Integrating Vectors for Genetic Studies in the Rare Actinomycete *Amycolatopsis marina*

Hong Gao[1,*], Buvani Murugesan[1], Janina Hossbach[1], Stephanie K. Evans[1], W. Marshall Stark[2] and Margaret C.M. Smith[1]

1 Department of Biology, University of York, York, North Yorkshire, UK, YO10 5DD

2 Institute of Molecular, Cell and Systems Biology, University of Glasgow, Glasgow, UK, G12 8QQ

* Hong Gao: hong.gao@york.ac.uk

Abstract:

Few natural product pathways from rare Actinomycetes have been studied due to the difficulty in applying molecular approaches in these genetically intractable organisms. In this study, we sought to identify integrating vectors, derived using phage *int/attP* loci, that would efficiently integrate site-specifically in the rare Actinomycete, *Amycolatopsis marina* DSM45569. Genomic analysis of the genome of *A. marina* DSM45569 indicated the presence of *attB*-like sites for TG1 and R4 integrases. The TG1 and R4 *attB*s were active in *in vitro* recombination assays with their cognate purified integrases and *attP* loci. Integrating vectors containing either the TG1 or R4 *int/attP* loci yielded a high frequency of exconjugants in conjugation assays from *E. coli* to *A. marina* DSM45569. Site-specific recombination of the plasmids into the host TG1 or R4 *attB* sites was confirmed by sequencing. The presence of homologous TG1 and R4 *attB* sites in other species of this genus indicated that vectors based on TG1 and R4 integrases could be widely applicable.

Keywords: Rare Actinomycetes, Amycolatopsis, integrating vectors, TG1 integrase, R4 integrase

Introduction

Streptomyces bacteria are widely exploited for their abundant bioactive natural products[1]. After 70 years of exploitation, however, the rate of discovery of new *Streptomyces* derived bioactive products has declined, and interest in other potential non-Streptomycete sources, such as rare Actinomycetes has grown[2, 3].

Amongst rare Actinomycetes, the *Amycolatopsis* genus generates particular interest for its production of important antibiotics such as vancomycin[4] and rifamycin[5], as well as a diverse range of active natural products[6-8]. The publicly available NCBI database contains nearly 70 genomes of *Amycolatopsis* strains, covering more than 40 species from this genus. Similar to *Streptomyces*, the genome of each *Amycolatopsis* contains, on average, over 20 secondary metabolic gene clusters[9]. The mining of these metabolic clusters, much like *Streptomyces*, offers great potential for novel antibiotic discovery as the gene products of these clusters are largely uncharacterised. The lack of widely available and efficient genetic tools for these rare species, however, impedes their potential in research and development.

Phage-encoded serine and tyrosine integrases catalyse site-specific integration of a circularized phage genome into the host chromosome as part of the process to establish a lysogen. DNA integration mediated by serine integrases occurs between short (approximately 50 bp) DNA substrates that are located on the phage genome, the phage attachment site (*attP*) and the host genome, the bacterial attachment site (*attB*). The product of *attP* x *attB* recombination is an integrated phage genome flanked by two new sites, *attL* and *attR*, each of which contain half sites from *attP* and *attB*. During phage induction integrase in the presence of a recombination directionality factor (RDF) again mediates site-specific recombination but this time between *attL* and *attR* to excise the phage genome, which can then be replicated during a lytic cycle. The mechanism of recombination and the factors that control integration versus excision have been elucidated in recent years[10-12].

57 Integrating vectors based on the *Streptomyces* phage ϕC31 integrase and *attP* locus
58 are the best known and most widely used in Actinomycete genome engineering[13-16].
59 In addition to the phage recombination machinery (*int/attP*), integrating vectors
60 contain a replicon for maintenance in *E. coli*, an *oriT* for conjugal transfer and a
61 marker or markers for selection in *E. coli* and in the recipient. They are powerful
62 genome engineering tools that act in an efficient, highly controllable and predictable
63 way[17].
64
65 The use of integrating vectors such as those derived from serine integrases requires
66 no additional phage or host functions for integration, thus makes it particularly
67 salient when the system is applied to other organisms that cannot be infected by the
68 phages. This hallmark makes serine integrases an promising tool used in
69 heterologous systems[11, 18]. However, these integration systems have not been fully
70 explored in rare Actinomycetes, e.g. *Amycolatopsis*. There is only one example of a
71 conjugation system based on ϕC31 integrase in *Amycolatopsis japonicum*
72 MG417-CF17[19], and it has been reported that other *Amycolatopsis* species lack ϕC31
73 *attB* sites in their chromosomes[20]. The ϕBT1 *attB* sites have been more commonly
74 identified in *Amycolatopsis*, and there have been successful conjugative transfer of
75 vectors based on ϕBT1 *int/attP* in *A. orientalis*[20] and *A. mediterranei*[21]. Despite these
76 successes, electroporation still remains the most applicable method used to transfer
77 integrative plasmids into this genus[20, 21].
78
79 In this paper, we chose to study *Amycolatopsis marina* DSM45569, a particular
80 *Amycolatopsis* species isolated from an ocean-sediment sample collected in the
81 South China Sea[22]. We explored the application of bacterial genetic engineering
82 using serine integrases, and developed conjugative and integrating vectors for use in
83 this species. We present evidence that suggests this vector could be applied to
84 various other species in this genus, opening up the prospect for more versatile
85 genetic manipulation of *Amycolatopsis*.
86

87 ## Materials and Methods

88 **Bacterial strains and culture conditions**

89 Plasmid propagation and subcloning was conducted using *E. coli* Top10 (F- *mcrA*
90 Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 *ΔlacX74 nupG recA1 araD139* Δ(*ara-leu)7697*
91 *galE15 galK16 rpsL*(Str^R) *endA1* λ^-). Plasmid conjugations from *E. coli* to *Streptomyces*
92 were carried out using *E. coli* ET12567(pUZ8002) containing the plasmid to be
93 transferred as the donor [23]. *E. coli* strains were grown in Luria-Bertani broth (LB) or
94 on LB agar at 37°C.

