Natural transformation with synthetic gene cassettes: new tools for integron research and biotechnology

Alicia M. Gestal, Elissa F. Liew and Nicholas V. Coleman

INTRODUCTION

Multiple antibiotic resistance in bacteria is a problem for human health (Davies & Davies, 2010), and is associated with the actions of mobile genetic elements (MGEs) such as plasmids, transposons and integrons (Frost et al., 2005). Integrons can assemble resistance genes from diverse sources at a single genetic locus (Hall & Collis, 1998; Mazel, 2006; Schlüter et al., 2007; Stokes & Hall, 1989; Gillings et al., 2008). The integron consists of the intI gene encoding integrase, an attachment site (attI), a promoter (Pc), and an array of mobile gene cassettes (Nunes-Duby et al., 1998; Recchia & Hall, 1997; Stokes & Hall, 1989; Lévesque et al., 1994; Jové et al., 2010). The cassettes typically consist of a single promoterless ORF and a recombination site (attC). The integron integrase recombines attC with attI, or attC with attC, enabling both integration and excision of gene cassettes, which shift between free circular and linear integrated forms. Integrons are common in bacteria, and diverse in their phylogeny and predicted functions (Boucher et al., 2007a; Elsaied et al., 2007; Gillings et al., 2005; Holmes et al., 2003a; Nield et al., 2001; Rowe-Magnus & Mazel, 2001; Rowe-Magnus et al., 2003). Integrons increasingly seem to have a role as general-purpose agents of adaptation, not just enablers of antibiotic resistance (Koenig et al., 2009, 2011; Rosewarne et al., 2010; Wright et al., 2008).

Previous studies have used excision assays (Collis & Hall, 1992; Drouin et al., 2002; Léon & Roy, 2003) or cointegrate assays (Biskri et al., 2005; Collis et al., 2002, 2001; Holmes et al., 2003b; Martinez & de la Cruz, 1988) to quantify integron recombination. Improvements to these assays include the use of mobilizable suicide plasmids to deliver single-stranded recombination templates (Bouvier et al., 2005), and using quantitative PCR to simplify the cointegrate assay (Shearer & Summers, 2009; Wei et al., 2011; Yang et al., 2009). Such approaches have given useful information, but it can be argued that they are not representative of integron behaviour in wild-type bacteria, since they use fragmented integrons, overexpressed integrase and/or plasmids with attC sites as proxies for true gene cassettes. We aimed to address this issue, by devising a new recombination assay that can be performed in wild-type bacteria.

Pseudomonas stutzeri strains play a role in pollutant biodegradation (Lewis et al., 2000; Shim et al., 2001) and are occasional opportunistic pathogens (Lalucat et al., 2006). Some P. stutzeri strains can be naturally transformed...
with DNA (Sikorski et al., 2002), and some contain chromosomal integrons (Holmes et al., 2003b). We previously developed an assay for integron recombination in P. stutzeri strain Q (PstQ), via electroporation of a plasmid (pUS23) containing an attC site as a model for a gene cassette (Coleman & Holmes, 2005). The assay had the advantage of being performed in the native integron host, and at wild-type levels of integrase expression, but was confounded by replication of pUS23 in P. stutzeri and insertion sequence-mediated activation of the reporter genes. Only one prior study has, to our knowledge, been successful in performing recombination in vivo in an integron (Collis et al., 1993), and in that case, the recombination frequencies were too low to quantify accurately. We aimed here to develop an improved assay, firstly by optimizing in vitro methods of cassette manufacture, and secondly by investigating natural transformation as a method of introducing the gene cassettes.

Several authors have pointed out the conceptual similarity between integrons and cloning/expression vectors (Boucher et al., 2006; Collis & Hall, 1992; Stokes et al., 1997; Vaisvil et al., 2001). Integrons permit site-specific integration of DNA without disrupting the host genome (Boucher et al., 2007b), have a flexible, modular nature (Recchia & Hall, 1995), and are found in diverse bacterial hosts. Integrons have also been envisioned as tools for directed evolution and metabolic engineering. Recently, the first successful example of this was reported (Bikard et al., 2010), in which a synthetic integron constructed in Escherichia coli was made to shuffle gene cassettes consisting of modules from the tryptophan biosynthesis pathway, yielding new variants of this pathway.

In this study, we report the development of new methods for studying integron-mediated recombination and expression in wild-type bacteria, and the use of these methods to engineer metabolic traits via the integron platform.

METHODS

Bacterial strains, media and growth conditions. Details of bacterial strains are given in Table 1. All cultures were grown aerobically in Luria–Bertani (LB) medium (Sambrook & Russell, 2001), unless indicated otherwise, with broths shaken at 200 r.p.m.

P. stutzeri strains were grown at 30 °C and E. coli strains were grown at 37 °C. The following antibiotics were used: gentamicin (Gm; 10 μg ml\(^{-1}\)), ampicillin (Ap; 100 μg ml\(^{-1}\)), trimethoprim (Tm; 50 μg ml\(^{-1}\)). For screening of lacZY activity on LB plates, X-Gal (80 μg ml\(^{-1}\)) was added. In experiments to select or maintain lacZY-containing strains, mineral salts medium (MSM) (Coleman et al., 2002) was used instead of LB, supplemented with lactose or glucose (20 mM) as the sole carbon source.

