Indian Journal of Microbiology Vol. 36, March 1996, pp 9-12.

A Rapid Method for Sequencing of rRNA Gene(s) Amplified by Polymerase Chain Reaction using an Automated DNA Sequencer

PP Dwivedi[1], BKC Patel*[1], GN Rees[2] and B Ollivier[3]

[1] Faculty of Science and Technology, Griffith University, Nathan Campus, Brisbane, Queensland - 4111, **Australia.**
[2] Microbiology Research Unit, Faculty of Applied Science, University of Canberra, Belconnen, ACT-2616, **Australia.**
[3] Laboratoire de Microbiologie ORSTOM, Universite de Provence, 13331, Marseille cedex 3, **France.**

A method for DNA sequencing of ribosomal RNA (rRNA) genes, amplified by polymerase chain reaction (PCR), using internal primers, designed on the basis of conserved regions of rRNA genes for determining a near complete sequence (99%) of the gene using an automated DNA sequencer (Applied Biosystem Incorporation, USA) is described. The procedure is extremely rapid as cloning of the gene is not required for sequence determination. In addition time consuming steps such as ethanol precipitation and hazardous steps such as phenol/chloroform extractions are excluded from the protocol for the purification of extension products after Taq cycle sequencing using the ABI dye terminator chemistry. The method has been successfully used for sequencing of the 16S and 18S rRNA genes of microbes which includes six members of domain Bacteria, one of domain Archaea and one belonging to Eukarya domain.

Key words : Polymerase chain reaction (PCR), DNA sequencing, ribosomal RNA (rRNA) genes.

Comparisons of 16S and 18S ribosomal RNA (rRNA) sequences have been useful for infering quantitative evolutionary relationship amongst numerous, diverse, cellular life forms (1, 2). For such studies ribosomal RNA gene(s), are amplified from the genome of any organism using the polymerase chain reaction (3), cloned and the nucleotide sequence determined (4-6). To facilitate cloning of rRNA gene(s), restriction enzyme sites (viz. SAI 1 and Bam H1) which occur rarely in the rRNA genes are engineered into the amplification primers. However, such infrequent sites, may be present within the gene and in such cases, incomplete sequence information is obtained. *Escherichia coli, Desulfoacinum infernum* and *Halothermothrix oreinii* (7), have been reported to have Sal I restriction enzyme site in their 16S rRNA gene. If this problem is encountered, then blunt end ligation or TA-cloning is the only way of cloning and obtaining the near complete nucleotide sequence of the 16S or 18S rRNA gene (8).

We have designed and used such amplification primers for amplifying the 16S rRNA and 18S rRNA genes of microorganisms. In addition we have also designed oligonucleotide sequencing primers for the conserved regions of these genes and have developed a method using these primers for DNA sequencing of such genes using an automated DNA sequencer from Applied Biosystem Incorporation (ABI), USA. This method avoids cloning and subsequent screening of recombinants for 16S or 18S rRNA genes. The procedure also avoids the use of phenol and chloroform extraction for separation of unincorporated dye terminators from extension products after Taq cycle sequencing reactions. The method described in present communication is very useful for rapidly obtaining the desired sequence of rRNA gene(s) as compared to other methods used to date.

Materials and Methods

Designing of amplification and sequencing primers

Amplification and sequencing primers, used for studying rRNA genes, have been described in Table 1. Of these, two bacterial and one archael primer[1] have an engineered polylinker region which contains rare restriction enzyme site useful for cloning of the 16S rRNA genes. A further ten internal sequencing primers were designed on the basis of conserved regions of the 16S and 18S rRNA gene sequences available in data base and from data available in the literature. The positions of the primers have been numbered on the basis of corresponding nucleotide numbers of the 16S rRNA gene of *Escherichia coli* and the 18S rRNA gene of *Candida albicans* (9).

*Corresponding Author

Amplification and purification of PCR product

16S and 18S rRNA gene (s) were amplified using various amounts (20 ng to 500 pg) of partially purified genomic DNA (10), 5 µl of 10 x Taq DNA polymerase buffer (500 mM KCl, 100 mM Tris-HCl pH 9, 1.75 mM magnesium chloride, 1% Triton X-100), 0.2 mM dNTPs and 50 p mole of each amplification primer (depending on the type of microbe under study, see Table-1). The total PCR volume was made upto 50 µl with sterile distilled water and overlaid with 40 µl of sterile mineral oil.

