Unmasking the ancestral activity of integron integrases reveals a smooth evolutionary transition during functional innovation

Jose Antonio Escudero\textsuperscript{1,2}, Celine Loot\textsuperscript{1,2}, Vincent Parissi\textsuperscript{3}, Aleksandra Nivina\textsuperscript{1,2,4}, Christiane Bouchier\textsuperscript{5} & Didier Mazel\textsuperscript{1,2}

Tyrosine (Y)-recombinases have evolved to deliver mechanistically different reactions on a variety of substrates, but these evolutionary transitions are poorly understood. Among them, integron integrases are hybrid systems recombining single- and double-stranded DNA partners. These reactions are asymmetric and need a replicative resolution pathway, an exception to the canonical second strand exchange model of Y-recombinases. Integron integrases possess a specific domain for this specialized pathway. Here we show that despite this, integrases are still capable of efficiently operating the ancestral second strand exchange in symmetrical reactions between double-stranded substrates. During these reactions, both strands are reactive and Holliday junction resolution can follow either pathway. A novel deep-sequencing approach allows mapping of the crossover point for the second strand exchange. The persistence of the ancestral activity in integrases illustrates their robustness and shows that innovation towards new recombination substrates and resolution pathways was a smooth evolutionary process.
Tyrosine (Y)-recombinases mediate recombination between two specific DNA sites to deliver a variety of biological functions (for a review see (ref. 1)). They are present in the three domains of life but are most widespread among prokaryotes, where they play fundamental roles such as the integration/excision of viral genomes, the alternation of gene expression and the resolution of deleterious chromosome dimers arising during replication. Members of this family show a conserved fold of the catalytic domain including a three-dimensional clustering of highly conserved RKRRH residues. Y-recombinases typically recognize specific DNA sites comprising two inverted binding domains separated by a 6–8 bp spacer. The recombination reaction is a stepwise process that starts with the cleavage and transfer of one strand, leading to the transient formation of a Holliday junction (HJ) that is further resolved through a second exchange of strands. This sequence of steps, which we will refer to as the classical pathway (Fig. 1a), is highly conserved among Y-recombinases. Type IB topoisomerases and telomere resolvases share the fold and catalytic chemistry, and can therefore be considered as distant relatives within the same superfamily, despite sequence divergence.15–7. This implies that recombinases have evolved to deliver mechanistically distinct reactions (for example, the cleavage of only one strand by topoisomerases), and to recognize structurally different substrates (for example, hairpins in telomeres). Such evolvability is likely facilitated by the robustness of the fold8 (its capacity to accommodate mutations while preserving its tertiary structure), as well as the modularity of protein function through the combination of different functional domains. Nevertheless, information on evolutionary pathways that enable functional innovation among these enzymes is scarce.

Integrons are recombination platforms that allow bacteria to evolve and adapt rapidly through the acquisition, stockpiling and reordering of genes embedded in mobile genetic elements named integron cassettes8–10 (Fig. 1b). They are ancient structures that have driven bacterial evolution for hundreds of millions of years11, and, through their mobilization onto conjugative plasmids, they have played a critical role in the rise of multidrug resistance among clinically relevant bacteria12. Recombination in integrons is governed by integron integrases13 (hereafter integrases), a distinct group of Y-recombinases closely related to XerCD recombinases8. Integras show the conserved fold of the family, as well as the RKRRH pentad of residues in the active site (Fig. 1c). Incorporation of cassettes into the integron platform occurs through site-specific recombination between the attC site in the cassette and the attI site in the integron14. Genes in cassettes are then expressed from the Pc promoter within the integron platform (Fig. 1b). Cassette recombination represents an exception to the archetypical reaction mediated by Y-recombinases, as a consequence of the structural peculiarities of attC sites. These sites show no sequence conservation, except for the crossover point on the bottom strand (bs) (5'−AAC−3'), that provides microhomology with attI sites, necessary for strand transfer15. Instead, attC sites share a conserved palindromic structure, and are only recognized as a folded single strand (ss)16 (Fig. 1d). Recombination between a ss-attC and the double strand (ds)-attI site forms an asymmetric HJ (aHJ) (Fig. 1a) that cannot be resolved through a second exchange of strands since it would linearize the replicon and be abortive17. During attC site folding, the imperfect pairing of both arms of the bs allows for the protrusion of a set of extrahalal bases (EHBs) (Fig. 1d) that are key in strand selectivity, favouring the recombination of the bs18,19. EHBs also induce conformational changes in one integrase monomer during the reaction, pulling apart the tyrosine residue and avoiding the nucleophilic attack on the L box of the attC site20 (Fig. 1d). This prevents the second strand exchange, and the aHJ is then resolved through replication21 (Fig. 1a). To accommodate EHBs, integrases possess a 20-residue-long domain dubbed I2 (ref. 22; Fig. 1c) that has not been identified elsewhere, suggesting that it could have emerged recently among Y-recombinases. Such domain insertions are well-characterized phenomena in protein evolution23, allowing for a modular organization of protein function.

ssDNA is central to the integron, also controlling integrase expression through the SOS response25. Integrons are, therefore, unique recombination platforms streamlined to recombine ssDNA through a distinct resolution pathway. They represent a molecular example of functional innovation26.

While attC sites have been extensively studied17,19,20,27,28, knowledge is lacking on the recognition and processing of attI sites. These are more canonical than attC sites, since they are recognized as dsDNA16. They have two binding domains, the L and R boxes, although one is degenerate (Fig. 1d), and are generally recognized specifically by their cognate integrases29. Some attI sites possess accessory binding domains30.

Integrase site recognition underscores a remarkable conformational dynamism31. Identical monomers recognize ds-attI sites by sequence and ss-attC sites through structural landmarks. From an evolutionary standpoint, this duality could reflect a single specialized activity (the recombination of ss/ds DNA partners) or two distinct ones: the canonical Y-recombinase activity of integrase ancestors on attI, and the functional innovation, that is, the recombination of ssDNA structures, on attC sites. For the latter to be true, both classical and innovative activities must be complete: integrases should be capable of resolving HJs through a second strand exchange.

To unmask the ancestral activity of integrases, we have studied the reaction between two attI sites. This is a symmetrical reaction involving exclusively ds partners for which resolution through a second strand exchange is theoretically possible, since it is neither abortive nor impeded by EHBs. We show that the top strand of attI1 is reactive and that HJ resolution can take place through a second strand exchange. We have determined the crossover point through a novel deep-sequencing approach that is applicable to other recombination systems, and validated this approach in vitro. Finally, we prove that attC sites govern the directionality of the reaction. To our knowledge, integrases are the first example among recombinases of a smooth evolutionary transition between substrates and processing pathways. We discuss the possible origins and the evolutionary implications of the dual activity of integrases, as well as the importance of attI1 × attI1 reactions in antibiotic resistance.

Results

For this work we chose to study attI1, the integration site of the most widespread and best-studied integron platform, the Class 1 integron.

