Detection of integron-encoded integrases in environments with moderate anthropogenic impact from Puerto Rico

By

Zenaida Castro Diaz

A Thesis Submitted in Partial

Fulfillment of the Requirements for the Degree of

Master In Science

in

Biology Science

University of Puerto Rico Mayaguez Campus 2012

Approved

Rafael Montalvo Rodríguez, Ph.D.	Date
Chairperson, Graduate Committee	
Carlos, Rodríguez Minguela Ph.D.	Date
Member, Graduate Committee	
Carlos Santos Flores Ph. D.	Date
Member, Graduate Committee	
Nannette Diffoot Carlo, Ph.D	Date
Member, Graduate Committee	
Rosa I. Román Pérez	Date
Representative of Graduate Studies	
Nannette Diffoot Carlo, Ph.D	Date
Chaiperson of the Department	

Abstract

Integrons are genetic elements with a site-specific recombination system which includes an integrase (IntI) capable of capturing and expressing promoterless, circular DNA molecules known as gene cassettes. Although integrons were initially discovered as elements involved in the transfer of multiple antibiotic resistance mechanisms within the clinical scene, their presence has been increasingly reported in environments with other integrated adaptative traits unrelated to antibiotic resistance functions. This project explores the presence of integrons in microbial communities from environments with moderate anthropogenic impacts in Puerto Rico. These include two extreme environments: hypersaline waters from Cabo Rojo and the Coamo Thermal Springs and soil from the Culebrones Cave in Arecibo. Samples from these places were analyzed to determine the presence of integrases by culture-dependent and culture-independent techniques. Metagenomic DNA was analyzed with an *int*-targeted PCR assay using degenerate primers. Alignments were performed for all sequences recovered and their relationship with other tyrosine recombinases was analyzed. Sequences were also compared against functional integrases for their characterization. Important motifs characteristic of the tyrosine recombinase family such as: Patch (II and III) and Box (I and II) were detected. Twenty eight novel integrase sequences were recovered from metagenomic DNA extracted from hypersaline water and thirty integron integrases from cave soil. Additionally, eight new integron-encoded integrases were recovered from the Coamo Thermal Springs. Furthermore, integrase genes similar to integrase class 2 that do not contain the premature stop codon X179 were detected by PCR in two halophilic isolates closely related to the genera *Salinivibrio* (γ -*Proteobacteria*) and *Bacillus*. One isolate from the Coamo Thermal Springs related to the genus Geobacillus showed integrase genes similar to integrase class 1. This study provides the first evidence of the presence of integrase integrons among the microbial community found in environments with moderate anthropogenic impacts from Puerto Rico and in the Caribbean. The results obtained also support the argument that the integron-based mechanism for horizontal gene transfer is widespread and can take place in solar saltern ponds, Coamo Thermal Springs and cave soil.

Resumen

Los integrones son elementos genéticos con un sistema de recombinación especifica el cual incluye una integrasa (IntI) capaz de capturar y expresar moléculas de DNA circulares y sin secuencias promotoras, conocidas como genes casetes. Aunque los integrones fueron inicialmente descubiertos en ambientes clínicos como elementos envueltos en la transferencia de múltiples mecanismos de resistencia, se ha reportado su presencia en ambientes con otras funciones adaptativas no relacionada a funciones de resistencia a antibióticos. Este proyecto explora la presencia y diversidad de integrones en comunidades microbianas de ambientes con impacto antropogénico moderado en Puerto Rico. Este incluye dos ambientes extremos las aguas hipersalinas de Cabo Rojo, los baños termales de Coamo y el suelo de Cueva Culebrones en Arecibo. Las muestras de cada lugar fueron analizadas para determinar la presencia de integrasas de integrones por técnicas dependientes e independientes de cultivos. Se analizó ADN metagenómico mediante un PCR especifico para Int utilizando cebadores degenerados. Se realizaron alineamientos para todas las secuencias recuperadas y se analizó su relación con otras recombinasas de tirosinas. Las secuencias fueron comparadas contra otras integrasas para su caracterización. Importantes "motifs" característicos de la familia de las tirosinas recombinasas fueron detectados como lo son: "Patch" (II y III) y "Box" (I y II). Veintinueve secuencias de integrasas fueron recolectadas a partir de ADN metagenómico extraído de las aguas hipersalinas y treinta integrasas de integrón del suelo de Cueva Culebrones. Ocho secuencias de integrón integrasas fueron recolectadas desde los Baños Termales de Coamo. Genes de integrasa relacionada a integrasa clase 2, que no contienen el codon de terminación prematuro X179, fueron detectados por PCR en dos aislados halofílicos relacionados al género Salinivibrio (γ -Proteobacteria) y Bacillus. Un aislado desde los Baños Termales de Coamo relacionado al género de Geobacillus presentó genes de integrasas similares a los de la integrasa clase 1. Este estudio provee la primera evidencia de la presencia de integrón integrasas entre comunidades microbianas encontradas en ambientes con impacto antropogénico moderado de Puerto Rico y el Caribe. El resultado obtenido, además, apoya el argumento de que los mecanismos de integrón para la transferencia horizontal de genes es esparcido y puede tomar lugar en ambientes halofílicos, termofílicos y en suelo de cuevas.

© Zenaida Castro Díaz, 2012

Dedication

Several important people are part of the success of my research and I want to dedicate this thesis to them, my Family. They are the most important people in my life (Biky, Mom, Dad, Na-talia), without their help and support I would not be the human that I am now. They were my inspiration for reaching my goal because they believed in me and in the hard moments they were always there to give me support and advice. Another important person I want to dedicate my thesis to is Josué Dumeng Valentín. He provide me peace and calmness during the last hard moments and frustration through love and patience.

Acknowledgements

I would like to thank God since without His help none of this work would have been possible. I would like thank my family (my mother Miriam Díaz, my father José L. Castro, my sister, Natalia Castro, and my little niece Biky) for their love, patience, companionship, advice, and support in all moments in my life, specially in my graduate years. I would like to thank them for their support and constant motivation, my aunt Alicia Castro, her husband Miguel A. Manzano, and my little cousins Fonsi and Andrea; they were very important during my graduate years.

I would like to thank Dr. Rafael Montalvo Rodríguez for his guidance and friendship and for giving me the opportunity to be his graduate student. He has been an excellent professor, mentor, and friend. I would like to thank Dr. Carlos Rodríguez Minguela for his advice, guidance, and dedication in my thesis and for inspiring me to discover the importance of integrons. I would also like to thank Dr. Carlos Santos and Dr. Nannette Diffoot for all their advice and guideance throughout this research.

Several students played huge roles in this research project: Lisaurie López, Ivelise Rodríguez, Frances Cervonii, Namir Massini, and Haydeliz Martínez.

I would like to thank Magaly Zapata for all her help, friendship and support during all these years. Thanks to Donato Seguí and José Almodóvar for their help in this research. Thank you both for your valuable time and support. Thanks to Prof. Ana Vélez for giving me also the oportunity to teach and for guiding me on how to be an educator.

A special thanks to all my undergraduate students from the cell physiology laboratory for being part of my professional growth as an educator in the biological science field.

I would like to thank all the members of the Extremophiles Laboratory for their support, friendship, and companionship throughout these years. Specially, I would like to thank my

friends who gave me support and valuable friendship, Alexa M. Báez Crespo, Carol V. López, Lisaurie López, and Gina Jimenez. Finally, I want to thank the "Integron Girls" (Odalys Álvarez and Glendalys Vargas) for their support.

Table of Contents

Abstract	iii
Resumen	v
Dedication	vii
Acknowledgements	ix
List of Figures	xiv
List of Table	xiv
Introduction	2
Literature Review	11
Objectives	12
Part I. Sampling and Detection of Integron-encoded Interases from	Metagenomic DNA
Material and Methods	20
Results	35
Discussion	
Part II. Sampling and Detection of Integron-encoded Integrases from	n Isolates
Material and Methods	42
Results	51
Discussion	53
Conclusions	54
Recommendations	55
Literature Cited	

List of Figures

- **Figure 1.1** Different mechanisms of horizontal gene transfer are presented. A) Transduction: bacteriophages can be integrated into a bacterial chromosome providing new phenotypic characteristic to the cell; B) Conjugation: plasmid DNA is transferred from a donor to a recipient cell throught the use of a pilus: C) Transformation: cell-free DNA is taken and incorporated into the bacterial genome. Adapted and modified from Levy and Marshall 2004.
- **Figure 1.2** Representation of a class 1 integron and its integration mechanism. The region common to all integrons (5'-CS) generally consists of these elements: an integrase coding gene (*intl*); the site-specific recombination sequence *att1* and two promoters, P, that drives the expression of integrase P(int) and the expression of the inserted gene cassette P(c). The 3'-CS is characteristic of class 1 integrons and contains permanent genes that confer resistance against quaternary ammonium compounds (qacE $\Delta 1$) and sulfonamide drugs (*sul1*). There is a second type of recombination site carried by gene cassettes that generally consist of a open reading frame (in this figure is *orf 2*) and presents at its 3' end a short imperfect inverted repeat sequence also called, *attC* which serves as a recognition site for the integrase (Collis, and Hall 1992).
- **Figure 1.3** Representation of the core sequences found in the 5' and the 3' inverted core sequences (ICS) at the *attC* recombination site.
- Figure 1.4 Location of sampling sites. Maps were obtained from Google Earth.
- **Figure 1.7** Diagram which demonstrates the common region for the integron-encoded integrase that was amplified in the PCR assay.
- **Figure 1.8** Metagenomic DNA extractions from the Solar Saltern Ponds (a,b,c), Frat =(Fraternidad), Can = (Candelaria), SP =(Salterns Ponds), Pt =(Puente), (d) soil from Culebrones Cave, CA1 and CA2= (soil within the cave), CE1 and CE2 = (soil at the entrance cave), and the Coamo Thermal Springs (e), Co1 and Co2 = (Coamo 1 and 2).
- **Figure 1.9** 16S rDNA amplification from metagenomic DNA using the universal primers Bact-1392R and Bact-27 from the Solar Saltern Ponds, Culebrones Cave and the Coamo Thermal Springs. The positive control is an 16S rDNA amplicon from a strain isolated from mangrove soil, and the negative control is a PCR sample without template DNA.
- **Figure 1.10** *Int-* PCR amplicons using universal, Hep35 and Hep36 of the total community DNA extracted from Fraternidad Lagoons as template. Positive control contained integrase genes as template from Mona Island soil, and negative control is a PCR sample without template DNA.
- **Figure 1.11** *Int-* PCR amplicons using universal, Hep35 and Hep36 of the total community DNA extracted from Culebrones Cave (a) and Coamo Thermal Springs (b). Posi-

tive control contained integrase genes as template from Mona Island soil, and negative control is a PCR sample without template DNA.

- Figure 1.12 Partial amino acid sequence alignment of environmental integrases from Culebrones Cave against other tyrosine recombinases, IntI class 1 from Escherichia coli, IntI 2 from Shigella sonnei, IntI 3 from Klebsiella pneumoniae, IntI 4 from Vibrio cholerae and tyrosine recombinases XerC and XerD from E. coli. The position of conserved motifs among integron-encoded integrases is indicated by segments labeled as Boxes I and II and Patches II and III and as additional domains (AD) under the alignment. The degree of conservation (percent identity) among equivalent residues is indicated as follows: blue (100%), yellow (95% and 85%). Conserved amino acids (100% identical) with functions related to protein folding, DNA binding, and recombination activity are highlighted in red, respectively. Green columns correspond to H, R, and Y residues from the conserved RHRY tetrad, which is characteristic of the entire tyrosine recombinase family (Nunes-Dubt et al. 1998). GenBank accession numbers for the protein sequences of reference integrases (classes 1 to 4) and tyrosine recombinases XerC and XerD are AAQ16665, AAT72891, AAO32355, 99031763, P0A8P6, and P0A8P8, respectively.
- Figure 1.13 Partial amino acid sequence alignment of environmental integrases from the Solar Saltern Ponds against other tyrosine recombinases, IntI class 1 from Escherichia coli, IntI 2 from Shigella sonnei, IntI 3 from Klebsiella pneumoniae, IntI 4 from Vibrio cholerae and tyrosine recombinases XerC and XerD from E. coli. The position of conserved motifs among integron-encoded integrases is indicated by segments labeled as Boxes I and II and Patches II and III and as additional domains (AD) under the alignment. The degree of conservation (percent identity) among equivalent residues is indicated as follows: blue (100%), yellow (95% and 85%). Conserved amino acids (100% identical) with functions related to protein folding, DNA binding, and recombination activity are highlighted in red, respectively. Green columns correspond to H, R, and Y residues from the conserved RHRY tetrad, which is characteristic of the entire tyrosine recombinase family (Nunes-Dubt et al. 1998). GenBank accession numbers for the protein sequences of reference integrases (classes 1 to 4) and tyrosine recombinases XerC and XerD are AAQ16665, AAT72891, AAO32355, 99031763, POA8P6, and POA8P8, respectively.
- Figure 1.14 Partial amino acid sequence alignment of environmental integrases from the Coamo Thermal Springs against other tyrosine recombinases, IntI class 1 from *Escherichia coli*, IntI 2 from *Shigella sonnei*, IntI 3 from *Klebsiella pneumoniae*, IntI 4 from *Vibrio cholerae* and tyrosine recombinases XerC and XerD from *E. coli*. The position of conserved motifs among integron-encoded integrases is indicated by segments labeled as Boxes I and II and Patches II and III and as additional domains (AD) under the alignment. The degree of conservation (percent identity) among equivalent residues is indicated as follows: blue (100%), yellow (95% and 85%). Conserved amino acids (100% identical) with functions related to protein folding, DNA binding, and recombination activity are highlighted in

red, respectively. Green columns correspond to H, R, and Y residues from the conserved RHRY tetrad, which is characteristic of the entire tyrosine recombinase family (Nunes-Dubt et al. 1998). GenBank accession numbers for the protein sequences of reference integrases (classes 1 to 4) and tyrosine recombinases XerC and XerD are AAQ16665, AAT72891, AAO32355, 99031763, P0A8P6, and P0A8P8, respectively.

- **Figure 1.15** Consensus neighbor-joining tree illustrating the phylogenetic relationship of IntI sequences recovered from from environments witht moderate anthropogenic impacts in Puerto Rico. Values represent the percentage of 1,000 boostrap replications that supported the branch. Outgroup references were designate with a thicker black branches highlight XerC, XerD, class IntI 1, IntI 2, IntI 3, and IntI4. Red triangle are sequences recovery from Solar Saltern Ponds, blue triangle circle from Coamo Thermal Springs and green circle are recovery sequences from Culebrones Cave. Numbers 1 to 10 denote the clusters formed by recovery sequences and the percent identities of IntI sequences associated are shown in parentheses. Roman numbers (I and II) are recovery sequences related to clinical integron. Values of boostrap support below 49% are not demonstrated and Genbank accession numbers follow the outgroup sequences.
- **Figure 2.1** Macroscopical and microscopical characteristics of strain SII5 and PII5 isolated from the Solar Saltern Ponds on Luria Bertanie medium.
- **Figure 2.2** Macroscopical and microscopical characteristic of strain CS91 isolated in a previous study (Burgos 2007) from the thermophilic springs, Coamo on Thermus medium, using light microscopy.
- **Figure 2.3** Genomic DNA extraction from prokaryotic cells isolated from the Solar Saltern Ponds PII2 and SII5 and CS91 isolated from Coamo Thermal Springs by Burgos, A., 2007.
- **Figure 2.4** 16S rDNA amplification using universal primers Bact-1392R and Bact-27F of the isolated strains from the Solar Saltern Ponds.
- **Figure 2.5** *Int-* PCR amplicons of PII5 strain isolated from the Solar Saltern Ponds of Cabo Rojo. Positive control contained integrase genes as template, and negative control without template DNA.
- **Figure 2.6** Partial amino acid sequence alignment of integrases from strains isolated against other tyrosine recombinases, IntI class 1 from *Escherichia coli*, IntI 2 from *Shigella sonnei*, IntI 3 from *Klebsiella pneumoniae*, IntI 4 from *Vibrio cholerae* and tyrosine recombinases XerC and XerD from *E. coli*. The position of conserved motifs among integron-encoded integrases is indicated by segments labeled as Boxes I and II and Patches II and III and as additional domains (AD) under the alignment. The degree of conservation (percent identity) among equivalent residues is indicated as follows: blue (100%), yellow (95% and 85%). Conserved amino acids (100% identical) with functions related to protein folding, DNA binding, and recombination activity are highlighted in red, respectively. Green col-

umns correspond to H, R, and Y residues from the conserved RHRY tetrad, which is characteristic of the entire tyrosine recombinase family (Nunes-Dubt et al. 1998). GenBank accession numbers for the protein sequences of reference integrases (classes 1 to 4) and tyrosine recombinases XerC and XerD are AAQ16665, AAT72891, AAO32355, 99031763, P0A8P6, and P0A8P8, respectively.

- **Figure 2.7** Consensus neighbor-joining tree (compared positions for 3 amino acid residues), illustrating phylogenetic relationship of IntI sequences recovery from halophilic and thermophilic strains. Values represent the percentage of 1,000 boostrap replications that supported the branch. Outgroup references were designate with a thicker black branches highlight (XerC, XerD, class IntI 1, IntI 2, IntI 3, and IntI4. Numbers select the clusters formed by recovery sequences and the percent identities of IntI sequences associated. Values of boostrap support below 20% are not demonstrated and Genbank accession numbers follow the outgroup sequences.
- Figure 2.8 Neighbour-joining distance tree using 16S rDNA sequences of SII5 isolated from solar salterns in Cabo Rojo and closely related species. Boostrap values higher than 50% are shown. *Halomonas marisflava* KCCM 80003T (AF2511431) was used as the outgroup
- **Figure 2.9** Neighbour-joining distance tree using 16S rDNA sequences of PII5 isolated from solar salterns in Cabo Rojo and closely related species. Boostrap values higher than 50% are shown. *Zymobacter palmae* DSM 10491T (D14555) was used as the outgroup.
- Figure 2.10 Neighbour-joining distance tree using 16S rDNA sequences of cs91 previously isolated by Burgos, A. from Coamo Thermal Springs and closely related species. Boostrap values higher than 50% are shown. *Chromohalobacter israelensis* DSM6768T (AF2118621) was used as the outgroup.
- **List of Tables**
- **Table 1.1**Summary of physical parameters from the Solar Saltern Ponds (Cabo Rojo).
- **Table 1.2**Summary of physical parameters from the Coamo Thermal Springs (Coamo).
- **Table 1.3**Summary of physical parameters from the Culebrones Cave (Arecibo).
- **Table 2.1**Oligonucleotides used for PCR analyses.

Introduction

Integrons are genetic elements encoding a site-specific recombination system capable of capturing and expressing promotorless DNA units, called gene cassettes (Hall 1989, Mazel 2007; Rodríguez-Minguela et al., 2009; Stokes 1989). Most of the integron-related studies focus their attention on hospital environments since they were discovered within the context of the antibiotic resistance problem. Fecal and industrial effluents have been linked to the presence of integronencoded antibiotic resistance in environmental settings (Rodríguez-Minguela et al. 2009; Elsaied et al., 2011). However, a worldwide scale study revealed that integrase from clinical integrons are not resctricted to the hospital scene (Rodríguez-Minguela et al. 2009). Integrons found in natural environment are important since they can be associated with the acquisition of new adaptive functions and the establishment of antibiotic resistance reservoirs (Koenig et. al., 2008; Stokes et al. 2001). However, there are few reports on the presence and role of integrons in extreme environments. Some of these reports include studies of the metagenome associated with uranium mines and deep sea hydrothermal vents, which revealed the presence of gene cassettes potentially encoding adaptative functions unrelated to antibiotic resistance genes as well as new integron integrase sequences (Ghauri et al. 2003; Elsaied et al. 2007).

So far, no studies concerning the diversity and distribution of integrons have been conducted in extreme environments across the Caribbean region. The present study was focused on determining the presence of integron integrases in solar salterns, thermal springs and cave ecosystems located at Cabo Rojo, Coamo and Arecibo, respectively. This was achieved by using PCR-based methods with metagenomic DNA retrieved from the above mentioned sites and with genomic DNA from three halophilic and thermophilic isolates related to the genus *Salinivibrio* (γ -Proteobacteria), *Bacillus* and *Geobacillus*. This information will provide further insight on the distribution of the module and its potential for driving evolutionary change and adaptation in extreme environments.

