epiXact®: Robust bacterial relatedness and outbreak detection pipeline for WGS data

Emma Briars[1], Mohamad Sater[1], Nicole Billings[1], Defne Surujon[1], Ian C. Herriott[1], Grace Cox[1], Alfred Wong[1], Mihae Platt[1], Miriam Huntley[1]

[1] Day Zero Diagnostics, Inc. (Boston, MA)

Introduction

Hospital acquired infections (HAI) pose a threat to patients and are of increasing concern as antibiotic resistance becomes more widespread. The most accurate method for identifying and tracking transmission involves whole genome sequencing (WGS), and subsequent single nucleotide polymorphism (SNP) analysis across clinical isolates. However, the speed and accessibility of WGS analyses has limited the translation of this technology from the research laboratory to reliable use in infection control settings. To enable accurate transmission analysis in a suspected outbreak, we have developed epiXact®, a CLIA-certified, rapid lab service that includes WGS and SNP analysis, with a turnaround time of 1-2 days for high resolution characterization of transmission events. Here, we performed a large-scale validation of the epiXact pipeline to measure its accuracy and reproducibility for clonality detection in both simulated and *in vitro* datasets, representing 5 species that span various Gram-stains, genome sizes, and GC-content.

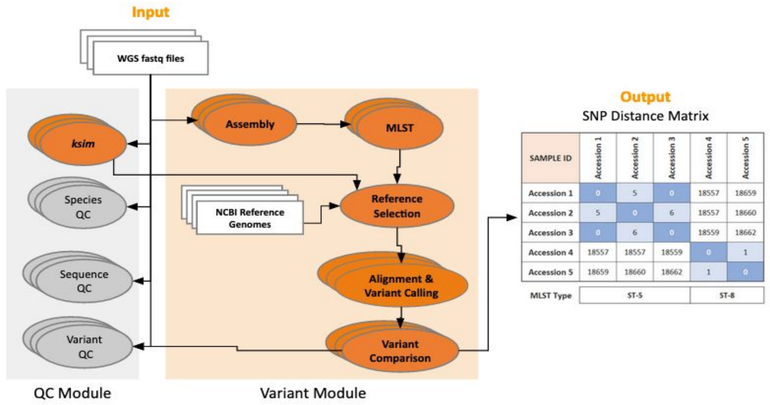
epiXact®

DZD's Illumina-based commercial HAI sequencing and analysis service, is used by multiple partnering hospitals to investigate a wide variety of suspected outbreaks. When infection control (IC) suspects transmission, cultured bacterial samples are sent to DZD. The epiXact test is a "rule-in/rule-out" test of an infection transmission event based on clonality of 2 or more bacterial isolates. This test is designed as a stand-alone test to provide clinicians with a definitive measure of the relatedness of bacterial isolates. Results are reported back in 1-2 days allowing IC to use the definitive genomic evidence to inform decisions regarding ward cleaning, staff screening, and equipment contamination.

epiXact® bioinformatic analysis pipeline

The epiXact bioinformatics pipeline is automated for the following analyses, providing rapid characterization of transmission events:

- Species & MLST Identification
- Assembly & Custom Reference Selection
- Alignment & Variant Calling



CLIA Validation Design

We designed and performed a validation to test the fidelity of the epiXact pipeline in accordance with CLIA regulations. We tested 7 total representative species frequently implicated in HAIs, with varied gram stain, genomic DNA GC content, and genome sizes.

***in vitro* data set:** Clonal and non-clonal bacterial samples from 5 species (Table 1) were sequenced with an Illumina iSeq100 to generate a blinded test dataset of 76 libraries.
***in silico* data set:** A total of 40 simulated sequenced libraries were generated across 5 different species (Table 2). For each species, we simulated 8 sequenced libraries from a representative NCBI reference genome, where 1 sequenced library represents the reference strain, and the other 7 sequenced libraries have between 5 and 200 simulated SNPs introduced to their sequences (relative to the reference).

Species	Validation	Description/ Characteristics	NCBI Reference Genome (*in silico*)
Enterococcus faecium	*in vitro, in silico*	Gram positive, low GC content, 2.9MB	GCF_001721005
Escherichia coli	*in vitro*	Gram negative, medium GC content, 4.6MB	N/A
Pseudomonas aeruginosa	*in vitro, in silico*	Gram negative, high GC content, 6.5MB	GCF_003028335
Staphylococcus aureus	*in vitro, in silico*	Gram positive, low GC content, 2.8MB	GCF_001278745
Clostridioides (Clostridium) difficile	*in vitro*	Gram positive, low GC content, 2.8MB	N/A
Klebsiella pneumoniae	*in silico*	Gram negative, high GC content, 5.2MB	GCF_002258055
Acinetobacter baumannii	*in silico*	Gram negative, low GC content, 4MB	GCF_001908295

Table 1: Species used for *in silico* and *in vitro* validation studies. Species span Gram-strains, GC content, and genome size.

