Diversity of Transposons in Mercury Resistant Bacteria

Thesis submitted in accordance with the requirements of the
University of Liverpool for the degree of
Doctor in Philosophy
by

Robert James Holt

June 1999

THE UNIVERSITY
of LIVERPOOL

CONTENTS.

FIGURES.

TABLES.

ABSTRACT.

The diversity of resolvase (*tnpR*) genes carried by a number of mercury resistant soil bacteria has been studied by DNA sequencing. The resulting DNA sequence was used for comparison to previously published restriction fragment length polymorphism (RFLP) data, permitting the relationships between DNA sequencing and RFLP typing approaches to be studied by the use of phylogenetic trees. DNA maximum likelihood and DNA parsimony were used to construct a variety of phylogenetic trees. DNA sequencing confirmed the validity of RFLP analysis and highlighted the importance of restriction endonuclease choice upon the resulting RFLP pattern and dendrogram topology. The *tnpR* genes of two previously uncharacterised mercury resistant bacteria, T2 7 and T2 12 were also studied. DNA sequence data placed T2 7 in a previously described gene class, *tnpR* D and T2 12 in a new gene class, *tnpR* F.

The presence of Class II transposon genes related to Tn*21* and Tn*501* and their structural arrangements have been determined in a collection of 124 mercury resistant Gram-negative bacteria. Seventy-five of the 124 isolates contained a *tnpA* (transposase) gene related to Tn*21* and Tn*501* and in all 64 isolates that contained both *tnpA* genes and plasmids, the *tnpA* gene was plasmid borne. The structural arrangement of the *tnpA* and *tnpR* genes within the isolates was determined. Sixty-nine of the 75 *tnpA* containing isolates had an arrangement of *tnpA* and *tnpR* genes similar to that found in the Tn*21* subgroup of transposons. Four strains did not produce a PCR product using *tnpR* primers. The remaining two isolates had undetermined arrangements of *tnpA* and *tnpR* genes. No Tn*3*-like arrangements of *tnpA* and *tnpR* were seen, despite being detected in DNA extracted directly from the isolation sites.

The relative orientation of the *tnp* genes and *mer* operon was studied and revealed the presence of two distinct structural groups. The *merC* gene was present in 44 isolates. Five isolates were found to carry integrase genes and these contained inserted gene cassettes varying in size from 1.1 kb to 4.5 kb. The significant lack of structural diversity amongst transposon associated genes in these isolates suggests that members of the Tn*21* subgroup of transposons may have a selective advantage in the natural environment. It is also apparent that the diversity of the transposon structures contained within these isolates has not been affected by recombination events which have been previously observed in these isolates.

ACKNOWLEGEMENTS.

I would like to say an enormous thank you to the following people who have supported me in various ways since my arrival in Liverpool. Most importantly, thank you to my two supervisors; Prof. Pete Strike and Dr Ken D. Bruce. This project was funded by the Natural Environment Research Council.

I would like to thank Prof. Don Ritchie for his continued support of my work and for all his useful comments and suggestions during our Friday morning meetings. I would also like Dr Hilary K. Young for her help with the integron study and Dr Andrea Pearson for her work on transposons prior to my arrival in Liverpool. This project used the SEQNET facility at Daresbury Laboratories. I would also like to thank my two internal examiners Dr Alan McCarthy and Prof. John Saunders for all their useful suggestions. All DNA sequencing work was carried out by the two Ang's; Angela Bardon and Angela Rosin.

A big thank you to all my mates in the Donnan Labs. Geoff Elliott: Thank you for sharing your uniquely sarcastic outlook on life with me and for all your help and advice. Dr Mark "Kitman" Hart: Thank you for being the butt of our jokes for so long, now you've left there's only Karl and it's not the same. Cheers for showing me all of your methods, you explained them like a true (Ale House) scientist. Jim Newton: top pop star look-alike and creator of the "Ale House Scientist" badge. Magic: drinking partner and flatmate. I'll get you dancing in the Razz if it's the last thing I do. Polly: Fellow gin drinker and founder member of the Pars pizza gang. Ang: Good luck with your continuing quest to be mentioned in everybodies thesis (twice in this one). Dr Andy Harrison: Thank you for providing me with a unique insight into what my hair will look like in 5 years. The Porn Meisters: Steve (all work and no play) Clapcott and Karl (all play and no work) Hodder. Dav: Thanks for lending us all your equipment over the last 3 years- you can have it back now.

A quick thank you to everyone else who's paths have crossed mine: The Giant and The Little Hawaiian, Blancey, Submarine Commander Matt, The Ever Talkative Charlotte, Dr Mark Osborn and everyone else who reads this that I've forgotten about. Thanks to everyone in The Vernon, Pars and The Razz for providing the essential ingredients for Friday nights.

Mum and Dad: Thank you for all your support, financial and otherwise (including the use of your aeroplane to take the picture in Figure 2.1.1). All you need now is a coffee table to put this thing on. Al: Thanks for all those amusing drunken phone calls and nights out in London and Peover (Social centre of Cheshire) and for being the best sister I've got!

ABBREVIATIONS

A:	Adenine
Ap:	Ampicillin
ATP:	Adenosine Triphosphate
bp:	base pair
C:	Cytosine
CsCl:	Caesium Chloride
dATP:	deoxy Adenosine Triphosphate
dCTP:	deoxy Cytosine Triphosphate
DGGE:	Denaturant Gradient Gel Electrophoresis
dGTP:	deoxy Guanosine Triphosphate
DNA:	Deoxyribonucleic Acid
DNase	Deoxyribonuclease
dNTP:	deoxy Nucleotide Triphosphate
EDTA:	Ethylene Diamine-Tetra-Acetic Acid
G:	Guanine
Hg:	Mercury
$HgCl_2$	Mercuric Chloride
Hg^R:	Mercury resistant
IPTG:	Isopropyl-β-D-galactosidase
IR:	Inverted repeat
IS:	Insertion sequence
kb:	kilobase pairs
kBq:	kiloBequerels
kDa:	kiloDaltons
LA:	Luria Bertani agar
LB:	Luria Bertani broth
PCR:	Polymerase Chain Reaction
PEG:	Polyethylene Glycol
RFLP:	Restriction Fragment Length Polymorphism
RNA:	Ribonucleic Acid
RNase:	Ribonuclease
SB:	Bacteria isolated from Salterbrook Bridge
SDW:	Sterile distilled water
SDS:	Sodium dodecyl sulphate
SE:	Bacteria isolated from Fiddlers Ferry sediment
SO:	Bacteria isolated from Fiddlers Ferry soil
T:	Thymine
TE:	Tris/EDTA
Tet:	Tetracyclin
Tris:	Tris (hydroxymethyl) methylamine
Tn:	Transposon
T2:	Bacteria isolated from Tipperary site
UV:	Ultraviolet
X-gal:	5-Bromo-4-chloro-3-indoyl-l-β-D-galactopyranoside
'93:	Bacteria isolated in 1993 by Osborn *et al.*, 1993
'96:	Bacteria isolated for this study

CHAPTER 1.

INTRODUCTION.

1.1. Gene transfer in the environment.

1.1.1. Bacterial diversity in the natural environment.

The study of bacterial diversity allows the characterisation of bacterial populations in a range of environments, and is important in assessing the evolution and adaptation of bacterial populations, e. g. due to selective pressures (Arber *et al.*, 1994, Bruce *et al.*, 1995, Holt *et al.*, 1996, Osborn *et al.*, 1995, Pearson *et al.*, 1996, Smit *et al.*, 1997). There are several factors which can be used to study bacterial diversity including species diversity, functional diversity and genetic diversity (Bruce *et al.*, 1995, Ellis *et al.*, 1995, Holt *et al.*, 1996, Osborn *et al.*, 1995, Pearson *et al.*, 1996, Roane and Kellogg, 1996, Sabry *et al.*, 1997).

The number and type of bacterial species present in an environment can be studied to ascertain whether or not the same bacterial populations are present in different systems. This approach has been used in studying the effects of selective pressures placed upon environmental systems (Osborn *et al.*, 1995, Roane and Kellogg, 1996, Sabry *et al.*, 1997). In general, selective pressures tend to decrease the number and diversity of the bacteria present in that particular environment (Hart *et al.*, 1998, Olson *et al.*, 1991, Osborn *et al.*, 1993, Osborn *et al.*, 1995, Rochelle *et al.*, 1991, Vaituzis *et al.*, 1975). The functional diversity of the bacteria present in a particular environment can be studied using a variety of techniques including BIOLOG and fatty acid methyl ester (FAME) analysis. Both these approaches can be used to compare the overall population in a given environment but are only useful in limited circumstances (Ellis *et al.*, 1995, Kennedy and Gewin, 1997, Tunlid and White, 1990).

The use of molecular tools is important as it allows the characterisation of genetic components within the bacterial population of a particular environment. Such tools can be used to study the individual genes involved in a particular metabolic process and to monitor the diversity of that gene in a particular environment. The study of gene transfer in the natural environment using molecular genetic techniques is of increasing importance as it allows an understanding of the mechanisms whereby genetic material is transferred between bacteria and other microorganisms. A wide range of mechanisms exist that can alter genetic diversity. These include mutation, DNA replication errors, recombination, conjugation, transformation, transduction, transposition and integration (Dahlberg *et al.*, 1998a, Grinsted *et al.*, 1990, Harding, 1996, Holt *et al.*, 1996, Liebert *et al.*, 1997, Stokes and Hall, 1989, Sundström, 1998, Trevors *et al.*, 1987, Yin and Stotzky, 1997).

1.1.2. The study of gene transfer.

There are several mechanisms of gene transfer including conjugation, transformation, transduction and transposition, all of which provide a mechanism for the generation and rapid dissemination of genetic diversity through a given population (Dahlberg *et al.*, 1998a, Grinsted *et al.*, 1990, Harding, 1996, Trevors *et al.*, 1987, Yin and Stotzky, 1997). Evidence for gene transfer has been provided by the presence of related gene sequences in distantly related bacteria (Kirby, 1992, Osborn *et al.*, 1993, Pearson *et al.*, 1996). Gene transfer has been studied in a variety of environments including water, soil, sediment, compost, industrial fermenters and clinical samples (Grinsted *et al.*, 1990, Hermansson and Linberg, 1994, Paget and Simonet, 1994, Trevors *et al.*, 1987, Yin and Stotzky, 1997). An understanding of the mechanisms involved in gene transfer in the natural environment has implications for the release of genetically engineered microorganisms (GEMS) into the environment as it will allow the processes involved in the release of such organisms to be understood (Leung *et al.*, 1994, Sayre *et al.*, 1991).

1.1.3. Methods of gene transfer.

Gene transfer was first observed in 1928 in *Streptococcus pneumoniae* (Griffith, 1928), although at this stage the nature of this transfer was not fully understood. Since this initial study gene transfer into bacteria either from external sources or from other organisms, including bacteria and viruses, has been characterised in a wide variety of host organisms and environments (Yin and Stotzky, 1997). The four major mechanisms of gene transfer in bacteria; conjugation, transformation, transduction and transposition, are discussed below;

1.1.3.1. Conjugation.

Conjugation is described as the process whereby plasmids are transferred between two bacterial cells (Ippen-Ihler., Trevors *et al.*, 1987, Yin and Stotzky, 1997). This process involves cell to cell contact, so that unlike transformation the DNA is never external to the cell. Conjugation has been observed in a wide range of environments including soil, plants, rhizosphere, fresh water, sea water and activated sludge (Yin and Stotzky, 1997).

Plasmids that are capable of conjugation, generally carry *tra* genes that allow transfer to occur. Such genes are normally carried on an operon. Certain plasmids, which do not carry *tra* genes, can be transferred via a mobilisation process if they are present in a cell which contains a conjugative plasmid (Lebaron *et al.*, 1994). The process of conjugation involves the use of extracellular structures called pili, which facilitate contact between the two cells and provide a means of transfer for the DNA to move between the cells (Trevors *et al.*, 1987, Yin and Stotzky, 1997). Within the diverse group of conjugative plasmids two distinct groups exist; the broad host range plasmids, which are capable of transfer between unrelated groups of bacteria and the narrow host range plasmids, which are only capable of transfering within a distinct group of bacterial species.

Incorporation of the plasmid into the bacterial chromosome can occur, however as plasmids are usually capable of self replication, this is often not necessary

for the survival of the plasmid DNA. The transfer of plasmids by conjugation is not limited to groups of closely related bacterial species, as conjugation has been observed between Gram-positive and Gram-negative bacteria, between bacteria and yeast cells and even between bacteria and plants (Harding, 1996, Trevors *et al.*, 1987, Yin and Stotzky, 1997). Once inside the recipient cell, the plasmids survival depends on the nature of any other plasmids in the cell, as certain plasmids are unable to coexist. Plasmids which are incapable of co-existence in the same cell are said to be in the same incompatibility group and are usually genetically related. Incompatibility is caused by the two plasmids containing similar replication or partitioning mechanisms (Austin and Nordström, 1990, Nocick, 1987).

Mobilisation of plasmids which do not contain the *tra* genes necessary for transfer can be achieved in two ways. The non conjugative plasmid can co-transfer with the conjugative plasmid by making use of the *tra* gene products of the conjugative plasmid. Alternatively, the non conjugative plasmid can recombine with the conjugative plasmid to form a single large plasmid structure. Once inside the cell, the plasmids may, or may not separate into their original forms.

1.1.3.2. Transformation.

Transformation is the uptake and incorporation of extracellular DNA into the bacterial genome (Trevors *et al.*, 1987, Yin and Stotzky, 1997). This can occur naturally or artificially, although natural transformation appears to only occur in bacteria. Artificial transformation is widely used as a molecular tool in a broad range of prokaryotic and eukaryotic organisms, including mammalian cells (Trevors *et al.*, 1987, Yin and Stotzky, 1997). The method of transformation first involves the release of DNA from either dead or viable cells, followed by the dispersal of this DNA into the environment surrounding the bacterial cell. The nature of this environment can affect the dispersal of the DNA, i.e. material which is released into an aquatic environment will be dispersed more efficiently than DNA released into soil where the DNA binds onto soil particles. The survival of the DNA will be affected by DNase

enzymes which are secreted by bacteria. However if the DNA is bound onto substances such as soil particles, humic acids or dead cellular material, these will protect the DNA by inhibiting DNase activity, therefore prolonging the survival of the DNA in the environment (Håvarstein, 1998, Paget and Simonet, 1994, Yin and Stotzky, 1997).

In order for the extracellular DNA to be taken up into the recipient bacterial cell, a physiological state know as competence must be achieved by the recipient cell. It has been suggested that the biological function of competence is to aid in the repair of damaged chromosomes (Håvarstein, 1998). Approximately 50 naturally competent species of bacteria have been described to date, including *Streptococcus pneumoniae*, *Acinetobacter calcoaceticus*, *Pseudomonas stutzeri* and *Streptococcus* sp (Håvarstein, 1998, Lawrence and Ochman, 1998, Lorenz and Wackernagel, 1994, Paget and Simonet, 1994, Yin and Stotzky, 1997). Competence can be induced artificially by the addition of calcium chloride to some bacterial cells making their cell walls permiable to DNA. This procedure is commonly used in molecular genetic techniques as a means of introducing foreign DNA into cells such as *Escherichia coli*. Bacteria which are capable of transforming DNA usually achieve this competent state using proteins encoded by genes contained within the bacterial chromosome. The extracellular DNA then binds onto the bacterial cell wall and is believed to be transported into the cell using cytoplasmic vesicles. Once inside the cell, the DNA undergoes homologous recombination into the bacteria chromosome. Depending on the nature and position of the novel DNA, expression of the genes contained on the transformed DNA strand may, or may not occur.

1.1.3.3. Transduction.

Transduction can be described as the transfer of genetic material between bacteria mediated by bacteriophages. This has been observed in a range of environments including animal guts, plants, soil and aquatic systems (Harding, 1996, Trevors *et al.*, 1987, Yin and Stotzky, 1997). This process was first described by

Zinder and Lederberg in 1951. There are two mechanisms of transduction which have been designated generalised transduction and specialised transduction (Harding, 1996, Trevors *et al.*, 1987, Yin and Stotzky, 1997). In *Escherichia coli* and related bacteria generalised transduction involves the transfer of any fragment of the bacterial genome, whereas specialised transduction involves the transfer of the genetic material directly adjacent to the site of phage insertion, and is brought about by imprecise excision of the phage DNA.

1.1.3.4. Transposition.

The existence of mobile genetic elements was first described in maize by B. McClintock in 1951, however this work was largely ignored until the description of further transposition events in both prokaryotes and eukaryotes in the late 1960s and 70s (Fiandt *et al.*, 1972, Hedges and Jacob, 1974, Jordan *et al.*, 1968, Shapiro *et al.*, 1969). Transposition is described as the process whereby discrete, self contained genetic entities, known as transposable elements, move from one position in the genome to another. Transposable elements have been observed in a wide range of genera in both prokaryotic and eukaryotic organisms and play a central role in evolution by providing a mechanism for the generation of diversity (Brown and Evans, 1991, Grinsted *et al.*, 1990, Harding, 1996, Kahn and Scharfer, 1995, Kleckner, 1981, Mahillon and Chandler, 1998, Mahillon, 1998). The rapid transfer of this diversity to other bacteria can be achieved by the association of transposable elements with other mechanisms of gene transfer such as conjugation and transformation, thus providing a huge potential for the dissemination of diversity within a given bacterial population.

There are several classes of transposable element, which are described in Section 1.1.4., along with their transposition mechanism (Kleckner, 1981, Brown and Evans, 1991). Transposition mediated transfer of DNA represents a major mechanism for gene transfer in the natural environment due to the increased mobility of those genes contained within the transposon. The nature of the genes carried is dependant

on the type of transposable element. The simplest elements only carry those genes necessary for transposition, whereas more complicated elements carry other, non essential genes, which may confer selective advantages to the host organism. Most transposable elements are flanked by a set of inverted repeat (IR) sequences which vary in length and play an important role in the transposition process.

1.1.3.4.1. Conjugative transposons.

Conjugative transposons were first described in a strain of *Enterococcus feacalis* by Franke and Clewell, 1981, following the discovery of Tn*916* which carries a tetracycline resistance gene, *tetM*. The mechanism whereby this transposon integrates and excises itself from the host bacterial genome is different to the transposition mechanism utilised by most other transposons. Conjugative transposons form a circular intermediate structure which is then transferred to the recipient cell in a process akin to conjugation, although these two processes are distinctly different (Scott and Churchward, 1995). Since the initial discovery of Tn*916*, conjugative transposons ranging in size from 10 to 150kb have been reported in a wide range of both Gram-negative and Gram-positive bacterial genera, (Scott and Churchward, 1995).

1.1.4. Classes of transposon.

Historically, a four class system has been used for the classification of transposable elements according to their size, the genes present and their transposition mechanism (Grinsted *et al.*, 1990, Kleckner, 1981). These four classes are:

Class I. IS elements and composite transposons.

Class II. Cointegrative (Tn*3*-like) transposons.

Class III. Transposable bacteriophages.

Class IV. Those elements that cannot be placed in classes I, II or III.

More recently a nine class system has been proposed, due to the large numbers of novel and unclassifiable elements which have now been described. This system

splits class II into several parts, and describes several novel elements (Mahillon, 1998).

Class I. IS elements and composite transposons.

Class II. Cointegrative (Tn*3*-like) transposons

Class III. Integrative transposons.

Class IV. Conjugative transposons.

Class V. Transposable bacteriophages.

Class VI. Integrative phages.

Class VII. Integrative plasmids

Class VIII. Integrons.

Class IX. Shufflons.

The four class system is described below. Novel elements such as conjugative transposons and integrons are discussed where relevant.

1.1.4.1. Class I transposable elements.

This group comprises insertion sequences (IS elements) and composite transposons. IS elements are generally less than 2 kb in length and contain only the genes necessary for their own transposition, whereas composite transposons comprise two IS elements flanking a variable region of DNA (Mahillon and Chandler, 1998, Mahillon, 1998).

IS elements are capable of insertion at multiple, non specific, sites within the recipient DNA molecule and are frequently associated with genes for biodegredation, catabolic pathways and pathogenic pathways. They have also been found to participate in chromosome rearrangement and plasmid integration (Mahillon and Chandler, 1998). The majority of IS elements have IR's of between 10 and 40 bp in length at either end of their sequence. Within these IR's, two domains exist with different functions, the first is involved in the binding of the recombinase (transposase) protein and the second is involved in the cleavage and strand transfer

reactions which ultimately lead to the transposition reaction occurring (Mahillon and Chandler, 1998, Mahillon, 1998).

Three mechanisms of transposition are known to occur in IS elements. Most elements utilise a "DDE"-type transposase gene, this is so called because of the conserved amino acid triad "DDE" on the protein, which is involved in the strand transfer reactions. The mechanism of transposition involves the cleavage of the DNA strands at the ends of the IS element, followed by the incorporation of the element into the target site (Mahillon, 1998). Depending on the exact pattern of strand cleavage, either conservative transposition will occur, whereby the IS element is merely transferred from one position in the genome to another, or replicative transposition will occur, resulting in two copies of the element, one at the original site and one at the target site. Those elements which utilise this mechanism of transpositional recombination are also flanked by direct repeats (DR) which correspond to the duplication of the target site during the transposition process (Mahillon, 1998).

Two IS families however, utilise different mechanisms of transposition, as they do not have the IR's that are essential for the "DDE" transpositional recombination to occur. The IS*91* family uses a rolling circle transposition mechanism, whereby DNA is transferred via a single stranded intermediate in a mechanism similar to rolling circle replication which occurs during conjugation (Mahillon, 1998). The IS*110* family however, uses a site specific recombination system and as such, it has been suggested that these elements should be considered as "site specific recombinational modules" rather than transposable elements (Mahillon, 1998). Neither the IS*110* family or the IS*91* family produce DR's during their transposition process. Over 500 of these elements have been described to date which have been classified into 17 families depending on their transposase gene sequence similarity, their IR sequences, the length of their DR and their transposition mechanism (Mahillon *et al.*, 1994, Mahillon and Chandler, 1998, Mahillon, 1998).

IS elements are capable of mobilising non-transposable regions of DNA by flanking it to create a composite transposon which then transposes as a single unit. Transposons such as Tn*5*, Tn*9* and Tn*10* were created using this mechanism (Mahillon, 1998). It is not necessary for composite transposons to carry two copies of the same IS element, nor do these have to have complete sets of IR's. In addition, the two IS elements flanking a composite transposon do not both have to be functional, e.g. Tn*10*, which carries a tetracycline resistance gene, is flanked by two copies of IS*10*; IS*10*L and IS*10*R. Of these, transposition is more efficient using IS10R, which is also capable of transposing as an independent unit. Composite transposons also create DR's during their transposition, e.g. Tn*10* creates DR's of 9bp.

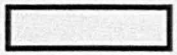
1.1.4.2. Class II transposable elements.

This class of transposable elements is often referred to as the Tn*3* family of transposons after the archetypal member of this class. These elements are larger than the IS elements, as they usually contain genes which are not essential for their basic transposition functions and as such, provide a greater potential for the transfer of genetic material in the environment than other transposable elements (Brown and Evans, 1991, Grinsted *et al.*, 1990, Kleckner, 1981, Mahillon, 1998). Most class II elements contain two genes which are necessary for transposition; the transposase genes (*tnpA*) which are similar in function, but not sequence, to the transposase genes found on IS elements and the resolvase genes (*tnpR*) which are not found in IS elements (Grinsted *et al.*, 1990, Kleckner, 1981, Mahillon, 1998, Pearson *et al.*, 1996). The majority of class II transposable elements have IR's of 38bp at either end of their sequence to allow recognition of the element by the transposase gene (Mahillon, 1998). A *res* site is also present in most class II elements, which is involved in the transposition process (Pearson *et al.*, 1996).

Class II transposons virtually always utilise a two step replicative transposition process which is mediated by the *tnpA* and *tnpR* gene products. This process is shown in Figure 1.1.4.2.1., along with the conservative transposition pathway. The

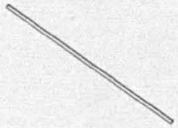
Figure 1.1.4.2.1. Schematic representation of replicative and conservative transposition via a Shapiro intermediate.

Key:



This represents one strand of the original transposon.

This represents one strand of the newly synthesised transposon.



DNA of original donor replicon.

DNA of recipient replicon.

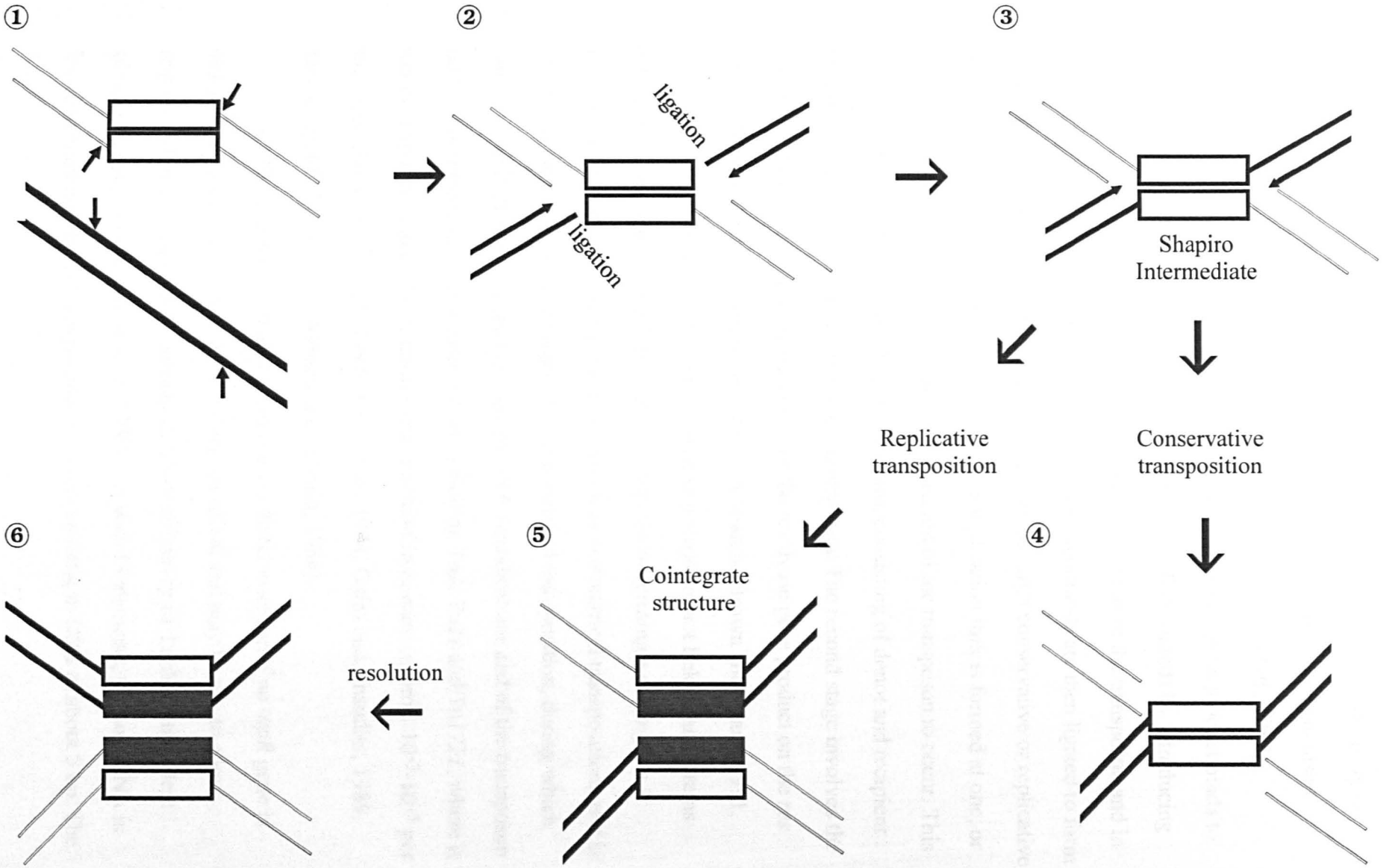
① Action of enzymes leads to cleavage of donor and recipient DNA.

③ Shapiro intermediate.

④ Conservative transposition leads to the formation of one copy of the transposon on the recipient DNA strand.

⑤ Replicative transposition leads to the formation of two copies of the transposon, on both the donor and recipient DNA strands.

⑥ This cointegrate structure is resolved into two seperate replicons by the action of the *tnpR* gene on the *res* site of the transposon.


①
②
③
ligation
ligation
Shapiro
Intermediate
Replicative
transposition
Conservative
transposition
⑥
⑤
④
Cointegrate
structure
resolution

transposition of class II elements usually occurs into random sites on the recipient DNA strand; however one class II transposon, Tn*502*, has been shown to insert preferentially into a specific site in the plasmid RP1 (Stanisich *et al.*, 1989).

During the first stage of transposition, the transposase gene product binds to the IR's and mediates the joining of host and recipient DNA strands by introducing single stranded nicks into the donor DNA strand at the ends of the transposon and in the opposite recipient strand at a random site. The two strands are then ligated to form a Shapiro intermediate. This intermediate can then undergo conservative or replicative transposition. During replicative transposition, a replication fork is formed at one, or both ends of the transposon, allowing the replication of the transposon to occur. This results in the formation of a cointegrate structure, consisting of donor and recipient replicons, each containing one copy of the transposon. The second stage involves the resolution of this cointegrate by the action of the resolvase gene product on the *res* site, resulting in the separation of the two DNA strands (Brown and Evans, 1991, Grinsted *et al.*, 1990, Kleckner, 1981). These two steps are not linked, and strains lacking a *res* site and/or a *tnpR* gene are still capable of forming a cointegrate structure. This method of transposition is known as symmetrical transposition, as it is initiated at both ends of the transposon. Asymmetrical transposition, during which transposition is initiated by the cleavage of DNA strands at one end of the transposon has been described in several transposons, including Tn*3*, Tn*21* and Tn*1721*, where it occurs at approximately 1% of normal transposition frequency (approx. 10^{-3}-10^{-5} per cell generation)(Arthur *et al.*, 1994, Avila *et al.*, 1984, Galas and Chandler, 1989, Harshey and Bulchan, 1981, Motsch and Schmitt, 1984).

Tn*3*-like elements can also transpose via direct insertion if no *tnpR* gene is present. This accounts for 5% of normal transposition and may be due to a low frequency direct transposition pathway or a low efficiency of TnpR independent cointegrate resolution (Bennett *et al.*, 1983). As with IS elements, the host DNA is also replicated during the transposition process, resulting in DR's of about 5 bp. The

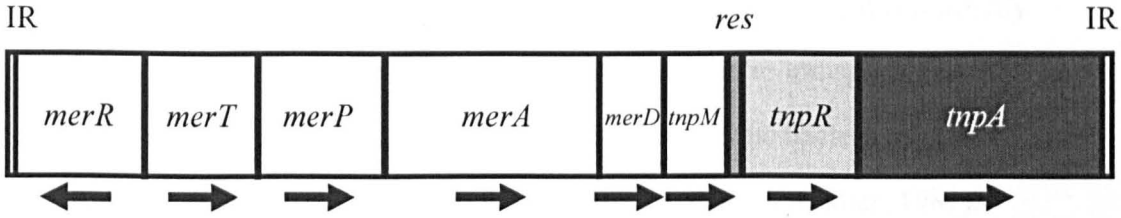
transfer mechanism of conjugative transposons differs to this and is described in 1.1.3.4.1.

Class II transposons carry a wide range of genes including those for heavy metal resistance, antibiotic resistance and degradation of compounds such as toluene and chlorobenzene (Bennet *et al.*, 1978, de la Cruz and Grinsted, 1982, Lebrun *et al.*, 1994, Mäe and Heinaru, 1994, van der Meer *et al.*, 1991, Tolmasky and Crosa, 1987). The structures of several class II transposons are shown in Figure 1.1.4.2.2. The class II elements Tn*4430* and Tn*5401* do not contain any non-essential genes and as such are smaller (approx. 4kb) than other members of this class (Baum, 1994, Lereclus *et al.*, 1986). Within the class II transposons it is not unusual for one element to transpose into another element, e.g. Tn*2501* is found within Tn*951*, Tn*4651* is found within Tn*4563* and Tn*5384* is found within Tn*5385* (Mahillon, 1998, Michiels *et al.*, 1987, Rice and Carias, 1998, Tsuda *et al.*, 1989). The potential for the acquisition of new genetic material is increased if the transposon contains an integron structure (Stokes and Hall, 1989). Integrons, which are usually found on class II transposons, are capable of transferring DNA into and out of a recombination hot spot (*rhs*) via circular intermediates (known as gene cassettes)(Stokes and Hall, 1989, Sundström, 1998). The genes contained within these cassettes generally encode for antibiotic resistance genes. It is believed that these circular intermediates are taken up into the cell by transformation, however transfer by conjugation and transposition is also possible depending on the precise location of the integron. Integrons are covered further in 1.2.5.2.

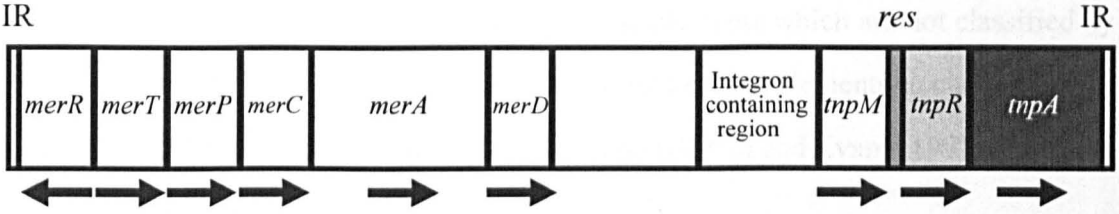
1.1.4.3. Class III transposable elements.

This class contains the transposable bacteriophages, which utilise transposition during their life cycle. The best characterised class III transposable element is the bacteriophage, Mu (Brown and Evans, 1991, Kleckner, 1981). Like class I and II elements, transpositional bacteriophages also produce DR's during their transposition process, in this case these are approximately 5 bp in length. The transpositional

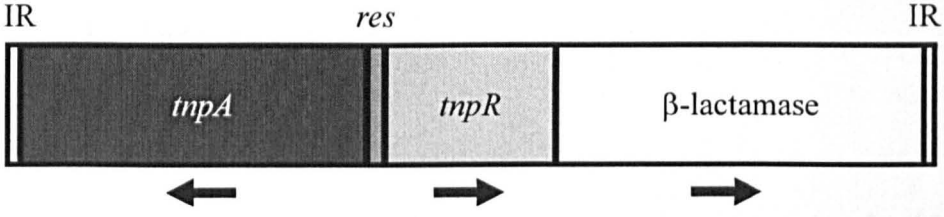
Tn*501*, 8.3kb:



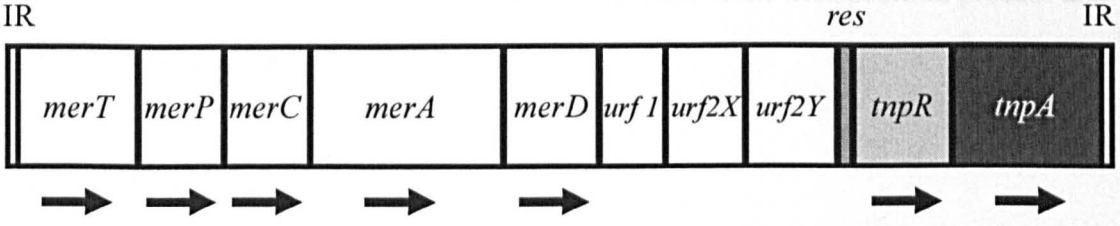
Tn*21*, 19.3kb:



Tn*3*, 5.0kb:



Tn*5036*, 8.0 kb:



Tn*4430*, 4.8kb:

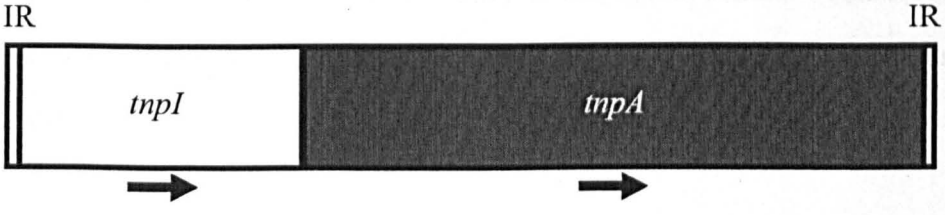


Figure 1.1.4.2.2. Schematic diagram showing the structures of Tn*501*, Tn*21*, Tn*3*, Tn*5036* and Tn*4430*. Arrows indicate the direction of translation. Not to scale.

mechanism utilised by class III elements is usually replicative and is normally achieved via a cointegrate DNA molecule. Non replicative transposition can occur however this only occurs during the initial stage where the bacteriophage genome is transferred into the hosts DNA (Brown and Evans, 1991, Kleckner, 1981).

1.1.4.4. Class IV transposable elements.

This group encompasses all transposable elements which are not classified by classes I, II and III. These include Tn*7*, which unlike most elements in classes I-III, transposes into a specific site in the host genome (Brown and Evans, 1991, Craig, 1991, Craig, 1996).

1.2. The Tn*3* family of transposons.

1.2.1. Genes and structures essential for transposition.

Most functional Class II transposons carry two genes involved in the transposition process; the transposase gene, *tnpA* and the resolvase gene, *tnpR*. Most transposons also contain a *res* site which is involved in the transposition process and are flanked by 38bp IR's:

1.2.1.1. IR sequences.

Most class II transposons are flanked by IR sequences which are recognised by the TnpA protein during the initial strand cleavage reaction of the transposition process. These IR's are generally 38bp in length and contain a conserved region in the centre of the sequence. The outer sequences of the IR's of different transposons are variable.

1.2.1.2. The *res* site.

These are approximately 130bp in length and are the site for the recombination of cointegrate transposon structures during the replicative transposition process (Grinsted *et al.*,1990, Kleckner, 1981). *res* sites comprise three highly conserved DNA binding sites, all of which are required for the resolution of the cointegrate DNA structure. Each site consists of two nine base pair repeats (TGTCRYTTA)(details of degeneracy codes are given in Table 2.15.1), which flank a central spacer region of 10, 16 and 7 bp in sites I, II and II respectively. The position of the *res* site relative to the *tnpA* and *tnpR* genes is of particular significance because if the *res* site is between the two genes, as it is in Tn*3*, then the two genes can recombine separately, whereas if the *res* site is next to the two genes then they are more likely to evolve as a a single unit (Grinsted *et al.*, 1990).

1.2.1.3. The transposase gene.

The transposase gene, *tnpA*, produces a protein which is involved in the recognition of IR sequences, the subsequent cleavage of donor and recipient DNA strands and their ligation to form a cointegrate structure. (Chou *et al.*, 1979, Pearson *et al.*, 1996, Grinsted *et al.*,1990). In class II transposons, this gene is approximately 3050 bp in length and produces a protein of approx. 110 kDa, whereas the transposase genes of IS elements are much smaller and in most cases, show no sequence similarity to the class II transposases (Grinsted *et al.*,1990, Mahillon and Chandler, 1998). The N-terminus (amino acids 1-216) of the protein is involved in the identification of the IR sequences. Within this, the amino acid residues 33-64 are highly conserved and it believed that this sequence corresponds to the conserved region of the IR sequences, whereas the variable region corresponds to the variable regions of the IR sequences (Evans and Brown, 1987). The C-terminal of the TnpA protein is believed to be responsible for cleavage and ligation of DNA strands (Brown *et al.*, 1985a, Turner, 1989, Ward and Grinsted, 1989).

The function of *tnpA* genes is such that often, the corresponding gene from one transposon can complement its homologue in an alternative transposon (Grinsted *et al.*,1982). However, not all Tn*3*-like transposons are capable of complementing each others *tnpA* functions. Complementation groups seem to exist within the Tn*3* group of transposons, i.e. Tn*501*, Tn*1721* and Tn*4653* can complement each others functions, as can Tn*21*, Tn*2603*, Tn*2613*, Tn*3926* and Tn*4000* (Grinsted *et al.*, 1990, Lett *et al.*, 1985, Schmitt, 1984, Schmitt *et al.*, 1979, Tanaka *et al.*, 1983, Tsuda *et al.*, 1989, Turner, 1985). It has also been shown that the IR's from several transposons are interchangeable, i.e. the transposase gene from Tn*21* is capable of recognising the IR sequence of Tn*501*, *1721* and Tn*2501* (Grinsted *et al.*, 1988, Martin *et al.*, 1989). Despite their conserved function, the DNA sequences of *tnp*A genes varies in different transposons. The *tnp*A genes of Tn*21* and Tn*501*, for example, show approximately 75% DNA sequence similarity and 72% amino acid identity (85% similarity)(Brown *et al.*,1985a, de la Cruz and Grinsted, 1982).

1.2.1.4. The resolvase gene.

The resolvase gene, *tnpR*, produces a protein which is involved in the resolution of the donor/recipient cointegrate structure formed by the TnpA protein (Gill *et al.*, 1978, Heffron *et al.*, 1981, Grinsted *et al.*,1990). This process involves recombination of the host and donor replicons at the *res* site. The *tnpR* gene is approximately 560 bp in length and is only found in class II elements. The protein formed by the gene contains a helix-turn-helix motif, which are common to many DNA binding proteins. The protein contains two domains, a smaller (5 kDa) C-terminal domain, which binds specifically to the *res* site and a larger (15.5 kDa) N-terminal domain, which is involved in strand exchange during recombination and protein/protein interactions (Abdel-Meguid *et al.*, 1984, Halford *et al.*, 1985). The function of *tnpR* genes is such that the corresponding gene from one transposon can often complement its homologue in an alternative transposon (Diver *et al.*,1983). Like the *tnpA* gene, the DNA sequences of *tnpR* genes from different transposons are

diverse. The *tnpR* genes of Tn*21* and Tn*501*, for example show 75% DNA sequence similarity and 80% amino acid identity (85% similarity), however most of this similarity is concentrated at the amino terminal of the gene (Brown *et al.*,1985a, de la Cruz and Grinsted, 1982).

1.2.1.5. *tnpM.*

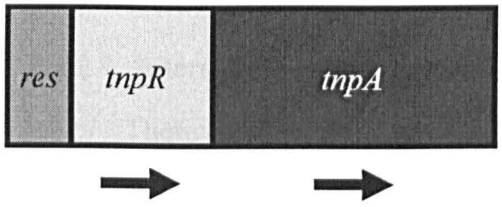
Certain transposons, e.g. Tn*21* and Tn*501* contain an additional gene, *tnpM*, of uncertain function. It has been proposed that this gene codes for a modulator protein which enhances transposition and suppresses cointegrate resolution (Hyde and Tu, 1985) however, more recently, the validity of this proposal has been questioned (Grinsted *et al.*, 1990).

1.2.2. Structural diversity of class II transposon genes.

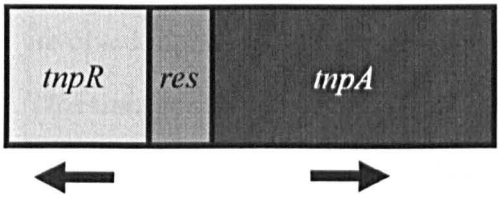
Although most class II transposons contain the *tnpA*, *tnpR* and *res* sites previously described, the structural arrangement of these components, as well as their relative orientation, varies in different transposons (Grinsted *et al.*,1990). The different structures of these elements in several transposons is shown in Figure 1.2.2.1., along with the possible remaining arrangement. Most transposons in this group contain one *tnpA* gene and one *tnpR* gene, which are present in different orientations, dependant on the individual transposon. As can be seen, Tn*3* has divergently transcribed *tnpA* and *tnpR* genes with the *res* site located in-between the two genes. The Tn*3*-like arrangement is also found on transposons such as Tn*2501* and Tn*1331* (Michiels *et al.*, 1987, Tolmasky and Crosa, 1987). Tn*4556* has the *tnpA* and *tnpR* genes both transcribed in the same direction again with the *res* site in-between the two genes.

Transposons in the Tn*21* subgroup have the *res* site upstream from the *tnpR* gene which is contiguous with, and transcribed in the same direction as the *tnpA* gene. The study of this structural diversity has important implications in the study of the evolutionary relationships between different transposons. As previously stated the

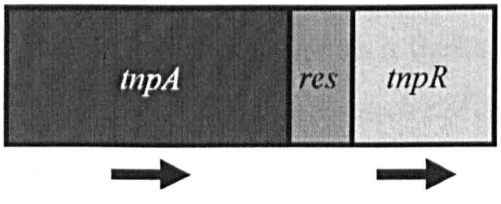
Tn*21*-like arrangement. This is the most common arrangement of the *res* site and *tnpA* and *tnpR* genes:



Tn*3*-like arrangement. This is present in a number of transposons including; Tn*3*, Tn*163*, Tn*1546*, Tn*2501*, Tn*5393*, Tn*5403* and Tn*5563*:



Tn*4556* arrangement. This arrangement has not been described in other transposons and, at present, is unique:



Apart from the position of the *res* site, there are four possible arrangements of *tnpA* and *tnpR* genes. The remaining arrangement is shown below along with possible *res* site locations:

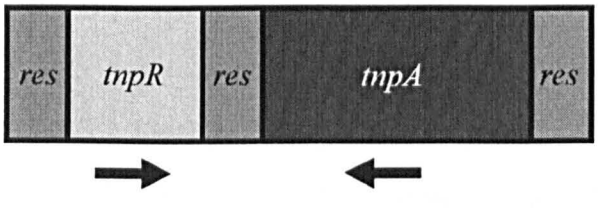


Figure 1.2.2.1. Structural arrangement of *tnpA, tnpR* and *res* sites.

position of the *res* site within the *tnpA* and *tnpR* genes is significant as it may allow recombination to occur between the two genes via the *res* site.

1.2.2.1. Alternative arrangements of *tnpA* and *tnpR* genes.

There are several transposons which contain novel arrangements of *tnpA* and *tnpR* genes. Tn*2610*, a 24kb transposon, carries genes for resistance to streptomycin and sulphonamide and has 3.5kb elements as inverted repeats at each end of the transposon. This transposon carries one functional *tnpA* gene and one non-functional *tnpA* gene. It also carries two functional copies of the *tnpR* gene, both of which are involved in the resolution of transposition intermediate structures (Yamamoto, 1989). This transposon may represent a class II composite transposon, i.e. a region of DNA flanked by two class II transposons.

Tn*5271*, a 17kb transposon carries the *cbaBC* gene complex, which allows the host organism to degrade 3- and 4-chlorobenzoate. The transposon is flanked by a directly repeated sequence of 3201 bp (IS*1071* itself an unusual transposable element, see 1.2.3.1.), which in turn is flanked by 110-bp inverted repeats. This transposon contains a *tnpA* gene, but does not contain a *tnpR* gene and is only capable of performing the first step in class II transposition (di Gioia *et al.*, 1998).

1.2.3. Alternative transposition systems.

There are several "Class II" transposons which do not contain the *tnpA*/*tnpR* transposition system described previously:

1.2.3.1. IS*1071*.

This IS element which was isolated from an *Alcaligenes* strain, contains a transposase gene which is similar to that found in class II elements (Nakatsu *et al.*, 1991). All previously described transposase genes isolated from IS elements were smaller and showed no sequence similarity to the class II transposase genes. Since this

discovery, several more examples of such elements have been described (Di Gioia *et al.*, 1998) leading to the characterisation of this new type of "class II" transposons.

1.2.3.2. Tn*4430* and Tn*5401*.

Tn*4430*, a 4.8kb transposon, was first described by Lereclus *et al.*, (1983) in a strain of *Bacillus thuringiensis* and has since been described in a wide range of plasmids, mostly those containing genes involved in the formation of crystal protein toxins in species of *Bacillus thuringiensis*. This transposon utilises a *tnpI*-encoded (284 amino acids) resolution system, which has a gene product similar to that of the integrase gene family (Mahillon and Seurinck, 1998). Tn*5401*, also found in strains of *Bacillus thuringiensis* also utilises this *tnpI* resolution system (Baum, 1994).

1.2.3.3. Tn*ISXc5*.

This transposon, originally isolated from a strain of *Xanthomonas campestris* (Tu *et al.*, 1989) has an resolvase gene with two features which differentiate it from the majority of class II resolvase genes. Firstly, the gene is more closely related to the invertase gene than other resolvases. Secondly, the carboxy terminal of the protein has been extended with a region of unknown function (Liu *et al.*, 1998). Despite these anomalies, the function of the gene remains akin to that of a "normal" resolvase gene.

1.2.3.4. Tn*4451*.

This 6.3kb chloramphenicol resistance transposon isolated from a strain of *Clostridium perfringens*, transposes via a circular intermediate in a mechanism comparable to conjugative transposition. However, unlike conjugative transposons, it is not self transmissible and can only transpose if mobilised. The transposon contains five genes which are believed to be involved in this process, *tnpV*, *tnpW*, *tnpX*, *tnpY* and *tnpZ*. *tnpX* is involved in the excision of the circular intermediate and *tnpZ* shows homology to the Mob/Pre family of mobilisation/recombination proteins. The functions of *tnpV*, *tnpW* and *tnpY* are unknown (Crellin and Rood, 1998).

1.2.3.5. Tn*5053* and Tn*5090* (Tn*402*).

Both these transposons utilise an operon containing *tniA*, *tniB*, *tniQ* and *tniR* genes for transposition. *tniA*, *tniB* and *tniQ* are involved in the formation of a cointegrate structure, whereas *tniR* is involved in resolution of this cointegrate at the *res* site (Kholodii *et al.*, 1993, Kholodii *et al.*, 1995, Rådstrom *et al.*, 1994).

1.2.3.6. Conjugative transposons.

Conjugative transposons move by excising themselves from the host DNA molecule, forming a circular intermediate and finally inserting into a target site in the recipient genome. An integrase gene similar to that found in integrons is responsible for this process (Scott and Churchward, 1995).

1.2.4. Transposition immunity.

This is the phenomenon whereby a target DNA strand, which is already carrying a transposable element, exhibits a significantly reduced affinity for the insertion of a second copy of that element. This was first described by Robinson *et al.*, 1977, and appears to be limited to complex transposons such as Mu, Tn7 and various members of the Tn*3* family of transposons. The immunity of the Tn*3* family is believed to involve the 38bp IR's and the *tnpA* gene although the precise mechanism is unclear (Arciszewska *et al.*, 1989, Grinsted *et al.*,1990, Mahillon and Chandler, 1998).

1.2.5. Transposon associated genes.

A wide range of non-essential genes are carried on class II transposons several examples of which are given below:

1.2.5.1. Mercury resistance (*mer*) operon.

Several transposons carry genes conferring resistance to mercury compounds including; Tn*21*, Tn*501*, Tn*502*, Tn*2603*, Tn*3926*, Tn*5036*, Tn*5041*, Tn*5053*, Tn*5059*, Tn*5086*, Tn*5384*, Tn*5385*, Tn*5467*, and Tn*5563* as well as those unnamed transposons carried on pKLH2, pPB, pMER610, pMER327/419, pMER330 and pMER05 (Brown *et al.*, 1985b, Bonafede *et al.*, 1997, Clennel *et al.*, 1995, Hobman *et al.*, 1994, Kholodii *et al.*, 1993, Kholodii *et al.*, 1995, Kholodii *et al.*, 1997, Kholodii *et al.*, 1997, Lett *et al.*, 1985, Reniero *et al.*, 1998, Rice and Carias, 1998, Stanisich *et al.*, 1998, Sundström *et al.*, 1993, Ward and Grinsted, 1987, Yamamoto *et al.*, 1981, Yeo *et al.*, 1998, Yurieva *et al.*, 1997). All these transposons carry mercury resistance genes on the *mer* operon which is discussed further in 1.3.3.

1.2.5.2. Integrons.

Integron structures have been identified on certain transposons (de la Cruz and Grinsted, 1982, Stokes and Hall, 1989, Sundström, 1998). These are capable of the insertion and excision of specific gene cassettes into/out of a site specific recombinational hot spot (*rhs*) or, more rarely, to a non-specific site (Recchia and Hall, 1995). This effectively provides a means whereby the transposon can acquire novel genetic material. Integrons contain an integrase gene similar to that found in a wide range of organisms, including conjugative transposons and the λ phage, which is responsible for the integration and excision of gene cassettes (Ouellette and Roy, 1987, Scott and Churchward, 1995). The gene cassettes involved in this process are discrete mobile units consisting of the gene itself and a 59-base pair element. These cassettes can exist as circular structures in the cell but are more commonly found inserted into the *rhs* of an integron. The 59-base pair elements are present in all gene cassettes and are involved in the integration/excision process (Hall *et al.*, 1991). During this process, which allows several gene cassettes to be present in the *rhs* of an integron structure, the 59-base pair element is split into two parts, one at each end of the gene. Recombination of gene cassettes may also occur due to the similar nature of

the 59-base pair elements, thereby increasing the potential for gene transfer within the integron structure.