Dispensability of known HJ-processing host factors. The attI1 site possesses two direct repeats upstream the L–R boxes that make the site ~65 bp long32. It is therefore possible that host factors specialized in homologous recombination, branch migration and resolution of HJ-like structures (such as stalled replication forks), could influence the resolution of the attI1 × attI1 HJ, biasing our results. To assess the impact of such host machinery, we have tested the influence of RecA and two well-characterized HJ-processing complexes, RecG33 and RuvABC33,34, using a classic suicide conjugation assay35. Recombination rates of mutants of these complexes were not statistically different from those of the wild-type (WT) strain, suggesting that known host HJ-processing systems do not

---

**ARTICLE**

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms10937

---

2 NATURE COMMUNICATIONS | 7:10937 | DOI: 10.1038/ncomms10937 | www.nature.com/naturecommunications
Figure 1 | Recombination in the integron. (a) Diagram showing the two recombination models among tyrosine recombinases. The left column shows the recombination pathway shared by a majority of Y-recombinases, except the integron integrase for which the pathway is represented in the right column. (b) Schematic representation of the integron architecture and recombination reactions. Open reading frames are represented as boxes, with arrowheads showing the direction of transcription. Arrows marked as Pc and Pint represent the promoters of the cassette array and the integrase, respectively. Acquisition of new cassettes involves the recombination of the integron attI site and the cassette attC site, while cassette excision involves two attC sites. Sequential coupling of excision and integration reactions shuffles the order of cassettes within the array. (c) Crystal structure of V. cholerae integrase (blue) and Cre (green). RKHRH residues are depicted as purple sticks. The catalytic tyrosine is shown in orange. Red residues represent the I2 domain in VchIntIA and the residues in Cre that align with those delimiting I2 in VchIntIA. Residues in VchIntIA that interact with EHBs are represented in green. (d) Sequence of the Class 1 integron attI site (attI1), and the attCaadA7 site. aHJ, asymmetric HJ; bs, bottom strand; DR, direct repeats; ds, double stranded; HJ, Holliday junction; ss, single stranded; ts, top strand.
significantly influence the attI1 × attI1 reaction (Supplementary Fig. 1).

Both strands of attI1 are reactive. In cassette integration reactions, only the bottom strands of attC and attI sites are exchanged, with the crossover point located within the R boxes of both sites. To test the hypothesis that the attI1 × attI1 reaction can take place through both the bottom and the top strand, we used an attI1 carried on a mismatched covalent circle. These synthetic circles mimic non-replicative plasmids assembled using complementary strands produced separately and hybridized together in vitro. This allowed us to specifically tag each strand with a mutation in a restriction site (SacII or NarI) so that the restriction pattern of cointegrates reveals which strand is transferred during recombination (Fig. 2a). In our setting, the transfer and replication of the bottom and the top strands leads, respectively, to a SacII+ /NarI− and SacII− /NarI+ restriction pattern. The classical pathway of double-strand exchange would yield cointegrates containing both mismatches that, once replicated, would produce a mixed population of plasmids and hence partial restriction profiles with both enzymes (Fig. 2b). We verified the correct assembly of the DNA circles (as in ref. 21 (Supplementary Fig. 2)) and transformed them into a mutS− MG1655 Escherichia coli (E. coli) strain containing an attI-bearing pSU plasmid (p929) as well as a plasmid with the integrase under the control of a PBAD promoter (pBAD-intI1). The mutS− background ensures that the mismatches are not repaired. By selecting for the circle resistance marker, we recovered clones in which these molecules had recombined with the resident pSU plasmid through an integrase-mediated attI1 × attI1 reaction. The mismatch-containing region was PCR amplified from 100 recombinant clones and used for the restriction pattern analysis (Fig. 2c), revealing that both strands of

Figure 2 | The top strand of attI1 is reactive. (a) Experimental set-up used for mismatched-covalent-circles recombination. Mismatched circles bearing an attI1 site are transformed in an E. coli MG1655 mutS− strain bearing a pSU-attI1 plasmid and expressing the integrase in trans. (b) Representation of the recombination products resulting from the three possible HJ resolution pathways. The mismatch-bearing regions were PCR amplified and subjected to NarI and SacII restriction. (c) Restriction pattern analysis of three clones obtained, representing the three resolution pathways. (d) Sequencing of clone 36 PCR product revealed the expected mismatches.
the attI site are reactive and can be transferred independently, without a clear bias on strand specificity. This is contrary to what is observed for the attI1 × attC reaction where only the bs of attI1 is reactive15. Some clones showed partial restriction patterns and mixed plasmid populations (Fig. 2d), suggesting the possibility of HJ resolution through second strand exchange. Nevertheless, it could not be ruled out that this pattern was the consequence of two independent recombination reactions involving two different plasmids.

**The integrase resolves HJs through a second strand exchange.** To further assess if both strands of the attI1 site could be transferred during the same reaction, we used a modified version of the dapA recombination reporter36. In this system, recombination between attI1 sites leads to reconstitution of the essential gene dapA, allowing recombinants to grow on media lacking diaminopimelic acid (DAP, the reticulating agent of peptidoglycan in *E. coli*). Comparison of the number of clones growing with and without DAP yields a recombination frequency for a given reaction. We produced two versions of this reporter system, one with two attI1 sites in direct orientation, and the other with sites in inverted orientation. Both structures were inserted ectopically in single copy in the chromosome of an *E. coli* MG1655 ΔdapA strain. The rationale of this experiment is that, when both sites are in inverted orientation, viable recombinants can only be recovered if HJ resolution takes place through a second strand exchange that inverts the region between both sites. In this orientation the replicative resolution pathway linearizes the chromosome and is hence abortive (Fig. 3a; Supplementary Movies 1 and 2). On the other hand, when sites are in direct orientation both resolution pathways result in the excision of the region between the sites and are hence productive, serving as a control for the experiment. Our results show that integrase expression increased significantly the yield of recombinant clones for sites in direct orientation (44-fold) but also for sites in inverted orientation (72-fold) (Fig. 3b) (verified by PCR). This demonstrates that the attI1 × attI1 HJ can be resolved through a second strand exchange, that is, through the classical Y-recombinase pathway. Furthermore, comparison of the increases in recombination frequencies for both orientations proves that the classical resolution pathway is extremely efficient.

To avoid expression of dapA before recombination, one attI1 site was modified 1 bp upstream of the L box to include a stop codon (attI1STOP) (Fig. 3c). As a consequence, resolution through branch migration is not possible, suggesting again that host factors were not necessary in the reaction. It implies too that resolution of the HJ had taken place not further than 1 bp upstream of the L box, as confirmed by the change of location of the stop codon after recombination.

