# The TLCΦ satellite phage harbors a Xer recombination activation factor

Caroline Midonet[a], Solange Miele[a], Evelyne Paly[a], Raphaël Guerois[a], and François-Xavier Barre[a,1]

[a]Institute for Integrative Biology of the Cell (I2BC), Université Paris-Saclay, Commissariat à l'Energie Atomique et aux Energies Alternatives, CNRS, Université Paris Sud, 91198 Gif sur Yvette, France

Edited by G. Balakrish Nair, Rajiv Gandhi Centre for Biotechnology, Kolkata, India, and approved July 10, 2019 (received for review February 19, 2019)

**The circular chromosomes of bacteria can be concatenated into dimers by homologous recombination. Dimers are solved by the addition of a cross-over at a specific chromosomal site, *dif*, by 2 related tyrosine recombinases, XerC and XerD. Each enzyme catalyzes the exchange of a specific pair of strands. Some plasmids exploit the Xer machinery for concatemer resolution. Other mobile elements exploit it to integrate into the genome of their host. Chromosome dimer resolution is initiated by XerD. The reaction is under the control of a cell-division protein, FtsK, which activates XerD by a direct contact. Most mobile elements exploit FtsK-independent Xer recombination reactions initiated by XerC. The only notable exception is the toxin-linked cryptic satellite phage of *Vibrio cholerae*, TLCΦ, which integrates into and excises from the *dif* site of the primary chromosome of its host by a reaction initiated by XerD. However, the reaction remains independent of FtsK. Here, we show that TLCΦ carries a Xer recombination activation factor, XafT. We demonstrate in vitro that XafT activates XerD catalysis. Correspondingly, we found that XafT specifically interacts with XerD. We further show that integrative mobile elements exploiting Xer (IMEXs) encoding a XafT-like protein are widespread in gamma- and beta-proteobacteria, including human, animal, and plant pathogens.**

cholera | site-specific recombination | lysogenic conversion | integrative mobile element | IMEX

Bacterial chromosomes are often circular. As a consequence, sister chromosomes can be concatenated into dimers by homologous recombination. Chromosome dimers physically impede the segregation of genetic information. They are separated at the time of cell division by a highly conserved chromosomally encoded site-specific recombination machinery, Xer (1). Xer is not essential but it allows maximal cell proliferation. In addition, Xer participates in the evolution of bacteria by horizontal gene transfer: Some plasmids rely on it for concatemer resolution; other mobile DNA elements exploit it to integrate into the genome of their host (1). Plasmids relying on Xer and integrative mobile elements exploiting Xer (IMEXs) participate in the acquisition of antibiotic resistance and pathogenicity genes (2–4). In particular, cholera toxin, the principal virulence factor of *Vibrio cholerae*, is encoded in the genome of an IMEX, the cholera toxin phage, CTXΦ (5–8). Nontoxigenic *V. cholerae* strains generally lack a suitable CTXΦ attachment site. In addition, CTXΦ integration is intrinsically irreversible. Nevertheless, epidemic clones carrying potentially more potent forms of the toxin are constantly created by the excision of CTXΦ prophages and the subsequent integration of newly evolved CTXΦ variants (9–13). Several IMEXs contribute to these excision and reintegration cycles (14–17). Foremost among those is a toxin-linked cryptic satellite phage, TLCΦ, whose integration corrects the unsuitable CTXΦ attachment site that most nontoxigenic strains carry and whose excision jointly leads to the elimination of CTXΦ prophages (14, 15).

Circular chromosomes harbor a unique 28-bp Xer recombination site, *dif* (1). Chromosome dimers are resolved by the addition of a cross-over between the 2 *dif* sites they carry. The reaction is performed in 2 sequential steps: The exchange of a first pair of strands creates a Holliday junction (HJ), which is converted into a cross-over by the exchange of the second pair of strands (1). By default, the Xer machinery is inactive. It can create low amounts of HJs, but they do not proceed to cross-overs (18). Chromosome dimer resolution requires the action of an essential cell-division protein, FtsK (19). FtsK plays 2 roles. First, it forms DNA pumps that are anchored in the division septum and bring together the 2 *dif* sites carried by chromosome dimers (20–25). Second, it activates the Xer machinery during septum constriction (25–32). Natural Xer-dependent plasmids and IMEXs rely on FtsK-independent Xer recombination reactions (33). Plasmid core sites are flanked by ∼200 bp of accessory DNA sequence for the binding of accessory proteins. The accessory proteins intertwine the accessory sequences harbored by a plasmid dimer into a highly organized nucleoprotein complex, which brings together the core recombination sites and allows Xer recombination to proceed (34, 35). TLCΦ and most CTXΦ variants integrate into the dimer resolution site of the primary chromosome of *V. cholerae*. CTXΦ relies on the FtsK-independent ability of Xer to create low levels of HJs between its attachment sites and *dif* (8, 16). The HJs are converted into product by replication. CTXΦ integration is facilitated by rolling-circle replication, which amplifies the number of copies of its genome, and by a ubiquitous base excision repair enzyme, Endo III, which stabilizes HJs (36–38). In contrast, TLCΦ integration depends on the addition of a full cross-over between its attachment site and *dif1* by a yet poorly understood FtsK-independent Xer recombination reaction (15).

## Significance

**Cholera toxin, the principal virulence factor of *Vibrio cholerae*, is encoded in the genome of an integrative mobile element exploiting Xer (IMEX), CTXΦ. Nontoxigenic strains generally lack a suitable CTXΦ attachment site. In addition, CTXΦ integration is intrinsically irreversible. Nevertheless, new epidemic clones carrying potentially more potent toxins are constantly created by CTXΦ excision and reintegration cycles. Previous studies suggested that these cycles depended on another IMEX, TLCΦ, whose integration corrected the attachment site of nontoxigenic strains and whose excision promoted the joint elimination of CTXΦ copies. Our work brings molecular understanding to the role played by TLCΦ and suggests how similar IMEXs might participate in the evolution of other pathogenic bacteria.**

Author contributions: C.M., S.M., E.P., R.G., and F.-X.B. designed research; C.M., S.M., E.P., and R.G. performed research; E.P. and F.-X.B. contributed new reagents/analytic tools; C.M., S.M., E.P., R.G., and F.-X.B. analyzed data; and C.M., S.M., R.G., and F.-X.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.



