USING DNA FINGERPRINT ANALYSIS TO UNDERSTAND THE MICROBIAL ECOLOGY OF WASTEWATER IN ON-SITE SYSTEMS

Michael Linich

School of Environmental and Life Sciences, The University of Newcastle

Abstract

This paper reports on the use of Terminal Restriction Fragment Length Polymorphism (TRFLP), a form of DNA fingerprinting, as a tool in beginning the analysis of the microbial ecology of an on-site wastewater treatment system. 16S DNA found in primary treated effluent was subjected to TRFLP and the results compared to the 16S DNA of common human bowel microorganisms and likely candidates in the present NCBI 16S DNA Database. The data was also subjected to TAP-TRFLP analysis at the Ribosomal Database Project (RDP).

Based on sequence analysis it was found that there is an abundance of unknown bacteria. There appears to be an association between the presence of Clostridia species and the patterns found by TRFLP and a weak association between common and abundant bowel bacteria and that found in primary treated effluent. *E.coli* and *B. subtilis* could be identified.

TAP –TRFLP analysis indicates that environmental bacteria related to species known only by their 16S DNA sequence database depositions are common, suggesting that a significant proportion of the microbial ecology of effluent remains unknown.

Key words

E.Coli, Gut associated bacteria, microbial ecology, TRFLP, TAP –TRFLP, 16S DNA,

1 Introduction

The composition and stability of microbial communities, generally, is not well understood. Microbial analysis by culture-dependent methods has underpinned our current knowledge and public health monitoring of wastewater. However, it will not be long before molecular fingerprinting methods are advanced enough to become routine methods in public health management. These methods are capable of providing a more detailed understanding of the microbial dynamics of wastewater treatment systems and a cheaper, more accurate, way of monitoring microbes of concern to public health.

Molecular microbial analysis has revealed that complex communities of microorganisms can change spatially within distances of less than a few centimetres and temporally over a very short period of time. Wastewater treatment systems provide fluctuating habitats for microbes with variations in living microbial and biomass inputs and hydraulic flows. Thus, with treatment of wastewater one aims for and expects a change in the microbial flora and fauna (eg. the removal of pathogens) before discharge to the natural environment.

The design of on-site wastewater treatment systems are now beginning to address particular microbial physiologies such as oxidation or aerobic metabolism, as occurs in aerated systems, and denitrification and phosphorus removal as occurs in larger biological nutrient removal (BNR) systems. Such systems can only be adequately understood by molecular technologies as many of the important species involved are either expensive to monitor or uncultivable. Achieving sustainable development in relation to the preservation and quality of public health is a most important objective in relation to on-site systems. Thus a better understanding of the survival and movement of pathogenic organisms through wastewater treatment systems and in the environment generally, along with the reduction of nitrates and phosphates by encouraging the growth of specific organisms are crucial to this objective.

DNA fingerprinting targets a gene or group of genes belonging to the individuals of interest and amplifies them by using a specific DNA primer in the polymerase chain reaction (PCR). This gives multiple copies of the DNA sequence of the target gene, often unique to the individual. In this investigation the individuals are bacteria and the gene being targeted is the 16S rRNA gene (16S DNA) that codes for the small subunit of the ribosome in the bacterial cell. 16S DNA a highly conserved gene common to all bacteria and it exhibits polymorphism in sequence identity between species. The same basic method can be used to target other genes in virus, protozoa or groups of organisms, which share a common gene. Once the PCR amplicons are available this DNA is then digested by DNA restriction enzymes that cut the DNA at specific sites into smaller fragments of varying size. These fragments can be separated from one another using gel electrophoresis, stained, to make visible and used to describe a DNA fingerprint unique to the individual or community.

In this experiment a process called Terminal Restriction Fragment Length Polymorphism (TRFLP) was used to study the effluent coming from an on-site wastewater treatment system. In essence TRFLP uses a fluorescent tag to label the DNA primer that initiates amplification in the PCR. This allows the terminal DNA fragment of the DNA digest to be detected by fluorescence spectrophotometry. In bacteria the terminal DNA digest fragment can often be used to identify the species involved (Khan *et a*l 2001), especially when a number of target specific restriction enzymes are used to digest the DNA. The benefit of TRFLP is that the results can be recorded, analysed transmitted and stored in electronic format.

