Physical map of the *Bacillus subtilis* 166 genome: evidence for the inversion of an approximately 1900 kb continuous DNA segment, the translocation of an approximately 100 kb segment and the duplication of a 5 kb segment

Mitsuhiro Itaya

Tel: +81 427 24 6254. Fax: +81 427 24 6316.

Mitsubishi Kasei Institute
of Life Sciences, 11
Minamiooya, Machida-shi,
Tokyo 194, Japan

An I-CeuI–NotI–SfiI endonuclease map of the *Bacillus subtilis* 166 genome was constructed. It was almost identical to that of *B. subtilis* 168 except for the inversion of an approximately 1900 kb DNA segment, the translocation of an approximately 100 kb segment and the duplication of a 5 kb segment. Continuity of the inverted segment was investigated by direct measurement of the distances between the two genomic loci where I-SceI recognition sites were created in the 168 and the 166 genomes. Size difference of the I-SceI fragments between the two strains fully demonstrated the inversion of an approximately 1900 kb long ‘continuous’ DNA segment and the ‘location’ of the two inversion junctions in the genome. The 100 kb DNA segment including the lysogenic SP β prophage was translocated close to one of the inversion junctions and was probably associated with the duplication of a 5 kb segment. These rearrangements are consistent with those indicated by genetic analyses.

Keywords: *Bacillus subtilis*, I-SceI, physical map, inversion, translocation

INTRODUCTION

The project of sequencing the *Bacillus subtilis* 168 genome has been completed (Ogasawara & Yoshikawa, 1996; Moszer *et al.*, 1996). Its restriction enzyme map will be revised and become a standard genome map for this species. As regards structural variants with large-sized DNA rearrangements, five *B. subtilis* Marburg derivative mutants have been reported in which orientation of given genomic DNA segments is reversed.

The genome structure of three mutants derived from strain 168 have been physically analysed: strain 60866, with an inversion of a 1700 kb DNA segment (Toda & Itaya, 1995); strain BEST2145, with an inversion of a 1631 kb segment (Itaya, 1994); and strain BEST5227, with an inversion of a 1652 kb segment (Toda *et al.*, 1996). The inversions in BEST2145 and BEST5227 were induced by homologous recombination between the two

repeat sequences in the genome. The inversion in strain 60866 was caused by gamma-ray irradiation, although a detailed scheme for the inversion remains to be demonstrated (Toda & Itaya, 1995). In contrast, inversions in *B. subtilis* 166 (Trowsdale & Anagnostopoulos, 1976) and a *trpE30* strain (Schneider & Anagnostopoulos, 1981) were suggested by genetic evidence.

Both strains 168 and 166 were isolated from a single ancestral strain, *B. subtilis* Marburg, together with other auxotrophic mutants, after being exposed to X-ray irradiation (Burkholder & Giles, 1947). The genomic structure of strain 166 deserved to be determined by restriction-enzyme-based analyses like that of the 168 mutants.

This communication describes a *SfiI*, *NotI* and I-CeuI endonuclease map of the *B. subtilis* 166 genome. The map construction of the 166 genome was done by aligning the *SfiI*, *NotI* and I-CeuI fragments independently of the genetic data. This approach was successful in the construction of the detailed *NotI* and *SfiI* map of

Abbreviation: CHEF, contour-clamped homogeneous electric field.