See Commentary on page 18159.

[1]To whom correspondence may be addressed. Email: francois-xavier.barre@i2bc.paris-saclay.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1902905116/-/DCSupplemental.

Published online August 16, 2019.

The Xer machinery generally consists of 2 related tyrosine recombinases, XerC and XerD (1). Each of them is in charge of the exchange of a specific pair of strands (1). A C and a D pathway of recombination can thus be defined depending on whether XerC or XerD catalyzes the formation of the HJ intermediate of the reaction. By default, XerD is inactive, whereas XerC can catalyze the formation and resolution of HJs (1). Chromosome dimer resolution follows the D pathway: A direct interaction with FtsK triggers the activity of XerD, leading to the formation of HJs that are resolved into product by XerC (26). On the contrary, Xer-dependent plasmids and CTXΦ exploit the default FtsK-independent C pathway (8, 16, 35). Surprisingly, however, TLCΦ integration and excision reactions follow the D pathway (15).

Here, we show that TLCΦ encodes for a Xer recombination activation factor, XafT, which specifically interacts with XerD and promotes complete Xer recombination reactions by the D pathway. We further show that XafT acts independent of the sequence context of the recombination sites. Thus, XafT can promote both the integration and the excision of TLCΦ. Our results explain how TLCΦ contributes to the integration and excision cycles of CTXΦ. The discovery of XafT further permitted us to search for TLCΦ-like IMEXs in the available databases, which revealed that they are widespread in gamma- and beta-proteobacteria, including human, animal, and plant pathogens.

## Results

**TLCΦ Integration Depends on a 1-kbp Region Flanking Its Attachment Site.** Plasmid core recombination sites, IMEX attachment sites, and *dif* sites are composed of an 11-bp XerC-binding arm and an 11-bp XerD-binding arm, which are separated by a 6- to 8-bp overlap region at the borders of which strand exchanges occur (1). The region immediately flanking the XerD side of the attachment site of TLCΦ, *attP*, is devoid of ORFs, suggesting that it might contain an accessory sequence (Fig. 1*A*). To test this possibility, we monitored the efficiency of integration of a series of truncations of a nonreplicative form of the phage genome, pTLC, in *V. cholerae*. In brief, rolling-circle replication of the phage genome was abolished by inactivating its HUH endonuclease, Cri. pTLC carries a conditional R6K origin of replication, RP4 origin of transfer, and chloramphenicol (Cm) resistance gene. Truncations were built and maintained in an *Escherichia coli* Π+ strain and delivered by conjugation to a Cm-sensitive N16961 *V. cholerae* reporter strain. Because the pTLC constructs are not replicative in *V. cholerae*, the apparition of Cm-resistant colonies of the reporter strain indicated integration. The reporter strain was equipped with a functional *lacZ-dif1* in place of *dif1* to verify the specificity of the integration events. Integration frequencies were normalized using the number of Cm-resistant colonies obtained with a replicative form of pTLC. Results indicated that the ∼1-kbp region flanking the XerD arm of *attP* was not necessary for integration (Fig. 1 *A*, pTLC7 and 8). However, we found that an ∼1-kbp region normally flanking the XerC arm of *attP* in the genome of TLCΦ was necessary and sufficient for integration (Fig. 1 *A*, pTLC4 and 5 and pTLC6 to 8, respectively).

**The Product of VC1465 Is Necessary and Sufficient for TLCΦ Integration.** The ∼1-kbp region flanking the XerC arm of *attP* was unlikely to be an accessory sequence, because ∼2 kbp of DNA separates it from *attP* in pTLC (Fig. 1*A*). However, it encompassed 2 complete ORFs, VC1465 and VC1466. Introduction of a stop codon by site-specific mutagenesis in VC1466 did not affect integration (Fig. 1 *A*, pTLC9), whereas introduction of a stop codon in VC1465 abolished it (Fig. 1 *A*, pTLC10). The importance of VC1465 for integration was confirmed in an N16961 derivative devoid of the colorimetric screen and of the IMEXs (*SI Appendix*, Fig. S1). Ectopic production of the protein encoded by VC1465 from an arabinose promoter restored the integration of a construct harboring a stop codon in VC1465 (Fig. 1 *B*, pTLC10) and a construct lacking all of the other ORFs of TLCΦ (Fig. 1*B*, pTLC1). In addition, it permitted recombination between *attP* and *dif1* sites harbored on a plasmid in an $FtsK_C^-$ *E. coli* strain

