Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Plasmid 74 (2014) 15–31

Contents lists available at ScienceDirect

Plasmid

journal homepage: www.elsevier.com/locate/yplas

Bringing them together: Plasmid pMV158 rolling circle replication and conjugation under an evolutionary perspective

Fabián Lorenzo-Díaz [a,1], Cris Fernández-López [b,1], M. Pilar Garcillán-Barcia [c,*], Manuel Espinosa [b,*]

[a] Unidad de Investigación, Hospital Universitario Nuestra Señora de Candelaria and Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias, Centro de Investigaciones Biomédicas de Canarias, Universidad de La Laguna, Santa Cruz de Tenerife, Spain
[b] Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu, 9, E-28040 Madrid, Spain
[c] Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), Universidad de Cantabria – CSIC-SODERCAN, Santander, Cantabria, Spain

ARTICLE INFO

Article history:
Received 24 March 2014
Accepted 22 May 2014
Available online 2 June 2014
Communicated by Saleem Khan

Keywords:
Conjugative transfer
Relaxases
Rolling-circle replicating plasmids
Origins of transfer
Firmicutes

ABSTRACT

Rolling circle-replicating plasmids constitute a vast family that is particularly abundant in, but not exclusive of, Gram-positive bacteria. These plasmids are constructed as cassettes that harbor genes involved in replication and its control, mobilization, resistance determinants and one or two origins of lagging strand synthesis. Any given plasmid may contain all, some, or just only the replication cassette. We discuss here the family of the promiscuous streptococcal plasmid pMV158, with emphasis on its mobilization functions: the product of the *mobM* gene, prototype of the MOB_V relaxase family, and its cognate origin of transfer, *oriT*. Amongst the subfamily of MOB_{V1} plasmids, three groups of *oriT* sequences, represented by plasmids pMV158, pT181, and p1414 were identified. In the same subfamily, we found four types of single-strand origins, namely *ssoA*, *ssoU*, *ssoW*, and *ssoT*. We found that plasmids of the rolling-circle Rep_2 family (to which pMV158 belongs) are more frequently found in Lactobacillales than in any other bacterial order, whereas Rep_1 initiators seemed to prefer hosts included in the Bacillales order. In parallel, MOB_{V1} relaxases associated with Rep_2 initiators tended to cluster separately from those linked to Rep_1 plasmids. The updated inventory of MOB_{V1} plasmids still contains exclusively mobilizable elements, since no genes associated with conjugative transfer (other than the relaxase) were detected. These plasmids proved to have a great plasticity at using a wide variety of conjugative apparatuses. The promiscuous recognition of non-cognate *oriT* sequences and the role of replication origins for lagging-strand origin in the host range of these plasmids are also discussed.

Abbreviations: AbR, antibiotic resistance(s); G−, Gram-negative; G+, Gram-positive; ICEs, integrative conjugative elements; LIC, Leading-strand initiation and control of replication; pRNA, primer RNA; RC, rolling circle; RCR-plasmids, rolling circle-replicating plasmids; RNAP, RNA polymerase; ss, single-stranded; ds, double-stranded; *sso*, single strand origin; T-DNA, transferred DNA strand; T4SSs, type IV secretion systems; T4CP, T4SS-coupling protein; VGT/HGT, vertical/horizontal gene transfer.

* Corresponding authors. Addresses: Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), Universidad de Cantabria – CSIC-SODERCAN, Albert Einstein 22, PCTCAN, 39011 Santander, Cantabria, Spain (M.P. Garcillán-Barcia). Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain (M. Espinosa). Fax: +34 915360432.

E-mail addresses: florenzo@ull.edu.es (F. Lorenzo-Díaz), cflopez@cib.csic.es (C. Fernández-López), maria.garcillan@unican.es (M.P. Garcillán-Barcia), mespinosa@cib.csic.es (M. Espinosa).

[1] Equal contribution.

http://dx.doi.org/10.1016/j.plasmid.2014.05.004
0147-619X/

1. Introduction

The concept developed by Marshall McLuhanin the early 1960s on considering planet Earth as becoming a Global Village (http://projects.chass.utoronto.ca/mcluhan-studies/v1_iss2/1_2art2.htm) has proved to be true fifty years afterwards. Humankind has become global, indeed, not only virtually but physically as well: in addition to the World Wide Web, economical, commercial, and touristic activities have led to unquestionable benefits for the exchange of cultures, but it has also imposed a heavy burden on the rest of the biosphere. A huge part of it is formed by the microbial world, so that going global has increased health risks in the way of outbreaks of epidemics, and the appearance of new transmissible diseases: the swine and avian flu, and the Severe Acute Respiratory Syndrome, to name the most known (see, for instance, the reports by the US-Center for Disease Prevention and Control: http://www.bt.cdc.gov/publications, and its European counterpart: http://www.ecdc.europa.eu/en/Pages/home.aspx). Within this global scenario, not only viruses but also bacterial plasmids have increased their relevance as vehicles to disseminate genetic information among different species: the human activities and human contacts have resulted to be brutal selection processes that have accelerated the horizontal transfer of genetic information among microorganisms (Baquero, 2004; Wellington et al., 2013). Selection of novel bacterial traits derives not only from the need of better starters for fermentation/food production but also from the abuse and misuse of broad-spectrum antibiotics. These last activities have led to the selection of bacteria harboring genetic elements with multiple antibiotic-resistances (AbR), which are rapidly spread (even as epidemic outbursts in hospitals) by horizontal gene transfer, HGT (Anderson and Seifert, 2011; Baquero, 2009). Genes responsible for AbR frequently cluster in the bacterial mobile elements (the mobilome) within the Integrative and Conjugative Elements (ICEs) and transmissible plasmids that, in turn, are platforms to recruit various smaller mobile elements, such as insertion sequences, transposons and integrons that can also encode AbR genes (Frost et al., 2005). Thus, HGT mediated by self-replicating plasmids or by transfer of the islands, plays an essential role in the bacterial, and consequently in the global biodiversity.

In the Plasmid Biology field, the processes pertaining to the dissemination of genetic information stored in plasmids are of special relevance. Those processes involve the two main ways of inheritance, namely vertical (VGT, from mother to daughter cells) and HGT (cell-to-cell) gene transfer (Thomas, 2000). Whereas the former includes DNA replication and partition, this latter being coupled to the cell division, HGT is usually achieved by conjugation between cells of the same or different species (del Solar et al., 1998; Grohmann et al., 2003; Lanka and Wilkins, 1995). Replication and conjugation are, in addition to genetic recombination, the most important sources of genetic variability among plasmids and their hosts. The entire genetic content of the bacterial mobilome amounts up to 25% of the total DNA circulating among bacteria, thus being a shared strong task force for the evolution of the bacterial populations (Ochman et al., 2000; Thomas and Nielsen, 2005). A substantial part of the mobilome is constituted by plasmids, the so-called plasmidome (Walker, 2012).

Plasmids are much more than cloning vectors or tools to over-produce proteins. In addition to their role in the spread of genetic information, bacterial plasmids are excellent models to study a number of biological processes, such as transactions involving macromolecular interactions: protein-DNA, protein–protein, and DNA–RNA (Espinosa, 2013). They also constitute a wealth of information on control of gene expression (del Solar and Espinosa, 2000; del Solar et al., 2002), intracellular distribution of DNA molecules (Reyes-Lamothe et al., 2014), and can be considered as useful models for system biology (Paulson and Ehrenberg, 1998). Excellent books dealing with various aspects of the biology of plasmids have been published (Funnell and Phillips, 2004; Thomas, 2000) and a new one is in press (Alonso and Tomalsky, 2014). In the present review we will concentrate on plasmids replicating by the rolling circle mechanism (RCR-plasmids) because they may encode up to two proteins with endonuclease/topoisomerase-like activity that cleave supercoiled DNA at two different regions on the same molecule, making their study thought-provoking.

2. Replication, conjugation, and mobilization: common themes

Replication and transfer generally require that plasmid-encoded protein(s) interact with their cognate DNA sites to initiate the process. The key initiator players in these two processes are: (i) the generically called Rep proteins involved in vegetative replication and that interact with their cognate origin of replication, *oriV*, and (ii) the relaxases, usually termed Tra or Mob proteins, involved in transfer, and that recognize their cognate origin of transfer, the *oriT*. Thus, interplays between Rep-*oriV* and Tra/Mob-*oriT* will define the VGT and the HGT processes, respectively. The early discovery of conjugation (Lederberg and Tatum, 1946, 1953) and the existence of plasmid-encoded proteins that relax the donor DNA as the first stage in the transfer process, led to a historically broad interest in plasmid transfer among Gram-negative (G−) bacteria. Conjugation in Gram-positive (G+) bacteria was studied relatively later and it is still a matter of active research (Goessweiner-Mohr et al., 2013; Grohmann et al., 2003). Replication involves a number of different strategies devoted to the melting of the DNA strands at the origin of replication (namely theta, strand-displacement and rolling-circle mechanisms), and is usually mediated by the Rep initiator alone or with the help of plasmid- or host-encoded proteins (del Solar et al., 1998). Although DNA melting to initiate the transfer is also required, conjugation, however, involves a single strategy devoted to the unidirectional transfer of mobile elements (plasmids or ICEs) from a donor bacterial cell to a recipient one through physical cell-to-cell contact (de la Cruz et al., 2010; Zechner et al., 2000). Based on their transfer machinery we can distinguish between: (i) self-transmissible (conjugative) plasmids and integrative conjugative elements

(ICEs), which codify all the functions required for their HGT, and (ii) mobilizable plasmids and integrative and mobilizable elements (IMEs), which 'travel light' because they encode only the relaxase and its cognate *oriT*. These latter plasmids and IMEs make use of functions provided by either the host chromosome or by other auxiliary (also termed 'helper') plasmid for their transference.

The assembly of plasmid- and host-encoded proteins in a specific DNA region initiates replication and transfer, thus initiation of both processes requires the generation of multiprotein-DNA-complexes, the replisome or the relaxosome, respectively (del Solar et al., 1998; Lanka and Wilkins, 1995; Pansegrau and Lanka, 1996a). In the case of RCR-plasmids (del Solar et al., 1987; Khan et al., 1981; Novick, 1998; Puyet et al., 1988; te Riele et al., 1986a,b), initiation of replication and mobilization are mechanistically similar processes. Both, Rep and Tra/Mob proteins have endonuclease/topoisomerase-like activities on their DNA targets, the double-strand origin (*dso*) or the *oriT*. These proteins cleave their cognate DNA at the phosphodiester bond of a specific di-nucleotide (the *nic* site) generating a stable amino acyl-DNA adduct (Chandler et al., 2013; de la Campa et al., 1990; Guasch et al., 2003; Khan, 2003; Koepsel et al., 1986; Moscoso et al., 1997; Pansegrau and Lanka, 1996a). The nick introduced by Rep or Tra/Mob proteins generates a free 3′-OH end, which acts as a primer for leading-strand synthesis in both cases, VGT and HGT. In the RCR-plasmids, proteins from the host replicative machinery, at least DNA-polymerases I and III, single-strand (ss) DNA-binding protein and PcrA helicase, participate in the elongation from the 3′-OH end generated by the plasmid-encoded Rep initiator (Anand et al., 2005; Anand and Khan, 2004; Díaz et al., 1994; Khan, 2003, 2005; Machón et al., 2010; Ruiz-Masó et al., 2006; Soultanas et al., 1999; Thomas et al., 2013). To process their DNA substrates, initiators of replication and transfer require the *nic* site being exposed in a single-stranded configuration, which can be achieved by DNA melting and generation of hairpin structures. Hairpin formation would be mediated either by the binding of auxiliary associated proteins, and/or by binding of the initiator or the relaxase. In both cases, the di-nucleotide to be cleaved will be exposed in an unpaired form (Jin and Novick, 2001; Lorenzo-Díaz et al., 2011; Lucas et al., 2010; Noirot et al., 1990; Ruiz-Masó et al., 2007).

Besides the initiator (relaxase) and its binding site (*oriT*), conjugation requires an additional machinery that RCR does not, namely the protein complex that completes the conjugation apparatus. This is a highly specialized protein machinery encoded by the donor plasmid DNA (or the host chromosome in some cases) that includes the coupling protein (T4CP) and the Type IV Secretion System (T4SS) which translocate the relaxase-DNA complex to the recipient cell (Draper et al., 2005; Garcillán-Barcia et al., 2007). T4SSs recruit their substrate by mechanisms still not fully understood for conjugation, and they also participate in other processes involving DNA trafficking between prokaryotic donors and recipient (prokaryotic or eukaryotic) cells, like in the case of Ti and related plasmids (Baron et al., 2002; Bhatty et al., 2013; Cascales et al., 2013; Chen et al., 2005; de Paz et al., 2005; Goessweiner-Mohr et al., 2013; Gomis-Ruth et al., 2004; Hamilton et al., 2000; Llosa et al., 2009; Zhang et al., 2012). The T4CP and the T4SS proteins participate in pumping the transferred DNA (T-DNA)-relaxase complex into the recipient cell (Gomis-Rüth and Coll, 2006; Guasch et al., 2003; Llosa et al., 2002; Matilla et al., 2010). In RCR as well as in transfer, the parental DNA strand will be displaced until either the *dso* or the *oriT* are reconstituted. This last stage involves DNA strand transfer reactions that will terminate either leading strand replication (in the donor cell) or plasmid transfer (in the recipient cell). Intermediates of both processes will be ssDNA molecules that correspond to the parental plus strand (del Solar et al., 1998; Grohmann et al., 2003; Pansegrau and Lanka, 1996a,b; Wilkins and Lanka, 1993). Synthesis of the lagging strand initiates from the so-called single-strand origins (*sso*), as described below.

