Explorative screening of complex microbial communities by real-time 16S rDNA restriction fragment melting curve analyses

Knut Rudi, Beate Skånseng, and Signe Marie Drømtorp
MATFORSK Norwegian Food Research Institute, Ås, Norway

BioTechniques 39:116-121(July 2005)

We have developed restriction fragment melting curve analyses (RFMCA), which is a novel method for the real-time analysis of microbial communities. The major advantage of RFMCA compared to, for example, terminal restriction fragment length polymorphism (T-RFLP) or temperature/denaturing gradient gel electrophoresis (TGGE/DGGE) is that the physical separation of DNA fragments is avoided. The RFMCA detection is done by melting point analyses in closed tube systems, which enables high-throughput applications. The robustness of RFMCA was demonstrated by analyzing both mixtures of known samples and the microbial communities in the cecal content of poultry. Our conclusions are that RFMCA is robust, gives a relatively high resolution, and has the potential for high-throughput explorative screenings of microbial communities and large clone libraries.

INTRODUCTION

We have developed a novel explorative approach to describe microbial communities. The basic idea of restriction fragment melting curve analyses (RFMCA) is to use fragment melting point differences rather than size separation to analyze patterns of restriction enzyme-cut DNA from complex samples. The major benefits of RFMCA are that the entire analysis can be done in a single tube and that the approach is suitable for high-throughput. RFMCA is explorative, unlike other real-time melting point assays, which are designed for detecting only specific bacteria or bacterial groups (1). Currently, the most widely used explorative methods to describe microbial communities are terminal restriction fragment length polymorphism (T-RFLP), temperature/denaturing gradient gel electrophoresis (TGGE/DGGE), analyses of clone libraries, or density gradient centrifugation (2–5). Common to these explorative methods is that they are based on the physical separation of DNA fragments. Methods based on physical separation, however, are relatively complicated and cannot be easily adapted for high-throughput applications.

Our general knowledge about microbial communities is still relatively limited (6,7). One of the major limiting factors is the type of method used for gaining information about the communities (8). What is still lacking are explorative screening methods to analyze large sample sets. Analyses of large sets of communities are necessary both for the generalization of observations and to span the diversity of microorganisms in a given habitat (9). Explorative screenings may also be used to identify samples with divergent microbial communities that need further characterization.

The aim of our work was to evaluate the robustness and discriminatory power of RFMCA with respect to its suitability as a high-throughput screening method. This was done both by the analyses of samples containing known mixtures of bacteria and by in-depth comparative analyses of two closely related microbial communities from the cecum of chicken.

MATERIALS AND METHODS

DNA Purification from Cecal Samples

Cecal samples from two chicken flocks raised in the eastern part of Norway in August 2003 were used for the optimization and evaluation of the robustness of the RFMCA method. The flocks were raised by two different producers (abbreviated W and M) under similar conditions (in standard broiler houses) and feeding regimes (Felleskjøpet AS, Oslo, Norway).

Immediately after slaughter, the ceca were transported on ice to the test laboratory and stored at -40°C. After thawing, 50 mg/mL cecum content was suspended in 4 M guanidine thiocyanate (GTC). Two-fold dilution series (0, 1:2, 1:4, and 1:8) in 4 M GTC were made, and each dilution was processed in duplicate by transferring 500 µL to sterile FastPrep® tubes (Qbiogene, Carlsbad, CA, USA) containing 250 mg glass beads (106 µm and finer; Sigma, Steinheim, Germany). The samples were homogenized for 80 s in a FastPrep Instrument (Qbiogene). DNA purification was performed using MagPrep® silica particles (Merck, Darmstadt, Germany), following the manufacturer's recommendations in a Biomek® 2000 Workstation (Beckman Coulter, Fullerton, CA, USA) (10).

PCR Amplification

16S rRNA gene sequences were amplified using universal primers 5′-TCCTACGGGAGGCAGCAGT-3′ (forward) and 5′-GGACTACCAGGGTATCTATTCCTGTT-3′ (reverse). The primers amplify the region from 331 to 797 in the *Escherichia coli* 16S rRNA sequence (11). The forward primer was labeled with 6-FAM, and the reverse primer was labeled with TAMRA for the T-RFLP analyses, while unlabeled primers were used for DNA sequencing and RFMCA.

