Chapter 3

Molecular characterization of antibiotic resistance gene cassettes associated with class 1 integrons in Gram-negative multiple-antibiotic-resistant copiotrophic bacterial isolates from river Torsa and determination of the carriage of class 1 integrons on conjugative plasmids

3.1 Introduction

Sequence based identification of antibiotic resistance genes along with the flanking DNA region from different mobile DNA elements, including plasmids and transposons, their alignments and comparison have eventually led to the discovery of more recent device for resistance gene acquisition and transfer- the "integrons". In 1986, the DNA sequences of several seemingly unrelated antibiotic resistance genes heralded the first hints regarding integrons. Common regions were noted upstream and downstream of various antibiotic resistance genes. These regions were found to be in different places on various plasmids, suggesting that, like transposons, these elements are mobile. However, the element differed from transposons in two important characteristics: (i) Transposons have direct or indirect repeat sequences at their ends, but the regions surrounding the antibiotic resistance genes in the new elements were not repeats, and (ii) the elements contained a site-specific integrase gene of the same family as those found in phage but lacked many gene products associated with transposition. Due to these differences, the elements were not grouped with transposons and were named integrons (Martinez and de la Cruz 1990, Ouellette and Roy 1987, Stokes and Hall 1989, Sundstrom *et al.* 1988).

3.1.1 Integrons as tool of natural genetic engineering system

The present day definition of integrons has been formulated by Hall and Collis (1995). Integron units are naturally occurring gene expression systems that can potentially take into custody one or more circularized open reading frames (the so called gene cassettes) and convert them into functionally expressed genes (Martinez and de la Cruz 1990, Collis and Hall 1992, Collis *et al.* 1993, Hall and Stokes 1993, Hall and Collis 1995, 98, Rowe-Magnus *et al.* 1999, Rowe-Magnus *et al.* 2002). It is these gene cassettes that encode the resistance determinants to several antimicrobial agents (Fluit and Schmitz 2004). The integrons themselves are defective for self-transposition but this defect is often complemented through their association with IS, transposons and/ or conjugative plasmids that can serve as vehicles for the intra and inter-species transmission of these genetic structures. These novel DNA elements are frequently found as part of chromosomally located or plasmid residing transposons, including *Tn21, Tn1403, Tn1404, Tn1696, Tn1412* and *Tn 2000* (Carattoli *et al.* 2001, Naas *et al.* 2001, Partridge *et al.* 2001, Patridge *et al.* 2002, Sundin 2002, Villa *et al.* 2002). The plasmids that harbor integrons are often large (>100kb) conjugative plasmids belonging to groups IncFI, IncFII or IncL/M. Class 1 integrons are found associated with a variety of insertion sequence elements, including IS26, IS1999, IS2000 and IS6100. Most frequently occurring IS element at the 3' - end of integrons is IS6100.

3.1.2 Types and structural organization

In general, integrons are of two types- resistance integrons (RI) and super integrons (SI). Three classes of resistance integrons, namely class 1, class 2 and class 3 have been defined on the basis of the divergence among their integrase genes. The integrase gene is responsible for the synthesis of the enzyme, integrase,

which helps in the integration of gene cassettes and excision of gene cassettes in and from the integron. In other words, the enzyme integrase is responsible for integration and excision of gene cassettes. The enzyme recognizes site-specific recombination sites for integration. Nearly 45-58% homology exists between the three-integrase classes suggesting that their evolutionary divergence has extended over a longer period than the 50 years of the antibiotic era (Rowe-Magnus and Mazel 1999). The class 1 integron platform is the most ubiquitous among multi-drug resistant bacterial populations and is found associated with *Tn21* transposon family (Hall 1997). *Tn7* transposon family is associated with the class 2 integron platform (Hall and Stokes 1993). Arakawa *et al.* (1995) identified class 3 integron platform on a large transferable plasmid in a *Serratia marcescens* strain. Recently, the work of Correia *et al.* (2003) revealed the presence of a new class 3 integron on p22K9. The integron, which was previously designated class 4, is now named *Vibrio cholerae* SI (Fluit *et al.* 2004). This distinct type of integron is now known to be an integral component of many γ-proteobacterial genomes (Rowe-Magnus *et al.* 2001).

Integrons possess two conserved segments separated by a variable region that includes different combinations of inserted gene cassettes. The essential components found within the 5' conserved segment include the *intI* gene of tyrosine recombinase family (Nunes-Duby *et al.* 1998), which encodes a polypeptide of 337 amino acids, *attI* site which is recognized by the integrase and acts as a receptor for gene cassettes and on the opposite strand, a common promoter region (P_{ANT}) from which integrated gene cassettes are expressed (Collis and Hall 1995). The integrase mediates a site-specific recombination between the *attI* site in the integron and a secondary target called *attC* (59 base element) found in the downstream of the integron associated gene cassettes.

3.1.3 Class 1 integrons

Class 1, the best-characterized integrons, has been frequently reported in clinical (Martinez Freijo *et al.* 1998, Martinez Freijo *et al.* 1999, Chang *et al.* 2000, Schmitz *et al.* 2001, White *et al.* 2001, Thungapathra *et al.* 2002, Jones *et al.* 2003) and environmental isolates (Gebreyes *et al.* 2002, Chen *et al.* 2004, Nandi *et al.* 2004). Like the other classes of integrons, their 5' CS contains the *intI*1, and *attI* loci, while their 3' CS is specific and usually contain a truncated antiseptic resistance gene (*qacEΔ1*), a sulfonamide resistance gene (*sulI*) and an open reading frame (ORF5) of unknown function.

Integron integrases, as stated earlier, being members of tyrosine recombinase family, possesses four invariant residues (RHRY) and conserved motifs (boxes I and II; patches I, II, and III). Multiple alignments of integron itegrases with tyrosine recombinases has revealed that the DNA binding and recombination properties of class 1 integron integrase variants carried mutations at residues that are well conserved among the tyrosine recombinases and at some residues from the additional motif that are conserved among the integron integrases. The well conserved residues studied were H277 (histidine) from conserved tetrad RHRY (about 90% conserved), E121 (glutamate) found in patch I motif (about 80% conserved in prokaryotic recombinases), K171 from the patch II motif (nearly

100% conserved), W229 (Tryptophan) and F233 (Phenylalanine) from the patch III motif, and G302 (Glycine) of box II (about 80% conserved in prokaryotic recombinases). Additional intI mutated residues were K219 and a deletion of the sequence ALER 215. It was observed that E121, K171 and G302 play a role in the recombination activity but can be mutated without disturbing binding to DNA. W229, F233 and the conserved histidine (H277) may be implicated in problem folding or DNA binding. Some of the extra residues of *intII* seem to play a role in DNA binding (K219) which others are implicated in the recombination activity (Stokes *et al.* 1994).

The 59 base element (*att*C) is the recombination site, present on the gene cassette, recognized by the integron integrase (Stokes *et al.* 1997). Cassettes are inserted at *attI*, a unique integrase recombination site located in the 5' conserved region of integron adjacent to the integrase gene with a consensus sequence, GTTRRRY (core site). Boundaries of each inserted gene cassettes are defined by two core sites in the same orientation with the sequence GTTRRRY (R = Purine, Y = Pyrimidine) that are the target of the recombination process (recombination occurs between the G and the first T). The 59 be which occurs at 3' end consists of an inverted imperfect repeat between 50 and 150 bp which has an inverse core site at the 5' end of the inverted repeat and a core site at the 3' end. The insertion of gene cassette into *attI*1 site results in the formation of a secondary site (*att*C) downstream of the cassette.

Three formally distinct reactions can be catalyzed by the class 1 integrase, *intI1*, which involve recombination either between *attI1* and a 59 base element (be), two 59 be or between two *attI1* sites. Experimental evidences have shown that events involving two *attI1* sites are less efficient than the reactions in which a 59-be participates (Partridge *et al.* 2000). Actually an unusual reaction between the *attI*1 site and a 59-be appears to be responsible for the loss of the central region of a 59-be to create a potential fusion of two adjacent gene cassettes. The full *attI*1 site, 65 bp in length, is required for high efficiency recombination with a 59-be site. Each integron structure carries only one *attI*1 site located at the 5' border of the cassette closest to the promoter. All sites further downstream belong to the *attC* type (Hanson *et al.* 1997). The structural difference between the *attI*1 sites from that of the 59-be sites plays a vital role in ensuring that cassettes are preferentially integrated adjacent to the *attI*1 site of a class 1 integron. However, the only common feature between *attI*1 and 59 be is the 7 bp core site (GTTAGGC or GTTRRRY). Both an identifiable 7 bp inverse core site and the extensive inverted repeats associated with 59 be is absent in *attI1* (Reechia *et al.* 1994). The degrees of conservation in 59-base elements are not high and their length vary from 57 to 141 bp. They can be identified by their location and the relationship of over 20 bp at their outer ends to consensus sequences that are imperfect inverted repeats of one another.

3.1.4 Gene cassettes

The antibiotic resistance genes that integrons capture are located on the mobile gene cassettes. The cassettes consist of a promoterless coding sequence (most commonly an antibiotic resistance gene) and at the 3' end of this sequence, a

so-called 59 base element (*attC* site). Hence, gene cassettes are not necessarily part of the integron. In nature they exist in a free circularized form but cannot be replicated or transcribed in this form. A recombination event that occurs between *attI*1 and *attC* follows integration of the cassettes into the integron variable region. The gene on the cassette is then bound by the *attI*1 site on the 5'-side and by *attC* on the 3'-side. The integrase, apart from catalyzing the recombination between *attI*1 and *attC* sites, also facilitates excessive recombination events that can lead to loss of cassettes from an integron and generate free circular cassettes. In general, Gene cassettes consist of one coding sequence but may also contain a variable number of non-translated nucleotides (Fluit and Schmitz 1999). The gene cassettes in an integron are all inserted in the same orientation with respect to their coding regions and are expressed from a common promoter (P_{ANT}) region located in the 5' CS of the integron. This promoter is located 214 bases from the inner boundary of the 5' conserved segment (Collis and Hall 1995). In fact, the P_{ANT} of class 1 integrons potentially contains two promoters, P1 and P2. Four different P1 and two different P2 promoters have been described (Stokes and Hall 1989, Bunny *et al.* 1995). The start codons of many gene cassettes largely remained undetermined, but the first in frame start codon is generally assumed to function as such. This codon often position near the 5' end of the gene cassette and the supposed ribosome binding sites are weak at best. Some gene cassettes appear to carry their own promoter sequences but most gene cassettes are transcribed from a common promoter region. The first gene cassette with its own promoter described was the chloramphenicol resistance determinant cmlA (Stokes *et al.* 1991, Bissonette *et al.* 1991). The *qacE* and *qacG* gene cassettes carry their own promoter sequences as well (Paulsen *et al.* 1993, Laraki *et al.* 1999, Guerineau *et al.* 1990). Summarily, the gene cassettes may be characterized as follows-

1. Gene cassettes are discrete genetic elements that may exist as free, circular, non- replicating DNA molecules when moving from one genetic site to other, but are normally found as linear sequence that contribute part of a larger DNA molecule such as plasmid or bacterial chromosome.
2. Gene cassettes normally contain only a single gene and an additional short sequence, called 59 be, that functions as a specific recombination site.
3. The cassettes are small, normally ranging from 500-1000 bp.
4. The genes carried on gene cassettes usually lack promoters and are expressed from a promoter on integron.
5. In rare cases, a cassette may carry two genes; these exceptions are likely to have been generated by the fusion of two individual cassettes, which at one time were side by side. The double gene cassette being generated by a deletion that recovered sequences on either side of the joint boundary, including the 59 be that was located at the end of the first gene, i.e, the one that reads towards where the joint boundary was.

In naturally occurring integrons there appear to be no restrictions on the number or order of inserted cassettes. Arrays of several different antibiotic resistance genes can be created following successive integration or deletion of the gene cassettes from the respective integron structures (Figure 3.1 and Figure 3.2).

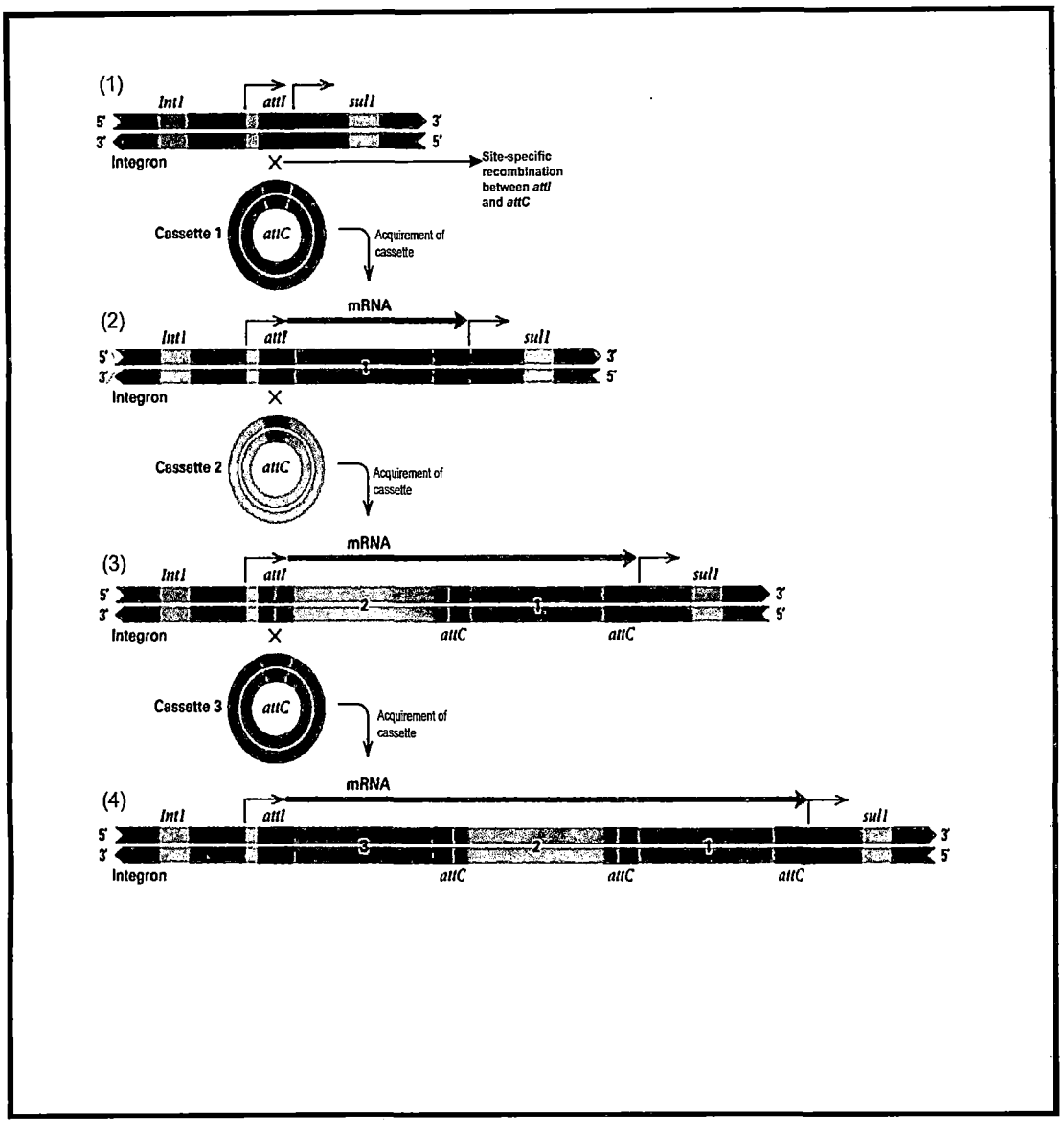


Figure 3.1. Sequential integration of antibiotic resistance gene cassettes within integron structure following site-specific recombination between *attI* site on integron and *attC* site on gene cassettes.

1. Showing site-specific recombination between *attI* and *attC* sites.
2. Acquirement of the second cassette by using the same *attI* site.
3. Addition of the third cassette into the integron.
4. Integron structure with three inserted gene cassettes

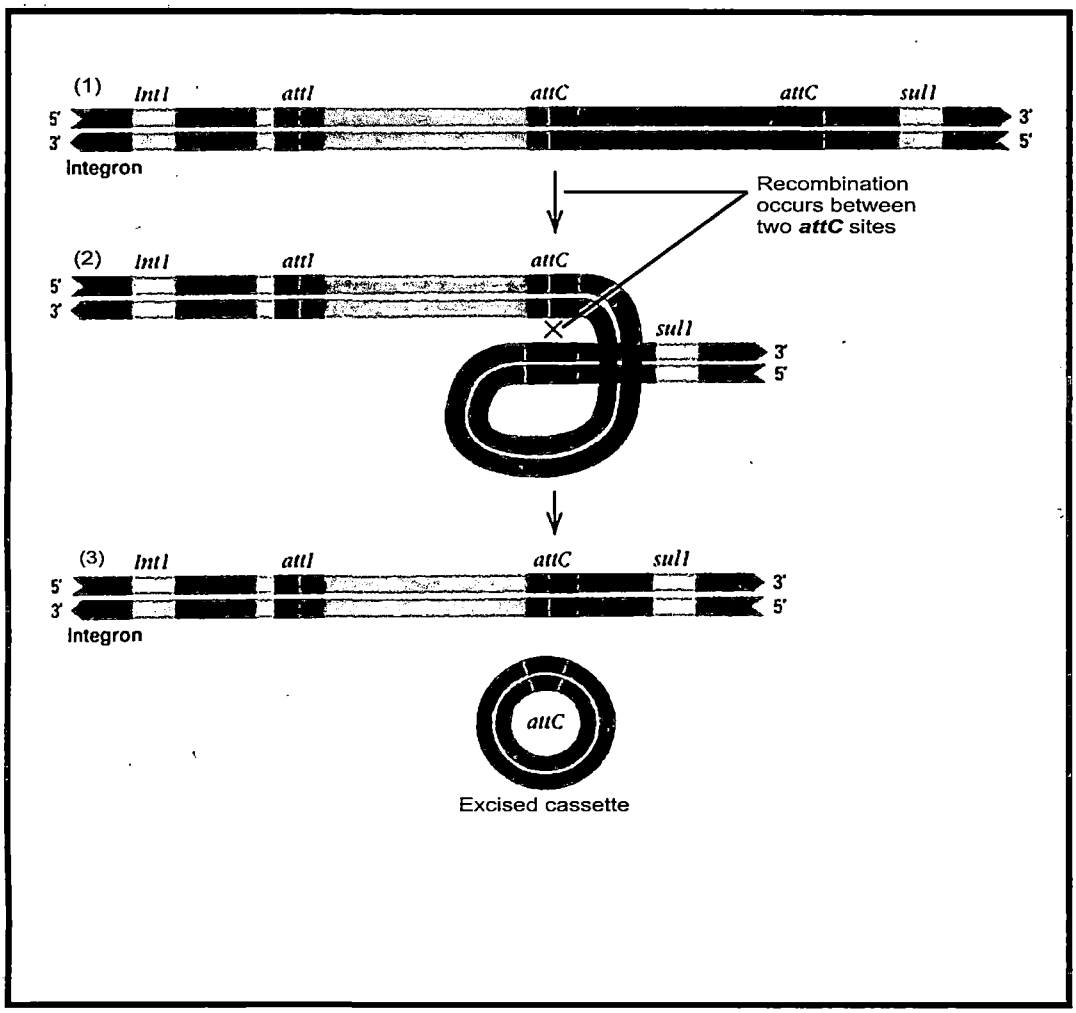


Figure 3.2. Diagrammatic sequence illustrating the excision of gene cassettes from an integron involving site-specific recombination between two *attC* sites

As an effect, the cassette containing the gene that encodes resistance to the antibiotic in the environment will be closest to the promoter. In this manner, the resistance gene necessary for the survival will be maximally expressed. Due to their ability to acquire new genes, integrons have a clear role in the evolution of the genomes of the plasmids and transposons that contain them. The most notable gene cassettes identified between the resistance integrons are those conferring resistance to antibiotics. More than 70 different antibiotic resistance genes, covering most classes of antimicrobials presently in use, are structured as gene cassettes (Fluit *et al.* 2004) and the list is growing.

3.1.4.1 Spread of gene cassettes

Dissemination of antibiotic resistance genes by horizontal transfer has led to the rapid emergence of bacterial multi-drug resistance (MDR), especially among the gram-negative enteric species (Chen *et al.* 2004). The spread of resistance is greatly enhanced when the resistance genes form part of a mobile gene cassette and reside themselves on resistance determinants like plasmids and transposons. The horizontal transfer of these gene cassettes occur by several mechanisms which include-

1. Mobilization of individual cassette by the integron encoded integrase.
2. Movement when the integron containing the cassette relocates-probably by targeted transposition.
3. Dissemination of large transposons such as *Tn21* carrying integrons and
4. Movement of conjugative plasmids containing integrons among different bacterial species.

Several studies have shown the role of integrons in mediating antibiotic resistance in a variety of enteric bacteria by lateral gene transfer. The spread of resistance genes is greatly enhanced following the movement of conjugative plasmids containing integrons among different bacterial species. In a study integrons from clinical isolates was chosen as a marker of resistance because of its association with multi-resistance. Conjugation experiments with these strains resulted in the transfer of complete resistance patterns at high frequencies (10^{-2} to 10^{-4}). These findings provide strong evidence in favor of the contribution of horizontal gene transfer to the emergence of multi-drug resistance (Leverstein van Hall *et al.* 2002). In a Korean study conjugal transfer and southern hybridization were performed to determine genetic localization of class 1 integrons found in *Escherichia coli* isolates from human and animal. The results revealed the role of horizontal transfer of class 1 integrons through conjugative plasmids for their wide dissemination (Kang *et al.* 2005). Shiga toxin producing *E. coli* (STEC) recovered from poultry, cattle, swine and humans were assayed for integron transfer by conjugation. Transfer of integrons by conjugation between strains of *E. coli* resulted in the transfer of antimicrobial resistance phenotypes. It was concluded that class 1 integrons, located on mobile plasmids, had facilitated the emergence and dissemination of antimicrobial resistance among STEC in humans and food animals (Singh *et al.* 2005). Another study reported that *Salmonella* genomic island 1 (SGI1) could be conjugally transferred from *Salmonella enterica* donor strains to non SGEI1 *S. enterica* and *E. coli* recipient strains (Doublet *et al.* 2005).

3.1.5 Integron Epidemiology

3.1.5.1 Integrons in clinical settings

Several groups have made systematic surveys on integron distribution in different environmental settings. A large proportion of such studies have dealt with the isolates from clinical settings where they have contributed significantly to the prevalence and dissemination of antibiotic resistance genes. One of the first such study was conducted by Sallen *et al.* (1995), who systematically screened 49 clinical enterobacterial isolates from one location in France and incidence of integrons was detected in 59% of them. Several studies have demonstrated wide distribution of class 1 integrons and their significant association with resistance to multiple classes of antibacterial

compounds in European hospitals. Schmitz *et al* (1999) tested 278 consecutive blood isolates belonging to 11 different gram-negative species and detected dominance of Class 1 integrons among them. 13% of these, belonging to six species, were shown to carry an integron. A similar result was reported from Netherlands where 135 strains belonging to seven species of Enterobacteriaceae carried an integron (Jones *et al.* 1997). The revelations from such studies could be extended to the rest of Western and Central Europe. It was shown that 42% of 163 strains (representing 13 species of Gram-negative bacteria), isolated from European countries, carried an integron. Another study on 900 blood culture isolates, representing five enterobacterial species, had revealed a gradual increase in the rate of pervasiveness of class 1 integrons among them over a period of seven years. The results of this study indicated that prevalence increased from 4.7% in 1993 to 9.7% in 1996 and finally to 17.4% in 1999 (Schmitz *et al.* 2001). A Chilean study investigated *Acinetobacter baumannii* isolates in which 17 integron carrying isolates were found (Gonzalez *et al.* 1998). In Taiwan, the presence of class 1 integrons was detected in 52% of the tested clinical isolates of *Escherichia coli* (Chang *et al.* 2000). Ploy *et al.* (2000) analyzed twenty *Acinetobacter baumannii* strains for integron content. The results indicated that integrons play a major role in development of multi drug resistance in *Acinetobacter*. The presence and spread of class 1 integrons among epidemiologically unrelated nontyphoidal *Salmonella* strains, isolated between 1989 and 1998 in a Spanish region had been ascertained by Guerra *et al.* (2000). Study by White *et al.* (2001) revealed the incidence of class 1 integrons in a collection of 120 urinary isolates collected from nine clinical settings in Sydney, Australia. A study, on the characterization of class 1 integrons in clinical strains of *Salmonella enterica* subsp. *enterica* serovars Typhimurium and Enteritidis, from Norwegian hospitals, was made by Lindstedt *et al.* (2003). Heir *et al.* (2004) demonstrated an overall low and stable prevalence of class 1 integron gene cassettes in clinical enterobacteriaceae and *E. coli* isolates in Norway.

3.1.5.2 Integrons in non-clinical samples

It is worth mentioning that human pathogenic samples are not the only reservoirs of integrons. Apart from their abundance in clinical samples, the distributions of class 1 integrons have also been evidenced from samples other than clinical interest. Integrons were found in gram-negative isolates from primates (Fluit *et al.* 1999). Integrons have even been detected in a *Pseudomonas* spp from an apple orchard (Schnabel *et al.* 1999). Mazel *et al.* (2000) conducted a study on clinically unselected enterobacteria to find out the prevalence of class 1 integrons among them. For this purpose, 72 *Escherichia coli* strains of the EcoR collection, isolated from a variety of animal hosts and a variety of geographic locations, were analyzed and only four of them were found to harbor the respective structure. In the United States, Goldstein *et al.* (2001) reported for the first time the distribution of class 1 to 4 integrases in veterinary Enterobacterial members isolated from livestock, companion animals, and exotics. Class 1 integrons have been identified in isolates obtained from cattle (Morabito *et al.* 2002), swine (Sunde *et al.* 2001, Gebreyes *et al.* 2002), chickens (Bass *et al.* 1999), and fish (L'abee-Lund *et al.* 2001). Incidence of

class 1 integrons had also been evidenced in *Salmonella* isolates recovered from retail meats purchased in United States and Peoples Republic of China (Chen *et al.* 2004). A large collection of Irish thermophilic *Campylobacter* isolates, from human and poultry sources, was investigated for the presence of integrons by O'Halloram *et al.* 2004. Their presence have also been described in isolates obtained from pets such as dogs (Sanchez *et al.* 2002) and zoo animals (Mazel *et al.* 2000). Integron sequences have also been detected from soils (Neild *et al.* 2001). Using culture independent methods, they detected several novel integrase genes. However, complete gene cassettes were not identified.

