Detection of class 1 integrons and molecular characterization of their gene cassettes in facultative oligotrophic bacteria of River Mahananda

2.1. Background

The localization of antibiotic-resistance determinants onto mobile entities such as plasmids and transposons readily explained the phenomenon of lateral gene transfer. On further examination of these resistance determinants located on mobile plasmids and transposons, a novel type of genetic element which later on termed as "INTEGRON" was discovered (Stokes and Hall 1989; Hall and Collis, 1995). Integrons are one of the genetic elements involved in the adaptation of bacteria. Integrons are found in plasmids, chromosomes and transposons. It consists of a gene encoding a site specific recombinase along with a specific site for recombination capable of recruiting gene cassettes. 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 integrons were not repeats, and (ii) the integrons 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 (Martinez and de la Cruz, 1990; Ouellette and Roy 1987, Stokes and Hall 1989, Sundstrom *et al.* 1988).

Integrons, a natural gene expression system: The present day definition of integrons was 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 (termed as 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; Hall and Collis, 98, Rowe-Magnus *et al.* 1999, Rowe-Magnus *et al.*, 2002). It is these gene cassettes that encode mostly (but not always) 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 *IS*26, *IS*1999, *IS*2000 and *IS*6100. Most frequently occurring IS element at the 3' - end of integrons is *IS*6100.

2.1.1. Types and structural organization of Integrons

Integrons have been categorized in two major groups: "chromosomal integrons" and "mobile integrons". Chromosomal integrons are located on the chromosome of bacterial species; *insilico* study demonstrates that 17% of the bacterial genomes exhibit such genetic arrangements (Cambray *et al.*, 2010). Chromosomal integrons are frequently described in bacteria isolated from marine or terrestrial ecosystems, for example *Vibrio* spp. and *Xanthomonas* spp. The chromosomal integrons have also been termed "super-integrons" (SIs) as they can carry large numbers (>100) of cassettes

that mainly encode proteins with unknown functions. While mobile integrons are located on mobile genetic elements such as transposons and plasmids, which promote their dissemination among bacteria but they are not self-movable. Mobile integrons contain a very limited numbers of gene cassettes (Naas *et al.*, 2001).

On the basis of amino-acid sequence identity of the IntI protein mobile integrons have been classified into several classes and of them classes 1-3 are more commonly studied (Hochhut *et al.*, 2001; Cambray *et al.*, 2010; Xu *et al.*, 2011; Stalder *et al.*, 2012). The gene cassettes (mobile DNA elements which is found in free circular form when not incorporated in integron platform) within these mobile integrons often encode antibiotic-resistance determinants. Integrase gene (*intI*) is responsible for the synthesis of the enzyme, integrase, which helps in the integration and excision of gene cassettes in the integron. The enzyme recognizes site-specific recombination sites for integration. Approximately 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) while class 2 integron are found in association of *Tn7* transposon family (Hall and Stokes, 1993). Arakawa *et al.* (1995) identified class 3 integron platform on a large transferable plasmid in a *Serratia marcescens* strain. 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 Gammaproteobacterial 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. Boucher *et al.* (2007) statistically analyzed about 603 completely or partially sequenced genomes, and revealed that 9% of them contained integrons, indicating that integrons are much more widely distributed than previously expected.

2.1.2. Class 1 integron

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; Lindstedt *et al.*, 2003; Betteridge *et al.*, 2011) and environmental isolates (Gebreyes *et al.*, 2002; Chen *et al.*, 2004; Nandi *et al.*, 2004; Nardelli *et al.*, 2012; Stalder *et al.*, 2012). 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 (*sul1*) and an open reading frame

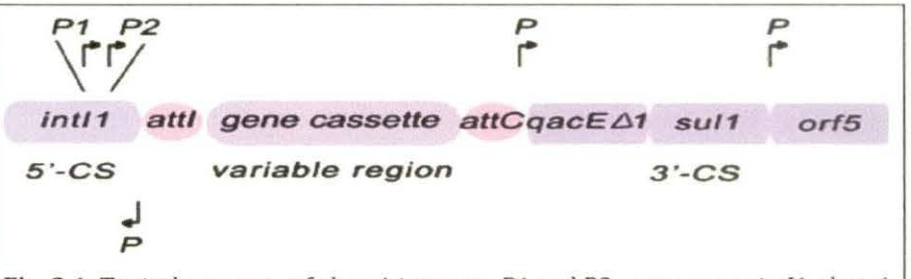


Fig. 2.1: Typical structure of class 1 integron. P1 and P2 , promoters; *intI1*, class 1 integron integrase gene; *attI*, integron associated attachment site; *attC*, cassette associated attachment site; *qacEΔ1*, Quaternary ammonium sulphate resistance gene; *sul1*, sulfonamide resistance gene, *orf5*, unknown function; CS, conserved segment. [Fig. adapted from Xu *et al.*, 2011].

(ORF5) of unknown function (Fig 2.1). Integron integrases, as described earlier, being members of tyrosine recombinase family, possesses four invariant amino-acid residues (RHRY) in conserved motifs (termed as boxes I and II) and mediates recombination between the *attI* site and secondary target called an *attC* site (also known as 59 base element or 59be site). Multiple alignments of integron integrases with tyrosine recombinases have 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 *intI* seem to play a role in DNA binding (K219) while others are implicated in the recombination activity (Reechia *et al.*, 1994).

The *att*C (59 be) is the recocmbination site, present in the gene cassette, recognized by the integron integrase (Stokes *et al.*, 1997). The *attC* sites comprise a family of diverse sequences which are not highly conserved and vary considerably in size from 57 to 141 bp (Collis *et al.*, 1993; Collis and Hall, 1995; Recchia and Hall, 1995; Mazel *et al.*, 2000). The *attC* region consists of four essential sites called 1R, 2R, 1L and 2L, with 1R and 2R as part of the RH consensus sequence, and 1L and 2L as part of the LH consensus sequence (Francia *et al.*, 1999; Fluit and Schmitz, 2004). The similarities of the *attC* sites are primarily restricted to their boundaries, which correspond to the inverse core site as RYYYAAC (R = Purine, Y = Pyrimidine) and the core site as GTTRRRY (Collis and Hall, 1995; Stokes *et al.*, 1997). The *attC* sites are generally associated with a single ORF in a structure termed gene cassettes, which are not necessarily observed in integrons, but once integrated they become part of the integron (Fluit and Schmitz, 1999). 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). Recombination process occurs between G and 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 (*attC*) 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 site 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 lengths 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.

2.1.3. Gene cassette

The antibiotic resistance genes that captured by integrons, are located on the mobile circular DNA element called as gene cassette. The cassettes usually consist of a promoter less coding gene sequence (most commonly an antibiotic resistance gene but may carry non-antibiotic resistance genes) 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 *attC* sites comprise a family of diverse sequences which are not highly conserved and vary considerably in size from 57 to 141 bp (Collis *et al.*, 1993; Collis and Hall, 1995; Recchia and Hall, 1995; Mazel *et al.*, 2000). The *attC* region consists of four essential sites called 1R, 2R, 1L and 2L, with 1R and 2R as part of the RH consensus sequence, and 1L and 2L as part of the LH consensus sequence (Francia *et al.*, 1999; Fluit and Schmitz, 2004). The similarities of the *attC* sites are primarily restricted to their boundaries, which correspond to the inverse core site as RYYYAAC and the core site as GTTRRRY (Collis and Hall, 1995; Stokes *et al.*, 1997). The integrase, apart from catalyzing the recombination between *attI*1 and *attC* sites, also facilitates excision 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 (Pc) region located 214 bases from the inner boundary of the 5' conserved segment (Collis and Hall, 1995) of integron. In fact, the Pc 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. Few 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 (i) Gene cassettes are discrete genetic elements that may exist as free, circular, non- replicating DNA molecules when moving from one genetic location to other, but are normally found as linear sequence that contribute part of a larger DNA molecule such as plasmid or bacterial chromosome, (ii) Gene cassettes normally contain only a single gene and an additional short sequence, called 59 be, that functions as a specific recombination site, (iii) The cassettes are small, normally ranging from 500-1000 bp, (iv) The genes carried on gene cassettes usually lack promoters and are expressed from a common promoter of integron, (v) A cassette may carry two genes; these are 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 (Fig. 2.2 and Fig. 2.3). 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. Mostly gene cassettes identified between the resistance integrons are those conferring resistance to antibiotics. The list is of antibiotic resistance genes, covering most classes of antimicrobials structured as gene cassettes are growing day by day (Fluit *et al.*, 2004; Gillings *et al.*, 2008a).

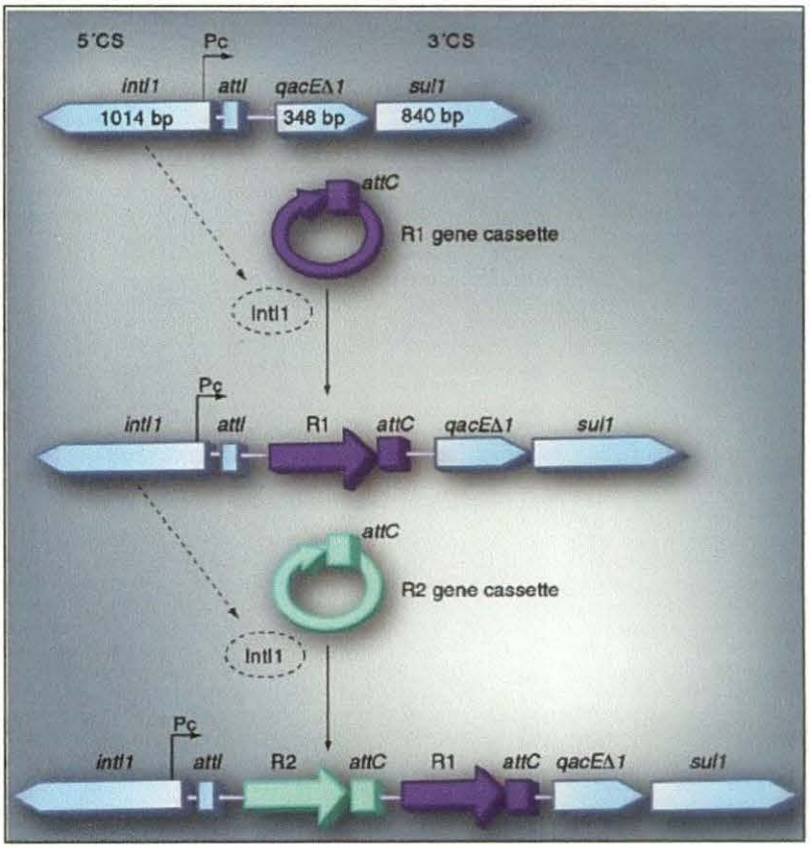


Fig. 2.2: The process of sequential integration of gene cassettes within integron platform following site-specific recombination between *attI* site on integron and *attC* site on gene cassettes. (i) Acquirement of the first gene cassette by using *attI* site. (ii) Acquirement of the second gene cassette by using the same *attI* site. (iii) Integron structure with two inserted gene cassettes. Symbols, IntI: Integrase; *attC* (59be): Recognition site for the integrase; Pc: Promoter. [Fig. adapted from Zhao and Hu, 2011].

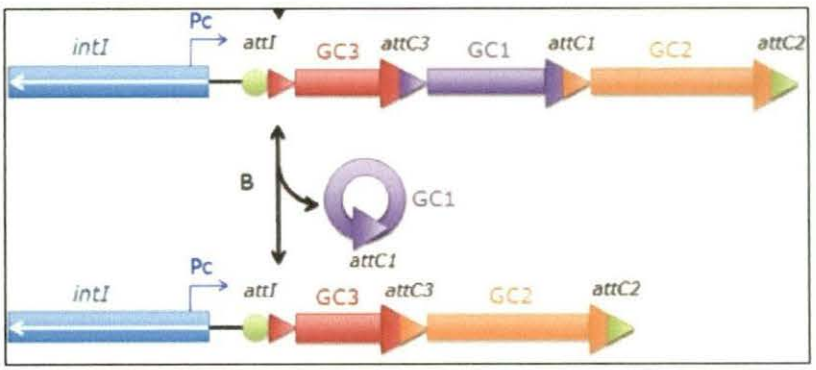


Fig. 2.3: The process of excision of gene cassette (GC) by site-specific recombination mechanism. The IntI1 protein catalyzes the insertion and excision (ii) of the GC in the integron. GC excision preferentially occurs between two *attC* sites. The GC1 is excised following the recombination between the two *attC1* and *attC3* sites. Symbols Pc: gene cassette promoter; *attI*: integron recombination site; *attC1*, *attC2*, and *attC3*: *attC* GC recombination sites; *intI*: the integrase gene; GC1, GC2, GC3 are the gene cassettes, and arrows indicate the direction of coding sequences. [Figure adapted from Stalder *et al.*, 2012].

Dissemination of gene cassettes: Spread 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 (i) Mobilization of individual cassette by the integron encoded integrase. (ii) Movement taking place when the integron containing the cassette relocates-probably by targeted

transposition. (iii) Dissemination of large transposons such as *Tn21* carrying integrons and (iv) Movement of conjugative plasmids containing integrons among different bacterial species.

2.1.4. Integron Epidemiology

Integrons in clinical samples: 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. Several studies have demonstrated broad 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. It was found 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 occurrences 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 isolates were found carrying integron (Gonzalez *et al.,* 1998). In Taiwan, 52% of the tested clinical isolates of *Escherichia coli* showed presence of class 1 integrons (Chang *et al.,* 2000). Ploy *et al.* (2000) analyzed twenty *Acinetobacter baumannii* strains containing integron. 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 noticed by Guerra *et al.* (2000). 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 by Lindstedt *et al.* (2003), reported presence of class 1 integrons in clinical strains of *Salmonella enterica* subsp. *enterica* serovars *Typhimurium* and *Enteritidis,* from Norwegian hospitals. 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. Put latest data. 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, 72 *Escherichia coli* strains 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), fish (L'Abée-Lund and Sørum, 2001; Nawaz *et al.,* 2010). 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'Halloran *et al.* (2004). Their presences have also been described in isolates obtained from farm and nonfarm animals such as swine, dogs, cats, horses, deers (Kadlec and Schwarz, 2008; Yang *et al.,* 2010), zoo animals (Mazel *et al.,* 2000; Ahmed *et al.,* 2007; Sato *et al.,* 2009) and Birds (Nebbia *et al.,* 2008). Recently class 1 integron detected in 67% of the total Gram-negative bacteria collected from animal samples like caecum,

faeces, intestine, oviduct, skin or vaginal fluid of cats, dogs, jennies, deer, eagles, emus, fox, owls, snake, squirrels, chickens, gilthead breams, rabbits and trouts (Domingues *et al.*, 2010). Recently in Portugal the *intI1* gene was identified in 72.5% of *S. Enterica* isolates recovered from Bísaro pigs and wild boars (Caleja *et al.*, 2011). Integron sequences have also been detected from soils (Neild *et al.*, 2001; Heringa *et al.*, 2010). Using culture independent methods, they detected several novel integrase genes.

Integrons in aquatic environment: Predominance of class 1 integrons have also been evidenced in various environmental samples including fish farms, lakes, estuary, irrigation water sources, and other aquatic environments. 109 gram-negative bacteria out of 3000 bacteria isolated from an estuarine environment analyzed, showed presence of *inti1* gene of class 1 integrons. 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 were 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). Study related to coliform bacteria isolated from the aquatic environment led the detection of class 1 integrons, and 24% of the isolates 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 multiple antibiotic resistant (MAR).

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 glumaticum* 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. In Guangzhou, China, Shi *et al.* (2006) isolated 46 gram positive bacteria from clinical specimen. PCR analyses of said isolated revealed the presence of class 1 integron in all the isolates. Very recently the consequences of surveillance and investigation on antibiotic resistance resulted in identification of class 1 integron in Methicillin-resistant staphylococci (MRS), the most prevalent pathogens that cause nosocomial infections throughout the world and labeled as a "super bug" in antibiotic resistance (Xu *et al.*, 2011).

2.1.5. Integron study in India

Above data suggested that integrons are evenly distributed among all members of bacterial families, and usually contribute antimicrobial resistance to them. Few studies in India have shown the prevalence and distribution of integrons from distinct clinical settings. Investigation of multidrug

resistance in *Vibrio cholerae* strains isolated from Calcutta, India, has established the association of antibiotic resistance phenotype with the presence of integrons (Thungapathra *et al.* 2002, Amita *et al.* 2003; Shi *et al.*, 2006). The gene cassettes identified in *V. cholerae* strains were *dfrA1, dfrA15, dfrA5, dfrA12, aac (6')-Ib, aadA1, aadA2* and *ereA2*. A new antibiotic resistance gene *aac(3)-Id* (aminoglycoside acetyltransferase gene) harbored in class 1 integron was reported in 2004 from a clinical isolate, *Vibrio fluvalis* H-O8942 (Ahmed *et al.* 2004). The isolate was also found to carry a class 1 integron borne aminoglycoside adenyltransferase gene, *aadA7*. In the same year a study conducted on *E. coli* isolates obtained from UTI (urinary tract infection) of pregnant woman at the Christian Medical College Hospital, Vellore, India, showed that all the tested isolates carried integrons (Mathai *et al.*, 2004). 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. A study conducted on copiotrophic bacteria isolated from river Torsa, India, showed that out of 100 MAR copiotrophic isolate, 40 were bearing the class 1 integron and the gene cassattes detected within class 1 integron were: *dfrA1, dfrA5, dfrA7, dfrA17,* a variant of *dfrA12, aac(6')-Ib, aadA1* and *aadA6* along with one novel ORF (Mukherjee and Chakraborty, 2006). Shahid (2010) reported *sul1* type integron in 32.5% *Citrobacter* spp. isolated clinical samples. A study on clinical *Klebsiella pneumoniae* strains revealed 92% carriage of integrons (Bhattacharjee *et al.*, 2010). The bla_{CTX-M} and bla_{TEM} resistance markers associated with the class 1 integron in ESBL-producing uropathogenic *E. coli* isolates have been reported from India (Akram *et al.*, 2011). The most prevalent gene cassette array detected in the above said study was "*aadA5-dfrA17-dfrA7*". In a recent investigation dfrA7 and *dfrA1-orfC* cassettes along with an adjacent dihydropteroate synthetase gene were detected in class 1 integron of nontyphoidal *Salmonella* isolates recovered from 40 seafood samples (Deekshit *et al.*, 2012). The same study has also demonstrated the presence of silent *catA1* gene in phenotypically sensitive bacteria.

In the present study, a total of 2188 purified facultative oligotrophic bacterial isolates were examined for the presence of class 1 integrons, using a highly reproducible PCR strategy. Selection of 2188 facultatively oligotrophic bacterial colonies/isolates was random (without any human biasness), because research randomizer tool was used which generated a random set from non-selective plates serving as an universal set of numbers (each isolated colonies bearing a number) irrespective of the knowledge of the phenotype (sensitive, single or multiple-antibiotic-resistant). Presence of class 1 integrons was confirmed in only 89 Gram negative isolates and one Gram positive isolate. The cloning and sequencing of amplicons enabled to characterize the nature of gene cassettes present in ninety integron-positive oligotrophic bacterial strains. The descriptions of ORFs encoding proteins for antibiotic resistance and other unrelated functions have been presented in this chapter.

2.2. Materials and Methods

2.2.1. Random selection of facultatively oligotrophic isolates and detection of class 1 integrons in them

2188 discrete unique numbered (generated using research randomizer; www.researchrandomizer.org) facultatively oligotrophic bacterial colony were selected from the oligotrophic bacterial population cultured on NPA (nutrient-poor agar) medium from 90 water samples, collected in three consecutive years (2007-2009). Each selected bacterial colony was purified and master plate (containing R2A agar) was constructed. Master plate constructed with purified facultative oligotrophic bacterial colonies were replicated on R2A agar plate amended with defined concentration of each antibiotic tested (please see chapter 1), and antibiotic-resistance-profiles were prepared on the basis of their responses to 12 different antibiotics. All the facultative isolates (2188 isolates) of second generation master plate were screened for the presence of class 1 integron by using polymerase chain reaction (PCR) method. The antibiotic resistance profile (ARP)

of integron positive isolates were cross checked after conducting PCR-based screening for detection of class 1 integrons. The isolates bearing class 1 integrons were then subjected for phenotypic/ and or genotypic analyses for identification of tentative genus.

2.2.2. Antibiotic resistance determination of oligotrophic bacteria

Antibiotic resistance was determined according to the method described in chapter 1.

2.2.3. Polymerase chain reaction (PCR)-based screening for detection of class 1 integrons

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. Primer Int_2F (5′-TCTCGGGTAACATCAAGG-3′), specific for the 3′ region of the integrase gene was also 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 location of primer Int2F is ~ 600 bp upstream to the position of 5′ CS primer (as it is expected to generate an additional length of ~ 600 bp compared to the corresponding product length of the CS-PCR)

The DNA templates for PCR were prepared as described by Levesque *et al.* (1995). Bacterial isolates were grown in 5 mL R2A broth in absence (for sensitive isolates)/ or presence of a selective antibiotic (in case of resistant ones) at 30 °C overnight, then 0.2 mL of the overnight grown culture was added to 0.8 mL of double distilled/or de-ionized water and boiled for 10 minutes. The bacterial suspension was then centrifuged at 6000 rpm for 5 minutes to remove cell debris and the supernatant obtained was used as the template.

Another method (this study) was also developed which was equally efficient to method described above for DNA preparation. In this method, 2-4 colonies grown for 24-72 hrs (depending on bacterial growth) on solid agar medium (R2A /or Luria agar) were scraped and transferred to the 1.5 mL microcentrifuge tube containing 0.2 mL sterile distilled water. The bacterial cells were suspended by tapping or by vortexing. The tubes containing cell suspension were closed tightly and placed on floater. The whole set was then placed in a beaker or plate containing double distilled water. The external water served as medium to transfer heat (conduction) to cell suspension present inside tubes it also prevent the bumping of bacterial cell suspension inside microcentrifuge tube. The cell suspension was then exposed to microwave for 90 seconds at 800 watt. The lysed cells were cooled to room temperature and centrifuged at 8000 rpm for 2 minutes. The supernatant used as DNA template in PCR reaction to amplify integron or 16S rRNA gene sequences.

PCR amplification was performed in 25 µL reaction volume. Each 25 µL PCR mix contains; 1.5 µL of 10mM dNTP mix, 2.5 µL of 10X buffer containing 15 mM $MgCl_2$, 6.0 pmol of each forward and reverse primer, 3 µL of template DNA and 0.8U *Taq* DNA Polymerase. Taq DNA polymerase was added after 12 minutes at 94°C (hot start method). However addition of Taq DNA polymerase along with the reaction mixture was equally efficient as in hot start method. The PCR was done in a GenAmp PCR system (Applied Biosystems) and 96 well DNA engine (BioRed, USA). The thermal programme followed for the amplification: denaturation for 1 minute at 94 °C, annealing at 55 °C for 1 minute, and 2-3 minutes of extension at 72 °C for a total of 30 cycles with final extension of 7 minutes at 72 °C. In all reactions, PCR set up containing whole cell DNA of *Morgnella* sp. TR90 (integron positive from our laboratory) was used positive control. The genomic DNA of *Escherichia coli* JM109 and sterile distilled water were used as negative controls. The chemical used in PCR reaction were purchased from Genei, India and Sigma Aldrich.

2.2.4. Gel electrophoresis and size determination of the CS-PCR products

After PCR amplification, 5-7 µL of the reaction mix containing amplified product were electrophoresed on 1% (w/v) agarose (SRL, India) gel containing 0.5 µg/mL ethidium bromide [1% (w/v) agarose in 1X Tris-acetate-EDTA (TAE) buffer, pH 7.8 (Sambrook and Russell, 2001)]. Electrophoresis was performed in 1X TAE buffer. A 500 bp ladder (Genei, India) was used as the molecular size marker. DNA bands were visualized under UV-transilluminator (Gibco-BRL, USA) and size of amplicons determined.

2.2.5. Cloning of CS-PCR products

2.2.5.1. Purification and Elution of the PCR products

The CS-PCR products which shows clear single prominent band were directly purified using GenElute™ PCR Clean-up kit (Sigma-Aldrich, St. Louis, MO) following manufacture's instruction. The CS-PCR products which shows two or more bands, first precipitated by dehydrated ethanol and then 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 ºC in 1X TAE. The required band (s) were sliced and taken in a 2 mL microcentrifuge tube, and were melted at 65 ºC 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 (10 mM Tris-Cl: 0.1 mM EDTA, pH 7.5) and quantified.

2.2.5.2. Preparation of competent *E. coli* by $CaCl_2$ method

Competent cells were prepared as described in Molecular cloning (Sambrook and Russell, 2001) volume 2.

2.2.5.3. Cloning of the purified PCR product

The pGEM–T Easy Vector System II (Promega Corporation, Madison, USA) was used for the cloning purpose of PCR products. For cloning ligation mixture were prepared according to the manufacturer's instructions and then transformed in competent *E. coli* (JM109, XL1 Blue and JM101) cells following heat shock method (Sambrook and Russell, 2001)

2.2.5.4. Selection of recombinants

The recombinants were selected on LB agar plate containing IPTG (isopropylthio-β-galactoside) and X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) and amicillin (100 mg/L) as selection marker. Blue-white screening enabled to discriminate recombinant plasmids (white colored colonies) from non-recombinant (blue colored colonies). Recombinants for the presence of the desired insert were confirmed by PCR reaction using 5′ CS and 3′ CS primers. Recombinant-plasmids were isolated by alkaline lysis method as described in Molecular cloning (Sambrook and Russell, 2001). Recombinant plasmid (clone) containing gene of interest was further confirmed by restriction digestion using *Eco*RI restriction enzyme. The fragments were separated on 1% agarose gel and visualized under UV trans-illuminator after staining with EtBr (ethidium bromide) dye.

Reagents used for plasmid isolation (alkaline lysis method)

Solution I	**Solution II**	**Solution III**
Glucose: 50 mM	NaOH: 0.2 N	5 M CH_3COOK: 60.0 mL
Tris-Cl (pH-8.0): 25 mM	SDS: 1% (w/v)	CH_3COOH: 11.5 mL
EDTA (pH-8.0): 10 mM		H_2O: 28.5 mL

Sol^{n} Ist : Solution was prepared and sterilized by autoclaving at 121 °C for 15 min, and stored at 4 °C.
Sol^{n} IInd: freshly prepared at the time use
Sol^{n} IIIrd: Prepared stock solution was filter sterilized and stored at 4 °C

2.2.6. DNA sequencing and computer analysis of the sequence data

The recombinant plasmids were used for sequencing of the inserts using primers for T7 and SP6 promoters. For large PCR product sequencing, primer walking method was used. Sequencing was performed at DBT-supported DNA sequencing facility at University of Delhi, South Campus, India.

2.2.7. Bioinformatic analyses

Analysis of nucleotide sequences were performed with several bioinformatics tools.

I. ***NEB CUTTER v. 2.0***: For determination of restriction map of the sequence, the software package hosted by New England Biolabs Inc (http://tools.neb.com/NEBcutter2) was used.

II. ***ORF finder***: was used for detection of open reading frame in the nucleotide sequence. This software package is freely available at www.ncbi.nlm.nih. Another package, 'Fast PCR v 3.3.67' of the Institute of Biotechnology, University of Helsinki, Finland, was also used for finding *orfs* in a given sequence.

III. ***BlastN and BlastP suite***: are the freely available software packages (http://www.ncbi.nlm.nih.) which were used for similarity search of nucleic acid and amino acid sequences (Altschul *et al.*, 1997).

IV. ***Conserved domain search (CDD)***: The domain analysis of the amino acid sequence was done using the freely available program of Marchler-Bauer and Bryant (2004) at 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.expasy.org/tools/ and 3D modeling was performed by automated mode 'SWISS-MODEL' (http://swissmodel.expasy.org/): a fully automated protein structure homology modeling server also accessible via ExPaSy web server (http://www.expasy.org). The function of the peptides was analyzed using the ProtFun 2.1 software of the website http://www.cbs.dtu.dk.

VI. ***CELLO v.2.5 and Psortb v3.0.2***: for detection of cellular localization of bacterial protein (available at http://cello.life.nctu.edu.tw; and http://www.psort.org/psortb/ respectively), used for bacterial protein localization prediction.

VII. ***ProtParam analyses***: a tool (available at http://web.expasy.org/protparam/) which allows the computation of various physical and chemical parameters for a given protein stored in Swiss-Prot or TrEMBL or for a user entered sequence. The computed parameters include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY).

VIII. ***TMMOD tool***: a free server for trans-membrane proteins topology prediction using a hidden Markov model. TMMOD uses TMHMM (Sonnhammer *et al.*, 1998) as a prototype, but differs from TMHMM by the architecture of the submodels for loops on both sides of the membrane. This tool is available at website, http://liao.cis.udel.edu/website/servers/TMMOD/. Another membrane protein secondary structure prediction server (http://split4.pmfst.hr/split/4/), SPLIT 4.0, was used for predicting and comparing the data obtained from one server.

