NoVAS Metagenomic Sequencing of Fresh Produce  
Report to Public Health England and the NoVAS Consortium

## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Summary</td>
<td>2</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>3</td>
</tr>
<tr>
<td>Background to the study</td>
<td>3</td>
</tr>
<tr>
<td><strong>Aims and objectives</strong></td>
<td>3</td>
</tr>
<tr>
<td>Structure of the report</td>
<td>3</td>
</tr>
<tr>
<td>2. Methodology adopted</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Sample Selection</td>
<td>4</td>
</tr>
<tr>
<td>2.2 Enrichment for viral RNA</td>
<td>5</td>
</tr>
<tr>
<td>2.3 Library preparation</td>
<td>5</td>
</tr>
<tr>
<td>2.4 Sequencing</td>
<td>5</td>
</tr>
<tr>
<td>2.5 Analysis</td>
<td>6</td>
</tr>
<tr>
<td>3. Results</td>
<td>7</td>
</tr>
<tr>
<td>3.1 Sequencing</td>
<td>7</td>
</tr>
<tr>
<td>3.2 Analysis</td>
<td>9</td>
</tr>
<tr>
<td>4. Conclusions and recommendations</td>
<td>12</td>
</tr>
<tr>
<td>Annex A: References</td>
<td>13</td>
</tr>
<tr>
<td>Annex B: SOP NGSL3b ScriptSeq Library Preparation</td>
<td>14</td>
</tr>
</tbody>
</table>
Executive Summary

Nucleic Acid extracts from twelve raspberry and lettuce samples that were determined to be positive for Norovirus contamination by qPCR assay were selected for metagenomic sequencing. Each sample was subdivided, and treated in two ways in an attempt to enrich for Norovirus RNA. These two methodologies were poly(A) capture (which attempts to enrich for poly(A)-tailed RNA molecules, including Norovirus RNA) and ribosomal depletion (which removes plant ribosomal RNA, increasing the proportion of viral RNA present). After enrichment Next Generation Sequencing libraries were prepared, and sequenced on the illumina MiSeq platform. Sequencing was successful, generating between 700,000 and 5.9 million sequence reads per sample. After analysis, sequence from the host plants, plant viruses, and the viral extraction control (Mengovirus) were detected in many of the samples. However no Norovirus sequences were detected in any sample. This could be explained by an extremely low titre of Norovirus being present on the fresh produce samples, combined with low quantities of input RNA.
1. Introduction

Background to the study

Aims and objectives

An aim of the NoVAS Norovirus attribution project was the generation of Norovirus genome sequence from the various samples tested. As Fera had received and processed produce samples (lettuce, and fresh and frozen raspberries), Fera was asked to undertake Next Generation Sequencing on these samples to contribute to this aim.

Fera’s work involved two separate enrichment protocols for Norovirus RNA, sequencing using our established RNA sequencing protocols (modified to account for low RNA inputs), and analysis to identify any Norovirus sequences present.

Structure of the report

The report has the following structure:

- Chapter 2 outlines the methods used in the study
- Chapter 3 describes the results arising from the study
- Chapter 4 provides conclusions and our recommendations

2. Methodology adopted

2.1 Sample Selection

Samples were selected to increase the likelihood of obtaining Norovirus sequence. To that end, nucleic acid extracts from the twelve samples with the highest number of positive reactions in duplicate of the Norovirus qPCR assays were identified (out of eight reactions per sample, each performed in duplicate). These were then taken forward for sequencing, as they were interpreted as being likely to have the highest titre of Norovirus RNA present. RNA concentrations were then determined using the Qubit HS RNA kit (Life Tech).

The samples selected for sequencing are show in Table 1.

Table 1. Samples selected for metagenomic sequencing.

