Open Access

OBM Genetics

Original Research

Deletion of Subtelomeric Regions in the Linear Chromosome of Variants Isolated from *Streptomyces avermitilis* NBRC 14893^T

Hisayuki Komaki *, Tomohiko Tamura

Biological Resource Center, National Institute of Technology and Evaluation (NBRC), 2-5-8 Kazusakamatari, Kisarazu, Chiba 292-0818, Japan; E-Mails: komaki-hisayuki@nite.go.jp, tamura-tomohiko@nite.go.jp

* **Correspondence:** Hisayuki Komaki; E-Mail: komak-hisayuki@nite.go.jp

Academic Editors: Ben Krause-Kyora, Sergey Pisarenko, Dmitry Kovalev and Alexander Kulichenko

Special Issue: Bacterial Genomes

OBM Genetics
2019, volume 3, issue 3
doi:10.21926/obm.genet.1903091

Received: April 19, 2019
Accepted: August 12, 2019
Published: August 23, 2019

Abstract

***Background*:** *Streptomyces avermitilis* is an actinomycete that produces avermectins. The complete genome sequence of *S. avermitilis* K139 was determined in 2003. In our previous study, we revealed that *S. avermitilis* NBRC 14893^T harbors two extra secondary metabolite-biosynthetic gene clusters (smBGCs), *com* and *ptx*, which are not present in strain K139.

***Methods*:** Whole-genome sequencing for *S. avermitilis* NBRC 14893^T and ATCC 31267^T was carried out by the SMRT DNA sequencing technology to determine the genomic loci encoding *com* and *ptx* gene clusters.

***Results*:** The genome size of both the strains, NBRC 14893^T, and ATCC 31267^T, was 10.5 Mb, which was 1.4 Mb-larger than that of strain K139. The chromosomes of these type strains included extra 1.1 Mb-long right-hand region that is not present in the K139 genome and encodes *com* and *ptx* gene clusters. The variants without *ptx* and/or *com* were spontaneously obtained while isolating single colonies from strain NBRC 14893^T. The genome analysis of the variants indicated that the chromosome has been deleted step-by-step from the subtelomeric region(s) in strains NBRC 14893^T.

***Conclusions*:** These results suggest that the linear chromosome of *S. avermitilis* is highly unstable. This study offers an experimental approach to glean the genome information and delineate the mechanism of genomic instability in the genus *Streptomyces*.

Keywords
Biosynthetic gene cluster; genome; instability; linear chromosome; secondary metabolite; *Streptomyces avermitilis*; subtelomeric region; type strain; variant

1. Introduction

The genus *Streptomyces,* a well-known representative actinomycete, includes Gram-positive bacteria that form filamentous mycelia, aerial hyphae, and conidial spores during their life cycle [1, 2]. The most attractive property of the genus is its ability to produce structurally diverse secondary metabolites, many of which have been developed into medicines such as antibiotics, antitumor agents, and immunosuppressants. Hence, the members of the genus *Streptomyces* are considered as a rich source of bioactive secondary metabolites for pharmaceutical developments [1, 3, 4]. *Streptomyces avermitilis* MA-4680^T (NRRL 8165^T) was isolated from a soil sample collected in Shizuoka, Japan in 1978 and was found to produce potent anthelmintic compounds (avermectins) [5]. A semisynthetic derivative, Ivermectin, is commercially used as an antiparasitic agent and agricultural pesticide. Nowadays, Ivermectin is also used to treat many tropical diseases such as onchocerciasis, strongyloidiasis, trichuriasis, and lymphatic filariasis. Therefore, *S. avermitilis* is one of the most important industrial microorganisms for the production of antiparasitic agents, veterinary medicine, and agricultural pesticides [6].

A draft genome sequence of *S. avermitilis* ATCC 31267^T was first reported by Omura *et al.* in 2001, covering 99% of the whole genome [7]. Subsequently, the same research group reported the complete genome sequence of *S. avermitilis* in 2003 [8] and registered it as that of *S. avermitilis* MA-4680^T in GenBank. The genome is composed of a linear chromosome with 9.02 Mb and a linear plasmid, SAP1, with 94 kb. *S. avermitilis* is also known to harbor another linear plasmid SAP2 in addition to SAP1 [9]; however, SAP2 has been reported to be easily eliminated during cultivation, while SAP1 is stably maintained [6]. Until now, in total 38 types of secondary metabolite-biosynthetic gene clusters (smBGCs) have been discovered from the complete genome sequence, all of which are encoded in the chromosomal DNA [6]. More than half of the smBGCs are located in the subtelomeric regions, while the biosynthetic gene clusters (BGCs) for the secondary metabolites, commonly produced by several *Streptomyces* species, such as geosmin (*geo*), pentalenolactone (*plt*) and oligomycin (*olm*), are present in the 6.5-Mb internal conserved region of the chromosome [8].

In our previous study, we revealed that *S. avermitilis* NBRC 14893^T harbors two extra smBGCs, a nonribosomal peptide synthetase (NRPS) gene cluster, and a hybrid NRPS/polyketide synthase (PKS) gene cluster, which were not observed in the already published *S. avermitilis* complete genome sequences. Our bioinformatic analysis suggested that the two gene clusters are likely involved in the synthesis of demethyl complestatin and a novel oxazolomycin analog, respectively [10]. In a recent study, the hybrid NRPS/PKS gene cluster has been reported to be involved in the

synthesis of phthoxazolin A, which corresponds to the partial structure of oxazolomycin [11]. In the present study, we conducted whole-genome sequencing for *S. avermitilis* type strains by the SMRT DNA sequencing technology using PacBio to determine the genomic locus encoding the two smBGCs and compared the genomes, focusing on the presence of smBGCs.

