

Fully Integrated Automation of NGS Library Preparation for Foodborne Pathogens

H. Zhu¹, B Kim¹, R. Yasmin¹, G. Kastanis², M. Allard², and R. Montagna¹

¹Rheonix, Inc., Ithaca, NY, ²Center for Food Safety & Applied Nutrition/FDA, College Park, MD

BACKGROUND

The identification of foodborne pathogens is important to prevent and track major outbreaks and to identify potential sources of contamination in food processing plants. Next Generation Sequencing (NGS) is a powerful tool to identify such organisms, but considerable effort is first required to isolate DNA and prepare the required DNA libraries. To simplify the process, DNA isolation and library preparation were integrated on a single instrument that automated both processes. Sequencing was performed on MiSeq instruments.

METHODS

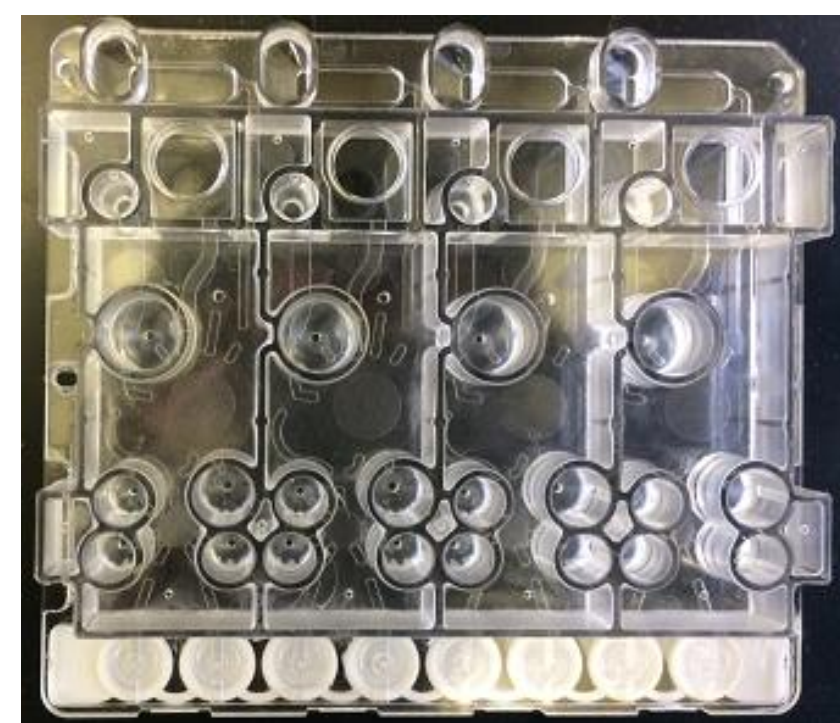
FDA's Center for Food Science & Applied Nutrition (CFSAN) provided 12 strains of *Salmonella* as well as its proficiency panel (PP) consisting of 4 strains of *Salmonella* and 2 strains of *E. coli* O157:H7 (Table I) used to evaluate the ability of various labs to correctly identify these pathogens using NGS. Nextera DNA Preparation reagents (Illumina, San Diego, CA) were used on the Rheonix Encompass *Optimum*TM workstation (Figure 1) to prepare the NGS libraries using the microfluidic Rheonix CARD[®] cartridge (Figure 2).

Figure 1: Rheonix Encompass *Optimum*TM Workstation



The Rheonix Encompass *Optimum*TM workstation's software can automatically control one to six CARD cartridges (i.e., 1-24 individual samples) without user intervention. Pneumatic signals are used to actuate pumps and valves and move paramagnetic magnets in and out of place. The workstation's robotic liquid handler delivers samples and reagents where necessary.

Figure 2: Rheonix CARD[®] Cartridge



The Rheonix CARD[®] Cartridge is an injection molded plastic microfluidic device that can process four individual samples. All pumps, valves, microchannels, reaction and reagent reservoirs are contained within the cartridge.

