Discrimination of Bacillus anthracis and closely related microorganisms by analysis of 16S and 23S rRNA with oligonucleotide microarray


Abstract

Analysis of 16S rRNA sequences is a commonly used method for the identification and discrimination of microorganisms. However, the high similarity of 16S and 23S rRNA sequences of Bacillus cereus group organisms (up to 99–100%) and repeatedly failed attempts to develop molecular typing systems that would use DNA sequences to discriminate between species within this group have resulted in several suggestions to consider B. cereus and B. thuringiensis, or these two species together with B. anthracis, as one species. Recently, we divided the B. cereus group into seven subgroups, Anthracis, Cereus A and B, Thuringiensis A and B, and Mycoides A and B, based on 16S rRNA, 23S rRNA and gyrB gene sequences and identified subgroup-specific markers in each of these three genes. Here we for the first time demonstrated discrimination of these seven subgroups, including subgroup Anthracis, with a 3D gel element microarray of oligonucleotide probes targeting 16S and 23S rRNA markers. This is the first microarray enabled identification of B. anthracis and discrimination of these seven subgroups in pure cell cultures and in environmental samples using rRNA sequences. The microarray bearing perfect match/mismatch (p/mm) probe pairs was specific enough to discriminate single nucleotide polymorphisms (SNPs) and was able to identify targeted organisms in 5 min. We also demonstrated the ability of the microarray to determine subgroup affiliations for B. cereus group isolates without rRNA sequencing. Correlation of these seven subgroups with groupings based on multilocus sequence typing (MLST), fluorescent amplified fragment length polymorphism analysis (AFLP) and multilocus enzyme electrophoresis (MME) analysis of a wide spectrum of different genes, and the demonstration of subgroup-specific differences in toxin profiles, psychrotolerance, and the ability to harbor some plasmids, suggest that these seven subgroups are not based solely on neutral genomic polymorphisms, but instead reflect differences in both the genotypes and phenotypes of the B. cereus group organisms.

Keywords: Bacillus anthracis; 16S rRNA; 23S rRNA; Microarray; SNP; Environmental sample

© 2007 Elsevier Ireland Ltd. All rights reserved.

Available online 12 September 2007
1. Introduction

*Bacillus anthracis*, the causative agent of the highly infectious disease anthrax, belongs to the *Bacillus cereus* group, which also contains six other closely related species: *B. cereus*, *Bacillus thuringiensis*, *Bacillus mycoides*, *Bacillus pseudomycoides*, *Bacillus weihenstephanensis* and *Bacillus medusa* [1–6]. *B. anthracis*, which in recent years has become known as a biological weapon and terrorist tool, can be isolated from the environment. For example, *B. anthracis* is periodically detected in sewer water and in soil, among animals and staff in agricultural settings and in zoos, and epizootics among bison are rather common in Canada. While there has been a general decrease in the number of *B. anthracis* outbreaks in the U.S., Canada and Europe, anthrax still is a significant problem in West Africa, Spain, Greece, Turkey, Albania, Romania and Central Asia [7,8].

Related organisms in the *B. cereus* group are common in samples of food, water or soil [2,9–15], and some of the non-anthrax *B. cereus* group bacteria can cause human infections and illness [2,10,16] which can be rather severe [17–19]. For example, *B. cereus* is a ubiquitous soil bacterium and opportunistic human pathogen, with six recognized types of clinical infections and disease: (i) local infections, particularly of burns, traumatic or post-surgical wounds, and eye; (ii) bacteremia and septicemia; (iii) central nervous system infections, including meningitis, abscesses, and shunt associated infections; (iv) respiratory infections; (v) endosub-circuititis and perisub-circuititis; and (vi) food poisoning, characterized by toxin-induced emetic and diarrheagenic syndromes (2). In the food industry, *B. cereus* and *B. weihenstephanensis* are important as food spoilage organisms, easily contaminating raw foods and processing equipment, including equipment used in dairies and paper mills [2,9–11,14,16,19–22].

*B. thuringiensis* is widely used as a biological insecticide, accounting for more than 90% of the biopesticides employed worldwide [23,24]. Despite a long safety record, however, there is increasing public concern about the potential pathogenicity of *B. thuringiensis* due to sporadic but sometimes severe infections caused by this species, including diarrhea, food poisoning, corneal ulcer, cellulites, and burn/war wound infections. Like *B. cereus*, *B. thuringiensis* produces a variety of virulence factors, which include phospholipase C, hemolysin and enterotoxins [23,25–34], and diarrheal enterotoxin production was detected in some commercial *B. thuringiensis*-based insecticides [35]. For these reasons, many crop customers require agricultural companies to control for the presence of *B. thuringiensis* in their products. There is also evidence that specific strains of *B. cereus* and *B. thuringiensis* are associated with periodontitis and other oral infections [36].

*B. mycoides* and *B. pseudomycoides* are ubiquitous in soil and water, and there are toxigenic strains within each of these species [37–43]. Recent findings indicated that *B. mycoides*, like *B. thuringiensis*, may produce pesticidal metabolites [44]. New genomic data indicated that *B. medusa* is a variant of *B. thuringiensis*, suggesting that *B. medusa* may possess or acquire pathogenic potential. Therefore, detection and surveillance of all microorganisms comprising the *B. cereus* group is of significant interest in many aspects of worldwide public health microbiology, from confirming acts of bioterrorism to disease surveillance and outbreak prevention.

Besides traditional microbiological techniques, a variety of molecular biological methods have been described for the selective identification of *B. anthracis*. Most of them use specific genes or proteins as targets for recognition and discrimination from related microorganisms [45–48]. Genomic methods target unique gene sequences located on plasmids or in bacterial chromosomes. These methods include: microarray detection of genomic markers [33,49–51]; direct testing of plasmid pXO1 DNA with specific probes [52]; real-time multiplex PCR strand specific detection of CAP-C gene on the pXO2 plasmid [53]; ribotyping of 16S/23S rRNA genes [54–56] and their internal transcribed spacers (16S–23S rRNA ITS) [49,57,58]; detecting variations in gyrA–gyrB intergenic spacer regions [57] and in rpoB gene [59]; identification markers in gyrA and gyrB genes [60,61] and in other chromosomal sequences [62]; analysis of putative genome-integrated prophages [63]; detection of single nucleotide and repetitive element polymorphisms [64–66]; AFLP analysis of plasmid and chromosomal sequences [67–69]; MME of chromosomal genes [70], PCR-ELISA [71]; and on chip PCR amplification of anthrax toxin genes [72]. Utilization of genomic markers located in single-copy genes for identification of *B. anthracis* and discrimination of it from related microorganisms is complicated by horizontal gene transfer (HGT), which affects up to 32% of bacterial genomes [73–75], and by other mechanisms of genome reorganization [76]. Toxin genes are the subject of HGT particularly often [77]. On the other hand, discriminative identification of *B. cereus* group members other then *B. anthracis* using genomic markers was impossible until very recently [18,58,69,70,78] due to the lack of a genomic-based classification of this group. Currently, rRNA gene sequences are commonly used in microbial systematics [79,80]. Bacterial genomes usually contain 10–15
copies of rRNA genes, which are involved in translation, a fundamental cell function, and which have coevolved with ribosomal proteins. These features would seem to make rRNA genes the least likely of all genes to experience interspecies HGT [81]. In addition, GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html) and Ribosomal Database Project II, release 9 (RDP II, r. 9) (http://rdp.cme.msu.edu/index.jsp) currently contain more than 300,000 bacterial rRNA sequences including several hundred belonging to B. cereus group organisms. However, it has been considered impossible to use rRNA gene sequences for B. cereus group classification and discrimination due to the high level of rRNA homology (more than 95%) among B. cereus group organisms [78,82–84].

Recently, we reported that the B. cereus group may be divided into seven subgroups (Anthracis, Cereus A and B, Thuringiensis A and B, and Mycoides A and B) based on 16S rRNA, 23S rRNA and gyrB gene sequences and each subgroup may be characterized with specific “signatures” (genomic markers) [85]. We also demonstrated that the currently defined species B. cereus, B. thuringiensis and B. mycoides actually represent mixtures of genetically distinct bacterial groups [85]. Here we have utilized rRNA markers in a 3D gel element oligonucleotide microarray to demonstrate for the first time that B. anthracis may be unambiguously discriminated from the most closely related organisms (subgroups Cereus A and B), as well as from bacteria comprising other, more divergent subgroups of the B. cereus group. This rRNA-based microarray was also capable of identification and discrimination of all seven subgroups within the B. cereus group.

