AN ANALYSIS OF MERCURY RESISTANCE GENES IN

STREAM SEDIMENT BACTERIA, USING TERMINAL RESTRICTION FRAGMENT

LENGTH POLYMORPHISM (TRFLP)

by

ELIZABETH ADRIENNE BURGESS

(Under the direction of J Vaun McArthur)

ABSTRACT

Resistance to mercury is carried on mobile genetic elements that can be passed between members of a bacterial community. It was thought that the diversity of mercury resistance genes would increase with increasing mercury contamination and reflect a spatial distribution of sampling sites in a stream. Three sites along the stream were sampled in triplicate. DNA was extracted from these samples, PCR with fluorescent primers was used to amplify genes of interest and PCR products were digested with different restriction enzymes. Terminal restriction fragment length polymorphism analysis (TRFLP) was used to investigate the diversity and similarity of mer and 16S ribosomal genes from the three sample sites. Results indicate that choice of restriction enzyme, labeled end, and method of analysis affect the observed patterns. These results indicate that TRFLP can be used to compare patterns among genes, but the technique is limited when comparing results from different restriction enzymes.

INDEX WORDS: terminal restriction fragment length polymorphism, mercury resistance, bacteria

AN ANALYSIS OF MERCURY RESISTANCE GENES IN STREAM SEDIMENT BACTERIA, USING TERMINAL RESTRICTION FRAGMENT LENGTH POLYMORPHISM (TRFLP)

by

ELIZABETH ADRIENNE BURGESS

B.S.F., Stephen F. Austin State University, 1999

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2003

AN ANALYSIS OF MERCURY RESISTANCE GENES IN STREAM SEDIMENT BACTERIA, USING TERMINAL RESTRICTION FRAGMENT LENGTH POLYMORPHISM (TRFLP)

by

ELIZABETH ADRIENNE BURGESS

Major Professor:	J Vaun McArthur
Committee:	Travis Glenn
	Larry Shimkets

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2003

ACKNOWLEDGEMENTS

I would like to thank my committee; Dr. McArthur for giving me the opportunity to be an MS student at SREL, Dr. Glenn for good beer and technical advice, and Dr. Shimkets for smiling when he sees me. I would also like to acknowledge the following people for their help; Cary Tuckfield, Cathy King, Jim Novak, Angela Lindell, Olga Tsyusko, Mandy Schable, and Alessandra Seccomandi. Finally, I would like to express my appreciation to the Environmental Remediation Sciences Division of the U.S. DOE for Financial Assistance Award DE-FC09-96SR to the UGA Research Foundation.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

SECTION 1

INTRODUCTION

Mercury is a ubiquitous compound in our environment. The global distribution of mercury in various forms, elemental, organic or inorganic, is the result of both environmental and anthropogenic processes. Although human activity has increased the atmospheric burden of mercury as a contaminant in recent time, ancient, environmental sources of mercury, such as volcanoes and geologic deposits are theorized to have released mercury at greater quantities than presently observed (Osborn et al. 1997). Mercury is potentially toxic to all living things. The mercury ion Hg(II) binds to thiol and sulfhydryl groups of proteins and disrupts the integrity of most biological functions (RTI 1999). Bacteria, recognized among the oldest extant organisms on earth, have been mitigating mercury toxicity since early in Earth's history. Bacteria play a crucial role in the biogeochemical cycling of this metal. The suite of genes that code for mercury resistance in bacteria is a well-studied and widespread system, which has been used frequently in analyses of bacterial ecology and evolution (Bruce 1997; Osborn et al. 1997; Rasmussen and Sorensen 1998; Liebert et al. 1999).

Collectively called the *mer* operon, bacterial mercury resistance genes are found in different combinations on various transposable or conjugative elements and can detoxify mercury by multiple mechanisms, depending on the form in which the contaminant occurs. Perhaps the most common mechanism is the reduction of the mercuric ion Hg(II) into Hg(0), which volatilizes from the cell environment. The genes that code for this mechanism, including *merR*, *merT*, *merP* and *merA*, have been extensively studied (Barkay et al. 1989; Nazaret et al. 1994; Bruce et al. 1995; Liebert et al. 1997; Osborn et al. 1997; Rasmussen and Sorensen 1998; Liebert et al. 1999; Sandaa et al. 2001; McArthur et al. unpublished data). MerP, the product of the gene *merP*, is found in the cell periplasm and it is believed to transfer Hg(II) to the MerT protein,

which carries the mercury to MerA, the product of the mercuric reductase gene *merA* where Hg(II) is reduced to Hg (0) and volatilized. MerR is a regulator of the operon. When mercury binds to MerR, repression of operon transcription is inhibited (Liebert et al. 1999).

Not only the structure and function, but also the ecology, distribution and evolution of the *mer* operon have been described. Studies have shown that the diversity of the mercury reducing genes may vary with the level and duration of mercury contamination. Bruce et al. (1995) demonstrated a lower level of diversity in *mer* genes analyzed by restriction fragment length polymorphism (RFLP) in a site recently contaminated with high levels of mercury. This was in comparison to higher diversity found at a site with a longer history of mercury contamination. Another study indicates that genetic variation is greater in *merR* genes amplified from contaminated soil versus genes amplified from pristine soils (Osborn et al. 1995). It has also been demonstrated that mercury resistance genes can be carried on conjugative bacterial plasmids in the microbial community and that these plasmids were more abundant in a contaminated versus a non-contaminated marine harbor environment (Rasmussen and Sorensen 1998). In an informative review, Osborn et al. (1997) presented a hypothetical lineage of Gram-negative *mer* operons within their associated transposable elements and stressed the importance of considering such operons in descriptions of bacterial evolution. In an analysis of *mer* genes examined from culturable organisms in the feces of primates, it was indicated that some genes or regions of the operon are more conserved than others (Liebert et al. 1997).

The importance of understanding the distribution of mercury resistance genes in bacteria may also be of concern in human health interests. Previous studies have shown a relationship between mercury concentrations and bacterial resistance to antibiotics in stream environments (McArthur and Tuckfield 2000; Alonso et al. 2001). Aviles et al. (1993) demonstrated an association between mercury resistance and resistance to multiple antibiotics in strains of *Pseudomonas aeruginosa* isolated from marine environments. The genes coding for resistance to mercury and for resistance to multiple antibiotics have been shown occur on the same broad-host

range, conjugative plasmids (Rasmussen and Sorensen 1998). Additional descriptions of the distribution and diversity of *mer* operons may provide further insight into the relationship between antibiotic resistance and mercury resistance and the potential for indirect selection for antibiotic resistance via mercury contamination.

In previous work on the Savannah River Site (SRS) (McArthur et al. unpublished data) aspects of the dynamics of mercury resistance genes from stream sediment bacteria were uncovered. Samples were collected from Four Mile Creek, a contaminated stream on the SRS. The samples were analyzed for heavy metal concentrations, prevalence of antibiotic resistance via plating isolation and diversity of mercury resistance genes via cloning and restriction fragment length polymorphism (RFLP). Unique RFLP haplotypes were designated as clones sharing the same restriction fragment pattern comparing two-enzyme digestions with *Nde*I and *Sau96*I, and *Hae*III and *Hind*III. The data demonstrated a peak in proportion of isolates resistant to antibiotics and number of RFLP haplotypes where a tributary, Castor Creek, enters Four Mile Creek. Subsequently, sequencing and phylogenetic analysis of 2 to 6 clones from each of five sites within the most common RFLP haplotype designation, Haplotype 1, demonstrated an unexpected relationship among sequences. Contrary to the anticipation that spatially closer sites would have more similar *mer* gene sequences, a clade was formed indicating a relationship among sequences from distant sites above and below the confluence of Castor Creek and Four Mile Creek, distinguishing sequences collected at distant sites from those recovered closer to the confluence. Perhaps near the confluence of the two streams the tributary stream introduces a selective force, maybe a unique spectrum of contamination, for a certain subclass of *mer* gene sequences. Bacterial populations in Four Mile Creek would have adapted to this selection, via vertical or horizontal transmission of the appropriate genes, and the response would persist in the main channel until the contaminants are no longer selective, possibly via degradation, sorption or dilution. Alternatively, perhaps the different sequences elsewhere in Four Mile Creek are the result of selection pressures to which the sequences near the confluence are not suited.

With the advent of molecular ecology, analysis of bacterial genes has revealed a whole host of hithertofore unknown bacteria, as well as enabled investigations into the diversity and distribution of functional genes. These discoveries, however, are a mixed blessing. For example, although generally thought to be the most accurate method for describing a community of bacteria, cloning and sequencing is a labor-intensive and time-consuming process. Several steps are involved in getting from sample to sequence. Additionally, with a diverse community or an especially variable gene, it is often difficult to obtain a representative sample of sequences, due to the prohibitive number of clones that would have to be sequenced. RFLP analysis has been used to reduce the necessity analyzing sequences of a large number of clones in order to get a picture of genetic variation in a bacterial community. The interpretation of RFLP data is generally more straightforward than that of sequence data, but an essentially equivalent number of steps are required to obtain the data. When analyzing clone libraries, it is only possible to look at one gene at a time. Again, because it is important to collect representative data to describe a community, it is possible that many clones will have to be screened using RFLP before an accurate description of community diversity can be provided. Even when often practiced, the steps taken to create clones, characterize a sufficient number, and compile and analyze the data from each can take a considerable amount of time.

Recently, terminal restriction fragment length polymorphism (TRFLP) has emerged as an alternative. The method attempts to capture the genetic variation within a bacterial community by looking at all members of the community at once. Characteristics of a community are assessed via the variation in restriction enzyme target sites nearest a fluorescent-primer labeled end, or terminal fragment, from the gene(s) of interest. Ideally, it provides a profile of the major members of a community of bacteria and enables an estimate of genetic diversity at a given sampling site, and may be used to compare bacterial communities from different sites. This method has been used with to characterize bacterial communities, generally looking at the 16S rRNA gene (Dunbar et al. 2000; Liu et al. 1997; Blackwood et al. 2003). Braker et al. (2001)

used TRFLP to determine the diversity of nitrate reductase genes and 16S rDNA within and among distant sample sites. TRFLP has also been used to identify different subclasses of *mer* genes in contaminated and uncontaminated soil samples (Bruce 1997). In comparison with other methods, Moeseneder et al. (1999)found TRFLP to provide a profile of higher resolution than that provided by denaturing gradient gel electrophoresis (DGGE) for marine bacterioplankton communities. In their review of the method, Osborn et al. (2000) expressed that TRFLP is a potentially powerful tool in microbial ecology.

Environmental bacteria often demonstrate genetic diversity in response to selective pressures (McArthur 1998; McArthur et al. 1992; Wise et al. 1995; Wise et al. 1996). We sought to determine if the diversity of *mer* genes could be calculated for sites along Four Mile Creek using TRFLP. We looked for a relationship between gene diversity and the presence and abundance of contaminants, especially mercury, in the stream. It was thought that highest levels of gene diversity would reflect the highest levels of mercury contamination. In addition, we anticipated that TRFLP profiles from a site near the confluence would distinguish this site from those above and below the confluence, as had been indicated by previous results. Finally, we felt that the terminal fragment of the haplotype that occurred most frequently in the RFLP analysis would be the most abundant fragment in the TRFLP profiles. This study was developed to provide new or unique insights into the relationship between bacterial gene diversity and heavy metal contamination and the utility of TRFLP for haplotype identification and diversity assessment.

SECTION 2

METHODS

Study Site

The Savannah River Site (SRS) was established in western South Carolina in 1950 as a production facility for nuclear weapons-grade tritium and plutonium-239. The 750 km^2 set aside for SRS operations has been closed to the public since the site was opened and encompasses both contaminated and essentially pristine areas for ecological studies.

Four Mile Creek (Four Mile) is a third-order stream on SRS that has been subjected to industrial contamination in the forms of thermal effluent from nuclear reactor operations and chemical effluent from seepage basins. Additionally, a sewage treatment facility established along Four Mile in 1992 discharges 1 million gallons of treated wastewater into the stream per day. Previous work published by McArthur and Tuckfield (2000) described samples taken from Four Mile on the SRS in June 1998 (Figure 1). The described sample sites are located approximately 200m apart and were identified using a Trimble™ GPS unit, which enables accurate relocation of each sample site to within 1.5m. On 13.June.2002 we returned to three of the sample sites described here as sites 1, 2 and 3 (Figure 1) approximately 1.8 km apart. Three sampling locations at each site, nine total, were sampled from using a 191 cm length of 3.2 cm diameter polyvinyl chloride (PVC) pipe. The pipe was used to take a series of sediment cores that were transferred to Fisherbrand Sterile Sampling Bags (Fisher Scientific) as described by Wise, et al. (1995). Samples were labeled by site and sample number, stored on ice for approximately 1 to 6 hours, and transported back to the lab for further analysis. Samples from the same site were analyzed independently throughout and results shown on a per site basis are the means of values calculated from three samples.

Metals Analysis

A portion of each sample was oven dried at 50°C (Fisher Isotemp Drying Oven) for 24 hours then homogenized by crushing and passed through a 1.0 mm sieve (USA Standard Testing #18). Approximately 60 g were analyzed for metals at the University of Georgia, College of Agricultural & Environmental Science, Soil, Plant and Water Laboratory. Results are reported in ppm for Cd, Cr, Cu, Mn, Ni, Pb, and Zn. Approximately 3 g of the dried, sieved portion of each sample was reserved for mercury concentration analysis by EPA Method 7473 using a DMA-80 Mercury Analyzer (Milestone Microwave Laboratory System). Results are reported in ppb. The detection limit of this method is 0.01 ng of total mercury.