DNA extraction, purification, sequencing and analysis. Purified genomic DNA (used for cassette construction and RAPD PCR) was extracted by a cetyltrimethylammonium bromide (CTAB)-based method (Coleman & Holmes, 2005). For PCR screening of colonies or patches on agar plates, biomass (~0.5 mm\(^3\)) was transferred to 50 μl Tris-EDTA buffer (Sambrook & Russell, 2001) and lysed by heating (99 °C, 5 min). The crude lysate was centrifuged (16 000 g, 5 min), and 1 μl of the supernatant was used as the template for PCR. Plasmid DNA was extracted by a standard alkaline lysis protocol (Sambrook & Russell, 2001). For sequencing or ligations, DNAs were purified with QIAquick columns (Qiagen). Gel electrophoresis was done using 1% agarose gels made with 0.5 × Tris-borate-EDTA buffer, stained with ethidium bromide. Sequencing was performed at the Australian Genome Research Facility (AGRF), Westmead node, via the Sanger method. Sequence alignments were made with CLUSTAL_X or MacVector software.

PCR and general molecular biology methods. All enzymes and molecular biology reagents were obtained from New England Biolabs, and used according to standard protocols (Sambrook & Russell, 2001), unless otherwise indicated. Quantification of DNA was done using a NanoDrop 1000 machine (Thermo Scientific). Standard PCRs (clone screening) were set up in 25 μl reactions containing 2.5 μl 10× Thermopolly buffer, 200 μM dNTPs, 1 μM of each primer and 1.25 U Taq polymerase. High-fidelity PCRs (cassette construction) used Phusion polymerase (0.02 U) (Finnzymes) in 50 μl reactions containing 3% DMSO. Thermocycling was done in an Eppendorf Mastercycler EP-S machine, with initial denaturation at 94 °C for 30 s (Tag) or 98 °C for 10 s (Phusion), then 30 cycles of denaturation at 94 °C, 98 °C for 30 s (Tag) or 10 s (Phusion), annealing at 48–66 °C for 30 s, extension at 72 °C for 1–8 min, then a final extension at 72 °C for 5 min. Thermocycling details for each primer pair are available on request.

Construction of pUS13, pUS23 and pBBR-lac plasmids. The construction of pUS23 has been described previously (Coleman & Holmes, 2005). Plasmid pUS13 was made as follows, using E. coli JM109 as the host. The aadB gene and attC\(\text{aaddB-oxa2}\) site were amplified from plasmid pMAQ105 using primers NVCS and NVC6, digested with KpnI and BamHI, and ligated into similarly digested pUC19 to yield plasmid pUS10. The attC\(\text{aaddB-oxa2}\) site from pMAQ105 was amplified using primers NVC10 and NVC8 (which added the Pho\(\text{X}\) promoter at the 3′ end of the PCR product), digested with HindIII and PstI, and ligated into pUS10 to yield plasmid pUS11. The gfp gene (mut3 variant) was amplified from pTGF2 using primers NVCS13 and NVCS14, digested with PstI and KpnI, and ligated into pUS11 to yield plasmid pUS13. The insert region of pUS13 was sequenced to confirm the correct structure. E. coli JM109 cells carrying pUS13 required propagation at room temperature (25 °C) to stably maintain the plasmid.

Plasmid pBBR-lac was made by amplifying the lacZY genes of E. coli P801 with primers NVCS86 and NVCS287, digesting with XbaI and NdeI, cloning into XbaI-digested pBBR1MCS-2, and transfer into E. coli JM108. The expected plasmid structure was confirmed by the blue phenotype conferred on the Lac}\(\text{C}\) strain JM108, by checking the size of junction PCR products (primers NVCS15/EFL7 and EFL8/EFL9), and the restriction pattern produced using EcoRI.