2. Materials and methods

2.1. Bacterial strains

Seven reference strains belonging to the B. cereus group, B. anthracis str. Ames, B. anthracis str. Sterne, B. thuringiensis str. B8, B. cereus str. NCTC9620, B. cereus str. T, B. thuringiensis str. 4Q281, and B. mycoides str. ATCC10206 were obtained as a generous gift from Dr. John Ezzell, U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, Frederick, MD, USA. Four other reference organisms, B. cereus NCTC11143, B. cereus ATCC43881, B. mycoides ATCC6462 and B. subtilis B-459, were obtained from the National Collection of Type Cultures and Pathogenic Fungi (NCTC), London, UK, and American Type Culture Collection (ATCC), Rockville, MD, respectively. Three other non-reference organisms, B. cereus NCTC3329, B. cereus HER1414 and B. mycoides ATCC23258 were also obtained from NCTC and ATCC. B. thuringiensis B8, related to subgroup Cereus A, apparently represents a misclassification, as it does not produce protein crystals or contain any cry genes [85] that are a fundamental characteristics of B. thuringiensis.

2.2. Oligonucleotide microarray design and construction

DNA microarrays were constructed with 11 pairs of oligonucleotide probes (ps1 through ps20, ps27 and ps28) (Table 1) targeting 16S rRNA sequence “signatures” (Fig. 1A) and four pairs of oligonucleotide probes (ps21 through ps24 and ps29 through ps32) (Table 2) targeting 23S rRNA sequence “signatures” (Fig. 1B) using 3D gel element microarray platform [86]. Probes ps25 (BCG1) and ps26 (BSG1) targeting the B. cereus group and the B. subtilis group, respectively, were selected previously [87]. Each probe was 20 nucleotides in length. Oligonucleotides were synthesized on an automatic DNA/RNA synthesizer (Applied Biosystems 394) using standard phosphoramidite chemistry. 5’-Amino-Modifier C6 (Glen Research, Sterling, VA) was linked to the 5’-end of the oligonucleotides. A micromatrix containing 100 μm × 100 μm × 20 μm polyacrylamide gel pads fixed on a glass slide and spaced 100 μm from each other was manufactured by photopolymerization [88], and activated as described earlier [89]. Six nanoliters of individual 1 nM amino-oligonucleotide solutions was applied to each gel pad containing aldehyde groups [88,90] according to the procedure described earlier [89].

2.3. Sample preparation and hybridization with microarray

Total RNA was isolated from frozen cell pellets by bead beating followed by phenol–chloroform extraction as previously described [91]. Isolated RNA was fragmented and labeled with direct magnesium-sodium periodate method and excess of the Lissamine rhodamine B ethylenediamine (LissRhod) (Molecular Probes, Eugene, OR) was removed with butanol treatment as described earlier [92]. Samples of labeled and fragmented nucleic acids from whole cells were obtained as described earlier [87,92] with some modifications. Fifty microliters of silica suspension containing 40 μl of dry silica was applied to a 25-mm-long sterile disposable centrifuge filter unit containing an Anopore membrane filter (aluminum dioxide) with a diameter of 6.5 mm and a pore size of 0.2 μm.
Fig. 1. Positions of subgroup-specific sequence differences in the 16S rRNA (A) and 23S rRNA (B) genes of seven *B. cereus* subgroups and corresponding sequences of reference microorganisms. The 16S rRNA and 23S rRNA sequences of *B. anthracis* Sterne (GenBank accession nos. AF176321 and AF267877, respectively) in 5′ to 3′ orientation have been used as the consensus sequences. Vertical lines denote nucleotides identical to the consensus sequence. Numbers indicate distance from 5′-end of consensus rRNA sequence. Principle of pair probe’s design is shown below sequences in (A). R = G or A; Y = T or C.
<table>
<thead>
<tr>
<th>Probe's target</th>
<th>Anthr; Cer A; Myc B</th>
<th>Anthr; Cer A</th>
<th>all other</th>
<th>Thur B</th>
<th>Myc B</th>
<th>all other</th>
<th>Anthracis A</th>
<th>Anthracis B</th>
<th>Cer B</th>
<th>Thur A</th>
<th>Thur B</th>
<th>Myc A</th>
<th>Myc B</th>
<th>Cer B</th>
<th>Cereus A</th>
<th>Cereus B</th>
<th>Thur B</th>
<th>Cereus A</th>
<th>Cereus B</th>
</tr>
</thead>
<tbody>
<tr>
<td>77</td>
<td>+/−</td>
<td>+/−</td>
<td>−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
</tr>
<tr>
<td>90, 92</td>
<td>2+ /−</td>
<td>+/−</td>
<td>2+ /−</td>
<td>+/−</td>
<td>+/−</td>
<td>2+ /−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
</tr>
<tr>
<td>192</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
</tr>
<tr>
<td>200, 208</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
</tr>
<tr>
<td>208</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
<td>+/−</td>
</tr>
</tbody>
</table>

Table 1

Degree of matching between oligonucleotide probes on microarray and the 16S rRNA sequences of reference microorganisms from the B. cereus group.

a) Positions of targeted subgroup-specific differences on 16S rRNA (Fig. 1A).

- “,” denotes perfect matching.
- “+,” “−,” “−,” and “−” denote one, two, and three mismatches, respectively.

Silica column containing nucleic acids was sealed at the bottom with a cap from a 1.5 ml micro-centrifuge tube and preheated in a sand bath at 95 °C for 3 min to remove traces of ethanol. To prepare labeling cocktail, 1220 μL 25 mM sodium phosphate (pH 7.0) was applied into an ampoule containing 1 mg of lissamine–rhodamine B ethylenediamine (LissRhod) (Molecular Probes, Inc., Eugene, OR). The dye was carefully suspended and then centrifuged at 14,000 × g for 30 s to precipitate insoluble components of the dye mixture. Seven hundred and thirty-two microliters of supernatant were mixed with 30 μL of 150 mM 1,10-phenanthroline-HCl (OP) (Fluka, Ronkonkoma, NY) and 30 μL of 15 mM CuSO4, and 132 μL of labeling cocktail was preheated for 30 s at 95 °C. Eighteen microliters of 33 mM H2O2 (Sigma–Aldrich, Milwaukee, WI, cat#: H1009) was added to the cocktail immediately before application of 150 μL of hot cocktail mixed with H2O2 to the preheated mini-column. The mini-column was sealed at the top to prevent evaporation and was incubated in the sand bath for 10 min at 95 °C. The reaction was stopped by removing the column from the sand bath and adding 20 μL of freshly prepared stop solution (2.6 M sodium acetate, pH 5.2 and 70 mM EDTA) and 540 μL of cold 100% ethanol. After 5-min incubation at room temperature, nucleic acids were precipitated on the column by centrifugation for 30 s at 14,000 × g. Excess fluorescent label was removed by successive washing of the column with 500 μL of 75% ethanol (twice) and 100% ethanol (once). For elution of labeled product, the silica in the column was suspended with 45–60 μL of 10 mM sodium carbonate (pH 8.5), the column was sealed from the top, placed into a micro-centrifuge tube, incubated for 2 min at 95 °C and centrifuged for 30 s at 14,000 × g. Elution was performed twice. Yield of labeled nucleic acids was about 20–50 μg.

Thirty-five microliters of hybridization solution containing 1–3 μg fragmented and labeled RNA, 1 M guanidine thiocyanate (GuSCN), 5 mM EDTA, and 40 mM HEPES (pH 7.5) was passed through a 0.22 μm filter to remove particulates, then heated at 95 °C for 3 min and placed on ice. Thirty microliters of the hybridization solution was added to a hybridization chamber (Grace Biolabs, Bend, OR), and the hybridization chamber was affixed to a microarray. The microarray was allowed to hybridize for an appropriate time at room temperature in the dark. After hybridization, the chamber and hybridization solution were removed from the microarray, and the microarray was washed twice for 10 s each with 100 μL washing buffer (0.9 M NaCl, 50 mM sodium phosphate (pH 7.0), 6 mM EDTA, and 1% Tween 20).
After hybridization the microarrays were analyzed in washing buffer with a custom made wide-field-high-aperture fluorescence microscope (GOI, St. Petersburg, Russia) equipped with a cooled CCD camera (Princeton Instruments, Acton, MA), a thermal table, and XY positioners [93,94]. Exposition time during the image capturing was 0.5–2.0 s. Alternatively, microarrays were washed for 2–3 s in H2O to remove washing buffer from the microarray surface and dried with dust-free compressed gas. Dry microarrays were analyzed with CCD camera equipped portable reader (Aurora Photonics, Lake Barrington, IL) with image capturing time 10 s.