IX. ***Sequence Alignment tools:*** To determine percentage homology between two sequences (DNA or protein) the tool, pairwise alignment (http://www.ebi.ac.uk) was used. For multiple sequence alignment (MSA) following tools were used: ClustalW (http://www.ebi.ac.uk, offline available with phylip package and Mega package), ClustalX, T-COFFEE

(http://tcoffee.vital-it.ch/cgi-bin/Tcoffee/tcoffee), multiple align show (http:/bioinformatics.org/sms/).

X. ***Ident and Sim:*** this tool was used to calculate similarity and identity ratio of two or more than two protein sequence by pairwise comparisons (http:/bioinformatics.org/sms/)

XI. ***InterProScan:*** InterPro is a resource that provides functional analysis of protein sequences by classifying them into families and predicting the presence of domains and important sites. To classify proteins in this way, InterPro uses predictive models, known as signatures, provided by several different databases that make up the InterPro consortium. For this purpose InterProScan software package (available at http://www.ebi.ac.uk) that allows sequences to be scanned against InterPro's signatures was run.

XII. ***ProQ - Protein Quality Predictor:*** a neural network based predictor working on a number of structural features predicts the quality of a protein model. It is optimized to find correct models in contrast to other methods which are optimized to find native structures. Two quality measures are predicted *LGscore* and *MaxSub*. This is free server available at http://www.sbc.su.se/~bjornw/ProQ/ProQ.cgi
LGscore is -log of a P-value and
MaxSub ranges from 0-1, [0 is insignificant and 1 very significant]

Different ranges of quality based on ProQ server (http://www.sbc.su.se/~bjornw/ProQ/ProQ.cgi):

Correct	Good	Very good
LGscore > 1.5	LGscore > 3	LGscore > 5
MaxSub > 0.1	MaxSub > 0.5	MaxSub > 0.8

2.2.9. Homology modeling of DfrA1 and AadA1 Proteins derived from bacterium MB47:

A homology model of DfrA1 and AadA was built with SWISS-MODEL using default parameters. In generating homology model, SWISS MODEL utilized the PDB file 3e0bB (crystal structure at resolution2.25 Å of dihydrofolate reductase from *Haemophilus influenzae*) showing 30.57% identity and possessing 0.00e-1 E-value with the query DfrA1 protein of MB47 and PDB file 1no5B (crystal structure at resolution 1.80 Å of HI0073, the nucleotide binding domain of the Hi0073/Hi0074 two protein nucleotidyl transferase from *Haemophilus influenzae*) showing 21.84% identity and 2.00e-9 E-value with the query AadA1. The template was automatically selected by the software for building 3D model of DfrA1 and AadA1 protein of MB47.

2.2.10. Accession numbers

The nucleotide sequences of gene cassettes were assigned EMBL and GenBank accession numbers (Table 2.3 A and B).

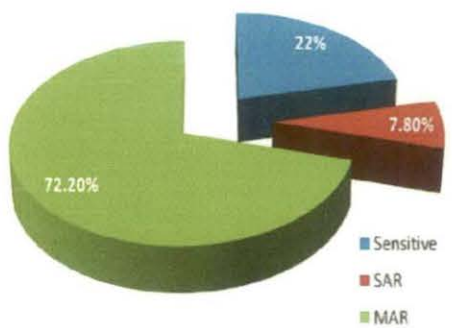


Fig. 2.4A: Frequency of sensitive, SAR (single-antibiotic resistant) and MAR (multiple-antibiotic resistant) among facultatively oligotrophic bacteria those yielded class 1 integron.

2.3. Results

2.3.1. Antibiotic resistance profile of integron positive isolates

The antibiotics-resistance-profile(s) of the ninety class 1 integron bearing facultative oligotrophic bacteria has been presented in Table 2.1. Of the 90 integron positive isolates, 18 (22%) were sensitive to all the twelve antibiotics, 07 (7.8%) were SAR (resistant to only one antibiotic), and 65 (72.2%) were MAR (resistant to two or more antibiotics) (Fig 2.4A).

Among multiple-antibiotic-resistant (65) integron positive isolates, 7 (10.77%) were resistant to two antibiotics, 8 (12.3%) were resistant to three, 5 (7.7%) were resistant to four, 9 (13.84%) were resistant to five, 12 (18.46%) were resistant to six, 9 (13.84%) were resistant to seven, 10 (15.4%) were resistant to eight, 1 (1.53%) was resistant to nine antibiotics, 3 (4.6%) were resistant to ten, 1 (1.53%) isolate was resistant to eleven antibiotics tested (Fig 2.4B). The facultative oligotrophic bacteria resistant to all twelve antibiotics used in this study did not yield any amplicon of class 1 integron. On the basis of resistance index (RI), the facultative oligotrophic bacteria bearing class 1 integrons can be categorized into 11 groups (Table 2.1 and 2.2).

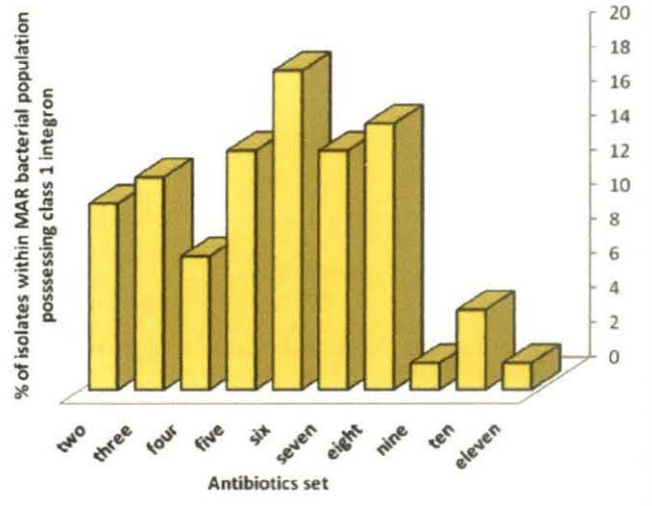


Fig 2.4B: Fractions of class 1 integron positive isolates exhibiting resistance to two and more than two antibiotics tested.

2.3.2. Detection of class 1 integrons

A total of 2188 facultative oligotrophic bacterial isolates, randomly selected (research randomizer tool was used to prevent any human biasness in random selection process) from the pool of oligotrophic bacteria which developed on NPA agar (for detail please see chapter 1), were screened for the presence of class 1 integron by specific PCR (CS-PCR) methodology. Of them, 90 isolates (4.1%) produced amplicon of variable lengths (as the methodology itself enables to amplify the variable region of the class 1 integrons; varied from 0.15 to 3.45 kb) (Fig. 2.5 and Table 2.3A, B, and C). Sequence analyses showed that majority of the amplicons of size >0.7 kb were found to carry antibiotic gene cassettes inserted in class 1 integron platform while in cases where the sizes of the amplicons varied between 0.1 to 0.7 kb, they were either related to empty class 1 integron or were coding for hypothetical proteins. The amplicon lengths ranging from >0.1 to < 0.5 kb of class 1 integron were recorded in five six oligotrophs (OB05, MB05, MB62, MB63, MB70, and MB83). The amplicons ranging between >0.5 kb and <0.7 kb were detected in seven bacteria (MR02, MB41, MB44, MB45, MB54, MB80, and MB81). Amplicons of lengths ranging from >0.7 to <1.0 kb of class 1 integron were recorded in sixteen oligotrophic bacteria (MR01, MR03, SR19, MB31, MB40B, MB42, MB43, MB49, MB51, MB55, MB56, MB57B, MB64, MB76, MB72, and MB77). Amplicon of lengths, ranging between ≥1.0 to ≤1.5 kb were noted in thirty three (OB12, MB12, MB16, MB22, MB48, MR04, MB19, MB39, MB46, MB20, MB24, MB50, MB57A, MB23, MB36, MB32, MB33, MB18, MB66, MB74, MB75, OD05, OD08, OC78, MB79A, MB79B, OD10, OD21, OD24, OC74, MB38, OC75, and NV66) of the 90 class 1 integron positive bacterial isolates. Twenty four isolates (MB52, MB03, MB21, MB26, MB67, MB47, MB40A, MB59, MB27, MB30, MB60, MB61, MB69, MB78, ºC16, MB29, MB34B, MB37A, MB53, MB73, MB25, MB65, OC24, and MB68) were found to carry amplicon of lengths, corresponding to the size which varied between >1.5 but ≤ 2.0 kb. A 2.0 kb and ~ 3.5 kb amplicon was recorded in MB82 and MB35 respectively (Table 2.3B). A very short variable region of 153 bp were amplified in two strains, MB62 and MB63, while 223bp amplicon was generated from isolates MB05 and MB70 (Table 2.3A). The amplicon of size ~1.0 kb was predominating and was detected in 24.4% (22/90) of the total integron positive isolates (Table 2.3 A and B).

2.3.3. Characterization of integron borne gene cassettes

Purified CS-PCR products cloned in pGEM-T Easy vector-II (Fig. 2.6) were subjected to DNA sequencing. Gene cassettes were identified by sequence analysis. The identity of the gene cassettes and their predicted functions are presented in Table 2.3 A and B. Facultative oligotrophic bacteria were found to carry both types of ORFs (i) not related to antibiotic resistance (Table 2.3A) and (ii) antibiotic resistance (Table 2.3B) gene cassettes.

2.3.3.1. Gene cassettes revealing ORFs unrelated to any known antibiotic resistance gene function

The cassettes, yielding ORFs unrelated to any known antibiotic resistance gene function, have been identified in 26 (~29%) of the total 90 isolates (Fig. 2.7). The features of such gene cassettes were analyzed and compared with previously existing data. The sequences derived from amplicons of OB05 and MB83 were found 75% identical (nucleotide-nucleotide) to the part of the genome of *Acidovorax avenae* subsp. *citrulli* (Ac. No. CP000512). The feature present in that part of *Acidovorax* genome included a portion of ribosomal large subunit pseudouridine synthase B. However the predicted polypeptide of the ORF generated from OB05 DNA sequence did not produce any significant similarity with existing sequences of the protein database. Sequence analysis of the amplicons of MB05 and MB70 revealed the presence of a truncated ORF of 64 amino acids that showed 73% identity with the methyl transferase protein from *Ralstonia eutropha* (Ac no. AAZ62060). A translated polypeptide from the sequence derived from bacterium MB09, signifying type I site-specific deoxyribonuclease, HsdR family, exhibited nearly 42% identity with a protein of the same family, found in *Thiomicrospira crunogena* (Ac. No. YP390604). Partial sequencing of the CS-PCR product (Ac. No. AM997273) obtained from MB12 showed the presence of a single ORF (*livM1*) of length 154 amino acids (Fig 2.8A). BlastP analysis (Table 2.4) of this putative polypeptide product exhibited best scores with ABC type branched chain amino acid transport system with 72% identity at amino acid level to the ABC transporter permease of bacterium *Aromatoleum aromaticum* EbN1 (Ac. No. NC_006513). The multiple sequence alignment of LivM1 derived from amplicon sequence of bacterium MB12 with its neighbouring ABC transporter proteins is shown in Fig. 2.8B. The transmemebrane (TM) regions predicted within the ABC transporter protein of MB12 (Protein ID: CAQ53856) are shown in Fig. 2.8C and 2.8D. Analysis revealed the presence of four transmembrane regions in the predicted LiVM1 protein of MB12. The predicted transmembrane (TM) regions (AA position 2-20, 50-74, 87-110, and 119-138) in ABC transporter protein of bacterium MB12 are presented in table 2.5. The conserved domain (CDD: conserved domain database tool available at www.ncbi.nlm.nih.gov) search for putative translated product obtained from translated ORF of MB12 integron sequence revealed that the predicted protein was one of the member of two TM subunits which play role in uptake of branched chain amino acids (Fig 2.8E). Mapping of signature motifs (Fig 2.8F) showed that 19 of 19 [TM_PBP1_LivM_like, a type of transporters which consist a PBP (periplasmic binding protein), two TMs (trans-membrane subunits) and two cytoplasmic ABCs (ATP-binding cassette)] residues which were recognized on the TM-ABC transporter protein sequences of other bacterial species were also found conserved on ABC protein derived of MB12. The theoretical pI and instability index of the putative ABC transporter of MB12 was computed 9.3 and 19.13. The protein was predicted as stable protein with an estimated half life of 10 h (*in vivo* in respect to *E coli*) with a high aliphatic index of 127.21. MB16 was found to carry gene cassette carrying *appA* gene encoding bacterial extracellular solute binding protein. ProtParam computing showed that the extracellular binding protein derived from bacterium was stable and its instability index was computed 18.94. The GRAVY (grand average hydropathicity), aliphatic index and theoretical pI was computed -0.087, 78.5, and 9.26 respectively. Psortb, a tool for subcellular localization prediction showed that the translated product of *appA* gene was periplasmic with localization score of 9.44. Sequence analysis of amplicon obtained from MB 19 has shown the

presence of a unique hybrid DNA sequence. In the 867 nucleotide long sequence, a continuous stretch of 242 nucleotides (from 3 to 244) produced 96% identity with a vertebrate (*Lepilemur dorsalis)* genomic fragment (Ac. No. AJ244007) and residual 623 nucleotide stretch produced (from nucleotides 245 to 867) 74% identities with *Burkholderia xenovorans* LB400 genomic DNA. This genomic DNA region of *B. xenovorans* coded for the putative reverse transcriptase maturase protein. MB 19 sequence yielded an ORF of 227 amino acids that has shown 65% identity with reverse transcriptase maturase protein from *Burkholderia cenocepacia* HI2424 (Ac. No. YP833935). An ORF of 164 amino acid residues (495 nucleotides) obtained following sequence analysis of 1010 bp amplicon of MB 22. The putative polypeptide shared 81% amino acid identity with phosphoribosylformylglycinamidine synthase (FGAM synthase) from *Acinetobacter baumannii* (Ac. No. YP001712860). The partial sequence of MB 28 (408 nucleotide) was 86% identical with *Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578 (Ac. No.CP000647) genomic DNA. The amino acid sequence (truncated protein) derived from amplicon of bacterium MB28 was encoding a bifunctional putative transcriptional regulator protein, this truncated protein of 84 amino acid shared homology of 96.3% with the similar protein present in *Klebsiella pneumoniae* (Ac. No.CP000647). The Complete sequence of 704 bp amplicon of MB 40B (Table 2.3B) yielded no significant homology with any nucleotide sequence available in the database. The same sequence yielded a complete ORF of 102 amino acids and BlastP analysis of this protein sequence revealed 58% identity with helicase domain protein of *Verminephrobacter eiseniae* EF01 (Ac. No. ABM585806). The sequences derived from amplicons of OB 12, MB 08, MB41, MB 44, MB 48, MB 51, MB 54, MB55, MB80 and MB81 did not produce any significant homology with any protein sequences available in the database. All of them were characterized as hypothetical proteins. All these polypeptides shared very low level of identity (2.03 to 25.19%) among either with each other or to pre-deposited hypothetical protein sequences. The translated product (Fig. 2.9A) of bacterium MB54 revealed a conserved domain belonging to the UPF0153 superfamily (Fig. 2.9B) when subjected to the BlastP search. However the function of proteins of this family is still unknown; since the proteins of this family contains 8 conserved cysteines hence it was expected that they may constitute a metal binding site and therefore might play a role in metal sequestering like Fe-S cluster as part of oxido-reductase complex. Protein homology and CDD search of putative polypeptide of 178 amino acid residues derived from the CS-PCR product of MB 56 isolate was found to carry a conserved protein domain, encoding hemolysin activator/ secretion protein that was involved in intracellular trafficking and secretion. This hemolysin activator protein of MB56 produced 55% identity with the hemolysin activator protein of *Acinetobacter* sp. (Ac. No. YP045656). Translation of DNA sequence obtained from MB 58 and MB71 yielded a truncated ORF of polypeptide of 181 amino acids. The BlastP analysis of the said truncated polypeptide yielded best score (97% identity) with transposase protein of *Acidovorax* sp. (Ac. No. YP987142).

2.3.3.2. Antibiotic resistance gene cassettes

A total of 2188 facultatively oligotrophic bacteria were selected for the detection of class 1 integron in their genome. To fulfill this target, a highly reproducible PCR methodology using specific primer pair (5′ CS and 3′ CS) was used. The class 1 integron was detected only in 90 isolates (4.1%) of the total oligotrophic isolates. Remaining 2098 bacterial strains failed to produce any amplicon with the primer set 5′ CS and 3′ CS.

A strong association between the nature of the gene cassettes and phenotype of the isolates was observed (Table 2.2 and 2.3 B and C). However, exception was noted in some cases of *aadA* gene cassettes where correspondence with the phenotype were not observed. About 71% (64/90) gene cassettes were found to carry antibiotic-resistance genes. The most common carriages were aminoglycoside adenyltransferase gene cassettes such as *aadA, aadA1, aadA2, aadA4,* and *aadA5* which encode aminoglycoside adenyltransferase enzyme conferring resistance to streptomycin/

spectinomycin antibiotics. Two types of dihydrofolate reductases, type-A (*dfrA1, dfrA5, dfrA7, dfrA12, dfrA16, dfrA17,* two novel *dfrA* genes, *dfrA28* and *dfrA30*) and type-B (*dfr-IIe*), conferring resistance to trimethoprim, were found.

Thirty three bacterial strains carried only one antibiotic-resistance gene cassette and twenty five were found to carry an array of two gene cassettes [Table 2.3 B and C]. An array of three gene cassette (*dfrA12-orf40A-aadA2*) was detected in three integron positive isolates (MB40A, MB61, and MB67) [Table 2.3 B]. One bacterium MB79A was having a gene cassette array of type B (also called as typeII) *dfr-IIe (dfrB)* dihydrofolate reductase, and *arr2* gene, responsible for resistance to rifampicin (a semi-synthetic antibiotic agent) [Table 2.3 C]. A gene cassette array of *Oxa1-aadA1* was detected from bacterium MB82 [Table 2.3 C]. This array was responsible for resistance towards β-lactam antibiotics such as ampicillin and aminoglycosides such as streptomycin/spectinomycin [Table 2.3 C]. The bacterium, MB35, revealed the largest array of gene cassettes, *blaIMP-9-aacA4-oxa10-aadA2,* among the all class 1 integron positive isolates. Maximum percentage (10%, of the total resistance cassettes) of two-gene-cassette array was observed for *dfrA-aadA1* type followed by *dfrA17-aadA5* (~8%). Several resistance phenotypes observed with the isolates (MB19, MB20, MB29, MB31, MB41, MB62, MB68, MB76 etc) for antibiotics like azithromycin, cefipime, cefotaxime, chloramphenicol, ciprofloxacin, levofloxacin, netilmicin, and oxytetracycline did not correspond to sequences of the gene cassettes amplified from them.

The details of gene sequences and types of resistance gene cassettes found within the variable region of class 1 integrons (amplicon generated though CS-PCR) from different bacterial isolates are shown in Table 2.3B. The sequence derived from the 1009 bp amplicon from bacterial strains MB20, MB50, MB23, MB36, MB32, MB33, MB66, OD05, OD08, showed single gene cassette bearing *aadA1* gene [Table 2.3B]. The *aadA2* gene cassette was seen alone in four isolates. Similarly aadA2 cassette was found alone in two isolates. The *aadA5* gene cassette was found singly in two isolates, MB18 and SR19 (Table 2.3.B and C). The sequence derived from MB 18 showed 99% identity at nucleotide level with *Riemerella anatipestifer* class 1 integron *aadA*5 gene (Ac. No. EF105289) and 98% identity at the amino acid level with aminoglycoside adenyltransferase from *Escherichia coli* (Ac. No. YP190214), *Klebsiella pneumoniae* (Ac. No. AF180469), *Salmonella enteritidis* (Ac. No. BAD02332). The isolate MB62 and MB63 produced amplicon of size 153 bp, sequence analysis revealed that these were empty class 1 integrons devoid of any gene cassette. Isolates, MB03 and MB52, were found to carry identical cassette array of two gene, *dfrA28* and *aadA1*. The gene dfrA28 (Ac. No. FN263373) was novel having 519-bp long ORF with shared homology (identity) of 76.4% at amino acid level to the *dfrA1* of *E. coli* (Ac. No. AJ419168). The details of this novel gene will be discussed in chapter 4. The 1913 bp amplicon carried two known genes, *dfrA12* and aadA2, as well as an ORF and was present in MB40A, MB61, and MB67 isolates. Aminoglycoside acetyltransferase gene cassette (*aac-6′-Ib*) was detected in only two isolates, OC74 and OC78 (Table 2.3B). The *aacA4* gene cassettes conferring resistance to kanamycin was detected from MB42, MB43, MB49, and MB35 isolates. In MB35, *aacA4* gene cassette was flanked by two gene cassettes conferring resistance to beta-lactam and aminoglycosides. A sequence of length 737bp from MR03 showed a single gene cassette, *dfrA16* coding for dihydrofolate reductase enzyme. The sequence derived from the 769 bp amplicon from three bacterial strains MB31, MB64, and MB76 showed single *dfrA7* gene cassette (Table 2.3B). The sequence derived from the 1543, 1556, and 1569 bp amplicons from seven isolates (MB26, MB47, MB59, MB60, MB69, MB65, and MB68) showed *dfrA1-aadA1* gene cassette array (Table 2.3B). The detailed sequence analyses (ORF features, *Insilico* restriction digestion pattern of nucleotide sequence, multiple alignment of translated ORF with the nearest sequence, CDD search, physic-chemical parameters, InterPro's signature search for DfrA1, protein secondary structure prediction, and homology modeling of DfrA1 and AadA1 proteins) derived from MB47 isolate are shown in Fig. 2.10A to 2.10Mz. Based on similarity of MB47 dihydrofolate reductase to the dihydrofolate reductase 3e0bB, the homology model was built. The

sequence identity between the query sequence and template sequence was found to be 30.57% with an expected value of 0.00e-1. The region of the enzyme from residue 1-151 was successfully modeled using SWISS-MODEL automated mode tool (Fig 2.10H). An alignment of query (MB47 DfrA1) and template sequence [dihydrofolate reductase of *Haemophilus influenzae* (PDB ID: 3e0bB)] is shown in Fig. 2.10H. The comparison of generated homology 3D model (Fig 2.10Hy) of MB47 DfrA1 protein (NCBI protein ID: CAQ76923) and 3e0bB three dimensional model (Fig 2.10Hx) of *Haemophilus influenzae* showed that the model of putative DfrA1 protein was similar to the dihydrofolate reductase of *Haemophilus influenzae.* The estimated Swiss-model quality showed that built model was reliable with Qmean score 0.7 (reliable range: 0-1) (Fig 2.10Hz). The model quality was also evaluated by ProQ tool. For homology modeling of AadA1 protein of MB47, template, 1no5B (crystal structure at resolution 1.80 Å of HI0073, the nucleotide binding domain of the Hi0073/Hi0074 two protein nucleotidyl transferase from *Haemophilus influenzae)* showing 21.84% identity and 2.00e-9 Evalue with query AadA1 was selected automatically by the software. The region of the enzyme from residue 10-96 (comprising nucleotide binding domain) was modeled as described above. Pairwise alignment between query (MB47 AadA1) and template sequence [nucleotidyl transferase from *Haemophilus influenzae* (PDB ID: 1no5B)] is shown in Fig. 2.10M. The comparative analysis between produced model (Fig 2.10My) of MB47 AadA1 protein (NCBI protein ID: CAQ76924) and 1no5B three dimensional model (Fig 2.10Mx) of *Haemophilus influenzae* showed that the model for nucleotide binding domain of AadA1 protein was similar to the nucleotide binding domain of *Haemophilus influenzae.* The obtained QMEAN score 0.51 exhibited that the generated homology model of AadA1 was reliable (Fig 2.10Mz). Further model quality was assessed by ProQ tool.

The sequence derived from the 1647 and 1694 bp amplicons from three bacterial isolates MB03, MB52 and MB78 respectively showed *dfrA28-aadA1* gene cassette array (Table 2.3B). The integron sequence derived from five bacterial isolates MB27, OC16, MB37A, MB53, and OC24 showed *dfrA17-aadA5* gene cassette array. Sequence features of variable region derived from MB53 are shown in Fig. 2.11. The *aad*A5 gene cassette was the second cassette identified from MB 53 and consisted of an ORF of 224 amino acids. AadA5 protein from MB 53 showed 100% identity with the same protein from a Gram positive bacterium *Staphylococcus epidermidis* (Ac. No. AB291061) and Gram negative bacterium *Enterobacter cloacae* (Ac. No. EF571855). The 1606 and 1663 bp amplicons of bacterial isolates (MB25 and MB34B) carried two tandemly inserted gene cassettes consisting of the *dfrA17* and the *aadA4* genes in class 1 integron sequence. The integron sequence derived from three isolates MB29, MB30, and MB73 showed two genes, *dfrA7* and *aadA5* in their gene cassette array. The sequence features (bearing *dfrA7* and *aadA5* ORFs) derived from MB29 are detailed in Fig. 2.12A - 2.12D. The gene cassette array, *dfrA1-orf* was detected in four bacteria, MB38, MB74, MB75, and MB57A (Table 2.3B). The sequence analyses of variable region derived from MB38 bacterium are detailed in Fig. 2.13A and 2.13B. Protein-protein homology of ORF38 derived from MB38 showed that it was 99% identical to the hypothetical protein of *Escherichia coli* (Protein id BAD08521) however, no conserved domain was observed in BlastP analysis. A single gene cassette, *dfrA30*, was observed in three isolates, MB45, MB72, and MB77. The gene, dfrA30 containing 471 bp long ORF, which shared maximum 93% identity at amino acid level with the closest known dfr (*dfrA5*) sequence of *E. coli* (Ac. No. AJ419169) was recognized as a novel (for reporting novel gene protein-protein identity should be <95%) gene cassette. The features of *dfrA30* gene and its source will be discussed in detail in chapter 5. The PROSITE motif search has revealed that the dihydrofolate reductase signature sequence, VIGngpdIPWsakg.EqllFkaiT, was intact in most of the Dfr protein sequences.

2.3.3.3. Integron carriages in samples collected from three different sampling sites (SS I, SS II, and SS III); and occurrences of class 1 integrons in different resistance index groups

The annual study of integron fishing in facultative-oligotrophic (%) isolates from each sampling site per year (2007-2009) is shown in Table Fig. 2.14. The study revealed that the frequency of class 1 integron was highest at SS II in all three years (Table 2.10; Fig 2.14).The maximum carriage of class 1 integron in isolates of SS III was in the year 2007 and at SS I, it was in year 2008. There was similar occurrence of class 1 integrons (7.7%) were observed at SS I, SS III and SS II during year 2007, 2008 and 2009 respectively (Table 2.10).

A thorough study of occurrences of integrons in different resistance index [Resistance index (RI) =number of antibiotics to which isolate exposed/ total number of antibiotics taken in study] ranges 0.08-1.0, were analyzed and results were interpreted. Results showed that incidence of integrons in resistant bacterial population were lowest at 0.08 RI and maximum at 0.66RI. A gradual increase of the presence of class 1 integron was noticed from 0.08RI to 0.5 and then an irregular pattern was observed (Fig. 2.15). It was found that incidence of class 1 integrons remained quite high between 0.5- 0.9; meaning that this category of resistance index is best for fishing integrons from oligotrophic bacteria.

