JOURNAL OF BACTERIOLOGY, May 2006, p. 3382–3390 Vol. 188, No. 9
0021-9193/06/$08.00+0 doi:10.1128/JB.188.9.3382–3390.2006

Pathogenomic Sequence Analysis of *Bacillus cereus* and *Bacillus thuringiensis* Isolates Closely Related to *Bacillus anthracis*†

Cliff S. Han,[1,3]‡ Gary Xie,[1,3]‡ Jean F. Challacombe,[1,3]‡* Michael R. Altherr,[3] Smriti S. Bhotika,[1,3]§ David Bruce,[1,3] Connie S. Campbell,[1,3] Mary L. Campbell,[1,3] Jin Chen,[1,3]‖ Olga Chertkov,[1,3] Cathy Cleland,[6] Mira Dimitrijevic,[1,3] Norman A. Doggett,[3] John J. Fawcett,[1,3] Tijana Glavina,[2,4] Lynne A. Goodwin,[1,3] Karen K. Hill,[3] Penny Hitchcock,[6]¶ Paul J. Jackson,[3,5] Paul Keim,[7] Avinash Ramesh Kewalramani,[1,3] Jon Longmire,[3] Susan Lucas,[2,5] Stephanie Malfatti,[2,5] Kim McMurry,[1,3] Linda J. Meincke,[1,3] Monica Misra,[1,3] Bernice L. Moseman,[1,3] Mark Mundt,[8] A. Christine Munk,[1,3] Richard T. Okinaka,[3] B. Parson-Quintana,[1,3] Lee Philip Reilly,[1,3] Paul Richardson,[2,4] Donna L. Robinson,[1,3] Eddy Rubin,[2,4] Elizabeth Saunders,[1,3] Roxanne Tapia,[1,3] Judith G. Tesmer,[1,3] Nina Thayer,[1,3] Linda S. Thompson,[1,3] Hope Tice,[2,4] Lawrence O. Ticknor,[6] Patti L. Wills,[1,3] Thomas S. Brettin,[1,3] and Paul Gilna[1,3]

DOE Joint Genome Institute Los Alamos National Laboratory, Los Alamos, New Mexico 87545[1]; DOE Joint Genome Institute Production Genome Facility, Walnut Creek, California 94598[2]; Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico 87545[3]; Lawrence Berkeley National Laboratory, Berkeley, California 94720[4]; Lawrence Livermore National Laboratory, Livermore, California 94550[5]; Los Alamos National Laboratory Decision Applications Division, Los Alamos, New Mexico 87545[6]; Northern Arizona University Department of Biological Sciences, Flagstaff, Arizona 86011-5640[7]; and Los Alamos National Laboratory, ESA Division, Los Alamos, New Mexico 87545[8]

Received 22 November 2005/Accepted 13 February 2006

Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* are closely related gram-positive, spore-forming bacteria of the *B. cereus* sensu lato group. While independently derived strains of *B. anthracis* reveal conspicuous sequence homogeneity, environmental isolates of *B. cereus* and *B. thuringiensis* exhibit extensive genetic diversity. Here we report the sequencing and comparative analysis of the genomes of two members of the *B. cereus* group, *B. thuringiensis* 97-27 subsp. *konkukian* serotype H34, isolated from a necrotic human wound, and *B. cereus* E33L, which was isolated from a swab of a zebra carcass in Namibia. These two strains, when analyzed by amplified fragment length polymorphism within a collection of over 300 of *B. cereus*, *B. thuringiensis*, and *B. anthracis* isolates, appear closely related to *B. anthracis*. The *B. cereus* E33L isolate appears to be the nearest relative to *B. anthracis* identified thus far. Whole-genome sequencing of *B. thuringiensis* 97-27and *B. cereus* E33L was undertaken to identify shared and unique genes among these isolates in comparison to the genomes of pathogenic strains *B. anthracis* Ames and *B. cereus* G9241 and nonpathogenic strains *B. cereus* ATCC 10987 and *B. cereus* ATCC 14579. Comparison of these genomes revealed differences in terms of virulence, metabolic competence, structural components, and regulatory mechanisms.

While *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* are closely related members of the *B. cereus* group (22), individual isolates exhibit differences in terms of host range and virulence. *B. anthracis* is the causal agent of anthrax, a zoonotic disease that can be lethal to humans. *B. cereus* is a ubiquitous soil organism and an opportunistic human pathogen most commonly associated with food poisoning (10). *B. thuringiensis* is an insect pathogen that is widely used as a biopesticide (36). Here we report the sequencing and comparative analysis of the genomes of two members of the *B. cereus* group, *B. thuringiensis* 97-27 subsp. *konkukian* serotype H34, isolated from a necrotic human wound (17), and *B. cereus* E33L, which was isolated from a swab of a zebra carcass in Namibia (P. C. B. Turnbull, personal communication). To facilitate the comparison of these two isolates with other members of the *B. cereus* group, we compiled a core genome of over 3,900 *B. cereus* group genes. Comparison of these genomes revealed differences in terms of virulence, metabolic competence, structural components, and regulatory mechanisms (see Table 1, below), supporting the idea that differential regulation modulates virulence rather than simple acquisition of virulence factor genes. Our analysis of the genome sequences of *B. thuringiensis* 97-27 and *B. cereus* E33L provides insight into the evolutionary relationships among these *B. cereus* group organisms, as well as the molecular mechanisms contributing to their host range and virulence.

* Corresponding author. Mailing address: Department of Energy Joint Genome Institute, Bioscience Division, MS M888, Los Alamos National Laboratory, Los Alamos, NM 87545. Phone: (505) 665-1485. Fax: (505) 665-3024. E-mail: jchalla@lanl.gov.

† Supplemental material for this article may be found at http://jb.asm.org/.

‡ C.S.H., G.X., and J.F.C. contributed equally to this study.

§ Present address: University of Florida, Gainesville, FL 32611.

¶ Present address: The Center for Biosecurity of UPMC, Baltimore, MD 21202.

‖ Present address: National Cancer Institute, Rockville, Md.

MATERIALS AND METHODS

Sequencing of the *B. thuringiensis* 97-27 and *B. cereus* E33L genomes. The random shotgun method of cloning, sequencing, and assembly was used. Large

(40-kb; *B. thuringiensis* 97-27 only), median (8-kb), and small (2.5- to 3.5-kb) insert libraries were sequenced for these genome projects with an average success rate of 90% and average high-quality read lengths of 643 and 621 nucleotides for *B. thuringiensis* 97-27 and *B. cereus* E33L, respectively. The completed genome sequences of *B. thuringiensis* 97-27 and *B. cereus* E33L contained 134,054 and 141,352 reads, respectively, achieving an average of 19.3- and 18.7-fold sequence coverage per base. After assembly, gaps between contigs were closed by editing, primer walking library clones, or PCR amplifications.

Annotation. Gene predictions were obtained using Glimmer (7, 35), and tRNAs were identified using tRNAScan-SE (25). Basic analysis of the gene predictions was performed by comparing coding sequences against the PFam, BLOCKS, and Prodom databases. Gene definitions and functional classes were added manually by a team of annotators using BLAST results in addition to information from the basic analysis.