Table 1. Bacterial strains and plasmids used in this study

<i>B. subtilis</i> strains	Genotype*	Construction	References
168 <i>trpC2</i> (= 1A1) GSY447	<i>trpC2</i> <i>trpE26 hisH2</i>		Itaya (1993a) Trowsdale & Anagnostopoulos (1976)
OA101	Prototroph from 168		Itaya & Tanaka (1991)
BEST2053	4::neo-I	pNEXT4F × 1A1	This study†
BEST2054	4::neo-I	pNEXT4F × GSY447	This study†
BEST2213	4::neo-I 34::spc-I	pNEXT34FA × BEST2053	This study‡
BEST2215	4::neo-I 5::spc-I	pNEXT5FA × BEST2053	
BEST2216	4::neo-I 38::spc-I	pNEXT38FA × BEST2053	
BEST2217	4::neo-I 33::spc-I	pNEXT33FA × BEST2053	
BEST2218	4::neo-I 41::spc-I	pNEXT41FA × BEST2053	
BEST2219	4::neo-I 24::spc-I	pNEXT24FA × BEST2053	
BEST2220	4::neo-I 46::spc-I	pNEXT46FB × BEST2053	
BEST2221	4::neo-I 55::spc-I	pNEXT55FA × BEST2053	This study‡
BEST2222	4::neo-I 34::spc-I	pNEXT34FA × BEST2054	
BEST2224	4::neo-I 5::spc-I	pNEXT5FA × BEST2054	
BEST2225	4::neo-I 38::spc-I	pNEXT38FA × BEST2054	
BEST2226	4::neo-I 33::spc-I	pNEXT33FA × BEST2054	
BEST2227	4::neo-I 41::spc-I	pNEXT41FA × BEST2054	
BEST2228	4::neo-I 24::spc-I	pNEXT24FA × BEST2054	
BEST2229	4::neo-I 46::spc-I	pNEXT46FB × BEST2054	
BEST2230	4::neo-I 55::spc-I	pNEXT55FA × BEST2054	
BEST2067§	<i>terC</i> ::neo-I	pBMAP106SCB × 1A1	This study†
BEST2068§	<i>terC</i> ::neo-I	pBMAP106SCB × GSY447	This study†
BEST2231	<i>terC</i> ::neo-I 46::spc-I	pNEXT46FB × BEST2069	This study‡
BEST2233	<i>terC</i> ::neo-I 46::spc-I	pNEXT46FB × BEST2073	This study‡
BEST3001	12::neo	pSOFT12A × OA101	Itaya & Tanaka (1991)†
BEST2235	12::neo	pSOFT12A × GSY447	This study†
<i>E. coli</i> plasmids	Insert*		References
pNEXT4F	4::neo-I (<i>NotI</i> -linking clone at 3503 kb)		This study†
pNEXT33FA	33::spc-I (<i>NotI</i> -linking clone at 1255 kb)		Toda <i>et al.</i> (1996)
pNEXT34FA	34::spc-I (<i>NotI</i> -linking clone at 3791 kb)		
pNEXT34FA	34::spc-I (<i>NotI</i> -linking clone at 3791 kb)		This study‡
pNEXT5FA	5::spc-I (<i>NotI</i> -linking clone at 205 kb)		
pNEXT38FA	38::spc-I (<i>NotI</i> -linking clone at 763 kb)		
pNEXT41FA	41::spc-I (<i>NotI</i> -linking clone at 1551 kb)		
pNEXT24FA	24::spc-I (<i>NotI</i> -linking clone at 1986 kb)		
pNEXT46FB	46::spc-I (<i>NotI</i> -linking clone at 2477 kb)		
pNEXT55FA	55::spc-I (<i>NotI</i> -linking clone at 3119 kb)		
pBMAP106SCB	<i>terC</i> ::neo-I		Itaya (1993a)†
pSOFT12A	12::neo		Itaya & Tanaka (1991)
pBEST513	neo-I		Itaya (1992)†
pBEST518	spc-I		Toda <i>et al.</i> (1996)‡

* neo-I and spc-I indicate neomycin resistance or spectinomycin resistance genes, respectively, carrying an I-SceI site. For details see text.

† *B. subtilis* transformants were selected with neomycin at 10 µg ml⁻¹. *E. coli* transformants were selected with kanamycin at 25 µg ml⁻¹.

‡ Selected with spectinomycin at 50 µg ml⁻¹.

§ An additional I-SceI site was created at the *NotI* site of pNEXT37 (357 kb) by insertion of the tetracycline resistance cassette from pBEST307 (Itaya, 1992).

|| The *SfiI* site at 2321 kb is disrupted by a *neo* insertion.

the *B. subtilis* 168 genome that was later correlated with the genetic map (Itaya & Tanaka, 1991; Itaya, 1993b). As 168 and 166 are isogenic strains, it was of great help that materials used for 168 map construction could be reused. Comparison with the 168 map indicated that the strain bears several DNA rearrangements: an inversion of an approximately 1900 kb long continuous DNA segment, a translocation of an approximately 100 kb segment and a duplication of a 5 kb segment. The

'continuity' of the long inverted segment was first proven by a new method using I-SceI endonuclease. These rearrangements were in agreement with those suggested by Anagnostopoulos and coworkers (Anagnostopoulos, 1990). After submission of the paper, it was announced at the 9th International Conference on Bacilli, Lausanne, Switzerland, 1997, that the entire nucleotide sequence of the *B. subtilis* 168 genome had been completed.

