J Forensic Sci, September 2006, Vol. 51, No. 5
doi:10.1111/j.1556-4029.2006.00208.x
Available online at: www.blackwell-synergy.com

TECHNICAL NOTE

Lorraine E. Heath,[1] *B.Sc. and Venetia A. Saunders,*[1] *Ph.D.*

Assessing the Potential of Bacterial DNA Profiling for Forensic Soil Comparisons*

ABSTRACT: A pilot study was undertaken to evaluate DNA profiling of the bacterial community in soil as an alternative to geological methods for forensic soil comparisons. Soil samples from three different ecosystems were compared, and the variation within and between ecologically different sites was determined by using terminal restriction fragment (TRF) analysis of 16S ribosomal DNA. Comparison of TRF profiles revealed that samples from within a specific ecosystem (e.g., a field) showed a significantly higher similarity to each other than to those from another ecosystem (e.g., a forest). In addition, some profile features were unique to specific ecosystems. These features may allow the determination of characteristic profiles that will facilitate identification of ecologically different sites, so that a given sample collected from a suspect could be identified as originating from, for example, a field, rather than a forest. The implications of these preliminary findings for forensic investigations are discussed.

KEYWORDS: forensic science, DNA typing, soil, terminal restriction fragment length polymorphism, microbial DNA profiling

The overall purpose of this research project was to evaluate the feasibility of comparing surface soil samples using a microbial (bacterial) DNA profiling method that could be performed on equipment available in all forensic laboratories that profile human DNA. Soil may be considered trace evidence with the potential to link a suspect with a crime scene, and current forensic soil comparisons are based on geological properties (1,2). However, as most forensic laboratories cannot afford an expert geologist, such analyses are rarely performed for routine casework, but are reserved for particularly high-profile, serious crimes (3). As an alternative method of forensic soil comparisons, Horswell et al. (4) suggested DNA profiling of the microbial community in soil. As most forensic laboratories have the facilities for DNA profiling, the use of such a technique to compare soil samples may be more practical for routine casework.

Of the many techniques being used for microbial DNA profiling, terminal restriction fragment length polymorphism (TRFLP) analysis (5) is the one most suitable for forensic applications because the terminal restriction fragments (TRFs) can be detected by equipment that has fluorescence detection capabilities, such as the Applied Biosystems Instruments (ABI) models 373/377/310 DNA Sequencers (Applied Biosystems Inc., Foster City, CA) commonly used in forensic laboratories. For TRFLPs of 16S ribosomal DNA (rDNA), each visible band (fragment) is shown as a peak after analysis with GeneScan software (Applied Biosystems Inc.) and represents a single "ribotype." In theory, the identification of ribotypes that are common and those that are unique to certain soils could be used to discriminate between different soil types. Horswell et al. (4) have demonstrated preliminary success with this technique for "matching" forensic soil samples.

The current project sought to develop a reliable method of microbial (bacterial) DNA profiling to differentiate between surface soil samples collected from locations with different ecological characteristics, namely uncultivated fields, forests, and dunes, and to evaluate the degree of discrimination between samples from different locations and from different areas within the same location. The results reported here are from a pilot study that investigated the inter- and intraecosystem variability of one field, one forest, and one dune system. 16S rDNA was the target and *Msp*I was the restriction enzyme used for TRFLP analysis of bacterial DNA from the soil samples. The implications of the findings for forensic applications are discussed.

Materials and Methods

Ecosystems and Soil Samples

One site of each ecosystem (field, forest, dune system) was chosen in Northwest England. The field and forest were both found on Catton Hall Farm, Frodsham, Cheshire, U.K., while the dune system was in the English Nature Reserve at Ainsdale, Merseyside, U.K. The forest was located on one side and the bottom of a valley, surrounded by fields used for livestock grazing, hay making, or left to grass, with a small stream at the bottom of the valley. The vegetation consisted of mature oak, ash, and holly trees with moderate low undergrowth consisting mainly of bluebells and ferns. The soil appeared fairly sandy, with extensive evidence of rabbit and badger burrowing activity on the side of the slope. The field was adjacent to the forest site and was of mixed grass. It had previously been used for livestock grazing, but had been unused for at least 1 year, and was occasionally mown but the grass was not harvested. A small road ran along the bottom edge of the field, which initially sloped shallowly away from the

[1]School of Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, U.K.

*This work was presented orally at the American Academy of Forensic Sciences 56th Annual Scientific Meeting, February 2004, Dallas, TX.

Received 12 March 2005; and in revised form 5 July 2005 and 4 March 2006; accepted 26 March 2006; published 7 Aug. 2006.

road before rising steeply to a flat plateau at the top. There were some valleys in the field, and some of the lower sections were quite damp with bog-like conditions and slightly different vegetation. The dune system was part of the "slacks" area composed of stable, mostly vegetated dunes behind the ocean-fronted, mostly unvegetated, mobile dunes. The vegetation consisted primarily of grasses, lichen, and mosses, with some aromatic shrubs in low-lying areas. The tops of the dunes were generally sandy and dry with little vegetation, whereas the areas between the dunes were damper with more vegetation and richer soil.

