APPLIED AND ENVIRONMENTAL MICROBIOLOGY, June 2003, p. 3317–3326 Vol. 69, No. 6
0099-2240/03/$08.00+0 DOI: 10.1128/AEM.69.6.3317–3326.2003

Fluorescent Heteroduplex Assay for Monitoring *Bacillus anthracis* and Close Relatives in Environmental Samples

Lori Merrill, Jennifer Richardson, Cheryl R. Kuske, and John Dunbar*

Biosciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico

Received 30 September 2002/Accepted 5 March 2003

A fluorescent heteroduplex method was developed to assess the presence of 16S rRNA gene (rDNA) sequences from *Bacillus anthracis* and close relatives in PCR-amplified 16S rDNA sequence mixtures from environmental samples. The method uses a single-stranded, fluorescent DNA probe, 464 nucleotides in length, derived from a *B. anthracis* 16S rRNA gene. The probe contains a unique, engineered deletion such that all probe-target duplexes are heteroduplexes with an unpaired G at position 343 (ΔG343). Heteroduplex profiles of sequences ≥85% similar to the probe were produced using an ABI 377 sequencer in less than 3 h. The method divides strains of the *Bacillus cereus-Bacillus thuringiensis-B. anthracis* group into two subgroups. Each subgroup is defined by a specific 16S rRNA gene sequence type. Sequence type A, containing one mismatch with the probe, occurs in *B. anthracis* and a small number of closely related clonal lineages represented mostly by food-borne pathogenic isolates of *B. cereus* and *B. thuringiensis*. Sequence type B, containing two mismatches with the probe, is found in the majority of *B. cereus* and *B. thuringiensis* strains examined to date. Sequence types A and B, when hybridized to the probe, generate two easily differentiated heteroduplexes. Thus, from heteroduplex profiles, the presence of *B. cereus-B. thuringiensis-B. anthracis* subgroups in environmental samples can be inferred unambiguously. The results show that fluorescent heteroduplex analysis is an effective profiling technique for detection and differentiation of sequences representing small phylogenetic or functional groups in environmental samples.

Microbiologists are often interested in monitoring functionally or phylogenetically related microbial species amidst the hundreds or thousands of species that occur in natural environments. In such cases, gene sequences from members of the group of interest can be obtained by PCR amplification, and the diversity of sequences within the amplified mixture can be either determined directly by cloning and sequencing or assessed indirectly by mutation scanning techniques. Cloning and sequencing provide the most definitive information. However, for comparative analysis this approach is often suitable for only one or a few samples due to the expense and the large number of clones required for adequate characterization of complex assemblages (12). Scanning techniques such as single-strand conformational polymorphism analysis, denaturing gradient gel electrophoresis (DGGE), and terminal restriction fragment (TRF) analysis offer less resolution than cloning and sequencing but enable rapid analysis of large sample sets by assessing the diversity of thousands of individuals (from a sample) in a single profile.

Heteroduplex analysis is a sensitive scanning technique for documenting differences between closely related sequences (14, 17, 20, 32, 38, 40, 46). In fact, a temperature-modulated, high-pressure liquid chromatography version of the basic heteroduplex method is rapidly becoming a tool of choice for discovering variation caused by single nucleotide polymorphisms (SNPs) (4, 19, 23, 30, 48). Heteroduplex analysis is a combination of DNA probe hybridization and electrophoresis. Hybridization of a probe with a complementary, nonidentical target creates heteroduplexes (duplexes with mismatches) that can be resolved from homoduplexes (perfectly matched duplexes) by electrophoresis through a gel that separates DNA molecules according to length and shape (7, 11, 25). Nucleotide mismatches that sufficiently distort the shape of a DNA molecule reduce the electrophoretic migration rate. Since more mismatches generally produce more distortion, the location of a heteroduplex in an electrophoretic profile coarsely reflects the relative similarity of the target sequence and probe (15).

Heteroduplex analysis has been widely used in viral epidemiology (5, 11) and analysis of human genetic disorders (14, 18, 23, 32). The method has also been used to identify genetic differences in closely related microbial strains (10, 15, 21, 24, 33, 37, 43, 45) and to assess the similarity of sequences in clone libraries (8, 16). However, use of the method has been largely confined to analysis of purified targets, or, in other words, pairwise comparisons of a double-stranded reference sequence (a probe) with a double-stranded sample sequence (target). In such pairwise comparisons, each profile typically contains two heteroduplexes corresponding to the distinct molecules formed by the sense and antisense strands of the target. This feature greatly complicates analysis of complex sequence mixtures (15) but can be circumvented by the use of fluorescent chemistry.

Here we describe a fluorescent heteroduplex method suitable for analyzing sequence mixtures. We applied the method to analysis of amplified rDNA to assess the presence of *Bacillus anthracis* and close relatives in environmental samples. Preliminary analysis of 16S rRNA gene sequences from a variety of *B. anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* isolates showed that a single polymorphism at position 980 differentiated *B. anthracis* from most *B. cereus* and *B. thuringiensis* strains. The region of the 16S gene containing this polymorphism was successfully used for heteroduplex analysis.

* Corresponding author. Mailing address: M888 Biosciences Div., Los Alamos National Laboratory, Los Alamos, NM 87545. Phone: (505) 667-8806. Fax: (505) 665-3024. E-mail: dunbar@lanl.gov.

MATERIALS AND METHODS

Bacterial strains. Genomic DNA stocks from a variety of *Bacillus* species were generously provided by Paul Jackson (Los Alamos National Laboratory). DNA stocks were obtained for *B. anthracis* K2991, *B. cereus* ATCC 43881, *B. thuringiensis* ATCC 33679, *B. subtilis* ATCC 6051, *B. mycoides* ATCC 6462, *B. megaterium* ATCC 15127, *B. amyloliquefaciens* ATCC 23350, *B. flexus* ATCC 49095, *Paenibacillus glucanolyticus* ATCC 49278, and *Brevibacillus laterosporus* ATCC 64. In addition, genomic DNA stocks were obtained for 67 *B. cereus* and *B. thuringiensis* strains (a list is available upon request). These 67 strains originated from three separate culture collections assembled by the Food Research Institute (University of Wisconsin—Madison, Madison), the U.S. Department of Agriculture (Peoria, Ill.), and Anna-Brit Kolstø (Biotechnology Centre of Oslo and Institute of Pharmacy, University of Oslo, Oslo, Norway).

DNA extraction from soil samples. Soil samples were collected from a soil near Las Cruces, N.Mex., a pinyon-juniper woodland in Cosnino, Ariz., and an urban lawn in Salt Lake City, Utah. DNA was extracted from the three samples by a modification of the method described by Kuske et al. (22). Briefly, 0.5 g of soil was homogenized with glass beads in 1 ml of Tris-EDTA and centrifuged to pellet insoluble debris, and the supernatant containing DNA was ethanol precipitated. The DNA was dissolved in 10 mM Tris (pH 8) and then purified further by passage through Sephadex G-200 (Sigma) microcolumns self-prepared in 96-well plates.

PCR. 16S rRNA gene fragments approximately 465 bp in length were amplified for heteroduplex analysis using the primers Bac629F (5′-AGGGTCATTG GAAACTGGG, *E. coli* positions 629 to 647 [K. H. Wilson, personal communication]) and Bac1091R (5′-AACCCAACATCTCACGAC, *E. coli* positions 1091 to 1073, designed in this study). This primer set enabled amplification of 16S rRNA gene sequences primarily from gram-positive species, in particular members of the *Bacillus-Lactobacillus-Staphylococcus* subgroup. Each 50-μl PCR mixture contained 30 mM Tris (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM deoxynucleoside triphosphates, a 0.1 μM concentration of each primer, and 1.25 U of Ampli*Taq* LD polymerase (Perkin-Elmer). For soil DNA templates, 1.875 U of Ampli*Taq* LD polymerase was used for amplification, and reaction mixtures were supplemented with 10 μg of bovine serum albumin (Boehringer Mannheim). Cycling conditions were as follows: 3 min of denaturation at 94°C; 35 cycles of 42°C for 45 s, 72°C for 110 s, and 94°C for 30 s; and a final cycle of 42°C for 45 s and 72°C for 5 min. Products were purified with a QIAquick PCR cleanup kit (Qiagen, Inc. Chatsworth, Calif.) and quantified by agarose gel electrophoresis and ethidium-bromide staining.