3. Modular construction of RCR-plasmids

A relevant part of the mobilizable plasmids replicate by the RC mechanism (del Solar et al., 1987; Khan et al., 1981; Novick, 1998; Puyet et al., 1988; te Riele et al., 1986a,b), although not all of them are mobilizable. In the case of small plasmids, isolated primarily from G+ bacteria, two pioneer discoveries were made: (i) the identification of specific single-stranded DNA molecules (ssDNA) as intermediates of plasmid replication (Puyet et al., 1988; te Riele et al., 1986a,b), and (ii) the finding that their Rep initiator proteins exhibited a sequence-specific relaxing activity on supercoiled DNA (Koepsel et al., 1985). These results led to the discovery of RCR-plasmids, which constituted a new class of plasmids that replicate by the asymmetric rolling circle mechanism that share similarities with the replication of ssDNA coliphages (Novick, 1998). Furthermore, the presence of RCR-plasmids in G− bacteria (Yasukawa et al., 1991) helped to cast off the idea of a genetic barrier between the two types of bacteria (del Solar et al., 1993). Thus, the conclusion that plasmid RCR is mechanistically similar to conjugative transfer was soon achieved (Waters and Guiney, 1993). It was interesting to learn that some of the RCR-plasmids encoded, in addition to the Rep topoisomerase-like initiator, another protein with the ability to relax DNA (Caryl et al., 2004; Grohmann et al., 2003; Guzmán and Espinosa, 1997; Smith and Thomas, 2004). These proteins were thought to be involved in inter-plasmidic recombination (plasmid recombination enzymes, Pre; (Projan and Novick, 1988)), although it was later shown that Pre proteins were required for plasmid mobilization (Priebe and Lacks, 1989). Further, the relaxase activity of the pMV158_Pre protein (renamed MobM) on supercoiled DNA, and its *nic* site were demonstrated *in vitro* and *in vivo* (Grohmann et al., 1997; Guzmán and Espinosa, 1997). However, even up today, these proteins are grouped into a family termed Mob-Pre at the database of protein families Pfam (PF01076), and no further investigations on their participation in plasmid recombination have been performed.

In general, RCR-plasmids appear to be constructed as gene cassettes (del Solar et al., 1993; Khan, 1997, 2005)

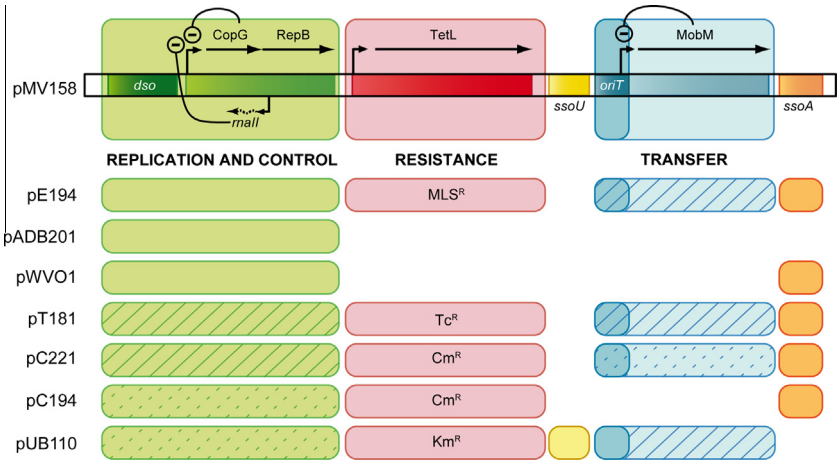

Fig. 1. Genetic organization of representatives of RCR-plasmids. The modular organization of pMV158 from *S. agalactiae* is depicted in the upper part. Genes and the identified promoters are indicated as arrows (arrowheads pointing to the transcription direction). Negative (−) regulatory elements within the replication and control (green) and the transfer (blue) modules are indicated. Below pMV158, other RCR-plasmids are depicted: pADB201 from *Mycoplasma mycoides*, pWV01 from *Lactococcus lactis* and pE194, pT181, pC221, pC194, and pUB110 from *S. aureus*. The corresponding antibiotic resistance (red module) is indicated with the following abbreviations: MLS^R, macrolide/lincosamide/streptogramin B; Tc^R, tetracycline; Cm^R, chloramphenicol; Km^R, kanamycin. Plasmids sharing similar genetic modules are presented in the same color and filling. See Supplementary Tables S1 and S2 for detailed information on the plasmids.

that may have up to four independent modules involved in: (i) Leading-strand initiation and control of replication (LIC); (ii) AbR determinant (DET); (iii) Mobilization (MOB), and (iv) One or two origins for lagging strand replication, *sso* (Fig. 1). This latter kind of origins varies among the different RCR described, but they have been categorized in four types: *ssoA*, *ssoU*, *ssoW*, and *ssoT* depending on their DNA sequence (Khan, 2000, 2005; Kramer et al., 1998a, 1997; Meijer et al., 1995b; van der Lelie et al., 1989), or in their role in plasmid promiscuity (Kramer et al., 1995; Lorenzo-Díaz and Espinosa, 2009). Their structure and roles will be described below (see Table 1). RCR-plasmids may harbor all the cassettes, like the streptococcal plasmids pMV158 and pRW35 (Priebe and Lacks, 1989; van der Lelie et al., 1989; Woodbury et al., 2008), or just the LIC cassette, as in the mycoplasma plasmids pADB201 (Bergemann et al., 1989) and pKMK1 (King and Dybvig, 1992), which are considered to be the smallest RCR-plasmids, unless we consider the hybrid phage-plasmid phasyl (Seufert et al., 1988) as one of them.

Many of these plasmids have left their 'print' on the chromosome of the hosts they might have colonized and, in fact, a bioinformatics survey detected RCR plasmids-related MOB_V-relaxase genes in 63 out of 1207 chromosomes analyzed (Guglielmini et al., 2011). The AbR trace could be considered as a thought-provoking approach to follow the RCR-plasmid 'fate' along evolutionary history. Although we have not performed any further search, a homolog of the pMV158-*tetL* determinant was found in the chromosome of *Bacillus subtilis* (Lacks et al., 1986), whereas a *cat* gene, homologous to the one harbored by plasmids pC194 and pC221 has been described to be present in the chromosome of G+ bacteria, like *Bacillus pumilus* (Harwood et al., 1983) or *Streptococcus pneumoniae* (Pepper et al., 1988), and even in the chromosome of *Clostridium perfringens* (Bannam and Rood, 1991). These findings argue in favor of the role of these RCR-plasmids in the integration and dispersion of resistance genes in the chromosome of hosts that they have colonized.

4. The family: pMV158 and relatives

In general, plasmids have been grouped according to their replicon, since this is the hallmark of the plasmid (Chang et al., 2000; del Solar et al., 1998; Khan, 1996; Nordström et al., 1984; Novick, 1989). As mentioned above, RCR-plasmids were grouped in several families, although the most studied are the Rep_1 family (PF01446), whose prototypes are the staphylococcal plasmids pC194, and pUB110, the Rep_2 family (PF01719) of which the streptococcal plasmid pMV158 and the staphylococcal plasmid pE194 are the representatives, and the Rep_trans family (PF02486) represented by the staphylococcal plasmids pT181 and pC221 (del Solar et al., 1998, 1993; Gruss and Ehrlich, 1989; Khan, 1997, 2000, 2005; Novick, 1989). These three families harbor all the cassettes with the exceptions of pC194 that does not contain a MOB module and the lagging-strand origin *ssoU*, which is reported to be present only in pUB110 and pMV158 (Fig. 1). Curiously, these two latter plasmids also share identical *oriT* sequences. *In vitro*, the MobM-protein from pMV158 was able to relax supercoiled DNA from both plasmids (Fernández-López et al., 2013a). Out of the representatives of the different families, pC194 has been reported to be one of the most promiscuous, since it was shown to replicate not only in staphylococci, streptococci and bacilli (Ballester et al., 1990; Horinouchi and Weisblum, 1982), but also in *Escherichia coli* and in the yeast *Saccharomyces* as well (Goursot et al., 1982). In the case of pMV158, it has been shown to exhibit an extraordinary host range. The plasmid was primarily isolated from *Streptococcus agalactiae* (Burdett, 1980) and, along the years, it has been

Table 1
MOB_{V1} plasmids.

Plasmid name [a]	Nucleotide GenBank Acc. No.	Plasmid size (kb)	Relaxase GenBank Acc. No. [b]	Relaxase tree cluster [c]	Host	Replication initiator family [d]	*sso* type (position related to *mob* gene) [e]	Virulence or antibiotic/metal resistance [f]
pMV158	NC_010096.1	5.541	YP_001586274	pMV158	*Streptococcus agalactiae*	Rep_2	*ssoA* (3′) + *ssoU* (5′)	Tc
pER13	NC_002776.1	4.139	NP_115336.1	pMV158	*Streptococcus thermophilus*	Rep_2	*ssoA* (3′)	
pSMQ172	NC_004958.1	4.230	NP_862547.1	pMV158	*Streptococcus thermophilus*	Rep_2	*ssoA* (3′)	
pGA2000	NC_019252.1	4.967	YP_006961085.1	pMV158	*Streptococcus pyogenes*	Rep_2	*ssoA* (3′)	MLS(B)
pGB2002	NC_015971.1	6.825	YP_004831084.1	pMV158	*Streptococcus agalactiae*	Rep_2	*ssoA* (3′)	MLS(B)
pRW35	NC_010423.2	4.968	YP_001716200.1	pMV158	*Streptococcus pyogenes*	Rep_2	*ssoA* (3′) + *ssoU* (5′)	MLS(B)
pDRPIS7493	NC_015876.1	4.727	YP_004769541.1	pMV158	*Streptococcus pseudopneumoniae*	Rep_2	ND	
pSSU1	NC_002140.1	4.975	NP_053061.1	pMV158	*Streptococcus suis*	Rep_2	*ssoA* (3′)	
pYSI8	NC_010936.1	4.973	YP_001967741.1	pYSI8	*Lactobacillus sakei*	Rep_2	*ssoT* (5′)	Lin
pK214	NC_009751.1	29.871	YP_001429536.1	pYSI8	*Lactococcus lactis*	Rep_trans and Rep_3 + L_lactis_RepB_C	ND	MEP, Strp, Chlr, Tc
			YP_001429523.1 (MOB_Q)					
pUR2941	HF583290.1	20.876	CCQ43999.1	pYSI8	*Staphylococcus aureus*	RepA_N + DnaB_2 and unknown and truncated	*ssoA* (5′)	Kan/Neo, Tc, MLS(B), Cd, Cu
pCPS49	NC_019142.1	5.292	YP_006958108.1	pYSI8	*Staphylococcus aureus*	Rep_2	*ssoA* (3′)	PLS(A)
pSTE1	NC_020237.1	11.951	YP_007419104.1	pSYI8	*Staphylococcus hyicus*	Rep_trans and HTH_Hin_like and truncated	ND	Strp, MLS(B), Tc
			YP_007419109.1	pKKS825				
pKKS825	NC_013034.2	14.363	YP_003084337.1	pKKS825	*Staphylococcus aureus*	Rep_1 and HTH_Hin_like and Rep_3	ND	Kan/Neo, Tc, Trim, PLS(A)
			YP_003084330.1 YP_004679012.1 (MOB_Q, truncated and fused to Rep_1)	pKKS825				
pDB2011	NC_021513.1	7.641	YP_008119849.1	pKKS825	*Listeria innocua*	Rep_1 and HTH_Hin_like	ND	MLS(B), Spec, Trim
pS130a	AUPT01000023.1	8.882	EPZ04218.1	pKKS825	*Staphylococcus aureus*	HTH_11	ND	Ery, Tc, Kan, Ble
pSCFS1	NC_005076.1	17.108	NP_899176.1	pKKS825	*Staphylococcus sciuri*	HTH_Hin_like and Rep_3	ND	Flr/Chlr, MLS(B), Spec
			NP_899168.1	pNM11				
pLB4	M33531.1	Incomplete	AAA25252.1	pLB4	*Lactobacillus plantarum*	Rep_2	*ssoT* (5′)	
pMRI_5.2	NC_019900.1	5.206	YP_007215174.1	pLB4	*Lactobacillus plantarum*	Rep_1 and Rep_2	*ssoT* (3′)	
pLAC1	NC_014164.1	3.478	YP_003650630.1	pLB4	*Lactobacillus acidipiscis*	Rep_1	ND	
pPLA4	AF304384.2	8.135	ABG23031.1	pLB4	*Lactobacillus plantarum*	Rep_3	ND	Bacteriocin
pPB1	NC_006399.1	2.899	YP_138221.1	pLB4	*Lactobacillus plantarum*	Rep_2	*ssoT* (5′)	
LkipL48	NC_014135.1	3.196	YP_003620509.1	pLB4	*Leuconostoc kimchii*	Rep_2	ND	
pMBLR00	NC_019353.1	3.370	YP_006964795.1	pLB4	*Leuconostoc mesenteroides*	Rep_2	ND	
pLAB1000	M55222.1	Incomplete	P35856.1	pLB4	*Lactobacillus hilgardii*	Rep_1	ND	
pSMA23	NC_010242.1	3.497	YP_001649176.1	pSMA23	*Lactobacillus casei*	Rep_1	ND	
pLC88	U31333.1	Incomplete	AAA74581.1	pSMA23	*Lactococcus casei*	Rep_1	ND	
p141	AB517606.1	Incomplete	BAH97325.1	pSMA23	*Lactobacillus plantarum*	Rep_1	ND	
pCD034-1	NC_016035.1	3.424	YP_004869658.1	pSMA23	*Lactobacillus buchneri*	Rep_1	*ssoT* (3′)	
pM4	NC_009666.2	3.320	YP_001621756.1	pSMA23	*Lactobacillus plantarum*	Rep_1	*sso-new*	
pF8801	NC_007593	5.558	YP_398641.1	pSMA23	*Pediococcus damnosus*	Rep_1	ND	