The 25-µL reactions contained 1× AmpliTaq Gold® reaction buffer (Applied Biosystems, Foster City, CA, USA), 1 mM $MgCl_2$, 1 mM dNTPs, 1 µM of each primer, and 1 U AmpliTaq Gold DNA polymerase (Applied Biosystems). The amplification profile used was 35 cycles of 95°C for 30 s, 65°C for 30 s, and 72°C for 45 s. The enzyme was activated and target DNA denaturated for 10 min at 95°C prior to amplification, and an extension step for 7 min at 72°C was included after the amplification. The reactions were

performed using a GeneAmp® PCR System 9700 (Applied Biosystems).

Cloning and DNA Sequencing

The TOPO TA Cloning® kit with TOP10 One Shot® chemically competent cells was used for cloning (both from Invitrogen, Carlsbad, CA, USA). Transformation of the cells was performed according to the manufacturer's instructions. The Rapid One Shot® chemical transformation protocol was used (Invitrogen). Plasmids from the positive colonies were isolated by resuspending a colony in 30 µL water, heating to 99°C for 5 min, removing the cell debris by centrifugation at 16,060× *g* (Biofuge® Fresco; Kendro Laboratory Products, Asheville, NC, USA) for 1 min, and transferring 25 µL to a new tube. The insert was amplified with the 5′-CGCCAGGGTTTTCCCAGTCACGACG-3′ (HU) and 5′-GCTTCCGGCTCGTATGTTGTGTGG-3′ (HR) primers, which are specific for the vector. Amplification was 30 cycles of 95°C for 4 min and then at 95°C for 15 s, 65°C for 30 s, and 72°C for 1 min. The reaction was ended with an extension step for 7 min at 72°C.

The presequencing reaction included treating 8 µL of the PCR product with 10 U exonuclease I (Amersham Biosciences, Piscataway, NJ, USA) and 2 U shrimp alkaline phosphatase (Amersham Biosciences) at 37°C for 15 min. The enzymes were inactivated by heating to 80°C for 15 min. Sequencing was performed using the Big Dye™ Terminator v 2.0 Cycle Sequencing Kit (Applied Biosystems) on an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems). The preparation of the sequencing mixture was performed as recommended by the manufacturer.

Restriction Enzyme Digestion

Five microliters of each of the amplification products were digested using a restriction enzyme mixture (10 U each of *Msp*I, *Alu*I, *Mse*I, and *Rsa*I) in a total volume of 20 µL 1× NEB buffer 2 (New England BioLabs, Beverly, MA, USA) at 37°C for 8 h, followed by an enzyme inactivation at 65°C for 5 min. The same