All hybridization experiments were performed two to four times. The signal intensity of each gel element on the microarray was calculated with MicroChip Imager software (Aurora Photonics, Lake Barrington, IL) as follows: the software set up an inner frame to define the area of the gel element and a larger outer frame covering the area immediately surrounding the gel element which was defined as the background for that gel element. The signal intensity of the gel element was calculated by subtracting the signal intensity per pixel in the outer frame (i.e. the background) from the signal intensity per pixel in the inner frame (i.e. the gel element) and multiplying the difference by the number of pixels within the inner frame. The average (\( a \)) and average deviation (\( Ad \)) for p/mm ratios were calculated between replicate hybridizations. Average deviation was calculated using equation \( Ad = \left( \frac{\sum |x - a|}{n} \right) / n \), where “\( x \)” is value, “\( a \)” denotes average and “\( n \)” indicates the number of replicate experiments. In our experiments summarized in Tables 3–5, \( Ad \) for p/mm signal ratios did not exceed 10% of the average, unless specifically indicated.

For identification of single nucleotide polymorphisms, signal ratios were calculated based on experimental data and compared to theoretical values. The regression coefficient was calculated using the formula \( y = Ax + k \), where \( x \), \( y \) and \( k \) represent theoretical value, experimental value and error, respectively. Error is random deviation from linear rule. Based on data table, \( A \) was calculated as \( A = \left( \frac{\sum_{i=1}^{N} x_i y_i}{\sum_{i=1}^{N} x_i^2} \right) \), where \( x_i \) and \( y_i \) are values and \( N \) is the size of data table.

3. Results

3.1. Design of subgroup-specific probes

Analysis of the 16S rRNA, 23S rRNA and gyrB gene sequences revealed subgroup-specific “signatures” in each of these three genes. Each “signature” represented a set of single- or double-nucleotide variations (differences) in gene sequence and differentiated one subgroup from all others [85]. We also found that discrimination of subgroups within the \( B. \) cereus group using rRNA sequences would require single base mismatch discrimination (Fig. 1). The microarray contained pairs of probes including one, two or sometimes three subgroup-specific single- or double-nucleotide differences. For example, pair ps1/ps2 was selected for identification of 16S rRNA sequence difference A/G (position 77) that is unique for subgroup Thuringiensis B (Fig. 1A). In this pair probe ps2 targeted nucleotide G that is specific for Thuringiensis B, and probe ps1 recognized nucleotide A that is present in all six other subgroups (Table 1 and Fig. 1A). Another pair, ps3/ps4, targeted two subgroup-specific variations, T/C (position 90) and T/A (position 92). Probe ps4 identified 16S rRNA sequences containing C at position 90 and A at position 92, and probe ps3 corresponded to T at position 90 and T at position 92 (Table 1 and Fig. 1A). In a similar manner, pair ps11/ps12 targeted three subgroup-specific differences, A/C (position 189), T/G (position 200) and G/C (position 208) (Table 1 and Fig. 1A). In addition, some redundancy was included in the probe set: pair ps5/ps6 targeted the same two subgroup-specific differences as pair ps3/ps4, but the targeted sequence for ps5/ps6 was shifted several nucleotides in comparison with pair ps3/ps4.

The microarray (Fig. 2) included three pairs of probes for identification of bacteria from subgroup Thuringiensis B (ps1/ps2, ps3/ps4 and ps5/ps6), and one pair, ps7/ps8, for identification of organisms from subgroups Thuringiensis A and B, Anthracis, and Cereus A and B. However, ps7/ps8 did not identify organisms from subgroups Mycoides A and B (Table 1). Among five pairs for identification of Mycoides B subgroup bacteria, pair (ps9/ps10) targeted variations A/C (position 189) and T/G (position 200); pair ps13/ps14 identified variations T/G (position 200) and G/C (position 208); pair ps15/ps16 recognized variation G/C (position 208), and pair ps19/ps20 detected variations T/C (position 1036) and A/G (position 1045) (Table 1 and Fig. 1A). Finally, pair ps17/ps18, targeting variation C/A at position 1015 discriminated subgroups Anthracis, Cereus A, and Mycoides B from subgroups Cereus B and Thuringiensis A and B (Table 1 and Fig. 1A). The last pair, ps27/ps28, recognized microorganisms of subgroup Mycoides A, discriminating them from organisms from all other subgroups, using A/G (position 133) as a target (Table 1, Fig. 1A). The combined application of these probes provided a basis for distinguishing subgroup Anthracis from all other subgroups except Cereus A.
Fig. 2. Identification of *B. cereus* subgroup reference microorganisms with 16S rRNA oligonucleotide microarray. Bulk RNA from reference microorganisms was isolated, fluorescently labeled with LissRhod, and hybridized with a microarray bearing 20 b oligonucleotides. Positions of the probes and targeted subgroups (in rectangles) are shown in the upper right corner. Reference organisms and their subgroups (in parentheses) are indicated under each hybridization image captured with stationary microscope. For more details about probe design and probe abbreviations, see Table 1.
Most bacteria of subgroup Cereus A revealed strain-specific sequence variations [85], although some isolates in this subgroup had 16S rRNA sequences identical to subgroup Anthracis (Fig. 1A). Organisms within subgroup Cereus A do not have any subgroup-specific differences in their 16S rRNA (Fig. 1A), but do differ from subgroup Anthracis in 23S rRNA sequence (Fig. 1B). Therefore, we designed probe pairs ps21/ps22 and ps23/ps24 (Table 2) for the discrimination of subgroups Anthracis and Cereus A using sequence differences located, respectively, at position 1559 (G/A) and insertion G at positions 1218–1219 in the 23S rRNA molecule (Table 2 and Fig. 1B). As another example of using 23S rRNA subgroup-specific variations for subgroup identification, we also demonstrated discrimination of subgroups Mycoides A and Mycoides B with probe pairs ps29/ps30 and ps31/ps32 that recognize 23S subgroup-specific substitutions AG/GA at positions 346 and 347, and T/A at position 174, respectively (Table 2 and Fig. 1B).

3.2. Identification strategy for subgroups and reference microorganisms

To demonstrate the ability of the microarray to differentiate all seven subgroups of B. cereus group, we selected ten reference organisms, B. anthracis Ames and B. anthracis Sterne (subgroup Anthracis), B. cereus NCTC11143 and B. thuringiensis B8 (subgroup Cereus A), B. cereus T and B. cereus NCTC9620 (subgroup Cereus B), B. mycoides ATCC43881 (subgroup Mycoides A), B. thuringiensis 4Q281 (subgroup Thuringiensis B), B. mycoides ATCC6462 (subgroup Mycoides A), and B. mycoides ATCC10206 (subgroup Mycoides B), whose subgroup affiliations were previously determined in accordance with their 16S rRNA and 23S rRNA gene sequences [85]. Our objectives were to determine if we could discriminate organisms from these closely related subgroups, to determine if we could achieve discrimination of bacteria whose rRNA sequences differed by only one base, and to determine if we could discriminate B. anthracis from all other closely related organisms in the B. cereus group. For these purposes we fabricated 16S rRNA microarrays for unambiguous identification of subgroups Cereus B, Thuringiensis A, Thuringiensis B, Mycoides A, and Mycoides B. This microarray may also identify subgroups Anthracis and Cereus A, and discriminate them from five above-mentioned subgroups, but may not discriminate subgroups Anthracis and Cereus A each from other (Fig. 2 and Table 1). Therefore, our 16S rRNA targeted microarray contained “subgroup Mycoides B-specific” probe Set #1, “Thuringiensis B-specific” probe Set #3, “Thuringiensis A and B-specific” probe ps8, “Mycoides A-specific” probe ps28 and probe pair ps17/ps18 that discriminates subgroups Cereus B, Mycoides A, Thuringiensis A and B from subgroups Anthracis, Cereus A and Mycoides B, respectively (map of microarray in Fig. 2).

3.2.1. Identification of subgroups and reference microorganisms with 16S rRNA gene sequences

3.2.1.1. Identification of subgroup Mycoides B. The 16S rRNA of B. mycoides ATCC10206, related to subgroup Mycoides B [85] (Fig. 1A), forms perfect duplexes with probes ps10, ps12, ps14, ps16 and ps20 (Set #1, “Mycoides B-specific”), with probes ps1, ps3 and ps5.

