

Endemic erythromycin resistant *Corynebacterium diphtheriae* in Vietnam in the 1990s

To Nguyen Thi Nguyen^{1,2,*}, Christopher M. Parry^{3,4,5}, James I. Campbell^{1,3}, Phat Voong Vinh¹, Rachel Kneen^{5,6} and Stephen Baker^{7,8}

Abstract

Diphtheria is a potentially fatal respiratory disease caused by toxigenic forms of the Gram-positive bacterium *Corynebacterium diphtheriae*. Despite the availability of treatments (antitoxin and antimicrobials) and effective vaccines, the disease still occurs sporadically in low-income countries and in higher income where use of diphtheria vaccine is inconsistent. Diphtheria was highly endemic in Vietnam in the 1990s; here, we aimed to provide some historical context to the circulation of erythromycin resistant organisms in Vietnam during this period. After recovering 54 *C. diphtheriae* isolated from clinical cases of diphtheria in Ho Chi Minh City between 1992 and 1998 we conducted whole genome sequencing and analysis. Our data outlined substantial genetic diversity among the isolates, illustrated by seven distinct Sequence Types (STs), but punctuated by the sustained circulation of ST67 and ST209. With the exception of one isolate, all sequences contained the *tox* gene, which was classically located on a corynebacteriophage. All erythromycin resistant isolates, accounting for 13% of organisms in this study, harboured a novel 18kb *erm(X)*-carrying plasmid, which exhibited limited sequence homology to previously described resistance plasmids in *C. diphtheriae*. Our study provides historic context for the circulation of antimicrobial resistant *C. diphtheriae* in Vietnam; these data provide a framework for the current trajectory in global antimicrobial resistance trends.

DATA SUMMARY

- (1) All supplementary data (Table S1, Table S2, Fig. S1, S2, and S3, available in the online version of this article) have been deposited in FigShare. (https://figshare.com/articles/figure/Supplementary_Figures/14728200)
- (2) All short reads used in this study were registered in the European Nucleotide Archive (ENA) under study accessions of PRJEB32654 (Accession numbers: ERR3331346- ERR3331399)
- (3) The plasmid sequence was deposited in GenBank (Accession number: MZ348427)

We confirm that all supporting data, code, and protocols have been provided within the article or through supplementary data files.

INTRODUCTION

Diphtheria is a life-threatening upper respiratory disease caused by the Gram-positive organism *Corynebacterium diphtheriae* (*C. diphtheriae*). The organism infects the pharynx, tonsils, and nasal passage, causing a sore throat, low-grade fever, and an inflammatory pseudomembrane on the tonsils 2–3 days post-infection [1, 2]. The most common transmission route of *C. diphtheriae* is

Received 23 December 2021; Accepted 10 June 2022; Published 19 October 2022

Author affiliations: ¹The Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam; ²Department of Microbiology, Monash University, Melbourne, Victoria, Australia; ³Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK; ⁴Clinical Sciences, Liverpool School of Tropical Medicine, Pembroke Pl, Liverpool L3 5QA, UK; ⁵Alder Hey Children's Hospital, NHS Foundation Trust, Liverpool, UK; ⁶Institute of Infection and Global Health, University of Liverpool, Liverpool L69 3BX, UK; ⁷University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge, UK; ⁸Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge, UK.

***Correspondence:** To Nguyen Thi Nguyen, To.NguyenThiNguyen@monash.edu

Keywords: antimicrobial resistance; *Corynebacterium diphtheriae*; diphtheria; erythromycin; outbreak.

Abbreviations: AMR, antimicrobial resistance; EPI, expanded programme immunization; MDR, multidrug resistant; MIC, minimum inhibitory concentration; ML, maximum likelihood; MLST, multilocus sequence typing; QC, quality checking; RAxML, randomized accelerated maximum likelihood; SNP, single nucleotide polymorphism; ST, sequence type; VFDB, virulence factors database; WGS, whole genome sequencing.

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Three supplementary figures and two supplementary tables are available with the online version of this article.

000861 © 2022 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License. This article was made open access via a Publish and Read agreement between the Microbiology Society and the corresponding author's institution.

Impact Statement

Infection caused by *Corynebacterium diphtheriae* has recently become a heavy burden on human health globally. Despite development and global introduction of diphtheria toxoid vaccine, diphtheria, the illness caused by *Corynebacterium diphtheriae* remains endemic in many low-income settings where there are issues with vaccine coverage. The most well-described virulence factor of *C. diphtheriae* is the toxin, which is encoded by the *tox* gene encoded on a beta-corynephage. Diphtheria antitoxin and antimicrobials (such as penicillin and erythromycin) are considered the most effective therapies for the treatment of diphtheria. However, there has been an increase in resistance to these antimicrobials that may limit treatment options. The introduction of the diphtheria vaccine was successful in Vietnam in the 1980s; however, there were several diphtheria outbreaks a decade later, some of which were characterised by a high proportion of erythromycin resistant isolates. In this study, we used whole genome sequencing to investigate the genome structure and diversity of the *C. diphtheriae* population from this period. We identified seven different sequence types (ST) among; the most common was ST67, which has been recently found in outbreaks in the central highlands of Vietnam. The toxigenic and virulence components of these isolates were also characterised. Specifically, we identified a novel plasmid carrying the *erm(X)* gene and conferring erythromycin resistance. Our results highlight the importance of historical collections in providing greater context during contemporary outbreaks and the role of drug resistance.

via contact with an infected individual or droplets contaminated with the organism [1]. Asymptomatic carriage is common, but infants, the elderly, and the immuno-compromised are at particular risk of severe disease and death (case-fatality rate of >20%) [2, 3]. Diphtheria toxoid vaccine was developed from the 1920s and introduced into Expanded Programme on Immunization (EPI) internationally in 1977 to reduce *C. diphtheriae* disease and community transmission. However, recent reports highlight that diphtheria remains endemic in several low-income settings. Outbreaks have been reported in India, Laos, Bangladesh and the Yemen and also in high-income settings regions with low immunisation coverage as seen in Norway [4–9]; in 2017 the estimated number of cases globally remained >16000 [10–12].

