

# Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of *Streptomyces kanamyceticus*

Koji Yanai\*<sup>†</sup>, Takeshi Murakami<sup>†</sup>, and Mervyn Bibb\*<sup>‡</sup>

\*Department of Molecular Microbiology, John Innes Centre, Colney, Norwich NR4 7UH, United Kingdom; and <sup>†</sup>Microbiological Resources and Technology Laboratories, Meiji Seika Kaisha Ltd., 788 Kayama, Odawara-shi, Kanagawa 2500852, Japan

Communicated by H. Boyd Woodruff, Soil Microbiology Associates, Watchung, NJ, April 21, 2006 (received for review March 6, 2006)

***Streptomyces kanamyceticus* 12-6 is a derivative of the wild-type strain developed for industrial kanamycin (Km) production. Southern analysis and DNA sequencing revealed amplification of a large genomic segment including the entire Km biosynthetic gene cluster in the chromosome of strain 12-6. At 145 kb, the amplifiable unit of DNA (AUD) is the largest AUD reported in *Streptomyces*. Striking repetitive DNA sequences belonging to the clustered regularly interspaced short palindromic repeats family were found in the AUD and may play a role in its amplification. Strain 12-6 contains a mixture of different chromosomes with varying numbers of AUDs, sometimes exceeding 36 copies and producing an amplified region >5.7 Mb. The level of Km production depended on the copy number of the Km biosynthetic gene cluster, suggesting that DNA amplification occurred during strain improvement as a consequence of selection for increased Km resistance. Amplification of DNA segments including entire antibiotic biosynthetic gene clusters might be a common mechanism leading to increased antibiotic production in industrial strains.**

aminoglycoside | antibiotic | amplifiable unit of DNA | clustered regularly interspaced short palindromic repeats | mobile genetic element

Microorganisms produce a wide variety of secondary metabolites, many with important applications in medicine (e.g., as antibiotics, immunosuppressants, and anticancer agents) and agriculture (e.g., as veterinary products and herbicides). The development of commercially viable fermentation processes for secondary metabolites generally involves the empirical isolation of overproducing mutants. The mechanisms underlying enhanced productivity in such mutants are generally unknown.

Kanamycin (Km), one of the most commercially successful antibiotics, is an aminoglycoside produced by *Streptomyces kanamyceticus* (1). Although the emergence of antibiotic resistance has limited the clinical use of Km to the treatment of tuberculosis, the semisynthetic Km derivatives amikacin, dibekacin, and arbekacin are used widely (2). In pursuit of rational approaches to increased Km production and new opportunities to create novel Km derivatives by genetic engineering, cosmid clones containing the Km biosynthetic gene cluster were sequenced (3, 4), and the functions of several genes were determined (4, 5). Interestingly, Southern analysis of genomic DNA from the wild-type strain and a Km overproducing mutant (strain 12-1) suggested additional copies of the Km cluster in the Km overproducing mutant (3).

DNA amplification in streptomycetes arises from sequences called amplifiable units of DNA (AUDs) and can result in tandem reiteration of several hundred copies of the AUD per chromosome (6). AUDs often lie in unstable chromosomal regions, such as those close to the ends of the linear chromosome, and a large deletion frequently accompanies the amplification (6). DNA amplification in streptomycetes usually occurs in the absence of any overt selective pressure, and the participating AUDs typically fail to reveal genes that might confer a growth advantage under the conditions used.

Here we describe amplification of an unusually large (144.9-kb) DNA segment including the entire Km biosynthetic gene cluster in an industrial Km overproducing derivative of *S. kanamyceticus*.

## Results

**Two Different Km Gene Clusters Occur in Km-Overproducing Mutants of *S. kanamyceticus*.** In an earlier work, a 4.95-kb SphI fragment containing at least one likely Km biosynthetic gene (*orf20*) hybridized to two distinct BamHI fragments (9.6-kb and 10.8-kb) in DNA from a Km-overproducing mutant (strain 12-1). Only the smaller fragment was found in the wild-type strain (3). Southern analysis of a further 11 overproducing mutants, including the current Meiji Seika Kaisha production strain, revealed the same two BamHI fragments, although the abundance of the 10.8-kb fragment varied from strain to strain. The 10.8-kb BamHI fragment, originally cloned in cosmid pKM9 (3), was subcloned in pUC118, yielding pKM92. To isolate the 9.6-kb BamHI fragment, a cosmid library of *S. kanamyceticus* (3) was screened by using the 4.95-kb SphI fragment and cosmid pKM7 identified. The sequences of the DNA fragments cloned in pKM7 and pKM92 were determined.

Eight genes, *orf25–32*, were identified by Frame analysis (7) in the 9.6-kb BamHI fragment and an adjacent 5.5-kb BamHI–AseI fragment (Fig. 1A), representing gene organization in the wild-type strain. *orf25–32* and the adjacent and previously sequenced *orf23* and *orf24* (3) resemble contiguous clusters in *Streptomyces coelicolor* (SCO6476–SCO6467; ref. 8) and *Streptomyces avermitilis* (SAV1908–SAV1917; ref. 9). Because these two strains do not produce Km, *orf23–32* were concluded not to be involved in Km biosynthesis. *orf20* encodes a glycosyltransferase probably involved in Km biosynthesis (3) and may define the right end of the Km gene cluster.

