JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 2005, p. 1662–1668 Vol. 43, No. 4
0095-1137/05/$08.00+0 doi:10.1128/JCM.43.4.1662–1668.2005

Efficient Discrimination within a *Corynebacterium diphtheriae* Epidemic Clonal Group by a Novel Macroarray-Based Method

Igor Mokrousov,* Olga Narvskaya,* Elena Limeschenko, and Anna Vyazovaya

Laboratory of Molecular Microbiology, St. Petersburg Pasteur Institute, St. Petersburg, Russia

Received 26 August 2004/Returned for modification 5 November 2004/Accepted 30 November 2004

A large diphtheria epidemic in the 1990s in Russia and neighboring countries was caused by a clonal group of closely related *Corynebacterium diphtheriae* strains (ribotypes Sankt-Peterburg and Rossija). In the recently published complete genome sequence of *C. diphtheriae* strain NCTC13129, representative of the epidemic clone (A. M. Cerdeño-Tarraga et al., Nucleic Acids Res. 31:6516–6523, 2003), we identified in silico two direct repeat (DR) loci 39 kb downstream and 180 kb upstream of the *oriC* region, consisting of minisatellite (27- to 36-bp) alternating DRs and variable spacers. We designated these loci DRA and DRB, respectively. A reverse-hybridization macroarray-based method has been developed to study polymorphism (the presence or absence of 21 different spacers) in the larger DRB locus. We name it spoligotyping (spacer oligonucleotide typing), analogously to a similar method of *Mycobacterium tuberculosis* genotyping. The method was evaluated with 154 clinical strains of the *C. diphtheriae* epidemic clone from the St. Petersburg area in Russia from 1997 to 2002. By comparison with the international ribotype database (Institut Pasteur, Paris, France), these strains were previously identified as belonging to ribotypes Sankt-Peterburg ($n = 79$) and Rossija ($n = 75$). The 154 strains were subdivided into 34 spoligotypes: 14 unique strains and 20 types shared by 2 to 46 strains; the Hunter Gaston discriminatory index (HGDI) was 0.85. DRB locus-based spoligotyping allows fast and efficient discrimination within the *C. diphtheriae* epidemic clonal group and is applicable to both epidemiological investigations and phylogenetic reconstruction. The results are easy to interpret and can be presented and stored in a user-friendly digital database (Excel file), allowing rapid type determination of new strains.

The diphtheria epidemic in Russia and neighboring countries in the 1990s (140,000 cases, 4,000 deaths in 1991 to 1996 [39]) stimulated research activities on *Corynebacterium diphtheriae*, a causative agent of the disease. A number of the typing methods available at that time (multilocus enzyme electrophoresis [MLE], pulsed-field gel electrophoresis [PFGE], ribotyping, and randomly amplified polymorphic DNA [RAPD] analysis) and newer methods (amplified fragment length polymorphism analysis) were applied for interstrain differentiation of the pathogen (6, 7, 8, 20, 22, 23, 28, 29, 33, 36, 40). These methods allowed the identification of a clonal group of closely related strains responsible for the epidemic in Russia and all other countries of the former Soviet Union and to trace strains exported into other countries (6, 22, 23, 32, 33). These strains were indistinguishable by PFGE, RAPD analysis, and amplified fragment length polymorphism analysis and very similar by ribotyping (there were two principal profiles, "Rossija" and "Sankt-Peterburg," which differed by one band [6, 14, 33]). Minor rare variants were identified by RAPD and ribotyping techniques (22), and a total of 27 types similar by >80% were identified by MLE typing of all strains of this clonal group studied to date (32, 33). However, MLE, PFGE, and ribotyping are time-consuming and rather cumbersome methods, while RAPD analysis lacks interlaboratory reproducibility and hence exchangeability of results. To identify and rapidly monitor subtle changes in the genome structure at an infraclonal level during and between epidemics, fast, simple, portable, and discriminatory molecular typing methods of *C. diphtheriae* are still needed.

Repetitive genome sequences present important sources of intraspecies variation. A new family of such loci (clustered, regularly interspaced, short palindromic repeats [CRISPR]) has recently been identified by in silico analysis of many bacterial species (19). This family is characterized by direct repeats (DR) varying in size from 21 to 37 bp, interspaced by similarly sized nonrepetitive sequences (variable spacers). DR and adjacent variable spacers form direct variant repeats (DVR) (20). The DNA reverse-hybridization method was developed to study variation in the *Mycobacterium tuberculosis* DR locus (the presence or absence of 43 different spacers) by using the macroarray format; this method was named "spoligotyping" (spacer oligonucleotide typing [20]) and has been widely used for epidemiological and phylogenetic purposes (12, 20, 35).

In 2003, a complete genome sequence of the *C. diphtheriae* epidemic strain of biotype gravis ribotype Sankt-Peterburg was published (4). This publication made possible a more thorough, precise, and comprehensive search of candidate polymorphic loci for the development of new typing methods for this pathogen. In the present study, we identified in silico a large DR region in the genome of *C. diphtheriae* and developed a reverse-hybridization macroarray-based method to study its polymorphism. Using this method, we evaluated clinical strains of the *C. diphtheriae* epidemic clone isolated in 1997 to 2002 in the St. Petersburg area in Russia.

* Corresponding author. Mailing address: Pasteur Institute, 14, Mira St., St. Petersburg 197101, Russia. Phone: 7 812 233 21 49. Fax: 7 812 232 92 17. E-mail for Igor Mokrousov: imokrousov@mail.ru. E-mail for Olga Narvskaya: onarvskaya@yahoo.com.

