Environmental Microbiology (2000) **2**(1), 39–50

# An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics

**A. Mark Osborn,*† Edward R. B. Moore and Kenneth N. Timmis**
*Department of Environmental Microbiology, GBF-National Research Centre for Biotechnology, Braunschweig, D-38124 Germany.*

## Summary

**A systematic evaluation of the value and potential of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure has been undertaken. The reproducibility and robustness of the method has been assessed using environmental DNA samples isolated directly from PCB-polluted or pristine soil, and subsequent polymerase chain reaction (PCR) amplification of total community 16S rDNA. An initial investigation to assess the variability both within and between different polyacrylamide gel electrophoresis (PAGE) runs showed that almost identical community profiles were consistently produced from the same sample. Similarly, very little variability was observed as a result of variation between replicate restriction digestions, PCR amplifications or between replicate DNA isolations. Decreasing concentrations of template DNA produced a decline in both the complexity and the intensity of fragments present in the community profile, with no additional fragments detected in the higher dilutions that were not already present when more original template DNA was used. Reducing the number of cycles of PCR produced similar results. The greatest variation between profiles generated from the same DNA sample was produced using different *Taq* DNA polymerases, while lower levels of variability were found between PCR products that had been produced using different annealing temperatures. Incomplete digestion by the restriction enzyme may, as a result of the generation of partially digested fragments, lead to an overestimation of the overall diversity within a community. The results obtained indicate that, once standardized, T-RFLP analysis is a highly reproducible and robust technique that yields high-quality fingerprints consisting of fragments of precise sizes, which, in principle, could be phylogenetically assigned, once an appropriate database is constructed.**

Received 16 June, 1999; revised 16 October, 1999; accepted 20 October, 1999. †Present address: Department of Biological Sciences, Essex University, Wivenhoe Park, Colchester CO4 3SQ, UK. *For correspondence. E-mail osborn@essex.ac.uk; Tel. (+44) 01206 87 3763; Fax (+44) 01206 87 2592.

© 2000 Blackwell Science Ltd

## Introduction

Investigations of natural microbial assemblages and their activities have increasingly necessitated the development and use of culture-independent methods, such as those involving polymerase chain reaction (PCR)-amplified rRNA genes (reviewed by Head *et al.*, 1998), as the majority of microorganisms in the environment cannot be cultivated on standard laboratory media. Whereas precise phylogenetic information on the dominant members of microbial consortia can be generated by the sequencing of cloned PCR products, the effort and cost involved hinders the analysis of multiple samples. This in turn hinders the investigation of changes in community structure in relation to changing physicochemical and/or biological conditions over space and time, which is essential for structure–function analyses aimed at shedding light on community functioning. Rapid profiling procedures, such as denaturing gradient gel electrophoresis (DGGE) (Muyzer *et al.*, 1993), temperature gradient gel electrophoresis (TGGE) (Felske *et al.*, 1997; Heuer *et al.*, 1997) and single-strand conformation polymorphism (SSCP) analysis (Lee *et al.*, 1996; Schwieger and Tebbe, 1998), all allow the analysis of multiple samples, but the community fingerprints they generate do not directly translate into taxonomic information. There should therefore be considerable interest in developing rapid community profiling methods, which provide such information. Theoretically, use of the resolution of procedures employed in DNA sequencing to create 16S rRNA gene fragment profiles could provide data of sufficient quality for use in database interrogation to obtain taxonomic information directly. As such methods are, or can be, automated, they are ideal for the processing of many samples in a reproducible fashion.

An increasing number of new methodologies incorporate automated sequencing systems for laser detection of fluorescently labelled DNA fragments: terminal-restriction fragment length polymorphism (T-RFLP) analysis

(Liu *et al*., 1997), length heterogeneity-PCR (LH-PCR) (Suzuki *et al*., 1998) and, most recently, the combination of randomly amplified polymorphic DNA (RAPD) and principal component analysis (Wikström *et al*., 1999). The high resolution of these methods, and the ability to include an internal size standard in each sample, should endow them with the necessary degree of reproducibility. In this report, we present an assessment of one of these methods, T-RFLP, which we consider to be potentially promising for microbial community analyses. Originally developed to identify Mycobacteria by Avaniss-Aghajani *et al*. (1996), the potential of T-RFLP to analyse variation between amplified 16S rRNA genes from different bacteria and obtain information on microbial community structure was shown by Liu *et al*. (1997) and Clement *et al*. (1998). Bruce (1997) exploited T-RFLP to analyse the diversity of mercury resistance genes in polluted soils. More recently, T-RFLP analysis has been used to study bacterial (Liu *et al*., 1998), archaeal (Van der Maarel *et al*., 1998) and eukaryal (Marsh *et al*., 1998) populations in natural habitats.

An ideal methodology will allow investigation of a microbial community in terms of both its qualitative and its quantitative composition. How realistic this is will depend upon a number of factors, including the detection limits, the degree of reproducibility offered by the particular method and the specific biological question to be addressed (e.g. determination of the overall diversity of a particular environment, identification of the predominant members of a community, monitoring population dynamics, etc.). Both Liu *et al*. (1997) and Bruce (1997) discussed the possibility that T-RFLP analysis may permit at least a semi-quantitative analysis of the relative proportions of dominant members/genotypes within a microbial community, while cautioning that T-RFLP analysis is subject to all the biases inherent in any PCR amplification approach. Even in simple artificial microbial communities, a bias in the amplification of particular sequences, caused by preferential annealing to particular primer pairs (Suzuki and Giovannoni, 1996), or an increase in the incidence of chimeric PCR products with increasing numbers of PCR cycles (Wang and Wang, 1996, 1997) reduces the quality of quantitative predictions. It is therefore critical to reduce to a minimum and clearly define the amount of experimental variability and/or bias inherent in PCR-based (and other) methods. While some of the pitfalls to be associated with such methods have been mentioned by individual investigators and, for DGGE and TGGE, collated in a review by Muyzer and Smalla (1998), we are not aware of any systematic evaluation of any of the current profiling methods. In this study, we present an evaluation of T-RFLP profiling, using DNA extracted from natural environmental samples, and investigate some of the factors believed to be important in sample handling (use of replicates, pooling of samples, dilution effects, choice of polymerase and reaction annealing temperature), which may, or may not, have a critical bearing on the final profile of the microbial community.

## Results and discussion

### *Experimental considerations*

T-RFLP analysis is based on the restriction endonuclease digestion (normally with 4 bp cutters) of fluorescently end-labelled PCR products. Either one or both primers used in the PCR can be labelled and, if both, each can be labelled with a different fluorescent dye. The digested product is mixed with a DNA size standard, itself labelled with a distinct fluorescent dye, and the fragments are then separated by electrophoresis using either gel- or capillary-based systems, with laser detection of the labelled fragments using an automated analyser (in our case, Applied Biosystems). Upon analysis, only the terminal, end-labelled restriction fragments are detected. The output from such an analysis is in two forms. First, an electropherogram is produced, which shows the profile of a microbial community as a series of coloured peaks of varying heights (Fig. 1). One set of coloured peaks, in this case red, shows the internal size standards, while two other colours show the two sets of terminal restriction fragments (T-RFs) of the digested product i.e. blue for the 5′ end and green for the 3′ end.