METHODS

Bacterial strains and plasmids. *B. subtilis* strain 1A1 (168: *trpC2*) was obtained from the *Bacillus* Genetic Stock Center (Columbus, OH, USA). The *NotI*-*SfiI* physical map (Itaya & Tanaka, 1991; Itaya, 1993a) and the I-*CeuI* map (Toda & Itaya, 1995) of this strain were reported earlier. *B. subtilis* strain GSY447 (166: *trpE26*) was obtained from C. Anagnostopoulos. Transformation of competent *B. subtilis* was done as described previously (Itaya & Tanaka, 1991). Transformation of competent *Escherichia coli* JA221 (F^- *hsdR* *hsdM*⁺ *trp* *leu* *lacY* *recA1*) cells was done by the method of Mandel & Higa (1970). Bacteria were grown in Luria-Bertani broth at 37 °C (Miller, 1972). Plasmids in *E. coli* were prepared by the method of Birnboim & Doly (1979) and purified by ultracentrifugation in the presence of CsCl and ethidium bromide.

In vitro DNA manipulations. Type II restriction enzymes, alkaline phosphatases and T4 DNA ligase were obtained from Toyobo, except for *SfiI* (New England Biolabs) and *NotI* (Takara Shuzo). Endonuclease I-*SceI* was purchased from Boehringer Mannheim and I-*CeuI* from New England Biolabs. DNA manipulations *in vitro* were done according to standard procedures (Maniatis *et al.*, 1982) or the manufacturers' instructions unless specified otherwise.

Insertion of I-*SceI* recognition sites in the *B. subtilis* genome.

The pivotal I-*SceI* site was created at 3503 kb of *B. subtilis* strains BEST2053 or BEST2054 genomes as listed in Table 1. A 1.3 kb neomycin resistance gene cassette carrying an I-*SceI* site (neo-I) was prepared from pBEST513 by *NotI* digestion. pBEST513 was constructed by insertion of the I-*SceI* linker supplied by Boehringer Mannheim into the *Bam*HI site of pBEST502 (Itaya, 1992). The neo-I cassette was inserted into the *NotI* site of pNEXT4, resulting in pNEXT4F. Transformation by pNEXT4F of strain 1A1 (168) resulted in the neomycin-resistant transformant, BEST2053. Similarly, the neomycin-resistant strain BEST2054 was isolated from GSY447 (Table 1).

The second I-*SceI* site was created in the BEST2053 or BEST2054 genome by using a spectinomycin resistance gene cassette carrying a I-*SceI* site (spc-I). For example, the spc-I cassette prepared from pBEST518 by *NotI* digestion was inserted in the *NotI* site of pNEXT34, resulting in pNEXT34FA. BEST2213 or BEST2222 were isolated as spectinomycin-resistant transformants from BEST2053 or BEST2054 by using pNEXT34FA. The I-*SceI* site was created at 3791 kb, the *NotI* site of pNEXT34. Genome structures were analysed by I-*SceI* as shown in Fig. 1. Similarly, the spc-I inserted *NotI*-linking clones were constructed using the pNEXT plasmids described in our previous report (Itaya & Tanaka, 1991). The plasmids indicated in Table 1 were used to isolate strains having the second I-*SceI* site in various *NotI* sites as listed in Table 1.

Preparation of genomic DNA. Intact unsheared genomic DNA for contour-clamped homogeneous electric field (CHEF) gel electrophoresis was prepared in agarose plugs as described previously (Itaya & Tanaka, 1991). Pulse time and running time for CHEF gel electrophoresis (Chu *et al.*, 1986) are specified in the figures. Other running conditions were constant: 3 V cm⁻¹ and 14 °C. After electrophoresis, the DNA

fragments were stained in ethidium bromide solution (6 ng ml⁻¹) for 15 min and photographed. I-*SceI* digestion did not exceed 30 min to minimize the non-specific cleavage which occurs during prolonged digestion (unpublished observations). To obtain sharp bands of I-*CeuI* fragments, gel blocks after enzyme digestion were treated with proteinase K (0.1 mg ml⁻¹) for 1 h at 37 °C as suggested by the manufacturer.

