Published in final edited form as:
J Microbiol Methods. 2013 March ; 92(3): 256–263. doi:10.1016/j.mimet.2012.12.021.

Analysis of Artifacts Suggests DGGE Should Not Be Used For Quantitative Diversity Analysis

Julia W. Neilson[a,*], **Fiona L. Jordan**[a], and **Raina M. Maier**[a]

Fiona L. Jordan: fjordan@m3eng.com; Raina M. Maier: rmaier@ag.arizona.edu

[a]Department of Soil, Water and Environmental Science, 1177 E. 4th Street, Shantz Bldg, Rm 429, The University of Arizona, Tucson AZ 85721

Abstract

PCR-denaturing gradient gel electrophoresis (PCR-DGGE) is widely used in microbial ecology for the analysis of comparative community structure. However, artifacts generated during PCR-DGGE of mixed template communities impede the application of this technique to quantitative analysis of community diversity. The objective of the current study was to employ an artificial bacterial community to document and analyze artifacts associated with multiband signatures and preferential template amplification and to highlight their impacts on the use of this technique for quantitative diversity analysis. Six bacterial species (three *Betaproteobacteria*, two *Alphaproteobacteria*, and one *Firmicutes*) were amplified individually and in combinations with primers targeting the V7/V8 region of the 16S rRNA gene. Two of the six isolates produced multiband profiles demonstrating that band number does not correlate directly with α-diversity. Analysis of the multiple bands from one of these isolates confirmed that both bands had identical sequences which lead to the hypothesis that the multiband pattern resulted from two distinct structural conformations of the same amplicon. In addition, consistent preferential amplification was demonstrated following pairwise amplifications of the six isolates. DGGE and real time PCR analysis identified primer mismatch and PCR inhibition due to 16S rDNA secondary structure as the most probable causes of preferential amplification patterns. Reproducible DGGE community profiles generated in this study confirm that PCR-DGGE provides an excellent high-throughput tool for comparative community structure analysis, but that method-specific artifacts preclude its use for accurate comparative diversity analysis.

Keywords

Multitemplate PCR; DGGE artifact; Preferential amplification; Microbial diversity

1.0 Introduction

Molecular approaches developed in the past decades for the analysis of environmental samples have vastly expanded the horizons of microbial ecology and environmental microbiology. Despite the versatility of many of these techniques, there are biases associated with each one that must be well characterized to facilitate optimal application of these methods to specific research objectives. Diversity studies typically use SSU rRNA genes as phylogenetic markers of microbial taxa present in a given microbial community. Techniques available for characterizing the products generated from the amplification of these marker genes range from high throughput sequencing strategies to low cost survey techniques such as denaturing gradient gel electrophoresis (DGGE). PCR-DGGE is widely used as a rapid

*corresponding author: Julia W. Neilson, Phone: 01 520-621-9759, FAX: 01 520-626-6782, jneilson@email.arizona.edu.

molecular fingerprinting tool for comparing the structure of microbial communities, evaluating temporal changes in those communities or characterizing the impact of environmental perturbations on community structure (Drees et al., 2006; Legatzki et al., 2012; Park and Crowley, 2010; Rosario et al., 2007; van Elsas et al., 2012).

Strengths of PCR-DGGE as a survey tool include the ability to detect changes in the community structure of dominant populations, the reproducibility of the profiles generated, and the option to excise bands of interest to identify specific operational taxonomic units (OTUs) (Drees et al., 2006; Hong et al., 2007; Legatzki et al., 2011). However, artifacts introduced during PCR or the subsequent DGGE analysis lead to skewed results when this method is used for quantitative analysis of α-diversity or relative OTU abundance. Despite consistent evidence for band artifacts associated with DGGE profiles that include 1) multiple bands associated with single isolates (Satokari et al., 2001; Nübel et al., 1996), 2) multiple sequences associated with a single band position (Legatzki et al., 2011; Sekiguchi et al., 2001; Yang and Crowley, 2000), and 3) heteroduplex bands (Hong et al., 2007), studies consistently use DGGE for quantitative assessment of community diversity. Whereas the limitations imposed by these artifacts must be acknowledged, PCR-DGGE still remains an excellent, highly reproducible, comparative community analysis tool when used appropriately (Grandlic et al., 2009; Hong et al.,2007; Legatzki et al., 2012; Nakatsu, 2007; Yang and Crowley, 2000).

The objective of this study was to document and characterize specific PCR-DGGE artifacts that impact the use of this tool for quantitative analysis of community diversity. Specifically, the study analyzes artifacts that skew quantitative associations between DGGE band number and α-diversity or band intensity and OTU abundance. A simple artificial bacterial community was created that incorporated a diversity of factors previously identified as potential sources of PCR-DGGE bias. The factors included: 1) phylogenetic diversity; 2) variable G+C content (Reysenbach et al., 1992); 3) variations in primer binding kinetics (Polz and Cavanaugh, 1998); and 4) potential variability in amplification efficiency due to stable secondary structure (Hansen et al., 1998). The community selected for this study included six environmental isolates (Table 1). A primer set targeting the V7/V8 16S rDNA region was selected for this analysis. This hypervariable region has been used successfully for characterizing the bacterial community structure of a wide range of environmental samples including desert soils (Drees et al., 2006), household dust (Maier et al., 2010) cave stalactites (Legatzki et al., 2011; Legatzki et al., 2012), hot springs (Ferris and Ward, 1997) and for monitoring remediation strategies in metal contaminated mine tailings (de Bashan et al., 2010; Grandlic et al., 2009; Rosario et al., 2007).

Previous studies have found conflicting results concerning the efficacy of PCR-DGGE for characterizing the diversity of microbial communities. First, Hong et al. (2007) found ten pairs of DGGE bands that migrated separately, but had the same sequence in a complex sulfate-reducing mine drainage community. This study concluded that future work was needed to explain this anomaly. Second, they found that a band in the same horizontal position on a DGGE gel represents the same organism less than 50% of the time. Additional studies have confirmed this observation by identifying multiple sequences from different organisms associated with a single band position (Legatzki et al., 2011; Sekiguchi et al., 2001).

In addition to band artifacts, preferential amplification patterns in the PCR-DGGE analysis of mixed template communities have also been identified. Suzuki and Giovannoni (1996) found that a mixture of bacterial templates amplified with the 519F/1406R primer-set produced a ratio of products equal to the initial template concentrations, but amplification of the same mixture with 27F/338R resulted in a strong bias towards a 1:1 product ratio

regardless of initial template ratios. A second template mixture amplified with the first primer set, produced a slight preferential amplification of one template over the other that could not be explained by the effect of G+C content on PCR efficiency. Hansen et al. (1998) observed strong preferential 16S rDNA amplification of one species from a diverse mixture of four isolates and the dominant species varied with the primer set used. Conversely, Lueders and Friedrich (2003) evaluated an artificial consortium consisting of four methanogens and found that the final amplicon frequencies accurately represented the initial templates over a range of template ratios regardless of PCR cycle number or annealing temperature. Similar results were obtained by Salles et al. (2002) using four *Burkholderia* strains. The inconsistency in these results clearly demonstrates the need for further investigation of potential mechanisms skewing the quantitative analysis of community diversity and relative OTU abundance using PCR-DGGE.

2.0 Materials and methods

2.1 Bacterial strains and media

An artificial bacterial community was created from phylogenetically variable environmental isolates (Table 1). The six isolates were all from the domain *Bacteria* and included three *Betaproteobacteria*, two *Alphaproteobacteria* and one *Firmicutes*. The isolation and identification of the organisms were described previously (Bodour et al., 2003). All of the bacteria were maintained as pure cultures on R2A medium (Difco Laboratories, Detroit, MI).