3.1.5.3 Integrons in aquatic environment

Prevalences of class 1 integrons have also been evidenced in various environmental samples including fish farms, estuary, irrigation water sources, and other aquatic environments. In a study, gram-negative bacteria isolated from an estuarine environment were analyzed to detect the presence of class 1 integrons. Three thousand isolates were examined for the said purpose and the presence of *intI*1 gene was detected among 109 of them. Characterization of the integrons revealed that majority lacked integrated gene cassettes in the variable region. The frequent identification of empty integrons led to support the view that antibiotic selective pressure might play a significant role in promoting the incorporation and maintenance of gene cassettes in integron variable regions (Rosser and Young 1999). A collection of 313 motile aeromonads isolated at Danish rainbow trout farms was analyzed to identify some of the genes involved in high levels of antimicrobial resistance. In 135 isolates combined sulfonamide/trimethoprim resistance appeared closely related to the presence of a class 1 integron (Schmidt *et al.* 2001). Coliform bacteria isolated from the aquatic environment were investigated for antibiotic susceptibilities and detailed structures of class 1 integrons. 24% of the isolates under study were found to carry *intI1* gene. Some of the isolates had the incomplete or nonfunctional class 1 integrons. The results indicated the possible role of antibiotic selective pressure for the maintenance of gene cassettes within the class 1 integron structure. In the absence of sustained antibiotic pressures, such as the aquatic environment, coliform bacteria may carry empty or non-functional class 1 integrons (Park *et al.* 2003). Irrigation water and sediments contaminated with fecal bacteria could serve as the source of class 1 and class 2 integron bearing *Escherichia coli* (Roe *et al.* 2003). Another study reported the incidence of enteric bacteria isolated from Mhlathuze River and the distribution of genetic elements that might be responsible for the observed antibiotic resistance. PCR based methods demonstrated the presence of class 1 integrons in more than 50% of those environmental bacteria that were also multiple antibiotic resistant (MAR). Conjugative plasmids were also isolated from a small percentage of the isolates. The study identified the Mhlathuze River as a reservoir of resistance genes and also as a medium for the spread of those genes (Biyela *et al.* 2004).

3.1.5.4 Integrons in Gram-positive bacteria

Furthermore, functional multi-resistance integrons are no longer restricted to the gram-negative bacteria. A survey by

Kazama *et al.* (1998) demonstrated the presence of *qacEΔ1* in both Staphylococcal and enterococcal isolates. A truncated integron in *M. fortuitum* and a complete and functional integron in another gram-positive bacterium, *Corynebacterium gluutamicum* have also been found. Antibiotic resistance genes and integrons in poultry house litter from commercial poultry farms have been quantified by Nandi *et al.* (2004). The authors found that integrons and associated resistance genes abound in several genera of gram-positive bacteria that constituted more than 85% of the litter community.

3.1.6 Integron study in India

All these data suggested that integrons were common worldwide, especially in Enterobacteriaceae, and that they contribute to antimicrobial resistance. Few studies in India have shown the prevalence and distribution of integrons from distinct clinical settings. Thungapathra *et al.* 2002, investigated the molecular mechanisms of multidrug resistance in *Vibrio cholerae* belonging to non-O1, non-O139 serogroups isolated during 1997 to 1998 in Calcutta, India. Twenty-two out of ninety-four strains were found to have class 1 integrons. The gene cassettes identified were *dfrA1*, *dfrA15*, *dfrA5*, *dfrA12*, *aac(6')-Ib*, *aadA1*, *aadA2* and *ereA2*. Their results indicated that besides class 1 integrons and plasmids, a conjugative transposon element, SXT, possibly contributed to the multiple antibiotic resistance. Other authors have also examined examined the distribution of class 1 integrons and SXT elements in *Vibrio cholerae* O1 E1 Tor strains, isolated in Calcutta, India, before and after the *V. cholerae* O139 outbreak in 1992 (Amita *et al.* 2003). Class 1 integrons, with *aadA1* gene cassettes, were detected primarily in the pre-O139 strains; the SXT element was found mainly in the post-O139 strains. The first report of identification of antibiotic resistance genes and a class 1 integron in *Vibrio fluvalis* H-O8942, isolated from a hospitalized infant aged 6 months suffering from cholera-like diarrhoea was made from India in 2002 (Ahmed *et al.* 2004). This isolate was found to carry a class 1 integron carrying a novel aminoglycoside adenyltransferase gene, *aac-(3')-Id*, and aminoglycoside adenyltransferase gene, *aadA7*. Phylogenetic analyses suggested that the *aac-(3')-Id* represents a fourth evolutionary lineage in the aminoglycoside acetyl transferase genes. The β-lactamase content of a multi-resistant strain of *Pseudomonas aeruginosa* clinical isolate from the Indian subcontinent was analyzed by Aubert *et al.* 2004 which further underlined its spread in Asia. In addition, the same report indicated that a bla_{VEB-1} like gene might be located outside a class 1 integron structure. The results pointed toward a fact that integrons may constitute a reservoir for the spread of antibiotic resistance genes that may be located and expressed either inside or outside an integron structure.

In this study, 100 Gram-negative multiple-antibiotic-resistant (MAR) isolates from the river Torsa which resisted 5 or more antibiotics were examined for the presence of class 1 integrons, employing a highly reproducible PCR strategy. Presence of class 1 integrons on conjugative plasmids of MAR isolates was demonstrated by conjugal transfer assay. The residence of class 1 integrons on conjugative R plasmids was confirmed in only six out of eleven integron-positive donor isolates. The sequences of the gene cassettes from the transconjugants were analyzed. Nucleotide sequence determination of 12

amplicons enabled to characterize the nature of gene cassettes of 20 isolates out of a total of 40 integron-positive MAR strains. The descriptions of ORFs, encoding proteins responsible for antibiotic resistance, have been presented.

3.2 Materials and Methods

3.2.1 MAR isolates for detecting the presence of class 1 integrons.

One hundred isolates were examined for the presence of class 1 integrons. The selected MAR isolates were primarily differentiated into pseudomonads and representatives of Enterobacteriaceae by oxidase and glucose fermentation tests before attempting the search for class 1 integron.

3.2.2 Antibiotic resistance determination

Antibiotic resistance was determined by the method described earlier in chapter 1. The antibiotics and the concentrations used were as follows: amikacin (25μg ml^{-1}), ampicillin (100μg ml^{-1}), cefotaxim (25μg ml^{-1}), cephalexin (25μg ml^{-1}), chloramphenicol (100μg ml^{-1}), gentamicin (25μg ml^{-1}), kanamycin (50μg ml^{-1}), netilmicin (25μg ml^{-1}), nitrofurantoin (25μg ml^{-1}), streptomycin (100μg ml^{-1}), tetracycline (20μg ml^{-1}), and tobramycin (25μg ml^{-1}). The isolates were considered multiple-antibiotic-resistant (MAR) if growth on at least two different antibiotic containing plates was at least equal to that on the growth control without antibiotics.

The MAR index of each individual isolate was scored by dividing the number of antibiotics (a) to which the isolate was resistant with the total number of antibiotics (b) to which the isolate was exposed, i.e., a/b (Krumperman 1983).

3.2.3 Detection of class 1 integrons by polymerase chain reaction (PCR) amplification

To identify the presence of class 1 integrons, a CS-PCR (conserved segment polymerase chain reaction) was performed according to the method described earlier (Levesque *et al.* 1995). Since primers 5′ CS (5′-GGCATCCAAGCAGCAAG -3′) and 3′ CS (5′-AAGCAGACTTGACCTGA-3′) used in this PCR anneal specifically in the 5′ and 3′ CS regions of class 1 integrons, the amplicons contained inserted gene cassettes flanked on both sides by small parts of the CSs [Leverstein van Hall *et al.* 2002]. Primer Int_2F (5′-TCTCGGGTAACATCAAGG-3′), specific for the 3′ region of the integrase gene (approximately 600 bp upstream from the 5′ CS primer site) was used in combination with the 3′ CS primer to show the proximity of the inserted gene cassettes to *intI* and to confirm the general structure of integron (Martinez Freijo *et al.* 1998, Schmitz *et al.* 2001).

The DNA templates for PCR were prepared as described by Levesque *et al.* 1995. MAR bacterial isolates were grown in 5ml LB in presence of a selective antibiotic at 37°C overnight, then 200μl of the overnight grown culture was added to 800μl of distilled water and boiled for 10 minutes. The bacterial suspension was then centrifuged at 6000 rpm for 5 minutes and the supernatant obtained was used as the template.

PCR amplification was performed using 'PCR Amplification Kit' (GENEI, India), in 50 μl reaction volume, following instructions provided by supplier. Each 50 μl PCR mix contains; 3 μl of 10mM dNTP mix, 5 μl of 10X buffer containing 15 mM $MgCl_2$, 12.5 pmol of each forward and

reverse primer, 15 µl of template DNA and 3U *Taq* DNA Polymerase. Taq DNA polymerase was added [1 µl of the 3U/ µl diluted solution] after 12 minutes at 94°C (hot start method). The PCR was done in a GenAmp PCR system (Applied Biosystems). The profile followed for the amplification: 1 minute of denaturation at 94°C, annealing for 1 minute at 55°C, and 2-3 minutes of extension at 72°C for a total of 30 cycles. In all reactions, PCR set up containing genomic DNA of *E. coli* XL1 Blue and sterile distilled water were used as target, which served as negative controls.

3.2.4 Size determination of the CS-PCR products

After PCR amplification, 10 µl of the amplification reactions were electrophoresed on 1% agarose (SRL, India) gel [1 gm agarose was added in 100 ml 1X TAE (Sambrook *et al.* 1989)] and run in 1X TAE buffer [Tris-acetate-EDTA, pH 7.8, prepared as recommended in Molecular Cloning Techniques (Sambrook *et al.* 1989)] containing 0.5 µg/ml ethidium bromide. A 500 bp ladder (Bangalore Genei, India) was used as the molecular size marker. After agarose gel electrophoresis, for visualization and documentation of the electrophoresed bands, a UV-transilluminator (Gibco-BRL, USA) and KD1SD-Software package (Kodak Digital Science, Japan) were used respectively.

3.2.5 Characterization of integrons by restriction fragment length polymorphism (RFLP) typing

To determine whether different isolates carried identical integrons, the amplicons of similar sizes were compared by RFLP typing. Selected amplified products generated by 5′ CS and 3′ CS primers were digested with *Eco*RI (New England Biolabs) restriction enzyme. Restriction analysis of the PCR products were done in a 20 µl reaction volume, containing 15 µl of the amplified PCR mix, 2 µl enzymes, 2 µl restriction buffer and 1 µl sterile distilled water, for 16 h at 30 ^{0}C. If the amplicons from two strains yielded the same RFLP pattern, two integrons were considered to be identical. If the PCR product contained a different RFLP pattern, the new product was sequenced as well.

3.2.6 Statistical analysis

The observations were classified simultaneously according to two attributes, MAR index [Low (0.41 – 0.58) and High (0.66–1.0)] and occurrence of integrons [presence (+) or absence (-)]. The frequencies in the different categories were arranged in a two-way table (known as 2 × 2 contingency table) (Table 3.1). The chi-square (χ^2) distribution was then used as test for independence of attributes, i.e., to test whether the two attributes are associated or not (Gupta 2001).

Table 3.1. 2×2 Contingency Table for χ^2 test

MAR index	*Incidence of class 1 integrons*		Total
	No. of integron + isolates	No. of integron – isolates	
Low (0.41 – 0.58)	6(a)	28 (b)	34 (R1)
High (0.66 – 1.0)	34 (c)	32 (d)	66 (R2)
	40 (C1)	60 (C2)	100 (N)

3.2.7 Identification of class 1 integron-positive MAR isolates

The cultures were identified according to Bergey's manual of systematic bacteriology [Brenner 1984]. Isolates were placed into genera or groups on the basis of cell and colonial morphology, Gram stain, motility, catalase and oxidase reactions, indole, Voges Proskauer, methyl red, citrate reactions, gelatin liquefaction, nitrate reduction, urease test, glucose oxidation and carbohydrate fermentations. The detailed descriptions of the biochemical tests and the analysis of the results have been made in chapter 4.

3.2.8 Detection of class 1 integrons on the conjugative plasmids of transconjugants by CS-PCR

The detailed description of selection of the potential donor isolates for conducting the conjugation experiment have been described in detail in section 2.2.8 of Chapter 2. The donors were the members of the pool of 81 MAR isolates that were grouped primarily as members of Enterobacteriaceae. The presence of class 1 integrons in both the donors and their respective transconjugants were detected by the same CS-PCR procedure as is described in section 3.2.3. The template DNA was prepared from both donors and transconjugants by boiling lysis. Plasmid DNAs isolated from the transconjugants was also used as template (~10ng) for the same PCR reaction. Genomic DNA prepared from the plasmid less *E.coli* DH5α was used as the negative control.

3.2.9 Identification of the amplicons generated from donors and their respective transconjugants

The amplicons were compared by RFLP typing to determine whether both donors and their respective transconjugants carried identical integrons. For this purpose enzymes *Eco*RI, *Pvu*I and *Nco*I were used. Restriction analysis of the PCR products was done in the same way as described in section 3.2.5. If the amplicons generated from both donor and transconjugant yielded the same RFLP pattern compared to that of the amplicon derived from the plasmid DNA isolated from the transconjugant, all three were considered to be identical.

3.2.10 Cloning of amplicons generated with 5′ CS and 3′ CS primers

3.2.10.1 Chemicals and reagents

All the chemicals used in the experiment were purchased from SRL Fine Chemicals (SRL India Ltd.). De-ionized double distilled water was used for the preparation of reagents. The reagents were either filter-sterilized or autoclaved, wherever required, before use and preserved in glass containers.

3.2.10.2 Elution of the PCR products from low melting point agarose

Approximately 100 µl of PCR products (generated by 5′ CS and 3′ CS primers) were directly precipitated by dehydrated ethanol and were suspended in 25 µl TE (10mM Tris-Cl : 1 mM EDTA, pH 7.5). 20 µl of the suspended DNA was loaded in an 'Low Melting Point Agarose' gel and were subjected to electrophoresis at 40–45 mV for at least 8 –10 h. Setting and running of the gel was performed at 4 ^{0}C in 1X TAE following techniques stated in 'Basic Methods in Molecular Biology' (Davis *et al.* 1987). The required band (s) were sliced and taken in an 2 ml microcentrifuge tube, and were melted at 65 ^{0}C for 5 min. To the melted agarose equal volume of TAE buffer was added and agarose particles were removed by repeated phenol extraction.

Finally DNA was precipitated, suspended in TE (10mM Tris-Cl: 0.1 mM EDTA, pH 7.5) and quantified.

3.2.10.3 Cloning of the PCR product

The pGEM-T Easy Vector System II, purchased form Promega Corporation, Madison, USA, was used for the cloning of PCR products following manufacturer's instructions.

I. Ligation

The vector (25 ng) and the insert DNA were taken in the molar ratio of 1:2 and were suspended in 4µl sterile double distilled water. 5 µl of 2X rapid ligation buffer and 1 µl of T4 DNA ligase were added and thoroughly mixed by vortexing and centrifugation. The mixture was kept at 4 ^{0}C for at least 16 h, heated at 60 ^{0}C for 10 min and 5-7 µl was used to transform competent *E. coli* cells.

II. Transformation

The competent cells of *Escherichia coli* XL1-Blue were transformed by the process described earlier in the section 2.2.13 of chapter 2.

III. Selection of recombinants

The recombinants were selected by α-complementation method (by blue-white screening). Recombinants were confirmed for the presence of the required insert by PCR reaction using 5′ CS and 3′ CS primers. Recombinant-plasmids were isolated and screened by alkaline lysis method (Birnboim and Doly, 1979). Purified clones were checked by *Eco*RI restriction digestion followed by agarose gel electrophoresis.

3.2.11 DNA sequencing and computer analysis of the sequence data

The recombinant plasmids were directly used for sequencing of the inserts using primers for T7 and SP6 promoters. Nucleotide sequencing was performed with the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kits (Parkin-Elmer) using specific primer and the reaction was analyzed in an 'AB1 PRISM 377 DNA Sequencer'.

3.2.12 Analysis of the sequence

Analysis of nucleotide sequences were done with several bioinformatics tools.

I. *Restriction analysis*: For determination of restriction map of the sequence, the software package NEB CUTTER v. 2.0 of New England Biolabs Inc. was used from the website http://tools.neb.com/NEBcutter2.

II. *Determination of open reading frame*: To determine the largest possible protein-coding region, the software package 'Fast PCR v 3.3.67' of the Institute of Biotechnology, University of Helsinki, Finland, was used.

III. *Similarity Search*: For similarity search studies of nucleic acid and amino acid sequences the BLAST N and BLAST P programs were used (Altschul *et al.* 1997) from the website http://www.ncbi.nlm.nih.

IV. *Conserve domain search* (CDD): The Domain analysis of the amino acid sequence of the inserts was done using the program of Marchler-Bauer and Bryant (2004) of the website http://www.ncbi.nlm.nih.

V. *Structure and function analysis*: For determination of probable secondary and three-dimensional structure of the peptides, the software packages, 'PredictProtein' from the website http://www.embl-heigelberg.de and 'SWISS-MODEL' (Schwede *et al.* 2003,

Guex and Peitsch, 1997) was used. The function of the peptides was analyzed using the ProtFun 2.1 software of the website http://www.cbs.dtu.dk.

3.2.13 Accession numbers

The nucleotide sequences of *dfrA1,* the *dfrA1* variant, *dfrA5, dfrA7,* the *dfrA12* variant, *dfrA17, aac-(6')-Ib, aadA1* and *aadA6* were assigned EMBL and GenBank accession numbers (Table. 3.4).

3.3 Results

3.3.1 Antimicrobial resistance pattern of MAR isolates

The antimicrobial resistance pattern of 100 MAR isolates was analyzed. All 100 isolates were resistant to both ampicillin and nitrofurantoin. The detail antimicrobial resistance pattern exhibited by the isolates has been presented in Table 3.2. Twenty nine were resistant to amikacin, 73 were resistant to cefotaxim, 96 were resistant to cephalexin, 76 were resistant to chloramphenicol, 57 were resistant to gentamicin, 70 were resistant to kanamycin, 43 were resistant to netilmicin, 53 were resistant to streptomycin, 89 were resistant to tetracycline and 49 were resistant to tobramycin. Among the 100 MAR bacteria distributed into 8 groups according to the MAR index, 9 isolates were resistant to all 12 antibiotics tested.

3.3.2 Detection of class 1 integrons and RFLP typing of the CS-PCR product

Class 1 integrons were sought in 100 MAR isolates. 40 (40%) of the isolates were identified to carry detectable class 1 integron structures. The remaining 60 isolates did not yield a product. The amplicon lengths, corresponding to the approximate sizes of the inserted cassette DNA, varied from 0.7 to 3.2 kb (Figure 3.3 and Table 3.3). The largest amplicon of 3.2 kb was detected singly in *Serratia* sp. TR 40. Among the 40 isolates, 32 yielded single amplicon and 8 yielded two amplicons of different sizes (Figure 3.3 and Table 3.3). PCR with a primer combination of Int_2F and 3' CS resulted in amplicons larger by an approximate size of 600 bp than the amplicons derived from 5' CS and 3' CS primers (Figure 3.4). Five isolates, TR 18, TR 40, TR 68,TR 76, TR 82 and TR 99 did not give any product when amplified with Int_2F and 3' CS primer set.

A 2.0 kb amplicon, recorded in eight strains (TR 06, TR 10, TR 11, TR 12, TR 13, TR 14, TR 15 and TR 17), yielded the same RFLP pattern upon digestion with *Eco*RI enzyme (Figure 3.5). Similarly, a 1.4 kb amplicon of TR 52 and TR 62 digested with *Eco*RI also yielded identical RFLP patterns.

3.3.3 Identification of integron-positive isolates

Biochemical characterization of integron-positive isolates revealed that 32 out of 40 isolates belong to the family Enterobacteriaceae. The other 8 gram-negative copiotrophic isolates did not belong to this family (Table 3.3). The isolates were identified up to the genus level following the principles of numerical taxonomy (data has been presented in chapter 4).

3.3.4 Test of significance

The number of integron positive isolates in low and high MAR index categories was 6 and 34; similarly, the number of integron negative isolates in the said two categories was 28 and 32 (Table 3.1). On the hypothesis of independence, the test statistic followed χ^2 distribution with 1

degree of freedom [χ^2 (corrected) = N {|ad – bc| - N/2}^{2} / R_1. R_2. C_1. C_2]. Since the observed value of the statistic, 12.18, was found greater than the tabulated value (value for χ^2 for 1 d.f at 5% level is 3.84), it was significant. We therefore rejected the null hypothesis at the 5% level of significance and concluded that the attributes were not independent; i.e. the data supported the alternative hypothesis that MAR index and integron carriage were associated.

3.3.5 Identification of integron-borne gene cassettes

Purified CS-PCR products of TR 02, TR 17, TR 40, TR 48, TR 52, TR 58, TR 59, TR 63, TR 85, TR 90, TR 95 and TR 97, cloned in pGEM-T Easy Vector (Figure 3.6), were subjected to DNA sequencing and gene cassettes thereby identified by sequence analysis (Table 3.3). Characterization of the gene cassettes revealed a significant association between the nature of the gene cassettes and the corresponding antibiotic resistance phenotype of the isolates. All integron-carrying strains expressed resistance to cotrimoxazole (trimethoprim + sulphonamide). The most common carriage by the integron-positive isolates involved dihydrofolate reductase cassettes conferring resistance to trimethoprim. These cassettes (*dfrA1, dfrA5, dfrA7, dfrA17* and two novel ORFs showing partial homology with *dfrA1* and *dfrA12* respectively) represented 70% of the cassettes detected (Table 3.3). Aminoglycoside adenyl transferase (*aad*A1 and *aad*A6) gene cassettes responsible for conferring resistance to streptomycin and spectinomycin were detected in four isolates, while the aminoglycoside acetyltransferase gene cassette (*aac*-(6')-*Ib*) was detected in only three isolates (Table 3.4). Resistance to other antibiotics like ampicillin, cephalosporins, chloramphenicol, nitrofurantoin and tetracycline did not correspond to identified gene cassettes. The sequence derived from 1.2 kb amplicon of TR 58 (Ac. No. AJ937774) showed 100% identity with *V.cholerae* class 1 integron *dfrA1* gene (Ac. No. AF221901). The sequence derivative from *E. coli* TR 02 (Ac. No. AJ620333) showed 100% identity with *E. coli* class 1 integron *dfrA5* gene (Ac. No. AJ419169), while the *Enterobacter* sp. TR 85 (Ac. No. AJ867252) revealed 98% identity with *E. coli* class 1 integron *dfrA5* gene (Ac. No. AJ419169); the sequence from *Salmonella* sp. TR 95 (Ac. No. AY604170) shared 98% identity with *Salmonella enterica* subsp. *enterica* serovar Typhi dihydrofolate reductase type *VII* gene (Ac. No. AY245101). The resulting sequence from *Serratia* sp. TR 48 (Ac. No. AJ868226) showed 98% identity with *Salmonella* sp. S126 class 1 integron *dfrA17* gene (Ac. No. AY263739). The BLASTP analysis of the sequence derived from the *Pseudomonas* sp. TR 52 (Ac. No. AJ620334) showed 100% identity with amino glycoside adenyl transferase AAD A6 of *Pseudomonas aeruginosa* (Ac. No. AF140629). The translated partial *aadA1* sequence of *Kluyvera* sp. TR 97 (Ac. No. AJ698461) showed 97% (189/193) identity with AAD A1 of *E. coli* isolate Ec 1484R (Ac. No. AY224185) and *S. typhimurium* (Ac. No. AJ496285). The partial CDS of *aadA1* gene as derived from *Serratia* sp. TR 40 (Ac. No.AJ938160) expressed 100% identity with the same gene from *Escherichia coli* isolate Ec1484R class 1 integron (Ac. No. AY224185). The complete CDS of *aac-(6')-Ib* from *Pseudomonas* sp. TR 59 (Ac. No. AJ697861) shared identity of 98 and 97% at nucleotide and protein level with amino glycoside 6'-N- acetyl transferase of *Vibrio*

cholerae class 1integron (Ac. No. AY103455) while the partial CDS of the same gene from *Yersinia* sp. TR 63 (Ac. No. AJ937775) shared 99% identity with the same *V. cholerae* (Ac. No. AY103455) class 1 integron. The sequence derived from the 0.8 kb amplicon of TR 58 (Ac. No.AJ938159) showed 99% identity (564/569) with *aac*-(6')-*Ib* gene of *Burkholderia cepacia* class 1 integron (Ac. No. AF371964).