Within each site, five representative areas were chosen and a 15 cm × 15 cm × 5 cm soil sample was collected. The forest samples (F1–F5) were collected on March 22, 2002; the field samples (G1–G5) were collected on August 7, 2002; and the dune samples (D1–D5) were collected on September 15, 2003. Once collected, all samples were treated in the same way. A small portion of each was used immediately for soil characterization, while the rest of the sample was air-dried, homogenized using a blender, sieved to remove large (>2 mm) debris, and stored in Ziploc plastic bags at −20°C.

It should be noted that for the most part, the ecosystem samples referred to herein are the five samples (as triplicate subsamples) from each of these three ecosystems. However, to check whether the time differences between the collection of the soil samples from the different ecosystems could have contributed to any observed differences between the DNA profiles, further samples from the field (Field L) and forest (Forest L) sites were collected on August 8, 2003 (a collection time similar to that of the dune samples). Field L and Forest L samples (as triplicate subsamples) were treated as for all the other samples and TRFLP profiles generated. These profiles were then compared by cluster analysis (see "Statistical Analysis") with those of the other ecosystem samples to determine whether they clustered with the samples deriving from the same collection site, even though collected at a different time.

Soil Characterization

Moisture content, pH, and organic matter content of the soil samples were determined by standard methods. The oven dry method was used for moisture content, weighing samples before and after drying. The percent of weight loss was used as the percent moisture content of the sample.

The pH of each sample was determined by using a pH meter method. Twenty-five milliliters of de-ionized water were added, the sample was thoroughly stirred, and then allowed to settle for 30 min. A pH meter was then used to determine the pH of the supernatant water without disturbing the settled soil.

The amount of carbon or organic matter was determined using the weight loss on ignition test (6).

DNA Extraction

Total soil DNA was extracted from three replicate subsamples of each soil sample collected from the different ecosystems using a modification of the method reported by Griffiths et al. (7). Samples of dry soil (0.1 g) were added to 2 mL tubes approximately half full of 0.1 mm zirconia/silica beads (Stratech, Soham, U.K.). 0.5 mL of phenol:chloroform:isoamylalcohol (25:24:1) and 0.5 mL of cetyltrimethylammonium bromide (CTAB) extraction buffer (equal volumes of 10% (w/v) CTAB in 0.7 M NaCl with 240 mM K_2HPO_4 [pH 8.0]) were added. Tubes were shaken in a Biospec Mini-BeadBeater (Stratech) at 2500 r.p.m. for 120 sec, and then centrifuged in an MSE Microcentaur centrifuge (Sanyo Gallenkamp, Leicestershire, U.K.) at 13,400 × *g* for 5 min. The aqueous layer was removed and residual phenol was extracted by mixing with an equal volume of chloroform:isoamyl alcohol (24:1), and centrifuging at 13,400 × *g* for 5 min. The aqueous layer was further purified by incubation with 50 μL of lysozyme solution (100 mg/mL sterile de-ionized water) for 30 min at 37°C. DNA was precipitated by incubation with two volumes of 30% (w/v) polyethylene glycol (PEG) 6000 in 1.6 M NaCl for 2 h at room temperature. After centrifugation at 13,400 × *g* for 10 min, the supernatant was discarded and the pellet was washed with 200 μL of 70% (v/v) ethanol, air-dried, and then re-suspended in 50 μL sterile de-ionized water. Extracted DNA was quantified by spectrophotometric analysis at 260 nm (8). In addition, DNA concentration was estimated by spotting on to 0.8% (w/v) agarose plates with ethidium bromide (1 μg/mL) and UV visualization comparing with DNA standards of known concentration.

DNA Amplification and Bacterial Community TRFLP Analysis

Terminal restriction fragment lengths were determined using a modification of the methods of Kuske et al. (9) and LaMontagne et al. (10). The bacterial 16S small subunit rRNA gene was amplified using universal eubacterial primers for 16S rRNA genes (from Invitrogen Life Technologies Ltd., Paisley, Scotland, U.K.). The forward primer 8-27f 5′-AGAGTTTGATCCTGGCTCAG-3′ (*Escherichia coli* positions 8–27) was fluorescently labelled with 5-carboxyfluorescein (FAM) during synthesis. The reverse primer 1507–1492r 5′-TACCTTGTTACGACTT-3′ (*E. coli* positions 1507–1492) was unlabelled. Extracted DNA samples were diluted 1 in 100 (forest and field samples) or 1 in 10 (dune samples) to facilitate pipetting of small quantities for analyses. All samples were further diluted 1 in 10 before addition to the PCR reactions to dilute PCR inhibitors and to ensure consistent PCR success. PCR reactions were performed in 50 μL volumes containing 25 μL Reddy PCR Master Mix (Abgene, Epsom, U.K.) (75 mM Tris-HCl, 20 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.01% (v/v) Tween 20, 0.2 mM each dNTP, 1.25 U Thermoprime Plus DNA Polymerase, and gel loading precipitant and red dye for electrophoresis), 50 pmol of each primer, and 5–15 ng of extracted DNA. To ensure that the total relative fluorescence (RF) in all samples was similar to facilitate later analyses and comparisons (see (11)), the amounts of DNA used in the PCR reactions were 5 ng for dune samples, 10 ng for field samples, and 15 ng for forest samples (as determined in a preliminary study). A positive control of 2.0 ng *E. coli* genomic DNA and a negative control containing no template DNA were amplified with each PCR batch. The thermocycling conditions were as follows: a hot start at 94°C for 3 min (1 cycle); 35 cycles of 94°C for 1 min, 50°C for 45 sec, 72°C for 1.5 min; and a final extension at 72°C for 5 min. To verify that the PCR products were of the correct size, amplified samples along with the DirectLoad Wide Range DNA Marker (50–10,000 bp) size standard (Sigma-Aldrich, Poole, Dorset U.K.) were resolved by electrophoresis on 1% (w/v) agarose gels at 80 V for 1–2 h and staining with ethidium bromide. Restriction enzyme digestion was performed in a final volume of 20 μL by adding 10 μL of PCR product to 6 μL sterile water, 20 U of restriction endonuclease *Msp*I (Cambio, Cambridge, U.K.), and 2 μL of corresponding enzyme buffer, according to the manufacturer's instructions. The reactions were incubated at 37°C for a minimum of 4 h and the restriction was terminated by heating at 65°C for 20 min. The length of the fluorescently labeled fragments was determined by electrophoresis in a denaturing 6% (w/v) polyacrylamide gel with an