95 *Amycolatopsis marina* DSM45569 was purchased from the German Collection of
96 Microorganisms and Cell Cultures (DSMZ, Germany), and maintained on Soya
97 Mannitol (SM) agar at 30°C. Harvested spores were maintained long term in 20%
98 glycerol at -80°C. Conjugations were plated on SM agar containing 10 mM $MgCl_2$ and
99 ISP2 medium[24] was used for the preparation of genomic DNA[25].

100 **DNA manipulation**

101 *E. coli* transformation and gel electrophoresis were carried out as described
102 previously[26]. Genomic DNA preparation from *Streptomyces* was performed following
103 the salting out procedure in the *Streptomyces* manual[25]. Plasmids from *E. coli* were
104 prepared using QIAprep® Spin Miniprep Kit (Qiagen, Germany) following the
105 manufacturer's instructions. Polymerase Chain Reaction (PCR) was carried out using
106 Phusion® High-Fidelity DNA Polymerase (NEB, USA) according to the manufacturer's
107 instructions. The primers used in this study are listed in Table 1. DNA samples were
108 purified by the QIAquick Gel Extraction Kit (Qiagen, Germany).

109 **Plasmid construction**

110 The integrating plasmid pHG4 contains the TG1 *int/attP* locus and the
111 apramycin-resistance gene (*aac3(IV)*) for selection (Figure 1). The fragment
112 containing *oriT*, *aac3(IV)* and TG1 *int/attP* was amplified from plasmid pBF20[27] using
113 the primer pair pHG4-for/pHG4-rev. The fragment was joined via In-Fusion cloning to
114 the 3344 bp HindIII-PacI fragment from pBF22[27] (containing the *E. coli* plasmid
115 replication origin, the *bla* gene encoding resistance to ampicillin and the

116 *actII-orf4/act1p* expression cassette) to form the plasmid pHG4.

117 To construct the integrating plasmid pJH1R4 (Figure 1), pSET152[28] was cut with AatII
118 and PvuI to remove the ϕC31 *attP* site and integrase gene. R4 phage lysate was used
119 as template in a PCR with the primers pJH1R4-for and pJH1R4-rev to amplify the R4
120 *attP* site and integrase coding region. The PCR product was joined to the AatII-PvuI
121 fragment from pSET152 via In-Fusion cloning.

122 The plasmid pHG1 was used as template to amplify attB-containing sequences for in
123 vitro recombination assays. This plasmid was originally constructed for the
124 expression of *EryF*. The eryF gene was amplified from *Saccharopolyspora erythraea*
125 BIOT-0666 genomic DNA using the primer pair pHG1A-for/pHG1A-rev, and inserted
126 by In-Fusion cloning into pBF20[27] cut with NheI and PacI to form the plasmid pHG1A.
127 The 3785 bp fragment containing the ϕC31 int/attP and hygromycin resistance gene
128 was amplified from plasmid pBF27C[27], using the primer pair pHG1-for and pHG1-rev.
129 Plasmid pHG1A was digested with XbaI and NheI, the 5668 bp fragment was ligated
130 with the 3785 bp PCR fragment from pBF27C by In-Fusion to give the plasmid pHG1
131 (Figure 1).

132 ***In Vitro* Recombination Assays**

133 *In vitro* recombination assays were performed using PCR amplified DNA fragments
134 containing the *attB* and *attP* attachment sites located at the ends. Recombination
135 between the *attP* and *attB* sites yielded a fusion of the two fragments to give a
136 product, whose length was almost the sum of the substrates (Figure 3A). To generate
137 the *attB* containing substrates, the forward primer, TG1-attB-Am-for, contained the
138 closest match in the *A. marina* genome to the characterised TG1 *attB* site from *S.*
139 *avermitilis,* TG1 $attB^{Sa}$ [29](Figure 2). TG1-attB-Am-for also had a sequence identical to
140 the 3' end of the act1p element from plasmid pHG1, which was used as a template
141 for PCR (Figure 1). Similarly, the forward primer R4-attB-Am-for contained the closest
142 match in the *A. marina* genome to the characterized R4 *attB* site from *S. parvulus,* R4
143 $attB^{Sp}$ [30] (Figure 2). R4-attB-Am-for also had a sequence identical to the 3' end of
144 ActII-orf4 element from the template plasmid pHG1 (Figure 1). Forward primers
145 TG1-attB-Sa-for and R4-attB-Sp-for were used to create positive control

146 recombination substrates containing the TG1 and R4 *attB* sites originally found in *S.*
147 *avermitilis*[29] and *S. parvulus*[30] respectively. The reverse primer used to generate all
148 the *attB* containing substrates (attB-rev) was located within the *hyg* gene of pHG1
149 and the amplified products were 1627 bp (TG1 $attB^{Am}$), 1035 bp (TG1 $attB^{Sa}$), 1854 bp
150 (R4 $attB^{Am}$) and 1855 bp (R4 $attB^{Sp}$). The DNA fragments containing the *attP* sites
151 were prepared as follows; the TG1-attP fragment (2471 bp) was amplified using the
152 primer pair TG1-attP-for/TG1-attP-rev with pHG4 as the template, and the R4-attP
153 fragment (990 bp) was amplified using the primer pair R4-attP-for/R4-attP-rev with
154 pJH1R4 as the template (Figure 1). Note that other than the *attB* and *attP* sites none
155 of the substrates contained any DNA that should interact with the integrases other
156 than non-specifically. Moreover each fragment was designed to be easily identifiable
157 by molecular weight.
158 The integrases were purified as described previously[31, 32]. All recombination reactions
159 were in 20 µl final volume. Recombination reaction of TG1 substrates was carried out
160 in TG1 RxE buffer (20 mM Tris [pH 7.5], 25 mM NaCl, 1 mM dithiothreitol [DTT], 10
161 mM spermidine, 10 mM EDTA, 0.1 mg/ml bovine serum albumin [BSA])[33], and
162 recombination reaction of R4 substrates was carried out in buffer containing 20 mM
163 Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM spermidine, 5 mM $CaCl_2$ and 50 mM DTT[32].
164 Integrase was added at the concentrations indicated. Recombination substrates were
165 used at 50 ng each per reaction. Reactions were incubated at 30°C overnight, and
166 then heated (10 min, 75°C) to denature integrase. The reaction mixtures were loaded
167 on a 0.8% agarose gel in Tris/Borate/EDTA (TBE) buffer (90 mM Tris base, 90 mM
168 boric acid and 2 mM EDTA) containing ethidium bromide for electrophoretic
169 separation.

