Analysis of a genomic island housing genes for DNA S-modification system in *Streptomyces lividans* 66 and its counterparts in other distantly related bacteria

Xinyi He\(^1\), Hong-Yu Ou\(^1\), Yu Qing\(^1\), Xiufen Zhou\(^1\), Jun Wu\(^1\), Jingdan Liang\(^1\), Wei Zhang\(^2\), Kumar Rajakumar\(^{3,4}\) and Zixin Deng\(^1\)*

\(^1\) Laboratory of Microbial Metabolism and School of Life Science & Biotechnology, Shanghai Jiaotong University, Shanghai 200030, People’s Republic of China.
\(^2\) Division of Biology, Illinois Institute of Technology, Chicago, Illinois 60616, United States of America.
\(^3\) Department of Infection, Immunity and Inflammation, Leicester Medical School, University of Leicester, Leicester LE1 9HN, United Kingdom.
\(^4\) Department of Clinical Microbiology, University Hospitals of Leicester NHS Trust, Leicester LE1 5WW, United Kingdom.

*For Correspondence.

E-mail zxdeng@sjtu.edu.cn; Tel. (+86) 21 62933404; Fax (+86) 21 6293 2418.
Summary

The complete sequence (92,770-bp) of a genomic island (GI) named SLG from *Streptomyces lividans* 66, encoding a novel DNA S-modification system (*dnd*), was determined. Its overall G+C content was 67.8%, lower than those of three sequenced *Streptomyces* genomes. Among eighty-five predicted open reading frames (ORFs) in SLG, twenty-two ORFs showed little homology with previously known proteins. SLG displays a mosaic structure composed of four modules, indicative of multiple recombination events in its formation. Spontaneous excision and circularization of SLG was observed, and the excision rate appeared to be induced at least five-fold by MNNG exposure. Using constructed mini-islands of SLG, we demonstrated that Slg01, a P4-like integrase, was sufficient to promote SLG integration, excision and circularization. Eleven counterpart *dnd* clusters, which also mapped to GIs in ten chromosomes and a plasmid, were found in taxonomically unrelated bacterial species from various geographic niches. Additionally, ca. 10% of actinomycetes were found to possess a *dnd* cluster in a survey involving 74 strains. Comparison of *dnd* clusters in the twelve bacteria strongly suggests that these *dnd*-bearing elements might have evolved from a common ancestor similar to plasmid-originated chromosome II of *Pseudoalteromonas haloplanktis* TAC125.
**Introduction**

*Streptomyces* species are soil-dwelling filamentous bacteria that produce most known natural antibiotics as well as many other secondary metabolites and secreted enzymes of economic and industrial importance (Hopwood, 2007). Considerable phenotypic variation is commonly observed in many genera of actinomycetes. This is predominantly attributed to high levels of chromosomal instability caused by homologous or illegitimate recombination events that frequently result in deletions, insertions, amplifications and/or rearrangements (Gunes *et al.*, 1999; Volff and Altenbuchner, 1998). Horizontal gene transfer events contribute to further genome instability via the integration of unstable alien mobile genetic elements into the *Streptomyces* linear chromosome. Examples of acquired elements include ISs, transposons, prophages, conjugative plasmids and genomic islands (GIs) (Bentley *et al.*, 2002; Choulet *et al.*, 2006; Ikeda *et al.*, 2003). The ‘mobilome’ (mobile genome) (Ou *et al.*, 2005) of an individual organism has been hypothesized to reflect the bacterium’s lifestyle, pathogenicity, adaptation to particular ecological niches and evolutionary history (Dobrindt *et al.*, 2004). Interestingly, a subset of GIs associated with secondary metabolism has recently been identified in many bacterial species and these are proposed as important players in the moulding of natural product biosynthesis (Dobrindt *et al.*, 2004; Piel *et al.*, 2004). Most integrases encoded by GIs recognize the 3’ termini of tRNA genes as integration hotspots (Ou *et al.*, 2006), whereas some integrases target other conserved loci such as *thdF* (in *Salmonella* SGI1) (Doublet *et al.*, 2005) and *bacA* (in *Streptomyces scabies* pathogenicity island PAI) (Kers *et al.*, 2005). Island-borne integrases or ‘cross-talking’ homologues typically mediate the process of site-specific integration and excision (Hochhut *et al.*, 2006). In some cases, additional excisionases or other auxiliary factors are required (Doublet *et al.*, 2005). In addition, some GIs are readily transmissible under standard laboratory conditions and are consequently thought to have been recently acquired. Examples of mobile GIs include the 500-kb symbiosis island in *Mesorizobium loti* (Sullivan and Ronson, 1998), SXT genomic island in *Vibrio cholerae* (Hochhut and Waldor, 1999) and *clc* genomic island of *Pseudomonas* sp. strain B-13 (Sentchilo *et al.*, 2003a; Sentchilo *et al.*, 2003b). In contrast, many chromosomally integrated GIs seem to have lost transmissibility. Several of these ‘immobile’ structures exhibit highly mosaic content, indicating the likely occurrence of multiple recombination events (Hsiao *et al.*
The sizes of linear *Streptomyces* chromosomes range from the 8.7-Mb chromosome in *S. coelicolor* A3(2) (Bentley et al., 2002) to the 10.1-Mb replicon in *S. scabies* (http://www.sanger.ac.uk/Projects/S_scabies/). These linear *Streptomyces* chromosomes appear to be compartmentalized into ‘core’ and ‘arm’ regions. The central core region contains the essential genes, whereas the arms carry conditionally adaptive genes and species-specific DNA (Bentley et al., 2002). Despite the high level of synteny between the core regions of actinomycete chromosomes, several GIs have been found in these regions (Bentley et al., 2002; Choulet et al., 2006; Ikeda et al., 2003). For example, 12 large insertions were identified in the core region of the *S. coelicolor* chromosome following comparison of its genome with that of a close relative, *S. ambofaciens* (Choulet et al., 2006). Most of these were probably acquired recently because they contain discernible mobility-associated features. Application of the Islander algorithm, which detects GIs adjacent to tRNA sites bounded by direct repeats and containing an integrase gene homologue (Mantri and Williams, 2004), led to the identification of five and three islands in the chromosomes of *S. coelicolor* A3(2) and *S. avermitilis* MA-4680, respectively. A large (325–660-kb), mobile PAI was reported to be conserved among three plant pathogenic *Streptomyces* species (*S. acidoscabies*, *S. scabies* and *S. turgidiscabies*), which encodes a pathogenicity determinant, the phytotoxin thaxtomin (Kers et al., 2005).

