Preparation and analysis of environmental DNA: optimisation of techniques for phylogenetic analysis of ATAD sludge

A.V. Piterina and J.T. Pembroke

Laboratory of Molecular and Structural Biochemistry, Department of Chemical and Environmental Sciences, Material and Surface Science Institute (MSSI), University of Limerick, Ireland

Non-culture based analysis of microbial populations via phylogenetic analysis is becoming an increasingly important tool in the determination of microbial communities and community flux in a large number of environmental niches. The results obtained are however only as good as the preparative techniques used to isolate the original templates and the PCR processing techniques used to amplify these templates. In thermal niches such as autothermal thermophilic aerobic digestion systems (ATAD) there are many features that limit the optimal recovery of microbial diversity. Such features include lysis and release of products from the microbial population such as nucleases and proteases which effect both template recovery and polymerase functionality, inhibitory substances from the environmental source and even the nature of the primers used to amplify the resultant templates. In optimisation of the extraction and amplification of environmental DNA from an ATAD system we have identified a number of issues that affect the recovery of diversity which may have widespread applicability to the recovery of diversity by molecular techniques from other environmental niches. Methods of optimisation and analysis of diversity recovery will be discussed.

Keywords: phylogenetic analysis; optimisation of diversity recovery; ATAD.

1. Introduction

Culture based techniques have been the mainstay of microbiology since their origins in the pioneering work of Robert Koch and Louis Pasteur in the 19th centaury. However culture based techniques have many limitations in analyzing the diversity of microorganisms inhabiting a particular environmental niche. Aerobic and anaerobic organisms cannot be cultured together, fastidious organisms will often not grow because essential nutrients for growth are not present and organisms requiring specific pH, optimal temperature, optimal gas mixtures, and unknown growth factors are often not recovered in culture based screens. Traditional methodologies of cultivation are also known to be inefficient in recovering symbiotic, stationary or slow growing organisms, and the so called viable but non-cultivable fraction which are believed to make up the bulk of environmental organisms [1]. The development of DNA-based techniques which rely on analysis of DNA extracted from microorganisms inhabiting a particular niche, has revolutionized the ability to characterise and identify the diversity and taxonomy of environmental organisms in a wide variety of niches [2]. DNA based techniques have been applied to diversity studies in a range of applications such as food [3], soil [4], water [5] waste water [6] and to organisms associated with the human body [7]. A distinct advantage of the use of molecular approaches is that they allow monitoring, detection and analysis of the genetic targets of interest directly from environmental samples, without the additional steps of cultivation and recovery [2]. Many molecular studies applied to soil and water have indicated that the choice of processing method and the design of extraction protocol may effect the degree of lysis of the microorganisms present in the sample and therefore recovery of their template DNA, which are necessary for subsequent analysis [8]. In addition factors such as the integrity and size of DNA obtained and the extent to which interfering organic and inorganic impurities are present pose a problem for molecular analyses. These issues are important as subsequent analysis by polymerase chain reaction (PCR) may be problematic with sheared DNA or may not occur at all in the presence of certain impurities. Any loss of template, interference with amplification of recovered template or bias for highly concentrated template over rare templates will have a dramatic effect on diversity recovery, which will thus effect interpretation of the true diversity and taxa present [1, 3, 4].

It is thus essential, when examining any new ecological niche to carefully apply methodology to evaluate the efficiency of sample extraction and determine the heterogeneity of the amplified DNA targets obtained. Although this seems self obvious many studies looking at diversity analysis have failed to analyse factors that would affect diversity recovery in a particular setting. Among the various methods applied, denaturing gradient gel electrophoresis (DGGE) has been shown to be effective in examining diversity of PCR amplified fragments from taxonomic targets like 16S rDNA [9]. DGGE separates similar size fragments based on their migration in a denaturing chemical gradient within a polyacrylamide gel and gives a snapshot of the amplification diversity. The position of each band is determined by its unique melting properties which are a function of the nucleotide sequence and GC content of the PCR amplicon [9]. PCR products differing even by one nucleotide will take a different position in the gel lane allowing the diversity of multiple amplicons from a particular niche to be examined as a community profile. To provide precise identification and quantification of the phylotypes present in different DNA samples, amplification of a near full-length taxonomic marker, such as the 16S rDNA gene, is necessary, followed by clone library generation and sequencing of each unique clone [10]. 16S rDNA sequences obtained can then be compared against a database of 16S rDNA sequences to observe similarity and taxonomic origin [10]. A key initial factor is template amplification and here multiple strategies have

