Review and re-analysis of domain-specific 16S primers (Review article)

G.C. Baker, J.J. Smith, D.A. Cowan*

Department of Biotechnology, University of the Western Cape, Bellville 7335, Cape Town, South Africa

Abstract

The Polymerase Chain Reaction (PCR) has facilitated the detection of unculturable microorganisms in virtually any environmental source and has thus been used extensively in the assessment of environmental microbial diversity. This technique relies on the assumption that the gene sequences present in the environment are complementary to the "universal" primers used in their amplification. The recent discovery of new taxa with 16S rDNA sequences not complementary to standard universal primers suggests that current 16S rDNA libraries are not representative of true prokaryotic biodiversity. Here we re-assess the specificity of commonly used 16S rRNA gene primers and present these data in tabular form designed as a tool to aid simple analysis, selection and implementation. In addition, we present two new primer pairs specifically designed for effective 'universal' Archaeal 16S rDNA sequence amplification. These primers are found to amplify sequences from Crenarchaeote and Euryarchaeote type strains and environmental DNA.

Keywords: Bacteria; Archaea; Korarchaeota; Nanoarchaeota; 16S; rRNA; Primer

1. Introduction

16S rRNA sequence analysis has been used to clarify the taxonomic affinities of a wide range of taxa (e.g., Baker et al., 1999; Mclnnery et al., 1995) and as a powerful tool for assessing the genetic diversity of environmental samples (e.g., Baker et al., 2001; Groflkopf et al., 1998; van Waasbergen et al., 2000; Whitehead and Cotta, 1999). PCR amplification of 16S rDNA from organisms which are as yet unculturable has provided an invaluable insight into our understanding of community structure, especially of communities inhabiting extreme environments, where growth conditions may be difficult to mimic in the laboratory. Over the past 25 years, a large number of primer sequences for amplification and sequencing of ssu rRNA genes have been published (e.g., DasSarma and Fleischmann, 1995; Elwood et al., 1985; Kolganova et al., 2002; Wata-nabe et al., 2001). Some of these primers have been designed as taxa specific, whilst others have been designed to amplify all prokaryotic ssu-rRNA genes and are referred to as 'universal'. As the database of 16S rRNA gene sequences has grown, new taxo-nomic groups have been discovered (Barns et al., 1994; Hugenholtz et al., 1998; Mclnnery et al., 1995; Nielsen et al., 1999). In particular, two novel sub-divisions of the domain Archaea, the Korarch-aeota (Barns et al., 1994, 1996) and Nanoarchaeota (Huber et al., 2002), have been discovered in the past decade.

These sub-divisions are based mainly on environmental DNA sequences from thermophilic ecotypes (Barns et al., 1994; Marteinsson et al., 2001; Hohn et al., 2002). Burggraf et al. (1997) have successfully cultivated a Korarchaeote in a mixed laboratory culture at 85 °C. The cells, identified by fluorescence hybridisation, are rods of 5 — 10 Am in length. The Nanoarchaea exhibit uncharacteristically small cell size and were first documented as exosymbionts of the Crenarchaeote, Ignicoccus at 70—98 °C (Huber et al., 2002). Both the Korarchaeota and Nanoarch-aeota are positioned near the base of the Archaeal branch of the universal phylogenetic tree (Barns et al., 1994; Kim et al., 2000; Huber et al., 2003), suggesting that these are relatively archaic lineages. Due to culturing constraints and the general perception that Archaea are limited to extreme environments, it is likely that representatives from these groups may well have been overlooked elsewhere. It is notable that phylotypic signals identified as low temperature Crenarchaeotes are almost ubiquitous in terrestrial and marine environments (DeLong et al., 1994; Jurgens et al., 1997; MacGregor et al., 1997; Schleper et al., 1998) but no examples of this group have yet been cultured.

Primers designed to be complementary to the conserved regions of the groups present in the original universal phylogenetic tree are not necessarily complementary to all those that exist in the database today. In fact, it has been demonstrated that *Nanoarchaeum* does not amplify using standard "universal archaeal primers" (Huber et al., 2002). New primers are thus required that are both universal and specific. Ideally, they must be specific to the domain in question, whilst complementary to sequences in all taxa within that domain. Lists of 16S rRNA primers have been published (e.g., DasSarma and Fleischmann, 1995; Rey-senbach and Pace, 1995) that are all uded to as universal or domain-specific, but little empirical evidence supporting these specificities is available.

The European Ribosomal RNA database (http:// www.silk.uia.ac.be) provides excellent secondary structure variability maps for 16S and 23S rRNAs and shows the superposition of primers on the 16S and 23S genes of *Escherichia coli*. Unfortunately, however, this resource is not readily available for the Archaea.

The objectives of this review are to summarise and critically review the information available on 16S rRNA primers, to provide an easy-to-use framework with which to examine the applicability of oligonu-cleotides on the basis of their specificities to various taxonomic groups and to offer up-to-date domain-specific primer sets. Previous papers (e.g., Wintzin-gerode et al., 1997) have discussed in general terms the pitfalls of PCR-based technologies, but here we focus on primer specificity and problems associated with primer-template mismatch.

2. Bacterial primer specificity

Regions of differing variability in the 16S rRNA variability map (Fig. 1) are colour-coded to facilitate the analysis of primer variability: totally conserved positions (v_i = 0), brown; highly conserved positions (v_i = 10" $^{0.925}$-10" $^{0.425}$), red; variable positions (v_i > 10+ $^{0.575}$), blue and nucleotides present in *E. coli* but absent in > 75% of other bacterial sequences, green; nucleotide positions with variability ranging from v_i = 10" $^{0.425}$ to v_i = 10+ $^{0.575}$, black. According to

this variability map, over 10% of bases in the 16S rRNA gene are totally conserved (within a sample of 500 bacterial sequences).

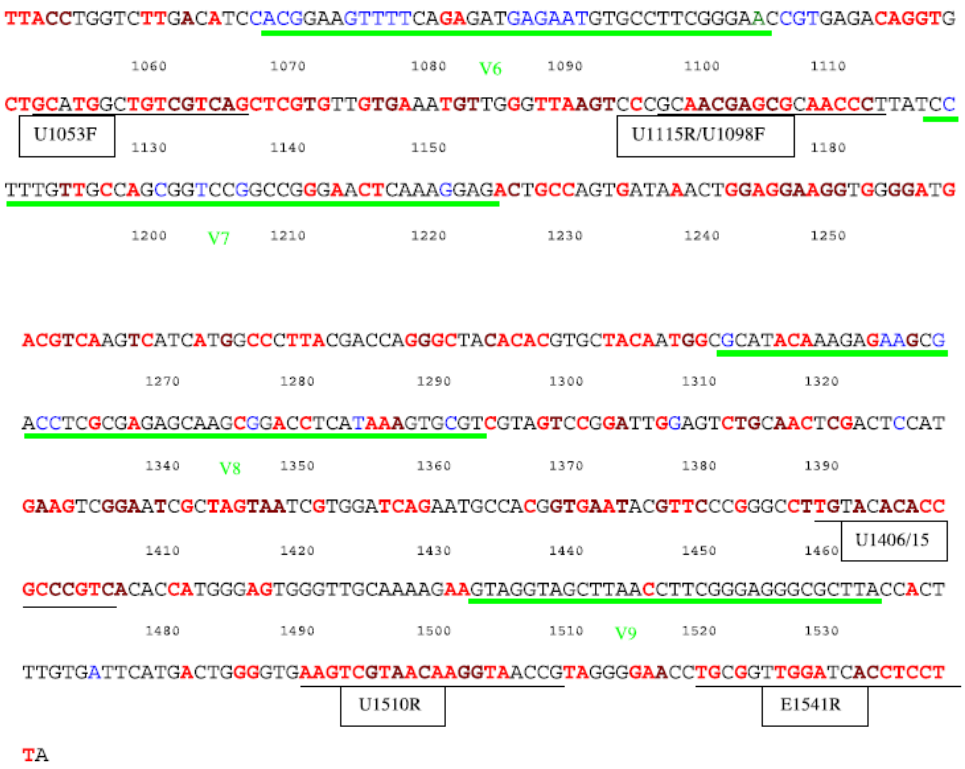

Table 1
Bacteria specific 16S rRNA gene sequence/PCR primers

Code	Sequence 5′–3′	N	Ta	B	A	E	Matches	Ref.
E8F	AGAGTTTGATCCTGGCTCAG	20	55	+++	+	–	2424	Reysenbach and Pace, 1995; Martinez-Murcia et al., 1995; Reysenbach et al., 1994
E9F	GAGTTTGATCCTGGCTCAG	19	53	+++	+	++	2741	McInnery et al., 1995; Hansen et al., 1998; Farelly et al., 1995
E334F	CCAGACTCCTACGGGAGGCAGC	21	65	+++	–	–	13,172	Rudi et al., 1997
E341F	CCTACGGGIGGCIGCA	16	51	+++	+	–	16,685	Watanabe et al., 2001
E786F	GATTAGATACCCTGGTAG	18	47	+++	+	–	12,616	Coloqhoun, 1997
E533R	TIACCGIIICTICTGGCAC	19	56	+++	+	++	18,724	Watanabe et al., 2001
E926R	CCGICIATTIITTTIAGTTT	20	50	+++	+ ++	++	19,950	Watanabe et al., 2001
E939R	CTTGTGCGGGCCCCCGTCAATTC	23	71	+++	–	–	8620	Rudi et al., 1997
E1115R	AGGGTTGCGCTCGTTG	16	47	+++	–	–	9052	Reysenbach and Pace, 1995
E1541R	AAGGAGGTGATCCANCCRCA	20	57	+++	+	–	1355	Suzuki and Giovanni, 1996