Construction of a *B. anthracis* reference probe. A DNA probe for heteroduplex analysis of *Bacillus* species was constructed from a *B. anthracis* 16S rRNA gene by introducing a 1-bp deletion at position 970. To accomplish this, two overlapping fragments of 16S ribosomal DNA (rDNA) from *B. anthracis* were separately amplified by PCR. The first fragment was amplified with primers Bac629F and Bac1214R (5′-ATAAGGGGCATGATGATTTGAC, *E. coli* positions 1214 to 1193; designed in this study). A second fragment was amplified with Bac962F (5′-CGAAGCAAGCGAAGAACCT, containing a C deletion relative to the *B. anthracis* sequence; designed in this study) and Bac1230R (5′-CGTGTGTAGC CCAGGTCATA, *E. coli* positions 1230 to 1212 [Wilson, personal communication]). These two amplicons were combined in a 1:50 ratio, amplified for 5 cycles without the addition of primers, and then amplified for 15 additional cycles with the addition of primers Bac629F and Bac1230R. The resulting PCR product was cloned into *Escherichia coli* by using a TOPO-TA cloning kit (Invitrogen Corporation, Carlsbad, Calif.). An *E. coli* clone with an appropriate insert was purified by subculturing on selective medium, and the constructed mutation was confirmed by sequencing. Plasmid DNA was subsequently extracted from the *E. coli* clone and used routinely as a template for PCR amplification of the heteroduplex probe.

For heteroduplex analysis, the cloned gene probe was PCR amplified with the primer Bac629F-PHO (containing a 5′ phosphate for recognition by lambda exonuclease) and Bac1091R-FAM (5′ end labeled with 6-carboxy fluorescein). Following purification, approximately 800 ng of amplified probe was digested in a 40-μl reaction mixture volume with 20 U of lambda exonuclease (Epicentre, Madison, Wis.) at 37°C for 2 h, yielding >95% fluorescently labeled single-stranded probe. The single-stranded probe DNA was purified with a QIAquick PCR cleanup kit prior to use.

Heteroduplex assay. For each hybridization reaction, approximately 15 ng of single-stranded fluorescent probe was mixed with a DNA target. NaCl was added to a final concentration of 100 mM. The mixture was heated to 94°C for 4 min, ramped to 60°C at 1°C per s, held at 60°C for 1.5 h, and then ramped to 25°C prior to electrophoresis. A 1-μl aliquot of the heteroduplex mixture was combined with 1.5 μl of loading buffer (1 μl of 15% Ficoll, 0.25 μl of 25 mM EDTA–3% blue dextran, and 0.25 μl of ABI GeneScan TAMRA [6-carboxytetramethyl-rhodamine] 2500 size standard). The samples were loaded onto a 0.48× concentration of mutation detection enhancement (MDE) gel (FMC Bioproducts, Rockland, Maine) containing 0.6× Tris-borate-EDTA buffer. Electrophoresis was performed with an ABI 377 DNA sequencer in GeneScan mode at 3,000 V and 40°C with 0.6× TBE buffer for 3.25 h. ABI GeneScan analytical software version 3.1 was used for data collection and fragment analysis. Analysis parameters included the use of a peak detection threshold of 25 fluorescence units, peak minimum half-width of 5 points, local Southern sizing method, and rightmost split peak correction.

TABLE 1. Mismatch scoring for estimation of duplex distortion

Mismatch[a]	Score
ΔG	58.27
ΔC	46.48
ΔT	35
ΔA	18
G:G	10
A:A	10
A:C	8.5
G:A	7.8
T:T	7
T:C	5.97
T:G	4
C:C	3
≥3 M	125
D	60

[a] Δ, gap; M, mismatches; D, destabilized 5- to 8-bp region containing mismatches separated by ≤1 bp.

Clone libraries. 16S rRNA gene fragments amplified with the primers Bac629F and Bac1092R from a New Mexico soil DNA sample were cloned into *E. coli* by using a TOPO-TA cloning kit. A set of 96 white colonies grown on selective medium were transferred to 96-well plates containing 15% glycerol for storage at −70°C. Plasmid DNA was extracted for cycle sequencing reactions. The 16S rRNA gene fragment insert in each plasmid vector was sequenced by using primer T3, Big-Dye terminator cycle sequencing reagents (ABI), and a 3700 capillary DNA sequencer (ABI).

DNA sequence analysis. Clone library sequences and 16S rRNA gene sequences from nearest neighbors were aligned using the Ribosomal Database Project (RDP) sequence alignment tool (27) and visually inspected for alignment errors. Phylogenetic analysis of the aligned sequences was performed with PAUP software (42). A distance matrix consensus tree was constructed by bootstrap analysis with 100 replicates of heuristic searching, 10 rounds of random sequence addition per replicate, and global branch swapping with the tree-bisection-reconnection algorithm. All best trees were saved.

To evaluate the relationship between heteroduplex migration and sequence variation, variation between the heteroduplex probe and 40 target sequences was compared in three ways. First, the total number of mismatches between the probe sequence and a given target sequence was determined from pairwise alignments of probe and target sequences. Second, genetic distance values were obtained with the RDP PHYLIP (3.5c; J. Felsenstein) tool (27) by using the Kimura two-parameter algorithm for calculation of a distance matrix. Third, heteroduplex migration was predicted by calculating distortion scores based on the postulated effects of mismatch types and the modulating impact of mismatch location and local melting temperature (T_m).

To calculate distortion scores, five steps were followed. First, a mismatch effect vector, *M*, was constructed for each DNA sequence. The position and types of mismatches in pairwise alignments of probe and target sequences were catalogued in a Microsoft Excel spreadsheet. Each type of mismatch (12 total) and two types of clustering effects (three or more contiguous mismatches or 5- to 8-bp regions containing mismatches separated by 1 bp or less) were assigned numerical scores indicating their relative impact on duplex distortion (see Table 1). Complementary base pairs were assigned a score of 0. With these values, each heteroduplex sequence was then converted to a vector, *M*, in which each sequence position was represented by a number indicating the degree of distortion expected due to the type of mismatch present.

Second, a location scaling vector, *L*, consisting of weight values between 0 and 1, was created to modulate the effects of mismatches based on their proximity to the end of the DNA sequence. Several different location scaling vectors were constructed and tested in this study. The most effective vector was constructed by

considering changes in the length of a DNA strand (i.e., the shortest distance between ends) as a bend migrates from the midpoint to one end of the strand. For convenience, a 90° angle was considered, so that changes in the length of the hypotenuse (the direct distance between DNA ends) could be easily calculated with the Pythagorean theorem. Location weights were expressed as fractions of the maximum distortion that could occur (i.e., by placing a bend at the midpoint of the DNA fragment), or in other words, as actual fragment length reduction divided by maximum fragment length reduction. The location scaling vector was expressed mathematically as follows:

$$L_i = (l - \sqrt{i^2 + (l - i)^2})/(l - \sqrt{(\tfrac{1}{2})^2 + (\tfrac{1}{2})^2}) \quad (1)$$

where L_i is the location weight at position i ($1 \leq i \leq 465$) and l is the sequence length (465 bp).

Third, a T_m scaling vector, T, consisting of weights between 0 and 1, was created to account for the modulating effect of local T_m. Predicted T_ms for each base pair position of the 465-bp *B. anthracis* 16S rRNA gene fragment were obtained with Oligo software (Molecular Biology Insights, Inc.). The local T_m for each sequence position was then calculated as the average of the T_ms of directly adjacent base pairs. To scale these values, several different scaling vectors were constructed and tested. The most effective vector used a logarithmic function to mimic the type of transitions that are typically observed in T_m profiles as duplex DNAs approach critical transition temperatures. Thus, the T_m scaling vector (T_i) was expressed as $1 + \log[(T_{m\max} - T_{mi})/(T_{m\max} - T_{m\min})]$, where T_{mi} is the local T_m for a given sequence position i and $T_{m\max}$ and $T_{m\min}$ are empirically selected maximum and minimum threshold values for the transition range (i.e., a range of T_ms across which mismatches are postulated to exert differential effects on a DNA duplex). A transition range of 48°C (the lowest local T_m in the *B. anthracis* 16S rRNA sequence fragment) to 61°C was selected as an optimal range for analysis of the sequences used in this study. Negative weights in the T_m scaling vector were set to zero. As a result, mismatches at locations with local T_ms greater than 59.8°C were represented as having no detectable effect on duplex distortion.