2.2 Preparation and quantification of genomic DNA

Cells were harvested from late log phase cultures of each isolate and lysed with three freeze-thaw cycles in boiling water and liquid nitrogen. Genomic DNA was then extracted following the procedure described by Ausubel et al. (1995), ethanol precipitated and resuspended in sterile deionized water. The concentration and purity of the genomic DNA were determined spectrophotometrically with a BioRad SmartSpec 3000 (Bio-Rad Laboratories, Hercules, CA). Template DNA purity for all isolates had a minimum A_{260}/A_{280} ratio of 1.8.

2.3 PCR amplification

Each of the isolates was amplified individually and amplicon profiles were generated using DGGE. Forward and reverse primers used for PCR amplification were: 1070F (5'ATGGCTGTCGTCAGCT3') and 1392R (5'ACGGGCGGTGTGTAC3') targeting the V7/V8 variable region of the 16S rDNA and amplifying a 346–351 bp product (Table 1). All isolates had 100% sequence identity with both primers with the exception of 71a, which had a 1 bp mismatch with 1070F. Products to be analyzed by DGGE were amplified using a 40-bp GC clamp (5'CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCCC3') attached to the 5' end of primer 1392R (Ferris et al., 1996). A 50 μl PCR reaction volume contained 1× Expand High Fidelity Buffer with 1.5 mM $MgCl_2$ (Roche Diagnostics, Mannheim, Germany), 200 μM concentration of each dNTP, 0.5 μM concentration of each primer, 5% dimethyl sulfoxide, 100 ng/μl bovine serum albumin, 1.5 U Expand High Fidelity DNA polymerase (Roche Diagnostics, Mannheim, Germany), and 5 μl template DNA (1–20 ng/μl). Amplification conditions were previously described (Rosario et al., 2007) and product size was confirmed on 2% GenePure LE agarose (ISC Bioexpress, Kaysville, UT) following ethidium bromide staining.

2.4 DGGE conditions and profile analysis

Amplicon profiles were analyzed with the D-Code Universal Mutation Detection System (Bio-Rad Laboratories, Hercules, CA) using 7% acrylamide gels with a 40 to 75% denaturing gradient (100% denaturing solution contains 7 M urea and 40% formamide). The gels were electrophoresed for 19 h in 0.5× TAE Buffer (pH 8.6) at a constant voltage and temperature of 50V and 60°C, respectively and then visualized with an AlphaInnotech Imaging System (Alpha Innotech Corporation, San Leandro, CA) following ethidium bromide staining. DNA from individual bands was extracted for downstream analysis by elution in 100 μl buffer (0.5 M NH_4OAc, 1 mM EDTA, pH 8.0) with shaking overnight at 37°C. Residual acrylamide was removed by passing extracts through a 1 cc syringe containing glass wool and DNA was recovered by ethanol precipitation. Excised bands were sequenced in both directions using primers 1070F and 1392R (University of Arizona Genetics Core, Arizona Research Labs, University of Arizona, Tucson, AZ).

2.5 Evaluation of multitemplate PCR

Multitemplate amplification was evaluated using pairwise mixtures of template DNA from each of the six isolates. Equal amounts (1:1) of template DNA from each isolate pair were amplified in 25 μl reactions. Template mixtures were prepared prior to amplification and added to triplicate PCR reaction tubes. Amplification of template DNA was confirmed by DGGE analysis and scored in terms of relative band intensity. The magnitude of bias in template pairs demonstrating preferential amplification was quantified by increasing the template concentration of the inhibited pair until equal band intensity was observed.

2.6 Real time PCR reaction conditions

Real-time PCR reactions were performed with the iCycler iQ system, using the iQ SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA). PCR efficiencies for each isolate were determined from five 2-fold dilutions of template DNA (5–100 ng/μl). Twenty-five microliter reactions contained 1× SYBR Green Supermix, 0.5μM of each primer and 1μl of template DNA from each of the serial dilutions. All reactions were performed in triplicate with triplicate template-free controls included on each plate to evaluate background fluorescence. Differences in PCR amplification patterns for each isolate were evaluated by comparing real time amplification curves using equal template concentrations (11.5 ± 0.6 ng DNA in 25 μl reactions). All isolates were amplified simultaneously on the same 96-well plate to minimize potential biases resulting from variations in reaction mixes or iCycler performance. All experiments were repeated to confirm results. The PCR cycle threshold and amplification efficiency analyses were done using the iCycler iQ Optical System Software version 3.0a (BioRad Laboratories, Inc. Hercules, CA).

3.0 Results

3.1 Multiband artifacts

Band patterns for each of the six individual isolates were characterized using DGGE. Isolates M22, 35, 50, and M16 had unique single bands, while 70-2 and 71a had multiband profiles containing two and four bands, respectively (Fig. 1). DGGE analysis of the individual isolates and of pooled amplicon mixtures revealed clear and reproducible resolution of isolate bands regardless of the combination of isolate amplicons pooled prior to electrophoresis (Fig. 1). Thus, migration patterns and band intensities of individual isolates were not affected when run on DGGE gels in combination with other isolates. The multiband profile patterns for isolates 70-2 and 71a were consistent over a range of electrophoresis conditions including variations in denaturing gradient, percent acrylamide and electrophoresis time (data not shown).

Isolate 70-2 was selected for further analysis of the multiband profile. The top and bottom bands were excised separately and the products were extracted for analysis. In the first experiment, both bands were re-amplified separately and the resulting products were evaluated by DGGE. Two-band patterns identical to the original 70-2 profile were generated from amplification of either the top or the bottom band. In a second trial, the excised bands were reanalyzed by DGGE without re-amplification. A sufficient quantity of DNA was obtained for this analysis by excising and pooling like bands from four consecutive lanes. In contrast to the previous results, single DGGE bands produced single top and bottom bands that migrated to the same position as their respective counterparts from the original 70-2 profile (Fig. 2). These results confirmed that the doublet configuration was an artifact of the PCR rather than the DGGE.

The two excised 70-2 bands were then sequenced in both the forward and reverse directions using primers 1070F and 1392R. Both the top and bottom band sequences were found to be identical to each other as well as to the initial PCR product and to the original 70-2 GenBank sequence (AY177375). Thus, the doublet-band profile of isolate 70-2 did not indicate the presence of two unique 16S rRNA gene sequences in this organism.

3.2 Artifacts related to heteroduplex bands in DGGE profiles

A comparison of the multitemplate DGGE profiles in Figs. 1 (lanes 1–5) and 3 (lanes 6–13) clearly demonstrates the presence of additional bands in the top third of the gel following amplification of multitemplate DNA mixtures (Fig. 3). These bands were not observed following DGGE of pooled products that had been amplified separately (Fig. 1). These artifacts were identified as heteroduplexes due to their proximity to the loading wells. Heteroduplex formation results from the reannealing of similar, but non-identical single-stranded products in PCR and the products formed are large with a decreased mobility in DGGE gels that is inversely proportional to the sequence similarity of the two template molecules (Hong et al., 2007; Kanagawa, 2003). The heteroduplex band patterns shown in lanes 6 through 13 (Fig. 3) were unique to each specific template pair and were produced reproducibly following replicate reactions. Thus, the bands formed were not random events.

