Chapter 17
Corynebacterium diphtheriae

Igor Mokrousov

17.1 Introduction

Diphtheria is a rare disease caused by toxigenic strains of *Corynebacterium diphtheriae* and, less often, *Corynebacterium ulcerans*. The *Corynebacterium* species is a rod-shaped bacterium having a high GC content and classified into *Actinomycetales*, an order of Gram-positive bacteria, containing *Mycobacterium tuberculosis* and *Streptomyces coelicolor*. *C. diphtheriae* species comprises four biotypes, var. *gravis*, var. *mitis*, var. *intermedius*, and var. *belfanti*. All biotypes, with the exception of the biotype *belfanti*, may produce the lethal diphtheria exotoxin. After infection, *C. diphtheriae* can colonize the skin and/or the upper respiratory tract where it releases the toxin, causing the symptoms of the disease (reviewed in ref. [1]).

Substantial endemic circulation of toxigenic *C. diphtheriae* is observed in Africa, the Eastern Mediterranean, South America, Southeast Asia, and the Indian subcontinent as manifested by outbreaks or large numbers of reported cases [1, 2]. Importation of the microorganism from these endemic regions poses a constant threat, particularly among subgroups of individuals with low vaccination levels. At the same time, persistent foci of diphtheria exist in developed countries eventually causing small outbreaks [3, 4]. The reemergence of diphtheria in the newly independent states (NIS) of the former Soviet Union demonstrated the continued threat of this thought to be rare disease. Although following mass immunization campaigns and additional control measures, this epidemic is under control, the currently observed reemergence of toxigenic genotypic variants and circulation of invasive nontoxigenic strains appear alarming. These reasons highlight the importance of understanding of population structure of this pathogen and development of rapid, reproducible, and discriminatory typing techniques for epidemiological

I. Mokrousov (✉)
Laboratory of Molecular Microbiology, St. Petersburg Pasteur Institute,
14 Mira street, St. Petersburg 197101, Russia
e-mail: imokrousov@mail.ru; igormokrousov@yahoo.com

I. de Filippis and M.L. McKee (eds.), *Molecular Typing in Bacterial Infections*, Infectious Disease, DOI 10.1007/978-1-62703-185-1_17,

surveillance and global monitoring of *C. diphtheriae*. The prospective genotyping by classical DNA fingerprinting methods (MLEE [multilocus enzyme electrophoresis], PFGE [pulsed-field gel electrophoresis], and ribotyping) of the circulating isolates permitted to monitor the evolution of the Russian epidemic clone in the 1990s. However, some of these methods are time-consuming and rather cumbersome (MLEE, PFGE, and ribotyping), while others lack interlaboratory reproducibility and hence exchangeability of results (RAPD [randomly amplified polymorphic DNA]). To rapidly identify and monitor subtle changes in the genome structure at an intraclonal level during and between epidemics, fast, simple, portable, and discriminatory molecular typing methods of *C. diphtheriae* are still needed. In 2003, the first complete genome sequence of *C. diphtheriae* was published [5] that became a milestone achievement in genome research of this important human pathogen and offered new possibilities in search for new polymorphic markers for *C. diphtheriae* strain typing.

17.2 Methods

In the past, epidemiologic surveillance of diphtheria was limited by traditional typing systems, such as serotyping, phage typing, and bacteriocin typing ([2] and references therein) with low discriminatory power and insufficient reproducibility. In the 1990s, faced to the NIS/Russian epidemics, several laboratories in Europe and the USA successfully applied new molecular techniques, already developed for other bacterial species, to epidemiologic studies of diphtheria strains, especially those circulating in the NIS and neighboring countries.

17.2.1 Ribotyping

Ribotyping had previously been shown to be an extremely useful tool for DNA profiling of many bacterial species [6] and has been recognized as a straightforward method for typing *C. diphtheriae* isolates [2, 7]. The basic principle of this method is Southern transfer and hybridization of the digested chromosomal DNA with specific rRNA genes derived probes. In many bacteria rRNA operons are present in several copies in genome and the resulting multi-band patterns are frequently strain-specific. The number of fragments generated by ribotyping is a reflection of the multiplicity of rRNA operons present in a bacterial species. Initially, the entire *rrn* operons, e.g., those of *Escherichia coli* or *Bacillus subtilis*, were cloned into plasmid and used as hybridization probes. At present, a mixture of five oligonucleotides representing phylogenetically conserved regions in 16S and 23S rRNA genes (OligoMix5) is used for this purpose making the method a really universal approach for (eu)bacterial strain typing [8]; example of *C. diphtheriae* riboprofiles is shown in Fig. 17.1a. At the same time, recent in silico analyses showed that resolved DNA

