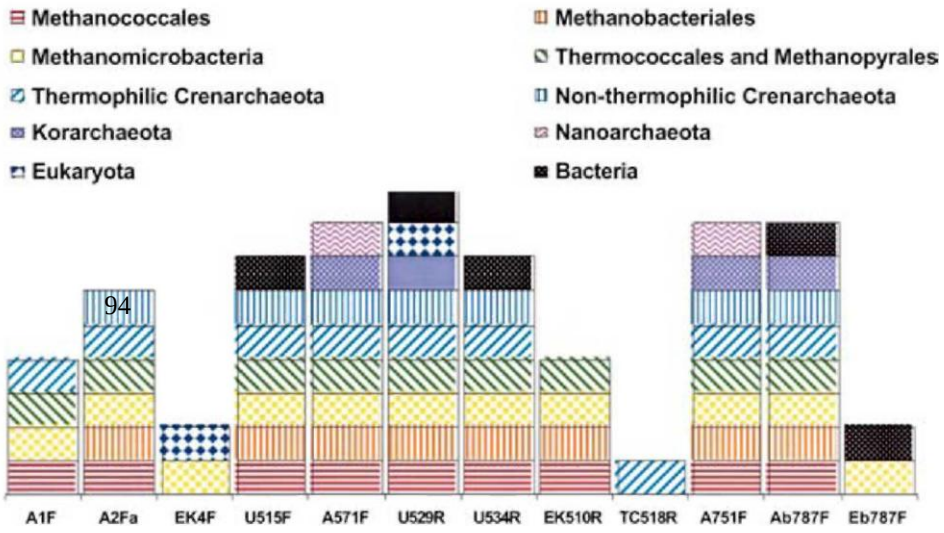


Figure 2 | Graph showing primer specificities to taxonomic groups



PCR using primer-template homology as low as 70% has been achieved [23]. However, in total community DNA samples, differential sequence complementarity to primers between taxa will lead to a significant bias in the amplification products. Incorporation of multiple bases at degenerate positions and the use of inosine residues have been used effectively to provide 'universal specificity', but excess use of these bases has been reported to have biased template-to-product ratios [24] and led to amplification of non-target groups [25].

In the literature there are many variations of 'Archaea-specific' primers with identical or similar annealing sites. The specificities of these primers differ considerably. An 'ideal' primer pair should have 100% homology to representatives from all archaeal taxa and have substantial mismatch to bacterial and eukaryotic sequences. None of the published primers analysed in our study [9] possessed these attributes. Therefore, careful modification of currently used primers, based on an up-to-date model of conserved nucleotides, or use of the primers described in this paper, is recommended in order to access a greater archaeal diversity, without the need for amplification with multiple primer sets.

We thank the Claude Harris Leon Foundation and the South African National Research Foundation for financial assistance, and Pia Wittwer for her valuable advice on the manuscript.

References

1 Theron, J. and Cloete, T.E. (2000) *Crit. Rev. Microbiol.* **26,** 37-57

2 Elwood, H.J., Olsen, G.J. and Sogin, M.L. (1985). *Mol. Biol. Evol.* **2,** 399

3 Barns, S.M., Delwiche, C.F., Palmer, J.D. and Pace, N.R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93,** 9188-9193

4 Huber, H., Hohn, M.J., Rachel, R., Fuchs, T., Wimmer, V.C. and Stetter, K.O. (2002) *Nature* (London) **417,** 63-67

5 Hugenholtz, P., Pitulle, C., Hershberger, K.L. and Pace, N.R. (1998) *J. Bacteriol.* **180,** 366-376

6 Nielsen, A.T., Liu, W.-T., Filipe, C., Grady, L., Molin, S. and Stahl, D.A. (1999) *Appl. Environ. Microbiol.* **65,** 1251-1258
7 Huber, H., Hohn, M.J., Stetter, K.O. and Rachel, R. (2003) *Microbiology* **154,** 165-171
8 Brunk, C.F. and Eis, N. (1998) *Appl. Environ. Microbiol.* **64,** 5064-5066
9 Baker, G.C., Smith, J.J. and Cowan, D.A. (2003) *J. Microbiol. Methods* **55,** 541-555
10 Reysenbach, A.-L. and Pace, N.R. (1995) in *Archaea: A Laboratory Manual - Thermophiles* (Robb, F.T. and Place, A.R., eds.), pp. 101-107, Cold Spring Harbour Press, Cold Spring Harbor
11 Liopez-Garcia, P., Gaill, P. and Moreira, D. (2002) *Env. Microbiol.* **4,** 204-215
12 Tajima, K., Nagamine, T., Matsui, H., Nakamura, M. and Aminov, R.I. (2001) *FEMS Microbiol. Lett.* **200,** 67-72
13 Martinez-Murcia, A.J., Acinas, S.G. and Rodriguez-Valera, F. (1995) *FEMS Microb. Ecol.* **17,** 247-256
14 Jurgens, G., Glockner, F.O., Amman, R., Saana, A., Mon tonen, L., Likolammi, M. and Munster, U. (2000) *FEMS Microb. Ecol.* **34,** 45-56
15 Vetriani, C., Jannasch, H.W., McGregor, B.J., Stahl, D.A. and Reysenbach, A.-L. (1999) *Appl. Env. Microbiol.* **65,** 4375-4384
16 Whitehead, T.R. and Cotta, M.A. (1999) *FEMS Microbiol. Lett.* **179,** 223-226
17 Hansen, M.C., Tolker-Neilson, T., Givskov, M. and Molin, S. (1998) *FEMS Microb. Ecol.* **26,** 141-149
18 Casamayor, E.O., Massana, R., Benlloch, S., Ovreas, L., Diez, B., Goddard, V., Gasol, J.M., Joint, I., Rodriguez-Valera, F. and Predros-Alio, C. (2002) *Env. Microbiol.* **4,** 338-348
19 Reysenbach, A.-L., Giver, L.J., Wickham, G.S. and Pace, N.R. (1992) *Appl. Environ. Microbiol.* **58,** 3417-3418
20 Suzuki, M.T. and Giovanni, S.J. (1996) *Appl. Environ. Microbiol.* **62,** 625-630
21 DasSarma, S. and Fleischmann, E.F. (1995) in *Archaea: A Laboratory Manual - Halophiles* (DasSarma, S. and Fleischmann, E.M., eds.), pp. 269-272, Cold Spring Harbor Press, Cold Spring Harbor
22 Barns, S.M., Fundyga, R.E., Jeffries, M.W. and Pace, N.R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91,** 1609-1613
23 Stern, C.D. and Holland, P.W.H. (1993) in *Essential Developmental Biology*, p. 324, Oxford University Press, Oxford
24 Polz, M.F. and Cavanaugh, C.M. (1998) *Appl. Environ. Microbiol.* **64,** 3724-3730
25 Watanabe, K., Kodama, Y. and Harayama, S. (2001) *J. Microb. Method.* **44,** 253-262