3.3.5.1 Sequence analysis of CS-PCR product derived from TR 02

A 0.8 kb amplicon was derived from the isolate TR 02. The amplicon was cloned (pTR02) and the primer T7 was used for sequencing of the cloned fragment. Sequence analysis of the cloned fragment revealed the presence of a single ORF. This ORF represented an uninterrupted polypeptide sequence of 147 amino acids. The start of the ORF began with a GTG start codon coding valine at positions 101 to 103 (Figure 3.7) [Valine may also act as start codon in some cases (Adrian *et al.* 1998, Chang *et al.* 2000)]. A typical *dfrV* gene cassette yields a polypeptide of 157 amino acids, as is found in plasmid pLMO2O. Sequence of the gene cassette of TR 02 was read upto 147th amino acid and the ORF was expected to extend further by 10 amino acids in the region yet to be sequenced. Even if the chain length is reduced by 10 amino acids, it will not hamper the activity or the stability of the functional protein, since these amino acids are not in contact with many other residues that participate in numerous interactions and thereby do not affect the reaction rate by several orders of magnitude (Saraf *et al.* 2003). The amino acid at the 27th position of the putative translated product of TR 02 sequence was glutamate like other members of family1 DHFRs and the typical signature sequence of DHFR was also present at the N terminal end. The Blast N analysis of the sequence revealed 100% homology at the nucleotide level with plasmid pLMO20 *dhfr* V gene cassette (Accession no: X12868, cds: from 1306 to 1779), *E. coli* class 1 integron *dfr* V gene cassette (Accession no: AJ419169; cds: from 117 to 590), *Vibrio cholerae* class 1 integron bearing *dfr* V gene cassette (Accession no: AJ512546; cds: from 92 to 565). The translated polypeptide sequence was compared with other amino acids sequences in the Swiss Prot database. The result revealed that the translation product bears 99% identity with *dfr* V.

3.3.5.2 Sequence analysis of CS-PCR product derived from TR 17

A 2.0 kb CS-PCR product of *Citrobacter* sp. TR 17 as insert in pTR17 was sequenced with T7 promoter primer. The analysis of the partial sequence of the insert revealed the presence of one unpunctuated ORF of 154 amino acids (Figure 3.8). The translated polypeptide in BlastP analysis yielded best scores with all DHFR sequences and highest identity was observed with *dfrA12* protein. However a typical *dfrA12* gene cassette codes for a protein of 165 amino acids. The TR 17 ORF was short of the first 13 amino acids that were present in *dfrA12*. Typically a *dfrA12* polypeptide starts with methionine and has a valine residue at 11th position followed by two successive A residues before another M at position 14. In case of TR 17, ATG codon at nucleotide position 112 to114 was predicted to be the start codon. Though there could be the coding of a V followed by two A residues situated immediate upstream of the ATG codon, but this valine coded by GTT is most unlikely to be the start codon because reports of

using alternative start codons for various type I DHFRs is restricted to the using of either GTG or TTG. The first 120 amino acid stretch of TR 17 ORF bears 100% identity when aligned with *dfrA12* protein sequences while the remaining 34 amino acids of the partial sequence did not match with any of the existing sequences of the protein databases. The nucleotide sequence of TR 17 when multiply aligned with *dfrA12* sequences of *V. cholerae* (Ac. No AY103459), *Salmonella enterica* (Ac. No. AY126944), *E. coli* (Ac. No. Z21672), and *Citrobacter freundii* (Ac. No. AF550415), revealed single base pair deletion(s) at nucleotide position after 470 and 505 and two transversion events at c.561 and 562 that has resulted in a frame-shift leading to the alteration of protein sequence coded after the 120th amino acid of the TR 17 ORF. The DHFR profile-based multiple alignment, generated by MaxHom algorithm (PredictProtein), using TR 17 protein sequence as a query against SWISS-PROT database was subsequently fed into the neural network for secondary structure prediction by PHD. The predicted secondary structure of TR 17 protein was composed of 22.73% helix, 22.08% extended sheet and others 55.19%. The predicted secondary structure of dfrA12 protein is composed of 16.36% helix, 34.55% extended sheet and others 49.09%. The protein appeared as compact as a globular domain.

3.3.5.3 Sequence analysis of CS-PCR product derived from TR 40

A 3.2 kb CS-PCR product of isolate TR 40 cloned in pGEM-T easy vector (pTR40) was sequenced with SP6 promoter primer. Nucleotide sequence analysis revealed the presence of a single ORF of 469 bp encoding a polypeptide of 155 amino acids (Figure 3.9). The 5′ end of the ORF was not complete. The ORF was terminated with a TAA stop codon at positions 467 to 469. A core site, GTTRRRY (GTTAGAT at positions 524 to 530), was present at the 3′ end of the said ORF. The putative coding sequence bears 100% identity with *aadA1* gene sequence.

3.3.5.4 Sequence analysis of CS-PCR product derived from TR 48

An amplicon of 1600 bp was obtained by CS-PCR and was cloned in T vector. The insert from the respective clone pTR48 was sequenced with T7 promoter primer. Sequence analysis of the insert revealed the presence of an ORF, yielding a putative protein of 134 amino acids. The ORF began with GTG codon (coding for valine) at nucleotide positions 81 to 83 and ended with the stop codon TGA at positions 483 to 485 (Figure 3.10). The translated polypeptide in BlastP analysis yielded best scores with all DHFR sequences and highest identity was observed with dfrA17 protein. This polypeptide sequence was 23 amino acids shorter than the normal dfrA17. A normal *dfrA17* encoded protein is 157 amino acids long. The 30 amino acid residues present at the C-terminal end (128th to 157th) of the normal dfrA17 protein were not present in the polypeptide under study. Moreover, the last four amino acid residues of this polypeptide were completely different. The estimated half-life and the instability index (II) of the predicted protein were computed to be 1.2 hours (mammalian reticulocytes, in vitro) and 32.42 respectively, which classified the protein as stable. Predicted secondary structure (PHD) composition of this protein was: *H* (denoting α- helix) = 18.94 %; *E* (denoting extended β- strand/sheet) = 31.06 % and *L* (denoting others, loop) =

50.0 %, which was comparable to typical dfrA17 protein composed of *H* =17.20 %, *E* =36.94 % and *L* =45.86 %.

3.3.5.5 Sequence analysis of CS-PCR product derived from TR 52

The partial sequencing of the cloned fragment of *Pseudomonas* sp. TR 52 was done by using the promoter primer T7. Sequence analysis revealed the presence of one ORF of 94 amino acids. ATG was the start codon for this ORF (positions 91 to 93) and the termination codon was TGA at positions 373 to 375 (Figure 3.11). The 5′ conserved region was represented with a stretch of 72 nucleotides (from positions 9 to 80) located 10 bp upstream of the start of the ORF. The instability index of this putative translated product was computed to be 41.43, which classified the protein as unstable. The predicted secondary structure of this protein was H =32.98%; E =21.28% and L =45.74%. The protein may be globular but not as compact as a globular domain. The protein bears 100% identity with aminoglycoside adenyltransferase AAD A6 protein. An AAD A6 protein sequence typically constituted of 281 amino acids with a predicted secondary structure composition of *H* =49.47%, *E* =7.83 % and *L* =42.70% and appeared as a compact globular domain.

3.3.5.6 Sequence analysis of CS-PCR product derived from TR 58

Providencia sp. TR 58 yielded two amplicons of different sizes (1.2 kb and 0.8 kb) with primers 5′ CS and 3′ CS. The amplicons were separately cloned in pGEM–T Easy Vector. Sequencing analysis confirmed the presence of different gene cassettes in these two amplicons.

The 1.2 kb insert from pTR5801 was sequenced with T7 primer. The sequence analysis showed the presence of an ORF of 445 bp encoding a polypeptide of 148 amino acids. The ORF began with GTG start codon at positions 93 to 95 (Figure 3.12). The 5′ conserved region was represented with a stretch of 75 nucleotides (from positions 1 to 74) located 20 bp upstream of the ORF. The 3′ end of this ORF was not complete. The BlastP analysis of the coding sequence revealed 99% identity with dihydrofolate reductase type I sequences of the existing protein database.

The primer SP6 was used for sequencing of the 0.8 kb insert of the clone pTR5802. A potential ORF of 441 bp encoding a polypeptide of 146 amino acids was identified. The ORF began with ATG start codon at positions 94 to 96 and terminated with a TAA stop codon at positions 532 to 534 (Figure 3.13). Blast analysis of the sequence identified it as *aac-6′-Ib*, showing 99% identity with the existing aminoglycoside-6'-N-acetyl transferase gene sequences.

3.3.5.7 Sequence analysis of CS-PCR product derived from TR 59

The 811 bp insert from pTR59 was sequenced with both T7 and SP6 promoter primers. A stretch of 67 nucleotides (from 12 to 78) represented the 5' conserved segment. Further 68 nucleotides downstream of that conserved region, a 579 bp ORF potentially encoding a polypeptide of 192 amino acids was identified. The ORF began with GTG start codon at positions 147 to 149 and terminated with a TAA stop codon at positions 723 to 725 (Figure 3.14). No typical *E. coli* ribosome binding sites were identified upstream of the ATG start codon. Blast analysis of the coding sequence identified it as *aac-6′-Ib*, showing 97% identity with the existing aminoglycoside-6'-N-acetyl transferase

gene sequences. The nucleotide sequence immediately after the 3' end of *aac-6'-Ib* encoded a structure of 59 nucleotides (725 to 783) which was recognizable as an *att*C site (59 base element). The predicted secondary structure of this putative translated product was *H*=25.00%, *E*= 29.17% and *L*=45.83% which was comparable to the predicted secondary structure of the AAC-6-Ib protein (*H*=26.56%, *E*=27.60% and *L*=45.83%). The instability index of this protein was computed to be 40.74 that classified the protein as instable.

3.3.5.8 Sequence analysis of CS-PCR product derived from TR 63

The primer T7 was used for the sequencing of the cloned insert of the 900 bp in the recombinant plasmid pTR63. Further analysis revealed the presence of one unpunctuated ORF of 140 amino acids (Figure 3.15). A 75 nucleotides long 5' conserved region was located 142 nucleotides upstream of the ORF. The ORF began with the start codon ATG at positions 143 to 145. The 3' end of this ORF was not complete. The translated polypeptide of this ORF revealed maximum homology (99%) with the *aac-6'-Ib* gene product.

3.3.5.9 Sequence analysis of CS-PCR product derived from TR 85

The primer T7 was used for sequencing the 800 bp insert of pTR85. Nucleotide sequence analysis of the insert revealed the presence of a potential ORF encoding a polypeptide of 157 amino acids. The ORF began with GTG start codon at positions 51 to 53 and terminated with a TAA stop codon at positions 522 to 524 (Figure 3.16). The translated polypeptide sequence was compared with other amino acid sequences in the Swiss-Prot database. Sequences with the best scores were all dhfr sequences. The amino acid sequence of this ORF shared 99% identity with the normal dfrV protein.

3.3.5.10 Sequence analysis of CS-PCR product derived from TR 90

The complete sequence of 1177 bp insert was derived from pTR90 by using T7 and SP6 promoter primers. BlastN analysis of that sequence showed highest identity of 98% (1131/1150) with *Vibrio cholerae* class 1 integron sequence (Ac. No. AF455254) reported from India. Two ORFs, ORF1 and ORF2 were predicted from the sequence. ORF1 was found to code for the longest uninterrupted polypeptide sequence of 218 amino acids while the translation product of ORF2 led to possible hypothetical protein sequence of 148 amino acids. The ORF1 began with the atypical *E. coli* start codon GTG at position 97 to 99 and ended with stop codon TGA at positions 751 to 753 (Figure 3.17). Despite the unusual start codon, which normally codes for valine, it was the only codon, which was preceded by a plausible SD sequence (TGAGGAAGA). Since the ATG codon that was located four amino acids downstream of the GTG codon was not preceded by such a sequence, it was unlikely that this was the start codon. The ORF1 was flanked by a core element (GTTAACC) located 14 nucleotides 5′ to the start of the reading frame and differed from the consensus sequence of the core element (GTTRRRY) by a single nucleotide. Like most gene cassettes no recognizable promoter was present between the core element and the start of the ORF. The translated polypeptide sequence of ORF1 was compared with amino acid sequences in the Swissprot database. The first 152 amino acid sequence of ORF1 which yielded best scores were all DHFR type I

sequences. It is interesting to note that all dfrA1 protein sequences are 157 amino acids long and the first 152 residues of this novel ORF shared 95% identity with the existing *dfrA1* type. A stretch of 33 amino acid residues, from 153rd to 185th residue, did not find any resemblance with the existing sequences of the protein databases. Furthermore, another stretch of 25 amino acid residues, from residue 186 to 210, bore 60% identity with E1 protein of *E. coli*. Again the remaining C-terminal amino acids did not match with any of the existing protein sequences (Figure 3.18A).

Complete sequence alignment between *V. cholerae* (Ac. No. AF455254; 1179 bp) and *Morganella* sp. TR 90 (Ac. No. AJ698460; 1177 bp) gene cassettes revealed a DNA region/ mutation window of 216 bp [494 to 709 nucleotide position of *V. cholerae* sequence aligned to nucleotide position- 497 to 697 in *Morganella* sp. TR 90] that has presumably suffered 16 single base pair deletion and one base replacement event giving rise to the CDS of ORF1 protein (Figure 3.18B). The CDS, which was restricted to coding of 157 amino acids of *dfrA1* protein in *V. cholerae*, was found extended to a length of 218 amino acid residues before encountering a stop codon in *Morganella* sp. TR 90 sequence. Within the mutation window recombination crossover region (hs1) covers the sequence block from c.590 – 644 where six separate single base pair deletion events have been noted in the pair-wise sequence alignment (Fig 3.18B). Intrinsic polymerase pausing or pausing induced by a region of DNA secondary structure may initiate or promote misalignment. Once initiated, misalignment is directed by DNA sequence complementarity and therefore involves repetitive DNA sequence elements, such as direct repeats, inverted repeats, or other repetitive tracts including runs of mono-, di-, and tri-nucleotide (or other) repeats. The possible frame shifts due to the above mentioned reasons can fit into a model where deletion events of a typical *V. cholerae* like sequence may give rise to *Morganella* sp. TR 90 type sequence, resulting into read through across the non-coding region and merging with N-terminal coding region of the next ORF. The possible mutational events due to replication slippage have been illustrated in the Box 3.1.

Using ProtParam (ExPASy) tool, the estimated half-life and the instability index (II) of the predicted protein was computed to be 10 h (*E. coli* in vivo) and 32.08 respectively, which classified the protein as stable. The dihydrofolate reductase signature sequence of this protein revealed from PROSITE motif search [Bairoch *et al.* 1997] was 'VIGngpdIPWsakg.EqIIFkaiT'. The predicted secondary structure (PHD) composition of ORF1 protein was: *H* (denoting α- helix) = 19.25%, *E* (denoting extended β- strand/sheet) = 29.58%, and *L* (denoting others, loop) = 51.17%, which was comparable to *dfrA1* protein composed of *H* = 19.75%, *E* = 31.21% and *L* = 49.04%. Both the ORF1 protein and *dfrA1* protein appeared as compact globular domains in Globe prediction algorithm and no bonded cysteine was found in them as revealed by CYSPRED algorithm (Predict Protein). The predicted secondary structure and the presence of a dihydrofolate reductase signature indicated that the TR 90 protein might function like other DHFRs, similarly to the dfrA1 protein and might also confer resistance to trimethoprim.

3.3.5.11 Sequence analysis of CS-PCR product derived from TR 95

Both T7 and SP6 promoter primers were used for sequencing of the 718 bp long insert of pTR 95. The nucleotide sequence analysis showed the presence of a single gene, *dfr* gene. A 5' CS region (from positions 21 to 77) and a 59 base element at the 3' end (from positions 566 to 693) flanked the gene. This suggests that *dfr* gene is present as a gene cassette inserted into the variable region between the 5' and 3' conserved segments of a class 1 integron. This gene cassette contained an ORF of 324 bp encoding 107 amino acids. The ORF began with the start codon ATG at positions 248 to 250 and terminated with TAA stop codon at positions 569 to 571(Fig 3.19). The coding sequence, designated *dfrA7*, showed 98% identity with the dihydrofolate reductase type VII gene product. This polypeptide lacked a 50 amino acids long stretch at the N-terminal end, which is present in the normal dfrA7 protein. The estimated half-life and the instability index (II) of the predicted protein was computed to be 30 h (mammalian reticulocytes invivo) and 33.79 respectively, which classified the protein as stable. Predicted secondary structure (PHD) composition for the putative translated product was: *H* (denoting α- helix) = 15.89%; *E* (denoting extended β- strand/sheet) = 40.19%; and *L* (denoting others, loop) = 43.93%. The protein appears as compact, as a globular domain. On the contrary, the ideal dfrA7 protein may be globular but is not as compact as a domain with the predicted secondary structure composed of *H*=17.20%, *E*=35.03% and *L*=47.77%. The instability index of the dfrA7 protein is computed to be 30.48, which classifies the protein as stable.

3.3.5.12 Sequence analysis of CS-PCR product derived from TR 97

The 2.0 kb insert of pTR97 was sequenced with SP6 promoter primer. Partial sequencing of the said insert followed by nucleotide sequence analysis revealed the presence of a single ORF of 579 bp, encoding 192 amino acids (terminated with TAA stop codon at positions 584 to 586) (Figure 3.20). The 5' end of this ORF is not complete. The 3' end of the said ORF was flanked by a GTTRRRY core site (GTTAGAT at positions 641 to 647). Nucleotides from positions 588 to 641 represented the 3' portion of the 59 be. An inverse core site, RYYAAC (GTCTAAC at positions 588 to 594, differed from the consensus sequence by a single nucleotide), was also detected at the 3' end. The coding sequence shared 97% identity with the existing aminoglycoside adenyl transferase (*aadA1*) gene sequences.

3.3.6 Location of class 1 integrons on the conjugative plasmids

Out of the eighteen successful donor isolates (Table 2.2 of chapter 2) (all of which were the members of the said pool of 81 MAR isolates), eleven (TR 02, TR 04, TR 10, TR 13, TR 17, TR 37, TR 48, TR 56, TR 68, TR 79, TR 81 and TR 85) were found to carry detectable class 1 integron structures. When plasmid DNAs from the transconjugants was used as the template DNA for the PCR reaction, specific amplicons were produced from six (TR 02, TR 04, TR 10, TR 13, TR 17 and TR 85) of them (Table 3.5). The result confirmed the location of class 1 integrons on the conjugative plasmid that was transferred from the donor to the recipient. Upon digestion with restriction enzymes, CS-PCR products derived from the donor, its

Table 3.5. Class 1 integrons in donors and respective transconjugants

Donor	Class 1 integron in donor	Presence of integron in transconjugant	Nature of the gene cassette and accession no.
TR 02	+	+	*dfrA5* AJ620333
TR 10	+	+	*dfrA12*
TR 13	+	+	*dfrA12*
TR 17	+	+	*dfrA12* AY604169
TR 37	+	-	
TR 48	+	+	*dfrA17* AJ868226
TR 56	+	-	
TR 68	+	-	
TR 79	+	-	
TR 81	+	-	
TR 85	+	+	*dfrA5* AJ867252

transconjugant and from the plasmid isolated from the respective transconjugant, yielded identical RFLP patterns. The complete sequence analysis of the cloned amplicons of transconjugants produced 100% identity with the class 1 integron sequences derived from the respective donor isolates.

Table 3.2 Multidrug resistance combinations in 100 gram-negative bacterial isolates of River Torsa

Mar Index	No. of strains[a]	Resistance[b]											
		Ami	Amp	Cef	Cep	Chl	Gen	Kan	Net	Nit	Str	Tet	Tob
0. 41	1	-	+	-	+	-	-	+	-	+	-	+	-
	2	-	+	-	+	+	+	-	-	+	-	-	-
	1	-	+	+	+	-	-	-	-	+	-	+	-
	1	-	+	-	+	-	-	-	-	+	-	+	+
	1	-	+	-	+	-	-	-	-	+	+	+	-
	3	-	+	-	+	+	-	-	-	+	-	+	-
	1	-	+	-	-	+	-	-	-	+	+	+	-
	1	-	+	-	-	+	-	+	-	+	-	+	-
0. 50	5	-	+	+	+	+	-	-	-	+	-	+	-
	1	-	+	+	+	+	-	+	-	+	-	-	-
	1	-	+	-	+	+	-	+	-	+	-	+	-
	1	-	+	+	+	-	-	-	-	+	-	+	+
	1	-	+	-	+	+	-	+	-	+	-	+	-
	1	-	+	+	+	-	-	-	-	+	+	+	-
	1	-	+	-	+	+	-	-	-	+	+	+	-
	1	-	+	+	-	+	-	-	-	+	+	+	-
0. 58	3	-	+	+	+	+	-	+	-	+	-	+	-
	2	-	+	-	+	+	-	+	-	+	+	+	-
	1	-	+	+	+	+	+	-	-	+	-	+	-
	1	+	+	+	+	+	-	-	-	+	-	+	-
	1	-	+	+	+	-	+	-	+	+	-	+	-
	1	-	+	-	+	+	-	+	-	+	-	+	+
	1	+	+	+	+	+	-	+	-	+	-	-	-
	1	-	+	-	-	-	-	+	+	+	+	+	+
0. 66	6	-	+	+	+	+	-	+	-	+	+	+	-
	2	-	+	+	+	-	+	+	-	+	+	+	-
	1	-	+	+	+	-	-	+	+	+	+	+	-
	1	-	+	+	+	-	+	-	-	+	+	+	+
	1	-	+	+	+	+	+	-	-	+	+	+	-
	1	-	+	-	+	+	+	+	+	+	-	-	+
	1	-	+	+	+	-	+	+	-	+	-	+	+
	1	+	+	+	+	-	-	+	-	+	-	+	+
	1	-	+	-	+	+	+	+	-	+	+	+	-
	1	-	+	-	+	-	+	+	-	+	+	+	+
	1	+	+	+	+	-	-	+	-	+	+	+	-
	1	-	+	+	+	+	+	+	-	+	-	+	-
0. 75	1	-	+	+	+	-	+	+	+	+	+	-	+
	1	-	+	+	+	+	+	+	+	+	-	-	+
	3	-	+	+	+	-	+	-	+	+	+	+	+
	1	+	+	+	+	-	+	-	+	+	-	+	+
	1	-	+	+	+	+	+	+	-	+	+	+	-
	1	-	+	+	+	+	+	+	+	+	-	-	-
	1	+	+	+	+	+	+	+	-	+	-	+	-
	1	-	+	+	+	+	+	-	-	+	+	+	+
	1	-	+	+	+	+	-	-	+	+	+	+	+
	2	+	+	+	+	+	-	+	-	+	+	+	-
	1	+	+	-	+	-	+	+	+	+	-	+	+
	2	-	+	+	+	+	+	+	+	+	-	+	-
	1	-	+	-	+	+	+	-	+	+	+	+	+
	1	-	+	+	+	-	-	+	+	+	+	+	+
	1	-	+	+	+	+	+	+	-	+	-	+	+
0. 83	1	-	+	+	+	+	+	+	-	+	+	+	+
	2	+	+	+	+	+	+	+	+	+	-	-	+
	1	+	+	+	+	+	+	+	-	+	+	+	-
	2	-	+	+	+	+	+	+	+	+	-	+	+
	2	+	+	-	+	+	+	+	+	+	-	+	+
	2	-	+	+	+	+	+	+	+	+	-	+	+
	1	+	+	-	+	+	+	+	-	+	+	+	+
	1	-	+	+	+	-	+	+	+	+	+	+	+
0. 91	3	-	+	+	+	+	+	+	+	+	+	+	+
	2	+	+	+	+	+	+	+	+	+	-	+	+
	2	+	+	-	+	+	+	+	+	+	+	+	+
	1	+	+	+	+	+	+	+	+	+	+	-	+
1. 0	9	+	+	+	+	+	+	+	+	+	+	+	+

[a] Total number of isolates with a particular combination of antibiotic resistance.
[b]Ami: Amikacin; Amp: Ampicilin; Cef: Cefotaxim; Cep: Cephalexin; Chl: Chloramphenicol; Gen:Gentamicin; Kan: Kanamycin; Net: Netilmicin; Nit: Nitrofurantoin; Str: Streptomycin; Tet: Tetracycline; Tob: Tobramycin.