been used in an attempt to overcome inhibition of amplification of problematic DNA samples. These strategies include serial dilution of the DNA [11] and titration with $MgCl_2$, however this can reduce the resolution power of DNA-based techniques and lead to lost of amplification of rare genetic targets by dilution while giving rise to low reproducibility and reduced diversity [11]. A range of chemical additives have also been shown to be beneficial to DNA polymerase performance in PCR reactions but their applicability and effectiveness depends on the nature and origin of the DNA samples and the type of inhibitors potentially present in the samples. The choice of chemical additive used to improve amplification can be made by screening numerous compounds in the PCR reaction until successful or more diverse target amplification occurs, which can then be monitored by DGGE. Thus for each new environmental niche a number of suitable chemical additives could be screened to eliminate interference or improve amplification based on the nature of impurities expected or found associated with that niche. Such an analysis should form a key component in optimizing a molecular examination of any environmental niche.
Although a number of commercial extraction and clean-up kits are available on the market which have been applied to environmental analysis [12], they have not been optimised to new niches, a full understanding of what they do is often lacking but most importantly the extent to which the full biodiversity of the niche under examination is being explored is rarely determined. We have utilised an unusual thermal environmental niche associated with autothermal thermophilic aerobic digestion (ATAD) to explore the factors that need to be addressed to optimise the recovery of microbial diversity which offers an insight into the general factors that need optimisation in any such analysis irrespective of the environmental source.

2. Materials and Methods

2.1 Extraction and processing methods for environmental sludge

ATAD sludge was collected from the Killarney ATAD plant, Co Kerry Ireland, treating domestic sewage, the temperature in the reactor ranged between 55-65°C. The plant has previously been described [13]. Two extraction methods were utilised, a commercial soil kit (MoBIO, UK) and a solvent based method [6]. PCR primers 27^f to amplify the V1 AGAGTTTGATCCTGGCTCAG and 1472^r to amplify the V9 region of the 16S rDNA gene GGTTACCTTGTTACGACTT were used [14] The thermal cycle conditions were 10-min activation of the polymerase at 94°C, 2 cycles of 1 min at 94°C, 1 min at 53°C and 2 min at 72°C. Annealing temperature was decreased by 1°C for every second cycle until 47°C, then 30 additional cycles were carried out before; a 9-min extension at 72°C. Amplification of PCR products was confirmed by electrophoresis in 1.2% (w/v) agarose 1xTBE buffer, with EtBr staining [14]. A variety of commercial *Taq*-polymerases were utilised for comparison. The DNA Polymerases used were Ampli*Taq* Gold DNA polymerase (*Thermus flavus*), Expand High Fidelity (Expand HF) PCR system (a mixture of two DNA polymerases, *Taq* (*Thermus aquaticus*) and Pwo (*Pyrococcus woese*), *Pwo* DNA polymerases (*Pyrococcus woesei*), *rTth* DNA polymerase (*Thermus thermophilus*); *Taq* DNA polymerase (*Thermus aquaticus*) (Bio-Line); Tfl DNA polymerase (*Thermus flavus*) (Promega,UK);(Promega); Red*Taq* Genomic DNA polymerase (Sigma) and Long-target genomic DNA *Taq* DNA polymerase (TaKaRa) and Hot*Tub* DNA polymerase (*Thermus ubiquatous*). PCR reaction mixtures were subjected to hot start where it was recommended by the manufacturer. A variety of chemical additives including acetylated Bovine Serum Albumin (aBSA); PVPP; glycerol, DMSO; non-acetylated bovine serum albumin (BSA) and formamide all known to enhance amplification were utilised to optimise amplification.

2.2. Monitoring efficiency of amplification

Sludge samples were periodically spiked with pGEM-TA (Promega) to monitor shearing and nuclease activity. DNA was routinely monitored for integrity by gel electrophoresis to determine the presence of nucleases in the sample.