3.3 Artifacts related to PCR bias

Paired template mixtures were created and amplified using equal template combinations of each of the six isolates. The resulting PCR products were evaluated by DGGE and the presence and absence of isolate bands as well as their relative intensity was recorded (Fig. 3, Table 2). The absence of an isolate band was recorded as no amplification for that isolate, despite the fact that the product may have been present at a level below the electrophoresis detection limit. All reactions were performed in triplicate and the results were reproducible.

Pairwise amplification results revealed that the band intensity of a specific isolate in a DGGE profile is a function of the other isolates present in a multitemplate PCR (Table 2). Isolate 71a was not amplified in any of the pairwise combinations except when amplified with isolate 35. Even in this combination the 71a band intensities were less than the isolate-35 band. Isolates 70-2, 50, M22, M16, and 35 produced bands when amplified in every pairwise combination; however, the resulting band intensities varied with the specific template combination. For example, isolate M16 produced an equivalent or greater band intensity when amplified in combination with 35 or 71a, but a lesser band intensity when amplified with 70-2, 50, or M22 (Fig. 3, lanes 10, 11, and 13; Table 2). In contrast, isolates 70-2 and 50 produced bands of equivalent or greater intensity regardless of the template mixtures. Amplification of all six isolates together using equal template concentrations produced bands for 70-2, 50, M22 and M16, but not for 35 or 71a (data not shown). The magnitude of observed preferential amplification was quantified by manipulating initial

PCR template concentrations to generate DGGE bands of comparable intensities. Equal band intensities were achieved using the following ratios: 1) 2:1 template ratio of M16 to M22, 70-2 or 50; 2) 2:1 ratio of 35 to M16; 3) 5:1 ratio of 35 with 50 or M22; and 4) 15:1 ratio of 71a with 70-2, 50, or M22. For none of these combinations did increases in template concentration of the secondary isolates compromise the band intensity of the primary isolate. In contrast, a 3:1 ratio of 71a to 35 produced bands of comparable intensity, whereas a template ratio of 6:1 produced a lesser band intensity for isolate 35. The cause of the observed multitemplate preferential amplification patterns was further investigated using quantitative real-time PCR (qPCR).

3.4 qPCR characterization of amplification dynamics

qPCR was used to investigate the amplification dynamics of the individual isolates. Isolate 71a was not included in this analysis because this isolate contained a single base primer-mismatch, whereas the remaining five isolates had 100% sequence identity with both primers. Previous research has identified primer-template mismatch as a cause of low amplification efficiency (Ahn et al., 2009; Kanagawa, 2003; Polz and Cavanaugh, 1998) and 71a was least competitive in all multitemplate mixtures. Thus, the poor amplification of 71a was attributed to primer mismatch and it was removed from further analysis.

Fluorescence data obtained from qPCR typically follows a sigmoidal curve as defined by Tichopad et al. (2003). An initial ground phase precedes the exponential amplification phase and is characterized by low product concentrations that are undetectable due to background fluorescence generated by the reaction mix. We hypothesize that the length of the ground phase for an individual isolate is related to factors inhibiting primer annealing and early extension of template DNA for a specific isolate. Cycle threshold (C_T) data from the five isolates were used to quantify the ground phase length for each isolate where a higher C_T value corresponded to a longer ground phase. The template concentrations selected for analysis were comparable to those used in the previous pairwise amplification experiments. The experiments were conducted both with and without the GC-clamp (using 1392R and 1392R-GC) to evaluate the influence of the clamp on reaction dynamics. Threshold points were set at 20 relative fluorescence units (RFU) for all reactions using 1392R-GC and 90 RFU for 1392R reactions.

Results from both sets of experiments are shown in Fig. 4A. Similar average C_T values were obtained for isolates M22, 50, and 70-2, when amplified with GC-clamp primers. In contrast, isolates M16 and 35 had C_T values 2.5 and 4.1 cycles greater, respectively, than the combined average of values for isolates M22, 50 and 70-2, indicating a longer ground phase associated with these isolates. The same pattern was observed for amplification using primers without the GC-clamp (Fig. 4A; 2.2 and 4.2 cycle increase for M16 and 35, respectively). These data reveal two patterns in early amplification dynamics. First, the GC-clamp caused a 1.1 ± 0.2 average cycle increase in ground phase length for all isolates evaluated, indicating that the clamp consistently inhibited primer annealing or early PCR extension in all reactions. Second, the ground phase was significantly longer for isolates M16 and 35, regardless of whether the clamp was present, indicating early-PCR inhibition for these two isolates. The reactions were repeated for all isolates and similar results were obtained.

Reaction fluorescence data for all isolates were then normalized with respect to the C_T value for isolate 35 (Fig. 4B) since this isolate had the longest ground phase, thus presumably the greatest early-cycle PCR inhibition. Normalization of C_T values with respect to isolate 35 provided a quantification of the competitive disadvantage suffered by 35 as a result of early cycle inhibition. Using data from amplification with the GC-clamp primers, the results indicated that at the onset of exponential amplification for isolate 35 (C_T = 20 RFU),

fluorescence values were 2.7-fold higher for isolate M16 and 6.7 to 8.2-fold higher for M22, 50 and 70-2. Similarly, at the M16 C_T-cycle, fluorescence levels for isolates M22, 50 or 70-2 were 4.0 to 5.5-fold higher. These patterns quantify the competitive disadvantage suffered by isolates 35 and M16 and mirror the patterns observed in the pairwise PCR-DGGE analyses.

PCR amplification efficiencies were also considered as possible explanations for differences in competitive amplification behavior. For 70-2, 50 and 35 variations of up to 20% in PCR efficiency were observed between replicated reactions (data not shown) although the variability in PCR efficiency was less for reactions without the GC-clamp. Such variability is not surprising given the nonspecific amplification conditions required for PCR-DGGE (personal communication from Bio-Rad technical support). Despite the variability in PCR efficiency, the relative differences in C_T values between isolates in replicated reactions (Fig. 4A) were consistent.

3.5 Evaluation of 16S rRNA gene secondary structure as a source of PCR inhibition

Previous studies have hypothesized that the early-cycle PCR inhibition observed for isolates M16 and 35 may result from inhibition of primer annealing and early extension by secondary structures at or near primer binding sites (Hansen et al., 1998; Polz and Cavanaugh, 1998). To test this hypothesis, we analyzed the most probable secondary structure of the 16S rDNA for each of the five isolates using the M-fold software of GCG/Seq Lab (Wisconsin Package v. 10.3 Accelrys Inc., San Diego, CA). The results revealed that both primer-binding sites contained hairpin loops for all five isolates at the annealing temperature of 55°C (Fig. 5). In contrast, at the extension temperature of 72°C one of the binding sites was located in a region free of secondary structure for isolates 70-2, 50 and M22 whereas both primer binding sites were still located in hairpin loops for isolates M16 and 35. Based on this analysis, we conclude that secondary structure at the primer binding site had a significant impact on the comparative primer binding and early extension kinetics of these isolates, providing one explanation for the observed extended ground phase impeding the competitive amplification of isolates M16 and 35. The Mfold analysis clearly demonstrates that the influence of secondary structure is both sequence and primer specific, thus the preferential amplification patterns observed for this community will vary with the primer sites selected.

4.0 Discussion

4.1 DGGE band artifacts

Multiband profiles associated with single isolates clearly represent artifacts that cause distortions in PCR-DGGE community diversity analyses; however authors continue to use band number as a proxy for α-diversity. The multiband profiles identified in this work were associated with isolates 70-2 and 71a. The profile bands were separated by significant distances and comprised a unique pattern specific to each isolate. Novel to this work, was the analysis of the isolate 70-2 dual-band profile which found that the sequence associated with each band was identical despite the fact that the bands had distinct migration patterns. We hypothesized that these individual bands represented different stable structural conformations for this single amplicon generated during PCR and that these two conformations had different denaturation dynamics when exposed to the denaturant in the polyacrylamide gel (Al Agellon, Arizona Research Labs, personal communication). This hypothesis was based on the observed difference in migration patterns for the first and second 70-2 trials.