Human gut associated microbes are already subject to molecular analysis and comparisons can be made between raw and treated effluent to assess changes in microbial ecology. This information will relate to the abundance and survival of microbes in treatment processes. The human gastrointestinal tract is colonized by a diverse microbial community of several hundred bacterial species, with 30–40 species accounting for 99% of the total microflora (Wang *et al* 2002) thus it should be possible to identify the most abundant organisms in raw human sewage and track their fate through the wastewater treatment process.

2 Materials and Methods

2.1 Sample source

One hundred and twenty millilitres of effluent was sampled from an on-site wastewater treatment system in sterile plastic screw-top containers. The treatment system is as described in (Geary *et al* 2001). Samples were taken from the sandfilter-dosing tank, above and below the clinoptilolite (zeolite) in the sand filter and from the header tank just prior to discharge.

2.2 DNA extraction

Duplicate samples were centrifuged in sterile 12 mL polypropylene tubes and solids concentrated into a pellet. Each sample was concentrated by centrifugation at 3000 rpm for 5 min. The supernatant was discarded and the pellet was resuspended in 1 mL of CLS-TC (Bio1101, CA, USA). Total DNA was isolated from the sample pellet using the FastPrep™ DNA System (Bio101, CA. USA) as described in Yeates and Gillings (1998). DNA quality and concentration was checked by agarose electrophoresis.

2.3 PCR amplification

DNA purified directly from effluent was used as the template for the polymerase chain reaction (PCR) amplification of bacterial 16S DNA (16S rRNA gene sequences). Duplicate PCRs were performed with a Hybaid Omn-E thermal cycler (Hybaid Ltd, Teddington, Middlesex, UK). Negative controls containing water only, and positive controls containing *E. coli* 16S rDNA, were included in each set of reactions.

The universal primers 27f and 1492r were used for the initial amplification of Eubacterial 16S rRNA genes from total biomass DNA. Primers Hex-27f (fluorescent) and 1492r amplify nearly full-length 16S rRNA genes from most Eubacteria (Lane 1991). Fifty microlitre PCR volumes were standard. Final concentrations of reagents were as follows: 20 mmol L^{-1} $(NH_4)2SO_4$, 75 mmol L^{-1} Tris-HCl (pH 9.0), 0.01% (w/v) Tween 20, 2 mmol L^{-1} $MgCl_2$, 0.5 mmol L^{-1} of each primer, 0.25 mmol L^{-1} of each deoxyribonucleotide triphosphate, 10 ug / mL RNAse A and 1 Unit of Red Hot DNA Polymerase (Advanced Biotechnologies, Surrey, UK).

The cycling parameters for 16S DNA amplifications were as follows; denaturation at 94ºC for 3 min, followed by 35 cycles of 94ºC for 30 seconds; 60ºC for 30 seconds and 72ºC for 1 minute and 30 secs; 72ºC for 5 minutes.

2.4 DNA Digests

Amplified products from the16S rRNA gene PCR from effluent were digested with the restriction enzymes *Hinf*1 and *Rsa*1 (Promega, NSW Aust.) according to the manufacturers instructions, in 50 microlitre reaction volumes overnight at 37ºC.

2.5 Electrophoresis and TRFLP – DNA Fingerprinting

DNA extracts, PCR products and PCR digests were resolved by electrophoresis in 1%, 1% and 2.5% agarose gels, respectively, and run in TBE buffer containing ethidium bromide (0.5 ug/ml) (Bio-Rad, Sydney, Australia) at 110 volts. DNA was visualised by UV excitation and recorded by Polaroid photography [12]. A 1Kb Plus DNA ladder (Life Technologies, MA, USA) was run on every gel.

