Extremophiles (2004) 8:351–356
DOI 10.1007/s00792-004-0394-3

ORIGINAL PAPER

**Hideto Takami · Shinro Nishi · Jei Lu
Shigeru Shimamura · Yoshihiro Takaki**

Genomic characterization of thermophilic *Geobacillus* species isolated from the deepest sea mud of the Mariana Trench

Received: 15 March 2004 / Accepted: 14 April 2004 / Published online: 27 May 2004

Abstract The thermophilic strains HTA426 and HTA462 isolated from the Mariana Trench were identified as *Geobacillus kaustophilus* and *G. stearothermophilus*, respectively, based on physiologic and phylogenetic analyses using 16S rDNA sequences and DNA–DNA relatedness. The genome size of HTA426 and HTA462 was estimated at 3.23–3.49 Mb and 3.7–4.49 Mb, respectively. The nucleotide sequences of three independent λ-phage inserts of *G. stearothermophilus* HTA462 have been determined. The organization of protein coding sequences (CDSs) in the two λ-phage inserts was found to differ from that in the contigs corresponding to each λ insert assembled by the shotgun clones of the *G. kaustophilus* HTA426 genome, although the CDS organization in another λ insert is identical to that in the HTA426 genome.

Keywords Deep-sea isolate · Genome analysis · *Geobacillus kaustophilus* · *Geobacillus stearothermophilus* · 16S rDNA sequence · Thermophile

Introduction

Aerobic, endospore-forming, Gram-positive *Bacillus* species have often been isolated from various terrestrial soils and deep-sea sediments (Sneath 1986; Bartholomew and Paik 1966). In 1996, strains showing similarity of more than 98% to the 16S rDNA sequences of mesophilic *Bacillus* species such as *Bacillus cereus* and *B. subtilis*, and thermophilic *Bacillus*-related species such as *Geobacillus kaustophilus*, *G. stearothermophilus*, and *G. thermocatenulatus* (Nazina et al. 2001) were isolated from the deep-sea sediment collected at a depth of 10,897 m in the Challenger Deep of the Mariana Trench (Takami et al. 1997).

The complete genome sequences of five mesophilic bacilli with different maximum temperatures for growth, *B. subtilis* (Kunst et al. 1997), *B. halodurans* (Takami et al. 2000), *Oceanobacillus iheyensis* (Takami et al. 2002), *B. anthracis* (Read et al. 2003), and *B. cereus* (Ivanova et al. 2003), have been determined, whereas the complete genome sequences of thermophilic *Bacillus*-related species have not been elucidated. Considering the diversity and commonality of *Bacillus*-related species, the complete genome sequences of thermophilic *Bacillus*-related species will provide much information for the further study of the thermostability of proteins at the molecular level.

Geobacillus species HTA426 and HTA462 are thermophilic isolates recovered from the deep-sea sediment of the Mariana Trench at 55°C. We focus on these two species as good candidates for comparative genomic analyses between mesophilic and thermophilic bacilli. With this research background, we initiated whole-genome analysis and routinized the systematic sequencing of the genome of *Geobacillus* sp. HTA426. In this study, we attempted to identify the thermophilic Mariana isolates, HTA426 and HTA462, both on the basis of conventional physiological and biochemical characteristics and through phylogenetic analysis based on 16S rDNA sequences and DNA–DNA hybridization patterns. We also attempted to estimate the genome size using pulsed-field gel electrophoresis (PFGE) and to characterize the genome by sequencing the λ-phage inserts of HTA462 and shotgun clones of the HTA426 genome.