Fourth, the vectors M, T, and L were multiplied to obtain a product vector, and the product vector was summed to obtain the net distortion score, d. This yields

$$d = \sum_{i=1}^{l} M_i \cdot L_i \cdot T_i \quad (2)$$

where i is the vector position (or sequence position) and l is the vector length. The predicted migration of a given heteroduplex was calculated by adding the distortion score to the average migration of the homoduplex (461 base pair equivalents [bpe]).

Fifth, predicted heteroduplex migration values were compared to observed migration values, and iterative rounds of trial and error parameter optimization were performed. Scores for mismatch types ΔG, ΔC, G:A, and T:C were defined first in each round of parameter optimization because these types occurred alone in *B. anthracis* and *B. cereus* heteroduplexes and their effects were clearly defined. Weights for other mismatch types were reiteratively adjusted to obtain the best correlation (highest R^2 value for a linear regression), ranking order, and lowest average deviation between predicted and observed migration behavior. The adjustment of scores and parameters was further guided (and constrained) by the need for consistency with previous experimental observations (17, 20, 34, 46) and by rational expectations of the effects of different mismatch types.

RESULTS

Convention for data analysis. The migration of heteroduplexes has typically been reported as a percentage of the migration distance of a homoduplex reference. To provide a clearer indication of the electrophoretic separation of the molecules, we report the location of each heteroduplex as a relative DNA fragment length based on comparison with fluorescent DNA size standards (GeneScan 2500; ABI) run in the same lane as the sample. The heteroduplex positions are reported as base pair equivalents instead of base pairs in order to differentiate heteroduplex positions from true DNA fragment lengths.

Analysis of *B. cereus*, *B. thuringiensis*, and *B. anthracis* strains. Figure 1 shows typical heteroduplex profiles of 16S rDNA amplicons from *B. anthracis* K2991 and *B. cereus* ATCC 43881. In the 16S rRNA gene region targeted for heteroduplex analysis (*E. coli* gene sequence positions 629 to 1091), these two strains differed by a single polymorphism at position 980. Amplicons from these strains produced distinct heteroduplex profiles. Each profile contained a homoduplex DNA fragment (the double-stranded probe) and a single predominant heteroduplex peak. The homoduplex was present as a result of incomplete lambda exonuclease digestion of the probe stock (the stock typically contained <5% residual double-stranded probe). The homoduplex probe was centered at position 461.1 bpe. The probe homoduplex was easily and reliably distinguished from the heteroduplex in each profile due to the engineered deletion at position 343 of the 464-bp probe sequence. For example, in Fig. 1 the heteroduplex consisting of the probe and *B. anthracis* K2991 containing a single mismatch (ΔG343) was centered at bpe position 467.8 while the heteroduplex consisting of the probe and *B. cereus* ATCC 43881 containing two mismatches (ΔG343 and G:A378) was centered at bpe 471.0. These results demonstrated the ability of the heteroduplex method to differentiate the *B. anthracis* K2991 16S rDNA sequence type (type A, containing one mismatch with the probe) and the *B. cereus* ATCC 43881 sequence type (type B, containing two mismatches).

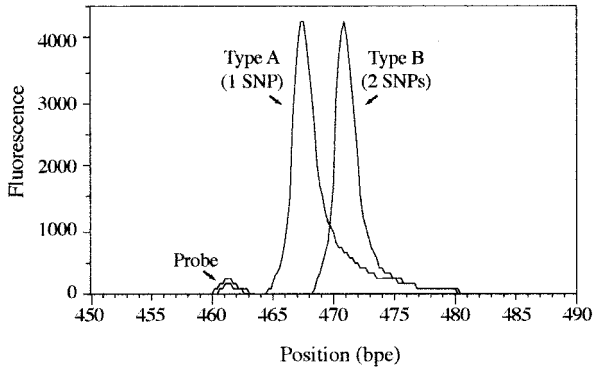

FIG. 1. Overlaid heteroduplex profiles of *B. anthracis* K2991 and *B. cereus* ATCC 43881. The *B. anthracis* 16S rDNA sequence is designated type A, and the *B. cereus* sequence is type B. The number of SNPs occurring in each sequence relative to the heteroduplex probe is indicated. The migration of each duplex is represented relative to DNA size standards and is therefore given in bpe.

The reproducibility of the type A and type B heteroduplex profiles was measured by analysis of 40 replicate profiles of *B. anthracis* K2991 and 14 replicate profiles of *B. cereus* ATCC 43881 (Fig. 2). Among the *B. anthracis* K2991 replicates, the location of the type A heteroduplex peak ranged from 467.0 to 469.0 bpe (average = 467.8, standard deviation [SD] = 0.44). The location of the heteroduplex peak exhibited a bimodal frequency distribution. The cause of the bimodal distribution is at present unknown. Data spanning the range of the distribution were obtained randomly over time within and between gels and were not the result of GeneScan software tracking or analysis errors. The location of the type B heteroduplex peak in 14 replicate profiles of *B. cereus* ATCC 43881 ranged from 470.4 to 471.4 bpe (average = 470.8; SD = 0.23).

To determine the occurrence of sequence types A and B among strains in the *B. cereus-B. thuringiensis-B. anthracis*

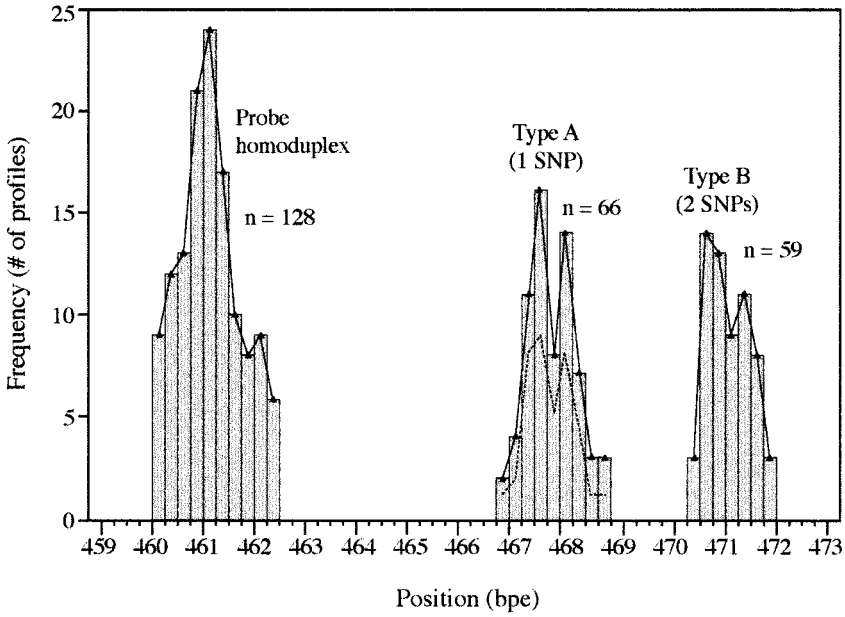

FIG. 2. Frequency distribution of peak locations in heteroduplex profiles of *B. anthracis*, *B. cereus*, and *B. thuringiensis* strains. The number of SNPs occurring in each sequence type relative to the heteroduplex probe is indicated. The dashed line shows the frequency distribution of the location of the heteroduplex peak in 40 replicate profiles of *B. anthracis* K2991.

group, heteroduplex profiles were produced for 65 different *B. cereus* and *B. thuringiensis* strains obtained from three culture collections. Twenty-four *B. cereus* strains yielded heteroduplex profiles indistinguishable from that of *B. anthracis* K2991. The relative location of the heteroduplexes from the *B. cereus* strains ranged from 467.13 to 468.98 bpe (average = 467.97, SD = 0.51). The data are shown in Fig. 2 combined with the *B. anthracis* K2991 data. Based on these results, the 24 *B. cereus* strains appeared to possess sequence type A. A total of 41 other *B. cereus* and *B. thuringiensis* strains produced heteroduplex profiles identical to that of *B. cereus* ATCC 43881 and were concluded to possess sequence type B. The relative location of heteroduplexes from these strains ranged from 470.3 to 471.9 bpe (average = 471.1, SD = 0.43). The data are shown in Fig. 2 combined with the data from 14 replicates of *B. cereus* ATCC 43881. Based on these results, the heteroduplex assay divided the collection of 69 *B. anthracis*, *B. cereus*, and *B. thuringiensis* strains into two subgroups defined by sequence types A and B.