Construction of gene cassettes by PCR and the enzymic ligation assisted by nucleases (ELAN) reaction. Gene cassettes were PCR-amplified as follows: type A from pUS23 with AMG5/AMG6, type B from pUS13 with AMG10/AMG11, type C from pUS13 with AMG10/NVC302, type F from E. coli P801 genomic DNA with NVC286/NVC287. Plasmid templates were linearized by restriction digestion using SpI (pUS23) or HindIII (pUS13) before PCR. The PCR products were column-purified, then self-ligated using the ELAN method (Cost & Cozzarelli, 2007). Briefly, restriction enzymes are chosen with compatible overhangs that yield a non-cleaveable ligation junction (e.g. NdeI/XbaI). The DNAs are then ligated in the presence of the enzymes used for digestion to enhance the yield of recombinant products. Specifically, ELAN should enhance the yield of circular products — these contain a non-cleaveable NdeI–XbaI junction. The ELAN reaction was set up in a 1 ml volume containing 3–5 μg
Table 1. Bacterial strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Genotype or description</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli JM109</td>
<td>recA1 supE44 endA1 hisD17 gmrA96 relA1 thi (λlac-proAB) F’[traD36 proAB+ lacIq lacZ Δ15]; Gm3, GFP+, Lac+</td>
<td>Yanisch-Perron et al. (1985)</td>
</tr>
<tr>
<td>E. coli JM108</td>
<td>recA1 supE44 endA1 hisD17 gmrA96 relA1 thi (λlac-proAB); Gm3, GFP+, Lac-</td>
<td>Yanisch-Perron et al. (1985)</td>
</tr>
<tr>
<td>E. coli P801</td>
<td>Wild-type E. coli K12; Hfr; Gm3, GFP+, Lac+</td>
<td>Utsumi et al. (1988)</td>
</tr>
<tr>
<td>P. stutzeri strain</td>
<td>Electrotransformable; contains integrone inPstQ, consisting of intPstQ, attPstQ and Q</td>
<td>Coleman &amp; Holmes (2005); Holmes et al. (2003b)</td>
</tr>
<tr>
<td>P. stutzeri ATCC 17587</td>
<td>Naturally transformable; contains integrone in17587, consisting of int17587, att17587 and four gene cassettes; Gm3, GFP-, Lac-</td>
<td>Sikorski et al. (1999); Tetu (2007)</td>
</tr>
<tr>
<td>P. stutzeri ATCC 17641</td>
<td>Naturally transformable; contains integrone in17641, consisting of int17641, att17641 and two gene cassettes; Gm3, GFP+, Lac-</td>
<td>Sikorski et al. (1999); Tetu (2007)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pUC19</td>
<td>2.7 kb, Ap R; general-purpose E. coli cloning and expression vector</td>
<td>Yanisch-Perron et al. (1985)</td>
</tr>
<tr>
<td>pBRR1MCS-2</td>
<td>5.1 kb, Km R; mobilizable broad-host-range cloning and expression vector</td>
<td>Kovach et al. (1995)</td>
</tr>
<tr>
<td>pBR-lac</td>
<td>9.6 kb, Km R; contains E. coli lacZY genes cloned into pBRR1MCS-2</td>
<td>This study</td>
</tr>
<tr>
<td>pSU2056</td>
<td>3.8 kb, Ap R, Int1+; pUC9 derivative containing 1176 bp BamHI–SalI fragment from Tn21; overexpresses IntI</td>
<td>Martinez &amp; de la Cruz (1990)</td>
</tr>
<tr>
<td>R388</td>
<td>33.9 kb, Tp R, Su R, Int1+, attC+; naturally occurring, conjugal, broad-host-range IncW plasmid; contains class 1 integron In3, consisting of intl1, att11, two gene cassettes (dfrB2, orfA) and an adjacent 3' conserved sequence (qacEA1, sul2)</td>
<td>Datta &amp; Hedges (1972); Hall et al. (1994); Ward &amp; Grinsted (1982)</td>
</tr>
<tr>
<td>pUS13</td>
<td>4.7 kb, Ap R, Gm3, Gfp+, attC+; pUC19 derivative containing two attC sites flanking P lac promoter, gfp and aadB</td>
<td>This study</td>
</tr>
</tbody>
</table>

PCR product, 1 × NEB buffer 4, 100 U each of XbaI and NheI, 500 U T4 DNA ligase and 1 mM ATP, incubated for 16 h at 33 °C and purified on a QiAquick column, eluting into 50 μl water. To estimate the proportion of circular molecules, 500 ng DNA from ELAN reactions was treated with Plasmid-Safe DNase (GeneResearch) before electrophoresis.

**Construction of cassette arrays by multiple displacement amplification.** Linear gene cassette arrays were constructed by multiple displacement amplification (MDA) using 29 plasmid DNA polymerase (Dean et al., 2001) with the REPLI-g UltraFast kit (Qiagen), with modifications. Purified DNA from ELAN mixtures (200–300 ng) was denatured then neutralized using the D1 and N1 kit buffers, then 8 μl of this template was amplified in 40 μl reaction mixtures containing 15 μl MilliQ water, 4 μl 10 × NEB buffer 4, 250 μM dNTPs, 60 μg BSA ml-1, 1.25 μM primer (NVC300 for aadB–gfp-containing cassettes, or EFL7/EFL8 for lacZY cassettes) and 1 μl 29 polymerase. The MDA reaction was incubated for 16 h at 30 °C, then used in natural transformation experiments without any further purification.

**Preparation of electrocompetent cell cultures and electroporation.** Electrocompetent E. coli cells were made by standard methods (Sambrook & Russell, 2001), while electrocompetent PstQ cells were made as described previously (Coleman & Holmes, 2005). Volumes of 50 μl of electrocompetent cells were mixed with ~0.1 μg plasmid vector (1–2 μl) or ~0.5 μg of gene cassettes (5–10 μl), transferred to chilled electroporation cuvettes (2 mm gap) and electroporated in a Bio-Rad Gene Pulser machine (2.5 kV, 25 μF, 200 Ω). Cells were recovered in 1 ml LB broth with shaking, either for 1 h at 37 °C (E. coli) or 2 h at 30 °C (PstQ), then plated on appropriate selective media. In the case of lacZY experiments, cells were washed three times in PBS before plating onto MSM with lactose as sole carbon source.