Communicated by K. Horikoshi

H. Takami (✉) · S. Nishi · J. Lu · S. Shimamura · Y. Takaki
Microbial Genome Research Group,
Japan Agency of Marine-Earth Science and Technology,
2-15 Natsushima, Yokosuka 237-0061, Japan
E-mail: takamih@jamstec.go.jp
Tel.: +81-46-867-9643
Fax: +81-46-867-9645

Materials and methods

Bacterial strains, media, and cultivation

Two deep-sea bacteria, designated HTA426 and HTA462, isolated from a deep-sea mud sample collected

at a depth of 10,897 m in the Mariana Trench (11°21.111′N, 142°25.949′E) were used in this study (Takami et al. 1997). These strains were grown aerobically at various temperature conditions in Luria-Bertani (LB) medium (pH 7.0).

Isoprenoid quinones and fatty acid analysis

Isoprenoid quinones were extracted from freeze-dried cells with chloroform:methanol (2:1) and purified using thin-layer chromatography. The purified isoprenoid quinones were analyzed using reverse-phase high-performance liquid chromatography (RP-HPLC, Komagata and Suzuki 1987), and the absorbance was measured at 270 nm, using menaquinone as a standard. Fatty acids were analyzed as methyl ester derivatives prepared from 10 mg freeze-dried cell materials by the method previously described (Kämpfer 1994).

G + C content and DNA–DNA hybridization

The G + C content was determined by RP-HPLC (Tamaoka and Komagata 1984). For analysis of relatedness, DNA–DNA hybridization was carried out at 40°C for 3 h and measured fluorometrically, using a previously described method (Ezaki et al. 1989).

Phylogenetic analysis based on 16S rDNA sequence

16S rDNA sequences were aligned using the Clustal multiple-alignment program (Clustal W, Thompson et al. 1994). A phylogenetic tree was inferred employing the neighbor-joining method (Saitou and Nei 1987), using the DNADIST and NEIGHBOR programs in the PHYLIP package, version 3.57 (University of Washington, Wash., USA).

PFGE

Chromosomal DNA for PFGE was prepared in agarose plugs by the method previously described (Takami et al. 1999a). DNA was digested with 100–200 U of *Apa*I or *Sse*8387I (Takara Shuzo, Ohtsu, Japan) at 37°C overnight in 500 µl of the restriction buffer recommended by the manufacturer. PFGE in 1% pulsed-field-certified agarose was performed by the method previously described (Takami et al. 1999a; Nakasone et al. 2000).

Preparation of λ-phage and shotgun libraries

The λ-phage library from HTA462 chromosome was prepared by the method described previously (Takami et al. 1999b). Three λ clones (GSA, GSB, and GSC) were randomly selected for sequencing analysis among the original recombinant phages prepared using the in vitro packaging system (Stratagene, La Jolla, Calif., USA). The DNA fragments cloned in λ phages were amplified by PCR using the LA PCR Kit, version 2 (Takara Shuzo), for preparation of the shotgun library. The shotgun libraries from the λ clones, and the HTA426 chromosome were prepared by the methods described previously (Takami et al. 1999b, 2000).

DNA sequencing

Three λ-phage clones prepared from the HTA462 genome and the whole genome of HTA426 were primarily sequenced using the whole-genome random sequencing method as described previously (Fleishmann et al. 1995). Sequencing was performed as described previously (Takami et al. 2002). The DNA fragments sequenced were assembled into contigs as described previously (Takami et al. 2000). The DDBJ/EMBL/GenBank accession numbers for the nucleotide sequences of the three λ inserts of the HTA462 genome and the contigs assembled by the shotgun closes of the HTA426 genome are as follows: AB126615 for λ GSA, AB126616 for λ GSB, AB126617 for λ GSC, AB126618 for contig GKA, AB126619 for contig GKB, and AB126620 for contig GKC.

Gene prediction and annotation

The predicted protein-coding regions were initially defined by searching for open reading frames longer than 50 codons using the Genome Gambler system (Sakiyama et al. 2000). Searches of the protein databases for amino acid similarities were performed using the same method as in the previous study (Takami et al. 2000, 2002). Annotation was performed as described in the previous study (Takami et al. 2000, 2002).

