· Article ·
doi: 10.13679/j.advps.2015.1.00030

Advances in Polar Science
March 2015 Vol. 26 No. 1: 30-37

Detection of integron integrase genes on King George Island, Antarctica

Verónica Antelo[1]*, Héctor Romero[2] & Silvia Batista[1]

[1] Unidad de Microbiología Molecular, Biogem, Instituto de Investigaciones Biológicas Clemente Estable (MEC), Montevideo, C.P. 11600, Uruguay;

[2] Departamento de Ecología y Evolución, Facultad de Ciencias, UDeLaR, Montevideo, C.P. 11400, Uruguay

Received 5 August 2014; accepted 25 December 2014

Abstract The presence and diversity of class 1 integrase gene (*intI*) sequences were evaluated by PCR using previously designed primers. Two clone libraries were constructed from DNA in sediment and microbial mat samples collected on Fildes Peninsula, King George Island, Antarctica.The libraries constructed from samples collected at Halfthree Point (HP) and Norma Cove (NC) contained 62 and 36 partial *intI* sequences, respectively. These sequences clustered into 10 different groups with < 95% amino acid identity. Alignment of the deduced amino acid sequences with those from recognized integron-encoded integrases demonstrated the presence of highly conserved motifs characteristic of *intI* integrases. The HP library contained 42 nucleotide sequences identical to the class 1 *intI* gene found in a collection of trimethoprim-resistant (Tmp^r) Antarctic *Enterobacter* sp. isolates, previously collected in the same area. These integrons, located on plasmids, had a genetic organization similar to that of pKOX105 from *Klebsiella oxytoca*. The 20 remaining HP and NC library sequences were similar to integrase sequences previously determined in a metagenomic analysis of environmental samples. We have demonstrated the presence of integron integrase genes in Antarctic sediment samples. About half these genes were very similar to the class 1 integrons found in human-associated microbiota, suggesting that they originated from human-dominated ecosystems. The remaining integrase genes were probably associated with endemic bacteria.

Keywords integrase, amplicon library, maritime Antarctica

Citation: Antelo V, Romero H, Batista S. Detection of integron integrase genes on King George Island, Antarctica. Adv Polar Sci, 2015, 26: 30-37, doi: 10.13679/j.advps.2015.1.00030

1 Introduction

Integrons are genetic elements capable of capturing and expressing promoterless open reading frames (ORFs), designated “gene cassettes”. These elements were first described in pathogenic bacteria, and were later identified in diverse environmental bacteria using both culture-dependent and culture-independent methods. Gene cassettes encode diverse functions associated with environmental adaptation[1-3]. Their incorporation into an integron is mediated by site-specific recombination, requiring an integron integrase, IntI. The capture or excision of a circularized gene cassette involves a recombination event between the integron *attI* site and an *attC* site (59-base element) on the gene cassette fragment[4-5] (Figure 1). IntIs recognize specific *attI* sites but can interact with different *attC* sites in different hosts, allowing the transfer of ORFs to genetically diverse bacteria[4-5]. Integrons are not mobile elements per se, but facilitate the transfer and spread of gene cassettes, allowing host adaptation to new environments[6]. Integrons can be found in chromosomes, but are also associated with mobile elements, such as plasmids and transposons, which can be exchanged much more readily between organisms in a given environment[5].

* Corresponding author (email: veronicaantelo@gmail.com)

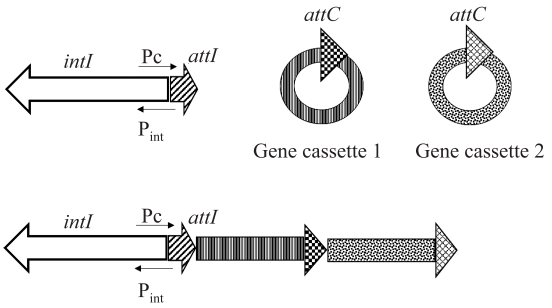

Figure 1 Basic structure of an integron including *intI* integrase gene, a promoter Pc controlling the expression of inserted gene cassettes. P_{int}: promoter of *intI* gene; *attI*: primary recombination site; circular gene cassettes: an ORF and a recombination site (*attC*).

