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

* Communicating author: gaohong221@gmail.com

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. Analysis of the genome of *A. marina* DSM45569 indicated the presence of *attB*-like sequences 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 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 indicates 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). However, after 70 years of exploitation, the rate of discovery of new *Streptomyces*-derived bioactive products has declined, and interest in other potential non-Streptomycete sources, such as the rare Actinomycetes, has grown(2, 3).

Amongst rare Actinomycetes, the *Amycolatopsis* genus is of 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 offers great potential for novel antibiotic discovery. However, the lack of widely available and efficient genetic tools for these rare species 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 contains 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

58 Integrating vectors based on the *Streptomyces* phage ϕC31 integrase and *attP* locus
59 are the best known and most widely used in Actinomycete genome
60 engineering(13-16). In addition to the phage recombination machinery (*int/attP*),
61 integrating vectors contain a replicon for maintenance in *E. coli*, an *oriT* for conjugal
62 transfer and a marker or markers for selection in *E. coli* and in the recipient. They are
63 powerful genome engineering tools that act in an efficient, highly controllable and
64 predictable way(17).

65

66 Integrating vectors using serine integrase-mediated recombination require no
67 additional phage or host functions for integration, an especially important feature
68 when they are to be used in other organisms that cannot be infected by the phages.
69 This property makes serine integrase-based vectors promising tools for use in
70 heterologous systems(11, 18). However, use of these integration vectors has not
71 been fully explored in rare Actinomycetes, e.g. *Amycolatopsis*. There is only one
72 reported example of a conjugation system based on ϕC31 integrase in *Amycolatopsis*
73 *japonicum* MG417-CF17(19), and it has been reported that other *Amycolatopsis*
74 species lack ϕC31 *attB* sites in their chromosomes(20). The ϕBT1 *attB* sites have been
75 more commonly identified in *Amycolatopsis*, and there has been successful
76 conjugative transfer of vectors based on ϕBT1 *int/attP* in *A. orientalis(20)* and *A.*
77 *mediterranei(21)*. Furthermore, electroporation remains the most widely applied
78 method for transfer of integrative plasmids into this genus, rather than
79 conjugation(20, 21).

80

81 In this paper, we chose to study *Amycolatopsis marina* DSM45569, a species isolated
82 from an ocean-sediment sample collected in the South China Sea(22). We explored
83 the application of bacterial genetic engineering using serine integrases, and
84 developed conjugative and integrating vectors for use in this species. We present
85 evidence suggesting that these vectors could be applied to various other species in

86 this genus, opening up the prospect for more versatile genetic manipulation of
87 *Amycolatopsis*.
88

89 Materials and Methods

90 Bacterial strains and culture conditions

91 Plasmid propagation and subcloning was conducted using *E. coli* Top10 (F- *mcrA*
92 Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 *ΔlacX74 nupG recA1 araD139* Δ(*ara-leu)7697*
93 *galE15 galK16 rpsL*(Str^R) *endA1* λ^-). Plasmid conjugations from *E. coli* to *Streptomyces*
94 were carried out using *E. coli* ET12567(pUZ8002) containing the plasmid to be
95 transferred as the donor (23). *E. coli* strains were grown in Luria-Bertani broth (LB) or
96 on LB agar at 37°C.
97 *Amycolatopsis marina* DSM45569 was purchased from the German Collection of
98 Microorganisms and Cell Cultures (DSMZ, Germany), and maintained on Soya
99 Mannitol (SM) agar at 30°C. Harvested spores were maintained long-term in 20%
100 glycerol at -80°C. Conjugations were plated on SM agar containing 10 mM $MgCl_2$, and
101 ISP2 medium(24) was used for the preparation of genomic DNA(25).

102 DNA manipulation

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

111 Plasmid construction

112 The integrating plasmid pHG4 contains the TG1 *int/attP* locus and the
113 apramycin-resistance gene (*aac3(IV)*) for selection (Figure 1A). The fragment
114 containing *oriT*, *aac3(IV)* and TG1 *int/attP* was amplified from plasmid pBF20(27)

115 using the primer pair pHG4-for/pHG4-rev. The fragment was joined via In-Fusion
116 cloning to the 3344 bp HindIII-PacI fragment from pBF22(27) (containing the *E. coli*
117 plasmid replication origin, the *bla* gene encoding resistance to ampicillin and the
118 *actII-orf4/act1p* expression cassette) to form the plasmid pHG4.
119 To construct the integrating plasmid pJH1R4 (Figure 1A), pSET152(28) was cut with
120 AatII and PvuI to remove the ϕC31 *attP* site and integrase gene. R4 phage lysate was
121 used as template in a PCR with the primers pJH1R4-for and pJH1R4-rev to amplify
122 the R4 *attP* site and integrase coding region. The PCR product was joined to the
123 AatII-PvuI fragment from pSET152 via In-Fusion cloning.
124 The plasmid pHG1 (Figure 1B) was used as template to amplify *attB*-containing
125 sequences for *in vitro* recombination assays. This plasmid was originally constructed
126 for the expression of *EryF*. The *eryF* gene was amplified from *Saccharopolyspora*
127 *erythraea* BIOT-0666 genomic DNA using the primer pair pHG1A-for/pHG1A-rev, and
128 inserted by In-Fusion cloning into pBF20(27) cut with NheI and PacI to form the
129 plasmid pHG1A. The 3785 bp fragment containing the ϕC31 *int/attP* and hygromycin
130 resistance gene was amplified from plasmid pBF27C(27), using the primer pair
131 pHG1-for and pHG1-rev. Plasmid pHG1A was digested with XbaI and NheI, and the
132 5668 bp fragment was ligated with the 3785 bp PCR fragment from pBF27C by
133 In-Fusion to give the plasmid pHG1.