2.6 Fragment comparison

TRFLP gene scanning was accomplished by submitting four microlitres of digested DNA to the gene sequencing facility at Macquarie University where the 16S rRNA gene fragments were determined using a Perkin Elmer automated DNA sequencer (model 377; Applied Biosystems, CA, USA).

Selected human gut associated microbes (Wang et al., 2002) and other common wastewater bacteria were chosen and their 16S DNA sequence obtained from the NCBI databank at the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST/)

through the National Library of Medicine. *In-silico* digests of this sequence data was then accomplished using 'Jellyfish™' (Labvelocity Inc USA) sequence analysis program.

The TAP-TRFLP online analysis program at the ribosomal database project (RDP) (http://rdp.cme.msu.edu/html) at the Centre for Microbial Ecology Michigan State University was then compared to these data. This approach combines restriction fragment analysis of a phylogenetic marker with automated sequencing gel technology and permits the rapid profiling of the microbial community in the effluent. The reference strains used for this analysis are listed in Table 1.

3 Results

3.1 DNA Extraction

Figure one shows the differences in quality of the DNA extracts from each sampling point.

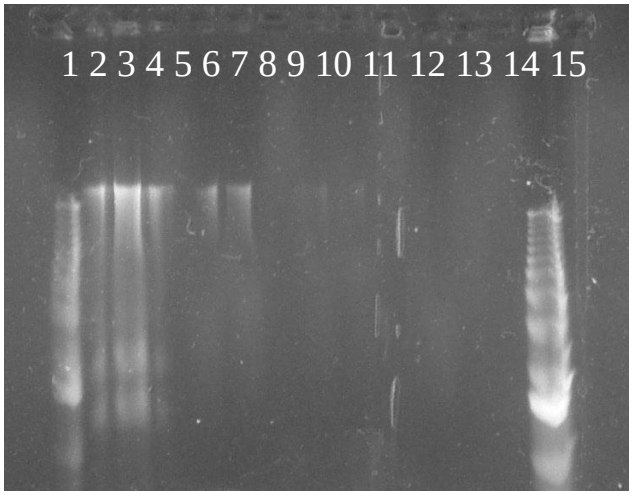

Figure 1: DNA extract from treatment system

Lane 1,2 Filtered effluent from septic tank.
Lane 3,4 Effluent above zeolite in sand filter
Lane 4,6 Effluent from below zeolite in sand filter
Lane 7,8 Effluent in last holding tank
Lane 14 1kbl DNA ladder.

Note the smearing in the DNA indicating DNA of poor quality (decomposing) and decreasing in quantity from the front to the rear end of the process.

3.2 PCR amplification

Figure 2 shows the differences in DNA amplification from each sampling point. Since the only amplified product recovered from PCR after several attempts came from the raw effluent it was decided to proceed with TRFLP analysis of this sample only.

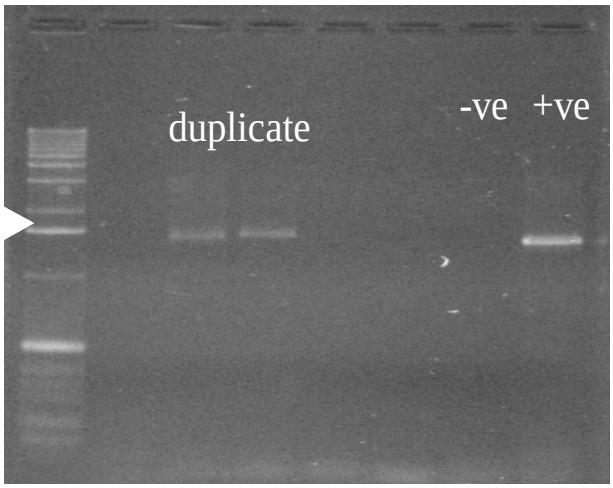

Figure 2: PCR of effluent DNA.

Lane 1 = 1kbL
Lane 2,3 septic tank filtered primary effluent
Lane 7 = negative control
Lane 8 = positive control.

3.3 DNA Digest

Figure 3 shows DNA restriction digest patterns with *Hinf*1 and *Rsa*1 enzymes respectively.