A similar experiment was performed using two attC sites, attCAdaA2 and attCereA2 (Supplementary Fig. 3). When sites were in direct orientation, we observed an increase in recombination frequencies of five orders of magnitude due to the presence of the integrase. Contrary to our observations for the attI1 × attI1 reaction, no clear increase in recombination frequency was observed for attC sites in inverted orientation. Sequencing of the few events (*n* = 8) in which the central structure was found inverted showed that the site had not been recombined in a double-stranded form. Indeed, the resulting sites were largely conserved, with the presence of moieties of the original sites only visible after the canonical 5′-AAC-3′ cleavage point. Although

---

**Figure 3 | The integrase can resolve HJs through second strand exchange.** (a) Graphic representation of the structural set-up of our experiment in which an attI1 × attI1 reaction reconstitutes the dapA gene, with attI1 sites in a direct orientation (left column) and in an inverted orientation (right column). Asterisks indicate attI sites modified to include a STOP codon (attI1STOP). The topology of HJs resulting from the first strand exchange is shown for both orientations. Recombination products following the two possible HJ resolution pathways are shown below. (b) Recombination frequency of all set-ups. Columns represent the mean and error bars the s.d. (*n* ≥ 3). Significance testing was performed with Kolmogorov-Smirnov test (*P value < 0.05; **P value < 0.01; NS, not significant). For the direct orientation set-up, strains used are B36, B82 and C162 (first, second and third column, respectively). For the inverted orientation set-up, strains used are B37, B83 and C163 (fourth, fifth and sixth column, respectively); their relevant genotypes are described in the table beneath the graph. (c) Sequence of the attI sites used in this experiment: attI1WT (the cleavage point on the bottom strand is marked with an arrow), attI1STOP, a variant not encoding an ORF, that allows to reduce background noise in the experiment while remaining recombionogenic; and attI1Ada, a variant in which the bottom strand is inactive.
these inversion events require a second cleavage that could not be elucidated, and therefore seem to be the consequence of illegitimate recombination reactions (possibly between sites on different replicating chromosomes), they are strong proof that attC sites are exclusively recombined in their structured form and confirms previous observations17.

To further prove that the second strand exchange occurs and is mediated by the integrase, we explored the possibility of impeding the bottom strand recombination. To do so, we mutated the cleavage point in the R box of one attI1 site in direct and inverted orientation from 5′-AAC-3′ to 5′-AAA-3′ (see attI1AAA in Fig. 3c). It has been observed that a C to A modification drops recombination rates for attI1 with both attC and attI1 partners15,37. Our results show a complete abolition of recombination when both sites are in the inverted orientation (that is, when the second strand exchange is needed to resolve the HJ), proving that the bs is no longer reactive. In contrast, recombination rates remain constant when the sites are in direct orientation, proving that integrase-mediated recombination through the ts is still possible for the attI1AAA site, and supporting the previously observed lack of clear bias in strand selection.

Mapping the ts crossover point using deep sequencing. Recombination between sites on different replicons results in co-integrate molecules bearing two sites. Each of them is composed of two moieties of the initial sites, but when the latter are identical, moieties are not easily identified and the crossover point of the reaction remains cryptic. Such is the case of the attI1 × attI1 reaction. Hence, to map the crossover point on the top strand of attI1, we built a library of attI1 sites in which we randomized all base pairs in the L box, the two adjacent base pairs within the spacer region and the C in the crossover point of the R box (Fig. 4a). Using N-containing primers, 10 attI1N site sets were produced, each bearing random bases in one of the ten positions. These sites were cloned in pir-dependent pSU plasmids and mixed to form a small library (n = 40) of pSU-attI1N plasmids to be used in our suicide conjugation assay. The rationale of this experiment is that the pattern of segregation of random bases between the two resulting sites in the co-integrate molecule will reveal the precise crossover point in the attI top strand (L box). Furthermore, since base composition of the L box can influence IntI1 binding, some attI1 mutants might be poorly recognized. A skew in the proportion of bases in a given position after recombination would underscore the importance of a specific base in the binding process (Fig. 4b). The shortcoming of this experimental set-up is that, as seen before (Fig. 2), the reaction can take place exclusively through the bottom strand (bs-exchange plus replication), without splitting the randomized L box. To overcome this, having confirmed the lack of bs exchange for attI1AAA, we used a pSU-attI1AAA plasmid as recombination partner for the library, forcing recombination to occur at the randomized L box. Deep sequencing of pSU-attI1N before recombination confirmed the expected distribution of 7.5% mutated bases in the randomized positions (explained in Methods) (Fig. 4c). Sequencing of the resulting left and right sites in the co-integrate revealed a complementary distribution of random bases among sites, allowing location of the crossover point in the L box between both adenines of the 5′-AAC-3′ triplet (Fig. 4c). It also allowed us to infer the importance of each base for protein–DNA interactions, revealing lower sequence stringency in the central 5′-CT-3′ base pairs of the L box (5′-CCCTAAA-3′), in accordance with previous interference data for the central T base65. To confirm the crossover position, we performed the attI1 × attI1 reaction in vitro using purified MBP-IntI1 protein, which is known to be fully active in vivo39, and attI1 fragments bearing a mismatch between the top and the bottom strand in the adenine that is transferred during the reaction (ts sequence 5′-ATC-3′, bs sequence 5′-GTT-3′) (Fig. 5a). Such a mismatch eliminates the microhomology necessary for docking of the transferred strand, and should abolish the transfer of the top strand. While the reaction between two WT attI1 sites is poorly detected in vitro40 (likely because HJ intermediates are quickly resolved), the reaction between L-box-mismatched substrates led to a significant accumulation of HJs (Fig. 5b), proving the role of the L-box cleavage in the resolution of these structures, and the dispensability of host factors in the reaction. The complementary effect (that is, the accumulation of HJs due to an R-box mismatch) was not observed, probably because mismatches in the R box led to a significant decrease in integrase binding (Supplementary Fig. 4a). The fact that binding is only affected by R-box mismatches is in accordance with the sequential and cooperative binding of monomers, starting on the R box38. As a control, all reactions were performed in parallel with the catalytically inactive version of the integrase, IntI1Y312F, for which no transfer activity was observed.

attC sites govern the directionality of the reaction. During attC site folding, the inverted repeats of the dsDNA molecule hybridize to form the R and L boxes in the hairpin (Fig. 1d). These boxes are present in hairpins arising from the folding of both the top and bottom strands. Despite a clear bias towards bs recombination, top strands are also recombined, although at much lower frequencies. In all cases, the crossover in the attI1 × attC reaction takes place in the R boxes of both sites. This means that top-strand recombination integrates cassettes opposite to the Pc promoter and are likely non-functional. Therefore, the similar frequency at which L and R boxes of attI1 recombine in the attI1 × attI1 reaction is in clear opposition to that found for the attI1 × attC reaction. An event that has never been reported is the recombination of attC sites through the L box. We sought to detect such recombination events with attI1, since they would presumably occur through the L box of attI1 ts. We reasoned that, since bottom strands of both sites recombine very efficiently through the R boxes, it would be easier to observe such events using the ts of attC. Using our conjugation assay, we delivered the ts of attCΔaA7 and checked a total of 607 recombinants. Among them, an estimate of 251 had recombined through the L (the rest having recombined through the resynthesized complementary bs) and only 1 (0.16%) had recombined through the L box with the top strand of attI1 (Supplementary Fig. 5). Since the EHB T23 on the bottom strand impedes the nucleophilic attack on the L box, we assessed if the EHB complementary to T23 on the top strand plays a role in this bias, using a ΔT23 mutant site. We found three L-box recombinants out of 642 colonies tested (0.47%) suggesting a role (although minor) for the complementary A base in avoiding L-box recombination of ts attC sites. The crossover point in these recombinants could not be determined precisely due to an adenine stretch, but it was not coherent with the cleavage site determined previously for the attI1 L box. This is probably related to conformational differences between attI1 and attC sites that force the crossover to shift. This need for a different crossover point could be itself at the basis of the low frequency of L-box events in the attI1 × attC reaction.