ABI 373A automated sequencer (Applied Biosystems Inc., Warrington, U.K.) as described by Liu et al. (5), with the modifications that 2.5 μL of digested DNA was used and the loading buffer was omitted as the PCR reaction already contained loading dye. GeneScan-2500 TAMRA (Applied Biosystems Inc.) was used for the internal size standard. TRFLP profiles were obtained for each of the three replicate subsamples from each of the 17 soil samples (five from each ecosystem and two additional samples [Field L and Forest L] from the field and forest obtained at a later date) collected. The amplified positive and negative controls were also subjected to *Msp*I digestion and denaturing polyacrylamide electrophoresis to ensure the validity of the profiles. The data were analyzed using GeneScan software version 2.1. The analysis parameters included a threshold for peak detection of 25 relative fluorescent units (RFU), a minimum peak half-width of three points, use of the Local Southern size calling method with light smoothing, and GeneScan 2500 split peak correction.

TRFs were considered identical if they differed by less than 0.5 bp. The first step of the profile analysis was to standardize the DNA quantity between replicate subsample profiles using the method devised by Dunbar et al. (11). The total relative fluorescence (RF) was calculated for each of the three subsamples (all peak heights greater than 25 RFU were added together). Each peak height in the two profiles with the larger total RF was standardized to the smallest by using a correction factor based on the proportion of the smallest total RF and a larger quantity. All the peak heights for the larger profile were then multiplied by this factor to give an adjusted height. The new total RF of all peaks in that subsample still greater than 25 RFU was calculated. This procedure was repeated until all three replicate subsample profiles had the same total RF. Subsample profiles were then compared and a composite profile for each soil sample was produced containing only those TRFs that were found in all three subsample profiles. The 17 composite profiles (five from each ecosystem, with an additional sample from the later collection date from the forest and field ecosystems) were then used to produce a single list of all TRF sizes found in one or more composite profiles. Each composite profile was then compared with this list and a binary vector was constructed for each sample representing the presence or absence of the TRFs in the list.

Statistical Analysis

Most statistical analyses were carried out using the computer package SPSS v. 11.0 (SPSS Inc., Chicago, IL), but Student's *t*-test (unpaired) or Mann–Whitney rank sum test were performed using SigmaStat (Jandel Scientific Software, San Rafael, CA). The Jaccard coefficient was used as a measure of similarity of binary vectors, and a matrix of pair-wise comparisons was constructed as per Dunbar et al. (11). Agglomerative hierarchical clustering was performed on this matrix and the results were displayed as a dendrogram. The percent dissimilarity between samples was rescaled (by SPSS v. 11.0) to numbers between 0 and 25, preserving the ratio of the distances between samples. The average similarity, expressed as a percentage, between and within ecosystems was examined by comparing the similarities between the five samples from within an ecosystem with the similarities between those five samples and the 10 samples from the other two ecosystems. Student's *t*-test was used for the comparison between the forest samples and the others, while the nonparametric Mann–Whitney rank sum test was used for the other two comparisons due to the unequal variance and nonnormal distribution observed in these two data sets.

TABLE 1—*Soil characteristics for each ecosystem.*

Ecosystem	Mean* (Range)		
	Moisture Content (%)†	pH	Organic Matter (%)‡
Forest	52.6 (22.8–67.0)	2.5 (2.3–2.6)	37.3 (7.0–61.9)
Field	33.4 (27.7–41.0)	3.9 (3.4–4.3)	5.3 (3.9–6.0)
Dune	17.5 (0.1–60.6)	5.8 (4.7–6.4)	7.5 (0.6–22.8)

*Mean from five sites within each ecosystem.
†Percent wet soil weight.
‡Percent dry soil weight.

Results

Characteristics of the soil samples with respect to moisture content, pH, and organic matter content for each ecosystem are presented in Table 1. The widest range for moisture content was in the samples from the dune site, reflecting the difference between the top of the dunes and the low-lying, damper regions between the dunes, while the narrowest range was seen in the samples from the field reflecting the more uniform nature of this site. The mean moisture content was the highest in the forest samples, with the wide range largely due to the lower moisture content seen in the F2 site, which was composed largely of sand. The pH range was fairly narrow for all samples, the forest soils being the most acidic. The widest range and highest mean organic matter content was found in the forest soils, with the lowest mean and narrowest range being exhibited by the field soils. The dune soils showed a relatively wide range, but a low mean organic matter content.

