Accepted Manuscript

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

Fabián Lorenzo-Díaz, Cris Fernández-López, M. Pilar Garcillán-Barcia, Manuel Espinosa

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

Fabián Lorenzo-Díaz[a]*, Cris Fernández-López[b]*, M. Pilar Garcillán-Barcia[c]§ and 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-IDICAN, Santander, Cantabria, Spain*

*: Equal contribution

§: Corresponding authors

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)

Key words: Conjugative transfer / Relaxases / Rolling-circle replicating plasmids / origins of transfer / Firmicutes

Running title: pMV158 family of plasmids

25 **Abstract**

26 Rolling circle-replicating plasmids constitute a vast family that is particularly

27 abundant in, but not exclusive of, Gram-positive bacteria. These plasmids are

28 constructed as cassettes that harbor genes involved in replication and its

29 control, mobilization, resistance determinants and one or two origins of lagging

30 strand synthesis. Any given plasmid may contain all, some, or just only the

31 replication cassette. We discuss here the family of the promiscuous

32 streptococcal plasmid pMV158, with emphasis on its mobilization functions: the

33 product of the *mobM* gene, prototype of the MOB_V relaxase family, and its

34 cognate origin of transfer, *oriT*. Amongst the subfamily of MOB_{V1} plasmids,

35 three groups of *oriT* sequences, represented by plasmids pMV158, pT181, and

36 p1414 were identified. In the same subfamily, we found four types of single-

37 strand origins, namely *ssoA*, *ssoU*, *ssoW*, and *ssoT*. We found that plasmids of

38 the rolling-circle Rep_2 family (to which pMV158 belongs) are more frequently

39 found in Lactobacillales than in any other bacterial order, whereas Rep_1

40 initiators seemed to prefer hosts included in the Bacillales order. In parallel,

41 MOB_{V1} relaxases associated with Rep_2 initiators tended to cluster separately

42 from those linked to Rep_1 plasmids. The updated inventory of MOB_{V1} plasmids

43 still contains exclusively mobilizable elements, since no genes associated with

44 conjugative transfer (other than the relaxase) were detected. These plasmids

45 proved to have a great plasticity at using a wide variety of conjugative

46 apparatuses. The promiscuous recognition of non-cognate *oriT* sequences and

47 the role of replication origins for lagging-strand origin in the host range of these

48 plasmids are also discussed.

49

Contents

Abbreviations: Antibiotic resistance(s), AbR; Gram-negative, G-; Gram-positive, G+; Integrative Conjugative Elements, ICEs; Leading-strand initiation and control of replication, LIC; primer RNA, pRNA; rolling circle, RC; rolling circle-replicating plasmids, RCR-plasmids; RNA polymerase, RNAP; single-stranded, ss; double-stranded, ds; single strand origin, *sso*; transferred DNA strand, T-DNA; Type IV secretion systems, T4SSs; T4SS-coupling protein, T4CP; Vertical/Horizontal Gene Transfer, VGT/HGT

§ **Corresponding authors at:** Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu, 9, E-28040 Madrid, Spain; Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC); Universidad de Cantabria–CSIC-IDICAN, Albert Einstein 22, PCTCAN, 39011 Santander, Cantabria, Spain

3

76 **Highlights**

77

78 - The replication and transfer of rolling circle-replicating plasmids is reviewed.

79 - Comparisons of replication and transfer cassettes are presented.

80 - The current understanding of the pMV158 DNA transfer mechanism is reviewed.

1. Introduction

The concept developed by Marshall McLuhan in the early 1960's 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

106 rapidly spread (even as epidemic outbursts in hospitals) by horizontal gene

107 transfer, HGT (Anderson and Seifert, 2011; Baquero, 2009). Genes responsible

108 for AbR frequently cluster in the bacterial mobile elements (the mobilome) within

109 the Integrative and Conjugative Elements (ICEs) and transmissible plasmids

110 that, in turn, are platforms to recruit various smaller mobile elements, such as

111 insertion sequences, transposons and integrons that can also encode AbR

112 genes (Frost et al., 2005). Thus, HGT mediated by self-replicating plasmids or

113 by transfer of the islands, plays an essential role in the bacterial, and

114 consequently in the global biodiversity.

115   In the Plasmid Biology field, the processes pertaining to the

116 dissemination of genetic information stored in plasmids are of special relevance.

117 Those processes involve the two main ways of inheritance, namely vertical

118 (VGT, from mother to daughter cells) and HGT (cell-to-cell) gene transfer

119 (Thomas, 2000). Whereas the former includes DNA replication and partition,

120 this latter being coupled to the cell division, HGT is usually achieved by

121 conjugation between cells of the same or different species (del Solar et al.,

122 1998; Grohmann et al., 2003; Lanka and Wilkins, 1995). Replication and

123 conjugation are, in addition to genetic recombination, the most important

124 sources of genetic variability among plasmids and their hosts. The entire

125 genetic content of the bacterial mobilome amounts up to 25% of the total DNA

126 circulating among bacteria, thus being a shared strong task force for the

127 evolution of the bacterial populations (Ochman et al., 2000; Thomas and

128 Nielsen, 2005). A substantial part of the mobilome is constituted by plasmids,

129 the so-called plasmidome (Walker, 2012).

130   Plasmids are much more than cloning vectors or tools to over-produce

131   proteins. In addition to their role in the spread of genetic information, bacterial

132   plasmids are excellent models to study a number of biological processes, such

133   as transactions involving macromolecular interactions: protein-DNA, protein-

134   protein, and DNA-RNA (Espinosa, 2013). They also constitute a wealth of

135   information on control of gene expression (del Solar and Espinosa, 2000; del

136   Solar et al., 2002), intracellular distribution of DNA molecules (Reyes-Lamothe

137   et al., 2014), and can be considered as useful models for system biology

138   (Paulson and Ehrenberg, 1998). Excellent books dealing with various aspects of

139   the biology of plasmids have been published (Funnell and Phillips, 2004;

140   Thomas, 2000) and a new one is in press (Alonso and Tomalsky, 2014). In the

141   present review we will concentrate on plasmids replicating by the rolling circle

142   mechanism (RCR-plasmids) because they may encode up to two proteins with

143   endonuclease/topoisomerase-like activity that cleave supercoiled DNA at two

144   different regions on the same molecule, making their study thought-provoking.

145

146   **2. Replication, conjugation, and mobilization: common themes**

147

148   Replication and transfer generally require that plasmid-encoded

149   protein(s) interact with their cognate DNA sites to initiate the process. The key

150   initiator players in these two processes are: i) the generically called Rep

151   proteins involved in vegetative replication and that interact with their cognate

152   origin of replication, *oriV*, and ii) the relaxases, usually termed Tra or Mob

153   proteins, involved in transfer, and that recognize their cognate origin of transfer,

154   the *oriT*. Thus, interplays between Rep-*oriV* and Tra/Mob-*oriT* will define the

155   VGT and the HGT processes, respectively. The early discovery of conjugation

156   (Lederberg and Tatum, 1946; Lederberg and Tatum, 1953) and the existence of

157   plasmid-encoded proteins that relax the donor DNA as the first stage in the

158   transfer process, led to a historically broad interest in plasmid transfer among

159   Gram-negative (G-) bacteria. Conjugation in Gram-positive (G+) bacteria was

160   studied relatively later and it is still a matter of active research (Goessweiner-

161   Mohr et al., 2013; Grohmann et al., 2003). Replication involves a number of

162   different strategies devoted to the melting of the DNA strands at the origin of

163   replication (namely theta, strand-displacement and rolling-circle mechanisms),

164   and is usually mediated by the Rep initiator alone or with the help of plasmid- or

165   host-encoded proteins (del Solar et al., 1998). Although DNA melting to initiate

166   the transfer is also required, conjugation, however, involves a single strategy

167   devoted to the unidirectional transfer of mobile elements (plasmids or ICEs)

168   from a donor bacterial cell to a recipient one through physical cell-to-cell contact

169   (de la Cruz et al., 2010; Zechner et al., 2000). Based on their transfer

170   machinery we can distinguish between: i) self-transmissible (conjugative)

171   plasmids and integrative conjugative elements (ICEs), which codify all the

172   functions required for their HGT, and ii) mobilizable plasmids and integrative

173   and mobilizable elements (IMEs), which 'travel light' because they encode only

174   the relaxase and its cognate *oriT*. These latter plasmids and IMEs make use of

175   functions provided by either the host chromosome or by other auxiliary (also

176   termed 'helper') plasmid for their transference.

177   The assembly of plasmid- and host-encoded proteins in a specific DNA

178   region initiates replication and transfer, thus initiation of both processes requires

179   the   generation   of   multiprotein-DNA-complexes,   the   replisome   or   the

180 relaxosome, respectively (del Solar et al., 1998; Lanka and Wilkins, 1995;

181 Pansegrau and Lanka, 1996a). In the case of RCR-plasmids (del Solar et al.,

182 1987; Khan et al., 1981; Novick, 1998; Puyet et al., 1988; te Riele et al., 1986a;

183 te Riele et al., 1986b), initiation of replication and mobilization are

184 mechanistically similar processes. Both, Rep and Tra/Mob proteins have

185 endonuclease/topoisomerase-like activities on their DNA targets, the double-

186 strand origin (*dso*) or the *oriT*. These proteins cleave their cognate DNA at the

187 phosphodiester bond of a specific di-nucleotide (the *nic* site) generating a stable

188 amino acyl-DNA adduct (Chandler et al., 2013; de la Campa et al., 1990;

189 Guasch et al., 2003; Khan, 2003; Koepsel et al., 1986; Moscoso et al., 1997;

190 Pansegrau and Lanka, 1996a). The nick introduced by Rep or Tra/Mob proteins

191 generates a free 3'-OH end, which acts as a primer for leading-strand synthesis

192 in both cases, VGT and HGT. In the RCR-plasmids, proteins from the host

193 replicative machinery, at least DNA-polymerases I and III, single-strand (ss)

194 DNA-binding protein and PcrA helicase, participate in the elongation from the

195 3'-OH end generated by the plasmid-encoded Rep initiator (Anand et al., 2005;

196 Anand and Khan, 2004; Díaz et al., 1994; Khan, 2003; Khan, 2005; Machón et

197 al., 2010; Ruiz-Masó et al., 2006; Soultanas et al., 1999; Thomas et al., 2013).

198 To process their DNA substrates, initiators of replication and transfer require the

199 *nic* site being exposed in a single-stranded configuration, which can be

200 achieved by DNA melting and generation of hairpin structures. Hairpin formation

201 would be mediated either by the binding of auxiliary associated proteins, and/or

202 by binding of the initiator or the relaxase. In both cases, the di-nucleotide to be

203 cleaved will be exposed in an unpaired form (Jin and Novick, 2001; Lorenzo-

204 Díaz et al., 2011; Lucas et al., 2010; Noirot et al., 1990; Ruiz-Masó et al., 2007).

205   Besides the initiator (relaxase) and its binding site (*oriT*), conjugation

206   requires an additional machinery that RCR does not, namely the protein

207   complex that completes the conjugation apparatus. This is a highly specialized

208   protein machinery encoded by the donor plasmid DNA (or the host chromosome

209   in some cases) that includes the coupling protein (T4CP) and the Type IV

210   Secretion System (T4SS) which translocate the relaxase-DNA complex to the

211   recipient cell (Draper et al., 2005; Garcillán-Barcia et al., 2007). T4SSs recruit

212   their substrate by mechanisms still not fully understood for conjugation, and

213   they also participate in other processes involving DNA trafficking between

214   prokaryotic donors and recipient (prokaryotic or eukaryotic) cells, like in the

215   case of Ti and related plasmids (Baron et al., 2002; Bhatty et al., 2013;

216   Cascales et al., 2013; Chen et al., 2005; de Paz et al., 2005; Goessweiner-Mohr

217   et al., 2013; Gomis-Ruth et al., 2004; Hamilton et al., 2000; Llosa et al., 2009;

218   Zhang et al., 2012). The T4CP and the T4SS proteins participate in pumping

219   the transferred DNA (T-DNA)-relaxase complex into the recipient cell (Gomis-

220   Rüth and Coll, 2006; Guasch et al., 2003; Llosa et al., 2002; Matilla et al.,

221   2010). In RCR as well as in transfer, the parental DNA strand will be displaced

222   until either the *dso* or the *oriT* are reconstituted. This last stage involves DNA

223   strand transfer reactions that will terminate either leading strand replication (in

224   the donor cell) or plasmid transfer (in the recipient cell). Intermediates of both

225   processes will be ssDNA molecules that correspond to the parental plus strand

226   (del Solar et al., 1998; Grohmann et al., 2003; Pansegrau and Lanka, 1996a;

227   Pansegrau and Lanka, 1996b; Wilkins and Lanka, 1993). Synthesis of the

228   lagging strand initiates from the so-called single-strand origins (*sso*), as

229   described below.

