

# International Sheep Genomics and the International Goat Genome Consortiums combined virtual meeting

June 8<sup>th</sup>/9<sup>th</sup> to June 10<sup>th</sup>/11<sup>th</sup> 2021



## Presenter abstracts

### 1. Functional annotation of the reference Rambouillet sheep genome

**Kimberly M. Davenport**<sup>1</sup>, Alisha T. Massa<sup>2</sup>, Michelle R. Mousel<sup>3,4</sup>, Maria K. Herndon<sup>2</sup>, Stephen N. White<sup>2,3,5</sup>, Mazdak Salavati<sup>6</sup>, Emily Clark<sup>6</sup>, Alan Archibald<sup>6</sup>, Suraj Bhattarai<sup>7</sup>, Stephanie D. McKay<sup>7</sup>, Kim C. Worley<sup>8</sup>, Brian Dalrymple<sup>9</sup>, James Kijas<sup>10</sup>, Alex Caulton<sup>11</sup>, Shannon Clarke<sup>11</sup>, Rudiger Brauning<sup>11</sup>, Tracy Hadfield<sup>12</sup>, Noelle E. Cockett<sup>12</sup>, Timothy P.L. Smith<sup>13</sup>, and Brenda M. Murdoch<sup>1,5</sup> on behalf of The Ovine FAANG Project Consortium

<sup>1</sup>Department of Animal, Veterinary, and Food Science, University of Idaho; <sup>2</sup>Department of Veterinary Microbiology and Pathology, Washington State University; <sup>3</sup>USDA, ARS, Animal Disease Research Unit; <sup>4</sup>Paul G. Allen School for Global Animal Health, Washington State University; <sup>5</sup>Center for Reproductive Biology, Washington State University; <sup>6</sup>The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh; <sup>7</sup>University of Vermont; <sup>8</sup>Baylor College of Medicine-Human Genome Sequencing Center; <sup>9</sup>University of Western Australia; <sup>10</sup>CSIRO Agricultural Flagship; <sup>11</sup>AgResearch; <sup>12</sup>Utah State University; <sup>13</sup>USDA, ARS, U.S. Meat Animal Research Center (USMARC)

Annotation of regulatory elements is a significant contributor to the investigation of complex phenotypic traits related to health and production in sheep. The Ovine Functional Annotation of Animal Genomes (FAANG) project, part of the broader livestock species FAANG initiative, aims to identify and characterize gene regulatory elements in domestic sheep. Our previous pilot study in sheep examined the relationship between histone modification ChIP-seq and DNA methylation in tissues from three different developmental lineages. The Ovine FAANG Project incorporated all core FAANG assays across a comprehensive set of tissues from the reference Rambouillet ewe Benz 2616. ChIP-seq analyses were performed on four histone modifications that are known as promoter (H3K4me3), active enhancer (H3K27ac), poised enhancer (H3K4me1), or repressed enhancer (H3K27me3) gene regulatory elements. ATAC-seq was performed on most of the same tissues to annotate regions of open chromatin. Sequences from ChIP-seq and ATAC-seq were mapped to the ARS-UI\_Ramb\_v2.0 genome for functional analyses. PCA clustering of ChIP-seq and ATAC-seq signals revealed separation by tissue type. Within a tissue type, active and poised histone modifications had a greater Spearman correlation with each other compared with the repressive mark, suggesting that active regulatory elements are similar within tissue types. Fourteen chromatin states including promoters, active enhancers, poised enhancers, repressed enhancers, and accessible chromatin were defined in each

tissue with ChromHMM based on combinations of ChIP-seq and ATAC-seq data. This high-resolution annotation of regulatory elements in the Rambouillet reference genome provides a valuable resource to facilitate a deeper understanding of biological mechanisms that contribute to complex traits important to global sheep production.

## 2. A gene expression atlas of the goat brain

**María Gracia Luigi-Sierra**<sup>1</sup>, Dailu Guan<sup>1</sup>, Manel López-Béjar<sup>2</sup>, Encarna Casas<sup>2</sup>, Sergi Olvera<sup>2</sup>, Jaume Gardela<sup>2</sup>, María Jesús Palomo<sup>3</sup>, Uchebuchi Ike Osuagwu<sup>3</sup>, Uchechi Linda Ohaneje<sup>3</sup>, Emilio Mármol-Sánchez<sup>1</sup>, Marcel Amills<sup>1,4</sup>

<sup>1</sup>Centre de Recerca Agrigenòmica (CRAG), Campus Universitat Autònoma de Barcelona, Bellaterra 08193, Spain; <sup>2</sup>Department of Health and Animal Anatomy, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain; <sup>3</sup>Department of Animal Medicine and Surgery, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain; <sup>4</sup>Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain

Brain is a complex organ formed by multiple highly specialized structures involved in the regulation of memory, behaviour, learning, sensory function, motor skills and body homeostasis. In this work, we aimed to build an atlas of gene expression of the goat brain. We collected samples from 12 regions (medulla oblongata, pons, cerebellar trunk, cerebellar hemisphere, rostral colliculus, olfactory bulb, hippocampus, frontal neocortex, hypothalamus, pineal gland, neurohypophysis and adenohipophysis) that were retrieved from seven Murciano-Granadina goats, three of which were 2 months pregnant at the time of slaughter (these animals were killed by reasons unrelated with this experiment). Total RNA was extracted from each tissue and subsequently sequenced. Gene expression levels were inferred and we carried out differential mRNA expression analyses between pregnant and non-pregnant goats. Through PCA and hierarchical clustering analyses based on gene expression estimates of 18,113 genes, we found that brain tissues group according to their embryonic origin and anatomical location i.e. (A) Telencephalon: olfactory bulb, hippocampus, and frontal neocortex; (B) Diencephalon: hypothalamus; (C) Mesencephalon: rostral colliculus; (D) Myelencephalon: pons and medulla oblongata; (E) Metencephalon: cerebellar trunk and hemisphere. The only exceptions to this trend were the pineal gland and hypophyseal tissues, which have a diencephalic origin but show a very differentiated pattern of gene expression. The cerebellum also displayed a highly differentiated mRNA profile. The olfactory bulb, which is deeply involved in the establishment of maternal behaviour in mice, was the organ showing the highest number of differentially expressed mRNAs between pregnant and non-pregnant goats (1,234 genes).

## 3. High-quality ovine reference genome assemblies

Kimberly M. Davenport<sup>1\*</sup>, Shavahn C. Loux<sup>2\*</sup>, Derek M. Bickhart<sup>3</sup>, Kim C. Worley<sup>4</sup>, Brad A. Freking<sup>5</sup>, Brian L. Vander Ley<sup>6</sup>, Shwetha C. Murali<sup>4</sup>, Steven G. Schroeder<sup>7</sup>, Noelle Cockett<sup>8</sup>, Michael P. Heaton<sup>5</sup>, Shannon Clarke<sup>9</sup>, Thomas W. Murphy<sup>5</sup>, Theodore S. Kalbfleisch<sup>2</sup>, Brenda M. Murdoch<sup>1</sup>, Timothy P.L. Smith<sup>5</sup>, and Benjamin D. Rosen<sup>7</sup>

<sup>1</sup> Department of Animal, Veterinary, and Food Science, University of Idaho, Moscow, ID

<sup>2</sup> Gluck Equine Research Center, University of Kentucky, Lexington, KY

<sup>3</sup> USDA, ARS, U.S. Dairy Forage Research Center, Madison, WI

4 Human Genome Sequencing Center and Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX

5 USDA, ARS, U.S. Meat Animal Research Center, Clay Center, NE

6 Great Plains Veterinary Educational Center, University of Nebraska-Lincoln, Lincoln, NE

7 USDA, ARS, Animal Genomics and Improvement Laboratory, Beltsville, MD

8 Utah State University, Logan, UT

9 AgResearch, Invermay Agricultural Centre, Mosgiel, New Zealand

Long-read sequencing technologies and new assembly methods have greatly increased the quality of genome assemblies. Taking advantage of these advances, we generated reference quality assemblies for the Rambouillet (Ramb), White Dorper (WD), and Romanov (Rom) breeds. We re-assembled the same Ramb ewe chosen for the Ovine Functional Annotation of Animal Genomes (FAANG) project. The assembly was submitted to NCBI GenBank as ARS-UI\_Ramb\_v2.0. It has improved continuity, with a contig N50 of 43.19 Mb (contig length needed to cover 50% of the genome), 84 gaps, and one contig spanning the entire length of chromosome 24. Importantly to the FAANG effort, it has greater per-base accuracy and fewer insertions and deletions identified from mapped RNA sequence than previous assemblies. The WD and Rom breeds form the foundation of a composite population displaying superior characteristics in reproduction, health, and welfare. We applied the trio-binning approach to obtain fully haplotyped-resolved assemblies of each parental genomes from a male offspring sired by a WD ram mated with a Rom ewe. Both assemblies are highly contiguous with contig N50s ~ 62Mb, each possessing multiple chromosome-spanning contigs (WD – 9 and 19; Rom – 5, 8, 14, 19, and 24) and multiple chromosomes with a single gap (WD – 5, 12, 13, 14, 16, 24, and 26; Rom – 9, 11, 12, 15, 17, 18, 23, and 25). These three assemblies will provide benefits as breed specific references now and form the foundation of future efforts where ovine genetic variation is captured in the Sheep Pangenome.

\*These authors contributed equally

Keywords: Sheep, Genome, Genome Assembly, Reference Genome, Rambouillet, White Dorper, Romanov

#### **4. Underpinning sheep genomics: variant calling from the “1000 Sheep Genome Project” (SheepGenomeDB).**

**R. Brauning**<sup>1</sup>, S. McWilliam<sup>2</sup>, H. Daetwyler<sup>3</sup>, A. McCulloch<sup>1</sup>, N. Cockett<sup>4</sup> and S. Clarke<sup>1</sup>

on behalf of the members of the International Sheep Genomics Consortium

<sup>1</sup>AgResearch, New Zealand, <sup>2</sup>CSIRO, Australia, <sup>3</sup>Agriculture Victoria, Australia, <sup>4</sup>Utah State University, USA

The long-term goals of the International Sheep Genomics Consortium (ISGC) to develop underpinning resources for the sheep research community has resulted in continued improvement of the sheep genome assembly and development of low, medium and high-density Illumina SNP chips. The ISGC members have continued to make available whole genome sequence data to the community that has been captured via the Sheep Genomes Database, an initiative of the ISGC that extends the consortiums recent achievements. SheepGenomesDB is an electronic warehouse containing sequence variants called from the expanding collection of sheep genomes. Through the application of a single harmonised pipeline for read QC, mapping, variant detection and annotation, SheepGenomesDB makes available variant collections derived in a standardised manner. Run 2 has seen ~1000 animals analysed with variant collections positioned on the OAR V3.1 with Run 3 reaching ~1500 animals analysed. For Run 3, the updated Rambouillet genome assembly was utilised and provides users with tools to obtain

variants defined by chromosomal location, SNP annotation results or via animals and breeds of interest.