MATERIALS AND METHODS

Bacterial strains. *C. diphtheriae* strains were recovered from diphtheria patients and carriers in the St. Petersburg area in Russia, 1997 to 2002; they were

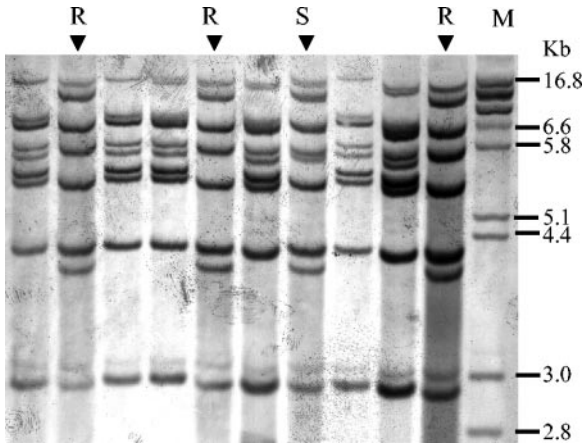


FIG. 1. BstEII riboprofiles of some *C. diphtheriae* strains. Arrowheads indicate ribotypes Sankt-Peterburg (S) and Rossija (R). M, molecular size markers: *Citrobacter koseri* CIP105177 DNA cleaved with MluI.

found to be unlinked by standard epidemiological investigation. Strain identification, biotyping, and toxigenicity determination were performed by standard microbiologic methods (9, 25). DNA was extracted as described previously (34).

Ribotyping. Ribotyping was done as described previously (34). Briefly, bacterial DNA was digested with BstEII, vacuum transferred onto positively charged nylon membranes (Hybond N+; Amersham Biosciences, Buckinghamshire, United Kingdom) and hybridized with a digoxigenin-labeled OligoMix5 (34) rRNA gene-derived hybridization probe. The hybridization profiles were visualized as banding patterns on a membrane with an alkaline phosphatase (Roche Applied Science)-catalyzed colorimetric reaction (Fig. 1). Further, the membranes were scanned and profiles were processed with the TAXOTRON package (15) and stored as a local database.

Identification of the direct repeat loci. A genome search for repeated sequences in the complete genome sequence of *C. diphtheriae* strain NCTC13129 (GenBank accession number NC_002935) was done using Tandem Repeats Finder software (2). Settings used were as follows: alignment parameters (match, mismatch, and indel), +2, −3, and −5, respectively; maximum period size, 100; minimum alignment score, 30. The obtained hits were manually searched for the presence of multiple, short (50- to 80-bp unit size), nonexact (homology, 50 to 80%) repeats. This process identified the location and structure of two regions that corresponded to the definition of the DR and CRISPR loci (19). They are situated downstream and upstream of the origin of replication (*oriC*), and we designated them the DRA and DRB loci, respectively. The DR sequences of these two loci are shown in Fig. 2.

The BLAST nucleotide search engine (www.ncbi.nlm.nih.gov/BLAST) and GeneDoc software (www.psc.edu/biomed/genedoc) were used for sequence searching in GenBank and for sequence alignment, respectively.

Reverse-hybridization spoligotyping assay. Analogously to the spoligotyping method used for *M. tuberculosis* analysis (20), we suggest using the same name for the developed macroarray assay for *C. diphtheriae* subtyping. The specific oligonucleotides (5′-amino labeled) were designed on the basis of the 22 different spacers sequences found in the DRB region in *C. diphtheriae* strain NCTC13129 (Fig. 3a) (for a locus description, see Results). The probes were chosen with OligoDesign software (16) to have similar melting temperatures (Table 1) and were covalently bound to a membrane as described previously (21).

```
gaagtctatcagggtttttgagaagtcaaccccagt---- DRA
---gtct-tctccgcacacgcggaggtatttc-------- DRB
---GTCT-TC---G-----G-G-AG--A---C-------- consensus

gaagtctatcagggtttttg-agaagtcaaccccagt     DRA
gaaatacctccgcgtgtgcggagaagac---------   cDRB
GAA-T---TC-G-GT-T--G-AGAAG-C---------     consensus
```

FIG. 2. DRA and DRB repeat motifs of *C. diphtheriae* and their alignment. cDRB, complementary DRB sequence.

A membrane (Biodyne C membrane; Pall Gelman Laboratory, Ann Arbor, Mich.) was activated by incubation with 16% (wt/vol) 1-ethyl-3-(3-dimethylaminopropyl)carbodiimid (Sigma, St. Louis, Mo.) for 15 min. The oligonucleotides were diluted to the appropriate concentration (Table 1) in 0.5 M $NaHCO_3$, pH 8.4, and applied to the membrane in parallel slots (channels) by using an MN45 miniblotter apparatus (Isogen Biosciences, Maarssen, The Netherlands). After 2 to 3 min of incubation at room temperature, the probes were removed from the slots and the membrane was inactivated with 0.1 M NaOH for 10 min, washed twice in 2× SSPE (0.36 M NaCl, 20 mM NaH2PO4, and 2 mM EDTA, pH 8.0) supplemented with 0.1% sodium dodecyl sulfate (SDS; BDH Laboratory Supplies, Poole, United Kingdom) for 10 min at 58°C, and rinsed for 20 min in 20 mM EDTA at room temperature.