2.3 DGGE analysis and monitoring of diversity recovery

For DGGE analysis primers 338f GCACTCCTACGGGAGGCAGCAG and 518R ATTACCGCGGCTGCTGG annealing to the V3-V5 region of 16s rDNA were used, the cycle involved 10-min activation of the polymerase at 93°C followed by 2 cycles consisting of 1 min at 94°C, 1 min at 56°C and 2 min at 72°C. The annealing temperature was subsequently decreased by 1°C for every second cycle until it reached 50°C, at which point, 30 additional cycles were carried out and a 9-min extension at 72°C was used. DGGE was carried out in a DGGE gel system (Shaw, USA) at 60°C in 0.5×TAE buffer. Gels contained 8% (w/v) polyacrylamide (37:1 acrylamide/bis-acrylamide), 0.5×Tris/Acetic acid/EDTA buffer (TAE) with gel dimension 1.5 mm / 22 cm / 22 cm. Gradients were formed between 30% and 80% denaturant and run for 15 min at 25 V, and then voltage maintained at 75 V for 18 hours. To avoid PCR bias we utilised touchdown PCR [15], a low cycle number and replicates.

3. Results and Discussion

In utilizing a molecular based approach to determine diversity there are a number of important characteristics that must be optimised if the recovered diversity is to be meaningful. The extraction procedure must as far as possible recover all the possible templates from the sampling site. This is particularly important in the case of rare templates from organisms that may be present in low numbers but which may be important to the functioning of that site. The DNA should be of good quality, high molecular weight and be free from potential PCR inhibitors, which is essential for subsequent cloning of amplicons for library generation or comparative analysis of resulting amplicons via DGGE. Initially we compared two extraction protocols applied to ATAD sludge followed by amplification of the resultant DNA extract via DGGE analysis Figure 1. We noted that the resultant DGGE profiles differed from each other. Although there were a number of common DGGE bands present there were also a number bands missing from the commercial extraction kit that were present in the solvent extraction protocol [6] used. This only came to light following direct comparison of two different extraction protocols. The patterns were stable for each protocol and indicated that to optimise diversity recovery from a novel niche such as ATAD sludge that more than one protocol should be used and that pooling of samples may offer an important tool in the optimisation of diversity recovery.

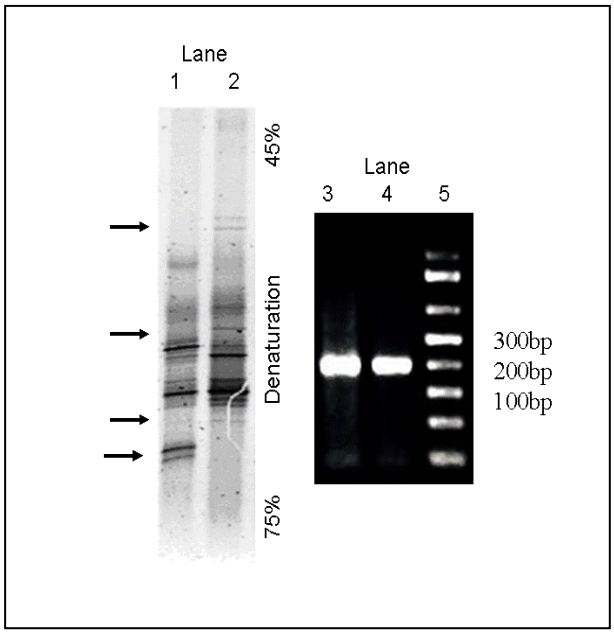

Figure 1 Comparative DGGE (45-75% denaturant) profiles of ATAD sludge DNA extracted by the solvent method DNA (lane 1) and the commercial soil extraction method (lane 2). Differences in DGGE profiles are indicated by arrows. DNA for this comparison was amplified by 16s rDNA V3-V5 primers and separated on 1.5% agarose. Lane 3 shows the PCR product of the solvent extract amplified by V3-V5 primers on non-denaturing agarose gel electrophoresis.. Lane 4 shows DNA extracted by the commercial kit amplified by V3-V5 primers on non denaturing agarose gel electrophoresis. Lane 5 shows molecular weight markers (100 bp DNA ladder) on non denaturing agarose gel electrophoresis.

In addition to methods used the site within the ATAD system analysed will also give rise to differences in DGGE profiles. ATAD sludge is initially mixed at 15°C and then aerated. The temperature rises with time in a mesophilic reactor to 45°C and in a final stage to 65°C. This increase in temperature with time should select initially for mesophilic organisms, then thermoduric and finally thermophilic populations. Figure 2 illustrates comparative DGGE profiles of DNA extracted from the ATAD reactor at different temperatures during processing. As the temperature increases there is an alteration in the DGGE profiles indicative of population shifts. This has implications for diversity analysis illustrating that in this case temperature has a major impact on the diversity profile returned. However, it illustrates that sampling is a key component and knowing the characteristics of the sampling site not just temperature but pH, chemical composition, degree of aeration may be key issues in the diversity recovered. Even slight changes in sample site characteristics can have a major impact on diversity recovery.

