CPHM-FRIDAY-293

A Universal Nucleic Acid Extraction Process for a Bacterial Gram-Agnostic Fully-Automated Isolate Whole Genome Sequencing Workflow

Clear Labs
(650) 257-3304
derreck.carter-house@clearlabs.com

Derreck Carter-House, Janki Patel, Andrew Pham, Sam Hoeffel, Michael Balamotis, Norma Ordaz-Yuan, Rick Kim, Shadi Shokralla, Anay Campos, James Maloney, Atul Singh, Andrew Lin, David Tran, Akshay Paropkari, Adam Allred, Charles Vaske, Prasanna Thwar, and Ramin Khaksar
Clear Labs, 1559 Industrial Rd., San Carlos, CA 94070

INTRODUCTION

Isolate whole genome sequencing (iWGS) is becoming more widely used as the method of choice to obtain genome-wide high-resolution data for investigating outbreaks and identification of antibiotic resistant bacteria in the public health and healthcare settings. Automation of the entire iWGS workflow, from sample extraction to library preparation and sequencing (isolate to results), would significantly alleviate the burden of laboratorians, freeing them for more critical functions. An automation-friendly and effective extraction method for diverse bacterial strains ensuring sufficient quality and quantity for downstream sequencing is crucial for a successful end-to-end automated workflow. Here, we describe a proprietary extraction method that has been developed and integrated into the Microbial Surveillance WGS application on our turnkey Clear Dx™ platform.

Figure 1. Clear Dx Microbial Surveillance platform

Comparison Features	Current Manual iWGS Workflow	Clear Dx™ Microbial Surveillance iWGS Workflow
Total run time	**44** hrs	**27** hrs
Hands-on-time	**5** Hours*	**40** mins (*> 80% reduction*)
Number of human touch points	> **20**	**1**
Molecular Biology and Bioinformatic Expertise	**Yes**	**No**
Quantification Steps	**3**	**0** (Built-in Normalizations and Quantification)
Number of Samples Per 8 Hour Shift	**Limited**	**24**

Table 1. Comparison of Hardware, Kits, and Time Reduction with Microbial Surveillance Product. *Estimate based on extraction and library preparation using Illumina DNA Prep Kit

The significant advantages of the Microbial Surveillance Product in terms of hands-on time reduction, expertise requirement and other technical steps are highlighted in Table 1. The product streamlines the entire workflow, replacing multiple devices and kits with a single integrated platform and a single kit from Clear Labs. The results demonstrate a reduction in required hardware from five separate components (NA extraction device, Qubit, Thermocycler, Magnet rack, Sequencer) to just one integrated platform. Similarly, the number of kits needed is reduced from three (Extraction kit, Quantification kit, and Sequencing kit) to only one kit from Clear Labs.

Moreover, the Microbial Surveillance Product significantly improves efficiency by reducing run time from 44 hours to 27 hours. Additionally, the hand-on-time is dramatically decreased from approximately 300 minutes to just 40 minutes, resulting in an impressive 80% reduction. These advancements contribute to increased productivity, robustness and streamlined operations in the context of microbial surveillance and analysis.

METHODS

We conducted an extensive analysis of 882 individual samples consisting of both gram-positive and gram-negative bacterial strains isolated from diverse sources as well as the American Type Culture Collection. Among these 160 unique strains, we used many Salmonella, Escherichia, and Listeria to optimize the extraction conditions which are some of the most commonly sequenced organisms in the CDC PulseNet Program.

To assess the required bacterial cell input of our lysis solution, we initially developed a prototype formulation that incorporated all the desired components for the final product. This allowed us to characterize the cell input into the assay. To optimize the universal lysis process, we employed the Design of Experiments (DOE) tool in JMP software (Version 17). By creating multiple unique conditions, we systematically investigated the effects of thermal, mechanical, enzymatic and chemical components on universal lysis. For each condition, we measured the yield of genomic DNA extracted from the bacteria and subsequently prepared a library for sequencing on Illumina iSeq platform. Additionally, we analyzed the mean sequencing depth obtained from each sample under the DOE conditions.