2. Materials and Methods

2.1 Bacterial Strains

Two *S. avermitilis* type strains, NBRC 14893^T and ATCC 31267^T, were obtained from the NBRC culture collection and the ATCC culture collection, respectively. The history of *S. avermitilis* type strains among culture collections is shown in Figure 1. The variants, MC3 and cw2, were isolated from *S. avermitilis* NBRC 14893^T by the procedure stated in Section 2.3.

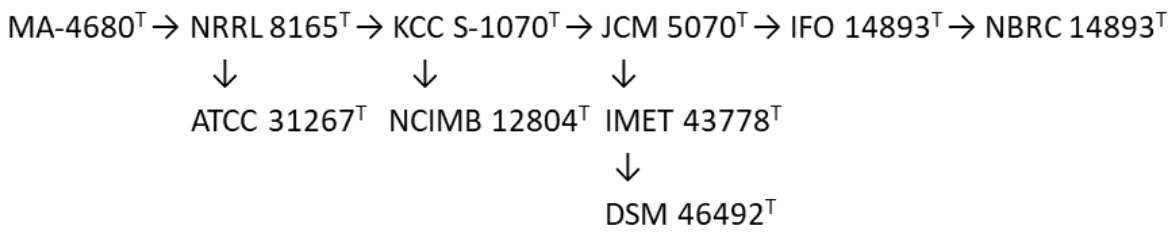

Figure 1 History of *S. avermitilis* type strains.

2.2 Genome Analysis

2.2.1 Genomic DNA Preparation

Each strain was inoculated into 15mL of YG medium (1% yeast extract, 1% glucose, pH 7.0) in three test tubes (diameter, 18mm; length, 16.5cm; 5mL YG/tube) and then cultivated with shaking (260 strokes/min) at 28°C using an incubator shaker TC-500R (Takasaki Kagaku Kiki Co., Ltd., Saitama, Japan). After one or two days, the cultured cells were harvested by centrifugation (5,255xg, 10min, 4°C) and preserved at –20°C until genome DNA preparation. Genomic DNAs were prepared according to the procedure reported previously [12].

2.2.2 Genome Sequencing

The whole-genome sequencing was carried out by the whole-genome *de novo* sequencing service of Macrogen Korea using the SMRT strategy with PacBio RSII with SMRT cell 8Pac V3 and DNA Polymerase Binding Kit P6. The reads of each strain were assembled using Canu v. 1.4.

2.2.3 Survey of Secondary Metabolite-Biosynthetic Gene Clusters (smBGCs)

SmBGCs in the genome sequences were searched using antiSMASH [13] and then manually checked. If antiSMASH failed to detect smBGCs in the genome data of Ikeda *et al*. [6, 8], we searched them by BLASTn using BioEdit. Annotation for each BGC was carried out according to previous reports [6] and the database of *Streptomyces avermitilis* Genome Project (http://avermitilis.ls.kitasato-u.ac.jp).

2.2.4 Accession Numbers

The genome sequences of *S. avermitilis* strains NBRC 14893^T, ATCC 31267^T, MC3 and cw2 have been deposited into the DDBJ under the accession numbers BJHX01000001-BJHX01000006, BJHY01000001-BJHY01000004, AP019621-AP019622, and BJHZ01000001-BJHZ01000004, respectively.

2.3 Isolation of Variants

2.3.1 Isolation of a Variant, MC3, from *S. avermitilis* NBRC 14893^T

S. avermitilis NBRC 14893^T was inoculated on an NBRC 802 agar plate (1% polypeptone, 0.2% yeast extract, 0.1% $MgSO_4 \cdot 7H_2O$, pH 7.0, 1.5% agar) and an ISP-4 agar plate (1% soluble starch, 0.1% K_2HPO_4, 0.1% $MgSO_4 \cdot 7H_2O$, 0.1% NaCl, 0.2% $(NH_4)_2SO_4$, 0.2% $CaCO_3$, 0.0001% $FeSO_4 \cdot 7H_2O$, 0.0001% $MnCl_2 \cdot 4H_2O$, 0.0001% $ZnSO_4 \cdot 7H_2O$, pH 7.0–7.4, 2% agar) and cultivated at 28°C for nine days. One bald colony (MC3) and two gray and two white colonies (MC4 to 7) were picked up from NBRC 802 and ISP-4 agar plates, respectively. The variant, MC3, has been deposited into the NBRC culture collection under the strain number NBRC 114117.

2.3.2 Detection of *com* and *ptx* Genes

Genome DNA preparation. The variants were inoculated into 3 mL of YG medium in a test tube and cultured with shaking at 28°C for two days. The cells were harvested by centrifugation and stored at –20°C until genome DNA preparation. The genomic DNAs were prepared using DNeasy Blood & Tissue Kit (Qiagen KK, Tokyo, Japan) according to the protocol for Gram-positive bacteria.

PCR for *com* and *ptx* genes. The nucleotide sequences of the primer set for *com* gene were TGAACATCCGTTGTCTTACG and TACTCACGAACACGAGAAGC, while those for *ptx* gene were GTCAACTTCCTGTGGACGAT and TTGTACAACTCCCCCACTTT. The reaction mixtures contained 1x GoTaq Green Master Mix (Promega KK, Tokyo, Japan), 0.4 μM of each primer, 5% DMSO and 1 μL of genomic DNA in a final volume of 20 μL. The reaction was started with denaturation at 94°C for 5 min followed by a total of 25 cycles that consisted of denaturation at 94°C for 30 sec, annealing at 55°C for 30 sec, and extension at 72°C for 30 sec. The reaction was terminated with a final extension at 72°C for 5 min. The amplification was confirmed by agarose gel electrophoresis.