The samples were placed in a thermal cycler (Corbett Research Limited, Australia) for amplification of the rRNA genes. The denaturation step was performed by incubating the samples at 95°C for 7 min. After denaturation, samples were taken out, 2.5 U of Taq DNA polymerase (Promega Corporation, USA) was added and the samples returned to the heating block of thermal cycler. These samples were then run through 30 cycles at 41°C (or 55°C) for 2 min 72°C for 4 min, 95°C for 1 min, followed by a final cycle of 41°C for 2 min and 72°C for 20 min for the complete extension of PCR-product. The oil overlay from the PCR-product was separated using the parafilm method (11). The PCR-product was electrophoresed on 1% agarose-TAE gels, the right size band was identified, excised under UV transilluminator and purified using Qiaex DNA gel extraction protocol (Qiagen Limited, USA).

Taq cycle sequencing and purification of extension products

The samples for Taq cycle sequencing were prepared to a final volume of 20 µl and contained purified PCR-product (300-350 ng), 3.2 p mole of sequencing primer and 9.5 ul of fluorescent dye deoxy terminators for each reaction, as supplied with the PRISM Ready Reaction Dye Deoxy Terminator Cycle Sequencing kit (Applied Biosystem Incorporation). Samples were placed in a thermal cycler (Corbett Research Limited, Australia) and run through 25 cycles at 96°C for 30 sec, 50°C for 15 sec and 60°C for 4 min.

The extension products, after cycle sequencing, were separated from oil overlay using the parafilm method (11) and purified from unincorporated dye terminators using Sephadex G-50 gravity columns prepared in a home made 24 column-unit which is a modified version of previously described stationary column-unit (12). For this, the columns were packed with 600 µl of Sephadex G-50, prepared in STE buffer (50 mM NaCl, 10 mM Tris-HCl pH 7.4, 1 mM EDTA) and allowed to settle for 5 min. The oil free extension products were loaded on to the column and allowed to flow under gravity as described earlier (12). The unincorporated dye terminators were adsorbed in Sephdex G-50 column and fluorolabelled DNA was eluted with 100 µl of sterile water. The eluted fluorolabelled DNA was dried in a speedyvac concentrator to complete dryness and finally resuspended in a suitable volume of formamide and loaded on a sequencing gel of the ABI Automated DNA sequencer following the recommended protocol (13).

Results and Discussion

Manual sequencing of 16S/18S rRNA genes using labelled ^{32}P, is the most commonly used approach for phylogenetic studies (4, 14–16). In this paper we have modified some of these protocols used (eg. sequencing primers, extraction protocol for extension products of cycle sequencing reactions) and successfully used these for automated DNA sequencing of rRNA genes.

The ten sequencing primers designated F-1 to F-4 and R-1 to R-6 worked extremely well for sequencing of 16S rRNA genes from members of the domain Bacteria. However, primers F-1, F-3, R-1 and R-5 failed to sequence 16S rRNA gene of Archaea and F-2, F-3, R-3 and R-5 failed to sequence the 18S rRNA gene of domain Eukarya (Table-1) but this is to be expected as there are sequence mismatches. From the ten sequencing primers, four universal sequencing primers (F-4, R-2, R-4 and R-6) capable of sequencing rRNA genes from all the three domains have been identified. With the use of these universal primers, and the additional Archaea amplification primer ARFD-1 and primer F-2, almost complete 16S rRNA gene can be sequenced from domain Archaea. In addition, the entire 18S rRNA gene from organisms of domain Eukarya can be sequenced using FF-1 and FR-1, the amplification primers for Eukarya, F-1, R-1 and the four universal sequencing primers (F-4, R-2, R-4 and R-6) described above.