Discussion
In this work we show that integron integrases can perform the second strand exchange in reactions involving exclusively
We found that integrases recombine the top strand of attI sites at frequencies matching those of the bottom strand and that there is no order in the use of strands during the reaction. The crossover point in the top strand of the attI site has been elucidated using a novel technique that involves deep sequencing. This approach has allowed overcoming...
technical difficulties found with classical in vitro methods due to the low recombination activity observed. With deep sequencing becoming more accessible, this technique could be useful for others working with recombinases that show low activity or for which optimal conditions in vitro have not been elucidated.

Our work reveals surprising aspects of the dynamics of attI \(\times\) attI reactions, with broad evolutionary implications that are discussed below. Still, other aspects of the attI–IntI interaction remain unclear. attI sites are recognized for their sequence, as proven by cognate IntI–attI recognition\(^29\). We (and others previously\(^{28,29}\)) have observed in our in vitro experiments a cooperative binding starting at the R box. It is also known that secondary recombination sites contain the 5'-AAC-3' signature of the R box of all attI sites\(^{37}\). Given that the rest of the R box of the site changes its sequence at every cassette integration event, one could presume that this triplet is fundamental for IntI binding.

The almost complete loss of binding activity in our experiments when a mismatch is located in this triplet (Supplementary Fig. 4) supports this idea. Nevertheless, this triplet is universal to all attI sites and would not allow integrase discrimination of cognate sites. Also, the sequence of the attI1 bears a 5’-AAC-3’ triplet at the border of the L box, but the crossover point occurs between different bases (A/A) compared with that of the bottom strand (A/C), suggesting that topology is more important than sequence in the cleavage of the top strand. Therefore, the relative influence of the universal crossover triplet and of the rest of the R box in the binding remains a subject of interest that would need further studying.

This work shows that the avoidance of the second cleavage on attC sites is not due to the mere presence of the integrase-specific I2 z-helix, but rather to its interaction with attC sites. The extremely low frequencies at which attC\(_{\text{randA2}}\) recombines through the L box highlights the strong constraints that attC sites impose on the reaction. It is noteworthy that in natural conditions L-box reactions between attC and attI sites would place the attI site downstream of the cassette. Hence, such tight control on directionality likely serves to ensure the correct functioning of the integron by keeping the attI site next to the Pc promoter after consecutive cassette insertions.

From a functional perspective, the reaction between attI sites has an unclear adaptive meaning. It could be argued that in chromosomal integrons these reactions have deleterious effects on bacterial fitness, since they lead to the formation of chromosomal dimers that have to be further resolved before cell division. This is probably the reason for the 3–4 orders of magnitude lower recombination frequency for attI \(\times\) attI reactions compared with those involving attC sites. Nevertheless, in the multicyclic world of mobile integrons, one cell can contain more than one integron on the same or different replicons\(^{41–43}\).

Given the possibility of having Pc’s of different strengths, an attI \(\times\) attI reaction can lead to the en bloc transfer of cassette arrays between integrons. This would alter the expression levels of multiple antimicrobial resistance genes at once, and could have important effects on chemotherapy outcomes. It could be argued that the importance of such rearrangements is probably minor, due to the low recombination frequencies for attI \(\times\) attI in laboratory conditions. Nevertheless, one aspect of this reaction that has been constantly overlooked is that, while attC sites are normally in their non-recombinogenic form, attI sites are recombinogenic during almost the entire cell cycle, with the brief exception of the passage of the replication fork. Despite this low recombination frequency, it is plausible that attI \(\times\) attI reactions are more prevalent among mobile integrons than previously envisaged, especially under transient, low-level (or even stochastic) expression of the integrase. These expression pulses typically occur during horizontal gene transfer events\(^{36,44}\).

This work shows that despite millions of years of evolution towards a system specialized in structured ssDNA recombination, and the acquisition of a 20-residue-long domain within the catalytic core, integrin integrons have conserved the structure of their canonical Y-recombinase ancestors. Many examples of DNA breakage and rejoining enzymes acting on dsDNA (like Y-recombinases) or ssDNA (like HUH endonucleases\(^{45}\)) can be found, but this is the first report of a recombinase with full dual activity on both types of sites, as well as a substrate-dependent switch in recombination pathways. Integrases are capable of recognizing distinct substrates (ds versus ss hairpins) in different ways (sequence versus structure specific) and processing the reaction through different pathways (double strand exchange versus replicative resolution). This duality has deep biological implications, since one pathway is semiconservative and dependent on host machinery, while the other is not. Integrons could function, at least theoretically, using only the classical recombination pathway on dsDNA to deliver cassette integration and excision reactions. It is therefore tempting to speculate that the force driving the evolution of integrons towards a mixed ss/ds-DNA system derives from the benefits of semiconservative recombinase. By producing recombined and non-recombined offspring, it allows testing the adaptive value of incoming DNA, while minimizing the deleterious effect of capturing maladaptive genes. Furthermore, it also represents a mechanism for gene duplication, as observed in the Vibrio cholerae superintegron\(^10\) and possibly in some mobile integrons\(^{46}\).

From an evolutionary perspective, the functional innovation of integrases towards the recombination of folded ss-attC sites, posed important mechanistic constraints in the recognition of the site and the resolution of a HJ to deliver productive reactions. The solution in integrases comes in the form of the additional I2 domain that recognizes the EHBs in attC sites. This impedes the second cleavage, allowing for the replicative resolution of the HJ. Interestingly, EHBs and the integrase I2 domain are a good example of coevolution. These structures were probably the first step in the evolutionary transition from the ancestral recombination system to the ss/ds-DNA recombination platform integrons are today. Nevertheless, the acquisition of the I2 domain raises a chicken-egg paradox: how could the domain be selected for in the absence of attC sites to provide a gain of function? And how could attC sites be selected for before the recognition domain existed? Such a conundrum is found in many evolutionary processes and can easily be understood with the discovery of functional intermediary forms that allow to infer smooth evolutionary transitions between states. We believe integron integrases are representatives of such a bi-functional state. In the light of our results, it now seems possible that an I2-like domain was acquired at a given time by the ancestor of today’s integron integrases and, to some degree, permitted the recognition of new ss-substrates while retaining the original activity (or part of it) on the initial site. From this starting point, the streamlining through natural selection, to reach what integrons are today, seems a more straightforward process.