2.3.3 Isolation of a Variant, cw2, from MC3

The variant, MC3, was cultivated on NBRC 802, YG, and ISP-4 agar plates at 28°C for four days. Four colonies were picked up from each plate and the presence of the *ptx* gene in these total 12 colonies was examined by PCR as stated above. All colonies were found to harbor the *ptx* gene. After cultivating 9 of the 12 colonies on YG agar plates at 28°C for six days, a single colony was picked up again from each plate, inoculated onto an HV agar plate [14] supplemented with 50 mg/L of cycloheximide, and grown at 28°C for two days. From each of the 9 plates, 4 colonies were picked up and cultured in 3 mL of YG medium. Since 25 of the 36 colonies grew well within four days, the presence of *ptx* gene in the 25 colonies was examined by PCR. Consequently, the gene was not amplified from two strains, cw2 and cw3, thus we chose cw2 for further study. The variants, cw2 and cw3, have been deposited into the NBRC culture collection under the strain number NBRC 114118 and NBRC 114119, respectively.

3. Results

3.1 Genome Sequencing of S. avermitilis *Type Strains*

The whole genome of *S. avermitilis* NBRC 14893^T and ATCC 31267^T was sequenced using PacBio (Table 1). The sequencing for *S. avermitilis* NBRC 14893^T yielded six contig sequences and the total genome size was 10.5 Mb. Based on the sizes and similarities (data not shown), contigs #4 and #5 were considered as SAP2 and SAP1 plasmids, respectively. The other contigs (#1, #2, #3 and #6) were derived from the chromosome sequence, though they could not be assembled into a single sequence. From *S. avermitilis* ATCC 31267^T, four contig sequences were obtained and the total genome size was 10.5 Mb. We considered that contigs #1 and #2 are the chromosomal sequence, while contigs #3 and #4 are SAP2 and SAP1 plasmids, respectively. These type strains shared almost the same genome size but unexpectedly this genome size was larger by 1.4 Mb than that of the already reported complete genome sequence of *S. avermitilis* (9.1 Mb). Additionally, SAP2 was present in both type strains whereas it was absent in the complete genome sequence. According to Suroto *et al.,* the previously sequenced strain by Ikeda *et al.* in 2003 was designated as K139, a progeny derived from KA–320^T (isogenic to MA-4680^T, ATCC 31267^T & NRRL 8165^T) [11, 15]. This accounts for the differences regarding the genome sizes and the presence of SAP2 between our results and already published one.

Table 1 Genome sequencing of *S. avermitilis* type strains.

Strain	Total genome size (bp)	G+C (%)	Contig	Length (bp)	G+C (%)
S. avermitilis NBRC 14893^T	10,484,508	70.5	#1	9,069,390	70.8
			#2	713,930	69.8
			#3	302,286	68.7
			#4^p	242,615	67.5
			#5^p	105,485	69.0
			#6	50,802	69.2
S. avermitilis ATCC 31267^T	10,503,935	70.5	#1	9,110,247	70.8
			#2	1,046,427	69.4
			#3^p	239,314	67.5
			#4^p	107,947	69.0

^{p} Plasmid

3.2 Difference among Plasmids of S. avermitilis *Type Strains and K139*

The two type strains retained two plasmids SAP1 and SAP2 but their size was different between the strains. As shown in Figure 2, the difference in the size of SAP1 between NBRC 14893^T and ATCC 31267^T was mainly derived from extra 2.4 kb in the left arm of ATCC 31267^T. Except for this difference, these two SAP1 sequences shared 99.9% similarity in the whole. The extra left arm showed 99.6% similarity with the region from 25 to 27 kb of NBRC 14893^T (green lines in Figure 2), suggesting that this region has been duplicated in ATCC 31267^T. SAP1 plasmids of NBRC 14893^T and ATCC 31267^T were 11.2 kb- and 13.7 kb-larger than that of K139, respectively. K139 SAP1

sequence showed 99.9% similarity to those of NBRC 14893^T and ATCC 31267^T. However, the region from 81.4 to 93.6 kb also showed 99.2% similarities to the left 12.2 kb of the type strains (blue lines). This suggests that the region has been duplicated in SAP1s of NBRC 14893^T and ATCC 31267^T but not in that of K139. As shown by red lines, 0.1 to 0.2 kb-length region near the left terminal showed 89% to 94% similarities to the right terminals of NBRC 14893^T and ATCC 31267^T. About a 120 bp-sequence from 40.47 to 40.59 kb was complementarily present in the SAP1s of NBRC 14893^T and ATCC 31267^T. The right terminal sequence from 0 to 0.23 kb is specific in K139.

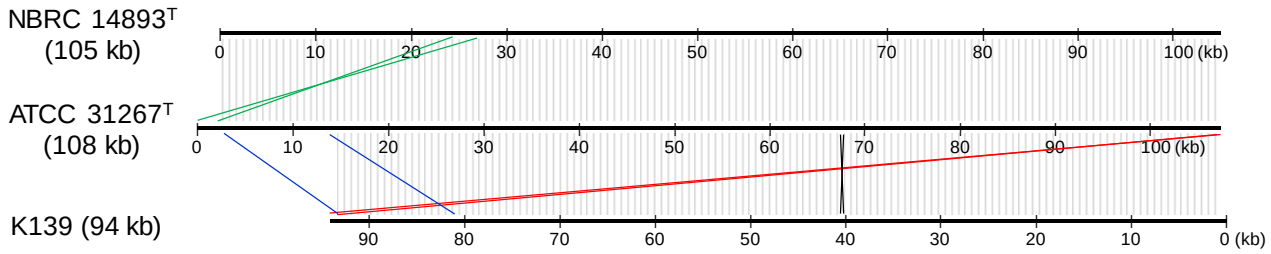

Figure 2 Comparison of SAP1 plasmids among *S. avermitilis* strains. SAP1 plasmids are shown as black bold lines. The value given in brackets is the size of the plasmid. Homologous regions are connected by vertical lines in a light gray color. The regions shown by green, blue, red colors, and thin black lines are also homologous.

