Targeted Nanopore Sequencing of the O-Antigen Region in *E. coli* using CRISPR Cleavage at *gndA*

Megan Lyons[1], Erika Biernbaum[2], Tyler Chandross-Cohen[2], Yara Cavalcante[2], Jasna Kovac[2], and Edward Dudley[2]

Department of Chemistry and Biochemistry[1], Duquesne University, Pittsburgh, PA 15282

Department of Food Science[2], The Pennsylvania State University, University Park, PA 16802

Abstract

The rise of antimicrobial resistant bacteria due to overuse of antibiotics, emphasizes a need for a rapid identification method which is critical for outbreak control. The O-antigen is a polysaccharide chain that is located in the lipopolysaccharide layer of *Escherichia coli*. The O-antigen is unique to each serogroup allowing for the identification of different strains of *E. coli*. To selectively sequence the O-antigen region in *E. coli*, we employed a novel system that allows for the use of a CRISPR-Cas9 targeted enrichment and Oxford Nanopore sequencing. We used CRISPR-Cas9 technology to create a double stranded break at a conserved housekeeping gene, *gndA*, which flanks the O-antigen region. This then allows for the sequencing of the O-antigen region using Oxford Nanopore sequencing. Single-guide RNAs (sgRNAs) were previously designed to cleave at *gndA* in the CRISPR reaction. After the CRISPR cleavage, Nanopore sequencing was completed, and the spectra were analyzed. Here, we first used a monoculture of *E. coli* O1 to optimize the protocol. This method allows for the rapid long-read sequencing of the O-antigen region distinguishing each serogroup. Future work will include using dual gRNAs to reduce off-target effects of cutting at *gndA* and using this system with more complex samples such as wastewater to potentially identify new O-antigens.

Background

O-Antigen in *E. coli*

The O-antigen is a repeating oligosaccharide unit located in Gram-negative bacteria.[1] It is important because it helps identify the serotype of bacteria present. The O-antigen aids in immune evasion and can attach to host cells during an infection. Sequencing the O-antigen can help identify the serogroups present in different environmental samples.

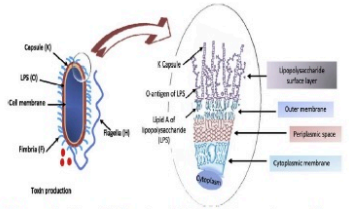


Figure 1. *E. coli* structure highlighting major surface components including the O-antigen . Figure adapted from Swelum et al.

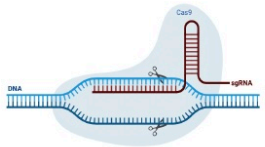


Figure 2. Targeted DNA cleavage by CRISPR-Cas9. Figure created with BioRender.com

CRISPR-Cas9

CRISPR-Cas9 is a bacterial adaptive immune response in which foreign DNA is integrated into its own. CRISPR utilizes a Cas9 enzyme which is an endonuclease that cuts DNA and a single guide RNA (sgRNA) that directs the Cas9 to the target DNA. This method allows for site-specific genetic modifications. Here, it was used for sequencing the O-antigen region.[2]

gndA

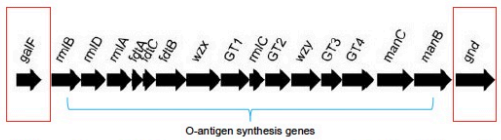
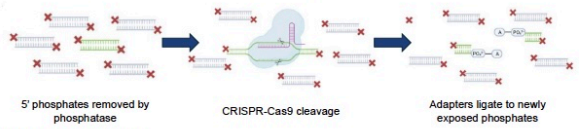


Figure 3. Example gene cluster of O-antigen synthesis genes in *E. coli*. Figure adapted from Perepelov et al.

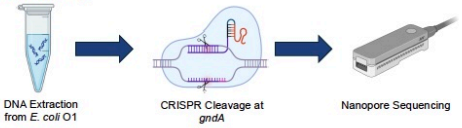
O-antigen synthesis genes are located between two housekeeping genes, *galF* and *gndA*.[3] The genes for O-antigen synthesis are located in a gene cluster. While the O-antigen synthesis genes are variable, these flanking housekeeping genes are conserved through most serogroups of *E. coli*.