Total community DNA was extracted from three subsamples from each of the soil samples from the three ecosystems and the 16S rDNA amplified using universal eubacterial primers. Amplicons, as resolved by agarose gel electrophoresis, were approximately 1500 bp (data not shown), the expected size for PCR products obtained using 16SrDNA primers 8-27f and 1507-1492r. TRFLP profiles were generated after digestion of products with *Msp*I. A profile for a soil subsample was composed of all the peaks (TRFs) identified by Genescan and included their size in base pairs (bp) and their height in RFUs. Representative electropherograms of the soil bacterial community TRFLP profiles for each ecosystem can be seen in Fig. 1. The baseline threshold/background noise level was low; therefore, peaks over 25 RFUs were considered to be true DNA fragments. Owing to poor resolution of the larger fragments, only TRFs from 70 to 470 bp were included in the comparisons. The TRF size ranges, number of TRFs, and the percent of reproducible peaks between the three subsamples by ecosystem can be seen in Table 2. The largest number of TRFs was observed in the field samples (both in the replicate and composite profiles), which had nearly twice as many TRFs as in the profiles of the forest samples, and these had the lowest number of TRFs (in both the replicate subsample and composite profiles). The reproducibility of peaks between the subsample profiles and the composite profiles was similar for all three ecosystems, ranging from 60% to 76%.

The results of similarity calculations for TRFLP profiles between and within each ecosystem are shown in Table 3. For all three ecosystems, profiles from within the ecosystem were significantly more similar to each other than to profiles from the other two ecosystems (the probability that the similarity of intra-ecosystem profiles is the same as that of interecosystem profiles being less than 1%). The hierarchical cluster analysis (Fig. 2) confirms this, with the five samples from each ecosystem clustering together, and separate from the other 10 samples. The distance

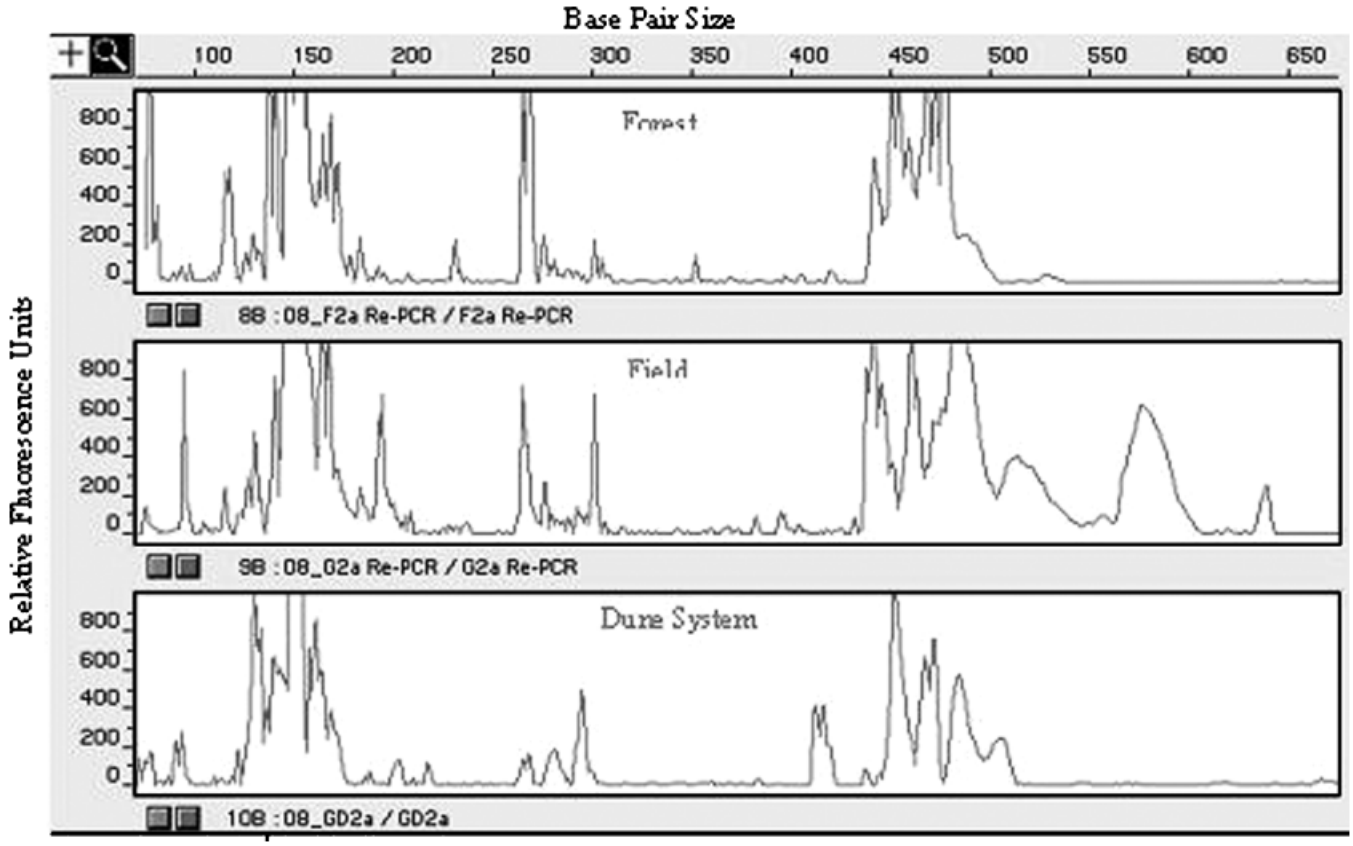

FIG. 1—*Representative electropherograms of terminal restriction fragment length polymorphism profiles from each of the ecosystems—forest, field, and dune. Electropherograms have been aligned by base pairs, with fragment size in base pairs on the X-axis and peak heights in relative fluorescence units on the Y-axis. Peak determination and counting were performed using Genescan software. Base pair sizes determined using GeneScan-2500 TAMRA (Applied Biosystems Inc.) internal size standard (size standard peaks not shown).*