Literature review

I. Mechanism of horizontal gene transfer

Horizontal gene transfer (HGT) is the flow of genetic information across species and between bacteria, HGT has played an important role in bacterial evolution, adaptation into communities (Sobecky and Hazen 2009: van-Elsas and Bailey 2002), and also has significantly contributed to the aggravation of the antibiotic resistance problem in the absence and presence of selective pressure (Figure 1.1). Recently, HGT has been implicated in the dispersal of genes involved in cellular processes such as photosynthesis (Igarashi et al., 2001), nitrogen fixation (Sullivan et al., 2001), carbon fixation (Utaker et al., 2002), sulfate reduction (Friedrich 2002), and pathogenicity (Hentschel, U., and J. Hacker, 2001; Gamieldien et al., 2002). Mechanisms for lateraltransfer between bacteria such as phages (Jiang and Paul 1998) transposons, and plasmids (Sobecky et al., 1998) have helped to widespread resistance without antibiotic exposure.

Transduction is a mechanism where genetic material can be transferred horizontally from a bacterial cell to another through bacteriophages. In this process, the phage genome can be integrated into the bacterial chromosome provoding new phenotypic characteristics. The phage excision process is not always exact, therefore bacterial genes can be incorporated into the virus and when it infects another cell, those genes can be integrated into the host genome (Ochman et al. 2000; Toussaint and Merlin 2002). In transformation, free DNA can be transferred horizontally from cell to cell and through a recombination process is incorporated into the host genome. The conjugation process is perfomed throught a sex pili where a conjugal tube is formed and a copy of single stranded DNA is passed to the recipient cell. The conjugation process is common in the intestinal tract of animals where bacterial populations are high (Garcia-Quintanilla, M., et al.,2008).



Figure 1.1 Different mechanisms of horizontal gene transfer are presented. A) Transduction: bacteriophages can be integrated into a bacterial chromosome providing new phenotypic characteristic to the cell; B) Conjugation: plasmid DNA is transferred from a donor to a recipient cell throught the use of a pilus: C) Transformation: cell-free DNA is taken and incorporated into the bacterial genome. Adapted and modified from Levy and Marshall, 2004.

II. Integrons

By the late 1980s Stokes and Hall identified the integron system as new type of mechanism in bacteria for the gain of genetic material encoding antibiotic resistance traits in bacteria. Integrons contain an integrase that mediates a site specific recombination reaction leading to the expression of open reading frames which generally lack promoters. Integrons integrases catalize the excision and rearrangement of resident cassettes (Stokes, H. W., and Hall, R.M. 1989). The gene that encodes the integrase is located at the 5' conserved region and it belongs to the superfamily of tyrosine recombinases. Members of this family are enzymes that perform site–specific recombination and require short and specific DNA sequences involving the formation of a Holliday junction intermediate. This mechanism involves cleavage and rejoining of DNA strands without ATP utilization or DNA synthesis (Gopaul, D. N., and G. D. Van Duyne, 1999) The integrase nucleophile residue consists of a tyrosine (Tyr–302) located at the 5' conserved region and is responsible for the cleavage and formation of a covalent intermediate by esterification of a DNA 3' phosphoryl group (Grainge, I., and M. Jayaram, 1999: Mazel, 2006). The tyrosine recombinase family is characterized by the presence of functional residues within conserved motifs (boxes I and II and patches I, II and III). However, integron integrases relative to other tyrosine recombinases contain distinctive additional domain residues around patch III (Gravel, A., N. Messier, and P. H. Roy. 1998; Nield B.S., et al. 2001; Nunes–Duby, et al. 1998). In addition to the integrase gene, the 5' conserved region has also DNA sequences that act as insertion (att I) site for the transcription (Pc) of gene cassettes (Figure 1.2). The integron integrase also recognizes a specific core sequence (GTTRRRYY) found at the 3' end of gene cassette in order to catalize the insertion.

Integrons class 1, 2 and 3 can be found in many gram-negative bacteria of clinical origin, specifically among members of the *Enterobactericiae* (Antunes, P., et al. 2006; Young-Kang H., et al., 2005; Martinez-Freijo, P. et al., 1998). Recently, integrons have been found in gram-positive bacteria such as *Corynebacterium glutamicum* (Fluit and Schmitz, 1999; Nesvera, J. et al., 1998) and *Sulfobacillus* genera (Ghauri, M.A., et al., 2003). The advance of genomics and the availability of numerous genome sequences from environmental bacteria allow to explain that integrons are ancient structures and widespread gene capture systems (Mazel, D., 2006).



Figure 1.2 Representation of a class 1 integron and its integration mechanism. The region common to all integrons (5'-CS) generally consists of the following elements: an integrase coding gene (*intl*) the site-specific recombination sequence *att1* and two promoters, P, that drives the expression of integrase P(int) and the expression of the inserted gene cassette P(c). The 3'-CS is characteristic of class 1 integrons and contains permanent genes that confer resistance against quaternary ammonium compounds (qacE ΔI) and sulfonamide drugs (*sul1*). There is a second type of recombination site carried by gene cassettes that generally consist of a open reading frame (in this figure is *orf 2*)and presents at its 3' end a short imperfect inverted repeat sequence also called, *attC* which serves as a recognition site for the integrase (Collis, C.M. and Hall, R.M. 1992).

The integrons are divided into two different groups, although both contain similar organization: mobile integrons and chromosomal integrons differ. The mobile integrons are associated with transposons and plasmids, which are elements responsible for integron spread. The gene cassettes present in the mobile integrons are mainly involved in antibiotic resistance (Mazel, D, et al., 2010). Chromosomal integrons or superintegrons were discovered with a distinctive integrase named VchIntA which is located in the chromosome of *Vibrio cholerae* and was not associated with mobile DNA elements.

Recent studies had demonstrated that integron integrases are activated during SOS response, resulting in an increase of cassette rearrangements (Baharoglu, Z., et al., 2010). SOS is a bacte-

rial stress response which is induced as a consequence of abnormal rate of single stranded DNA (ssDNA) present in the cell, probably caused by UV radiation or exposure to antibiotics (Baharoglu, Z., et al., 2010). LexA is a repressor of the SOS response and the integron integrase RecA is an activator of both by stimulating the auto-proteolysis of LexA (Guerin, E., et al., 2009: Baharoglu, Z., et al., 2010).

III. Gene Cassettes

Gene cassettes exist as promotorless DNA molecules outside of the integron, in a closed circular form, and only are functional when inserted at specific locations in the host genome (defined by the presence of an integron) (Stokes, H.W. and Hall, R.M., 1989: Collis et al., 1993). The source of gene cassette remains unknown, but their expression is influenced by their position reactive to the Pc promoter. Their structure consists of a single open reading frame and short sequences of flanking DNA on both the 3' and 5' ends, called cassette attachment site (attC) or 59 base elements (59be). The attC sequence serves as integrase recognition site during the recombination event. These sequences vary in length from 54 bp to 141 bp, although there are two consensus regions. The core sequence found in the 5' and consist of the nucleotides GTTRRRY (R = purine; Y = pyrimidine) and the 3' inverted core sequences (ICS) consists of the nucleotides RYYYAAC (Figure 1.3). However, inverse core site in the *attI* site not is found therefore it cannot form a palindromic structure like *attC*. The open reading frame in the vast majority of the cases encodes a protein necessary for antibiotic resistance, however gene cassette encoding novel and unknow functions are being increasingly reported (Nield, S, et al., 2001). Recent studies suggests have shown other adaptative functions not associated with antibiotic resistance such as degradation of industrial wastes associated resistance to heavy metals, nitroaromatic catabolism, biosynthesis and DNA modification (Koenig, J. E., et al., 2009; Nemergut,

D.R. et al., 2004). Moreover, virulence factors, toxin-antitoxin systems and polysaccharide biosynthesis are associated with gene cassette functions (Ghauri, M.A., et al., 2003; Koenig et al., 2008).



Figure 1.3 Representation of the core sequences found in the 5['] and the 3['] inverted core sequences (ICS) at the *attC* recombination site.

IV. Integron classification

Integrons are classified based on the difference in the percentage of identity of the integrase gene (Mazel, D., 2006) or their predicted amino acid sequences. The most common integrons found in many bacterial species are class 1 and 2 and are associated at clinical settings. Integrons class 4, 5, 6, 7 and 9 are associated with a novel class of integrons, called super-integrons.

Class 1 integrons are found in many pathogenic and commensal bacteria associated with humans and animals (Cambray, G., et al., 2010). Environments exposed to fecal contamination show prevalence for integrons class 1 and are typically associated to replicative Tn21 transposons (Rodríguez, Minguela, et al., 2009). Also, they are the most widespread and studied of all integron classes. Two signature traits of class 1 integrons are the presence of $qacE\Delta$ and sulf1genes at the 3'-CS region, which confer resistance to quaternary ammonium compounds and to sulphonamides drugs respectively Mazzel, D., 2006). However, there are reports where class 1 integrons have been detected lacking this typical 3'-CS region (Saenz, Y., et al., 2010). Class 2 integrons are characterized primarily for the presence of an internal stop codon at amino acid position 179 in the *intl2* gene. Class 2 elements are mobilized by non replicative transposons *Tn7*. Recently, studies have described intI2 variants without the internal stop codon, in environmental samples (Rodriguez-Minguela, C, et al. 2009). In the terminal 3'-CS there are gene cassettes that conferred resistance against trimethroprim (*dfrAI*a o *dfrAI*b), streptothricin (*sat*) and streptomycin/spectinomycin (*aadA*1) (Sabaté, M. and Prats, G., 2002; Sundström, L., P., et al., 1998; Sundström, L. et al., 1991; Sundström, L. et al., 1990).

Classes 3 integrons are associated with transposons located into uncharacterized plasmids and are less frequently detected than class 2. Class 3 integrons had been described with gene cassettes encoding a metalobeta-lactamase in isolates of *Serratia marcescens* from Japan (Osano, E., et al., 1994). Two environmental strains, *Delfia acidovorans* C17 and *Delfia tsurruhatensis* A90 were found to contain class 3 integrons gene with gene cassetts encoding unidentified functions (Xu, H. et al., 2010).

Other integron classes include a class 4 integron identified in *Vibrio cholera* associated with mobile genetics element integrated in the chromosome via conjugation denominated, SXT. Also a Class 5 element found in compound transposon associated with a pRSV1 plasmid of *Alivibrio salmonicida* has been described (Mazel et al., 2010).

V. Super integrons or chromosomal integrons

Another group of distinctive integrons, called superintegrons or chromosomal integrons, have been found in many γ -proteobacteria such as Vibrionaceae and Xanthomonads (Cambray et al., 2010). Rowe-Magnus et al. 1999 postulate that the super-integron is an ancient element, possibly pre-dating the antibiotic era by thousands of years. These gene cassettes contain high similarity between their *attC* sequences and are located in the chromosome without association to mobile elements. Class 4 was the first chromosomal superintegron (*Vibrio cholera*) and is known to contain 176 gene cassettes (Mazel et al., 1998; Rowe-Magnus et al., 2001). The largest superintegrons identified to date are in *Vibrio vulnificus* CMCP6 and YJO16; these are 153 and 138 kb long, and contain 219 and 188 gene cassettes respectively (Chen et al., 2003). It has been estimated that as little as five *Vibrio spp*. contain a gene cassette load equivalent in size to a small genome (Chen et al., 2003). Relative to clinical integrons, super–integrons contain a minor load of cassettes encoding antibiotic resistance; most of the genes associated with super–integrons encode functions unrelated to resistance traits (Rowe–Magnus, 2009).

Several superintegron cassettes contain homology to enzymes involved in energy metabolism and detoxification, such as 4-oxalocrotonate tautomerase which enables bacteria to use various aromatic hydrocarbons as their sole energy source (Whitman, 2002), a carbohydrate transport protein of the major facilitator superfamily, an adenylate kinase that interconverts ADP to AMP and ATP for driving energy requiring processes, and a lactoylglutathionelyase detoxification enzyme that converts toxic methylglyoxal (MG) to lactic acid (Ferguson et al., 1998). However, cassettes encoding virulence factors and DNA modification functions have been also described (Ogawa and Takeda, 1993).

VI. Integrons in extreme environments

Extreme environments are habitats distinguished by physical parameters such as salinity, temperature, pH, pressure, and oxygen concentrations that are extreme in relation to those present where eukaryotic organisms preferentially live (Stetter, 1998; Konings et al., 2002). Microorganisms found in extreme environments have considerable industrial application and potential for biotechnology development. Enzymes produced by extreme microorganism have unique features to tolerate extreme conditions such as high temperatures, extreme pH values, high pressure and high salt (Van der Burg, 2003). However, the distribution and diversity of the integron platform across these extreme habitats remain relatively unexplored.

In genomic studies of halophilic bacteria, *Salinibacter ruber* demonstrated the ability to conduct lateral gene transference with halophilic members of the Archaea domain through transposons and possible presence of a integron integrase class 1 (Mongodin et al., 2005). Additional studies in man made salt mines were preformed in Pakistan where different gene cassettes were recovered from halotolerant strains related to *Halomonas magadiensis*, *Virgibacillus halodenitrificans* and *Yania flava*, and from metagenomic samples (Ghauri et al., 2006). The possible functions of these gene cassettes were associated with protein families of putative transcriptional regulators and a putative betaketoacyl reductase.

Integrons have been reported among acidophilic moderate thermophilic bacteria related to *Sulfobacillus thermosulfidooxidans* recovered from uranium mines, as well as in strains isolated from deep–sea hydrothermal vent (Ghauri et al., 2003: Elsaved, et al., 2007). Class 1 integron has been characterized in a thermophilic *Campylobacter* (O'Halloran et al., 2004). Moreover, integron integrases were detected in the acidophilic bacterium *Acidithiobacillus ferrooxidans* whereas 14 novel integron integrases and new gene cassette implicated in the metabolism of nitroaromatic compound have been detected in metal–contamined mine tailings (Rowe-Magnus and Mazel, 2001; Nemergut et al., 2004).

To date, no studies concerning the diversity and distribution of integrons platforms have been conducted from environments with moderate anthropogenic impact and extreme environments in Puerto Rico. The present study focused on determining the presence of integron integrases in Solar Saltern Ponds, thermal springs and a cave ecosystem located at Cabo Rojo, Coamo and Arecibo, respectively. **Objectives.** Explore the presence and diversity of integron–encoded integrases in environments with moderadte anthropogenic impact using culture-independent and culture-dependent techniques; the two major objectives were the following:

- **1.** Determine the presence and diversity of integrin encoded integrases by constructing and analyzing clone libraries of PCR-amplified metagenomic DNA.
- 2. Determine the presence of integrin encoded integrases via PCR from strains isolated from the Solar Saltern Ponds (Cabo Rojo), and Coamo Thermal Springs (Coamo).

Part I. Sampling Procedure

Material and Methods

I. Sampling sites: Description of the Solar Salterns in Cabo Rojo, Coamo Thermal Springs in Coamo, and Culebrones Cave in Arecibo, Puerto Rico.

1.1 Solar Saltern Ponds

The solar saltern are located on the southwest coast of Cabo Rojo. This system is a thalassohaline extreme environment which includes an estuary surrounded by natural mats, a mangrove forest, artificial ponds and saline lagoons (Fraternidad and Candelaria) with minimal exposure to anthropogenic impact (Fig. 1.4). The principal reason for this is due to its designation as a National Wildlife Refuge by the U.S. Fish and Wildlife Service since 1974. The Candelaria Lagoon is connected to the sea through a channel which we denominated: Puente. The solar saltern has limited contact with the sea and the constant high solar radiation together with the lack of rain during most of the year promotes the formation of a hypersaline environment (Departamento de Recursos Naturales, 1986). Additionally, the solar saltern are the most important stop site for migratory shorebirds in the Caribbean (U.S Fish Wildlife Service). Previously, researches have been studying the prokaryotic diversity of these salterns (Montalvo–Rodriguez et al., 1997, 1998, 2000), nevertheless no studies have been performed to determine the diversity and distribution of integron integrases.

1.2 Coamo Thermal Spring

There is at least one natural environment in Puerto Rico which can be considered alkalithermophilic. This habitat is the Coamo Thermal Springs located in the municipality of Coamo (south central Puerto Rico) which forms at the east-early trending of the Cordillera Central (Glover, 1971) (Fig. 1.4). Most of this area lies on the flank of a much-faulted geanticlinal volcanic core that is partly overlapped by a sequence of carbonate rocks and sediments (Glover, 1971). It is an unique thermal spring located on the island and its origin is not totally clear, although some theories about the origin do exist (Giusti, 1971). Until this moment, the most accepted theory is based on a geological fault in this area and the natural thermal gradient that exist on the terrestrial cortex (Quiñonez and Guzmán, 1983). Prokaryotic diversity of these thermal springs has been studied (Burgos, 2007), however no studies have been conducted with respect to detection of integron-encoded integrases.

1.3 Culebrones Cave

The Mata de Plátano Field Station is located 7 kms southwest of Arecibo, Puerto Rico, and inside of the reserve, a cave denominated Culebrones is found (Fig. 1.4). Culebrones Cave is a refugee for approximately 13 different bats species (Rodríguez-Durán and Soto-Centeno, 2003). Large numbers of native boas congregate to feed on bats. Culebrones Cave is considered a hot cave with temperatures ranging from 26–40°C and relative humidity exceeding 99% (Rodriguez-Duran, 1998; Rodríguez-Durán and Soto-Centeno, 2003). Despite the ecological importance of cave ecosystems and their features, no studies have been performed in these environments regarding integrases integrons.

II. Sampling Sites.

To confirm our experimental hypothesis, in August 2008 different environments with moderate anthropogenic impacts in Puerto were analyzed to detect integron-encoded integrases. The environments analyzed with moderate anthropogenic impacts in Puerto include Culebrones Cave, Solar Saltern Ponds, Coamo Thermal Springs and Culebrones Cave. The samples were collected in duplicate using Whirl Pack®. Culebrones Cave soil samples were retrieved from two distinct points, soil at the entrance cave and soil within the cave. The temperature outside was $31.5 \,^{\circ}$ C and inside of Culebrones Cave was $38 \,^{\circ}$ C. The samples were stored at $-20 \,^{\circ}$ C. In the case of Solar Saltern Ponds, water samples were collected from Candelaria and Fraternidad lagoons, artificial ponds (solar ponds) and "puente" (Table 1.1). The Coamo Thermal Springs water was collected from the mainstream where anthropogenic effect are minimal. Thermal water temperature was $39 \,^{\circ}$ C and pH was 8.9. Solar saltern (25 ml) and thermal springs (500ml) water samples were homogenized with a pipette following the filtration in 47mm, $0.45 \,\mu$ m (pore size) cellulose nitrate membrane filters (Whatman, Germany). Filtered membranes were stored at $-20 \,^{\circ}$ C.



Fig 1.4 Location of sampling sites. Maps were obtained from Google Earth.

Table 1.1 Summary of physical parameters from the Solar Saltern Ponds (Cabo Rojo)

Sampling site	Temperature °C	Salinity % w/v
Fraternidad	40	25
Candelaria	38	8
"Puente"	36	4
Solar ponds	40	35

Sampling site	Temperature °C	pH
Mainstream	39 ℃	8.9

Table 1.2 Summary of physical parameters from the Coamo Thermal Springs (Coamo)

Table 1.3 Summary of physical parameters from the Culebrones Cave (Arecibo)

Sampling site	Temperature °C	Humidity
Cave entrance (Cueva entrada)	31.5	90
Inside cave (Cueva adentro)	38	80

III. Metagenomic DNA extraction:

Metagenomic DNA extractions from cave soil, solar saltern ponds and thermal spring water were performed using the MoBio DNA extraction kit @ according to manufacturers instructions. Modification were performed to cause cell lysis; beat beading technique was used through vortex for 5 minutes at maximum speed followed by centrifugation at 10,000 rpm for 10 minutes. The isolated metagenomic DNA was resuspended in 50 µl molecular water and treated with RNAse (20 µg/µl) for 30 minutes at 37°C. The DNA quality was verified on 0.8% agarose gels and staining with ethidium bromide. All metagenomic DNA's were used as templates for PCR amplification.