All spacers of the DRB region were amplified with a single primer pair, the reverse primer being 5′-biotin labeled (Fig. 3a). Amplification was performed in a PTC-100 thermal controller (MJ Research, Inc.) with 15 pmol of each primer (forward, CDRF; 5′-CACGCGGAGGTATTTC; reverse, CDRR; biotin-5′-CG TGTGCGGAGAAGAC) in 30 μl of a PCR mixture (1.5 mM $MgCl_2$, 1 U of *Tth* polymerase [Eurobio, Les Ulis, France], and a 200 μM concentration of each deoxynucleoside triphosphate) under the following conditions: initial denaturation 95°C for 3 min; 33 cycles of 94.5°C for 45 s, 53°C for 45 s, and 72°C for 45 s; and a final elongation at 72°C for 3 min. The PCR products were verified in 1.5% agarose gel. A bright, wide, ~65- to 70-bp band was observed in all samples (data not shown); this band represented all primary short sequences of single spacers (Fig. 3a). In some samples, it was accompanied by weaker larger bands containing several DVR, seen as a ladder (data not shown). This PCR result in gel is similar to that usually observed in *M. tuberculosis* spoligotyping. The biotin-labeled PCR fragments of the *C. diphtheriae* DRB region were hybridized to the set of the 22 spacer-derived probes by using the MN45 miniblotter, providing a macroarray format. For this purpose, 25 μl of PCR products was diluted in 150 μl of 2× SSPE–0.1% SDS, denatured for 8 min, and cooled on ice for 10 min. The heat-denatured single-stranded PCR products were applied to the membrane with immobilized probes in the miniblotter slots (perpendicular to the probe lines) and hybridized at 58 to 62°C for 60 min. The membrane was then washed twice with gentle shaking in 100 ml of 2× SSPE–0.5% SDS for 8 min at the same temperature, incubated at 42°C with a 1:4,000 dilution of streptavidin-peroxidase conjugate in 2× SSPE–0.5% SDS for 60 min, washed twice with 100 ml of 2× SSPE–0.5% SDS at 42°C for 10 min, rinsed with 2× SSPE at room temperature for 5 min, and subjected to luminescent detection of hybrids with enhanced-chemiluminescence (ECL) liquid, followed by exposure to the light-sensitive film (ECL Hyperfilm; Amersham Biosciences). After development of the ECL films, the autoradiographs (hybridization profiles) were visually assessed for the presence or absence of signals. For reuse, the membranes were stripped in 1% SDS solution at 80°C (twice for 40 min) and rinsed in 20 mM EDTA, pH 8.0, at room temperature. The membranes were reused up to seven times without reduction of signal strength. Spacer 22, located 1.4 kb downstream of the principal DR region (Fig. 3a), was analyzed in the preliminary experiments. Unlike other spacers, it was amplified as a part of only one PCR fragment, and therefore its hybridization signal was weak. It appeared to be present in all 58 strains studied in initial experiments, and we excluded spacer 22 from further analysis. The obtained hybridization profiles of 21 signals (e.g., Fig. 3b) were entered into a Microsoft Excel spreadsheet using Monotype Sorts police (Table 2). This method allowed the simple schematic presentation of hybridization profiles as black or white boxes.

Different concentrations of the probes (10, 30, 80, and 150 pmol per slot) and hybridization and washing temperatures (58, 60, or 62°C) were initially tested on a set of 16 strains of the *C. diphtheriae* epidemic clone. These variations did not affect the specificity of hybridization signals (invariably these were either present or absent in particular profiles) (Fig. 3b). The spacer sequences were checked against the complete genome sequence of *C. diphtheriae*. No significant homology with other regions was found, and this uniqueness of spacers accounts for high specificity of hybridization signals under the assay conditions used. This finding is similar to those of *M. tuberculosis* spoligotyping, where probes varied in temperature from 59 to 72°C but the hybridization and washing temperature was 60°C and the produced signals were perfectly specific (20). Reproducibility of the method was confirmed by repeating analyses of 40 strains studied with different profiles; no variation in the profile of the same strain was observed in different experiments. A control experiment to determine possible contamination with previously amplified amplicons was performed by including a negative control sample (distilled water) in each PCR and a subsequent hybridization experiment; no contamination was detected (Fig. 3b).

Statistical analysis. The Hunter Gaston discriminatory index (HGDI) was used to evaluate the discriminatory power of the typing method. The HGDI is a probability that two strains consecutively taken from a given population would be