170

171 ## Results

172 **Identification of possible *attB* sequences from the genome of *Amycolatopsis***
173 ***marina* DSM45569**

174 The sequences of *attB* sites recognised by a variety of integrases (ϕ31[34], ϕJoe[35],

175 Bxb1[36], R4[32], SPBc[37], SV1[38], TG1[39] and TP901[40]) were used in BLAST searches of the
176 genome sequence of *Amycolatopsis marina* DSM45569 (NCBI Genome Database
177 NZ_FOKG00000000)(Table 2). The most significant hits for R4 and TG1 *attB* sites had
178 the highest identities and lowest *E*-value. The R4 *attB* site is located within a gene
179 predicted to encode a fatty-acyl-CoA synthase (SFB62308.1) and the TG1 *attB* site is
180 located within a gene predicted to encode a putative succinyldiaminopimelate
181 transaminase (WP_091671332.1). The BLAST analysis was extended to other species
182 of *Amycolatopsis* to assess the conservation of *attB* sites in the genus (Figure 2). Both
183 R4 and TG1 *attB* sites were highly conserved to the *attB* sites (84% for R4 and 62%
184 for TG1) originally identified from *S. parvulus*[30] and *S. avermitilis*[29].

185 ***In Vitro* Recombination**

186 In each recombination reaction, substrates containing *attP* and the putative *attB* site
187 were mixed in cognate pairs with different concentrations of purified R4 or TG1 in
188 the corresponding buffer and incubated overnight at 30°C, as described in Materials
189 and Methods. The expected recombination events and the nature of the products
190 are shown in Figure 3A-3C. TG1 catalysed the recombination between substrates
191 more efficiently than R4 (Figure 3D). As expected because neither phage is an
192 *Amycolatopsis* phage, the recombination efficiencies for each integrase were
193 observably better when the *Streptomyces attB* sites were used (Figure 3E) compared
194 to the *A. marina attB* sites (Figure 3D), particularly for TG1 integrase. Nevertheless,
195 the presence of recombination activity indicated that both *A. marina attB* sites were
196 functional in a recombination reaction and were likely to be active integration sites
197 for integrative conjugation vectors.

198 ***In Vivo* Integration**

199 *A. marina* DSM45569 was unable to grow in the presence of apramycin so integrating
200 plasmids, pHG4 and pJH1R4, containing the apramycin resistance determinant,
201 aac3(IV) were constructed. Following the standard *Streptomyces* conjugation
202 protocol (see Materials and Methods), a frequency of approximately 160
203 exconjugants/10^8 spores was obtained for transfer of pHG4 (encoding TG1 integrase),
204 while the conjugation efficiency of pJH1R4 (R4 integrase) was only 20

205 exconjugants/10^8 spores (Table 3). For each integration, six exconjugants were picked
206 at random and streaked on MS agar containing apramycin first, then genomic DNA
207 was prepared and used as template in PCR reactions, in which the primer pairs of
208 TG1-attL-Am-for/rev and R4-attL-Am-for/rev were used to test the occurrence of
209 recombination at the expected TG1 and R4 *attB* sites (Figure 4). All PCR reactions
210 using exconjugants as templates gave the expected band sizes. Sequencing (GATC,
211 Germany) of the PCR products with the primers TG1-attL-Am-for and R4-attL-Am-for
212 confirmed that the plasmids had integrated into the predicted *attB* sites for TG1 or
213 R4 integrase within *A. marina* DSM45569 (Figure 5).

214 Discussion

215 The lack of effective genetic engineering tools is considered one of the greatest
216 hindrances in the search for novel natural products from rare Actinomycetes[41-43].
217 Previous studies have largely focused on the use of the well-characterised ϕC31
218 based integration vectors, and have mostly overlooked other phage integrases in rare
219 Actinomycete species[44-46]. Additionally, the conjugation methods used widely in
220 *Streptomyces* gene transfer have shown little success in rare Actinomycetes,
221 including the genus *Amycolatopsis* species, thus electroporation has been the
222 long-preferred method of gene transfer in this species[5, 9, 47, 48]. However, the growing
223 interest in the use of serine integrases for synthetic biology applications[11] has led to
224 the further research into expanding the pool of available enzymes and their
225 potentials as genetic tools[49-51]. Therefore, within this study, we explored whether
226 integrating vectors derived from eight different serine integrases could be employed
227 for genetic engineering of *A. marina* DSM45569. Sequence analysis of the *A. marina*
228 DSM45569 genome identified close matches to the *attB* sites used by TG1 and R4
229 integrases. Although conjugation frequencies were low, integrating plasmids based
230 on the TG1 and R4 recombination systems successfully integrated into the expected
231 *attB* sites in *A. marina* DSM45569. Conservation between the *attB* sites for TG1 and
232 R4 in a number of *Amycolatopsis* species is high suggesting that the plasmids derived
233 from these phages should be widely useful in this genus.

234

235 As is common with serine integrase mediated recombination, the attB sites are
236 located within open reading frames and potentially disrupt the gene. TG1 *attB*^{Am} site
237 is located within a gene predicted to encode an aminotransferase, and the R4 *attB*^{Am}
238 site is located within a gene predicted to encode a fatty-acyl-CoA synthase.
239 Compared to the wild type (unintegrated) strain, the strains with integrated pHG4 or
240 pJH1R4 did not show any difference in growth. However, further study is required to
241 investigate the effects of TG1 or R4 plasmid recombination on both of primary and
242 secondary metabolism as, for example, the integration of ϕC31 derived plasmids has
243 been shown to have pleiotropic effects on bacterial physiology[52].
244

245 Currently, the two most commonly used methods of bacterial gene transfer are
246 conjugation and electroporation, and both come with several advantages and
247 disadvantages. Electroporation is the introduction of pores within bacterial
248 membranes via an electric current to allow mass, unrestricted transfer of genetic
249 material into species[53, 54]. The efficiency of transformation, however, is species
250 dependent[53]. Unlike electroporation, conjugation is not limited by the size of vectors
251 that can be transformed, and has been used to successfully transfer entire genomes
252 in *E. coli*[55]. Additionally, conjugation uses the transfer of DNA as a single strand that is
253 relatively insensitive to the majority of the restriction systems of the cells[56]. This
254 demonstrates the advantageous use of conjugation over electroporation. In this
255 study, we successfully integrated a plasmid via the *attB* sites for TG1 and R4
256 integrases by conjugation, thus supplementing the potential gene transfer methods
257 that could be used in the genus *Amycolatopsis*.
258

