

# Long-non coding RNAs repertoires in liver and two T lymphocyte cell types in four livestock species

Kévin Muret, Sarah Djebali Quelen, Thomas Derrien, Cédric Cabau, Christophe C. Klopp, Diane Esquerre, Kylie Munyard, Gwenola Tosser-Klopp, Hervé Acloque, Elisabetta Giuffra, et al.

▶ **To cite this version:**

Kévin Muret, Sarah Djebali Quelen, Thomas Derrien, Cédric Cabau, Christophe C. Klopp, et al.. Long-non coding RNAs repertoires in liver and two T lymphocyte cell types in four livestock species. 36. Conference of the International Society for Animal Genetics (ISAG), Jul 2017, Dublin, Ireland. 2017, ISAG 2017 - Genomes to Phenomes - Abstract Book. hal-01621915

**HAL Id: hal-01621915**
**https://hal.science/hal-01621915**

Submitted on 2 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

tium[2], M. H. Pinard-Van der Laan[2], S. Lagarrigue[3], and E. Giuffra[2], [1]*GenPhySE, INPT, ENVT, INRA, Université de Toulouse, Toulouse, France;* [2]*GABI, AgroParisTech, INRA, Université Paris Saclay, Paris, France;* [3]*PEGASE, INRA, Agrocampus Ouest, Rennes, France.*

Functional annotation of livestock genomes is a critical and obvious next step to derive maximum benefit for agriculture, animal science, animal welfare and human health. The aim of the Fr-AgENCODE project is to generate multi-species functional genome annotations by applying high-throughput molecular assays on three target tissues/cells relevant to the study of immune and metabolic traits. An extensive collection of stored samples from other tissues is available for further use (FAANG Biosamples 'FR-AGENCODE'). From each of two males and two females per species (pig, cattle, goat, chicken), strand-oriented RNA-seq and chromatin accessibility ATAC-seq assays were performed on liver tissue and on two T-cell types (CD3+CD4+ & CD3+CD8+) sorted from blood (mammals) or spleen (chicken). Chromosome Conformation Capture (*in situ* Hi-C) was also carried out on liver. Sequencing reads from the 3 assays were processed using standard processing pipelines. While most (50–70%) RNA-seq reads mapped to annotated exons, thousands of novel transcripts and genes were found, including extensions of annotated protein-coding genes and new lncRNAs (see abstract #69857). Consistency of ATAC-seq results was confirmed by the significant proportion of called peaks in promoter regions (36–66%) and by the specific accumulation pattern of peaks around gene starts (TSS) *v.* gene ends (TTS). Principal Component Analyses for RNA-seq (based on quantified gene expression) and ATAC-seq (based on quantified chromatin accessibility) highlighted clusters characterised by cell type and sex in all species. From Hi-C data, we generated 40kb-resolution interaction maps, profiled a genome-wide Directionality Index and identified from 4,100 (chicken) to 12,100 (pig) topologically-associating domains (TADs). Correlations were reported between RNA-seq and ATAC-seq results (see abstract #71581). In summary, we present here an overview of the first multi-species and -tissue annotations of chromatin accessibility and genome architecture related to gene expression for farm animals.

**Key Words:** multispecies, Functional Annotation of Animal Genomes (FAANG), ATAC-seq, RNA-seq, Hi-C

**20 Long-non coding RNAs repertoires in liver and two T lymphocyte cell types in four livestock species [FAANG pilot project "FR-AgENCODE"].** K. Muret*[1], S. Djebali[2], T. Derrien[3], C. Cabau[2], C. Klopp[4], D. Esquerré[2,5], K. Munyard[6], G. Tosser-Klopp[2], H. Acloque[2], E. Giuffra[7], S. Foissac[2], and S. Lagarrigue[1], [1]*UMR PEGASE INRA, Agrocampus Ouest UMR PEGASE, Rennes, France;* [2]*UMR GenPhySE, INRA, INPT, ENVT, Université de Toulouse, Castanet-Tolosan, France;* [3]*IGDR, CNRS-University Rennes 1, Rennes, France;* [4]*SIGENAE, INRA, Castanet-Tolosan, France;* [5]*Plateforme GENOTOUL, INRA, Castanet-Tolosan, France;* [6]*School of Biomedical Sciences, Curtin Health Innovation Research Institute, Western Australian Biomedical Research Institute, Curtin University, Perth, Western Australia, Australia;* [7]*UMR GABI, INRA, AgroParisTech, Université Paris Saclay, Jouy-en-Josas, France.*

