PATHOGEN CHARACTERISTICS

- Aerobic gram-positive bacillus
- There are toxigenic and nontoxigenic isolates
- Known to cause large outbreaks
- Short incubation period, usually 1 week
- The species *C. diphtheriae* is divided into four biochemical biovars: belfanti, gravis, intermedius, and mitis

DETECTION

- Preferred test for toxin production is a culture based test called the Elek test.
- PCR to detect toxin gene (tox A or tox B) and *dtxR* gene are available
- Antimicrobial susceptibility testing to determine the minimum inhibitory concentration
- Test for susceptibility or resistance to: benzylpenicillin, tetracycline, vancomycin, erythromycin, gentamicin, clindamycin, linezolid, rifampicin, ciprofloxacin, cotrimoxazole

GENOMIC ATTRIBUTES

- Genome length: 2.45 Mbp
- Core genes: 1,630
- Evolutionary rate: 1.67×10^{-6} substitutions per site per year or 0.40915 subs per year
- Genome transfer: horizontal transfer substantially helps shape the bacterial genome
- The diphtheria toxin is encoded on a prophage and can also be carried by *Corynebacterium ulcerans* and *Corynebacterium pseudotuberculosis*
- The ß-corynephage encodes the diphtheria toxin, and can be transmitted between isolates. The Omega gene also encodes the toxin, but Beta is the most common.
- If the ß-homolog gene is detected, it will be worth more investigation. It is called a homologue because the gene resembles the beta toxin, may function like the beta toxin, but evolved separately.
- Prophages are genetically more similar within specific clusters of bacterial isolates than between clusters, suggesting that prophages do not randomly mix between isolates, but rather cluster within specific clades.
- Pilus expression may strongly influence the virulence of a strain, especially in combination with the presence of the tox gene.
- Recombination plays an important role in bacterial evolution and has been linked to increased virulence in some pathogens. Among *C. diphtheriae,* recombination can be observed in the upper respiratory tract, where *C. diphtheriae* can form a colonizing state that fosters horizontal gene transfer.

WHOLE GENOME SEQUENCING (WGS)

- Pathogen needs to be cultured
- WGS allows for a detailed comparison on the genomic level with high resolution. In the case of *C. diphtheriae*, high-resolution typing is helpful to provide the epidemiological broader context and include or exclude transmission events between patients.
- Future surveillance programs may not only incorporate baseline features of an isolate such as sequence type and presence or absence of the tox gene, but also more detailed genomic analysis and a virulence factor profile. The aim of this would be to better assess the potential of a strain to cause outbreaks with more severe clinical phenotypes.
- Through WGS analysis, the presence of virulence factors such as the toxin gene (and ß-corynephage) and pili, and genes encoding antimicrobial resistance determinants can be determined.

- Comparison of WGS data across a species generally uses one of two approaches: cgMLST, or SNP-based variant calling across the whole genome based on a reference, which provides more information and higher resolution.
- Reference used is NCTC13129

SNP Phylogeny

- "Reports suggest that isolates recovered from outbreak clusters of diphtheria frequently differ by fewer than 150 SNPs, whereas unlinked sporadic isolates differ by an average of 30,000 SNPs."
- Isolates within transmission networks typically differ by <150 SNPs. Isolates from epi linked cases may differ by <40 SNPs.

PEOPLE AT INCREASED RISK

People in the same household, people with a history of frequent, close contact with the patient, people directly exposed to secretions from the suspected infection site (e.g., mouth, skin) of the patient. Across Europe, infections are mainly diagnosed in travelers and refugees from regions where diphtheria is more endemic, patients from urban areas with poor hygiene, and intravenous drug users. About half of the cases are non-toxin producing isolates.

EPIDEMIOLOGICAL OBJECTIVES

Rapid identification of the bacterial pathogen and toxin production for patient and outbreak management.

Diphtheria is a serious infection caused by the strains of bacteria called *Corynebacterium diphtheriae* that are toxigenic. Diphtheria spreads from person to person, usually through respiratory droplets from coughing or sneezing. People can also get sick from touching infected open sores or ulcers. The more common presentation is wound infections. It is imperative to differentiate toxigenic from non-toxigenic isolates as it is the toxin that can cause people to get very sick. The toxin attacks healthy tissues in the respiratory system. Within two to three days, the dead tissue forms a thick, gray coating, "pseudomembrane," that can build up in the throat or nose. It can cover tissues in the nose, tonsils, voice box, and throat, making it very hard to breathe and swallow. If the toxin gets into the bloodstream, it can cause heart, nerve, and kidney damage. Immunized patients are still prone to *C. diphtheriae* bacteremia and endocarditis.

