JOURNAL OF BACTERIOLOGY, July 2008, p. 4959–4970 Vol. 190, No. 14
0021-9193/08/$08.00+0 doi:10.1128/JB.00229-08

A Family of Insertion Sequences That Impacts Integrons by Specific Targeting of Gene Cassette Recombination Sites, the IS*1111-attC* Group▽†

Sasha G. Tetu and Andrew J. Holmes*

School of Molecular and Microbial Biosciences, University of Sydney, Sydney, New South Wales, Australia

Received 15 February 2008/Accepted 6 May 2008

Integrons facilitate the evolution of complex phenotypes by physical and transcriptional linkage of genes. They can be categorized as chromosomal integrons (CIs) or mobile resistance integrons (MRIs). The significance of MRIs for the problem of multiple antibiotic resistance is well established. CIs are more widespread, but their only demonstrated significance is as a reservoir of gene cassettes for MRIs. In characterizing CIs associated with *Pseudomonas*, we discovered a subfamily of insertion sequences, termed the IS*1111-attC* group, that insert into the recombination sites of gene cassettes (*attC* site) by site-specific recombination. IS*1111-attC* elements appear to have recently spread from *Pseudomonas* species to clinical class 1 integrons. Such elements are expected to significantly impact integrons. To explore this further, we examined CIs in 24 strains representing multiple levels of evolutionary divergence within the genus *Pseudomonas*. Cassette arrays frequently had a degenerated "footprint" of an IS*1111-attC* group element at their terminus and in three cases were occupied by multiple functional IS*1111-attC* elements. Within *Pseudomonas* spp. the IS-integron interaction appears to follow an evolutionarily rapid cycle of infection, expansion, and extinction. The final outcome is extinction of the IS element and modification of the right-hand boundary of the integron. This system represents an unusual example of convergent evolution whereby heterologous families of site-specific recombinases of distinct genetic elements have adopted the same target site. The interactions described here represent a model for evolutionary processes that offer insights to a number of aspects of the biology of integrons and other mosaic genetic elements.

Mobile genetic elements (MGEs) facilitate the movement of DNA within or between genomes and have significant evolutionary impact (35). The evolutionary significance of MGEs can be broadly summarized as gene rearrangement (intragenomic gene transfer), gene dissemination (horizontal gene transfer), and altered expression patterns (pleiotropic effects due to knockout or regulatory changes as a consequence of insertion). MGEs found in bacteria include transposons, insertion sequences (IS), gene cassettes, and bacteriophages, among others. These elements are often found together as complex mosaic structures in genomic islands (53, 55), where their interaction may give rise to higher-order properties. Such genomic islands have tremendous biological impact, including xenobiotic degradation (27), pathogenicity (20), multiple antibiotic resistance (17), and ecological niche differentiation (11). However, our conceptual understanding of how they arise remains limited. It is evident that the capacity to form mosaic structures by "combinatorial genetic evolution" would be advantageous to biological entities at many scales of complexity, but particularly to those that do not autonomously result in a phenotype. The classic example of this is the capture and "activation" of gene cassettes by class 1 integrons (21, 34). Here a specific process (integron capture of gene cassettes) gives rise to assemblages of diverse genes that are often borne on simply organized mobile elements. At the other end of the spectrum, large genomic islands that encode functions, such as pathogenicity, symbiosis, and degradative pathways, are also definable genetic entities but have a far more complex structure than class 1 integrons. The underlying processes governing the formation of such complex islands are unclear but may also involve specific interactions between different genetic entities. Here we have investigated interaction between integrons and a family of insertion sequences.

Insertion sequences are the smallest autonomously mobile elements (generally <2.5 kb). The definition of what constitutes an IS remains relatively broad, and the features that unite them are simple organization, the capacity to insert at multiple sites in a target DNA molecule (degree of specificity varies), and a phenotypically cryptic nature (encoded proteins have no function other than mobility). Most characterized IS can be divided into about 20 families (33, 48). IS can exert impact directly, either through insertional inactivation of genes or influencing the expression of adjacent genes (1, 28, 47, 50). Of greater interest, however, are impacts that emerge through their interaction with other elements. Acting in concert, IS can further alter bacterial genomes by forming compound transposons which can excise, flip, or move, leading to gene deletion, inversion, or rearrangement. By interacting with other types of genetic elements, IS can have even broader consequences, and they are clearly involved in arranging genes into complex modular assemblies that have higher-order properties (48). Here we focus on their interaction with integrons.

The interaction between integrons and mobile elements un-

* Corresponding author. Mailing address: School of Molecular and Microbial Biosciences, University of Sydney, Sydney 2006, New South Wales, Australia. Phone: 11 (612) 9351-2530. Fax: 11 (612) 9351-4571. E-mail: a.holmes@usyd.edu.au.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 16 May 2008.

TABLE 1. PCR primers

Primer	Sequence (5′–3′)	Orientation[a]	Location (position)
ST87	ATATGGACTCTCCCCACAAG	R	ISPst6 (1351–1370)
ST88	ACVGGCGAGTCCTGCARCGG	F	ISPst6 (1176–1196)
ST89	GAACGCTTGGCCCCYGCTGC	F	ISPst6 (1235–1254)
ST117	TAATGGACTCTCCCCGCACC	F	ISPst6 (6–25)
ST123	CGGCAACCTGGTAAACGGAC	R	ISPst6 (102–121)
ST148	GGTGAGAAGCGAGACTGTCG	F	ISPst6 (1155–1174)
ST139	AATATGACACCAAGCACTCC	F	ATCC 17587 cassette array
ST96	TTTCACGAGTATTCCCGAGG	R	ATCC 17587 cassette array
pgemF	CCGACGTCGCATGCTCC	F	pGEM-T Easy vector
pgemR	CTCCCATATGGTCGACCTG	R	pGEM-T Easy vector

[a] R, reverse; F, forward.

derlies many questions regarding integron biology. Integrons were discovered through their role in the assembly of antibiotic resistance genes, and there are now approximately 100 distinct integron classes known (4). Five classes are known to confer antibiotic resistance phenotypes and typically occur as part of complex mosaic elements with transposons and plasmids. In stark contrast, the vast majority of integron classes occur on chromosomes and have no obvious phenotype. Since all integrons thus far experimentally tested have been shown to have similar functions (3, 12, 14, 15, 24, 26, 30), the implication is that there are strong biological consequences from interaction between integrons and mobile elements. This has led many workers to distinguish between mobile resistance integrons (MRIs) and chromosomal integrons (CIs). (For an exchange of opinions as to the significance of differences between integrons, see references 22 and 34.)

Interactions between integrons and IS elements are a likely forerunner to incorporation of either the integron or assemblies of gene cassettes into larger mosaic structures with higher-order biological properties (7, 38). They potentially explain the origin of MRI involvement in multiple antibiotic resistance and could also explain various aspects of the biology of CIs. Of the various IS reported in association with integrons, one of the most interesting is ISPa21 (37). Since this element is a member of an IS grouping (IS*1111* section within the IS*110* family) considered to typically show target site specificity, it raises the prospect of specific IS-integron interaction (36). We discovered another member of this group, ISPst6, in a *Pseudomonas stutzeri* strain. Interestingly, ISPst6 was within an integron that is a member of a family of CIs strongly associated with *Pseudomonas* species (26, 54). Such genus-associated CI families are predicted to have been continuously present and strongly influenced the evolution of their respective genera (18, 44). The *P. stutzeri* species complex is a well-studied model for bacterial evolutionary genetics, and numerous distinct genomovars have been described. These subtaxa are genomically distinct subgroups within the species complex with well-defined evolutionary relationship (2, 19, 29, 43). Consequently, sampling *P. stutzeri* genomovars provides a unique opportunity to explore the outcomes of interaction between chromosomal integrons and insertion sequences over evolutionary significant periods of time. We confirmed that ISPst6 and relatives specifically target cassette arrays and show that these interactions have influenced the fate of cassette arrays in pseudomonads. This direct interaction between distinct genetic entities is the result of convergent evolution between distinct recombinases for the same target site. We postulate this may represent a general principle contributing to assembly of more complex mosaic elements.

MATERIALS AND METHODS

Bacterial strains and growth conditions. A total of 24 strains representative of diversity within the *P. stutzeri* species complex were screened for the presence of ISPst6 by Southern hybridization. These included representatives of genomovars 1, 2, 4, 5, 7, and 8 and strains of *P. mendocina*, *P. straminea*, *P. aeruginosa*, *P. fluorescens*, and *P. putida* as outgroups. A conjugative plasmid carrying a class 1 integron (R388) and an *Escherichia coli* strain (JM109) that does not contain an integron were also included in Southern analyses. Details of strains used are given in the supplemental material. All strains were routinely grown in Luria-Bertani (LB) broth or agar (1.5%). When required, 100 μg of ampicillin and/or 12.5 μg of chloramphenicol per ml was added. All *Pseudomonas* strains were grown at 30°C, and all *E. coli* strains were grown at 37°C.