Results and discussion

Identification of strains HTA426 and HTA462

The phylogenetic tree based on 16S rDNA sequences of the two Mariana isolates, HTA426 and HTA462, and 15 other related thermophilic species was constructed (Fig. 1). HTA426, showing 97.5% identity with the 16S rDNA sequence of *Geobacillus kaustophilus* DSM7263^T, formed a large cluster with four other species, *Bacillus caldolyticus*, *B. caldovelox*, *B. caldotenax*, and *G. thermoleovorans*, in the phylogenetic tree. On the other hand, HTA462 formed a branch with *G. stearothermophilus* ATCC12980^T, whose 16S rDNA sequence is shared by HTA462 with 98.8% identity. DNA–DNA hybridization analysis was carried out to compare HTA426, HTA462, and four other related strains (Table 1). HTA426 showed the highest DNA–DNA

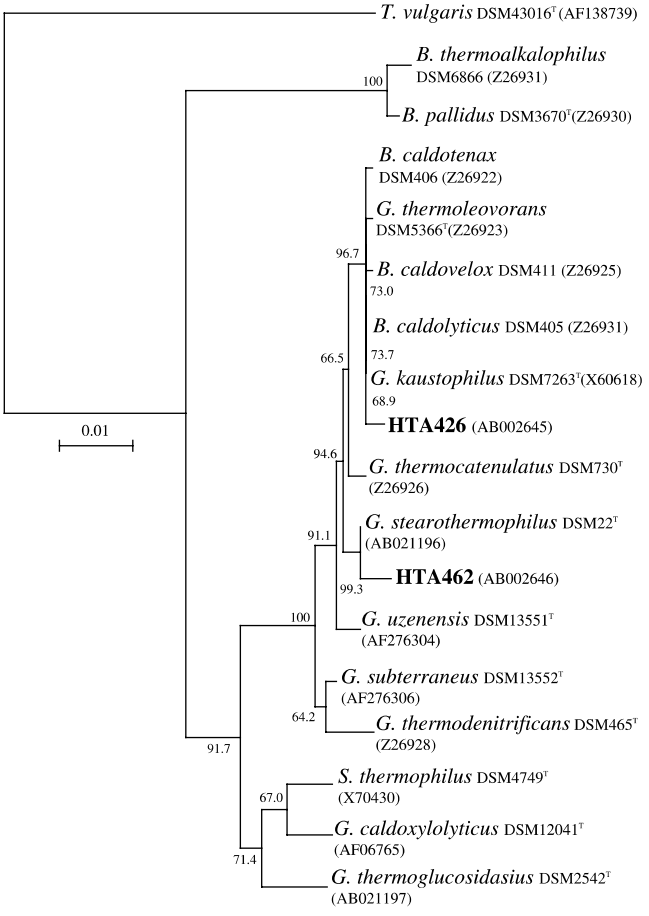

Fig. 1 Unrooted phylogenetic tree based on 16S rDNA sequence comparison showing the relationship of strains HTA426, HTA462, and other related strains. The *numbers next to nodes* indicate the percentages of bootstrap samples, derived from 1,000 samples, which supported the internal branches. Bootstrap probability values of less than 50% were omitted from this figure. The sequence of *Thermoactinomyces vulgaris* DSM43016^T has been included to serve as an outgroup. The accession number for each sequence is shown in *parentheses*. The *bar* indicates 0.01 Knuc unit. *S*, *Saccharococcus*

Table 1 DNA–DNA hybridization among strains HTA426 and HTA462 and other related strains