The entire SAP2 sequence of ATCC 31267^T showed 99.97% similarity with that of NBRC 14893^T. SAP2 plasmid of NBRC 14893^T included extra 2.2 kb of the left arm and 1.1 kb of the right arm, which are not present in ATCC 31267^T (data not shown).

3.3 Secondary Metabolite-Biosynthetic Gene Clusters

We next surveyed smBGCs in the genomes of strains NBRC 14893^T and ATCC 31267^T and then compared them with that of strain K139. No smBGC was found in SAP1 and SAP2 plasmids. As shown in Figure 3, NBRC 14893^T and ATCC 31267^T encoded 41 different types of smBGCs in the chromosomes, all of which were shared with the type strains. BGCs for demethyl complestatin (*com*) and phthoxazolin A (*ptx*), which we identified in the previous work [10], were encoded in the right subtelomeric region. In contrast, the chromosome of K139 is 1.1 Mb-shorter than those of type strains. This might have been caused by the loss of right subtelomeric region and consequently lacking the type-I PKS (*t1pks*), *ptx,* and *com* gene clusters. However, *pks11* and *ams* gene clusters, which are also encoded in the right subtelomeric region in the type strains, were present in the left subtelomeric region in strain K139.

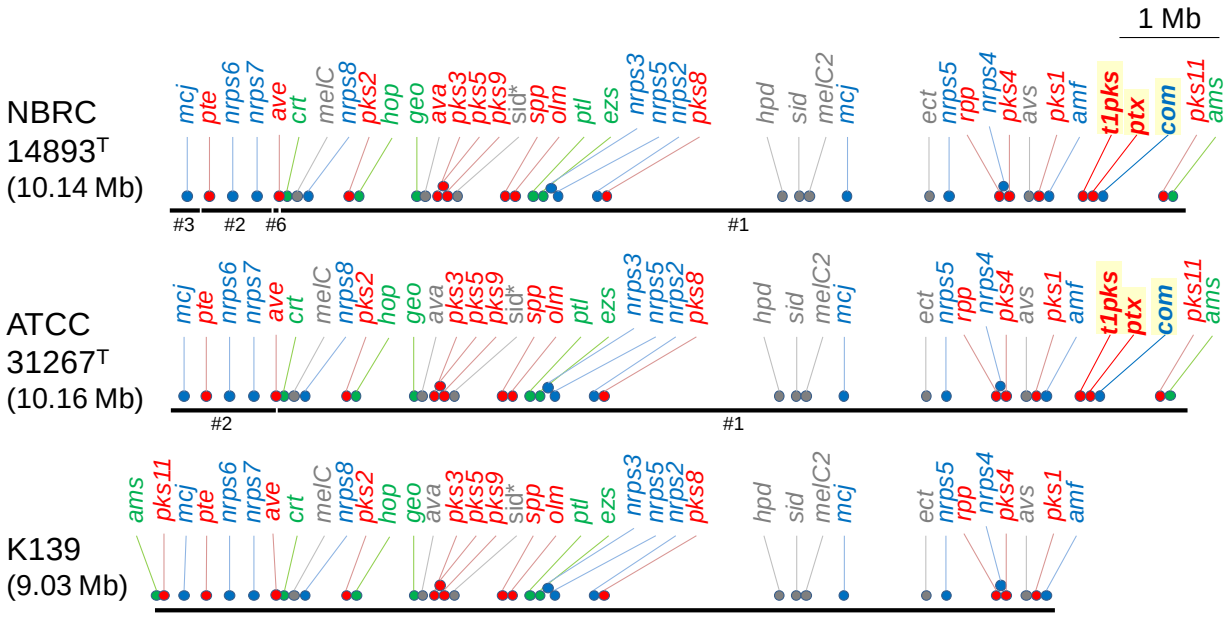

Figure 3 Secondary metabolite-biosynthetic gene clusters (smBGCs) in *S. avermitilis* chromosomes. The chromosomes are indicated as black bold lines. Contig numbers in Table 1 are shown below the lines. The contig arrangements of #3, #2 and #6 in NBRC 14893^T and #2 and #1 in ATCC 31267^T were verified by bridge PCRs amplifying the gaps. All arrangements of contigs were also confirmed by comparison with the complete genome sequence of K139. The values in brackets are sizes of chromosomes. SmBGCs are color-coded according to the putative products as follows: red, polyketide; blue, peptide; green, terpenoid; gray, others. The experimentally identified products of smBGCs are as follows: *ams*, avermitilol and avermitilone; *ave*, avermectin; *crt*, isorenieratene; *ect*, ectoine and 5-hydroxyectoine; *ezs*, albaflavenol and albaflavenone; *geo*, geosmin and germacradienol; *hpd*, ochronotic pigment; *hop*, squalene; *melC*, melanin; *olm*, oligomycin; *pte*, filipin; *ptl*, neopentalenoketolactone; *ptx*, phthoxazolin; *sid*, nocardamine and desferrioxamine B: *spp*, spore pigment. The predicted products: *amf*, lantipeptide; *ava*, 2-alkyl-4-hydroxymethylfuran-3-carboxylic acids (γ-butyrolactone?); *avs*, vibrioferrin-like siderophore; *com*, demethyl complestatin; *melC2*, melanin; *mcj*, lassopeptide; *rpp,* tetrahydroxynaphthalene. SmBGCs specific in type strains are boldfaced and light yellow-shaded. *sid, Since the cluster is annotated for siderophore synthesis but no specific gene name is suggested, it is putatively indicated as not italic "sid" in this figure.