## 5. An atlas of variant ages in sheep genome.

**Charlotte HER**, François POMPANON (supervisor)

Laboratoire d'Ecologie Alpine, Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS, Grenoble, F-38000, FRANCE

Domestication represents a major turning point in human evolution, leading to the emergence of farming during the Neolithic. By providing a series of independent long-term evolutionary experiments where animals were selected for specific traits, this process has been of longstanding interest to evolutionary biologists. It is well studied in sheep thanks to genomics. Modern DNA is widely used to find out which traits have been selected by humans, to identify their genetic basis or to distinguish populations from each other (etc.). However, the chronology of these selective and demographic events, and even the history of mutations remains poorly understood. In this context, we have constructed an atlas of variant ages for about 40 million mutations in the sheep genome from a sample of almost 400 worldwide whole genomes with the GEVA method (Albers and McVean, 2020). It could allow us to compare the age distributions of mutations in different populations, to trace the connections between them over time or to estimate the time of appearance of variants selected by humans. We are also completing this atlas by dating the selection of these variants with other methods such as CLUES (Stern, Wilton and Nielsen, 2019) or startmrc (Smith et al., 2018). This atlas therefore represents a valuable database for the community, which could grow up with the arrival of new samples. It is a new asset for understanding the processes and history of domestication but also for characterizing genomic regions or populations constituting a potentially ancient heritage for conservation purposes.

Albers, P.K. and McVean, G. (2020) Dating genomic variants and shared ancestry in population-scale sequencing data. *PLOS Biology*. 18 (1), pp. e3000586. doi:10.1371/journal.pbio.3000586.

Smith, J., Coop, G., Stephens, M. and Novembre, J. (2018) Estimating Time to the Common Ancestor for a Beneficial Allele. *Molecular Biology and Evolution*. 35 (4), pp. 1003–1017. doi:10.1093/molbev/msy006.

Stern, A.J., Wilton, P.R. and Nielsen, R. (2019) An approximate full-likelihood method for inferring selection and allele frequency trajectories from DNA sequence data [online]. Available from: <http://biorxiv.org/lookup/doi/10.1101/592675> doi:10.1101/592675 [Accessed 15 July 2019].

## 6. FarmGTEx and the characterisation of transcriptomic variation in sheep and goat.

**Pau Navarro**<sup>1</sup>, Shuli Liu<sup>2</sup>, Yahui Gao<sup>2</sup>, George Liu<sup>2</sup>, Albert Tenesa<sup>1,3</sup> and Lingzhao Fang<sup>1</sup> on behalf of the FarmGTEx Consortium

<sup>1</sup>MRC Human Genetics Unit at the Institute of Genetics and Cancer, The University of Edinburgh, Edinburgh EH4 2XU, UK

<sup>2</sup>Animal Genomics and Improvement Laboratory, Henry A. Wallace Beltsville Agricultural Research Center, Agricultural Research Service, Agricultural Research Service, USDA, Beltsville, Maryland 20705, USA

<sup>3</sup>The Roslin Institute, Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Midlothian EH25 9RG, UK

The systematic characterization of genomic variants regulating the transcriptome of livestock is essential for interpreting the molecular mechanisms underlying complex traits, understanding adaptive

evolution and domestication, optimising genetic improvement programmes and allowing comparative genomic studies.

The Farm Animal Genotype-Tissue Expression (**FarmGTE<sub>x</sub>**) consortium is a collaborative endeavour that includes over 20 Universities and Institutes around the world. Its aim is to provide a comprehensive **atlas of tissue-specific gene expression and genetic regulation in livestock**.

Our first work has focussed on cattle, and we have shown that by uniformly analysing publicly available RNA sequence data ( $n > 10K$ ) via our newly developed transcriptome pipeline, we can get new insights on the transcriptome across 24 tissues, including the discovery of hundreds of thousands of *cis*- and *trans*- genetic associations with gene expression and alternative splicing (eQTL and sQTL) and the evaluation of the tissue specificity/similarity of these genetic regulatory effects. We also linked gene expression in different tissues to 43 economically important traits in cattle using a large transcriptome-wide association study (TWAS) and colocalization analysis to provide novel biological insights into the molecular regulatory mechanisms underpinning agronomic traits (Liu et al, 2020).

We are following-up this work in other livestock species, including sheep and goats. We will discuss the opportunities such research using publicly available data presents, and how it could be enhanced through collaboration with other international efforts such as the International Sheep and Goat Genomics Consortiums, as well as and how we plan to further characterise transcriptomic variation and its relationship to phenotype by inviting contributions from interested parties.