Integrons were initially described in clinical isolates, and were classified into one of three classes based on the IntI sequences encoded[6]. These integrons have some characteristics in common: they are associated with plasmids and transposons, and contain a relatively small number of gene cassettes (ca. six). In general, gene cassettes are shared between these classes, including determinants of antibiotic resistance[1].

The inclusion of new sequences from genomic and metagenomic data has allowed a more complex picture of integrons to be established. Integrons have been grouped according to the level of homology among their *intI* genes, although specific values that distinguish each integron class have not been established[5].

Class 1 integrons are the most commonly found and studied[5]. This class is usually recovered from clinical samples, and contributes to the dissemination of antibiotic resistance in this context. The most probable evolution of the class 1 integrons has been inferred from the high diversity of *intI* sequences detected in the environment with metagenomic DNA analyses[7-8], which contrasts with the conserved sequences of clinical *intI* genes. Accordingly, the inferred scenario is that a chromosomal integron was captured by a Tn402-like transposon by recombination, generating a hybrid mobile element[8]. The original transposition functions of the Tn402-like element were eventually lost, and only portions of the original sequences are present in clinical integrons. These Tn402-associated integrons could have some features that expedite their mobilization into different plasmids, increasing their ability to incorporate and spread gene cassettes, and possibly facilitating their adaptation to different environments.

Metagenomic studies continue to recover new integrons, demonstrating that much of the diversity of this family of genetic elements is still unknown[7,9]. Early studies in clinical environments relied on culture-dependent methods. However, novel culture-independent methods, with a greater capacity to detect these elements in their corresponding hosts, might generate new information if used to analyze clinical samples.

To describe the prevalence of class 1 integrons in a presumed low-human-impact environment, we constructed PCR amplicon libraries of fragments of *intI* genes from metagenomic DNAs isolated from two sites on Fildes Peninsula, King George Island, in maritime Antarctica. The presence of integrases in samples from King George Island has not been reported previously. However, Stokes et al.[10] have previously reported integron gene cassettes amplified from total DNA purified from ornithogenic soil samples from Cape Denison on continental Antarctica[10].

In a previous study using culture-dependent methods to analyze samples collected from different sites on Fildes Peninsula, we identified "classical" class 1 integron genes (*intI*) in several psychrotolerant bacteria displaying high levels of resistance to antibiotics, including trimethoprim (Tmp), streptomycin (Str), and ampicillin (Amp). In some cases, the integrase gene and adjacent regions were on a conjugative plasmid with sequences highly similar to those of pKOX105 of *Klebsiella oxytoca* and other clinical bacteria[11]. In that study, we used a primer pair designed by Mazel et al.[12] to amplify class 1 *intI* genes (int1.F/int1.R). This primer pair was also successfully used by Nemergut et al.[9] to detect both "classical" *intI1* and 12 new *intI* genes from heavy-metal-contaminated mine tailings. These new sequences displayed high diversity compared with the diversity of class 1 *intI*, with less than 90% protein sequence similarity. That study[9] suggested that culture-independent methods using these primers might detect greater *intI* diversity than that detectable with culture-dependent methods.

2 Materials and methods

2.1 Sampling procedure

Samples of sediment/soil and bryophyte fragments associated with a benthic microbial mat were collected (Table 1) from two sites on Fildes Peninsula, King George Island, in maritime Antarctica during the Antarctic summer (Figure 2). The samples were collected in 2008 and 2013, during campaigns organized by the Instituto Antártico Uruguayo. Samples for total DNA purification were collected in sterile 50 mL tubes, which were stored at –20°C until processing, except during their transfer between Antarctica and the laboratory in Montevideo, when they were stored on ice. The samples from Halfthree Point (HP) were brown microbial mats associated with bryophytes. The samples collected from Norma Cove (NC) were sediment/soil associated with bryophytes. The NC sampling site, located at a relatively high altitude, is covered by bryophytes during the summer.