Table 3.3 MAR indices and sizes of inserted gene cassettes for forty gram-negative integron positive isolates of River Torsa

Mar Index	Isolates carrying class 1 integron	Identification	Approx. length of the amplicons generated with 5'cs and 3'cs (bp)
0.41	TR 53	*Escherichia* sp.	2700
0.50	TR 97	*Kluyvera* sp.	2000
	TR 76	*Proteus* sp.	2000
	TR 48	*Serratia* sp.	1600
0.58	TR 77	*Proteus* sp.	1200,1800
	TR 39	*Serratia* sp.	1900,3000
0.66	TR 6, TR 10 and TR 14	*Citrobacter* spp.	2000
	TR 81	*Providencia* sp.	1300
	TR 40	*Serratia* sp.	3200
0.75	TR 11, TR 12 , TR 13 and TR 15	*Citrobacter* spp.	2000
	TR 85	*Enterobacter* sp.	800
	TR 90	*Morganella* sp.	1177
	TR 56	*Proteus* sp.	1200,1800
	TR 18	*Providencia* sp.	1200
	TR 95	*Salmonella* sp.	718
	TR 27 and TR 92	*Pseudomonas* spp.	1200,1800
	TR 99	*Pseudomonas* sp.	1200
0.83	TR 91	*Citrobacter* sp.	800
	TR 02	*Escherichia* sp.	750
	TR 73	*Providencia* sp.	900
	TR 79	*Salmonella* sp.	900
	TR 63	*Yersinia* sp.	900
	TR20	*Acinetobacter* sp.	1300
0.91	TR 04	*Enterobacter* sp.	800
	TR 68	*Providencia* sp.	1200, 800
	TR 78	*Serratia* sp.	900
	TR 82	*Serratia* sp.	1200
	TR 49	*Acinetobacter* sp.	1100,900
	TR 59	*Pseudomonas* sp.	811
1.0	TR 17	*Citrobacter* sp.	2000
	TR 37	*Kluyvera* sp.	3000
	TR 58	*Providencia* sp.	1200,800
	TR 52	*Pseudomonas* sp.	1400
	TR 62	*Pseudomonas* sp.	1400

Table 3.4 Characterization of the gene cassettes in MAR isolates of river Torsa

Isolate	Antimicrobial resistance profile[a]	Nature of the gene cassette	Cassette encoded antibiotic resistance	Accession No.
TR 97	AMP, CEF, CEP, CHL, COT, NIT, TET,	*aad*A1	STR	AJ698461
TR 48	AMP, CEF, CEP, COT, NIT, TET, TOB	*dfrA17*	TMP	AJ868226
TR 6	AMP, CEF, CEP, CHL, COT, KAN, NIT, STR, TET	*dfrA12*	TMP	-[b]
TR 10	AMP, CEF, CEP, COT, GEN, NIT, STR, TET, TOB	*dfrA12*	TMP	-[b]
TR 14	AMP, CEF, CEP, CHL, COT, GEN, NIT, STR, TET	*dfrA12*	TMP	-[b]
TR 40	AMP, CEP, COT, GEN, KAN, NIT, STR, TET, TOB	*aad A1*	STR	AJ938160
TR 11, TR 12, TR 13	AMP, CEF, CEP, COT, GEN, NET, NIT, STR, TET, TOB	*dfrA12*	TMP	-[b]
TR 15	AMI, AMP, CEF, CEP, COT, GEN, NET, NIT, TET, TOB	*dfrA12*	TMP	-[b]
TR 85	AMI, AMP, CEP, COT, GEN, KAN, NET, NIT, TET, TOB	*dfrA5*	TMP	AJ867252
TR 90	AMP, CEF, CEP, CHL, COT, GEN, KAN, NET, NIT, TET	*dfrA1* variant	TMP	AJ698460
TR 95	AMP, CEF, CEP, COT, KAN, NET, NIT, STR, TET, TOB	*dfrA7*	TMP	AY604170
TR 02	AMP, CEF, CEP, CHL, COT, GEN, KAN, STR, TET, NIT, TOB	*dfrA5*	TMP	AJ620333
TR 63	AMI, AMP, CEP, CHL, COT, GEN, KAN, NET,NIT, TET, TOB	*aac*-(6')- *Ib*	AMI	AJ937775
TR 59	AMI, AMP, CEF, CEP, CHL, COT, GEN, KAN, NET,NIT, TET, TOB	*aac*-(6')- *Ib*	AMI	AJ697861
TR 17	AMI, AMP, CEF, CEP, CHL, COT, GEN, KAN, NET,NIT, STR, TET, TOB	*dfrA12*	TMP	AY604169
TR 58	AMI, AMP, CEF, CEP, CHL, COT, GEN, KAN, NET,NIT, STR, TET, TOB	*dfrA1*, *aac*-(6')- *Ib*	TMP AMI	AJ937774, AJ938159
TR 52	AMI, AMP, CEF, CEP, CHL, COT, GEN, KAN, NET,NIT, STR, TET, TOB	*aadA6*	STR	AJ620334
TR 62	AMI, AMP, CEF, CEP, CHL, COT, GEN, KAN, NET,NIT, STR, TET, TOB	*aadA6*	STR	-[c]

[a]AMI: Amikacin; AMP: Ampicillin; CEF: Cefotaxim; CEP: Cephalexin; CHL: Chloramphenicol; COT: Cotrimoxazole, GEN: Gentamicin; KAN:Kanamycin; NET: Netilmicin; NIT: Nitrofurantoin; STR: Streptomycin; TET: Tetracycline; TMP: Trimethoprim; TOB: Tobramycin

[b]RFLP pattern similar to TR 17 amplicon

[c]RFLP pattern similar to TR 52 amplicon

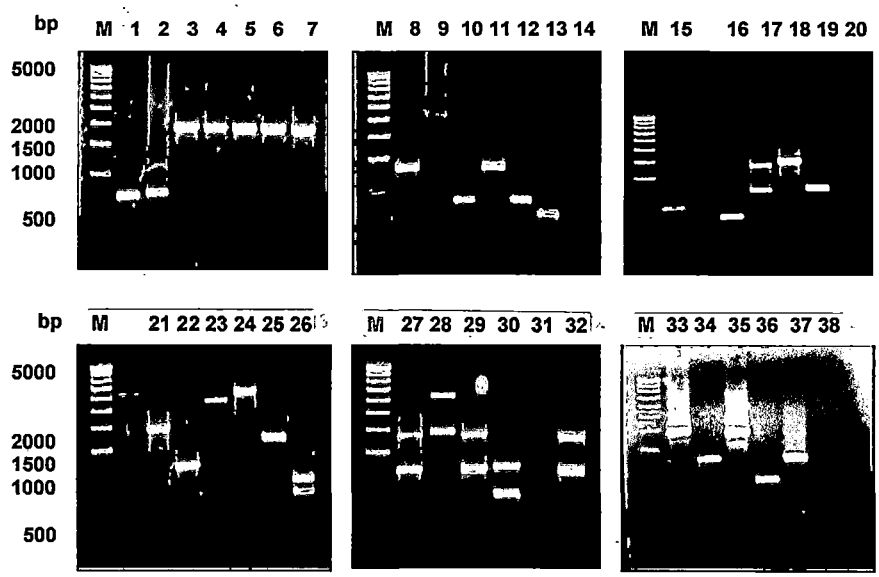


Figure 3.3. PCR amplification, using the 5' CS and 3' CS primers, of variable regions of class 1 integrons from Torsa isolates. The PCR products were separated by electrophoresis in 1.0%agarose.

Lane M, 500 bp DNA ladder; 1, TR 02; 2, TR 04; 3, TR 06; 4, TR 10; 5, TR 11; 6, TR 12; 7, TR 13; 8, TR 52; 9, TR 53; 10, TR 59; 11, TR 62; 12, TR 63; 13, TR 91; 14, TR 73; 15, TR 79; 16, TR 85; 17, TR 77; 18, TR 97; 19, TR 82; 20, TR 95; 21, TR 17; 22, TR 20; TR 37; 24, TR 40; 25, TR 48; 26, TR 49; 27, TR 27; 28, TR 39; 29, TR 56; 30, TR 58; 31, TR 68; 32, TR 92; 33, TR 14; 34, TR 18; 35, TR 76; 36, TR 78; 37, TR 90; 38, TR 73.

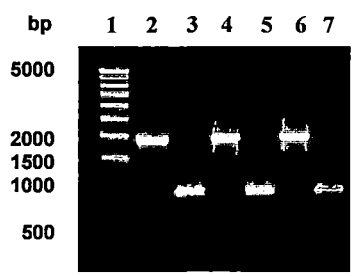


Figure 3.5. Agarose gel electrophoresis of *Eco*RI restricted 2000 bp amplicons generated with 5' CS and 3' CS primers. Lane 1, 500 bp DNA ladder; 2, TR 06 ; 3, *Eco* RI digest of TR 06 ; 4, TR 10 ; 5, *Eco* RI digest of TR 10 ; 6, TR 17 ; 7, *Eco* RI digest of TR 17

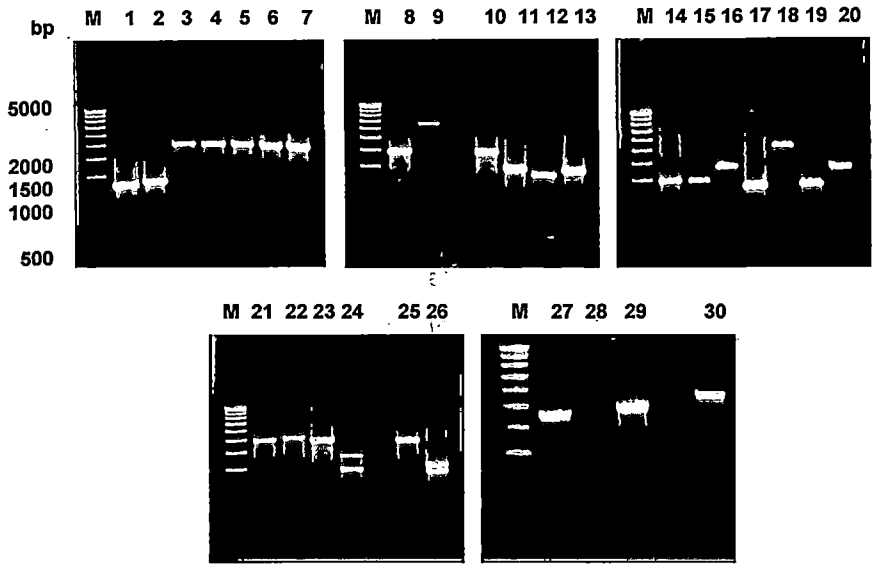
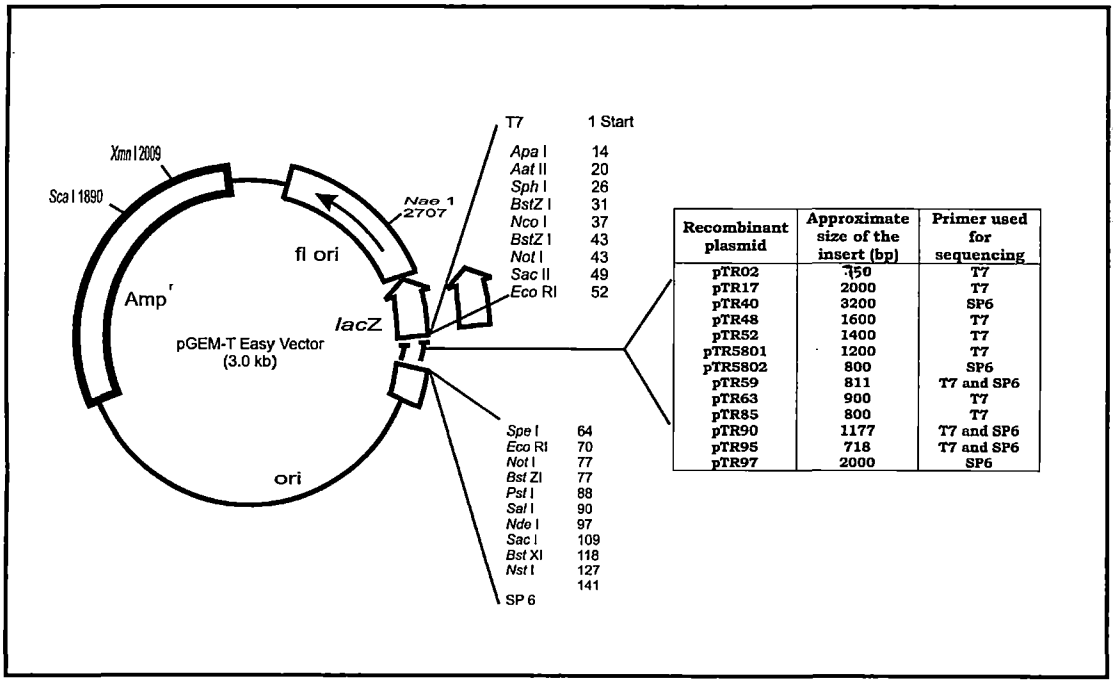


Figure 3.4. PCR amplification, using the Int_2F and 3' CS primers, of variable regions of class 1 integrons from Torsa isolates. The PCR products were separated by electrophoresis in 1.0% agarose. Lanes M, 500 bp DNA ladder; 1,TR 02 ; 2, TR 04 ; 3, TR 06 ; 4, TR 10 ; 5, TR 11 ; 6, TR 12 ; 7, TR 13 ; 8, TR 52 ; 9, TR 53 ; 10, TR 62 ; 11, TR 63 ; 12, TR 91 ; 13, TR 73 ; 14, TR 78 ; 15, TR 79 ; 16, TR 81 ; 17, TR 85 ; 18, TR 97 ; 19, TR 95 ; 20, TR 90 ; 21, TR 27 ; 22, TR 39 ; 23, TR 56 ; 24, TR 58 ; 25, TR 77 ; 26, TR 49 ; 27, TR 48; 28, TR 37; 29, TR 92; 30, TR 20.



Recombinant plasmid	Approximate size of the insert (bp)	Primer used for sequencing
pTR02	150	T7
pTR17	2000	T7
pTR40	3200	SP6
pTR48	1600	T7
pTR52	1400	T7
pTR5801	1200	T7
pTR5802	800	SP6
pTR59	811	T7 and SP6
pTR63	900	T7
pTR85	800	T7
pTR90	1177	T7 and SP6
pTR95	718	T7 and SP6
pTR97	2000	SP6

Figure 3.6. Cloning of CS-PCR product(s) in pGEM-T Easy Vector
Amp^r, gene for ampicillin resistance; *ori*, origin of replication; *lacZ*, gene for the synthesis of β-galactosidase.

```
CGCTGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAG 60

GGCAGTCGCCCTAAAACAAAGTTAACCCGGAACCAAAATTGTGAAAGTATCATTAATGGC 120
                                          V  K  V  S  L  M  A
TGCAAAAGCGAAAAACGGAGTGATTGGTTGCGGTCCACACATACCCTGGTCCGCGAAAGG 180
 A  K  A  K  N  G  V  I  G  C  G  P  H  I  P  W  S  A  K  G
AGAGCAGCTACTCTTTAAAGCCTTGACGTACAACCAGTGGCTTTTGGTGGGCCGCAAGAC 240
 E  Q  L  L  F  K  A  L  T  Y  N  Q  W  L  L  V  G  R  K  T
GTTCGAATCTATGGGAGCACTCCCTAATAGGAAATACGCGGTCGTTACTCGCTCAGCCTG 300
 F  E  S  M  G  A  L  P  N  R  K  Y  A  V  V  T  R  S  A  W
GACGGCCGATAATGACAACGTAATAGTATTCCCGTCGATCGAAGAGGCCATGTACGGGCT 360
 T  A  D  N  D  N  V  I  V  F  P  S  I  E  E  A  M  Y  G  L
GGCTGAACTCACCGATCACGTTATAGTGTCTGGTGGCGGGGAGATTTACAGAGAAACATT 420
 A  E  L  T  D  H  V  I  V  S  G  G  G  E  I  Y  R  E  T  L
GCCCATGGCCTCTACGCTCCATATATCGACGATTGATATTGAGCCGGAAGGAGATGTTTT 480
 P  M  A  S  T  L  H  I  S  T  I  D  I  E  P  E  G  D  V  F
CTTTCCGAATATTCCCAATACCTTCGAAGTTGTTTTTGAGCAACACTTTAGCTCAAACAT 540
 F  P  N  I  P  N  T  F  E  V  V  F  E  Q  H  F  S  S  N  I
```

Figure 3.7 [A]. Partial nucleotide sequence of CS-PCR product of TR 02 showing ORF and the putative translated product

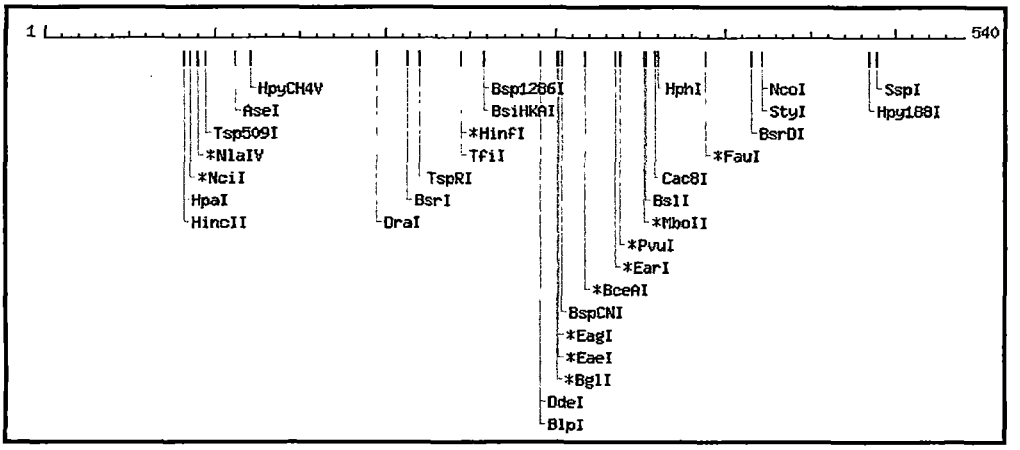


Figure 3.7 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 02. *, cleavage affected by methylations

```
GGCAAGTCTGCTAAACCCGTAANCAGGNTCNGGNTGCANAACAACAGGGTNTGTAGTCAG 60


GTCAAGTCTGCTTAAACTCGGGAATCAGTACGCATTCATCTCGTTGCTGCGATGGGAGCC 120
                                                      M  G  A
AATCGGGTTATTGGCAATGGTCCTAATATCCCCTGGAAAATTCCGGGTGAGCAGAAGATT 180
N  R  V  I  G  N  G  P  N  I  P  W  K  I  P  G  E  Q  K  I
TTTCGCAGACTCACTGAGGGAAAAGTCGTTGTCATGGGGCGAAAGACCTTTGAGTCTATC 240
F  R  R  L  T  E  G  K  V  V  V  M  G  R  K  T  F  E  S  I
GGCAAGCCTCTACCGAACCGTCACACATTGGTAATCTCACGCCAAGCTAACTACCGCGCC 300
G  K  P  L  P  N  R  H  T  L  V  I  S  R  Q  A  N  Y  R  A
ACTGGCTGCGTAGTTGTTTCAACGCTGTCGCACGCTATCGCTTTGGCATCCGAACTCGGC 360
T  G  C  V  V  V  S  T  L  S  H  A  I  A  L  A  S  E  L  G
AATGAACTCTACGTCGCGGGCGGAGCTGAGATATACACTCTGGCACTACCTCACGCCCAC 420
N  E  L  Y  V  A  G  G  A  E  I  Y  T  L  A  L  P  H  A  H
GGCGTGTTTCTATCTGAGGTACATCAAACCTTCGAGGGTGACGCCTTCTTCCAATGCTCA 480
G  V  F  L  S  E  V  H  Q  T  F  E  G  D  A  F  F  Q  C  S
ACGAAACAGAATTCGAGCTTGTCTCACCGAAACCATTCAAGCTGTAATTCCGTACACCCA 540
T  K  Q  N  S  S  L  S  H  R  N  H  S  S  C  N  S  V  H  P
CTCCGTTTATGCGCGTCGAACGGGCTAACATTCCG                          575
L  R  L  C  A  S  N  G  L  T  F
```

Figure 3.8 [A]. Partial nucleotide sequence of CS-PCR product of TR 17 showing ORF and the putative translated product.

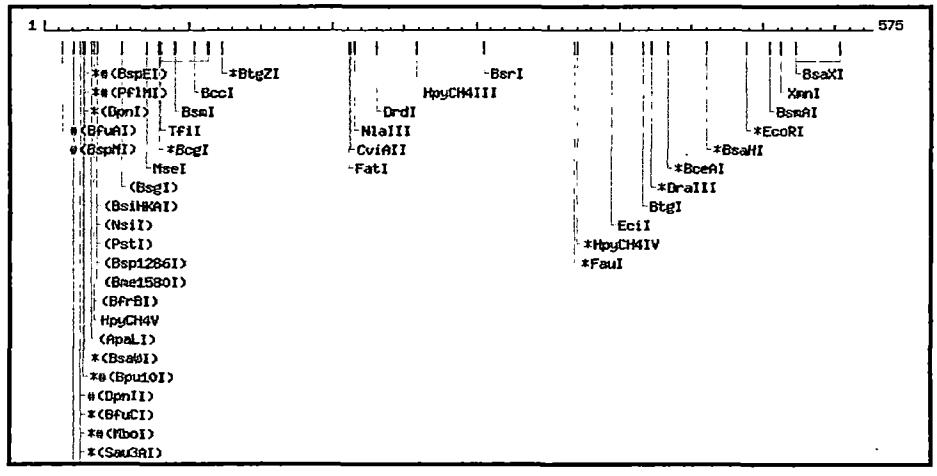
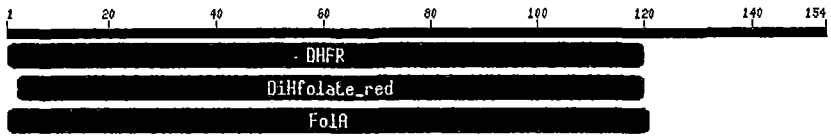


Figure 3.8 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 17. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
        ....,....1....,....2....,....3....,....4....,....5....,....6
AA      MGANRVIGNGPNIPWKIPGEQKIFRRLTEGKVVVMGRKTFESIGKPLPNRHTLVISRQAN
PHD_sec                    HHHHHHHHH      EEEE               EEEEEEE

        ....,....7....,....8....,....9....,....10.1.,....11.1.,....12.1
AA      YRATGCVVVSTLSHAIALASELGNELYVAGGAEIYTLALPHAHGVFLSEVHQTFEGDAFFQCS
PHD_sec      EEEE  HHHHHHHHH      EEEEE HHHHHHHHHH    EEEEEEE

        .,....13.1.,....14.1.,....15.1.
AA      TKQNSSLSHRNHSSCNSVHPLRLCASNGLTF
PHD_sec      EEEEEE
```

Figure 3.8 [C]. Predicted secondary structure of the putative translated product of ORF of TR 17 amplicon

Result of conserve domain search for the putative translated product obtained from ORF of TR 17




```
gnl|CDD|27775 cd00209, DHFR, Dihydrofolate reductase (DHFR).
              CD-Length = 158 residues, 77.2% aligned
              Score =  166 bits (423), Expect = 9e-43
Query:  1     MGANRVIGNGPNIPWKIPGEQKIFRRLTEGKVVVMGRKTFESIGK-PLPNRHTLVISRQA  59
Sbjct:  7     VDENGVIGKDNKLPWHLPEDLKHFKKTTTGNPVIMGRKTFESIPRRPLPGRTNIVLSRQL  66

Query:  60    NYR-ATGCVVVSTLSHAIALASELGNELYVAGGAEIYTLALPHAHGVFLSEVHQTFEGDA  118
Sbjct:  67    DYQDAEGVEVVHSLEEALELAENTVEEIFVIGGAEIYKQALPYADRLYLTRIHAEFEGDT  126

Query:  119   FF  120
Sbjct:  127   FF  128


gnl|CDD|22940 pfam00186, DiHfolate_red, Dihydrofolate reductase..
              CD-Length = 173 residues,  75.1% aligned
              Score =  140 bits (353), Expect = 1e-34
Query:  3     ANRVIGNGPNIPWKIPGEQKIFRRLTEGK---VVVMGRKTFESIG---KPLPNRHTLVIS  56
Sbjct:  10    KNGGIGKDGDLPWRLPNDLKYFKAVTTGTPRNAVIMGRKTWESIPEKFRPLPGRLNIVLS  69

Query:  57    RQANYRATGCVVVSTLSHAIALAS------ELGNELYVAGGAEIYTLALPHAHGVFLSEV  110
Sbjct:  70    RSEDYDAQGDNVVVSSSIEAALDLLAEPPEASIERVFVIGGAQLYAAALPLADRLYLTRI  129

Query:  111   HQTFEGDAFF  120
Sbjct:  130   DGEFEGDTFF  139


gnl|CDD|10137 COG0262, FolA, Dihydrofolate reductase [Coenzyme metabolism].
              CD-Length = 167 residues,  77.2% aligned
              Score =  110 bits (277), Expect = 7e-26
Query:  1     MGANRVIGNGPNIPWKIPGEQKIFRRLTEGKVVVMGRKTFESIG---KPLPNRHTLVISR  57
Sbjct:  8     VSLDGVIGRDNSLPWHLPEDLAHFKATTLGKPVIMGRKTYESLPGEWRPLPGRKNIVLSR  67

Query:  58    QANYRATGCVVVS-TLSHAIALA-SELGNELYVAGGAEIYTLALPH--AHGVFLSEVHQT  113
Sbjct:  68    NPDLKTEGGVEVVDSIEEALLLLLKEEGEDIFIIGGGELYRQFLPAGLADELILTIIPVL  127

Query:  114   F-EGDAFFQ  121
Sbjct:  128   LGEGDTLFP  136
```

```
ATTTGGAGAATGGCAGCGCAATGACATTCTTGCAGGTATCTTCGAGCCAGCCACGATCGA 60
 F  G  E  W  Q  R  N  D  I  L  A  G  I  F  E  P  A  T  I  D
CATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCC 120
 I  D  L  A  I  L  L  T  K  A  R  E  H  S  V  A  L  V  G  P
AGCGGCGGAGGAACTCTTTGATCCGGTTCCTGAGCAGGATCTATTTGAGGCGCTAAATGA 180
 A  A  E  E  L  F  D  P  V  P  E  Q  D  L  F  E  A  L  N  E
AACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCT 240
 T  L  T  L  W  N  S  P  P  D  W  A  G  D  E  R  N  V  V  L
TACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGC 300
 T  L  S  R  I  W  Y  S  A  V  T  G  K  I  A  P  K  D  V  A
TGCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAG 360
 A  D  W  A  M  E  R  L  P  A  Q  Y  Q  P  V  I  L  E  A  R
ACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGTGCAGATCAGTTGGAAGA 420
 Q  A  Y  L  G  Q  E  E  D  R  L  A  S  R  A  D  Q  L  E  E
ATTTGTTCACTACGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAATGTCTAACAAT 480
 F  V  H  Y  V  K  G  E  I  T  K  V  V  G  K  *(3' end of ORF)
TCGTTCAAGCCGACGCCGCTTCGCGGCGCGGCTTAACTCAAGCGTTAGATGCACTAAGCA 540

CATAATTGCTCACAGCCAAACTATCAGGTCAAGTCTGCTTA                     581
```

Figure 3.9 [A]. Partial nucleotide sequence of CS-PCR product of TR 40 showing ORF and the putative translated product (* indicate stop codon).

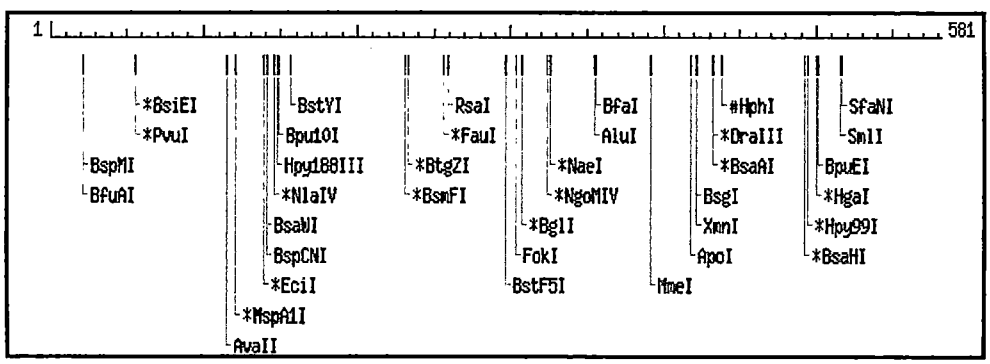


Figure 3.9 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 40. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
CGCGGTCGACGTATGGAGCAGCACGATCGTGTACTGCAGCAGGGCAGGTCGCCCTCNACA 60

CAGAGTTAGCCATTAAGGGAGTGAAATTGAAAATATCATTGATTTCTGCAGTGTCAGAAA 120
                    V  K  L  K  I  S  L  I  S  A  V  S  E  N
ATGGCGTAATCGGTAGTGGTCCTGATATCCCGTGGTCAGTAAAAGGTGAGCAACTACTCT 180
  G  V  I  G  S  G  P  D  I  P  W  S  V  K  G  E  Q  L  L  F
TTAAAGCGCTCACATATAATCAATGGCTCCTTGTCGGAAGAAAAACATTTGACTCTATGG 240
  K  A  L  T  Y  N  Q  W  L  L  V  G  R  K  T  F  D  S  M  G
GTGTTCTTCCAAATCGCAAATATGCAGTAGTGTCAAAGAACGGAATTTCAAGCTCAAATG 300
  V  L  P  N  R  K  Y  A  V  V  S  K  N  G  I  S  S  S  N  E
AAAACGTCCTAGTTTTTCCTTCAATAGAAAATGCTTTGAAAGAGCTATCAAAAGTTACAG 360
  N  V  L  V  F  P  S  I  E  N  A  L  K  E  L  S  K  V  T  D
ATCATGTATATGTCTCTGGCGGGGGTCAAATCTATAATAGCCTTATTGAAAAAGCAGATA 420
  H  V  Y  V  S  G  G  G  Q  I  Y  N  S  L  I  E  K  A  D  I
TAATTCATTTGTCTACTGTTCACGTTGAAGTCGAAGGTGATATCAAAATTCCCTATAATG 480
  I  H  L  S  T  V  H  V  E  V  E  G  D  I  K  I  P  Y  N  A
CCTGAGAATTTCAATTTGGTTTTTGAACAGTTTTTTATGTCTAATATAAATTATACATAC 540
  *(3' end of ORF)
CAGATTTGGAAAAAAGGCTAACAATGCGTTGCAGCACCAGTCGCTTCGCTCCTTGGACAG 600

CTTTTAAGTCGCGTCTTTGGGGGTTTN                                   627
```

Figure 3.10 [A]. Partial nucleotide sequence of CS-PCR product of TR 48 showing ORF and the putative translated product (* indicate stop codon).