134 ***In Vitro* Recombination Assays**

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

145 *attB*^{Sp} (30) (Figure 2). R4-attB-Am-for also had a sequence identical to the 3' end of
146 ActII-orf4 element from the template plasmid pHG1 (Figure 1). Forward primers
147 TG1-attB-Sa-for and R4-attB-Sp-for were used to create positive control
148 recombination substrates containing the TG1 and R4 *attB* sites originally found in *S.*
149 *avermitilis*(29) and *S. parvulus*(30) respectively. The reverse primer used to generate
150 all the *attB*-containing substrates (attB-rev) was located within the *hyg* gene of pHG1;
151 the amplified products were 1627 bp (TG1 *attB*^{Am}), 1035 bp (TG1 *attB*^{Sa}), 1854 bp (R4
152 *attB*^{Am}) and 1855 bp (R4 *attB*^{Sp}). The DNA fragments containing the *attP* sites were
153 prepared as follows; the TG1-*attP* fragment (2471 bp) was amplified using the primer
154 pair TG1-attP-for/TG1-attP-rev with pHG4 as the template, and the R4-*attP* fragment
155 (990 bp) was amplified using the primer pair R4-attP-for/R4-attP-rev with pJH1R4 as
156 the template (Figure 1). Note that other than the *attB* and *attP* sites, none of the
157 substrates contained any DNA that should interact specifically with the integrases.
158 Moreover, each fragment was designed to be easily identifiable by molecular weight.
159 The integrases were purified as described previously(31, 32). All recombination
160 reactions were in 20 µl final volume. Recombination reactions of TG1 substrates were
161 carried out in TG1 RxE buffer (20 mM Tris [pH 7.5], 25 mM NaCl, 1 mM dithiothreitol
162 [DTT], 10 mM spermidine, 10 mM EDTA, 0.1 mg/ml bovine serum albumin [BSA])(33),
163 and recombination reactions of R4 substrates were carried out in buffer containing
164 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM spermidine, 5 mM $CaCl_2$ and 50 mM
165 DTT(32). Integrase was added at the concentrations indicated. Recombination
166 substrates were used at 50 ng each per reaction. Reactions were incubated at 30°C
167 overnight, and then heated (10 min, 75°C) to denature integrase. The reaction
168 mixtures were loaded on a 0.8% agarose gel in Tris/Borate/EDTA (TBE) buffer (90 mM
169 Tris base, 90 mM boric acid and 2 mM EDTA) containing ethidium bromide for
170 electrophoretic separation.

171

172 ## Results

173 **Identification of possible *attB* sequences from the genome of *Amycolatopsis***

174 ***marina* DSM45569**

175 The sequences of *attB* sites recognised by a variety of integrases (ϕC31(34), ϕJoe(35),
176 Bxb1(36), R4(32), SPBc(37), SV1(38), TG1(29) and TP901(39)) were used in BLAST
177 searches of the genome sequence of *Amycolatopsis marina* DSM45569 (NCBI
178 Genome Database NZ_FOKG00000000) (Table 2). The most significant hits for R4 and
179 TG1 *attB* sites had the highest identities and lowest *E*-value. The predicted R4 *attB*
180 site is located within a gene predicted to encode a fatty-acyl-CoA synthase
181 (SFB62308.1) and the TG1 *attB* site is located within a gene predicted to encode a
182 putative succinyldiaminopimelate transaminase (WP_091671332.1). The BLAST
183 analysis was extended to other species of *Amycolatopsis* to assess the conservation
184 of these *attB* sites in the genus (Figure 2). Both R4 and TG1 *attB* sites were highly
185 conserved relative to the *attB* sites originally identified from *S. parvulus*(30) and *S.*
186 *avermitilis*(29) (84% for R4 and 62% for TG1).

187 ***A. marinum attB*-like sequences for TG1 and R4 are both active in *in vitro***
188 **recombination**

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

201 ***In vivo* integration**

202 *A. marina* DSM45569 was unable to grow in the presence of apramycin, so
203 integrating plasmids pHG4 and pJH1R4, containing the apramycin resistance

204 determinant *aac3(IV)*, were constructed. Following the standard *Streptomyces*
205 conjugation protocol (see Materials and Methods), a frequency of approximately 160
206 exconjugants/10^8 spores was obtained for transfer of pHG4 (encoding TG1 integrase),
207 while the conjugation efficiency of pJH1R4 (R4 integrase) was only 20
208 exconjugants/10^8 spores (Table 3). For each integration, six exconjugants were picked
209 at random and streaked on MS agar containing apramycin. Genomic DNA was then
210 prepared and used as the template in PCR reactions, in which the primer pairs of
211 TG1-attL-Am-for/rev and R4-attL-Am-for/rev were used to test for the occurrence of
212 recombination at the expected TG1 and R4 *attB* sites (Figure 4). All PCR reactions
213 using exconjugants as templates gave the expected band sizes. Sequencing (GATC,
214 Germany) of the PCR products with the primers TG1-attL-Am-for and R4-attL-Am-for
215 confirmed that the plasmids had integrated into the predicted *attB* sites for TG1 or
216 R4 integrase within *A. marina* DSM45569 (Figure 5).