### Table 2

<table>
<thead>
<tr>
<th>Probe’s target</th>
<th>Anth/Thur A; Myc A,B</th>
<th>Anth/Cer A; Cer B</th>
<th>Myc B</th>
<th>Thur B; Myc A</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. anthracis Ames/Sterne (Anthr)</td>
<td>+/−</td>
<td>+/−</td>
<td>−2/+</td>
<td>−/+</td>
</tr>
<tr>
<td>B. cereus NCTC11143 (Cer A)</td>
<td>−/+</td>
<td>−/+</td>
<td>−2/+</td>
<td>−/+</td>
</tr>
<tr>
<td>B. thuringiensis B8 (Cer A)</td>
<td>−/+</td>
<td>−/+</td>
<td>−2/+</td>
<td>−/+</td>
</tr>
<tr>
<td>B. cereus NCTC9620 (Cer B)</td>
<td>−/+</td>
<td>−/+</td>
<td>−2/+</td>
<td>−/+</td>
</tr>
<tr>
<td>B. thuringiensis 4Q281 (Thur B)</td>
<td>N/A</td>
<td>N/A</td>
<td>−/+</td>
<td>+/−</td>
</tr>
<tr>
<td>B. mycoides ATCC6462 (Myc A)</td>
<td>+/−</td>
<td>N/A</td>
<td>−2/+</td>
<td>+/−</td>
</tr>
<tr>
<td>B. mycoides ATCC10206 (Myc B)</td>
<td>+/−</td>
<td>N/A</td>
<td>−/+</td>
<td>−/+</td>
</tr>
</tbody>
</table>

“+”, denotes perfect matching; “−”, “−2”, and “−3”, denote one, two and three mismatches, respectively; Anthr, Cer A, Cer B, Thur A, Thur B, Myc A and Myc B, correspond to subgroups Anthracis, Cereus A, Cereus B, Thuringiensis A, Thuringiensis B, Mycoides A and Mycoides B.

a Probe’s name.

b Positions of targeted subgroup-specific differences on 23S rRNA (Fig. 1B).
Table 3

<table>
<thead>
<tr>
<th>Probe pairs</th>
<th>B. thuringiensis B/B. thuringiensis 4Q281 (1:1.5)</th>
<th>B. thuringiensis B/B. thuringiensis 4Q281 (1:5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Theoretical</td>
<td>Experimental</td>
</tr>
<tr>
<td>ps3/ps4</td>
<td>4.245</td>
<td>0.653</td>
</tr>
<tr>
<td>ps5/ps6</td>
<td>1.860</td>
<td>0.182</td>
</tr>
<tr>
<td>ps7/ps8/ps18</td>
<td>6.697</td>
<td>0.256</td>
</tr>
</tbody>
</table>

3.2.1.2. Identification of subgroup Thuringiensis B. B. thuringiensis 4Q281, affiliated to subgroup Thuringiensis B [85], forms a perfect match with all probes from Set #3 (“Thuringiensis B-specific”), with probes ps13, ps15 and ps19 from Set #2, and with probes ps8 (“Thuringiensis A, B-specific”), ps18 and ps27 (Fig. 2). Each probe in pairs ps9/ps10 and ps11/ps12 included one additional mismatch with this isolate (Table 1), but also recognized this target as a “not Mycoides B” organism and, despite the additional mismatch, produced pronounced signals (Fig. 2).

3.2.1.3. Identification of subgroup Thuringiensis A. The combination of probe pair ps7/ps8 and probes ps1 through ps6 (Set #4 and “Thuringiensis B-specific” Set #3) (Table 1, Fig. 2) unambiguously identified subgroup Thuringiensis A and discriminated it from subgroup Thuringiensis B. Hybridization of 16S rRNA of B. cereus ATCC43881, which is affiliated with subgroup Thuringiensis A [85] revealed perfect matches with probe ps8 (specific to subgroups Thuringiensis A and B), with probes ps1, ps3 and ps5 (Set #4, “not Thuringiensis B”) and with probe ps27 (Fig. 2). Therefore, hybridization of 16S rRNA from bacteria of the B. cereus group with probes from region 77-92 (Fig. 1A, Table 1) was used to discriminate microorganisms of subgroup Thuringiensis B from bacteria from all other subgroups, and it was especially important for the discrimination between subgroups Thuringiensis A and B.

3.2.1.4. Identification of subgroups Anthracis and Cereus A. B. anthracis Ames (subgroup Anthracis) and B. cereus NCTC11143 (subgroup Cereus A) [85] (Fig. 1A) form perfect matches with probe Set #2 and Set #4, with probes ps7, ps17 and ps27 (Fig. 2 and Table 1).Probe ps17 is specific for subgroups Anthracis, Cereus and with probes ps17 and ps27. At the same time, this organism includes mismatches with probes ps9, ps11, ps13, ps15 and ps19 (Set #2), ps2, ps4 and ps6 (Set #3) and with probes ps18 and ps28 (Table 1, Figs. 1A and 2). For this reason, Mycoides B isolates hybridized with the microarray revealed fluorescent signals from Set #1, Set #4, and probes 17 and 27 that were higher than the signals from Set #2, Set #3, and probes 18 and 28 (Fig. 2). Pair ps7/p8 recognized B. mycoides ATCC10206 as a “not Thuringiensis A, B” organism revealing signal from ps7 stronger than ps8 one, but because both probes have one additional mismatch with this target (Table 1), pair ps7/ps8 signals were too weak and were seen only with significant overexposure (not shown).
A, and Mycoides B, forming perfect duplexes with 16S rRNA from B. anthracis Ames, B. cereus NCTC11143 and B. mycoides ATCC10206 (subgroup Mycoides B), and mismatches with all other reference microorganisms (Fig. 2 and Table 1). In contrast, probe ps18 contains a mismatch (at position 1015) for B. anthracis Ames, B. cereus NCTC11143 and B. mycoides ATCC10206, and is a perfect match with all other references microorganisms (Table 1 and Fig. 1A). Discrimination of B. anthracis Ames and B. cereus NCTC11143 from B. mycoides ATCC10206 was based on a “perfect” signal for probe ps17 (compare with ps18) in combination with “mismatch” signals for probe Set #1 (“Mycoides B-specific”) (Fig. 2 and Table 1).

3.2.1.5. Identification of subgroup Cereus B. Identification of B. cereus T, which is related to subgroup Cereus B [85], was established based on perfect match signals for probes ps7, ps18, ps27, and for probe Sets #2 and #4 (Fig. 2 and Table 1). Specifically, pair ps17/ps18 provided differentiation of subgroup Cereus B from subgroups Anthracis and Cereus A (Fig. 2 and Table 1).

3.2.1.6. Identification of subgroup Mycoides A. B. mycoides ATCC6462 forms perfect duplexes with probes ps28 (“Mycoides A-specific”), ps1, ps3, ps5 (Set #4), ps15 and ps19 (Fig. 2 and Table 1). Each probe in pairs ps7/ps8, ps9/ps10, ps11/ps12, ps13/ps14 and ps17/ps18 forms additional mismatches with this isolate (Table 1) and for this reason revealed weak signals or no signals after hybridization (Fig. 2).

3.2.2. Identification of subgroups and reference microorganisms with 23S rRNA gene sequences

Although subgroups Anthracis and Cereus A could not be differentiated with 16S rRNA sequences (Fig. 2), utilization of subgroup-specific variations in 23S rRNA genes helped to resolve this problem (Fig. 3 A and Table 2). Here we also demonstrated differentiation of isolates from subgroups Mycoides A and Mycoides B as an example of the usefulness of 23S rRNA variations for identification of different subgroups (Fig. 3 B and Table 2).

3.2.2.1. Differentiation of subgroup Anthracis from subgroup Cereus A. Organisms that belong to subgroup Cereus A contain 16S rRNA sequences that are identical to subgroup Anthracis or that differ by strain-specific sequence variation only [85] (Fig. 1A). Thus, we used 23S rRNA sequences to differentiate subgroups Cereus

Table 5
Influence of environmental admixtures on probe signal ratio

<table>
<thead>
<tr>
<th>Probe pairs</th>
<th>K1 pure cells (p/mm)</th>
<th>K2 environmental samples (p/mm)</th>
<th>K1/K2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ps3/ps4</td>
<td>3.100</td>
<td>2.302</td>
<td>1.347</td>
</tr>
<tr>
<td>ps5/ps6</td>
<td>2.719</td>
<td>1.960</td>
<td>1.387</td>
</tr>
<tr>
<td>ps14/ps13</td>
<td>7.404</td>
<td>6.424</td>
<td>1.153</td>
</tr>
<tr>
<td>ps17/ps18</td>
<td>3.462</td>
<td>1.925</td>
<td>1.798</td>
</tr>
<tr>
<td>ps27/ps28</td>
<td>4.088</td>
<td>3.111</td>
<td>1.314</td>
</tr>
<tr>
<td>ps31/ps32</td>
<td>1.974</td>
<td>1.862</td>
<td>1.060</td>
</tr>
</tbody>
</table>

*Logarithmic B. mycoides ATCC10206 cells were used for these experiments. Images were captured with portable reader after 1 h hybridization of 3µg labeled sample with microarray.