<table>
<thead>
<tr>
<th>Number of duplicate reactions positive for Norovirus by qPCR (out of 8)</th>
<th>Fera Sample Reference</th>
<th>Commodity</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>429</td>
<td>Lettuce</td>
</tr>
<tr>
<td>3</td>
<td>143</td>
<td>Raspberry</td>
</tr>
<tr>
<td>3</td>
<td>155</td>
<td>Raspberry</td>
</tr>
<tr>
<td>3</td>
<td>196</td>
<td>Raspberry</td>
</tr>
<tr>
<td>3</td>
<td>409</td>
<td>Lettuce</td>
</tr>
<tr>
<td>2</td>
<td>145</td>
<td>Raspberry</td>
</tr>
<tr>
<td>2</td>
<td>180</td>
<td>Lettuce</td>
</tr>
<tr>
<td>2</td>
<td>214</td>
<td>Lettuce</td>
</tr>
<tr>
<td>2</td>
<td>272</td>
<td>Lettuce</td>
</tr>
<tr>
<td>2</td>
<td>415</td>
<td>Raspberry</td>
</tr>
<tr>
<td>2</td>
<td>416</td>
<td>Lettuce</td>
</tr>
<tr>
<td>2</td>
<td>188</td>
<td>Lettuce</td>
</tr>
</tbody>
</table>
2.2 Enrichment for viral RNA

Each of the nucleic acid extracts from the samples identified in 2.1 was subdivided in two. Each subdivision was then subjected to one of two enrichment methodologies designed to increase the proportion of viral RNA in the sample.

Method 1 was poly(A)-capture, designed to enrich for RNA molecules with a poly(A) tail. This includes eukaryotic mRNA as well as some viral genomes, including Norovirus. Poly(A)-capture was carried out using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (NEB #E7490S) according to manufacturer’s instructions.

Method 2 was ribosomal depletion, designed to remove eukaryotic ribosomal RNA from the extracts, increasing the relative proportion of viral RNA. Ribosomal depletion was carried out using the illumina ScriptSeq™ Complete Kit (Plant Leaf) (illumina BPL1224), according to SOP NGSL3b (ScriptSeq Library Prep) (Annex B), with the following modification;

In step 1 of the Ribo-Zero process 4μl of rRNA Removal Solution was added instead of 8μl. This was to enable the maximum volume of nucleic acid possible to be processed, while simultaneously using the ratio of rRNA Removal Solution to Sample recommended for low RNA inputs.

2.3 Library preparation

Enriched extracts were then processed for sequencing as per the illumina ScriptSeq™ Complete Kit (Plant Leaf) (illumina BPL1224), according to SOP NGSL3b (ScriptSeq Library Prep). Extracts that had been enriched by poly(A) capture were processed from the “Prepare ScriptSeq library (Box 3)” Step.

Library preparation methodology was performed according to SOP NGSL3b (ScriptSeq Library Prep) (Annex B), with the following modification;

20 cycles of PCR were performed, instead of 15. This was to enable sufficient sequencing library to be prepared from low levels of input RNA.

2.4 Sequencing

Prepared libraries were pooled and sequenced on an illumina MiSeq, using MiSeq Reagent Kit v3 (illumina MS-102-3003), according to manufacturer's instructions.
2.5 Analysis

Two different analyses were performed to attempt to detect Norovirus sequences.

First, sequences were trimmed for quality and length (to remove sequence below Q20 and 100nt) using sickle v1.33, and assembled using trinity v2.1.1 to form contigs. These contigs were aligned against the NCBI nucleotide database using blast v2.2.29 to assign taxa of origin. This approach provides a broader examination of sequences present in the sample, and is useful when the target sequence is very different from the known reference.

A second approach to identify sequences of viral origin was to align untrimmed reads to a reference genome (Norovirus, NC_001959.2, and Mengovirus DQ294633.1, with a spurious 50bp region of tandemly repeated cytosine (C) nucleotides removed from the start of the sequence). Alignment was performed using bwa mem v0.7.10, and alignments were analysed using samtools 0.1.19. This approach is a more sensitive method for detecting sequences from known viruses of interest.

A paired t-test (Microsoft Excel 2010) was used to compare the total number of reads generated, and the number of reads assigned to Mengovirus, by each enrichment method.
3. Results

3.1 Sequencing

The RNA concentrations of the samples ranged from 12.8 ng/µl at the highest to below the limit of detection (eight samples) at the lowest, which is below the 2ng/µl minimum recommend for the Scriptseq complete protocol, even with the modifications for low input RNA.

Using an increase in cycle number, sequence was successfully obtained for each enrichment methodology on each sample. The number of sequence reads generated for each sample are shown in Table 2. The control sample consisted of RNA extracted from a health tobacco plant spiked with artificial mRNA. It acts as a positive control allowing the success of the sequencing of the artificial RNA to be assessed as well as a negative control as no virus should be detected in this sample.

Table 2. Number of individual sequence reads generated for each enrichment method for each sample. illumina technology generates paired read data, so the total number of read pairs for each sample is the total number of reads / 2.