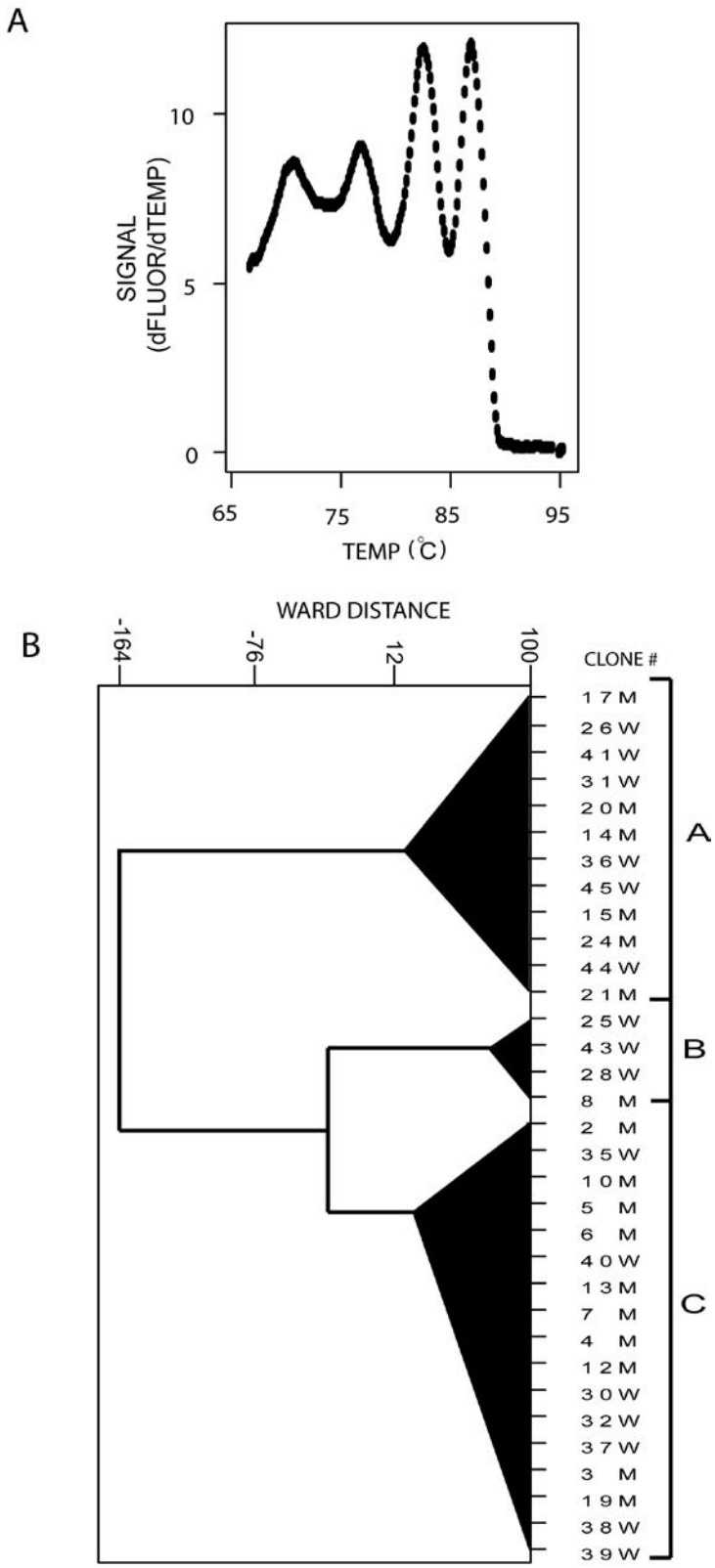

Figure 1. Cluster analyses for the RFMCA patterns for cloned 16S rDNA sequences. (A) The RFMCA pattern for clone 17 M is shown as an example of the data used for cluster analyses. dFLUOR/dTEMP, change in fluorescence signal relative to temperature. (B) The clustering was performed using the Ward algorithm for linkage and correlation distances measures. The CLONE # indicates from which sample the clone was obtained. RFMCA, restriction fragment melting curve analyses.

Table 1. Evaluation and Optimization of RFMCA Parameters

Parameters	Conditions Tested	Optimum	Comments
Temperature (°C)	4–95	65–92	Irreproducible signals below 65°C, and all fragments were melted above 92°C.
DSMO (%)	0, 0.5, 1, 3	0	DSMO gave diffuse peak patterns.
SSC (×)	0, 0.5, 1, 10	0	SSC gave diffuse peak patterns.
Restriction Enzymes	*Alu*I, *Msp*I, *Mse*I, and *Rsa*I	All enzymes	Combination of all four enzymes gave the best resolution.

The optimization was done on a random set of six DNA segments cloned from cecal samples. The analyses were run in triplicate. DMSO, dimethyl sulfoxide; SSC, standard saline citrate.

approach was used for both the RFMCA and T-RFLP samples.

RFMCA Melting

For RFMCA, SYBR® Green I stain 10,000× stock solution (Molecular Probes, Eugene, OR, USA) was added to the restriction enzyme-cut reactions to a concentration of 10× in a total volume of 25 µL. The melting reactions were performed using either an ABI PRISM 7700 Sequence Detection System or the 7900HT system (both from Applied Biosystems). Dissociation Curves 1.0 software (Applied Biosystems) was used to analyze the melting patterns for the 7700 data, while SDS 2.2 software (Applied Biosystems) was used to analyze the data generated with the 7900HT system.