**Natural transformation of strains 17587 and 17641.** A single colony of 17587 or 17641 (~5 × 106 cells) was transferred onto LB agar to give a 1 cm diameter patch. The patches were grown for 16 h, then DNA (200–600 ng in 1–5 μl) was dispensed onto the centre of the patch. After incubating for a further 24 h, the cells were scraped off the plate, resuspended in 1 ml PBS, and spread-plated on appropriate selective medium. In the case of lacZY experiments, cells were washed three times in PBS before plating onto MSM with lactose.

**Analysis of frequencies and locations of gene cassette insertion.** For each combination of cassette type and host strain, at least two separate transformation experiments were performed, and colonies appearing with the expected phenotype (either GmR or Lac+) were counted. For aadB–gfp cassettes, approximately 20 colonies from each of two experiments were patched to new medium, while for lacZY cassettes, two colonies from each of two experiments were patched. Patches were screened by PCR for the expected left and right junctions between the integron and the gene cassette, assuming insertion at the attl site. All screening PCRs were done twice, with the second PCR done using a long extension time (10 min) to detect left junctions of cassettes inserted at attC sites close to attl. The sizes of the junction PCR products were recorded, and representative PCR products were sequenced. The phenotype of one or two representatives of each type of recombinant was investigated in more detail by fluorometric analysis (aadB–gfp cassettes, details below) or by monitoring growth in MSM broth with lactose as sole carbon source (lacZY cassettes).

**Fluorometry and fluorescence microscopy.** Selected GmR recombinants were grown for 16 h in LB-Gm broth to OD600 1.0–1.2, washed twice in buffer containing 10% sucrose, 1 mM MgSO4, 1 mM HEPES (pH 7), and resuspended in the same buffer to give an OD600 between 0.5 and 0.9. Fluorometric readings were made in a Cary
Eclipse Fluorescence spectrophotometer (10 mm path length, excitation 490 nm, emission 520 nm, excitation slit width 5 nm, emission slit width 20 nm). Two fluorometric readings were made per sample and the experiment was performed three separate times. The data were expressed as average fluorescence units (AFU) divided by the density of the cell suspensions (AFU/OD<sub>600</sub>). The response of the fluorometer was confirmed to be linear over the OD<sub>600</sub> range used in these experiments. Selected samples, including any displaying weak fluorescence in fluorometer assays, were also examined by fluorescent microscopy to confirm GFP, as described previously (Coleman & Holmes, 2005).

**RESULTS**

**Construction of synthetic gene cassettes**

Synthetic gene cassettes made and tested in this study are shown in Fig. 1. These include both promoter-containing and promoterless types, circular and linear types, and elements with and without attC sites. Four different selection/detection methods are possible with these cassettes: gentamicin resistance (Gm<sup>R</sup>), green fluorescence (GFP<sup>+</sup>), growth on lactose (Lac<sup>+</sup>) and production of a blue colour from X-Gal (X-Gal<sup>+<sup>). The attC site used in all cases was attC<sub>aadD1-oxa2</sub>, which is a small (60 bp) site known to be recombinogenic in <i>E. coli</i> and <i>Pseudomonas</i> (Coleman & Holmes, 2005; Collis et al., 2001; Holmes et al., 2003b). Cassette types A–E were made by PCR amplification of the marker genes and attC site from plasmid templates, while F and G were made by amplification of the marker genes from chromosomal DNA, with the attC site added during PCR (types F and G) (Fig. 1). The latter strategy gives great flexibility: any gene can be converted into a gene cassette in two steps (PCR, ligation).

The first phase of cassette construction used the ELAN approach (Cost & Cozzarelli, 2007) to make circular gene cassettes from PCR products (see Methods for details). The products of the ELAN reaction were examined by electrophoresis with and without treatment by an enzyme that degrades linear DNA (Plasmid-Safe DNase). Five main bands were seen (Supplementary Fig. S1), three of which were removed by the enzyme. Our interpretation is that two circular forms are present (monomer, dimer), and three linear forms (monomer, dimer, trimer). Due to the ELAN method used, the linear ligation products are likely to be head–tail type concatemers.

During natural transformation, linear dsDNA is taken up (Carlson et al., 1983) and converted to ssDNA. Since the recombinogenic form of the gene cassette is ssDNA, we reasoned that a large linear dsDNA containing an array of identical gene cassettes might be a good substrate for DNA uptake, then integron-mediated recombination. We created such arrays by MDA (Blanco et al., 1989; Dean et al., 2001) using the ELAN mixture as the template. This method of cassette construction yielded very large products (>23 kb by gel electrophoresis; Supplementary Fig. S2), which we expect to be large tandem gene cassette arrays.

