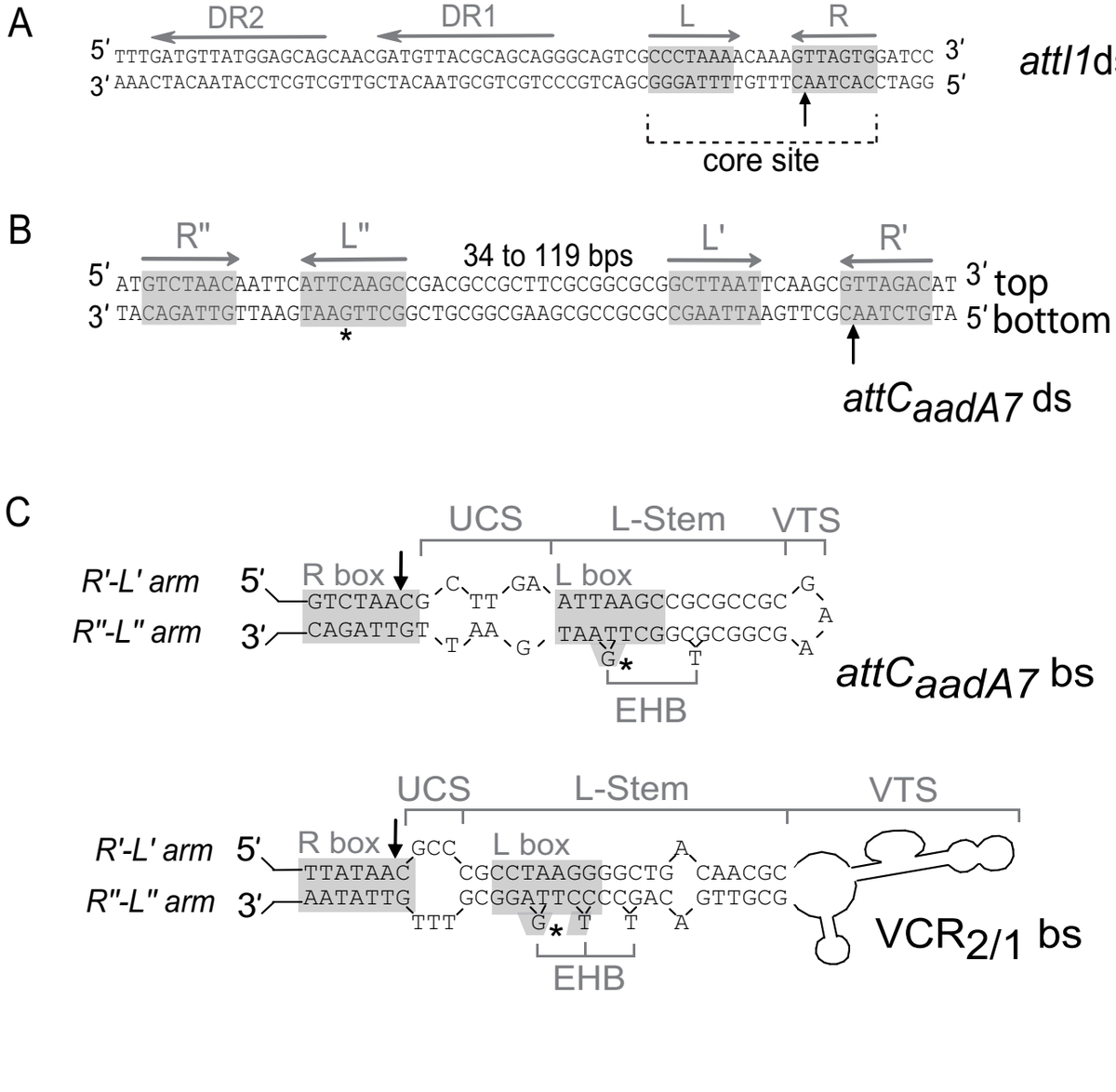

Figure S1: Integron recombination sites

A) Sequence of the double strand (ds) *attI1* site.

