Analysis of RecA-independent recombination events between short direct repeats related to a genomic island and to a plasmid in *Escherichia coli* K12

María F. Azpiroz* and Magela Laviña

Sección Fisiología y Genética Bacterianas, Facultad de Ciencias, Iguá 4225, Montevideo 11.400, Uruguay

* Corresponding author:
María Fernanda Azpiroz
E-mail: fernanda@fcien.edu.uy
Tel. 598 25258618 ext. 233
Fax. 598 25258629

Abstract

RecA-independent recombination events between short direct repeats, leading to deletion of the intervening sequences, were found to occur in two genetic models in the *Escherichia coli* K12 background. The first model was a small *E. coli* genomic island which had been shown to be mobile in its strain of origin and, when cloned, in the *E. coli* K12 context too. However, it did not encode a site-specific recombinase as mobile genomic island usually do. Then, it was deduced that the host cells should provide the recombination function. This latter was searched for by means of a PCR approach to detect the island excision in *E. coli* K12 mutants affected in a number of recombination functions, including the 16 *E. coli* K12 site-specific recombinases, the RecET system, and multiple proteins that participate in the RecA-dependent pathways of homologous recombination. None of these appeared to be involved in the island excision. The second model, analyzed in a RecA deficient context, was a plasmid construction containing a short direct repeat proceeding from *Saccharomyces cerevisiae*, which flanked the *cat* gene. The excision of this gene by recombination of the DNA repeats was confirmed by PCR and through the detection, recovery and characterization of the plasmid deleted form.

In sum, we present new evidence on the occurrence of RecA-independent recombination events in *E. coli* K12. Although the mechanism underlying these processes is still unknown, their existence suggests that RecA-independent recombination may confer mobility to other genetic elements, thus contributing to genome plasticity.

Introduction

Bacterial genomes contain short repeated DNA sequences, which are prone to recombine and generate genetic rearrangements. Some of them are associated with mobile genetic elements such as prophages and genomic islands, whose insertion generates short, direct repeats on either side of the element. These flanking repeats, known as attachment sites, are usually the target of dedicated site-specific recombinases that excise these elements so that they remain as independent covalently closed circular molecules. In the inverse reaction, the same recombinases integrate the circular phage or island into the corresponding attachment site in the genome (Hallet & Sherratt, 1997; Hochhut et al., 2006; Lin et al., 2008; Murphy & Boyd, 2008; Juhas et al., 2009).

The present work continues previous studies on the mobility of a genomic island, called H47 (H47 GI), contained in the chromosome of an *Escherichia coli* strain (H47). It is a small 13 kb-genetic element devoted to the production of the antibacterial peptide microcin H47 (Laviña, Gaggero & Moreno, 1990; Poey, Azpiroz & Laviña, 2006; Azpiroz, Bascuas & Laviña, 2011). The H47 GI was found to be an unstable mobile element, able to excise from the chromosome by recombination between its attachment sites (*attL* and *attR*) in a process generating two products: the chromosome without the island and the island as an independent covalently closed circular molecule (Figure 1A). A DNA segment containing the island had been cloned into several multi-copy plasmid vectors and introduced into *E. coli* K12 cells. In this foreign background, the H47 GI retained its mobility as in the original H47 strain. In addition, it was found that the excised island was able to integrate into another plasmid provided that it contained the corresponding attachment site. These findings were attained by means of PCR experiments, which allowed the detection of both recombinant sequences generated from each type of recombination event, excision and integration (Azpiroz, Bascuas & Laviña, 2011). The H47 GI is flanked by two extensive and imperfect direct repeats, *attL*

and *attR*, which share four main regions of homology (Figure 1B). In the case of excision, *attL* and *attR* recombine and the two resulting products were detected by the appearance of amplicons that included the recombined attachment sites: *attC* in the deleted replicon and *attI* in the excised H47 GI (Figure 1C). Amplicons containing the *attC* site were a mix of four sequences corresponding to recombination at the four main regions of homology between *attL* and *attR*. This recombination pattern remained unchanged when the analysis was carried out in RecA-deficient *E. coli* K12 cells, indicating that the H47 GI mobility is a process independent from the homologous recombination pathway (Figures 1C and S1) (Azpiroz, Bascuas & Laviña, 2011).

Many reports refer that genomic islands usually undergo mobility events, and that these are catalyzed by site-specific recombinases encoded by the islands themselves, a feature that is shared with temperate phages (Hochhut et al., 2006; Juhas et al., 2009). However, the H47 GI does not encode any recombinase: when most of its DNA was deleted, recombination between its attachment sites remained unchanged (Azpiroz, Bascuas & Laviña, 2011). Therefore, the genomic background of *E. coli* H47 and of *E. coli* K12 should provide the recombination function responsible for the H47 GI mobility. The island would thus employ a sort of parasitic strategy and, given the characteristics of the rearrangements observed, a site-specific recombination process was expected to accomplish the genetic exchange.

In this work, we searched for the function underlying the H47 GI mobility, working with the cloned island in *E. coli* K12 cells. The H47 GI excision was PCR-assayed in a broad collection of mutants affected in the site-specific and homologous recombination pathways encoded by *E. coli* K12. None of these functions affected the island mobility, indicating that the process under study could be included into the poorly-known group of RecA-independent recombination events (Bzymek & Lovett, 2001). However, we could not succeed to obtain a

clone carrying the deleted form with the *attC* site. To gain further insight into this type of process, a plasmid construction containing the repeated FRT (Flp recognition target) site from *Saccharomyces cerevisiae* was similarly analyzed in the context of *E. coli* K12. In this model, RecA-independent deletion of the intervening sequence between the repeats was detected in vitro by PCR and in vivo by the appearance of the deleted form in sufficient amounts to be recovered from plasmid DNA preparations.