**Type A and B cassette DNAs integrate at attl in the mobilized integron In<sup>3</sup>**

The integron In<sup>3</sup> on plasmid R388 was used to test the activity of the synthetic gene cassettes; this integron has three possible recombination sites (attI<sub>1</sub>, attIC<sub>drf2/orfa</sub> and attIC<sub>orfA/qacE</sub>). The <i>E. coli</i> cells used in this test also contained pSU2056, which overexpresses IntI1 integrase. We hypothesized that type A cassette DNA (promoterless) would yield insertions close to P<sub>C</sub> (i.e. at the attI site), while type B cassette DNA (containing P<sub>ta</sub>) would yield more diverse insertions, including some distant from P<sub>C</sub> (i.e. at attC sites).

Electroporation of both type A and B cassette DNAs into strain JM109 (pSU2056) (R388) yielded similar numbers of Gm<sup>R</sup> colonies [3 × 10<sup>3</sup> (µg DNA)<sup>−1</sup> (Table 2)]. No Gm<sup>R</sup> colonies were seen when plasmid-free JM109 was transformed with the type A or B cassette DNA, or when JM109 (pSU2056) (R388) was electroporated with type C cassette DNA (lacking an attC site). Twenty Gm<sup>R</sup> colonies from two separate experiments of each type were screened by PCRs spanning the predicted recombination junctions at the attI<sub>1</sub> site. The PCRs confirmed that both type A and B cassette DNAs had integrated at the attI<sub>1</sub> site of In<sup>3</sup> in nearly all cases. Increasing the extension time of the PCR did not yield any further products. Most junction PCR products were of the expected size, but a minority displayed size variations (± ~100 bp) (Supplementary Table S2). This was analysed in more detail in the experiments with Pseudomonas (below). The sequences of junction PCR products were consistent with site-specific recombination at the GTTRRRRY core site of attI<sub>1</sub> (Fig. 2). The hypothesis that promoter-containing cassettes would integrate at more diverse sites was disproved, since most Gm<sup>R</sup> colonies (88–95%) were due to insertions at the attI<sub>1</sub> site, and the remainder due to cassette insertion at unknown locations.

**Type A and B cassette DNAs integrate at attl in the PstQ integron**

The wild-type integron-containing bacterium <i>P. stutzeri</i> (PstQ) is a more challenging test of the synthetic gene cassette approach, because there is no overexpression of integrase in this case. Electroporation of PstQ with ELAN
mixtures yielded GmR colonies with cassette types A and B, but not C (Table 2). The frequency of appearance of GmR colonies in PstQ was approximately 50-fold lower than in JM109(R388)(pSU2056), but was well above the threshold of detection [2 c.f.u. (µg DNA)^-1]. PCR screening indicated that integration was independent of whether the cassette contained a promoter, and was almost exclusively at attI. Only two GmR PstQ derivatives out of a total of 56 screened did not yield PCR products consistent with attI integration.

Sequencing of junction PCR products from three independently derived GmR colonies confirmed that site-specific integration at the attI site of the PstQ integron had occurred (Supplementary Fig. S3). Recombination in clones 4B and 6B was between the expected attI and attC core sites, but in the case of clone 1B, the junction sequence indicated recombination between attI and a secondary site in the gene cassette (TTTAGGC). There was variation in the sizes (±100 bp) of the junction PCR products obtained from GmR PstQ clones (Supplementary Table S2). The shorter sequences were consistent with recombination at a secondary site in the cassette, while the longer sequences contained extra copies of primers at the cassette ligation junction (data not shown).

Natural transformation of type A, B and D cassette DNAs leads to integration at the attI sites of the 17587 and 17641 integrons

P. stutzeri strains 17587 and 17641 are naturally transformable (Lorenz & Sikorski, 2000) and contain chromosomal
integrons (Tetu, 2007; Wilson, 2007) (GenBank accession no. EF648209). The integrons in 17641 and 17587 have not previously been tested for recombination or expression activity. Both type A and B gene cassette DNAs were capable of natural transformation of 17587 and 17641 to yield GmR colonies at efficiencies similar to that of electroporation into

<table>
<thead>
<tr>
<th>Type of gene cassette DNA</th>
<th>Combined transformation and recombination frequency* † [c.f.u. (μg DNA)⁻¹] and insertion location‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control JM109 Mobile integron in E. coli JM109 (R388) (pSU2056)§ Chromosomal integrons in P. stutzeri</td>
<td></td>
</tr>
<tr>
<td>No DNA</td>
<td>0</td>
</tr>
<tr>
<td>Type A (aadB gfp attC; ELAN)</td>
<td>3430 ± 24</td>
</tr>
<tr>
<td>Type B (P₅lac gfp aadB attC; ELAN)</td>
<td>3280 ± 111</td>
</tr>
<tr>
<td>Type C (aadB gfp; ELAN)</td>
<td>0</td>
</tr>
<tr>
<td>Type D (P₅lac gfp aadB attC; MDA)</td>
<td>ND</td>
</tr>
</tbody>
</table>

*Data are the average of two (types F and G) or three (types A–E) separate experiments ± SD.
†Natural transformation with gene cassettes for 17587 and 17641, and electroporation in all other cases.
‡Insertion location is given as (number of clones with gene cassettes at the attI site/total number screened).
§pUS2056 was present at the start of the experiment, but was most likely not maintained.
‖Values of 0 indicate no colonies observed. The limit of detection was 2 c.f.u. (μg DNA)⁻¹.