Sensitivity. As little as 10 to 20 pg of sequence type A per μl was detectable when *B. anthracis* 16S rDNA amplicons were mixed alone with the probe (data not shown). To test the sensitivity of detecting sequence type A in a background of sequence type B, heteroduplex assays were performed with mixtures containing various dilutions of *B. anthracis* 16S rDNA and a constant amount of *B. cereus* ATCC 43881 16S rDNA (Fig. 3). In a background of 5 ng of sequence type B, the detection limit for sequence type A was approximately 100 pg. In other words, sequence type A was detectable in a mixture with sequence type B at a 1:50 ratio. It is important to note that this detection limit describes the ratio of sequences that occur in a mixture after a PCR amplification has been performed. The detection sensitivity for mixtures of genomic DNA prior to PCR has not yet been determined.

Differentiation of sequence types A and B from other sequence types. Replicate heteroduplex profiles were generated separately for *B. mycoides* ATCC 6462, *B. subtilis* ATCC 6051, *B. amyloliquefaciens* ATCC 23350, and *B. flexus* ATCC 49095 to determine the extent to which sequences from these strains could be differentiated from sequence types A and B. The 16S rDNA sequence amplified from *B. mycoides* ATCC 6462 for heteroduplex analysis contained three polymorphisms (ΔG343, A:C399, and G:T408) relative to the probe. In three replicate profiles, the average location of the *B. mycoides* heteroduplex was 471.9 bpe (data not shown), difficult to distinguish from the type B heteroduplex (at 471.07 bpe) without adequate replication. The more distantly related 16S rDNA sequences amplified from *B. subtilis* ATCC 6051, *B. amyloliquefaciens* ATCC 23350, and *B. flexus* ATCC 49095 were easily differentiated from sequence types A and B by heteroduplex analysis. The amplified 16S rDNA from these strains contained 21, 21, and 20 mismatches with the probe, respectively. The *B. subtilis* ATCC 6051 heteroduplex had an average location of 483.9 bpe, while the *B. amyloliquefaciens* heteroduplex had an average location of 484.3 bpe. The *B. flexus* ATCC 49095 heteroduplex was located at position 488.8 bpe. Thus, *B. subtilis* ATCC 6051, *B. amyloliquefaciens* ATCC 23350, and *B. flexus* ATCC 49095 collectively produced three distinct heteroduplexes easily differentiated from the sequence type A and B heteroduplexes.

Analysis of sequence mixtures. Heteroduplex analysis of a constructed mixture of sequence types yielded results similar to analyses of the different sequence types alone. Figure 4A shows a heteroduplex profile of a mixture of 16S rDNA amplicons from *B. anthracis* K2991 (type A sequence), *B. cereus* ATCC 43881 (type B sequence), *B. thuringiensis* ATCC 33679 (type B sequence), *B. subtilis* ATCC 6051, and *B. flexus* ATCC 49095. The four predominant heteroduplex peaks observed in

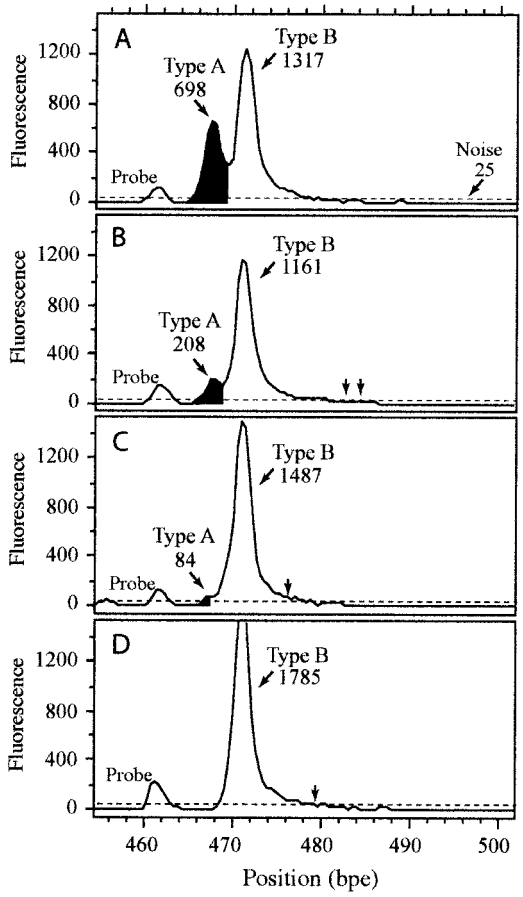

FIG. 3. Detection sensitivity for 16S rDNA amplicons of *B. anthracis* K2991 (sequence type A) diluted in a constant background of 5 ng of 16S rDNA amplicons from *B. cereus* ATCC 43881 (sequence type B). (A) 1:2 ratio; (B) 1:20 dilution; (C) 1:50 dilution; (D) *B. cereus* only. The numbers in each panel indicate peak heights. Arrows show peaks detected above the noise threshold, but only the type A and type B peaks were reproducible.

the profile at positions 468, 471, 483, and 488 bpe corresponded to the locations of peaks observed in the separate heteroduplex profiles of the individual strains (see data above). These results demonstrated the ability of the method to produce a comprehensible profile of a constructed sequence mixture.

To evaluate the complexity of heteroduplex profiles from environmental samples, profiles were obtained for 16S rDNA amplicons from three soil samples from New Mexico, Arizona, and Utah (Fig. 4B and C). The New Mexico, Arizona, and Utah heteroduplex profiles contained 12, 11, and 2 peaks, respectively. None of the profiles contained a heteroduplex corresponding to sequence type A or B from *B. anthracis, B. cereus*, or *B. thuringiensis* strains. Most of the peaks in the profiles did not correspond to heteroduplex data from any of the 10 *Bacillus* species examined thus far. In the New Mexico soil heteroduplex profile, the first and second heteroduplex peaks were located at bpe 485 and 489, closely corresponding to the location of *B. subtilis-B. amyloliquefaciens* heteroduplexes (average, 484 bpe) and *B. megaterium-B. flexus* heteroduplexes (489 bpe). The remaining 10 heteroduplex peaks in the New Mexico profile included peaks with novel locations ranging from 492 to 537 bpe and one distant peak at 605 bpe. The Arizona soil heteroduplex profile contained a set of 11 peaks located between 484 and 503 bpe, 3 of which had locations matching the location of peaks in the New Mexico soil profile. Based on comparison of three or four replicate profiles for each soil, the average standard deviation of the peak locations was 0.3 bpe, although for peaks located above 570 bpe the standard deviation was 3 bpe. Only two heteroduplex peaks (at 491 and 570 bpe) occurred in the profile from the Utah soil sample, and neither peak corresponded to heteroduplexes from known standards.

Comparison of the profiles clearly showed differences in the composition of *Bacillus* species in the three soils. Although some of the heteroduplex peak locations in the New Mexico and Arizona soil profiles were similar, the relative abundance of the peaks differed markedly. For example, *B. subtilis-B. amyloliquefaciens* 16S rDNA sequences appeared to be the most abundant sequence types amplified from *Bacillus* species in the New Mexico soil sample, whereas in the Arizona soil profile, 16S rDNA from an unknown species (or set of species) produced the largest heteroduplex peak (at position 501 bpe). These results demonstrated the capacity of the heteroduplex method to detect the presence or absence of particular targets (sequence types A and B) while simultaneously providing information about more distantly related (possibly undescribed) relatives in environmental samples.

Soil 16S rDNA clone library. A clone library was constructed from 16S rDNA amplicons from the New Mexico soil sample to evaluate the diversity of sequences comprising the observed heteroduplex profile. Single-pass sequence data were obtained from 45 clones. Comparison of the sequences (each approximately 465 bp in length) with the RDP database yielded similarity scores (S_{ab} values) ranging from 0.53 to 1.00 (median = 0.90), suggesting that some of the sequences may have corresponded to new species not currently represented by sequences in the RDP database.