230

3. Modular construction of RCR-plasmids

232

233   A relevant part of the mobilizable plasmids replicate by the RC

234 mechanism (del Solar et al., 1987; Khan et al., 1981; Novick, 1998; Puyet et al.,

235 1988; te Riele et al., 1986a; te Riele et al., 1986b), although not all of them are

236 mobilizable. In the case of small plasmids, isolated primarily from G+ bacteria,

237 two pioneer discoveries were made: i) the identification of specific single-

238 stranded DNA molecules (ssDNA) as intermediates of plasmid replication

239 (Puyet et al., 1988; te Riele et al., 1986a; te Riele et al., 1986b), and ii) the

240 finding that their Rep initiator proteins exhibited a sequence-specific relaxing

241 activity on supercoiled DNA (Koepsel et al., 1985). These results led to the

242 discovery of RCR-plasmids, which constituted a new class of plasmids that

243 replicate by the asymmetric rolling circle mechanism that share similarities with

244 the replication of ssDNA coliphages (Novick, 1998). Furthermore, the presence

245 of RCR-plasmids in G- bacteria (Yasukawa et al., 1991) helped to cast off the

246 idea of a genetic barrier between the two types of bacteria (del Solar et al.,

247 1993). Thus, the conclusion that plasmid RCR is mechanistically similar to

248 conjugative transfer was soon achieved (Waters and Guiney, 1993). It was

249 interesting to learn that some of the RCR-plasmids encoded, in addition to the

250 Rep topoisomerase-like initiator, another protein with the ability to relax DNA

251 (Caryl et al., 2004; Grohmann et al., 2003; Guzmán and Espinosa, 1997; Smith

252 and Thomas, 2004). These proteins were thought to be involved in inter-

253 plasmidic recombination (plasmid recombination enzymes, Pre; (Projan and

254 Novick, 1988), although it was later shown that Pre proteins were required for

255 plasmid mobilization (Priebe and Lacks, 1989). Further, the relaxase activity of

256 the pMV158_Pre protein (renamed MobM) on supercoiled DNA, and its *nic* site

257 were demonstrated *in vitro* and *in vitro* (Grohmann et al., 1997; Guzmán and

258 Espinosa, 1997). However, even up today, these proteins are grouped into a

259 family termed Mob-Pre at the database of protein families Pfam (PF01076), and

260 no further investigations on their participation in plasmid recombination have

261 been performed.

262   In general, RCR-plasmids appear to be constructed as gene cassettes

263 (del Solar et al., 1993; Khan, 1997; Khan, 2005) that may have up to four

264 independent modules involved in: i) Leading-strand initiation and control of

265 replication (LIC); ii) AbR determinant (DET); iii) Mobilization (MOB), and iv) One

266 or two origins for lagging strand replication, *sso* (Fig. 1). This latter kind of

267 origins varies among the different RCR described, but they have been

268 categorized in four types: *ssoA*, *ssoU*, *ssoW*, and *ssoT* depending on their DNA

269 sequence (Khan, 2000; Khan, 2005; Kramer et al., 1998a; Kramer et al., 1997;

270 Meijer et al., 1995b; van der Lelie et al., 1989), or in their role in plasmid

271 promiscuity (Kramer et al., 1995; Lorenzo-Díaz and Espinosa, 2009). Their

272 structure and roles will be described below (see Table 1). RCR-plasmids may

273 harbor all the cassettes, like the streptococcal plasmids pMV158 and pRW35

274 (Priebe and Lacks, 1989; van der Lelie et al., 1989; Woodbury et al., 2008), or

275 just the LIC cassette, as in the mycoplasma plasmids pADB201 (Bergemann et

276 al., 1989) and pKMK1 (King and Dybvig, 1992), which are considered to be the

277 smallest RCR-plasmids, unless we consider the hybrid phage-plasmid phasyl

278 (Seufert et al., 1988) as one of them.

279   Many of these plasmids have left their 'print' on the chromosome of the

280   hosts they might have colonized and, in fact, a bioinformatics survey detected

281   RCR plasmids–related MOB$_V$-relaxase genes in 63 out of 1207 chromosomes

282   analyzed (Guglielmini et al., 2011). The AbR trace could be considered as a

283   thought-provoking approach to follow the RCR-plasmid 'fate' along evolutionary

284   history. Although we have not performed any further search, a homolog of the

285   pMV158-*tetL* determinant was found in the chromosome of *Bacillus subtilis*

286   (Lacks et al., 1986), whereas a *cat* gene, homologous to the one harbored by

287   plasmids pC194 and pC221 has been described to be present in the

288   chromosome of G+ bacteria, like *Bacillus pumilus* (Harwood et al., 1983) or

289   *Streptococcus pneumoniae* (Pepper et al., 1988), and even in the chromosome

290   of *Clostridium perfringens* (Bannam and Rood, 1991). These findings argue in

291   favor of the role of these RCR-plasmids in the integration and dispersion of

292   resistance genes in the chromosome of hosts that they have colonized.

293

294   **4. The family: pMV158 and relatives**

295

296   In general, plasmids have been grouped according to their replicon, since

297   this is the hallmark of the plasmid (Chang et al., 2000; del Solar et al., 1998;

298   Khan, 1996; Nordström et al., 1984; Novick, 1989). As mentioned above, RCR-

299   plasmids were grouped in several families, although the most studied are the

300   Rep_1 family (PF01446), whose prototypes are the staphylococcal plasmids

301   pC194, and pUB110, the Rep_2 family (PF01719) of which the streptococcal

302   plasmid pMV158 and the staphylococcal plasmid pE194 are the

303   representatives, and the Rep_trans family (PF02486) represented by the

304 staphylococcal plasmids pT181 and pC221 (del Solar et al., 1998; del Solar et

305 al., 1993; Gruss and Ehrlich, 1989; Khan, 1997; Khan, 2000; Khan, 2005;

306 Novick, 1989). These three families harbor all the cassettes with the exceptions

307 of pC194 that does not contain a MOB module and the lagging-strand origin

308 *ssoU*, which is reported to be present only in pUB110 and pMV158 (Figure 1).

309 Curiously, these two latter plasmids also share identical *oriT* sequences. *In*

310 *vitro*, the MobM-protein from pMV158 was able to relax supercoiled DNA from

311 both plasmids (Fernández-López et al., 2013a). Out of the representatives of

312 the different families, pC194 has been reported to be one of the most

313 promiscuous, since it was shown to replicate not only in staphylococci,

314 streptococci and bacilli (Ballester et al., 1990; Horinouchi and Weisblum, 1982),

315 but also in *Escherichia coli* and in the yeast *Saccharomyces* as well (Goursot et

316 al., 1982). In the case of pMV158, it has been shown to exhibit an extraordinary

317 host range. The plasmid was primarily isolated from *Streptococcus agalactiae*

318 (Burdett, 1980) and, along the years, it has been transferred in the laboratory

319 (either by mobilization or by transformation) to more than 20 bacterial species,

320 from Firmicutes to α-Proteobacteria and in all of them it replicates stably (del

321 Solar et al., 1998; Espinosa, 2013; Lacks et al., 1986).

322 After a BlastP search using prototypes of the three RCR-plasmid

323 families, we found the following distribution: i) in Rep_1 (prototype RepA from

324 pC194), 334 non-redundant homologs in Bacillales, and 157 in Lactobacillales;

325 ii) in Rep_2 (prototype taken, RepE initiator from pE194), 44 non-redundant

326 homologs in Bacillales and 242 in Lactobacillales; iii) in Rep_trans (RepC from

327 pT181 as prototype), 176 non-redundant homologs in Bacillales and 241 in

328 Lactobacillales. These figures suggest that whereas the Rep_1 family of RCR-

14

329   plasmids would prefer to colonize Bacillales and the Rep_2 family of pMV158

330   would be better fitted to Lactobacillales, the Rep_trans family of RCR-plasmids

331   would be distributed more evenly. In general, we found a good correlation

332   between the G+C content of RCR-plasmids and their respective hosts

333   (Espinosa et al., 1995), with the exception of the staphylococcal plasmid

334   pUB110 which has a G+C content close to 45%, making it a plasmid which has

335   been considered more like a bacilli than a staphylococci replicon (Alonso et al.,

336   1988; McKenzie et al., 1986).

337   Traditionally, RCR-plasmids have been classified according their LIC

338   module (del Solar et al., 1998; del Solar et al., 1993). According to this criterion,

339   we performed a PSI-BLAST search for plasmids of the pMV158 family using the

340   initiator $RepB_{pMV158}$ as the query. The results are compiled in Supplementary

341   Table S1. The 78 plasmids retrieved belong to the Rep_2 family. Out of them,

342   55 did not encode any relaxase whereas the remaining 23 contained relaxase

343   genes belonging to the MOB_V family (see below).

344   Many of the retrieved plasmids lack any distinguishable marker, although

345   resistance to a variety of antibiotics, and even one instance of resistance to

346   arsenate was found. A phylogenetic tree of the 78 plasmids retrieved was

347   constructed (Supplementary Figure S1), using the protein GpA initiator from

348   phage ΦX174 as the out-group representative. There will be a detailed in-depth

349   analysis on the LIC module of plasmids of the pMV158 family to be published

350   elsewhere (G. del Solar and J.A. Ruiz-Masó, personal communication).

351   Furthermore, most of the retrieved plasmids were primarily isolated from

352   Lactobacillales, but as shown in Supplementary Figure S1 they are also widely

353   distributed in other taxonomic orders of Firmicutes, Tenericutes and

354 Proteobacteria, with notable examples of plasmids from *Mycoplasma* (eight

355 plasmids). Although not retrieved in this search, there is an RCR-plasmid from

356 *Mycoplasma yeatsii*, pMyBK1, which deserves a mention here because of two

357 features (Kent et al., 2012; Breton et al., 2012). Firstly, it is unique at encoding

358 an initiator of RCR different from the ones described here. Secondly, pMyBK1 is

359 the only example of mobilizable plasmid in genus *Mycoplasma*, and it is

360 precisely a MOB$_V$ plasmid. Inspecting the sequence of the *mob* gene of

361 pMyBK1, we found that it harbors 39 UGA codons (out of a total of 520 codons).

362 Since in *Mycoplasma* UGA specifies "tryptophan" instead of "stop" (Halbedel

363 and Stülke, 2007; Inamine et al., 1990; Yamao et al., 1985), we can conclude

364 that the transfer of pMyBK1 would be limited to *Mycoplasma* species.

365

366 **5. The MOB$_V$ family**

367

368 Classification of plasmids according to the relaxases and origins of

369 transfer they carry has supposed a novel definition of plasmid families (Francia

370 et al., 2004; Garcillán-Barcia et al., 2009). This classification provided a more

371 global view of HGT by conjugation, whereas classification by replicons would

372 point more to the vertical transfer. The HGT-based classification has showed

373 that there are many plasmids that do not encode transfer functions (Garcillán-

374 Barcia et al., 2011; Smillie et al., 2010). RCR-plasmids of Firmicutes fall mostly

375 in the so-called MOB$_V$ family, being the MobM relaxase from the streptococcal

376 plasmid pMV158 the representative of the family (Garcillán-Barcia et al., 2009).

377 It was believed that plasmids are DNA molecules that result from shuffling of

378 various gene cassettes, evolving independently one of each other (Osborn et

379 al., 2000). As we will discuss below, analysis of evolutionary roots of these

380 cassettes motivates to think that there could be crosstalk between them, as

381 some module combinations prevail among others.

382   To update the inventory of elements composing the MOB_{V1} subfamily we

383 have performed a PSI-BLAST search using $MobM_{pMV158}$ as the query. The

384 retrieved MOB_{V1} plasmids ranged from 2.7 to 30 kb (median = 5.1 kb; Figure 2).

385 Roughly half of them coded for antibiotic resistance genes (mainly to a wide

386 variety of protein synthesis inhibitors, such as macrolides, lincosamides,

387 aminoglycosides, streptogramins, amphenicols and tetracyclines) or other

388 virulence traits like resistance to heavy metals and production of bacteriocins

389 (Table 1). The organization of the mobilization region was found to be similar in all

390 members: the *oriT* located upstream, close to the *mob* gene. Furthermore, the *nic*

391 site was placed on the same strand as the *mob* gene, as expected.

