



The Different Faces of Rolling-Circle Replication and Its Multifunctional Initiator Proteins

Paweł Wawrzyniak^{1,2*}, Grażyna Plucienniczak¹ and Dariusz Bartosik²

¹ Department of Bioengineering, Institute of Biotechnology and Antibiotics, Warsaw, Poland, ² Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, Warsaw, Poland

OPEN ACCESS

Edited by:

Zhe Zhao,
Hohai University, China

Reviewed by:

Gregory Marczyński,
McGill University, Canada
Hao Luo,
Tianjin University, China
Alan Leonard,
Florida Institute of Technology,
United States

*Correspondence:

Paweł Wawrzyniak
wawrzyniakp@iba.waw.pl

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 19 September 2017

Accepted: 15 November 2017

Published: 30 November 2017

Citation:

Wawrzyniak P, Plucienniczak G and
Bartosik D (2017) The Different Faces
of Rolling-Circle Replication and Its
Multifunctional Initiator Proteins.
Front. Microbiol. 8:2353.
doi: 10.3389/fmicb.2017.02353

Horizontal gene transfer (HGT) contributes greatly to the plasticity and evolution of prokaryotic and eukaryotic genomes. The main carriers of foreign DNA in HGT are mobile genetic elements (MGEs) that have extremely diverse genetic structures and properties. Various strategies are used for the maintenance and spread of MGEs, including (i) vegetative replication, (ii) transposition (and other types of recombination), and (iii) conjugal transfer. In many MGEs, all of these processes are dependent on rolling-circle replication (RCR). RCR is one of the most well characterized models of DNA replication. Although many studies have focused on describing its mechanism, the role of replication initiator proteins has only recently been subject to in-depth analysis, which indicates their involvement in multiple biological process associated with RCR. In this review, we present a general overview of RCR and its impact in HGT. We focus on the molecular characteristics of RCR initiator proteins belonging to the HUH and Rep_*trans* protein families. Despite analogous mechanisms of action these are distinct groups of proteins with different catalytic domain structures. This is the first review describing the multifunctional character of various types of RCR initiator proteins, including the latest discoveries in the field. Recent reports provide evidence that (i) proteins initiating vegetative replication (Rep) or mobilization for conjugal transfer (Mob) may also have integrase (Int) activity, (ii) some Mob proteins are capable of initiating vegetative replication (Rep activity), and (iii) some Rep proteins can act like Mob proteins to mobilize plasmid DNA for conjugal transfer. These findings have significant consequences for our understanding of the role of RCR, not only in DNA metabolism but also in the biology of many MGEs.

Keywords: rolling-circle replication, transposition, conjugal transfer, multifunctional protein, mobile genetic elements, horizontal gene transfer

INTRODUCTION

Rolling-circle replication (RCR) was first described nearly half a century ago following analysis of the replication of *Escherichia coli* bacteriophage Φ X174 (Gilbert and Dressler, 1968; Dressler, 1970). Since then, many other genetic elements replicating by the rolling-circle (RC) mechanism have been identified, including bacteriophages, plasmids of Gram-positive and Gram-negative bacteria, archaeal plasmids and eukaryotic viruses (Baas, 1985; Koonin and Ilyina, 1992; Ruiz-Masó et al., 2015). A specific feature of RCR is its asymmetric character manifested by the separation

of leading (+) and lagging (–) DNA strand synthesis. Replication of the (+) and (–) strands is initiated at different replication *origins*. The RCR initiation protein plays a key role in (+) strand replication. It initiates DNA replication by cleaving the leading strand within the corresponding *origin* (Figure 1A).

Interestingly, RCR was also found to play an important role in horizontal gene transfer (HGT). For example, it is the initial step in the conjugal transfer of plasmids and integrative and conjugative elements (ICEs) (Llosa et al., 2002), as well as the basis of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells (Zupan and Zambryski, 1995). It is also thought that some transposable elements, such as bacterial IS91-like insertion sequences and eukaryotic helitrons, use RCR to replicate in host genomes (Mendiola et al., 1994; Thomas and Pritham, 2015). According to their biological roles, RCR initiator proteins are divided into three groups: Rep proteins (vegetative replication), Mob proteins (mobilization for conjugal transfer) and Tnp proteins (transposition).

A number of reports concerning RCR replication have recently been published (e.g., Wang et al., 2013; Boer et al., 2016; Carr et al., 2016; Wright and Grossman, 2016). These have revealed new facts about RCR and especially the role of the replication initiator proteins. In this review, different RCR variants are compared, and features and functions of selected multifunctional RCR initiator proteins encoded by mobile genetic elements (MGEs) are described.

RCR INVOLVEMENT IN VARIOUS BIOLOGICAL PROCESSES

RC Mechanism in Vegetative DNA Replication

The best known RCR replicons are bacteriophages and bacterial plasmids, including the aforementioned prototype *E. coli* phage Φ X174 (Henry and Knippers, 1974; Ikeda et al., 1976; Eisenberg and Kornberg, 1979) and the plasmid pT181 isolated from *Staphylococcus aureus* (Koepsel et al., 1985). In both replicons the replication initiator protein interacts with two DNA regions, called the *bind* site and the *nick* site, located within the double-stranded *origin* (*dso*) (Figure 1B). The initiator protein interacts with the *bind* locus in a non-covalent manner. The exposed *nick* site of the *dso* is cleaved by the catalytic tyrosine of the initiator protein (Figure 1B). The released 3'-OH end then acts as the primer for synthesis of the new DNA strand. The initiator protein remains covalently attached (through a phosphotyrosine bond) to the 5' end of the nicked strand until leading strand replication is terminated, and the circular ssDNA intermediate is released from the newly replicated dsDNA molecule (Figure 1A).

Conversion of ssDNA into dsDNA is initiated at the lagging strand replication *origin*, referred to as the single-strand initiation (*ssi*) sequence in Φ X174 (Masai et al., 1990) or the single-strand *origin* (*sso*) in pT181 (Birch and Khan, 1992). The Rep protein is not involved in this stage of DNA replication; only the enzymatic machinery of the host is required. In Φ X174, primase complex is

recruited within the *ssi* and is responsible for the synthesis of the RNA primers. Starting from the *ssi*, primase then moves along the DNA strand, periodically synthesizing additional RNA primers. In pT181, only a single RNA primer is synthesized at the *sso* site by the RNA polymerase.

The RCR mechanisms of Φ X174 and pT181 differ significantly at the leading strand termination stage. This is a consequence of the different life strategies of bacteriophages and bacterial plasmids. In phage Φ X174, termination of leading (+) strand replication initiates another round of replication, enabling the phage to rapidly amplify its genome before lysis and escape from the bacterial cell (Novick, 1998). In contrast, the number of pT181 copies per bacterial cell is tightly regulated. Therefore, after a single replication cycle, RepC, the replication initiator protein of pT181, is inactivated and cannot initiate another round of replication (Figure 1A).

Although the crystallographic structures of Protein A of Φ X174 and RepC of pT181 have yet to be determined, detailed molecular analyses have helped to describe their role in the initiation of replication (Figure 1). The Rep proteins exhibit nucleolytic activity and they cleave the leading DNA strand to initiate the replication process. The active sites of both these enzymes contain conserved tyrosine residues that are directly involved in the catalytic functions. Protein A has two tyrosines, Y1 and Y2, which are responsible for the initiation and termination of replication, respectively (Mansfeld et al., 1986) (Figure 1A). In the active site of RepC of pT181 there is only one tyrosine, but as this Rep protein acts as a dimer, the tyrosine residue of one subunit (Y1) initiates replication, while that of the other (Y2) terminates it (Rasooly and Novick, 1993) (Figure 1A). These two structural arrangements seem to be preferred since prokaryotic RCR initiator proteins are mainly purified as a monomers or dimers (Ruiz-Masó et al., 2015).

Interestingly, the RepB protein of the extensively studied promiscuous plasmid pMV158 (Lorenzo-Díaz et al., 2014, 2017; López-Aguilar et al., 2015) can form a hexameric ring (RepB₆). RepB is the only plasmid-encoded RCR initiator with a determined atomic structure (Boer et al., 2009). This protein is composed of two domains: (i) an N-terminal *origin*-binding domain (OBD) (with catalytic tyrosine residue Y⁹⁹), and (ii) a C-terminal oligomerization domain (OD), responsible for formation of the hexameric structure. In this regard, RepB resembles Rep proteins of eukaryotic viruses in which hexameric ring formation determines helicase activity (Boer et al., 2016). However, RepB does not contain a helicase domain. It is believed that the RepB₆ ring fastens the DNA strand, enabling recruitment of helicase and DNA polymerase. The replisome so formed is highly efficient in duplex unwinding and DNA synthesis, and allows pMV158 to replicate in a broad spectrum of bacterial hosts (Boer et al., 2009). It is not clear how RepB₆ terminates replication. In one of the proposed models, the flexible structure of RepB₆ OBDs enables the use of the Y⁹⁹ from one monomer for initiation cleavage and the Y⁹⁹ from another monomer for termination cleavage (Boer et al., 2009). OBD flexibility also facilitates *origin* recognition by RepB₆ (Boer et al., 2016).