Primers described as bacteria-specific were submitted to the 'check probe' facility of the Ribosomal Database Project (http://www.rdp.cme.msu.edu/) to check for bacterial specificity. The total number of matches are given, and matches for Bacteria (B), Archaea (A) and Eukarya (E) are represented as follows: (–, no matches; +, <25 matches; ++, 25–100 matches; +++, <100 matches). Primer numbering relates to *E. coli* position complementary to the 5′ end of the primer. Last letter denotes direction: Forward and Reverse. Ta=(4 × (G+C)+2 × (A+T)) – 5 M=A or C; R=A or G; W=A or T; S=C or G; Y=C or T; K=G or T; N=A or G or C or T.

3. Archaeal primer specificity

The Archaeal alignment shows substantial mismatch between primers that are designed to be generally Archaea-specific or universally prokaryotic and the published sequences of some Archaeal taxa, especially the Korarchaeotes and Nanoarchaeum. Several of the Archaea-specific primers (e.g., EK4F, M704) are homologous to one taxonomic group within the Archaea or have a Euryarchaeota or Cren-archaeota bias. Of the 51 primers described only 18 are 100% complementary to Korarchaeote sequences and only 11 are complementary to Nanoarchaeum (Table 2, see after References).

Table 3 shows examples of the specificity differences of some 'taxa-specific' primers. For instance A1F to EK4F are all designed to anneal to the same part of the gene, but differ in sequence and specificity. All of these primers, except N3F, are said to be Archaea-specific, but EK4F has a very high specificity for methanogens-specific sequences and does not match 100% with any other Archaeal group. A2Fa has a more general specificity to Archaeal sequences. Similarly, Eb787F and Ab787F, which are both described as universal primers, differ by only one base (5 nt from the 3' end), but have entirely different specificities. Of the primers described as universal, many have poor specificity for the Crenarchaeotes (529R, 1053F, 515F), some have poor Euryarchaeota specificity (519F, 534R) and others have poor Ar-chaea-specificity over-all (926R).

4. Primer specificity

Universal primer design is a compromise between universal complementarity and other primer attributes, such as primer—primer complementarity, annealing temperature and G/C ratio (McPherson et al., 1995). Ideal annealing temperatures and lack of self-complementation are often sacrificed in 16S rRNA gene primer design in order to obtain optimal specificity across whole domains.

In Tables 1 and 3, primer specificity has been judged on the basis of 100% complementarity of primer to template. In reality, in the case of amplification from pure cultures, 70% identity to target sequence is sufficient for successful annealing and amplification (Stern and Holland, 1993). However, poor complementarity of "universal" primers, especially at the 3 ' end, to particular taxa in environmental samples will lead to under-representation of these genotypes within 16S rRNA gene libraries.

Primers that include multiple nucleotides at degenerate positions have been used to provide "universal specificity", but such primers are reported to have biased template-to-product ratios (Polz and Cava-naugh, 1998) and excessive degeneracies can lead to amplification of non-target genes or domains. Wata-nabe et al. (2001) reviewed the use bacteria-specific primers containing inosine residues and concluded that these primers are useful in the detection of diverse environmental populations. In our analysis, primers containing inosine residues had broader specificity, but excessive use can lead to amplification of non-target groups. Practically, inosine residues are considerably more expensive than standard bases.

5. New universal Archaeal primers

The majority of Archaea-specific primers were designed prior to the identification of the Korarch-aeote and Nanoarchaeote sub-divisions and thus Ar-chaea-specific primers need to be reassessed in the light of the discovery of these taxonomic groups. Specific primers have been designed to be effective for Korarchaeote and Nanoarchaeote identification (Brunk and Eis, 1998; Huber et al., 2002), but for total Archaeal community analysis, a set of primers specific to all Archaea would be essential. On the basis of our 1300 bp Archaeal 16S rRNA gene sequence alignment, there are clearly regions of conservation common to all four sub-divisions ofArchaea that have not previously been utilised.

6. Primer design

New primers have been designed on the basis of the Archaeal alignment (Fig. 2, see after References) by manual searching for areas of sequence conservation across all four Archaeal sub-divisions. Given that no regions of sufficient length with 100% complementarity to all Archaeal sub-divisions exist in the alignment, primers were designed to maximise complementarity at the 3' end. Where mismatches occurred degenerate bases were added. Where there were three different bases in a specific position, an inosine residue was added. No more than 25% total degeneracy, and less than 10% inosine residues were allowed per primer. Primers were designed with annealing temperatures of between 50 and 60 °C

7. Laboratory assessment of new primers

New primers (A571F/UA1204R and A751F/ UA1406R) were assessed in the laboratory in comparison to a published primer pair (A2Fa/U1510R), which are frequently used for amplification of Archaea ssu rDNA from environmental samples (Hugenholtz et al., 1998; Reysenbach and Pace, 1995; Lopez-Garciaet al., 2002; Martinez-Murcia et al., 1995; Jurgens et al., 2000). Primer pairs were tested against DNA extracted from two Crenarchaeote and two Euryarchaeote type-strains, *E. coli* and two hydrothermal sediment samples from the Tokaanu and Waiotapu thermal region in New Zealand. DNA extraction was conducted using the modified Zhou method (Stach et al., 2001).

PCR was conducted using 0.5 AM primer, 0.2 mM dNTPs, 1.5 mM Mg+, and Roche *Taq* polymerase (1 unit) under the following conditions: Initial Dena-turation-2 min at 94 °C; 30 cycles-1 min at 94 °C, 1 minat55 °C (A571F/UA1204RandA751/UA1406R) or 1 min at 50 °C (A2Fa/U1510R), 1 min at 72 °C; Extension-10 min at 72 °C.

The two new primer pairs were used successfully in amplifying DNA from *Sulfolobus solfataricus*, *S. shibatii, Pyrococcus woesii* (data not shown) and *Thermococcus litoralis* and did not amplify *E. coli* DNA. Both the new primer sets and the published set, A2Fa/U1510R, also amplified DNA extracted from Waiotapu hydrothermal sediment samples, NZ, and the amplicon yield using the new primers was higher than that with the published set. The new primer sets were also effective in amplifying a product from environmental DNA extracted from the Tokaanu hy-drothermal region, NZ, where the published primer set failed (Fig. 3). The PCR product amplified using A571 and A1204R was cloned into pGEM-T Easy System 1 (Promega) and transformed into MC1061 competent *E. coli.* colonies containing different *Hinf1* restriction fragment patterns were sequenced using M13 primers. Unique partial 16S rRNA Archaeal sequences were obtained which clustered within the sub-division Crenarchaeota.

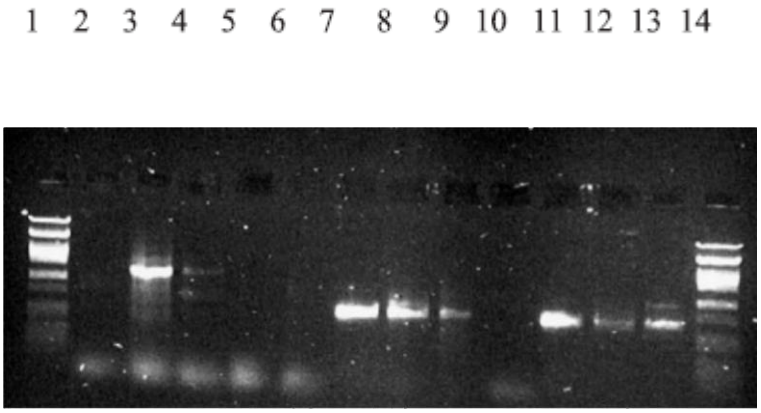

Fig. 3. One percent agarose TBE gel showing PCR products run against a E *Pst* 1 digest. Primers—Lanes 2-5: A2Fa/U1510R; 6-9: A571/ UA1204R; 10-13: A751/UA1406R. Templates—Lanes 2, 6, 10: *E. coli;* 3,7,11: *T. litoralis;* 4, 8, 12: Waiotapu environmental sample; 5, 9, 13: Tokaanu environmental sample.