IV. PCR amplification of 16S rDNA and gel electrophoresis:

To prove the quality of the extracted DNA, the gene encoding the 16S rRNA was amplified by PCR using the forward primer Bact-27F and the reverse primer Bact-1392R (Table 2.1). The reaction mixture contained the following concentrations for a final volume 30 μ l : 1X Buffer, 2.5 mM Mg, 1X BSA New England®, 2.5 mM dNTP's, 1 pMol of BACT 1392 and BACT 27, 0.026 U *Taq* polymerase (Promega) and 1 μ l of template (15 ng/ μ l). The PCR reaction consisted of initial denaturation of 5 minutes at 94° C; 30 cycles of denaturation 1 minute at 94° C; annealing 1 minute at 50 ° C and extension 3 minutes at 72° C, and the final extension at 72° C for 10 minutes (Hezayen *et al.*, 2002). The PCR product quality was verified on 1.0 % agarose gels and staining with ethidium bromide.

V. Int-PCR and gel electrophoresis:

The conserved regions for the integrase genes were amplified using the following reagents for a final volume of 12.5 μ l : 1X Buffer, 3 mM Mg, 1X BSA New England®, 1.0 mM dNTP's, 10 pMol of degenerate primers hep35 and hep 36 (Table 2.1), 0.0104 U *Taq* polymerase (Promega) and 1 μ l of template (15 ng/ μ l).

Amplification cycles were the following: initial denaturation at 94 °C by 5 minutes; 35 cycles of denaturation at 95 °C; annealing at 52 °C and extension at 72 °C; the final extension was at 72 °C for 7 minutes. The PCR product quality was verified in 1.5 % agarose gels stained with eth-idium bromide.



Figure 1.7 Diagram which demonstrates the common region for the integron-encoded integrase that was amplified in the PCR assay.

VI. Ligation and transformation:

The PCR product was purified using Wizard SV Gel and PCR clean up system (Promega, Madison, WI, USA) according to manufactures instructions and was ligated into pGEM Vector Systems (Promega, Madison, WI, USA). The ligation mixture was transformed into JM109 *Escherichia coli* by heat shock according to manufacturer's instructions.

VII. Alkaline – SDS Plasmid Minipreps and sequencing

Clones containing inserts were incubated overnight at 37°C (1 ml). The plasmid isolation was performed using the Hanahan procedure. Basically, cells were resuspended in lysis buffer (25 mM Tris-HCl pH 8.0, 50 mM Glucose and 10 mM EDTA) followed by denaturalization with 0.2 N NaOH and 1% SDS solution of 5 M Potassium Acetate. The isolated plasmids were resuspended in 50 μ l of TAE and treated with RNAse (20 μ g/ μ l). The plasmid was resolved by electrophoresis in agarose (1.5%) and visualized after staining with ethidium bromide (10ug/mL). Plasmid concentration was determined using the nanodrop.

VIII. DNA sequencing:

Fifty clones were randomly chosen and sent for sequencing at the UPR-Río Piedras Sequencing Facility and UW-htseq Sequencing Facility (Washington University at St. Louis, MO) using the T7 primer. Samples were prepared according to facility specifications.

IX. Characterization and Phylogenetic analyses of intI sequences:

The environmental integron integrases were edited in Bioedit and aligned against sequences in GenBank databases using ClustaX and Bioedit. The sequences from Int1, Int2, Int3, Int4, XerC, and XerD (GenBank accession number AAQ16665.1, AAT72891.1, AA2A3VO32355.1, 99031763 P0A8P6.1, and P0A8P8.1) were used as references for the identification of functional residues and domains. The tyrosine recombinase: XerC and Xer D were chosen to identify additional domains characteristic of integron integrases. Comparative amino acid sequence analyses were performed by constructing consensus neighbor-joining dendrograms using Mega4 software package as previously described (Rodríguez–Minguela et al., 2009). The dendograms were generated with p-distance model and tested by 2,000 boostrap replications.

Results:

Metagenomic DNA extractions from environments with moderate anthropogenic impacts from Puerto Rico:

Culebrones Cave soil, Solar Saltern Ponds and Coamo Thermal Springs samples were used for total metagenomic DNA. Figure 1.8 demonstrates the quality of ten metagenomic DNA extractions performed from Solar Saltern Ponds, Culebrones Cave, and Coamo Thermal Springs. Metagenomic DNA extracted using membrane and bead beading technique was quantified using the Nanodrop ND-1000 Spectrophotometer© and adjusted so there were 15 and 12ng of template for 16S rDNA region and integrase integron PCR amplification, respectively. Universal primers were used for the partial amplification of integron integrases (491bp) and 16S rDNA genes (900bp).



10K

Figure 1.8 Figure 1.8 Metagenomic DNA extractions from the Solar Saltern Ponds (a,b,c), Frat =(Fraternidad), Can = (Candelaria), SP =(Salterns Ponds), Pt =(Puente), (d) soil from Culebrones Cave, CA1 and CA2= (soil within the cave), CE1 and CE2 = (soil at the cave entrance), and the Coamo Thermal Springs (e), Co1 and Co2 = (Coamo 1 and 2).
Metagenomic 16S rDNA from environments with moderate anthropogenic impact from Puerto Rico:

The quality of these metagenomic DNA's were determined and then used as templates for 16S rDNA amplification by PCR. PCR amplicons were of approximately of 900 pb and the DNA markers were 1kb (Figure 1.11).



900

Figure 1.9 16S rDNA amplification from Metagenomic DNA using the universal primers Bact 1392R and BACT 27 from the Solar Saltern Ponds, Cueva Culebrones and the Coamo Thermal Springs. The positive control (+) is an 16S rDNA amplicon from a strain isolated from mangrove soil, and the negative control (-) is a PCR sample without template DNA. Frat =(Fraternidad), Can = (Candelaria), SP =(Salterns Ponds), Pt =(Puente), CA1 and CA2= (soil within the cave), CE1 and CE2 = (soil at the cave entrance), and Co1 and Co2 = (Coamo 1 and 2).

Int-PCR from metagenomic DNA from environments with moderate anthropogenic impacts from Puerto Rico:

The metagenomic DNA's extracted were used as a templates for integrase gene amplification using degenerate primers hep35 and hep 36 nucleotide sequences complementary to the Box I and Box II domains found in integron integrases. The amplification was performed in triplicate and PCR products of approximately of 491 pb were obtained from of Solar Saltern Ponds, Culebrones Cave, and Coamo Thermal Springs (Figure 1.10 and 1.11). These were cloned for each place. It was necessary to perform optimization of the PCR annealing temperature in order to achieve optimal amplification of the desire products. The purified PCR products of each place



was cloned and nine clone libraries were constructed and 100 clones were obtained. Only 50 clones from each clone libraries were selected randomly for sequencing purposes.

Figure 1.10 *Int-* PCR amplicons using universal, Hep35 and Hep36 of the total community DNA extracted from Fraternidad Lagoons as template. Positive control (+) contained integrase genes as template from Mona Island soil, and negative control (–) is a PCR sample without template DNA. Frat =(Fraternidad), SP =(Saltern Ponds), Pt =(Puente), Can = (Candelaria).



Figure 1.11. *Int*- PCR amplicons using universal, Hep35 and Hep36 primer and total community DNA extracted from Cueva Culebrones (a) CA1 and CA2= (soil within the cave), CE1 and CE2 = (soil of the entrance cave) and Coamo Thermal Springs (b) Co = (Coamo). Positive control (+) contained integrase genes as template from Mona Island soil, and negative control (-) is a PCR sample without template DNA.

Int Alignment analysis from environments with moderate anthropogenic impacts from Puerto Rico:

Although, the full length of IntI putative sequences were not retrieved by the restriction primer pair used, the integron integrase conserve motifs and residues critical for binding and recombination mechanisms could be indicative that recovery Int sequences can be functional. Environmental IntI sequences recovered were aligned against functional integron integrases of a class1 integrase described in *Escherichia coli*, IntI2 described in *Shigella sonnei* and IntI3 de scribed in *Klebsiella pneumonia* (GenBank accession numbers: AAQ166651, AAT728911 and AAO323551, respectively). The numbers in the frames of the alignment were determined using Int4 (99031763) numbering because their crystal structure and recombination site were determined (MacDonald et al., 2006). Residue R143 corresponds to an arginine residue found within the Box I region of Int4.

Functional motifs which are characteristic of the tyrosine recombinase family such as: Patch (II and III) and Box (I and II) were detected in all recovery sequences from extreme environments (Fig. 1.12, Fig.1.13, and Fig. 1.14). The signature tetrad RHRY characteristic of the tyrosine recombinase family was detected. Also, these residues are enclosed within the Box II motif and corresponded to positions: H268, R271 and Y303. H268 is implicated in the appropriate orientation and stabilization of the transition state during the integrative reaction (Messier and Roy, 2001). Y303 is the nucleophile responsible for cleavage and forms a covalent intermediate by esterification. Moreover, amino acids related to catalytic and DNA binding functions were 100% conserved: K160, (Box I), A205, L206, K209 and F223 (additional domain), and G293 (Box II). Residues A205 and L206 play a role in positioning and stabilization of the active site with respect to the attC site. Residue K160 (Patch II) is an important amino acid in maintaining an also important function in the cleavage activity (Rodríguez-Minguela, et al., 2009).

The P211 (additional domain) was conserved 100% in all IntI sequences from Culebrones Cave and Coamo Thermal Springs (Fig. 1.12 and Fig.1.14), while 80% of the sequences from Solar Saltern Ponds were conserved. Residue A205 (additional domain) was conserved 100% in samples from Culebrones Cave and Coamo Thermal Springs. However, 80% of the residue 205 was conserved from Solar Saltern Ponds (Fig. 1.13), with an exception of p57 and can 47, where substitutions were observed by another hydrophobic amino acid (leucine) and polar residue (serine). Residue P211 (additional domain) was conserved 100% in all IntI recovery sequences from Culebrones Cave and Coamo Thermal Springs. Meanwhile, 80% of the residue P211 in the se-

quences from Solar Saltern Ponds was conserved and substituted by acid residue lysine and hydrophobic amino acid glycine (Fig. 1.13).

The crystal structure of VchInt4 and the role residues important for attachment and stabilization of the synaptic complex have been recently described (MacDonal et al., 2006). The residue W219 associated with attachment and residues P232, H240, and H241 are conserved 80 and 100% across the IntI sequences from Culebrones Cave and conserved 100% in sequences recovery from the Solar Saltern Ponds and Coamo Thermal Springs (Fig. 1.12, Fig.1.13, and Fig. 1.14). Residues P232, H240 and H241 are implicated in stabilizing the complex. Residue K145 was conserved 100% in all environment sequences except in the case of Int4 residue Q145 which appeared to be exclusive of the *Vibrio* clade (Rodriguez-Minguela, 2009).

Residues D191 and P211 (additional domain) are conserved 100% in IntI sequences recovery from the Culebrones Cave and thermal springs and 90% conserved in IntI sequences from Solar Saltern Ponds. Two other residues, L202 and D231, (additional domain) 80 and 90% were conserved in all sequences.

Residues L291, L292, G293, and H294 were also 100% conserved in all sequences from Culebrones Cave and Coamo Thermal Springs and have been implicated in having a key role in the recombination mechanism (Messier and Roy, 2001).

The IntI sequences recovery from Culebrones Cave (Fig.1.12) contains residues as P224 and V222 and were identified within the Patch III region with 90% conservation. A high degree of sequence conservation was observed within the Patch II region in which residues G159, G162, R165, and L169 between 80% and 90% were conserved. In the Box II the residues: H272, F274, A275, T276, L278, L279, D284, I285, R286, V288, Q289, L292 G 293 H294, T299, T300, I302,

T304, H305, V306 were conserved 100% . The residues H277, G282, T287, and D296 were conserved 99%. In addition, the residues E290, L307 were conserved 80% within of the Box II.

Similarly the Int sequences in the residues from Solar Saltern Ponds were identified as P224 (Patch III) with conserved 80%. Other residues were found in the Patch II; G162, R165 were 100% conserved and L169 was conserved 80%. A high degree of sequence conservation was observed within the Box II region in which residues H272, F274, A275, T276, I285, L292, T299, T300, I302, T304, H 305, V306 were conserved 100% and L307 was conserved 99%. L279 and D284 were conserved 90% within the Box II and Q289 was conserved 99%, s34 was substituted by the hydrophobic amino acid glycine. R286 is 99% conserved in all sequences from Solar Saltern Ponds. G293 and H294 were found conserved 100% and L292 was conserved 99%; all are important amino acids related to the recombination mechanism (Messier, N and Roy, P.H., 2001). Residue H277 was conserved 80% and substitution by other polar and hydrophobic amino acid residue as arginine, glutamine and alanine were detected. G282 and L307 were conserved 99% within of the Box II.

Other residues found in Coamo Thermal Springs (Fig.1.14) were found within the Patch II: G159, G162, K163, R165, L169 that were conserved 100%. In the Box II conserved 100% and residues were identified residues across all the sequences: H272, S273, F274, A275, T276, H276, H277, L278, L279, G282, D284, R286, V288, Q289, T299, T300, I302, T304, H305, V306 and L307. Within Box II, residues: I285, T287, D296, E290 and L307 were conserved 99%.