259 In conclusion, we have identified the highly conserved sequences of the *attB* sites for
260 TG1 and R4 integrases within the genus *Amycolatopsis* and demonstrated their use in
261 conjugative DNA transfer. The *A. marina* DSM45569 *attB* sites showed slightly lower
262 recombination efficiencies *in vitro* than the previously identified *attB* sites from
263 *Streptomyces* spp. although this is not likely enough to explain the order of

magnitude reductions in conjugation frequencies observed with *A. marina* compared to *Streptomyces* spp. (Table 3). By optimising conjugation conditions the conjugation frequencies could be increased. Alternatively, efficiently used *attB* sites for the widely used vectors, such as those based on phiC31 *int/attP* could be incorporated in to the *Amylcolatopsis* genome using TG1 or R4 integrating plasmids described here. In short this work shows that integrative vectors are viable and promising tools for the genetic engineering of rare actinomycetes.

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council project grant BB/K003356/1, and Buvani Murugesan acknowledges the receipt of a summer studentship from the Department of Biology, University of York.

References

1. Jones, S.E. et al. Streptomyces exploration is triggered by fungal interactions and volatile signals. *Elife* **6**, e21738 (2017).
2. Zarins-Tutt, J.S. et al. Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Natural product reports* **33**, 54-72 (2016).
3. Jose, P.A. & Jebakumar, S.R.D. Non-streptomycete actinomycetes nourish the current microbial antibiotic drug discovery. *Frontiers in microbiology* **4** (2013).
4. Jung, H.-M., Kim, S.-Y., Moon, H.-J., Oh, D.-K. & Lee, J.-K. Optimization of culture conditions and scale-up to pilot and plant scales for vancomycin production by Amycolatopsis orientalis. *Applied microbiology and biotechnology* **77**, 789-795 (2007).
5. Li, C. et al. RifZ (AMED_0655) is a pathway-specific regulator for rifamycin biosynthesis in Amycolatopsis mediterranei. *Applied and Environmental Microbiology* **83**, e03201-03216 (2017).
6. Hashizume, H. et al. Valgamicin C, a novel cyclic depsipeptide containing the unusual amino acid cleonine, and related valgamicins A, T and V produced by Amycolatopsis sp. ML1-hF4. *The Journal of Antibiotics*, ja2017135 (2017).
7. Li, X., Wu, X., Zhu, J. & Shen, Y. Amexanthomycins A–J, pentangular polyphenols produced by Amycolatopsis mediterranei S699Δ rifA. *Applied Microbiology and Biotechnology*, 1-14 (2017).
8. Schwarz, P.N. et al. The Immunosuppressant Brasilicardin: Determination of the Biosynthetic Gene Cluster in the Heterologous Host Amycolatopsis japonicum. *Biotechnology Journal* (2017).
9. Kumari, R., Singh, P. & Lal, R. Genetics and genomics of the genus Amycolatopsis. *Indian journal of microbiology* **56**, 233-246 (2016).
10. Colloms, S.D. et al. Rapid metabolic pathway assembly and modification using serine

301 integrase site-specific recombination. *Nucleic acids research* **42**, e23-e23 (2013).
302 11. Fogg, P.C., Colloms, S., Rosser, S., Stark, M. & Smith, M.C. New applications for phage
303 integrases. *Journal of molecular biology* **426**, 2703-2716 (2014).
304 12. Smith, M.C. in Mobile DNA III 253-272 (American Society of Microbiology, 2015).
305 13. Sosio, M. et al. Artificial chromosomes for antibiotic-producing actinomycetes. *Nature*
306 *biotechnology* **18**, 343-345 (2000).
307 14. Hong, Y. & Hondalus, M.K. Site-specific integration of Streptomyces ΦC31 integrase-based
308 vectors in the chromosome of Rhodococcus equi. *FEMS microbiology letters* **287**, 63-68
309 (2008).
310 15. Saha, S. et al. Activation and characterization of a cryptic gene cluster reveals a cyclization
311 cascade for polycyclic tetramate macrolactams. *Chemical Science* **8**, 1607-1612 (2017).
312 16. Thorpe, H.M. & Smith, M.C. In vitro site-specific integration of bacteriophage DNA catalyzed
313 by a recombinase of the resolvase/invertase family. *Proceedings of the National Academy of*
314 *Sciences* **95**, 5505-5510 (1998).
315 17. Rutherford, K. & Van Duyne, G.D. The ins and outs of serine integrase site-specific
316 recombination. *Current opinion in structural biology* **24**, 125-131 (2014).
317 18. Baltz, R.H. Streptomyces temperate bacteriophage integration systems for stable genetic
318 engineering of actinomycetes (and other organisms). *Journal of industrial microbiology &*
319 *biotechnology* **39**, 661-672 (2012).
320 19. Stegmann, E., Pelzer, S., Wilken, K. & Wohlleben, W. Development of three different gene
321 cloning systems for genetic investigation of the new species Amycolatopsis japonicum
322 MG417-CF17, the ethylenediaminedisuccinic acid producer. *Journal of biotechnology* **92**,
323 195-204 (2001).
324 20. Lei, X. et al. Complete genome sequence of Amycolatopsis orientalis CPCC200066, the
325 producer of norvancomycin. *Journal of biotechnology* **247**, 6-10 (2017).
326 21. Li, C., Zhou, L., Wang, Y., Zhao, G. & Ding, X. Conjugation of ϕBT1-derived integrative plasmid
327 pDZL802 in Amycolatopsis mediterranei U32. *Bioengineered* **8**, 549-554 (2017).
328 22. Bian, J. et al. Amycolatopsis marina sp. nov., an actinomycete isolated from an ocean
329 sediment. *International journal of systematic and evolutionary microbiology* **59**, 477-481
330 (2009).
331 23. MacNeil, D.J. Characterization of a unique methyl-specific restriction system in Streptomyces
332 avermitilis. *Journal of bacteriology* **170**, 5607-5612 (1988).
333 24. Shirling, E.t. & Gottlieb, D. Methods for characterization of Streptomyces species1.
334 *International Journal of Systematic and Evolutionary Microbiology* **16**, 313-340 (1966).
335 25. Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. & Hopwood, D.A. Practical Streptomyces
336 Genetics. (The John Innes Foundation, Norwich; 2000).
337 26. Sambrook, J. & Russell, D.W. Molecular cloning: a laboratory manual. (Cold Spring Harbor
338 Laboratory Press, New York; 2001).
339 27. Fayed, B. et al. Multiplexed integrating plasmids for engineering the erythromycin gene
340 cluster for expression in Streptomyces and combinatorial biosynthesis. *Applied and*
341 *environmental microbiology*, AEM. 02403-02415 (2015).
342 28. Wilkinson, C.J. et al. Increasing the efficiency of heterologous promoters in actinomycetes.
343 *Journal of molecular microbiology and biotechnology* **4**, 417-426 (2002).
344 29. Morita, K. et al. The site-specific recombination system of actinophage TG1. *FEMS*