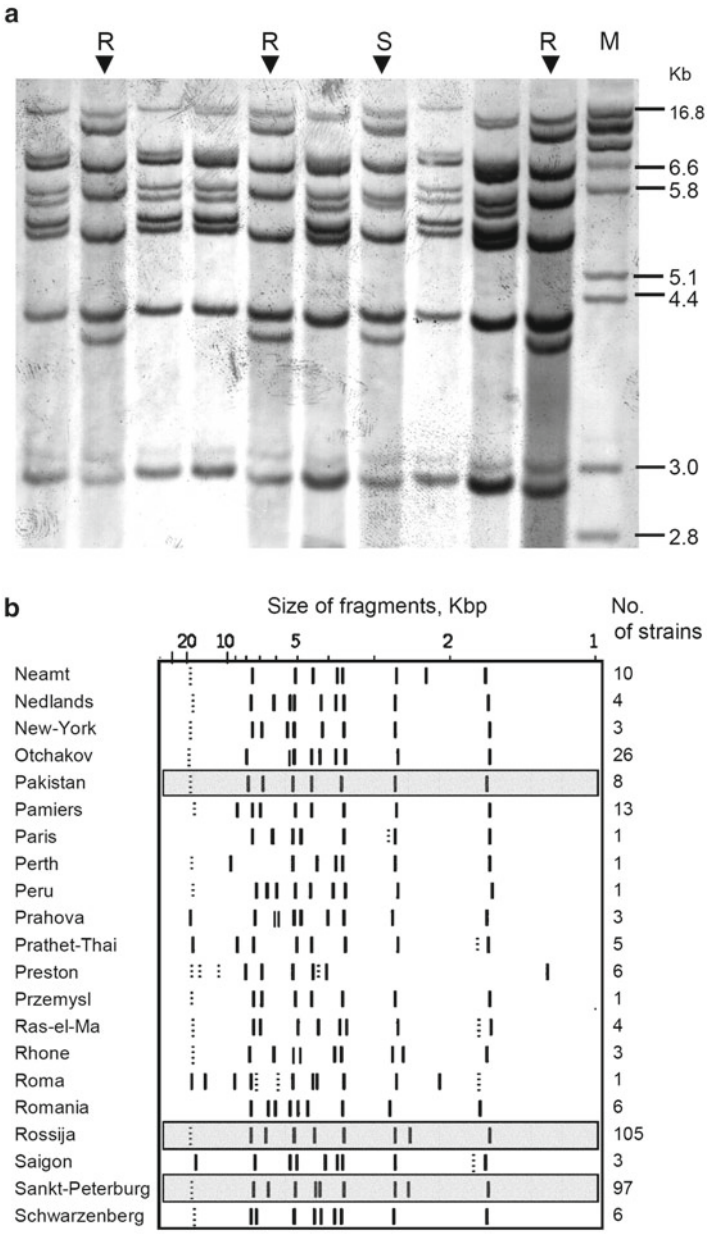

Fig. 17.1 *Corynebacterium diphtheriae* ribotyping. (**a**) *Bst*EII riboprofiles of some *C. diphtheriae* strains. *Arrows* indicate ribotypes Sankt-Peterburg (S) and Rossija (R) of the Russian epidemic clonal group. M: molecular weights marker, *Citrobacter koseri* CIP105177 DNA cleaved with *Mlu*I. (**b**) Extract of the international ribotype database in Institut Pasteur, Paris [10]: a schematic view obtained after computer processing of the profiles. Ribotypes Rossija, Sankt-Peterburg, and likely related Pakistan are gray-shaded. Reprinted from ref. [23] by permission of Elsevier ©2009

polymorphisms rather reflect restriction fragment length polymorphisms (RFLPs) of the neutrally evolving housekeeping genes typically found to flank chromosomal rRNA gene sequences [9]; this, however, does not reduce but rather highlights the phylogenetic value of ribotyping.

Detailed technical description of the method is given by Regnault et al. [8]. Some practical issues must be mentioned. First, the hybridization buffer may be either in-house traditional solution (SSC, SDS, sarkosyl, and blocking reagent [8]) or commercially available ready to use DIG EasyHyb (Roche). The latter permits using reduced temperatures for hybridization (42°C); however, it is rather sensitive to storing conditions and may generate strong background when expired. Second, five oligonucleotides included in the OligoMix5 were initially labeled at 3′-end with digoxigenin (DIG) using labeling kit. At present, it seems more practical and reliable to use commercially DIG-labeled oligonucleotides. Regarding the probe itself, OligoMix5 is a recommended one although previously used complete rDNA operon probe was shown to produce comparable results [8]. Third, a choice of restriction enzyme is a critical issue for development of the discriminatory and reliable scheme. *Pvu*II and *Bst*EII were shown to generate similar discrimination of *C. diphtheriae* strains, but *Bst*EII was ultimately retained since it generated much better interpretable patterns presenting a wide range of fragment sizes.

The hybridization profiles may be visualized as banding patterns on a membrane with an alkaline phosphatase (Roche Applied Science)-catalyzed colorimetric reaction (Fig. 17.1a); *Citrobacter koseri* CIP105177 DNA cleaved with *Mlu*I is used as molecular weights marker. Further, the membranes are scanned and profiles can be processed with TAXOTRON (Institute Pasteur, Paris) or Bionumerics (Applied Maths, Belgium) packages. One should always keep in mind a problem of comparison of banding profiles obtained in different gel runs and laboratories, but this is inherent to all analyzes dealing with banding profiles.

Efficient surveillance of the circulating *C. diphtheriae* variants would not be possible without international ribotype database that makes a good example of the long-term concerted efforts of many collaborating laboratories [10] within the frames of the WHO supported European Laboratory Working Group on Diphtheria (ELWGD) and the European Commission DIPNET project (http://www.dipnet.org). The nomenclature of *C. diphtheriae* ribotypes was published in 2004: a total of 86 ribotypes obtained after *Bst*EII digestion and hybiridization to OligoMix5 were given a geographic name chosen to reflect the place where one of the strains was isolated or studied [10] (Fig. 17.1b). Two patterns are considered identical when they are composed of the same number of fragments and, for homologous fragments, size differences are below a 5% threshold error value. In each reference pattern, each fragment size is calculated as the average of sizes of homologous fragments observed in the corresponding cluster. This numeric approach is not intended for phylogenetic inferences or taxonomic definitions but only to identify identical or similar patterns. The ELWGD principles for naming *C. diphtheriae* ribotypes were agreed as follows: (1) a ribotype name

should be unique; (2) each ribotype name is associated with a reference strain made available from a culture collection affiliated with the World Federation of Culture Collections; (3) once validated, a *C. diphtheriae* ribotype is given a geographic name taken from the place where the strain was isolated; (4) names are labels only and do not imply that a ribotype originated in the area used for naming it [10].

A traditional, manual ribotyping performed in many laboratories is labor-intensive and time-consuming. To solve these problems, an automated ribotyping has been established using RiboPrinter (Dupont Qualicon) automated ribotyping system that became a technological breakthrough with respect to convenience, reproducibility, and speed. At the same time, the higher speed inherently results in the relatively shorter-run of agarose gel format thus reducing quality of bands separation and discrimination. In this sense, traditional larger (16-cm length) gels generate much better resolution. In addition, the cost of ribotyping is significantly greater using the automated system.