values displayed along the bottom horizontal axis represent a relative dissimilarity between composite profiles (determined from three replicate profiles) from each sample. Samples whose dendrogram arm split at low distance values (e.g., F3 and F4) are more similar to each other than those that split at higher values (e.g., D1 and D3). In addition, Field L and Forest L samples, despite being collected from the field and forest sites at the later date, clustered with the other samples from their respective ecosystems.

Overall, 255 TRFs were identified in one or more composite profiles. Of these, seven were ubiquitous in all 15 composite profiles. Table 4 compares the number of TRFs found in one or more profiles within each ecosystem. In addition to the seven unique TRFs (five unique to the field, two to the dune, and none to the forest soil profiles), three TRFs were only found in all five samples of two of the three ecosystems (one was unique to the forest and the dune samples, while two were unique to the field and dune samples).

TABLE 2—*Size and number of terminal restriction fragments in profiles from each ecosystem.*

Ecosystem	Size Range (bp)	Mean* Number (Range) TRFs in Subsample Profiles	Mean† Number (Range) TRFs in Composite Profiles	Percent Reproducible TRFs (%)
Field	75.09–468.34	143 (122–165)	108 (99–113)	76
Forest	72.64–469.14	84 (69–103)	50 (40–62)	60
Dune	71.52–469.14	114 (99–132)	76 (67–88)	67

*Mean of the replicate profiles from the three subsamples from each of the five sites in each ecosystem.

†Mean of composite profiles generated from the three subsample profiles from each of the five sites in each ecosystem.

TRF, terminal restriction fragment.

Discussion

The purpose of this study was to examine the ability of TRFLP analysis to discriminate between soil samples taken from areas with different ecological properties. The goal was to determine whether samples taken from within the same ecosystem (e.g., a

TABLE 3—*Inter- and intraecosystem similarity of terminal restriction fragment length polymorphisms.*

Ecosystem	Mean* (Range) Intraecosystem Similarity (%)†	Mean‡ (Range) Interecosystem Similarity (%)†	Significance (p Value)§
Field	54.9 (45.2–62.2)	25.9 (16.8–34.3)	<0.0001 (*t*-test)
Forest	55.4 (43.4–67.8)	23.8 (16.8–30.9)	<0.0001 (Mann–Whitney)
Dune	36.0 (29.4–54.7)	26.3 (18.3–34.3)	<0.0001 (Mann–Whitney)

*Mean of similarity between the five composite profiles from the five sites within each ecosystem.

†Percent similarity between binary vectors determined by using the Jaccard coefficient.

‡Mean of similarity between the five composite profiles within the ecosystem and the five composite profiles from each of the other two ecosystems.

§*t*-test/Mann–Whitney rank sum probability values for the difference between the inter- and intraecosystem similarity percentages for each ecosystem. The criterion for statistical significance was set at $p = 0.05$.