To date, the only recombination reaction that shares some similarities with cassette recombination is the integration of phage CTX\(_{\Phi}\) of Vibrio cholerae. Nevertheless, CTX\(_{\Phi}\) cannot be considered a complete system since it does not encode an integrase but instead hijacks the host XerCD recombinases\(^{47}\). As a consequence, no specialization of the system towards the recombination of ss-DNA is observed: XerCD lack structural adaptations for the recognition and processing of ss-substrates (I2-like domains), and the phage site stem does not have EHBs but rather mimics a ds-DNA molecule. Nevertheless, the resemblances between cassettes and CTX\(_{\Phi}\) integration, the close phylogenetic relation between integrases and XerCD
recombinases and the fact that both systems can coexist in the same bacterial species, makes it tempting to speculate about a viral origin for integrons. Interestingly, the link to the SOS response found in all integrons is also present in many phages, including CTXΦ (ref. 48). To date, no chromosome-dimer resolution recombinase of the XerCD family is known to possess a domain similar to I2 and directionality is ensured through their interaction with other host proteins such as PslK (J. Bouvier, personal communication).

Nevertheless, one could imagine that acquisition of an I2-like domain by a Xer-like ancestor could have provided the function of avoiding the second strand exchange during phage integration. This would protect the integrity of the chromosome while favouring correct integration events, and would hence be of adaptive value to the cell and the phage. We can now imagine that the chromosome dimer resolution activity of the hybrid recombinase would still be conserved. Phages, or any other organism, would be able to preserve this activity in order to ensure the integrity of the chromosome during the infection process.

In silico work. Protein modeling was performed using Phyre2 (sbg.bio.ic.ac.uk/phyre2) and MacPyMol on templates c2a23vA (VchIntIa) and c1ma7A (Crec). Sequence analysis was performed using Geneious 7.1.

Influence of host machinery in the attI × attI reaction. To test whether RecA, RecG and RuvABC impact the attI × attI reaction, we performed a suicide conjugation assay based on that of Bisker et al. (35) and previously implemented in Bouvier et al. (ref. 17). Briefly, the attI site provided by conjugation is carried on a suicide vector from the R6K-based pS9 family that is known to use the Pin protein to initiate its own replication. This plasmid also contains an RP4 origin of transfer (oriTR4). The donor strain B2163 carries the transfer functions in its chromosome, requires DAP to grow in rich medium and can sustain pS9 replication through the expression of a chromosomally integrated pin gene. The MG1655 recipient strains, which contain the pBAD::intI1 (Ap+) expression (the IntI1 integrase) and the pSU38::attI1 (Sp+) carrying the attI1 site, lacks the pin gene and therefore cannot sustain replication of the incoming attI-containing pS9 vector. The only way for this vector to be maintained in the recipient cell is to form a cointegrate by attI × attI recombination. We used the 010001 strain as the donor and as recipient strains the MG1655 E. coli oA197 (control), and its derivatives oA9898 (recA- : recG-) and oA9899 (recA- : recG- : ruvABC-). The recombination frequency is calculated as the ratio of transconjugants expressing the pS9 marker (Cm+) to the total number of recipient clones (Ap+, Sp+). Recombination frequencies correspond to the average of at least three independent trials. The attI × attI co-integrate formation was checked by PCR with the appropriate SWbeg and MFDR primers (on eight randomly chosen clones per experiment).

Mismatched covalent circles preparation. Construction of the p8669 and p8670 phagemid vectors to produce mismatched covalent circles was performed as in ref. 21 with some modifications. Briefly, we exchanged attC sites in plasmids p7770 and p7771 for attI sites, giving rise to plasmids p8669 and p8670 (phagemids, carrying the oriD in both orientations ensuring the production of respectively the bottom (oriD-) and top strand (oriD+)), single stranded DNA production was performed using M13K07 Helper Phage as in ref. 21. Both phagemid vectors (p8669 and p8670) are introduced by transformation into F' carrying pir strain cells (R2150, o4446) to respectively obtain the o8675 and o8676 strains. M13 infection and purification of said DNA were performed according to the manufacturer (M13K07 Helper Phage and QIAprep Spin M13 kit). The complementary single strand DNA molecules were annealed, digested by both EcoRI and MfeI restriction enzymes to eliminate oriFi (which does not anneal due to the inverted orientation in each molecule), and self-ligated.

Non-replicative recombination assay. This assay was developed previously in the laboratory (27). We transform with 1–2 µg mismatched covalent circles containing the attI site into MG1655mut1 (carrying the pSU38) (p3938) and the pSU38::attI1 (p929) plasmids (strain o7994). Circles cannot replicate in this genetic background (o7994). Note that this assay was first envisaged in a mutS-, recA-strain but transformation efficiency was too low due to a filamenting (sick) phenotype in the double mutant. We hence studied the influence of recA and found it to be responsible for <0.5% of all recombinants in our assay (see below for more information).

RecA oligonucleotide (1 µg) was added to transformation reactions performed by deleting the absence of integrase. RecA oligonucleotide (1 µg) was added to transformation reactions performed by deleting

Methods

Bacterial media and conditions. E. coli strains were grown in Luria Bertani broth (LB) at 37 °C. Growth at 42 °C was used to induce the expression and force the loss of the plasmid bearing lambda phage integrase. Antibiotics were used at the following concentrations: carbenicillin (Ap), 100 µg ml⁻¹, chloramphenicol (Cm), 25 µg ml⁻¹, kanamycin (Km), 25 µg ml⁻¹, spectinomycin (Sp), 40 µg ml⁻¹, Thymidine (Thy) and DAP were supplemented when necessary to a final concentration of 0.3 mM. Glucose and 1-arabinose were added at 10 and 2 mg ml⁻¹ final concentration, respectively. Chemicals were obtained from Sigma-Aldrich (France).

Bacterial strains, plasmids and primers. Bacterial strains in this study are E. coli DH5α, MG1655 (laboratory collection), 111, and B2163 (ref. 56). Details on strains, plasmids and primers are described, respectively, in Supplementary Tables 1–3.

DNA procedures. Standard techniques were used for DNA manipulation and cloning (27). Restriction and DNA-modifying enzymes were purchased from New England Biolabs and Fermentas (Thermo Scientific). PCRs were performed with DreamTaq DNA polymerase, and Phusion polymerase (Thermo Scientific) according to the manufacturer’s instructions. Agarose (1%) electrophoresis gels were used to visualize DNA. DNA purification from PCR products and gels, as well as plasmid extractions, were performed using Qiagen kits. When necessary, DNA sequencing was verified using an ABI BigDye Terminator v.3.1 sequencing kit and an ABI Prism 3100 Capillary Genetic Analyzer (Applied Biosystem). GATC and EUROFINS sequencing services were also used.