17.2.2 Pulsed-Field Gel Electrophoresis

In this classical DNA fingerprinting method the genomic DNA is prepared to remain intact using lysis (lysozyme, RNase, and proteinase K) of bacterial cells directly in agarose blocks, followed by deproteination in proteolysis buffer. For *C. diphtheriae* typing, the DNA is cleaved with rare-cutting enzyme *Sfi*I and PFGE is carried out in TBE 1.5% agarose gels at 14°C by using a CHEF DRII system (Bio-Rad) with pulse times 5–20 s over 20 h and 1–5 s over 18 h. A lambda DNA concatemer (Bio-Rad) is used as a molecular size marker. The bands are visualized by staining the gel in ethidium bromide. Difference of three or more bands is used to distinguish PFGE types. *C. diphtheriae* PFGE typing can be performed according to the procedure described by De Zoysa et al. [7].

In principle, PFGE is highly reproducible and produces well-resolved fragments representing the entire bacterial chromosome in a single gel. It has been assumed to be the most discriminating of the currently available genotypic methods. However, in reality, PFGE was reported to be less discriminating than ribotyping when applied to isolates of *C. diphtheriae*: PFGE was not able to distinguish between the two main ribotype patterns in Russia [7, 11]. Furthermore, unlike conventional ribotyping, the molecular genetic basis for a detected PFGE profile is inherently imprecise for a number of reasons: (1) restriction endonuclease sites are unpredictably scattered throughout the chromosome and (2) unlike ribotyping, detected polymorphisms may involve any functional category of nonneutral genes, including those under the pressure of diversifying antigenic selection [9]. Although PFGE results can address the question "identical or not?," indexing the degree of identity between any two isolates with variant PFGE profiles remains uncertain.

17.2.3 Multilocus Enzyme Electrophoresis

The other important method used for *C. diphtheriae* strain typing is MLEE. This method detects amino acid substitutions affecting charge and conformation in cellular housekeeping enzymes. Such mobility variants, or electromorphs, of the same enzyme can be visualized by their activity in a starch gel matrix as bands of different migration rates. Each electromorph is considered to represent a distinct allele of that enzyme. A profile of electromorphs that define the electrophoretic type (ET) of each strain is obtained by testing 27 different enzymes. Finally, on the basis of these data, the genetic distance between the strains is calculated and presented as a dendrogram [12]. Although MLEE is a powerful and valid technique, it is not ideal, since comparisons of the results obtained in different laboratories are problematic, and the relationship between nucleotide sequence variation and isoenzyme variation typically is unknown.

17.2.4 PCR-Based Genotyping Methods for Rapid Screening

17.2.4.1 RAPD Typing

PCR with random (also named arbitrary or universal) primers has become a very popular approach for typing many pro- and eukaryotic species since 1990. In this polymerase chain reaction-based method, palindromic DNA structures are amplified from different genome regions using a single nonspecific primer. Different primers target different genome regions and may disclose different degree of interstrain genetic variation (see examples of RAPD profiles in Fig. 17.2). The use of crude DNA preparations of *C. diphtheriae* cultures was shown to result in poor amplification and the RAPD patterns were not reproducible. Furthermore, different thermal cycler models produced different RAPD patterns from the same DNA sample. However the reproducibility of the technique was good when the same thermal cycler was used throughout [13]. When applied to the Russian epidemic clonal group, RAPD had the lowest discriminatory power compared to MLEE and ribotyping [14]. At the same time, this discriminatory "inability" implied a possibility to use RAPD typing for preliminary screening of *C. diphtheriae* isolates and rapid and unambiguous identification of the epidemic clone [14–17].

17.2.4.2 AFLP Typing

Another method also widely used for typing bacterial pathogens and subsequently applied to *C. diphtheriae* is amplified fragment length polymorphism (AFLP) analysis that is based on the selective PCR amplification of genomic restriction fragments

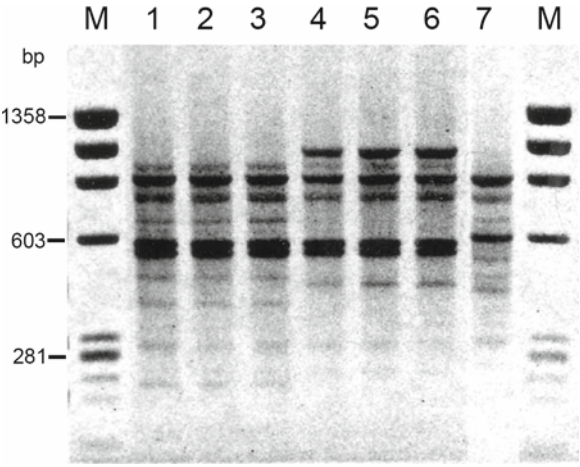

Fig. 17.2 Example of RAPD profiles of *Corynebacterium diphtheriae* strains from St. Petersburg, Russia. M: molecular weight marker phiX174/*Hae*III

of the whole genome [18]. Classically, the technique involved three steps: (1) restriction of the DNA and ligation of oligonucleotide adapters, (2) selective amplification of sets of restriction fragments, and (3) gel analysis of the amplified fragments. PCR amplification of restriction fragments is achieved by using the adapter and restriction site sequence as target sites for primer annealing. The selective amplification is achieved by the use of primers that extend into the restriction fragments, amplifying only those fragments in which the primer extensions match the nucleotides flanking the restriction sites. Using this method, sets of restriction fragments may be visualized by PCR without knowledge of nucleotide sequence. The method allows the specific co-amplification of high numbers of restriction fragments. The number of fragments that can be analyzed simultaneously, however, is dependent on the resolution of the detection system. Typically 50–100 restriction fragments are amplified and detected on denaturing polyacrylamide gels.

De Zoysa and Efstratiou [19] used a simplified version of that technique, utilizing a one-step digestion-ligation reaction with one enzyme, and the PCR with a single primer; consequently, a relatively small number of amplified bands could be separated by agarose gel electrophoresis. Strains with ribotypes Sankt-Peterburg and Rossija (epidemic clone) and ribotype Pakistan could not be distinguished; however, the technique discriminated isolates of ribotypes Vladimir, Lyon, and Otchakov. The reproducibility of the method was examined by using two different thermal cyclers and duplicate AFLP runs for each isolate with two separate DNA extractions. Under all these different conditions, the fragments for each AFLP profile were identical, although variations in the intensities of some of the bands were observed with different PCR runs. The technique analyzes the whole genome, requires only a small amount of DNA, and requires no prior sequence information about the target DNA.