DNA manipulations. High-molecular-weight DNA was extracted from all strains by using a CTAB (cetyltrimethylammonium bromide)-phenol-chloroform method, with a preliminary lysozyme and RNase A treatment (46). Plasmid preparations were carried out according to a standard alkaline lysis protocol (46). Restriction enzymes were used according to the manufacturer's instructions (New England Biolabs). PCR amplifications were carried out by using *Taq* DNA polymerase (New England Biolabs) in 25-μl reactions, unless otherwise stated. Primer sequences are given in Table 1. DNA purification was carried out with Qiaquick PCR purification or gel extraction kits (Qiagen). Automated sequencing was carried out by the Australian Genome Research Facility (Brisbane, Australia) on an ABI 3730xl 96-capillary automated DNA sequencers using BigDye terminator chemistry.

Detection of ISPst6 minicircles. PCR primers ST89 and ST123 were used in 50-μl reactions with DNA extracted from each of the genomovar 2 strains. They enable recovery of a 336-bp section of ISPst6 containing the junction of left and right ends formed when this element is in its circular form. PCR products of the correct size were excised from 2% Tris-acetate-EDTA gels, cleaned, and cloned into pGEM-T Easy vector (Promega) according to standard procedures and sequenced with primers in the plasmid backbone.

Promoter activity assays. Three sections of ISPst6 from ATCC 17595 were tested for promoter activity: the 336-bp ISPst6 junction region (this section contains 49 bp of the transposase gene and the 287 bp upstream in circular form, 72 bp from left-hand side [LHS] and 215 bp from right-hand side [RHS]; ST89 to ST123); a subsection of this section containing only P_{IRL} (49 bp of transposase gene and 67 bp from LHS [ST123 to ST117]); and this same 116-bp region in the opposite direction, which would affect expression of downstream sequence (P_{out}). All PCR products were amplified in 200-μl reactions with Phusion polymerase (New England Biolabs) and ligated into the SmaI site of the promoter analysis vector pKK232-8 (6), upstream of the promoterless chloramphenicol acetyltransferase gene to create pUS84 ($P_{junc+IRL}$), pUS90 (P_{IRL}), and pUS92 (P_{out}). The *E. coli lacZ* gene promoter was similarly cloned into this vector in forward and reverse orientation as positive and negative controls (pUS85 and pUS86). The insert regions of all constructs were confirmed by sequencing. The chloramphenicol MIC was determined by serial twofold dilution in Mueller-Hinton agar (BD), following the guidelines of the National Committee for Clinical Laboratory Standards. Inocula of 10^4 CFU per spot were applied with an

antibiotic sensitivity replicator. After 16 h of incubation at 37°C, the chloramphenicol MIC was defined as the lowest concentration that prevented the visible growth of colonies on the plate.

Transcript mapping. For pUS84, where more than one putative promoter was present within the construct, transcript start points were mapped by using the method of Lloyd et al. (31) with minor alterations to the reverse transcription (RT) reaction. Briefly, RNA was extracted from log-phase *E. coli* cells carrying these constructs by using Qiagen RNeasy mini-columns according to the manufacturer's standard protocol (Qiagen). Approximately 5 mg of RNA was used in RT-PCR with a FAM-labeled reverse primer situated in the chloramphenicol acetyltransferase gene (see reference 31 for the sequence) using 400 U of Superscript III (Invitrogen) and a 3-h incubation at 50°C. The products were treated with RNase H (New England Biolabs) for 20 min and then ethanol precipitated and resuspended in 2 μl of distilled water. The size of the fluorescently labeled RT products was directly determined by using an ABI Prism 377 GeneScan analyzer (Macquarie University Sequencing Facility, Sydney, Australia) and compared to the insert sequence to determine the transcript start point.

Southern hybridizations. PstI and PvuII digests of approximately 5 μg of genomic DNA of each strain were separated via electrophoresis in a 0.8% agarose gel, visualized, and then transferred to a nylon membrane (Hybond N+; Amersham) using capillary transfer under alkaline conditions. Probes for Southern hybridization were labeled with digoxigenin-6-dUTP by PCR according to the Roche protocol. All hybridizations were performed in DIG Easy-Hyb buffer (Roche) containing 25 ng of digoxigenin-labeled probe/ml, followed by incubation overnight at 55°C. Detection and visualization of hybridization products was performed by using a Roche DIG detection kit according to the manufacturer's instructions, with washing conditions of 68°C in 0.5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–1% sodium dodecyl sulfate. Two IS probes were tested; the first consisted of 1,365 bp of the element amplified from ATCC 14405 genomic DNA with the ST87-ST117 primer pair. The second probe consisted of 194 bp from the 3′ end of ISPst6 amplified from ATCC 14405 with the primers ST87 and ST88; this probe was designed to bind to only one fragment of ISPst6 digested with PvuII, so that the copy number could be estimated.

Construction of *attC* containing plasmid as an ISPst6 "trap." A 219-bp section of *P. stutzeri* ATCC 17587 (gv2), containing a typical *Pseudomonas*-type *attC* (also known as the PAR element) (54), with a small amount of flanking sequence was amplified by PCR (ST139 and ST96) and then cloned into pGEM-T Easy vector (Promega) according to the manufacturer's instructions to make pST001. This construct was transformed into electrocompetent *E. coli* Epi300 (Epicenter) cells, together with a large insert Fosmid clone carrying multiple copies of ISPst6 within ATCC 14405 array 1. Cells were plated out on LB-ampicillin-chloramphenicol agar. A small number of clones containing both Fosmid and pSGT001 were recovered; one was chosen for subsequent work. Cells were screened for transposition events via PCR with one primer in pGEM-T backbone and one in ISPst6 (pgemR-ST123). Once PCR products indicated transposition had occurred, dilutions were plated onto LB-ampicillin-chloramphenicol agar to screen individual colonies. After 1 week at 37°C followed by 3 weeks at 4°C, individual colonies gave PCR products spanning the IS-plasmid boundary; these products were sequenced confirming the presence of ISPst6 in pST001. PCR with primers spanning the other boundary (pgemF-ST148) did not give a product with either the broth or the single colony template.

Sequence analysis. The nucleotide sequence of the putative IS were used to search the National Centre for Biotechnology Information BLAST website (www.ncbi.nlm.nih.gov:80/BLAST) utilizing both BLASTN and BLASTX programs. In addition to searching the default NCBI database, searches were made of the env_nt and unfinished genome databases. The analysis function in the IS finder database (http://www-is.biotoul.fr/is.html) was used to search for characterized IS with homology to the ISPst6 element and its putative transposase. The ISPst6 sequence in integrated, and the circular form was screened for potential promoter elements by using the Neural Network Promoter Prediction service for prokaryotes (41), found on the Berkeley *Drosophila* Genome Project website (http://www.fruitfly.org/seq_tools/promoter.html), with a minimum score of 0.7 for the promoter prediction. Vector NTI (version 10.0, Invitrogen) was used to view and analyze sequence information.

Multiple sequence alignment was performed using the default settings of the AlignX module of the Vector NTI software package, which is based on the CLUSTAL W algorithm. Alignments were manually optimized in the GeneDoc alignment editor. Phylogenetic analyses were performed by using programs from the Phylip package through Biomanager. Trees constructed from protein alignments were based on a restricted alignment that removed positions affected by frameshifts in ISUnCu1 and ISAvX1. Trees constructed from nucleotide sequence alignments included these corresponding regions.

GenBank accession numbers. Sequences generated in the present study have been deposited in GenBank under accession numbers EF439838, EF648209, EF648210, EF648212, EF648211, EF648213, EF648215, EF648214, and EF648216. Other sequences cited here are given in Table 2.

RESULTS

ISPst6 was discovered in the gene cassette array of an integron in *P. stutzeri* strain ATCC 14405. The sequence of ISPst6 has been deposited in the GenBank database (EF439838). ISPst6 is 1,371 bp in length, has a G+C content of 59%, and is flanked by subterminal inverted repeats. A putative ribosome-binding site (5′-GGAGAGG-3′) and putative sigma 70 promoter sequence (5′-TCTCCC-N16-TATAAC-3′) can be seen upstream of a single open reading frame (ORF; 347 amino acids). The translated amino acid sequence of this ORF contains a conserved domain (spanning amino acids 183 to 296) seen in the pfam02371-transposase 20 family (Conserved Domain database, National Center for Biotechnology Information). The insertion site of ISPst6 was within the *attC* site of a gene cassette. BLAST searches and sequence comparison of the putative transposase protein with other elements in the IS finder database (http://www-is.biotoul.fr/is.html) indicated that ISPst6 most likely belongs to the IS*1111* family.