A sequence different from that of the 9.6-kb BamHI fragment was observed 262 bp downstream of the putative translation initiation site of *orf25* in the 10.8-kb BamHI fragment, indicating that a recombination event had occurred at this point during derivation of the Km-overproducing mutants. Six genes, *orf1083–1088*, were identified downstream of the truncated *orf25* (Fig. 1A). Only Orf1083 [27% identity to a hypothetical protein (SAV167) of *S. avermitilis*] and the C-terminal half of Orf1088 [35% identity to a hypothetical protein of SLP2 of *Streptomyces lividans* (10)] showed homology to proteins in the National Center for Biotechnology Information database.

*S. kanamyceticus* 12-6 is the earliest strain in the Km improve-

Conflict of interest statement: No conflicts declared.

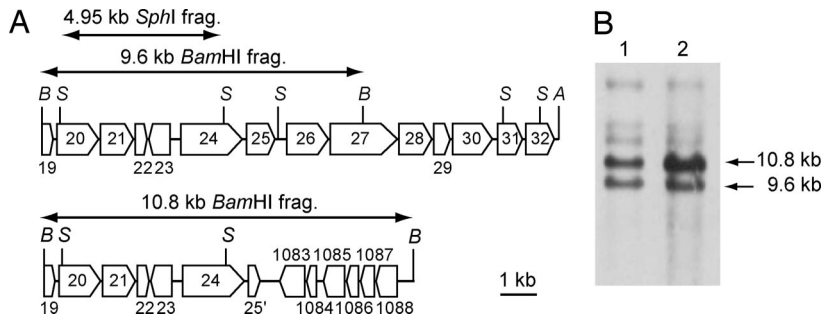
Freely available online through the PNAS open access option.

Abbreviations: AUD, amplifiable unit of DNA; Rs, recombination site; PFGE, pulsed-field gel electrophoresis; CRISPR, clustered regularly interspaced short palindromic repeats; Km, kanamycin.

Data deposition: The sequences reported in this paper have been deposited in the Genbank database (accession nos. AB254080, AB254081, and AB254082).

<sup>†</sup>To whom correspondence should be addressed. E-mail: mervyn.bibb@bbsrc.ac.uk.

© 2006 by The National Academy of Sciences of the USA



**Fig. 1.** Two different but overlapping BamHI fragments encode part of the Km biosynthetic gene cluster in strain 12-6. (A) ORFs deduced from sequences derived from pKM7 (the 9.6-kb fragment) and pKM92 (the 10.8-kb fragment). A, B, and S indicate AseI, BamHI, and SphI sites, respectively. (B) Genomic DNAs of strain 12-6 were prepared from mycelium grown without (lane 1) or with (lane 2) Km (1,000  $\mu\text{g}/\text{ml}$ ), digested with BamHI, and separated on a 0.8% agarose gel. The 4.95-kb SphI fragment (A) was used as probe.

ment pedigree to show the two hybridizing bands. Whereas the intensities of the two bands derived from mycelium grown without Km were similar, the intensity of the 10.8-kb band was at least twice that of the 9.6-kb band in DNA from cultures containing Km (1000  $\mu\text{g}/\text{ml}$ ) (Fig. 1B). Thus strain 12-6 contained more than three copies of at least some of the Km biosynthetic genes when grown in the presence of Km.

**Chromosome Organization in the Wild-Type and Km-Overproducing Strains.** To elucidate the origin of strain 12-6, chromosome walking was carried out by using a cosmid library of the wild-type strain. Two probes were used that corresponded to sequences on either side of the novel recombination site in the 10.8-kb BamHI fragment found only in the Km-overproducing mutants. Inspection of the resulting contiguous 13 cosmid clones (Fig. 2) revealed that strain 12-6 had arisen by recombination of DNA sequences immediately to the right of the Km gene cluster [recombination site B (RsB)] with sequences (RsA)  $\approx 145$  kb to the left.

Restriction fragments from the left ends of the cloned inserts of cosmids 1-3 and 2-1 failed to hybridize in Southern blots with DNA of strain 12-6, whereas a fragment from the left end of cosmid 4-5 hybridized to a band of the same size in DNA from both the wild type and 12-6. Thus, strain 12-6 had undergone a deletion whose endpoint (RsC) was in the region of the chromosome covered by cosmid 4-5.