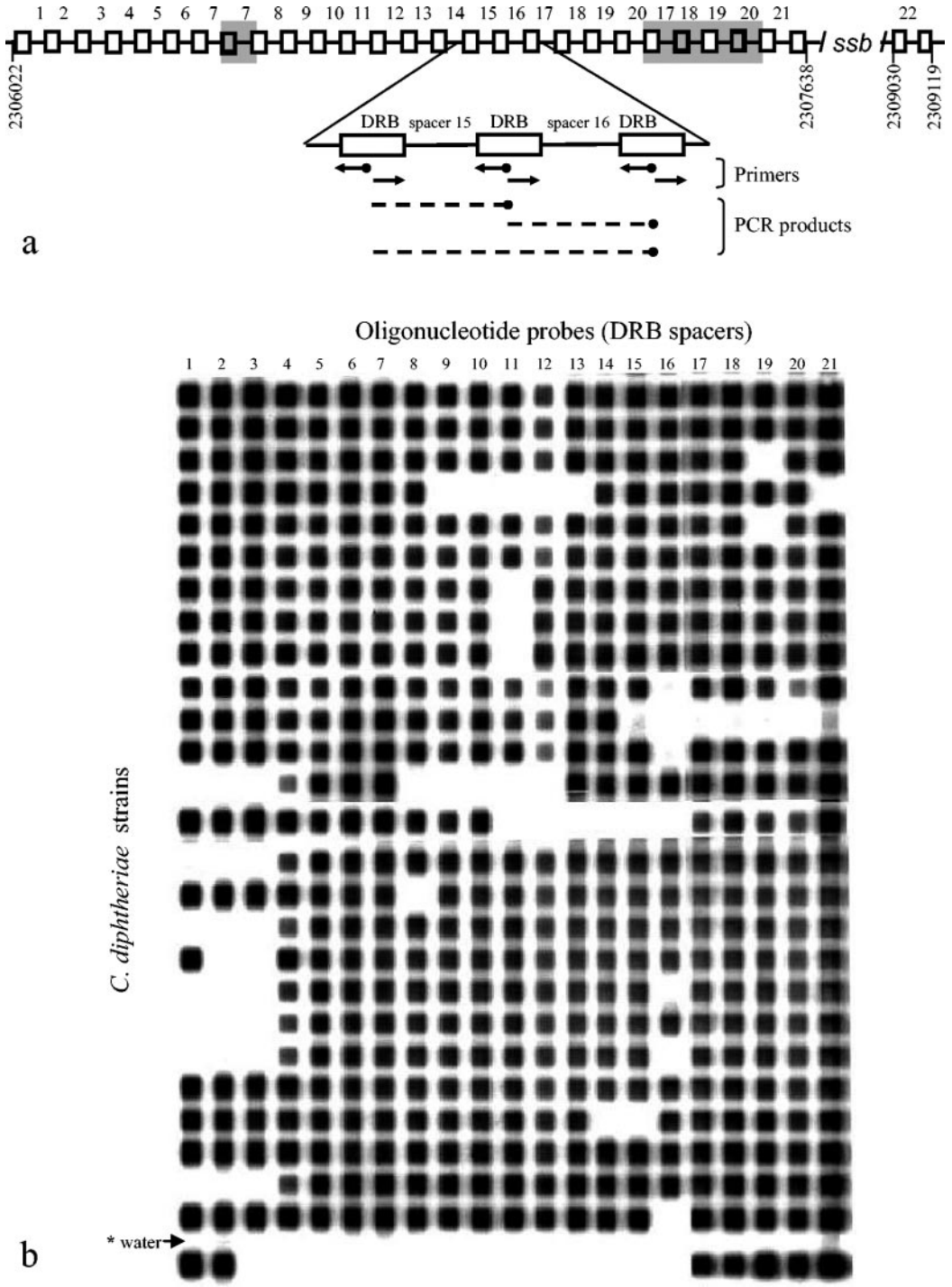


FIG. 3. DRB spoligotyping of *C. diphtheriae* strains. (a) Schematic structure of the DRB locus. Boxes are exact direct repeats; lines indicate variable spacers. Some spacers (and hence DVR) are duplicated (shown as shaded areas). Biotin labels are shown as black dots. (b) Example of reverse hybridization of the membrane containing the 21 different DRB spacer probes with amplified spacers of *C. diphtheriae* clinical strains of the epidemic clone. *, PCR negative control (distilled water).

placed into different types by the typing method; the lower the index value, the less discriminative the typing method. The HGDI was calculated as described previously (18).

Odds ratios were calculated using EpiCalc 2000 version 1.02 software (13) with a 95% confidence interval.

RESULTS

An in silico search of the complete genome sequence of the *C. diphtheriae* biotype gravis epidemic strain NCTC13129 (4) identified the location and structure of two DR (CRISPR) loci.

TABLE 1. Sequences of oligonucleotide probes used in the *C. diphtheriae* spoligotyping assay

DRB spacer	Probe sequence (5′-amino labeled)	Probe concn (pmol/slot)
1	GGTGAGTAAAACTCCGAAAGTTACCG	10
2	GGAAATGTTTCTTAAACTCCCCACCA	10
3	TACCGTGGCGAACCAGTGC	10
4	CCGATGTGACCGCAGCAC	80
5	TCATATCCCTGAAATAAACTGTTCCAGTG	10
6	AGTACCGGTCTGCCATGAGG	10
7	GGTGGATAATAGGCTAGGTGATTCTTGG	10
8	AGGCGATCGCCGGTTCTAT	150
9	GCTCCACAGGGCCAAAAGTTC	30
10	GCCGCTGGAGGAGCTTACTC	80
11	TTTGGTTGTGCAGTCCCCGAA	10
12	GCGGCCTTGTGCAGGTCA	10
13	GAAATCGCTGAACGCCGACTC	30
14	ATGTCCCAAAGCTAAGGTAGCGG	30
15	GCTTGTCGGGCCGCATATG	30
16	GAGGATTCGACGGGTTTCCATTG	30
17	CTTGTGAGGGTCCTCGGATGAC	80
18	GCACCATCCACGGGTGTTATTTG	30
19	CTCAGATTGCGGGGATTCTCCC	80
20	AAAGCGAGGTCGCTGCGC	80
21	CATGCCATCAGACCCGACC	30
22	GTCCTCACCCTCGTAGGACCA	150

For neither of the two repeat motifs (Fig. 2) was a significant homology in GenBank found. The first locus (DRA) is located in the first quadrant of the chromosome (positions 39014 to 39484) and consists of seven units (DVRs); the DR size is 36 bp, while spacer sizes are 27 to 28 bp. This locus is identical to that described previously by Jansen et al. (19), who analyzed an incomplete sequence of the same strain. The second DR locus (DRB) is located in the fourth quadrant of the chromosome (positions 2306022 to 2309119) and consists of 27 DRs and 26 spacers. Additionally, ca. 1.4 kb downstream of these 26 DVRs, there are two DRs separated by one spacer (Fig. 3a). The DRB repeat size is shorter than in the DRA locus (29 bp versus 36 bp), while spacers are 32 to 33 bp in length. The DRA spacers are unique, whereas some of the DRB spacers are duplicated (Fig. 3a, shaded areas), a phenomenon observed also in some *M. tuberculosis* strains (38).

Further detailed analysis of Russian strains of the *C. diphtheriae* epidemic clone was done on a larger and presumably more polymorphic DR locus (DRB) consisting of 27 spacers. Since some spacers are duplicated and spacer 22 produced permanently weak signal (see above), the final number of the unique different spacers targeted in the assay was 21. Design and optimization of the reverse hybridization macroarray-based spoligotyping method is given in Materials and Methods.

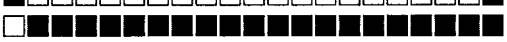
A total of 512 *C. diphtheriae* strains from the St. Petersburg area in Russia isolated from 1997 to 2002 have been analyzed by ribotyping in our laboratory (5, 30; O. Narvskaya et al., unpublished data). Comparison with an international ribotype database established at the Institut Pasteur in Paris (14) identified 257 strains as belonging to the epidemic clone (ribotype Sankt-Peterburg, 142 strains, and ribotype Rossija, 115 strains) (Fig. 1). Of these 257 strains of the epidemic clone, 154 strains (ribotypes Sankt-Peterburg, 79 strains, and Rossija, 75 strains) were randomly selected for evaluation by the DRB-based spoligotyping method. A total of 34 distinct spoligoprofiles were revealed and entered into a Microsoft Excel spreadsheet. Further, these profiles were sorted automatically by using a respective Excel function and were assigned consecutive type numbers (ST used for the spoligotype abbreviation). Type ST1 comprising all signals in the spoligoprofile included strains of both ribotypes of the epidemic clone, while the other 33 spoligotypes included strains of either the Sankt-Peterburg or Rossija ribotype, but not both (Table 2). Twenty spoligoprofiles were shared by 2 to 46 strains, while 14 strains had unique profiles (Table 2). The discriminatory power of the spoligotyping assessed with the HGDI estimate was 0.85 for the entire sample of strains and survey period, 1997 to 2002. Generally speaking, higher diversity is observed within the Sankt-Peterburg than within the Rossija ribotype, which is manifested as a higher number of allelic variants (22 versus 13) (Table 2) and a higher HGDI (0.83 versus 0.71) for the total survey period.