3.4 Isolation of Variants Lacking com *and* ptx *Gene Clusters*

It is speculated that 1.1 Mb in the right subtelomeric region has have been eliminated during the isolation of K139 from KA-320^T but we wondered whether such a large region can easily be deleted from the chromosome. In order to clarify the question, we tested if a variant whose chromosome is shortened can occur spontaneously. At first, we isolated five single colonies from NBRC 14893^T and then examined if they harbor *com* and *ptx* genes, encoded in the right subtelomeric region, by PCR. Four isolates were found to harbor both genes. However, one isolate, namely MC3, harbored *ptx* gene but lacked *com* gene. Next, we examined if the variants excluding

ptx gene can be obtained from MC3 by single colony isolation. We isolated 12 single colonies from MC3 but no isolate devoid of the *ptx* gene was observed. Subsequently, we further repeated single colony isolation twice as stated in 2.3.3 and examined 25 colonies. Among them, only two colonies, namely cw2 and cw3, lacked *ptx* gene. For further study, we chose MC3 and cw2 as variants lacking only *com* gene and both genes, respectively.

3.5 Genome Analysis of Variants without ptx *and/or* com *Gene Cluster (s)*

We sequenced the whole genomes of *S. avermitilis* variants MC3 and cw2 using PacBio (Table 2). Two contig sequences were obtained from MC3 and the total genome size was 9.6 Mb. Contig #1 with 9.4 Mb and contig #2 with 108 kb corresponded to the chromosome and SAP1 plasmid, respectively. From cw2, four contig sequences were obtained and the total genome size was 8.3 Mb. As cw2 possessed no plasmid, all the sequences belonged to only the chromosome, although they could not be assembled into a single sequence.

Table 2 Genome sequencing of *S. avermitilis* variants.

Strain	Total genome size (bp)	G+C (%)	Contig	Length (bp)	G+C (%)
S. avermitilis MC3	9,557,539	70.7	#1	9,449,866	70.7
			#2[p]	107,673	68.9
S. avermitilis cw2	8,300,465	70.9	#1	7,644,811	70.9
			#2	458,943	70.6
			#3	103.764	69.3
			#4	92.947	69.4

[p] Plasmid

We next surveyed smBGCs in the genomes of these variants and compared them with those of strain NBRC 14893^T. As shown in Figure 4, MC3 and cw2 chromosomes encoded 40 and 33 different types of smBGCs, respectively. The chromosome of MC3 was 0.7 Mb-shorter than that of NBRC 14893^T and devoid of *com* gene cluster. However, *pks11* and *ams* gene clusters, encoded in the right-arm terminal in NBRC 14893^T, were present in the left subtelomeric region. The assembly of smBGCs in MC3 was more similar to K139 except for the presence of *t1pks* and *ptx* (Figure 3). The chromosome of cw2 was 1.1 Mb-shorter than that of MC3. The left subtelomeric region was more deleted than the right one. The seven smBGCs indicated in boldfaces in the figure were found to disappear from MC3 to cw2. Among the seven, *ams, pks11, mcj, pte,* and *nrps11* were encoded in the left subtelomeric region, while *t1pks* and *ptx* were in the right in MC3. The other 33 smBGCs, from *nrps7* to *amf*, encoded in the chromosomal core were stably conserved among type strains, the variants and K139.