Southern hybridization and DNA probes. DNA in agarose gels was transferred onto nylon membranes (Nytran 13N; Schleicher & Schuell) by capillary blotting for 15–17 h (Southern, 1975). A non-radioactive labelling nucleotide, digoxigenin-11-dUTP, was used for preparing DNA probes. The random primer labelling technique, pre-hybridization and hybridization procedures and colour development were according to the protocol of the DNA Labelling and Detection kit from Boehringer Mannheim.

RESULTS AND DISCUSSION

A long continuous DNA segment is inverted in the *B. subtilis* 166 genome

Although the existing genetic evidence suggested that about 36% or about 1800 kb of the *B. subtilis* 166 genome is inverted compared with that of 168 (Trowsdale & Anagnostopoulos, 1976), the continuity of the inverted segment has not been clarified. Because strain 166 is able to become competent it was possible to create I-*SceI* sites in the genome by transformation. Creation of two I-*SceI* sites in the genome enables direct measurement of the physical distance between the two loci (Itaya *et al.*, 1992). This measurement was done to verify the continuity of the inverted segment and also to locate the two inversion junctions in the physical map. As described in Methods, two I-*SceI* sites were created in both the 168 and the 166 genomes, one at a fixed *NotI* site at 3503 kb and the other at various *NotI* sites as indicated in Fig. 1.

For example, from strain BEST2213 with two I-*SceI* sites at 3503 kb (designated 0 in Fig. 1) and 3791 kb (designated 8 in Fig. 1) two I-*SceI* fragments are generated, one of 283 kb from the region 0–8, and the other of 3905 kb from the rest of the genome. *B. subtilis* 168-derived strains (listed in Table 1) were constructed that had the second I-*SceI* site at designated *NotI* sites (1–8) as shown in Fig. 1. As the position of the second I-*SceI* site was moved clockwise in the 168 genome (8 → 1 → 2 → 3 → 4 → 5 → 6 → 7), the small fragment increased in size whilst the large fragment decreased in size accordingly. The fragment sizes were calculated from the *SfiI*-*NotI* map of 168. The array of these I-*SceI* fragments resolved in the gel had the shape of an X, as seen in Fig. 1(b).

Strains having two I-*SceI* sites were similarly constructed in the 166 genetic background, as listed in Table 1. A similar increase and decrease of the I-*SceI* fragments was obtained as the second I-*SceI* site was moved clockwise (8 → 1 → 2), indicating that the 166 genomic region 0–2 is continuous as is that of the 168 genome. The X-shape

Fig. 1. Physical distances between the two I-SceI sites created in the 168 and 166 genomes. (a) The 4188 kb *SfiI* restriction map of the *B. subtilis* 168 genome. *SfiI* fragments AS–ZS are aligned in the inner circle (Itaya & Tanaka, 1991). The location of *oriC* and *terC* are indicated by an open circle and filled box, respectively (Itaya, 1993a). The ten *rrn* operons are indicated by single capital letters. One pivotal I-SceI site was created at position 0 (3503 kb; pNEXT4). The second I-SceI site was created at the indicated *NotI* sites at positions 1 (205 kb; pNEXT5), 2 (763 kb; pNEXT38), 3 (1255 kb; pNEXT33), 4 (1551 kb; pNEXT41), 5 (1986 kb; pNEXT24), 6 (2477 kb; pNEXT46), 7 (3119 kb; pNEXT55) and 8 (3791 kb; pNEXT34). Calculated I-SceI sizes for strain 168 are shown on the outer arcs with double arrowheads, with the size of the other fragment in parentheses. Inversion junctions are indicated by thick bars. (b) I-SceI digests from strains 168 and 166 with the two I-SceI sites resolved. The intensities of the larger bands are less than expected due to non-specific cleavage during I-SceI digestion. Size markers on the left are from I-CeuI digests of the 168 genome (Toda & Itaya, 1995).

was abruptly interrupted when the second I-SceI site was created at the *NotI* site at 1255 kb (designated 3 in Fig. 1) and again increase and decrease of the two I-SceI fragments followed as the second I-SceI site was moved clockwise (3 → 4 → 5 → 6 → 7). These observations indicated that one inversion junction is located between 763 kb (2) and 1255 kb (3) or in the *SfiI* fragments between CS and WS, and the other is located between 3119 kb (7) and 3503 kb (0) or in the *SfiI* fragments MS, HS or PS.