Sequence analysis. We compared the genomes at the nucleotide level using genome alignment tools such as MUMmer2 (8), ACT (http://www.sanger.ac.uk/Software/ACT/), and Pipmaker (37). To obtain a list of orthologs in the *B. thuringiensis* 97-27 and *B. cereus* E33L genomes, we wrote a perl script that determines bidirectional best hits as follows. Genes *g* and *h* are considered orthologs if *h* is the best BLASTP hit for *g* and vice versa, with *e*-values less than or equal to 10^{-15}. A gene is considered strain specific if it has no hits with an *e*-value of 10^{-15} or less.

To identify IS elements in *B. thuringiensis* 97-27 and *B. cereus* E33L and compare them to IS elements present in other *B. cereus* group members, all known IS elements were used as query sequences and used with BLAST against the genomes of three strains of *B. anthracis* (Ames, A2012, and Sterne), *B. thuringiensis* 97-27, *B. cereus* E33L, and *B. cereus* (ATCC 14579).

Tandem repeats were identified in *B. thuringiensis* 97-27 and *B. cereus* E33L genomes using the Tandem Repeats Finder (4) with the threshold set for a minimum alignment score of 50.

AFLP. Amplified fragment length polymorphism (AFLP) analysis of the microbial DNAs was accomplished as previously described (18). Briefly, each of the DNA preparations was digested with EcoRI and MseI, and the resulting fragments were ligated to double-stranded adapters and then amplified by PCR using +0/+0 primers. Selective amplifications using the +1/+1 primer combination of 6-carboxyfluorescein-labeled EcoRI-C and MseI-G resulted in products that were mixed with a solution containing DNA size standards (Genescan-500 from Applied Biosystems Inc., Foster City, CA; and MapMarker-400 from BioVentures, Inc., Murfreesburo, TN), both labeled with *N*,*N*,*N*,*N*-tetramethyl-6-carboxyrhodamine. Following a 2-min heat denaturation at 90°C, the reaction mixtures were loaded onto a 5% Long Ranger DNA sequencing gel (Cambrex Bio Science, Rockland, ME) and visualized on an ABI 377 automated fluorescent sequencer (Applied Biosystems, Inc., Foster City, CA). Each set of AFLP experiments also included as a sample *B. anthracis* Vollum DNA, which was used as an internal control to allow comparison of results from different gels run at different times. GeneScan analysis software (Applied Biosystems, Inc., Foster City, CA) was used to determine the lengths of the sample fragments by comparison to the DNA fragment length size standards included within each sample.

Data analysis of the microbial DNAs was as previously described (39). DNA fragment sizes between 100 and 500 bp from triplicate data (derived from three lanes from three different gels) for each sample were combined. Fragment sizes that appeared in all three replicates were used to represent the sample, and the peak heights for the fragment sizes were averaged. This "averaged" sample was then used to compare to other "averaged" samples. A hierarchical agglomerative clustering routine using group averages was used to determine which fragments among the samples had similar lengths. A decision rule was added to this clustering routine that limited the allowable number of fragments within a cluster to equal the number of samples being compared and limited the maximum acceptable range of fragments sizes for a cluster to a preset value. Similarities between samples were measured using the Jaccard coefficient. Dendrograms were produced using the similarity matrix and the unweighted pair-group mean average method (F. J. Rohlf, NTSYS-PC numerical taxonomy and multivariate analysis system, version 1.8; Exeter Software, Setauket, N.Y.).

Nucleotide sequence accession numbers. The sequences of the *B. thuringiensis* 97-27 and *B. cereus* E33L genomes and plasmids can be accessed using the GenBank accession numbers AE017355, CP000001, CP000040, CP000041, CP000042, CP000043, CP000044, and CP000047.

RESULTS

General genome features. The 5.31-Mb genome of *B. thuringiensis* 97-27 comprises two replicons: a circular chromosome, encoding at least 5,198 open reading frames, and the pBT9727 plasmid (see Fig. S1A in the supplemental material). The 5.84-Mb genome of *B. cereus* E33L comprises six replicons: a circular chromosome, encoding at least 5,682 open reading frames, and five plasmids (see Fig. S1B in the supplemental material). *B. thuringiensis* 97-27 and *B. cereus* E33L have broad similarities to and share a high degree of synteny with *B. anthracis* Ames (33), *B. cereus* ATCC 14579 (21), and *B. cereus* ATCC 10987 (32). Within the *B. cereus* group, *B. anthracis*, *B. cereus* E33L, and *B. thuringiensis* 97-27 are part of a distinct cluster which contains many pathogenic organisms (18) (Fig. 1).

As illustrated in Fig. S2 in the supplemental material, a total of 3,917 putative proteins are shared among *B. anthracis* Ames, *B. cereus* ATCC 14579, *B. thuringiensis* 97-27, and *B. cereus* E33L using as a criterion whether genes were bidirectional best hits in BLAST searches. Comparison of the genomes of *B. cereus* E33L and *B. thuringiensis* 97-27 with *B. anthracis* Ames and *B. cereus* 14579 also identified strain-specific genes in each organism. Of the 5,682 predicted *B. cereus* E33L proteins, 253 of the chromosomally encoded genes and 416 of the plasmid genes are unique. Of the 5,197 predicted *B. thuringiensis* 97-27 proteins, 307 of the chromosomally encoded genes and 66 of the plasmid genes are unique. *B. cereus* E33L and *B. anthracis* Ames are the closest pair and share the highest number (221) of common proteins.

Virulence genes. The chromosomally encoded virulence genes in *B. thuringiensis* 97-27 and *B. cereus* E33L are common to the *B. cereus* group of bacteria (14). Neither *B. thuringiensis* 97-27 nor *B. cereus* E33L has the highly characterized *B. anthracis* toxin genes (*pag*, *lef*, and *cya*) encoded on pX01 or the *cap* genes encoded by pXO2 (26, 27). Nonetheless, our results indicate that both *B. thuringiensis* 97-27 and *B. cereus* E33L share a set of virulence factors common to the members of the *B. cereus* group. These common virulence genes include the three nonhemolytic enterotoxin genes (*nheABC*), two channel-forming type III hemolysins, a perfringolysin O (listeriolysin O), a phosphatidyl-inositol-specific and a phosphatidyl-choline-preferring phospholipase, RNA polymerase sigma-B factor, and a p60 family extracellular protease. These last five genes are homologous to virulence genes encoded by the gram-positive pathogen *Listeria monocytogenes* 12. *B. thuringiensis* 97-27 and *B. cereus* E33L also have a gene encoding cytotoxin K, which was previously identified in *B. cereus* (ATCC 14579). While *B. cereus* E33L lacks the *hbl* operon, which is suspected as a primary factor in diarrheal *B. cereus* food poisoning (3), *B. thuringiensis* 97-27 has the *hbl* operon containing the hemolytic enterotoxin genes *hblCDBA* also found in *B. cereus* (ATCC 14579) (Fig. 2A). Interestingly, the *hbl* gene cluster consists of the hemolytic enterotoxin (*hblCDBA*) and other genes encoding the spore germination proteins *gerIABC*, as well as other related proteins, that are ordered and oriented in a way that suggests their expression is coordinated by the transcriptional regulator TrrA. This gene cluster is part of a large, approximately 17.7-kb, 11-gene insertion (Fig. 2B). A degenerate IS*Rso11* transposase fragment is found at the presumed insertion boundary region, and direct repeats that overlap with the C-terminal UvrC-like protein were identified. This observation suggests a mechanism for the acquisition of these virulence