Diphtheria toxin is the key virulence factor of *C. diphtheriae* and one of the most well-characterised bacterial toxins. The toxin triggers the characteristic disease features in the respiratory tract and also more severe symptoms including myocarditis, peripheral neuropathy, and acute renal injury [1, 13, 14]. The toxin is encoded by the *tox* gene and is carried on a beta-corynephage [1, 15]. *Tox* is regulated by the metallo-regulatory transcriptional regulator DtxR, which is repressor of *tox*, and is active when iron is not limited [16]. When secreted, the diphtheria exotoxin induces local tissue destruction and facilitates the formation of the pseudomembrane [1, 15]. Penicillin and erythromycin have long been the primary treatments for diphtheria, but recent reports describe an increase in resistance against these antimicrobials, potentially posing a problem for diphtheria treatment [17–20]. Several studies have described resistance mechanisms against antimicrobials used for the treatment of diphtheria [21–23]. Specifically, erythromycin resistance in *C. diphtheriae* is commonly encoded by the plasmid mediated *erm(X)* gene [24]. *Erm(X)* gene encodes an rRNA methyl transferase enzyme catalysing demethylation of adenine in the 23S rRNA gene, which prevents effective binding of the macrolide to the target rRNA [24].

Diphtheria is endemic in Vietnam, although disease incidence declined rapidly after the introduction of the vaccine into the EPI in the late 1980s [25, 26]. The incidence of diphtheria increased in Vietnam in the mid-1990s and there were several outbreaks, with some isolates exhibiting resistant to erythromycin [25, 27, 28]. There are little available data from this period, and no molecular characterisation of bacterial isolates from this period. Here, through whole genome sequencing (WGS) and analysis we aimed to investigate the population structure, diversity, and antimicrobial resistance gene composition in *C. diphtheriae* isolated from Vietnamese patients with symptomatic disease in the 1990s.

METHODS

Bacterial isolates

We revived 54 *C. diphtheriae* isolated from the nose or throat of children with a clinical diagnosis of respiratory diphtheria attending the Hospital for Tropical Diseases in Ho Chi Minh City between 1992 and 1998. Some of them were part of two studies reported in Clinical Infectious Diseases [29, 30]. Briefly, at the time of isolation, throat and nasal swabs from clinical cases were cultured on Hoyle's tellurite agar and sheep blood agar and incubated at 37 °C for 24–48 h. Single colonies suspected to be *C. diphtheriae* were identified morphologically using methylene blue and then confirmed by the cysteinase and pyrazinamidase test and by API Coryne (bioMérieux, France) [31]. Toxin production was confirmed by using a standard Elek test. Organisms were stored at –20 °C in protect bottles with the BHI+glycerol 10% medium until their attempted revival and recovered on sheep blood agar and Brain Heart Infusion broth supplemented with yeast extract (0.4%) and Tween 80 (0.2%) (BHI/YE/Tween).

Antimicrobial susceptibility testing

The Minimum Inhibitory Concentrations (MICs) of the organisms were performed by the agar plate dilution method [32]. Single colonies were sub-cultured into Brain Heart Infusion broth supplemented with yeast extract (0.4%) and Tween 80 (0.2%) (BHI/YE/Tween) before incubation at 37 °C for 18 h. A multipoint inoculator (Mast Laboratories Ltd, Merseyside, UK) was used to apply 300 µl inoculum (10^5 to 10^6 c.f.u. ml⁻¹) of each organism to the surface of Mueller Hinton agar (Unipath Ltd, Basingstoke, UK) containing saponin-lysed sheep blood (5%) and serial two-fold dilutions of Erythromycin, Tetracycline, Ceftriaxone, Rifampicin, Ampicillin and Chloramphenicol. All antimicrobials were sourced from Sigma Aldrich, except ceftriaxone, which was from Rocephin (Rocephin, Roche, Hong Kong, China). Toxigenic *C. diphtheriae* var *gravis* NCTC10356, non-toxigenic *C. diphtheriae* var *gravis* NCTC11397, and *Staphylococcus aureus* ATCC25923 were included as controls. Plates were incubated at 37 °C for 18 h. Antimicrobial susceptibility was determined following guideline for infrequently isolated or fastidious bacteria [33].

Genome sequencing

Genomic DNA from all isolates was extracted by using the Wizard Genomic DNA Extraction Kit (Promega, USA) [34] and quality assessed using a Qubit dsDNA system (Invitrogen) [35]. The genomic DNA was diluted and 1 ng from each isolate was used for library preparation in accordance with the guidelines of the manufacturer (New England Biolab and NexteraXT) before being subjected to WGS on an Illumina Miseq System [36]. FastQC was used for assessing the quality of reads and passed reads were progressed for mapping and assemblies [37]. The FASTQ generated were submitted to the ENA (<https://www.ebi.ac.uk/ena>) under the project number of PRJEB32654, and accession numbers from ERR3331346 to ERR3331399 (Table S1).

Genome analysis

All reads passing QC checking were mapped to the NCTC13129 reference genome (Accession no. BX248353.1) using Bowtie2 v2.2.9 and SAMTools v1.3.1. All single nucleotide polymorphisms (SNPs) with the Phred quality score of ≥ 30 of each isolate were determined and called [38, 39]. A total of 66761 conserved SNPs were identified after removing those located in the recombinogenic regions and those arising in $<5\%$ of all isolates [40, 41]. The pseudogenome alignment was generated and then subjected to statistical scanning for potential recombinations using Gubbins. These regions were excluded from further analysis [42]. Variable sites that had minimum read depth ≤ 5 and a genome coverage $\leq 50\%$ were also excluded.

A phylogeny of 54 *C. diphtheriae* genome sequences was reconstructed from core SNPs determined from mapping using Randomized Accelerated Maximum Likelihood (RAxML) [43]. RAxML was run five times for SNP allele table by using the generalized time-reversible model with Γ distribution (GTR+ Γ) to model site-specific rate variation with one hundred bootstrap pseudo-replicate analyses to assess the maximum likelihood (ML) tree supports.

