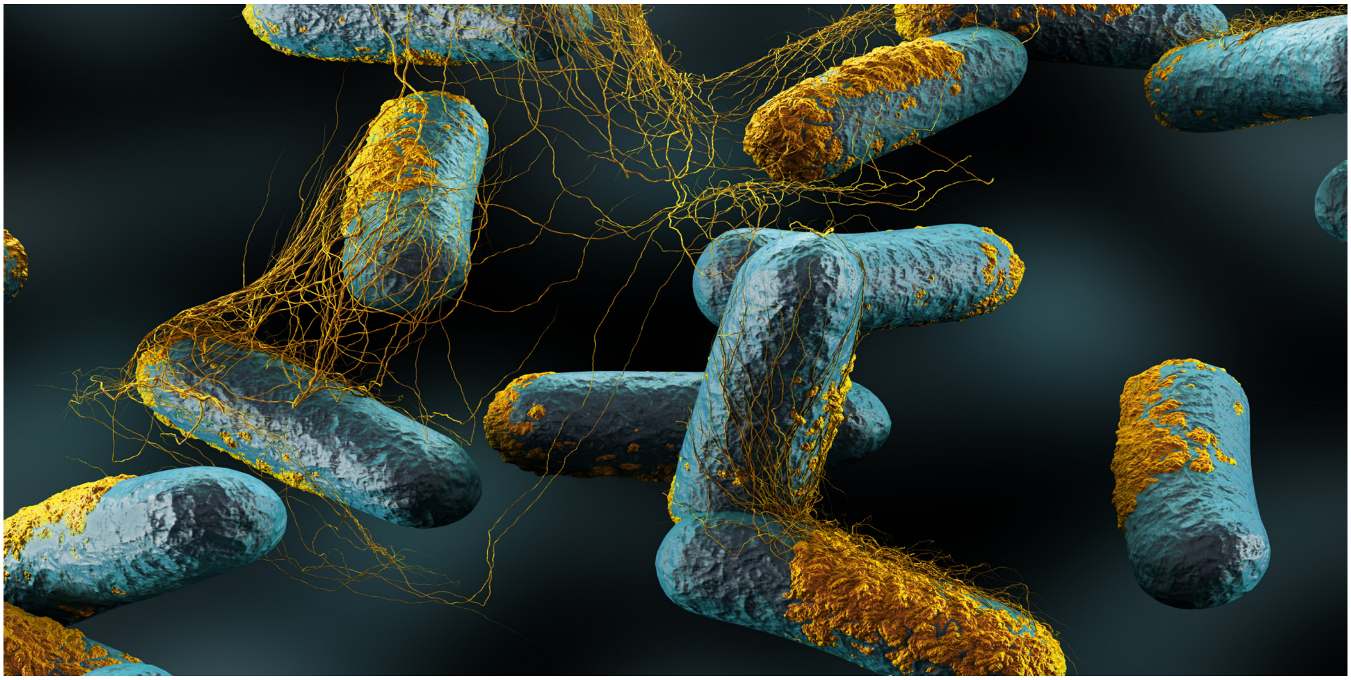
LEVERAGING BIOINFORMATICS TOOLS TO DIVE DEEP INTO *CLOSTRIDIUM PERFRINGENS*



Bill Wolfgang is usually knee-deep in whole genome sequencing data while working in the Bacteriology Laboratory in New York Department of Health's Wadsworth Center. So when he was reviewing laboratory data on a *Clostridium perfringens* sample that had not been sequenced, he began to wonder why.

While *C. perfringens* doesn't get the publicity afforded to *Salmonella* or *Listeria*, this bacterium is one of the most common causes of foodborne illness in the country. There is no national surveillance system for *C. perfringens*—these cases are investigated locally. Investigations begin with Wadsworth epidemiologists—they identify a potential outbreak and send samples to Wadsworth laboratorians for testing. When the laboratory confirms that a sample is *C. perfringens*, epidemiologists typically find a linkage to a contaminated product, and that's that. Why bother with sequencing?

Wadsworth has been a contributor to PulseNet since 1997 and has been a GenomeTrakr laboratory since 2012—two surveillance networks that use the power of whole genome sequencing to link human illnesses with contaminated product. Bill and his team wondered if, by not sequencing *C. perfringens*, they could be missing critical information that could verify New York's illness case clustering. They decided to investigate this hypothesis that WGS-based subtyping can benefit epidemiological investigations of *C. perfringens*, as it has with other enteric organisms. They used FDA funds to support this investigation, using GalaxyTrakr and NCBI Pathogen Detection for data analysis; GalaxyTrakr is a suite of open-source bioinformatics tools that are operated with limited bioinformatics skills and do not require a centralized national surveillance database for comparison. NCBI Pathogen Detection is a centralized national database that can be used for surveillance.

LEVERAGING BIOINFORMATICS TOOLS TO DIVE DEEP INTO *CLOSTRIDIUM PERFRINGENS*

While GalaxyTrakr and NCBI Pathogen Detection had historically been used only for GenomeTrakr organisms (e.g., *Salmonella*, *E. coli*, *Listeria*), Bill and his team wanted to show their broader applicability in other pathogen investigations.

Bill and his team retrospectively analyzed 52 patient and food samples associated with 13 foodborne outbreaks and 24 sporadic samples collected over the past decade and built phylogenetic trees using both NCBI Pathogen Detection and GalaxyTrakr's CFSAN Single Nucleotide Polymorphism (SNP) pipeline. This analysis helped determine if the phylogenetic clustering aligns with the epidemiologically-defined outbreaks.

They found that the genomic clustering lined up well with the epidemiologic clustering. The phylogenetic trees generated from both the SNP pipeline and Pathogen Detection agreed with one another. They discovered a restricted SNP distance between 0-20 SNPs within a cluster that is associated with an outbreak, which can help identify a sample included in a cluster from a sporadic case. Out of the 13 outbreaks associated with 52 of these samples, there were four outbreaks that had samples included epidemiologically that were not genetically related, as well as two outbreaks that had no samples that were genetically related. This WGS data was of great value to their epidemiological colleagues, helping to better define outbreaks and hopefully find related cases and sources of contamination more quickly.

For Bill and his team, this successful project relied on the great support received from GenomeTrakr and those supporting the GalaxyTrakr SNP pipeline. Their willingness to quickly and efficiently answer questions or issues that occurred allowed Bill's team to successfully move forward. Their findings were published in the January 2021 issue of Journal of Clinical Microbiology.[1] Bill and his team plan to use GalaxyTrakr tools during future *C. perfringens* investigations to assist their epidemiologists and protect public health.

REFERENCE

1. Carey J, Cole J, Venkata SLG, Hoyt H, Mingle L, Nicholas D, Musser KA, Wolfgang WJ. Determination of Genomic Epidemiology of Historical Clostridium perfringens Outbreaks in New York State by Use of Two Web-Based Platforms: National Center for Biotechnology Information Pathogen Detection and FDA GalaxyTrakr. J Clin Microbiol. 2021 Jan 21;59(2):e02200-20. doi: 10.1128/JCM.02200-20. PMID: 33177125; PMCID: PMC8111140

This publication is supported by the Food and Drug Administration (FDA) of the US Department of Health and Human Services (HHS) as part of a financial assistance award U18FD006937 totaling $XX with 100 percent funded by FDA/HHS. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by FDA/HHS, or the US Government.