A comparison of genotyping and toxigenicity data showed that, in general, most strains were toxin positive (116 of 154). However, the distribution of the toxin-negative strains varies significantly among two ribotypes and different spoligotypes (Table 2): from only 11.4% of Sankt-Peterburg strains to 38.9% of Rossija strains. Furthermore, toxin-negative ribotype Rossija strains were found mainly within the ST4 type (50%) and ST4 lineage (related types ST4, ST5, ST6, and ST15, 52.1%), unlike other Rossija strains (14.8%).

DISCUSSION

The DR regions consisting of alternating minisatellite repeats and nonrepeats are intriguing loci in bacterial genomes. Their evolutionary history and, especially, biological function remain unclear. Although DR sequences are very dissimilar among different species, a recent in silico analysis identified such loci in many bacterial lineages (19). Previously, van Embden (38) hypothesized that such a locus in *M. tuberculosis* might have initially presented a region consisting of hundreds of short (36-bp) tandem repeats. Variable spacers emerged and accumulated further during evolution, and subsequent changes in the DR locus in *M. tuberculosis* have occurred and are still occurring via consecutive deletions of either single units or contiguous blocks, occasionally including insertion sequence-mediated disruption and recombination (1, 11, 27). Such a scenario reasonably excludes the possibility of a common ancestry for all DR in bacterial evolution and rather suggests their independent emergence in different species and hence a biological function, albeit obscure. The first hypothesis about the role that such loci may play—replicon partitioning—was made on the model of *Haloferax*, *Archaea* (26). Later, based on the analysis of the adjacent genes, Jansen et al. (19) proposed their putative role in DNA metabolism. Generally, the order of single DVR in the *M. tuberculosis* DR locus is strictly conserved (with a possibility of rare duplications [38]), and its changes (deletions of spacers) appear evolutionarily neutral. A large number of variable characters (i.e., particular spacers that may be present or absent in the locus) provide sufficient variation to differentiate clinical strains (12, 20, 38). The robustness of the spoligotyping (DR locus)-based *M. tuberculosis* phylogeny was confirmed by other independent molecular markers (35). To date, DR-based strain typing (spoligotyping) has been used only for *M. tuberculosis* (12, 20); in one

TABLE 2. Schematic presentation of the DRB spoligotype database of the *C. diphtheriae* epidemic-clone strains and their yearly distribution

Ribotype and spoligotype	Spoligoprofile	No. of strains isolated/no. of toxin-negative strains						
		Total	1997	1998	1999	2000	2001	2002
Sankt-Peterburg								
ST1		32/5	4	5	9/3	4	3	7/2
ST2		2/1			2/1			
ST3		4	2	1	1			
ST8		1				1		
ST9		1			1			
ST12		3		1	2			
ST13		1/1					1/1	
ST16		1		1				
ST17		1	1					
ST18		5			2			3
ST19		1						1
ST22		3					2	1
ST23		1			1			
ST24		2				2		
ST27		2			2			
ST28		1						1
ST29		3			2	1		
ST30		1					1	
ST31		7/2			2	4/1	1/1	
ST32		3		1	2			
ST33		3				3		
ST35		1						1
Rossija								
ST1		14/1			8	2	1	3/1
ST4		38/19		4/4	15/2	13/7		6/6
ST5		2/1			1			1/1
ST6		6/3			1	2/1	1/1	2/1
ST7		1/1						1/1
ST10		2/1					2/1	
ST11		3	3					
ST14		2/1				2/1		
ST15		2/2				1/1	1/1	
ST20		1			1			
ST21		1				1		
ST25		1						1
ST26		2				2		

instance, interstrain variation in a similar locus was shown for group A *Streptococcus* strains (17). Here, we identified a large DR locus in another important human pathogen, *C. diphtheriae*, and evaluated its variation for epidemiological subtyping of the *C. diphtheriae* clonal group that caused a severe epidemic in the 1990s in the former Soviet Union countries.

Although only one, if any, DR locus per genome is usually found in different species (19), two such loci are present in the genome of *C. diphtheriae*. Equal spacer sizes within particular loci (32 to 33 bp in the DRB locus and 27 to 28 bp in the DRA locus) suggest that their extensions (the nucleotide accumulations between DRs) occurred at the same pace. Consequently, the longer the spacers are, the earlier their generation started, and hence the DRB locus might have had a longer history. Intriguingly, both loci are situated in the general vicinity of *oriC*, specifically, 39 kb downstream (DRA) and 180 kb upstream (DRB). The sequence alignment of the DRA and DRB sequences (Fig. 2) suggests more similarity for the DRA-complementary DRB pair than for the DRA-DRB pair. Thus, assuming that both DR are orthologous and that DRB is a primary sequence, we speculate that the DRA sequence may have been produced by a large intragenomic rearrangement like X inversion around the *oriC* region (10) during the early evolution of the *C. diphtheriae* genome, although further studies are undoubtedly required to test this hypothesis. Interestingly, an open reading frame encoding a putative SSB protein

is located within the DRB locus, separating its principal (spacers 1 to 21) and minor (spacer 22) subregions (Fig. 3a). The SSB protein is known to be an important component of DNA replication, recombination, and repair, which corroborates the aforementioned hypothesis (19) about the possible role of the CRISPR and DR regions in DNA metabolism.