Previous studies have identified alternate explanations for isolate-specific multiband profiles, but this is the first evidence revealing that multiple bands can be produced as PCR artifacts from the amplification of a single sequence. Janse et al. (2004) analyzed artifactual double bands that could be eliminated by extending the length of the final 72°C extension, but these secondary bands were shadow bands traveling consistently in close proximity to each prominent band. Others have argued that multiband profiles result from multiple heterogeneous 16S rDNA copies within a single organism (Cocolin et al., 2001; Nübel et al., 1996; Salles et al., 2002), but one can argue that these copies have distinct sequences and thus should logically migrate as separate bands.

Additional artifacts observed in this study that skew diversity quantification were the heteroduplex bands identified in Fig. 3; artifacts that have been documented in numerous other papers (Ferris and Ward, 1997; Syvyk et al., 2008). Syvyk et al. described heteroduplexes as “phantom bands” that form during the final PCR cycles when amplicon concentrations are high relative to primer concentrations. Studies have shown that these bands can be minimized if not eliminated by reducing the number of PCR cycles (Syvyk et al., 2008) or reducing the starting template concentration (Legatzki et al., 2011), but bands located in the top centimeter of DGGE gels should be excluded from profiles to minimize overestimation of diversity as a result of heteroduplex anomalies. A final artifact observed in numerous studies is the co-migration of bands with divergent sequences (Hong et al., 2007; Legatzki et al., 2011; Sekiguchi et al., 2001), an artifact leading to an underestimation of diversity. In summary, multiple PCR-DGGE artifacts skew diversity estimates based on DGGE band number, thus DGGE should be used as a tool for analyzing comparative community structure, not as a means of quantifying α-diversity,

4.2 Preferential Amplification in Multitemplate PCR

Multitemplate PCR amplification of the artificial community designed for this study clearly demonstrated that DGGE band intensity does not always correlate with initial template concentration. Two factors were identified as probably causes of preferential amplification. First, the one isolate with a primer:template sequence mismatch (71a) amplified most poorly in all template combinations. Second, early-cycle PCR inhibition resulting from secondary-structure inhibition at primer binding sites was found to be the most probable explanation for preferential amplification among isolates with 100% primer:template sequence identity. Isolates 70-2 and 50 demonstrated the most consistent amplification, producing DGGE bands of equal intensity when amplified in all pairwise combinations with other isolates in the community. These isolates were both *Betaproteobacteria* with high levels of sequence similarity in the region being amplified (96.5 %). The next most consistent result was obtained for isolate M22, also a member of the *Betaproteobacteria* with 98.3 and 96% sequence identity to 50 and 70-2, respectively. M22 produced a brighter band when amplified with 50 than when amplified with 70-2 suggesting that templates with greater sequence identity compete more successfully in multitemplate PCR amplifications. The *Alphaproteobacteria*, M16 and 35, produced faint bands when amplified with any of the *Betaproteobacteria*, but stronger bands when amplified with each other. The sequence identity of these isolates with each other was 93.7%, while the average identity with the *Betaproteobacteria* was 84.9 and 83.2%, respectively.

The results suggest that isolates with greater phylogenetic and sequence similarity appeared to compete more favorably in multitemplate mixtures. Previous studies have also found that that more closely related organisms (with greater sequence similarity) show less evidence of preferential amplification (Leuders and Friedrich, 2003; Salles et al., 2002). These observations are supported by the M-fold analysis that found the secondary structure patterns to be most consistent among closely related organisms. In summary, the results

suggest that preferential amplification within this artificial community was influenced most by primer mismatch and second by PCR inhibition due to secondary structure inhibition at primer binding sites during early amplification cycles.

A second explanation that must be considered is the potential influence of the rRNA gene copy number on the qPCR results. A standardized DNA template concentration does not guarantee equal rDNA concentrations for all isolates. Klappenbach et al.(2000) reported a range of 1 to 6 copies of the rRNA operon among different *Alpha*- and *Betaproteobacteria* isolated from a soil sample and up to 10–13 rDNA copies for low G + C organisms such as *Bacillus* sp and *Clostridium* sp. The five isolates analyzed in this study by qPCR were all *Alpha*- and *Betaproteobacteria* and thus could potentially have up to 6-fold differences in rDNA copy number. A one cycle C_T difference in real time PCR corresponds to a 2-fold difference in initial template concentration. Thus, the 4.4 cycle C_T difference observed between isolate 35 and isolate 50, and the 3.9 cycle difference between 35 and 70-2 and M22 would represent $2^{3.9}$ and $2^{4.4}$ or 15- and 21-fold differences in copy number, respectively. Since copy numbers greater than 10 have only been found in low G+C gram positive bacteria (Klappenbach et al., 2000), a 15–21 fold difference in copy number among *Alpha* and *Betaproteobacteria* is highly improbable. Thus, this study concludes that early-cycle PCR inhibition caused the variation in qPCR C_T values following amplification of equal template concentrations of the five isolates.

In summary, the pairwise amplification and qPCR results clearly demonstrated that band intensity does not consistently represent initial template concentration. In fact, experiments revealed that a 15×s concentration of isolate 71a was required to obtain a band of equal intensity with that of any of the *Betaproteobacteria* (70-2, 50 or M22) analyzed. Thus, PCR-DGGE band intensity is not an accurate indicator of community OTU abundance.

Conclusions

This study demonstrates that PCR-DGGE provides a highly reproducible assessment of microbial community structure, but a biased quantification of species richness and relative species abundance. First, a novel analysis of a multiband isolate profile demonstrated that two unique, but reproducible bands were generated from the amplification of a single sequence. This phenomenon produces overestimation in DGGE-based community diversity analyses. Further, preferential amplification biases between phylogenetically diverse members of this community caused by early-cycle PCR inhibition and primer mismatch were shown to skew quantification of relative species abundance based on band intensity. Numerous improvements have been proposed to reduce the impact of PCR-DGGE artifacts on quantitative diversity analysis, but we contend that this molecular tool should be reserved for comparative community structure analysis, not comparative diversity. PCR-DGGE is best adapted for generating community snapshots in applications monitoring and identifying microbial community shifts in response to seasonal changes, bioremediation applications or environmental perturbations. Specifically, PCR-DGGE can be applied to screen large sample sets to identify a subset of interest for further examination of α- or functional diversity.

Acknowledgments

This research was supported by the National Institute of Environmental Health Sciences Superfund Basic Research Program Grant 2 P42 ES04940 and the National Science Foundation Microbial Observatory Grant MCB60404300.

We thank Ryan Sprissler of the Genomic Analysis and Technology Core and Al Agellon, former Education and Training Program Coordinator of the Arizona Research Labs, Biotechnology Division, at the University of Arizona, Tucson, AZ for technical support in the data analysis.

References

Ausubel, FM.; Brent, R.; Kingston, RE.; Moore, DD.; Seidman, JG.; Smith, JA.; Struhl, K. Preparation of genomic DNA from bacteria. In: Ausubel, FM., editor. Current Protocols in Molecular Biology. New York: John Wiley & Sons, Inc.; 1995. p. 2.4.1

Ahn JK, Kim Y, Kim T, Song H, Kang C, Ka J. Quantitative improvement of 16S rDNA DGGE analysis of soil bacterial community using real-time PCR. J. Microbiol. Meth. 2009; 78:216–222.