Labeled DNA	Unlabeled DNA	
	DNA–DNA hybridization (%)	
	Geobacillus sp. HTA426	*Geobacillus* sp. HTA462
Geobacillus sp. HTA426	100	60.9
Geobacillus sp. HTA462	61.5	100
G. stearothermophilus	65.4	79.6
G. thermocatenulatus	71.4	56.6
G. kaustophilus	82.7	50.4
G. uzenesis	59.3	62.3

relatedness of 82.7% to *G. kaustophilus* among the four reference strains, although the strain showed comparatively high relatedness to *G. stearothermophilus* (65.4%) and *G. thermocatenulatus* (71.3%). These findings appear to indicate that the organization of the genomes of the three *Geobacillus* species is similar to that of the HTA426 genome. On the other hand, HTA462 showed lower homologies of 50.4% with *G. kaustophilus* and of 56.6% with *G. thermocatenulautus* compared with the 79.6% homology with *G. stearothermophilus* (Table 1). Likewise, HTA462 showed a comparatively low homology of 60.9–61.5% with HTA426. Therefore, HTA426 and HTA462 are identified as members of *G. kaustophilus* and *G. stearothermophilus*, respectively, based on the phylogenetic analysis and DNA–DNA hybridization patterns.

Strains HTA426 and HTA462 are Gram-positive, strictly aerobic, and motile by means of peritrichous flagella. Cell growth of HTA426 and HTA462 was observed at temperatures of 42–74 and 37–72°C, respectively, and the optimum growth temperature for both strains was 60°C in LB medium (Table 2). The pH range for growth of HTA426 and HTA462 was 4.5–8.0 and 4.5–8.2, respectively. The growth patterns of HTA426 and HTA462 in terms of temperature and pH are very similar to those of the type strains *G. kaustophilus* (DSM7263^T) and *G. sterothermophilus* (ATCC12980^T) (Nazina et al. 2001). The major isoprenoid quinone in both HTA426 and HTA462 is menaquinone-7 as in related species showing thermophilic phenotypes (Sneath 1986). The G + C content of the DNA of HTA426 and HTA462 was 51.7 and 51.8 mol%, and these values were very similar to those of DSM7263^T and ATCC12980^T (Sneath 1986), respectively, as shown Table 2. The major cellular fatty acids of HTA426 and HTA462 are iso-15:0, iso-16:0, iso-17:0, and anteiso-17:0, and the major fatty acid pattern of HTA462 was the same as that of *G. stearthermophilus* ATCC12980^T. Through a series of analyses of the Mariana isolates, we concluded that HTA426 and HTA462 should be identified as *G. kaustophilus* HTA426 and *G. stearothermophilus* HTA462, respectively.

Estimation of genome size of strains HTA426 and HTA462

The genomes digested with I-*Ceu*I were used for estimation of the genome size and the copy number of the *rrn* operon because this enzyme is known to recognize a specific sequence of 26 bases within the *rrn* operons in bacterial genomes. As shown in Fig. 2, digestion of the HTA426 genome with I-*Ceu*I yielded nine fragments ranging in size from 19.4–1,621 kb, and the total size of the HTA426 genome was estimated to be 3,485 kb. Likewise, I-*Ceu*I generated nine fragments from the type strain of *G. kaustophilus* (DSM7263^T), and the genome size of DSM7263^T was estimated to be 3,681 kb. Accordingly, the genome size estimated by the digestion pattern of I-*Ceu*I was very similar between HTA426 and DSM7263^T. On the other hand, the genome of HTA462 was cut into 11 fragments (17.9–1,249 kb) by the digestion with I-*Ceu*I, differing from the ten fragments

Table 2 Characteristics of Mariana isolates and some related thermophilic *Geobacillus* species. *ND* No data, *W* weak growth

Characteristics	*G. kaustophilus* DSM7263^T	HTA426[a]	*G. stearothermophilus* ATCC12980^T	HTA462[a]
Genome size (Mb)	3.27–3.68[a]	3.23–3.49	2.06–2.72[a]	3.7–4.49
Morphology	Rod[a]	Rod	Rod[a]	Rod
Spore	+[c]	+	+[e]	+
Anaerobic growth	–[c]	–	W[d]	–
pH range for growth	6.2–7.5[b]	4.5–8.0	6.0–8.0[d]	4.5–8.2
Growth at 3% NaCl	ND	+	+[b]	+
Temperature range (°C)	40–75[b]	42–74	37–70[a]	37–72
Major isoprenoid quinone	MK-7[b,f]	MK-7	MK-7[b]	MK-7
Major cellular fatty acid	ND	Iso-15:0	Iso-15:0[d]	Iso-15:0
		Iso-16:0	Iso-16:0	Iso-16:0
		Iso-17:0	Iso-17:0	Iso-17:0
		Anteiso-17:0	Anteiso-17:0	Anteiso-17:0
G + C content (mol%)	52–58[b]	51.7	51.9[b]	51.8