B) Sequence of the double strand (ds) $attC_{aadA7}$ site.

C) Proposed secondary structures for single-stranded bottom strands (bs) of $attC_{aadA7}$ and $VCR_{2/1}$ sites.
The inverted repeats L and R, L' and R' and L'' and R'' are indicated with grey arrows. The putative IntI1 binding domains are marked with grey boxes. Vertical arrows indicate cleavage positions. Secondary structures of $attC_{aadA7}$ and $VCR_{2/1}$ sites were determined using the MFOLD web application from the Institut Pasteur. All known folded single-stranded *attC* sites present an almost canonical core site, R and L boxes (formed by the pairing between R''/R' and L''/L') separated by an unpaired central spacer (UCS), two or three extrahelical bases (EHB) and a variable terminal structure (VTS) (12). The asterisk (*) shows the position of the extrahelical G base present in L'' relative to L' that we have shown to be responsible for the strand selectivity (12). The recognition of structural determinants of the *attC* bs folded site by the integrase is in agreement with the structure of the integrase bound to the *attC* x *attC* synaptic complex (13).

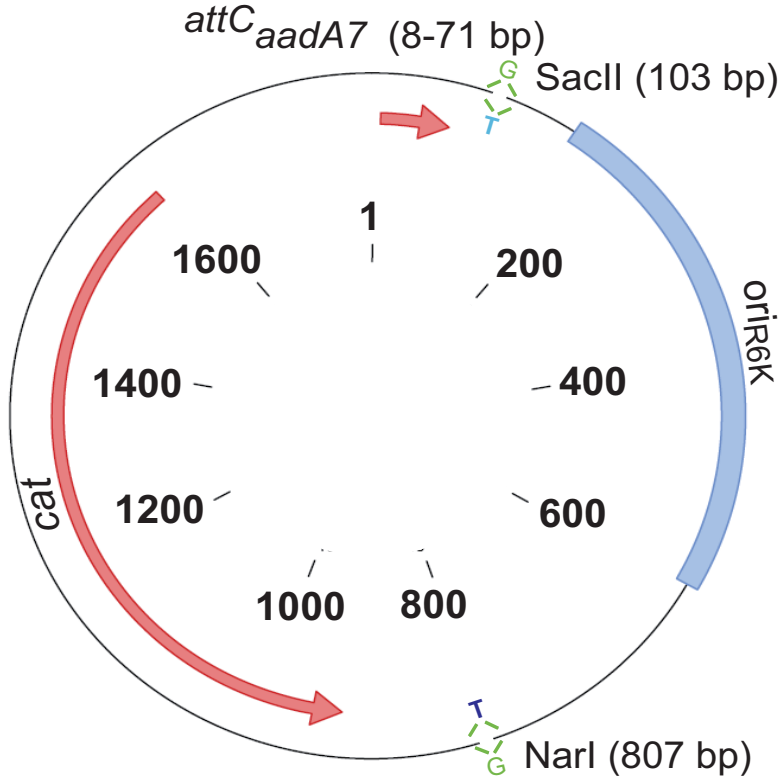

Figure S2: Representation of the mismatched covalent circle
Positions of the *attC*$_{aadA7}$ recombination site (8-71 bp), GT mismatched nucleotides in the SacII (103 bp) and NarI (807 bp) restriction sites, the ori$_{R6K}$ origin of replication and the chloramphenicol resistance gene (*cat*) are indicated.