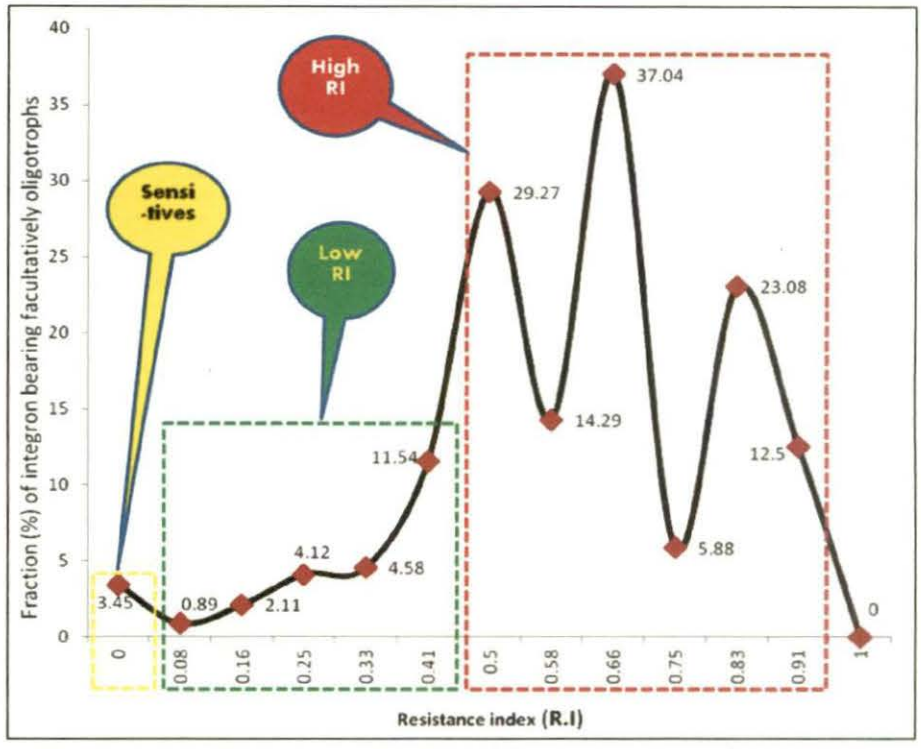


Fig. 2.15: Incidence of class 1 integrons in different resistance index groups

2.4. Discussion

The notion that the majority of bacteria existing in nature are yet to be discovered has gained its support from the large-scale employment of new technologies in microbiological research of the past decades. A marked discrepancy between the relatively low number of microorganisms cultivated under laboratory conditions and the number and diversity of organisms actually existing in the environment has been revealed from the analysis of environmental specimens conducted by molecular biological methods or microscopy. Genetic diversity of this majority and unidentified gene families remains still obscure. Bacteria are differentiated from other living systems by a striking plasticity which allows them to adapt easily to various frequently changing environmental conditions. Specific mechanisms by which changes are brought about in the amount or order of genes in bacterial genomes contribute to the high adaptability of bacteria (Ilyina, 2006). Bacteria undergo selection pressures in the presence of various toxic substances; hence developments of resistance against the inhibitory agents, commonly the resistance to metal and antibiotics, provide selective advantages (Hideomi *et al.*, 1977; Timoney *et al.*, 1978; Watkinson *et al.*, 2007). Resistant bacteria are ubiquitous as they are isolated from diverse geographical locations (Mudryk *et al.* 2000; DeSouza *et al.*, 2006; Xi *et al.*, 2009; Chattopadhyay and Grossart, 2011; Nageswaran *et al.*, 2012). With increasing bacterial resistance to almost all the discovered antibiotics, the phenomenon has been recognized by microbiologists from all over the world. A sensitive bacterium can become

resistant either through mutation or by acquiring resistance genes by means of horizontal transfer through transformation, transduction, or conjugation. It is assumable that amongst three modes of gene transfer mechanisms, conjugation may be the most common mechanism of transferring antibiotic resistance genes. By dint of conjugal gene transfer, plasmids and transposons carrying antibiotic resistance genes can easily mobilized from one cell to another. From the last quarter of the 20th century, another class of mobile DNA elements bearing antibiotic resistance gene came into prominence of research.The discovery of integrons came into light during 1986, when researchers meticulously analyzed the DNA sequence of numerous apparently unrelated antibiotic resistance genes and found that the regions flanking various antibiotic resistance genes were conserved. Furthermore these common regions were observed in different places on various plasmids, this suggested that, like transposons, these elements were also mobile. However, these elements differed from transposons in two important ways: (i) the regions flanking the antibiotic resistance genes in the new elements were not direct or indirect repeats, as found in transposons and (ii) the elements were holding a site-specific integrase gene of the same family (tyrosine recombinase) as found in λ phages but deficient in many gene products associated with transpositon. Due to these dissimilarities, the new elements were separated from transposons and were called integrons.

Integrons are the genetic elements with an ability to take into custody one or more non-functional genetic determinants [called as gene cassettes, a promoter-less (leaving exception) non-replicating mobile DNA elements which can also exist in free-circular form] often encoding antibiotic resistance genes by site-specific recombination and convert them in functional (Hall and Collis, 1995). The locations of these determinants have been traced in the chromosome as well as in different plasmids and transposons by several researchers (Hall and Stokes, 1993; Ozgumus *et al.*, 2009). On comparing the amino acid sequences of the integrases, encoded by *int*I genes, integrons have been classified into several classes and class 1 to 3 are termed as mobile integrons. Of these, the well studied class, class 1 integron is now ubiquitous and playing important role in the dissemination of antibiotic resistance (often) genes in both Gram-negative (Hall and Stokes, 1993; Mukherjee and Chakraborty, 2006) and Gram-positive bacteria (Nandi *et al.*, 2004; Shi *et al.*, 2006; Xu *et al.*, 2008; Xu *et al.*, 2011). Typically the integrons have been identified on plasmids or on transposable elements such as mercury resistance transposon *Tn21* (Grinsted *et al.*, 1990; Lindstedt *et al.*, 2003; Rodríguez *et al.*, 2006). Researchers have shown that most of the class1 integrons are derivative of defective transposon (Brown *et al.*, 1996; Xu *et al.*, 2009). Analyses of whole genome sequences resulting from genome sequencing projects (Mazel, 2006; Rowe-Magnus *et al.*, 2001; Boucher *et al.*, 2007) have revealed that about 10% of the sequenced genome of distantly related lineage possess a variety of *intI* genes (Boucher *et al.*, 2007). To date integrons in the Eucarya or Archaea have not been reported (Nemergut *et al.*, 2008)

Class 1 integrons, the commonly studied class of integrons consists of three DNA segments, (i) 5''-conserved segment, (ii) variable region which contain gene cassette (mostly coding for antibiotic resistance, but not always) of different lengths and sequences and the last (iii) 3'-conserved segment (Fig. 2.1). The 5'-conserved segment (5'-CS) includes the *intI1* gene (class 1 integrase) encoding the integrase (as discussed above) belonging to the tyrosine recombinase family which catalyzes site-specific recombination between *attI1* and *attC* (also called as 59 base element as the first described *attC* was of 59 bp long) sites with site preferences. The protein, IntI1, in bacteria can bring about recombination between, either two *attC*, one attI1 and one *attC*, or two attI1 sites (Collis *et al.*, 2001). Integration and excision reactions, involving attI1xattI1 or *attCxattC* sites, appear to be the most relevant biological events (Stokes *et al.*, 1997). This 5'CS also contains a common promoter (Pc) region. In fact Pc potentially contains two promoters, P1 and P2 (Bissonnette *et al.*, 1991; Stokes and Hall, 1989) which play role in expression of captured promoter-less gene cassettes. Thus, integrons are "natural cloning and expression systems" that are able to capture

cassettes with open reading frames (ORFs) and convert them into actively functioning genes (Rowe-Magnus and Mazel, 1999, 2001).

The 3′-conserved segment (3′-CS) is generally composed of a truncated *qacE* (*qacED1*) gene encoding low-level resistance to disinfectants, the *sul1* gene encoding sulfonamide resistance, and in some occasions two open reading frames (ORFs), ORF4 and ORF5 (Stokes and Hall, 1989)of unknown function. Integrons cannot be mobilized by itself but contain gene cassettes that can be mobilized to other integrons or to secondary sites in the bacterial genome. The class 1 integron system, being an effective gene-capture and assembling system, renders bacteria to incorporate gene cassettes and transforms them to functional entities by modulating their correct expression. Therefore, it has been seen as key player in the propagation and dissemination of resistance genes leading to simplistic spread of resistance genes and an agent for rapid evolution of resistance to a wide range of unrelated antibiotics among diverse bacteria (Mazel, 2006; Nemergut *et al.*, 2008).

The gene cassettes are important in the swapping of antibiotic resistance genes. The origin of these cassettes is still unclear but it is expected that they might have originated from chromosomal or super integrons (Mazel *et al.*, 1998; Fluit and Schmitz, 2004). Most of the known gene cassettes are identified as resistance cassettes carrying genes for antibiotic resistance determinant, notwithstanding the fact that they can also transfer a diverse suite of genes unrelated to antibiotic resistance (Holmes *et al.*, 2003; Nemergut *et al.*, 2004; Moura *et al.*, 2007; Gillings *et al.*, 2008b).

Class 1 integrons have been reported from bacteria isolated from diverse niches like human and animal clinical settings, agricultural environments, urban and pristine areas, food animals, pets, wild animals and also in aquatic environments (Rosser and Young, 1999; Nandi *et al.*, 2004; Agerso and Sandvang, 2005;Gaze *et al.*, 2005; Hardwick *et al.*, 2008; Leverstein-van Hall *et al.*, 2001; Sunde, 2005; Roe *et al.*, 2003; Moura *et al.*, 2007; Wright *et al.*, 2008; Binh *et al.*, 2009; Moura *et al.*, 2010; Byrne-Bailey *et al.*, 2010; Betteridge *et al.*, 2011; Nardelli *et al.*, 2012). However, reports on the incidence of integrons from bacteria isolated from fresh-water ecosystems like creeks or rivers are very few (Petersen *et al.*, 2000; Mukherjee and Chakraborty, 2006; Ozgumus *et al.*, 2007; Ozgumus *et al.*, 2009). There are few studies that are reported from India, particularly outside the clinical settings (Mukherjee and Chakraborty, 2006; Pathak and Gopak, 2008, Deekshit *et al.*, 2012).

Surveillance for the presence of integrons in bacteria has become a tool for clinical infection control and the study of antibiotic resistant mechanism (Jones *et al.*, 2003). Hence, such studies extended to environments outside the clinical settings would enable to assess the antibiotic-resistance gene pool in the environment as such. Bacteria from diverse ecological niches (environmental, animal and human) are able to mix in aquatic system and consequently resistance in sensitive bacteria may develop as a result of illegal exchanges and shuffling of genes using different genetic platforms and vectors (Baquero *et al.*, 2008).Water is the principal carrier of pollutants in the environment and therefore has been attracted by the environmental biologists. Furthermore, water bodies have been identified as ideal carriers for the dissemination of antibiotic-resistance genes (Lupo *et al.*, 2012). When compared to the waters less affected by human habitation (i.e pristine water bodies), the incidence and occurrences of class 1 integrons bearing strains is higher in polluted water sources (Stalder *et al.*, 2012). Hence, detection and determination of integron constituents at the molecular level is an important scientific issue.

The present study was out lined in the following steps: (i) enumeration of oligotrophic bacteria on nutrient poor medium (discussed in previous chapter); (ii) screening the isolates for obligate and facultative bacteria (discussed in previous chapter); (ii) random selection of facultative oligotrophs for determination of antibiotic susceptibility/resistance (described in previous chapter); (iv) molecular analyses of class 1 integron borne determinants (in the present chapter); and (v) the provisional identification of bacteria bearing class 1 integron (to be detailed in chapter 3).

Mostly the integron studies have been conducted on gram negative bacteria but it is also documented that gram positive bacteria do also bear resistance gene associated with integrons (Nesvera *et al.*, 1998; Nandi *et al.*, 2004; Shi *et al.*, 2006; Yan *et al.*, 2010; Xu *et al.*, 2011).

In this study, the test strains were selected by the use of research randomizer (a tool to generate unique random numbers) to prevent any human biasness, hence the collection of randomly selected bacteria (for screening of class 1 integrons) is an overall representation of culturable facultatively oligotrophic bacteria of the River Mahananda at Siliguri, irrespective of gram-reaction and sensitivity towards any of the antibiotics used. Most of the earlier studies on integron were knowingly or unknowingly restricted to the copiotrophs (bacteria that are cultivated on nutrient-rich media). As per existing literature, we do not find any reference of similar work motivated to explore class 1 integrons in oligotrophic bacteria from the environmental samples. Only few references are available regarding antibiotic-resistance in oligotrophic bacteria (Niktin *et al.*, 1988; Zlatkin *et al.*, 1991; Kimura *et al.*, 1995; Oh *et al.*, 1995; Kumar *et al.*, 2011; Bhullar *et al.*, 2012). The current study showed that 76.2% (1667/2188) of the oligotrophic bacterial population recovered was resistant to at least one of the antibiotic (12 antibiotics) tested. These results were similar to those published by Kimura (1995), a study on oligotrophic bacteria of soil in their case, where it was found that 83.7% oligotrophic bacteria exhibited antibiotic resistance. In this study, it was observed that MAR (multiple antibiotic resistant) bacteria were dominant over SAR (single antibiotic resistance) by 6% (Fig. 1.12, chapter 1) and the distribution of MAR population were as follows: 19.86% (331) were resistant to two, 11.64% (194) to three, 6.54% (109) to four, 4.67% (78) to five, 2.46% (41) to six, 3.78% (63) to seven, 1.62% (27) to eight, 1.02% (17) to nine, 0.78% (13) to ten, 0.48% (8) to eleven antibiotics respectively, and 0.24%(4) to all the 12 antibiotics tested (Fig. 1.13, chapter 1). Closer examination of data showed that facultatively oligotrophic MAR bacteria (in total) were dominant at sampling station II in comparison to SS I and SS III (Fig. 1.15, chapter 1). The SAR bacteria were more frequent amongst isolates of SS I and SS III.

A hypothesis, in light of previous existing data on antibiotic resistances in oligotrophic bacteria, was proposed that "oligotrophic bacteria may be potential reservoir of antibiotic-resistance genes in the aquatic (riverine) environment". The results of this study reflect that 3/4th of the oligotrophic bacterial population were resistant to at least one of the 12 different antibiotics tested. Studies done on river water copiotrophs revealed that MAR bacteria were dominant (Ash *et al.*, 2002; Mukherjee and chakraborty, 2006, Pathak and Gopak, 2008; Ozgumus *et al.*, 2009 Li *et al.*, 2009; Dong *et al.*, 2010) in the copiotrophic population. In this study, the MAR isolates were predominant within the enumerated oligotrophic bacterial population. In addition to above data, isolates screened for the presence of class 1 integrons were found to possess antibiotic-resistance genes; however other unknown/and unrelated function encoded by gene cassettes were also traced. Hence, this study has its own merit in terms of antibiotic resistance and data generated on antibiotic-resistance-gene-cassettes which will allow filling the gap and opening a new dimension in the study of oligotrophic bacteria.

A total of 2188 randomly selected facultatively oligotrophic bacteria were used for detection of class 1 integron; of them 90 (4.1%) produced amplicons, corresponding to the variable region of a class 1 integron, with the primer set 5'CS (specifically bind at 5' conserved segment of class 1 integron) and 3'CS (specifically bind at 3'conserved segment of class 1 integron) (Levesque *et al.*, 1995) and remaining 2098 bacterial strains failed to produce any amplicon. This result was comparable to an earlier study where non-selected gram-negative bacteria from an estuarine environment revealed the presence of class 1 integron in 3.6% of the total bacteria tested (Rosser and Young 1999). A study conducted on Rio Grande River showed that 10% of the *E. coli* isolates were MAR and 13% of them were carried class 1 integron sequences (Roe *et al.*, 2003). Class 1 integrons was observed in 3.8% of the isolates recovered from environments polluted by quaternary ammonia compounds (Gaze *et al.*, 2005). It was observed that within the facultatively oligotrophic

MAR bacteria of Mahananda River, the incidence of class 1 integrons remained quite high in the resistance index group which ranged between 0.5- 0.9, meaning that this category of resistance index is best for fishing integrons amongst oligotrophic bacteria (Fig 3.11).

The length of variable region amplified by CS-PCR ranged between 0.15 to ~3.5 kb (Fig.3.2 A and B). Most of the amplicons of length >0.7 kb were found to carry resistance gene cassette. In case of absence of any gene-cassette within the integron platform, the expected size was found 153 bp (Gebreyes and Thakur, 2005,) and these were termed as empty class 1 integron as noticed in MB62 and MB63 isolates (Ac. No. FM998811 and FM958478 respectively). It was seen that amplicons of size ~1.0 kb predominated and was detected in 24.4% (22/90) of the total integron positive isolates (Table 3.3 A, B and C). Maximum carriages of class 1 integron were documented in MAR oligotrophic bacteria; Moura *et al.* (2007) found that 47.6% bacteria resistant to five or more antibiotics (isolated from slaughter house waste water treatment plant) possessed class 1 integron. In clinical settings, 43-75% antibiotic resistant bacteria were found to carry class 1 integrons (Levesque *et al.*, 1995; Jones *et al.*, 1997; Martinez-Freijo *et al.*, 1998; Chang *et al.*, 2000; Jones *et al.*, 2003). On examining the antibiotic-resistance-profiles of the class 1integron positive isolates, it was found that 22% (18/90) isolates were sensitive to all the twelve antibiotics, 7.8% (07/90) were SAR (resistant to only one of the 12 antibiotics), and 72.2% (65/90) were MAR (resistant to two or more antibiotics) (Fig 3.1A, Table 3.11). Within 65 multiple-antibiotic-resistant isolates, 10.77% (7/65) were resistant to two antibiotics, 12.3% (8/65) were resistant to three, 7.7% (5/65) were resistant to four, 13.84% (9/65) were resistant to five, 18.46% (12/65) were resistant to six, 13.84% (9/65) were resistant to seven, 15.4% (10/65) were resistant to eight, 1.53% (1/65) was resistant to nine antibiotics, 4.6% (3/65) were resistant to ten, 1.53% (1/65) isolate was resistant to eleven antibiotics tested (Fig 3.1B). The facultative oligotrophic bacteria resistant to all twelve antibiotics used in this study did not yield any amplicon with primers for class 1 integron. On the basis of resistance index (RI), the isolates could be categorized into 11 groups as shown in Table 3.1 and 3.2. Observations led from this study revealed that incidence of integrons in antibiotic-resistant bacteria were lowest at 0.08 RI and maximum at 0.66RI (Figure3.11).

The sequence analyses revealed that function encoded by gene cassettes were strongly associated to the antibiotic-resistance-phenotype of the respective integron positive isolates. Similar to earlier studies (Bissonnette *et al.*, 1992) majority (71%) of the integron-borne cassettes corresponded to antibiotic- resistance-genes while 29% of the gene cassettes were found to carry genes unrelated to antibiotic resistance. The common carriage of antibiotic resistance genes were observed for aminoglycoside-modifying enzymes and trimethoprim-resistant dihydrofolate reductases. Only a few antibiotic-resistance-gene-cassettes were found to carry genes for other function like resistance for β-lactams (e.g ampicillin, cephalothin, and oxacillin etc), kanamycin and gentamycin. The gene cassettes coding for trimethoprim-resistant dihydrofolate reductase were found to carry two types of dihydrofolate reductases, type A (*dfrA1, dfrA5, dfrA7, dfrA12, dfrA16, dfrA17,* two novel *dfrA* genes, *dfrA28* and *dfrA30*) and type B (*dfr-IIe*), conferring resistance to trimethoprim, were found. Type A and typeB DFRs are also called as type I and type II. Two distinct types of R-plasmid dihydrofolate reductases are known. Plasmid-borne R483 enzyme (Type I enzymes) (Sköld and Widh, 1974), are produced in amounts several fold higher than the chromosomal enzyme. The calculated I_{50} (50% inhibitory concentrations) values of trimethoprim, methotrexate, and aminopterin were found to be increased by over several thousand folds from the corresponding values obtained with the chromosomal enzyme. On the other hand, plasmid-encoded Type II dihydrofolate reductases are synthesized in about the same amount, or less, than the chromosomal counterpart. Both types of plasmid encoded dihydrofolate reductase in terms of binding of dihydrofolate, NADPH, folic acid, and 2,4-diaminopyrimidine showed little difference from that exhibited by the chromosomal enzyme (Pattishall *et al.*, 1977).

A single gene cassette (antibiotic-resistance or unrelated to antibiotic-resistance) was recognized in 65.55% (59/90) and an array of two gene-cassettes were noticed in 31% (28/90) isolates (Table 3.3 A, B and C) while only 3.3% (3/90) isolates were found to carry an array of three gene cassette on amplified amplicons. A single isolate, MB35 was the only representative which carried an array of four gene-cassettes in a single amplicon generated by CS-PCR technique. Sequence analysis of gene cassettes revealed that 7.8% isolates (MB57A, MB67, MB40A, MB74, MB75, MB61, and MB38) of the total integron positive isolate carried a combination of two or more than two gene cassettes where one was coding for unknown protein and other were coding to known antibiotic resistance. It was noted that the gene cassette coding for uncharacterized protein was either located in between the two antibiotic-resistance-gene-cassettes or at second position to the antibiotic-resistance-gene-cassette. The presence of gene-cassette coding for unknown protein, at second position, to the antibiotic-resistance-gene-cassette was probably due to their earlier recruitment or due to later reshuffling which transferred them at second position or due to acquisition of new antibiotic resistance gene-cassette for selective advantage. These types of combination also have been reported by other workers (Nemergut *et al.*, 2001; Kadlec and Schwarz, 2008) but their presence in facultatively oligotrophic bacteria hint another possibility of function related to oligotrophy and may be active at low nutrient condition to provide sustainability in poor nutrient medium. So, one can speculate that these bacteria may have the necessary genetic tool to engineer protein in the face of emergency or threat. An array of three gene cassette (*dfrA12-orf40A-aadA2*) was detected in three integron positive isolates (MB40A, MB61, and MB67) (Table 3.3 B). One bacterium MB79A was having a gene cassette array of type B (also called as typeII) *dfr-IIe (dfrB)* dihydrofolate reductase, and *arr2* gene, responsible for resistance to rifampicin (a semi-synthetic antibiotic agent) (Table 3.3 C). Maximum percentage (10%, of the total antibiotic-resistance-gene-cassettes) of two-gene-cassette array was observed for *dfrA-aadA1* type followed by *dfrA17-aadA5* (~8%). The gene encoding beta-lactamases were rarely observed in this study. Only MB82 and MB35 isolates were found to carry *oxa1* and *oxa10* sequence respectively. The *oxa1* and *oxa10* genes codes for class D enzymes which possess a distinct catalytic mechanism for the beta-lactam hydrolysis than the other β- lactamases. The *aadA* gene cassette derived from MB 18 was found 99% identical at nucleotide level to the class 1 integron borne *aad*A5 gene (Ac. No. EF105289) of *Riemerella anatipestifer*, and 98% identical at amino acid level to the aminoglycoside adenyltransferase of *Escherichia coli, Klebsiella pneumoniae,* and *Salmonella enteritidis* (Ac. No. YP190214; AF180469; and BAD02332 respectively). It depicts that the *aadA5* gene in strain MB18 (gram positive bacterium which was classified as *Brevibacterium siliguriense*) probably transferred via horizontal gene transfer from one of these source bacteria. The unoccupied integration site of class 1 integron (empty class 1 integron) which was detected in two oligotrophic bacterium, MB62 and MB63, was assumed as ancestor of resistance/nonresistance integron. The noticeable point was that the MB62 was MAR having no resistance-gene cassettes in integron structure (resistance gene functions were rendered by determinants present elsewhere in its genome) while MB63 corresponding to the same genus having class 1 integron devoid of any cassette was sensitive to all antibiotics. It is speculated that these integron structures are the future recipients of gene cassette(s) of special function to get selective advantage in response to upcoming environmental stresses. The present study also describes two new dihydrofolate reductase genes, *dfrA28* and *dfrA30,* harboured in class 1 integron from the isolate MB52 and MB45 respectively (for detail please see chapter 4 and 5). Sequence analyses of Dfr proteins exhibited that most of the proteins carried an intact "VIGngpdIPWsakg.EqllFkaiT" dihydrofolate reductase signature sequence in their sequences. The homology modelling of DfrA and AadA1 proteins derived from MB47 isolate was found similar to the dihydrofolate reductase (PDB ID: 3e0bB) and nucleotide binding domain of nucleotidyl transferase protein derived from *Haemophilus influenzae* (PDB ID: 1no5B) (Fig 2.10H to 2.10Hz and 2.10N to 2.10Nz).

Besides antibiotic resistance gene cassettes, a number of gene cassettes (~29% of the total) bearing ORFs for unrelated function than the antibiotic resistance were observed (Fig. 2.7). Reports on gene cassettes unrelated to antibiotics are very rare from culturable bacteria (Nemergut *et al.*, 2004; Stokes *et al.*, 2006; Gillings *et al.*, 2008a). Most of the uncharacterized gene cassettes have been identified from gene-cassette-metagenome sequences (Holmes *et al.*, 2003; Gillings *et al.*, 2008b). A class 1 integron borne gene cassette harbouring an ORF of a hypothetical protein similar to hypothetical protein of *Acidovorax* strain MUL2G8 was detected from bacteria *Hydrogenophaga, Imtechium,* and *Aquabacterium* (Stokes *et al.*, 2006). This ORF was also homologus to the hypothetical protein VP1784 of *Vibrio parahaemolyticus* (Gillings *et al.*, 2008a). The said authors also detected a second gene cassette containing an ORF with no homology to any known gene, ORF186 (coding for hypothetical protein) and six cassettes of which first encoded putative IS*91*-like transposition function and rest for hypothetical or unknown function from the bacteria *Hydrogenophaga,* bacterium B4 and bacterium E7 respectively. Diverse gene cassettes coding for protein function other than antibiotic resistance were noticed in the present study. The CS-PCR product sequencing of OB05 and MB83 showed that the derived sequences was 75% identical (nucleotide-nucleotide) to the part of the genome of *Acidovorax avenae* subsp. *citrulli* (Ac. No. CP000512). The feature present in that part of *Acidovorax* genome included a portion of ribosomal large subunit pseudouridine synthase B. However the predicted polypeptide of the ORF generated from OB05 DNA sequence did not produce any significant similarity with existing sequences of the protein database. Sequence analyses of the amplicons of MB05 and MB70 revealed the presence of a truncated ORF of 64 amino acids that showed 73% identity with the methyl transferase protein from *Ralstonia eutropha* (Ac no. AAZ62060). Partial sequencing of the CS-PCR product obtained from MB12 showed the presence of a single ORF (*livM1*) of length 154 amino acids (Fig 2.8A). BlastP analysis (Table 2.4) of this putative polypeptide product exhibited best scores with ABC type branched chain amino acid transport system with 72% identity at amino acid level to the ABC transporter permease of bacterium *Aromatoleum aromaticum* EbN1 (Ac. No. NC_006513). The multiple sequence alignment of LivM1 derived from amplicon sequence of bacterium MB12 with its neighbouring ABC transporter proteins is shown in Fig. 2.8B. The transmemebrane (TM) regions predicted within the ABC transporter protein of MB12 are shown in Fig. 2.8C and 2.8D. Analysis revealed the presence of four transmembrane regions in the predicted LiVM1 protein of MB12. The predicted transmembrane (TM) regions (AA position 2-20, 50-74, 87-110, and 119-138) in ABC transporter protein of bacterium MB12 (Protein ID: CAQ53856) are presented in table 2.5. The conserved domain search for putative translated product obtained from translated ORF of MB12 integron sequence revealed that the predicted protein was one of the members of two TM subunits which play role in uptake of branched chain amino acids (Fig 2.8E). Mapping of signature motifs (Fig 2.8F) showed that 19 of 19 [TM_PBP1_LivM_like, a type of transporters which consist a PBP (periplasmic binding protein), two TMs (trans-membrane subunits) and two cytoplasmic ABCs (ATP-binding cassette)] residues which were recognized on the TM-ABC transporter protein sequences of other bacterial species were also found conserved on ABC protein derived of MB12. Hence, it was assumed that these types of transporters probably play a vital role in the import of solutes from environment. Finding of these transporters in class 1 integrons has a greater significance since oligotrophic bacteria thriving in low nutrient milieu may require improved transporter molecules to concentrate and channelize nutrients from outside to the interior of the cells. The theoretical pI and instability index of the putative ABC transporter of MB12 was computed 9.3 and 19.13. The protein was predicted as stable protein with an estimated half life of 10 h (*in vivo* in respect to *E coli*) with a high aliphatic index of 127.21. MB16 was found to carry gene cassette carrying *appA* gene encoding bacterial extracellular solute binding protein. ProtParam computing showed that the extracellular binding protein derived from bacterium MB16 was stable and its instability index was computed 18.94. The GRAVY (grand average hydropathicity), aliphatic index

and theoretical pI was computed -0.087, 78.5, and 9.26 respectively. Psortb, a tool for subcellular localization prediction showed that the translated product of *appA* gene was periplasmic with localization score of 9.44.

The Complete sequence of 704 bp amplicon of MB 40B (Table 2.3B) yielded no significant homology with any nucleotide sequence available in the database. The same sequence yielded a complete ORF of 102 amino acids and BlastP analysis of this protein sequence revealed 58% identity with helicase domain protein of *Verminephrobacter eiseniae* EF01 (Ac. No. ABM585806). The sequences derived from amplicons of OB 12, MB 08, MB41, MB 44, MB 48, MB 51, MB 54, MB55, MB80 and MB81 did not produce any significant homology with any protein sequences available in the database. All of them were characterized as hypothetical proteins. All these polypeptides shared very low level of identity (2.03 to 25.19%) among either with each other or to pre-deposited hypothetical protein sequences. Some of the predicted proteins were found homologous to certain protein sequences present in database, for example, translated product of bacterium MB54 was carrying a conserved domain belonged to the UPF0153 super family(Fig. 2.9A and B). The proteins of this family contain 8 conserved cysteines, hence it was expected that they may constitute a metal binding site and therefore might play a role in metal sequestering like Fe-S cluster as part of oxido-reductase complex. Molecular iron is an important cofactor for growth, because it is used by bacteria during aerobic respiration, DNA replication, and other biological processes. Due to requirement of iron as a cofactor, bacteria must have to develop mechanisms for the acquisition of iron from their environment, for example some bacteria secrete siderophores molecules to bind free iron and then the iron molecules are captured by receptors present on bacterial surface (surface receptors that transport iron for transport across the cytoplasmic membrane). However iron also plays a crucial role in establishment of host pathogen relation that has been proved by many researchers. Such proteins in oligotrophic bacteria might play an important role in concentrating essential elements from the nutrient limited environment.