Table 4
Probe signal ratio for B. mycoides ATCC10206 sample hybridization with the microarray

<table>
<thead>
<tr>
<th>Probe pairs</th>
<th>Degree of probe matching</th>
<th>Hybridization time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.5 µg a</td>
</tr>
<tr>
<td>ps3/ps4</td>
<td>+/-2</td>
<td>1.53</td>
</tr>
<tr>
<td>ps5/ps6</td>
<td>+/-2</td>
<td>1.37</td>
</tr>
<tr>
<td>ps14/ps13</td>
<td>+/-2</td>
<td>2.80</td>
</tr>
<tr>
<td>ps17/ps18</td>
<td>+/-</td>
<td>1.52</td>
</tr>
<tr>
<td>ps27/ps28</td>
<td>+/-</td>
<td>1.80</td>
</tr>
<tr>
<td>ps31/ps32</td>
<td>+/-</td>
<td>1.49</td>
</tr>
</tbody>
</table>

*a* Amount of labeled sample.
Fig. 3. Discrimination of subgroups Anthracis and Cereus A (A), or Mycoides A and Mycoides B (B) isolates with 23S rRNA probes. Image capture was carried out with stationary microscope for *B. anthracis* Sterne, *B. thuringiensis* B8 and *B. cereus* HER1414, and with portable reader for all other isolates (A and B). Probe signal ratio is shown on the right side of the images. For more details about probe design and probe abbreviations, see Table 2.

A and Anthracis. The 23S rRNA sequences of *B. thuringiensis* B8 and *B. cereus* NCTC11143 differ from *B. anthracis* Ames at three sites, Y/C (594), insertion G (1218–1219) and G/A (1559) (Fig. 1B). We designed two pairs of probes, ps21/ps22 and ps23/ps24 to target site 1559 and insertion 1218–1219, respectively. Probes ps21 and ps23 formed a perfect duplex with the 23S rRNA of subgroup Anthracis but not subgroup Cereus A 23S rRNA (Table 2 and Fig. 1B). Probes ps22 and ps24 provided complementary information; they have a mismatch with subgroup Anthracis 23S rRNA and are complementary with Cereus A 23S rRNA (Table 2 and Fig. 1B).
To validate the discriminative feasibility of 23S rRNA probes ps21/ps22 and ps23/ps24, we used them for testing of reference microorganisms from subgroup Anthracis (B. anthracis Ames and B. anthracis Sterne) and subgroup Cereus A (B. cereus NCTC11143 and B. thuringiensis B8), as well as for determination of subgroup affiliation for organisms B. thuringiensis B8, as well as for determination of subgroup Cereus A (B. cereus NCTC11143 and B. cereus NCTC3329), whose rRNA genes were not sequenced. We analyzed hybridization signals from these two probe pairs only when 16S rRNA targeted probes revealed a pattern typical for organisms of subgroup Anthracis and Cereus A, B. anthracis Ames and B. cereus NCTC11143 revealed this pattern (Fig. 2), as did B. anthracis Sterne, B. thuringiensis B8, B. cereus HER1414 and B. cereus NCTC3329 (Fig. 3A, ps18/ps17 are shown only). 23S rRNA probe pairs ps21/ps22 and ps23/ps24 successfully discriminated reference isolates related to subgroup Cereus A (B. cereus NCTC11143 and B. thuringiensis B8) from reference microorganisms of subgroup Anthracis (B. anthracis Ames and B. anthracis Sterne), demonstrating hybridization signal ratios for these two pairs 0.9, 0.4 (ps21/ps22, Cereus A), 1.8, 2.2 (ps21/ps22, Anthracis), and 0.4, 0.4 (ps23/ps24, Cereus A), 1.7, 1.7 (ps23/ps24, Anthracis), respectively (Fig. 3A). The same two pairs, ps21/ps22 and ps23/ps24, helped to identify isolates B. cereus HER1414 and B. cereus NCTC3329, whose rRNA genes have not yet been sequenced, as belonging to subgroup Cereus A (Fig. 3A).

3.3.1. SNP in the sites of subgroup-specific differences in the isolates with sequenced rRNA

Complete sequencing of the genome of several isolates of B. anthracis and B. cereus indicated that organisms from the B. cereus group may contain at least 10 copies of the RNA operon [78]. Comparative analysis revealed that a SNP is rather rare event in rRNA sequences in general, but, nevertheless, may be found even at some sites of subgroup-specific differences, for example at positions 594 and 1559 of 23S rRNA (Fig. 1B) in B. anthracis AMES, B. cereus ATCC10987 and B. cereus ATCC14579 [85]. Our recent sequencing [85] indicated that B. cereus T, B. thuringiensis 4Q281 and B. cereus NCTC9620 also have a SNP in their 23S rRNA genes at site 1559 with G:A ratios equal to 1.5:1, 1:1 and 1:3.5, respectively (Fig. 4). At the same time, B. anthracis Ames and B. mycoides ATCC10206 have G at this site, and B. thuringiensis B8 has A in this position [85]. We demonstrated the possibility of recognizing polymorphisms in this site by using a pair of probes, ps21/ps22 (Table 2). For these probes the perfect-mismatch (p/mm) ratios for B. cereus T, B. thuringiensis 4Q281 and B. cereus NCTC9620 were intermediate (1.0, 0.72 and 0.64, respectively) between B. anthracis Ames, or B. mycoides 10206 (1.8 and 1.5, respectively) and B. thuringiensis B8 (0.45) (Fig. 4).

3.3.2. Identification and quantification of “polymorphic” sites in mixtures

To demonstrate the ability of our microarray to identify SNP sites in individual bacterial strains we performed hybridizations of mixtures of rRNA samples isolated from pure cultures and compared them with hybridization images of homogeneous rRNA samples from the same isolates. For these purposes we used the 16S rRNA of B. thuringiensis B8 (subgroup Cereus A) and B. thuringiensis 4Q281 (subgroup Thuringiensis B) and a slightly modified 16S rRNA microarray (Fig. 5). B. thuringiensis B8 has a 16S rRNA sequence that is completely identical to B. anthracis and differs from B. thuringiensis 4Q281 by five subgroup-specific differences (Fig. 1A). Both organisms shared hybridization signals with probe Sets #1 and #2, but generated different signals for probe pairs ps3/ps4, ps5/ps6, ps7/ps8 and ps17/ps18. B. thuringiensis B8 forms perfect duplexes with probes ps3, ps5, ps7 and ps17, but B. thuringiensis 4Q281 forms perfect duplexes with probes ps4, ps6, ps8 and ps18 (Fig. 5, Tables 1 and 3). Therefore, we selected probe pairs ps3/ps4, ps5/ps6, ps7/ps8 and ps17/ps18 as reference pairs for testing of the microarray for its ability to differentiate quantitatively these two organisms in mixtures to imitate polymorphism estimation.
### 3.4. Influence of hybridization conditions on mismatch recognition

The time required for hybridization and the amount of labeled sample are important parameters in microbial diagnostics. We estimated changes in perfect-mismatch ratios of hybridization signals for selected probe pairs with the amount of *B. mycoides* ATCC10206 labeled RNA varying from 1 to 5 μg and the hybridization time varying from 5 min to 18 h (Table 4). Images were captured with stationary microscope. For this experiments we selected both 16S and 23S rRNA probe pairs containing one (ps17/ps18, ps27/ps28 and ps31/ps32) or two mismatches (ps3/ps4, ps5/ps6 and ps14/ps13) (Tables 1, 2 and 4). Results indicated that p/m ratio is significantly decreased with decreased hybridization time and amount of labeled sample. However, all six selected pairs demonstrated correct perfect-mismatch recognitions and therefore revealed correct bacteria recognition even after 5 min of hybridization at room temperature and with labeled sample value as small as 2.5 μg. The p/mm ratio was minimal (1.37) for pair ps5/ps6 (Table 4), but nonetheless provided correct recognition assuming that average deviation in this experiment did not exceed 10% of p/mm ratio values between different microarrays.

#### Table 4.

<table>
<thead>
<tr>
<th>Probe Pair</th>
<th>p/m Ratio</th>
<th>p/mm Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>ps21/ps22</td>
<td>1.80</td>
<td>-</td>
</tr>
<tr>
<td><em>B. anthracis</em> Ames (Anthracs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>B. mycoides</em> ATCC10206 (Mycoides B)</td>
<td>1.50</td>
<td>-</td>
</tr>
<tr>
<td><em>B. cereus</em> T (Cereus B)</td>
<td>1.00</td>
<td>1.5</td>
</tr>
<tr>
<td><em>B. thuringiensis</em> 4Q281 (Thuringiensis B)</td>
<td>0.72</td>
<td>1</td>
</tr>
<tr>
<td><em>B. cereus</em> NCTC9620 (Cereus B)</td>
<td>0.64</td>
<td>0.3</td>
</tr>
<tr>
<td><em>B. thuringiensis</em> B8 (Cereus A)</td>
<td>0.45</td>
<td>-</td>
</tr>
</tbody>
</table>

Fig. 4. Identification of single-base polymorphisms in reference microorganisms using hybridization of fluorescently labeled total RNA from *B. cereus* group bacteria to probes targeting the 23S rRNA. Images were captured with stationary microscope. R = G or A.