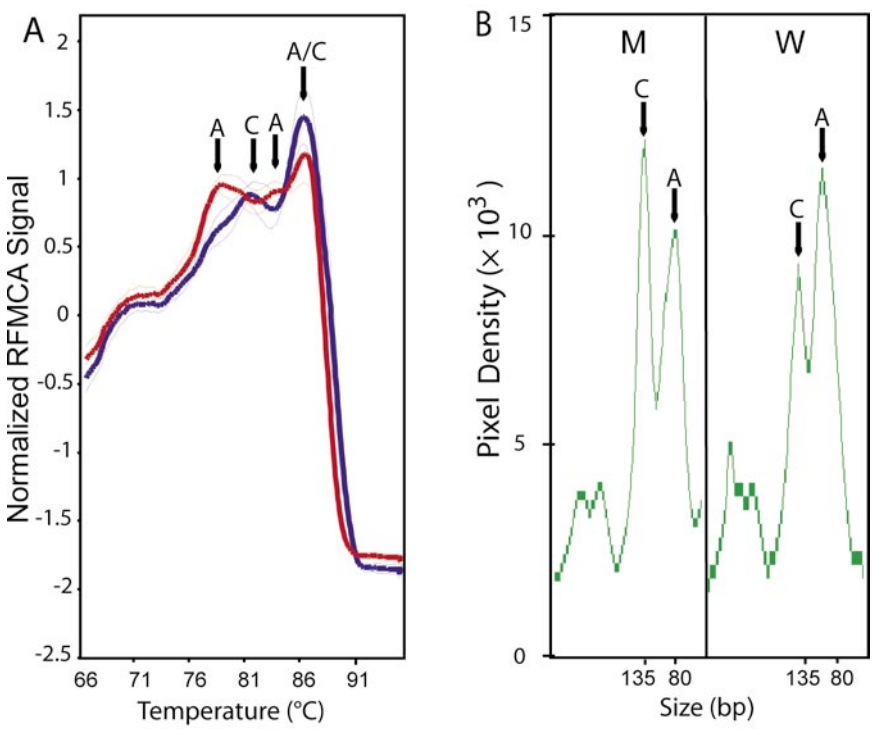

Figure 2. RFMCA and T-RFLP for the W and M samples. (A) RFMCA melting pattern for the W (red) and the M (blue) samples. The thin lines represent the standard deviation (eight samples for both M and W). The peaks for bacteria belonging to the A and C groups are marked with arrows. (B) The T-RFLP results (TAMRA-labeled reverse primer) for the W and the M samples are shown. The two main discriminatory bands for the A and C groups are marked. RFMCA, restriction fragment melting curve analyses; T-RFLP, terminal restriction fragment length polymorphism.

T-RFLP Size Separation

The T-RFLP samples were separated in a 3% agarose gel at 100 volts for 1 h. The detection was performed using a Typhoon™ 8600 Variable Mode Imager (Amersham Biosciences). Finally, quantification was performed using ImageMaster™ Total Lab software (Amersham Biosciences).

Phylogenetic Reconstruction and Cluster Analyses

Sequences of representative strains were selected from the GenBank® nucleotide sequence database (March 2004) based on searches with the Basic Local Alignment Search Tool (BLAST) program (www.ncbi.nlm.nih.gov) and aligned with sequences obtained in this study using Clustal X (12). The alignments were then manually edited using the program BioEdit (13). A phylogenetic tree was constructed using Tamura Nei distances (14) and the Minimum Evolution algorithm provided in the MEGA 2 software package (15). Statistical support for the branches in these trees was obtained by bootstrap analysis with 500 replicates.

The RFMCA data were clustered using correlation coefficient distances, and Ward linkage for dendrogram construction (Minitab v. 14; Minitab, State College, PA, USA). The RFMCA input data were normalized by subtracting the mean and dividing by the standard deviation for each data point prior to the cluster analyses.

Statistical Analyses

We used the two-tailed *t*-tests and the tests for standard deviation provided in the Minitab v. 14 software package. The multivariate statistical analyses were done using the Unscrambler® v. 9.0 software (Camo, Woodbridge, NJ, USA). We used principal component analyses (PCA) and partial least square regression (PLSR) in combination with the prediction tools provided in the Unscrambler software. We performed PCA and PLSR analyses using full cross-validation with centered data. The variables were weighted according to their standard deviations. The prediction was done by first building a PLSR model using a calibration set, and then the model was validated using an independent sample set. The input data were normalized by subtracting the mean and dividing by the standard deviation. The loading for the initial solution was computed from the data.

RESULTS AND DISCUSSION

Optimizing the Resolution of RFMCA

We wanted to identify the window for RFMCA with the highest possible resolution. The parameters tested were restriction enzyme combinations, melting temperature range, and stringency. The results are summarized in Table 1.