Since all the earlier studies on integrons were focused on copiotrophic bacteria (grow in rich nutrient media) while no such study was found concerning oligotrophic bacteria, a focused study was required to fill this gap. The present study uncovers the diversity of gene cassettes associated with class 1 integrons in oligotrophic bacteria of riverine origin.

2.5. Conclusion

The results have revealed that Class 1 integrons are prevalent in facultatively oligotrophic bacteria irrespective of being resistant or sensitive to antibiotics. Oligotrophic bacteria isolated from River Mahananda are good source of novel genes as well as potential reservoir of antibiotic gene cassettes. The high percentage of occurrence of MAR bacteria (and high incidence of class 1 integron in them) at sampling station II (where Mahananda flows within the Siliguri city) is possibly due to selection induced by pollution generated from anthropogenic activities. The presence of class 1 integrons in oligotrophic bacteria is indicative of the spread of stress-combating-genes including antibiotic resistance and other putative genes coding for known or unknown function in the aquatic environment. The predominance of aminoglycoside and trimethoprim resistance gene cassettes are probably sign of abuse of these drugs or the high frequency of horizontal gene transfer in that environment. This study has also demonstrated that integron typing can be a useful tool for studying the dissemination of resistance genes among oligotrophic bacteria. 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 as shown by earlier studies (Rowe-Magnus and Mazel, 2001; Holmes *et al.*, 2003; Mazel, 2006). The study of integrons and their associated gene cassettes could provide information about antibiotics which should be used most carefully to prevent further dissemination of resistance. Hence, now it has become more relevant to design a tighter antibiotic policy or other measures that can lower the amount of horizontal transfer in different environmental settings.

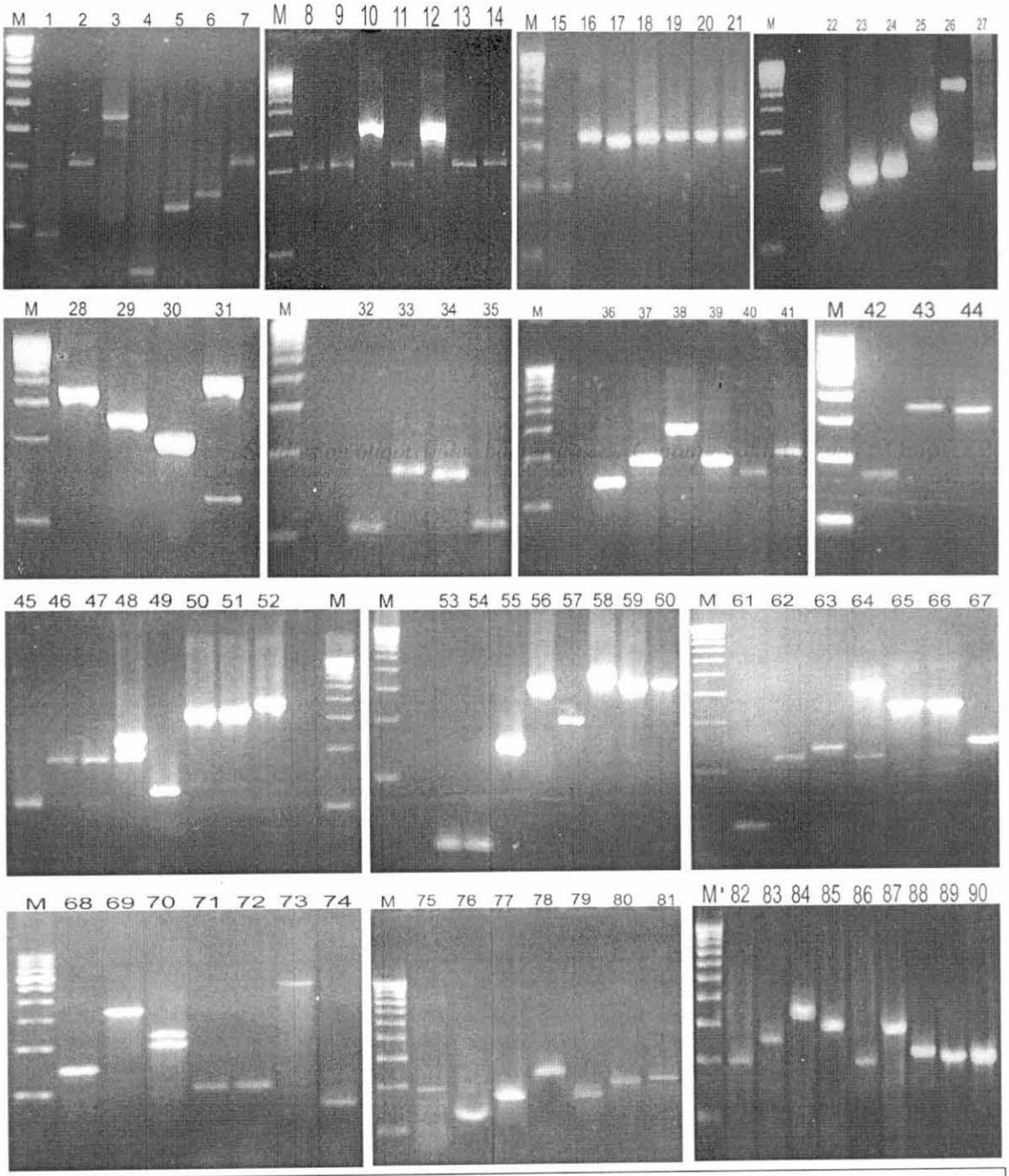


Fig. 2.5: Variable regions of class 1 integrons amplified using 5' CS and 3' CS primers. The PCR products were electrophoresed on 1% agarose gel.

Lane: M, 500 bp DNA ladder; 1, OB05; 2, OB12; 3, MB03; 4, MB05; 5, MB08; 6, MB09; 7, MB12; 8, MB16; 9, MB18; 10, MB19; 11, MB20; 12, MB21; 13, MB22; 14, MB23; 15, MB24; 16, MB25; 17, MB26; 18, MB27; 19, MB28; 20, MB29; 21, MB30; 22, MB31; 23, MB32; 24, MB33; 25, MB34; 26, MB35; 27, MB36; 28, MB37; 29, MB38; 30, MB39; 31, MB40; 32, MB41; 33, MB42; 34, MB43; 35, MB44; 36, MB45; 37, MB46; 38, MB47; 39, MB48; 40, MB49; 41, MB50; 42, MB51; 43, MB52; 44, MB53; 45, MB54; 46, MB55; 47, MB56; 48, MB57; 49, MB58; 50, MB59; 51, MB60; 52, MB61; 53, MB62; 54, MB63; 55, MB64; 56, MB65; 57, MB66; 58, MB67; 59, MB68; 60, MB69; 61, MB70; 62, MB71; 63, MB72; 64, MB73; 65, MB74; 66, MB75; 67, MB76; 68, MB77; 69, MB78; 70, MB79; 71, MB80; 72, MB81; 73, MB82; 74, MB83; 75, MR 01; 76, MR 02; 77, MR 03, 78, MR 04; 79, SR19; 80, NV66; 81, OD 05; 82, OD 08; 83, OD 10; 84, OD 16; 85, OD 24; 86, OC 74; 87, OC 75; 88, OC78; 89, OD 21; 90, OD 24

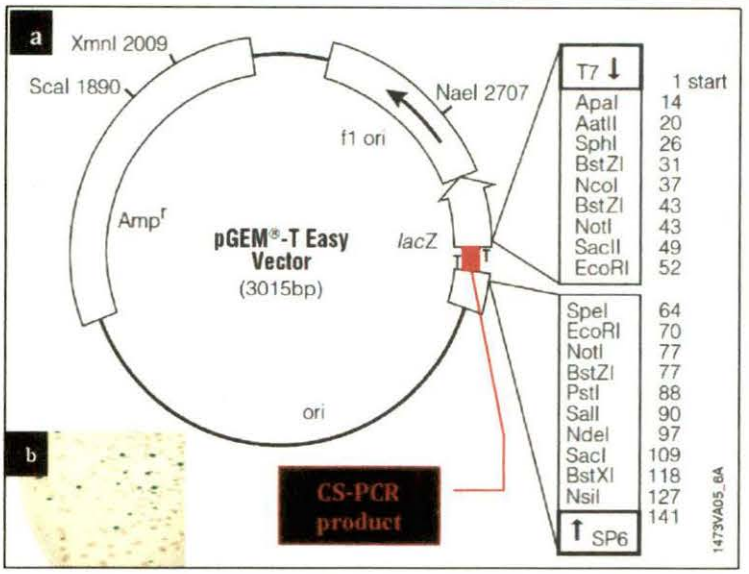


Fig. 2.6: a, Cloning of CS-PCR product in pGEMT-Easy vector and b, screening (Blue-white) of recombinant clone on Luria agar plate amended with ampicillin (100mg/L). Gene of interest ligated in the multiple cloning site of T-vector is shown in red colour.
Amp^r, gene for ampicillin resistance; *ori*, orgin of replication; *lacZ*, gene for beta-galactosidase synthesis.

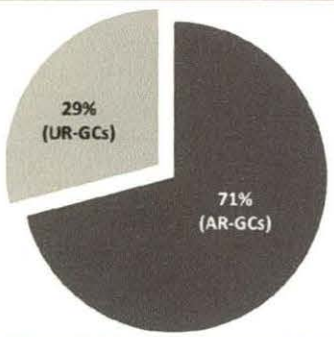


Fig. 2.7: Distribution of gene cassettes identified within the amplified variable regions of class 1 integron.

UR-GCs, cassette unrelated to antibiotics resistance;
AR-GCs, cassette related to antibiotic resistance.

```
TAAGCAGACTTGACCTGACCGACAACACCCACatgtacctgttcgtactggtcatctttctggcgggctt
                          livM1->  M  Y  L  F  V  L  V  I  F  L  A  G  F
cctgctgctgtggcgcgtggttcattcacccttggccagatcctcaaggcgattcgtgaacatcaggac
  L  L  L  W  R  V  V  H  S  P  F  G  Q  I  L  K  A  I  R  E  H  Q  D
cgcgccacgtccccgggctatgacaccgacaccttcaagctgctggcctttgtcatctcgggcgcactgg
R  A  T  S  P  G  Y  D  T  D  T  F  K  L  L  A  F  V  I  S  G  A  L  A
ccgggacggccggagcgaccaaggcgctggtgtttcagctggcctcgctcaccgatgtgcactggagcat
  G  T  A  G  A  T  K  A  L  V  F  Q  L  A  S  L  T  D  V  H  W  S  M
gtcgggcgaggtggtgctgatgaccctggtgggtggcatgggcacgctgtttggtccggtggctggtgcg
  S  G  E  V  V  L  M  T  L  V  G  G  M  G  T  L  F  G  P  V  A  G  A
gccgtgatcgtgtccatgcagaactacctggcacagctgggggcatgggtgacggtggtgcagggtgtga
A  V  I  V  S  M  Q  N  Y  L  A  Q  L  G  A  W  V  T  V  V  Q  G  V  I
tcttcgtggtctgcgtgctcgccttccgccgcggcatcattggcgagatcgccaacctcatcaagaaacc
  F  V  V  C  V  L  A  F  R  R  G  I  I  G  E  I  A  N  L  I  K  K  P
cctctgaCCGGTTGGTGCCATGGGCCCGGCTAGACTGCAGGTTTTCCTGTAGTCGGCGCCTCGCATGTCT
 L  #(stop)
TCTTCCCGCGTTTCCGGTTTGTGGCTGATCAGCCTGGCCGCTTTCAGCAGCATGGCTTCCATGCGCGTGT
GCGACCCGATGCTGGTGGCGCTCGCGACGGAGTTCCAGGTCAGCGTGGGGGAGGCCTCTCGTGTGATTTC
GGCGT
```

Fig. 2.8A: Partially sequenced variable region (Ac. No. AM997273) of class 1 integron derived from MB12 showing ORF and the putative translated product.

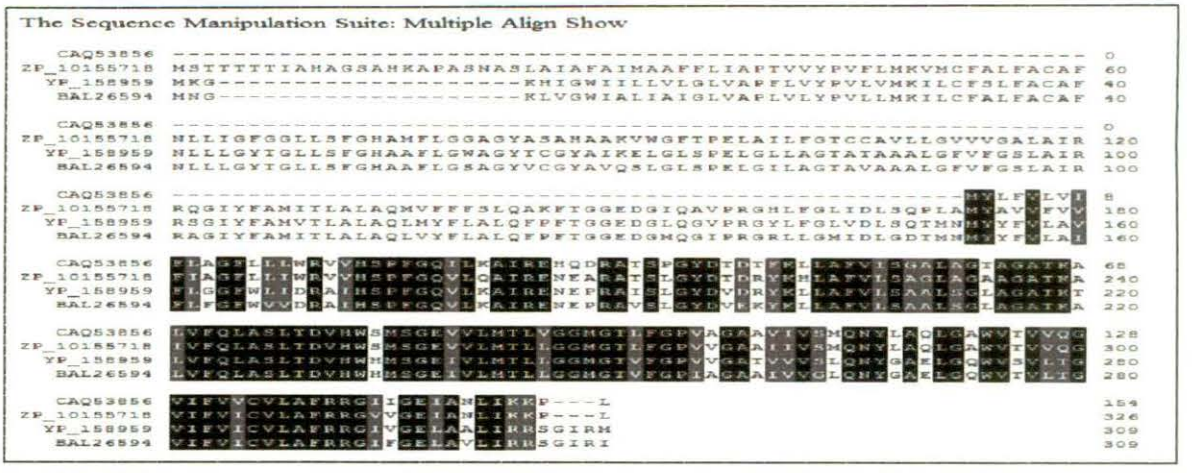


Fig. 2.8B: Multiple sequence alignment between ABC transporter protein obtained from MB12 and the nearest neighboring sequences. Black, identical amino acids and gray, similar amino acids present in all the protein sequences.

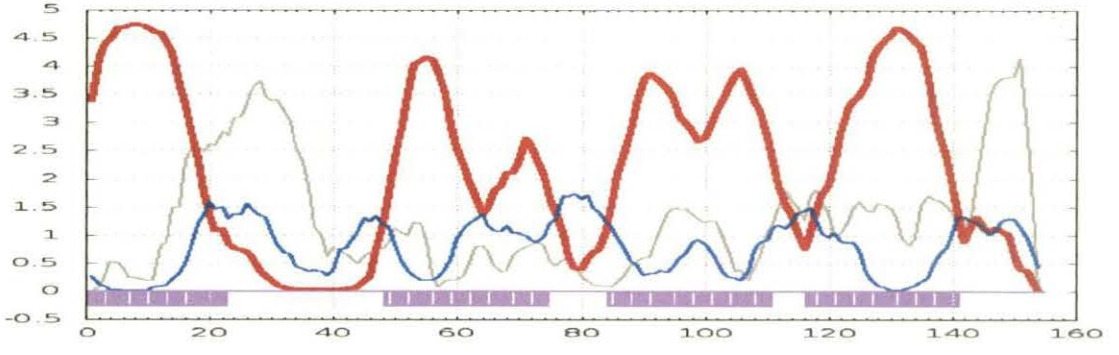


Fig. 2.8C: Graphical presentation of tans-membrane (tm) region within LivM1 protein of MB12.
Colour coding: Red line: Transmembrane helix preference; Blue line: Beta preference; Gray line: Modified hydrophobic moment index; Violet boxes (below abscisa): Predicted trans-membrane helix position.

Predicted amino acid locations of Tm regions: 1-22, 49-74, 85-110, and 117-140 (Protein ID: CAQ53856)

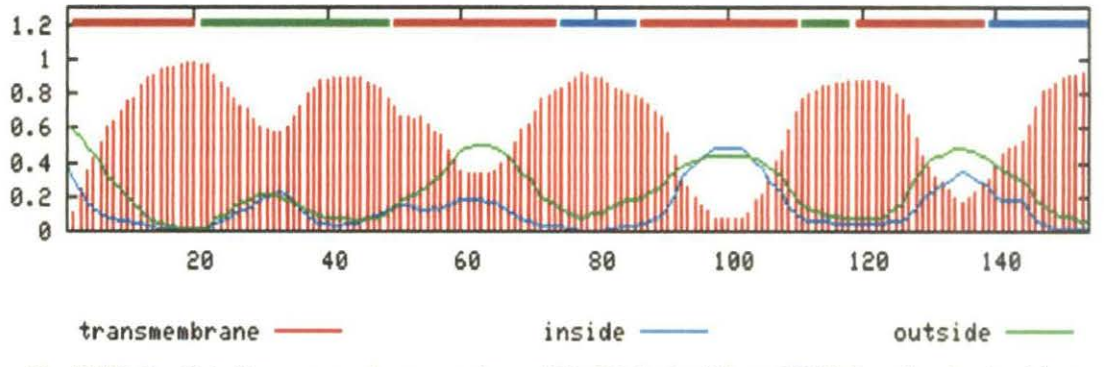


Fig. 2.8D: Predicted trans-membrane regions of LivM1 derived from MB12 (results obtained from Protein Topology Prediction Hidden Markov Model TMMOD tool)

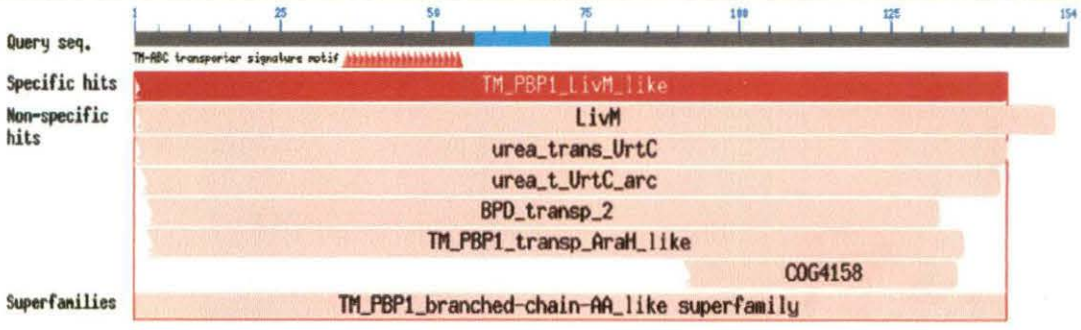


Fig. 2.8E: Result of conserved domain search for the putative translated product obtained from ORF of MB12

Description of CDD search: Result and feature of Specific hit of query sequence with the existing database (cd06581): Transmembrane subunit (TM) of *Escherichia coli* LivM and related proteins. LivM is one of two TMs of the *E. coli* LIV-1/LS transporter, a Periplasmic Binding Protein (PBP)-dependent ATP-Binding Cassette (ABC) transporter involved in the uptake of branched-chain amino acids (AAs). These types of transporters generally bind type 1 PBPs. PBP-dependent ABC transporters consist of a PBP, two TMs, and two cytoplasmic ABCs, and are mainly involved in importing solutes from the environment. The solute is captured by the PBP, which delivers it to a gated translocation pathway formed by the two TMs. The two ABCs bind and hydrolyze ATP and drive the transport reaction. *E. coli* LivM forms a heterodimer with another TM, LivH, to generate the transmembrane pore. LivH is not included in this subgroup. The LIV-1/LS transporter is comprised of two TMs (LivM and LivH), two ABCs (LivG and LivF), and one of two alternative PBPs, LivJ (LIV-BP) or LivK (LS-BP). In addition to transporting branched-chain AAs including leucine, isoleucine and valine, the *E. coli* LIV-1/LS transporter is involved in the uptake of the aromatic AA, phenylalanine

Signature motif on conserved domain

```
Feature 1                                                  ********************
gi 585411     118 .[150].ALLLVVLSLFVINPLLRMPLGRAWEALRED.[2].IACRSLGLSPRRIKLTAFTISAAFAGFAGTLFAARQGF 336
query           1 .[  4].VLVIFLAGFLLLWFVVHSPFGQILKAIREH.[2].DFATSPGYDTDTFKLLAFVISGALAGTAGATKALVFQL  73
gi 27378466    32 .[145].VLVVALIAYLLIARLVHSHFGRAMRGLMEN.[1].TLAVSVGIDVTKTLTLAAVISAAIAGAAGSLYAHYIRI 245
gi 91786264    20 .[156].CLAVAIGSFAAFHLLVSSPFGRALQAIKLD.[2].DAARASGINVTLYKSQAFVTAAVFTGWAGALFTYHNQY 214
gi 55380262   361 .[141].ALLIAVLSFLFARPMMNAPFGSVLQSIRES.[1].ERAEFIGYDVNRYKRRAFVISGGMAGLAGGLLAVSPST 570
gi 51893498   350 .[127].TLTLLLGGIWVAEPLSGSYLGRALRAVRED.[1].VAARAYGVGLERYKSIAFGLSAFIAGLAGVLLAHAYTY 545
gi 15792343    38 .[136].TGGIAIVAVILILNIVYSKYGRAMKAIRDD.[1].DAASAMGINTFWIRTLAFSTSAFLEGVGGLLACLLTT 242
gi 154245652   54 .[133].ILAMLALAILATMVMERSKLGYCLKAIRGD.[1].DAAEALGISPTRYTQAAFAISAFMTALGGSFYAQFFRY 255
gi 92113544    17 .[134].VALFVAGVYFLVERAYASPWGPVLRAIREN.[1].PATAAAGKSIAGFFLQAFVLGSAIMGLGGGLYAHFFGF 219
gi 91974534    45 .[127].TLAIFMVALWLIWRITNSPFGSLLTAIKSN.[1].VRTQYLGYDTALYKLAAIVISGSFSGLAGGLYAWFQYA 240

Feature 1
gi 585411     337 VSP.[1].SFTFAESAFVLAIVVLGGM.[1].SQFAVILAAILLVVSRELMRDF.[ 4].MLMLGGLMVLMMIWRPQGLL 406
query          74 ASL.[1].DVHWSMSGEVVLMTLVGGM.[1].TLFGPVAGAAVIVSMQNYLAQL.[ 4].TVVQGVIFVVCVLAFRRGII 143
gi 27378466   246 IDP.[1].VEAFINTVTMVIMVISGGK.[1].SLAGPVVGGLIFGLLPVALRPI.[ 6].WIAYGGVLIAILFVLPRGIV 317
gi 91786264   215 LNP.[1].DFTFWKSIEIVIMVAVGGI.[1].SLAGAVFGAFAVTLLPEVLGDL.[ 4].MIIFGSLVVLCMGLGHSGLA 284
gi 55380262   571 VIV.[4].TLNWIHSGEVIVIALFGGM.[1].TLYGPMIGSGVFFGAEEFLSSY.[ 4].RLIIGTMFILFVLFVPRGIV 643
gi 51893498   546 IAP.[1].TFTHAMSLQVLTMVVLGGL.[1].NTIGAVAGALALIVIPESARFL.[ 4].FLVYGLLLLVMLRFRPRGLL 615
gi 15792343   243 VSP.[1].QFDFLLTFQLLIIIVLGGL.[1].STTGAIIGAILVIGGSEWLRFL.[11].MVVFSIVLILVMLFARKGIM 329
gi 154245652  256 INP.[2].NMGLELSIDMALMSIIGGQ.[1].TPFGPLVGALLLMPLGEISRGY.[ 5].LVIYGVVLMITVLYFPKGLI 331
gi 92113544   220 LSP.[2].FMPLYGTFLVWVMLIAGGS.[1].NNRGAILGAVLVWGVWSGTELL.[16].VLLIGVLLQVILVTRPEGVL 302
gi 91974534   241 AYP.[1].NLFWIESGNIVILTLLGGG.[2].SFFGPILGAAVFVGAQDLISGY.[ 4].MFFFGLIFIVVVTTFPNGLP 311

Feature 1
gi 585411     407 P 407
query         144 G 144
gi 27378466   318 P 318
gi 91786264   285 G 285
gi 55380262   644 S 644
gi 51893498   616 S 616
gi 15792343   330 G 330
gi 154245652  332 A 332
gi 92113544   303 P 303
gi 91974534   312 E 312
```

Fig. 2.8F: TM-ABC transporter signature motif on conserved domain TM_PBP1_LivM_like. 19 of 19 of the residues that compose this conserved feature have been mapped in the query sequence.

```
GGCATCCAAGCAGCAAGTACCCGACTGTTTACAACTAACACCGGAAACTTTAAAAACCATACACTGG
                              *(start codon)
TTGCCATCGAGTTGTGCCTATAAACGTttgccatcgagttgtgcctataaacgtttgaatgaaggca
            1R                L  P  S  S  C  A  Y  K  R  L  N  E  G  K
aaaatttaccttcttggcactatttaaatacgggttccagacaaagcgtagtgaaagcgagaaagtc
  N  L  P  S  W  H  Y  L  N  T  G  S  R  Q  S  V  V  K  A  R  K  S
ggtagcagggcgttgtattcctgaaactgatgtatatgaagatgatatcgaagattatgtggtgcgt
 V  A  G  R  C  I  P  E  T  D  V  Y  E  D  D  I  E  D  Y  V  V  R
tgggtgcgttaaTACGCTCCCGATGAGCATTTTTAATTACAGCGTATAATAAAGTAACATTTTTATC
W  V  R  #(stop codon, end of ORF)                          1L
GTAAATTTTTTAATCATCTCCGTCATATAAAAATAAAAGCAAAAGGAGATAAAAATGATGAGACGAT
TAGCAGCCCCCTTATTATGTAGTAGTTTCTTCTTG//TTAATGGCCTGTGGCTCTAATAATACTAAT
TCAAAAAGTCCTGAACAAAATACAACTACAAAAACAGAGCAGAAGAGTGCTGCCAAACAAATTTATC
AGGTCAAGTCTGCTT
```

Fig. 2.9A: Nucleotide sequence (Ac. No. FM955254) and translated product of variable region obtained from MB54. The ORFs, putative translated product and other features of class 1 integron borne gene cassette have been shown in the sequence. Symbols: *, start codon; #, stop codon; 1R, core site; //, beginning of 3' Conserved segment

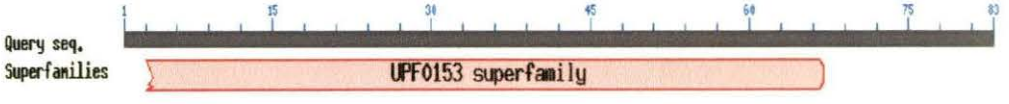


Fig. 2.9B: Result of conserved domain search for the putative translated product obtained from ORF of MB54.