Another important example is the ca. 90-kb *S. lividans* SLG, found inserted into the 3’-end of the chromosomal murA1 gene, and absent from *S. coelicolor* A3(2) (Zhou et al., 2004). The SLG island contains a phage φHAU3 resistance gene (φHAU3r) (Zhou et al., 1994b) and a 8.2-kb five-gene cluster involved in site-specific incorporation of sulphur (Zhou et al., 2005) into *S. lividans* DNA (Dyson et al., 1998, Liang et al., 2007). We named this SLG-mediated *in vivo* DNA modification the Dnd phenotype. Although the role of this modification is not yet known, modified DNA is sensitive to *in vitro* oxidative double-strand cleavage and degradation during normal and pulsed-field gel electrophoresis (PFGE) (Zhou et al., 1988).

In order to better understand the function and evolution of SLG in
**Streptomyces**, we sequenced and characterized SLG in *S. lividans* 66. We found that SLG has a mosaic composition and shows features of multiple horizontal acquisition events. We demonstrated that an SLG-encoded P4-type integrase mediated site-specific integration, excision, and circularization of SLG, while native SLG itself is apparently non-transmissible. The *dnd* cluster is widely distributed among distantly related bacterial species. Most strikingly, the identified *dnd*-clusters appear to be borne on large integrated elements, supporting the existence of a diverse 'family' of GIs that carry this novel and intriguing DNA modification system.

**Results**

**SLG, a mosaic-like region characteristic of an unusual genomic island in *S. lividans***

The 93-kb DNA is thought to be inserted into the *S. lividans* 66 chromosome at the 3’ terminus of *murA1* with a recognizable flanking direct repeat (DR) that matched the 15-bp 3’-end of *murA1* (Zhou *et al.*, 2004). Analysis of the SLG sequence revealed an intact integrase gene homologue (*slg01*) and a truncated transposase gene (*slg02*) in the *murA1*-proximal end (Zhou *et al.*, 2004). Eighty-five putative protein-coding genes, *slg01*-slg85, were predicted with FramePlot beta 3.0 (Ishikawa and Hotta, 1999), and found to occupy 76% of the island sequence (Fig. 1B). The results of a BLASTP (Schaffer *et al.*, 2001), Pfam (Finn *et al.*, 2006) and TMHMM (Krogh *et al.*, 2001) searches for predicted proteins and its putative functions are shown in Table S1 (supplementary material). Twenty-two of them showed little homology with proteins in databases. SLG has a markedly lower G+C content (67.8%) than the genome of *S. coelicolor* A3(2) (72.1%), a very close relative of *S. lividans* 66 (Kieser *et al.*, 1992; Leblond *et al.*, 1993), suggesting that SLG could have been acquired laterally from elsewhere. A distinct dinucleotide bias of the island region compared with the complete sequence of *S. coelicolor* A3(2) also strongly supported a foreign origin of SLG (Table 1).

For the convenience of description and discussion, the SLG island was divided into four putative modules defined on the basis of G+C content, clustered functional organization, sequence homology and/or AT-rich valleys (Fig. 1): module I (11.9-kb, 64.3% G+C) harbours a putative recombination hotspot (Fig. 1A; Fig. S1), module II (20.4-kb, 65.4% G+C) putatively encodes a nucleotide-related metabolic pathway.
(green in Fig. 1B; Table S1), module III (46.2-kb, 68.8% G+C) seemed to cover a large syntenic DNA fragment (slg45-71) with 95% full-length nucleotide identity to a region in S. coelicolor A3(2) (SCO3509-34, highlighted in red in Fig. 1B) and module IV (14.2-kb, 65.1% G+C) encompasses the dnd gene cluster (blue in Fig. 1B; Table S1). (Detailed bioinformatics analysis of these four modules is provided in the supplementary results.)

SLG can spontaneously excise from the S. lividans chromosome

As a putative GI, can SLG excise from the S. lividans chromosome? The failure to detect a putatively excised DNA band specific to SLG by PFGE analysis of total DNA of S. lividans 66 (Table 2) prompted us to examine this possibility with a more sensitive PCR method using total DNA of S. lividans 66 as template, and oligonucleotides flanking both sides of attL (Lp1F and Lp1R) and attR (Rp2F and Rp1R) as primers (Table S2, Fig. 2). If SLG excises from the chromosome and circularizes between the attL and attR, primers Rp2F and Lp1R would only amplify a product of 467-bp; Likewise, another pair of primer Lp1F and Rp1R targeting to chromosome fusion between attL and attR would amplify a 725-bp coupled with the SLG excision. Two PCR fragments with the expected sizes were obtained (Fig. 2), demonstrating that the chromosomally-integrated SLG was capable of spontaneous excision from the chromosome of the wild-type 66 at low frequency.

To determine the frequency of spontaneous excision of SLG, we randomly selected 1,015 colonies of S. lividans HXY1 (dndA::aadA) (Zhou et al., 2005) and patched onto spectinomycin-containing medium after growth on SFM medium to identify putative SLG-minus derivatives. None of the colonies exhibited spectinomycin sensitivity, indicating no SLG was excised. However, when spores were treated with MNNG at a concentration of 1 mg/ml, pH 8 at 30°C for 1 hour, as described in (Kieser et al., 2000), we found that 1 out of 200 (0.5%) randomly selected S. lividans HXY1 colonies lost SLG (Fig. S2) via excision at the attB site coupled with the loss of resistance to spectinomycin. Therefore, the exposure to MNNG might have increased the frequency of SLG excision in comparison to growth in standard condition. To estimate the rate of SLG excision, we used quantitative real-time PCR analysis and calculated the ratios of the cells containing excised-SLG verse all cells in the different media including TSBY (34% sucrose), YEME (34%
sucrose), and YMG. The copy number of the excised SLG (as determined by PCR using primers P21F and P21R, Table S2) was compared with the copy number of reference locus (3196th-nt - 3343th-nt upstream of attL, whose copy number was arbitrarily set to 100%) (determined by PCR using primers LC2F and LC2R, Table S2). We observed that the ratios of cells harbouring the excised SLG ranged from 0.016% to 0.027% (Fig. 2C) in different media.

Role of an integrase in site-specific excision and integration of SLG

To localize SLG-borne features that might be required for site-specific integration into and/or excision/self-circularization from the chromosome, we constructed a thermosensitive Streptomyces replicon pSG5-derived (Maas et al., 1998) plasmid (pJTU1514) carrying the attL/attR sites, an intact integrase gene (slg01) and a truncated transposase gene (slg02), but with replacement of all other SLG internal genes by pUC18 carrying an E. coli replicon and an apramycin resistance gene (aac(3)IV) (Fig. 3A), using an approach based on the method proposed by (Ubeda et al., 2003). As a result, we generated a smaller circular molecule with precise site-specific excision via attL/attR as a circular SLG-derived mini-island by simple loop-out of pJTU1515 (Fig. 3B) from pJTU1514. (see supplementary results.)