Bodour AA, Wang J, Brusseau ML, Maier RM. Temporal change in culturable phenanthrene degraders in response to long-term exposure to phenanthrene in a soil column system. Environ. Microbiol. 2003; 5:888–895. [PubMed: 14510842]

Cocolin L, Manzano M, Cantoni C, Comi G. Denaturing gradient gel electrophoresis analysis of the 16S rRNA gene V1 region to monitor dynamic changes in the bacterial population during fermentation of Italian sausages. App. Environ. Microbiol. 2001; 67:5113–5121.

de Bashan LE, Hernandez J, Nelson KN, Bashan Y, Maier RM. Growth of quailbush in acidic, metalliferous desert mine tailings: effect of *Azospirillum brasilense* Sp6 on biomass production and rhizosphere community structure. Microb. Ecol. 2010; 60:915–927. [PubMed: 20632001]

Drees KP, Neilson JW, Betancourt JL, Quade J, Henderson DA, Pryor BM, Maier RM. Bacterial community structure in the hyperarid core of the Atacama Desert, Chile. App. Environ. Microbiol. 2006; 72:7902–7908.

Ferris MJ, Muyzer G, Ward DM. Denaturing gradient gel electrophoresis profiles of 16S rRNA-defined populations inhabiting a hot spring microbial mat community. Appl. Environ. Microbiol. 1996; 62:340–346. [PubMed: 8593039]

Ferris MJ, Ward DM. Seasonal distributions of dominant 16S rRNA-defined populations in a hot spring microbial mat examined by denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. 1997; 63:1375–1381. [PubMed: 9097434]

Grandlic CJ, Palmer MW, Maier RM. Optimization of plant growth-promoting bacteria-assisted stabilization of mine tailings. Soil. Biol. Biochem. 2009; 41:1734–1740. [PubMed: 20161141]

Hansen MC, Tolker-Nielsen T, Givskov M, Molin S. Biased 16S rDNA PCR amplification caused by interference from DNA flanking the template region. FEMS Microbiol. Ecol. 1998; 26:141–149.

Hong H, Pruden A, Reardon KF. Comparison of CE-SSCP and DGGE for monitoring a complex microbial community remediating mine drainage. J. Microbiol. Meth. 2007; 69:52–64.

Janse I, Bok J, Zwart G. A simple remedy against artifactual double bands in denaturing gradient gel electrophoresis. J. Microbiol. Meth. 2004; 57:279–281.

Kanagawa T. Bias and artifacts in multitemplate polymerase chain reactions (PCR). J. Biosci. Bioeng. 2003; 96:317–323. [PubMed: 16233530]

Klappenbach JA, Dunbar JM, Schmidt TM. rRNA operon copy number reflects ecological strategies of bacteria. App. Environ. Microbiol. 2000; 66:1328–1333.

Legatzki A, Ortiz M, Neilson JW, Casavant RR, Palmer MW, Rasmussen C, Pryor BM, Pierson LS, Maier RM. Factors influencing observed variations in the structure of bacterial communities on calcite formations on Kartchner Caverns, AZ, USA. Geomiccrobiol. J. 2012; 29:422–434.

Legatzki A, Ortiz M, Neilson JW, Dominguez S, Andersen GL, Toomey RS, Pryor BM, Pierson LS, Maier RM. Bacterial and archaeal community structure of two adjacent calcite speleothems in Kartchner Caverns, Arizona, USA. Geomicrobiol. J. 2011; 28:99–117.

Lueders T, Friedrich MW. Evaluation of PCR amplification bias by terminal restriction fragment length polymorphism analysis of small-subunit rRNA and *mcrA* genes by using defined template mixtures of methanogenic pure cultures and soil DNA extracts. Appl. Environ. Microbiol. 2003; 69:320–326. [PubMed: 12514011]

Maier RM, Palmer MW, Andersen GL, Halonen MJ, Josephson KC, Maier RS, Martinez FD, Neilson JW, Stern DA, Vercelli D, Wright AL. Environmental determinants of and impact on childhood asthma by the bacterial community in household dust. Appl. Environ. Microbiol. 2010; 76:2663–2667. [PubMed: 20154107]

Nakatsu CH. Soil microbial community analysis using denaturing gradient gel electrophoresis. Soil Sci. Soc. Am. J. 2007; 71:562–571.

Nübel U, Engelen B, Felske A, Snaidr J, Wieshuber A, Amann RI, Ludwig W, Backhaus H. Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. J. Bacteriol. 1996; 178:5636–5643. [PubMed: 8824607]

Park J, Crowley DE. Nested PCR bias: a case study of *Pseudomonas* spp. in soil microcosms. J. Environ. Mon. 2010; 12:985–988.

Polz MF, Cavanaugh CM. Bias in template-to-product ratios in multitemplate PCR. Appl. Environ. Microbiol. 1998; 64:3724–3730. [PubMed: 9758791]

Reysenbach AL, Giver LJ, Wickham GS, Pace NR. Differential amplification of rRNA genes by polymerase chain reaction. Appl. Environ. Microbiol. 1992; 58:3417–3418. [PubMed: 1280061]

Rosario K, Iverson SL, Henderson DA, Chartrand S, McKeon C, Glenn EP, Maier RM. Community changes during plant establishment at the San Pedro River mine tailings site. J. Environ. Qual. 2007; 36:1249–1259. [PubMed: 17636285]

Salles JF, de Souza FA, van Elsas JD. Molecular method to assess the diversity of *Burkholderia* species in environmental samples. Appl. Environ. Microbiol. 2002; 68:1595–1603. [PubMed: 11916673]

Satokari RM, Vaughan EE, Akkermans ADL, Saarela M, de Vos WM. Bifidobacterial diversity in human feces detected by genus-specific PCR and denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. 2001; 67:504–513. [PubMed: 11157210]

Sekiguchi H, Tomioka N, Nakahara T, Uchiyama H. A single band does not always represent single bacterial strains in denaturing gel electrophoresis analysis. Biotechnol. Lett. 2001; 23:1205–1208.

Suzuki MT, Giovannoni SJ. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. 1996; 62:625–630. [PubMed: 8593063]

Syvyk A, Nalian A, Hume M, Martynova-vanKley A. A positive control for detecting heteroduplexes in DGGE for microbial community fingerprinting. Texas J Sci. 2008; 60:33–44.

Tichopad A, Dilger M, Schwarz G, Pfaffl MW. Standardized determination of real-time PCR efficiency from a single reaction set-up. Nucleic Acids Res. 2003; 31:6688.

van Elsas JD, Chiurazzi M, Mallon CA, Elhottova D, Kristufek V, Salles JF. Microbial diversity determines the invasion of soil by a bacterial pathogen. Proc. Nat. Acad. Sci. U.S.A. 2012; 109:1159–1164.