The profiles generated by T-RFLP can vary in two ways. First, there can be variation in the number and size (in basepairs) of T-RFs present in the profile. For example, the profiles from two different soils, CD2 and A (Fig. 1A and D), are clearly different. Secondly, variation can be found in the height (and consequently the area) of any particular peak (as will become clear, this variation can have a major influence in estimates of the biodiversity represented in the numerically rarer members of a community). Such variation is clearly seen when comparing the heights of either blue or green peaks in any of the four profiles in Fig. 1. The height of each peak should provide a measure of the relative proportion of each component of a population, although biases caused by preferential annealing of the primer to certain templates (Suzuki and Giovannoni, 1996) means that the absolute values should be treated with caution. It is also possible that some T-RFs will be common in size to many different organisms, which may mean that these particular peaks are more likely to be found in a profile. This is particularly true for T-RFs generated from the 3′ end of the 16S rRNA gene. It is likely that the greater discrimination provided by the 5′ region of the gene (i.e. an increase in number of T-RFs) is a consequence of the length heterogeneities at the 5′ end of the gene, within the V1, V2 and V3 regions (Suzuki *et al*., 1998). Such variation is shown in Fig. 1, in which the

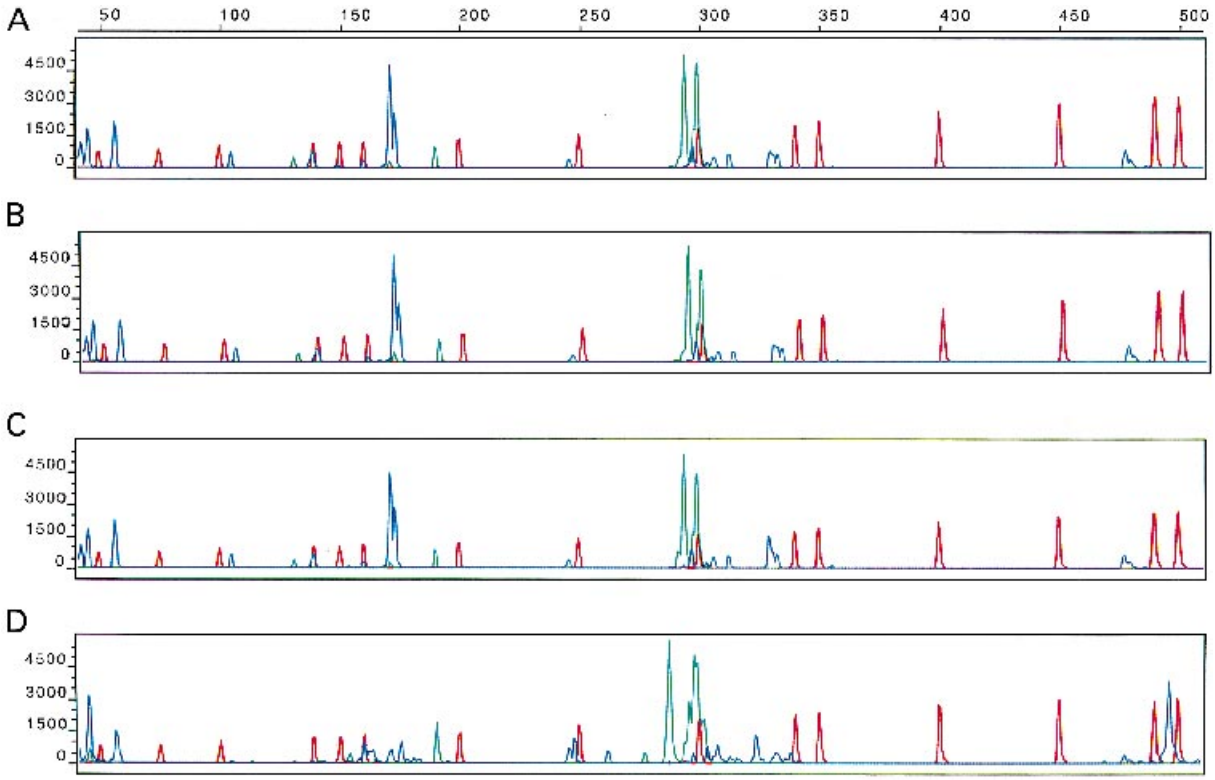


**Fig. 1.** T-RFLP profiles from *Hha*I analysis of 16S rRNA gene PCR products amplified from DNA isolated directly from polluted (CD2) and unpolluted (A) soils.
A. Soil CD2 (DNA isolation a).
B. Soil CD2 (DNA isolation b).
C. Soil CD2 (DNA isolation c).
D. Soil A.
Fragment size in base pairs is shown at the top, while peak heights are shown as fluorescent units detected. Red peaks show the GS500 DNA internal size standard, while 5′ and 3′ terminal restriction fragments are shown as blue and green peaks respectively.

number of blue 5′ T-RFs is clearly greater than the number of green 3′ T-RFs. Correspondingly, the peak heights of the smaller number of 3′ T-RFs tends to be greater, on average, than those of the more numerous 5′ T-RFs.

The second output generated from the analysis program is numerical and consists of a table, which includes, most importantly, the size, in basepairs, of each of the peaks (calculated by reference to the internal standard) and the height of each peak (relative to the amount of fluorescence detected) (Table 1). When analysing any particular profile, a minimum threshold of fluorescence is first defined to exclude background noise. In the present study, the minimum peak height was set at 100 fluorescent units.

**Table 1.** Example of numerical T-RFLP output of 5′ fragments generated by *Hha*I digestion of 16S rRNA gene PCR products amplified from DNA isolated directly from soil CD2a.

| Dye/sample peak[a] | Minutes[b] | Size[c] | Peak height[d] | Peak area[e] | Data point[f] |
|---|---|---|---|---|---|
| B, 5 | 75.43 | 138.86 | 833 | 6810 | 1460 |
| B, 6 | 83.55 | 159.88 | 357 | 2752 | 1617 |
| B, 7 | 87.21 | 168.62 | 110 | 880 | 1688 |
| B, 8 | 88.19 | 170.96 | 4833 | 39 009 | 1707 |
| B, 9 | 89.07 | 173.05 | 2633 | 22 453 | 1724 |
| B, 10 | 119.97 | 245.93 | 408 | 4124 | 2322 |

**a.** B indicates blue, i.e. 5′ fragments. 5–10 represents the number of each fragment in the profile.
**b.** Time in minutes of migration of the fragment during electrophoresis before it reaches the detector.
**c.** Size of fragment in basepairs.
**d.** Peak height is given as fluorescent units.
**e.** Peak area is a function of the peak height and the relative spread of each peak.
**f.** Each peak is assigned a data point.