Next, we tested whether pJTU1515 could be site-specifically integrated into the S. lividans ZX1 core chromosome (“backbone”) at the attB site to form a strain identical to HXY10 (Table 2). By protoplast transformation, pJTU1515 was introduced into ZX1, and a strain (HXY10-1) selected by apramycin resistance was subjected to PCR using five primers designed to anneal in the vicinity of the attL and attR sites: Lp1F and Rp1R annealed to backbone DNA, whereas Lp1R, Rp1F and Rp2F bound to sequences in pJTU1515 (in the directions indicated in Fig. 4). Using primers Lp1F and Lp1R targeting the left junction, a 0.8-kb specific product was amplified from HXY10-1 (Fig. 4A). Similarly, using primers Rp1F and Rp1R that target the right junction, a 0.7-kb specific product was amplified (Fig. 4C). The data suggested that the mini-island had integrated site-specifically at the attB site. In addition, the circularly excised mini-island was detected by PCR with primers Rp2F and Lp1R in HXY10-1-derived DNA (Fig. 4B), this suggested that the mini-island existed dynamically in chromosomally-integrated and non-replicative, episomal forms in the host strain.
Finally, the three elements at the murA1-proximal end of SLG were investigated for their requirement for island excision and integration, including: (i) \textit{slg01 (int)}), encoding an intact integrase homologue; (ii) \textit{slg02 (tnp)}), encoding a likely truncated transposase; and (iii) the 15-bp predicted \textit{attP} site. Plasmids derived from pJTU1515 (Fig. 4; Table 2) were constructed and analyzed: (i) pJTU1520, in which the 5’ 590 nucleotides of \textit{slg01} and the entire \textit{slg02} were removed from pJTU1515, failed to integrate into ZX1 by transformation, but its integration/excision capability could be restored in \textit{trans} by firstly introducing pJTU1522 carrying an intact \textit{slg01 (int)}, resulting in strain HXY12 (Fig. 4). After a round of non-selective growth of HXY12, the replicative but highly unstable pHZ1358-derived pJTU1522 was cured from the mycelium, leading to HXY13 (Fig. 4). PCR with primers Lp1R and Rp2F failed to detect free pJTU1520 in HXY13, although all the expected PCR fragments could be detected in HXY11, HXY12 and HXY13 by PCR using the same primers used for HXY10-1 (Fig. 4A-C). (ii) pJTU1516, in which the 3’ 306 nucleotides of \textit{slg02 (tnp)} were deleted from pJTU1515, was introduced by transformation into ZX1 to generate HXY11, the resultant plasmid was found to be capable of integration, excision and circularization, as pJTU1515 (in HXY10-1). (iii) pJTU1517 was constructed by removing a 175-bp \textit{attP} region from pJTU1515, and was found to be still able to integrate into the ZX1 at a similar rate as pJTU1515, the exconjugants were confirmed by PCR using primer pairs: P1F&R, P2F&R and aac(3)IV-T F&R (Table S2, data not shown), suggesting the availability of a secondary \textit{attP} site other than \textit{attP} (15-bp) locating within the pJTU1517. The secondary attachment site is yet to be determined. Hence, the functional integrase encoded by \textit{slg01 (int)} on SLG was found to be necessary to mediate site-specific integration, excision, and circularisation of the mini-island.

The native SLG seems to be non-transmissible in Streptomyces

To determine whether the SLG island is transmissible between\cite{Schwartz2000} \textit{Streptomyces} strains, an inter-species mating experiment was first performed with \textit{S. lividans} HXY1 (\textit{dndA::aadA}, resistant to streptomycin) harboring a helper conjugative plasmid pJJ101 \cite{Kieser1982} as a donor, and \textit{S. coelicolor} M145-derived recipient strain ZH3 (\textit{SCO3930::aac(3)IV}, resistant to apramycin) \cite{Li2006} as a recipient. No ZH3-derived exconjugants with acquired SLG,
however, could be obtained, as monitored from an intensive screening of ca. 100 streptomycin-apramycin double resistant colonies (with production of pigmental actinorhodin characteristic of the recipient strain ZH3) by PCR amplification using a pair of primers (M1F & M1R) targeting to module I of SLG (exemplified in Fig. S3). Meanwhile, in a second intra-specific mating experiment involving use of *S. lividans* HXY18 (*dndA::aadA, slg10::tsr*) with the helper conjugative plasmid pIJ101 as a donor, and HXY19 (ZX1 derivative with *aac(3)IV* inserted at 17090th nt upstream of the *attL*) as a recipient, no HXY19-specific exconjugants with acquired SLG was detected. It seemed, therefore, that the native SLG may have lost its capabilities of inter-, and even intra-species transmission between *Streptomyces* strains.

*Widespread distribution of dnd clusters among distantly related bacterial species*

Interrogation of bacterial genome sequences (available in GenBank up to October 2006) identified additional likely homologues of the *dnd* cluster in another 11 strains belonging to phylogenetically diverse bacterial species (Fig. 5; Fig. S4). These strains represented species and genera of variable origin and diverse habitats, such as GC-rich *S. lividans* versus AT-rich *Pelagibacter ubique*, soil-dwelling organisms versus marine microbes, non-pathogenic saprophytes versus human pathogens.

Eight of the 12 genomes bearing the *dnd* cluster contained homologues of five genes. Examination of gene order and spacing suggested that *dndA* and *dndB-E* comprised independent transcription units (Fig. 5). Indeed, reverse transcription-PCR on *S. lividans* 66 confirmed that *dndB-E* formed a single operon (Liang et al., unpublished data). DndA protein could function as an IscS-like cysteine desulfurase (Schwartz *et al.*, 2000; You *et al.*, 2007) and is essential for Dnd phenotype (Zhou *et al.*, 2005). The *dndA* gene is divergently transcribed relative to *dndB-E* in Strain A, B, D, I and J (Fig. 5) but lies downstream of and in the same orientation as *dndB-E* in H, K and L (Fig. 5). No putative *dndA* ortholog was identified in the immediate vicinity of *dndB-E* in C, E, F and G (Fig. 5). However, an *iscS* homologue was present elsewhere in these bacterial genomes and the cognate proteins may have served as functional homologues of DndA. We are currently exploiting these data to further examine the functional nature of the Dnd proteins.

Notably, 5 of the 12 strains with a genomic G+C content of less than 56% all
contain another highly conserved three-gene cluster tightly linked with the dnd cluster (Fig. 5D-H). Additionally, the gene designated PSHAb0092 (Fig. 5D) shares 35% amino-acid identity with Pfl_0743 (Fig. 5C). SAV2928 (Fig. 5B) shares 38% identity with RD1_0805 (Fig. 5K), 28% identity with Meso_4564 (Fig. 5L). Finally, at the right flanking region of the dnd loci, Slg81 (Fig. 5A) resembles SAV2934 (Fig. 5B) with 72% amino-acid identity.