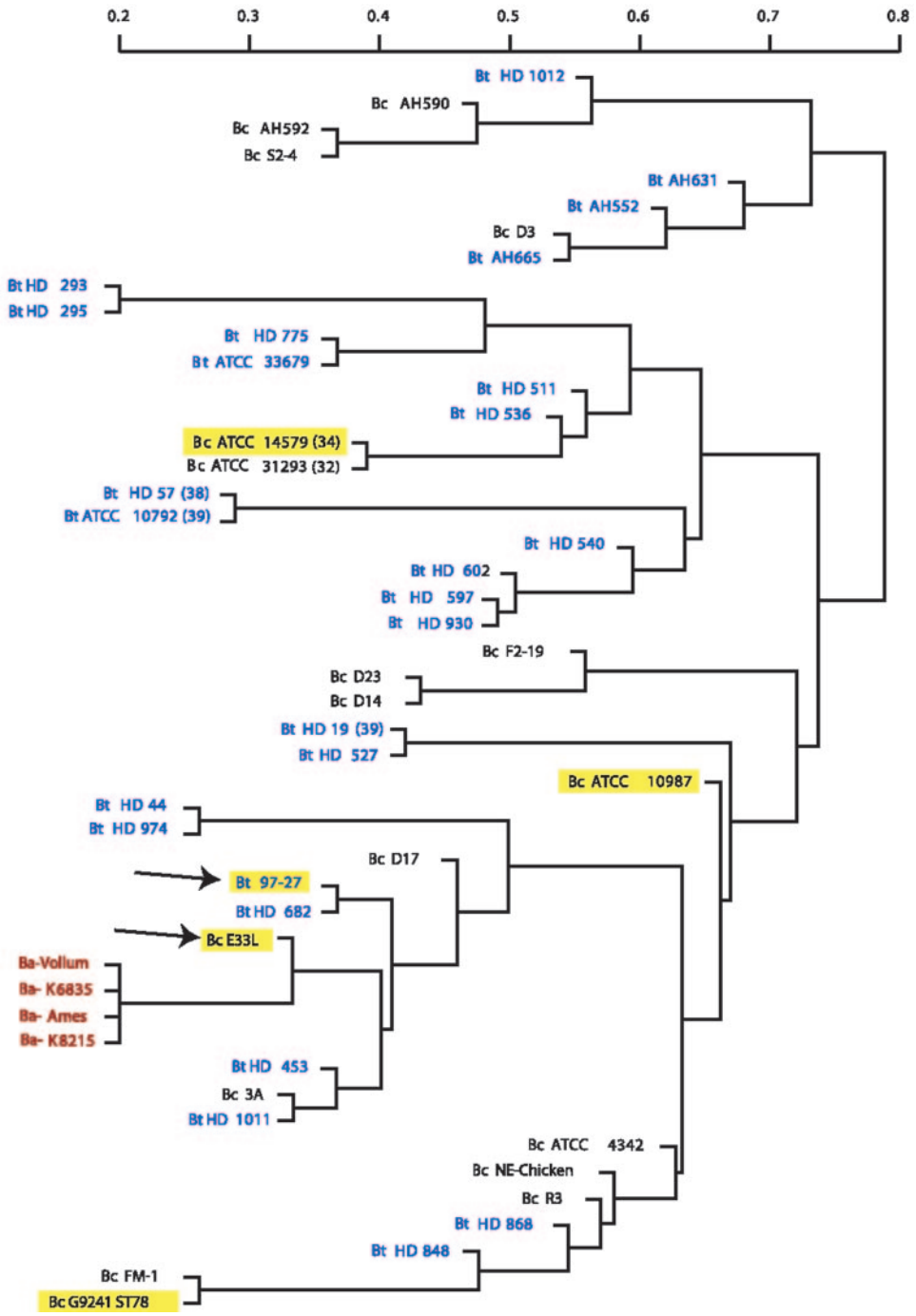

FIG. 1. An AFLP-based tree of *B. anthracis*, *B. cereus*, and *B. thuringiensis* isolates. These 48 isolates are representative of the branches identified when over 300 isolates of *B. anthracis*, *B. thuringiensis*, and *B. cereus* were examined by AFLP. Yellow highlighted isolates have fully sequenced genomes, blue indicates *B. thuringiensis* isolates, black shows *B. cereus* isolates, and red indicates *B. anthracis* isolates.

factors in *B. thuringiensis* 97-27, *B. cereus* 14579, and *B. cereus* G9241.

The opportunistic pathogenicity of *B. cereus* and *B. thuringiensis* may depend on the secretion of nonspecific extracellular virulence factors in response to transcriptional activation by PlcR (34). However, in all *B. anthracis* strains, the *plcR* gene is inactivated by a frameshift mutation which creates an early stop codon (1). In other *B. cereus* isolates, the *plcR* gene product up regulates the transcription of genes encoding enterotoxins, proteases, phospholipases, metabolic enzymes, proteins involved in motility and chemotaxis, proteins involved in sporulation, DNA metabolism, transcriptional regulators, and a variety of transporters by binding to a specific upstream motif (1, 21, 34). The genes encoding PlcR appear intact in *B. thu-*