DNA Extraction

An additional portion of each sample was used for genetic analysis. DNA was extracted by a bead beating procedure from 300-400 mg stream sediment per sample using the FastPrep System (Bio101) and the FastDNA SPIN Kit for Soil (Bio101). The manufacturer's methods were modified to include washes of the Binding Matrix with 500 µL of 5.5 M Guanidine Thiocyanate 2-3 times and the repetition of the SEWS-M wash six times in order to reduce humic acid contamination. Humic acids, organic acids generated during the decomposition of organic matter, inhibit PCR. Two independent extractions were run for each sample. Each extraction resulted in approximately 300 µL of DNA of unknown concentration. These were stored at –20°C. The extraction product was treated with Ribonuclease A to eliminate RNA by adding 1 µL of 10 mg/mL stock Ribonuclease A solution (made from 0.001 g Ribonuclease A (Sigma RNAse) to 100 µL of sterile, nanopure water) to each extract tube. The contents were mixed by inversion and flicking and then incubated in a dry bath at 37°C for 1 hour (Type 16500 Dri-Bath from Thermolyne) and stored at –20°C. The extracted product was quantified against known concentrations of calf thymus DNA via agarose gel electrophoresis and ethidium bromide UV illumination using an Eagle Eye II Still Video System and the associated software, version 3.2 (Stratagene). When necessary to achieve sufficient concentrations of DNA for PCR (~25 ng/µL),

extract product was ethanol precipitated by adding 1µL glycogen solution (20 mg/mL) to extract, mixing by pipetting. Then 30 µL of 4 M LiCl and 993 µL of ice-cold 95% EtOH were added and mixed by pipetting or inverting. Tubes were incubated at –70°C for 30 minutes, then centrifuged at 14000 rpm (Eppendorf MicroCentrifuge 5415 C) for 15 min to pelletize DNA, and the supernatant was discarded. Samples were then washed with 100 µL ice-cold 70% EtOH, centrifuged again at 14000 rpm for 5 min, and the supernatant discarded. The resulting pellet was dried in a fume hood for ~1 hour and resuspended in nanopure water.

Due to difficulty establishing optimal PCR conditions, DNA from the original FastPrep extraction was used quickly for Sites 1 and 3 and it was necessary to run another extraction. DNA from Sites 1 and 3 for 16S gene analysis was from an extraction of approximately 8 g per sample using another bead beating procedure, the UltraClean Mega Soil DNA Kit (MoBio Laboratories, Inc.). In this case, only one extraction was made per sample. Manufacturer's methods were modified by homogenizing soil and bead solution before mixing with beads and included the addition of Inhibitor Removal Solution, an optional step. Centrifuging was done using a Sorvall centrifuge with an SLA-1000 rotor. The alternative extraction method was chosen because it yields a higher volume of product. Sufficient extraction product from the FastPrep system remained for 16S ribosomal gene analysis of Site 2. An equal aliquot from each parallel extraction was combined on a per sample basis. Again, for Sites 1 and 3, only one MoBio extract was made per sample.

PCR

Mercury Resistance Genes

Following successful DNA extraction, the product was used as template for PCR. The PCR served to amplify an approximately 1000 bp portion of the *mer* operon described as *merRTΔP* (Bruce et al. 1995) using the fluorescently labeled primers RXf:FAM (5'-/56-FAM/ATA AAG CAC GCT AAG GCR TA-3') labeled with 6-FAM™ (Fluorescein), which

fluoresces blue, and PXr:HEX (5'-/5HEX/TTC TTG ACW GTG ATC GGG CA-3') labeled with HEX™ (Hexachloro-fluorescein), which fluoresces green (Integrated DNA Technologies) (Figure 2). Terminal fragments from the RXf:FAM and PXr:HEX primers are called fam ends and hex ends respectively. Both ends were labeled to provide more data for analysis. These primers, designed by Bruce et al (1995), are based on conserved regions in the Tn*501*, Tn*21* and pMER419 sequences. All results are from a single PCR event (8.21.02). Each 25 µL reaction, 1 per extract, made from a PCR Master Mix, had final concentrations of reagents as follows: 0.4 µM each primer, 2 mM $MgCl_2$ (Life Technologies), 1X PCR buffer minus Magnesium (Life Technologies), 1X BSA, 0.15 mM dNTP (Promega), 5% DMSO, 1 U Taq (Life Technologies). The reactions were run on an Eppendorf Mastercycler Gradient thermocycler for the following program: an initial warming cycle at 94°C for 00:20; 5 cycles of 96°C for 00:30, 65°C for 00:30 and 72°C for 01:00; 21 cycles of 94°C for 00:20, 65°C for 00:30, -0.5 each cycle, and 72°C for 01:00; 15 cycles of 94°C for 00:30, 55°C for 00:30, and 72°C for 01:00; a final extension cycle at 70°C for 00:20; then held at 15°C. This type of thermocycler program, called touch-down, requires high-stringency annealing of primers to target regions as the reaction begins, but reduces stringency as the reaction progresses and the primers are increasingly likely to encounter previously produced product to amplify rather than erroneous regions of template DNA (Palumbi 1996)

The product was identified as *merRTΔP* by size at approximately 1000 bp (1 kb) as compared to a Hi-Lo™ DNA Marker (Minnesota Molecular catalog) using agarose gel electrophoresis as described above. Negative controls were water. PCRs were done for each extract and the resulting amplified DNA from replicate extractions were combined to make a single PCR product per sample, three per site. This process resulted in 50 µL PCR product volumes at concentrations of 1-10 ng/µL, which were stored at –80°C.

16S Ribosomal Genes

As part of a supplemental survey to test the effect of different restriction enzymes on estimations of diversity using TRFLP, we repeated much of our methodology on a portion of the 16S ribosomal RNA gene (16S rDNA). MoBio and FastPrep DNA extraction products were used for PCR to amplify a portion of the 16S ribosomal gene. The fluorescently labeled primers 27f:FAM (5'-/56-FAM/AGA GTT TGA TCM TGG CTC AG-3', where M=A or C) labeled with 6-FAM™ (Fluorescein) and 1518r:HEX (5'-/5HEX/AAG GAG GTG ATC CAN CCR CA-3', where N=any base) labeled with HEX™ (Hexachloro-fluorescein) (Integrated DNA Technologies) were used to amplify an approximately 1500 bp region of the 16S rDNA known to represent most eubacteria. Each 25 µL reaction, 1 per extract, made from a PCR Master Mix, had final concentrations of reagents as follows: 0.5 µM each primer, 1.5 mM MgCl (Sigma), 1X PCR buffer minus Magnesium (Sigma), 0.2 mM dNTP (Promega), 0.625 U Jump Start Taq (Sigma). The reactions were run on an Eppendorf Mastercycler Gradient thermocycler for the following program: an initial warming cycle at 95°C for 05:00; 30 cycles of 95°C for 00:20, 60°C for 00:20 and 73°C for 01:30; a final extension cycle at 72°C for 10:00; then held at 15°C. Success of amplification was determined by size at approximately 1500 bp versus a Hi-Lo marker and a lack of product in negative controls using agarose gel electrophoresis and UV illumination.

Enzyme Digest

Enzyme digest reactions were designed based on the manufacturer (New England Biolabs, Inc.) recommendations for a 10-fold over-digestion factor to ensure completeness. The over-digestion factor (OF) was calculated as OF=(units of enzyme) x (hours of reaction)/(µg of DNA). Effort was made to design the reaction using the smallest volumes possible for conservation without sacrificing accuracy due to pipettor error. The approximate quantities of DNA after amplification varied from 1-10 ng/µL (0.001-0.01 µg/µL). Master mixes were made for 10-20 U/µL enzymes using (0.1 µL enzyme + 0.9 µL 1X PCR buffer) times the number of reactions and for 5 U/µL using (0.2 µL enzyme + 0.8 µL 1X PCR buffer) times the number of

reactions. Each 5 µL reaction was made with 1µL digest master mix and 4 µL PCR product and incubated at 37°C for 1-2 hours.

A series of enzymes were selected based on their availability in the lab and their activity in PCR buffer. Results are reported for *mer* genes and 16S rDNA from the enzymes *Bsr*I (5'-ACTGGN^/TGAC^CN-5'), *Hae*III (5'-GG^CC/CC^GG-5'), *Hinf*I (5'-G^ANTC/CTNA^G-5'), and *Sau96*I (5'-G^GNCC/CCNG^G-5'). For *mer* gene analysis, *Rsa*I (5'-GT^AC/CA^TG-5') was also included. Digest products were stored in the –80°C freezer.

TRFLP

Mercury Resistance Genes

TRFLP methods were based on those used by Bruce (1997) in combination with methods used to analyze microsatellites (Glenn laboratory protocols). One µL of sample (digest product) was combined with 3.5 µL of a Master Mix made from 0.5 µL MapMarker® 1000 (BioVentures, Inc.) with an X-Rhodamine label, which fluoresces red, 0.5 µL blue tracking dye and 2.5 µL deionized formamide per reaction as per BioVenture's specifications. Samples were heated to 95°C for 5 min on an Eppendorf Mastercycler Gradient thermocycler, then were kept on ice and in dark until 1.5 µL of each were loaded on a 12 cm polyacrylamide (5 mL 5X TBE, 9 g urea, 10 mL diH_2O, 9 mL Acryl/Bis, a pinch resin beads, 15 µL Temed, and 125 µL 10% APS) gel. Gels were run in XL mode on a Perkin Elmer ABI Prism 377 DNA Sequencer for 3 hours at 2400 volts using filter D against a matrix developed in-house and GeneScan software v. 3.1.2 (Applied Biosystems). Due to the higher sensitivity of the ABI Prism analysis as compared to agarose gel/ethidium bromide imaging, enzyme digest negative controls, sample PCR and multiple enzymes per sample were run. This enabled elimination of erroneous peaks during later analysis (e.g. peaks occurring in negative controls as well as sample lanes were removed from sample lanes).

Data extraction was done using the local southern method, including all information above 50 fluorescence units in the analysis. This fluorescence limit was chosen by assessing the

lowest level at which most noise dropped out of the electropherograms. The local southern method consistently gave the most accurate size calling for the internal ladder. The size standard was defined for each run as well as by lane when necessary to get an accurate assessment of size. Fragments were analyzed using Genotyper v. 2.5 (Applied Biosystems). After binning all peaks within +/- 1 bp, peaks occurring in multiple enzymes, blanks or PCRs within a sample were removed from analysis. The remaining peaks were identified as individual fragments of associated abundance (peak height) for diversity calculations or converted to binary code by presence or absence of a peak of a particular size for similarity calculations. Results are reported as the averages of at least two or three TRFLP replicates per sample and three samples per site, generally from a single enzyme digest reaction.

16S Ribosomal Genes

TRFLP gels for the 16S rDNA analysis were essentially the same as that for *mer* gene analysis. Although the amplicon is larger, the same ladder was used. For these genes, analysis of data was done as described by Stepanasukas et al. (2003) using TRFLP Tools, a program written in Visual Basic for Microsoft Excel. Again, peaks occurring in negative controls were removed from analysis.

Data Analysis

Site Profiles

Data from TRFLP gels of *mer* gene digests with *Hae*III and *Hinf*I were used to create profiles for each site. The relative peak heights were complied into bar graphs to illustrate the distribution and abundance of fragments within a site. The average number of fragments counted, the average total height, the average fluorescence of individual peaks and the variance of normalized peak height per site were calculated.

Diversity

Data from both the 16S ribosomal genes and the *mer* genes were analyzed to determine the diversity among sites and whether the diversity values were correlated with the concentrations

of the heavy metals. These calculations were done for the fam ends of all the enzymes mentioned above. For the 16S rDNA, diversity was also calculated from hex ends for all the enzymes used to digest the ribosomal genes. Hex ends were analyzed for *mer* gene diversity from *Hae*III and *Hinf*I digests only.

Diversity was calculated using the Simpson index as H_{si}, where

$H_{si}=1-\Sigma p_i^2$, and

$p_i=n_i/N$,

or the proportion of individuals i in the total population N. The value of n_i was the peak height of an individual fragment and N was the total of all peak heights of the same color in a given gel lane. The Simpson index has a range of 0 to 1 where 1 is the most diverse. Per site diversity of *mer* genes presented in the results are averages of two or three TRFLP replicates per sample and three samples per site. A General Linear Model (GLM) was used to test the significance of differences in *mer* gene diversity among sites and enzymes and between ends for the enzymes *Hae*III and *Hinf*I (SAS 2003). Results of diversity calculations are the only data analysis reported from our investigation of 16S rDNA.

Similarity

Using the data from the *mer* gene TRFLP gels, similarity was calculated from binary matrices based on the presence or absence of a given peak in a profile and the NTSYSpc Numerical Taxonomy and Multivariate Analysis System v. 2.1, Similarity for Qualitative Data module. The similarity coefficient chosen was the R-T coefficient (Rogers and Tanimoto 1960). In this case the similarity coefficient is calculated as S_{ij} where S_{ij}=the ratio of the number of fragments in common in sample i and j to the number of distinct fragments possessed by i and j. Similarity matrices were calculated from the average binary matrices of TRFLP replicates within samples and of samples within sites. Generally there were three TRFLP replicates for Site 1 and only two for Sites 2 and 3. When this was the case, one replicate from Site 1 was dropped at random. Proximity values were calculated from the R-T coefficient matrices to investigate the

differences in results from different enzymes using Mantel tests in the NTSYSpc module for matrix comparisons (Mantel 1976). In addition, UPGMA dendrograms were generated, using the R-T coefficient matrices, in an attempt to illustrate the similarity relationships among samples within an enzyme.