To test the prevalence of the dnd cluster in the same genera, we performed a Dnd phenotype survey on 74 actinomycete strains collected from geographically distinct regions. The DNA of five Streptomyces strains were found to possess Dnd phenotype as that of S. lividans 66 and S. avermitilis MA-4680 (S. acrimycini 2236, S. caneus, S. griseoplanes 92023, S. verticillus ATCC 15003 and Streptomyces sp. As-01) (data not shown). We also investigated the linkage between dnd and integrase in these surveyed actinomycete strains by Southern hybridization using slg01 as a probe. Five newly identified Dnd+ strains showed a negative signal, whereas one Dnd- strain (Streptomyces sp. 30214) produced a positive result (data not shown). This observation suggests that dnd-bearing GIs do not necessarily have to be linked with a specific integrase gene, but can be equipped with one or more alternatives (Fig. 5). Dnd phenotype has also been discovered recently in other bacterial genus or species including two E. coli isolates causing blood stream infections in humans (Rajakumar et al., personal communication). Two isolates of Salmonella enterica serovar Livingstone and two isolates of S. enterica serovar Cerro have been identified by normal agarose gel electrophoresis to display the Dnd phenotype (Murase et al., 2004), whose counterpart gene cluster had been isolated (Tiegang Xu, personal communication). Similarly, 11 out of 34 clinical isolates of Pseudomonas aeruginosa have the Dnd phenotype (Romling and Tummler, 2000). These findings suggested that dnd clusters are more common among different strains than had been anticipated based on available genome sequence data.

All dnd-clusters in other 11 sequenced genomes locate on mobile elements

As mentioned above, the S. lividans 66 dnd cluster lies in module IV of the large, mosaic SLG. This discovery prompted us to examine further the genomic context of the dnd clusters in the other 11 completely or partially sequenced bacterial strains. Remarkably, all eleven other identified dnd clusters seemed to lie on mobile genetic
elements, ten within the chromosome and one on a large plasmid (Fig. 5L, Table 1). Sequence analysis showed that these putative dnd-encoding islands shared key features typical of GIs on the basis of G+C content (Column 4, Table 1), dinucleotide bias (Column 5&6, Table 1), context homology and genomic island characteristic elements, such as integration into tRNA, or possession of DRs, integrase and/or transposase (Column 7, Table 1; Fig. 5). The prediction that the dnd-bearing fragments are on genomic islands in *E. coli* B7A and *Hahella chejuensis* KCTC 2396 agrees well with the predication (Ou *et al.*, 2007) and (Jeong *et al.*, 2005), respectively. The fragment encoding the dnd homologue in *S. avermitilis* MA-4680 is also a putative island as predicted by ISLANDPATH (Hsiao *et al.*, 2003). In addition, dnd cluster is also found on the plasmid in *Mesorhizobium* sp. BNC1 (Fig. 5L) or on plasmid-derived chromosome in the *Pseudoalteromonas haloplanktis* TAC125 (Fig. 5D). Collectively, we demonstrated that the dnd clusters identified in all known cases up to now are located on the mobile elements.

**Discussion**

The 93-kb SLG of *S. lividans* 66, absent from its close relative *S. coelicolor* A3(2) was found to integrate into the 3’ terminus of *murA1*, a well conserved locus in three sequenced *Streptomyces* genomes. Like many functional GIs (Williams, 2002), the direct relevance of *att* and *int* for mediating site-specific excision and integration of SLG into the *S. lividans* 66 chromosome was demonstrated using a constructed mini-island with *int* sandwiched between *attL* and *attR* in a temperature-sensitive plasmid. Across the mosaic-like, four-module SLG, the dnd gene cluster is present on module IV, a relatively small portion of this large island. We propose that module I, flanked by *att* sites and including *int*, is the basic mobile entity of SLG. Module II encodes hypothetical proteins involved in a putative nucleotide-related metabolic pathway. Module III shows greater than 95% nucleotide identity to a contiguous region of *S. coelicolor* A3(2); however, two internal ORFs are replaced by a putative IS element, including a transposase gene and two tightly linked genes, *ea31* and *ea59* (known as *φHAU3r*, Zhou *et al.*, 1994a) in six bacterial species (Fig. S5), such organization is reminiscent of the IS element containing *ea31* and *ea59* in *Pseudomonas syringae* pv. *tomato* DC3000 (Alfano *et al.*, 2000). Such a mosaic compilation of diverse functional modules could involve an initial invasion by a
mobile element, containing, for example, a component of module I and followed by recombinational promiscuity in disparate organisms, and suggests that SLG could be an example of how a bacterial host has successfully captured peripatetic genetic information from multiple sources. Thus, given the highly mosaic structure of SLG, we suggest that the present entity represents a relic of multiple past recombination events that may have since lost its native capability of self-transmission.

Horizontal transfer of GIs is often initiated by excision of a linear form from the chromosome to form a circular, mobilizable episome. This process can be induced by a degradable substrate in the growth medium, as in activation of \textit{clc} excision (Sentchilo \textit{et al.}, 2003a); regulated in a cell density-dependent manner, as in excision of ICE\textit{Ml}SymR7A (Ramsay \textit{et al.}, 2006); or more often by activation of the promoter of an independent excisionase or integrase, which also leads to GI integration into the genome of the recipient strain (Doublet \textit{et al.}, 2005; Ramsay \textit{et al.}, 2006; Sentchilo \textit{et al.}, 2003a). Apparently, the excision of SLG falls into the last case as a result of the absolute necessity of the \textit{int} gene (Fig. 4). The currently reported excision rate of the SLG in different media, which was kept at a stable level (0.016%~0.027%) after standard 36 hours of incubation, is a suggestion that the regulation on the integrase is not so dependent on properties of the medium component, such as osmotic pressure (34% sucrose), and/or the availability of the iron (YMG is a low-iron medium compared with YEME). By contrast, the likely up-regulation of GI integrase activity, as detected by the increased SLG excision after MNNG exposure, is reminiscent of increased prophage induction from lysogens when hosts are subjected to UV irradiation (Tomizawa and Ogawa, 1967), but its exact mechanisms have not been investigated any further.

Horizontal gene transfer of the excised GI usually involves process of conjugation, transformation, or phage transduction. Conceivably, the conjugation would need genes related to DNA transfer, whose presence was apparently not detected bioinformatically in SLG, in agreement with the detected absence of its self-transmissibility. However, the lack of the self-transmissibility of a GI could also result from the deletion of its transfer functions after conjugation, a case of which had been described in \textit{E. coli} ECOR31 that a 35-kb transfer region of a conjugative plasmid was deleted from the Yps HPI but present on the \textit{Escherichia coli} HPI.
As well, the failed detection of the mobilized transfer of the SLG by the conjugative plasmid pIJ101 does not necessarily conclude that the SLG is non-mobilizable as transfer systems between GI and the helper plasmid have to match well each other, and some of the GIs may need phage transduction system(s) other than conjugal system(s) of the helper plasmid(s) for its/their mobilization. Examples of the mobilization by phage transduction systems had been reported by Ruzin et al., (2001) for phage SaPI1, and by O’Shea and Boyd, (2002) for phage VPI, respectively. In addition, even the mobilization of the GIs by the phage transduction systems has their matching specificities. For example, phages 13 and 80α were reported able to mobilize SaPI1 (Ruzin et al., 2001), but failed to mobilize SaPIbov in *Staphylococcus aureus* (Fitzgerald et al., 2001). Furthermore, while SaPlbov2 was reported capable of excision from the chromosome *Staphylococcus aureus*, its mobility is still unclear till now (Ubeda et al., 2003). Given the above considerations, and the mosaic structure of the SLG, it is tempting to speculate that SLG may have lost genes required for its natural and/or mobilized transfer in order to maintain a relatively stable inheritance with the host during evolution. To our knowledge, no other *Streptomyces* GIs had been described or demonstrated in details of being able to integrate into and/or excise from their chromosomes, as demonstrated here as a result of a specific control by a discrete *int* gene via a localized 15-bp direct repeats, in SLG.