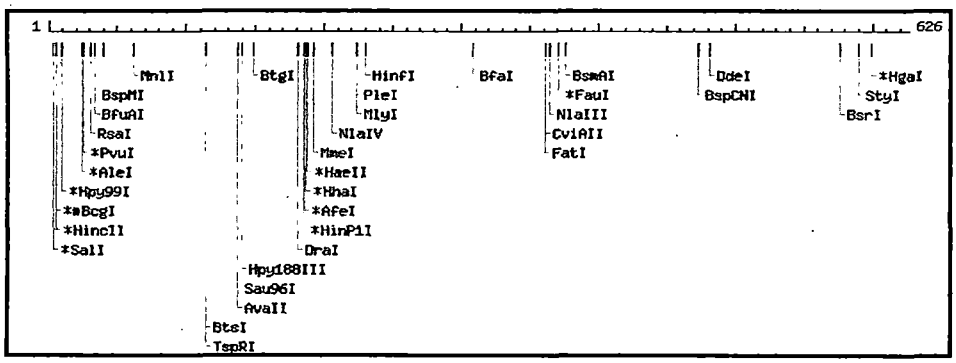
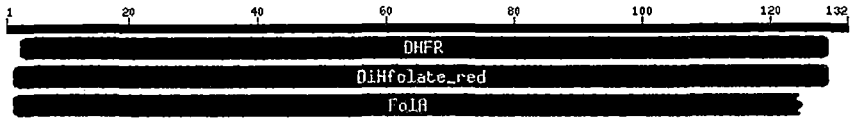


Figure 3.10 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 48. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
        ....,....1....,....2....,....3....,....4....,....5....,....6
AA      LKISLISAVSENGVIGSGPDIPWSVKGEQLLFKALTYNQWLLVGRKTFDSMGVLPNRKYA
PHD_sec  EEEEEEEE    EEE           HHHHHHHH     EEE               EEE

        ....,....7....,....8....,....9....,....10.1.,....11.1.,....12.1
AA      VVSKNGISSSNENVLVFPSIENALKELSKVTDHVYVSGGGQIYNSLIEKADIIHLSTVHVEVE
PHD_sec EEEE         EEEE  HHHHHHHHH    EEEEE  HHHHHHHH    EEEEEEEEEEE

        .,....13.
AA      GDIKIPYNA
PHD_sec
```

Figure 3.10 [C]. Predicted secondary structure of the putative translated product of ORF of TR 48 amplicon

Result of conserve domain search for the putative translated product obtained from ORF of TR 48



gnl|CDD|27775 cd00209, DHFR, Dihydrofolate reductase (DHFR).
CD-Length = 158 residues, 81.6% aligned
Score = 137 bits (346), Expect = 5e-34

```
Query:  3    ISLISAVSENGVIGSGPDIPWSVKGEQLLFKALTYNQWLLVGRKTFDSMG--VLPNRKYA  60
Sbjct:  1    ISLIVAVDENGVIGKDNKLPWHLPEDLKHFKKTTTGNPVIMGRKTFESIPRRPLPGRTNI  60

Query:  61   VVSKNGISSSNENVLVFPSIENALKELSKVTDHVYVSGGGQIYNSLIEKADIIHLSTVHV  120
Sbjct:  61   VLSRQLDYQDAEGVEVVHSLEEALELAENTVEEIFVIGGAEIYKQALPYADRLYLTRIHA  120

Query:  121  EVEGDIKIP  129
Sbjct:  121  EFEGDTFFP  129
```

gnl|CDD|22940 pfam00186, DiHfolate_red, Dihydrofolate reductase..
CD-Length = 173 residues, 80.9% aligned
Score = 132 bits (333), Expect = 1e-32

```
Query:  2    KISLISAVSENGVIGSGPDIPWSVKGEQLLFKALTYNQW---LLVGRKTFDSMGV----L  54
Sbjct:  1    MISLIVAVDKNGGIGKDGDLPWRLPNDLKYFKAVTTGTPRNAVIMGRKTWESIPEKFRPL  60

Query:  55   PNRKYAVVSKNGISSSNE-NVLVFPSIENALKELSKVT----DHVYVSGGGQIYNSLIEK  109
Sbjct:  61   PGRLNIVLSRSEDYDAQGDNVVVSSSIEAALDLLAEPPEASIERVFVIGGAQLYAAALPL  120

Query:  110  ADIIHLSTVHVEVEGDIKIP  129
Sbjct:  121  ADRLYLTRIDGEFEGDTFFP  140
```

gnl|CDD|10137 COG0262, FolA, Dihydrofolate reductase [Coenzyme metabolism].
CD-Length = 167 residues, only 79.0% aligned
Score = 103 bits (259), Expect = 6e-24

```
Query:  2    KISLISAVSENGVIGSGPDIPWSVKGEQLLFKALTYNQWLLVGRKTFDSMG----VLPNR  57
Sbjct:  1    MIILIVAVSLDGVIGRDNSLPWHLPEDLAHFKATTLGKPVIMGRKTYESLPGEWRPLPGR  60

Query:  58   KYAVVSKNGISSSNENVLVFPSIENALKELSKVTD-HVYVSGGGQIYNSLIE--KADIIH  114
Sbjct:  61   KNIVLSRNPDLKTEGGVEVVDSIEEALLLLLKEEGEDIFIIGGGELYRQFLPAGLADELI  120

Query:  115  LSTVHVEV-EGD  125
Sbjct:  121  LTIIPVLLGEGD  132
```

```
CCGCGCTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAG 60

GGCAGTCGCCCTAAAACAAAGTTAGACATCATGAGTAACGCAGTACCCGCCGAGATTTCG 120
                                M  S  N  A  V  P  A  E  I  S
GTACAGCTATCACTGGCTCTCAACGCCATCGAGCGTCATCTGGAATCAACGTTGCTGGCC 180
V  Q  L  S  L  A  L  N  A  I  E  R  H  L  E  S  T  L  L  A
GTGCATTTGTACGGCTCTGCACTGGACGGTGGCCTGAAGCCATACAGTGATATTGATTTG 240
V  H  L  Y  G  S  A  L  D  G  G  L  K  P  Y  S  D  I  D  L
CTGGTTACTGTGGCTGCACGGCTCGATGAGACTGTCCGACAAGCCCTGGTCGTAGATCTC 300
L  V  T  V  A  A  R  L  D  E  T  V  R  Q  A  L  V  V  D  L
TTGGAAATTTCTGCCTCCCCTGGCCAAAGTGAAGCTCTCCGCGCCTTGGAAGTTACCATC 360
L  E  I  S  A  S  P  G  Q  S  E  A  L  R  A  L  E  V  T  I
GTCCGTGCATGGTGATGTTGTCCCTTGGCGTTATCCGGCCAGACGGGAACTGCAATTCGG 420
V  R  A  W  *  (3' end of ORF)
GGAGTGGCAGCGTAAGGACATTCTTGGCGGGCATCTTCGAGCCCGCCACAACCGATGTTG 480

ATCTTGGCTATTCTGGCTTAACTTAAAGTAAGCAACATAAGCCTTGCATTTGCAGGGTCG 540
```

Figure 3.11 [A]. Partial nucleotide sequence of CS-PCR product of TR 52 showing ORF and the putative translated product (* indicate stop codon).

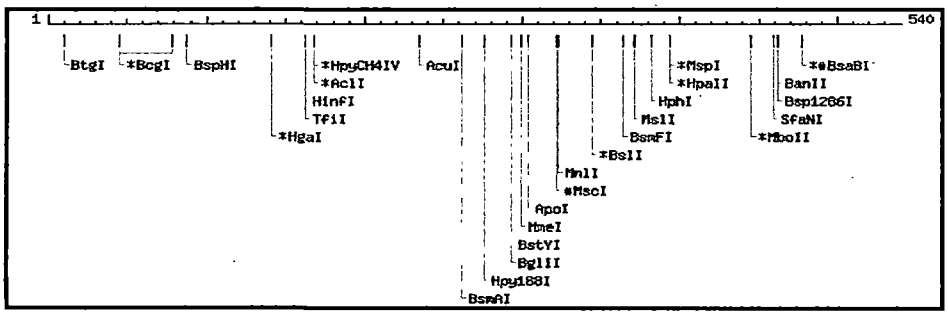
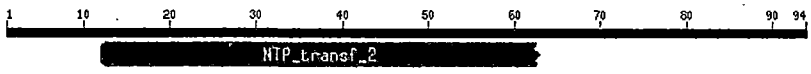


Figure 3.11 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 52. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
          ....,....1....,....2....,....3....,....4....,....5....,....6
AA        MSNAVPAEISVQLSLALNAIERHLESTLLAVHLYGSALDGGLKPYSDIDLLVTVAARLDE
PHD_sec         HHHHHHHHHHHHHHHHHH    EEEEEEE              EEEEE     H

          ....,....7....,....8....,....9....
AA        TVRQALVVDLLEISASPGQSEALRALEVTIVRAW
PHD_sec HHHHHHHHHHHH            EEEEEEEE
```

Figure 3.11 [C]. Predicted secondary structure of the putative translated product of ORF of TR 52 amplicon

Result of conserve domain search for the putative translated product obtained from ORF of TR 52



gnl|CDD|25840 pfam01909, NTP_transf_2, Nucleotidyltransferase domain. Members of this family belong to a large family of nucleotidyltransferases. This family includes kanamycin nucleotidyltransferase (KNTase) which is a plasmid-coded enzyme responsible for some types of bacterial resistance to aminoglycosides. KNTase in-activates antibiotics by catalysing the addition of a nucleotidyl group onto the drug..

CD-Length = 94 residues, only 53.2% aligned
Score = 38.2 bits (88), Expect = 2e-04

```
Query:  12  QLSLALNAIERHLESTLLAVHLYGSALDGGLKPYSDIDLLVTVAARLDETVR  63
Sbjct:  4   ILSKIKEALKELYG--VAEVYLFGSYARGTYLPGSDIDLLVVVPDPVELFLE  53
```

```
GCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGC 60

CCTAAAACAAAGTTAACCTCTGAGGAAGAATTGTGAAACTATCACTAATGGTAGCTATAT  120
                                  V  K  L  S  L  M  V  A  I  S
CGAAGAATGGAGTTATCGGGAATGGCCCTGATATTCCATGGAGTGCCAAAGGTGAACAGC  180
  K  N  G  V  I  G  N  G  P  D  I  P  W  S  A  K  G  E  Q  L
TCCTGTTTAAAGCTATTACCTATAACCAATGGCTGTTGGTTGGACGCAAGACTTTTGAAT  240
  L  F  K  A  I  T  Y  N  Q  W  L  L  V  G  R  K  T  F  E  S
CAATGGGAGCATTACCCAACCGAAAGTATGCGGTCGTAACACGTTCAAGTTTTACATCTG  300
  M  G  A  L  P  N  R  K  Y  A  V  V  T  R  S  S  F  T  S  D
ACAATGAGAACGTATTGATCTTTCCATCAATTAAAGATGCTTTAACCAACCTAAAGAAAA  360
  N  E  N  V  L  I  F  P  S  I  K  D  A  L  T  N  L  K  K  I
TAACGGATCATGTCATTGTTTCAGGTGGTGGGGAGATATACAAAAGCCTGATCGATCAAG  420
  T  D  H  V  I  V  S  G  G  G  E  I  Y  K  S  L  I  D  Q  V
TAGATACACTACATATATCTACAATAGACATCGAGCCGGAAGGTGATGTTTACTTTCCTG  480
  D  T  L  H  I  S  T  I  D  I  E  P  E  G  D  V  Y  F  P  E
AAATCCCCAGCAATTTTAGGCCAGTTTTTACCCAAGACTTCGCCTCTAACATAAATT     537
  I  P  S  N  F  R  P  V  F  T  Q  D  F  A  S  N  I  N
```

Figure 3.12 [A]. Partial nucleotide sequence of CS-PCR product of TR 5801 showing ORF and the putative translated product.

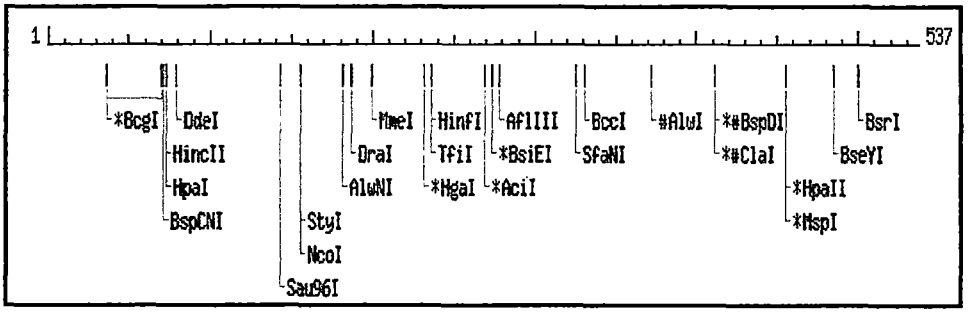


Figure 3.12 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 5801. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
GGCGGAGAAGAAGCACGCCCGACACTTGCTGACGTACAGGAACAGTACTTGCCAAGCGAT 60

TTAGCGCAAGAGTCCGTCACTCCATACATTGCAATGCTGAATGGAGAGCCGATTGGGTAT 120
                                  M  L  N  G  E  P  I  G  Y
GCCCAGTCGTACGTTGCTCTTGGAAGCGGGGACGGATGGTGGGAAGAAGAAACCGATCCA 180
A  Q  S  Y  V  A  L  G  S  G  D  G  W  W  E  E  E  T  D  P
GGAGTACGCGGAATAGACCAGTTACTGGCGAATGCATCACAACTGGGCAAAGGCTTGGGA 240
G  V  R  G  I  D  Q  L  L  A  N  A  S  Q  L  G  K  G  L  G
ACCAAGCTGGTTCGAGCTCTGGTTGAGTTGCTGTTCAATGATCCCGAGGTCACCAAGATC 300
T  K  L  V  R  A  L  V  E  L  L  F  N  D  P  E  V  T  K  I
CAAACGGACCCGTCGCCGAGCAACTTGCGAGCGATCCGATGCTACGAGAAAGCGGGGTTT 360
Q  T  D  P  S  P  S  N  L  R  A  I  R  C  Y  E  K  A  G  F
GAGAGGCAAGGTACCGTAACCACCCCAGATGGTCCAGCCGTGTACATGGTTCAAACACGC 420
E  R  Q  G  T  V  T  T  P  D  G  P  A  V  Y  M  V  Q  T  R
CAGGCATTCGAGCGAACACGCAGTGATGCCTATCCCTTCCATCGAGGGGGACGTCCAAGG 480
Q  A  F  E  R  T  R  S  D  A  Y  P  F  H  R  G  G  R  P  R
GCTGGCGCCCTTGGCGGCCCCTCATGTCAAACGCTAGATGCACTAAGCACATAATTGCTC 540
A  G  A  L  G  G  P  S  C  Q  T  L  D  A  L  S  T  *(3'end of ORF)
ACAGCCAAACTATCAGGTCAAGTCTGCTTA
```

Figure 3.13 [A]. Partial nucleotide sequence of CS-PCR product of TR 5802 showing ORF and the putative translated product (* indicate stop codon).

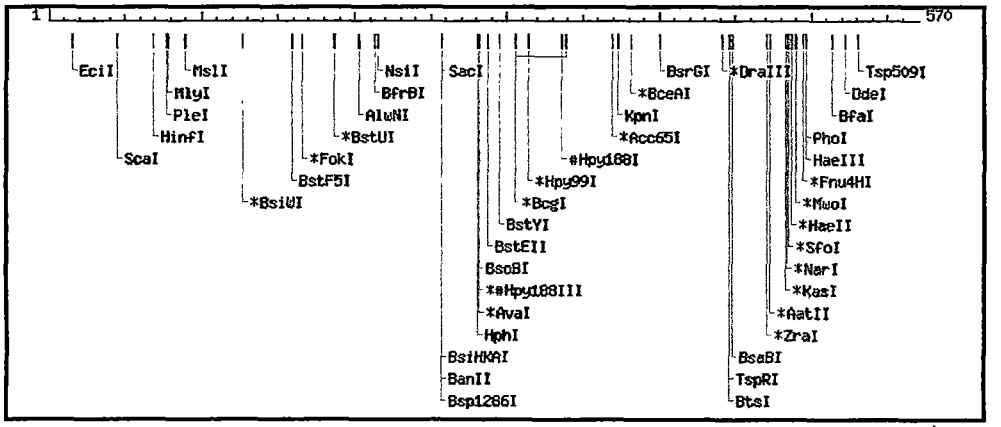


Figure 3.13 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 5802. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
GTTNGCCGTGCGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAG 60

TCGCCCTAAAACAAAGTTGGGCGAACCCGGAGCCTCATTAATTGTTAGCCGTTAAAATTA 120

AGCCCTTTACCAAACCAATACTTATTATGAAAAACACAATACACAGCATCGTGACCAACA 180
                          M  K  N  T  I  H  S  I  V  T  N  S
GCAACGATTCCGTCACACTGCGCCTCATGACTGAGCATGACCTTGCGATGCTCTATGAGT 240
  N  D  S  V  T  L  R  L  M  T  E  H  D  L  A  M  L  Y  E  W
GGCTAAATCGATCTCATATCGTCGAGTGGTGGGGCGGAGAAGAAGCACGCCCGACACTTG 300
  L  N  R  S  H  I  V  E  W  W  G  G  E  E  A  R  P  T  L  A
CTGACGTACAGGAACAGTACTTGCCAAGCGTTTTAGCGCAAGAGTCCGTCACTCCATACA 360
  D  V  Q  E  Q  Y  L  P  S  V  L  A  Q  E  S  V  T  P  Y  I
TTGCAATGCTGAATGGAGAGCCGATTGGGTATGCCCAGTCGTACGTTGCTCTTGGAAGCG 420
  A  M  L  N  G  E  P  I  G  Y  A  Q  S  Y  V  A  L  G  S  G
GGGACGGNNTGTGGGAAGAAGATACCGATCCAGGAGTACGCGGAATAGACCAGTTTATGG 480
  D  G  X  W  E  E  D  T  D  P  G  V  R  G  I  D  Q  F  M  A
CGAATGCATCACAACTGGGCAAAGGCTTGGGAACCAAGTTGGTTCGAGCTCTGGTTGAGT 540
  N  A  S  Q  L  G  K  G  L  G  T  K  L  V  R  A  L  V  E  L
TGCTGTTCAATGATCCCGAGGTCACCAAGATCCAAACGGACCCGTCGCCGAGCAACTTGC 600
  L  F  N  D  P  E  V  T  K  I  Q  T  D  P  S  P  S  N  L  R
GAGCGATCCGATGCTACGAGAAAGCGGGGTTTGAGAGGCAAGGTACCGTAACCACCCCAG 660
  A  I  R  C  Y  E  K  A  G  F  E  R  Q  G  T  V  T  T  P  D
ATGGTCCAGCCGTGTACATGGTTCAAACACGCCAGGCATTCGAGCGAACACGCAGTGATG 720
  G  P  A  V  Y  M  V  Q  T  R  Q  A  F  E  R  T  R  S  D  A
CCTAACCCTTCCATCGAGGGGGACGTCCAAGGGCTGGCGCCCTTGGCCGCCCCTCATGTC 780
  *(3'end of ORF)
AAACGTTAGATGCAGTAAGCACATAATTGTA                               811
```

Figure 3.14 [A]. Complete nucleotide sequence of CS-PCR product of TR 59 showing ORF and the putative translated product (* indicate stop codon).