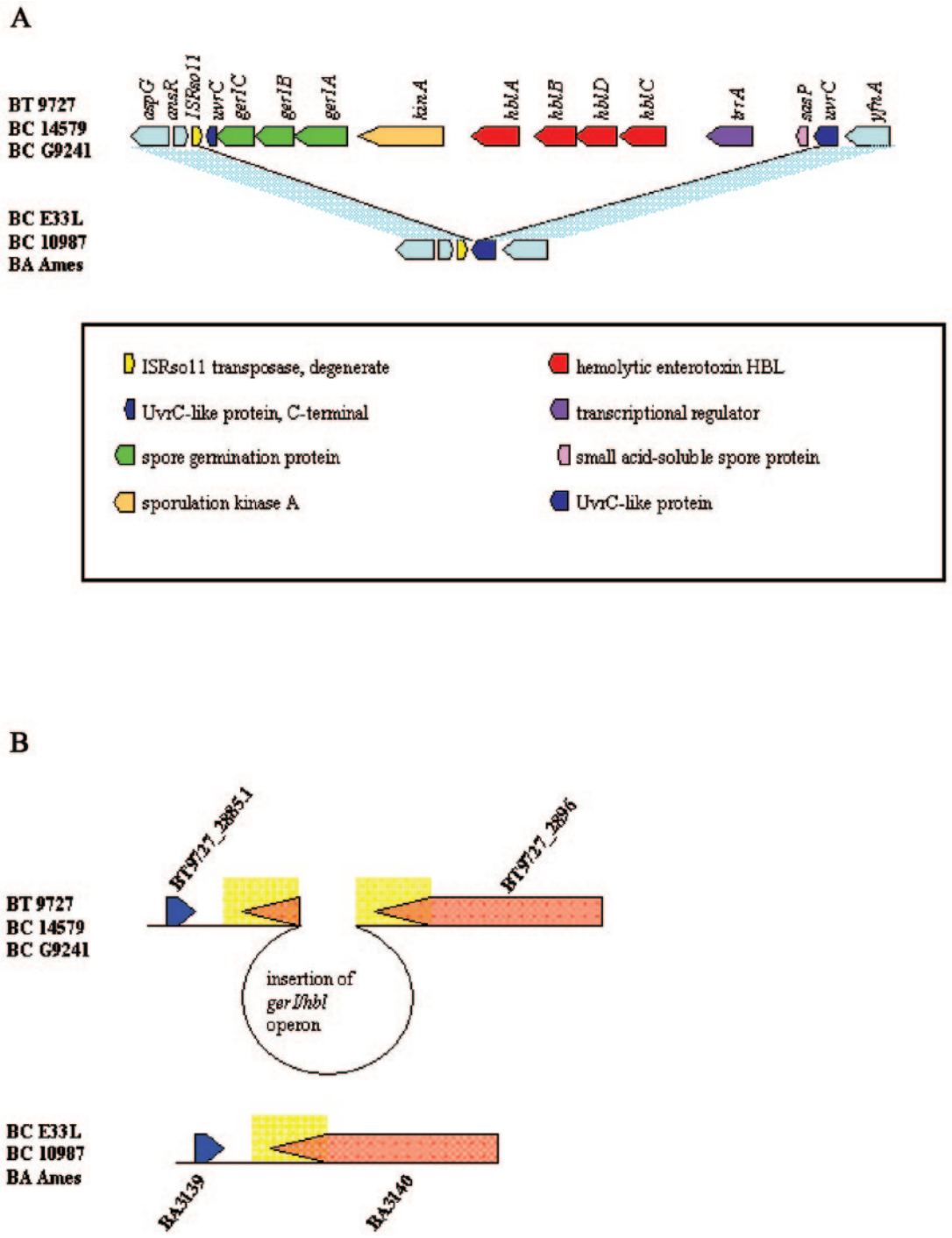

FIG. 2. A. Comparison of the *gerI* and *hbl* operon regions in *B. cereus*, *B. thuringiensis*, and *B. anthracis*. The light blue area between the two groups indicates that these regions share a high level of identity. A conserved region consists of five contiguous genes in *B. anthracis* Ames, including L-asparaginase (BA3137), *ans* operon repressor (BA3138), degenerate IS*Rso11* transposase (BA3139), UvrC-like protein (BA3140), and amino acid permease (BA3141). B. Flanking region of insertion boundary. The orthologs of genes are shown as arrows of the same color. BT9727_2896/BA3140 encode UvrC-like proteins. BT9727_2885.1/BA3139 encode degenerate IS*Rso11* transposase. Yellow blocks denote the direct repeats found around the insertion boundary. The red triangle indicates the genes of the C-terminal UvrC-like protein fragment.

ringiensis 97-27 and in *B. cereus* E33L. Analyzing the upstream sequences of coding regions for PlcR binding motifs identified genes likely to be activated by PlcR in *B. thuringiensis* 97-27 and *B. cereus* E33L. We found motifs upstream of most of the genes previously identified as potential members of a *plcR* regulon in *B. cereus* (21) (see Table S1 in the supplemental material). Of particular interest are genes encoding probable virulence factors. In this respect, we found that the nonhemo-

TABLE 1. Major phenotypic characteristics of *B. cereus* group genomes

Characteristic	*B. anthracis* Ames	*B. cereus* ATCC 10987	*B. cereus* ATCC 14579	*B. thuringiensis* 97-27	*B. cereus* E33L
Plasmid	pXO1 (189 kb), pXO2 (96 kb)	pBc10987 (208 kb)	Linear phage-like pBClin15 (15 kb)	pBT9727 (77 kb)	pE33L466 (466 kb), pE33L54 (54 kb), pE33L9 (9 kb), pE33L8 (8 kb), pE33L5 (5 kb)
Tripartite lethal toxin	Present	Absent	Absent	Absent	Absent
rRNA	33 copies	12 copies	39 copies	39 copies	39 copies
Urease gene cluster	Absent	Present	Absent	Absent	Absent
Xylose utilization genes	Absent	Present	Absent	Absent	Absent
Capsule synthesis					
Polysaccharide capsule	Absent	Present[a]	Present[a]	Present[a]	Absent
Polyglutamic acid capsule	Present[a]	Absent	Absent	Absent	Absent
Flagellar genes	Fragmental	Intact	Intact	Intact	Intact
N-Acetylgalactosamine degradation	Absent	Present	Absent	Absent	Present
Functional PlcR	Absent	Present[b]	Present[b]	Present[b]	Present[b]
Phage[d]	4	4 (1 degenerate)	6 (1 linear plasmid)	7	18
Arginine degradation					
Arginase genes	Present	Present	Present	Present	Present
Arginine deiminase genes	Absent	Present	Present	Present	Absent
B.cereus repeat 1 (*bcr1*)[e]	10 copies	72 copies	56 copies	19 copies	22 copies
Restriction enzymes	Absent	Present[c]	Present[c]	Partial	Absent
Enterotoxins					
Hemolytic enterotoxin HBL	Absent	Present	Absent	Present	Absent
Nonhemolytic enterotoxin *NHE*	Present	Present	Present	Present	Present
Diarrheal toxin BceT	Absent	Absent	Present	Present	Absent
cry genes	Absent	Absent	Absent	Absent	Absent

[a] The capsule biosynthesis region of 20 kb is partially similar in both *B. cereus* genomes.
[b] *plcR* in *B. anthracis* Ames contains a frameshift that results in a truncated and nonfunctional protein. The *plcR* gene in the *B. cereus* strains is full length, apparently functional, and can act as a regulatory protein.
[c] There are four unique restriction-modification systems in *B. cereus* ATCC 10987 and three in *B. cereus* ATCC 14579. *B. thuringiensis* 97-27 and *B. anthracis* have a CDS weakly similar to the 5-methylcytosine-specific Mrr endonuclease.
[d] The phages are not conserved in sequence or genomic location in the genomes studied.
[e] *bcr1* is a 160-bp repeated DNA sequence with unknown function overwhelmingly overrepresented in intergenic regions of the *B. cereus* group organisms.

lytic enterotoxin genes (*nheA*, *nheB*, and *nheC*) in both *B. thuringiensis* 97-27 and *B. cereus* E33L contained upstream PlcR motifs. In *B. thuringiensis* 97-27, *B. cereus* E33L, and *B. cereus* (ATCC 14579), there are PlcR motifs upstream of cytotoxin K and several proteases, including collagenase, bacillolysin, enhancin, aminopeptidase Y, and peptidase T. In addition, we found PlcR motifs upstream of the phospholipase C and phosphatidylinositol-specific phospholipase C genes in all three genomes. Another gene that has an upstream PlcR motif in *B. thuringiensis* 97-27, *B. cereus* E33L, and *B. cereus* (ATCC 14579) is error-prone DNA polymerase IV. This gene was previously suggested to induce adaptive point mutations that may affect pathogenicity. These observations (21) support the hypothesis that differences in virulence among *B. anthracis*, *B. cereus*, and *B. thuringiensis* are predominately due to alterations in gene expression rather than simple gain or loss of gene functions.