Materials & Methods

Bacterial strains, plasmids, and growth conditions

E. coli K12 BZB1011 is a spontaneous nalidixic acid resistant derivative of W3110, a strain whose genome has been sequenced (GenBank: AP_009048). Most strains used are BZB1011 derivatives, carrying deletions proceeding from selected mutants of the Keio Collection (Baba et al., 2006). In these latter, gene deletions with replacement by a kanamycin resistance (*kan*) cassette were introduced into BZB1011 by P1 transduction, selecting the kanamycin resistant clones (Table S1). When required, double deletion mutants were constructed in two steps: the kanamycin cassette of the first mutated gene was eliminated as indicated previously (Datsenko & Wanner, 2000) and the cured mutant was then transduced with a P1 grown on the second deletion mutant, selecting the kanamycin resistant clones. *E. coli* K12 JC8679 is a *recB recC sbcA* mutant (Coli Genetic Stock Center strain 6490). Two mutant derivatives of BW25113, the parent *E. coli* K12 strain for the Keio Collection, were used: BW25113 *recA56* and BW25113 *topB recA56*, this latter bearing the *ΔtopB761::kan* mutation from the Keio Collection. The *recA56* allele was introduced by transducing BW25113 and BW25113 *topB* (JW1752) with P1 grown on a *srl::*Tn*10 recA56* strain from our laboratory collection,

selecting the tetracycline resistant clones and then assaying the *recA* defficiency by testing their UV sensitivity.

Plasmid pEX2000 is a pBR322 derivative that carries the H47 GI included in a 16,823 bp-chromosomal DNA segment from strain H47 (GenBank: AJ_009631). Plasmid pΔint1carries the same DNA segment but deleted for most of the H47 GI, keeping the attachment sites and the external chromosomal adjacent sequences (Azpiroz, Bascuas & Laviña, 2011). pUY-FRT was constructed by cloning a 1,182 bp HindIII fragment from pKD3 (GenBank: AY_048742) into the pUC13 vector. The insert contains the chloramphenicol resistance gene (*cat*) flanked by the FRT sites from *S. cereviseae*. Since pKD3 requires the *pir* gene product to be provided in trans for its replication (Datsenko & Wanner, 2000), pUY-FRT was constructed to avoid this requirement.

Bacteria were grown in LB medium at 37º C. Antibiotics were added at the following final concentrations: kanamycin, 30 µg/ml; ampicillin, 50 µg/ml, chloramphenicol, 60 µg/ml, and tetracycline, 20 µg/ml.

Plasmid purification, PCR assays and DNA sequencing

Plasmid DNA was extracted with "QIAprep Spin Miniprep Kit" (Qiagen). DNA bands were extracted from gels using "QIAEXII Gel extraction kit" (Qiagen). For PCR reactions, primers and conditions employed are explained in Table 1. For sequencing purposes, PCR products were extracted from gels as explained. Both strands of the OUT and IN amplicons from the pUY-FRT model were sequenced using the primers employed for their amplification. DNA sequencing was performed at the "Molecular Biology Unit" of the Pasteur Institute of Montevideo.

Results

Excision of the H47 genomic island

Recombination underlying the H47 GI mobility was studied through the detection of the island excision from a multi-copy recombinant plasmid, pEX2000, in the context of *E. coli* K12 cells. Plasmid pEX2000 carries a DNA segment from the *E. coli* H47 chromosome containing the H47 GI flanked by its direct repeats and by adjacent chromosomal sequences on both sides. The direct repeats, *attL* (148 bp) and *attR* (144 bp), are imperfect and share four main segments of homology of 13, 11, 20 and 17 bp (Fig. 1B). Recombination between them leads to the excision of the H47 GI, which is assessed by PCR using primers "out" in order to detect the occurrence of the empty *attC* site (amplicon OUT_{H47}) (Figure 1A). The assays were performed in the context of the *E. coli* K12 strain BZB1011and in a set of BZB1011 derivative mutants affected in recombination functions (Table S1). The purpose was to identify the mutant/s where excision did not occur, thus revealing the gene/s involved in the recombination process. Plasmid pEX2000 was introduced into each of these strains and then was extracted to be used as template in PCR reactions. The sequence of the OUT amplicons corresponding to the mutant strains was compared with that from the wild-type BZB1011 strain, i.e. with a mix of four overlapping sequences with the predominance of that corresponding to recombination in site III (Figure S1).

Considering that genomic islands usually employ site-specific recombinases to mediate their mobility, the search first concentrated in mutant strains deficient for such enzymes. Although the H47 GI attachment sites did not reveal significant homology with any known target for a site-specific recombinase, we did not discard this possibility since it has been reported that some of these enzymes are able to work -although with low efficiency- on secondary sites that differ in their sequence from the primary one (Menard & Grossman, 2013). Therefore, the

166 excision of the H47 GI was analyzed in mutants deficient for each of the 16 site-specific
167 recombinases encoded by the *E. coli* K12 chromosome, most of them related to prophages
168 (*intF, intD, ybcK, intE, pin, intR, pinR, pinQ, intQ, intS, intA, xerD, xerC, intB, fimB* and
169 *fimE*). Two prophage genes coding for site-specific recombinases, *pinQ* and *pinR*, are highly
170 similar (99% of identity) and for that reason the double mutant *pinQ pinR* was constructed
171 and included in the assays. In all cases the OUT_{H47} amplicon appeared and always revealed
172 the same pattern of overlapping sequences as that proceeding from the wild-type BZB1011
173 strain (data not shown).

174 Given the previous results, the involvement of site-specific recombination in the H47 GI
175 mobility was discarded. An important fact was that recombination always took place in
176 regions of total homology between the attachment sites, indicating that this condition would
177 be necessary. Therefore, we examined the possible involvement of several genes that
178 participate in homologous recombination.