During the extraction process, we resuspended approximately 10^9 cells or one 1 µl bacterial loopful in 400 µl of our cell suspension buffer, which was designed to protect the DNA against exonuclease activity. Subsequently, we vortexed the mixture at a ratio of 2 parts cell suspension to 1 part beads at approximately 3000 RPM for 30 seconds. Next, we transferred 100 µl of the resulting lysate to a 96-well plate and loaded it onto the Clear Labs Microbial Surveillance System. Within this system, 40 µl of lysate was combined with 40 µl of lysis buffer, followed by a heat incubation step. The lysis buffer balanced pH, linearized, digested, and emulsified proteins and hydrophobic components within the cells. After the second lysis step, genomic DNA was captured using Beckman Coulter SPRIselect beads, washed with ethanol, and then eluted using a pH-balanced resuspension buffer. This eluted DNA was subsequently ready for quantification and library preparation. The outputs from the extraction process included DNA yield determined using PicoGreen fluorescence-based quantification, fragment size assessed through gel electrophoresis on Agilent Tapestation, as well as various sequencing metrics such as the number and quality of reads, median insert size, and depth of genome coverage.

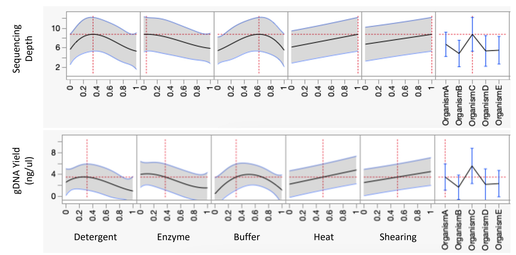

Figure 2 -Example of JMP's DOE results Response Profiler

RESULTS

Our findings indicated that, with few exceptions, most gram-negative bacteria required 1 loopful, while gram-positive bacteria necessitated 4 loopfuls (data not shown). These loopfuls were added to a 400 µl resuspension buffer specifically designed to minimize endonuclease activity following the initial mechanical shearing. We tested the winning DOE condition's performance on numerous different organisms. Figure 2 shows a simulated output of the DOE winning condition and the recommended buffer components or conditions to achieve maximum genomic DNA yield and sequencing depth.

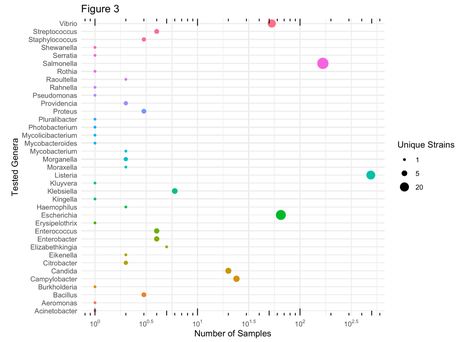

Figure 3 - Counts of all tested genera (n=37) and the number of unique strains within each genus, a total of 160.

Figure 3 provides an overview of the diversity of genera included in our sequencing performance summary. Among the 882 samples, Listeria and Salmonella accounted for several hundred samples each, comprising various unique strains. The figure illustrates the distribution of samples across 37 genera. Additionally, the accompanying table offers a summary of this distribution. Sequencing data were filtered using illumina's guidelines for run quality, e.g. Read 2 quality scores should be greater than 80% over Q30.

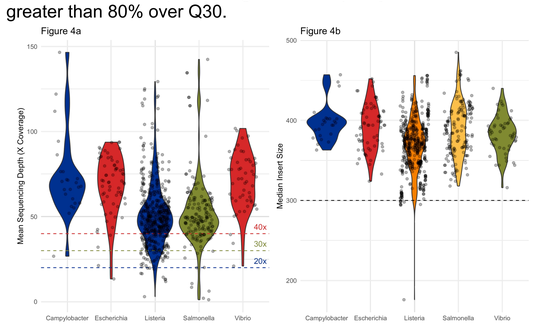

Figure 4 - Mean Sequencing Depth, Median Insert Size and Quality Scores showing sequencing. For Mean Sequencing Depth dashed line color corresponds violin plot color to pulsenset minimum coverage requirement.