Table 1 Sample sites in King George Island

Sample-site-campaign	Location	Description
HP-Halfthree Point (March—April, 2008)	62°13′29″ S 58°57′9″ W	Brown microbial mat with bryophytes
NC-Norma Cove (January, 2013)	62°11′18″ S 58°54′32″ W	Sediment / Soil with bryophytes

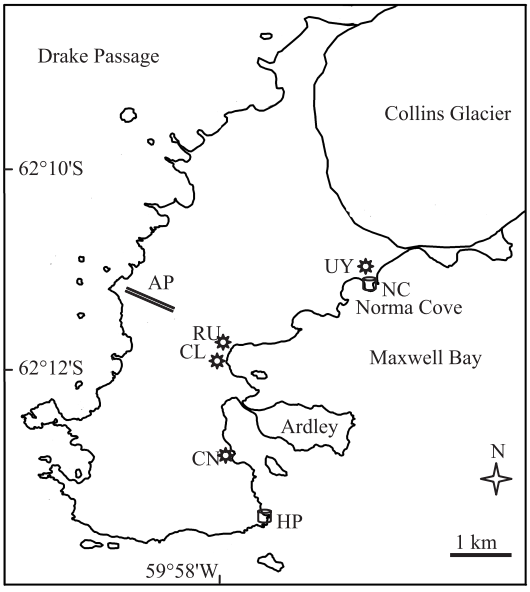

Figure 2 Map of Peninsula Fildes. Antarctic bases (✿): UY, Uruguay; RU, Russia; CL, Chile; CN, China; AP, Chilean airport. Sampling locations (⛁)

2.2 DNA purification

Total DNA was purified from approximately 0.5 g of each sample using the ZR Soil Microbe DNA Microprep™ (Zymo Research Corporation, Irvine, CA, USA), according to the instructions of the manufacturer.

2.3 Library construction

A single set of primers was used, designed for the specific amplification of 484 bp of the integrase gene (Table 2). Amplification was performed in a Thermo® Electron Px2R Thermocycler (Thermo Fisher Scientific, Waltham, MA, USA). The reaction mixtures included (in a final volume of 25 μL): 2.5 μL of 10 X reaction buffer, 2.5 μL of 10 μM each primer, 2.5 μL of dNTPs (2 mM each), 0.5 μL of *Taq* DNA polymerase (5 U·μL^{-1}; SBS Genetech, Beijing, China), 1 μL of DNA template (approximately 50 ng of purified DNA), and 13.5 μL of ultrapure water. The reactions were initially maintained at 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 1 min, with a final extension at 72°C for 10 min. Template prepared from *Escherichia coli* JM83 containing kanamycin-resistant (Km^r) plasmid pAT674 was used as the positive control[13].

The amplification products were purified with agarose gel electrophoresis using the PureLink™ Quick Gel Extraction Kit (Invitrogen, Life Technologies, Carlsbad, CA, USA), according to the instructions provided. The purified DNA was ligated to the pCR®2.1 TOPO vector provided with the TOPO® TA Cloning® Kit (Invitrogen). The ligation products generated were used to transform chemically competent *E. coli* TOP10 cells (Invitrogen). White colonies growing on LB plates with 50 μg·mL^{-1} Amp and 20 μg·mL^{-1} X-gal were isolated. Each library consisted of cells transformed with pCR®2.1 TOPO containing the cloned fragments.

Table 2 Oligonucleotides employed in this study

Oligonucleotides	Nucleotide sequence (5′ to 3′)	Target gene	Amplicon size/pb	Reference
IntI1_fwd	GGGTCAAGGA TCTGGATTTCG	*intI1*	484	Mazel et al[12]
IntI1_rev	ACATGCGTGTA AATCATCGTCG	*intI1*		

2.4 Selection of clones from each library for DNA sequence analysis

The plasmids were prepared, the correct sizes of the inserted fragments were verified, and the inserts were sequenced. The plasmids were isolated from overnight cultures in 5 mL of LB broth containing Amp, using a FlexiPrep Kit (Amersham Pharmacia Biotech, Inc., Piscataway, NJ, USA), according to the manufacturer's instructions. The correct sizes of the insert fragments were verified with agarose gel electrophoresis after the plasmids were digested with ***Eco***RI. Selected plasmids were subjected to DNA sequence analysis.

The DNA sequences of both libraries were determined by the Sequencing Service of the Instituto Pasteur (Montevideo, Uruguay) using primers M13 Forward and M13 Reverse, which recognize both sides of the inserts in pCR® 2.1 TOPO (Invitrogen).

2.5 DNA sequence analysis

The sequences obtained were analyzed with VecScreen to identify and remove segments of vector origin (http://www.ncbi.nlm.nih.gov/tools/vecscreen/). The DNA sequences were also analyzed with MEGA 5.2 (http://megasoftware.net/).