The requirements for the restriction enzymes applied for RFMCA should be that they are compatible with the same buffer system and that they are frequent cutters. The four restriction enzymes *Msp*I (C▼CGG), *Alu*I, (AG▼CT), *Mse*I (T▼TAA), and *Rsa*I (GT▼AC) met these criteria. These enzymes were used accordingly in the optimization of the RFMCA method. The resolution for samples cut with single enzymes was lower than the samples cut with all four enzymes. The theoretical average fragment size of 256 bp for the samples cut by single enzymes is probably too large to be separated by melting point analyses. The theoretical average size of the fragments for the combinations of all four enzymes is 64 bp, which is probably within the range that can be

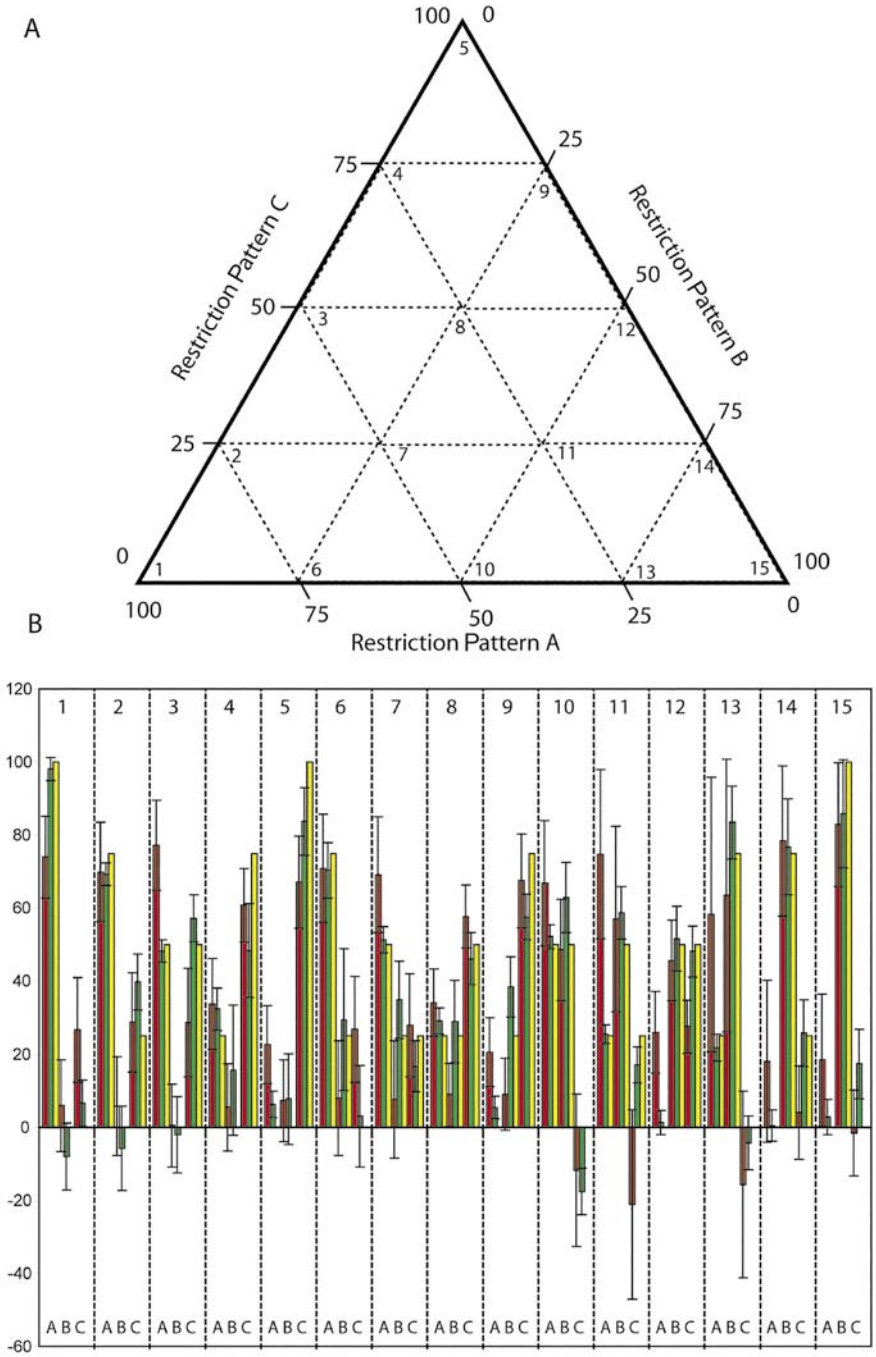

Figure 3. Comparison of RFMCA and T-RFLP for mixtures of known components. (A) Clones with restriction patterns corresponding to the major groups of patterns A (17 M), B (43 W), and C (13 M) identified in Figure 1 were mixed according to the experimental design shown. The numbers within the triangle indicate the numbering of the samples. (B) Predictions for the validation set of samples of the T-RFLP (red bars) and the RFMCA (green bars) data for the restriction patterns A, B, and C. The yellow bars show the expected values. The numbering corresponds to the numbers in panel A. The standard deviations are determined from jackknife cross-validation. RFMCA, restriction fragment melting curve analyses; T-RFLP, terminal restriction fragment length polymorphism.

separated by melting point analyses.

The best differentiation and reproducibility with boundaries of melting peaks of ±2.5°C were obtained for the melting temperatures ranging 65°–92°C (see Figure 1A for a typical pattern). The melting patternas obtained below 65°C were relatively unstable, possibly due to variable accumulation of small fragments such as primer dimers for these temperatures. All the fragments were melted above 92°C, and thus no useful information was obtained above that temperature.

We investigated whether modifying the stringency could increase the resolution of RFMCA. The stringency was lowered by the addition of high salt standard saline citrate (SSC) solution, while the stringency was increased by adding the cosolvent dimethyl sulfoxide (DMSO) (16). Both SSC and DMSO led to less distinct melting peak patterns and lowered resolution (Table 1). We concluded that SSC and DMSO did not improve the performance of RFMCA. These compounds were therefore not used further.