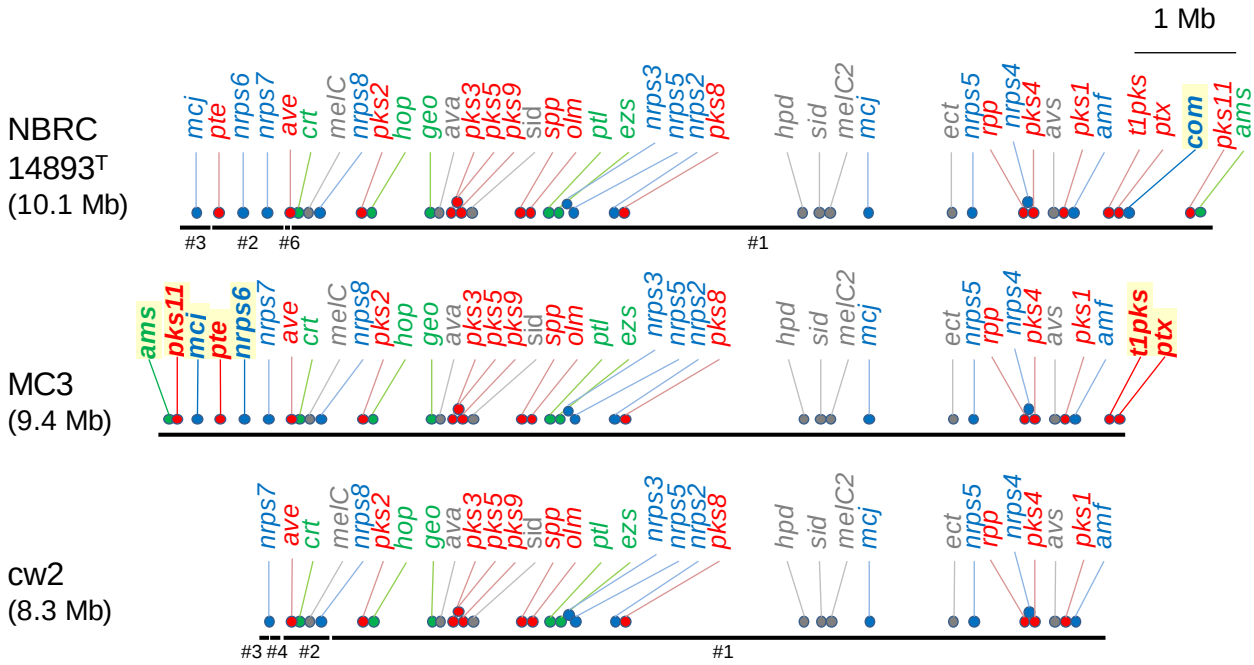

Figure 4 SmBGCs in the chromosomes of *S. avermitilis* NBRC 14893^{T} and the variants. The chromosomes are indicated as black bold lines. Contig numbers in Table 2 are shown below the chromosomes. The contig arrangement of #3, #4, and #2 in cw2 was verified by bridge PCRs amplifying the gaps and all arrangements of contigs were also confirmed by the comparison with complete chromosome sequences of K139 and MC3. The values in brackets are sizes of chromosomes. SmBGCs are color-coded according to the putative products as follows: red, polyketide; blue, peptide; green, terpenoid; gray, others. The experimentally identified products of smBGCs are as follows: *ams*, avermitilol and avermitilone; *ave*, avermectin; *crt*, isorenieratene; *ect*, ectoine and 5-hydroxyectoine; *ezs*, albaflavenol and albaflavenone; *geo*, geosmin and germacradienol; *hpd*, ochronotic pigment; *hop*, squalene; *melC*, melanin; *olm*, oligomycin; *pte*, filipin; *ptl*, neopentalenoketolactone; *ptx*, phthoxazolin; *sid*, nocardamine and desferrioxamine B: *spp*, spore pigment. The predicted products: *amf*, lantipeptide; *ava*, 2-alkyl-4-hydroxymethylfuran-3-carboxylic acids (γ-butyrolactone?); *avs*, vibrioferrin-like siderophore; *com*, demethyl complestatin; *melC2*, melanin; *mcj*, lassopeptide; *rpp,* tetrahydroxynaphthalene. SmBGCs deleted in the variant(s) are boldfaced and yellow-shaded in this figure. *sid, Since the cluster is annotated for siderophore synthesis but no specific gene name is suggested, it is putatively indicated as not italic "sid" in this figure.

4. Discussion

In general, replication origins (*oriC*) of most *Streptomyces* species are in the middle of the linear chromosomes. It has been reported that subtelomeric regions of *S. avermitilis* K139 are 2.0 Mb from the left telomere and 0.5 Mb from the right one. Such asymmetry has been considered to be specific to *S. avermitilis*. Consequently, *oriC* of *S. avermitilis* K139 is not located at the middle but has shifted 0.8 kb away from the center toward the right (upper, Figure 5) [8]. Our present study revealed that *S. avermitilis* type strains have much longer right subtelomeric regions in the

chromosomes than *S. avermitilis* K139. Therefore, like the other *Streptomyces* species, the *S. avermitilis* chromosome is actually symmetric on the length of both subtelomeric regions and its *oriC* is located at the middle of the chromosome (lower, Figure 5).

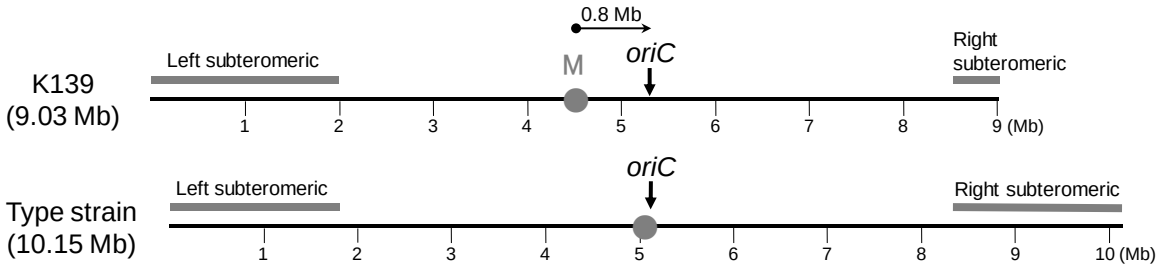

Figure 5 The position of *oriC* and middle region in the linear chromosomes of *S. avermitilis* strains. The chromosomes are indicated as black bold lines. Gray circles (M) are the centers of each chromosome. Subtelomeric regions are highlighted by gray bars. The position of *oriC* has shifted 0.8 Mb away from the center toward the right in K139 chromosome, whereas the same is at the center in chromosomes of the type strains.

Genetic instability of *Streptomyces* strains has been known and large DNA rearrangements involving gene duplication, elimination, and acquisition have frequently been observed in subtelomeric regions [16, 17]. In *Streptomyces ambofaciens*, up to 2.5 Mb located at the ends of the chromosome can be lost [18]. In contrast, a phenotype of *S. avermitilis* has been reported to be more stable than others [6, 19]. In a present study, we compared plasmid sequences among various *S. avermitilis* strains. Consequently, the elimination and/or duplication were observed in the subtelomeric regions, suggesting terminal instability of the linear plasmids. Next, we spontaneously obtained a variant, MC3, whose right subtelomeric region was shorter than the type strains. This may suggest instability of the linear chromosome also in *S. avermitilis*. Chen *et al.* reported chromosomal instability in *S. avermitilis* but focused on major deletion in the central region [20]. Therefore, our study is the first report on major deletion in the subtelomeric regions of *S. avermitilis* chromosome. The type strains encode *ams* and *pks11* gene clusters in the right arms of the chromosomes, whereas the right arm shortened variants such as MC3 and K139 encode them in the left arms (Figure 3 and Figure 4). We checked the presence of some regions involved in the chromosomal rearrangement between *pks11* and *mcj* gene clusters of MC3 and K139. Over 10 kb of the region just after the *pks11* gene cluster showed about 99.6% sequence similarities with a region upstream of *mcj* gene cluster and near the left end in the type strain (pink arrows, upper panel, Figure 6). The possible mechanism of the relocation of *ams* and *pks11* gene clusters from right to left could be shown in Figure 6. The linear chromosome is circularized by the terminal proteins [21]. Afterward, homologous recombination occurs between the regions shown in pink. Consequently, these two gene clusters move from right to left subtelomeric regions in the linear chromosome.