Yang CH, Crowley DE. Rhizosphere microbial community structure in relation to root location and plant iron nutritional status. Appl. Environ. Microbiol. 2000; 66:345–351. [PubMed: 10618246]

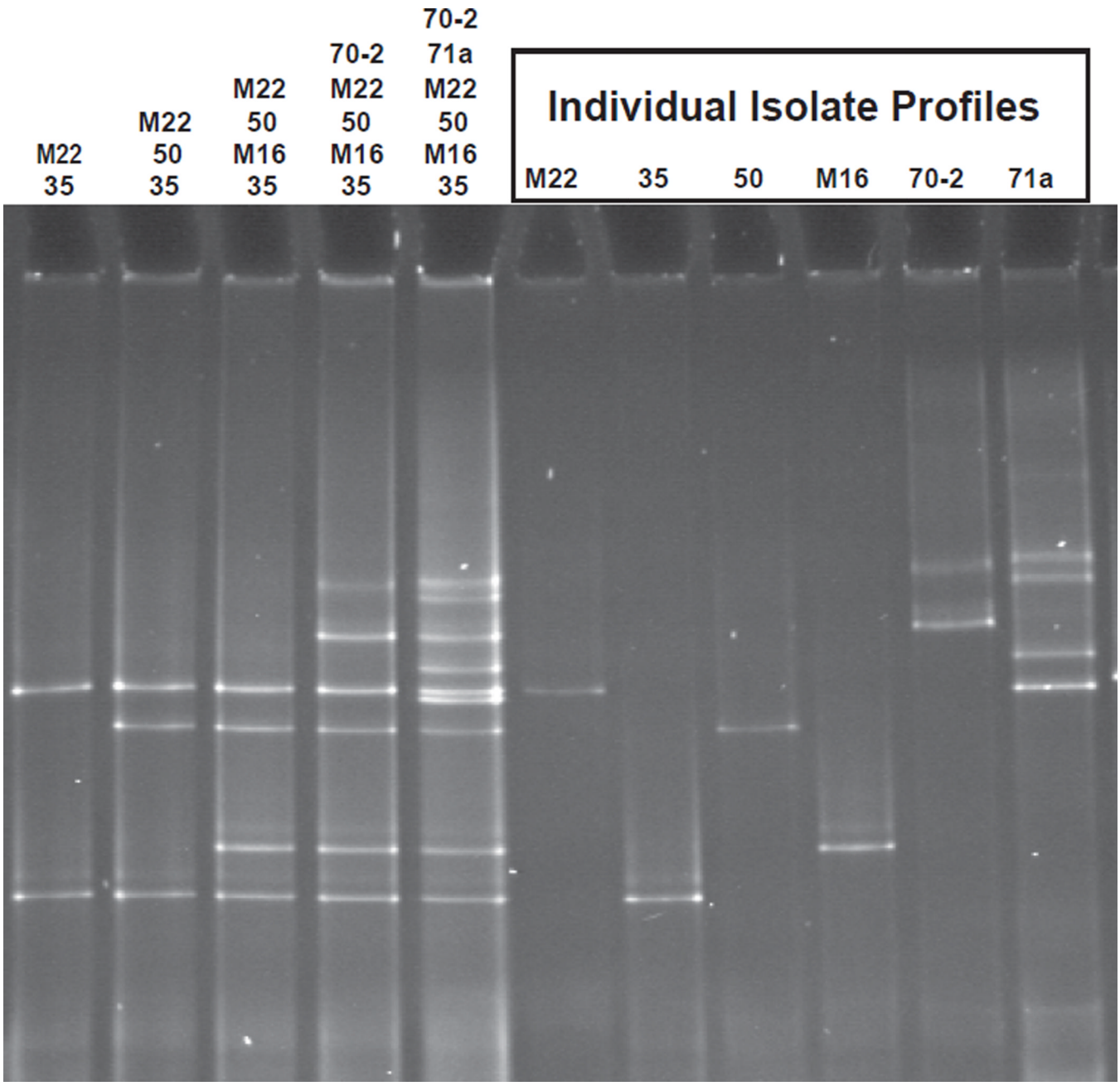


Fig. 1.
DGGE profiles of individually amplified isolates. Lanes with multiple isolates represent post-PCR combinations of equal volumes of amplified products.

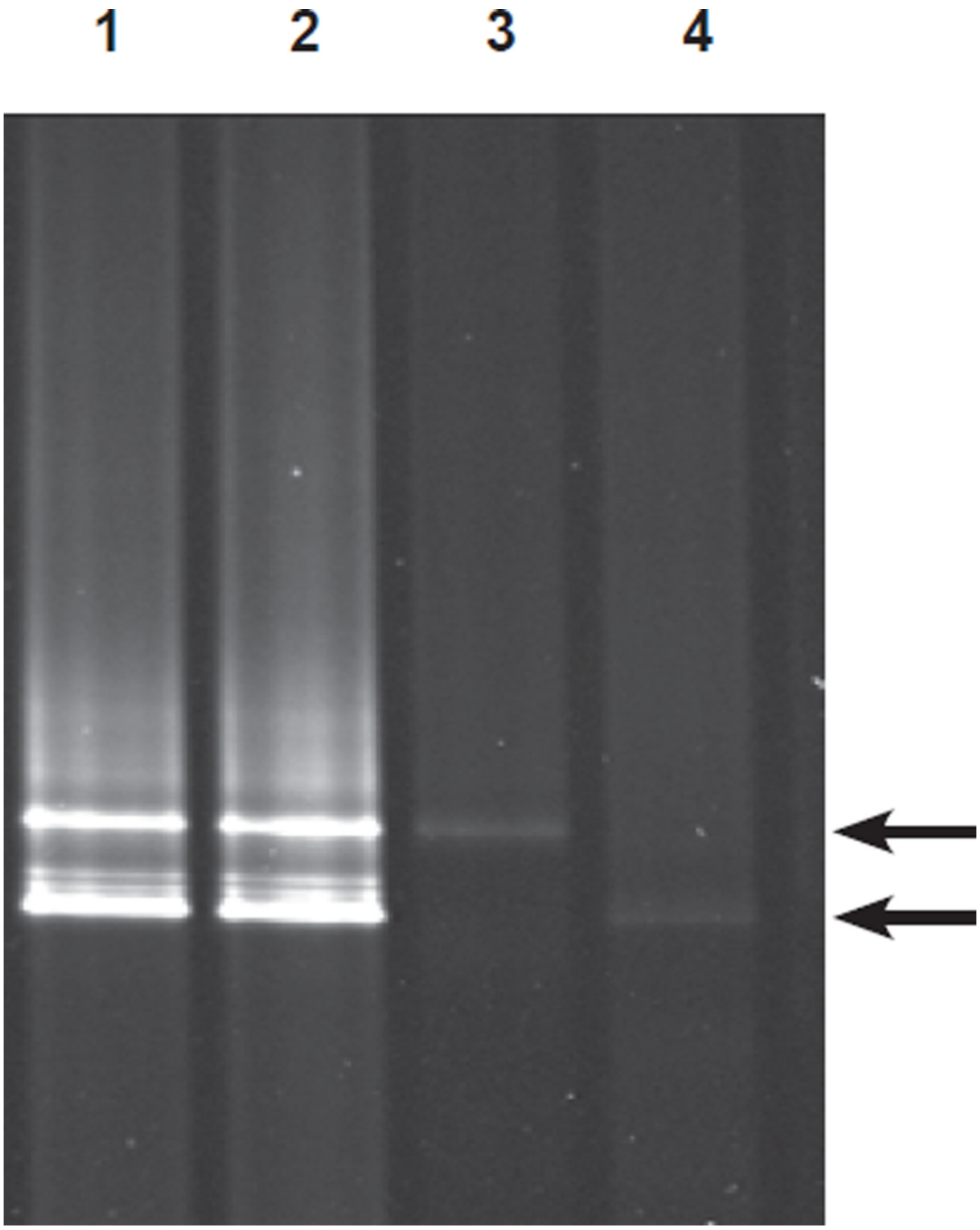


Fig. 2.