Of the 45 total sequences, two sequences closely matched *Proteobacteria* sequences while five others appeared to be affiliated with the *Acidobacteria* class. These seven sequences each contained between 75 and 86 mismatches with the probe. In experiments with one of the seven clones, a heteroduplex could not be detected in a 3-h run profile (extending to 1,100 bpe) due either to poor hybridization between the highly mismatched probe and target or to extremely slow heteroduplex migration. Thus, these seven sequences were probably not represented by the New Mexico soil heteroduplex profile. This observation suggests a potential advantage of the heteroduplex method: phylogenetically specific heteroduplex profiles can be created even for sequence mixtures containing extraneous sequences.

Excluding the 7 clones described above, the remaining 38 clones most closely matched sequences from the gram-positive division, and 34 of these sequences appeared to be affiliated with the *Bacillus-Lactobacillus-Streptococcus* subdivision (S_{ab} values of 0.786 to 1; median = 0.916). The 34 *Bacillus-Lactobacillus-Streptococcus*-affiliated clones contained between 13 and 40 mismatches with the probe. The relative migration of the 34 clone heteroduplexes is shown in Fig. 4B, overlaid with the New Mexico soil heteroduplex profile. Most of the clones (25 of 34) produced heteroduplexes that corresponded to the

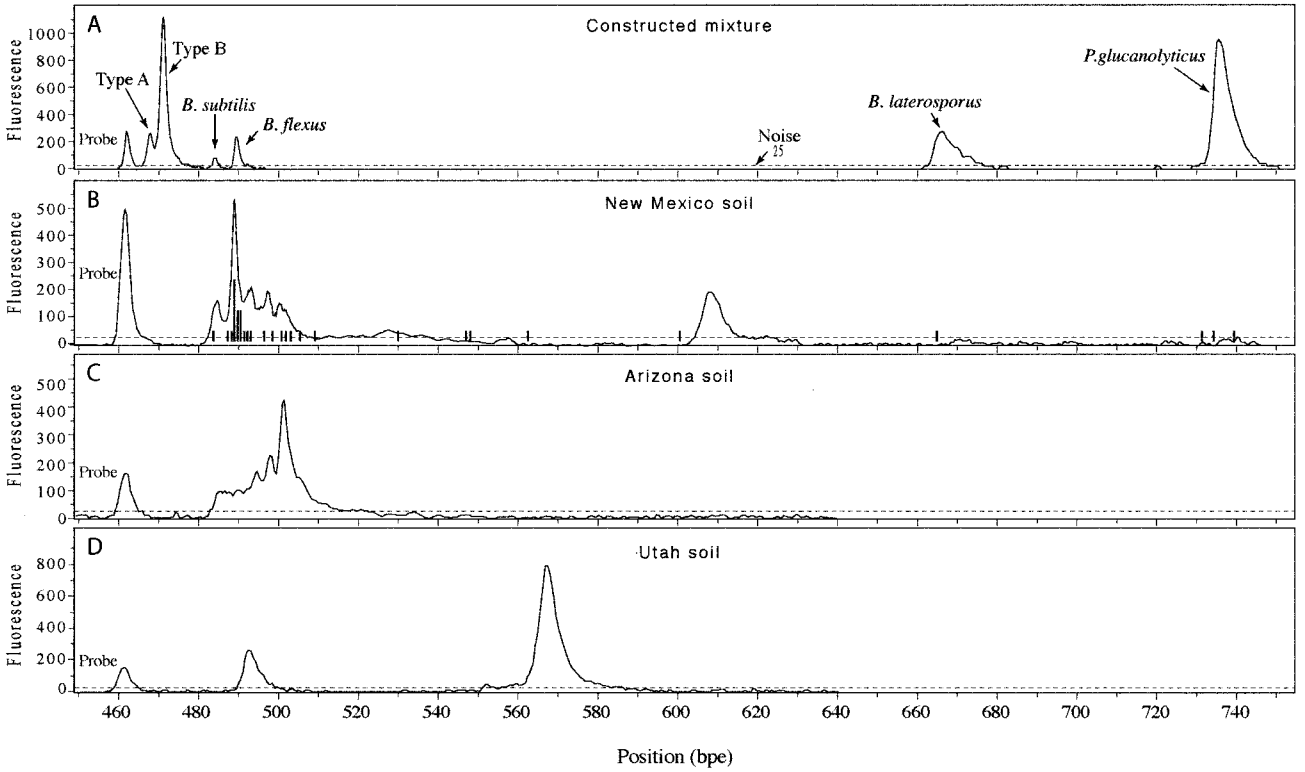

FIG. 4. Heteroduplex profiles of sequence mixtures. (A) Profile of a mixture of 16S rDNA from *B. anthracis* K2991 (type A sequence), *B. cereus* ATCC 43881 (type B sequence), *B. thuringiensis* ATCC 33679 (type B sequence), *B. subtilis* ATCC 6051, and *B. flexus* ATCC 49095. Overlaid with this profile are profiles of *B. laterosporus* ATCC 64 and *P. glucanolyticus* ATCC 49278 for reference. The dashed line indicates the noise threshold. (B) New Mexico soil heteroduplex profile. The overlaid bars indicate the location of heteroduplexes formed by 16S rDNA clones. The height of the smallest bars represents 1 clone. Larger bars represent proportionally more clones. (C and D) Arizona and Utah soil profiles.

location of peaks in the soil heteroduplex profile. However, 9 of the 34 clones yielded heteroduplexes located between 505 and 740 bpe that did not match the location of peaks detected above the noise threshold in the original soil heteroduplex profile.

Figure 5 shows phylogenetic relationships between 31 of the 34 cloned sequences (the missing three sequences were of inadequate length) and closely related *Bacillus* species. Branching order was unresolved for many of the sequences in Fig. 6, as was observed previously (6) in a phylogenetic analysis of nearly full-length (>1,300-bp) 16S rRNA gene sequences from 20 *Bacillus* species (including all the species shown in Fig. 5). Clades with unresolved branching order were therefore shown as paraphyletic lineages. Of the 31 cloned sequences, 13 were closely related to sequences from known *Bacillus* species as follows: "*B. macroides*" (one similar clone), "*B. pseudomegaterium*" (four similar clones), *B. subtilis-B. amyloliquefaciens* (one similar clone) *B. flexus-B. megaterium* (five similar clones), and *B. circulans* (two similar clones). Comparison of heteroduplex mobility data with the phylogenetic tree showed that some sequences belonging to different clades produced indistinguishable heteroduplexes. For example, 12 clones representing four separate clades produced heteroduplexes located between 489 and 491 bpe. These results show that the current heteroduplex assay is incapable of resolving every *Bacillus* clade. In other words, with the breadth of species detected by the present assay, most of the peaks occurring in a profile cannot be confidently attributed to a single species but instead must be considered to indicate groups of possible species.

Prediction of heteroduplex migration. As shown in Fig. 6A and B, heteroduplex mobility is poorly correlated with the number of SNPs or the genetic distance between the probe sequence and various target sequences. To develop a tool for more accurate prediction of heteroduplex migration, we attempted to estimate relative heteroduplex distortion based on the types, locations, and local T_ms of mismatches occurring between a probe and target. Among the set of 47 sequences analyzed, 12 types of mismatches occurred. Two additional categories were created to account for (i) the magnified effect of three or more contiguous mismatches and (ii) clustered mismatches (separated by 1 bp) contributing to destabilization of 5- to 8-bp stretches of DNA. Each category was assigned a weight to account for postulated contributions to duplex distortion. Complementary base pairs were assigned a score of 0. Table 1 lists the categories and weights used (as described in Materials and Methods) to calculate heteroduplex distortion scores.