345 *microbiology letters* **297**, 234-240 (2009).
346 30. Shirai, M., Nara, H., Sato, A., Aida, T. & Takahashi, H. Site-specific integration of the
347 actinophage R4 genome into the chromosome of Streptomyces parvulus upon lysogenization.
348 *Journal of bacteriology* **173**, 4237-4239 (1991).
349 31. Rowley, P.A., Smith, M.C., Younger, E. & Smith, M.C. A motif in the C-terminal domain of ϕC31
350 integrase controls the directionality of recombination. *Nucleic acids research* **36**, 3879-3891
351 (2008).
352 32. Miura, T. et al. In vivo and in vitro characterization of site-specific recombination of
353 actinophage R4 integrase. *The Journal of general and applied microbiology* **57**, 45-57 (2011).
354 33. Morita, K. et al. Site-specific genome integration in alphaproteobacteria mediated by TG1
355 integrase. *Applied microbiology and biotechnology* **93**, 295-304 (2012).
356 34. Combes, P., Till, R., Bee, S. & Smith, M.C. The Streptomyces genome contains multiple
357 pseudo-attB sites for the ϕC31-encoded site-specific recombination system. *Journal of*
358 *bacteriology* **184**, 5746-5752 (2002).
359 35. Fogg, P.C., Haley, J.A., Stark, W.M. & Smith, M.C. Genome integration and excision by a new
360 Streptomyces bacteriophage, ϕJoe. *Applied and environmental microbiology*, AEM.
361 02767-02716 (2016).
362 36. Ghosh, P., Pannunzio, N.R. & Hatfull, G.F. Synapsis in phage Bxb1 integration: selection
363 mechanism for the correct pair of recombination sites. *Journal of molecular biology* **349**,
364 331-348 (2005).
365 37. Lazarevic, V. et al. Nucleotide sequence of the Bacillus subtilis temperate bacteriophage
366 SPβc2. *Microbiology* **145**, 1055-1067 (1999).
367 38. Fayed, B., Younger, E., Taylor, G. & Smith, M.C. A novel Streptomyces spp. integration vector
368 derived from the S. venezuelae phage, SV1. *BMC biotechnology* **14**, 51 (2014).
369 39. Morita, K. et al. In vitro characterization of the site-specific recombination system based on
370 actinophage TG1 integrase. *Molecular genetics and genomics* **282**, 607 (2009).
371 40. Christiansen, B., Johnsen, M., Stenby, E., Vogensen, F. & Hammer, K. Characterization of the
372 lactococcal temperate phage TP901-1 and its site-specific integration. *Journal of bacteriology*
373 **176**, 1069-1076 (1994).
374 41. Li, X., Zhou, X. & Deng, Z. Vector systems allowing efficient autonomous or integrative gene
375 cloning in Micromonospora sp. strain 40027. *Applied and environmental microbiology* **69**,
376 3144-3151 (2003).
377 42. Marcone, G.L., Carrano, L., Marinelli, F. & Beltrametti, F. Protoplast preparation and reversion
378 to the normal filamentous growth in antibiotic-producing uncommon actinomycetes. *The*
379 *Journal of antibiotics* **63**, 83-88 (2010).
380 43. Wolf, T. et al. Targeted genome editing in the rare actinomycete Actinoplanes sp. SE50/110 by
381 using the CRISPR/Cas9 System. *Journal of biotechnology* **231**, 122-128 (2016).
382 44. Wagner, N., Oßwald, C., Biener, R. & Schwartz, D. Comparative analysis of transcriptional
383 activities of heterologous promoters in the rare actinomycete Actinoplanes friuliensis. *Journal*
384 *of biotechnology* **142**, 200-204 (2009).
385 45. Anzai, Y. et al. Production of rosamicin derivatives in Micromonospora rosaria by introduction
386 of d-mycinose biosynthetic gene with ΦC31-derived integration vector pSET152. *Journal of*
387 *industrial microbiology & biotechnology* **36**, 1013 (2009).
388 46. Kim, D.-Y., Hwang, Y.-I. & Choi, S.-U. Cloning of metK from Actinoplanes teichomyceticus