Estimation of the size of the *B. subtilis* 166 genome as 4186 kb

SfiI fragments from the 166 genome separated by CHEF

gel electrophoresis look similar to those of the 168 genome as shown in Fig. 2(a). Apparently, four fragments, FS (238 kb), GS (192 kb), HS (192 kb) and KS (160 kb) were lost, but four new fragments, GKS (250 kb), FS' (232 kb), HS' (202 kb) and CHS (102 kb) appeared. The CS fragment was named CS' regardless of the apparent lack of size alteration because the *rrn*-specific probe did not hybridize to fragment CS' as shown in Fig. 2(a). The gain and loss of the *SfiI* fragments of the 166 genome resulted in an estimation of 4186 kb for the total size, a net decrease of 2 kb compared to the 168 genome (Fig. 2a).

NotI fragments from strain 166 were nearly identical to those of strain 168 except for the loss of at least two *NotI*

(Fig. 1). Exchange of the flanking DNA segments at the junction should form two chimeric fragments comprised of CS and HS. The right junction region was investigated first. The *rrn*-specific probe which originally identified the *rrnD* in the CS fragment no longer hybridized to CS' but identified the CHS fragment (102 kb) (Fig. 2a), indicating that *rrnD* resides in CHS. This observation led to the conclusion that CHS was formed by a connection of the *rrnD* region of CS and part of HS. The *rrnD* located in the 13N fragment of the 168 genome and the 13N fragment of the 166 genome was unaltered (Fig. 2b). This indicated that the 13N fragment was left unchanged upon the formation of the inversion, that is, the inversion junction must reside outside 13N.

A *SfiI*–*NotI* map of the left inversion junction region

The left junction was more complicated. Probes specific to the CS (*cotA*, pNEXT38) hybridized to CS' as expected, indicating that most of the CS fragment was retrieved in CS'. However, the HS-specific probes hybridized differently. pNEXT56 and pNEXT25 hybridized to CS' as expected whilst the *degQ* probe hybridized to HS'. This observation indicated that the HS fragment was divided into three parts, CHS, CS' and HS'. There must be a hypothetical *SfiI* site that separated the HS fragment into CS' and HS'. It was puzzling how the *SfiI* site had been acquired, but it was demonstrated that translocation had taken place, as described below.

Translocation of an approximately 100 kb segment

It was found that CS' hybridized to the *cotD* probe that was used to search for the lost GS and 1N fragments (Figs 2a, b). This observation revealed that the CS' fragment contained part of the GS fragment. Furthermore, the pSOFT12 probe, a *SfiI*-linking clone that hybridizes to both GS and KS, hybridized to CS' and HS' and the shortened 1N fragment (145 kb) (data not shown). Thus, the *SfiI* site of pSOFT12 must be the hypothetical one that links the CS' and the HS' as mentioned above. Proximity of HS' and CS' was directly proved by elimination of the *SfiI* site from the 166 genome. pSOFT12A was used to insert a neomycin resistance gene at the *SfiI* site. The neomycin-resistant transformant from GSY447 (BEST2235) generated a new approximately 550 kb *SfiI* fragment, with concomitant loss of HS' (350 kb) and CS' (202 kb) in contrast to the creation of a new 352 kb *SfiI* fragment comprised of the fragments GS (192 kb) and KS (160 kb) in the 168 genome (BEST3001) (data not shown).

The loss of the GS, KS and 1N fragments of GSY447 was caused by removal of the segment including pSOFT12, leaving the GKS (250 kb) fragment in the original *SPβ* locus. The insertion site of the removed segment was located in the HS fragment, between the *NotI* site of 12N (3367 kb) and the *NotI* site of pNEXT25 (3394 kb),

because the size of the 12N was unaltered and the pNEXT25 probe failed to hybridize to the original 27 kb *NotI* fragment (3367–3394 kb). Another as yet unidentified rearrangement might be present due to the translocation and the insertion.