392

393 **5.1. The origin of transfer (*oriT*)**

394

395   Plasmid conjugation initiates through the assembly of the relaxosome on

396 the *oriT*, a region that contains inverted (IR) and/or direct (DR) repeats, A+T-

397 rich tracts and, most importantly, the *nic* site. The relaxase requires that its

398 target is presented as ssDNA to cleave it and generate the relaxase-DNA

399 adduct, which will be pumped through a T4SS. Once in the recipient, the *oriT* is

400 reconstituted by a transesterification reaction mediated by the relaxase to close

401 the incoming molecule (reviewed in (Chandler et al., 2013).

402   The *oriT* of pMV158 spans 41 bp upstream of the *mobM* gene, exhibits a

403 high A+T content (75.6%) and has three IRs (IR1 to IR3; Lorenzo-Díaz et al.,

17

404 2011). Furthermore, a 6-bp sequence (5'-ACTTTA-3') is repeated in the IR1/IR3

405 left-arm and the IR2 right-arm (Figure 3). The *nic* site was firstly mapped *in vitro*

406 between coordinates 3595-3596 (dinucleotide 5'-GpT-3') in the pMV158 sequence

407 (GenBank Acc. No. NC_010096; (Guzmán and Espinosa, 1997)). Then, the *nic*

408 sites for the pMV158 and pE194 *oriT*s were determined in their respective hosts *in*

409 *vivo* (Grohmann et al., 1997), mapping exactly in the same position as already

410 suggested by the previous *in vitro* results. *In silico* analysis revealed that the IRs

411 could generate three alternative stem-loop structures in which the position of

412 the *nic* site would be placed in different positions: i) located 8-bp upstream to

413 IR1, ii) in the IR2 inter-arm region, or iii) at the 3'-end of the IR3 (Lorenzo-Díaz

414 et al., 2011). IR3 includes the IR1 sequence and, since IR1/3 and IR2 partially

415 overlap (see Figure 3), the generation of cruciform secondary structures by one

416 of them would hinder the formation of the other, indicating that the target DNA

417 accessibility by the relaxase could depend on the plasmid DNA superhelicity

418 (Fernández-López et al., 2014). Our *in vitro* analysis demonstrated that MobM

419 binds specifically to ssDNA encompassing IR1/3 with high affinity (Lorenzo-Díaz

420 et al., 2011), which allowed the MobM protein to repress its own synthesis

421 (Lorenzo-Díaz et al., 2012). Functional relevance of the IRs in the different steps

422 of the conjugative process is currently under exploration. We hypothesize that

423 IR1/3 may be involved in the Mob-recognition of the *oriT* at the initiation of the

424 relaxosome formation (in the donor cell) and IR2 at the termination reaction to

425 close the T-strand (in the recipient cell).

426   Based on the sequence and structure of *oriT*$_{pMV158}$, we inspected the *oriT*

427 regions of the 93 MOB$_V$-plasmids listed in Table 1. A total of 97 sequences were

428 manually identified upstream to their respective relaxase encoded genes (five of

429   them exhibiting two different MOB_V-related *oriT*s: pSTE1, pKKS285, pSCFS1,

430   unnamed (GenBank Acc. No GG692894.1), and pTB19. Analysis of the region

431   showed a high degree of sequence conservation in the majority of the *oriT*s,

432   being composed by three IRs and exhibiting the consensus sequence

433   'GTGBG↓T' for the *nic* site (B denoting a G, C or T following IUPAC code; '↓'

434   being the *nic* site). Two other minor plasmid groups, represented by pT181 and

435   p1414, grouped in different clusters given their differences respect to the *oriT*

436   sequence of pMV158 (Figure 3 and Supplementary Figure S2). However, these

437   two clusters maintain the number and distribution of the IRs. Unfortunately, the

438   experimental evidences for mapping the specific *nic* site in these plasmids are

439   scarce and only a few *oriT* predictions, such as for plasmids like pSMQ172 and

440   pPB1, have been published (de las Rivas et al., 2004; Turgeon and Moineau,

441   2001).

442   Given that the $oriT_{pMV158}$ is the only element required in *cis* for pMV158 to

443   be transferred (Farías and Espinosa, 2000), it is plausible to assume that any

444   RCR-plasmid containing an *oriT*-like sequence would be potentially mobilizable.

445   Such is the case of pXY3 (Zhou et al., 2010) and pCI411 (Coffey et al., 1994)

446   plasmids (not included in Table 1), which only seem to contain an orphan and

447   non-canonical *oriT*. In fact, the prototype RCR-plasmid pC194, which lacks both a

448   *mob* gene and a canonical *oriT*, has been mobilized by using the conjugative

449   transposon *Tn*916 as a helper (Naglich and Andrews, 1988; Showsh and Andrews

450   Jr, 1999). It is assumed that the relaxase of the helper should be able to recognize

451   a suitable *oriT* in the Δ*mob* plasmid. It is more intriguing the finding that pC194

452   and Δ*oriT*-Δ*mob*-derivatives of plasmids pUB110 and pTA1060 were efficiently

453   mobilized by the mating apparatus of ICE*Bs1*, but without the intervention of the

454 ICE*Bs1* NicK relaxase (Lee et al., 2012). Thus, in this case, the plasmids were not

455 transferred by cross-recognition of an *oriT* by the relaxase of the helper. They

456 were neither transferred by forming cointegrates with ICE*Bs1*. Unexpectedly, the

457 RCR initiation proteins of the mobilized plasmids were crucial for transfer (Lee et

458 al., 2012). Thus, the strategy we have followed to search for plasmids containing

459 MOB_{V1} relaxases might underestimate the real population of mobile elements.

460

461 **5.2. The MOB_{V1} relaxases**

462

463 The second element that belongs to the transfer module of the pMV158

464 RCR-family is the *mobM* gene, which codifies the MobM relaxase (Guzmán and

465 Espinosa, 1997). Thus, in addition to using the homology of the Rep proteins to

466 define the pMV158 RCR-plasmid family (Supplementary Figure S1 and

467 Supplementary Table S1), we decided to find out whether any relationship

468 between the Rep- and the Mob-proteins of the plasmid family existed. The

469 classification system for mobilizable plasmids based on the amino acid sequence

470 of the relaxases allowed defining MobM from pMV158 as the prototype of the

471 MOB_V superfamily (Francia et al., 2004; Garcillán-Barcia et al., 2011; Garcillán-

472 Barcia et al., 2009). This superfamily is composed of more than 200 relaxases, of

473 which about 140 were located in plasmids and the rest in bacterial chromosomes

474 (Guglielmini et al., 2011).

475 Five subfamilies were described, and MobM from pMV158 was taken as

476 the prototype of the MOB_{V1} subfamily (Garcillán-Barcia et al., 2009). Members of

477 this subfamily showed three conserved motifs: i) Motif I (HxxR), of yet unknown

478 function; ii) Motif II (NY(D/E)L), which is the proposed catalytic domain, and iii)

479   Motif III (HxDE...PHxH), which corresponds to the metal-coordination motif, also

480   known as the 3H motif (Chandler et al., 2013). The MOB_{V2} family, represented by

481   the Mob protein of the theta-replicating plasmid pBBR1 (Szpirer et al., 2001),

482   exhibited Motifs I and III but lacked Motif II. The three motifs are located in the N-

483   terminal moiety of the MOB_V relaxases. This moiety harbors the DNA binding and

484   nicking activities, since a truncated version of MobM, which only contains the first

485   N-terminal 199 residues, retained the relaxase activity on supercoiled DNA

486   (Fernández-López et al., 2013a; Fernández-López et al., 2013b; Lorenzo-Díaz et

487   al., 2011). The C-terminal moiety of MobM could be involved in, at least, two

488   functions: i) protein-protein interactions (dimerization and interactions with the

489   auxiliary plasmid-encoded coupling protein) and ii) association with the cell

490   membrane, through a proposed coiled-coil region located between residues 400

491   and the C-terminal end of MobM. Disruption of the alpha helical-rich region by

492   mutations (changes to Pro residues) resulted in failure of MobM-association with

493   membranes; further, the pMV158-derivative harboring these mutations lost its

494   ability to be transferred (de Antonio et al., 2004). There was no indication of a

495   helicase activity in the C-terminal moiety of MobM (our unpublished observations).

496   We have updated the inventory of elements that cluster with pMV158 into

497   the MOB_{V1} subfamily. 97 non-redundant plasmid MOB_{V1} relaxases were retrieved

498   from a PSI-BLAST search using the N-terminal 300 amino acids of MobM from

499   pMV158 as a query (Table 1). All of them shared the three Motifs described above

500   (Figure 4A). The inferred phylogeny of MOB_{V1} relaxases, rooted by a MOB_{V2}

501   relaxase, showed well-supported external clusters of highly related sequences

502   and poorly-supported internal nodes (Figure 4B). It is a fact that reflects the low

503   overall similarity of the taxa, mainly circumscribed to the three relaxase motifs

504 described (Francia et al., 2004; Garcillán-Barcia et al., 2009). The plasmids coding

505 these relaxases are primarily distributed in several genera of Lactobacillales and

506 Bacillales, with a few members out of the phylum Firmicutes. Curiously, the

507 plasmids encoding relaxases comprised in a monophyletic clade in the

508 phylogenetic tree shown in Figure 4B are generally hosted in a single taxonomic

509 order (either Lactobacillales or Bacillales), suggesting less inter-order transfer than

510 expected.

511   No genes associated with conjugative transfer others than the relaxase

512 ones were encoded by these plasmids. Thus, they are classified as mobilizable,

513 requiring the conjugative machinery of other plasmids to be transferred (namely, a

514 coupling protein and a mating-pair formation apparatus composed by a T4SS and

515 a conjugative pilus). Eight mating-pair formation (MPF) types have been

516 phylogenetically described (Guglielmini et al., 2011); three of them, MPF_{FATA},

517 MPF_{FA} and MPF_T were able to mobilize MOB_{V1} plasmids. Specifically, pMV158

518 was mobilized between G+ bacteria by functions supplied by helper MPF_{FATA}

519 plasmids of the Inc18 family, like pIP501 and pAMβ1 (Grohmann et al., 1999;

520 Priebe and Lacks, 1989; van der Lelie et al., 1990), and even to the G− bacterium

521 *Escherichia coli* by MPF_T plasmids RP4 and R388 but not by MPF_F plasmid F

522 (Farías and Espinosa, 2000). Other elements, such as pC194 and pUB110, were

523 mobilized by auxiliary MPF_{FA} elements such as *Tn*916 (Lee et al., 2012; Naglich

524 and Andrews, 1988), and MPF_{FATA} plasmids, like pLS20, mobilized pUB110 and

525 pBC16 (Koehler and Thorne, 1987; Selinger et al., 1990).

526   Most of the MOB_{V1} plasmids replicate by the rolling circle mechanism

527 (Figure 5 and Table 1). MOB_{V1} relaxases are predominantly linked to RCR

528 initiators of the three different subgroups: Rep_1 (PF01446), Rep_2 (PF01719),

529 and Rep_trans (PF02486). A small fraction of MOB_{V1} relaxases is linked to a wide

530 variety of theta replication families, of which Rep_3 is the most abundant.

531 Congruently, with the abovementioned taxonomic bias in the abundance of RCR

532 initiators and the MOB_{V1} relaxase distribution, most relaxases encoded in Rep_2

533 RCR plasmids grouped separately from those encoded in Rep_1 RCR plasmids

534 (clade Lactobacillales vs. Bacillales in Figure 6 and Supplementary Table S2). An

535 example of this bias is found in Enterococci (Lactobacillales), where Rep_1

536 initiators are commonly found in multireplicon plasmids, and may not be

537 functional, as it is the case for plasmid pAMα1 (Clewell et al., 2014).

538 Nevertheless, a few functional Rep_1 initiators can be also found in the

539 Lactobacillales clade, such as those grouped in the cartooned pSMA23 cluster of

540 Figure 4B. Curiously, within the pSMA23 group some exceptions also exist:

541 plasmid pCD034-2 encodes a Rep_2 instead of a Rep_1 initiator, an indication of

542 recent recombination events that led to new backbone combinations. Besides,

543 more ancient recombination events could give rise to arrangements present in

544 plasmids of the Bacillales MOB_{V1} clade, since the highly divergent relaxases

545 included in that group are linked to initiators of the three RCR subgroups (Figure

546 6A).