389 ATCC31121 and effect of its high expression on antibiotic production. *Journal of microbiology*
390 *and biotechnology* **21**, 1294-1298 (2011).
391 47. Shen, Y., Huang, H., Zhu, L., Luo, M. & Chen, D. Type II thioesterase gene (ECO-orf27) from
392 Amycolatopsis orientalis influences production of the polyketide antibiotic, ECO-0501 (LW01).
393 *Biotechnology letters* **34**, 2087-2091 (2012).
394 48. Lee, K. et al. Increased vancomycin production by overexpression of MbtH-like protein in
395 Amycolatopsis orientalis KFCC10990P. *Letters in applied microbiology* **63**, 222-228 (2016).
396 49. Mandali, S., Dhar, G., Avliyakulov, N.K., Haykinson, M.J. & Johnson, R.C. The site-specific
397 integration reaction of Listeria phage A118 integrase, a serine recombinase. *Mobile DNA* **4**,
398 2-2 (2013).
399 50. Yoon, B., Kim, I., Nam, J.-A., Chang, H.-I. & Ha, C.H. In vivo and in vitro characterization of
400 site-specific recombination of a novel serine integrase from the temperate phage EFC-1.
401 *Biochemical and biophysical research communications* **473**, 336-341 (2016).
402 51. Yang, L. et al. Permanent genetic memory with> 1-byte capacity. *Nature methods* **11**,
403 1261-1266 (2014).
404 52. Talà, A. et al. Pirin: A novel redox-sensitive modulator of primary and secondary metabolism
405 in Streptomyces. *Metabolic engineering* **48**, 254-268 (2018).
406 53. Bekker, V., Dodd, A., Brady, D. & Rumbold, K. Tools for metabolic engineering in Streptomyces.
407 *Bioengineered* **5**, 293-299 (2014).
408 54. Cadoret, F., Soscia, C. & Voulhoux, R. Gene transfer: transformation/electroporation.
409 *Pseudomonas Methods and Protocols*, 11-15 (2014).
410 55. Isaacs, F.J. et al. Precise manipulation of chromosomes in vivo enables genome-wide codon
411 replacement. *Science* **333**, 348-353 (2011).
412 56. Dominguez, W. & O'Sullivan, D.J. Developing an efficient and reproducible conjugation-based
413 gene transfer system for bifidobacteria. *Microbiology* **159**, 328-338 (2013).
414

415 **Figures**

416 **Figure 1**. Plasmids used in this study. The primer binding sites are indicated.

417 **Figure 2**. Alignment of R4 and TG1 *attB* sites in *A. marina* DSM45569 and other
418 *Amycolatopsis* species. GenBank accession nos. of DNA sequences: *A. alba*
419 (NMQU01000019.1), *A. antarctica* (NKYE01000021.1), *A. azurea* (MUXN01000005.1),
420 *A. kentuckyensis* (MUMI01000226.1), *A. lurida* (FNTA01000004.1), *A. nigrescens*
421 (ARVW01000001.1), *A. niigatensis* (PJMY01000003.1), *A. orientalis*
422 (ASXH01000007.1), *A. pretoriensis* (MUMK01000092.1), *A. regifaucium*
423 (LQCI01000034.1), *A. rubida* (FOWC01000001.1), *A. sacchari* (FORP01000010.1), *A.*
424 *thailandensis* (NMQT01000219.1), *A. thermoflava* (AXBH01000004.1), *A.*
425 *tolypomycina* (FNSO01000004.1), *A. xylanica* (FNON01000002.1), *Streptomyces*
426 *avermitilis* (NC_003155.5) and *S. parvulus* (CP015866.1).

427 **Figure 3**. (A) Design and synthesis of DNA substrate *attB*. The *attB* sites were fused at
428 the end of forward primers, to amplify a sequence flanked with attB from a
429 unrelevant DNA template, pHG1 in this case. (B) Design and synthesis of DNA
430 substrate attP. The attP sites were amplified directly from integrating vectors carrying
431 TG1/attP (pHG4) or R4/attP (pJH1R4). (C) Recombination substrates and their
432 expected products. (D) *In vitro* recombination between DNA fragments containing
433 TG1 $attB^{Am}$ (1627 bp) and TG1 *attP* (2471 bp; left) and R4 $attB^{Am}$ (1854 bp) and R4
434 *attP* (990 bp; right). The expected product of the TG1 integrase mediated reaction
435 was a 4.1 kb DNA fragment containing the $attR^{Am}$ site and a 53 bp fragment
436 containing $attL^{Am}$ (not observed). For the R4 integrase recombination the expected
437 product was a 2.8 kb containing $attR^{Am}$ fragment and an 51 bp $attL^{Am}$ fragment (not
438 observed). (E) *In vitro* recombination between DNA fragments containing TG1 $attB^{Sa}$
439 (1035 bp) and TG1 *attP* (2471 bp; left) and R4 $attB^{Sp}$ (1855) and R4 *attP* (990 bp;
440 right). The expected product fragments were 3.5 kb containing $attR^{Sa}$ for the TG1
441 reaction and 2.8 kb containing $attR^{Sp}$ for the R4 reaction. M: Fast DNA Ladder (NEB,
442 USA).

443 **Figure 4**. PCR confirmation of site-specific integration in the exconjugants. (A) Design

444 of primers. (B) PCR (using primers TG1-attL-Am-for/rev) of the expected TG1 *attL*
445 containing fragment from *A. marina* DSM45569:pHG4. (C) PCR (using primers
446 R4-attL-Am-for/rev) of the expected R4 *attL* containing fragment from *A. marina*
447 DSM45569:pJH1R4. M: Fast DNA Ladder. Colonies 1 to 6 are independent
448 exconjugants.
449 **Figure 5.** The insertion sites of TG1 and R4 integration plasmids in *A. marina*
450 DSM45569. Sequencing (using primers TG1-attL-Am-for or R4-attL-Am-for) of PCR
451 products containing *attL* from exconjugants validated the site-specific recombination
452 of the TG1 and R4 *attB* sites in *A. marina* DSM45569 after introduction of pHG4 or
453 pJH1R4, respectively.
454

455 **Figure 1**

456

457 **Figure 2.**

A. R4 *attB* sites

```
A marina        GGTTGCCCATCACCATGCCGAAGCAGTGATAGAAGGGAACCGGGATGCAG
A kentuckyensis GGTTGCCCATCACCATGCCGAAGCAGTGGTAGAAGGGCACCGGGATGCAG
A nigrescens    GGTTGCCCATCACCATGCCGAAGCAGTGGTAGAAGGGCACCGGGATGCAG
A niigatensis   GGTTGCCCATCACCATGCCGAAGCAGTGGTAGAAGGGCACCGGGATGCAG
A orientalis    GGTTGCCCATCACCATGCCGAAGCAGTGGTAGAAGGGCACCGGGATGCAG
A pretoriensis  GGTTGCCCATCACCATGCCGAAGCAGTGGTAGAAGGGCACCGGGATGCAG
A rubida        GGTTGCCCATCACCATGCCGAAGCAGTGGTAGAAGGGCACCGGGATGCAG
A tolypomycina  GGTTGCCCATCACCATGCCGAAGCAGTGGTAGAAGGGCACCGGGATGCAG
A xylanica      GGTTGCCCATCACCATGCCGAAGCAGTGATAGAAGGGCACCGGGATGCAG
S parvulus      AGTTGCCCATGACCATGCCGAAGCAGTGGTAGAAGGGCACCGGCAGACAC
```

