E181945 mRNA-Seq Summary Report

1. Project Information

Project: E181945
Investigator: Gavuthami Murugesan
Report by: Louise MacGillivray
Report Date: 09.05.2019

Number of Samples: 8

2. Summary of Sequencing Protocol

Quality Control

Total RNA samples were quantified using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific Inc, #Q32866) and the Qubit RNA BR assay kit (#Q10210).

Library Preparation

Libraries were prepared from 500ng of each total-RNA sample using the NEBNEXT Ultra II Directional RNA Library Prep kit with Poly-A mRNA magnetic isolation (NEB #E7490) according to the provided protocol.

Poly-A containing mRNA molecules were purified using poly-T oligo attached magnetic beads. Following purification the mRNA was fragmented using divalent cations under elevated temperature and primed with random hexamers. Primed RNA fragments were reverse transcribed into first strand cDNA using reverse transcriptase and random primers. RNA templates were removed and a replacement strand synthesised incorporating dUTP in place of dTTP to generate ds cDNA. The incorporation of dUTP in second strand synthesis quenches the second strand during amplification as the polymerase used in the assay is not incorporated past this nucleotide. AMPure XP beads (Beckman Coulter, #A63881) were then used to separate the ds cDNA from the second strand reaction mix, providing blunt-ended cDNA. A single 'A' nucleotide was added to the 3' ends of the blunt fragments to prevent them from ligating to another during the subsequent adapter ligation reaction, and a corresponding single 'T' nucleotide on the 3' end of the adapter provided a complementary overhang for ligating the adapter to the fragment. Multiple indexing adapters were then ligated to the ends of the ds cDNA to prepare them for hybridisation onto a flow cell, before 11 cycles of PCR were used to selectively enrich those DNA fragments that had adapter molecules on both ends and amplify the amount of DNA in the library suitable for sequencing. After amplification libraries were purified using AMPure XP beads.

Library QC

Libraries were quantified by fluorometry using the Qubit dsDNA HS assay and assessed for quality and fragment size using the Agilent Bioanalyser with the DNA HS Kit (#5067-4626).

Sequencing

Sequencing was performed using the NextSeq 500/550 High-Output v2 (150 cycle) Kit (# FC-404-2002) using a High Out v2.5 Flow Cell on the NextSeq 550 platform (Illumina Inc,

#SY-415-1002). 8 libraries were combined in an equimolar pool based on the library quantification results and run across one High-Output Flow Cell.

3. Summary of Analysis Pipeline

Basecall data produced by the NextSeq 550 is automatically uploaded to BaseSpace, a cloud-based data management and analysis service provided by Illumina. Here it is converted into FASTQ files in order to allow analysis using a number of apps accessible directly through BaseSpace, or to download so that alternative analysis pipelines can be used.

4. Summary

A 2x75bp sequencing run on the Nextseq 550 using a high output flow cell is expected to generate up to 400M reads (50-60Gb) with a data quality of >80% higher than Q30, based on a cluster density of 170-230K/mm². When multiplexing 8 samples per flow cell we would therefore expect to see up to 50M reads per sample.

Table 1: Sequencing statistics

Run Name	Flow Cell ID	Sequence Date	Cluster Density (K/mm²)	% Clusters PF	Yield (Gb)	% Data >Q30
E181945_TM_RNAseq_APR2019_060519	HG552BGAB	06.05.2019	216	89.2	77.6	93.2

Table 2: Clusters PF per sample

Sample ID	Sequence Date	Clusters Passing Quality Filters
Treg_Rest_1	06.05.2019	61,533,633
Treg_Rest_2	06.05.2019	52,210,160
Treg_Rest_3	06.05.2019	62,035,728
Treg_Rest_4	06.05.2019	63,673,054
Treg_Act_1	06.05.2019	59,316,804
Treg_Act_2	06.05.2019	61,478,690
Treg_Act_3	06.05.2019	64,048,408
Treg_Act_4	06.05.2019	59,741,647

Coverage of each sample was even across the flow cells, and all samples generated well in excess of 50M reads (Min:52M, Max:64M, Mean:60M).

Louise MacGillivray
9th May, 2019