28

		150	160		170	180	190	200	2	10	220	230	240	D
01029	Distant in the	EACHUTT	VDHCMONO	DETTM	POATAFPI	OPHIADUW	TENEN	TTPC VC PUT		Dega py gut				
ca1238	RV KDP1	FASHHII	VRHGRGAQ	DRUTM	DEST TOPI	QRHLARVK	CT UED	- LLDGYGEVVL	TALAK	U PSAEKSWL	HIV	PASKKS ID	RSGQERRL	÷
ca120	DUVDP	FAVHOLT	VRDGRGNR	DEVIM	POSLETT	OPOLAPTY	T VHVE	- DANGIGLVIL	VA PAD	DENA CT SUP	×.	PDA SUDSTD	DRCOPPTI	÷
ca1233	DUVDU	FAROFIL	VEDGUGSK	DUTU	POST PTAT	OKOTOFAT	FT HDE	LASCHONUM	HATAR	V PKG PFEVG	č.	PRA AG PSUD	LTGNRGRO	÷
ca28	RVKDT	FAOROII	VRDGKGNK	DRVTM	PESLIOPI	OAHLOOVK	SLHEDD-	- LANGYGEVYL	P FA LDR	V PNA SE EWS	ov	PARNISUD	RSGUTER	T.
ca213	RVKDFI	FAYHOLI	VRDAKGAO	DRVTM	POSLEETI	OROLARTK	LVHVED-	- RAAGYGEVYL	YAFAR	D PHAGT SWE	DOY V	PASKRSID	RSGOERRL	Т
ca229	RVKDLI	FARQEIL	VRDGKGSK	DRVTV	POSLRTAI	QKQIQFAI	ELHRED-	- LASGHGNVWL	HALAR	Y PKG PFEVG	ÔY	PAAGFSVD	LIGNRGRO	Ē
ca232	RVKDVI	FASHHIT	VRHGKGAO	DRLTM	POALAEPL	ORHLARVK	LFHEAD-	- LLDGYGEVLL	YA FAR	DPSAEKSWL	BÔYV	PATKRSVD	RAGVERR	I
ca238	RVKDFI	FASHHIT	VRHGKGAQ	DRLTM	LPQAIAEPL	QRHLARVK	LFHEAD-	- LLDGYGEVVL	YA FAR	DPSAEKSWL	ĤYV	FRATKRSVD	RAGVE RR HI	Ι
ca218	RVKDVI	FASHHIT	VRHGKGAQ	DRLTM	PQAIAEPL	QRHLARVK	LFHEAD-	- LLDGYGEVLL	P YA FAR	DPPAEKSWL	RQYV	PRATKRSVD	RAGVE RR H	Ι
ca23	RVKDVI	FAYHQLI	VRDAKGAH	DRVTM	PQSLEETI	QRQLARTK	LVHVED-	- RAAGYGEVYL	P YA FAR	D PHAGT SWE	DQYV	PPASKRS ID	RSG QE RR HI	Ι
ca217	RVKDLI	FASHHIT	VRHG <mark>KG</mark> AQ	DRLTM	LPQAIAEPI	QRHLARVK	LFHEAD-	 LLDGYGEVLL 	P YA FAR P	D PSAEKSWL	N QY V	PRATKRSVD	RAGVE RR HE	I
ca219	RVKDLI	FASHHIT	VRHGRGAQ	DRLTM	LPQAIAEPI	QRHLARVK	LFHEAD-	- LLDGYGEVLL	P YA FAR	D PSAEKSWL	QYV	PATKRSVD	PRAGVE RR H	I
ca220	RVKDM	FASHHIT	VRHGRGAQ	DRLTM	LPQAIAEPI	QRHLARVK	LFHEAD-	- LLDGYGEVLL	P YA FAR P	DPSAEKSWL	QYV	PATKRSVD	RAGVERRH	Ι
ca25	RV KDVI	FAYHQLI	VRDAKGAH	DRVTM	LPQSLEETI	QRQLARTK	TAHAE -	- RAAGYGEVYL	YAFAR	D PHAGT SWE	QYV	PASKRSID	RSGQERRH	I
cel3	RVKUM	FAYRQIV	IRDGROOK	DRVIM	PVFLVDAL	RKQIAKVA	ELHRED-	- LSAGAGEVNL	E A SK	ITSAARELG	N X X	PASTIAND	TISKEQR	÷
cell3	RV KLIVI	FAVDOTU	TRDGRGAK	DRVIM	PVELVDAL	RKQIAKVA	ELIKE.	- LSAGAGEVNL	PEALSKI	Y PARARELG	NY I	PASTIAND	DTEDODD	÷.
cel 21	RV ADVI	FAIRQIV	VPDGVGAO	DEVIM	POSLEEAL	HDOLADTK	TVHEF	- LANGAGEVAL	VA FAD	D PHA CT SWP	P IV	PDA ST ISUD	Deepppp	Ť
ce1 39	DUVDU	FAVROTV	TRDGVGOK	DEVUD	DODDVETI	OOHLENVK	DEHOED.	LONGYGAVEL	VAMAR	V POA POEMA	ŏv.	PRSTISUT	DRSS SUPP H	÷
ce147	RVKDVI	FDROOLV	VRDGKGAK	DRITM	PTRLLEPI	ODHLRIVR	RT HVED	- LARGYGAVYL	DALER	Y PNANE EWP	ÖYU	PRAETISSD	RTSGVRRH	Ľ
ce1 62	RVKDVI	FDROOLV	VRDGKGAK	DRVTM	PVFLVDAL	RKOIAKVA	ELHRED-	- LSAGAGEVNL	FALSR	YPSAARELG	IYON	PASTIAND	ITSRE OR HE	I
ce246	RVKDLI	FSRKEIV	VREGKGFR	DRVTM	PEKLSGEL	ARHIERVQ	ILHRAD-	- LERGHGSVYL	YALDR	Y PNA OT DWG	D ÔYV	PRSDRLSTD	RSGVTRRH	А
ce243	RVKDLI	FERGEIL	IRDGKGFK	DRVTM	PKILADHL	RLQLLRVK	LFHEAD-	- LLDGHGEVLL	YAFDR	DPSAGKAWL	QYV	PATKRSVD	RAGVERRH	I
ce236	RVKDVI	LERLQVT	VRFGRGDK	DRVTL	SENLVERI	SEHRSKLR	KL FAQ	- RGACLAGVWI	PEALER	Y PRAGV SWE	QWL	PSKQLMHD	HSG LQRR H	v
ce227	RVKDFI	FAYHQLI	VRDAKGAQ	DRVTM	PQSLEEAL	QRQLARTK	LVHTED-	- LAAGYGEVY <mark>L</mark>	YA FAR	D PHAGT SCE	QYV	PPA SKRS LD	RSGRERR H	v
ce26	RVKDLI	FERGEIL	IRDGKGFK	DRVTM	PKILADHI	RLQLLAAQ	RLHGED-	- LEAGFGSVYL	YALAE	Y PGREKDWG	QYV	PPS PLRS ID	PRSDVE RR HI	L
ce224	RVKDFI	FERGEIL	IHDGRGFK	DRVTM	PKILADHI	RLQLLAAQ	RLHDED-	- LEAGFGSVYL	YA LAE	Y PGR GK DWG	QYV	PPS PLRS ID	RSDVERRH	м
ce221	RVKDFI	FELAQIT	VRDAKGGK	DRVTM	PVNL SEPI	RRHLLRIK	AQHEQD-	- LEDGFGTVHL	PFALSR	IL PNA AR EWA	QYV	PPS SR IS ID	RIGKKORH	I
ce220	RVKDVI	LERLQVT	VRFGKGDK	DRVTL	SENLVERI	SEHRSKLR	KL FAQ	- RGACLAGVWI	PEALER	Y PRAGV SWE	QWL	PSKQLMHD	HSG FORR H	Ľ
ce214	KA KDTI	FARKEIT	VRSGRGAK	DRRVP	PERLREAL	QRQLEQAR	LLHDGD-	- LAAGLGETRL	HALGR P	Y PNAASEWG	C	PAORLSRD	RDGTMRRH	ĸ
ce210	RVKD	FAQRQII	VRDGRGNK	DRVIM	PESLIQPI	QAHLQQVK	SLHEDD-	- LANGIGEVIL		I PNASKEWS		PARNLSVD	RSGVIRR H	÷
00021762 To+TA	RV KULI	FDYCATR	TWOCHCCK	MOTUT	AVET VOUT	VEODALAN	DVVDD1	HOW YC CUM		V DALA DV P PD	2.1	PRESCRESVE	PRDIMPD IN	Ľ,
AA0166651 IntI4	DURDIT	FDHGTIT	VDFCVCSK	DPATM	DESLADSI	PFOLSPAD	AWWIK	- OFCRECUT		V DDA CH SWD	PW	CHTHSTO	DSCIA/PD H	2
AAT728911 IntT2	DUVDE	FDNGCIT	VHDGYCCK	SPNST	PTPL TPAT	KXLIFOAR	LTOOD	LOGVGPSL	FAIDH	V PSA VROAA	MET	PRSTICNE	VNG KT. CR HI	T.
AA0323551 IntI3	RVKDVI	FDRHATT	VRSGKGDK	DRVVM	PRALVPRI	RAOLIOVR	AVWGOD	RATGROGUYL	HALER	Y PRAGE SWA	FW	PRAKLSVD	OTGVERRH	T.
POA8P61 XERC	DIKHLI	LESGEVW	VM-GKGSK	ERRLP	GRNAVAWI	EHWLDLRD)		LFGSEDI		AL	LS	KLGK	RI
POA8P81 XERD	TMSDIS	LRQGVVR	VI-GRGNK	ERLVP	GEEAVYWL	ETYLEHGR	PW	L	LNGVSII)	VL	PS	QRAQQ	QM
		•		Patch	∏ →				ما ما نه : م به		▶ ◄	┝╸╺╴	n -	
	Box I							F	Addition		п	atch III		
		250	260	27	0 2	80	290	300	Domain	(AD)	P	attim		
10.20						1								
ca12.50	DESELO	A LEAD ALL	NO NG LURKS	SCHIL	BUCEL TUTT	PACYNTRE	WEL GRY							
ca1213	TETMIO	A TERAT	POACLOKPS	ROWTI	DUSPLTHIT	ADEVATOTI	NOPI CHO							
ca1233	DPKR TOT	AVKOAT -	IT AG THELV	TEHVI	RHSFATHLL	FTGSDIRT	JOEL GH	WKTTOTYTHUT.						
ca28	DESSLOT	A VER AV-	VA AGLNK PA	SCHTE	RHCFATHLL	ENGYDIRT	JOEL GHY	USTTNIVTHUT.						
ca213	TETMLO	AIKHAI-	ROAGLDKRG	SCHTL	RHCFATHLL	EDGYDIRT	JOEL GHR	URTIMITHUL						
ca229	DPKRIO	RAVKOAT -	ITAGIHKLV	TPHVL	RHSFATHLL	ETGSDIRTY	OEL GHS	VKTTOIYTHVL						
ca232	AATVLQ	RAVINAI-	HRAGVQKRS	SCHTL	RHSFATHLL	ADGYDIRT	QEL GHO	VRTTMITTHVL						
ca238	AATVLQI	RAVKNAI-	HRAGVOKRS	SCHTL	RHSFATHLL	ADGYDIRT	VQEL GHO	DVR TTMIYTHVL						
ca218	AATVLOI	RAVKNAI-	HRAGVQKRS	SCHTL	RHSFATHLL	ADGYDIRT	VQEL GHO	VRTTMITHVL						
ca23	TETMLQI	KAVKHAI-	RQAGLDKRG	SCHTL	RHCFATHLL	EDGYDIRTV	VQEL GHR	THE R. P. LEWIS CO., LANSING MICH.						
ca217	AATVLQI	RAVENAI-	HRAGVQKRS	SCHTL	DHOFA THIL	ADOVDTOT!		CONKTIMITINUL						
ca219	AATVLQI	RAVENAI-				CONTRACTOR OF	VQEL GHQ	DVRTTMITTHV-						
ca220	AATVLQI		HRAGVQKRS	SCHTL	RHSFATHLL	ADGYDIRT	NQEL GHQ NQEL GHQ	OVRITMITHV-						
ca25		RAVENAI-	hr ag vokrs hr ag vokrs	SCHTL	RH SFATHLL RH SFATHLL	ADG YDI RTV ADG YDI RTV	NQEL GHQ NQEL GHQ NQEL GHQ	DVRITMITHV- DVRITMITHV- DVRITMITHVV						
L L L J	TETMLQI	KAVKNAI- KAVKHAI-	HR AG VQKRS HR AG VQKRS RQ AG LDKRG	SCHTL SCHTL SCHTL	RHSFATHLL RHSFATHLL RHCFATHLL	ADG YDIRT ADG YDIRT EDG YDIRT	AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHQ	UVRITMIYTHV- UVRITMIYTHVV UVRITMIYTHVV UVRITMIYTHVV						
112	TETMLQI DERI LQI	RAVKNAI- KAVKHAI- RAAKQAI-	HR AG VQKRS HR AG VQKRS RQ AG LDKRG RE AG INKRA	SCHTL SCHTL SCHTL SCHTL	RHSFATHLL RHSFATHLL RHCFATHLL RHSFATHLL	ADGYDIRT ADGYDIRT EDGYDIRT ENGYDIRT	AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHR AQEL GHR	UVRITMIYTHV- DVRTTMIYTHV- DVRTTMIYTHVV DVRTTMIYTHVV UVRTTMIYTHVL UVTTTMIYTHVL						
cell3	TETMLQI DERILQI DERILQI	RAVKNAI KAVKHAI RAAKQAI RAAKQAI	HR AG VOKRS HR AG VOKRS RO AG LDKRG RE AG INKRA RE AG INKRA	SCHTL SCHTL SCHTL SCHTF	RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT	AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHR AQEL GHR AQEL GHR	OVRITMIYTHV DVRITMIYTHV DVRITMIYTHVV OVRITMIYTHVL OVTITMIYTHVL OVTITMIYTHVL						
cell3 cell6 cel21	TETMLQI DERILQI DERILQI SRLVLQI	RAVKNAI- KAVKHAI- RAAKQAI- RAAKQAI- RAVKAAI-	HR AG VQKRS HR AG VQKRS RQ AG LDKRG RE AG INKRA RE AG INKRA RKAK IAKAA	SCHTL SCHTL SCHTL SCHTF SCHTF	RH SFATHLL RH SFATHLL RH CFATHLL RH SFATHLL RH SFATHLL RH SFATHLL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ETG YDIRT	AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHR AQEL GHR AQEL GHR AQEL GHR	URINIAL DURITMIYTHU DURITMIYTHU DURITMIYTHU DURITMIYTHU CURITMIYTHU CURITMIYTHU DUNITMIYTHU						
cell3 cell6 cel21 cel39	TETMLÖ DERILO SRLVLO SETVLO	RAVKINAI KAVKHAI RAAKQAI RAAKQAI RAVKAAI RAVKAAI KSVKHAI	HR AG VQKRS HR AG VQKRS RQ AG LDKRG RE AG INKRA RE AG INKRA RK AK IAK AA RP AG LDKRG	SCHTL SCHTL SCHTL SCHTF SCHTF SCHTF	RH SPATHLL RH SPATHLL RH SPATHLL RH SPATHLL RH SPATHLL RH SPATHLL RH SPATHLL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ETG YDIRT EDG YDIRT	AGEL GHO AGEL GHO AGEL GHO AGEL GHO AGEL GHO AGEL GHO AGEL GHO AGEL GHO	UVRITMIYTHU DVRITMIYTHU DVRITMIYTHU DVRITMIYTHU DVRITMIYTHU DVITTMIYTHU DVITTMIYTHU DVITTMIYTHU DVITTMIYTHU						
cell3 cell6 cel21 cel39 cel47	TETMLQI DERILQI SRLVLQI SETVLQI HENGLQI DESCLOI	RAVKNAI KAVKHAI RAAKQAI RAAKQAI RAVKAAI KSVKHAI KAVQKAAR KAVQKAAR	HR AG VQKRS HR AG VQKRS RQ AG LDKRG RE AG INKRA RE AG INKRA RKAK IAKAA RP AG LDKRG EA AG LAKRV DA TG T EK DA	SCHTL SCHTL SCHTL SCHTF SCHTF SCHTF SCHTL NPHTL	RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT	AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHX AQEL GHX AQEL GHX AQEL GHX AQEL GHX AQEL GHX AQEL GHX	UVRIMIYIHV DVRIMIYIHV DVRIMIYIHV UVRIMIYIHV UVRIMIYIHVL UVRIMIYIHVL UVRIMIYIHVL UVRIMIYIHVL UVRIMIYIHVL						
cell3 cell6 cel21 cel39 cel47 cel62	TETMLQI DERILQI SRLVLQI SETVLQI HENGLQI DESGLQI DEBLLQI	RAVKNAI KAVKHAI RAAKQAI RAAKQAI RAVKAAI KAVQKAAR KAVQKAAR KAVQKAAR KAVRAAV	HR AG VQKRS HR AG VQKRS RQ AG LDKRG RE AG INKRA RE AG INKRA RKAK IAKAA RP AG LDKRG EA AG LAKRV RA TG ISK PA RP AG INKRA	SCHTL SCHTL SCHTL SCHTF SCHTF SCHTF SCHTL SCHTF SCHTF	RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL RH SFATHLL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT ESG YDIRT ESG YDIRT	AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHQ	UVRITMITHU DVRITMIYHU DVRITMIYHU OVRITMIYHU OVRITMIYHU OVTITMIYHU DVTITMIYHU DVTITMIYHU DINITMIYHU DVSITMIYHU DVSITMIYHU DVSITMIYHU						
cell3 cel16 cel21 cel39 cel47 cel62 cel62 ce246	TETMLÖI DERILQI SRLVLQI SETVLQI HENGLQI DESGLQI DERILQI DERILQI	RAVKNAI	HR AG VQKRS HR AG VQKRS RQ AG LDKRQ RE AG INKRA RE AG INKRA RP AG LDKRQ EA AG LAKRV RA TG ISK PA RE AG INKRA	SCHTL SCHTL SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF	RH SFATHLL RH SFATHLL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ETG YDIRT EDG YDIRT EDG YDIRT ESG YDIRT ENG YDIRT ENG YDIRT	AQEL GHQ AQEL GHQ	UVRITMITHU UVRITMIVHU UVRITMIVHU UVRITMIVHU UVRITMIVHU UVRITMIVHU UVRITMIVHU UVITMIVHU UVITMIVHU UVITMIVHU UVITMIVHU UVITMIVHU UVITMIVHU UVITMIVHU						
cell3 cell6 cel21 cel39 cel47 cel47 cel62 ce246 ce243	TETMLÖI DERILQI SRLVLQI SETVLQI HENGLQI DESGLQI DERILQI DPQAVQI AATVLOI	KAVKNAI	HR AG VOKRS RQ AG LDKRG RE AG INKRA RE AG INKRA RKAK IAK AA RP AG LDKRG EA AG LAKRV RA TG ISK PA RE AG INKRA RT AG IAR PA DR AG VOKRG	SCHTL SCHTL SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF TPHTL SCHTF	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT NDS YDIRT	AQEL GHQ AQEL GHQ AQEL GHH AQEL GHH AQEL GHH AQEL GHH AQEL GHH AQEL GHH AQEL GHH AQEL GHH AQEL GHH AQEL GHH	UNIT MITTU UNIT MITTU						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce243 ce236	TETMLQI DERILQI SRLVLQI SETVLQI HENGLQI DESGLQI DERILQI DPQAVQI AATVLQI LDATFOI	KAVKINAI KAVKHAI KAVKHAI KAAKQAI KAAKQAI KAVKAAI KAVKAAI KAVKAAI KAVKAAK KAVKAAK KAVKAAI KAVKDAI HAIROAA	HR AG VOKRS HR AG VOKRS RQ AG LDKRG RE AG INKRA RE AG INKRA RK AK IAK AA RP AG LDKRG LA AG LARRV RA TG ISK PA RE AG INKRA RT AG IAR PA QR AG VOKRG	SCHTL SCHTL SCHTL SCHTF SCHTF SCHTF SCHTL SCHTF TFHTL SCHTF	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT ENG YDIRT NG YDIRT ADG YDIRT	AQEL GHQ AQEL GHQ AQEL GHQ AQEL GHH AQEL GHH	UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU UNTIMITAU						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce246 ce243 ce236 ce227	TETMLQI DERILQI SRLVLQI SETVLQI HENGLQI DESGLQI DERILQI D PQAVQI AATVLQI SETVLQI	KAVKNAI KAVKHAI KAVKHAI KAAKQAI KAVKAAI KAVKAAI KAVKAAI KAVKAAX KAVKAAX KAVKDAI KAVKDAI KAVKHLI	HR AG VORRS HR AG VORRS RE AG INKRA RE AG INKRA RE AG INKRA RE AG INKRA RE AG INKRA RI AG IAR DA RE AG INKRA RT AG IAR PA QR AG VORRG LA AR LINKRV LQ AG LAM 30	SCHTL SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF TFHTL SCHTF	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT ENG YDIRT DSG YDIRT ADG YDIRT XDG YDIRT ENG YDIRT	ACEL CHO ACEL CHO ACE	U VRITMITHU D VRITMIYHU D VRITMIYHU						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce243 ce236 ce227 ce226	TETMLQI DERILQI SRLVLQI SETVLQI HENGLQI DESGLQI DERILQI DERILQI DATVLQI LDATFQI SETVLQI QDQAVOI	2A VKN AI (A VKH AI) 2A A KQ AI) 2A A KQ AI) 2A VKA AI 2A VKA AI 2A VKA AI 2A VKA AI 2A A KQ AI 2A VKA AI 2A VKA AV) 2A VKA AV)	HR AG VORRS HR AG VORRS RQ AG LDKRG RE AG INKRA RK AK IAKAA RK AK IAKAA RP AG LDKRG EA AG LDKRG EA AG LDKRG EA AG LDKRG INKRA RT AG IAR PA QR AG VORRG RA AR LDKRA RK AN LDKRA	SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF TFHTL SCHTF SCHTF SCHTF	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT ESG YDIRT ESG YDIRT ADG YDIRT ESG YDIRT ESG YDIRT EGG YDIRT		UVRITMITHU- DVRITMITHU DVRITMITHU DVRITMITHU OVRITMITHU OVRITMITHU OVRITMITHU OVRITMITHU OVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce243 ce236 ce227 ce26 ce227	TETMLQI DERILQI SELVLQI SETVLQI SETVLQI DESGLQI DESGLQI DERILQI DEQAVQI AATVLQI SETVLQ QDQAVQI QDQAVQI	KAVKNAI (RVKHAI KAVKAAI KAVKAAI KAVKAAI KAVKAAI KAVKAAI KAVKAAI KAVKDAI KAVKDAI HAIRQAA KAVKHLI KAVKAAV KAVKAAV	HR AG VORES RQ AG LDKRG R2 AG LDKRG R2 AG LDKRG R2 AG INKRA R2 AG INKRA R4 AG LDKRG EA AG LDKRG R4 AG LDKRG R4 AG LDKRA R7 AG LDR PA R6 AR LDKRA LDKRU LQ AG LDM SG RK AN LVK PA	SCHTL SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF TPHTL SCHTF TPHTL SCHTF SCHTL SCHTL	RH SPATBLL RH SPATBLL	ADG YDIRT ADG YDIRT EDG YDIRT EDG YDIRT ETG YDIRT ETG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT ENG YDIRT ADG YDIRT ADG YDIRT EGG YDIRT EGG YDIRT		UNIT MITHU UNIT MITHU						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce246 ce236 ce236 ce227 ce26 ce227 ce26 ce224 ce221	TETMLÖİ DERILQİ DERILQİ SRLVLQİ SETVLQİ HENGLQİ DESGLQİ DERILQİ DEQAVQİ AATVLQİ QEQAVQİ QEQAVQİ AEGILQİ	AAVKNAI KAVKNAI KAVKAAI KAKQAI KAVKAAI KAVKAAI KAVKAAI KAVKAAI KAVKAAI KAVKAAI KAVKAI KAVKHII KAVKAAV KAVKAAV KAVKAAV KAKAAV	HR AG VORES HR AG VOR S RQ AG LDKRG RE AG INKRA RP AG LDKRG EA AG INKRA RP AG LDKRG EA AG LDKRG RA TG ISK PA RE AG INKRA RT AG IAR PA QR AG VORRG RA AR LDKR PA LQ AG LAMSG RK AN LVK PA RKEN LVK PA	SCHTL SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF TFHTL SCHTF TFHTL SCHTL SCHTL SCHTL SCHTL	RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT DSG YDIRT EDG YDIRT EDG YDIRT EGG YDIRT EGG YDIRT		U VRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce246 ce243 ce236 ce227 ce26 ce221 ce221 ce221	TETMLÖI DERILQI DERILQI SRLVLQI SRLVLQI SETVLQI DESGLQI DERILQI DEQAVQI AATVLQI QDQAVQI QDQAVQI ALGILQI LDATFQI	AAVKON AI KA VKCH AI KA NKQ AI KA VKQ AI KA VKQ AI KA VQK AA RI KA VQK AA RI KA VQK AA RI KA VKCH AI KA VKCH AI KA VKCH LI KA VKCH LI KA VKCH LI KA VKCH LI KA VKCH LI KA VKCH AV KA VKQ AX KA IRQ AX	HR AG VQERS RQ AG LDKRG RE AG IMKRA RE AG IMKRA RF AG IMKRA RF AG LDKRG RA AG LDKRG RA AG LDKRA RE AG IMKRA RE AG IMKRA RA AG IMKRA RA AL IMKRA RKAM LUK PA RKAM LUK PA RKAM LUK PA	SCHTL SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF TFHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL TFHTL	RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT DSG YDIRT ADG YDIRT EGG YDIRT SKG YDIRT SKG YDIRT		UVRITMITHU- UVRITMITHU- UVRITMITHU- UVRITMITHU- UVRITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVRITMITHU- UVRITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU-						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce243 ce226 ce227 ce26 ce227 ce26 ce221 ce221 ce220 ce214 ce214	TETMLQI DERILQI SERULQI SETVLQI HENGLQI DESGLQI DERILQI DATVLQI LDATFQI QDQAVQI QDQAVQI DEACLQI DEACLQI	KA VKO AI – KA VKO AI – KA VKO AI – KA VKA AI – KA VKA AI – KS VKO AI – KS VKO AI – KS VKO AI – KA VKA AI – KA VKO AI – KA VKO AI – KA VKO AI – KA VKO AI – KA VKO AI – KA VKO AI – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV –	HR AG VQERS RQ AG LDKRG R2 AG LDKRG R2 AG INKRA R2 AG INKRA R2 AG LDKRG RA AG LDKRG RA AG LDKRG RA TG ISK PA RC AG LDKRA RC AG LDKRA RC AG LDKRA RC AG LDKRA RC AL LNKRV HR AG ILK PA	SCHTL SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF TFHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTFHTL	RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL RH SPATHIL	ADG YDI RT' ADG YDI RT' EDG YDI RT' EDG YDI RT' EDG YDI RT' EDG YDI RT' EDG YDI RT' EDG YDI RT' EDG YDI RT' ENG YDI RT' ADG YDI RT' EDG YDI RT' EGG YDI RT' EGG YDI RT' EKG ADI RT' EKG ADI RT'		UVRITMITHU- DVRITMITHU- DVRITMITHU DVRITMITHU DVRITMITHU DVITMITHU DVITMITHU DVITMITHU DVITMITHU DVSITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU DVRITMITHU						
cell3 cell6 cel21 cel39 cel47 cel42 ce246 ce246 ce243 ce226 ce227 ce26 ce227 ce26 ce221 ce221 ce220 ce214 ce211 ce210 ce23	TETMLQI DERILQI DERILQI SETVLQI HENGLQI DESILQI DERILQI DERILQI DERILQI SETVLQI QDQAVQI QDQAVQI QDQAVQI LDATFQI DEACLQI DESSLQI	KA VKO AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKA AV – KA VKA AV – KA VKH AV – KA VKH LI – KA VKH LI – KA VKH AV –	HR AG VQFR B RQ AG LDKRG RQ AG LDKRG RE AG INKRA RP AG LDKRG RKAK IAR AA RP AG LDKRG EA AG LDKRG LA AG LDKRG AG LARRO LQ AG VQKRG RK AN LVK PA DE AG IVKRA RA AR LDKRV DE AG IVKRA RA AR LDKRV	SCHTL SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF TFFTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTF	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT ADG YDIRT ENG YDIRT EGG YDIRT SKG YDIRT SKG YDIRT EKG ADIRT EKG ADIRT		UVET MITHU- DVET MITHU- DVET MITHU DVET MITHU DVET MITHUL						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce246 ce243 ce236 ce227 ce26 ce227 ce26 ce221 ce220 ce224 ce221 ce220 ce224 ce221 ce220 ce224 ce220 ce224 ce221 ce220 ce224 ce220 ce224 ce220 ce224 ce220 ce224 ce220 ce224 ce220 ce224 ce220 ce224 ce221 ce220 ce224 ce221 ce226 ce224 ce221 ce226 ce224 ce227 ce226 ce224 ce226 ce224 ce227 ce226 ce224 ce226 ce227 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce227 ce226 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce227 ce226 ce226 ce226 ce226 ce226 ce227 ce226 ce226 ce226 ce226 ce226 ce227 ce226 ce	TETMLQI DERILQI SETVLQI SETVLQI HENGLQI DESGLQI DESGLQI DESCLQI DESCLQI QQQAVQI QDQAVQI QDQAVQI LDATFQI LDATFQI DESCLQI DESSLQI HENGLQI	KA VKO AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AV – KA VKH AV – KA VKH AV – KA VKH AK – KA VKH AK – KA VKH AK – KA VKH AK – KA VKH AK – KA VKH AK – KA VKH AK – KA VKH AK – KA VKH AK – KA VKH AK – KA VKH AK – KA VKH AK –	HR AG VQERS RQ AG LDKRG RE AG IMKRA RE AG IMKRA RE AG IMKRA RP AG LDKRG RA AG LDKRG RA AG LDKRA RA TG ISK PA RE AG IMKRA RA AG IMKRA RA AG IMKRA RA AR LMKRV HR AG IMKRA	300 TL 300 TL 300 TF 300 TF 300 TF 300 TF 300 TF 707 TL 300 TL	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT EDG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT DSG YDIRT EGG YDIRT EGG YDIRT EGG YDIRT EGG YDIRT EGG ADIRT ESG ADIRT ESG SDIRT		UVRITMITHU DVRITMITHU						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce246 ce246 ce223 ce227 ce226 ce227 ce226 ce221 ce220 ce214 ce220 ce214 ce220 ce214 ce220 ce214 ce223 99031763 Intl4	TETMLQI DERILQI SELVLQI SELVLQI SELVLQI DESGLQI DESGLQI DESGLQI DESGLQI SELVLQI QDQAVQI AATVLQI LDATFQI SELVLQI DEASLQI DESSLQI HENGLQI NETVLQI	KA VKO A.I KA VKH A.I KA VKH A.I KA VKH A.I KA VKA A.I KA VKA A.I KA VKA A.I KA VKA A.I KA VKA A.I KA VKA A.I KA VKA A.I KA VKA A.I KA VKA A.I KA VKA A.V I.A.IKA A.V I.A.IKA A.V I.A.IKA A.V KA VKA A.V - KA VKA A.V - - - - - - - - - - - - -	HR AG VORES RA AG LOKRO RE AG INKRA RE AG INKRA RE AG INKRA RE AG INKRA RE AG LOKRO RA AG LOKRO RA AG LOKRO RA AG INKRA RA TG ISK PA RA AG INKRA RKAN LUK PA RKAN LUK PA RKAN LUK PA RKAN LUK PA AG INKRA HR AG INK PA HR AG INK PA HR AG INK PA AG LOKR PA CAG AG AG AG A CAG AG A CAG A	300 TL 300 TL 300 TL 300 TF 300 TF 300 TF 300 TF 100 TF 100 TF 100 TF 100 TC 100 TL	RH SPATHIL RH SFATHIL	ADG YDIRT ADG YDIRT EDG YDIRT EDG YDIRT ENG YDIRT ETG YDIRT ETG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT DSG YDIRT EGG YDIRT SKG YDIRT ESG ADIRT ESG ADIRT ESG SDIRT ESG SDIRT		UVRIT MITHU- UVRITMITHU UVRITMITHU UVRITMITHU UVRITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU UVIITMITHU						
cell3 cell6 cel21 cel39 cel47 cel47 ce246 ce243 ce224 ce224 ce224 ce221 ce221 ce214 ce210 ce214 ce210 ce23 Intl4 AAQ166651 Intl1	TETMLQ DERILQ DERILQ SRLVLQ SRLVLQ SRLVLQ SETVLQ HENGLQ DESGLQ DESGLQ DESGLQ LDATFQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ LDATFQ DESCLQ HENGLQ HENGLQ YDQTFQ	KAUKO AL - KAUKO AL -	HR AG VQFR.B RQ AG LDKRG R2 AG INKRA R2 AG INKRA R2 AG INKRA R2 AG INKRA R2 AG INKRA R4 AG LDKRG EA AG LDKRG R4 AG LDKRG R4 AG LDKRA R4 AG LDKRA R4 AR LDKRA R4 AR LDKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R4 AG INKRA R5 AG	SCHTL SCHTL SCHTF SCHTF SCHTF SCHTF SCHT SCHT SCHT SCHT SCHT SCHT SCHT SCHT	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT EGG YDIRT EGG YDIRT EKG ADIRT EKG ADIRT EKG ADIRT EKG ADIRT EKG ADIRT EKG ADIRT EKG ADIRT EKG ADIRT		UVET MITHU- DVET MITHU- DVET MITHU DVET MITHU DVET MITHUL						
cell3 cell6 cel21 cel39 cel47 cel47 cel47 cel42 ce246 ce243 ce236 ce227 ce26 ce224 ce221 ce220 ce214 ce210 ce210 ce23 99031763 Intl4 AAQ166651 Intl1 AAT726911 Intl2	TETMLQ DERILQ DERILQ SRLVLQ SRLVLQ SETVLQ DESGLQ DERILQ DERILQ DERILQ DERILQ DERILQ SETVLQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ AXTVLQ QDQAVQ ALGILQ DESSLQ HENGLQ HENGLQ HENGLQ	KA VKO AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKA AI – KA VKA AI – KA VKA AI – KA VKA AI – KA VKA AI – KA VKH LI – KA VKH LI – KA VKH LI – KA VKH AV – KA VKH AV – KA VKA AV – KA VKA AV – KA VKA AV – KA VKA AV – KA KA KA AV – KA VKA AK – KA VKA AK – KA VKA AK – KA VKA AK – KA VKA AK – KA VKA AK – KA VKA AK – KA VKA AK – KA VKA AK – KA VKA AK – KA VKA AV – KA VKA VKA VKA VKA VKA VKA VKA VKA VKA V	HR AG VORES HR AG VORES RQ AG LDKRG RE AG INKRA RE AG INKRA RP AG LDKRG EA AG INKRA RP AG LDKRG EA AG LDKRG AG LDKRG AG LDKRA AG LAN SG RK AN LUK PA DE AG IVKRA RA AR LIMKRU LQ AG LAM SG RK AN LUK PA DE AG IVKRA RA AR LIMKRU VT VG LIKQM EA G - IEKTV VT VG LIKQM EA G - IEKTV	SCHTL SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF SCHTF SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTL SCHTF	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT DSG YDIRT DSG YDIRT EGG YDIRT EGG YDIRT EGG YDIRT ESG ADIRT ESG ADIRT EGG SDIRT EGG SDIRT CARBON		UVET MITHUL UVET MITHU UVET MITHU UVET MITHUL						
cell3 cell6 cel21 cel39 cel47 cel47 cel62 ce246 ce246 ce243 ce226 ce227 ce226 ce221 ce220 ce214 ce220 ce214 ce220 ce214 ce220 ce214 ce220 ce214 ce220 ce211 ce220 ce221 ce221 ce220 ce220 ce221 ce220	TETMLQ DERILQ DERILQ SRLVLQ SRLVLQ SRLVLQ DESGLQ DESGLQ DESGLQ DESGLQ DESCLQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ DESSLQ HENGLQ NETVLQ NE	KA VKO AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AI – KA VKH AJ – KA VKH AJ – KA VKH AJ – KA VKH AJ – KA VKH AJ – KA VKH AJ – KA VKH AJ – KA VKH AJ – KA VKH AJ – KA VKH AJ –	HR AG VORES RQ AG LDKRG RE AG IMKRA RE AG IMKRA RE AG IMKRA RF AG LDKRG RA AG LDKRG RA AG LDKRG RA AG LDKRA RE AG IMKRA RE AG IMKRA RA AG IMKRA YA AG ILM FA GINKRA AG ILM FA CALING RA AR LDMRU HR AG ILM FA CALING RA AR LDMRU HR AG ILM FA CALING RA AR LDMRU HR AG ILM FA CALING RA AR LDMRU HR AG ILM FA AG INKRA MA AG INKRA MA AG INKRA MA AG INKRA MA AG INKRA MA AG INKRA	300 TL 300 TL 300 TL 300 TF 300 TF 300 TF 300 TF 300 TF 300 TF 300 TF 300 TF 300 TF 300 TF 300 TL 300 TF 300 TL 300 TF 300 TL 300 TF 300 TL 300 TF 300 TL 300 TF 300 TC 300 TF	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT ADG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EGG YDIRT EGG YDIRT EGG YDIRT EGG YDIRT EGG SDIRT EGG SDIRT EGG SDIRT EGG SDIRT CAG SDIRT CAG SDIRT CAG SDIRT		UVRITMITHU- UVRITMITHU- UVRITMITHU- UVRITMITHU- UVRITMITHU- UVRITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU- UVIITMITHU-						
cell3 cell6 cel21 cel39 cel47 cel62 ce246 ce246 ce246 ce223 ce227 ce226 ce227 ce221 ce220 ce214 ce220 ce214 ce210 ce23 99031763 IntI4 AAQ166651 IntI1 AAT728911 IntI2 AAQ323551 IntI3 POA8P61 XERC POA8P61 XERC	TETMLQ DERILQ DERILQ SRLVLQ SRLVLQ SRLVLQ SETVLQ DESGLQ DESGLQ DERILQ DESGLQ DEQAVQ AXTVLQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ DESCLQ DESSLQ HESSLQ HESSLQ HESSLQ HESSLQ SETVLQ SRLVQ	KA VKO AI – KA VKO AI – KA VKI AI KA AI – KA VKI AI – KA VKI AI KA AI – KA VKI AI – KA V	HR AG VORES RA AG VORES RQ AG LDKRG RE AG INKRA RE AG INKRA RE AG INKRA RE AG INKRA RE AG LDKRO RA AG LDKRO RA AG LDKRO RA AG INKRU LQ AG INKRU LQ AG LDM SG RKAN LUK PA RKAN LUK PA RKAN LUK PA RKAN LUK PA RKAN LUK PA RKAN LUK PA RKAN LUK PA RKAN LUK PA RKAN LUK PA RA AR INKRU HR AG I IK PA YA AG INKRU VI VG LIK QM EA G I VSKRU VO QAG IAKHU KQ GL NARHU	SC T11 SC T21 SC T2 SC T	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT ADG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EGG YDIRT EGG YDIRT EGG YDIRT EGG YDIRT EGG SDIRT EGG SDIRT EGG SDIRT EGG SDIRT EGG SDIRT EGG SDIRT EGG SDIRT		UVRITMITHU- UVRITMITHU- UVRITMITHU UVRITMITHU UVRITMITHU UVITMITHU						
cell3 cell6 cel21 cel39 cel47 cel42 ce246 ce243 ce224 ce224 ce221 ce220 ce224 ce221 ce220 ce214 ce210 ce23 90031763 Int14 AAQ166651 Int112 AAQ323551 Int13 P0A8P61 XERC P0A8P61 XERD	TETMLQ DERILQ DERILQ SRLVLQ SRLVLQ SRLVLQ SETVLQ HENGLQ DESGLQ DESGLQ DESGLQ LDATFQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ QDQAVQ DESCLQ DESSLQ HENGLQ HENGLQ TESSLQ HENGLQ TRQTFW	KAVKONAI – KAVKONAI – KAVKONAI – KAVKONAI – KAVKONAI – KAVKONAI – KAVKANI – KAVKA	HR AG VQFR.B RQ AG LDKRG R2 AG INKRA R2 AG INKRA R2 AG INKRA R2 AG INKRA R2 AG INKRA R4 AG INKRA R4 AG LDKRC R4 AG LDKRC R4 AG LMRA R4 AR LNKRV LQ AG LMNG R4 AR LNKRV LQ AG INKRA R4 AR LNKRV HR AG IIK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA YA AG INK PA	SC TL SC TL SC TL SC TL SC TF SC TL SC TF SC TL SC TL TF TL SC TL TF TL SC TL TF TL TF TL TF TL TF TL TF TL TC TF TC TL TC TL TC TL TC TL TC TL SC TL	RH SPATHIL RH SPATHIL	ADG YDIRT ADG YDIRT ADG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT EDG YDIRT EDG YDIRT EDG YDIRT ENG YDIRT DSG YDIRT ENG YDIRT EGG YDIRT EGG YDIRT EKG ADDRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT RSG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT ENG YDIRT RSG YDIRT ENG YDIRT		UVET MITHU- UVET MITHU UVET MITHU UVET MITHU UVET MITHUL						