ISPst6 is active and has all of the typical features of the IS*1111* family. IS*1111* family elements are typically comprised of a single ORF encoding a transposase flanked by relatively long noncoding sequences, and bounded by conserved subterminal 12- to 13-bp inverted repeats (consensus, ATGGACgC cCCC). In ISPst6 the noncoding regions were 53 and 240 bp at the left and right ends, respectively. The inverted repeat boundaries were precisely determined to be 5′-GGCTGTAA TGGACTCTCCC-3′ (LHS) and 5′-GGGAGAGTCCATAT A-3′ (RHS). The transposase of ISPst6 does not show a particularly close relationship to IS*1111A*, the archetypical member of the family (only 35% amino acid identity), but the percent identity to another previously described member of the family, ISPa21 (37) was 84%. Notably, ISPst6 possesses the conserved D-E-D-D motif (D_7 E_{51}, D_{88}, D_{91}) that has been demonstrated to be essential for activity in the Piv/MooV family of DNA recombinases (8, 52) and is inferred to be essential for IS*1111*, IS*110*/IS*492* family members (10, 36).

Previously studied members of the IS*1111* family have been shown to move via circular intermediates (36, 39). We looked for similar ISPst6 minicircles by PCR and sequenced products of the expected size from two independent *P. stutzeri* strains containing ISPst6 (ATCC 17587 and ATCC 17595). In each case the recovered ISPst6 sequence contained the expected junction of 10 bp between the two inverted repeats and no additional sequence, a finding consistent with amplification from a circular intermediate (Fig. 1.). Interestingly, another typical IS*1111* family feature was seen in this junction sequence: the formation of a putative sigma 70 promoter (36, 39). The region upstream of the transposase gene in the circular form contains two putative promoters, one formed in the junction region (P_{junc}) and one entirely at the LHS, expected to drive the expression of the gene in the integrated form (P_{IRL}). This section was recovered by PCR and tested for promoter activity in the promoter analysis vector pKK232-8. The pUS84 construct containing both putative transposase

TABLE 2. IS*1111-attC* subfamily elements

Database accession no.	IS[a]	Strain and genomic context	Insertion site (*attC* length in bp)[b]	IS length (bp)[b]	Available sequence***
EF439838 (gp:ABS31633)	**ISPst6**	*P. stutzeri* ATCC 14405, within array 1	BGC100 (77)	1,371	
EF439838 (gp:ABS31637)	**ISPst6**	*P. stutzeri* ATCC 14405, within array 1	BGC101 (76)	1,371	
EF439838	**ISPst6**	*P. stutzeri* ATCC 14405, within array 1	BGC099 (76)	1,368	
EF439838	**ISPst6**	*P. stutzeri* ATCC 14405, array 1	BGC102 (Inc)	Inc	
EF648209	**ISPst6**	*P. stutzeri* ATCC 17587, within array 1	BGC094 (75)	1,371	
EF648209	**ISPst6**	*P. stutzeri* ATCC 17587, array 1	BGC095 (Inc)	Inc	
EF648210	**ISPst6**	*P. stutzeri* ATCC 17595, within array 1	BGC096 (76)	1,371	
EF648210	**ISPst6**	*P. stutzeri* ATCC 17595, within array 1	BGC097 (Inc)	1,369	
EF648212	**ISPst6**	*P. stutzeri* ATCC 14405, within array 2	BGC103 (76)	1,371	
EF648212	ISPstX1	*P. stutzeri* ATCC 14405, array 2 terminus	BGC104 (Inc)	Degen	
EF648211	ISPstX2	*P. stutzeri* ATCC 17595, within array 2	BGC098 (76)	Degen	
EF648211	ISPstX5	*P. stutzeri* ATCC 17595, array 2 terminus	BGC106 (Inc)	Degen	
EF648213	ISPstX3	*P. stutzeri* RNAIII, array 1 terminus	BGC105 (Inc)	Degen	
EF648215	ISPstX4a	*P. stutzeri* ATCC 17593, array 1 terminus	Unidentified (NA)	Degen	
EF648214	ISPstX4b	*P. stutzeri* ATCC 17684, array 1 terminus	Unidentified (NA)	Degen	
EF648216	ISPmeX1	*P. mendocina* NW1 array 1 terminus	BGCXXX (Inc)	Degen	
AY920928 (gp:AAY43212)	**ISPa21**	*P. aeruginosa* DEJ, In109 (class 1)	bla_{GES-9} (110)	1,374	
AY920928 (gp:AAY43212)	**ISPa21**	*P. aeruginosa* DEJ, In109 (class 1)	*aacA7* (60)	1,374	
AY660529 (gp:AAT74616)	**ISPa21**	*P. aeruginosa* In163 (class 1)	*aadA7* (60)	1,375	
AM296017 (gp:CAL30105)	ISPaX1	*P. aeruginosa* InXX (class 1)	*aac(6)-II* (60)	1,376	
DQ522236 (gp:ABF70510)	ISPaX2	*P. aeruginosa* 257, InXX (class 1)	*aadB* (60)	1,378	
AY139602 (gp:AAN41438)	**ISUnCu1**	UnCul plasmid Psp7 InXX (class 1)	*aadA1* (60)	1,381	
DQ302723 (gp:ABC17634)	**ISUnCul**	*P. aeruginosa* InXX (class 1)	*aadA1* (60)	1,382	
AAAU03000002 (gp:EAM06815)	ISAvX1	*A. vinelandii* AvOP, chromosome	Unidentified (60)	Degen	
DQ393782 (gp:ABD58912)	ISPaX3	*P. aeruginosa* N2-11; InXX (class 1)	bla_{OXA-17} (117)	1,409	
DQ902344 (gp:ABI63580)	ISPaX3	*K. pneumoniae* ChaK36; InXX (class 1)	bla_{OXA-17} (117)	1,409	
AY038186	ISPalX1	*P. alcaligenes*; InPal55044 terminus	PAR33 (Inc)	Degen	
AACY01190092	ISenv2	Sargasso sea clone (SKAXM26TF)	Unidentified (Inc)	Inc	
AACY01163073	ISenv3	Sargasso sea clone (SHAAG73TR)	Unidentified (Inc)	Inc	

[a] IS names deposited in the IS finder database are in boldface. Names shown in a regular typeface are of elements that have not been deposited into this database because they are either nonfunctional or are identified in unpublished sequence depositions.

[b] Data are incomplete (Inc) in some instances or not applicable (NA) in others if no site was identified. The exact length of degenerate (Degen) elements is not always determinable. ISUnCu1 is probably nonfunctional due to frameshifts that introduce significant protein divergence.

[c] IS elements are shown in the conventional IS orientation. A dotted line indicates no sequence is available. A space indicates an alignment gap. An arrowhead indicates the immediately adjoining sequence is part of an *attC* site. A vertical line indicates the adjoining sequence is not related to gene cassettes in any way (e.g., chromosomal gene) and marks the terminus of the associated cassette array.