The sequence analyses were performed with the BLAST program in the National Center for Biotechnology Information nonredundant protein sequence database (http://blast.ncbi.nlm.nih.gov/Blast.cgi)[14] and metagenomic databases from King George Island (https://img.jgi.doe.gov/cgi-bin/m/main.cgi) containing marine microbial communities from chronically polluted sediments at four geographic locations (IDs 3300000129, 3300000132, 3300000119, and 3300000123). We used an e-value threshold of 10^{-5} and a minimum coverage for our amplified sequences of 100 amino acids.

Sequences similar to IntI proteins from the eight named classes were obtained with a BLAST search of all completely sequenced genomes of the U.S. Department of Energy Joint Genome Institute (DOE JGI), using the same criteria as used in the previous searches.

The deduced amino acid sequences from each library were aligned against functional integrase sequences (classes 1, 3, and 4) and tyrosine recombinases XerC and XerD (GenBank accession numbers: IntI1, AAQ16665; IntI2, AAT72891; IntI3, AAO32355; XerC *E.coli*, P0A8P6; XerD *E.coli*,

P0A8P8; XerC *Thiobacillus denitrificans*, 499632306; XerD *T. denitrificans*, 516743720) to identify the motifs conserved among the integron-encoded integrases[15].

The alignments were generated with the Muscle software program (version 3.7)[16] and edited with Jalview (version 2.8.1)[17].

2.6 Phylogenetic analyses

A phylogenetic analysis based on the alignment of the amino acid sequences deduced from the *intI* gene fragments in the libraries, the best BLAST hits selected from the GenBank database, and the BLAST hits from the Integrated Microbial Genomes and Metagenomes database (IMG/M) was conducted with the maximum likelihood method, at the web server http://phylogeny.fr[18] with the default options in the PhyML software.

3 Results

Some areas of Fildes Peninsula have been designated Antarctic Special Protected Areas (ASPAs) because of their archaeological or ecological value. Fildes Peninsula is ASPA No. 125, and specific protected areas within it are designated ASPA 125 zones under the regulations of the Antarctic Protected Areas System. One of these areas is located within Halfthree Point (zone 125d), very close to one of the sites analyzed in this study (HP). The HP sample location was selected because of its longer distance from scientific stations compared with NC site, located adjacent to zone ASPA 125f, next to the Uruguayan Station (Artigas Base).

The total DNA extracted from each sample was used as the template for PCR with primers IntI1_fwd and IntI1_rev (Table 2), generating single products of the expected size (ca. 500 bp). Clone libraries were prepared using the DNA amplification products generated from samples HP and NC. Seventy-five clones from each library, with inserts of the expected size, were randomly selected and the insert of each was sequenced.

In total, 98 high-quality sequences were obtained: 62 from the HP samples and 36 from the NC samples. A DNA sequence analysis with BLASTn indicated that 42 HP library clones had the same sequence and were very similar (99%–100% identity) to those of class 1 *intI* genes. The sequences of these 42 inserts were identical to that found in seven *Enterobacter* sp., resistant to Amp, Tmp, and Str, that were recovered on LB solid medium from soil samples collected from several sites throughout Fildes Peninsula[11]. These isolates were all psychrotrophs and the class 1 *intI* gene identified in six of the seven isolates analyzed had a genetic organization similar to that of pKOX105 from *Klebsiella oxytoca* (HM126016)[19]. The 42 *intI* gene fragments showed 100% identity (with BLASTn) to a fragment of the *intI1* gene isolated from *Pseudomonas putida* strain DLL-E4 (CP007620.1), *K. pneumoniae* Kp15 plasmid pENVA (HG918041), and *Xanthomonas oryzae* pv. *oryzae* strain YNA12-2 (HQ662557), of both environmental and clinical origins.

The 20 remaining HP *intI* sequences and the 36 *intI* from NC displayed various levels of similarity (with BLASTx) to integrase sequences obtained in metagenomic analyses of environmental samples[20-23], and in one study using the same primer pair for amplification as used in our study[9].