(*continued on next page*)

Table 1 (*continued*)

Plasmid name [a]	Nucleotide GenBank Acc. No.	Plasmid size (kb)	Relaxase GenBank Acc. No. [b]	Relaxase tree cluster [c]	Host	Replication initiator family [d]	*sso* type (position related to *mob* gene) [e]	Virulence or antibiotic/metal resistance [f]
pCD034-2	NC_016034.1	2.707	YP_004869655.1	pSMA23	*Lactobacillus buchneri*	Rep_2	*ssoT* (3′)	
pG6301	NC_019372.1	3.516	YP_006965557.1	pSMA23	*Lactobacillus plantarum*	Rep_1	ND	
pLB925A02	NC_012549.1	3.524	YP_002790952.1	pSMA23	*Lactobacillus brevis*	Rep_1	ND	
pSD11	NC_014919.1	3.225	YP_004134615.1	pSMA23	*Lactobacillus brevis*	Rep_1	ND	
pGL2	NC_016981.1	4.572	YP_005352352.1	pGL2	*Lactococcus garvieae*	Rep_2	*ssoW* (3′)	Bacteriocin
pAMalpha1	NC_005013.1	9.759	NP_863358.1 NP_863352.1	pAMalpha1 pUB110	*Enterococcus faecalis*	Rep_1 + Rep_1 and Rep_3	*ssoU* (5′)	Tc
Unnamed	GG670384.1	Incomplete	EEU18290.1	pAMalpha1	*Enterococcus faecalis*	Rep_3	ND	
Unnamed	GG692894.1	Incomplete	EEU66435.1 EEU66441.1	pAMalpha1 pUB110	*Enterococcus faecalis*	Rep_1 + Rep_1 and Rep_3	ND	Tc
EF62pA	NC_017314.1	5.143	YP_005706998.1	pAMalpha1	*Enterococcus faecalis*	Rep_3		
pBMO2	NC_004930.1	3.854	NP_862027.1	pBM02	*Lactococcus lactis*	Rep_2	*ssoW* (3′)	
pI4	AF300457.1	Incomplete	AAG28767.1	pI4	*Bacillus coagulans*	–	*ssoT* (5′)	Coagulin
pTXW	NC_013952.1	3.178	YP_003517730.1	pTXW	*Lactobacillus paracasei*	Rep_2	*ssoW* (5′)	
pWCZ	NC_019669.1	3.078	YP_007027014.1	pTXW	*Lactobacillus paracasei*	Rep_2	ND	
pLA106	NC_004985.1	2.862	NP_862697.1	pTXW	*Lactobacillus acidophilus*	Rep_2	ND	
pRCEID2.9	NC_017466.1	2.952	YP_005849229.1	pTXW	*Lactobacillus casei*	Rep_2	ND	
pT181	NC_001393.1	4.439	NP_040472.1	pT181	*Staphylococcus aureus*	Rep_trans	*ssoA* (3′)	
pSEQU3	AVBD01000026.1	4.846	ERH33926.1	pT181	*Staphylococcus equorum*	Rep_trans	ND	Tc
pKH17	NC_010284.1	4.441	YP_001654074.1	pT181	*Staphylococcus aureus*	Rep_trans	ND	Tc
pSE-12228-01	NC_005008.1	4.439	NP_863257.1	pT181	*Staphylococcus epidermidis*	Rep_trans	ND	Tc
pKH6	NC_001767.1	4.439	NP_053796.1	pT181	*Staphylococcus aureus*	Rep_trans	ND	Tc
pS0385-1	NC_017334.1	5.246	YP_005735514.1	pT181	*Staphylococcus aureus*	Rep_trans	ND	Tc
SAP095B	NC_013312.1	4.439	YP_006937497.1	pT181	*Staphylococcus aureus*	Rep_trans	ND	Tc
pSBK203	U35036.1	Incomplete	AAA79055.1	pT181	*Staphylococcus aureus*	Rep_trans	ND	Chlr
pKH7	NC_002096.1	4.118	NP_052168.1	pT181	*Staphylococcus aureus*	Rep_trans	ND	Chlr
pS1c	AUPS01000031.1	3.899	EQM91159.1	pT181	*Staphylococcus aureus*	Rep_trans	ND	
SAP047A	NC_013331.1	28.974	YP_006938074.1	pT181	*Staphylococcus aureus*	Rep_1 and Rep_3 and RepA_N	ND	Cd, β-lac, enterotoxin G
pPCZ1	NC_013539.1	4.738	YP_003329162.1	pPCZ1	*Planococcus* sp.	Rep_3	*ssoA* (3′)	
pGI1	NC_004335.1	8.254	NP_705753.1	pGI1	*Bacillus thuringiensis*	Rep_1	*ssoT* (5′)	
pCT8513	NC_017207.1	8.513	YP_005569975.1	pGI1	*Bacillus thuringiensis*	Rep_1	ND	
pB52y	AVEZ01000046.1	6.283	EQM25212.1	pGI1	*Bacillus licheniformis*	Unknown	ND	
pBMB9741	NC_001272.2	6.578	YP_724461.1	pGI1	*Bacillus thuringiensis*	Rep_1	ND	
pIS56-8	NC_020377.1	8.251	YP_007482091.1	pGI1	*Bacillus thuringiensis*	Rep_1	ND	
BTB_7p	NC_018882.1	7.635	YP_006931124.1	pGI1	*Bacillus thuringiensis*	Rep_1	ND	
pHT7	NC_020243.1	7.635	YP_007425204.1	pGI1	*Bacillus thuringiensis*	Rep_1	ND	
BTB_9p	NC_018886.1	8.513	YP_006931150.1	pGI1	*Bacillus thuringiensis*	Rep_1	ND	
pE33L5	NC_007104.1	5.108	YP_245942.1	pE33L5	*Bacillus cereus*	HTH_36	*ssoT* (5′)	
pW_3	ABCZ02000102.1	Incomplete	EDX54234.1	pE33L5	*Bacillus cereus*	Truncated	ND	
pH308197_11	NC_011340.1	11.567	YP_002267516.1	pE33L5	*Bacillus cereus*	HTH_CRP	ND	
p1414	NC_002075.1	7.949	NP_049443.1	p1414	*Bacillus subtilis*	Rep_1	*ssoT* (3′)	
pBamNAU-B3a	NC_022531.1	8.438	YP_008628645.1	p1414	*Bacillus amyloliquefaciens*	–	ND	
pBA45-1	NC_020273.1	8.009	YP_007447244.1	p1414	*Bacillus amyloliquefaciens*	Rep_1	ND	
pPL1	NC_013537.1	6.704	YP_003329154.1	p1414	*Bacillus subtilis*	Rep_1 and unknown	ND	
pTA1015	NC_001765.1	5.807	NP_053784.1	p1414	*Bacillus subtilis*	Rep_1	*ssoT* (3′)	
pBS608	NC_006825.1	6.611	YP_195753.1	p1414	*Bacillus subtilis*	Rep_1	ND	

pTA1060	NC_001766.1	8.737	NP_053788.1	p1414	*Bacillus subtilis*	Rep_1 and unknown	*ssoT* (3′)	
pBSG3	NC_014104.1	8.439	YP_003600423.1	p1414	*Bacillus amyloliquefaciens*	Rep_1	*ssoT* (3′)	
pSD853_7.9	NC_015392.1	7.860	YP_004376195.1	p1414	*Salmonella enterica*	Rep_1	ND	
pTRACA20	NC_013279.1	3.780	YP_003208332.1	pTRACA20	Uncultured bacterium	DNA_primase_S	*ssoW* (3′)	
pUB110	NC_001384.1	4.548	NP_040431.1	pUB110	*Staphylococcus aureus*	Rep_1 + Rep_1	*ssoU* (5′)	Neo, Ble
pSES22	NC_007621.1	4.040	YP_415518.1	pUB110	*Staphylococcus saprophyticus*	Rep_1	ND	MLS(B)
pERGB	JN970906.1	Incomplete	AEW23141.1	pUB110	*Staphylococcus aureus*	Rep_1 and Rep_1	ND	PLS(A), Tb, Tc, Trim
pTB19	M63891.1	Incomplete	AAA98305.1	pUB110	*Geobacillus stearothermophilus*	Rep_1	*ssoU* (5′)	Tc, Ble
			AAA98307.1	pUB110				
pV7037	HF586889.1	Incomplete	CCQ71694.1	pUB110	*Staphylococcus aureus*	RepA_N and truncated	ND	Tc, Cd
pBC16	NC_001705.1	4.630	NP_043522.1	pUB110	*Bacillus cereus*	Rep_1 + Rep_1	*ssoU* (5′)	Tc
pSWS47	NC_022618.1	28.743	YP_008719890.1	pUB110	*Staphylococcus epidermidis*	Rep_3 and truncated and truncated and RepA_N	ND	PLS(A), Kan/Neo, Tc, Trim
			YP_008719902.1 (MOB_P)					
pTB53	D14852.1	Incomplete	BAA03580.1	pUB110	*Bacillus* sp.	–	ND	
pIP1714	AF015628.1	Incomplete	AAC61672.1	pUB110	*Staphylococcus cohnii*	Rep_1 + Rep_1	ND	PLS(A), MLS(B)
pNM11	NC_019558.1	11.383	YP_007016413.1	pNM11	*Planococcus citreus*	Rep_3	ND	
pBS-03	JQ394981.1	Incomplete	AFJ49144.1	pNM11	*Bacillus* sp.	Rep_1	ND	Flr/Chlr, Strp
			AFJ49142.1 (MOB_V, truncated)					
pSS-03	NC_016054.1	7.122	YP_004888092.1	pNM11	*Staphylococcus arlettae*	Rep_1	ND	Flr/Chlr, MLS(B)
			YP_004888090.1 (MOB_V, truncated)					
pJ612	NC_019186.1	5.048	YP_006959664.1	pJ612	*Haemophilus influenzae*	Rep_3	ND	β-lac
pA1606	NC_019180.1	5.646	YP_006959644.1	pJ612	*Haemophilus influenzae*	Rep_3	ND	β-lac

[a] Plasmids whose relaxases were retrieved by a PSI-BLAST using MobM_pMV158 (300-N terminal residues) as a query are listed.

[b] Underlined accession numbers denote those relaxase genes that are probably misannotated in the GenBank database (*i.e.* extended N-terminal sequence respect to that of MobM_pMV158).

[c] It locates the corresponding plasmid in one of the cartooned clusters of Fig. 4, for which a prototype was selected.

[d] Replication initiation protein family. When more than one, their names are separated by "and". "+" is used for initiators that contain more than one pfam domain. Further details on replication initiation protein families can be found at http://pfam.sanger.ac.uk/.

[e] Only the previously identified *ssos* are annotated. Most of the *sso* sequences span 200–300 bp and are located in close proximity to the transfer module, with the exception of plasmid pUR2941 (*ssoA* was mapped 7 kb upstream *mob*). pM4 plasmid has a new type of *sso* as described in (Yin et al., 2009). ND, not determined.

[f] Antibiotic or metal resistance to: Tc, tetracycline; MLS(B), macrolide/lincosamide/streptogramin B; Lin, lincosamide; MEP, macrolide efflux protein; Strp, Streptomycin; Chlr, chloramphenicol; Kan, kanamycin; Neo, neomycin; Cu, copper; Cd, cadmium; PLS(A), pleuromutilins/lincosamide/streptogramin A; Trim, trimethoprim; Spec, spectinomycin; Ery: erythromycin; Ble, bleomycin; Flr, florfenicol; β-lac, beta-lactam; Tb, tobramycin.

transferred in the laboratory (either by mobilization or by transformation) to more than 20 bacterial species, from Firmicutes to α-Proteobacteria and in all of them it replicates stably (del Solar et al., 1998; Espinosa, 2013; Lacks et al., 1986).

After a BlastP search using prototypes of the three RCR-plasmid families, we found the following distribution: (i) in Rep_1 (prototype RepA from pC194), 334 non-redundant homologs in Bacillales, and 157 in Lactobacillales; (ii) in Rep_2 (prototype taken, RepE initiator from pE194), 44 non-redundant homologs in Bacillales and 242 in Lactobacillales; (iii) in Rep_trans (RepC from pT181 as prototype), 176 non-redundant homologs in Bacillales and 241 in Lactobacillales. These figures suggest that whereas the Rep_1 family of RCR-plasmids would prefer to colonize Bacillales and the Rep_2 family of pMV158 would be better fitted to Lactobacillales, the Rep_trans family of RCR-plasmids would be distributed more evenly. In general, we found a good correlation between the G + C content of RCR-plasmids and their respective hosts (Espinosa et al., 1995), with the exception of the staphylococcal plasmid pUB110 which has a G + C content close to 45%, making it a plasmid which has been considered more like a bacilli than a staphylococci replicon (Alonso et al., 1988; McKenzie et al., 1986).

Traditionally, RCR-plasmids have been classified according their LIC module (del Solar et al., 1998, 1993). According to this criterion, we performed a PSI-BLAST search for plasmids of the pMV158 family using the initiator $RepB_{pMV158}$ as the query. The results are compiled in Supplementary Table S1. The 78 plasmids retrieved belong to the Rep_2 family. Out of them, 55 did not encode any relaxase whereas the remaining 23 contained relaxase genes belonging to the MOB_V family (see below).

Many of the retrieved plasmids lack any distinguishable marker, although resistance to a variety of antibiotics, and even one instance of resistance to arsenate was found. A phylogenetic tree of the 78 plasmids retrieved was constructed (Supplementary Fig. S1), using the protein GpA initiator from phage ΦX174 as the out-group representative. There will be a detailed in-depth analysis on the LIC module of plasmids of the pMV158 family to be published elsewhere (G. del Solar and J.A. Ruiz-Masó, personal communication).