<table>
<thead>
<tr>
<th>Enrichment method</th>
<th>Sample</th>
<th>Sequence reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>poly(A)-capture</td>
<td>143</td>
<td>1199920</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>145</td>
<td>1115124</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>155</td>
<td>1006296</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>180</td>
<td>995802</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>188</td>
<td>956490</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>196</td>
<td>1218258</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>214</td>
<td>835690</td>
</tr>
<tr>
<td>Enrichment method</td>
<td>Sample</td>
<td>Sequence reads</td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------</td>
<td>----------------</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>272</td>
<td>686446</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>409</td>
<td>1260962</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>415</td>
<td>613204</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>416</td>
<td>822908</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>429</td>
<td>1221984</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>Control</td>
<td>782886</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>143</td>
<td>1571760</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>145</td>
<td>1095900</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>155</td>
<td>1473212</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>180</td>
<td>2139938</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>188</td>
<td>3051470</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>196</td>
<td>2966918</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>214</td>
<td>5857586</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>272</td>
<td>3975074</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>409</td>
<td>1553518</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>415</td>
<td>3610274</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>416</td>
<td>2802394</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>429</td>
<td>1033246</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>Control</td>
<td>1051562</td>
</tr>
</tbody>
</table>
3.2 Analysis

Assembly and comparison to the nucleotide database

Sequence of plant origin was detected in every sample. In addition, a number of samples contained identifiable plant virus or Mengovirus sequence (see Table 3).

Table 3. The number of contigs identified to each plant virus or Mengovirus from each sample. Note that identification is based on closest hit by blast, which may be influenced by the completeness of the reference database.

<table>
<thead>
<tr>
<th>Enrichment Method</th>
<th>Sample</th>
<th>Virus</th>
<th>Number of contigs identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>poly(A)-capture</td>
<td>145</td>
<td>Tomato ringspot virus</td>
<td>3</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>145</td>
<td>Tomato ringspot virus</td>
<td>4</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>145</td>
<td>Black raspberry necrosis virus</td>
<td>4</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>145</td>
<td>Mengovirus</td>
<td>1</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>180</td>
<td>Mengovirus</td>
<td>1</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>188</td>
<td>Mirabilis mosaic virus</td>
<td>1</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>196</td>
<td>Mengovirus</td>
<td>1</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>196</td>
<td>Black raspberry necrosis virus</td>
<td>3</td>
</tr>
<tr>
<td>poly(A)-capture</td>
<td>214</td>
<td>Mengovirus</td>
<td>1</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>214</td>
<td>Mirabilis mosaic virus</td>
<td>1</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>272</td>
<td>Mirabilis mosaic virus</td>
<td>1</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>272</td>
<td>Mengovirus</td>
<td>1</td>
</tr>
</tbody>
</table>
### Mapping to reference viral genomes

No sequences mapping to the Norovirus genome were found. However eight of the samples which underwent ribosomal depletion and one of the samples which underwent poly(A)-capture contained sequences mapping to the Mengovirus genome. Samples containing detectable Mengovirus reads are shown in Table 4.

**Table 4. Numbers of reads which mapped to the Mengovirus reference genome from each sample. Samples not listed had no reads mapping to Mengovirus.**

<table>
<thead>
<tr>
<th>Enrichment Method</th>
<th>Sample</th>
<th>Number of reads mapping to Mengovirus</th>
</tr>
</thead>
<tbody>
<tr>
<td>poly(A)-capture</td>
<td>214</td>
<td>4</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>145</td>
<td>4</td>
</tr>
<tr>
<td>Enrichment Method</td>
<td>Sample</td>
<td>Number of reads mapping to Mengovirus</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>155</td>
<td>2</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>180</td>
<td>4</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>188</td>
<td>2</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>196</td>
<td>4</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>272</td>
<td>2</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>409</td>
<td>2</td>
</tr>
<tr>
<td>ribosomal depletion</td>
<td>416</td>
<td>2</td>
</tr>
</tbody>
</table>

**Comparison of numbers of reads**

The mean total number of reads generated in samples processed by poly(A)-capture is 994,424 (standard deviation = 218,694), and the mean total number of reads generated in samples processed by ribosomal depletion is 2,594,274 (standard deviation = 1,422,326). A paired two sample t-test shows a significant difference between the groups at the 95% confidence level (p=0.0049).