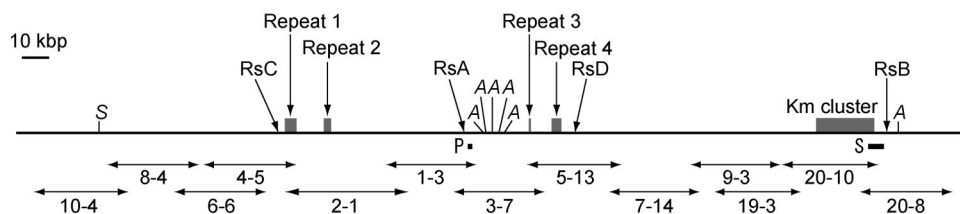
To isolate a DNA fragment containing RsC from strain 12-6, a cosmid library of strain 12-6 was probed with a DNA fragment from the left end of cosmid 4-5. The resulting cosmid, 12L1-4, contained a 2.2-kb BamHI fragment not found in cosmid 4-5. Southern analysis of BamHI-digested genomic DNAs using this fragment as a probe revealed bands of 12.4 kb and 5.8 kb in the wild-type strain, and of 2.2 kb and 5.8 kb in strain 12-6. The 2.2-kb BamHI fragment hybridized not only to cosmid 4-5 but also to a 5.8-kb BamHI fragment from cosmid 5-13. This finding suggested that the 2.2-kb BamHI fragment contained RsC of strain 12-6, and that recombination had occurred between RsC in the 12.4-kb BamHI fragment of the wild-type strain and a recombination site (RsD) in the 5.8-kb BamHI fragment (Fig. 3). This deletion event occurred in addition to recombination between RsA and RsB. Taken together, these results are con-

sistent with the chromosome structures shown in Fig. 4A and B; that is, in strain 12-6, the region between RsC and RsA in the wild-type strain was replaced by a duplication of the region between RsD and RsB, yielding the new junction sequences RsA/RsB and RsC/RsD.

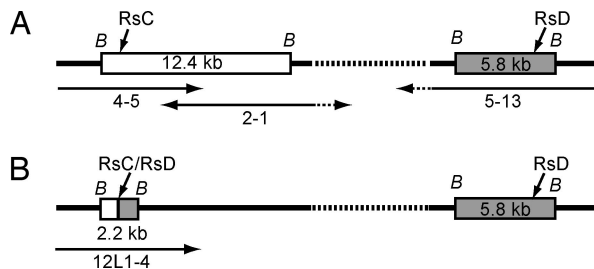
On the basis of this interpretation, Southern analyses were carried out to compare the chromosome structures of the wild-type and 12-6 strains cultivated with Km (1,000  $\mu\text{g}/\text{ml}$ ) (Fig. 5A and B). The PCR fragment P (Fig. 2) was used as a probe. The wild-type strain yielded hybridizing bands of 511 kb (AseI) and 138 kb (AseI plus SspI), whereas strain 12-6 gave 553-kb (AseI) and 180-kb (AseI plus SspI) bands, i.e., 42 kb larger than their counterparts in the wild-type strain, entirely consistent with the chromosome structure proposed for strain 12-6 (Fig. 4B). In addition to these bands, a 133-kb band of higher intensity was detected in both digested samples of strain 12-6 DNA. This observation suggested that the DNA between RsA and RsB in strain 12-6 was present as a tandemly reiterated AUD (Fig. 4C) and that the 133-kb band was an internal AseI fragment of the AUD.

Because there are no SspI sites in the AUD, its copy number in strain 12-6 could be estimated from the size of the bands detected in Southern analysis of SspI-digested DNA. A single band of 498 kb was observed in the wild-type strain. In contrast, at least four bands (532 kb, 658 kb, 802 kb, and 964 kb) were detected in strain 12-6 (Fig. 5C), corresponding to SspI fragments containing two, three, four, and five copies of the AUD, respectively. This finding indicates that strain 12-6 consists of a mixture of genomes with a varying number of tandemly arrayed AUDs.

In addition to these bands, a strong hybridization signal was detected at the origin in the wells containing strain 12-6 DNA (Fig. 5C), indicative of circular or large (>2 Mb) linear DNA molecules (11, 12). To determine the topology of this hybridizing DNA, agarose plugs were treated with SspI and then with Plasmid-Safe ATP-Dependent DNase (Epicentre Technologies, Madison, WI), which is active on double-stranded linear but not on circular DNA. Subsequent pulsed-field gel electrophoresis (PFGE) and Southern analysis failed to reveal any hybridization, either at the origin or elsewhere in the lane (data not shown), indicating a linear topology. To estimate the size of the linear SspI fragments at the origin, PFGE



**Fig. 2.** Restriction map of the wild-type *S. kanamyceticus* chromosome around the proposed sites of recombination. A and S, AseI and SspI recognition sites, respectively. The four recombination sites are indicated with RsA, RsB, RsC, and RsD. The positions of the repeat sequences (Repeats 1–4) and of the Km biosynthetic gene cluster (Km cluster) are indicated. The arrangement of the isolated cosmids is shown below the restriction map. P and S indicate the positions of the 1,167-bp PCR and 4.95-kb SphI probe fragments, respectively.