547   Figure 6 also provides an interesting example of the switch of plasmids

548 from and to integrative elements. ICE*Sgal1* is an integrative element highly similar

549 to plasmid pUB110 (both in relaxase and initiator) but located in the chromosome

550 of *Streptococcus gallolyticus* UCN34 (Lactobacillales order). It is tempting to

551 speculate that this element integrated in the chromosome once transferred from

552 Bacillales to the Lactobacillales background (maybe helped by a Tn*916*-like

553 element located close to ICE*Sgal1*), where it was not able to replicate and

554 became an integrative and mobilizable element (IME).

555   A significant proportion (one third) of MOB_{V1} plasmids recorded in Table 1

556 encoded more than one replication initiation protein (Figure 5). Besides, 10 out of

557 the 93 plasmids coded more than one relaxase gene. Both facts suggest the

558 frequent arising of plasmid cointegrates. Precisely, it is known the ability of RCR-

559 plasmid encoded relaxases to promote site-specific DNA recombination at *oriT*

560 rendering plasmid cointegrates during conjugative mobilization, where the host *rec*

561 system may function to stimulate the recombination process (Gennaro et al.,

562 1987; Novick et al., 1984; Projan and Novick, 1988). Further, it is worth recalling

563 that the staphylococcal plasmid pE194 was able to integrate into the *Bacillus*

564 *subtilis* chromosome by a RecA-independent recombination mechanism, and

565 using as little as 6-14 bp homologies (Dempsey and Dubnau, 1989). This, in

566 conjunction with the finding of RCR-plasmids integrated into other bigger plasmids

567 (Oskam et al., 1991), raises the question of whether RCR-plasmids that are

568 integrated on IMEs would participate in their replication and mobilization; in fact,

569 we have found a putative initiator of replication of the Rep_1 family within an

570 streptococcal island which also encode a MOB_{V1} protein (our unpublished

571 observations). Besides, the cross-recognition of heterologous *oriT*s by MOB_V

572 relaxases (Fernández-López et al., 2013a) could be favored in the multi-relaxase

573 cointegrates, potentiating their spreading.

574   Recombination seems to be the most likely cause of the different

575 topologies exhibited in the dendrograms of initiators and relaxases (compare trees

576 in panels A and B of Figure 6). Despite the absence of a strong coevolution

577 between initiators and relaxases, there is a clear tendency to find stable

24

578 backbones Rep_2-MOB$_{V1}$ and Rep_1-MOB$_{V1}$ differentially adapted to different

579 taxonomic orders.

580

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

582

583   Rolling circle replication in ssDNA coliphages and in plasmids requires

584 the existence of origins that are involved in lagging strand synthesis (Kornberg

585 and Baker, 1992; Novick, 1998). When RCR-plasmids were discovered (te

586 Riele et al., 1986a; te Riele et al., 1986b), it was apparent that the ssDNA

587 intermediates should harbor signals for the conversion of the ssDNA to plasmid

588 dsDNA. This was first reported for plasmid pT181 (Gruss et al., 1987) and soon

589 afterwards for the pMV158-derivative plasmid pLS1 (del Solar et al., 1987). In

590 these two plasmids, it was shown that their *ssoA*s were located in non-coding

591 200-300 base-pair-long regions that have the potential to generate one or

592 several secondary structures on ssDNA. These signals were orientation-

593 dependent (this means that they were functional only when placed in the

594 displaced strand). Deletion of the region encompassing the *ssoA* led to

595 accumulation of ssDNA intermediates and to plasmid instability, but the

596 plasmids were still able to replicate (del Solar et al., 1987; Dempsey et al.,

597 1995; Gruss et al., 1987; Kramer et al., 1995; Murray; Seegers et al., 1995).

598 This finding led to the hypothesis that alternative, albeit less efficient, *sso*s could

599 replace the genuine conversion signal (del Solar et al., 1987; Kramer et al.,

600 1998a; Meijer et al., 1995a; Meijer et al., 1995b). Although accumulation of

601 ssDNA intermediates and plasmid unstable inheritance by VGT were thought to

602 be related phenomena (Meijer et al., 1995a; Meijer et al., 1995b), cloning of

603 signals that allowed an efficient ssDNA → dsDNA conversion, at least in

604 plasmid pLS1, did not lead to stable plasmid inheritance (Hernández-Arriaga et

605 al., 2000). Apart from *ssoA*, other three *sso* types have been described based

606 on sequence similarity and structure analysis (Khan, 1996; Khan, 2000; Kramer

607 et al., 1998a): *ssoU* (pUB110, pMV158), *ssoT* (pTA1060, pBAA1), and *ssoW*

608 (pWV01). A fifth type was described in plasmid pM4 from *Lactobacillus*

609 *plantarum*, which exhibited no significant sequence or structural similarity with

610 any of the four classical *sso* (Yin et al., 2009). Plasmids pMV158 and pRW35

611 have the unusual feature of harboring two *sso*, *ssoU* and *ssoA*. Out of the two

612 origins, the former has a more complex structure than the latter (Figure 7).

613 Two regions were mapped in the *ssoA*: the recombination *s*ite B (RS$_B$)

614 and a 6-nucleotide consensus sequence (CS-6) (del Solar et al., 1987; Gruss et

615 al., 1987). These two regions acted as efficient signals only in ssDNA

616 configuration (replicative intermediates). In these molecules, the *ssoA* would

617 adopt a long stem-loop structure where the RS$_B$ would be located at the stem

618 and the CS-6 in the loop of the hairpin (Figure 7). Biochemical and genetic

619 analyses demonstrated that RS$_B$ was recognized as the binding site of the host

620 RNA polymerase (RNAP), thus acting as ssDNA promoter (Kramer et al., 1997).

621 These promoters were shown to generate within paired secondary structures on

622 ssDNA molecules and harbor sequences that are recognized by the host RNAP

623 to initiate lagging strand synthesis (Glucksmann-Kuis et al., 1992; Masai and

624 Arai, 1997). From the *ssoA*$_{pMV158}$ promoter, the RNAP synthesized a short 20

625 nt-long primer RNA (pRNA) that stopped at the CS-6 sequence, which acted as

626 a transcription terminator (Kramer et al., 1997). The pRNA was processed by

26

627 DNA polymerase I and was proposed to be elongated by the host DNA

628 polymerase III to finish the lagging DNA strand synthesis (Kramer et al., 1998b).

629   Efficient ssDNA → dsDNA conversion is also needed in the conjugative

630 process. Within the recipient cell the transferred ssDNA should be converted

631 into dsDNA either prior or after the circularization of the T-DNA by a strand-

632 transfer reaction. During transfer, a fast conversion of ssDNA intermediates

633 would be critical to finish the process, since no proteins essential for vegetative

634 replication and control would be synthesized in the recipient cell until the first

635 dsDNA plasmid copy is generated (Lorenzo-Díaz and Espinosa, 2009). Despite

636 its importance, the synthesis of the lagging strand in the recipient cell remains

637 an unresolved issue. For plasmids F and Collb-P9 single-stranded promoters

638 were identified in the leading strand region, which is the one that enters first into

639 the recipient cell, and that includes genes that promote the establishment of the

640 incoming plasmid (Bates et al., 1999; Masai and Arai, 1997; Nasim et al., 2004).

641   We have explored the *sso* diversity and distribution among MOB$_{V1}$

642 plasmids. Based on the *sso*s previously identified by homology or characterized

643 by *in vivo* and/or *in vitro* approaches, we annotated 35 *sso* elements in 33 out of

644 93 plasmids in Table 1. This limited number is due to the little overall homology

645 among *sso*s, even those of the same type. All the reported *sso* types were

646 found: *ssoT* (13 plasmids), *ssoA* (11 plasmids), *ssoU* (6 plasmids) and *ssoW* (4

647 plasmids). Two plasmids (pMV158 and pRW35) contained two *sso*s (*ssoA* and

648 *ssoU*). No new members arose in our search of the new type of *sso* reported for

649 plasmid pM4 (Zhai et al., 2009). Curiously enough, we have observed that while

650 *ssoA* and *ssoW* are generally found downstream and close to the *mob* gene,

651 the *ssoU* is always upstream with respect to the *oriT* sequence. In the case of

652   the *ssoT* element we have found that it can be located upstream (n=5) or

653   downstream (n=8) of the MOB module.

654   It has been demonstrated that the *sso* type is one of the key elements in

655   determining the host range of a plasmid: whereas *ssoU* and *ssoT* support a

656   wide-host range, *ssoA* and *ssoW* seem to evolve in the other direction, working

657   efficiently only in their natural hosts *in vivo* (Khan, 2005). From the analysis

658   abovementioned, it seems apparent that highly related plasmids do not always

659   contain the same type of *sso*, whereas distantly related do. These facts

660   complicate even more the picture of the putative host range a MOB_{V1} plasmid can

661   reach. Whether the location of the *sso* affects the efficiency of a given plasmid

662   for replication and/or conjugative transfer and, consequently, its host range has

663   not been explored yet.

664

665   **7. Conclusions and perspectives**

666

667   RCR-plasmids represent a pool of genetic information that is shared by

668   many bacteria. We have found them in Lactobacillales, Bacillales, Mollicutes,

669   etc. Among the bacteria that host them, we have found several of the most

670   relevant G+ pathogens, namely *Staphylococcus aureus*, *S. pneumoniae*, and

671   *Enterococcus faecalis*. These three bacterial species are important for human

672   health because they: i) exhibit very high rates of human morbidity and mortality;

673   ii) are the cause of many health-care associated infections; iii) have acquired

674   elevated resistance to antibiotics; iv) play an important role as reservoir of AbR

675   and virulence genes, and v) carry conjugative or mobilizable broad host-range

676   plasmids, thus contributing to the spread of resistances. Infections caused by

677   the above G+ bacteria represent, in addition to a high economic impact, a threat

678   to hospitals, children, elder population and immuno-compromised people

679   (Jones, 2001). Furthermore, plasmid-encoded genes related not only to AbR

680   but also to replication and/or transfer, are also found within the ICEs. However,

681   and in spite of the relevance of these bacteria, very little is known on the

682   genetics and biochemistry of the transfer functions of the RCR-plasmids studied

683   here with the exception of the streptococcal plasmid pMV158 (reviewed in

684   (Espinosa, 2013; Fernández-López et al., 2014) and the staphylococcal plasmid

685   pC221 (Caryl et al., 2004; Caryl and Thomas, 2006; Smith and Thomas, 2004).

686   And even in these two plasmids there is little, if any, information on the

687   interactions of the relaxases encoded by pMV158 or pC221 with the machinery

688   provided by auxiliary plasmids (Arends et al., 2013; Goessweiner-Mohr et al.,

689   2013; Grohmann et al., 2003).

690   In addition to the above, it would be interesting to explore how the

691   conjugative transfer is regulated in plasmids that contain more than one MOB_V

692   cassette as well as the level of cross-recognition between the MobM-like

693   relaxases and their non-cognate *oriT*s. We have shown that MobM is able to

694   relax supercoiled DNAs from plasmids with *oriT*s that share total (plasmid

695   pUB110) or partial (plasmid pDL287) homology with the *oriT* of pMV158

696   (Fernández-López et al., 2013a), a phenomenon that could play an important

697   role in the plasmid spreading between bacteria in natural environments.

698   Furthermore, the fact that RCR initiators relax DNA in a way highly similar to

699   conjugative relaxases and the recent finding of RCR initiators involvement in

700   plasmid mobilization (Lee et al., 2012) open a new research field in the

701   conjugative transmission of RCR plasmids.

702

703

Acknowledgements

713

8. References

Alonso, J. C., et al., 1988. Functional analysis of the leading strand replication origin of plasmid pUB110 in *Bacillus subtilis*. Nucl. 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. Nucl. 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. Nucl. 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. Nature 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, 1359-65.

Bates, S., et al., 1999. Expression of leading region genes on Incl 1 plasmid Collb-P9: genetic evidence for single-stranded DNA transcription. Microbiology. 145, 2665-2662.

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.

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

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

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

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

759   Cascales, E., et al., 2013. DNA substrate-induced activation of the *Agrobacterium*
760   VirB/VirD4 Type IV Secretion System. J. Bacteriol. 195**,** 2691-2704.

761   Clewell, D. B., et al., 2014. Extrachromosomal and mobile elements in Enterococci:
762   transmission, maintenance, and epidemiology. In: C. D. Gilmore MS, Ike Y, et
763   al., editors., (Ed.), Enterococci: From Commensals to Leading Causes of Drug
764   Resistant Infection. Massachusetts Eye and Ear Infirmary, Boston.

765   Coffey, A., et al., 1994. Nucleotide sequence and structural organization of the small
766   broad-host-range plasmid pCI411 from *Leuconostoc lactis*. Microbiology. 140**,**
767   2263-2269.