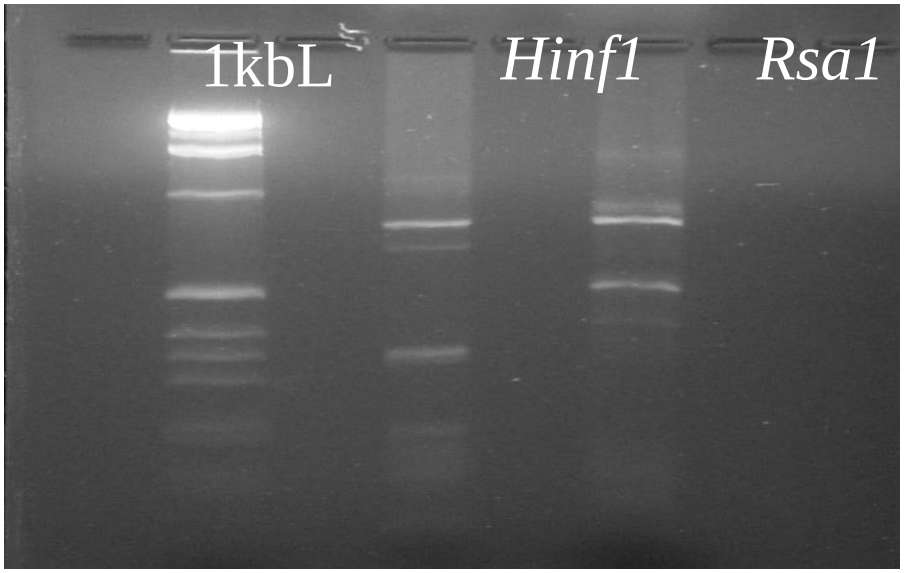

Figure 3: PCR digest with Hinf1 (Lane 4) and Rsa1 (Lane 6) and 1kb DNA ladder (Lane 2)

3.4 TRFLP – DNA Fingerprinting

The reference strains used for this analysis are listed in table one and a summary of the TAP-TRFLP results are listed in Table 1. Figure 4 and 5 shows TRFLP patterns obtained from each digest.

Table 1: Common and abundant bowel bacteria and their Terminal Restriction Fragment length.

Common Human Bowel Associated Microbes	**Restriction Enzyme *Hinf1* Terminal Fragment Length**	**Restriction Enzyme *Rsa1* Terminal Fragment Length**
Lactobacillus acidophilus	266	794
Escherichia coli	329	425
Enterococcus faecalis	344	898
Enterococcus faecium	361	890
Shigella boydii	220	316
Ruminococcus albus	129	442
Aeromonas hydrophila	330	883
Bifidobacterium longum	320	454
Fusobacterium prausnitzii	323	446
Camplyobacter jejuni	333	457
Pseudomonas aeruginosa	81	641
Salmonella species	325	442
Ruminococcus callidus	150	422
Clostridium perfringens	330	453
Clostridium clostridiiforme	97	475
Bacteroides fragilis	184	*no fragment*
Bacteroides vulgatus	323	*446*