Furthermore, most of the retrieved plasmids were primarily isolated from Lactobacillales, but as shown in Supplementary Fig. S1 they are also widely distributed in other taxonomic orders of Firmicutes, Tenericutes and Proteobacteria, with notable examples of plasmids from *Mycoplasma* (eight plasmids). Although not retrieved in this search, there is an RCR-plasmid from *Mycoplasma yeatsii*, pMyBK1, which deserves a mention here because of two features (Kent et al., 2012; Breton et al., 2012). Firstly, it is unique at encoding an initiator of RCR different from the ones described here. Secondly, pMyBK1 is the only example of mobilizable plasmid in genus *Mycoplasma*, and it is precisely a MOB_V plasmid. Inspecting the sequence of the *mob* gene of pMyBK1, we found that it harbors 39 UGA codons (out of a total of 520 codons). Since in *Mycoplasma* UGA specifies "tryptophan" instead of "stop" (Halbedel and Stülke, 2007; Inamine et al., 1990; Yamao et al., 1985), we can conclude that the transfer of pMyBK1 would be limited to *Mycoplasma* species.

5. The MOB_V family

Classification of plasmids according to the relaxases and origins of transfer they carry has supposed a novel definition of plasmid families (Francia et al., 2004; Garcillán-Barcia et al., 2009). This classification provided a more global view of HGT by conjugation, whereas classification by replicons would point more to the vertical transfer. The HGT-based classification has showed that there are many plasmids that do not encode transfer functions (Garcillán-Barcia et al., 2011; Smillie et al., 2010). RCR-plasmids of Firmicutes fall mostly in the so-called MOB_V family, being the MobM relaxase from the streptococcal plasmid pMV158 the representative of the family (Garcillán-Barcia et al., 2009). It was believed that plasmids are DNA molecules that result from shuffling of various gene cassettes, evolving independently one of each other (Osborn et al., 2000). As we will discuss below, analysis of evolutionary roots of these cassettes motivates to think that there could be crosstalk between them, as some module combinations prevail among others.

To update the inventory of elements composing the MOB_{V1} subfamily we have performed a PSI-BLAST search using $MobM_{pMV158}$ as the query. The retrieved MOB_{V1} plasmids ranged from 2.7 to 30 kb (median = 5.1 kb; Fig. 2). Roughly half of them coded for antibiotic resistance genes (mainly to a wide variety of protein synthesis inhibitors, such as macrolides, lincosamides, aminoglycosides, streptogramins, amphenicols and tetracyclines) or other virulence traits like resistance to heavy metals and production of bacteriocins (Table 1). The organization of the mobilization region was found to be similar in all members: the *oriT* located upstream, close to the *mob* gene. Furthermore, the *nic* site was placed on the same strand as the *mob* gene, as expected.

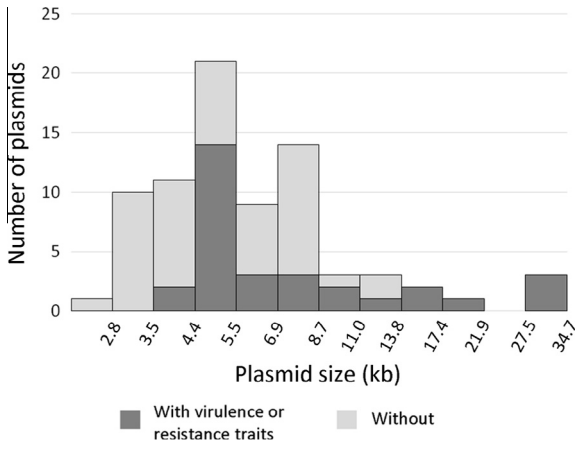

Fig. 2. Distribution of MOB_{V1} plasmids according to their size. The *X* axis was built by using the $\log_{10}$ of plasmid size values. Each bar represents the abundance of plasmids for a given size range, which is indicated at each side of the bar. Plasmids encoding genes for metal or antibiotic resistance and/or bacteriocin production are indicated in dark gray. The rest are indicated in light gray. Data was obtained from Table 1.

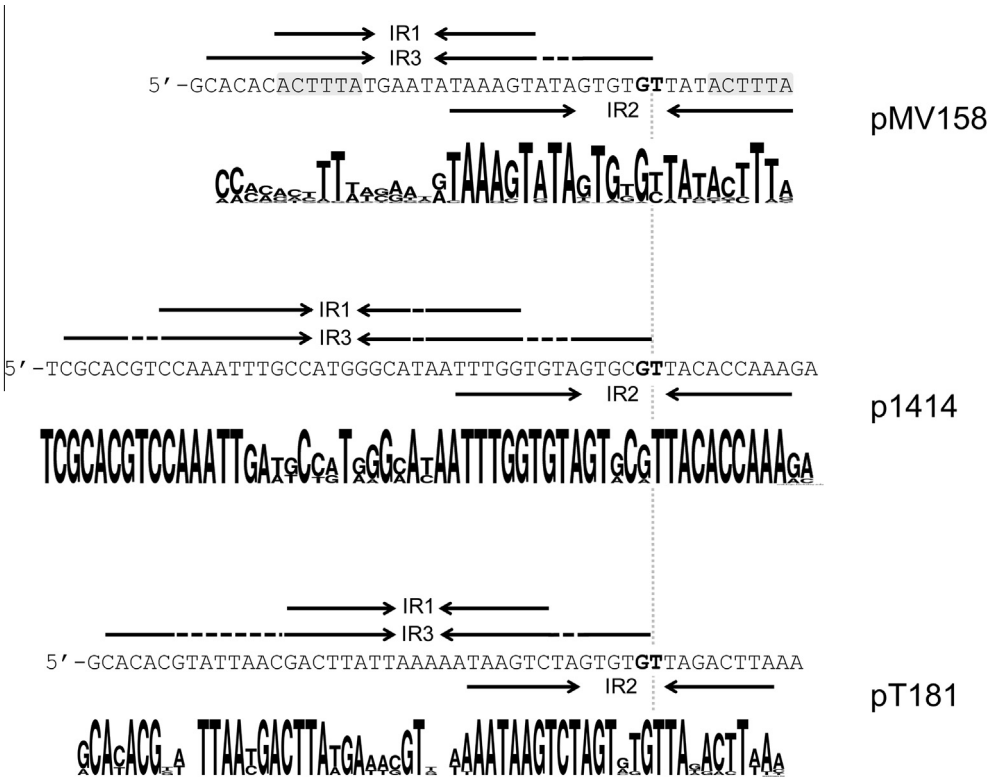

Fig. 3. The origin of transfer (*oriT*) in MOB_{V1} representative plasmids. Comparison of the *oriT* sequences located in the pMV158, p1414 and pT181 plasmids, lined up by the position of the *nic* site (boldface letters). The three overlapping inverted repeats (IR1, IR2 and IR3) are depicted by arrows, and dashed lines indicate the position of those unpaired bases in the predicted secondary structures they could form. Gray background in the pMV158-*oriT* sequence denotes the position of a conserved repeated region. Consensus of the *oriT* sequences aligned in Supplementary Fig. S2 were prepared using WebLogo (version 2.8.2; Crooks et al., 2004).

5.1. The origin of transfer (oriT)

Plasmid conjugation initiates through the assembly of the relaxosome on the *oriT*, a region that contains inverted (IR) and/or direct (DR) repeats, A + T-rich tracts and, most importantly, the *nic* site. The relaxase requires that its target is presented as ssDNA to cleave it and generate the relaxase-DNA adduct, which will be pumped through a T4SS. Once in the recipient, the *oriT* is reconstituted by a transesterification reaction mediated by the relaxase to close the incoming molecule (reviewed in (Chandler et al., 2013).

The *oriT* of pMV158 spans 41 bp upstream of the *mobM* gene, exhibits a high A + T content (75.6%) and has three IRs (IR1 to IR3; Lorenzo-Díaz et al., 2011). Furthermore, a 6-bp sequence (5′-ACTTTA-3′) is repeated in the IR1/IR3 left-arm and the IR2 right-arm (Fig. 3). The *nic* site was firstly mapped *in vitro* between coordinates 3595–3596 (dinucleotide 5′-GpT-3′) in the pMV158 sequence (GenBank Acc. No. NC_010096; (Guzmán and Espinosa, 1997)). Then, the *nic* sites for the pMV158 and pE194 *oriT*s were determined in their respective hosts *in vivo* (Grohmann et al., 1997), mapping exactly in the same position as already suggested by the previous *in vitro* results. *In silico* analysis revealed that the IRs could generate three alternative stem-loop structures in which the position of the *nic* site would be placed in different positions: (i) located 8-bp upstream to IR1, (ii) in the IR2 inter-arm region, or (iii) at the 3′-end of the IR3 (Lorenzo-Díaz et al., 2011). IR3 includes the IR1 sequence and, since IR1/3 and IR2 partially overlap (see Fig. 3), the generation of cruciform secondary structures by one of them would hinder the formation of the other, indicating that the target DNA accessibility by the relaxase could depend on the plasmid DNA superhelicity (Fernández-López et al., 2014). Our *in vitro* analysis demonstrated that MobM binds specifically to ssDNA encompassing IR1/3 with high affinity (Lorenzo-Díaz et al., 2011), which allowed the MobM protein to repress its own synthesis (Lorenzo-Díaz et al., 2012). Functional relevance of the IRs in the different steps of the conjugative process is currently under exploration. We hypothesize that IR1/3 may be involved in the Mob-recognition of the *oriT* at the initiation of the relaxosome formation (in the donor cell) and IR2 at the termination reaction to close the T-strand (in the recipient cell).

Based on the sequence and structure of $oriT_{pMV158}$, we inspected the *oriT* regions of the 93 MOB_V-plasmids listed in Table 1. A total of 97 sequences were manually identified upstream to their respective relaxase encoded genes (five of them exhibiting two different MOB_V-related *oriT*s: pSTE1, pKKS285, pSCFS1, unnamed (GenBank Acc. No GG692894.1), and pTB19. Analysis of the region showed

a high degree of sequence conservation in the majority of the *oriT*s, being composed by three IRs and exhibiting the consensus sequence 'GTGBG↓T' for the *nic* site (B denoting a G, C or T following IUPAC code; '↓' being the *nic* site). Two other minor plasmid groups, represented by pT181 and p1414, grouped in different clusters given their differences respect to the *oriT* sequence of pMV158 (Fig. 3 and Supplementary Fig. S2). However, these two clusters maintain the number and distribution of the IRs. Unfortunately, the experimental evidences for mapping the specific *nic* site in these plasmids are scarce and only a few *oriT* predictions, such as for plasmids like pSMQ172 and pPB1, have been published (de las Rivas et al., 2004; Turgeon and Moineau, 2001).

Given that the $oriT_{pMV158}$ is the only element required in *cis* for pMV158 to be transferred (Farías and Espinosa, 2000), it is plausible to assume that any RCR-plasmid containing an *oriT*-like sequence would be potentially mobilizable. Such is the case of pXY3 (Zhou et al., 2010) and pCI411 (Coffey et al., 1994) plasmids (not included in Table 1), which only seem to contain an orphan and non-canonical *oriT*. In fact, the prototype RCR-plasmid pC194, which lacks both a *mob* gene and a canonical *oriT*, has been mobilized by using the conjugative transposon *Tn*916 as a helper (Naglich and Andrews, 1988; Showsh and Andrews, 1999). It is assumed that the relaxase of the helper should be able to recognize a suitable *oriT* in the Δ*mob* plasmid. It is more intriguing the finding that pC194 and Δ*oriT*-Δ*mob*-derivatives of plasmids pUB110 and pTA1060 were efficiently mobilized by the mating apparatus of ICE*Bs1*, but without the intervention of the ICE*Bs1* NicK relaxase (Lee et al., 2012). Thus, in this case, the plasmids were not transferred by cross-recognition of an *oriT* by the relaxase of the helper. They were neither transferred by forming cointegrates with ICE*Bs1*. Unexpectedly, the RCR initiation proteins of the mobilized plasmids were crucial for transfer (Lee et al., 2012). Thus, the strategy we have followed to search for plasmids containing MOB_{V1} relaxases might underestimate the real population of mobile elements.

5.2. The MOB_{V1} relaxases

The second element that belongs to the transfer module of the pMV158 RCR-family is the *mobM* gene, which codifies the MobM relaxase (Guzmán and Espinosa, 1997). Thus, in addition to using the homology of the Rep proteins to define the pMV158 RCR-plasmid family (Supplementary Fig. S1 and Table S1), we decided to find out whether any relationship between the Rep- and the Mob-proteins of the plasmid family existed. The classification system for mobilizable plasmids based on the amino acid sequence of the relaxases allowed defining MobM from pMV158 as the prototype of the MOB_V superfamily (Francia et al., 2004; Garcillán-Barcia et al., 2011, 2009). This superfamily is composed of more than 200 relaxases, of which about 140 were located in plasmids and the rest in bacterial chromosomes (Guglielmini et al., 2011).

