A Custom Assay for a Sub-Population of Salmonella Heidelberg

Gilbert Ortiz, Michael Karberg, Adam Allred, Angela Burrell, Michelle Swimley, Amanda Manolis, Thermo Fisher Scientific, 2130 Woodward Street, Austin, Texas 78744

ABSTRACT

Salmonella enterica serovar Heidelberg is among the top ten most common serovars of *Salmonella* (1) associated with poultry products and human gastroenteritis. Because of the risk of sickness from contaminated food, there is regulatory and industrial interest in detecting and eliminating this serovar from poultry processing facilities. In addition, the ability to track contamination from supplier to processor, and between or within multiple processing facilities, would be a valuable tool in outbreak investigation or prevention. Currently, detection methods can discriminate between different serovars of *Salmonella*, but discrimination between strains or sub-populations within the same serovar is often impossible using these methods. To address this gap, we used next generation DNA sequencing to identify the genetic "fingerprint" of a sub-population of *Salmonella* Heidelberg found within a processing facility, and then designed a quantitative PCR assay for its detection.

Using the Ion PGM™ Sequencer and bioinformatics tools, we compared the genomes of 10 different *Salmonella* Heidelberg isolates showing the same PFGE pattern to 116 other Heidelberg genomes available in GenBank. We found unique SNPs within these isolates that were absent from the other Heidelberg genomes. Based on these differences, we designed, developed, and manufactured a novel TaqMan® Heidelberg SNP Genotyping assay that specifically detects the Heidelberg sub-population.

INTRODUCTION

Thermo Fisher engaged its Custom Assay Development process at the request of a collaborator to develop a qPCR assay to detect *S*. Heidelberg. The process consists of 1) obtaining samples from the collaborator, 2) sequencing, 3) using in-house bioinformatics capabilities to design a TaqMan® qPCR assay, 4) testing and optimizing the assay, and 5) manufacturing the assay to meet the customer's unique needs (Figure 1). Standard completion times range from 4 to 12 weeks depending on complexity and format of final manufactured assay.

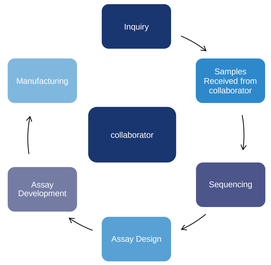

Figure 1. Custom Assay Development Workflow

MATERIALS AND METHODS

Sequencing: 18 *Salmonella* samples were received from a collaborator and nucleic acid was obtained from 10 of the samples using the PureLink® Genomic DNA Mini kit. 100 ng of nucleic acid was input into the Ion Xpress™ Plus Fragment Library Kit. The recommended protocol for a 300 bp sequencing reaction was followed. During library preparation, samples were ligated with barcoded adapters, allowing for pooling of multiple samples in a single sequencing reaction. The resulting libraries were quantified using the Ion Library TaqMan® Quantitation Kit, followed by template preparation using the Ion OneTouch™ 200 Template Kit v2 DL and the Ion OneTouch™ instrument and enrichment system. After template enrichment, samples were loaded on an Ion 318™ chip and sequenced on the Ion PGM™ System using the Ion PGM™ Sequencing 300 Kit (640 flows).

Bioinformatic analysis and assay design: Sequence reads for the 10 samples were assembled into draft genomes using MIRA (2). 108 additional Heidelberg genomes were collected from Genbank and other previous sequencing projects and analyzed for conserved regions that were common to all Heidelberg genomes. SNPs were identified in these regions and evaluated for their ability to distinguish the customer's genome sequences from all others. Three such SNPs were identified, and TaqMan® assays were designed to specifically target these SNPs.

qPCR: All 18 collaborator samples were analyzed. Allelic discrimination qPCR was performed using the Applied Biosystems 7500 Fast Real-Time PCR System and analyzed using the Applied Biosystems Sequence Detection Software v1.4.0. Reactions were performed using the standard conditions for TaqMan® Environmental Master Mix, and contained up to 10,000 copies of genomic DNA per reaction. The Manual Ct, Threshold was set at 0.2, and the Auto Baseline set to automatic. The FAM detector will identify a negative result, while the VIC detector at an early Ct will identify a positive sample.