**Table 2.** Variation in sizes of T-RFs as estimated by different size calling algorithms.

| | Observed fragment size by size calling method[b] | | | | |
|---|---|---|---|---|---|
| Expected fragment size[a] | 2nd Order | 3rd Order | Cubic Spline | Local Southern | Global Southern |
| 42 | 46.42 | 40.72 | 39.68 | 39.73 | **41.75** |
| 61 | 63.60 | **60.95** | 61.59 | **61.49** | 60.35 |
| 108 | 104.56 | 107.19 | **108.13** | **108.15** | 104.86 |
| 143 | 137.99 | **142.80** | **143.20** | **143.19** | 140.22 |
| 173 | 168.46 | 173.66 | **173.08** | **173.28** | 172.07 |
| 199 | 194.16 | **198.83** | 197.77 | 197.78 | **198.74** |
| 241 | 239.47 | **241.70** | **241.04** | **240.93** | 244.66 |

**a.** The expected size in bp of the 5′ T-RF products that would be generated after *Alu*I digestion of amplification products generated using FAM63F and 1389R were calculated from the DNA sequence of a series of cloned 16S rRNA genes.
**b.** The observed sizes in bp of 5′ T-RFs were calculated using the five size calling algorithms in the GENESCAN 3.1 software, after *Alu*I digestion of amplification products generated using FAM63F and HEX1389R from the same series of cloned 16S rRNA genes.
Numbers in bold indicate that the observed fragment size was the same (± 0.5 bp) as would be expected from the DNA sequence.

The calculation of T-RF size after comparison with the internal standard can be estimated using five different algorithms on the GENESCAN software. To determine which algorithm results in the most accurate estimation of the actual T-RF size, the size of a number of 5′ T-RFs was calculated using the different size calling algorithms, after *Alu*I digestion of amplification products generated using FAM63F and HEX1389 from a series of cloned 16S rRNA genes. These T-RF sizes were then compared with the expected T-RF sizes calculated from the DNA sequences (A. M. Osborn, unpublished) of the same cloned 16S rRNA genes (Table 2). Two size calling methods (2nd Order and Global Southern) resulted in estimated T-RF sizes that differed in the majority of cases from the predicted T-RF sizes. T-RF sizes estimated using the other three size calling algorithms showed good agreement with the predicted T-RF sizes, with only the smaller T-RFs, e.g. 42 bp, showing variation between experimentally derived and predicted values of greater than 1 bp. On the basis of these experiments, the Local Southern size calling method was adopted as providing the best estimate of experimentally derived T-RF size under the conditions used in this study. Our experience with other systems would suggest that the choice of size calling algorithm should be determined empirically for each set of running conditions or type of equipment used.

To assess the reproducibility and potential of T-RFLP analysis for the study of complex microbial communities, we analysed real environmental DNA samples, rather than artificial mixtures of bacterial isolates. The DNAs used in this study were extracted from samples of PCB-contaminated sandy soil (4.7 g of PCB $kg^{-1}$; soil samples CD2a, b and c) near Wittenburg, Germany, and from a soil 5 m distant, containing only trace levels of PCB (2.4 mg of PCB $kg^{-1}$; soil sample A). The structure of the metabolically active component of the microbial community in the polluted area has been previously analysed by sequence analysis of a reverse transcriptase (RT)–PCR-generated library (Nogales *et al*., 1999) and constitutes a useful resource for comparison with the results from T-RFLP analysis.

In our assessment of T-RFLP analysis, we have tried to define the extent of variability in profiles that are generated from replicate DNA isolations, replicate PCRs, replicate digests and the variation observed both within and between gels when running the same sample. To this end, in order to reduce as much as possible any variation resulting from pipetting inaccuracies, master mixes of reagents and enzymes were used for all PCRs, restriction digestions and denaturation steps. In addition to assessing the reproducibility of the system, we have also investigated and evaluated a number of other factors, such as the concentration of template DNA, primer annealing temperature, number of PCR cycles, variation between *Taq* polymerases or the concentration of restriction endonuclease, which may be critical in the generation of microbial community profiles by PCR-based methodologies.

### *Rationale in primer and enzyme choice*

The oligonucleotide primers 63F and 1389R (*Escherichia coli* numbering) (Brosius *et al*., 1978) were chosen for PCR amplification of 16S rDNA before endonuclease digestion. This primer pair was chosen for a number of reasons. First, based on systematic testing with a variety of bacterial species, Marchesi *et al*. (1998) suggested that the primers 63F and 1387R were more useful for the amplification of 16S rRNA genes for ecological studies than previously described primer pairs. In our study, instead of 1387R, we used 1389R to avoid a potential mismatch at the 3′ end of the primer (position 1388), which occurs across a range of organisms (Marchesi *et al*., 1998). Secondly, the primer pair 63F and 1389R amplifies most of the 16S rRNA gene, including all but one of the nine highly variable regions identified in the 16S rRNA molecule by Van de Peer *et al*. (1996). Finally, the positions of 63F

and 1389R, which lie internal to the primer sites 27F and 1492R/1525R that are commonly used for amplification of nearly complete 16S rRNA genes, facilitates the *in silico* prediction (Moyer *et al*., 1996) of the expected T-RFs for most full-length PCR-amplified 16S rRNA gene sequences in the DNA databases. This may ultimately allow the taxonomic ordering of the predominant members within a total microbial community profile.

Restriction endonucleases for use in T-RFLP analysis of PCR products generated using the primers 63F and 1389R were selected following computer-based predictions of the expected T-RFs from a range of 16S rRNA gene sequences from representatives of the different phylogenetic groupings of the domain Bacteria. Of the nine 4 bp cutters investigated, some enzymes, e.g. *Taq*I and *Rsa*I, result in very little variation in the predicted sizes of T-RFs using this primer pair. The three enzymes, *Alu*I, *Hha*I and *Tha*I, were found to provide the greatest discrimination, in terms of the number of different predicted T-RFs that would be produced from different species. *Alu*I and *Hha*I were subsequently chosen for these studies on the basis of enzyme cost.

### *Reproducibility*

One important parameter in the comparison of microbial community profiles generated by any profiling methodology is the degree of reproducibility between replicate samples run both on the same electrophoresis gel (intragel variation) and between different gels (intergel variation). A major advantage of fluorescent-based analysis is the presence of an internal standard in each sample, which can be used to quantify variation, not only in terms of the size, in basepairs, of terminal restriction fragments, but also in terms of the relative proportions of each fragment in a community profile. Any attempt to exploit this methodology in a semi-quantitative manner, as suggested by Liu *et al*. (1997), would be dependent upon being able to eliminate or quantitatively define any variation seen between replicate samples.