Liu et al. (2020) A comprehensive catalogue of regulatory variants in the cattle transcriptome. <https://doi.org/10.1101/2020.12.01.406280>

## 7. Constructing the Ovine Pan-genome

**Brenda M. Murdoch**, Noelle Cockett, Tracy Hadfield, Michelle Mousel, Stephen White, Emily Clark, Mazdak Salavati, Shannon Clarke, Andrew Hess, Rudiger Brauning, Darren Hagen, Theodore S. Kalbfleisch, Michael Heaton, Derek Bickhart, Benjamin Rosen, Timothy P.L. Smith

A wealth of scientific advancements, research tools, and discoveries have emerged based on establishment of reference genomes for livestock species, enabling our abilities to link genome to phenome. Advances in sequencing technologies now enable construction of *de novo* genome assemblies at much greater quality than previously possible, presenting the possibility of using breed-specific reference genomes in genomic studies. Moreover, single representative reference genomes of a species are inadequate to describe the full genome variation across populations within that species. This project aims to bring together the resources necessary to describe inter-breed genetic diversity in sheep. Our objectives are to produce high-quality reference genomes, with the recently described trio binning approach. Different breeds and species of sheep were chosen based on their divergence for important traits. We aim to identify breed-specific genomic regions associated with a range of phenotypically distinct traits. Additionally, we will use Iso-Seq and RNA-Seq to provide breed-specific transcriptomes for fetal tissues from F1 crosses generated for the trio binning approach. Finally, this project will generate and make available genetic resource tools for studying the ovine genome through online databases, including those supported by the international FAANG Consortium and, the Ensembl genome portal for farmed and companion animals. This project will increase our understanding of how genomic complexity within a species can result in phenotypic differences and will empower scientists to predict the physiological outcomes more effectively from genomic variation toward the improvement of breeding strategies.

## 8. Expression Profile of Ghrelin gene in goat breeds in Kerala

Charlotte Coretta Rodricks<sup>1</sup>, Thirupathy Venkatachalapathy<sup>1</sup>, T.V. Aravindakshan<sup>1</sup>, Manoj M<sup>1</sup>

<sup>1</sup>Department of Animal Genetics and Breeding College of Veterinary and Animal Sciences, Mannuthy

Ghrelin, a 28 amino acid orexigenic peptide, is an endogenous ligand for the growth hormone secretagogue receptor. It is secreted by the oxyntic cells present in the gastric mucosa and plays a vital role in regulating energy homeostasis and feed intake. It acts as adiposity inducer, appetite stimulant and also has been reported to regulate gastric motility. The present study was carried out to find out the expression profile of ghrelin gene in different tissues in native goat breeds in Kerala. Tissue samples (ovary, fallopian tube, abomasum, lung, spleen and liver) were collected from three Malabari and Attappady Black breed goats. Total RNA was isolated, and Real Time-PCR was performed to find out the relative expression. The expression of  $\beta$ -Actin was taken as reference and the  $2^{-\Delta\Delta Ct}$  values were calculated. It was observed that there was a higher expression of ghrelin in the abomasal tissues and least expression in the liver tissues. Significantly higher level of expression was noticed in Attappady Black than Malabari goats. It was observed that there was a fold increase for lung (3.85), abomasum (3.80) and ovary (4.07) in Attappady Black breed over Malabari breed.

KEYWORDS : Ghrelin, Malabari, qRT-PCR, Attappady Black, expression

## 9. Determining developmental transcriptome differences in tissues related to growth traits in sheep using RNA-sequencing

Shernae A. Woolley<sup>1</sup>, Mazdak Salavati<sup>1</sup>, Emily L. Clark<sup>1</sup>

<sup>1</sup>The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, United Kingdom

The sheep transcriptome offers numerous opportunities for understanding the biological function and the roles of genes in maintaining health and productivity. Growth traits in sheep are important for production and are often polygenic, making the identification of trait-specific genomic markers for genomic selection challenging. This study aims to identify transcriptionally active genes associated with growth through the generation and analysis of RNA-sequencing data across different developmental stages and tissue types in Texel x Scottish Blackface sheep. A dataset consisting of 48 samples from three prenatal and four postnatal developmental stages was generated from whole embryos (n=3), skeletal bicep muscle (n=13), liver (n=16), ovary (n=8), placentome (n=4) and caruncle (n=4) tissue. Transcript per million expression estimates were generated using Kallisto based on the gene models from the Oar\_rambouillet\_v1.0 genome (Ensembl v103), and network cluster analysis was performed in Graphia. Gene-to-gene network analysis revealed large clusters of genes expressed in ovary, liver and skeletal bicep tissues. One of the largest clusters showed high expression levels in newborn to 8 week old lambs in skeletal bicep tissue, and were associated with GO terms for muscle fibre organisation and contraction. Tissue- and developmental stage-specific differences in gene expression were observed, with the *MSTN (GDF8)* gene displaying increased expression in skeletal bicep tissue in day 100 fetuses relative to the other developmental stages. These results provide a foundation for further analysis of the transcriptional differences across developmental stages in sheep, and will be integrated with information from GWAS for growth traits to generate a set of functional variants that could be integrated in genomics enabled breeding programmes.