Known alleles associated with antimicrobial resistance (AMR) and virulence genes were directly detected from read sets mapping approach based on SRST2 [26]. The ARG-Annot database was used for the detection of AMR genes. MLST and virulence genes were determined using the *C. diphtheriae* BIGSdb database at Institute Pasteur (<http://bigsdbsdb.web.pasteur.fr>) and Virulence Factors Database (VFDB) for *Corynebacterium*. Additionally, fastq reads were further *de novo* assembled using SPAdes v3.14.1 to generate contigs. Contaminated sequences were removed according to the size of contigs compared that of the reference genome. Genomes with evidence of contamination were excluded from the study, according to the following criteria: total assembly length >2.5 –3 Mb, with evidence of $>1\%$ read contamination as determined by MetaPhlAn and applied for other MDR bacteria, or $<50\%$ reads mapping to the NCTC13129 reference chromosome (accession number: BX248353.1); or a ratio of heterozygous/ homozygous single nucleotide polymorphism (SNP) calls compared to the reference chromosome exceeding 20% [44]. The presence of AMR carrying plasmid was resolved and visualized using assembled contigs in Bandage v0.8.1 [45].

Plasmids

Isolates carrying the *erm(X)* gene were cultured on sheep blood agar supplemented with Erythromycin (2 mg l⁻¹) and plasmid DNA was extracted using the Plasmid Miniprep Kit (QIAprep Kit). Plasmid DNAs were subjected to library preparation prior to sequencing on an Illumina Miseq system. All raw reads were *de novo* assembled and visualized in Bandage to determine and resolve the *erm(X)* carrying plasmid. The circularised plasmid was compared with the reference plasmid pNG2 (Accession no. AF492560.1) using Blast Ring Image Generator (BRIG) v0.95 [46]. The plasmid sequences were deposited in Genbank (Accession number: MZ348427).

RESULTS

Genomic diversity of *C. diphtheriae*

Of the 54 recovered *C. diphtheriae* isolated from 1992 to 1998 from patients with diphtheria, the overwhelming majority (53/54; 98.15%) were confirmed to be toxigenic (Table S1). Mapping reads from the WGS of 54 isolates to the *C. diphtheriae*

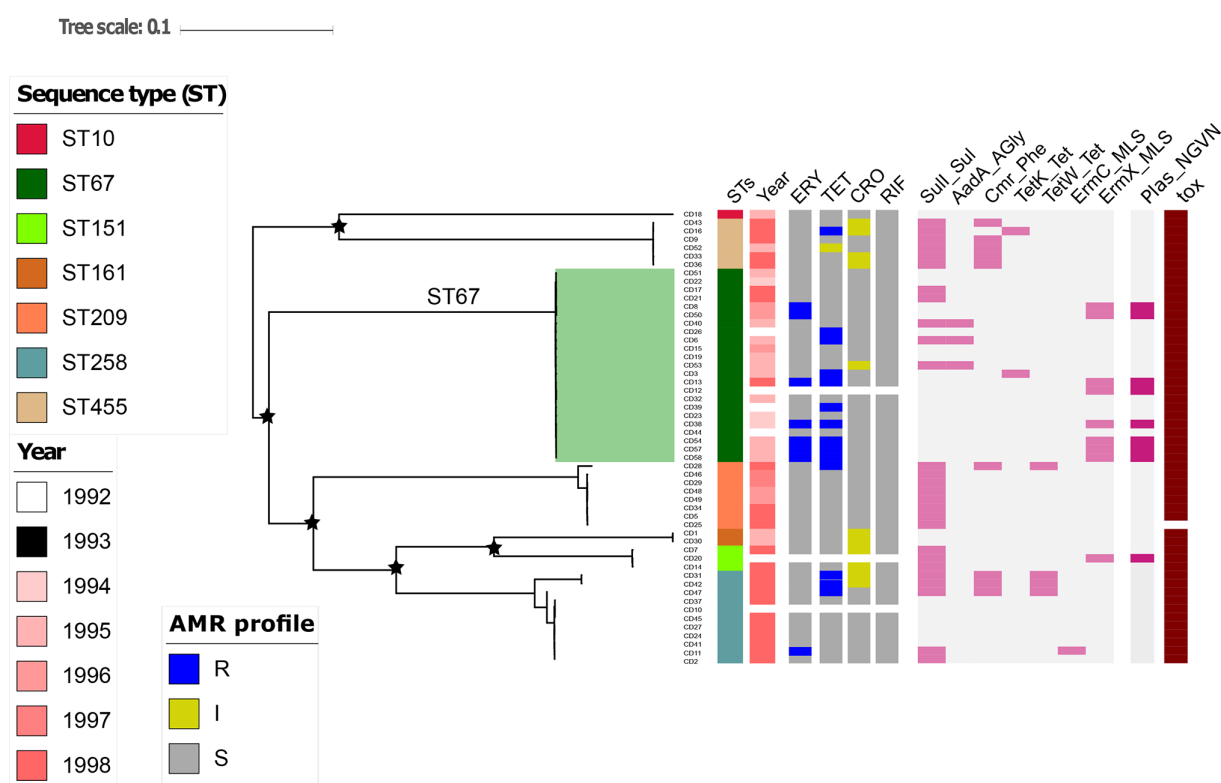


Fig. 1. Phylogeny, antimicrobial resistance phenotype and genotype of the *C. diphtheriae* isolates. Maximum likelihood phylogeny of 54 *C. diphtheriae* was reconstructed from 66 761 core SNPs. The corresponding columns were metadata: Sequence Type (STs), the year of isolation (Year), antimicrobial susceptible profile (blue: resistant; yellow: intermediate; grey: susceptible) using agar plate dilution method for Erythromycin (ERY), Tetracycline (TET), Ceftriaxone (CRO), Rifampicin (RIF). The heat map was the antimicrobial resistant genes (pink: gene presence; green: gene absence) and the last columns was a determined plasmid carrying *erm(X)* gene encoding for erythromycin resistance (pink: plasmid presence; green: plasmid absence). The stars show bootstrap support values 100% on internal nodes. The shading colour was noted for the most common sequence type of ST67.

mitis NCTC13129 reference genome determined an average coverage of 90.7%, an average mapping of 86.19%, and an average read depth coverage of 59x. A phylogenetic tree was reconstructed from 66761 core SNPs and identified seven independent lineages. Multilocus sequence typing (MLST) analysis delineated seven independent STs; of which ST67 was the most common (23/54; 42.6%), followed by ST258 (11/54; 20.4%), ST209 (8/54; 14.8%), ST455 (6/54, 11.1%), ST151 (3/54, 5.5%), ST161 (2/54, 3.7%) and ST10 (1/54, 1.8%) (Fig. 1).

