Comparative Analysis of Bacillus anthracis, Bacillus cereus, and Related Species on the Basis of Reverse Transcriptase Sequencing of 16S rRNA

CAROL ASH, JOHN A. E. FARROW, MATTHIAS DORSCH, ERKO STACKEBRANDT, and MATTHEW D. COLLINS*

Department of Microbiology, AFRC Institute of Food Research, Shinfield, Reading, RG2 9AT, United Kingdom, and Department of Microbiology, University of Queensland, Queensland 4072, Australia

The primary structures of the 16S rRNAs of Bacillus anthracis, Bacillus cereus, Bacillus mycoides, and Bacillus thuringiensis were determined by using the reverse transcription-dideoxy sequencing method. All of the strains exhibited very high levels of sequence similarity (>99%) that were consistent with the close relationships shown by previous DNA hybridization studies. The sequences of B. anthracis Sterne and B. cereus emetic strain NCTC 11143 were found to be identical for a continuous stretch of 1,446 bases and differed from the sequence of B. cereus NCDO 1771T (T = type strain) by only a single nucleotide. The 16S rRNA sequences of B. mycoides and B. thuringiensis differed from each other and from the sequences of B. anthracis and B. cereus by four to nine nucleotides.

The species Bacillus anthracis, Bacillus cereus, Bacillus mycoides, and Bacillus thuringiensis were originally described on the basis of their habitats, their pathogenicity for mammals or insects, and their morphological and physiological characteristics. However, the taxonomic interrelationships of these species are equivocal. All four species share many phenotypic properties, and several workers have questioned their status as separate species (6, 11, 12). Small-subunit rRNA is now recognized as a powerful molecular chronometer (16). Degrees of sequence conservation, ranging from highly variable to highly conserved regions, enable systematists to measure small as well as great genealogical distances. In this study we determined partial primary 16S rRNA sequences of B. anthracis, B. cereus, B. mycoides, and B. thuringiensis in order to investigate the genealogical interrelationships of these organisms.

MATERIALS AND METHODS

Cultures and cultivation. Details concerning the test strains which we examined are shown in Table 1. Strains were grown in shake flasks containing nutrient broth no. 2 (Oxoid) to late exponential phase at 30°C.

Extraction and sequence determination of 16S rRNA. Total cellular rRNA was extracted from ca. 2 g of wet cells by mechanical disruption, using glass beads and a Braun homogenizer, and was purified as described by Embley et al. (5). Nucleotide sequences were determined by the Sanger dideoxynucleotide method (9) directly from cellular rRNA, using avian myeloblastosis virus reverse transcriptase (8). The sequences of the oligonucleotide primers were the same as those described by Embley et al. (5). In addition, the following primer was used: 5'TCACCAACTAGCTAATG, which is complementary to positions 258 to 242 (Escherichia coli nomenclature). This primer was included to enable determination of the sequence between positions 100 and 150, which could not always be fully established by using the primer at position 357 described by Lane et al. (8). Nucleotide sequence accession numbers. The 16S rRNA sequences have been deposited in the EMBL Data Library (accession numbers X55059 to X55063).

RESULTS AND DISCUSSION

The 16S rRNA sequences of B. anthracis Sterne, B. cereus NCDO 1771T (T = type strain), B. cereus NCTC 11143 (emetic strain), B. mycoides DSM 2048T, and B. thuringiensis NCDO 1771T were determined by using the reverse transcription-dideoxy sequencing method. All of the strains exhibited very high levels of sequence similarity (>99%) that were consistent with the close relationships shown by previous DNA hybridization studies. The sequences of B. anthracis Sterne and B. cereus emetic strain NCTC 11143 were found to be identical for a continuous stretch of 1,446 bases and differed from the sequence of B. cereus NCDO 1771T (T = type strain) by only a single nucleotide. The 16S rRNA sequences of B. mycoides and B. thuringiensis differed from each other and from the sequences of B. anthracis and B. cereus by four to nine nucleotides.