The final optimized RFMCA protocol involved cutting with all four restriction enzymes and melting in the range 65°–95°C for 20 min, while only data for the temperature range of 65°–92°C were used for the subsequent discrimination analyses.

Application of RFMCA for Characterization of Complex Communities in Chicken Cecal Samples

The reproducibility and discriminatory power of RFMCA were evaluated by in-depth comparisons of the two closely related microbial communities W and M (see Materials and Methods for details). An initial characterization of the diversity in the samples was performed by cloning and sequencing of partial 16S rRNA gene sequences. The cloned fragments were subsequently subjected to RFMCA. Three major RFMCA patterns (A to C) were identified from these clones using correlation coefficient distances, and Ward linkage for dendrogram construction (Figure 1B).

There was a good correspondence between RFMCA and DNA sequence classification (data not shown). Basically, RFMCA pattern A corresponded to *Clostridiales*, B corresponded to *Bacteroidales*, while C corresponded to *Bacillales*, *Lactobacillales*, and uncultured Gram-positive bacteria.

The RFMCA principle was further evaluated by direct analyses of the microbial communities in the cecal content from the W and M samples. Eight independent DNA purifications consisting of duplicate analyses of each of the dilutions (0, 1:2, 1:4, and 1:8) described in the Materials and Methods section were analyzed for each of the samples (Figure 2A). Diagnostic peaks for the A groups of bacteria were identified in the W sample, while there were peaks corresponding to the C group of bacteria in the M sample (see arrows in Figure 2A). We were also able to detect clear differences in the microbial communities using PCA. The first principal component gave an average score of 1.82 ± 0.71 for W and -2.64 ± 0.38 for M, respectively. These scores were found significantly different using a two-tailed *t*-test (t = 15.37 and $P < 0.0005$).

We did a theoretic evaluation of the expected restriction fragments identified by T-RFLP. We identified fragments of 146 and 124 bp for clones belonging to cluster C, while the expected fragments for clones belonging to cluster A were 87 and 72 bp. We identified two T-RFLP bands that were discriminatory between the W and M samples (Figure 2B, t = 4.87 and P = 0.001), which probably correspond to the theoretically identified 146 and 124 bp, and the 87 and 72 bp fragments, respectively. A resolution of approximately ±10 bp was determined for our T-RFLP by comparison with known molecular weight standards (data not shown).