Several well documented transposons contain an integron including Tn7, Tn*21*, Tn*1696*, Tn*5086* and Tn*5090* (Bissonnette *et al.*, 1991, Rådstrom *et al.*, 1994, Sundström *et al.*, 1991, Sundström *et al.*, 1993, Ward and Grinsted, 1987). Remnants of integron structures have also been found in Tn*610* and Tn*1331* (Hall and Collis, 1998, Tolmasky, 1993). Three classes of integrons have been documented to date (Hall and Collis, 1998). Type 1 integrons are the most commonly encountered, e.g. the integron found on Tn*21* which carries an *aadA* gene coding for resistance to streptomycin/ spectinomycin is a Type I integron (de la Cruz and Grinsted, 1982). Type 2 integrons, such as that found on the class IV transposable element Tn7, contain non-functional integrase genes (Sundström *et al.*, 1991). The Type 3 group of integrons, at present, only contains one member. This integron carries an integrase gene with a sequence similarity of 61% to that found on Type 1 integrons. It also contains unusual 59 bp elements, which have sequences inserted into them. This unusual combination of integrase gene and 59bp elements may complement each other to allow successful integration of gene cassettes (Arakawa *et al.*, 1995). The genes which have been found inserted into the *rhs* site are most commonly antibiotic resistance genes, with over 40 types being classified to date (Hall and Collis, 1998) including genes for resistance to β-lactams, chloramphenicol, trimethoprim, aminoglycosides, quaternary ammonium compounds and erythromycin (Hall and Collis, 1998). This high occurrence of antibiotic resistance genes may be due, in part, to the high level of research into integrons in clinical isolates. The number of bacteria which contain integrons in estuarine environments has been estimated at 5% (Young, H. K. Personal communication). This represents an enormous potential for gene transfer in the natural environment.

1.2.5.3. Other transposon associated genes.

Other genes carried on transposons include genes for resistance to antibiotics e.g. Tn*3* contains a gene for resistance to β-lactams (Maekawa *et al.*, 1993). Tn*5422* carries genes for resistance to cadmium compounds (Lebrun *et al.*,1994) and Tn*5280* carries chlorobenzene degradation genes (van der Meer, 1991).

1.2.6. Evolution of transposons.

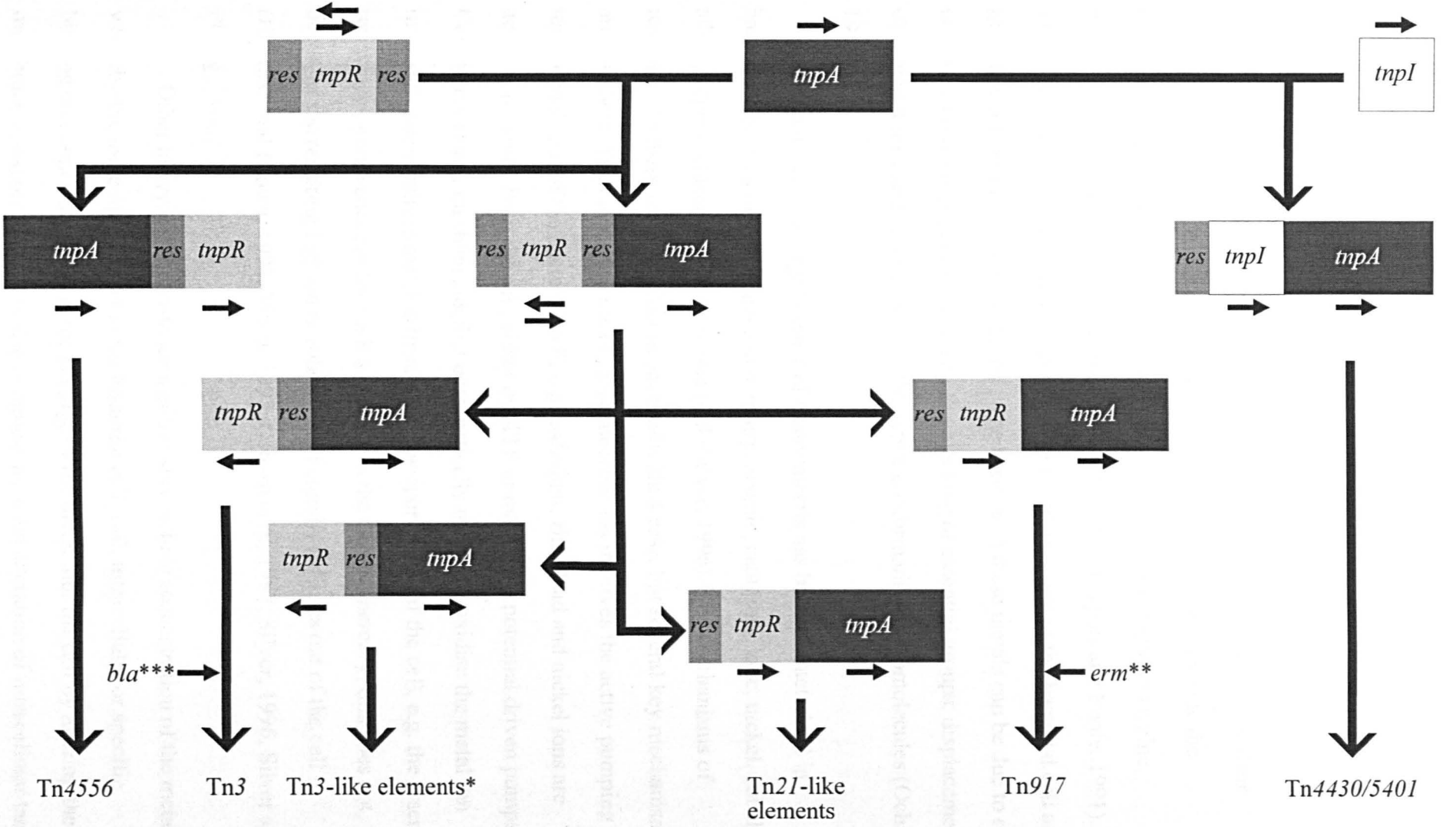
A number of different pathways have been proposed for the evolution of class II transposons (Grinsted *et al.*, 1990, Schmidt *et al.*, 1989, Tanaka *et al.*, 1983). These hypothetical pathways are generally based on genetic or physical maps of a small number of transposons and as such, provide a limited insight into the evolution of class II transposon structures (Schmidt *et al.*, 1989, Tanaka *et al.*, 1983). The pathway of evolution proposed by Grinsted *et al.*, 1990, differs from these previous schemes, in that it does not deal with the insertion and deletion events leading to the formation of variants of the same transposon, rather than considering the initial formation of the basic *tnpA/tnpR/res* transposition unit which forms the basic skeleton of the transposon to which other non essential genes can then be added. This proposed scheme for the evolution of transposable elements from these basic sub-units is shown in Figure 1.2.6.1. A more accurate scheme will only be obtained when the sequences from a large number of transposon structures are available, along with an indication of the time scale during which these structures have arisen.

Figure 1.2.6.1. Proposed evolution of transposons from *tnpA*, *tnpR*, *tnpI* and *res* elements. Adapted from Grinsted *et al.*, 1990.

* Several transposons are known to carry this Tn*3*-like arrangement of *tnpA*, *tnpR* and *res* site, including Tn*3*, Tn*163*, Tn*1546*, Tn*2501*, Tn*5393*, Tn*5403* and Tn*5563*.

** *erm* indicates the insertion of erythromycin resistance genes into the transposon structure.

*** *bla* indicates the insertion of β-lactamase resistance genes into the transposon structure.


res tnpR res
tnpA
tnpI
tnpA res tnpR
res tnpR res tnpA
res tnpI tnpA
tnpR res tnpA
res tnpR tnpA
tnpR res tnpA
res tnpR tnpA
bla***
erm**
Tn4556
Tn3
Tn3-like elements*
Tn21-like elements
Tn917
Tn4430/5401

1.3. Mercury resistance transposons.

1.3.1. Heavy metal resistance mechanisms in bacteria.

Bacteria require a large number of elements which are essential for their continued growth and survival (Hughes and Poole, 1991). These include the macronutrients; potassium, sodium and calcium, and the micronutrients; zinc, molybdenum, nickel and iron (Higgins and Burns, 1975, Hughes and Poole, 1991). Those metals which have no biological function and are toxic to the bacterial cell are generally referred to as "heavy metals". The toxicity of these metals can be due to one or more of three mechanisms, namely; the blocking of essential groups, displacement of essential ions and modification of the active conformation of biomolecules (Ochiai, 1977).

Resistance to a large number of heavy metals has been characterised in bacteria. These include resistance to mercury, arsenic, cadmium, zinc, nickel, cobalt, silver, copper, chromium, tellurite and lead (Silver, 1996). The mechanisms of resistance to these metals differs in each individual case, but several key mechanisms are involved. The most common of these mechanisms involves the active pumping of toxic metals out of the bacterial cell; e.g. cadmium, zinc, lead and nickel ions are actively removed from the cell, either by ATP or membrane potential driven pumps. Certain resistance mechanisms first enzymatically reduce or oxidise the metal ion involved. The resulting metal is then either transported out of the cell, e.g. the arsenic resistance operon encodes for such a system. In the case of mercury, this does not occur, as the resulting Hg^0 ion is volatile and naturally diffuses out of the cell (Hobman and Brown, 1998, Misra, 1992, Osborn *et al*, 1997, Silver, 1996, Silver and Phung, 1996).

Other heavy metal resistance mechanisms include sequestration of the metal, whereby the toxic metal binds to the bacterial cell wall, intracellular or specific binding components. Blocking the passage of the metal into the cell by altering the membrane transport systems is also common, as is the alteration of intracellular target

sites (Cooksey, 1994, Hobman and Brown, 1997, Silver and Phung, 1996, Wood and Wang, 1983).

1.3.2. Mercury compounds in the environment.

Mercuric compounds in natural environments can arise from a variety of natural and anthropogenic sources such as volcanic activity, mining operations and other industrial activities although the precise chemical speciation and concentration varies between different environments. (Hart *et al.*, 1998, Misra, 1992, Osborn, *et al.*, 1993). Bacteria are known to play an important role in the cycling of mercuric compounds in the environment. The presence of mercury compounds in a particular environment often correlates with an increase in the number of cultivable mercury resistant bacteria isolated from the environment along with an associated decrease in bacterial diversity (Hart *et al.*, 1998, Olson *et al.*, 1991, Osborn *et al.*, 1993, Osborn *et al.*, 1995, Rochelle *et al.*, 1991, Vaituzis *et al.*, 1975). The toxicity of mercurial compounds to bacteria arises from the affinity of these compounds for the thiol groups in bacterial proteins (Misra, 1992).

Several different mechanisms of bacterial mercury resistance have been described. Pah-Hou *et al.*, 1981, described the reduced uptake of Hg^{2+} ions by the decreased permeability of the membrane to Hg^{2+} ions. The enzyme mediated demethylation of methylmercury was also described by Pan-Hou and Imura, 1981. The most widely studied method for resistance to mercury is the reduction of divalent mercury. This has been characterised in a large number of both Gram-negative and Gram-positive bacteria from a wide range of environments (Brown, 1985, Brown *et al.*, 1991, Foster, 1987, Misra, 1992, Summers, 1986, Walsh *et al.*, 1988). The genes conferring this type of resistance mechanism are found on *mer* operons. Bacterial mercury resistance is commonly used as a model system for studying the mechanisms involved in the flux of genetic material during the adaptation of natural bacterial populations to selective pressures. Mercury resistance is used as it is a toxic environmental pollutant with a well characterised environmental cycle and the

resistance genes are carried on an operon, which is commonly found on both transposons and plasmids, therefore greatly increasing the potential for the transfer of these resistance genes (Barkay *et al.*, 1985).

1.3.3. Mercury resistance operons.

A large number of *mer* operons have been characterised in a wide range of bacterial genera including both Gram-negative and Gram-positive bacteria (Hobman and Brown, 1997, Osborn *et al.*, 1997, Silver, 1996). In Gram-negative bacteria the genes on the *mer* operon are closely related and have the same basic structure in most organisms. The transcriptional regulation of the *mer* operon is carried out by the protein coded for by the *merR* gene which is responsible for the activation of the operon. This 16 kDa protein activates the operon in the presence of Hg^{2+} and negatively regulates the transcription of both itself and the other genes in the operon in the absence of Hg^{2+} (Heltzel *et al.*, 1987, Lund *et al.*, 1986). The N-terminal end of the protein contains two helix-turn-helix DNA binding motifs and the C-terminal domain contains three cysteine residues that are involved in binding of Hg^{2+} ions. The MerR protein forms a dimer which is capable of binding one Hg^{2+} ion (Skewchuk *et al.*, 1989).

Once the operon is active, mercuric ions (Hg^{2+}) pass through the outer membrane of the cell where they are bound by the 6.5 kDa MerP protein. The Hg^{2+} ions are then transferred to the MerT protein (12.4 kDa), which contains three membrane spanning alpha helices. The Hg^{2+} ions are transported across the inner membrane into the cytoplasm of the cell where they are reduced to Hg^0 by the MerA protein (Brown, 1985, Ni'Bhriain and Foster, 1986). The MerA protein forms a dimer with a molecular weight of 112 kDa, containing two active site electron acceptors and a redox active disulphide residue. Once reduced by MerA, the volatile Hg^0 diffuses out of the cell due to its low vapour pressure (Hobman and Brown, 1997, Misra, 1992, Osborn *et al.*, 1997, Silver, 1996).

In addition to this basic "skeleton operon", several additional genes are also present in certain *mer* operons. The *merD* gene encodes for a 13.5 kDa regulatory

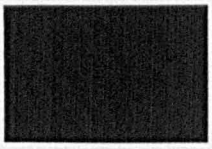
protein, which, like MerR, has an amino terminal helix turn-helix DNA binding motif. The *merC* gene is believed to encode for an additional transport protein responsible for the transport of Hg^{2+} ions into the cell, although the precise function of this gene is unclear. *merF*, which shows homology to *merT*, may encode a secondary mercury transport system. "Broad spectrum" resistance to organomercurial compounds such as phenyl mercuric acetate differs from the "narrow spectrum" resistance described above and resistance to these compounds is usually conferred out by the 23kDa protein coded for by the *merB* gene. This protein is an organomercurial lyase (Hobman and Brown, 1997, Osborn *et al.*, 1997). The *merG* gene is believed to be involved in reducing the permeability of the cell to phenyl mercury and has been found inserted between the *merA* and *merB* genes in plasmid pMR26, which was isolated from a *Pseudomonas* strain (Kiyono and Pan-Hou, 1999). In addition to these genes which have well characterised functions, several operons contain genes which have unknown functions. The operons contained on Tn*21*, Tn*501*, pMER419, pKLH2 and Tn*5053* contain 2 genes of unknown function known as orf I and orf II. The operons of Tn*3926*, Tn*5036*, Tn*5059* and pMER610 contain 3 genes designated urf1, urf2X and urf2Y (Kholodii *et al.*, 1997, Lett *et al.*, 1985, Yurieva *et al.*, 1997). In general, the position of these genes on the operon remains identical in different bacteria. The structures of several Gram-negative and Gram-positive *mer* operons is shown are Figure 1.3.3.1., which also shows the positions at which additional genes insert into the operon.

1.3.3.1. Alternative mercury resistance determinants.

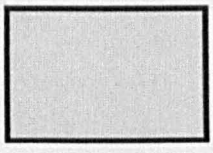
The *mer* operons of Gram-positive bacteria differ from those found in Gram-negative bacteria by virtue of both their sequence and their structural organisation. Unlike Gram-negative *mer* operons, Gram-positive *mer* operons are seldom found on plasmids (Beliveau and Trevors, 1990, Bogdanova *et al.*, 1998). Several mercury resistance determinants from Gram-positive bacteria have been sequenced to date, namely the operons from *Bacillus* sp. strain RC607 (Wang *et al.*, 1989), *Streptomyces*

Figure 1.3.3.1. Structures of *mer* operons in Tn*501*, Tn*21*, Tn*5053,* Tn*5036* and *Bacillus* strain RC607. Not to scale.

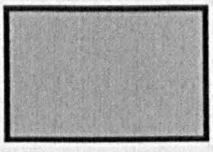
Key:



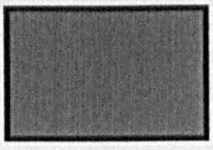
Regulatory gene, i.e. *merR* and *merD*.



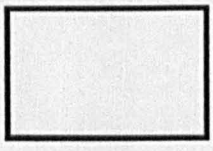
Transport gene, i.e. *merT*, *merP*, *merF* and *merC*.



Gene with unknown function, i.e. *orf 1-4*, *urf1, 2X* and *2Y*.



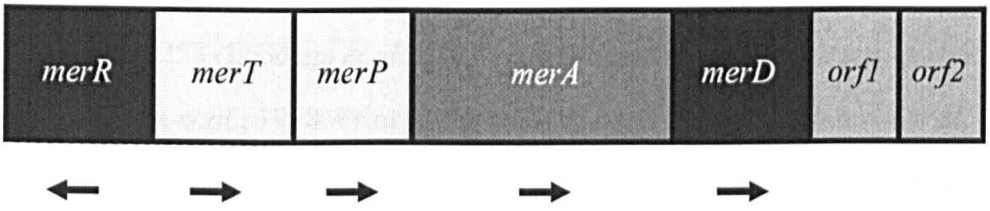
mercuric reductase gene (*merA*).



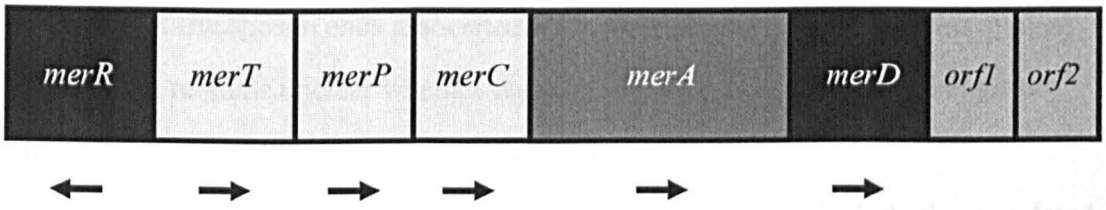
organomercurial lyase (*merB*).

Arrows indicate the direction of transcription.

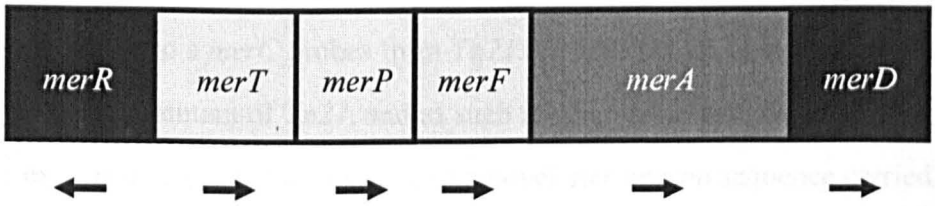
Tn*501*:



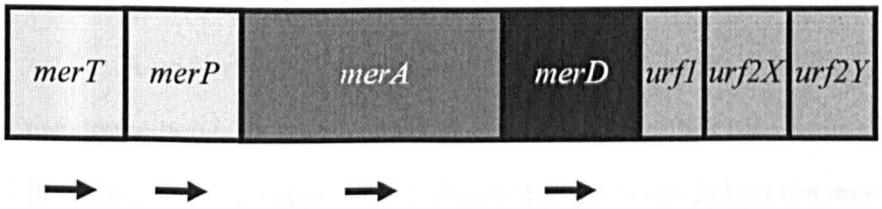
Tn*21*:



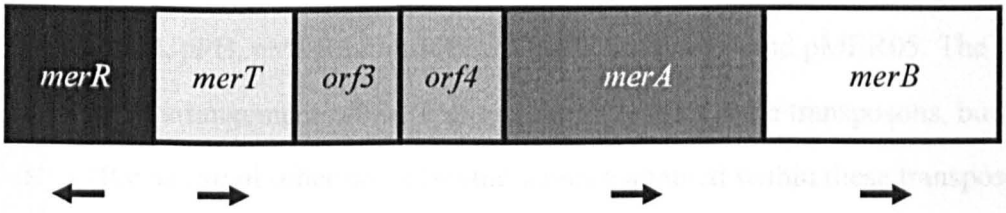
Tn*5053*:



Tn*5036*:



Bacillus strain RC607:



lividans strain 1326 (Sedlmeier and Altenbuchner, 1992), *Staphylococcus aureus* plasmid pI258 (Laddaga *et al.*, 1987) and the *mer* determinants sequenced by Bogdanova *et al.*, 1998. Prior to the study carried out by Bogdanova *et al.*, 1998, the association of class II transposons with *mer* genes had only been reported in Tn*4556*, which carries ORFs with low homology to *mer* genes, but does not confer resistance to mercury compounds on the host bacteria. Bogdanova *et al.*, 1998 characterised the presence of transposon ends associated with *mer* operons in three *Bacillus* strains, although it remains unclear whether these transposons contain a complete set of transposition genes or whether they are defective.

Gram-negative bacteria containing novel *mer* operons include those isolated from *Thiobacillus* (Inoue *et al.*, 1991) that has its *merR* gene spaced apart from the operon and pMERPH (Peters *et al.*, 1991) which carries a *mer* operon which does not hybridise to a *merC* probes from Tn*21*. pMERPH also failed to complement a *merR* defective mutant of Tn*21*, and as such is believed to contain no *merR* gene. This evidence suggests the presence of a novel *mer* operon sequence carried on pMERPH, which is rare in the natural environment.

1.3.4. Mercury resistance transposons.

A wide range of transposons have been described which contain mercury resistance genes although some of these contain either incomplete operons and/or non functional genes. Table 1.3.4.1 shows the genes carried on the *mer* operons in Tn*21*, Tn*501*, Tn*502*, Tn*2603*, Tn*3926*, Tn*5036*, Tn*5041*, Tn*5053*, Tn*5059*, Tn*5086*, Tn*5384*, Tn*5385*, Tn*5467*, and Tn*5563* as well as those unnamed transposons carried on pKLH2, pPB, pMER610, pMER327/419, pMER330 and pMER05. The table also shows the arrangement of the *tnpA* and *tnpR* genes in these transposons, but does not show the nature of other non-essential genes contained within these transposon structures. Sequence data was not available for all of these transposons. Where this is the case, it is indicated in Table 1.3.4.1. Certain mercury resistance transposons have not been characterised in detail, including Tn*2613*, Tn*3401*, Tn*3402* and Tn*3403* and

Table 1.3.4.1. Mercury resistance transposons.

Tn*21*: This indicates an arrangement of *tnpA/tnpR* identical to that found on Tn*21*.

Tn*3*: This indicates an arrangement of *tnpA/tnpR* identical to that found on Tn*3*.

NSA: No sequence data available.

* The presence/absence of a *tnpR* gene in this strain in unclear at present.

∞ This transposon does not utilise the *tnpA/tnpR* transposition system.

Ψ This transposon contains non functional *tnpA* and *tnpR* genes in the Tn*21*-like arrangement.

° Other *mer* genes may be present on these transposons.

§These transposons are believed to represent Tn*21* variants, with minor insertions and deletions.

Transposon	Mercury resistance	*tnpA/tnpR* arrangement	Genes present								Reference
			merR	*merT*	*merP*	*merA*	*merD*	*merC*	*merF*	*merB*	
Tn*21*	+	Tn*21*	+	+	+	+	+	+	-	-	Ward and Grinsted,1987
Tn*501*	+	Tn*21*	+	+	+	+	+	-	-	-	Brown *et al.*,1985b
Tn*502*	+	Tn*21*	+	+	+	+	+	-	-	-	Stanisich *et al.*,1989
Tn*1935*§	+	Tn*21*	+	+	+	+	+	+	-	-	Colonna *et al.*,1988
Tn*2410*§	+	Tn*21*	+	+	+	+	+	+	-	-	Kratz *et al.*,1983
Tn*2411*§	+	Tn*21*	+	+	+	+	+	+	-	-	Kratz *et al.*,1983
Tn*2603*§	+	NSA	NSA								Yamamoto *et al.*,1981
Tn*3926*	+	Tn*21*	+	+	+	+	+	+	-	-	Lett *et al.*,1985
Tn*5036*	+	Tn*21*	-	+	+	+	+	+	-	-	Yurieva *et al.*,1997
Tn*5041*	+	No *tnpR**	+	+	+	+	+	+	-	-	Yurieva *et al.*,1997
Tn*5053*	+	*tniA* system$^{\infty}$	+	+	+	+	+	-	+	-	Kholodii *et al.*,1995
Tn*5059*	+	Tn*21*	+	+	+	+	+	-	-	-	Kholodii *et al.*,1997
Tn*5086*	+	Tn*21*	NSA								Sundström *et al.*,1993
Tn*5384*	+°	NSA	+	-	-	+	-	-	-	+	Bonafede *et al.*,1997

Tn*5385*	+°	NSA	+	-	-	+	-	-	-	+	Rice and Carias., 1998
Tn*5467*	-	Tn*21*$^{\Psi}$	+	-	-	-	-	-	-	-	Clennel *et al.*,1995
Tn*5563*	-	Tn*3*	-	+	+	-	-	-	-	-	Yeo *et al.*,1998
pKLH2	+	NSA	+	+	+	+	+	+	-	-	Kholodii *et al.*,1993a
pMER610	+	Tn*21*	+	+	+	+	+	+	-	-	Kholodii *et al.*,1997
pMER327/419	+	NSA	+	+	+	+	+	-	+	-	Hobman *et al.*,1994
pMER330	+	NSA	+	+	+	+	+	-	+	-	Hobman *et al.*,1994
pMER05	+	NSA	+	+	+	+	+	-	+	-	Hobman *et al.*,1994
pPB	+	Tn*21*	+	+	+	+	+	+	-	+	Reniero *et al.*,1998

as such are not included in this list (Radford *et al.*, 1981, Tanaka *et al.*, 1983). In those instances where sequences of the *tnpA* or *tnpR* genes was not available, the location of the *mer* genes on a transposon structure has usually been deduced by the sequencing of IRs similar to that of a known transposon adjacent to the sequenced *mer* genes (Brown *et al.*, 1985b, Bonafede *et al.*, 1997, Clennel *et al.*, 1995, Hobman *et al.*, 1994, Kholodii *et al.*, 1993a, Kholodii *et al.*, 1995, Kholodii *et al.*, 1997, Lett *et al.*, 1985, Reniero *et al.*, 1998, Rice and Carias, 1998, Stanisich *et al.*, 1998, Sundström *et al.*, 1993, Ward and Grinsted, 1987, Yamamoto *et al.*, 1981, Yeo *et al.*, 1998,Yurieva *et al.*, 1997). Transposons carrying *mer* operons have been discovered in a wide range of environments including clinical specimens, soil, sediment, milk, water and the guts of animals (Lett *et al.*,1985, Jobling *et al.*, 1988a, Rice and Carias, 1998, Yeo *et al.*,1998, Yurieva *et al.*,1997).

1.3.5. Plasmid-borne transposons.

Due to the potentially mobile nature of transposons, they are found at both chromosomal and plasmid sites in the bacterial cell. The potential for transposon mobility is increased if the transposon is present on a conjugative or mobilisable plasmid. Transposons such as Tn*21* and Tn*501* have been found on plasmids, as have a wide variety of others including the mercury resistance transposons Tn*502*, Tn*2603*, Tn*5036*, Tn*5059*, Tn*5086*, Tn*5467* and Tn*5563* (Brown *et al.*, 1985a, Brown *et al.*, 1985b, Clennel *et al.*, 1995, Stanisich *et al.*, 1998, Sundström *et al.*, 1993, Ward and Grinsted, 1987, Yamamoto *et al.*, 1981, Yeo *et al.*, 1998,Yurieva *et al.*, 1997). The plasmids carrying transposons differ, not only in the transposon that they carry, but also in their conjugative/mobilisable ability, their size and the other genes which they carry (Dahlberg *et al.*, 1997, Jobling and Ritchie, 1987). Information regarding the distribution of plasmids, their evolutionary relationships and the diversity of these plasmids in relation to natural selective pressures is needed to understand the role of plasmids in the flow of genetic information in natural bacterial communities.

1.4. Methods for studying genetic diversity.

The study of genetic diversity in the natural environment requires the extraction of genetic material, either RNA or DNA. This extraction can be either from bacteria or directly from the environment itself (Pearson *et al.*, 1996, Hart *et al.*, 1998, Hermansson *et al.*, 1987).

1.4.1. Isolation of DNA from bacterial isolates.

A number of different protocols can be used for the initial selection of bacterial isolates, depending on the individual characteristics which are to be studied, e.g. mercury resistant bacteria are commonly isolated by selective growth on media containing $HgCl_2$ (Barkay *et al.*, 1985, Barbieri *et al.*, 1996, Hart *et al.*, 1998, Osborn *et al.*, 1993, Radford *et al.*, 1991, Rochelle *et al.*, 1991, Trajanovska *et al.*, 1997). Following this, a wide range of different methods for the extraction of genomic DNA, plasmid DNA, or RNA from the bacteria are available (Ampe *et al.*, 1998, Bertin, 1995, Birnboim and Doly, 1979, Domenico *et al.*, 1992, Olsen., 1990, Pitcher *et al.*, 1989, Trevors, 1985, Wheatcroft and Williams, 1981). The choice of method will depend on the quality and quantity of DNA/RNA required for analysis, as well as the nature of the bacteria being studied. Once suitable material has been obtained, a number of different approaches for the study of diversity exist.

1.4.1.1. Methods for analysis of bacterial DNA.

Most of the molecular methods for studying genetic diversity at the DNA level involve an initial polymerase chain reaction (PCR) step. PCR utilises DNA primers which are specific for the gene to be studied, and allows the amplification of a large amount of DNA from that region. DNA/DNA hybridisation experiments are commonly used to ascertain whether or not a particular gene is present in the DNA sample by using DNA probes for that particular gene. The presence of *mer* genes in bacteria isolated from polluted and pristine environments has been studied using this

technique (Barkay *et al.*, 1985, Peters *et al.*, 1991, Rochelle *et al.*, 1991). Restriction Fragment Length Polymorphism (RFLP) analysis of PCR products using restriction endonuclease enzymes is commonly used to classify the PCR products from different strains of bacteria by virtue of their DNA sequence (Hart *et al.*, 1998, Osborn *et al.*, 1993, Pearson *et al.*, 1996). RFLP reactions using PCR products from bacterial 16S rDNA genes are common and this has been designated "amplified ribosomal DNA restriction analysis" (ARDRA) (Smit *et al.*, 1997). An alternative to RFLP analysis is DNA sequencing, whereby the sequence of the PCR product is determined. Depending on the length of the PCR product this may require additional sub-cloning into a suitable vector in order to allow the whole sequence to be determined. DNA sequencing yields more information than RFLP analysis, as RFLP analysis is dependant upon the choice of enzyme used, i.e. an accurate indication of diversity may not be produced if certain restriction enzymes are used. The amino acid sequence of the PCR product can also be ascertained if the DNA sequence is known, a distinct advantage over RFLP analysis.

1.4.1.1.1 Phylogenetic trees.

In order to allow the comparison of RFLP patterns or sequencing data from different bacterial isolates, phylogenetic analysis is commonly used (Gray and Herwig, 1996, Head *et al.*, 1998, Osborn *et al.*, 1993, Pearson *et al.*, 1996). This involves using data to construct phylogenetic trees which show the relationships between the different data sets in a visual form. A wide range of computer programs exist for the construction of such trees and the choice of program is dependant on the nature of the data available and the type of phylogenetic tree which is required (Felsenstein, 1981, Fitch, 1971, Fukami and Tateno, 1991, Hillis *et al.*, 1994, Hillis, 1997, Page, 1996, Tateno *et al.*, 1994).

There are two basic types of phylogenetic tree, rooted and unrooted. Rooted trees allow the comparison of different data sets and root these data to an imaginary evolutionary origin which is determined by the computer program. Unrooted trees

however only compare the data from the different sets to each other. The choice of program depends on the type of analysis which is required (Felsenstein, 1981, Fitch, 1971, Hillis *et al.*, 1994, Hillis, 1997). Within these two subsets there exists a wide range of different programs using different conditions and assumptions to produce phylogenetic trees, although the actual trees produced do not differ greatly (Holt *et al.*, 1996).

1.4.1.1.2. Plasmid diversity.

The diversity of plasmids isolated from the natural environment can be studied using a number of methods. Plasmid classification has traditionally involved incompatibility group (Inc) typing by PCR or hybridisation; however recent studies have suggested that mercury resistance plasmids from the natural environment do not conform to the standard Inc groups which were predominantly characterised during the study of plasmids isolated from clinical samples (Dahlberg *et al.*, 1997, Smit *et al.*, 1998, Wilmotte *et al.*, 1996). More recently, plasmids have been isolated from the environment using exogenous isolation whereby plasmids are extracted from the environment by the addition of a recipient bacterial strain which "captures" conjugative plasmids directly from the environment. This recipient strain is then isolated from the environment by virtue of antibiotic resistance markers carried on its chromosome. This approach allows plasmids that are active in the environment to be studied (Bale *et al.*, 1988, Smit *et al.*, 1998, Top *et al.*, 1994). The level of bacterial conjugation within the environment has been studied by the labelling of plasmids with green fluorescent protein (*gfp*). This system is sensitive enough to detect a single conjugative event (Dahlberg *et al.*, 1998a, Dahlberg *et al.*, 1998b). Other methods of studying diversity include size analysis and restriction analysis using restriction endonucleases (Jobling *et al.*, 1988b, Lilley *et al.*, 1996).

1.4.2. Isolation of DNA directly from the environment.

The isolation of DNA directly from the environment ("total" DNA) has both advantages and disadvantages compared to the methods involving cultivation

described previously (Bruce *et al.*, 1995, Leff *et al.*, 1995, Kuske *et al.*, 1998, Ogram *et al.*, 1987, Pilia *et al.*, 1991, Rosado *et al.*, 1997, Yeates *et al.*, 1997). It is estimated that only 0.1-1% of the bacteria in soil are cultivable under laboratory conditions (Bruce *et al.*, 1995, Head *et al.*, 1998, Rosado *et al.*, 1997, Woese, 1996). The extraction of DNA from the environment should allow the DNA from "non cultivable" species to be characterised. However, a number of problems exist with "total" DNA analysis. Firstly, due to the nature of the material obtained, PCR of the DNA is usually required to increase the amount of DNA available for study. As there may be more than one variant of the gene present in the environment the possibility for the production of PCR chimeras exists, which may lead to a false impression of diversity (Wang and Wang, 1997). PCR bias may also be a problem, i.e. if one particular variant is common in the sample, this may be amplified preferentially over some of the less common variants, which may remain uncharacterised (Polz and Cavanaugh, 1998). The nature of certain environments, such as soil and sediment, may prevent PCR amplification from occurring, i.e. soil contains high levels of humic acids, which are similar in their chemical nature to DNA and will therefore be coextracted with the DNA. The extraction protocol must therefore be altered to allow removal of these inhibitory substances to a suitably low level in order to allow PCR amplification to occur (Wilson, 1997). Studies using "total" DNA must involve adequate tests and control reactions to minimise these effects and to ensure that the observed diversity is not due to the experimental procedure.

1.4.2.1. Methods for analysis of "total" DNA.

Once DNA has been isolated from the environment, it can be analysed in a number of ways. Hybridisation has been used to compare the similarity of soils; this involves digesting the extracted DNA followed by subsequent labelling with a random primed reaction. This analysis was used to show that the similarity of four agricultural soils was between 35 and 75% (Griffiths *et al.*, 1996). Cot curves have also used to

analyse DNA extracted directly from soil (Torsvik *et al.*, 1990). RFLPs and DNA sequencing can also be carried out as previously described, although these methods can be time consuming, as they require the PCR products to be cloned into a suitable vector so that the individual variants of the mixed PCR products can be separated. To avoid cloning procedures, denaturing gradient gel electrophoresis (DGGE) has been successfully used to allow the characterisation of mixed PCR products (Head *et al.*, 1998, Muyzer and Smalla, 1998, Øvreäs *et al.*, 1998, Torsvik *et al.*, 1998). This utilises an initial PCR reaction using modified primers, one of which contains a 40 bp GC clamp, which is designed to anneal to itself. This PCR product is then electrophoresed on an acrylamide gel containing a gradient of denaturant, which denatures the PCR product depending on its GC content. Once denatured the PCR product is held in place on the gel by the GC clamp. This procedure allows the rapid separation of a large number of samples on the basis of their DNA sequence. To date, this procedure has mostly been used for the study of 16s RNA genes and, with one notable exception, has not been used for the study of functional genes (Muyzer and Smalla, 1998).

Other methods for studying the genetic diversity of an environment include single strand conformation polymorphisms (SSCP) which involves PCR of "total" DNA, followed by heating of the sample and subsequent reannealing to ascertain the diversity of a particular sample. This is observed by the incorrect reannealing of different variants of the PCR product producing different running forms when electrophoresed on a gel (Lee *et al.*, 1996, Schweiger and Tebbe, 1998). Fluorescent in situ hybridisation (FISH) uses DNA probes tagged with a fluorescent molecule to study bacteria in their environment using microscopy without relying on PCR reactions (Torsvik *et al.*, 1998). Analysis of mRNA isolated from the natural environment is difficult as the mRNA degrades quickly and is often present at relatively low levels. RNA analysis will allow functional genes in a population to be studied and will provide a mechanism whereby the active variants of a particular gene can be identified (Duarte *et al.*, 1998).

1.4.3. Aims of project.

This project is concerned with the study of the genetic diversity of transposon associated genes in a collection of 39 mercury resistant Gram-negative bacteria which were isolated from a number of polluted and pristine sites. These isolates have been studied in detail with respect to the diversity of their *mer* operons, as has DNA extracted from the isolation sites (Bruce *et al.*, 1992, Bruce *et al.*, 1995, Osborn *et al.*, 1993, Osborn *et al.*, 1995). The diversity of the *tnpA* and *tnpR* genes in 30 of these isolates was previously determined by RFLP analysis (Pearson *et al.*, 1996). This study showed the presence of 3 classes of *tnpR* gene and 6 classes of *tnpA* gene in the isolates. There was no observed linkage between the two *tnp* genes, or between these and the *mer* genes, suggesting that extensive recombination has occurred between the *tnp* genes and the *mer* genes in these isolates.

The first aim of this study was to characterise the remaining nine T2 isolates which were not studied by Pearson *et al.*, 1996, by PCR and DNA sequencing. The DNA sequence of the *tnpR* genes from members of each RFLP class was determined and phylogenetic trees were produced to allow the comparison of the data from the different isolates. This study allowed the comparison of two different approaches for the study of genetic diversity, namely RFLP and DNA sequencing.

The second aim of this study was to determine the structural diversity of the *tnp* and *mer* genes in these 39 strains, together with 85 new strains isolated from the same sites as the original 39 isolates. The intention was to determine whether there was any correlation between the structural arrangement of the genes on the transposon and the RFLP/DNA sequence data previously obtained (Pearson *et al.*, 1996). In this study, the relative arrangement of the *tnpA* and *tnpR* genes was determined in the 124 Gram-negative isolates to observe if any Tn*3*-like arrangements of *tnpA* and *tnpR* genes were associated with *mer* genes. All previous mercury resistance transposons had contained Tn*21*-like arrangements of these genes, with the notable exception of Tn*5563*, which contains *merT* and *merP* genes associated with a Tn*3*-like arrangement of *tnpA and tnpR* genes (Yeo *et al.*, 19980. However, this transposon does not confer

resistance to mercury on the host micro-organism (Yeo *et al.*, 1998). The approximate size of the plasmids contained within the 124 isolates, together with the location of the *tnpA* gene was determined, as was the presence of integron structures within the isolates and the size of their inserted gene cassettes. The presence of genes between *merP* and *merA* was determined to ascertain whether *merF* or *merC* genes were present in the *mer* operon. Finally, PCR reactions using primers for *tnp* and *mer* genes were used to ascertain the relative location of the *mer* operon and the *tnp* genes in the isolates.

CHAPTER 2.

MATERIALS AND METHODS.

2.1. Bacterial strains.

A list of all the strains used in this study is shown in Table 2.1.1.

2.2. Sites for isolation of '93 and '96 strains.

The thirty-nine '93 strains used in this study were previously cultured from samples taken at polluted and pristine sites (Osborn *et al.*, 1993) using the method described in Section 2.5. Thirty nine strains were isolated on the basis of their mercury resistant phenotype and their ability to hybridise to a *mer*RTΔP probe from Tn*501* (Osborn *et al.*, 1993). The identity of the '93 strains was determined by API analysis and the results are shown in Table 2.2.1. The eighty-five '96 strains were isolated in this study on the basis of their mercury resistant phenotype alone. All 124 strains were Gram-negative. The sites used for bacterial isolation were as follows:

SO strains were isolated from soil at a mercury polluted site at Fiddlers Ferry, Merseyside. Sediment at this site was used to isolate the SE strains. This site is shown in Figure 2.2.1 and 2.2.2. and is situated near Fiddlers Ferry power station, Merseyside. SB strains were isolated from soil at pristine site at Salterbrook Bridge in the Peak District National Park (Figure 2.2.3.) The '93 T2 strains were isolated from soil taken at a disused copper mine containing high levels of mercury in Tipperary, Eire (no photograph available). Soil samples were routinely taken from the top 2 inches of soil at the appropriate site. This soil was then homogenised by shaking and bacteria were then isolated (Section 2.5). Levels of mercury pollution at the Fiddlers Ferry and Salterbrook Bridge sites have been previously determined by the Centre for Analytical Research in the Environment, Imperial College, Ascot:

Table 2.1.1. Bacterial strains used in this study:

† *E.coli* Genetic Stock Centre: http://cgsc.biology.yale.edu
* Isolation of these strains is described in Section 2.2 and 2.5.

Strain	Source/Reference	Characteristics
Tn*21*	*E.coli* Genetic Stock Centre†	pACYC184::Tn*21*
Tn*501*	*E.coli* Genetic Stock Centre†	pACYC184::Tn*501*
CGSC# 6217	*E.coli* Genetic Stock Centre†	zia-2006::Tn*3*
UWC1	Lilley *et al.*, 1996	
XL1-BLUE	Stratagene	
CGSC# 6198	*E.coli* Genetic Stock Centre† (Guyer et al, 1981)	KL (3X)::Tn*1000*, S^-
'93 collection*	Isolated from polluted and pristine soil (Osborn *et al.*, 1993)	39 Hg^R Gram-negative bacteria
'96 collection*	Isolated from polluted and pristine soil	85 Hg^R Gram-negative bacteria
Lm74	Lebrun *et al.*, 1994	*Listeria monocytogenes* carrying Tn*5422*
MKD3601	Kholodii *et al.*, 1997	pUC19::Tn*5036*
MKD1605	Kholodii *et al.*, 1997	RP1::Tn*5041*
G122	Ulrich and Pühler, 1994	*Rhizobium* carrying Tn*163*
PaW85	Mäe and Heinaru, 1994	*Pseudomonas putida* carrying Tn*4652*
DS16	Shaw and Clewell, 1985	*Enterococcus facaelis* carrying pAD1::Tn*917*
BT1	Lereclus *et al.*, 1986	*Bacillus thuringiensis* kurstaki carrying Tn*4430*
BM4147	Arthur *et al.*, 1993	*Enterococcus feacium* carrying Tn*1546*

Table 2.2.1. API identification of the '93 collection of mercury resistant bacterial strains. These strains were isolated by Osborn *et al.*, 1993. Table is taken from Osborn *et al.*, 1993.

Isolate	API identification	Level of certainty for identification*
T2 7	*Acinetobacter calcoaceticus*	Good
T2 12	*Aeromonas hydrophila*	Very good
T2 13	*Aeromonas hydrophila*	Very good
T2 17	*Agrobacterium radiobacter*	Very good
T2 19	*Aeromonas hydrophila*	Good
T2 23	*Aeromonas hydrophila*	Excellent
T2 37	*Aeromonas salmonicida*	Acceptable
T2 38	*Enterobacter aerogenes*	Very good
T2 41	*Aeromonas hydrophila*	Very good
T2 46	*Aeromonas hydrophila*	Excellent
SE3	*Pseudomonas testosteroni*	Acceptable
SE6	*Acinetobacter calcoaceticus*	Good
SE9	*Alcaligenes faecalis*	Good
SE11	*Acinetobacter calcoaceticus*	Good
SE12	*Acinetobacter calcoaceticus*	Good
SE18	*Pseudomonas testosteroni*	Good
SE20	*Alcaligenes faecalis*	Good
SE23	*Pseudomonas testosteroni*	Good
SE31	*Klebsiella oxytoca*	Very good
SE35	*Alcaligenes faecalis*	Good
SO1	*Enterobacter cloacae*	Excellent
SO2	*Enterobacter cloacae*	Excellent
SO3	*Enterobacter cloacae*	Good
SO5	*Enterobacter cloacae*	Excellent
SO6	*Enterobacter cloacae*	Excellent
SO7	*Enterobacter cloacae*	Excellent
SO8	*Enterobacter cloacae*	Excellent
SO9	*Enterobacter cloacae*	Excellent
SO12	*Pseudomonas* sp.	Good
SO13	*Enterobacter cloacae*	Good
SB2	*Pseudomonas* sp.	Good
SB3	*Pseudomonas* sp.	Good
SB4	*Pseudomonas fluorescens*	Good
SB5	*Pseudomonas* sp.	Good
SB8	*Pseudomonas* sp.	Good
SB12	*Pseudomonas* sp.	Good
SB13	*Pseudomonas* sp.	Good
SB22	*Pseudomonas* sp.	Good
SB24	*Pseudomonas* sp.	Good
SB29	*Pseudomonas* sp.	Good

* Reliability of identification classifications are as follows: Excellent, 99.9%; Very good, 99%; Good, 90%; Acceptable, 80%.

Figure 2.2.1. Photograph of Fiddlers Ferry sampling site. Site indicated by arrow.

Figure 2.2.2. Photograph of Fiddlers Ferry sampling site. Site indicated by arrow.

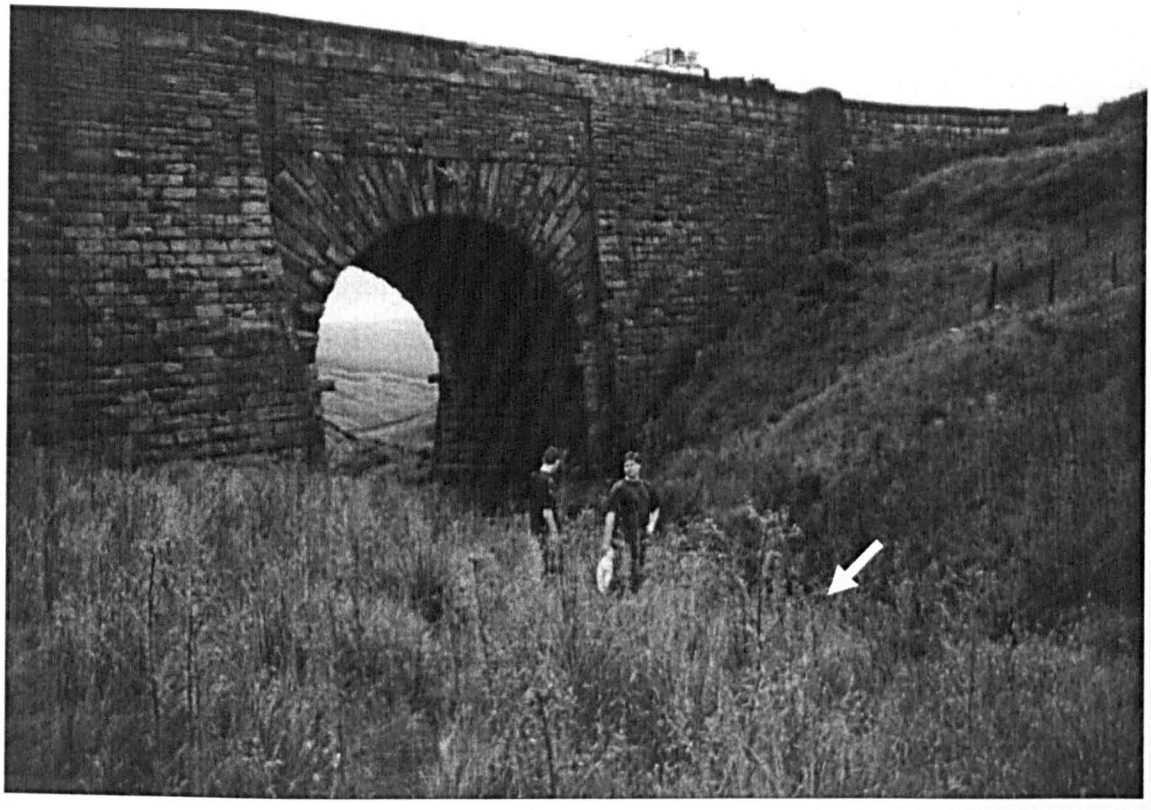


Figure 2.2.3. Photograph of Salter Brook Bridge sampling site. Site indicated by arrow.

Mercury concentration at the SO, SE and SB sampling sites:

SO: 0.441 ppm ± 0.039

SE: 0.161 ppm ± 0.029

SB: <0.12 ppm. (Limit of detection <0.12 ppm).

T2: Data unavailable.

2.3. Media.

Luria Bertani agar (LA) was used for bacterial growth on solid media and Luria Bertani broth (LB) was used for bacterial growth in liquid culture. Both were autoclaved prior to use:

LB: 10g l^{-1} Bactotryptone (Difco)
5g l^{-1} Bactoyeast extract (Difco)
10g l^{-1} NaCl.
To pH 7.5 with 7.5M NaOH.

LA: As for LB with 15g l^{-1} Agar No2 (Amersham).

2.3.1. Selective reagents.

Antibiotic, heavy metal and other supplements (supplied by Sigma, unless stated) were added where needed at the following concentrations; ampicillin (100μg ml^{-1}), methicillin (40μg ml^{-1}), tetracyclin (20μg ml^{-1}), $HgCl_2$ (50μM, AnalaR). X-Gal and IPTG (Metford Laboratories Ltd) were also used at final concentrations of 52.5μg ml^{-1} and 40μg ml^{-1} respectively.

2.4. Solutions.

Chemicals were purchased mostly from Sigma, Fisons and BDH Laboratory Supplies. Unless otherwise stated, reagents were General Purpose Grade. All solutions were made according to Sambrook *et al.*, 1989 unless indicated below.

- Extraction buffer (Osborn *et al.*, 1993) (Section 2.5):

  50mM Tris-HCl, pH 7.5.

- Extraction buffer (Bruce *et al.*, 1992): 1% SDS

  0.12M Na_2HPO_4 (pH 8.0)

- 10X TBE: 108.0 g l^{-1} Tris (hydroxymethyl) methylamine

  55.0 g l^{-1} Orthoboric acid

  7.4 g l^{-1} EDTA.

- 20X SSC: 2M NaCl, 0.3M tri-sodium citrate.

- GET: 50mM glucose, 10mM EDTA, 25mM Tris-HCl pH8.0.

- TE: 10mM Tris-HCl, pH8, 1mM EDTA pH8.0.

- GES: 60% Guanidium thiocyanate, 0.5% sarcosyl, 0.1M EDTA (Pitcher *et al.*, 1989).

- 25% Ringers:

  250 mg l^{-1} NaCl

  150 mg l^{-1} KCl

  120 mg l^{-1} $CaCl_2$

  50 mg l^{-1} $NaHCO_3$

- Phenol/chloroform/isoamyl alcohol. These were mixed in a ratio of 25:24:1.

- 1X TAE: 40mM Tris (hydroxymethyl) methylamine

  20mM Acetic Acid

  1mM EDTA (pH 8.3).

- Wheatcroft and Williams solution (Wheatcroft and Williams, 1981):

  5% (v/v) Antifoam A (Sigma)

  0.1mg ml^{-1} xylene cyanole in TE, pH 8.

- Type II DNA loading buffer: 0.25% bromophenol blue

  0.25% xylene cyanole

  25% Ficoll 400 in SDW.

- Prehybridisation buffer (2.12.4): 8.46ml of SDW

  1g of dextran sulphate

  1ml of 10% SDS, this was heated at 56°C for 30 minutes, at which time 0.58g of NaCl was added and the buffer incubated at 56°C for a further 15 minutes.

- Chloroform/isoamyl alcohol: 20 ml chloroform

  0.8 ml isoamyl alcohol.

- Crystal violet: 0.8% (w/v) ammonium oxalate

  20% (w/v) crystal violet

  20% (v/v) ethyl alcohol.

- Grams iodine: 0.33% (v/v) iodine

  0.66% (v/v) potassium iodide.

- Olsen *et al.*, 1990 solutions:

  Solution I: 50 mM Tris

  10 mM EDTA (pH 8)

Solution II: 3% SDS

50 mM Tris (pH 12.45)

- Birnboim and Doly, 1979 solutions:

Solution I: 50mM glucose

10mM EDTA

25mM Tris-HCl (pH8).