<table>
<thead>
<tr>
<th>Organism</th>
<th>B. anthracis Sterne</th>
<th>B. cereus NCDO 1771T</th>
<th>B. cereus NCTC 11143</th>
<th>B. mycoides DSM 2048T</th>
<th>B. thuringiensis NCDO 1771T</th>
<th>B. subtilis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>99.9</td>
<td>100</td>
<td>99.5</td>
<td>99.7</td>
<td>94.1</td>
<td></td>
</tr>
<tr>
<td>B. anthracis Sterne</td>
<td>99.9</td>
<td>100</td>
<td>99.5</td>
<td>99.7</td>
<td>94.1</td>
<td></td>
</tr>
<tr>
<td>B. cereus emetic strain</td>
<td>0.1</td>
<td>1</td>
<td>99.5</td>
<td>99.7</td>
<td>94.0</td>
<td></td>
</tr>
<tr>
<td>NCTC 11143</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>94.1</td>
<td>85</td>
</tr>
<tr>
<td>B. mycoides DSM 2048T</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>94.1</td>
<td>85</td>
</tr>
<tr>
<td>B. thuringiensis NCDO 1771T</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>94.1</td>
<td>88</td>
</tr>
<tr>
<td>B. subtilis</td>
<td>85</td>
<td>87</td>
<td>87</td>
<td>88</td>
<td>86</td>
<td></td>
</tr>
</tbody>
</table>

* The values in the upper right are percentages of similarity, and the values on the lower left are numbers of nucleotide differences. The actual numbers of base differences between strains may be marginally higher because approximately 2% of the nucleotides were not determined.

* Corresponding author.
I
B.an
B.ce
B.my
B.th
B.c.em
CGGCGGACGGGUGAGUAU~UMCCUGCCCAUAA
CGGCGGACGGGUGAGUM~CGU~UMCCUGA~~UMGACU~AUMCUC~AAAC~ACACUGGMUACCGGAUM~~GMC~U~W
CGGCGGACGGGUGAGUMCU~UMCCUACCCAUAA
CGGCGGACGGGUGAGUMCACGU~UMCCUGC~U~ACUGGGAUMCUC~AAAC~ACACUGGMUACCGGAUM~~GMACACUGG~U~W
CGGCGGACGGGUGAGUM~CGUAACCUGCCCAUAA
CGGCGGACGGGUGAGUAU~UMCCUGCCCAUAA
CGGCGGACGGGUGAGUM~CGU~UMCCUGA~~UMGACU~AUMCUC~AAAC~ACACUGGMUACCGGAUM~~GMC~U~W
CGGCGGACGGGUGAGUMCU~UMCCUACCCAUAA
CGGCGGACGGGUGAGUMCACGU~UMCCUGC~U~ACUGGGAUMCUC~AAAC~ACACUGGMUACCGGAUM~~GMACACUGG~U~W

101

B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em

201

B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em

301

B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em

401

B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em

501

B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em

601

B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em

701

B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em

801

B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em

901

B.an
B.ce
B.my
B.th
B.c.em
B.an
B.ce
B.my
B.th
B.c.em
COMPARATIVE ANALYSIS OF BACILLUS SPECIES

1001

B. an UGACACCCCGCCCGUCACACCACGAGAGACACCACACAGUMCACCCGMGUCGGNGG
B. ce UGACACCCCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG
B. my UGACACCCCGCCCGUCACACCACGAGAGACACCACACAGUAACACCGGAAGUCGGUGG
B. th UGACACCCCGCCCGUCACACCACGAGAGACACCACACAGUACACCGGAAGUCGGUGG
B. c. em UGACACCCCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG

1101

B. an UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUMCACCCGMGUCGGNGG
B. ce UGACACCCCGCGCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG
B. my UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUAACACCGGAAGUCGGUGG
B. th UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUACACCGGAAGUCGGUGG
B. c. em UGACACCCCGCGCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG

1201

B. an UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUMCACCCGMGUCGGNGG
B. ce UGACACCCCGCGCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG
B. my UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUAACACCGGAAGUCGGUGG
B. th UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUACACCGGAAGUCGGUGG
B. c. em UGACACCCCGCGCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG

1301

B. an UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUMCACCCGMGUCGGNGG
B. ce UGACACCCCGCGCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG
B. my UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUAACACCGGAAGUCGGUGG
B. th UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUACACCGGAAGUCGGUGG
B. c. em UGACACCCCGCGCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG

1401

B. an UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUMCACCCGMGUCGGNGG
B. ce UGACACCCCGCGCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG
B. my UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUAACACCGGAAGUCGGUGG
B. th UGACACCCCGCGCGCCCGUCACACCACGAGAGACACCACACAGUACACCGGAAGUCGGUGG
B. c. em UGACACCCCGCGCGCCCGUCACACCACGAGAGACGAGAGUMCACCCGMGUCGGUGG

FIG. 1. Primary structures of 16S rRNAs of B. anthracis Sterne (B. an), B. cereus NCDO 1771T (B. ce), B. cereus NCTC 11143 (B. c. em), B. mycoides DSM 2048T (B. my), and B. thuringiensis NCIMB 9134T (B. th) as determined by reverse transcription. The first and last nucleotides in the sequences are analogs to the nucleotides at positions 1 and 1,440 of the E. coli 16S rRNA sequence (2). N, undetermined nucleotide.

thuringiensis NCIMB 9134T were investigated by performing reverse transcription experiments. Each of the sequences shown in Fig. 1 comprises a continuous stretch of 1,446 bases (ranging from position 1 to position 1,440 on the E. coli numbering system), which corresponds to approximately 94% of the total 16S rRNA primary sequence. A comparison of the sequence alignments revealed very high levels of sequence similarity among all of the strains examined (Table 1). Only 11 base substitution points on the sequences were identified, and these occurred in variable domains V1, V2, and V6. Of the variable positions, four involved base pairs (77-92, 1009-1020 [E. coli numbering]). Table 2 shows the locations of the base differences in the strains. The sequences of B. anthracis Sterne and B. cereus emetic strain NCTC 11143 were found to be identical and differed from the B. cereus NCDO 1771T sequence in only a single nucleotide (position 1,005 [E. coli numbering]). In contrast, the type strains of B. mycoides and B. thuringiensis differed from each other and from B. anthracis and B. cereus strains by a slightly greater number of nucleotides (ca. four to nine bases) (Tables 1 and 2). The high overall levels of sequence similarity, together with the somewhat scattered distribution of strain-specific nucleotides, eliminate the possibility that diagnostic species-specific oligonucleotide probes could be designed.

Previous chromosomal DNA-DNA hybridization investigations indicated that B. anthracis, B. cereus, B. mycoides, and B. thuringiensis are closely related and probably prese-
sent a single species (10, 13). Our 16S rRNA sequence data also clearly demonstrate that these "species" form a genealogically tight group of microorganisms, with the number of base differences comparable to the numbers found at the intraspecific level for many other gram-positive species (for example, zero to six differences in Lactococcus lactis [3] and Streptococcus parasanguis [14] and zero to nine differences in Mycobacterium intracellulare [1]). It must be emphasized however that it is not possible to draw firm conclusions regarding species differentiation on the basis of rRNA sequence relatedness data alone (because of differing rates of sequence divergence). Several other distinct genospecies are known which have similarly small numbers of nucleotide differences (e.g., members of the Enterococcus avium group of species [15], Enterococcus casseliflavus and Enterococcus gallinarum [15], and Aerococcus viridans genospecies 1 and 2 [4]). Clearly a comprehensive reexamination of the levels of DNA-DNA relatedness among members of the B. cereus complex will be required to resolve the specific interrelationships of these organisms.

ACKNOWLEDGMENTS

We are grateful to R. A. Hutson (Porton Down, United Kingdom) for assistance in the preparation of B. anthracis rRNA. E.S. and M.D. are grateful to Gene-Trak Systems for support.

REFERENCES