3.1 Alignment analysis

Integron integrases are members of the tyrosine recombinase family. They contain four very highly conserved residues (RHRY), regions of conserved motifs (boxes I and II, patches I, II, and III)[15], and an additional domain (AD) of about of 35 residues, specific to integrons[15,24]. These conserved elements allow the identification of potentially "new" *intI* genes[20]. In this study, the inferred amino acid sequences (160 amino acids) from both libraries were aligned with representative integron-encoded integrases and the tyrosine recombinases XerC and XerD (GenBank accession numbers P0A8P6 and P0A8P8, respectively) using Jalview (Figure 3 shows representative clones from the libraries; complete data not shown). This allowed us to identify the conserved motifs noted above that are characteristic of *intI* sequences[15] (Figure 3). However, patch I is not shown in the 160 amino acid sequences analyzed. The motifs typical of the tyrosine recombinase family, such as boxes I and II, patches II and III, and AD, were identified in all the sequences, including those from the HP and NC libraries[15,24] (Figure 3). Residues essential for the DNA-binding[24] and recombination activities[24] of the enzymes were present, suggesting that they represent functional proteins[24] (Figure 3).

3.2 Phylogenetic analysis

A maximum likelihood phylogenetic analysis was performed of the deduced amino acid sequences from our Antarctic libraries, and all those sequences that were most similar (see Methods) to our library sequences, using nonredundant protein sequences from the GenBank database (http://blast.ncbi.nlm.nih.gov/Blast.cgi)[9,19-21] and a collection of sequences obtained from metagenomic studies undertaken at King George Island. We also included representative integron-encoded integrases from different classes (from classes 1–8) obtained from GenBank. The analysis showed that the IntI sequences displayed great diversity, and those from our libraries, isolated by PCR amplification with primers designed to amplify class 1 *intI* genes were dispersed throughout the phylogenetic tree. However, some of our sequences clustered in three separate groups, distinct from the recognized "classical" classes of integrases (Figure 4).

The sequences on the phylogenetic tree clustered into different groups. One group contained class 1 IntI sequences, usually recovered from clinical samples. This branch included 42 HP sequences, all identical to the IntI protein previously

Figure 3 Alignment of deduced amino acid *intI* sequences (160aa) of representative HP and NC clones with known integrases (IntI1, IntI2, IntI3, IntI4). The alignment also includes: IntIG. met from *Geobactermethallireducens* (WP_004514318), IntIN. eur from *Nitrosomonaseuropea* (WP_011111085), IntIX.camp from *Xanthomonascampestris* (WP_011035598), and tyrosine recombinases XerC (P0A8P6) and XerD (P0A8P8) from *E. coli*. Typical motifs from tyrosine recombinase family (Box I and II, Patches II and III) are shown within black border boxes. Additional domain (AD) typical of integron integrases are highlighted within a pink border box. Three residues from the conserved RHRY tetrad from the tyrosine recombinase family are indicated within red border boxes. Essential residues for binding and recombination activities are highlighted within blue border boxes. IntI1, from *E. coli* (AAQ16665), IntI2, from *Shigellasonei* (AAT72891), IntI3, from *Klebsiellapneumoniae* (AAO32355), IntI4, from *Vibrio cholerae* (AAC38424), IntI5, from *Vibrio mimicus* (AF180939). IntI6 (AAK00307), IntI7 (AAK00305) and IntI8 (AF314189).

found on this peninsula[11]. The tree shows one sequence (HP7) representative of these 42 sequences, to simplify the visualization of the results.

Thirty-one of the 36 NC sequences clustered as a single group, clearly separate from other recognized integrases.

Five sequences from the HP library (HP21, HP31, HP56, HP84, and HP133) and NC15 also clustered together. These HP sequences, NC15, and the 31 sequences from NC discussed above shared a common origin and seemed to be related to sequences recovered with metagenomic methods by Rodríguez-Minguela et al.[23].

Three sequences from the NC library (NC11, NC37, and NC38) clustered on a branch with integrases recovered from different undisturbed environments, including a dry forest (Puerto Rico), rainforests (Hawaii), and agricultural soils (South Africa and Uruguay)[23].

Eleven HP sequences formed a single cluster closely related to three IntI sequences recovered from heavy-metal-contaminated mine tailings[9]. All these sequences formed a clade with an environmental class 6 IntI[20].