INVESTIGATION QUESTIONS

- Relatedness among specific communities, for example people experiencing homelessness
- Are there differences in the strains and sequences observed among drug users vs. non drug users?
- Are there particular patterns that are notable in the I-5 corridor?
- Are there differences (sequencing diversity) across wound sites? For example, when there are sequences available for the same case from different locations

From the literature

- Include or exclude transmission events between patients.
- Define clusters
- Identify which strains cause outbreaks with more severe clinical phenotypes (i.e. endocarditis)

TOOLS

Elek test: used to confirm toxin production

Etest: is used to detect antibiotic resistance

pupMLST: can be used to determine MLST allele profiles

Roary: defines protein-coding gene clusters

Snippy: to calculate SNPs by mapping trimmed reads to the reference genome

Phylogeny: RaxML has been used to estimate the phylogeny using maximum-likelihood using SNP alignments. Neighbor-joining algorithms have also been used in the literature. IQTree with best-fit model GTR+F+R10.

Tools used at WA DOH

PHoeNIx (https://github.com/CDCgov/phoenix?tab=readme-ov-file): bioinformatics analysis pipeline built using Nextflow.It performs: quality control, checks for contamination, confirms taxa ID, performs sequence typing, assembles reads into scaffolds, detects antimicrobial resistance and hypervirulence genes, and searches for plasmid markers.

BigBacter (https://github.com/DOH-JDJ0303/bigbacter-nf): pipeline developed by a PHL Bioinformatician for bacterial genomic surveillance. This pipeline pre-clusters isolates into closely related subtypes prior to phylogenetic analysis, selects and archives cluster-specific reference genomes for SNP analysis, identifies and excludes low quality samples, archives samples and automatically includes them when samples from the same cluster are identified, automatically generates figures needed for phylogenetic analysis (i.e., trees and SNP matrices).

Gubbins (https://github.com/nickjcroucher/gubbins)**:** algorithm that iteratively identifies loci containing elevated densities of base substitutions, which are marked as recombinations, while concurrently constructing a phylogeny based on the putative point mutations outside of these regions.

RESOURCES

Dangel A, Berger A, Konrad R, Bischoff H, Sing A. **Geographically Diverse Clusters of Nontoxigenic Corynebacterium diphtheriae Infection, Germany, 2016-2017.** Emerg Infect Dis. 2018 Jul;24(7):1239-1245. doi: 10.3201/eid2407.172026. PMID: 29912709; PMCID: PMC6038752.

Hennart M, Panunzi LG, Rodrigues C, Gaday Q, Baines SL, Barros-Pinkelnig M, Carmi-Leroy A, Dazas M, Wehenkel AM, Didelot X, Toubiana J, Badell E, Brisse S. **Population genomics and antimicrobial resistance in Corynebacterium diphtheriae.** Genome Med. 2020 Nov 27;12(1):107. doi: 10.1186/s13073-020-00805-7. PMID: 33246485; PMCID: PMC7694903.

Julien Guglielmini, Melanie Hennart, Edgar Badell, Julie Toubiana, Alexis Criscuolo, Sylvain Brisse. **Genomic epidemiology and strain taxonomy of Corynebacterium diphtheriae.**bioRxiv 2021.07.18.452871; doi: https://doi.org/10.1101/2021.07.18.452871

Meinel DM, Kuehl R, Zbinden R, Boskova V, Garzoni C, Fadini D, Dolina M, Blümel B, Weibel T, Tschudin-Sutter S, Widmer AF, Bielicki JA, Dierig A, Heininger U, Konrad R, Berger A, Hinic V, Goldenberger D, Blaich A, Stadler T, Battegay M, Sing A, Egli A. **Outbreak investigation for toxigenic Corynebacterium diphtheriae wound infections in refugees from Northeast Africa and Syria in Switzerland and Germany by whole genome sequencing.** Clin Microbiol Infect. 2016 Dec;22(12):1003.e1-1003.e8. doi: 10.1016/j.cmi.2016.08.010. Epub 2016 Aug 30. PMID: 27585943.

Seth-Smith HMB, Egli A. **Whole Genome Sequencing for Surveillance of Diphtheria in Low Incidence Settings.** Front Public Health. 2019 Aug 21;7:235. doi: 10.3389/fpubh.2019.00235. PMID: 31497588; PMCID: PMC6713046.

Xiaoli L, Benoliel E, Peng Y, Aneke J, Cassiday PK, Kay M, McKeirnan S, Duchin JS, Kawakami V, Lindquist S, Acosta AM, DeBolt C, Tondella ML, Weigand MR. **Genomic epidemiology of nontoxigenic Corynebacterium diphtheriae from King County, Washington State, USA between July 2018 and May 2019.** Microb Genom. 2020 Dec;6(12):mgen000467. doi: 10.1099/mgen.0.000467. Epub 2020 Dec 4. PMID: 33275088; PMCID: PMC8116682.