The shortened *SPβ* area

A *NotI*–*SfiI* map of the *SPβ* region was constructed by using as probes pNEXT46, pNEXT36, pFAST8 and *terC*. To determine the size of the loss in this region, the genomic segment between *terC* and the *NotI* site of pNEXT46 was isolated by creation of *I-SceI* sites at these two loci. The *spc-I* cassette was inserted at the *NotI* site (2477 kb) and the neo-I cassette was inserted at *terC* (2012 kb). A 465 kb *I-SceI* fragment was obtained from the 168 genome, as predicted by the physical map (Fig. 3). In contrast, a 367 kb *I-SceI* fragment was obtained from the 166 genome (data not shown). The region of the 166 genome was shortened by 98 kb (465 kb minus 367 kb). The estimated 98 kb decrease in size of this region was inconsistent with other indirect estimations, for example, a 121 kb decrease of the 1N fragment (265 kb minus 144 kb; Fig. 2b) or a 102 kb decrease of the GS+KS to GKS fragments (352 kb minus 250 kb). This inconsistency remains to be investigated.

Duplication of a fragment in the *B. subtilis* 166 genome

An *Sse8387I*-linking clone, pFAST8, mapped at 2230 kb, hybridized to the GS and the 1N fragments (Itaya, 1995b). This probe fortuitously was found to hybridize to the 'two' *SfiI* fragments CS' and GKS, and the 'two' *NotI* fragments of 144 kb and 145 kb, of the 166 genome (Fig. 2a, b). As CS' and GKS are separated in the 166 genome, this result indicated that a segment of the 5 kb insert of pFAST8 is duplicated in the 166 genome. Probably, the duplication was formed in association with the translocation, as suggested by genetic analyses (Trowsdale & Anagnostopoulos, 1975, 1976).

I-CeuI map of the *B. subtilis* 166 genome

Endonuclease *I-CeuI* recognizes and cleaves a 26 bp sequence in all ten *rrn* operons of the 168 genome and the *rrn* map is useful to detect large-sized DNA rearrangements (Toda & Itaya, 1995). *I-CeuI* fragments were analysed to confirm the relocation of the *rrn* operons included in the *SfiI* map of the 166 genome. Four *I-CeuI* fragments, DB (2220 kb), BO (1037 kb), ED (306 kb) and JW (5.9 kb), were lost in the 166 genome as shown in Fig. 4. New fragments named DB', OD' and BE' were observed. The DB fragment was shortened by 124 kb in the *SPβ* region, resulting in DB'. The BO and ED fragments in which the inversion junctions are

Fig. 4. I-CeuI fragments from *B. subtilis* 168 and 166 genomes. Digests of about 3 µg DNA from *B. subtilis* 168 *trpC2* (168) and GSY447 (166) were separated by CHEF gel electrophoresis. Loss of the 5.9 kb JW fragment was attributed to the deletion between the operons *rrnJ* and *rrnW* (Toda & Itaya, 1995). Other changes are described in the text.

located gave rise to the BE' and OD' fragments upon inversion and translocation, although their sizes were not estimated.

The loss of the JW fragment was caused by the deletion of the 5.9 kb segment between *rrnJ* and *rrnW* and is consistent with the results in the *SfiI* (Fig. 2a) and *NotI* (Fig. 2b) analyses.

Consistency with the genetic and nucleotide sequence data

Anagnostopoulos and coworkers proposed that strain 166 has two distinctive structural traits: a DNA segment covering the genetic loci from *trpD* to *ilvA* is translocated to a locus between *thr* and *cysB*, and a DNA segment of one-third of the chromosome between *cysB* and *glyB* loci is inverted (Trowsdale & Anagnostopoulos, 1976). The I-CeuI-NotI-SfiI map of the 166 genome (Fig. 3) appears consistent with that indicated by genetic analyses. It remains uncertain if the deletion between *rrnJ* and *rrnW* occurred in association with the other DNA rearrangements since this type of deletion also occurs spontaneously during genetic translocation (unpublished observations).