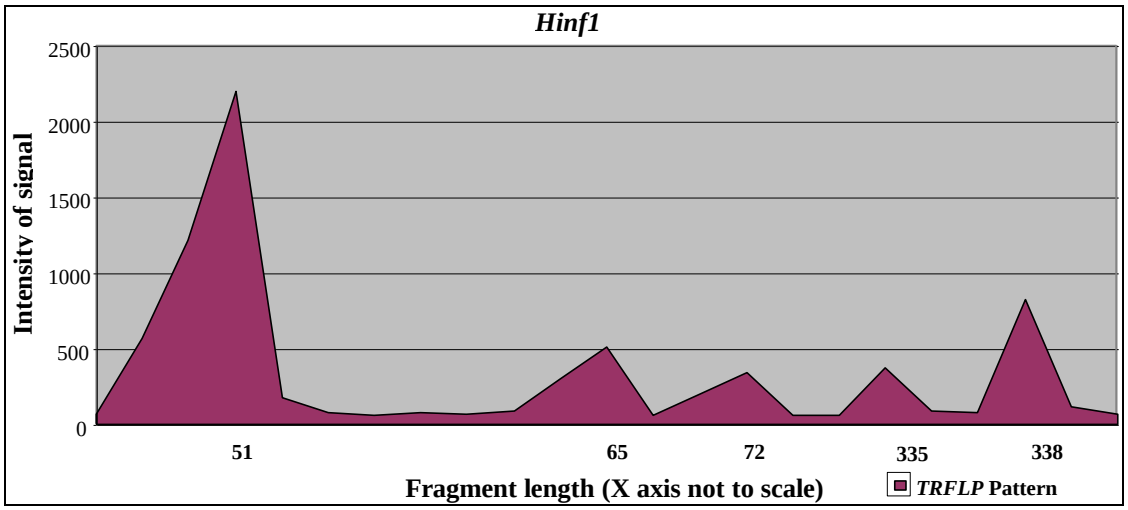

Figure 4: *Hinf1* TRFLP analysis.
Peaks show dominance of fragment length and area relative abundance.

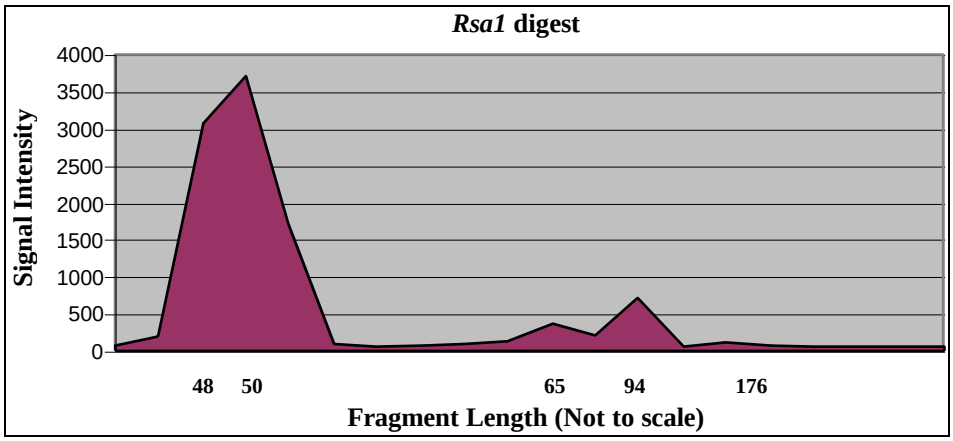

Figure 5. *Rsa1* TRFLP analysis.
Peaks show dominance of fragment length and area relative abundance.

4 Discussion

It could be argued, based on current knowledge of cultured bacteria, that we would expect some similarity between human bowel bacteria and bacteria known to be found in or associated with primary treated filtered wastewater effluent from a small domestic household. However, this does not appear to be the case. TRFLP analysis indicates there are at least two major phylogenetically related clusters of microorganisms in the final effluent with the possibility of three to four more minor clusters. TAP–TRFLP analysis of over 3000 database sequence gives no correlation between organisms known to have *Hinf1* and *Rsa1* terminal restriction fragments from at least two of the major groups (first two peaks on Figures 4 and 5). This finding supports the understanding that at least 95% of environmental bacteria are unknown by culture methods (Bergquist and Saul 1996). This finding highlights the relative absence of expected organisms and the presence of other species of bacteria that are becoming known only by their sequence identity in the NCBI and RDP databases as 'environmental clones.'

The presence of bacteria in the effluent identified by TAP –TRFLP whose *insilico* restriction fragments also correlate to the observed TRFLP patterns, but not noted amongst common human gut associated bacteria in the published literature, supports the idea that significant changes in microbial diversity and abundance takes place in wastewater both temporally and spatially. This process begins as the microbial environment changes when human waste is released into the wastewater stream. Anecdotal evidence also suggests that an individual's bowel flora varies temporally depending on the diet and state of health. Such changes in diversity, temperature and chemistry results in changing species diversity and abundance that facilitates a microbial succession that stimulates further changes in diversity and abundance patterns.

