# Chapter 5: Gastrointestinal Genomics and Clinical Microbiology 2017 - Shiga Toxin-producing Escherichia coli (STEC)

## Background

Shiga toxin-producing Escherichia coli (STEC), also known as Verocytotoxin-producing E. coli (VTEC), is defined by the presence of the phage encoded Shiga toxin (Stx) genes (stx1, stx2 or both). There are more than 400 different serotypes of STEC and over 100 of these are known to cause symptoms of gastrointestinal (GI) disease in humans, including severe bloody diarrhoea and haemolytic uraemic syndrome (HUS). Previous studies have indicated that the presence of stx2, specifically the stx2a subtype, is more frequently associated with severe disease. Many STEC associated with human disease also have the intimin-encoding gene eae (E. coli attaching and effacing), located on a pathogenicity island called the locus of enterocyte effacement (LEE), and associated with the intimate attachment of the bacteria to the human gut mucosa. Recently, strains of STEC that do not have the eae gene but carry a plasmid encoding aggR, associated with the enteroaggregative *E. coli* group, have also been associated with causing HUS. Shiga toxin-producing *Escherichia coli* (STEC) are considered to be a significant threat to public health due to the severity of gastrointestinal symptoms associated with human infection and the risk of cases developing Haemolytic Uraemic Syndrome (HUS). STEC are zoonotic; transmission occurs by direct contact with animals or their environment, or by consumption of contaminated food or water. The infectious dose is low (<10 organisms) and person-to-person spread is common.

In England, the current Standards for Microbiology Investigations protocols are specific for the isolation of non-sorbitol fermenting colonies of *E. coli* serogroupO157 on cefixime tellurite sorbitol MacConkey (CT-SMAC) agar. STEC serogroups other than O157 (non-O157 STEC) are not detected using this method. However, since 2012 the implementation of commercial PCR assays for the detection of STEC in faecal specimens from cases with symptoms of gastrointestinal infection, at a twelve local hospital laboratories, has resulted in an increase in the detection of non-O157 STEC in the UK. Faecal specimens that are PCR positive for the Shiga Toxin (stx) genes at the local hospital laboratories in England are sent to the Gastrointestinal Bacterial Reference Unit (GBRU) at Public Health England (PHE) for isolation of STEC and subsequent serotyping. Recent advances in whole genome sequencing (WGS) have led to the development of a method for high throughput sequencing of bacterial genomes at low cost. WGS was implemented at GBRU for real-time surveillance of STEC in June 2015.

## Scenario - story

This scenario is designed to give you an insight into how whole genome sequencing may be useful from different clinical and public health perspectives, as well as the practical details of DNA extraction, whole genome sequencing (WGS) library preparation and bioinformatics analysis.

### Chapter 1

You are a clinical microbiologist in a large UK hospital. It is Monday 23rd January and since Friday 20th January, five patients have been admitted to ITU with symptoms of Haemolytic Uraemic Syndrome (HUS).

Faecal specimens submitted to the local laboratories were culture negative for the common bacterial gastrointestinal (GI) pathogens (Salmonella, Campylobacter, Shigella and STEC O157) and were couriered to GBRU for further testing. This morning the reference laboratory reported detection of non-O157 STEC from three cases identified Friday 13th January. You need to perform whole genome sequencing on these isolates and analyse the results in order to determine the serotype and virulence profile of the strain associated with the outbreak.

This will involve:

* WGS library prep
* Using the Galaxy bioinformatics platform
* KmerID to speciate from raw sequencing data
* Identification of serotype and virulence profile from the genome
* Performing whole genome SNP analysis & interpretation of the phylogeny of the outbreak strain

#### Chapter 2

You are an epidemiologist at Public Health England. It is Wednesday 25th January and what started as a cluster of cases linked to a hospital in one region has become a large national outbreak. You will analyse phylogenetic trees to determine the relationship between cases in the national outbreak and generate epidemiological hypotheses from based on the data. This will involve:

* Interpretation of national phylogenetic trees

#### Chapter 3

You are the clinical microbiologist from chapter 1 again. Public Health colleagues in France and Germany have both reported an increase in the number of HUS cases over the week-end. The French have reported PCR data suggesting that the Shiga toxin subtype is a highly pathogenic type and the Germans have reported that the isolate is multidrug resistant. You know that both these properties are difficult to analyse using short read Illumina data so you decide to try out the new-fangled MinION machine to sequence the outbreak strain so determine whether the data enables you to improve your analysis of the accessory genome. This will involve the following:

* MinION™ nanopore sequencing Library Preparation
* MinION data analysis

### Scenario – practical

#### Isolating Genomic DNA from Gram negative bacteria

Story:

This morning the reference laboratory reported detection of non-O157 STEC from three cases identified Friday 13th January. You need to perform whole genome sequencing on these isolates and analyse the results in order to determine the serotype and virulence profile of the strain associated with the outbreak.

Principle of the procedure

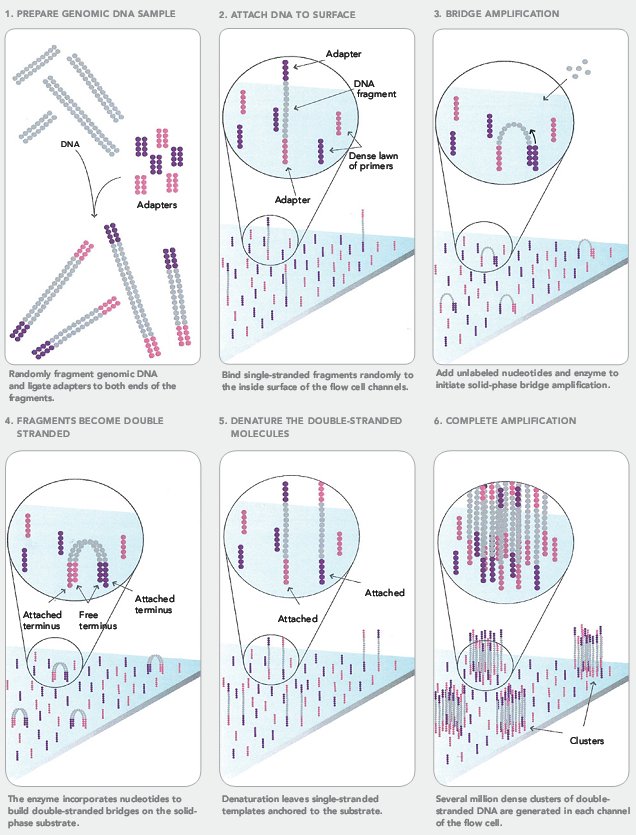
The Wizard Genomic DNA Purification Kit is designed for isolation of DNA is based on a four step process. The first step in the purification procedure lyses the cells and the nuclei. The cellular proteins are then removed by a salt precipitation step, which precipitates the proteins but leaves the high molecular weight genomic DNA in solution. Finally, the genomic DNA is concentrated and desalted by isopropanol precipitation.

1. Add 1ml of an overnight culture to a 1.5ml microcentrifuge tube.
2. Centrifuge at 13, 000 rpm for 2 min to pellet the cells and remove the supernatant.
3. Add 600ul Nuclei Lysis Solution and gently pipet until the cells are resuspended.
4. Incubate at 80oC for 5 mins to lyse the cells; then cool to room temperature.
5. Add 3ul of RNase solution to the cell lysate. Invert the tube 2-5 times to mix.
6. Incubate at 3 oC for 15-60 mins. Cool the sample to room temperature.
7. Add 20 ul of protein precipitation solution to the lysate. Vortex vigorously at high speed for 20 seconds to mix the protein precipitation solution with the cell lysate.
8. Incubate the sample on ice for 5 mins.
9. Centrifuge at 13, 000 rpm for 3 mins.
10. Transfer the supernatant containing the DNA to a clean microcentrifuge tube containing 600ul of isopropanol. Some supernatant may remain in the original tube containing the protein pellet. Leave this residual liquid in the tube to avoid contaminating the DNA solution with the precipitating protein.
11. Gently mix by inversion until the thread like strands of DNA form a visible mass.
12. Centrifuge at 13,000 rpm for 2 mins.
13. Carefully pour off the supernatant and drain the tube on clean absorbent paper. Add 600 ul of room temperature 70% ethanol and gently invert the tube several times to wash the DNA pellet.
14. Centrifuge at 13, 000 rpm for 2 mins and carefully aspirate the ethanol.
15. Drain the tube on clean absorbent paper and allow the pellet to air-dry for 10-15 min.
16. Add 100ul of DNA Rehydration solution to the tube and rehydrate the DNA by incubating at 65oC for 1 h. Periodically mix the solution gently.
17. Store the DNA at 4oC.

#### Whole Genome Sequencing – Nextera XT protocol

**PRINCIPLE OF PROCEDURE**

Using a single transposase enzymatic reaction, sample DNA is simultaneously fragmented and tagged with sequencing adapters. Short sample specific oligonucleotide barcodes are attached to the fragmented DNA. Libraries containing different indexed adapters are then constructed, quantified, pooled in equimolar amounts, and sequenced.  Deconvoluting the bar codes informatically then allows multiple libraries to be sequenced on a single flow cell.



Overview of Illumina sequencing. Reproduced from SEQanswers.com

**Fragmentation and Tagmentation of Genomic DNA**

**Personnel Protective Equipment required**

Lab coat, nitrile gloves and safety specs

**Hazardous substances**

**80% Ethanol** – Flammable, Irritant

**Sodium Hydroxide** – Corrosive, Irritant, Skin Sensitizer

**TD buffer** (contains Formamide) – Toxic, Irritant, Teratogen – expectant or new mothers should avoid handling this chemical.

**LDR Formamide** (MiSeq reagent kit) – Toxic, Teratogen – expectant or new mothers should avoid handling this chemical.

**PR2 Incorporation Buffer** (MiSeq reagent kit) – Irritant, Skin Sensitizer

**Library Normalisation Wash 1** (Nextera XT sample prep. kit) – Mutagen, Teratogen – expectant or new mothers should avoid handling this chemical.

**Library Normalisation Additives 1** (Nextera XT sample prep. kit) – Mutagen, Teratogen – expectant or new mothers should avoid handling this chemical.

**Dye Concentrate** (Agilent high sensitivity DNA kit) (contains DMSO) – Irritant, Flammable

|  |  |  |
| --- | --- | --- |
| Item | Quantity | Storage |
| Amplicon Tagment  Mix (ATM) | 1 tube  (7μl) | Ice  Bucket |
| Tagment DNA Buffer  (TD) | 1 tube  (12μl) | Ice  Bucket |
| Neutralize Tagment  Buffer (NT) | 1 tube  (7μl) | Room  Temp |
| 1ng Input DNA | Provided at:  5μl @ 0.2ng/μl | -15°to  -25° |

1. Ensure all reagents are adequately mixed by gently inverting the tubes 3–5 times, followed by a brief spin in a microcentrifuge.
2. Retrieve the tube containing your genomic DNA (Labelled as your group number).
3. Add 10μl of TD Buffer to the DNA sample.
4. Add 5μl of ATM and gently pipette up and down 5 times to mix.
5. Place the sample in a thermocycler and run the following program (with heated lid):

|  |  |
| --- | --- |
| Thermal Cycler Setting | |
| Program Name: | Nextera XT Tagmentation |
| Total Sample Volume: | 20µl |
| Parameters: | 55°C for 5 minutes  Hold at 10°C |

1. Once the tubes have reached 10°C immediatelyremove from the thermal cycler.
2. Add 5µl of NT buffer and pipette mix gently 5 times to ensure that the sample is thoroughly mixed.
3. Pulse spin in a microcentrifuge.
4. Incubate the sample at room temperature for 5 minutes.

**PCR Amplification**

Retrieve the following reagents and consumables from the ice bucket:

|  |  |  |
| --- | --- | --- |
| Item | Quantity | Storage |
| Nextera PCR Master  Mix (NPM) | 1 tube  (17μl) | Ice  Bucket |
| Index 1(i7) & 2 (i5)  Primer Mix (index) | 1 tube  (10μl) | Ice  bucket |

1. To your DNA sample, add 15µl of NPM and pipette mix 5 times.
2. Add 10µl of the Index Primer Mix.
3. Replace the lid and pulse spin in a microcentrifuge.
4. Place the sample tube onto a thermal cycler using the following parameters:

|  |  |
| --- | --- |
| Thermal cycler settings | |
| Program Name: | Nextera XT PCR |
| Total Volume: | 50 µl |
| Parameters: | 72°C for 3 minutes  95°C for 30 seconds  12 cycles of:  95°C for 10 sec  55°C for 30 sec  72°C for 30 sec  72°C for 5 minutes  Hold at 10°C |

**Ampure Clean-up**

|  |  |  |
| --- | --- | --- |
| Item | Quantity | Storage |
| Resuspension Buffer | 1 tube (55μl) | Room Temperature |
| AMPure XP beads (XP) | 1 tube | Room Temperature |
| Fresh 80% ethanol | 1 x 15 ml tube | N/A |

1. Retrieve your sample from the thermal cycler. Briefly vortex and pulse spin in a microcentrifuge.
2. Label a new tube with your group number.
3. Transfer 50µl of the PCR product from the original sample tube to the new clean tube.
4. Vortex the AMPure XP beads for 30 seconds to ensure that the beads are evenly dispersed.
5. Add 30μl of AMPure XP beads to the tube. Pipette mix 10 times.
6. Incubate at room temperature for 5 minutes.
7. Place the tube on a magnetic stand for 2 minutes or until the supernatant has cleared.
8. Carefully remove and discard all the supernatant from each well.