To investigate intragel variation, a *Hha*I digestion of a PCR product amplified from a 1 μl aliquot of a $10^{-1}$ dilution of DNA (≈ 16.5 ng of DNA) extracted from soil CD2a was electrophoresed, in triplicate, on the same gel. Intergel variation was investigated by electrophoresis of the same sample on three different gels. Twenty-five (19 5′ and six 3′) T-RFs were common to all the profiles. Only one or two additional T-RFs were observed for any of the replicate runs (Table 3). Of the T-RFs that were present, in addition to the 25 common fragments, one was a 5′ digestion product (169 bp) and the other was a 3′ product (289 bp) and, when present in the profile, they had very low peak heights (between 100 and 155 units) (see Table 4; and also Table 1, peak B7). If there is variation in either loading or the detection system, then the T-RFs that are most likely to vary in their presence or absence in any particular profile will be those with the smallest peak heights. Consistent with this is the finding that all of the T-RFs that have significant peak heights are found in all of the replicate profiles.

We therefore attempted to assess what extent of variation in peak height occurs within and between gels. Table 4 shows the percentage standard deviation from the mean peak height between T-RFs from the replicate T-RFLP profiles electrophoresed on the same gel. For the majority of T-RFs, this variation is less than 11%, although smaller T-RFs show slightly greater variation in peak height. Similar levels of variation were seen between replicate T-RFLP

**Table 3.** Reproducibility of T-RFLP profiling.

| Source of variation | T-RF type | Number of T-RFs (replicate no.) 1 | 2 | 3 | Number of common T-RFs |
|---|---|---|---|---|---|
| Intragel | 5′ | 20 | 19 | 20 | 19 |
| | 3′ | 7 | 6 | 6 | 6 |
| | Total | 27 | 25 | 26 | 25 |
| Intergel | 5′ | 20 | 20 | 19 | 19 |
| | 3′ | 7 | 6 | 7 | 6 |
| | Total | 27 | 26 | 26 | 25 |
| Digests | 5′ | 19 | 19 | 21 | 19 |
| | 3′ | 6 | 7 | 6 | 6 |
| | Total | 25 | 26 | 27 | 25 |
| PCR | 5′ | 20 | 21 | 19 | 19 |
| | 3′ | 7 | 7 | 6 | 6 |
| | Total | 27 | 28 | 25 | 25 |
| DNA isolation | 5′ | 20 | 20 | 20 | 20 |
| | 3′ | 7 | 7 | 8 | 7 |
| | Total | 27 | 27 | 28 | 27 |

T-RFLP patterns were generated by *Hha*I digestion of PCR products amplified from DNA isolated from soil CD2. The number of T-RFs from each replicate profile and the number of common T-RFs are given.

**Table 4.** Intragel variation in T-RF peak heights between replicate T-RFLP profiles.

| T-RF | Peak height (replicate no.) 1 | 2 | 3 | Average peak height | SD (as %) |
|---|---|---|---|---|---|
| *Hha*I 5′ | | | | | |
| 43 | 950 | 1012 | 1235 | 1065.7 ± 149.9 | 14.0 |
| 46 | 1342 | 1524 | 1915 | 1593.7 ± 292.8 | 18.3 |
| 57 | 1873 | 1935 | 2357 | 2055 ± 263.3 | 12.8 |
| 105 | 497 | 485 | 586 | 522.7 ± 55.2 | 10.6 |
| 139 | 566 | 543 | 658 | 589 ± 60.9 | 10.3 |
| 160 | 269 | 250 | 297 | 272 ± 23.6 | 8.7 |
| 169 | 117 | | 101 | | |
| 171 | 4393 | 4211 | 4584 | 4396 ± 186.5 | 4.2 |
| 173 | 2047 | 1954 | 2329 | 2110 ± 195.3 | 9.3 |
| 246 | 273 | 257 | 305 | 278 ± 24.4 | 8.8 |
| 296 | 135 | 130 | 154 | 139.7 ± 12.7 | 9.0 |
| 297 | 807 | 782 | 945 | 844.7 ± 87.8 | 10.4 |
| 304 | 198 | 189 | 229 | 205.3 ± 21.0 | 10.2 |
| 307 | 378 | 357 | 437 | 390.7 ± 41.5 | 10.6 |
| 313 | 381 | 359 | 430 | 390 ± 36.3 | 9.3 |
| 330 | 736 | 669 | 779 | 728 ± 55.4 | 7.6 |
| 331 | 551 | 509 | 595 | 551.7 ± 43.0 | 7.8 |
| 333 | 532 | 493 | 572 | 532.3 ± 39.5 | 7.4 |
| 478 | 606 | 570 | 696 | 624 ± 64.9 | 10.4 |
| 480 | 265 | 246 | 301 | 270.7 ± 27.9 | 10.3 |
| *Hha*I 3′ | | | | | |
| 131 | 307 | 305 | 368 | 326.7 ± 35.8 | 11.0 |
| 190 | 706 | 685 | 848 | 746.3 ± 88.7 | 11.9 |
| 289 | 129 | | | | |
| 292 | 381 | 369 | 451 | 400.3 ± 44.2 | 11.0 |
| 294 | 4603 | 4579 | 5297 | 4826.3 ± 407.8 | 8.4 |
| 298 | 1281 | 1268 | 1535 | 1261.3 ± 150.5 | 11.1 |
| 299 | 3450 | 3385 | 4137 | 3657.3 ± 416.7 | 11.4 |

A *Hha*I-digested PCR product, amplified using FAM63F and HEX1389R from DNA isolated from soil CD2, was electrophoresed in triplicate on the same gel.

profiles electrophoresed on different gels (data not shown). This variation observed in peak height will generally be responsible for the presence or absence (i.e. detection) of minor T-RFs in any given community profile.

Probably the greatest source of variation is the result of uneven sample loading into wells of a gel. This loading variation is equally important for other profiling methods, such as TGGE and DGGE, but it is more difficult to quantify for these systems. Moreover, the detection systems for DGGE and TGGE analyses (e.g. ethidium bromide, SYBR Green I or silver staining) are not as sensitive as laser detection of fluorescent dyes and, as a consequence, estimates of any variation for these methods will be less precise. In preliminary studies carried out using a capillary-based electrophoresis system (ABI Prism 310) with automated sample loading, the degree of variation between replicate runs of the same sample was lower than that experienced with manual loading of acrylamide gels (data not shown), suggesting that this technology will provide even more reproducible profiles. The observed variation in peak height and, consequently, in peak area is of considerable importance if the relative proportions of particular T-RFs are to be compared, as was attempted by Bruce (1997) in his study of *mer* gene diversity. Unless the extent of variation between replicate samples is quantified, conclusions about relative gene dosage within complex communities should be treated with caution. Moreover, when investigating quantitative relationships, it is important to consider the effects of inherent biases during the PCR amplification stage (Suzuki and Giovannoni, 1996) and variations in the copy number of the gene in different bacterial species (e.g. 16S rRNA gene copy number can range between 1 and 14; Cole and Girons, 1994), while genes carried on plasmids or transposons can vary widely in their copy number, even between individual strains of the same species.