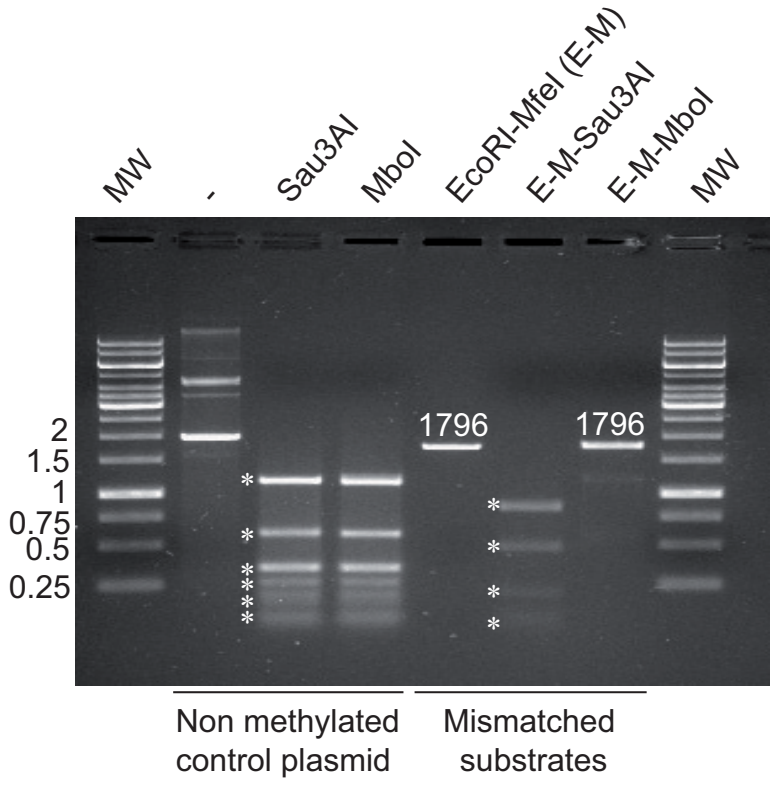

Figure S3: Methylation status of mismatched covalent circles

Non-methylated control plasmid (produced in the GM48 dam- strain) and mismatched substrates were subjected to Sau3AI and MboI restriction digest in order to determine their methylation status. While Sau3AI is not blocked by *dam* methylation, MboI is. Undigested (-), Sau3AI and MboI digested control plasmid are shown in the left panel. The control plasmid was sensitive to both Sau3A and MboI digestions according to its « non-methylated status ». The digestion is expected to generate 21 bands from 8 to 1138 bps but in our running conditions, we could only visualize the six largest bands (1138, 585, 341, 258, 200 and 153 bps).
We tested the mismatched EcoRI and MfeI digested substrates (1796 bps). Note that at this step of mismatched circles production, the *oriFd* has been eliminated but the subtrate is still not circularized (see Materials and methods). The EcoRI-MfeI (E-M) linearized substrates were submitted to Sau3AI and MboI digestion (right panel). As expected, the substrates are sensitive to the Sau3AI digestion but resistant to the MboI digestion. The Sau3AI digestion is expected to generate 15 bands from 5 to 814 bps but in our running conditions, we could only visualize the four largest bands (814, 463, 191 and 84 bps). These results show a Sau3AI sensitivity and a MboI resistance of the mismatched substrates indicating a high degree of methylation of the GATC sequences (22).

MW, Molecular Weight Marker DNA.