(Please note: Your DNA is now bound to the Ampure XP beads, avoid disturbing the beads. If any beads are inadvertently aspirated into the tips, dispense the beads back into the tube and let it rest on the magnet for 2 minutes and confirm that the supernatant has cleared)

1. Add 200µl of 80% ethanol to the tube whilst still on the magnetic stand. Incubate for approx 30 secs and carefully remove and discard the supernatant. Do not remove the beads.
2. Repeat the ethanol wash in step 9.
3. If required use a P10 pipette to remove excess ethanol, so as to not remove the beads.
4. With the samples still on the magnetic stand, allow the beads to air-dry for 5 minutes.
5. Remove the tube from the magnetic stand and add 52.5μl of RSB.
6. Gently pipette mix up and down 10 times.
7. Incubate at room temperature for 2 minutes.
8. Place the tube back on the magnetic plate for 2 minutes (or until the supernatant has cleared).
9. Label a new tube ‘CAN’(Clean Amplified NTA) & your group number.
10. Carefully transfer 30μl of the supernatant to the CAN tube.

**Library Quality and Quantity Check (Bioanalyzer)**

* The size distribution of your library can be checked by running 1μl of it on an Agilent Technologies 2100 Bioanalyzer using a High Sensitivity DNA chip.
* Typical libraries show a broad size distribution from ~250-1000bp, with an average of ~400–500bp.

Example of DNA Library Size Distribution

**Setting up the Chip Priming Station**

1. Insert the syringe into the clip
2. Slide it into the hole of the luer lock adapter and screw it tightly to the chip priming station.
3. Check base plate is in position C.
4. Adjust the syringe clip to the lowest position.

**Checking the Chip Priming Station for Good Seal — Seal Test**

1. Make sure the syringe is tightly connected to the Chip Priming Station.
2. Pull the plunger of the syringe to the 1.0ml position (plunger pulled back).
3. Place an empty chip in the Chip Priming Station.
4. Close the Chip Priming Station and make sure to lock it by pressing the cover, the lock of the latch will audibly click when it closes.
5. Press the plunger down until it is locked by the clip.
6. Wait for 5 seconds and press the side of the clip to release the plunger.
7. Appropriate sealing is verified if the plunger moves back up to the 0. ml mark within less than 1 second.

**Preparing the Gel-Dye Mix**

1. Allow High Sensitivity DNA dye concentrate (blue) and High Sensitivity DNA gel matrix (red) to equilibrate to room temperature for 30 min.
2. Add 15μl of High Sensitivity DNA dye concentrate (blue) to a High Sensitivity DNA gel matrix vial (red).
3. Vortex solution well and spin down. Transfer to spin filter.
4. Centrifuge at 2240 g ± 20 % for 10 min. Protect solution from light. Store at 4 °C.

**Loading the Gel-Dye Mix**

1. Allow the gel-dye mix equilibrate to room temperature for 30 min before use.
2. Put a new High Sensitivity DNA chip on the chip priming station.
3. Pipette 9.0μl of gel-dye mix in the well, marked G.
4. Make sure that the plunger is positioned at 1ml and then close the chip priming station.
5. Press plunger until it is held by the clip.
6. Wait for exactly 60seconds then release clip.
7. Wait for 5seconds, and then slowly pull back the plunger to the 1 ml position.
8. Open the chip priming station and pipette 9.0μl of gel-dye mix in the two additional wells marked G.

**Loading the Marker**

1. Pipette 5μl of marker (green) in all sample and ladder wells. Do not leave any wells empty.

**Loading the Ladder and the Samples**

1. Pipette 1μl of High Sensitivity DNA ladder (yellow) in the ladder well.
2. In each of the 11 sample wells pipette 1μl of sample (used wells) or 1μl of marker (unused wells).
3. Put the chip horizontally in the adapter and vortex for 1 min at the indicated setting (2400 rpm).
4. Run the chip in the Agilent 2100 Bioanalyzer within 5 min.

**Starting the Run**

1. Select the High Sensitivity assay from the Assay menu.
2. Enter details in the sample name table.
3. Click the Start button in the upper right of the window to start the chip run.
4. The incoming raw signals are displayed in the Instrument context.
5. After the run is finished, remove the chip.

**Average Library Fragment Size**

1. When viewing the results of the run, navigate to the ‘Region Table Bar’
2. Move the blue bars to either side of the curve.
3. The average length in bp will be displayed
4. Use this value to calculate the molarity using the values from the Qubit quantitation assay.

**Library Normalisation**

**NB: REAGENTS CONTAIN FORMAMIDE: TO BE CONDUCTED INSIDE FUME HOOD**

|  |  |  |
| --- | --- | --- |
| Item | Quantity | Storage |
| Library Normalisation Additives1 (LNA1) | 1 tube | Ice bucket |
| Library Normalisation  Beads 1 (LNB1) | 1 tube | Ice bucket |
| Library Normalisation  Wash 1 (LNW1) | 1 tube | Ice bucket |
| Library Normalisation  Storage Buffer 1 (LNS1) | 1 tube | Room Temperature |
| Fresh 0.1 N NaOH | 1 tube | Room Temperature |
| 96-well plate | 1 plate | N/A |
| 15 ml conical tube | 1 tube | N/A |

1. Remove the LNA1, LNB1 and LNW1 from the ice bucket (and LNS1 from its storage location) and bring to room temperature. Vortex for approximately 1 minute prior to use, ensuring any precipitate is fully resuspended.
2. Retrieve the ‘CAN’ tubes and, pulse-spin in a microcentrifuge. Pipette mix the sample thoroughly.
3. Label a 96-well plate ‘Normalisation’
4. Transfer 20µl of supernatant from each ‘CAN’ tube to a separate well in the ‘Normalisation’ plate.
5. Volumes of LNA1 and LNAB required:

|  |  |  |
| --- | --- | --- |
| Sample Number | Reagents | Volume |
| 12 | LNA1 | 550µl |
| LNB1 | 100µl |

1. Ensure the LNA1 and LNB1 is thoroughly mixed just prior to use.
2. Combine the required volumes of LNA1 and LNB1 in a 15ml conical tube. Vortex thoroughly for 30 seconds and invert 15-20 times.
3. Add 45μl of LNA1/LNB1 mix to each sample in the ‘Normalisation’ plate.
4. Seal the plate and shake at 1800rpm for 30 minutes.
5. Place the plate on the magnetic stand and remove the plate seal. Incubate on the magnet for 2 mins and carefully remove all of the supernatant (with multichannel).

(Please note: Avoid disturbing the beads during this step).

1. Remove the plate from the magnetic stand and add 45μl of LNW1 to each sample well.
2. Seal the plate and place onto a plate shaker at 1800rpm for 5mins.
3. Place the plate on the magnetic stand and remove the plate seal. Incubate on the magnet for 2mins and carefully remove the supernatant (with multichannel).
4. Repeat wash steps 11 – 13.

(Please note: Ensure excess LNW1 is removed, if required use a P10 pipette or multi-channel to remove residues).

1. Remove the plate from the magnetic stand and add 30μl of 0.1 N NaOH to each sample well to elute sample.

(Please note: Only use freshly prepare 0.1N NaOH and do not store).

1. Seal the plate with and shake at 1800rpm for 5 minutes.
2. Label a new 96-well plate ‘FINAL’.
3. Pipette 30μl of LNS1 to the appropriate wells on the ‘FINAL’ plate.
4. Remove the ‘Normalisation’ plate from the shaker and check to ensure all samples are completely resuspended, if not, pipette mix and place back onto the shaker for a further 5minutes @1800rpm.
5. Place the ‘Normalisation’ plate on a magnetic stand and remove the plate seal. Incubate for 2 minutes.
6. Transfer 30μl the supernatant from the ‘Normalisation’ plate to the ‘FINAL’ plate.
7. Seal the ‘FINAL’ plate and centrifuge at 1000xg for 1 minute.

**Loading the MiSeq**

**Preparing the 20pM PhiX (control) Library**

* Gently vortex and pulse the 10nM PhiX Library.
* Combine 2μL 10nM PhiX Library and 3μL EBT buffer to make a 4nM PhiX library.
* Add 5μL 0.2M NaOH to the above Eppendorf to make 2nM PhiX library.
* Vortex the Eppendorf and pulse spin.
* Incubate for 5 minutes at room temp to denature the PhiX library.
* Add 990μL HT1 to the Eppendorf to make 20pM denatured PhiX library.

Store the 20pM denatured PhiX library between -15°C to -25°C. Dispose after 3weeks.

**Preparing the ‘POOLED AMPLICON LIBRARY’ (PAL) tube**

* Retrieve the ‘FINAL’ plate and place plate shaker and mix for 1 minute at

1500rpm and pulse-spin in a plate centrifuge

* Place onto a magnetic stand (to collect any remaining beads)
* Pool 5μL of sample from the ‘FINAL’ plate into an eppendorf labelled ‘Pool’.

**Preparing the ‘DILUTED AMPLICON LIBRARY’ (DAL) tube**

* Set a heat block to 96°C and prepare an ice water bath
* Gently vortex and pulse spin the PAL tube.
* Vortex HT1 to remove all trace of precipitate.
* Label a clean 1.5mL LoBind Eppendorf with ‘DAL’.
* Add 22μL of ‘PAL’ and 588μL HT1 to the ‘DAL’ tube.
* Gently vortex and pulse-spin.
* Incubate the ‘DAL’ tube at 96°C for 2 minutes.
* Invert the ‘DAL’ tube twice and immediately place into the ice water bath.
* Incubate the ‘DAL’ tube for 5 minutes.
* Add 18μL denatured 20pM PhiX library to the ‘DAL’ tube.
* Gently vortex and pulse spin – keep on ice
* The ‘DAL’ tube is now ready for loading onto the MiSeq

**Preparing MiSeq Reagent Cartridge**

**Allow approximately one hour for thawing cartridge**

1. Place the reagent cartridge in a water bath containing enough room temperature deionized water to submerge the base of the reagent cartridge up to the water line printed on the reagent cartridge. Do not allow the water to exceed the maximum water line.
2. Allow the reagent cartridge to thaw in the room temperature water bath for approximately one hour or until completely thawed.
3. Remove the cartridge from the water bath and gently tap it on the bench to dislodge water from the base of the cartridge. Dry the base of the cartridge. Make sure that no water has splashed on the top of the reagent cartridge.
4. Invert the reagent cartridge to mix the thawed reagents, and then visually inspect that all positions are thawed.
5. Visually inspect the reagent marked IMF (Position 1) to make sure that it is fully mixed and free of precipitates.
6. NOTE: The MiSeq sipper tubes go to the bottom of each reservoir to aspirate the reagents, so it is important that the reservoirs are free of air bubbles.
7. Place the reagent cartridge on ice or set aside at 2° to 8°C until you are ready to set up your run.

**MiSeq Instrument Prep:**

**Clean the Flow Cell**

1. Wash the flow cell with Millipore water
2. Dry any excess water with a lint-free lens cleaning tissue, and visually inspect to make sure that the flow cell ports are free of obstructions and that the gasket is well seated around the flow cell ports.
3. If the gasket appear to be dislodged, gently press it back into place until it sits securely around the flow cell ports.

**Loading the Flow Cell**

1. Raise the flow cell compartment door, and then press the release button to the right of the flow cell latch. The flow cell latch opens.
2. Visually inspect the flow cell stage to make sure it is free of lint. If lint or other debris is present, clean the flow cell stage using an alcohol wipe or a lint-free tissue moistened with ethanol or isopropanol. Carefully wipe the surface of the flow cell stage until it is clean and dry.
3. Hold the flow cell by the edges of the flow cell cartridge near the Illumina label.
4. Make sure the label is facing upward and place the flow cell on the flow cell stage.
5. Gently press down on the flow cell latch to close it over the flow cell. You will hear a click when the flow cell latch is secure.
6. As you close the flow cell latch, two alignment pins near the hinge of the flow cell latch properly align and position the flow cell.
7. Check the lower-left corner of the screen to confirm that the flow cell RFID was successfully read.
8. Close the flow cell compartment door.
9. Select Next on the Load Flow Cell screen. The Load Reagents screen opens

**Loading Reagents**

1. Remove the bottle of PR2 from fridge. Gently invert the bottle to mix the PR2 bottle, and then remove the lid.
2. Open the reagent compartment door.
3. Raise the sipper handle until it locks into place.
4. Place the PR2 bottle in the indentation to the right of the reagent chiller.
5. Make sure that the waste bottle is empty. If required, empty the contents into the appropriate waste container.
6. Slowly lower the sipper handle. Make sure that the sippers lower into the PR2 and waste bottles.
7. Check the lower-left corner of the screen to confirm that the RFID of the PR2 bottle was read successfully.
8. Select Next on the Load Reagents screen.