Although it is yet undefined how the *dnd* gene cluster evolved and was disseminated across different bacterial species based on the limited genome sequence data, the diversity of the *dnd*-bearing hosts, the markedly different *dnd* sequence signatures and the lack of a functionally mobile *dnd* GI, collectively suggest that the *dnd* cluster was organized into a functional locus on a conjugative plasmid and/or other mobile element in very ancient times, prior to extensive spread and sequence diversification over the eons. Plasmids have been proposed to play a role in the evolution and dissemination of some GIs, such as the *exoU* locus of *Pseudomonas aeruginosa* (Kulasekara et al., 2006). In this study, the *dnd* cluster islands were found in *Mesorhizobium sp. BNC1* plasmid 3 (Fig. 5L) and the plasmid-derived Chromosome II of *Pseudoalteromonas haloplanktis* TAC125 (Fig. 5D). In addition, the degree of gene organization conservation extends beyond the *dnd* clusters for several of these GIs (Fig. 5D-H), suggesting these elements might have been evolved from a common ancient ancestor. Similarly, we hypothesize that the plasmid-derived,
mosaic-like small chromosome of *P. haloplanktis* TAC125 could have served as a ‘natural depository’ for various accessory genetic elements, such as the *dnd* loci and its linked genes that could potentially be sourced from diverse bacterial or phage donors (Fig. 5D).

The widespread occurrence of highly conserved *dnd* gene clusters in the bacterial kingdom that are however only sporadically represented amongst members of a species, is reminiscent of classical DNA methylation-based restriction-modification systems, which frequently play a key role in preventing the uptake of foreign DNA or in altering the way in which the genetic blueprint of an organism is decoded and translated into proteins (Brezellec *et al.*, 2006). It is reasonable to assume that the sulphur-based DNA modification system in widespread bacterial species conferred by *dnd* clusters on GIs may play a role in DNA maintenance (Bolden *et al.*, 1984), replication (Schmidt *et al.*, 1985) and/or code for novel DNA uptake barrier systems (Bair and Black, 2006) that ultimately help promote the survival and dispersal of ‘selfish islands’. We are currently actively exploring these hypotheses.

**Experimental procedures**

*Bacterial strains and plasmids, growth conditions, and genetic manipulations*

The construction of plasmids and strains are listed in Table 2. Culture and standard bacteriological methods were generally as described by (Sambrook *et al.*, 1989) and (Kieser *et al.*, 2000). *E. coli* strains were grown at 37°C in Luria-Bertani (LB) medium. *Streptomyces* strains were routinely grown at 28°C on SFM (Kieser *et al.*, 2000) agar for sporulation or for conjugation between *E. coli* ET12567/pUZ8002 and *Streptomyces*; on R2YE (Kieser *et al.*, 2000) for protoplast transformation; on solid MM (Kieser *et al.*, 2000) medium for thermo-sensitivity tests; or in YEME and TSBY (Kieser *et al.*, 2000) liquid media supplemented with 34% (w/v) sucrose or YMG (glucose 4.0, malt extract 10.0, yeast extract 4.0, pH 7.2.) for mycelial growth.

Plasmid and total DNA was prepared from *Streptomyces* strains according to (Kieser *et al.*, 2000). Unmethylated DNA was prepared from *E. coli* ET12567. *In vivo* generation of targeted mutations in *Streptomyces* was achieved by conjugation between *E. coli* ET12567 containing the RP4 derivative pUZ8002 (Flett *et al.*, 1997)
and *S. lividans* according to (Kieser et al., 2000).

**DNA sequencing and sequence analysis**

Four overlapping cosmids, 16C3, 16C2, 16E3 and 16H2 (Zhou et al., 2004), were purified (Sambrook *et al.*, 1989) and sheared with a 550 Sonic Dismembrator (Fisher Scientific) into ca. 2-3-kb fragments. These random fragments were blunted with T4 polymerase (Fermentas) at 11°C for 20 min, gel purified and ligated into pUC18 for shotgun sequencing. Some gaps were filled using PCR amplifications as detailed in (Zhang *et al.*, 1999). The assembled sequence covers the complete 92,770-bp SLG, flanked by 24,872-bp and 19,223-bp at the left and right boundary, respectively.

The nucleotide sequence of SLG and the flanking regions has been deposited in GenBank under accession number EF210454. Putative protein coding sequences larger than 150-bp were predicted using FramePlot beta 3.0 (Ishikawa and Hotta, 1999), with a 120-bp sliding window and a step of 15-bp. Homology searches were performed by using BLAST at National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). The G+C plot (Fig.1A) was drawn using the Freak software within the EMBOSS package (Altschul *et al.*, 1997) with a 300-bp sliding window and a step of 10-bp. Dinucleotide bias analysis was performed using the method proposed by Karlin (Karlin, 2001). The dinucleotide relative abundance value δ° (Karlin, 2001) was calculated with the δρ–web program (http://deltarho.amc.nl/) (van Passel *et al.*, 2005).

**PCR primers and PCR reaction conditions**

Primers used in this study are listed in Table S2. PCR reactions (50µl) containing 5 ng template DNA, 25pM of each primer, 1.5mM MgCl₂, 100µM dNTPs, 1 unit *Taq* polymerase, in 1×PCR buffer (Sangon, Shanghai, China) were performed. PCR cycling conditions were as follows: initial denaturation at 94°C for 180 s, followed by 30 cycles of denaturation at 94°C for 30 s, annealing at a primer-specific temperature for 30 s, and final extension at 72°C for a duration dependent on the length of the expected amplicon (50 s kb⁻¹). PCR products were purified from 0.8% agarose gels using the DNA Gel Extraction Kit (V-gene Biotechnology Limited, China) and subsequently inserted into pMD18-T vector (TaKaRa, Dalian, China) for sequencing.
**Real-time PCR**