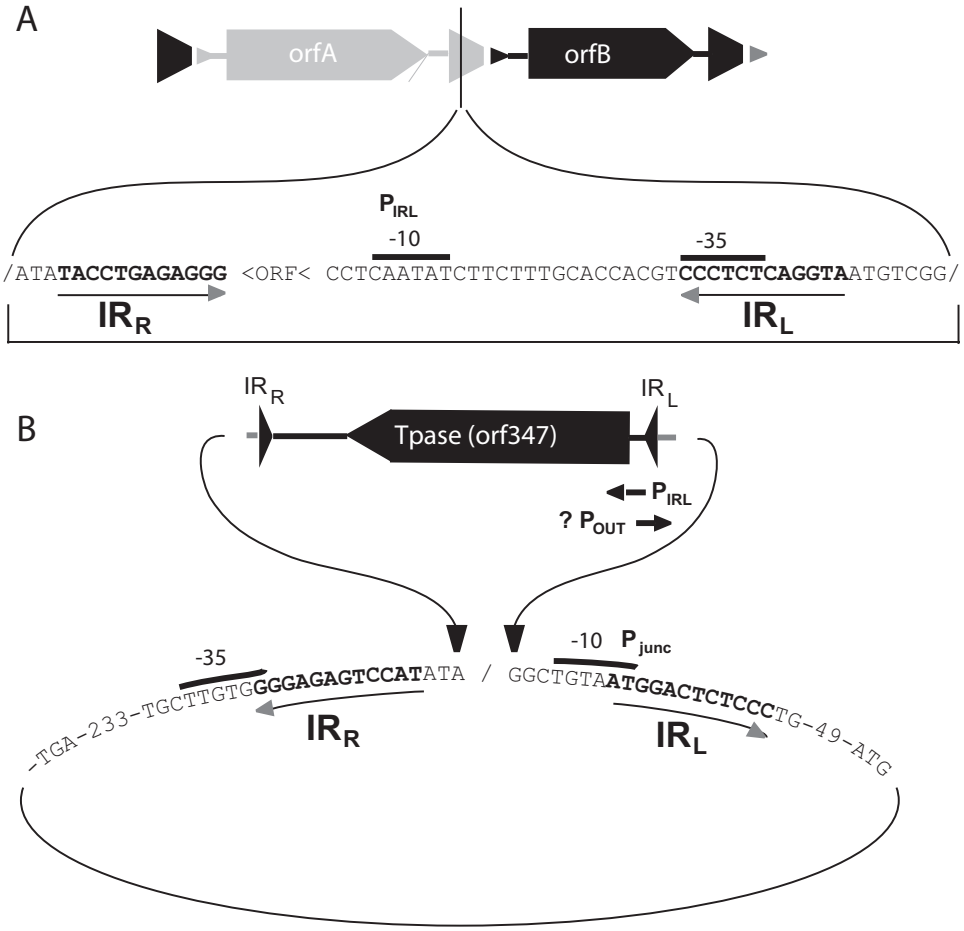

FIG. 1. Structural features of ISPst6 in integrated and circularized forms. (A) Schematic representation of a segment of an integron-assembled gene cassette array showing the inverted repeat and P_{IRL} sequences for an inserted IS*1111-attC* element. Note that the IS inserts in reverse orientation and the displayed sequence are for the sense strand of the IS but written 3′→5′ (i.e., as shown in panel B). Cassette-associated genes are represented by block arrows, and *attC* sites are represented by triangles. Components of the first gene cassette are colored gray, and components of the second gene cassette are colored black. The integron (not shown) is located upstream of the *orfA* gene cassette. (B) Linear diagram of ISPst6 as it occurs in cassette arrays. The location and orientation of the putative promoters P_{IRL} and P_{out}, which were tested for activity in pUS90 and pUS92 constructs, are also shown. The sequence of the junction formed between IR_R and IR_L upon excision of ISPst6 as a circle is shown below. The recombination junction between the left and right termini (slash), the subterminal nature of the IR sequences (boldface letters with underline arrows), and the putative −35 and −10 sequences of the junction promoter (overlined) are all indicated. The start of the major transcript mapped in the pUS84 construct lies within the sequence CCCTG at the end of IR_L.

promoters showed significant chloramphenicol resistance (MIC = 128 mg/liter). Transcript start-point mapping indicated that the junction promoter P_{junc} was responsible for the expression of the downstream gene in this construct (data not shown). A subsection containing only P_{IRL} was also tested to determine whether expression in the circular and integrated forms is likely to be different for this element. Neither the P_{IRL}-only construct (pUS90) nor another putative promoter P_{OUT} (pUS92) conferred significant chloramphenicol resistance.

Distribution and diversity of ISPst6 and relatives. To explore the distribution of ISPst6 within *P. stutzeri*, we used Southern hybridizations with ISPst6 fragments as probes to examine 19 strains, representing six genomovars within this species complex, and strains of another five pseudomonad species as outgroups. Four strains had one weakly hybridizing band, and the three genomovar 2 strains showed multiple strongly hybridizing sequences. We estimate that there are four copies of ISPst6 within ATCC 17587, four to eight copies within ATCC 17595, and six to nine copies within ATCC 14405 (Fig. 2.). We predict that there is a single ISPst6-like sequence in each of the two genomovar 1 strains (ATCC 17593 and ATCC 17684), a genomovar 7 strain (RNAIII) and *P. mendocina* NW1 (data not shown). Cloning and sequencing efforts targeting these hybridizing fragments recovered a total of nine IS elements and three IS-like fragments from the three genomovar 2 strains and one IS-like fragment from each of the other four probe-positive strains (the data are summarized in Table 2). We consider this a representative sample of the ISPst6-related sequences that occur within this strain collection. BLAST searches of public databases identified further examples of ISPst6-related sequences in other species, all of which were also associated with integrons, or sequences consistent with the presence of a gene cassette. The

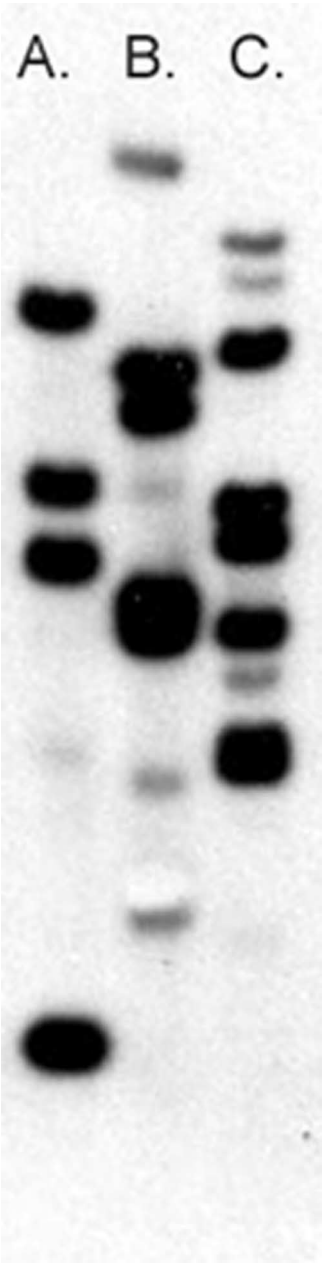

FIG. 2. Detection of multiple copies of ISPst6 in *P. stutzeri* genomovar 2 strains by Southern hybridization. PvuII digests of chromosomal DNA from genomovar 2 strains ATCC 17587, ATCC 17595, and ATCC 14405 are in lanes A, B, and C, respectively. The probe used was a 194-bp sequence from within the RH noncoding region of ISPst6. This probe is predicted to yield only one band per IS copy when DNA is digested with PvuII. A single very faint band was also seen in PvuII digests of the four other strains (see the text).

total data set comprised our 16 sequences and 13 database sequences.

Phylogenetic analyses of transposase sequences revealed a strongly supported clade (91%) within the IS*1111* family (subgroup 1 in Fig. 3A). A subgroup within this clade (57% bootstrap support) included all of the elements associated with *attC* sites. We refer to this latter set as the IS*1111-attC* subgroup. Our data set also included several partial sequences and IS-like fragments associated with *attC* sites that could not be included in the transposase tree. However, in all cases the available sequence included a portion of the right-hand noncoding region (Table 2). This feature enabled the construction of additional trees that included nearly all ISPst6-related sequences (Fig. 3B). These analyses supported the conclusion that all degenerated IS-like fragments associated with *attC* sites derive from members of the IS*1111-attC* subgroup. In the transposase sequence tree the deepest branching member of the IS*1111-attC* subgroup was ISPaX3 and the closest outgroup was ISHhaX1.

If elements with >95% sequence identity are considered as isoforms, then the 29 sequences in the data set represent 16 distinct members of the IS*1111-attC* subgroup (Table 2). Several points are worth noting. First, there is a strong association of the IS*1111-attC* group with the family *Pseudomonadaceae*. The only exception is on a plasmid-borne integron in *Klebsiella pneumoniae*. Second, there is a trend for near-identical IS*1111-attC* elements to occur in closely related contexts (equivalent arrays of closely related strains) and for divergent IS*1111-attC* elements to occur in more distantly related contexts (distinct arrays of the same strain and/or divergent strains). In the case of CI arrays there is also a trend for each cassette array to contain either multiple copies of intact near-identical IS*1111-attC* elements or a single, degenerated copy. This is consistent with independent acquisition of IS followed by rapid expansion and extinction of the IS lineages in each bacterial genome.

The IS*1111-attC* subgroup specifically target gene cassette *attC* sites. Target specificity is thought to be a general feature of the IS*1111* family (36). Although only ISPst6 and ISPa21 were observed in multiple sites, a good sample of target sites used by the IS*1111-attC* subgroup was obtained since collectively the 16 different IS*1111-attC* subgroup elements (or fragments thereof) were observed in at least 27 independent loci (Table 2). All intact elements were inserted within the recombination site of a gene cassette (*attC*) in inverse orientation with respect to the gene cassette array (Fig. 1). Our data set also included a number of partial sequences and degenerated sequences. In all cases where an IS boundary (IR_R or IR_L) was still identifiable the adjoining sequence was within an *attC* site. Two degenerated elements (ISPstX4a and ISPstX4b) had lost both IS boundaries. These elements were located within 300 bp of a cassette array and are likely to also derive from an ancestral element that inserted into an *attC* site. Significantly, a total of seven degenerated elements had lost the left boundary (IR_L), and in all cases this sequence fragment marked the boundary of the cassette array. The adjoining sequence was a chromosomal framework gene. The insertion sites of the most closely related outgroups of the IS*1111-attC* group (ISHhaX, ISPpu11, and ISPnaX) were also examined. None of these were located in the proximity of integron features (*intI* homolog or gene cassette array).