Having established that the inherent variation resulting from loading and/or the detection system is low and, more importantly, limited to the presence/absence of minor members of the profile, the variability generated during three stages of the T-RFLP protocol was investigated: (i) variation between different DNA isolations from the same soil sample; (ii) variation between different PCRs from a single DNA sample; and (iii) variation between different digestions of the product from a single PCR from a single DNA sample. Again, variability between samples was assessed from three replicates in terms of the variation observed in the number (Table 3) and

heights of peaks generated after *Hha*I digestion. As with the previous experiments, 25 T-RFs were common to all the profiles. Replicate digests of the same PCR product showed no more variation in peak heights than was seen for intra- or intergel variation, suggesting that variation between digestions was not important. The greatest variability observed in the three experiments was in the number of T-RFs present and the variation in peak heights seen between replicate PCRs. It is not possible to say exactly how much of this variation was caused by intragel variation, and what component was attributable to variation between PCRs. As with the experiments investigating intra- and intergel variation, it is only the T-RFs with small peak heights that vary in their presence/absence between profiles. When three replicate PCR products were pooled before restriction digestion, the resulting profile was very similar to those from the individual PCRs (data not shown). Therefore, at least for this environment, it is not essential to use pooled replicate PCRs when conducting T-RFLP analysis, at least when studying the predominant members of a microbial community. Consequently, this may facilitate more rapid and less expensive analysis of multiple environmental samples.

Finally, the degree of variation between three replicate DNA isolations (CD2a, b and c) from the same soil sample was investigated. Figure 1A–C shows the overall profiles generated from these three DNA isolations; clearly, these profiles are very similar. For comparison, the T-RFLP profile of DNA extracted from soil A (Fig. 1D), in which levels of PCB are negligible, differs considerably in both number and sizes of T-RFs and in their peak heights. The variation in the number (Table 3) and peak heights of the T-RFs from different DNA isolations was no greater than that resulting from intra- or intergel variation. Similar levels of variability have been observed with replicate isolations from another soil sample (data not shown). These results suggest that, at least for this environment, it is not necessary to pool multiple DNA extracts from the same soil sample before T-RFLP analysis.

### *Experimental parameters affecting the composition of T-RFLP profiles*

Having established the overall reproducibility of the method, we have investigated the effects of altering a number of experimental parameters on the profiles generated by T-RFLP analysis. First, the effect of varying the initial template DNA concentration used in the PCR was assessed. Farrelly *et al.* (1995) has suggested previously that, for an artificial community, the ratio of amplified rRNA genes reflects the ratio of template DNA in the starting mixture, despite contrary evidence of bias resulting from factors such as preferential primer annealing (Suzuki and Giovannoni, 1996) or differences in reannealing kinetics when concentrations of PCR products exceed certain levels (Suzuki *et al.*, 1998).

In the current study, PCR products were generated using 1 μl aliquots of 1:10, 1:50, 1:100, 1:500 and 1:1000 dilutions of the CD2a DNA sample (16.5, 3.3, 1.65, 0.33 and 0.165 ng of DNA respectively), digested with *Alu*I and *Hha*I. As would be expected, diluting the initial template concentration resulted in a decline in both the number of peaks and the height of each peak present in a profile, after digestion with either enzyme. Figure 2 shows this decline for 5′ T-RFs after digestion with *Alu*I. Using either enzyme, the number of T-RFs in a profile was seen to decline upon dilution of the initial amount of template DNA, with no additional T-RFs observed in the profiles generated from the higher dilutions of template that were not present in profiles from less diluted template.

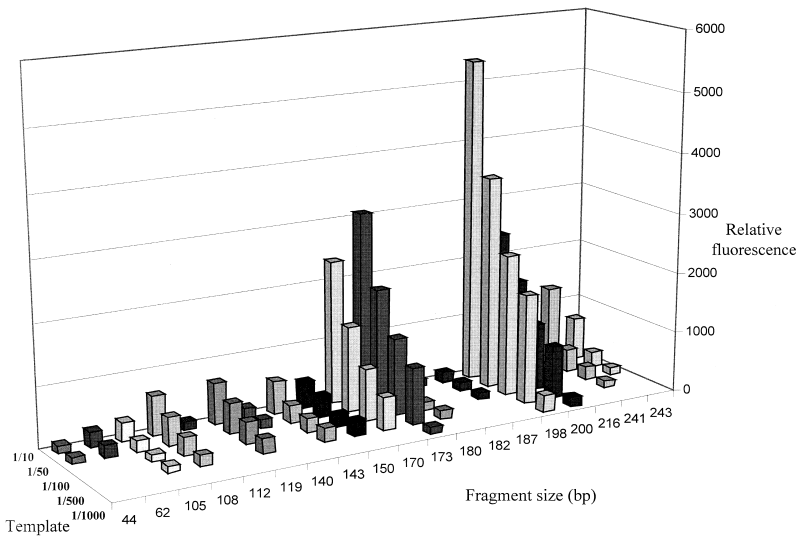


**Fig. 2.** Effect of decreasing DNA concentrations in the PCR on T-RFLP profiles. The 5′-terminal restriction fragments generated after *Alu*I digestion of 16S rRNA gene PCR products amplified from decreasing concentrations of CD2a soil DNA are shown. The size in basepairs of each terminal restriction fragment is shown, together with its peak height in terms of units of fluorescence.

Earlier DGGE- (Ferris and Ward, 1997), TGGE- (Felske *et al.*, 1998) and SSCP-(Schwieger and Tebbe, 1998) based studies investigating the effects of reducing the template concentration also showed such a decline in both number and intensity of bands in these profiles. In the T-RFLP profile generated from a 1:1000 dilution of CD2a DNA, few T-RFs were detected, and these were of the same size, in basepairs, as the T-RFs with the greatest peak heights in the profiles generated from the less diluted DNA. Consequently these peak heights, at the very least, reflect the relative abundance of those ribotypes that are preferentially amplified during PCR and may, in addition, reflect the relative abundance of those ribotypes in the community, subject to any inherent bias during PCR amplification. Furthermore, the use of serial dilution may permit the subsequent possible predictive identification of these dominant ribotypes, after referral to a database of restriction fragment sizes (Brunk *et al.*, 1996).

Reducing the number of PCR cycles, as expected, leads to a reduction in both the number and the peak heights of T-RFs in a profile. For example, after *Hha*I digestion, 36, 26 and 14 T-RFs were produced after 35, 30 and 25 PCR cycles respectively. Every T-RF in the profiles from 25 cycles was also found in the profiles generated after 30 cycles. Similarly, all the T-RFs observed in the profiles from 30 cycles were again found in the profiles generated after 35 cycles for both *Alu*I and *Hha*I digests. This would suggest that the increase in the number of bands, as the number of cycles is increased, results from the appearance of minor T-RFs in the PCR as they are amplified to a level at which they can be detected. We have adopted 30 cycles of PCR before T-RFLP analysis for our standard protocol as, although increased cycling will generate more complex profiles, such profiles will be subject to increasing PCR bias (Suzuki and Giovannoni, 1996), while the use of 25 cycles will provide a more limited community profile.