179 First, the RecA-independent RecET homologous recombination pathway, encoded by the
180 defective Rac prophage present in *E. coli* K12, was analyzed (Kolodner, Hall & Luisi-
181 DeLuca, 1994). It was considered that the *recET* genes might have a basal level of expression,
182 as has been proposed for *E. coli* K12 crytic prophage genes (Wang et al., 2010). The excision
183 of the H47 GI was assayed in BZB1011 mutant derivatives *recE* (deficient for exonuclease
184 VIII) and *recT* (deficient for the recombinase function). None of these contexts affected the
185 recombination pattern. The same result appeared using the *recB recC sbcA* strain JC8679, a
186 genetic background where the RecET pathway is known to be induced (Kolodner, Hall &
187 Luisi-DeLuca, 1994).

188 It was also considered that recombination under study, although RecA-independent, could
189 share some functions with the homologous recombination pathways. Following this idea, the

190 excision of the H47 GI was surveyed in a set of mutants defective for proteins that participate
191 in homologous recombination (*recB, recC, recD, recF, recO, recR, recN, recJ, recQ, recG,*
192 *sbcB, sbcC, sbcD, uvrD, ruvA, ruvB, ruvC, radA, recX, exoX, rarA, seqA* and *helD*). The
193 double mutant *recT recA*, deficient for all known homologous recombination pathways, was
194 constructed and also included in the assays. The result was that in neither mutant the excision
195 of the island was detectably affected: the OUT_{H47} product always appeared as a weak band in
196 gels and its sequence exhibited the same pattern as that of the amplicon from the wild-type
197 strain (data not shown). Therefore, none of the functions analyzed related to homologous
198 recombination appeared to be involved in the H47 GI mobility.

199 **In vivo search for the H47 GI excision**

200 Although in vitro PCR experiments consistently supported the existence of recA-independent
201 recombination events between the *attL* and *attR* sites flanking the H47 GI, this phenomenon
202 still awaited an in vivo confirmation. This would imply the detection and characterization of
203 at least one of the recombination products, the deleted molecule and/or the excised circular
204 form. Obviously, the deleted form appeared as more likely to be detected because of its
205 replicative condition. However, the model of the H47 genomic island was rather complex
206 since it encodes the production of an antibiotic activity, which would exert a deleterious effect
207 upon cells that had lost the island in a recombination event. For this reason, we investigated
208 this phenomenon not only in strains carrying the entire island (pEX2000) but also in strains
209 with a plasmid derivative in which most of the intervening sequence between the *att* sites had
210 been deleted (pΔint1).
211 After being propagated in *E. coli* K12 cells (BZB1011 *recT recA*), plasmids were extracted,
212 and important amounts of DNA (1-2 μg) were run in gels in an attempt to perceive new bands

that could correspond to the recombination products. We did not succeed to see anything other than the plasmids´ original forms.

It has been described that *E. coli* mutants affected in the *topB* gene, encoding topoisomerase III, have an increased frequency of spontaneous deletions occurring at short direct repeats (Whoriskey, Schofield & Miller, 1991; Schofield et al., 1992; Uematsu, Eda & Yamamoto, 1997). Therefore, we repeated the assays in the context of strains carrying a *topB* deficient allele. BZB1011 *topB* was constructed, but its *recA* derivative resulted nonviable. Considering that the *topB* context would be affected in DNA supercoiling, the nalidixic acid resistant strain BZB1011 might not be suitable for this study since it most probably contains a mutant DNA gyrase. For this reason, the *gyr* wild type BW25113 context was employed. BW25113 *recA* and BW25113*topB recA* were constructed and transformed with pEX2000 and pΔint1. In all these strains, although recombination between the *att* H47 sites was PCR-confirmed, no extra-bands appeared when plasmids´ DNA was analyzed in gels.

H47 GI-related *attC* sequence survey in data banks

Since no product of the H47 GI excision could be detected in vivo, a clue in this sense was searched for in data banks. This type of survey had been done before, mainly focusing on the presence of the H47 GI in other strains besides *E. coli* H47 (Azpiroz, Bascuas & Laviña, 2011). Now, we looked for the four types of *attC* sequences (I-IV), corresponding to recombination events in each of the four segments of identity between the *attL* and *attR* sites. The search revealed that these sequences are widespread among pathogenic *E. coli* strains and also appear in two strains of *Salmonella enterica*, always being present in the chromosome. Although these sequences exhibited a certain degree of variability, several *attC* sites could be recognized and distinguished from *attL* and *attR*. They were further confirmed as *attC* by analyzing their adjacent sequences (about 500 nucleotides on each side), which should

237 correspond to those surrounding the H47 GI. Most of them were *attC*$_{III}$, resulting from
238 recombination at the most extensive site of homology between the direct repeats. Among
239 these, six identical matches were found (*E. coli* strains Sanji, SEC470, RS76, PCN033, Santai
240 O157:H16, and *S. enterica* Heidelberg str. SL476). There was a match with *attC*$_{II}$ (*E. coli*
241 ETEC H10407) and another single match with *attC*$_{IV}$ (*E. coli* FHI23). No sequence of the
242 *attC*$_{I}$ type appeared. Therefore, these findings strongly supported the idea that the H47 GI is
243 indeed a mobile element due to recombination between its attachment sites.