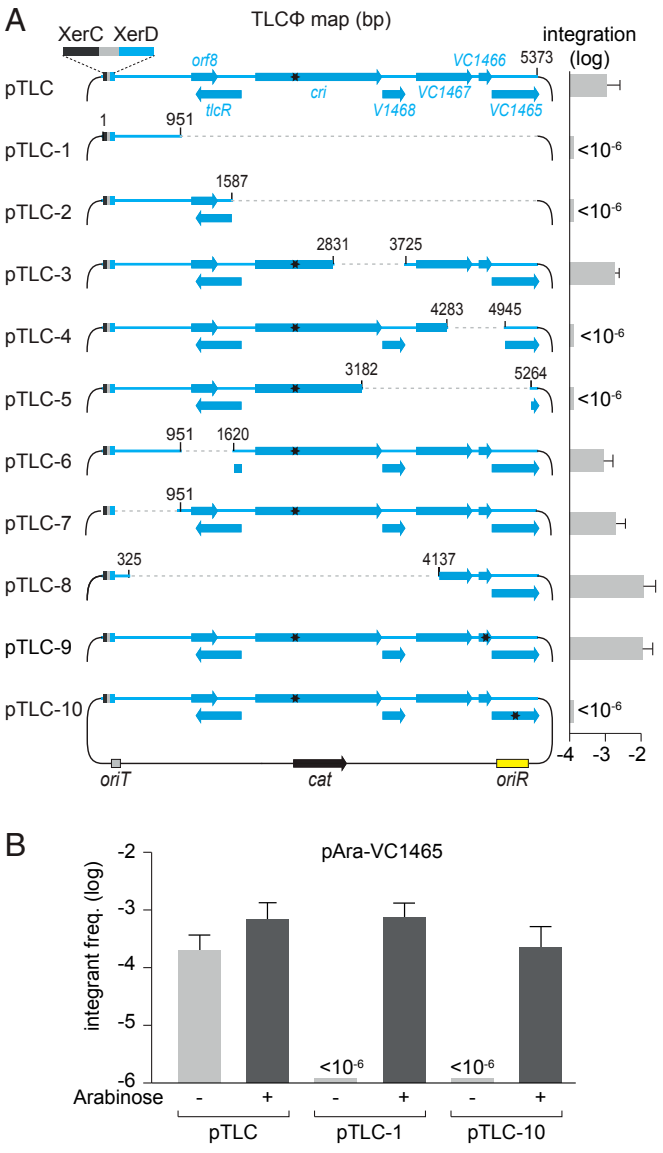


**Fig. 1.** TLCΦ integration depends on the peptide encoded by the VC1465 ORF. (*A*) In vivo integration of chloramphenicol-marked nonreplicative forms of TLCΦ. TLCΦ-specific and nonspecific features are shown in blue and black, respectively. Lines: dsDNA; arrows: ORFs; black stars: ORF-inactivating point mutations; *cat*, chloramphenicol resistance marker; *cri*, TLCΦ rolling-circle replication HUH endonuclease; *tlcR*, *cri* transcriptional repressor; black boxes: *attP* XerC arms; gray boxes: *attP* central regions; blue boxes: *attP* XerD arms. (*B*) Integration frequency of chloramphenicol-marked nonreplicative forms of TLCΦ under production of the peptide encoded by the VC1465 ORF in *trans*. + and − signs indicate whether the compound was included or not in the assay. Bar charts show the mean and SD of 3 independent experiments.

producing the *V. cholerae* Xer recombinases in place of the *E. coli* recombinases (*SI Appendix*, Fig. S2). Taken together, these results suggested that TLCΦ encoded for its own Xer activation factor, XafT.

**XafT Promotes the Recombination of *attP* and *dif1* In Vitro.** Most *Vibrio* species carry a *dif* site with the same sequence on their primary chromosome, *dif1* (Fig. 2*A*) (25). However, the ancestors of the strains at the origin of the previous (sixth) and present (seventh) cholera pandemics and many of their nontoxigenic descendants are equipped with a variant, $dif1_{AT}$, which precludes the integration of CTXΦ (Fig. 2*A*) (14). In addition, most

toxigenic descendants of the present pandemic are equipped with a third variant, $dif1_{GT}$ (Fig. 2*A*) (8). We reconstituted the Xer recombination reaction leading to the integration of TLCΦ using 34-bp synthetic double-stranded DNA (dsDNA) fragments containing any of these sites (Fig. 2*B*, S2), a 152-bp dsDNA fragment containing *attP* (Fig. 2*B*, S1), and purified XafT and *V. cholerae* XerC and XerD peptides. The *attP* substrate was created by PCR using pTLC as a template. The *dif1* substrates were assembled by annealing complementary oligonucleotides. The 5′ and the 3′ ends of the oligonucleotides corresponding to the strand normally processed by XerD were labeled with Cy3 and Cy5, respectively. Denaturing gel electrophoresis served to monitor the apparition of singly labeled DNA strands resulting from XerD catalysis (Fig. 2*B*, P1 and P2). Recombination was only observed when XerC, XerD, and XafT were added to the reactions (Fig. 2*B*). Taken together, these results suggested that no other protein factor than XafT was necessary to promote Xer recombination reactions between *attP* and the 3 most common *V. cholerae dif1* variants.

**TLCΦ Integration Follows the D Pathway.** Singly labeled dsDNA products with a size corresponding to the addition of a cross-over between *attP* and *dif1* (Fig. 3*A*, S1, S2, P1, and P2 schematic) were observed on nondenaturing gels (Fig. 3 *A*, *Top*). They were not detected when XerC or XerD was replaced by catalytically inactive mutants, suggesting that they resulted from the combined action of XerC and XerD (Fig. 3 *A*, *Top*, KQ lanes). We also observed the apparition of a highly retarded product carrying both the Cy3 and Cy5 labels (Fig. 3 *A*, *Top*). This product corresponds to an HJ between *attP* and *dif1* (Fig. 3*A*, HJ schematic). The amount of HJs was much greater than the amount of cross-over products (Fig. 3 *A*, *Top*). It remained very high when the catalytic activity of XerC was abolished, but no HJs could be detected when the catalytic activity of XerD was abolished (Fig. 3 *A*, *Top*, KQ lanes). The very faint amount of *attP*/*dif1* HJs created by XerC and XerD in the absence of XafT probably corresponds to the default activity of XerC (Fig. 3 *A*, *Top*). Taken together, these results suggested that XafT promoted the formation of *attP*/*dif1* HJs by XerD catalysis, which were resolved into product by XerC catalysis (Fig. 3*A*).