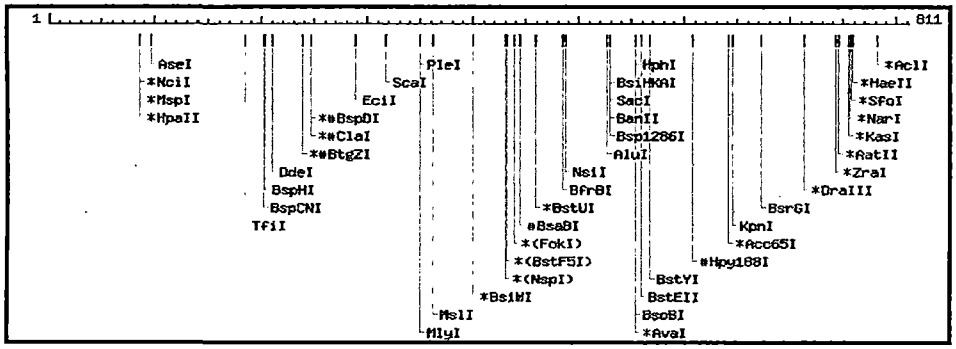
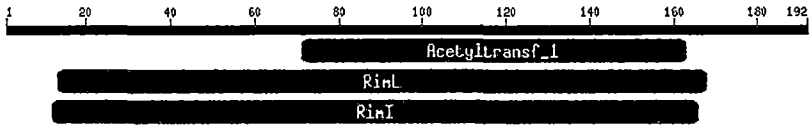


Figure 3.14 [B]. Restriction site analysis of the complete nucleotide sequence generated from cloned CS-PCR product of TR 59. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
          ....,....1....,....2....,....3....,....4....,....5....,....6
AA        MKNTIHSIVTNSNDSVTLRLMTEHDLAMLYEWLNRSHIVEWWGGEEARPTLADVQEQYLP
PHD_sec       EEEEE      EEEEEE      HHHHHHHH                  HHHHHHHHHH

          ....,....7....,....8....,....9....,....10.1.,....11.1.,....12.1
AA        SVLAQESVTPYIAMLNGEPIGYAQSYVALGSGDGXWEEDTDPGVRGIDQFMANASQLGKGLGT
PHD_sec   HH       EEEEEEE    EEEEEEEEEEE              EEEEEEEEE        HHH

          .,....13.1.,....14.1.,....15.1.,....16.1.,....17.1.,....18.1
AA        KLVRALVELLFNDPEVTKIQTDPSPSNLRAIRCYEKAGFERQGTVTTPDGPAVYMVQTRQ
PHD_sec   HHHHHH     EEEEEE     HHHHHHHHH   EEE EEE     EEEEEE  HHHH

          .,....19.
AA        SDA
PHD_sec   HHH
```

Figure 3.14 [C]. Predicted secondary structure of the putative translated product of ORF of TR 59 amplicon

Result of conserve domain search for the putative translated product obtained from ORF of TR 59



gnl|CDD|25558 pfam00583, Acetyltransf_1, Acetyltransferase (GNAT) family. This family contains proteins with N-acetyltransferase functions..

CD-Length = 82 residues, 100.0% aligned
Score = 58.0 bits (140), Expect = 7e-10

```
Query:  71   YIAMLNGEPIGYAQSYVALGSGDGXWEED--TDPGVRGIDQFMANASQLGKGLGTKLVRA  128
Sbjct:  1    LVAEEDGELVGFASLRPIDEEGNVAEIEGLAVDPEYRG-----------KGIGTALLEA  48

Query:  129  LVELLFNDPEVTKIQTDPSPSNLRAIRCYEKAGFE  163
Sbjct:  49   LLEYAR-ELGLKRIELEVLEDNEAAIALYEKLGFK  82
```

gnl|CDD|11381 COG1670, RimL, Acetyltransferases, including N-acetylases of ribosomal proteins [Translation, ribosomal structure and biogenesis].

CD-Length = 187 residues, 81.8% aligned
Score = 69.2 bits (168), Expect = 3e-13

```
Query:  13   NDSVTLRLMTEHDLAMLYEWLNRSHIVEWWGGEEARPTLADVQEQYLPSVLAQESVTPYI  72
Sbjct:  7    TLRLLLREVDLEDLELLAEWANDPEVMLFWWLPPPLTPPTSDEE--LLRLLAEAWEDLGG  64

Query:  73   AMLNGEPIGYAQSYVALGSGDGXWEEDTDPGVRGIDQFMANASQLGKGLGTKLVRALVEL  132
Sbjct:  65   GAFAIELKATGDGELIGVIGLSDIDRAANGDLAEIGYWLDPEYW-GKGYATEALRALLDY  123

Query:  133  LFNDPEVTKIQTDPSPSNLRAIRCYEKAGFERQGTV  168
Sbjct:  124  AFEELGLHRIEATVDPENEASIRVYEKLGFRLEGEL  159
```

gnl|CDD|10329 COG0456, RimI, Acetyltransferases [General function prediction only].

CD-Length = 177 residues, 82.5% aligned
Score = 47.3 bits (111), Expect = 1e-06

```
Query:  12   SNDSVTLRLMTEHDLAMLYEWLNRSHIVEWWGGEEARPTLADVQEQYLPSVLAQESVTPY  71
Sbjct:  8    SEDKVTIREAINKDLLDVALAALEARTFDIRLPWSREYFEKDLTQAPELLLVAETG--GL  65

Query:  72   IAMLNGEPIGYAQSYVALGSGDGXWEED-----TDPGVRGIDQFMANASQLGKGLGTKLV  126
Sbjct:  66   DGLLDGKVVGFLLVRVVDGRPSADHEGHIYNLAVDPEYRG-----------RGIGRALL  113

Query:  127  RALVELLFNDPEVTKIQTDPSPSNLRAIRCYEKAGFERQG  166
Sbjct:  114  DEALERLRERGLADKIVLEVRESNEAAIGLYRKLGFEVVK  153
```

```
GCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGC 60

CCTAAAACAAAGTTGGGCGAACCCGGAGCCTCATTAATTGTTAGCCGTTAAAATTAAGCC 120

CTTTACCAAACCAATACTTATTATGAAAAACACAATACACAGCATCGTGACCAACAGCAA 180
                      M  K  N  T  I  H  S  I  V  T  N  S  N
CGATTCCGTCACACTGCGCCTCATGACTGAGCATGACCTTGCGATGCTCTATGAGTGGCT 240
 D  S  V  T  L  R  L  M  T  E  H  D  L  A  M  L  Y  E  W  L
AAATCGATCTCATATCGTCGAGTGGTGGGGCGGAGAAGAAGCACGCCCGACACTTGCTGA 300
 N  R  S  H  I  V  E  W  W  G  G  E  E  A  R  P  T  L  A  D
CGTACAGGAACAGTACTTGCCAAGCGTTTTAGCGCAAGAGTCCGTCACTCCATACATTGC 360
 V  Q  E  Q  Y  L  P  S  V  L  A  Q  E  S  V  T  P  Y  I  A
AATGCTGAATGGAGAGCCGATTGGGTATGCCCAGTCGTACGTTGCTCTTGGAAGCGGGGA 420
 M  L  N  G  E  P  I  G  Y  A  Q  S  Y  V  A  L  G  S  G  D
CGGATGGTGGGAAGAAGAAACCGATCCAGGAGTACGCGGAATAGACCAGTTACTGGCGAA 480
 G  W  W  E  E  E  T  D  P  G  V  R  G  I  D  Q  L  L  A  N
TGCATCACAACTGGGCAAAGGCTTGGGAACCAAGCTGGTTCGAGCTCTGGTTGAGTTGCT 540
 A  S  Q  L  G  K  G  L  G  T  K  L  V  R  A  L  V  E  L  L
GTTCAATGATCCCGAGGTCACCA                                      563
 F  N  D  P  E  V  T
```

Figure 3.15 [A]. Partial nucleotide sequence of CS-PCR product of TR 63 showing ORF and the putative translated product.

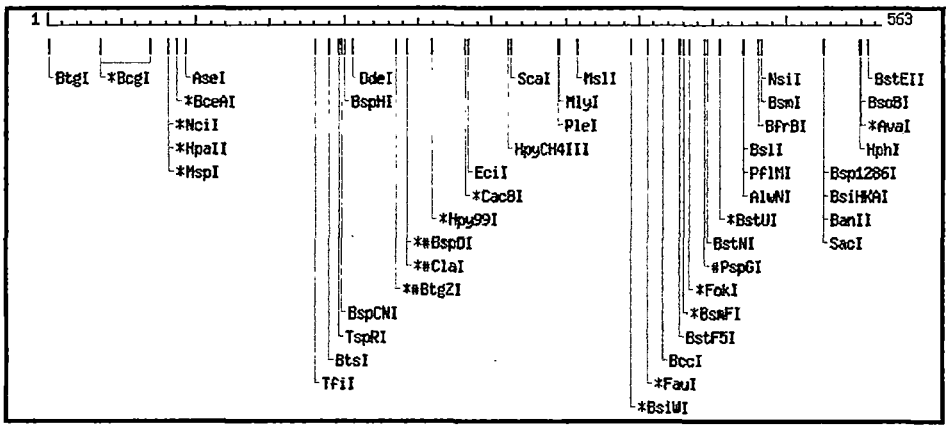


Figure 3.15 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 63. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
NCGCAGCAGGGCAGTCGCCCTAAAAACAAAGTTAACCCGGAACCAAAATTATGAAAGTAT 60
                                                      M  K  V  S
CATTAATGGCTGCAAAAGCGAAAAACGGAGTGATTGGTTGCGGTCCACACATACCCTGGT 120
  L  M  A  A  K  A  K  N  G  V  I  G  C  G  P  H  I  P  W  S
CCGCGAAAGGAGAGCAGCTACTCTTTAAAGCCTTGACGTACAACCAGTGGCTTTTGGTGG 180
  A  K  G  E  Q  L  L  F  K  A  L  T  Y  N  Q  W  L  L  V  G
GCCGCAAGACGTTCGAATCTATGGGAGCACTCCCTAATAGGAAATACGCGGTCGTTACTC 240
  R  K  T  F  E  S  M  G  A  L  P  N  R  K  Y  A  V  V  T  R
GCTCAGCCTGGACGGCCGATAATGACAACGTAATAGTATTCCCGTCGATCGAAGAGGCCA 300
  S  A  W  T  A  D  N  D  N  V  I  V  F  P  S  I  E  E  A  M
TGTACGGGCTGGCTGAACTCACCGATCACGTTATAGTGTCTGGTGGCGGGGAGATTTACA 360
  Y  G  L  A  E  L  T  D  H  V  I  V  S  G  G  G  E  I  Y  R
GAGAAACATTGCCCATGGCCTCTACGCTCCATATATCGACGATTGATATTGAGCCGGAAG 420
  E  T  L  P  M  A  S  T  L  H  I  S  T  I  D  I  E  P  E  G
GAGATGTTTTCTTTCCGAATATTCCCAATACCTTCGAAGTTGTTTTTGAGCAACACTTTA 480
  D  V  F  F  P  N  I  P  N  T  F  E  V  V  F  E  Q  H  F  S
GCTCAAACATTAACTATTGCTATCAAATTTGGCAAAAGGGTTAACAAAGCTATGCAATTG 540
  S  N  I  N  Y  C  Y  Q  I  W  Q  K  G  *(3'end of ORF)
ACGGTAAAAAGCTTCGTTCGCTTCGCTTGCTACGCTTCTTACCGCAATTGATAACGGCGT 600

TAGATGCACTAAGCACATAAT                                         621
```

Figure 3.16 [A]. Partial nucleotide sequence of CS-PCR product of TR 85 showing ORF and the putative translated product (* indicate stop codon).

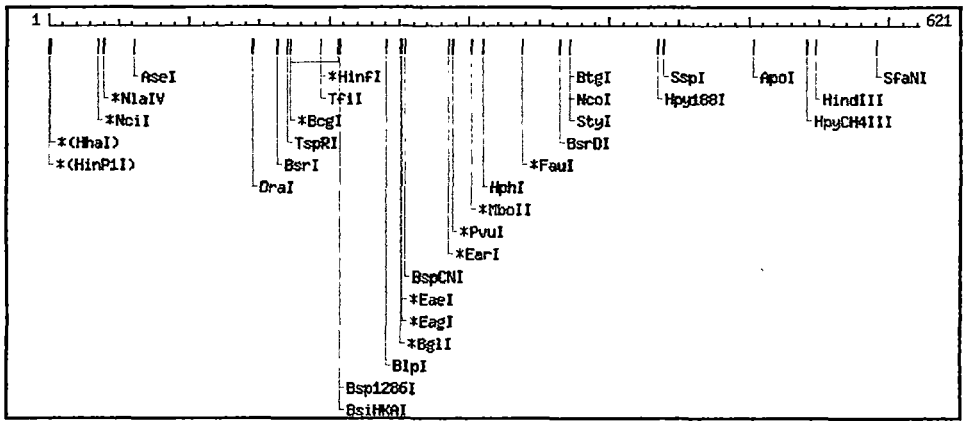


Figure 3.16 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 85. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
GTTNGACGTGTGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAG 60

TCGCCCTAAAACAAAGTTAACCTCTGAGGAAGAATTGTGAAACTATCACTAATGGTAGCT 120
                                        V  K  L  S  L  M  V  A
ATATCGAAGAATGGAGTTATCGGGAATGGCCCTGATATTCCATGGAGTGCCAAAGGTGAA 180
I  S  K  N  G  V  I  G  N  G  P  D  I  P  W  S  A  K  G  E
CAGCTCCTGTTTAAAGCTATTACCTATAACCAATGGCTGTTGGTTGGACGCAAGACTTTT 240
Q  L  L  F  K  A  I  T  Y  N  Q  W  L  L  V  G  R  K  T  F
GAATCAATGGGAGCATTACCCAACCGAAAGTATGCGGTCGTAACACGTTCAAGTTTTACA 300
E  S  M  G  A  L  P  N  R  K  Y  A  V  V  T  R  S  S  F  T
TCTGACAATGAGAACGTATTGATCTTTCCATCAATTAAAGATGCTTTAACCAACCTAAAG 360
S  D  N  E  N  V  L  I  F  P  S  I  K  D  A  L  T  N  L  K
AAAATAACGGATCATGTCATTGTTTCAGGTGGTGGGGAGATATACAAAAGCCTGATCGAT 420
K  I  T  D  H  V  I  V  S  G  G  G  E  I  Y  K  S  L  I  D
CAAGTAGATACACTACATATATCTACAATAGACATCGAGCCGGAAGGTGATGTTTACTTT 480
Q  V  D  T  L  H  I  S  T  I  D  I  E  P  E  G  D  V  Y  F
CCTGAAATCCCCAGCAATTTAGGCCAGTTTTACCAAGACTTCGCCTCTAACATAAATTAT 540
P  E  I  P  S  N  L  G  Q  F  Y  Q  D  F  A  S  N  I  N  Y
AGTTACAAATCTGGCAAAGGGTTACAAGTGGCAGCAACGGATCGCAACCTGTCACGCCTT 600
S  Y  K  S  G  K  G  L  Q  V  A  A  T  D  R  N  L  S  R  L
TGTACCAAAGCCGCGCCAGGTTGCGATCCGCTGTGCAGCGTTAGGCTACATGAAATCGTA 660
C  T  K  A  A  P  G  C  D  P  L  C  S  V  R  L  H  E  I  V
CATTACGAAGCGAATGCACCATGGATAGGAAGAATGAATGCCCAAACCCAAAGTGTGGGA 720
H  Y  E  A  N  A  P  W  I  G  R  M  N  A  Q  T  Q  S  V  G
                     M  D  R  K  N  E  C  P  N  P  K  C  G  K
AGGAAACTCCTGCCTGGCAATCGAGCGGCATGAGCGACAGTTGCCCGCATTTTTTCTGTG 780
R  K  L  L  P  G  N  R  A  A  *(end of ORF 1)
   E  T  P  A  W  Q  S  S  G  M  S  D  S  C  P  H  F  F  C  D
ATACTTGCTCGAATGTAATCCATAGAGAGCAGGACCATGCATTACTGTATGAAAATGAAA 840
   T  C  S  N  V  I  H  R  E  Q  D  H  A  L  L  Y  E  N  E  I
TCAATCAAGAGCTCTTGGATCGAATAGCAGCAACTCTTCCAGATTGCCCTTGCGGGGGTA 900
   N  Q  E  L  L  D  R  I  A  A  T  L  P  D  C  P  C  G  G  R
GGTTTGTTCCTGGTGCAAACCCAAAGTGTCCGAGTTGCAAGACCGAGTACGTGCACCAAT 960
   F  V  P  G  A  N  P  K  C  P  S  C  K  T  E  Y  V  H  Q  W
GGGATGCAGTGAAAAGGTTGAATGTACCTTTTATGCCAATCTTGGATGGTTCCTGCTTGA 1020
   D  A  V  K  R  L  N  V  P  F  M  P  I  L  D  G  S  C  L  I
TTCGAGATAGGCTGTATTCGTATGAAGTATGCATTGGTTCTAAACCAAAATACTGGTGGC 1080
   R  D  R  L  Y  S  Y  E  V  C  I  G  S  K  P  K  Y  W  W  R
GTTTGTTCACAAATGCCTTAACAAGTTTAGGCAAGGGACGCTCCTGACGTCGCGCCCCTG 1140
   L  F  T  N  A  L  T  S  L  G  K  G  R  S  *  (end of ORF 2)
GTAACACATATTGTNAGATATGCCATATGGTCCAGCA                         1177
```

Figure 3.17 [A]. Complete nucleotide sequence of CS-PCR product of TR 90 showing two ORFs and two putative translated products [the putative DHFR protein and hypothetical protein]. The translated sequence of dfrA1 comprises nucleotides 97 to 753. (* indicate stop codons).

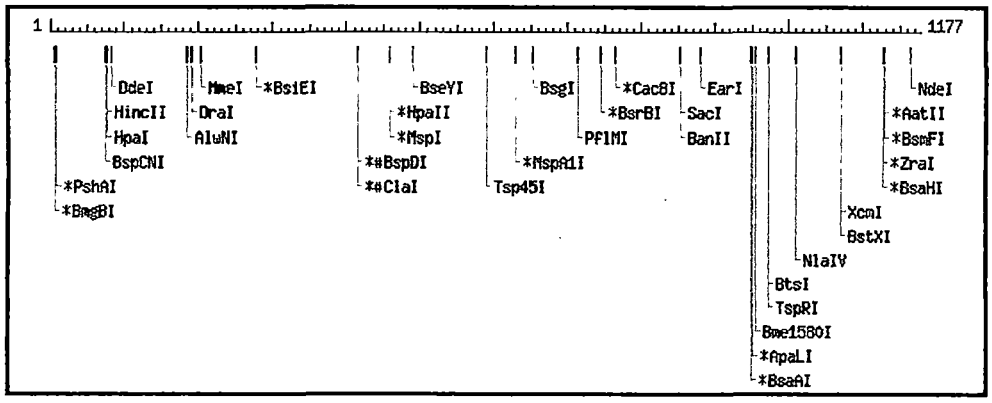
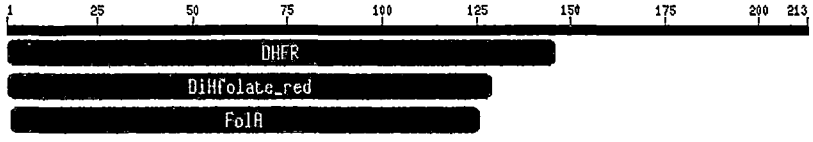


Figure 3.17 [B]. Restriction site analysis of the complete nucleotide sequence generated from cloned CS-PCR product of TR 90. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
         ...,....1...,....2...,....3...,....4...,....5...,....6
AA       MVAISKNGVIGNGPDIPWSAKGEQLLFKAITYNQWLLVGRKTFESMGALPNRKYAVVTRS
PHD_sec   EEEEE              HHHHHHHHHH    EEEEE              EEEEEEE

         ....,....7...,....8...,....9...,....10...,....11...,....12
AA       SFTSDNENVLIFPSIKDALTNLKKITDHVIVSGGGEIYKSLIDQVDTLHISTIDIEPEGDV
PHD_sec          EEEEE HHHHHHHHHH   EEEEE  HHHHHHHHHH EEEEEEEE

         ...,....13...,....14...,....15...,....16...,....17...,....18
AA       YFPEIPSNLGQFYQDFASNINYSYKSGKGLQVAATDRNLSRLCTKAAPGCDPLCSVRLHE
PHD_sec            EEEEE       EEEE   EEEE     HHHHEEE          EEEEEE

         ...,....19...,....20...,....21...,....22
AA       IVHYEANAPWIGRMNAQTQSVGRKLLPGNRAA
PHD_sec  EEEEEE  HHHH      HHH
```

Figure 3.17 [C]. Predicted secondary structure of the putative translated product of ORF 1 of TR 90 amplicon

Result of conserve domain search for the putative translated product obtained from ORF 1 of TR 90




```
gnl|CDD|27775 cd00209, DHFR, Dihydrofolate reductase (DHFR).
              CD-Length = 158 residues,  94.9% aligned
              Score =  152 bits (385), Expect = 3e-38
Query:  1     MVAISKNGVIGNGPDIPWSAKGEQLLFKAITYNQWLLVGRKTFESMG--ALPNRKYAVVT  58
Sbjct:  4     IVAVDENGVIGKDNKLPWHLPEDLKHFKKTTTGNPVIMGRKTFESIPRRPLPGRTNIVLS  63

Query:  59    RSSFTSDNENVLIFPSIKDALTNLKKITDHVIVSGGGEIYKSLIDQVDTLHISTIDIEPE  118
Sbjct:  64    RQLDYQDAEGVEVVHSLEEALELAENTVEEIFVIGGAEIYKQALPYADRLYLTRIHAEFE  123

Query:  119  GDVYFPEIPSNLGQ--FYQDFASNINYSYK  146
Sbjct:  124   GDTFFPEIDESEWELVSEEEVFEEDGYSYT  153


gnl|CDD|22940 pfam00186, DiHfolate_red, Dihydrofolate reductase..
              CD-Length = 173 residues,  81.5% aligned
              Score =  133 bits (335), Expect = 2e-32
Query:  1     MVAISKNGVIGNGPDIPWSAKGEQLLFKAITYNQW---LLVGRKTFESMGA----LPNRK  53
Sbjct:  5     IVAVDKNGGIGKDGDLPWRLPNDLKYFKAVTTGTPRNAVIMGRKTWESIPEKFRPLPGRL  64

Query:  54    YAVVTRS-SFTSDNENVLIFPSIKDALTNLKKIT----DHVIVSGGGEIYKSLIDQVDTL  108
Sbjct:  65    NIVLSRSEDYDAQGDNVVVSSSIEAALDLLAEPPEASIERVFVIGGAQLYAAALPLADRL  124

Query:  109  HISTIDIEPEGDVYFPEIPSN  129
Sbjct:  125  YLTRIDGEFEGDTFFPEIDAD  145


gnl|CDD|10137 COG0262, FolA, Dihydrofolate reductase [Coenzyme metabolism].
              CD-Length = 167 residues,  79.6% aligned
              Score =  108 bits (272), Expect = 4e-25
Query:  2     VAISKNGVIGNGPDIPWSAKGEQLLFKAITYNQWLLVGRKTFESMGA----LPNRKYAVV  57
Sbjct:  6     VAVSLDGVIGRDNSLPWHLPEDLAHFKATTLGKPVIMGRKTYESLPGEWRPLPGRKNIVL  65

Query:  58    TRSSFTSDNENVLIFPSIKDALTNLKKITD-HVIVSGGGEIYKSLIDQ--VDTLHISTID  114
Sbjct:  66    SRNPDLKTEGGVEVVDSIEEALLLLKEEGEDIFIIGGGELYRQFLPAGLADELILTIIP  125

Query:  115  IE-PEGDVYFPEI  126
Sbjct:  126  VLLGEGDTLFPEG  138
```

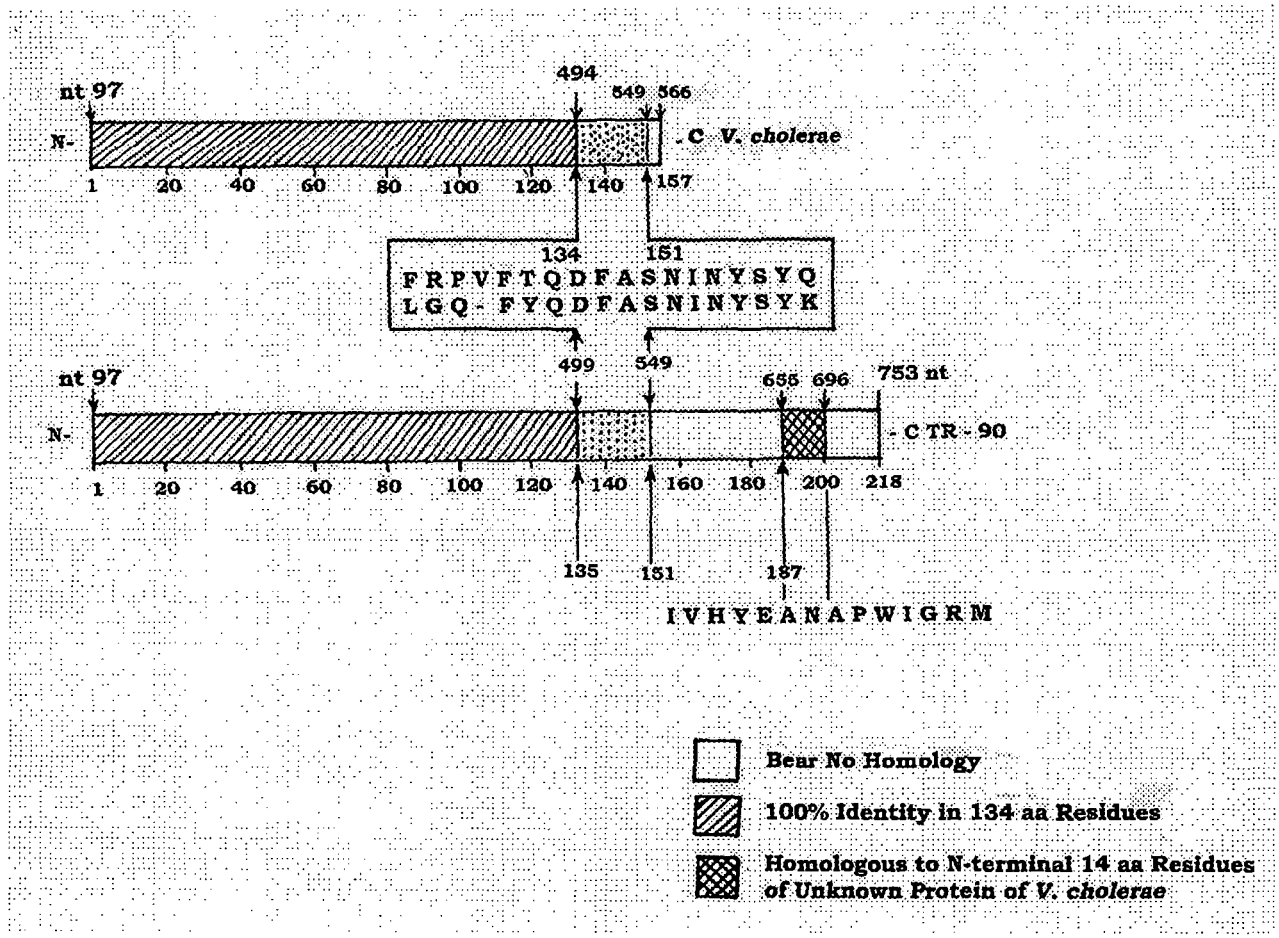


Figure 18A. Comparative analysis of putative peptide sequences of dihydrofolate reductase type I of *V. cholerae* and *Morganella* sp. TR 90.