Solution II: 0.2M NaOH

1% SDS.

Solution III: 3M Sodium acetate, pH4.8.

2.5. Strain isolation from soil.

The method for isolating of bacteria from soil and sediment has been previously described (Osborn *et al.*, 1993):

Ten grams of soil was added to 90 ml of Extraction buffer (Section 2.4). This mixture was blended (using a Braun food blender) at room temperature for 1 minute, then stored on ice for 1 minute. This cycle was repeated three times, at which time a serial dilution of the soil suspension was carried out using 25% Ringers solution. Spread plates (LA + 50μM $HgCl_2$) were incubated at room temperature until bacterial colonies had grown (normally 1-2 days).

2.6. Growth of bacterial strains.

Typically, bacteria isolated from soil were grown at room temperature (approx. 22 °C), usually overnight, or longer if this was required. Laboratory *E.coli*

strains were incubated overnight at 37 °C. Liquid cultures in LB were carried out in a shaking incubator at the appropriate temperature.

2.7. Gram staining.

Gram staining of bacterial cultures was carried out using material from plate (LA) cultures. Control strains were used to confirm the validity of Gram-staining reactions. A loopful of sterile water was placed in the centre of a microscope slide and a small quantity of bacterial cells mixed into this using a sterile loop. The cells were then heat fixed onto the slide, which was then flooded with Crystal violet for 30 seconds, then rinsed using tap water. Grams iodine was added to the slide for 1 minute, and rinsed with water. Acetone was added to the slide for 1-2 seconds and immediately washed off with water. 0.5% (w/v) Safranin was then added to the slide for 30 seconds, washed off with water and the slide blotted dry before examination under a microscope. Gram-positive organisms appeared purple, whereas Gram-negative organisms appeared pink.

2.8. Isolation of plasmid DNA.

A variety of plasmid isolation methods were employed throughout this study (Bertin, 1995., Birnboim and Doly, 1979, Holmes and Quigley, 1981, Kado and Lui, 1981., Olsen *et al.*, 1990, Wheatcroft and Williams, 1981). The Olsen, 1990 and the Wheatcroft and Williams, 1981, were the most commonly used methods due to their ability to extract large and small plasmids in sufficient quantities for visualisation and Southern blotting. For extraction of small plasmids such as cloning vectors, the method of Birnboim and Doly, 1979 was used.

2.8.1. Olsen *et al.*, 1990.

Isolation of plasmid DNA for use in PCR was carried out using this technique, as the Wheatcroft and Williams, 1981, method yields DNA which is often unsuitable for PCR due to the high amounts of phenol/chloroform present in some samples:

A microcentrifuge tube was filled with 1.5 ml of an overnight culture, which was spun at 13000 g at room temperature in a microcentrifuge for five minutes at which time the supernatant was removed and the cell pellet was re suspended in 40µl of Solution I. Four hundredµl of Solution II was added to the tube, which was inverted to mix the contents. This was incubated at 56°C for 30 minutes and 300µl of 1.5 M potassium acetate, pH 5.2 was added and the tube inverted to mix the contents. This was then placed on ice for 20 minutes then spun at 13000 g at room temperature for 10 minutes.

The resulting supernatant was added to 600µl of ice cold isopropyl alcohol and left at room temperature for 15 minutes. This was spun at 13000 g at room temperature for a further 10 minutes, the supernatant removed, and the pellet re suspended in 200µl TE. Two hundredµl of phenol/chloroform (1:1, unbuffered) was added, the tube inverted and spun at 13000 g at room temperature for 10 minutes. The upper layer (150µl) was then transferred to a new microcentrifuge tube, 100µl 7.5 M ammonium acetate and 600µl 100% ethanol were added, the tube inverted and left on ice for 30 minutes. This was spun at 13000 g at room temperature for 6 minutes and the supernatant was removed. The DNA pellet was then dried and re suspended in 100 µl TE.

Plasmid DNA was visualised on 0.7% TAE agarose gels, run at 40V for approximately 24 hours followed by visualisation using ethidium bromide staining (1 µg ml^{-1}) (Section 2.16).

2.8.2. Wheatcroft and Williams, 1981.

This method of plasmid DNA isolation was used to detect plasmids in bacterial strains. This was due to the larger amount of samples which may be processed in a set time period using this method compared to that of Olsen *et al.*, 1990. This method was also used to prepare plasmid DNA for Southern hybridisations as the quantity of DNA produced was higher than that of the Olsen *et al.*, 1990 method.

One ml of LB was inoculated with the bacterial culture to be studied and incubated overnight. The resulting culture was placed in a microcentrifuge tube and spun at 13000 g at room temperature for 5 minutes to pellet the cells. After removal of the supernatant, the pellet was re suspended in 100µl of Wheatcroft and Williams solution. Twentyµl of 1M NaOH saturated with SDS was added and the tube inverted 20 times and vortexed for 1 minute. At this time 20µl of the sample was loaded onto an agarose gel to check for the presence of DNA. Eightyµl of phenol/chloroform (1:1, unbuffered) was added to the remaining reaction mix. This was vortexed for 5 seconds and centrifuged at 13000 g at room temperature for 5 minutes, at which time the top layer in the tube, which contained the plasmid DNA was transferred to a new microcentrifuge tube. The DNA was stored at -20 °C.

2.8.3. Birnboim and Doly, 1979.

A plasmid extraction technique based on the method described by Birnboim and Doly, 1979 was used in the preparation of plasmid DNA from transformants. This method is suitable for the extraction of the small plasmids used in cloning techniques and yields DNA suitable for PCR and long term storage:

A 1.5 ml microcentrifuge tube was filled with an overnight culture (LB) and spun at 13000 g at room temperature for 3 minutes to pellet the cells. After removal of the supernatant, the pellet was re suspended in 100µl of ice cold Solution I and mixed by vortexing. 200µl of freshly made Solution II was then added, the tube inverted 5

times, 150µl of Solution III added and the tube inverted to mix the contents. The tube was then kept on ice for 15 minutes and spun at 13000 g at room temperature for 5 minutes. The supernatant (650µl) was removed and transferred to a new microcentrifuge tube, an equal volume of unbuffered phenol/chloroform added, the tube mixed by vortexing and spun at 13000 g at room temperature for 3 minutes. The top layer of liquid in the tube was placed in a new microcentrifuge tube and 390µl of isopropyl alcohol added. This was kept at -20 °C overnight and then spun at 13000 g at room temperature for 5 minutes. After removal of the supernatant, the DNA pellet was washed in 1ml 70% ethanol and spun at 13000 g at room temperature for 5 minutes. Following removal of the supernatant, the pellet was dried and re suspended in 50µl TE, pH8.

2.9. Isolation of genomic DNA.

Genomic DNA was isolated from overnight bacterial cultures using two methods: A boiling method was used for the isolation of PCR templates (Güssow and Clackson, 1989, Osborn *et al.*, 1993). The method of Pitcher *et al*, 1989, was used in the isolation of DNA for restriction digests and for long term storage of DNA samples.

2.9.1. Güssow and Clackson, 1989., Osborn *et al.*, 1993.

This method allows fast extraction of DNA from overnight cultures, suitable for use as DNA template in conventional PCR. This DNA, however cannot be stored for long periods of time (>6hrs) as samples begin to degrade after this time:

Two hundredµl of an overnight culture were added to a 0.5 ml microcentrifuge tube and spun at 13000 g at room temperature for 3 minutes to pellet the cells. This pellet was re suspended in 200µl of SDW and placed in a boiling water bath for 10 minutes. The tube was spun at 13000 g at room temperature for 3 minutes and 50µl of the supernatant was used as a PCR template.

2.9.2. Pitcher *et al.*, 1989.

The DNA produced using this method is suitable for long term storage (4°C) and can also be used as a template for PCR. This method also yields much higher quantities of DNA:

A pellet of cells was first obtained, either from a plate culture re suspended in 200µl TE or from an overnight liquid culture. Following centrifugation at 13000 g at room temperature for 5 minutes, the cell pellet was re suspended in 200µl of TE and 500µl of GES was added and the tube was vortexed to mix the contents. The tube was placed on ice for 10 minutes and 250µl of 7.5 M ammonium acetate added. The contents of the tube were mixed by vortexing and placed on ice for 10 minutes. Following the addition of 750µl chloroform/isoamyl each sample was vortexed for 2 minutes and spun at 13000 g at room temperature for 20 minutes. The top 600µl of each sample was then added to 0.54 vols. of isopropyl alcohol, the tube inverted for 1 minute, and spun at 6000 g at room temperature for 30 seconds. The supernatant was removed and the pellet washed 3 times in 1 ml 70% ethanol. The resulting pellet was dried and re suspended in 100 µl SDW. Samples were stored at 4°C.

2.10. Extraction and purification of DNA from soil.

DNA was extracted from soil and sediment using the method described by Bruce *et al.*, 1992. This DNA was then treated using the Wizard® DNA Clean-Up System. (Promega). This approach yields high quantities of DNA which is of suitable quality to allow PCR reactions to be carried out.

2.10.1. Extraction of DNA from soil, Bruce *et al.*, 1992.

A 2g sample of soil/sediment was added to 2ml of glass beads (0.17-0.18 mm diameter, B. Braun Biotech International) and 2ml of extraction buffer. This was bead beaten for 30 seconds (B.Braun Biotech International 853034 bead beater) and the resulting solution was washed out with a further 5ml of extraction buffer. This was centrifuged at 2800g for 10 minutes, the supernatant removed and stored on ice. A

further 5ml of extraction buffer was then added and the solution centrifuged as before. The resulting supernatant was also kept on ice. This process was repeated with a further 5 ml of buffer at which time the 3 supernatant solutions were pooled, stored on ice for 10 minutes and centrifuged at 3000 g at room temperature for 30 minutes at 4°C. The supernatant was then added to 15% PEG and 1/10 vols. 5M NaCl and incubated at 4°C overnight.

This was then spun at 3000 g for 30 minutes at 4 °C, the supernatant discarded and the DNA pellet re suspended up to 2ml with TE. This was spun down a CsCl gradient at 500,000 g at 18 °C for 16 hours. Following removal of the DNA band from the gradients, the DNA was washed with CsCl saturated propan-2-ol to remove ethidium bromide. The DNA solution was then dialysed overnight to remove any CsCl present in the solution.

Following dialysis, the solution containing the DNA was added to an equal volume of buffered phenol/chloroform, the mixture vortexed, then spun at 13000 g at room temperature for 5 minutes. The top 0.4ml was transferred to a new microcentrifuge tube and 1 vol. of propan-2-ol and 0.25 vols. of sodium acetate were added. This was incubated at -20°C overnight, spun at 13000 g at room temperature for 15 minutes, the supernatant removed and 1 ml 70% ethanol was added. Following centrifugation at 13000 g at room temperature for 5 minutes the supernatant was removed. This process was repeated twice more, at which time the pellet was dried and re suspended in 100µl of SDW.

2.10.2. Purification of DNA using the Wizard® DNA Clean-Up System.

The DNA isolated from soil was purified to allow more efficient amplification of DNA sequences. This was used in accordance with the manufacturers instructions.

2.11. Storage of bacterial cultures.

Long term storage of bacterial strains was carried out using Protect Bacterial Preserves (Technical Service Consultants Ltd). These were used in accordance with the manufacturers instructions.

2.12. DNA hybridisation.

Both Southern and dot blot DNA hybridisations were utilised in this study, depending on the nature of the DNA sample being studied. Dot blots were used to confirm the identity of PCR products, whereas Southern hybridisations were used to determine the whether or not the sequence of interest was present on a particular DNA fragment on an agarose gel, e.g. where more than one size band was present.

2.12.1. Radionucleotides.

^{32}P labelled I-dCTP (370kBq) was used for labelling of probes for DNA hybridisation.

2.12.2. Dot blot hybridisation.

Dot blots were carried out using a Bio-Rad Bio-Dot vacuum manifold and DNA samples were prepared for blotting onto positive membrane (Appligene) in accordance with manufacturers instructions:

One sheet of positive membrane and two sheets of 3MM (Whatman) filter paper were cut to the size of the membrane and soaked in 2X SSC. 10µl of each PCR product (including appropriate controls) was added to 10µl 0.2M NaOH and incubated at room temperature for 10 minutes, before the tubes were transferred onto ice. The dot blot manifold was assembled and samples were loaded onto the membrane and left at room temperature for 30 minutes. Suction was then applied to the blotting apparatus for 30 seconds, the apparatus dismantled and the membrane left to air dry before prehybridisation.

2.12.3. Southern hybridisation (Southern, 1975).

Agarose gels were prepared for Southern hybridisation by denaturing the DNA in 0.25 M HCl for 15 minutes, or until the type II DNA loading buffer turned green. The gel was then rinsed in SDW and the transfer apparatus set up as in Figure 2.12.3.1, using 0.4M NaOH as the transfer buffer, and left overnight. Southern blots of plasmid DNA differed in that the transfer time was extended to 48 hours. Once transfer was complete the positive membrane was rinsed in 2X SSC to remove any agarose and prehybridisation was carried out.

2.12.4. Prehybridisation.

Prehybridisation of the positive membrane was carried out in accordance with the manufacturers instructions (Appligene). Prehybridisation buffer was prepared and added to the membrane, which had been placed in a clean hybridisation tube. Prehybridisation was carried out at 65 °C for at least 1 hour.

2.12.5. Preparation of DNA probes.

PCR products were commonly used as probes. These were prepared for hybridisation by an initial electrophoresis step in 1% (w/v) low melting point agarose (NuSieve GTG). The band was excised from the gel using a scalpel blade and the agarose containing the probe re suspended in 3 volumes of SDW. This was heated in a boiling water bath for 7 minutes and 11µl was used for the labelling reaction.

Probes were radiolabelled with ^{32}P dCTP using a random prime labelling kit (Boehringer Mannheim) as follows: 1 µl of dTTP, dATP and dGTP were added to 2 µ l of hexanucleotide mix, 3 µl of ^{32}P dCTP, 1 µl Klenow enzyme and 11µl of boiled probe. The reaction was carried out at 37°C for 1-2 hours.

The probe was then purified by passing it through a Sephadex G50 column (Sambrook *et al.*, 1989): 130 µl of SDW was added to the probe/labelling solution and placed into the top of the column. Successive 150µl additions of SDW were carried out and the liquid from the bottom of the column was collected for each

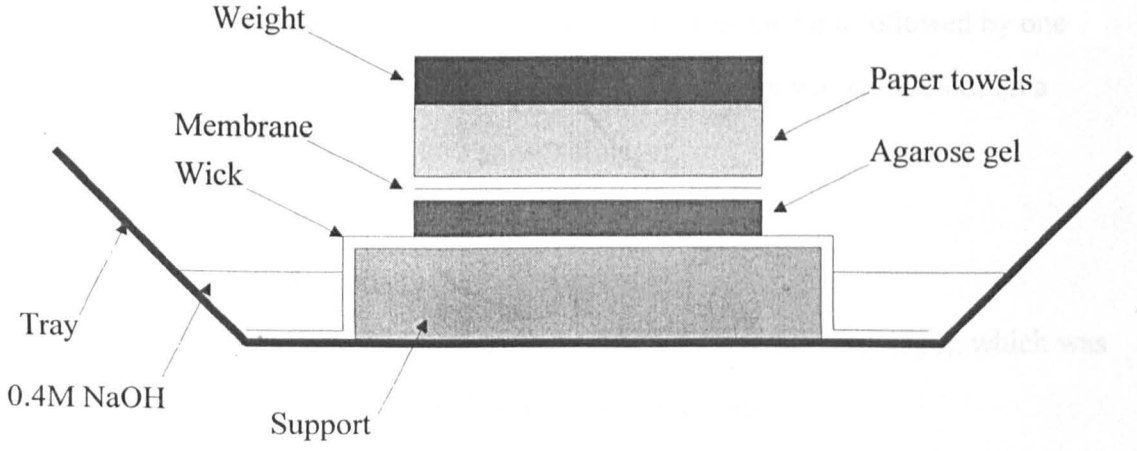


Figure 2.12.3.1. Apparatus used for Southern blotting of agarose gels.

addition. The two tubes with the highest detectable level of radioactivity were pooled, added to 100µl of denatured salmon sperm DNA (Sigma) and placed over a boiling water bath for 10 minutes. This probe mixture and 1ml SDW were added to the hybridisation tube containing the membrane/buffer and hybridised at 65 °C overnight.

2.12.6. Hybridisation and signal detection.

After overnight hybridisation at 65 °C, a 70% stringency wash was carried out as follows: The prehybridisation buffer/probe mix was disposed of and two five minute washes with 2X SSC at room temperature were carried out. Following this two 15 minute washes with 2X SSC/1% SDS (w/v) were carried out followed by one wash at room temperature with 0.1X SSC. Signal detection was carried out on a Molecular Dynamics STORM 860 phosphoimager.

2.13. Cloning of PCR products (Section 4.2.8 only).

Cloning was carried out using the pGEM®-T Easy Kit (Promega), which was used in accordance with the manufacturers recommendations:

2.13.1. Ligation reactions.

The PCR products were purified using Pharmacia Biotech MicroSpin S-400 HR columns, which were used in accordance with manufacturers instructions. Vector:PCR product concentration ratios of 3:1, 1:1 and 1:3 were used. Positive and negative DNA controls were set up using the control DNA supplied. Each ligation reaction contained: 1 µl T4 DNA ligase 10X buffer, 50ng pGEM®-T Easy vector, 1 µl T4 DNA ligase (3 Weiss units), 2 µl of control DNA, or the appropriate amount of the PCR product for cloning. The reaction volume was then made up to 10µl with SDW and mixed by pipetting. Ligation was carried out overnight at 4°C.

2.13.2. Preparation of competent cells.

Competent cells were prepared according to the method described by Nishimura *et al.*, 1990:

Half a ml of an overnight (LB) culture of XL1-BLUE was added to 10ml of LB and incubated at 37 °C with shaking for 2 hours. The culture was then spun at 8000 g for 4 minutes at 2°C at which time the supernatant was removed and 5ml of chilled 100mM $CaCl_2$ (4 °C)was added. The cells were gently re suspended, placed on ice for 20 minutes and again spun at 8000 g for 4 minutes (2°C). After removal of the supernatant, the tubes were left on ice for 1 hour before use in transformation reactions.

2.13.3. Transformation reactions.

Twoμl of each ligation reaction was added to a 1.5ml microcentrifuge tube and 50μl of competent cells were added. A negative control containing 0.1ng of pGEM®-T Easy vector and 100μl of competent cells was also set up. The transformation reactions were placed on ice for 20 minutes, at which time they were heat shocked at 42°C for 45 seconds then placed on ice for 2 minutes. 950μl (900μl for the negative control) LB was added to the tubes which were incubated at 37°C for 1.5 hours. The reaction mixes were then plated out onto MAXIT (methicillin, ampicillin, X-Gal, IPTG and tetracyclin) plates and incubated at 37°C overnight. All white colonies were then sub-cultured onto fresh MAXIT plates in preparation for plasmid extraction or PCR.

2.14. PCR.

PCR reactions comprised of the following; Template DNA, 10µl of 10X PCR buffer and 1.5mM $MgCl_2$ (GIBCO BRL) were added to the reaction mix along with 20pmol of each primer. 1 mM (final concentration) of dNTPs (Pharmacia) along with 2.5U of Taq DNA polymerase (GIBCO BRL) were added and the reaction volume was brought up 100µl using SDW and overlaid with mineral oil (Sigma). Cycling reactions were carried out in a Perkin Elmer 480 Thermal Cycler. The typical conditions used were 95°C for 4 mins, followed by 28 cycles of 95°C for 1min, 62°C for 1min and 72°C for 2mins followed by 10mins at 72°C. Different annealing temperatures were employed depending on the nature of the primer sequences.

To amplify regions of DNA longer then 3kb, a long template PCR system was used (Expand PCR, Boehringer Manheim) in accordance with the manufacturers instructions (Barnes, 1994). This typically utilised twice the normal concentrations of primers dNTPs and high quality template DNA. Cycling conditions were also routinely altered.

2.15. Oligonucleotides.

The oligonucleotide primers used in this study are shown in Table 2.15.1.

2.16. Visualisation of DNA samples.

This was carried out using agarose gel electrophoresis. Typically 0.7% TBE agarose gels were used. These were electrophoresed at 100V until adequate sample separation had occurred. For plasmid DNA samples, TAE buffer was used as this allowed the plasmid bands to be distinguished better than was possible on TBE gels. Ethidium bromide stain was used to allow DNA visualisation under UV (wavelength 254nm). This was either added to the agarose as the gel was poured (0.1µg ml gel^{-1}) or the gel was stained in ethidium bromide after electrophoresis had occurred. Gels containing plasmid DNA were stained after electrophoresis as ethidium bromide can alter the electrophoresis of the samples down the gel. Gel images were visualised on

Table 2.15.1. Oligonucleotide primers used in this study:

* All primers written 5'-3'. Standard IUB degeneracy code:
M = AC, R = AG, W = AT, S = GC, Y = CT, K = GT, V = ACG, H = ACT, D = AGT, B = CGT, N = GATC.
† These primers were designed to be homologous to the *tnpA* genes found on a variety of transposons (4.2.2. [c, f & m])

N/A Not Applicable.

[a & e] These 4 primers have been previously described, (Pearson *et al.*, 1996) having been designed to complement the DNA sequences of the Tn*21* and Tn*501 tnpR* and *tnpA* sequences. 501R1/C and 501R2/C were designed to correspond to the published DNA sequences of Tn*21* and Tn*501 tnpR* genes. They are positioned at either end of this 560 bp gene and produce a PCR product of approximately 500 bp. Primers 1406 and 2638 were designed to correspond to the published DNA sequences of Tn*21* and Tn*501 tnpA* genes and produce a PCR product of approximately 1.2 kb.

[b] This primer is the reverse complement of primer 1406.

[c] These primers were designed to the consensus sequences of the *tnpA* genes of Tn*1,1000,1412,1721,21,2501,3,3926,501,5036,5401,ISXC5.*

[d & i] These primers have previously been described, (Young, H. K. and Rosser, S. J., Personal communication). int21A/int21B were used as an indicator of integrase gene presence and 4127/4128 allow the size of the gene cassettes within an integron to be determined.

[f] These primers were designed to the consensus sequences of the *tnpA* genes of Tn*163,5393,5403,4652,5041.*

[g] This primer was designed to correspond to the sequence approximately 930bp from the start codon of the *tnpA* gene of Tn*501*/Tn*21.*

[h] These primers complement the DNA sequence in *merP* and *merA* respectively allowing the nature of any genes present in-between *merP* and *merA* to be determined.

[j] These primers have been designed to the *tnpA* and *tnpR* sequences of Tn*3* respectively.

[k] This primer was designed to be homologous to *merD* Tn*21/501* sequence.

[l] This primer is designed to complement the *merP* sequence of Tn*501.*

[m] These primers were designed to the consensus sequences of the *tnpA* genes of Tn*1546,4430,5422,917,Bac.*

[n] tnpR1/C is an extended version of 501R1/C.

[o] Universal and reverse primers were used to sequence from cloned PCR products in pGEM®-T Easy vector.

Primer	Sequence*	Accession No	Position
1406[a]	TGC GCT CCG GCG ACA TCT GG	X04891	1462-1481
1406REV[b]	CCA GAT GTC GCC SGA GCG CA	X04891	1462-1481
2501[c]	CAT TGG GAC GAG ATG ATG CGG	Various†	2501-2520
2638[a]	TCA GCC CGG CAT GCA CGC G	X04891	2676-2694
2850[c]	GCT CCA TAT ACA CCG TGT TCC A	Various†	2770-2792
4127[d]	TGA TCC GCA TGC CCG TTC CAT ACA G	X12870	709-733
4128[d]	GGC AAG CTT AGT AAA GCC CTC GCT AG	X12870	2321-2346
501R1/C[e]	GTT CAG CAS CTT CGA CCA G	X01298	539-557
501R2/C[e]	TAS AGG GTT TCS CGR CTG AT	X01298	1026-1045
62351[f]	TGG GCT TCA ATT TCG TTC CGC G	Various†	2316-2338
62601[f]	TTC AAT GCG GCG CAT TTC CCG	Various†	2564-2585
950[g]	YCT GGA ACT GCT GCT GAT GCT T	X04891	987-1008
mercP[h]	CCC GAT CAC WGT CAA GMA VGC	Z00027	1091-1110
mercA[h]	CGC TCG ATC AGC GWG ACV YG	Z00027	1696-1715
int21A[i]	GTC AAG GTT CTG GAC CAG TTG C	X12870	1090-1111
int21B[i]	ATC ATC GTC GTA GCG ACG TCG G	X12870	219-240
tn3R[j]	CCC TGC ATC TTT GAG CGC TCT	V00613	3270-3290
mercD[k]	GTT CGT CGA GCG TCG GCG	Z00027	3302-3319
pback[l]	TGC CCG ATC ACT GTC AAG AA	Z00027	1088-1108
pos2000[m]	GGA ATG AAT ATT GGT CTT ACC AAA ATG	Various†	1868-1892
pos2400[m]	CAG TAT AGC CAG CTG TGT CTG	Various†	2268-2289
tn3A[j]	GTA TCA GCG CTG CAT GCT CAC	V00613	2721-2741
tnpR1/C[n]	MGG GTC AGC ASC TTC GAC CAG	X01298	537-557
universal[o]	CAG GAA ACA GCT ATG AC	N/A‡	N/A‡
reverse[o]	TTT TCC CAG TCA AGC C	N/A‡	N/A‡

an Alpha Innotech IS500 Gel Imaging System. All agarose gels were loaded with GIBCO BRL size markers.

2.17. Restriction digestion of DNA.

Restriction digests of DNA samples were carried out using restriction enzymes (GIBCO BRL) which were used in accordance with manufacturers recommendations. Typically 1.5µl of the appropriate enzyme buffer was added to 10µl DNA sample, 2.5 µl SDW and 1µl restriction enzyme. This reaction mix was placed in a microcentrifuge tube at 37 °C for a sufficient time as to allow for complete digestion of the DNA sample. Inactivation of the restriction enzymes was carried out by the addition of 0.1 vols. of Type II DNA loading buffer.

2.18. DNA sequencing.

PCR products were purified for DNA sequencing using MicroSpin S-400 HR columns used in accordance with manufacturers instructions (Pharmacia Biotech). DNA sequencing (both DNA strands) was carried out on an ABI 373A automated sequencer. For longer sequences a LiCor 4200 DNA sequencer was used.

2.19. Analysis of DNA sequence data.

Sequence analysis was carried out using the GCG programs at the SEQNET facility, Daresbury Laboratories. The most commonly used programs include; FASTA, GAP, PILEUP, MAPSORT, REVERSE and TRANSLATE (Genetics Computer Group, 1994). The GenBank and EMBL DNA databases were also used along with the SwissProt protein database .

2.20. Construction of phylogenetic trees.

Alignment of sequences and conversion into suitable formats was carried out using PILEUP, LINEUP, TOPHYLIP, TOPIR and READSEQ programs in the PHYLIP 3.5C facility of the SEQNET database, Daresbury Laboratories which was also used for all subsequent analysis (Pearson *et al.*, 1996., Fitch, 1971., Olsen, 1990). DNAML and Treeview (Page, 1996) were used in the construction of Maximum Likelihood trees and DNA parsimony trees utilised DNAPARS and Treeview. Bootstrapping of phylogenetic trees was carried out using SEQBOOT and CONSENSE.

2.21. Accession numbers.

DNA sequences of class *tnpR* A-F (this study), have been assigned accession numbers U61277-U61281. The *merD/tnpR* fragment in 4.2.8. has been assigned accession number AF134211.

The DNA sequences of the *tnpR* genes used in 3.2.3 and the *tnpA* genes used in 4.2.2.1. have been assigned accession numbers as shown in Table 2.22.1.

Transposon	Resolvase (*tnpR*) sequence accession number	Transposase (*tnpA*) sequence accession number
Tn*1*	L10085	L10085
Tn*3*	V00613	V00613
Tn*21*	X01298	X04981
Tn*163*	N/A	L14931
Tn*501*	K01725	Z00027
Tn*917*	M11180	M11180
Tn*1000*	X60200	X60200
Tn*1331*	M55547	N/A
Tn*1412*	N/A	L36547
Tn*1546*	M97297	M97297
Tn*1721*	X02590	X61367
Tn*2501*	M15197	Y00502
Tn*3926*	X78059	X14236
Tn*4556*	M29297	M29297
Tn*4430*	N/A	X13481
Tn*4652*	N/A	X83686
Tn*4653*	D90148	N/A
Tn*5036*	N/A	Y09025
Tn*5041*	N/A	X98999
Tn*5053*	L40585	Z23093
Tn*5393*	M96392	M95402
Tn*5401*	N/A	U03554
Tn*5403*	X75779	X75779
Tn*5422*	N/A	L28104
Tn*5469*	N/A	U33002
pMER05*	L20694	N/A
Tn*B. firmus*	M90749	M90749
TnISXC5	N/A	Z73593

Table 2.21.1. Accession numbers of resolvase and transposase genes used in this study:

* resolvase related to that contained within Tn*552*.
N/A: Not used in this study.

CHAPTER 3.

PHYLOGENETIC ANALYSIS OF *tnpR* GENES IN MERCURY RESISTANT SOIL BACTERIA: THE RELATIONSHIP BETWEEN DNA SEQUENCING AND RFLP TYPING APPROACHES.

3.1. Introduction.

The diversity of *tnpA* genes, *tnpR* genes and mercury resistance genes (merRT ΔP) contained within a collection of 39 bacteria cultured from both mercury polluted and pristine sites has been examined by RFLP analysis (Bruce *et al.*, 1995., Osborn *et al.*, 1993., Pearson *et al.*, 1996) and in the case of the mercury resistance genes, a limited number have been examined by DNA sequencing (Osborn *et al.*, 1995). The *tnpR* genes from these mercury resistant Gram-negative bacteria have been characterised by RFLP analysis (Pearson *et al.*, 1996) and five classes of *tnpR* gene have been identified and designated *tnpR* class A-E based on their RFLP profiles. In this study (Holt *et al.*, 1996), representatives from each of these *tnpR* classes were sequenced and the resulting sequence data was compared to the RFLP data previously described. This analysis provides a useful insight into the relative merits of DNA sequencing and RFLP typing approaches. The sequence diversity of the *tnpR* genes in these strains was also studied by the use of phylogenetic trees (Felsenstein, 1993). Three methods of dendrogram construction have been used in this study (Holt *et al.*, 1996), which produce both rooted and unrooted phylogenetic trees; DNA maximum likelihood (Felsenstein, 1981), which compares the percentage similarity between sequences, the DNA parsimony method (Fitch, 1971) which calculates the minimum number of base changes needed on a given tree, and Neighbour joining rooted trees which utilises a method similar to the maximum likelihood dendrograms. The effects of DNA sequence changes on the amino acid sequence of *tnpR* are studied to ascertain

whether conserved DNA and amino acid sequence changes have occurred, and two previously uncharacterised *tnpR* genes were sequenced and allocated to RFLP classes.

In this study DNA sequence data and RFLP data were compared using using rooted and unrooted phylogenetic trees. This allows the relative merits of the two typing approaches to be considered and also reveals the factors influencing tree topology. The effects of DNA sequence changes on the amino acid sequence of *tnpR* are studied to ascertain whether conserved DNA and amino acid sequence changes have occured and two previously uncharacterised *tnpR* genes are sequenced and allocated to RFLP classes.

3.2. Results and Discussion.

3.2.1. Presence of *tnpR* in T2 isolates.

In a previous study carried out by Pearson *et al.*, 1996, the presence of *tnpR* genes was determined in 30 out of the 39 mercury resistant strains isolated by Osborn *et al.* (1993) and these genes were characterised by RFLP analysis. The presence of *tnpR* genes in the nine uncharacterised T2 strains, which were isolated from a polluted copper mine in Tipperary, Eire, was determined. PCR reactions were carried out using primers 501R1/C and 501R2/C on DNA from all 9 T2 isolates. This showed that the *tnpR* gene was present in all of the T2 isolates. T2 7 and T2 12 were chosen for further study.

3.2.2. Sequencing of PCR products.

A representative from each of the previously described *tnpR* class (Pearson *et al.*, 1996) together with one additional class A and four additional class C strains were selected for DNA sequencing. More than one class A and C were chosen in order to acertain the sequence similarity within RFLP classes. Class *tnpR* A was represented by '93 SO1 and '93 SO2, *tnpR* B by '93 SE31 and *tnpR* C by '93 SB3, '93 SB29,

'93SB22 and '93 SE23. Two previously uncharacterised strains, '93 T2 7 and '93 T2 12 were also studied. The PCR products of approximately 500 bp were sequenced directly using primers 501R1/C and 501R2/C (Figures 3.2.2.1 and 3.2.2.2) and DNA sequences were obtained from the representatives of each *tnpR* class. These sequences are shown in Figures 3.2.2.3-11. An alignment of the sequences of *tnpR* A, B, C, T2 7, T2 12, Tn*21* and Tn*501* is shown in Figure 3.2.2.12. The sequences obtained in this study have been assigned GenBank Accession numbers U61277-U61281.

3.2.3. Sequence similarity within RFLP classes.

The similarity between the 4 class C *tnpR* sequences, from '93 SB3,22,29 and '93 SE23 was determined and found to be 100% between '93 SB3 and '93 SB29 over the 372bp analysis region and 98.4% between these two and the 310bp of sequence obtained from '93 SE23. The similarity between '93 SB3 and 29 and '93 SB22 was 99.4% over the 364bp of sequence available for this strain.

'93 SO1 and SO2 Class A *tnpR* sequences were found to be identical to each other over a length of 383bp. For the purposes of this study, '93 SO1 and '93 SB3 sequences were chosen as group representatives of classes A and C respectively because of the greater amount of sequence data which was available for these strains. This high DNA sequence similarity within each RFLP class (ranging from 97.6% to 100%) supports the previously published RFLP data.

3.2.4. Phylogenetic analysis.

Initially an unrooted parsimony tree was constructed showing the relationships between the *tnpR* genes sequenced in this study and those for which corresponding sequence data is available. This tree shows the relative position of the sequences obtained from these isolates as compared to all known *tnpR* sequences. The sequence data from classes *tnpR* A, B and C were aligned with the published sequences of 18 different *tnpR* genes and these data used to construct the dendrogram shown in Figure 3.2.4.1. This shows four major groupings of resolvase genes. The sequences obtained

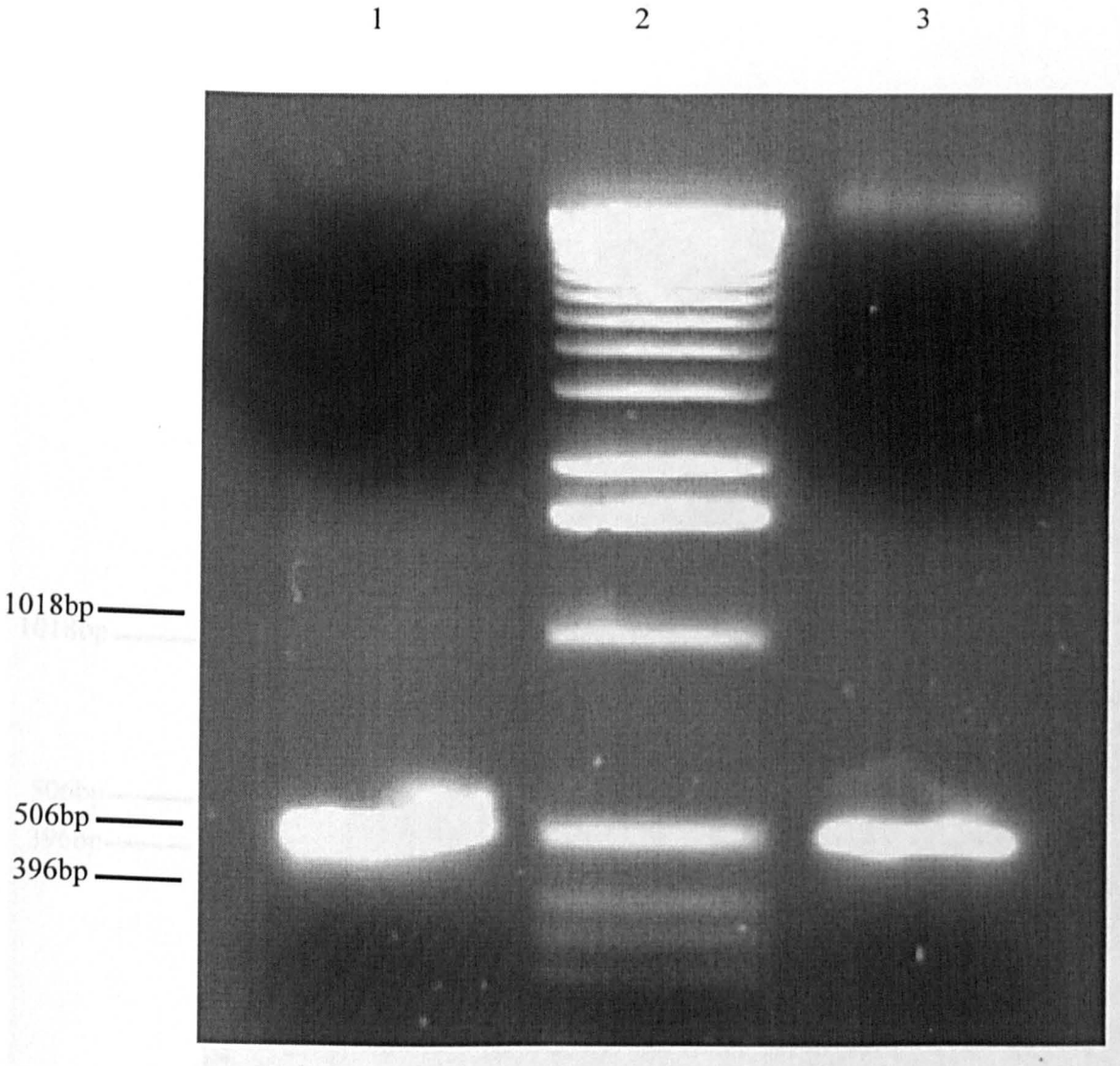


Figure 3.2.2.1. Agarose gel showing PCR products produced using primers 501R1/C and 501R2/C designed to allow amplification of *tnpR* genes. These PCR products were sequenced and the resulting data used to construct phylogenetic trees.

Lane 1 PCR product from '93 SO1
2 kb ladder
3 PCR product from '93 SE31

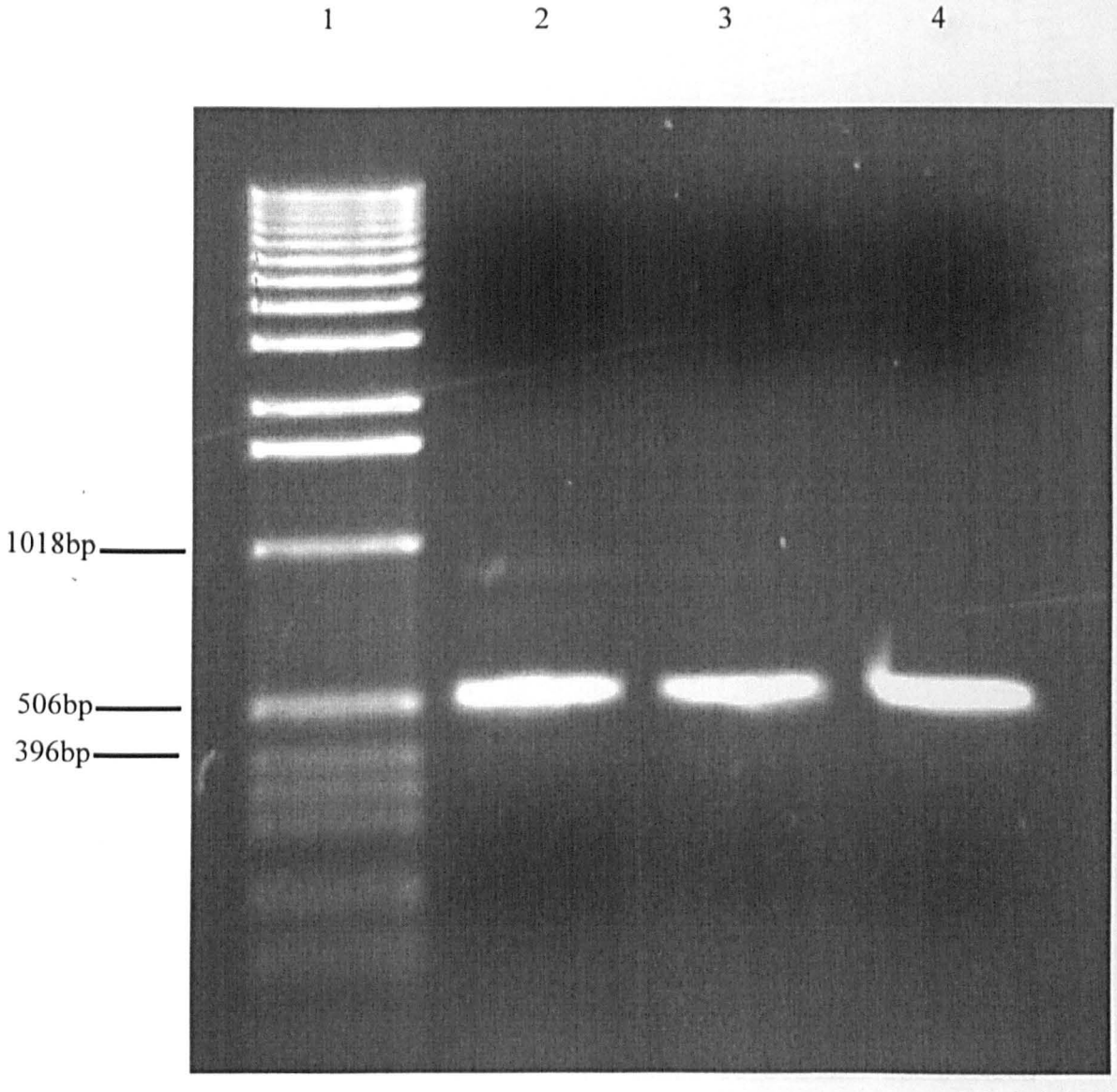


Figure 3.2.2.2. Agarose gel showing PCR products produced using primers 501R1/C and 501R2/C designed to allow amplification of of *tnpR* genes. These PCR products were sequenced and the resulting data used to construct phylogenetic trees.

Lane 1 kb ladder
2 PCR product from '93 SB2
3 PCR product from '93 SB3
4 PCR product from '93 SB29

Figure 3.2.2.3. Nucleotide sequence from '93 SO1, Class A *tnpR* gene. This sequence was assigned accession number U61277.

```
  90  TTCACCGAAC AAGGCTTCTG GCAAGGACAC CCAGCGTCCC GAGCTGGAAA
 141  GGCTGCTGGC CTTCGTCCGC GAGGGCGACA CCGTGGTGGT GCATAGCATG
 190  GACAGGCTGG CACGCAACCT TGATGACCTG CGCCGCATCG TCCAAGGGCT
 241  GACACAACGG GGCGTGCGCA TGGAGTTCGT CAAAGAAGGG CTGAAGTTCA
 290  CCGGCGAGGA CTCACCGATG GCCAATCTGA TGCTGTCGGT CATGGGAGCC
 341  TTCGCTGAGT TCGAGCGCGC CCTGATCCGC GAACGTCAGC GCGAGGGAAT
 390  CGTGCTGGCC AAGCAGCGCG GTGCCTACCG GGGACGAAAG AAATCGCTGA
 441  ACAGCGAACA AATTGCCGAG TTGAAACGGC GAGTTGCGGC GAACCAAAAA
 490  AC
```

Figure 3.2.2.4. Nucleotide sequence from '93 SE31, Class B *tnpR* gene. This sequence was assigned accession number U61278.

```
  74  TCGACCTCGG TAAGGTATTC ACCGATAAGG CTTCCGGCAA GGACACGCAA
 124  CGTCCCGAAA CTTGAAAGGC TGCTGGCCTT TGTGCGCGAG GGCGACACCG
 174  TGGTGGTGCA CAGCATGGAC AGGTTGGCAC GCAATCTCGA TGACCTGCGC
 224  CGCATCGTTC AGGAGCTGAC ACAACGGGGC GTGCGGATGG AGTTTGTCAA
 274  AGAAGGGCTG GCGTTCACCG GCGATGACTC ACCGATGGCC AATTTGATGC
 324  TGTCGGTCAT GGGGGCTTTT GCGGAGTTCA GCGCGCACTG TATCCGCGAA
 374  CGCCAGCGCG AGGGAATCGT GCTGGCCAAG CAGCGCGGTG CCTACCGGGG
 424  ACGAAAGAAA TCGCTGAACA GCGAACAAAT TGCCAAGTTG AAACAGCGAG
 474  TCGCGGCAGG CGATCAAAAA AC
```

Figure 3.2.2.5. Nucleotide sequence from'93 SB3, Class C, *tnpR* gene. This sequence was assigned accession number U61279.

```
  81  GGATTAAGTG TTCACCGACA AGCGTCGGGC AAGGACACAC GGCGGCCCGA
 131  ACTGGAACGG CTGCTCGCCT TCGTGCGCGA AGGCGACACG GTCGTGGTGC
 181  ACAGCATGGA TCGTCTGGCG CGCAACCTCG ACGACCTGCG CCGCCTGGTG
 231  CAGGGCCTCA CCCAGCGCGG CGTACGCATC GAGTTCCTTA AGGAGCATTT
 281  GACCTTCACC GGCGAGGACT CGCCGATGGC GAACCTGATG CTGTCGGTAA
 331  TGGGCGCGTT CGCCGAGTTC GAACGCGCCT TGATCCGCGA GCGGCAGCGC
 381  GAGGGCATCG CGCTCGCCAA GCAGCGCGGG GCCTACCGTG GCAGAAAGAA
 431  ATCCCTGTCG TCTGAACGTA TTGCCGAACT GCGCCAACGT TTT
```

Figure 3.2.2.6. Nucleotide sequence from '93 T2 7, Class D *tnpR* gene. This sequence was assigned accession number U61280.

```
 108  GGGCAAGGAC ACCCAGCGCC CCCAGGTTGG GAAGCGCTTG CTGAGCTTCG
 158  TCCGCGAAGG CGATACAGTG GTGGTGCACA GCATGGACCG GCTGGCCCGC
 208  AACCTCGATG ACCTGCGTTC GCTTGGTACA GAAGCTGACT CAACGCGGCG
 258  TGCGCATCGA GTTTCCTGAA GGAGGGCCTG GTGTTTCACT GGCGAAGGAC
 308  TCGCCGATGG CCAACCTGAT GCTGTCGGTG ATGGGGGCCT TCGCTGAGTT
 358  TCGAACGCGC CCTGATCCGC GAGCGGCAGC GTGAGGGCAT CACCTTGGCC
 408  AAGCAGCGTG GCGCGTACCG GGGCCGCAAG AAATCCCTGT CC
```

Figure 3.2.2.7. Nucleotide sequence from '93 T2 12, Class F *tnpR* gene. This sequence was assigned accession number U61281.

```
  85  CCCAGCGGCC CGAGTCTTGA TTCGATGACT GGCTTCGTAC GTCGAAGGGA
 135  CACCGTGGTG GTTTCATAGC ATGGATCAGC CTGGCGCCGC AACCTTCGAT
 185  GACTTGCGTC CGCCTCGTGC AAAAGCTCAC CAAGCGCGGT GTTGGGTTAT
 235  CGAGTTCGTC AAGGAAAGCC TGTACCTTCA CCGGGCGAGG AATTCACCCG
 285  ATGGACGCAA CCTAATGCTG TCTGTTCATG GGGGCGTTTC GCCCAATTCT
 335  AGCGGGCCTT GTATCCCCGA CGGCAGAGGG TAAGGCATCG CGTTCGTCCA
 385  AAACAGCGCC GGTAACCCTT ATTCGGGGCC GCAAAGAAGA AAGATG
```

Figure 3.2.2.8. Nucleotide sequence from '93 SO2, Class A *tnpR* gene.

```
  90  TTCACCGAAC AAGGCTTCTG GCAAGGACAC CCAGCGTCCC GAGCTGGAAA
 140  GGCTGCTGGC CTTCGTCCGC GAGGGCGACA CCGTGGTGGT GCATAGCATG
 190  GACAGGCTGG CACGCAACCT TGATGACCTG CGCCGCATCG TCCAAGGGCT
 240  GACACAACGG GGCGTGCGCA TGGAGTTCGT CAAAGAAGGG CTGAAGTTCA
 290  CCGGCGAGGA CTCACCGATG GCCAATCTGA TGCTGTCGGT CATGGGAGCC
 340  TTCGCTGAGT TCGAGCGCGC CCTGATCCGC GAACGTCAGC GCGAGGGAAT
 390  CGTGCTGGCC AAGCAGCGCG GTGCCTACCG GGGACGAAAG AAATCGCTGA
 440  ACAGCGAACA AATTGCCGAG TTGAAACGGC GAGTTGCGGC GAACCAAAAA
 490  AC
```

Figure 3.2.2.9. Nucleotide sequence from '93 SB29, Class C *tnpR* gene.

```
 85  AAAGTGTTCA CCGACAAGCG TCGGGCAAGG ACACACGGCG GCCCGAACTG
135  GAACGGCTGC TCGCCTTCGT GCGCGAAGGC GACACGGTCG TGGTGCACAG
185  CATGGATCGT CTGGCGCGCA ACCTCGACGA CCTGCGCCGC CTGGTGCAGG
235  GCCTCACCCA GCGCGGCGTA CGCATCGAGT TCCTTAAGGA GCATTTGACC
285  TTCACCGGCG AAGGACTCGC CGATGGCGAA CCTGATGCTG TCGGTAATGG
335  GCGCGTTCGC CGAGTTCGAA CGCGCCTTGA TCCGCGAGCG GCAGCGCGAG
385  GGCATCGCGC TCGCCAAGCA GCGCGGGGCC TACCGTGGCA GAAAGAAATC
435  CCTGTCGTCG GAACGTTATT GCCGAACTGC GCC
```

Figure 3.2.2.10. Nucleotide sequence from '93 SE23, Class C *tnpR* gene.

```
 85  AAAGTGTTCA CCGACAAGGC GTCGGGCAAG GACACACGGC GGCCCGAACT
135  GGAACGGCTG CTCGCCTTCG TGCGCGAAGG GCGAACACGG TTCGTGGTGC
185  ACAGCATGGA TCGTCTGGCG CGCAACCTCG ACGACCTGCG CCGCCTGGTG
235  CAGGGCCTCA CCCAGCGCGG CGTACGCATC GAGTTCCTTA AGGAGCATTT
285  GACCTTCACC GGCGAGGACT CGCCGATGGC GAACCTGATG CTGTCGGTAA
335  TGGGCGCGTT CGCCGAGTTC GAACGCGCCT TGATCCGCGA GCGGCAGCGC
385  GAGGGCATCG
```

Figure 3.2.2.11. Nucleotide sequence from '93 SB22, Class C *tnpR* gene.

```
 67  GAACAGATCC AGGTGGATAA AGTGTTCACC GACAAGGCGT CGGGCAAGGA
117  CACACGGCGG CCCGAACTGG AACGGCTGCT CGCCTTCGTG CGCGAAGGCG
167  ACACGGTCGT GGTGCACAGC ATGGATCGTC TGGCGCGCAA CCTCGACGAC
217  CTGCGCCGCC TGGTGCAGGG CCTCACCCAG CGCGGCGTAC GCATCGAGTT
267  CCTTAAGGAG CATTTGACCT TCACCGGCGA GGACTCGCCG ATGGCGAACC
317  TGATGCTGTC GGTAATGGGC GCGTTTCGCC GAGTTCGAAC GCGCCTTGAT
367  CCGCGAGCGG CAGCGCGAAG GCATCGCGCT CGCCCAAGCA GCGCGGGGCC
417  TACCGTGGCA GGAA
```