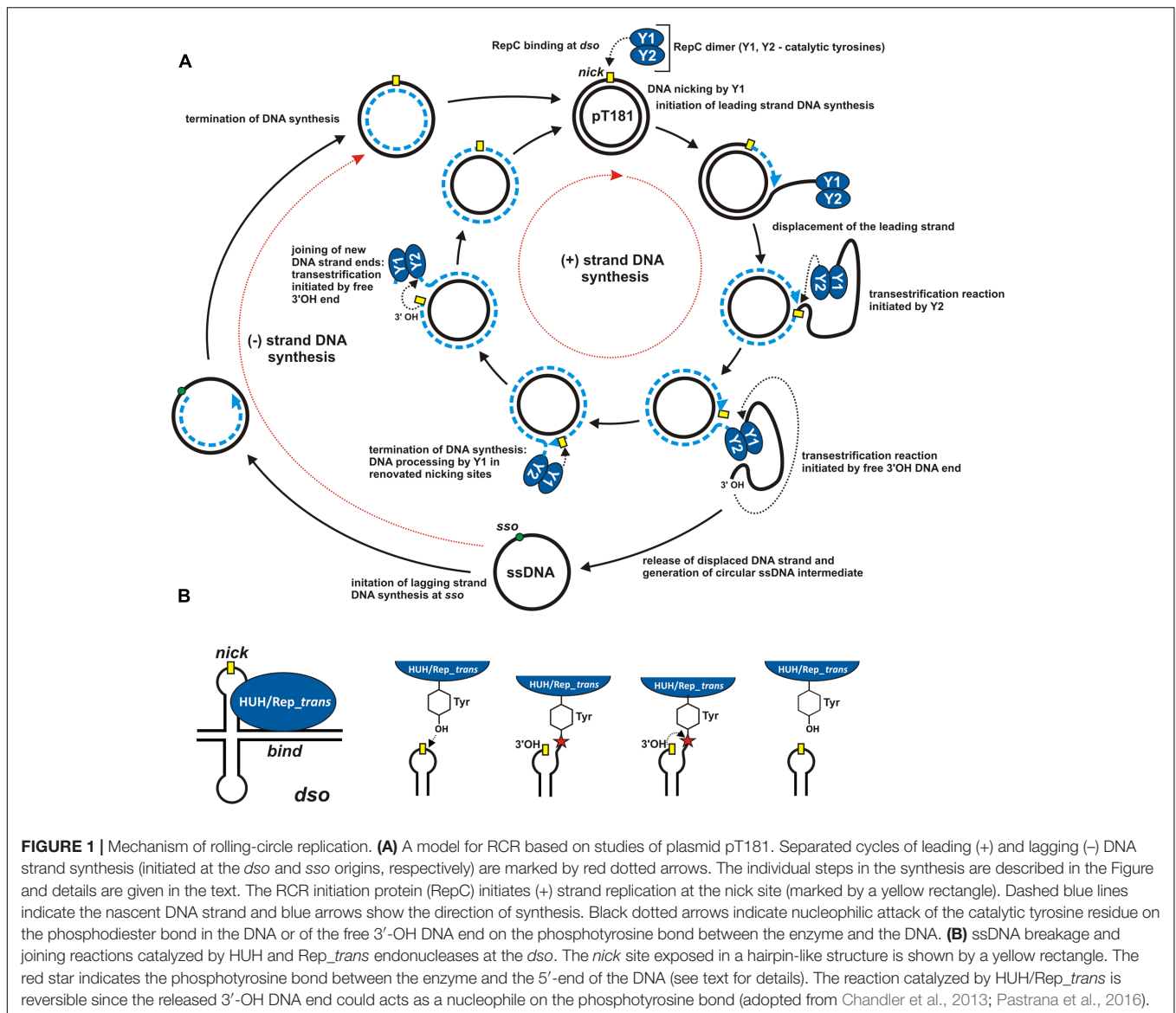


FIGURE 1 | Mechanism of rolling-circle replication. **(A)** A model for RCR based on studies of plasmid pT181. Separated cycles of leading (+) and lagging (-) DNA strand synthesis (initiated at the *dso* and *sso* origins, respectively) are marked by red dotted arrows. The individual steps in the synthesis are described in the Figure and details are given in the text. The RCR initiation protein (RepC) initiates (+) strand replication at the nick site (marked by a yellow rectangle). Dashed blue lines indicate the nascent DNA strand and blue arrows show the direction of synthesis. Black dotted arrows indicate nucleophilic attack of the catalytic tyrosine residue on the phosphodiester bond in the DNA or of the free 3'-OH DNA end on the phosphotyrosine bond between the enzyme and the DNA. **(B)** ssDNA breakage and joining reactions catalyzed by HUH and Rep_{trans} endonucleases at the *dso*. The *nick* site exposed in a hairpin-like structure is shown by a yellow rectangle. The red star indicates the phosphotyrosine bond between the enzyme and the 5'-end of the DNA (see text for details). The reaction catalyzed by HUH/Rep_{trans} is reversible since the released 3'-OH DNA end could act as a nucleophile on the phosphotyrosine bond (adopted from Chandler et al., 2013; Pastrana et al., 2016).

The replicons described above have circular genomes. However, non-circular genomes may also replicate using an RC-like mechanism. For example, a variation of RCR named rolling-hairpin replication (RHR) (Figure 3) was observed in the adenovirus-associated virus, AAV, a member of the family *Parvoviridae* (Tattersall and Ward, 1976). The genome of this virus is relatively small (~4700 bp) and comprises only two genes: *rep* – encoding the replication initiator, and *cap* – encoding the capsid formation protein. These genes are located between two inverted terminal repeats (ITRs), whose palindromic sequences promote the formation of characteristic T-shaped structures at both ends of the DNA molecule (Figure 3). Interestingly, the AAV *rep* gene may encode up to four peptides, two of which, Rep78 and Rep68, contain motifs that are characteristic of Rep proteins of RCR plasmids.

During infection, the single-stranded DNA of the virus is introduced into a eukaryotic cell and then converted to dsDNA.

Synthesis of the complementary strand does not require an RNA primer because DNA replication is initiated from the free 3'-OH end of the ITR acting as a primer (Figure 3). However, replication initiated in this way does not lead directly to duplication of the entire AAV genome, and the RHR mechanism must be used to complete the replication round (Figure 3). RHR is initiated at a *dso* located within the 3' ITR. Similarly to the double-stranded *origin* of RCR plasmids and bacteriophages, the *dso* of AAV also comprises two conserved regions: a binding site for the replication protein (RBS) and a DNA cleavage site, known as the terminal resolution site (TRS) (Balakrishnan and Jayandharan, 2014).

RCR in Conjugal Transfer

Many plasmids and ICEs may be transmitted to other bacterial cells by conjugal transfer (Smillie et al., 2010; Carraro and Burrus, 2014; Cabezón et al., 2015). This process requires the

involvement of two gene modules responsible for (i) mating pair formation (MPF) and (ii) DNA transfer and replication (DTR). In the case of elements containing only the DTR module (mobilizable plasmids and integrative and mobilizable elements, IMEs), conjugal transfer is possible only when a compatible MPF module is present *in trans*.

The MPF region usually comprises genes encoding the type 4 secretion system (T4SS) involved in DNA transport and biogenesis of the pilus connecting bacterial conjugation pairs (Garcillán-Barcia et al., 2009; Guglielmini et al., 2014). The DTR region encodes the Mob protein, a relaxase that cleaves one strand of DNA within a specific site known as the *origin of transfer* (*oriT*). The *oriT* region comprises functional *nick* and *bind* sites, which are analogous to those of the *dso* RCR plasmids. Moreover, it has been demonstrated that cleaving one DNA strand in *oriT* also initiates RCR.

The key role of RCR in conjugal transfer is to separate the replication of the two strands of DNA so that it occurs simultaneously in different bacterial cells. The leading strand (+) is replicated in the donor cell and the lagging strand (–) in the recipient cell. This process is more complex than the vegetative plasmid replication described above.

In conjugal transfer, the relaxase protein (Mob) acts like the Rep protein of plasmids and viruses during RC vegetative replication. The Mob protein also contains tyrosine residues in its active site, which determine the catalytic functions. For example, the TrwC relaxase of the plasmid R388 contains two catalytic tyrosines (Y¹⁸ and Y²⁶) that are essential for DNA transfer (marked Y1 and Y2 in Figure 2). According to the model proposed by Gonzalez-Perez et al. (2007), the first tyrosine (Y¹⁸) is responsible for the initiation of replication at *oriT* in the donor cell, while the second residue (Y²⁶) promotes termination of leading strand replication in the recipient cell. The relaxase TraI of the plasmid F also contains two conserved tyrosine residues within its catalytic domain, although only one (Y¹⁶) is active in DNA processing (marked Y1 in Figure 2). This suggests the involvement of two protein molecules acting independently at the initiation and termination sites of the lagging strand (Gonzalez-Perez et al., 2007).

After the initiation of replication, relaxase remains covalently bound to the 5' end of *oriT*, while the free 3'-OH end is used as a primer for the synthesis of the complementary strand in the donor cell. DNA relaxation induces the formation of a nucleoprotein complex called the relaxosome, which is then targeted via the T4 coupling protein (T4CP) to the components of the T4SS. As a result, the leading DNA strand, along with the relaxosome, is transferred to the recipient cell through the pores formed in the membrane by the T4SS. The single-stranded DNA transferred from the host is recognized by the replication engine of the recipient strain, and it serves as the template for the synthesis of the complementary strand. This completes the round of replication.

Interestingly, some relaxases have additional domains responsible for primase activity. Such multifunctional proteins are transferred to the recipient cell along with the relaxosome complex, where they are involved in the RNA primer synthesis. As a result, the transferred ssDNA is converted to dsDNA even