3.5. Influence of environmental admixtures on p/mm signal ratio

We tested the influence of environmental contaminants on microarray hybridization. To maximize inhibition activity, we did not use pre-filtration to remove organic debris present in the collected samples. Instead, 10 ml of ground water samples collected from the bottom of wells in Waterfall Forest Preserve (Darien, IL)
Fig. 5. Identification of “polymorphic” sites imitated in mixtures. Samples of *B. thuringiensis* B8 and *B. thuringiensis* 4Q281 labeled total RNA were hybridized with microarray separately (A and B) or mixed in proportions 1:2 (C), or 1:5 (D). Positions of the probes and targeted subgroups (in rectangles) are shown on scheme (E). Images were captured with stationary microscope. For probe signal ratio in pairs selected for SNP quantization, see Table 3. Correlation between actual and predicted data is indicated on chart (F). Values of theoretical (x) vs. experimental (y) signal ratios for four selected pairs obtained for two different mixtures, 1:2 and 1:5 (Table 3) determine each data point (blue rhombuses) position on the chart. Approximation line is described by equation $y = Ax + k$, where $A$ and $k$ are regression coefficient and error, respectively. One of the data point obtained for pair ps3/ps4 (rhombus in upper right corner) was not representative due to low hybridization signals. Calculations performed without these data (ps3/ps4, for mixture 1:2) revealed volumes of $A$ to be equal to 0.96. The error values for all other data, where differences between experimental and theoretical ratios did not exceed 10% of the experimental ratio (Table 3), were 0.04 or less.

were mixed with $3.52 \times 10^8$ *B. mycoides* ATCC10206 cells and precipitated by centrifugation. The pellets, which included microbial cells together with debris, had an average volume of 100 $\mu$l which exceeded the volume of pure bacterial cells approximately 20-fold. These “environmental” samples and control pure *B. mycoides* ATCC10206 cells were used for microarray sample preparation in accordance with our standard Silica mini-column labeling-fragmentation procedure and 3 $\mu$g of labeled samples were hybridized for 1 h with the microarray. Hybridization images were captured with portable reader. Results indicated that environmental contamination decreased p/mm pair ratios only slightly, 1.3-fold on average (Table 5), and did not affect bacterial recognition.

3.6. Cross-hybridization of selected probes with non-target bacteria

Subgroup-specific differences in rRNA sequences for bacteria from the *B. cereus* group are few and they are localized. This creates difficulties in the selection of probes specific to individual subgroups of bacteria. For example, the sequences of 2 of the 20 probes on the microarray (ps17 and ps20) (Table 1) selected for identification of *B. cereus* group reference microorganisms also match the 16S rRNA sequences of a number of bacteria that belong to other groups of the genus *Bacillus*. The use of probes targeting the entire *B. cereus* group and for some other groups may resolve this problem. To demonstrate this we used probes ps25 and ps26, respec-
Fig. 6. Identification of microbial groups. Microarray containing probes ps25 and ps26 targeting 16S rRNA of the B. cereus group and the B. subtilis group, respectively, was hybridized with labeled bulk RNA of the reference microorganisms. Probe signal ratios are shown in the far right column. Images were captured with stationary microscope.

<table>
<thead>
<tr>
<th>Microorganism</th>
<th>B. cereus group</th>
<th>B. subtilis group</th>
<th>Probe signal ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. anthracis Ames (Anthracis)</td>
<td></td>
<td></td>
<td>9.5</td>
</tr>
<tr>
<td>B. cereus T (Cereus B)</td>
<td></td>
<td></td>
<td>8.3</td>
</tr>
<tr>
<td>B. mycoides ATCC10206 (Mycoides B)</td>
<td></td>
<td></td>
<td>7.8</td>
</tr>
<tr>
<td>B. thuringiensis 4Q281 (Thuringiensis B)</td>
<td></td>
<td></td>
<td>8.1</td>
</tr>
<tr>
<td>B. subtilis B-459</td>
<td></td>
<td></td>
<td>0.4</td>
</tr>
</tbody>
</table>

4. Discussion

4.1. Molecular typing of B. cereus group as a tool for identification and discrimination of B. anthracis and related microorganisms

A wide spectrum of DNA typing systems developed recently have not been successful in discriminating between B. cereus and B. anthracis, or B. cereus and B. thuringiensis, or B. cereus, B. weihenstephanensis and B. mycoides [18,58,69,70,78]. These systems included 16S and 23S rRNA genes, 16S–23S rRNA ITS and gyrB gene analysis [43,49,55,58,61,95–101], MEE [36,70,102–104], MLST [78,104–109] and AFLP [18,62,67,69,97] among others. These findings resulted in several suggestions to consider B. cereus and B. thuringiensis [102], or these two bacteria together with B. anthracis [36,58,67,70], as one species.

Our recent findings [85] indicated that the B. cereus group could be divided into several genomic groups based on 16S and 23S rRNA sequences, and that genomic group Anthracis, which unifies all strains of B. anthracis, could be unambiguously discriminated from other members of the B. cereus group based on these sequences. This grouping was confirmed by analysis of gyrB gene sequences and indicated that the genomogroups defined by rRNA and gyrB sequences are inconsistent with the seven currently defined species [85]. We found that three of the largest and most widely distributed B. cereus group species (B. cereus, B. thuringiensis and B. mycoides), which had been defined based on morphophysiological criteria [9], actually represent mixtures of microorganisms related to several genetically distinct bacterial genomogroups [85]. We called such groups “subgroups” and at present time have identified seven of them: Anthracis, Cereus A and B, Thuringiensis A and B, and Mycoides A and B [85]. We also performed molecular typing and identified subgroup-specific genomic markers in 16S rRNA, 23S rRNA and gyrB sequences [85]. Our grouping results confirm AFLP-based findings that B. cereus and B. thuringiensis species reveal a high degree of polymorphism, suggesting that these two species may be polyphyletic [69,97]. Our findings as well as many others mentioned above, demonstrated that current taxonomy, which divides the B. cereus group into seven species (B. anthracis, B. cereus, B. thuringiensis, B. mycoides, B. pseudomycoides, B. weihenstephanensis and B. medusa [1–6]) needs serious corrections.

The B. cereus group was divided into seven subgroups based on sequences of three different evolutionarily conserved genes (16S rRNA, 23S rRNA, and gyrB) which were analyzed by two different methods: identification of “subgroup-specific signatures” and phylogenetic tree construction. Phylogenetic analysis was performed for each of these three genes separately with four different methods of phylogenetic tree construction: minimum evolution, neighbor joining, UPGMA, and maximum parsimony [85]. Therefore, we demonstrated the existence of these seven subgroups by several genomic tests, which independently confirmed the presence of seven new genomic taxa within the B. cereus group. Several studies by other laboratories have produced phylogenetic trees for B. cereus group organisms using a variety of target genes and phylogenetic tools. For example, B. cereus group organisms have been compared based on 16S rRNA sequences [97], 16S–23S
Several other targets [17,78,99,104,105,109] have been used to compare [106], seven housekeeping gene fragments [107], and (see Table 6), we found good agreement between our ing at specific strains that were analyzed in common our results were compared to these studies by look-

Branches bearing clades: A; B-II, III, IV; C-I, II and D-

Clades based phylogenetic tree constructed for the concatenated sequences of six chromosomal genes produced four main branches bearing clades: A; B-II, III, IV; C-I, II and D- I, II [106], which contained isolates found in four of our subgroups Anthracis (clade A), Cereus A (clades B-II, B-III and B-IV), Cereus B (clades C-I and C-II), and Thuringiensis B (clades D-I and D-II), respectively. Similarly, an MLST-based tree generated using concatenated sequences of seven housekeeping gene fragments [107] produced clades “Anthracis”, “Cereus I”, “Tolworthi”, “Kurstaki” and “Others”, which included organisms from our subgroups Anthracis, Cereus A, Cereus B, Thuringiensis B and Mycoides A, respectively. In the same manner, phylogenetic grouping based on distance analysis of 16S–23S rRNA ITS sequences [49] revealed groups I, V and VII, which correspond to our subgroups Anthracis/Cereus A, Thuringiensis B, Mycoides A and Mycoides B, respectively, as well as a phylo-
ge"{\text{c}}\text{gnetic tree branch bearing groups II, III and IV which contain isolates related to subgroup Cereus B.}