Haplotype Survey

To corroborate results from this study with previous research in Four Mile Creek, we sought to identify the terminal fragment of the most frequent haplotype identified in RFLP analysis of clones, called Haplotype 1. Sequencher (v4.1) was used to assemble a consensus sequence from sequences of two to six Haplotype 1 clones from five sites and develop a cut map showing the fragment sizes that would result following digestion with *Hae*III or *Hinf*I (Figure 3). The sequences used to draw the cut map did not have the primer sequences on either end, so 20 bp must be added to determine the terminal fragment sizes anticipated for Haplotype 1. The fam end of Haplotype 1 after digestion with *Hae*III should be 123 bp, the hex end, 70 bp. The terminal fragments from *Hinf*I digestion should be 456 bp on the fam end and 236 bp on the hex end.

SECTION 3

RESULTS

Metals

Table 1 shows the concentrations of heavy metals known to be contaminants on the Savannah River Site in each of the samples. No standard error values are available for mercury concentration because mercury concentrations were below detection in at least one sample per site. The instrument detection limit (IDL) of the DMA-80 Mercury Analyzer is 0.01 ng/g total mercury or 0.01 ppb. Samples collected from the same stream in 1998 showed higher levels of mercury, from <10 (IDL) to 228 ppb measured by a different method. In 1998, the site described as Site 1 in this study had a mercury concentration of 84.5 ppb, greater than the concentration of <10 ppb at Site 3. Data for Site 2 in 1998 is not available, however mercury at the next closest site, approximately 200 m downstream, was at 35.3 ppb.

Figures 4, 5 and 6 show the differences in concentrations of different metals in different years. Overall, the change in concentrations of heavy metals between 1998 and 2002 was usually a decrease at Site 1 and an increase in concentration at sites downstream. For concentrations of Cd, the opposite was true, for Ni, concentrations increased at all sites and for Mn, concentrations decreased at Sites 1 and 3 and increased at Site 2 as compared to a 1998 sample site approximately 200m downstream. The spatial pattern of metal concentrations in 1998 generally decreased with distance downstream, except for Cr and Mn. In 2002, most metal concentrations, except Cr, Cu and Zn, showed a peak at Site 2. Cr, Cu and Zn all decreased with distance downstream.

Site Profiles

A profile was generated for each site mapping TRFLP haplotypes as the proportion of total fluorescence of a given fragment size for either the hex or fam end of the *Hae*III or *Hinf*I

enzymes (Figures 7, 8, 9 and 10). Statistics calculated for each site as averages across replicates and samples are presented in Table 2. Different colored bands in the profile figures represent different samples within a site. The error bars of the samples are calculated from TRFLP gel replicates. Both ends of the *Hae*III digest seem to give a similar profile showing both fragments common to all three sites and some unique fragments within each site. *Hinf*I digest profiles are more different among sites and between opposite primer ends.

Diversity

Mercury Resistance Genes

Initial analyses were done for the fam end terminal fragment. It can be observed that results for nearly every enzyme indicate a different pattern of diversity among sites (Figure 11). As calculated from *Bsr*I and *Sau96*I digests, diversity decreased from Site 1 to Site 2 and changed little between Site 2 and Site 3. Diversity changed negligibly among all sites as determined from *Rsa*I digests. However, diversity increased between Sites 1 and 2 and decreased drastically at Site 3 for *Hinf*I digests. *Hae*III digests showed the opposite pattern among sites, decreasing between Sites 1 and 2 and increasing drastically at Site 3.

Error bars for Site 1 *Rsa*I values are based on two, rather than three, samples from Site 1. *Rsa*I data was included to represent an additional enzyme with A's and T's in the cutting sites, in case the phenomenon of differences between enzymes was the result of AT- versus GC-rich cutting sites. *Rsa*I estimates of diversity do not follow the same pattern across samples sites as those of another AT-targeting enzyme, *Hinf*I.

In looking closer at the two enzymes demonstrating the most noticeably different results, box-and-whisker plots were generated for *Hae*III and *Hinf*I estimations of diversity (Figures 12 and 13). These figures illustrate the distribution of values within a sample site. The whiskers and box designate the range and quartiles of the data, while the symbols represent replicates of TRFLP gels within a sample represented by the symbol. Although some TRFLP gel replicates seem to group, there are some distribution-skewing outliers.

We also calculated the Simpson diversity index for the enzymes *Hae*III and *Hinf*I using the hex ends (Figure 14). For the fam ends of the *Hinf*I digest, diversity increases from Site 1 to Site 2 while the hex ends demonstrate a slight decrease from Site 1 to Site 2. Both ends show a decrease from Site 2 to Site 3 in the *Hinf*I digest. However, an analysis of only the hex-labeled fragment would indicate no diversity at Site 3 whereas the fam-labeled fragment reveals a moderate level of diversity at this site. *Hae*III hex ends show an increase in diversity among sites from Site 1 to Site 3. On the other end, *Hae*III fam-labeled fragments, diversity decreases from Site 1 to Site 2, then increases from Site 2 to Site 3. In some samples from Site 2, diversity obtained using the fam end of the amplicon was zero for the enzyme *Hae*III. Using the hex end gave a positive value for diversity in all samples.

A GLM was used to test the significance of differences among sites and enzymes, and between ends for the enzymes *Hae*III and *Hinf*I. In *Hae*III fam ends, although Sites 1 and 2 are not significantly different from one another, both are significantly different from Site 3 ($p<0.05$). From *Hinf*I fam ends, Site 2 differed significantly from Sites 1 and 3 ($p<0.05$), which did not differ from one another. As for differences between fam ends of *Hae*III and *Hinf*I, only those for Site 3 were significant ($p<0.05$). When looking at the hex ends of *Hae*III and *Hinf*I, they also indicate significantly different diversity values for Site 3 ($p<0.05$). Within an enzyme, *Hae*III, there are no significant differences among sites. Sites 1 and 2 were significantly ($p<0.05$) different from Site 3, but not from each other as determined from analyzing the hex ends of *Hinf*I. In addition, for Sites 1 and 3, alternate ends of the *Hae*III digest are significantly ($p<0.05$) different from each other. When comparing opposite ends of the *Hinf*I digest, only those from Site 2 gave significantly different results.

Because the diversity indices calculated from opposite terminal fragments were different, we made a brief examination to determine if a linear relationship might exist between the ends. A scatter plot with xy pairs of fam and hex ends does not indicate a relationship between diversities

calculated from opposite ends of the amplicons from *Hae*III and *Hinf*I digests (Figure 15). The CORR procedure in SAS (v.8.0) was used to determine the r and p values.

16S Ribosomal Genes

Diversity in 16S rDNA did not differ much between sites for most enzymes and both terminal fragments (Figures 16 and 17). Diversities for Site 2 were calculated from only one sample due to failed PCR. Although it may appear that diversity drops at Site 2 for the *Hinf*I hex end, the accuracy of zero diversity at this data point should be questioned. Error bars for Sites 1 and 3 in these figures represent a single standard error of three samples per site and many of the values overlap. The fact that most values are within one standard error of one another, it is not possible to consider them different. This indicates that there are likely no significant differences in 16S ribosomal gene diversity among our sample sites.

Similarity

The similarity analysis of *mer* gene TRFLP profiles among sites also revealed an inconsistency between enzymes for fam ends (Figure 18). R-T coefficient matrices were compared among enzymes to calculate proximity values. No significant similarities were found between R-T coefficient matrices from different enzymes calculated by samples or by replicates. When comparing matrices for alternate ends, *Hae*III fam and hex ends had a significant positive relationship ($p<0.05$) when comparing matrices generated across replicates. The ends from both enzymes had significant ($p<0.05$) positive relationships when comparing matrices across samples only. This indicates that fam and hex ends should show similar relationships among samples within an enzyme when determined from matrices based on the averages of replicates.

Haplotype Survey

Using the cut map (Figure 3) developed from consensus sequences of the most common haplotype found in analysis of clone libraries obtained previously from these study sites, an attempt was made to identify RFLP Haplotype 1 in TRFLP profiles. The expected fragment sizes from *Hae*III digestion were 123 bp and 70 bp for fam and hex ends respectively. Referring to

Figures 7 and 8 shows an absence of these fragments at all three sites. The most frequent *Hae*III hex end fragment was 193 bp in all three samples from all three sites. The most abundant *Hae*III fam end fragment varied by site.

The fragment sizes expected to be most abundant in *Hinf*I digests were a 456 bp fam end and a 236 bp hex end. Again, fragments of these sizes were not readily observed in any sample. However, fragments of very close (3-6 bp difference) sizes were found at least once at all sites in fam end profiles and hex end fragments of 235 bp, 1 bp less than the anticipated size, were found in all three samples from Site 1.

SECTION 4

DISCUSSION

Our first objective in this study was to determine the relationship, if any, between *mer* gene diversity and concentrations of mercury. Mercury concentrations were generally below detection. Therefore it can be posited that just because mercury genes were present it doesn't mean they are active. So, why are they present? It is possible that mercury resistance genes present now are the result of mercury concentrations in previous years, which were, as observed, higher. Also, higher mercury concentrations upstream may have selected for mercury resistant bacteria that have been transported via stream flow to sites with lower levels of mercury contamination. There are other potential mechanisms of mercury resistance that have not been considered due to the relative dearth of information available about the genes that code for them (Osborn et al. 1997). One note of interest is that the concentration of DOC and the level of pH in the stream may affect the path of mercury detoxification. Higher DOC and a lower pH result in a skew toward production of methylmercury rather than mercury reduction (RTI 1999). These parameters may be of interest in future studies. At low mercury levels, mercury often occurs in the form of methyl-mercury, which was not quantified separately from inorganic mercury. In the presence of methyl-mercury, the operon may contain the gene *mer*B, which codes for organomercurial lyase. An investigation for the presence of this gene might prove interesting. The regulatory and transport functions of the genes examined should still be part of the operon function in the presence of methyl-mercury (Osborn et al. 1997). There is also the possibility that the *mer* genes are under selection for alternate functions. For example, the proteins designed to trap and transport mercury may serve to mediate other, similar contaminants that may occur at higher concentrations than mercury at these sites. In addition, *mer* genes may be present at these

sites as a result of past selection. It was noted in the results that mercury occurred at these sites in higher concentrations in previous years.

Although there are notable limitations of this study, it can be said there is worthwhile information to be interpreted from our results. No study is able to account for all variables. It is possible to create more rigorous tests of our conclusions, but it is not possible to know how much we don't know.

Our sample size is very small, only three replicates at three sites. This is, unfortunately, often the case in microbial ecology, due in part to the difficulty in analyzing the large amount of information to be obtained from a small sampling effort. Even if a small sample size would be accurate to capture large differences as a result of environmental conditions, if variance is introduced from sample analysis methods, more samples may be required to provide an accurate representation of the effects of the environment. However, with the advancement of techniques like the one used here, it should be possible to process more samples.

Sources of variance in our sample analysis methods could include PCR artifacts. An inherent bias in PCR has become an underlying assumption in analysis of microbial communities. Primers will anneal more readily to sequences that are more frequent or that have a closer match with the primer sequence. Even sequencing every entry in a clone library can only capture those molecules that were amplified from the original sample. In their review, Wintzingerode et al. (1997) attempted to enumerate sources of PCR bias and present possible was to account for them in analysis. Also, our results are only applicable to the *mer* genes that our primers amplified (Bruce et al. 1995; Liebert et al. 1997), which although generally conserved, are designed from a subset of *mer* genes i.e., those occurring in Tn501, Tn21 and pMER419 (Bruce et al. 1995).

Another aspect of uncertainty that may be introduced by sample analysis is differences between replicate TRFLP gels. In a paper describing methods to optimize TRFLP as a tool for analysis of microbial communities, Osborn (Osborn et al. 2000) found replicate TRFLP gels to be generally consistent. In general, only the smallest peaks were affected by inter-gel variations,

often lost to differences in the volume or concentration of sample loaded (Osborn et al. 2000). It is uncertain where the differences between replicate TRFLP gels arise in this study however, time between replicate runs may have been a factor. Each gel was run from a single enzyme digest reaction. Although care was taken to keep the digested samples out of light and in ultracold conditions between replicate gels, the fluorescent molecules may have faded. Changes in fluorescence occurring during sample storage would alter the appearance of a fluorescence-based profile. Again, in this case, the smallest peaks would be lost from analysis, resulting in underestimations of diversity.

There are recognized shortcomings in TRFLP analysis without the addition of operator error. It is possible that different sequences will have the same terminal fragment size. Sequence differences that occur anywhere on the amplicon other than at the last cutting site before the labeled primer will not be observed in TRFLP analysis. In such a case, diversity would be underestimated. Also, there is no way to recover fragments for sequence analysis to eliminate the possibility of mismaplifications, possibly resulting in an overestimation of diversity.

One especially interesting result we obtained were differences in diversity of *mer* genes estimated using different restriction enzymes or terminal fragments from different ends. The lack of differences between enzymes for estimation of 16S ribosomal gene diversity is probably the result of the high diversity of these genes in environmental samples. It is likely that there are cutting sites aplenty for any enzyme chosen to analyze these genes. As a result of high within site diversity, there is a limitation on the ability to distinguish differences between sites. A TRFLP profile may not provide enough data to capture subtle differences between similar, diverse samples sites. Osborn et al. (2000) noted that terminal fragments from the 5' end of the 16S rDNA amplicon should reveal more diversity than those from the more conserved 3' end of the amplicon. If restriction enzyme cutting sites are conserved among eubacteria at the 3' end of the 16S ribosomal gene, taxa or diversity may be underestimated when looking only at the 3' end.