Fig. 2. Alignment of R388 integron In3 and gene cassette sequences with recombination junction sequences amplified from four GmR recombinant E. coli JM109(R388) clones (6A, 8A, 1B, 4B). The GTTRRRY core site is underlined. Positions conserved in the majority of sequences are in bold type with a grey background.
PstQ. Natural transformation of 17587 and 17641 with MDA-generated linear cassette DNA (type D) yielded GmR colonies at very high frequencies \(10^5-10^6\) (μg DNA)\(^{-1}\), which were orders of magnitude higher than those seen with the ELAN ligation mixture. The appearance of GmR colonies depended absolutely on the presence of \(attC\) sites in both ELAN-derived and MDA-derived cassettes, since type C and E cassette DNAs gave no colonies. This indicates that integron-mediated events were occurring rather than illegitimate or homologous recombination. The rate of spontaneous mutation to GmR was below the detection limit under these experimental conditions (<10\(^{-9}\) per cell).

Integration of type A, B and D gene cassette DNAs at \(attI\) sites in 17587 and 17641 was confirmed by PCR in the vast majority of cases tested (Table 2 and Supplementary Table S2). Out of 240 GmR colonies screened in these experiments, only 11 failed to give \(attI\) junction PCR products. Further PCR testing using longer extension times again gave the same negative results. The PCR products obtained from 17641 and 17587 were variable in size (± ~200 bp) (Supplementary Table S2). Sequencing of representative PCR products (Fig. 3, Supplementary Figs S4 and S5) revealed similar phenomena to those described above, i.e. all sequences were consistent with IntI-mediated recombination, some indicated secondary site recombination, and some contained extra copies of the PCR primers.

**Fluorometric analysis of strains carrying type A and B cassettes**

All the GmR strains assayed had higher fluorescence than their corresponding GmS parental strain (Table 3). The differences were statistically significant by \(t\) test at \(P<0.1\), but not at \(P<0.01\). In all cases, strains with integrated type B cassette DNA were more fluorescent than the corresponding strain with integrated type A cassette DNA \((P<0.001)\): the difference ranged from twofold in \(E. coli\) to 15-fold in 17587, after correction for wild-type autofluorescence. The plasmid-borne integron In3 gave stronger expression than the chromosomal integrons \((P<0.001, \text{vs data from PstQ})\); this difference was more noticeable with type A cassettes (3.8-fold) than with type B cassettes (1.9-fold).

Based solely on fluorescence data, the \(P_C\) promoters of PstQ and 17641 were 4.5 and 3.1 times stronger than that of 17587 (Table 3). This observation matches sequence predictions, since 17641 and PstQ contain identical putative \(P_C\) sequences (TTGAGCCGAATGGCTACGCTGAT) (Coleman & Holmes, 2005), while the putative \(P_C\) sequence in 17587 is different (TTGAGCCGAATGGCTACGCTGAT). The putative −10 sequence of the 17587 \(P_C\) promoter (CTCGGT) is a notably poorer match to the consensus \((TATAAT)\) compared with the −10 sequences of PstQ and 17641 (TCTGAT).

---

**Fig. 3.** Alignment of representative \(P. stutzeri\) 17641 integron and gene cassette sequences with recombination junction sequences amplified from GmR clones (4A, 9A, 4B, 7B). The GTTRRRY core site and proposed secondary site TTTAGGC are underlined. Positions conserved in the majority of sequences are in bold type with a grey background.
Table 3. Fluorescence analysis of gfp-containing recombinants and parental strains

| Strain      | Cassette DNA type | Fluorescence (AFU/OD₆₀₀)* | t test 1‡‡ | t test 2$ | t test 3|| |
|-------------|------------------|---------------------------|------------|----------|----------|
| PstQ        | None             | 138 ± 26                  | NA         |          |          |
| PstQ-3A     | A                | 253 ± 47                  | 1.8 × 10⁻² |          |          |
| PstQ-6A     | A                | 211 ± 48                  | 2.3 × 10⁻² |          |          |
| PstQ-1B     | B                | 734 ± 48                  | 2.5 × 10⁻⁴ | 1.9 × 10⁻⁴ |          |
| PstQ-6B     | B                | 1020 ± 22                 | 4.6 × 10⁻⁶ |          |          |
| 17587       | None             | 101 ± 28                  | NA         |          |          |
| 17587-2A    | A                | 122 ± 42                  | 2.3 × 10⁻¹ |          |          |
| 17587-8A    | A                | 122 ± 22                  | 3.1 × 10⁻² |          |          |
| 17587-6B    | B                | 403 ± 116                 | 1.9 × 10⁻² | 9.5 × 10⁻⁵ |          |
| 17587-16B   | B                | 428 ± 95                  | 1.3 × 10⁻² |          |          |
| 17641       | None             | 144 ± 32                  | NA         |          |          |
| 17641-4A    | A                | 212 ± 37                  | 6.2 × 10⁻² |          |          |
| 17641-9A    | A                | 207 ± 27                  | 5.1 × 10⁻² |          |          |
| 17641-4B    | B                | 789 ± 83                  | 1.8 × 10⁻³ | 3.0 × 10⁻⁵ |          |
| 17641-7B    | B                | 812 ± 157                 | 1.3 × 10⁻² |          |          |
| JM109(R388)(pSU2056) | None | 129 ± 23                  | NA         |          |          |
| JM109(R388)-6A | A       | 875 ± 131                | 4.7 × 10⁻³ |          | 7.3 × 10⁻⁵ |
| JM109(R388)-8A | A       | 869 ± 128                | 6.2 × 10⁻³ |          |          |
| JM109(R388)-1B | B       | 1680 ± 222               | 2.7 × 10⁻³ | 3.7 × 10⁻⁴ | 6.0 × 10⁻⁴ |
| JM109(R388)-4B | B       | 1730 ± 373               | 7.9 × 10⁻³ |          |          |