When TAP –TRFLP results are correlated with peak heights and TRFLP patterns most organisms are found to be organisms related to Clostridia, Mycobacteria, sulphur metabolising bacteria, Pseudomonas, expected microbes such as Bacteroides and *E. coli* and sequence data associated with uncultured environmental clones known only from gene sequences deposited in the database. Many of the groups mentioned have pathogenic or opportunistic pathogens in their genera (e.g. Vibrio species *Hinf1*/625, *Rsa1*/428). One can identify organisms represented by sequence data and bowel related flora (eg. *E.coli Hinf1*/330, *Rsa1*/428) as well as bacteria that one would expect to inhabit soils (eg *Bacillus subtilis Hinf1*/336, *Rsa1*/476) and anaerobic biomass fermenting environments (Lactobacillus spp, *Desulphobacterium baculatum Hinf1*/335, *Rsa1*/59).

There are technical points to be made about the nature of the data obtained. PCR does not discriminate between cultivable and non-cultivable bacteria so all targeted viable DNA would have been amplified. However, there are some bacteria that have multiple copies of the 16s rRNA gene. The Clostridia bacteria are amongst this group and as such their 16S DNA may be more preferentially amplified and if they were also abundant this amplification would amount to several orders of magnitude. The DNA of non-growing but viable spores of spore forming microbes will also be amplified so the patterns observed may not be a true reflection of the relative abundance of organisms participating in that community.

The Clostridia are anaerobic spore forming bacteria that only produce spores when the oxygen levels increase or the water chemistry changes. The detection of several Clostridia clusters in this study indicated the importance of this diverse group and as they are of interest to public health may serve as suitable targets for future molecular analysis. *Clostridium perfringens*, whilst not detected because of the limitations of this dataset, may prove to be a suitable candidate for further analysis because it is an organism correlated specifically with human faecal matter. Whilst this study did not detect *Clostridium perfringens* by TRFLP it was expected that it would be found. This indicates the method may be sensitive to the slight phylogenetic sequence variations that exists from region to region and suggests the use of a third restriction enzyme to improve the accuracy by allowing better prediction with respect to observed TRFLP peaks.

These findings have emphasised the importance of molecular studies to enlighten our understanding of the immense diversity of microbial life in nutrient rich environments. Such knowledge will allow us to better manipulate microcosms for sustainability and better understand and monitor them for the protection of public health.

References

Bergquist P. L. and Saul D. (1996) Microbial Diversity- Its Unexplored Territory. *Microbiology Australia 22-24*.

Geary P., Stafford D., Spinks A. and Whitehead J. (2001) The use of clinoptilolite in a low cost aerobic sandfilter. ***In***: *Proceedings of On-site'01 Conference Advancing On-site wastewater treatment systems*. Armidale.

Khan A, Nawaz M, Robertson L, Khan SA, Cerniglia C. (2001) Identification of predominant human and animal anaerobic intestinal bacterial species by terminal restriction fragment patterns (TRFPs): a rapid, PCR-based method. *Mol Cell Probes.15 (6) : 349-55*.

Lane D. (1991) 16S/23S rRNA sequencing. ***In*** Stackenbrandt E and Goodfellow M. (Eds) *Nucleic acid techniques in Bacterial Systematics*. John Wiley and Sons New York p 115-176.

Sambrook J. and Russel D. (2001) Molecular Cloning. Cold Spring Harbour laboratory Press.

Wang RF, Beggs MF, Robertson LH and Cerniglia CE. (2002) Design And Evaluation Of Oligonucleotide-Microarray Method For The Detection Of Human Intestinal Bacteria ***In*** *Fecal Samples*. FEMS Microbiology Letters 213, 2 pp175-182

Yeates C. and Gillings M. (1998) Rapid purification of DNA from soil for molecular biodiversity analysis. *Letters in Applied Microbiology 27. 49-53*.