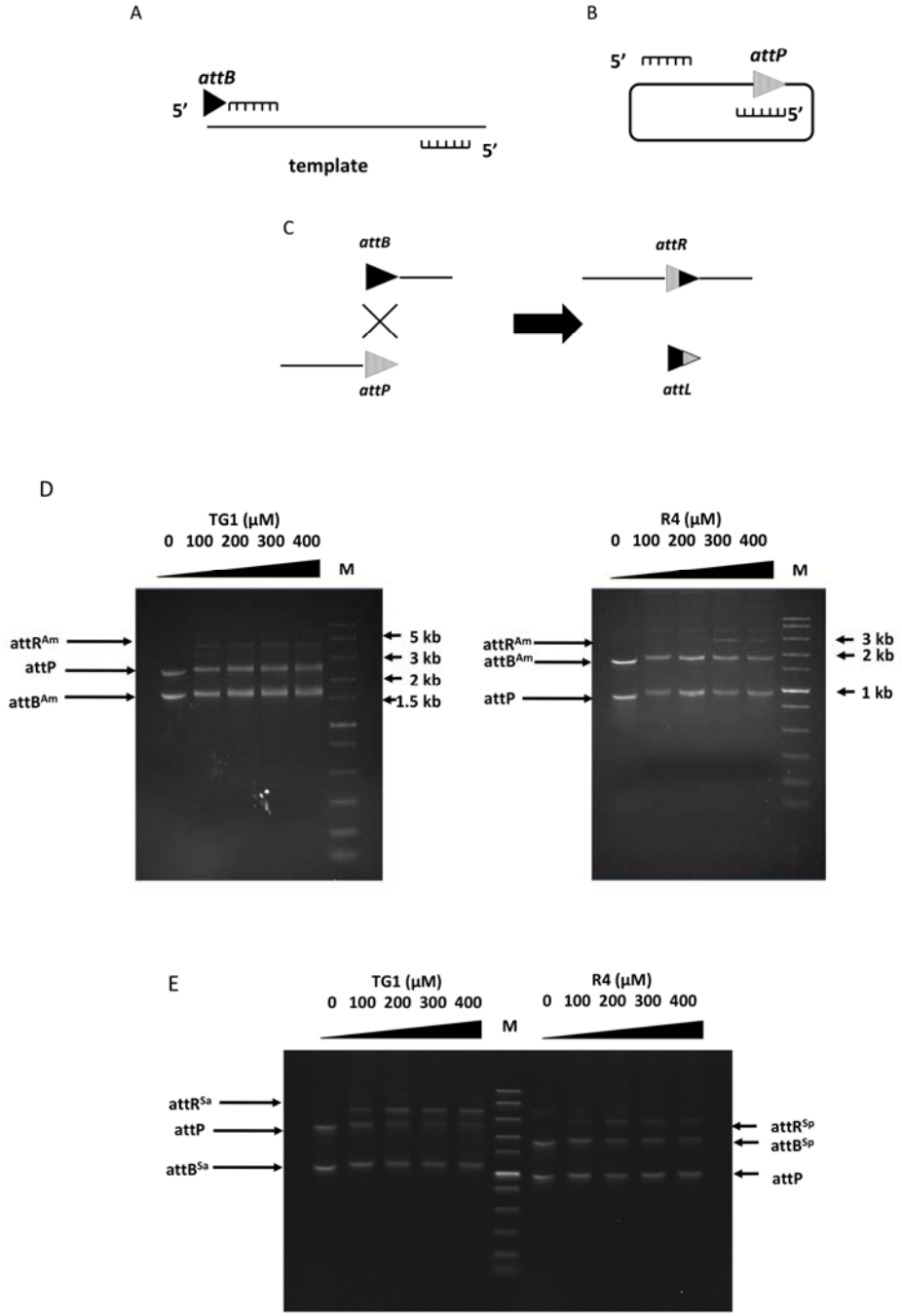
B. TG1 *attB* sites

```
A marina        TCGATCTCCAGTGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGG
A alba          TCGATCTCCAGCGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGG
A antarctica    AGCATCTCCAGTGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGG
A azurea        TCGATCTCCAGCGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGG
A lurida        TCGATCTCCAGCGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGG
A regifaucium   TCGATCTCCAGCGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGG
A sacchari      TCGATCTCCAGCGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGG
A thaitandensis TCGATCTCCAGCGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGG
A thermoflava   TCGATCTCCAGCGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGG
S avermitilis   TCGATCAGCTCCGCGGGCAAGACCTTCTCCTTCACGGGGTGGAAGGTCGG
```