[a]This study
[b]Nazina et al. (2001)
[c]Sunna et al. (1997)
[d]Manachini et al. (2000)
[e]Sneath (1986)
[f]*MK-7* Menaquinone-7

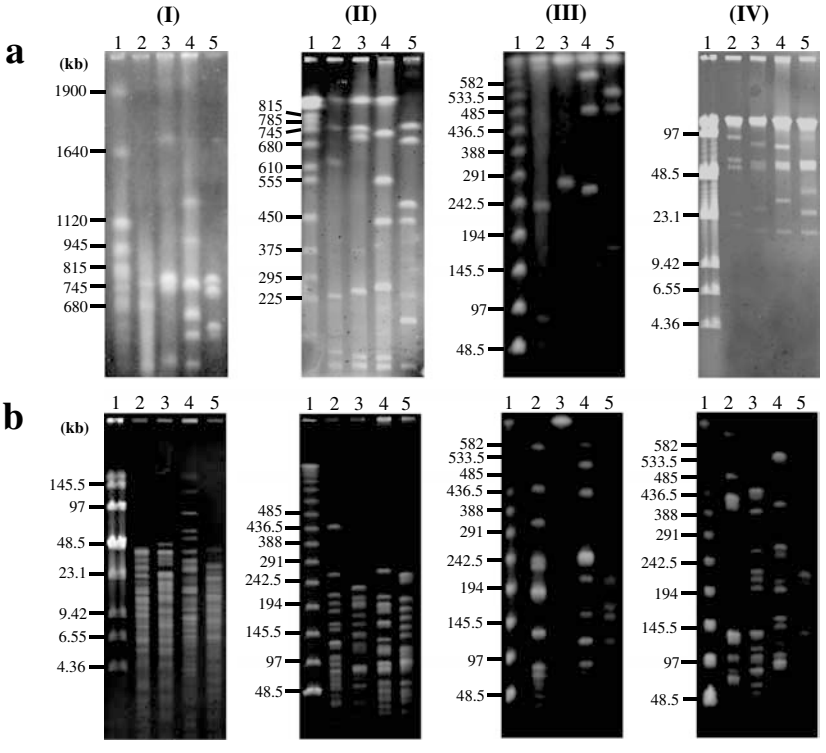

Fig. 2a, b Pulsed-field gel electrophoresis patterns of the chromosomal DNA of HTA426 and HTA462 and comparison with those of the related strains. **a** Digestion patterns obtained with I-*Ceu*I.*I* Separation of fragments ranging in size from 700–1,900 kb, *II* separation of fragments ranging in size from 200–800 kb, *III*separation of fragments ranging in size from 50–500 kb,*IV* separation of fragments ranging in size from 5–75 kb. **b** Digestion patterns obtained with *Sgr*AI, *Pme*I, *Sse*8387I, and *Pac*I. *ISgr*AI digestion,*II Pme*I digestion, *III Sse*8387I digestion, *IV Pac*I digestion. *Lane 1* Molecular size marker, *lane 2*HTA426, *lane 3 Geobacillus kaustophilus*DSM7263^T, *lane 4*HTA462, *lane 5 G. stearothermophilus* ATCC12980^T

(18.6–764.5 kb) in the case of the type strain of *G. stearothermophilus* (ATCC12980^T). Also, there is a significant difference in the genome size between HTA462 and ATCC12980^T. The genome size of the two strains was estimated to be 4,490 kb (HTA462) and 2,716 kb (ATCC12980^T), respectively, but these estimated sizes differed from that of another strain of the same species (strain 10) for which the genome size was estimated to be 3,400 kb (Lewis 2002).