In silico work. Protein modeling was performed using Phyre2 (sbg.bio.ic.ac.uk/phyre2) and MacPyMol on templates c2a23vA (VchIntIa) and c1ma7A (Crec). Sequence analysis was performed using Geneious 7.1.

Influence of host machinery in the attI × attI reaction. To test whether RecA, RecG and RuvABC impact the attI × attI reaction, we performed a suicide conjugation assay based on that of Bisker et al. (35) and previously implemented in Bouvier et al. (ref. 17). Briefly, the attI site provided by conjugation is carried on a suicide vector from the R6K-based pS9 family that is known to use the Pin protein to initiate its own replication. This plasmid also contains an RP4 origin of transfer (oriTR4). The donor strain B2163 carries the transfer functions in its chromosome, requires DAP to grow in rich medium and can sustain pS9 replication through the expression of a chromosomally integrated pin gene. The MG1655 recipient strains, which contain the pBAD::intI1 (Ap+) expression (the IntI1 integrase) and the pSU38::attI1 (Sp+) carrying the attI1 site, lacks the pin gene and therefore cannot sustain replication of the incoming attI-containing pS9 vector. The only way for this vector to be maintained in the recipient cell is to form a cointegrate by attI × attI recombination. We used the 010001 strain as the donor and as recipient strains the MG1655 E. coli oA197 (control), and its derivatives oA9898 (recA- : recG-) and oA9899 (recA- : recG- : ruvABC-). The recombination frequency is calculated as the ratio of transconjugants expressing the pS9 marker (Cm+) to the total number of recipient clones (Ap+, Sp+). Recombination frequencies correspond to the average of at least three independent trials. The attI × attI co-integrate formation was checked by PCR with the appropriate SWbeg and MFDR primers (on eight randomly chosen clones per experiment).

Mismatched covalent circles preparation. Construction of the p8669 and p8670 phagemid vectors to produce mismatched covalent circles was performed as in ref. 21 with some modifications. Briefly, we exchanged attC sites in plasmids p7770 and p7771 for attI sites, giving rise to plasmids p8669 and p8670 (phagemids, carrying the oriD in both orientations ensuring the production of respectively the bottom (oriD-) and top strand (oriD+)), single stranded DNA production was performed using M13K07 Helper Phage as in ref. 21. Both phagemid vectors (p8669 and p8670) are introduced by transformation into F' carrying pir strain cells (R2150, o4446) to respectively obtain the o8675 and o8676 strains. M13 infection and purification of said DNA were performed according to the manufacturer (M13K07 Helper Phage and QIAprep Spin M13 kit). The complementary single strand DNA molecules were annealed, digested by both EcoRI and MfeI restriction enzymes to eliminate oriFi (which does not anneal due to the inverted orientation in each molecule), and self-ligated.

Non-replicative recombination assay. This assay was developed previously in the laboratory (27). We transform with 1–2 µg mismatched covalent circles containing the attI site into MG1655mut1 (carrying the pSU38) (p3938) and the pSU38::attI1 (p929) plasmids (strain o7994). Circles cannot replicate in this genetic background (o7994). Note that this assay was first envisaged in a mutS-, recA-strain but transformation efficiency was too low due to a filamenting (sick) phenotype in the double mutant. We hence studied the influence of recA and found it to be responsible for <0.5% of all recombinants in our assay (see below for more information).
chromosomal inversion test proves a second strand exchange. In this experiment we used an assay in which recombination between two sites leads to the reconstitution of the essential dapa gene allowing for growth in media not supplemented with DAP, DNA molecules containing the sites in direct orientation were designed in silico and synthesized by GenART (Life Technologies). To avoid background noise (expression and translation of dapa) before detection of the right att site (the initially unpaired site) was modified to include a STOP codon 1 bp upstream of the L box (Fig. 2). The design also included two XhoI and SmaI restriction sites surrounding the left att site and the Pmam promoter (registry part J23100) that drives transcription of dapa after recombination. Through XhoI/Smal double digestion we inverted the left site to obtain the inverted orientation. These constructions were subsequently cloned EcolR/NruI into the pir-dependent plasmid pA669 (Sp ρ) containing a lambda phage att site, and transformed into pir competent cells (o1628). A FRT recombination site was added in EcolR to allow for the resolution of dimers after integration in the chromosome. Sequence verification of these plasmids was performed before transforming the entire strain, an overnight incubation of dapa-. coli strain c5535 carrying the replication/expression-thermosensitive plasmid pC1533 bearing the lambda phage integrase (Sp ρ). In this strain, pir-dependent plasmids do not replicate and are only maintained through their integration in the lambda attB site of the chromosome. Integration events where selected for using the Sp resistance marker of pC1533. Loss of the pC1533 plasmid after overnight incubation at 42 °C was verified by replica plating on media containing carbenicillin (selecting for the Ap resistance marker of pC1533). Monomeric integration was verified by PCR with primers (o1329 and o1698). Sequence was verified in all strains. Strains oB36 and oB37 containing the constructions in direct and inverse orientation were selected for the study and transferred to the pBAD-attl plasmid pC252 and pA669 (pBAD::pir plasmid giving rise to strains oB82 and oB83. To test the recombination rate of the genetic structures, strains were grown overnight in LB supplemented with DAP, carbenicillin, spectinomycin and glucose, then diluted 1:50 into fresh media containing DAP, carbenicillin (to avoid the loss of p9398) and glucose or arabinose to repress or induce the expression of the integrase and the integrase, respectively. Plasmids were purified from the resulting colonies by phenol/chloroform/isoamyl-alcohol (24/25/1 v/v/v). Aqueous phase was run on an Agilent 2100 Bioanalyzer to determine the size of the resulting circles. Sequencing of random clones and IGV 2.3 (https://www.broadinstitute.org/igv/home). Base frequency was analyzed at each position in the library before recombination matched the expected distribution. Sequencing of the resulting sites after recombination showed a shift distribution of randomized bases in both sites, allowing to infer the crossover point.

In vitro experiments. Translational fusions with a maltose-binding protein were obtained of Infl1 and the catalytically inactive Infl1Δ1212 by cloning in the commercial vector pMAL-C5X and transforming E. coli Top10 strain. Plasmids were further purified and transformed in the BL21 strain for protein production. Colonies from this transformation were collected and grown overnight in 25 ml of LB containing carbenicillin and 0.2% glucose. After incubation, cells were collected by centrifugation at 4 °C, resuspended in 25 ml of fresh media, and inoculated in 31 of LB containing carbenicillin. When the culture reached an OD of 6.0, IPTG was added for induction of the fusion protein and the culture was incubated at 14 °C overnight. The cells were then collected by centrifugation and frozen. Cells were thawed to room temperature under cold tap water, and lysed by passage through an EmulsiFlex C-5 cell breaker (Avantium Europe GmBH) at 13,000 p.s.i. at 4 °C. The cell lysate was centrifuged at 23,500 rpm for 40 min in an RC-5C using an ss-34 rotor (Sorvall, Thermo Scientific). The Infl1-MBP fusion protein was purified with 5 ml Amsoyl column (GE Healthcare) on an AKTA-prime PLC machine (GE Healthcare).