458

459

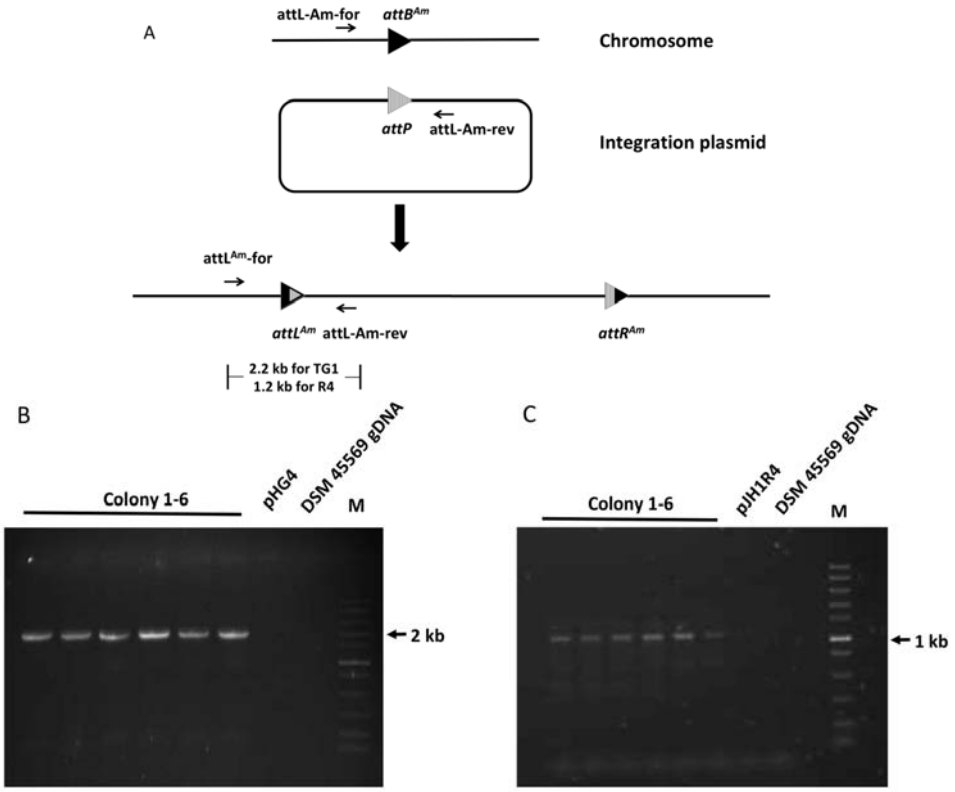
460 **Figure 3.**



461

462

463 **Figure 4.**



464

465

466 **Figure 5.**

A

```
  attB - A. marina GGTTGCCCATCACCATGCCGAAGCAGTG----------------------
attL - exconjugant GGTTGCCCATCACCATGCCGAAGCAGTGGTACTGCTTGTGGGTACACTCT
     attP - pJH1R4 -------------------GAAGCAGTGGTACTGCTTGTGGGTACACTCT
```

B

```
  attB - A. marina GATCTCCAGTGCGGGCAAGACGTT--------------------------
attL - exconjugant GATCTCCAGTGCGGGCAAGACGTTGCTCTTACCCAGTTGGGCGGGATAGC
       attP - pHG4 ----------------------TTGCTCTTACCCAGTTGGGCGGGATAGC
```

467

468

469 **TABLE 1** Oligonucleotides used in this study

Oligonucleotide	Sequence (5'-3')
pHG1A-for	CGAACGCATCGATTAATTAAGGAGGATCGTATGACGACCGTTCCCG
pHG1A-rev	CGTGGTGGGCGCTAGCCTCCTCTAGTCATCCGTCG
pHG1-for	ACTAGAGGAGGCTAGCTTCAATGGAGGAGATGATCGAGG
pHG1-rev	GCAGGTCGACTCTAGATCTCGCTACGCCGCTACG
pHG4-for	CGAACGCATCGATTAATTAAGCGGCCGCCATATGGAATTCGGTACCGCATGCAGAT CTAGGAACTTCGAAGTTCCCGC
pHG4-rev	TGATTACGCCAAGCTTTCGACTCTAGAGTAAGCGTCACGG
pJH1R4-for	CTAGCGATTGCCATGACGTCGGAGCTGCTTACCAATGTC
pJH1R4-rev	AAGAGGCCCGCACCGATTCCAAGAGGCCGGCAACTAC
TG1-attB-Am-for	TCGATCTCCAGTGCGGGCAAGACGTTCAACTGCACCGGCTGGAAGATCGGGACCA CCGGACGAACGCA
TG1-attB-Sa-for	TCGATCAGCTCCGCGGGCAAGACCTTCTCCTTCACGGGGTGGAAGGTCGGCGGTG GAGCTCGGAGA
R4-attB-Am-for	GGTTGCCCATCACCATGCCGAAGCAGTGATAGAAGGGAACCGGGATGCAGGTGA GAAGGTGCTCGTGT
R4-attB-Sp-for	AGTTGCCCATGACCATGCCGAAGCAGTGGTAGAAGGGCACCGGCAGACACGG
attB-rev	CTGCATCTCAACGCCTTCCGG
TG1-attP-for	AACCTTCACGCTCATGCC
TG1-attP-rev	GTCGAGATTCTCCGTCTCCTG
R4-attP-for	GATCGGTCTTGCCTTGCTC
R4-attP-rev	ACCCGCAGAGTGTACCCA
TG1-attL-Am-for	ACAACCCCACCGGCACCGTCTTCA
TG1-attL-Am-rev	AGTATAGGAACTTCGAAGCAGCTC
R4-attL-Am-for	CGGCCGGTGATGTTGACGT
R4-attL-Am-rev	TCGGCCGTCACGATGGTCA

470 The possible *attB* sequences identified from the genome sequence of *Amycolatopsis*
471 *marina* DSM45569 are shown underlined.
472

473 **TABLE 2**. The original *attB* sites for integrases and results of BLAST search

Integrase	*attB* sites	Reference	Best Hit from BLAST	
			Homology (%)	*E*-value
ϕ31	CGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACC	33	41	0.015
ϕJoe	ATCTGGATGTGGGTGTCCATCTGCGGGCAGACGCCGCAGTCGAAGCACGG	34	30	0.60
Bxb1	TCGGCCGGCTTGTCGACGACGGCGGTCTCCGTCGTCAGGATCATCCGGGC	35	48	0.014
R4	AGTTGCCCATGACCATGCCGAAGCAGTGGTAGAAGGGCACCGGCAGACAC	31	84	3e-11
SPBc	AGTGCAGCATGTCATTAATATCAGTACAGATAAAGCTGTATCTCCTGTGAACACAATGGGTGCCA	36	No hits	
SV1	ATCAGGGCGGTCAGGCCGTAGATGTGGAAGAACGGCAGCACGGCGAGGAC	37	32	0.17
TG1	TCGATCAGCTCCGCGGGCAAGACCTTCTCCTTCACGGGGTGGAAGGTCGG	38	62	0.001
TP901	ATGCCAACACAATTAACATCTCAATCAAGGTAAATGCTTTTTGCTTTTTTTGC	39	25	8.1

474

475

476 **TABLE 3**. Conjugation efficiency of pHG4 and pJH1R4 in different species

Exconjugants/10^8 spores	**pHG4**	**pJH1R4**
A. marina	160	20
S. coelicolor	1.47×10^3	3.28×10^4
S. lividans	1.56×10^3	3.33×10^4

477

478