## 10. Transcriptomic analysis of the early mucosal host response in goats infected with *Haemonchus contortus*

**Jean-Christophe Bambou**<sup>1</sup>, Hadeer M. Aboshady<sup>2</sup>, Y. Félicité<sup>1</sup>, Nathalie Mandonnet<sup>1</sup>

<sup>1</sup>*INRAE, URZ, 97170, Petit-Bourg (Guadeloupe), France*

<sup>2</sup>*Department of animal production, Faculty of Agriculture, Cairo University, Cairo, Egypt*

In small ruminant production, gastrointestinal nematode (GIN) infection remains one of the major economic constraints affecting animal health and welfare. The aim of this study was to analyse the abomasal mucosa transcriptome of naïve and immune goats at early time points after GIN infection to identify the different pathways and regulators involved in the early host protective immune response. Naïve and immune Creole kids were orally infected with 10,000 *Haemonchus contortus* infective larvae, and abomasal mucosa was sampled at 0, 4 and 6 days post-infection (dpi). High-throughput sequencing of cDNA from the abomasal mucosa was used to quantitatively analyse the transcriptome. Number of differentially expressed genes (DEG) was very low for both naïve and immune animals when comparing day 0 versus day 4 post-infection. A total of 2237 and 3206 DEGs were identified comparing 0 versus 6 dpi in naïve and immune animals, respectively. Meanwhile comparing immune versus naïve at 6 dpi, showed only 18 DEGs. Ingenuity pathway analysis (IPA) showed that several immune responses were activated in immune compared with naïve animals at 0 and 4 dpi such as Th2 and Th1 pathways, natural killer cell, B cell receptor, IL-2 and IL-15 signaling. In the other hand, both naïve and immune animals showed activation for those pathways comparing 0 versus 6 dpi. Similar pattern was recorded for upstream regulator genes which were related to immunity like TNF, IL1 $\beta$ , IL2, IL5, TGF $\beta$ 1, IFN $\gamma$ , TCR, IL18, IL6 and IL4. In conclusion, our results indicated that the immune response remains activated in immune animals even after several weeks in parasite-free conditions.

## **11. Genomic tools for the identification of loci associated with facial eczema in New Zealand sheep**

**KM McRae**<sup>1</sup>, SJ Rowe<sup>1</sup>, PL Johnson<sup>1</sup>, HJ Baird<sup>1</sup>, NG Cullen<sup>1</sup>, MJ Bixley<sup>1</sup>, JE Plowman<sup>1</sup>, S Deb-Choudhury<sup>1</sup>, R Brauning<sup>1</sup>, NC Amyes<sup>1</sup>, KG Dodds<sup>1</sup>, S-AN Newman<sup>1</sup>, JC McEwan<sup>1</sup> & SM Clarke<sup>1</sup>

*1. AgResearch, New Zealand 2. University of Otago, New Zealand*

Facial eczema (FE) is a metabolic disease of great importance in New Zealand ruminants. Liver and bile duct damage, caused by ingestion of the mycotoxin sporidesmin, results in photosensitisation, and reduced production. There is considerable genetic variation in the response to the ingestion of sporidesmin sheep, and breeding programs to increase FE tolerance have been successful. The objective of this study was to utilise a dataset of high- and low-density genotyped and phenotyped animals, to interrogate the sheep genome for regions associated with variability in tolerance to FE in New Zealand sheep.

A novel QTL on chromosome 15 at the  $\beta$ -globin locus is reported which explains 5% of the phenotypic variance in the response to FE. Mass spectrometry of haemoglobin from animals with differing genotypes at the  $\beta$ -globin locus indicated that the markers are associated with different forms of adult  $\beta$ -globin. Type A animals appeared to be more tolerant to FE, however, the overall frequency of haplotype A in genotyped animals was 0.5, indicating that the locus may be under balancing selection in these flocks. Haemoglobin haplotypes have previously been associated with variation in several health-related traits in sheep, and therefore warrant further investigation regards their role in tolerance to FE in sheep.

## **12. Functional genome and microbiome in blood of goats affected by the gastrointestinal pathogen *Haemonchus contortus***

**Yonathan Tilahun**, Jessica Quijada-Pinango, Qunhui Yang, Felicia Johnson, Charles Lett, Kayla Smith, Malcolm McCallum, Terry Gipson, Zaisen Wang - School of Agriculture and Applied Sciences, Langston University, Langston, Oklahoma 73050, USA

Peter Hoyt - Institute for Biosecurity and Microbial Forensics, Oklahoma State University, Stillwater, Oklahoma 74078, USA

Andrew Tritt - Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

The Alpine goat *Capra aegagrus hircus* is parasitized by the barber pole worm (*Haemonchus contortus*). This relationship results in changes that affect the gene expression of the host, the pest, and the microbiome of both. Hematological parameters indicating genes that are expressed and the percent composition of abundant and diverse microbial flora are reflective of infestation. We explored the similarity between blood samples of non-infected, infected, infected zoledronic acid-treated, and infected antibody (anti- $\gamma\sigma$  T cells) treated wethers under controlled conditions. We identified responses to barber pole worms using blood-based analysis of transcripts and the microbiome. Seven days post-inoculation (dpi) we identified 7,627 genes associated with different treatment types. Across all treatments we identified fewer raw read counts and a reduced diversity in microbial flora on 7 dpi than in 21 dpi wethers. We also identified that there were differences in percent composition of microbial flora known to be associated with inflammation. This study identifies treatment specific genes, and an increase in microflora abundance and diversity as wethers age post infestation. Further, *Firmicutes/Bacteroidetes* (F/B) ratio reflect metabolic health, based on depression or elevation above thresholds defined by the baseline of non-infected hosts depending on the type of intrusion exhibited by the pest.