The mean number of reads mapping to the Mengovirus reference genome from the samples processed by poly(A)-capture is 0.33 (standard deviation = 1.15), and the mean number of reads mapping to the Mengovirus reference genome from the samples processed by ribosomal depletion is 1.83 (standard deviation = 1.59). A paired two sample t-test shows a significant difference between the groups at the 95% confidence level (p=0.04).
4. Conclusions and recommendations

Sequencing was successfully performed for all samples, and with both enrichment methodologies attempted. On average more sequence reads were produced from samples processed with ribosomal depletion than with poly(A)-capture. Of the reads produced, no Norovirus sequence was detected. However, plant sequences, plant virus sequences, and Mengovirus extraction control sequences were detected in many samples. Again on average more Mengovirus sequences were detected from samples treated with ribosomal depletion than with poly(A)-capture. We conclude that Norovirus was not present on the fresh produce at high enough levels to be detectable using the illumina MiSeq platform.

Recommendations for improvements or alternative approaches are:

- Utilise a higher throughput sequencer (e.g. illumina HiSeq), to attempt to make use of the greater depth of sequence to detect viruses at extremely low titre. However, given the low volumes of RNA and low amount of starting Norovirus RNA, it may be the case that any Norovirus RNA is being lost during sample prep.

- Attempt to enrich for Norovirus particles in sample prior to extraction, e.g. using porcine mucin (Tian et al., 2008).

- Concentrate on samples likely to possess higher yields of Norovirus than fresh produce, e.g. shellfish, sewage, stool. Then inferences about source of infection could be drawn for at least some foodstuffs.
Annex A: References

Annex B: SOP NGSL3b ScriptSeq Library Preparation

Title: ScriptSeq library preparation (25 samples) NGS3LB Version 3
Owner: Aimee Fowkes
Approver: Ian Adams
ISSUE DATE: 25/01/2017

Introduction
This SOP describes the production of MiSeq ready sequencing libraries from RNA extracted from plants. It is used for the assessment of viral infection in plants.

Scope
This SOP uses the EppiBio ScripSeq Complete library protocol to:

a. Ribosome deplete the RNA with ribozero reagents
b. Clean up RNA with RNA clean
c. Make first strand cDNA
d. Terminally tag with linkers
e. Clean with Ampure XP beads
f. PCR amplify and add indexes
g. Clean up with Ampure XP

Safety
Please reference to laboratory COSHH and safety procedures

Materials

- Agilent Tapestation and D1000 HS kit
- Thermal cycler
- Pipettes and filter tips (separate sets for pre and post PCR stages)
• Magnetic separation stand
• Microfuge
• Qubit double stranded DNA fluorimeter and reagents
• Ampure XP magnetic beads
• RNAclean XP magnetic cleanup beads
• ScriptSeq Complete (plant leaf or root) library preparation kit (Illumina)
• ScriptSeq index kit (Illumina)
• Failsafe DNA polymerase (illumina)
• PCR plates, Eppendorf tubes
• 1ug DNA free RNA from infected plant
• EACC spike in control (LifeTech)

Procedure
The ScriptSeq Complete Kit (Plant Leaf) is composed of three boxes:
• Box 1 -80oC Ribo-Zero™ rRNA Removal Reagents (Plant Leaf)
• Box 2 4oC Magnetic Core Kit
• Box 3 -20oC ScriptSeq™ v2 RNA-Seq Library Preparation Kit
The reagents are colour coded and these colours have been used in the SOP

Controls / Samples
For the purposes of virus detection this method has been validated using 24 samples and a control of RNA extracted from a healthy White burley tobacco plant (negative control / extraction control) spiked with 1 μl EACC spike control mix (positive control)

Ribosome deplete the RNA with ribozero reagents (Box 1 and 2)

Location 10G04 mastermix cabinet
Allow the Magnetic Core Kit components to equilibrate to room temperature for 30 minutes before proceeding
1. Prepare Magnetic Beads (Box 2). A 24 reaction kit contains enough reagents to do 25 reactions.

2. Mix the Magnetic Beads tube (5.4 ml or 225 μl/sample) well by pipetting or gentle vortexing.

3. Place the magnetic beads tube (5.4mL) in a magnetic stand and wait for 2 minutes or until complete separation.

4. Remove and discard the supernatant carefully without disturbing the pellet.

5. Add 5.4 ml of RNase-Free Water and mix by repeated pipetting or by vortexing at medium speed.

6. Place the tube in the magnetic stand for at least 1 minute (until the solution appears clear), repeat wash steps (IV-VI).