Sequences HP83 and NC30 formed a clade with a class 8 IntI[20] and with integrases recovered from different environments, including sediments from the Arctic Ocean[23] and a gas-hydrate-bearing core from Vancouver Island (Canada), among others[22].

Sequences HP4, HP8, and HP154 clustered as a single group, separated from the other integrases identified in this study.

4 Discussion

A tree based on the amino acid sequences of IntI proteins deduced from all the genomes recorded in the DOE JGI was constructed to investigate the diversity of these proteins (Supplementary Figure 1S). Our results show that from a phylogenetic perspective, the known classes of IntI are not clustered as defined groups. Further support for this idea is given by the phylogenetic tree constructed from environmental IntI sequences (Figure 4). Some of the classes are phylogenetically relatively close, but large groups of sequences do not share a clear ancestor with the known

Figure 4 Phylogenetic tree of IntI sequences obtained from HP and NC libraries and best BLAST hits selected from Genbank and from IMG/M database. The tree was constructed using web server: http://phylogeny.fr[18] with PhyML software and default options. Sequences from HP library are indicated in blue and those from NC library in red. Sequences from known IntI class 1 to class 8 are shown in pink and IntI sequences from some isolates, previously found in the same area, are shown in green.

classes. It is important to note that our analysis is based on an exclusively phylogenetic perspective, and other factors, such as sequence signatures, specific functions, or other relevant traits, may be used for sequence classification.

A specific primer pair for class 1 *intI* genes was used to amplify the genes from the total DNAs from two sites on Fildes Peninsula. We thus isolated the amplicons, used to construct the two clone libraries, HP and NC. The sampling locations were selected partly because of their different distances from scientific stations. Ninety-eight high-quality sequences were analyzed, including 62 from HP and 36 from NC. In the HP library, ca. 68% of the sequences were identical to those previously isolated from the same area with culture-dependent methods[11]. These class 1 *intI*-carrying isolates, from the genus *Enterobacter*, were isolated from different sites across Fildes Peninsula. The origin of these *intI1* genes is unknown, although the analysis of those enterobacterial isolates, which were resistant to Amp, Tmp, and Str[11], suggested that animals or humans could have been involved in the dissemination of these *intI* class 1 integrons[25-26].

Without considering the bias imposed by PCR, we do not know whether these 42 sequences were amplified from different hosts, from a single type of organism, or from siblings of a single strain that was relatively abundant in the sample. Additional studies are required to determine whether these integrons were introduced from anthropogenic, bird, or

mammalian sources.

Similar class 1 *intI* sequences were not recovered from the NC library. This sample was collected at a site located on the coast, on a cliff, so presumably it was not easily accessible to mammals, although presumably it would have been visited by birds and humans.

In contrast, the other 56 sequences from the HP and NC sites displayed much more diversity. The sequences also clustered into groups that were clearly separated from those isolated in a metagenomic analysis on King George Island. However, some of them were similar to other environmental integrase sequences isolated at other sites outside Antarctica, and some sequences formed three groups clearly separate from other previously described integrase sequences. A study of the genetic organization of the newly predicted integrons containing these integrase genes might provide new insight into the evolution of these elements and their bacterial hosts.

Acknowledgments We thank Dr. Paul R. Gill for his insightful discussions and the technical reading of the manuscript. This work was supported by Instituto Antártico Uruguayo, ANII (Agencia Nacional de Investigación e Innovación), PEDECIBA-Biología (Programa de Ciencias Básicas) and Fondo Conjunto de Cooperación México-Uruguay (Uru/13/001).