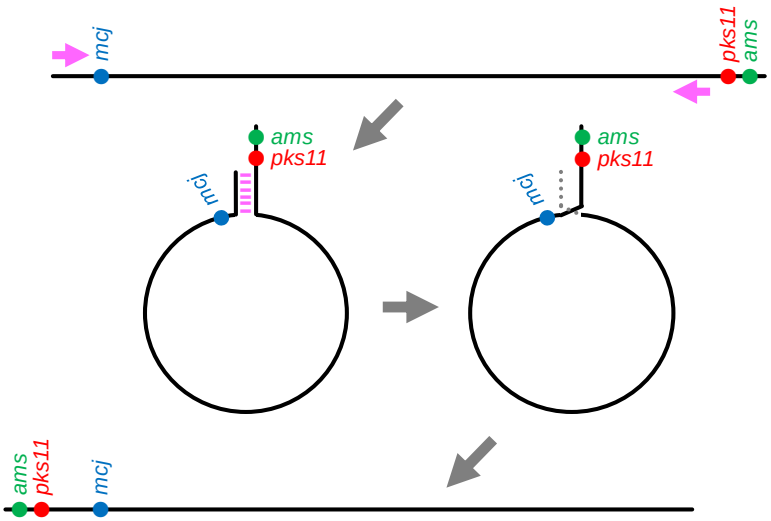

Figure 6 Mechanism on the relocation of *ams* and *pks11* gene clusters from the right to left arm in the linear chromosome. Homologous regions are shown as pink arrows and the hybridization is illustrated using pink short lines between them.

In the variants MC3 and cw2, all the BGC-missing regions are terminals of their chromosomes. We examined similarities of the regions between each variant and its parental strain (Figure 7). The 127 kb from the left terminal of MC3 chromosome showed 99.94% similarity to the right terminal region in its parental strain NBRC 14893^T (blue arrows, Figure 7a). The relocation can be explained by the mechanism shown in Figure 6. In contrast, 103 kb from the right terminal of MC3 chromosome (red arrow) showed 99.96% similarity to SAP2. The deletion of the inner left subtelomeric region of NBRC 14893^T may be coupled with the recombination of SAP2. Contig #3 corresponding to the left terminal region of cw2 showed 99.95% similarity to the inner left subtelomeric region of the chromosome in its parent strain MC3. The regions of 16 kb from the right terminal region (green arrow) and subsequent 31 kb (light brown arrow) showed 99.76% and 99.91% similarities to a part in the left arm of MC3 as shown in Figure 7b. When both the subtelomeric regions get deleted from MC3, the part might duplicate and relocate to the right terminal in the chromosome of cw2.

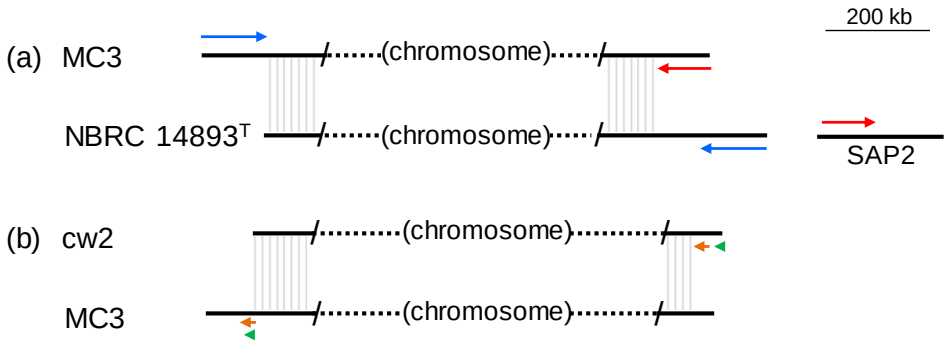

Figure 7 Comparisons of terminal regions between each variant and its parental strain. NBRC 14893^T is the parent of MC3 (a). MC3 is the parent of cw2 (b). Homologous regions are connected by vertical lines in the light gray (more than 99.94% sequence similarity) or shown using colored arrows for each pair. In this figure, the homologous regions within the chromosome are not shown.

We isolated MC3 only by examining five single colonies from NBRC 14893^T. In contrast, further chromosome-shortened variants were not observed by isolating 12 single colonies from MC3 and then we had to repeat additional single colony isolation twice to obtain a variant lacking *ptx*. Therefore, once a single colony is isolated, the sub-strain may genetically be stable for a few generations. If this assumption is correct, it is reasonable to say that the phenotype of K139 is stable because K139 is a progeny that is probably isolated from KA-320^T as a single colony.

During this study, draft genome sequence of *S. avermitilis* DSM 46492^T was published in GenBank under the accession numbers LMWT01000001 to LMWT01000124. The genome size is 8.95 Mb and thus it is 1.6 Mb smaller than those of ATCC 31267^T and NBRC 14893^T. It is unclear why the genome of DSM 46492^T is much smaller, although DSM 46492^T must be the type strain isogenic to ATCC 31267^T and NBRC 14893^T. We are speculating that IMET or DSM might have incidentally isolated a variant whose chromosome was deleted from JCM 5070^T and then preserved it as DSM 46492^T or the research group who sequenced the whole genome of DSM 46492^T may have isolated such a variant from DSM 46492^T and sequenced it.