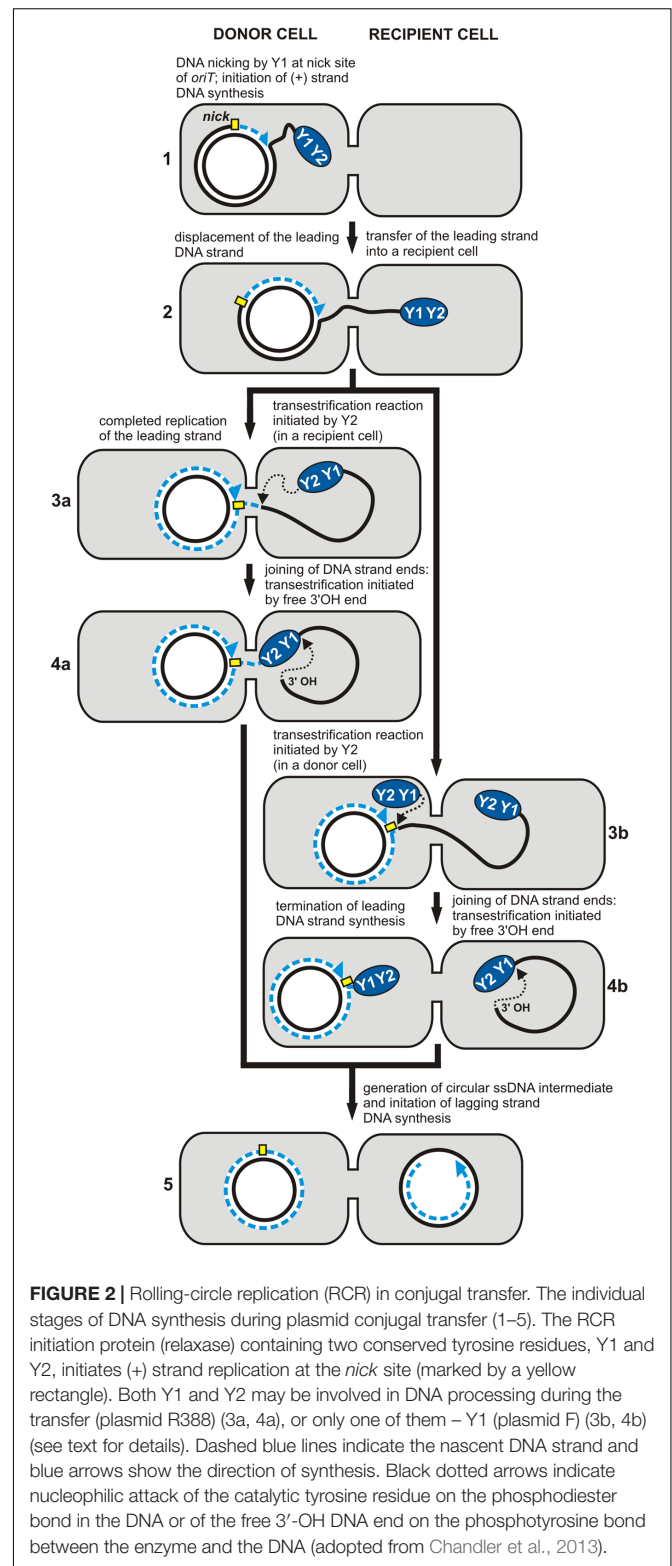
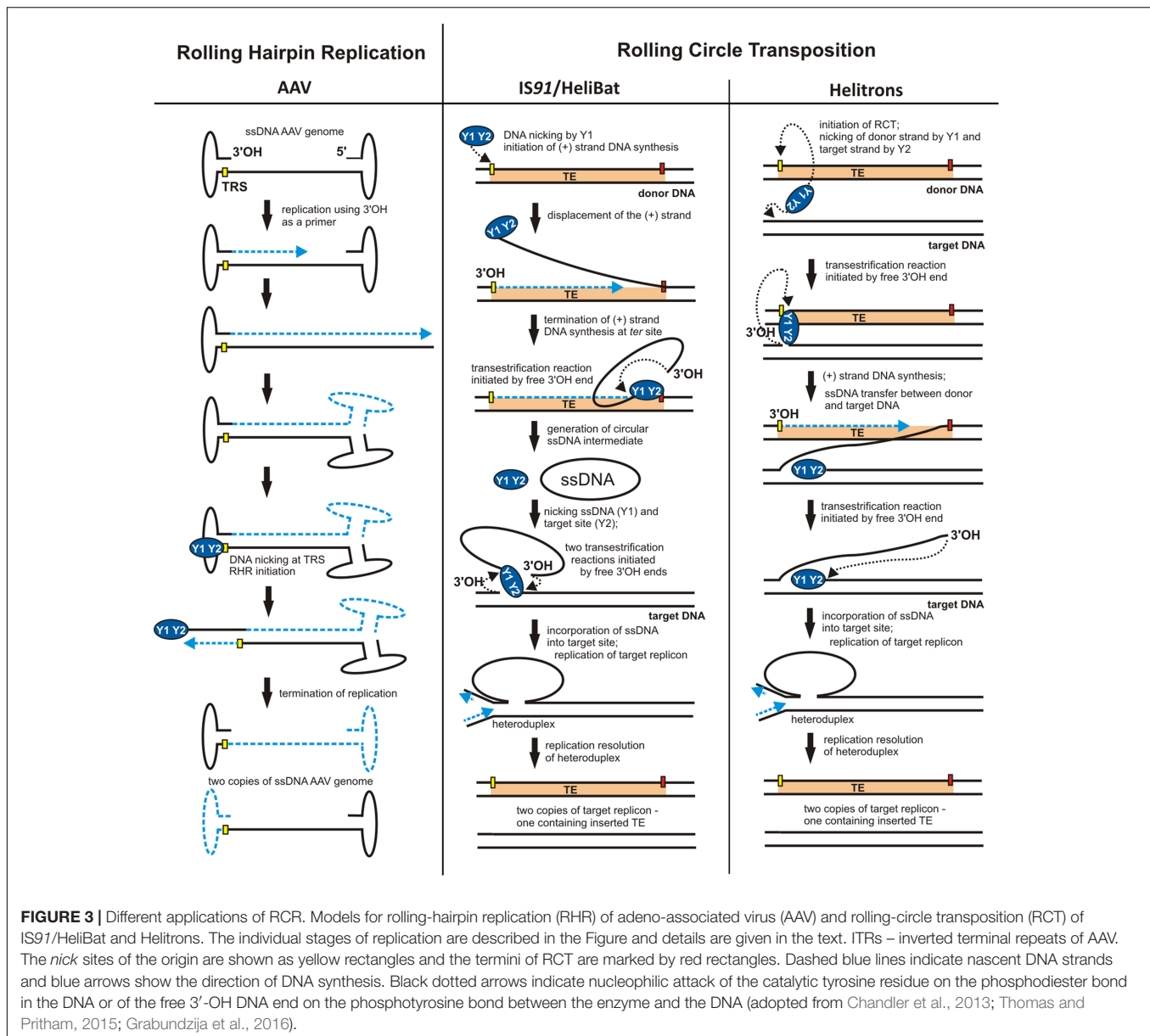


FIGURE 2 | Rolling-circle replication (RCR) in conjugal transfer. The individual stages of DNA synthesis during plasmid conjugal transfer (1–5). The RCR initiation protein (relaxase) containing two conserved tyrosine residues, Y1 and Y2, initiates (+) strand replication at the *nick* site (marked by a yellow rectangle). Both Y1 and Y2 may be involved in DNA processing during the transfer (plasmid R388) (3a, 4a), or only one of them – Y1 (plasmid F) (3b, 4b) (see text for details). Dashed blue lines indicate the nascent DNA strand and blue arrows show the direction of synthesis. Black dotted arrows indicate nucleophilic attack of the catalytic tyrosine residue on the phosphodiester bond in the DNA or of the free 3'-OH DNA end on the phosphotyrosine bond between the enzyme and the DNA (adopted from Chandler et al., 2013).

in strains where the lagging strand *origin* is not recognized by the replication engine of the recipient. It is thought that this primase activity of the Mob protein might extend the host range of the MOB module (Henderson and Meyer, 1999).



RCR in Transposition

Transposable elements (TEs) are the largest group of mobile genetic elements found in bacteria. Usually they are integrated within the genome of the host, but they may change their location by recombination, known as transposition. The transposition is catalyzed by a transposase enzyme encoded by TEs.

A characteristic feature of the majority of TEs is the presence of inverted repeats (IRs) at their termini. These sequences are responsible for interaction with the transposase. During transposition, most TEs also generate direct repeats (DRs) in the target DNA. Among the exceptions to this general model is the insertion sequence IS91 identified in a hemolytic *Escherichia coli* strain (Zabala et al., 1982). This TE does not contain IR sequences nor does it generate DRs during transposition. Moreover, the amino acid sequence of the IS91-encoded transposase (TnpA)

does not resemble those of other TE transposases (Mendiola et al., 1992), but it contains motifs typical for RCR initiation proteins, including two strongly conserved tyrosine residues (Y²⁴⁹ and Y²⁵³). Mutations in Y²⁴⁹ and Y²⁵³ result in the loss of TnpA *in vivo* activity (Mendiola and de la Cruz, 1992). Therefore, TnpA is phylogenetically closer to the RCR initiator proteins of plasmids and bacteriophages than to typical TE transposases. These similarities are reflected in the transposition of IS91 which follows the RCR model (Figure 3).

The IS91 transposition model, called rolling-circle transposition (RCT), involves excision of only one IS91 strand from the donor site and its integration within the target DNA (sequential cut-out, copy-in mechanism) (Mendiola et al., 1994; Curcio and Derbyshire, 2003). The terminal regions of IS91 that initiate and terminate the first transposition step

are named *ori91* and *ter91*, respectively. TnpA protein cleaves one strand of DNA within *ori91*, which initiates replication of the leading IS91 DNA strand using the 3'-OH terminus as a primer (Figure 3). The displaced DNA strand is cleaved by TnpA at *ter91* and converted into a RCR-specific circular ssDNA intermediate (Garcillán-Barcia et al., 2001). In the second step, transposase catalyzes the incorporation of the single-stranded intermediate into the new target site. There, following synthesis of the complementary strand (during target DNA replication), a complete copy of IS91 is reproduced (Figure 3) (Curcio and Derbyshire, 2003).

IS91 is the archetype of the IS91 family of insertion sequences, whose representatives, also capable of RCT, are found in the genomes of bacteria belonging to various taxonomic groups. Interestingly, other TEs, unrelated to IS91, but following the RCT model, have been identified in plant and animal genomes. These elements, known as helitrons, belong to the 2nd class of eukaryotic transposons, first detected by *in silico* analysis of genomic sequences. In some cases, these elements represent as much as 6% of the genetic material of the host. They are particularly abundant in the genomes of *Heliconius melpomene* (butterfly), *Myotis lucifugus* (bat) and *Zea mays* (corn).

Helitrons, like members of the IS91 family, do not generate DR sequences during transposition and their transposases resemble RCR initiator proteins (Thomas and Pritham, 2015). Moreover, ssDNA intermediates were identified for one of the helitrons of *M. lucifugus* (HeliBat), suggesting that its transposition mechanism is analogous to that of IS91 (Grabundzija et al., 2016). For other helitrons, a different RCT mechanism is postulated where the transposase simultaneously cleaves single DNA strands within the donor and target sequences, which is consistent with the copy-in model. The DNA strand removed during replication is then directly transposed to the target site without generating ssDNA intermediates (Figure 3) (Curcio and Derbyshire, 2003; Thomas and Pritham, 2015).