8. Conclusions

PCR has been used to great effect to identify organisms that are as yet unculturable in vitro (e.g.,Hill et al., 2000; Polz and Cavanaugh, 1998; Theron and Cloete, 2000; Ward et al., 1990). Bias in phylogenetic analysis is introduced through differential amplification caused by differences in the efficiency of primer binding, interference by sequences flanking primer regions (Hansen et al., 1998) and differences in the kinetics of the PCR reaction. (e.g., Brunk and Eis, 1998; Reysenbach and Pace, 1995; Suzuki and Giovanni, 1996). As a consequence, many, if not all, 16S rDNA libraries will not be totally representative of microbial communities, especially on a quantitative level (Farelly et al., 1995; Reysenbach et al., 1992).

None of the primers in current use are truly "universal" and no single set of primers can be recommended that are guaranteed to amplify all prokaryotes. We have designed two primer pairs that potentially amplify representatives from all Archaeal groups and may access a greater Archaeal diversity than is possible with previously published primers. In addition, we provide data in Figs. 1 and 2 and their associated tables as a tool to aid the choice of the most appropriate primers for specific objectives. We also emphasise, as have others (Polz and Cavanaugh, 1998; Suzuki and Giovanni, 1996; Wintzingerode et al., 1997), that the pooling of several PCR reactions utilising slightly different primers may significantly reduce bias and provide a more accurate understanding of microbial community structure.

Acknowledgements

We would like to thank BBSRC, the Claude Harris Leon Foundation and NRF-IRDP for financial assistance. Archaeal biomass kindly provided by CAMR, Porton Down, UK.

References

Baker, G.C., Beebee, T.J.C., Ragan, M.A., 1999. *Prototheca ri-chardsi,* a pathogen of anuran larvae, is related to a clade of protistan parasites near the animal-fungal divergence. Microbiology 145, 1777-1784.

Baker, G.C., Gaffar, S., Cowan, D.A., Suharto, A.R., 2001. Bacterial community analysis of Indonesian hot springs. FEMS Microbiol. Lett. 200, 103 -109. Barns, S.M., Fundyga, R.E., Jeffries, M.W., Pace, N.R., 1994. Remarkable archaeal diversity detected in Yellowstone national Park hot spring environment. Proc. Natl. Acad. Sci. U. S. A. 91, 1609-1613.

Barns, S.M., Delwiche, C.F., Palmer, J.D., Pace, N.R., 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. PNAS 93 (17), 9188-9193.

Brunk, C.F., Eis, N., 1998. Quantitative measure of small subunit rRNA gene sequences of the Kingdom Korarchaeota. Appl. Environ. Microbiol. 64, 5064-5066.

Burggraf, S., Heyder, P., Eis, N., 1997. A pivotal archaea group. Nature (London) 385, 780.

Casamayor, E.O., Massana, R., Benlloch, S., Ovreas, L., Diez., B., Goddard, V., Gasol, J.M., Joint, I., Rodriguez-Valera, F., Predros-Alio, C., 2002. Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparision of genetic fingerprinting methods in a multipond solar saltern scaling plots. Environ. Microbiol. 4, 338 -348.

Coloqhoun, J.A., 1997. Discovery of deep-sea actinomycetes. PhD Dissertation. Research School of Biosciences. University of Kent, Canterbury.

DasSarma, S., Fleischmann, E.F., 1995. Archaea: A Laboratory Manual—Halophiles. Cold Spring Harbour Laboratory Press, New York, pp. 269-272.

DeLong, E.F., Wu, K.Y., Prezelin, B.B., Jovine, R.V.M., 1994. High abundance of Archaea in Antarctic marine picoplankton. Nature 371, 695-697.

Elwood, H.J., Olsen, G.J., Sogin, M.L., 1985. The small-subunitribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata.* Mol. Biol. Evol. 2, 399.

Farelly, V., Rainey, F.A., 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. Appl. Environ. Microbiol. 61, 2798-2801.

Groflkopf, R., Janssen, P.H., Liesack, W., 1998. Diversity and structure of the methanogenic community in anoxic rice paddy soil microcosms as examined by cultivation and direct 16S rRNA gene sequence retrieval. Appl. Environ. Microbiol. 64, 960-969.

Hansen, M.C., Tolker-Neilson, T., Givskov, M., Molin, S., 1998. Biased 16S rDNA PCR amplication caused by interference from DNA flanking template region. FEMS Microbiol. Ecol. 26, 141-149.

Hill, G.T., Mitkowski, N.M., Aldrich-Wolfe, L., Emele, L.R., Jurkonie, D.D., Ficke, A., Maldonado-Ramirez, S., Lynch, S.T., Nelson, E.B., 2000. Methods for assessing the composition and diversity of soil microbial communities. Appl. Soil Ecol. 15,25 -36.

Hohn, M.J., Hedlund, B.P., Huber, H., 2002. Detection of 16S rDNA sequences representing the novel phylum "Nanoarchaeota": indication for a wide distribution in high temperature biotopes.Syst. Appl. Microbiol. 25 (4), 551 -554.

Huber, H., Hohn, M.J., Rachel, R., Fuchs, T., Wimmer, V.C., Stetter, K.O., 2002. A new phylum of Archaea represented by a nano-sized hyperthermophilic symbiont. Nature 417, 63-67.

Huber, H., Hohn, M.J., Stetter, K.O., Rachel, R., 2003. The phylum Nanoarchaeota: present knowledge and future perspectives of a unique form of life. Res. Microbiol. 154 (3), 165-171.

Hugenholtz, P., Pitulle, C., Hershberger, K.L., Pace, N.R., 1998. Novel division level bacterial diversity in a Yellowstone hot spring. J. Bacteriol. 180, 366-376.

Jurgens, G., Lindstroem, K., Saano, A., 1997. Novel group within the kingdom Crenarchaeota from boreal forest soil. Appl. Environ. Microbiol. 63, 803-805.

Jurgens, G., Glockner, F.O., Amman, R., Saana, A., Montonen, L., Likolammi, M., Munster, U., 2000. Identification of novel Arch-aea in bacterioplankton of a boreal forest lake by phylogenetic analysis and fluorescent in situ hybridisation. FEMS Microbiol.Ecol. 34, 45 -56.

Kim, H., Honda, D., Hanada, S., Kanamori, N., Shibata, S., Miyaki, T., Nakamura, K., Oyaizu, H., 2000. A deeply branched novel phylotype found in Japanese paddy soils. Microbiology 146, 2309-2315.

Kolganova, T.V., Kuznetsov, B.B., Tourova, T.P., 2002. Designing and testing oligonucleotide primers for amplification and sequencing of Archaeal 16S rRNA genes. Mikrobiologiya 71, 283-286.

Lee, S.-Y., Bollinger, J., Bezdicek, D., Ogram, A., 1996. Estimation of the abundance of an uncultured soil bacterial strain by a competitive quantitative PCR method. Appl. Environ. Microbiol. 62, 3787-3793.

Lopez-Garcia, P., Moreira, D., Lopez-Lopez, A., Rodriguez-Valera, F., 2001. A novel haloarchaeal-related lineage is widely distributed in deep oceanic regions. Environ. Microbiol. 3, 72-78.

L(Spez-Garcia, P., Gaill, P., Moreira, D., 2002. Wide bacterial diversity associated with tubes of the vent worm *Riftia pachyptila*. Environ. Microbiol. 4, 204-215.

MacGregor, B.J., Moser, D.P., Alm, E.W., Nealson, K.H., Stahl, D., 1997. Crenarchaeota in Lake Michigan sediment. Appl. Environ. Microbiol. 63, 1178-1181.

Marteinsson, V.T., Kristjansson, K.J., Kristmanndottin, M.D., Sae-mundsson, K., Hannington, M., Petursdottin, S.K., Geptner, A., Stoffers, P., 2001. Discovery and description of giant submarine smectite submarine cones on the seafloor, Eyjafjordur, Northern Iceland, a novel thermal microbial habitat. App. Environ. Microbiol. 67 (2), 827-833.

Martinez-Murcia, A.J., Acinas, S.G., Rodriguez-Valera, F., 1995. Evaluation of prokaryotic diversity by restrictase digestion of 16S rDNA directly amplified from hypersaline environments.FEMS Microbiol. Ecol. 17, 247-256.

McInnery, J.O., Wilkinson, M., Patching, J.W., Embley, T.M., Powell, R., 1995. Recovery and phylogenetic analysis of novel archaeal rRNA sequences from deep-sea deposit feeder. Appl. Environ. Microbiol. 61, 1646-1648.

McPherson, M.J., Hames, B.D., Taylor, G.R., 1995. PCRII. Oxford Univ. Press, UK.