Table 1. Sequencing and amplification primers for rRNA genes

Name of Primer	Corresponding sequence in *E. coli* for AD, BD & in *C. albicans* for ED	Sequence (5' - 3')	Primers used		
			BD*	AD*	ED*
FF-1	1-21	5'-AACCTGGTTGATCCTGCCAGT-3'	nt	nt	+
FR-1	1772-1795	5'-GATCCTTCTGCAGGTTCACCTAC-3'	nt	nt	+
ARFD-1	8-28	5'-**CCGAATTCGTCGACAAC**UCCGGUUGAUCCUGCCGG(A)AG(C)-3'	nt	+	nt
Fd-1	8-27	5'-**CCGAATTCGTCGACAAC**AGACTTTGATCCTGGCTCAG-3'	+	-	-
F-1	339-357	5'-CTCCTACGGGAGGCAGCAG-3'	+	-	+
F-2	783-803	5'-CAGGATTAGATACCCTGGTAG-3'	+	+	-
F-3	907-926	5'-AAACTCAAAGGAATTGACGG-3'	+	-	-
F-4	1390-1405	5'-TGTACACACCGCCCGT-3'	+	+	+
Rd-1	1542-1525	5'-**CCCGGGATCCAAGCTT**AAGGAGGTGATCCAGCC-3'	+	-	-
R-1	357-342	5'-CTGCTGCCTCCCGTAG-3'	+	-	+
R-2	536-519	5'-GTATTACCGCGGCTGCTG-3'	+	+	+
R-3	800-783	5'-CCAGGGTATCTAATCCTG-3'	+	nt	-
R-4	926-907	5'-CCGTCAATTCCTTTGAGTTT-3'	+	+	+
R-5	1115-1100	5'-GGGGTTGCGCTCGTTG-3'	+	-	-
R-6	1513-1494	5'-TACGGTTACCTTGTTACGAC-3'	+	+	+

Fd1 and Rd1 : Amplification primers for bacteria domian (BD). ARFD-1 and R-6 : Amplification primers for Archae domain (AD). Bold Letters : Polylinker region engineered with restriction enzyme sites. ARFD-1 & FD-1 (EcoR1 & Sal I), Rd-1 (Xma I, BamHI and Hind III). F1 - F4 : Forward sequencing primers. R1 - R-6 : reverse sequencing primers. + : Primers worked for sequencing. - : Primers did not work for sequencing. nt : Not tested for sequencing. * : See table-2 for organisms for this study.

The purification of extension products of cycle sequencing away from the unincorporated dye terminators is normally carried out using phenol/ chloroform extraction method or spin columns in ABI protocols. However, we have routinely separated the extension products by using a modified Sephadex G-50 gravity columns (12). This method is extremely quick because it does not use ethanol precipitation. Ethanol precipitation is required after phenol/chloroform extraction and is not only a tedious process but may result in significant loss of extension products leading to poor quality sequence data. The gravity fed Sephadex G-50 multi-columns are more advantageous to use as compared to spin columns since these are more convenient in handling many samples in parallel. In addition, the results obtained by gravity G-50 columns are more consistent than spin columns. Using this method 24 samples are handled very efficiently, but the number can be increased by adding more columns to the multi-column unit.

The sequence data obtained using the protocol described above have allowed us to read 400 to 450 nucleotides and analysis of the data has shown that accuracy often exceeds 98%. However, sometimes signal intensities are insufficient in 400-600 bp region for reliable base callings. Comparable sequence data are also obtainable

Table 2. 16S or 18S rRNA genes studied from various domains

Name of domain	Organisms studied	Source	Reference
Eukarya	*Neocallimastrix* (strain aei)	Australia	Denman unpublished results.
Archae	*Thermococcus* sp (strain Rt 3)	New Zealand	Jones *et. al.* -do-
Bacteria	*Clostridium indicus* sp. nov. (strain Indi B4)	India	Chrisotomas *et. al.* -do-
Bacteria	*Desulfotomaculum thermosapovorans* sp. nov. (strain MLF)	France	Fardeau *et. al.* -do-
Bacteria	*Desulfoacinum infernum* gen. nov. sp. nov. (strain B αG1)	Australia	Rees *et. al.* -do-
Bacteria	Thermophilic spirochete (strain R118B1)	New Zealand	Jonsson *et. al.* -do-
Bacteria	*Desulfotomaculum* sp (strain T93B)	Norway	Unpublished results. -do-
Bacteria	*Lactobacter thermoamylovorans* gen. nov. sp. nov. (strain DKP)	France	Combet *et. al.* -do-

from the 16S rRNA genes cloned in p-bluescript (-) vector (sequence data not shown). We have been successful in sequencing various new rRNA gene(s) from Bacterial, Archael and Eukaryotic organisms but the data are yet to be published (Table 2).