Analysis of the RCT mechanism has shown that in many cases RC transposases do not recognize the termination sites located at the end of each TE. This may result in the transposition of genomic DNA fragments located in the immediate vicinity of the TE. Such a phenomenon is frequently observed in the case of helitrons and ISCRs (insertion sequence common regions) (Toleman et al., 2006). Moreover, a correlation between the occurrence of IS91 family members and the proliferation of virulence genes in bacteria was reported (Garcillán-Barcia and de la Cruz, 2002). This indicates the importance of RCT in shaping genomes and its role in the dissemination of non-viable genetic information via HGT (Thomas and Pritham, 2015).

CHARACTERISTICS OF RCR INITIATOR PROTEINS

As explained above, RCR initiator proteins break the continuity of one DNA strand at the *nick* site. This is a multi-step transesterification reaction (Figure 1B) which is reversible and does not require ATP (Byrd and Matson, 1997). To date, two main phylogenetically distinct families of initiator proteins with

endonuclease activity have been described: HUH and Rep_*trans*. Although the mechanism of these endonucleases is very similar, crystallographic studies have revealed significant differences in the structure of their catalytic domains (Chandler et al., 2013; Carr et al., 2016).

HUH Endonucleases

Endonucleases containing the HUH motif (H-H – conserved histidine residues, U – hydrophobic amino acid) constitute a diverse group of proteins. They are involved in various processes including (i) the initiation of replication of bacterial plasmids (Rep), bacteriophages and eukaryotic viruses, (ii) conjugal transfer of DNA (Mob), and (iii) transposition of TEs (Tnp).

The RCR Rep proteins of plasmids are divided into two families: (i) Rep_1 (PF01446), comprising the pC194 and pUB110 initiator proteins, and (ii) Rep_2 (PF01719), the pMV158 and pE194 Rep proteins (Punta et al., 2012). HUH endonucleases from viruses are more diverse. They are classified into four families: (i) Phage_GPA (PF05840), whose prototype is the phage ΦX174 protein A; (ii) Viral_Rep (PF02407), which includes eukaryotic replication proteins, e.g., circovirus PCV2; (iii) Gemini_AL1 (PF00799), responsible for the replication of geminiviruses; and (iv) Rep_N (PF08724), comprising proteins that initiate the replication of linear parvovirus genomes, including AAV (Punta et al., 2012).

The majority of relaxases (Mob proteins) encoded in the DTR regions of plasmid transfer systems, ICE and IME, are also HUH endonucleases. These proteins are classified into six families (Garcillán-Barcia et al., 2009), four of which (MOB_F, MOB_Q, MOB_P and MOB_V) group enzymes with catalytic domains typical of HUH endonucleases. The two remaining families, MOB_H and MOB_C, are significantly different. Although MOB_H proteins share some similarities with the HUH enzymes at the amino acid sequence level (e.g., the presence of a three-histidine HHH motif), there are no crystallographic data to confirm their structural similarity (Garcillán-Barcia et al., 2009; Chandler et al., 2013). This is not the case for members of the MOB_C family, which possess a PD-(D/E)XK-like nuclease domain and lack motifs typical of HUH enzymes. Therefore, they constitute an evolutionarily independent family (Francia et al., 2013).

Apart from their involvement in replication and conjugation, HUH endonucleases can also act as transposases for (i) IS91, ISCR and helitrons, (ii) IStrons – mosaic elements combining the features of TEs and introns (Tourasse et al., 2014), and (iii) repetitive extragenetic palindromes (REP) classified as bacterial interspersed mosaic elements (BIMEs) (Ton-Hoang et al., 2012). However, not all transposases containing the HUH motif perform RCT, e.g., those belonging to the IS200-IS608 bacterial transposon family (He et al., 2015).

The activity of all HUH endonucleases is similar. They first cleave DNA at a specific location and covalently bind to the free 5'-end (Chandler et al., 2013). This produces a stable nucleoprotein complex, and the released 3'-OH end is then used as a replication primer. Three characteristic motifs present in the catalytic domain of HUH endonucleases determine the activity of these proteins (Ilyina and Koonin, 1992): (i) motif I, responsible for the recognition of the DNA sequence; (ii) motif II, with two

histidine residues, responsible for the binding of divalent cations (Mg^{2+} or Mn^{2+}) (metal ions are directly involved in catalysis, plus they participate in catalytic domain stabilization and *origin* recognition), (iii) motif III, containing a catalytic tyrosine residue that carries out a nucleophilic attack on the phosphodiester bond (Yang, 2008; Lorenzo-Díaz et al., 2011; Chandler et al., 2013; Ruiz-Masó et al., 2016). Modified forms of these motifs occur in several HUH endonucleases (Koonin and Ilyina, 1992). For example, HUH proteins may contain either one or two catalytic tyrosine residues in their active site, and can thus be divided into two groups: 1Y and 2Y (Chandler et al., 2013). There may also be variations in the order of the conserved motifs within the catalytic domain (Ilyina and Koonin, 1992). For example, plasmid Rep proteins have motifs arranged in the order I-II-III, while in Mob proteins the scheme is reversed (III-II-I). This difference, however, is insignificant at the level of the tertiary protein structure because the active sites of proteins from both groups are similar (Dyda and Hickman, 2003).

Rep_{trans} Endonucleases (PF02486)

The Rep_{trans} family comprises proteins involved in the initiation of (i) replication of multiple bacterial plasmids, e.g., pT181, pC221 and pUB112 (Projan and Novick, 1988; Balson and Shaw, 1990), (ii) replication of bacteriophages, including CTXΦ (Waldor and Mekalanos, 1996), and (iii) conjugal transfer of some integrative elements, such as ICEBs1 (Lee and Grossman, 2007) and Tn916 (Rocco and Churchward, 2006). Recently, the Rep_{trans} proteins have been classified within the MOB_T family (Punta et al., 2012), even though they do not share significant similarity with HUH endonucleases (homology of $\leq 10\%$) and they lack conserved domains characteristic of HUH proteins (Carr et al., 2016). However, Rep_{trans} and HUH proteins have some common features, as they both (i) use tyrosine as a nucleophile, (ii) catalyze a transesterification reaction resulting in the formation of a phosphotyrosine bond between the protein and the 5'-end of the DNA, and (iii) are dependent on divalent cations (Koepsel et al., 1985; Dempsey et al., 1992; Zhao and Khan, 1997; Khan, 2003; Pastrana et al., 2016).

Until recently there were no crystallographic data available for Rep_{trans} family proteins. The first three-dimensional structures of the main domain of the Rep protein of plasmid pSTK1 from *Geobacillus stearothermophilus* and fusion proteins comprised of domains of *Staphylococcus* spp. Rep proteins were determined by Carr et al. (2016), and they confirmed the difference between the Rep_{trans} and HUH proteins. The catalytic tyrosine residues of Rep_{trans} proteins are located within a β -sheet structure, whereas the YxxK/H motif of HUH endonucleases is always part of an α -helix. Crystallographic analysis also revealed that the binding of divalent cations (Mg^{2+} or Mn^{2+}) by Rep_{trans} proteins is mediated by an arrangement of three acidic amino acids – aspartate (D), aspartate (D) and glutamate (E) (Carr et al., 2016). Moreover, the active sites of the analyzed Rep_{trans} proteins have an almost identical structure consisting of five regions: (i) motif N', containing a highly conserved aspartic acid residue (D), (ii) motif I – RxD, (iii) motif II – T/SxY/ExG, (iv) motif III – YxKxxE, and (v) motif IV – W/LxRxE (the DDE motif residues and the catalytic tyrosine Y are marked

primary function		secondary function			
PROTEIN	PROTEIN FAMILY	ELEMENT			
● ●	Rep68/78	HUH (Rep _N)	AAV	virus	
● ●	Rep60	HUH (Rep ₁)	pTA1060		
● ●	RepA	HUH (Rep ₁)	pC194	plasmids	
● ●	RepB	HUH (Rep ₁)	pUB110		
● ●	Pre	HUH (MOB _v)	pT181		
● ●	Pre	HUH (MOB _v)	pE194		
● ●	MobM	HUH (MOB _v)	pMV158		
● ●	TraX	MOB _c	pAD1		
● ●	TrwC	HUH (MOB _F)	R388		
● ●	Mob02281	HUH (MOB _v)	pBMB0228		
● ●	Mob02282	HUH (MOB _v)	pBMB0228		
● ●	NikB	HUH (MOB _F)	R64		
● ●	ORF50240	MOB _H	ICE <i>c/c</i>	ICES	
● ●	NikK	Rep _{trans}	ICEBs1		
● ●	TecH	Rep _{trans}	Tn916		
● ●	OrfJ	Rep _{trans}	ICES <i>t3</i>		
● ●	Tral	MOB _H	R391		

●	mobilization for conjugal transfer
●	vegetative replication
●	recombination

FIGURE 4 | Multifunctional RCR initiator proteins encoded by different types of genetic elements.

in bold). Similar motifs are also present in the replication proteins of bacteriophages M13 and Cri, representing the Phage_CRI family (PF05144). This may indicate their close relationship to the Rep_{trans} family of proteins (Carr et al., 2016).

THE MULTIFUNCTIONAL CHARACTER OF RCR-INITIATING ENZYMES

RCR initiators are encoded in the genomes of many bacteria, eukaryotic MGEs and viruses. For a long time it was thought that except for the bifunctional Mob/Pre proteins (described below), these enzymes were specialized in performing dedicated biological processes, and so were named Rep, Mob or Tnp proteins. Surprisingly, recent reports have demonstrated the multifunctionality of these RCR initiators (e.g., Wang et al., 2013; Wright and Grossman, 2016; González-Prieto et al., 2017). It is now evident that these proteins, depending on their mechanism of action, can conduct different biological processes associated with RCR (Figure 4).