Five subfamilies were described, and MobM from pMV158 was taken as the prototype of the MOB_{V1} subfamily (Garcillán-Barcia et al., 2009). Members of this subfamily showed three conserved motifs: (i) Motif I (HxxR), of yet unknown function; (ii) Motif II (NY(D/E)L), which is the proposed catalytic domain, and (iii) Motif III (HxDE. . .PHxH), which corresponds to the metal-coordination motif, also known as the 3H motif (Chandler et al., 2013). The MOB_{V2} family, represented by the Mob protein of the theta-replicating plasmid pBBR1 (Szpirer et al., 2001), exhibited Motifs I and III but lacked Motif II. The three motifs are located in the N-terminal moiety of the MOB_V relaxases. This moiety harbors the DNA binding and nicking activities, since a truncated version of MobM, which only contains the first N-terminal 199 residues, retained the relaxase activity on supercoiled DNA (Fernández-López et al., 2013a,b; Lorenzo-Díaz et al., 2011). The C-terminal moiety of MobM could be involved in, at least, two functions: (i) protein–protein interactions (dimerization and interactions with the auxiliary plasmid-encoded coupling protein) and (ii) association with the cell membrane, through a proposed coiled-coil region located between residues 400 and the C-terminal end of MobM. Disruption of the alpha helical-rich region by mutations (changes to Pro residues) resulted in failure of MobM-association with membranes; further, the pMV158-derivative harboring these mutations lost its ability to be transferred (de Antonio et al., 2004). There was no indication of a helicase activity in the C-terminal moiety of MobM (our unpublished observations).

We have updated the inventory of elements that cluster with pMV158 into the MOB_{V1} subfamily. 97 non-redundant plasmid MOB_{V1} relaxases were retrieved from a PSI-BLAST search using the N-terminal 300 amino acids of MobM from pMV158 as a query (Table 1). All of them shared the three Motifs described above (Fig. 4A). The inferred phylogeny of MOB_{V1} relaxases, rooted by a MOB_{V2} relaxase, showed well-supported external clusters of highly related sequences and poorly-supported internal nodes (Fig. 4B). It is a fact that reflects the low overall similarity of the taxa, mainly circumscribed to the three relaxase motifs described (Francia et al., 2004; Garcillán-Barcia et al., 2009). The plasmids coding these relaxases are primarily distributed in several genera of Lactobacillales and Bacillales, with a few members out of the phylum Firmicutes. Curiously, the plasmids encoding relaxases comprised in a monophyletic clade in the phylogenetic tree shown in Fig. 4B are generally hosted in a single taxonomic order (either Lactobacillales or Bacillales), suggesting less inter-order transfer than expected.

No genes associated with conjugative transfer others than the relaxase ones were encoded by these plasmids. Thus, they are classified as mobilizable, requiring the conjugative machinery of other plasmids to be transferred (namely, a coupling protein and a mating-pair formation apparatus composed by a T4SS and a conjugative pilus). Eight mating-pair formation (MPF) types have been phylogenetically described (Guglielmini et al., 2011); three of them, MPF_{FATA}, MPF_{FA} and MPF_T were able to mobilize MOB_{V1} plasmids. Specifically, pMV158 was mobilized between G+ bacteria by functions supplied by helper MPF_{FATA} plasmids of the Inc18 family, like pIP501 and pAMβ1 (Grohmann et al., 1999; Priebe and Lacks, 1989; van der Lelie et al., 1990), and even to the G− bacterium *E. coli* by MPF_T plasmids RP4 and R388 but not by MPF_F

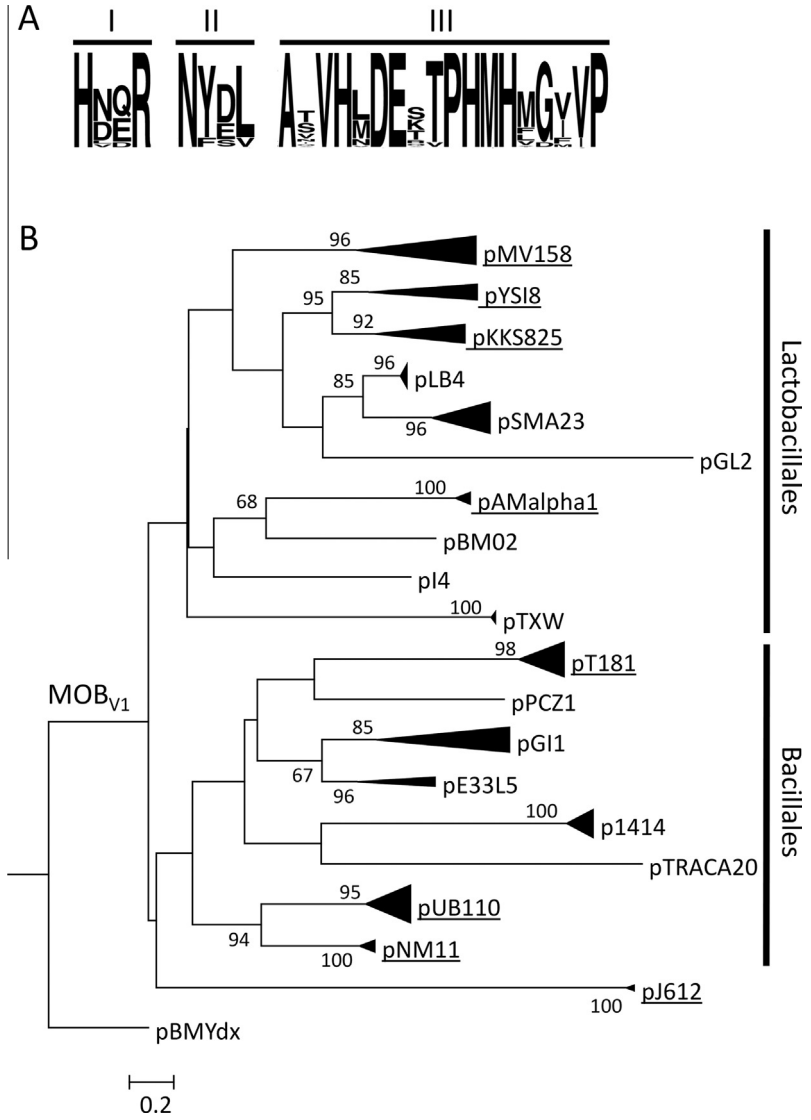

Fig. 4. Phylogeny of MobM$_{pMV158}$ homologs. (A) Logos of the MOB$_{V1}$ relaxase motifs. MOB$_{V1}$ relaxases were analyzed using WebLogo (version 2.8.2) (Crooks et al., 2004). (B) The 300 N-terminal residues of the MobM relaxase of plasmid pMV158 were used as query in a PSI-BLAST search (Altschul et al., 1997) (e-value: 1 × E−6 and limited to 100 non-redundant plasmid hits). The search converged in the sixth iteration. The 300 N-terminal residues of the homologs were aligned using MUSCLE (Edgar, 2004). The phylogenetic reconstruction was carried out by maximum likelihood (ML), using RAxML version 7.2.7 (Stamatakis, 2006). 100 ML trees were executed using the JTTGAMMA model. 1000 bootstrap trees were then inferred to obtain the confidence values for each node of the best ML tree. Only bootstrap values >50% are indicated. The MOB$_{V2}$ relaxase of plasmid pBMYdx (GenBank Acc. No. NP_981974.1) was used as outgroup. Highly related clusters are compressed and a prototype member is indicated. The names and features of all members included in the tree are recorded in Table 1. Clusters grouping plasmids that encode antibiotic resistance traits are underlined. Vertical bars delimit clades for which most of their members are hosted either in Lactobacillales or in Bacillales.

plasmid F (Farías and Espinosa, 2000). Other elements, such as pC194 and pUB110, were mobilized by auxiliary MPF$_{FA}$ elements such as *Tn*916 (Lee et al., 2012; Naglich and Andrews, 1988), and MPF$_{FATA}$ plasmids, like pLS20, mobilized pUB110 and pBC16 (Koehler and Thorne, 1987; Selinger et al., 1990).

Most of the MOB$_{V1}$ plasmids replicate by the rolling circle mechanism (Fig. 5 and Table 1). MOB$_{V1}$ relaxases are predominantly linked to RCR initiators of the three different subgroups: Rep_1 (PF01446), Rep_2 (PF01719), and Rep_trans (PF02486). A small fraction of MOB$_{V1}$ relaxases is linked to a wide variety of theta replication families, of which Rep_3 is the most abundant. Congruently, with the abovementioned taxonomic bias in the abundance of RCR initiators and the MOB$_{V1}$ relaxase distribution, most relaxases encoded in Rep_2 RCR plasmids grouped separately from those encoded in Rep_1 RCR plasmids (clade Lactobacillales vs. Bacillales in Fig. 6 and Supplementary Table S2). An example of this bias is found in Enterococci (Lactobacillales), where Rep_1 initiators are commonly found in multireplicon plasmids, and may not be functional, as it is the case for plasmid pAMα1 (Clewell et al., 2014).

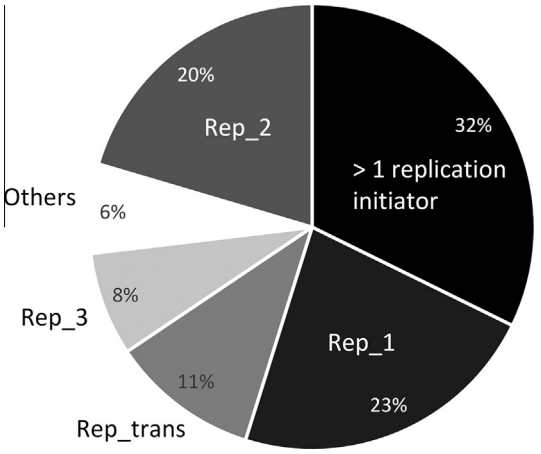

Fig. 5. Distribution of replication initiation protein families in MOB_{V1} plasmids. The percentage of replication initiator families included in Table 1 is presented. Plasmids with more than one initiator are included in "≥1 replication initiator". Plasmids with a single initiator are grouped in "Rep_1", "Rep_2" and "Rep_trans" families (when RCR), or in "Rep_3" and "Others" (non-RCR).

Nevertheless, a few functional Rep_1 initiators can be also found in the Lactobacillales clade, such as those grouped in the cartooned pSMA23 cluster of Fig. 4B. Curiously, within the pSMA23 group some exceptions also exist: plasmid pCD034-2 encodes a Rep_2 instead of a Rep_1 initiator, an indication of recent recombination events that led to new backbone combinations. Besides, more ancient recombination events could give rise to arrangements present in plasmids of the Bacillales MOB_{V1} clade, since the highly divergent relaxases included in that group are linked to initiators of the three RCR subgroups (Fig. 6A).

Fig. 6 also provides an interesting example of the switch of plasmids from and to integrative elements. ICE*Sgal1* is an integrative element highly similar to plasmid pUB110 (both in relaxase and initiator) but located in the chromosome of *Streptococcus gallolyticus* UCN34 (Lactobacillales order). It is tempting to speculate that this element integrated in the chromosome once transferred from Bacillales to the Lactobacillales background (maybe helped by a Tn*916*-like element located close to ICE*Sgal1*), where it was not able to replicate and became an integrative and mobilizable element (IME).

A significant proportion (one third) of MOB_{V1} plasmids recorded in Table 1 encoded more than one replication initiation protein (Fig. 5). Besides, 10 out of the 93 plasmids coded more than one relaxase gene. Both facts suggest the frequent arising of plasmid cointegrates. Precisely, it is known the ability of RCR-plasmid encoded relaxases to promote site-specific DNA recombination at *oriT* rendering plasmid cointegrates during conjugative mobilization, where the host *rec* system may function to stimulate the recombination process (Gennaro et al., 1987; Novick et al., 1984; Projan and Novick, 1988). Further, it is worth recalling that the staphylococcal plasmid pE194 was able to integrate into the *B. subtilis* chromosome by a RecA-independent recombination mechanism, and using as little as 6–14 bp homologies (Dempsey and Dubnau, 1989). This, in conjunction with the finding of RCR-plasmids integrated into other bigger plasmids (Oskam et al., 1991), raises the question of whether RCR-plasmids that are integrated on IMEs would participate in their replication and mobilization; in fact, we have found a putative initiator of replication of the Rep_1 family within an streptococcal island which also encode a MOB_{V1} protein (our unpublished observations). Besides, the cross-recognition of heterologous *oriT*s by MOB_V relaxases (Fernández-López et al., 2013a) could be favored in the multi-relaxase cointegrates, potentiating their spreading.

Recombination seems to be the most likely cause of the different topologies exhibited in the dendrograms of initiators and relaxases (compare trees in panels A and B of Fig. 6). Despite the absence of a strong coevolution between initiators and relaxases, there is a clear tendency to find stable backbones Rep_2-MOB_{V1} and Rep_1-MOB_{V1} differentially adapted to different taxonomic orders.