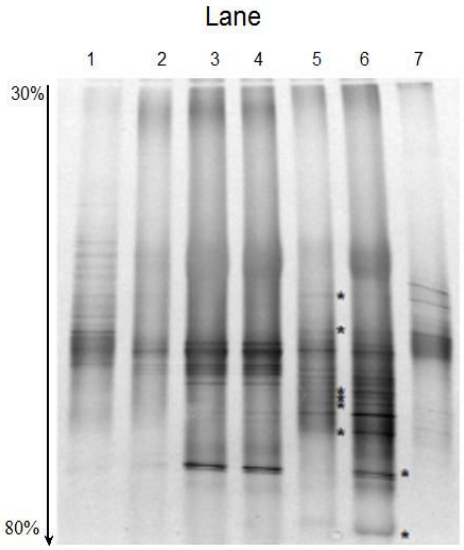

Figure 2 Comparative DGGE profiles (between 30–80% denaturant) of ATAD sludge samples amplified via 16s rDNA V6-V8 primers sampled at different operating temperatures within the ATAD process. Lane 1, 15°C, Lane 2, 30°C; Lane 3, 45°C; Lane 4, 50°C, Lane 5, 60°C; Lane 6, 65°C and Lane 7, biosolids post treatment (25°C).

During the amplification of ATAD DNA extracted from ATAD sludge at 60°C we noted that there were a number of factors that resulted in poor amplification of the extracted and pooled templates. Initially we noted a large amount of sheared DNA. We hypothesized that this might be due to nuclease action as with the succession of microbial populations as a function of process temperature it might be assumed that lysis of mesophilic and thermoduric populations would result in the liberation of nucleases. To examine this possibility we spiked the ATAD extracts with pGEM-TA plasmid vector DNA and found that the spiked DNA was degraded upon incubation at 65°C. This indicated that there were thermostable nucleases present which would be particularly problematic given that the extracted DNA was to be used for PCR where elevated temperatures are used in amplification. Various chemical additions were then utilised to examine their effect on inhibiting nuclease action. Addition of up to 5mM EDTA, some 5-folds more than used in TE buffer had little effect. Heating the ATAD DNA for 95°C for 20 minutes also had no effect indicating the presence of highly thermostable nucleases, which only become manifest at elevated temperature conditions that are necessary for subsequent PCR. Addition of 1% formamide however did result in inhibition of this ATAD associated nuclease activity and allowed recovery of spiked pGEM-TA. Formamide was subsequently added to all ATAD DNA preparations. Many environmental DNA samples such as food may contain thermoduric or thermostable organisms that release similar thermostable nucleases whose activity may only become manifest upon subsequent use of PCR conditions. Their inhibition is important in optimal diversity recovery.

ATAD DNA extracts often suffered from poor amplification by PCR. To determine the purity of recovered ATAD DNA we routinely examined its ability to undergo amplification by PCR using a number of universal primers. Lack of product was used as an indication of problems with the template. To verify that these were problems with the amplification we routinely utilised multiple samples of ATAD sludge and samples prepared by different extraction methods. In addition to nuclease activity several other factors appeared to effect amplification and these were analysed in some detail in an attempt to overcome the inhibition caused. Dilution of target DNA appeared to have a positive effect resulting in improved amplification in almost all cases. However, dilution of the DNA template can have an effect on the recovery of rare templates in the extraction mix and care should be utilised applying dilution as a first tool to improve amplification. We also noted that different universal primer sets were more efficient in amplification. Primer set 25f-1497r appeared to be least effected compared to the other primer sets utilised indicating that there may be factors in ATAD sludge and indeed in other environmental samples that effect the affinity of primers for target during PCR and that these warrant investigation when examining a new environmental niche. ATAD sludge like many environmental samples is a complex mixture of not only complex organic materials such as carbohydrates and humic substances but also synthetic materials that contaminate the sludge via being present in domestic waste. For this reason we analysed a number of potential amplification additives (Table 1), reported to enhance PCR amplification in an attempt to improve amplification of ATAD sludge extracted DNA.