The *attC* sites of gene cassettes are unambiguously recognizable by their imperfect inverted repeat (pseudo-palindrome) structure with two characteristic "flipped-out" positions (21, 26, 32, 49). They are however highly variable in length and sequence. Length variation (range, 57 to 145 bp) arises due to one or more insertions to the common pseudo-palindromic backbone structure of 51 bp. Broad patterns in

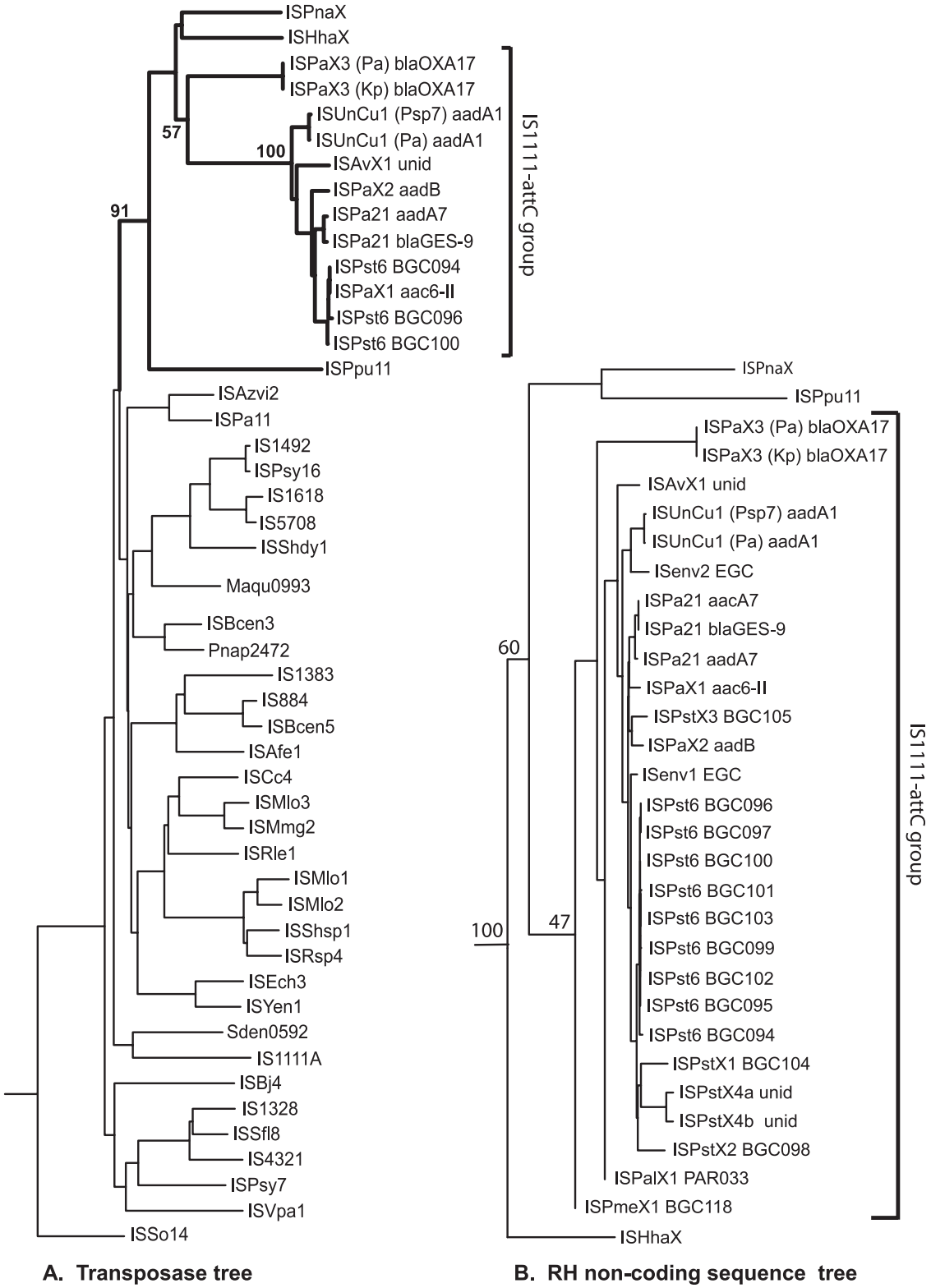

FIG. 3. Dendrograms showing the relationship of the IS*1111-attC* group to other members of the IS*1111* family. The numbers shown at the nodes are those from a bootstrap analysis of 100 replicates. Tree topologies with respect to the indicated groups were identical in parsimony and maximum-likelihood analyses. (A) Phylogeny based on the transposase sequence. All sequences associated with *attC* sites comprise a set (IS*1111-attC* group) within the strongly supported subgroup 1 (indicated by heavy lines). Note the predicted sequence of ISUnCu1 after correction for frameshifts was used. The tree was from an alignment of 308 amino acids. The tree was rooted using IS*110* as an outgroup. (B) Phylogeny based on an alignment of 305 bp from the right hand noncoding region of the IS. All IS-like fragments associated with *attC* sites cluster with sequences that were within the IS*1111-attC* group in transposase trees. The tree was rooted by using ISMlo1as an outgroup.

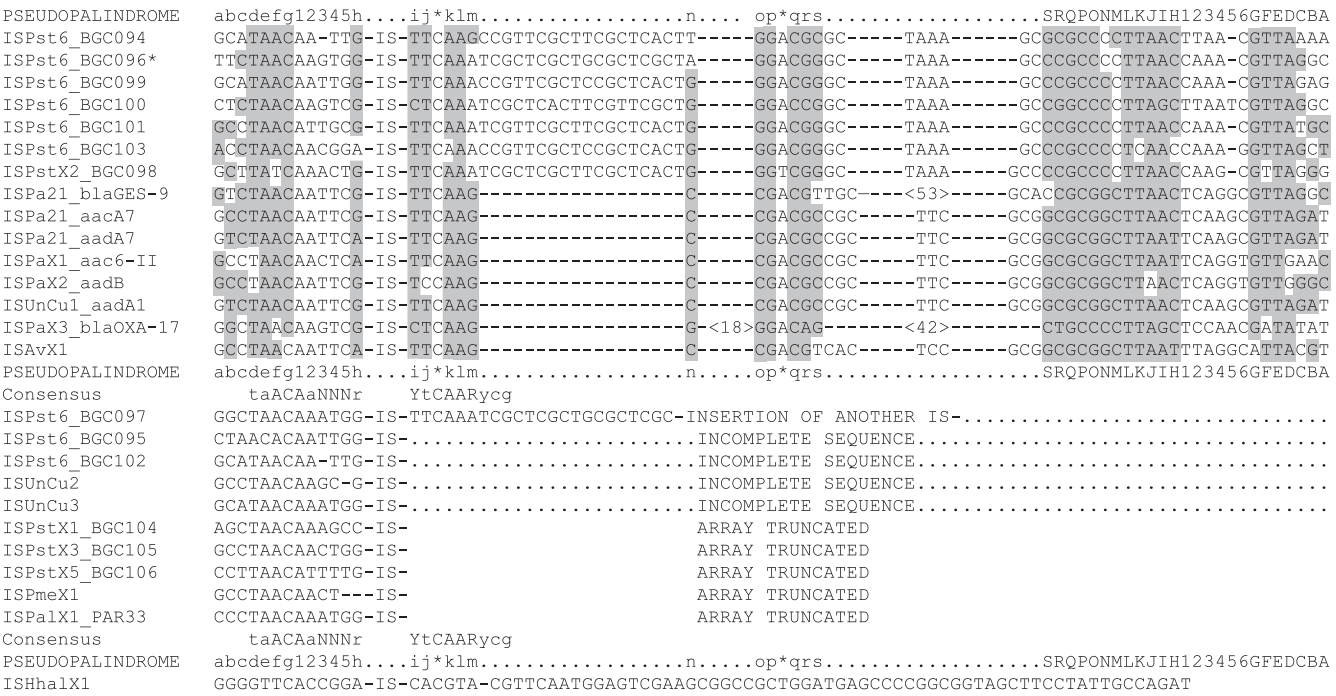

FIG. 4. Comparison of all available sequences harboring IS*1111-attC* elements (or fragments thereof) aligned according to secondary structure. The 51 conserved positions in the *attC* pseudopalindrome are indicated; upper- and lowercase letters are complementary. Agreement with the inverted repeat predicted by the pseudopalindrome model is indicated by shading (note that positions F-A derive from the next *attC* site in the array, so mismatches between a-f and F-A are expected [see also Fig. 1]). Partial sequences inferred to derive from *attC* sites are also shown. The sequence surrounding the most closely related nonmember of IS*1111-attC* (ISHhaX1) is shown below the alignment to illustrate the lack of relationship to *attC* sites. Dots represent unavailable sequence, dashes represent alignment gaps, and a space represents the absence of a gene cassette-related sequence (array truncation). The numbers shown in the bla_{GES-9} and bla_{OXA-17} *attC* sites represent additional nucleotides not shown.

attC diversity can be recognized according to the location (left, central, right, none, etc.) and the sequence of these insertions. Of the 25 *attC* sequences harboring IS, 10 were only partial sequences. The 15 sites for which the complete sequence could be inferred include examples with left and central insertions of various length and sequence. We categorize them as six *aadB*-like sites (~60 bp, with a 9-bp insertion between positions s and S), seven *Pseudomonas*-like sites (~76 bp, with ~18-bp insertion at m-n and an 8-bp insertion at s-S), one blaGES-like site (~110 bp with a 59-bp insertion at s-S), and one blaOXA-like site (111 bp with a 18-bp insertion at n-o and a 42-bp insertion at s-S). The pairwise divergence between two members of different subcategories (e.g., *aadB*-like and *Pseudomonas*-like) is equivalent to the extremes of divergence seen for all *attC* sites. In all cases the IS element had precisely inserted into an equivalent position within the *attC* site, corresponding to between positions h and i (Fig. 4). In the five cases where we could not reconstruct the complete *attC* site, we could identify an IS-*attC* junction consistent with position h (unambiguous identification of position h is not always possible from a partial sequence [error of ±1 nucleotide]). Position h was also the immediately adjacent nucleotide to five of the degenerate IS-like fragments. The sequences immediately surrounding ISHhaX, ISPpu11, and ISPnaX did not show the diagnostic secondary structural features of *attC* sites, nor did they show local primary sequence conservation (Fig. 4).