As shown in Fig. 6, the calculated distortion scores were better correlated ($R^2 = 0.987$) with the observed heteroduplex migration data than with mismatch number or sequence similarity values. The average difference between predicted and observed heteroduplex migration results was −2.85 bpe. The mismatch factors and the weight scales (location and T_m scales) presented here significantly improved the match be-

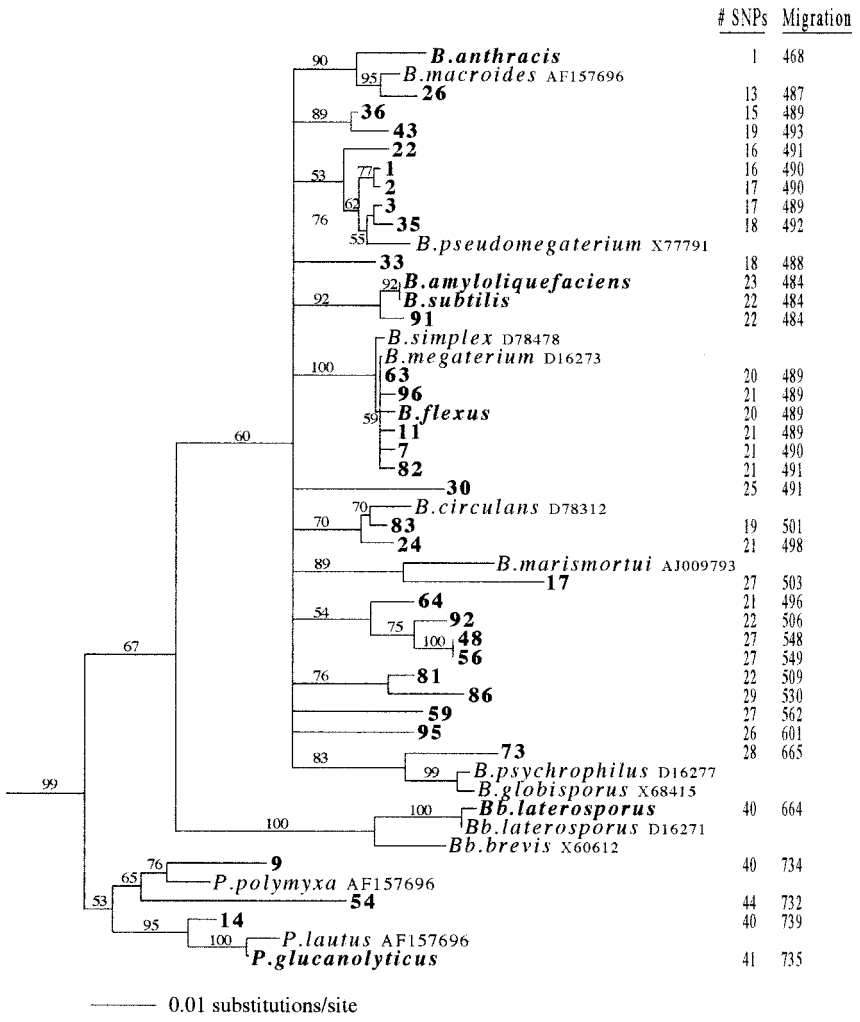

FIG. 5. Phylogenetic reconstruction of 16S rRNA gene sequences amplified from a New Mexico soil sample. The number of SNPs occurring in each sequence relative to the heteroduplex probe is indicated along with the observed average location of the heteroduplex formed by each sequence. Bold numbers are clone designations; bold species names are reference strains examined by heteroduplex analysis.

tween predicted and observed results compared to previous iterations. Previous iterations included use of the following: (i) no location or T_m weight scales, (ii) a location scale with zero weighting for mismatches within 50 bp of an end and full weighting for all other mismatches, (iii) a location scale with a linear decrease in weight from the midpoint to 50 bp from an end, (iv) a T_m scale with a linear decrease in weight from the lowest observed T_m to various $T_{m\max}$ values, and (v) all combinations of the above location and T_m scale options. As shown in Fig. 6C, there are still some large deviations between predicted and observed results. These results demonstrate that although a better fit between predicted and observed results can be obtained by considering mismatch type, mismatch location, and local T_ms, an accurate model predicting heteroduplex distortion will require more complexity than the calculation presented here.

DISCUSSION

Comparative analysis of microbial communities in large sets of environmental samples currently depends on use of scanning techniques (e.g., TRF analysis, DGGE, and heteroduplex analysis) that offer high-throughput analysis. Although scanning techniques provide less resolution than cloning and sequencing, they are capable of assessing the diversity of thousands of individuals in a sample simultaneously and summarizing the diversity in a single profile. To adapt heteroduplex profiling as a scanning technique for complex sequence mixtures, we modified the traditional method by using DNA strand-specific fluorescent labeling, probe engineering, and strand-specific exonuclease digestion of probe DNA. Using the modified format, we obtained a simple, rapid assay for monitoring the occurrence of *B. anthracis* and its close relatives in environmental samples.

The probe described in this study enables confident identification of two subgroups of strains within the *B. cereus-B. thuringiensis-B. anthracis* group. These subgroups correspond to subgroups defined by amplified fragment length polymorphism (AFLP) analysis. AFLP data have shown that the *B. cereus-B. thuringiensis-B. anthracis* assemblage consists of five major strain clusters (44). One cluster, referred to as AFLP

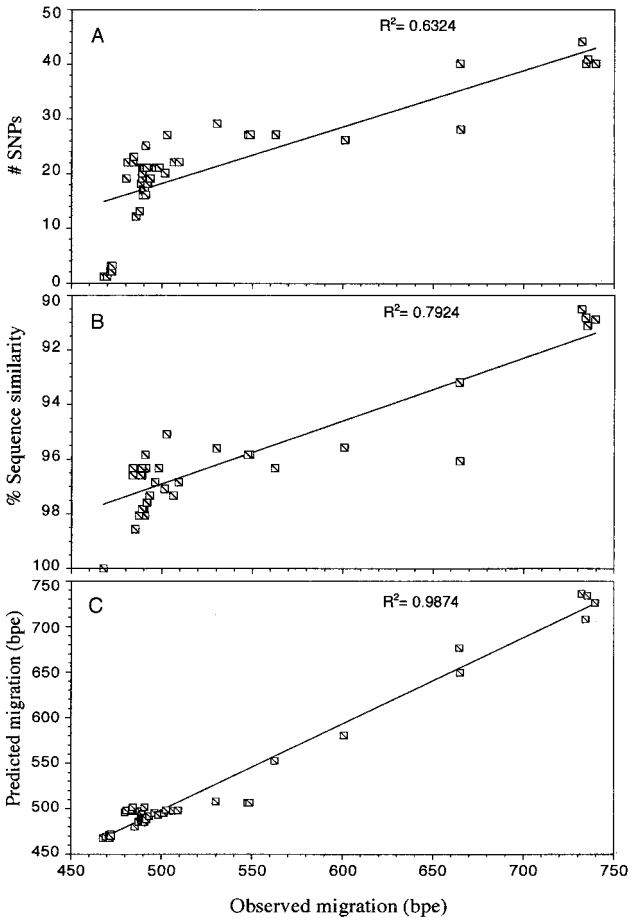

FIG. 6. Comparison of heteroduplex migration with various indices of probe-target sequence identity. Best-fit linear regression lines are shown in each panel.

cluster 1 (35), contains all *B. anthracis* strains and a relatively small number of *B. cereus* and *B. thuringiensis* strains (K. K. Hill and P. J. Jackson, unpublished data). The other four AFLP clusters contain the vast majority of *B. thuringiensis* strains and *B. cereus* strains. With the heteroduplex assay described in this study, AFLP cluster 1 strains generate a heteroduplex that migrates at 468 bp, while all other strains examined thus far produce a heteroduplex at 471 bp. Thus, the present heteroduplex assay generally distinguishes AFLP cluster 1 strains from more distantly related *B. cereus* and *B. thuringiensis* strains.

The heteroduplex method as a tool for microbial community analysis suffers from some of the disadvantages inherent in all scanning techniques. Scanning techniques assess the diversity of a sequence mixture based on indirect detection of sequence variation. For example, TRF analysis is based on differences in the position of restriction sites among sequences, DGGE relies on differences in the melting behavior of duplex DNA molecules, and heteroduplex analysis exploits differences in the shape and length of DNA heteroduplexes. Not all polymorphisms are detected. Thus, these methods typically underestimate the true diversity of a sequence mixture (13, 26, 29, 31). With our heteroduplex assay, only 12 defined heteroduplex peaks were observed in a profile of a sequence mixture amplified from a New Mexico soil sample (Fig. 4B), although by cloning and sequencing, the mixture appeared to contain at least twice as many sequence similarity groups (Fig. 5). The assay described here has relatively low resolution due to the limited amount of variation occurring in the 16S rRNA gene. That is, the assay divides *B. cereus-B. thuringiensis-B. anthracis* strains into only two subgroups, and although the heteroduplex peaks generated by these two subgroups can be interpreted unambiguously, the other peaks cannot be so easily interpreted. This problem can be overcome to a great extent by creating assays that target other genes with greater sequence variability (data not shown).