Restriction endonucleases that recognize an 8-bp sequence were tested for their ability to digest the chromosomes of *G. kaustophilus* (HTA426 and DSM7263^T) and *G. stearothermophilus* (HTA462 and ATCC12980^T). *Sgr*AI generated at least 20 resolvable fragments, and the digestion patterns of HTA426 and DSM7263^T were similar between the two strains, although the digestion patterns with three other enzymes, *Pme*I, *Sse*8387I, and *Pac*I, differed from each other. It was particularly notable that the DSM7263 genome was not digested with *Sse*8387I, whereas this enzyme generated at least 11 resolvable fragments from the HTA426 genome (Fig. 2). Similarly, the digestion of the HTA462 and ATCC12980^T genomes with *Pme*I yielded at least 15 fragments, and the digestion patterns of their genomes were comparatively similar to each other. However, three other enzymes showed different digestion patterns between the two strains. The genome size of both strains was underestimated based on the digestion pattern with *Sse*8387I and *Pac*I compared with the case of I-*Ceu*I digestion to be 3,270–3,677 kb (HTA462) and 1,949–2,060 kb (ATCC12980^T), respectively. Through a series of PFGE analyses, it became

clear that thermophilic *G. stearothermophilus* may have variable genome sizes, and the digestion patterns of the genomes with various restriction endonucleases are different depending on the strains, indicating their flexible genome organization or different DNA modification system.

Comparison of gene organization between HTA426 and HTA462 genomes

There were 14 protein coding sequences (CDSs) comprising more than 50 codons in the sequenced fragment λ GSA of the HTA462 genome, and 15 CDSs were identified in the region of the HTA426 genome (contig GKA) corresponding to the region of λ GSA (Fig. 3). Of 14 putative proteins deduced from the CDSs identified in the λ GSA fragment, 9 showed significant amino acid sequence similarities ranging from 82.4–100% to those identified in the contig GKA, but the CDSs were organized in a different order. Half of the CDSs identified in the λ GSA and the contig GKA were functionally unknown, although most of them were conserved in other species (see accession number).

Seventeen CDSs were identified in the λ-GSB fragment of HTA462, and they were compared with 16 CDSs identified in the contig GKB of HTA426 corresponding to this region. The amino acid sequences of the first seven putative proteins in the λ GSB showed no similarity to those identified in the contig GKB. On the other hand, the amino acid sequences of the latter ten putative proteins showed remarkable similarity of 97.2 to 100% to those identified in the contig GKB, and the organization of these ten CDSs was identical between the two fragments (Fig. 3).

There were 16 CDSs in the sequenced fragment of λ GSC of HTA462 genome, and the same number of CDSs was identified in the contig GKC. The amino acid sequences of all CDSs identified in the λ GSC showed significant similarity—ranging from 91.3–99.5%—to those identified in the contig GKC of HTA426. The organization of the genes in this region was completely identical between *G. kaustophilus* HTA426 and *G. stearothermophilus* HTA462.

As the first step in analysis of the genomes of thermophilic bacilli, we determined the nucleotide sequences of three independent λ inserts of the genome of *G. stearothermophilus* HTA462 and the shotgun clones of the genome of *G. kaustophilus* HTA426 corresponding to the three λ inserts and characterized them for comparative study. The CDSs identified in these λ inserts of the HTA462 genome were found to be very similar to those identified in the contigs of HTA426 but were organized in a different order depending on the region of the genome. The findings obtained in this study will be helpful for further genomic study of thermophilic bacilli and comparative study with the genomes of mesophilic bacilli.