The target population of our study included toxigenic and nontoxigenic *C. diphtheriae* strains of the biotype gravis epidemic clone (ribotypes Sankt-Peterburg and Rossija). During the diphtheria epidemic from 1990 to 1996 these closely related toxigenic strains were isolated in high proportions (70 to 90%) of patients in all former Soviet Union countries, including Russia (6, 23, 30, 33), Belarus (37), Central Asian countries (32), Georgia (36), and Moldova (5); few strains were identified in other European countries as imported cases (32). Kombarova et al. (23) reported that ribotype Rossija was first identified in their laboratory in a strain isolated in 1987 in the Vladimir Province in central Russia, where the main source of infection was soldiers who had arrived from Soviet Central Asia. Further, a riboprofile very similar to those of the epidemic clone was identified in Pakistan in 1994 (14). On the other hand, strains of ribotypes Sankt-Peterburg and Rossija were identified in 15 to 22% (33) to 28% (23) of Russian *C. diphtheriae* strains before the epidemic (1985 to 1990) and are still circulating in this country (23, 30). It has been suggested that persistent foci of diphtheria in Russia could be a possible source of the epidemic strains since Russia was never totally free of reported cases of diphtheria (39). Reports of persistent endemic foci in the United States (24, 31) and Canada (24) suggest that the circulation of toxigenic strains of *C. diphtheriae* can occur for prolonged periods even in the absence of recognized clinical cases, at least in certain communities. Finally, we suggest that such permanent isolation of *C. diphtheriae* strains of ribotypes Sankt-Peterburg and Rossija in Russia reflects a stable endemicity of this clonal group within this geographic area. A hypothesis about the Central Asian origin of this clone and its importation to Russia by returning military units from Afghanistan between 1979 and 1990 (23) is intriguing but requires experimental confirmation by analysis of representative strain samples from diverse geographic locations, including possible source areas.

The DRB locus polymorphism and spoligotype distribution within two ribotypes of the *C. diphtheriae* epidemic clone allow us to speculate about its evolutionary history. Genetically, these strains were described in many studies as homogeneous and indistinguishable by different DNA-based methods targeting different genome regions. Strikingly, although 34 types were identified in our study by the spoligotyping method, only one primordial and apparently ancestral type, ST1, was shared by strains of both ribotypes (Table 2). Other types, derived from ST1 by successive single- or multiple-deletion events, are confined to one ribotype, not to both (Table 2). This and the above findings taken together confirm the monophyletic origin of the epidemic clone and, at the same time, demonstrate the clear divergence between ribotypes Sankt-Peterburg and Rossija in the survey area. If we assume that the ST1 type is ancestral and that the evolution of the DRB locus is neutral and occurs mainly via successive deletions of either single spacers or contiguous blocks, then profiles with a single disruption (e.g., types ST2, ST3, and ST25) (Table 2) rank at the same evolutionary level and next to the ancestral type, ST1. We define them as one-step types (with ST1 being considered the zero type). Other types with two disruptions (e.g., ST19 and ST21) are defined as two-step types and have probably emerged more recently. Consequently, ribotype Sankt-Peterburg has 16 one-step types and 5 two-step types compared to the 8 one-step types and 4 two-step types of ribotype Rossija (Table 2). It is noteworthy that the sample sizes were almost the same for the two ribotypes in our study (79 versus 75). However, under DRB spoligotyping analysis, compared to ribotype Rossija, ribotype Sankt-Peterburg is characterized by a higher HGDI value, a larger number of types (allelic variants), and a larger number of "older" one-step types. Assuming that more diversity is generated due to a longer evolutionary history, the Sankt-Peterburg ribotype appears to be evolutionarily older than and ancestral to ribotype Rossija, which may have originated from one particular subpopulation (ancestral ST1 type) of presumably already heterogeneous ribotype Sankt-Peterburg strains, followed by subsequent independent evolution of the DRB locus in both ribotype lineages. Comparison with toxin production data reveals an additional line of divergence between ST4 cluster strains (ST4, ST5, ST6, and ST15) and all other spoligotypes. This ST4 cluster includes 48 of 75 Rossija strains and is marked with a significantly higher proportion of toxin-negative strains than other Rossija spoligotypes (52.1% versus 14.8%; 95% confidence interval, 6.25 [1.88 to 20.82]; $P < 10^{-3}$) and Sankt-Peterburg spoligotypes (52.1% versus 11.4%; 95% confidence interval, 8.45 [3.45 to 20.71]; $P < 10^{-6}$). This finding looks unexpected, since DR locus evolution is apparently neutral, unlike that of biologically meaningful toxin production. However, it has recently been shown that the DtxR protein not only regulates the expression of the diphtheria toxin but also binds in an iron-dependent way to operators of many genes scattered throughout the *C. diphtheriae* chromosome (3, 4). Therefore, we feel that further studies of both *tox* and *dtxR* genes and DRB locus-adjacent regions are needed to elucidate their possible functional and evolutionary links.

To sum up, the developed reverse-hybridization macroarray-based method targeting the polymorphic DRB region in the genome of *C. diphtheriae* allows rapid and efficient discrimination of the closely related strains of the epidemic clone and is applicable for both epidemiological investigations and phylogenetic reconstruction. Technically, the method is fast, reproducible, and portable; it is not demanding, since consumables and equipment are relatively inexpensive even in low-income countries, and many strains may be analyzed at a time. Because of the inherently discrete unit composition of the DR locus, the spoligotyping results are easy to interpret and can be presented and stored in a straightforward and user-friendly digital format. The database may be maintained as an Excel file, allowing easy type determination of a new strain by automatic sorting.

ACKNOWLEDGMENTS

We thank E. V. Timofeeva, E. V. Loseva, N. A. Avsyukevich, N. M. Abakumova, V. G. Zhavoronkov, L. A. Lipatova, T. E. Demakova, A. S. Kvetnaya, and D. Bicenko for providing clinical isolates and microbiological data.

We acknowledge partial support from Institut Pasteur, Paris, France.