6. The single-strand origins (*sso*)

Rolling circle replication in ssDNA coliphages and in plasmids requires the existence of origins that are involved in lagging strand synthesis (Kornberg and Baker, 1992; Novick, 1998). When RCR-plasmids were discovered (te Riele et al., 1986a,b), it was apparent that the ssDNA intermediates should harbor signals for the conversion of the ssDNA to plasmid dsDNA. This was first reported for plasmid pT181 (Gruss et al., 1987) and soon afterwards for the pMV158-derivative plasmid pLS1 (del Solar et al., 1987). In these two plasmids, it was shown that their *sso*As were located in non-coding 200–300 base-pair-long regions that have the potential to generate one or several secondary structures on ssDNA. These signals were orientation-dependent (this means that they were functional only when placed in the displaced strand). Deletion of the region encompassing the *ssoA* led to accumulation of ssDNA intermediates and to plasmid instability, but the plasmids were still able to replicate (del Solar et al., 1987; Dempsey et al., 1995; Gruss et al., 1987; Kramer et al., 1995; Murray et al., 1989; Seegers et al., 1995). This finding led to the hypothesis that alternative, albeit less efficient, *sso*s could replace the genuine conversion signal (del Solar et al., 1987; Kramer et al., 1998a; Meijer et al., 1995a,b). Although accumulation of ssDNA intermediates and plasmid unstable inheritance by VGT were thought to be related phenomena (Meijer et al., 1995a,b), cloning of signals that allowed an efficient ssDNA → dsDNA conversion, at least in plasmid pLS1, did not lead to stable plasmid inheritance (Hernández-Arriaga et al., 2000). Apart from *ssoA*, other three *sso* types have been described based on sequence similarity and structure analysis (Khan, 1996, 2000; Kramer et al., 1998a): *ssoU* (pUB110, pMV158), *ssoT* (pTA1060, pBAA1), and *ssoW* (pWV01). A fifth type was described in plasmid pM4 from *Lactobacillus plantarum*, which exhibited no significant sequence or structural similarity with any of the four classical *sso* (Yin et al., 2009). Plasmids pMV158 and pRW35 have the unusual feature of harboring two *sso*, *ssoU* and *ssoA*. Out of the

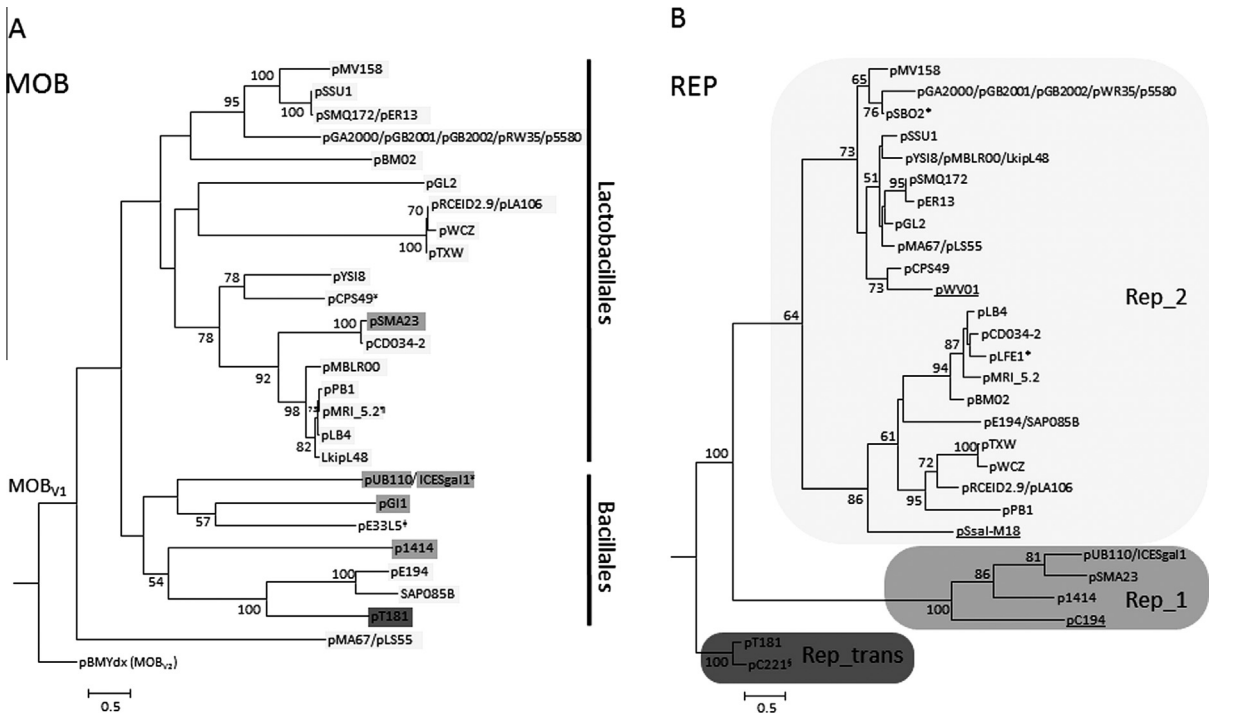

Fig. 6. Phylogeny of the Rep and Mob (MOB_{V1}) proteins of relevant RCR-plasmids. Representative RCR plasmids and MOB_{V1} elements included in Supplementary Table S2 were used to trace the evolutionary relationships of their relaxase and replication initiation proteins. The homologs were aligned using MUSCLE (Edgar, 2004). The phylogenetic reconstruction was carried out by maximum likelihood (ML), using RAxML version 7.2.7 (Stamatakis, 2006). 100 ML trees were executed using the JTTGAMMA model. 1000 bootstrap trees were then inferred to obtain the confidence values for each node of the best ML tree. Only bootstrap values >50% are indicated. (A) Phylogenetic tree of the N-terminal 300 residues of MOB_{V1} relaxases. Each plasmid is shadowed in gray according to the RCR initiator subgroup as indicated in the legend of panel B. [#]Indicates an exceptional MOB_{V1} plasmid, pE33L5, which does not encode an RCR initiator but a HTH_36 (PF13730) replication initiation protein. Vertical bars delimit clades for which most of their members are hosted either in Lactobacillales or in Bacillales. [¥]Indicates an element not hosted in the taxonomic order indicated by the bars. [¶]Indicates that plasmid pMRI_5.2 also encodes a Rep_1 RCR initiator. (B) Phylogenetic tree of the RCR initiators. Plasmids pT181 and pC221 were used as outgroups. A gray color palette was used to indicate clades containing different RCR initiators families: Rep_1 (PF01446) and Rep_2 (PF01719), as well as the Rep_trans (PF02486) used to root the tree. *According to their GenBank annotated sequences, plasmids pSBO2 and pLFE1 encode truncated MOB_V relaxases and thus were not included in the MOB phylogeny, neither were the underlined plasmids (pWV01, pSsal-M18 and pC194) since they do not encode relaxases. [§]pC221 is a mobilizable RCR plasmid, but it encodes a MOB_{P7} instead of a MOB_V relaxase.

two origins, the former has a more complex structure than the latter (Fig. 7).

Two regions were mapped in the *ssoA*: the recombination site B (RS_B) and a 6-nucleotide consensus sequence (CS-6) (del Solar et al., 1987; Gruss et al., 1987). These two regions acted as efficient signals only in ssDNA configuration (replicative intermediates). In these molecules, the *ssoA* would adopt a long stem-loop structure where the RS_B would be located at the stem and the CS-6 in the loop of the hairpin (Fig. 7). Biochemical and genetic analyses demonstrated that RS_B was recognized as the binding site of the host RNA polymerase (RNAP), thus acting as ssDNA promoter (Kramer et al., 1997). These promoters were shown to generate within paired secondary structures on ssDNA molecules and harbor sequences that are recognized by the host RNAP to initiate lagging strand synthesis (Glucksmann-Kuis et al., 1992; Masai and Arai, 1997). From the $ssoA_{pMV158}$ promoter, the RNAP synthesized a short 20 nt-long primer RNA (pRNA) that stopped at the CS-6 sequence, which acted as a transcription terminator (Kramer et al., 1997). The pRNA was processed by DNA polymerase I and was proposed to be elongated by the host DNA polymerase III to finish the lagging DNA strand synthesis (Kramer et al., 1998b).

Efficient ssDNA → dsDNA conversion is also needed in the conjugative process. Within the recipient cell the transferred ssDNA should be converted into dsDNA either prior or after the circularization of the T-DNA by a strand-transfer reaction. During transfer, a fast conversion of ssDNA intermediates would be critical to finish the process, since no proteins essential for vegetative replication and control would be synthesized in the recipient cell until the first dsDNA plasmid copy is generated (Lorenzo-Díaz and Espinosa, 2009). Despite its importance, the synthesis of the lagging strand in the recipient cell remains an unresolved issue. For plasmids F and ColIb-P9 single-stranded promoters were identified in the leading strand region, which is the one that enters first into the recipient cell, and that includes genes that promote the establishment of the incoming plasmid (Bates et al., 1999; Masai and Arai, 1997; Nasim et al., 2004).

We have explored the *sso* diversity and distribution among MOB_{V1} plasmids. Based on the *ssos* previously identified by homology or characterized by *in vivo* and/or

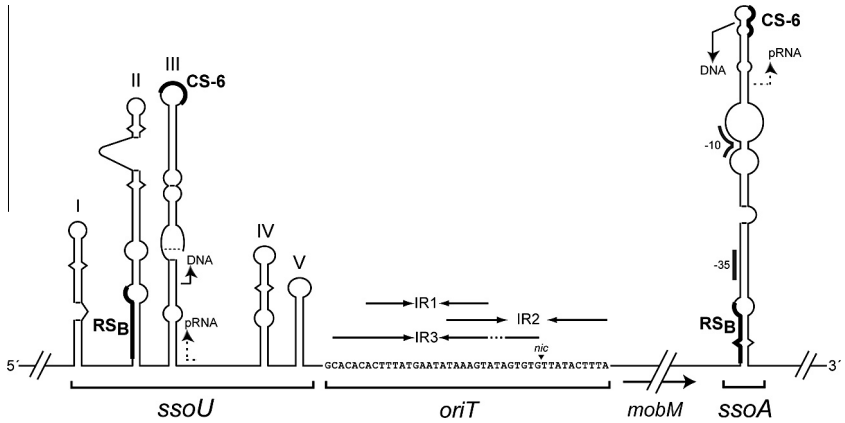

Fig. 7. Predicted secondary structures of the lagging-strand origins of replication *ssoA* and *ssoU*. The *oriT* and the *mobM* gene of plasmid pMV158 are flanked by two lagging-strand origins of replication (*ssoA* and *ssoU*). $oriT_{pMV158}$ sequence (coordinates 3564–3605 from pMV158; GenBank Acc. No. NC_010096) is shown at the center of the image. Its three inverted repeats are represented by arrows and the *nic* site by a vertical arrowhead. Both *sso*s can generate long hairpin-loop structures that function as 'ssDNA promoters' (Kramer et al., 1999, 1997; Masai and Arai, 1997). The RNAP-binding site (RS_B), located in the base of the hairpin is recognized by the RNAP to synthesize a short pRNA. A consensus sequence (CS-6), located in the loop of the hairpin, acts as the termination point for the pRNA synthesis. The pRNA is then used by DNA polymerase I for limited extension synthesis, followed by replication of the lagging strand by DNA Pol III. The figure was modified with permission from the American Society for Microbiology from (Fernández-López et al., 2014). No further reproduction or distribution is permitted without the prior written permission of American Society for Microbiology.

in vitro approaches, we annotated 35 *sso* elements in 33 out of 93 plasmids in Table 1. This limited number is due to the little overall homology among *sso*s, even those of the same type. All the reported *sso* types were found: *ssoT* (13 plasmids), *ssoA* (11 plasmids), *ssoU* (6 plasmids) and *ssoW* (4 plasmids). Two plasmids (pMV158 and pRW35) contained two *sso*s (*ssoA* and *ssoU*). No new members arose in our search of the new type of *sso* reported for plasmid pM4 (Zhai et al., 2009). Curiously enough, we have observed that while *ssoA* and *ssoW* are generally found downstream and close to the *mob* gene, the *ssoU* is always upstream with respect to the *oriT* sequence. In the case of the *ssoT* element we have found that it can be located upstream ($n = 5$) or downstream ($n = 8$) of the MOB module.

It has been demonstrated that the *sso* type is one of the key elements in determining the host range of a plasmid: whereas *ssoU* and *ssoT* support a wide-host range, *ssoA* and *ssoW* seem to evolve in the other direction, working efficiently only in their natural hosts *in vivo* (Khan, 2005). From the analysis abovementioned, it seems apparent that highly related plasmids do not always contain the same type of *sso*, whereas distantly related do. These facts complicate even more the picture of the putative host range a MOB_{V1} plasmid can reach. Whether the location of the *sso* affects the efficiency of a given plasmid for replication and/or conjugative transfer and, consequently, its host range has not been explored yet.

7. Conclusions and perspectives

RCR-plasmids represent a pool of genetic information that is shared by many bacteria. We have found them in Lactobacillales, Bacillales, Mollicutes, etc. Among the bacteria that host them, we have found several of the most relevant G+ pathogens, namely *Staphylococcus aureus*, *S. pneumoniae*, and *Enterococcus faecalis*. These three bacterial species are important for human health because they: (i) exhibit very high rates of human morbidity and mortality; (ii) are the cause of many health-care associated infections; (iii) have acquired elevated resistance to antibiotics; (iv) play an important role as reservoir of AbR and virulence genes, and (v) carry conjugative or mobilizable broad host-range plasmids, thus contributing to the spread of resistances. Infections caused by the above G+ bacteria represent, in addition to a high economic impact, a threat to hospitals, children, elder population and immunocompromised people (Jones, 2001). Furthermore, plasmid-encoded genes related not only to AbR but also to replication and/or transfer, are also found within the ICEs. However, and in spite of the relevance of these bacteria, very little is known on the genetics and biochemistry of the transfer functions of the RCR-plasmids studied here with the exception of the streptococcal plasmid pMV158 (reviewed in (Espinosa, 2013; Fernández-López et al., 2014) and the staphylococcal plasmid pC221 (Caryl et al., 2004; Caryl and Thomas, 2006; Smith and Thomas, 2004). And even in these two plasmids there is little, if any, information on the interactions of the relaxases encoded by pMV158 or pC221 with the machinery provided by auxiliary plasmids (Arends et al., 2013; Goessweiner-Mohr et al., 2013; Grohmann et al., 2003).

In addition to the above, it would be interesting to explore how the conjugative transfer is regulated in plasmids that contain more than one MOB_V cassette as well as the level of cross-recognition between the MobM-like relaxases and their non-cognate *oriT*s. We have shown that MobM is able to relax supercoiled DNAs from plasmids with *oriT*s that share total (plasmid pUB110) or partial (plasmid pDL287) homology with the *oriT* of pMV158 (Fernández-López et al., 2013a), a phenomenon that could play an important role in the plasmid spreading between bacteria in natural environments. Furthermore, the fact

that RCR initiators relax DNA in a way highly similar to conjugative relaxases and the recent finding of RCR initiators involvement in plasmid mobilization (Lee et al., 2012) open a new research field in the conjugative transmission of RCR plasmids.