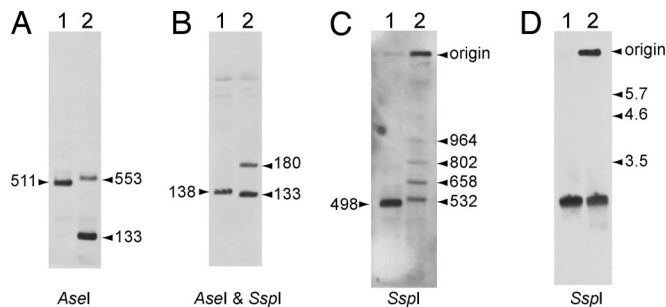
**Fig. 3.** Chromosome structures around RsC and RsD in wild-type (A) and 12-6 (B) strains. Regions covered by individual cosmids are indicated with arrows. B, BamHI site.

conditions were used that allowed gel entry and resolution of the chromosomes of *Schizosaccharomyces pombe*, but these also failed to move the hybridizing DNA from the origin (Fig. 5D). Thus, the SspI fragments at the origin were deduced to be >5.7 Mb, suggesting that they contain >36 copies of the AUD. The relative intensities of the different SspI fragments (Fig. 5C) suggest that the chromosome containing two copies of the AUD (Fig. 4B) is the predominant species in strain 12-6, consistent with the earlier estimate of an average copy number for the amplified region of about three (Fig. 1B).

**Sequence Analysis of the Deleted and Amplified Regions in the Km-Overproducing Strain.** The above analysis indicated that the region between RsC and RsA was deleted in strain 12-6, whereas the region between RsA and RsB was tandemly reiterated. To define the nature of these changes at the nucleotide level, seven cosmids of wild-type DNA (4-5 to 9-3, Fig. 2) were sequenced, generating 205,447 nt that overlapped the previously published sequence (4). The 2.2-kb BamHI fragment was also sequenced to reveal the RsC/RsD junction.

Sequence alignments of the recombination sites and new junctions revealed a common 6-nt sequence (5'-TGGTCC-3') at RsA (94,708–94,713) and RsB (239,609–239,614) (Fig. 6A). The RsA/RsB junction, determined from the sequence of the 10.8-kb BamHI fragment of pKM92, confirmed that recombination had occurred at this 6-nt sequence. In contrast, incomplete direct repeats of 666 bp, including 645 bp of identical sequence, were found at RsC and RsD (28,269–28,934 and 135,582–136,247, respectively; Fig. 6B). The sequence of the RsC/RsD junction was the same as that of RsC, indicating that recombination occurred in the perfectly repeated 38-nt sequence located at the right end of the homologous region.

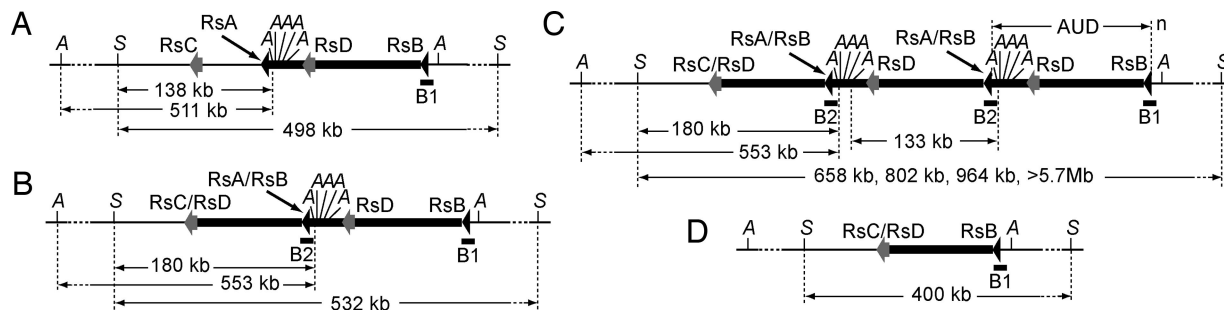
Because the distance between RsD and RsB is 104,027 bp, and that between RsC and RsA is 66,439 bp, replacing the latter segment with the former would result in an increment of 37,588 bp,



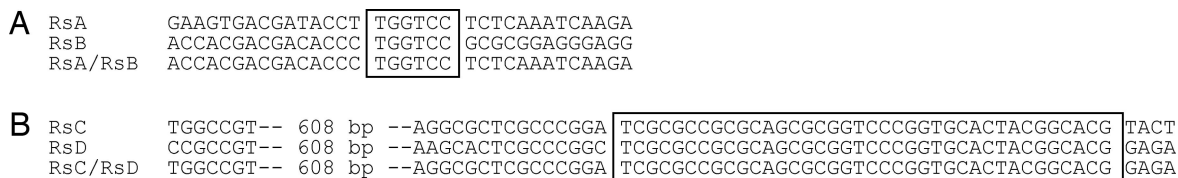
**Fig. 5.** Southern analyses of PFGE gels. Samples from the wild-type and 12-6 strains were loaded in lanes 1 and 2, respectively. Restriction endonucleases used were AseI (A), AseI and SspI (B), and SspI (C and D). The 1,167-bp PCR fragment (Fig. 2) was used as a probe. PFGE conditions were as follows. (A and C) Agarose concentration, 1.0%; voltage gradient, 6 V/cm; angle, 120°; switch time, 50–90 s; run time, 22 h. (B) Agarose concentration, 1.0%; voltage gradient, 6 V/cm; angle, 120°; switch time, 5–20 s; run time, 22 h. (D) Agarose concentration, 0.8%; voltage gradient, 2 V/cm; block 1 switch time, 1,200 s; block 1 run time, 24 h; block 2 switch time, 1,500 s; block 2 run time, 24 h; block 3 switch time, 1,800 s; block 3 run time, 24 h. Fragment sizes in A, B, and C are given in kb. AseI-digested DNA of *S. coelicolor* was used as a size marker (A–C). The positions and sizes (Mb) of the *S. pombe* chromosomes are indicated to the right of D.