The restriction enzyme site(s) described in polylinker region of amplification primers (Table 1) have been used earlier in our laboratory for the directional cloning of rRNA genes (17-20) but the occurence of rare restriction enzyme sites in the 16S rRNA genes of the microbes, e.g. *Desulfoacinum infernum* (Table 2) and *Halothermothrix oreinii* (7), provides incomplete sequence data which if available may be useful for identification of probe sites or group specific signature sequences. We do not clone 16S or 18S rRNA genes for their sequence determination, but instead routinely sequence rRNA genes from their PCR-product by the method described in this paper. The method is also useful from a health safety point of view and to inexperienced users, as it avoids the use of phenol and chloroform during sequencing manipulations. This procedure has been used for sequencing of 16S/18S genes from the limited range of organisms from domains bacteria, archae and eukarya described in Table-2. The sequencing of the genes from other phyla/subphyla is underway.

Acknowledgements

The financial assistance for these studies from Australian Research Council and Queensland Water Resource Commission is gratefully acknowledged. We thank Brenda Cheung for her assistance in running the sequencing gels on ABI seqencer.

References

1. **Olsen GJ, Lane DJ, Giovannoni SJ & Pace NR,** *Ann Rev Microbiol,* **40** (1986) 337.
2. **Woese CR,** *Microbiol Rev,* **51** (1987) 221.
3. **Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB & Erlich HA,** *Science,* **239** (1988) 487.
4. **Weisburg WG, Barns SM, Pelletier DA & Lane DJ,** *J Bacteriol,* **173** (1991) 697.
5. **Eisen JA, Smith SW & Cavanaugh CM,** *J Bacteriol,* **174** (1992) 3416.
6. **Patel BKC, Saul DS, Reeves RA, Williams LC, Cavanagh JE, Nichols PD & Bergquist PL,** *FEMS Microbiol Lett,* **115** (1994) 313.
7. **Cayol JL, Ollivier B, Patel BKC, Prensier G, Guezennec J & Gracia JL,** *Int J System Bacteriol,* **44** (1994) 534.
8. **Sambrook J, Fritsch EF & Maniatis T,** *Molecular Cloning : A Laboratory Manual,* 2nd ed. Cold Spring Harbor Laboratory Press, USA, 1989.
9. **Neefs JM, Van de Peer Y, De Rijk P, Goris A & De Wachter R,** *Nucleic acid Res,* **19** (1991) 1987.
10. **Marmur J,** *J Mol Biol,* **3** (1961) 208.
11. **Whitehouse, EH & Spears T,** *Biotechniques,* **11** (1991) CRS No.145.
12. **Rosenthal A & Charnock-Jones DS,** *J DNA Sequencing and Mapping,* **3** (1992) 61.
13. **Annonym,** *Applied Biosystem Model* 373A : Foster City, California, USA.
14. **Bottger EC,** *FEMS Microbiol Lett,* **65** (1989) 171.
15. **Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML & Pace NR,** *Proc Natl Acad Sci* USA, **82** (1994) 6955.
16. **Lane DJ, In** *Nucleic Acid Techniques in Bacterial Systematics* (Stackebrandt E & Goodfellow M eds.) John Wiley and Sons, USA, (1991) p 115.
17. **Love CA, Patel BKC, Ludwig W & Stackebrandt E,** *FEMS Microbiol Lett,* **107** (1993) 317.
18. **Love CA, Patel BKC, Nichols PD & Stackebrandt E,** *Syst Appl Microbiol,* **16** (1993) 244.
19. **Redburn AC & Patel BKC,** *FEMS Microbiol Lett,* **113** (1993) 81.
20. **Redburn AC and Patel BKC,** *FEMS Microbiol Lett,* **115** (1994) 33.