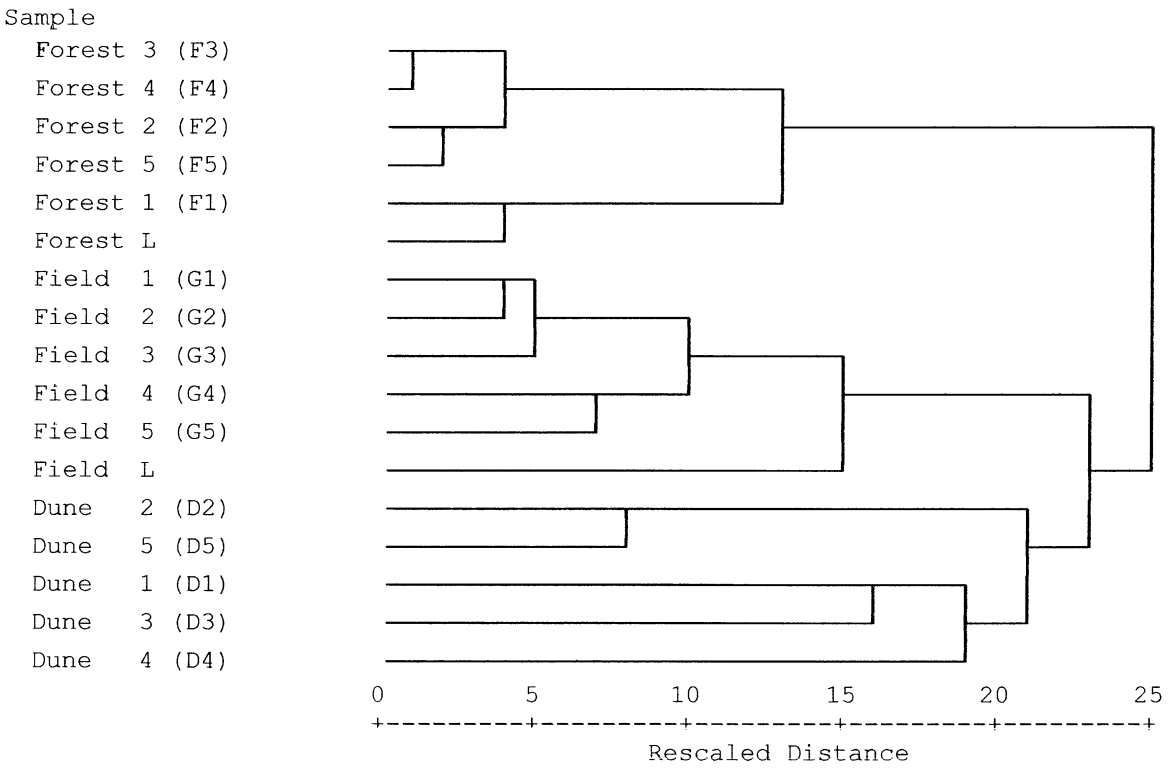

FIG. 2—*Dendrogram of the agglomerative hierarchical cluster analysis of the ecosystem discrimination samples showing the separate clustering of the samples from each ecosystem. The forest samples (F1–F5); field samples (G1–G5); dune samples (D1–D5); and one sample from each of the field (Field L) and forest (Forest L) sites collected at a different time are shown. The distance bar represents relative dissimilarity between composite profiles (determined from three replicate profiles) from each sample rescaled by SPSS v. 11.0 (0 = 32.2%, 25 = 83.2%). Note Field L clusters with the field samples and Forest L with the forest samples.*

forest) were more or less similar to each other than to soil samples taken from other ecosystems (e.g., a field or sand dunes). In addition, this study provided preliminary data regarding the possibility of identifying ecosystem-specific TRFs, or ribotypes, that could potentially allow a soil sample to be identified as to which type of ecosystem it had originated from.

A prerequisite for this study was an extraction protocol that generated DNA of suitable quality for PCR and restriction analysis from all the soil types used. Accordingly, a preliminary evaluation of a number of DNA extraction kits and modifications of the bead beating method of Griffiths et al. (7) was undertaken for extraction of total community DNA from the soils (data not shown). While a kit would be more suitable for use in a forensic laboratory, a modification of the bead beating method proved more reliable in terms of DNA yield and purity, reproducibility of TRFLP profiles, and extracting DNA from the largest variety of bacteria in the different soils tested and so was used here. Furthermore, in line with other workers (12) it was found that the TRFLP profiles generated depended on the DNA extraction procedure used. This underlines the need to handle all samples in the same way, using the same extraction protocol, for forensic soil comparisons and consider data obtained in conjunction with the specific protocol used. Clearly, before any recommendation could be made about an appropriate method of extracting DNA for routine use in forensic laboratories, it would be necessary to evaluate further the changes in TRFLP profiles as a function of extraction procedure.

TABLE 4—*Comparison of the diversity of TRFs in the three ecosystems.*

Ecosystem	Number TRFs in one or more composite profiles (% of total*)	Number TRFs in all five composite profiles (% of total†)	Number TRFs unique to ecosystem‡
Field	185 (73%)	53 (29%)	5
Forest	91 (36%)	27 (30%)	0
Dune System	161 (63%)	15 (9%)	2

*Percent of total TRFs (255) found in one or more composite profiles from all ecosystems.

†Percent of total TRFs (185, 91, and 161 for field, forest, and dune, respectively) found in one or more composite profiles for that ecosystem.

‡Number of TRFs found only in all five composite profiles from that ecosystem.

TRF, terminal restriction fragment.

Using the bead beating method for DNA extraction and restriction of the amplified 16S rDNA with *Msp*I to generate the TRFLPs, the average number of TRFs in the original profiles, after standardization, ranged from 84 for the forest profiles to 143 for the field profiles, with the dune samples at 114. This is comparable with data of other workers such as Kuske et al. (9), who obtained an average of 132 TRFs from field samples using a sodium dodecyl sulfate-based extraction protocol, with the same primers and restriction enzyme used in this study. However, Clement et al. (12) obtained 56 TRFs from a sand sample using a nonbead-beating, solvent-based extraction protocol, but using the same or similar primers and the restriction enzyme *Msp*I as used in this study. The lower apparent diversity found by Clement et al. (12) may be due to the different extraction protocol, or to the sand sample being collected from a mobile dune or beach area rather than the more stable dune system used here. The mean percent of reproducibility in peak number between replicate samples in the present study varied from 60% to 76% in the different ecosystems. This is similar to the reproducibility found by Dunbar

et al. (11), who determined the average reproducibility in their studies to be 74%. Dunbar et al. (11) and Osborn et al. (13) also found that a large amount of the irreproducibility could be eliminated by raising the threshold for interpretation (the RF value above which a peak is identified as truly resulting from a TRF, rather than baseline noise), but this could also eliminate some reproducible peaks and potentially alter profile interpretations and comparisons.