217 **Discussion**

218 The lack of effective genetic engineering tools is considered one of the greatest
219 hindrances in the search for novel natural products from rare Actinomycetes(40-42).
220 Previous studies in rare Actinomycete species have largely focused on the use of the
221 well-characterised ϕC31-based integration vectors, and have mostly overlooked tools
222 based on other phage integrases(43-45). Additionally, the conjugation methods used
223 widely in *Streptomyces* gene transfer because of their ease have shown little success
224 in rare Actinomycetes, including species in the genus *Amycolatopsis*, so
225 electroporation has been the long-preferred method of gene transfer for species in
226 this genus(5, 9, 46, 47). However, the growing interest in the use of serine integrases
227 for synthetic biology applications(11) has led to further research into expanding the
228 pool of available enzymes and their potentials as genetic tools(48-50). Therefore,
229 within this study, we explored whether integrating vectors based on eight different
230 serine integrases could be employed for genetic engineering of *A. marina* DSM45569.
231 Sequence analysis of the *A. marina* DSM45569 genome identified close matches to
232 the *attB* sites used by TG1 and R4 integrases. Although conjugation frequencies were

233 low, integrating plasmids based on the TG1 and R4 recombination systems
234 successfully integrated into the expected *attB* sites in *A. marina* DSM45569.
235 Conservation between the *attB* sites for TG1 and R4 in a number of *Amycolatopsis*
236 species is high, suggesting that plasmids with the integration systems from these
237 phages should be widely useful in this genus.
238
239 As is common with serine integrase-mediated recombination, the *attB* sites in *A.*
240 *marina* are located within open reading frames and potentially disrupt the gene. The
241 TG1 $attB^{Am}$ site is located within a gene predicted to encode a putative
242 succinyldiaminopimelate transaminase (WP_091671332.1), and the R4 $attB^{Am}$ site is
243 located within a gene predicted to encode a fatty-acyl-CoA synthase (SFB62308.1).
244 Compared to the wild-type (unintegrated) strain, the strains with integrated pHG4 or
245 pJH1R4 did not show any difference in growth. However, further study is required to
246 investigate the effects of TG1 or R4 plasmid recombination on both primary and
247 secondary metabolism as, for example, the integration of ϕC31 integrase-based
248 plasmids has been shown to have pleiotropic effects on bacterial physiology(51).
249
250 Currently, the two most commonly used methods of bacterial gene transfer are
251 conjugation and electroporation, both of which come with several advantages and
252 disadvantages. Electroporation involves the introduction of pores within bacterial
253 membranes via an electric current to allow mass, unrestricted transfer of genetic
254 material into species(52, 53). The efficiency of transformation, however, is
255 species-dependent(52). Unlike electroporation, conjugation is not limited by the size
256 of vectors that can be transformed, and has been used successfully to transfer entire
257 genomes in *E. coli*(54). Additionally, conjugation uses the transfer of DNA as a single
258 strand that is relatively insensitive to the majority of the restriction systems of the
259 cells(55). Thus conjugation may have advantages over electroporation. In this study,
260 we successfully integrated a plasmid into the *attB* sites for TG1 and R4 integrases by
261 conjugation, thus supplementing the potential gene transfer methods that could be
262 used in the genus *Amycolatopsis*.

263

264 In conclusion, we have identified highly conserved sequences of the *attB* sites for
265 TG1 and R4 integrases within the genus *Amycolatopsis* and demonstrated their use in
266 conjugative DNA transfer. The *A. marina* DSM45569 *attB* sites showed slightly lower
267 recombination efficiencies *in vitro* than the previously identified *attB* sites from
268 *Streptomyces* spp. However, this slight reduction is not enough by itself to explain
269 the order of magnitude reductions in conjugation frequencies observed with *A.*
270 *marina* compared to *Streptomyces* spp. (Table 3). The conjugation frequencies might
271 be increased by optimising conjugation conditions. Alternatively, efficiently used *attB*
272 sites for the widely used vectors, such as those based on phiC31 *int/attP* could be
273 incorporated into the *Amylcolatopsis* genome using TG1 or R4 integrating plasmids as
274 described here. In short, this work shows that integrative vectors are viable and
275 promising tools for the genetic engineering of rare Actinomycetes.