Proteins Responsible for Mobilization for Conjugation Transfer (Mob) Also Act as Recombinases/Integrases (Int)

In the 1980s, it was noted that some HUH endonucleases may initiate two independent processes. These included the

plasmid recombination (Pre) proteins of three Gram-positive bacterial plasmids: pT181 and pE194 of *Staphylococcus aureus*, and pMV158 of *Streptococcus agalactiae* (Gennaro et al., 1987). These plasmids have the ability to form co-integrates within a strictly defined RSA (recombination site A) sequence. The formation of such structures is independent of the RecA protein, which usually plays a major role in homologous recombination, but requires the presence of the plasmid-encoded Pre protein (Gennaro et al., 1987). Further studies revealed that Pre proteins initiate the transduction of pT181, pE194 and pMV158 (Priebe and Lacks, 1989). Since these proteins can act as both relaxases (Mob) and recombinases they were initially classified within the Mob/Pre family, but they are now included in the MOB_V family (Garcillán-Barcia et al., 2009).

Pre proteins are not an isolated example of multifunctional endonucleases. Similar properties were reported for several other distantly related relaxases, including TraX from the plasmid pAD1 (MOB_C family) (Francia and Clewell, 2002), NikB (MOB_P) from the plasmid R64 (Furuya and Komano, 2003), TrwC from the plasmid R388 (MOB_F family) (César et al., 2006) and a relaxase from an ICE_{clc} element (MOB_H family) (Miyazaki and van der Meer, 2011; Bellanger et al., 2014). However, not all relaxases can act as recombinases despite their similarity, e.g., the TraI proteins from plasmids F and pKM101, which are related to TrwC (MOB_F family). It may be concluded that the recombinase activity of some relaxases is not dependent on the catalytic properties of these enzymes (César et al., 2006).

The factors that actually determine the recombinogenic properties of relaxases have yet to be identified. However, it has been shown that in some cases such activity may be stimulated by other proteins, e.g., NikA from plasmid R64 (Furuya and Komano, 2003) or TrwA from R388 (Agúndez et al., 2012). The structure of the *oriT* itself and the genetic context in which it occurs are also important. This was demonstrated by studies of the ICE_{clc} model, which contains two *origins* of transfer (*oriT1* and *oriT2*) recognized by the same relaxase. In this case, relaxation-dependent recombination exclusively involves the *oriT1* sequence, despite the fact that both *origins* are fully functional and can be used interchangeably during the initiation of conjugal transfer (Miyazaki and van der Meer, 2011).

Relaxation-catalyzed recombination may also lead to the formation of fusion replicons composed of various co-occurring plasmids. This was observed for the plasmid pBMB0228 of *Bacillus thuringiensis*, a cointegrate of two RCR replicons bearing related DTR modules. Each of the Mob proteins encoded by these modules (Mob02281 and Mob02282, 86% amino acid sequence identity) may interact with both *oriT* regions and can catalyze the formation or separation of co-integrates (Wang et al., 2013). These recombination events occur within a short (24 bp) sequence shared by both *oriTs*, containing nearly identical *nick* sites (Wang et al., 2013).

The generation of transient fusion replicons by site-specific recombination has significant implications for the evolution of plasmids. Replicons with a mosaic structure and unique properties (e.g., increased host range or conjugal transferability) might be created in this way.

Relaxase-catalyzed recombination can also occur during conjugal transfer. For example, it was demonstrated that the TrwC protein (HUH endonuclease) is transported to the recipient cell, where it catalyzes the integration of the transferred plasmid within the genome of the new host. Integration occurs only when the target sequence *oriT* R388 is present within the chromosome or a plasmid already residing in the recipient cell. More intriguingly, *oriT* R388-like sequences were also identified in the human genome (on chromosomes 5 and X) and these sequences could be the target for TrwC-catalyzed integration (Agúndez et al., 2012). Therefore, it might be possible to use TrwC relaxase to construct vectors that are potentially useful in gene therapy (González-Prieto et al., 2013, 2017).

Proteins Initiating Replication of Viral Genomes (Rep) Also Promote Their Chromosomal Integration (Int)

The *rep* gene of the previously mentioned adenovirus-associated virus can encode four Rep proteins formed by alternative splicing of transcripts and the activity of two internal promoters. The two largest proteins, Rep78 and Rep68, participate in the RHR initiation of the autonomous form of the AAV genome. In this case, the presence of an adenovirus in the cell is necessary for AAV replication. In the absence of virus, the AAV genome is integrated into the host chromosome and the lysogenic cycle is initiated. AAV is the only known eukaryotic virus that preferentially integrates at a strictly defined site, AAVS1 (adeno-associated virus integration site 1), located within chromosome 19 of the human genome (Smith, 2008). Because of the similarity between the AAVS1 sequence and the RBS and TRS regions of AAV, this virus has become a convenient model for gene therapy vectors (Balakrishnan and Jayandharan, 2014). However, results obtained using high-throughput DNA sequencing have shown that only 45% of AAV insertions occur within AAVS1. The other sites of integration, apart from being similar to the RBS sequence of the virus, were characterized by high transcriptional activity (Janovitz et al., 2013).

The incorporation of the wild-type viral genome into the host cell chromosome depends on the activity of the Rep78 and Rep68 proteins (Rep78/68) (Balagúe et al., 1997; Surosky et al., 1997). However, it was found that even a mutant virus, in which the DNA is not processed by Rep78/68, can integrate into the genome of the host with comparable efficiency (Young and Samulski, 2001). Further studies showed that integration of the virus occurs not at the position where the Rep78/68 nicks the DNA, but at some distance from this cleavage site. In addition, several copies of the viral genome are often integrated at the same target site in the form of head-to-tail concatamers.

Interestingly, recombination events within the target sequence may occur in the absence of AAV DNA. In such cases, the only requirement is the presence of Rep78/68 protein in the human cell (Young et al., 2000). The mechanism of this process is different from reactions catalyzed by previously characterized integrases. According to the model proposed by Young and Samulski (2001), Rep78/68 initiates local RCR

replication by nicking DNA at the TRS in the target sequence. Successive initiations lead to the formation of single-stranded DNA intermediates with Rep attached to the 5'-ends, which are the "hot spots" for recombination events. It is postulated that Rep78/68 recruits cellular proteins involved in DNA recombination, including components of the non-homologous end joining system (NHEJ) (Daya et al., 2009). This system, found in both prokaryotic and eukaryotic cells, is responsible for the repair of double-stranded breaks in DNA (Nash et al., 2009). According to an alternative model, the Rep78 and Rep68 proteins of AAV are contacted by two DNA molecules (viral genome and host chromosome) while interacting with AAVS1 and the viral DNA. This results in the transition of DNA polymerase from the chromosome template to the AAV DNA template and, consequently, in the integration of the viral genome (Henckaerts et al., 2009; Krupovic and Forterre, 2015).

Proteins Initiating ICE Conjugal Transfer (Mob) Are Also Involved in Vegetative Replication (Rep)

Integrative and conjugative elements are a heterogeneous group of MGEs (Guglielmini et al., 2011). They are characterized by the presence of (i) a recombination system encoding an integrase responsible for ICE integration-excision, and (ii) a conjugal transfer system (Carraro and Burrus, 2014).

Integrative and conjugative elements usually occur in the chromosome-integrated form. They may, however, be excised from the genome and form autonomous circular forms (dsDNA) that mimic plasmids, although they lack their own replication systems. Such circular forms of ICE may be reintegrated into the host genome or transferred via conjugal transfer to another bacterial cell, where they integrate within the recipient genome. These elements combine the features of (i) transposons (ability to integrate within DNA, changing location within a host genome), (ii) plasmids (circularity, coding for their own conjugal transfer system), and (iii) bacteriophages (recombination systems involved in DNA integration and excision).

ICEBs1 (20.5 kb) from a strain of *Bacillus subtilis* is one of the model elements in ICE biology research (Auchtung et al., 2016). Surprising results have been reported indicating that the free circular form of ICEBs1, despite the absence of a replication system, can replicate autonomously (Lee et al., 2009; Grohmann, 2010). It was also shown that the replication of this ICE is initiated within the *oriT* by the ICEBs1 relaxase NicK, which initiates both conjugal transfer and vegetative replication of this element. Thus, the dogma that DNA replication initiated within DTR modules occurs only during conjugal transfer was disproved.

Autonomous vegetative replication depending on *oriT* and relaxase has also been demonstrated for other ICEs, e.g., Tn916 (18 kb) the archetype of the Tn916/Tn1545 family (Wright and Grossman, 2016), ICESt3 (28 kb) a representative of the ICESt1/ICESt3 family (Carraro et al., 2016), and R391 an 89-kb element from the STX/R391 family (Carraro and Burrus, 2014). The STX/R391 family is comprised of very large ICEs (sometimes > 100 kb) isolated from *Gammaproteobacteria*

(Carraro and Burrus, 2014). The vegetative RC replication of such large DNA molecules has not been previously observed. It was believed that this replication mechanism was specific only for small (up to ~10 kb) high copy-number replicons (Khan, 1997; Ruiz-Masó et al., 2015). It should be emphasized that the R391 relaxase (TraI) is not phylogenetically related to analogous proteins encoded by elements from other ICE families, and therefore it was classified to a distinct family of relaxases – MOB_H (Garcillán-Barcia et al., 2009). It is noteworthy that the autonomous replication of ICE circular forms is not limited to elements encoding Rep_{trans} relaxases, as was once thought, and it undoubtedly plays an important role in ICE biology.

If the ICE does not excise from the host genome its maintenance in daughter cells will be guaranteed. However, following excision of an ICE from the host genome, replication is critical for its maintenance as a single copy in an autonomous form. Without replication, every cell division could result in the formation of segregated ICE-free clones. Duplication of ICE circular forms by RCR enables a copy of the element to be passed to the daughter cells, which stabilizes their presence in the bacterial population.