DGGE profile of amplified DNA from isolate 70-2. Lanes 1 and 2: PCR products from duplicate amplifications of template DNA from isolate 70-2. Lanes 3 and 4: DNA excised from the top (lane 3) and bottom (lane 4) bands of the isolate 70-2 DGGE profile. Excised DNA was subjected to DGGE without re-amplification. Arrows indicate the location of the top and bottom bands.

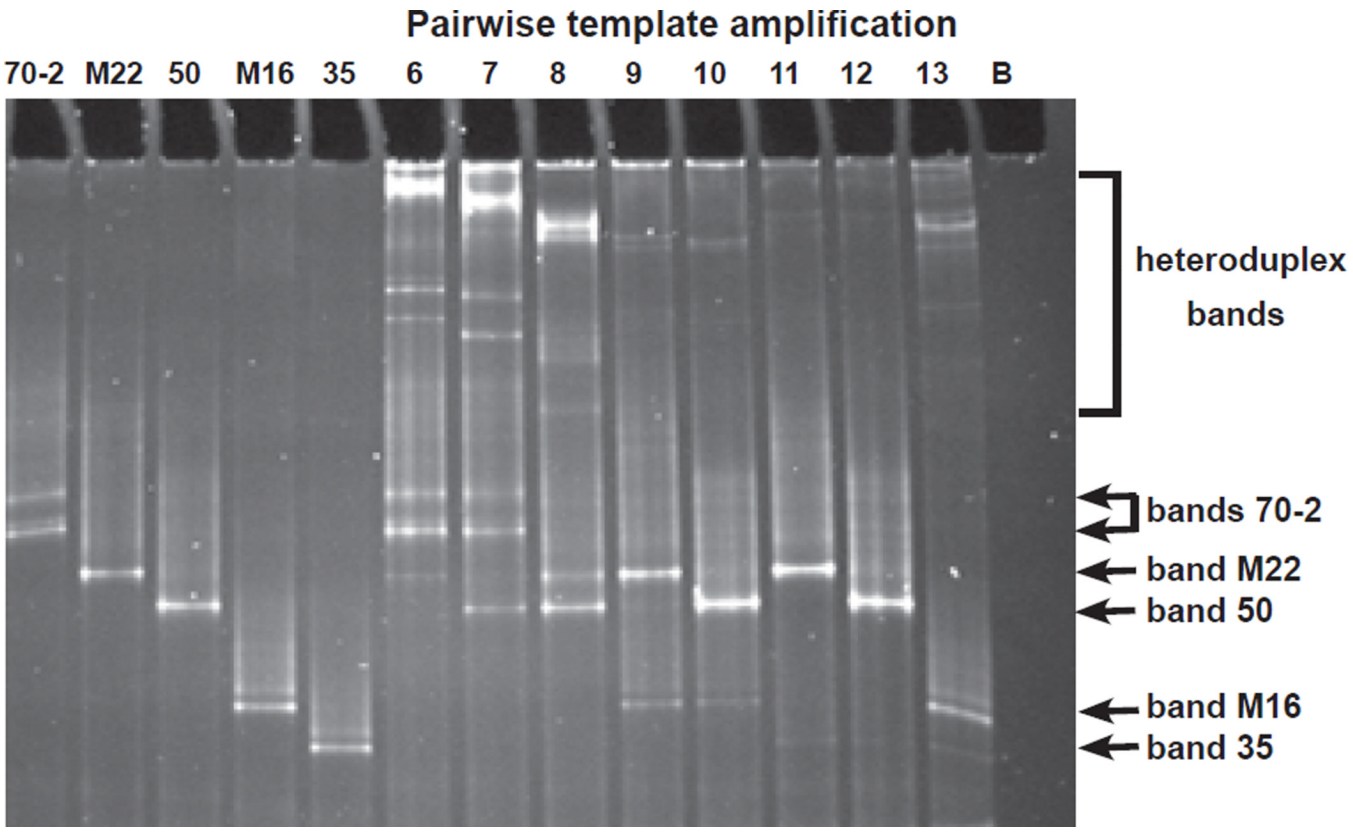
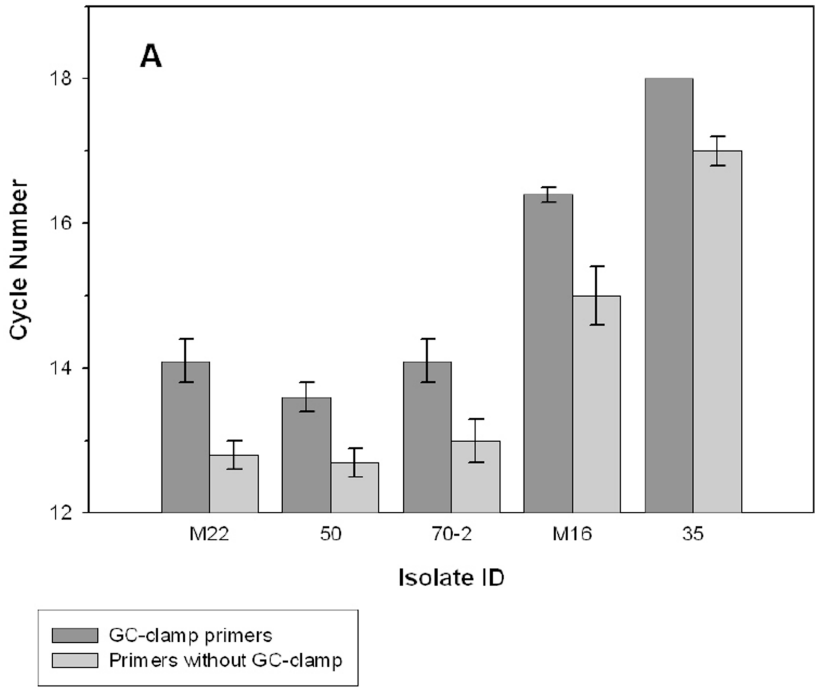


Fig. 3.

DGGE profile from amplifications of equal amounts of template DNA from paired isolates. Lanes 1–5: Amplicons of isolates amplified individually included to identify migration position of isolate-specific bands as labeled. Pairwise amplifications in lanes 6–13: 6, 70-2 and M22; 7, 70-2 and 50; 8, M22 and 50; 9, M22 and M16; 10, 50 and M16; 11, M22 and 35; 12, 50 and 35; 13, M16 and 35; B, blank. Nonspecific bands in the upper portion of the gel in lanes 6–13 are presumed to be heteroduplexes formed as a result of the reannealing of similar, but nonidentical single-stranded products.



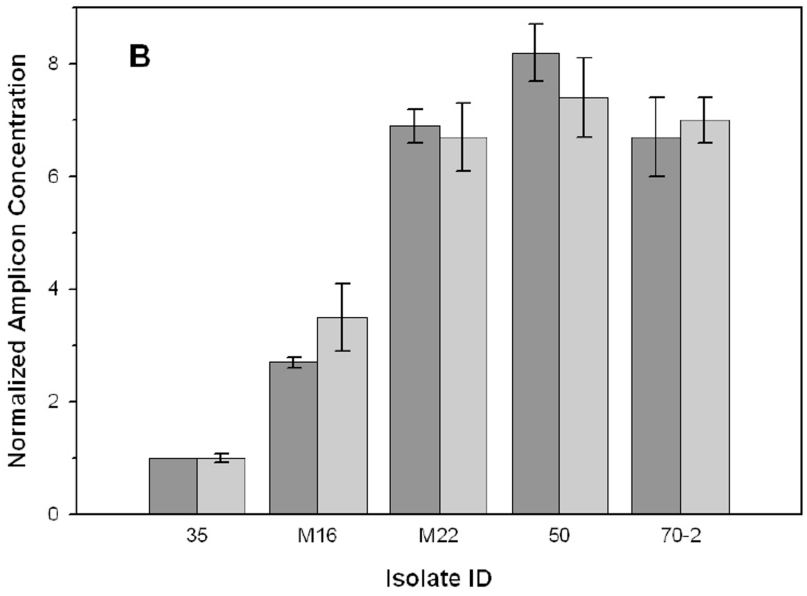


Fig. 4.