Methods

CRISPR Cleavage Reaction for Sequencing[4]



Nanopore Sequencing



Bioinformatic Analysis Tools:

- Bowtie2
- Minimap2
- IGV
- Samtools

Created with BioRender.com

Results

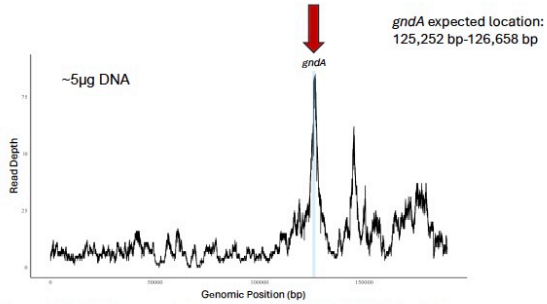


Figure 4. Oxford Nanopore Sequencing Depth of the CRISPR reaction using 5 µg of DNA. The red arrow represents the *gndA* gene. It is expected to see the highest peak at the gene of interest.

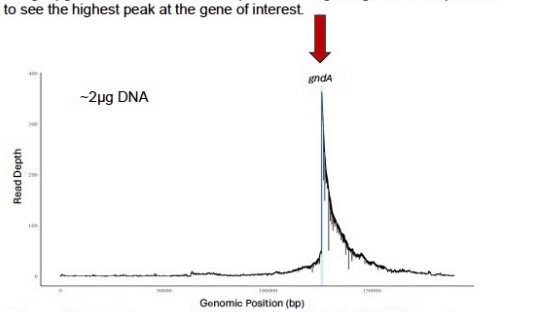


Figure 5. Oxford Nanopore Sequencing Depth of the CRISPR reaction using 2 µg of DNA. The red arrow represents the *gndA* gene. It is expected to see the highest peak at the gene of interest.

Conclusions

- The CRISPR protocol for *gndA* was confirmed.
- The Oxford Nanopore Sequencing protocol was optimized with ~2 µg DNA.
- The sgRNA was designed to cleave at *galF* but this did not work efficiently.
- Because the *galF* CRISPR reaction was unsuccessful, more than the O-antigen region was sequenced.
- The CRISPR reaction has a high specificity that allows for precise targets which will create an accurate sequencing of the O-antigen region
- Methods for the CRISPR cleavage and Nanopore sequencing are still being optimized, however this provides evidence that this novel method of using CRISPR to enrich a gene for sequencing will produce promising results as an efficient and rapid identification technique.

Future Directions

- Since the CRISPR reaction targeting *galF* was unsuccessful, we plan to redesign the sgRNA.
- Optimize the *galF* CRISPR protocol to allow for complete sequencing of the O-antigen region in *E. coli*.
- Use this novel method in wastewater, food, and environmental samples to detect which serogroup of *E. coli* is present.
- Use this method to detect and identify new O-antigen serogroups.

Acknowledgements

- Funding from USDA-NIFA grant 2021-67037-32628, Engaging Students from Undergraduate-Centric Institutions in Research, Informatics, and Experimental Opportunities in Food Microbiology
- USDA Food Microbiology REEU Program
- The Pennsylvania State University Department of Food Science
- Duquesne University Women in STEM

PennState

USDA National Institute of Food and Agriculture
U.S. DEPARTMENT OF AGRICULTURE

DUQUESNE UNIVERSITY

References

(1)Swelum, A. A.; Elbestawy, A. R.; El-Saadony, M. T.; Hussein, E. O. S.; Alhotan, R.; Suliman, G. M.; Taha, A. E.; Ba-Awadh, H.; El-Tarabily, K. A.; Abd El-Hack, M. E. Ways to Minimize Bacterial Infections, with Special Reference to Escherichia Coli, to Cope with the First-Week Mortality in Chicks: An Updated Overview. *Poultry Science* **2021**, *100* (5), 101039. https://doi.org/10.1016/j.psj.2021.101039.

(2) Li, T.; Yang, Y.; Qi, H.; Cui, W.; Zhang, L.; Fu, X.; He, X.; Liu, M.; Li, P.; Yu, T. CRISPR/Cas9 Therapeutics: Progress and Prospects. *Signal Transduction and Targeted Therapy* **2023**, *8* (1), 1–23. https://doi.org/10.1038/s41392-023-01309-7.

(3) Perepelov, A. V.; Song, Y.; Zhu, Y.; Shashkov, A. S.; Filatov, A. V.; Hu, B. Structure and Gene Cluster of the O-Antigen of Escherichia Coli Strain SDLZB008. *Carbohydrate Research* **2020**, *498*, 108154–108154. https://doi.org/10.1016/j.carres.2020.108154.

(4) Cottingham, H.; Judd, L. M.; Wisniewski, J. A.; Wick, R. R.; Stanton, T. Holt, K. E.; Hawkey, J. Targeted Sequencing of *Enterobacterales* Bacteria Nanopore Technologies. *mSystems* **2025**, *10* (2). https://doi.org/10.1128/m