References

1 Hall R M, Collis C M. Antibiotic resistance in gram-negative bacteria: the role of gene cassettes and integrons. Drug Resist Updates, 1998, 1: 109–119
2 Mazel D, Davies J. Antibiotic resistance in microbes. Cell Mol Life Sci, 1999, 56: 742–754
3 Rowe-Magnus D A, Mazel D. Integrons: natural tools for bacterial genome evolution. Curr Opin Microbiol, 2001, 4: 565–569
4 Hall R M. Integrons and gene cassettes: hotspots of diversity in bacterial genomes. Ann N.Y. Acad Sci, 2012, 1267: 71–78, doi: 10.1111/j.1749-6632.2012.06588.x
5 Gillings M R. Integrons: past, present and future. Microbiol Mol Biol Rev, 2014, 78(2): 257–277, doi:10.1128/MMBR.00056-13
6 Boucher Y, Labbate M, Koenig J E, et al. Integrons: mobilizable platforms that promote genetic diversity in bacteria. Trends Microbiol, 2007, 15(7): 301–309, doi:10.1016/j.tim.2007.05.004
7 Gillings M R, Krishnan S, Worden P J, et al. Recovery of diverse genes for class 1 integron-integrases from environmental DNA samples. FEMS Microbiol Lett, 2008, 287: 56–62, doi: 10.1111/j.1574-6968.2008.01291.x
8 Gillings M, Boucher Y, Labbate M, et al. The evolution of class 1 integrons and the rise of antibiotic resistance. J Bacteriol, 2008, 190: 5095–5100
9 Nemergut D R, Martin A P, Schmidt S K. Integron diversity in heavy-metal-contaminated mine tailings and inferences about integron evolution. Appl Environ Microbiol, 2004, 70: 1160–1168
10 Stokes H W, Holmes A J, Nield B, et al. Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA. Appl Environ Microbiol, 2001, 67(11): 5240–5246
11 Antelo V B. Mecanismos de transferencia horizontal genética asociados a la degradación de gasoil y de resistencia a antibióticos en bacterias heterótrofas de suelos de la Antártida. 2012, Tesis de Maestría. PEDECIBA. UdelaR
12 Mazel D, Dychinco B, Webb V A, et al. Antibiotic resistance in the ECOR collection: integrons and identification of a novel *aad* gene. Antimicrob Agents Chemother, 2000, 44: 1568–1574
13 Ploy M C, Courvalin P, Lambert T. Characterization of In40 of *Enterobacter aerogenes* BM2688, a class 1 integron with two new gene cassettes, *cmlA2* and *qacF*. Antimicrob Agents Chemother, 1998, 42(10): 2557–2563
14 Zhang Z, Schwartz S, Wagner L, et al. A greedy algorithm for aligning DNA sequences. J Comput Biol, 2000, 7(1–2): 203–14
15 Nunes-Dubuy S, Kwon H J, Tirumalai R S, et al. Similarities and differences among 105 members of the Int family of site-specific recombinases. Nucleic Acids Res, 1998, 26(2): 391–406
16 Edgard R C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res, 2004, 32(5): 1792–1797
17 Waterhouse A M, Procter J B, Martin D M A, et al. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. Bioinformatics, 2009, 25(9):1189–1191
18 Dereeper A, Guignon V, Blanc G, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res, 2008, Jul 1: 36 (Web Server issue):W465-9. Epub 2008 Apr 19
19 Carattoli A, Aschbacher R, March A, et al. Complete nucleotide sequence of the IncN plasmid pKOX105 encoding VIM-1, QnrS1 and SHV-12 proteins in Enterobacteriaceae from Bolzano, Italy compared with IncN plasmids encoding KPC enzymes in the USA. J Antimicrob Chemother 2010, 65(10): 2070–2075
20 Nield B S, Holmes A J, Gillings M R, et al. Recovery of new integron classes from environmental DNA. FEMS Microbiol Lett, 2001: 195: 59–65
21 Elsaied H, Stokes H W, Nakamura T, et al. Novel and diverse integron integrase genes and integron-like genes cassettes are prevalent in deep-sea hydrothermal vents. Environ Microbiol, 2007, 9(9): 2298–2312
22 Elsaied H, Stokes H W, Yoshioka H, et al. Novel integrons and gene cassettes from a cascadian submarine gas-hydrate-bearing core. FEMS Microbiol Ecol, 2013, 87: 343–356
23 Rodríguez-Minguela C M, Apajalahti J H A, Chai B et al. Worldwide prevalence of class 2 integrase outside the clinical settings is associated with human impact. Appl Environ Microb, 2009, 75(15): 5100–5110
24 Messier N, Roy P H. Integron integrases posses a unique additional domain necessary for activity. J Bacteriol, 2001, 183(22): 6699–6706
25 Saikia S, Saikia D, Ramteke P W. Use of microbes from seabird faeces to evaluate heavy metal contamination in Antarctic region. Appl Ecol Env Res, 2008, 6(3): 21–31
26 Miller R V, Gammon K, Day M J. Antibiotic resistance among bacteria isolated from seawater and penguin fecal samples collected near Palmer Station, Antarctica. Can J Microbiol, 2009, 55: 37–45