We aligned supplementary metadata with the phylogeny, which included year of sampling and antimicrobial susceptibility against to erythromycin, tetracycline, ceftriaxone, and rifampicin. All isolates were susceptible to rifampicin ($\text{MIC} \leq 0.016 \text{ mg l}^{-1}$). Almost all isolates were susceptible to ceftriaxone ($\text{MIC} \leq 1.0 \text{ mg l}^{-1}$), with 20% (11/54) exhibiting intermediate resistance ($\text{MIC} = 2 \text{ mg l}^{-1}$). Notably, 13% (7/54) of isolates were resistant to erythromycin, the traditional second line antimicrobial for diphtheria treatment. The proportion of isolates that were resistant to tetracycline, an alternative second line antimicrobial, was 18.5% (10/54). The sequences were screened for genes encoding resistance to antimicrobials; 16.7% (9/54) of organisms carried *erm(X)*, conferring resistance to erythromycin (Fig. 1). In addition, half of the sequenced isolates possessed the *sul1* gene, encoding sulphonamide resistance, and 16.7% (9/54) of sequences carried the *cmr* gene, encoding resistance to chloramphenicol, correlating with an $\text{MIC} \geq 2 \text{ mg l}^{-1}$ against chloramphenicol. The MIC values against ampicillin ranged from 0.06 mg l^{-1} to 1 mg l^{-1} .

Virulence factors of *C. diphtheriae*

Three distinct pili structures have been described in *C. diphtheriae* [47, 48]. These structures are encoded by three different gene clusters and facilitate the attachment of the organism to host cells. Analysis of the three pilus encoding regions in our *C. diphtheriae* genomes collection found that all organisms lacked the SpaH-type pilus (Fig. S1); however, most isolates (96.3%; 52/54) possessed the SpaA-type pilus. Additionally, the SpaD-type pilus was conserved in all the ST455, ST10, and ST161 isolates only. The SapD surface-anchored pilus proteins were present in all isolates, whereas only six isolates of ST455 contained SapA.

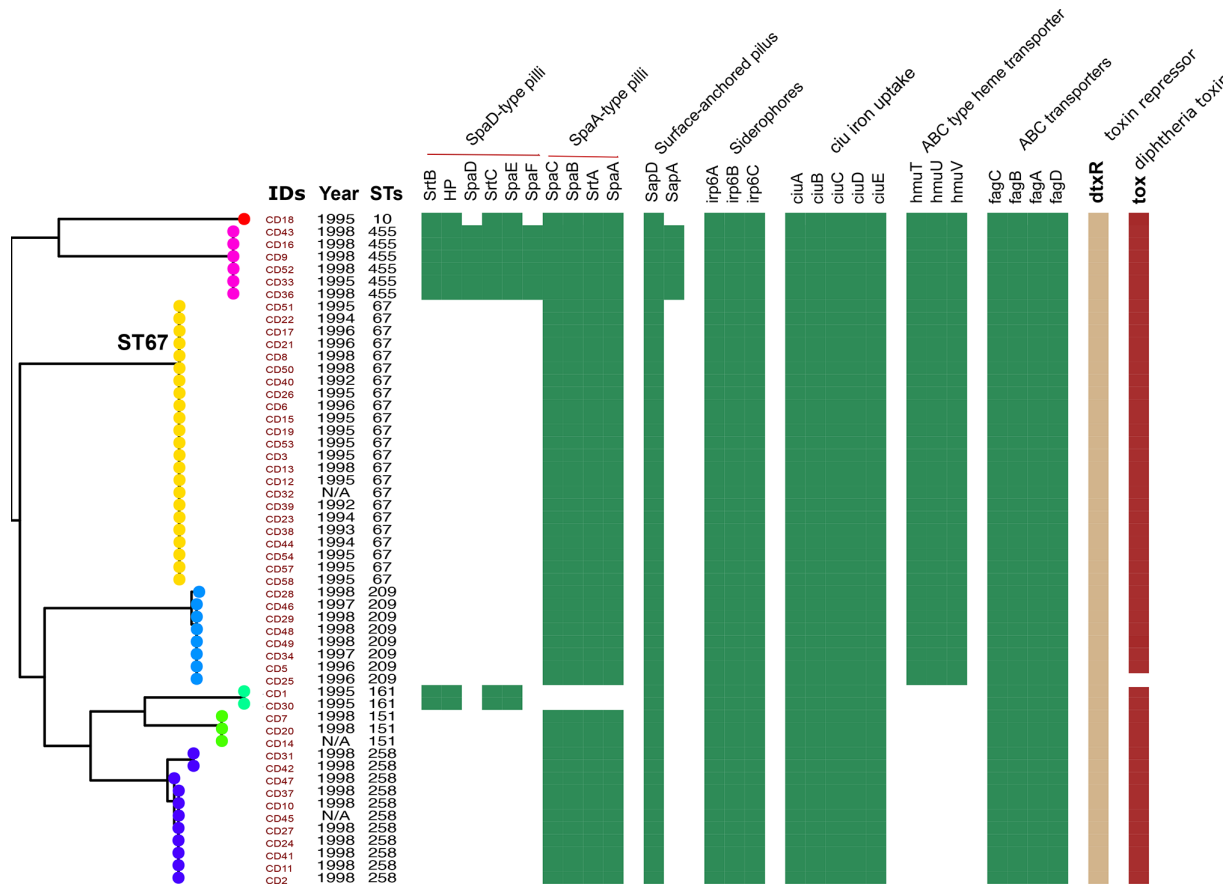


Fig. 2. Virulence factors in the 54 *C. diphtheriae* isolates. SpaD and SpaA-type pili gene clusters encoding for pili proteins that were employed to target to human pharyngeal cells. STs: sequence types of isolates. Colours: gene presence; white: gene absence.

Other gene clusters related to iron uptake was present in all isolates, with the exception of ABC-type haem transporters which were lacking in the ST161, ST151 and ST258 organisms (Fig. 2). In concordance with the phenotypic data, we found almost all (98.1%; 53/54) organisms possessed *tox* (associated with the corynebacteriophage). Notably, the non-toxigenic isolate missing the *tox* gene was also lacking the entire prophage (Fig. S2). Despite the prophage being comparatively well conserved in the ST67 and ST258 isolates, the genes encoding the tail proteins and prophage anti-repressor were absent. Other organisms in the collection were found to be lacking some regions encoding exported proteins and various transcriptional regulators (Table S2).