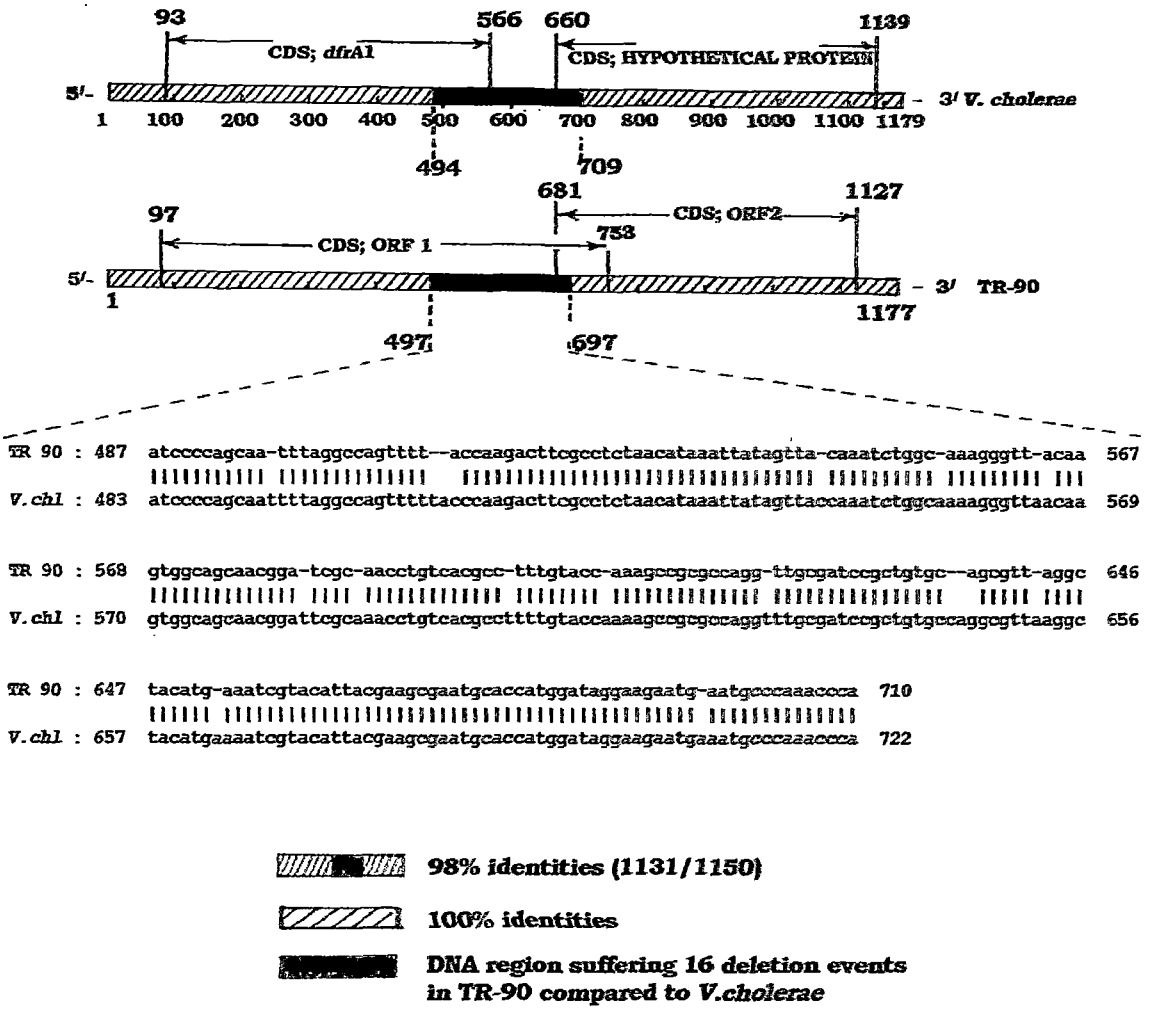


Figure 18B. Probable genesis of dfrA1 ORF in TR 90
A comparative analysis of *dfrA1* sequences of *V. cholerae* and *Morganella* sp. TR 90 showing mutations intensive DNA regions

```
GTTACGACGGCGTCGATGTGTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCA 60

GTCGCCCTAAAACAAAGTTAGCCATTACGGGGGTTGAATTGAAAATTTCATTGATTTCTG 120

CAACGTCAGAAAATGGCGTAATCGGTAATGGCCCTGATATCCCATGGTCAGCAAAAGGTG 180

AGCAGTTACTCTTTAAAGCGCTCACATATAATCAGTGGCTCCTTGTGGGAAGGAAACATT 240

TGACTCTATGGGTGTTCTTCCAAATCGAAAATATGCAGTAGTGTCGAGGAAAGGAATTTC 300
      M  G  V  L  P  N  R  K  Y  A  V  V  S  R  K  G  I  S
AAGCTCAAATGAAAATGTATTAGTCTTTCCTTCAATAGAAATCGCTTTGCAAGAACTATC 360
 S  S  N  E  N  V  L  V  F  P  S  I  E  I  A  L  Q  E  L  S
GAAAATTACAGATCATTTATATGTCTCTGGTGGCGGTCAAATCTACAATAGTCTTATTGA 420
 K  I  T  D  H  L  Y  V  S  G  G  G  Q  I  Y  N  S  L  I  E
AAAAGCAGATATAATTCATTTGTCTACTGTTCACGTTGAGGTTGAAGGTGATATCAATTT 480
 K  A  D  I  I  H  L  S  T  V  H  V  E  V  E  G  D  I  N  F
TCCTAAAATTCCAGAGAATTTCAATTTGGTTTTTGAGCAGTTTTTTTTGTCTAATATAAA 540
 P  K  I  P  E  N  F  N  L  V  F  E  Q  F  F  L  S  N  I  N
TTACACATATCAGATTTGGAAAAAAGGCTAACAAGTCGTTCCAGCACCAGTCGCTGCGCT 600
 Y  T  Y  Q  I  W  K  K  G  *
CCTTGGACAGTTTTTAAGTCGCGGTTTTATGGTTTTGCTGCGCAAAAGTATTCCATAAAA 660

CCACAACTTAAAAACTGCCGCTNAACTCGGCGTTAGATGCACTAAACACATAATTGGC   718
```

Figure 3.19 [A]. Complete nucleotide sequence of CS-PCR product of TR 95 showing ORF and the putative translated product (* indicate stop codon).

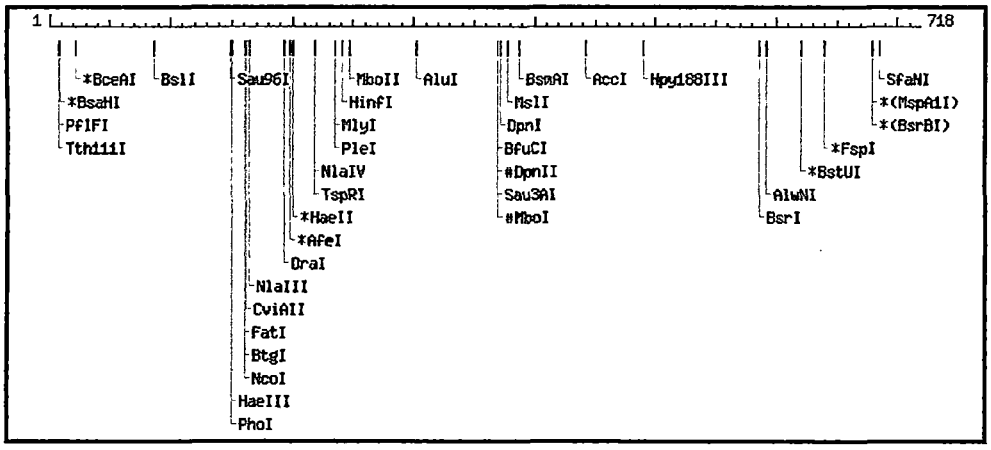
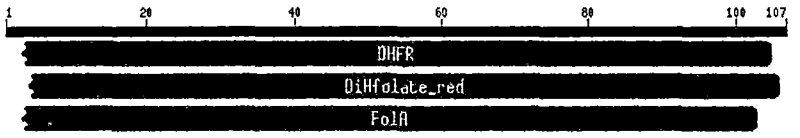


Figure 3.19 [B]. Restriction site analysis of the complete nucleotide sequence generated from cloned CS-PCR product of TR 95. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

```
        ....,....1....,....2....,....3....,....4....,....5....,....6
AA      MGVLPNRKYAVVSRKGISSSNENVLVFPSIEIALQELSKITDHLYVSGGGQIYNSLIEKA
PHD_sec         EEEEEE           EEEE  HHHHHHHHH    EEEEE  HHHHHHHH

        ....,....7....,....8....,....9....,....10.1.,..
AA      DIIHLSTVHVEVEGDINFPKIPENFNLVFEQFFLSNINYTYQIWKKG
PHD_sec  EEEEEEEEEE              EEEEEEEEE  EEEEEEEEE
```

Figure 3.19 [C]. Predicted secondary structure of the putative translated product of ORF of TR 95 amplicon

Result of conserve domain search for the putative translated product obtained from ORF of TR 95



gnl|CDD|27775 cd00209, DHFR, Dihydrofolate reductase (DHFR).
CD-Length = 158 residues, only 67.1% aligned
Score = 87.1 bits (216), Expect = 4e-19

```
Query:  3    VLPNRKYAVVSRKGISSSNENVLVFPSIEIALQELSKITDHLYVSGGGQIYNSLIEKADI  62
Sbjct:  53   PLPGRTNIVLSRQLDYQDAEGVEVVHSLEEALELAENTVEEIFVIGGAEIYKQALPYADR  112

Query:  63   IHLSTVHVEVEGDINFPKI-PENFNLVF--EQFFLSNINYTYQIWK  105
Sbjct:  113  LYLTRIHAEFEGDTFFPEIDESEWELVSEEEVFEEDGYSYTFETYE  158
```

gnl|CDD|22940 pfam00186, DiHfolate_red, Dihydrofolate reductase..
CD-Length = 173 residues, only 65.9% aligned
Score = 96.5 bits (240), Expect = 6e-22

```
Query:  4    LPNRKYAVVSRKGISSSNE-NVLVFPSIEIALQELSKIT----DHLYVSGGGQIYNSLIE  58
Sbjct:  60   LPGRLNIVLSRSEDYDAQGDNVVVSSSIEAALDLLAEPPEASIERVFVIGGAQLYAAALP  119

Query:  59   KADIIHLSTVHVEVEGDINFPKIP-ENFNLVFEQFFL-----SNINYTYQIWKK  106
Sbjct:  120  LADRLYLTRIDGEFEGDTFFPEIDADDWELVSSSEGVEEDKDNGYEYTFETWVK  173
```

gnl|CDD|10137 COG0262, FolA, Dihydrofolate reductase [Coenzyme metabolism].
CD-Length = 167 residues, only 63.5% aligned
Score = 60.7 bits (147), Expect = 4e-11

```
Query:  3    VLPNRKYAVVSRKGISSSNENVLVFPSIEIALQELSKITD-HLYVSGGGQIYNSLIE--K  59
Sbjct:  56   PLPGRKNIVLSRNPDLKTEGGVEVVDSIEEALLLLKEEGEDIFIIGGGELYRQFLPAGL  115

Query:  60   ADIIHLSTVHVEV-EGDINFPKI-PENFNLVFEQFFLSNINYTYQI  103
Sbjct:  116  ADELILTIIPVLLGEGDTLFPEGDPADWELVSSEDADEKGGYFYTF  161
```

```
GACCTTTGAAACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCAC 60
      E  T  S  A  S  P  G  E  S  E  I  L  R  A  V  E  V  T
CATTGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAGCTAAGCGCGAACTGCAATT 120
 I  V  V  H  D  D  I  I  P  W  R  Y  P  A  K  R  E  L  Q  F
TGGAGAATGGCTGCGCAATGACATTCTTGCAGGTATCTTCGAGCCAGCCACGATCGACAT 180
 G  E  W  L  R  N  D  I  L  A  G  I  F  E  P  A  T  I  D  I
TGATCTGGCTATCTTGCTGTCAAAAGCAAGAGGACATAGCGTTGCCTTGGTAGGTCCAGC 240
 D  L  A  I  L  L  S  K  A  R  G  H  S  V  A  L  V  G  P  A
GGCGGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGCTAAATGAAAC 300
 A  E  E  L  F  D  P  V  P  E  Q  D  L  F  E  A  L  N  E  T
CTTAACGCTATGGAACTCGCCGCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTAC 360
 L  T  L  W  N  S  P  P  D  W  A  G  D  E  R  N  V  V  L  T
GTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGC 420
 L  S  R  I  W  Y  S  A  V  T  G  K  I  A  P  K  D  V  A  A
CGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAGACA 480
 D  W  A  M  E  R  L  P  A  Q  Y  Q  P  V  I  L  E  A  R  Q
GGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAAGAATT 540
 A  Y  L  G  Q  E  E  D  R  L  A  S  R  A  D  Q  L  E  E  F
TGTTCACTACGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAATGTCTAACAATTCG 600
 V  H  Y  V  K  G  E  I  T  K  V  V  G  K  *
TTCAAGCCGACGCCGCTTCGCGGCGCGGCTTAACTAAAGCGTTAGATGCACTAAGCACAT 660

AAGTGCACCAGCC                                                  673
```

Figure 3.20 [A]. Partial nucleotide sequence of CS-PCR product of TR 97 showing ORF and the putative translated product (* indicate stop codon).

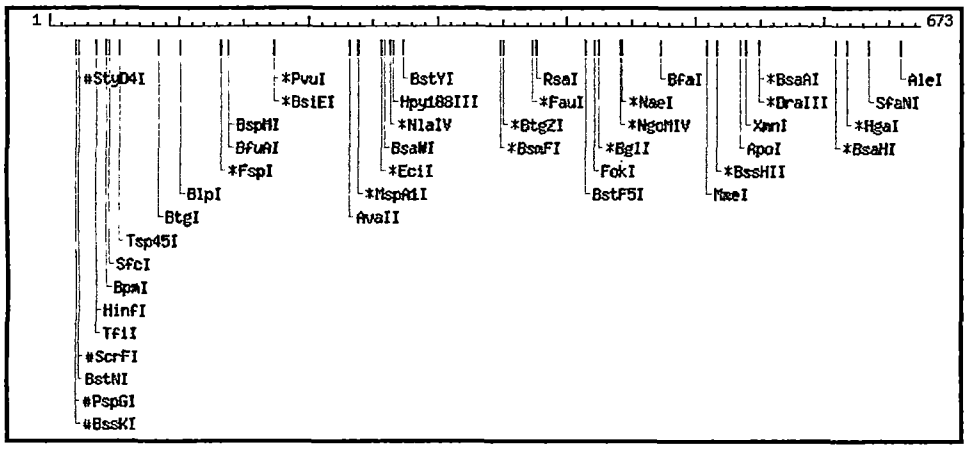


Figure 3.20 [B]. Restriction site analysis of the partial sequence generated from cloned CS-PCR product of TR 97. #, cleavage affected by CpG methylation; and *, cleavage affected by other methylations

Box 3.1. Mutational events due to replication slippage

490 520

1. ***V.cholerae***: 5′ G**CAA|T|TTTA**GGC**CAGTTTT|T|A|**CCC*AAGACTT* 3′

TR 90: 5′ G**CAATTTA**GGC**CAG**T**TTTA**CC*AAGACTT* 3′

494 521

540 571

2. ***V.cholerae***: 5′ A**GTTA|C|CAAA**TCTGG**C|A|AAA**GG**GTT|A|ACAA**GT 3′

TR 90: 5′ A**GTTACAAA**TCTGG**CAAA**GG**GTTACAA**GT 3′[direct repeat]

541 569

581 593

3. ***V.cholerae***: 5′ GGA|T|TCGC|A|AACC 3′

TR 90: 5′ GGATCGCAACC 3′[Quasipalindrome]

579 589

601 619

4. ***V.cholerae***: 5′ C**C|T|TTTG**TAC**C|A|AAAG**CCG 3′

TR 90: 5′ C**CTTTG**TAC**CAAAG**CCG 3′[inverted repeat]

597 613

640 657

5. ***V.cholerae***: 5′ GTGC|C|A|**g**GCGTT|A|AGGCT 3′

TR 90: 5′ GTGC**a**GCGTTAGGCT 3′[Quasipalindrome]

633 647

659 671

6. ***V.cholerae***: 5′ CATG|A|AAATCGTA 3′

TR 90: 5′ CATGAAATCG 3′

649 660

703 712

7. ***V.cholerae***: 5′ GAATG|A|AATG 3′

TR 90: 5′ GAATGAATG 3′[direct repeat]

692 700

3.4 Discussion

River waters are the main receptacle of urban effluents that contain high levels of antibiotic resistant bacteria. We found that culturable antibiotic resistant bacteria were widespread in non-concentrated water samples from the sampling stations of river Torsa (Chapter 1). It has been known for quite some time that cultivable bacteria represent only a small portion of the vast number of bacteria present in the environment and the number that can be enumerated is higher than can be cultivated. Usually, bacteria from a sample are cultivated first. Nutrient typical for known (pathogenic) bacteria or groups of bacteria with similar nutrient requirements are used for this purpose. That is, bacteria, which are able to grow on the nutrient used, are selected and then isolated. In the second step, these isolates are classified according to their susceptibility to a certain antibiotic or a mixture of antibiotics to detect resistant bacteria or multi-resistant ones. For this purpose, antibiotic concentrations up to the milligram per liter range are used. These concentrations are orders of magnitude higher than the environmental ones. The bacteria identified as resistant are investigated further by monitoring the genetic material encoding the resistance found. It has already been accepted that the input of bacteria already resistant following the use of antibiotics in human and veterinary medicine seems to be more important source of resistance genes in the environment. Therefore, more data are needed with respect to the input of already resistant bacteria into the environment for answering the questions related to the spread of resistance genes in nature. Integron gene sequences have been identified as a primary source of resistance genes within microbial populations.

In this study, the results have limitations and must be considered in light of the fact that many aquatic organisms are probably non-culturable. Another important consideration of the study was the exclusion of the gram-positive organisms. Although gram-positive organisms may be important as reservoir of resistance genes, only those antibiotic-resistant isolates shown to be gram-negative were used for further analysis. The ampicillin resistant isolates were predominantly gram-negative. The resistances of ampicillin resistant isolates to other antibiotics, that too on a rational basis, were performed to select MAR isolates for integron-assay.

Ampicillin-resistant MAR isolates were selected for the detection of class 1 integrons. Antibiotic ampicillin has been found to be the most abused antibiotic of this region (Mukherjee *et al.* 2005). The data on the incidence and recovery of antibiotic resistant bateria from the waters of Torsa, all through the year, revealed maximum frequency of ampicillin-resistant (which have been available therapeutically for a long time) bacteria. Earlier reports have shown that about 90% of the integron-positive bacteria were found resistant to ampicillin (White *et al.* 2001). On the basis of all these observations, it was thought that fishing integrons in ampicillin-resistant isolates would be more effective.

Twelve different antibiotics were used for selecting the MAR isolates. Earlier studies have shown that integrons were significantly associated with resistance to certain antibiotics including gentamicin, kanamycin, streptomycin, tobramycin, sulfafurazole, ampicillin, chloramphenicol, and tetracycline. The association of older antibiotics, ampicillin, chloramphenicol and

tetracycline with the presence of an integron was also likely to be the genetic linkage between integrons and conjugative plasmids and transposons. Amikacin and netilmicin were selected with a rationale that genes conferring resistance to newer antibiotics (modern aminoglycosides such as amikacin and netilmicin) are already part of gene cassettes. Antibiotics like cefotaxim, cephalexin and nitrofurantoin, in earlier studies also, revealed insignificant association with integron carriage.

In this study, 100 gram-negative MAR bacteria from river Torsa were examined and found that 40% contained detectable class 1 integron structure. The proportion of strains in this collection of antibiotic resistant bacteria carrying integrons is comparable to that of other studies. For example, 43-75% of antibiotic-resistant clinical bacteria contained class 1 integrons [Levesque *et al.* 1995, Jones *et al.* 1997, Martinez-Freijo *et al.* 1998, Chang *et al.* 2000, Jones *et al.* 2003], while non-selected gram-negative bacteria from an estuarine environment revealed the presence of *int1* 1 gene among 3.6% of the isolates [Rosser and Young 1999]. A study conducted on Rio Grande River (which separates the United States from Mexico) showed that 10% of the *E. coli* isolates were MAR and 13% of MAR contained class 1 integron sequences [Roe *et al.* 2003]. These results, together with those obtained in the present study, indicate that class 1 integrons are widespread in clinical as well as in environmental samples.

The MAR index of the 100 isolates under study ranged from 0.41-1.0. The rationality behind the criterion of selecting MAR isolates with such MAR indices for detecting the class 1 integrons lay in the fact that calculated MAR index exhibited by 82-100% of the class 1 integron bearing isolates was in the range of 0.4-0.9 [Guerra *et al.* 2000, Mazel *et al.* 2000, Thungapathra *et al.* 2002].

The selected 100 gram-negative MAR isolates were primarily differentiated into Pseudomonads and enteric bacteria by oxidase and glucose fermantation tests. 19 out of 100 isolates were oxidase positive and could not ferment glucose. This study did not find any significant difference in the incidence of integrons among the members of enterobacteriaceae (32 out of 81 MAR) and non-enterobacterial members (8 out of 19 MAR). This is similar to the observations of Rosser and Young [1999] who found no significant difference in the incidence of integrons among the coliforms, *Pseudomonas* spp. and *Vibrio* spp. The presence of class 1 integrons in six *Pseudomonas* spp. and two *Acinetobacter* spp. is significant in the context of gene transfer and dissemination of resistance gene cassettes in the environment. An earlier study on two European rivers expressed clinical concern over the development of drug resistance in mesophilic *Aeromonas* spp. [Goni-Urizza *et al.* 2000]. Among the enteric members, the genus *Citrobacter* represented highest proportion of the integron-positive isolates (8 of 40). All these isolates have been characterized as different strains depending on various biochemical characteristics. This study also revealed the presence of class 1 integron in one *Yersinia* sp. TR 63. Presence of class 1 integrons in *Yersinia enterocolitica* was reported for the first time by Soto *et al.* (2003). This is an enteroinvasive bacterium, mainly associated with enterocolitis and less frequently with wide

variety of clinical and immunologic manifestations. Although in this study the copiotrophs were cultured from river water samples, and were deemed to be environmental samples, it is possible that some of the strains, particularly those identified as coliforms, may have originated from human and animal wastes as fluvial waters receive human and wastewater discharges. The physico-chemical and bacteriological investigation on the Torsa River revealed that the ratio of fecal coliform to fecal streptococci ranged from 1.77 to 3.75 and the ratio of free ammonia to albuminoid ammonia ranged from 42 to 761 in the pre-monsoon months indicating pollution by human and agricultural wastes [Bhadra *et al.* 2003].

The observations made from this study, revealed a significant association between integron carriage and higher MAR index. This confirms the results of the earlier studies where a significant relationship between multiresistance and presence of integrons was found (Martinez Freijo *et al.* 1998, Schmitz *et al.* 2001, Leverstein van Hall *et al.* 2003).

Under the PCR conditions used in this study, the size range of inserted gene cassettes varied between 718 and 3200 bp. The different sizes of the gene cassettes inserted between the CS regions of the isolates demonstrate the variable nature of these structures, presumably reflecting differences in the number and type of inserted gene cassettes similar to the observations made by earlier authors (Schmitz 2001). In the present study, integron positive bacteria having single amplicon were the most predominant, although a few (20%) contained two regions of inserted DNA of different sizes (Table 3.3). Yielding two amplicons of different sizes with primers 5′ CS and 3′ CS (Table 3.3) have also been noted by earlier authors where they have confirmed by sequence analysis the co-existence of two distinct integrons carrying different gene cassettes in such isolates [Chang *et al.* 2000, Peters *et al.* 2001, Leverstein van Hall *et al.* 2002]. Additionally, many inserted regions of DNA, indistinguishable with respect to size, were detected in different isolates, as well as from the different strains of the same isolate, which is suggestive of horizontal gene transfer. PCR aimed at determining the proximity of the inserted gene cassettes to an integrase gene demonstrated some sequence heterogeneity among the 5′ CS regions in some of the isolates under study. Five isolates, TR 18, TR 40, TR 68, TR 82 and TR 99, gave a product with the 5′ CS and 3′ CS primers but no product with Int_2F and 3′ CS primer set. Martinez-Freijo *et al.* also made similar type of observations. They found that ten of the seventy integron-positive isolates gave a product with the 5′ CS and 3′ CS primers but no amplification product with Int_2F and 3′ CS primers. In addition, seven isolates gave a product with the 5′ CS and 3′ CS primers but no product with Int_2F and 3′ CS primer set. Several studies have shown the variable nature of the 3′ CS region of an integron with respect to size and genetic structure (Rosser and Young 1999). This study did not attempt to investigate this.

All the isolates positive for class 1 integrons were resistant to multiple antibiotics. However, as the data showed, the presence of class 1 integrons would account for the resistance to only a few drugs. Similar observations have been previously addressed which led to the obvious conclusion that there exist other determinants responsible for mediating

multi resistance in many of the isolates (Thungapathra *et al.* 2002, Chang *et al.* 2000). In this study *dfr* gene cassettes were most frequently found in the variable region of integrons and was often the only cassette present (Table. 3.4). The *dfrA1, dfrA5, dfrA7, dfrA12* and *dfrA17* gene cassettes were found among the isolates included in this study. Taxonomic identification of these isolates revealed that not a single genus but many different genera were involved in the carriage of these gene cassettes (Table 3.3). The *dfrA17* gene cassette was reported for the first time in an *Escherichia coli* isolate EC107 isolated from a clinical urine specimen in Taiwan (Chang *et al.* 2000). All of these five different *dfr* cassettes (*dfrA1, dfrA5, dfrA7, dfrA12* and *dfrA17*) were identified among the urinary *E. coli* isolates from the last two decades in Korea (Yu *et al.* 2004). According to the results of their study, the *E. coli* isolates that carried *dfrA17* associated with class 1 integrons were found to be phylogenetically unrelated, indicating that *dfrA17* was widely distributed in the different clones of *E.coli*. The prevalence of *dfrA17* was mainly due to the horizontal transfer of class 1 integrons through conjugative plasmids. The *dfrA17* gene cassette characterized in this study was found in *Serratia* sp. and did not produce 100% sequence identity with that of the previously described one from *E.coli*. The observation of this study also revealed the possibility of the spread of *dfrA17* gene cassette through conjugal transfer. Thungapathra *et al.* (2002) reported the presence of *dfrA1, dfrA5, dfrA12* gene cassettes among *Vibrio cholerae* isolates belonging to non-O1 and non-O139 serogroups isolated during 1997 to 1998 in Calcutta, India.