Figure 3.2.2.12. DNA sequence alignment of *tnpR* classes A-E. Class A is represented by '93 SO1 and 2, Class B by '93 SE31, Class C by '93 SB29, SB3, SB22 and SE23. Class D sequence from Tn*501* and '93 T2 7 is shown. The complete *tnpR* open reading frames from Tn*21* (Class E) and Tn*501* (Class D) are also shown.

```
                    1                                                  50
'93 T27      Class D  .......... .......... .......... .......... ..........
Tn501        Class D  GTGCAGGGGC ACCGCATCGG CTACGTCCGG GTCAGCAGCT TCGACCAGAA
'93 SO1      Class A  .......... .......... .......... .......... ..........
'93 SO2      Class A  .......... .......... .......... .......... ..........
'93 SE31     Class B  .......... .......... .......... .......... ..........
'93 SB29     Class C  .......... .......... .......... .......... ..........
'93 SB3      Class C  .......... .......... .......... .......... ..........
'93 SB22     Class C  .......... .......... .......... .......... ..........
'93 SE23     Class C  .......... .......... .......... .......... ..........
Tn21         Class E  ATGACTGGAC AGCGCATTGG GTATATCAGG GTCAGCACCT TCGACCAGAA

                   51                                                 100
'93 T27      Class D  .......... .......... .......... .......... ..........
Tn501        Class D  CCCGGAACGC CAGCTGGAAC AGACACAGGT GAGCAAGGTG TTCACCGACA
'93 SO1      Class A  .......... .......... .......... .........T TCACCGAACA
'93 SO2      Class A  .......... .......... .......... .........T TCACCGAACA
'93 SE31     Class B  .......... .......... ...TCGACCT CGGTAAGGTA TTCACCGATA
'93 SB29     Class C  .......... .......... .......... ....AAAGTG TTCACCGACA
'93 SB3      Class C  .......... .......... .......... GGATTAAGTG TTCACCGACA
'93 SB22     Class C  .......... ......GAAC AGATCCAGGT GGATAAAGTG TTCACCGACA
'93 SE23     Class C  .......... .......... .......... ....AAAGTG TTCACCGACA
Tn21         Class E  CCCGGAACGG CAACTGGAAG GCGTCAAGGT TGATCGCGCT TTTAGCGACA

                   101                                                150
'93 T27      Class D  .......GGG CAAGGACACC CAGCGCCCCC AGGTTGGGAA GCGCTTGCTG
Tn501        Class D  AGGCATCGGG CAAGGACACC CAGCGCCCCC AGCTCGAAGC GC...TGCTG
'93 SO1      Class A  AGGCTTCTGG CAAGGACACC CAGCGTCCCG .AGCTGGAAA G..GCTGCTG
'93 SO2      Class A  AGGCTTCTGG CAAGGACACC CAGCGTCCCG .AGCTGGAAA G..GCTGCTG
'93 SE31     Class B  AGGCTTCCGG CAAGGACACG CAACGTCCCG AAACTTGAAA G..GCTGCTG
'93 SB29     Class C  A.GCGTCGGG CAAGGACACA CGGCGGCCCG .AACTGGAAC G..GCTGCTC
'93 SB3      Class C  A.GCGTCGGG CAAGGACACA CGGCGGCCCG .AACTGGAAC G..GCTGCTC
'93 SB22     Class C  AGGCGTCGGG CAAGGACACA CGGCGGCCCG .AACTGGAAC G..GCTGCTC
'93 SE23     Class C  AGGCGTCGGG CAAGGACACA CGGCGGCCCG .AACTGGAAC G..GCTGCTC
Tn21         Class E  AGGCATCCGG CAAGGATGTC AAGCGTCCGC .AACTGGAAG C..GCTGATA
```

```
                    151                                                 200
'93 T27      Class D  AGCTTCGTCC GCGAAGGCGA TACA...GTG GTGGTGCACA GCATGGACCG
Tn501        Class D  AGCTTCGTCC GCGAAGGCGA TACA...GTG GTGGTGCACA GCATGGACCG
'93 SO1      Class A  GCCTTCGTCC GCGAGGGCGA CAC...CGTG GTGGTGCATA GCATGGACAG
'93 SO2      Class A  GCCTTCGTCC GCGAGGGCGA CAC...CGTG GTGGTGCATA GCATGGACAG
'93 SE31     Class B  GCCTTTGTGC GCGAGGGCGA CAC...CGTG GTGGTGCACA GCATGGACAG
'93 SB29     Class C  GCCTTCGTGC GCGAAGGCGA CAC...GGTC GTGGTGCACA GCATGGATCG
'93 SB3      Class C  GCCTTCGTGC GCGAAGGCGA CAC...GGTC GTGGTGCACA GCATGGATCG
'93 SB22     Class C  GCCTTCGTGC GCGAAGGCGA CAC...GGTC GTGGTGCACA GCATGGATCG
'93 SE23     Class C  GCCTTCGTGC GCGAAGGGCG AACACGGTTC GTGGTGCACA GCATGGATCG
Tn21         Class E  AGCTTCGCCC GCACCGGCGA CAC...CGTG GTGGTGCATA GCATGGATCG

                    201                                                 250
'93 T27      Class D  GCTGGCCCGC AACCTCGATG ACCTGCGTTC GCTTGGTACA GAAGCTGACT
Tn501        Class D  GCTGGCCCGC AACCTCGATG ACCTGCG.TC GCTTGGTACA GAAGCTGACT
'93 SO1      Class A  GCTGGCACGC AACCTTGATG ACCTGCG.CC GCATCGTCCA AGGGCTGACA
'93 SO2      Class A  GCTGGCACGC AACCTTGATG ACCTGCG.CC GCATCGTCCA AGGGCTGACA
'93 SE31     Class B  GTTGGCACGC AATCTCGATG ACCTGCG.CC GCATCGTTCA GGAGCTGACA
'93 SB29     Class C  TCTGGCGCGC AACCTCGACG ACCTGCG.CC GCCTGGTGCA GGGCCTCACC
'93 SB3      Class C  TCTGGCGCGC AACCTCGACG ACCTGCG.CC GCCTGGTGCA GGGCCTCACC
'93 SB22     Class C  TCTGGCGCGC AACCTCGACG ACCTGCG.CC GCCTGGTGCA GGGCCTCACC
'93 SE23     Class C  TCTGGCGCGC AACCTCGACG ACCTGCG.CC GCCTGGTGCA GGGCCTCACC
Tn21         Class E  CCTGGCGCGC AATCTCGATG ATTTGCG.CC GGATCGTGCA AACGCTGACA

                    251                                                 300
'93 T27      Class D  CAACGCGGCG TGCGCATCGA GTTTCCTGAA GGAGGGCCTG GTGTTTCACT
Tn501        Class D  CAACGCGGCG TGCGCATCGA G.TTCCTGAA GGAGGGCCTG GTG.TTCACT
'93 SO1      Class A  CAACGGGGCG TGCGCATGGA G.TTCGTCAA AGAAGGGCTG AAG.TTCACC
'93 SO2      Class A  CAACGGGGCG TGCGCATGGA G.TTCGTCAA AGAAGGGCTG AAG.TTCACC
'93 SE31     Class B  CAACGGGGCG TGCGGATGGA G.TTTGTCAA AGAAGGGCTG GCG.TTCACC
'93 SB29     Class C  CAGCGCGGCG TACGCATCGA G.TTCCTTAA GGAGCATTTG ACC.TTCACC
'93 SB3      Class C  CAGCGCGGCG TACGCATCGA G.TTCCTTAA GGAGCATTTG ACC.TTCACC
'93 SB22     Class C  CAGCGCGGCG TACGCATCGA G.TTCCTTAA GGAGCATTTG ACC.TTCACC
'93 SE23     Class C  CAGCGCGGCG TACGCATCGA G.TTCCTTAA GGAGCATTTG ACC.TTCACC
Tn21         Class E  CAACGCGGCG TGCATATCGA A.TTCGTCAA GGAACACCTC AGT.TTTACT
```

```
                        301                                                 350
'93 T27      Class D  GGCGAAGGAC TCGCCGATGG CCAACCTGAT GCTGTCGGTG ATGGGGGCCT
Tn501        Class D  GGCG.AGGAC TCGCCGATGG CCAACCTGAT GCTGTCGGTG ATGGGGGCCT
'93 SO1      Class A  GGCG.AGGAC TCACCGATGG CCAATCTGAT GCTGTCGGTC ATGGGAGC.C
'93 SO2      Class A  GGCG.AGGAC TCACCGATGG CCAATCTGAT GCTGTCGGTC ATGGGAGC.C
'93 SE31     Class B  GGCG.ATGAC TCACCGATGG CCAATTTGAT GCTGTCGGTC ATGGGGGC.T
'93 SB29     Class C  GGCGAAGGAC TCGCCGATGG CGAACCTGAT GCTGTCGGTA ATGGGCGCG.
'93 SB3      Class C  GGCG.AGGAC TCGCCGATGG CGAACCTGAT GCTGTCGGTA ATGGGCGCG.
'93 SB22     Class C  GGCG.AGGAC TCGCCGATGG CGAACCTGAT GCTGTCGGTA ATGGGCGCGT
'93 SE23     Class C  GGCG.AGGAC TCGCCGATGG CGAACCTGAT GCTGTCGGTA ATGGGCGCG.
Tn21         Class E  GGCGAA.GAC TCTCCGATGG CGAACCTGAT GCTCTCGGTG ATGGGCGCG.

                        351                                                 400
'93 T27      Class D  TCGCTGAGTT TCGAACGCGC CCTGATCCGC GAGCGGCAGC GTGAGGGCAT
Tn501        Class D  TCGCTGAG.T TCGAGCGCGC CCTGATCCGC GAGCGGCAGC GTGAGGGCAT
'93 SO1      Class A  TTCGCTGAGT TCGAGCGCGC CCTGATCCGC GAACGTCAGC GCGAGGGAAT
'93 SO2      Class A  TTCGCTGAGT TCGAGCGCGC CCTGATCCGC GAACGTCAGC GCGAGGGAAT
'93 SE31     Class B  TTTGCGGAGT TCAGCGCGCA CTGTATCCGC GAACGCCAGC GCGAGGGAAT
'93 SB29     Class C  TTCGCCGAGT TCGAACGCGC CTTGATCCGC GAGCGGCAGC GCGAGGGCAT
'93 SB3      Class C  TTCGCCGAGT TCGAACGCGC CTTGATCCGC GAGCGGCAGC GCGAGGGCAT
'93 SB22     Class C  TTCGCCGAGT TCGAACGCGC CTTGATCCGC GAGCGGCAGC GCGAAGGCAT
'93 SE23     Class C  TTCGCCGAGT TCGAACGCGC CTTGATCCGC GAGCGGCAGC GCGAGGGCAT
Tn21         Class E  TTCGCCGAGT TCGAGCGCGC CCTGATCCGC GAGCGTCAGC GCGAGGGTAT

                        401                                                 450
'93 T27      Class D  CAC.CTTGGC CAAGCAGCGT GGCGCGTACC GGGGCCGCAA GAAATCCCTG
Tn501        Class D  CAC.CTTGGC CAAGCAGCGT GGCGCGTACC GGGGCCGCAA GAAAGCCCTG
'93 SO1      Class A  CGTGCT.GGC CAAGCAGCGC GGTGCCTACC GGGGACGAAA GAAATCGCTG
'93 SO2      Class A  CGTGCT.GGC CAAGCAGCGC GGTGCCTACC GGGGACGAAA GAAATCGCTG
'93 SE31     Class B  CGTGCT.GGC CAAGCAGCGC GGTGCCTACC GGGGACGAAA GAAATCGCTG
'93 SB29     Class C  CGCGCTCG.C CAAGCAGCGC GGGGCCTACC GTGGCAGAAA GAAATCCCTG
'93 SB3      Class C  CGCGCTCG.C CAAGCAGCGC GGGGCCTACC GTGGCAGAAA GAAATCCCTG
'93 SB22     Class C  CGCGCTCGCC CAAGCAGCGC GGGGCCTACC GTGGCAGGAA ..........
'93 SE23     Class C  CG........ .......... .......... .......... ..........
Tn21         Class E  TGCGCTCG.C CAAGCAACGC GGGGCTTACC GTGGCAGGAA GAAATCCCTG
```

```
                    451                                                 500
'93 T27      Class D  TCC..... ..........  ..........  ..........  ..........
Tn501        Class D  TCCGATGAGC .AGGCTGCTA CCCTGCGGCA GCGAGCGACG GCCGGCGAGC
'93 SO1      Class A  AACAGCGAAC .AAATTGCCG AGTTGAAACG GCGAGTTGCG GCGAACCAAA
'93 SO2      Class A  AACAGCGAAC .AAATTGCCG AGTTGAAACG GCGAGTTGCG GCGAACCAAA
'93 SE31     Class B  AACAGCGAAC .AAATTGCCA AGTTGAAACA GCGAGTCGCG GCAGGCGATC
'93 SB29     Class C  TCGTCGGAAC GTTATTGCCG AACTGCGCC. ..........  ..........
'93 SB3      Class C  TCGTCTGAAC G.TATTGCCG AACTGCGCCA ACGTTTT...  ..........
'93 SB22     Class C  ..........  ..........  ..........  ..........  ..........
'93 SE23     Class C  ..........  ..........  ..........  ..........  ..........
Tn21         Class E  TCGTCTGAGC G.TATTGCCG AACTGCGCCA ACGTGTCGAG GCTGGCGAGC

                    501                                                 550
'93 T27      Class D  ..........  ..........  ..........  ..........  ..........
Tn501        Class D  CCAAGGCGCA GCTTGCCCGC GAGTTCAACA TCAGCCGGGA AACCCTCTAC
'93 SO1      Class A  AAAC......  ..........  ..........  ..........  ..........
'93 SO2      Class A  AAAC......  ..........  ..........  ..........  ..........
'93 SE31     Class B  AAAAAAC...  ..........  ..........  ..........  ..........
'93 SB29     Class C  ..........  ..........  ..........  ..........  ..........
'93 SB3      Class C  ..........  ..........  ..........  ..........  ..........
'93 SB22     Class C  ..........  ..........  ..........  ..........  ..........
'93 SE23     Class C  ..........  ..........  ..........  ..........  ..........
Tn21         Class E  AAAAGACCAA GCTTGCTCGT GAATTCGGAA TCAGTCGCGA AACCCTGTAT

                    551                 574
'93 T27      Class D  ..........  ..........  ....
Tn501        Class D  CAGTACCTCC GCACGGACGA CTGA
'93 SO1      Class A  ..........  ..........  ....
'93 SO2      Class A  ..........  ..........  ....
'93 SE31     Class B  ..........  ..........  ....
'93 SB29     Class C  ..........  ..........  ....
'93 SB3      Class C  ..........  ..........  ....
'93 SB22     Class C  ..........  ..........  ....
'93 SE23     Class C  ..........  ..........  ....
Tn21         Class E  CAATACTTGA GAACGGATCA GTAA
```

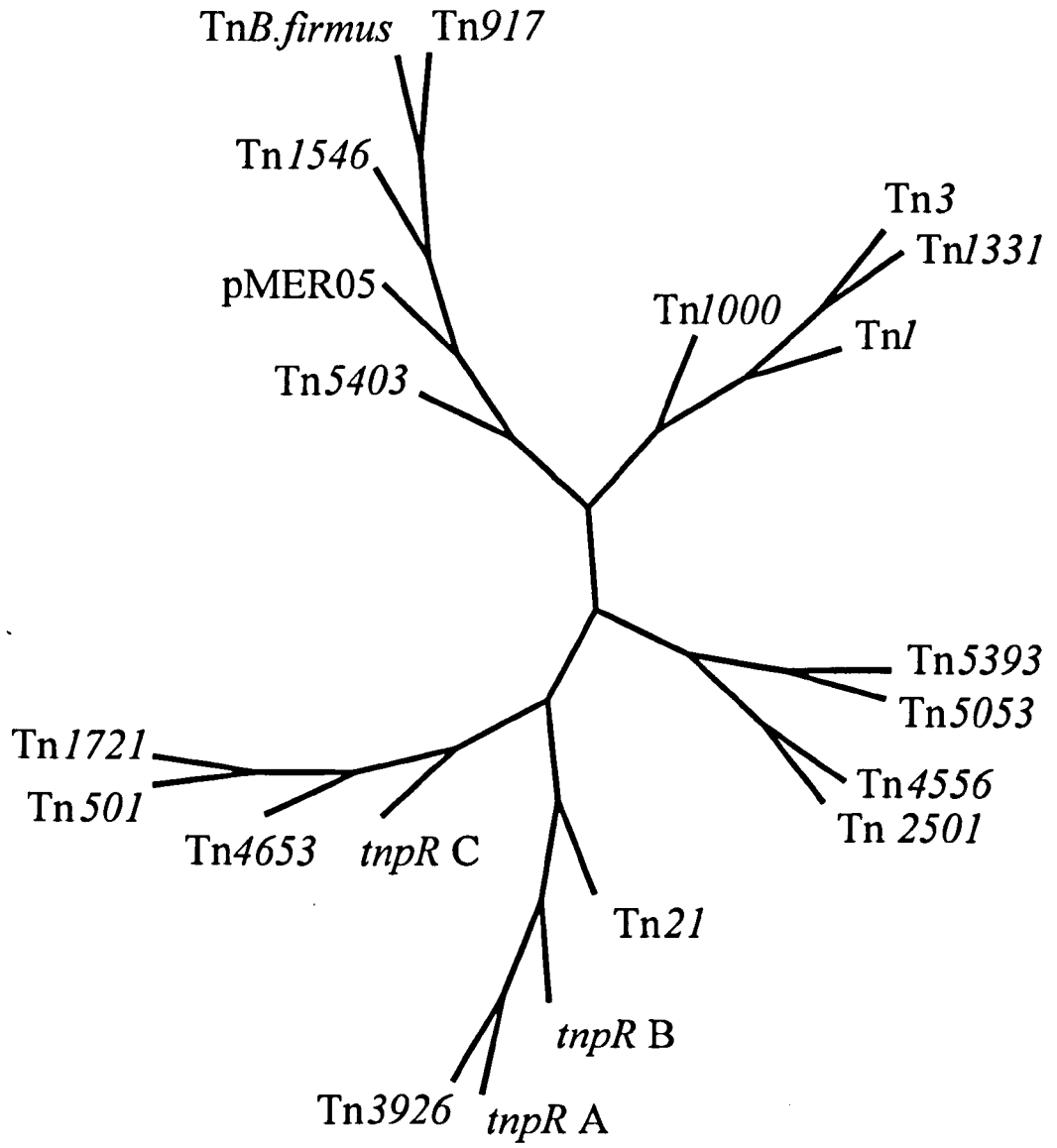


Figure 3.2.4.1. DNA Parsimony tree showing *tnpR* genes from classes *tnpR* A,B, C,D,E and common transposons. *tnpR* A,B and C correspond to RFLP classes previously described by Pearson *et al.*, 1996 and sequenced in this study. *tnpR* D is represented by Tn*501 and tnpR* E by Tn*21*. Bootstrap values (not shown) were between 98 and 100.

in this study fell into one of these groups, which also contains Tn*501* and Tn*21* from which the primers used in this study were designed.

The sequence data obtained from each of the *tnpR* classes and the published sequences of Tn*501* and Tn*21* (i.e. classes A-E) were aligned, resulting in a 372bp analysis region. This 372bp region represents the amount of sequence data which was available from the shortest sequence, i. e. more sequence data was available for other isolates, but this region of sequence was not present in all isolates. This alignment was used to construct a number of phylogenetic trees:

Initially, a rooted tree was constructed to allow comparison of the data from this study with that from the study of Pearson *et al*, 1996 (Figure 3.2.4.2). This tree is consistent with that of Pearson *et al*, 1996 (Figure 3.2.4.3). As unrooted trees are preferable for showing the differences/similarities between species, the remainder of the analysis was carried out using two methods of tree construction; maximum likelihood and DNA parsimony. The relative advantages and disadvatages of each method have been discussed at length (DeBry and Abele, 1995, Felsenstein, 1981, Fitch, 1971, Fukami and Tateno, 1991, Hills *et al.*, 1994, Huelsenbeck, 1995, Tateno *et al.*, 1994, Tillier and Collins, 1995) and the two methods used to construct unrooted trees produced comparable results.

An unrooted tree was constructed from the 372bp analysis region using the maximum likelihood method (Figure 3.2.4.4) and the DNA parsimony method (Figure 3.2.4.5). Both of these were consistant with the dendrogram derived by Pearson *et al*, 1996, (Figure 3.2.4.2) and were also compatible with each other. Four further unrooted trees derived from the first and second 186bp halves of the 372bp sequence were constructed. (Figures 3.2.4.6, 3.2.4.7, 3.2.4.8 and 3.2.4.9.). The dendrogram corresponding to the first half of the initial fragment differed from that of the overall analysis region, in that the positions of class *tnpR* D and *tnpR* E are reversed whereas the dendrogram representing the second half of the fragment is similar to that of the overall analysis region.

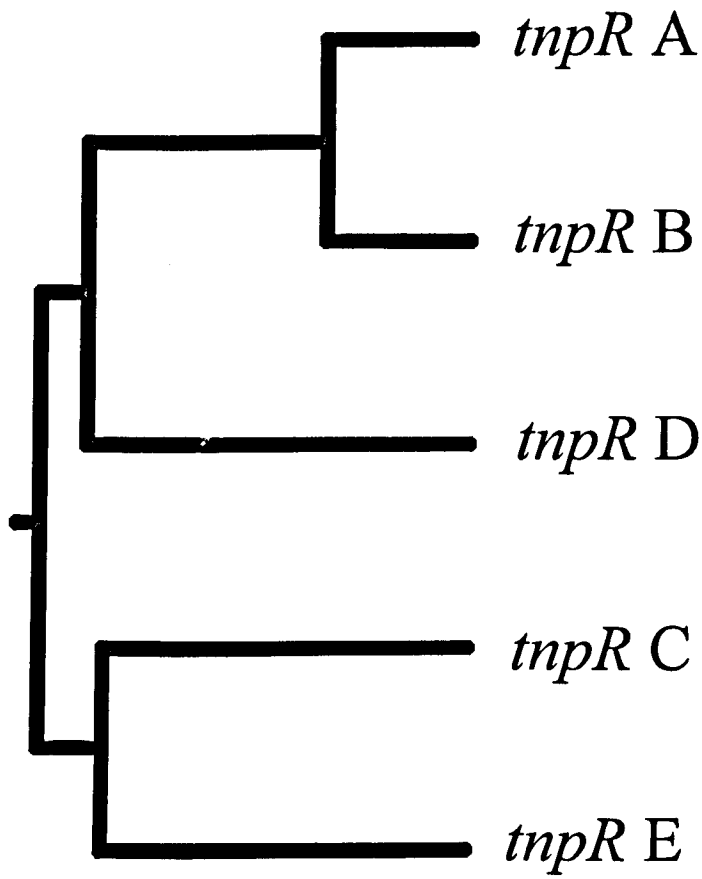


Figure 3.2.4.2. Rooted maximum likelihood tree of 372 bp resolvase analysis region. This 372 bp region represents the sequence data which was available for all five classes. This tree was constructed to allow comparison of this data with that previously published by Pearson *et al.*, 1996. *tnpR* classes as described in Figure 3.2.4.1. Bootstrap values (not shown) were between 99 and 100.

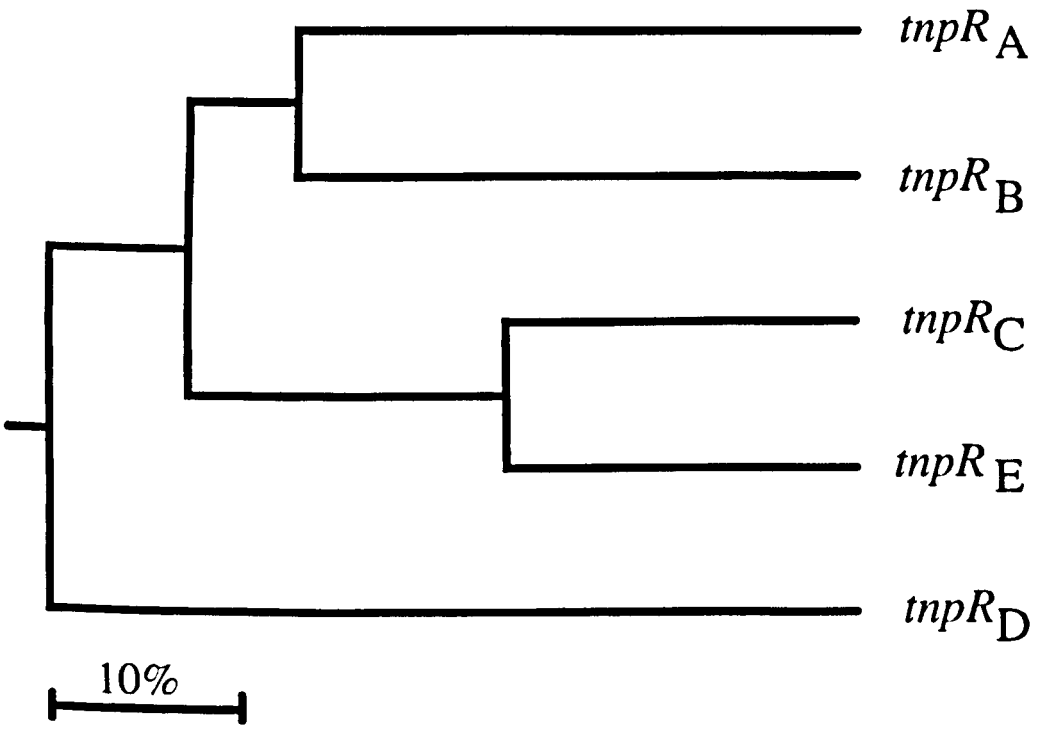


Figure 3.2.4.3. Rooted UPGMA tree of *tnpR* classes A,B,C,D and E from Pearson *et al.*, 1996. *tnpR* classes as described in Figure 3.2.4.1. Bootstrap values (not shown) were between 99 and 100.

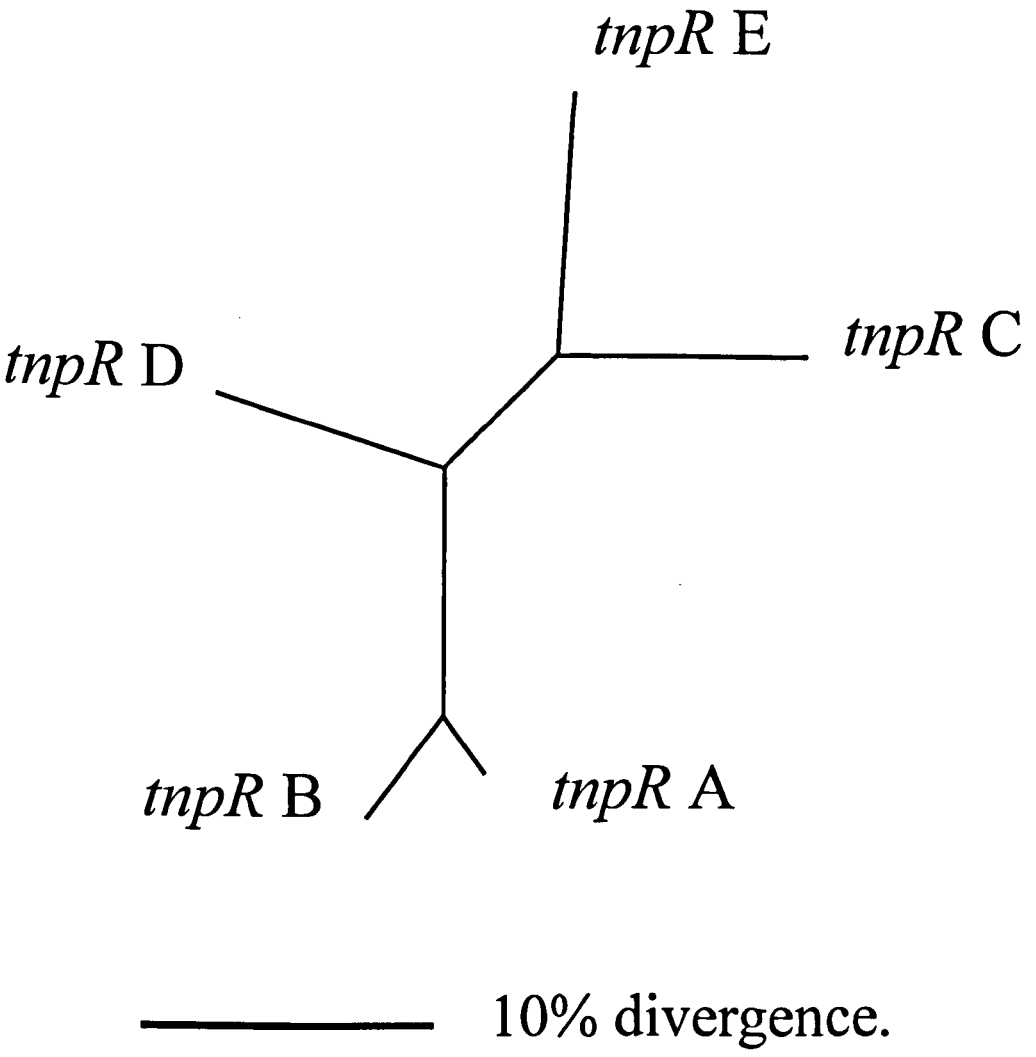


Figure 3.2.4.4. Maximum likelihood tree of complete 372bp resolvase analysis region. *tnpR* classes as described in Figure 3.2.4.1. Bootstrap values (not shown) were between 99 and 100.

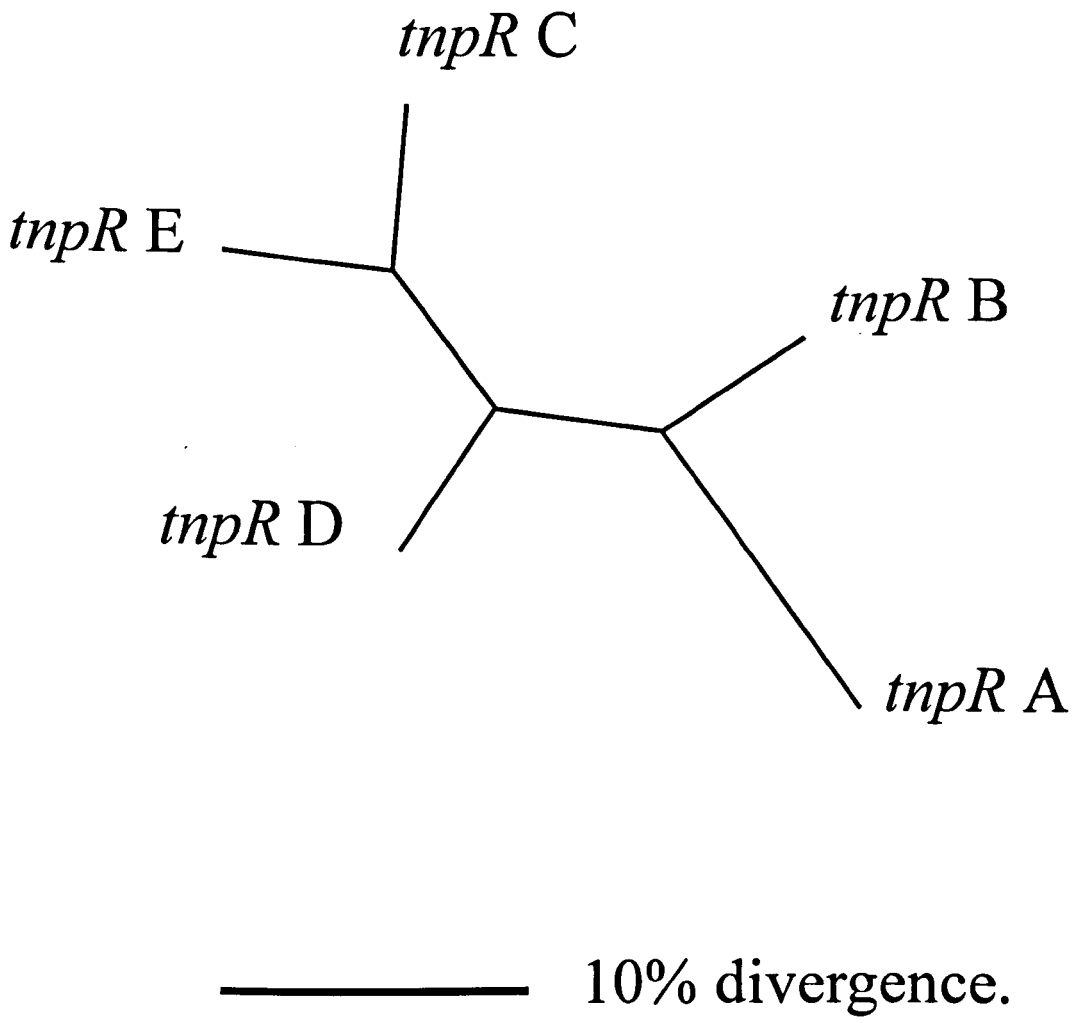


Figure 3.2.4.5. DNA parsimony tree of complete 372bp resolvase analysis region. *tnpR* class groupings comparable to the Maximum likelihood tree (Figure 3.2.4.4). *tnpR* classes as described in Figure 3.2.4.1. Bootstrap values (not shown) were between 99 and 100.

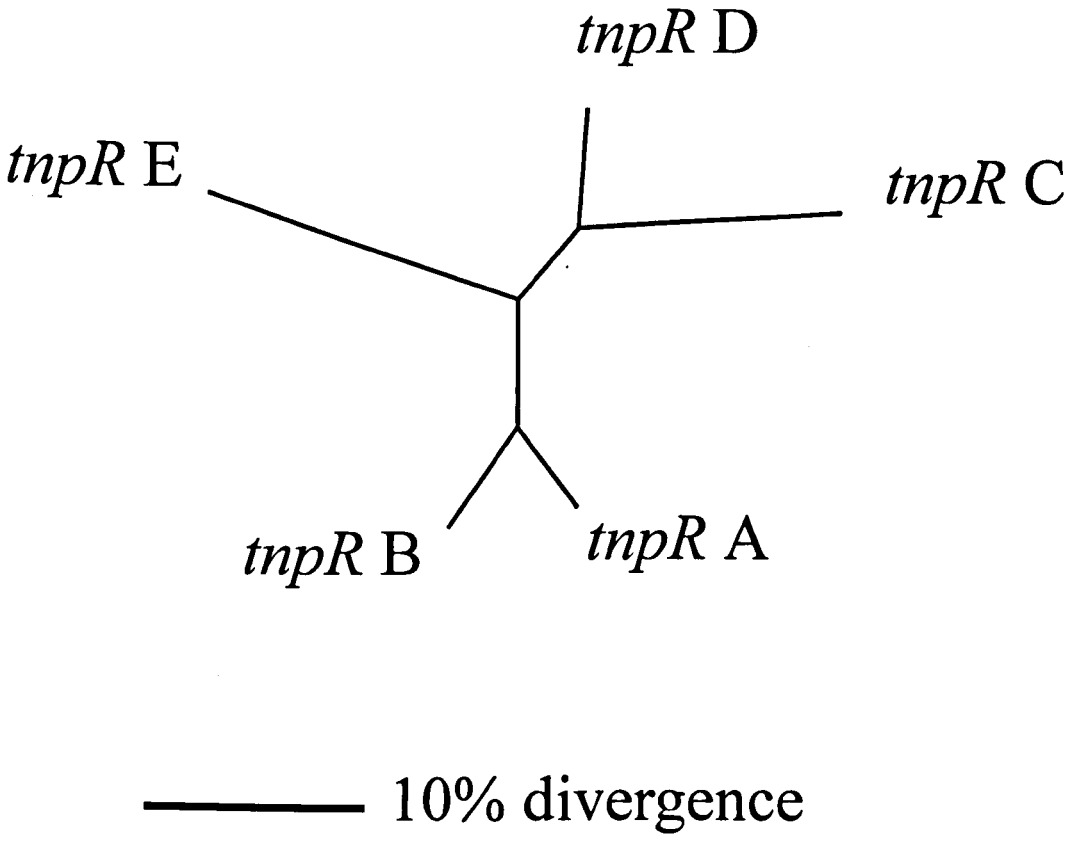


Figure 3.2.4.6. Maximum likelihood tree of the first 186bp of resolvase analysis region. *tnpR* classes as described in Figure 3.2.4.1. Bootstrap values (not shown) were between 99 and 100.

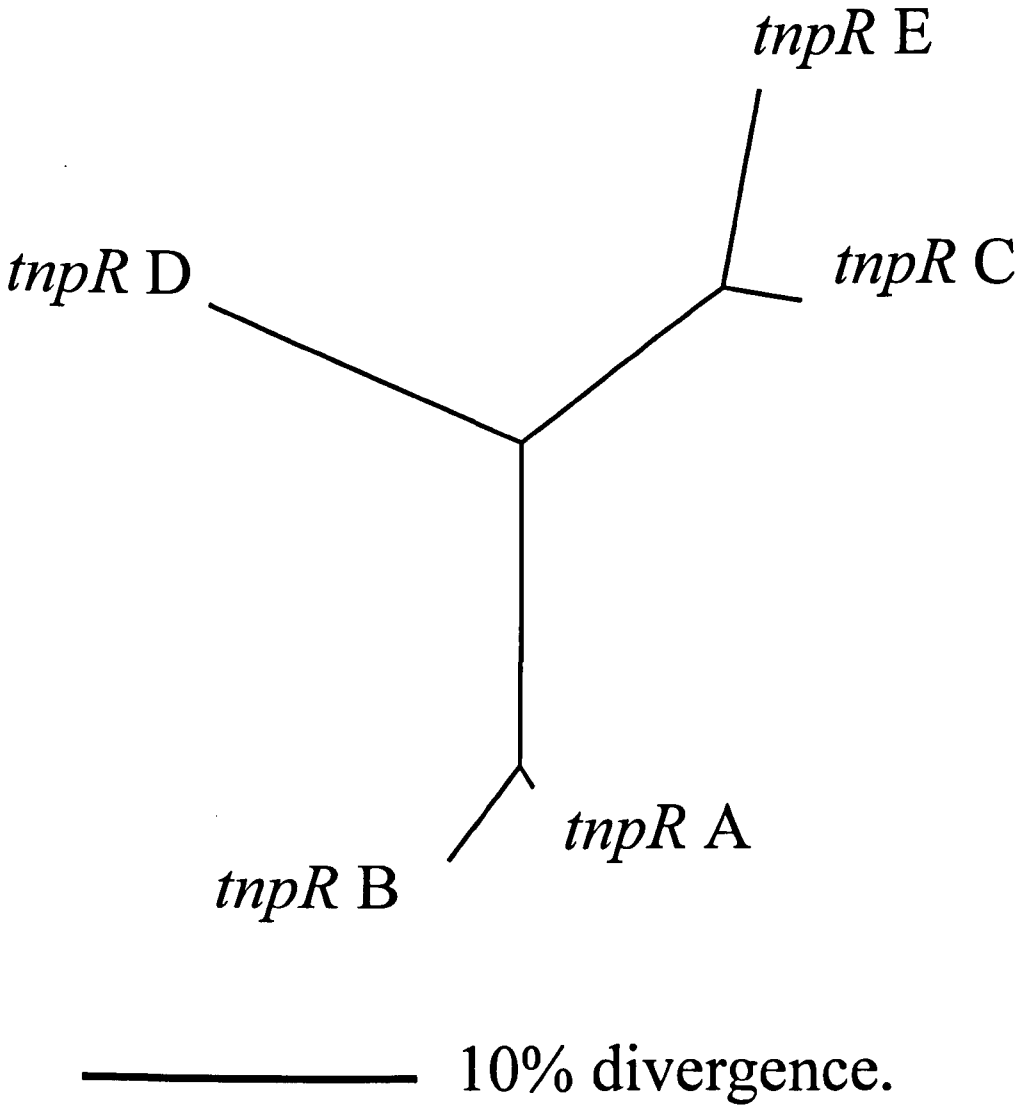


Figure 3.2.4.7. Maximum likelihood tree of the second 186bp of resolvase analysis region. *tnpR* classes as described in Figure 3.2.4.1. Bootstrap values (not shown) were between 99 and 100.

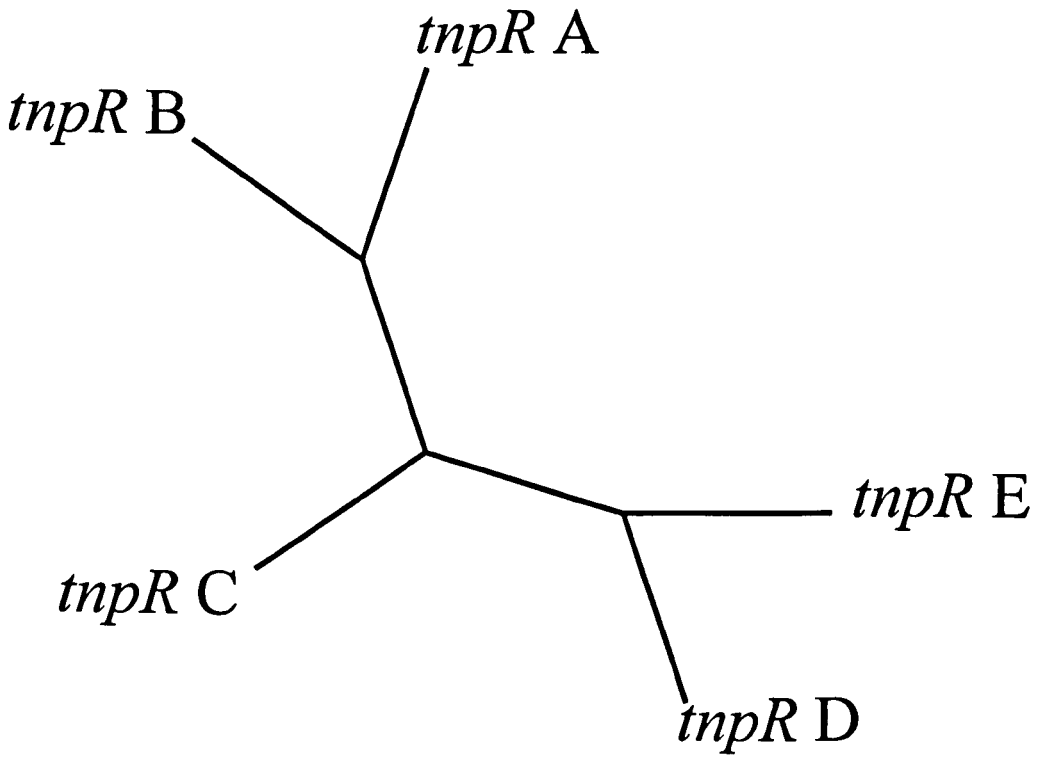


Figure 3.2.4.8. DNA parsimony tree showing 1st 186bp of resolvase analysis region. *tnpR* classes A-E as shown in Figure 3.2.4.1. Bootstrap values (not shown) were between 99 and 100.

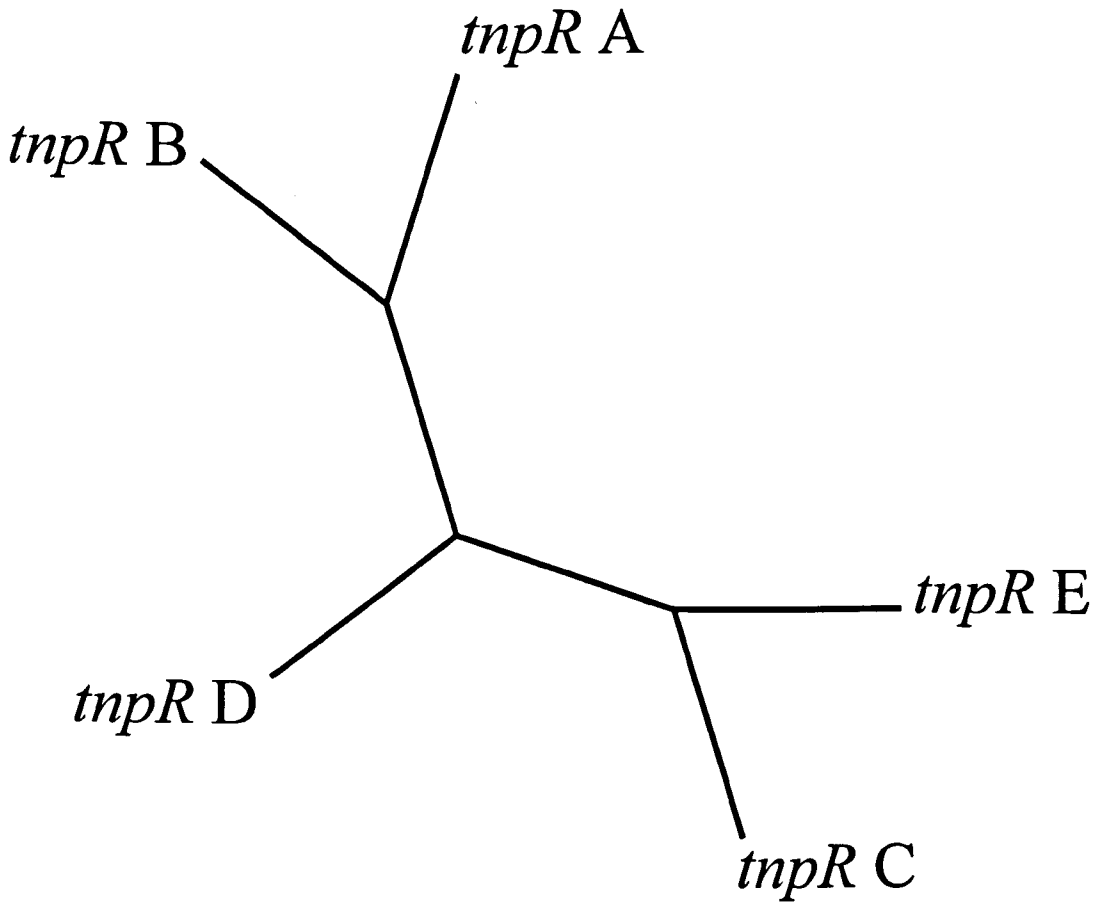


Figure 3.2.4.9. DNA parsimony tree showing 2nd 186bp of resolvase analysis region.*tnpR* classes A-E as shown in Figure 3.2.4.1. Bootstrap values (not shown) were between 99 and 100.

3.2.5. Restriction endonuclease patterns of sequenced analysis regions.

A restriction map of the original *tnpR* A-E sequences was constructed from sequence data using the restriction endonucleases used in the original RFLP study. The sizes of predicted bands generated matched those observed in Pearson *et al*, 1996.The majority of the restriction endonuclease sites observed (sixteen out of twenty) are contained within the second half of the region (Figure 3.2.5.1). As the second half of the sequence contains a larger number of the restriction sites originally used to classify these genes, the second half of the sequence has provided a more accurate representation of the overall analysis region. The dendrogram of the first 186bp still however shows marked structural similarity to that of the overall analysis region with the exception of the class D and E branches. This emphasises the importance of restriction enzyme choice when carrying out RFLP analysis as the position and frequency of restriction sites will affect the RFLP patterns which are observed. DNA parsimony trees of the complete fragment as well as the two halves of sequence were also constructed. Tree topology was identical to those produced using the maximum likelihood method.

3.2.6. Insertion of '93 T2 isolates into unrooted tree.

Two previously uncharacterised strains, '93 T2 7 and '93 T2 12 (Osborn *et al.*, 1995) were chosen for further study. Again, approximately 500bp PCR products from *tnpR* were generated and sequenced. '93 T2 7 DNA sequence was found to be 97.6% similar to Tn*501* and gave an identical restriction endonuclease digestion pattern to Tn*501* placing it in class *tnp*R-D. '93 T2 12 was assigned a new resolvase class, *tnpR* F by virtue of its DNA sequence. This sequence information was used to construct a maximum likelihood tree which is shown in Figure 3.2.6.1.

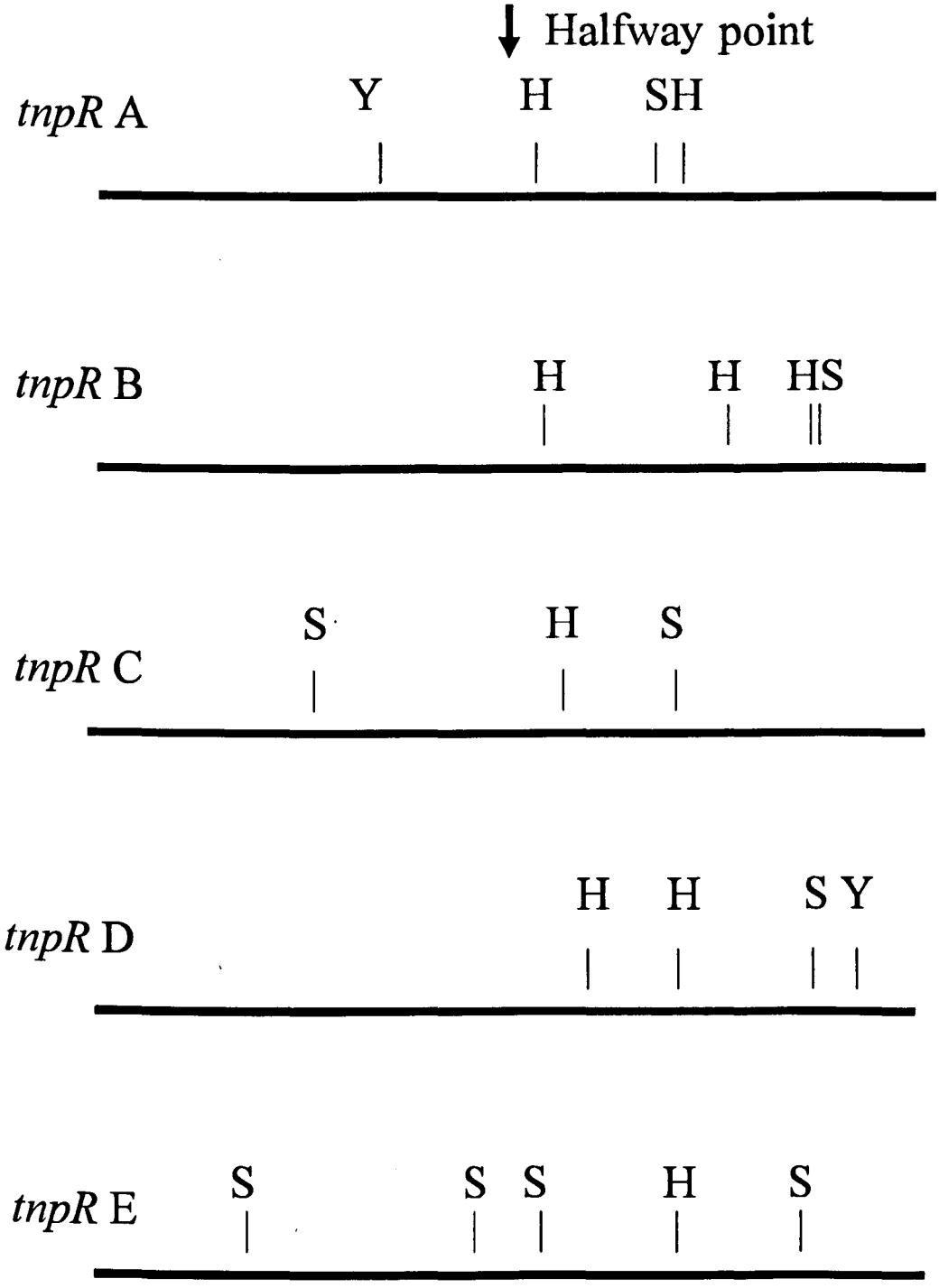


Figure 3.2.5.1. Position of restriction sites used for RFLP analysis in *tnpR* A-E: H: *Hinf*I S: *Sau*3AI Y: *Sty*I (Pearson *et al.*, 1996).