A novel plasmid carrying *erm(X)* gene

To better characterise the plasmid carrying the *erm(X)* gene, we performed a plasmid DNA extraction and sequenced the raw plasmid DNA. Reads were *de novo* assembled and generated three large contigs with a coverage of 248×, 235×, and 480×, respectively (Fig. S3). Comparing the contigs from the plasmid with 480× coverage with sequences available on GenBank, we identified that these contigs encoded an 842 bp mobile element. Using Bandage to create an interactive visualization of the assembly graph from the *de novo* assembly, we resolved these contigs and assembled a novel circular plasmid of 18651 kb, which we named pNGVN (Fig. 3b). A pairwise comparison between pNGVN and previously described plasmid pNG2 (AF_492560.1) revealed that the pNGVN had 65% sequence coverage and an average of identity of 88.78% (85.44–98.23%) to pNG2 (Fig. 3a). We described a total of 23 putative coding sequences in pNGVN; of which 10/23 coding sequences were functionally annotated. Plasmid pNGVN contained several functional genes that were comparable to those of pNG2 including *repA*, *traA*, *parA*, and *parB*. These genes are a replicase protein, a conjugal transfer protein, and type Ib plasmid partitioning proteins, respectively. Additionally, the *erm(X)* gene on pNGVN was 98% identical to that on pNG2 and located next to transposable elements that were previously identified in *C. striatum* and *C. resistans*. In the pNG2 plasmid sequence, the *erm(X)* gene had an *erm(LP)* located upstream. This gene was annotated as the 23S rRNA adenine N-6-methyltransferase leader peptide. We also identified an identical sequence to the *erm(LP)* gene in pNGVN; however, the output from the Prokka tools for our sequence was unable to annotate this gene. Six out of nine organisms carrying *erm(X)* possessed pNGVN

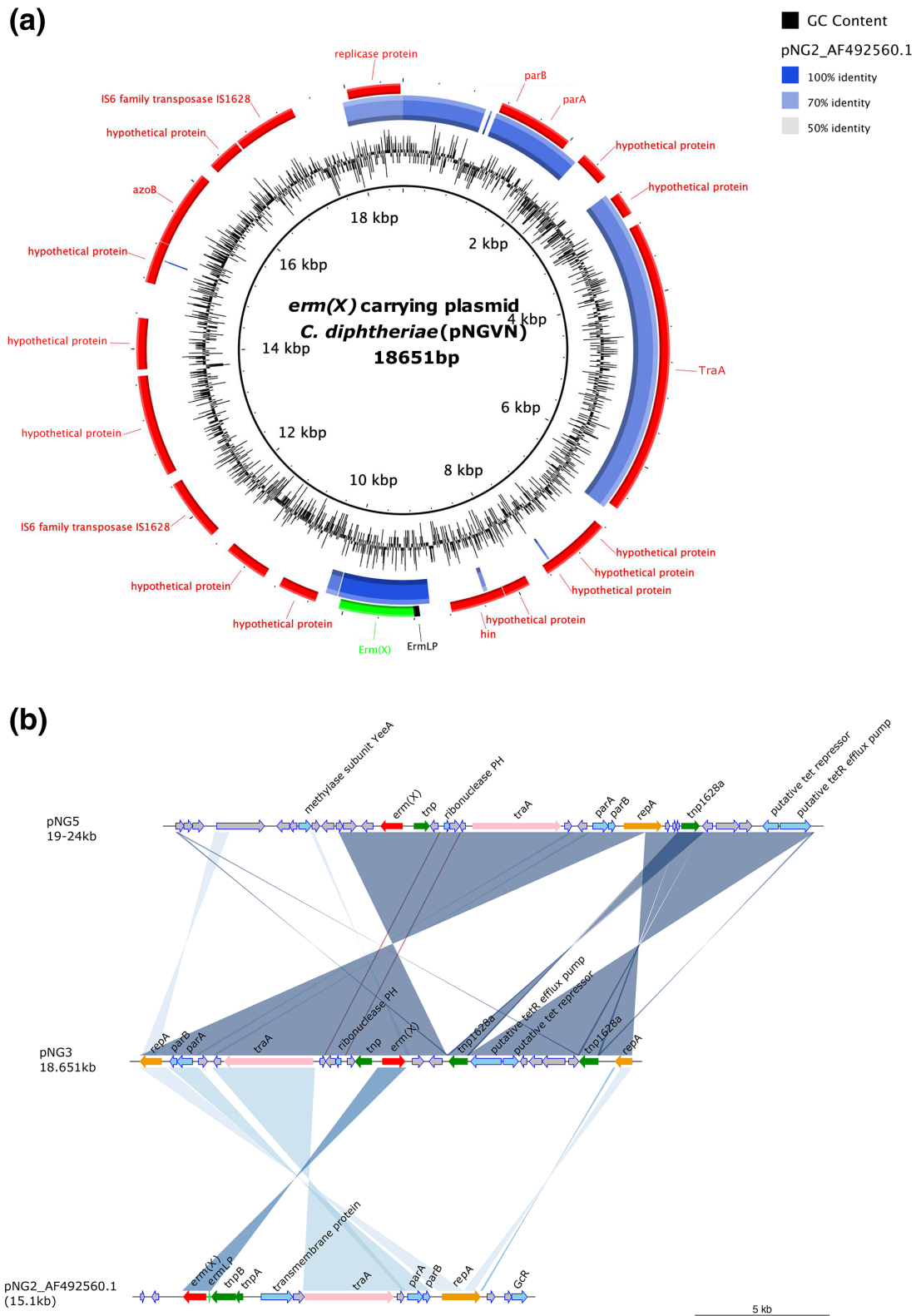


Fig. 3. The genetic structure of *erm(X)*-carrying plasmids in *C. diphtheriae*. a. Pairwise comparison of *erm(X)*-carrying plasmids: described plasmid (pNG2, AF_492560.1, 15kb) and our novel plasmid (pNGVN, 18.651 kb) detected in six out of nine *C. diphtheriae* strains and a larger plasmid (pNG5, 19–24 kb) possessed by three other isolates. b. The genetic structure of the *erm(X)*-carrying plasmid pNGVN. BLASTN comparison between pNGVN (central ring) identified from an erythromycin resistant ST67 outbreak isolate and pNG2 plasmid from GenBank (Accession number: AF_492560.1). The outermost ring indicates the gene annotations of the pNGVN plasmid. *erm(X)* gene is highlighted in green colour.

like plasmids. The additional three isolates possessed a larger plasmid (pNG5, 19–24 kb) with a pNGVN backbone with an additional cassette gene (~6 kb) carrying the methylase subunit, *YeeA* and a further nine hypothetical proteins (Fig. 3a).