Nielsen, A.T., Liu, W.-T., Filipe, C., Grady, L., Molin, S., Stahl, D.A., 1999. Identification of a novel group of bacteria in sludge from a deteriorated biological phosphorous removal reactor. Appl. Environ. Microbiol. 65, 1251 -1258.

Polz, M.F., Cavanaugh, C.M., 1998. Bias in template-to-product ratios in multi-template PCR. Appl. Environ. Microbiol. 64,3724-3730.

Reysenbach, A.-L., Pace, N.R., 1995. In: Robb, F.T., Place, A.R. (Eds.), Archaea: A Laboratory Manual—Thermophiles. Cold Spring Harbour Laboratory Press, New York, pp. 101 -107.

Reysenbach, A.L., Giver, L.J., Wickham, G.S., Pace, N.R., 1992. Differential amplification of rRNA genes by polymerase chain reaction. Appl. Environ. Microbiol. 58, 3417-3418.

Reysenbach, A.-L., Wickham, G.S., Pace, N.R., 1994. Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. Appl. Environ.Microbiol. 60, 2113-2119.

Rudi, K., Skulberg, O.M., Larsen, F., Jacoksen, K.S., 1997. Strain classification of oxyphotobacteria in clone cultures on the basis of 16S rRNA sequences from variable regions V6, V7 and V8. Appl. Environ. Microbiol. 63, 2593-2599.

Schleper, C., Delong, E.F., Preston, C.M., Feldman, R.A., Wu, K.-Y., Swanson, R.V., 1998. Genomic analysis reveals chromosomal variation in natural populations of the uncultured psychrophilicarchaeon cenarchaeum symbiosum. J. Bacteriol. 180, 5003-5009.

Stach, J.E.M., Bathe, S., Clapp, J.P., Burns, R.G., 2001. PCR-SSCP comparison of 16S rDNA sequence diversity in soil obtained using different isolation and purification methods. FEMS Microbiol. Ecol. 36, 139-151.

Stern, C.D., Holland, P.W.H., 1993. Essential Developmental Biology. Oxford Univ. Press, Oxford. 324 pp.

Suzuki, M.T., Giovanni, S.J., 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. 62, 625 -630.

Tajima, K., Nagamine, T., Matsui, H., Nakamura, M., Aminov, R.I., 2001. Phylogenetic analysis of archaeal 16S rRNA libraries from the rumen suggests the existence of a novel group of archaea not associated with known methanogens. FEMS Microbiol. Lett. 200, 67-72.

Theron, J., Cloete, T.E., 2000. Molecular techniques for determining microbial diversity and community structure in natural environments. Crit. Rev. Microbiol. 26, 37 -57.

Van de Peer, Y., Chapelle, S., De Wachter, R., 1996. A quantitative map of nucleotide substitution rates in bacterial rRNA. Nucleic Acids Res. 24, 3381 -3391.

van Waasbergen, L.G., Balkwill, D.L., Crocker, F.H., Bjornstad, B.N., Miller, R.V., 2000. Genetic diversity among Arthrobacter species collected across a heterogeneous series of terrestrial deep-subsurface sediments as determined on the basis of 16S rRNA and recA gene sequences. Appl. Environ. Microbiol. 66, 3454-3463.

Vetriani, C., Jannasch, H.W., McGregor, B.J., Stahl, D.A., Reysen-bach, A.-L., 1999. Population structure and phylogenetic characterisation of marine benthic Archaea in deep-sea sediments. Appl. Environ. Microbiol. 65, 4375-4384.

Ward, D.M., Weller, R., Bateson, M.M., 1990. 16S rRNA sequences reveal numerous uncultured micro-organisms in a natural community. Nature 345, 63 -65.

Watanabe, K., Kodama, Y., Harayama, S., 2001. Design and evaluation of PCR primers to amplify 16S ribosomal DNA fragments used for community fingerprinting. J. Microbiol. Methods 44, 253-262.

Whitehead, T.R., Cotta, M.A., 1999. Phylogenetic diversity of methanogenic Archaea in swine waste storage pits. FEMS Mi-crobiol. Lett. 179, 223-226.

Wintzingerode, F.V., Gobel, U.B., Stackebrandt, E., 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol. Rev. 21, 213-229.

Table 1
Bacteria specific 16S rRNA gene sequence/PCR primers

Code	Sequence 5′–3′	N	Ta	B	A	E	Matches	Ref.
E8F	AGAGTTTGATCCTGGCTCAG	20	55	+++	+	–	2424	Reysenbach and Pace, 1995; Martinez-Murcia et al., 1995; Reysenbach et al., 1994
E9F	GAGTTTGATCCTGGCTCAG	19	53	+++	+	++	2741	McInnery et al., 1995; Hansen et al., 1998; Farelly et al., 1995
E334F	CCAGACTCCTACGGGAGGCAGC	21	65	+++	–	–	13,172	Rudi et al., 1997
E341F	CCTACGGGIGGCIGCA	16	51	+++	+	–	16,685	Watanabe et al., 2001
E786F	GATTAGATACCCTGGTAG	18	47	+++	+	–	12,616	Coloqhoun, 1997
E533R	TIACCGIIICTICTGGCAC	19	56	+++	+	++	18,724	Watanabe et al., 2001
E926R	CCGICIATTIIITTTIAGTTT	20	50	+++	+ ++	++	19,950	Watanabe et al., 2001
E939R	CTTGTGCGGGCCCCCGTCAATTC	23	71	+++	–	–	8620	Rudi et al., 1997
E1115R	AGGGTTGCGCTCGTTG	16	47	+++	–	–	9052	Reysenbach and Pace, 1995
E1541R	AAGGAGGTGATCCANCCRCA	20	57	+++	+	–	1355	Suzuki and Giovanni, 1996

Primers described as bacteria-specific were submitted to the 'check probe' facility of the Ribosomal Database Project (http://www.rdp.cme.msu.edu/) to check for bacterial specificity. The total number of matches are given, and matches for Bacteria (B), Archaea (A) and Eukarya (E) are represented as follows: (–, no matches; +, <25 matches; ++, 25–100 matches; +++, <100 matches). Primer numbering relates to *E. coli* position complementary to the 5′ end of the primer. Last letter denotes direction: Forward and Reverse. Ta=(4 × (G+C) + 2 × (A+T)) – 5 M=A or C; R=A or G; W=A or T; S=C or G; Y=C or T; K=G or T; N=A or G or C or T.

Table 2
List of published archaeal 16S sequences used to represent the major taxonomic groups

Crenarchaeotes
Sulfolobales; *Sulfolobus solfataricus* (X03235)
Thermoproteales; *Thermoproteus tenax* (M35966)
Desulfurococcales; *Desulfurococcus mobilis* (M36474)
Unclassified Crenarchaeotes; *Cenarchaeum symbiosum* (U51469)

Euryarchaeotes
Archaeaglobales; *Archaeoglobus fulgidus* (X05567)
Halobacteriales; *Haloarcula argentinensis* (D50849)
Methanococcales; *Methanococcus jannaschii* (M59126)
Thermococcales; *Pyrococcus furiosus* (U20163)

Korarchaeotes
Uncultured Korarchaeote pBA5 (AF176347)
Unidentified Korarchaeote pJP27 (L25852)

Nanoarchaeota
Nanoarchaeum equitans (AJ318041)