consistent with the difference between the wild type and strain 12-6 estimated by Southern analyses (42 kb; Figs. 4 and 5). Similarly, the size of the AUD is calculated to be 144,901 bp, whereas the calculated AseI fragment size in the AUD is 135,809 bp, again consistent with observed results (Figs. 4 and 5). The largest AUD reported previously in *Streptomyces* is AUD2 from *S. lividans* with a size of 92 kb (13). Thus, to our knowledge, the AUD described in this study is the largest yet reported in *Streptomyces*.

One hundred and eighty five ORFs (*orf1001–1185*) were annotated in the 205,447-bp *S. kanamyceticus* sequence by using ARTEMIS (14). Comparison of the predicted encoded proteins with those of *S. coelicolor* showed excellent synteny in the regions flanking RsC and RsD; SCO6541–SCO6525 correspond to *S. kanamyceticus* Orf1008–1023 to the left of RsC, whereas SCO6525–SCO6486 correspond to *S. kanamyceticus* Orf1120–1149 to the right of RsD. The direct repeats found at RsC and RsD lie within two *S. kanamyceticus* homologues of SCO6525, which encodes a conserved hypothetical protein. RsC lies in Orf1023, and RsD lies in Orf1120; whereas Orf1120 shows homology throughout its length to SCO6525, only the N-terminal 209 aa of Orf1023 resemble SCO6525 (255 aa). The interruption in synteny; the nature of many of the 96 proteins predicted to be encoded by the intervening 106.6-kb sequence, which includes homologues of proteins derived from insertion



**Fig. 4.** Variations in chromosome structure around the Km gene cluster. Shown are the wild type (A), duplicated AUD (B), highly amplified AUD (C), and a remnant of AUD after deletion of the region between RsC and RsD (D). Sizes of restriction fragments (A, AseI; S, SspI) detected in the hybridization experiments (Figs. 4 and 6) are indicated. n, multiple copies of the AUD. B1 and B2, positions of the 9.6-kb and 10.8-kb BamHI fragments (Fig. 1), respectively. The thick line indicates the AUD between RsA and RsB.



**Fig. 6.** Nucleotide sequences of the recombination and junction sites. (A) R<sub>s</sub>A, R<sub>s</sub>B, and the R<sub>s</sub>A/R<sub>s</sub>B junction. (B) R<sub>s</sub>C, R<sub>s</sub>D, and the R<sub>s</sub>C/R<sub>s</sub>D junction. Note conservation of 6 (A) and 38 (B) nt at the recombination and junction sites, respectively.

sequence elements (*orf1050*, *orf1106*, and *orf1116*), conjugative plasmids (*orf1061*, *orf1082*, *orf1088*, and *orf1089*), and bacteriophage (*orf1095* and *orf1119*); and the lower GC content of this region [68.4 mol% versus 71.8 mol% to the left of R<sub>s</sub>C (nucleotides 1–28934) and 72.5 mol% to the right of R<sub>s</sub>D (nucleotides 135582–205447)] strongly imply that the DNA between R<sub>s</sub>C and R<sub>s</sub>D is a mobile genetic element.

Strikingly, a pair of highly repeated sequence elements occurs toward each end of the R<sub>s</sub>C–R<sub>s</sub>D segment: repeat 1 (29639–33328) and repeat 2 (43210–45495) toward the left end, and repeat 3 (118626–119142) and repeat 4 (127726–130558) toward the right end. Each repeat consists of a highly conserved 29-nt sequence followed by a 32-nt nonconserved sequence, with the conserved 29-nt sequences showing weak palindromic motifs (Table 1). The conserved 29-nt sequences of repeats 1 and 2 are almost complementary, as are those of repeats 3 and 4. These characteristics resemble those of CRISPRs (clustered regularly interspaced short palindromic repeats), which occur in prokaryotes but are absent from eukaryotes and viruses (15, 16). A search for homologous sequences using BLASTN (17) revealed similar repeated sequences in the chromosome of *S. avermitilis* (8993350–8996980 and 8981247–8983654; ref. 9) and in two streptomycete plasmids, SCP1 (18) and pSV1 (19). Eight ORFs (Orf1024–1031) lie between repeats 1 and 2, and six of their products show homology to proteins encoded by CRISPR-associated (*cas*) genes, which include helicases and exonucleases (16). Seven ORFs (Orf1108–1114) lie between repeats 3 and 4, and five of their products also resemble Cas proteins. CRISPRs have been implicated in large-scale chromosome rearrangements (20), DNA condensation (21), replicon partitioning, DNA repair, and RNAi-based immunity against foreign DNA (ref. 22 and references therein). To our knowledge, the compound nature of the apparently mobile R<sub>s</sub>C–R<sub>s</sub>D segment, with its terminally located pairs of CRISPRs, is unique. Whereas repeats 1 and 2 were deleted in strain 12-6, repeats 3 and 4 persist as part of the AUD, and it is conceivable that these CRISPRs and their associated Cas proteins are crucial for amplification of the AUD.