Description of CDD search: This family of proteins contains 8 conserved cysteines that may form a metal binding site. The function of these proteins is unknown but presumably Fe-S cluster as part of an oxidoreductase complex.

```
GGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACG
                              *(start)
CAGCAGGGCAGTCGCCCTAAAACAAAGatgaaactatcactaatggtagctatatcgaagaatggagtta
                  dfrA1->   M  K  L  S  L  M  V  A  I  S  K  N  G  V  I
tcgggaatggccctgatattccatggagtgccaaaggtgaacagctcctgtttaaagctattacctataa
  G  N  G  P  D  I  P  W  S  A  K  G  E  Q  L  L  F  K  A  I  T  Y  N
ccaatggctgttggttggacgcaagacttttgaatcaatgggagcattacccaaccgaaagtatgcggtc
 Q  W  L  L  V  G  R  K  T  F  E  S  M  G  A  L  P  N  R  K  Y  A  V
gtaacacgttcaagttttacatctgacaatgagaacgtagtgatctttccatcaattaaagatgctttaa
V  T  R  S  S  F  T  S  D  N  E  N  V  V  I  F  P  S  I  K  D  A  L  T
ccaacctaaagaaaataacggatcatgtcattgtttcaggtggtggggagatatacaaaagcctgatcga
  N  L  K  K  I  T  D  H  V  I  V  S  G  G  G  E  I  Y  K  S  L  I  D
tcaagtagatacactacatatatctacaatagacatcgagccggaaggtgatgtttactttcctgaaatc
 Q  V  D  T  L  H  I  S  T  I  D  I  E  P  E  G  D  V  Y  F  P  E  I
cccagcaattttaggccagtttttacccaagacttcgcctctaacataaattatagttaccaaatctggc
P  S  N  F  R  P  V  F  T  Q  D  F  A  S  N  I  N  Y  S  Y  Q  I  W  Q
aaaaggttaaCAAGTGGCAGCAACGGATTCGCAAACCTGTCACGCCTTTTGTACCAAAAGCCGCGCCAG
  K  G  #(stop)          1R      *
GTTTGCGATCCGCTGTGCCAGGCGTTAAACATCatgagggaagcggtgatcgccgaagtatcgactcaac
                          aadA1->M  R  E  A  V  I  A  E  V  S  T  Q  L
tatcagaggtagttggcgtcatcgagcgccatctcgaaccgacgttgctggccgtacatttgtacggctc
  S  E  V  V  G  V  I  E  R  H  L  E  P  T  L  L  A  V  H  L  Y  G  S
cgcagtggatggcggcctgaagccacacagtgatattgatttgctggttacggtgaccgtaaggcttgat
 A  V  D  G  G  L  K  P  H  S  D  I  D  L  L  V  T  V  T  V  R  L  D
gaaacaacgcggcgagctttgatcaacgacctttggaaacttcggcttcccctggagagagcgagattc
E  T  T  R  R  A  L  I  N  D  L  L  E  T  S  A  S  P  G  E  S  E  I  L
tccgcgctgtagaagtcaccattgttgtgcacgacgacatcattccgtggcgttatccagctaagcgcga
  R  A  V  E  V  T  I  V  V  H  D  D  I  I  P  W  R  Y  P  A  K  R  E
actgcaatttggagaatggcagcgcaatgacattcttgcaggtatcttcgagccagccacgatcgacatt
 L  Q  F  G  E  W  Q  R  N  D  I  L  A  G  I  F  E  P  A  T  I  D  I
gatctggctatcttgctgacaaaagcaagagaacatagcgttgccttggtaggtccagcggcggaggaac
D  L  A  I  L  L  T  K  A  R  E  H  S  V  A  L  V  G  P  A  A  E  E  L
tctttgatccggttcctgaacaggatctatttgaggcgctaaatgaaaccttaacgctatggaactcgcc
  F  D  P  V  P  E  Q  D  L  F  E  A  L  N  E  T  L  T  L  W  N  S  P
gcccgactgggctggcgatgagcgaaatgtagtgcttacgttgtcccgcatttggtacagcgcagtaacc
 P  D  W  A  G  D  E  R  N  V  V  L  T  L  S  R  I  W  Y  S  A  V  T
ggcagaatcgcgccgaaggatgtcgctgccgactgggcaatggagcgcctgccggcccagtatcagcccg
G  R  I  A  P  K  D  V  A  A  D  W  A  M  E  R  L  P  A  Q  Y  Q  P  V
tcatacttgaagctagacaggcttatcttggacaagaagaagatcgcttggcctcgcgcgcagatcagtt
  I  L  E  A  R  Q  A  Y  L  G  Q  E  E  D  R  L  A  S  R  A  D  Q  L
                                                        1L
ggaagaatttgttcactacgtgaaaggcgagatcaccaaggtagtcggcaaataTGTCTAACAATTCGT
 E  E  F  V  H  Y  V  K  G  E  I  T  K  V  V  G  K  #(stop)
TCAAGCCGACGCCGCTTCGCGGCGCGGCTTAACTCAAGCG//TTAGATGCACTAAGCACATAATTGCTCA
CAGCCAAACTATCAGGTCAAGTCTGCTT
```

Fig. 2.10A: Nucleotide sequence (Ac. No FM179327) of variable region obtained from MB47 isolate showing ORFs, putative translated product and other features of class 1 integron borne gene cassette.
Symbols: *, start codon; #, stop codon; 1R, core site; 1L, inverse core site; //, beginning of 3' Conserved segment

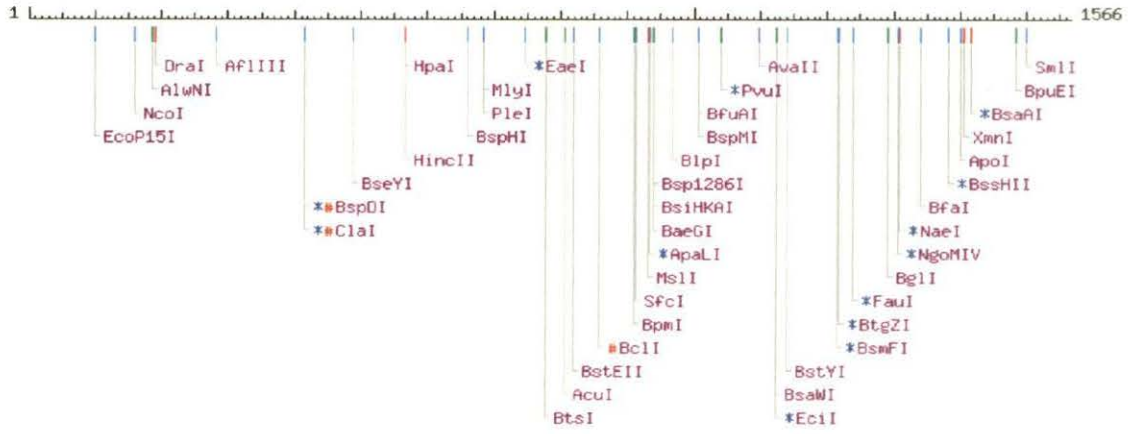


Fig. 2.10B: Restriction map of 1566 bp nucleotide sequence CS-PCR product of MB47.
Symbols: *, cleavage affected by CpG methylation; and #, cleavage affected by other methylations

Fig. 2.10C: Multiple sequence alignment of DfrA1 protein derived from MB47 and it's nearest neighboring sequences. Black, identical amino acids and gray, similar amino acids present in all the Dfr proteins.

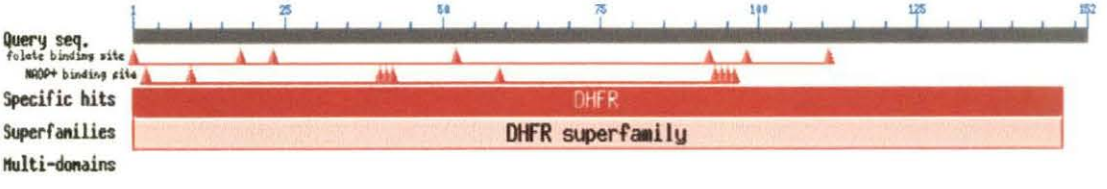
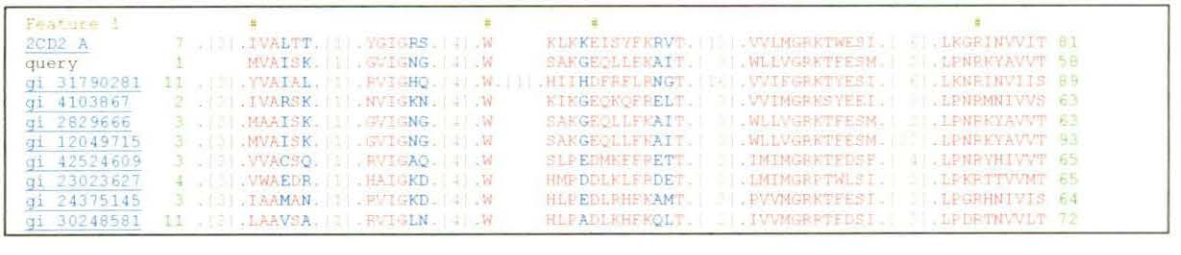
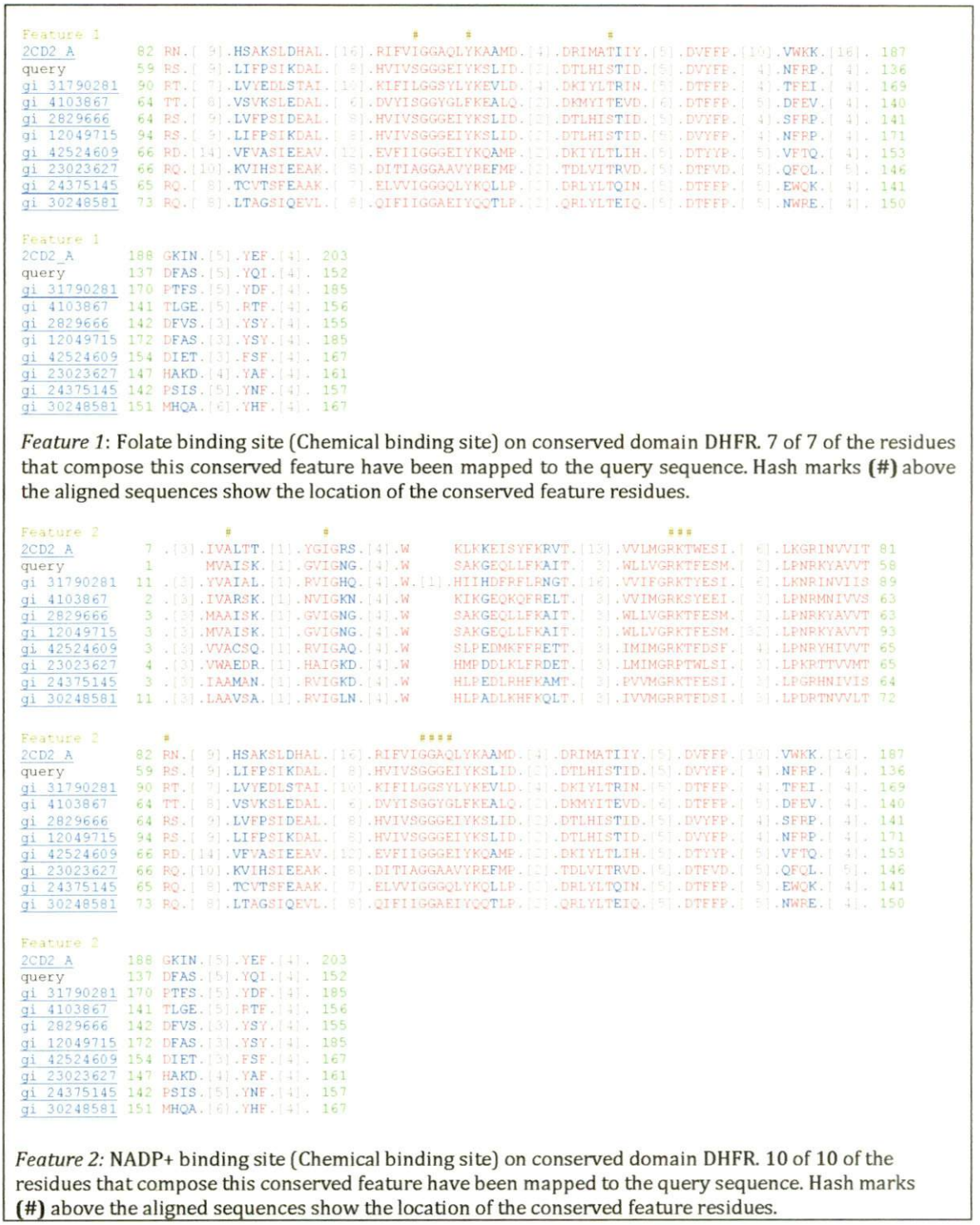


Fig. 2.10D: Detected putative conserved domain in translated product of *orf1* derived MB47 isolate.

Description of CDD search: [Specific hit] cd00209, Dihydrofolate reductase (DHFR). Reduces 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate with NADPH as a cofactor. This is an essential step in the biosynthesis of deoxythymidine phosphate since 5,6,7,8-tetrahydrofolate is required to regenerate 5,10-methylenetetrahydrofolate which is then utilized by thymidylate synthase. Inhibition of DHFR interrupts thymidilate synthesis and DNA replication, inhibitors of DHFR (such as Methotrexate) are used in cancer chemotherapy. 5,6,7,8-tetrahydrofolate also is involved in glycine, serine, and threonine metabolism and aminoacyl-tRNA biosynthesis.





Feature 1: Folate binding site (Chemical binding site) on conserved domain DHFR. 7 of 7 of the residues that compose this conserved feature have been mapped to the query sequence. Hash marks **(#)** above the aligned sequences show the location of the conserved feature residues.

Feature 2: NADP+ binding site (Chemical binding site) on conserved domain DHFR. 10 of 10 of the residues that compose this conserved feature have been mapped to the query sequence. Hash marks **(#)** above the aligned sequences show the location of the conserved feature residues.

Fig. 2.10E: Features that mapped on query sequence in CDD search

Computation determination (using ProtParam tool available at, http://web.expasy.org/protparam/) of various physical and chemical parameters of DfrA1 protein of isolate, MB47 coding for trimethoprim resistance

Computed parameters of DfrA1 are following:

Number of amino acids: *152*

Molecular weight: *16974.2*

Theoretical pI: *5.64*

Total number of negatively charged residues (Asp + Glu): 16

Total number of positively charged residues (Arg + Lys): 14

Ext. coefficient 25440 M^{-1} cm^{-1}, at 280 nm measured in water.
Abs 0.1% (=1 g/l) 1.499
Estimated half-life:
The N-terminal of the sequence considered is M (Met).
The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).
Instability index:
The instability index (II) is computed to be 29.84
This classifies the protein as stable.

Aliphatic index: 90.99
Grand average of hydropathicity (GRAVY): -0.163

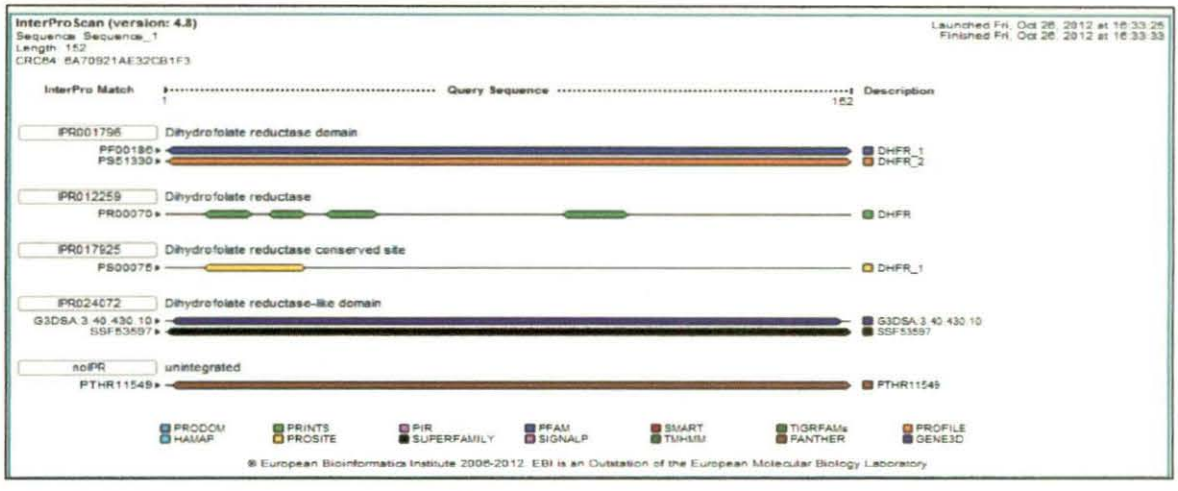


Fig. 2.10F. InterPro's signatures of DfrA1 protein derived from MB47 isolate. The map generated through InterProScan software package.

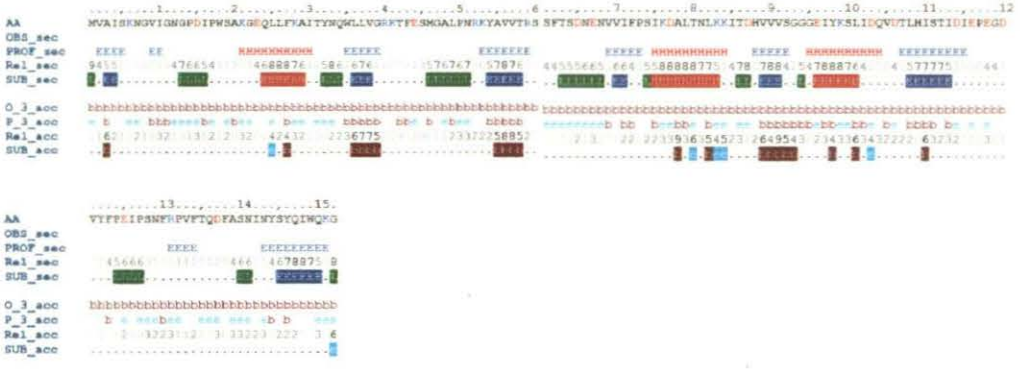


Fig 2.10G: Predicted secondary structures in DfrA1 protein coding for trimethoprim resistant dihydrofolate reductase from bacterium, MB47.

Abbreviations:

AA:	amino acid sequence
OBS_sec:	observed secondary structure [H=helix, E=extended (sheet), blank=other (loop)]
PROF_sec:	predicted secondary structure [[H=helix, E=extended (sheet), blank=other (loop)] PROF: profile network prediction HeiDelberg
Rel_sec:	reliability index for PROF secondary prediction [0=low to 9=high]
SUB_sec:	subset of PROF secondary prediction for all residues with an expected average accuracy >82% (please see table) [L=loop]
O_3_sec:	observed relative solvent accessibility (acc) in 3 states, b=0-9%; i=9-36%; e=36-100%
P_3_sec:	PROF predicted relative solvent accessibility (acc) in 3 states, b=0-9%; i=9-36%; e=36-100%

Homology modeling (automated mode software, http://swissmodel.expasy.org/) of DfrA1 protein derived from class 1 integron borne gene cassette from a facultatively oligotrophic bacterium, MB47.

Fig. 2.10H: graphical presentation of pairwise alignment between target protein dihydrofolate reductase derived from isolate, MB47 (green bar) and *Bacillus anthracis (blue bar)*.

Model information: Modeled residue range: 1 to 151 Based on template: [3e0bB] (2.25 Å) **Quaternary structure information:** Template (3e0b): Monomer Model built: Single chain **Ligand information:** Ligands in the template: N22: 1, NAP: 1. Ligands in the model: none.	

Fig. 2.10Hx: 3D model (ribbon model) of the template: 3e0bB (crystal structure at resolution2.25 Å of dihydrofolate reductase derived from *Haemophilus influenzae*)

Model information: Template: [3e0bB] (2.25 Å) Modeled residue range: 1 to 151 Sequence Identity [%]:30.57 Evalue: 0.00e-1 Quality information: QMEAN score: 0.7 Z-Score: -1.01 Predicted LGscore : 3.971 Predicted MaxSub : 0.455	

Fig 2.10Hy: 3D model (ribbon model) of the DfrA1 protein of bacterium MB47

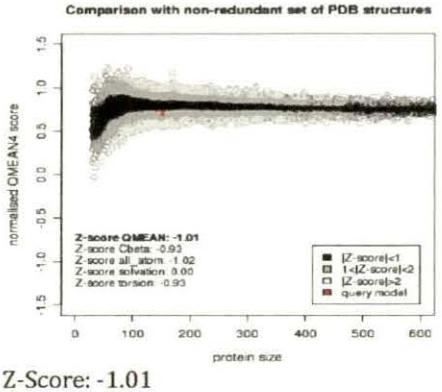
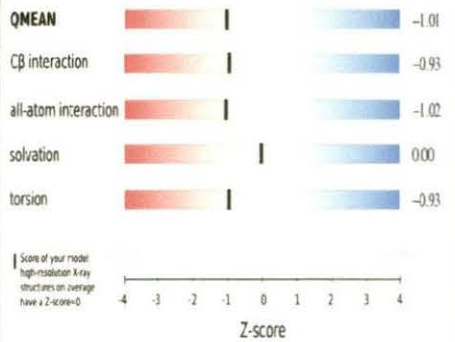
QMEAN4 score	Estimated absolute model quality	Score components
0.7 (Range for reliable model: 0-1)	 Z-Score: -1.01	

Fig 2.10Hz: Global model quality estimation of DfrA1 protein of bacterium MB47.

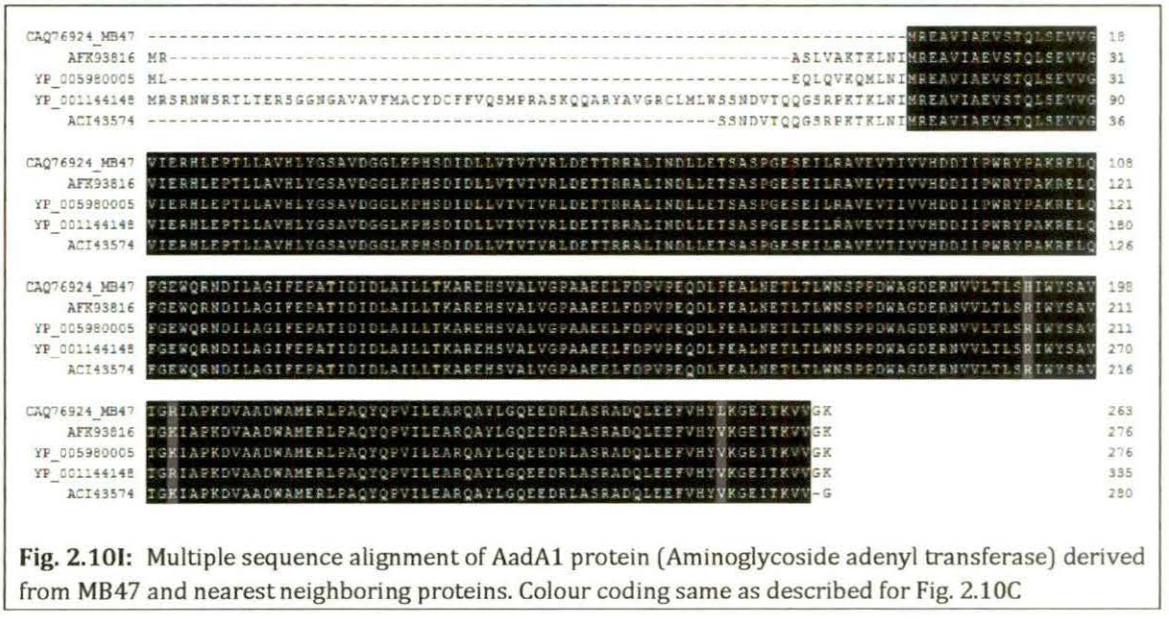


Fig. 2.10I: Multiple sequence alignment of AadA1 protein (Aminoglycoside adenyl transferase) derived from MB47 and nearest neighboring proteins. Colour coding same as described for Fig. 2.10C

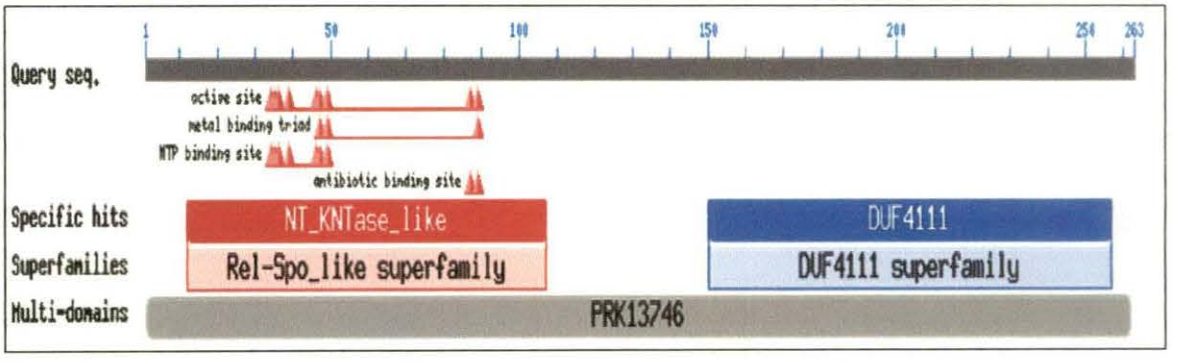
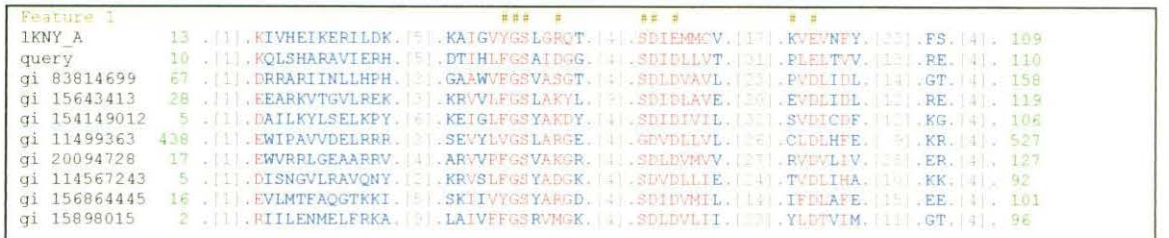
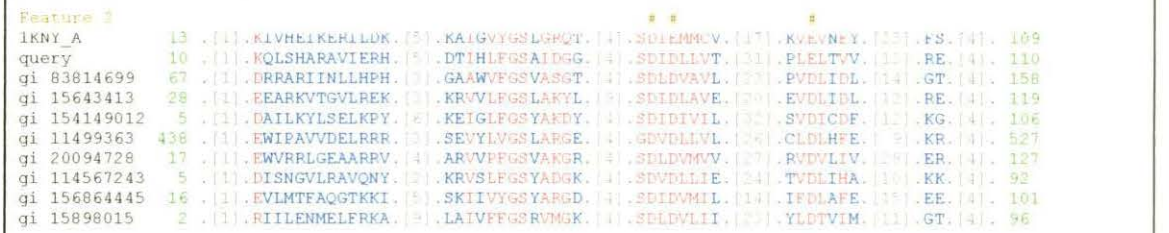


Fig. 2.10J: Detected putative conserved domain in translated nucleotide sequence derived from MB47 bacterium.

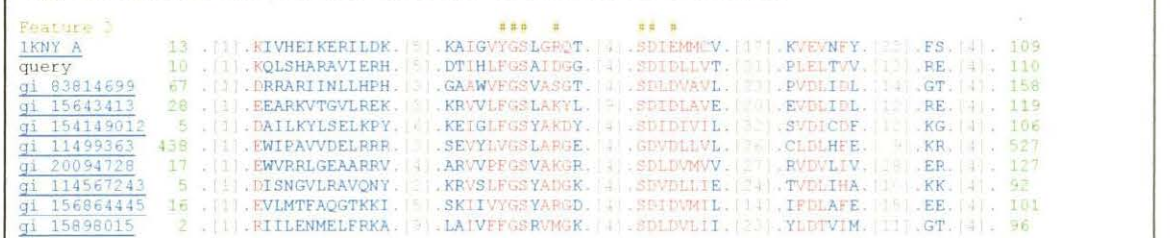
Description of CDD search: Specific hit result cd05403, Nucleotidyltransferase (NT) domain of *Staphylococcus aureus* kanamycin nucleotidyltransferase, and similar proteins. *S. aureus* KNTase is a plasmid encoded enzyme which confers resistance to a wide range of aminoglycoside antibiotics which have a 4'- or 4''-hydroxyl group in the equatorial position, such as kanamycin A. This enzyme transfers a nucleoside monophosphate group from a nucleotide (ATP, GTP, or UTP) to the 4'-hydroxyl group of kanamycin A. This enzyme is a homodimer, having two NT active sites. The nucleotide and antibiotic binding sites of each active site include residues from each monomer. Included in this subgroup is *Escherichia coli* AadA5 which confers resistance to the antibiotic spectinomycin and is a putative aminoglycoside-3'-adenylyltransferase. It is part of the aadA5 cassette of a class 1 integron. This subgroup also includes *Haemophilus influenzae* HI0073 which forms a 2:2 heterotetramer with an unrelated protein HI0074. Structurally HI0074 is related to the substrate-binding domain of S. aureus KNTase. The genes encoding HI0073 and HI0074 form an operon. Little is known about the substrate specificity or function of two-component NTs. The characterized members of this subgroup may not be representative of the function of this subgroup. This subgroup belongs to the Pol beta-like NT superfamily. In the majority of enzymes in this superfamily, two carboxylates, Dx [D/E], together with a third more distal carboxylate, co-ordinate two divalent metal cations involved in a two-metal ion mechanism of nucleotide addition. These carboxylate residues are conserved in this subgroup.



Feature1: Active site on conserved domain. 9 of 9 of the residues that compose this conserved feature have been mapped to the query sequence. Hash marks **(#)** above the aligned sequences show the location of the conserved feature residues.



Feature 2: metal binding triad [ion binding site] on conserved domain. 3 of 3 of the residues that compose this conserved feature have been mapped to the query sequence. Hash marks **(#)** above the aligned sequences show the location of the conserved feature residues.



Feature 3: NTP binding site [chemical binding site] on conserved domain. 7 of 7 of the residues that compose this conserved feature have been mapped to the query sequence. Hash marks **(#)** above the aligned sequences show the location of the conserved feature residues.