AFLP was suggested to have the potential to replace ribotyping as the "gold standard" within the ELWGD [19].

17.2.4.3 New Generation Molecular Markers

Publication of the first complete genome sequence of *C. diphtheriae* strain NCTC13129 [5], along with use of high-throughput 454/Solexa technologies for partial genome sequencing [20] have greatly accelerated the development of new typing approaches for *C. diphtheriae*.

17.2.4.4 Multilocus Sequence Typing

Multilocus sequence typing (MLST) is proposed as a procedure for characterizing isolates of bacterial species using the sequences of internal fragments of seven housekeeping genes. MLST is based on the well-established principles of MLEE, but differs in that it assigns alleles at multiple housekeeping loci directly by DNA sequencing, rather than indirectly via the electrophoretic mobility of their gene products (http://www.mlst.net). Several criteria are used in the selection of all potential loci. Genes included are those encoding putative housekeeping products necessary for biological roles in DNA repair, replication, and amino acid biosynthesis. Genes that are either located near or implicated as being putative virulence factors and mobile elements should be avoided, since these may come under greater selective evolutionary pressures than other genes. The selected loci should be distributed as much as possible across the chromosome to ensure that each locus was genetically unlinked. For each housekeeping gene, the different sequences present within a bacterial species are assigned as distinct alleles and, for each isolate, the alleles at each of the seven loci define the allelic profile or sequence type (ST). Each isolate of a species is therefore characterized by a series of seven integers which correspond to the alleles at the seven housekeeping loci.

The number of genes used for MLST is seven but they vary for different species, e.g., *adk*, *atpG*, *frdB*, *fucK*, *mdh*, *pgi*, *recA* are used for *Haemophilus influenzae*, *atpD*, *gltB*, *gyrB*, *recA*, *lepA*, *trpB,* and *phaC* for *Burkholderia cepacia* complex (http://www.mlst.net). Bolt et al. [21] recently reported MLST scheme (genes *atpA*, *dnaE*, *dnaK*, *fusA*, *leuA*, *odhA*, *rpoB*) to characterize global populations of *C. diphtheriae*, *C. ulcerans* and *Corynebacterium pseudotuberculosis*. In particular, 149 isolates of *C. diphtheriae* from 18 countries isolated during 50 years were analyzed by MLST and strain discrimination was generally in accordance with ribotyping data and clonal complexes associated with disease outbreaks were identified. Moreover MLST showed divisions between *gravis*/*mitis* and *belfanti* and evidence of novel veterinary subgroups.

17.2.4.5 MLVA Typing

Multi-locus VNTR (variable number of tandem repeats) analysis MLVA is based on variation in copy number in the polymorphic VNTR loci scattered throughout the bacterial chromosome; this approach is widely used for various bacterial pathogens including relatively close relative of *C. diphtheriae*, *M. tuberculosis*. The number of repeat copies per locus may vary among strains, and the use of several such loci allows sufficient interstrain differentiation. The VNTR profiles are presented as multi-digit numerical codes ("complex haplotypes"), each digit representing the copy number in a locus. In fact, the VNTR loci present multiple independent genetic markers and therefore ideally suit for phylogeographic analysis. At present, MLVA typing of *C. diphtheriae* is still under development [22, 23, 23a].

17.2.4.6 CRISPR-(spoligo)typing

CRISPR (clustered regularly interspaced short palindromic repeats) regions are found in many bacteria and consist of 20–40 bp direct repeats (DR), interspaced by similarly sized non-repetitive variable spacers (Fig. 17.3a). In the complete genome sequence of *C. diphtheriae* strain NCTC13129, two CRISPR loci are located 39 kb downstream and 180 kb upstream of the *oriC* region [24]. The first locus (DRA) consists of seven units; the second DR locus (DRB) consists of 27 DRs and 26 spacers whereas some of them are duplicated.

The recently developed method to study polymorphism in the DRB and DRA CRISPR loci is based on a reverse hybridization in macroarray format [24, 25] (Fig. 17.3b). Analogously to the spoligotyping method used for *M. tuberculosis* analysis, we suggested using the same name for this assay for *C. diphtheriae* subtyping. The specific oligonucleotides (5′-amino labeled) were designed on the basis of the 21 and 6 different spacers sequences found in the DRB and DRA regions in *C. diphtheriae* strain NCTC13129. The probes were chosen to have similar melting temperatures and are covalently bound to a membrane. All spacers of each DR locus are amplified with a locus-specific single primer pair, the reverse primer being 5-biotin labeled (Fig. 17.3a). The biotin-labeled PCR fragments of the both CRISPR regions are co-hybridized to the set of the 27 spacer-derived probes (21 DRB and 6 DRA spacers) by using the MN45 miniblotter providing a macroarray format. The autoradiographs (hybridization profiles on the chemiluminescence-sensitive ECL film [Amersham Biosciences, UK]) are visually assessed for presence/absence of signals (Fig. 17.3b). The obtained hybridization profiles of 21 (DRB locus) and six (DRA locus) signals are entered into spreadsheet in binary format; this allows further simple comparison of profiles by automatic sorting function.