Another possible source of variation for T-RFLP analysis is the choice of *Taq* polymerase. To investigate this parameter, the profiles resulting from *Alu*I and *Hha*I digestions of PCR products of a 1:10 dilution of CD2A DNA, prepared using either Perkin-Elmer (AmpliTaq) or Qiagen *Taq* polymerases, were compared. A greater number of T-RFs was generated from PCR products with the *Taq* polymerase from Qiagen than from products amplified using AmpliTaq, e.g. after *Hha*I digestion, 41 and 27 T-RFs respectively. Every T-RF generated from the AmpliTaq DNA product was also found in the profile from the Qiagen *Taq* PCR product, suggesting that the choice of *Taq* polymerase did not lead to bias, at least for the predominant members of the population. What is clearly important from these experiments is that, for comparisons of T-RFLP, or any other fingerprinting-based profiles, the same reagents should be used for all experiments. Similar considerations should be made in the choice of oligonucleotide primers, restriction enzymes and PCR product purification systems to ensure that any variation introduced during sample handling is minimized.

Examination of variation in T-RFLP profiles produced by *Alu*I and *Hha*I digestion using different annealing temperatures (55°C and 60°C) in the PCR showed that more T-RFs (65 in total) were generated if an annealing temperature of 60°C was used in the PCR than with an annealing temperature of 55°C (58 T-RFs). This is perhaps surprising, as it might have been anticipated that a higher or more stringent annealing temperature would result in fewer T-RFs. One possible explanation for this apparent anomaly would be that a lower annealing temperature might lead to an increase in the number of different sequences in a particular reaction but, as a consequence, these are produced in lower concentration and are not detected in the T-RFLP profile. Interestingly, all the additional T-RFs formed after annealing at 60°C were also found in a T-RFLP profile generated using Qiagen *Taq* polymerase, and all but two of these T-RFs were found in the profile produced from DNA by 35 instead of 30 cycles of PCR.

The variation observed between T-RFLP profiles upon altering the number of PCR cycles or annealing temperature or using different *Taq* polymerases underlines that diversity is easily missed using this method, although of course this criticism is equally true of all molecular methodologies currently in use. From earlier T-RFLP-based studies, it is apparent that one other potential problem is the possible presence of partially digested products in the T-RFLP profile (Clement *et al.*, 1998; see also Fig. 2 in Bruce, 1997). If digestion of PCR products amplified from complex microbial community DNA is incomplete, then this may lead to the presence of additional peaks in the T-RFLP profile, which may in turn lead to an overestimation of the diversity in the community. Moreover, the sequence context of cleavage sites constitutes a source of bias for certain restriction endonucleases, a bias that is revealed under conditions of limiting enzyme concentration. More importantly, if it is planned to predict on the basis of fragment size, by reference to a database or by comparison with clones from a library, an actual genus or species for any particular T-RF, such partial fragments may lead to erroneous conclusions.

In this study, we have looked at the effects of varying the concentration of restriction enzyme used in the digestion reaction, in terms of both the number of fragments present in a profile and the heights of the peaks in a profile. Reduction of the amount of *Hha*I enzyme from 20 U to 10 U to 5 U had two predictable consequences. First, although the overall number of peaks present in a profile does not increase, a shift in the size of fragments is observed, with the disappearance of some of the smaller fragments coupled to an increase in the number of larger fragments. The fragments that disappear are those that had relatively

**Table 5.** Predicted *in silico* 5′-terminal and partial restriction fragment sizes from 16S rRNA gene PCR products for different species of the genus *Sphingomonas*.

| Species | Accession number | 5′-terminal fragment | 1st partial fragment[a] | 2nd partial fragment[a] | 3rd partial fragment[a] |
|---|---|---|---|---|---|
| *S. yanoikuyae* | D16145 | 108 | **173** (65) | **215** (42) | 359 (144) |
| *S. chlorophenicola* | X87162 | 173 | **215** (42) | 359 (144) | 730 (371) |
| *S. flava* | X87164 | 174 | 216 (42) | 360 (144) | 731 (371) |
| *S. agrestis* | Y12803 | 173 | **215** (42) | 359 (144) | 730 (371) |
| *S. mali* | Y09638 | 173 | 192 (19) | **215** (23) | 359 (144) |
| *S. pruni* | Y09637 | 173 | 192 (19) | **215** (23) | 359 (144) |
| *S. assacharolytica* | Y09639 | 173 | 192 (19) | **215** (23) | 359 (144) |
| *S. trueperi* | X97776 | 173 | **215** (42) | 359 (144) | 518 (159) |
| *S. parapaucimobilis* | X72721 | 98 | 218 (120) | 362 (144) | 521 (159) |
| *S. paucimobilis* | X72722 | 95 | **215** (120) | 519 (304) | 731 (212) |
| *S. adhaesiava* | X72720 | 215 | 359 (144) | 401 (42) | 445 (44) |
| *S. capsulata* | D16147 | 214 | 358 (144) | 400 (42) | 444 (44) |
| *S. subarctica* | X94102 | 215 | 359 (144) | 401 (42) | 445 (44) |

**a**. Predicted partially digested fragments are given with the corresponding product of complete digestion shown in brackets. Fragments shown in bold highlight predicted partially digested fragments, which are the same size as those of the terminal restriction fragments from other *Sphingomonas* species.

low peak heights in the 20 U profile (i.e. between 100 and 200 fluorescence units). Secondly, there is a decrease in the peak heights of some of the smaller fragments, with a concomitant increase in the peak height of some of the larger fragments.

Even when restriction endonuclease is not limiting (e.g. 20 U of the *Hha*I used here), there may still be incomplete digestion of the PCR product. One way to determine which of the bands in a T-RFLP profile may be the result of incomplete digestion is to carry out parallel experiments with a reduced amount of enzyme to identify those T-RFs that increase in peak height with decreasing enzyme concentration. Such T-RFs are potentially composed of partially digested products. One option would be to exclude such fragments from subsequent phylogenetic- (Liu *et al*., 1997) or principal component analysis-based (Clement *et al*., 1998) comparisons of diversity in different samples. However, by doing so, we may actually be ignoring some of the natural variation that exists among different closely related members of the same genus. Table 5 shows both the predicted 5′ T-RF and the partially digested fragment sizes that would be expected after *Alu*I digestion of PCR-amplified 16S rRNA genes from different species of the genus *Sphingomonas.* It is clear from this table that some of the expected T-RFs from certain species are the same size as those of the predicted partially digested fragments from other species. Thus, if such fragments are excluded from the analyses, we risk removing genuine T-RFs that represent different members of the same genus or group. Consequently, such fragments should be retained in the analysis.