276 Acknowledgments

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

280 References

281 1. Jones SE, Ho L, Rees CA, Hill JE, Nodwell JR, Elliot MA. 2017. Streptomyces exploration is
282 triggered by fungal interactions and volatile signals. Elife 6:e21738.
283 2. Zarins-Tutt JS, Barberi TT, Gao H, Mearns-Spragg A, Zhang L, Newman DJ, Goss RJM. 2016.
284 Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating
285 and upregulating the biosynthesis of bacterial cryptic or silent natural products. Natural
286 product reports 33:54-72.
287 3. Jose PA, Jebakumar SRD. 2013. Non-streptomycete actinomycetes nourish the current
288 microbial antibiotic drug discovery. Frontiers in microbiology 4.
289 4. Jung H-M, Kim S-Y, Moon H-J, Oh D-K, Lee J-K. 2007. Optimization of culture conditions and
290 scale-up to pilot and plant scales for vancomycin production by Amycolatopsis orientalis.
291 Applied microbiology and biotechnology 77:789-795.
292 5. Li C, Liu X, Lei C, Yan H, Shao Z, Wang Y, Zhao G, Wang J, Ding X. 2017. RifZ (AMED_0655) is a
293 pathway-specific regulator for rifamycin biosynthesis in Amycolatopsis mediterranei. Applied
294 and Environmental Microbiology 83:e03201-16.
295 6. Hashizume H, Iijima K, Yamashita K, Kimura T, Wada S-i, Sawa R, Igarashi M. 2017. Valgamicin
296 C, a novel cyclic depsipeptide containing the unusual amino acid cleonine, and related

297 valgamicins A, T and V produced by Amycolatopsis sp. ML1-hF4. The Journal of
298 Antibiotics:ja2017135.
299 7. Li X, Wu X, Zhu J, Shen Y. 2017. Amexanthomycins A–J, pentangular polyphenols produced by
300 Amycolatopsis mediterranei S699Δ rifA. Applied Microbiology and Biotechnology:1-14.
301 8. Schwarz PN, Buchmann A, Roller L, Kulik A, Gross H, Wohlleben W, Stegmann E. 2017. The
302 Immunosuppressant Brasilicardin: Determination of the Biosynthetic Gene Cluster in the
303 Heterologous Host Amycolatopsis japonicum. Biotechnology Journal.
304 9. Kumari R, Singh P, Lal R. 2016. Genetics and genomics of the genus Amycolatopsis. Indian
305 journal of microbiology 56:233-246.
306 10. Colloms SD, Merrick CA, Olorunniji FJ, Stark WM, Smith MC, Osbourn A, Keasling JD, Rosser SJ.
307 2013. Rapid metabolic pathway assembly and modification using serine integrase site-specific
308 recombination. Nucleic acids research 42:e23-e23.
309 11. Fogg PC, Colloms S, Rosser S, Stark M, Smith MC. 2014. New applications for phage integrases.
310 Journal of molecular biology 426:2703-2716.
311 12. Smith MC. 2015. Phage-encoded serine integrases and other large serine recombinases, p
312 253-272, Mobile DNA III. American Society of Microbiology.
313 13. Sosio M, Giusino F, Cappellano C, Bossi E, Puglia AM, Donadio S. 2000. Artificial chromosomes
314 for antibiotic-producing actinomycetes. Nature biotechnology 18:343-345.
315 14. Hong Y, Hondalus MK. 2008. Site-specific integration of Streptomyces ΦC31 integrase-based
316 vectors in the chromosome of Rhodococcus equi. FEMS microbiology letters 287:63-68.
317 15. Saha S, Zhang W, Zhang G, Zhu Y, Chen Y, Liu W, Yuan C, Zhang Q, Zhang H, Zhang L. 2017.
318 Activation and characterization of a cryptic gene cluster reveals a cyclization cascade for
319 polycyclic tetramate macrolactams. Chemical Science 8:1607-1612.
320 16. Thorpe HM, Smith MC. 1998. In vitro site-specific integration of bacteriophage DNA catalyzed
321 by a recombinase of the resolvase/invertase family. Proceedings of the National Academy of
322 Sciences 95:5505-5510.
323 17. Rutherford K, Van Duyne GD. 2014. The ins and outs of serine integrase site-specific
324 recombination. Current opinion in structural biology 24:125-131.
325 18. Baltz RH. 2012. Streptomyces temperate bacteriophage integration systems for stable genetic
326 engineering of actinomycetes (and other organisms). Journal of industrial microbiology &
327 biotechnology 39:661-672.
328 19. Stegmann E, Pelzer S, Wilken K, Wohlleben W. 2001. Development of three different gene
329 cloning systems for genetic investigation of the new species Amycolatopsis japonicum
330 MG417-CF17, the ethylenediaminedisuccinic acid producer. Journal of biotechnology
331 92:195-204.
332 20. Lei X, Zhang C, Jiang Z, Li X, Shi Y, Liu M, Xie Y, Wang L, Hong B. 2017. Complete genome
333 sequence of Amycolatopsis orientalis CPCC200066, the producer of norvancomycin. Journal
334 of biotechnology 247:6-10.
335 21. Li C, Zhou L, Wang Y, Zhao G, Ding X. 2017. Conjugation of ϕBT1-derived integrative plasmid
336 pDZL802 in Amycolatopsis mediterranei U32. Bioengineered 8:549-554.
337 22. Bian J, Li Y, Wang J, Song F-H, Liu M, Dai H-Q, Ren B, Gao H, Hu X, Liu Z-H. 2009.
338 Amycolatopsis marina sp. nov., an actinomycete isolated from an ocean sediment.
339 International journal of systematic and evolutionary microbiology 59:477-481.
340 23. MacNeil DJ. 1988. Characterization of a unique methyl-specific restriction system in