**The Chromosome Complement of the Km-Overproducing Strain Changes in the Presence of Km.** The copy number of the 10.8-kb BamHI fragment containing the R<sub>s</sub>A/R<sub>s</sub>B junction differed in cultures grown in the presence and absence of Km (Fig. 1B), suggesting that selection for increased Km resistance might induce amplification. To investigate this further, Southern analysis was carried out on SspI-digested genomic DNA isolated from strain 12-6 grown in the presence and absence of Km (Fig. 7A).

**Table 1. Features of the repeated sequence elements**

Repeat	CRISPR sequence*	No. of repeats
1	cggttacCTCCGctcgcGCGGAGagcac	61
2	gtgctCTCCGcgcgaGCGGAGgtgggtcg	38
3	gggtccaTCCCCGCGggCGCGGGGAgcac	9
4	ctgctCCCCGCGcgcGCGGGGttggtccc	47

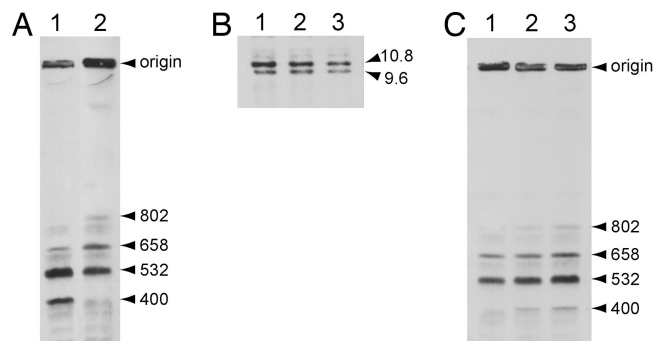
\*Capital letters indicate palindromic sequences.

The relative intensities of the 658-kb and 802-kb SspI bands, containing three and four copies of the Km gene cluster, respectively, and of the origin region were greater in cultures grown in the presence of Km (Fig. 7A, lane 2) than in its absence (Fig. 7A, lane 1), explaining the increase in copy number of the 10.8-kb BamHI fragment (Fig. 1B). In addition to these bands, and the 532-kb band containing two copies of the Km gene cluster, a new SspI fragment of 400 kb was detected in the culture grown in the absence of Km (Fig. 7A, lane 1). Ten single colonies obtained from the latter culture were subjected to PCR analysis for detection of the R<sub>s</sub>A/R<sub>s</sub>B junction; 40% lacked the sequence. These strains retained the R<sub>s</sub>C/R<sub>s</sub>D junction, confirmed by both PCR and Southern analysis. It was therefore concluded that the 400-kb SspI band arose by excision of the AUD from the chromosome of strain 12-6 by homologous recombination, yielding the structure shown in Fig. 4D.

To examine the stability of the chromosome population acquired after growth with Km, a portion of the resulting 12-6 culture was grown in the absence of the antibiotic. Southern analysis of BamHI-digested DNAs obtained after one and three rounds of growth in nonselective medium indicated a decrease in the copy number of the 10.8-kb BamHI fragment (Fig. 7B). Moreover, SspI digestion revealed the simultaneous appearance of the 400-kb band, increased intensity of the 532-kb and 658-kb bands, and a decrease in the intensity of the signal at the origin (Fig. 7C), all consistent with a decrease in the average copy number of the Km gene cluster.

**Km Production Varies with the Copy Number of the Km Gene Cluster.**

Strain 12-6 contains, on average, three copies of the Km gene cluster per genome. Isolate 12-6-4 contains a single copy of the



**Fig. 7.** Southern blot analyses of standard and PFGE gels. Fragment sizes in A, B, and C are given in kb. The 4.95-kb SphI fragment (Fig. 1A) was used as a probe. (A) Agarose plugs were prepared from mycelium of strain 12-6 grown without (lane 1) or with (lane 2) Km (1,000 μg/ml) and digested with SspI, and the DNA fragments were resolved by PFGE with the following conditions: agarose concentration, 1.0%; voltage gradient, 6 V/cm; angle, 120°; switch time, 50–90 s; run time, 22 h. (B) Genomic DNAs of strain 12-6 were prepared from mycelium grown with Km (lane 1) or after subculturing in fresh medium without Km once (lane 2) or three times (lane 3), digested with BamHI, and separated on a 0.8% agarose gel. (C) Agarose plugs for PFGE were made from the same preparations of mycelia used in B and separated on a PFGE gel after SspI digestion, using the same conditions as for A.



