Development of a qPCR-based assay for the detection of *Escherichia coli*

Rajesh Vasani
Mentored by William Harty

Introduction

Testing samples for bacteria is a process that has been done for several decades. Popular methods include multiple-tube fermentation (MTF) and membrane filter (MF) (Baudart et al., 2002).

The MTF method streaks water samples onto petri dishes and looks for gas production. For the MF method, water samples are collected and passed through a special membrane with 0.45-micron sized pores that trap microorganisms for extraction. Both methods take the extracted samples and apply them to a growth medium to grow over a period of 48 hours. The drawback of these tests is that they can take several days to produce results and potential contamination of the tested bacteria could have easily spread while waiting for the 48-hour growth time of the bacteria. A more modern, faster, and more specific technology for testing for the presence of bacteria is the use of quantitative Polymerase Chain Reaction (qPCR). This study analyzed the efficacy of using qPCR to detect the presence of *E. coli* in water samples. The benefit of using qPCR for bacteria detection instead of MTF and MF is its significantly decreased processing time for obtaining results.

Materials & Methods

The targeted bacteria for this project was *Escherichia coli*, more commonly known as *E. coli*. The first step taken was to locate potential genes in the *E. coli* bacteria that could be targeted in the qPCR reaction. Since the primer and probes used in qPCR target specific DNA regions, it was important to locate a region of DNA that is unique to *E. coli* bacteria. The uidA gene was chosen because it is the gene detected in 97.7% of the 435 *E. coli* colonies commonly found in water sources (Martins et al., 1993). Next, the uidA gene sequence was used in a primer blast search, an online tool by the National Center for Biotechnology Information (NCBI), to find suitable sites for the primers to be placed on a gene sequence. The primer blast search resulted in three suitable primer pairs.

Next Polymerase Chain Reaction (PCR) verification of the primers was performed. The three primer pairs were prepared with a master mix and run through PCR to determine the most optimal primer pair to use for the qPCR assay . In this process of PCR, spiked samples containing *E. coli* were used to test the primers.

During PCR the DNA in the reaction splits from double to single strand during the denaturization phase. Next, in the annealing phase, primers attach to their targeted region within the targeted gene. In the last phase, elongation, an enzyme called DNA polymerase extends the DNA sequence from the primer. This effectively doubles the amount of the targeted region of DNA in the reaction. This is repeated until there are hundreds of thousands of copies of the targeted DNA region.

Materials & Methods (cont.)

The resulting PCR reactions of each of the three primer pairs were put through gel electrophoresis. Gel electrophoresis uses electricity to move the contents of DNA by the number of base pairs it has, with smaller base pair amounts traveling farther than larger base pair amounts. Primer pair three was removed due to issues with repeating sections of DNA. The result of the gel electrophoresis (Figure 1) show that primer set two's results have less smear and are more aligned than the results of primer set one. Due to its favorable testing outcome set two was used in the qPCR test.

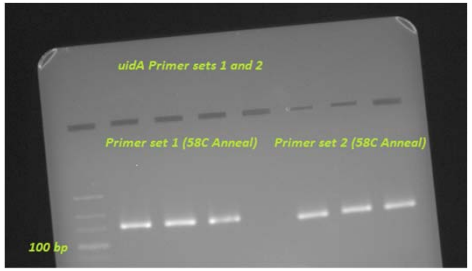


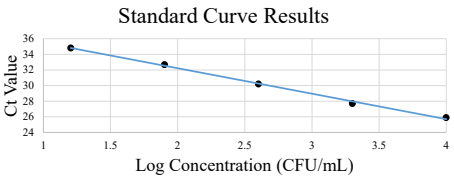
Figure 1: Results of gel electrophoresis. Set one shows slight misalignment and smearing. While set two has much less smearing and better alignment than set one does. Thus, primer set two was determined to be a better candidate for the PCR testing.

A fluorescent probe was designed and ordered for primer set two. The water sample and probe were added to the master mix. In qPCR, the process is the same as the PCR steps previously mentioned with the exception that probes anneal next to the primer during the annealing phase of qPCR. When DNA Polymerase elongates the DNA, the probes fluoresce in the reaction. This fluorescence is recorded by the qPCR machine, which produces a graph of the fluorescence, measured in Relative Fluorescence Units (RFU), at each cycle of the qPCR reaction. The cycle threshold (Ct) value is the cycle of PCR at which the concentration of *E. coli* was "detected" past a predefined cycle threshold value. A standard curve was developed by running qPCR tests with varying amounts of initial *E. coli* concentrations and recording their Ct values.

To ensure the validity of the standard curve, a One-way ANOVA test was conducted. Using each starting concentration group of 10,000 Colony Forming Units (CFU), 2,000 CFU, 400 CFU, 80 CFU, and 16 CFU. Each group had three tests conducted. The Ct value (in cycles) was recorded for each test. The null hypothesis was that there is no mean difference between any of the concentration groups. The One-way ANOVA showed there was a significant difference in the mean Ct value between the 10,000 CFU (M = 25.9, SD = 0.119), 2,000 CFU (M = 27.7, SD = 0.049), 400 CFU (M = 30.2, SD = 0.049), 80 CFU (M = 32.7, SD = 0.041), 16 CFU (M = 34.8, SD = 0.143), $F(1,10) = 4770$, $p < 0.0001$. Using an alpha level of 0.05 the null hypothesis was rejected, indicating the desired result, that there was a statistically significant difference in the Ct values of varying concentration groups.

Results

The curves produced from the qPCR data were analyzed. The cycle number at which a sample reached the threshold line, referred to as the Ct value, was recorded along with the log value of the bacterial concentration at that point. Plotting this data produced a standard curve for the assay (Graph 1). Water samples taken from Deer Creek and Gunpowder River produced no fluorescence, suggesting zero bacterial presence. Tests with samples spiked with *E. coli* returned Ct values which suggested a positive result, confirming the accuracy of the standard curve.



Graph 1: The standard curve plot of Ct count vs Log amount. The equation for the linear model is $y = -3.2654x + 38.749$ with a correlation coefficient of 0.9986.

Conclusion

A qPCR assay targeting *E.coli* was developed and shown to be quick and highly efficient. This fast and reliable test allows for quick response to potential contaminations of target diseases. The basic principles in this assay design could be applied to other gene targets of various organisms/targets. Additional modifications to this assay could include multiplexing to detect and differentiate between multiple different targets within the same reaction.

References

Baudart, J., de-Roubin, M., Laurent, P., Rompre, A., & Servais, P. (2002) Detection and enumeration of coliforms in drinking water: Current methods and emerging approaches. *Journal of Microbiological Methods*, *49*(1), 31–54. https://pubmed.ncbi.nlm.nih.gov/11777581/

Martins, M. T., Rivera, I. G., Clark, D. L., Stewart, M. H., Wolfe, R. L., & Olson, B. H. (1993). Distribution of uidA gene sequences in Escherichia coli isolates in water sources and comparison with the expression of beta-glucuronidase activity in 4-methylumbelliferyl-beta-D-glucuronide media. *Applied and Environmental Microbiology*, *59*(7), 2271–2276. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC182268/