The sequence of the oligonucleotides used for the strand transfer assay is reported in Supplementary Table 3. A total of 100 pmol of oligonucleotides coding for top and bottom strands were labeled using T4 DNA kinase (Promega) and 25 nM γ-[32P] ATP for 1 h at 37 °C. Once radiolabeled, top and bottom strands were then annealed (85 °C for 3 min and at room temperature for overnight). The strand transfer assay was performed using conditions derived from ref 40.59. Briefly, purified MBP-Infl1 was diluted in 1 M NaCl, 20 mM HEPES, 1 M, pH 7, 10 mM DTT for 30 min on ice. Then incubated with the different recombinase strains (25 nM final) for 30 min on ice in a final volume of 5 μl. The reaction was then started by adding 5 μl of the reaction buffer (final concentrations 100 mM NaCl, 15% DMSO, 8% PEG, 10 mM MgCl2, 100 mM NaCl, 10 mM DTT). Samples were incubated for 2 h at 37 °C. After reaction the products were treated h at 35 °C with protease K (Promega) 1 mg ml−1 for 60 min and deproteinized using phenol/chloroform/isooctyl alcohol (24:25:1 v/v/v). A gel phase was run on 12% polyacrylamide gel at 1,000 V for 5–6 h. Gels were then radiographed and quantified using Image J software (Supplementary Fig. 4b).

A total of 3 pmol of MBP-Infl1 were incubated for 20 min at room temperature with 1 pmol of radiolabeled double-stranded DNA containing the recombination sites attR, 10% DMSO, 50 μM MgCl2, 10% DTT. The cultures were then filtered on nitrocellulose filters. Filters were washed three times with PBS and counted using a Wallac 1409 liquid scintillation counter.

L-box recombination in the attCaddA-attl reaction. This experiment is based on the suicidal conjugation assay explained above. This time the plasmids delivered by conjugation carried the top or bottom strand of the wild type or the AT23 attCaddA sites. Maintenance of these plasmids in the recipient cell is only possible through an attl-attl recombination. We used the oD206, oD361, oD805 and oD806 Infl1 strains as donors and oD969 as the recipient strain. The recombination frequencies in R′ (bs) and (rt) and the corresponding mean deviations were deduced by determining the orientation of cassette integration for at least 20 independent clones for each bs transfer (and multiplying the total recombination frequency by the ratio of R′ and (rt) recombination) and for 100 clones in the ts transfer assays. The orientation of cassette integration was determined by performing two PCR reactions: the first reaction with SW23bgin/MFD primers produced a product only when the recombination took place in the delivered attC strand (in R′ when bs is delivered, 530 bp; in R′ when ts is delivered, 520 bp); the second reaction with SW23end/MFD primers produced a product only when the recombination occurred at the right site (in R′ when ts is delivered, 240 bp; in R′ when ts is delivered, 210 bp). In addition, the recombination in the L box was detected by performing two PCR reactions on all clones. The first reaction with RplP/S71 produced a product when the recombination took place in the L′ box (340 bp), the second reaction with RplP/S71 produced a product when the recombination took place in the L′ box (340 bp). Recombination through the L′ box of attl sites was only detected when the top strand was delivered (data for bs is hence not shown).
References

1. Grindley, N. D., Whiteson, K. L. & Rice, P. A. Mechanisms of site-specific recombination. Annu. Rev. Biochem. 75, 567–605 (2006).

2. Landy, A. The lambda integrase site-specific recombination pathway. Microbiol. Spectr. 3, MDNA3-0051-2014 (2015).

3. Golden, W. J., Robinson, S. J. & Haselkorn, R. Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium Anabaena. Nature 314, 419–423 (1985).

4. Val, M. E., Skovgaard, O., Ducos-Galand, M., Bland, M. J. & Mazel, D. Genome engineering in Vibrio cholerae: a feasible approach to address biological issues. PLoS Genet. 8, e1002472 (2012).

5. Sherratt, D. J. & Wigley, D. B. Conserved themes but novel activities in recombinases and topoisomerases. Cell 95, 149–152 (1998).

6. Toth-Petrocz, A. & Tawfik, D. S. The robustness and inovvability of protein binding specificity and target cooperativity. Curr. Opin. Struct. Biol. 26, 131–138 (2014).

7. Kobryn, K. & Chaconas, G. Hairpin Telomere Resolvases. Microbiol. Spectr. 2, MDNA3-0023-2014 (2014).

8. Mazel, D. Integrons: agents of bacterial evolution. Nat. Rev. Microbiol. 4, 608–620 (2006).

9. Recuca, G. J., Gross, F. M. & Hall, R. M. Gene cassettes: a new class of mobile element. Microbiology 141, 3015–3027 (1995).

10. Escudero, J. A., Loot, C., Nivina, A. & Mazel, D. The Integron: Adaptation On Demand. Microbiol. Spectr. 3, MDNA3-0019-2014 (2015).

11. Rowe-Magnus, D. A. & Mazel, D. The role of integrons in antibiotic resistance demand. J. Bacteriol. 196, 762–771 (2014).

12. Hall, R. M., Brookes, D. E. & Stokes, H. W. Site-specific insertion of genes into single-stranded integron cassette sites and their role in strand selection. PLoS Genet. 5, e1000632 (2009).

13. Macdonald, D., Demarre, G., Bouvier, M., Mazel, D. & Gopaul, D. N. Structural features of single-stranded integron cassette attC sites and their role in strand selection. PLoS Genet. 26, 39–46 (1998).

14. Toll-Riera, M. & Alba, M. Emergence of novel domains in proteins. BMC Evol. Biol. 13, 47 (2013).

15. Aroul-Selvam, R., Hubbard, T. & Sadashiran, R. Domain insertions in protein structures. J. Mol. Biol. 338, 633–641 (2004).

16. Garfinkel, E. et al. The SOS response controls integron recombination. Science 324, 1034 (2009).

17. Wagner, A. The molecular origins of evolutionary innovations. Trends Genet. 27, 397–410 (2011).

18. Demarre, G., Frumerie, C., Gopaul, D. N. & Mazel, D. Identification of key structural determinants of the IntI1 integrase integrase that influence attI1 x attI1 recombination efficiency. Nucleic Acids Res. 35, 6475–6489 (2007).

19. Looit, C. et al. The integrase efficiently prevents the melting effect of Escherichia coli single-stranded DNA-binding protein on folded attC sites. J. Bacteriol. 196, 762–771 (2014).