Analysis of cycle threshold (CT) values from qPCR of equal concentrations of template DNA from each isolate. (A) CT values for isolates amplified using the primer sets with and without the GC-clamp. CT values represent the average of triplicate reactions for each isolate using 11.5 ± 0.6 ng template DNA in 25 μl reactions and analyzed at the user defined threshold points of 20 RFU for the GC-clamp primer set and 90 RFU for the primer set without the GC-clamp. Error bars represent the standard deviations and the standard deviation for isolate 35 amplified with the GC-clamp primers was 0. (B) Amplicon concentrations of each isolate normalized with respect to the fluorescence value for 35 at the CT value identified in Fig. 4A.

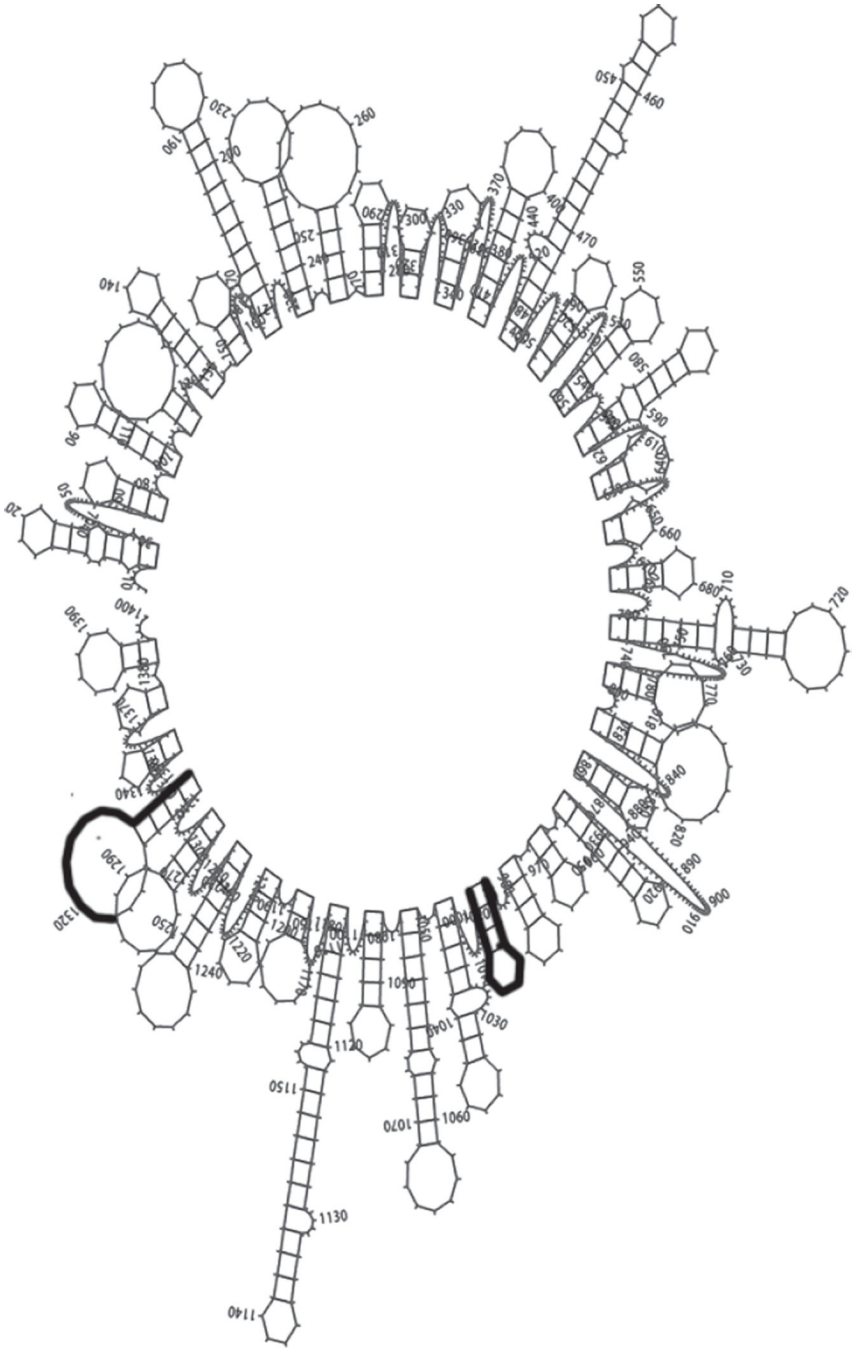


Fig. 5.
Most probable secondary structure for isolate 70-2 at the annealing temperature of 55°C. Highlighted areas represent the primer binding sites. Similar secondary structures were present at the primer binding sites for isolates 50, M22, M16 and 35 at 55°C.

Table 1

Genetic characterization of environmental isolates evaluated in this study

Isolate ID (accession no.)	Closest Isolate ID[a] (% homology)	DGGE amplicon		Phylum
		GC content (%)	Length (bp)	
71a (AY177362)	*Bacillus pumilis* (99%) Glacial ice bacterium G500K-16 (99%)	53.6	351	*Firmicutes*
70-2 (AY177375)	β-*proteobacterium* Wuba68 (96%) *Janthinobacterium agaricidamnosum* SAFR-022 (96%)	54.1	346	β-*Proteobacteria*
M22 (AY864081)	*Ralstonia detusculanense* (99.93%) *Ralstonia* sp. APF11 (99.93%)	55.0	346	β-*Proteobacteria*
50 (AY177368)	*Ralstonia* sp. S23 (99%)	56.2	346	β-*Proteobacteria*
M16 (AY177365)	*Rhizobia giardinii* (99%) Uncultured soil bacteria clone G9-284-4 (99%)	56.5	351	α-*Proteobacteria*
35 (AY177358)	*Methylobacterium* sp.F18 (96%)	57.0	351	α-*Proteobacteria*

[a]: Closest identification based on similarity to 16S rDNA sequences in the GenBank database using the BLAST algorithm.

Table 2

Summary of mixed template amplification patterns. The effect of the presence of an equal concentration of a secondary isolate on amplification of the primary isolate. Product quantification based on DGGE results.

Primary isolate	Secondary isolate	Amplification Status of Primary Isolate[a]	Amplicon Sequence Identity 1° and 2° Isolates[b] (%)
71a	70-2, M22, 50, M16	-	83.7–87.7
71a	35	+	86.3
35	71a	++	86.3
35	70-2, M22, 50, M16	+	82.4–93.7
70-2	M22, 50, M16, 35, 71a	++	83–96.5
M22	50, M16, 35, 71a	++	83.7–98.3
M22	70-2	+	96
50	70-2, M22, M16, 35, 71a	++	83–98.3
M16	71a, 35	++	87.7, 93.7
M16	70-2, M22, 50	+	84.7–85.3

[a]: symbols represent relative band intensity in pairwise amplifications; -, no visible band for primary isolate when amplified with any of secondary isolates; +, primary isolate amplified, but band intensity less than secondary isolates; ++, primary isolate amplified with band intensity equivalent to or greater than secondary isolates

[b]: sequence identity of 346–351 bp amplicons