Most likely, 16S rDNA gene diversity is so high at these sites that any enzyme or end used for analysis would demonstrate similar answers.

If cutting sites in a given amplicon may be randomly distributed, meaning that opposite ends may be considered independent samples of genetic variation within the system under study. This would increase the sample size of the genetic diversity in the community. The assumption of independent cutting sites in an RFLP profile is fundamental to the utility of restriction enzymes to map genetic variance (Nei and Li 1979). Multiple enzyme digests are more likely to violate this assumption.

Despite different diversity estimates obtained using different restriction enzymes, TRFLP may be a useful tool for illuminating differences in the diversity of functional genes from environmental samples. The results from restriction enzymes will depend on the rates of mutations within the enzyme target sequences. Sequences that are more conserved will demonstrate less diversity than those with high mutation rates. Digests run with multiple enzymes and caution used when comparing results from different enzymes may enhance the accuracy of conclusions drawn from TRFLP data. Differences in diversity estimates depending on which restriction enzymes are used and which terminal fragment is analyzed indicate the need for careful consideration in selecting genes, primers and enzymes to be used for TRFLP of environmental samples. We hope it is understood that the results described here pertain to a specific segment of a specific operon and may not be consistent when working with other genes.

As terminal fragments of the 16S gene have been used to identify taxonomic groups in a community, separation of terminal fragments of the *mer* genes into groups of Tn*21*-like and Tn*501*-like may be interesting (Bruce 1997). The major difference between these two groups is the presence of the *merC* gene in operons of the Tn*21* family. The *merC* gene is generally absent in Tn*501*-like operons (Liebert et al. 1997). In addition, these two groups may indicate two evolutionarily significant lineages of the *mer* operon (Osborn et al. 1997). Of course, any attempt to identify taxa or other group designations will be limited by the number of sequences available

for fragment size comparison. Our restriction enzymes were not chosen to be diagnostic for sequence differences in subclasses of *mer* operons.

Traditionally, analysis of environmental microbial communities was limited to those individuals of the community that are culturable. With the advent of molecular techniques, analyses could include non-culturable members of the community as well. The emerging methodologies attempt to capture a broad snapshot of microbial communities, but concurrently must sacrifice image details. As demonstrated by previous work in our lab, clones identified as the same RFLP haplotype may have different sequences. TRFLP provides even less data per individual. A single peak may represent multiple individuals. The robustness of new molecular methods still must be rigorously tested before conclusions can be drawn from the results. Some sequencing may always be necessary to determine if an analysis is sufficiently fine-toothed to answer the question at hand. It is anticipated that TRFLP can be a useful tool in microbial ecology, as long as its limitations are understood. Comparisons of TRFLP profiles to determine patterns diversity among sampling sites or between studies should take into consideration variations revealed by comparing results from different restriction enzymes and alternate ends of the target amplicon. It is especially important that studies done using TRFLP include details of enzyme digest reactions.

TRFLP may be suitable as a reconnaissance method. If it can be used to obtain an idea of the diversity in a gene of interest, it should provide an idea of how many clones of a library must be sequenced for later, more detailed analyses. In addition, in a community that is already highly characterized, TRFLP could be a way to compare communities under different experimental treatments or over time within a treatment.

Additional questions unanswered by this study include whether the diversity of haplotypes occurring at the confluence is the result of genetic exchange, contaminant selection or the presence of bacteria native to Castor Creek. Determining the locations of the genes within the bacteria (i.e., on conjugative plasmids) could help provide an answer. If the phenomenon is the

result of bacterial response to contaminant selection, further analysis could be of evolutionary significance. Finally, correlations between unique haplotypes and contamination conditions may indicate the potential utility of bacteria as biomarkers of contamination. Further study should include sampling and sequencing along Castor Creek.

SECTION 5

LITERATURE CITED

Alonso, A., P. Sanchez, J.L. Marinez. (2001). "Environmental selection of antibiotic resistance genes." Environmental Microbiology **3**(1): 1-9.

Aviles, M., J. C. Codina, A. Perez-Gracia, F. Cazorla, P. Romero, A.de Vicente. (1993). "Occurrence of resistance to antibiotics and metals and of plasmids in bacterial strains isolated from marine environments." Water Science and Technology **27**(3-4): 475-478.

Barkay, T., C. Liebert, M. Gillman. (1989). "Environmental significance of the potential for *mer*(Tn*21*)-mediated reduction of Hg^{2+} to Hg^{0} in natural waters." Applied and Environmental Microbiology **55**(5): 1196-1202.

Blackwood, C. B., T. Marsh, S.-H. Kim, E.A. Paul. (2003). "Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities." Applied and Environmental Microbiology **69**(2): 926-932.

Braker, G., H. J. Ayala-del-Rio, A.H. Devol, A. Fesefeldt, J.M Tiedje. (2001). "Community structure of denitrifiers, *Bacteria* and *Archaea* along redox gradients in Pacific Northwest marine sediements by terminal restriction fragment length polymorphism analysis of amplified nitrite reductase (*nirS*) and 16rRNA genes." Applied and Environmental Microbiology **67**(4): 1893-1901.

Bruce, K. D. (1997). "Analysis of *mer* gene subclasses within bacterial communities in soils and sediments resolved by fluorescent-PCR-restriction fragment length polymorphism profiling." Applied and Environmental Microbiology **63**(12): 4914-4919.

Bruce, K. D., A. M. Osborn, A.J. Pearson, P. Strike, D.A. Ritchie. (1995). "Genetic diversity within *mer* genes directly amplified from communities of noncultivated soil and sediment bacteria." Molecular Ecology **4**: 605-612.

Dunbar, J., L.O. Ticknor, C.R. Kuske. (2000). "Assessment of microbial diversity in four southwestern United States soils by 16S rRNA gene terminal fragment analysis." Applied and Environmental Microbiology **66**(7): 2943-2950.

Liebert, C., R. M. Hall, A.O. Summers. (1999). "Transposon Tn*21*, flagship of the floating genome." Microbiology and Molecular Biology Reviews **63**(3): 507-522.

Liebert, C., J. Wireman, T. Smith, A.O. Summers. (1997). "Phylogeny of mercury resistance (*mer*) operons of gram-negative bacteria isolated from the fecal flora of primates." Applied and Environmental Microbiology **63**(3): 1066-1076.

Liu, W.-T., T. L. Marsh, H. Cheng, L.J. Forney. (1997). "Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA." Applied and Environmental Microbiology **63**(11): 4516-4522.

Mantel, N. (1967). "The detection of disease clustering and a generalized regression approach." Cancer Research **27**:209-220

McArthur, J. V. (1998). "Physiologic and genetic determinants of lotic bacterial distribution and abundance." Proceedings of the International Association for Theoretical and Applied Limnology **26**: 1080-1082.

McArthur, J. V., L. G. Leff, M.H. Smith. (1992). "Genetic diversity of bacteria along a stream continuum." Journal of the North American Benthological Society **11**(3): 269-277.

McArthur, J. V. and R. C. Tuckfield (2000). "Spatial patterns in antibiotic resistance among stream bacteria: effects of industrial pollution." Applied and Environmental Microbiology **66**(9): 3722-3726.

Moeseneder, M. M., J. M. Arrieta, G. Muyzer, C. Winter, G.J. Herndl. (1999). "Optimization of terminal-restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis." Applied and Environmental Microbiology **65**(8): 3518-3525.

Nazaret, S., W. H. Jeffery, E. Saouter, R. von Haven, T. Barkay. (1994). "*mer*A gene expression in aquatic environments measured by mRNA production and Hg(II) volatilization." Applied and Environmental Microbiology **60**(11): 4059-4065.

Nei, M. and W.-H. Li (1979). "Mathematical model for studying genetic variation in terms of restriction endonucleases." Proceedings of the National Academy of Science **76**(10): 5269-5273.

Osborn, A. M., K. D. Bruce, P. Strike, D.A. Ritchie. (1995). "Sequence conservation between regulatory mercury resistance genes in bacteria from mercury polluted and pristine environments." Systematic and Applied Microbiology **18**: 1-6.

Osborn, A. M., K. D. Bruce, P. Strike, D.A. Ritchie. (1997). "Distribution, diversity and evolution of the bacterial mercury resistance (mer) operon." FEMS Microbiology Reviews **19**: 239-262.

Osborn, A. M., E. R. B. Moore, K. Timmins. (2000). "An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics." Environmental Microbiology **2**(1): 39-50.

Palumbi, S. R. (1996). Nucleic acids II: The polymerase chain reaction. Molecular Systematics. D. M. Hillis, C. Moritz and B. K. Mable. Sunderland, Massachusetts, Sinauer Associates, Inc.**:** 205-247.

Rasmussen, L. D. and S. J. Sorensen (1998). "The effect of longterm exposure to mercury on the bacterial community in marine sediment." Current Microbiology **36**: 291-297.

Research Triangle Institute (RTI). (1999). "Toxicological Profile for Mercury."

Rogers, J. R. and T. T. Tanimoto (1960). "A computer program for classifying plants." Science **132**(3434): 1115-1118.

Sandaa, R.-A., V. Torsvik, O. Enger. (2001). "Influence of long-term heavy-metal contamination on microbial communities in soil." Soil Biology and Biochemistry **33**: 287-295.

Stepanauskas, R., M. A. Moran, B.A. Bergamaschi, J.T. Hollibaugh. (2003). "Covariance of bacterioplankton composition and environmental variables in a temperate delta system." Aquatic Microbial Ecology **31**: 85-98.

Wintzingerode, F. von, U.B. Gobel, E. Stackebrandt. (1997). "Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis." FEMS Microbiology Reviews **21**: 213-229

Wise, M. G., J. V. McArthur, C. Wheat, L.J. Shimkets. (1996). "Temporal variation in genetic diversity and structure of a lotic population of *Burkholderia* (*Pseudomonas*) *cepacia*." Applied and Environmental Microbiology **62**(5): 1558-1562.

Wise, M. G., L. J. Shimkets, JV. McArthur. (1995). "Genetic structure of a lotic population of *Burkholderia* (*Pseudomonas*) *cepacia*." Applied and Environmental Microbiology **61**(5): 1791-1798.

SECTION 6

TABLES

Table 1: Concentrations of heavy metals in ppm (standard error), an average of three samples per site, or ppb for Hg, the highest of three samples per site. Samples were collected from Four Mile Creek on the Savannah River Site, 13.June.2002.

	Site 1	*Site 2*	*Site 3*
Cd	0.31 (0.06)	0.38 (0.04)	0.15 (0.03)
Ni	0.40 (0.06)	0.71 (0.11)	0.25 (0.05)
Pb	0.52 (0.05)	0.73 (0.16)	0.54 (0.07)
Cr	0.10 (0.03)	0.16 (0.03)	0.09 (0.02)
Cu	1.24 (0.13)	1.07 (0.22)	0.63 (0.06)
Zn	14.24 (2.50)	9.35 (2.94)	4.02 (0.42)
Mn	101.97 (23.19)	184.08 (54.41)	14.93 (3.01)
Hg	15.53	22.45	Bd*

*Bd=below instrument detection limit (10 ppb)

Table 2: Statistics associated with TRFLP profiles for *Hae*III and *Hinf*I digests considering both ends of the amplicon. Average number of bands, average total peak height, average fragment peak height for individual fragments and variance of normalized peak heights were calculated using values from replicate TRFLP gels across samples

	Site 1	*Site 2*	*Site 3*
Hae*III *fam ends (Fig. 7)			
Avg. no. of bands	5	3	7
Avg. total peak height	1098.67	301.89	1342.33
Avg. fragment peak height	205.29	171.01	189.09
Normalized variance	0.0363	0.0468	0.0092
Hae*III *hex ends (Fig. 8)			
Avg. no. of bands	2	3	6
Avg. total peak height	301.89	387.33	821.83
Avg. fragment peak height	106.43	165.81	156.40
Normalized variance	0.0437	0.0568	0.0919
Hinf*I *fam ends (Fig. 9)			
Avg. no. of bands	6	5	4
Avg. total peak height	2975.22	593.17	2448.33
Avg. fragment peak height	521.37	132.34	522.54
Normalized variance	0.0775	0.0173	0.1192
Hinf*I *hex ends (Fig. 10)			
Avg. no. of bands	2	2	1
Avg. total peak height	425.56	213.50	172.50
Avg. fragment peak height	169.80	135.03	172.50
Normalized variance	0.0913	0.0270	0

