BioSpyder

Service Project Report and Deliverables
SP0438 – US EPA CCTE – Purified RNA in Mouse S1500+ v1.2

Overview and Purpose

Thank you for contracting with BioSpyder to run your samples using the TempO-Seq assay.

This report describes the results of this study (both QC and samples). Gene count files output from the assays performed in this study will be transferred by AWS S3 bucket. The work described in this report is a study to use targeted sequencing-based RNA expression analysis.

If you have any questions about this report or the results, please contact services@biospyder.com.

Sample Names and Plate Layout

Samples were named according to the plate name provided by US EPA CCTE and well position.

Example: samples for plate C9H21X16 is in the SP0438_gene_counts.csv file. Each column represents one well, named 83_d6_M_B through 166_d6_F_L.

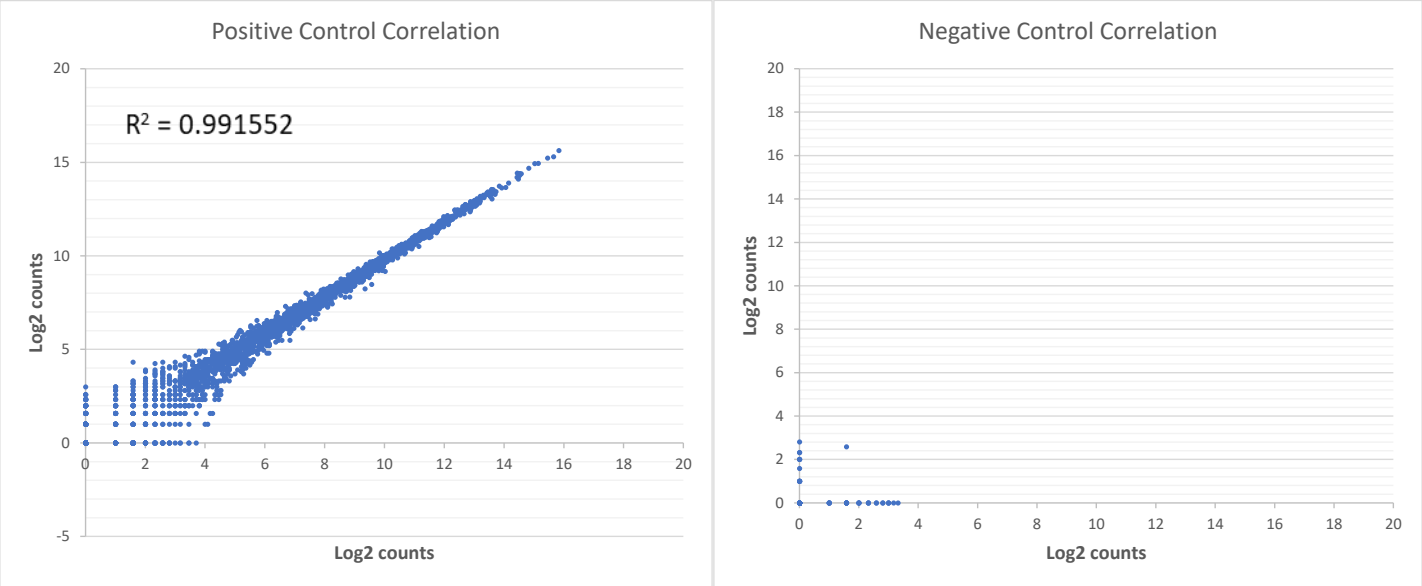
QA/QC test results

QA/QC of BioSpyder data is comprised of the following steps, for which results are shown below. We also verify that each sequencing run has met or exceeded Illumina sequencing metrics.

1. Read Count Analysis

Metric	Value	Status	Significance
Average number of mapped reads in positive controls	2,253,278	Pass (>1.07 M mapped reads)	Sufficient number of mapped reads are required
The signal to noise ratio	22421:1	Pass (>20:1)	Total number of mapped reads in the positive control divided by the total number of mapped reads in negative controls
The percentage of mapped reads in positive control	93.83%	Pass (>80%)	

2. Replicate Analysis of Positive Controls (Reference RNA) and No-Sample Controls



Positive controls show a high degree of reproducibility. Negative controls demonstrate a negligible number of reads. Shown above are selected graphs of reproducibility of positive and negative control samples within plates.

3. Exceptions

No exceptions noted.

Additional Considerations

4. Attenuated Genes

A standard attenuation panel was used for this experiment, as agreed upon. Data is delivered raw, without correction for these attenuation factors. Back-calculated data can be provided by BioSpyder upon request, but are not needed for differential expression analysis.