The problem of potential false positives is, of course, not limited to T-RFLP analysis. In other profiling methods, single-stranded bands may be seen in TGGE profiles (Heuer *et al*., 1997), while doublets are found in both DGGE (Rosado *et al*., 1998) and TGGE gels, possibly because of incomplete extension of the same template as a result of hairpin formation in the GC clamp region (Nübel *et al*., 1996). Also observed in some DGGE profiles are the presence of additional heteroduplex bands, consisting of one strand from each of two closely related species (Ferris and Ward, 1997), while in SSCP profiles, double-stranded products, representing homo- and heterogeneous products, may be present (Lee *et al*., 1996). This latter problem can, however, be prevented by digestion of a single phosphorylated strand of the PCR product, using lambda exonuclease, when a 5′-terminal phosphate is included on one of the original primers (Schwieger and Tebbe, 1998).

### *Concluding remarks*

The observation of Wintzingerode *et al*. (1997), who state that ‘each physical, chemical and biological step involved in the molecular analysis of an environment is a source of bias which will lead to a distorted view of the real world’, is as applicable to this technique as to any of the other methodologies currently used in microbial ecology. Obviously, any experimental approach has potential problems associated with it, which makes it more or less appropriate for different applications. Judging the appropriateness and knowing the limitations of the selected apparatus necessitates assessment of the inherent sources of bias and variation. There is, however, little quantitative information of this type available for nucleic acid profiling techniques. The present report aims to rectify this situation with regard to T-RFLP analysis.

In this study, we have demonstrated that T-RFLP is a robust and reproducible methodology for the rapid analysis of microbial community structure in different samples

and for the study of community dynamics and changes in community structure in response to changes in prevailing physicochemical parameters. Furthermore, we have attempted to highlight parameters that can significantly influence variability and attempted to define and discuss both the merits and limitations of the method. The high quality of the information provided, namely the precise basepair length of the T-RFs generated, allows the comparison of profiles for any gene system of interest, generated from whole communities and/or from isolates or clones, with databases of *in silico* predicted T-RFs and opens up the possibility of immediate phylogenetic assignment of certain fragments in a profile. We conclude therefore that T-RFLP analysis will develop into a powerful tool in microbial ecology.

## Experimental procedures

### *Isolation of community DNA*

Soil samples (≈ 40 g), transported on ice, were mixed by shaking, divided into 5 g aliquots within 5 h of sampling and frozen at −70°C until DNA was to be isolated. DNA was extracted, essentially according to the method of Zhou *et al*. (1996) with minor modifications. Soil (5 g wet weight) was resuspended in 12 ml of extraction buffer: 100 mM Tris, 100 mM EDTA, 100 mM sodium phosphate buffer (pH 8.0). Proteinase K (100 μl; 10 mg ml$^{-1}$) and 180 μl of lysozyme (100 mg ml$^{-1}$) were added, and the sample was shaken at 37°C for 30 min, followed by the addition of 3 ml of 10% SDS, 4.5 ml of 5 M NaCl and 1.5 ml of 5% CTAB (hexadecylmethylammonium bromide)/1.5 M NaCl and further incubation at 65°C for 15 min. The samples were then frozen in liquid nitrogen and subsequently thawed at 65°C for a further 15 min. This freeze–thaw cycle was repeated twice. After centrifugation at 4200 r.p.m. for 10 min at 4°C (all subsequent centrifugations were carried out under these conditions unless stated otherwise), the resulting supernatant was transferred to a fresh sterile tube. The remaining pellet was resuspended by vortexing for 10 s in 4 ml of extraction buffer, 1 ml of 10% SDS and incubated at 65°C for 10 min. After further centrifugation, the supernatant was removed and combined with that from the first centrifugation. An equal volume of phenol:-chloroform:isoamylalcohol (25:24:1) was added to the combined supernatant and mixed by inversion. The sample was then centrifuged and the aqueous upper layer transferred to a fresh sterile tube. An equal volume of chloroform:isoamylalcohol (24:1) was added, mixed by inversion and, after further centrifugation, the aqueous phase was transferred to a sterile oak ridge tube. Isopropanol (0.7 vol.) at room temperature and 0.3 vol. of 10 M ammonium acetate (pH 7.5) was added to the supernatant with mixture by inversion before centrifugation at 12 000 r.p.m. for 30 min. After careful discarding of the supernatant, the resulting pellet was resuspended in 1.5 ml of 70% ethanol, transferred to an Eppendorf tube and centrifuged at 14 000 r.p.m. for a further 15 min. The supernatant was again discarded and the pellet air dried before resuspension in 250 μl of sterile distilled water.

### *Oligonucleotide primers and PCR amplification*

Oligonucleotide primers designed to anneal at consensus sequences in bacterial 16S rRNA genes (Marchesi *et al*., 1998), FAM63F (5′-CAGGCCTAACACATGCAAGTC-3′) and HEX1389R (5′-ACGGGCGGTGTGTACAAG-3′) were labelled at the 5′ end with the phosphoramidite dyes 6-FAM and HEX (Applied Biosystems) respectively. PCR products were amplified using 2.5 U of *Taq* DNA polymerase (AmpliTaq, Perkin-Elmer, unless otherwise stated) with the primers FAM63F and HEX1389R. Community DNA (1 μl) was used as template (actual masses of DNA are stated in the *Results and discussion*). Reactions (50 μl final volume) were initially denatured for 2 min at 94°C followed by 30 cycles (unless otherwise stated) of 94°C for 1 min, 55°C (unless stated) for 1 min and 72°C for 2 min. This was followed by a final extension step of 72°C for 10 min. Primer and dNTP (dATP, dCTP, dGTP and dTTP) concentrations were 20 pmol and 50 μM respectively. All reagents were prepared as a master mix, before addition to template DNA. Each PCR product (3 μl) was visualized after electrophoresis on 0.8% agarose TAE gels and subsequent staining in ethidium bromide for 15 min. The PCR product (45 μl) was then purified using QIAquick columns (Qiagen) to remove unincorporated nucleotides and labelled primers, and the DNA was eluted in a final volume of 50 μl.

### *T-RFLP analysis*

For both restriction digestion and subsequent denaturation, reagents were prepared as a master mix before addition to the sample. Purified PCR product (10 μl) was digested with 20 U (unless stated otherwise) of either *Alu*I or *Hha*I in a total volume of 15 μl at 37°C for 3 h. This restriction digestion (2 μl) was mixed with 2 μl of deionized formamide, 0.5 μl of ROX-labelled GS500 internal size standard (Applied Biosystems) and 0.5 μl of loading buffer. Each sample was then denatured by heating at 95°C for 5 min and immediately transferred to ice. Samples (1.5 μl of the denatured digest) were electrophoresed in a 36 cm 5% polyacrylamide gel containing 7 M urea for 6 h at 3000 V on an ABI377 genetic analyser with filter set A and a well-to-read length of 36 cm. T-RFLP profiles were analysed using GENESCAN software (version 2.1) (Applied Biosystems). The size, in basepairs, of terminal restriction fragments (T-RFs) was estimated by reference to the internal standard using the Local Southern method, unless stated otherwise. T-RFs with a peak height of less than 100 fluorescence units were excluded from analyses.

Computer-based, *in silico* predictions of the expected T-RFs to be generated from 16S rRNA sequences in the sequence databases were performed using MAPSORT (GCG, Version 8, August 1994, Genetics Computer Group, Madison, WI, USA 53711).