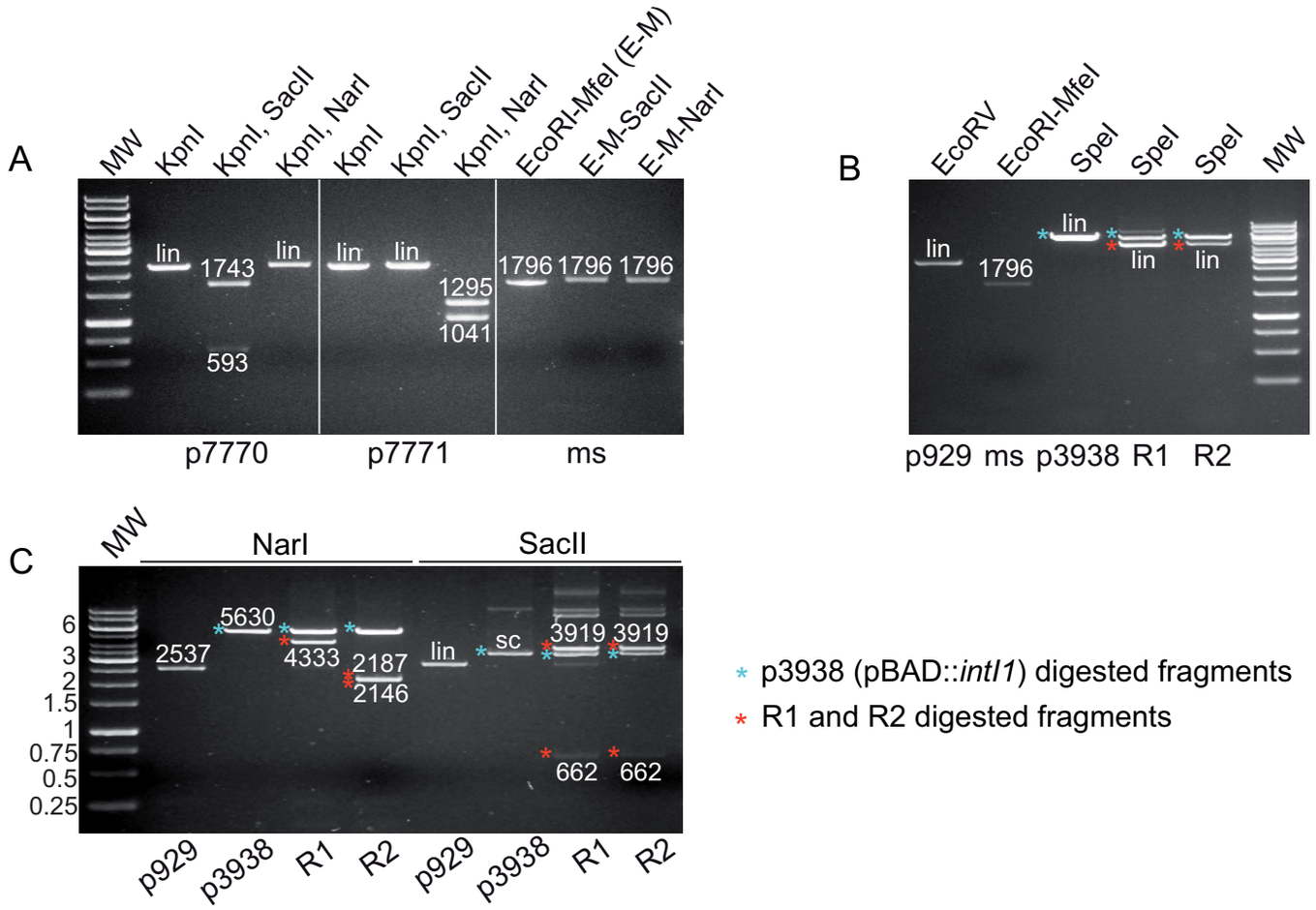

Figure S4: Electrophoretic analysis of reactional substrates and products in the non-replicative recombination assay.
MW, Molecular Weight Marker DNA; lin, linear; sc, supercoiled; ms, mismatched substrates. « R1 » corresponds to a recombinant plasmid which contains mutated (resistant) NarI and wild type (sensitive) SacII sites. « R2 » corresponds to a recombinant plasmid which contains both wild type (cleavable) NarI and SacII sites. Fragment size is indicated.

A) NarI and SacII sensitivity analysis of the phagemid vectors and the mismatched resulting substrate.
The p7770 and p7771 vectors are linearized (lin, 2336 bps) by KpnI restriction digest. We checked the SacII sensitivity of p7770 (1743 and 593 bps) and the NarI resistance (lin). Inversely, we checked the SacII resistance (lin) and NarI sensitivity (1295 and 1041 bps) of p7771. We also confirmed that the EcoRI-MfeI linearized mismatched substrates (ms) are SacII and NarI resistant (lin, 1796 bps) due to the presence of a GT mismatched dinucleotide in each restriction site.