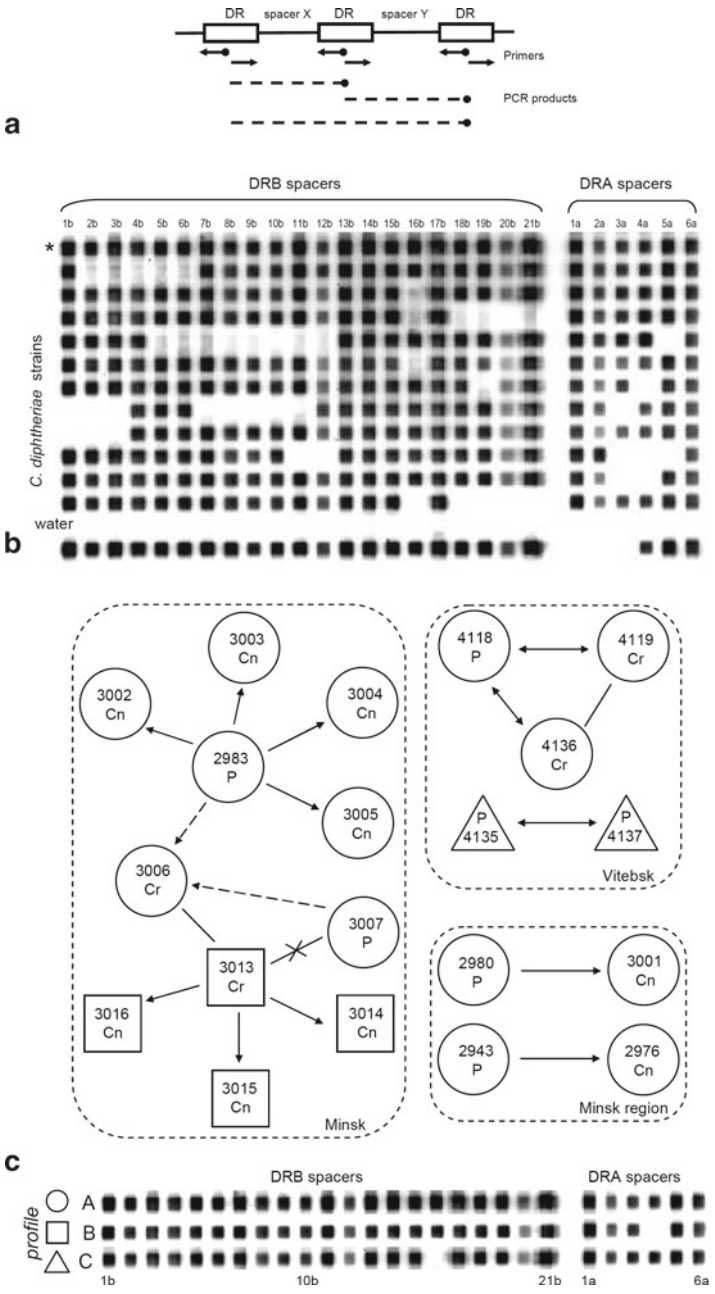

Fig. 17.3 *Corynebacterium diphtheriae* spoligotyping. (**a**) PCR of a CRISPR locus with a locus-specific DR-sequence-defined one primer pair that amplifies all present variable spacers. Reverse primer is labeled with biotin (shown as *black dot*); this permits detection of hybridization signals via streptavidine-peroxidase mediated chemiluminescence on a light-sensitive film. (**b**) Examples of spoligoprofiles of the Russian strains of the epidemic clone; *asterisk* designates the ancestral profile T1 with all 27 signals present. Reprinted from ref. [23] by permission of Elsevier ©2009. (**c**) Combined conventional and molecular epidemiological investigation of the *C. diphtheriae* ribotype Rossija foci in Belarus. In the epidemiological linkage network individuals are linked based on standard investigation; these links were further rejected (*cross*) or confirmed (*solid line*) by spoligotyping analysis of strains. P, patient, Cr, carrier, Cn, contact. Reprinted from ref. [28] by permission of Springer ©2009

17.3 Applications

17.3.1 Origin and Dynamics of the Epidemic Clone in the 1990s

In the 1990s molecular typing methods allowed identification of a clonal group of closely related strains responsible for the epidemic in Russia and NIS countries and to trace strains imported to other countries [2, 7, 14]. These strains were indistinguishable by PFGE, RAPD, AFLP and very similar in ribotyping two principal profiles, Rossija and Sankt-Peterburg, differing in one band (Fig. 17.2) [2, 7, 10, 19]. Minor rare variants were identified by RAPD and ribotyping techniques [14] and a total of 27 ET types similar at >80% were described by MLEE typing in all strains of this clonal group [2]. During diphtheria epidemic in 1990–1996 these closely related toxigenic strains were isolated in a high proportion (70–90%) of patients in all NIS countries and Russia; few strains were also identified in other European countries as imported cases [2, 7, 14, 17]. Prospective and retrospective studies using MLEE, ribotyping and PFGE showed that (1) the pre-epidemic period was characterized by the simultaneous presence of many different ETs; (2) the epidemic clonal group had a unique PFGE profile and comprised MLEE-defined ET8 complex strains of the ribotypes Sankt-Peterburg and Rossija [2]. These ribotypes were less frequently seen in Russia before the epidemic. However, since 1991, they have accounted for an increasing proportion of the isolates studied and by 1994, they accounted for 80% of all identified ribotypes.

The population of *C. diphtheriae* was not constant throughout the epidemic process. As the epidemic became widespread, new geographic variants within the ET8 complex emerged, while more evolutionarily stable ribotypes remained unchanged [2]. In Russia, a steady decline of diphtheria incidence since 1997 was accompanied by increasing heterogeneity of the *C. diphtheriae* population. Although different ribotypes were identified in Russian archival strains recovered in 1940–1990, only one ribotype was prevailing in each particular period: Lyon in 1940–1960, Otchakov in 1980s, Sankt Peterburg and Rossija since mid-1980s until peak in mid-1990s (96% in 1996). Since 1997, during the period of lower incidence, Sankt Peterburg and Rossija became less common (77% in 2001), and other ribotypes became more prominent (Otchakov, Lyon, Cluj, Buzau) [26].

In Belarus, another NIS country affected by the diphtheria epidemic, mass immunization caused visible changes in the circulating population of *C. diphtheriae* [27]. The *gravis* biotype which prevailed in 1996–2000 was replaced with the *mitis* biotype in 2001–2005. Simultaneously, the proportion of toxigenic *C. diphtheriae* strains decreased from 47.1% (1996) to 6.8% (2005). Ribotyping analysis revealed the elimination of rare ribotypes (both toxigenic and nontoxigenic) during the period of decreased morbidity. In 2001–2005 not only "toxigenic" rare ribotypes were eliminated but also the proportion of "toxigenic" prevalent ribotypes decreased from 55.6% to 27.0%. However, the strains of these ribotypes that continued to circulate remained toxigenic. The proportion of Sankt-Peterburg ribotype in the total population decreased from 24.3% to 2.3%, in contrast, the proportion of the Rossija ribotype increased from 25.1% to 49.1% [27].