Table 1. Additives used in PCR to optimise amplicon recovery [16, 17, 18]

Additive	**Role**
PVPP	Binder of polar molecules, contaminants in natural extracts
DMSO	Reduces DNA secondary structure, facilitates strand separation. Can effect *Taq* polymerase activity at high levels
Formamide	Enhances *Taq* activity and may possess DN*ase* inhibitory activity.
BSA	May bind color compounds e.g. melanin, bile salts, humic substances which inhibit polymerase. May bind anions and act as substrate for proteases.
Betaine	Iso-stabilizing agent, equalizes the contribution of GC- and AT-base pairing to the stability of the DNA duplex
DTT	Effect secondary structure allowing enzymatic processing.
$MgCl_2$	Effects primer template interaction and non specific annealing
Glycerol	Enhanced association between enzyme and template.
Non-ionic detergents Triton X100/Nonidet P-40	Suppresses formation of secondary structure, stabilises *Taq*-polymerase

Many commercial PCR reaction mixes have some of these compounds added but generally the compositions are proprietary. In some instances nuclear magnetic resonance analyses has indicated that compounds such as betaine are present [16] however mostly the compositions remain unknown. It may therefore be important to analyse the conditions present in niches being examined to identify potential inhibitory components as in ATAD sludge [13], which can inform the combinations of additives that may be necessary. This must then be followed by analysis to determine if such additives actually enhance the PCR reaction. Using this approach with ATAD DNA we found that addition of formamide and BSA had a dramatic effect (Fig. 4) on improving the amplification of ATAD sludge DNA by PCR. It is clear, that certain additives had an enhancement effect on the amplification where as others had no apparent effect (Figure 3). Therefore it is essential that when analysing amplification of DNA from different sources that the optimal additive mix be determined to optimise recovery of all templates. Another factor that may be important in amplification from templates from thermal niches is the nature of the DNA polymerase used. We utilised a number of commercial polymerases following optimisation of extraction conditions. Each polymerase used was optimised in terms of additives and reaction conditions and each performed well. However, we noted that certain polymerases were more robust in the absence of additives. It is suggested that the choice of polymerase should be examined and optimised to enhance the recovery of amplicons and not necessarily taken for granted as is often the case.

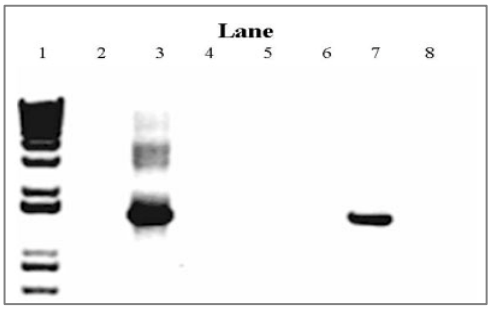

Figure 3. Negative image of EtBr stained agarose gel electrophoresis of amplified ATAD sludge DNA extracted and amplified using 27f-1491r primers and *Taq*-polymerase. Lane 1, MW marker, Lane 2,,1% formamide, Lane 3, BSA and 1% Formamide; Lane 4m 1%w/v PVPP, Lane 5, 10% v/v glycerol, Lane 6, 5% DMSO, Lane 7, 50 ng BSA, Lane 7, 50ng BSA; Lane 8, 200ng ATAD DNA no additive

A comparative profile (Figure 4) of DGGE banding patterns obtained for the total ATAD bacterial community (Figure 4, Lane 4). by direct DNA extraction and the bacterial community (Figure 4, Lanes 1, 2, 3 ,5, 6 & 7) recovered on different types of agar by culture techniques was carried out to show the utility of the molecular approach. These patterns were found to be very different from each other and from total community profiles generated from direct extraction of community DNA (Figure 4), indicating that the culturable community was not representative of the total ATAD community and would misrepresent the true microbial diversity. Thus there are a number of factors that are important in optimisation of amplicon recovery from thermal niches, such as ATAD, that need to be examined to insure optimal diversity recovery when applying molecular culture-independent techniques for microbial examination. It should not be taken for granted that applying standard extraction and amplification protocols will be sufficient. Unless techniques are specifically optimised the diversity recovered will be compromised.