As *B. thuringiensis* isolates usually contain *cry*, *cyt*, and/or *vip* genes encoding insecticidal crystalline toxins, we compiled a list of 131 *cry* and *cyt* gene sequences (6), as well as the sequences of *vip3A* (11), *vip3V* (9), *vip1Ac* and *vip2Ac* (38), and other more recently identified *cry* gene sequences and blasted these sequences against the *B. thuringiensis* 97-27 genome. There were no full-length hits to any of the query sequences. For the most part, the partial hits had low identities (under 40 to 50%) with the query sequences. Manual examination of the annotated *B. thuringiensis* 97-27 genes that were partial hits and further analysis of these genes did not reveal any obvious candidates for *cry*, *cyt*, or *vip* genes. Most of these *B. thuringiensis* 97-27 genes had >80% amino acid sequence identity to other *B. cereus* group (noninsecticidal) genes. So, we are confident that the genome of our current isolate contains no homologs of the known *cry*, *cyt*, or *vip* genes. However, another possibility is that the plasmid encoding these genes was lost during culture.

Capsule biosynthetic genes. Many microbial pathogens produce polymeric capsules that provide protection against host immune systems during the invasion process. *Bacillus* species can produce both polysaccharide capsules that are common to many gram-positive and gram-negative species and the less common polyglutamic acid capsule. A summary of the capsule biosynthetic content in the sequenced members of the *B. cereus* group is provided in the supplemental material (see Fig. S3). In *B. anthracis*, three pXO2-encoded genes, *capB*, *capC*, and *capA*, are required for synthesis of the polyglutamic acid capsule, and this structure plays a key role in the virulence of this organism (12). To date, all *B. cereus* group strains appear to contain a weak homolog of the pXO2 *capA* gene. This is also true in *B. cereus* E33L, which contains a putative protein with 32% identity to *capA*. However, *B. cereus* E33L does not have a polyglutamic acid capsule (P. C. B. Turnbull, personal communication), nor does it appear to encode any genes involved in polysaccharide capsule synthesis (Table 1). Interestingly, the *B. thuringiensis* 97-27 genome and *B. cereus* 14579 encode a homolog of a member of a polysaccharide capsule synthesis

TABLE 2. Protein and sugar utilization genes in the *B. cereus* group

Utilization and function	No. of genes[a]				
	B. subtilis	*B. anthracis* Ames	*B. cereus* 14579	*B. cereus* E33L	*B. thuringiensis* 97-27
Amino acid and peptide utilization					
Peptide ABC transporter-ATP binding protein	7	18	23	20	18
Branched chain amino acid transporter	4	11	11	13	14
LysE/RhtB/CadD amino acid efflux system	2	6	8	9	8
Peptidase	30	64	91	92	90
Protease	24	50	49	61	52
Amino acid and amine catabolism	34	52	55	55	55
(BA0242) tyrosine degradation	No	Yes	Yes	Yes	Yes
Epr Bpr AprX protease	Yes	No	No	No	No
Sugar utilization					
PTS-sugar transporter	25	19	18	23	20
Carbohydrate polymer degradation	41	12	12	23	12
Mannose, arabinose, rhamnose catabolic pathway	Yes	No	No	No	No

[a] No and yes indicate the absence and presence, respectively, of genes in the pathway.

pathway recently identified on a plasmid of the pathogenic *B. cereus* G9241 (19).

Sporulation and germination. In anthrax, spores are the agent of infection. Spore formation occurs in response to nutrient limitation in the environment. In *B. subtilis*, sporulation is initiated by a deficiency in carbon or nitrogen (29) and is linked to changes in the expression of genes for degradative enzymes, such as alpha amylase, neutral protease, and alkaline protease (20). The *B. subtilis* spore coat is composed of at least 30 polypeptides, homologs of many of the *B. subtilis* spore coat protein genes present in *B. anthracis* (23). We found differences in the number and composition of genes encoding spore coat proteins among *B. thuringiensis* 97-27, *B. cereus* E33L, *B. cereus* ATCC 14579, and *B. anthracis* (see Table S2 in the supplemental material). Both *B. anthracis* and *B. cereus* spores germinate in response to L-alanine and ribosides. The germination response to L-alanine and ribosides requires proteins of the *gerA* family (5). *B. thuringiensis* 97-27, *B. cereus* E33L, *B. cereus* 569, and *B. cereus* ATCC 14579 have a *gerI* operon that is involved in an inosine-induced germination. The *gerI* operon is homologous to the *gerA* family operons of *B. subtilis* (5) and the *gerH* operon in *B. anthracis* (41). Similarly, the *gerQ* operon encodes germinant receptors that respond to inosine (2). *B. thuringiensis* 97-27 encodes *gerQ*, while *B. cereus* E33L does not.

Carbohydrate and amino acid utilization. Like *B. cereus* 14579 and *B. anthracis* (21, 33), *B. thuringiensis* 97-27 and *B. cereus* E33L appear predisposed to an environment rich in protein, having fewer genes for carbohydrate catabolism and more genes for amino acid metabolism (Table 2). For example, there are 12 carbohydrate polymer degradation genes in *B. anthracis* Ames, *B. cereus* 14579, and *B. thuringiensis* 97-27 and 23 in *B. cereus* E33L, compared to 41 in *B. subtilis*. The *B. cereus* group also appears to have reduced numbers of sugar-specific phosphoenolpyruvate-dependent phosphotransferase system genes. In contrast, members of the *B. cereus* group have an expanded capacity for amino acid and peptide utilization. For example, there are 18 to 23 genes encoding peptide/amino acid ABC transporter-ATP binding proteins in the *B. cereus* group, compared to 7 in *B. subtilis*. There are six to nine genes for LysE family amino acid efflux system proteins in *B. cereus* group members and only two in *B. subtilis*. In addition to the expanded number of peptidase and protease genes in *B. cereus* group species, 52 to 55 genes encode proteins involved in amino acid and amine catabolic pathways compared to 34 in *B. subtilis*. These observations suggest that proteins, peptides, and amino acids may be a preferred nutrient source for all members of the *B. cereus* group, which is consistent with the observations made previously (21, 32).

Although the *B. cereus* group species are closely related, variations in the sugar catabolism pathways are observed. Of particular note, *B. cereus* E33L has 11 extra genes for carbohydrate polymer degradation compared with *B. anthracis* Ames. Most of these are located on the large plasmid pE33L466. One of the most significant differences between *B. cereus* E33L and other isolates in the *Bacillus cereus* group is the large number of carbohydrate utilization gene clusters organized as operons on the pE33L466 plasmid (Fig. 3). These genes encode enzymes for myo-inositol degradation, galactose utilization, and pectin and gellan degradation. It is worth noting that the region of pE33L466 containing these four interlocked gene clusters is flanked by IS elements that were probably involved in their mobilization and integration into pE33L466. Figure 3 illustrates the metabolic pathways in which the products of these genes participate.