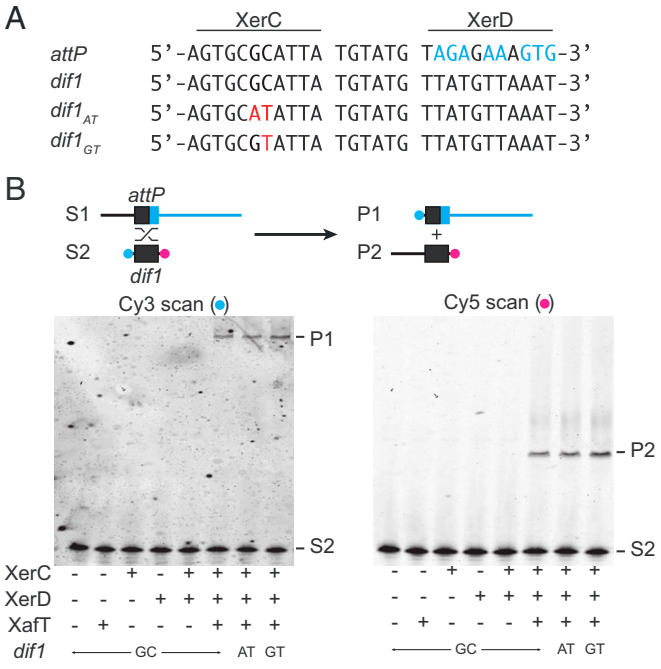


**Fig. 2.** TLCΦ encodes for a Xer recombination activation factor, XafT. (*A*) Sequence alignment of the attachment site of TLCΦ and of the 3 most common *V. cholerae dif1* variants. *attP*: TLCΦ attachment site; *dif1*: most common *dif1* in *Vibrios*; $dif1_{AT}$: *dif1* site of the ancestors of the strains at the origin of the sixth and seventh cholera pandemics; $dif1_{GT}$: hybrid *dif1* site resulting from the integration of CTXϕ and/or of the RS1 satellite phage. A single of the 2 DNA strands is represented in the 5′-to-3′ orientation from left to right. Bases of *attP* that differ from *dif1* are indicated in blue. Bases of $dif1_{AT}$ and $dif1_{GT}$ that differ from *dif1* are indicated in red. (*B*) Denaturing PAGE analysis of in vitro reconstituted Xer recombination reactions between a 152-bp *attP* substrate and the 3 most common *V. cholerae dif1* variants, as indicated. Schematics of the *attP* (S1) and *dif1* (S2) substrates and of the recombination products (P1 and P2) are drawn above the gel images. TLCΦ-specific and nonspecific features are shown in blue and black, respectively. Black boxes: *attP* XerC arms and central regions; blue boxes: *attP* XerD arms. Cyan and magenta dots indicate the positions of the Cy3 and Cy5 fluorescent labels, respectively. (*B*, *Left* and *Right*) Scans with the appropriate settings for the specific detection of the Cy3 and Cy5 labels, respectively. + and – signs indicate whether XerC, XerD, and XafT were added to the reactions or not.

**XafT Promotes Recombination between 2 *dif1* Sites.** The XerD-binding arm of *attP* significantly differs from the canonical XerD-binding arm of *dif* sites (Fig. 2*A*). To check the possibility that it might contribute to the activation of XerD by XafT, we analyzed the products of in vitro recombination reactions between a short, 34-bp Cy3- and Cy5-labeled *dif1* substrate and a long, 152-bp *dif1* substrate. The 152-bp substrate was produced by PCR using as a template pTLC*dif1*, a pTLC plasmid in which *attP* had been replaced by *dif1*. XerC and XerD promoted the formation of *dif1*/*dif1* HJs and cross-over products in the presence of XafT (Fig. 3*A*, *Bottom*). No *dif1*/*dif1* cross-over products were observed when the catalytic activity of XerC or XerD was suppressed by KQ mutations (Fig. 3 *A*, *Bottom*). Inactivation of XerC catalysis barely affected the amount of *dif1*/*dif1* HJs, but inactivation of XerD catalysis dramatically decreased it (Fig. 3 *A*, *Bottom*). The low amount of HJs observed with the XerD-catalytic mutant or in the absence of XafT probably results from the default basal XerC activity on *dif1* sites (18, 36). Thus, the XerD-binding arm of *attP* does not contribute to the integration mechanism of TLCΦ. On the contrary, the XerD-binding arm of *attP* seemed to be detrimental to recombination, since a higher amount of HJ intermediate and cross-over products was observed in the *dif1*/*dif1* reactions than in the *attP*/*dif1* reactions (Fig. 3*A*, *Top* and *Bottom*). This is probably explained by the poor binding of XerD to the XerD-binding arm of *attP* (15). Taken together, these results confirmed that XafT acts as a XerD activation factor.

**No Flanking DNA Is Required for XafT-Mediated Recombination.** Our in vivo integration assays suggested that no specific flanking DNA was required for XafT-mediated recombination (Fig. 1*A*, pTLC7). Correspondingly, the *dif1*, $dif1_{AT}$, or $dif1_{GT}$ sites of the short synthetic recombination substrates that we used for our in vitro recombination assays were only flanked by 3-bp GC clamps. However, the XerD and XerC sides of *attP* and *dif1* in the 152-bp PCR substrates were flanked by 89 bp of the phage genome and 35 bp of the pTLC vector, respectively. To determine the minimal DNA requirements for the action of XafT, we tested recombination between the short Cy3- and Cy5-labeled synthetic *dif1* substrate and synthetic *dif1* substrates with only 3 bp on the XerC side and 5 to 28 bp on the XerD side (Fig. 3*B*, S1 and S2 schematics). The expected Cy5-labeled recombination product has the same length as the Cy3- and Cy5-labeled substrate (Fig. 3*B*, S2 and P2 schematics). Nevertheless, we could detect its apparition using a sequencing gel because Cy3 and Cy5 labels retard the migration of DNA: The XerD cleavage strand of the Cy5-labeled 34-bp product migrated ∼5 nt below the doubly labeled XerD cleavage strand of the substrate (Fig. 3 *B*, *Right*). The lengths of the XerD cleavage strand of the Cy3-labeled recombination products ranged from 36 to 59 nt (Fig. 3*B*, P1 schematic). The 5 longer ones migrated above the doubly labeled XerD cleavage strand of the substrate (Fig. 3 *B*, *Left*). However, the shorter one migrated ∼1 nt below it (Fig. 3 *B*, *Left*).