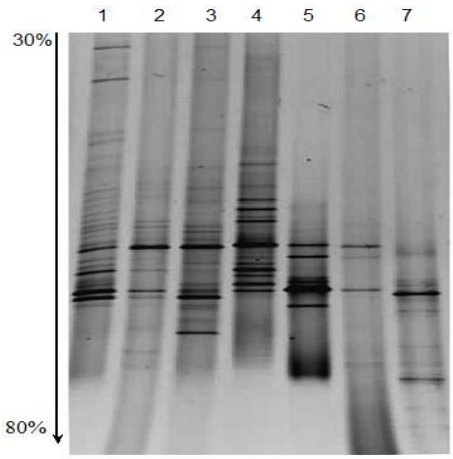

Figure 4. DGGE analysis of bacterial diversity recovered on different microbiological media (Lanes 1, 2, 3, 5, 6&7) compared to the ATAD community diversity recovered by amplification of DNA extracted by direct extraction from ATAD sludge at the same sampling point (Lane 4).

The techniques applied here to optimise diversity recovery to the ATAD thermal niche have general applicability to the examination of any environmental niche.

References

[1] Colwell R, Grimes J. eds. *Non-culturable Microorganisms in the Environment.* Washington. DC: ASM; 2000**.**
[2] Nocker A, Burr M, Camper A.K. Genotypic microbial community profiling: a critical technical review. *Microbial. Ecology*. 2007; 54:276-893.
[3] Giraffa G, Neviani E. DNA-based, culture-independent strategies for evaluating microbial communities in food-associated ecosystems. *International Journal of Food Microbiology*, 2001; 20:19-34.
[4] Ranjard L, Poly F, Nazaret S. Monitoring complex bacterial communities using culture-independent molecular techniques: application to soil environment. *Research in Microbiology* 2000;51:167-177.
[5] Kawai M, Matsutera E. 16S ribosomal DNA-based analysis of bacterial diversity in purified water used in pharmaceutical manufacturing processes by PCR and denaturing gradient gel electrophoresis. *Applied Environmental Microbiology* 2000;68:699-704.
[6] Piterina AV, Barlett J, Pembroke JT. Molecular Analysis of Bacterial Community DNA in Sludge Undergoing Autothermal Thermophilic Aerobic Digestion (ATAD): Pitfalls and Improved Methodology to Enhance Diversity Recovery. *Diversity* 2010; 2:505-526.
[7] Zhou X, Bent SJ, Schneider MG, Davis CC, Islam MR, Forney LJ. Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. *Microbiology* 2004;150:2565-2573.
[8] Frostegard A, Courtois S, Ramisse V, Clerc S, Bernillon D, Le Gall F, Jeannin P, Nesme X, Simonet P Quantification of bias related to the extraction of DNA directly from soils. *Applied Environmental Microbiology* 1999;65:5409-5420.
[9] Muyzer G, de Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied Environmental Microbiology* 1993;59:695-700.
[10] Relman A. Universal bacterial 16S rDNA amplification and sequencing. In: Persing DH, Smith TF, Tenover FC, White TJ eds *Diagnostic Molecular Microbiology, Principles and Applications.* Washington, DC:ASM 1993;489–495.
[11] Chandler DP, Frederickson JK, Brockman J. Effect of PCR template concentration on the composition and distribution of total community 16S rDNA clone libraries. *Molecular Ecology* 1997;6:475–482.
[12] Klammer S, Mondini C, Insam H. Microbial community fingerprint of compost stored under different condition. *Annual Review Microbiology* 2005;55:299-305.
[13] Piterina AV, Barlett J, Pembroke JT. ^{13}C-NMR Assessment of the pattern of organic matter transformation during domestic wastewater treatment by autothermal aerobic digestion (ATAD). *International Journal of Environmental Research and Public Health* 2009;6:2288-2306.
[14] Lane DJ. 16S/23S rRNA Sequencing. In: Stackebrandt E, Goodfellow M. eds. *Nucleic acid techniques in bacterial systematics*. New York, NY: John Wiley & Sons, 1991;115-175.
[15] Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS. "Touchdown" PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* 1991;19:4008.
[16] Frackman S, Kobs G, Simpson D, Storts D. Betaine and DMSO: Enhancing Agents for PCR *Promega Notes* 1998;65:27.
[17] Kreader CA. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Applied Environmental Microbiology* 1996;62 :1102-1106.
[18] Sarkar G Kapelner S, Sommer SS. Formamide can dramatically improve the specificity of PCR. *Nucleic Acids Research* 1990;18:7465.