Antibiotic resistance. The ecological niche and potential virulence of *B. thuringiensis* 97-27 and *B. cereus* E33L may be expanded through the presence of two lantibiotic resistance operons that are not present in *B. cereus* or *B. anthracis*. These include a mersacidin resistance operon consisting of *mrsR2*, *mrsK2*, *mrsF*, *mrsG*, and *mrsE* and a salivaricin resistance operon consisting of *salY*, *salK*, and *salR*. *B. thuringiensis* 97-27 and *B. cereus* E33L have all of the genes in the mersacidin operon, while *B. anthracis* strains A2012, Ames, and Sterne only have *mrsF*. Although *B. thuringiensis* 97-27 and *B. cereus* E33L have all of the genes in the salivaricin and mersacidin resistance operons, they do not encode the *mrsA* gene to produce mersacidin or the *salA* gene to produce salivaricin. Therefore, these organisms can detect the presence of mersacidin and salivaricin produced by other bacteria but do not encode

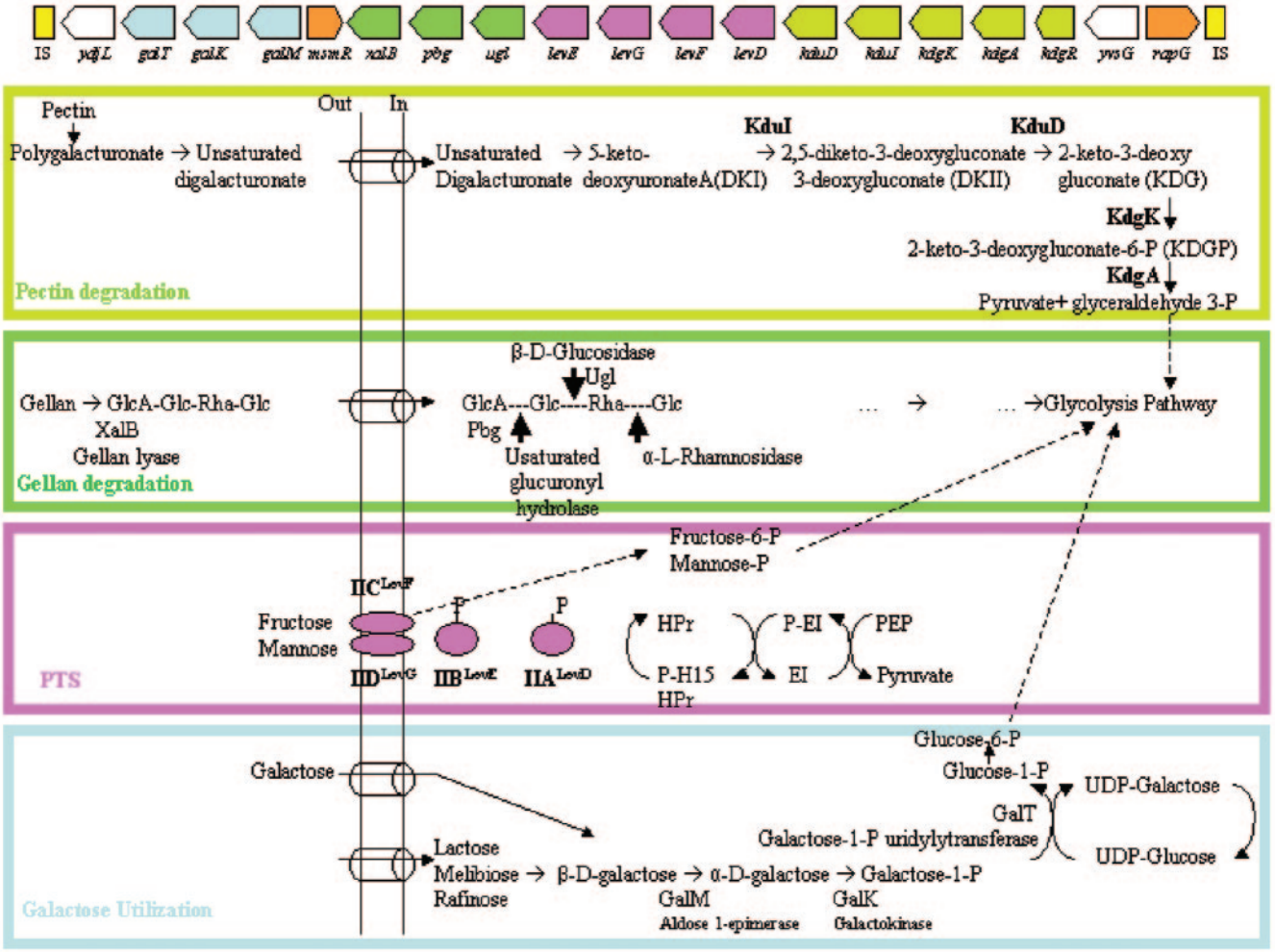

FIG. 3. Schematic presentation of phosphotransferase system-catalyzed sugar uptake and phosphorylation in *B. cereus* E33L, showing possible metabolic pathways catalyzed by the products of genes in this polymorphic locus. Steps along the pathways are catalyzed by the gene products specified near the corresponding arrow.

the capability to produce these lantibiotics themselves. Instead, the response may include increased expression of genes encoding other lantibiotics or virulence factors as previously suggested (40).

DISCUSSION

There is considerable debate in regard to the systematic classification of members of the *B. cereus* group. Historically, these organisms were classified into three species (*B. cereus*, *B. thuringiensis*, and *B. anthracis*) on the basis of distinct phenotypic differences that defined them. For example, the isolation of an organism from an animal with anthrax resulted in the designation of *B. anthracis*. While the relationship between these organisms is still not clearly understood, recent molecular approaches (15, 18, 31, 33) have revealed extensive similarities between genomes and relatively few consistent differences warranting the segregation of isolates into discrete species classified as *B. anthracis*, *B. cereus*, and *B. thuringiensis*. One unifying concept that has emerged from nucleic acid sequence analyses is that the *B. cereus* group has evolved as asexually derived clonal populations (15, 18, 28, 30). This has allowed most of the vast number of isolates from this group to be subdivided into consistent phylogenetic clusters.

In this classification scheme (30), *B. thuringiensis* 97-27 and *B. cereus* E33L are both members of the anthracis lineage and are descended from ancestral clones that are very distinct from the tolworthi, kurstaki, sotto, and thuringiensis lineages. Importantly, the anthracis lineage provides a molecular-based distinction that separates commercially important *B. thuringiensis* strains from pathogenic *B. anthracis*.

The *B. thuringiensis* 97-27 chromosome and plasmid lacked the typical *cry*, *cyt*, and *vip* genes encoding the insecticidal proteins characteristic of strains that are known to produce entomopathic toxins. The original *B. thuringiensis* designation for this isolate was due to the discovery of crystals in the initial characterization of the strain (17). However, in a subsequent publication, a second isolate (from the same patient) lacking crystalline toxin was mentioned as a spontaneous mutant (16). It is possible that the plasmid(s) encoding the crystalline toxin genes was lost spontaneously during culture; this has been documented for other *B. thuringien-*

sis plasmids (13, 24). Certainly, *B. thuringiensis* 97-27 is distinct from other known *B. thuringiensis* isolates in that it is suspected of causing human morbidity (17) resulting in severe tissue necrosis. It was subsequently demonstrated to cause lethal infection in laboratory mice (16). The phylogenetic lineage placement, subsequent laboratory diagnostics, and our comparative sequence analysis suggest that *B. thuringiensis* 97-27 is more like a pathogenic *B. cereus* strain than an insecticidal strain.