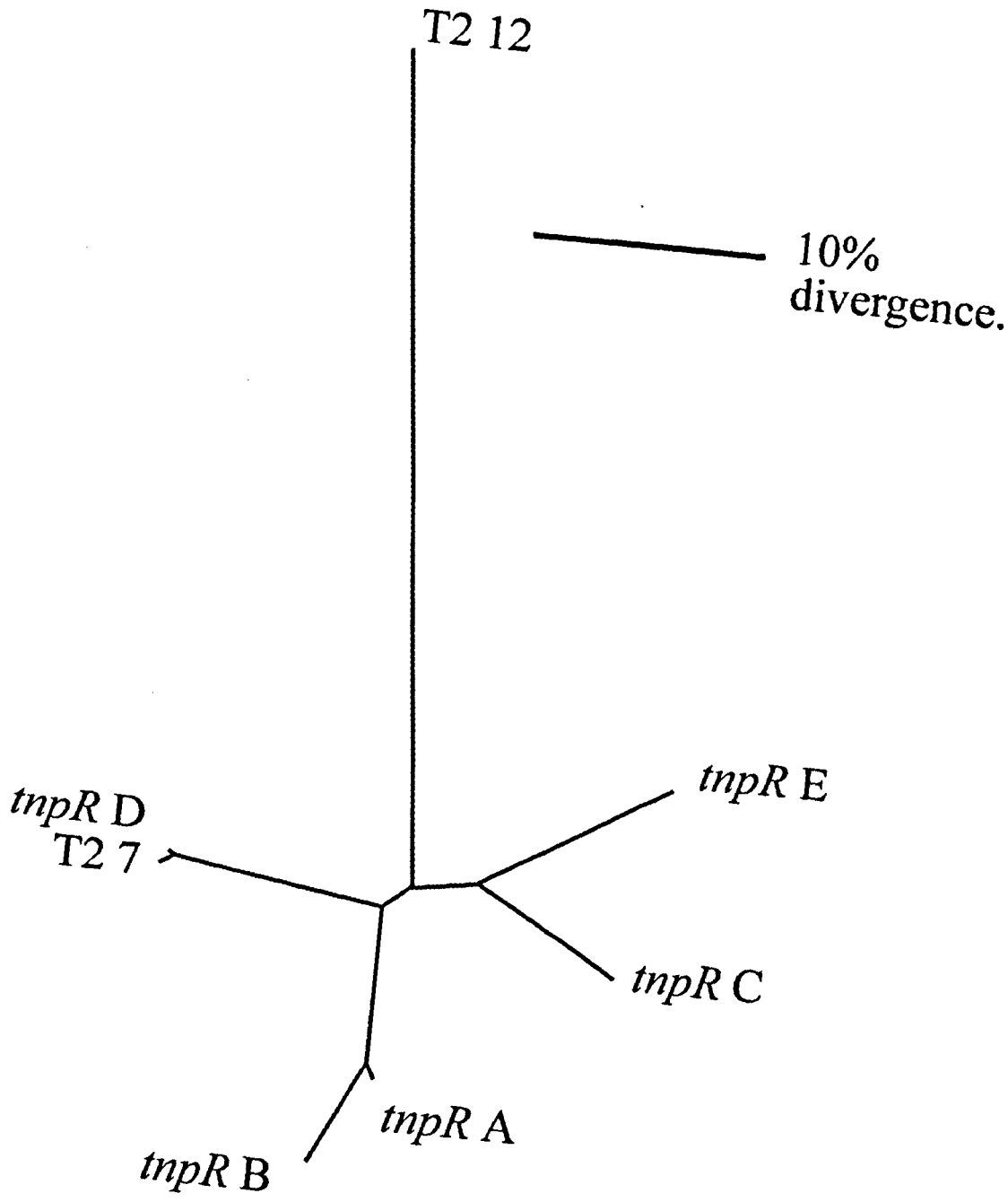


Figure 3.2.6.1. Maximum likelihood tree showing *tnpR* A-E, T2 7 and T2 12. *tnpR* classes as described in Figure 3.2.4.1. Bootstrap values (not shown) were between 99 and 100.

3.2.7. Differences in DNA sequence between classes.

The DNA sequences of the *tnpR* genes from all 8 isolates along with the sequences of Tn*21* and Tn*501* were aligned to observe the nature of the DNA changes leading to the differences between classes. This alignment is shown in Figure 3.2.2.12.

3.2.8. Amino acid sequences.

The partial DNA sequences were translated to obtain partial amino acid sequences from all the *tnpR* genes sequenced. These sequences are shown in Figures 3.2.8.1-8. An alignment of the amino acid sequences from Classes A-E is shown in Figure 3.2.8.9 which also shows the residues which are believed to be involved in the catalytic activity of the resolvase protein (Arnold *et al.*, 1999, Boocock *et al.*, 1995). The DNA sequences in this study showed a number of base changes in the 3rd codon position with conserved 1st and 2nd codon positions based on the presumed reading frame. The partial protein sequence of *tnpR* A-E indicates that a large number of these 3rd base shifts were not translated into amino acid changes. Amongst amino acids which had been altered due to mutation, a significant number are such that an amino acid of comparable physicochemical structure had replaced the original residue. The presumed amino acid sequences obtained from '93 T2 7 and '93 T2 12 suggested that the *tnpR* genes contained within these isolates may be non-functional. The amino acid residues which are believed to be involved in the catalytic activity of this protein remain conserved throughout the 5 different classes of gene (Figure 3.2.8.9.) with the exception of the E124 residue in *tnpR* class B which appears to have undergone a E124X mutation. It has been suggested that DNA rearrangments may be responsible for the majority of polymorphisms that arise (Hall, 1994). However, this does not appear to be the case in this study as could be seen from the DNA sequences, which suggested that point mutations were responsible for the observed differences between RFLP classes.

Figure 3.2.8.1. Amino acid sequence of '93 SO1, Class A *tnpR* gene.

SPXKASGKDT QRPELERLLA FVREGDTVVV HSMDRLARNL DDLRRIVQGL
TQRGVRMEFV KEGLKFTGED SPMANLMLSV MGAFAEFERA LIRERQREGI
VLAKQRGAYR GRKKSLNSEQ IAELKRRVAA NQK

Figure 3.2.8.2. Amino acid sequence of '93 SO2, Class A *tnpR* gene.

SPXKASGKDT QRPELERLLA FVREGDTVVV HSMDRLARNL DDLRRIVQGL
TQRGVRMEFV KEGLKFTGED SPMANLMLSV MGAFAEFERA LIRERQREGI
VLAKQRGAYR GRKKSLNSEQ IAELKRRVAA NQK

Figure 3.2.8.3. Amino acid sequence of '93 SE31, Class B *tnpR* gene.

STSVRYSPIR LPARTRNVPK LERLLAFVRE GDTVVVHSMD RLARNLDDLR
RIVQELTQRG VRMEFVKEGL AFTGDDSPMA NLMLSVMGAF AEFERALYPR
TPARGNRAGQ AARCLPGTKE IAEQRTNCQV ETASRGRRSK N

Figure 3.2.8.4. Amino acid sequence of '93 SE3, Class C *tnpR* gene.

GLSVHRQASG KDTRRPELER LLAFVREGDT VVVHSMDRLA RNLDDLRRLV
QGLTQRGVRI EFLKEHLTFT GEDSPMANLM LSVMGAFAEF ERALIRERQR
EGIALAKQRG AYRGRKKSLS SERIAELRQR F

Figure 3.2.8.5. Amino acid sequence of '93 SE23, Class C *tnpR* gene.

KVFTDKASGK DTRRPELERL LAFVREGRTR FVVHSMDRLA RNLDDLRRLV
QGLTQRGVRI EFLKEHLTFT GEDSPMANLM LSVMGAFAEF ERALIRERQR
EGI

Figure 3.2.8.6. Amino acid sequence of '93 SB29, Class C *tnpR* gene.

SVHRQASGKD TRRPELERLL AFVREGDTVV VHSMDRLARN LDDLRRLVQG
LTQRGVRIEF LKEHLTFTGE GLADGEPDAV GNGRVRRVRT RLDPRAAARG
HRARQAARGL PWQKEIPVVW TLLPNCA

Figure 3.2.8.7. Amino acid sequence of '93 T2 7, Class D *tnpR* gene.

GQGHPAPPGW EALAELRPRR RYSGGAQHGP AGPQPR*PAF AWYRS*LNAA
CASSFLKEGL VFHWRRTRRW PT*CCR*WGP SLSFERALIR ERQREGITLA
KQRGAYRGRK KSLS

Figure 3.2.8.8. Amino acid sequence of '93 T2 12, Class F *tnpR* gene.

PSGPSLDSMT GFVRRRDTVV VS*HGSAWRR NLR*LASASC KSSPSAVLGY
RVRQGKPVPS PGEEFTRWTQ PNAVCSWGRF AQF*RALYPR RQRVRHRVRP
KQRR*PLFGA AKKKD

```
tnpR A  LERLLAFVRE GDTVVVHSMD RLARNLDDLR RIVQGLTQRG VRMEFVKEGL
tnpR C  LERLLAFVRE GDTVVVHSMD RLARNLDDLR RLVQGLTQRG VRIEFLKEHL
tnpR D  LEALLSFVRE GDTVVVHSMD RLARNLDDLR RLVQKLTQRG VRIEFLKEGL
tnpR E  LEALISFART GDTVVVHSMD RLARNLDDLR RIVQTLTQRG VHIEFVKEHL
tnpR B  LERLLAFVRE GDTVVVHSMD RLARNLDDLR RIVQELTQRG VRMEFVKEGL

        ==*=CC=C=* ========== ========== =C==*===== =C*==C==*=


tnpR A  KFTGEDSPMA NLMLSVMGAF AEFERALIRE RQREGIVLAK QRGAYRGRKK
tnpR C  TFTGEDSPMA NLMLSVMGAF AEFERALIRE RQREGIALAK QRGAYRGRKK
tnpR D  VFTGEDSPMA NLMLSVMGAF AEFERALIRE RQREGITLAK QRGAYRGRKK
tnpR E  SFTGEDSPMA NLMLSVMGAF AEFERALIRE RQREGIALAK QRGAYRGRKK
tnpR B  AFTGDDSPMA NLMLSVMGAF AEFERALYRX RQREGIVLAK QRGAYRGRKK

        C========= ========== =======*=* ======C=== ==========
        #


tnpR A  SLNSEQIAEL KRR
tnpR C  SLSSERIAEL RQR
tnpR D  ALSDEQAATL RQR
tnpR E  SLSSERIAEL RQR
tnpR B  SLNSEQIAKL KQR

        C=**=*C=*= C==
```

Figure 3.2.8.9. Translated partial protein sequence of *tnpR* A-E. Symbols indicate the similarity of the amino acids present in that particular position. Amino acids in bold represent those residues which are believed to be involved in the catalytic process.

=: Identical amino acid.
C: Conserved amino acid change. i.e. The amino acids present at this position are of a similar chemical structure.
*: Different amino acid. i.e. The amino acids present at this position are not chemically related.
#: K is different, T,V,S & A are conserved.

3.3. Conclusions.

In this study the relative advantages and disadvantages of DNA sequencing and RFLP typing approaches were studied. The data obtained indicated that DNA sequencing provides a large amount of information over that which can be acquired by RFLP analysis e.g. determination of protein sequences. The use of phylogenetic analysis showed that the different methods of tree construction produced comparable results and also indicated that the choice of restriction endonuclease enzymes for use in RFLP analysis may bias results. The sequence diversity of the *tnpR* genes in these strains was relatively high and comparison of the DNA sequences suggested that diversity may have been due to point mutations, mostly in the assumed third codon position.

CHAPTER 4.

CONSERVATION OF TRANSPOSON STRUCTURES IN MERCURY RESISTANT SOIL BACTERIA.

4.1. Introduction.

Previous studies on a collection of 39 mercury resistant bacteria ('93 isolates) have concentrated on the study of sequence diversity within the *mer* and *tnp* genes. The diversity has been studied by RFLP analysis and by DNA sequencing (Holt *et al.*, 1996, Osborn *et al.*, 1993, Osborn *et al.*, 1995, Pearson *et al.*, 1996). A significant conclusion from these studies was that recombination between transposon genes and between transposon and *mer* genes, was common. The identity of the '93 isolates has been determined by Osborn *et al.* (Osborn *et al.*, 1993) using API (Table 2.2.1). The species identified included *Enterobacter cloacae*, *Alcaligenes faecalis*, *Acinetobacter calcoceticus*, *Klebsiella oxytoca*, *Agrobacterium radiobacter*, *Aeromonas spp* and *Pseudomonas spp* (Osborn *et al.*, 1993). Prior to this study no information was available regarding the structural diversity of the transposon genes carried in these isolates. The presence of plasmids had not been studied and as such, the location of the transposons contained within these isolates was not known. In this study, the relationship between the transposition genes and their associated *mer* operons has been investigated further. The number of mercury resistant strains has been expanded by the isolation of 85 new isolates ('96 isolates). The genus/species of these isolates was not characterised in this study as, due to the diverse taxonomy of the '93 isolates it was felt that API typing of the '96 isolates would not yield any further information. The presence of transposon associated genes has been studied in all 124 isolates, and PCR reactions have been carried out to determine the presence of integrons, and the size of any genes present within the *rhs* of those integrons. The presence and

structural organisation of the *mer* genes in the isolates has been studied, as have the relative orientations of the *tnp* genes and *mer* operons in the isolates. The presence and approximate size of plasmids present within the isolates was also determined to ascertain whether any correlation existed between the structural diversity of the transposon and the plasmid diversity.

The primary aim of this study was to determine the presence and structural diversity of the *tnpA* and *tnpR* genes contained within the isolates and to seek evidence, or otherwise, of genetic rearrangements and/or recombination occurring in the transposon associated region.

4.2. Results and Discussion.

4.2.1. Isolation of '96 isolates.

To allow the study of a greater number of bacterial strains, samples of soil were taken from 3 of the 4 sites used by Osborn *et al.*, 1993, namely the SO and SE sites at Fiddlers Ferry and the SB site at Salterbrook Bridge. (Figures 2.2.1, 2.2.2 and 2.2.3). The T2 site was not available at the time of sampling. Mercury resistant bacteria were isolated using the same method as used Osborn *et al.*, 1993. The relative numbers of bacteria isolated on $HgCl_2$ selective media compared to non-selective media are shown in Table 4.2.1.1. Thirty isolates were chosen from the SO site, 31 from the SE site and 24 from the SB site. All these isolates were Gram stained and found to be Gram-negative. These strains were designated the '96 isolates in order to distinguish them from the original 39 strains isolated by Osborn *et al.*, 1993, which were designated the '93 isolates.

Table 4.2.1.1. Numbers of bacteria isolated from SO, SE and SB sites on selective and non-selective media.

Isolation site	Number of bacteria g soil^{-1}	
	No $HgCl_2$	$HgCl_2$
SO	4.80×10^7	1.48×10^5
SE	8.20×10^7	6.77×10^5
SB	1.53×10^5	2.60×10^2

Soil samples were taken from three sites, SO, SE and SB (Section 2.2). Soil was sampled from the top two inches of soil at each site. Bacteria were then isolated as described in Section 2.5 (Osborn *et al.*, 1993). For the determination of bacterial numbers, soil solutions were plated out in triplicate on both selective media (containing 50μM $HgCl_2$) and on standard LA plates. After growth had occured, appropriate dilutions were counted (plates containing between 30 and 300 bacterial colonies) and the mean number of colonies was calculated.

4.2.2. Design of oligonucleotide primers.

4.2.2.1. Transposase and resolvase gene primers.

An initial DNA sequence database search was carried out to obtain nucleotide sequences from as many transposase genes as were available. This search yielded 20 sequences, the accession numbers of which are shown in Table 2.21.1. Due to the highly diverse nature of these sequences it proved impossible to design one set of primers capable of detecting all 20 sequences. A DNA parsimony phylogenetic tree was constructed using 2.6 kb of aligned sequence to allow the relative similarities of the sequences to be observed and to assist in primer design (Figure 4.2.2.1.1). Three sets of primers were designed, these are shown in Figure 2.15.1 along with all other primers used in this study:

Primers 2501 and 2850 were designed to concensus sequences from the *tnpA* genes of Tn*1*, Tn*3*, Tn*21*, Tn*501*, Tn*1000*, Tn*1721*, Tn*2501*, Tn*3926*, Tn*5036* and Tn*5401* (Allmeier *et al.*, 1992, Baum, 1994, Bennett *et al.*, 1978, Bishop and Sherrat, 1984, Broom *et al.*, 1995, de la Cruz and Grinsted, 1982, Lett *et al.*, 1985, Maekawa *et al.*, 1993, Michiels *et al.*, 1987, Yurieva *et al.*, 1997). Primers pos2000 and pos2400 were designed to concensus sequences from the *tnpA* genes of Tn*917*, Tn*1546*, Tn*4430* and Tn*5422* which were all isolated from Gram-positive bacteria (Arthur *et al.*, 1993, Lebrun *et al.*, 1994, Lereclus *et al.*, 1986, Shaw and Clewell, 1985). Primers 62351 and 62601, were designed to concensus sequences from the *tnpA* genes of Tn*163*, Tn*4652*, Tn*5041*, Tn*5393*, and Tn*5403*. (Chiou and Jones, 1993, Kholodii *et al.*, 1997, Mäe and Heinaru, 1994, Rinkel *et al.*, 1994, Ulrich and Pühler, 1994).

All of these primers were also suitable for use in DGGE experiments, as they are non degenerate and produce PCR products of less than 400bp. To ascertain the specificity of these primers PCR reactions were carried out using control strains (strains listed in Table 2.1.1). Primer pairs 2501/2850 and pos2000/pos2400 produced PCR products of the correct size when PCR reactions were carried out using DNA

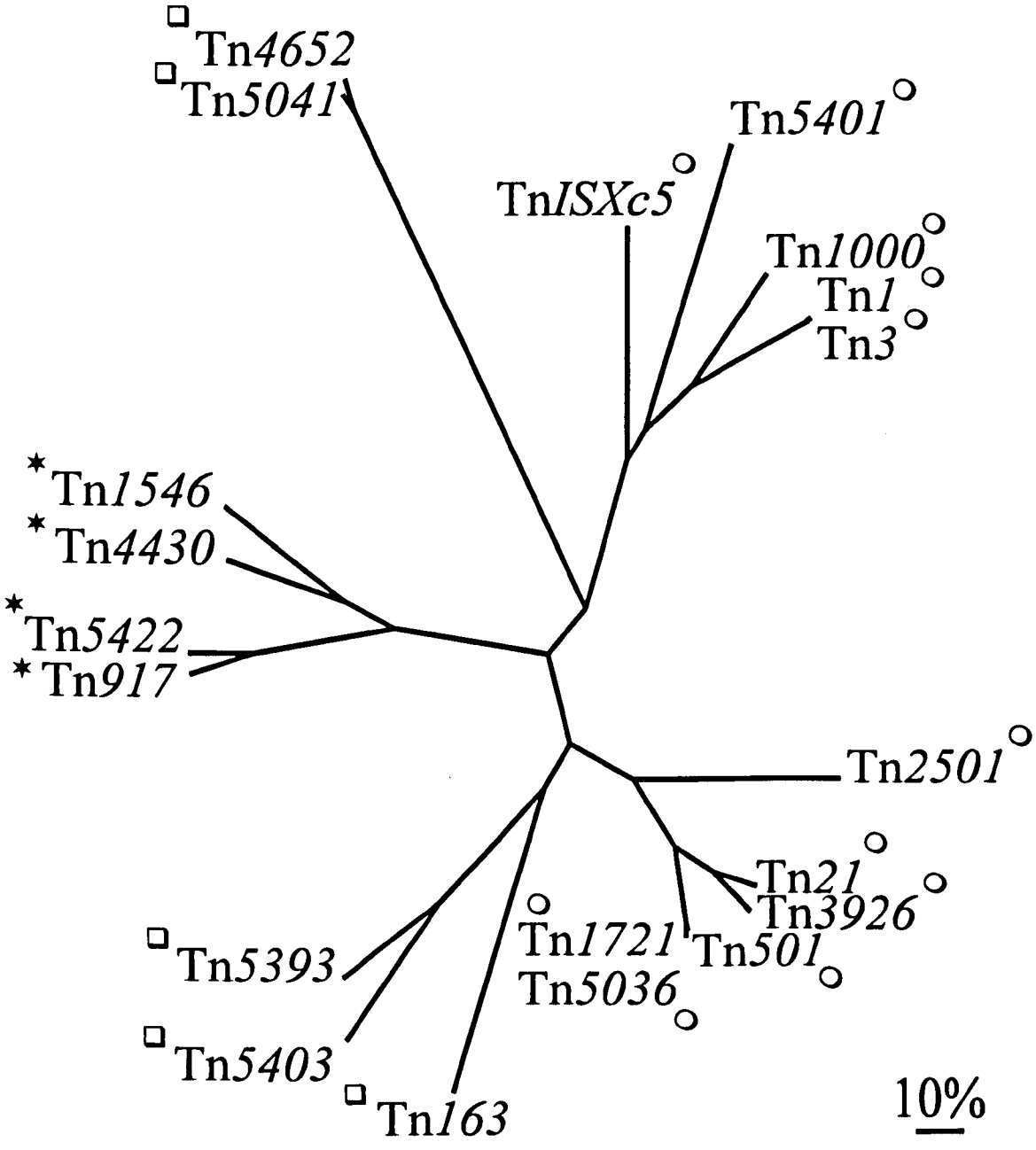


Figure 4.2.2.1.1. DNA parsimony tree showing relationships between all available DNA sequences from *tnpA* genes. Bootstrap values (not shown) were between 97 and 100.

Primer sets are indicated by symbols:
□: Primers 2501/2850
○: Primers pos2000/pos2400
✶: Primers 62351/62601 (not studied).

isolated from appropriate control strains as templates. The 62351/62601 primer set failed to yield any PCR products using DNA from control strains as PCR template. However due to the highly divergent nature of these *tnpA* genes, this is perhaps unsurprising (Figure 4.2.2.1.1). Primer pair 62351 and 62601 were not used in this study.

Other primers specific for *tnpA* and *tnpR* genes were also used in this study (Table 2.15.1). Primers 1406 and 2638 and primers 501R1/C and 501R2/C have been previously described, (Pearson *et* al., 1996) having been designed to complement the DNA sequences of the Tn*21* and Tn*501* *tnpA* and *tnpR* sequences respectively. Primer 950 was designed to correspond to the sequence approximately 930bp from the start codon of the *tnpA* gene of Tn*501*/Tn*21*. To allow detection of DNA sequences specific to Tn*3* *tnpA* and *tnpR* genes, primers tn3A and tn3R were designed to the *tnpA* and *tnpR* sequenes of Tn*3* respectively.

4.2.2.2. Integron primers.

To allow detection of integron structures, two sets of oligonucleotides were designed (Table 2.15.1). Primers int21A and int21B were designed to the integrase gene of Tn*21* and were used as an indicator of integron gene presence (Young, H. K. and Rosser, S. J. Personal communication). Primers 4127 and 4128 allow the size of any gene cassettes inserted within the *rhs* of an integron to be determined (Young, H. K. and Rosser, S. J. Personal communication).

4.2.2.3. *mer* operon primers.

Structural variations in the *mer* operons of these strains was studied using the following primers. Primers mercP and mercA complement the DNA sequence in *merP* and *merA* respectively allowing the presence of any genes present inbetween *merP* and *merA* to be determined. Primer mercD was designed to be homologous to the *merD* sequence present in Tn*21* and Tn*501* and allowed the structural diversity between the *mer* operon and the *tnp* genes to be studied.

4.2.3. Presence of *tnpA* genes.

One hundred and twenty-four mercury resistant strains were examined for the presence of *tnpA* genes using 2 sets of primers; pos2000/pos2400 and 2501/2850. No strain produced a PCR product using primers pos2000 and pos2400, despite *tnpA* amplification from control strains. Given that each of the 124 mercury resistant strains are Gram-negative and the transposase genes to which these primers were designed were originally found in Gram-positive bacteria, this is perhaps unsurprising.

Seventy-five strains produced a PCR product using primers 2501 and 2850 (Tables 4.2.3.1 and 4.2.3.2). The presence of *tnpA* genes in twenty-two out of thirty of the '93 SO, SE and SB isolates has been previously demonstrated by hybridisation to *tnpA* probes from both Tn*21* and Tn*501* (Pearson *et al*, 1996). Those isolates which were previously identified as containing *tnpA* sequences by Pearson *et al.*, 1996 also produced PCR products using primers 2501 and 2850. To allow the comparison of the '93 and '96 isolates, PCR reactions were carried out on all 124 isolates using the primers 1406 and 2638, used in the characterisation of the '93 isolates (Pearson *et al*, 1996). All 75 strains produced a PCR product using these primers (Tables 4.2.3.1 and 4.2.3.2.). These PCR products hybridised to the corresponding *tnpA* PCR products from both Tn*21* and Tn*501*. The number of isolates in this study containing transposase genes is high compared to similar studies on marine and clinical isolates which indicated the presence of the gene in 3/40 marine isolates and 116/807 clinical isolates (Dahlberg and Hermansson, 1995, Zuhlsdorf and Weidermann, 1992). These high numbers may have been caused by the initial selection for resistance to $HgCl_2$. This initial selection may also explain the higher numbers of *tnpA* containing strains observed in the '93 isolates (Table 4.2.3.2).

Forty-nine of the strains did not produce a *tnpA* PCR product; 22 of these were in the '96 SB group. This may indicate the lack of this gene in these isolates, or may be due to sequence variation at the primer annealing sites. The differences observed between the '93 and '96 SB isolates may be due to the additional selection of the '93 isolates for their ability to hybridise to a *merRTΔP* probe (Osborn *et al.*, 1993). This

Table 4.2.3.1. Table of results for '93 and '96 isolates.

Key:

Location of *tnpA* gene:

Plasmid:	Gene located on plasmid
Chromosome:	Gene located on chromosome
-:	Strain either contains no plasmid and/or no *tnpA* gene

No of plasmids:

Large plasmids were defined as those which electrophoresed above the bacterial chromosome, whereas small plasmids electrophoresed below the chromosomal band. This is discussed further in Section 4.2.4.

tnpA/*tnpR* arrangement:

Tn*21/501*:	Strain contains an arrangement of genes similar to that found in the Tn*21* subgroup of transposons
No *tnpR*:	Strain did not produce a PCR product with any *tnpR* primers
UNKNOWN:	Strain contains both *tnpA* and *tnpR* genes in an undetermined arrangement
-:	Strain contains no *tnpA* gene

merD/*tnpR* arrangement:

1.3:	1.3kb PCR product produced
2:	2kb PCR product produced
-:	No PCR product produced

merC:

YES:	Strain produces PCR product which blots to *merC* probe.
NO:	Strain produces a PCR product which blots to Tn*501* PCR product. i.e. No *merC* gene
UNKNOWN:	Strain produces a PCR product of similar size to Tn*21*, but does not blot to a *merC* probe.
-:	Strain produces no PCR product

STRAIN	*tnpA* gene	No of plasmids	Location of *tnpA* gene	*tnpA/tnpR* arrangement	Integrase gene	Integron insert size	*merD/tnpR* arrangement	*merC*
'93 SO1	+	1	Plasmid	Tn*21/501*	-	-	2	YES
'93 SO2	+	1	Plasmid	Tn*21/501*	-	-	2	YES
'93 SO3	+	1	Plasmid	Tn*21/501*	-	-	2	YES
'93 SO5	+	1	Plasmid	Tn*21/501*	-	-	2	YES
'93 SO6	+	1	Plasmid	Tn*21/501*	-	-	2	YES
'93 SO7	+	1	Plasmid	Tn*21/501*	-	-	2	YES
'93 SO8	+	1	Plasmid	Tn*21/501*	-	-	2	YES
'93 SO9	+	1	Plasmid	Tn*21/501*	-	-	2	YES
'93 SO12	+	1	Plasmid	No *tnpR*	-	-	-	YES
'93 SO13	+	1	Plasmid	Tn*21/501*	-	-	2	YES
'93 SE3	+	2	Plasmid	Tn*21/501*	-	-	-	-
'93 SE6	-	1	-	-	-	-	-	YES
'93 SE9	-	1 small	-	-	-	-	1.3	YES
'93 SE11	-	1	-	-	-	-	-	YES
'93 SE12	-	1 & 1 small	-	-	-	-	-	-
'93 SE18	+	1 & 1 small	Plasmid	Tn*21/501*	-	-	1.3	-
'93 SE20	+	1	Plasmid	Tn*21/501*	-	-	1.3	YES
'93 SE23	+	2	Plasmid	Tn*21/501*	-	-	-	-
'93 SE31	+	1	Plasmid	Tn*21/501*	-	-	1.3	UNKNOWN
'93 SE35	+	1	Plasmid	Tn*21/501*	-	-	-	-
'93 SB2	+	1	Plasmid	Tn*21/501*	-	-	2	NO
'93 SB3	+	2	Plasmid	Tn*21/501*	-	-	2	NO
'93 SB4	+	1	Plasmid	Tn*21/501*	-	-	2	NO
'93 SB5	-	0	-	-	-	-	-	-
'93 SB8	-	1	-	-	-	-	2	NO
'93 SB12	-	2	-	-	-	-	1.3	-
'93 SB13	+	1 & 3 small	Plasmid	Tn*21/501*	-	-	1.3	-
'93 SB22	+	1	Plasmid	Tn*21/501*	-	-	2	NO
'93 SB24	-	0	-	-	-	-	-	-
'93 SB29	+	1	Plasmid	Tn*21/501*	-	-	2	-

'93 T27	+	2	Plasmid	Tn*21/501*	-	-	1.3	YES
'93 T212	+	1 & 2 small	Plasmid	Tn*21/501*	-	-	-	-
'93 T217	+	1	Plasmid	Tn*21/501*	-	-	1.3	YES
'93 T219	+	0	Chromosomal	Tn*21/501*	-	-	1.3	YES
'93 T223	+	0	Chromosomal	Tn*21/501*	-	-	-	YES
'93 T237	+	0	Chromosomal	Tn*21/501*	+	1.2kb	-	UNKNOWN
'93 T238	+	1	Plasmid	Tn*21/501*	-	-	1.3	UNKNOWN
'93 T241	+	0	Chromosomal	Tn*21/501*	-	-	-	YES
'93 T246	+	0	Chromosomal	Tn*21/501*	-	-	-	-
'96 SO1	-	1	-	-	-	-	-	-
'96 SO2	+	1	Plasmid	Tn*21/501*	-	-	-	-
'96 SO3	-	1	-	-	-	-	-	-
'96 SO4	-	2	-	-	-	-	-	-
'96 SO5	+	1	Plasmid	Tn*21/501*	-	-	-	-
'96 SO6	+	1	Plasmid	Tn*21/501*	-	-	-	-
'96 SO7	+	1 & 1 small	Plasmid	Tn*21/501*	-	-	-	-
'96 SO8	+	1 & 1 small	Plasmid	Tn*21/501*	-	-	-	-
'96 SO9	-	1	-	-	-	-	-	-
'96 SO10	-	1 & 1 small	-	-	-	-	-	-
'96 SO11	+	1 & 1 small	Plasmid	Tn*21/501*	-	-	-	-
'96 SO12	+	0	Chromosomal	Tn*21/501*	-	-	-	-
'96 SO13	-	1	-	-	-	-	-	-
'96 SO14	-	2 small	-	-	-	-	-	-
'96 SO15	+	1	Plasmid	Tn*21/501*	-	-	-	-
'96 SO16	+	1	n/d	Tn*21/501*	-	-	-	-
'96 SO17	+	1	Plasmid	Tn*21/501*	-	-	-	-
'96 SO18	+	1	Plasmid	Tn*21/501*	-	-	-	-
'96 SO19	-	1	-	-	-	-	-	-
'96 SO20	+	2	n/d	Tn*21/501*	-	-	-	-
'96 SO21	+	2	Plasmid	Tn*21/501*	-	-	-	-
'96 SO22	-	0	-	-	-	-	-	-
'96 SO23	-	1	-	-	-	-	-	-

STRAIN	*tnpA* gene	No of plasmids	Location of *tnpA* gene	*tnpA/tnpR* arrangement	Integrase gene	Integron insert size	*merD/tnpR* arrangement	*merC*
'96 SO24	+	1 & 1 small	Plasmid	Tn*21/501*	-	-	-	-
'96 SO25	+	1 small	n/d	Tn*21/501*	-	-	-	-
'96 SO26	+	1 & 2 small	n/d	Tn*21/501*	-	-	-	-
'96 SO27	+	1	Plasmid	Tn*21/501*	-	-	-	-
'96 SO28	+	0	Chromosomal	Tn*21/501*	-	-	-	-
'96 SO29	+	1	Plasmid	Tn*21/501*	-	-	-	-
'96 SO30	+	0	Chromosomal	Tn*21/501*	-	-	-	-
'96 SE1	+	1	Plasmid	Tn*21/501*	-	-	1.3	YES
'96 SE2	-	1	-	-	-	-	-	YES
'96 SE3	+	3	Plasmid	Tn*21/501*	-	-	1.3	YES
'96 SE4	+	2 & 2 small	Plasmid	Tn*21/501*	-	-	1.3	YES
'96 SE5	-	3 small	-	-	-	-	-	-
'96 SE6	+	0	Chromosomal	Tn*21/501*	+	1.1kb	-	-
'96 SE7	-	2 small	-	-	-	-	-	YES
'96 SE8	+	1	Plasmid	Tn*21/501*	-	-	1.3	YES
'96 SE9	+	1 & 4 small	Plasmid	Tn21/501	+	1.4kb	1.3 & 2	YES
'96 SE10	+	0	Chromosomal	No *tnpR*	-	-	-	-
'96 SE11	-	1	-	-	-	-	-	-
'96 SE12	+	1	Plasmid	Tn*21/501*	-	-	1.3	YES
'96 SE13	+	1 & 1 small	Plasmid	Tn*21/501*	-	-	1.3	YES
'96 SE14	+	1 & 1 small	Plasmid	Tn*21/501*	-	-	-	-
'96 SE15	+	2 & 2 small	Plasmid	Tn*21/501*	+	4.5kb	-	-
'96 SE16	-	1 & 1 small	-	-	-	-	-	-
'96 SE17	-	2 small	-	-	-	-	-	YES
'96 SE18	+	2	Plasmid	Tn*21/501*	-	-	1.3	YES
'96 SE19	+	1 & 2 small	Plasmid	UNKNOWN	+	1.3kb	-	-
'96 SE20	-	2	-	-	-	-	-	YES
'96 SE21	-	1 & 1 small	-	-	-	-	-	-
'96 SE22	+	2 & 2 small	Plasmid	No *tnpR*	-	-	-	-
'96 SE23	+	0	Chromosomal	Tn*21/501*	-	-	-	-

'96 SE24	+	1	Plasmid	Tn*21/501*	-	-	1.3	YES
'96 SE25	-	1 & 2 small	-	-	-	-	-	-
'96 SE26	+	2 small	Plasmid	Tn*21/501*	-	-	-	NO
'96 SE27	+	1 & 3 small	Plasmid	Tn*21/501*	-	-	-	YES
'96 SE28	+	1	Plasmid	Tn*21/501*	-	-	-	NO
'96 SE29	+	1 & 1 small	Plasmid	Tn*21/501*	-	-	-	YES
'96 SE30	+	1	Plasmid	UNKNOWN	-	-	-	-
'96 SE31	+	2 & 2 small	Plasmid	Tn*21/501*	-	-	-	YES
'96 SB1	-	1	-	-	-	-	-	-
'96 SB3	-	1	-	-	-	-	-	-
'96 SB4	-	0	-	-	-	-	-	-
'96 SB5	-	0	-	-	-	-	-	-
'96 SB7	-	1	-	-	-	-	-	-
'96 SB8	-	1 small	-	-	-	-	-	-
'96 SB9	-	0	-	-	-	-	-	YES
'96 SB11	-	0	-	-	-	-	-	YES
'96 SB12	-	0	-	-	-	-	-	YES
'96 SB13	-	0	-	-	-	-	-	YES
'96 SB14	+	1 & 1 small	Plasmid	No *tnpR*	-	-	-	-
'96 SB15	-	1	-	-	-	-	-	-
'96 SB16	-	3 small	-	-	-	-	-	YES
'96 SB17	-	1 small	-	-	-	-	-	-
'96 SB18	-	1 small	-	-	-	-	-	YES
'96 SB19	-	0	-	-	-	-	-	YES
'96 SB20	-	0	-	-	-	-	-	-
'96 SB21	+	1	Plasmid	Tn*21/501*	-	-	-	-
'96 SB22	-	1	-	-	-	-	-	-
'96 SB24	-	0	-	-	-	-	-	YES
'96 SB25	-	1	-	-	-	-	-	-
'96 SB28	-	1	-	-	-	-	-	-
'96 SB31	-	1	-	-	-	-	-	-
'96 SB37	-	0	-	-	-	-	-	YES

Table 4.2.3.2. Presence of *tnpA* gene and arrangement of *tnpA* and *tnpR* genes. Numbers represent the number of strains containing the described characteristic.

ISOLATION GROUP	No of strains in group	*tnpA* gene	*tnpA/tnpR* arrangement		
			Tn*21*/Tn*501*	No *tnpR*	UNKNOWN
'93 SO	10	10	9	1	0
'93 SE	10	6	6	0	0
'93 SB	10	6	6	0	0
'93 T2	9	9	9	0	0
'96 SO	30	20	20	0	0
'96 SE	31	22	18	2	2
'96 SB	24	2	1	1	0
TOTAL	124	75	69	4	2

would select for those strains which contain *mer* genes similar to Tn*501* and which are therefore potentially more likely to contain associated transposons.

4.2.4. Plasmid extractions.

Several different plasmid extraction techniques were tested in order to allow isolation of plamids from these environmental isolates (Birnboim and Doly, 1979., Holmes and Quigley, 1981., Kado and Lui, 1981., Wheatcroft and Williams, 1981 and Olsen *et al.*, 1990). The methods of Birnboim and Doly, 1979., Kado and Lui, 1981 and Holmes and Quigley, 1981 were found to be unsuitable for the isolation of large plasmids. The techniques described by Olsen *et al.*, 1990 and Wheatcroft and Williams, 1981 were used in this instance as they allow the isolation of large plasmids from a variety of different host bacteria.

Plasmid extractions were carried out on the collection of 124 isolates, and the samples were analysed on agarose gels to allow the visualisation of plasmid DNA. Results are shown in Table 4.2.3.1. One hundred of the isolates were found to contain plasmids, with both large plasmids (which electrophoresed above the chromosomal DNA band) and small plasmids (which electrophoresed below the chromosomal DNA band) being identified (Figures 4.2.4.1. and 4.2.4.2.). This classification system used the assumption that those plasmids which electrophored above the chromosomal band on a gel were classified as "large", however in certain circumstances these plasmids may not be "large" due to factors such as running forms of plasmid DNA and the migration of the chromosomal DNA band. This is also true of the assumption that small plasmids will electrophorese below the chromosomal band on a gel. Of the 75 strains containing *tnpA* genes, 64 were found to contain plasmids, while 11 apparently did not. Whilst large and small plasmids were observed in both the '93 and '96 collections, the '96 isolates contained a higher proportion of smaller plasmids. No incompatibility group data was available for these plasmids, however recent studies suggest that mercury resistance plasmids isolated from the environment do not

Figure 4.2.4.1. Agarose gel showing plasmid DNA from a selection of '93 and '96 isolates. DNA was extracted by the method of Olsen *et al.*, 1990. The chromosomal DNA band is shown (Ch'somal). Plasmids which electrophoresed above this chromosomal DNA band were classified as large plasmids and those which electrophoresed below the chromosomal DNA band were classified as small plasmids. This is discussed further in Section 4.2.4. Examples of both large and small plasmids are indicated (Lanes 5 and 22). Size markers are also shown. A number of different plasmid extractions were used to ascertain the presence of plasmids in these strains due to the unclear nature of some of the data. This agarose gel represents one such experiment.

Lane 1	kb ladder
2	Tn*21*
3	Tn*501*
4	UWC1
5	'93 SO1
6	'93 SE3
7	'93 SE18
8	'93 SE20
9	'93 SE23
10	'93 SE31
11	'93 SE35
12	'93 SB4
13	'93 SB13
14	'93 SB29
15	'93 T2 7
16	'93 T2 12
17	'93 T2 17
18	'93 T2 38
19	'96 SE3
20	'96 SE4
21	'96 SE8
22	'96 SE9
23	'96 SE12
24	'96 SE13

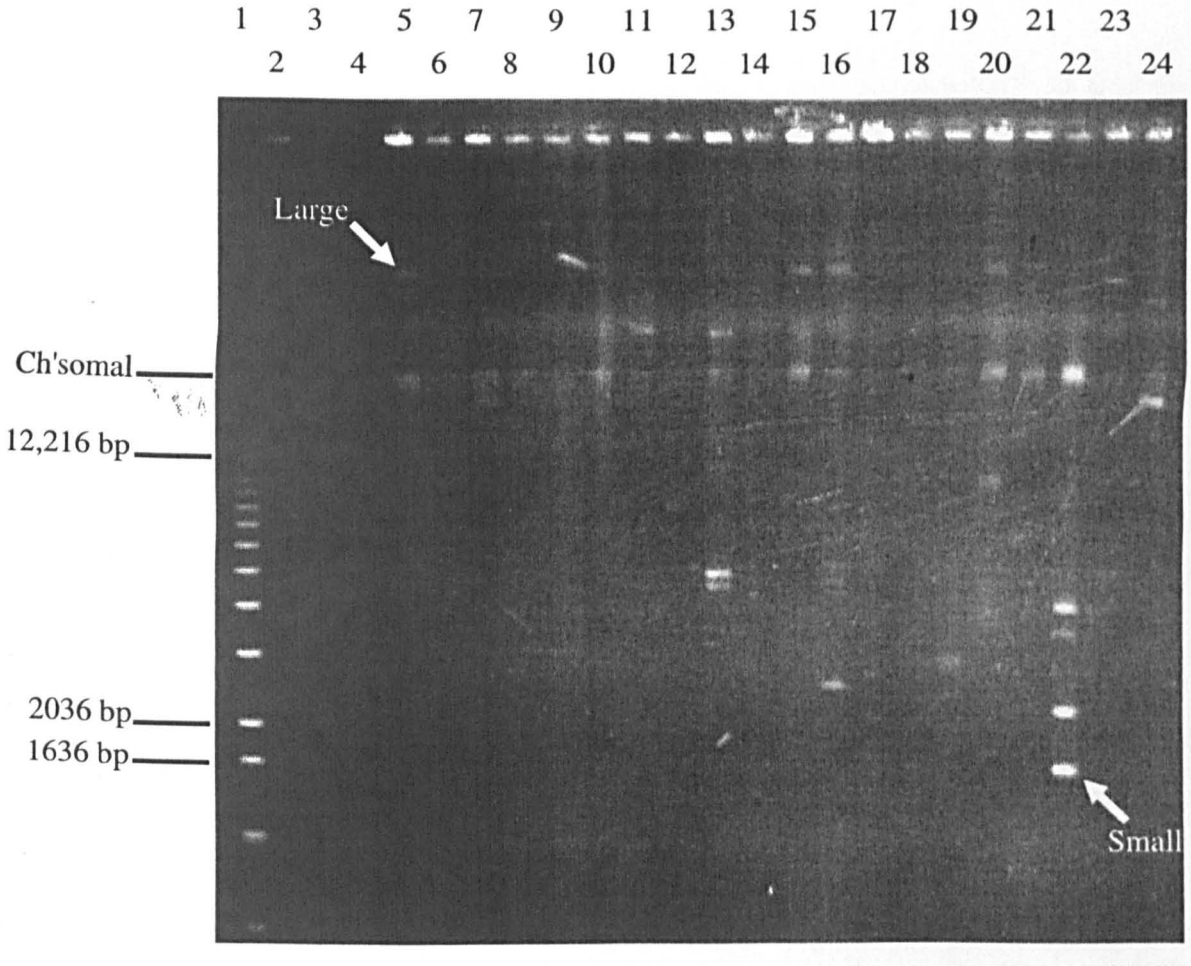
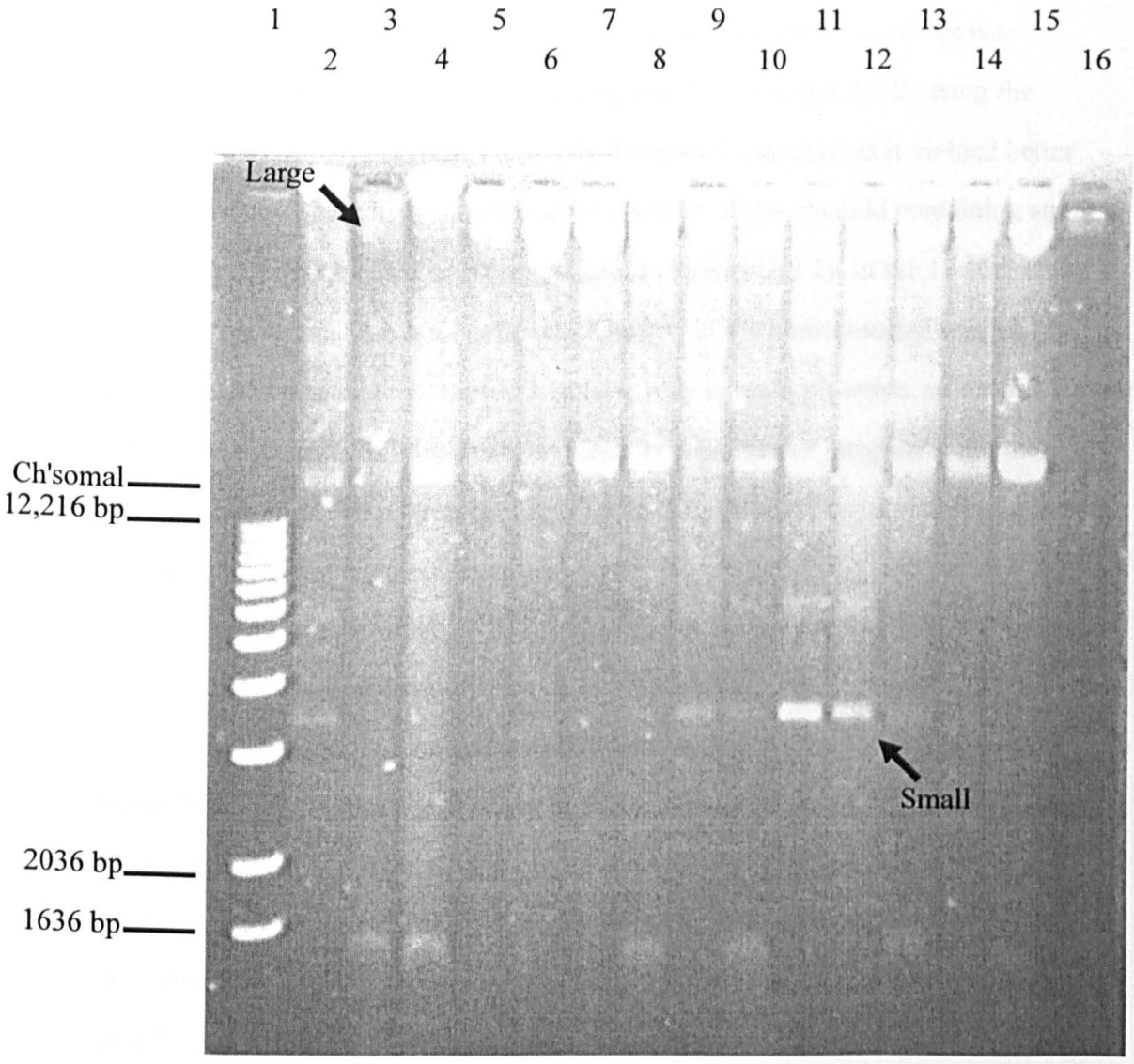

1 3 5 7 9 11 13 15 17 19 21 23
2 4 6 8 10 12 14 16 18 20 22 24
Large
Ch'somal
12,216 bp
2036 bp
1636 bp
Small

Figure 4.2.4.2. Agarose gel showing plasmid DNA from a selection of '96 isolates. DNA was extracted by the method of Olsen *et al.*, 1990. The chromosomal DNA band is shown (Ch'somal). Plasmids which electrophoresed above this chromosomal DNA band were classified as large plasmids and those which electrophoresed below the chromosomal DNA band were classified as small plasmids. Examples of both large and small plasmids are indicated (Lanes 3 and 12). This is discussed further in Section 4.2.4. Size markers are also shown. A number of different plasmid extractions were used to ascertain the presence of plasmids in these strains due to the unclear nature of some of the data. This agarose gel represents one such experiment.

Lane 1	kb ladder
2	'96 SO7
3	'96 SO8
4	'96 SO11
5	'96 SO15
6	'96 SO17
7	'96 SO18
8	'96 SO21
9	'96 SO24
10	'96 SO29
11	'96 SE15
12	'96 SE19
13	'96 SE22
14	'96 SB14
15	'96 SB21
16	*tnpA* PCR product


1 3 5 7 9 11 13 15
2 4 6 8 10 12 14 16
Large
Ch'somal
12,216 bp
Small
2036 bp
1636 bp

conform to existing incompatibility groupings (Dahlberg *et al.* 1997, Smit *et al.*, 1998, Wilmotte *et al.*, 1996).

4.2.5. Location of *tnpA* genes.

The location of the *tnpA* genes contained within the 75 isolates was determined by Southern hybridisation (Figures 4.2.5.1 and 4.2.5.2) using the 1406/2638 PCR product from Tn*21*. Alkali transfer was used as it yielded better results than the alternative salt transfer method. In all 64 plasmid containing strains, the *tnpA* gene was located on a large plasmid (Table 4.2.3.1). In the 11 apparently plasmid-free strains, the *tnpA* gene was detected in the chromosomal material. However, it is possible that these 11 isolates may contain plasmids, either of a similar size to the bacterial chromosomal fragments or which have integrated into the bacterial chromosome. Chromosomally located *tnp*A genes were only found in those isolates which contained no plasmids.

4.2.6. Structural arrangement of *tnpA* and *tnpR* genes.

The structural arrangement of the *tnpA* and *tnpR* genes was determined by the use of PCR. Of the four possible arrangements of these genes indicated in Figure 4.2.6.1, the most commonly encountered is that of the Tn*21*-subgroup. Initially, reactions were carried out using 950 and 501R1/C primers, (Pearson *et al.*, 1996) corresponding to the Tn*21* arrangement of genes. Results are shown in Tables 4.2.3.1 and 4.2.3.2. The PCR products produced from the '93 T2 isolates are also shown in Figure 4.2.6.2. Of the 75 isolates which produced a *tnpA* PCR product, 69 of these yielded an arrangement of genes similar to that found in the Tn*21*-subgroup of transposons, this fragment being of the same size as expected from Tn*21* (1.5 kb) (Figure 4.2.6.2.). The nature of these PCR products was verified by their hybridisation to *tnpA*/*tnpR* PCR products from Tn*21* and Tn*501* .

The 6 strains which produced no PCR products using these primers were then examined using the three remaining combinations of primers to determine whether

Figure 4.2.5.1. Southern blot of plasmid DNA (from Figure 4.2.4.2) with a Tn*21* *tnpA* probe. This shows the *tnpA* probe hybridising to the large plasmids, but not to the small plasmids or the chromosomal DNA band. The classification of "large" and "small" plasmids is discussed further in Section 4.2.4. A number of different plasmid extractions and Southern blots were used to ascertain the location of *tnpA* genes in these strains due to the unclear nature of some of the data. This Southern blot represents one such experiment.