Proteins Initiating the Replication of Plasmids (Rep) Also Initiate Conjugal Transfer (Mob)

Analysis of the ICEBs1 conjugation element from *Bacillus subtilis* led to the observation that it is capable of mobilizing various RC replicons including pC194, pBS42 (bearing the REP module from pUB110) and pHP13 (with a pTA1060-like REP). Interestingly, these replicons do not carry DTR systems characteristic of other mobilizable plasmids, and thus lack relaxase genes and *oriT* regions. A more detailed analysis failed to demonstrate a role for the ICEBs1 NicK relaxase in plasmid mobilization. Instead it was shown that the vegetative RCR-initiating proteins (Rep) of these plasmids can act as relaxases (Mob) involved in conjugal transfer. To initiate transfer, such Rep proteins cleave the DNA at the *nick dso*, which then serves as an *oriT* (Lee et al., 2012). The coupling protein ConQ encoded by ICEBs1 plays an important role in the transfer, as it probably binds to the DNA-Rep/Mob relaxosome and directs this complex to the components of the T4SS. As part of the ICEBs1 transfer system the ConQ protein interacts with the NicK parental relaxase, a member of the Rep_{trans} endonuclease family, as well as the pC194, pUB110 and pTA1060 Rep endonucleases belonging to the HUH Rep₁ family. This is particularly noteworthy since the Rep proteins of pC194, pUB110 and pTA1060 are not phylogenetically related to NicK, and there are no characteristic translocation signals (TSs) in their sequences enabling relaxosome transfer via the T4SS machinery (Lang et al., 2010; Alperi et al., 2013). The molecular basis of these interactions has yet to be determined, and the function of the Rep proteins of pC194, pUB110 and pTA1060 in plasmid conjugal transfer remains uncharacterized.

The aforementioned results raise doubts concerning the validity of plasmid classification based on their mobilization potential, if this characteristic is strictly determined by the

presence of DTR or *oriT* systems. According to these criteria the plasmids pC194, pBS42 and pHP13 are non-mobilizable, which is refuted by the experimental data. Most previously analyzed plasmids do not carry DTR systems. However, if a DTR system is not always necessary for conjugal transfer, these plasmids might be mobilizable and play a larger role in HGT than was previously thought. The demonstration of a dual role for Rep/Mob proteins, being capable of initiating vegetative and conjugal RCRs, sheds new light on the biology of the RCR plasmid group as well as ICE elements commonly found in many bacterial genomes (Lee et al., 2012).

CONCLUSION

Rolling-circle replication is involved in various processes conducted by genetic elements, including (i) replication of their autonomous forms (plasmids, bacteriophages, viruses and ICEs), (ii) conjugal transfer (plasmids, ICEs), (iii) transposition (TEs), and (iv) recombination. It should be noted that parts of the host chromosome may also be multiplied in a RC manner. Such replication is mediated by helitrons and ISCRs (able to mobilize adjacent chromosomal genes for transposition) as well as by some conjugal plasmids, which temporarily integrate into the host chromosome. In the latter scenario (as shown in the case of plasmid F), conjugation initiated by the integrated plasmid may result in replication and transfer of a large part of the host chromosome to a recipient cell (Griffiths et al., 1999).

In this review, we have summarized various RCR models developed for individual MGE groups. We have paid particular attention to the multifunctional nature of the RCR initiator proteins, which has not been comprehensively described before. These versatile proteins can determine the course of different biological processes based on RCR. The examples of Mob and Rep proteins that we have discussed show integrase (Int) activity. Other Mob proteins are capable of initiating vegetative RC replication (Rep activity) and some Rep proteins can mobilize plasmid DNA for conjugal transfer (Mob activity) (Figure 4). There are currently no empirical data concerning multifunctional proteins representing the third group of RCR initiators – the Tnp proteins. However, only single examples of this group have been characterized in detail and most of these proteins were analyzed using only bioinformatics tools.

It appears that two factors are critical for the multifunctional activity of RCR initiators: (i) the universal character of the DNA breaking/joining reaction they catalyze, and (ii) the genetic context in which the RCR initiator is present. The type of protein seems to be less important since HUH, Rep_*trans* as well as MOB_C and MOB_H endonucleases can all act as multifunctional proteins.

It may be concluded that many MGEs possess RCR genetic modules that can have additional features, which remain to be discovered, and the eventual characterization of these features may lead to the overturning of widely accepted dogma. For

example, plasmids and viral genomes are no longer considered to be the only genetic elements capable of duplicating their genomes by autonomous replication. Recent studies have shown that integrative and conjugal elements should now be included in this group because their DTR modules also initiate vegetative RC replication. Moreover, this ability is likely to be a characteristic of ICEs (Burrus, 2017), which redefines the properties of this group of MGEs.

Further discoveries in this field may lead to the reclassification of some plasmids. Until now, two main plasmid groups have been distinguished: mobilizable (containing a DTR module, auxiliary transfer system-dependent), and non-mobilizable (i.e., unable to transfer). However, since some Rep proteins were shown to be involved in mobilization, a new subgroup of mobilizable plasmids, the RCR plasmids without DTR modules, should be created. The above examples highlight the need for detailed functional characterization of a wider range of MGEs containing RCR modules.

Research on RCR has a long history. Nonetheless there are still many unanswered questions about this model of DNA replication, as well as the nature and diversity of proteins involved in the replication process. The vast majority of identified RCR initiators are classified within two relatively well-characterized families, HUH and Rep_*trans*. Little is known about the molecular characteristics of other RCR initiators belonging to the MOB_C and MOB_H families.

Current research on RCR is mainly focused on (i) mechanistic studies involving advanced biophysical methods, (ii) determination of the three-dimensional structures of RCR initiators, and (iii) involvement of host factors in this process (Ruiz-Masó et al., 2015; Pastrana et al., 2016). This review highlights the need for in-depth studies on the multifunctional activity of RCR initiators in order to understand the molecular basis of this phenomenon. It will be of particular interest to define the specific factors and conditions responsible for the switch between the distinct biological functions of initiator proteins. It is equally important to identify how common this phenomenon is among different types of MGEs. Thus further studies are required to fill the many remaining gaps in our knowledge concerning the process of RCR – one of the main driving forces of HGT.

AUTHOR CONTRIBUTIONS

PW and DB designed the study. PW wrote the draft manuscript. DB and GP discussed, revised and modified the manuscript to its final version. All authors approved the manuscript for publication.

ACKNOWLEDGMENT

This review was written within the framework of a grant funded by the National Science Centre: UMO-2012/07/N/NZ1/03098 (PRELUDIUM 4).