**Load Sample Libraries onto Cartridge**

1. Use a clean 1 ml pipette tip to pierce the foil seal over the reservoir labelled Load Samples.

NOTE: Do not pierce any other reagent positions. Other reagent positions are pierced automatically during the run.

1. Pipette 600μl of your sample libraries into the Load Samples reservoir. Take care to avoid touching the foil seal as you dispense your sample.



1. Proceed directly to the run setup steps using the MiSeq Control Software (MCS) interface.

**Load the Reagent Cartridge**

NOTE Do not leave the reagent chiller door open for extended periods of time.

1. Open the reagent chiller door.
2. Hold the reagent cartridge on the end with the Illumina label, and slide the reagent cartridge into the reagent chiller until the cartridge stops.
3. Close the reagent chiller door.
4. Check the lower-left corner of the screen to confirm that the RFID of the reagent cartridge was read successfully.
5. Close the reagent compartment door.

Select Next on the Load Reagents screen. The Review screen opens.

**Sample Sheet Set-up (Illumina Experiment Manager)**

The Illumina Experiment Manager is a wizard-based application that guides you through the steps to create your sample sheet.

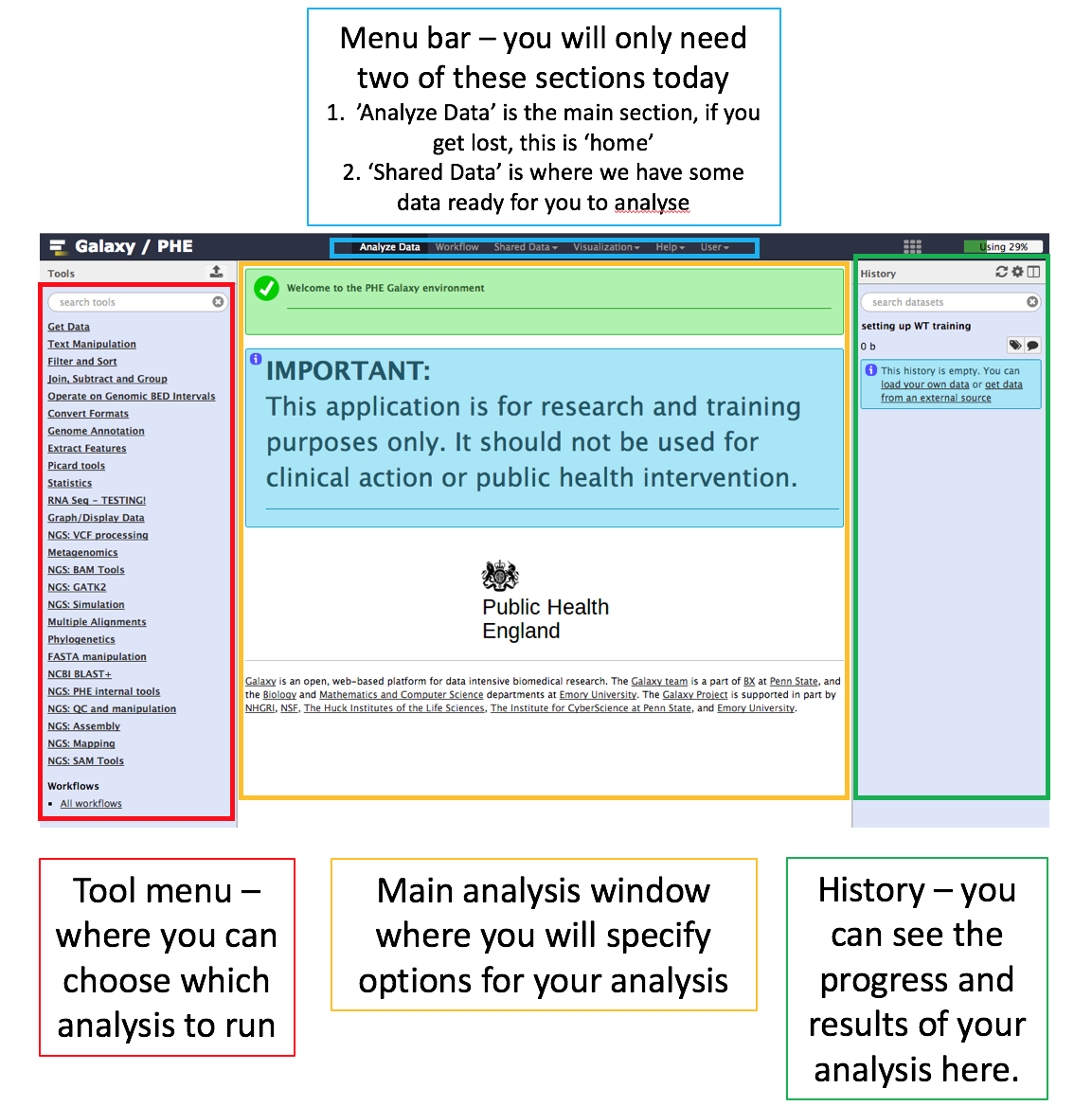
**Starting the Run**

* After you have loaded the flow cell and reagents, the MCS interface prompts you to review run parameters and perform a pre-run check before beginning the run.

#### Getting started with Galaxy bioinformatics platform

For the bioinformatics section of the analysis, we will be using the Galaxy bioinformatics website. Galaxy is an application that allows bioinformaticians to host their own applications for users to interact with via a web browser rather than the command line. There is a public instance available at <https://usegalaxy.org/>, today we are going to be using the PHE instance.

* 1. Open the Chrome internet browser via the Start menu or desktop shortcut.
  2. In Chrome navigate to <http://bioinformatics-galaxy.phe.org.uk/root>
  3. You should be able to see something like the below, without the coloured boxes.



Story:

It is Tuesday 24th January and the Illumina sequencing data from the run you prepared yesterday will be available tomorrow morning. Meanwhile, the Germans have asked for help analysing the WGS data they have from their outbreak strain. First of all, you will confirm that you have sequenced an isolate of *E. coli*, and then you want to identify the serotype and virulence profile*.*

#### KmerID to speciate from raw sequencing data

Principle of the procedure

For this we will use a PHE method called KmerID, the code for this process is available from <https://github.com/phe-bioinformatics> but we will use Galaxy to run this software, rather than the command line. KmerID determines a similarity index between the FASTQ reads and each of the 1769 published reference genomes by calculating the percentage of 18-mers in the reference that are also present in the FASTQs. Only 18-mers that occur at least twice in the FASTQ are considered present. Mixed cultures are detected by comparing the list of similarities between the sample and the references with the similarities of the references to each other, and filtering this comparison for inconsistencies. Other publicly available tools that perform a similar function are Kraken (<https://ccb.jhu.edu/software/kraken/>) and One Codex (<https://www.onecodex.com/>), which I would recommend looking at because you don’t need to install anything and it has a very nice user interface.

Questions to answer during this section

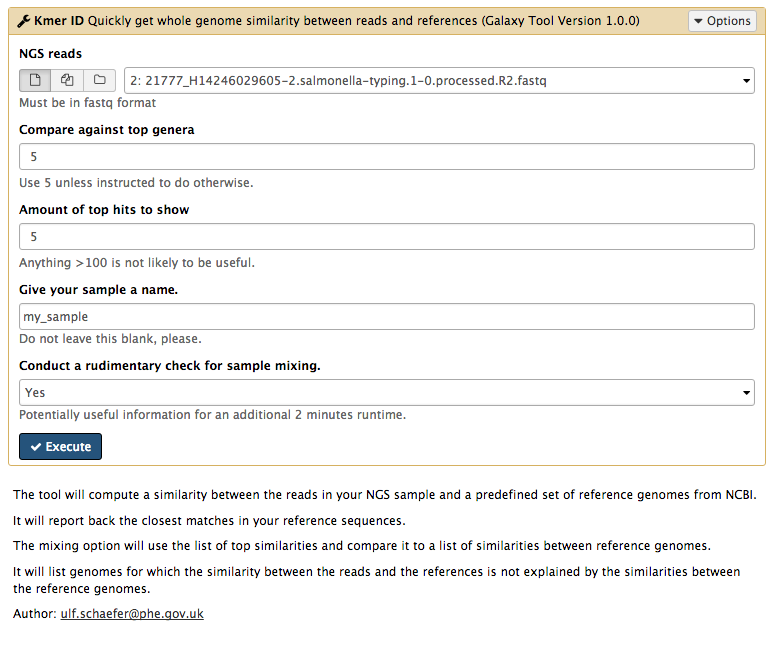
1. What is the top hit in the Kmer ID analysis
2. Is the sample mixed?

Practical steps

1. Go to ‘Shared Data’ on the menu bar.
2. Select ‘Data Libraries’ from the drop-down menu
3. On the new page that loads, select ‘WTAC’ for Wellcome Trust Advanced Course and then ‘2017’
4. Click on the small blue arrow next to ‘German Data’
5. Check the tick boxes next to **both** fastqs (these are sequence data files and end with .fastq)
6. Then, check that the current action next to ‘For selected datasets’ is ‘import to current history’ and click ‘Go’.
7. You should get a green notification bar saying that 2 datasets have been imported into your current history.
8. On the menu bar, click ‘Analyze data’ to return to the Home screen. You should now see two items in your history bar on the right hand side of the screen, these are your two fastqs.
9. On the left hand side of the Home screen is the tool menu, near the bottom there is an option ‘NGS: PHE internal tools’, clicking on this option will reveal a lot of options, including ‘Kmer ID’. Click on the underlined Kmer ID section (as below).

../../../Desktop/Screen%20Shot%202016-01-06%20at%2010.54.52.pn

1. The main analysis window should now look like the below. By default, one of the fastqs will be selected for analysis. It does not matter that only one fastq is being analysed, or which half of the pair you analyse. You can leave all the options as default and ‘Execute’. This will add two items to your history bar which will initially be grey and then turn yellow.



1. After about 10 minutes the yellow boxes will turn green and the analysis will be complete. Click on the eye symbol ../../../Desktop/Screen%20Shot%202016-01-06%20at%2011.11.07.pn next to the item in your history that says ‘Kmer ID on data 2: Mixing check’. This should change the contents of the main analysis window. What is the top hit?

#### Characterising the Genome

Principle of the procedure

The presence or absence of specific loci or specific allelic variants are important to characterise the strain and its pathogenic potential.

MLST

Achtman et al. proposed a sequenced based approach, multilocus sequence typing (MLST), based on the sequences of multiple house-keeping genes. Isolates that possess identical alleles for the seven gene fragments analysed are assigned a common sequence type (ST) and related STs from clonal complexes (CCs).

SEROTYPE

Pathogenic *E. coli* strains can be categorized based on elements that can elicit an immune response in animals, namely: the O antigen (part of the lipopolysaccharide layer) and the H antigen (the flagellin). The O antigen is used for serotyping *E. coli* and these O group designations go from O1 to O181 and are encoded by the rfb gene cluster. The H antigen is a major component of flagella, involved in *E. coli* movement. It is generally encoded by the *fliC* gene. There are 53 identified H antigens, numbered from H1 to H56.

VIRULENCE FACTORS

Several virulence factors have been identified in *E. coli* and are used to characterise strains into pathotypes. Virulence factors include toxins (e.g. shiga toxin and enterotoxins) and adherence mechanisms (e.g. intimin and fimbriae).

ANTIMICROBIAL RESISTANCE

Antimicrobial resistance is a growing problem in microbiology and especially significant in gram negative bacteria such as *E. coli*. Antimicrobial sensitivity of isolates to a battery of therapeutics is often used to guide treatment of infection.

Questions to answer during this section

1. What MLST sequence type is the isolate?
2. What is the serotype?
3. What is the pathotype?
4. What is the Shiga toxin subtype?
5. What is the genotypic antimicrobial profile of the isolate?
6. Do these characteristics tell you anything about the origin of this strain?

Practical steps

MLST

1. As we already have the fastqs, we can go straight to running the analysis. On the tool menu, in NGS: PHE internal tools, select MLST.