Both *B. thuringiensis* 97-27 and *B. cereus* E33L have homologs of chromosomal virulence genes found in other members of the *B. cereus* group. Consequently, the isolation of *B. thuringiensis* 97-27 from a rare case of disease and the presence of common *B. cereus* group chromosomal virulence genes make it likely that this organism is an opportunistic pathogen. While *B. cereus* E33L came from a carcass swab, it is probably an environmental isolate and not the cause of death. The relationships between members of the *B. cereus* group are nonlinear and complex, likely resulting from cycles of isolation and niche expansion facilitated, at least in part, by horizontal gene transfer mechanisms. While the germination of *Bacillus anthracis* spores or its vegetative growth may be limited to nutritionally rich environments like that found in a mammalian host, the rapid death of the host resulting from vegetative growth would limit the opportunity for genetic exchange and would result in the homogeneity observed in sequenced strains of this species. In contrast, the capacity for vegetative growth outside of an infected host or nonlethal infection provides an opportunity for genetic exchange and niche expansion. The sequences of the two *B. cereus* group members presented here provide fertile ground to study the evolution of host range and virulence.

ACKNOWLEDGMENTS

We thank Peter Turnbull for supplying the isolation and identification information for *B. cereus* E33L and for providing constructive criticism of the manuscript. We also thank Martin Hugh-Jones for providing information about the natural history of *B. anthracis* and Mich Chandler at LMGM for providing us with a list of insertion sequences.

This program is supported by the U.S. Department of Energy under contract no. W-7405-ENG-36.

REFERENCES

1. **Agaisse, H., M. Gominet, O. A. Okstad, A. B. Kolsto, and D. Lereclus.** 1999. PlcR is a pleiotropic regulator of extracellular virulence factor gene expression in *Bacillus thuringiensis*. Mol. Microbiol. **32:**1043–1053.
2. **Barlass, P. J., C. W. Houston, M. O. Clements, and A. Moir.** 2002. Germination of *Bacillus cereus* spores in response to L-alanine and to inosine: the roles of gerL and gerQ operons. Microbiology (Reading) **148:**2089–2095.
3. **Beecher, D. J., J. L. Schoeni, and A. C. L. Wong.** 1995. Enterotoxic activity of hemolysin BL from *Bacillus cereus*. Infect. Immun. **63:**4423–4428.
4. **Benson, G.** 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. **27:**573–580.
5. **Clements, M. O., and A. Moir.** 1998. Role of the gerI operon of *Bacillus cereus* 569 in the response of spores to germinants. J. Bacteriol. **180:**6729–6735.
6. **Crickmore, N., D. Zeigler, J. Feitelson, E. Schnepf, J. Van Rie, D. Lereclus, J. Baum, and D. Dean.** 1998. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. **62:**807–813.
7. **Delcher, A., D. Harmon, S. Kasif, O. White, and S. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636–4641.
8. **Delcher, A. L., A. Phillippy, J. Carlton, and S. L. Salzberg.** 2002. Fast algorithms for large-scale genome alignment and comparison. Nucleic Acids Res. **30:**2478–2483.
9. **Doss, V., K. Kumar, R. Jayakumar, and V. Sekar.** 2002. Cloning and expression of the vegetative insecticidal protein (vip3V) gene of Bacillus thuringiensis in Escherichia coli. Protein Expr. Purif. **26:**82–88.
10. **Drobniewski, F. A.** 1993. *Bacillus cereus* and related species. Clin. Microbiol. Rev. **6:**324–338.
11. **Estruch, J., G. Warren, M. Mullins, G. Nye, J. Craig, and M. Koziel.** 1996. Vip3A, a novel Bacillus thuringiensis vegetative insecticidal protein with a wide spectrum of activities against lepidopteran insects. Proc. Natl. Acad. Sci. USA **93:**5389–5394.
12. **Ezzell, J. W., and S. L. Welkos.** 1999. The capsule of *Bacillus anthracis*, a review. J. Appl. Microbiol. **87:**250.
13. **Gonzalez, J. M., H. T. Dulmage, and B. C. Carlton.** 1981. Correlation between specific plasmids and delta-endotoxin production in *Bacillus thuringiensis*. Plasmid **5:**351–365.
14. **Guttmann, D. M., and D. J. Ellar.** 2000. Phenotypic and genotypic comparisons of 23 strains from the *Bacillus cereus* complex for a selection of known and putative *B. thuringiensis* virulence factors. FEMS Microbiol. Lett. **188:**7–13.
15. **Helgason, E., O. A. Okstad, D. A. Caugant, H. A. Johansen, A. Fouet, M. Mock, I. Hegna, and A. B. Kolsto.** 2000. *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*: one species on the basis of genetic evidence. Appl. Environ. Microbiol. **66:**2627–2630.
16. **Hernandez, E., F. Ramisse, T. Cruel, R. le Vagueresse, and J. D. Cavallo.** 1999. *Bacillus thuringiensis* serotype H34 isolated from human and insecticidal strains serotypes 3a3b and H14 can lead to death of immunocompetent mice after pulmonary infection. FEMS Immunol. Med. Microbiol. **24:**43–47.
17. **Hernandez, E., F. Ramisse, J. P. Ducoureau, T. Cruel, and J. D. Cavallo.** 1998. *Bacillus thuringiensis* subsp. *konkukian* (serotype H34) superinfection: case report and experimental evidence of pathogenicity in immunosuppressed mice. J. Clin. Microbiol. **36:**2138–2139.
18. **Hill, K. K., L. O. Ticknor, R. T. Okinaka, M. Asay, H. Blair, K. A. Bliss, M. Laker, P. E. Pardington, A. P. Richardson, M. Tonks, D. J. Beecher, J. D. Kemp, A. B. Kolsto, A. C. L. Wong, P. Keim, and P. J. Jackson.** 2004. Fluorescent amplified fragment length polymorphism analysis of *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* isolates. Appl. Environ. Microbiol. **70:**1068–1080.
19. **Hoffmaster, A. R., J. Ravel, D. A. Rasko, G. D. Chapman, M. D. Chute, C. K. Marston, B. K. De, C. T. Sacchi, C. Fitzgerald, L. W. Mayer, M. C. J. Maiden, F. G. Priest, M. Barker, L. X. Jiang, R. Z. Cer, J. Rilstone, S. N. Peterson, R. S. Weyant, D. R. Galloway, T. D. Read, T. Popovic, and C. M. Fraser.** 2004. Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax. Proc. Natl. Acad. Sci. USA **101:**8449–8454.
20. **Honjo, M., A. Nakayama, K. Fukazawa, K. Kawamura, K. Ando, M. Hori, and Y. Furutani.** 1990. A novel *Bacillus subtilis* gene involved in negative control of sporulation and degradative enzyme production. J. Bacteriol. **172:**1783–1790.
21. **Ivanova, N., A. Sorokin, I. Anderson, N. Galleron, B. Candelon, V. Kapatral, A. Bhattacharyya, G. Reznik, N. Mikhailova, A. Lapidus, L. Chu, M. Mazur, E. Goltsman, N. Larsen, M. D'Souza, T. Walunas, Y. Grechkin, G. Pusch, R. Haselkorn, M. Fonstein, S. D. Ehrlich, R. Overbeek, and N. Kyrpides.** 2003. Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. Nature **423:**87–91.
22. **Jensen, G. B., B. M. Hansen, J. Eilenberg, and J. Mahillon.** 2003. The hidden lifestyles of *Bacillus cereus* and relatives. Environ. Microbiol. **5:**631–640.
23. **Lai, E., N. Phadke, M. Kachman, R. Giorno, S. Vazquez, J. Vazquez, J. Maddock, and A. Driks.** 2003. Proteomic analysis of the spore coats of *Bacillus subtilis* and *Bacillus anthracis*. J. Bacteriol. **185:**1443–1454.
24. **Lereclus, D., O. Arantes, J. Chaufaux, and M. Lecadet.** 1989. Transformation and expression of a cloned delta endotoxin gene in Bacillus thuringiensis. FEMS Microbiol. Lett. **60:**211–218.
25. **Lowe, T., and S. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
26. **Okinaka, R. T., K. Cloud, O. Hampton, A. R. Hoffmaster, K. K. Hill, P. Keim, T. M. Koehler, G. Lamke, S. Kumano, J. Mahillon, D. Manter, Y. Martinez, D. Ricke, R. Svensson, and P. J. Jackson.** 1999. Sequence and organization of pXO1, the large *Bacillus anthracis* plasmid harboring the anthrax toxin genes. J. Bacteriol. **181:**6509–6515.
27. **Pannucci, J., R. T. Okinaka, R. Sabin, and C. R. Kuske.** 2002. *Bacillus anthracis* pXO1 plasmid sequence conservation among closely related bacterial species. J. Bacteriol. **184:**134–141.
28. **Pearson, T., J. D. Busch, J. Ravel, T. D. Read, S. D. Rhoton, J. M. U'Ren, T. S. Simonson, S. M. Kachur, R. R. Leadem, M. L. Cardon, M. N. Van Ert, L. Y. Huynh, C. M. Fraser, and P. Keim.** 2004. Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing. Proc. Natl. Acad. Sci. USA **101:**13536–13541.
29. **Piggot, P. J., and J. G. Coote.** 1976. Genetic aspects of bacterial endospore formation. Bacteriol. Rev. **40:**908–962.
30. **Priest, F. G., M. Barker, L. W. J. Baillie, E. C. Holmes, and M. C. J. Maiden.** 2004. Population structure and evolution of the *Bacillus cereus* group. J. Bacteriol. **186:**7959–7970.
31. **Radnedge, L., P. G. Agron, K. K. Hill, P. J. Jackson, L. O. Ticknor, P. Keim, and G. L. Andersen.** 2003. Genome differences that distinguish *Bacillus*