DISCUSSION

Here, we sequenced whole genome of 54 *C. diphtheriae* isolates collected from patients admitted to a major tertiary hospital in the south of Vietnam in the 1990s. Our results outlined substantial genetic diversity among the Vietnamese *C. diphtheriae* population, with seven different STs identified. The most common STs (ST67) and the ST209 have also been identified in diphtheria outbreaks in the western and central highlands of Vietnam recently, implying the sustained circulation of these STs across the country [28]. ST67 was the most common genotype associated with diphtheria in an observational studies from San Lazaro Hospital in the Philippines and in Europe [49, 50], whereas ST258 and ST209 were both reported in Thailand in 2012 [51]. Estimating within-group SNP variation for different STs in our study showed SNP numbers varied substantially within these STs; the average of SNP difference in ST67 was 65 (range: 51–83 SNPs), 39 in ST151 (range: 32–49 SNPs), 32 in ST258 (range: 29–36 SNPs) and 20 in ST455 (range: 17–26 SNPs). There are different thresholds for genetic associations suggesting whether organisms originated from outbreaks. The threshold for a recent diphtheria outbreak in the US was 11.6 SNPs (range: 0–24 SNPs) and 2–12 SNPs has been suggested for other MDR bacteria [52–55]. The diverse nature of the organisms within this study suggest that these organisms were indicative of endemic circulation rather than an outbreak.

Almost all *C. diphtheriae* isolates in this were susceptible to rifampicin and ceftriaxone (some organisms higher MICs at 2 mg l⁻¹), whilst 13 and 18.5% of isolates were resistant to erythromycin and tetracycline, respectively. The organisms described generally exhibited lower MIC (0.06 mg l⁻¹ ≤ MIC ≤ 1 mg l⁻¹) against chloramphenicol than those described by Maple *et al.* (all at MIC at 1 mg l⁻¹) [56]. Notably, despite these organisms being isolated in the 1990s, the prevalence of erythromycin resistance was higher than the reports from Indonesia and Brazil (5.3%, 4.2%, respectively). Alternatively, the observed proportion of tetracycline resistance was higher than in the aforementioned study in Brazil (12.8%), but substantially lower than that described in Indonesia (84.2%) [17, 18]. Screening for AMR genes identified a variety of genes conferring AMR phenotypes. All isolates were susceptible to penicillin; therefore, we found no *pbp2m* gene or presence of pLRPD-like plasmid in this collection. Although some organisms had higher MICs against ceftriaxone (MIC=2 mg l⁻¹), we did not detect any genes/chromosomal mutations associated with this phenotype. A principal observation was the identification of the *erm(X)* gene, encoding erythromycin resistance located on a novel 18651 kb plasmid (primarily found in the CD54 strain). The plasmid exhibited 65% DNA sequence identity to plasmid pNG2 (AF_492560.1, 15 kb), which was described by Tauch and colleagues in 2003. In total we identified 23 coding sequences from the novel plasmid, 47.8% of these had been functionally annotated, necessitating some further work to determine the function of the identified sequences.

We additionally found that almost all isolates (98%) were toxigenic, which was associated with the *tox* gene located on the classical beta corynebacterium. However, we did identify some signs of genetic degradation around the *tox* gene, with genes encoding bacteriophage tail proteins, prophage anti-repressor, exported and transcriptional proteins were missing. To our knowledge this is the first time such degradation has been observed around the *tox* gene, but genetic variation with the corynebacterium sequence is relatively common [57–59]. Additionally, most isolates possessed the SpaA pilus. The presence of the SpaA type pili had been determined to be essential for corynebacterial adherence to the host cells and known as pilus mediated adherence. The SpaD-type pili was identified in only nine isolates, while SpaH was not detected [60]. This has also been observed in other *C. diphtheriae* and is therefore not unique to these isolates from Vietnam. These data suggest some redundancy in the pili required to attach to host cells and induce infection of the pathogen. However, we did not perform phenotypic characterisation and the loss (or rearrangement) of some genetic material may be an artefact of long term storage after the original isolation, which is a potential limitation.

Diphtheria has been largely controlled by the sustained use of the vaccine since the 1980s; however, in locations where vaccine uptake is less complete the organism still has the potential to trigger outbreaks. This scenario is more common in low-middle income countries, such as Vietnam, where the cases of diphtheria arise sporadically for sustained periods and ongoing surveillance is required. Here, using a historic collection of organisms and data we have investigated a collection of *C. diphtheriae* isolated in Vietnam in the 1990s, when erythromycin resistance was an emerging problem. Our data provide some additional context to *C. diphtheriae* in Vietnam and we can show that identical STs (67 and 209) have remained in circulation for >30 years and are capable of seeding new outbreaks [28]. Our results emphasize that historic collections of organisms improve the framework for ongoing surveillance for diseases that have not yet been eradicated and provide better context for the current trajectory in global antimicrobial resistance trends.

Funding information

This project was supported by a Wellcome senior research fellowship to Stephen Baker (215515/Z/19/Z). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgements

The authors would like to thank the staff at the Hospital for Tropical Diseases, Ho Chi Minh City, Vietnam for their assistance in isolation, storage, and provision of samples and data collection for this study. 'Dr. James I Campbell was not able to verify his authorship, but To Nguyen Thi Nguyen confirms that Dr. James I Campbell should be credited as the author and that Dr. James I Campbell contributed to the paper in providing the bacterial collection and conceptualizing for this paper.'