Figure 1.12 partial amino acid sequence alignment of environmental integrases from Culebrones Cave against other tyrosine recombinases, IntI class 1 from *Escherichia coli*, IntI 2 from *Shigella sonnei*, IntI 3 from *Klebsiella pneumoniae*, IntI 4 from *Vibrio cholerae* and tyrosine recombinases XerC and XerD from *E. coli*. The position of conserved motifs among integron-encoded integrases is indicated by segments labeled as Boxes I and II and Patches II and III and as additional domains (AD) under the alignment. The degree of conserved amino acids (100% identical) with functions related to protein folding, DNA binding, and recombination activity are high-lighted in red, respectively. Green columns correspond to H, R, and Y residues from the conserved RHRY tetrad, which is characteristic of the entire tyrosine recombinase family (Nunes-Dubt, S., et al. 1998). GenBank accession numbers for the protein sequences of reference integrases (classes 1 to 4) and tyrosine recombinases XerC and XerD are AAQ16665, AAT72891, AAO32355, 99031763, P0A8P6, and P0A8P8, respectively.

	150					1.00							
	150	10	50 1	70	180	190	200	2.	10 2	20	2 30	2	140
	PHENO I PHENO CO	OTTERDO	TOWNER LETALT	DURATE	ODWING	TRROWOODT	CD CDC CUVI				CD THEFT	DRC FWDT	
can28	PUTT DEAC	OTTUPDO	TOWN TTMI	DADALTE	ODITIVE	UPPOROODI	CD_CPCSUVIL		PRANE DALW	QIV PA	CP TUND	ROG EMPS	
can48	DICTO UDEA EN	TTTWDC	TOUTO SUDI	DVDIVEDI	VY COL	URDITDODI	UNDOVCCUP		PRARKDALW		GRIVED	DECIVER	
can49	PUVDUD PACE	OTTUPDO	TOWNS TTMT	DIGITT	ODITUS	VERCHOODI	CD CDCCVCM	PRO LOK R	ARE ARDIQU	Quil PA	ARE LI LV	DEGLIKS	
can44	RUND VDENVE	TTIPDC	CARDA TIME	PASALILI	COD LLV S	VRRYBYYDL	GR-GFGSVYL	PFALAR	PNANRDWIW	QWV PA	GRIVED	RSGEMRE	
Fr4	RVRDVDFARS	CINKDG	CHARDE VINL	PUSLVPA	2Q1QLLH	ARSLAADL	AQ-GLGEVWL	PHALIGVE.	PNAGKEWGW	QTI PA	AGLSID	REGERRE	
Fr6	RURDLDEDNI	QLQV666	GRAN LITL	AILLVPL	LRIQIRK	VEFLLEQUN	REKDYAGVWM	PUALAK A	GRAAKLLNW	QYL PA	SRVSVE	GISLLR	CHART
Fr7	RVRDLDEGI	Q I I VICE G	GRUE TIML	PVSLVEL	CREQIET	VQ DQ HKRDL	AA-GYGEVWL	PHALEVE	PIASREPAW	QYV PA	YRRSKD	RSAKER	- V
FT9	RVRDVDESME	CITTE	CONRUGATIVE	PETEVEL	LORGVER	VKDLHEDDL	AN-GREAVIL	PNALAR .	PNAEVELAW	QIV PA	OT ISAD	RSGQLRB	. nov
Ft10	RVRDVDFARG	Q II VKL G	ORDEVINL	PERLESE	LOGPLLR	VRLLHDRDL	TE-GEGRVYL	PRALAK .	PNADRQWPW	QWV P3	ARKSSD	RSKALK	(HHV
Fr61	RVRDLDFLR	QIRVIRG	GRADETIVM	DRALVIP	LENHLLE	MKRLFDEDV	LN-GFGELHA	PPALARE	PNAELRWGD	Y YV AL	DISLTLE	GMRVASE	REL
F=79	RVRD VDEDQS	LVYVRGG	GDKDRTSLL	SPRLVNP	LKQHLEK	VKALHEQDL	KQ-GHGEVYL	PGALSRE	PNAGY EWAW	QYV PA	GRLAVD	RSGVVR	CHHV
Ft82	RVRDLDFERE	EITLENT	GKRDRRIML	AANVVEPI	LKLHLNV	VRAMWEQDL	AR-GYGTVQL	PNALARE	PNA PREWGW	Q YV PA	SRISTD	QTG FR GE	WHIL
FT93	RVKDLDFERM	QIRVYRG	GRACER ITVM	DRALVTPI	LENHLLE	MRKLFDEDD	LN-GFGEVYF	PQALERH	KNSGT DWR W	Q YV PA	KGRSVD	RSGRTRE	CHHV
p53	RVRDFDFASI	RILVSSG	CREAT TANK	PESALTE.	IDYERT	VERLHITER	GR-GFGSVYL	PFALARE	PNANR DWI W	Q YV PA	GQIRRD	RSGEMRE	THHI
p 57	KVRDVDLI AS	QITVRDG	GRENEITML	PASALTE	LODLLVS	VRRQHHQDV	GE-GFGSMYL	PALVIRE	PNANR DWV W	SAALby	PNICKD	RCGEVRE	CHHL
p 60	RVRDVDEAQE	QIIVRDG	GRENEVIML	PVSAATE.	IQDHLVG	IRRQHQQDL	GR-GFGSVYL	PFALEQ	PNANR DWIW	Q YV PA	GRIVKD	RSG EMRF	(HHL
p 99	RVRDVDEAQE	QITVRDG	GRENEVIML	PVSAATE.	IQDHIVG	IRRQHQQDL	GR-GFGSVYL	PFALEQ	PNANR DWIW	Q YV PA	GRIVKD	RSGEMRE	HHL
p23	RVRDVDEAQE	QITVRDG	GREARISML	PASALTE	LQDLLVS	VRRQHQQDL	GR-GFGSVYL	PFALAR	PNANR DWIW	QWV PA	GRIVKD	RSG EMRR	HHL
p25	RVKDVDFENE	LIEVHOS	GKSRLVPM	PENLVKP	LKRWVAS	RKVLHRHDV	AD-GTASVYL	PFAIQRE	PNAHREFRW	QYL AS	PRFSRD	KTRRLHE	HHI
p31	KVRD VDE AQI	QTIVIDG	CREAR TIML	PASALILI	LODLLVS	VKRQHQQDL	GR-GFGSVYL	PFALARE	PNANRDWIW	QWV PA	GRIVED	RSGEMRE	HHL
p75	KVRD VDEAQE	QTIARDA	SESEVIML	PDILVAP	LORHLER	VKKLHQLDL	DQ-GYGSVYL	PFALERE	PHADR SWPW	QEV PS	MNRSRD	RSGIVRE	CHHL
p76	RVRDVDEAQE	QTIVRDG	GRENEVIML	PVSAATE.	LODHLAG	TKKQHQQDL	GR-GFGSVYL	PFALEQ	PNANR DWIW	Q YV PA	GRIVKD	RSGEMRE	(HHL
p33	RVKDLDEAQE	QITVRDG	GREAR ITML	PASALTE	LQDLLVS	VRRQHQQDL	GR-GFGSVYL	PFALAR	PNANR DWIW	QWV PA	GRIVKD	RSGEMRE	HHL
p42	RVEDIDEAGE	QIIVRDG	GRENEVIML	PVSAATE.	IODHLVG	IRRQHQQDL	GR-GFGSVYL	PFALEQ	PNANR DWIW	Q YV PA	GREVKD	RSG EMRP	HHL
pb	RVRDVDEAQE	QIIVRDG	GRENEVIML	PVSAAAE.	IQDHLVG	IRRQHLQDL	GR-GFGSVYL	PFALEQ	PNANR DWIG	Q YV PA	GRIVKD	RSG EMRP	(HHL
\$33	RVKDVDEDYE	QLWVRSG	GORDEVSLL	PORCIPE	KDQIER	VALLHQKDL	AG-GYGKALL	PKALSKK	PDEDKELRW	QYL PS	KI IYED	RSGLKHE	YHL
\$ 34	RVKDLDEDYE	QLWVRLG	CORDEVELL	PORCIPE	KDQIER	VALLHQKDL	AG-GYGKALL	PKALSKK	PDEDKELRG	QYL PS	KIIYED	RSGLKHR	(YHL
\$73	R VKD FD FD YE	QLWVRSG	GQKDRVSLL	POKCIPE	VKDQIEK	VALLHQKDL	AG-GYGKALL	PKALSKK	PDEDKELRW	QYL PS	KIISED	RSGLKHR	(YHL
\$79	R VKD FD FD YE	QLWVRSG	COKDRUSLL	POKCIPE	KDQIEK	VALLHQKDL	AG-GYGKALL	PKALSK	PDE DKELRW	QYL PP3	KI IYED	RSGLKHE	(YHL
\$92	RVKDLDENHN	AITVHNG	GKSRSTLL	PNQLAGE	LHSLCER	AADIQQGDN	QK-GIG-PSM	PYALDR	PNA YR KTGW	MFI PS	VN YC HH	VINKL CR	(HHL
\$125	RVKDLDEDYE	QLWVRSG	GQKDRVSLL	POKCIPE	VKDQIEK	VALLHQKDL	AG-GYGKALL	PKALSK	PDE DKELRW	QYL PS	KIISED	RSGLKHE	(YHL
99031763 Intl4	RVQDIDEDYG	AIRIWQG	GKNRTVTL	AKELYPH	LKEQIAL	AKRYYDRDL	H QRN YG GV WL	PTALKE K	PNA PYEFRW	HYLPPS	FQLSLD	ESDVMRR	(HHM
AAQ100051 Intil	RVHDLDFDHO	TIIVREG	CSKDR ALML	PESLAPS	LREQLSR	AR AWWL KDQ	AE-GRSGVAL	PDALER	PRAGH SWPW	FWV AQ	HTHSTO	RSGVVRF	(HHM
AAT728911 Int12	R VKD FD FD NG	CITVHDG	GGKSRNSLL	PTRLIPA	IKXLIEQ	ARLIQQDDN	LQ-GVG-PSL	PFALDHK	PSAYRQAAW	MEV	STLCNH	YNG KL CF	(HHL
AAU323551 Intl3	RVKDVDEDRI	AIIVRSG	GORDRVVML	PRALVPRI	LRAQLIQ	VR AV WG QDR	AT-GRGGVYL	PHALER	PRAGE SWAW	FWV PS	AKLSVD	QTG VE RE	HHL
PUASPEI XERC	DINHLDLESG	E VWVMG -	GSKER RLPI	GRNAVAW	TEHWLDL	RDLFGS	ED			DAL	K	LG	KRI
PUROPOI KERD	TMSDISLRQO	VVRVIG-	GNKERLVPL	GEEAVYW	LETYLEH	GR PWLLNGV	SI			DVL	Q	RA	LOOM
	Day I	→	Dotch !!	-▶			م ∧ ما ما :+!	al Darr		-	▶ ← ,	AD 🔶	
	BOXI		Patch II				Audition	ai Dom	ain (AD) 🗕	Patrl	h III		
										ratu			