Both adenyl transferase (*aadA1* and *aadA6*) and acetyl transferase gene cassettes (*aac-6'-Ib*) were found to be present at roughly the same frequency. These cassettes have also been reported in various bacterial genera from different regions of the world.

The sequence analysis of the amplicon from *Morganella* sp. TR 90 enabled to characterize a novel ORF of 218 amino acids of which the first 152 amino acids bear 95% identities with the normal 157 amino acid long *dfrA1* protein. The sequence alignment data of *Morganella* sp. TR 90 and *V. cholerae* class 1 integron gene cassettes (AJ698460 vs AF455254) has helped to identify sequence contexts that may be important for promoting specific types of *in vivo* mutations in the gene cassettes leading to an extended version of DFR A1 protein(Figure 18.A) Accumulation and fixation of multiple mutations resulting in sequence divergence can ultimately lead to the emergence of a novel protein. Exquisite and insightful sequence analysis of mutation in genes e (lysozyme), *rII, ac,* and others have shown that many frameshift mutations, base substitutions, deletions or duplication can be explained by STREISINGERs slipped mispairing model (Streisinger *et al.* 1966) and its variations (Drake and Ripley 1994). Slipped mispairing that is potentially mutagenic can occur between simple or complicated repeats, in imperfect palindromes, or between nearly homologous sequences located at considerable distances, i.e, called ectopic sequences (Ripley 1990). It has been demonstrated directly that misaligned bases in potential stem-loop structures with no direct repeats are also hot spots of frameshift mutations (Ripley 1990). Earlier

authors have suggested that bulged mispairing of any single-stranded DNA with any complementary sequence, regardless of the reasons for single strandedness, can generate mispaired heteroduplex and thus be mutagenic (Drake and Ripley 1994).

Several studies have addressed the role of integrons in the spread of antibiotic resistance genes by lateral gene transfer (Leverstein van-Hall *et al.* 2002, Chen *et al.* 2004, Aubert *et al.* 2004). The spread of resistance genes is greatly enhanced following the movement of conjugative plasmids containing integrons among different bacterial species. In this study, class 1 integron borne *dfrA5, dfrA12* and *dfrA17* gene cassettes was found to be located on the conjugative plasmids of some MAR bacterial isolates. These isolates belonged to the genus *Escherichia, Citrobacter* and *Serratia* respectively. The earlier authors have also confirmed the location of class 1 integrons on conjugative plasmids by the similar PCR assay (Chen *et al.* 2004). The horizontal transfer of such plasmids enabled the rapid spread of the gene cassettes among a wide variety of bacterial species (Girlich *et al.* 2001, White *et al.* 2001, Fluit *et al.* 2004). It has also been shown that horizontal transfer of integron-carrying elements play a dominant role in the development of multiresistance by Enterobacteriaceae independent of species or isolate origin (Leverstein van-Hall *et al.* 2003). The diversity of species carrying integrons means that stage suitable for the storage, acquisition, rearrangement and expression of gene cassettes may be prevalent in nature. The same cassettes can spread in different bacterial species, in different human or animal hosts and in different areas of the globe. Such observations emphasize the apprehension that human, animal and environmental microbial ecosystems are inextricably entangled, with antibiotic resistance readily surpassing ecological boundaries and thinning out widely.

3.5. Conclusion

The results of this study revealed that integron structures were prevalent in the gram-negative, multiple-antibiotic-resistant copiotrophic bacterial population of the Torsa river. A similar study from a South African river has been reported recently [Biyela *et al.* 2004]. The present study indicated that River Torsa of India is contaminated with class 1 integron bearing copiotrophic bacterial population and therefore it may act as a reservoir as well as a medium for the spread of bacterial antibiotic resistance genes in the environment. The river water isolates carrying class 1 integron borne gene cassettes in their conjugative plasmids is a significant observation in the context of gene transfer and the dissemination of resistance genes in the environment. The predominance of selected aminoglycoside and trimethoprim resistance gene cassettes may, however, indicate that some gene cassettes are more stably incorporated within the integron structures than others. This study has also demonstrated that integron typing can be a useful tool for studying the dissemination of resistance genes among gram-negative bacteria. The accumulation of resistance genes by integrons is one explanation for the emergences of multiple resistant strains of Enterobacteriaceae and pseudomonads. Integron-driven gene capture is likely to be an important factor in the more general process of horizontal gene transfer in the evolution of bacterial

genomes. Since many integrons possess more than one antibiotic resistance-conferring gene cassette and are often located on genetic elements that carry the other resistance determinants, selection for one antimicrobial resistance determinant selects for many. The study of integrons and their associated gene cassettes could provide information about antibiotics, which should be used most carefully to prevent further accumulation of resistance. Further studies are needed, however, to determine whether antibiotic policies or other measures can halt or lower the amount of horizontal transfer in different environmental settings.

3.6 Summary of chapter 3

The presence of class 1 integrons, in multiple- antibiotic- resistant (MAR), gram-negative, copiotrophic bacteria of River Torsa was detected by using a polymerase chain reaction (PCR) based screening method. Of 100 isolates that were resistant to at least five of the twelve antibiotics tested, 40 carried class 1 integrons, with inserted DNA regions of 0.7 to 3.2 kb. Carriage of integrons in strains of higher MAR index was found to be statistically significant. DNA sequencing was used to identify the genetic content of the integron-variable regions. In addition to the identification of gene cassettes, *dfrA1, dfrA5, dfrA7, dfrA17* and a variant of *dfrA12* for trimethoprim; *aac-(6')-Ib* for amikacin and tobramycin; and *aadA1* and *aadA6* for streptomycin and spectinomycin resistance, a novel ORF, predicted from sequence of *Morganella* sp. TR 90, bearing homology with *Vibrio cholerae dfrA1* gene cassette, has been characterized. The residence of class 1 integrons on conjugative R plasmids was confirmed in six integron positive isolates. Gene cassettes identified from the transconjugants were found to be 100% identical in sequence context with the *dfrA5, dfrA12* and *dfrA17* cassettes identified previously from the respective donor isolates. All of these cassettes conferred resistance to trimethoprim. The results of the study has revealed that the waters of river Torsa in West Bengal of India, could eventually become a major reservoir for antibiotic resistant microbes due to dissemination of antibiotic resistance genes carried by class 1 integrons through conjugal transfer.

3.7 References

Adrian, P.V., Plesis, M.D., Klugman, K.P. and Amyes, S.G.B. 1998. New Trimethoprim-resistant Dihydrofolate reductase cassette, *dfr* XV inserted in a class I integron. *Antimicrob. Agents. Chemother.* **42** : 2221-2224.

Ahmed, A. M., T. Nakagawa., E. Arakawa., T. Ramamurthy., S. Shinoda, and T. Shimamoto. 2004. New aminoglycoside acetyltransferase gene, *aac(3')-Id*, in a class 1 integron from a multiresistant strain of *Vibrio fluvalis* isolated from an infant aged 6 months. J. Antimicrob.Chemother. **53** : 947-951.

Altschul, S. F., T. L. Madden., A. A. Schaffer., J. Zhang., Z. Zhang., W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new gwnwration of protein database search programs. Nucleic. Acid. Res. **25** : 3389-3402.

Amita., S. Roy Chowdhury., M. Thungapathra., T. Ramamurthy., G. B. Nair, and A. Ghosh. 2003. Class I integrons and SXT elements in E1 Tor strains isolated before and after 1992 *Vibrio cholerae* O139 outbreak, Calcutta, India. Emerg. Infect. Dis. **9** : 500-502.

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

Aubert, D., D. Girlich., T. Naas., S. Nagarajan, and P. Nordmann. 2004. Functional and structural characterization of the genetic environment of the extended-spectrum β-lactamase bla_{VEB-1} gene from a *Pseudomonas aeruginosa* isolate obtained in India. Antimicrob. Agents. Chemother. **48** : 3284-3290.

Bairoch, A., P. Bucher, and K. Hofmann. 1997. The PROSITE database, its status in 1997. Nucleic. Acid. Res. **25** : 217-221.

Bass, L., C. A. Liebert., M. D. Lee., A. O. Summers., D. G. White., S. G. Thayer and J. J. Maurer. 1999. Incidence and characterization of integrons, genetic elements mediating multiple-drug resistance, in avian *Escherichia coli*. Antimicrob. Agents. Chemother. **43** : 2925-2929.

Bhadra, B., S. Mukherjee., R. Chakraborty, and A. K. Nanda. 2003. Physicochemical and bacteriological investigation on the river Torsa of North Bengal. J. Environ. Biol. **24** : 125-133.

Bissonnette, L., S. Champetier., J. P. Buisson, and P. H. Roy. 1991. Characterization of the nonenzymatic chloramphenicol resistance (*cmlA*) gene of the In4 integron of *Tn1696*: similarity of the product to transmembrane transport proteins. J. Bacteriol. **173** : 4493-4502.

Biyela, P. T., J. Lin, and C. C. Bezuidenhout. 2004. The role of aquatic ecosystems as reservoirs of antibiotic resistant bacteria and antibiotic resistance genes, Water. Sci. Technol. **50** : 45-50.

Brenner, D. J. in: Noel R. Kreig (Ed.), *Bergey's manual of systematic bacteriology*, 1984, pp. 409-516 (Vol. I.)

Birnboim, H. C, and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic. Acid. Res. **7**: 1513-1523.

Bunny, K. L., R. M. Hall, and H. W. Stokes. 1995. New mobile gene cassettes containing an aminoglycoside resistance gene, *aacA7*, and a chloramphenicol resistance gene, *catB3*, in an integron in pBWH301. Antimicrob. Agents. Chemother. **39** : 686-693.

Carattoli, A., L. Villa., C. Pezzella., E. Bordi, and P. Visca. 2001. Expanding drug resistance through integron acquisition by IncFI plasmids of *Salmonella enterica* Typhimurium. Emerg. Infect. Dis. **7**: 444-447.

Chen, S., S. Zhao., D. G. White., C. M. Shroeder., R. Lu., H. Yang., P. F. McDermott., S. Ayers, and J. Meng. 2004. Characterization of multiple-antimicrobial-resistant *Salmonella* serovars isolated from retail meats. Appl. Environ. Microbiol. **70** : 1-7.

Chang, C. Y., L. L. Chang., Y. H. Chang., T. M. Lee, and S. F. Chang. 2000. Characterisation of drug resistance gene cassettes associated with class 1 integrons in clinical isolates of *Escherichia coli* from Taiwan, ROC. J. Med. Microbiol. **49** : 1097-1102.

Collis, C. M., G. Grammaticopoulos., J. Briton., H. W. Stokes, and R. M. Hall. 1993. Site specific insertion of gene cassettes into integrons. Mol. Microbiol. **9** : 41-52.

Collis, C. M, and R. M. Hall. 1992. Gene cassettes from the insert regions of integrons are excised as a covalently closed circles. Mol. Microbiol. **6** : 2875-2885.

Collis, C. M, and R. M. Hall. 1995. Expression of antibiotic resistance genes in the integrated

cassettes of integrons. Antimicrob. Agents. Chemother. **39** : 155-162.

Correia, M., F. Boavida., F. Grosso., M. J. Salgado., L. M. Lito., J. M. Cristino., S. Mendo, and A. Duarte. 2003. Molecular characterization of a class 3 integron in *Klebsiella pneumoniae*. Antimicrob. Agents. Chemother. **47** : 2838-2843.

Davis, L. G., L. D. Dibner, and J. F. Battery. 1986. Basic Methods in Molecular Biology. Elsevier Science Publishing Co., Inc., ISBN 0-444-01082-3, Amsterdam, The Netherlands.

Doublet, B., D. Boyd., M. R. Mulvey., and A. Cloeckaert. 2005. The *Salmonella* genomic island 1 is an integrative mobilizable element. Mol. Microbiol. **55** : 1911-1924.

Drake, J. W, and L. S. Ripley. 1994. Mutagenesis, pp. 98-124 in Molecular Biology of Bacteriophage T4, edited by J. D. Karam, J. W. Drake, K. N. Kreuzer, G. Mosig, D. H. Hall, F. A. Eiserling, L. W. Black, E. K. Spicer, E. Kutter, K. Carlson and E. S. Miller. ASM Press, Washington, DC.

Fluit, A. C, and F. -J. Schmitz. 2004. Resistance integrons and super integrons. Clin Microbiol Infect. **10** : 272-288.

Fluit, A. C, and F. -J. Schmitz. 1999. Class 1 integrons, gene cassettes, mobility, and epidemiology. Eur. J. Clin. Microbiol. Infect. Dis. **18** : 761-770.

Gebreyes, W. A, and C. Altier. 2002. Molecular characterization of multi-drug resistant Salmonella enterica subsp. Enterica serovar Typhimurium isolates from swine. J. Clin. Microbiol. **40** : 2813-2822.

Girlich, D., L. Poirel., A. Leelaporn., A. Karim, and C. Tribuddharat. 2001. Molecular epidemiology of the integron-located VEB-1 extended spectrum beta-lactamase in nosocomial enterobacterial isolates in Bangkok, Thailand. J. Clin. Microbiol. **39** : 175-182.

Goldstein, C., M. D. Lee., S. Sanchez., C. Hudson., B. Phillips., B. Register., M. Grady., C. Liebert., A. O. Summers., D. G. White, and J.J. Maurer. 2001. Incidence of class1 and class 2 integrases in clinical and commensal bacteria from livestock, companion animals, and exotics. Antimicrob. Agents. Chemother. **45** : 723-726.

Goni-Urizza, M., L. Pineau., M. Capdepuy., C. Roques., P. Caumette, and C. Quentin. 2000. Antimicrobial resistance of mesophilic *Aeromonas spp*. isolated from two European rivers, J. Antimicrob. Chemother. **46** : 297-301.

Gonzalez, G., K. Sossa., H. Bello., M. Dominguez., S. Mella, and R. Zemelman. 1998. Presence of integrons in isolates of different biotypes of *Acinetobacter baumanii* from Chilean hospitals. FEMS. Microbiol. Lett. **161**: 125-128.

Guerineau, F., L. Brooks, and P. Mullineaux. 1990. Expression of the sulfonamide resistance gene from plasmid R46. Plasmid. **23** : 34-41.

Guerra, B., S. Soto., S. Cal, and M. C. Mendoza. 2000. Antimicrobial resistance and spread of class 1 integrons among *Salmonella* serotypes. Antimicrob. Agents. Chemother. **44** : 2166-2169.

Guex, N, and M. C. Peitsch. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comperative protein modeling. Electrophoresis. **18** : 2714-2723.

Gupta, S. P. 2001. Statistical Methods, Sultan Chand and Sons, Educational Publishers, New Delhi. pp. 953-1003.

Hall, R. M, and C. M. Collis. 1995. Mobile gene cassettes and integrons: capture and spread by site- specific recombination. Mol Microbiol. **15** : 593-600.

Hall, R. M. 1997. Mobile gene cassettes and integrons: moving antibiotic resistance genes in gram-negative bacteria. Ciba. Found. Symp. **207** : 192-202.

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

Hall, R. M, and H. W. Stokes. 1993. Integrons: novel DNA elements which capture genes by site-specific recombination. Genetica. **90** :115-132.

Hanson, K., O. Skold, and L. Sindstrom. 1997. Non palindromic attI sites of integrons are capable of site-specific recombination with one another and with secondary targets. Mol. Microbiol. **26** : 441-453.

Heir, E., B-A. Lindstedt., T. M. Leegaard., E. Gjernes, and G. K. Kapperud. 2004. Prevalence and characterization of integrons in blood culture Enterobacteriaceae and gastrointestinal *Escherichia coli* in Norway and reporting of a novel class 1 integron-located lincosamide resistance gene. Annals of Clinical Microbiology and Antimicrobials. **3** : 12.

Jones, L. A., C. J. McIver., W. D. Rawlinson, and P. A. White. 2003. Polymerase chain reaction screening for integrons can be used to complement resistance surveillance programs. Commun Dis Intell. **27** : S103-S110.

Jones, M. E., E. Peters., A. M. Weeraink., A. Fluit, and J. Verhoef. 1997. Widespread occurrence of integrons causing multiple antibiotic resistances in bacteria. Lancet. **349** : 1742-1743.

Kang, H. Y., Y. S. Jeong., J. Y. Oh., S. H. Tae., C. H. Choi., D. C. Moon., W. K. Lee., Y. C. Lee., S. Y. Seol., D. T. Cho, and J. C. Lee. 2005. Characterization of antimicrobial resistance and class 1 integrons found in *Escherichia coli* isolates from humans and animals in Korea. J. Antimicrob. Chemother. **55** : 639-644.

Kazama, H., H. Hamashima., M. Sasatsu, and T. Arai. 1998. Distribution of the antiseptic-resistance gene qacEΔ1 in gram-positive bacteria. FEMS. Microbiol. Lett. **165**: 295-299.

Krumperman, P. H. 1983. Multiple antibiotic resistance indexing of *Escherichia coli* to identify high risk sources of fecal contamination of foods. Appl. Environ. Microbiol. **46** : 165-170.

L'Abee-Lund, T. M, and H. Sorum. 2001. Class 1 integrons mediate antibiotic resistance in the fish pathogen *Aeromonas salmonicida* worldwide. Microb. Drug. Resist. **7** : 263-272.

Laraki, N., M. Galleni., I. Tham., M.L. Riccio., G. Amicosante., J. Frere, and G. M. Rossolini. 1999. Structure of In31, a bla –containing *Pseudomonas aeruginosa* integron phyletically related to In5, which carries an unusual array of gene cassettes. Antimicrob. Agents. Chemother. **43** : 890-901.

Leverstein–van Hall, M. A., A. T. A. Box., H. E. M. Blok., A. Paauw., A. C. Fluit, and J. Verhoef. 2002. Evidence of extensive interspecies transfer of integron-mediated antimicrobial resistance genes among multidrug-resistant enterobacteriaceae in a clinical setting. J. Infect. Dis. **186** : 49-56.

Leverstein-van Hall, M. A., H. E. M. Blok., A. R. Donders., A. Paauw., A. C. Fluit, and J. Verhoef. 2003. Multidrug resistance among enterobacteriaceae is strongly associated with the presence of integrons and is independent of species and isolate origin. J. Infect. Dis. **187** : 251-259.

Levesque, C., L. Piche., C. Larose, and P. H. Roy. 1995. PCR Mapping of integrons reveals several novel combinations of resistance genes. Antimicrob. Agents. Chemother. **39** : 185-91.

Lindstedt, B. A., E. Heir., I. Nygard, and G. Kapperud. 2003. Characterization of class 1 integrons in clinical strains of *Salmonella enterica* subsp. *enterica* serovars Typhimurium and Enteritidis from Norwegian hospitals. J. Med. Microbiol. **52** : 141-149.

Martinez, E, and F. de la Cruz. 1990. Genetic elements involved in *Tn21* site specific integration, a novel mechanism for the dissemination of antibiotic resistance genes. EMBO J. 9 : 1275-1281.

Marchler-Bauer, A, and S. H. Bryant. 2004. "CD-search : protein domain annotations on the fly". Nucleic. Acids. Res. **32** : W327-W331.

Martinez Freizo, P., A. C. Fluit., F. J. Schmitz., V. S. C. Grek., J. Verhoef, and M. E. Jones. 1998. Class I integrons in gram negative isolates from different European hospitals and association with decreased susceptibility to multiple antibiotic compounds. J. Antimicrob. Chemother. **42** : 689-696.

Martinez Freizo, P., A. C. Fluit., F. J. Schmitz., J. Verhoef, and M. E. Jones. 1999. Many class I integrons comprise distinct stable structures occurring in different species of enterobacteriaceae isolated from widespread geographic regions in Europe. Antimicrob. Agents. Chemother. **43** : 686-689.

Mazel, D., B. Dychinco., V. A. Webb, and J. Davies. 2000. Antibiotic resistance in ECOR collection: integrons and identification of a novel *aad* gene. Antimicrob. Agents. Chemother. **44** : 1568-1574.

Morabito, S., R. Tozzoli., A. Caprioli., H. Karch, and A. Carattoli. 2002. Detection and characterization of class 1 integrons in enterohemorrhagic *Escherichia coli*. Microb. Drug. Resist. **8** : 85-91.

Mukherjee, S., B. Bhadra., R. Chakraborty., A. Gurung., S. Some, and R. Chakraborty. 2005. Unregulated use of antibiotics in Siliguri city vis-à-vis occurrence of MAR bacteria in community waste water and river Mahananda, and their potential for resistance gene transfer. J. Environ. Biol. **26:** 229-238.

Naas, T., Y. Mikami., T. Imai., L. Poirel, and P. Nordmann. 2001. Characterization of In53, a class 1 plasmid- and composite transposon-located integron of *Escherichia coli* which

characterizes an unusual array of gene cassettes. J. Bacteriol. **183** : 235-249.

Nandi, S., J. J. Maurer., C. Hofacre, and A. O. Summers. 2004. Gram positive bacteria are a major reservoir of class 1 antibiotic resistance integrons in poultry litter. PNAS. **101**: 7118-7122.

Neild, B. S., A. J. Holmes., M. R. Gillings., G. D. Reechia., B. C. Mabbutt, K. M. H. Nevalainen and H. W. Stokes. 2001. Recovery of new integron classes from environmental DNA. FEMS Microbiol. Lett. **195** : 59-65.

Nunes-Duby, S. E., H. J. Kwon., R. S. Tirumalai., T. Ellenberger, and A. Landy. 1998. Similarities and differences among 105 members of the Int family of site specific recombinases. Nucleic Acids Res. **26** : 391-406.

O'Halloran, F., B. Lucey., B. Cryan., T. Buckley, and S. Fanning. 2004. Molecular characterization of class 1 integrons from Irish thermophilic *Campylobacter* spp. J. Antimicrob. Chemother. **53** : 952-957.

Ouellette, M, and P. H. Roy. 1987. Homology of ORFs from Tn2603 and from R46 to site specific recombinases. Nucleic Acids Res. **15** : 10055.

Park, J. C., J. C. Lee., J. Y. Oh., Y. W. Jeong., J. W. Cho., H. S. Joo., W. K. Lee, and W. B. Lee. 2003. Antibiotic selective pressure for the maintenance of antibiotic resistant genes in coliform bacteria isolated from the aquatic environment. Water. Sci. Technol. **47** : 249-253.

Partridge, S. R., G. D. Reechia., C. Scaramuzzi., C. M. Collis., H. W. Stokes, and R. M. Hall. 2000. Definition of the *att1* site of class 1 integrons. Microbiology. **146** : 2855-2864.

Partridge, S. R., H. J. Brown., H. W. Stokes, and R. M. Hall. 2001. Transposons *Tn1696* and *Tn21* and their integrons In4 and In2 have independent origins. Antimicrob. Agents. Chemother. **45** : 1263-1270.

Partridge, S. R., G. D. Reechia., H. W. Stokes, and R. M. Hall. 2001. Family of class 1 integrons related to In4 from *Tn1696*. Antimicrob. Agents. Chemother. **45** : 3014-3020.

Partridge, S. R., H. J. Brown, and R. M. Hall. 2002. Characterization and movement of the class 1 integron known as *Tn2521* and *Tn1405*. Antimicrob. Agents. Chemother. **46** : 1288-1294.

Paulsen, I. T., T.G. Littlejohn., P. Randstrom., L. Sundstrom., O. Skold., G. Swedberg, and R. A. Skurray. 1993. The 3′ conserved segment of integrons contains a gene associated with multidrug resistance to antiseptics and disinfectants. Antimicrob. Agents. Chemother. **37** : 761-768.

Peters, E. D. J., M. A. Leverstein–van Hall., A.T. A. Box., J. Verhoef., and A. C. Fluit. 2001. Novel gene cassettes and integrons, Antimicrob. Agents. Chemother. **45** : 2961-2964.

Ploy, M. C., F. Denis., P. Courvalin, and T. Lambert. 2000. Molecular characterization of integrons in *Acinetobacter baumannii*: description of a hybrid class 2 integron. Antimicrob. Agents. Chemother. **44** : 2684-2688.

Reechia, G. D., H. W. Stokes, and R. M. Hall. 1994. Characterisation of specific and secondary recombination sites recognized by the integron DNA integrase. Nucleic Acids. Res. **22** : 2071-2078.

Ripley, L. S. 1990. Frameshift mutations-determinants of specificity. Ann. Rev. Genet. **24** : 189-213.

Roe, M. T., E. Vega, and S. D. Pillai. 2003. Antimicrobial resistance markers of class 1 and class 2 integron bearing *Escherichia coli* from irrigation water and sediments, Emerg. Infect. Dis. **9** : 822-826.

Rosser, S. J, and H. K. Young. 1999. Identification and characterization of class 1 integrons in bacteria from an aquatic environment. J. Antimicrob. Chemother. **44** : 11-18.

Rowe-Magnus, D. A, and D. Mazel. 1999. Resistance gene capture. Current Opinion in Microbiology. **2** : 483-488.

Rowe-Magnus, D. A., A. M. Guerout., P. Ploncard., B. Dychinco., J. Davies, and D. Mazel. 2001. The evolutionary history of chromosomal super integrons provides an ancestry for multi-resistant integrons. Proc. Natl. Acad. Sci. USA. **98** : 652-657.

Rowe-Magnus, D. A, and D. Mazel. 2002. The role of integrons in antibiotic resistance gene capture. Int. J. Med. Microbiol. 292 : 1-11.

Sallen, B., A. Rajoharison., S. Desvarenne, and C. Mabilat. 1995. Molecular epidemiology of integron associated antibiotic resistance genes in clinical isolates of Enterobacteriaceae. Microb. Drug. Resist. 1: 195-202.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: A laboratory manual. 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.