To confirm the precise nature of insertion, an "IS-trap" experiment was performed. A single *Pseudomonas*-like *attC* (BGC093) was amplified and cloned from *P. stutzeri* strain ATCC 17587 to give pST001. This plasmid was transformed into an *E. coli* strain carrying a fosmid harboring the entire ATCC 14405 cassette array 1, with its multiple copies of ISPst6. A PCR product consistent with insertion into the test *attC* site was seen in a population of cells after 1 day at 37°C. Sequencing of this amplicon confirmed that it included the junction between IR_R and the test *attC* site of pST001 with the crossover occurring between positions h and i. Interestingly, we could recover single colonies yielding this PCR product but were unable to recover either a PCR product of the IR_L junction or a plasmid consistent with pST001 containing a single intact copy of ISPst6. This may reflect a recombination mechanism that proceeds via an intermediate structure requiring additional, possibly host-specific, factors for its resolution (16).

DISCUSSION

Integrons may not be equally susceptible to IS*1111-attC* element interaction. Interaction will be influenced by both the IS activity and the size and composition of an integrons cassette array. Factors controlling the size and composition of cassette arrays are not fully understood but appear to reflect integron location (chromosomal versus mobile) and cassette properties (51). Cassette arrays typically encompass *attC* heterogeneity, and there is a tendency for this to be much greater

in MRI arrays (<30% conservation) than in CI arrays (~70 to 80% conservation in most characterized examples) (22). Therefore, the ability of IS*1111-attC* elements to exploit integrons, and particularly MRIs, will be dependent on their activity in the host bacterial species and their ability to tolerate *attC* sequence variation.

Integron integrases (IntIs) appear to tolerate *attC* variation by recognition of a specific structure rather than a specific sequence (5, 32). All known *attC* sites form a similar stem-loop with conserved structure and extrahelical bases (25, 49), and it is proposed that the key to IntI-mediated recombination is a DNA template formed by self-pairing of a uniplex DNA molecule. This mechanism is consistent with integron biology since the majority of DNA is acquired in single-stranded form (34). IS elements also seem to tolerate remarkable levels of *attC* variation, but the notion of a single-stranded DNA target is more difficult to reconcile with IS biology. ISPst6 was found in nine distinct target sites that showed up to 23% sequence divergence, ISPa21 was found in three target sites (encompassing 70% divergence), and collectively the target sites of the family encompassed *attC* site variation equivalent to that of IntIs. It is probable that all IS*1111-attC* elements are capable of specifically inserting into a wide range of *attC* sites. However, they may only be active under specific conditions since their distribution was limited and we could not demonstrate full function in *E. coli*.

Impact of IS*1111-attC* elements on integron-related phenotypes. Integrons depend on a host replicon for survival. Where an integron's maintenance in its host replicon is dependent on selective advantage conferred by integron-related phenotypes, IS insertion in gene cassettes is potentially detrimental. Integron-related phenotypes result from their ability to capture, and then direct the expression of, cassette-associated genes through their P_C promoter. Integrons can coexpress multiple cassette-associated genes via readthrough transcription from P_C to at least some extent (13), and it is therefore useful to distinguish between simple and complex integron-related phenotypes. Simple phenotypes are those that result from activation of a single cassette-encoded gene. The direct interaction between IS*1111-attC* and gene cassettes is predicted to have minimal impact on simple phenotypes. By specifically targeting the *attC* site, IS*1111-attC* elements avoid disruption of the ORF and any direct interference with its transcription. This elegant strategy is reminiscent of the interaction between group II retroelements and terminators (42), and it is noteworthy that a subset of group II retroelements also targets *attC* sites (9, 40).

Complex phenotypes are those that are encoded by multiple genes. These could be additive, as in the case of multiple drug resistance, or emergent as in the case of multicomponent enzymes or degradative pathways. The expression of complex phenotypes is dependent on the simultaneous expression of cassette genes (13), and IS insertion is highly likely to uncouple P_C-directed expression of downstream gene cassettes. An IS could minimize this impact by restoring simultaneous expression if it carried outward-facing promoters (1, 28, 47, 50). The presence of such a promoter was proposed in ISPa21, although no confirmation of expression of the downstream aminoglycoside resistance genes was reported (37). In the case of ISPst6, an equivalent putative P_{OUT} promoter was tested and found to be inactive in *E. coli* (pUS92 construct). In the *P. stutzeri* arrays examined here no phenotype is known, so further testing of this aspect was not feasible. In summary, the IS are unlikely to impact simple phenotypes but are predicted to influence complex phenotypes. To the extent that complex phenotypes are important to the maintenance of an integron, IS*1111-attC* elements may impact its evolutionary fitness. It is important to recognize here that complex phenotypes are a product of integron function over time, so we also need to consider ongoing activity of the integron.

Impact of IS*1111-attC* elements on integron functions. There is a strong correlation between genomic context and observable complex phenotypes in integrons. Multiple resistance phenotypes have only been reported from integrons on mobile elements, despite CIs typically containing more gene cassettes. This may reflect that for MRIs, their success is linked to that of their host plasmids. Resistance plasmids exploit a niche of frequently changing antibiotic challenges. In this selective context, it is important not only to express resistance genes but also to maintain the capacity to rapidly change the resistance phenotype. Indeed, it has been proposed that superior integron functions contribute to the relative success of class 1 integrons over class 2 and class 3 integrons in environments of intensive antibiotic use (23, 24).

The insertion of IS*1111-attC* elements is predicted to impact integron functions via the disruption of cassette mobility. It was therefore surprising to find examples where identical (ISPaX3_blaOXA-17) or near-identical (IsUnCu1_aadA1) copies of an IS-gene cassette fusion existed in two distinct integron contexts (Table 2). These observations could reflect either IntI-mediated mobilization of the fusion cassette or, alternately, integron-mediated capture or excision of adjacent gene cassettes. Irrespective of the mechanism, we note that the presence of a single fusion cassette has not prevented the capacity to develop new combinations of gene cassettes.

Long-term outcomes of IS-integron interaction. We also considered that the impact on integron function may increase as occupancy rates of *attC* sites increases or only become evident over long periods of time in a suitable host for IS activity. Our sample design enabled us to examine long-term outcomes by comparison of orthologous CIs in *Pseudomonas* strains separated by a range of evolutionary distances. We detected evidence for eight distinct IS elements (at the 95% identity level), each of which was confined to a single *Pseudomonas* lineage. We interpret this as independent and recent acquisition of a distinct IS in each lineage. Nonfunctional versions of all eight elements were found, and only for ISPst6 did we also find fully functional copies. Multiple near-identical copies of ISPst6 were present in nonidentical gene cassettes. We interpret this as ISPst6 acquisition by a common ancestor of the three genomovar 2 strains, followed by independent expansion to available *attC* sites in each array. A consequence of this expansion is that the operon-like structure of the cassette array is heavily disrupted. Perhaps the most significant observations were that in all cases where an integron contained an IS*1111-attC* element the cassette arrays were different and that in all cases where the available sequence information extended to the chromosomal framework at the right-hand end of the integron, a degenerated IS*1111-attC* element marked the cassette array boundary.