	250	260	270	280	290	300
			····I ····I		I I	
can28	HESGIQKALKQ.	AVRAVGIA-KKV	GCHTERH SEA	THLLENGYD	RT IQELL	HKDVKTTMIYTHVL
can48	HESG IQKAVKQ	AMRQTGIA-KKV	GCH TERH SEA	THLLEDGYD	RT IQELL	HKDVKTTMIYTHVL
can49	YRONLSLE IRR.	AT RRAKV P-KRV	TAN TERNSEA	SHLLAANYD	HT IQKLL	HSDVKTTMIYTHVL
can44	HESG IQKAVKQ	AMRQTGIA-KKV	GCHTFRHGFA	THLLEDGYD	RT IQELL	HKDVKTTMIYTHVL
Fr4	DEKQVQRGMKR.	AA AWV GMV-KPV	SPHTLRHSFA	THLLEGG YD	RTVQELL	HKDVSTTMIYTHVL
Fr6	DETGLNRV IRH.	AARRAQIQ-KDV	SCHTLRH SFA	THLLQAGTD	RTVQQQL	HTDVKTTMIYTHV-
Fr7	DESL PORAVKA	AI KLAKVR-KAG	TCH SFRH SFA	THLLEAGYD	RTVQELL	HKDVSTTQIYTHVL
Fr9	HVNT IQRAVKA	AVRKANVD-KNA	SCHITFRHSFA	TRMFESGYD	RT IQELL	HADIATTMIYTHVL
Fr10	HELPVQRAMKE.	AVLLARLS-KPA	TCHTLRH SFA	THLLEAGYD	RTVQELL	HODVATTOIYTHVL
Fr61	HRNR FSRVLTE.	AAERAQIT- <mark>K</mark> NV	TPHVLRHAFA	THML EMG TD	RKIQKLL	HAFVS TTMIYTHVL
Fr79	SIMVLORAMKE.	AMRRAGIA-KHA	SVHTLRHSFA	THLLMNGVN	REVQEYL	HENVE TTOIYTHVL
Ft82	YHTG IQRAVKR.	AAEAAGLA-KRV	TPH TL RH SFA	THLLEAGCD	RTVQELL	HKDVKTTŇIYTHVL
Fr93	D PSP IN KAMNN	AVREAGIQ-KSV	SGHTLRH SFA	THFLQPGTD	RT IQYLL	HEDVOTTMIYTHVL
p 53	HPASCRSIMER.	AVNEAGIS-KKV	SMH SL RH SFA	THLLEGGTD	RFIQKLL	HSS IKTT AIYTHVL
p 57	HDEN LERS FKI	FVAAPQEQ-RGV	GCHVLRN SFA	THFLEDGYD	RT IQELL	HKDVKTTMIYTHVL
p 60	HESG IQKALKQ	AVRAVGIA-KKV	GCH TFRH SFA	THLLENG YD	RT IQELL	HKDVKTTMIYTHVL
p 99	HESGIQKALKQ.	AVRAVGI A- <mark>K</mark> KI	GCHTFRHSFA	THLLENG YD	RT IQELL	HKDVKTTMIYTHVL
p23	HESG IQKAVKQ	AMRQTGIA-KKV	GCHTERH SFA	THLLEDGYD	RT IQEVL	HEDVETTMIYTHVL
p25	HET-FSSHLRR.	AVETTGI IAKHE	TSHTFRHSFA	THLIWSGTD	RT IQELL	HPD IKTTMIYTHVL
p 31	HESG IQKAVKQ	AMRQTGIA-KKV	GCHTERH SEA	THLLEDGYD	RT IQELL	HKDVKTTMIYTHVL
p75	HEGSVQKSLKL.	AVRMARIT-KRV	GCHTLRH SFA	THLLEGGYD	RTVQELL	HKDVKTTMIYTHVL
p76	HESG IQKALKQ	AVRAVGIA- <mark>K</mark> KV	GCHTFRHSFA	THLLENG YD	RMIQELL	HKDVKTTMIYTHVL
p 33	HESGIQKAVKQ.	AMRQTGIA-KKV	GCHTERH SFA	THLLEDGYD	RT IQELL	HKDVKTTMIYTHVL
p42	HESG IQKALKQ	AVRAVGIA-KKV	GCHTERN SFA	THLLENG YD	RT IQELL	HKDVKTTMIYTHVL
p5	HESG IQKALKQ	AVRAVGIA- <mark>K</mark> KV	GCHTERH SFA	THLLENGYD	RT IQELL	HKDVKTTMIYTHVL
\$ 33	SNRYVQKR IKK	AAKKAGLN-KKI	TCHTLRH SFA	THLLQNGYD	RTVQELL	HKSLKTTMIYTHVL
s 34	SNRYVQKRLKK.	AAKKAGLN- <mark>K</mark> KI	TCHTLRH SFA	THLLQNLLR	-TIGRIG	HKSLKTTMIYTHVL
s73	SNRYVQKR LKK	AAKKAGLN-KKI	TCHTLRH SFA	THLLQNGYD	RTVQELL	HKSLKTTMIYTHVL
\$79	SNRYVQKR IKK	AAKKAGLN- <mark>K</mark> KI	TCHTLRH SFA	THLLQNGYD	RTVQELL	HKSLKTTMIYTHVL
s 92	HDSNVRKALNR.	AVQRAGI TH <mark>K</mark> HV	TCH TFRH SFA	TQML LAG YD	RT IQTLL	HTDLS TTOIYTHVL
\$125	SNRYVQKRLKK	AAKKAGLN-KKI	TCHTLRHSFA	THLLQNGYD	RTVQELL	HKSLKTTMIYTHVL
99031763 Int	14 NETVLQKAVRR	SAQEAGIE-KTV	TCHTLRH SFA	THLLEVGAD	RTVQEQL	HTDVKTTOIYTHVL
AAQ166651 Int	11 YDOT FORAFKR.	AVEQAGIT-KPA	TPHTL RH SFA	TALLRSGYD	RTVODLL	HSDVS TTMIYTHVL
AAT728911 Int	12 HDSVARKALKA	AVQKAGIVSKRV	TCHTFRHSFA	THLLQAGRD	RTVQELL	HNDVKTTOIYTHVL
AA0323551 Int	13 FEERLNRQLKK	AVVQAGIA-KHV	SVHTLRH <mark>SFA</mark>	THLLQAGTD	RTVQELL	HSDVSTTMITTHUL
POA8P61 XERC	SARNVOKR FAE	NG IKQGLNN-HV	HPH KL RH SFA	THMLESSGD	RGVQELL	HANLSTTOIYTH-L
POASPS1 KERD	TR FWHR IKH	YAVLAGI DSE KL	SPHULRHAFA	THLLNHG AD	RVVQMLL	HSDLSTTOIYTHVA
			◀	Bo	x II 👘 📖	

FFigure 1.13 Partial amino acid sequence alignment of environmental integrases from the Solar Saltern Ponds against other tyrosine recombinases, IntI class 1 from *Escherichia coli*, IntI 2 from *Shigella sonnei*, IntI 3 from *Klebsiella pneumoniae*, IntI 4 from *Vibrio cholerae* and tyrosine recombinases XerC and XerD from *E. coli*. The position of conserved motifs among integron-encoded integrases is indicated by segments labeled as Boxes I and II and Patches II and III and as additional domains (AD) under the alignment. The degree of conservation (percent identity) among equivalent residues is indicated as follows: blue (100%), yellow (95% and 85%). Conserved amino acids (100% identical) with functions related to protein folding, DNA binding, and recombination activity are highlighted in red, respectively. Green columns correspond to H, R, and Y residues from the conserved RHRY tetrad, which is characteristic of the entire tyrosine recombinase family (Nunes-Du⁻bt, S., et al. 1998). GenBank accession numbers for the protein sequences of reference integrases (classes 1 to 4) and tyrosine recombinases XerC and XerD are AAQ16665, AAT72891, AAO32355, 99031763, POA8P6, and POA8P8, respectively.



Figure 1.14 Partial amino acid sequence alignment of environmental integrases from the Coamo Thermal Springs against other tyrosine recombinases, IntI class 1 from *Escherichia coli*, IntI 2 from *Shigella sonnei*, IntI 3 from *Klebsiella pneumoniae*, IntI 4 from *Vibrio cholerae* and tyrosine recombinases XerC and XerD from *E. coli*. The position of conserved motifs among integron-encoded integrases is indicated by segments labeled as Boxes I and II and Patches II and III and as additional domains (AD) under the alignment. The degree of conservation (percent identity) among equivalent residues is indicated as follows: blue (100%), yellow (95% and 85%). Conserved amino acids (100% identical) with functions related to protein folding, DNA binding, and recombination activity are highlighted in red, respectively. Green columns correspond to H, R, and Y residues from the conserved RHRY tetrad, which is characteristic of the entire tyrosine recombinase family (Nunes-Dubt, S., et al. 1998). GenBank accession numbers for the protein sequences of reference integrases (classes 1 to 4) and tyrosine recombinases XerC and XerD are AAQ16665, AAT72891, AAO32355, 99031763, POA8P6, and POA8P8, respectively.

Phylogenetic analysis from integron integrases from environments with moderate anthropogenic impact from Puerto Rico:

Thirty integron integrases were found in soil from the two different places of Culebrones Cave in Arecibo. Additionally, twenty eight integron integrases were recovered in the screening of four places from the Solar Saltern Ponds in Cabo Rojo (Candelaria and Fraternidad Lagoon, "puente" and saltern ponds). Also, eight integron-encoded integrases were recovered from the Coamo thermal spring. The primers set used to amplify approximately 491bp of integron integrases were hep35 and hep36 complimentary to nucleotide sequences founded in the Box I and Box II domains of integron integrases (Table 2.1). All intI-like clones identified as integron integrases recovered from Culebrones Cave, the Solar Saltern Ponds, and the Coamo Thermal Spring created clusters in which the tyrosine recombinases represented by XerC and XerD (P0A8P6, and P0A8P8) were excluded. Most of the environmental integrases recovered from different environments with moderate anthropogenic impacts were different from the other previously described. The sequences corresponding to the clone s92 were the only genotype related to the integrase class2 (AAT28911) described in pathogenic strain Shigella sonneii (AAT728911) (Figure 1.15, cluster I, 66% identity). Clinical integrases class1 (AAQ166651) and 3 (AAO323551) formed one cluster where the most similar clones were from Coamo Thermal Springs (Figure 1.15, cluster 10) with 46 and 98% identity respectively. Related to IntI4 (99031763) is the clone Fr6 (Figure 1.15, cluster II, 55% identity) which was collected from the Cabo Rojo Solar Saltern Ponds.

In the dendogram constructed, one cluster exclusively consisted of sequences recovered from "Puente" and Candelaria (Figure 1.15, cluster 1, 63%-88% identity) and very distant from the clinical integron integrases. However, the sequences are similar to the integrase of uncultured

33

bacterium (ABD625711) whose origin is Mona Island's dry forest soil. Nevertheless, discrete cluster are observed from solar ponds (Figure 1.15, cluster 5, 98%-63% identity) related to integron integrase of manure soil origin, while sequences recovered from Fraternidad are disperse without any cluster formed. Discrete cluster consisting of sequences from Coamo Thermal Springs were detected (Figure 1.15). One cluster is observed to be related to integron class 1 and class 3, meanwhile other clusters are distant from the clinical integron integrases (Figure 1.15, clusters 10 and 2). Culebrones Cave sequences showed different discrete clusters (Figure 1.15, clusters 3, 4, 6 and 7) whose sequences are distant to the clinical integron integrases with identity percentage of 98% up to 66% and all related to integron integrase of uncultured bacterium from dry forest soil from Mona Island and peat soil.

The following sequences recovered from the Culebrones Cave and the Solar Saltern Ponds are nearest to an Int of cultured bacteria. Clone ce139 was related to integron integrases described as a representative of the reductive sulfur cycle, *Desulfurivibrio alkaliphilus* (YP00336917161) having a 71% *IntI* gene sequence similarity (Figure 1.15, cluster 4). Can49 related to Int described in *Verrucomicrobiae bacterium* (ZPO50599471) having 53% *IntI* gene sequence similarity (Figure 1.17). Sequences recovered from Culebrones Cave and Fraternidad Laggon were related to integron integrase of Gram negative and the rod shape opportunistic pathogen, *Pseudomonas stutzeri* (ACB468491), with identity percentage from 59 and 66% (Fig. 1.15, cluster 9).



Figure 1.15 Consensus neighbor-joining tree illustrating the phylogenetic relationship of IntI sequences recovered from different extreme environments in Puerto Rico. Values represent the percentage of 1,000 boostrap replications that supported the branch. Outgroup references were designate with a thicker black branches highlight XerC, XerD, class IntI 1, IntI 2, IntI 3, and IntI4. Red triangle are sequences recovery from Solar Saltern Ponds, blue triangle circle from Coamo Thermal Springs and green circle are recovered sequences from Cueva Culebrones. Numbers 1 to 10 denote the clusters formed by recovery sequences and the percent identities of IntI sequences associated are shown in parentheses. Roman numbers (I and II) are recovery sequences related to clinical integron. Values of boostrap support below 49% are not demonstrated and Genbank accession numbers follow the outgroup sequences.

Discussion

It is important to gain information about the negative and positive effects of horizontal gene transfer and understating the role of the integron in these environments. This study represents the first attempt to determine the distribution and presence of the integrase of integron from environments with moderate anthropogenic impact and extreme conditions in Puerto Rico and the Caribbean. Validation of DNA quality was performed for all environmental samples, using specific primers for amplified the 16S rDNA. Optimizations were performed during IntI PCR, where annealing temperature range were 50-59°. We selected 52°C as the annealing where an intense band was obtained and with less accessory bands. In previous work the annealing of 52° was selected to chose environment integron integrases not related to clinical settings (Rodríguez-Minguela et al., 2009); on the contrary, high annealing did not generate specific bands as the annealing of 56°, instead bands associated to integron integrases class 1 were generated. Although, the degenerate primers used amplify approximately 491bp of integrase integron genes, they were constructed using sequences of IntI class 1, 2 and 3 (White et al., 2000), which allowed amplified environments integron integrase not associated to clinical settings. Meanwhile the primers used by Nield et al., (2001) were primers which allowed to amplify integron platforms directly from the environment.

The molecular screening of integron integrases from environments with moderate anthropogenic impact in Puerto Rico demonstrate 68 novel integron integrases (thirty from Culebrones Cave, twenty eight from Solar Saltern Ponds and eight from Coamo Thermal Springs). The recovery of integron integrases from environments with moderate anthropogenic impact demonstrates that this mechanism not is restricted to clinical environments. The integron integrases found shared domain and motifs features of integron integrases. Moreover, the high conservation of important motifs related to recombination mechanisms, stabilization, and binding activity show those probably are functional integron integrases. The low identities value (Fig.1.15) found in the recovery integron integrases with other IntI previously described indicate the diversity of integron integrase.

All sequences were aligned with previously described tyrosine integrases and functional integron integrases. The comparative phylogy and alignment performed demonstrated the integron integrase exclude other members of the tyrosine recombinases family, as are the XerD and XerC with a range of identity between 44–66%. XerD and XerC were used as reference for the identification of additional domain regions as principal feature of the integron integrases. Nevertheless, making inferences about the origins of environmental sequences recovered and their relationship with Xer are difficult because their evolution and origins are poorly understood. Clusters were observed for each environment when a consensus tree was performed for all environmental sequences. Sequences similar to integrase class 1, 2, and 3 were rare with exception Coamo Thermal Springs where these are closer to this class of integron integrase (46 and 98% identity).

The most distant to clinical integron integrase were found in majority in the Culebrones Cave, puente and salterns pond (Solar Saltern Pond) int–like. Fraternidad clones were very different and not forming any clusters within them. Only one int–like clone from Fraternidad Fr6 was the most closely related to IntI4 with 55% identity. Little sequences were closely related to integrase integron from cultured bacterium, an important reason to continue understanding the importance of the culture dependent technique at the moment of performing an analysis. Nine environment recovery sequences were related to culture bacterium: ce23 and ce139 related to *Desulfurivibrio alkaliphilus* (YP00336917161), ce214 and Fr10 related to *Nitrococcus mobilis* (ZP011289111); ca1233, ca1229 and ca229 to integrase of Gram negative and rod shape oppor-

tunistic pathogen, *Pseudomonas stutzeri* (ACB468491) and s92 to integron integrase class 2 from *Shigella sonnei*. and can49 related to Int described in *Verrucomicrobiae bacterium* (ZPO50599471).