Jarvis *et al.* (1990) reported nucleotide sequences including the two junctions of the inversion. Their sequence data revealed that one of the inversion junctions resides about 5 kb upstream of the *rrnD* operon, proximal to *oriC*. The right junction in the 166

map was located close to the *NotI* site at 947 kb. *rrnD* was mapped at 952 kb of the 168 physical map (Toda & Itaya, 1995). Thus, the order in the 166 genome is junction-*NotI* (947 kb)-*rrnD*. Although the two independent data sets seem remarkably similar, the *NotI* site at 947 kb predicted by the map was not found in the sequence data (Jarvis *et al.*, 1990). The loss of the *NotI* site may be attributed to strain differences or a loss during the cloning process.

Little sequence similarity was observed at the two junctions (Jarvis *et al.*, 1990). This implied that the inversion was unlikely to have been caused by homologous recombination between two hypothetical regions of the genome. Along with the simultaneous duplication and the translocation of the 98 kb genomic segment, rearrangements in the 166 genome were likely to have been formed during the repair process after X-ray irradiation. An inversion in a mutant, 60866, isolated after gamma-ray irradiation also accompanied a deletion (Toda & Itaya, 1995).

Occurrence of large DNA inversions has been reported in many other bacteria, for example *E. coli* (Rebollo *et al.*, 1988), *Salmonella typhimurium* (Segall *et al.*, 1988), *Pseudomonas aeruginosa* (Römling *et al.*, 1997) and *Streptomyces* species (Dary *et al.*, 1993). Systematic investigation of inversion formation was done for the first two species and plasticity of these bacterial genomes was suggested.

As regards *B. subtilis*, four examples have now been established by restriction-enzyme-based physical analy-

Fig. 5. *B. subtilis* inversion mutants. The junctions of the inverted segments are indicated by arrows and the lengths of the inversions are given in parentheses. BEST2145-type (Itaya, 1994), BEST5227-type (Toda *et al.*, 1996), C/S-type (Toda & Itaya, 1995) and *trpE26*-type (this study) mutants are shown. The first two of these mutations are caused by homologous recombination.

sis as shown in Fig. 5. Many detrimental situations have to be tolerated for the inversion mutants to survive (Toda *et al.*, 1996). The mutants must survive despite the asymmetric lengths of DNA between *oriC* and *terC*, different localization of genes which might influence specific gene expression, three-dimensional hindrance of genome structure as suggested by the study of *E. coli* (Rebollo *et al.*, 1988) and *S. typhimurium* (Segall *et al.*, 1988), and gain or loss of genes at the junction points. No particular phenotypes have been ascribed to inversion mutations, although changes in global gene expression were suggested in *Streptomyces* species (Dary *et al.*, 1993). One of the *B. subtilis* inversion mutants in Fig. 5 (BEST5227) was constructed by a method to invert a particular region of the genome via homologous recombination between two inverted repeats created in the 168 genome (Toda *et al.*, 1996). A project is now in progress to invert other regions of the *B. subtilis* genome and investigate genetic and physical constraints upon DNA inversion. The results will not only be important in understanding the nature of the present bacterial genome structure but also useful in designing the *B. subtilis* 168 genome (Itaya 1995a, b).

Finally, the sets of *I-SceI*-carrying antibiotic resistance genes could be used to verify a large DNA rearrangement when the mutant can develop competence, and the *I-SceI* strains will also be useful for isolating altered regions for nucleotide sequencing.

ACKNOWLEDGEMENTS

I especially thank Dr C. Anagnostopoulos for *B. subtilis* strains, related documents, careful reading of the manuscript and useful discussions. I also thank Drs T. Tanaka and T. Toda for their help in constructing plasmids and discussion.