RESULTS

Purified, well-characterized DNA was initially used to optimize the library preparation portion of the fully automated process. Sequence-ready DNA libraries were prepared from these DNA samples on the Encompass workstation within approximately 3 hours, which included about 30 minutes of hands-on-time. Sequencing of these libraries by CFSAN on MiSeq instruments yielded excellent quality metrics. In the case of the 12 strains of *Salmonella* received from CFSAN, quality scores exceeded Q30 with depth of coverage exceeding 112X (Table I). For the PP samples, the quality scores for both forward and reverse reads exceeded Q30 with the depth of reads in the range of 30X – 40X. In addition, 95-98% of the reads could be mapped to the reference genomes (Table II). Then, additional studies were performed starting with cultured *Salmonella* and *E. coli* microbes to optimize the combined DNA isolation and library preparation on a single instrument. Sequencing of these libraries on MiSeq instruments yielded similar results and quality scores.

Table I: Depth of Coverage and QA Status

| Organism | Depth of Coverage | QA Status |
|---|-------------------|-----------|
| <i>Salmonella enterica subsp. enterica serovar Cubana</i> | 112X | Pass |
| <i>Salmonella enterica subsp. enterica serovar Javiana</i> | 167X | Pass |
| <i>Salmonella enterica subsp. enterica serovar Montevideo</i> | 144X | Pass |
| <i>Salmonella enterica subsp. enterica</i> | 189X | Pass |
| <i>Salmonella enterica subsp. enterica serovar Agona</i> | 152X | Pass |
| <i>Salmonella bongori</i> | 137X | Pass |
| <i>Salmonella enterica subsp. enterica serovar Typhi</i> | 43X | Pass |
| <i>Salmonella enterica subsp. enterica serovar Typhi</i> | 146X | Pass |
| <i>Salmonella enterica subsp. enterica serovar Heidelberg</i> | 54X | Pass |
| <i>Salmonella enterica subsp. enterica serovar Heidelberg</i> | 133X | Pass |
| <i>Salmonella enterica subsp. enterica</i> | 52X | Pass |
| <i>Salmonella enterica subsp. enterica</i> | 136X | Pass |

Table II: Annotation Matrix for Proficiency Samples*

| Expected | SAP17-7299 | SAP17-7399 | SAP17-7699 | SAP17-8290 | ECP17-1298 | ECP17-46 |
|-----------------|------------|------------|------------|------------|------------|-----------|
| Annotated | SAP17-7299 | SAP17-7399 | SAP17-7699 | SAP17-8290 | ECP17-1298 | ECP17-46 |
| Reads | 2,359,472 | 2,306,434 | 2,219,094 | 2,107,160 | 1,058,210 | 2,238,070 |
| % Mapped | 98.19 | 94.66 | 99.03 | 98.06 | 98.81 | 98.81 |
| Seq Length | 35-151 | 35-151 | 35-151 | 35-151 | 35-151 | 35-151 |
| Mean For Q | 33.80 | 35.70 | 35.30 | 35.80 | 35.70 | 35.70 |
| Mean Rev Q | 31.50 | 33.90 | 33.50 | 30.80 | 32.10 | 32.80 |
| Mean Depth | 39.68 | 40.33 | 41.04 | 39.66 | 19.27 | 35.61 |
| SNPs | 0 | 0 | 0 | 0 | 0 | 0 |
| Mean Insert | 204.28 | 214.01 | 219.99 | 241.37 | 267.76 | 224.66 |
| NG50 | 123,951 | 157,127 | 96,251 | 169,719 | 100,019 | 63,498 |
| Genome Fraction | 98.52 | 98.63 | 98.62 | 98.60 | 94.89 | 94.70 |
| Contigs | 189 | 175 | 182 | 202 | 610 | 684 |

*Identity of 2017 PulseNet-GT harmonized Proficiency strains:

SAP17-7299 (*Salmonella enterica enterica* Typhimurium) SAP17-8290 (*Salmonella enterica enterica* Typhimurium)
 SAP17-7399 (*Salmonella enterica enterica* Typhimurium) ECP17-1298 (*Escherichia coli* O157:H7)
 SAP17-7699 (*Salmonella enterica enterica* Typhimurium) ECP17-46 (*Escherichia coli* O157:H7)

CONCLUSIONS

The ability to reduce the hands-on time and improve quality metrics using automated DNA isolation and subsequent NGS library preparation will help to reduce costs, improve turn-around time, and provide consistently reliable data. Automation of the entire process on a single instrument will also significantly reduce the amount of training required as well as the capital equipment costs and required laboratory space as compared to manual processes using multiple pieces of equipment.