244 **RecA-independent recombination in the model of plasmid pUY-FRT**

245 In view that our efforts to detect in vivo the H47 GI excision were unsuccessful, the studies
246 were extended to a simpler genetic model. It was a multi-copy plasmid, pUY-FRT, carrying
247 the *cat* gene (for chloramphenicol resistance) flanked by a perfect 46 bp-direct repeat. The
248 repeated sequences proceeded from *S. cereviseae* and contained the FRT target for the site-
249 specific recombinase Flp. Recombination between the repeats would lead to the deletion of
250 the intervening *cat*-containing DNA segment. The assays were carried out in the absence of
251 the Flp enzyme in *E. coli* K12 cells. Specifically, plasmid pUY-FRT was propagated and then
252 extracted from BZB1011 *recT recA*, BW25113 *recA* and BW25113 *topB recA*.
253 The experiments began by transforming with pUY-FRT the three *E. coli* K12 mutant
254 backgrounds and chloramphenicol resistant transformants were selected so as to ensure that
255 the assays started with clones carrying the original non-deleted form of the plasmid. After
256 isolation in the same medium, clones were grown in liquid LB Ap, which was supposed to be
257 a permissive condition for the propagation of the deleted plasmid.

258 First, plasmid DNA was used as template in PCR reactions devoted to detect the excision of
259 the DNA segment between the repeats, i.e. the *cat* gene. As with the H47 GI, two types of
260 reactions were carried out to detect the two possible products: an OUT$_{FRT}$ amplicon of 381 bp

261 in the deleted plasmid, and an IN_{FRT} amplicon of 335 bp in the excised circular form. In both
262 types of reactions, amplicons of the expected size were produced and appeared in the three
263 contexts assayed (data not shown). The DNA sequence of these PCR products exhibited the
264 expected recombination pattern, i.e. a single repeat flanked by recombined sequences (data
265 not shown).

266 For in vivo analysis, ca. 500 ng of plasmid DNA were loaded in gels. This time, a new small
267 band was barely seen in plasmid preparations proceeding from the *topB*$^+$ backgrounds, being
268 more evident in the BW25113 *recA* context, and this same band appeared clearly stronger
269 when the plasmid came from the *topB* deficient cells. We presumed that it could correspond
270 to the deleted plasmid molecule, which would be called pUY-FRTΔ (Figure 2A). Then, this
271 smaller band, proceeding from BW25113 *recA* and BW25113 *topB recA*, was extracted from
272 gels and used to transform BW25113 *recA* with selection for ampicillin-resistance. In each
273 case, a few clones grew and, when assayed in the presence of chloramphenicol, all proved to
274 be sensitive. Plasmid DNA was extracted from a clone of each type of transformant and was
275 then digested with EcoRI and HindIII. Both plasmids were found to be identical in their size
276 and restriction profile, which corresponded to those of a deleted plasmid derived from pUY-
277 FRT by recombination between its direct repeats (Figure 2B). DNA sequencing confirmed
278 this structure: the sequence was read through the recombined repeat and beyond it more than
279 300 bp on each side.

280 Finally, it should be mentioned that no trace of the excised DNA could be seen.

281

282

283

Discussion

The analysis of the mobility of an *E. coli* small genomic island led us to study RecA-independent recombination phenomena. Specifically, the H47 GI presented the peculiarity of being able to excise from its site of insertion although it did not code for a cognate integrase. Moreover, its mobility was not affected in a RecA deficient background (Azpiroz, Bascuas & Laviña, 2011). It should be mentioned that a few other genetic elements have been described to be mobile and to lack a recombinase gene (Santiviago et al., 2010; Palmiere, Mingoia & Varaldo, 2013).

The fact that the H47 GI mobility could be analyzed in the *E. coli* K12 genetic context appeared as an excellent opportunity to identify the host function involved in the genetic exchange. The availability of the genome sequence of *E. coli* K12 as well as of an exhaustive collection of derivative deletion mutants encouraged this presumption. However, after assaying the H47 GI excision in a number of mutants affected in recombination functions, including the 16 *E. coli* K12 site-specific recombinases, the RecET system, and multiple proteins that participate in the RecA-dependent pathways of homologous recombination, it was concluded that none of these processes appeared to be involved and that an unknown RecA-independent recombination mechanism would be responsible for the H47 GI mobility. Anyway, our results could harbor some uncertainty since the PCR technique has been described to generate chimeric products when the template contains repeated sequences, an artifact that has been detected in the amplification of 16S rRNA genes in metagenomic studies (Wang & Wang, 1997). Although this possibility seems rather unlikely, we cannot completely discard it.

Therefore, an in vivo demonstration of the existence of the H47 GI mobility was searched for. This was particularly difficult given the lack of selection for the recombined clones and that

the frequency of recombination appeared to be low. In fact, we were unable to obtain a clone carrying a recombination product, i.e. an *attC* site in a replicon that had previously carried the H47 GI.

Other authors have analyzed the mobility of some genetic elements that lack a recombinase gene by means of PCR experiments similar to those employed in this work (Palmiere, Mingoia & Varaldo, 2013). These elements also have in common that their attachment sites are very extensive, of several hundred bp. In two cases, a defective prophage-like element from *S. enterica* serovar Enteritidis and a transposon from *Streptococcus pneumonia*, clones that spontaneously lost these elements could be isolated. Interestingly, their excision was PCR-detected in a RecA-deficient background (Santiviago et al., 2010; Palmieri et al., 2012). In view of these previous descriptions, we presume that the length of the attachment sites and their level of identity may be an important factor that determines the frequency at which recombination takes place and, thus, the likelihood of its in vivo detection. The two mentioned examples from *S. enterica* and *S. pneumonia* have attachment sites of 308 bp (98% of identity) and ca. 1200 bp (>99% of identity), respectively. For their part, the H47 GI repeats have clearly smaller numbers: 146/8 bp in length with 77% of identity, the most extensive perfect homology being of 20 bp (see Figure 1B). In other previous descriptions of RecA-independent recombination events, it was also found that homology was required and that recombination rates increased with the length of the repeats (Albertini et al., 1982; Bi & Liu, 1994). Another factor that could hinder the detection of recombined molecules was that the H47 GI encodes an antibacterial activity, microcin H47, together with it cognate immunity. Those cells harboring deleted molecules by loss of the island would be in a clear disadvantage compared to other cells because they would be more sensitive to the microcin.