../../../Desktop/Screen%20Shot%202016-01-06%20at%2012.11.00.pn

1. In the main analysis window, put the fastq that ends R1.fastq into the ‘Read dataset for direction 1’ box, and the fastq that ends R2.fastq into the ‘Read dataset for direction 2’. Under ‘Select the organism’ choose *Escherichia coli*. Click Execute.
2. After 10-15 minutes, the three boxes in your history should have turned Green. Click the eye next to the box with the title that ends ‘MLST result XML’. The ST result is the fourth line in the main analysis screen <result type="MLST" value="X"> - what is the MLST value?

SEROTYPE & PATHOTYPE

1. On the tool menu, in NGS: PHE internal tools, select Gastro Serotyping. This will scan our FASTQ reads against a database of O & H determinant genes and some virulence genes.

Screen%20Shot%202017-01-12%20at%2014.26.03.png

1. In the main analysis window, put the fastq that ends R1.fastq into the ‘Read dataset for direction 1’ box, and the fastq that ends R2.fastq into the ‘Read dataset for direction 2’. Click Execute.
2. After 10-15 minutes, the box in your history should have turned Green. Click the eye next to the box with the title that ends ‘gastro serotyping XML’. What is are the results in the “o” and “h” fields? Are there any other positive matches?

SHIGA TOXIN SUBTYPE

1. On the tool menu, in NGS: PHE internal tools, select Stx subtyping.

Screen%20Shot%202017-01-12%20at%2014.35.21.png

1. In the main analysis window, put the fastq that ends R1.fastq into the ‘Read dataset for direction 1’ box, and the fastq that ends R2.fastq into the ‘Read dataset for direction 2’. Click Execute.
2. After 10-15 minutes, the box in your history should have turned Green. Click the eye next to the box with the title that ends ‘stx subtyping XML’. How many *stx* genes does this strain have and what are the subtypes?
3. What pathotype of *E. coli* is this sample?

ANTIMICROBIAL GENOTYPE

1. On the tool menu, in NGS: PHE internal tools, select Gastro resistance finder.

Screen%20Shot%202017-01-12%20at%2014.49.14.png

1. In the main analysis window, put the fastq that ends R1.fastq into the ‘Read dataset for direction 1’ box, and the fastq that ends R2.fastq into the ‘Read dataset for direction 2’. Click Execute.
2. After 10-15 minutes, the box in your history should have turned Green. Click the eye next to the box with the title that ends ‘resistance finder XML’. What classes of antibiotic have resistance mechanisms identified? What antibiotic would you recommend for treatment?

#### Performing whole genome SNP analysis & interpretation of the phylogeny

Story:

Scientists at Public Health England (PHE) curate a database containing over 3,000 isolates of STEC from clinical cases, animals and food. The majority of the cases acquired their infection in the UK but between 20-30% of cases travelled abroad in the week prior to onset of infection. Travel history and other epidemiological data are stored in the STEC Enhanced Surveillance System (SESSy). In the next part of the practical, you will compare the sequence from the German case with those sequences in the PHE database.

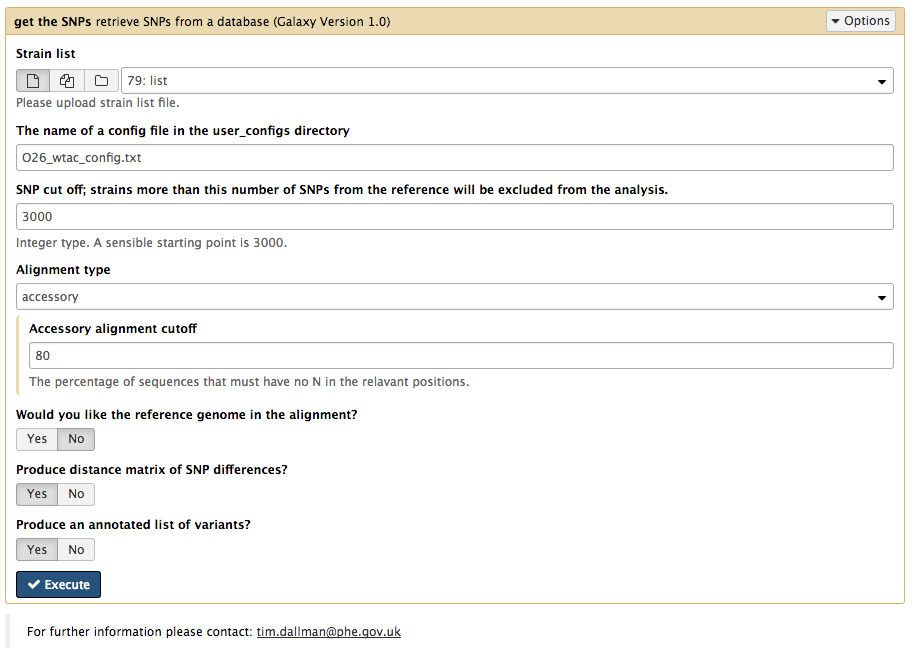
Principle of the procedure

Whole genome Single Nucleotide Polymorphism analysis involves ‘mapping’ the raw fastqs of a sample against an appropriate reference genome (i.e. same or closely related Sequence Type). The result of the mapping is then analysed to find differences (specifically, single nucleotide mutations), between the reference genome and the strain being analysed (the fastq). All the differences between a group of strains and the reference genome are collated and used to create a ‘pseudo-sequence’ of variant positions. This sequence can then be processed by phylogenetic algorithms to produce a phylogenetic tree. This process involves three main bioinformatics steps (mapping, SNP calling, collating variant positions) which are going to be carried out for you by SnapperDB. You will use two functions of SnapperDB, ‘fastq\_to\_db’ to upload the variant information into the database, and ‘get\_the\_snps’ to collate the variant positions on a set of interest.

Questions to answer during this section

1. How many SNPs is the outbreak from the most closely related isolate in the PHE database
2. Using the epidemiological data linked to the other closely related isolates, can you infer anything about the source of the outbreak strain?

Practical steps

1. First, we need to upload our data to the SNP database. On the tool menu, go to NGS: PHE internal tools and select ‘fastq to db’ from the SnapperDB section. In the main analysis window, select the R1 reads in the forward menu and the R2 read in the reverse menu. In the config file box, type ‘O26\_wtac\_config.txt’ – this must be **EXACTLY** correct (don’t include the quotation marks), or SnapperDB will not work. Then execute.
2. To see the progress and the different sub programs called click on the ../../../Desktop/Screen%20Shot%202016-01-06%20at%2011.11.07.pn icon. After 15-30 mins the boxes in the history window will turn green and the job will be completed, we are now ready to ‘get\_the\_snps’.
3. On the menu bar go to ‘shared data’ -> Data libraries -> WTAC -> 2017->German Data and add the ‘strain\_list\_1’ file to your current history. If you need more details on how to do this, consult the first part of the KmerID practical steps as they are very similar.
4. Return to the ‘Analyze Data’ page. On the tool menu, go to NGS: PHE internal tools and then ‘get the snps’. In the main analysis window select ‘strain\_list\_1’ as the strain list, set ‘O26\_wtac\_config.txt’ as the config file, select alignment type as ‘accessory’. Select ‘No’ to the question ‘Would you like the reference genome in alignment’ and ‘Yes’ to produce a distance matrix of SNP distances and list of annotated variants. Click execute.
5. After 1-2 minutes, the items in your history should turn from yellow to green. Click on the eye to inspect the results. There is a ‘pseudo-sequence’ of all the variant positions, a distance matrix of the pairwise SNP differences and a list that tells you where those SNPs are in the genome.
6. How many SNPs are there across the alignment? What percentage are in genes? What percentage are synonymous, non-synonymous?
7. How many SNPs between the German isolate and the closest isolate?
8. Click on the history item title to expand the history item. You should now be able to see a floppy disk two thirds down the left hand side of the expanded history item. Click on this floppy disk to download the pseudosequence, save it somewhere easily retrievable e.g. your Desktop.
9. Now we need to analyse the pseudo-sequence to generate a phylogenetic tree. On your computer, open the free application MEGA , (available from <http://www.megasoftware.net/>) which is a Graphic User Interface (or GUI, pronounced gooey, you may hear bioinformaticsy types referring to software ‘having a gooey’) for phylogenetic analysis.
10. In MEGA, go to file -> open a file/session -> select your saved pseudo-sequence file -> select ‘Analyze’ in the pop up box -> ensure ‘nucleotide sequences’ selected in the next pop up box, press ok -> select No when MEGA asks you ‘Protein coding nucleotide sequence data?’. Then you are ready to analyse your data.
11. In the MEGA window select the ‘Phylogeny’ box -> Construct/Test Maximum Likelihood Tree -> click yes in the dialogue box -> leave all the options as their defaults -> click on ‘Compute’
12. After 3-5 minutes (this is why bioinformaticians are always on twitter, lots of 5 minute breaks) your tree should be finished. Now we need to analyse the tree.

**MinION™ nanopore sequencing Library Preparation**

Story:

Public Health colleagues in France and Germany have both reported an increase in the number of HUS cases over the week-end. The French have reported PCR data suggesting that the Shiga toxin subtype is a highly pathogenic type and the Germans have reported that the isolate is multidrug resistant. You know that both these properties are difficult to analyse using short read Illumina data so you decide to try out the new-fangled MinION machine to sequence the outbreak strain so determine whether the data enables you to improve your analysis of the accessory genome. This will involve the following:

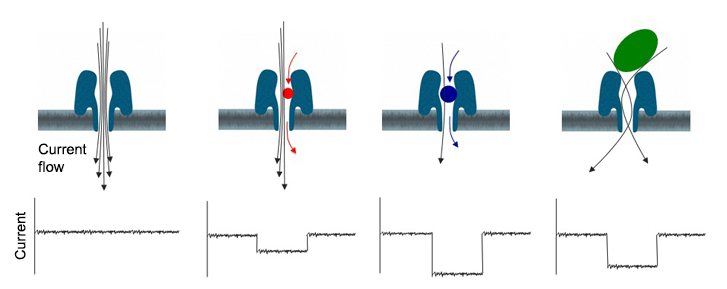
* MinION™ nanopore sequencing Library Preparation
* MinION data analysis

In contrast to the Illumina protocol, the Oxford nanopore MinION uses intact DNA strands and data is analysed in real time.

**PRINCIPLE OF PROCEDURE**

Sample DNA is prepared so that it has a hairpin structure at its end, allowing both stands to be read (sense and anti-sense).

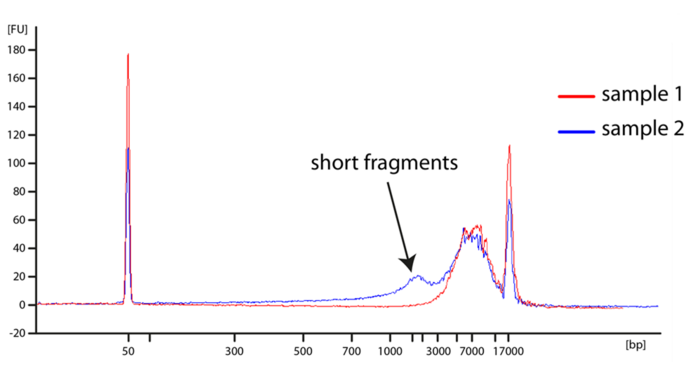
Samples are loaded onto the flowcell which contain nanopores set in a polymer membrane. An ionic current passes through the nanopores and as single molecules enter the pore this current is disrupted. Each of the DNA bases, G, A, T and C, creates a characteristic disruption in current allowing the molecule to be identified.



**DNA Fragmentation (Optional)**

|  |  |  |
| --- | --- | --- |
| Item | Quantity | Storage |
| <1µg DNA sample | 1 tube | Ice bucket |
| Covaris g-TUBE | - | Room Temperature |

1. Transfer <1µg genomic DNA in 46µl to the Covaris g-TUBE
2. Spin the g-TUBE for 1 minute
3. Remove and check all the DNA has passed through the g-TUBE
4. If DNA remains in the upper chamber, spin again for 1 minute at the same speed
5. Remove g-TUBE, invert the tube and replace into the centrifuge
6. Spin the g-TUBE for 1 minute to collect the fragmented DNA
7. Remove and check the DNA has passed into the lower chamber
8. If DNA remains in the upper chamber, spin again for 1 minute
9. Transfer the 46 µl fragmented DNA to a clean 1.5 ml Eppendorf DNA LoBind tube
10. Analyse 1 µl of the fragmented DNA for fragment size, quantity and quality using the Agilent Bioanalyzer (optional)
11. Below is an example of a successful fragmentation (sample 1) and an unsuccessful fragmentation (sample 2). The trace obtained for sample 2 shows a shoulder as a result of the presence of smaller fragments; this is indicative of substantial shearing/degradation of the input material and is likely to reduce the quality of the library preparation and the read length distribution.