768   Crooks, G. E., et al., 2004. WebLogo: a sequence logo generator. Genome Res. 14**,**
769   1188-1190.

770   Chandler, M., et al., 2013. Breaking and joining single-stranded DNA: the HUH
771   endonuclease superfamily. Nature Rev. Microbiol. 11**,** 625-538.

772   Chang, T.-L., et al., 2000. Role of individual monomers of a dimeric initiator protein in
773   the initiation and termination of plasmid rolling circle replication. J. Biol. Chem.
774   275**,** 13529-13534.

775   Chen, I., et al., 2005. The ins and outs of DNA transfer in bacteria. Science. 310**,** 1456-
776   1460.

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

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

781   de la Cruz, F., et al., 2010. Conjugative DNA metabolism in Gram-negative bacteria.
782   FEMS Microbiol. Rev. 34**,** 18-40.

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

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

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

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

792 del Solar, G., et al., 2002. A genetically economical family of plasmid-encoded
793   transcriptional repressors in control of plasmid copy number. J. Bacteriol. 184,
794   4943-4951.

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

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

800 Dempsey, L. A., Dubnau, D. A., 1989. Identification of plasmid and *Bacillus subtilis*
801   chromosomal recombination sites used for pE194 integration. J. Bacteriol. 171,
802   2856-2865.

803 Dempsey, L. A., et al., 1995. Localization of the start sites of lagging-strand replication
804   of rolling circle plasmids from Gram-positive bacteria. Mol. Microbiol. 15, 679-
805   687.

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

808 Draper, O., et al., 2005. Site-specific recombinase and integrase activities of a
809   conjugative relaxase in recipient cells. Proc. Natl. Acad. Sci USA. 102, 16385-
810   16390.

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

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

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

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

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

824 Fernández-López, C., et al., 2013a. Nicking activity of the pMV158 MobM relaxase on
825   cognate and heterologous origins of transfer. Plasmid. 120-130.
826 Fernández-López, C., et al., 2013b. Functional properties and structural requirements
827   of the plasmid pMV158-encoded MobM relaxase domain J. Bacteriol. 195,
828   3000-3008.
829 Francia, M. V., et al., 2004. A classification scheme for mobilization regions of bacterial
830   plasmids. FEMS Microbiol. Rev. 28, 79-100.
831 Frost, L., et al., 2005. Mobile genetic elements: the agents of open source evolution.
832   Nature Rev. Microbiol. 3, 722-732.
833 Funnell, B. E., Phillips, G. J. Eds.), 2004. Plasmid Biology. ASM Press, Washington,
834   D.C.
835 Garcillán-Barcia, M. P., et al., 2011. Identification of bacterial plasmids based on
836   mobility and plasmid population biology. FEMS Microbiol. Rev. 35, 936-956.
837 Garcillán-Barcia, M. P., et al., 2009. The diversity of conjugative relaxases and its
838   application in plasmid classification. FEMS Microbiol. Rev. 33, 657-687.
839 Garcillán-Barcia, M. P., et al., 2007. Conjugative transfer can be inhibited by blocking
840   relaxase activity within recipient cells with intrabodies. Mol. Microbiol. 63, 404-
841   416.
842 Gennaro, M. L., et al., 1987. A site-specific recombination function in *Staphylococcus*
843   *aureus* plasmids. J. Bacteriol. 169, 2601-2610.
844 Glucksmann-Kuis, M. A., et al., 1992. Specific sequences and a hairpin structure in the
845   template strand are required for N4 virion RNA polymerase promoter
846   recognition. Cell. 70, 491-500.
847 Goessweiner-Mohr, N., et al., 2013. Conjugative type IV secretion systems in Gram-
848   positive bacteria. Plasmid. 70, 289-302.
849 Gomis-Rüth, F. X., Coll, M., 2006. Cut and move: protein machinery for DNA
850   processing in bacterial conjugation. Curr. Opin. Struct. Biol. 16, 744-752.
851 Gomis-Ruth, F. X., et al., 2004. Coupling factors in macromolecular type-IV secretion
852   machineries. Curr Pharm Des. 10, 1551-65.
853 Goursot, R., et al., 1982. Plasmids from *Staphylococcus aureus* replicate in yeast
854   *Saccharomyces cerevisiae.* Nature. 298, 488-490.
855 Grohmann, E., et al., 1999. Mobilisation of the streptococcal plasmid pMV158:
856   interactions of MobM protein with its cognate *oriT* DNA region. Mol. Gen.
857   Genet. 261, 707-715.
858 Grohmann, E., et al., 2003. Conjugative plasmid transfer in Gram-positive bacteria.
859   Microbiol. Mol. Biol. Rev. 67, 277-301.

860 Grohmann, E., et al., 1997. Determination of specific DNA strand discontinuities with
861   nucleotide resolution in exponentially growing bacteria harbouring rolling circle-
862   replicating plasmids. FEMS Microbiol. Lett. 152, 363-369.

863 Gruss, A. D., Ehrlich, S. D., 1989. The family of highly interrelated single-stranded
864   deoxyribonucleic avid plasmids. Microbiol. Rev. 53, 231-241.

865 Gruss, A. D., et al., 1987. Functional analysis of a palindromic sequence required for
866   normal replication of several staphylococcal plasmids. Proc. Natl. Acad. Sci.
867   USA. 84, 2165-2169.

868 Guasch, A., et al., 2003. Recognition and processing of the origin of transfer DNA by
869   conjugative relaxase TrwC. Nat. Struct. Biol. 10, 1002-1010.

870 Guglielmini, J., et al., 2011. The repertoire of ICE in prokaryotes underscores the unity,
871   diversity, and ubiquity of conjugation. PLoS Genet. 7, e1002222.

872 Guzmán, L., Espinosa, M., 1997. The mobilization protein, MobM, of the streptococcal
873   plasmid pMV158 specifically cleaves supercoiled DNA at the plasmid *oriT*. J.
874   Mol. Biol. 266, 688-702.

875 Halbedel, S., Stülke, J., 2007. Tools for the genetic analysis of *Mycoplasma*. Int. J.
876   Med. Microbiol. 297, 37-44.

877 Hamilton, C. M., et al., 2000. TraG from RP4 and TraG and VirD4 from Ti plasmids
878   confer relaxosome specificity to the conjugal transfer system of pTiC58. J.
879   Bacteriol. 182, 1541-1574.

880 Harwood, C. R., et al., 1983. Nucleotide sequence of a *Bacillus pumilus* gene
881   specifying chloramphenicol acetyltransferase. Gene. 24, 163-169.

882 Hernández-Arriaga, A. M., et al., 2000. A functional lagging strand origin does not
883   stabilize plasmid pMV158 inheritance in *Escherichia coli*. Plasmid. 43, 49-58.

884 Horinouchi, S., Weisblum, B., 1982. Nucleotide sequence and functional map of
885   pC194, a plasmid that specifies inducible chloramphenicol resistance. J.
886   Bacteriol., 815-825.

887 Inamine, J. M., et al., 1990. Evidence that UGA is read as a tryptophan codon rather
888   than as a stop codon by *Mycoplasma pneumoniae*, *Mycoplasma genitalium*,
889   and *Mycoplasma gallisepticum*. J. Bacteriol. 172, 504-506.

890 Jin, R., Novick, R. P., 2001. Role of the double-strand origin cruciform in pT181
891   replication. Plasmid. 46, 95-105.

892 Jones, R. N., 2001. Resistance patterns among nosocomial pathogens: trends over the
893   past few years. Chest. 119(2 Suppl), 397S-404S.

894 Kent, B. N., et al., 2012. Development of a novel plasmid as a shuttle vector for
895   heterologous gene expression in *Mycoplasma yeatsii*. J. Microbiol. Methods.
896   91, 121-127.

897   Khan, S. A., 1996. Mechanism of replication and copy number control of plasmids in
898   Gram-positive bacteria. In: J. K. Setlow, (Ed.), Genetic Engineering. vol. 18.
899   Plenum Press, New York.

900   Khan, S. A., 1997. Rolling-circle replication of bacterial plasmids. Microbiol. Mol. Biol.
901   Rev. 61, 442-455.

902   Khan, S. A., 2000. Plasmid rolling-circle replication: recent developments. Mol.
903   Microbiol. 37, 477-484.

904   Khan, S. A., 2003. DNA-protein interactions during the initiation and termination of
905   plasmid pT181 rolling-circle replication. Prog. Nucleic Acid Res. Mol. Biol. 75,
906   113-137.

907   Khan, S. A., 2005. Plasmid rolling-circle replication: highlights of two decades of
908   research Plasmid. 53, 126-136.

909   Khan, S. A., et al., 1981. Replication of plasmid pT181 in vitro: requirement for a
910   plasmid-encoded product. . Proc. Natl. Acad. Sci. USA. 78, 4902-4906.

911   King, K. W., Dybvig, K., 1992. Nucleotide sequence of *Mycoplasma mycoides*
912   subspecies *mycoides* plasmid pKMK1. Plasmid. 86-91.

913   Koehler, T. M., Thorne, C. B., 1987. *Bacillus subtilis* (natto) plasmid pLS20 mediates
914   interspecies plasmid transfer. Journal of Bacteriology. 169, 5271-5278.

915   Koepsel, R. R., et al., 1986. Sequence-specific interaction between the replication
916   initiator protein of plasmid pT181 and its origin of replication. Proc. Natl. Acad.
917   Sci. USA. 83, 5484-5488.

918   Koepsel, R. R., et al., 1985. The replication initiator protein of plasmid pT181 has
919   sequence-specific endonuclease and topoisomerase-like activivties. Proc. Natl.
920   Acad. Sci. USA. 82, 6845-6849.

921   Kornberg, A., Baker, T., 1992. DNA replication. Freeman and Co. New York.

922   Kramer, M. G., et al., 1995. Lagging-strand origins of the promiscuous plasmid
923   pMV158: physical and functional characterization. Microbiology. 141, 655-662.

924   Kramer, M. G., et al., 1998a. Lagging strand replication of rolling-circle plasmids:
925   specific recognition of the *ssoA*-type origins in different gram-positive bacteria.
926   Proc. Natl. Acad. Sci. USA. 95, 10505-10510.

927   Kramer, M. G., et al., 1999. Characterization of a single-strand origin, *ssoU*, required
928   for broad host range replication of rolling-circle plasmids. Mol. Microbiol. 33,
929   466-475.

930   Kramer, M. G., et al., 1997. Plasmid rolling circle replication: identification of the RNA
931   polymerase-directed primer RNA and requirement of DNA polymerase I for
932   lagging strand initiation. EMBO J. 16, 5784-5795.

933 Kramer, M. G., et al., 1998b. Lagging-strand replication from the *ssoA* origin of plasmid
934 pMV158 in *Streptococcus pneumoniae*: *in vivo* and *in vitro* influences of
935 mutations in two conserved *ssoA* regions. J. Bacteriol. 180**,** 83-89.

936 Lacks, S. A., et al., 1986. Identification and analysis of genes for tetracycline resistance
937 and replication functions in the broad-host-range plasmid pLS1. J. Mol. Biol.
938 192**,** 753-765.

939 Lanka, E., Wilkins, B. M., 1995. DNA processing reactions in bacterial conjugation.
940 Annu. Rev. Biochem. 64**,** 141-169.

941 Lederberg, J., Tatum, E., 1946. Gene recombination in *E. coli*. Nature. 158**,** 558.

942 Lederberg, J., Tatum, E., 1953. Sex in bacteria; genetic studies, 1945-1952. Science.
943 118**,** 169-175.

944 Lee, C. A., et al., 2012. The *Bacillus subtilis* conjugative transposon ICEBs1 mobilizes
945 plasmids lacking dedicated mobilization functions. J. Bacteriol. 194**,** 3165-3172.

946 Lorenzo-Díaz, F., et al., 2011. The MobM-relaxase domain of plasmid pMV158:
947 thermal stability and activity upon Mn^{2+}-and DNA specific-binding. Nucleic Acids
948 Res. 39**,** 4315-4329.

949 Lorenzo-Díaz, F., Espinosa, M., 2009. Lagging strand DNA replication origins are
950 required for conjugal transfer of the promiscuous plasmid pMV158. J. Bacteriol.
951 191**,** 720–727.

952 Lorenzo-Díaz, F., et al., 2012. Autoregulation of the synthesis of the MobM relaxase
953 encoded by the promiscuous plasmid pMV158. J. Bacteriol. 194**,** 1789-1799.

954 Lucas, M., et al., 2010. Relaxase DNA binding and cleavage are two distinguishable
955 steps in conjugative DNA processing that involve different sequence elements
956 of the *nic* site. J. Biol. Chem. 285**,** 8918-8926.