341 Streptomyces avermitilis. Journal of bacteriology 170:5607-5612.
342 24. Shirling Et, Gottlieb D. 1966. Methods for characterization of Streptomyces species1.
343 International Journal of Systematic and Evolutionary Microbiology 16:313-340.
344 25. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. 2000. Practical Streptomyces Genetics.
345 The John Innes Foundation, Norwich.
346 26. Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor
347 Laboratory Press, New York.
348 27. Fayed B, Ashford DA, Hashem AM, Amin MA, El Gazayerly ON, Gregory MA, Smith MC. 2015.
349 Multiplexed integrating plasmids for engineering the erythromycin gene cluster for
350 expression in Streptomyces and combinatorial biosynthesis. Applied and environmental
351 microbiology:AEM. 02403-15.
352 28. Wilkinson CJ, Hughes-Thomas ZA, Martin CJ, Bohm I, Mironenko T, Deacon M, Wheatcroft M,
353 Wirtz G, Staunton J, Leadlay PF. 2002. Increasing the efficiency of heterologous promoters in
354 actinomycetes. Journal of molecular microbiology and biotechnology 4:417-426.
355 29. Morita K, Yamamoto T, Fusada N, Komatsu M, Ikeda H, Hirano N, Takahashi H. 2009. The
356 site-specific recombination system of actinophage TG1. FEMS microbiology letters
357 297:234-240.
358 30. Shirai M, Nara H, Sato A, Aida T, Takahashi H. 1991. Site-specific integration of the
359 actinophage R4 genome into the chromosome of Streptomyces parvulus upon lysogenization.
360 Journal of bacteriology 173:4237-4239.
361 31. Rowley PA, Smith MC, Younger E, Smith MC. 2008. A motif in the C-terminal domain of ϕC31
362 integrase controls the directionality of recombination. Nucleic acids research 36:3879-3891.
363 32. Miura T, Hosaka Y, Yan-Zhuo Y, Nishizawa T, Asayama M, Takahashi H, Shirai M. 2011. In vivo
364 and in vitro characterization of site-specific recombination of actinophage R4 integrase. The
365 Journal of general and applied microbiology 57:45-57.
366 33. Morita K, Morimura K, Fusada N, Komatsu M, Ikeda H, Hirano N, Takahashi H. 2012.
367 Site-specific genome integration in alphaproteobacteria mediated by TG1 integrase. Applied
368 microbiology and biotechnology 93:295-304.
369 34. Combes P, Till R, Bee S, Smith MC. 2002. The Streptomyces genome contains multiple
370 pseudo-attB sites for the ϕC31-encoded site-specific recombination system. Journal of
371 bacteriology 184:5746-5752.
372 35. Fogg PC, Haley JA, Stark WM, Smith MC. 2016. Genome integration and excision by a new
373 Streptomyces bacteriophage, ϕJoe. Applied and environmental microbiology:AEM. 02767-16.
374 36. Ghosh P, Pannunzio NR, Hatfull GF. 2005. Synapsis in phage Bxb1 integration: selection
375 mechanism for the correct pair of recombination sites. Journal of molecular biology
376 349:331-348.
377 37. Lazarevic V, Düsterhöft A, Soldo B, Hilbert H, Mauel C, Karamata D. 1999. Nucleotide
378 sequence of the Bacillus subtilis temperate bacteriophage SPβc2. Microbiology
379 145:1055-1067.
380 38. Fayed B, Younger E, Taylor G, Smith MC. 2014. A novel Streptomyces spp. integration vector
381 derived from the S. venezuelae phage, SV1. BMC biotechnology 14:51.
382 39. Christiansen B, Johnsen M, Stenby E, Vogensen F, Hammer K. 1994. Characterization of the
383 lactococcal temperate phage TP901-1 and its site-specific integration. Journal of bacteriology
384 176:1069-1076.