Author contributions

Conceptualization S.B., C.M.P., J.I.C. Methodology T.N.T.N., P.V.V., R.K. Investigation T.N.T.N., P.V.V., R.K. Writing-original draft T.N.T.N. Writing-review and editing S.B., C.M.P. Resources C.M.P., J.I.C. P.V.V.; Supervision S.B., C.M.P. All authors have read and accepted the final version of the manuscript.

Conflicts of interest

The authors declare that there are no conflicts of interests.

Ethical statement

This was a retrospective study from stored isolates, given that no patient data were accessed ethical approval was not required.

References

- Hadfield TL, McEvoy P, Polotsky Y, Tzinserling VA, Yakovlev AA. The pathology of diphtheria. *J Infect Dis* 2000;181 Suppl 1:S116–20.
- Hong K-W, Asmah Hani AW, Nurul Aina Murni CA, Pusparani RR, Chong CK, et al. Comparative genomic and phylogenetic analysis of a toxigenic clinical isolate of *Corynebacterium diphtheriae* strain B-D-16-78 from Malaysia. *Infect Genet Evol* 2017;54:263–270.
- Centres for Disease Control and Prevention. Diphtheria; 2016
- Parande MV, Parande AM, Lakkannavar SL, Kholkute SD, Roy S. Diphtheria outbreak in rural North Karnataka, India. *JMM Case Rep* 2014;1:3–5.
- Saikia L, Nath R, Saikia NJ, Choudhury G, Sarkar M. A diphtheria outbreak in Assam, India. *Southeast Asian J Trop Med Public Health* 2010;41:647–652.
- Rasmussen I, Wallace S, Mengshoel AT, Høiby EA, Brandtæg P. Diphtheria outbreak in Norway: lessons learned. *Scand J Infect Dis* 2011;43:986–989.
- Nanthavong N, Black AP, Nouanthong P, Souvannaso C, Vilivong K, et al. Diphtheria in Lao PDR: insufficient coverage or ineffective vaccine? *PLoS One* 2015;10:e0121749.
- Moghalles SA, Aboasba BA, Alamad MA, Khader YS. Epidemiology of diphtheria in yemen, 2017–2018: surveillance data analysis. *JMIR Public Health Surveill* 2021;7:e27590.
- Badell E, Alharazi A, Criscuolo A, Almoayed KAA, Lefrancq N, et al. Ongoing diphtheria outbreak in Yemen: a cross-sectional and genomic epidemiology study. *Lancet Microbe* 2021;2:e386–e396.
- Zakikhany K, Efstratiou A. Diphtheria in Europe: current problems and new challenges. *Future Microbiol* 2012;7:595–607.
- Daskalaki I. *Corynebacterium diphtheriae*. In: *Principles and Practice of Pediatric Infectious Diseases*, 4th edn. 2012.
- Vitek CR, Wharton M. Diphtheria in the former Soviet Union: reemergence of a pandemic disease. *Emerg Infect Dis* 1998;4:539–550.
- Havaladar PV, Sankpal MN, Doddannavar RP. Diphtheritic myocarditis: clinical and laboratory parameters of prognosis and fatal outcome. *Ann Trop Paediatr* 2000;20:209–215.
- Solders G, Nennesmo I, Persson A. Diphtheritic neuropathy, an analysis based on muscle and nerve biopsy and repeated neurophysiological and autonomic function tests. *J Neurol Neurosurg Psychiatry* 1989;52:876–880.
- Sharma NC, Efstratiou A, Mokrousov I, Mutreja A, Das B, et al. Diphtheria. *Nat Rev Dis Primers* 2019;5:81.
- Zajdowicz SLW, Holmes RK. Phage Conversion and the Role of Bacteriophage and Host Functions in Regulation of Diphtheria Toxin Production by *Corynebacterium diphtheriae*. In: *The Mechanistic Benefits of Microbial Symbionts*. 2016. pp. 15–45.
- Pereira GA, Pimenta FP, dos Santos FRW, Damasco PV, Hirata Júnior R, et al. Antimicrobial resistance among Brazilian *Corynebacterium diphtheriae* strains. *Mem Inst Oswaldo Cruz* 2008;103:507–510.
- Sariadji K, Sunarno S, Pusparani N, Sembiring M. Antibiotic susceptibility pattern of *Corynebacterium diphtheriae* isolated from outbreaks in Indonesia 2010–2015. *Indones Biomed J* 2016;10:51.
- Mina NV, Burdz T, Wiebe D, Rai JS, Rahim T, et al. Canada's first case of a multidrug-resistant *Corynebacterium diphtheriae* strain, isolated from a skin abscess. *J Clin Microbiol* 2011;49:4003–4005.
- Rockhill RC, Sumarmo, Hadiputranto H, Siregar SP, Muslihun B. Tetracycline resistance of *Corynebacterium diphtheriae* isolated from diphtheria patients in Jakarta, Indonesia. *Antimicrob Agents Chemother* 1982;21:842–843.
- Will RC, Ramamurthy T, Sharma NC, Veeraraghavan B, Sangal L, et al. Spatiotemporal persistence of multiple, diverse clades and toxins of *Corynebacterium diphtheriae*. *Nat Commun* 2021;12:1500.
- Forde BM, Henderson A, Playford EG, Looke D, Henderson BC, et al. Fatal respiratory Diphtheria Caused by β -Lactam-Resistant *Corynebacterium diphtheriae*. *Clin Infect Dis* 2021;73:e4531–e4538.
- Hennart M, Panunzi LG, Rodrigues C, Gaday Q, Baines SL, et al. Population genomics and antimicrobial resistance in *Corynebacterium diphtheriae*. *Genome Med* 2020;12:107.
- Olender A. Mechanisms of Antibiotic Resistance in *Corynebacterium spp.* Causing Infections in People. In: *Antibiot Resist Bact - A Contin Chall New Millenn*. 2012.
- Murakami H, Phuong NM, Thang HV, Chau NV, Giao PN, et al. Endemic diphtheria in Ho Chi Minh City; Viet Nam: A matched case-control study to identify risk factors of incidence. *Vaccine* 2010;28:8141–8146.
- Inouye M, Dashnow H, Raven L-A, Schultz MB, Pope BJ, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med* 2014;6:90.
- Doanh PV, Pham T D, Vu N L, Chu V T, Do T H H, et al. An outbreak of diphtheria in K'Bang District, Gia Lai, Vietnam, October 2013 – July 2014. *Int J Infect Dis* 2016;45:172.
- Kitamura N, Le TTT, Le LT, Nguyen LD, Dao AT, et al. Diphtheria outbreaks in schools in Central Highland Districts, Vietnam, 2015–2018. *Emerg Infect Dis* 2020;26:596–600.
- Kneen R, Nguyen MD, Solomon T, Pham NG, Parry CM, et al. Clinical features and predictors of diphtheritic cardiomyopathy in Vietnamese children. *Clin Infect Dis* 2004;39:1591–1598.
- Kneen R, Pham NG, Solomon T, Tran TM, Nguyen TT, et al. Penicillin vs. erythromycin in the treatment of diphtheria. *Clin Infect Dis* 1998;27:845–850.
- Zasada AA, Mosiej E. Contemporary microbiology and identification of *Corynebacteria spp.* causing infections in human. *Lett Appl Microbiol* 2018;66:472–483.
- EUCAST. EUCAST Definitive Document E.DEF 3.1, June 2000: Determination of minimum inhibitory concentrations (MICs) of antibacterial agents by agar dilution. *Clin Microbiol Infect* 2000;6:509–515.
- Clinical and Laboratory Standards Institute (CLSI). *M45 - Methods for Antimicrobial Dilution and Disk Susceptibility Testing of Infrequently Isolated or Fastidious Bacteria - 3rd edition*. 2016.
- Corporation P. Wizard® genomic DNA purification kit. *Tech Man Wizard® Genomic DNA Purif Kit* 2012:608–277.
- Thermo Fisher Scientific. (n.d) User guide: qubit dsdna HS assay kits.
- Illumina. *Nextera XT DNA library prep reference guide*. 2016.