Our seven subgroups also correspond to some phenotypic properties of B. cereus group organisms. For example, organisms from different subgroups differ in toxin profiles: B. cereus BCRC 17039 (synonym of B. cereus NCTC11143, affiliated to subgroup Cereus A) showed toxin pattern VII; B. cereus BCRC 10603 (synonym of B. cereus ATCC14579, subgroup Cereus B) demonstrated toxin pattern I; and B. mycoides BCRC 10604 (synonym of B. mycoides ATCC6462, subgroup Mycoides A) demonstrated toxin pattern IV [112]. Similarly, several lines of evidence suggest that significant phenotypic differences exist between subgroups Mycoides A and Mycoides B. First, von Stetten et al. [100] developed PCR primers that could discriminate psychrotolerant and mesophilic strains of the B. cereus group based on 16S rDNA sequences. Comparison of their primer sequences with the subgroup specific mutations identified in our previous study [85] revealed that the psychrotolerant primers matched sequences from subgroup Mycoides A and the mesophilic primers matched sequences from subgroup Mycoides B. We also found that the last subgroup, Mycoides B, included B. pseudomycoides [85], which recently was described as a new species [5]. B. pseudomycoides revealed mesophilic phenotype and genetically was rather far from psychrotolerant strains of B. mycoides and B. wei-

Table 6
Subgroup affiliation of type strains and reference microorganisms that are common in the cited papers

<table>
<thead>
<tr>
<th>Subgroups</th>
<th>Strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthracis–Cereus A</td>
<td>All B. anthracis strains</td>
</tr>
<tr>
<td>Cereus A</td>
<td>B. cereus G9241</td>
</tr>
<tr>
<td></td>
<td>B. cereus ATCC10987b; B. cereus WSBC10030;</td>
</tr>
<tr>
<td></td>
<td>B. thuringiensis 97-2b; B. cereus E33Lb; B.</td>
</tr>
<tr>
<td></td>
<td>cereus NCTC11143c; B. cereus IMSNU 12078</td>
</tr>
<tr>
<td>Cereus B</td>
<td>B. cereus ATCC14579a (DSM31T)c; B. cereus</td>
</tr>
<tr>
<td></td>
<td>ATCC11778c (ATCC9634; DMS345); B. thuringi-</td>
</tr>
<tr>
<td></td>
<td>ensis HD1e</td>
</tr>
<tr>
<td>Thuringiensis A</td>
<td>B. thuringiensis ATCC33679d; B. cereus ACTC43881</td>
</tr>
<tr>
<td>Thuringiensis B</td>
<td>B. thuringiensis ATCC10792e (IMSNU 1208;</td>
</tr>
<tr>
<td></td>
<td>DSM2046c); B. cereus AH527; B. thuringi-</td>
</tr>
<tr>
<td></td>
<td>ensis KCTC 1509f</td>
</tr>
<tr>
<td>Mycoides A</td>
<td>All B. weihenstephanensis strains; B. cereus</td>
</tr>
<tr>
<td></td>
<td>AH521c; B. cereus WSBC10204c; B. mycoides</td>
</tr>
<tr>
<td></td>
<td>ATCC6462c (DSM2048T)</td>
</tr>
<tr>
<td>Mycoides B</td>
<td>All B. pseudomycoides strains</td>
</tr>
</tbody>
</table>

a For more detailed information about these strains, see Refs. [85,111].

b Affiliation of these isolates to correspondent subgroups were obtained from analysis of rRNA and gyrB gene sequences extracted from whole genome sequences of these bacteria (AAEK00000000; AE017355; NC_006274).

c Sequences of these isolates perfectly match with sequences of their subgroups.

d Synonyms of strain title are shown in parenthesis.

e Relation of the isolate was established in accordance with its gyrB gene sequence only, GeneBank accession number AF136390.

f Isolate contains substitution C/Y (192) in its 16S rRNA sequence that makes it to be intermediate form between Thuringiensis A and Cereus B [85]. Finally this dilemma may be solved after sequencing of their 23S rRNA and gyrB genes.

rRNA ITS sequences [49,98], and gyrB sequences [61]. In addition, AFLP-based phylogenetic tree construction [18], MLST based on six chromosomal genes [106], seven housekeeping gene fragments [107], and several other targets [17,78,99,104,105,109] have been used to compare B. cereus group organisms. When our results were compared to these studies by looking at specific strains that were analyzed in common (see Table 6), we found good agreement between our seven subgroups and clades found in all of these studies [17,18,61,78,97–99,104–110]. For example, an MLST-based phylogenetic tree constructed for the concatenated sequences of six chromosomal genes produced four main branches bearing clades: A; B-II, III, IV; C-I, II and D-
subgroup Mycoides A, as well as that B. weihenstephanensis represents one of the B. mycoides strains affiliated to subgroup Mycoides A [85]. The B. cereus subgroups may also differ in the presence of certain types of plasmids. For example, extensive analysis of B. cereus group isolates with sequenced rRNA genes indicated that plasmids harboring cry genes have never been found in subgroups Anthracis, Mycoides A and Mycoides B [85], but may be present in all of the other four subgroups (S. Bavykin, unpublished). The fact that our seven subgroups, which are based on rRNA and gyrB sequences, correlate well with groupings based on AFLP, MME and MLST analysis of a wide spectrum of different genes, and the demonstration of subgroup-specific differences in toxin profiles, psychrotolerance, and the ability to harbor some plasmids, suggest that these seven subgroups are not based solely on neutral genomic polymorphisms, but instead reflect differences in both the genotypes and phenotypes of the B. cereus group organisms. Further investigations that we are performing now will indicate the final number of subgroups and will group them in genomospecies, in accordance with genetic principles [79].

Unique chromosomal sequences, as well as anthrax toxin genes, are the well established genetic markers used in genomic methods of B. anthracis identification and discrimination from related organisms of B. cereus group [33,49–53,57,59–62,64–66]. However, recently a number of publications revealed the presence of plasmids with high levels of homology (up to 99.6%) to B. anthracis plasmids pXO1 and pXO2 in a number of non-anthrax isolates from the B. cereus group, such as B. cereus ATCC10987, B. cereus F4810/72, B. cereus G9241, B. thuringiensis HD73, B. thuringiensis 97-27, B. cereus D-17, B. cereus ATCC43881, B. thuringiensis ATCC33679, as well in several isolates from outside the B. cereus group, such as B. circulans CBD118 and Bacillus sp. CBD119 (related to B. luciferences) [17,78,105,113–117]. Therefore, traditional targets used for B. anthracis identification must be supplemented with new genomic markers. From this point of view, multicycoped rRNA genes, as standardized targets commonly used for most bacteria classification and discrimination, represent a highly useful target for emergent B. cereus organism detection. In addition, the presence of 20,000–50,000 ribosomes in one bacterial cell [118] makes rRNA a valuable target for B. cereus organism identification without preliminary PCR amplification. The goal of this study was to confirm the existence and demonstrate the practical usefulness of previously developed genetic markers for identification and discrimination of all seven subgroups of the B. cereus group.

4.2. Discrimination of subgroups in B. cereus group with rRNA probes

A set of 16S/23S rRNA targeted oligonucleotide probes was designed in the current study (Figs. 1 and 2, Tables 1 and 2) and these probes were immobilized within an oligonucleotide microarray. This microarray enabled us for the first time to differentiate all seven subgroups in the B. cereus group from each other (Figs. 2 and 3), including subgroup Anthracis that contains all B. anthracis isolates whose rRNA was sequenced [85] (http://www.ncbi.nlm.nih.gov/Genbank/index.html; http://rdp.cme.msu.edu/index.jsp). We successfully tested our microarray with a set of previously identified [85] reference strains representing all seven subgroups. Using our microarray we also identified subgroup affiliation for B. cereus HER1414 and B. cereus NCTC3329 (Cereus A) (Fig. 3A) and B. mycoides ATCC23258 (Mycoides A) (Fig. 3B), whose rRNA genes have not been sequenced yet.

Previous studies have shown that 16S rRNA sequences of organisms from subgroups Anthracis and Cereus A, e.g., B. anthracis Sterne and B. cereus NCTC11143, have 99.9–100% similarity and do not contain subgroup-specific differences [83,85]. However, we demonstrated that these organisms, as well as other isolates that belong to Anthracis and Cereus A subgroup, may be differentiated using subgroup-specific signatures located in their 23S rRNA sequences (Fig. 3A and Table 2).

