Forensic soil comparisons based on bacterial population profiles

F.C.A. Quaak, L. Joosten and I. Kuiper[1], Netherlands Forensic Institute

Introduction

Soil is encountered as trace evidence in various forensic cases and can contain biotic and abiotic parameters. These parameters (e.g. soil type, pollen, diatoms, microbes) can be used to predict the source of origin or to compare a sample with other soil samples.
The bacterial population is such a parameter which can be used to compare soil samples. Bacterial populations are adapted to the local biotic and abiotic conditions, can fluctuate over the seasons and during severe weather conditions, but are usually quite stable within shorter time intervals.
By direct DNA extraction from soil, bacterial DNA-marker amplification and modification, followed by visualization of the terminal restriction fragment length polymorphism (tRFLP), a bacterial profile from soil samples can be generated in a fast and cost effective way.

tRFLP method design

- DNA extraction from soil: various commercial extraction kits were tested. In our hands the FastDNA Spin kit for Soil (QBiogene) showed highest yield and reproducibility.

- From the DNA extract the 16S rDNA marker was amplified using universal primers[2]. Digestion of the product with enzymes *Hha*I and *Alu*I showed best results for *Hha*I, giving fragments in the range of 100-500bp.

- The fragments were analysed using capillary electrophoresis (ABI Genetic Analyzer 3130XL, POP4 and 30 sec. injection time) which results in a profile as shown in figure 1, panel A.

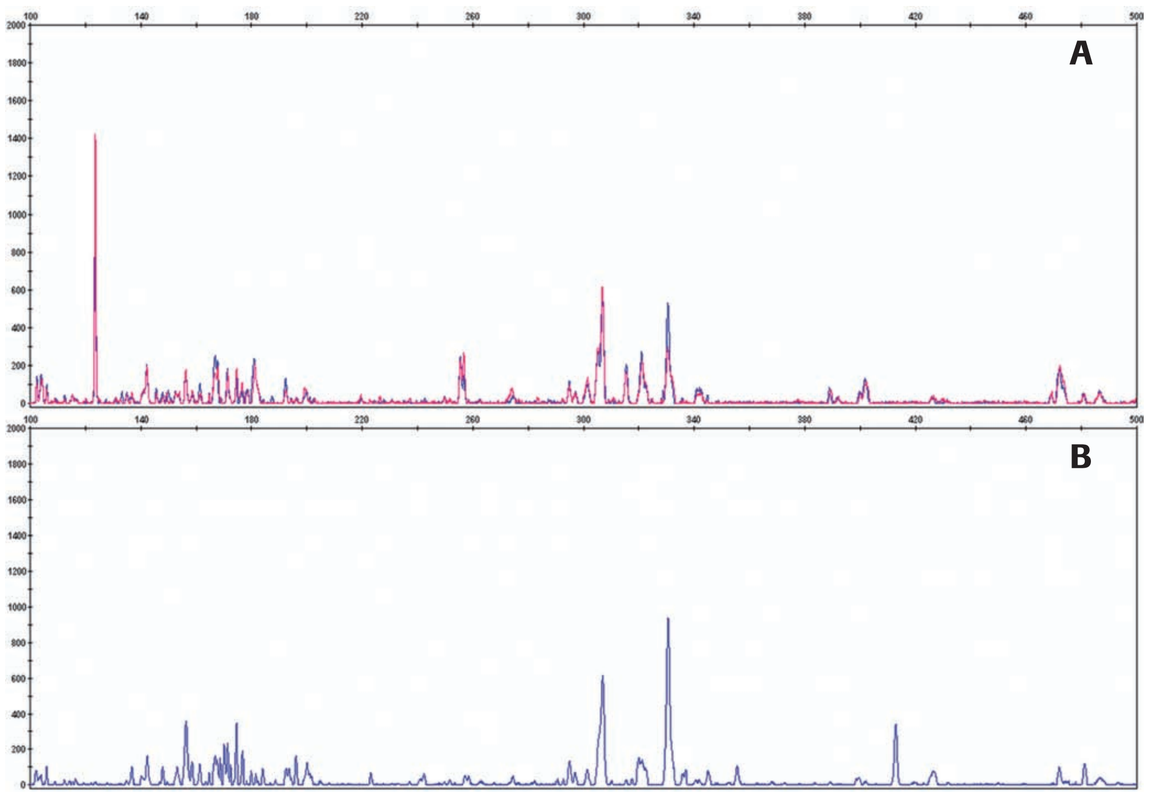

Figure 1: Bacterial profile using the tRFLP method. Panel A shows profiles from duplicate extractions of the same soil sample (red and blue line). Panel B shows a profile of a second soil sample.