anthracis from *Bacillus cereus* and *Bacillus thuringiensis*. Appl. Environ. Microbiol. **69:**2755–2764.

32. **Rasko, D. A., J. Ravel, O. A. Okstad, E. Helgason, R. Z. Cer, L. X. Jiang, K. A. Shores, D. E. Fouts, N. J. Tourasse, S. V. Angiuoli, J. Kolonay, W. C. Nelson, A. B. Kojsto, C. M. Fraser, and T. D. Read.** 2004. The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1. Nucleic Acids Res. **32:**977–988.
33. **Read, T. D., S. N. Peterson, N. Tourasse, L. W. Baillie, I. T. Paulsen, K. E. Nelson, H. Tettelin, D. E. Fouts, J. A. Eisen, S. R. Gill, E. K. Holtzapple, O. A. Okstad, E. Helgason, J. Rilstone, M. Wu, J. F. Kolonay, M. J. Beanan, R. J. Dodson, L. M. Brinkac, M. Gwinn, R. T. DeBoy, R. Madpu, S. C. Daugherty, A. S. Durkin, D. H. Haft, W. C. Nelson, J. D. Peterson, M. Pop, H. M. Khouri, D. Radune, J. L. Benton, Y. Mahamoud, L. X. Jiang, I. R. Hance, J. F. Weidman, K. J. Berry, R. D. Plaut, A. M. Wolf, K. L. Watkins, W. C. Nierman, A. Hazen, R. Cline, C. Redmond, J. E. Thwaite, O. White, S. L. Salzberg, B. Thomason, A. M. Friedlander, T. M. Koehler, P. C. Hanna, A. B. Kolsto, and C. M. Fraser.** 2003. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. Nature **423:**81–86.
34. **Salamitou, S., F. Ramisse, M. Brehelin, D. Bourguet, N. Gilois, M. Gominet, E. Hernandez, and D. Lereclus.** 2000. The plcR regulon is involved in the opportunistic properties of *Bacillus thuringiensis* and *Bacillus cereus* in mice and insects. Microbiology **146:**2825–2832.
35. **Salzberg, S., A. Delcher, S. Kasif, and O. White.** 1998. Microbial gene identification using interpolated Markov models. Nucleic Acids Res. **26:**544–548.
36. **Schnepf, E., N. Crickmore, J. Van Rie, D. Lereclus, J. Baum, J. Feitelson, D. R. Zeigler, and D. H. Dean.** 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. **62:**775–806.
37. **Schwartz, S., Z. Zhang, K. A. Frazer, A. Smit, C. Riemer, J. Bouck, R. Gibbs, R. Hardison, and W. Miller.** 2000. PipMaker: a web server for aligning two genomic DNA sequences. Genome Res. **10:**577–586.
38. **Shi, Y., W. Xu, M. Yuan, M. Tang, J. Chen, and Y. Pang.** 2004. Expression of vip1/vip2 genes in Escherichia coli and Bacillus thuringiensis and the analysis of their signal peptides. J. Appl. Microbiol. **97:**757–765.
39. **Ticknor, L. O., A. B. Kolsto, K. K. Hill, P. Keim, M. T. Laker, M. Tonks, and P. J. Jackson.** 2001. Fluorescent amplified fragment length polymorphism analysis of Norwegian *Bacillus cereus* and *Bacillus thuringiensis* soil isolates. Appl. Environ. Microbiol. **67:**4863–4873.
40. **Upton, M., J. R. Tagg, P. Wescombe, and H. F. Jenkinson.** 2001. Intra- and interspecies signaling between *Streptococcus salivarius* and *Streptococcus pyogenes* mediated by SalA and SalA1 lantibiotic peptides. J. Bacteriol. **183:** 3931–3938.
41. **Weiner, M., T. D. Read, and P. C. Hanna.** 2003. Identification and characterization of the *gerH* operon of *Bacillus anthracis* endospores: a differential role for purine nucleosides in germination. J. Bacteriol. **185:**1462–1464.