In this study, thirty integron integrases were discovered were of in soil from four different places in the Culebrones Cave in Arecibo; twenty eight from Solar Saltern Ponds in Cabo Rojo (Candelaria and Fraternidad Lagoon, Puente and Saltern Ponds) and eight from molecular screening performed in Coamo Thermal Springs. Nemergut, D.R., et al. 2004 and Nield et al. 2001. Used a similar molecular approach and identified 14 and 3 new integrase genes from diverse environments.

The present work demonstrates that integron integrases not is restricted to clinical setting; in addition, the integron integrases may be a common feature of bacterial from caves, halophilic and thermophilic environments. Besides, the presence of novel intI genes implies the occurence of uncharacterized integrons and strengthens the importance of exploring these environments. In view of high diversity of integron integrase from different environment samples, it can create an extensive scenario of genetic functions where the prokaryotic cell can adapt quickly to environmental changes.

Part II. Sampling Procedure, Isolation, Characterization and *Int*-PCR of Isolates

Strains

I. Sample Processing and Isolation of Halophilic Bacteria:

At the Solar Saltern Ponds, eight samples containing 50 ml of saline water were collected (in duplicate) from the four different sampling points (Fraternidad and Candelaria Lagoons, Puente and solar ponds). The sampling was performed using Whirl Pack® bags and then samples were processed using the membrane filter technique. Saline water samples were homogenized with a pipette following the filtration of a volume of 25 ml in 47 mm, 0.45 µm (pore size) cellulose nitrate membrane filters (Whatman, Germany). Filtered membranes were inoculated into Luria-Bertani (LB) with different NaCl concentrations (0, 1, 5, and 8 %) at pH 7.5 and adjusted with 1M NaOH. All plates were incubated at 37°C. After three days of incubation colonies were selected and purified by the quadrant streak.

II. Alkalithermophilic bacteria culture media:

The *Thermus* medium (TM) was used for alkalithermophilic bacteria reactivation. TM consists of (in g/L) 8.0 peptone; 4.0 yeast extract and 2.0 NaCl, pH to 7.5 adjusted with 1M NaOH. The strains were incubated at 70° C.

III. Morphological and cultural characterization of halophilic strains:

Gram stains were performed using both heat-fixed smears and smears fixed in acetic acid 5% (Dussault, 1955). Macroscopic characteristics were documented using the classical characterization of colony appearance. Electron microscopy procedures were performed as previously described (Díaz-Muñoz and Montalvo-Rodríguez, 2005). The samples were examined using JEOL JSM-541 OL SEM microscope at 15 kv.

IV. Cryopreservation of halophilic strains:

Isolated strains were grown in 1.5 mL microtubes containing 930 μ L of Luria Bertani at different NaCl concentrations (5 and 8 % w/v) and when cells were at logarithmic phase, 30 μ L Dimethyl Sulfoxide (DMSO) were added and the sample was quickly frozen at - 80°C. The culture was stored at - 80°C.

V. DNA extraction

Isolated strains were grown in Luria Bertani at different NaCl concentrations (1, 5 and 8 % w/v) and incubated at 37 °C. DNA was extracted using the bead beating technique suspended in lysis buffer (40 mM Tris-Acetate pH 7.8-8.0, 20 mM Sodium acetate pH 8.0, 1.0 mM EDTA pH 8.0, and 1% SDS) with lysozyme followed by chloroform extraction and ethanol precipitation. The DNA quality was verified on 0.8% agarose gels and later stained with ethidium bromide. All genomic DNA were used for PCR amplification.

VI. PCR amplification of 16S rDNA and gel electrophoresis:

Amplification of the gene encoding the 16S rRNA was perfomed by PCR using the forward primer Bact-27F and the reverse primer Bact-1392R (Table 2.1). The reaction mixture contained the following reagents at a final volume of 30 μ l : 1X Buffer, 2.5 mM Mg, 1X BSA New England®, 2.5 mM dNTP's, 1 pMol of BACT 1392R and BACT 27F, 0.026 U *Taq* polymerase (Promega) and 1 μ l of DNAtemplate (15 ng/ μ l).

For 16S rRNA gene amplification, the PCR cycles consisted of initial denaturation for 5 minute at 94° C; 30 cycles of denaturation 1 minutes at 94° C, annealing 1 minute at 50 ° C and extension 3 minutes at 72° C, and final extension at 72° C for 10 minutes (Hezayen *et al.*, 2002). PCR amplicons were purified using the MinElute PCR purification kit (USA QIAGEN Inc.) according to manufacturer instructions and concentration determined with nanodrop.

VII. *Int*-PCR:

The PCR reaction was performed using the following reagents at a final volume 12.5 μ l : 1X Buffer, 3 mM Mg, 1X BSA New England®, 1.0 mM dNTP's, 5 pMol of hep35 and hep 36, 0.0104 U *Taq* polymerase (Promega) and 1 μ l of template (15 ng/ μ l).

PCR cycles were the following: initial denaturation at 94 °C by 5 minutes; 30 cycles of denaturation at 95 °C, annealing of 52 °C and extension of 72 °C, the final extension is of 72 °C by 7 minutes. The PCR product quality was verified on 1.5 % agarose gels and staining with ethidium bromide.

VIII. DNA sequencing:

Purified PCR products were sent for sequencing to UPR-Río Piedras Sequencing Facility and UW-htseq Sequencing Facility (Washington University at St. Louis, MO). Samples were prepared according to facility specifications.

IX. Sequence and Phylogenetic analyses:

Genomic integron integrases were edited in Bioedit and aligned against sequences in BLASTX using ClustaX. The sequences from Int1, Int2, Int3, XerC, and XerD (GenBank accession number AAQ16665.1, AAT72891.1, AA2A3VO32355.1, POA8P6.1, and POA8P8.1 were used as reference for identification of functional group and domains. Comparative amino acid sequences were performed constructing consensus neighbor-joining dendrograms using Mega 4 software package as previously described (Rodríguez–Minguela et al., 2009). The dendograms were generated with p-distance model and tested by 2,000 boostrap replications.

Application	Primer	Sequence (5' to 3')	Product size (pb)	Annealing Temperatu °C	are Reference
Amplification of 16S rDNA	Bact- 27F 1392R	AGAGTTTGATCMTGGCTCAG ACG GGC GGT GTG TAC A	1365	50	Amman et al., 1995
Amplification of integron	hep35 (FWD)	TGC GGG TYA CCT ACA AAG CC	491	52	White et al., 2000
encoded inte- grases	hep36 (REV)	CAR CAD ATG CGT RTA RAT	491	52	
Promoter re- gion of T7	T7	TAA TAC GAC TCA CTA TAG GG	variable	50	
Promotor re- gion of SP6	SP6	ATT TAG GTG ACA CTA TAG AA	variable	50	

 Table 2.1 Oligonucleotides used for PCR analyses.

Results:

Morphological characterization

Samples from Solar Saltern Ponds, Cabo Rojo, showed growth on agar plates containing Luria Bertani medium with 5 and 8% (w/v) NaCl after a 3 days at 37°C. A total of 25 strains were isolated from Candelaria and Fraternidad lagoons, "puente" and the salterns ponds. Two strains PII5 and SII5 were positive for integron encoded integrases. Morphological characterization was based on the classical macroscopic technique of colony appearance. Microscopical characterization was performed using the Gram staining procedure with acetic acid fixation as described by Dussault (1995). Strain PII5 is a Gram positive curved rod and grew in LB with 5% (w/v) NaCl (Fig. 2.1). Strain SII5 is a Gram negative rod and grew in LB with 8% (w/v) NaCl (Fig. 2.1). Both isolates showed colonies with entire and flat morphologies.



Figure 2.1 Macroscopical and microscopical characteristics of strain SII5 and PII5 isolated from the Solar Saltern Ponds on Luria Bertanie medium, curved rods strain using light microscope.

Strain CS91 was isolated in a previous study from the Thermophilic Springs, Coamo on TM medium (Burgos, A., 2007). Cells are gram positive long rods (Fig.2.1) with irregular and entire, flat, white colonies.



Figure 2.2 Macroscopical and microscopical characteristics of strain CS91 isolated in a previous study (Burgos, 2007) from the Thermophilic Springs, Coamo on TM, using light microscopy.

Genomic DNA extractions from prokaryotic strains isolated from Solar Saltern Ponds in Cabo Rojo and Coamo Thermal Springs.

Genomic DNA from strains isolated from the Solar Saltern Ponds and Coamo Thermal Springs was extracted using the organic extraction protocol with phenol/chloroform. Figure 2 shows the quality of genomic DNA extraction. Genomic DNA was quantified using the Nanodrop ND-1000 Spectrophotometer© and concentration was adjusted to 15ng of template for 16S rDNA and integrase integron PCR amplification.



Figure 2.3 Genomic DNA extraction from prokaryotic cells isolated from the Solar Saltern Ponds PII2 and SII5 and CS91 isolated from the Coamo Thermal Springs by Burgos (2007).

Amplification of the 16S rDNA from strains from Solar Saltern Ponds and Coamo Thermal Springs:

The quality of the metagenomic DNA extraction was determined through 16S rDNA amplification by PCR. PCR amplicons were approximately 873 pb in size (Figure 1.11). It was necessary to perform annealing optimization of the PCR parameters for some strains in order to achieve optimal amplification. 1

873bp



Figure 2.4 16S rDNA amplification using universal primers Bact-1392R and Bact 27-F of the isolated strains from the Solar Saltern Ponds.

Int-PCR from genomic DNA from strains isolated from Solar Saltern Ponds and Thermal **Springs:**

A PCR product of 500 pb corresponding to a putative integrase was obtained by PCR from two prokaryotic strains isolated from Solar Saltern Ponds and one strain previously isolated from the Thermal Spring (Burgos, 2007). Amplicons were purified and sent for sequencing (Figure

2.4).



Figure 2.5 Int- PCR amplicons of PII5 strain isolated from the Solar Saltern of Cabo Rojo. Positive control contained integrase genes as template, and negative control without template DNA.

Integrase alignments analysis from strains PII5, SII5 and CS91:

Although the entire IntI putative sequences could not be retrieved using the primers hep35 and hep36 (Table 2.1), the integron integrase conserve motifs and residues critical for binding and recombination mechanism which could be indicative that recovered Int sequences can be functional. In addition, these sequences were compared against functional integron integrase in order to characterize the IntI sequences obtained from the halophilic and thermophilic strains. IntI sequences recovered from isolates were aligned against functional integron integrase of a class1 integrase described in Escherichia coli, IntI2 described in Shigella sonnei and IntI3 described in Klebsiella pneumonia (GenBank accession numbers: AAQ166651, AAT728911 and AAO323551, respectively). The numbers in the frames of the alignment were determined using Int4 (99031763) numbering, because their crystal structure and recombination site were determined (MacDonald et al., 2006). Residue R143 correspond to an arginine residue found within the Box I region of Int4.

Important motifs which are characteristic of the tyrosine recombinase family, such as Patch (II and III) and Box (I and II), were detected. Residues of the conserved tetrad RHRY, characteristic of all tyrosine recombinase family, were detected in the following positions: H268, R271 and Y303, and each amino acid was located inside of the Box II motif. The principal roles of these residues, specifically H268, are appropriate orientation of the DNA and stabilization of transition state (Messier and Roy, 2001). However, Y303 is the nucleophile responsible for the cleavage; in addition it forms a covalent intermediate by esterification.

Important residues relationship with catalytic and DNA binding functions were identified with 100% conserved: K160 inside of Box I, A205, L206, K209 and F223 inside of additional domain, G293 inside of Box II. The residue K160 is an important amino acid found in Patch II and maintains a possible important function in the cleavage activity (Rodríguez-Minguela et al., 2009). The residues A205 and L206 play a role in positioning and stabilization of active site of attC. The P211 was 100% conserved in all IntI sequences inside an additional domain.

The crystal structure of VchInt4 was determined and important residues for attachment and stabilization of synaptic complex during recombination process were identified. Residue W219 was associated with attachment and residues P232, H240 and H24 were implicated in stabilizing the complex during recombination process and were 100% conserved. Residue, K145 is present in all genomic sequences recovered except in the case of Int4 where the presence of Q145 appeared to be exclusive of the *Vibrio* clade (Rodríguez-Minguela et al., 2009).

Residue D191 is 100% conserved across the IntI sequences from isolates. Residues, G159, K160, G161, K163, R165 within the Patch II region were 100% conserved. A high degree of sequences conservation was observed within the Box II region in which residues: R271,H272, S273, F274, A275, T276, G282, D284, I285, R286 G293, H294, T299, T300, I302, Y303 T304, H305, V306 and L307 were 100% conserved. Residues L291, L292, G293, and H294 were 99-100% conserved in all Int sequences from isolates and have been implicated in having a key role in the recombination mechanism (Messier and Roy, 2001).



Figure 2.6 Partial amino acid sequence alignment of integrases from strains isolated against other tyrosine recombinase, IntI class 1 from *Escherichia coli*, IntI 2 from *Shigella sonnei*, IntI 3 from *Klebsiella pneumoniae*, IntI 4 from *Vibrio cholerae* and tyrosine recombinases XerC and XerD from *E. coli*. The position of conserved motifs among integron-encoded integrases is indicated by segments labeled as Boxes I and II and Patches II and III and as additional domains (AD) under the alignment. The degree of conservation (percent identity) among equivalent residues is indicated as follows: blue (100%), yellow (95%)

and 85%). Conserved amino acids (100% identical) with functions related to protein folding, DNA binding, and recombination activity are highlighted in red, respectively. Green columns correspond to H, R, and Y residues from the conserved RHRY tetrad, which is characteristic of the entire tyrosine recombinase family (Nunes-Du⁻bt, S., et al. 1998). GenBank accession numbers for the protein sequences of reference integrases (classes 1 to 4) and tyrosine recombinases XerC and XerD are AAQ16665, AAT72891, AAO32355, 99031763, POA8P6, and POA8P8, respectively.

Phylogenetic analysis of IntI from halophilic and thermophilic strains:

Three integron–encoded integrases were recovered from the isolated strains. These sequences were analyzed using multiple-sequence alignment made by using the Clustal W program and neighbor joining analysis with Mega 4 (Tamura et al., 2007). The sequences identified as integron integrases recovered from halophilic and thermophilic strains grouped clusters which excluded the tyrosine recombinases represented by XerC and XerD. Two clusters were formed with these sequences. The subsequent sequences formed one cluster; PII5 and SII5 the sequences obtained from two halophilic strains were closely related to IntI2 of an uncultured bacterium clone (ACM887541) and IntI2 from *Shigella sonnei*, with 64 and 68% identity respectively. The sequence CS91 found in thermophilic strain was closely related to IntI 1 described in *Escherichia coli* (AAQ166510) with 85% identity. Inside this cluster, Int3 from *Klebsiella pneumoniae* and IntI4 from *V. cholerae* formed their own independent branch (Fig. 2.7).



Figure 2.7 Consensus neighbor-joining tree (compared positions for 3 amino acid residues), illustrating phylogenetic relationship of IntI sequences recovery from halophilic and thermophilic strains. Values represent the percentage of 1,000 boostrap replications that supported the branch. Outgroup references were designate with a thicker black branches highlight (XerC, XerD, class IntI 1, IntI 2, IntI 3, and IntI4. Numbers select the clusters formed by recovery sequences and the percent identities of IntI sequences associated. Values of boostrap support below 20% are not demonstrated and Genbank accession numbers follow the outgroup sequences.

Phylogenetic analysis of 16S rDNA from halophilic and thermophilic strains:

The phylogenetic relationship of these sequences was determined using Clustal W program and neighbor joining analysis with Mega 4 (Tamura et al., 2007). Isolated strains from Solar Saltern Ponds were analyzed using the 16S rDNA gene and demonstrated that strain SII5 is closely related with the genus *Bacillus* and the closest relative was *Bacillus safensis* FO 036B^T with a boostrap value of 99%. Strain PII5 belonged to the *gamma-Proteobacteria* and is phylogenetically close to *Salinivibrio costicola* DSM 16359^T with boostrap value of 99% (Fig. 2.9).

Phylogenetic analysis using the 16S rDNA gene demonstrated that strain cs91 is closely related with to the genus *Geobacillus*, and the closest relative was *Geobacillus kaustrophilus* NCIMB 8547^T, with boostrap value of 95% (Fig. 2.10).



Figure 2.8 Neighbour-joining distance tree using 16S rDNA sequences of SII5 isolated from Solar saltern Ponds in Cabo Rojo and closely related species. Boostrap values higher than 50% are shown. *Halomonas marisflava* KCCM 80003T (AF2511431) was used as the outgroup.



0.01

Figure 2.9 Neighbour-joining distance tree using 16S rDNA sequences of PII5 isolated from Solar Saltern Ponds in Cabo Rojo and closely related species. Boostrap values higher than 50% are shown. *Zymobacter palmae* DSM 10491T (D14555) was used as the outgroup.



Figure 2.10 Neighbour-joining distance tree using 16S rDNA sequences of cs91 previously isolated by Burgos, A. from Coamo Thermal Springs and closely related species. Boostrap values higher than 50% are shown. *Chromohalobacter israelensis* DSM6768T (AF2118621) was used as the outgroup.

Discussion:

The objective of this part of the research was focused on the distribution and presence of integrases in strains isolated from environments with moderate anthropogenic impact and extreme conditions in Puerto Rico such as hypersaline water of the Solar Saltern Ponds in Cabo Rojo and the Coamo Thermal Springs. This study represents the first attempt to determine the presence of the integrase of integron into strains isolated from environments with moderate anthropogenic impact and extreme conditions in Puerto Rico and the Caribbean. Validation of DNA quality was performed for all strains, using 16S rDNA primers. Two bacterial isolates from Solar Saltern Ponds and one reactivated from one collection from Coamo Thermal Springs were screened for the presence of integron integrases by PCR. The genomic DNA of each bacterial was subject to 16S rDNA for phylogenetic analyze and determination of genus. SII5 a gram positive and rod shape bacterium isolated from Solar Saltern Ponds with 30% NaCl related phylogenetically to Bacillus safensis. B. safensis is a bacterium highly resistant to gamma and UV radiation and was the first isolated from spacecrafy-assembly facility of the Jet Propulsion Laboratory, Pasadena, California (Satomi et al., 2006). In addition, Gram-positive bacteria are major reservoirs of integron class 1 from aviar fecal material (Nandi et al., 2004). The strains SII5 and PII5 contain integrase related to integron integrase class 2 with domain and motifs feature of integron integrases, although the premature stop codon X179 was not.

Strain PII5 strain isolated hypersaline ponds with 8% NaCl was phylogenetic related to the genus *Salinivibrio* (Gamma-proteobacteria) which is gram negative curved rod shaped bacterium. Many bacterial isolated from sea water are found in hypersaline ponds as *Salinivibrio costicola subsp. vallismortis* isolated from a hypersaline pond located in Death Valley, California (Huang et al., 2000). The creation of this genus was a consequence of difference between *Vibrio*

The strain previously isolate from Coamo Thermal Springs is a gram positive rod- shaped bacterium close to genus *Geobacillus*. Their integrase integron is close to integron integrase class 1 with domain and motif features of integron integrases. The existence of high conservation of important motifs related with recombination mechanism, stabilization and binding activity show those probably are integron integrase functional. The genus of *Geobacillus* provokes the industrial interest for their thermostable enzyme used in biotechnological processes. Previous work show integron integrase in thermophilic bacteria of the genus of *Sulfobacillus*, although this work amplification of possible cassette with specific primers was performed (Ghauri et al., 2003), integron integrases have been reported in other extreme bacteria as *Acidithiobacillus ferooxidans* (Rowe-Magnus and Mazel, 2001).

The present work demonstrates that integron integrases are not restricted to pathogenic and multidrug resistant bacteria, in addition the integron integrases may be a common feature of bacterial isolated from halophilic and thermophilic environments. Mechanism of horizontal gene transfer, as integron can take place in halophilic, thermophilic in and soil cave environments. This is the first attermpt to study the integron integrases from environments with moderate anthropogenic impact and extreme conditions in Puerto Rico and the Caribbean. Moreover the recovery of integron integrase from isolated areas is a great opportunity to advance investigation in biotechnology.