Real-time PCR was performed using a Mastercycler® ep realplex 4S (Eppendorf) in the presence of SYBR-green. The PCR was performed using SYBR® Premix ExTaq™ (TakaRa Biotechnology (Dalian) Co., ltd) according to the manufacturer instruction. Cells of *S. lividans* 1326 were harvested after 36 hours of standard growth in TSBY (34%), YEME (34%) (Kieser et al., 2000) and YMG to investigate whether there exists variation of the excision rate influenced by the media components, i.e. osmotic pressure, ion. Amplification products designed to be less than 200-bp in size and primers are listed in Table S2. The left flanking fragment (nt-3343-nt-3196; LC2F & LC2R) of SLG was used as the reference loci. Primer P21F & P21R (nt-92686-nt-56) is targeting for quantification of ratios of cells bearing excised SLG; Reactions was performed in 20 µl volumes. PCR condition was set as follows: initial denaturation at 95°C for 5 minutes, followed by 40 amplification cycles at 95°C for 15s, and final extension at 60°C for 30s. Melting-curves were analyzed at the end of each elongation step to validate the amplification specificity. All PCR amplifications were performed in duplicates on different days to validate the reproducibility of the assays. The relative copy number for each size of DNA molecule was calculated using the comparative Ct method with the formula $2^{\Delta\Delta Ct}$ ($\Delta\Delta Ct=\Delta Ct$ sample-$\Delta Ct$ reference). The copy number for reference fragment was assigned a value of 100%, others were presented as calculated percentage relative to the copy numbers of the reference locus.

**Determination of integrase activity through excision of the mini-island from pJTU1514**

pJTU1514 was introduced into *S. lividans* ZX1, a dnd-mutant of *S. lividans* 66 (Zhou et al., 2004), by conjugation according to (Kieser et al., 2000). After cultivation on SFM at 28°C for 16 hours, exconjugant HXY7 was transferred to grow at 42°C for 72 hours on MM agar supplemented with apramycin (30 μg ml⁻¹). These exconjugants were propagated for sporulation on SFM agar at 28°C for 4 days. Total DNA was prepared from mycelium of HXY10 and used to transform *E. coli* DH10B by electroporation. Plasmid DNA was extracted from *E. coli* transformants exhibiting resistance to apramycin and digested with ApaI, Nhel, SspI, and BglIII to determine
which part was deleted. Subsequently, the fusion junction was amplified by PCR with primers P-pJTU1515F and P-pJTU1515R and the resulting product inserted into pMD18-T, leading to pJTU1511.

Conjugal mating experiments between Streptomyces

Conjugal mating experiments were carried out using the S. lividans HXY1 (dndA::aadA) or S. lividans HXY18 (dndA::aadA, slg10::tsr) as the SLG donor strain and the S. coelicolor M145 mutant ZH3 (SCO3930::aac(3)IV) as a recipient for the inter-specific conjugal mating or S. lividans HXY19 (ZX1 derivative with aac(3)IV inserted at 17090th nt upstream of the attL) as a recipient for the intra-species conjugal mating. Spores of both strains were mixed and co-cultured on the SFM medium at 28°C for 20 hours. The mating plates were overlaid with 50 μg ml\(^{-1}\) streptomycin, 100 μg ml\(^{-1}\) apramycin and/or 15 μg ml\(^{-1}\) thiostrepton to select the potential exconjugants. The donor-to-recipient ratios range from 10:1, 1:1 and 1:10. The number of either donor or recipient spores was approximately \(10^7-10^9\) cfu per 90 mm plate on SFM agar.

Acknowledgements

We thank Prof. Sir. D. A. Hopwood, FRS for critical reading of the manuscript and many valuable comments. We are grateful to Dr. James M. Fleckenstein of University of Tennessee Health Sciences Center, USA, for providing Escherichia coli B7A and the Institute for Genomic Research (TIGR) for their policy of making preliminary sequence data publicly available and acknowledge the use in this study of unpublished genome data corresponding to E. coli B7A. This work received support from the National Science Foundation of China, the 863 and 973 programs from the Ministry of Science and Technology, the Funds from the Ministry of Education and the Shanghai Municipal Council of Science and Technology.

References

genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. *Proc Natl Acad Sci U S A* **97**: 4856-4861.


New York: Oxford University Press.


Ou, H.Y., Chen, L.L., Lonnen, J., Chaudhuri, R.R., Thani, A.B., Smith, R., Garton,


Zhang, J., Voss, K.O., Shaw, D.F., Roos, K.P., Lewis, D.F., Yan, J., Jiang, R., Ren, H.,


Figure Legends

Fig. 1. (A) G+C plot of the SLG DNA sequence. The numbers next to the troughs in the plot represent the four low G+C content valleys. (B) Organizational map of the 85 predicted ORFs on the chromosomally integrated SLG element extending from the left junction *attL* to the right junction *attR*. The four modules (I-IV) referred to in the text are shown as thick lines; boxes above and below the axis represent ORFs in the forward and reverse frame, respectively. The ORFs shown in green are predicted to be involved in nucleotide metabolism and biosynthetic pathways; red boxes indicate ORFs with >95% DNA identity to syntenic genes in *S. coelicolor* A3(2); and blue boxes represent homologues of *S. avermitilis*. The inset schematically shows an expanded view of the *S. coelicolor* A3(2) region related to SLG-formation (*slg* 45-71) and highlights the limited nature of the modular swap that has occurred. There are no *S. coelicolor* A3(2) SCO3521 and SCO3522 homologues in *S. lividans* 66, while there is another copy of this tandem gene-pair (SCO0874-0875) in *S. coelicolor* A3(2).

Fig. 2. Spontaneous excision of SLG from the chromosome of *S. lividans* 66. Thick straight lines represent the *S. lividans* 66 chromosome; *attL* and *attR* are shown as grey arrowheads, whilst primer annealing sites are represented by short arrows. (A) PCR analysis of *S. lividans* 66 genomic DNA using primers Lp1F and Rp1R detected the 467-bp fusion fragment between *attL* and *attR*, confirming the existence of excised and circularized forms of SLG. (B) PCR analysis of *S. lividans* 66 genomic DNA using primers Rp2F and Lp1R, which would amplify across the predicted deletion points, detected a 725-bp junction fragment, supporting the existence of an SLG-minus variant in the population. (C) Real-time PCR analysis of the ratios of the cells harboring the excised SLG. All values were shown as the relative copy numbers to the reference fragment. DNA was prepared from three different media to monitor the difference of excision rates. All data represent the mean ±SD.

Fig. 3. Integration and excision of a SLG mini-island in *S. lividans* 66. The mini-island was introduced on a thermo-sensitive plasmid, pJTU1514; derivatives that had undergone an island-swapping event were selected by growth at 42°C in the presence of apramycin, resistance to which was encoded by the mini-island-borne
**Fig. 4.** Localization of essential elements for SLG excision and integration by constructing SLG derivatives. The symbols are the same as in Fig. 3. Primer annealing sites are indicated by short arrows. pJTU1515, pJTU1516, pJTU1517 and pJTU1520 are shown in the linear form; dotted lines indicate regions deleted. The replicative plasmid pJTU1522 carried an intact copy of the SLG integrase gene. HXY10-1, HXY11 and HXY13 are ZX1 derivatives containing an integrated copy of pJTU1515, pJTU1516 and pJTU1520, respectively. HXY12, the parent strain of HXY13, contains an integrated copy of pJTU1520 and episomal pJTU1522. (A) PCR analysis using primers Lp1F and Lp1R and genomic DNA as indicated, targeting the left junctions of integrated mini-islands. (B) PCR analysis using primers Rp2F and Lp1R, targeting the fused junctions of excised circular mini-islands. No PCR product was observed with HXY13 DNA. (C) PCR analysis using primers Rp1F and Rp1R, targeting the right junctions of integrated mini-islands.