957 Llosa, M., et al., 2002. Bacterial conjugation: a two-step mechanism for DNA transport.
958 Mol. Microbiol. 45**,** 1-8.

959 Llosa, M., et al., 2009. Bacterial type IV secretion systems in human disease. Mol.
960 Microbiol. 73**,** 141-151.

961 Machón, C., et al., 2010. RepD-mediated recruitment of PcrA helicase at the
962 *Staphylococcus aureus* pC221 plasmid replication origin, oriD. Nucleic Acids
963 Res. 38**,** 1874-1888.

964 Masai, H., Arai, K., 1997. F*rpo*: a novel single-stranded DNA promoter for transcription
965 and for primer RNA synthesis of DNA replication. Cell. 89**,** 897-907.

966 Matilla, I., et al., 2010. The conjugative DNA translocase TrwB is a structure-specific
967 DNA-binding protein. J. Biol. Chem. 285**,** 17537-17544.

968 McKenzie, T., et al., 1986. The nucleotide sequence of pUB110: some salient features
969 in relation to replication and its regulation. Plasmid. 15**,** 93-103.

970 Meijer, W. J. J., et al., 1995a. Effects of the generation of single-stranded DNA on the
971 maintenance of plasmid pMV158 and derivatives in different *Bacillus subtilis*
972 strains. Plasmid. 33, 79-89.

973 Meijer, W. J. J., et al., 1995b. Effects of the generation of single-stranded DNA on the
974 maintenance of plasmid pMV158 and derivatives in *Lactococcus lactis*.
975 Plasmid. 33, 91-99.

976 Moscoso, M., et al., 1997. Initiation of replication of plasmid pMV158: mechanisms of
977 DNA strand transfer reactions mediated by the initiator RepB protein. J. Mol.
978 Biol. 268, 840-856.

979 Murray, R. W., R.R. Koepsel, and S.A. Khan. 1989. Synthesis of single-stranded
980 plasmid pT181 DNA *in vitro*. J. Biol. Chem. 264:1051-1057.

981 Naglich, J. G., Andrews, R. E. J., 1988. *Tn916*-dependent conjugal transfer of pC194
982 and pUB110 from *Bacillus subtilis* into *Bacillus thuringiensis* subsp. *israelensis*.
983 Plasmid. 20, 113-126.

984 Nasim, M. T., et al., 2004. The activity of a single-stranded promoter of plasmid ColIb-
985 P9 depends on its secondary structure. Mol Microbiol. 53, 405-417.

986 Noirot, P., et al., 1990. Initiation of rolling circle replication in pT181 plasmid: initiator
987 protein enhances cruciform extrusion at the origin. Proc. Natl. Acad. Sci. USA.
988 87, 8560-8564.

989 Nordström, K., et al., 1984. Control of replication of bacterial plasmids: genetics,
990 molecular biology, and physiology of the plasmid R1 system. Plasmid. 12, 71-
991 90.

992 Novick, R. P., 1989. Staphylococcal plasmids and their replication. Annu. Rev.
993 Microbiol. 43, 537-565.

994 Novick, R. P., 1998. Contrasting lifestyles of rolling-circle phages and plasmids. Trends
995 Biochem. Sci. 23, 434-438.

996 Novick, R. P., et al., 1984. Staphylococcal plasmid cointegrates are formed by host-
997 and phage-mediated general *rec* systems that act on short regions of
998 homology. Mol. Gen. Genet. 195, 374-377.

999 Ochman, H., et al., 2000. Lateral gene transfer and the nature of bacterial innovation.
1000 Nature. 405, 229-304.

1001 Osborn, M., et al., 2000. The evolution of bacterial plasmids. In: C.M.Thomas, (Ed.),
1002 The horizontal gene pool. Harwood Academic Publishers, Amsterdam.

1003 Oskam, L., et al., 1991. The large *Bacillus* plasmid pTB19 contains two integrated
1004 rolling-circle plasmids carrying mobilization functions. Plasmid. 26, 30-39.

1005 Pansegrau, W., Lanka, E., 1996a. Enzymology of DNA strand transfer by conjugative
1006 mechanisms. Prog. Nucl. Acid Res. Mol. Biol. 54, 197-251.

1007 Pansegrau, W., Lanka, E., 1996b. Mechanisms of initiation and termination reactions in
1008   conjugative DNA processing: independence of tight substrate binding and
1009   catalytic activity of relaxase (TraI) of IncPα plasmid RP4. J. Biol. Chem. 271,
1010   13068-13076.

1011 Paulson, J., Ehrenberg, M., 1998. Trade-off between segregational stability and
1012   metabolic burden: a mathematical model of plasmid ColE1 replication control. J.
1013   Mol. Biol. 279, 73-88.

1014 Pepper, K., et al., 1988. Heterogeneity of chromosomal genes encoding
1015   chloramphenicol resistance in streptococci. Plasmid. 19, 71-74.

1016 Priebe, S. D., Lacks, S. A., 1989. Region of the streptococcal plasmid pMV158
1017   required for conjugative mobilization. J. Bacteriol. 171, 4778-4784.

1018 Projan, S. J., Novick, R. P., 1988. Comparative analysis of five related staphylococcal
1019   plasmids. Plasmid. 19, 203-221.

1020 Puyet, A., et al., 1988. Identification of the origin and direction of replication of the
1021   broad-host-range plasmid pLS1. Nucleic Acids Res. 16, 115-133.

1022 Reyes-Lamothe, R., et al., 2014. High-copy bacterial plasmids diffuse in the nucleoid-
1023   free space, replicate stochastically and are randomly partitioned at cell division.
1024   Nucleic Acids Res. 42, 1042-1051.

1025 Ruiz-Masó, J. A., et al., 2006. Genetic and biochemical characterization of the
1026   *Streptococcus pneumoniae* PcrA helicase and its role in plasmid rolling circle
1027   replication. J. Bacteriol. 188, 7416-7425.

1028 Ruiz-Masó, J. A., et al., 2007. Interactions between the RepB initiator protein of
1029   plasmid pMV158 and two distant DNA regions within the origin of replication.
1030   Nucleic Acids Res. 35, 1230-1244.

1031 Seegers, J. F. M. L., et al., 1995. Structural and functional analysis of the single-strand
1032   origin of replication from the lactococcal plasmid pWV01. Mol. Gen. Genet. 249,
1033   43-50.

1034 Selinger, L. B., et al., 1990. Mobilization of closely related plasmids pUB110 and
1035   pBC16 by *Bacillus* plasmid pXO503 requires trans-acting open reading frame
1036   beta. J. Bacteriol. 172, 3290-3297.

1037 Seufert, W., et al., 1988. A novel replicon occurring naturally in *Escherichia coli* is a
1038   phage-plasmid hybrid. EMBO J. 7, 4005-4010.

1039 Showsh, S. A., Andrews Jr, R. E., 1999. Analysis of the requirement for a pUB110 *mob*
1040   region during Tn*916*-dependent mobilization. Plasmid. 41, 179-186.

1041 Smillie, C., et al., 2010. Mobility of plasmids. Microbiol. Mol. Biol. Rev. 74, 434-452.

1042 Smith, M. C. A., Thomas, C. D., 2004. An accessory protein is required for relaxosome
1043   formation by small staphylococcal plasmids. J. Bacteriol. 186, 3363-3373.

1044 Soultanas, P., et al., 1999. Plasmid replication initiator protein RepD increases the
1045   processivity of PcrA DNA helicase. Nucleic Acids Res. 27**,** 1421-1428.

1046 Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic
1047   analyses with thousands of taxa and mixed models. Bioinformatics. 22**,** 2688-
1048   2690.

1049 Szpirer, C. Y., et al., 2001. Mobilization function of the pBHR1 plasmid, a derivative of
1050   the broad-host-range plasmid pBBR1. J. Bacteriol. 183**,** 2101-2110.

1051 te Riele, H., et al., 1986a. Are single-stranded circles intermediates in plasmid DNA
1052   replication? EMBO J. 5**,** 631 - 637.

1053 te Riele, H., et al., 1986b. Single-stranded plasmid DNA in *Bacillus subtilis* and
1054   *Staphylococcus aureus*. Proc. Natl. Acad. Sci. USA. 83**,** 2541-2545.

1055 Thomas, C. M., 2000. The horizontal gene pool. Harwood Academic Publishers,
1056   Amsterdam, p. 419.

1057 Thomas, C. M., Nielsen, K. M., 2005. Mechanisms of, and barriers to, Horizontal Gene
1058   Transfer between bacteria. Nat. Rev. Microbiol. 3**,** 711-721.

1059 Thomas, J., et al., 2013. A conserved helicase processivity factor is needed for
1060   conjugation and replication of an Integrative and Conjugative Element. PLoS
1061   Genet. 9**,** e1003198.

1062 Turgeon, N., Moineau, S., 2001. Isolation and characterization of a *Streptococcus*
1063   *thermophilus* plasmid closely related to the pMV158 family. Plasmid. 45**,** 171-
1064   183.

1065 van der Lelie, D., et al., 1989. Similarity of minus origins of replication and flanking
1066   open reading frames of plasmids pUB110, pTB913 and pMV158. Nucleic Acids
1067   Res. 17**,** 7283-7294.

1068 van der Lelie, D., et al., 1990. Conjugal mobilization of streptococcal plasmid pMV158
1069   between strains of *Lactococcus lactis* subsp. *lactis*. J. Bacteriol. 172**,** 47-52.

1070 Walker, A., 2012. Welcome to the plasmidome. Nature Rev. Microbiol. 10**,** 379.

1071 Waters, V. L., Guiney, D. G., 1993. Processes at the nick region link conjugation, T-
1072   DNA transfer and rolling-circle replication. Mol. Microbiol. 9**,** 1123-1130.

1073 Wellington, E. M., et al., 2013. The role of the natural environment in the emergence of
1074   antibiotic resistance in Gram-negative bacteria. Lancet Infect. Dis. . 13**,** 155-
1075   165.

1076 Wilkins, B., Lanka, E., 1993. DNA processing and replication during plasmid transfer
1077   between Gram-negative bacteria. Plenum Press, New York.

1078 Woodbury, R. L., et al., 2008. Plasmid-borne *erm*(T) from invasive, macrolide-resistant
1079   *Streptococcus pyogenes* strains. Antimicrob. Agents Chemother. 52**,** 1140-
1080   1143.

1081 Yamao, F., et al., 1985. UGA is read as tryptophan in *Mycoplasma capricolum*. Proc
1082   Natl Acad Sci U S A. 82**,** 2306-2309.
1083 Yasukawa, H., et al., 1991. Rolling-circle replication of the plasmid pKYM isolated from
1084   a Gram-negative bacterium. Proc. Natl. Acad. Sci. USA. 88**,** 10282-10286.
1085 Yin, S., et al., 2009. Functional analysis of the plasmid pM4 replicon from *Lactobacillus*
1086   *plantarum* M4: Determination of the minimal replicon and functionality
1087   identification of the putative *sso*. Plasmid. 62**,** 166-171.
1088 Zechner, E. L., et al., 2000. Conjugative-DNA transfer processes. In: C. M. Thomas,
1089   (Ed.), The horizontal gene pool. Harwood Academic publishers, Amsterdam.
1090 Zhai, Z., et al., 2009. Characterization of a novel rolling-circle replication plasmid pYSI8
1091   from *Lactobacillus sakei* YSI8. Plasmid. 62**,** 30-34.
1092 Zhang, W., et al., 2012. Type-IVC Secretion System: a novel subclass of Type IV
1093   Secretion System (T4SS) common existing in Gram-positive genus
1094   *Streptococcus*. PLoS ONE. 7**,** e46390.
1095 Zhou, H., et al., 2010. Characterization of a rolling-circle replication plasmid pXY3 from
1096   *Lactobacillus plantarum* XY3. Plasmid. 64**,** 36-40.
1097
1098
1099
1100

1101 **Legends to the figures**

1102

1103 **Figure 1. Genetic organization of representatives of RCR-plasmids.** The

1104 modular organization of pMV158 from *S. agalactiae* is depicted in the upper

1105 part. Genes and the identified promoters are indicated as arrows (arrowheads

1106 pointing to the transcription direction). Negative (-) regulatory elements within

1107 the replication and control (green) and the transfer (blue) modules are indicated.