Lane 1	*tnpA* PCR product
2	'96 SB21
3	'96 SB14
4	'96 SE22
5	'96 SE19
6	'96 SE15
7	'96 SO29
8	'96 SO24
9	'96 SO21
10	'96 SO18
11	'96 SO17
12	'96 SO15
13	'96 SO11
14	'96 SO8
15	'96 SO7
16	kb ladder

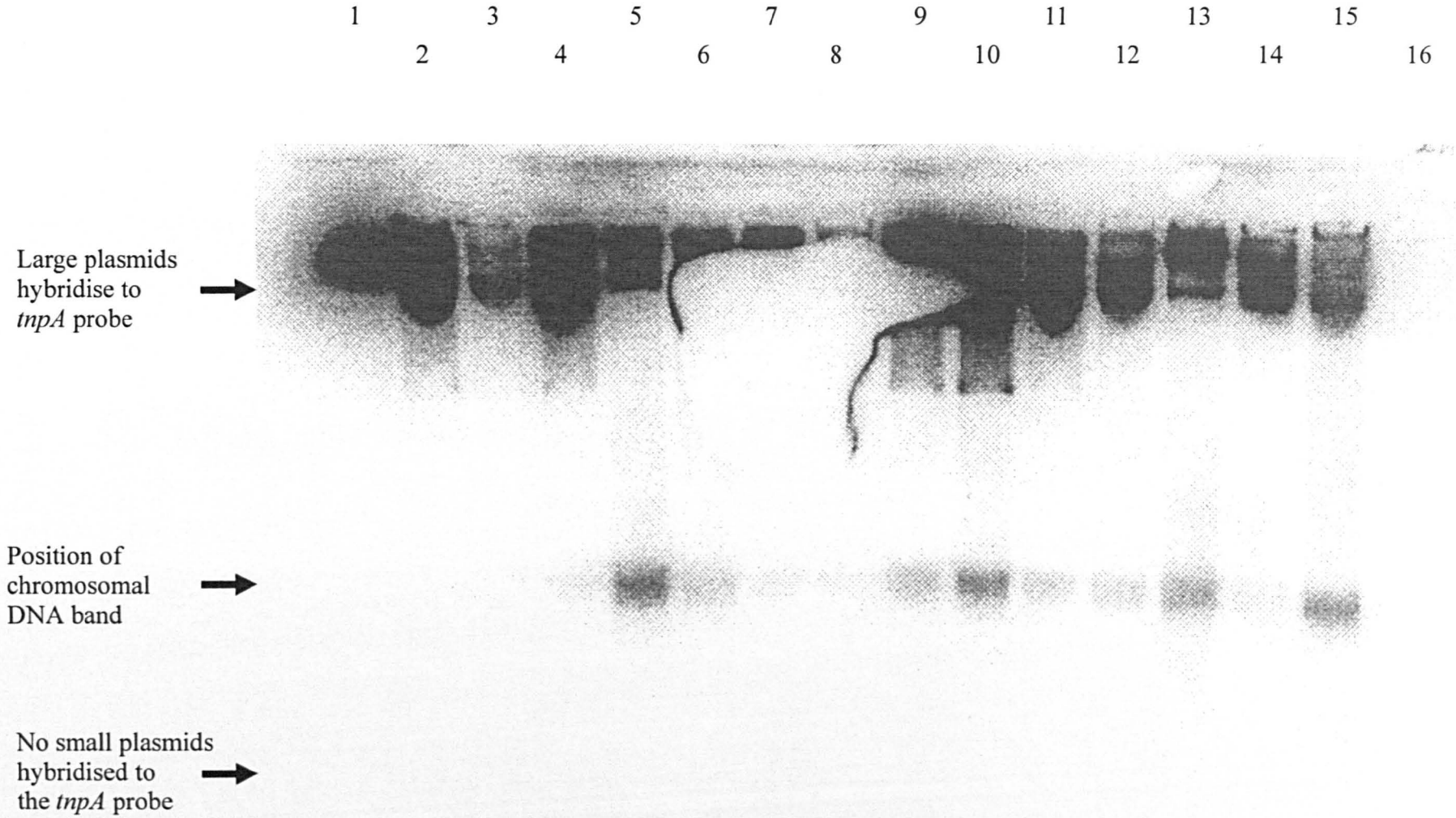
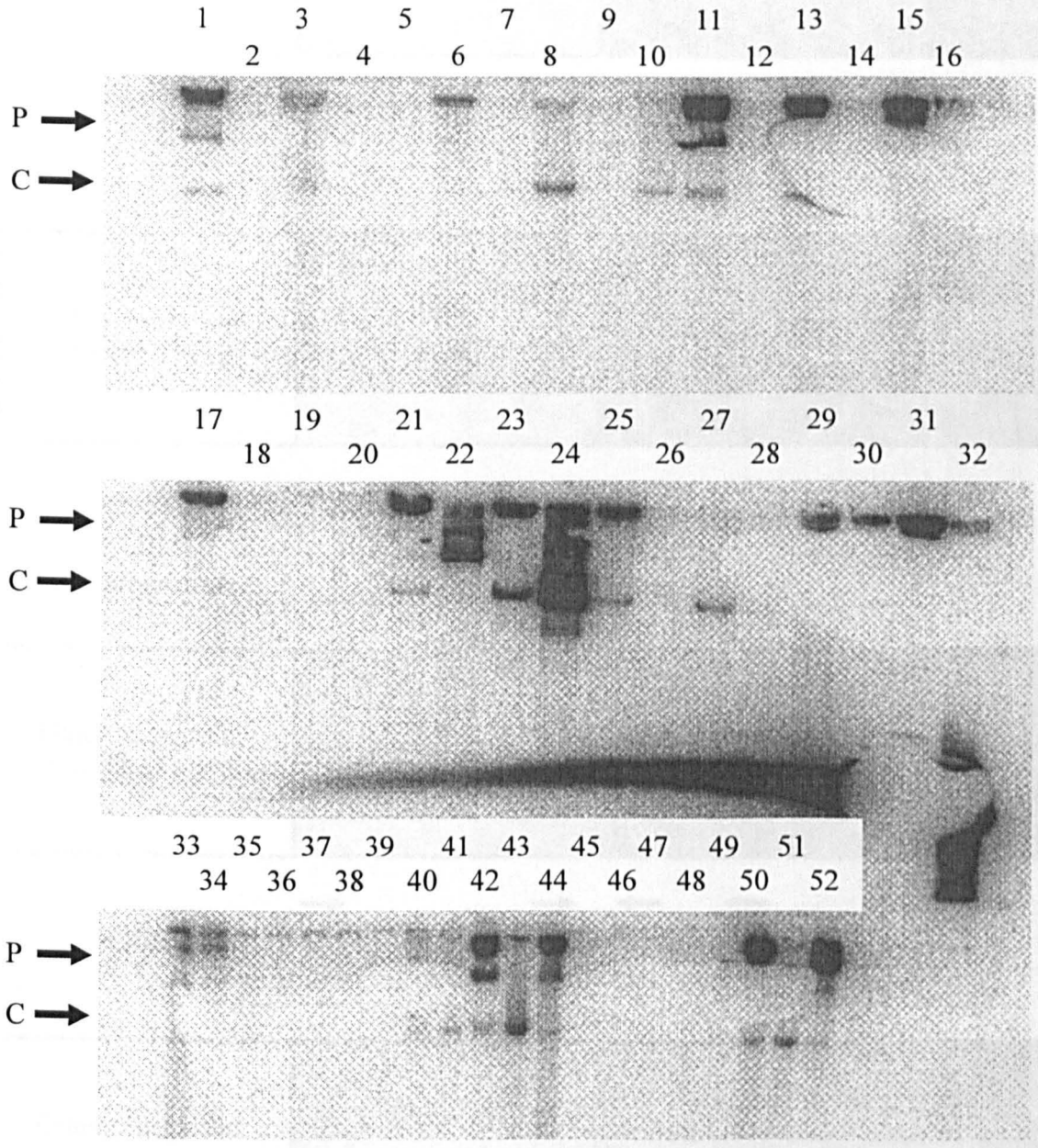
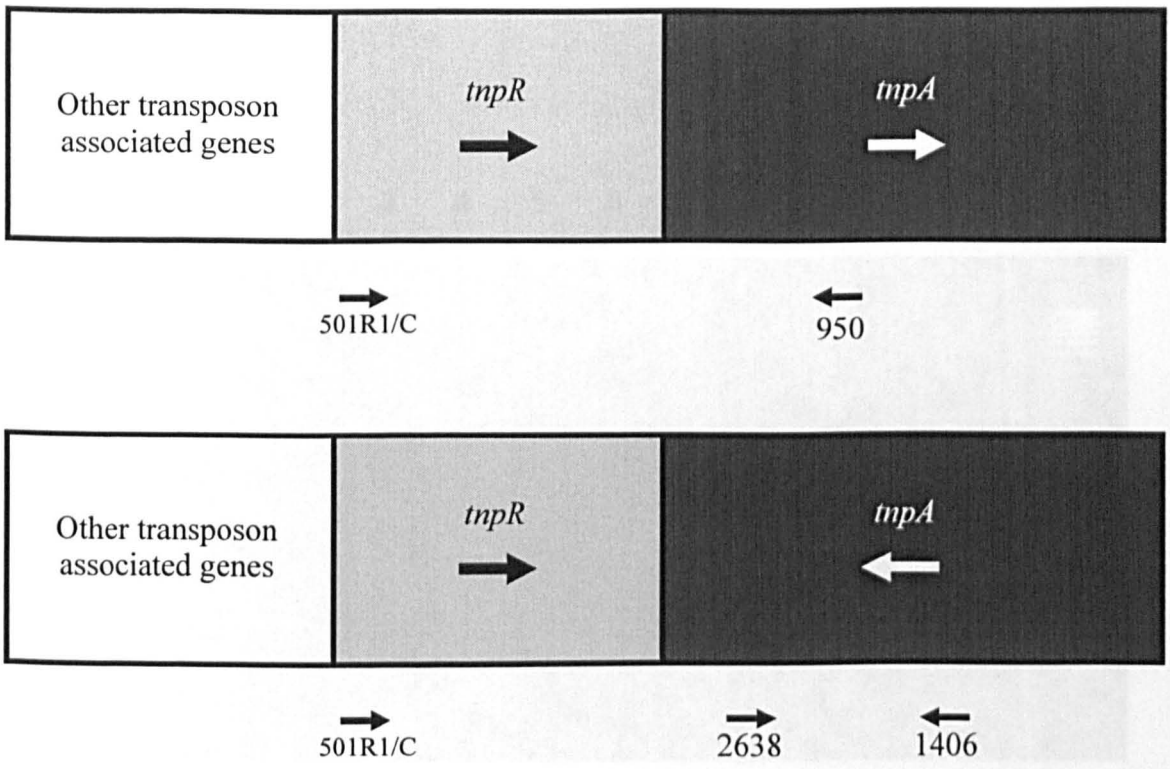

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Large plasmids hybridise to *tnpA* probe
Position of chromosomal DNA band
No small plasmids hybridised to the *tnpA* probe

Figure 4.2.5.2. Southern blot of plasmid DNA with Tn*21* *tnpA* probe. Positions of plasmid (P) DNA and chromosomal (C) DNA are shown. The classification of "large" and "small" plasmids is discussed further in Section 4.2.4. A number of different plasmid extractions and Southern blots were used to ascertain the location of *tnpA* genes in these strains due to the unclear nature of some of the data. This Southern blot represents one such experiment.

Lane 1	'96SE4	Lane 26	'96 SE 22
2	'96 SO29	27	'96 SE19
3	'96 SO27	28	'96 SE15
4	'96 SO24	29	'96 SE14
5	'96 SO15	30	'96 SE9
6	'96 SO11	31	'96 SE8
7	'96 SO8	32	Tn*501*
8	'96 SO7	33	'96 SO22
9	'93 T2 17	34	'96 SO27
10	'93 T2 12	35	'96 SE7
11	'93 T2 7	36	'96 SB1
12	'93 SB13	37	'96 SB3
13	'93 SE31	38	'96 SB4
14	'93 SE18	39	'96 SB5
15	'93 SO1	40	'96 SO17
16	kb ladder	41	'96 SO18
17	Tn*21*	42	'96 SO21
18	'96 SB37	43	'93 SE35
19	'96 SB21	44	'93 SB3
20	'96 SB14	45	'96 SO21
21	'96 SE31	46	'96 SO19
22	'96 SE30	47	'96 SO18
23	'96 SE29	48	'96 SO17
24	'96 SE27	49	'96 SO15
25	'96 SE26	50	'96 SO13


1 3 5 7 9 11 13 15
2 4 6 8 10 12 14 16
P
C
17 19 21 23 25 27 29 31
18 20 22 24 26 28 30 32
P
C
33 35 37 39 41 43 45 47 49 51
34 36 38 40 42 44 46 48 50 52
P
C

Tn*21*-like arrangement:



Tn*3*-like arrangement:

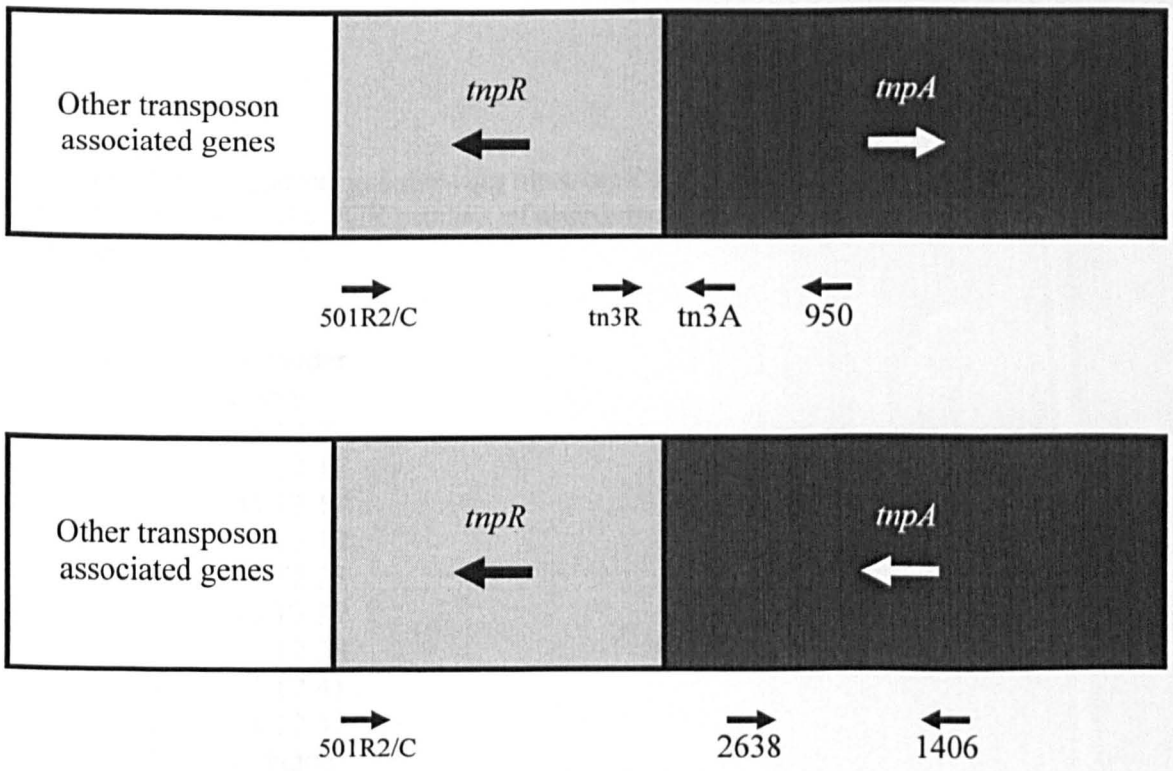


Figure 4.2.6.1. Schematic representation of the 4 possible orientations of *tnpA* and *tnpR* genes and the primers used to detect these arrangements. Not to scale. *res* sites not shown

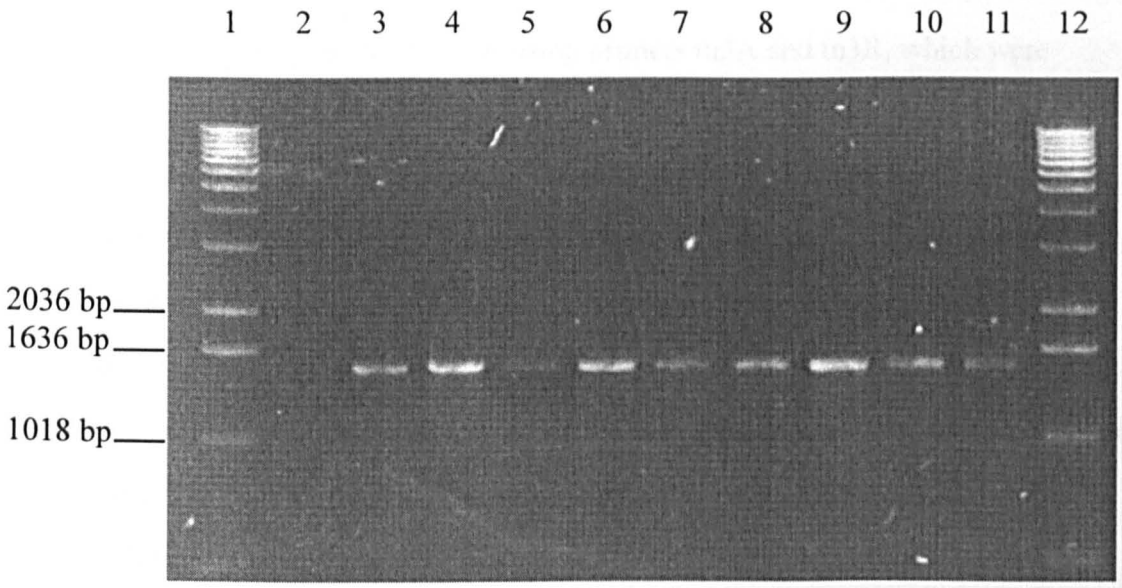


Figure 4.2.6.2. Agarose gel showing *tnpA/tnpR* PCR products from '93 T2 strains. All strains produced a PCR product of approximately 1.5kb. Size markers are also shown.

Lane 1	kb ladder
2	empty
3	'93 T2 7
4	'93 T2 12
5	'93 T2 17
6	'93 T2 19
7	'93 T2 23
8	'93 T2 37
9	'93 T2 38
10	'93 T2 41
11	'93 T2 47
12	kb ladder

any other arrangements of genes were present. This was carried out using the following primer combinations: 501R1/C and 1406, 501R2/C and 950, and 501R2/C and 1406, thus covering all 4 possible arrangements of genes as indicated in Figure 4.2.6.1. No PCR products were produced using these 3 combinations of primers. The 6 strains were also subjected to PCR using primers tn3A and tn3R, which were designed to Tn*3* *tnpA* and *tnpR* genes respectively. The results shown in Tables 4.2.3.1 and 4.2.3.2 indicate that no strains have Tn*3*-like gene arrangements. Subsequently the presence of *tnpR* genes in those isolates not producing any viable PCR products was determined by PCR using primers 501R1/C and 501R2/C. Four strains produced no PCR products using these primers, suggesting that these strains either have no *tnpR* gene, or that the specificity of the primers is such that they did not allow the detection of diverse gene sequences. Two strains, '96 SE19 and '96 SE30 had both *tnpR* and *tnpA* genes, in undetermined configurations. Long template PCR using the Expand system (2.14) using all five sets of primers yielded no PCR products for these two strains. All strains producing a *tnpA* PCR product have had their *tnpA/tnpR* arrangement determined except for '96 SE19 and 30; these two isolates may contain transposon structures which are distinct in evolutionary terms from the Tn*3* group of Class II transposons despite having gene sequences which are similar to the other transposons studied.

4.2.6.1. RFLP of *tnpA/R* products from '93 SE35, SB4 and SO2.

RFLP reactions were carried out to ascertain whether the PCR products produced were of an identical nature. Reactions were carried out using *StyI* and *PvuII*. Results shown in Figure 4.2.6.1.1 indicate that the 3 PCR products yielded different restriction patterns suggesting that they are not identical.

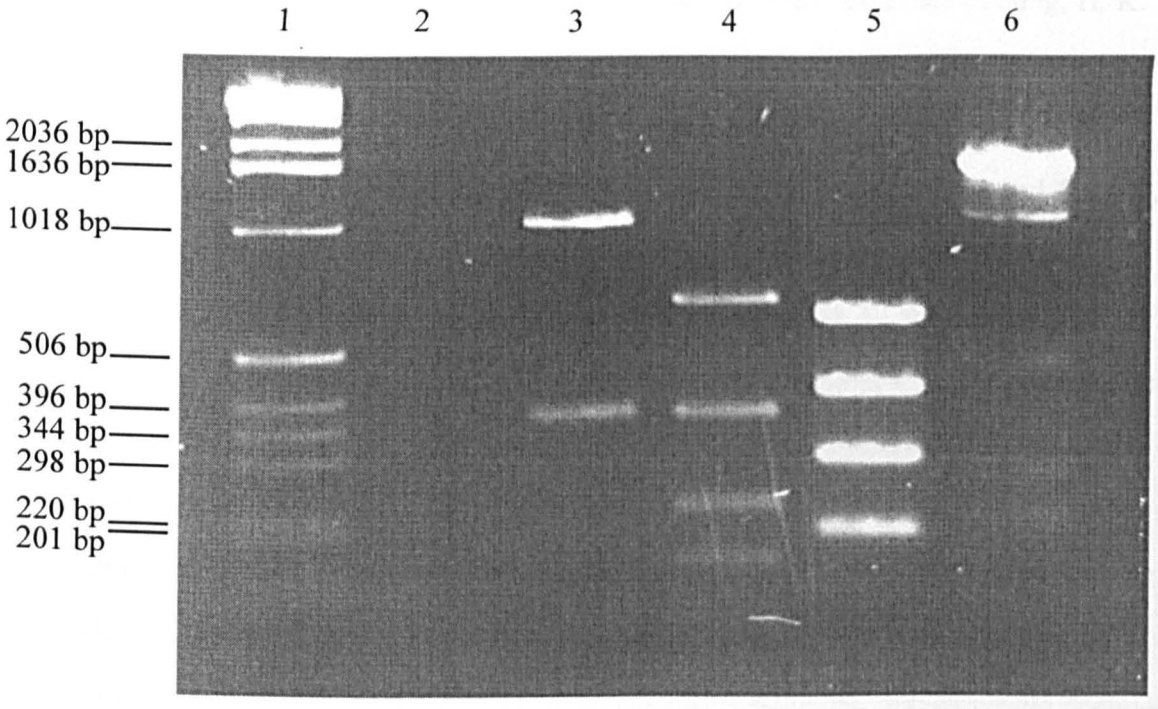


Figure 4.2.6.1.1. Agarose gel showing restriction digest of *tnpA/tnpR* PCR products (produced using primers 501R1/C and 950) using *Sty*I. This indicates that the 3 PCR products yielded different restriction patterns suggesting that the DNA sequences of the products are not identical. Restriction analysis using *Pvu*II indicated no restriction sites for this enzyme in any of the PCR products.

Lane 1	kb ladder
2	empty
3	Tn*501*
4	'93 SE35
5	'93 SB4
6	'93 SO2

4.2.7. Presence of integrase genes and size of inserted gene cassettes.

The presence of integrase genes serves as an indicator of the presence of integron elements within the transposon. PCR using primers *int21A and int21B* designed to allow amplification from a wide range of integrase genes (Young, H. K. and Rosser, S. J. Personal communication) was carried out on the collection of isolates. Five of the 124 strains produced PCR products of the correct size which hybridised to the corresponding PCR product from Tn*21* (Tables 4.2.3.1 and 4.2.7.1). The size of the gene cassettes inserted into the *rhs* of the integron elements was determined by PCR using primers 4127 and 4128 (Young, H. K. and Rosser, S. J. Personal communication). Insert size varied between 1.1 kb and 4.5 kb. Although the nature of these inserts is currently undetermined, such cassettes commonly encode antibiotic resistance genes. With the exception of '93 T2 37, all the isolates found to contain integron structures were members of the '96 SE group, totalling 13% of those isolates. This is in contrast to the '93 SE group in which no PCR products were seen using these primers.

4.2.8. Relative arrangement of *merD* and *tnpR* genes.

The arrangement of *mer* and *tnp* genes in the 75 (*tnpA* amplifying) isolates was studied by PCR using primers mercD and 501R2/C, corresponding to the Tn*21/501* orientation of the *merD* and *tnpR* genes. All strains failing to produce a PCR product were tested using primers mercD and 501R1/C in order to ascertain whether any structural diversity (ie inversion of *tnpR*) was present in the isolates. No such arrangements were observed. Results are shown in Tables 4.2.3.1 and 4.2.8.1.

Reactions using the Tn*21*/Tn*501*-like primers yielded 2 sizes of product; a 2 kb product, which is the expected size from Tn*501*, and a 1.3 kb band. The PCR product from Tn*501* hybridised to all the 2kb PCR products indicating that the gene arrangement in these strains was similar to that of Tn*501* (Figure 4.2.8.1).

The DNA sequence of the 1.3kb PCR product from '96 SE13 was determined by cloning of the PCR product into a pGEM vector and subsequent DNA sequencing

Table 4.2.7.1. Integron containing isolates.

STRAIN	Integrase gene*	Integron insert size†
'93 T237	+	1.2kb
'96 SE6	+	1.1kb
'96 SE9	+	1.4kb
'96 SE15	+	4.5kb
'96 SE19	+	1.3kb
Other 119 isolates	-	N/A

* +: Integrase gene present.
-: No PCR product.
† N/A: Not applicable.

Table 4.2.8.1. Table showing results of PCR reactions using primers for *merD* and *tnpR* genes. Also shown are results from PCR reactions to detect the presence and size of genes inserted between *merA* and *merP*.

ISOLATION GROUP	*merD/tnpR* PCR			*merA/merP* PCR			
	2 kb	1.3 kb	No product	YES	NO	No product	Unknown
'93 SO	9	0	1	10	0	0	0
'93 SE	0	4	3	4	0	5	1
'93 SB	6	1	0	0	5	5	0
'93 T2	0	4	5	5	0	2	2
'96 SO	0	0	20	0	0	30	0
'96 SE	1*	9*	13	16	2	13	0
'96 SB	0	0	2	9	0	15	0
TOTAL	15*	17*	44	44	7	70	3

* '96 SE9 produces both 2 kb and 1.3 kb PCR products.

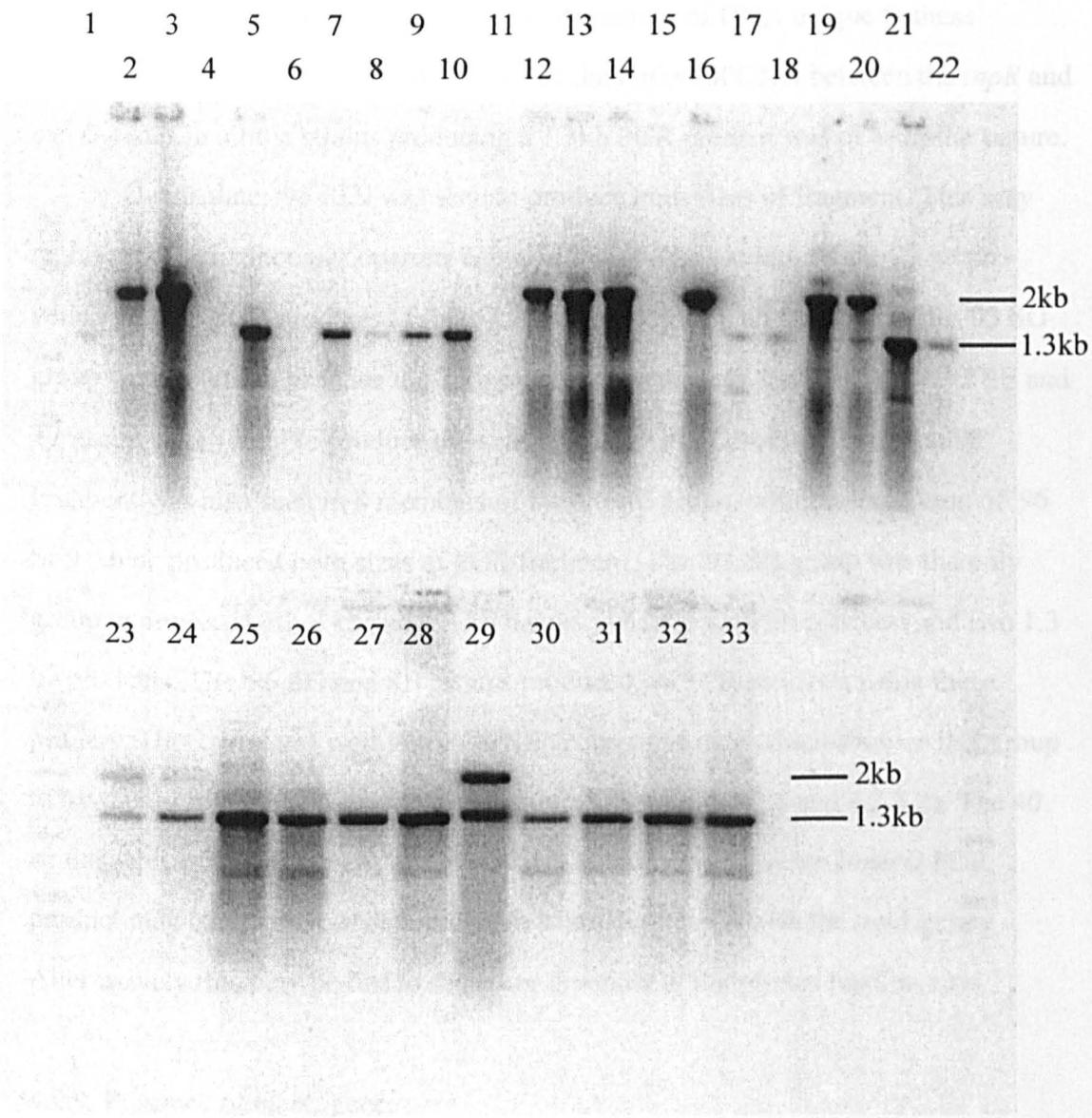
merD/tnpR PCR was carried out on the 75 strains which positively amplified using *tnpA* primers.

merC PCR was carried out on all 124 strains.

Numbers represent the number of strains with the described characteristic.

Figure 4.2.8.1. Southern blot of mercD/501R2/C PCR products. The mercD/501R2/C PCR product from Tn*501* was used as a probe. The PCR reaction using these primers yielded two sizes of PCR product; a 2 kb product, which is the expected size from Tn*501*, and a smaller 1.3 kb product. The Southern blot shows both sizes of PCR product hybridising to the 2 kb Tn*501* probe. This suggests that the two products contain similar gene sequences. One strain, '96 SE9 (Lane 29), produced both sizes of PCR product. In order to determine the nature of the smaller, 1.3 kb PCR product, DNA sequencing was carried out.

Lane 1	Tn*21*
2	Tn*501*
3	Tn*501*
4	empty
5	'93 SO1
6	'93 SE3
7	'93 SE9
8	'93 SE18
9	'93 SE20
10	'93 SE31
11	'93 SE35
12	'93 SB2
13	'93 SB3
14	'93 SB4
15	'93 SB5
16	'93 SB8
17	'93 SB12
18	'93 SB13
19	'93 SB22
20	'93 SB29
21	'93 T2 7
22	'93 T2 17
Lane 23	'93 T2 19
24	'93 T2 38
25	'96 SE1
26	'96 SE3
27	'96 SE4
28	'96 SE8
29	'96 SE9
30	'96 SE12
31	'96 SE13
32	'96 SE18
33	'96 SE24


1 3 5 7 9 11 13 15 17 19 21
2 4 6 8 10 12 14 16 18 20 22
2kb
1.3kb
23 24 25 26 27 28 29 30 31 32 33
2kb
1.3kb

(Figure 4.2.8.2). This indicated the presence of a region of DNA corresponding to that found between the *tnpR* and *merD* genes of Tn*3926*, Tn*5036*, Tn*5059* and pMER610, which contains 3 open reading frames of unknown function (Osbourn, *et* al., 1995, Yurieva *et al.*, 1997). DNA sequence information was used to obtain a probe from the sequenced PCR product, corresponding to the region of DNA unique to these transposons. This was used to ascertain that the region of DNA between the *tnpR* and *merD* genes in all the strains producing a 1.3kb PCR product was of a similar nature.

One isolate, '96 SE9 was seen to produce both sizes of fragment. This may represent two distinct *mer* operons contained within this strain. Of the 35 strains which yielded PCR products (Tables 4.2.3.1 and 4.2.8.1), all 9 strains in the '93 SO group were found to produce the 2 kb product, whereas all 8 strains in the '93 SE and T2 group were found to produce the smaller 1.3 kb PCR product. This smaller fragment was also seen in 8 members of the '96 SE group, with the exception of '96 SE9 which produced both sizes of PCR fragment. The '93 SB group was the only group to produce both 2 kb and 1.3 kb bands, yielding six 2 kb products and two 1.3 kb products. The '96 SO and SB groups produced no PCR products using these primers. This correlates well to the '96 SB group *tnpA* data which showed this group to have only two transposon containing strains (Tables 4.2.3.1 and 4.2.3.2). The 40 strains which contain a *tnpA* gene but which do not produce a *tnpA/merD* PCR product may contain a *mer* operon which is not associated with the *tnpA* gene. Alternatively this may be due to sequence diversity at the primer binding sites.

4.2.9. Presence of *merC* gene.

Using PCR primers mercP and mercA the presence and size of any genes contained within the *merP/merA* interval was determined for all 124 isolates. PCR reactions using these primers produced two sizes of product using Tn*21* or Tn*501* as PCR templates. Tn*21* yielded a 1 kb PCR product due to the presence of the *merC* gene and Tn*501* yields a 600 bp product.

```
1     TCGTCGAGCG TCGGCGCGAG GCCCTGGCCA GCCTCGAAAT GCAACTGGCC GCCATGCCAA

                                          <Urf1 start
61    CCGAACCGGC ACAGCACGCG GAGAGTCCGC CATGAACAGC CCAGAGCACT TGCCGTCTGA
                                end of merD>

121   GACGCACAAA CCGATCACCG GCTACTTGTG GGGCGCGCTG GCCGTGCTCA CCTGTCCCTG

181   CCATTTGCCG ATTCTCGCCA TTGTGCTAGC CGGCACGACG GCCGGCGCGT TCATCGGGGA

241   GCACTGGGGT ATTGCAGCCC TCACGCTGAC CGGCTTGTTT GTCCTGTCTG TGACGCGGCT

                         end of Urf1>
301   GCTGCGGGCC TTCAAGGGAA GATCATGACC GCTTCCCAGC CAGCCGAGAG TGGGCAGCTT
                                <Urf2X start

                                             <Urf2Y start
361   TGAGCTTCGC TACCAATCTG GAGGAGTACC ACCATGAACG CAAACGCCCC GAACACTGCC
       >end of Urf2X

421   AGTTGCACCA CCTGCTGCGT ATGCTGCAAA GAAATTCCGC TCGATGCCGC CTTCACCCCG

481   GAAGGCGCGG AATACGTCGA ACATTTCTGC GGGCTGGATT GCTATGAACG CTTCCAGGCA

541   CGCGCCAAGG CCGCGACAGA ATCTGACATT GCGCCTGTCC CTGGCGGTTC GCAGCCGTCA

          >end of Urf2Y
601   GATTGAGGCA TACCCTAACC TGATGTCAGA TGCTTGGCGC AAAGACGTCA GAATAGAGTT

661   GTAGTTTGTA TTTATTGATA CAAGCCGCCA AGGGTAATGG ATTTCATCCT GACACTTTTA

                      <tnpR start
721   CCTTTGGAGG CACCTTGCAA GGTCAACGCA TTGGCTATAT CCGCGTCAGC AGCTTCGACC

781   AGAACCCTGA TCGGCAACTG GAGCAAATCG AAGTCGGTAA GGTATTCACC GATAAGGCTT

841   CCGGCAAGGA CACGCAACGT CCCGAACTTG AAAGGCTGCT GGCCTTTGTG CGCGAGGGCG

901   ACACCGTGGT GGTGCACAGC ATGGACAGGT TGGCACGCAA TCTCGATGAC CTGCGCCGCA

961   TCGTTCAGGG GCTGACACAA CGGGGCGTGC GGATGGAGTT TGTCAAAGAA GGGCTGGCGT

1021  TCACCGGCGA TGACTCACCG ATGGCCAATT TGATGCTGTC GGTCATGGGG GCTTTTGCGG

1081  AGTTCGAGCG CGCACTGATC CGCGAACGCC AGCGCGAGGG AATCGTGCTG GCCAAGCAGC

1141  GCGGTGCCTA CCGGGGACGA AAGAAATCGC TGAACAGCGA ACAAATTGCC AAGTTGAAAC

1201  AGCGAGTCGC GGCAGGCGAT CAAAAAACCT TGGTGGCCCG TGACTTCGGC ATCAGCCGCG

1261  AAACCC
```

Figure 4.2.8.2. DNA sequence of mercD/501R2/C PCR product from '96SE13. Positions of open reading frames are indicated. This sequence has been assigned Accession No AF13421.

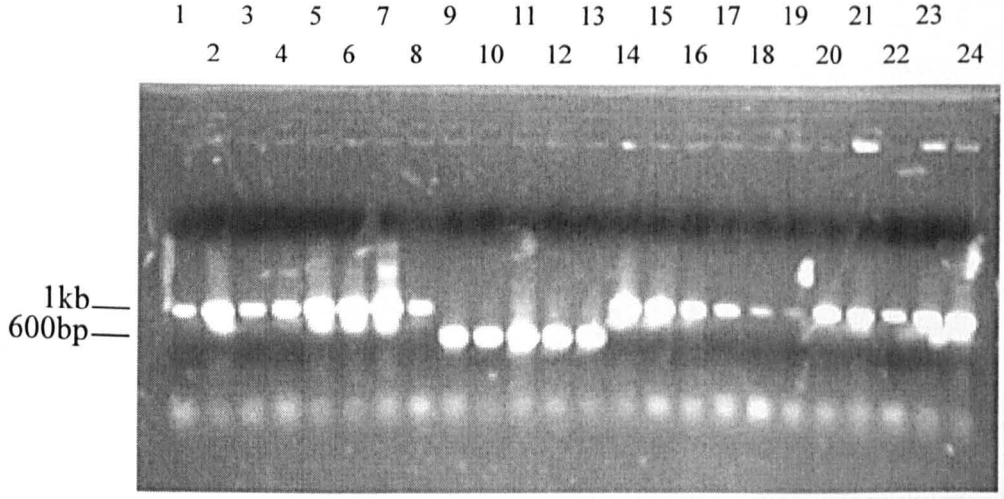
PCR products were obtained from 54 of the isolates, (Tables 4.2.3.1 and 4.2.8.1). Southern hybridisations were carried out on these PCR products using mercP/mercA PCR products from both Tn*21* and Tn*501* (examples of these are shown in Figures 4.2.9.1 and 4.2.9.2). To distinguish between those strains which produce a PCR product containing *merC* and those which do not, PCR products of both sizes were hybridised to a *merC* gene probe. Forty-four of the forty-seven 1 kb PCR products hybridised to the *merC* probe, i. e. they contained a *merC* gene. The three strains that did not contain *merC* genes; '93 SE31, T2 37 and T2 38 apparently contain a gene between the *merP* and *merA* genes that is of similar size to *merC*, but which remains of an undetermined nature. Seven isolates were seen to produce a 600 bp PCR product, which did not hybridise to the *merC* probe and were therefore assumed to be Tn*501*-like.

All isolates in the '93 SO, SE and T2 group that produced a PCR product (except the three strains with unknown inserts) contained a *merC* gene, whereas all the '93 SB isolates were Tn*501*-like, containing no genes between *merP* and *merA*. The '96 SO isolates produced no PCR products, which correlates well to the *merD/tnpR* data for this group. The observed mercury resistance of these isolates may not be due to archetypal Gram negative *mer* genes, i.e. mercury resistance may be confered by a non-*mer* operon system or could be due to an operon with large sequence diversity to that found on Tn*21* and Tn*501*.

All nine of the '96 SB strains that produced a PCR product contained the *merC* gene, as did all of the '96 SE strains except '96 SE 26 and 28. This is interesting as the '96 SB group did not produce a *merD/tnpR* PCR product. This may be due to the *mer* genes being located at a position removed from the transposase gene or contained on non-transposon structures. There is also a difference in the SB strains in that the '93 SB isolates which gave a PCR product all contained a Tn*501*-like *mer* operon while the '96 SB isolates which gave a PCR product, all contained Tn*21*-like *mer* operons containing *merC*. The observed frequency of *merC* genes is higher than previously

Figure 4.2.9.1. Agarose gel showing all mercA/mercP PCR products from '93 and '96 isolates. PCR product sizes are indicated. This shows that some strains produced a 600bp PCR product containing no *merC* gene (equivalent to the PCR product from Tn*501*), whereas other strains produced a 1kb PCR product (equivalent to the PCR product from Tn*21*). These isolates may or may not contain a *merC* gene inserted between their *merP* and *merA* genes. This was confirmed by Southern blot hybridisation using a *merC* gene probe derived from the Tn*21* PCR product (Figure 4.2.9.2).

Lane 1	Tn*21*	Lane 25	Tn*501*
2	'93 SO1	26	'96 SE7
3	'93 SO12	27	'96 SE8
4	'93 SE6	28	'96 SE9
5	'93 SE9	29	'96 SE12
6	'93 SE11	30	'96 SE13
7	'93 SE20	31	'96 SE17
8	'93 SE31	32	'96 SE18
9	'93 SB2	33	'96 SE20
10	'93 SB3	34	'96 SE24
11	'93 SB4	35	'96 SE26
12	'93 SB8	36	'96 SE27
13	'93 SB22	37	'96 SE28
14	'93 T2 7	38	'96 SE29
15	'93 T2 17	39	'96 SE31
16	'93 T2 37	40	'96 SB9
17	93 T2 38	41	'96 SB11
18	'93 T2 19	42	'96 SB12
19	"93 T2 23	43	'96 SB13
20	'93 T2 41	44	'96 SB16
21	'96 SE1	45	'96 SB18
22	'96 SE2	46	'96 SB19
23	'96 SE3	47	'96 SB24
24	'96 SE4	48	'96 SB37


1 3 5 7 9 11 13 15 17 19 21 23
2 4 6 8 10 12 14 16 18 20 22 24
1kb
600bp

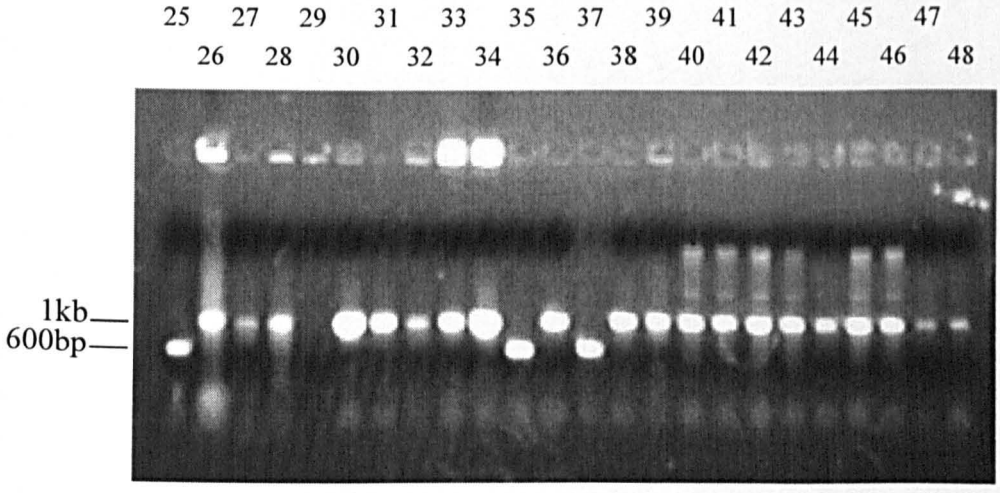
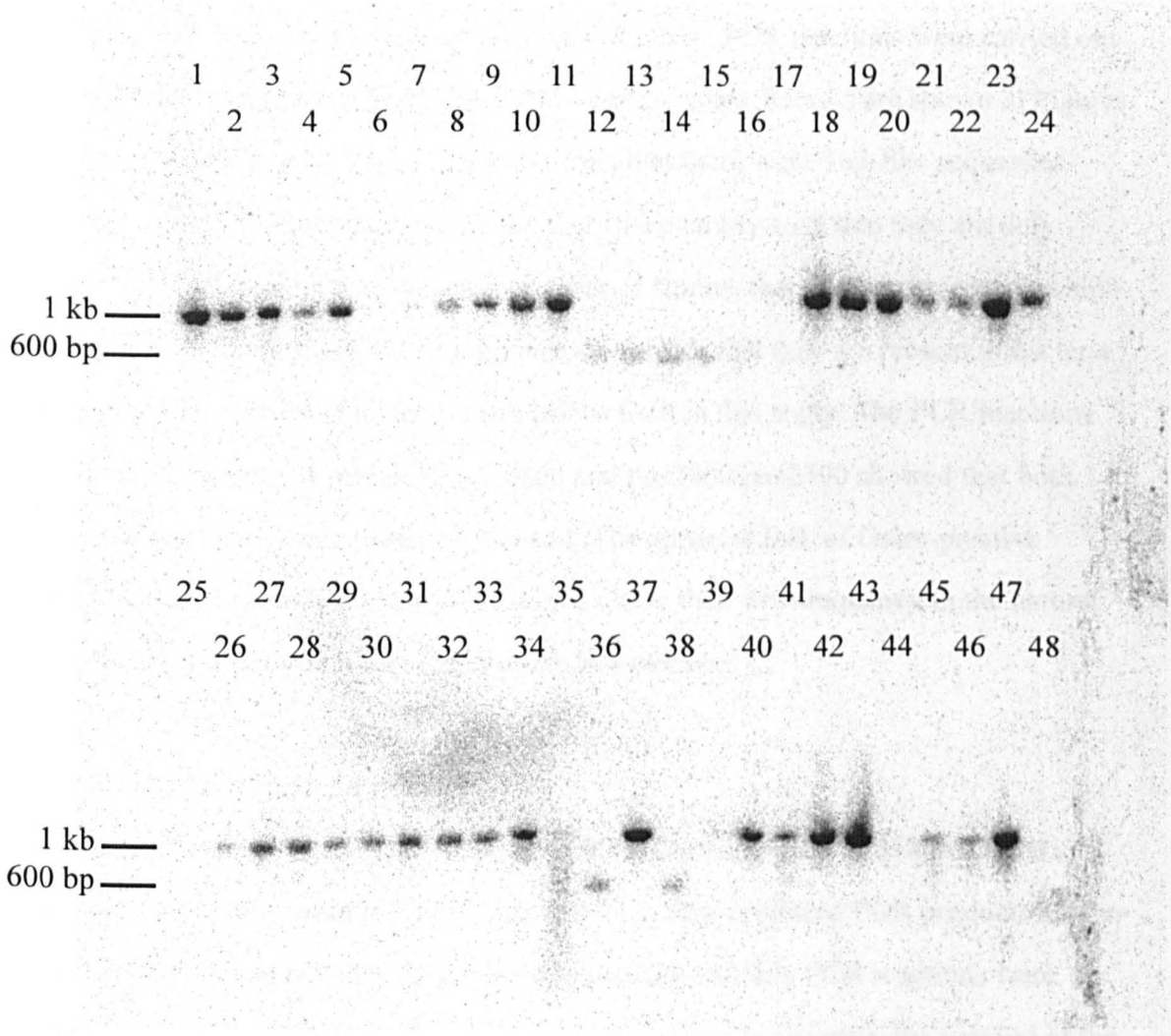

25 27 29 31 33 35 37 39 41 43 45 47
26 28 30 32 34 36 38 40 42 44 46 48
1kb
600bp

Figure 4.2.9.2. Southern blot of mercA/mercP PCR products (from Figure 4.2.9.1) with *merC* gene probe. Approximate PCR product sizes are indicated. As expected, those strains which produced a 600 bp PCR product, did not hybridise to the probe. Of those strains producing a 1 kb product, only three did not not hybridise to the *merC* probe afrter repeated Southern blots. This indicates that these strains contain a gene other than *merC* inserted between their *merA* and *merP* genes.

Lane 1	Tn*21*	Lane 25	Tn*501*
2	'93 SO1	26	'96 SE7
3	'93 SO12	27	'96 SE8
4	'93 SE6	28	'96 SE9
5	'93 SE9	29	'96 SE12
6	'93 SE11	30	'96 SE13
7	'93 SE20	31	'96 SE17
8	'93 SE31	32	'96 SE18
9	'93 SB2	33	'96 SE20
10	'93 SB3	34	'96 SE24
11	'93 SB4	35	'96 SE26
12	'93 SB8	36	'96 SE27
13	'93 SB22	37	'96 SE28
14	'93 T2 7	38	'96 SE29
15	'93 T2 17	39	'96 SE31
16	'93 T2 37	40	'96 SB9
17	'93 T2 38	41	'96 SB11
18	'93 T2 19	42	'96 SB12
19	'93 T2 23	43	'96 SB13
20	'93 T2 41	44	'96 SB16
21	'96 SE1	45	'96 SB18
22	'96 SE2	46	'96 SB19
23	'96 SE3	47	'96 SB24
24	'96 SE4	48	'96 SB37


1 3 5 7 9 11 13 15 17 19 21 23
2 4 6 8 10 12 14 16 18 20 22 24
1 kb
600 bp
25 27 29 31 33 35 37 39 41 43 45 47
26 28 30 32 34 36 38 40 42 44 46 48
1 kb
600 bp

described and *merC* genes are seen here in a wider range of bacterial species and plasmids (Gilbert and Summers, 1988).

4.2.10. PCR reactions using DNA extracted directly from soil.

As the 124 isolates in this study did not contain either Gram-positive *tnpA* sequences or Tn*3*-like arrangements of *tnpA/R* genes, PCR reactions were carried out on DNA isolated directly from soil at the sampling sites. Results are shown in Figures 4.2.10.1, 4.2.10.2 and 4.2.10.3. This confirms that there were Tn*3*-like sequences present in the soil, but that these were either in a quantity such that they are only detectable by isolating a much larger number of strains, that they are present at a high frequency but are not associated with *mer* operons or that they are present in bacteria which are not culturable under the conditions used in this study. The PCR reactions carried out using *tnpA* primers 2501/2850 and pos2000/pos2400 showed that both types of *tnpA* gene were present in the soil. The apparent lack of Gram-positive sequences in this study may be due again to either their low frequency in the natural environment or their non-association with *mer* operons.

4.2.11. DGGE using *tnpA* primers.

Primer pairs: 2501 and 2850, pos2000 and pos2400 and 62351 and 62601 were designed to be used in DGGE reactions. i.e. they produced PCR products shorter than 400bp and contained no degenerate sequences. Initially PCR reactions were carried out using DNA from appropriate control strains to ascertain the validity of the primers for use in PCR reactions using DNA extracted directly from the environment. Primer pair 62601/62601 did not amplify a PCR product when used in control PCR reactions despite repeated attempts to optimise PCR conditions. This primer set was not used in further experiments. This set of primers was only designed to amplify products from five sequences, which as can be seen from Figure 4.2.2.1.1. are very diverse, so this lack of amplification is perhaps unsurprising.

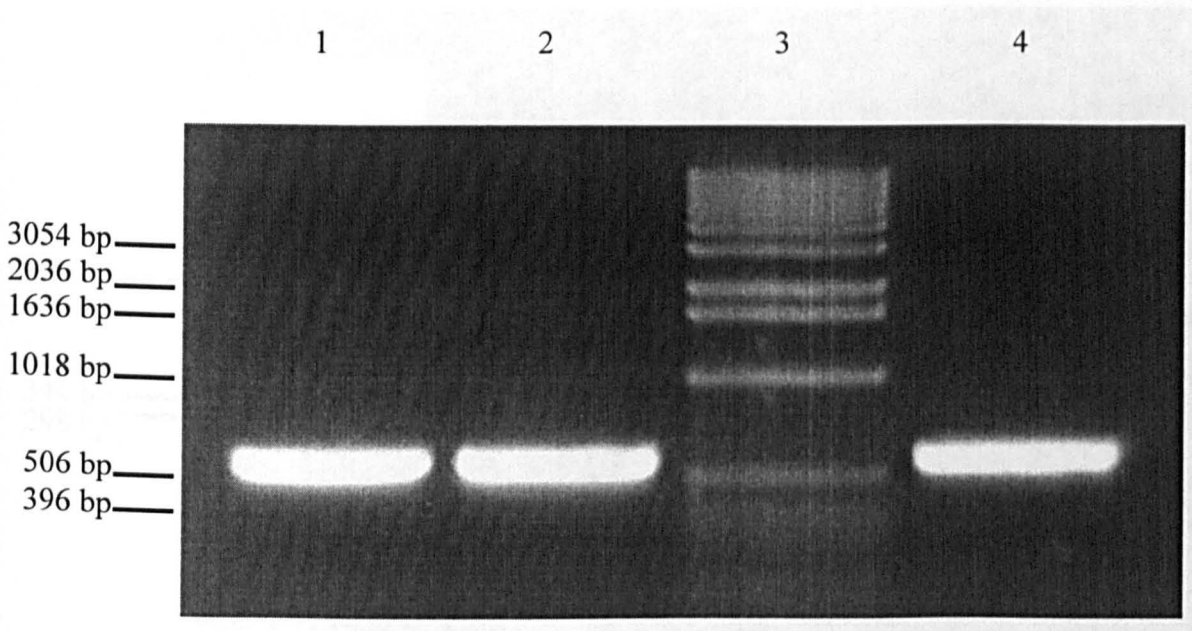


Figure 4.2.10.1. Agarose gel showing PCR from "total" DNA using tn3A and tn3R primers. This indicates the prsence of Tn*3*-like arrangements of *tnpA* and *tnpR* genes at the sampling site.

Lane 1	Tn*3* control DNA
2	Fiddlers Ferry soil DNA
3	kb ladder
4	Fiddlers Ferry sediment DNA

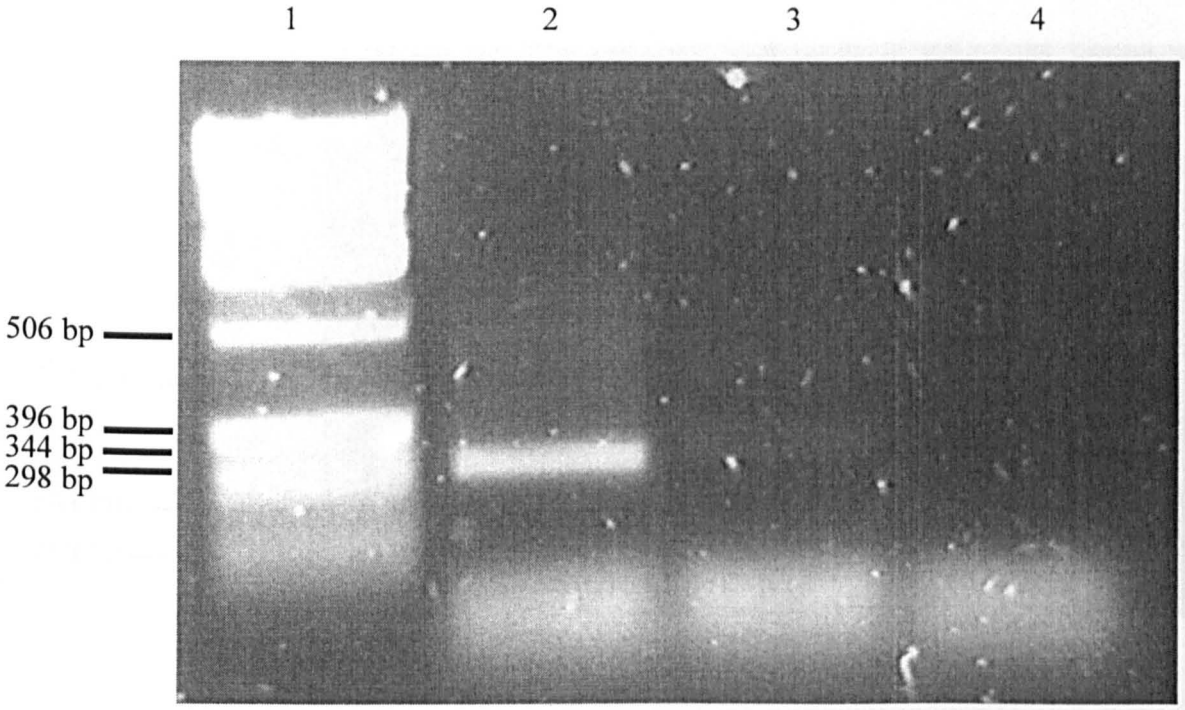


Figure 4.2.10.2. Agarose gel showing PCR from "total" DNA using 2501 and 2850 *tnpA* primers. This shows the presence of *tnpA* sequences at the sampling site.