**Fig. 5.** Maps of putative dnd-encoding GIs and backbone in 11 bacterial chromosomes and one plasmid. The dnd gene clusters are shown as blue arrows, with the labels referring to the matching dnd gene in *S. lividans* 66. Red, light purple and
yellow arrows indicate genes coding for functionally similar proteins that do not necessarily exhibit sequence similarity, while the remaining filled arrows indicate sets of genes sharing sequence homology. Unfilled arrows indicate genes coding for hypothetical proteins. Hatched boxes represent flanking genomic backbone. The G+C content of islands and chromosomes and the taxonomic division to which the organisms belong are shown in curved and square parentheses, respectively. GenBank accession numbers are shown.
Table 1. *dnd* clusters identified in twelve bacterial strains present in putative GIs

<table>
<thead>
<tr>
<th>Species &amp; strain designation</th>
<th>DNA coordinates</th>
<th>Size (kb)</th>
<th>G+C content (%)</th>
<th>Genome fragmen ts with lower δ* (%)</th>
<th>Other features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Streptomyces lividans 66 SLG</strong></td>
<td>1-92770 (slg01-slg85)</td>
<td>92.7</td>
<td>67.8</td>
<td>47.5</td>
<td>100</td>
</tr>
<tr>
<td><strong>Streptomyces avermitilis MA-4680</strong></td>
<td>3655943-3684643 (SCO2922-2937)</td>
<td>28.7</td>
<td>62.8</td>
<td>76.8</td>
<td>99.4</td>
</tr>
<tr>
<td><strong>Pseudomonas fluorescens PfO-1</strong></td>
<td>865302 – 879219 (Pfl_0738-0747)</td>
<td>14.9</td>
<td>51.4</td>
<td>100.9</td>
<td>99.1</td>
</tr>
<tr>
<td><strong>Pseudoalteromonas haloplanktis TAC125(ChrII)</strong></td>
<td>101629-119302 (PSHAb0089-0099)</td>
<td>17.4</td>
<td>36.1</td>
<td>78.1</td>
<td>100</td>
</tr>
<tr>
<td><strong>Escherichia coli</strong> B7A</td>
<td>9671-27566 ([NZ_AAJT01000066] 1-33601 [NZ_AAJT01000025])</td>
<td>54.3</td>
<td>49.9</td>
<td>56.6</td>
<td>100</td>
</tr>
<tr>
<td><strong>Hahella chejuensis</strong> KCTC 2396</td>
<td>7145272 – 7193152 (HCH_07029-07068)</td>
<td>47.9</td>
<td>47.0</td>
<td>67.5</td>
<td>100</td>
</tr>
<tr>
<td><strong>Oceanobacter sp. RED65</strong></td>
<td>38472-57669 ([NZ_AAQH01000003.1])</td>
<td>19.2</td>
<td>40.7</td>
<td>[ca. 76.4]</td>
<td>100</td>
</tr>
<tr>
<td><strong>Bacillus cereus</strong> E33L</td>
<td>930024-963899 ([BCZK0812-0835])</td>
<td>33.9</td>
<td>32.6</td>
<td>60.4</td>
<td>99.4</td>
</tr>
<tr>
<td><strong>Geo bacterium uraniumreducens</strong> Rf4</td>
<td>1-12996 ([NZ_AAON01000085])</td>
<td>13.6</td>
<td>51.4</td>
<td>65.2</td>
<td>92.8</td>
</tr>
<tr>
<td><strong>Pelagibacter ubique</strong> HTCC1002</td>
<td>59035-97062 ([NZ_AAPV01000002])</td>
<td>37.9</td>
<td>28.0</td>
<td>[ca. 39.5]</td>
<td>100</td>
</tr>
</tbody>
</table>

Note: δ* is the fraction of DNA bases where G+C content is lower than genome G+C content. A superscript letter indicates the cluster alignment with *dnd* cluster.
**Roseobacter denitrificans** OCh (RD1_0795-0808)

| 768003 – 792504 | 24.5 | 52 [59%] | 114.8 | 100 | *dnd*-cluster, two clustered transposition protein at right boundary

*a* Dinucleotide bias analysis was adapted from the method proposed by (Karlin et al., 2001). The value $\delta^*$ denotes the dinucleotide relative abundance difference between the island fragment and the reference genome. The $\delta^*$ value was calculated with the $\delta\rho$-web program (http://deltarho.amc.una.nl) (van Passel et al., 2005). The high $\delta^*$ values of these fragments indicate a likely heterologous origin.

*b* The percentage distribution of $\delta^*$ is plotted using the $\delta\rho$-web tool with random host genomic fragments of equal length as input sequences (van Passel et al., 2005).

*c* The complete sequence of *S. coelicolor* A3(2) was employed as the reference genome as its genome is closely related to that of *S. lividans* 66 (Kieser et al., 1992).

*d* The complete sequence of *E. coli* K-12 MG1655 was employed as the reference genome as sequencing of the *E. coli* B7A genome is ongoing and 198 contigs are available to date.

*e* The contig sequence of the ongoing sequenced *Oceanobacter* sp. RED65 genome (NZ_AAQR01000031; 215045-bp) was employed as the reference. The RED65 genome has an estimated size of 3.53 Mb.

*f* The complete sequence of *Geobacter metallireducens* GS-15 was employed as the reference genome as sequencing of the *Geobacter uraniumreducens* Rf4 genome is ongoing and 189 contigs are available to date.