It would not be surprising to find similar amplifications of DNA segments, including entire antibiotic biosynthetic gene clusters in other industrial strains. Moreover, the insights gained in this study might prove valuable in engineering significant increases in productivity in future strain improvement programs.

## Materials and Methods

**Strains.** *S. kanamyceticus* NBRC13414 (wild-type) and its derivative 12-6 (Km-overproducing mutant) were obtained from the Meiji Seika Kaisha culture collection.

**Preparation of Genomic DNA and Chromosome Walking.** Genomic DNAs were prepared as in ref. 11. Cosmid libraries were made in SuperCos1 (Stratagene). Probe P was generated by using oligonucleotides KM-16' (5'-CCGGCACTTCCGCTCCAA-3') and KM-17' (5'-GCGGGTTCGCCAACTCCA-3') as primers and pKM92 as a template. For chromosome walking, a DNA library of the wild-type strain was constructed in SuperCos1 and probed by using a 1167-bp PCR fragment (Fig. 2, P), yielding cosmid 1-3. Chromosome walking gave five overlapping cosmids to the left (2-1, 4-5, 6-6, 8-4, and 10-4) and four to the right (3-7, 5-13, 7-14, and 9-3) of cosmid 1-3 (Fig. 2). A separate chromosome walking experiment using the 4.95-kb SphI fragment (Fig. 2, S) from the right end of the Km gene cluster identified three further cosmid clones (19-3, 20-8, and 20-10; Fig. 2). Clone 19-3 overlapped with 9-3 (demonstrated by restriction enzyme and hybridization analyses).

**PFGE.** *S. kanamyceticus* strains were grown in yeast extract malt extract medium containing 2% glycine (11). Agarose plugs were prepared and restriction endonuclease digestions performed as in ref. 11. A CHEF-DR III Pulsed-Field Electrophoresis System (Bio-Rad Laboratories) was used to resolve DNA digests. Agarose gels were made in 0.5× TBE buffer (45 mM Tris-borate/1 mM EDTA), also used as running buffer for electrophoresis. Electrophoresis conditions varied with the size of the DNA fragments to be resolved (see figure legends).

**Hybridizations.** Hybridizations were carried out by using an ECL Direct Detection System (Amersham Pharmacia Biosciences).

**DNA Sequencing and ORF Annotation.** Sequencing of pKM7 and pKM92 was carried out by Shimadzu Corporation (Kyoto); other cosmids were sequenced by Lark Technologies (Takeley, U.K.).

Putative ORFs were identified by using ARTEMIS (Release 7, Sanger Institute), and potential homologues were identified by using BLAST at National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>).

**Construction of Conjugative Cosmids.** For pMJCOS1, a 4.4-kb fragment containing the apramycin-resistance gene, *oriT*, and the  $\phi$ C31 integrase gene and attachment site was amplified by PCR with EcoRI-digested pSET152 (29) as a template and the oligonucleotides Super152-F (5'-CTGGGTCATTTTCGGC-GAGGACCGCTTTCGCTGGAGCGCGGTTTCATGTGC-AGCTCCATC-3') and Super152-R (5'-ACCACAGAAGTA-AGGTTTCCTTACAAAGATCCGGACCAAAATTCCC-CAATGTCAAGCAC-3') as primers, gel-purified, and introduced into *E. coli* BW25113/pIJ790/SuperCos1 (30) by electroporation. Transformants were selected on LB agar medium (11) containing carbenicillin (100  $\mu$ g/ml) and apramycin (50  $\mu$ g/ml), yielding pMJCOS1.

For pMJ20-10-1, a 5.2-kb SspI fragment of pMJCOS1 containing the apramycin-resistance gene, *oriT*, and the  $\phi$ C31 integrase gene and attachment site was gel-purified and introduced into *E. coli* BW25113/pIJ790/cosmid 20-10 by electroporation. Transformants were selected as above, yielding pMJ20-10-1.

**Quantitation of Km Production.** *S. kanamyceticus* strains were grown as described (1). Culture broths were mixed with an equal volume of 4M NaOH, boiled for 2 h, and neutralized by adding an equal volume of 2M HCl. After filtering, 5- $\mu$ l samples of the filtrate were subjected to HPLC on a L-column ODS (4.6 × 150 mm; Chemicals Evaluation and Research Institute, Tokyo, Japan) at a column temperature of 40°C using a mobile phase of 87% (vol/vol) 0.5 M sodium perchlorate, 6.9 mM 1-octansulfonate (adjusted to pH 2.5 with H<sub>2</sub>SO<sub>4</sub>), and 13% (vol/vol) acetonitrile at a flow rate of 1.6 ml/min. Fractions of the eluate were mixed with *o*-phthalaldehyde (31) pumped at a flow rate of 0.8 ml/min in a stainless-steel reaction coil (0.5 mm × 6 m) at 40°C. Derivatized Km was detected by measuring absorbance at 340 nm and eluted from the column at 16.1–16.4 min.