## Acknowledgements

The authors would particularly like to thank Dr Balbina Nogales and Dietmar Böltner for many useful discussions, and would also like to thank Applied Biosystems (Weiterstadt, Germany) for facilitating preliminary experiments on loading variability with an ABI310 capillary-based system. This work

was supported by a grant from the German Federal Ministry for Education and Research (BMBF) BEO project FKY03119433C. K.N.T gratefully acknowledges generous support from the Fonds der Chemischen Industrie.

## References

Avaniss-Aghajani, E., Jones, K., Holtzmann, A., Aronson, T., Glover, N., Boian, M., *et al*. (1996) Molecular technique for rapid identification of Mycobacteria. *J Clin Microbiol* **34:** 98–102.

Brosius, J., Palmer, J.L., Kennedy, H.P., and Noller, H.F. (1978) Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. *Proc Natl Acad Sci USA* **75:** 4801–4805.

Bruce, K.D. (1997) Analysis of *mer* gene subclasses within bacterial communities in soils and sediments resolved by fluorescent-PCR-restriction fragment length polymorphism profiling. *Appl Environ Microbiol* **63:** 4914–4919.

Brunk, C.F., Avaniss-Aghajani, E., and Brunk, C.A. (1996) A computer analysis of primer and probe hybridization potential with bacterial small-subunit rRNA sequences. *Appl Environ Microbiol* **62:** 872–879.

Clement, B.G., Kehl, L.E., DeBord, K.L., and Kitts, C.L. (1998) Terminal restriction fragment patterns (TRFPs), a rapid, PCR based method for the comparison of complex bacterial communities. *J Microbiol Methods* **31:** 135–142.

Cole, S.T., and Girons, I.S. (1994) Bacterial genomics. *FEMS Microbiol Rev* **14:** 139–160.

Farrelly, V., Rainey, F.A., and Stackebrandt, E. (1995) Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Appl Environ Microbiol* **61:** 2798–2801.

Felske, A., Rheims, H., Wolterink, A., Stackebrandt, E., and Akkermanns, A.D.L. (1997) Ribosome analysis reveals prominent activity of an uncultured member of the class Actinobacteria in grassland soils. *Microbiology* **143:** 2983–2989.

Felske, A., Akkermans, A.D.L., and De Vos, W.M. (1998) Quantification of 16S rRNAs in complex bacterial communities by multiple competitive reverse transcription-PCR in temperature gradient gel electrophoresis fingerprints. *Appl Environ Microbiol* **64:** 4581–4587.

Ferris, M.J., and Ward, D.M. (1997) Seasonal distribution of dominant 16S rRNA-defined populations in a hot spring microbial mat examined by denaturing gradient gel electrophoresis. *Appl Environ Microbiol* **63:** 1375–1381.

Head, I.M., Saunders, J.R., and Pickup, R.W. (1998) Microbial evolution, diversity and ecology: a decade of ribosomal RNA analysis of uncultured microorganisms. *Microb Ecol* **35:** 1–21.

Heuer, H., Krsek, M., Baker, P., Smalla, K., and Wellington, E.M.H. (1997) Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. *Appl Environ Microbiol* **63:** 3233–3241.

Lee, D.-H., Zo, Y.-G., and Kim, S.-J. (1996) Nonradioactive method to study genetic profiles of natural bacterial communities by PCR-single-strand-conformation polymorphism. *Appl Environ Microbiol* **62:** 3112–3120.

Liu, W.T., Marsh, T.L., Cheng, H., and Forney, L.J. (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphism of genes encoding 16S rRNA. *Appl Environ Microbiol* **63:** 4516–4522.

Liu, W.T., Marsh, T.L., and Forney, L.J. (1998) Determination of the microbial diversity of anaerobic-aerobic activated sludge by a novel molecular biological technique. *Water Sci Technol* **37:** 417–422.

Marchesi, J.R., Sato, T., Weightman, A.J., Martin, T.A., Fry, J.C., Hiom, S.J., *et al*. (1998) Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. *Appl Environ Microbiol* **64:** 795–799.

Marsh, T.L., Liu, W.T., Forney, L.J., and Cheng, H. (1998) Beginning a molecular analysis of the eukaryal community in activated sludge. *Water Sci Technol* **37:** 455–460.

Moyer, C.L., Tiedje, J.M., Dobbs, F.C., and Karl, D.M. (1996) A computer-simulated restriction fragment length polymorphism analysis of bacterial small-subunit rRNA genes: efficacy of selected tetrameric restriction enzymes for studies of microbial diversity in nature. *Appl Environ Microbiol* **62:** 2502–2507.

Muyzer, G., De Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59:** 695–700.

Muyzer, G., and Smalla, K. (1998) Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie Van Leeuwenhoek* **73:** 127–141.

Nogales, B., Moore, E.R.B., Abraham, W.-R., and Timmis, K.N. (1999) Identification of the metabolically-active members of a bacterial community in a PCB-polluted moorland soil. *Environ Microbiol* (in press).

Nübel, U., Engelen, B., Felske, A., Snaidr, J., Wieshuber, A., Amann, R.I., *et al*. (1996) Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. *J Bacteriol* **178:** 5636–5643.

Rosado, A.S., Duarte, G.F., Seldin, L., and van Elsas, J.D. (1998) Genetic diversity of *nifH* gene sequences in *Paenibacillus azotofixans* strains and soil samples analyzed by denaturing gradient gel electrophoresis of PCR-amplified gene fragments. *Appl Environ Microbiol* **64:** 2770–2779.

Schwieger, F., and Tebbe, C.C. (1998) A new approach to utilize PCR-single-strand-conformation polymorphism for 16S rRNA gene-based microbial community analysis. *Appl Environ Microbiol* **64:** 4870–4876.

Suzuki, M.T., and Giovannoni, 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.

Suzuki, M., Rappé, M.S., and Giovannoni, S.J. (1998) Kinetic bias in estimates of coastal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicon length heterogeneity. *Appl Environ Microbiol* **64:** 4522–4529.

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

Van der Maarel, M.J.E.C., Artz, R.R.E., Haanstra, R., and Forney, L.J. (1998) Association of marine Archaea with the digestive tracts of two marine fish species. *Appl Environ Microbiol* **64:** 2894–2898.

Wang, G.C.-Y., and Wang, Y. (1996) The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. *Microbiology* **142:** 1107–1114.

Wang, G.C.-Y., and Wang, Y. (1997) Frequency of formation of chimeric molecules as a consequence of PCR coamplification of 16S rRNA genes from mixed bacterial genomes. *Appl Environ Microbiol* **63:** 4645–4650.

Wikström, P., Andersson, A.-C., and Forsman, M. (1999) Biomonitoring complex microbial communities using random amplified polymorphic DNA and principal component analysis. *FEMS Microbiol Ecol* **28:** 131–139.

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

Zhou, J., Bruns, M.A., and Tiedje, J.M. (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62:** 316–322.