7. Resuspend in 1500 μl (60 μl/Sample) of Magnetic bead Resuspension solution and mix well by repeated pipetting or by vortexing at medium speed.

8. Add 25 μl of RiboGuard RNase inhibitor to the 1500μl resuspended Magnetic beads and mix briefly by vortexing.

9. Aliquot 65 μl of the washed magnetic beads/RiboGuard RNase inhibitor into a 96-well plate.

**Location 10G04 mastermix and spiking cabinet**

Remove the Ribo-Zero rRNA Removal Reagents from --70°C to --80°C storage, thaw the tubes, and place them on ice.

The Ribo-Zero process uses 1-5 μg of DNA-free RNA (at least 500 ng). The sample should be dissolved in RNase-Free Water or TE Buffer

1. Treat sample with rRNA Removal Solution (box 1)

Mix in 40 μl total volume:

1. In the mastermix cabinet, add to one well per sample on a 96-well plate:
   - 8 μl rRNA Removal Solution
   - 4 μl Reaction Buffer
   - Transfer the plate to the spiking cabinet and add:
1. 1-5 μg total RNA (28 μl, adjust the volume with water)
   - Incubate 10 min @ 68°C, then 5 minutes @ RT (PLH/922 thermocycler, Ummey's Folder, 'Ribo' programme)

2. Location 10G04 spiking cabinet

3. Using a pipette, add the treated RNA from step 1 to the 96-well plate containing the washed Magnetic Beads and, without changing the pipette tip, immediately and thoroughly mix the contents of the tube by pipetting at least 10 times and leave at room temperature. Repeat this process for each sample.

   Important! Always add the treated RNA sample to the washed Magnetic Beads and immediately mix by pipetting. Never add the Magnetic Beads to the treated RNA sample. Immediate and thorough mixing prevents the beads from forming clumps that can significantly impact the efficiency of the rRNA removal

1.
   - Incubate 5 minutes @ RT
   - Incubate 5 minutes @ 50°C (PLH/922 thermocycler, Ummey's Folder, 'RRNA' programme), remove the plate and immediately place it on a magnetic stand for at least 1 minute (until the solution appears clear)
   - Transfer the supernatant (85-90 μl rRNA-depleted sample) to a labelled RNase-free plate
   - Immediately proceed to next step or, alternatively, the supernatant may be stored at -70°C to -80°C.

2.
   1. Purify rRNA-depleted sample

      Bring the RNAClean XP bead solution to room temperature for 30 minutes and vortex the solution very well before use.

      1.
         - Add 160 μl RNAClean XP solution to each well containing 85-90 μl of rRNA depleted RNA and mix by pipetting up and down 10 times
         - Stand for 15 min and then place on magnetic stand for approximately 2 min
- Remove supernatant
- Wash with 2x 200 μl fresh 80% ethanol
- Air dry beads for 15 min
- Remove plate from the stand and add 12 μl water, mix by pipetting up and down 10 times
- Spin down plate on plate centrifuge
- Stand for 2 min and then place on magnetic stand
- Transfer 9 μl supernatant to a fresh plate for next step. Sample can be stored at -80°C at this stage

**Prepare ScriptSeq library (Box 3)**

1. XII. Fragment RNA
   - Mix the following per sample (multiply by number of samples and make a mastermix):
     1 μl RNA Fragmentation Solution
     2 μl cDNA Synthesis Primer
     Add to the 9 μl Ribo-Zero treated RNA in its' 96 well plate
     - Incubate at 85°C for 2 minutes in thermocycler then hold at 4°C.

1.

1. Synthesize cDNA
   - Mix the following master mix (per sample):
     3.0 μl cDNA Synthesis PreMix
     0.5 μl 100 mM DTT
     0.5 μl StarScript Reverse Transcriptase (leave in freezer until required)
     - Add 4 μl of Master Mix to each reaction. Mix by pipetting.
     - Incubate at 25°C for 5 min followed by 42°C for 20 min. Cool reaction to 37°C hold (PLH/922 thermocycler, Ummey's Folder, 'Script1' programme)
     - Remove plate from thermocycler to the spiking cabinet and quickly add 1.0 μl of Finishing Solution to each reaction. Mix by pipetting and replace on thermocycler.
• Incubate at 37°C for 10 min.
• Incubate at 95°C for 3 min, cool to 25°C and Pause the thermocycler.