**End-prep**

|  |  |  |
| --- | --- | --- |
| Item | Quantity | Storage |
| NEBNext End repair / dA-tailing | 1 tube | Ice bucket |
| Freshly prepared 70% ethanol | 500µl | Room Temperature |
| Agencourt AMPure XP beads | 1 tube | Room Temperature |

1. In a 1.5ml DNA LoBind Eppendorf tube mix together the following:

|  |  |
| --- | --- |
| Reagent | Volume |
| <1 ug DNA | 45µl |
| Nuclease Free water | 5µl |
| Ultra II End-prep reaction buffer | 7µl |
| Ultra II End-prep enzyme mix | 3µl |
| Total | **60µl** |

1. Mix gently by inversion and spin down.
2. Using a thermal cycler, incubate for 5 minutes at 20°C and 5 minutes at 65°C
3. Resuspend AMPure XP beads by vortexing.
4. Add 60µl of resuspended AMPure XP beads to the end-prep reaction and mix by pipetting.
5. Incubate on a rotator mixer for 5 minutes.
6. Spin down the sample and pellet on a magnet. Keep the tube on the magnet, and pipette off the supernatant.
7. Keep on magnet, wash beads with 200µl of freshly prepared 70% ethanol without disturbing the pellet. Remove the ethanol using a pipette and discard. Repeat.
8. Spin down and place the tube back on the magnet. Pipette off any residual ethanol. Briefly allow to dry.
9. Remove the tube from the magnetic rack and resuspend pellet in 31µl nuclease-free water. Incubate for 2 minutes.
10. Pellet beads on magnet until the eluate is clear and colourless.
11. Remove and retain 31µl of elute into a clean 1.5ml Eppendorf DNA LoBind tube.
12. Quantify 1µl of end-prepped DNA using a Qubit fluorimeter – aiming to recover >700ng.

**Ligation of Barcode Adapter**

|  |  |  |
| --- | --- | --- |
| Item | Quantity | Storage |
| NEB Blunt/TA ligase master mix | 1 tube | Ice bucket |
| Barcode Adapter | 1 tube | Ice bucket |
| Freshly prepared 70% ethanol | 500µl | Room Temperature |
| Agencourt AMPure XP beads | 1 tube | Room Temperature |

1. Label a 1.5ml Eppendorf DNA LoBind tube with your group number.
2. Add the following reagents, mixing by inversion between each sequential addition:

|  |  |
| --- | --- |
| Reagent | Volume |
| End prep DNA | 30µl |
| Barcode Adapter | 20µl |
| Blunt/TA Ligase Master Mix | 50µl |
| Total | **100µl** |

1. Mix gently by inversion and spin down.
2. Incubate the reaction for 10 minutes at room temperature.
3. Resuspend the AMPure XP beads by vortexing.
4. Add 40µl of the resuspended beads to the reaction and mix by pipetting.
5. Incubate on a rotator mixer for 5minutes.
6. Spin down the solution and pellet on a magnet. Once it is clear and colourless aspirate off supernatant.
7. Keep on the magnet, wash beads with 200µl of 70% ethanol without disturbing the pellet.
8. Remove the 70% ethanol using a pipette and discard. Repeat.
9. Spin down and place the tube back on the magnet. Pipette off any residual 70% ethanol. Briefly allow to dry.
10. Remove the tube from the magnetic rack and resuspend pellet in 25µl nuclease-free water. Incubate for 2 minutes.
11. Pellet beads on magnet until the elute is clear and colourless.
12. Remove and keep 25µl of elute into a clean 1.5ml Eppendorf DNA LoBind tube labeled with your group number.
13. Quantify 1µl of end-prepped DNA using a Qubit fluorimeter.
14. Dilute the library to a concentration of 10ng/µl with nuclease-free water.

**Barcoding PCR**

1. In a fresh 1.5ml Eppendorf tube labeled with your group number, set up a barcoding PCR reaction as follows for each library:

|  |  |
| --- | --- |
| Reagent | Volume |
| PCR Barcode (one of BC1-BC12) | 2µl |
| 10ng/µl adapter ligated template | 2µl |
| LongAmp Taq 2x master mix | 50µl |
| Nuclease-free water | 46µl |
| Total | **100µl** |

1. Mix gently by inversion and spin down.
2. Amplify using the following cycling conditions:

|  |  |
| --- | --- |
| Thermal cycler settings | |
| Total Volume: | 100 µl |
| Parameters: | 95°C for 3 minutes  15 cycles of:  95°C for 15 sec  62°C for 15 sec  65°C for 24 sec  65°C for 10 minutes  Hold at 4°C |

1. Quantify the barcoded libraries using standard techniques, and pool all barcoded libraries in the desired ratios in a 1.5mL DNA LoBind Eppendorf tube.
2. Prepare 1µg of pooled barcoded libraries in 45µ nuclease-free water.

**End-prep**

1. In a 1.5ml Eppendorf DNA LoBind tube, set up the end-prep reaction using NEBNext Ultra II End Repair/dA-tailing module and 1µg pooled DNA in 45µl as follows:

|  |  |
| --- | --- |
| Reagent | Volume |
| DNA | 45µl |
| Ultra II End-Prep buffer | 7µl |
| Ultra II End-Prep enzyme mix | 3µl |
| DNA CS (control strand) | 5µl |
| Total | **60µl** |

1. Mix gently by inversion and spin down.
2. Using a Thermal cycler incubate at 20°C for 5 minutes and at 65°C for 5 mins.
3. Resuspend the AMPure XP beads by vortexing.
4. Add 60 µl of the resuspended beads to the End-Prep reaction and mix by pipetting.
5. Allow the DNA to bind to the beads by rotating for 5 minutes on a Rotating wheel.
6. Spin down the solution and pellet on a magnet. Once it is clear and colourless aspirate off supernatant.
7. Keep on the magnet, wash beads with 200µl of freshly prepared 70% Ethanol without disturbing the pellet. Remove the 70% ethanol using pipette and discard. Repeat.
8. Briefly spin the tube to collect residual liquid at the bottom of the tube. Return the tube to the magnet and aspirate residual wash solution. Briefly allow to air dry.
9. Remove the tube from the magnetic rack and resuspend the pelleted beads in 31µl nuclease-free water, than incubate for 2 minutes at room temperature.
10. Place the tubes on a magnet to pellet the beads, once it is clear and colourless remove and keep 31µl of the eluate into a clean 1.5ml Eppendorf DNA LoBind tube.
11. Quantify 1µl of eluted sample using the Qubit Fluorometer. There should be more than 700ng of material.

**Adaptor Ligation**

1. Ensure that the Blunt/TA master mix is mixed thoroughly before use.
2. In a DNA LoBind 1.5ml tube add the following, mix by inversion after each:

|  |  |  |
| --- | --- | --- |
| Reagent | | Volume |
| Nuclease free water | 8µl | |
| End-Prepped DNA | 30µl | |
| Adapter Mix | 10µl | |
| HP Adapter | 2µl | |
| Blunt/TA Ligase Master Mix | 50µl | |
| Total | **100µl** | |

1. Briefly spin down in a microfuge.
2. Incubate for 10mins at room temperature.
3. Add 1µl HP tether, mix by inversion, briefly spin down in a microfuge and incubate for 10 mins at room temperature.

**MyOne C1 bead preparation**

1. Resuspend MyOne C1 beads by vortexing until homogenous.
2. Take 50µl of resuspended MyOne C1-beads and transfer to a clean 1.5ml DNA LoBind tube. Pellet the beads on a magnet and aspirate off and discard the supernatant.
3. Add 100µl Bead Binding buffer to the pelleted beads. Resuspend beads by vortexing. Place the tube on a magnet, allow beads to pellet and aspirate off and discard supernatant.
4. Repeat the wash in step 3.
5. Add 100µl Bead Binding buffer to the pelleted washed beads. Resuspend beads by vortexing. These are the ‘Washed beads’ required for the subsequent purification.

**Library Purification**

1. To the adapter-ligated, tether-bound DNA add 100µl ‘washed beads’, carefully mix by pipetting and incubate at room temperature for 5 mins on a rotating wheel.
2. Place the tube on a magnetic rack; allow beads to pellet and pipette off the supernatant.
3. Resuspend the pelleted beads in 150µl Bead Binding Buffer. Place the tube on a magnet, allow beads to pellet. Aspirate off the Bead Binding Buffer using a pipette and discard.
4. Repeat step 3.

**Elution of library form MyOne C1-beads**

1. Resuspend the pelleted beads in 25µl of Elution Buffer by pipetting up and down. Incubate for 10 minutes at 37°C.
2. Pellet the beads on a magnet until the eluate is clear and colourless.
3. Remove and retain 25µl of eluate which contains the library into a clean 1.5ml Eppendorf DNA LoBind tube.
4. Place the tube of library on ice until required for library loading.
5. Quantify 1µl of eluted sample using the Qubit fluorometer. One should expect to retain more than 100ng of material.

**Priming the Flow Cell**

**IMPORTANT:** Thoroughly mix the contents of the RBF tube by inversion or pipetting, and spin down briefly. The library is loaded dropwise without putting the pipette tip firmly into the port.

Take care to avoid introducing air during pipetting.

1. Flip back the MinION lid and slide the sample port cover clockwise to that the sample port is visible.



1. Ensure that the sample port cover is fully opened (a 90 ° clockwise turn).
2. Check for small bubbles under the cover. Draw back a small volume to remove any bubbles. Check that there is continuous buffer from the priming port, across the sensor array to the outlet channel of the flow cell.



1. Prepare the flow cell priming mix in a clean 1.5 ml Eppendorf DNA LoBind tube.

|  |  |
| --- | --- |
| Reagent | Volume |
| RBF | 500µl |
| nuclease-free water | 500µl |
| Total | **1000µl** |

1. Load 800µl of the priming mix into the flow cell via the priming port and wait 5 minutes, avoiding the introduction of air bubbles.
2. Gently lift the activator to make the SpotON sample port accessible.
3. Load 200µl of the priming mix into the Flow Cell via the priming port.

**Loading a Library**

1. Prepare the library for loading as follows:

|  |  |
| --- | --- |
| Reagent | Volume |
| RBF | 37.5µl |
| LLB | 25.5µl |
| Adapted and tethered library | 12µl |
| Total | 75µl |

1. Mix gently by pipetting just prior to loading.
2. Load 75µl of the prepared library into the flow cell via the SpotON sample port in a dropwise fashion. Ensure each drop lows into the port before adding the next.
3. Gently replace the activator, making sure the bung enters the SpotON port, close the sample port and replace the MinION lid.

**Starting a sequencing run**

1. Once a MinION and Flow Cell are connected, a Label Experiment dialogue box appears. Click into the Sample ID box and name your sample using free text in alphanumeric format only.
2. Click into the FlowcellID box and enter the Flow Cell ID, which is the code found on a sticker on the top side of a Flow Cell.
3. Select the appropriate protocol script.

* Experiment type: **Choose Operation**
* Flow Cell product code: Choose the Flow Cell type under **Flow cell product code**
* Sequencing kit: Choose the sequencing kit version you have used to prepare the library
* Whether or not live basecalling is enabled
* A drop-down menu will appear with the most relevant sequencing scripts. Select the script you need.

1. Start the script using the Execute button at the bottom of the Connections page.
2. Select the appropriate protocol script using the Start Protocol dialogue box.

* In the MinKNOW GUI click Start Protocol and the Run Protocol Script window opens
* Click on the down arrow in the Select Protocol Script box to reveal the protocol script choice
* Select the protocol script required

**Progression of MinKNOW protocol script**

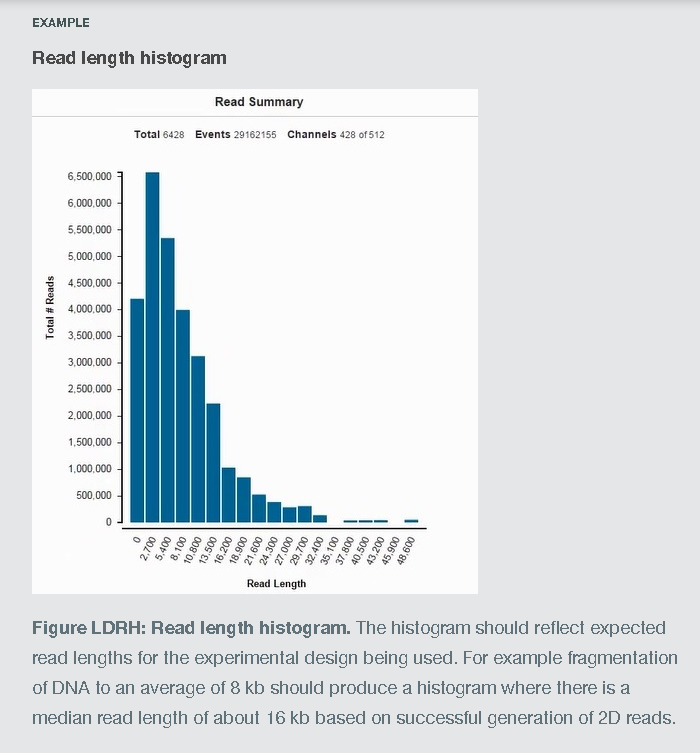
1. Check that the number of active pores reported in the MUX scan are similar to those reported at the end of the Platform QC.