Table 3
"Universal" and Archaea-specific 16S rRNA gene sequence/PCR primers

Code	Pub spec.	Sequence 5′–3′	N	Tm	Matches	EK1	EK2	EK3	EK4	TC	NTC	K	N	E	B	REF
A1F	A+	ATTCCGGTTGATCCTGC	17	49	89	+	–	++	+	+	–	–	–	–	–	Tajima et al., 2001
A2Fa	A+	TTCCGGTTGATCCYGCCGGA	20	60	215	+	+	+++	+	++	++	–	–	–	–	Reysenbach and Pace, 1995; Martinez-Murcia et al., 1995; Jurgens et al., 2000
A2Fb	A+	TTCCGGTTGATCCTGCCGGA	20	61	189	+	+	++	+	++	++	–	–	–	–	López-García et al., 2002
A3Fa		TCCGGTTGATCCYGCCGG	18	56	265	+	+	+++	+	++	++	–	–	–	–	McInnery et al., 1995
A3Fb		TCYGKTTGATCCYGSCRGAG	20	61	263	+	++	+++	+	+	–	–	–	–	–	Lopez-Garcia et al., 2001
N3F	N+	TCCCGTTGATCCTGCG	16	52	1	–	–	–	–	–	–	–	+	–	–	Huber et al., 2002
EK4F		CTGGTTGATCCTGCCAG	17	49	1483	–	–	++	–	–	–	–	–	+	–	Vetriani et al., 1999
A109F		ACKGCTCAGTAACACGT	17	46	512	+	+++	+++	++	+	++	–	–	–	–	Whitehead and Cotta, 1999
Kb228F	K/A	GAGGCCCCAGGRTGGGACCG	20	66	8	+	–	–	–	+	–	+	–	–	–	Brunk and Eis, 1998
Eb246F	A+B+	AGCTAGTTGGTGGGGT	16	45	3021	–	–	–	–	–	–	–	–	–	+++	DasSarma and Fleischmann, 1995
A333F	A+	TCCAGGCCCTACGGG	15	51	495	+	+++	+++	++	+	+	–	–	–	–	Reysenbach and Pace, 1995
A340F	A+	CCCTACGGGGYGCASCAG	18	60	766	++	+++	+++	+++	++	+	–	–	–	–	Vetriani et al., 1999
U341F		CCTACGGGRSGCAGCAG	17	54	16776	++	+++	+++	++	++	+	–	–	–	+++	Hansen et al., 1998
A344F	A+	ACGGGGTGCAGCAGGCGCGA	20	68	709	++	+++	+++	++	+	+	–	–	–	–	Casamayor et al., 2002
Kb366F	K/A	CTCCGCAATRCGCGMAAG	18	53	24	+	+	–	+	+	+	+	–	–	–	Brunk and Eis, 1998
U515F	A+B+	GTGCCAGCMGCCGCGGTAA	19	60	18695	++	+++	+++	+	+	+	–	–	–	+++	Reysenbach and Pace, 1995; Reysenbach et al., 1992
U519F		CAGCMGCCGCGGTAATWC	18	54	18018	+	+	+++	++	+++	++	–	–	–	+++	Suzuki and Giovanni, 1996
A571F		GCYTAAAGSRICCGTAGC	18	52	675	++	++	+++	++	+++	+++	+	+	–	–	This review
K604F	K/A	GTTAAATCCGCCTGAAGACA	20	53	2	–	–	–	–	–	–	+	–	–	–	Brunk and Eis, 1998
A751F		CCGACGGTGAGRGRYGAA	18	54	456	+	+	+++	++	+++	++	+	+	–	–	This review
Ab779F	A/K	GCRAASSGGATTAGATACCC	20	56	1635	++	+++	+++	++	+++	+++	+	–	–	+++	Brunk and Eis, 1998
Eb787F	A+B+	ATTAGATACCCTGGTA	16	39	12682	–	–	+	–	–	–	–	–	–	+++	DasSarma and Fleischmann, 1995
Ab787F	A+B+	ATTAGATACCCGGGTA	16	41	1304	++	+++	+++	++	++	+++	+	–	–	+++	DasSarma and Fleischmann, 1995
Ab789F		TAGATACCCSSGTAGTCC	18	51	2035	++	+++	+++	++	+++	+++	+	–	–	+++	Barns et al., 1994
Ab906F	A+	GAAACTTAAAKGAATTG	17	38	5070	++	+++	+++	++	+++	+++	+	–	–	+++	Reysenbach and Pace, 1995
A1040F	A+	GAGAGGWGGTGCATGGCC	18	55	699	+	+++	+++	++	+++	+++	–	+	–	–	Reysenbach and Pace, 1995
U1053F	A++B+	GCATGGCYGYCGTCAG	16	49	9988	+	+++	+++	++	+	+	–	+	–	+++	DasSarma and Fleischmann, 1995
A1098F	A+B+	GGCAACGAGCGMGACCC	17	54	437	+	+++	+++	+	+++	+	+	–	–	–	Reysenbach and Pace, 1995

Mb1225F	A+B+	ACACGCGTGCTACAAT	16	43	73	–	–	+	–	–	–	–	–	–	+	DasSarma and Fleischmann, 1995
Ab127R	A++	CCACGTGTTACTSAGC	16	45	570	+	+++	+++	+	+++	+	–	–	–	+	DasSarma and Fleischmann, 1995
A348R		CCCCGTAGGGCCYGG	15	50	709	+	+++	+++	++	+++	+	+	–	–	–	Barns et al., 1994
EK510R	EK	CTTGCCCRGCCCTT	14	42	406	+	++	+++	++	–	–	–	–	–	–	DasSarma and Fleischmann, 1995
TC518R	C	ACACCAGRCTTGCCCCCCGCTT	22	68	50	–	–	–	–	++	–	–	–	–	–	Barns et al., 1994
U529R	A+B++	ACCGCGGCKGCTGGC	15	50	19143	++	+++	+++	+	+	+	+	–	–	+++	DasSarma and Fleischmann, 1995
U534R	A+B+	GWATTACCGCGGCKGCTG	18	54	18018	+	+	+++	++	++	++	–	–	–	+++	DasSarma and Fleischmann, 1995
K624R	K/A	TGTCTTCAGGCGGATTTAAC	20	53	2	–	–	–	–	–	–	+	–	–	–	Brunk and Eis, 1998
M704R	A++	TTACAGGATTTCACT	15	35	166	–	–	+++	–	–	–	–	–	–	–	DasSarma and Fleischmann, 1995
Ab909R	A/K	TTTCAGYCTTGCGRCCGTAC	20	49	1939	++	+++	+++	++	+++	++	+	+	–	+++	Brunk and Eis, 1998
U926R	A+B+	CCGTCAATTCCTTTRAGTTT	20	50	14386	–	+	+	–	–	–	+	–	–	+++	Reysenbach and Pace, 1995
Ab927R		CCCGCCAATTCCTTTAAGTTTC	22	59	920	++	+++	+++	++	+++	+++	–	–	–	+	Jurgens et al., 2000
A934R	A+	GTGCTCCCCCGCCAATTCCT	20	61	830	++	+++	+++	++	+++	+++	–	–	–	–	DasSarma and Fleischmann, 1995
A976R	A+	YCCGGCGTTGAMTCCAATT	19	53	498	+	+++	+++	++	++	++	–	–	–	–	Reysenbach and Pace, 1995
A1115R	A+	GGGTCTCGCTCGTTG	15	45	465	+	+++	+++	+	+++	+	+	–	–	–	Reysenbach and Pace, 1995
EKb1242R	A+	CCATTGTAGCSCGCGTG	17	51	454	+	+++	+++	++	++	–	–	–	–	–	DasSarma and Fleischmann, 1995
UA1204R		TTMGGGGCATRCIKACCT	18	53	645	+	++	+++	++	++	++	+	+	–	–	This review
UA1406R		ACGGGCGGTGWGTRCAA	17	52	17237	+	++	+++	++	++	+	+	+	–	+++	This review
N1406R	N+	ACGGGCGGTGAGTGCAA			1	–	–	–	–	–	–	–	+	–	–	Huber et al., 2002
U1406R	A+B+	GACGGGCGGTGTGTRCA	17	52	16819	+	++	+++	++	++	++	–	–	–	+++	Reysenbach and Pace, 1995; Hansen et al., 1998
Eb1415R		ACGGCGGTGTGTRC	14	42	29	–	–	+	–	–	+	–	–	–	+++	Lee et al., 1996
N1510R	N+	ACGGCTACCTTGTGTCGACTT			1	–	–	–	–	–	–	–	+	–	–	Huber et al., 2002
U1510R	A+B+	GGTTACCTTGTTACGACTT	19	49	950	+	–	+	+	+	–	–	–	–	+++	Reysenbach and Pace, 1995; Tajima et al., 2001; Lopez-Garcia et al., 2001

Primers described as archaeal-specific were submitted to the 'check probe' facility of the Ribosomal Database Project (http://www.rdp.cme.msu.edu/) to check for archaeal specificity. All forward primers were submitted as 'target sequence' and reverse primers were submitted as probes. The search was based on a 100% match. The annealing temperature, number of matches to the database and specificity to major taxonomic groups are listed. Primers have been re-named according to their specificity as judged by results of the Probe Match and complementarity to the 1300 bp archaeal alignment. The number of total matches is noted as a general measure of primer specificity, accepting that this measure is totally dependent on the bias of sequences present in the database. Primers specific to either ends of the gene have a low number of matches, because these areas are less often sequenced, with the majority of sequence information being available for the V2–V6 regions. Certain taxa, such as the Korarchaeotes and Nanoarchaea have a low hit rate as there are currently very few representatives from these taxa in the database. Column headings: EK1 = Methanococcales; EK2 = Methanobacteriales; EK3 = Methanomicrobacteria and relatives; EK4 = Thermococcales and Methanopyrales; TC = Thermophilic Crenarchaeotes; NTC = Non-thermophilic crenarchaeotes; K = Korarchaeotes; N = Nanoarchaea; B = Bacteria; E = Eukaryotes. Ribosomal RNA database matches: – = no matches; + = < 25 matches; ++ = 25–100 matches; +++ = >100 matches.