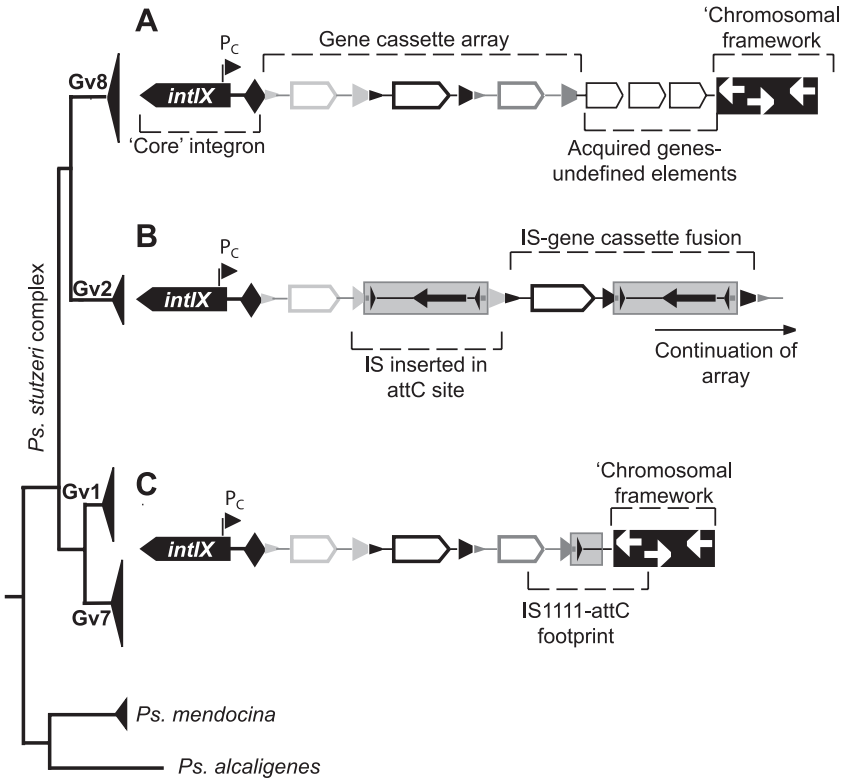

FIG. 5. Representation of the cycle of interaction between IS*1111-attC* elements and chromosomal integrons in *Pseudomonas*. A schematic 16S rRNA phylogenetic tree is shown on the left with four *P. stutzeri* genomovars and selected other species shown (the evolutionary distance within genomovars is exaggerated). (A) The preinfection phase was observed in genomovar 8 strains. In all five strains, no IS*1111-attC* sequences were seen, and integrons had the typical CI array structure. A small number of ORFs of unknown origin separate the last cassette from chromosomal framework genes. (B) The expansion phase was observed in genomovar 2 strains. This process is probably rapid since the IS must have been acquired after divergence from Gv8 and has given rise to multiple gene cassette/IS*1111-attC* element fusions by independent expansion in all three strains in the time since acquisition. (C) The postextinction state was observed in genomovars 1 (two of three strains) and 7 (one of five strains) and in *P. mendocina* (one of two strains) and *P. alcaligenes* (one strain). No intact IS were observed, and IS fragments defined an abrupt boundary between the last cassette and chromosomal framework genes.

When viewed collectively, these data suggest that a cycle of "infection," expansion, and extinction of IS*1111-attC* elements occurs within CI arrays (Fig. 5). The IS invasion of CIs does not directly result in loss of the integron or stabilization of the cassette array. It also does not result in loss of simple phenotypes, but it is expected to influence the expression of complex phenotypes (if these exist). The eventual outcomes are extinction of the IS, truncation of the cassette array, and modification of the right-hand boundary of the integron. This extinction/truncation process may involve other genetic elements (five unrelated IS elements were also seen) and is sufficiently rapid that it can be completed much faster than speciation in the *P. stutzeri* complex (Fig. 5).

Our observations have wider implications for the evolution of structures that are mosaics of different genetic elements and how we approach their study. The discovery that CIs can coevolve with their bacterial host (45) and are far more diverse than MRIs (4) has given rise to the hypothesis that CIs are the ancestors of MRIs (34). However, this hypothesis does not account for the observations that all known CIs appear to be acquired genes (based on codon usage analyses) and that even in cases where a family of CIs is strongly correlated to a bacterial genus there is actually only limited phylogenetic congruence (4, 44). An alternate hypothesis is that the ancestral integrons are mobile elements that can be readily fixed after acquisition by bacterial chromosomes. Here, the major outcome of the interaction between IS and integrons is array truncation and modification of the right-hand integron boundary. If the ancestral integrons are part of integrative elements, then the IS cycle of interaction represents a plausible means of integron fixation in a chromosome. The rapidity of the process means careful selection of strains for study (replication at multiple levels of evolutionary divergence) would be necessary to observe such phenomena.

Conclusions. The "DNA translocases" of integrons and IS*1111-attC* are heterologous proteins with different catalytic mechanisms. Comparison of the structural basis of their binding with the same recombination site is likely to shed significant insight into the nature of specific DNA-protein interactions. We consider that the recognition of gene cassette *attC* sites by both enzyme families represents convergent evolution. Both the IS*1111-attC* transposases and the IntIs are monophyletic groups in which all known members share target specificity for *attC* sites and the most closely related outgroups of the same protein superfamily do not. This implies that *attC* specificity is an ancestral feature of both the IS*1111-attC* and IntI

clades that has retained its biological relevance to both types of genetic element over the course of their evolutionary radiation. The advantage to IntIs of targeting *attC* sites is obvious given the relationship between integrons and gene cassettes. At first impression, the advantage to IS*1111-attC* transposases would seem to be that it gives access to large numbers of insertion sites. However, given the extinction cycle of IS-integron interactions and the rarity of functional IS, this is arguably not a successful strategy. We postulate that this unusual example of convergent evolution represents a broader biological theme, namely, that there is long-term selective advantage in interaction between genetic elements, and shared recognition of specific, extragenic target sites enables this. Insertion sequences are known to be involved in the formation of mosaic genetic elements (48), and the interaction between integrons and IS*1111-attC* group elements in *P. stutzeri* represents a model for how IS elements may initially come into contact with diverse assemblies of genes. As genomic data sets with greater replication at multiple levels of evolutionary divergence are assembled and analyzed, we predict that further examples of families of simple genetic elements sharing target sites will be recognized and ultimately shown to contribute to the formation of mosaic genetic elements.

ACKNOWLEDGMENTS

This study was funded by the Australian Research Council. S.G.T. is the recipient of an Australian Postgraduate Award.

Strains used in this study were provided by Johannes Sikorski and Neil Wilson. We acknowledge Nick Coleman and Ruth Hall for helpful discussions.