Conclusions

- Molecular screening from environments with moderade antropogenic impact and extreme conditions (Solar Saltern Ponds, Culebrones Cave and Coamo Thermal Springs) demostrated the presence of 66 novel integron encoded integrases with motifs features of integron integrases. This might suggest integron encoded intregrases can be found at non clinical related environments. The ecological role of these novel integrases is still unknown.
- Integron encoding integrases were found on the gram-positive isolates: SII5 related to Bacillus (99% similarity) and CS91 related to Geobacillusis (98% similarity). This provides additional evidence of the presence of this system in gram-positive bacteria.
- Isolate PII5 related to the family *Vibrionaceae* showed the presence of a novel integron encoded integrase this could be one of the first report inside of the genus *Salinivibr*io (99% similarity).

Recommendations

- This study demonstrated that integron integrases are distributed in environments with moderate anthropogenic impact from Puerto Rico, although the bacterial that carry these integrases are unknown. Important development techniques for detected the prokaryotic population associated to this integron integrases, were used without the need to perform culture technique dependent.
- The metagenomic technique is a useful technique to obtain complete and intact genes, without requiring PCR technique and primer design.
- Design different media that allow the isolation of different bacteria and determine preence or absence of integron integrases.
- Design a temporal study to compare prokaryotic diversity during different seasons of the year and associate it with the presence or absence of integron integrases.
- Determine the stability of integrons during a period of time in environments with modeate anthropogenic impact.
- Perform a gene cassette recovery and determine their functions using culture dependent and independent techniques.

Literatured Cited

- Altschul, S.F., Madden, T.L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25; p.# 3389-3402.
- Amman, R.I., Ludwig, W., and Schleifer, K.H. 1995. Phylogenetic identification and in situ detection of individual microbial cell without cultivation. Microbiol. Rev. 59: p.# 143-169.
- Antunes, P., Machado, J., and Peixe, L. 2006. Characterization of antimicrobial resistance and class 1 and 2 integrons in *Salmonella enterica* isolates from different sources in Portugal. J. Antimicrob Chemother. 58; p.# 297-304
- Avila P., and de la Cruz, F. 1988. Physical and genetic map of the IncW plasmid R388. Plasmid 20: p.# 155-157.
- Baharoglu, Z. Bikard, D., and Mazel, D. 2010. Conjugative DNA Transfer Induces the Bacterial SOS Response and Promotes Antibiotic Resistance Development through Integron Activation. Plos genetic 6; p. # 1-10
- Benlloch, S., Acinas, S.G., Antón, J., López-López, S.P., Rodríguez-Valera, L. F. 2001. Archaeal Biodiversity in Crystalliezer Ponds from Solar Salterns: Culture versus PCR. *Microb. Ecol* 41; p. # 12-19.
- Cambray, G., Guerot, A., and Mazel, D. 2010. Integrons. Annual Reviews 44:p. 141-166
- Chen, C.Y., Wu, K.M., Chang, Y.C., Chang, C.H., Tsai, H.C., Liao, T.L., Liu, Y.M., Chen, H.J., Shen, A.B., and Li, J.C. 2003. Comparative genome analysis of Vibrio vulnificus, a marine pathogen. Genome Res., 13, p. # 2577–2587.
- Collis, C. M., Grammaticopoulos, G., Briton, J., Stokes, H. W and Hall, R. M. 1993. Sitespecific insertion of gene cassettes into integrons. Molecular Microbiology 9;p. # 41–52.
- Collis, C.M. and Hall, R.M. 1992. Site-specific Deletion and Rearrangement of Integron Insert Genes Catalyzed by the Integron DNA Integrase. J Bacteriol 174; p.# 1574-1585

- Díaz-Muñoz, G. Montalvo-Rodríguez, R. (2005). Halophilic Black Yeast *Hortaea werneckii* in the Cabo Rojo Solar Salterns: Its First Record for this Extreme Environment in Puerto Rico. *Caribb. J. Sci* 41, p. # 360-365.
- Díaz-Muñoz G. 2006. Fungal Diversity Present at a Hypersaline Environment in Cabo Rojo, Puerto Rico Determined by Characterization of Isolates and Analysis of Environmental rRNA Genes Clones Libraries. M.Sc. Thesis, Department of Biology, University of Puerto Rico, Mayagüez, PR, USA; p .# 140
- **Dillon, B., L. Thomas, G. Mohmand, A. Zelynski, and J. Iredell.** 2005. Multiplex PCR for screening of integrons in bacterial lysates. J. Microbiol. Methods. 62, p.# 221-232.
- Elsaied, H., Stokes, H.W., Nakamura, T., Kitamura, K., Fuse, H., and Maruyama, A. 2007. Novel and diverse integron integrase genes and integron-like gene cassettes are prevalent in deep-sea hydrothermal vents. Environ Microbiol, 9; p. # 2298-312
- Elsaied, H., Stokes, H.W., Nakamura, T., Kitamura, K., Kurusu, Y.,Kamagata, Y., and Maruyama, A. 2011. Marine integrons containing novel integrase genes attachment sites, *attl*, and associated gene cassettes in polluted sediments from Suez and Tokyo Bays. ISME Journal, 5:p. #1162-1177
- **Felsenstein, J.1993**. PHYLIP (Phylogeny Inference Package), version. 3.6. Seattle: Department of Genetics, University of Washington.
- Ferguson,G.P., Totemeyer,S., MacLean,M.J. and Booth,I.R. 1998. Methylglyoxal production in bacteria: suicide or survival? Arch. Microbiol., 170, p. # 209–218.
- Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A., Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G. and Kelley,J.M. 1995. The minimal gene complement of Mycoplasma genitalium. Science, 270, p.# 397–403.
- **Friedrich, M. W. 2002.** Phylogenetic analysis reveals multiple lateral transfers of adenosine-5_phosphosulfate reductase genes among sulfate-reducing microorganisms. J. Bacteriol. 184., p. # 278–289.

- Fluit, A.C., and Schmitz, F.J. 1999. Class 1 integrons gene cassette, mobility, and epidemiology. Eur J Clin Microbio Infect Dis. 18; p. # 761-770
- Gamieldien, J., A. Ptitsyn, and W. Hide. 2002. Eukaryotic genes in *Mycobacterium tuberculosis* could have a role in pathogenesis and immunomodulation. Trends Genet. 18, p. # 5–8.
- Garcia-Quintanilla M, Ramos-Morales F, Casadesus J (2008) Conjugal transfer of the Salmonella enterica virulence plasmid in the mouse intestine. J Bacteriol 190: p. # 1922–1927.
- Ghauri, M.A., Khalid, A.M, Grant, S., Grant, W., and Heaphy, S. 2006. High-salt environments and identification of gene cassette asociated open reading frame. Current Microbiology. 52; p. # 487–492
- Ghauri, M.A., Khalid, A.M, Grant, S., Heaphy, S. and Grant, W. 2003. Phylogenetic analysis of different isolates of *Sulfobacillus* spp. Isolated from uranium-rich environments and recovery of genes using integron-specific primers. Extremophiles. 7; p. # 341-345
- **Glover, L. 1971.** Geology of the Coamo Area, Puerto Rico and its relation to the volcanic Arc-Trench association. Geological Survey Professional Paper. United States Government Printing Office, Washington.
- Guerin, E., Cambray, G., Sanchez-Alberola, N., Campoy, S., Erill, I., Da Re, S., Gonzelez-Zorn, B., Barbe, J., Marie-Cecile, P., and Mazel, D. 2009. The Sos Response Controls Integron Recombination. Science 324; p.# 1034
- Gopaul, D. N., and G. D. Van Duyne. 1999. Structure and mechanism in site-specific recombination. Curr. Opin. Struct. Biol. 9: p. #14–20.
- Grainge, I., and M. Jayaram. 1999. The integrase family of recombinase: organization and function of the active site. Mol. Microbiol. 33: p. # 449–456.
- Güssow D. and Clackson T. 1989. Direct clone characterization from plaques and colonies by the polymerase chain reaction. *Nucleic Acids Res.*, 17, 4000
- Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Doson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R.D., Nierman, W.C. and White, O. 2000. DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae. Nature 406, p. # 477- 483.
- Hentschel, U., and J. Hacker. 2001. Pathogenicity islands: the tip of the iceberg. Microbes Infect. 3. p. # 545–548.
- Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B.C. and Herrmann, R. 1996. Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae. Nucleic Acids Res., 24,
- Igarashi, N., J. Harada, S. Nagashima, K. Matsuura, K. Shimada, and K. V. P. Nagashima. 2001. Horizontal transfer of the photosynthesis gene cluster and operon rearrangement in purple bacteria. J. Mol. Evol. 52, p. # 333–341.
- Jiang, S., and Paul, J. H. 1998. Gene transfer by transduction in the Marine Environment. Appl Environ Microbiol. 64; p. # 2780-2787
- Koenig, J. E., Boucher, Y., Charlebois, R.L., Nesba, C., zhaxybayeva, O., Bapteste, E., Spencer, M., Joss, M., Stokes, H.W. and Dooittle, W.F. 2008. Integron-associated gene cassettes in Halifax Harbour: assessment of a mobile gene pool in marine sediments. Environ Microbiol, 10; p. # 1462-2920
- Konings, W., Sonja-Verena A., Koning, S., and Driessen, A. 2002. The cell membrane plays a crucial role in survival of bacteria and archaea in extreme environments. Annie van Leewenhook 81; p. # 61-72
- Maidak, B.L., Olsen, G.J., Larsen, N., Overbeek, R., McCaughey, M.J., and Woese, C.R. 1996. The ribosomal database project (RDP). Nucleic acids Res 24; p. # 82-85
- Marchler-Bauer A., and Bryant, S.H. 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32; p. # 327-331.

- Martinez-Freijo, P. Fluit, A.C., Scmitz, F.J., Grek, JV.S., Verhoef, J., and Jones, M.E. 1998. Class 1 integrons in Gram-negative isolates from different European hospitals and association with decreased susceptibility to multiple antibiotic compounds. J Antimicrob Chemoter 42; p. # 689-696
- Mazel, D. 2006. Integrons: agents of bacterial evolution. Nat Rev Microbiol 4; p. # 608-620
- Mazel, D. Dychinco, B., Webb, V.A., and Davies, J. 1998. A distinctive class of integron in the *Vibrio cholera* genome. Science 280; p. # 605-608
- Mongodin, E.F., Nelson, K.E., Daugherty, S., DeBoy, R. T. Wister, J., Khouri, H., Weidman, J., Walsh, D.A, Papke, R.T., Sanchez-Perez, G., Sharma, A. K., Nesbo, C. L., Bapteste, E., Doolittle, W. F., Charleboiss, R.L., Legault, B. and Rodriguez-Valera, F. 2005. The genome of *Salinibacter ruber*: Convergence and gene exchange among hyperhalophilic bacteria and archaea. PNAS 102; p. #18147-18152
- Montalvo-Rodríguez R., López-Garriga J., Vreeland RH., Oren A., Ventosa A., and Kamekura M. 2000. *Haloterrigena thermotolerans* sp. Nov., a halophilic archaeon from Puerto Rico. Int J Syst Evol Microbiol 50; p. # 1065-1071
- Montalvo-Rodríguez, R., Ruiz-Acevedo, A., and López-Garriga, J. 1997. New isolates of extremely halophilic Archaebacteria (Halobacteria) from Puerto Rico and the Caribbean. Caribb. J. Sci. 33; p.98-104
- Montalvo-Rodriguez R., Vreeland R., Oren A., Kessel M., Betancourt C., and López-Garriga. 1998. *Halogeometricum borinquense* gen. nov., sp. Nov., a novel halophilic archaeon from Puerto Rico. Int J Syst Bacteriol 48; p. # 1305-1312
- Nandi, S. Maurer, J.J. Hofacre, C., and Summers, A.O. 2004. Class 1 antibiotic resistance integrons in poultry litter. PNAS 101; p. # 7118-7122
- Nield, B. S., Holmes, A. J., Gillings, M. R., Recchia, G. D., Mabbutt, B. C. Nevalainen, K. M., and Stokes, H. W. 2001. Recovery of new integron classes from environmental DNA. FEMS Microbiol. Lett. 195; p.# 59-65

- Nemergut, D.R, Martin, A.P. and Schmidt, S.K. 2004. Integron Diversity in Heavy-Metal-Contamined Mine Tailings and Inferences about Integron Evolution. Appl Environ Microbiol; 70; p. #1160-1168
- Nesvera, J. Hochmannova, J., and Patek, M. 1998. An integron of class 1 is present on the plasmid pCG4 from gram-positive bacterium *Corynebacterium glutamicum*. FEMS Micriobiol Lett. 169; p. # 391-395
- Nunes-Du^{*}by, S. E., H. J. Kwon, R. S. Tirumalai, T. Ellenberger, and Landy, A. 1998. Similarities and differences among 105 members of the Int family of site-specific recombinases. Nucleic Acids Res. 26; p. # 391–406.
- Ochman, H., Lawrence, J. G., and Groisman, E.A. 2000. Lateral gene transfer and the nature of bacterial innovation. Nature 405; p. # 299-304
- O'Halloran, F., Lucey, B. Cryan, B. Buckley, T., and Fanning, S. 2004. Molecular characterization of class 1 integrons from Irish thermophilic *Campylobacter* spp. J Antimicrob Chemother 53; p.# 952-957
- **Oren, A., and Francisco-Rodriguez V. 2001**. The contribution of halophilic Bacteria to the red coloration of saltern crystallizer ponds. FEMS Microbiol Ecol 36; p.# 123-130
- Osano, E., Y. Arakawa, R. Wacharotayankun, M. Ohta, T. Horii, H. Ito, F. Yoshimura, and N. Kato. 1994. Molecular characterization of an enterobacterial metallo-β-lactamase found in a clinical isolate of *Serratia marcescens* that shows imipenem resistance. Antimicrob. Agents Chemother. 38; p. # 71-78
- Quiñones, F. and Guzmán, S. 1983. Geochemistry of the thermal springs at Coamo, Puerto Rico. Water resources investigations in Puerto Rico and the U.S. Virgin Islands: U.S. Geological Survey. 10, p. #66-72.
- Rodríguez-Durán, A. 1998. Nonrandom aggregations and distribution of cave-dwelling bats of Puerto Rico. Journal of Mammalogy 79: p. # 141-146.

- Rodríguez-Durán, A. and J. A. Soto-Centeno. 2003. Temperature selection by tropical bats roosting in caves. Journal of Thermal Biology, in press.
- Rodríguez-Minguela, C.M., 2005. A Molecular Approach for the Detection and Characterization of Novel Tetracycline Resistance Genes, Integrons, and Integrons-Encoded Antibiotic Resistance Determinants in the Environment. Thesis, Department of Crop and Soil Sciences, Michigan State University, USA.
- **Rodríguez-Minguela, C.M., J.H.A. Apajalahti, B. Chai, J.R.Cole, and J.M. Tiedje**. Integron-encoded integrases are widely distributed across diverse terrestrial and deep sea environment around the globe. (submitted to PNAS).
- Rowe-Magnus, D.A., and Mazel, D. 2001. Integrons: natural tools for bacterial genome evolution. Current Opinion in Microbiology; 4 p.# 565-569
- Rowe-Magnus, D.A., Guerot, A.M. and Mazel, D. 1999. Super-integrons. Research in Microbiology; 150 p.# 641-651
- **Rowe-Magnus, Guerout, A.M., Ploncard, P., Dychinco, B., Davies, J., and Mazel, D. 2001**. The Evolutionary history of chromosomal super-integrons provides an ancestry for multireistant integrons. PNAS 98; p. # 652-657
- Saano, A., Tas, S., Pippola, K., Lidstrom, K., and Van-Elsas, J. 1995. Nucleic acids in the environment: methods and application. Trevors, J.T. and Van Elsas, J. (eds). Berlin: Springer Verlag. p.# 49-67
- Sabaté, M. and Prats, G. 2002. Estructura y función de los integrones. Enferm Infecc Microbiol Clin, 20; p.# 341-345
- Sáenz, Y., Vinué, L., Ruiz, E., Somalo, S., Bezares-Rojo, B., Zarazaga, M., and Torres, C.
 2010. Class 1 integrons lacking qacE ; and sul1 genes in Escherichia coli isolates of food, animal and humans origen. VetMicrobiol

- Satomi M., La Duc M.T. and Venkateswaran K. 2006. *Bacillus safensis* sp. nov., isolated from spacecraft and assembly-facility surfaces. *International Journal of Systematic and Evolutionary Microbiology*, 56(8): p. #1735-1740.
- Sobecky, P.A., and Hazen, TH. 2009. Horizontal gene transfer and mobile genetic elements in marine systems. Methods Mol Biol: 532: p. #435-453
- Sobecky, P. A., Mincer, T. J., Chang, A. T., and Helinski, D. R. 1998. Isolation of Broad-Host-Range Replicons from Marine Sediment Bacteria. Appl Environ Microbiol. 64; p. # 2822-2830
- Stokes, H. W., and Hall, R.M. 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. Mol. Microbiol. 3; p.# 1669– 1683.
- Stokes, H.W., Holmes, A.J., Nield, B.S., Holley, M.P. Helena-Nevalainen, K.M., Mabbutt, B.C., and Gillings, M.R. 2001. Gene Cassette PCR: Sequence-Independent Recovery of Entire Genes from Environmental DNA. Appl Environ Microbiol 67; p.# 5240-5246
- Stokes, H.W., Nesbo, C. L., Holley, M.,Bahl, M. I., Gillings, M. R. and Boucher, Y. 2006. Class 1 Integrons Potentially Predating the Association with Tn402-Like Transposition Genes Are Present in a Sediment Microbial Community. J Bacteriol. 188; p. # 5722-5730
- Sullivan, J. T., S. D. Brown, R. R. Yocum, and C. W. Ronson. 2001. The bio operon on the acquired symbiosis island of *Mesorhizobium* sp. strain R7A includes a novel gene involved in pimeloyl-CoA synthesis. Microbiology 147 :p. # 1315–1322
- Sundström L., Rådström, P., Swedberg, G., and Sköld. 1998. Site-specific recombination promotes linkage between trimethoprim-and sulfonamide resistance genes. Sequences characterization of *dhfrV* and *sull* and a recombination active locus of *Tn21*. Mol. Gen. Genet. 213; p. #191-201
- Sundström L., and Sköld. 1991. Site-specific insertion of three structural gene cassettes in transposon Tn7. J. Bacteriol. 173; p. # 3025-3028

- Sundström L., and Sköld. 1990. The dhfrl trimethoprim resistance gene of Tn7 can be found at specific sites in other genetic surroundings. Antimicrob. Agents Chemother. 34; p.# 642-650
- Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: p.# 1596-1599
- **Tiedje, J.M. and J.L. Stein. 1999.** Microbial Biodiversity: Strategies for its recovery. In A.L. Demain and J. Davies (eds.). Manual of Industrial Microbiology and Biotechnology. Amer. Soc. Microbiol., Washington, D; p. # 682-692
- **Toussaint, A. and Merlin, C. 2002.** Mobile elements as a combination of functional modules. Plasmid, 47; p. # 26-35
- Utaker, J. B., K. Andersen, A. Aakra, B. Moen, and I. F. Nes. 2002. Phylogeny and functional expression of ribulose 1,5-bisphosphate carboxylase/oxygenase from the autotrophic ammonia-oxidizing bacterium *Nitrosospira sp.* isolate 40KI. J. Bacteriol. 184; p. # 468-478.
- van den Burg, B. 2003. Extremophiles as a source for novel enzymes. Curr Opi Microbiol, 3; p. # 213-8
- van-Elsas, J. D., and Bailey, M.J. 2002. The Ecology of transfer of mobile genetic elements. FEMS Microbiology Ecology. 42; p. 187-197
- Ventosa A., Nieto J., and Oren A. 1998. Biology of Moderately Halophilic Aerobic Bacteria. Microbiol Mol Biol Rev 62; p.# 504-544
- Whitman, C.P. 2002. The 4-oxalocrotonate tautomerase family of enzymes: how nature makes new enzymes using a beta-alpha-beta structural motif. Arch. Biochem. Biophys., 402, p. # 1–13.

- Xu, H., Davies, J. and Miao, V. 2010. Molecular Characterization of Class 3 Integrons from Delfia spp. J. Bacteriol. 189 (17): p.# 6276-6283.
- Young-Kang H., Sook-Jeong, Y., Young-Oh, J., Ho-Tae, S., Hee-Choi, C., Chan-Moon, D., Kil-Lee, W., Chul-Lee, Y., Yong-Seol, S., Taek-Dong, C., and Chul-Lee, J. 2005. Characterization of antimicrobial resistance and class 1 integrons found in *Escherichia coli* isolates from humans and animals in Korea. J Antimicrob Chemother 55; p.# 639-644