Nevertheless, when we worked with a plasmid construction lacking most of the island content, recombinant molecules could not be detected either.

Finally, a survey performed in data banks searching for H47 GI-related *attC* sequences revealed that these empty attachment sites exist in natural strains. Representatives of three of the four possible versions of recombination between the direct repeats were found. These findings show that the H47 GI is indeed a mobile genetic element whose presence in the bacterial chromosome is unstable. Although these empty attachment sites would represent traces of recombination events between the direct repeats, they do not ensure the widespread occurrence of this type of process in the *E. coli* species: the possibility that recombination could only take place in some genetic backgrounds providing a specific recombinase cannot be discarded.

At this point, we wondered how general this recombination phenomenon could be and then decided to study a different genetic model, plasmid pUY-FRT, in the context of *E. coli* K12. This model is much simpler than the H47 GI and contains perfect and longer direct repeats of 46 bp flanking the *cat* gene. As in the case of the island, excision events were detected by a PCR approach in wild type and in *recA*- and *recT*-deficient contexts, and the genetic exchange always occurred at the pUY-FRT repeats. However, unlike what happened with the island, this model allowed the detection of a deleted derivative molecule which had lost the *cat* gene by recombination between the direct repeats. In fact, this is probably the first time that a RecA-independent recombination product is readily detected in a plasmid preparation from a culture that had been seeded with a strain carrying the original plasmid form. This phenomenon appeared to be enhanced in a *topB* mutant, deficient for topoisomerase III, a context that has been previously described to increase some RecA-independent recombination events (Whoriskey, Schofield & Miller, 1991; Schofield et al., 1992; Uematsu, Eda &

Yamamoto, 1997).Therefore, in the model of pUY-FRT we succeeded to demonstrate that a RecA-independent recombination mechanism is able to work on short direct repeats in *E. coli* K12. These results support the idea that similar events could indeed happen in the H47 GI but, in this case, their frequency would be too low to be detected in vivo. There are several differences between the two models analyzed that could determine differences in their excision rate. For instance, it has been claimed that the longer the extent of perfect homology between the repeats and the shorter the distance separating them, the higher the rate of recombination (Albertini et al., 1982; Bi & Liu, 1994; Lovett et al., 1993; Lovett et al., 1994). These two factors could explain the higher frequency of excision in pUY-FRT in relation to that in H47 GI. pUY-FRT has 46 bp-long perfect direct repeats separated by an intervening sequence of 886 bp, while the H47 GI has perfect repeats of 13, 11, 20 and 17 bp, separated by 12,635 bp. Apparently, the condition of the DNA repeats would be predominant, considering that recombination between the H47 GI repeats in plasmid pΔint1 could not be detected in vivo even though the intervening sequence was shortened to 554 bp, and even in a *topB* background.

Following the results attained by PCR experiments, the deleted material in the excision events would remain as a non-replicative circle, as has been reported for other genetic elements that exhibit RecA-independent mobility (Palmieri, Mingoia & Varaldo, 2013). However, we could not detect the excised form in vivo in none of the two models analyzed. Obviously, this aspect deserves more studies.

In sum, we provide new evidence on the occurrence of spontaneous events of excision in *E. coli* K12 which are not mediated by the known mechanisms of recombination. The rearrangements under study appear to be related to previous descriptions of RecA-independent recombination events in *E. coli* K12 (Azpiroz, Bascuas & Laviña, 2011; Bzymek

& Lovett, 2001; Kingston et al., 2015) and in other bacterial organisms (Santiviago et al., 2010; Palmieri et al., 2012). This type of genetic exchange occurs between regions of homology, including very short ones on which RecA-dependent recombination would not be able to work or would not be efficient. Since its frequency appears to be particularly low, PCR-detection of the recombination products is being routinely used by several authors, while obtaining clones carrying a recombined form has been achieved in very few cases. In this work we present PCR-based evidence of RecA-independent recombination in two genetic models while in vivo evidence of this phenomenon was attained in only one of them. Finally, it should be kept in mind that if RecA-independent recombination were able to operate on a wide spectrum of repeated sequences, including very short ones, then it would provide mobility to different genetic elements, thus broadening the repertoire of possible rearrangements in the bacterial cell.

Acknowledgements

We are indebted to María Parente for excellent technical assistance. This work was supported by the Agencia Nacional de Investigación e Innovación FCE_2_2011_1_6019 and by Programa de Desarrollo de las Ciencias Básicas, Uruguay.

References

1. Albertini AM, Hofer M, Calos MP, Miller JH. 1982. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. *Cell* 29: 319-328.

2. Azpiroz MF, Bascuas T, Laviña M. 2011. Microcin H47 system: an *Escherichia coli* small genomic island with novel features. *PloS One* 6: e26179.

3. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2: 2006.0008

4. Bi X & Liu LF. 1994. *recA*-independent and *recA*-dependent intramolecular plasmid recombination: differential homology requirement and distance effect. *J Mol Biol* 235: 414-423.

5. Bzymek M & Lovett ST. 2001. Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proc Natl Acad Sci* 98: 8319-8325.

6. Datsenko KA & Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97: 6640-6645.

7. Hallet B & Sherratt DJ. 1997. Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. *FEMS Microbiol Rev* 21: 157-178.

8. Hochhut B, Wilde C, Balling G, Middendorf B, Dobrindt U, Brzuszkiewicz E, Gottschalk G, Carniel E, Hacker J. 2006. Role of pathogenicity island-associated

integrases in the genome plasticity of uropathogenic *Escherichia coli* strain 536. *Mol Microbiol* 61:584-595.

9. Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. 2009. Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* 33: 376-393.

10. Kingston AW, Roussel-Rossin C, Dupont C, Raleigh EA. 2015. Novel *recA*-independent horizontal gene transfer in *Escherichia coli* K-12. *PloS One* 10: e0130813.

11. Kolodner R, Hall SD, Luisi-DeLuca C. 1994. Homologous pairing proteins encoded by the *Escherichia coli recE* and *recT* genes. *Mol Microbiol* 11: 23-30.

12. Laviña M, Gaggero C, Moreno F. 1990. Microcin H47, a chromosome-encoded microcin antibiotic of *Escherichia coli*. *J Bacteriol* 172: 6585-6588.

13. Lin T, Lee C, Hsieh P, Tsai S, Wang J. 2008. Characterization of integrative and conjugative element ICE*Kp1*-associated genomic heterogeneity in a *Klebsiella pneumoniae* strain isolated from a primary liver abscess. *J Bacteriol* 190: 515-526.

14. Lovett ST, Drapkin PT, Sutera VA, Gluckman-Peskind TJ.1993. A sister-strand exchange mechanism for *recA*-independent deletion of repeated DNA sequences in *Escherichia coli*. *Genetics* 135: 631-642.

15. Lovett ST, Gluckman TJ, Simon PJ, Sutera Jr VA, Drapkin PT. 1994. Recombination between repeats in *Escherichia coli* by a *recA*-independent, proximity-sensitive mechanism. *Mol Gen Genet* 245: 294-300.

16. Menard KL & Grossman AD. 2013. Selective pressures to maintain attachment site specificity of integrative and conjugative elements. *PLoS Genet* 9: e1003623.

17. Murphy RA & Boyd EF. 2008. Three pathogenicity islands of *Vibrio cholerae* can excise from the chromosome and form circular intermediates. *J Bacteriol* 190: 636-647.

18. Palmieri C, Mingoia M, Massidda O, Giovanetti E, Varaldo PE. 2012. *Streptococcus pneumoniae* transposon Tn*1545*/Tn*6003* changes to Tn*6002* due to spontaneous excision in circular form of the *erm*(B)- and *aphA3*-containing macrolide-aminoglycoside-streptothricin (MAS) element. *Antimicrob Agents Chemother* 56: 5994-5997.

19. Palmieri C, Mingoia M, Varaldo PE. 2013. Unconventional circularizable bacterial genetic structures carrying antibiotic resistance determinants. *Antimicrob Agents Chemother* 57: 2440-2441.

20. Poey ME, Azpiroz MF, Laviña M. 2006. Comparative analysis of chromosome-encoded microcins. *Antimicrob Agents Chemother* 50: 1411-1418.

21. Santiviago CA, Blondel CJ, Quezada CP, Silva CA, Tobar PM, Porwollik S, et al. 2010. Spontaneous excision of the *Salmonella enterica* serovar Enteritidis-specific defective prophage-like element φSE14. *J Bacteriol* 192: 2246-2254.

22. Schofield MA, Agbunag R, Michaels ML, Miller, JH. 1992. Cloning and sequencing of *Escherichia coli mutR* shows its identity to *topB*, encoding topoisomerase III. *J Bacteriol* 174: 5168-5170.

23. Uematsu N, Eda S, Yamamoto K. 1997. An *Escherichia coli topB* mutant increases deletion and frameshift mutations in the *supF* target gene. *Mutat Res* 383: 223-230.

24. Wang X, Kim Y, Ma Q, Hong SH, Pokusaeva K, Sturino JM, Wood TK. 2010. Cryptic prophages help bacteria cope with adverse environments. *Nat Commun* 1: DOI: 10.1038/ncomms1146.

25. Wang GCY and Wang Y. 1997. Frequency of formation of chimeric molecules as a consequence of PCR coamplification of 16S rRNA genes from mixed bacterial genomes. *Appl Environ Microbiol* 63: 4645-4650.

26. Whoriskey SK, Schofield MA, Miller JH. 1991. Isolation and characterization of *Escherichia coli* mutants with altered rates of deletion formation. *Genetics* 127:21-30.

500 **Table 1. Conditions of PCR reactions**[(a)]
501

Model	Primer name[(b)]	Primer sequence 5´3´	Annealing temperature	DNA template (ng)	Product name and expected size (bp)
H47 GI excision from pEX2000	out1	CCGTTCATTTTCCTGCTGACCC			
	out2	TCTGTTGCCCGTTGATGTTTCCT	58º	200-400	OUT_{H47} (314-317)
	in1	GTTTGTAGGAGCTTTCTTTTTTG			
	In2	CGCTGATGACTGTTTTTATGTTG	53º	200-400	IN_{H47} (761-764)
cat excision from pUY-FRT	F	GTTGTAAAACGACGGCCAGT			
	R	CACAGGAAACAGCTATGACC	58º	200-300	OUT_{FRT} (381)
	II-in1	AAGGCGACAAGGTGCTGATG			
	II-in2	GGAACCTCTTACGTGCCGAT	58º	200-300	IN_{FRT} (335)

502
503 [(a)] PCR amplifications were performed using U-Taq DNA polymerase (SBS Genetech) in a total volume of 30 µl. Reaction mixes
504 contained 1x buffer, 200 µM of each deoxynucleotide triphosphate, 500 nM of each primer, 1.25 U of DNA polymerase and
505 template DNA. Conditions for amplification were: 2 min at 94°C, 30 cycles of incubation at 94°C for 30 sec, annealing
506 temperature for 30 sec, 72°C for 30 sec, and a final extension step at 72°C for 2 min.
507 [(b)] Primers used to assay the H47 GI model were presented in Azpiroz et al. 2011. F and R are M13 forward and reverse
508 primers. II-in1 and II-in2 were designed in this work.
509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