Lane 1	kb ladder
2	Tn*21* control DNA
3	Fiddlers Ferry soil DNA
4	Negative water control

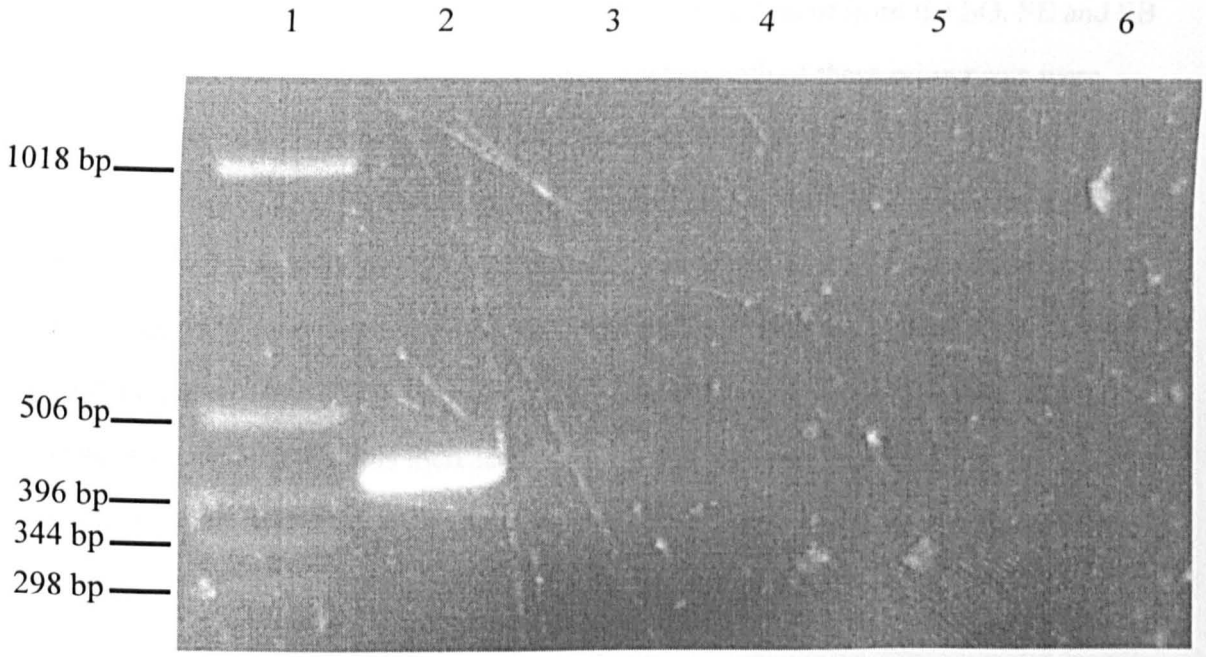


Figure 4.2.10.3. Agarose gel showing PCR from "total" DNA using pos2000/pos2400 *tnpA* primers. This indicates the presence of *tnpA* sequences previously associated with Gram-positive bacteria at the sampling site.

Lane 1 kb ladder
2 DNA extracted from soil at Fiddlers Ferry
3 DNA extracted from soil at Fiddlers Ferry (1/10 conc used in Lane 2)
4 DNA extracted from sediment at Fiddlers Ferry
5 Empty
6 Negative water control

Primer pairs 2501/2850 and pos2000/pos2400 succesfully amplified the correct size product from a number of different control strains therefore validating their use in PCR using DNA extracted directly from the environment. PCR reactions carried out on DNA isolated directly from soil and sediment from the SO, SE and SB sites yielded PCR product of the correct size when both of these primer sets were used. This required 72 cycles of PCR to be carried out on DNA which had been purified using Wizard® spin columns. If no purification step was carried out then no amplification was seen, despite repeated attempts to optimise PCR conditions. However, PCR reactions using primers containing the 45 bp GC clamp necessary for DGGE also required 72 cycles of PCR and produced weak products which were unsuitable for DGGE. This increased number of cycles also increased the contamination risk in the PCR reactions. Several different approaches were used to increase the efficiency of these PCR rections but with limited success.

CHAPTER 5.

DISCUSSION.

In the initial study DNA sequencing was used to examine *tnpR* gene diversity within a collection of mercury resistant soil bacteria, from which the *tnpR* and *mer* regions had been previously characterised by RFLP analysis (Pearson *et al.*, 1996). The virtually complete congruence which was observed between DNA sequence and RFLP data validates the use of RFLP for typing studies. The relative merits of both typing procedures and the factors affecting RFLP studies were also examined.

Analysis of sequence data was carried out by construction of phylogenetic trees using two methods; DNA maximum likelihood and DNA parsimony analysis. The relative advantages and disadvatages of each method have been discussed at length (DeBry and Abele, 1995, Felsenstein, 1981, Fitch, 1971, Fukami and Tateno, 1991, Hills *et al.*, 1994, Huelsenbeck, 1995, Tateno *et al.*, 1994, Tillier and Collins, 1995) and the two methods used to construct unrooted trees produced comparable results.

The DNA sequences of eight *tnpR* PCR products from various strains along with the published sequences of Tn*501* and Tn*21* were used to construct phylogenetic trees. Similarity within classes was seen to be high, all members of class *tnpR* A were seen to be identical to each other although this may be due to the possible clonal nature of the '93 SO isolates (Pearson *et al.*, 1996). Class *tnpR* C sequences were highly conserved, dropping down to a lower level of 98.4% similarity between '93 SE23 and '93 SB3 and 29 sequences. '93 SE3 and '93 SO1 were chosen as representatives of their respective classes and all dendrograms were constructed using these sequences.

A DNA parsimony tree showing the relationship between the *tnpR* sequences obtained in this study and those contained within transposons for which DNA

sequence is available was constructed. This shows four major groupings of resolvase genes. The sequences obtained in this study fell into one of these groups, which also contains Tn*501* and Tn*21* from which the primers used in this study were designed.

In order to compare DNA sequence data with previously described RFLP data, rooted and unrooted maximum likelihood trees were constructed, indicating that the two approaches produced comparable results (Pearson *et al.*, 1996). Phylogenetic trees were also contructed using sequence from the first 186 bp of sequence and also from the second 186 bp of sequence. It was observed that the topology of the tree for the complete 372bp analysis region was similar to that of the second 186bp of sequence. The tree representing the first 186bp of sequence differed from these trees in that the positions of *tnpR* D and *tnpR* E were reversed. The restriction maps of the analysis region showed the position of the restriction endonuclease sites observed in the RFLP analysis. Sixteen of the twenty restriction endonuclease sites are contained within the second half of the region. As the second half of the sequence contains a larger number of the restriction sites originally used to classify these genes, the second half of the sequence has provided a more accurate representation of the overall analysis region. The dendrogram of the first 186bp still however shows marked structural similarity to that of the overall analysis region with the exception of the class D and E branches. This emphasises the importance of restriction enzyme choice when carrying out RFLP analysis as the position and frequency of restriction sites will affect the RFLP patterns which are observed. DNA parsimony trees of the complete fragment as well as the two halves of sequence were also constructed. Tree topology was identical to those produced using the maximum likelihood method. The DNA sequence similarity within each RFLP class was found to be high, ranging from 97.6% to 100%. This supported the previously published RFLP data.

The characterisation of *tnpR* genes contained within '93 T2 7 and '93 T2 12 indicated that '93 T2 7 should be placed into class *tnpR*-D by virtue of its 97.6% similarity to Tn*501* and its identical restriction endonuclease pattern and that '93 T2 12 should be allocated a new *tnpR* gene class, *tnpR* F on the basis of the data

presented here. For DNA sequencing to be used as a reliable alternative to RFLP for the type classification of genes, a similarity value over which any two sequences are considered to be in the same class must be established. For this study, the data presented here showed variation within classes to be as little as 2.4% and the variation within the gene to be as high as 20%. More studies are required in order to determine the general applicability of such values upon different sequences, with further analysis needed in each individual case (Osborn *et al.*, 1995). Sequencing of variants within a given gene population provides a large amount of information over that which can be acquired by RFLP data including e.g. allowing the determination of the protein sequence of the fragment. The DNA sequences in this study showed a number of base changes in the 3rd codon position with conserved 1st and 2nd codon positions based on the presumed reading frame. The partial protein sequence of *tnpR* A-E indicates that a large number of these 3rd base shifts were not translated into amino acid changes. Amongst amino acids which had been altered due to mutation, a significant number are such that an amino acid of comparable physicochemical structure had replaced the original residue. The presumed amino acid sequences obtained from '93 T2 7 and '93 T2 12 suggested that the *tnpR* genes contained within these isolates may be non-functional. The amino acid residues which are believed to be involved in the catalytic activity of this protein remain conserved throughout the 5 different classes of gene (Figure 3.2.8.9.) with the exception of the E124 residue in *tnpR* class B which appears to have undergone a E124X mutation. It has been suggested that DNA rearrangments may be responsible for the majority of polymorphisms that arise (Hall, 1994). However, this does not appear to be the case in this study as could be seen from the DNA sequences, which suggested that point mutations were responsible for the observed differences between RFLP classes.

The second aspect of this study was concerned with determining the structural arrangements of transposon associated genes in an expanded collection of 124 mercury resistant isolates. The majority of isolates contained arrangements of *tnpA* and *tnpR* genes similar to that found in the Tn*21* subgroup of bacteria. Those strains

which differed from this basic gene arrangement were '96 SE 19 and 30 which had undetermined arrangements of *tnpA* and *tnpR* genes, the four strains which did not appear to contain a *tnpR* gene and those strains which did not produce PCR products with the primers used. The majority of *mer* operons contained in the isolates fell into 2 major structural groups, the Tn*21/501*-like structures and the shorter Tn*3926*-like structures. Those isolates which did not produce PCR products using the primers described in this study may contain novel transposon associated genes which were not detectable due to the sequence variation at the primer annealing sites. The differences observed between the '93 and '96 isolates may be due to the selection of the '93 isolates by their ability to hybridise to a *mer* probe (Osborn *et al.*, 1993).

This study shows that the predominant transposon gene structures contained in a collection of mercury resistant isolates were Tn*21*-like, whereas the region of DNA between the *mer* operon and the transposase genes fell into 2 structural groups. It is interesting to note that these isolates show a distinct lack of structural diversity compared to that which might be expected from such a study if genetic recombination and rearrangement were common place. The Tn*3*-like arrangement of *tnpA* and *tnpR* genes was not observed in any of the strains using primers tn3A and tn3R, despite being detected in the soil/sediment, suggesting that Tn*3*-like structures are not associated with mercury resistance genes in this environment. Although Tn*3* transposons do not contain a *mer* operon, it would be possible to obtain a Tn*3*-like *tnpA*/R arrangement within a mercury resistance transposon via recombination between a Tn*3* transposon and a Tn*21*-like transposon. If recombination is frequent within the bacterial community, it might be expected that such events would have occured.

It is possible that the Tn*21* subgroup of transposons have a selective advantage which results in their increased abundance in polluted environments. This advantage may be due to the arrangement of *tnpA* and *tnpR* genes in these transposons, in that the genes are present on one side of the *res* site, unlike the Tn*3* arrangement of genes where the *tnpA* and *tnpR* genes are on either side of the *res* site. The Tn*21*

arrangement may allow the *tnpA/R* transposition subunit to be transferred as a complete unit whereas the Tn3 arrangement may promote recombination between the two genes. Alternatively, there may be some synergistic advantage in the relationship between the mercury resistance genes and the arrangement of genes in the Tn*21* subgroup which outcompetes alternative arrangements. This would explain the apparent lack of any other transposon structures in the isolates studied. The position of the *res* site in members of the Tn*21*-subgroup of transposons is such that it may allow recombination to occur with greater frequency between transposition genes and the genes with which they are associated, whereas the Tn*3*-like arrangement may favour recombination between the transposition genes themselves (Grinsted *et al.*, 1990). This may explain the apparent success of Tn*21*-like mercury resistance transposons in the environment.

Previous studies on transposons in the environment have concentrated either on the study of sequence diversity of transposon genes or on the isolation of novel transposon structures (Bogdanova *et al.*, 1998, Pearson *et al.*, 1996, Yurieva *et al.*, 1997). The isolation of novel sequences is important to gain an understanding of transposon diversity however this does not provide information about the relative abundance of different transposon sequences which are present in a particular environment (Bogdanova *et al.*, 1998, Pearson *et al.*, 1996, Yurieva *et al.*, 1997) Studies which have previously attempted to quantify transposon numbers in the natural environment have done so by PCR or hybridisation studies however the actual structures of the transposons in these isolates in unknown (Dahlberg and Hermansson, 1995, Zuhlsdorf and Weidemann, 1992). This study represents the first instance where the arrangement of transposon genes has been studied in bacteria isolated from soil. This information, combined with that obtained from sequence diversity studies can be used to provide an indication of the recombination and transfer events which may occur in polluted soil environments.

The effects of heavy metal contamination on gene transfer and plasmid diversity has been studied in the natural environment (De Rore *et al.*, 1994, Dronen *et*

al., 1998, Lilley *et al.*, 1994, Top *et al.*, 1995). These studies were carried out by the addition of heavy metals to soil microcosms or field sites, followed by subsequent isolation of plasmid DNA by exogenous isolation. Although mercury resistance plasmids were detected in some soils with no heavy metal pollution, the number of resistant transconjugants were generally low. The addition of heavy metals to the environment increased the amount of resistant transconjugants isolated. The diversity of the isolated plasmids was determined by RFLP analysis and was found to be high, i. e. the heavy metal resistance was not carried on a single plasmid which had transferred within the bacterial population (De Rore *et al.*, 1994, Dronen *et al.*, 1998, Lilley *et al.*, 1994, Top *et al.*, 1995). This data suggests that the observed mercury resistance may be carried on a transposon structure, capable of transferring between diverse plasmids by transposition. The relatively low structural diversity of transposons structures observed in this study is consistant with these observations. i. e. the resistance genes are carried on a transposon, which has a selective advantage in the environment and is capable of transfering between bacteria by incorporation into plasmid structures. The genes which are carried on these transposons may undergo mutation and recombination events leading to the establishment of a diverse population of transposon sequences with similar structural arrangements.

The genetic diversity of the '93 isolates has previously been studied in detail (Bruce *et al.*, 1997, Holt *et al.*, 1996, Osborn *et al.*, 1993, Osborn *et al.*, 1995, Pearson *et al.*, 1996). The identity of the '93 strains has been previously determined by API analysis (Osborn *et al.*, 1993). No correlation was apparent between the bacterial species and the arrangement of genes contained within the strains. The sequence diversity of both *tnpA* and *tnpR* genes has been studied by RFLP and DNA sequencing (Holt *et al.*, 1996, Pearson *et al.*, 1996). The previous RFLP study carried out by Pearson *et al*, 1996, indicated the presence of 3 classes of *tnpR* gene and 6 classes of *tnpA* gene within the strains isolated from soil and sediment. There was no observed linkage between different classes of the two genes, suggesting that recombination is frequent between the *tnpA* and *tnpR* genes and between *mer* and *tnp*

genes within the '93 isolates (Pearson *et al.*, 1996). The data obtained from the *tnpR* sequencing study suggested that the diversity in the *tnpR* genes may have been due to point mutations, however this appears not to be the case. These data compared with the data presented in this study suggests that recombination may have occurred between closely related transposons, but that this has not affected the actual structural arrangement of the genes.

CHAPTER 6.

REFERENCES.

Abdel-Meguid, S. S., Gringley, N. D. F., Templeton, N. S. and Steitz, T. A. (1984). Cleavage of the site specific recombination protein γδ resolvase: the smaller of two fragments binds DNA specifically. Proc. Nat. Aced. Sci. U.S.A. 81: 2001-2005.

Allmeier, H., Cresnar, B., Greck, M. and Schmitt, R. (1992). Complete nucleotide sequence of Tn*1721*: gene organisation and a novel gene product with features of a chemotaxis protein. Gene. 111: 11-20.

Ampe, F. ben Omar, N. and Guyot, J. P. (1998). Recovery of total microbial RNA from lactic acid fermented foods with a high starch content. Lett. Appl. Microbiol. 27: 270-274.

Arakawa, Y., Murakami, M., Suzuki, K., Ito, H., Wacharotayankun, R., Ohsuka, S., Kato, N. and Ohta, M. (1995). A novel integon-like element carrying the metallo-β-lactamase gene bla_{IMP}. Antimicrob. Chemother. 39: 1612-1615.

Arber, W., Naas, T. and Blot, M. (1994). Generation of genetic diversity by DNA rearrangements in resting bacteria. FEMS Microbiol. Ecol. 15: 5-14.

Arciszewska, L. K., Drake, D. and Craig, N. L. (1989). Transposon Tn*7* cis acting sequences in transposition and transposition immunity. J. Mol. Biol. 207: 35-52.

Arnold, P. H., Blake, D. G., Grindley, N. D. F., Boocock, M. R. and Stark, W. M. (1999). Mutants of Tn*3* resolvase which do not require accessory binding sites for recombinational activity. EMBO. 18: 1407-1414.

Arthur, A., Nimmo., E. R., Hettle, S. J. H. and Sherratt, D. J. (1984). Transposition and transposition immunity of transposon Tn*3* derivatives having different ends. EMBO. 3: 1723-1729.

Arthur, M., Molinas, C., Depardieu, F. and Courvalin, P. (1993). Characterisation of Tn*1546*, a Tn*3*-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptidoglycan precursors in *Enterococcus faecium* BM4147. J. Bacteriol. 175:117-127.

Austin, S. and Nordström, K. (1990). Partition mediated incompatibility of bacterial plasmids. Cell. 60: 351-354.

Avila, P., de la Cruz, F., Ward, E. and Grinsted, J. (1984). Plasmids containing one end of Tn*21* can fuse with other plasmids in the presence of Tn*21* transposase. Mol. Gen. Genet. 195: 288-293.

Bale, M. J., Day, M. J. and Fry, J. C. (1988). Novel method for studying plasmid transfer in undisturbed river epilithon. Appl. Environ. Microbiol. 54: 2756-2758.

Barbieri, P., Bestetti, G., Reniero, D and Galli, E. (1996). Mercury resistance in aromatic compound degrading *Pseudomonas* strains. FEMS Microbiol. Ecol. 20: 185-194.

Barkay, T., Fouts, D.L. and Olson, B.H. (1985). Preparation of a DNA gene probe for detection of mercury resistance genes in Gram-negative bacterial communities. Appl. Environ. Microbiol. 49: 686-692.

Barnes, W. M. (1994). PCR amplification of up to 35kb DNA with high fidelity and high yield from λ bacteriophage templates. Proc. Natl. Acad. Sci. USA. 91: 2216-2220.

Baum, J. A. (1994). Tn*5401*, a new Class II transposable element from *Bacillus thuringiensis*. J. Bacteriol. 176: 2835-2845.

Beliveau, B. H. and Trevors, J. T. (1990). Mercury resistance determined by a self transmissible plasmid in *Bacillus cereus*. Bio. Meth. 3: 188-196.

Bennett, P.M., Grinsted, J., Choi, C.L and Richmond, M.H. (1978). Characterisation of Tn*501*, a transposon determining resistance to mercuric ions. Mol. Gen. Genet. 159: 101-106.

Bennett, P.M., de la Cruz, F. and Grinsted, J. (1983). Cointegrates are not obligatory intermediates in transposition of Tn*3* and Tn*21*. Nature. 305: 743-744.

Bertin, A. (1995). Comparison of several procedures for plasmid profile determination in *Escherichia Coli*. J. Microbiol. Meth. 22: 109-117.

Birnbiom, H.C and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7: 1513-1523.

Bishop, R. and Sherrat, D. (1984). Transposon Tn*1* intra-molecular transposition. Mol. Gen. Genet. 196: 117-122.

Bissonnette, L., Champetier, S., Buisson, J. P. and Roy, P. H. (1991). Characterisation of the non enzymatic chloramphenicol resistance (*cmlA*) gene of the In4 integron of Tn*1696*: similarity of the product to transmembrane transport proteins. J. Bacteriol. 173: 4493-4502.

Bogdanova, E. S., Bass, I. A., Minakhin, L. S., Petrova, M. A., Mindlin, S. Z., Volodin, A. A., Kalyaeva, E. S., Tiedje, J. M., Hobman, J. L., Brown, N. L and Nikiforov, V. G. (1998). Horizontal spread of *mer* operons among Gram-positive bacteria in natural environments. Microbiology. 144: 609-620.

Bonafede, M. E., Carias, L. L. and Rice, L. B. (1997). Enterococcal transposon Tn*5384*: evolution of a composite transposon through cointegration of enterococcal and staphylococcal plasmids. Antimicrob. Chemother. 41: 1854-1858.

Boocock, M. R., Zhu., X. and Grindley, N. (1995). Catalytic residues of γδ resolvase activity *in cis*. EMBO. 14: 5129-5140.

Brown, N. L. (1985). Bacrerial resistance to mercury-*reductio ad absurdum?* Trends in Biochem. Sci. 10: 400-403.

Brown, N. L., Winnie, J. N., Fritzinger, D and Pridmore, R. D. (1985a). The nucleotide sequence of the *tnp*A gene completes the sequence of the *Pseudomonas* transposon Tn*501*. Nucleic Acids Res. 13: 5657-5669.

Brown, N. L., Misra, T. K., Winnie, J. N., Schmitt, A., Seiff, M. and Silver, S. (1985b). The nucleotide sequence of the mercuric resistance operons of plasmid R100 and transposon Tn*501*: further evidence for *mer* genes which enhance the activity of the mercuric ion detoxification system. Mol. Gen. Genet. 202: 143-151.

Brown, N. L., Camakaris, J., Lee, B. T. O., Williams, T., Morby, A. P., Parkhill, J. and Rouch, D. A. (1991). Bacterial resistances to mercury and copper. J. Cell. Biochem. 46: 106-114.

Brown, N. L. and Evans, L. R.. (1991). Transposition in prokaryotes- Transposon Tn*501*. Res. Microbiol. 142: 689-700.

Bruce, K. D., Hiorns, W. D., Hobman, J. L., Osborn, A. M., Strike, P. and Ritchie, D. A. (1992). Amplification of DNA from native populations of soil bacteria by using the polymerase chain reaction. Appl. Environ. Microbiol. 58: 3413-3416.

Bruce, K. D., Osborn, A. M., Pearson, A. J., Strike, P. and Ritchie, D. A. (1995). Genetic diversity within *mer* genes directly amplified from communities of non-cultivated soil bacteria. Mol. Ecol. 4: 605-612.

Bruce, K.D., Osborn, A. M., Harrison, J. F., Strike, P. and Ritchie, D. A. Diversity within the *mer*ΔPA region of bacterial mercury resistance determinants; implications for gene evolution in environments. Submitted for publication to Applied and Environmental Microbiology.

Chiou, C. S. and Jones, A. L. (1993). Nucleotide sequence analysis of a transposon (Tn*5393*) carrying streptomycin resistance genes in *Erwinia amylovora* and other Gram- negative bacteria. J. Bacteriol. 175: 732-740.

Chou, J., Casadaban, M. J., Lemaux, P. G. and Cohen, S. N. (1979). Identification and characterisation of a self regulated repressor of translocation of the Tn*3* element. Proc. Natl. Acad. Sci. U.S.A. 76: 4020-4024.

Clennel, A. M., Johnston, B. and Rawlings, D. E. (1995). Structure and function of Tn*5467*, a Tn*21*-like transposon located on the *Thiobacillus ferrooxidans* broad host range plasmid pTF-FC2. Appl. Environ. Microbiol. 61: 4223-4229.

Colonna, B., Bernardini, M., Micheli, G., Maimone, F., Nicoletti, M. and Casalino, M. (1988). The *Salmonella wein* virulence plasmid pZM3 carries Tn*1935*, a multi resistance transposon containing a composite IS*1936*-kanamycin resistance transposon. Plasmid. 20: 221-231.

Cooksey, D. A. (1994). Molecular mechanisms of copper resistance and accumilation in bacteria. FEMS Microbiol. Rev. 14: 381-386.

Craig, N. L. (1991). Tn7 a target site specific transposon. Mol. Microbiol. 11: 2569-2573.

Craig, N. L. (1996). Transposon Tn*7*. Curr. Top. Microbiol. Immun. 204: 27-48.

Crellin, P. K. and Rood, J. I. (1998). Tn*4451* from *Clostridium perfringens* is a mobilisable transposon that encodes the functional Mob protein, TnpZ. Mol. Microbiol. 27: 631-642.

de la Cruz, F. and J. Grinsted. (1982). Genetic and molecular characterisation of Tn*21*, a multiple resistance transposon from R100.1. J. Bacteriol. 151: 222-228.

Dahlberg, C. and M. Hermansson. (1995). Abundance of Tn*3*, Tn*21* and Tn*501* transposase (*tnpA*) sequences in Bacterial Community DNA from Marine Environments. Appl. Environ. Microbiol. 61: 3051-3056.

Dahlberg, C., Linberg, C., Torsvik, V. L. and Hermansson, M. (1997). Conjugative plasmids isolated from bacteria in marine environments show various degrees of homology to each other and are not closely related to well characterised plasmids. Appl. Environ. Microbiol. 63: 4692-4697.

Dahlberg, C., Bergström, M., Andreasen, M., Christensen, B. B., Molin, S and Hermansson, M. (1998a). Interspecies bacterial conjugation by plasmids from marine environments visualised by *gfp* expression. Mol. Biol. Evol. 15: 385-390.

Dahlberg, C., Bergström, M and Hermansson, M. (1998b). In situ detection of high levels of horizontal plasmid transfer in marine bacterial communities. Appl. Environ. Microbiol. 64: 2670-2675.

Davis, B. D., Dulbecco, R., Eisen, H. N and Ginsberg, H. (1980). Microbiology. 3rd edition. Harber and Row Publishers, Philadelphia.

DeBry, R.W. and Abele, L.G . (1995). The relationship betwen parsimony and maximum likelihood analyses: Tree scores and confiidence estimates for three data sets. Mol. Biol. and Evol. 12: 291-297.

Diver, W.P., Grinsted, J., Fritzinger, D.C., Brown, N.L., Altenbuchner, J., Rogowsky, P. and Schmitt, R. (1983). DNA sequences of and complementation by the *tnp*R genes of Tn*21*, Tn*501* and Tn*1721*. Mol. Gen. Genet. 191: 189-193.

Domenico, P., Marx, J. L., Schoch, P. E. and Cunha, B. A. (1992). Rapid plasmid DNA isolation from mucoid Gram-Negative bacteria. J. Clinical. Microbiol. 30: 2859-2863.

Dronen, A. K., Torsvik, V., Goksoyr, J. and Top, E. M. (1998). Effect of mercury addition on plasmid incidence and gene mobilising capacity in bulk soil. FEMS Microbiol. Ecol. 27: 381-394.

Duarte, G. F., Rosado, A. S., Seldin, L. and Keijzer-Wolters, A. C. (1998). Extraction of ribosomal RNA and genomic DNA from soil for studying the diversity of the indigenous bacteria community. J. Microbiol. Meth. 32: 21-30.

Ellis, R. J., Thompson, I. P. and Bailey, M. J. (1995). Metabolic profiling as a means of characterising plant associated microbial communities. FEMS Microbiol. Ecol. 16: 9-18.

Evans, L. R. and Brown, N. L. (1987). Construction of hybrid Tn*501*/Tn*21* transposases *in vivo*: identification of a region of transposase conferring specificity of recognition of the 38bp terminal repeat sequences. EMBO. 6: 2849-2853.

Felsenstein, J. (1981). Evolutionary trees from DNA sequence: A maximum likelihood approach. J. Mol. Evol. 17: 368-376.

Felsenstein, J. (1993). PHYLIP (Phylogeny Inference Package) version 3.5C, Department of Genetics, University of Washington, Seattle.

Fiandt, M., Szybalski, W. and Malxmy, M. H. (1972). Polar mutations in *lac*, *gal* and phage λ consist of a few DNA sequences inserted in either orientation. Mol. Gen. Genet. 119: 223.

Fitch, W.M. (1971). Toward defining the course of evolution: minimum change needed for a specific tree topology. System. Zoo. 20: 406-416.

Foster, T. J. (1987). The genetics and biochemistry of mercury resistance. CRC Crit. Rev. Microbiol. 15: 117-140.

Franke, A. E. and Clewell, D. B. (1981). Evidence for a chromosome borne resistance transposon (Tn*916*) in *Streptococcus faecalis* that is capable of conjugal transfer in the absence of a conjugative plasmid. J. Bacteriol. 145: 494-502.

Fukami, K. and Tateno,Y. (1991). Robustness of maximum likelihood tree estimation against different patterns of base substitutions. J. Mol. Evol. 32: 79-91.

Galas, D. J. and Chandler, M. (1989). Bacterial insertion sequences. pp109-162. In: Mobile DNA. Eds D. E. Berg and M. M. Howe. Amer. Soc. Microbiol., Washington D.C.

Genetics Computer Group. Program for the Wisconsin Package. Version 8, Sept 1994. Genetics Computer Group, 575 Science Drive, Madison, Wisconsin. USA, 53711.

Gilbert, M. P and Summers, A. O. (1988). The distribution and divergence of DNA sequences relating to the Tn*21* and Tn*501 mer* operons. Plasmid. 20: 127-136.

Gill, R., Heffron, E., Dougan, G. and Falkow, S. (1978). Analysis of sequences transposed by complementation of two classes of transposition defficient mutants of transposable element Tn*3*. J. Bacteriol. 136: 742-756.

di Gioia, D., Peel, M., Fava, F and Wyndham, R. C. (1998). Structures of homologous transposons carrying *cbaABC* genes from Europe and North America. Appl. Environ. Microbiol. 64: 1940-1946

Götz, A., Pukall, R., Smit, E., Tietze, E., Prager, R., Tschäpe, H., van Elsas, J. D. and Smalla, K. (1996). Detection and characterisation of broad host range plasmids in environmental bacteria by PCR. Appl. Environ. Microbiol. 62: 2621-2628.

Gray, J. P and Herwig, R. P. (1996). Phylogenetic analysis of the bacterial communities in marine sediments. Appl. Environ. Microbiol. 62: 4049-4059.

Griffith, F. (1928). Significance of pneumococcal types. J. Hyg. 27: 113-159.

Griffiths, B. S., Ritz, K. and Glover, L. A. (1996). Broad scale approaches to the determination of soil microbial community structure: Application of the community DNA hybridisation technique. Microbial Ecology. 31: 269-280.

Grindley, N. D. F., Lauth, M. R., Wells, R. G., Wityk, R. J., Salvo, J. J. and Reed, R. R. (1982). Transposon mediated site specific recombination: identification of three binding sites for resolvase at the *res* sites of γδ and Tn*3*. Cell: 19-27.

Grinsted, J., de la Cruz, F., *Altenbuchner J. and Schmitt, R. (1982). Complementation* of transposition of *tnp*A mutants of Tn*3*, Tn*21*, Tn *501* and Tn*1721*. Plasmid. 8: 276-286.

Grinsted, J., Martin, C and de la Cruz, F. (1988). Factors that affect transposition mediated by the Tn*21* transposase. Plasmid. 20: 54-60.

Grinsted, J. De la Cruz, F and Schmidt, R. (1990). The Tn*21* subgroup of bacterial transposable elements. Plasmid. 24: 163-189

Güssow, D. and Clackson, T. (1989). Direct clone characterisation from plaques and colonies by the polymerase chain reaction. Nucleic. Acids. Res. 17: 4000.

Guyer, M. S., Reed, R. R., Steitz, J. A. and Low, K. B. (1980). Identification of a sex factor affinity site in *E.Coli* as gamma-delta. Cold. Spr. Harb. Symp. Quant. Biol. 45: 135-140.

Halford, S., Jordan, S. L. and Kirkbride, E. A. (1985). The resolvase protein from the transposon Tn*21*. Mol. Gen. Genet. 200: 169-175.

Hall, L.M.C. (1994). Are point mutations or DNA rearrangements responsible for the restriction fragment length polymorphisms that are used to type bacteria? Microbiology. 140: 197-204.

Hall, R.M., Brookes, D. E. and Stokes, H.W. (1991). Site specific insertion of genes into integrons: role of the 59-base element and determination of the recombination crossover point. Mol. Microbiol. 5: 1941-1959.

Hall, R. M. and Collis, C. M. (1998). Antibiotic resistance in gram-negative bacteria: the role of gene cassettes and integrons. Drug. Resistance Updates. 1: 109-119.

Harding, K. (1996). The potential for horizontal gene transfer within the environment. Agro. Food Industry Hi-tech. 7: 31-35.

Harshey, R. M. and Bukhari, A. I. (1981). A mechanism of DNA transposition. Proc. Natl. Acad. Sci. U.S.A. 78: 1090-1094.

Hart, M. C., Elliott, G. N., Osborn, A. M., Ritchie, D. A. and Strike, P. (1998). Diversity amongst *Bacillus merA* genes amplified from mercury resistant isolates and directly from mercury polluted soil. FEMS Microbiol. Ecol. 27: 73-84 and 395.

Håvarstein, L. S. (1998). Bacterial gene transfer by natural genetic transformation. APMIS. 106: 43-46.

Head, I. M., Saunders, J. R and Pickup, R. W. (1998). Microbial evolution, diversity and ecology: A decade of ribosomal RNA analysis of uncultivated microorganisms. Microb. Ecol. 35: 1-21.

Hedges, R. W. and *Jacob, A. (1974).* Transposition of ampicillin from RP4 to other replicons. Mol. Gen. Genet. 132: 31-40.

Heffron, F., Kostriken, R., Morita, C. and Parker, R. (1981). Tn*3* encodes a site specific recombination system: identification of essential sequences, genes and actual site of recombination. Cold Spring Harbor Sypm. Quant. Biol. 45: 259-268.

Heltzel, A., Gambill, D., Jackson, W. J., Totis, P. A. and Summers, A. O. (1987). Overexpression and DNA binding properties of the *mer* encoded regulatory proteins from plasmid NR1 (Tn*21*). J. Bacteriol. 169: 3379-3384.

Hermansson, M., Jones, G. W. and Kjelleberg, S. (1987). Frequency of antibiotic and heavy metal resistance, pigmentation and plasmids in bacteria of the marine air-water interface. Appl. Environ. Microbiol. 53: 2338-2342.

Hermansson, M. and Linberg, C. (1994). Gene transfer in the marine environment. FEMS Microbiol. Ecol. 15: 47-54.

Higgins, I. J. and Burns, R. G. (1975). Metals. In: The chemistry and microbiology of pollution. Ch 7: pp189-210. Academic Press.

Hillis, D.M., Huelsenbeck, J.P. and Cunningham, C.W. (1994). Application and accuracy of molecular phylogenies. Science. 264: 671-677.

Hillis, D. M. (1997). Phylogenetic analysis. Curr. Biol. 7: R129-R131.

Hobman, J. L., Kholodii, G., Nikiforov, V., *Ritchie, D. A., Strike, P. and Yurieva, O.* (1994). The sequence of the *mer* operon of pMER327/419 and transposon ends of pMER327/419, 330 and 05. Gene. 146: 73-78.

Hobman, J. L. and Brown, N. L. (1997). Bacterial mercury resistance genes. Met. Ions. Biol. Syst. 34: 527-568.

Holmes, D. S. and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Analyt. Biochem. 114: 193-197.

Holt, R. J., Strike, P. and Bruce, K. D. (1996). Phylogenetic analysis of *tnpR* genes in mercury resistant soil bacteria: the relationships between DNA sequencing and RFLP typing approaches. FEMS Microbiol. Letts. 144: 95-102.

Holt, R. J., Bruce, K. D. and Strike, P. (1999). Conservation of transposon structures in soil bacteria. FEMS Microbiol. Ecol. 30: 25-37.

Huelsenbeck, J.P. (1995). The robustness of two phylogenetic methods: Four-taxon simulations reveal a slight superiority of maximum likelihood over neighbor joining. Mol. Biol. and Evol. 12: 843-849.

Hughes, M. N. and Poole, R. K. (1991). Metal speciation and microbial growth: the hard (and soft) facts. J. Gen. Microbiol. 137: 725-734.

Hyde, D.R. and Tu, C. P. D. (1985). *tnp*M, a novel regulatory gene that enhances Tn*21* transposition and supresses cointegrate resolution. Cell. 42: 629-638.

Inoue, C., Kusano, T. and Silver, S. (1996). Mercuric ion uptake by *Escherichia coli* cells producing *Thiobacillus ferrooxidans merC*. Biosci. Biotechnol. Biochem. 60: 1289-1292.

Ippen-Ihler, K. (1988) Bacterial conjugation. In: Gene transfer in the environment. Levy, S.B. and Miller, R.V.) p33-72. McGraw Hill Inc.

Ishizaki, K. and Ohtsubo, E. (1985). Cointegration and resolution mediated by IS101 present in plasmid pSC101. Mol. Gen. Genet. 199: 388-395.

Ivey, D. M., Guffanti, A. A., Shen, Z., Kudyan, N. and Krulwich, T. A. (1992). The *cadC* gene product of alkaliphilic *Bacillus firmus* OF4 partially restores Na^+ resistance to an *Escherichia coli* strain lacking an Na^+/H^+ antiporter (NhaA). J. Bacteriol. 174: 4878-4884.

Jobling, M. G. and Ritchie, D. A.(1987). Genetic and physical analysis of plasmid genes expressing inducible resistance of tellurite in *Escherichia Coli*. Mol. Gen. Genet. 208: 288-293.

Jobling, M. G., Peters, S. E. and Ritchie, D. A. (1988a). Plasmid borne mercury resistance in aquatic bacteria. FEMS Microbiol. Letts. 49: 31-37.

Jobling, M. G., Peters, S. E. and Ritchie, D. A. (1988b). Restriction pattern and polypeptide homology among plasmid borne mercury resistance determinants. Plasmid. 21: 106-112.

Jordan, T., Saedler, H. and Starlinger, P. (1968). O^c and strong polar mutations in the *gal* operon are insertions. Mol. Gen. Genet. 102: 353.

Kado, C. I. and Liu, S. T. (1981). Rapid procedure for detection and isolation of large and small plasmids. J. Bacteriol. 145: 1365-1373.

Kahn, K. and Scharfer, M. R. (1995). Characterisation of transposon Tn*5469* from the cyanobacterium *Fremyella diplosiphon*. J. Bacteriol. 177: 7026-7032.

Kennedy, A. C and Gewin, V. L. (1997). Soil microbial diversity: present and future considerations. Soil Science. 162: 607-617.

Kholodii, G. Ya., Gorlenko, Z. M., Lomovskaya, O. L., Mindlin, S. Z., Yurieva, O. V. and Nikiforov, V. G. (1993a). Molecular characterisation of an abberant mercury resistance transposable element from an environmental *Acinetobacter* strain. Plasmid 30: 303-308.

Kholodii, G. Ya., Yurieva, O. V., Lomovskaya, O. L., Gorlenko, Zh. M., Mindlin, S. Z. and Nikiforov, V. G. (1993b). Tn*5053*, a mercury resistance transposon with integrons ends. J. Mol. Biol. 230: 1103-1107.

Kholodii, G. Ya., Mindlin, S. Z., Bass, I. A., Yurieva, O. V., Minakhina, S. V. and Nikiforov, V. G. (1995). Four genes, two ends, and a *res* region are involved in transposition of Tn*5053*: a paradigm for a novel famliy of transposons carrying either a *mer* operon or an integron. Mol. Microbiol. 17: 1189-1200.

Kholodii, G. Ya., Yurieva, O. V., Gorlenko, Zh. M., Mindlin, S. Z., Bass, I. A., Lomovskaya, O. L., Kopteva, A. V. and Nikiforov, V. G. (1997). Tn*5041*: a chimeric mercury resistance transposon related to the toluene degradative transposon Tn*4651*. Microbiology. 143: 2549-2556.

Khrishnapillai, V., Royle, P. and Lehrer, J. (1981). Insertions of the transposon Tn*1* into the *Pseudomonas aeruginosa* chromosome. Genetics. 97: 495-511.

Kirby, R. (1992). Evolutionary origins of the class A and B β-lactamases. J. Mol. Evol. 34: 345-350.

Kiyono, M. and Pan-Hou, H. (1999). The *merG* gene product is involved in phenyl mercury resistance in *Pseudomonas* strain K-62. J. Bacteriol. 181: 726-730.

Kleckner, N. (1981). Transposable elements in prokaryotes. Ann. Rev. Genet. 15: 341-404.

Kratz, J., Schmidt, F. and Wiedemann, B. (1983). Characterization of Tn*2411* and Tn*2410*, two related transposon derived from R-plasmid R1767 and related to Tn*2603* and Tn*21*. J. Bacteriol. 155: 1333-1342.

Kuske, C. R., Banton, K. L., Adorada, D. L., Stark, P. C., Hill., K. K and Jackson, P. J. (1998). Small scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. Appl. Environ. Microbiol. 64: 2463-2472.

Ladagga, R. A., Chu, L., Misra, T. K. and Silver, S. (1987). Nucleotide sequence and expression of the mercury resistance operon from *Staphylococcus aureus* plasmid pI258. Proc. Natl. Acad. Sci. USA. 84: 5106-5110.

Lawrence, J. G. and Ochman, H. (1998). Molecular archaeology of the *Escherichia coli* genome. Proc. Nat Acad. Sci.USA. 95: 9413-9417.

Lebaron, P., Batailler, N. and Baleux, B. (1994). Mobilization of a recombinant nonconjugative plasmid at the interface between wastewater and the marine coastal environment. FEMS Microbiol. Ecol. 15: 61-70.

Lebrun, M., Audurier, A. and Cossart, P.(1994). Plasmid borne cadmium resistance genes in *Listeria monocytogenes* are present on Tn*5422*, a novel transposon closely related to Tn*917*. J. Bacteriol. 176: 3049-3061.

Lee, D. H., Zo, Y. G. and Kim, S. J. (1996). Non radioactive method to study genetic profiles of natural bacterial communities by PCR single stranded conformation polymorphism. Appl. Environ. Microbiol. 62: 3112-3120.

Leff, L. G., Dana, J. R., McArthur, J. V and Shimkets, L. J. (1995). Comparison of methods of DNA extraction from stream sediments. Appl. Environ. Microbiol. 61: 1141-1143.

Lereclus, D., Mahillon, J., Menou, G. and Lecadet, M. M. (1986). Identification of Tn*4430*, a transposon in *Bacillus thuringiensis* functional in *Escherichia coli*. Mol. Gen. Genet. 204: 52-57.

Lett, M. C., Bennet, P. M. and Vidon, D. J. M. (1985) Characterisation of Tn*3926* a new mercury resistance transposon from *Yersinia enterocolitica*. Gene. 40:79-91.

Leung, K., England, L. S., Cassidy, M. B., Trevors, J. T. and Weir, S. (1994). Microbial diversity in soil: effect of releasing genetically engineered microorganisms. Mol. Ecol. 3: 413-422

Liebert, C. A.,Wireman, J., Smith, T. and Summers, A. O. (1997). Phylogeny of mercury resistance (*mer*) operons of gram-negative bacteria isolated from the faecal flora of primates. Appl. Environ. Microbiol. 63: 1066-1076.

Lilley, A. K., Bailey, M. J., Day, M. J. and Fry, J. C. (1996). Diversity of mercury resistance plasmids obtained by exogenous isolation from the bacteria of sugar beet in tree successive years. FEMS Microbiol. Ecol. 20: 211-227.

Liu, C.-C., Hühne, R., Tu, J., Lorbach, E. Dröge. (1998). The resolvase encoded by *Xanthomonas campestris* transposable element *ISXc5* constitutes a new subfamily closely related to DNA invertases. Genes to Cells. 3: 221-233.

Lorenz, M. G. and Wackernagel, W. (1994). Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Rev. 58: 563-602.

Lund, P. A., Ford, S. J. and Brown, N. L. (1986). Transcriptional regulation of the mercury resistance genes of transposon Tn*501*. J. Gen. Microbiol. 132: 465-480.

Mäe, A. A. and Heinaru, A. L. (1994). Transposon-mediated mobilisation of chromosomally located catabolic operons of the CAM plasmid by TOL plasmid transposon Tn*4652* and CAM plasmid transposon Tn*3614*. Microbiology. 140: 915-921.

Maekawa, T., Amemura-Maekawa, J. and Ohtsubo, E. (1993). DNA binding domains in Tn*3* transposase. Mol. Gen Genet. 236: 267-274.

Mahillon, J., Rezsöhazy, R., Hallet, B. and Declour, J. (1994). IS*231* and other *Bacillus thuringiensis* transposable elements: a review. Genetica. 93: 13-26.

Mahillon, J. and Seurinck, J. (1988). Complete sequence of pGI2, a *Bacillus thuringiensis* plasmid containing Tn*4430*. Nucleic Acids Res. 15: 11827-11828.

Mahillon, J. (1998). Transposons as gene haulers. APMIS. 106: 29-36.

Mahillon, J and Chandler, M. (1998). Insertion sequences. Microbiol. Mol. Biol. Rev. 62: 725-774.

Martin, C., Grinsted, J. and de la Cruz, F. (1989). Effects of variation of inverted repeat sequences on reactions mediated by the transposase of Tn*21*. J. Bacteriol. 171: 3996-4001.

McClintock, B. (1951). Chromosome organisation and gene expression. Cold Spr. Harb. Symp. Quant. Biol. 16: 13-43.

van de Meer, J. R., Zehnder, A. J. B and de Vos, W. M. Identification of a novel composite transposable element, Tn*5280*, carrying chlorobenzene dioxygenase genes of *Pseudomonas* sp. strain P51. J. Bacteriol. 173: 7077-7083.

Michiels, T., Cornelis, G., Ellis, K. and Grinsted, J. (1987). Tn*2501*, a component of the lactose transposon Tn*951*, is an example of a new catagory of Class II transposable elements. J. Bacteriol. 169: 624-631.

Misra, T. K. (1992). Bacterial resistance to inorganic mercury salts and organomercurials. Plasmid. 27: 4-16.

Motsch, S. and Schmitt, R. (1984). Replicon fusion mediated by a single ended derivative of transposon Tn*1721*. Mol. Gen. Genet. 195: 281-287.

Muyzer, G and Smalla, K. (1998). Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. Antonie van Leeuwenhoek. 73: 127-141.

Nakatsu, C., Ng, J., Singh, R., Straus, N. and Wyndham, C. (1991). Chlorobenzoate catabolic transposon Tn*5271* is a composite Class I element with flanking Class II insertion sequences. Proc. Natl. Acad. Sci. 88: 8312-8316.

Ni'Bhriain, N. N., Silver, S. and Foster, T. J. (1986). Polypeptides specified by the mercuric resistance (*mer*) operon of plasmid R100. Gene. 42: 323-330.

Nocick, R. P. (1987). Plasmid incompatibility. Microbiol. Rev. 51: 381-395.

Nurk, A., Tamm, A., Hôrak, R. and Kivissar, M. (1993). In-vivo generated fusion promoters in *Pseudomonas putida*. Gene: 23-29.

Ochiai, E. I. (1977). Bioinorganic chemistry: An introduction. Allyn and Bacon. Boston.

Ogram, A., Sayler, G.S. and Barkay, T. (1987). The extraction and purification of microbial DNA from sediments. J. Microbiol. Meth. 7: 57-66.

Olsen, J. E. (1990). An improved method for rapid isolation of plasmid DNA from wild type Gram-negative bacteria for plasmid restriction profile analysis. Letts. Appl. Microbiol. 10: 209-212.

Olson, B.H., Lester, J.N., Cayless, S.M. and Ford, S. (1989). Distribution of mercury resistance determinants in bacterial communities of river sediments. Water. Res. 23: 1209-1217.

Olson, B.H., Cayless, S.M., Ford, S. and Lester, J.N. (1991). Toxic element contamination and the occurrence of mercury resistant bacteria in Hg- contaminated soil, sediments and sludges. Arch. Environ. Cont. Tox. 20: 226-233.

Osborn, A. M., Bruce, K. D., Strike,P. and Ritchie, D. A. (1993). Polymerase chain reaction - Restriction fragment length polymorphism analysis shows divergence among *mer* determinants from Gram-negative bacteria indistinguishable by DNA - DNA hybridisation. Appl. Environ. Microbiol. 59: 4024-4030.

Osborn, A. M., Bruce, K. B., Strike, P. and Ritchie, D. A. (1995). Sequence conservation between regulatory mercury resistance genes in bacteria from mercury polluted and pristine environments. System. Appl. Microbiol. 18: 1-6.

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

Osbourn, S. E. V., Turner, A. K. and Grinsted, J. (1995). Nucleotide sequence within Tn*3926* confirms this as a Tn*21*-like transposable element and provides evidence for the origin of the *mer* operon carried by the plasmid pKLH2. Plasmid. 33: 65-69.

Oullette, M. and Roy, P. H. (1987). Homology of orfs from Tn*2603* and from R46 to site specific recombinases. Nucleic Acids Res. 15: 10055.

Øvreås, L., Jensen, S., Daae, F. L. and Torsvik, V. (1998). Microbial community changes in a perturbed agricultural soil investigated by molecular and physiological approaches. Appl. Environ. Microbiol. 64: 2739-2742.

Page, R. D. M. (1996). Treeview: An application to display phylogenetic trees on personal computers. Comp. Apps. 12: 357-358.

Paget, E. and Simonet, P. (1994). On the track of natural transformation in soil. FEMS Microbiol. Ecol. 15: 109-118.

Pan-Hou, H. S. K. and Imura, N. (1981). Role of hydrogen sulfide in mercury resistance determined by plasmid of *Clostridium cochlearium* T2. Arch. Microbiol. 130: 49-52.

Pan-Hou, H. S. K., Nishimoto, M. and Imura, N. (1981). Possible role of membrane proteins in mercury resistance of *Enterobacter aerogenes*. Arch. Microbiol. 130: 93-95.

Peters, S. E., HobmanJ. L., Strike, P. and Ritchie, D. A. (1991). Novel mercury resistance determinants carried by IncJ plasmids pMERPH and R391. Mol. Gen. Genet. 228: 294-299.

Pearson, A. J., Bruce, K.D., Osborn, A.M., Strike, P and Ritchie, D. A. (1996). The distribution and association of class II transposase and resolvase genes with *mer* genes in soil bacteria. Appl. Environ. Microbiol. 62: 2961-2965.

Pillai, S. D., Josephson, K. L., Bailey, R. L., Gerba, C. P. and Pepper, I. L. (1991). Rapid method for processing soil samples for polymerase chain reaction amplification of specific gene sequences. Appl. Environ. Microbiol. 57: 2283-2286.

Pitcher, D.G., Saunders, N.A. and Owen, R.J. (1989). Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. Lett. Appl. Microbiol. 8: 151-156.

Polz, M. F. and Cavanaugh, C. M. (1998). Bias in template to product ratios in multitemplate PCR. Appl. Environ. Microbiol. 64: 3724-3730.

Radford, A. J., Oliver, J., Kelley, W. J. and Reanney, D. C. (1981). Translocatable resistance to mecuric and phenylmercuric ions in soil bacteria. J. Bacteriol. 147: 1110-1112.

Rådstrom, P., Skold, O., Swedberg, G., Flensburg, J., Roy, P. H. and Sundström, L.(1994). Transposon Tn*5090* of plasmid R751, which carries an integron, is related to Tn*7*, Mu and the retroelements. J. Bacteriol. 176: 3257-3268.

Recchia, G. D. and Hall, R. M.. (1995). Gene cassettes: a new class of mobile element. Microbiology. 141: 3015-3027.

Reniero, D., Mozzon, E., Galli, E. and Barbieri, P. (1998). Two aberrant mercury resistance transposons in the *Pseudomonas stutzeri* plasmid pPB. Gene: 208: 37-42.