4.1. Normalized Counts per Gene per Sample

No normalization was applied to the data files provided

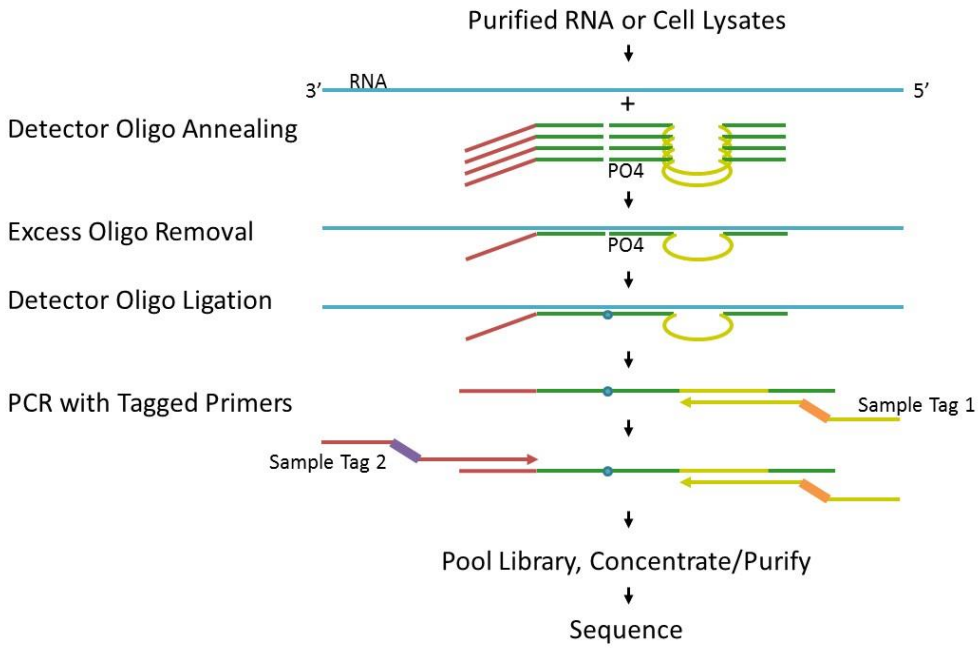
Sequencing Output: Raw data in FASTQ file format.

Arrangements have been made to transfer the raw sequencing data files generated on NovaSeqX in FASTQ format via AWS S3 bucket.

5. Protocols

5.1. TempO-Seq Assay Protocol Summary

Sequencing libraries for targeted panels were generated as described briefly and depicted in the figure below.



In TempO-Seq, each Detector Oligo (DO) consists of a sequence complementary to an mRNA target plus a universal (i.e. same for every targeted gene) primer binding site. They anneal in immediate juxtaposition to each other on the targeted RNA template such that they can be ligated together. Ligated detector oligos are PCR-amplified using a primer set (single-plex PCR reaction, with a single primer pair for each sample) that introduces both the adaptors required for sequencing and a sample-specific barcode. The barcode sequences flank the target sequence and are inserted appropriately into the standard Illumina adaptors to permit standard dual-index sequencing of the barcodes and deconvolution of sample-specific reads from the sequencing data using the standard Illumina software. All the PCR-amplified and barcoded samples are pooled into a single library for sequencing. Sequencing reads are demultiplexed using the standard sequencing instrument software for each sample using the barcodes to give a FASTQ file for each.

6. Data Analysis Protocol

TempO-Seq data are analyzed using the Tempo-SeqR software package using the statistical computing language R. The input for TempO-Seq data analysis is a folder of zipped FASTQ files. Each FASTQ file contains the reads and quality scores for one sample. Each FASTQ file is aligned using the Bowtie algorithm to a pseudo-transcriptome input by the user. The output for TempO-Seq data analysis is a table of counts with each column representing a sample and each row representing a gene.

6.1. Data Normalization

Data normalization is commonly not necessary. A variety of normalization methods may be used at your discretion, including normalizing to total reads per sample.

6.2. Probe Manifest

ProbeID is composed of gene name along with a numeric identifier, separated by an underscore. For additional information about probe location, including target sequence and RefSeq and ENSEMBL IDs, contact BioSpyder at support@biospyder.com.

7. Return of Remaining Materials

If you would like the remaining lysates or libraries returned at your cost, please make that request to support@biospyder.com (unless arrangements have already been made). You will need to supply a FedEx account number to charge against. We cannot guarantee that any excess materials remain or warrant their quality, but will attempt to satisfy requests.

If desired for use with other experimental platforms, RNA may be isolated from BioSpyder Lysates using a Trizol-based method. DNA can be extracted using most standard kits, or through phenol/chloroform extraction. Request protocols by email to support@biospyder.com.