1108 Below pMV158, other RCR-plasmids are depicted: pADB201 from *Mycoplasma*

1109 *mycoides*, pWV01 from *Lactococcus lactis* and pE194, pT181, pC221, pC194,

1110 and pUB110 from *S. aureus*. The corresponding antibiotic resistance (red

1111 module) is indicated with the following abbreviations: MLS^R, Macrolide /

1112 Lincosamide / Streptogramin B; Tc^R, Tetracycline; Cm^R, Chloramphenicol; Km^R,

1113 Kanamycin. Plasmids sharing similar genetic modules are presented in the

1114 same color and filling. See Supplementary Tables S1 and S2 for detailed

1115 information on the plasmids.

1116

1117 **Figure 2. Distribution of MOB$_{V1}$ plasmids according to their size.** The X

1118 axis was built by using the $\log_{10}$ of plasmid size values. Each bar represents the

1119 abundance of plasmids for a given size range, which is indicated at each side of

1120 the bar. Plasmids encoding genes for metal or antibiotic resistance and/or

1121 bacteriocin production are indicated in dark grey. The rest are indicated in light

1122 grey. Data was obtained from Table 1.

1123

1124 **Figure 3. The origin of transfer (*oriT*) in MOB$_{V1}$ representative plasmids**.

1125 Comparison of the *oriT* sequences located in the pMV158, p1414 and pT181

1126   plasmids, lined up by the position of the *nic* site (boldface letters). The three

1127   overlapping inverted repeats (IR1, IR2 and IR3) are depicted by arrows, and

1128   dashed lines indicate the position of those unpaired bases in the predicted

1129   secondary structures they could form. Grey background in the pMV158-*oriT*

1130   sequence denotes the position of a conserved repeated region. Consensus of

1131   the *oriT* sequences aligned in Supplementary Figure S2 were prepared using

1132   WebLogo (version 2.8.2; (Crooks et al., 2004)).

1133

1134   **Figure 4. Phylogeny of MobM$_{pMV158}$ homologs**. A: Logos of the MOB$_{V1}$

1135   relaxase motifs. MOB$_{V1}$ relaxases were analyzed using WebLogo (version 2.8.2)

1136   (Crooks et al., 2004). B: The 300 N-terminal residues of the MobM relaxase of

1137   plasmid pMV158 were used as query in a PSI-BLAST search (Altschul et al.,

1138   1997) (e-value: 1xE-6 and limited to 100 non-redundant plasmid hits). The

1139   search converged in the sixth iteration. The 300 N-terminal residues of the

1140   homologs were aligned using MUSCLE (Edgar, 2004). The phylogenetic

1141   reconstruction was carried out by maximum likelihood (ML), using RAxML

1142   version 7.2.7 (Stamatakis, 2006). 100 ML trees were executed using the

1143   JTTGAMMA model. 1000 bootstrap trees were then inferred to obtain the

1144   confidence values for each node of the best ML tree. Only bootstrap values >

1145   50% are indicated. The MOB$_{V2}$ relaxase of plasmid pBMYdx (GenBank Acc.

1146   No. NP_981974.1) was used as outgroup. Highly related clusters are

1147   compressed and a prototype member is indicated. The names and features of

1148   all members included in the tree are recorded in Table 1. Clusters grouping

1149   plasmids that encode antibiotic resistance traits are underlined. Vertical bars

1150   delimit clades for which most of their members are hosted either in

1151   Lactobacillales or in Bacillales.

1152

1153   **Figure 5. Distribution of replication initiation protein families in MOB$_{V1}$**

1154   **plasmids.** The percentage of replication initiator families included in Table 1 is

1155   presented. Plasmids with more than one initiator are included in "> 1 replication

1156   initiator". Plasmids with a single initiator are grouped in "Rep_1", "Rep_2" and

1157   "Rep_trans" families (when RCR), or in "Rep_3" and "Others" (non-RCR).
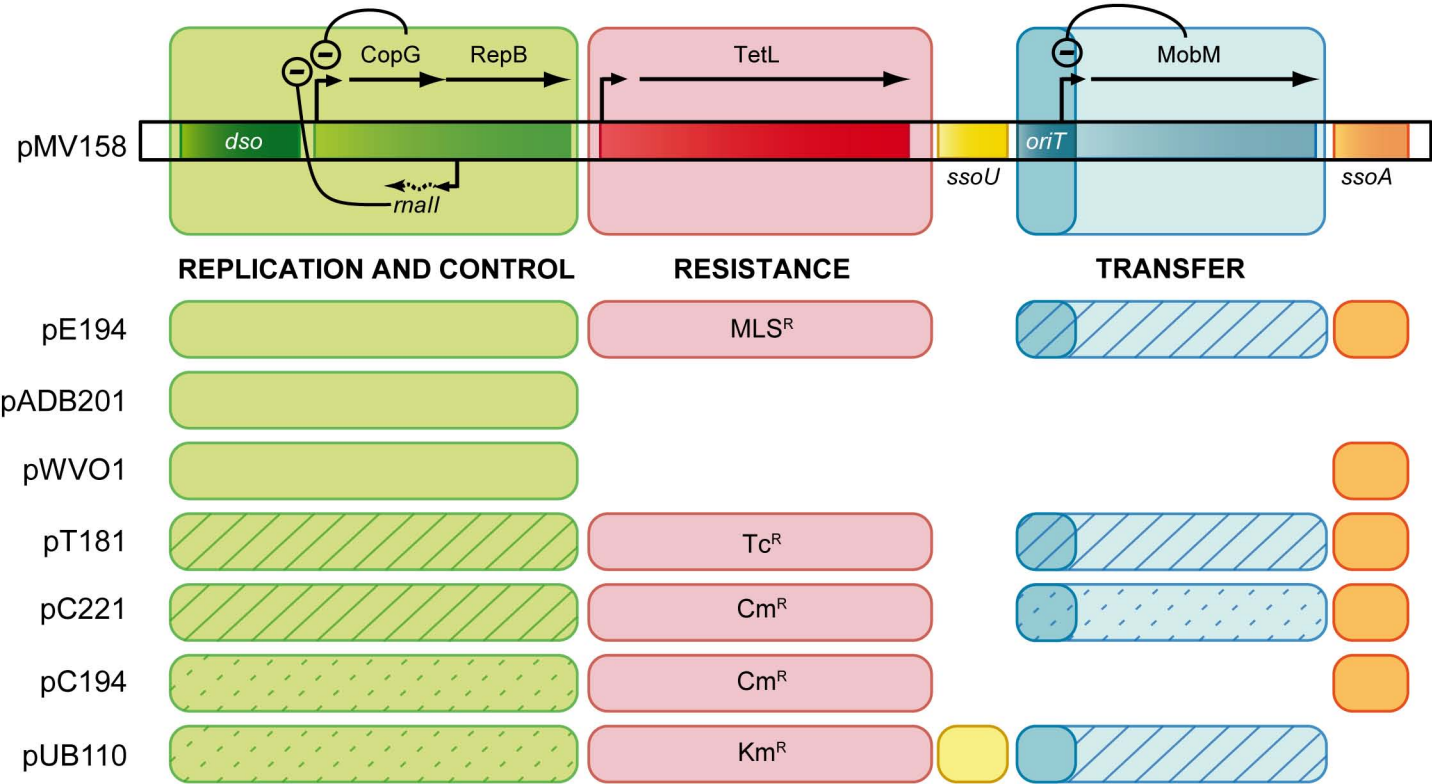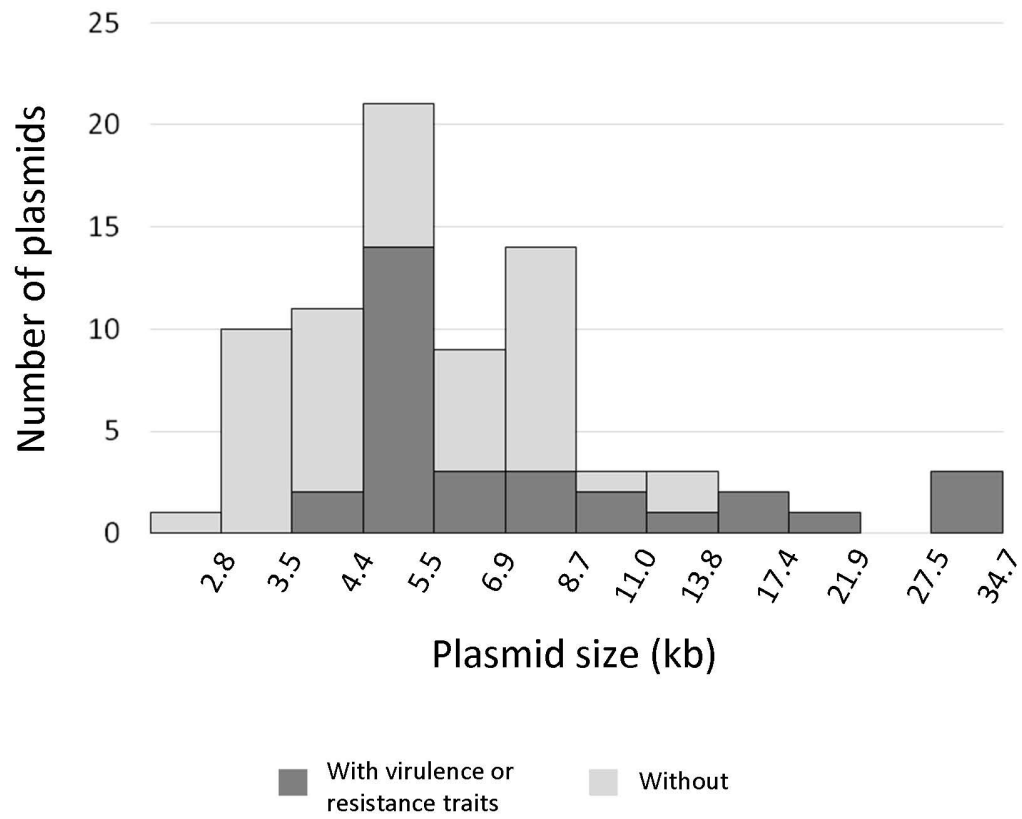
1158

1159   **Figure 6. Phylogeny of the Rep and Mob (MOB$_{V1}$) proteins of relevant RCR-**

1160   **plasmids.** Representative RCR plasmids and MOB$_{V1}$ elements included in

1161   Supplementary Table S2 were used to trace the evolutionary relationships of their

1162   relaxase and replication initiation proteins. The homologs were aligned using

1163   MUSCLE (Edgar, 2004). The phylogenetic reconstruction was carried out by

1164   maximum likelihood (ML), using RAxML version 7.2.7 (Stamatakis, 2006). 100 ML

1165   trees were executed using the JTTGAMMA model. 1000 bootstrap trees were

1166   then inferred to obtain the confidence values for each node of the best ML tree.

1167   Only bootstrap values > 50% are indicated. A: Phylogenetic tree of the N-terminal

1168   300 residues of MOB$_{V1}$ relaxases. Each plasmid is shadowed in grey according to

1169   the RCR initiator subgroup as indicated in the legend of panel B. ǂ indicates an

1170   exceptional MOB$_{V1}$ plasmid, pE33L5, which does not encode an RCR initiator but

1171   a HTH_36 (PF13730) replication initiation protein. Vertical bars delimit clades for

1172   which most of their members are hosted either in Lactobacillales or in Bacillales. ¥

1173   indicates an element not hosted in the taxonomic order indicated by the bars. ¶

1174   indicates that plasmid pMRI_5.2 also encodes a Rep_1 RCR initiator. B:

44

1175   Phylogenetic tree of the RCR initiators. Plasmids pT181 and pC221 were used as

1176   outgroups. A grey color palette was used to indicate clades containing different

1177   RCR initiators families: Rep_1 (PF01446) and Rep_2 (PF01719), as well as the

1178   Rep_trans (PF02486) used to root the tree. *: According to their GenBank

1179   annotated sequences, plasmids pSBO2 and pLFE1 encode truncated MOB_V

1180   relaxases and thus were not included in the MOB phylogeny, neither were the

1181   underlined plasmids (pWV01, pSsal-M18 and pC194) since they do not encode

1182   relaxases. §: pC221 is a mobilizable RCR plasmid, but it encodes a MOB_{P7} instead

1183   of a MOB_V relaxase.

1184

1185   **Figure 7. Predicted secondary structures of the lagging-strand origins of**

1186   **replication *ssoA* and *ssoU*.** The *oriT* and the *mobM* gene of plasmid pMV158

1187   are flanked by two lagging-strand origins of replication (*ssoA* and *ssoU*). $oriT_{pMV158}$

1188   sequence (coordinates 3564-3605 from pMV158; GenBank Acc. No. NC_010096)

1189   is shown at the center of the image. Its three inverted repeats are represented by

1190   arrows and the *nic* site by a vertical arrowhead. Both *sso*s can generate long

1191   hairpin-loop structures that function as 'ssDNA promoters' (Kramer et al., 1999;

1192   Kramer et al., 1997; Masai and Arai, 1997). The RNAP-binding site (RS_B), located

1193   in the base of the hairpin is recognized by the RNAP to synthesize a short pRNA.