### **13. A recessive stop-gain mutation in *CCDC65* is associated with lamb mortality in French Lacaune dairy sheep**

**Maxime Ben Braiek**<sup>1</sup>, Carole Moreno-Romieux<sup>1</sup>, Charlotte Allain<sup>2</sup>, Philippe Bardou<sup>3</sup>, Arnaud Bordes<sup>1</sup>, Frédéric Debat<sup>1</sup>, Florence Plisson-Petit<sup>1</sup>, David Portes<sup>2</sup>, Julien Sarry<sup>1</sup>, Némuel Tadi<sup>1</sup>, Florent Woloszyn<sup>1</sup> and Stéphane Fabre<sup>1</sup>

<sup>1</sup>*GenPhySE, Université de Toulouse, INRAE, ENVT, 31326 Castanet-Tolosan, France*

<sup>2</sup>*UE Domaine de La Fage, INRAE, 12250 Saint Jean et Saint Paul, France*

<sup>3</sup>*Sigenae, INRAE, 31326 Castanet-Tolosan, France*

[maxime.ben-braiek@inrae.fr](mailto:maxime.ben-braiek@inrae.fr)

*Funded by SMARTER (EU H2020 n°772787), PRESAGE (CASDAR n°20ART1532777) and HOMLET (APIS-GENE, Région Occitanie) projects*

In a previous work using 50k SNP data from Lacaune dairy sheep genotypes, we have identified 11 haplotypes with homozygous deficiency (LDHH) supposed to host recessive lethal mutations. Among them, LDHH6 located on OAR3 was the most frequent haplotype (12.1% of heterozygous carriers) and it was associated with an increase in the stillbirth rate.

In the present study, we have analyzed the whole-genome sequences of two Lacaune rams heterozygous carriers of LDHH6, and we have compared them to the sequences of 23 non-carriers Lacaune rams and 68 other animals from 14 different sheep breeds. After variant filtering, we have identified a SNP in the two LDHH6 carriers whose variant allele induced a premature stop-gain mutation in the Coiled-Coil Domain Containing 65 (*CCDC65*) gene. *CCDC65* is involved in the assembly of the nexin-dynein complex for the formation of microtubules in ciliated cells and mutations in *CCDC65* are associated with “*Ciliary dyskinesia, primary 27*” causing respiratory distress in human.

In order to identify the affected phenotype in sheep, we generated 17 at-risk matings between carrier rams and carrier ewes genotyped for the candidate mutation in *CCDC65*. A total of 16 lambs were born alive, and among them, five were identified as homozygous carriers and died between the age of one week and one month, all suffering from respiratory problems as tachypnea. At the autopsy, we particularly observed a broad hepatization of lungs involving infectious pneumonia.

Management of this causal mutation in the Lacaune dairy sheep selection scheme through reasoned



mating of carrier rams and ewes could improve overall lamb viability by 2%.

## **14. Genome-wide association studies for meat quality traits in New Zealand sheep imputed to whole-genome sequence**

**A. Hess**<sup>1</sup>, L. Brito<sup>2</sup>, M. Lee<sup>3</sup>, W. Bain<sup>1</sup>, N. Pickering, S. Newman<sup>1</sup>, R. Brauning<sup>1</sup>, M. Hess<sup>1</sup>, J. McEwan<sup>1</sup>, S. Clarke<sup>1</sup>

1. AgResearch, New Zealand 2. Purdue University, USA 3. University of Otago, New Zealand

Genotyping for DNA-based parentage and genomic selection is routine in the New Zealand sheep industry. To-date, ~23,000 New Zealand sheep have been genotyped with the OvineHD BeadChip array (600K SNPs). Imputation to whole-genome sequence (WGS) has the potential to identify causative mutations associated with traits of economic importance to the New Zealand sheep industry. The sheep population in New Zealand is very diverse, with over 50 purebred or composite breeds observed in the Sheep Improvement Limited database, and frequent crossbreeding. The goal of this study was to utilize whole-genome sequence data to impute individuals with HD genotypes to WGS using Beagle 5.1. Individuals with WGS data from Run2 of the International Sheep Genomics Consortium (ISGC), including 213 individuals from New Zealand, were used as a reference to impute individuals with HD genotypes to WGS. Overall imputation accuracy was high (average individual accuracy 0.97) and was consistently high among all major breeds with high density genotype data, showing the ability to utilize the individuals with WGS sequence data from the ISGC to impute the diverse New Zealand sheep population. Genotypes imputed to WGS were used in genome-wide association studies on meat quality traits, which identified QTL with strong candidate genes. Very few of the top SNPs were in the intragenic regions of candidate genes. Currently, we are undergoing work to perform targeted long-read sequencing of these regions with the aim of identifying putative causative variants for these QTL. This information can be used to improve the design of the available genotyping arrays.