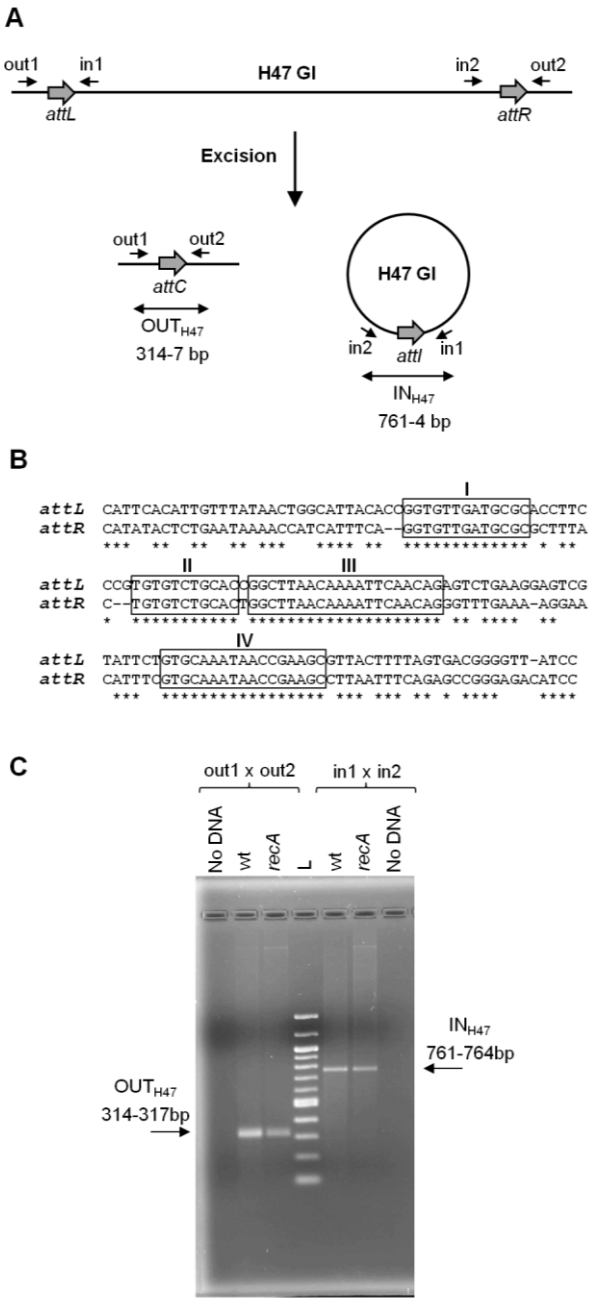


Figure 1. Excision of the H47 genomic island. A, PCR approach performed to detect the excision of the H47 GI. Attachment sites are shown with grey thick arrows and primers with small arrows. Products of excision were detected by the appearance of amplicons obtained with the indicated primer pairs: amplicon OUT_{H47}, containing the *attC* site, and amplicon IN_{H47}, containing the recombined attachment site *attI* in the excised H47 GI. **B,** Alignment of

543 the *attL* and *attR* sites. Identical nucleotides are indicated with asterisks. The four main
544 regions of homology are boxed and named with roman numbers. **C,** Detection of the H47 GI
545 excision using as template pEX2000 propagated in *E. coli* K12 BZB1011 (wt), and in its
546 derivative mutant *recA*. L: 100 pb DNA Ladder (BioLabs).

547

548

549

550

551

552

553

554

555

556

557

558

559

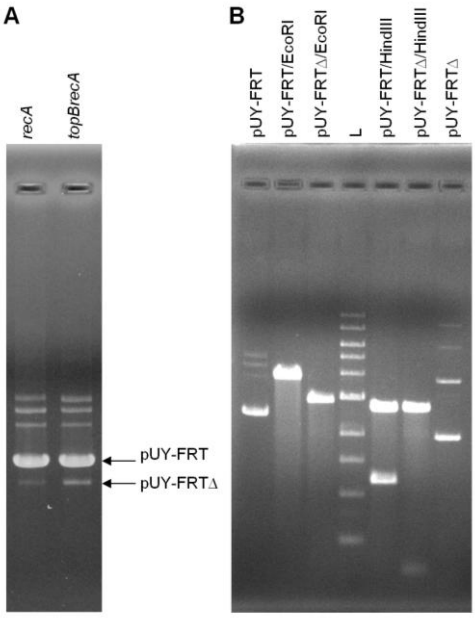
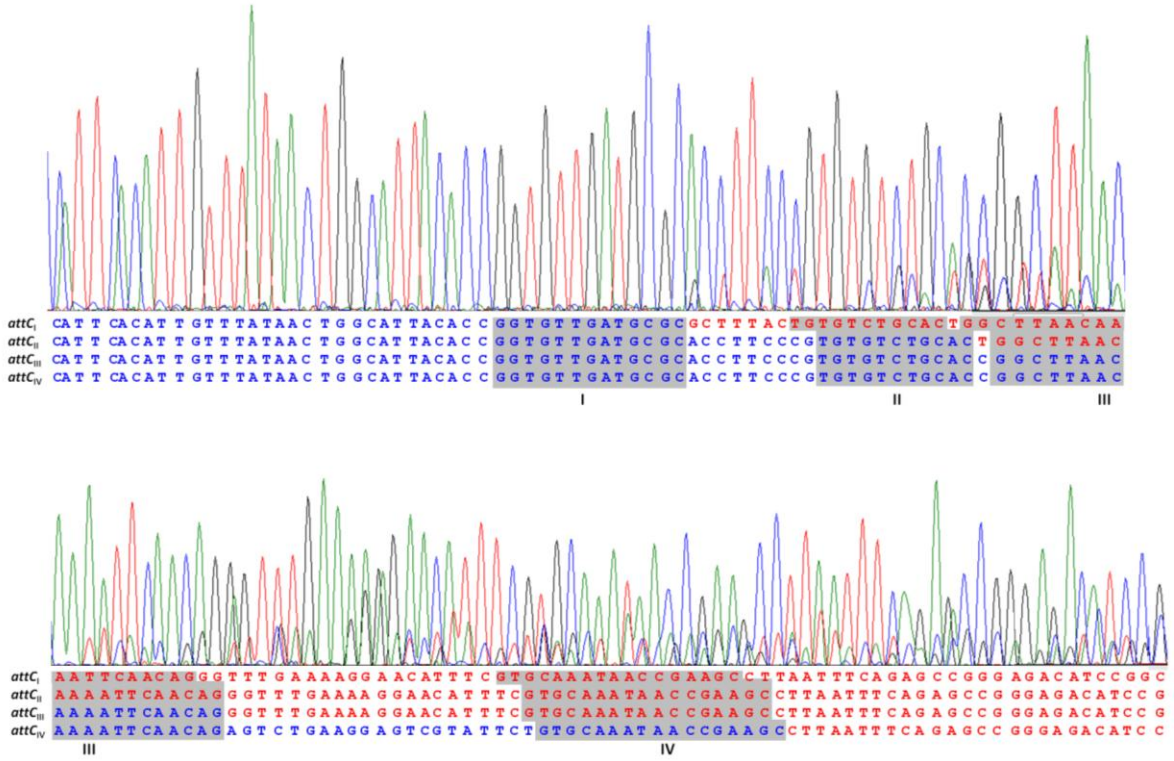