Genus	Sample Count	Mean Coverage Pass Rate	Median Insert Size Pass Rate	Quality Score Pass Rate
Campylobacter	24	100.00	100.00	100
Escherichia	65	93.85	100.00	100
Listeria	490	97.96	98.16	100
Salmonella	167	94.01	100.00	100
Vibrio	53	94.34	100.00	100

Table 2. below shows the performance of samples extracted with the optimal DOE Extraction Condition.

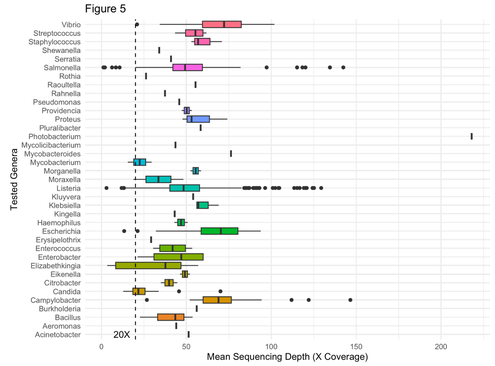

Figure 5 - Mean sequencing depth of PulseNet and HAI pathogens using the optimal DOE extraction condition. 854 of 882 samples had greater than the required 20x Mean Sequencing Depth.

Figure 4a and b present the performance of the top 5 PulseNet organisms using our DOE-formulated lysis. We achieved an impressive passing rate of over 90% for the required sequencing depth. For PulseNet organisms, the CDC mandates a mean sequencing depth of 40x for Escherichia and Vibrio strains, 30x for Salmonella strains, and 20x for Campylobacter and Listeria strains. Furthermore, PulseNet requirements encompass a median insert size of approximately 300 bp or larger, as well as an average quality score of 30 or higher for each base. Nearly 100% of the samples met the median insert size requirement, and all samples achieved a q-score above thirty.

In Figure 5, we expand the application of our DOE-derived lysis to organisms commonly associated with hospital-acquired infections (HAI). The required sequencing depth was determined based on the expected coverage of the genome in the final pooled library, typically aiming for 20 to 30 reads per base. 96.8% of samples successfully achieved this coverage, aligning with our expectations.

CONCLUSION

In conclusion, our study successfully analyzed over 800 bacterial samples, including both gram-negative and gram-positive strains. Through the implementation of our optimized DNA extraction method, which involved a combination of thermal, mechanical, enzymatic and chemical components, we established a baseline cell input requirement for several common healthcare-associated infection (HAI) genera.

We validated the performance of our extraction method on a diverse range of organisms, as illustrated by the sequencing performance summary presented in Figure 2. The top 5 PulseNet organisms, representing significant public health concerns, exhibited a high passing rate for the required sequencing depth, meeting or surpassing the CDC's recommended thresholds. Additionally, our method showed promising results for organisms commonly associated with hospital-acquired infections (HAI), achieving the expected sequencing depth for optimal genome coverage.

These findings highlight the efficacy and reliability of our optimized DNA extraction method utilized in the Clear Labs fully automated system, in providing high-quality genomic DNA for downstream applications. The success of our approach offers valuable insights for the field of microbial surveillance and contributes to the improvement of diagnostic and surveillance techniques in public health and healthcare settings.

For more information regarding the Clear Dx products, please contact inquiries@clearlabs.com

Clear Labs
1559 Industrial Road
San Carlos, CA 94070

Disclaimer: The Clear Dx instrument is for Research Use Only (RUO) and is not intended for diagnostic purposes.

© Copyright 2022, American Association for Clinical Chemistry. All Rights Reserved.