SECTION 7

FIGURES

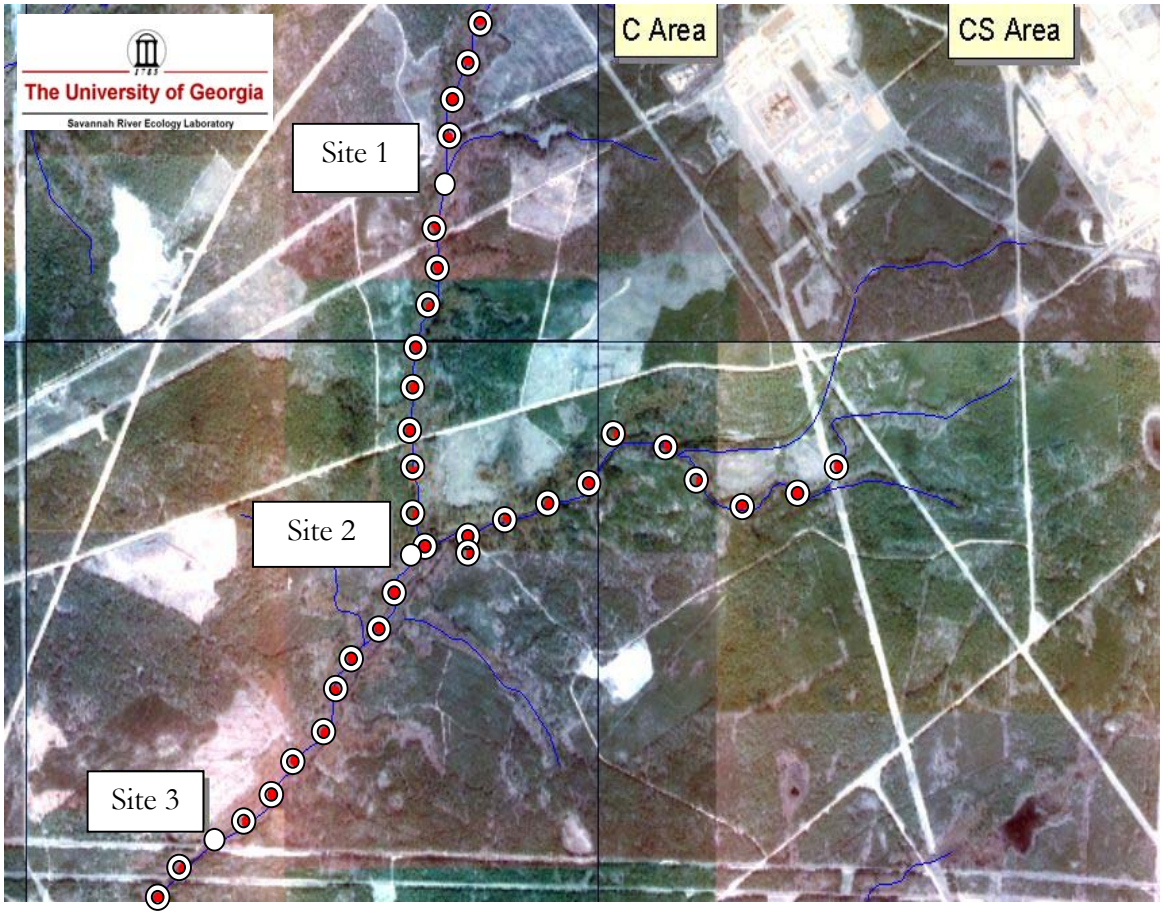

Figure 1: Aerial photograph with overlay depicting GPS designated sample sites on Four Mile Creek and Castor Creek tributary. The closed-circle, labeled locations are those examined in this study.

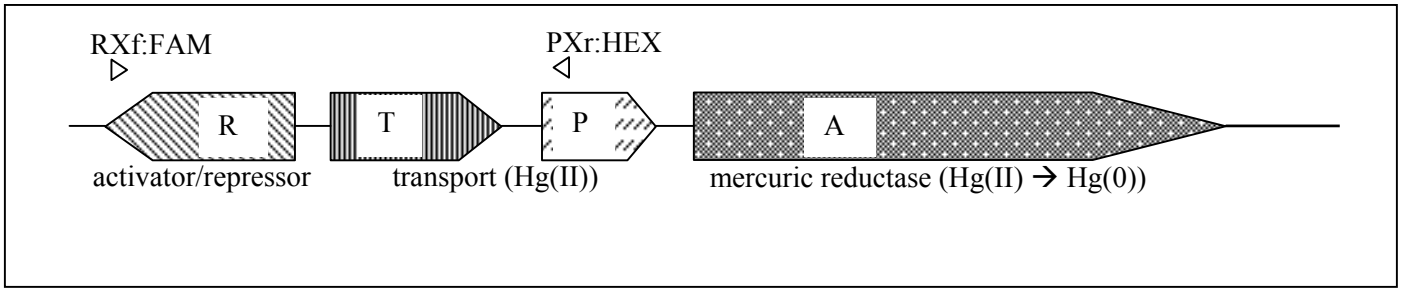

Figure 2: The generally conserved loci of the *mer* operon and the functions or products for which the associated genes code. The small arrows indicate the direction and approximate locations of primers used for PCR and reactions (from Liebert et al 1997).

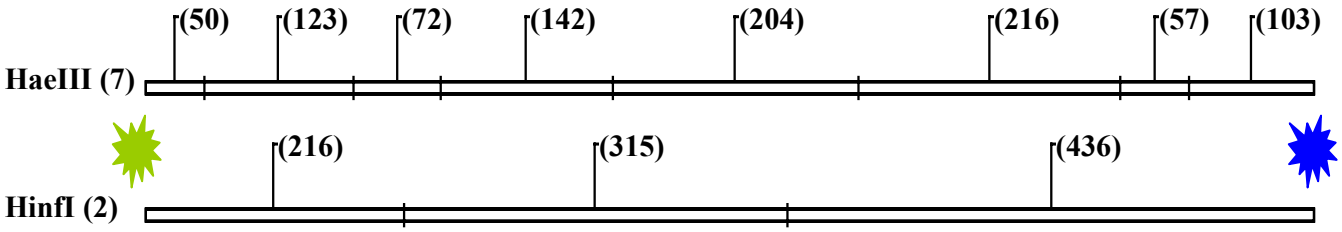

Figure 3: Restiction enzyme cutting site maps for *Hae*III and *Hinf*I from a consensus sequence generated by the assembly of sequences from two to six *mer* clones designated as haplotype 1 from each of five sites along Four Mile Creek. The stars at either end of the amplicon represent the color of the anticipated fragment from either end.

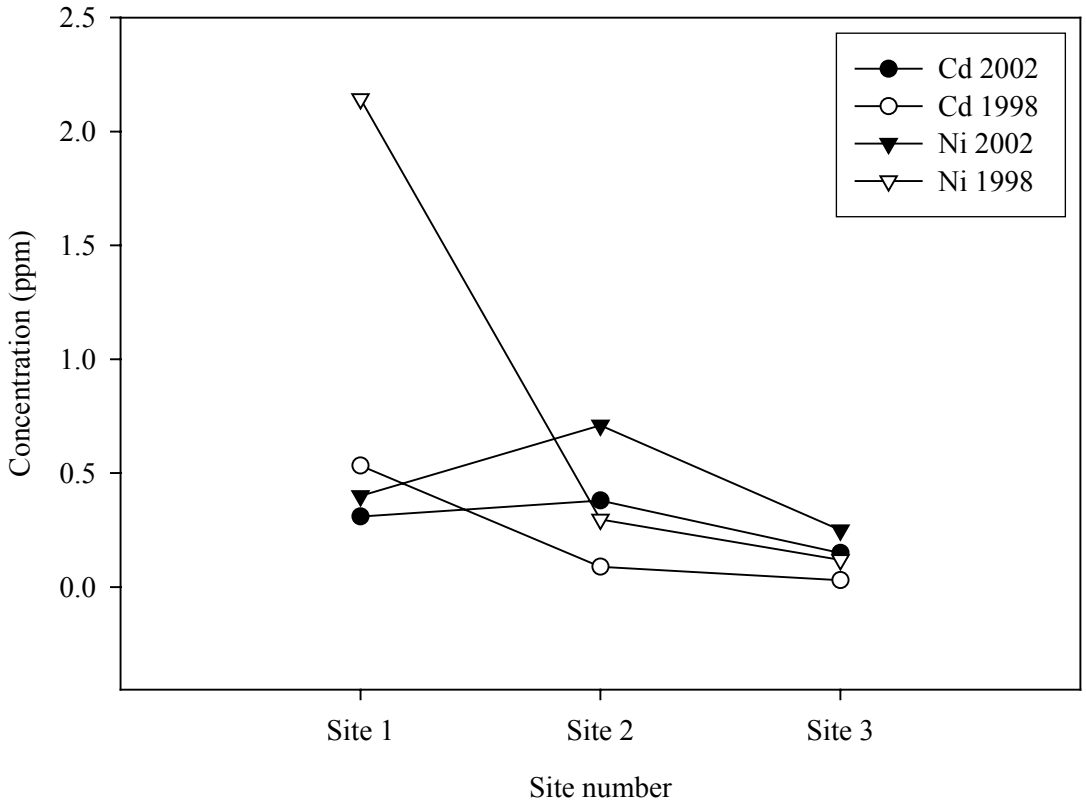

Figure 4: Concentrations of two metals for two years by site. Open symbols represent samples taken in 1998. Closed circles represent samples collected for the present study on 13.June.2002.

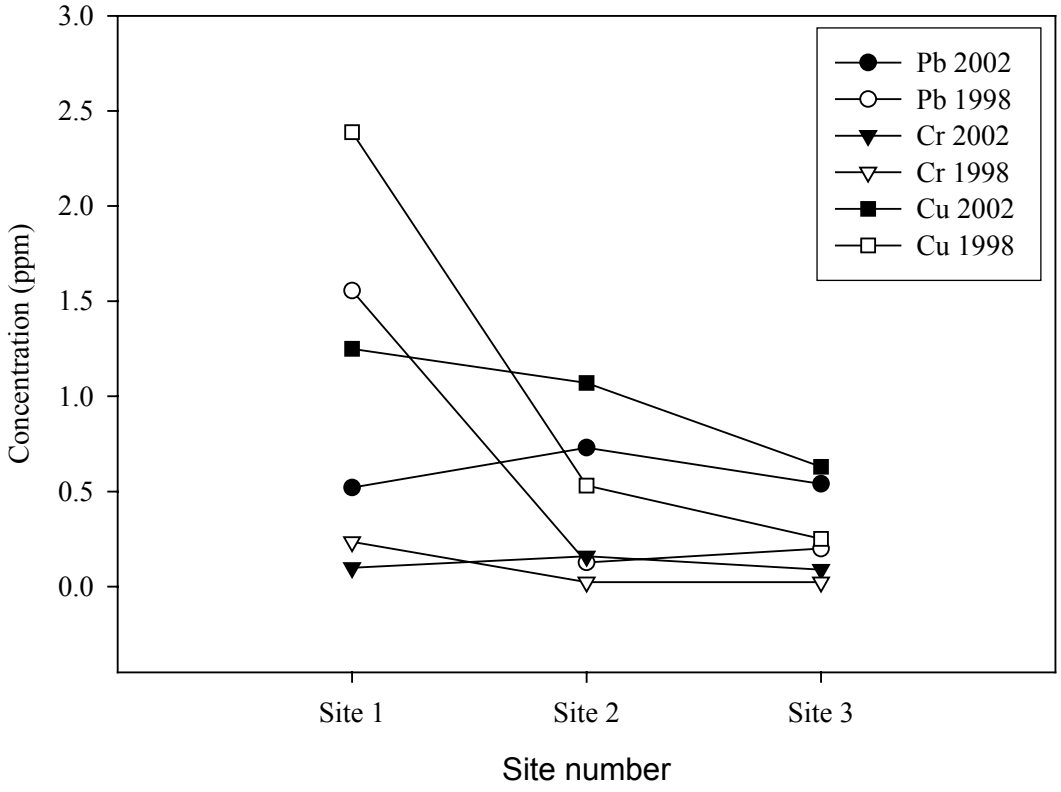

Figure 5: Concentrations of three metals for two years by site. Open symbols represent samples taken in 1998. Closed symbols represent samples collected for the present study on 13.June.2002.

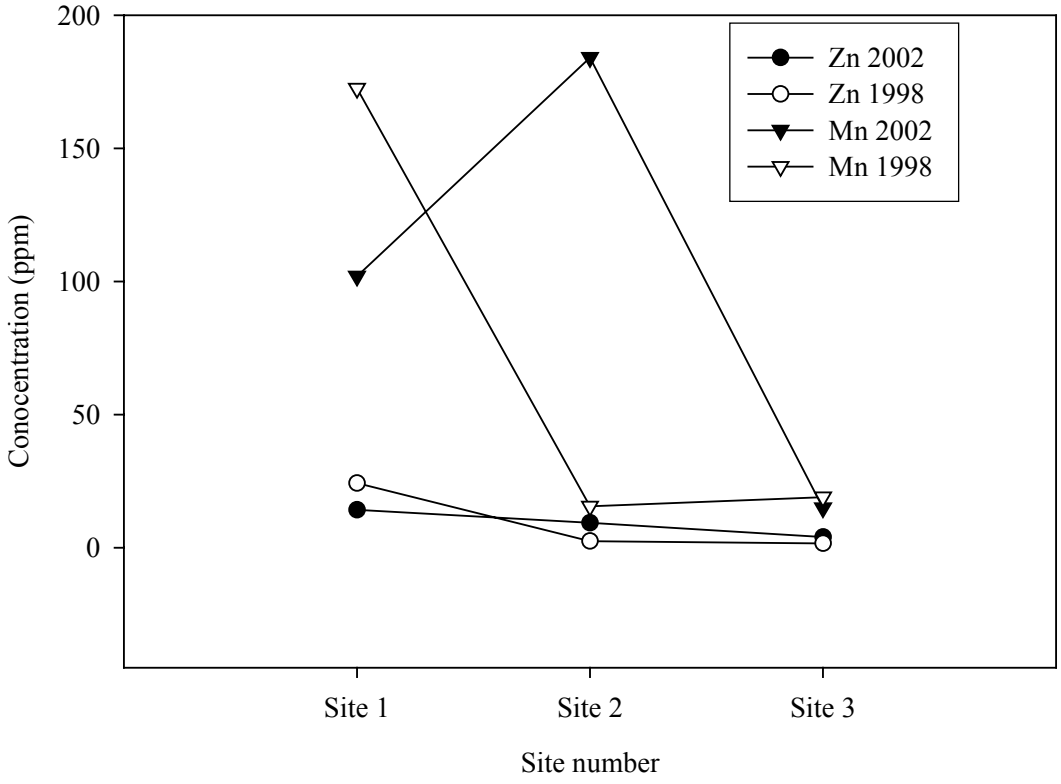

Figure 6: Concentrations of two metals for two years by site. Open symbols represent samples taken in 1998. Closed symbols represent samples collected for the present study on 13.June.2002.