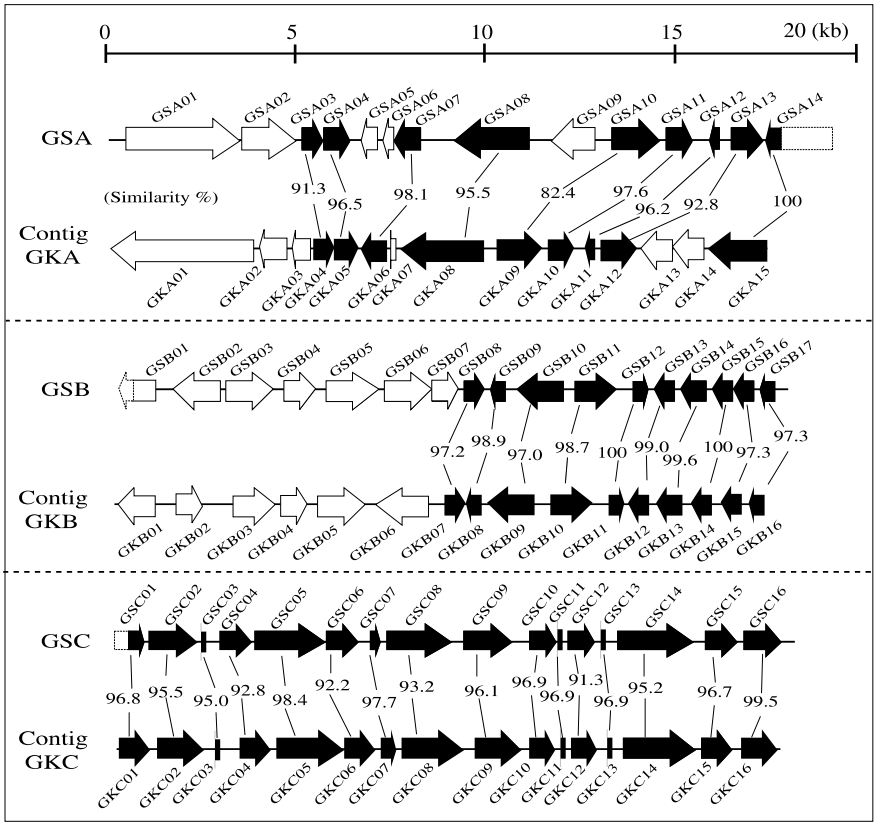

Fig. 3 Organization of the protein coding sequences (CDSs) in the λ-phage inserts (*GSA*, *GSB*, and *GSC*) of HTA462 and comparison with the corresponding region of HTA426 among the contigs (*GKA*,*GKB*, and *GKC*) constructed by assembling the shotgun clones. The *black arrows* show commonly conserved CDSs in DNA fragments between HTA462 and HTA426. The *white arrows* indicate CDSs that have no mutual relationship between HTA462 and HTA426 in each DNA fragment. The *dashed boxes* indicate the truncated CDSs in the cloned fragment in the λ phage. Each *number* shows the similarity value at the amino acid sequence level

Acknowledgements We thank H. Suzuki, S. Gohda and Y. Shen for their technical assistance.

References

Bartholomew JW, Paik G (1966) Isolation and identification of obligate thermophilic sporeforming bacilli from ocean basin cores. J Bacteriol 92:635–638

Ezaki T, Hashimoto Y, Yabuuchi E (1989) Fluorometric deoxyribonucleic acid-deoxyribonucleic acid hybridization in microdilution wells as an alternative to membrane filter hybridization in which radioisotopes are used to determine genetic relatedness among bacterial strains. Int J Syst Bacteriol 39:224–229

Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness, EF, Kerlavage AR, Bult CJ, Tomb J, Dougherty BA, Merrick JM, et al (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496–512

Ivanova N, Sorokin A, Anderson I, Galleron N, Candelon B, Kapatral V, Bhattacharyya A, Reznik G, Mikhallova N, Lapidus A, et al (2003) Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. Nature 423:87–91

Kämpfer P (1994) Limits and possibilities of total fatty acid analysis for classification and identification of *Bacillus* species. Syst Appl Microbiol 17:86–98