1194   A consensus sequence (CS-6), located in the loop of the hairpin, acts as the

1195   termination point for the pRNA synthesis. The pRNA is then used by DNA

1196   polymerase I for limited extension synthesis, followed by replication of the lagging

1197   strand by DNA Pol III. The figure was modified with permission from the American

1198   Society for Microbiology from (Fernández-López et al., 2014). No further

1199   reproduction or distribution is permitted without the prior written permission of

1200   American Society for Microbiology.

pMV158

CopG RepB

dso rnaII

TetL

ssoU

oriT MobM

ssoA

REPLICATION AND CONTROL

RESISTANCE

TRANSFER

pE194

MLS^R

pADB201

pWVO1

pT181

Tc^R

pC221

Cm^R

pC194

Cm^R

pUB110

Km^R

```
                    ┌──────────→  IR1  ←──────────┐
                    ┌──────────→  IR3  ←── ── ───┐
5′-GCACAC ACTTTATGAATATAAAGTATAGTGT GT TATACTTTA
                          ┌──────────→  IR2  ←──────────┐
```

pMV158

```
              ┌──────────→  IR1  ←── ── ───┐
       ┌── ── ──────────→  IR3  ←── ── ── ── ──┐
5′-TCGCACGTCCAAATTTGCCATGGGCATAATTTGGTGTAGTGC GT TACACCAAAGA
                          ┌──────────→  IR2  ←──────────┐
```

p1414

```
                    ┌──────────→  IR1  ←──────────┐
       ┌── ── ── ── ──────→  IR3  ←── ── ───┐
5′-GCACACGTATTAACGACTTATTAAAAATAAGTCTAGTGT GT TAGACTTAAA
                          ┌──────────→  IR2  ←──────────┐
```

pT181

I II III CS-6 IV V

RS_B

DNA

pRNA

ssoU

IR1
IR2
IR3
nic

GCACACACTTTATGAATATAAAGTATAGTGTGTTATACTTTA

oriT

mobM

CS-6

DNA

pRNA

-10

-35

RS_B

ssoA

5′

3′
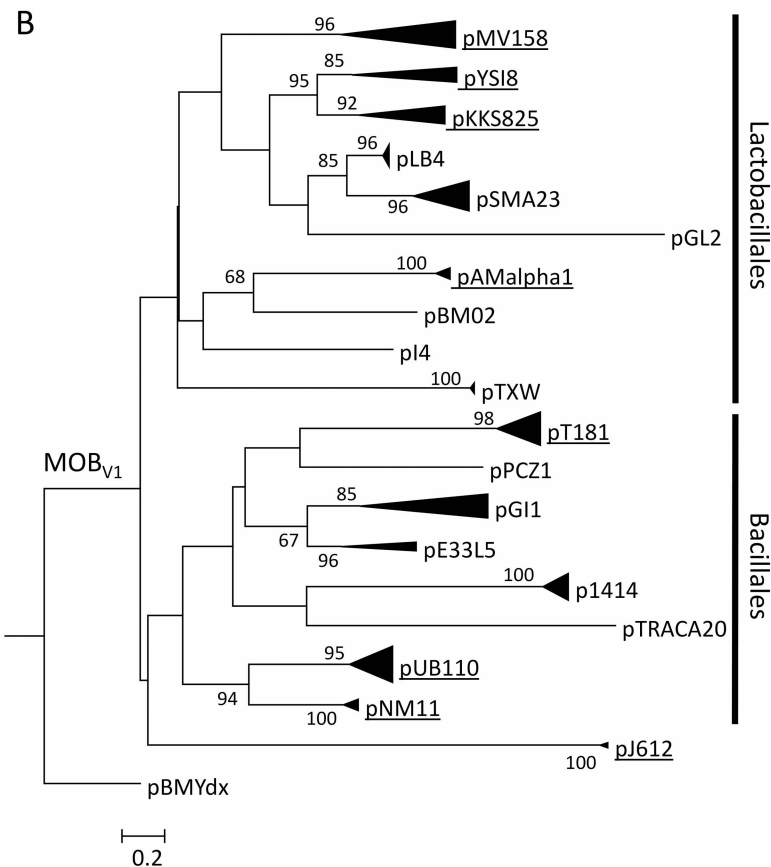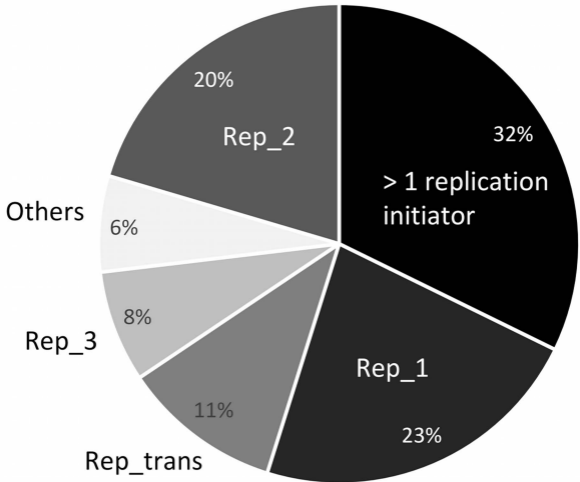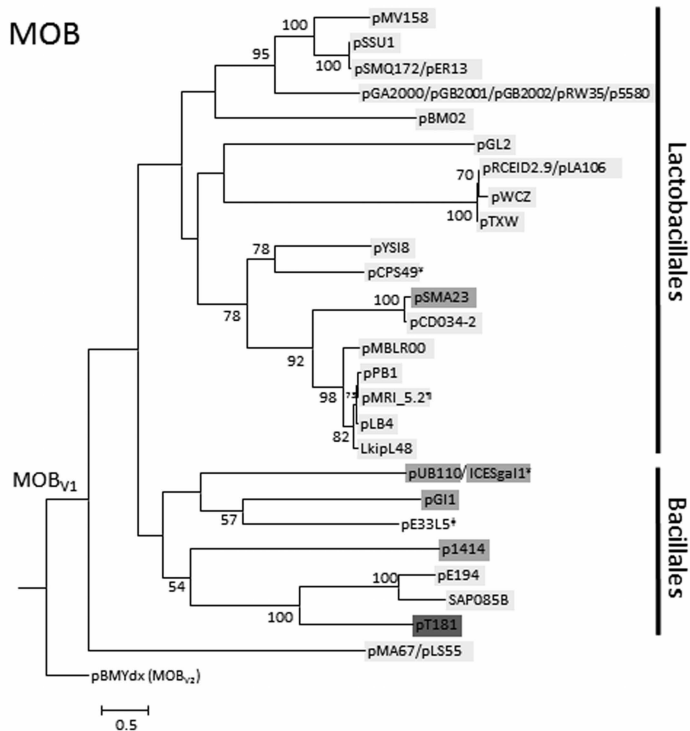
Table 1: MOB$_{V1}$ plasmids

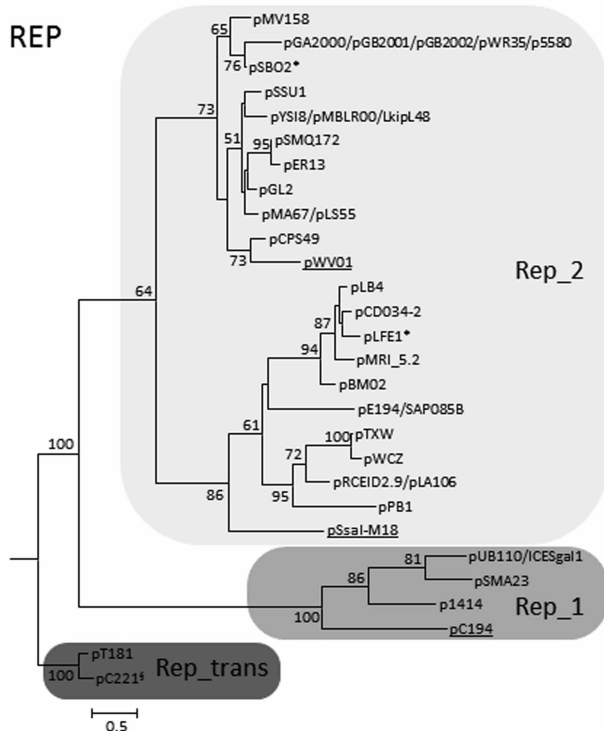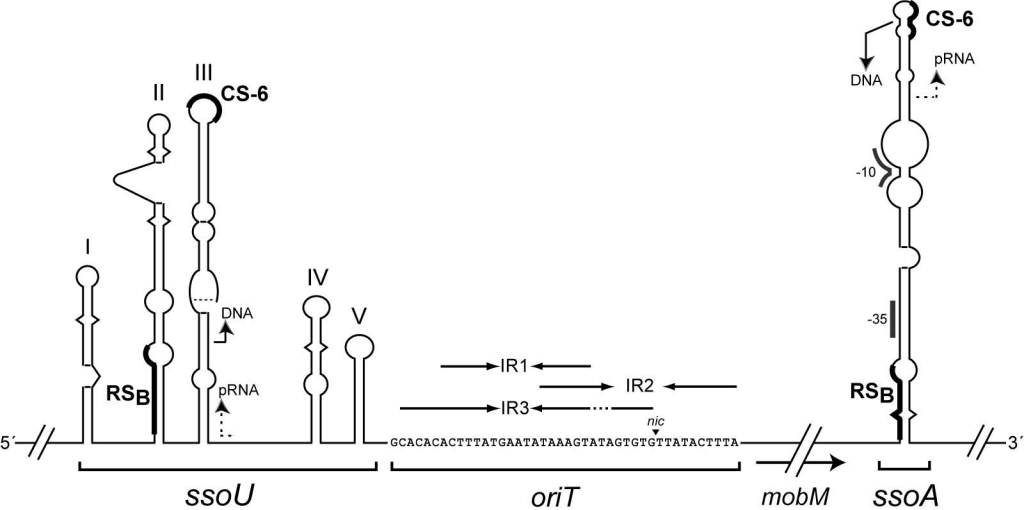| Plasmid name [a]	Nucleotide GenBank Acc. No.	Plasmid size (kb)	Relaxase GenBank Acc. No. [b]	Relaxase tree cluster [c]	Host	Replication initiator family [d]	sso type (position related to mob gene) [e]	Virulence or antibiotic / metal resistance [f]
pMV158	NC_010096.1	5.541	YP_001586274	pMV158	*Streptococcus 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 YP_001429523.1 (MOB$_Q$)	pYSI8	*Lactococcus lactis*	Rep_trans and Rep_3+L_lactis_RepB_C	ND	MEP, Strp, Chlr, Tc
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 YP_007419109.1	pSYI8 pKKS825	*Staphylococcus hyicus*	Rep_trans and HTH_Hin_like and truncated	ND	Strp, MLS(B), Tc
pKKS825	NC_013034.2	14.363	YP_003084337.1 YP_003084330.1 YP_004679012.1 (MOB$_Q$, truncated and fused to Rep_1)	pKKS825 pKKS825	*Staphylococcus aureus*	Rep_1 and HTH_Hin_like and Rep_3	ND	Kan/Neo, Tc, Trim, PLS(A)
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 NP_899168.1	pKKS825 pNM11	*Staphylococcus sciuri*	HTH_Hin_like and Rep_3	ND	Flr/Chlr, MLS(B), Spec
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	
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	pAMalpha1	*Enterococcus faecalis*	Rep_1+Rep_1 and Rep_3	ND	Tc

			EEU66441.1	pUB110				
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 thuringiensi*	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 AFJ49142.1 (MOB$_V$, truncated)		pNM11	*Bacillus* sp.	Rep_1	ND	Flr/Chlr, Strp
pSS-03	NC_016054.1	7.122	YP_004888092.1 YP_004888090.1 (MOB$_V$, truncated)		pNM11	*Staphylococcus arlettae*	Rep_1	ND	Flr/Chlr, MLS(B)
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 Figure 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

55

1214 **Highlights**

1215

1216 - The replication and transfer of rolling circle-replicating plasmids is reviewed.

1217 - Comparisons of replication and transfer cassettes are presented.

1218 - The current understanding of the pMV158 DNA transfer mechanism is reviewed.

1219

1220