Acknowledgments

Thanks are due to members of our labs for information and fruitful discussions. We regret that we have been unable to include all the contributions of people working in our field due to space limitations, and apologize for missing references. Research was funded by the Spanish Ministry of Economy and Competitiveness (grants CSD-2008-00013-INTERMODS to M.E., Sara Borrell CD13/00304 to F.L.-D., and BFU2011-26608 to M.P.G.-B.), and by the 7*th* Framework Programme (FP7-REGPOT-2012-CT2012-31637-IMBRAIN to F.L.-D. and by 282004/FP7-HEALTH-2011-2.3.1-2 to M.P.G.-B.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.plasmid.2014.05.004.

References

Alonso, J.C. et al., 1988. Functional analysis of the leading strand replication origin of plasmid pUB110 in *Bacillus subtilis*. Nucleic Acids Res. 16, 9127–9145.

Alonso, J.C., Tomalsky, M., 2014. Plasmids: biology and impact in biotechnology and discovery. In: A. Press, (Ed.).

Altschul, S.F. et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Anand, S.P. et al., 2005. The PcrA3 mutant binds DNA and interacts with the RepC initiator protein of plasmid pT181 but is defective in its DNA helicase and unwinding activities. Plasmid 54, 104–113.

Anand, S.P., Khan, S.A., 2004. Structure-specific DNA binding and bipolar helicase activities of PcrA. Nucleic Acids Res. 32, 3190–3197.

Anderson, M.T., Seifert, H.S., 2011. Opportunity and means: horizontal gene transfer from the human host to a bacterial pathogen. MBio 2, e00005–e00011.

Arends, K. et al., 2013. TraG encoded by the pIP501 type IV secretion system is a two-domain peptidoglycan-degrading enzyme essential for conjugative transfer. J. Bacteriol. 195, 4436–4444.

Ballester, S. et al., 1990. Comparative expression of the pC194-*cat* gene in *Streptococcus pneumoniae*, *Bacillus subtilis* and *Escherichia coli*. Gene 86, 71–79.

Bannam, T.L., Rood, J.I., 1991. Relationship between the *Clostridium perfringens catQ* gene product and chloramphenicol acetyltransferases from other bacteria. Antimicrob. Agents Chemother. 35, 471–476.

Baquero, F., 2004. From pieces to patterns: evolutionary engineering in bacterial pathogens. Nat. Rev. Microbiol. 2, 510–518.

Baquero, F., 2009. Environmental stress and evolvability in microbial systems. Clin. Microbiol. Infect. 15, 5–10.

Baron, C. et al., 2002. Bacterial secrets of secretion: EuroConference on the biology of type IV secretion processes. Mol. Microbiol. 43, 1359–1365.

Bates, S. et al., 1999. Expression of leading region genes on Incl 1 plasmid ColIb-P9: genetic evidence for single-stranded DNA transcription. Microbiology 145, 2665–2672.

Bergemann, A.D. et al., 1989. Homology of mycoplasma plasmid pADB201 and staphylococcal plasmid pE194. J. Bacteriol. 171, 593–595.

Bhatty, M. et al., 2013. The expanding bacterial type IV secretion lexicon. Res. Microbiol. 164, 620–639.

Breton, M. et al., 2012. Distribution and diversity of mycoplasma plasmids: lessons from cryptic genetic elements. BMC Microbiol. 12, 257.

Burdett, V., 1980. Identification of tetracycline-resistant R-plasmids in *Streptococcus agalactiae* (group B). Antimicrob. Agents Chemother. 18, 753–760.

Caryl, J.A. et al., 2004. Reconstitution of a staphylococcal plasmid–protein relaxation complex in vitro. J. Bacteriol. 186, 3374–3383.

Caryl, J.A., Thomas, C.D., 2006. Investigating the basis of substrate recognition in the pC221 relaxosome. Mol. Microbiol. 60, 1302–1318.

Cascales, E. et al., 2013. DNA substrate-induced activation of the *Agrobacterium* VirB/VirD4 type IV secretion system. J. Bacteriol. 195, 2691–2704.

Clewell, D.B. et al., 2014. Extrachromosomal and mobile elements in enterococci: transmission, maintenance, and epidemiology. In: Gilmore, M.S., Clewell, D.B., Ike, Y, et al. (Eds.), Enterococci: From Commensals to Leading Causes of Drug Resistant Infection. Massachusetts Eye and Ear Infirmary, Boston.

Coffey, A. et al., 1994. Nucleotide sequence and structural organization of the small broad-host-range plasmid pCI411 from *Leuconostoc lactis*. Microbiology 140, 2263–2269.

Crooks, G.E. et al., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.

Chandler, M. et al., 2013. Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. Nat. Rev. Microbiol. 11, 625–638.

Chang, T.-L. et al., 2000. Role of individual monomers of a dimeric initiator protein in the initiation and termination of plasmid rolling circle replication. J. Biol. Chem. 275, 13529–13534.

Chen, I. et al., 2005. The ins and outs of DNA transfer in bacteria. Science 310, 1456–1460.

de Antonio, C. et al., 2004. Features of the plasmid pMV158-encoded MobM, a protein involved in its mobilization. J. Mol. Biol. 335, 733–743.

de la Campa, A.G. et al., 1990. Initiation of replication of plasmid pLS1. The initiator protein RepB acts on two distant DNA regions. J. Mol. Biol. 213, 247–262.

de la Cruz, F. et al., 2010. Conjugative DNA metabolism in gram-negative bacteria. FEMS Microbiol. Rev. 34, 18–40.

de las Rivas, B. et al., 2004. Complete nucleotide sequence and structural organization of pPB1, a small *Lactobacillus plantarum* cryptic plasmid that originated by modular exchange. Plasmid 52, 203–211.

de Paz, H.D. et al., 2005. Functional interactions between type IV secretion systems involved in DNA transfer and virulence. Microbiology 151, 3505–3516.

del Solar, G., Espinosa, M., 2000. Plasmid copy number control: an ever-growing story. Mol. Microbiol. 37, 492–500.

del Solar, G. et al., 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. 62, 434–464.

del Solar, G. et al., 2002. A genetically economical family of plasmid-encoded transcriptional repressors in control of plasmid copy number. J. Bacteriol. 184, 4943–4951.

del Solar, G. et al., 1993. Rolling circle-replicating plasmids from gram-positive and gram-negative bacteria: a wall falls. Mol. Microbiol. 8, 789–796.

del Solar, G. et al., 1987. Initiation signals for the conversion of single stranded to double stranded DNA forms in the streptococcal plasmid pLS1. Nucleic Acids Res. 15, 5561–5580.

Dempsey, L.A., Dubnau, D.A., 1989. Identification of plasmid and *Bacillus subtilis* chromosomal recombination sites used for pE194 integration. J. Bacteriol. 171, 2856–2865.

Dempsey, L.A. et al., 1995. Localization of the start sites of lagging-strand replication of rolling circle plasmids from gram-positive bacteria. Mol. Microbiol. 15, 679–687.

Díaz, A. et al., 1994. Multiple roles for DNA polymerase I in establishment and replication of the promiscuous plasmid pLS1. Mol. Microbiol. 14, 773–783.

Draper, O. et al., 2005. Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. Proc. Natl. Acad. Sci. U.S.A. 102, 16385–16390.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Espinosa, M., 2013. Plasmids as models to study macromolecular interactions: the pMV158 paradigm. Res. Microbiol. 164, 199–204.

Espinosa, M. et al., 1995. Plasmid rolling circle replication and its control. FEMS Microbiol. Lett. 130, 111–120.

Farías, M.E., Espinosa, M., 2000. Conjugal transfer of plasmid pMV158: uncoupling of the pMV158 origin of transfer from the mobilization gene *mobM*, and modulation of pMV158 transfer in *Escherichia coli* mediated by IncP plasmids. Microbiology 146, 2259–2265.

Fernández-López, C. et al., 2014. Mobilizable rolling-circle replicating plasmids from Gram-positive bacteria: a low-cost conjugative transfer. In: Alonso, J.C., Tomalsky, M. (Eds.). vol. ASM Press, in press.

Fernández-López, C. et al., 2013a. Nicking activity of the pMV158 MobM relaxase on cognate and heterologous origins of transfer. Plasmid, 120–130.
Fernández-López, C. et al., 2013b. Functional properties and structural requirements of the plasmid pMV158-encoded MobM relaxase domain. J. Bacteriol. 195, 3000–3008.
Francia, M.V. et al., 2004. A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol. Rev. 28, 79–100.
Frost, L. et al., 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. 3, 722–732.
Funnell, B.E., Phillips, G.J. (Eds.), 2004. Plasmid Biology. ASM Press, Washington, DC.
Garcillán-Barcia, M.P. et al., 2011. Identification of bacterial plasmids based on mobility and plasmid population biology. FEMS Microbiol. Rev. 35, 936–956.
Garcillán-Barcia, M.P. et al., 2009. The diversity of conjugative relaxases and its application in plasmid classification. FEMS Microbiol. Rev. 33, 657–687.
Garcillán-Barcia, M.P. et al., 2007. Conjugative transfer can be inhibited by blocking relaxase activity within recipient cells with intrabodies. Mol. Microbiol. 63, 404–416.
Gennaro, M.L. et al., 1987. A site-specific recombination function in *Staphylococcus aureus* plasmids. J. Bacteriol. 169, 2601–2610.
Glucksmann-Kuis, M.A. et al., 1992. Specific sequences and a hairpin structure in the template strand are required for N4 virion RNA polymerase promoter recognition. Cell 70, 491–500.
Goessweiner-Mohr, N. et al., 2013. Conjugative type IV secretion systems in gram-positive bacteria. Plasmid 70, 289–302.
Gomis-Rüth, F.X., Coll, M., 2006. Cut and move: protein machinery for DNA processing in bacterial conjugation. Curr. Opin. Struct. Biol. 16, 744–752.
Gomis-Ruth, F.X. et al., 2004. Coupling factors in macromolecular type-IV secretion machineries. Curr. Pharm. Des. 10, 1551–1565.
Goursot, R. et al., 1982. Plasmids from *Staphylococcus aureus* replicate in yeast *Saccharomyces cerevisiae*. Nature 298, 488–490.
Grohmann, E. et al., 1999. Mobilisation of the streptococcal plasmid pMV158: interactions of MobM protein with its cognate *oriT* DNA region. Mol. Gen. Genet. 261, 707–715.
Grohmann, E. et al., 2003. Conjugative plasmid transfer in gram-positive bacteria. Microbiol. Mol. Biol. Rev. 67, 277–301.
Grohmann, E. et al., 1997. Determination of specific DNA strand discontinuities with nucleotide resolution in exponentially growing bacteria harbouring rolling circle-replicating plasmids. FEMS Microbiol. Lett. 152, 363–369.
Gruss, A.D., Ehrlich, S.D., 1989. The family of highly interrelated single-stranded deoxyribonucleic avid plasmids. Microbiol. Rev. 53, 231–241.
Gruss, A.D. et al., 1987. Functional analysis of a palindromic sequence required for normal replication of several staphylococcal plasmids. Proc. Natl. Acad. Sci. U.S.A. 84, 2165–2169.
Guasch, A. et al., 2003. Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. Nat. Struct. Biol. 10, 1002–1010.
Guglielmini, J. et al., 2011. The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. PLoS Genet. 7, e1002222.
Guzmán, L., Espinosa, M., 1997. The mobilization protein, MobM, of the streptococcal plasmid pMV158 specifically cleaves supercoiled DNA at the plasmid *oriT*. J. Mol. Biol. 266, 688–702.
Halbedel, S., Stülke, J., 2007. Tools for the genetic analysis of *Mycoplasma*. Int. J. Med. Microbiol. 297, 37–44.
Hamilton, C.M. et al., 2000. TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. J. Bacteriol. 182, 1541–1574.
Harwood, C.R. et al., 1983. Nucleotide sequence of a *Bacillus pumilus* gene specifying chloramphenicol acetyltransferase. Gene 24, 163–169.
Hernández-Arriaga, A.M. et al., 2000. A functional lagging strand origin does not stabilize plasmid pMV158 inheritance in *Escherichia coli*. Plasmid 43, 49–58.
Horinouchi, S., Weisblum, B., 1982. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bacteriol., 815–825.
Inamine, J.M. et al., 1990. Evidence that UGA is read as a tryptophan codon rather than as a stop codon by *Mycoplasma pneumoniae*, *Mycoplasma genitalium*, and *Mycoplasma gallisepticum*. J. Bacteriol. 172, 504–506.
Jin, R., Novick, R.P., 2001. Role of the double-strand origin cruciform in pT181 replication. Plasmid 46, 95–105.
Jones, R.N., 2001. Resistance patterns among nosocomial pathogens: trends over the past few years. Chest 119 (2 Suppl.), 397S–404S.
Kent, B.N. et al., 2012. Development of a novel plasmid as a shuttle vector for heterologous gene expression in *Mycoplasma yeatsii*. J. Microbiol. Methods 91, 121–127.
Khan, S.A., 1996. Mechanism of replication and copy number control of plasmids in gram-positive bacteria. In: Setlow, J.K. (Ed.), Genetic Engineering, vol. 18. Plenum Press, New York.
Khan, S.A., 1997. Rolling-circle replication of bacterial plasmids. Microbiol. Mol. Biol. Rev. 61, 442–455.
Khan, S.A., 2000. Plasmid rolling-circle replication: recent developments. Mol. Microbiol. 37, 477–484.
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.
Khan, S.A., 2005. Plasmid rolling-circle replication: highlights of two decades of research. Plasmid 53, 126–136.
Khan, S.A. et al., 1981. Replication of plasmid pT181 in vitro: requirement for a plasmid-encoded product. Proc. Natl. Acad. Sci. U.S.A. 78, 4902–4906.
King, K.W., Dybvig, K., 1992. Nucleotide sequence of *Mycoplasma mycoides* subspecies *mycoides* plasmid pKMK1. Plasmid, 86–91.
Koehler, T.M., Thorne, C.B., 1987. *Bacillus subtilis* (natto) plasmid pLS20 mediates interspecies plasmid transfer. J. Bacteriol. 169, 5271–5278.
Koepsel, R.R. et al., 1986. Sequence-specific interaction between the replication initiator protein of plasmid pT181 and its origin of replication. Proc. Natl. Acad. Sci. U.S.A. 83, 5484–5488.
Koepsel, R.R. et al., 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.
Kornberg, A., Baker, T., 1992. DNA Replication. Freeman and Co., New York.
Kramer, M.G. et al., 1995. Lagging-strand origins of the promiscuous plasmid pMV158: physical and functional characterization. Microbiology 141, 655–662.
Kramer, M.G. et al., 1998a. Lagging strand replication of rolling-circle plasmids: specific recognition of the *ssoA*-type origins in different gram-positive bacteria. Proc. Natl. Acad. Sci. U.S.A. 95, 10505–10510.
Kramer, M.G. et al., 1999. Characterization of a single-strand origin, *ssoU*, required for broad host range replication of rolling-circle plasmids. Mol. Microbiol. 33, 466–475.
Kramer, M.G. et al., 1997. Plasmid rolling circle replication: identification of the RNA polymerase-directed primer RNA and requirement of DNA polymerase I for lagging strand initiation. EMBO J. 16, 5784–5795.
Kramer, M.G. et al., 1998b. Lagging-strand replication from the *ssoA* origin of plasmid pMV158 in *Streptococcus pneumoniae*: *in vivo* and *in vitro* influences of mutations in two conserved *ssoA* regions. J. Bacteriol. 180, 83–89.
Lacks, S.A. et al., 1986. Identification and analysis of genes for tetracycline resistance and replication functions in the broad-host-range plasmid pLS1. J. Mol. Biol. 192, 753–765.
Lanka, E., Wilkins, B.M., 1995. DNA processing reactions in bacterial conjugation. Annu. Rev. Biochem. 64, 141–169.
Lederberg, J., Tatum, E., 1946. Gene recombination in *E. coli*. Nature 158, 558.
Lederberg, J., Tatum, E., 1953. Sex in bacteria; genetic studies, 1945–1952. Science 118, 169–175.
Lee, C.A. et al., 2012. The *Bacillus subtilis* conjugative transposon ICEBs1 mobilizes plasmids lacking dedicated mobilization functions. J. Bacteriol. 194, 3165–3172.
Lorenzo-Díaz, F. et al., 2011. The MobM-relaxase domain of plasmid pMV158: thermal stability and activity upon Mn^{2+} and DNA specific-binding. Nucleic Acids Res. 39, 4315–4329.
Lorenzo-Díaz, F., Espinosa, M., 2009. Lagging strand DNA replication origins are required for conjugal transfer of the promiscuous plasmid pMV158. J. Bacteriol. 191, 720–727.
Lorenzo-Díaz, F. et al., 2012. Autoregulation of the synthesis of the MobM relaxase encoded by the promiscuous plasmid pMV158. J. Bacteriol. 194, 1789–1799.
Lucas, M. et al., 2010. Relaxase DNA binding and cleavage are two distinguishable steps in conjugative DNA processing that involve different sequence elements of the *nic* site. J. Biol. Chem. 285, 8918–8926.
Llosa, M. et al., 2002. Bacterial conjugation: a two-step mechanism for DNA transport. Mol. Microbiol. 45, 1–8.
Llosa, M. et al., 2009. Bacterial type IV secretion systems in human disease. Mol. Microbiol. 73, 141–151.
Machón, C. et al., 2010. RepD-mediated recruitment of PcrA helicase at the *Staphylococcus aureus* pC221 plasmid replication origin, oriD. Nucleic Acids Res. 38, 1874–1888.
Masai, H., Arai, K., 1997. Frpo: a novel single-stranded DNA promoter for transcription and for primer RNA synthesis of DNA replication. Cell 89, 897–907.