*g* The sequence of the largest *Pelagibacter ubique* HTCC1002 contig (NZ_AAPW0100001; 1012886-bp) among four available contigs was employed as the reference. The HTCC genome has an estimated size of 1.33 Mb.
<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant properties</th>
<th>Comment</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. lividans</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>Wild type, Dnd⁺, φHAU3⁺, SLP2⁺, SLP3⁺</td>
<td></td>
<td>(Feitelson &amp; Hopwood, 1983; Zhou et al., 2004)</td>
</tr>
<tr>
<td>JT46</td>
<td>rec-46, str-6, pro-2, Dnd⁺, φHAU3⁺</td>
<td></td>
<td>(Tsai and Chen, 1987)</td>
</tr>
<tr>
<td>ZX1</td>
<td>JT46 derivative, selected for DNA stability during electrophoresis, dnd, φHAU3⁺</td>
<td></td>
<td>(Zhou et al., 2004)</td>
</tr>
<tr>
<td>HXY1</td>
<td>1326 derivatives with insertion of a aadA gene into dndA, Spc/Stren</td>
<td></td>
<td>(Zhou et al., 2005)</td>
</tr>
<tr>
<td>HXY7</td>
<td>ZX1 derivative obtained following conjugal acquisition at 28°C of pJTU1514</td>
<td></td>
<td>Fig. 3A; This study</td>
</tr>
<tr>
<td>HXY10</td>
<td>ZX1 derivative containing a chromosomally integrated copy of pJTU1515</td>
<td></td>
<td>Fig. 3C &amp; 4; This study</td>
</tr>
<tr>
<td>/HXY10-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HXY11</td>
<td>ZX1 derivative containing a chromosomally integrated copy of pJTU1516</td>
<td></td>
<td>Fig. 4; This study</td>
</tr>
<tr>
<td>HXY12</td>
<td>ZX1 derivative containing episomal pJTU1522 and a chromosomally integrated copy of pJTU1520</td>
<td></td>
<td>Fig. 4; This study</td>
</tr>
<tr>
<td>HXY13</td>
<td>HXY12 derivative obtained following spontaneous loss of pJTU1522</td>
<td></td>
<td>Fig. 4; This study</td>
</tr>
<tr>
<td>HXY16</td>
<td><em>S. lividans</em> 66 derivative with SLG precisely excised at the attB site</td>
<td></td>
<td>Fig. S2; This study</td>
</tr>
<tr>
<td>HXY18</td>
<td><em>S. lividans</em> HXY1-derived mutant, slg10 is disrupted by ori (pUC18)-tsr mediated by pJTU1519</td>
<td></td>
<td>This study</td>
</tr>
<tr>
<td>HXY19</td>
<td>ZX1 derivative with aac(3)IV cassette inserted at 17090th nt upstream of the attL (the distal end to dnd cluster) of SLG</td>
<td></td>
<td>This study</td>
</tr>
<tr>
<td><em>S. coelicolor</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M145</td>
<td>Wild type</td>
<td>Sequenced strain</td>
<td>(Kieser et al., 2000)</td>
</tr>
<tr>
<td>ZH3</td>
<td>M145 derivative with SCO3930 partially replaced with aac(3)IV</td>
<td></td>
<td>(Li et al., 2006)</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>F⁻ recA lacZ ΔM15</td>
<td>General cloning host</td>
<td>(Hanahan, 1983)</td>
</tr>
<tr>
<td>ET12567</td>
<td>Strain used for conjugation between <em>E. coli</em> and <em>Streptomyces</em> spp. recF, dam, dcm, hsdS, Cml⁺, Str⁺, Tef⁺, Km⁺</td>
<td>Genomic library cosmid</td>
<td>(Flett et al., 1997)</td>
</tr>
<tr>
<td>/pUZ8002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16C3</td>
<td>SuperCos1-derived cosmid with insert from <em>S. lividans</em> 66 carrying dnd and attR</td>
<td></td>
<td>(Zhou et al., 2004)</td>
</tr>
<tr>
<td>16H2</td>
<td>SuperCos1-derived cosmid with insert from <em>S. lividans</em> 66 carrying attL</td>
<td>Genomic library cosmid</td>
<td>This study</td>
</tr>
<tr>
<td>16C2</td>
<td>SuperCos1-derived cosmid with insert from <em>S. lividans</em> 66</td>
<td>Genomic library cosmid</td>
<td>This study</td>
</tr>
</tbody>
</table>

30
**lividans 66**

*SuperCos1-derived cosmid with insert from S. lividans 66*  
*Genomic library cosmid*

**pHZ132**

Bifunctional pSG5 derivative, *vph*, *oriT*, *ori*(pAT153), *bla*, *tsr*, *cos*, *rep*, *ori*(pSG5)  
*(Bao et al., 1997)*

**pSET152**

*aac(3)IV*, *lacZ*, *rep*<sup>pUC</sup>, *att<sup>pC31</sup>*

**pOJ260**

*aac(3)IV*, *oriT*, *rep*<sup>pUC</sup>, *lacZ*

**pMD18-T**

Shuttle cosmid derived from pHZ1358 (Sun et al., 2002), *oriT*, *ori*(ColE1), *bla*, *tsr*, *cos*, *rep*(pIJ101), *ori*(pIJ101)  
*(Kieser et al., 2000)*

**pJTU1510**

pMD18-T containing 722-bp PCR product carrying *attR* from 16C3  
*Fig. 3A*  
*This study*

**pJTU1511**

pMD18-T containing the PCR amplicon generated from pJTU1515 using primers P-pJTU1515F and P-pJTU1515R  
*Fig. 3C*  
*This study*

**pJTU1512**

A 3575-bp ApaI fragment carrying *attL-int-tnp* from 16H2 was cloned into the unique *ApaI* site of pSET152  
*This study*

**pJTU1513**

pJTU1512 containing an insert bearing the *oriT* and thermo-sensitive *ori* of pSG5  
*This study*

**pJTU1514**

Mini-island-carrying vector constructed by ligation of a HincII-BamHI fragment carrying *attR* from pJTU1510 with EcoRV-BamHI-cut pJTU1513  
*Fig. 3A*  
*This study*

**pJTU1515**

The excised and re-circularized mini-island derived from HXY10 DNA following replicative passage in *E. coli* DH10B  
*Fig. 3C*  
*This study*

**pJTU1516**

pJTU1515 mini-island derivative lacking *tnp*  
*Fig. 4*  
*This study*

**pJTU1517**

The region from 148nt upstream to 12nt downstream of the *attP* was removed from pJTU1515 with PstI and NheI digestion, blunted and self-relagated.  
*This study*

**pJTU1518**

SmaI DNA fragment (10364-14128) was recovered from 16H2 and ligated with pSET152  
*This study*

**pJTU1519**

6173-bp *tsr-bla-ori* (pAT153) containing NotI fragment from pHZ132 was inserted into NotI site of pJTU1518  
*This study*

**pJTU1520**

pJTU1515 mini-island derivative lacking *tnp* and *int*  
*Fig. 4*  
*This study*

**pJTU1521**

pOJ260 (*Streptomyces* suicide plasmid) containing an intact copy of the *int* gene from pJTU1515  
*This study*

**pJTU1522**

pJTU412 containing an intact *int* gene from pJTU1521  
*Fig. 4*  
*This study*

---

*oriT*, origin of transfer of plasmid RK2; *tsr*, thiostrepton resistance gene; *aac3(IV)*, apramycin resistance gene; *aadA*, streptomycin/spectinomycin (*str/spc*) resistance
gene; *bla*, ampicillin resistance gene; *vph*, viomycin resistance gene; *dnd*, gene cluster encoding DNA degradation phenotype; *int*, integrase of SLG; φHAU3', resistant to phage φHAU3.
Figure 2
Figure 3
Figure 4
Figure 5