REFERENCES

- Agúndez, L., González-Prieto, C., Machón, C., and Llosa, M. (2012). Site-specific integration of foreign DNA into minimal bacterial and human target sequences mediated by a conjugative relaxase. *PLOS ONE* 7:e31047. doi: 10.1371/journal.pone.0031047
- Alperi, A., Larrea, D., Fernández-González, E., Dehio, C., Zechner, E. L., and Llosa, M. (2013). A translocation motif in relaxase TrwC specifically affects recruitment by its conjugative Type IV secretion system. *J. Bacteriol.* 195, 4999–5006. doi: 10.1128/JB.00367-13
- Auchtung, J. M., Aleksanyan, N., Bulku, A., and Berkmen, M. B. (2016). Biology of ICEBs1, an integrative and conjugative element in *Bacillus subtilis*. *Plasmid* 86, 14–25. doi: 10.1016/j.plasmid.2016.07.001
- Baas, P. D. (1985). DNA replication of single-stranded *Escherichia coli* DNA phages. *Biochim. Biophys. Acta* 825, 111–139. doi: 10.1016/0167-4781(85)90096-X
- Balagúe, C., Kalla, M., and Zhang, W. W. (1997). Adeno-associated virus Rep78 protein and terminal repeats enhance integration of DNA sequences into the cellular genome. *J. Virol.* 71, 3299–3306.
- Balakrishnan, B., and Jayandharan, G. R. (2014). Basic biology of adeno-associated virus (AAV) vectors used in gene therapy. *Curr. Gene Ther.* 14, 86–100. doi: 10.3389/fnmol.2014.00076
- Balson, D. F., and Shaw, W. V. (1990). Nucleotide sequence of the rep gene of staphylococcal plasmid pCW7. *Plasmid* 24, 74–80. doi: 10.1016/0147-619X(90)90027-A
- Bellanger, X., Payot, S., Leblond-Bourget, N., and Guédon, G. (2014). Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. *FEMS Microbiol. Rev.* 38, 720–760. doi: 10.1111/1574-6976.12058
- Birch, P., and Khan, S. A. (1992). Replication of single-stranded plasmid pT181 DNA in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 89, 290–294. doi: 10.1073/pnas.89.1.290
- Boer, D. R., Ruíz-Masó, J. A., López-Blanco, J. R., Blanco, A. G., Vives-Llàcer, M., Chacón, P., et al. (2009). Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains. *EMBO J.* 28, 1666–1678. doi: 10.1038/emboj.2009.125
- Boer, D. R., Ruíz-Masó, J. A., Rueda, M., Petoukhov, M. V., Machón, C., Svergun, D. I., et al. (2016). Conformational plasticity of RepB, the replication initiator protein of promiscuous streptococcal plasmid pMV158. *Sci. Rep.* 6:20915. doi: 10.1038/srep20915
- Burrus, V. (2017). Mechanisms of stabilization of integrative and conjugative elements. *Curr. Opin. Microbiol.* 38, 44–50. doi: 10.1016/j.mib.2017.03.014
- Byrd, D. R., and Matson, S. W. (1997). Nicking by transesterification: the reaction catalysed by a relaxase. *Mol. Microbiol.* 25, 1011–1022. doi: 10.1046/j.1365-2958.1997.5241885.x
- Cabezón, E., Ripoll-Rozada, J., Peña, A., de la Cruz, F., and Arechaga, I. (2015). Towards an integrated model of bacterial conjugation. *FEMS Microbiol. Rev.* 39, 81–95. doi: 10.1111/1574-6976.12085
- Carr, S. B., Phillips, S. E., and Thomas, C. D. (2016). Structures of replication initiation proteins from staphylococcal antibiotic resistance plasmids reveal protein asymmetry and flexibility are necessary for replication. *Nucleic Acids Res.* 44, 2417–2428. doi: 10.1093/nar/gkv1539
- Carraro, N., and Burrus, V. (2014). Biology of three ICE families: SXT/R391, ICEBs1, and ICES1/ICES3. *Microbiol. Spectr.* 2:MDNA3-0008-2014. doi: 10.1128/microbiolspec.MDNA3-0008-2014
- Carraro, N., Libante, V., Morel, C., Charron-Bourgoin, F., Leblond, P., and Guédon, G. (2016). Plasmid-like replication of a minimal streptococcal integrative and conjugative element. *Microbiology* 162, 622–632. doi: 10.1099/mic.0.000219
- César, C. E., Machón, C., de la Cruz, F., and Llosa, M. (2006). A new domain of conjugative relaxase TrwC responsible for efficient *oriT*-specific recombination on minimal target sequences. *Mol. Microbiol.* 62, 984–996. doi: 10.1111/j.1365-2958.2006.05437.x
- Chandler, M., de la Cruz, F., Dyda, F., Hickman, A. B., Moncalian, G., and Ton-Hoang, B. (2013). Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat. Rev. Microbiol.* 11, 525–538. doi: 10.1038/nrmicro3067
- Curcio, M. J., and Derbyshire, K. M. (2003). The outs and ins of transposition: from mu to kangaroo. *Nat. Rev. Mol. Cell Biol.* 4, 865–877. doi: 10.1038/nrm1241
- Daya, S., Cortez, N., and Berns, K. I. (2009). Adeno-associated virus site-specific integration is mediated by proteins of the nonhomologous end-joining pathway. *J. Virol.* 83, 11655–11664. doi: 10.1128/JVI.01040-09
- Dempsey, L. A., Birch, P., and Khan, S. A. (1992). Six amino acids determine the sequence-specific DNA binding and replication specificity of the initiator proteins of the pT181 family. *J. Biol. Chem.* 267, 24538–24543.
- Dressler, D. (1970). The rolling circle for phiX174 DNA replication. II. Synthesis of single-stranded circles. *Proc. Natl. Acad. Sci. U.S.A.* 67, 1934–1942. doi: 10.1073/pnas.67.4.1934
- Dyda, F., and Hickman, A. B. (2003). A mob of reps. *Structure* 11, 1310–1311. doi: 10.1016/j.str.2003.10.010
- Eisenberg, S., and Kornberg, A. (1979). Purification and characterization of phiX174 gene A protein. A multifunctional enzyme of duplex DNA replication. *J. Biol. Chem.* 254, 5328–5332.
- Francia, M. V., and Clewell, D. B. (2002). Transfer origins in the conjugative *Enterococcus faecalis* plasmids pAD1 and pAM373: identification of the pAD1 *nic* site, a specific relaxase and a possible TraG-like protein. *Mol. Microbiol.* 45, 375–395. doi: 10.1046/j.1365-2958.2002.03007.x
- Francia, M. V., Clewell, D. B., de la Cruz, F., and Moncalián, G. (2013). Catalytic domain of plasmid pAD1 relaxase TraX defines a group of relaxases related to restriction endonucleases. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13606–13611. doi: 10.1073/pnas.1310037110
- Furuya, N., and Komano, T. (2003). NikAB- or NikB dependent intracellular recombination between tandemly repeated *oriT* sequences of plasmid R64 in plasmid or single-stranded phage vectors. *J. Bacteriol.* 185, 3871–3877. doi: 10.1128/JB.185.13.3871-3877.2003
- Garcillán-Barcia, M. P., Bernales, I., Mendiola, M. V., and de la Cruz, F. (2001). Single-stranded DNA intermediates in IS91 rolling-circle transposition. *Mol. Microbiol.* 39, 494–501. doi: 10.1046/j.1365-2958.2001.02261.x
- Garcillán-Barcia, M. P., and de la Cruz, F. (2002). Distribution of IS91 family insertion sequences in bacterial genomes: evolutionary implications. *FEMS Microbiol. Ecol.* 42, 303–313. doi: 10.1111/j.1574-6941.2002.tb01020.x
- Garcillán-Barcia, M. P., Francia, M. V., and de la Cruz, F. (2009). The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.* 33, 657–687. doi: 10.1111/j.1574-6976.2009.00168.x
- Gennaro, M. L., Kornblum, J., and Novick, R. P. (1987). A site-specific recombination function in *Staphylococcus aureus* plasmids. *J. Bacteriol.* 169, 2601–2610. doi: 10.1128/jb.169.6.2601-2610.1987
- Gilbert, W., and Dressler, D. (1968). DNA replication: the rolling circle model. *Cold Spring Harb. Symp. Quant. Biol.* 33, 473–484. doi: 10.1101/SQB.1968.033.01.055
- Gonzalez-Perez, B., Lucas, M., Cooke, L. A., Vyle, J. S., de la Cruz, F., and Moncalián, G. (2007). Analysis of DNA processing reactions in bacterial conjugation by using suicide oligonucleotides. *EMBO J.* 26, 3847–3857. doi: 10.1038/sj.emboj.7601806
- González-Prieto, C., Agúndez, L., Linden, R. M., and Llosa, M. (2013). HUH site-specific recombinases for targeted modification of the human genome. *Trends Biotechnol.* 31, 305–312. doi: 10.1016/j.tibtech.2013.02.002
- González-Prieto, C., Gabriel, R., Dehio, C., Schmidt, M., and Llosa, M. (2017). The conjugative relaxase TrwC promotes integration of foreign DNA in the human genome. *Appl. Environ. Microbiol.* 83, e207–e217. doi: 10.1128/AEM.00207-17
- Grabundzija, I., Messing, S. A., Thomas, J., Cosby, R. L., Bilic, I., Miskey, C., et al. (2016). A Helitron transposon reconstructed from bats reveals a novel mechanism of genome shuffling in eukaryotes. *Nat. Commun.* 7:10716. doi: 10.1038/ncomms10716
- Griffiths, A. J. F., Miller, J. H., Suzuki, D. T., Lewontin, R. C., and Gelbart, W. M. (1999). *An Introduction to Genetic Analysis*, 7th Edn. San Francisco, CA: W. H. Freeman.
- Grohmann, E. (2010). Autonomous plasmid-like replication of *Bacillus ICEBs1*: a general feature of integrative conjugative elements? *Mol. Microbiol.* 75, 261–263. doi: 10.1111/j.1365-2958.2009.06978.x
- Guglielmini, J., Néron, B., Abby, S. S., Garcillán-Barcia, M. P., de la Cruz, F., and Rocha, E. P. (2014). Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res.* 42, 5715–5727. doi: 10.1093/nar/gku194
- Guglielmini, J., Quintais, L., Garcillán-Barcia, M. P., de la Cruz, F., and Rocha, E. P. (2011). The repertoire of ICE in prokaryotes underscores the unity, diversity,

