Transcriptome analysis of crossbred and purebred cattle using CAGE and mRNA-Seq for the BovReg Project

Mapping of transcription start sites (TSS) is a key first step in understanding transcript regulation and diversity and how these might influence phenotypic plasticity in cattle. Including data from multiple tissues and non-reference breeds or crosses will increase our understanding of TSS usage, transcription complexity and will enhance the annotation of the reference genome for cattle (ARS-UCD1.2). For the BovReg project, we performed Cap Analysis Gene Expression (CAGE) sequencing and mRNA-Sequencing of 105 tissue samples (from 24 different tissues) from dairy (Holstein, n=43), composite beef (KC-composite, n=31) and beef/dairy cross (Charolais x Holstein, n=31) cattle. A *de novo* assembled transcriptome of the tissue samples was produced using the mRNA-Seq data and isoform expression estimates were calculated using Kallisto (TPM). Network cluster analysis was performed in Graphia and large clusters of genes with tissue-specific expression patterns were identified. CAGE libraries were mapped to ARS-UCD1.2_Btau5.0.1Y. The mapped reads were analysed using the CAGEfightR package to generate uni-directional (TSS) and bi-directional (TSS-Enhancer) clusters. Cross population analysis revealed that the greatest number of population-specific TSS were detected in the KC-composite (3,102), relative to the Holstein (1,152) and the Charolais x Holstein (1,092). A similar pattern was observed in the TSS-Enhancer clusters (419 KC-composite, 286 Charolais x Holstein and 202 Holstein) indicating that the total number of TSS observed was greater in composite relative to purebred cattle. The gene annotation datasets we have generated from crossbred and purebred cattle, with BovReg partners, create a new high-resolution map that captures additional transcriptional complexity in the bovine genome.

==========

Mazdak Salavati[1], Richard Clark[2], Doreen Becker[3], Christa Kühn[3,4], Graham Plastow[5], Gabriel Costa Monteiro Moreira[6], Carole Charlier[6,7] and Emily L. Clark[1] on behalf of the BovReg consortium

[1]The Roslin Institute, University of Edinburgh, Edinburgh, UK
[2]Genetics Core, Edinburgh Clinical Research Facility, The University of Edinburgh, Edinburgh, UK
[3]Institute of Genome Biology, Research Institute for Farm Animal Biology (FBN), Dummerstorf, Germany
[4]Faculty of Agricultural and Environmental Sciences, University Rostock, Rostock, Germany
[5]Livestock Gentec, Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, Canada
[6]Unit of Animal Genomics, GIGA Institute, University of Liège, Liège, Belgium.
[7]Faculty of Veterinary Medicine, University of Liège, Liège, Belgium.

==========