Fig. 2. Alignment of representative taxa from each of major Archaeal subdivisions. Tabulated Clustal alignment of 1300 bp (from base 106–1406; *E. coli* numbering) of archaeal 16S rRNA gene sequences. Row 1 = Crenarchaeotes; Row 2 = Euryarchaeotes; Row 3 = Korarchaeotes; Row 4 = Nanoarchaeotes (see Table 2). Sequence data from base numbers 0–105 and >1407 was not available for all 11 taxa, so has been omitted from the table. The sequence is annotated with "universal" and Archaea-specific primer positions. Bases not complementary to the primers are highlighted in blue (where there were several variants of one primer the non-consensus bases are highlighted).

	TTCCCGRCCCGTTC EK510R TTCGCCCCCCGTTCR-GACCACA TC518R GTGCCAGCMGCCGCGGTAA U515F A571F GCYTAAAGSRICCGTAGC CAGCMGCCGCGGTAATWC U519F CAGCCGCCGCGGGAACAC N519F CGGTCGKCGGCGCCA U529R GTCGKCGGCGCCATTAWG U534R
X03235 M36474 M35966 U51469	CCCGCTCTAAAAAGGCGGGGG-AATAAGCGGGGGGGCAAGT-CTGGTGTCAGCCGCCGCGGTAATACCAGCTCCGCGAGTGGTCGGGGTGATTACTGGGCCTAAAGCGCCTGTAGCCGG CCCGCTGTAGGAAGGCGGGGG-AATAAGCGGGGGGGCAAGT-CTGGTGTCAGCCGCCGCGGTAATACCAGCCCCGCGAGTGGTCGGGACGATTATTGGGCCTAAAGCGCCCGTAGCCGG CCCGGTGTAAGGAGCCGGGGCG-AATAAGCGGGGGGTAAGT-CTGGTGTCAGCCGCCGCGGTAATACCAGCCCCGCGAGTGGTCAGGGTGATTACTGGGCTTAAAGCGCCCGTAGCCGG ACCGGTCTTAAAACCACCGGCGAATAAGGGGTGGGCAAGTTCTGGTGTCAGCCGCCGCGGTAAAACCAGCACCTCAAGTGGTCAG^^
X05567 D50849 M59126 U20163	GGGTGCCTAAAAAGCACCCCACAGCAAGGGCCGGGCAAGG-CCGGTGGCAGCCGCCGCGGTAATACCGGCGGCCCGAGTGGCGGCCACTTTTATTGGGCCTAAAGCGTCCGTAGCCGG TGTACCGTAAGGTGGTACAGG-AACAAGGACTGGGCAAGA-CCGGTGCCAGCCGCCGCGGTAATACCGGCAGTCCAAGTGATGGCCGATATTATTGGGCCTAAAGCGTCCGTAGCTTG CGGAGTGTAAACAGC - TCCGGGAATAAGGGCTGGGCAAGT - CCGGTGCCAGCAGCCGCGGTAATACCGGCGGCCCAAGTGGTGGCC7ACTGTTATTGGGCCTAAAGCGTCCGTAGCCGG CGGAGTGTAAAAAGC- TCCGGGAATAAGGGCTGGGCAAGG-CCGGTG CAGC GCCGCGGTAATACCGGCGGCCCGAGTGGTGGCCACTATTATTGGGCCTAAAGCGGCCGTAGCCGG
AF176347 L25B52	CCCTGTGTAAAAAGCAGGGGGTAGGAAGGGGAGGGTAAGG-CTGGTGCCAGCCGCCGCGGTAAAACC7AGCTCCCCGAGGGGTTCCC7ACGCATACTGGGCCTAAAGCGTCCGTAGCTGG CCCTGTGTAAAAAGCAGGGGGTAGGAAGGGGAGGGTAAGG-CTGGTGCCAGCCGCCGCGGTAAAACCAGCTCCCCGAGGGGTTCCCACGCATACTGGGCCTAAAGCGTCCGTAGCCGG
AJ318041	GGGAGTGTAAGTAGCTCCCCG-AATAAG G GGGCAAGA-GGGGTGGCAGCCGCCGCGGAACACCCCCACCGCGAGCGGTGGCCGTGATTATTGGGCCTAAAGGGGCCGTAGCCGG

	GTTAAATCCGCCTGAA GACA K604F CAATTTAGGCGGACTT CTGT K624R M704R TCACTTTAGGACATT
X03235 M36474 M35966 U51469	CCCACCAAGTCGCCCCTTAAAGTCCCCGGCTCAACCGGGGAACTGG-GGGCGATACTGGTGGGCTAGGGGGCGGGAGAGGCCGGGGGGTACTCCCGGAGTAGGGGCGAAATCCTTAGAT CCCGGCAAGTCCCCTCCTAAATTCCCGGGCTCAACCCGGGGACTGG-AGGGGATACTGCCGGGCTAGGGGGTGGGAGAGGCCGAGGGTACTCCCGGGGTAGGGGCGAAATCCTATAAT CCCGGCAAGTCGCTCCTGAAATCCCCAGGCTCAACCTGGGGGCAGG-GGGCGATACTGCCGGGCTAGGGGGCGGGAGAGGCCGCCGGTACTCCGGGGGTAGGGGCGAAATCCTATAAT CCGTGTAAGTTTTCGGTTAAATCTATGCGCTCAACGCATAGGCTGCC-GGAAATACTGCACGGCTAGGGAGTGGGAGAGGTAGACGGTACTCGGTAGGAAGGGGTAAAATCCTTTGAT
X05567 D50849 M59126 U20163	GCTGGTAAGTCCTCCGGGAAATCTGGCGGCTTAACCGTCAGACTGCCGGAGGATACTGCCAGCCTAGGGACCGGGAGAGGCCGGGGGTATTCCCGGAGTAGGGGTGAAATCCTGTAAT CTGTGTAAGTCCIATTGGGAAATCGACCIAGCTCIAACTGGTCGGCGTCCGGTGGAAACTACACAGCTTGGGGCCGAGAGACTCAACGGGTACGTCCGGGGTAGGAGTGAAATCCTGTAAT CCCNGTAAGTCTCTGCTTAAATCGTGCGGCTCAACCGCAGGGCTGC- -AGAGATACTGCCNNNNTTGGNACCGGGAGAGGCCNGGGGTACCCCAGGGGTAGCGGTGAAATGCGTTGAT GCCCGTAAGTCCCTGGCGAAATCCCACGGCTCAACCGTGGGGCTCGCTGGGGATACTGCGGGCCTTGGGACCGGGAGAGGCCGGGGGTACCCCCGGGGTAGGGGTGAAATCCTATAAT
AF176347 L25852	CCCCGTAAGTCCTCGGTTAAATCCGCCTGAA-GACAGGCGGACCGC-CGAGGATACTGCGGGGCTAGGGAGCGGGAGGGGCCGAGGGTATTCCGGGGGGAGCGGTAAAATGCGTAGAT CCCCGTAAGTCCTCGGTTAAATCCGCCTGAA-GACAGGCGGACCGC-CGAGGATACTGCGGGGCTAGGGAGCGGGAGGGGCCGAGGGTATTCCGGGGGGAGCGGTAAAATGCGTAGAT
AJ318041	GCCGGTGTGGCTCCGGTGAAATCCTCGGGCTCAACCCGAGGGCGCGCCGGAGCTACTACCGGCCTAGGGACCGGGAGGGGCCGACCGTACTCCCGGGGGAGCGGTGAAATGCTGTAAT
	CCGACGGTGAGRGRYGAA A751F ATTAGATACCCTGGTA Eb787F ATTAGATACCCGGGTA Ab787F TAGATACCCSSGTAGTCC Ab789F GCRAASSGGATTAGATACCC Ab779F
X03235 M36474 M35966 U51469	ACCGGGAGGACCACCAGTGGCGGAAGCGCCCCGCTAGAACGCGCCCGACGGTGAGAGGCGAAAGCCGGGGCAGCAAACGGGATTAGATACCCCGGTAGTCCCGGCTGTAAACGATGCG CCCGGGAGGACCACCAGTGGCGAAGGCGCTCGGCTGGAACACGCCCGACGGTGAGGGGCGAAAGCCGGGGGAGCGAACCGGATTAGATACCCGGGTAGTCCCGGCTGTAAACGATGCG CCCCGGAGGACCACCAGTGGCGAAAGCGGGCGGCCAGAACGCGCCCGACGGTGAGGGGCGAAAGCCGGGGGAGCAAAGGGGATTAGATACCCCTGTAGTCCCGGCCGTAAACGATGCG CTATCGATGACCCCCTGTGGCGAAGGCGGTCTACTAGAACACGTCCGACGGTGAGGGATGAAAGCTGGGGGAGCAAACCGGATTAGATACCCGGGTAGTCCCAGCTGTAAACAATGCA
X05567 D50849 M59126 U20163	CCCGGGAGGACCACCTGTGGCGAAGGCGCCCGGCTGGAACGGGTCCGACGGTGAGGGACGAAGGCCAGGGGAGCGAACCGGATTAGATACCCGGGTAGTCCTGGCTGTAAACGATGCG CCTGGACGGACCACCAATGGGGAAACCACGTTGAGAGACCGGACCCGACAGTGAGGGACGAAAGCCAGGGTCTCGAACCGGATTAGATACCCGGGTAGTCCTGGCTGTAAACAATGCT CCCTGGGGGACCACCTGTGGCGAAGGCGCCCGGCTGGAACGGGTCCGACGGTGAGGGACGAAGGCCAGGGGAGCAAACCGGATTAGATACCCGGGTAGTCCTGGCTGTAAACTCTGCG CCCGGGGGGACCGCCAGTGGCGAAGGCGCCCGGCTGGAACGGGTCCGACGGTGAGCGCGGAAGGCCAGGGGAGCGAACCGGATTAGATACCCGGGTAGTCCTGGCTGTAAAGGATGCG
AF176347 L25852	CCCCGGAGGACCACCAGTGGCGAAGGCGCTCGGCTGGAACGCGTCCGACGGTGAGGGACGAAAGCTGGGGGAGCAAACCGGATTAGATACCCGGGTAGTCCCAGCCGTAAACGATGCC CCCCGGAGGACCACCAGTGGCGAAGGCGCTCGGCTGGAACGCGTCCGACGGTGAGGGACGAAAGCTGGGGGAGCAAACCGGATTAGATACCCGGGTAGTCCCAGCCGTAAACGATGCC
AJ318041	CCCGGGAGGACGACCCGTGGCGAAAGCGGTCGGCCAGAACGGGTCCGACGGTGAGGG CGAAGGCCGGGGGC GAACGGGATTAGA ACCC GTA TCCCGGCTGTCAACGCTGCG