* If there is a significant reduction in the numbers try to restart MinKNOW from the control panel.
* If the numbers are still significantly different close down the host computer and reboot.
* When the numbers are similar to those reported at the end of the Platform QC the experiment being carried out can be restarted on the Connection page; there is no need to load any additional library after restart.

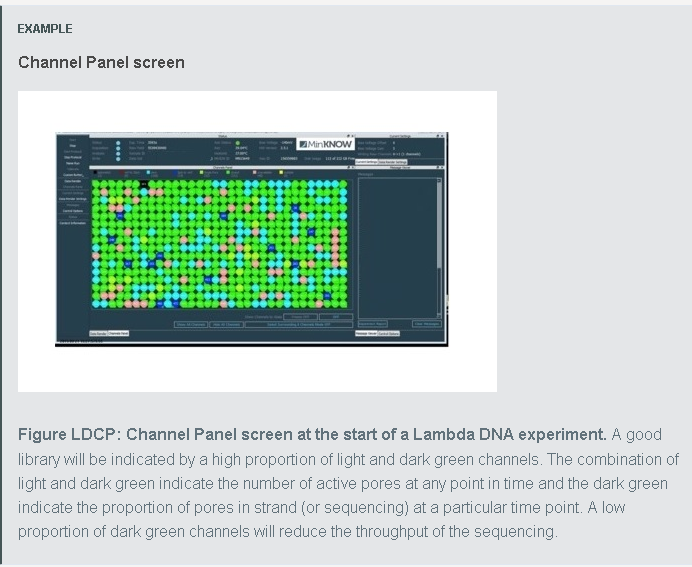
1. Check the heatsink temperature is approximately 34 °C.

* The MinION is able to maintain a heatsink temperature of 34°C on a typical lab bench when the local ambient conditions are between 19.5 °C and 24.5 °C. However, there are a number of external factors which can disrupt the local conditions and which need to be taken into account, for example warm air expelled from laptops, or cool air from a fan or air conditioning system increasing airflow around the MinION.
* The MinION takes approximately 10 minutes to get to temperature.

1. Monitor the development of the read length histogram.



1. Check pore occupancy by looking at the panel at the top of the Status or Physical Layout views.



1. Ongoing monitoring of the experiment is best achieved using the VIEW REPORT in Desktop Agent.

Story:

It is Wednesday 25th January and the Illumina sequencing data from the run you prepared on Monday is ready for analysis. Isolates from nine cases of HUS from five different hospitals in the UK are ready for analysis.

**Quality Assessment & Assembly of Sequencing Data**

Next generation sequencing is already having a profound impact on clinical microbiology with the potential to resolve pathogens to the resolution of a single strain. Assemblies of pathogen whole genomes allows for the genetic repertoire to be unveiled providing insights on pathogenicity, antibiotic resistance and other clinically important features. The genetic fingerprint derived from the whole genome sequence can also be used to compare strains to provide unparalleled resolution for elucidating outbreaks and transmission routes.

1. **The FASTQ format**

Data from Illumina machines (HiSeq/MiSeq) will usually be received as FASTQ files.

@M00882:12:A1CJ2:1:1101:14996:1321 1:N

NCGGATACACCAGAGAACGATGCAGCCTTCCCCCGCCAGACACATGCCGGGAACTCGGCGCTCTACCCGCAGGTCAAAATGGTCTGCCAGATGGAACTGACCAGCCATCTGCTGACGGCTGCAAGAACAGCGAAAATGAGCTTGCTGAGCA

+

#5<????BDDDDDDDDFFGGGGHIHHIIIIIHIHHHHHHIIHHIHHHIHHHHHHIHEHHHHHHFHHHHEHHHH@FFGGGF?FFGGGGGGGGGGEGEGGCCECEECGGGGGGGGGGGEGGGDGGGGGGCCEE<GAADC?CCEGEGEGEEEGG

The first line contains the sequence identifiers:

* **M00882**  The unique instrument name
* **12** The run id
* **A1CJ2** The flowcell id
* **1** Flowcell lane
* **1101** Tile number within the flowcell lane
* **14996** 'x'-coordinate of the cluster within the tile
* **1321** 'y'-coordinate of the cluster within the tile
* **1** The member of a pair, 1 or 2 (paired-end or mate-pair reads only)
* **N** Y if the read fails filter (read is bad), N otherwise

The second line contains the sequence of the read (ACTG or N for unresolved)

The final line is the quality score for each base. The quality is the probability that the corresponding base call is incorrect. It is referred to as a Phred quality score and is derived as follows:

Q_\text{sanger} = -10 \, \log_{10} p

The Phred score is encoded as an ASCII character by adding 64 to the Phred value.

Q=30 means probability base is incorrect is 1/1000

Q=20 means probability base is incorrect is 1/100

For paired end libraries (as prepared yesterday) you will receive two FASTQ files per sample one for each pair.

**Exercise 1. Assessing the quality with FastQC**

**Load the FASTQs**

1. Log on to the Public Health England Galaxy Server
2. Click on the ‘Shared Data’ Tab
3. Click on ‘Data libraries drop down’
4. Click on the ‘WTAC’ & ‘2017’ & ‘MiSeq’ Folder
5. Find your groups FASTQ files that you produced on Monday (if in the unlikely scenario your MiSeq run failed grab the two FASTQs STEC.R1.fastq.gz and STEC.R2.fastq.gz)
6. Click the check boxes next to them and select ‘Import to current history’
7. Click on the ‘Analyze Data’
8. Notice the FASTQ files in your history – Let’s have a look at one.

**Assess the quality with FASTQC**

1. From the Tools on the left hand side click ‘NGS: QC and manipulation’
2. Click ‘FastQC:Read QC’
3. Select one of the FASTQ reads and Click ‘Execute’
4. Note the job running in your History – this will take a couple of minutes…..
5. When it’s finished (GREEN) click on the ‘eye’ icon to see the results

**Q1. How many sequences in the file?**

**Q2. What is the %GC?**

**Q3. At what position does the average quality fall below Q30?**

Explore the other statistics provided by FASTQC

It may be necessary to perform operations on the FASTQ files to exclude or correct low quality regions. This can involve trimming the end of reads where the quality is low (see FASTX toolkit) or using K-mer frequency distributions to correct or discard reads (see QUAKE/MUSKET/trimmomatic).

**Remove bad quality reads with Trimmomatic**

1. From the Tools on the left hand side click ‘NGS: QC and manipulation’
2. Click Trimmomatic
3. Select the **R1.fastq** as Direction 1 fastq reads to trim
4. Select the **R2.fastq** as Direction 2 fastq reads to trim
5. Change the Quality encoding to **phred33**
6. Deselect Perform Sliding Window trimming
7. Change the minimum quality to trim leading bases to **30**
8. Change the minimum quality to trim trailing bases to **30**
9. Change the minimum read length to **50**
10. Click ‘Execute’
11. Note the job running in your History – this will take a couple of minutes…..

**Let’s reassess the quality with FASTQC**

1. From the Tools on the left hand side click ‘NGS: QC and manipulation’
2. Click ‘FastQC:Read QC’
3. Select one of the Trimmomatic paired output FASTQ reads and Click ‘Execute’
4. Note the job running in your History – this will take a couple of minutes…..
5. When it’s finished (GREEN) click on the ‘eye’ icon to see the results

**Q1. How many sequences in the file?**

**Q2. What is the %GC?**

**Q3. At what position does the average quality fall below Q30?**

Explore the other statistics provided by FASTQC – Our sequence data is much nicer now!

**Analysing the outbreak isolates through the PHE *E.* *coli* pipeline**

Now we have high quality data it’s time to process the FASTQs as previously completed for the German data. This will involve as before; k-mer identification, MLST typing, *in silico* serotyping, pathotyping, Shiga toxin sub-typing and AMR typing. If you can’t remember how to run these components refer to yesterday’s section in the manual.

* 1. Based on these results do you think your isolate is likely to be the part of the same outbreak as the German isolate?
  2. Based on these results do you think that all the isolates from the UK are likely to be part of the same outbreak?

**Phylogenetic Analysis of Outbreak Data**

Questions to answer during this section

1. How many SNPs different from each other are the UK isolates?
2. How many SNPs different are they from the German outbreak strain?
3. What can you conclude about (i) the links between all the UK cases and (ii) the relationship between the UK cases and the German outbreak?

Practical steps

1. We are going to assume that we have already carried out the mapping and SNP calling for all these isolates, so they are already in the database waiting to be queried.
2. On the menu bar go to ‘shared data’ -> Data libraries -> WTAC -> 2017-> Outbreak. From the outbreak folder add the ‘strain\_list\_2’ file to your current history Return to the ‘Analyze Data’ page.
3. On the tool menu, go to NGS: PHE internal tools and then ‘get the snps’. In the main analysis window select ‘strain\_list\_2’ as the strain list, set ‘O26\_wtac\_config.txt’ as the config file, select alignment type as ‘accessory’. Select ‘No’ to the question ‘Would you like the reference genome in alignment’ and ‘Yes’ to produce a distance matrix of SNP distances and list of annotated variants. Click execute.
4. After 1-2 minutes, the item in your history should turn from yellow to green. Click on the eye to inspect the results. This is your ‘pseudo-sequence’ of all the variant positions.
5. Download the pseudo-sequence and load it into MEGA. Follow the same steps as in the previous section to derive a tree.

***de novo* assembly and MinION data analysis**

1. ***de novo* Assembly**

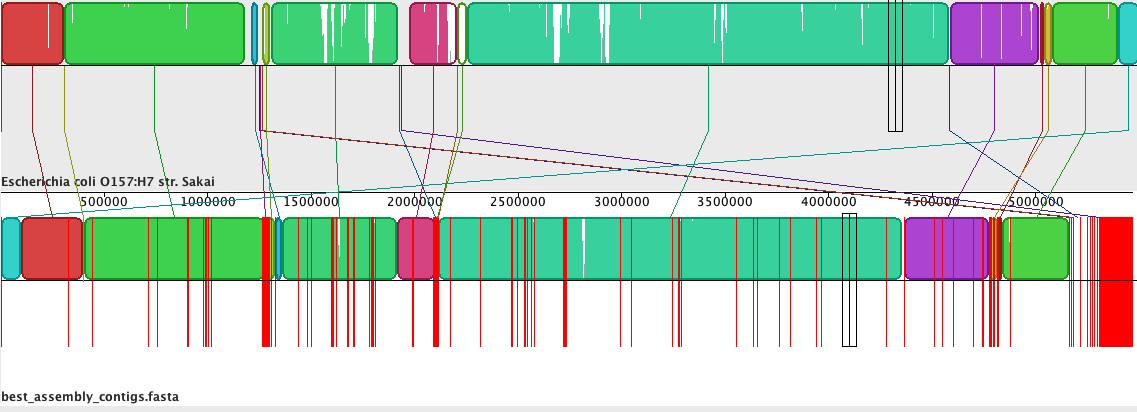
*de novo* assembly refers to the assembly of a genome with no reference to guide the process. A common analogy is putting a jigsaw of several hundreds thousands of pieces together without the picture to refer to.

**Vocabulary:**

* **Read** Any sequence that comes out of the sequencer
* **Paired read** read1, gap < 500 bp, read2
* **Mate-pair** read1, gap > 1 kbp, read2
* **k-mer** Any sequence of length k
* **Contig** gap-less assembled sequence
* **Scaffold** sequence which may contain gaps (N)

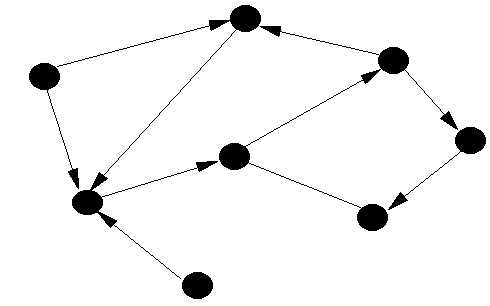
Example of a reference genome and an assembly aligned to it.

Nb. The assembly is fragmented into many contigs



**Assembly Theory**

*de novo* assembly uses graph theory. A graph is a set of nodes and a set of edges. The edges can have direction.