385 40. Li X, Zhou X, Deng Z. 2003. Vector systems allowing efficient autonomous or integrative gene
386 cloning in Micromonospora sp. strain 40027. Applied and environmental microbiology
387 69:3144-3151.
388 41. Marcone GL, Carrano L, Marinelli F, Beltrametti F. 2010. Protoplast preparation and reversion
389 to the normal filamentous growth in antibiotic-producing uncommon actinomycetes. The
390 Journal of antibiotics 63:83-88.
391 42. Wolf T, Gren T, Thieme E, Wibberg D, Zemke T, Pühler A, Kalinowski J. 2016. Targeted genome
392 editing in the rare actinomycete Actinoplanes sp. SE50/110 by using the CRISPR/Cas9 System.
393 Journal of biotechnology 231:122-128.
394 43. Wagner N, Oßwald C, Biener R, Schwartz D. 2009. Comparative analysis of transcriptional
395 activities of heterologous promoters in the rare actinomycete Actinoplanes friuliensis. Journal
396 of biotechnology 142:200-204.
397 44. Anzai Y, Iizaka Y, Li W, Idemoto N, Tsukada S-i, Koike K, Kinoshita K, Kato F. 2009. Production of
398 rosamicin derivatives in Micromonospora rosaria by introduction of d-mycinose biosynthetic
399 gene with ΦC31-derived integration vector pSET152. Journal of industrial microbiology &
400 biotechnology 36:1013.
401 45. Kim D-Y, Hwang Y-I, Choi S-U. 2011. Cloning of metK from Actinoplanes teichomyceticus
402 ATCC31121 and effect of its high expression on antibiotic production. Journal of microbiology
403 and biotechnology 21:1294-1298.
404 46. Shen Y, Huang H, Zhu L, Luo M, Chen D. 2012. Type II thioesterase gene (ECO-orf27) from
405 Amycolatopsis orientalis influences production of the polyketide antibiotic, ECO-0501 (LW01).
406 Biotechnology letters 34:2087-2091.
407 47. Lee K, Lee B, Ryu J, Kim D, Kim Y, Lim SK. 2016. Increased vancomycin production by
408 overexpression of MbtH-like protein in Amycolatopsis orientalis KFCC10990P. Letters in
409 applied microbiology 63:222-228.
410 48. Mandali S, Dhar G, Avliyakulov NK, Haykinson MJ, Johnson RC. 2013. The site-specific
411 integration reaction of Listeria phage A118 integrase, a serine recombinase. Mobile DNA
412 4:2-2.
413 49. Yoon B, Kim I, Nam J-A, Chang H-I, Ha CH. 2016. In vivo and in vitro characterization of
414 site-specific recombination of a novel serine integrase from the temperate phage EFC-1.
415 Biochemical and biophysical research communications 473:336-341.
416 50. Yang L, Nielsen AA, Fernandez-Rodriguez J, McClune CJ, Laub MT, Lu TK, Voigt CA. 2014.
417 Permanent genetic memory with> 1-byte capacity. Nature methods 11:1261-1266.
418 51. Talà A, Damiano F, Gallo G, Pinatel E, Calcagnile M, Testini M, Fico D, Rizzo D, Sutera A,
419 Renzone G. 2018. Pirin: A novel redox-sensitive modulator of primary and secondary
420 metabolism in Streptomyces. Metabolic engineering 48:254-268.
421 52. Bekker V, Dodd A, Brady D, Rumbold K. 2014. Tools for metabolic engineering in
422 Streptomyces. Bioengineered 5:293-299.
423 53. Cadoret F, Soscia C, Voulhoux R. 2014. Gene transfer: transformation/electroporation.
424 Pseudomonas Methods and Protocols:11-15.
425 54. Isaacs FJ, Carr PA, Wang HH, Lajoie MJ, Sterling B, Kraal L, Tolonen AC, Gianoulis TA, Goodman
426 DB, Reppas NB. 2011. Precise manipulation of chromosomes in vivo enables genome-wide
427 codon replacement. Science 333:348-353.

428 55. Dominguez W, O'Sullivan DJ. 2013. Developing an efficient and reproducible
429 conjugation-based gene transfer system for bifidobacteria. Microbiology 159:328-338.

430

431 ## Figures

432 **Figure 1**. Plasmids used in this study. The primer binding sites are indicated. (A)
433 Integrating plasmids pHG4 and pJH1R4; (B) PCR template plasmid pHG1.

434 **Figure 2**. Alignment of R4 and TG1 *attB* sites in *A. marina* DSM45569 and other
435 *Amycolatopsis* species. GenBank accession nos. of DNA sequences: *A. alba*
436 (NMQU01000019.1), *A. antarctica* (NKYE01000021.1), *A. azurea* (MUXN01000005.1),
437 *A. kentuckyensis* (MUMI01000226.1), *A. lurida* (FNTA01000004.1), *A. nigrescens*
438 (ARVW01000001.1), *A. niigatensis* (PJMY01000003.1), *A. orientalis*
439 (ASXH01000007.1), *A. pretoriensis* (MUMK01000092.1), *A. regifaucium*
440 (LQCI01000034.1), *A. rubida* (FOWC01000001.1), *A. sacchari* (FORP01000010.1), *A.*
441 *thailandensis* (NMQT01000219.1), *A. thermoflava* (AXBH01000004.1), *A.*
442 *tolypomycina* (FNSO01000004.1), *A. xylanica* (FNON01000002.1), *Streptomyces*
443 *avermitilis* (NC_003155.5) and *S. parvulus* (CP015866.1).

444 **Figure 3**. (A) Design and synthesis of DNA substrate *attB*. The *attB* sites were fused at
445 the end of forward primers, to amplify a sequence flanked with *attB* from an
446 unrelated DNA template, pHG1 in this case. (B) Design and synthesis of DNA
447 substrate *attP*. The *attP* sites were amplified directly from integrating vectors
448 carrying TG1/*attP* (pHG4) or R4/*attP* (pJH1R4). (C) Recombination substrates and
449 their expected products. (D) *In vitro* recombination between DNA fragments
450 containing TG1 $attB^{Am}$ (1627 bp) and TG1 *attP* (2471 bp; left), and R4 $attB^{Am}$ (1854 bp)
451 and R4 *attP* (990 bp; right). The expected products of the TG1 integrase-mediated
452 reaction were a 4.1 kb DNA fragment containing the $attR^{Am}$ site, and a 53 bp
453 fragment containing $attL^{Am}$ (not observed). For the R4 integrase recombination
454 reaction, the expected products were a 2.8 kb fragment containing $attR^{Am}$, and a 51
455 bp $attL^{Am}$ fragment (not observed). (E) *In vitro* recombination between DNA
456 fragments containing TG1 $attB^{Sa}$ (1035 bp) and TG1 *attP* (2471 bp; left), and R4 $attB^{Sp}$
457 (1855 bp) and R4 *attP* (990 bp; right). The expected products were a 3.5 kb fragment
458 containing $attR^{Sa}$ for the TG1 reaction, and a 2.8 kb fragment containing $attR^{Sp}$ for the
459 R4 reaction. M: Fast DNA Ladder (NEB, USA).