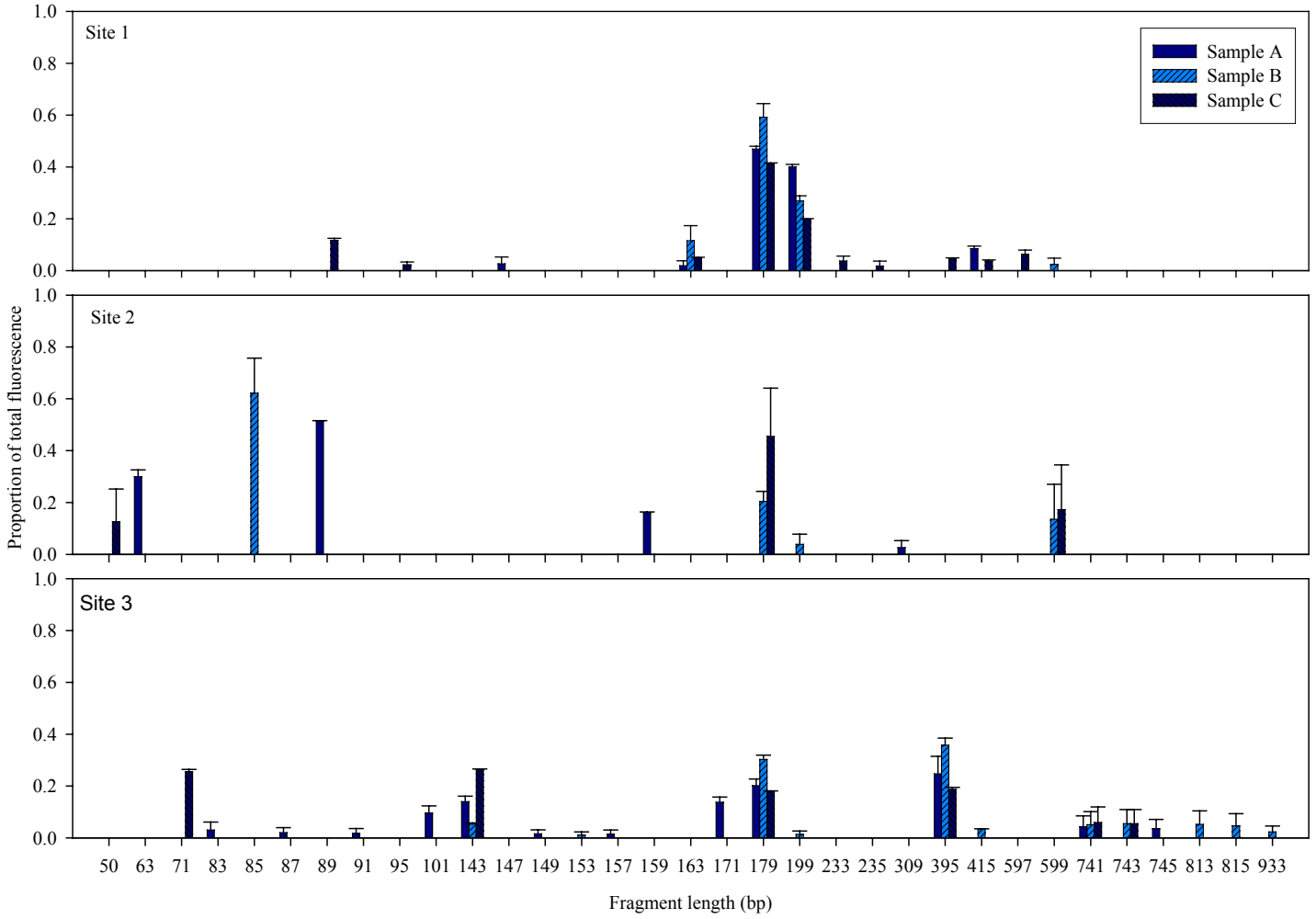

Figure 7: Bar graphs depicting the distribution of fragment size and abundance from *mer* genes within sites from *Hae*III digested fam ends on a per sample basis. Error bars represent the standard error from replicate TRFLP gels

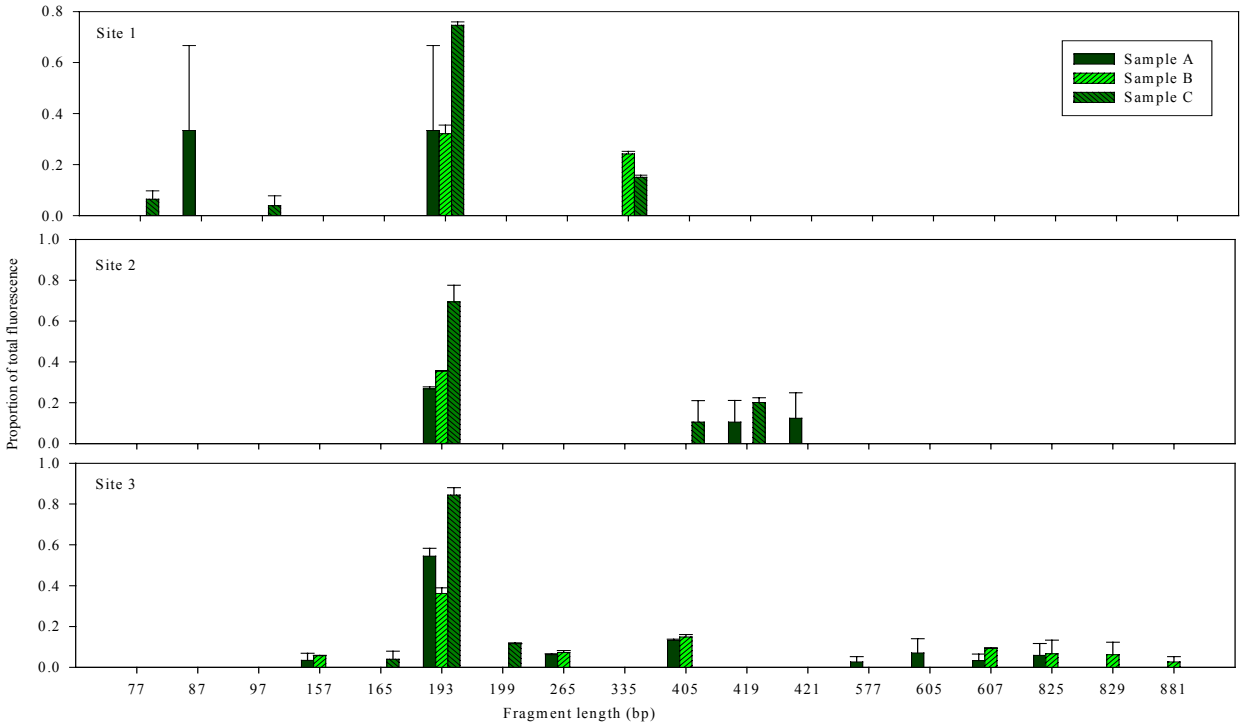

Figure 8: Bar graphs depicting the distribution of fragment size and abundance from *mer* genes within sites of *Hae*III digested hex ends on a per sample basis. Error bars represent the standard error from replicate TRFLP gels

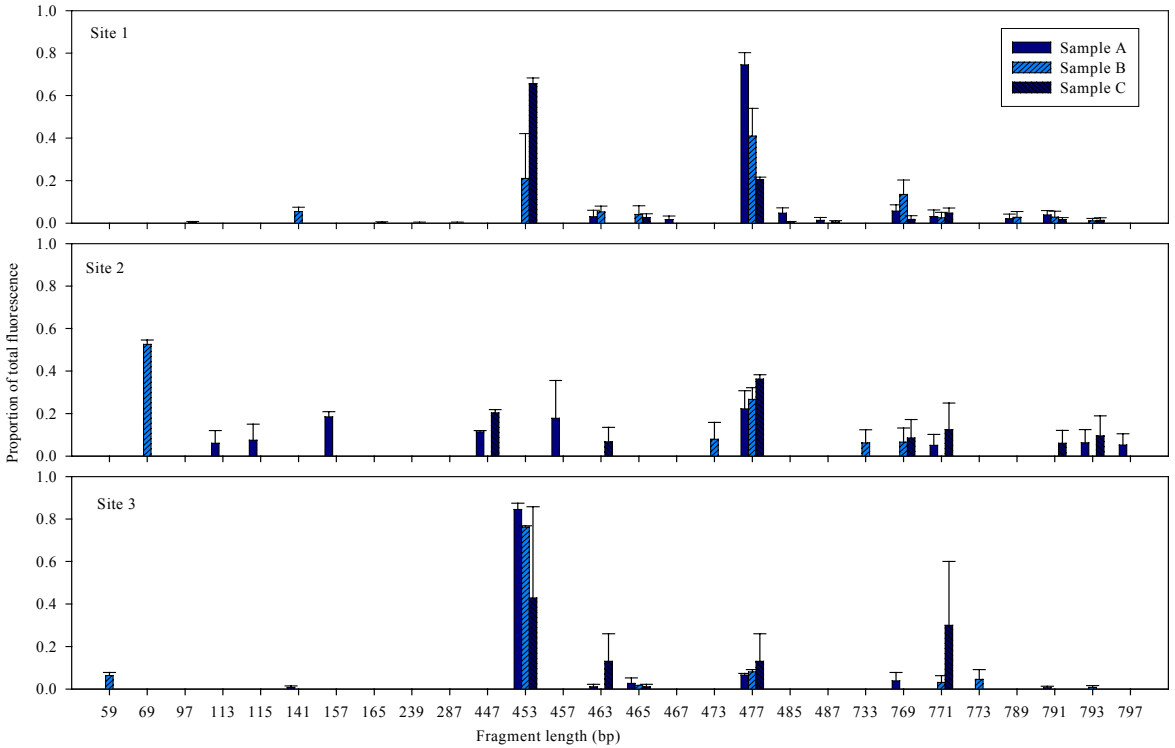

Figure 9: Bar graphs depicting the distribution of fragment size and abundance from *mer* genes within sites of *Hinf*I digested fam ends on a per sample basis. Error bars represent the standard error from replicate TRFLP gels

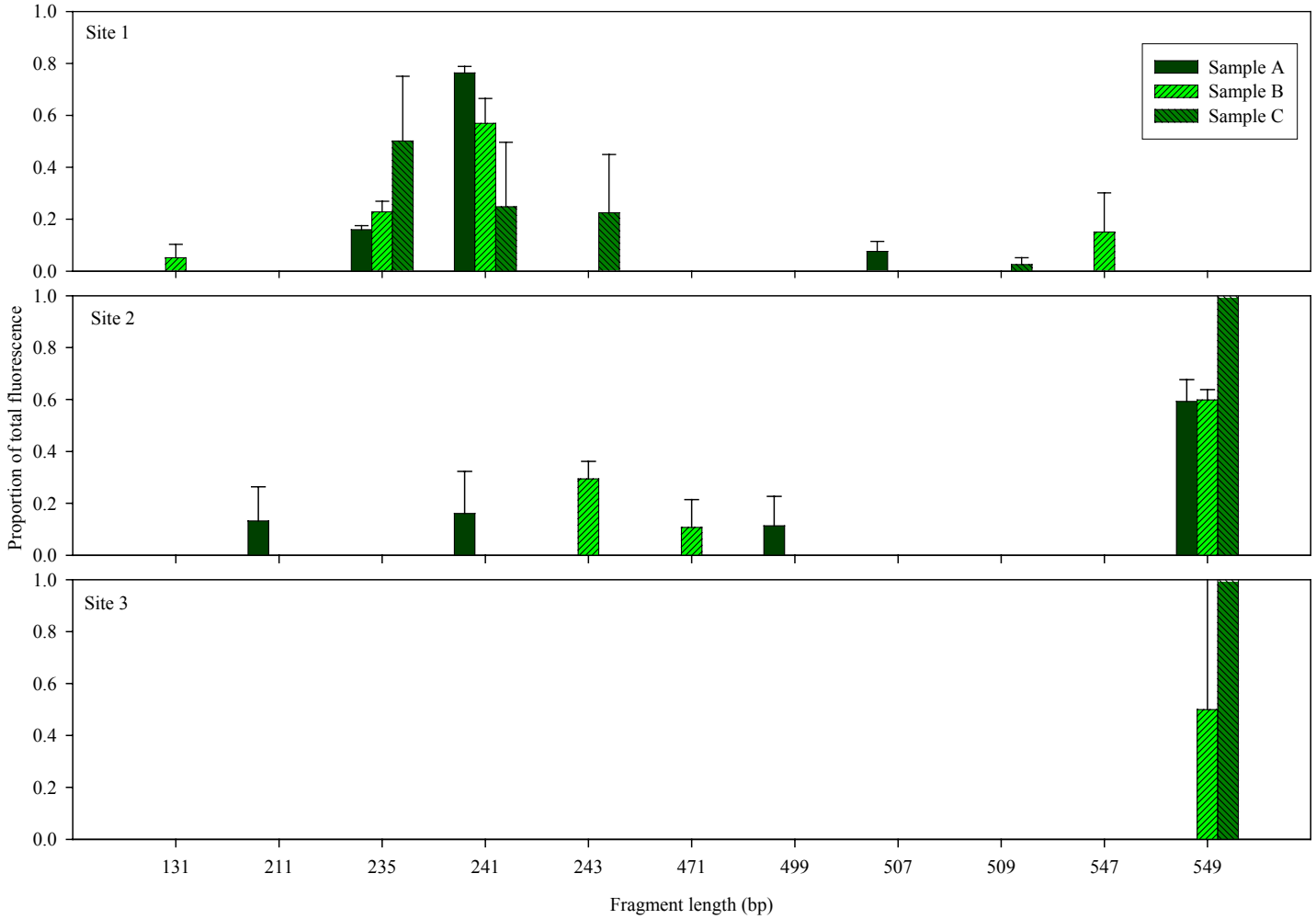

Figure 10: Bar graphs depicting the distribution of fragment size and abundance from *mer* genes within sites of *Hinf*I digested hex ends on a per sample basis. Error bars represent the standard error from replicate TRFLP gels