The most significant limitation of heteroduplex analysis arises from the inability of current separation technology to produce tight bands with highly precise migration behavior. For example, the location of a heteroduplex in an MDE gel often spans a 2-bp range when replicate profiles are compared (Fig. 3). The large margin of error reduces the discriminatory power of heteroduplex profiles. Migrating heteroduplex DNA bands also have broad tailing edges. Due in part to the broad bandwidth, large peaks can easily obscure adjacent small peaks and thereby decrease detection sensitivity. Without improvements in heteroduplex separation technology, the sensitivity of detection of a specific target in the presence of a closely related sequence will continue to be relatively low.

Determining the detection limits of heteroduplex assays in terms of numbers of cells in an environmental sample requires additional work. Such limits depend primarily on the sensitivity of PCR. With the PCR primers described in this study, we measured the 35-cycle PCR detection limit as approximately 0.05 pg of purified *B. anthracis* genomic DNA (i.e., approximately 10 to 20 cell equivalents per reaction yield a visible PCR amplicon on an agarose gel) but have not yet determined the detection limit for *B. anthracis* DNA in a complex environmental sample. Following PCR, the limit for detecting a *B. anthracis* heteroduplex depends both on the absolute quantity of amplified *B. anthracis* 16S rDNA (at least 10 pg of the amplified *B. anthracis* sequence is needed) and on its abundance relative to closely related sequences differing by a single polymorphism. We observed a detection limit of 1:50 for *B. anthracis* 16S rDNA amplicons in the presence of *B. cereus* 16S rDNA amplicons differing by one polymorphism. This limit appears typical for heteroduplex assays (19) and may arise either from the confounding effects of broad bandwidth during electrophoresis or from competitive hybridization dynamics during probe annealing. For comparison with other techniques, the detection limit for adjacent TRFs differing by only 1 bp is about 1:100 or less (also limited by bandwidth), and a detection limit of approximately 1:50 has been reported for DNA microarrays (47). Thus, the sensitivity of heteroduplex detection is comparable to that of other profiling techniques.

While the heteroduplex method has limitations, it also has significant advantages. Since heteroduplex assays do not depend on polymorphisms occurring in specific locations, these assays are very flexible in documenting variation in sequence mixtures. The combination of electrophoresis with DNA probe hybridization also turns a traditionally confounding problem (i.e., probe cross talk) into an exploitable asset. Conventional hybridization-based assays are easily confounded by hybridization between a probe and sequences that are closely related to

a desired target. This problem is typically unavoidable when highly similar sequences are present in a mixture. However, the electrophoresis component of a heteroduplex assay can clearly differentiate the primary target(s) from secondary targets by sorting probe-target hybrids according to size and shape. A single probe can therefore be used to find sequences closely related to a specific target in environmental samples. This feature makes heteroduplex assays such as the one presented here inherently multiplex. Moreover, heteroduplex assays (unlike other profiling methods) are not hampered by nonspecific PCR amplification of extraneous DNA. These factors make heteroduplex profiling easier to perform and interpret than other profiling methods.

Predicting heteroduplex migration from DNA sequence data is an essential component for enhancing the use of heteroduplex profiling of sequence mixtures. A predictive tool would enable in silico evaluation of the resolution of heteroduplex assays, thereby speeding the design of assays appropriate for different applications. Highsmith et al. (20) suggested that G+C content and mismatch type are the overriding factors determining heteroduplex migration. However, the effect of a given type of mismatch can vary depending on flanking sequence context (1–3, 9, 17, 20, 28, 34, 36, 39, 41, 46, 49). In addition, the location (and orientation) of mismatches relative to helical turns and to natural sequence motifs that cause DNA bending can profoundly alter the effect of a particular mismatch type (34). Nonetheless, we attempted to calculate heteroduplex distortion scores using a relatively simple formula (developed in this study) based on the distribution and types of mismatches expected to occur between probe and target sequences. Our approach was based on the assumption that although the effects of a given type of mismatch may vary, the behavior of heteroduplexes containing many mismatches might be generally modeled based on the average effect of individual types of mismatches. As shown in Fig. 6C, for the sequences used in this study, our approach provided an improved prediction of heteroduplex behavior compared to the use of sequence similarity (15). Our approach still provides predictions that are too coarse for effective use in the design and evaluation of heteroduplex assays. However, the approach presented here should provide a useful reference point for development of more sophisticated predictive models of heteroduplex migration.

ACKNOWLEDGMENTS

This research was supported by a grant from the DOE Chemical & Biological Nonproliferation Program.

We thank Paul Jackson and Karen Hill for providing DNA samples and unpublished AFLP data.

REFERENCES

1. **Allawi, H. T., and J. SantaLucia.** 1998. Nearest-neighbor thermodynamics of internal A · C mismatches in DNA: sequence dependence and pH effects. Biochemistry **37:**9435–9444.
2. **Allawi, H. T., and J. SantaLucia.** 1998. Nearest-neighbor thermodynamics of internal G · A mismatches in DNA. Biochemistry **37:**2170–2179.
3. **Allawi, H. T., and J. SantaLucia.** 1997. Thermodynamics and NMR of internal G · T mismatches in DNA. Biochemistry **36:**10581–10594.
4. **Arnold, N., E. Gross, U. Schwarz-Boeger, J. Pfisterer, W. Jonat, and M. Kiechle.** 1999. A highly sensitive, fast, and economical technique for mutation analysis in hereditary breast and ovarian cancers. Hum. Mutat. **14:**333–339.
5. **Barlow, K. L., J. Green, and J. P. Clewley.** 2000. Viral genome characterisation by the heteroduplex mobility and heteroduplex tracking assays. Rev. Med. Virol. **10:**321–335.
6. **Barns, S. M., K. K. Hill, P. J. Jackson, and C. R. Kuske.** 1999. 16S ribosomal RNA sequence-based phylogeny of *Bacillus* species and the *B. cereus* group 99–5628. Los Alamos National Laboratory, Los Alamos, N.Mex.
7. **Bhattacharyya, A., and D. M. J. Lilley.** 1989. The contrasting structures of mismatched DNA sequences containing looped-out bases (bulges) and multiple mismatches (bubbles). Nucleic Acids Res. **17:**6821–6841.
8. **Bowyer, J., N. Verrils, M. R. Gillings, and A. J. Holmes.** 2000. Heteroduplex mobility assay as a tool for predicting phylogenetic affiliation of environmental ribosomal RNA clones. J. Microbiol. Methods **41:**155–160.
9. **Crothers, D. M.** 1998. DNA curvature and deformation in protein-DNA complexes: a step in the right direction. Proc. Natl. Acad. Sci. USA **95:**15163–15165.
10. **Daffonchio, D., C. Cherif, and S. Borin.** 2000. Homoduplex and heteroduplex polymorphisms of the amplified ribosomal 16S-23S internal transcribed spacers describe genetic relationships in the "*Bacillus cereus* group." Appl. Environ. Microbiol. **66:**5460–5468.
11. **Delwart, E. L., E. G. Shpaer, J. Louwagie, F. E. McCutcham, M. Grez, H. Rubsamen-Waigmann, and J. I. Mullins.** 1993. Genetic relationships determined by a DNA heteroduplex mobility assay analysis of HIV-1 *env* genes. Science **262:**1257–1261.
12. **Dunbar, J., S. M. Barns, L. O. Ticknor, and C. R. Kuske.** 2002. Empirical and theoretical bacterial diversity in four Arizona soils. Appl. Environ. Microbiol. **68:**3035–3045.
13. **Dunbar, J., L. O. Ticknor, and C. R. Kuske.** 2001. Phylogenetic specificity and reproducibility and new method for analysis of terminal restriction fragment profiles of 16S rRNA genes from bacterial communities. Appl. Environ. Microbiol. **67:**190–197.
14. **Edwards, S. M., R. H. Z Kote-Jarai, and R. A. Eeles.** 2001. An improved high throughput heteroduplex mutation detection system for screening BRCA2 mutations: fluorescent mutation detection (F-MD). Hum. Mutat. **17:**220–232.
15. **Espejo, R. T., C. G. Feijoo, J. Romero, and M. Vasquez.** 1998. PAGE analysis of the heteroduplexes formed between PCR-amplified 16S rRNA genes: estimation of sequence similarity and rDNA complexity. Microbiology **144:**1611–1617.
16. **Fack, F., S. Deroo, S. Kreis, and C. P. Muller.** 2000. Heteroduplex mobility assay (HMA) pre-screening: an improved strategy for the rapid identification of inserts selected from phage-displayed peptide libraries. Mol. Diversity **5:**7–12.
17. **Ganguly, A., M. J. Rock, and D. J. Prockop.** 1993. Conformation sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes. Proc. Natl. Acad. Sci. USA **90:**10325–10329.
18. **Glavac, D., and M. Dean.** 1995. Application of heteroduplex analysis for mutation detection in disease genes. Hum. Mutat. **6:**281–287.
19. **Hecker, K. H., P. D. Taylor, and D. T. Gjerde.** 1999. Mutation detection by denaturing DNA chromatography using fluorescently labeled polymerase chain reaction products. Anal. Biochem. **272:**156–164.
20. **Highsmith, W. E., Q. Jin, A. J. Nataraj, J. M. O'Connor, V. D. Burland, W. R. Baubonis, F. P. Curis, N. Kusukawa, and M. M. Garner.** 1999. Use of a DNA toolbox for the characterization of mutation scanning methods. I. Construction of the toolbox and evaluation of heteroduplex analysis. Electrophoresis **20:**1186–1194.
21. **Jensen, M. A., and R. J. Hubner.** 1996. Use of homoduplex ribosomal DNA spacer amplification products and heteroduplex cross-hybridization products in the identification of *Salmonella* serovars. Appl. Environ. Microbiol. **62:**2741–2746.
22. **Kuske, C. R., K. L. Banton, D. L. Adorada, P. C. Stark, K. K. Hill, and P. J. Jackson.** 1998. Small-scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. Appl. Environ. Microbiol. **64:**2463–2472.
23. **Le Maréchal, C., M. P. Audrézet, I. Quéré, O. Raguénès, S. Langonné, and C. Férec.** 2001. Complete and rapid scanning of the cystic fibrosis transmembrane conductance regulator (CFTR) gene by denaturing high-performance liquid chromatography (D-HPLC): major implications for genetic counselling. Hum. Genet. **108:**290–298.
24. **Leys, E. J., J. H. Smith, S. R. Lyons, and A. L. Griffen.** 1999. Identification of *Porphyromonas gingivalis* strains by heteroduplex analysis and detection of multiple strains. J. Pathol. **168:**A122.
25. **Lilley, D. M. J.** 1995. Kinking of DNA and RNA by base bulges. Proc. Natl. Acad. Sci. USA **92:**7140–7142.
26. **Liu, W., T. L. Marsh, H. Cheng, and L. J. Forney.** 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl. Environ. Microbiol. **63:**4516–4522.
27. **Maidak, B. L., J. R. Cole, T. G. Lilburn, C. T. Parker, P. R. Saxman, R. J. Farris, G. M. Garrity, G. J. Olsen, T. M. Schmidt, and J. M. Tiedje.** 2001. The RDP-II (Ribosomal Database Project). Nucleic Acids Res. **29:**173–174.
28. **Marathias, V. M., B. Jerkovic, H. Arthanari, and P. H. Bolton.** 2000. Flexibility and curvature of duplex DNA containing mismatched sites as a function of temperature. Biochemistry **39:**153–160.
29. **Marsh, T. L.** 1999. Terminal restriction fragment length polymorphism (T-