For all three ecosystems, the five samples from within the ecosystem were significantly more similar to each other than to the other ecosystem samples. Both the field and forest samples had similar average intra- and interecosystem similarities: 54.9% and 55.4% intraecosystem, and 25.9% and 23.8% interecosystem, respectively. The dune samples had a much wider range and a lower average of intraecosystem similarity. This is most likely due to the higher heterogeneity (subjectively determined by the appearance of the area including soil composition, moisture level, and vegetation) of the dune system. The field and forest habitats appeared relatively homogeneous (i.e., there appeared to be relatively little variation in the moisture level, soil composition, and vegetation throughout the area), with the exception of the soil composition of forest sample 2, while the dune system had very different vegetation, moisture levels, and soil composition throughout. Nevertheless, the dune samples were significantly more similar to each other than to the samples from the other two ecosystems. The results of the agglomerative hierarchical cluster analysis also indicated that the samples from each ecosystem group together. Interestingly, other reports have suggested that soil microbial DNA profiles are primarily dependent on the physical and chemical properties of the soil, rather than the associated vegetation (14). Furthermore, Franklin and Mills (15) suggested that microbial DNA profiles can be very heterogeneous, even within a habitat that appears relatively homogenous, possibly due to variations in soil properties. However, these reports (14,15) examined frequently disturbed agricultural fields with artificially controlled vegetation. Kang and Mills (16), in contrast, found a clear relationship between the soil bacterial community structure and the associated vegetation in their 2-year study of ecosystem succession in a grass field.

Of the three ecosystems studied here, the field had the highest diversity with a total of 185 TRFs found in one or more composite profiles, while the lowest diversity was found in the forest ecosystem, with only 91 TRFs in one or more composite profiles. While it is possible that this is due to lower numbers and diversity of microbes in the forest soil, the PCR-based methodology used may also have affected these profiles. Humic acid, a PCR inhibitor (17), is known to be present in higher quantities in soils with higher organic matter (18) like the forest soils in this investigation (with an average of 37.4% organic matter compared with 5.3% and 7.5% in the field and dune samples, respectively, see Table 1). Furthermore, humic acid has been shown to introduce a bias toward lower diversity estimates (10). The possibility of identifying unique ribotypes at vegetatively different sites was suggested by both Kuske et al. (9) and Brodie et al. (19), who found TRFs, or ribotypes, unique to each of the sites with different vegetation that they studied. The current study identified five TRFs unique to the five field composite profiles and two unique to the five dune composite profiles. No unique TRFs were found for the forest ecosystem. These results are in keeping with the larger number of TRFs in the field and dune samples and the lower "apparent" diversity in the forest profiles. Nonetheless, these results suggest that the presence or absence of these unique TRFs may be used to identify an unknown soil sample as originating in either a field or a dune system. In addition, one specific TRF was found in all forest and dune composite profiles, but not in any of the field composite profiles. Possibly, the presence of this TRF and the absence of the two unique dune TRFs could be used to indicate a forest as the origin of a sample. The utility and reliability of these "identifying" TRFs needs to be investigated further by examining other field, forest, and dune sites all over the country. It may, in turn, be possible for forensic laboratories to construct databases of the inter- and intraecosystem variability and of ecosystem-specific TRFs for their local area, to be used in forensic soil comparisons.

To check for possible temporal/seasonal effects on TRFLPs, further field and forest samples (L) were collected at a later time, but from the same location as the orginal samples. In both cases, the sample collected later clustered with its respective originating ecosystem: Field L with field samples, Forest L with forest samples (see Fig. 2). This might suggest that the nature of the ecosystem from which a soil sample was collected had a greater impact on its bacterial DNA profile than the time of collection. Although Chabrerie et al. (20) did not investigate seasonal effects in their study of grassland successional transects, they suggested that the apparent stability of the soil microbial community structure may be due to bacterial mechanisms of resistance and dormancy. As environmental conditions change throughout the seasons, different species may alternate between being active and dormant. However, they will all still be present in the soil and thus will be detected by DNA profiling methods (such as those used in the current study) that examine the entire soil bacterial community rather than only the active members. Griffiths et al. (21) examined the same soils in the four different seasons, and found that only the samples collected in the summer were significantly different from those collected in all the other seasons (which were not significantly different from each other). Horswell et al. (4) found that although samples collected from the same site 8 months apart were less similar to each other than those collected in a very short time frame, they still showed a high degree of similarity (70% compared with 90%). Although seasonal and temporal variations in soil microbial DNA profiles need to be investigated further, these previous studies (4,20,21) and the preliminary results from the current study would support the potential of the ecosystem-discrimination-based approach to compare forensic soil samples. Stable soil microbial DNA profiles may allow comparisons to be made between soil on, for example, a suspect's shoe and a possible crime scene despite significant time lapses between the collection of samples. In the current study, TRFLP profiles have been generated using 0.1 g soil samples. Moreover, Horswell et al. (4) obtained profiles from soil samples recovered from the sole of a shoe and from soil-stained clothing. The use of such small samples confirms the applicability of this assay to analysis of trace evidence including for example, soil from the tread of a vehicle's tyres or from under the fingernails of a suspect.

The results of this pilot study lend support to the notion that forensic soil samples originating in specific ecological sites may be distinguished by their bacterial DNA profiles. A more extensive study is now required using all the different ecosystems to establish whether specific TRFs can indeed be used as potential indicators of a particular ecosystem. In this way, it may be possible to determine the origin of a forensic soil sample. Additionally, it may be possible to use the relative abundance of individual TRFs to distinguish between different ecosystems and different sites within a given ecosystem, as suggested by Brodie et al. (19). This would increase the strength of this type of soil evidence. However, the stability of bacterial DNA profiles needs to be evaluated to determine the maximum time that can elapse between

sample deposition on the suspect and reliable comparison with collected samples and still allow the detection of TRFs characteristic of a particular ecosystem.