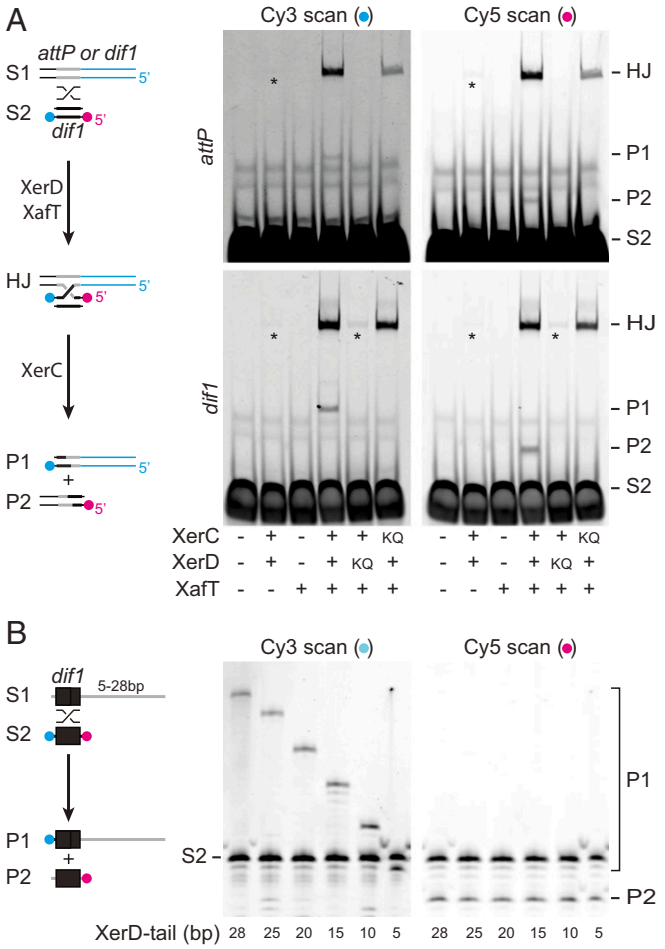


**Fig. 3.** XafT action is sequence-independent. (*A*) Native PAGE analysis of in vitro reconstituted Xer recombination reactions between a short synthetic labeled *dif1* substrate and a 152-bp *attP* (*Top*) or *dif1* (*Bottom*) substrate. A schematic of the 152-bp (S1) and *dif1* (S2) substrates, the Holliday junction (HJ) intermediate, and the recombination products (P1 and P2) is drawn (*Left*). DsDNA is represented by 2 straight lines to depict the HJ intermediate. Cyan and magenta dots indicate the positions of the Cy3 and Cy5 fluorescent labels, respectively. (*A*, *Left* and *Right*) Scans with the appropriate settings for the specific detection of the Cy3 and Cy5 labels, respectively. + and – signs indicate whether XerC, XerD, and XafT were added to the reactions or not; KQ, amino acid substitution inactivating the catalytic activity of Xer recombinases. Faint HJ products are indicated by asterisks. (*B*) Denaturing PAGE analysis of in vitro reconstituted Xer recombination reactions between short synthetic *dif1* substrates. The legend is as in Fig. 2*B*. Gray: nonspecific DNA.

Taken together, these results confirmed that no flanking DNA was required for the action of XafT.

**XafT Directly Interacts with XerD.** As no accessory DNA was required for XafT-mediated Xer recombination, we suspected that XafT might directly interact with the recombinases. We tested this possibility using the yeast 2-hybrid assay. The XafT gene was cloned in-frame with the GAL4 activation domain (AD) in a vector carrying the *LEU2* gene and introduced into a MATα *his3-200 ade2-101 trp1-901 leu2-3 Gal4Δ gal80Δ* yeast strain. The XafT, XerC, and XerD genes were cloned in-frame with the GAL4 DNA-binding domain (DBD) in a vector carrying the *TRP1* gene and introduced into a MATa *trp1-901 leu2-3 his3-200 Gal4Δ gal80ΔLYS2::GAL1*$_{UAS}$*-Gal1-TATA-His3 GAL2*$_{UAS}$*-Gal2-TATA-Ade2* yeast strain. The resulting strains were crossed and the diploids were selected on media lacking leucine and tryptophan. The diploids carrying the AD-XafT and DBD-XafT or DBD-XerD production vectors were white, indicating that the fusions permitted transcription of the *ADE2* gene from the *GAL2*$_{UAS}$ promoter (Fig. 4*A*, *Left*). In addition, they could grow on a medium lacking adenine and histidine, indicating that they also restored transcription from the *GAL1*$_{UAS}$ promoter (Fig. 4*A*, *Right*). Taken together, these results suggested that XafT formed multimers that interacted with XerD. We next tested whether XafT could directly interact with XerD using an in vitro pull-down assay. We used as baits MBP fusions to XerC, XerD, and a protein from VGJΦ unrelated to tyrosine recombinases (Cont). The fusions were bound to magnetic beads covalently coupled with amylose. Our purified XafT peptide slightly stuck to the amylose beads itself (Fig. 4*B*). We recovered 3 times more XafT with MBP-XerD–coated beads (Fig. 4*B*). No significant enrichment in XafT was observed in the MBP-XerC and MBP-Cont controls (Fig. 4*B*). Taken together, these results suggested that XafT directly and specifically interacts with XerD.

**XafT Homologs Are Found in the Genome of Many IMEXs.** There is no sequence homology between FtsK and XafT. XafT consists of an HTH domain from the XRE family of transcriptional regulators and a domain of unknown function, DUF3653. XerD activation was independent of DNA binding, suggesting that its action was due to DUF3653 (Fig. 3*B*). We recovered 179 different DUF3653-containing proteins in the NCBI, InterPro, and UniProt databases using a sensitive profile–profile search procedure (*SI Appendix*, Fig. S3). DUF3653-containing proteins are spread over the major orders of the γ- and β-proteobacteria, including many human, animal, and plant pathogens (*SI Appendix*, Fig. S4). Two hundred

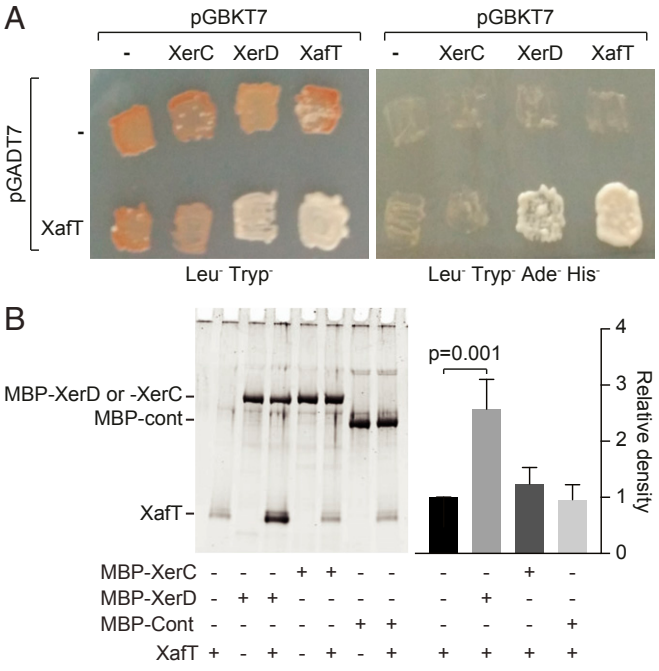


**Fig. 4.** XafT activates XerD catalysis by a direct interaction. (*A*) Yeast 2-hybrid assay. Pictures of patches of the Mat a/α yeast strains carrying the pGAD7 and pGBKT7 on media lacking leucine (Leu) and tryptophan (Tryp) (*Left*) and on media further lacking histidine (His) and adenine (Ade) (*Right*). XafT and - indicate whether the Gal4 activation domain of pGADT7 was in-fusion with XafT or not. XerC, XerD, XafT, and - indicate whether the Gal4 DNA-binding domain of pGBKT7 was in-fusion with XerC, XerD, and XafT or not. (*B*) In vitro pull-down assay. (*B*, *Left*) Gel scan showing the in vitro retention of XafT on magnetic beads coated with MBP-XerD, MBP-XerC, and a control protein (MBP-Cont). (*B*, *Right*) Relative density quantification of XafT retention. + and – signs indicate whether XafT, MBP-XerC, -XerD, and -Cont were added to the reactions. The bar chart shows the mean and SD of 3 independent experiments.