B) Size analysis of both R1 and R2 recombinant products.
p929, ms, p3938 and both R1 and R2 recombinant products were respectively digested by EcoRV (lin, 2785 bps), EcoRI-MfeI (1796 bps), SpeI (lin, 5765 bps) and SpeI (lin, 4581 bps). We confirmed the presence of both R1 and R2 linearized bands (4581 bps), products of recombination between p929 (2785 bps) and ms (1796 bps) substrates. We can also visualize, in both R1 and R2 lanes, the SpeI linearized p3938 plasmid.

C) NarI and SacII sensitivity analysis of both R1 and R2 recombinants products.
p929, p3938, R1 and R2 were digested by both NarI and SacII restriction enzymes.
When cut with NarI, expected band sizes for p929 are 2537 and 248 bps, for p3938, 5630, 114 and 21 bps, for R1, 4333 and 248 bps and for R2, 2146, 2187 and 248 bps. In our running conditions, the 248, 114 and 21 bps bands are too small to be visualized on the gel and both 2146 and 2187 bands are below resolution capacity of agarose electrophoresis, hence they are merged.
When cut by SacII, for p929, expected band size is 2785 bps (lin), for p3938 plasmid, we expected a non-digested supercoiled (sc) plasmid and for R1 and R2, we expected two bands (3919 and 662 bps). In both R1 and R2 lanes, we can also visualize the presence of the sc p3938 plasmid. These digestions permit to confirm the presence of a mutated NarI and wild type SacII site in the R1 recombinant and both wild type NarI and SacII sites in the R2 recombinant.

Supplementary Table 1 : Primers

Name	*Sequence*
o221 (Seq*NarI*)	GCTTAATGAATTACAACAGTACTGC
o1073 (Seq*SacII*)	AGCAACTTAAATAGCCTCTAA
o1074 (dfrBcat1)	GGGGCGTAATTAAGGAACGTAATATGGAACGAAGTAGCAATGAAGTCAG
o1075 (dfrBBclI)	AGATTGATCAAAAAAATTTCTAGGCACCAATAACTGCCTAGAATTTAGTTGATGCGTTCAAGCGCCG
o2157 (dfrBcat2)	TCGTTCCATATTACGTTCCTTAATTACGCCCCGCCCTGCCACTCATCGCAG
o1761 (3634*Bcl*I)	TTTTTTGATCAATCTGATTAATAAGATGATCTTCTTG
o1076 (Fd2-1)	AATTGGGTACCCTCGAGACGCGCCCTGTAGCGGCGCATTAAGC
o1077 (Fd2-2)	TACCGTCGACAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCG
o607 (5'-7560-1)	AATTCCAGATCTAACAATTCATTCAAGCCGACGCCGCTTCGCGGCG
o608 (3'-7560-1)	GAATTAAGCCGCGCCGCGAAGCGGCGTCGGCTTGAATGAATTGTTAGATCTGG
o610 (5'-6978-2)	CGGCTTAATTCAAGCGTTAGATCTGGATCCACTAGTTCTAGAGCGGCCGCCACCGTGGTGGAGCT
o611 (3'-6978-2)	CCACCACGGTGGCGGCCGCTCTAGAACTAGTGGATCCAGATCTAACGCTT
o1762 (5'*NarI*m)	GCGTGCGCAAGAAATCCGGTGCCG
o1763 (3'*NarI*m)	CGGTGGAGCTCCAGCTTTTG
o1895 (MFD)	CGCCAGGGTTTTCCCAGTCAC
o1897 (Swbeg)	CCGTCACAGGTATTTATTCGGCG

Primers were obtained from Sigma-Aldrich and are given in 5' → 3' direction.