The graphs used to assemble short read data (illumina) uses de Bruijn graphs. Graphs allow the representation of overlaps between reads.

**Example – de Brujin graphs**

**Nodes** represent all k-mers (k-length sub-strings) present in the reads

**Edges** represent the overlap between the k-mers observed.

**E.G. 1 k=3 Single read**

Read ACTG

Graph ACT CTG

**E.G. 2 k=3 3 reads**

Read1 ACTG

Read2 CTGC

Read3 TGCT

Graph ACT CTG TGC GCT

**E.G. 3 k=3 4 reads – 1 has an error**

Read1 ACTG

Read2 CTGC

Read3 CTGA

Read4 TGCT

Graph ACT CTG TGC GCT

TGA

**E.G. 4 k=3 6 reads – What is the effect of repeats?**

Read1 A**CTG**

Read2 **CTG**C

Read3 TGCT

Read4 G**CTG**

Read5 **CTG**A

Read6 TGAC

Graph ACT CTG TGC

GAC TGA GCT

**How does one assemble using a graph?**

To generate contigs from our graph we calculate all node-disjoint paths through the graph.

Node-disjoint means that two different path cannot share a node.

**Example – Contig construction**

Graph ACT CTG TGC

GAC TGA GCT

Contig1 ACT **CTG** **TGC** CTGCT

GAC TGA **GCT**

Contig2 ACT CTG TGC TGAC

**GAC** **TGA** GCT

Contig3 **ACT** CTG TGC ACT

GAC TGA GCT

Choosing the k-mer value is important! Smaller k-mers increase sensitivity as more likely to observe an overlap, large k-mers increase accuracy.

**Velvet (other *de novo* assemblers are available!)**

Zerbino, D. R.; Birney, E. (2008). "Velvet: Algorithms for de novo short read assembly using de Bruijn graphs". Genome Research 18 (5): 821–829.

**Exercise 2.** Assemble your genome.

**Load the FASTQs**

1. Click on the ‘Shared Data’ Tab
2. Click on ‘Data libraries drop down’
3. Click on the ‘WTAC’->’2017’ & ‘Assembly’ Folder
4. There are two STEC FASTQs called “illumina” – click the check boxes next to them and select ‘Import to current history’

**Run Velvet**

1. From the Tools on the left hand side click ‘NGS: Assembly’
2. Click ‘Velvet Optimiser
3. Set Start k-mer value as **45**
4. Set End k-mer value as **45**
5. Set k-mer search step size as **2**
6. Click Add new Input read libraries
7. Select File type: **shortPaired**
8. Click the tick box are the reads paired and in to different files
9. Select **\_1.fastq** for Read dataset for direction 1:
10. Select **\_2.fastq** for Read dataset for direction 2:
11. Click execute

**How do we assess the assembly?**

- Number of contigs/scaffolds

- Total length of the assembly

- Length of the largest contig/scaffold

- Percentage of gaps in scaffolds (’N’)

- N50 of contigs/scaffolds

-Internal consistency

- Number of predicted genes

**N50**

*N50 length is defined as the length N for which half of all bases in the sequences are in a sequence of length L < N*

or

*Half of the assembled bases reside in contig having a length of at least the n50 contig*

**Assess assembly**

1. How many contigs were produced for each assembly?
2. What was the n50 reported at the end of the logfile
3. What is the size of the largest contig?
4. What is the total size of the assembly?

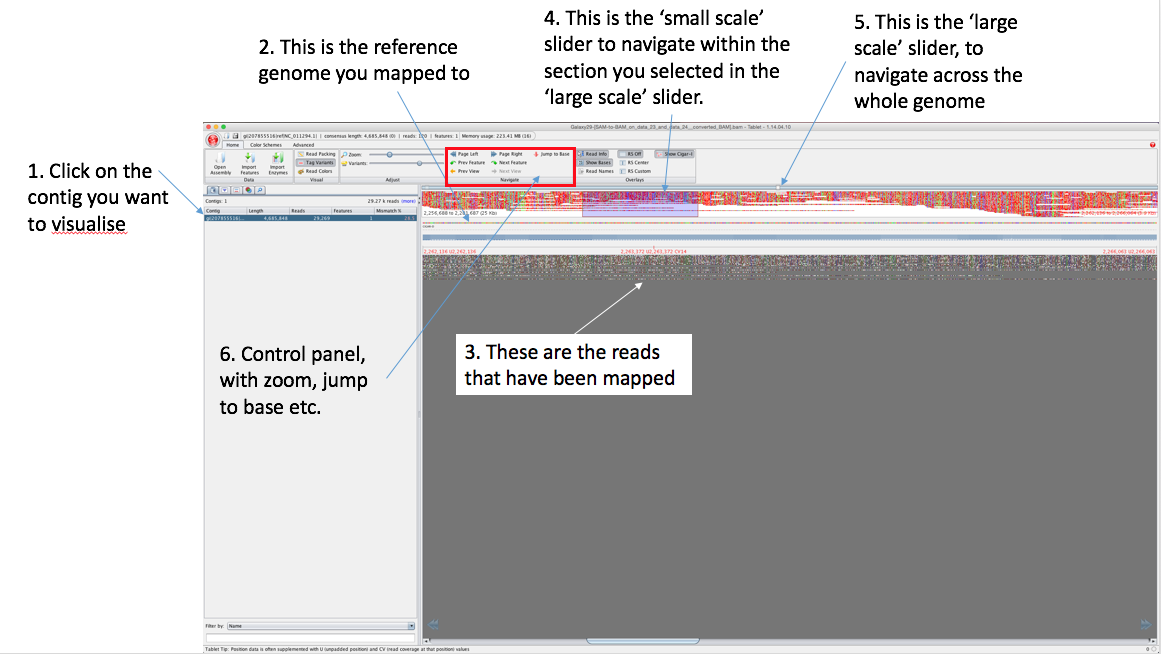
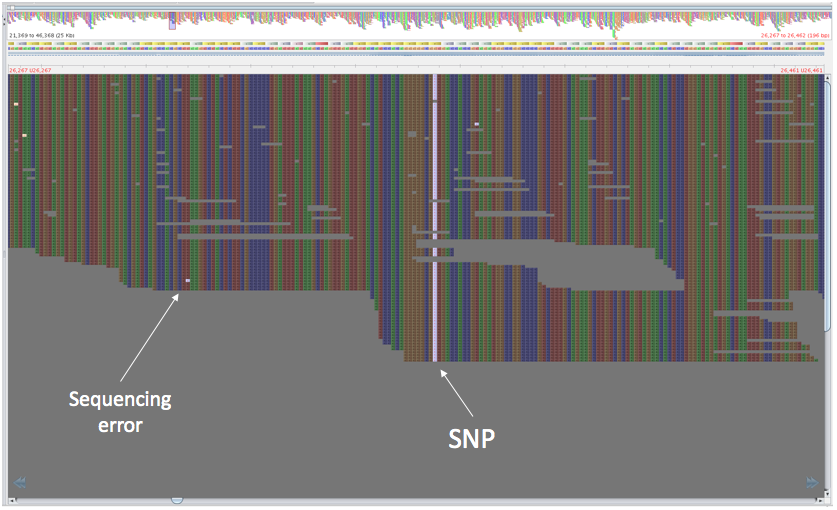
**MinION read analysis**

We are also going to compare Illumina and MinION reads that have been mapped to the same reference as the best way to understand MinION data. We will do this using the Tablet alignment viewer <https://ics.hutton.ac.uk/tablet/>.

Questions to answer during this section

1. How many Illumina reads are mapped to the reference genome? How many MinION reads?
2. Find at least 5 SNPs in the Illumina mapping. Are these also SNPs in the MinION analysed sample? Do the two technologies agree?
3. Would you rather call SNPs directly from the MinION data, or just look for ones you have previously characterised from the Illumina data?

Practical steps

1. If you don’t already have them, from Shared data -> WTAC -> 2017 -> Assembly add illumina.R2.fastq and illumina.R1.fastq to your current history.
2. From Shared data -> WTAC -> 2017 -> Assembly add stec\_reference.fa to your current history
3. In the tool menu, go to NGS:Mapping -> map with BWA -> ‘Will you select a reference genome from your history or use a built-in index?’ - Use one from history -> ‘Select a reference from history’ – ensure you have stec\_reference.fa selected -> ‘Is this library mate-paired?’ – paired-end -> ‘forward fastq’ - illumina .R1.fastq -> ‘reverse fastq’ illumina.R2.fastq -> Execute
4. After 2-3 minutes, your job will be finished. Then, we need to convert the output file (a SAM file) to a BAM file (remember – bioinformatics is basically just advanced file conversion) so that we can view it in Tablet. In the tools menu -> NGS: SAM Tools -> SAM-to-BAM -> ‘Choose the source for the reference list’ – History -> ‘Convert SAM file’ – the output of ‘Map with BWA’ -> ‘Using reference file’ - stec\_reference.fa -> Execute
5. When the SAM -> BAM conversion has finished, click on the output -> then click on the floppy disk to download -> then download the dataset and the bam\_index. An index is a ‘helper file’ that lets other programs quickly access the main file. You also need to download the stec\_reference.fa in the same way.
6. Then, open Tablet  via the Start menu. From the menu bar select ‘open assembly’ -> in ‘Primary assembly’ click ‘Browse’ then navigate to the bam file you just downloaded -> in ‘Reference’ click ‘Browse’ then navigate to the stec\_reference.fa file you just downloaded -> Open
7. See below for a quick guide to the key parts of the Tablet software. This is one of my favourite pieces of bioinformatics software, I hope you like it too!
8. Scroll across the genome, looking for SNPs. Try to find at least 5 SNPs, and not the same ones as the person sat next to you. Mapped bases that are ‘lighter’ are different from the reference genome.
9. See below for an example.
10. Go back to Galaxy, from Shared data -> WTAC -> 2017 -> Assembly add minion.fasta to your current history.
11. From the tool menu -> NGS: PHE internal tools -> mapwithlast -> in ‘Read data from your current history’ select minion.fasta -> in ‘Reference genome from your history’ select stec\_reference.fa -> Execute.
12. After 5 minutes, the alignment should be finished, we need to convert to BAM again. In the tools menu -> NGS: SAM Tools -> SAM-to-BAM -> ‘Choose the source for the reference list’ – History -> ‘Convert SAM file’ – the output of ‘Map with BWA’ -> ‘Using reference file’ - stec\_reference.fa -> Execute
13. Once the job has finished, download the BAM and bam\_index and open in Tablet.
14. Take the variant positions you identified in step 8 and navigate to the same locations in this bamfile using Tablet’s ‘jump to base’ function. Does the MinION data agree with the Illumina data?

**MinION assembly analysis**

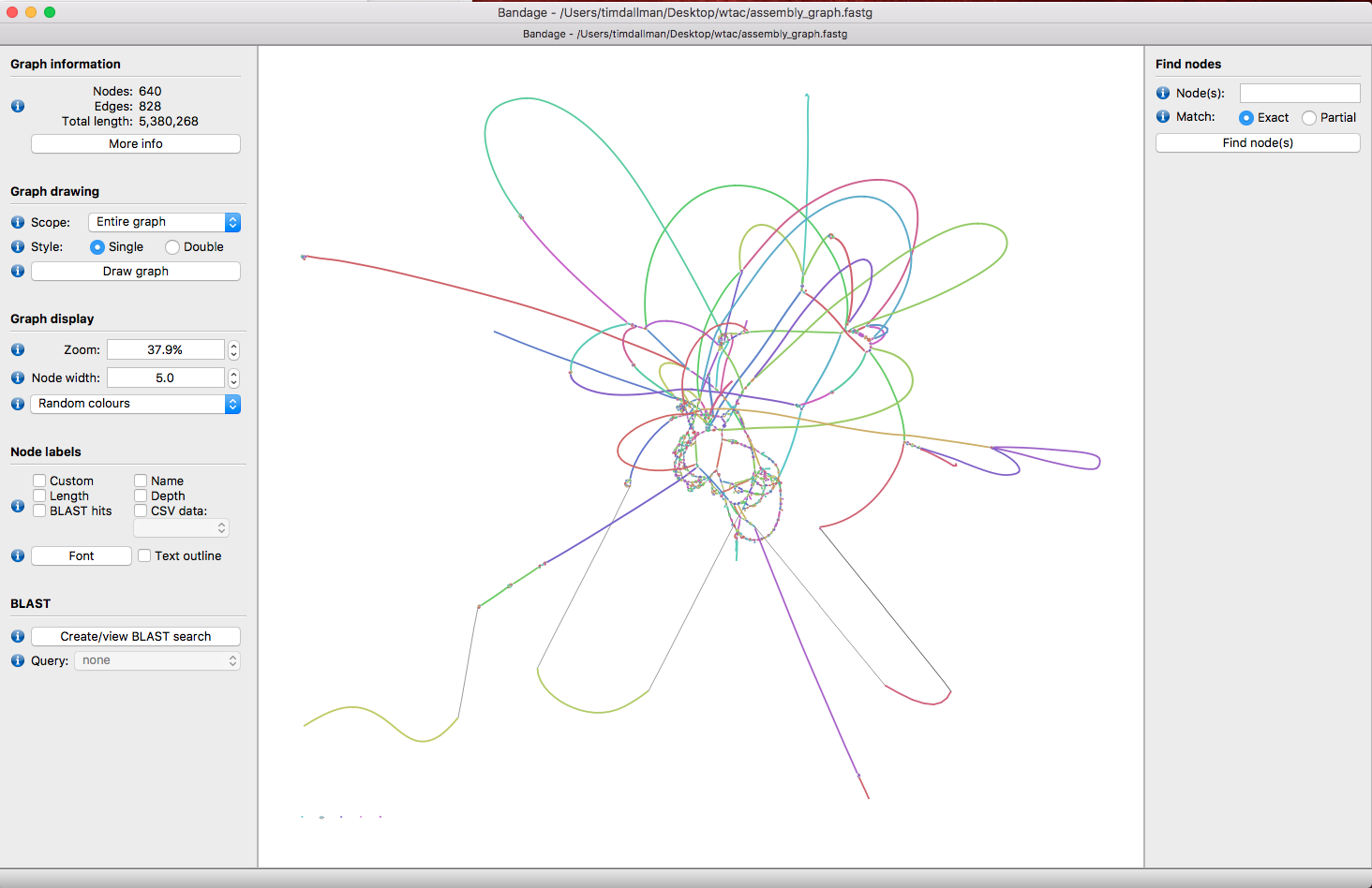
We are also going to compare Illumina and MinION assemblies to see if the longer reads help us understand the accessory genome. To do this we are going to use the program Bandage <http://rrwick.github.io/Bandage/>. Bandage is a program for visualising *de novo* assembly graphs. By displaying connections which are not present in the contigs file, Bandage opens up new possibilities for analysing *de novo* assemblies.