Application of the method

- Reproducibility: 50 soil samples were analysed in duplicate (both extraction and PCR); all duplicate extractions and PCRs gave similar profiles (figure 1A).

- Robustness: profiles from almost any soil type (e.g. sand, silt, clay, potting soil) were generated. In order to test the influence of storage conditions, samples were kept at room temperature, at 4°C and frozen. For the samples tested, no differences were observed in the derived bacterial profiles.

- Discriminatory value: All 50 samples were collected from different locations throughout The Netherlands. All samples gave different bacterial profiles (Figure 1, panel A and B).

Comparison of bacterial profiles

In order to compare bacterial profiles in an objective and reliable way, information is required on minimal and maximal similarity of profiles derived form the same site and from different sites. For this purpose a method for data handling resulting in a database of reference samples was constructed.

Figure 2 represents the various possibilities and steps which are followed to derive a method for data handling:

- After generating a duplicate profile, a consensus profile can be constructed or both separate profiles can be taken into account. Working with single profiles is preferential because no information is lost.

- Both binary and continuous data can be used in the comparison of bacterial profiles. Using continuous data prevents loss of information by taking peak area into account.

- When using an alignment of single profiles, normalised or standardised data is required. Here normalisation is chosen, since this is the most common method.

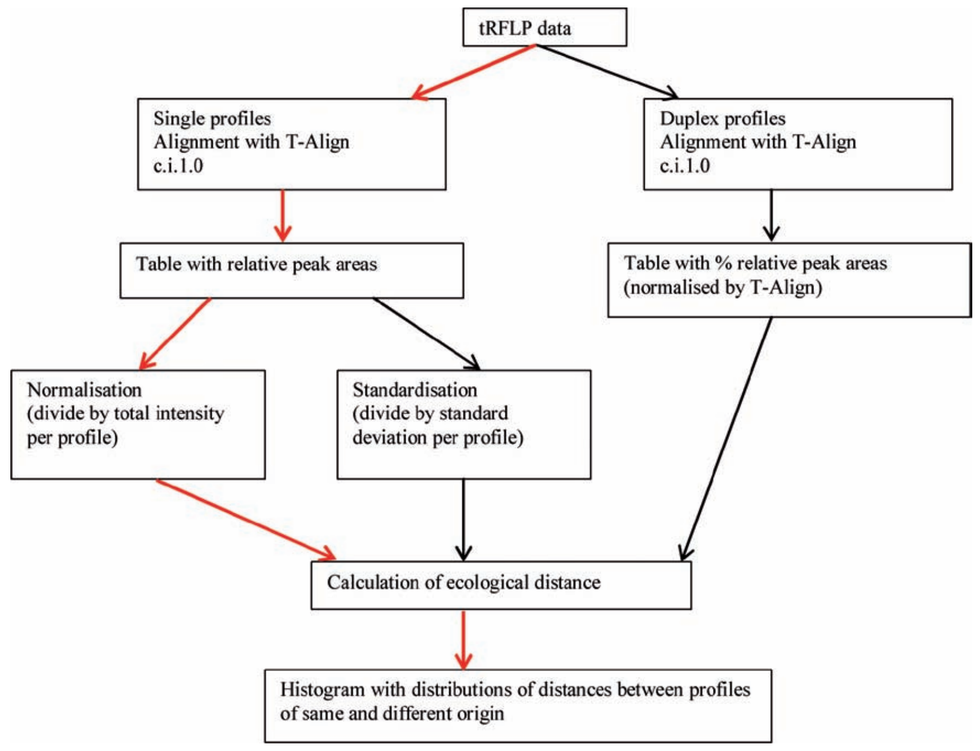

Figure 2: Flowchart for tRFLP data handling.