RFLP): an emerging method for characterizing diversity among homologous populations of amplification products. Curr. Opin. Microbiol. **2:**323–327.
30. **Meldrum, C. J., R. Crooks, and R. J. Scott.** 2001. D-HPLC detection of APC mutations. Am. J. Hum. Genet. **69:**269.
31. **Moeseneder, M. M., J. M. Arrieta, G. Muyzer, C. Winter, and G. J. Herndl.** 1999. Optimization of terminal-restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. **65:** 3518–3525.
32. **Nataraj, A. J., O. Olivos-Glander, N. Kusukawa, and W. E. J. Highsmith.** 1999. Single-strand conformation polymorphism and heteroduplex analysis for gel-based mutation detection. Electrophoresis **20:**1177–1185.
33. **Olicio, R., C. A. S. Almeida, and H. N. Seuanez.** 1999. A rapid method for detecting and distinguishing clinically important yeasts by heteroduplex mobility assays (HMAs). Mol. Cell. Probes **13:**251–255.
34. **Oussatcheva, E. A., L. S. Shlyakhtenko, R. Glass, R. R. Sinden, Y. L. Lyubchenko, and V. N. Potaman.** 1999. Structure of branched DNA molecules: gel retardation and atomic force microscopy studies. J. Mol. Biol. **292:**75–86.
35. **Pannucci, J., R. T. Okinaka, R. Sabin, and C. R. Kuske.** 2002. *Bacillus anthracis* pX01 plasmid sequence conservation among closely related bacterial species. J. Bacteriol. **184:**134–141.
36. **Peyret, N., P. A. Seneviratne, H. T. Allawi, and J. SantaLucia.** 1999. Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A · A, C · C, G · G, and T · T mismatches. Biochemistry **38:**3468–3477.
37. **Pinar, A., S. Ahkee, R. D. Miller, J. A. Ramirez, and J. T. Summersgill.** 1997. Use of heteroduplex analysis to classify legionellae on the basis of 5S ribosomal-RNA gene sequences. J. Clin. Microbiol. **35:**1609–1611.
38. **Rossetti, S., S. Corra, M. O. Biasi, A. E. Turco, and P. F. Pignatti.** 1995. Comparison of heteroduplex and single-strand conformation analyses, followed by ethidium fluorescence visualization, for the detection of mutations in four human genes. Mol. Cell. Probes **9:**195–200
39. **SantaLucia, J., H. T. Allawi, and P. A. Seneveratne.** 1996. Improved nearest-neighbor parameters for predicting DNA duplex stability. Biochemistry **35**: 3555–3562.
40. **Skopek, T. R., W. E. Glaab, J. J. Monroe, K. L. Kort, and W. Schaefer.** 1999. Analysis of sequence alterations in a defined DNA region: Comparison of temperature modulated heteroduplex analysis and denaturing gradient gel electrophoresis. Mutat. Res. **430:**13–21.
41. **Sugimoto, N., M. Nakano, and S. Nakano.** 2000. Thermodynamics-structure relationship of single mismatches in RNA/DNA duplexes. Biochemistry **39:** 11270–11281.
42. **Swofford, D. L.** 1998. PAUP: phylogenetic analysis using parsimony (and other methods). Sinauer Associates, Sunderland, Mass.
43. **Thomas, G. A., D. L. Williams, and S. A. Soper.** 2001. Capillary electrophoresis-based heteroduplex analysis with a universal heteroduplex generator for detection of point mutations associated with rifampin resistance in tuberculosis. Clin. Chem. **47:**1195–1203.
44. **Ticknor, L. O., A. B. Kolstø, K. K. Hill, P. Keim, M. T. Laker, M. Tonks, and P. J. Jackson.** 2001. Fluorescent amplified fragment length polymorphism analysis of Norwegian *Bacillus cereus* and *Bacillus thuringiensis* soil isolates. Appl. Environ. Microbiol. **67:**4863–4873.
45. **Uribe, P., B. A. Suarezlsla, and R. T. Espejo.** 1999. Ribosomal RNA heterogeneity and identification of toxic dinoflagellate cultures by heteroduplex mobility assay. J. Phycol. **35:**884–888.
46. **White, M. B., M. Carvalho, D. Derse, S. J. O'Brien, and M. Dean.** 1992. Detecting single base substitutions as heteroduplex polymorphisms. Genomics **12:**301–306.
47. **Wu, L., D. K. Thompson, G. Li, R. A. Hurt, J. M. Tiedje, and J. Zhou.** 2001. Development and evaluation of functional gene arrays for detection of selected genes in the environment. Appl. Environ. Microbiol. **67:**5780–5790.
48. **Xiao, W., D. Stern, M. Jain, C. G. Huber, and P. J. Oefner.** 2001. Multiplex capillary denaturing high-performance liquid chromatography with laser-induced fluorescence detection. BioTechniques **30:**1332–1338.
49. **Zuccheri, G., A. Scipioni, V. Cavaliere, G. Gargiulo, P. DeSantis, and B. Samori.** 2001. Mapping the intrinsic curvature and flexibility along the DNA chain. Proc. Natl. Acad. Sci. USA **98:**3074–3079.