460 **Figure 4**. PCR confirmation of site-specific integration in the exconjugants. (A) Design
461 of primers. (B) PCR (using primers TG1-attL-Am-for/rev) of the expected TG1
462 *attL*-containing fragment from *A. marina* DSM45569:pHG4. (C) PCR (using primers
463 R4-attL-Am-for/rev) of the expected R4 *attL*-containing fragment from *A. marina*
464 DSM45569:pJH1R4. M: Fast DNA Ladder. Colonies 1 to 6 are independent
465 exconjugants.
466 **Figure 5.** The insertion sites of TG1 and R4 integration plasmids in *A. marina*
467 DSM45569. Sequencing (using primers TG1-attL-Am-for or R4-attL-Am-for) of PCR
468 products containing *attL* from exconjugants validated the site-specific recombination
469 of the TG1 and R4 *attB* sites in *A. marina* DSM45569 after introduction of pHG4 or
470 pJH1R4, respectively.
471

472 **Figure 1**

473

474 **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
```

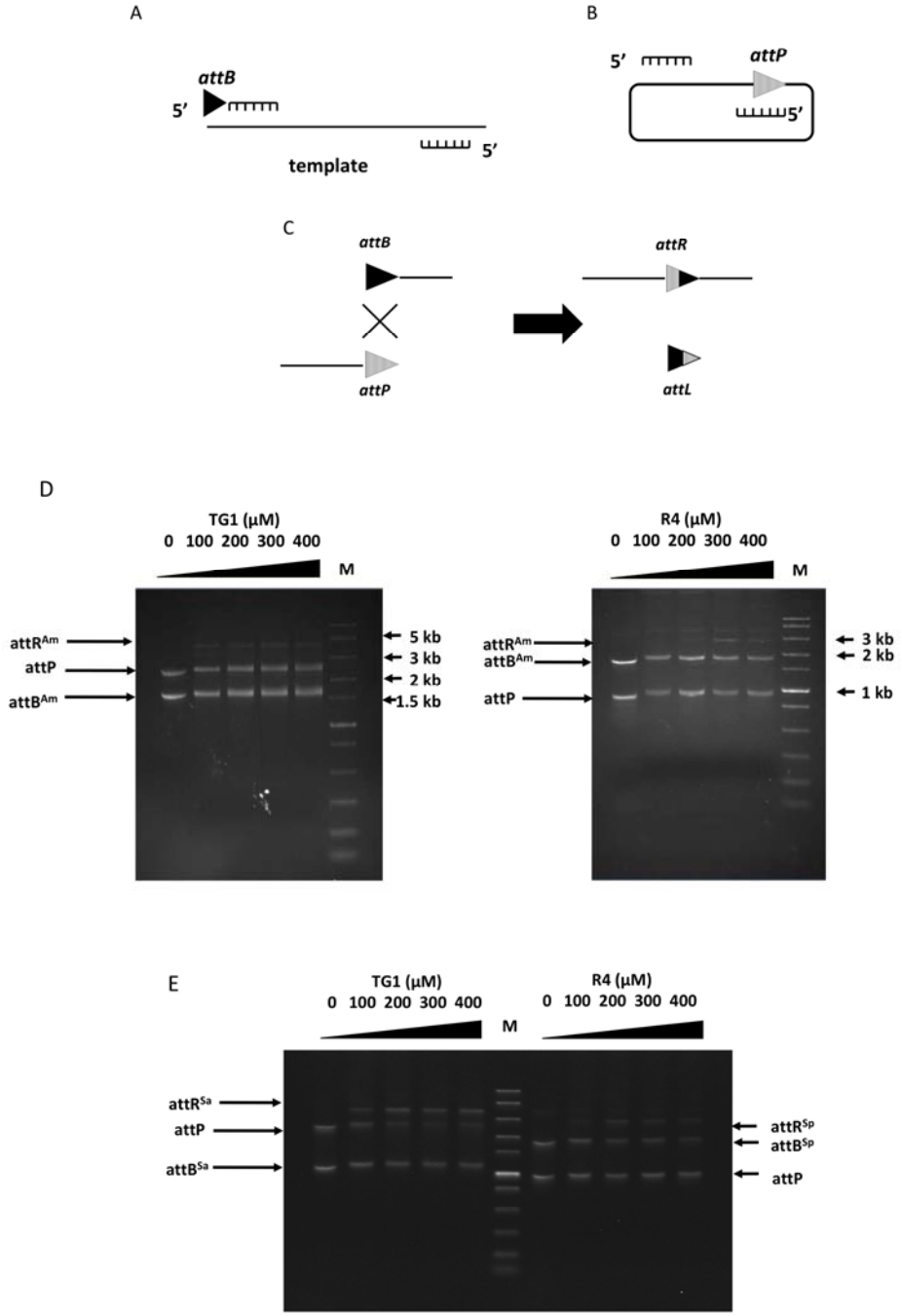
B. TG1 *attB* sites

