

Investigating Environmental Microbes in Newcastle Using Nanopore Sequencing

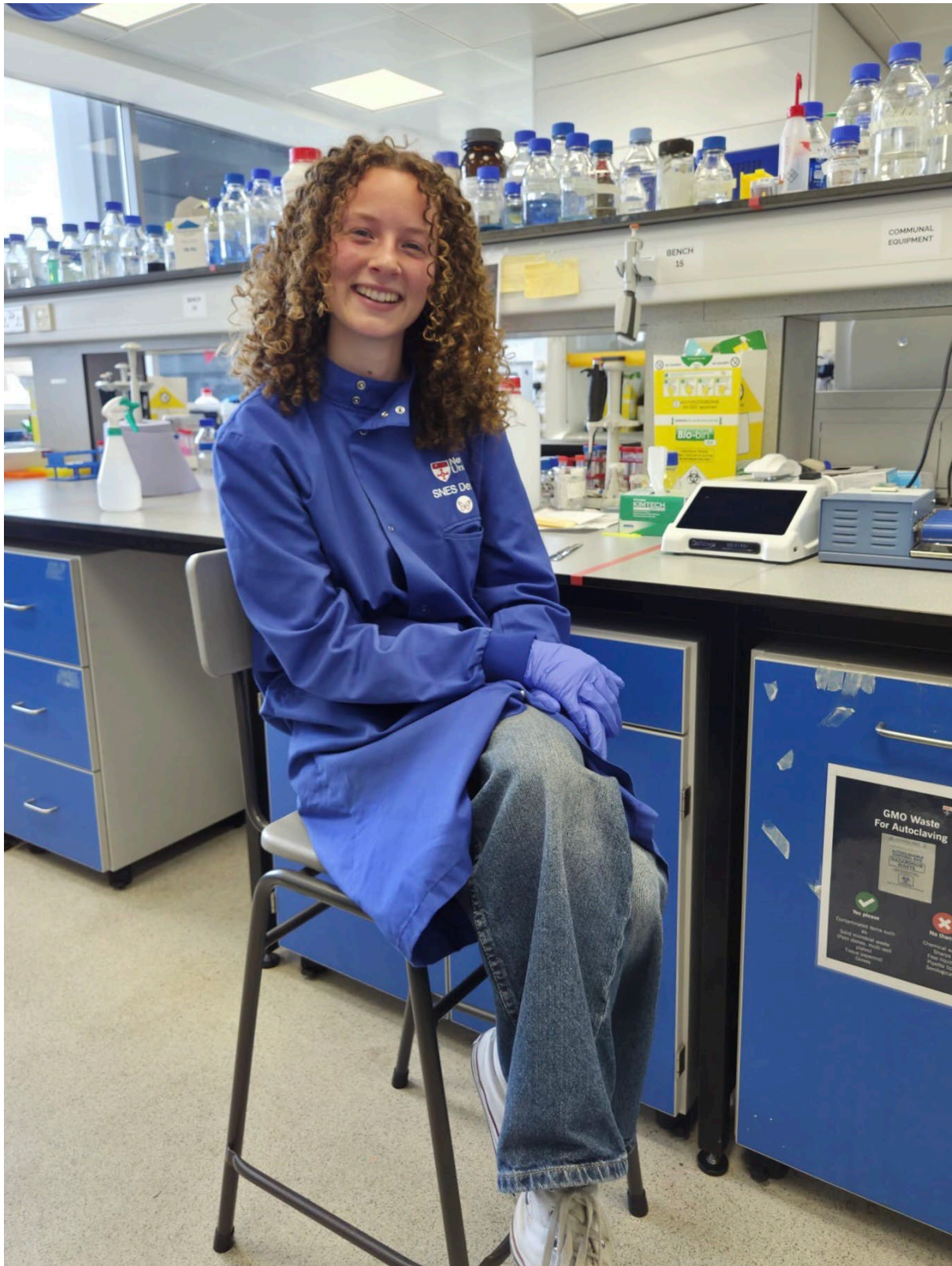
Student: Mia Johnson

Supervisor: Matt Bawn

Project Location: School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne

Lay Summary:

Rivers support complex microbial ecosystems which drive nutrient cycling and form the foundations of food chains. Microbes are essential for maintaining the health of river ecosystems, however the diversity of them below the surface is often overlooked. Current methods of determining water quality remain predominantly measured by instrumental analysis, which requires specialist equipment and extensive training. This report aims to exemplify that molecular methods such as 16S DNA sequencing and metagenomics, derived from Oxford Nanopore Sequencing, can be utilised alongside widely used techniques to test water quality. I aim to show that Nanopore produces an accurate microbial profile, in real time, and can be completed with minimal training. The area of focus are tributary microbial communities in the Tyne catchment; there is limited characterisation of these ecosystems, this project hopes to address that gap. Metagenomics will be performed to create a picture of the diversity, and whole genome sequencing will be used to identify specific species, utilising next generation sequencing tools; both techniques will be reinforced with qualitative data. The results will contribute to a more complete picture of river health and demonstrate how molecular tools can be used alongside established methods to monitor water quality.



Aim

I aimed to investigate the microbial diversity of public water ecosystems in Newcastle Upon Tyne, using DNA sequencing in conjunction with culture-based and microscopy techniques,

to exemplify that next-generation sequencing tools are a reliable way to determine water quality, specifically to identify harmful pathogens present in recreational waters.