REFERENCES

1. **Aubert, D., T. Naas, and P. Nordmann.** 2003. IS1999 increases expression of the extended-spectrum beta-lactamase VEB-1 in *Pseudomonas aeruginosa*. J. Bacteriol. **185:**5314–5319.
2. **Bennasar, A., R. Rossello-Mora, J. Lalucat, and E. R. B. Moore.** 1996. 16S rRNA gene sequence analysis relative to genomovars of *Pseudomonas stutzeri* and proposal of *Pseudomonas balearica* sp. nov. Int. J. Syst. Bacteriol. **46:**200–205.
3. **Biskri, L., M. Bouvier, A. M. Guerout, S. Boisnard, and D. Mazel.** 2005. Comparative study of class 1 integron and *Vibrio cholerae* superintegron integrase activities. J. Bacteriol. **187:**1740–1750.
4. **Boucher, Y., M. Labbate, J. E. Koenig, and H. W. Stokes.** 2007. Integrons: mobilizable platforms that promote genetic diversity in bacteria. Trends Microbiol. **15:**301–309.
5. **Bouvier, M., G. Demarre, and D. Mazel.** 2005. Integron cassette insertion: a recombination process involving a folded single strand substrate. EMBO J. **24:**4356–4367.
6. **Brosius, J.** 1984. Plasmid vectors for the selection of promoters. Gene **27:**151–160.
7. **Brown, H. J., H. W. Stokes, and R. M. Hall.** 1996. The integrons In0, In2, and In5 are defective transposon derivatives. J. Bacteriol. **178:**4429–4437.
8. **Buchner, J. M., A. E. Robertson, D. J. Poynter, S. S. Denniston, and A. C. Karls.** 2005. Piv site-specific invertase requires a DEDD motif analogous to the catalytic center of the RuvC Holliday junction resolvases. J. Bacteriol. **187:**3431–3437.
9. **Centron, D., and P. H. Roy.** 2002. Presence of a group II intron in a multiresistant *Serratia marcescens* strain that harbors three integrons and a novel gene fusion. Antimicrob. Agents Chemother. **46:**1402–1409.
10. **Choi, S., S. Ohta, and E. Ohtsubo.** 2003. A novel IS element, IS*621*, of the IS*110*/IS*492* family transposes to a specific site in repetitive extragenic palindromic sequences in *Escherichia coli*. J. Bacteriol. **185:**4891–4900.
11. **Coleman, M. L., M. B. Sullivan, A. C. Martiny, C. Steglich, K. Barry, E. F. DeLong, and S. W. Chisholm.** 2006. Genomic islands and the ecology and evolution of *Prochlorococcus*. Science **311:**1768–1770.
12. **Coleman, N. V., and A. J. Holmes.** 2005. The native *Pseudomonas stutzeri* strain Q. chromosomal integron can capture and express cassette-associated genes. Microbiology **151:**1853–1864.
13. **Collis, C. M., and R. M. Hall.** 1995. Expression of antibiotic resistance genes in the integrated cassettes of integrons. Antimicrob. Agents Chemother. **39:**155–162.
14. **Collis, C. M., M. J. Kim, S. R. Partridge, H. W. Stokes, and R. M. Hall.** 2002. Characterization of the class 3 integron and the site-specific recombination system it determines. J. Bacteriol. **184:**3017–3026.
15. **Drouin, F., J. Melancon, and P. H. Roy.** 2002. The IntI-like tyrosine recombinase of *Shewanella oneidensis* is active as an integron integrase. J. Bacteriol. **184:**1811–1815.
16. **Duval-Valentin, G., B. Marty-Cointin, and M. Chandler.** 2004. Requirement of IS*911* replication before integration defines a new bacterial transposition pathway. EMBO J. **23:**3897–3906.
17. **Fournier, P. E., D. Vallenet, V. Barbe, S. Audic, H. Ogata, L. Poirel, H. Richet, C. Robert, S. Mangenot, C. Abergel, P. Nordmann, J. Weissenbach, D. Raoult, and J. M. Claverie.** 2006. Comparative genomics of multidrug resistance in *Acinetobacter baumannii*. PLoS Genet. **2:**e7.
18. **Gillings, M. R., M. P. Holley, H. W. Stokes, and A. J. Holmes.** 2005. Integrons in *Xanthomonas*: a source of species genome diversity. Proc. Natl. Acad. Sci. USA **102:**4419–4424.
19. **Ginard, M., J. Lalucat, B. Tuemmler, and U. Romling.** 1997. Genome organization of *Pseudomonas stutzeri* and resulting taxonomic and evolutionary considerations. Int. J. Syst. Bacteriol. **47:**132–143.
20. **Hacker, J., and E. Carniel.** 2001. Ecological fitness, genomic islands and bacterial pathogenicity: a Darwinian view of the evolution of microbes. EMBO Rep. **2:**376–381.
21. **Hall, R. M., D. E. Brookes, and H. W. Stokes.** 1991. Site-specific insertion of genes into integrons: role of the 59-base element and determination of the recombination cross-over point. Mol. Microbiol. **5:**1941–1959.
22. **Hall, R. M., A. J. Holmes, P. H. Roy, and H. W. Stokes.** 2007. What are superintegrons? Nat. Rev. Microbiol. **5:**C1. (Author reply, **5:**C2.)
23. **Hanau-Bercot, B., I. Podglajen, I. Casin, and E. Collatz.** 2002. An intrinsic control element for translational initiation in class 1 integrons. Mol. Microbiol. **44:**119–130.
24. **Hansson, K., L. Sundstrom, A. Pelletier, and P. H. Roy.** 2002. IntI2 integron integrase in Tn7. J. Bacteriol. **184:**1712–1721.
25. **Holmes, A. J., M. R. Gillings, B. S. Nield, B. C. Mabbutt, K. M. Nevalainen, and H. W. Stokes.** 2003. The gene cassette metagenome is a basic resource for bacterial genome evolution. Environ. Microbiol. **5:**383–394.
26. **Holmes, A. J., M. P. Holley, A. Mahon, B. Nield, M. Gillings, and H. W. Stokes.** 2003. Recombination activity of a distinctive integron-gene cassette system associated with *Pseudomonas stutzeri* populations in soil. J. Bacteriol. **185:**918–928.
27. **Janssen, D. B., I. J. Dinkla, G. J. Poelarends, and P. Terpstra.** 2005. Bacterial degradation of xenobiotic compounds: evolution and distribution of novel enzyme activities. Environ. Microbiol. **7:**1868–1882.
28. **Kallastu, A., R. Horak, and M. Kivisaar.** 1998. Identification and characterization of IS*1411*, a new insertion sequence which causes transcriptional activation of the phenol degradation genes in *Pseudomonas putida*. J. Bacteriol. **180:**5306–5312.
29. **Lalucat, J., A. Bennasar, R. Bosch, E. Garcia-Valdes, and N. J. Palleroni.** 2006. Biology of *Pseudomonas stutzeri*. Microbiol. Mol. Biol. Rev. **70:**510–547.
30. **Leon, G., and P. H. Roy.** 2003. Excision and integration of cassettes by an integron integrase of *Nitrosomonas europaea*. J. Bacteriol. **185:**2036–2041.
31. **Lloyd, A. L., B. J. Marshall, and B. J. Mee.** 2005. Identifying cloned *Helicobacter pylori* promoters by primer extension using a FAM-labeled primer and GeneScan analysis. J. Microbiol. Methods **60:**291–298.
32. **MacDonald, D., G. Demarre, M. Bouvier, D. Mazel, and D. N. Gopaul.** 2006. Structural basis for broad DNA-specificity in integron recombination. Nature **440:**1157–1162.
33. **Mahillon, J., and M. Chandler.** 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. **62:**725–774.
34. **Mazel, D.** 2006. Integrons: agents of bacterial evolution. Nat. Rev. Microbiol. **4:**608–620.
35. **McAdams, H. H., B. Srinivasan, and A. P. Arkin.** 2004. The evolution of genetic regulatory systems in bacteria. Nat. Rev. Genet. **5:**169–178.
36. **Partridge, S. R., and R. M. Hall.** 2003. The IS*1111* family members IS*4321* and IS*5075* have subterminal inverted repeats and target the terminal inverted repeats of Tn*21* family transposons. J. Bacteriol. **185:**6371–6384.
37. **Poirel, L., L. Brinas, N. Fortineau, and P. Nordmann.** 2005. Integron-encoded GES-type extended-spectrum beta-lactamase with increased activity toward aztreonam in *Pseudomonas aeruginosa*. Antimicrob. Agents Chemother. **49:**3593–3597.
38. **Poirel, L., J. W. Decousser, and P. Nordmann.** 2003. Insertion sequence ISEcp1B is involved in expression and mobilization of a bla_{CTX-M} beta-lactamase gene. Antimicrob. Agents Chemother. **47:**2938–2945.
39. **Prosseda, G., M. C. Latella, M. Casalino, M. Nicoletti, S. Michienzi, and B. Colonna.** 2006. Plasticity of the P junc promoter of ISEc11, a new insertion sequence of the IS*1111* family. J. Bacteriol. **188:**4681–4689.
40. **Quiroga, C., P. H. Roy, and D. Centron.** 2008. The S.ma.I2 class c group II intron inserts at integron *attC* sites. Microbiology **154:**1341–1353.
41. **Reese, M. G.** 2001. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. Comput. Chem. **26:**51–56.
42. **Robart, A. R., W. Seo, and S. Zimmerly.** 2007. Insertion of group II intron

retroelements after intrinsic transcriptional terminators. Proc. Natl. Acad. Sci. USA **104:**6620–6625.
43. **Rossello, R., E. Garcia-Valdes, J. Lulucat, and J. Ursing.** 1991. Genotypic and phenotypic diversity of *Pseudomonas stutzeri*. Syst. Appl. Microbiol. **14:**150–157.
44. **Rowe-Magnus, D. A., A. M. Guerout, L. Biskri, P. Bouige, and D. Mazel.** 2003. Comparative analysis of superintegrons: engineering extensive genetic diversity in the *Vibrionaceae*. Genome Res. **13:**428–442.
45. **Rowe-Magnus, D. A., A. M. Guerout, P. Ploncard, B. Dychinco, J. Davies, and D. Mazel.** 2001. The evolutionary history of chromosomal super-integrons provides an ancestry for multiresistant integrons. Proc. Natl. Acad. Sci. USA **98:**652–657.
46. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
47. **Schneider, D., and R. E. Lenski.** 2004. Dynamics of insertion sequence elements during experimental evolution of bacteria. Res. Microbiol. **155:** 319–327.
48. **Siguier, P., J. Filee, and M. Chandler.** 2006. Insertion sequences in prokaryotic genomes. Curr. Opin. Microbiol. **9:**526–531.
49. **Stokes, H. W., D. B. O'Gorman, G. D. Recchia, M. Parsekhian, and R. M. Hall.** 1997. Structure and function of 59-base element recombination sites associated with mobile gene cassettes. Mol. Microbiol. **26:**731–745.
50. **Szczepanowski, R., I. Krahn, A. Puhler, and A. Schluter.** 2004. Different molecular rearrangements in the integron of the IncP-1 beta resistance plasmid pB10 isolated from a wastewater treatment plant result in elevated beta-lactam resistance levels. Arch. Microbiol. **182:**429–435.
51. **Szekeres, S., M. Dauti, C. Wilde, D. Mazel, and D. A. Rowe-Magnus.** 2007. Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. Mol. Microbiol. **63:**1588–1605.
52. **Tobiason, D. M., J. M. Buchner, W. H. Thiel, K. M. Gernert, and A. C. Karls.** 2001. Conserved amino acid motifs from the novel Piv/MooV family of transposases and site-specific recombinases are required for catalysis of DNA inversion by Piv. Mol. Microbiol. **39:**641–651.
53. **Toussaint, A., and C. Merlin.** 2002. Mobile elements as a combination of functional modules. Plasmid **47:**26–35.
54. **Vaisvila, R., R. D. Morgan, J. Posfai, and E. A. Raleigh.** 2001. Discovery and distribution of super-integrons among pseudomonads. Mol. Microbiol. **42:** 587–601.
55. **Walsh, T. R.** 2006. Combinatorial genetic evolution of multiresistance. Curr. Opin. Microbiol. **9:**476–482.