Evaluation of RFMCA for Defined Samples

Representative samples with restriction digestion patterns resembling pattern A, B, and C were chosen for evaluating the performance of RFMCA and T-RFLP (Figure 3). The samples were mixed according to the experimental design shown in Figure 3A. Models were first built using a calibration set of data. These models were then evaluated using a new set of independent validation data (Figure 3B). These analyses showed that RFMCA overall gave a good accuracy and precision (Figure 3B). The misclassification for the RFMCA data was <15%. This example also shows that it should be possible to quantify the composition of mixed bacterial populations, given that the patterns for the pure components are known. Such an application would be particularly important in process or quality control where known mixtures of bacteria are used, such as in food fermentation. The reason for the relatively high error rate for the T-RFLP data, however, may be due to relatively low resolution of the agarose gel electrophoresis applied. Our T-RFLP results may not be representative for other separation techniques such as high-throughput capillary gel electrophoresis.

Future Potential

Future implementation of the method includes databases of patterns from pure cultures or cloned samples. These can be matched with the patterns obtained for the communities, such as by multivariate regression (e.g., principal component regression) in order to obtain a semiquantitative description of the composition of the communities analyzed. Hopefully, the implementation of novel explorative screening methods will help to better understand the important and emerging field of explorative microbial community analyses (17).

ACKNOWLEDGMENTS

This work was supported by a Norwegian research levy on agricultural products and by the grant nos. 153088/110 and 157596/130 from the Research Council of Norway.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. **Apajalahti, J.H., A. Kettunen, M.R. Bedford, and W.E. Holben.** 2001. Percent G+C profiling accurately reveals diet-related differences in the gastrointestinal microbial community of broiler chickens. Appl. Environ. Microbiol. *67*:5656-5667.
2. **Acinas, S.G., V. Klepac-Ceraj, D.E. Hunt, C. Pharino, I. Ceraj, D.L. Distel, and M.F. Polz.** 2004. Fine-scale phylogenetic architecture of a complex bacterial community. Nature *430*:551-554.
3. **Fukushima, H., Y. Tsunomori, and R. Seki.** 2003. Duplex real-time SYBR green PCR assays for detection of 17 species of food- or waterborne pathogens in stools. J. Clin. Microbiol. *41*:5134-5146.
4. **Domann, E., G. Hong, C. Imirzalioglu, S. Turschner, J. Kuhle, C. Watzel, T. Hain, H. Hossain, et al.** 2003. Culture-independent identification of pathogenic bacteria and polymicrobial infections in the genitourinary tract of renal transplant recipients. J. Clin. Microbiol. *41*:5500-5510.
5. **Muyzer, G. and K. Smalla.** 1998. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. Antonie Van Leeuwenhoek *73*:127-141.
6. **Pace, N.R.** 1997. A molecular view of microbial diversity and the biosphere. Science *276*:734-740.
7. **Venter, J.C., K. Remington, J.F. Heidelberg, A.L. Halpern, D. Rusch, J.A. Eisen, D. Wu, I. Paulsen, et al.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science *304*:66-74.
8. **Theron, J. and T.E. Cloete.** 2000. Molecular techniques for determining microbial diversity and community structure in natural environments. Crit. Rev. Microbiol. *26*:37-57.
9. **Amann, R.I., W. Ludwig, and K.H. Schleifer.** 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. Microbiol. Rev. *59*:143-169.
10. **Skånseng, B. and K. Rudi.** 2004. Automated DNA purification from bacteria in the poultry gastrointestinal tract for large-scale microbial community analyses. AFAC workshop, Alternatives to feed antibiotics and anticoccidials in the pig and poultry meat production, 19-20 September 2004, Århus, Denmark.
11. **Nadkarni, M.A., F.E. Martin, N.A. Jacques, and N. Hunter.** 2002. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. Microbiology *148*:257-266.
12. **Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins.** 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *25*:4876-4882.
13. **Hall, T.A.** 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. *41*:95-98.
14. **Tamura, K. and M. Nei.** 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. *10*:512-526.
15. **Kumar, S., K. Tamura, I.B. Jakobsen, and M. Nei.** 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics *17*:1244-1245.
16. **Sambrook, J. and D. Russel.** 2000. Molecular Cloning: a Laboratory Manual, 3rd ed. CSH Laboratory Press, Cold Spring Harbor, NY.
17. **Curtis, T.P. and W.T. Sloan.** 2004. Prokaryotic diversity and its limits: microbial community structure in nature and implications for microbial ecology. Curr. Opin. Microbiol. *7*:221-226.

Received 29 November 2004; accepted 12 February 2005.

Address correspondence to Knut Rudi, MATFORSK Norwegian Food Research Institute,Osloveien 1, 1430 Ås, Norway. e-mail: knut.rudi@matforsk.no