Fig. 2.10K: Features showing conserved domain (NT_KNTase_like) in query sequence

Computational determination using ProtParam tool (http://web.expasy.org/protparam/) of various physical and chemical parameters of AadA1 protein (responsible for streptomycin /spectinomycin resistance) of isolate, MB47

Computed parameters of AadA1 are following:

Number of amino acids: 263; Molecular weight: 29355.3; Theoretical pI: 4.60

Total number of negatively charged residues (Asp + Glu): 45
Total number of positively charged residues (Arg + Lys): 23

Ext. coefficient 41940 M^{-1} cm^{-1}, at 280 nm measured in water.
Abs 0.1% (=1 g/l) 1.429

Estimated half-life:
The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:
The instability index (II) is computed to be 41.28
This classifies the protein as unstable.
Aliphatic index: 110.49; Grand average of hydropathicity (GRAVY): -0.075

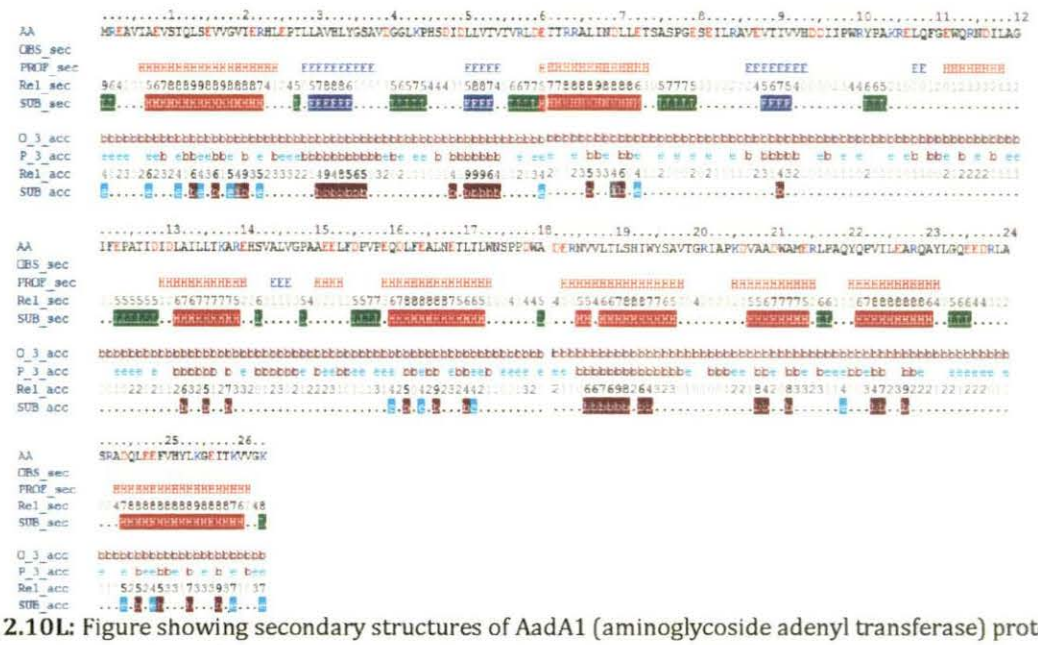


Fig 2.10L: Figure showing secondary structures of AadA1 (aminoglycoside adenyl transferase) protein coding for streptomycin/spectinomycin resistance derived from bacterium MB47.

Abbreviations: same as stated for Fig. 2.9F.

Homology modeling (automated mode, website *http://swissmodel.expasy.org/*) of AadA1 protein derived from class 1 integron borne gene cassette from a facultatively oligotrophic bacterium, MB47.

Fig 2.10M: graphical presentation of pairwise alignment between target nucleotide binding domain of the protein nucleotidyl transferase of AadA1 (green) derived from MB47 and *Haemophilus influenzae (blue)*. The blue bar with red outer line showed the modeled residues range and identity of amino acids present in-between two protein sequences.

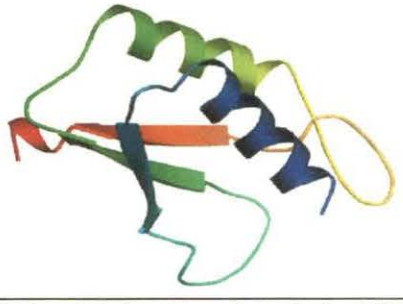
Model information: Modeled residue range: 10 to 96 Template: [1no5B] (1.80 Å) **Quaternary structure information:** Template (1no5): MONOMER Model built: SINGLE CHAIN **Ligand information:** Ligands in the template: SO4:1, ZN: 4. Ligands in the model: none.	

Fig. 2.10Mx: 3D model (ribbon model) of the template: 1no5B (crystal structure at resolution1.80 Å of HI0073, the nucleotide binding domain of the Hi0073/Hi0074 two protein nucleotidyl transferase from *Haemophilus influenzae*)

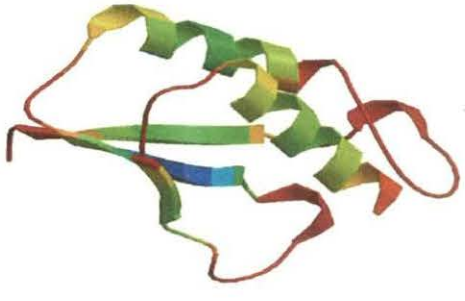
Model information: Template: [1no5B] (1.80 Å) Modeled residue range: 10 to 96 Sequence Identity [%]:21.84 Evalue: 2.00e-9 QMEAN score: 0.51 Z-Score: -2.21 Predicted LGscore : 2.112 Predicted MaxSub : 0.312	

Fig. 2.10My: 3D model (ribbon model) of the AadA1 protein of bacterium MB47

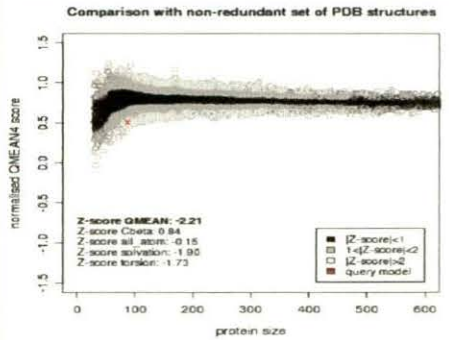
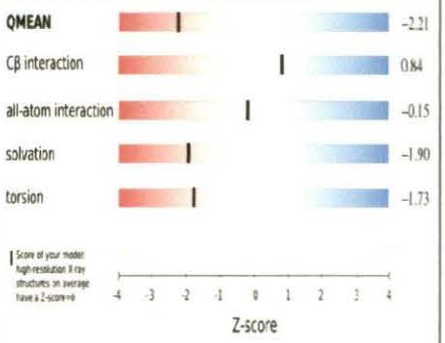
QMEAN4 score	Estimated absolute model quality	Score components
0.51 (Range for reliable model:0-1)	 Z-Score: -2.21	

Fig. 2.10Mz: Global model quality estimation of AadA1 protein of bacterium MB47

```
GGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACG
                                1R                          *(start)
CAGCAGGGCAGTCGCCCTAAAACAAAGTTAGCCATTAAGGGAGTTAAAatgaaaatatcattgatttctg
                                          dfrA17-- >      M  K  I  S  L  I  S  A
cagtgtcagaaaatggcgtaatcggtagtggtcctgatatcccgtggtcagtaaaaggtgagcaactact
  V  S  E  N  G  V  I  G  S  G  P  D  I  P  W  S  V  K  G  E  Q  L  L
ctttaaagcgctcacatataatcaatggctccttgtcggaagaaaaacatttgactctatgggtgttctt
 F  K  A  L  T  Y  N  Q  W  L  L  V  G  R  K  T  F  D  S  M  G  V  L
ccaaatcgcaaatatgcagtagtgtcaaagaacggaatttcaagctcaaatgaaaacgtcctagtttttc
P  N  R  K  Y  A  V  V  S  K  N  G  I  S  S  S  N  E  N  V  L  V  F  P
cttcaatagaaaatgctttgaaagagctatcaaaagttacagatcatgtatatgtctctggcggggtca
  S  I  E  N  A  L  K  E  L  S  K  V  T  D  H  V  Y  V  S  G  G  G  Q
aatctataatagccttattgaaaaagcagatataattcatttgtctactgttcacgttgaagtcgaaggt
 I  Y  N  S  L  I  E  K  A  D  I  I  H  L  S  T  V  H  V  E  V  E  G
gatatcaaattccctataatgcctgagaatttcaatttggtttttgaacagtttttttatgtctaatataa
D  I  K  F  P  I  M  P  E  N  F  N  L  V  F  E  Q  F  F  M  S  N  I  N
                                1L
attatacataccagatttggaaaaaaggctaaCAATGCGTTGCAGCACCAGTCGCTTCGCTCCTTGGACA
  Y  T  Y  Q  I  W  K  K  G  #(stop
GCTTTTAAGTCGCGTCTTTGTGGTTTTGCTGCGCAAAAGTATTCCACAAAGCCGCAACTTAAAAGCTGCC
               1R       *(start)
GCTGAACTTAACGTTAGGCATCatggtgaatttttccctgcacaagttttcaagcagctgtcccacgct
            aadA5-->   M  G  E  F  F  P  A  Q  V  F  K  Q  L  S  H  A
cgcgcggtgatcgagcgccatctggctgcgacactggacacaatccacctgttcggatctgcgatcgatg
R  A  V  I  E  R  H  L  A  A  T  L  D  T  I  H  L  F  G  S  A  I  D  G
gagggctgaagccggacagcgacatagacttgctcgtgaccgtcagcgccgcacctaacgattcgctccg
  G  L  K  P  D  S  D  I  D  L  L  V  T  V  S  A  A  P  N  D  S  L  R
gcaggcgctaatgctcgatttgctgaaagtctcatcaccgccaggcgatggcggaacatggcgaccgctg
 Q  A  L  M  L  D  L  L  K  V  S  S  P  P  G  D  G  G  T  W  R  P  L
gagctaactgttgtcgctcgaagcgaagtagtgccttggcgctatccggcgcggcgtgagcttcagttcg
E  L  T  V  V  A  R  S  E  V  V  P  W  R  Y  P  A  R  R  E  L  Q  F  G
gtgagtggctccgccacgacatcctttccggaacgttcgagcctgccgttctggatcacgatcttgcgat
  E  W  L  R  H  D  I  L  S  G  T  F  E  P  A  V  L  D  H  D  L  A  I
tttgctgaccaaggcgaggcaacacagccttgcgcttctaggcccatccgcagccacgtttttcgagccg
 L  L  T  K  A  R  Q  H  S  L  A  L  L  G  P  S  A  A  T  F  F  E  P
gtgccgaaggagcatttctccaaggcgcttttcgacactattgcccagtggaatgcagagtcggattgga
V  P  K  E  H  F  S  K  A  L  F  D  T  I  A  Q  W  N  A  E  S  D  W  K
agggtgacgagcggaacgtcgttcttgctcttgctcgcatttggtacagcgcttcaactggtctcattgc
  G  D  E  R  N  V  V  L  A  L  A  R  I  W  Y  S  A  S  T  G  L  I  A
tcctaaggacgttgctgccgcatgggtatcggagcgtttgcctgccgagcatcggcccctcatctgcaag
 P  K  D  V  A  A  A  W  V  S  E  R  L  P  A  E  H  R  P  L  I  C  K
gcacgcgcggcgtacctgggtagcgaggacgacgacctagcaatgcgcgtcgaagagacggccgcgttcg
A  R  A  A  Y  L  G  S  E  D  D  D  L  A  M  R  V  E  E  T  A  A  F  V
ttcgatatgccaaagcaacgattgagagaatcttgcgttgaGCGGCATGTGCGAAAAGTGCATCGACCCG
  R  Y  A  K  A  T  I  E  R  I  L  R  #  (stop)
                     1R
CGCCGAGGGCATCTGATGCCTAACTCGGCGTTCAAGCGGACGGGCTGCGCCCGCCGCTCAACTATGCG//T
(59 base element)
TAGATGCACTAAGCACATAATTGCTCACAGCCAAACTATCAGGTCAAGTCTGCTT
```

Fig. 2.11: Nucleotide sequence (Ac. No. FM179325) of variable region obtained from MB53 showing ORFs, putative translated product and other features of class 1 integron borne gene cassette (symbols: *, start codon; #, stop codon; 1R, core site; 1L, inverse core site; //, beginning of 3' Conserved segment).

```
GGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGCTATGGAGCAGCAACGATGTTA
                                  1R                    * (start)
CGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGCCATTAAGGGAGTTAAAttgaaaatatcattgatt
                                                  dfrA7-> L  K  I  S  L  I
tctgcagtgtcagaaaatggcgtaatcggtagtggtcctgatatcccgtggtcagtaaaaggtgagca
S  A  V  S  E  N  G  V  I  G  S  G  P  D  I  P  W  S  V  K  G  E  Q
actactctttaaagcgctcacatataatcaatggctccttgtcggaagaaaaacatttgactctatgg
 L  L  F  K  A  L  T  Y  N  Q  W  L  L  V  G  R  K  T  F  D  S  M  G
gtgttcttccaaatcgcaaatatgcagtagtgtcaaagaacggaatttcaagctcaaatgaaaacgtc
  V  L  P  N  R  K  Y  A  V  V  S  K  N  G  I  S  S  S  N  E  N  V
ctagtttttccttcaatagaaaatgctttgaaagagctatcaaaagttacagatcatgtatatgtctc
L  V  F  P  S  I  E  N  A  L  K  E  L  S  K  V  T  D  H  V  Y  V  S
tggcgggggtcaaatctataatagccttattgaaaaagcagatataattcatttgtctactgttcacg
 G  G  G  Q  I  Y  N  S  L  I  E  K  A  D  I  I  H  L  S  T  V  H  V
ttgaagtcgaaggtgatatcaaattccctataatgcctgagaatttcaatttggtttttgaacagttt
  E  V  E  G  D  I  K  F  P  I  M  P  E  N  F  N  L  V  F  E  Q  F
                                                  1L
tttatgtctaatataaattatacataccagatttggaaaaaaggctaaCAATGCGTTGCAGCACCAGT
F  M  S  N  I  N  Y  T  Y  Q  I  W  K  K  G  #(stop)
CGCTTCGCTCCTTGGACAGCTTTTAAGTCGCGTCTTTGTGGTTTTGCTGCGCAAAAGTATTCCACAAA
                                  1R          *(start)
GCCGCAACTTAAAAGCTGCCGCTGAACTTAACGTTAGGCATCatggtgaatttttccctgcacaagt
                                                  aadA5-> M  G  E  F  F  P  A  Q  V
tttcaagcagctgtcccacgctcgcgcggtgatcgagcgccatctggctgcgacactggacacaatcc
 F  K  Q  L  S  H  A  R  A  V  I  E  R  H  L  A  A  T  L  D  T  I  H
acctgttcggatctgcgatcgatggagggctgaagccggacagcgacatagacttgctcgtgaccgtc
  L  F  G  S  A  I  D  G  G  L  K  P  D  S  D  I  D  L  L  V  T  V
agcgccgcacctaacgattcgctccggcaggcgctaatgctcgatttgctgaaagtctcatcaccgcc
S  A  A  P  N  D  S  L  R  Q  A  L  M  L  D  L  L  K  V  S  S  P  P
aggcgatggcggaacatggcgaccgctggagctaactgttgtcgctcgaagcgaagtagtgccttggc
 G  D  G  G  T  W  R  P  L  E  L  T  V  V  A  R  S  E  V  V  P  W  R
gctatccggcgcggcgtgagcttcagttcggtgagtggctcgcacgacatctttcggaacgttcgagc
  Y  P  A  R  R  E  L  Q  F  G  E  W  L  A  R  H  L  S  E  R  S  S
ctgccgtctggatcacgatctgcgattttgctgaccaaggcgaggcaacacagccttgcgcttctagg
L  P  S  G  S  R  S  A  I  L  L  T  K  A  R  Q  H  S  L  A  L  L  G
cccatccgcagccacgtttttcgagccggtgccgaaggagcatttctccaaggcgcttttcgacacta
 P  S  A  A  T  F  F  E  P  V  P  K  E  H  F  S  K  A  L  F  D  T  I
ttgcccagtggaatgcagagtcggattggaagggtgacgagcggaacgtcgttcttgctcttgctcgc
  A  Q  W  N  A  E  S  D  W  K  G  D  E  R  N  V  V  L  A  L  A  R
atttggtacagcgcttcaactggtctcattgctcctaaggacgttgctgccgcatgggtatcggagcg
I  W  Y  S  A  S  T  G  L  I  A  P  K  D  V  A  A  A  W  V  S  E  R
tttgcctgccaagcatcggcccctcatctgcaaggcacgcgcggcgtacctgggtagcgaggacgacg
 L  P  A  K  H  R  P  L  I  C  K  A  R  A  A  Y  L  G  S  E  D  D  D
acctgcaatgcgcgtcgaagagacggccgcgttcgttcgatatgccaaagcaacgattgagagaatct
  L  Q  C  A  S  K  R  R  P  R  S  F  D  M  P  K  Q  R  L  R  E  S
                                                  1L
tgcgttgagcggcatgtgcgaaaagtgcatcgacccgcgccgagggcatctgatgcctaaCTCGGCGT
C  V  E  R  H  V  R  K  V  H  R  P  A  P  R  A  S  D  A  #(stop)
TCAAGCGGACGGGCTGCGCCCGCCGCTCAACTATGCG//TTAGATGCACTAAGCACATAATTGCTCAC
AGCCAAACTATCAGGTCAAGTCTGCTT
```

Fig. 2.12A: Nucleotide sequence (Ac. No. HE650979) of variable region obtained from MB29 showing ORFs, putative translated product and other features of class 1 integron borne gene cassette.
Symbols: *, start codon; #, stop codon; *1R*, core site; *1L*, inverse core site; //, beginning of 3' conserved segment.

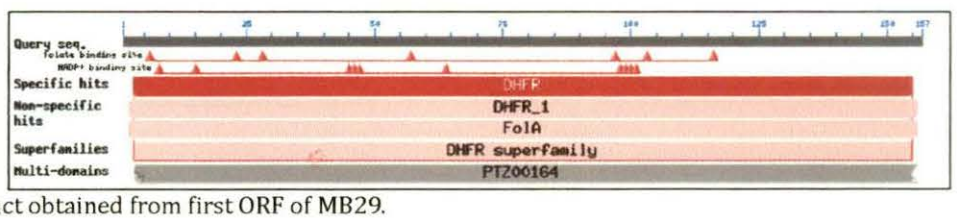


Fig. 2.12B: Result of conserved domain search for the putative translated product obtained from first ORF of MB29.

Description of CDD search: same as described in Fig. 2.9C.

BlastP: Homology of predicted orf1 of MB29 with the existing protein sequences database

gb|EGB89678.1| dihydrofolate reductase [*Escherichia coli* MS 117-3]
Length=210
Score = 320 bits (820), Expect = 3e-109, Method: Compositional matrix adjust.
Identities = 157/157 (100%), Positives = 157/157 (100%), Gaps = 0/157 (0%)

```
Query  1    LKISLISAVSENGVIGSGPDIPWSVKGEQLLFKALTYNQWLLVGRKTFDSMGVLPNRKYA  60
            LKISLISAVSENGVIGSGPDIPWSVKGEQLLFKALTYNQWLLVGRKTFDSMGVLPNRKYA
Sbjct  54   LKISLISAVSENGVIGSGPDIPWSVKGEQLLFKALTYNQWLLVGRKTFDSMGVLPNRKYA  113

Query  61   VVSKNGISSSNENVLVFPSIENALKELSKVTDHVYVSGGGQIYNSLIEKADIIHLSTVHV  120
            VVSKNGISSSNENVLVFPSIENALKELSKVTDHVYVSGGGQIYNSLIEKADIIHLSTVHV
Sbjct  114  VVSKNGISSSNENVLVFPSIENALKELSKVTDHVYVSGGGQIYNSLIEKADIIHLSTVHV  173

Query  121  EVEGDIKFPIMPENFNLVFEQFFMSNINYTYQIWKKG  157
            EVEGDIKFPIMPENFNLVFEQFFMSNINYTYQIWKKG
Sbjct  174  EVEGDIKFPIMPENFNLVFEQFFMSNINYTYQIWKKG  210
```

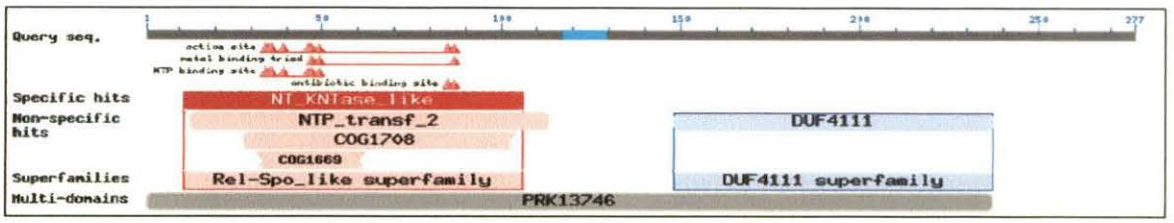


Fig. 2.12C: Result of conserved domain search for the putative translated product obtained from second ORF derived from MB29 bacterium.

Description of CDD search: same as described earlier in fig. 2.9l.

BlastP result of ORF2 of MB29

gb|AEQ26263.1| aminoglycoside adenyltransferase [Riemerella anatipestifer]
Length=262

Score = 439 bits (1129), Expect = 3e-153, Method: Compositional matrix adjust.
Identities = 221/239 (92%), Positives = 225/239 (94%), Gaps = 2/239 (1%)

```
Query  1    MGEFFPAQVFKQLSHARAVIERHLAATLDTIHLFGSAIDGGLKPDSDIDLLVTVSAAPND  60
            MGEFFPAQVFKQLSHARAVIERHLAATLDTIHLFGSAIDGGLKPDSDIDLLVTVSAAPND
Sbjct  1    MGEFFPAQVFKQLSHARAVIERHLAATLDTIHLFGSAIDGGLKPDSDIDLLVTVSAAPND  60

Query  61   SLRQALMLDLLKVSSPPGDGGTWRPLELTVVARSEVVPWRYPARRELQFGEWLARHLSER  120
            SLRQALMLDLLKVSSPPGDGGTWRPLELTVVARSEVVPWRYPARRELQFGEWL   +
Sbjct  61   SLRQALMLDLLKVSSPPGDGGTWRPLELTVVARSEVVPWRYPARRELQFGEWLRHDILSG  120

Query  121  SSLPS--GSRSAILLTKARQHSLALLGPSAATFFEPVPKEHFSKALFDTIAQWNAESDWK  178
            +  P+       AILLTKARQHSLALLGPSAATFFEPVPKEHFSKALFDTIAQWNAESDWK
Sbjct  121  TFEPAVLDHDLAILLTKARQHSLALLGPSAATFFEPVPKEHFSKALFDTIAQWNAESDWK  180

Query  179  GDERNVVLALARIWYSASTGLIAPKDVAAAWVSERLPAKHRPLICKARAAYLGSEDDDL  237
            GDERNVVLALARIWYSASTGLIAPKDVAAAWVSERLPA+HRPLICKARAAYLGSEDDDL
Sbjct  181  GDERNVVLALARIWYSASTGLIAPKDVAAAWVSERLPAEHRPLICKARAAYLGSEDDDL  239
```

```
gi 123545237    3 ASELFDPVPKESFVKALSDTAAQW.[2].ESDWQGDERNVVLALARIWFSVSTGEIAPKDAAASWLLERLPHEHRAVLMV 79
query         148 AATFFEPVPKEHFSKALFDTIAQW.[2].ESDWKGDERNVVLALARIWYSASTGLIAPKDVAAAWVSERLPAKHRPLICK 224
gi 75401522    11 AATFFEPVPKEHFSKALFDTIAQW.[2].ESDWKGDERNVVLALARIWYSASTGLIAPKDVAAAWVSERLPAEHRPLICK 87
gi 123565762  150 AAALFEQVPARDFVAALLATVAQW.[2].EPDWRGDACNVVLALARIWYSAATGRIAPKDVAATWVLERLPETYRPVVAA 226
gi 123074532  150 ATDLFDPVPRTDLARAFYDTALQW.[2].PADWRGDERTVVLALARIWFSLSTGGIAPKDVAAQWVIERLPDEHQSIMRS 226
gi 123450561  150 AQKVFMPVPEHDFLQVLSDTLKLW.[2].HEDWENEERNIVLTLARIWYSTETGGIVPKDVAAEWVLERLPAEHKPILVE 226
gi 75449966   149 AEEFFDPVPEQDLFEALRETLKLW.[2].QPDWAGDERNVVLTLSRIWYSAITGKIAPKDVAADWAIKRLPAQYQPVLLE 225
gi 239820329  150 ADELFDPVPRADFSRSLLDTVAQW.[2].APDWQGDERNVVLALARIWFSACTGQIAPKDTAAAWALERLPSEHRPVLAS 226
gi 146307205  178 AAELFEPVPEADMIHALHDTAAQW.[2].PDDWLGDECNVVLALARIWLTLSSGEIAAKDVAADWLLERLPPEHRPLLET 254
gi 146282523  145 AAELFDPVPEEDMRQALRDTLAQW.[2].PADWQGDECNVVLAVARIWLTLTTDAIAPKDVAAIWLLERLPDEHRAVLEK 221

gi 123545237   80 ALAAYLGTARDDL.[18]. 110
query         225 APAAYLGSEDDDL.[18]. 255
gi 75401522    88 APAAYLGSEDDDL.[18]. 118
gi 123565762  227 APAAYLGGEEEAA.[18]. 257
gi 123074532  227 APASYLGETQDDL.[16]. 257
gi 123450561  227 APQAYLGLCKDSL.[18]. 257
gi 75449966   226 APQAYLGQKEDHL.[18]. 256
gi 239820329  227 ALAAYLGRAPDNL.[18]. 257
gi 146307205  255 APDVYLGQARDDL.[18]. 285
gi 146282523  222 ASQAYRGDAIEKP.[18]. 252
```

Fig. 2.12D: Sequence alignment of part of protein sequence of orf2 (coding for Aad) derived from MB29 bacterium showing domain similarity with DUF4111 family.

Domain of unknown function (DUF4111): the protein-protein homology of downstream sequence of *orf 2* which was coding for AadA5 protein showed similarity with an unknown function domain. Although the exact function of this domain is not known however it frequently appears downstream of sequences of the family, Nucleotidyl transferase, pfam01909. It is also found in species associated with methicillin-resistant bacteria.

```
GGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTAC
                                                                      *
GCAGCAGGGCAGTCGCCCTAAAACAAAGTTAACCTCTGAGGAAGAATTGTGAAACTATCACTAatggta
                             1R                                DfrA1->  M  V
gctatatcgaagaatggagttatcgggaatggccctgatattccatggagtgccaaaggtgaacagctc
A  I  S  K  N  G  V  I  G  N  G  P  D  I  P  W  S  A  K  G  E  Q  L
ctgtttaaagctattacctataaccaatggctgttggttggacgcaagacttttgaatcaatgggagca
L  F  K  A  I  T  Y  N  Q  W  L  L  V  G  R  K  T  F  E  S  M  G  A
ttacccaaccgaaaatatgcggtcgtaacacgttcaagttttacatctgacaatgagaacgtattgatc
L  P  N  R  K  Y  A  V  V  T  R  S  S  F  T  S  D  N  E  N  V  L  I
tttccatcaattaaagatgctttaaccaacctaaagaaaataacggatcatgtcattgtttcaggtggt
F  P  S  I  K  D  A  L  T  N  L  K  K  I  T  D  H  V  I  V  S  G  G
ggggagatatacaaaagcctgatcgatcaagtagatacgctacatatatctacaatagacatcgagccg
G  E  I  Y  K  S  L  I  D  Q  V  D  T  L  H  I  S  T  I  D  I  E  P
gaaggtgatgtttactttcctgaaatccccagcaattttaggccagtttttacccaagacttcgcctct
E  G  D  V  Y  F  P  E  I  P  S  N  F  R  P  V  F  T  Q  D  F  A  S
aacataaattatagttaccaaatctggcaaaaggttaaCAAGTGGCAGCAACGGATTCGCAAACCTGT
N  I  N  Y  S  Y  Q  I  W  Q  K  G  #(end of dfrA1)
                                                  1R                *
CACGCCTTTTGTACAAAAGCCGCGCCAGGTTTGCGATCCGCTGTGCCAGGCGTTAAGGCTACatgaaaa
                                                  ORF38->  M  K  I
tcgtacattacgaagcgaatgcaccatggataggaagaatgaaatgcccaaacccaaagtgtgggaagg
  V  H  Y  E  A  N  A  P  W  I  G  R  M  K  C  P  N  P  K  C  G  K  E
aaactcctgcctggcaatcgagcggcatgagcgacagttgcccgcatttttctgtgatacttgctcga
  T  P  A  W  Q  S  S  G  M  S  D  S  C  P  H  F  F  C  D  T  C  S  N
atgtaatccatagagagcaggaccatgcattactgtacgaaaatgaaatcaatcaagagctcttggatc
  V  I  H  R  E  Q  D  H  A  L  L  Y  E  N  E  I  N  Q  E  L  L  D  R
gaatagcagcaactcttccagattgcccttgcggggataggtttgttcctggtgcaaacccaaagtgtc
  I  A  A  T  L  P  D  C  P  C  G  D  R  F  V  P  G  A  N  P  K  C  P
cgagttgcaagaccgagtacgtgcaccaatgggatgcagtgaaaaggttgaatgtacctttatgccaa
  S  C  K  T  E  Y  V  H  Q  W  D  A  V  K  R  L  N  V  P  F  M  P  I
tcttgtatggttcctgcttgattcgagataggctgtattcgtatgaagtatgcattggttctaaaccaa
  L  Y  G  S  C  L  I  R  D  R  L  Y  S  Y  E  V  C  I  G  S  K  P  K
                          1L
aatactggtggcgtttgttcacaaatgccttaacaagtttaggcaagggacgctcctgaCGTCGCGCCC
  Y  W  W  R  L  F  T  N  A  L  T  S  L  G  K  G  R  S  #
CTGCTAAAAGCG//TTAGATGCACTAAGCACATAATTGCTCACAGCCAAACTATCAGGTCAAGTCTGCT
T
```

Fig. 2.13A: Nucleotide sequence (Ac. No. AM997276) of variable region amplified through CS-PCR from MB38 isolate, showing putative translated product and other features of class 1 integron borne gene cassette. Symbols: *, start codon; #, stop codon; 1R, core site; 1L, inverse core site; //, beginning of 3' Conserved segment.