Sample Number	*S. enterica* serovar	Sample Source	Year
1	Heidelberg	Environmental	2012
2	Heidelberg	Environmental	2012
3	Heidelberg	Environmental	2012
4	Heidelberg	Environmental	2012
5	Heidelberg	Environmental	2012
6	Heidelberg	Environmental	2012
7	Heidelberg	Environmental	2012
8	Heidelberg	Environmental	2012
9	Heidelberg	Environmental	2012
10	Heidelberg	Environmental	2012
11	Heidelberg	Environmental	2012
12	Heidelberg	Environmental	2012
13	Heidelberg	Environmental	2012
14	Heidelberg	Environmental	2012
15	Heidelberg	Environmental	2012
16	Heidelberg	Environmental	2012
17	Heidelberg	Environmental	2012
18	Heidelberg	Environmental	2012

Figure 2. Samples Used for This Study

Eighteen samples were received from a collaborator that were shown to have the same PFGE pattern. All strains were isolated from food production samples in 2012 and were identified as *Salmonella enterica* ser. Heidelberg. Of the 18 *S.* Heidelberg isolates, 10 were sequenced and bioinformatically compared to 116 other Heidelberg genomes. All of the collaborator samples were shown to have similar genomes to each other but were different from the *S.* Heidelberg genomes (Figure 5). Based on the similarities a TaqMan® qPCR assay was designed and tested on all 18 collaborator samples, along with selected exclusionary *S.* Heidelberg strains.

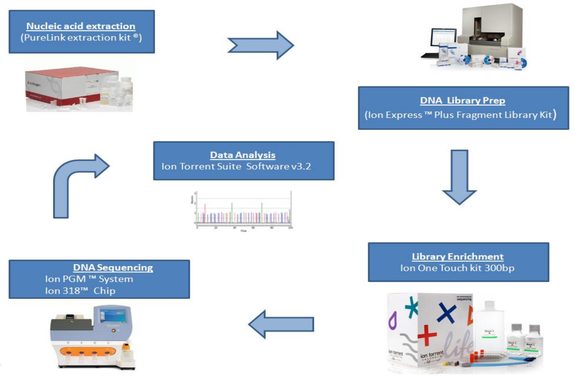

Figure 3. Sequencing Workflow

RESULTS

Sequencing Metric	Min	Max	Average
Read Length	192	217 bp	207
Coverage	41x	57x	49x
Number of Contigs	71	157	107
N50	85,952	2.0x 10^5	152,399

Figure 4. Sequencing Data

The above table shows results from all sequencing runs. Read length ranged from 192-217 base pairs (bp). Sequencing coverage was greater than 40X for all samples resulting in an average of 107 contigs for each sample after assembly. The N50 value was calculated for all samples as well. N50, reported in base pairs, is a weighted median statistic such that 50% of the entire assembly is contained in contigs equal to or larger than this value. For all assemblies N50 values were greater than 85,000 bp.

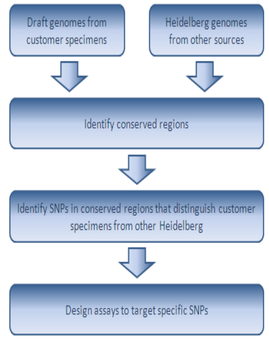

Figure 4. Strategy for design of specific TaqMan® assays

Sequences obtained from the Ion PGM™ sequencer, as well as sequences downloaded from Genbank and from previous sequencing projects, were used to design TaqMan® primers and probes for targeting specific SNPs within conserved regions.