5. Conclusions

Whole-genome sequencing for *S. avermitilis* type strains revealed that the linear chromosomes possess extra 1.1 Mb-long right subtelomeric region that is not present in K139 genome and encodes secondary metabolite-biosynthetic gene clusters. The variants with much shorter subtelomeric region(s) were spontaneously obtained from *S. avermitilis* NBRC 14893^T by single colony isolation(s). These results suggest the instability of subtelomeric regions in the linear chromosome of *S. avermitilis*. The information gathered herein could be useful in further experiments delineating their genome information to clue the mechanism for genomic instability in the genus *Streptomyces*.

Acknowledgments

We are grateful to Mr. Shinpei Ino for preparing genomic DNAs. We also thank Dr. Natsuko Ichikawa for registering the genome sequences on the DDBJ.

Author Contributions

HK designed the study, conducted experiments and analysis, and wrote the manuscript. TT advanced the genome sequencing and edited the manuscript.

Funding

This study was supported in part by the commissioned project from the Japan Patent Office.

Competing Interests

The authors have declared that no competing interests exist.

References

1. Waksman SA, Henrici AT. The nomenclature and classification of the actinomycetes. J Bacteriol. 1943; 46: 337-341.
2. Witt D, Stackebrandt E. Unification of the genera *Streptoverticillum* and *Streptomyces*, and amendation of *Streptomyces* Waksman and Henrici 1943, 339^{AL}. Syst Appl Microbiol. 1990; 13: 361-371.
3. Demain AL. Pharmaceutically active secondary metabolites of microorganisms. Appl Microbiol Biotechnol. 1999; 52: 455-463.
4. Watve MG, Tickoo R, Jog MM, Bhole BD. How many antibiotics are produced by the genus *Streptomyces*? Arch Microbiol. 2001; 176: 386-390.
5. Burg RW, Miller BM, Baker EE, Birnbaum J, Currie SA, Hartman R, et al. Avermectins, new family of potent anthelmintic agents: Producing organism and fermentation. Antimicrob Agents Chemother. 1979; 15: 361-367.
6. Ikeda H, Shin-ya K, Omura S. Genome mining of the *Streptomyces avermitilis* genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. J Ind Microbiol Biotechnol. 2014; 41: 233-250.
7. Omura S, Ikeda H, Ishikawa J, Hanamoto A, Takahashi C, Shinose M, et al. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites. Proc Natl Acad Sci USA. 2001; 98: 12215-12220.
8. Ikeda H, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T, et al. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. Nat Biotechnol. 2003; 21: 526-531.
9. Evans M, Kaczmarek FS, Stutzman-Engwall K, Dyson P. Characterization of a *Streptomyces-lividans*-type site-specific DNA modification system in the avermectin-producer *Streptomyces avermitilis* permits investigation of two novel giant linear plasmids, pSA1 and pSA2. Microbiology. 1994; 140: 1367-1371.
10. Komaki H, Ichikawa N, Sekine M, Kitahashi Y, Fujita N, Suzuki K. *Streptomyces avermitilis* NBRC 14893^{T} has complestatin and oxazolomycin analog biosynthetic gene clusters, which do not exist in the reported genome sequence of *S. avermitilis* MA-4680^{T}. Microbiol Cult Coll. 2014; 30: 1-11.
11. Suroto DA, Kitani S, Arai M, Ikeda H, Nihira T. Characterization of the biosynthetic gene cluster for cryptic phthoxazolin A in *Streptomyces avermitilis*. PLoS One. 2018; 13: e0190973
12. Saito, H, Miura, K. Preparation of transforming deoxyribonucleic acid by phenol treatment. Biochim Biophys Acta 1963; 72: 619-629.
13. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, et al. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res. 2017; 45: W36-W41.
14. Hayakawa M, Nonomura H. Humic acid-vitamin agar, a new medium for the selective isolation of soil actinomycetes. J Ferment Technol. 1987; 65: 501-509.
15. Suroto DA, Kitani S, Miyamoto KT, Sakihama Y, Arai M, Ikeda H, Nihira T. Activation of cryptic phthoxazolin A production in *Streptomyces avermitilis* by the disruption of autoregulator-receptor homologue AvaR3. J Biosci Bioeng. 2017; 124: 611-617.

16. Chen CW, Huang CH, Lee HH, Tsai HH, Kirby R. Once the circle has been broken: Dynamics and evolution of *Streptomyces* chromosomes. Trends Genet. 2002; 18: 522-529.
17. Choulet F, Aigle B, Gallois A, Mangenot S, Gerbaud C, Truong C, et al. Evolution of the terminal regions of the *Streptomyces* linear chromosome. Mol Biol Evol. 2006; 23: 2361-2369.
18. Fischer G, Decaris B, Leblond P. Occurrence of deletions, associated with genetic instability in *Streptomyces ambofaciens*, is independent of the linearity of the chromosomal DNA. J Bacteriol. 1997; 179: 4553-4558.
19. Komatsu M, Komatsu K, Koiwai H, Yamada Y, Kozone I, Izumikawa M, et al. Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. ACS Synth Biol. 2013; 2: 384-96.
20. Chen W, He F, Zhang X, Chen Z, Wen Y, Li J. Chromosomal instability in *Streptomyces avermitilis*: Major deletion in the central region and stable circularized chromosome. BMC Microbiol. 2010; 10: 198.
21. Yang MC, Losik R. Cytological evidence for association of the ends of the linear chromosome in *Streptomyces coelicolor*. J Bacteriol. 2001; 183: 5180-5186.