Fig. 2 *(continued)*.

	GAAACTTAAAK6AATT6 Ab906F CATGCCRGCGTTCYGACTTT Ab909R CTTTGAATTTCCTTAACCGCCC Ab927R TTTGARTTTMCTTAACTGCC U926R A934R TCCTTAACCGCCCCCTCGTG
X03235 M36474 M35966 U51469	GGCTAGGTGTCGAGTA-GGCTTAGAGCCTACTCGGTGCCGCAGGGAAGCCGTTAAGCCCGCCGCCTGGG<^AGTACGGTCGCAAGACTGAAACTTAAAGGAATTGGCGGGGGAGCACCA GGCTAGGTGTTGGGTG-GGCTTAGAGCCCACCCAGTGCCGCAGGGAAGCCGTTAAGCCCGCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGGCGGGGGAGCACCA GGCTAGCTGTCGGCCG-GGCTTAGGGCCCGGCCGGTGGCGTAGGGAAACCGTTAAGCCCGCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTG CGGGGGGGCACCA AACTCAGTGATGCATTGGTGTAAGAGCCAGTGCAGTGCCGCAGGGAAGCCGTTAAGTTTGCCGCCTGGGAAGTACGTACGCAAGTATGAAACTTAAAGGAATTGGCGGGGGAGCACCA
X05567 D50849 M59126 U20163	GACTAGGTGTCACCGAAG-CTACGAGCTTCGGTGGTGCCGGAGGGAAGCCGTTAAGTCCGCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGGGGGAGCACTA CGCTAGGTAT- GTCACGCGCCATGAGCACGTGATGTGCCGTAGTGAAGACGATAAGCGAGCCGCCTGGGAAGTACGTCCGCAAGGATGAAACTTAAAGGAATTGGCGGGGGAGCACCA GACTAGGTGTCGCGTCGG- CTTCGGGCNGACGCGGTGCCGAAGGGAAGCCGTTAAGTCCGCCGCCTGGGGAGTACGGTCGCAAGACTGAAACTTAAAGGAATTGGCGGGGGAGCACTA GGCTAGGTGTCGGGCGAG- CTTCGAGCTCGCCCGGTGCCGTAGGGAAGCCGTTAAGCCCGCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGGGGGAGCACTA
AF176347 L25852	GGCTAGGTGCCGGCTGAGGTTTCGGCCTCAGCCGGTGTCGAAGCGAAGGCATTAAGCCGGCCGCCTGAGGAGTACAGCCGCAAGGCCGAAACTTAAAGGAATTGACGGGGGGGCACCA GGCTAGGTGCCGGCTGAGGTTTCGGCCTCAGCCGGTGTCGAAGCGAAGGCATTAAGCCGGCCGCCTGAGGAGTACAGCCGCAAGGCCGAAACTTAAAGGAATTGACGGGGGGGCACCA
AJ318041	GGCTACCTGCTGGGCG-GGCTACGAGCCCGCCCAGTGGGGTAGGGAAGCCGTTAAGCCCGCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTTAAAGGAATAGGCGGGGGAGCAC-A
	GAGAGGWGGTGCATG TTAACCTMAGTTGCGGCCY A976R GCATG
X03235 M36474 M35966 U51469	CAAGGGGTGGA-ACCTGCGGCTCAATTGGAGTCAACGCCTGGAATCTTACCGGGGGAGACCGC-AGTATGACGGCCAGGCTAACGACCTTGCCTGACTC-GCGGAGAGGAGGTGCATG CAAGGGGTGGA- GCCTGCGGTTCAATTGGAGTCAACGCCGGGAATCTCACCGGGGGAGACAGC-AGGATGACGGCCAGGTTAAAGGCCTTGCCTGACGC-GCTGAGAGGAGGTGCATG CAAGGGGTGAA- GCTTGCGGCTTAATTGGAGTCAACGCCGGAAACCTTACCCGGGGCGACAGC-AGGATGAAGGCCAGGCTAACGACCTTGCCGGACGA-GCTGAGAGGAGGTGCATG CAAGGGGTGAA- - GCTGCGGTTCAATTGGAGTCAAGGCNAGAAATCTTACCCGGAGAGACAGCCAGTGTGAAGGTCAAGGTGAAGACTTTACCAGACAA-GCTGAGAGGTGGTGCATG
X05567 D50849 M59126 U20163	CAACGGGTGGA-GCCTGCGGTTTAATTGGATTCAACGCCGGGAAGCTTACCGGGGGAGACAGC-GGGATGAAGGTCGGGCTGAAGACCTTACCAGACTA-GCTGAGAGGTGGTGCATG CAACCGGAGGA- GCCTGCGGTTTAATTGGACTCAACGCCGGACATCTCACCGGTCCCGACAGTAGTAATGACAGTCAGGTTGACGACTTTACTCGACGCTACTGAGAGGAGGTGCATG CAACGGGTGGN- NACCGCGGTTTAATTGGATTCAACGCCGGGCATCTTACCAGGGGCGACGGC-AGGATGAAGGCCAGGTTGACGACCTTGCCAGACGC-GCCGAGAGGTGGTGCATG CAAGGGGTGGA- GCGTGCGGTTTAATTGGATTCAACGCCGGGAACCTCACCGGGGGCGACGGC-AGGATGAAGGCCAGGCTGAAGGTCTTGCCGGACGC-GCCGAGAGGAGGTGCATG
AF176347 L25852	CAAGGGGTGAATGCCTGCGGCTCAATTGGACTCAACGCCGGGAATCTTACCGGGGGCGACAGC-AGGATGAAGGTCAGGCTGAAGACCTTACCTGACGC-GCTGAGGGGTGGTGCATG CAAGGGGTGAATGCCTGCGGCTCAATTGGACTCAACGCCGGGAATCTTACCGGGGGCGACAGC-AGGATGAAGGTCAGGCTGAAGACCTTACCTGACGC-GCTGAGGGGTGGTGCATG
AJ318041	CAAGAGGTGGG-GTGCGCGGTTTAATTGGATTC ACGCCGGGAACCTCACCGGGGCTGACAGC-ACAATGATGGTCGGCCTGAAGGGCCTACCGGAGGC-GCTGAGAGGAGGTGCATG
	GCC A1040F GCYGYCGTCAG U1053F A1115R GTTGCTCGCTCTGGG A1098F GGCAACGAGCGMGACCC
X03235 M36474 M35966 U51469	GCCGTCGCCAGCTCGTGTTGTGAAATGTCCGGTTAAGTCCGGCAACGAGCGAGACCCCCACCCCTAGTTGG - TTTCTGGACT - C - CGGTCCAGAAC - CACACTAGGGGGACTGCCGGC GCCGTCGCCAGCTCGTGCTGTGAAGTGTCCGGTTAAGTCCGGAAACGAGCGAGACCCCCACCCCTAGTTGCTACCCGGGGCT-A-CGGCTCCGGGG-CACACTAGGGGGACTGCCGCC GCCGTCGTCAGCTCGTGCCGTGAGGTGTCCGGTTAAGTCCGGCAACGAGCGAGACCCCCACCCCTAGTTGCTACCCCGCTCT-T-CGGGGCGGGGGGCACACTAGGGGGACTGCCGGC NCCGTCGCCAGCTCGTGCCGTGAGATGTCCTGTTAAGTCAGGTAACGAGCGAGATCCTTGCCTCTAGTTGCCACCATTGCTCTCCGGAGCAGTGGGGCGAATTAGCGGGACCGCCG-C
X05567 D50849 M59126 U20163	GCCGCCGTCAGTTCGTACTGTGAAGCATCCTGTTAAGTCAGGCAACGAGCGAGACCCGCGCCCCCAGTTGCCAGCGGTTCCCTTCGGGGAAGCCGGGCACACTGGGGGGACTGCCGGC GCCGCCGTCAGCTCGTACCGTGAGGCGTCCTGTTAAGTCAGGCAACGAGCGAGACCCACACTTCTAGTTGCCAGCAACACCCCTGCGG-TGGTTGGGTACACTAGGAGGACTGCCATT GCCGTCGTCAGCTCGTACCGTGAGGCGTCCTGTTAAGTCAGGTAACGAGCGAGACCCGTGCCCCATGTTGCTA-CCTCCTCCGNCCGGGAGGA-GNGCACTCATGGGGNACCGNCGGC GCCGCCGTCAGCTCGTACCGTGAGGCGTCCACTTAAGTGTGGTAACGAGCGAGACCCGCGCCCCCAGTTGCCAGTCCCTCCCGCTCGGGAGG--AGGCACTCTGGGGGGACTGCCGGC
AF176347 L25852	GCCGTCGCCAGCTCGTGCCGTGAGGTGTCCTGTTAAGTCAGGCAACGAGCGAGACCCCCGCCCTCAGTTGCCAGCGGGGCCT-TACGGCT- -CCGGGCAAACTGGGGGGACTGCCGGC GCCGTCGCCAGCTCGTGCCGTGAGGTGTCCTGTTAAGTCAGGCAACGAGCGAGACCCCCGCCCTCAGTTGCCAGCGGGGCCT-TACGGCTGGCCGGGCAAACTGGGGGGACTGCCGGC
AJ318041	GCCGCCGTCAGCCTGTGCCGTGAGGTGCCCTGTTAAGTCAGGAAACAGGCGAGACCCGCGCCCGCAGTTGCGA--------CGGCCGAAAGGCCG - GCACACTGCGGGGACTGCCGGG ACACGCGTGCTACAAT Mbl225F DA1204R TCCAKICRTACGGGGMTT GTGCGCSCGATGTTACC EKbl242R