```
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
```

475

476

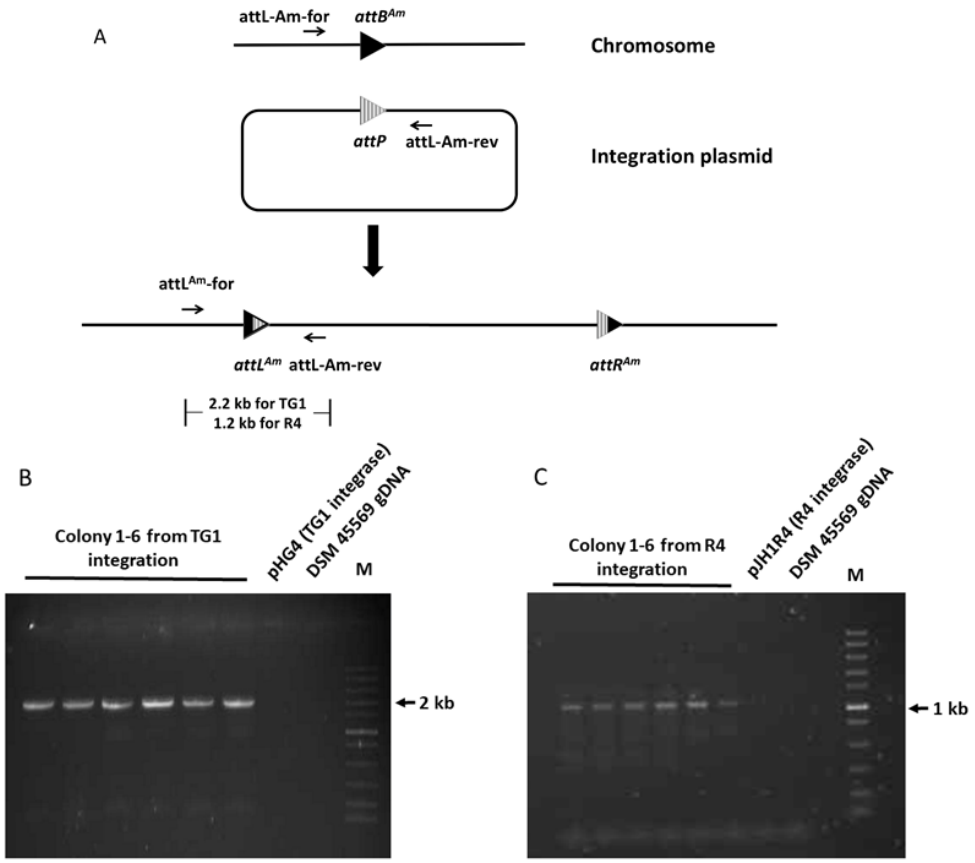
477 **Figure 3.**



478

479

480 **Figure 4.**



481

482

483 **Figure 5.**

A

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

B

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

484

485

486 **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

487 The possible *attB* sequences identified from the genome sequence of *Amycolatopsis*
488 *marina* DSM45569 are shown underlined.
489

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

Integrase	***attB* sites and the best hit from BLAST**	**Homology (%)**	***E*-value**
ϕ31	S.coelicolor CGGTGCGGGGGCAGGGCGTGCCCTTGGGTCCGCGGG----CGCGAGTCCACC hit from A.marina GGGTGCGGGGCACAGGGCGTGCCC---GATCCGCGGATCGCCACGAACGCCGA-	41	0.015
ϕJoe	S.venezuelae ATCTGGATGTGGGTGTCCATGTGCGGGCAGACGCCGCAGTCGAAGCACGG hit from A.marina AGGTGTCCATAGGTGTGGGCGGACCAGTCCGAGCCGCAGTCGAAGCAGCC	30	0.60
Bxb1	Mycobacterium smegma TCGGCCGGCTTGTCGACGACGGCGGTCTCCG-TCGTCAGGATCATCCGGGC hit from A.marina TCGGCCTGCTGCTCGACGACGGCGGCTGTGGCCGTGGTGGGTCCGCG-	48	0.014
R4	S.parvulus AGTTGCCCATCACCATGCCGAAGCAGTGGTAGAAGGGCACCGGCACACAC hit from A.marina GGTTGCCCATCACCATGCCGAAGCAGTGATAGAAGGGAACCGGGATGCAG	84	3e-11
SPBc	No hit		
SV1	S.venezuelae ATCAGGGCGGTCAGGCCG--TAGATGTGGAAGAACGGCAGCACGGCGAGGAC hit from A.marina --CTTCCGGACTCGGCTGGTGGCTGTGCCAGGCCGGCAGCACGGCGAGGCC	32	0.17
TG1	S.avermitilis TCGATCAGCTCCGCGGGCAAGACCTTCTCCTTCACCGGGTGGAAGGTCGG hit from A.marina TCGATCTCCAGTGCGGGCAAGACCTTCAACTCCACCGGCTGGAAGATCGG	62	0.001
TP901	Lactococcus lactis s ATGCCAACACAATTAACATCTCAATCAAGCAAATGCTTTTGCTTTTTTGC hit from A.marina ATCCGGCCTACCGGAACATCTCAATCACCGAACTCGCCCACCGCTGTGGATTC	25	8.1

493 **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

494

495