In the mid-1990s, imported cases of diphtheria contracted in Russia and Ukraine by foreign travelers were diagnosed in Finland, Estonia, Norway, Poland, Latvia, Lithuania, and Germany ([7], and references therein). However, such importation of toxigenic clones not always occurred—even through the borders of the neighboring countries—thus underlying a critical importance of the vaccination coverage and host susceptibility. For example, Romania has a sufficiently long border with the former Soviet Union countries (Ukraine and Moldova), hence opportunities for strain importation. However, the Russian/NIS epidemic did not spread to Romania owing to enforced prevention and control measures [17].

17.3.2 Global Diversity of C. diphtheriae

More distant world regions, such as UK, USA, and Canada were less influenced by Russian epidemics, but instead, via epidemiological links with other world areas reflecting long-term historical links. From 1986 to 1997 in the UK, there were 45 cases of toxigenic *C. diphtheriae* infections that were imported from Africa, Asia, Europe, and Eastern Europe. Most imported cases were caused by the biotype *mitis* and comprised many ribotypes not seen within the Eastern European region. Those ribotypes appeared to be unique to those particular countries, e.g., the ribotypes from cases imported from Thailand were identical to those of the epidemic isolates in Thailand and Laos. To compare the isolates collected during the recent Russian diphtheria epidemic with those circulating worldwide, isolates from Russian and previous epidemic areas (Thailand, Vietnam, Sweden, and the USA) were examined by ribotyping along with sporadic isolates from cases and contacts in the West Indies, France, Italy, Denmark, Romania, Rwanda, and Australia. An assessment of the transcontinental spread of the organism showed that several genotypes of *C. diphtheriae* circulated on different continents of the world and that each outbreak was caused by a distinct clone. The ribotypes seen in Europe appeared to be distinct from those seen elsewhere, and certain ribotypes appeared to be unique to particular countries [7, 11].

Another study used higher-resolution MLEE typing to determine the genetic relatedness of *C. diphtheriae* strains from Russia and NIS and worldwide (Australia, Bangladesh, Ecuador, Finland, Somalia, Sweden, Tunisia, and the USA) [2]. In addition to the ET8 complex (approximate epidemic isolates from Russia and NIS), two more clusters were observed. The first cluster (0.22 dissimilarity), comprised isolates from diphtheria patients and carriers in several US states from 1973 to 1996. These *C. diphtheriae* isolates were clearly distinct from the current epidemic isolates from Russia and NIS. Clustering of older and recent US isolates suggested an endemic focus of toxigenic *C. diphtheriae* in the USA. The other cluster (0.14 dissimilarity) included isolates collected in 1992 in Australia. The rest of the isolates were spread throughout the dendrogram without particular association between ET and geographic or temporal origin of strains [2].

At the same time, persistent foci of diphtheria do exist in developed countries eventually causing small outbreaks. Despite the virtual elimination of diphtheria in

the USA and Canada, toxigenic strains continue to circulate in some communities within the two countries. Molecular characterization of USA (South Dakota) and Canadian (Ontario) *C. diphtheriae* isolates showed that strains with characteristic molecular subtypes have persisted in these areas for at least 25 years [3, 4]. The enhanced surveillance in South Dakota revealed that toxigenic *C. diphtheriae* is circulating among American Indian populations [3].

17.3.3 Comparison of Methods

The traditional molecular methods used for various bacteria and applied to *C. diphtheriae* typing were ribotyping, PFGE, and MLEE [2, 7]. In spite of their wide use, these techniques are time-consuming, require specialized equipment and technical expertise, and, therefore, cannot be performed in all laboratories (Table 17.1). In contrast, the PCR-based methods are faster and simpler, although they frequently lack sufficient discriminatory power and reproducibility and their standardization is a challenge. De Zoysa et al. [11] compared four typing methods and found ribotyping to be highly discriminatory and reproducible. The statistical analysis data calculated for the different typing methods indicated that ribotyping is the most suitable technique and the method of choice for the typing of *C. diphtheriae*. The two PCR-based techniques, RAPD and AFLP, proved to be rapid and easier to perform than ribotyping and PFGE. The diversity indices indicated that AFLP and PFGE are less discriminatory than ribotyping and RAPD. RAPD and AFLP were suggested as rapid methods which can be used as screening techniques, prior to ribotyping, during outbreak investigations [11].

Unlike the above traditional typing methods, CRISPR-based spoligotyping of *C. diphtheriae* was evaluated in only two settings. A large number of variable characters (i.e., particular spacers that may be present or absent in a CRISPR locus) should provide sufficient level of variation to differentiate clinical strains. Indeed, a spoligotyping study of the Russian strains of the epidemic clone revealed an astonishingly high diversity. The 156 Russian strains of the epidemic clone from St. Petersburg, 1997–2002, were subdivided into 45 spoligotypes compared to only two ribotypes (Sankt-Peterburg and Rossija). The larger DRB locus (21 spacers) demonstrated higher polymorphism. Nevertheless, a combined use of the two loci additionally contributed to further strain differentiation (compared to use of the DRB locus alone) [25]. Interestingly, ribotype Sankt-Peterburg was found to be more heterogeneous than Rossija, in both CRISPR loci alone and in their combination [24, 25]. The polymorphism in both CRISPR loci and spoligotype distribution within two ribotypes of the *C. diphtheriae* epidemic clone led us to suggest: (1) a monophyletic origin of the epidemic clone, (2) presently clonal evolution of these CRISPR loci in *C. diphtheriae* genome, and (3) divergence between ribotypes Sankt-Peterburg and Rossija in northwestern Russia. Assuming that more diversity is generated due to longer evolutionary history, ribotype Sankt-Peterburg appears to be evolutionarily older and ancestral to ribotype Rossija. This latter may have originated