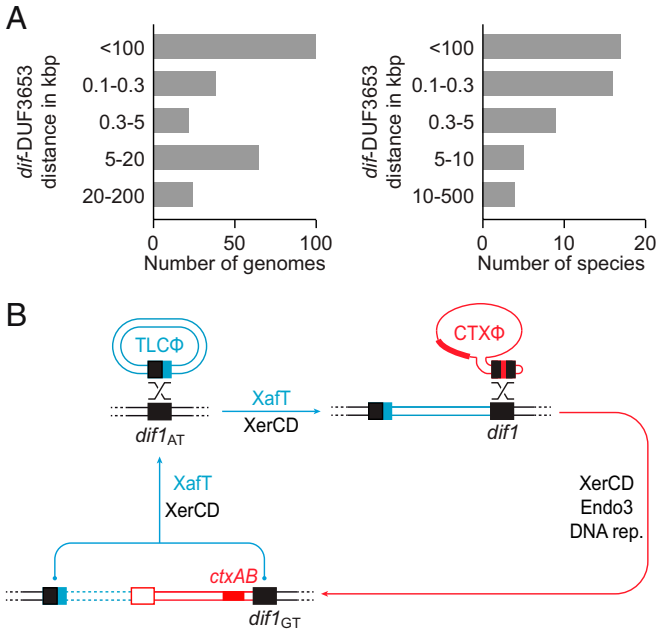


**Fig. 5.** XafT-like proteins are widespread and associated with IMEXs. (*A*, *Left*) Table showing the number of genomes harboring a DUF3653 at the specified distances from *dif1*. (*A*, *Right*) Table showing the number of species harboring a DUF3653 at the specified distances from *dif*. (*B*) Schematic of the integration reaction of TLCΦ into $dif1_{AT}$, the irreversible integration reaction of CTXϕ into *dif1*, and the *attP* x $dif1_{GT}$ recombination reaction leading to the joint excision of TLCΦ and CTXϕ. DNA strands are represented by thin lines. Xer recombination sites are represented by colored boxes. TLCΦ-specific and CTXϕ-specific features are shown in blue and red, respectively. Nonspecific or bacterial features are shown in black. The attachment site of CTXϕ is assembled from the phage ssDNA genomes. White boxes with a red contour indicate that no Xer recombination substrate can assemble from the integrated dsDNA prophage. (*C*) Sequence alignment of the *Neisseria gonorrhoeae dif* site and of the *dif*-like sites that flank the other end of the GGIs. One of the 2 DNA strands is represented in the 5′-to-3′ orientation from left to right. Bases of the GGI sites that are identical to TLCΦ *attP* are indicated in blue. Bases of the GGI sites that differ from $dif1_{Ng}$ are indicated in red.

and fifty genome hits containing at least one homolog of a DUF3653 were obtained from 51 species in which *dif* had been annotated (Fig. 5*A*). Most of the hits were located in the immediate vicinity of *dif*, suggesting that they belonged to IMEXs (Fig. 5*A*). In particular, we found all of the lysogenic phages integrated at *dif* in Xanthomonales plant pathogens (6). Most of the hits were not associated with HTH domains (as seen from the architectures of DUF3653-containing proteins in PFAM), strengthening the proposition that DUF3653 but not the HTH domain is responsible for the activation of XerD in XafT.

## Discussion

**TLCΦ Harbors Its Own XerD Activation Factor, XafT.** Some mobile elements codify their own integrases, but others rely on recombinases from the host for integration. The Xer machinery is highly conserved and extremely versatile, representing an advantage for the mobile elements that can make use of it. Most mobile elements exploit FtsK-independent Xer recombination reactions initiated by XerC. The discovery of XafT, which interacts with XerD and promotes recombination by the D pathway, unveils a new paradigm for the exploitation of Xer recombination (Figs. 1, 2, 3, and 4). Chromosome dimer resolution is promoted by a direct contact between XerD and the extreme C-terminal domain of FtsK, FtsKγ. However, FtsKγ can only efficiently activate XerD when it is directed toward the complex by FtsK translocation or when it is fused to the recombinases, which limited the analysis of the molecular switch it operates in the reciprocal control exerted by XerC and XerD on their catalytic activities (27, 31, 32). The discovery of XafT, which can efficiently promote the recombination between *dif* sites without being fused to the recombinases, opens new possibilities for dissecting the mechanisms of activation of XerD (Fig. 3).

**XafT Promotes Recombination between Apparently Defective Sites.** The XerD-binding arm of the attachment site of TLCΦ precludes the efficient binding of the recombinase (15). Nevertheless, XafT can promote its recombination with the 3 most common *dif* sites found on the primary chromosome of *V. cholerae*. It includes $dif1_{AT}$, the apparently defective site of the primary chromosome of the ancestors of the sixth and seventh pandemics and most of their nontoxigenic derivatives (Fig. 2) (14). XafT permits the integration of TLCΦ into $dif1_{AT}$, which corrects it into a site suitable for the integration of CTXΦ, *dif1* (Fig. 5*B*). The subsequent integration of CTXΦ then replaces *dif1* by $dif1_{GT}$ (Fig. 5*B*).