REFERENCES

1. **Beggs, M. L., K. D. Eisenach, and M. D. Cave.** 2000. Mapping of IS*6110* insertion sites in two epidemic strains of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **38:**2923–2928.
2. **Benson, G.** 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. **27:**573–580.
3. **Canchaya, C., G. Fournous, and H. Brüssow.** 2004. The impact of prophages on bacterial chromosomes. Mol. Microbiol. **53:**9–18.
4. **Cerdeño-Tarraga, A. M., A. Efstratiou, L. G. Dover, M. T. G. Holden, M. Pallen, S. D. Bentley, G. S. Besra, C. Churcher, K. D. James, A. De Zoysa, T. Chillingworth, A. Cronin, L. Dows, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, S. Moule, M. A. Quail, E. Rabbinowitsch, K. M. Rutherford, N. R. Thomson, L. Unwin, S. Whitehead, B. G. Barrel, and J. Parkhill.** 2003. The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. Nucleic Acids Res. **31:**6516–6523.
5. **Damian, M., F. Grimont, O. Narvskaya, M. Straut, M. Surdeanu, R. Cojocaru, I. Mokrousov, A. Diaconescu, C. Andronescu, A. Melnic, L. Mutoi, and P. A. D. Grimont.** 2002. Study of *Corynebacterium diphtheriae* strains isolated in Romania, northwestern Russia and the Republic of Moldova. Res. Microbiol. **153:**99–106.
6. **De Zoysa, A., A. Efstratiou, R. C. George, M. Jahkola, J. Vuopio-Varkila, S. Deshevoi, G. Tseneva, and Y. Rikushin.** 1995. Molecular epidemiology of *Corynebacterium diphtheriae* from northwestern Russia and surrounding countries studied by using ribotyping and pulse-field gel electrophoresis. J. Clin. Microbiol. **33:**1080–1083.
7. **De Zoysa, A., and A. Efstratiou.** 1999. PCR typing of *Corynebacterium diphtheriae* by random amplification of polymorphic DNA. J. Med. Microbiol. **48:**335–340.
8. **De Zoysa, A., and A. Efstratiou.** 2000. Use of amplified fragment length polymorphism for typing *Corynebacterium diphtheriae*. J. Clin. Microbiol. **38:**3843–3845.
9. **Efstratiou, A., and P. A. Maple.** 1994. W.H.O. manual for the laboratory diagnosis of diphtheria. Reference no. ICP-EPI038(C). World Health Organization, Geneva, Switzerland.
10. **Eisen, J. A., J. Heidelberg, O. White, and S.-L. Salzberg.** 2000. Evidence for symmetric chromosomal inversions around the replication origin in bacteria. Genome Biol. **1:**research0011. [Online.]
11. **Filliol, I., C. Sola, and N. Rastogi.** 2000. Detection of a previously unamplified spacer within the DR locus of *Mycobacterium tuberculosis*: epidemiological implications. J. Clin. Microbiol. **38:**1231–1234.
12. **Filliol, I., J. R. Driscoll, D. van Soolingen, B. N. Kreiswirth, K. Kremer, G. Valetudie, D. A. Dang, R. Barlow, D. Banerjee, P. J. Bifani, K. Brudey, A. Cataldi, R. C. Cooksey, D. V. Cousins, J. W. Dale, O. A. Dellagostin, F. Drobniewski, G. Engelmann, S. Ferdinand, D. Gascoyne-Binzi, M. Gordon, M. C. Gutierrez, W. H. Haas, H. Heersma, E. Kassa-Kelembho, M. L. Ho, A. Makristathis, C. Mammina, G. Martin, P. Moström, I. Mokrousov, V. Narbonne, O. Narvskaya, A. Nastasi, S. N. Niobe-Eyangoh, J. W. Pape, V. Rasolofo-Razanamparany, M. Ridell, M. L. Rossetti, F. Stauffer, P. N. Suffys, H. Takiff, J. Texier-Maugein, V. Vincent, J. H. de Waard, C. Sola, and N. Rastogi.** 2003. Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assessed by spoligotyping in an international study. J. Clin. Microbiol. **41:**1963–1970.
13. **Gilman, J., and M. Myatt.** 1998. EpiCalc 2000, version 1.02. Brixton Books, Brixton, United Kingdom.
14. **Grimont, P. A. D., F. Grimont, A. Efstratiou, A. De Zoysa, I. Mazurova, C. Ruckly, M. Lejay-Collin, S. Martin-Delautre, B. Regnault, and European Laboratory Working Group on Diphtheria.** 2004. International nomenclature for *Corynebacterium diphtheriae* ribotypes. Res. Microbiol. **155:**162–166.
15. **Grimont, P. A. D.** 2000. Taxotron package. Taxolab, Institut Pasteur, Paris.
16. **Herold, K. E., and A. Rasooly.** 2003. Oligo Design: a computer program for development of probes for oligonucleotide microarrays. BioTechniques **35:**1216–1221.
17. **Hoe, N., K. Nakashima, D. Grigsby, X. Pan, S. J. Dou, S. Naidich, M. Garcia, E. Kahn, D. Bergmire-Sweat, and J. M. Musser.** 1999. Rapid molecular genetic subtyping of serotype M1 group A *Streptococcus* strains. Emerg. Infect. Dis. **5:**254–263.
18. **Hunter, P. R., and M. A. Gaston.** 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J. Clin. Microbiol. **26:**2465–2466.
19. **Jansen, R., J. D. A. van Embden, W. Gaastra, and L. M. Schouls.** 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. **43:**1565–1575.
20. **Kamerbeek, J., L. Schouls, A. Kolk, M. van Agterveld, D. van Soolingen, S. Kuijper, A. Bunschoten, H. Molhuizen, R. Shaw, M. Goyal, and J. D. A. van Embden.** 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. **35:**907–914.
21. **Kaufhold, A., A. Podbielski, G. Baumgarten, M. Blokpoel, J. Top, and L. Schouls.** 1994. Rapid typing of group A streptococci by the use of DNA amplification and non-radioactive allele-specific oligonucleotide probes. FEMS Microbiol. Lett. **119:**19–25.
22. **Kombarova, S., C. Kim, V. Melnikov, M. Reeves, O. Borisova, I. Mazurova, and T. Popovic.** 2001. Rapid identification of *Corynebacterium diphtheriae* clonal group associated with diphtheria epidemic, Russian Federation. Emerg. Infect. Dis. **7:**133–136.