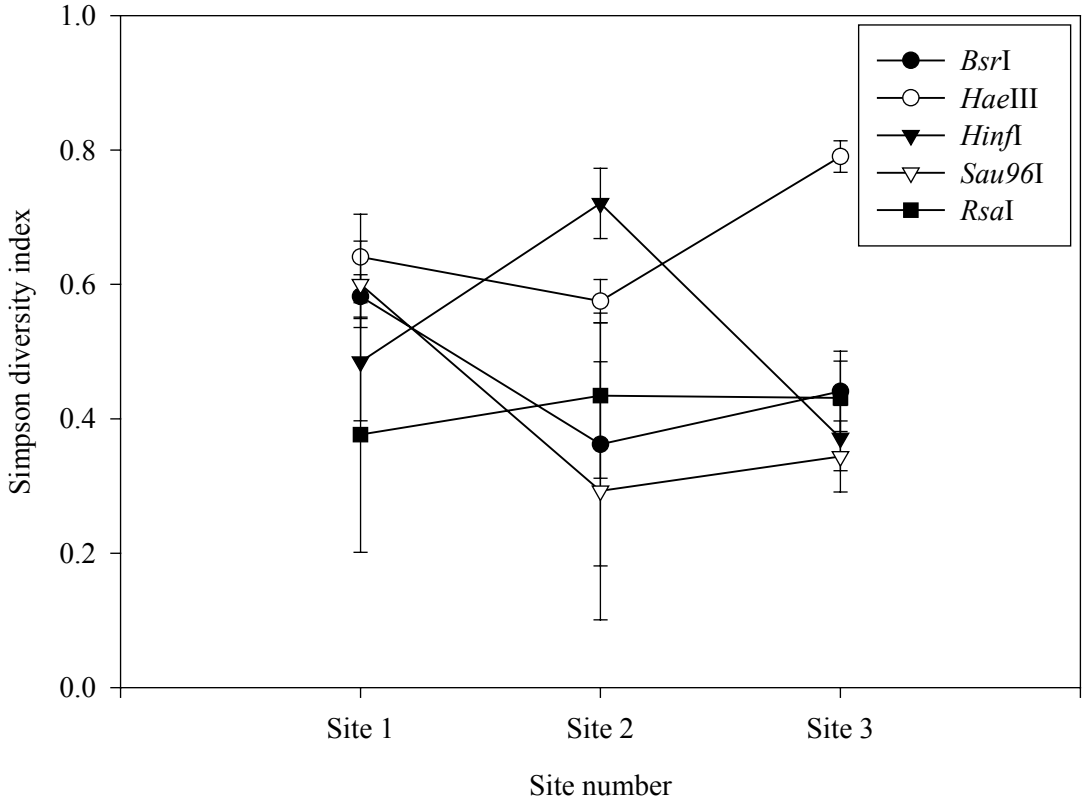

Figure 11: Simpson diversity index of *mer* genes by site for five restriction enzymes. Values are averages of estimates from two or three gels per sample and three samples per site. Error bars are standard error from the averages

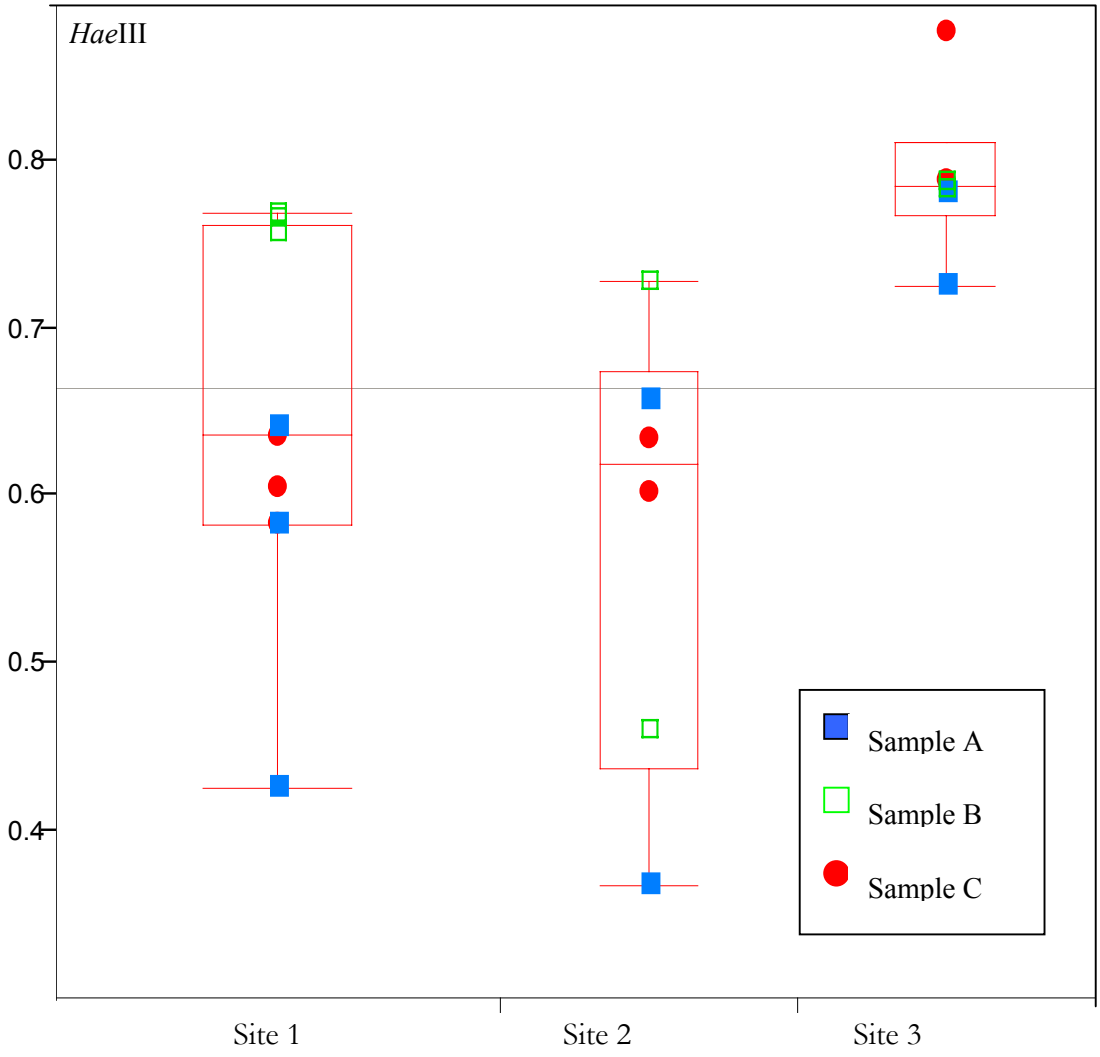

Figure 12: Box and whisker plot to show the distribution of *mer* gene diversity estimates within a site for the enzyme *Hae*III. Each symbol represents a TRFLP replicate. Each type of symbol represents a sample at a given site.

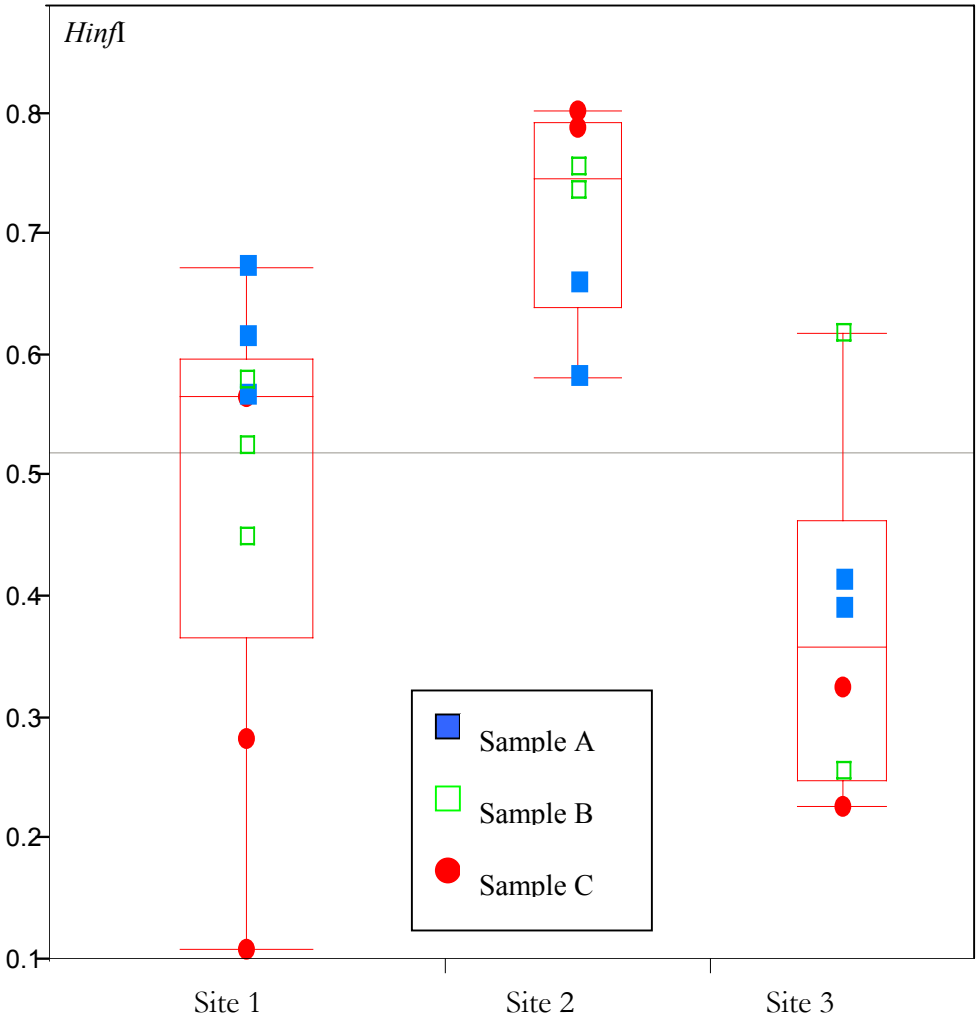

Figure 13: Box and whisker plot to show the distribution of *mer* gene diversity estimates within a site for the enzyme *Hinf*I. Each symbol represents a TRFLP replicate. Each type of symbol represents a sample at a given site.

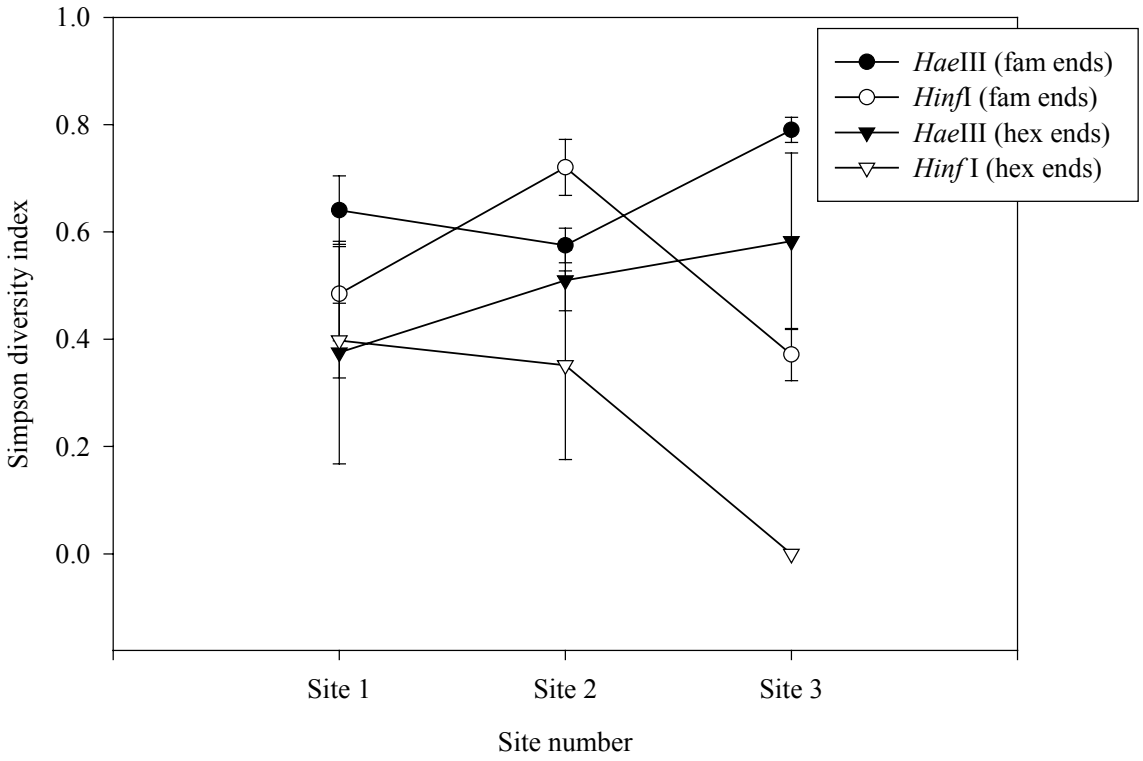

Figure 14: Simpson diversity index of *mer* genes by site for fam and hex ends. Values are averages of two to three gels per sample and three samples per site from enzyme digests using HaeIII and HinfI. Error bars are standard error from the averages.