Matilla, I. et al., 2010. The conjugative DNA translocase TrwB is a structure-specific DNA-binding protein. J. Biol. Chem. 285, 17537–17544.
McKenzie, T. et al., 1986. The nucleotide sequence of pUB110: some salient features in relation to replication and its regulation. Plasmid 15, 93–103.
Meijer, W.J.J. et al., 1995a. Effects of the generation of single-stranded DNA on the maintenance of plasmid pMV158 and derivatives in different *Bacillus subtilis* strains. Plasmid 33, 79–89.
Meijer, W.J.J. et al., 1995b. Effects of the generation of single-stranded DNA on the maintenance of plasmid pMV158 and derivatives in *Lactococcus lactis*. Plasmid 33, 91–99.
Moscoso, M. et al., 1997. Initiation of replication of plasmid pMV158: mechanisms of DNA strand transfer reactions mediated by the initiator RepB protein. J. Mol. Biol. 268, 840–856.
Murray, R.W., Koepsel, R.R., Khan, S.A., 1989. Synthesis of single-stranded plasmid pT181 DNA *in vitro*. J. Biol. Chem. 264, 1051–1057.
Naglich, J.G., Andrews, R.E.J., 1988. *Tn916*-dependent conjugal transfer of pC194 and pUB110 from *Bacillus subtilis* into *Bacillus thuringiensis* subsp. *israelensis*. Plasmid 20, 113–126.
Nasim, M.T. et al., 2004. The activity of a single-stranded promoter of plasmid ColIb-P9 depends on its secondary structure. Mol. Microbiol. 53, 405–417.
Noirot, P. et al., 1990. Initiation of rolling circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin. Proc. Natl. Acad. Sci. U.S.A. 87, 8560–8564.
Nordström, K. et al., 1984. Control of replication of bacterial plasmids: genetics, molecular biology, and physiology of the plasmid R1 system. Plasmid 12, 71–90.
Novick, R.P., 1989. Staphylococcal plasmids and their replication. Annu. Rev. Microbiol. 43, 537–565.
Novick, R.P., 1998. Contrasting lifestyles of rolling-circle phages and plasmids. Trends Biochem. Sci. 23, 434–438.
Novick, R.P. et al., 1984. Staphylococcal plasmid cointegrates are formed by host- and phage-mediated general *rec* systems that act on short regions of homology. Mol. Gen. Genet. 195, 374–377.
Ochman, H. et al., 2000. Lateral gene transfer and the nature of bacterial innovation. Nature 405, 229–304.
Osborn, M. et al., 2000. The evolution of bacterial plasmids. In: Thomas, C.M. (Ed.), The Horizontal Gene Pool. Harwood Academic Publishers, Amsterdam.
Oskam, L. et al., 1991. The large *Bacillus* plasmid pTB19 contains two integrated rolling-circle plasmids carrying mobilization functions. Plasmid 26, 30–39.
Pansegrau, W., Lanka, E., 1996a. Enzymology of DNA strand transfer by conjugative mechanisms. Prog. Nucleic Acid Res. Mol. Biol. 54, 197–251.
Pansegrau, W., Lanka, E., 1996b. Mechanisms of initiation and termination reactions in conjugative DNA processing: independence of tight substrate binding and catalytic activity of relaxase (TraI) of IncPα plasmid RP4. J. Biol. Chem. 271, 13068–13076.
Paulson, J., Ehrenberg, M., 1998. Trade-off between segregational stability and metabolic burden: a mathematical model of plasmid ColE1 replication control. J. Mol. Biol. 279, 73–88.
Pepper, K. et al., 1988. Heterogeneity of chromosomal genes encoding chloramphenicol resistance in streptococci. Plasmid 19, 71–74.
Priebe, S.D., Lacks, S.A., 1989. Region of the streptococcal plasmid pMV158 required for conjugative mobilization. J. Bacteriol. 171, 4778–4784.
Projan, S.J., Novick, R.P., 1988. Comparative analysis of five related staphylococcal plasmids. Plasmid 19, 203–221.
Puyet, A. et al., 1988. Identification of the origin and direction of replication of the broad-host-range plasmid pLS1. Nucleic Acids Res. 16, 115–133.
Reyes-Lamothe, R. et al., 2014. High-copy bacterial plasmids diffuse in the nucleoid-free space, replicate stochastically and are randomly partitioned at cell division. Nucleic Acids Res. 42, 1042–1051.
Ruiz-Masó, J.A. et al., 2006. Genetic and biochemical characterization of the *Streptococcus pneumoniae* PcrA helicase and its role in plasmid rolling circle replication. J. Bacteriol. 188, 7416–7425.
Ruiz-Masó, J.A. et al., 2007. Interactions between the RepB initiator protein of plasmid pMV158 and two distant DNA regions within the origin of replication. Nucleic Acids Res. 35, 1230–1244.
Seegers, J.F.M.L. et al., 1995. Structural and functional analysis of the single-strand origin of replication from the lactococcal plasmid pWV01. Mol. Gen. Genet. 249, 43–50.
Selinger, L.B. et al., 1990. Mobilization of closely related plasmids pUB110 and pBC16 by *Bacillus* plasmid pXO503 requires trans-acting open reading frame beta. J. Bacteriol. 172, 3290–3297.
Seufert, W. et al., 1988. A novel replicon occurring naturally in *Escherichia coli* is a phage-plasmid hybrid. EMBO J. 7, 4005–4010.
Showsh, S.A., Andrews Jr., R.E., 1999. Analysis of the requirement for a pUB110 *mob* region during Tn*916*-dependent mobilization. Plasmid 41, 179–186.
Smillie, C. et al., 2010. Mobility of plasmids. Microbiol. Mol. Biol. Rev. 74, 434–452.
Smith, M.C.A., Thomas, C.D., 2004. An accessory protein is required for relaxosome formation by small staphylococcal plasmids. J. Bacteriol. 186, 3363–3373.
Soultanas, P. et al., 1999. Plasmid replication initiator protein RepD increases the processivity of PcrA DNA helicase. Nucleic Acids Res. 27, 1421–1428.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Szpirer, C.Y. et al., 2001. Mobilization function of the pBHR1 plasmid, a derivative of the broad-host-range plasmid pBBR1. J. Bacteriol. 183, 2101–2110.
te Riele, H. et al., 1986a. Are single-stranded circles intermediates in plasmid DNA replication? EMBO J. 5, 631–637.
te Riele, H. et al., 1986b. Single-stranded plasmid DNA in *Bacillus subtilis* and *Staphylococcus aureus*. Proc. Natl. Acad. Sci. U.S.A. 83, 2541–2545.
Thomas, C.M., 2000. The Horizontal Gene Pool. Harwood Academic Publishers, Amsterdam, p. 419.
Thomas, C.M., Nielsen, K.M., 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. 3, 711–721.
Thomas, J. et al., 2013. A conserved helicase processivity factor is needed for conjugation and replication of an integrative and conjugative element. PLoS Genet. 9, e1003198.
Turgeon, N., Moineau, S., 2001. Isolation and characterization of a *Streptococcus thermophilus* plasmid closely related to the pMV158 family. Plasmid 45, 171–183.
van der Lelie, D. et al., 1989. Similarity of minus origins of replication and flanking open reading frames of plasmids pUB110, pTB913 and pMV158. Nucleic Acids Res. 17, 7283–7294.
van der Lelie, D. et al., 1990. Conjugal mobilization of streptococcal plasmid pMV158 between strains of *Lactococcus lactis* subsp. *lactis*. J. Bacteriol. 172, 47–52.
Walker, A., 2012. Welcome to the plasmidome. Nat. Rev. Microbiol. 10, 379.
Waters, V.L., Guiney, D.G., 1993. Processes at the nick region link conjugation, T-DNA transfer and rolling-circle replication. Mol. Microbiol. 9, 1123–1130.
Wellington, E.M. et al., 2013. The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria. Lancet Infect. Dis. 13, 155–165.
Wilkins, B., Lanka, E., 1993. DNA Processing and Replication During Plasmid Transfer Between Gram-Negative Bacteria. Plenum Press, New York.
Woodbury, R.L. et al., 2008. Plasmid-borne *erm*(T) from invasive, macrolide-resistant *Streptococcus pyogenes* strains. Antimicrob. Agents Chemother. 52, 1140–1143.
Yamao, F. et al., 1985. UGA is read as tryptophan in *Mycoplasma capricolum*. Proc. Natl. Acad. Sci. U.S.A. 82, 2306–2309.
Yasukawa, H. et al., 1991. Rolling-circle replication of the plasmid pKYM isolated from a gram-negative bacterium. Proc. Natl. Acad. Sci. U.S.A. 88, 10282–10286.
Yin, S. et al., 2009. Functional analysis of the plasmid pM4 replicon from *Lactobacillus plantarum* M4: determination of the minimal replicon and functionality identification of the putative *sso*. Plasmid 62, 166–171.
Zechner, E.L. et al., 2000. Conjugative-DNA transfer processes. In: Thomas, C.M. (Ed.), The Horizontal Gene Pool. Harwood Academic publishers, Amsterdam.
Zhai, Z. et al., 2009. Characterization of a novel rolling-circle replication plasmid pYSI8 from *Lactobacillus sakei* YSI8. Plasmid 62, 30–34.
Zhang, W. et al., 2012. Type-IVC secretion system: a novel subclass of type IV secretion system (T4SS) common existing in gram-positive genus *Streptococcus*. PLoS ONE 7, e46390.
Zhou, H. et al., 2010. Characterization of a rolling-circle replication plasmid pXY3 from *Lactobacillus plantarum* XY3. Plasmid 64, 36–40.