Understanding genome-to-phenome relationships requires deep and cross-disciplinary genetic analyses among which functional annotation provides crucial insights. The development of High Throughput Sequencing and RNA-seq now help us to find a large number of heterogeneous and low-expressed transcripts known to be long non-coding RNAs (lncRNAs). One of the aims of the FAANG pilot project 'FR-AgENCODE' is to identify and characterise the long non-coding RNAs of multiple tissues and cell lines in 4 farm animals (chicken, bovine, pig and goat) of both sexes. Here, we focus our analysis on the liver tissue and two blood T-cell types (CD3+CD4+, CD3+CD8+) where samples were collected through 4 biological replicates (2 males and 2 females). It allows us to compare lncRNA repertoires between tissues, sex and species in relation with fundamental biological functions like energy storage and immunity. High depth strand-specific RNA-seq produced ~200M paired-end reads for each of the 16 RNA-seq datasets. After transcriptome reconstruction, we used the recently published FEELnc program (Wucher *et al*., 2017, Nucleic Acid Research) to identify lncRNAs longer than 200 bp and without protein-coding capabilities. FEELnc also classifies lncRNAs based on their genomic localizations with respect to the ENSEMBL protein-coding annotation: intergenic lncRNAs are categorized depending on the distance and orientation with respect to the closest mRNAs and the intragenic lncRNAs are extracted based on their overlap with mRNAs exons and introns. We will report these lncRNA repertoires in terms of intergenic/intragenic lncRNA class, structure and expression and comparing these features between livestock species, tissues and sexes. By profiling the transcriptional landscape of lncRNAs in these 4 species, this data will further contribute to the global action for annotating functional elements of livestock genomes.

**Key Words:** long non-coding RNAs, multispecies, Functional Annotation of Animal Genomes (FAANG), comparative genomics, RNA-seq

**21 Genome-wide CRISPR knockout screening for host factors involved in bovine herpes virus type 1 infection.** W. Tan*[1], I. Dry[1], S. Lillico[2], C. B. A. Whitelaw[2], and B. Dalziel[1], [1]*Division of Infection and Immunity, the Roslin Institute, University of Edinburgh, Edinburgh, Scotland;* [2]*Division of Developmental Biology, the Roslin Institute, University of Edinburgh, Edinburgh, Scotland.*

Bovine herpes virus type 1(BHV-1) causes infectious bovine rhinotracheitis, fatalities in calves and pregnancy abortions in cows, leading to huge economy loss in Ireland and the UK. Unfortunately, little is known about how host factors interact with this virus, and our lack of knowledge is impeding vaccine and drug developments. CRISPR/Cas9 are novel molecular scissors that cut DNA in a site-specific manner and have been used to edit many genes in livestock. It relies on base pairing between the small CRISPR guide RNA(gRNA) and target DNA, an activity that leads the gRNA-bound Cas9 protein to exert DNA cutting, often resulting in gene inactivation or conversion changes. Since the specificity of CRISPR/Cas9 is largely determined by the base pairing, it is straightforward to introduce many guides together to achieve knockout of multiple genes in parallel. Various groups have fully utilised this strategy and generated CRISPR libraries that target every gene in given organisms. Cells treated with such libraries lead to dissections of host factors involved in infections such as those from flaviviruses and HIV. To enable whole genome screening in cattle, we are creating a CRISPR library with ~120,000 guides targeting every gene in the bovine genome. We will then use it to screen for host factors involved in BHV-1 infection. To achieve this, MDBK cells that stably express Cas9 from the *rosa26* locus will be transduced by a lentivirus library that expresses all guides at low titre, ensuring single knockout events in the majority of the cells. The resulting cell pool encompassing knockout events in all genes will then be challenged by recombinant BHV-1 virus with mCherry. Cell populations with early and late mCherry signals and cells devoid of mCherry post infection will be obtained by FACS sorting. Integrated gRNA in these populations will be amplified and sequenced by Illumina Hiseq which will reveal copy numbers of guides; enriched guides indicate inhibitory roles of corresponding genes whereas depleted guides imply facilitator roles of targeted genes. These candidates will then be validated by transfecting individual guides into MDBK cells which will also be challenged by the virus.

**Key Words:** BHV-1, CRISPR-Cas9, whole genome screening, gene discovery