Table 17.1 Comparison of methods used for *C. diphtheriae* strain typing

Method	Target	Rationale/principle	Type of data	Advantages	Disadvantages	References or website
Ribotyping	16S and 23S rRNA genes and flanking regions	Total genomic digestion followed by Southern hybridization with rDNA probe	Restriction fragment length polymorphism: patterns based on size and no. of fragments that contain rRNA genes	Good discrimination, reproducible	Time consuming, limited portability	[8, 10] http://www.dipnet.org/ribo.public.php
PFGE	Whole genome	Total genomic digestion with rarely cutting endonuclease followed by electrophoresis characterized by frequent changes in orientation of electric field	Restriction fragment length polymorphisms of total genomic DNA	Reproducible	Less discriminatory than ribotyping; labor-intensive; limited portability	[7, 17]
MLEE	Various loci, usually housekeeping genes	Allelic differences in mobility revealed by starch gel electrophoresis	Electrophoretic type: composite of allele types	Reproducible, discriminatory	Labor-intensive; limited portability	[2, 12]
AFLP	Whole genome	Selective PCR amplification of genomic restriction fragments	Patterns in agarose gel based on size and no. of amplicons	Rapid, inexpensive, reproducible	Less discriminatory than ribotyping	[19]
RAPD	Whole genome	Differences in abundance and distribution of annealing sites revealed by polymerase chain reaction	Patterns in agarose gel based on size and no. of amplicons	Rapid, inexpensive	Insufficient reproducibility	[13–16]

Spoligotyping	Two CRISPR loci	Reverse hybridization macroarray-based detection of spacers present in CRISPR loci	Digital binary profile (0/1 for presence/absence of particular spacers)	Highest discrimination; rapid, portable, inexpensive	Reproducibility and stability to be checked; use with other ribotypes beyond epidemic clone yet to be established	[24, 25]
MLST	Seven selected housekeeping genes	Partial sequencing	Multi-digit allelic profile	Rapid, portable, reproducible	Discrimination similar to ribotyping	[21] http://www.mlst.net/misc/new_schemes.asp
MLVA	Multiple VNTR loci	Size detection of PCR fragments by agarose gel or capillary sequencer to detect number of repeat units in a locus	Multi-digit allelic profile	Portable, easy, reproducible	Discrimination and stability should be tested	[22, 23, 23a]

from one particular subpopulation (ancestral type T1 [Fig. 17.3b]) of the presumably already heterogeneous ribotype Sankt-Peterburg, followed by subsequent independent non-homoplasious evolution of the DRB and DRA loci in both ribotype sublineages.

In another study, 20 *C. diphtheriae* biotype *gravis* toxigenic isolates collected in Belarus from suspected foci of diphtheria infection (diphtheria cases, carriers, or contacts) were subjected to ribotyping and spoligotyping [28]. All strains had identical ribotype profile Rossija based on comparison with international ribotype database in Institut Pasteur of Paris. However, spoligotyping based on analysis of two CRISPR loci differentiated these strains into three spoligotypes (Fig. 17.3c). Comparison of the spoligotyping results with the epidemiological linkage network helped us to resolve suspected links in the chains of transmission.

17.4 Conclusions

Since 1993, application of the traditional molecular subtyping methods and continuous monitoring of the spread of the epidemic clones had a significant public health impact making it possible to distinguish between epidemic, endemic, and imported cases and allowing for implementation of timely and adequate preventive measures. To improve such continuous monitoring of the circulation of existing clones and to be able to rapidly detect the appearance of new and unusual clones, the *C. diphtheriae* ribotype database has been established in Institut Pasteur, Paris [10] and is also available at DIPNET web site (http://www.dipnet.org). More recently, publication of the *C. diphtheriae* complete genome sequence and partial genome sequencing permitted to select seven housekeeping genes for MLST [21] and to identify two CRISPR regions [24, 25] and multiple VNTR loci [22, 23] as possible candidates for development of new generation typing formats of *C. diphtheriae*.

General and/or potential advantages of the CRISPR- and VNTR-based methods are (1) digital presentation of data as discrete binary or multistate characters and portability; (2) high discriminatory power; (3) possibility to use cell lysates instead of purified DNA; and (4) high-throughput capacity. The novel CRISPR-(spoligo)typing [25] and MLST [21] methods may become a powerful tool for portable and high-resolution epidemiological monitoring and phylogenetic analysis of *C. diphtheriae.*

The critically important issue regarding these new methods is their field evaluation in both particular settings and in global collection. For example, it is not surprising that a collection of the geographically diverse strains was well differentiated by MLST [21]. On the other hand, although spoligotyping achieved excellent discrimination within the Russian epidemic clone [24, 25], the method should be validated in other settings. Furthermore, development of the CRISPR typing should include new spacers and other circulating variants of *C. diphtheriae* beyond the epidemic clone and investigate the spoligoprofile stability in short- and middle-term natural and laboratory-modeled evolution. Indeed, MLST, VNTR, and CRISPR

encoded information is easily databasable. Standardization of these new methods and building of a comprehensive database representing various markers and different levels of genetic diversity of *C. diphtheriae* and including microbiological/clinical strain/patient data is a task and challenge for near future. Availability of the very recently published new complete genomes of *C. diphtheriae* strains representing different biotypes [29–31] will provide new in-depth insights into pathogenomics and evolutionary history of this important human pathogen.