Figure 2. Excision of the *cat* gene in pUY-FRT and generation of pUY-FRTΔ. A, Plasmid DNA extracted from strains BW25113 *recA* (pUY-FRT) and BW25113 *topB recA* (pUY-FRT). Bands corresponding to the original plasmid pUY-FRT (3,864 bp) and to its deletion derivative pUY-FRTΔ (2,932 bp) are indicated with arrows. **B,** Restriction analysis of pUY-FRT and pUY-FRTΔ. L: 1 kb DNA Ladder (BioLabs).

Table S1. Deletion mutations carried by *E. coli* K12 BZB1011 derivative strains

Mutant alleles [a]	Keio knockout strains carrying the allele
ΔintD746::kan	JW0525 [b]
ΔintF725::kan	JW0275 [b]
ΔintS756::kan	JW2345 [b]
ΔpinQ766::kan	JW1538 [b]
ΔpinR737::kan	JW1368 [b]
ΔpinQ766 ΔpinR737::kan	JW1538 and JW1368
ΔybcK752::kan	JW0532 [b]
Δpin-746::kan	JW1144 [b]
ΔintE729::kan	JW1126 [b]
ΔintR772::kan	JW1339 [b]
ΔintQ726::kan	JW1571 [b]
ΔintA777::kan	JW2602 [b]
ΔintB741::kan	JW4227 [b]
ΔxerC757::kan	JW3784 [b]
ΔxerD745::kan	JW2862 [b]
ΔfimB780::kan	JW4275 [b]
ΔfimE781::kan	JW4276 [b]
ΔrecE787::kan	JW1344 [b]
ΔrecT786::kan [d]	JW1343 [b]
ΔrecA774::kan [d]	JW2669 [c]
ΔrecT786 ΔrecA774::kan	JW1343 and JW2669
ΔtopB761::kan	JW1752 [c]
ΔrecB745::kan	JW2788 [c]
ΔrecC747::kan	JW2790 [c]
ΔrecD744::kan	JW2787 [c]
ΔrecF735::kan	JW3677 [c]
ΔrecR776::kan	JW0461 [c]
ΔrecO737::kan	JW2549 [c]
ΔrecJ743::kan	JW2860 [c]
ΔrecN772::kan	JW5416 [c]
ΔrecQ767::kan	JW5855 [c]
ΔrecG756::kan	JW3627 [c]
ΔsbcB780::kan	JW1993 [c]
ΔsbcC761::kan	JW0387 [c]
ΔsbcD762::kan	JW0388 [c]
ΔruvA786::kan	JW1850 [c]
ΔruvB785::kan	JW1849 [c]
ΔruvC789::kan	JW1852 [c]
ΔrecX773::kan	JW2668 [c]
ΔradA785::kan	JW4352 [c]
ΔexoX769::kan	JW1833 [c]
ΔrarA788::kan	JW0875 [c]
ΔseqA735::kan	JW0674 [c]
ΔhelD777::kan	JW0945 [c]
ΔuvrD769::kan	JW3786 [c]

[a] Allele numbers following the Coli Genetic Stock Center
[b] Provided by NBRP (NIG, Japan): E. coli
[c] Provided by the Coli Genetic Stock Center
[d] Deletion mutations confirmed by PCR using primers external to the genes



590
591 **Figure S1. Chromatogram of the OUT_{H47} amplicon using as template pEX2000**
592 **propagated in *E. coli* K12 BZB1011 strain.** DNA sequencing was performed with primer
593 out1. The *attC* portion of the chromatogram is shown. The four overlapping sequences
594 corresponding to the products of recombination at the four segments of identity between the
595 *attL* (blue) and *attR* (red) sites are indicated below ($attC_I$ to $attC_{IV}$). The recombination
596 segments are shaded in grey and named with roman numbers. After each recombination
597 segment, a new sequence appears that overlaps the previous one/s.