Methodology

Figure 1: Satellite map of Newcastle Upon Tyne indicating the 5 sample locations and images of where the samples were collected from.

I collected samples from 5 water sources of interest within the Tyne catchment: Leazes Park, Lort Dene, Jesmond Dene, River Tyne and the Victorian Tunnels (Figure 1). I then vacuum filtered the samples with the intention of trapping as many microbes as possible; the filters were placed onto nutrient agar plates and were subsequently left to incubate for 48 hours. The filters were then scraped and streaked onto 5 different types of selective agar to look for pathogens such as *E.coli*, *Salmonella* and *Pseudomonas*.

My second experiment provided quantitative data to accompany the first. 16S DNA extractions and quality control checks were performed. The DNA was amplified using PCR and sequenced with the 16S Barcoding Kit 24 V14. Metagenomic analysis was undertaken to build profiles of microbial communities. Whole genome sequences were also completed using Oxford Nanopore Technologies to identify pathogens at species-level resolution and to look for antimicrobial resistance genes.

Electron microscopy was undertaken to visualise bacterial isolates to accompany the whole genome sequences and support the identification of microbes.

Results

Figure 2: Results of selective agar plating for each sample location. EMB selects for gram negative bacteria that ferment lactose, such as E. coli, which produce a green sheen. SS selects for Salmonella and Shigella, CUA selects for citrate positive Enterobacteriaceae resulting in a colour change to pink due to a pH change. Pseudo agar selects for Pseudomonas.

Selective agar plating in figure 2, revealed faecal contamination across most sites. *E. coli* was detected at all locations except the Victorian Tunnel Wall, with the highest levels in the Lort Dene and River Tyne. Growth on Salmonella–Shigella agar suggested the presence of *Salmonella* in Leazes Park and the Lort Dene, while pseudomonas-selective agar showed high levels of *Pseudomonas* in the River Tyne.

Metagenomics, shown in figure 3, revealed that the sequencing data largely mirrored the selective agar results. High levels of Enterobacter in Leazes Park, Jesmond Dene, and the

Figure 3: Relative abundance of top 25 genera, across 5 sample locations. Classified by minKNOW. Unknown = minKNOW software could not classify sequence. Other = <25 genera that would have made the plot difficult to read. Results from Oxford Nanopore 16S Barcoding Kit 24 V14. Graph made on RStudio.

Lort Dene were consistent with culture data. However, *Escherichia* was underrepresented in the River Tyne, likely due the Nanopore producing significantly fewer reads for this location, meaning the data is not representative of the entire sample. Diversity analysis showed the Victorian Tunnel as the least diverse site, dominated by *Aneurinibacillus*, a genus linked to heavy metal contamination, suggesting industrial influence. The River Tyne was dominated by *Brevibacillus*, while Leazes Park, Jesmond Dene, and the Lort Dene displayed the highest diversity. All three contained *Escherichia* and *Shigella* among their most abundant genera, indicating significant pollution. I was able to comfortably use the Nanopore after 2 attempts, demonstrating its accessibility. If I was to repeat this experiment, I would aim to remove culture bias by extracting microbe DNA directly from the water samples.

Scanning electron microscopy (Figure 4) confirmed rod-shaped morphologies, indicating *Bacillus*, across 3 random samples, supporting that *Bacillus* is the most abundant genus in the Tyne catchment. Whole genome sequencing identified species such as *Bacillus pseudomycoides* and *Bacillus cereus*, demonstrating the resolution achieved when combining culture, sequencing, and imaging.

Figure 4: Scanning electron microscopy images taken at ~2KX, ~5KX and ~8KX magnification depicting the morphology and distribution of 3 random bacterial samples. The top and bottom rows are from the Tyne, the middle row is from the Victorian Tunnels. Plugs of agar were fixed, dehydrated gradually in ethanol and gold plated for visualising. Images were taken using a VEGA3 TESCAN SEM.

To conclude, whole genome sequencing to identify pathogenic strains of bacteria, alongside metagenomics for genus level identification of microbes, has proved an appropriate method to measure water quality, opening the field for the use of next-generation sequencing tools.

Reflection

Over the course of my six-week project, I have had numerous opportunities to develop new skills both in and out of the lab. A highlight was being able to take scanning electron microscopy images of the bacteria I had worked on. Learning how the microscope worked and then taking my own images was incredibly rewarding. Another major highlight was working in a fabulous team of researchers, where experiences, techniques and advice was readily shared. I had many offers to assist with their PhD projects, meaning I gained experience in protein purification using tangential flow filtration, and used a rotary evaporator to obtain crude extracts. Outside of the lab, I enjoyed being involved in scientific communications through my six-week Instagram series for the School of Natural and Environmental Sciences at Newcastle. I also completed a series of project updates on LinkedIn, delivered a presentation on my findings in a departmental lab meeting, and presented a poster at the GenerationResearch conference at the University of York.

Immersing myself in my own research project has exemplified my love for investigation; having the opportunity to problem solve and contribute to science has been incredible. Networking with experienced and up-and-coming scientists has resulted in valuable relationships and taught me the importance of collaboration to solve global issues. I hope to continue this journey of discovery through a master's by Research or a PhD. Working with microbes has given me an appreciation for the hidden world, which microscopes can open our eyes to; I would like to continue utilising this technology throughout my career.