Fig. 2 *(continued)*.

X03235 M36474 M35966 U51469	-GT-AAGCCGGAGGAAGGAGGGGGCCACGGCAGGTCAGCATGCCCCGAAACTCCCGGGCCGCACGCGGGTTACAATGGCAGGGACAACGGGATGCTACCTCGAAAGGGGGAGCCAATC GTTTAAGGCGGAGGAAGGAGGGGGCCACGGCAGGTCAGCATGCCCCGAACCCCCCGGGCTACACGCGGGCTACAATGGCGGGGACAGCGGGATCCGACCCCGAAAGGGGGAGGCAATC -GT- AAGCCGGAGGAAGGAGGGGGCGACGGCAGGTCAGTATGCCCCGAAACCCCGGGGCTGCACGCGAGCTGCAATGGCGGGGACAGCGGGATCCGACCCCGAAAGGGGGAGGCAATC AGTTAATGCGGAGAAAGGAAAGGGCCACGGCAGGTCAGTATGCCCCGAAACTCTGGGGCCACACGCGGGCTGCAATGGTAGTGACAATGGGTTCCATATCCGAANGGAGGAGGTAATC
X05567 D50849 M59126 U20163	GC-TAAGCCGGAGGAAGGTGCGGGCAACGGCAGGTCCGTATGCCCCGAATCCCCCGGGCTACACGCGGGCTACAATGGCCGGGACAATGGGTACCGACCCCGAAAGGGGTAGGTAATC GC- TAAAATGGAGGAAGGAATGGGCAACGGTAGGTCAGTATGCCCCGAATGGACCGGGCAACACGCGGGCTACAATGGCTATGACAGTGGGACGCAACGCCGAGAGGCGAAGCTAATC GC- TAAGCCGGAGGAAGGTGCGGGCAACGACAGGTCCGCATGCCCCGCAATCCCCTGGGCTACACGCGGGTACAATGGCCGGGACAATGGGACGGGACCGCGAAAGGGGGAGCGAATC GA- TAAGCCGGAGGAAGGGGCGGGCGACGGTAGGTCAGTATGCCCCGAAACCCCCGGGCTACACGCGCGCTACAATGGGCGGGACAATGGGACCCGACCTCGAAAGGGGAAGGGAATC
AF176347 L25852	GAA-GAGCCGGAGGAAGGAGGGGGCTACGGCAGGTCAGTATGCCCCTAATCCCCCGGGCCGCACGCGGGCTGCAATGGGCGGGACAGCGGGATGCGACCCCGAGAGGGGGAGCTAATC GAA- GAGCCGGAGGAAGGAGGGGGCTACGGCAGGTCAGTATGCCCCTAATCCCCCGGGCCGCACGCGGGCTGCAATGGGCGGGACAGCGGGATGCGACCCCGAGAGGGGGAGCAAGTC
AJ318041	GA--AACCCGGAGGAAGGTGCGGGCGACGGCAGGTATGCATGCCCCGAATGCCCCGGGC TACACGCGCGCATCAATGGGCGGGACAGGGGGCCGCGACCCCGAAAGGGGGAGCAAATC
	U1406R CGGAACATGTGTGGCGGGCAG UA1406R AACRTGWGTGGCGGGCA N1406R AACGTGAGTGGCGGGCA
X03235 M36474 M35966 U51469	CTT-AAACCCTGCCGCAGTTGGGATCGAGGGCTGAAACCCGCCCTCGTGAACGAGGAATCCCTAGTAACCGCGGGTCAACAACCCGCGGTGAATACGTCCCTGCTCCTTGCACACACCGCCCGTC CCTCAAACCCCGCCGTGGTTGGGATCGAGGGCTGCAACTCGCCCTCGTGAACGAGGAATCCCTAGTAACCGCGCGTCAACATCGCGCGGTGAATACGTCCCTGCTCCTTGCACACACCGCCCGTC CCGTAAACCCCGCCCCAGTAGGGATCGAGGGCTGCAACTCGCCCTCGTGAACGTGGAATCCCTAGTAACCGCGTGTCACCAACGCGCGGTGAATACGTCCCTGCCCCTTGCACACACCGCCCGTC CCC- AAACGCTACCACAGTTATGACTGAGGGCTGCAACTCGCCCTCACGAATCTGGAATCCCTAGTAACCGCGCGTCATTACCGCGCAGTGAATACTGCCCTGCTCCTTGCACACACCGCCCGTC
X05567 D50849 M59126 U20163	CCCTAAACCCGGTCTAACCTGGGATCGAGGGCTGCAACTCGCCCTCGTGAACCTGGAATCCGTAGTAATCGCGCCTCAAAATGGCGCGGTGAATACGTCCCTGCTCCTTGCACACACCGCCCGTC TCC- AAACGTAGTCGTAGTTCGGATTGCGGGCTGAAACCCGCCCGCATGAAGCTGGATTCGGTAGTAATCGCGTGTCAGAAGCGCGCGGTGAATACGTCCCTGCTCCTTGCACACACCGCCCGTC CCCTAAACCGGGTCGTAGTCCGGATCGAGGGCTGTAACTCGCCCTCGTGAAGCCGGAATCCGTAGTAATCGCGCCTCACCATGGCGCGGTGAATGCGTCCCTGCTCCTTGCACACACCGCCCGTC CCCTAAACCCGCCCTCAGTTCGGATCGCGGGCTGCAACTCGCCCGCGTGAAGCTGGAATCCCTAGTACCCGCGTGTCATCATCGCGCGGCGAATACGTCCCTGCTCCTTGCACACACCGCCCGTC
AF176347 L25852	CCTGAAACCCGCCCGTGGTTGGGATCGAGGGTTGCAACTCGCCCTCGTGAACCCGGAATCCCTAGTAACCGCGGTTCTCCATACCGCGGTGAATACGTCCCTGCCCCTTGTACACACCGCCCGTC CCTGAAACCCGCCCGTGGTTGGGATCGAGGGTTGCAACTCGCCCTCGTGAACCCGGAATCCCTAGTAACCGCGGTTCTCCATACCGCGGTGAATACGTCCCTGCCCCT
AJ318041	CCC-AAACCCGCTCTCAGTCCAGATCGAGGGCTGCAACTCGCCCTCGTGACGGCGGAATCTCTAGTAGTCGGACGTCACCAGCGTCCGGCGAATACGTCCCTGCTCCTTGCACTCACCGCCCGTC

Fig. 2 (*continued)*