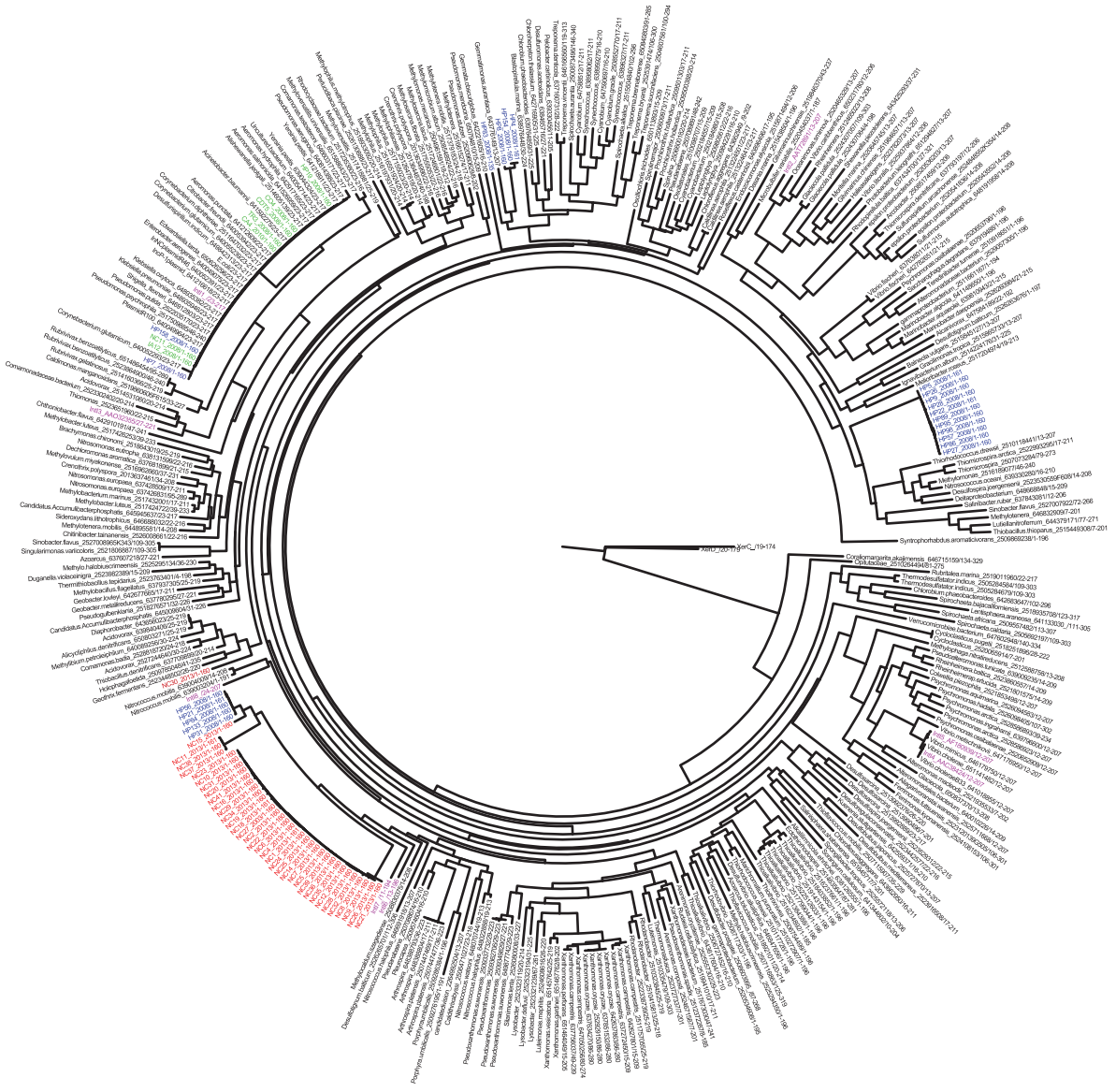

Figure 1S Phylogenetic tree estimated by maximun likelihood method, including IntI amino acid sequences from HP and NC libraries, representative class 1 to class 8 integrases and IntI sequences identified in complete genomes of DOE JGI database. Color codes are the same as those of Figure 4.