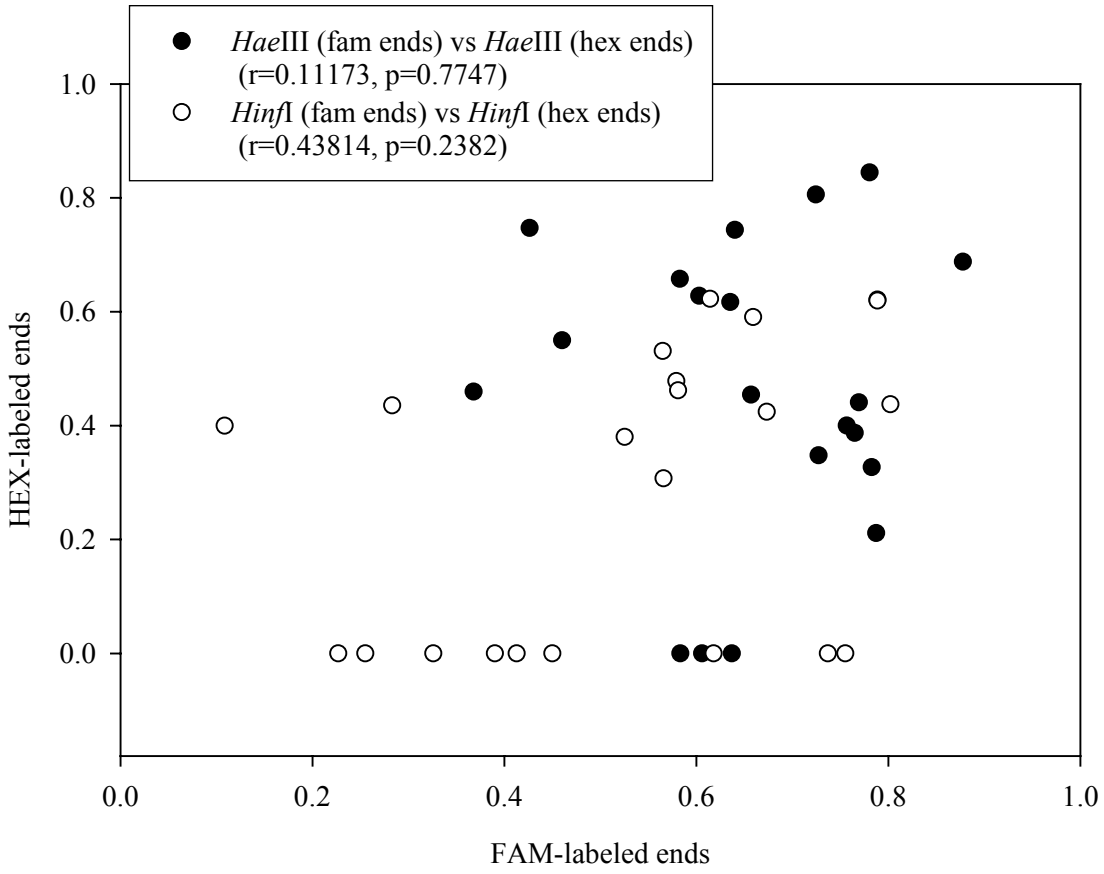

Figure 15: A scatter plot of Simpson diversity index of *mer* genes estimated using alternately labeled terminal fragments from digestions with HaeIII or HinfI as xy pairs. r and p values were calculated using the correlation procedure in SAS v.8

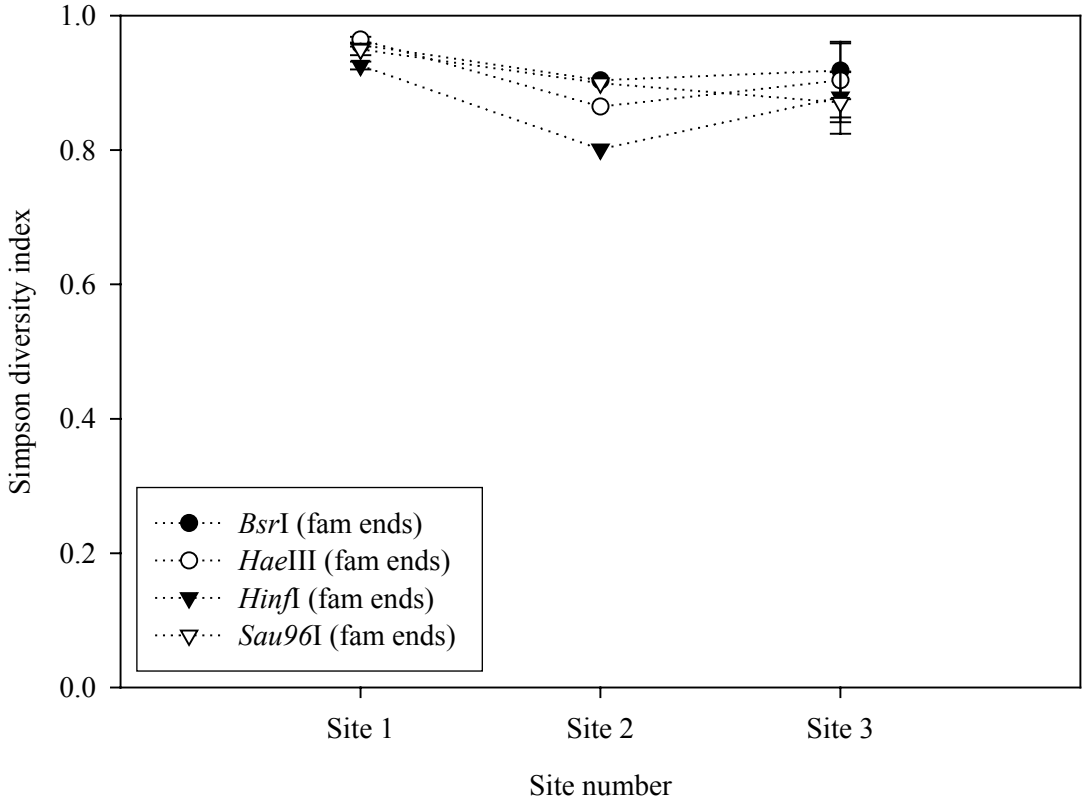

Figure 16: Simpson diversity index by site from fam ends of four restriction enzyme digests of 16S rDNA. Lines between points are dotted to indicate that these relationships are uncertain due to the presence of only one sample replicate for Site 2

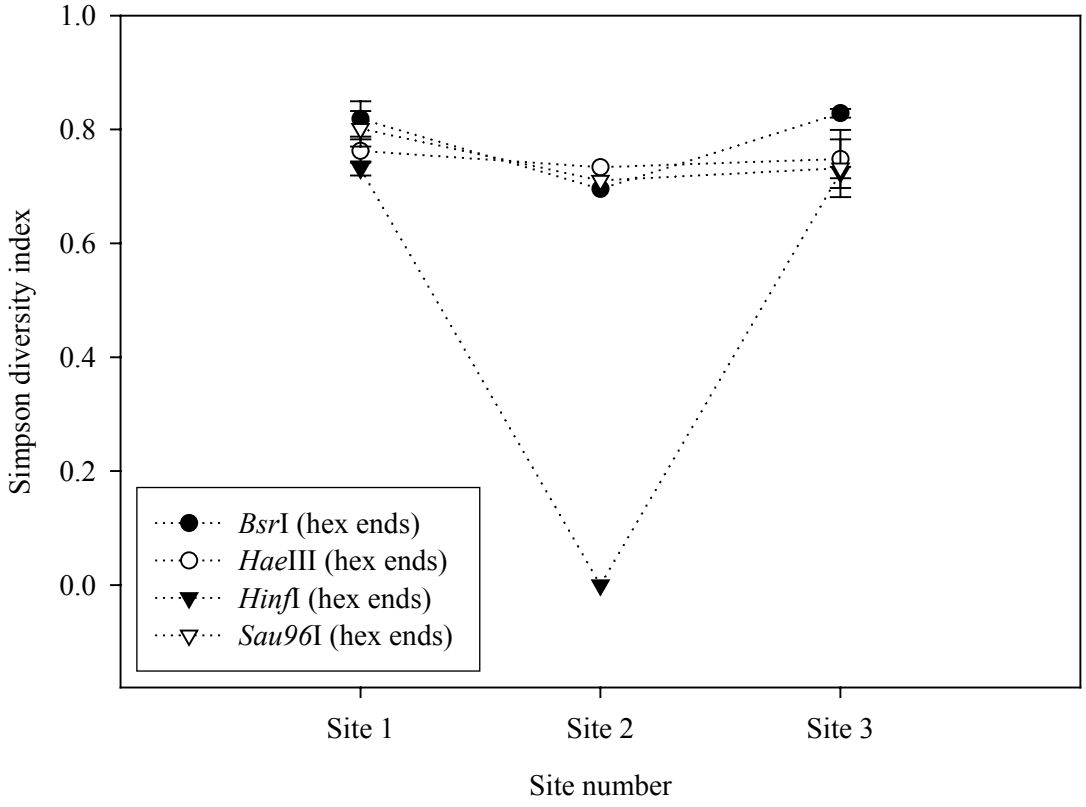

Figure 17: Simpson diversity index by site from hex ends of four restriction enzyme digests of 16S rDNA. Lines between points are dotted to indicate that these relationships are uncertain due to the presence of only one sample replicate for Site 2

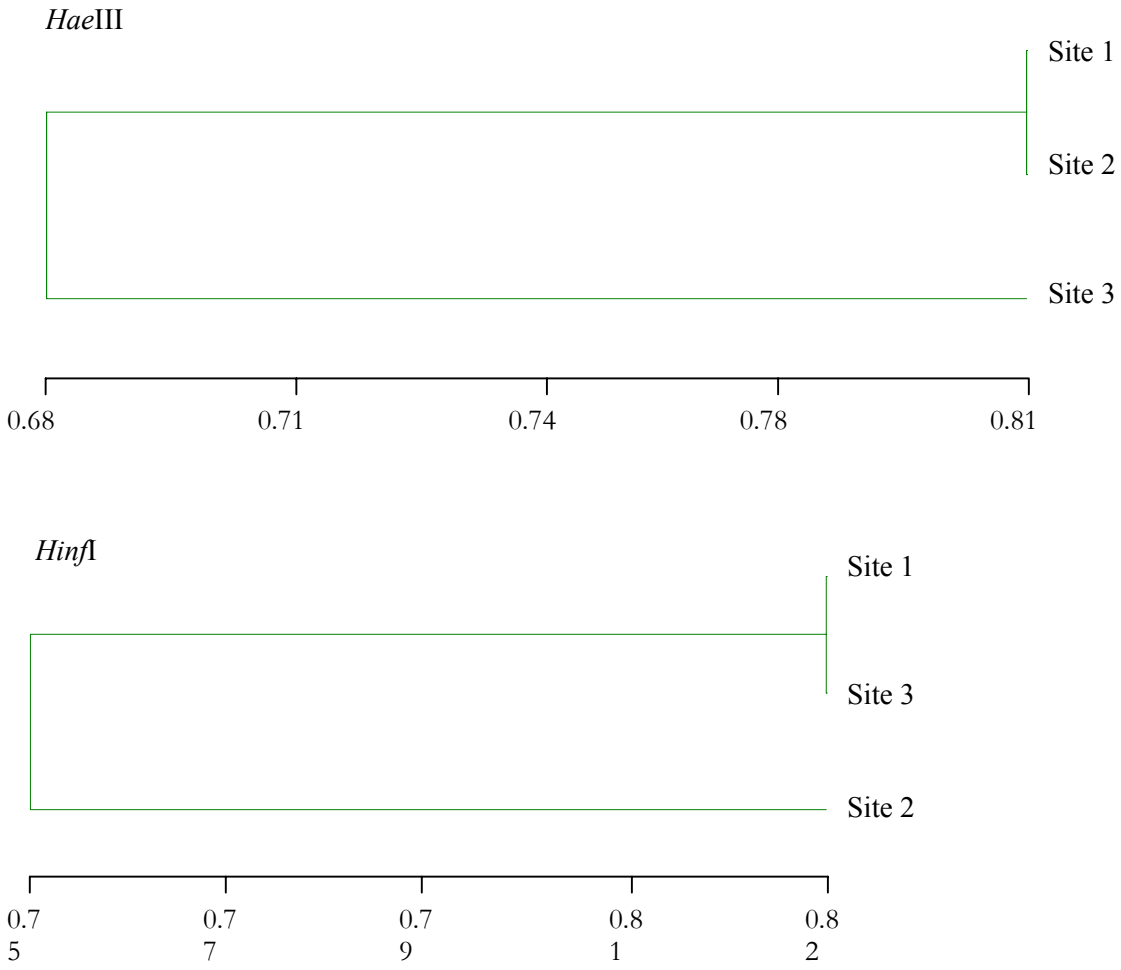

Figure 18: UPGMA trees for *mer* genes digested with two enzymes, *Hae*III and *Hinf*I, based on R-T coefficient of similarity. Scales represent the coefficient distance between sites for each enzyme.