Pipeline Evaluation: The data was input into the automated epiXact pipeline, which computed the number of SNP differences between pairs of samples to quantify relatedness. Pairs of samples were classified as clonally related, closely related, or not closely related depending on the pairwise SNP distances.

epiXact® Validation Performance

epiXact demonstrated high accuracy for determining clonality between bacterial isolates in both the *in vitro* and *in silico* data sets, as well as species identification and MLST typing (Table 4, Figure 1).

Test Output	Metric	Value
Clonality	**Analytical Sensitivity**	100.00%
Clonality	**Analytical Specificity**	*98.53%**
Clonality	**Accuracy**	98.73%
Clonality	**Precision**	100.00%
Species Identification	**Accuracy**	100.00%
MLST Assignment	**Accuracy**	98.2%
***in silico* Clonality**	**Average Pos. Pred. Value**	100.00%
***in silico* Clonality**	**Average Recall**	100.00%
***in silico* SNP Distance**	**Mean Absolute Error**	2.6 SNPs

Table 4: Summary of epiXact performance across large scale validation including *in vitro* and *in silico* simulated datasets. **Additional analysis showed that the false positive calls were due to mis-labelled data in the CDC AR Isolate Bank, suggesting that epiXact was able to successfully identify an unmarked case of clonal pairs in the CDC AR Isolate Bank dataset (communicated to the CDC).*

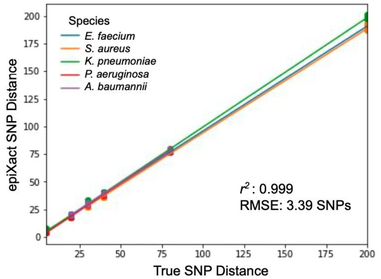


Figure 1: Performance of epiXact in determining pairwise SNP distance between samples (*in silico*)

epiXact® in Action

epiXact uncovers source of deadly *Burkholderia* cluster

BWH Infection Control Notices an Increase in *Burkholderia cepacia* complex infections
In December 2020, 3 cardiothoracic intensive care unit (CTICU) patients at Brigham and Women's Hospital were identified with hospital-onset *Burkholderia cepacia* complex infections. Infection control enacted preventative interventions, but soon after, 2 additional CITCU patients with *B. cepacia* complex infections were identified.

epiXact® Utilized to Find Clonally Related Patient Samples
Whole genome sequencing was performed on 13 CTICU-associated *B. cepacia* complex clinical specimens (November 2020 – February 2021). All samples were processed with an RUO epiXact pipeline to determine clonality. **epiXact identified 6 clonally related isolates**, speciated as *Burkholderia contaminans.* A key commonality amongst the WGS-confirmed cluster cases was that all infections occurred in patients on extracorporeal membrane oxygenation (ECMO).

Investigation of ECMO Water Heater Devices
Specimens from the water and surfaces of all 9 of the hospital's ECMO heater devices in service were cultured, and samples that grew *B. contaminans* were sent for WGS. **epiXact identified that 9 water samples were clonal to the cluster clinical isolates** (Figure 2).

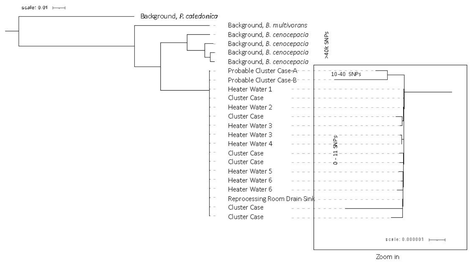


Figure 2: Maximum likelihood phylogeny of *B. cepacia* complex cluster and non-cluster isolates. Branch length and scale bar represents the number of substitutions per site. Same species isolates are grouped in same color boxes. The SNP count indicates the min. or range of pairwise SNP distance between the isolates.

Summary: Clonal *Burkholderia contaminans* samples eradicated in ECMO patients
epiXact identified 6 clonally related *B. cepacia* complex isolates from patients in the CTICU that were also clonal to samples isolated from ECMO heater devices. Infection control was able to use this information to implement targeted infection control measures. In the 12 months since then, no new cases of clonal samples have been reported.

Conclusions

Overall, epiXact demonstrated high accuracy for determining clonality between bacterial isolates in *in silico* and *in vitro* validations. This CLIA validated pipeline offers rapid, accurate, and actionable results, therefore making it a highly relevant tool for infection control. We have also shown (before CLIA validation) how epiXact can be used to link cases of hospital-onset infections and aid infection control in identifying the transmission link.

References

Rhee et al. Clinical Infectious Diseases 2022 March. *B. cepacia* investigation
Zellmer et al. Clinical Infectious Diseases 2021 June, Vol 72, Issue 11.
Shenoy et al. Infection Control & Hospital Epidemiology 2020 May;41(5):531-538
DeFilipp et al. The New England Journal of Medicine 2019 October.

Contact

Emma Briars, PhD: emma@dayzerodiagnostics.com
epiXact Team: epiXact@dayzerodiagnostics.com
URL: https://dayzerolab.com/

DAYZERO DIAGNOSTICS