References

1. Efstratiou A, Engler KH, De Zoysa A (1998) Diagnosis and epidemiology of diphtheria. Methods Mol Med 15:191–212
2. Popovic T, Mazurova IK, Efstratiou A et al (2000) Molecular epidemiology of diphtheria. J Infect Dis 181:S168–S177
3. Popovic T, Kim C, Reiss J et al (1999) Use of molecular subtyping to document long-term persistence of *Corynebacterium diphtheriae* in South Dakota. J Clin Microbiol 37:1092–1099
4. Marston CK, Jamieson F, Cahoon F et al (2001) Persistence of a distinct *Corynebacterium diphtheriae* clonal group within two communities in the United States and Canada where diphtheria is endemic. J Clin Microbiol 39:1586–1590
5. Cerdeño-Tarraga AM, Efstratiou A, Dover LG et al (2003) The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. Nucleic Acids Res 31:6516–6523
6. Grimont F, Grimont PA (1986) Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools. Ann Inst Pasteur Microbiol 137B:165–175
7. De Zoysa A, Efstratiou A, George RC et al (1995) Molecular epidemiology of *Corynebacterium diphtheriae* from northwestern Russia and surrounding countries studied by using ribotyping and pulsed-field gel electrophoresis. J Clin Microbiol 33:1080–1083
8. Regnault B, Grimont F, Grimont PA (1997) Universal ribotyping method using a chemically labelled oligonucleotide probe mixture. Res Microbiol 148:649–659
9. Bouchet V, Huot H, Goldstein R (2008) Molecular basis of ribotyping. Clin Microbiol Rev 21:262–273
10. Grimont PA, Grimont F, Efstratiou A et al (2004) International nomenclature for *Corynebacterium diphtheriae* ribotypes. Res Microbiol 155:162–166
11. De Zoysa A, Hawkey P, Charlett A, Efstratiou A (2008) Comparison of four molecular typing methods for characterization of *Corynebacterium diphtheriae* and determination of transcontinental spread of *C. diphtheriae* based on *BstE*II rRNA gene profiles. J Clin Microbiol 46:3626–3635
12. Reeves MW, Evins GM, Heiba AA et al (1989) Clonal nature of *Salmonella typhi* and its genetic relatedness to other salmonellae as shown by multilocus enzyme electrophoresis, and proposal of *Salmonella bongori* comb. nov. J Clin Microbiol 27:313–320
13. De Zoysa AS, Efstratiou A (1999) PCR typing of *Corynebacterium diphtheriae* by random amplification of polymorphic DNA. J Med Microbiol 48:335–340
14. Kombarova SI, Mazurova IK, Mel'nikov VG et al (2001) Genetic structure of *Corynebacterium diphtheriae* strains isolated in Russia during epidemics of various intensity. Zh Mikrobiol Epidemiol Immunobiol 3:3–8, In Russian
15. Mokrousov I (1995) Arbitrary PCR typing of *Corynebacterium diphtheriae* strains. PHLS Microbiol Digest 12:89

16. Nakao H, Popovic T (1998) Use of random amplified polymorphic DNA for rapid molecular subtyping of *Corynebacterium diphtheriae*. Diagn Microbiol Infect Dis 30:167–172
17. Damian M, Grimont F, Narvskaya O et al (2002) Study of *Corynebacterium diphtheriae* strains isolated in Romania, northwestern Russia and the Republic of Moldova. Res Microbiol 153:99–106
18. Vos P, Hogers R, Bleeker M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
19. De Zoysa A, Efstratiou A (2000) Use of amplified fragment length polymorphisms for typing *Corynebacterium diphtheriae*. J Clin Microbiol 38:3843–3845
20. Dallman T, Neal S, De Zoysa A et al (2008) A comparative genomic study of several European *Corynebacterium diphtheriae* strains. In: Proceeding 2nd annual meeting of DIPNET and 10th international meeting of the European Laboratory Working Group on Diphtheria, Larnaca, Cyprus, 5–7 November 2008, p 53
21. Bolt F, Cassiday P, Tondella ML, Dezoysa A, Efstratiou A, Sing A, Zasada A, Bernard K, Guiso N, Badell E, Rosso ML, Baldwin A, Dowson C (2010) Multilocus sequence typing identifies evidence for recombination and two distinct lineages within *Corynebacterium diphtheriae*. J Clin Microbiol 48:4177–4185
22. Neal S, Efstratiou A (2007) DIPNET—establishment of a dedicated surveillance network for diphtheria in Europe. Euro Surveill 12:E9–E10
23. Mokrousov I (2009) *Corynebacterium diphtheriae*: genome diversity, population structure and genotyping perspectives. Infect Genet Evol 9:1–15
23a. Zasada AA, Jagielski M, Rzeczkowska M, Januszkiewicz A. (2011) The use of MLVA for *Corynebacterium diphtheriae* genotyping - preliminary studies. Med Dosw Mikrobiol 63: 209-218. In Polish
24. Mokrousov I, Narvskaya O, Limeschenko E, Vyazovaya A (2005) Efficient discrimination within *Corynebacterium diphtheriae* epidemic clone by the novel macroarray-based method. J Clin Microbiol 43:1662–1668
25. Mokrousov I, Limeschenko E, Vyazovaya A, Narvskaya O (2007) *Corynebacterium diphtheriae* spoligotyping based on combined use of two CRISPR loci. Biotechnol J 2:901–906
26. Kombarova SY, Melnikov VG, Borisova OY et al (2002) Genotype characteristics of *Corynebacterium diphtheriae* strains in a period of low diphtheria incidence in Russia: 1997 to 2001. In: Proceedings of seventh international meeting of the European Laboratory Working Group on Diphtheria and Diphtheria Surveillance Network DIPNET, Vienna, Austria, 12–14 June 2002
27. Kolodkina V, Titov L, Sharapa T et al (2006) Molecular epidemiology of *C. diphtheriae* strains during different phases of the diphtheria epidemic in Belarus. BMC Infect Dis 6:129
28. Mokrousov I, Vyazovaya A, Kolodkina V et al (2009) Novel macroarray-based method of *Corynebacterium diphtheriae* genotyping: evaluation in a field study in Belarus. Eur J Clin Microbiol Infect Dis 28:701–703
29. Trost E, Blom J, Soares S de C, et al (2012) Pangenomic study of *Corynebacterium diphtheriae* that provides insights into the genomic diversity of pathogenic isolates from cases of classical diphtheria, endocarditis, and pneumonia. J Bacteriol 194:3199–3215
30. Sangal V, Tucker NP, Burkovski A, Hoskisson PA. (2012) The draft genome sequence of *Corynebacterium diphtheriae* bv. mitis NCTC 3529 reveals significant diversity between the primary disease-causing biovars. J Bacteriol 194:3269
31. Sangal V, Tucker NP, Burkovski A, Hoskisson PA. (2012) Draft genome sequence of *Corynebacterium diphtheriae* biovar intermedius NCTC 5011. J Bacteriol 194:4738