## **15. Genome-wide association study of reproductive efficiency in Florida dairy goats**

**C. Ziadi**<sup>1</sup>, N. Laseca<sup>1</sup>, M. Sánchez<sup>2</sup>, M. Ramón<sup>3</sup>, S. Demyda-Peyrás<sup>4</sup>, E. Muñoz<sup>5</sup>, MD. López<sup>6</sup>, A. Molina<sup>1</sup>

<sup>1</sup>Departamento de Genética, Universidad de Córdoba, Campus de Rabanales, Ctra, Madrid–Cadiz, km 396, 14071 Córdoba, Spain. <sup>2</sup>Departamento de Producción Animal, Universidad de Córdoba, Córdoba, Spain. <sup>3</sup>Centro Regional de Selección y Reproducción Animal (CERSYRA-IRIAF), Valdepeñas, Ciudad Real, Spain. <sup>4</sup>Facultad de Ciencias Veterinarias, Universidad Nacional de la Plata, La Plata, Argentina. <sup>5</sup>Departamento de Patología Animal, Producción Animal, Bromatología y Tecnología de los Alimentos, Universidad de Las Palmas de Gran Canaria, Campus Universitario Cardones de Arucas, Arucas, Spain. <sup>6</sup>Edificio de Producción Animal, ACRIFLOR, Campus de Rabanales, Córdoba, Spain.

Reproductive efficiency is a key factor to ensure farm profitability in dairy goats. Although goat genome annotation has not been completed, genome-wide association (GWAS) analyses have been performed for several economically important traits, except for fertility. In this study we aimed to identify genomic regions associated with reproductive efficiency (RE), defined as the deviation between optimal and real parity number at age of each doe in Florida goats. RE values were estimated as phenotype and deregressed pseudophenotypes (breeding values) in 611 animals. Genomic data was obtained using

the Illumina 55 K Goat Bead-Chip (53,347 SNPs) and filtered using PLINK v1.9. SNP markers with call-rate  $\leq 0.95$ , with a minor allele frequency  $< 0.01$  and linkage disequilibrium pruning ( $r^2 > 0.5$ ) were excluded. 46,715 SNPs were retained for analysis. The GWAS analysis was performed using several alternative approaches (using RE phenotype or RE pseudophenotypes, taking or not taking into account the genomic relationship matrix and considering or not genomic principal components (PC) as a covariate) using the GEMMA software; aiming to determine the best fitting model for fertility in Florida dairy goats. To our knowledge, this study is the first attempt to identify regions associated with female fertility in dairy goats and could be useful to improve goat female fertility.

## 16. Genome-wide analysis of copy number variation identifies regions associated with fat deposition in thin and fat-tailed sheep breeds

Kobra Taghizadeh<sup>1</sup>, **Mohsen Gholizadeh**<sup>1\*</sup>, Ghodrath Rahimi-Mianji<sup>1</sup>, Mohammad Hossein Moradi<sup>2</sup>, Roy Costilla<sup>3</sup>, Stephen Moore<sup>3</sup>, Rosalia Di Gerlando<sup>4</sup>

<sup>1</sup> Department of Animal Science, Faculty of Animal and Aquatic Science, Sari Agricultural Sciences and Natural Resources University, SARI-IRAN

<sup>2</sup> Department of Animal Science, Faculty of Agriculture and Natural Resources, Arak university, ARAK-IRAN

<sup>3</sup> Queensland Alliance for Agriculture and Food Innovation. University of Queensland, Australia

<sup>4</sup> Dipartimento di Scienze Agrarie, Alimentari e Forestali, Università degli Studi di Palermo, Italy

In this study, to investigate the copy number variation (CNV) in thin and fat-tailed sheep breeds, we used OvineSNP50K genotype data from three Iranian sheep breeds including 96 Baluchi sheep, 47 Lori-Bakhtiari sheep (as fat-tailed breeds) and 47 Zel sheep (as thin-tailed breed). For association test, to identify genomic regions associated with fat deposition, 70 samples of Valle del Belice sheep were also added to the association test as thin-tailed sheep breed. PennCNV and CNVRuler software were, respectively, used to study the copy number variation and genomic association analyses. We detected 573 and 242 CNVs in the fat and thin tailed breeds, respectively. In terms of CNV regions (CNVRs), these represented 328 and 187 CNVRs that were within or overlapping with 790 known Ovine genes. The CNVRs covered approximately 73.85 Mb of the sheep genome with average length 146.88 kb, and corresponded to 2.6% of the autosomal genome sequence. Five CNVRs were randomly chosen for validation, of which 4 were experimentally confirmed using Real time qPCR. Functional enrichment analysis showed that genes harbouring CNVs in thin-tailed sheep were involved in the adaptive immune response, regulation of reactive oxygen species biosynthetic process and response to starvation. In fat-tailed breeds these genes were involved in cellular protein modification process, regulation of heart rate, intestinal absorption, olfactory receptor activity and ATP binding. Association test identified one copy gained CNVR on chromosomes 6 harbouring two protein-coding genes *HGFAC* and *LRPAP1*. Our findings provide information about genomic structural changes and their association to the interested traits including fat deposition and environmental compatibility in sheep.

Keywords: thin-tailed, fat-tailed, Sheep, SNP, CNV

## 17. Development of epigenetic clocks for New Zealand livestock

**Alex Caulton**<sup>\*1,2</sup>, Ken G. Dodds<sup>1</sup>, Amin Hagani<sup>3</sup>, Joseph A. Zoller<sup>4</sup>, Christine Couldrey<sup>5</sup>, Steve Horvath<sup>3</sup>, Shannon M. Clarke<sup>1</sup>

<sup>1</sup> AgResearch, Invermay Agricultural Centre, Mosgiel, Otago, New Zealand

<sup>2</sup> Department of Biochemistry, University of Otago, Dunedin, Otago, New Zealand

<sup>3</sup> Department of Human Genetics, David Geffen School of Medicine, University of California Los

Angeles, Los Angeles, CA 90095, USA

<sup>4</sup>Department of Biostatistics, Fielding School of Public Health, University of California Los Angeles, Los Angeles, California, USA

<sup>5</sup>Livestock Improvement Corporation, Hamilton, New Zealand

Robust biological biomarkers of chronological age have been developed in humans and model mammalian species such as rat and mice using DNA methylation data. The concept of these so called “epigenetic clocks” has emerged from a large body of literature describing the correlation between genome-wide methylation levels and age. Epigenetic clocks exploit this phenomenon and use small targeted panels of differentially methylated CpG sites to make robust predictions of chronological age, independent of tissue type.