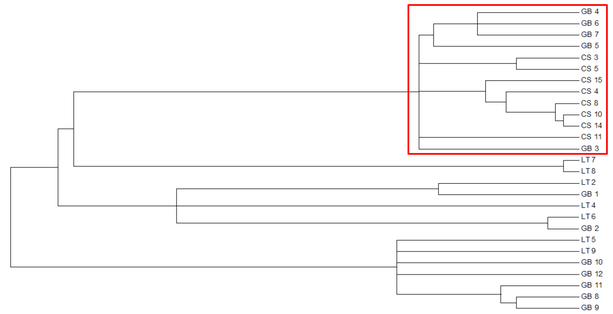

Figure 5. Phylogenetic tree based on SNPs

470 non-singleton SNPs were identified in core regions which were common to all 116 Heidelberg genomes analyzed. A phylogenetic tree was generated for a subset of these genomes using the neighbor-joining method in MEGA5 (3). The cluster outlined in red includes all the genomes obtained from sequencing collaborator specimens as well as closely related genomes available in Genbank.

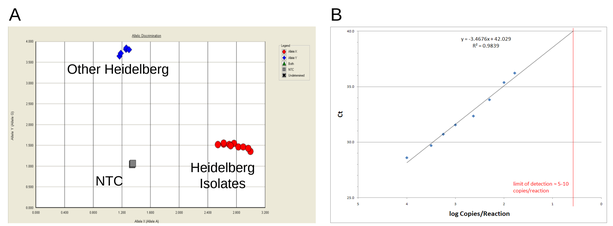

Figure 6. Assay Characteristics

A) qPCR analysis. All 18 confirmed *S.* Heidelberg samples were detected using the assay (red dataset). The assay also includes an internal positive control (IPC) that is used to help gauge the quality of the qPCR data. B) Limit of Detection (LOD). The *S.* Heidelberg assay was used to determine the LOD using DNA isolated from the strains supplied by the collaborator. The results are displayed as an average of all 18 strains. For this assay, a CT cutoff of 40 would equate to an LOD of ~5-10 CFUs detected per reaction. We recommend additional experiments be performed to set an appropriate CT cutoff for each sample of interest.

Inclusivity	Assay Result
S. Heidelberg sample 1	Positive
S. Heidelberg sample 2	Positive
S. Heidelberg sample 3	Positive
S. Heidelberg sample 4	Positive
S. Heidelberg sample 5	Positive
S. Heidelberg sample 6	Positive
S. Heidelberg sample 7	Positive
S. Heidelberg sample 8	Positive
S. Heidelberg sample 9	Positive
S. Heidelberg sample 10	Positive
S. Heidelberg sample 11	Positive
S. Heidelberg sample 12	Positive
S. Heidelberg sample 13	Positive
S. Heidelberg sample 14	Positive
S. Heidelberg sample 15	Positive
S. Heidelberg sample 16	Positive
S. Heidelberg sample 17	Positive
S. Heidelberg sample 18	Positive

Exclusivity	Assay Result
Cronbacter sakazakii	Negative
Salmonella enterica ser. Typhimurium	Negative
Campylobacter coli	Negative
Oncorhyncus keta	Negative

Figure 7. Specificity

The assay was tested against an exclusion panel of various serotypes of *Salmonella* and related pathogens and found to be specific for *Salmonella* Heidelberg.

CONCLUSIONS

- The Custom Assay Development process was used to design a TaqMan® qPCR SNP assay for a specific strain of *S.* Heidelberg at the request of a collaborator.
- 10 samples were sequenced within 3 days of receiving pure cultures from the collaborator on the Ion PGM™ System.
- Data analysis, including genome assembly and design of specific TaqMan® primers and probes, was completed in eight days.
- Assay development and manufacturing was completed in 10 days.
- The selected assay detects less than 10 CFU per reaction of *S.* Heidelberg genomic DNA.
- The assay discriminates between collaborator-provided strains of *S.* Heidelberg and other *Salmonella* serovars.

REFERENCES

1) National Salmonella Surveillance Annual Report, 2011. Retrieved August 12, 2014, from http://www.cdc.gov/ncezid/dfwed/PDFs/salmonella-annual-report-2011-508c.pdf

2) Chevereux, B et al. Genome sequence assembly using trace signals and additional sequence information. Proceedings of the German Conference on Bioinformatics (1999).

3) Tamura K et al. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution (2011)

TRADEMARKS/LICENSING