20. Collins, C. M., Kim, M. J., Stokes, H. W. & Hall, R. M. Integron-encoded IntI1 integrases preferentially recognize the adjacent cognate attI site in recombinase with a 59-bp site. Mol. Microbiol. 46, 1415–1427 (2002).

21. Pirttikka, L., Bouvier, M., Guarout, A. M., Boisard, S. & Mazel, D. Comparative study of class 1 Integron and Vibrio cholerae superintegron integrase activities. J. Bacteriol. 187, 1740–1750 (2005).

22. Saharoglu, Z., Bikard, D. & Mazel, D. Conjugative DNA transfer induces the bacterial SOS response and promotes antibiotic resistance development fok the whole genome. EMBO J. 6, e1001165 (2010).

23. Hanson, K., Skold, O. & Persson, L. Non-pathogenic attI sites of integrons are capable of site-specific recombination with one another and with secondary targets. Mol. Microbiol. 26, 441–453 (1997).

24. Fagre, A., Fournier, B. & Roy, P. H. DNA complexes obtained with the integron integrase IntI1 at the attI1 site. Nucleic Acids Res. 26, 6347–6355 (1998).

25. Mesier, N. & Roy, P. H. Integron integrases possess a unique additional domain necessary for activity. J. Bacteriol. 183, 6699–6706 (2001).

26. Dubois, V., Debreyer, C., Litvak, S., Quentin, C. & Parissi, V. A new in vitro strand transfer assay for monitoring bacterial class 1 integron recombine IntI1 activity. PLoS ONE 2, e1315 (2007).

27. González-Zorr, B. et al. Genetic basis for dissemination of armA. J. Antimicrob. Chemother. 56, 583–585 (2005).

28. Boyd, D., Cloeckaert, A., Chaslus-Dancla, E. & Mulve, M. R. Characterization of variant Salmonella genomic island 1 multidrug resistance regions from serovars Typhimurium DT104 and Agona. Antimicrob. Agents Chemother. 46, 1714–1722 (2002).

29. Hamidian, M., Holt, K. E. & Hall, R. M. Genomic resistance island AG1 carrying a complex class 1 integron in a multiply antibiotic-resistant ST25 Acinetobacter baumannii isolate. J. Antimicrob. Chemother. 70, 2519–2523 (2015).

30. Baharoglu, Z., Krin, E. & Mazel, D. Connecting environment and genome plasticity in the characterization of transformation-induced SOS regulation and carbon catabolite control of the Vibrio cholerae integron integrase. J. Bacteriol. 194, 1659–1667 (2012).

31. Chandler, M. et al. Novel integrase and attI integrase: the HUH endonuclease superfamily. Nat. Rev. Microbiol. 11, 525–538 (2013).

32. San Millan, A. et al. Sequencing of plasmids pMAMBl and pMAMBl2 from Pseudomonas aeruginosa reveals a blaVIM-1 amplification causing high-level carbapenem resistance. J. Antimicrob. Chemother. 70, 3000–3003 (2015).

33. Val, M. E. et al. The single-stranded genome of phage CTX is the form used for integration into the genome of Vibrio cholerae. Mol. Cell 19, 559–566 (2015).

34. Quinones, M., Kinsey, H. H. & Walder, M. K. LexA cleavage is required for CTX prophage induction. Mol. Cell 17, 291–300 (2005).

35. Das, B., Martinez, E., Midonet, C. & Barre, F. X. Integrative mobile elements exploiting Xer recombination. Trends Microbiol. 21, 23–30 (2013).

36. Midonet, C. & Barre, F. X. Xer site-specific recombination: promoting vertical and horizontal transmission of genetic information. Microbiol. Spectr. 2, MDNA3-0056-2014 (2014).

37. Walder, M. K. & Mekalanos, J. J. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science 272, 1910–1914 (1996).

38. Campbell, A. del-Campillo-Campbell, A. & Ginsberg, M. L. Specificity in DNA detection by phage integrases. Gene 300, 13–18 (2002).

39. Gottfried, P., Yagi, E. & Kolet, M. Core-binding specificity of bacteriophage integrases. Mol. Gen. Genet. 263, 619–624 (2000).

40. Yagi, E., Dorgai, L. & Weisberg, R. A. Identifying determinants of recombination specificity: construction and characterization of chimeric bacteriophage integrases. J. Mol. Biol. 252, 163–177 (1995).

41. Tawk, D. S. Accuracy-rate tradeoff-how do enzymes meet demands of selectivity and catalytic efficiency? Curr. Opin. Chem. Biol. 21, 73–80 (2014).

42. Demarre, G. et al. A new family of mobilizable suicide plasmids based on the broad host range R388 plasmid (IncW) or RP4 plasmid (IncPa) conjugative machineries and their cognate E.coli host strains. Res. Microbiol. 156, 245–255 (2005).

43. Beekman, J., Fritsch, E. F. & Maniatis, T. Molecular Cloning: A Laboratory Manual (Cold Spring Harbour Laboratory Press, 1989).

44. Looit, C., Bikard, D., Ratchlin, A. & Mazel, D. Cellular pathways controlling integron cassette site folding. EMBO J. 29, 2632–2634 (2010).

45. Dubois, V., Debreyer, C., Quentin, C. & Parissi, V. In vitro recombination catalyzed by bacterial class 1 integron integrase IntI1 involves cooperative binding and specific oligomers. PLoS ONE 4, e3228 (2009).

46. Martin, S. S., Chu, V. C. & Baldwin, E. Modulation of the active complex through integron activation. PLoS Genet. 3, e1000044 (2007).
Acknowledgements
This work was supported by the Institut Pasteur, the Centre National de la Recherche Scientifique (CNRS-UMR3525), the European Union Seventh Framework Programme (FP7-HEALTH-2011-single-stage), the ‘Evolution and Transfer of Antibiotic Resistance’ (EvoTAR) and the French Government’s Investissement d’Avenir program Laboratoire d’Excellence ‘Integrative Biology of Emerging Infectious Diseases’ (ANR-10-LABX-62-IBEID) and the French National Research Agency (ANR-12-BLAN-DynamINT). J.A.E. is supported by the Marie Curie Intra-European Fellowship for Career Development (FP-7-PEOPLE-2011-IEF, ICADIGE). We thank Mick Chandler, Dan Tawfik and Sebastian Aguilar Pierlé for critical reading of the manuscript.

Author contributions
J.A.E., C.L. and D.M. designed the experiments. J.A.E., C.L., V.P. and A.N. performed the experiments. J.A.E. and C.B. produced and analysed the deep-sequencing data. J.A.E. and D.M. wrote the paper.

Additional information
Accession codes: The crossover point sequencing data for the library, left site and right site has been deposited at the Sequence Read Archive under the accession code SRP071193.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Escudero, J. A. et al. Unmasking the ancestral activity of integron integrases reveals a smooth evolutionary transition during functional innovation. Nat. Commun. 7:10937 doi: 10.1038/ncomms10937 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/