Komagata K, Suzuki K (1987) Lipid and cell wall analysis in bacterial systematics. Meth Microbiol 19:161–207

Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, Bertero MG, Bessieres P, Bolotin A, Borchert S, et al (1997) The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. Nature 390:249–256

Lewis SA (2002) The working draft sequence of the Gram-positive bacteria *Bacillus stearothermophilus* strain 10 genome, its analysis and annotation. PhD Thesis, University of Oklahoma, Norman, pp 62–66

Manachini PL, Mora D, Nicastro G, Parini C, Stackebrandt E, Pukall R, Fortina G (2000) *Bacillus thermodenitrificans* sp. nov., nom. Rev. Int J Syst Evol Microbiol 50:1331–1337

Nakasone K, Masui N, Takaki Y, Sasaki R, Maeno G, Sakiyama, T, Hirama C, Fuji F, Takami H (2000) Characterization and comparative study of the *rrn* operons of alkaliphilic *Bacillus halodurans* C-125. Extremophiles 4:209–214

Nazina TN, Tourova TP, Poltaraus AB, Novikova EV, Grigoryan, AA, Ivanova AE, Lysenko AM, Petrunyaka VV, Osipov GA, Belyaev SS, et al (2001) Taxonomic study of aerobic thermophilic bacilli: description of *Geobacillus subteaneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus, Bacillus thermocatenulatus, Bacillus thermoleovolans, Bacillus kaustophilus, Bacillus thermoglusidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus, G. thermocatenulatus, G. thermoleovolans, G. kaustophilus, G. thermoglusidasius* and *G. thermodenitrificans*. Int J Syst Evol Microbiol 51:433–446

Read TD, Peterson SN, Tourasse N, Bailli LW, Paulsen IT, Nelson KE, Tettelin H, Fouts DE, Eisen JA, Gill SR, et al (2003) The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. Nature 423:81–86

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sakiyama T, Takami H, Ogasawara N, Kuhara S, Kozuki T, Doga, K, Ohyama A, Horikoshi K (2000) An automated system for genome analysis to support microbial whole-genome shotgun sequencing. Biosci Biotech Biochem 64:670–673

Sneath PHA (1986) Endospore-forming Gram-positive rods and cocci. In: Sneath PHA, Mair NS, Sharp ME, Holt JG (eds) Bergey's manual of systematic bacteriology, vol 2. Williams and Wilkins, Baltimore, pp 1104–1139

Sunna A, Tokajian S, Burghardt J, Rainey F, Antranikian G, Hashwa F (1997) Identification of *Bacillus kaustophilus*, *Bacillus thermocatenulatus* and *Bacillus* strain HSR as members of *Bacillus thermoleovorans*. Syst Appl Microbiol 20:232–237

Takami H, Inoue A, Fuji F, Horikoshi K (1997) Microbial flora in the deepest sea mud of the Mariana Trench. FEMS Microbiol Lett 152:279–285

Takami H, Nakasone K, Ogasawara N, Hirama C, Nakamura Y, Masui N, Fuji F, Takaki Y, Inoue A, Korikoshi K (1999a) Sequencing of three λ clones from the genome of alkaliphilic *Bacillus* sp. strain C-125. Extremophiles 3:29–34

Takami H, Nakasone K, Hirama C, Takaki Y, Masui N, Fuji F, Nakamura Y, Inoue A (1999b) An improved physical and genetic map of the genome of alkaliphilic *Bacillus* sp. C-125. Extremophiles 3:21–28

Takami H, Nakasone K, Takaki Y, Maeno G, Sasaki R, Masui N, Fuji F, Hirama C, Nakamura Y, Ogasawara N, et al (2000) Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. Nucleic Acids Res 28:4317–4331

Takami H, Takaki Y, Uchiyama I (2002) Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments. Nucleic Acids Res 30:3927–3935

Tamaoka J, Komagata K (1984) Determination of DNA base composition by reverse-phase high-performance liquid chromatography. FEMS Microbiol Lett 25:125–128

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680