4.3. Perfect match–mismatch discrimination

This study also indicated that our 3D gel-based oligonucleotide microarray format [86] was able to discriminate microbial strains which differed by only one base in their rRNA molecules based on hybridization signal ratios between pairs of perfect match and mismatched probes. For example, using pairs ps7/ps8 and ps17/ps18 we discriminated B. cereus NCTC11143 (Cereus A) from B. cereus T (Cereus B) and B. cereus T from B. cereus ATCC43881 (Thuringiensis A) (Fig. 2), strains that differ from each other in their 16S rRNA gene sequences by only one substitution (A/C (1015) and T/C (192), respectively) (Fig. 1A and Table 1). Successful one base mismatch discrimination was also demonstrated for 16S rRNA targeted probe pairs ps1/ps2, ps15/ps16, ps27/ps28 (Table 1 and Fig. 2) and for pairs ps21/ps22, ps23/ps24 and ps31/ps32 related to 23S rRNA (Table 2 and Fig. 3). The ability of our technology to discriminate single base differences suggests that organisms affiliated to the same subgroup,
but differing by one nucleotide strain-specific mutation [85], may be also differentiated with our 3D microarray technology.

4.3.1. Efficacy of perfect–mismatch paired microarray for emergent detection

One of the obstacles for nucleic acid detection with oligonucleotide microarrays is the considerable difference between A-T and G-C pair stability that results in sequence-dependent intensity of fluorescent signals from different oligonucleotide probes of the same length. Due to this sequence dependent hybridization stability, mismatch-forming duplexes can be more stable than perfect duplexes of the same length but with different sequences. Therefore, perfectly matched oligonucleotide probes may produce signals of lower intensity than other probes containing a mismatch with the target. Sequence-dependent hybridization impacts the time required for efficient hybridization of all probes, and as a result most microarray hybridization protocols hybridize for at least several hours [119]. However, microarrays constructed with p/mm pairs in which the mismatch is located in the middle of the probe effectively overcome the problem of sequence dependent hybridization stability because within a p/mm pair a perfectly matched sequence should always produce a duplex that is more stable than one containing mismatch. For those cases in which both members of the p/mm probe pair generate signals of equal intensity, these probes can simply be eliminated from further consideration. The p/mm pair strategy has been widely used in microarray technology for one base mismatch discrimination and as a control for nonspecific hybridization [120,121]. In our case, we used p/mm probe pairs to discriminate single- or double-nucleotide variations in rRNA sequences and differentiate one subgroup from all others. Use of p/mm pairs permitted us to dramatically decrease hybridization time to as little as 5 min while still enabling perfect–mismatch discrimination (Table 4). In addition, we have recently developed a free radical based Silica mini-column method for microarray sample preparation that requires only 1 min for labeling-fragmentation of nucleic acids [92]. Therefore, our methods can produce highly purified end-labeled nucleic acids samples with easy-controlled fragment length, in roughly 15 min starting from whole bacterial cells. In addition, these methods can be utilized for both DNA and RNA and should be easily automated. Therefore, combination of the 3D gel-based p/mm paired microarray with free radical-based method of sample preparation may be effectively used in development of detectors for emergent bacteria identification.

4.3.2. SNP sites quantification

We have demonstrated above that our microarray technology enabled us to discriminate single nucleotide variations in rRNA sequences. At the same time, variation of p/mm ratios in some sites of subgroup-specific differences in different organisms of the same subgroup suggested that several copies of rRNA genes (total number of rRNA operons may vary from 13 to15 in B. cereus group [78]) may contain SNPs. Our recent analysis of the complete genome sequences of B. cereus group isolates confirmed this. SNPs were found at positions 594 and 1559 of 23S rRNA in several isolates [85]. In current study we demonstrated in vivo (Fig. 4) and in artificial mixtures (Fig. 5 and Table 3) that our microarray platform was able to identify the presence of SNPs when single nucleotide substitutions were present in at least one of five copies, with an accuracy of p/mm ratio not smaller then 10% (Table 3). Moreover, based on the data presented in Fig. 5 it is easy to hypothesize that it would be possible to identify individual species within mixed communities using a database of hybridization images of individual B. cereus group organisms.

4.3.3. Bacterial identification in environmental samples

Analysis of nucleic acids in environmental samples generally requires removal of organic matter (e.g., humic acids), debris, etc., because many environmental contaminants can inhibit reactions employed for analysis. Membrane filtration with a pre-filter that accumulates insoluble environmental contaminants helps but does not completely prevent PCR inhibition due to co-elution of compounds such as humic substances, metal ions, etc., that are concentrated on the filter along with targeted bacteria [123]. Our successful testing of our microarray with environmental samples without pre-filtration in presence of organic contaminants (Table 5), which exceeded volume of pure bacterial cells approximately 20-fold, indicated that even at this extreme condition our microarray technology in combination with our previously developed free-radical based silica mini-column sample preparation method [87,92], worked properly and has a significant potential for direct analysis of environmental samples without preliminary cell culturing or PCR amplification.

4.4. Hierarchical approach in microarray construction

In the current study we demonstrated that our microarray unambiguously discriminated B. anthracis from all other members of the B. cereus group based on 16S/23S
Fig. 7. Two strategies for microbial identification. Hierarchically nested probe design and analysis strategy (C and D) and endpoint probe design and analysis strategy (A and B). The hierarchical strategy reduces false positives due to concordance of positive microarray signals (red dots) at predecessor nodes in the phylogenetic tree, resulting in a “branch” analysis (D) instead of an “endpoint” analysis (B) for data interpretation and reporting.

rRNA sequences, and we demonstrated discrimination of seven subgroups within the *B. cereus* group (Anthracis, Cereus A and B, Thuringiensis A and B, and Mycoides A and B) that had previously been identified only theoretically based on “subgroup-specific signatures” [85]. Subgroup discrimination was achieved using probe pairs selected in accordance with these subgroup-specific markers. However, the limited number of subgroup-specific differences between organisms within the *B. cereus* group made it difficult to select subgroup-specific probes that did not match any organisms from closely related, non-*B. cereus* group species. To address this difficulty, the microarray included probes ps25 and ps26, which were specific for the *B. cereus* group and the *B. subtilis* group, respectively, and which successfully discriminated organisms from these two groups (Fig. 6). The combination of group-specific probes ps25 and ps26 with probe pairs specific for the *B. cereus* subgroups represents a hierarchical probe design strategy that offers significant advantages over an “endpoint” strategy that focuses on designing specific probes for each individual target. The “endpoint” approach requires the design of probes that are absolutely unique for each individual target (Fig. 7A). The requirement for absolutely unique probes is labor-consuming and necessitates regular renovation of probe sets in response to newly described organisms. The “endpoint” strategy also generates only “endpoint” information (Fig. 7B). Alternatively, “hierarchical” probe design (Fig. 7C) allows one to concentrate probe selection efforts on the key nodes and, in practice, produces more reliable identification assuming perfect signals in a chain of nodes as a major requirement of correct group identification (Fig. 7D). The “hierarchical” approach also does not necessitate frequent re-design of probes characterizing each node, but in fact calls for doing this rather rarely (in response to considerable
changes in structure of phylogenetic tree only) in comparison with the “endpoint” approach. Therefore, the hierarchical microarray approach we have taken combining probe pairs differentiating closely related species/isolates (“determinative” probe set) with specific (non-paired) probes targeting particular nodes on the target-related branch of the phylogenetic tree (“phylogenetic” probe set) should significantly improve the reliability of microarray-based microbial identification. Recently, we successfully tested a hierarchically organized “phylogenetic” probe set targeting “all life”, eubacteria, B. subtilis group and B. cereus group cells with human, Escherichia coli, B. subtilis and B. thuringiensis targets [87]. From this point of view, further development of hierarchically complicated bacterial phylogenetic trees, where each particular branch consists of dozens of nodes, will lead us to development of highly reliable microarray-based microbial diagnostics. Therefore, combination of high multiplicity of rRNA targets in bacterial cells with simultaneous using of “phylogenetic” and “discriminative” probe sets provide highly sensitive and highly reliable identification of B. anthracis and related B. cereus group organisms. The developed subgroup-specific markers may be useful for B. anthracis and other B. cereus group pathogens identification in terroristic and environmental events, food outbreaks and bacterial control in agriculture, where B. thuringiensis strains are used as a bio-pesticide, as well as in basic studies of molecular mechanisms of toxin actions and their distribution among B. cereus group isolates.

Acknowledgements

We express our gratitude to Gennadiy Yershov; Boris Chernov; Julia Golova; Eduard Timofeev, Irina Taran, Sergey Surzhikov and Ann Gammell for microarray fabrication and Alexey Chernyi for helpful consultations in analysis of rRNA sequences and hybridization images. We are also grateful to John Ezzell, who kindly provided us with reference strains of microorganisms, Wentso Liu and Jodi Flax for their help in the early stages of this project.

This research was supported by NIH grants 5R01AI 059517 and 5R43AI 069627. Defense Advanced Research Project Agency under Interagency Agreement AO-E428 and by the Russian Human Genome Program grant 5/2000.

References