37. Andrews S. FastQC: A quality control tool for high throughput sequence data. *Babraham Bioinformatics* 2010.
38. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–359.
39. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–2079.
40. Lam MMC, Wyres KL, Wick RR, Judd LM, Fostervold A, et al. Convergence of virulence and MDR in a single plasmid vector in MDR *Klebsiella pneumoniae* ST15. *J Antimicrob Chemother* 2019;74:1218–1222.
41. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc Natl Acad Sci U S A* 2015;112:E3574–81.
42. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015;43:e15.
43. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312–1313.
44. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, et al. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods* 2012;9:811–814.
45. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 2015;31:3350–3352.
46. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 2011;12:402.
47. Ton-That H, Marraffini LA, Schneewind O. Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. *Mol Microbiol* 2004;53:251–261.
48. Ton-That H, Schneewind O. Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol Microbiol* 2003;50:1429–1438.
49. Saito N, Dimapilis VO, Fujii H, Suzuki M, Telan EFO, et al. Diphtheria in Metro Manila, the Philippines 2006–2017: a clinical, molecular, and spatial characterization. *Clin Infect Dis* 2021;72:61–68.
50. Czajka U, Wiatrzyk A, Mosiej E, Formińska K, Zasada AA. Changes in MLST profiles and biotypes of *Corynebacterium diphtheriae* isolates from the diphtheria outbreak period to the period of invasive infections caused by nontoxigenic strains in Poland (1950–2016). *BMC Infect Dis* 2018;18:1–8.
51. Paveenkittiporn W, Sripakdee S, Koobkratok O, Sangkitporn S, Kerdsin A. Molecular epidemiology and antimicrobial susceptibility of outbreak-associated *Corynebacterium diphtheriae* in Thailand, 2012. *Infect Genet Evol* 2019;75:104007.
52. Xiaoli L, Benoliel E, Peng Y, Aneke J, Cassidy PK, et al. Genomic epidemiology of nontoxigenic *Corynebacterium diphtheriae* from King County, Washington State, USA between July 2018 and May 2019. *Microb Genom* 2020;6.
53. Stimson J, Gardy J, Mathema B, Crudu V, Cohen T, et al. Beyond the SNP Threshold: identifying outbreak clusters using inferred transmissions. *Mol Biol Evol* 2019;36:587–603.
54. Chorlton SD, Ritchie G, Lawson T, Romney MG, Lowe CF. Whole-genome sequencing of *Corynebacterium diphtheriae* isolates recovered from an inner-city population demonstrates the predominance of a single molecular strain. *J Clin Microbiol* 2020;58:e01651–19.
55. Berger A, Dangel A, Schober T, Schmidbauer B, Konrad R, et al. Whole genome sequencing suggests transmission of *Corynebacterium diphtheriae*-caused *Cutaneous diphtheria* in two siblings, Germany, 2018. *Euro Surveill* 2019;24.
56. Maple PA, Efstratiou A, Tseneva G, Rikushin Y, Deshevoi S, et al. The in-vitro susceptibilities of toxigenic strains of *Corynebacterium diphtheriae* isolated in northwestern Russia and surrounding areas to ten antibiotics. *J Antimicrob Chemother* 1994;34:1037–1040.
57. Sekizuka T, Yamamoto A, Komiya T, Kenri T, Takeuchi F, et al. *Corynebacterium ulcerans* 0102 carries the gene encoding diphtheria toxin on a prophage different from the C. diphtheriae NCTC 13129 prophage. *BMC Microbiol* 2012;12:12–72.
58. Brüssow H, Canchaya C, Hardt W-D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* 2004;68:560–602.
59. Liu Y, Harrison PM, Kunin V, Gerstein M. Comprehensive analysis of pseudogenes in prokaryotes: widespread gene decay and failure of putative horizontally transferred genes. *Genome Biol* 2004;5:R64.
60. Mandlik A, Swierczynski A, Das A, Ton-That H. *Corynebacterium diphtheriae* employs specific minor pilins to target human pharyngeal epithelial cells. *Mol Microbiol* 2007;64:111–124.

Five reasons to publish your next article with a Microbiology Society journal

1. When you submit to our journals, you are supporting Society activities for your community.
2. Experience a fair, transparent process and critical, constructive review.
3. If you are at a Publish and Read institution, you'll enjoy the benefits of Open Access across our journal portfolio.
4. Author feedback says our Editors are 'thorough and fair' and 'patient and caring'.
5. Increase your reach and impact and share your research more widely.

Find out more and submit your article at microbiologyresearch.org.