*Data are means ± SD of three separate experiments.
†All paired t tests, single-tailed.
‡‡t test 1: data from GmR GFP+ strain versus corresponding GmR GFP− parental strain.
§t test 2: data from type A cassettes versus type B cassettes in the same host strain.
||t test 3: data from plasmid-borne integron versus a chromosomal integron (PstQ), for each cassette type.

Metabolic engineering using type G cassette DNA

*P. stutzeri* strains cannot grow on lactose, but can grow on glucose (Rossello-Mora et al., 1994). Acquisition of *lacY* and *lacZ*, encoding lactose permease and β-galactosidase, should enable *P. stutzeri* to grow on lactose via the glucose released. To test this, *lacZY* genes cloned in the broad-host-range vector pBBR1MCS2 (construct pBBR-lac) were introduced into *P. stutzeri* strains. No KmR transformants of 17641 were obtained, but KmR and X-Gal+ transformants of PstQ and 17587 were obtained. Both of these grew poorly on lactose in MSM compared with glucose, but displayed more growth on lactose than in medium with no added carbon source (data not shown). Since plasmid-borne *lacZY* genes could transform *P. stutzeri* strains to an X-Gal+ and Lac+ phenotype, we proceeded to study *lacZY* gene cassettes (types F+G).

Introduction of type F cassette DNA into PstQ via electroporation or into 17587 and 17641 by natural transformation gave no Lac+ recombinants. Electroporation of type G cassette DNA into PstQ also failed to give any Lac+ recombinants. However, natural transformation of 17587 and 17641 with type G cassette DNA yielded Lac+ colonies at low frequencies [17–67 c.f.u. (µg DNA)⁻¹] (Table 2). PCR screening revealed that 3/4 of the 17587 colonies and 4/4 of the 17641 colonies contained the *lacZY* genes at the attl site. Sequencing confirmed that the insertion events were consistent with integron integrase-mediated site-specific recombination (Supplementary Fig. S5).

A RAPD PCR with the BOXA1R primer (Supplementary Fig. S6) confirmed that the Lac+ strains were genetically related to the expected *P. stutzeri* strains, i.e. they were not *E. coli* or other contaminants. All the Lac+ recombinants were dark blue on LB-X-Gal agar, while the parental strains 17587 and 17641 were orange on LB-X-Gal. Two Lac+ recombinants (17587-lac1 and 17641-lac1) were chosen for further investigation: both grew to high densities on lactose as sole carbon and energy source in mineral salts broth medium (Fig. 4), reaching final OD₆₀₀ values of 1.3 (17587-lac1) and 1.8 (17641-lac1). Doubling times on lactose were 10 h (17587-lac1) and 1.8 h (17641-lac1), which was much slower than growth on glucose (e.g. doubling time of 3.3 h for wild-type 17641). The stability of the *lacZY* cassette in the integron of strain 17587-lac1 was tested under non-selective conditions (plain LB broth), and compared with the stability of a plasmid (pBBR-lac) carrying the same genes (Supplementary Fig. S7). The *lacZY* genes in 17587-lac1 were very stable, and 99% of cells were still Lac+ after 20 generations in plain LB. Under the same conditions, the
percentage of Lac+ colonies in 17587(pBBR-lac) cultures declined to zero after eight generations.

DISCUSSION

Our study is the first to show directly that gene cassettes can enter bacterial cells via natural transformation. This has important implications, since many clinically significant bacterial genera, such as Acinetobacter, Campylobacter, Helicobacter, Pseudomonas and Vibrio, are both naturally transformable and integron-containing (Crespo et al., 2005; Etchuuya et al., 2011; Johnsborg et al., 2007; Meibom et al., 2005; O’Halloran et al., 2004). Our data provide a conjugation-independent mechanism for the transfer of gene cassettes between chromosomal integrons; there is good circumstantial evidence that this occurs [i.e. the finding of identical cassettes in different hosts (Rowe-Magnus et al., 2003)]. We propose that the phenomenon of combined natural transformation and integron-mediated recombination that we have observed here in *P. stutzeri* is significant in other species contexts and in the environment at large.