**XafT Is Both an Integration and an Excision Factor.** The attachment site of CTXΦ consists of the stem of a hairpin formed by its single-stranded DNA (ssDNA) genome. Conversion to dsDNA masks the site in the integrated prophage, which ensures the directionality of the reaction: Xer can integrate the ssDNA genome of CTXΦ but cannot excise the integrated prophage. No such process can impose directionality on the reaction promoted by XafT, since the attachment site of TLCΦ is composed of dsDNA (Fig. 2). Xer-mediated plasmid dimer resolution requires the assembly of a nucleoprotein complex with a specific topology. It ensures directionality of the recombination reaction because the complex can only be formed if the resolution sites belong to the same DNA molecule. No such process can impose directionality on the reaction promoted by XafT, which directly interacts with XerD (Fig. 4) and activates it independent of the sequence context of the sites (Fig. 3). Thus, the Xer reaction promoted by XafT is intrinsically reversible. Correspondingly, we found that TLCΦ is able to excise from its host genome (15). However, excision was less efficient than integration (15). Future work will need to investigate the amount of XafT produced by unintegrated and integrated copies of TLCΦ and the possible role of the sequences surrounding *attP* and *dif1* in the regulation of recombination.

**XafT Contributes to the Rapid Drift of the Cholera Toxin Region.** *V. cholerae* strains can be grouped into 12 distinct lineages, out of which only 1 gave rise to pandemic clones (9). The transition from the previous (sixth) to the current (seventh) pandemic was associated with a shift between 2 different phyletic subclades, the so-called classical and El Tor biotypes (9). The seventh pandemic isolates are rapidly drifting (9–13). In particular, cycles of excision and reintegration of CTXΦ promote the continuous apparition of clones producing new potentially more active forms of cholera toxin, which is a major clinical concern. The presence of an integrated copy of CTXΦ limits the possibilities for integration of new phage variants because it limits rolling-circle amplification of CTXΦ phages from the same incompatibility group (36–39). Thus, excision of previously integrated CTX copies is crucial for the spreading of new toxin variants. However, CTXΦ integration is intrinsically irreversible (7, 8). Infection by the RS1 satellite phage was found to favor CTXΦ excision (17). Homologous recombination between CTXΦ and a newly integrated copy of RS1, which contains ~3 kbp of homology with CTXΦ, could explain some of the excision events (17, 40). However, many others are accompanied by the joint elimination of TLCΦ and preexisting RS1 copies (14). Our discovery that

XafT can promote recombination between *attP* and $dif1_{GT}$ provides an explanation for these events (Fig. 5*B*).

**TLCΦ-Like IMEXs Are Widespread.** We tracked the presence of XafT-like homologs comprising the DUF3653 superfamily in available databases. The proportion of DUF3653 identified in close vicinity to a *dif* site is remarkable and strongly suggests that the function of DUF3653 in Xer recombination is common to most members of the superfamily (Fig. 5*A*). The details of the ORFs' ID number, position of the *dif* site, and genome ID number can be accessed from the information provided in an interactive tree of life (*SI Appendix*, Fig. S5; https://itol.embl.de/tree/13216652179197671505892051) (41). The position and orientation of the DUF3653 domains with respect to *dif* consistently vary with the different DUF3653 subclasses, which could correspond to subclass-specific mechanistic constraints. Our findings will help monitor IMEXs' dynamics. For instance, multiple occurrences of DUF3653 could be detected next to *dif* in *Xanthomonas oryzae*, suggesting multiple integration events (*SI Appendix*, Fig. S5; GenBank accession no. CP011955). Generalization of the role of XafT-like proteins may also help account for the integration of the Neisseriales gonococcal genetic islands (GGIs), which has long remained elusive (33). The GGIs are inserted between the *dif* site of their host chromosome and a *dif*-like site with a degenerate XerD-binding arm, which prevents FtsK-dependent Xer-mediated excision events (42, 43). Strikingly, 4 out of 8 bp of the XerD-binding arm of the *dif*-like site of one of the GGIs, $dif_{GGI3}$, are identical to the bases found in the *attP* XerD arm (Fig. 5*C*). The GGIs do not carry a XafT-like protein, but a XafT-like protein encoded by another IMEX could have promoted their integration in *trans* in the same manner as production of XafT from an ectopic vector can complement for the integration of pTLC1 (Fig. 1). Alternatively, the GGIs could harbor a Xer recombination activation factor from a superfamily different from DUF3653.

## Materials and Methods

Relevant strains, plasmids, and oligonucleotides are described in *SI Appendix*, Tables S1–S3, respectively. In vivo integration assay in *V. cholerae*, plasmid recombination assays in *E. coli*, protein purification, in vitro recombination, and pull-down assays, and the sensitive profile–profile search procedure for DUF3653 sequences in available databases are detailed in *SI Appendix*.

**ACKNOWLEDGMENTS.** We thank E. Espinosa, E. Galli, and C. Possoz for helpful discussions. This work had financial support from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement 281590) and the Agence Nationale pour la Recherche (PhenX/16-CE12-0030-01). C.M. was the recipient of a L'Oreal fellowship for Women in Science.