Rice, L. B. and Carias, L. L. (1998). Transfer of Tn*5385*, a composite, multiresistance chromosomal element from *Enterrococcus faecalis*. J. Bacteriol. 180: 714-721.

Rinkel, M., Hubert, J. C., Roux, B. and Lett, M. C. (1994). Transposon Tn*5403*, a mobilisation-helper element: Complete nucleotide sequence and distribution in aquatic strains. FEMS Microbiol. Ecol. 15: 89-96.

Roane, T. M. and Kellogg, S. T. (1996). Characterisation of bacterial communities in heavy metal contaminated soils. Can. J. Microbiol. 42: 593-603.

Robinson, M. K., Bennett, P. M. and Richmond, M. H. (1977). Inhibition of Tn*A* translocation by Tn*A*. J. Bacteriol. 129: 407-414.

Rochelle, P.A., Wetherbee, M.K. and Olson, B.H. (1991). Distribution of DNA sequences encoding narrow spectrum and broad spectrum mercury resistance. Appl. Environ. Microbiol. 57: 1581-1589.

Rosado, A. S., Duarte, G. F., Seldin, L. and van Elsas, J. D. (1997). Molecular microbial ecology: a minireview. Revista de Microbiologia. 28: 135-147.

Sabry, S. A., Ghozlan, H. A and Abou-Zeid, D. -M. (1997). Metal tolerance and antibiotic resistance patterns of a bacterial population isolated from sea water. J. Appl. Microbiol. 83: 245-252.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular Cloning. A laboratory manual. 2nd edition. Cold Spring Harbour Laboratory Press.

Sayre, P. and Miller, R. V. (1991). Bacterial mobile genetic elements: Importance in assessing the environmental fate of genetically engineered sequences. Plasmid. 26: 151-171.

Schewchuk, L. M., Hellman, J. D., Ross, W., Park, S. J., Summers, A. O. and Walsh, C. T. (1989). Transcriptional switching by the MerR protein: activation and repression mutants implicate distinct DNA and mercury binding domains. Biochem. 28: 2340-2344.

Schmidt, F. R. J., Nucken, E. J. and Henscke, R. B. (1989). Structure and function of hot spots providing signals for site directed specific recombination and gene expression in Tn21 transposons. Mol. Microbiol. 3: 1545-1555.

Schmitt, R. and Klopfer-Kaul, I. (1984). Evolutionary relationship between Tn*21*-like elements and pBP201, a plasmid from *Klebsiella pneumoniae* mediating resistance to gentamycin and eight other drugs. Mol. Gen .Genet. 197: 109-119.

Schmitt, R., Bernhard, E. and Mattes, R. (1979). Characterisation of Tn*1721*, a new transposon containing tetracycline resistance gene capable of amplification. Mol. Gen. Genet. 172: 53-65.

Schweiger, F. and Tebbe, C. C. (1998). A new approach to utilise PCR single stranded conformation polymorphism for 16s rRNA gene based microbial community analysis. Appl. Environ. Microbiol. 64: 4870-4876.

Scott, J. R. and Churchward, G. G. (1995). Conjugative transposition. Ann. Rev. Microbiol. 49: 367-397.

Sedlmeier, R. and Altenbuchner, J. (1992). Cloning and DNA sequence analysis of the mercury resistance genes of *Streptomyces lividans*. Mol. Gen. Genet. 236: 76-85.

Shapiro, J. A. (1969). Mutations caused by the insertion of genetic material into the *gal* operon of *Escherichia coli*. J. Mol. Biol. 40: 93.

Shaw, J. H. and Clewell, D. B. (1985). Complete nucleotide sequence of macrolide-lincosamide-streptogramin B-resistance transposon Tn*917* in *Streptococcus faecalis*. J. Bacteriol. 164: 782-796.

Siemieniak, D. R., Slightom, J. L. and Chung, S. T. (1990). Nucleotide sequence of *Streptomyces fradiae* transposable element Tn*4556*: a class II transposon related to Tn*3*. Gene. 86: 1-9.

Silver, S. (1996). Bacterial resistances to toxic heavy metals- a review. Gene: 179: 9-19.

Silver, S. and Phung, Le. T. (1996). Bacterial heavy metal resistance: New surprises. Ann. Rev. Microbiol. 50: 753-789.

Smit, E., Leeflang, P and Wernars, K. (1997). Detection of shifts in microbial community structure and diversity in soil caused by copper contamination using amplified ribosomal DNA restriction analysis. FEMS Microbiol. Ecol. 23: 249-261.

Smit, E., Wolters, A and van Elsas, J. D. (1998). Self transmissible mercury resistance plasmids with gene-mobilising capacity in soil bacterial populations: influence of wheat roots and mercury addition. Appl. Environ. Microbiol. 64: 1210-1219.

Stanisich, V. A., Arwas, R., Bennett, P. M. and de la Cruz, F. (1989). Characterisation of a *Pseudomonas* mercury resistance transposon Tn*502*, which has a preferred insertion sit in RP1. J. Gen. Microbiol. 135: 2909-2915.

Stokes, H. W. and Hall, R. M. (1989). A novel family of potentially mobile DNA elements encoding site specific gene integration functions: integrons. Mol. Microbiol. 3: 1669-1683.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments seperated by gel electrophoresis. J. Mol. Biol. 98: 503-517.

Summers, A. O. (1986). Organisation, expression and evolution of genes for mercury resistance. Ann. Rev. Microbiol. 40: 607-634.

Sundin, G.W., Demezas, D. H. and Bender, C. L. (1994). Genetic and plasmid diversity within natural populations of *Pseudomonas syringiae* with various exposures to copper and streptomycin bactericides. Appl. Environ. Microbiol. 60: 4421-4431.

Sundström, L. (1998). The potential of integrons and connected programmed rearrangements for mediating horizontal gene transfer. APMIS. 106: 37-42.

Sundström, L., Roy, P. H. and Sköld, O. (1991). Site specific insertion of three structural gene cassettes in transposon Tn*7*. J. Bacteriol. 173: 3025-3028.

Sundström, L., Swedberg, G. and Sköld, O. (1993). Characterisation of transposon Tn*5086*, carrying the site specifically inserted gene *dhfrVII* mediating trimethoprim resistance. J. Bacteriol. 175: 1796-1805.

Tanaka, M., Yamamoto, T. and Sawai, T. (1983). Evolution of complex resistance transposons from an ancestral mercury transposon. J. Bacteriol. 153: 1432-1438.

Tateno, Y., Takezaki, N. and Nei, M. (1994). Relative efficiencies of the maximum likelihood, neighbor-joining, and maximum parsimony methods when substitution rate varies with site. Mol. Biol. and Evol. 11: 261-277.

Tillier, E.R.M. and Collins, R.A. (1995). Neighbor joining and maximum likelihood with RNA sequence: Addressing the interdependance of sites. Mol. Biol. and Evol. 12: 7-15.

Tolmasky, M. E. and Crosa, J. H. (1987). Tn*1331*, a novel multiresistance transposon encoding resistance to amikacin and ampicillin in *Klebsiella pneumoniae*. Antimicrob. Chemother. 31: 1955-1060.

Tolmasky, M. E. (1993). Multiresistance transposon Tn*1331*, an overview. An. Asoc. Quím. Argent. 81: 261-267.

Top, E., de Smet, I., Verstraete, W., Dijkmans, R. and Mergeay, M. (1994). Exogenous isolation of mobilising plasmids from polluted soils and sludges. Appl. Environ. Microbiol. 60: 831-839.

Top, E., De Rore, H., Collard, J. M., Gellens. V., Slobodkina, G., Verstraete, W. and Mergeay, M. (1995). Retromobilisation of heavy metal resistance genes in unpollutes and heavy metal polluted soil. FEMS Microbiol. Ecol. 18: 191-203.

Torsvik, V., Salte, K., Sorheim, R. and Goksoyr, J. (1990). Comparison of phenotypic diversity and DNA heterogeneity in a population of soil bacteria. Appl. Environ. Microbiol. 56: 776-781.

Torsvik, V., Daae, F. L., Sandaa, R. A. and Ovrese, L. (1998). Novel techniques for analysing microbial diversity in natural and perturbed environments. J. Biotechnol. 64: 53-62.

Trajanovska, S., Britz, M. L. and Bhave, M. (1997). Detection of heavy metal ion resistance genes in Gram-positive and Gran-negative bacteria isolated from a lead contaminated site. Biodegredation. 8: 113-124.

Trevors, J. T. (1985). Bacterial plasmid isolation and purification. J. Microbiol. Meth. 3: 259-271.

Trevors, J.T., Barkay, T. and Bourquin, A.W. (1987). Gene transfer among bacteria in soil and aquatic environments: a review. Can. J. Microbiol. 33: 191-198.

Tsuda, M. and Iino, T. (1987). Genetic analysis of a transposon carrying toluene degrading genes on a TOL plasmid pWWO. Mol. Gen. Genet. 210: 270-276.

Tsuda, M., Minegishi, K. -I. and Iino, T. (1989). Toluene transposons Tn*4651* and Tn*4653* are class II transposons. J. Bacteriol. 171: 1385-1393.

Tu, J., Wang, H. R., Chang, S. F., Charng, Y. C., Lurz, R., Dobrinski, B. and Wu, W. C. (1989). Transposable elements of *Xanthomonas campestris* pv. *citri* originating from indigenous plasmids. Mol. Gen. Genet. 217: 505-510.

Tunlid, A. and White, D. C. (1990). Lipid biomarkers in environmental samples. In: Analytical microbial methods, Fox, A., Morgan, S. L., Lennart, L. and Odham, G. (Eds). Plenium Press, New York. p259-274.

Turner, A. K. (1989). Analysis of the class II bacterial transposable elements Tn*2501* and Tn*3926*. Ph.D thesis, University of Bristol, UK.

Turner, A. K. and Grinsted, J. (1987). DNA sequence of the transposase gene of the new catagory of Class II transposon, Tn*2501*. Nucleic Acids Res. 15: 10049-10049.

Turner, A. K. and Grinsted, J. (1989). DNA sequence of the transposase gene of the Class II transposon, Tn*3926*. Nucleic Acids Res. 17: 1757-1757.

Ulrich, A. and Pühler, A. (1994). The new class II transposon Tn*163* is plasmid borne in two unrelated *Rhizobium leguminosarum* biovar *viciae* strains. Mol. Gen. Genet. 242: 505-516.

Vaituzis, Z., Nelson, J. D., Wan, L. W. and Colwell, R. R. (1975). Effects of mercuric chloride on growth and morphology of selected strains of mercury resistant bacteria. Appl. Microbiol. 29: 275-286.

Walsh, C. T., Distefano, M. D., Moore, M. J., Shewchuk, L. M. and Verdine, G. L. (1988). Molecular basis of bacterial resistance to organomercurial and inorganic mercuric salts. FASEB. 2: 124-130.

Wang, G. C. and Wang, Y. (1997). Frequency of formation of chimeric molecules as a consequence of PCR coamplification of 16s rRNA genes from mixed bacterial genomes. Appl. Environ. Microbiol. 63: 4645-4650.

Wang, Y., Moore, M., Levison, H. S., Silver, S., Walsh, C. and Mahler, I. (1989). Nucleotide sequence of a chromosomal mercury resistance determinant from a *Bacillus* sp. with broad spectrum mercury resistance. J. Bacteriol. 171: 83-92.

Ward, E. and Grinsted, J. (1987). The nucleotide sequence of the *tnpA* gene of Tn*21*. Nucleic Acids Res. 15: 1799-1799.

Wells, R. G. and Grindley, N. D. F. (1984). Analysis of the γδ *res* site sites required for site specific recombination and gene expression. J. Mol. Biol. 179: 667-687.

Wheatcroft, R. and Williams, P. A. (1981).Rapid methods for the study of both stable and unstable plasmids in *Pseudomonas*. J. Gen. Microbiol. 124: 433-437.

Wilmotte, A., Hennen, A., van der Lelie, D. and Mergeay, M. (1996). The broad host range plasmid pMOL96 contains a putative transposon similar to Tn*402* (Tn*5090*) and Tn*5053*. Archives of Physiology and Biochemistry. 104: B55.

Wilson, I. G. (1997). Inhibition and facilitation of nucleic acid amplification. Appl. Environ. Microbiol. 63: 3741-3751.

Woese, C. R. (1996). Phylogenetic trees: Wither microbiology. Curr. Biol. 6: 1060-1063.

Wood, J. M. and Wang, H-K. (1983). Microbial resistance to heavy metals. Environ. Sci. Technol. 17: A583.

Yamamoto, T., Tanaka, M., Baba, R. and Yamagishi, S. (1981). Physical and functional mapping of Tn*2603*, a transposon encoding ampicillin, streptomycin, sulphonamide and mercury resistance. Mol. Gen. Genet. 181: 464-469.

Yamamoto, T. (1998). Organisation of complex transposon Tn*2610* carrying two copies of *tnpA* and *tnpR*. Antimicrob. Chemother. 33: 746-750.

Yeates, C., Gillings, M. R., Davison, A. D. Altavilla, N and Veal, D. A. (1997). PCR amplification of crude microbial DNA extracted from soil. Letts. Appl. Microbiol. 25: 303-307.

Yeo, C.C., Tham, J. M., Kwong, S.M., Yiin, S. and Poh, C. L. (1998). Tn*5563*, a transposon encoding putative mercuric ion transport proteins located on plasmid pRA2 of *Pseudomonas alcaligenes*. FEMS Microbiol. Letts. 165: 253-260.

Yin, X. and Stotzky, G. (1997). Gene transfer amongst bacteria in natural environments. Adv. Appl. Microbiol. 45: 153-212

Young, H-K and Rosser, S. J. Personal communication.

Yurieva, O., Kholodii, G., Minakhin, L., Gorlenko, Z., Kalyaeva, E., Mindlin, S. and Nikiforov, V. (1997). Intercontinental spread of promiscuous mercury resistance transposons in environmental bacteria. Mol. Microbiol. 24: 321-329.

Zinder, N. D. and Lederberg, J. (1952). Genetic exchange in *Salmonella*. J. Bacteriol. 64: 679-699.

Zuhlsdorf, M. T. and Weidermann, B. (1992). Tn*21*-specific structures in Gram negative bacteria from clinical isolates. Antimicrob. Chemother. 36: 1915-1921.

ELSEVIER

FEMS Microbiology Letters 144 (1996) 95–102

FEMS MICROBIOLOGY LETTERS

Phylogenetic analysis of *tnp*R genes in mercury resistant soil bacteria: the relationship between DNA sequencing and RFLP typing approaches

Robert J. Holt *, Peter Strike, Kenneth D. Bruce

Department of Genetics and Microbiology, Donnan Laboratories, University of Liverpool, Crown St, Liverpool, L69 3BX, UK

Received 3 July 1996; revised 16 August 1996; accepted 21 August 1996

Abstract

The diversity of resolvase (*tnp*R) genes carried by a number of mercury resistant soil bacteria has been investigated by DNA sequencing. The resulting DNA sequence information was compared to previously published *tnp*R DNA sequences and to previously published restriction fragment length polymorphism (RFLP) data, permitting the relationships between DNA sequencing and RFLP approaches to be studied by the use of phylogenetic trees. DNA maximum likelihood and DNA parsimony were used to construct a variety of phylogenetic trees. DNA sequencing confirmed the validity of RFLP analysis and highlighted the importance of restriction endonuclease choice upon the resulting RFLP patterns and dendrogram topology. The *tnp*R genes of two previously uncharacterised mercury resistant bacteria, T2-7 and T2-12 were also studied. DNA sequence data placed T2-7 in a previously described gene class, *tnp*R-D and T2-12 in a new gene class, *tnp*R-F. The significance of this data with respect to the recombination and evolution events occurring within bacterial populations are discussed.

Keywords: DNA sequencing; Phylogeny; Restriction fragment length polymorphism; Soil bacteria; Resolvase

1. Introduction

Mercuric compounds in natural environments can arise from a variety of natural and anthropogenic sources such as mining operations and other industrial sites although the precise chemical speciation and concentration vary between different environments. The presence of mercury is often associated with an increase in the number of cultivable mercury resistant bacteria isolated from the environment [1].

Bacterial mercury resistance, typically found on the *mer* (mercury resistance) operon, is frequently identified on Class II (Tn*3*-like) transposons [1–5]. Given the role of mechanisms such as conjugation and transformation in the spread of DNA within bacterial populations [6,7], the potential for transfer of these genes is increased if the transposon itself becomes incorporated into a conjugative plasmid. A wide range of transposons are known to carry *mer* genes [2–5] with the most widely studied ones being Tn*501* and Tn*21* [2–4]. All functional Class II transposons, including those containing mercury resistance genes, carry two genes involved in the transpo-

* Corresponding author. Tel.: +44 (151) 794 3621; Fax: +44 (151) 794 3655; E-mail: R.J.Holt@liverpool.ac.uk

PII S0378-1097(96)00345-X

sition process. The transposase gene, *tnp*A, is involved in the formation of the donor/recipient cointegrate [8]. The resolvase gene, *tnp*R, codes for a protein involved in the resolution of the two parts of the cointegrate structure [9]. The function of *tnp*A and *tnp*R genes is such that often the corresponding gene from one transposon can complement its homologue in an alternative transposon [8,9].

Previous studies have examined the diversity of specific genes [10–16]. Information obtained from such studies is important in order to characterise genetic components within the bacterial population of a particular environment and is important in assessing the evolution of bacterial genomes and the mutation and recombination events that lead to diversity within such populations. The diversity of transposon genes, *tnp*A and *tnp*R and mercury resistance genes, *mer*RTΔP contained within bacteria cultured from both mercury polluted and pristine sites has been examined by RFLP analysis [1,11,12] and in the case of the mercury resistance genes, a limited number have been examined by DNA sequencing [10]. The *tnp*R genes from cultivated mercury resistant Gram-negative soil bacteria have been characterised by RFLP analysis [12] and five classes of *tnp*R gene have been identified and designated *tnp*R class A–E based on their RFLP profiles. In the present study, representatives from each of these *tnp*R classes are sequenced and sequence data is compared to the RFLP data previously described. This analysis provides a useful insight into the relative merits of DNA sequencing and RFLP typing approaches. Sequence variation can now be studied both within and between classes and assisted by the construction of phylogenetic trees [17], allowing the similarities between different DNA sequences to be visualised. Two methods of dendrogram construction have been used in this study, both of which produce unrooted trees; DNA maximum likelihood [18], which compares the percentage similarity between sequences and DNA parsimony method [19] which calculates the minimum number of base changes needed on a given tree. The effects of DNA sequence changes on the amino acid sequence of *tnp*R are studied to ascertain whether conserved DNA and amino acid sequence changes have occurred and two previously uncharacterised *tnp*R genes are sequenced and allocated to RFLP classes.

2. Materials and methods

2.1. Bacterial strains

All mercury resistant strains used in this study have been previously described [1,10–12]. SO strains were isolated from mercury polluted soil at Fiddlers Ferry, Merseyside. SE strains, also from Fiddlers Ferry, were isolated from sediment samples. SB strains were isolated from a pristine soil site at Salter Brook Bridge. T2 strains were isolated from soil taken at a disused copper mine containing high levels of mercury in Tipperary, Eire [1]. The *tnp*R, *tnp*A and *mer*RTΔP classes of these isolates have been previously described [1,10–12].

2.2. Preparation of template DNA for PCR

DNA template for PCR was isolated using a variety of methods. These were the plasmid extraction method described by Olsen [20], a method combining the lysis step of Kado and Liu [21] and a neutralization and purification step modified from Birnboim and Doly [22]. Total genomic DNA was isolated according to Pitcher et al. [23] and the boiling preparation method described in Osborn et al. [1] was also used.

2.3. PCR reactions

PCR reactions were made up as follows; 10 μl of 10×PCR buffer and 1.5 mM $MgCl_2$ (Gibco BRL) were added to the reaction mixture along with 20 pmol of each primer; 501R1/C and 501R2/C. 1 mM (final concentration) of dNTPs (Pharmacia) along with 2.5 U of Taq DNA polymerase (Gibco BRL) were used with the reaction volume brought up to 100 μl using sterile distilled water and overlaid with mineral oil (Sigma). Cycling reactions were carried out in a Perkin Elmer 480 Thermal Cycler. The typical conditions used were 95°C for 4 min, followed by 28 cycles of 95°C for 1 min, 62 or 63°C for 1 min and 72°C for 2 min followed by 10 min at 72°C. Depending on the exact nature of the template strain, slightly altered annealing temperatures were employed.

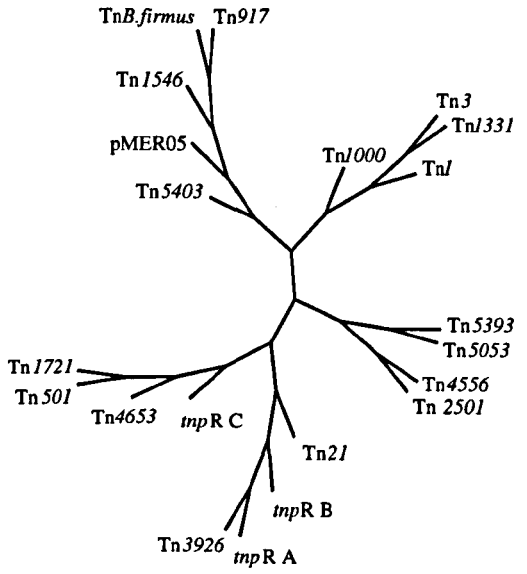


Fig. 1. DNA parsimony tree showing *tnp*R genes from classes *tnp*R-A–E and common transposons. *tnp*R-A–C correspond to RFLP classes previously described [12]. *tnp*R-D: Tn*501*, *tnp*R-E; Tn*21*. Tn1–Tn5403, as described in Section 2.7.

2.4. Primers

The two primers used in this study have been previously described [12] and were designed to correspond to the published DNA sequences of Tn*21* and Tn*501* *tnp*R genes. They are positioned at either end of this 560 bp gene and produce a PCR product of approx. 500 bp. These primers were initially used in the RFLP study [12] and were used here in order to allow the same fragment to be analysed:

501R1/C:
5'-GTT CAG CA[GC] CTT CGA CCA G -3'

501R2/C:
5'-TA[CG] AGG GTT TC[GC] CG[AG] CTG AT -3'

2.5. DNA sequencing

PCR products were purified for DNA sequencing employing Pharmacia Biotech MicroSpin S-400 HR columns used in accordance with the manufacturer's instructions. DNA sequencing was carried out on an ABI 373A automated sequencer.

2.6. Sequence analysis and construction of phylogenetic trees

Alignment of sequences and conversion into suitable formats was carried out using PILEUP, LINEUP, TOPIR and READSEQ programs in the PHYLIP 3.5C facility of the SEQNET database, Daresbury Laboratories which was also used for all subsequent analysis [17,24,25]. DNAML and DRAWTREE were used in the construction of maximum likelihood trees and DNA parsimony trees utilised DNAPARS and DRAWTREE. All trees were bootstrapped using SEQBOOT and CONSENSE.

2.7. DNA sequences

The previously published DNA sequences of the Tn*501* and Tn*21* resolvase genes have been assigned Genbank accession numbers K01725 and X01298, respectively. DNA sequences of class *tnp*R-A,B,C,D (T2-7, this study) and *tnp*R-F (T2-12, this study) have been assigned accession numbers U61277–U61281.

The DNA sequences of the *tnp*R genes used in Fig. 1 have been assigned accession numbers as follows: Tn*B. firmus*, M90749; Tn*917*, M11180;

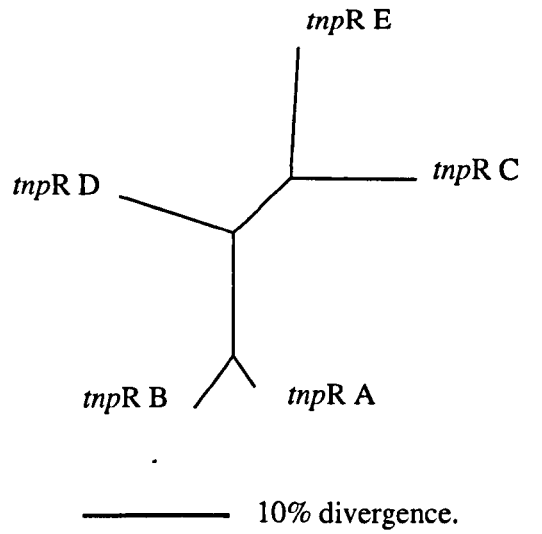


Fig. 2. Maximum likelihood tree of 372 bp resolvase analysis region. *tnp*R classes as described in Fig. 1.

Tn*1546*, M97297; pMER05 (resolvase related to that contained within Tn*552*), L20694; Tn*5403*, X75779; Tn*1000*, X60200; Tn*1*, L10085; Tn*1331*, M55547; Tn*3*, V00613; Tn*5393*, M96392; Tn*5053*, L40585; Tn*4556*, M29297; Tn*2501*, M15197; Tn*3926*, X78059; Tn*4653*, D90148; Tn*1721*, X02590.

3. Results

3.1. Sequencing of PCR products

A representative from each previously described *tnp*R class [12] together with one additional class A and two additional class C strains were selected for DNA sequencing. More than one class A and C were chosen in order to ascertain the sequence similarity within RFLP classes. Class *tnp*R-A was represented by SO1 and SO2, *tnp*R-B by SE31 and *tnp*R-C by SB3, SB29 and SE23. Two previously uncharacterised T2 strains, T2-7 and T2-12 were also studied. The PCR products of approx. 500 bp were sequenced directly and a partial DNA sequence from a representative of each *tnp*R class was obtained. These sequences are shown in Fig. 6, with the complete *tnp*R sequences from Tn*21* and Tn*501*.

3.2. Sequence similarity within RFLP classes

The similarity between the 3 class C *tnp*R sequences, from SB3, SB29 and SE23 was determined and found to be 100% between SB3 and SB29 over the 372 bp analysis region and 98.4% between these two and the 310 bp of sequence obtained from SE23. SO1 and SO2 class A *tnp*R sequences were found to be identical to each other over a length of 383 bp. For the purposes of this study, SO1 and SB3 sequences were chosen as group representatives of classes A and C respectively.

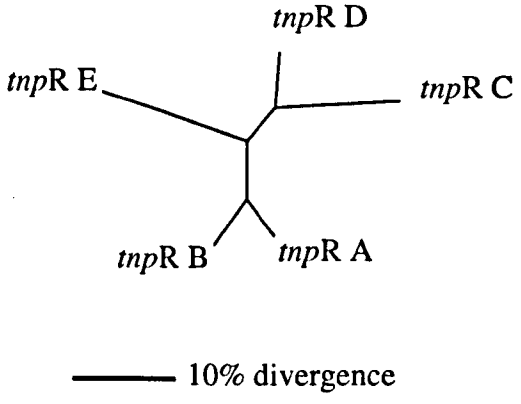


Fig. 3. Maximum likelihood tree of the first 186 bp of resolvase sequence. *tnp*R classes as described in Fig. 1.

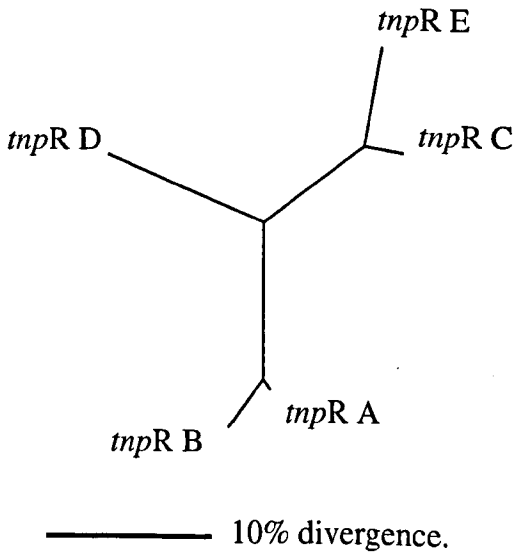


Fig. 4. Maximum likelihood tree of the second 186 bp of resolvase sequence. *tnp*R classes as described in Fig. 1.

3.3. Phylogenetic analysis

Initially, an unrooted parsimony tree was constructed showing the relationships between the *tnp*R genes sequenced in this study and those for which corresponding sequence data is available. The sequence data from classes *tnp*R-A, B and C were aligned with the published sequences of 18 different *tnp*R genes and this data used to construct the dendrogram shown in Fig. 1.

The sequence data obtained from each of the PCR products and the published sequences of Tn*501* and Tn*21* were aligned resulting in a 372 bp analysis region from *tnp*R-A–E. An unrooted tree was constructed from the 372 bp analysis region using the maximum likelihood method (Fig. 2) and the DNA parsimony method (data not shown). Both of these were consistent with the dendrogram derived by Pearson et al. [12].

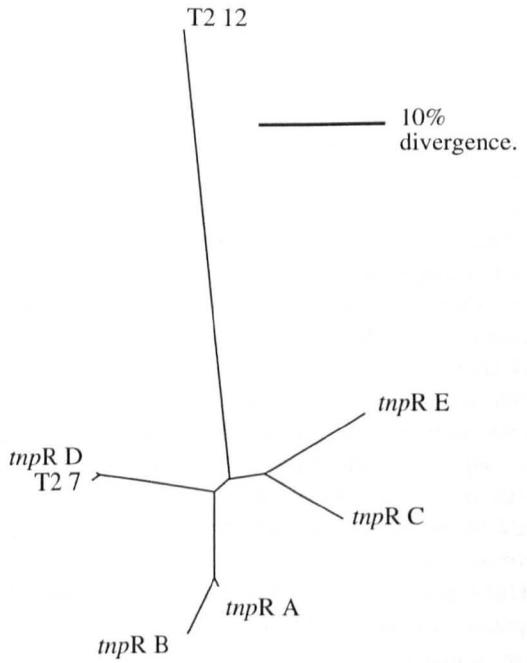


Fig. 5. Maximum likelihood tree showing *tnp*R-A–E and T2-7 and T2-12. *tnp*R classes as described in Fig. 1.

Two further unrooted trees derived from the first and second 186 bp halves of the 372 bp sequence were constructed. (Figs. 3 and 4, respectively). The dendrogram corresponding to the first half of the initial fragment differed from that of the overall analysis region, in that the positions of *tnp*R-D and *tnp*R-E are reversed whereas the dendrogram representing the second half of the fragment is similar to that of the overall analysis region.

3.4. Restriction endonuclease patterns of sequenced analysis regions

A restriction map of the original *tnp*R-A–E sequences was constructed from sequence data using the restriction endonucleases used in the original RFLP study. The sizes of predicted bands generated matched those observed in Pearson et al. [12]. The majority of the sites, 16/20, were located in the second half of the 372 bp analysis region (data not shown).

```
tnpRD  GTGCAGGGGC ACCGCATCGG CTACGTCCGG GTCAGCAGCT TCGACCAGAA
T2 7   .......... .......... .......... .......... ..........
tnpRA  .......... .......... .......... .......... ..........
tnpRB  .......... .......... .......... .......... ..........
tnpRC  .......... .......... .......... .......... ..........
tnpRE  ATGACTGGAC AGCGCATTGG GTATATCAGG GTCAGCACCT TCGACCAGAA
tnpRF  .......... .......... .......... .......... ..........

tnpRD  CCCGGAACGC CAGCTGGAAC AGACACAGGT GAGCAAGGTG TTCACCGACA
T2 7   .......... .......... .......... .......... ..........
tnpRA  .......... .......... .......... .........T TCACCGNACA
tnpRB  .......... .......... ...TCGACCT CGGTAAGGTA TTCACCGATA
tnpRC  .......... .......... .......... GGATTAAGTG TTCACCGACA
tnpRE  CCCGGAACGG CAACTGGAAG GCGTCAAGGT TGATCGCGCT TTTAGCGACA
tnpRF  .......... .......... .......... .......... ..........

tnpRD  AGGCATCGGG CAAGGACACC CAGCGCCCCC AGCTCGAAGC GC...TGCTG
T2 7   .......GGG CAAGGACACC CAGCGCCCCC AGGTTGGGAA GCGCTTGCTG
tnpRA  AGGCTTCTGG CAAGGACACC CAGCGTCCCG .AGCTGGAAA G..GCTGCTG
tnpRB  AGGCTTCCGG CAAGGACACG CAACGTCCCG AAACTTGAAA G..GCTGCTG
tnpRC  A.GCGTCGGG CAAGGACACA CGGCGGCCCG .AACTGGAAC G..GCTGCTC
tnpRE  AGGCATCCGG CAAGGATGTC AAGCGTCCGC .AACTGGAAG C..GCTGATA
tnpRF  .......... ........CC CAGCGGCCCG AGTCTTGATT C..GATGACT

tnpRD  AGCTTCGTCC GCGAAGGCGA TACAGTGGTG G.TGCACAGC ATGGACCGGC
T2 7   AGCTTCGTCC GCGAAGGCGA TACAGTGGTG G.TGCACAGC ATGGACCGGC
tnpRA  GCCTTCGTCC GCGAGGGCGA CACCGTGGTG G.TGCATAGC ATGGACAGGC
tnpRB  GCCTTTGTGC GCGAGGGCGA CACCGTGGTG G.TGCACAGC ATGGACAGGT
tnpRC  GCCTTCGTGC GCGAAGGCGA CACGGTCGTG G.TGCACAGC ATGGATCGTC
tnpRE  AGCTTCGCCC GCACCGGCGA CACCGTGGTG G.TGCATAGC ATGGATCGCC
tnpRF  GGCTTCGTAC GTCGAAGGGA CACCGTGGTG GTTTCATAGC ATGGATCAGC

tnpRD  TGGCCCGC.. .AACCTCGAT GACCTGCG.T CGCTTGGTAC AGAAGCTGAC
T2 7   TGGCCCGC.. .AACCTCGAT GACCTGCGTT CGCTTGGTAC AGAAGCTGAC
tnpRA  TGGCACGC.. .AACCTTGAT GACCTGCG.C CGCATCGTCC AAGGGCTGAC
tnpRB  TGGCACGC.. .AATCTCGAT GACCTGCG.C CGCATCGTTC AGGAGCTGAC
tnpRC  TGGCGCGC.. .AACCTCGAC GACCTGCG.C CGCCTGGTGC AGGGCCTCAC
tnpRE  TGGCGCGC.. .AATCTCGAT GATTTGCG.C CGGATCGTGC AAACGCTGAC
tnpRF  CTGGCGCCGC AACCTTCGAT GACTTGCGTC CGCCTCGTGC AAAAGCTCAC

tnpRD  TCAACGCGGC GT..GCGCAT CGAG.TTCCT GAAGGAGGGC CTGGTG.TTC
T2 7   TCAACGCGGC GT..GCGCAT CGAGTTTCCT GAAGGAGGGC CTGGTGTTTC
tnpRA  ACAACGGGGC GT..GCGCAT GGAGTTCGTC AAAGAAGGGC ..TGAAGTTC
tnpRB  ACAACGGGGC GT..GCGGAT GGAGTTTGTC AAAGAAGGGC ..TGGCGTTC
tnpRC  CCAGCGCGGC GT..ACGCAT CGAGTTCCTT AAGGAGCATT ..TGACCTTC
tnpRE  ACAACGCGGC GT..GCATAT CGAATTCGTC AAGGAACACC ..TCAGTTTT
tnpRF  CAAGCGCGGT GTTGGGTTAT CGAGTTCGTC AAGGAAAGCC TGTACCTTCA

tnpRD  ACTGGCG.AG GACTCGCCGA TGG...CCAA CCTGATGCTG TC.GGTGATG
T2 7   ACTGGCGAAG GACTCGCCGA TGG...CCAA CCTGATGCTG TC.GGTGATG
tnpRA  ACCGGCG.AG GACTCACCGA TGG...CCAA TCTGATGCTG TC.GGTCATG
tnpRB  ACCGGCG.AT GACTCACCGA TGG...CCAA TTTGATGCTG TC.GGTCATG
tnpRC  ACCGGCG.AG GACTCGCCGA TGG...CGAA CCTGATGCTG TC.GGTAATG
tnpRE  ACTGGCG.AA GACTCTCCGA TGG...CGAA CCTGATGCTC TC.GGTGATG
tnpRF  CCGGGCGAGG AATTCACCCG ATGGACGCAA CCTAATGCTG TCTGTTCATG

tnpRD  GGGGCCTTCG CTGAG.TTCG AGCGCGCCCT G.ATCCGCGA GCGGCAGCGT
T2 7   GGGGCCTTCG CTGAGTTTCG AACGCGCCCT G.ATCCGCGA GCGGCAGCGT
tnpRA  GGAGCCTTCG CTGAG.TTCG AGCGCGCCCT G.ATCCGCGA ACGTCAGCGC
tnpRB  GGGGCTTTTG CGGAG.TTCG AGCGCGCACT GTATCCGCGA ACGCCAGCGC
tnpRC  GGCGCGTTCG CCGAG.TTCG AACGCGCCTT G.ATCCGCGA GCGGCAGCGC
tnpRE  GGCGCGTTCG CCGAG.TTCG AGCGCGCCCT G.ATCCGCGA GCGTCAGCGC
tnpRF  GGGGCGTTTC GCCCAATTCT AGCGGGCCTT GTATCCCCGA CGGCAGAGGG

tnpRD  GAGGGCATCA CCTTGGCCAA GCAGCGTGGC GCGTAC.... ...CGGGGCC
T2 7   GAGGGCATCA CCTTGGCCAA GCAGCGTGGC GCGTAC.... ...CGGGGCC
tnpRA  GAGGGAATCG TGCTGGCCAA GCAGCGCGGT GCCTAC.... ...CGGGGAC
tnpRB  GAGGGAATCG TGCTGGCCAA GCAGCGCGGT GCCTAC.... ...CGGGGAC
tnpRC  GAGGGCATCG CGCTCGCCAA GCAGCGCGGG GCCTAC.... ...CGTGGCA
tnpRE  GAGGGTATTG CGCTCGCCAA GCAACGCGGG GCTTAC.... ...CGTGGCA
tnpRF  TAAGGCATCG CGTTCGTCCA AAACAGCGCC GGTAACCCTT ATTCGGGGCC

tnpRD  GCAAGAAAGC CCTGTCCGAT GAGCAGGCTG CTACCCTGCG GCAGCGAGCG
T2 7   GCAAGAAATC CCTGTCC... .......... .......... ..........
tnpRA  GAAAGAAATC GCTGAACAGC GAACAAATTG CCGAGTTGAA ACGGCGAGTT
tnpRB  GAAAGAAATC GCTGAACAGC GAACAAATTG CCAAGTTGAA ACAGCGAGTC
tnpRC  GAAAGAAATC CCTGTCGTCT GAACGTATTG CCGAACTGCG CCAACGTTTT
tnpRE  GGAAGAAATC CCTGTCGTCT GAGCGTATTG CCGAACTGCG CCAACGTGTC
tnpRF  GCAAAGAAGA AAGATG.... .......... .......... ..........

tnpRD  ACGGCCGGCG AGCCCAAGGC GCAGCTTGCC CGCGAGTTCA ACATCAGCCG
T2 7   .......... .......... .......... .......... ..........
tnpRA  GCGGCGAACC AAAAAAC... .......... .......... ..........
tnpRB  GCGGCAGGCG ATCAAAAAAC .......... .......... ..........
tnpRC  .......... .......... .......... .......... ..........
tnpRE  GAGGCTGGCG AGCAAAAGAC CAAGCTTGCT CGTGAATTCG GAATCAGTCG
tnpRF  .......... .......... .......... .......... ..........

tnpRD  GGAAACCCTC TACCAGTACC TCCGCACGGA CGACTGA
T2 7   .......... .......... .......... .......
tnpRA  .......... .......... .......... .......
tnpRB  .......... .......... .......... .......
tnpRC  .......... .......... .......... .......
tnpRE  CGAAACCCTG TATCAATACT TGAGAACGGA TCAGTAA
tnpRF  .......... .......... .......... .......
```

Fig. 6. Complete DNA sequences of *tnp*R genes contained within Tn*21* and Tn*501* and partial sequences of class *tnp*R-A–D and F.

3.5. Insertion of T2 isolates into unrooted tree

Two previously uncharacterised strains, T2-7 and T2-12 [10], were chosen for further study. Again approx. 500 bp PCR products from *tnp*R were generated and sequenced. T2-7 DNA sequence was found to be 97.6% similar to Tn*501* and gave an identical restriction endonuclease digestion pattern to Tn*501* placing it in class *tnp*R-D. T2-12 was assigned a new resolvase class, *tnp*R-F by virtue of its DNA sequence. This sequence information was used to construct a maximum likelihood tree which is shown in Fig. 5.

3.6. DNA sequence alterations between classes

The complete DNA sequences of each of the 7 classes were aligned to observe the nature of the DNA changes leading to the differences between classes. The complete DNA sequences obtained in this study are shown in Fig. 6. These DNA sequences were used to obtain amino acid sequences from all *tnp*R classes (data not shown).

4. Discussion

In this study, DNA sequencing was used to examine *tnp*R gene diversity within a collection of mercury resistant soil bacteria from which the *tnp*R and *mer* regions had been previously characterised by RFLP analysis [12]. The virtually complete congruence which was observed between DNA sequence and RFLP data validates the use of RFLP for typing studies. The relative merits of both typing procedures and the factors affecting RFLP studies have also been examined.

Analysis of sequence data was carried out by construction of phylogenetic trees using two methods; DNA maximum likelihood and DNA parsimony analysis. The relative advantages and disadvantages of each method have been discussed at length [18,19,25–30] and the two methods, used to construct unrooted trees, produced comparable results.

The DNA sequences of eight *tnp*R PCR products from various strains along with the published sequences of Tn*501* and Tn*21* were used to construct phylogenetic trees. Similarity within classes was seen to be high; all members of class *tnp*R-A were observed to be identical to each other although this may be due to the possible clonal nature of these SO isolates [12]. Class *tnp*R-C sequences were highly conserved, dropping to a lower level of 98.4% similarity between SE23 and SB3 and 29 sequences. SE3 and SO1 were chosen as representatives of their respective classes and all dendrograms were constructed using these sequences.

A DNA parsimony tree showing the relationship between the *tnp*R sequences obtained in this study and those contained within transposons for which DNA sequence is available was constructed. This shows four major groupings of resolvase genes. The sequences obtained in this study fall into one of these groups, which also contains Tn*501* and Tn*21* from which the primers used in this study were designed.

Unrooted trees constructed to compare the DNA sequences of the previously described RFLP groups show similarity between DNA sequence and RFLP data [12]. It was found, however, that the topology of the maximum likelihood tree for the complete 372 bp analysis region was similar to that of the second 186 bp of sequence, whereas in the tree representing the first 186 bp of sequence the positions of *tnp*R-D and *tnp*R-E differ being reversed. The restriction maps of the analysis region, (data not shown) shows the position of the restriction endonuclease sites observed in the RFLP analysis where 16 of the 20 restriction endonuclease sites are contained within the second half of the region. As the second half of the sequence contains a larger number of the restriction sites originally used to classify these genes than the first half of the sequence and these sites vary, the second half of the sequence has provided a more accurate representation of the overall analysis region. However, the dendrogram of the first 186 bp still shows marked structural similarity to that of the overall analysis region with the exception of the class D and E branches. This emphasises the importance of restriction enzyme choice when carrying out RFLP analysis as the position and frequency of restriction sites will affect the RFLP patterns which are observed. DNA parsimony trees of the two halves of sequence were constructed (data not shown). Tree topology was identical to those produced using the maximum likelihood method. The DNA sequence

similarity within each RFLP class was found to be high, ranging from 97.6 to 100%. This supported the previously published RFLP data.

The characterisation of *tnp*R genes contained within T2-7 and T2-12 indicated that T2-7 should be placed into class *tnp*R-D by virtue of its 97.6% similarity to Tn*501* and its identical restriction endonuclease pattern and that T2-12 should be allocated a new *tnp*R gene class, *tnp*R-F on the basis of the data presented here (Fig. 5). For DNA sequencing to be used as a reliable alternative to RFLP for the type classification of genes, a similarity value over which any two sequences are considered to be in the same class must be established. For this study, the data presented here showed variation within classes to be as little as 2.4% and the variation within the gene to be as high as 20% (data not shown). More studies are required in order to determine the general applicability of such values upon different sequences, with further analysis needed in each individual case [10]. Sequencing of variants within a given gene population provides a large amount of information over that which can be acquired by RFLP data including, e.g. allowing the determination of the protein sequence of the fragment. The DNA sequences in this study showed a number of base changes in the 3rd codon position with conserved 1st and 2nd codon positions based on the presumed reading frame (Fig. 6). The partial protein sequence of *tnp*R-A–E (data not shown) indicates that a large number of these 3rd base shifts were not translated into amino acid changes. Amongst amino acids which had been altered due to mutation, a significant number are such that an amino acid of comparable physicochemical structure had replaced the original residue. The amino acid sequences obtained from T2-7 and T2-12 suggested that the *tnp*R genes contained within these isolates may be non-functional. It has been suggested that DNA rearrangements may be responsible for the majority of polymorphisms that arise [31], however, this not the case in this study as could be seen from the DNA sequence obtained which indicated that point mutations were responsible for the observed differences between RFLP classes. Evidence for recombinational events was not seen within these sequences; this may be due to the relatively small length of the *tnp*R gene enforcing physical constraints on the recombination system.

The percentage of bacteria contained within the soil which are cultivable under laboratory conditions may be as low as 0.3% [11]. The ability to extract DNA directly from soil [11,32] will therefore permit the analysis of diversity within genes in their bacterial communities whether from cultivable or non-cultivable organisms. The study of diversity amongst genes from the natural environment may lead to an understanding of the recombination and exchange of DNA responsible for the evolution of the gene in that particular environment. This and other diversity studies may ultimately lead to the characterisation of individual recombination and mutational events within bacteria and provide an explanation for the linkage between genes such as *tnp*A, *tnp*R and *mer*RTΔP and the events leading to their evolution.

Acknowledgments

This work was funded by NERC. R.J.H. is supported by a NERC Ph.D. studentship and K.D.B. by an NERC fellowship. We would like to thank Angela Bardon for DNA sequencing and Prof. D.A. Ritchie and Andrea Pearson for continued support. This work was carried out with the use of the SEQNET facility, Daresbury.

References

[1] Osborn, A.M., Bruce, K.D., Strike, P. and Ritchie, D.A. (1993) Polymerase chain reaction – Restriction fragment length polymorphism analysis shows divergence among *mer* determinants from Gram-negative bacteria indistinguishable by DNA-DNA hybridisation. Appl. Environ. Microbiol. 59, 4024–4030.

[2] Bennett, P.M., Grinstedt, J., Choi, C.L and Richmond, M.H. (1978) Characterisation of Tn*501*, a transposon determining resistance to mercuric ions. Mol. Gen. Genet. 159, 101–106.

[3] Brown, N.L., Winnie, J.N., Fritzinger, D and Pridmore, R.D. (1985) The nucleotide sequence of the *tnp*A gene completes the sequence of the *Pseudomonas* transposon Tn*501*. Nucleic Acids Res. 13, 5657–5669.

[4] De la Cruz, F. and Grinsted, J.(1982) Genetic and molecular characterisation of Tn*21* , a multiple resistance transposon from R100.1. J. Bacteriol. 151, 222–228.

[5] Kholodii, G.Ya., Yurieva, O.V., Lomovskaya, O.L., Gorlenko, Zh.M., Mindlin, S.Z. and Nikiforov, V.G. (1993) Tn*5053*, a mercury resistance transposon with integrons ends. J. Mol. Biol. 230, 1103–1107.

[6] Ippen-Ihler, K. Bacterial conjugation. In: Gene Transfer in the Environment (Levy, S.B. and Miller, R.V., Eds.), pp. 33–72, McGraw Hill, New York.
[7] Trevors, J.T., Barkay, T. and Bourquin, A.W. (1987) Gene transfer among bacteria in soil and aquatic environments – a review. Can. J. Microbiol. 33, 191–198.
[8] Grinstead, J., De la Cruz, F., Altenbuchner J. and Schmitt, R. (1982) Complementation of transposition of *tnp*A mutants of Tn*3*, Tn*21*, Tn *501* and Tn*1721*. Plasmid 8, 276–286.
[9] Diver, W.P., Grinstead, J., Fritzinger, D.C., Brown, N.L., Altenbuchner, J., Rogowsky, P. and Schmitt, R. (1983) DNA sequences of and complementation by the *tnp*R genes of Tn*21*, Tn*501* and Tn*1721*. Mol. Gen. Genet. 191, 189–193.
[10] Osborn, A.M., Bruce, K.D., Strike, P. and Ritchie, D.A. (1995) Sequence conservation between regulatory mercury resistance genes in bacteria from mercury polluted and pristine environments. Syst. Appl. Microbiol. 18, 1–6.
[11] Bruce, K.D., Osborn, A.M., Pearson, A.J., Strike, P and Ritchie, D.A. (1995) Genetic diversity within *mer* genes directly amplified from communities of non-cultivated soil bacteria. Mol. Ecol. 4, 605–612.
[12] Pearson, A.J., Bruce, K.D., Osborn, A.M., Ritchie, D.A. and Strike, P. (1996) The distribution and association of class II transposase and resolvase genes with *mer* genes in soil bacteria. In press.
[13] Barkay, T., Fouts, D.L. and Olson, B.H. (1985) Preparation of a DNA gene probe for detection of mercury resistance genes in Gram-negative bacterial communities. Appl. Environ. Microbiol. 49, 686–692.
[14] Olson, B.H., Lester, J.N., Cayless, S.M. and Ford, S. (1989) Distribution of mercury resistance determinants in bacterial communities of river sediments. Water. Res. 23, 1209–1217.
[15] Olson, B.H., Cayless, S.M., Ford, S. and Lester, J.N. (1991) Toxic element contamination and the occurrence of mercury resistant bacteria in Hg-contaminated soil, sediments and sludges. Arch. Environ. Contam. Toxicol. 20, 226–233.
[16] Rochelle, P.A., Wetherbee, M.K. and Olson, B.H. (1991) Distribution of DNA sequences encoding narrow spectrum and broad spectrum mercury resistance. Appl. Environ. Microbiol. 57, 1581–1589.
[17] Felsenstein, J. (1993) PHYLIP (Phylogeny Inference Package) version 3.5C, Department of Genetics, University of Washington, Seattle.
[18] Felsenstein, J. (1981) Evolutionary trees from DNA sequence: A maximum likelihood approach. J. Mol. Evol. 17, 368–376.
[19] Fitch, W.M. (1971) Toward defining the course of evolution: minimum change needed for a specific tree topology. Syst. Zool. 20, 406–416.
[20] Olsen, J.E. (1990) An improved method for rapid isolation of plasmid DNA from wild type Gram-negative bacteria for plasmid restriction profile analysis. Lett. Appl. Microbiol. 10, 209–212.
[21] Kado, C.I. and Liu, S.T. (1981) Rapid procedure for detection and isolation of large and small plasmids. J. Bacteriol. 145, 1365–1373.
[22] Birnbiom, H.C and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7, 1513–1523.
[23] Pitcher, D.G., Saunders, N.A. and Owen, R.J. (1989) Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. Lett. Appl. Microbiol. 8, 151–156.
[24] Genetics Computer Group. Program for the Wisconsin Package. Version 8, Sept. 1994, Genetics Computer Group, Madison, WI.
[25] Hills, D.M., Huelsenbeck, J.P. and Cunningham, C.W. (1994) Application and accuracy of molecular phylogenies. Science 264, 671–677.
[26] Fukami, K. and Tateno,Y. (1991) Robustness of maximum likelihood tree estimation against different patterns of base substitutions. J. Mol. Evol. 32, 79–91.
[27] Tateno, Y., Takezaki, N. and Nei, M. (1994) Relative efficiencies of the maximum likelihood, neighbor-joining, and maximum parsimony methods when substitution rate varies with site. Mol. Biol. Evol. 11, 261–277.
[28] Tillier, E.R.M. and Collins, R.A. (1995) Neighbor joining and maximum likelihood with RNA sequence: Addressing the interdependence of sites. Mol. Biol. Evol. 12, 7–15.
[29] DeBry, R.W. and Abele, L.G. (1995) The relationship betwen parsimony and maximum likelihood analyses: Tree scores and confidence estimates for three data sets. Mol. Biol. Evol. 12, 291–297.
[30] Huelsenbeck, J.P. (1995) The robustness of two phylogenetic methods: Four-taxon simulations reveal a slight superiority of maximum likelihood over neighbor joining. Mol. Biol. Evol. 12, 843–849.
[31] Hall, L.M.C. (1994) Are point mutations or DNA rearrangements responsible for the restriction fragment length polymorphisms that are used to type bacteria? Microbiology. 140, 197–204.
[32] Ogram, A., Sayler, G.S. and Barkay, T. (1987) The extraction and purification of microbial DNA from sediments. J. Microbiol. Methods 7, 57–66.