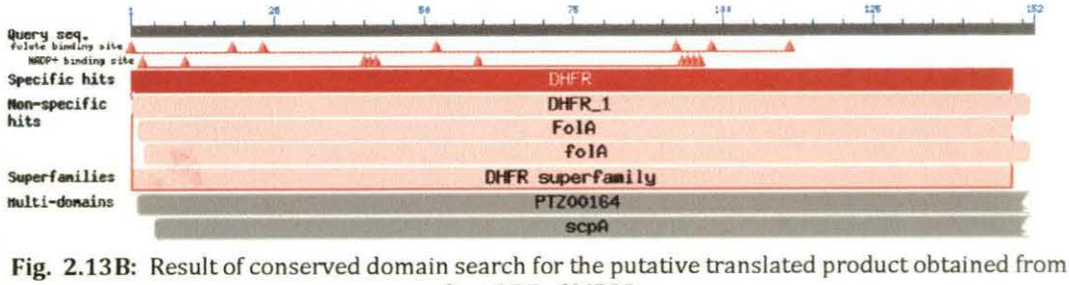


Fig. 2.13B: Result of conserved domain search for the putative translated product obtained from first ORF of MB38

BlastP result of second ORF38 of MB38,

gb|AAM33365.1|AF455254_2 unknown [*Vibrio cholerae*]
dbj|BAD02391.1| unknown [*Vibrio cholerae* non-O1/non-O139]
dbj|BAD08521.1| hypothetical protein [*Escherichia coli*]
36 more sequence titles
Length=159

```
Score =  330 bits (846),  Expect = 1e-115, Method: Compositional matrix adjust.
Identities = 157/159 (99%), Positives = 157/159 (99%), Gaps = 0/159 (0%)

Query  1    MKIVHYEANAPWIGRMKCPNPKCGKETPAWQSSGMSDSCPHFFCDTCSNVIHREQDHALL  60
            MKIVHYEANAPWIGRMKCPNPKCGKETPAWQSSGMSDSCPHFFCDTCSNVIHREQDHALL
Sbjct  1    MKIVHYEANAPWIGRMKCPNPKCGKETPAWQSSGMSDSCPHFFCDTCSNVIHREQDHALL  60

Query  61   YENEINQELLDRIAATLPDCPCGDRFVPGANPKCPSCKTEYVHQWDAVKRLNVPFMPILY  120
            YENEINQELLDRIAATLPDCPCG RFVPGANPKCPSCKTEYVHQWDAVKRLNVPFMPIL
Sbjct  61   YENEINQELLDRIAATLPDCPCGGRFVPGANPKCPSCKTEYVHQWDAVKRLNVPFMPILD  120

Query  121  GSCLIRDRLYSYEVCIGSKPKYWWRLFTNALTSLGKGRS  159
            GSCLIRDRLYSYEVCIGSKPKYWWRLFTNALTSLGKGRS
Sbjct  121  GSCLIRDRLYSYEVCIGSKPKYWWRLFTNALTSLGKGRS  159
```

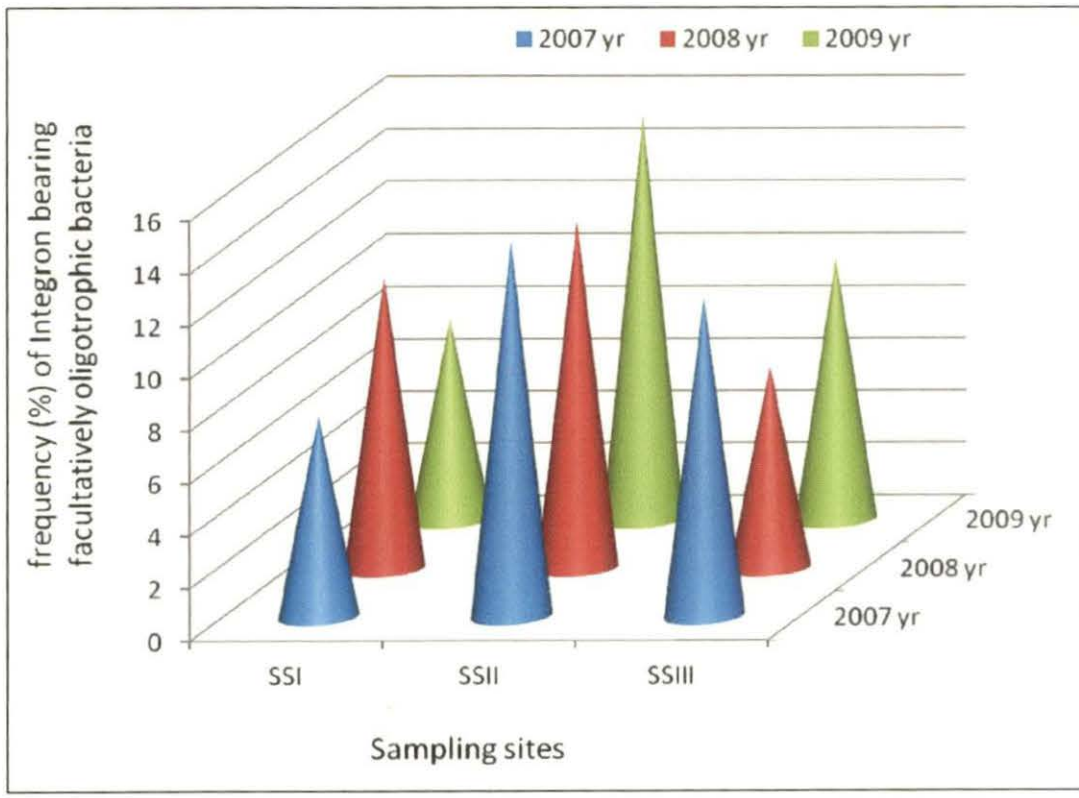


Fig. 2.14: Integron carriage in facultative oligotrophic bacteria at three different sampling sites in three consecutive years

Table 2.1: Susceptible/resistance profile of ninety class-1 integron positive oligotrophic bacterial isolates (symbols: please see table 2.2 footnote)

Group	RI	No. of strains[a]	Phenotypes[b]											
			Amp	Azi	Cef	Cft	Chl	Cip	Cot	Kan	Lev	Net	Str	Tet
0	0.00	18	-	-	-	-	-	-	-	-	-	-	-	-
		1	+	-	-	-	-	-	-	-	-	-	-	-
		1	-	+	-	-	-	-	-	-	-	-	-	-
1	0.08	1	-	-	-	-	-	-	+	-	-	-	-	-
		3	-	-	-	-	-	-	-	-	-	+	-	-
		1	-	-	-	-	-	-	-	-	-	-	-	+
		2	+	-	-	-	-	-	-	-	-	+	-	-
		1	+	-	-	-	-	-	-	-	-	-	+	-
2	0.16	1	-	-	+	+	-	-	-	-	-	-	-	-
		1	-	-	-	-	-	-	+	-	-	-	+	-
		1	-	-	-	-	-	-	+	-	-	-	-	+
		1	-	-	-	-	-	-	-	-	-	-	+	+
		1	+	-	-	-	-	-	+	-	-	-	+	-
		1	+	-	-	-	+	-	+	-	-	-	-	-
3	0.25	1	+	-	-	-	-	-	+	-	-	-	-	+
		2	-	-	-	-	-	-	-	-	-	+	+	+
		1	+	-	-	-	-	-	-	+	-	-	-	+
		2	+	-	-	-	-	-	+	-	-	-	-	+
		1	-	-	-	-	-	-	+	-	+	+	-	+
		1	+	-	-	-	-	-	+	-	-	+	-	+
4	0.33	1	+	-	-	+	+	-	+	-	-	-	-	-
		1	+	-	-	+	+	-	-	-	-	-	+	-
		1	-	+	-	-	-	-	-	-	+	+	-	+
		2	+	-	+	-	+	-	+	-	-	+	-	-
		2	+	-	-	-	+	-	+	-	-	-	+	+
5	0.41	1	+	-	-	+	-	-	+	+	-	+	-	-
		1	+	-	+	-	-	+		+	-	+	-	-
		2	+	-	-	-	-	-	+	+	+	-	-	+
		1	-	-	-	+	+	-	+	-	-	-	+	+
		1	+	-	+	+	-	+	+	-	-	-	-	+
		1	+	+	-		-	+	+	-	+	-	-	+
		1	+	-	-	+	+	-	-	-	-	+	+	+
		1	+	-	-	-	-	+	+	+	-	+	-	+
6	0.50	1	+	+	-	+		+	+	-	-	+	-	-
		1	+	-	-	+	+	-	+	-	-	-	+	+
		1	+	-	+	+	+	-	+	-	+	-	-	-
		2	+	-	+	+	-	-	+	-	-	-	+	+
		1	+	-	-	+	-	-	+	-	-	+	+	+
		1	-	-	-	-	+	-	+	-	+	+	+	+
		1	-	+	+	-	+	+	-	+	+	-	-	-
		1	+	-	+	+	-	+	+	-	+	-	-	+
		1	+	-	+	+	-	+	+	+	-	+	-	+
		1	+	-	+	-	-	+	+	+	-	+	-	+
7	0.58	1	+	+	-	-	-	+	+	-	+	+	-	+
		1	+	-	+	+	-	+	+	-	-	+	-	+
		1	+	-	+	+	+	+	+	-	-	+	-	-
		1	+	-	+	+	+	+	+	+	-	-	+	-
		1	+	-	+	+	+	+	-	-	-	+	-	+
		1	+	-	+	+	-	-	+	-	+	+	-	+
		2	+	-	+	+	-	+	+	-	+	+	-	+
		1	+	-	+	+	-	+	+	+	-	+	-	+
		1	+	-	+	+	-	+	+	-	-	+	+	+
8	0.66	1	+	-	-	+	+	+	+	-	+	-	+	+
		1	+	+	+	+	-	-	+	-	-	+	+	+
		1	+	+	+	+	+	+	+	-	+	-	-	-
		1	+	+	+	+	-	+	+	-	+	-	-	+
		1	+	+	+	+	-	+	+	+	-	+	-	-
		1	+	+	+	-	-	+	+	-	+	-	+	+
9	0.75	1	+	-	+	+	+	+	+	-	+	+	+	-
10	0.83	1	+	-	+	+	+	+	+	+	+	+	-	+
		2	+	+	+	+	+	+	+	-	+	-	+	+
11	0.91	1	+	+	+	+	+	+	+	-	+	+	+	+

Table 2.2: Resistance phenotype of class 1integron positive oligotrophic bacterial isolates

Group	Resistance index (RI)	Isolate	Phenotype[a]
		OB 05	Amp
		MB 09	Tet
1		MR 01	Azi
	0.08	MB 20	Net
		MB 24	Net
		MB 50	Net
		MB 57	Cot
		MR 02	Cef, Cft
2	0.16	MB 23	Amp, Str
		MB 36	Str, Tet
		MB 49	Amp, Net
		MB 52	Cot, Str
		MB 51	Amp, Net
		MB 64	Cot, Tet
		MB 03	Amp, Cot, Str
3	0.25	MB 21	Amp. Chl, Cot
		MB 26	Amp, Cot, Tet
		MB 32	Net, Str, Tet
		MB 33	Net, Str, Tet
		MB 42	Amp, Kan, Tet
		MB 67	Amp, Cot, Tet
		MB 77	Amp, Cot, Tet
		MR 03	Cot, Liv, Net, Tet
4	0.33	MB 47	Amp, Cot, Net, Tet
		MB 62	Amp, Cft, Chl, Cot
		MB 81	Amp, Cft, Chl, Str
		MR 04	Azi, Lev, Net, Tet
		MB 18	Amp, Cef, Chl, Cot, Net
		MB 31	Amp, Chl, Cot, Str, Tet
5		MB 40	Amp, Cft, Cot, Kan, Net
	0.41	MB 59	Amp, Cef, Chl, Cot, Net,
		MB 66	Amp. Chl, Cot, Str, Tet
		MB 72	Amp, Cef, Cip, Kan, Net
		MB 74	Amp, Cot, Kan, Liv, Tet
		MB 75	Amp, Cot, Kan, Liv, Tet
		MB 80	Cft, Chl, Cot, Str, Tet
		MB 27	Amp, Cef, Cft, Cip, Cot, Tet
		MB 30	Amp, Azi, Cip, Cot, Liv, Tet
6		MB 41	Amp, Cft, Chl, Net, Str, Tet
	0.50	MB 60	Amp, Cip, Cot, Kan, Net, Tet
		MB 61	Amp, Azi, Cft, Cip, Cot, Net
		MB 69	Amp, Cft, Chl, Cot, Str, Tet
		MB 78	Amp, Cef, Cft, Chl, Cot, Liv
		SR 19	Amp, Cef, Cft, Cot, Str, Tet
		OD 05	Amp, Cft, Cot, Net, Str, Tet
		OD 08	Amp, Cef, Cft, Cot, Str, Tet
		OC 16	Chl, Cot, Liv, Net, Str, Tet
		MB 82	Azi, Cef, Chl, Cip, Kan, Liv
		MB 19	Amp, Cef, Cft, Cip, Cot, Liv, Tet
		MB 28	Amp, Cef, Cft, Cip, Cot, Kan, Net,
7	0.58	MB 29	Amp, Cef, Cip, Cot, Kan,Net, Tet
		MB34B	Amp, Azi, Cip, Cot, Liv, Net, Tet
		MB37A	Amp, Cef, Cft, Cip, Cot, Net, Tet
		MB 53	Amp, Cef, Cft, Chl, Cip, Cot, Net
		MB 70	Amp, Cef, Cft, Chl, Cot, Kan, Str
		MB 73	Amp, Cef, Cft, Chl, Cip, Net, Tet
		OC 78	Amp, Cef, Cft, Cot, Liv, Net, Tet
		MB 25	Amp, Cef, Cft, Cip, Cot, Liv, Net, Tet
		OC 24	Amp, Cef, Cft, Cip, Cot, Liv, Net, Tet
8	0.66	MB 35	Amp, Cef, Cft, Cip, Cot, Kan, Net, Tet,
		MB 45	Amp, Cef, Cft, Cip, Cot, Net, Str, Tet
		MB 79	Amp, Cft, Chl, Cip, Cot, Liv, Str, Tet
		OD 10	Amp, Azi, Cef, Cft, Cot, Net, Str, Tet
		OD 21	Amp, Azi, Cef, Cft, Chl, Cip, Cot, Lev
		OD 24	Amp, Azi, Cef, Cft, Cip, Cot, Lev, Tet
		MB 65	Amp, Azi, Cef, Cft, Cip Cot, Kan, Net
		MB 76	Amp, Azi, Cef, Cip, Cot, Liv, Str, Tet

Table 2.2: continue.....

Group	Resistance index (RI)	Isolate	Phenotype[a]
9	0.75	OC 74	Amp, Cef, Cft, Chl, Cip, Cot, Liv, Net, Str
		MB 38	Amp, Cef, Cft, Chl, Cip, Cot, Kan, Liv, Net, Tet
10	0.83	MB 68	Amp, Azi, Cef, Cft, Chl, Cip, Cot, Liv, Str, Tet
		OC 75	Amp, Azi, Cef, Cft, Chl, Cip, Cot, Liv, Str, Tet
11	0.91	NV 66	Amp, Azi, Cef, Cft, Chl, Cip, Cot, Lev, Net, Str, Tet

[a] number of strains sensitive/or resistant to a particular antibiotics or set of antibiotics.
[b]Amp: ampicillin; Azi: azithromycin; Cef: cefipime; Cft: cefotaxime; Chl: chloramphenicol; Cip: ciprofloxacin; Cot: co-trimoxazole; Kan: kanamycin; Lev: levofloxacin; Net: netilmicin; Str: streptomycin; Tet: oxytetracycline

Table 2.3A: Isolates bearing gene cassettes unrelated to antibiotic resistance.

RI[a]	Isolate	Approx length (bp)	Nature of the gene cassette	Cassette encoded Function	Accession Number
0.0	OB 12	1000	*orf1*	Hypothetical protein	AM997272
	MB 05	223	*mts*	Ribosomal methyl transferase	AM991331
	MB 08	608	*orf1*	Hypothetical protein	FN178516
	MB 12	1000	*livM1*	Branched chain ABC transporter	AM997273
	MB 16	1000	*appA*	Oligopeptide ABC transporter	AM991327
	MB 22	1010	*fgam*	Phosphoribosyl formyl glycinamide synthase	AM991334
	MB 44	522	*orf1*	Hypothetical protein	MB41[b]
	MB 48	1095	*orf1*	Hypothetical protein	FN178520
	MB 54	527	*orf1*	Hypothetical protein	FM955254
	MB 55	959	*orf55, fgam*	Hypothetical protein, Phosphoribosyl formyl glycinamide synthase	HE653232
	MB 56	900	*fhaC*	Putative hemolysin activator protein	FM955483
	MB 58	624	*tnp*	Transposase, IS4 family protein	FM955255
	MB 71	627	*tnp*	Transposase, IS4 family protein	MB58[b]
	MB 83	435	*rsu*	Ribosomal large subunit pseudouridine synthase B	OB05[b]
0.08	OB 05	489	*rsu*	Ribosomal large subunit pseudouridine synthase B	AM997271
	MB 09	682	*hsdR*	DNA degradation	AM991332
	MR 01	951	*fgam*	Phosphoribosyl formyl glycinamide synthase	FN561626
0.16	MB 51	794	*orf1*	Hypothetical protein	AM997281
	MR 02	513	GGDEF	Signal transduction	FN561627
0.33	MB 81	522	*orf1*	Hypothetical protein	MB 41[b]
	MR 04	1300	SNF2 family	ATP-dependent helicase	FN561629
0.41	MB 80	522	*orf1*	Hypothetical protein	MB41[b]
0.50	MB 41	522	*orf1*	Hypothetical protein	HE653229
0.58	MB 28	1620	*ydcR*	Bifunctional putative Transcriptional regulator	FM179326
	MB 19	1500	*yfkc*	Putative reverse transcriptase maturase	AM997282
	MB 70	223	*mts*	Ribosomal methyl transferase	MB05[b]

[a]Resistance index; [b] gene sequence identical to that isolate

Table 2.3B: Isolates bearing gene cassettes related to antibiotic resistance, and bearing no cassettes (empty class 1 integron).

RI[a]	Isolate	Approx length (bp)	Nature of the gene cassette	Cassette encoded Function	Accession Number
0.0	MB 39	1009	*aadA2*	STR	AM997277
	MB 43	792	*aacA4*	KAN	HE653231
	MB 46	1009	*aadA2*	STR	AM997280
	MB 63	153	*In*0	Empty class 1 integron	FM958478
0.08	MB 20	1009	*aadA1*	STR	MB 32[b]
	MB 24	1003	*aadA1*	STR	AM991326
	MB 50	1009	aadA1	STR	MB 32[b]
	MB 57A MB 57B	1242 (i) 973 (ii)	*dfrA1-orfB* (i) *aadA1* (ii)	TMP-Hypothetical (i) STR (ii)	HE653233 (i) HE653234 (ii)
0.16	MB 23	1009	*aadA1*	STR/SPEC	MB 32[b]
	MB 36	1009	*aadA1*	STR/SPEC	AM937245
	MB 49	792	*aacA4*	KAN	MB43[b]
	MB 52	1694	*dfrA28- aadA1*	TMP-STR	FN263373
	MB 64	769	*dfrA7*	TMP	MB31[b]
0.25	MB 03	1647	*dfrA28-aadA1*	TMP-STR	AM937241
	MB 21	1569	*dfrA1- aadA*	TMP-STR	AM937243
	MB 26	1571	*dfrA1-aadA1*	TMP-STR	AM991328
	MB 32	1009	*aadA1*	STR/SPEC	AM991330
	MB 33	1009	*aadA1*	STR/SPEC	AM991333
	MB 42	862	*aacA4*	KAN	HE653230
	MB 67	1913	*dfrA12-orf40A- aadA2*	TMP-Hypothetical-STR	MB40A[b]
	MB 77	730	*dfrA30*	TMP	MB72[b]
0.33	MR 03	737	*dfrA16*	TMP	FN561628
	MB 47	1569	*dfrA1- aadA1*	TMP-STR	FM179327
	MB 62	153	*In*0	Empty class 1 integron	FM998811
0.41	MB 18	1048	*aadA5*	STR/SPE	AM937242
	MB 31	769	*dfrA7*	TMP	HE650981
	MB 40A MB 40B	1913 (i) 704 (ii)	*dfrA12- orf40A-aadA2* (i) *yrf1* (ii)	TMP-Hypothetical-STR(i) Helicase (ii)	FM179328 (i) AM997278 (ii)
	MB 59	1543	*dfrA1-aadA1*	TMP-STR	HE653235
	MB 66	1009	*aadA1*	STR/SPE	HE650982
	MB 72	730	*dfrA30*	TMP	HE650983
	MB 74	1242	*dfrA1-orf38*	TMP-Hypothetical	MB38[b]
	MB 75	1242	*dfrA1-orf38*	TMP-Hypothetical	MB38[b]
0.5	MB 27	1614	*dfrA17-aadA5*	TMP-STR	AM937244
	MB 30	1664	*dfrA7-aadA5*	TMP-STR	HE650980
	MB 60	1556	*dfrA1-aadA1*	TMP-STR	HE653236
	MB 61	1913	*dfrA12-orf40A-aadA2*	TMP-Hypothetical-STR	MB40A[b]
	MB 69	1543	*dfrA1-aadA1*	TMP-STR	MB59[b]
	MB 78	1694	*dfrA28-aadA1*	TMP-STR	MB52[b]
	MB 82	2013	*Oxa1-aadA1*	Bla-STR	HE650986
	SR 19	729	*dfrA5*	TMP	FN396373
	OD 05	1009	*aadA*	STR	FN396375
	OD 08	1009	*aadA*	STR	FN396376
	OC 16	1664	*dfrA17-aadA5*	TMP-STR	FN396368
0.58	MB 29	1657	*dfrA7-aadA5*	TMP-STR	HE650979
	MB 34B	1663	*dfrA17-aadA4*	TMP-STR	AM997275
	MB 37A	1661	*dfrA17-aadA5*	TMP-STR	AM991329
	MB 53	1664	*dfrA17-aadA5*	TMP-STR	FM179325
	MB 73	1664	*dfrA7-aadA5*	TMP-STR	[b]MB30
	OC 78	1170	*aac-6'-Ib*	STR	FN396372

Table 2.3B: continue.....

RI[a]	Isolate	Approx length (bp)	Nature of the gene cassette	Cassette encoded Function	Accession Number
0.66	MB 25	1606	*dfrA17-aadA4*	TMP-STR	AM997274
	MB 35	3454	*blaIMP-9-aacA4-oxa-10-aadA2*	Kan-Bla-STR	FN178517
	MB 45	667	*dfrA30*	TMP	AM997279
	MB 65	1556	*dfrA1-aadA1*	TMP-STR	MB60[b]
	MB 76	769	*dfrA7*	TMP	MB31[b]
	MB 79A MB 79B	1167 (i) 1009 (ii)	*dfrIIe-arr2* (i) *aadA1* (ii)	TMP-RMP (i) STR (ii)	HE650984 (i) HE650985 (ii)
	OD 10	1350	*aadA*	STR	FN396374
	OC 24	1521	*dfrA17-aadA5*	TMP-STR	FN396369
	OD 21	1009	*aadA2*	STR	FN396378
	OD 24	1000	*dfrA12*	TMP	FN396377
0.75	OC 74	1000	*aac-6'-Ib*	STR	FN396370
0.83	MB 38	1241	*dfrA1-orf38*	TMP-Hypothetical	AM997276
	MB 68	1543	*dfrA1-aadA1*	TMP-STR	MB59[b]
	OC 75	1400	*dfrA17*	TMP	FN396371
0.91	NV 66	1009	*aadA2*	STR	FN396367

[a]Resistance index (RI)
[b] gene sequence identical to that isolate
TMP, trimethoprim; STR, streptomycin; KAN, kanamycin; SPEC, spectinomycin
Gene name, described in main text of the manuscript

Table 2.4: The nearest protein sequence to the ABC transporter protein derived from MB12 obtained through protein-protein homology searching (BlastP).

Protein id	Description and source of origin
CAQ53856	ABC type branched chain amino acid transport system [bacterium AK-**MB12**]
ZP_10155718	branched-chain amino acid ABC transporter, permease protein [*Hydrogenophaga* sp. PBC
	ABC transporter permease [*Aromatoleum aromaticum* EbN1]
YP_158959	branched-chain amino acid ABC transporter system, permease protein [*Azoarcus* sp. KH32C]
BAL26594	

Table 2.5: Predicted trans-membrane regions of LivM1 protein (154AA) derived from bacterium, MB12.

Annotation	TM PROTEIN	
Number of predicted TMHs	4	
inside	1	1
TMhelix	2	20
outside	21	49
TMhelix	50	74
inside	75	86
TMhelix	87	110
outside	111	118
TMhelix	119	138
inside	139	154

TMH, trans-membrane helices;
AA, amino acid

Table 2.6: *Protein sequences with protein IDs that produces significant alignment score with the putative protein sequence derived from first orf1 (dfrA1) of class 1integron sequence of facultative oligotrophic bacterium, MB47.*

Protein id	Description and source of origin
CAQ76923	dihydrofolate reductase [bacterium AK-**MB47**]
YP_002527542	dihydrofolate reductase type I DhfrA1 [*Escherichia coli*]
YP_133853	dihydrofolate reductase type 1 [uncultured bacterium]
BAD99259	dihydrofolate reductase [*Escherichia coli*]
ACJ22953	dihydrofolate reductase type 1 [*Salmonella* sp. D76]

Table 2.7: Data summary of DfrA1_MB47 protein analyses [predicted secondary structure composition and solvent accessibility composition (core/surface ratio)] obtained through Protparam computing.

Sec. structure type	H	E	L	Accessibility type	b	e
% in protein	19.74	32.89	47.37	% in protein	35.53	64.47

H: helix, E=extended, L=loop

Classes used in Predicted solvent accessibility composition (core/surface ratio)

- e: residues exposed with more than 16% of their surface
- b: all other residues

Table 2.8: Protein sequences with protein IDs that produces significant alignment score with the putative protein sequences derived from *orf2* (*aadA1*) of class 1integron sequence of facultative oligotrophic bacterium, MB47.

Protein id	Description and source of origin
CAQ76924	Aminoglycoside adenyl transferase [bacterium AK-**MB47**]
AFK93816	aac(3'')-Ia [Proteus mirabilis]
YP_005980005	Streptomycin adenyltransferase [*Pseudomonas aeruginosa* NCGM2.S1]
YP_001144148	Aminoglycoside resistance protein [*Aeromonas salmonicida* subsp. *salmonicida* A449]
ACI43574.1	AadA1 [*Escherichia coli*]

Table 2.9: Data summary of AadA1_MB47 protein analyses [predicted secondary structure composition and solvent accessibility composition (core/surface ratio)] obtained through Protparam computing.

Sec. structure type	H	E	L	accessibility type	b	e
% in protein	49.43	10.65	39.92	% in protein	42.97	57.03

H: helix, E=extended, L=loop

Classes used in Predicted solvent accessibility composition (core/surface ratio)

- e: residues exposed with more than 16% of their surface
- b: all other residues

Table 2.10: Fraction (in %) of facultative oligotrophic bacteria of river Mahananda carrying integron [summarized data of 90 samples].

Year	Integron frequencies at three sampling sites (%)		
	SSI	**SSII**	**SSIII**
2007 yr	7.7	14.4	12.2
2008 yr	11.2	13.3	7.7
2009 yr	7.7	15.5	10

REFERENCES

Agersø Y, Sandvang D. Class 1 integrons and tetracycline resistance genes in *Alcaligenes, Arthrobacter,* and *Pseudomonas* spp. isolated from pigsties and manured soil. *Appl Environ Microbiol* 2005; 71:7941-7.

Ahmed AM, Motoi Y, Sato M *et al.* Zoo animals as reservoirs of gram-negative bacteria harboring integrons and antimicrobial resistance genes. *Appl Environ Microbiol* 2007; 73: 6686-90.