23. **Kombarova, S., I. K. Mazurova, V. G. Melnikov, N. N. Kostyukova, K. I. Volkovoi, O. Yu. Borisova, T. V. Platonova, and A. Efstratiou.** 2001. Genetic structure of *Corynebacterium diphtheriae* strains isolated in Russia during diphtheria epidemic process of different intensity. Zh. Mikrobiol. Epidemiol. Immunobiol. **N3:**3–8. (In Russian.)
24. **Marston, C. K., F. Jamieson, F. Cahoon, G. Lesiak, A. Golaz, M. Reves, and T. Popovic.** 2001. Persistence of distinct *Corynebacterium diphtheriae* clonal group within two communities in the United States and Canada where diphtheria is endemic. J. Clin. Microbiol. **39:**1586–1590.
25. **Mazurova, I. K., V. Melnikov, and S. Kombarova.** 1995. Manual for laboratory diagnosis of diphtheria infection. Russian Federation State Committee on Sanitary Epidemiologic Surveillance, Moscow, Russia.
26. **Moijica, F. J. M., G. Ferrer, and F. Rodruguez-Valera.** 1995. Long stretches of short tandem repeats are present in the largest replicons of the archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. Mol. Microbiol. **17:**85–93.
27. **Mokrousov, I., O. Narvskaya, E. Limeschenko, T. Otten, and B. Vyshnevskyi.** 2002. Novel IS*6110* insertion sites in the direct repeat locus of *Mycobacterium tuberculosis* clinical strains from the St. Petersburg area of Russia, and evolutionary and epidemiological considerations. J. Clin. Microbiol. **40:**1504–1507.
28. **Mokrousov, I., O. Narvskaya, G. Tseneva, N. Mikhailov, and I. Popel.** 1996. Genetic typing of *Corynebacterium diphtheriae* strains by the polymerase chain reaction with universal primers. Zh. Mikrobiol. Epidemiol. Immunobiol. **N5:**73–75. (n Russian.)
29. **Nakao, H., and T. Popovic.** 1998. Use of random amplified polymorphic DNA for rapid molecular subtyping of *Corynebacterium diphtheriae*. Diagn. Microbiol. Infect. Dis. **30:**167–172.
30. **Narvskaya, O., I. Mokrousov, N. Abakumova, N. Avsyukevich, E. Timofeeva, L. Loseva, A. Kvetnaya, and D. Bicenko.** 2000. Molecular epidemiology of *Corynebacterium diphtheriae* infection in St. Petersburg: a six-year study, p. 58–59. *In* Proceedings of the Sixth International Meeting of the European Laboratory Working Group on Diphtheria. European Commission, Brussels, Belgium.
31. **Popovic, T., C. Kim, J. Reiss, M. Reeves, H. Nakao, and A. Golaz.** 1999. Use of molecular subtyping to document long-term persistence of *Corynebacterium diphtheriae* in South Dakota. J. Clin. Microbiol. **37:**1092–1099.
32. **Popovic, T., I. K. Mazurova, A. Efstratiou, J. Vuopio-Varkila, M. W. Reeves, A. De Zoysa, T. Glushkevich, and P. Grimont.** 2000. Molecular epidemiology of diphtheria. J. Infect. Dis. **181:**S168–S177.
33. **Popovic, T., S. Y. Kombarova, M. W. Reeves, H. Nakao, I. K. Mazurova, M. Wharton, I. K. Wachsmuth, and J. D. Wenger.** 1996. Molecular epidemiology of diphtheria in Russia, 1985–1994. J. Infect. Dis. **174:**1064–1072.
34. **Regnault, B., F. Grimont, and P. A. Grimont.** 1997. Universal ribotyping method using a chemically labelled oligonucleotide probe mixture. Res. Microbiol. **148:**649–659.
35. **Sola, C., I. Filliol, E. Legrand, I. Mokrousov, and N. Rastogi.** 2001. *Mycobacterium tuberculosis* phylogeny reconstruction based on combined numerical analysis with IS*1081*, IS*6110*, VNTR and DR-based spoligotyping suggests the existence of two new phylogeographical clades. J. Mol. Evol. **53:**680–689.
36. **Sulakvelidze, A., M. Kekelidze, T. Gomelauri, Y. Deng, N. Khetsuriani, K. Kobaidze, A. De Zoysa, A. Efstratiou, J. G. Morris, Jr., and P. Imnadze.** 1999. Diphtheria in the Republic of Georgia: use of molecular typing techniques for characterization of *Corynebacterium diphtheriae* strains. J. Clin. Microbiol. **37:**3265–3270.
37. **Titov, L., V. Kolodkina, A. Dronina, F. Grimont, P. A. D. Grimont, M. Lejay-Collin, A. De Zoysa, C. Andronescu, A. Diaconescu, B. Marin, and A. Efstratiou.** 2003. Genotypic and phenotypic characteristics of *Corynebacterium diphtheriae* strains isolated from patients in Belarus during an epidemic period. J. Clin. Microbiol. **41:**1285–1288.
38. **van Embden, J. D. A., T. van Gorkom, K. Kremer, T. Jansen, B. A. M. van der Zeijst, and L. M. Schouls.** 2000. Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. J. Bacteriol. **182:**2393–2401.
39. **Vitek, C. R., and M. Wharton.** 1998. Diphtheria in the former Soviet Union: reemergence of a pandemic disease. Emerg. Infect. Dis. **4:**539–549.
40. **von Hunolstein, C., G. Alfarone, F. Scopetti, M. Pataracchia, R. La Valle, F. Franchi, L. Pacciani, A. Manera, A. Giammanco, S. Farinelli, K. Engler, A. De Zoysa, and A. Efstratiou.** 2003. Molecular epidemiology and characteristics of *Corynebacterium diphtheriae* and *Corynebacterium ulcerans* strains isolated in Italy during the 1990s. J. Med. Microbiol. **52:**181–188.

JOURNAL OF CLINICAL MICROBIOLOGY, July 2005, p. 3590 Vol. 43, No. 7
0095-1137/05/$08.00+0 doi:10.1128/JCM.43.7.3590.2005

ERRATUM

Efficient Discrimination within a *Corynebacterium diphtheriae* Epidemic Clonal Group by a Novel Macroarray-Based Method (print version)

Igor Mokrousov, Olga Narvskaya, Elena Limeschenko, and Anna Vyazovaya

Laboratory of Molecular Microbiology, St. Petersburg Pasteur Institute, St. Petersburg, Russia

Volume 43, no. 4, p. 1662–1668, 2005. Page 1662, corresponding author footnote, line 4: "imikrousov@mail.ru" should read "imokrousov@mail.ru."