We thank Govind Chandra for assistance in sequence analysis, Naomi Sumida for technical assistance, and David Hopwood and Keith Chater for comments on the manuscript. K.Y. and T.M. were funded by Meiji Seika Kaisha Ltd., and M.B. was funded by a grant to the John Innes Centre from the Biotechnology and Biological Sciences Research Council.

- Umezawa, H., Ueda, M., Maeda, K., Yagishita, K., Kondo, S., Okami, Y., Utahara, R., Osato, Y., Nitta, K. & Takeuchi, T. (1957) *J. Antibiot.* **A10**, 181–188.
- Kondo, S. & Hotta, K. (1999) *J. Infect. Chemother.* **5**, 1–9.
- Yanai, K. & Murakami, T. (2004) *J. Antibiot.* **57**, 351–354.
- Kharel, M. K., Subba, B., Basnet, D. B., Woo, J. S., Lee, H. C., Liou, K. & Sohng, J. K. (2004) *Arch. Biochem. Biophys.* **429**, 204–214.
- Matsuhashi, Y., Murakami, T., Nojiri, C., Toyama, H., Anzai, H. & Nagaoka, K. (1985) *J. Antibiot.* **38**, 279–282.
- Volff, J.-N. & Altenbuchner, J. (1998) *Mol. Microbiol.* **27**, 239–246.
- Bibb, M. J., Findlay, P. R. & Johnson, M. W. (1984) *Gene* **30**, 157–166.
- Bentley, S. D., Chater, K. F., Cerdeño-Tarraga, A.-M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., et al. (2002) *Nature* **417**, 141–147.
- Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. & Omura, S. (2003) *Nat. Biotechnol.* **21**, 526–531.
- Huang, C.-H., Chen, C.-Y., Tsai, H.-H., Chen, C., Lin, Y.-S. & Chen, C. W. (2003) *Mol. Microbiol.* **47**, 1563–1576.
- Kieser, P., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000) *Practical Streptomyces Genetics* (The John Innes Foundation, Norfolk, U.K.)
- Sambrook, J. & Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY), 3rd Ed.
- Eichenseer, C. & Altenbuchner, J. (1994) *J. Bacteriol.* **176**, 7107–7112.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.-A. & Barrell, B. (2000) *Bioinformatics* **16**, 944–945.
- Mojica, F. J. M., Díez-Villaseñor, C., Soria, E. & Juez, G. (2000) *Mol. Microbiol.* **36**, 244–246.
- Jansen, R., van Embden, J. D. A., Gaastera, W. & Schouls, L. M. (2002) *Mol. Microbiol.* **43**, 1565–1575.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
- Bentley, S. D., Brown, S., Murphy, L. D., Harris, D. E., Quail, M. A., Parkhill, J., Barrell, B. G., McCormick, J. R., Santamaria, R. I., Losick, R., et al. (2004) *Mol. Microbiol.* **51**, 1615–1628.
- Yamasaki, M., Ikuto, Y., Ohira, A., Chater, K. & Kinashi, H. (2003) *Microbiology* **149**, 1351–1356.
- DeBoy, R. T., Mongodin, E. F., Emerson, J. B. & Nelson, K. E. (2006) *J. Bacteriol.* **188**, 2364–2374.
- Peng, X., Brügger, K., Shen, B., Chen, L., She, Q. & Garrett, R. A. (2003) *J. Bacteriol.* **185**, 2410–2417.
- Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. (2006) *Biol. Direct* **1**, 7.
- Cramer, R. & Davies, J. E. (1986) *J. Antibiot.* **39**, 128–135.
- Fierro, F., Barredo, J. L., Díez, B., Gutierrez, S., Fernández, F. & Martín, J. F. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 6200–6204.
- Peschke, U., Schmidt, H., Zhang, H.-Z. & Piepersberg, W. (1995) *Mol. Microbiol.* **16**, 1137–1156.
- Schmid, E., Büchler, C. & Altenbuchner, J. (1999) *Microbiology* **145**, 3331–3341.
- Young, M. & Cullum, J. (1987) *FEBS Lett.* **212**, 10–14.
- Ishikawa, T. (1961) Ph.D. thesis (University of Tokyo, Tokyo).
- Bierman, M., Logan, R., O'Brien, K., Seno, E. T., Rao, R. N. & Schoner, B. E. (1992) *Gene* **116**, 43–49.
- Gust, B., Challis, G. L., Fowler, K., Kieser, T. & Chater, K. F. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 1541–1546.
- Kubo, H., Kobayashi, Y. & Nishikawa, T. (1985) *Antimicrob. Agents Chemother.* **28**, 521–523.