- Ecological distances are suitable to express the similarity between tRFLP-profiles. We compared different distance coefficients (Jaccard, Bray-Curtis, Euclidean) between profiles from the same and different origin by making histograms. Using Bray-Curtis distance, the range of distances between samples with the same origin and with different origin are best separated (smallest overlap, see figure 3).

- When the histogram is built upon a database representing a high number of samples and with samples which represent sampling strategies which can be encountered in case-work (e.g. close proximity, collected with different time intervals), it can be used as a 'decision' model to determine whether a soil sample and reference sample may have the same origin.

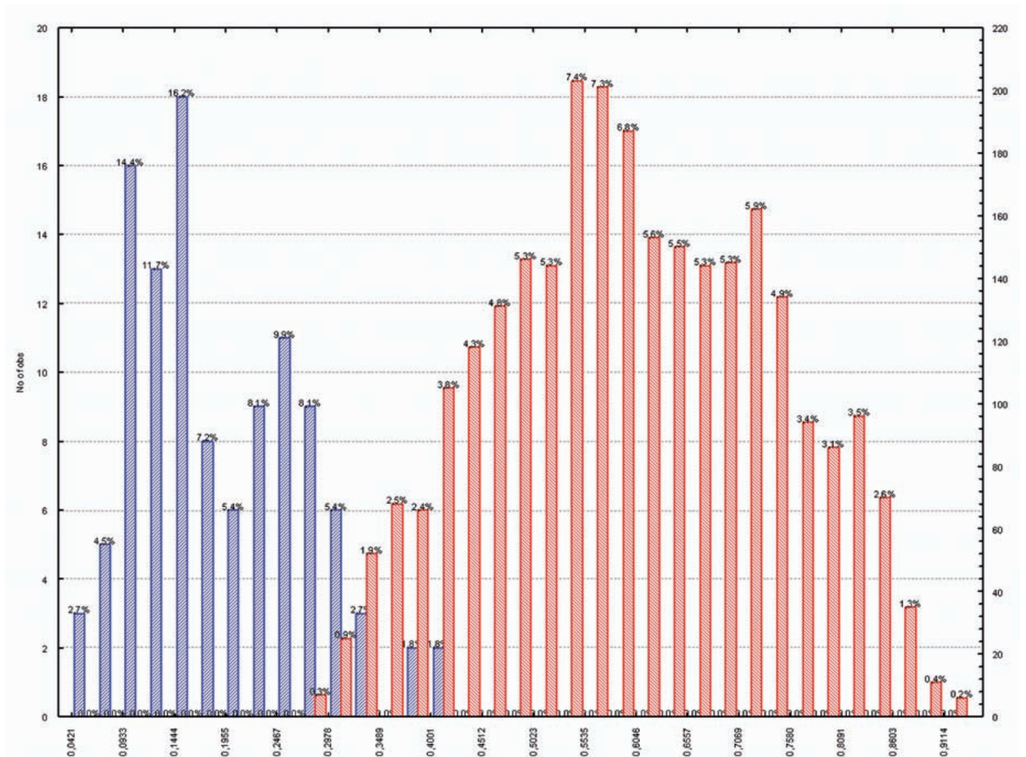

Figure 3: Histogram with Bray-Curtis distances calculated after comparison of bacterial profiles from different locations (red bars) and from the same location (blue bars). The y-axis gives the number of observations.

Conclusions

- tRFLP is a reproducible and robust method to generate bacterial profiles from soil.
- Data analysis using single profiles, normalisation and calculation of Bray-Curtis distances shows good results in the comparison of bacterial profiles.
- So far the method looks promising for comparison of forensic soil samples.

Ongoing and future projects

- Increase the number of profiles in the reference database
- Apply data analysis method on other dataset, e.g. pollen data

Notes

[1] Corresponding author: i.kuiper@nfi.minjus.nl

[2] FAM63f: 5'-CAG GCC TAA CAC ATG CAA GTC-3'
1389r: 5'-ACG GGC GGT GTG TAC AAG-3'