Questions to answer during this section

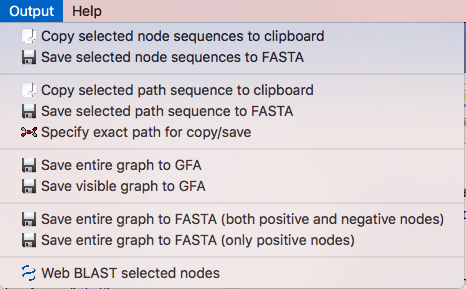
* + 1. How many contigs (nodes) are there in the illumine assembly compared to the minion assembly?
    2. What is the longest node and the n50 from each assembly method?
    3. How big is the bacteriophage containing the Stx toxin?
    4. How many plasmids does the isolate have?
    5. Are the AMR elements on a plasmid or the chromosome?
    6. Does the accessory genome analysis tell us anymore about the possible origin of the strain?

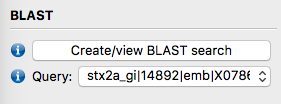
Practical Steps

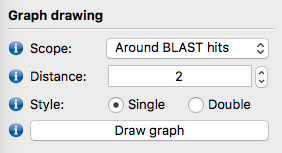
* + - 1. Add the illumina.fastg and minion.fastg assembly graphs and the blast database blast.fa from Shared data -> WTAC -> 2017 -> Assembly to your current history.
      2. Click back to Analyze data and as previous click on the floppy disk icon to download the files to your Desktop.
      3. Open Bandage and load the Illumina assembly graph by clicking ‘Load Graph’ from the File dropdown.
      4. Click on the ‘More Info’ button to get some information about the assembly.
      5. Let’s take a look at the assembly graph – by clicking the ‘Draw Graph’ button – see below for an example. Zoom in examine the connectivity of the contigs.



1. You can select nodes and save the underlying sequence to the clipboard or to a file, we can then BLAST this sequence against the NCBI database by clicking the ‘Output’ tab and clicking ‘Web BLAST selected nodes’.



1. We can also provide our own BLAST database to search for features in the assembly. Click on the ‘Create/view BLAST search button’ to bring up a new window, as below. Click ‘Build BLAST database’ first. Then click ‘from FASTA file’ and select blast.fa that you downloaded to your desktop and you will see the list of AMR and virulence genes loaded. Finally click the ‘Run BLAST search button’ and the BLAST results will appear.
2. Let’s see where some of these genes are in our assembly. Back on the main viewer on the side menu select the ‘stx2a’ gene in the BLAST query drop down menu. 
3. Under the Graph drawing section change the Scope: to ‘Around BLAST hits’ and the Distance to 2. Click the ‘Draw graph’ button.



1. The node with the BLAST match is coloured blue. How big is it? Let’s BLAST the node and surrounding ones against the NCBI database? What are the matches?
2. Let’s look at the other BLAST matches – can you spot any AMR genes?
3. Now let’s repeat with the minion data and load the minion.fastg assembly graph.
4. How has the assembly changes with the different sequence data? What is the biggest contig now?
5. Let’s repeat the BLAST analysis. How big is the contig with the *stx2a* gene now? Can you spot any plasmids?
6. BLAST a contig containing an AMR gene against NCBI?

**Appendix**

**Protocol for Real-time multiplex PCR assay for the detection of *VTEC, Shigella and Campylobacter***

PRINCIPLES OF THE RT-PCR PROCEDURE

This procedure is for the molecular detection of *verocytotoxin-producing E. coli, Shigella and Campylobacter* using a multiplex real-time PCR assay. Each reaction contains an amplification internal control (IC). The purpose of this control is to identify potential inhibition from the processed specimen.

Simultaneous amplification and detection of the different targets in the multiplex PCR assay is achieved through the use of target specific primers and dual-labelled probes. Dual-labelled probes, also known as Taqman™ probes, are oligonucleotides designed to the internal region of a target. They are labelled with two different dyes, a reporter dye at the 5’ end and a quencher dye at the 3’ end. When excited the reporter dye emits fluorescence that is quenched by the quencher dye by Förster-type energy transfer (FRET), resulting in no fluorescence. However, following hybridisation of the probe to target DNA during PCR, the 5’-3’ exonuclease activity of *Taq* polymerase cleaves the hybridised probe. This separates the reporter dye from the quencher dye, preventing FRET and resulting in the emission of fluorescence (Figure 5.1). Fluorescence increases in proportion to the rate of probe cleavage during each PCR cycle, which is displayed on the real-time PCR instrument (Figure 5.2). In order to compare and interpret the fluorescent data a threshold is set. This enables a cycle threshold (ct) value to be assigned to each sample, which is the point at which the fluorescence of the sample crosses the threshold. A sample with no ct value is considered negative and a sample with a ct value is considered positive. The probes for the different targets in this assay are labelled with different reporter dyes so that they can be detected on different channels of the real-time PCR instrument (see table 5.1). There are a number of different real-time PCR instruments (platforms) on the market including Smartcyclers, Lightcyclers, Taqmans. Each instrument has different advantages and disadvantages. For this assay a RotorGene is used.

**Table 5.1:** Labels used for the multiplex probes and the channels they are detected in.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Organism | Label | Max Em (nm) | Max Abs (nm) | Channel |
| Campylobacter | Fam | 520 | 494 | Green |
| VTEC | Yak | 549 | 530 | Yellow |
| Shigella | Cy5 | 662 | 646 | Red |
| GFP | Rox | 605 | 575 | Orange |

Fluorescing

Reporter

Target DNA

Taq polymerase

5’ Reporter

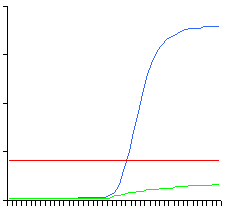
3’ Quencher

1. Dual-labelled probe – no fluorescence due to FRET

2. Dual-labelled probe binds to target DNA

3. Taq polymerase cleaves the probe – fluorescence, no FRET

**Figure 5.1 Diagram of a dual-labelled probe**



Number of cycles

Amount

of

fluorescence

Threshold

Ct

value

Positive sample

Negative sample

**Figure 5.2 Diagram illustrating RT-PCR data**

1. The Y-axis is the amount of fluorescence detected, the X-axis is the number of PCR cycles and the ct (cycle threshold) value is the point at which the positive sample’s (in blue) fluorescence crosses the set threshold (straight line in red). The negative sample’s (green) fluorescence does not cross the threshold and therefore does not have a ct value.

Real-time probe based PCR assays can be performed using exactly the same reagents used for conventional PCR assays, however there are a number of companies that produce specifically designed mixes, referred to as quantitative PCR (QPCR) mixes. These are most commonly supplied as two-times concentrates that solely require the addition of the primers and probe, thus reducing pipetting errors. The Invitrogen Platinum® QPCR SuperMix-UDG is used for this assay. It contains dUTPs instead of dTTPs and uracil DNA Glycosalase (UDG). These components reduce contamination by amplimer cross over. If amplimers generated with dUTPs are carried over into a new reaction the UDG is able to render the amplimer unamplifiable by cleaving the N-glycosidic bonds between the uracil bases and the phosphodiester backbone (Longo *et al*, 1990).

Enteric pathogens multiplex PCR assay protocol

Equipment:

* Gilson pipettes (P1000, P200, P20, P10)
* Plugged tips (1 ml, 200 μl, 20 μl & 10 μl)
* 0.2 ml flat cap thin walled PCR tubes
* 1.5 ml Eppendorf tubes
* Tube racks
* 4°C refrigerator
* 20°C freezer
* Disposable gloves
* Rotor-Gene (or other Real-time PCR instrument)
* Discard jar

Reagents:

* DNA extracts (samples and controls)
* Invitrogen Platinum® QPCR SuperMix-UDG (-20°C)
* PCR grade sterile distilled water
* Primers
* Probes

Procedure:

1. Remove a tube of Invitrogen Platinum QPCR mix and probes from the freezer to defrost. (The probes should be in an amber tube or kept in the dark as they are light sensitive).
2. Determine the number (N) of PCR reactions as below:

N = the number of samples + 4 controls (1 negative control, 1 Campylobacter positive control, 1 Salmonella positive control, 1 Shigella positive control) + 2 extra (to allow for mix loss during pipetting).

Each group will test two samples, therefore N = 8 (enough for 6 reaction tubes)

1. Place 6 x 0.2 ml flat capped tubes into a tray and label them appropriately as below including your group number:

S 1 = Unknown sample

S2 = Unknown sample

NEG = Negative control

CAM = *Campylobacter species* positive control

SAL = *Salmonella species* positive control

SHI = *Shigella species* positive control

In a sterile 1.5 ml eppendorf prepare the mastermix:

|  |  |  |
| --- | --- | --- |
|  | x 1 | x 8 |
| Invitrogen Plat Supermix | 12.5 µl |  |
| MgCl2 | 1.5 µl |  |
| Primer mix (2.5µM) | 2.5 µl |  |
| Probe mix (2 µM) | 2.5 µl |  |
| IC | 1 µl |  |

Mix the mastermixes by inverting 10 times

Pipette 20 µl of mastermix into the 0.2 ml tubes.

Add 5 µl of the sample / controls to the appropriately labelled reaction tubes.

Place the tubes into the Rotor-Gene. Ensure that the tubes are at an angle so they flat in the rota, secure them with the metal ring that screws on top of them and then close the RotorGene.

Then set up the RotorGene:

* Select new run
* Select the ‘Capsular screen’ program.
* Select the 36 well run and confirm there are no 0.2 ml domed cap tubes in the rotor.
* Ensure the reaction volume is set to 25 µl

• Go to edit profile to check the parameters of the program are correct:

Hold: 95°C 3 minutes

Cycling: 95°C 15 seconds

60°C 45 seconds (acquiring on green, yellow, red, and orange channels)

• Click on ‘Start run’.

• Go to the sample option on the main menu and enter the appropriate reaction tube details for each position on the rota.

The RotorGene will then display the program details, the stage it is at and the amount of fluorescence obtained for each sample.

When the run is complete remove the tubes from the RotorGene and discard them in to a discard jar.

Then select ‘Analysis’ from the main menu, followed by the ‘Quantification’ option and then click on show. Set the threshold to 0.05. Repeat this process for each channel.

Record which samples have ct values and for which channels (it is useful to look at the raw data as well as the analysed data to ensure amplification has really occurred). Determine if the assay has worked (all controls correct) and if the sample was positive for any of the targets.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Campy  PCR | Salmonella  PCR | Shigella  PCR | IC | Result |
| S1 |  |  |  |  |  |
| S2 |  |  |  |  |  |
| NEG |  |  |  |  |  |
| CAM |  |  |  |  |  |
| SAL |  |  |  |  |  |
| SHI |  |  |  |  |  |