Suggested Reference


Copyright

Items in ResearchSpace are protected by copyright, with all rights reserved, unless otherwise indicated. Previously published items are made available in accordance with the copyright policy of the publisher.

For more information, see General copyright.

Copyright

Items in ResearchSpace are protected by copyright, with all rights reserved, unless otherwise indicated. Previously published items are made available in accordance with the copyright policy of the publisher.

For more information, see General copyright.
Detection of tissue- and sex-specific gene expression in *Bos taurus* using high depth RNA sequencing

**Thomas Lopdell, Matt Littlejohn**
Livestock Improvement Corporation, Hamilton, New Zealand

**Introduction**
As part of a larger sequencing project, LIC is using RNA-Seq data to aid in the identification of genes and variants underlying major dairy traits. This includes defining transcript structures (qualitative assessment), as well as looking for differentially-expressed genes between groups of samples (quantitative analysis, e.g. eQTL mapping). Two tissues have been examined so far, consisting of 29 lactating mammary samples, and eight pituitary samples (including three males and five females; all less than six weeks old). This poster summarises the two data sets, and examines a potential method for determining sex from RNA-Seq data.

**Methods**
Illumina Truseq libraries were generated from mammary and pituitary samples, and were sequenced using the Illumina HiSeq instrument. Since casein mRNA was anticipated to comprise a large proportion of the transcripts sequenced from mammary tissue, a high-depth approach was used, allocating two samples per lane. Pituitary samples were sequenced four per lane. Reads were mapped to the UMD3.1 bovine genome using TopHat2. Transcripts and exons were assembled from the mapped reads using Cufflinks, then merged with the reference gene set using Cuffmerge. Finally, the exon boundaries output by Cuffmerge were fed back into Tophat2 in order to improve the alignments around the splice junctions. DESeq was used to calculate a set of genes which are differentially-expressed by sex.

**Mammary Sample Results**
The depth of reads mapping to each chromosome is shown below. Chromosome 6 has much higher read depth, due to the extreme expression of the casein cluster of genes. Approximately 50% of all mapped reads were derived from the four genes at this locus, far exceeding the 45S rRNA locus (chromosome 25) for relative transcript abundance. The table to the right shows the top five most highly expressed genes in lactating mammary tissue. All of these genes encode proteins which are secreted in milk.

**Pituitary Sample Results**
The reads from the pituitary extract samples mapped more evenly across the genome, despite also representing a highly specialised secretory tissue. The table to the right lists the most highly expressed genes in pituitary tissue. GH1 and PRL encode peptide hormones which are secreted by the pituitary gland. POMC is a precursor to several peptide hormones, including ACTH and MSH which are also produced in the pituitary gland. CGA encodes the alpha subunit of several glycoprotein hormones (gonadotrophins), such as LH and FSH. The protein encoded by the GNAS gene is the stimulatory G-protein alpha subunit, which associates with members of the large and functionally-diverse family of G-protein coupled receptors.

**Differentially-Expressed Genes**
In pituitary tissue, DESeq identified 136 genes with differential expression by sex, after correcting the p-values using the false discovery rate. These genes are highlighted in blue in the MA plot below. Genes above the red line have higher expression in females, while genes below the line have higher expression in males. The five most significant genes, labelled on the MA plot, all map to the X and Y bovine sex chromosomes. The gene locations and p-values of differential expression are indicated in the table below. Four of the genes (EIF2S3Y, UTY, USP9Y and ZRSR2Y) are Y-linked, with significantly higher expression in males. XIST, however, maps to the X chromosome and is involved in X inactivation, having significantly higher expression in females.

**Conclusions**
- The diversity of transcripts representing individual cell and tissue types can vary massively, with highly specialised tissues such as lactating mammary producing large numbers of transcripts derived from a small number of genes. Unlike microarray methods, heavy skewing in the relative abundance of transcripts may affect the sensitivity to detect less abundant transcripts, requiring increased sequencing depth, or the use of transcript depletion technologies prior to sequencing.
- It appears to be possible to determine sex from RNA-Seq data. However, further validation with more male animals is required.

**References**