REFERENCES

- Anagnostopoulos, C. (1990). Genetic rearrangements in *Bacillus subtilis*. In *The Bacterial Chromosome*, pp. 361–371. Edited by K. Drlica & M. Riley. Washington, DC: American Society for Microbiology.
- Birnboim, H. C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* 7, 1513–1523.
- Burkholder, P. R. & Giles, N. H. (1947). Induced biochemical mutations in *Bacillus subtilis*. *Am J Bot* 34, 345–348.
- Chu, G., Vollrath, D. & Davis, R. W. (1986). Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* 234, 1582–1585.
- Dary, A., Kaiser, P., Bourget, N., Thompson, C. J., Simonet, J. M. & Decaris, B. (1993). Large genomic rearrangements of the unstable region in *Streptomyces ambofaciens* are associated with major changes in global gene expression. *Mol Microbiol* 10, 759–769.
- Itaya, M. (1992). Construction of a novel tetracycline resistance gene cassette useful as a marker on the *Bacillus subtilis* chromosome. *Biosci Biotech Biochem* 56, 685–686.
- Itaya, M. (1993a). Stability and asymmetric replication of the *Bacillus subtilis* 168 chromosome. *J Bacteriol* 175, 741–749.
- Itaya, M. (1993b). Physical map of the *Bacillus subtilis* 168 chromosome. In *Bacillus subtilis and Other Gram-positive Bacteria*, pp. 463–471. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.
- Itaya, M. (1994). First evidence for homologous recombination-mediated large DNA inversion on the *Bacillus subtilis* chromosome. *Biosci Biotech Biochem* 58, 1836–1841.
- Itaya, M. (1995a). Toward a bacterial genome technology: integration of the *Escherichia coli* prophage lambda genome into the *Bacillus subtilis* 168 chromosome. *Mol Gen Genet* 248, 9–16.
- Itaya, M. (1995b). An estimation of minimal genome size required for life. *FEBS Lett* 362, 257–260.
- Itaya, M. & Tanaka, T. (1991). Complete physical map of the *Bacillus subtilis* 168 chromosome constructed by a gene-directed mutagenesis method. *J Mol Biol* 220, 631–648.
- Itaya, M., Laffan, J. J. & Sueoka, N. (1992). Physical distance between the site of type II DNA binding to the membrane and *oriC* on the *Bacillus subtilis* chromosome. *J Bacteriol* 174, 5466–5470.
- Jarvis, E. D., Cheng, S. & Rudner, R. (1990). Genetic structure and DNA sequences at junctions involved in the rearrangements of *Bacillus subtilis* strains carrying the *trpE26* mutation. *Genetics* 126, 785–797.
- Mandel, M. & Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *J Mol Biol* 53, 159–162.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Moszer, I., Kunst, F. & Danchin, A. (1996). The European *Bacillus*

subtilis genome sequencing project: current status and accessibility of the data from a new World Wide Web site. *Microbiology* **142**, 2987–2991.

Ogasawara, N. & Yoshikawa, H. (1996). The systematic sequencing of the *Bacillus subtilis* genome in Japan. *Microbiology* **142**, 2993–2994.

Rebollo, J.-E., Franco, V. & Louarn, J. M. (1988). Detection and possible role of two large nondividable zones on the *Escherichia coli* chromosome. *Proc Natl Acad Sci USA* **85**, 9391–9395.

Römling, U., Schmidt, K. D. & Tümmler, B. (1997). Large chromosomal inversions occur in *Pseudomonas aeruginosa* clone C strains isolated from cystic fibrosis patients. *FEMS Microbiol* **150**, 149–156.

Schneider, A. M. & Anagnostopoulos, C. (1981). Linkage map and properties of a *Bacillus subtilis* strain carrying a non-tandem duplication of the *purB*–*trc* region of the chromosome. *J Gen Microbiol* **125**, 241–256.

Segall, A., Mahan, M. J. & Roth, J. R. (1988). Rearrangement of the bacterial chromosome: forbidden inversions. *Science* **241**, 1314–1318.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**, 503–517.

Toda, T. & Itaya, M. (1995). I-CeuI recognition sites in the *rrn* operons of the *Bacillus subtilis* 168 chromosome: inherent landmarks for genome analysis. *Microbiology* **141**, 1937–1945.

Toda, T., Tanaka, T. & Itaya, M. (1996). A method to invert DNA segments of the *Bacillus subtilis* 168 genome by recombination between two homologous sequences. *Biosci Biotech Biochem* **60**, 773–778.

Trowsdale, J. & Anagnostopoulos, C. (1975). Evidence for the translocation of a chromosome segment in *Bacillus subtilis* strains carrying the *trpE26* mutation. *J Bacteriol* **122**, 886–898.

Trowsdale, J. & Anagnostopoulos, C. (1976). Differences in the genetic structure of *Bacillus subtilis* strains carrying the *trpE26* mutation. *J Bacteriol* **126**, 609–618.

Received 24 June 1997; revised 21 August 1997; accepted 3 September 1997.