Bring the Ampure XP beads to room temperature

1. **Terminal-Tag cDNA**
   - Mix the following master mix (per sample):
     - 7.5 μl Terminal Tagging Premix (warm to room temp to allow easier pipetting a sit is viscous)
     - 0.5 μl DNA Polymerase
     - Solution is viscous! Mix thoroughly by pipetting.
     - Add 8.0 μl of Master Mix to each reaction. Mix by pipetting.
     - Incubate reaction at 25°C for 15 minutes, followed by 95°C for 3 minutes. Cool to 4°C. (PLH/922 thermocycler, Ummey's Folder, 'Script2' programme)

1. **Purify di-tagged sample**

2. **PCR**
   - Mix in a 0.2-ml PCR plate (50 μl total volume):
22.5 μl di-tagged cDNA

25 μl FailSafe PCR PreMix E

1 μl Forward PCR Primer

1 μl index PCR Primer (see note)

0.5 μl FailSafe PCR Enzyme

PCR cycle conditions: 95°C for 1 minute

15 cycles of: 95°C for 30 seconds

55°C for 30 seconds

68°C for 3 minutes

Then 68°C for 7 minutes and hold at 12°C. Sample can be stored at -80°C at this stage

**Index primers (sets 1-4):** If doing standard 25 samples use two sets of ScriptSeq index primers and one primer from a third set for the control. Avoid using the sets that were used on the previous run. If doing 12 or more use a complete set plus part of a second set. If doing less than 12 please refer to the ScriptSeq index manual for details on maintaining correct diversity.

**Location 10G01**

**QC Library**

- Quantify libraries by Qubit and visualize a selection on Agilent Tapestation (see separate SOP: NGS1 Library QC & acceptance).

**Critical decision point**

1. Check concentration of libraries: Concentrations above 5 ug/ml for most of the libraries are acceptable

2. Check size of libraries: Libraries with a visible tapestation peak between 200-800bp are acceptable.

**Discuss with at least 1 other member of the sequencing team about the quality of the samples and record and decisions to stop / repeat / proceed**

Pool equal quantities of libraries based on the qubit values.

1.
1. Purify pooled library sample

Bring the AmpureXP solution to room temperature for 30 minutes and vortex very well prior to use

1. Take 200 μl of the pooled library in a 1.5 mL tube. Add 120 μl AmpureXP solution to the library pool, mix by pipetting up and down 10 times
2. Stand for 15 min and then place on magnetic stand
3. Remove supernatant
4. Wash with 2x 200 μl fresh 80% ethanol
5. Air dry beads for 15min
6. Add 50 μl water, mix by pipetting up and down 10 times
7. Stand for 2 min and then place on magnetic stand
8. Transfer 45 μl supernatant to fresh tube
9. Quantify library pool by Qubit and visualize on Agilent Tapestation (NGS1 Library QC & acceptance).

Documents/Records referred to


http://support.illumina.com/sequencing/sequencing_kits/scriptseq-v2-rna-seq/documentation.html

NGS1 Library QC & acceptance
Fera hereby excludes all liability for any claim, loss, demands or damages of any kind whatsoever (whether such claims, loss, demands or damages were foreseeable, known or otherwise) arising out of or in connection with the preparation of any technical or scientific report, including without limitation, indirect or consequential loss or damage; loss of actual or anticipated profits (including loss of profits on contracts); loss of revenue; loss of business; loss of opportunity; loss of anticipated savings; loss of goodwill; loss of reputation; loss of damage to or corruption of data; loss of use of money or otherwise, and whether or not advised of the possibility of such claim, loss demand or damages and whether arising in tort (including negligence), contract or otherwise. This statement does not affect your statutory rights.

Nothing in this disclaimer excludes or limits Fera liability for: (a) death or personal injury caused by Fera negligence (or that of its employees, agents or directors); or (b) the tort of deceit; [or (c) any breach of the obligations implied by Sale of Goods Act 1979 or Supply of Goods and Services Act 1982 (including those relating to the title, fitness for purpose and satisfactory quality of goods);] or (d) any liability which may not be limited or excluded by law (e) fraud or fraudulent misrepresentation.

The parties agree that any matters are governed by English law and irrevocably submit to the non-exclusive jurisdiction of the English courts.