Ahmed AM, Nakagawa T, Arakawa E *et al.* 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* 2004; 53: 947-51.

Akram M, Shakil S, Khan AU. Prevalence of integrons, *bla*CTX-M and *bla*TEM resistance markers among ESBL-producing uropathogenic *Escherichia coli* isolates: first report of genomic *bla*CTX-M from India. *J Chemother* 2011; 23:131-4.

Altschul SF, Madden TL, Schaffer AA *et al.* Gapped BLAST and PSI-BLAST: a new gwnwration of protein database search programs. *Nucleic Acid Res* 1997; 25: 3389-402.

Amita, Chowdhury S R, Thungapathra M *et al.* Class I integrons and SXT elements in E1 Tor strains isolated before and after 1992 *Vibrio cholerae* O139 outbreak, Calcutta, India. *Emerg Infect Dis* 2003; 9: 500-2.

Arakawa Y, Murakami M, Suzuki K *et al.* A novel integron like element carrying the metallo-beta-lactamase gene bla_{IMP}. *Antimicrob Agents Chemother* 1995; 39: 1612-5.

Ash RJ, Mauck B, Morgan M. Antibiotic resistance of gram-negative bacteria in rivers, United States. *Emerg Infect Dis* 2002; 8: 713-6.

Aubert D, Girlich D, Naas T *et al.* 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* 2004; 48: 3284-90.

Baquero F, Martinez JL, Canton R. Antibiotics and antibiotic resistance in water environments. *Curr Opin Biotechnol* 2008; 19: 260-5.

Bass L, Liebert CA, Lee MD *et al.* Incidence and characterization of integrons, genetic elements mediating multiple-drug resistance, in avian *Escherichia coli. Antimicrob Agents Chemother* 1999; 43: 2925-9.

Betteridge T, Partridge SR, Iredell JR *et al.* Genetic context and structural diversity of class 1 integrons from human commensal bacteria in a hospital intensive care unit. *Antimicrob Agents Chemother* 2011; 55: 3939–43.

Bhattacharjee A, Sen MR, Prakash P *et al.* Observation on integron carriage among clinical isolates of *Klebsiella pneumoniae* producing extended-spectrum β-lactamases. *Indian J Med Microbiol* 2010; 28: 207-10.

Binh CTT, Heuer H, Kaupenjohann M *et al.* Diverse *aadA* gene cassettes on class 1 integrons introduced into soil via spread manure. *Res Microbiol* 2009; 160: 427–33.

Bissonnette L, Champetier S, Buisson J-P *et al.* Characterization of the nonenzymatic chloramphenicol resistance (*cmLA*) gene of the In4 integron of *Tn*1696: similarity of the product to transmembrane transport proteins. *J Bacteriol* 1991; 173: 4493-502.

Bissonnette L, Roy PH. Characterization of In0 of *Pseudomonas aeruginosa* plasmid pVS1, an ancestor of integrons of multiresistant plasmids and transposons of gram-negative bacteria. *J Bacteriol* 1992; 174: 1248–57.

Boucher Y, Labbate M, Koenig JE *et al.* Integrons: mobilizable platforms that promote genetic diversity in bacteria. *Trends Microbiol* 2007; 15: 301–9.

Brown HJ, Stokes HW, Hall RM. The integrons In0, In2 and In5 are defective transposon derivatives. *J Bacteriol* 1996; 178: 4429-37.

Bunny KL, Hall RM, Stokes HW. New mobile gene cassettes containing an aminoglycoside resistance gene, *aacA7*, and a chloramphenicol resistance gene, *catB3*, in an integron in pBWH301. *Antimicrob Agents Chemother* 1995; 39: 686-93.

Byrne-Bailey KG, Gaze WH, Zhang L *et al.* Integron prevalence and diversity in manured soil. *Appl Environ Microbiol* 2010; 77: 684–7.

Caleja C, Toro M de, Gonçalves A *et al.* Antimicrobial resistance and class I integrons in *Salmonella enterica* isolates from wild boars and Bísaro pigs. *Int Microbiol* 2011; 14:19-24.

Cambray G, Guerout AM, Mazel D. Integrons. *Annu Rev Genet* 2010; 44: 141–66.

Carattoli A, Villa L, Pezzella C *et al.* Expanding drug resistance through integron acquisition by IncFI plasmids of *Salmonella enterica* Typhimurium. *Emerg Infect Dis* 2001; 7: 444-7.

Chang CY, Chang LL, Chang YH *et al.* Characterisation of drug resistance gene cassettes associated with class 1 integrons in clinical isolates of *Escherichia coli* from Taiwan, ROC. *J Med Microbiol* 2000; 49: 1097-102.

Chattopadhyay MK, Grossart H-P. Antibiotic and heavy metal resistance of bacterial isolates obtained from some lakes in northern Germany. NSHM *J Pharm Healthcare Manag* 2011; 02: 74-5.

Chen S, Zhao S, White DG *et al.* Characterization of multiple-antimicrobial-resistant *Salmonella* serovars isolated from retail meats. *Appl Environ Microbiol* 2004; 70: 1-7.

Collis CM, Grammaticopoulos G, Briton J *et al.* Site specific insertion of gene cassettes into integrons. *Mol Microbiol* 1993; 9: 41-52.

Collis CM, Hall RM. Expression of antibiotic resistance genes in the integrated cassettes of integrons. *Antimicrob Agents Chemother* 1995; 39:155–62.

Collis CM, Hall RM. Gene cassettes from the insert regions of integrons are excised as a covalently closed circles. *Mol Microbiol* 1992; 6: 2875-85.

Collis CM, Recchia GD, Kim MJ *et al.* Efficiency of recombination reactions catalyzed by class 1 integron integrase IntI1. *J Bacteriol* 2001; 183: 2535–42.

Correia M, Boavida F, Grosso F *et al.* Molecular characterization of a class 3 integron in *Klebsiella pneumoniae. Antimicrob Agents Chemother* 2003; 47: 2838-43.

Deekshit VK, Kumar BK, Rai P *et al.* Detection of class 1 integrons in *Salmonella* Weltevreden and silent antibiotic resistance genes in some seafood-associated nontyphoidal isolates of *Salmonella* in south-west coast of India. *J App Micrbiol* 2012; 112: 1113-22.

DeSouza MJ, Nair S, Loka Bharathi PA *et al.* Metal and antibiotic-resistance in psychrotrophic bacteria from Antarctic Marine waters. *Ecotoxicology* 2006; 15: 379–84.

Domingues SS, Dias C, Saavedra MJ. Detection and characterization of class 1 integrons of Gram-negative bacteria isolates from domestic, wild and food-producing animals. 20th European Congress of Clinical Microbiology and Infectious Diseases (ECCMID). 2010.

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

Fluit AC, Schmitz F-J. Resistance integrons and super integrons. *Clin Microbiol Infect* 2004; 10: 272-88.

Francia MV, Zabala JC, de la Cruz F *et al.* The IntI1 integron integrase preferentially binds singlestranded DNA of the *attC* site. *J Bacteriol* 1999; 181: 6844–9.

Gaze WH, Abdouslam N, Hawkey PM *et al.* Incidence of class 1 integrons in a quaternary ammonium compound-polluted environment. *Antimicrob Agents Chemother* 2005; 49: 1802–7.

Gebreyes WA, Altier C. Molecular characterization of multi-drug resistant *Salmonella enterica* subsp. *Enterica* serovar *Typhimurium* isolates from swine. *J Clin Microbiol* 2002; 40: 2813-22.

Gillings M, Boucher Y, Labbate M *et al.* The evolution of class 1 integrons and the rise of antibiotic resistance. *J Bacteriol* 2008a; 190: 5095–100.

Gillings MR, Krishnan S, Worden PJ *et al.* Recovery of diverse genes for class 1 integron-integrases from environmental DNA samples. *FEMS Microbiol Lett* 2008b; 287: 56–62.

Goldstein C, Lee MD, Sanchez S *et al.* Incidence of class1 and class 2 integrases in clinical and commensal bacteria from livestock, companion animals, and exotics. *Antimicrob Agents Chemother* 2001; 45: 723-6.

Gonzalez G, Sossa K, Bello H *et al.* Presence of integrons in isolates of different biotypes of *Acinetobacter baumanii* from Chilean hospitals. *FEMS Microbiol Lett* 1998; 161: 125-8.

Grinsted J, Cruz FDL, Schmitt R. The *Tn*21 subgroup of bacterial transposable elements. *Plasmid* 1990; 24: 163-89.

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

Guerra B, Soto S, Cal S *et al.* Antimicrobial resistance and spread of class 1 integrons among *Salmonella* serotypes. *Antimicrob Agents Chemother* 2000; 44: 2166-9.

Hall RM, Collis CM. Antibiotic resistance in gram negative bacteria: the role of gene cassettes and integrons. *Drug Resist Updates* 1998; 1: 109-19.

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

Hall RM, Stokes HW. Integrons: novel DNA elements which capture genes by site-specific recombination. *Genetica* 1993; 90: 115-32.

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

Hanson K, Sköld O, Sindstrom L. Non palindromic attI sites of integrons are capable of site-specific recombination with one another and with secondary targets. *Mol Microbiol* 1997; 26: 441-53.

Hardwick SA, Stokes HW, Findlay S *et al.* Quantification of class 1 integron abundance in natural environments using real-time quantitative PCR. *FEMS Microbiol Lett* 2008; 278: 207–12.

Heir E, Lindstedt B-A, Leegaard TM *et al.* 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. *Ann Clin Microbiol Antimicrob* 2004; 3: 12.

Heringa S, Kim J, Shepherd MW *et al.* The presence of antibiotic resistance and integrons in *Escherichia coli* isolated from compost. *Foodborne Pathog Dis* 2010; 7:1297–304.

Hideomi N, Ishikawa T, Yasunaga S *et al.* Frequency of heavy-metal resistance in bacteria from inpatients in Japan. *Nature* 1977; 266: 165-7.

Hochhut B ,Lotfi Y, Mazel D *et al.* Molecular analysis of antibiotic resistance gene clusters in *Vibriocholera* O139 and O1 SXTconstins. *Antimicrob Agents Chemother* 2001; 45: 2991–3000.

Holmes AJ, Gillings MR, Nield BS. The gene cassette metagenome is a basic resource for bacterial genome evolution. *Environ Microbiol* 2003; 5: 383-94.

Ilyina TS. Bacterial Superintegrons, a Source of New Genes with Adaptive Functions *Russian Journal of Genetics* 2006; 42: 1294–302.

Jones LA, McIver CJ, Rawlinson WD *et al.* Polymerase chain reaction screening for integrons can be used to complement resistance surveillance programs. *Commun Dis Intell* 2003; 27: S103–S10.

Kadlec K, Schwarz S. Analysis and distribution of class 1 and class 2 integrons and associated gene cassettes among *Escherichia coli* isolates from swine, horses, cats and dogs collected in the BfT-GermVet monitoring study. *J Antimicrob Chemother* 2008; 62: 469-73.

Kazama H, Hamashima H, Sasatsu M *et al.* Distribution of the antiseptic-resistance gene qacEΔ1 in gram-positive bacteria. *FEMS Microbiol Lett* 1998; 165: 295-9.

Kimura Y, Nakamura T, Kajikawa K *et al.* Antibiotic sensitivity of oligotrophic and eutrophic bacteria. Technical Bulletin of Faculty of Agriculture-Kagawa University 1995; 47: 121-6.

L'Abée-Lund TM, Sørum H. Class 1 integrons mediate antibiotic resistance in the fish pathogen Aeromonas salmonicida worldwide. *Microb Drug Resist* 2001; 7: 263-72.

Laraki N, Galleni M, Tham I *et al.* Structure of In31, a *bla*-containing *Pseudomonas aeruginosa* integron phyletically related to In5, which carries an unusual array of gene cassettes. *Antimicrob Agents Chemother* 1999; 43: 890-901.

Leverstein-van Hall MA, Box ATA, Blok HEM *et al.* Evidence of extensive interspecies transfer of integron-mediated antimicrobial resistance genes among multidrug-resistant Enterobacteriaceae in a clinical setting. *J Infect Dis* 2001; 186: 49-56.

Levesque C, Piche L, Larose C *et al.* PCR Mapping of integrons reveals several novel combinations of resistance genes. *Antimicrob Agents Chemother* 1995; 39: 185-91.

Li D, Yu T, Zhang Y, Yang M *et al.* Antibiotic resistance characteristics of environmental bacteria from an oxytetracycline production wastewater treatment plant and the receiving river. *Appl Environ Microbiol* 2010; 76: 3444–51.

Li, D, Yang M, Hu J *et al.* Antibiotic resistance profile in environmental bacteria isolated from penicillin production wastewater treatment plant and the receiving river. *Environ Microbiol* 2009; 11: 1506-17.

Lindstedt B-A, Heir E, Nygård I *et al.* Characterization of class I integrons in clinical strains of *Salmonella enteric* subsp. *enterica* serovars *Typhimurium* and *Enteritidis* from Norwegian hospitals. *J Med Microbiol* 2003; 52:141-9.

Lupo A, Coyne S, Berendonk TU. Origin and evolution of antibiotic resistance: the common mechanisms of emergence and spread in water bodies. *Front Microbiol* 2012; 3:18.

Marchler-Bauer A, Bryant SH. *"CD-Search: protein domain annotations on the fly. Nucleic Acids Res* 2004; 32: 327-31.

Martinez E, de la Cruz F. Genetic elements involved in Tn*21* site-specific integration, a novel mechanism for the dissemination of antibiotic resistance genes. *EMBO J* 1990; 9: 1275–81.

Martinez-Freizo P, Fluit AC, Schmitz FJ *et al.* Class I integrons in gram negative isolates from different European hospitals and association with decreased susceptibility to multiple antibiotic compounds. *J Antimicrob Chemother* 1998; 42: 689-96.

Martinez-Freizo P, Fluit AC, Schmitz FJ *et al.* Many class I integrons comprise distinct stable structures occurring in different species of enterobacteriaceae isolated from widespread geographic regions in Europe. *Antimicrob Agents Chemother* 1999; 43: 686-9.

Mathai E, Grape M, Kronvall G. Integrons and multidrug resistance among *Escherichia coli* causing community-acquired urinary tract infection in southern India. *APMIS* 2004; 112: 159-64.

Mazel D, Dychinco B, Webb VA *et al.* A distinctive class of integron in the Vibrio cholerae genome. *Science* 1998; 280:605-8.

Mazel D, Dychinco B, Webb VA *et al.* Antibiotic resistance in the ECOR collection: integrons and identification of a novel aad gene. *Antimicrob Agents Chemother* 2000; 44:1568–74.

Mazel D. Integrons: agents of bacterial evolution. *Nat Rev Microbiol* 2006; 4: 608-20.

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

Moura A, Henriques I, Smalla K *et al.* Wastewater bacterial communities bring together broad-host range plasmids, integrons and a wide diversity of uncharacterized gene cassettes. *Res Microbiol* 2010; 161: 58–66.

Moura A, Henriques I, Ribeiro R *et al.* Prevalence and characterization of integrons from bacteria isolated from a slaughterhouse wastewater treatment plant. *J Antimicrob Chemother* 2007; 60:1243-50.

Mudryk Z, Donderski W, Skorczewski P *et al.* Effect of some heavy metals on neustonic and planktonic bacteria isolated from the deep of Gdansk. *Oceanological Studies* 2000; 29: 89-99.

Mukherjee S, Chakraborty R. Incidence of class-1 integrons in multiple antibiotic-resistant Gram-negative copiotrophic bacteria from the River Torsa in India. *Res Microbiol* 2006; 157: 220-6.

Naas T, Mikami Y, Imai T *et al.* 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* 2001; 183: 235-49.

Nageswaran N, Ramteke PW, Verma OP *et al.* Antibiotic Susceptibility and Heavy Metal Tolerance Pattern of *Serratia Marcescens* Isolated From Soil and Water. *J Bioremed Biodeg* 2012; 3: 158.

Nandi S, Maurer JJ, Hofacre C *et al.* Gram positive bacteria are a major reservoir of class 1 antibiotic resistance integrons in poultry litter. *PNAS* 2004; 101: 7118-22.

Nardelli M, Scalzo PM, Ramírez MS *et al.* Class 1 Integrons in Environments with Different Degrees of Urbanization. *PLoS ONE* 2012; 7(6): e39223.

Nawaz M, Khan SA, Khan AA *et al.* Detection and characterization of virulence genes and integrons in *Aeromonas veronii* isolated from catfish. *Food Microbiol* 2010; 27: 327–31.

Nebbia P, Tramuta C, Giammarino M *et al.* Antimicrobial resistance, class 1 and 2 integrons and gene cassettes in avian *Escherichia coli. Ital J Anim Sci* 2008; 7: 391-5.

Neild BS, Holmes AJ, Gillings MR *et al.* Recovery of new integron classes from environmental DNA. *FEMS Microbiol Lett* 2001; 195: 59-65.

Nemergut DR, Martin AP, Schmidt SK. Integron diversity in heavy-metal-contaminated mine tailings and inferences about integron evolution. *Appl Environ Microbiol* 2004; 70:1160-8.

Nemergut DR, Robeson MS, Kysela RF *et al.* Insights and inferences about integron evolution from genomic data. *BMC Genomics* 2008; 9:1–12.

Nesvera J, Hochmannova J, Patek M. An integron of class 1 is present on the plasmid pCG4 from gram-positive bacterium *Corynebacterium glutamicum. FEMS Microbiol Lett* 1998; 169: 391-5.

Nikitin DI, Vishnevetskaia OIu, Zlatkin IV. Clustering of oligotrophic microorganisms on the basis of their antibiotic resistance and dynamic membrane characteristics. *Mikrobiologiia* 1988; 57: 262-7.

Oh H, Lee J, Kim K *et al.* A novel laboratory cultivation method to examine antibiotic-resistance-related microbial risks in urban water environments. *Water Sci Technol* 2009; 59: 346–52.

O'Halloran F, Lucey B, Cryan B *et al.* Molecular characterization of class 1 integrons from Irish thermophilic *Campylobacter* spp. *J Antimicrob Chemother* 2004; 53: 952–7.

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

Ozgumus OB, Celik-Sevim E, Alpay-Karaoglu S *et al.* Molecular characterization of antibiotic resistant *Escherichia coli* strains isolated from tap and spring waters in a coastal region in Turkey. J Microbiol 2007; 45: 379-87.

Ozgumus OB, Sandalli C, Sevim A *et al.* Class 1 and Class 2 Integrons and Plasmid-Mediated Antibiotic Resistance in Coliforms Isolated from Ten Rivers in Northern Turkey. *J Microbiol* 2009; 47: 19-27.

Park JC, Lee JC, Oh JY *et al.* Antibiotic selective pressure for the maintenance of antibiotic resistant genes in coliform bacteria isolated from the aquatic environment. *Water Sci Technol* 2003; 47: 249-53.

Partridge SR, Brown HJ, Hall RM. Characterization and movement of the class 1 integron known as *Tn2521* and *Tn1405. Antimicrob Agents Chemother* 2002; 46: 1288-94.

Partridge SR, Reechia GD, Scaramuzzi C *et al.* Definition of the *att1* site of class 1 integrons. *Microbiology* 2000; 146: 2855-64.

Partridge SR, Reechia GD, Stokes HW *et al.* Family of class 1 integrons related to In4 from *Tn1696. Antimicrob Agents Chemother* 2001; 45: 3014-20.

Pathak SP, Gopak K. Prevalence of bacterial contamination with antibiotic-resistant and enterotoxigenic fecal coliforms in treated drinking water. *J Toxicol Environ Health A* 2008; 71: 427-33.

Pattishall KH, Acar J, Burchall JJ *et al.* Two distinct types of trimethoprim-resistant dihydrofolate reductase specified by R-plasmids of different compatibility groups. *J Biol Chem* 1977; 252: 2319-23.

Paulsen IT, Littlejohn TG, Randstrom P *et al.* The 3′ conserved segment of integrons contains a gene associated with multidrug resistance to antiseptics and disinfectants. *Antimicrob Agents Chemother* 1993; 37: 761-8.

Petersen A, Guardabassi L, Dalsgaard A *et al.* Class 1 integrons containing a *dhfrI* trimethoprim resistance gene cassette in aquatic *Acinetobacter* spp. *FEMS Microbiol Lett* 2000; 182: 73-6.

Ploy MC, Denis F, Courvalin P *et al.* Molecular characterization of integrons in *Acinetobacter baumannii*: description of a hybrid class 2 integron. *Antimicrob Agents Chemother* 2000; 44: 2684-8.

Recchia GD, Hall RM. Gene cassettes, a new class of mobile element. *Microbiol* 1995; 141: 3015–27.

Reechia GD, Stokes HW, Hall RM. Characterisation of specific and secondary recombination sites recognized by the integron DNA integrase. *Nucleic Acids Res* 1994; 22: 2071-8.

Rodríguez I, Martín MC, Mendoza MC *et al.* Class 1 and class 2 integrons in non-prevalent serovars of *Salmonella enterica*: structure and association with transposons and plasmids. *J Antimicrob Chemother* 2006; 58: 1124-32.

Roe MT, Vega E, Pillai SD. Antimicrobial resistance markers of class 1 and class 2 integron bearing *Escherichia coli* from irrigation water and sediments. *Emerg Infect Dis* 2003; 9: 822-6.

Rosser SJ, Young HK. Identification and characterization of class 1 integrons in bacteria from an aquatic environment. *J Antimicrob Chemother* 1999; 44: 11-8.

Rowe-Magnus DA, AM Guerout, P Ploncard *et al.* The evolutionary history of chromosomal super integrons provides an ancestry for multi-resistant integrons. *Proc Natl Acad Sci USA* 2001; 98: 652-7.

Rowe-Magnus DA, Mazel D. Integrons: Natural Tools for Bacterial Genome Evolution. *Curr Opin Microbiol* 2001; 4: 565–9.

Rowe-Magnus DA, Mazel D. Resistance Gene Capture. *Curr Opin Microbiol* 1999; 2: 483–8.

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

Sambrook J, Russell. Molecular Cloning–A Laboratory Manual (3rd ed.). Cold Spring harbor Laboratory Press. New York. 2001.

Sato M, Ahmed AM, Noda A *et al.* Isolation and molecular characterization of multidrug-resistant Gram-negative bacteria from imported flamingos in Japan. *Acta Vet Scand* 2009; 51: 46.

Schmidt AS, Bruun MS, Dalsgaard I *et al.* Incidence, distribution and spread of tetracycline resistance determinants and integron associated antibiotic resistance genes among motile aeromonads from a fish farming environment. *Appl Environ Microbiol* 2001; 67: 5675-82.

Schmitz FJ, Hafner D, Geisel R *et al.* Increased prevalence of class 1 integrons in *Escherichia coli, Klebsiella* species, and *Enterobacter* species isolates over a 7-year period in a German university hospital. *J Clin Microbiol* 2001; 39: 3724-6.

Schmitz FJ, Martinez-Freijo P, Theis S *et al.* Prevalence of class 1 integrons and association with decreased antibiotic susceptibility in German gram-negative blood culture isolates. *Clin Microbiol Infect* 1999; 5: 496-8.

Schnabel EL, Jones AL. Distribution of tetracycline resistance genes and transposons among phylloplane bacteria in Michigan apple orchards. *Appl Environ Microbiol* 1999; 65:4898-907.

Shahid M. Citrobacter spp. simultaneously harboring *bla*CTX-M, *bla*TEM, *bla*SHV, bla_{ampC}, and insertion sequences IS26 and *orf513*: an evolutionary phenomenon of recent concern for antibiotic resistance. *J Clin Microbiol* 2010; 48: 1833-8.

Shi L, Fujihara K, Sato T *et al.* Distribution and characterization of integrons in various serogroups of Vibrio cholerae strains isolated from diarrhoeal patients between 1992 and 2000 in Kolkata, India. *J Med Microbiol* 2006; 55: 575-83.

Shi L, Zheng M, Xiao Z *et al.* Unnoticed spread of class 1 integrons in gram-positive clinical strains isolated in Guangzhou, China. *Microbiol Immunol* 2006; 50:463-7.

Sköld O, Widh A. A new dihydrofolate reductase with low trimethoprim sensitivity induced by an R factor mediating high resistance to trimethoprim. *J Biol Chem* 1974; 249: 4324–5.

Smith JL, Drum DJV, Dai Y *et al.* Impact of antimicrobial usage on antimicrobial resistance in commensal *Escherichia coli* strains colonizing broiler chickens. *Appl Environ Microbiol* 2007; 73: 1404–14.

Stalder T, Barraud O, Casellas M *et al.* Integron Involvement in Environmental Spread of Antibiotic Resistance. *Front Microbiol* 2012; 3: 119.

Stokes HW, Hall RM. A novel family of potentially mobile DNA elements encoding site-specific gene integration functions: integrons. *Mol Microbiol* 1989; 3:1669-83.

Stokes HW, Hall RM. Sequence analysis of the inducible chloramphenicol resistance determinant in the *Tn*1696 integron suggests regulation by translational attenuation. *Plasmid* 1991; 26: 10-9.

Stokes HW, O'Gorman DB, Recchia GD *et al.* Structure and function of 59-base element recombination sites associated with mobile gene cassettes. *Mol Microbiol* 1997; 26:731–45.

Sunde M, Sorum H. Self-transmissible multidrug resistance plasmids in *Escherichia coli* of the normal intestinal flora of healthy swine. *Microb Drug Resist* 2001; 7: 191-6.

Sunde M. Prevalence and characterization of class 1 and class 2 integrons in *Escherichia coli* isolated from meat and meat products of Norwegian origin. *J Antimicrob Chemother* 2005; 56: 1019-24.

Sundin GW. Distinct recent lineages of the *strA-strB* streptomycin- resistance genes in clinical and environmental bacteria. *Curr Microbiol* 2002; 45: 63-9.

Sundstrom L, Radstrom P, Swedberg G *et al.* Site-specific recombination promotes linkage between trimethoprim and sulfonamide resistance genes. Sequence characterization of dhfrV and sulI and a recombination active locus of *Tn21. Mol Gen Genet* 1988; 213: 191-201.

Thungapathra M, Amita, Sinha KK *et al.* Occurrence of antibiotic resistance gene cassettes *aac*(6')-*Ib*, *dfr A5*, *dfr A12*, and *ereA2* in class I integrons in Non-O1, Non-O139 *Vibrio cholerae* strains in India. *Antimicrob Agents Chemother* 2002; 46: 2948-55.

Timoney JF, Port J, Giles J *et al.* Heavy-metal and antibiotic resistance in the bacterial flora of sediments of New York Bight. *Appl Environ Microbiol* 1978; 36: 465–72.

Villa L, Visca P, Tosini F *et al.* Composite integron array generated by insertion of an ORF341-type integron within a Tn 21-like element. *Microb Drug Resist* 2002; 8: 1-8.

Watkinson AJ, Micalizzi GB, Graham GM *et al.* Antibiotic-resistant *Escherichia coli* in wastewaters, surface waters, and oysters from an urban riverine system. *Appl Environ Microbiol* 2007; 73: 5667-70.

White PA, McIver CJ, Rawlinson WD. Integrons and gene cassettes in the *Enterobacteriaceae*. *Antimicrob Agents Chemother* 2001; 45: 2658-61.

Wright MS, Baker-Austin C, Lindell AH *et al.* Influence of industrial contamination on mobile genetic elements: class 1 integron abundance and gene cassette structure in aquatic bacterial communities. *ISME J* 2008; 2: 417-28.

Xi, C, Zhang Y, Marrs CF *et al.* Prevalence of antibiotic resistance in drinking water treatment and distribution systems. *Appl Environ Microbiol* 2009; 75: 5714-8.

Xu H, Su Z, Wang S *et al.* Four Novel Resistance Integron Gene-Cassette Occurrences in Bacterial Isolates from Zhenjiang, China. *Curr Microbiol* 2009; 59: 113-7.

Xu Z, Li L, Alam MJ, *et al.* First Confirmation of Integron-Bearing Methicillin-Resistant *Staphylococcus aureus*. *Curr microbiol* 2008; 57: 264-8.

Xu Z, Li L, Shi L *et al.* Class 1 integron in staphylococci. *Mol Biol Rep* 2011; 38: 5261-79.

Yan H, Li L, Zong M *et al.* Occurrence and characteristics of class 1 and 2 integrons in clinical bacterial isolates from patients in South China. *J Health Sci* 2010; 56: 442–50.

Yang H, Byelashov OA, Geornaras I *et al.* Characterization and transferability of class 1 integrons in commensal bacteria isolated from farm and nonfarm environments. *Foodborne Pathog Dis* 2010; 7: 1441–51.

Zhao WH, Hu ZQ. IMP-type metallo-beta-lactamases in Gram-negative bacilli: distribution, phylogeny, and association with integrons. *Crit Rev Microbiol* 2011; 37:214–26.

Zlatkin IV, Vishnevetskaia Olu, Nikitin DI. Some aspects of antibiotic resistance of oligotrophic bacteria. *Antibiot Khimioter* 1991; 36: 34-7.