- and ubiquity of conjugation. *PLoS Genet.* 7:e1002222. doi: 10.1371/journal.pgen.1002222
- He, S., Corneloup, A., Guynet, C., Lavatine, L., Caumont-Sarcos, A., Siguier, P., et al. (2015). The IS200/IS605 family and “peel and paste” single-strand transposition mechanism. *Microbiol. Spectrum* 3:MDNA3-0039-2014. doi: 10.1128/microbiolspec.MDNA3-0039-2014
- Henckaerts, E., Dutheil, N., Zeltner, N., Kattman, S., Kohlbrenner, E., Ward, P., et al. (2009). Site-specific integration of adeno-associated virus involves partial duplication of the target locus. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7571–7576. doi: 10.1073/pnas.0806821106
- Henderson, D., and Meyer, R. (1999). The MobA-linked primase is the only replication protein of R1162 required for conjugal mobilization. *J. Bacteriol.* 181, 2973–2978.
- Henry, T. J., and Knippers, R. (1974). Isolation and function of the gene A initiator of bacteriophage phiX174, a highly specific DNA endonuclease. *Proc. Natl. Acad. Sci. U.S.A.* 71, 1549–1553. doi: 10.1073/pnas.71.4.1549
- Ikedo, E. J., Yudelevich, A., and Hurwitz, J. (1976). Isolation and characterization of the protein coded by gene A of bacteriophage phiX174 DNA. *Proc. Natl. Acad. Sci. U.S.A.* 73, 2669–2673. doi: 10.1073/pnas.73.8.2669
- Ilyina, T. V., and Koonin, E. V. (1992). Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaeobacteria. *Nucleic Acids Res.* 20, 3279–3285. doi: 10.1093/nar/20.13.3279
- Janovitz, T., Klein, I. A., Oliveira, T., Mukherjee, P., Nussenzweig, M. C., Sadelain, M., et al. (2013). High-throughput sequencing reveals principles of adeno-associated virus serotype 2 integration. *J. Virol.* 87, 8559–8568. doi: 10.1128/JVI.01135-13
- Khan, S. A. (1997). Rolling-circle replication of bacterial plasmids. *Microbiol. Mol. Biol. Rev.* 61, 442–455.
- Khan, S. A. (2003). DNA-protein interactions during the initiation and termination of plasmid pT181 rolling-circle replication. *Prog. Nucleic Acid Res. Mol. Biol.* 75, 113–137. doi: 10.1016/S0079-6603(03)75004-1
- Koepsel, R. R., Murray, R. W., Rosenblum, W. D., and Khan, S. A. (1985). The replication initiator protein of plasmid pT181 has sequence-specific endonuclease and topoisomerase-like activities. *Proc. Natl. Acad. Sci. U.S.A.* 82, 6845–6849. doi: 10.1073/pnas.82.20.6845
- Koonin, E. V., and Ilyina, T. V. (1992). Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. *J. Gen. Virol.* 73, 2763–2766. doi: 10.1099/0022-1317-73-10-2763
- Krupovic, M., and Forterre, P. (2015). Single-stranded DNA viruses employ a variety of mechanisms for integration into host genomes. *Ann. N.Y. Acad. Sci.* 1341, 41–53. doi: 10.1111/nyas.12675
- Lang, S., Gruber, K., Mihajlovic, S., Arnold, R., Gruber, C. J., Steinlechner, S., et al. (2010). Molecular recognition determinants for type IV secretion of diverse families of conjugative relaxases. *Mol. Microbiol.* 78, 1539–1555. doi: 10.1111/j.1365-2958.2010.07423.x
- Lee, C. A., Babic, A., and Grossman, A. D. (2009). Autonomous plasmid-like replication of a conjugative transposon. *Mol. Microbiol.* 75, 268–279. doi: 10.1111/j.1365-2958.2009.06985.x
- Lee, C. A., and Grossman, A. D. (2007). Identification of the origin of transfer (*oriT*) and DNA relaxase required for conjugation of the integrative and conjugative element *ICEBs1* of *Bacillus subtilis*. *J. Bacteriol.* 189, 7254–7261. doi: 10.1128/JB.00932-07
- Lee, C. A., Thomas, J., and Grossman, A. D. (2012). The *Bacillus subtilis* conjugative transposon *ICEBs1* mobilizes plasmids lacking dedicated mobilization functions. *J. Bacteriol.* 194, 3165–3172. doi: 10.1128/JB.00301-12
- Llosa, M., Gomis-Rüth, F. X., Coll, M., and de la Cruz, F. (2002). Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol.* 45, 1–8. doi: 10.1046/j.1365-2958.2002.03014.x
- López-Aguilar, C., Romero-López, C., Espinosa, M., Berzal-Herranz, A., and Del Solar, G. (2015). The 5'-tail of antisense RNAII of pMV158 plays a critical role in binding to the target mRNA and in translation inhibition of repB. *Front. Genet.* 6:225. doi: 10.3389/fgene.2015.00225
- Lorenzo-Díaz, F., Dostál, L., Coll, M., Schildbach, J. F., Menéndez, M., and Espinosa, M. (2011). The MobM relaxase domain of plasmid pMV158: thermal stability and activity upon Mn²⁺ and specific DNA binding. *Nucleic Acids Res.* 39, 4315–4329. doi: 10.1093/nar/gkr049
- Lorenzo-Díaz, F., Fernández-López, C., Garcillán-Barcia, M. P., and Espinosa, M. (2014). Bringing them together: plasmid pMV158 rolling circle replication and conjugation under an evolutionary perspective. *Plasmid* 74, 15–31. doi: 10.1016/j.plasmid.2014.05.004
- Lorenzo-Díaz, F., Fernández-López, C., Lurz, R., Bravo, A., and Espinosa, M. (2017). Crosstalk between vertical and horizontal gene transfer: plasmid replication control by a conjugative relaxase. *Nucleic Acids Res.* 45, 7774–7785. doi: 10.1093/nar/gkx450
- Mansfeld, A. D., van Teeffelen, H. A., Baas, P. D., and Jansz, H. S. (1986). Two juxtaposed tyrosyl-OH groups participate in phiX174 gene A protein catalysed cleavage and ligation of DNA. *Nucleic Acids Res.* 14, 4229–4238. doi: 10.1093/nar/14.10.4229
- Masai, H., Nomura, N., Kubota, Y., and Arai, K. (1990). Roles of phi X174 type primosome- and G4 type primase-dependent primings in initiation of lagging and leading strand syntheses of DNA replication. *J. Biol. Chem.* 265, 15124–15133.
- Mendiola, M. V., Bernales, I., and de la Cruz, F. (1994). Differential roles of the transposon termini in IS91 transposition. *Proc. Natl. Acad. Sci. U.S.A.* 91, 1922–1926. doi: 10.1073/pnas.91.5.1922
- Mendiola, M. V., and de la Cruz, F. (1992). IS91 transposase is related to the rolling-circle-type replication proteins of the pUB110 family of plasmids. *Nucleic Acids Res.* 20, 3521. doi: 10.1093/nar/20.13.3521
- Mendiola, M. V., Jubete, Y., and de la Cruz, F. (1992). DNA sequence of IS91 and identification of the transposase gene. *J. Bacteriol.* 174, 1345–1351. doi: 10.1128/jb.174.4.1345-1351.1992
- Miyazaki, R., and van der Meer, J. R. (2011). A dual functional origin of transfer in the ICE_{cl} genomic island of *Pseudomonas knackmussii* B13. *Mol. Microbiol.* 79, 743–778. doi: 10.1111/j.1365-2958.2010.07484.x
- Nash, K., Chen, W., Salganik, M., and Muzyczka, N. (2009). Identification of cellular proteins that interact with the adeno-associated virus rep protein. *J. Virol.* 83, 454–469. doi: 10.1128/JVI.01939-08
- Novick, R. P. (1998). Contrasting lifestyles of rolling-circle phages and plasmids. *Trends Biochem. Sci.* 23, 434–438. doi: 10.1016/S0968-0004(98)01302-4
- Pastrana, C. L., Carrasco, C., Akhtar, P., Leuba, S. H., Khan, S. A., and Moreno-Herrero, F. (2016). Force and twist dependence of RepC nicking activity on torsionally-constrained DNA molecules. *Nucleic Acids Res.* 44, 8885–8896. doi: 10.1093/nar/gkw689
- Priebe, S. D., and Lacks, S. A. (1989). Region of the streptococcal plasmid pMV158 required for conjugative mobilization. *J. Bacteriol.* 171, 4778–4784. doi: 10.1128/jb.171.9.4778-4784.1989
- Projan, S. J., and Novick, R. (1988). Comparative analysis of five related staphylococcal plasmids. *Plasmid* 19, 203–221. doi: 10.1016/0147-619X(88)90039-X
- Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., et al. (2012). The Pfam protein families database. *Nucleic Acids Res.* 40, D290–D301. doi: 10.1093/nar/gkr1065
- Rasooly, A., and Novick, R. P. (1993). Replication-specific inactivation of the pT181 plasmid initiator protein. *Science* 262, 1048–1050. doi: 10.1126/science.8235621
- Rocco, J. M., and Churchward, G. (2006). The integrase of the conjugative transposon Tn916 directs strand- and sequence-specific cleavage of the origin of conjugal transfer, *oriT*, by the endonuclease Orf20. *J. Bacteriol.* 188, 2207–2213. doi: 10.1128/JB.188.6.2207-2213.2006
- Ruiz-Masó, J. A., Bordanaba-Ruiseco, L., Sanz, M., Menéndez, M., and Del Solar, G. (2016). Metal-induced stabilization and activation of plasmid replication initiator RepB. *Front. Mol. Biosci.* 3:56. doi: 10.3389/fmolb.2016.00056
- Ruiz-Masó, J. A., Machó, N. C., Bordanaba-Ruiseco, L., Espinosa, M., Coll, M., and Del Solar, G. (2015). Plasmid rolling-circle replication. *Microbiol. Spectr.* 3:PLAS-0035-2014. doi: 10.1128/microbiolspec.PLAS-0035-2014
- Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P., and de la Cruz, F. (2010). Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* 74, 434–452. doi: 10.1128/MMBR.00020-10
- Smith, R. H. (2008). Adeno-associated virus integration: virus versus vector. *Gene Ther.* 15, 817–822. doi: 10.1038/gt.2008.55
- Surosky, R. T., Urabe, M., Godwin, S. G., McQuiston, S. A., Kurtzman, G. J., Ozawa, K., et al. (1997). Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome. *J. Virol.* 71, 7951–7959.

- Tattersall, P., and Ward, D. C. (1976). Rolling hairpin model for replication of parvovirus and linear chromosomal DNA. *Nature* 263, 106–109. doi: 10.1038/263106a0
- Thomas, J., and Pritham, E. J. (2015). Helitrons, the eukaryotic rolling-circle transposable elements. *Microbiol. Spectr.* 3:MDNA3-0049-2014. doi: 10.1128/microbiolspec.MDNA3-0049-2014
- Toleman, M. A., Bennett, P. M., and Walsh, T. R. (2006). ISCR elements: novel gene-capturing systems of the 21st century? *Microbiol. Mol. Biol. Rev.* 70, 296–316. doi: 10.1128/MMBR.00048-05
- Ton-Hoang, B., Siguier, P., Quentin, Y., Onillon, S., Marty, B., Fichant, G., et al. (2012). Structuring the bacterial genome: Y1-transposases associated with REP-BIME sequences. *Nucleic Acids Res.* 40, 3596–3609. doi: 10.1093/nar/gkr1198
- Tourasse, N. J., Stabell, F. B., and Kolsto, A. B. (2014). Survey of chimeric IStron elements in bacterial genomes: multiple molecular symbioses between group I intron ribozymes and DNA transposons. *Nucleic Acids Res.* 42, 12333–12351. doi: 10.1093/nar/gku939
- Waldor, M. K., and Mekalanos, J. J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272, 1910–1914. doi: 10.1126/science.272.5270.1910
- Wang, P., Zhang, C., Zhu, Y., Deng, Y., Guo, S., Peng, D., et al. (2013). The resolution and regeneration of a cointegrate plasmid reveals a model for plasmid evolution mediated by conjugation and *oriT* site-specific recombination. *Environ. Microbiol.* 15, 3305–3318. doi: 10.1111/1462-2920.12177
- Wright, L. D., and Grossman, A. D. (2016). Autonomous replication of the conjugative transposon Tn916. *J. Bacteriol.* 198, 3355–3366. doi: 10.1128/JB.00639-16
- Yang, W. (2008). An equivalent metal ion in one- and two-metal ion catalysis. *Nat. Struct. Mol. Biol.* 15, 1228–1231. doi: 10.1038/nsmb.1502
- Young, S. M. Jr., McCarty, D. M., Degtyareva, N., and Samulski, R. J. (2000). Roles of adeno-associated virus Rep protein and human chromosome 19 in site-specific recombination. *J. Virol.* 74, 3953–3966. doi: 10.1128/JVI.74.9.3953-3966.2000
- Young, S. M. Jr., and Samulski, R. J. (2001). Adeno-associated virus (AAV) site-specific recombination does not require a Rep-dependent origin of replication within the AAV terminal repeat. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13525–13530. doi: 10.1073/pnas.241508998
- Zabala, J. C., de la Cruz, F., and Ortiz, J. M. (1982). Several copies of the same insertion sequence are present in alpha-hemolytic plasmids belonging to four different incompatibility groups. *J. Bacteriol.* 151, 472–476.
- Zhao, A. C., and Khan, S. A. (1997). Sequence requirements for the termination of rolling-circle replication of plasmid pT181. *Mol. Microbiol.* 24, 535–544. doi: 10.1046/j.1365-2958.1997.3641730.x
- Zupan, J. R., and Zambryski, P. (1995). Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiol.* 107, 1041–1047. doi: 10.1104/pp.107.4.1041

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Wawrzyniak, Plucienniczak and Bartosik. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.