

A Globin Blocker to Increase Sequencing Efficiency for QuantSeq 3' mRNA-Seq in Porcine Blood

A.S. Leaflet R3275

Kyu-Sang Lim, Post-doctoral Research Associate, ISU;
Qian Dong, Graduate Student, ISU;
Pamela Moll, Lexogen GmbH, Austria;
Jana Vitkowska, Lexogen GmbH, Austria;
Gregor Wiktorin, Lexogen GmbH, Austria;
Stephanie Bannister, Lexogen GmbH, Austria;
Dalia Daujotyte, Lexogen GmbH, Austria;
Christopher K. Tuggle, Professor, ISU;
Jack C. M. Dekkers, Distinguished Professor, ISU;

Summary and Implications

Blood is an extremely valuable source to investigate animal biology, including whole genome gene expression. The efficiency of gene expression profiling is hampered by the high proportions of hemoglobin (HB) mRNA of erythrocytes, which decreases the resolution of QuantSeq 3' mRNA-sequencing in blood. Here, we used different concentrations of a globin blocker (GB), which includes a HB-specific oligonucleotide mix, as a novel method to decrease sequencing of HB RNA. While reads for HBA and HBB accounted for 19.4 and 36.9% of total reads in non-GB samples, GB under optimal concentrations reduced HBA and HBB reads on average to 8.7 and 2.3% of total reads, respectively. The number of genes that could be reliably detected was of approximately 2,200 genes greater in GB samples than in non-GB samples. In conclusion, the use of this GB procedure increases the resolution of gene expression data from QuantSeq in porcine blood samples.

Introduction

The whole blood transcriptome is a potential source of biomarkers for many biological processes. RNA sequencing is a powerful method to query the blood transcriptome but is expensive. In this regard, QuantSeq 3'RNA-seq, which aims to generate one 3' end fragment per transcript, reduces sequencing costs and data analysis time. However, the hemoglobin mRNAs, such as HBA and HBB, occupy most of the blood transcriptome and are a limiting factor for gene expression analysis in blood. Several protocols for globin depletion of blood mRNA samples have been proposed. These protocols must be used prior to the RNAseq library preparation protocols. Our objectives were to evaluate a novel globin blocker procedure for pigs that can be seamlessly integrated into the QuantSeq workflow without adding another step and to evaluate the effect of the GB in porcine blood RNA sequence data.

Materials and Methods

A blood sample was collected in a Tempus blood RNA tube from a Large-White female pig at weaning (27 day of age) and stored at -80 °C. RNA was extracted using the Norgen Preserved Blood RNA Purification Kit I. The RNA integrity number (RIN) of extracted RNA was 6.5, which was assessed by an Agilent 2100 Bioanalyzer. An RNA Library was generated using the QuantSeq 3' mRNA-Seq Library Prep Kit FWD, without (Non-GB)/with 4 different concentrations of the GB solution using a 10-fold dilution series (GB-c1, GB-c2, GB-c3 and GB-c4) and sequenced by the Illumina NextSeq 500 Sequencing System. Raw sequencing reads were trimmed by Bbduk to remove adapter/poly-A sequences and low quality bases. The processed reads were aligned to pig reference genome Sscrofa 10.2 using STAR. Read counts per gene were calculated by Htseq with the pig genome GTF file from the 10.2 pig genome assembly.

Results and Discussion

On average, 15.91 ± 1.83 million clean reads were obtained after trimming the raw QuantSeq 3' mRNA-seq data, of which 88.01 ± 0.02 % of reads were mapped to the pig reference genome. Mapping rates in GB libraries were slightly lower than in non-GB libraries but more than 85% of reads were aligned in the GB libraries. Reads counts per gene among two technical replicates of each GB concentration treatment were highly reproducible ($r = 0.98$ in GB-c1 and $r = 0.99$ in other treatments). The proportion of HBA reads over total read counts was much lower in the GB libraries (7.6 to 16.2%) compared with non-GB libraries (19.4%). HBB reads in non-GB libraries accounted for 36.9% of all reads but this was dramatically reduced in the GB libraries (to 2.5 to 13.6%). In addition, the numbers of genes with counts per million (CPM) greater than 1, which provides a reliable estimate of expression of the gene, were greater in the GB libraries (7,234 to 8,010) compared with the non-GB libraries (5,782).

In conclusion, the GB assay was effective in blocking globin RNA reads in QuantSeq 3' mRNA-seq and increases the sensitivity of QuantSeq data from blood. In addition, because the GB is included in an existing step of the QuantSeq library prep workflow, it has limited impact on work time and costs.

Acknowledgments

This work was funded by Lexogen GmbH and USDA-NIFA grant number 2017-67007-26144. Dr. Claire Rogel-Gaillard is acknowledged for providing the RNA samples.