Clearly, this study provides a baseline for assessing the validity of using TRFLP profiles of the bacterial community of soil for forensic investigations. However, much needs to be done before any such method might be accepted by the courts. In addition to the recommendations for further work already discussed, an investigation of the impact of DNA extraction procedure on the nature and reproducibility of TRFLP profiles is needed. Magnetic beads may be useful here to improve the purity of extracted DNA, in turn removing PCR inhibitors that might otherwise bias diversity estimates and hence influence the profiles generated. Moreover, a study on small-scale variability of TRFLP profiles is required in order to assess how localized a profile is. We are currently undertaking such an investigation using soil transects from each of the three ecosystems studied. When the above-mentioned work has been completed, the validity and reliability of the TRFLP approach would need to be tested through blind proficiency tests before its use in forensic casework.

Overall, the results of this preliminary investigation suggest that bacterial DNA profiling using TRFLP analysis may, with further refinement, provide a practical method for comparing forensic soil samples, which could be easily applied by all forensic laboratories currently performing human DNA profiling without the need for geological specialists or additional equipment.

Acknowledgment

This work was funded by Liverpool John Moores University, Liverpool, U.K.

References

1. Junger EP. Assessing the unique characteristics of close-proximity soil sample; just how useful is soil evidence? J Forensic Sci 1996;41(1):27–34.
2. Sugita R, Marumo Y. Validity of colour examination for forensic soil identification. Forensic Sci Int 1996;83(3):201–10.
3. Fraysier HD, van Hoven H. A simpler mineralogical soil analysis method using dispersion staining. Microscope 1993;40(3):107–10.
4. Horswell J, Cordinger SJ, Maas EW, Martin TM, Sutherland BW, Speir TW, et al. Forensic comparisons of soils by bacterial community DNA profiling. J Forensic Sci 2002;47(2):350–3.
5. Liu WT, Marsh TL, Cheng H, Forney LJ. Characterisation of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol 1997;63(11): 4516–22.
6. Dean WE. Determination of carbonate and organic matter in calcareous sediments and sedimentary rocks by loss on ignition; comparison with other methods. J Sediment Petro 1974;44:242–8.
7. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. Appl Environ Microbiol 2000;66(12):5488–91.
8. Sambrook J, Russell D. Molecular cloning. A laboratory manual. 3rd ed. New York: Cold Spring Harbor, 2000.
9. Kuske CR, Ticknor LO, Miller ME, Dunbar JM, Davis JA, Barns SM, et al. Comparison of soil bacterial communities in rhizospheres of three plant species and the interspaces in an arid grassland. Appl Environ Microbiol 2002;68(4):1854–63.
10. LaMontagne MG, Michel FC, Holden PA, Reddy CA. Evaluation of extraction and purification methods for obtaining PCR-amplifiable DNA from compost for microbial community analysis. J Microbiol Methods 2002;49:225–264.
11. Dunbar J, Ticknor LO, Kuske CR. Phylogenetic specificity and reproducibility and new method for analysis of terminal restriction fragment profiles of 16S rRNA gene from bacterial communities. Appl Environ Microbiol 2001;67(1):190–7.
12. Clement BG, Kehl LE, DeBord KL, Kitts CL. Terminal restriction fragment patterns (TRFPs), a rapid, PCR-based method for the comparison of complex bacterial communities. J Microbiol Methods 1998;31:135–42.
13. Osborn AM, Moore RB, Timmis KN. An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. Environ Microbiol 2000;2:39–50.
14. Girvan MS, Bullimore J, Pretty JN, Osborn AM, Ball AS. Soil type is the primary determinant of the composition of the total and active bacterial communities in arable soils. Appl Environ Microbiol 2003;69(3):1800–9.
15. Franklin RB, Mills AL. Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. FEMS Microbiol Ecol 2003;44(3):335–46.
16. Kang SH, Mills AL. Soil bacterial community structure changes following disturbance of the overlying plant community. Soil Science 2004;169(1): 55–65.
17. Tebbe CC, Vahjen W. Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. Appl Environ Microbiol 1993;59(8):2657–65.
18. Paul EA, Clark FE. Soil microbiology and biochemistry. San Diego: Academic Press, 1989.
19. Brodie E, Edwards S, Clipson N. Soil fungal community structure in a temperate upland grassland soil. FEMS Microbiol Ecol 2003;45(2): 105–14.
20. Chabrerie O, Laval K, Puget P, Desaire S, Alard D. Relationship between plan and soil microbial communities along a successional gradient in a chalk grassland in north-western France. Appl Soil Ecol 2003;24:43–56.
21. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. Influence of depth and sampling time on bacterial community structure in an upland grassland soil. FEMS Microbiol Ecol 2003;43:35–43.

Additional information and reprint requests:
Lorraine E. Heath, B.Sc.
Arizona Department of Public Safety
Central Regional Crime Laboratory
2102 W. Encanto Blvd, Mail Drop #1150
Phoenix, AZ 85009
E-mail: leheath@cox.net