We believe that this study is the first to demonstrate integron-mediated capture of a catabolic gene. The method that we have used to generate *lacZY* cassettes (incorporation of *attC* sites in the PCR primers) is novel, and offers a general-purpose method for rapidly constructing gene cassettes from any desired target gene. These experiments showed that the foreign genes were integrated into the genetic backbone of the integron and that the encoded enzymes were incorporated into the existing metabolic framework of the cell. This provides evidence to support the idea that integrons are general-purpose agents of bacterial evolution (Rowe-Magnus & Mazel, 2001), and also supports the idea of integrons as general-purpose tools for biotechnology (Bikard et al., 2010). Integron-based cloning as we have applied it has many advantages: it can be used with wild-type environmental bacteria, it does not require a vector, it can yield high frequencies of recombinants in the right circumstances, and the recombinants are stable in the absence of selection. The lack of a cloning vector and associated antibiotic resistance genes is advantageous from a biosafety perspective, since the amount of foreign DNA is minimized. We anticipate that elements of the method we have outlined here will be useful for cloning genes in diverse bacteria, and may offer particular benefits as a cloning system in bacteria that currently lack effective plasmid or phage vectors.

The DNA preparations that we used to transform cells contained molecules of varied topology and size. The recombinogenic molecules in the ELAN mixture may differ in experiments using electroporation versus natural transformation, since in the latter case, linear dsDNA is thought to be the form preferentially taken up (Carlson et al., 1983) rather than circular DNA. In the case of circular gene cassettes, the mechanism by which ssDNA (the active recombinogenic intermediate) is formed after electroporation is unclear: it could be due to endogenous nucleases in the bacteria, or due to extrusion of *attC* sites from the dsDNA (Loot et al., 2010). The MDA-derived cassettes were naturally transformed and integrated at much higher efficiencies than the ELAN preparations. This may be because natural transformation of linear dsDNA would yield ssDNA (Johnsborg et al., 2007), which is the recombinationally active form of the gene cassette (Bouvier et al., 2005). The presence of ssDNA in the cell due to the natural transformation process could also trigger the expression of the *intI* gene via the SOS response (Baharoglu et al., 2010; Guerin et al., 2009) – this possibility needs to be tested.

Integration nearly always occurred at the *attI* site, regardless of whether the gene cassettes contained promoters. We found no evidence for *attC*×*attC* reactions between incoming cassettes and resident integrons. Our results differ from earlier *in vitro* and *in vivo* studies, where both *attC*×*attI* and *attC*×*attC* reactions were seen (Dubois et al., 2007; Collis et al., 2002, 2001; Demarre et al., 2007). We believe that our data measured in the wild-type integron hosts more accurately reflect the true
recombination preferences of the integron system. The few non-attI strains detected may represent attC×attC events outside the PCR range, insertion at secondary sites, insertion at attI sites in unknown integrons, or non-integron-mediated recombination. Cloning of the activated reporter genes and sequencing are needed to determine the insertion locations and likely mechanisms in these cases. It was notable that in the current study, the assay for integron recombination was not strongly influenced by other MGEs such as insertion sequence elements, unlike a previous method developed in PstQ (Coleman & Holmes, 2005). The truly non-replicative nature of the gene cassette DNAs in the current study is one reason for this difference.

The Lac+ P. stutzeri recombinants were obtained at much lower frequencies compared with GmR recombinants with the corresponding MDA-derived gfp–aadB template. This could be due to the minor changes in the attC sequence engineered in PCR of the type G cassette DNAs, but we believe it is more likely that this low frequency was due to the need for a secondary mutation to give good growth on lactose. Since P. stutzeri lacks galactose catabolic enzymes, this sugar would accumulate in lacZ-expressing P. stutzeri cells given lactose. We propose that the Lac+ cells obtained in these experiments contain a secondary mutation, allowing them to remove, modify or consume galactose. Growth inhibition due to galactose accumulation has been seen in other bacteria engineered with the lac genes (Pries et al., 1990).

In conclusion, we have developed a versatile approach for detecting and quantifying integron-mediated recombination and expression, which also provides a new way to use integrons for molecular cloning and metabolic engineering.

ACKNOWLEDGEMENTS

We thank Hatch Stokes, University of Technology Sydney, for pMAQ105, pBBR1MCS-2, pSU2056 and R388, Johannes Sikorski, DSMZ, for P. stutzeri strains, Peter Reeves, The University of Sydney, for E. coli P801, and Lars Hansen, University of Copenhagen, for pTGF-2. We thank Andrew Holmes for co-supervision of A. M. G., and Neil Wilson and Sasha Tetu for initial mapping of the P. stutzeri integrons. Preliminary work with metabolic integrons was supported by ARC Discovery grant DP0450056 (Michael Gillings, Hatch Stokes and Andrew Holmes).

REFERENCES