1. C. Midonet, F.-X. Barre, Xer site-specific recombination: Promoting vertical and horizontal transmission of genetic information. *Microbiol. Spectr.* **2**, MDNA3-0056-2014 (2014).
2. B. Das, E. Martínez, C. Midonet, F.-X. Barre, Integrative mobile elements exploiting Xer recombination. *Trends Microbiol.* **21**, 23–30 (2013).
3. M. S. Ramirez, G. M. Traglia, D. L. Lin, T. Tran, M. E. Tolmasky, Plasmid-mediated antibiotic resistance and virulence in gram-negatives: The *Klebsiella pneumoniae* paradigm. *Microbiol. Spectr.* **2**, PLAS-0016-2013 (2014).
4. T. H. Koh *et al.*, Putative integrative mobile elements that exploit the Xer recombination machinery carrying $bla_{IMI}$-type carbapenemase genes in *Enterobacter cloacae* complex isolates in Singapore. *Antimicrob. Agents Chemother.* **62**, e01542-17 (2017).
5. M. K. Waldor, J. J. Mekalanos, Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**, 1910–1914 (1996).
6. K. E. Huber, M. K. Waldor, Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* **417**, 656–659 (2002).
7. M.-E. Val *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 (2005).
8. B. Das, J. Bischerour, M.-E. Val, F.-X. Barre, Molecular keys of the tropism of integration of the cholera toxin phage. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4377–4382 (2010).
9. J. Chun *et al.*, Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15442–15447 (2009).
10. A. Mutreja *et al.*, Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* **477**, 462–465 (2011).
11. F.-X. Weill *et al.*, Genomic history of the seventh pandemic of cholera in Africa. *Science* **358**, 785–789 (2017).
12. D. Domman *et al.*, Integrated view of *Vibrio cholerae* in the Americas. *Science* **358**, 789–793 (2017).
13. E. J. Kim *et al.*, Molecular insights into the evolutionary pathway of *Vibrio cholerae* O1 atypical El Tor variants. *PLoS Pathog.* **10**, e1004384 (2014).
14. F. Hassan, M. Kamruzzaman, J. J. Mekalanos, S. M. Faruque, Satellite phage TLCφ enables toxigenic conversion by CTX phage through dif site alteration. *Nature* **467**, 982–985 (2010).
15. C. Midonet, B. Das, E. Paly, F.-X. Barre, XerD-mediated FtsK-independent integration of TLCϕ into the *Vibrio cholerae* genome. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 16848–16853 (2014).
16. B. Das, J. Bischerour, F.-X. Barre, VGJphi integration and excision mechanisms contribute to the genetic diversity of *Vibrio cholerae* epidemic strains. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2516–2521 (2011).
17. M. Kamruzzaman *et al.*, RS1 satellite phage promotes diversity of toxigenic *Vibrio cholerae* by driving CTX prophage loss and elimination of lysogenic immunity. *Infect. Immun.* **82**, 3636–3643 (2014).
18. F. X. Barre *et al.*, FtsK functions in the processing of a Holliday junction intermediate during bacterial chromosome segregation. *Genes Dev.* **14**, 2976–2988 (2000).
19. W. Steiner, G. Liu, W. D. Donachie, P. Kuempel, The cytoplasmic domain of FtsK protein is required for resolution of chromosome dimers. *Mol. Microbiol.* **31**, 579–583 (1999).
20. S. Bigot, O. A. Saleh, F. Cornet, J.-F. Allemand, F.-X. Barre, Oriented loading of FtsK on KOPS. *Nat. Struct. Mol. Biol.* **13**, 1026–1028 (2006).
21. S. Bigot *et al.*, KOPS: DNA motifs that control *E. coli* chromosome segregation by orienting the FtsK translocase. *EMBO J.* **24**, 3770–3780 (2005).
22. N. Dubarry, C. Possoz, F.-X. Barre, Multiple regions along the *Escherichia coli* FtsK protein are implicated in cell division. *Mol. Microbiol.* **78**, 1088–1100 (2010).
23. N. Dubarry, F.-X. Barre, Fully efficient chromosome dimer resolution in *Escherichia coli* cells lacking the integral membrane domain of FtsK. *EMBO J.* **29**, 597–605 (2010).
24. O. A. Saleh, C. Pérals, F.-X. Barre, J.-F. Allemand, Fast, DNA-sequence independent translocation by FtsK in a single-molecule experiment. *EMBO J.* **23**, 2430–2439 (2004).
25. M.-E. Val *et al.*, FtsK-dependent dimer resolution on multiple chromosomes in the pathogen *Vibrio cholerae*. *PLoS Genet.* **4**, e1000201 (2008).
26. L. Aussel *et al.*, FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* **108**, 195–205 (2002).
27. A. N. Keller *et al.*, Activation of Xer-recombination at dif: Structural basis of the FtsKγ-XerD interaction. *Sci. Rep.* **6**, 33357 (2016).
28. S. P. Kennedy, F. Chevalier, F.-X. Barre, Delayed activation of Xer recombination at dif by FtsK during septum assembly in *Escherichia coli*. *Mol. Microbiol.* **68**, 1018–1028 (2008).
29. G. Demarre *et al.*, Differential management of the replication terminus regions of the two *Vibrio cholerae* chromosomes during cell division. *PLoS Genet.* **10**, e1004557 (2014).
30. E. Galli, C. Midonet, E. Paly, F.-X. Barre, Fast growth conditions uncouple the final stages of chromosome segregation and cell division in *Escherichia coli*. *PLoS Genet.* **13**, e1006702 (2017).
31. L. Bonné, S. Bigot, F. Chevalier, J.-F. Allemand, F.-X. Barre, Asymmetric DNA requirements in Xer recombination activation by FtsK. *Nucleic Acids Res.* **37**, 2371–2380 (2009).
32. P. F. J. May, P. Zawadzki, D. J. Sherratt, A. N. Kapanidis, L. K. Arciszewska, Assembly, translocation, and activation of XerCD-dif recombination by FtsK translocase analyzed in real-time by FRET and two-color tethered fluorophore motion. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5133–E5141 (2015).
33. C. Midonet, F.-X. Barre, How Xer-exploiting mobile elements overcome cellular control. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8343–8345 (2016).
34. S. D. Colloms, J. Bath, D. J. Sherratt, Topological selectivity in Xer site-specific recombination. *Cell* **88**, 855–864 (1997).
35. M. Bregu, D. J. Sherratt, S. D. Colloms, Accessory factors determine the order of strand exchange in Xer recombination at psi. *EMBO J.* **21**, 3888–3897 (2002).
36. E. Martínez, E. Paly, F.-X. Barre, CTXφ replication depends on the histone-like HU protein and the UvrD helicase. *PLoS Genet.* **11**, e1005256 (2015).
37. E. Martínez, J. Campos-Gómez, F.-X. Barre, CTXϕ: Exploring new alternatives in host factor-mediated filamentous phage replications. *Bacteriophage* **6**, e1128512 (2016).
38. J. Bischerour, C. Spangenberg, F.-X. Barre, Holliday junction affinity of the base excision repair factor Endo III contributes to cholera toxin phage integration. *EMBO J.* **31**, 3757–3767 (2012).
39. H. H. Kimsey, M. K. Waldor, CTXphi immunity: Application in the development of cholera vaccines. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7035–7039 (1998).
40. E. J. Kim, C. H. Lee, G. B. Nair, D. W. Kim, Whole-genome sequence comparisons reveal the evolution of *Vibrio cholerae* O1. *Trends Microbiol.* **23**, 479–489 (2015).
41. I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
42. F. Fournes *et al.*, FtsK translocation permits discrimination between an endogenous and an imported Xer/dif recombination complex. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7882–7887(2016).
43. N. Kono, K. Arakawa, M. Tomita, Comprehensive prediction of chromosome dimer resolution sites in bacterial genomes. *BMC Genomics* **12**, 19 (2011).