Here we present highly accurate livestock epigenetic clocks whereby we have utilised the custom mammalian methylation array “HorvathMammalMethyl40” to construct the first epigenetic clock for domesticated goat, cattle, red and wapiti-breed deer and composite-breed sheep, in addition to one New Zealand livestock clock for all four species. The livestock clock shows similarly high accuracies to the individual species clocks ( $r > 0.93$ ), utilising only ~215 CpG sites to make predictions of chronological age with a single mathematical model.

The applications of this livestock clock could extend well beyond the scope of chronological age estimates and we envision its application in breeding programmes as a predictor for age-related production traits. Many independent studies in human health demonstrating that a deviation between true age and clock derived molecular age is indicative of lifespan and age-related phenotypes. There is, therefore, untapped potential to utilise livestock clocks as a molecular phenotype for age related disease and productivity traits for selective breeding purposes.

## 18. Whole genome bisulfite sequencing reveals tissue specific DNA methylation profiles in Sheep.

**Suraj Bhattarai**<sup>1</sup>, Kimberly M. Davenport<sup>2</sup>, Alisha T. Massa<sup>3</sup>, Michelle R. Mousel<sup>4,5</sup>, Maria K. Herndon<sup>3</sup>, Stephen N. White<sup>3,4,6</sup>, Noelle Cockett<sup>7</sup>, Timothy P.L. Smith<sup>8</sup>, Brenda M. Murdoch<sup>2,6</sup> and Stephanie D. McKay<sup>1\*</sup> on behalf of The Ovine FAANG Project Consortium

<sup>1</sup>Department of Animal & Veterinary Sciences, University of Vermont, Burlington, VT

<sup>2</sup>Department of Animal, Veterinary, and Food Science, University of Idaho, Moscow, ID

<sup>3</sup>Department of Veterinary Microbiology and Pathology, Washington State University, Pullman, WA

<sup>4</sup>USDA, ARS, Animal Disease Research Unit, Pullman, WA

<sup>5</sup>Paul G. Allen School for Global Animal Health, Washington State University, Pullman, WA

<sup>6</sup>Center for Reproductive Biology, Washington State University, Pullman, WA

<sup>7</sup>Utah State University, Logan, UT

<sup>8</sup>USDA, ARS, U.S. Meat Animal Research Center (USMARC), Clay Center, NE

Epigenetic research is important in livestock species as it captures environmental cues with specific genomic information to further explain economically important traits. To identify and characterize regulatory features in the sheep genome, the Ovine Functional Annotation of Animal Genomes (FAANG) project was developed. In this study, we have performed genome wide DNA methylation analysis in three different tissues from four sheep. These tissues originate from three different primordial tissue types and are uniquely suited to characterize tissue specific DNA methylation profiles. Whole genome bisulfite sequencing (150 bases paired end) of cerebellum, liver, and spleen samples generated approximately 986 million, 1.07 billion and 904 million paired end reads, of which, 84.24%, 78.86% and 82.48% uniquely mapped to the reference genome (*Oar\_rambouillet\_v1.0*; genome size ~2.87 Gb) for the cerebellum, liver, and spleen samples, respectively, at an average depth of 21x. We found that the same tissues from different animals cluster together based on CG methylation. Moreover,

tissues differed in their proportion of hypomethylated regions (HMRs) with cerebellum, liver, and spleen containing an average 1.66%, 4.45% and 0.83% HMRs in their genomes, respectively. Further, HMRs which were shared between the three tissues were enriched to promoter regions. Interestingly, cerebellum exhibited five-seven folds higher methylation at non-CG sites compared to the liver and spleen. Finally, inspection of non-CG sites in cerebellum revealed that more than two-thirds of the methylated sites were in CA context. Thus, this study provides an extensive characterization of genome wide DNA methylation profile in three different tissues of sheep.

## 19. Industry implementation of genomic improvement in dairy goats using DNA sequencing

**John C McEwan**<sup>1</sup>, Mary Wheeler<sup>2</sup>, Shannon M. Clarke<sup>1</sup>, Rudiger Brauning<sup>1</sup>, Alex Caulton<sup>1</sup>, Tracey C. Van Stijn<sup>1</sup>, Hayley J. Baird<sup>1</sup>, Rayna M. Anderson<sup>1</sup>, Andrew Hess, Barry J. Foote<sup>3</sup>, Judy F. Foote<sup>3</sup>, A.W.N. (Sandy) Cameron<sup>4</sup>, Simon McLachlan<sup>5</sup>, Thor Blichfeldt<sup>6</sup>, Jette Jakobsen<sup>6</sup>, Sally-Anne Turner<sup>7</sup>, Ken G. Dodds<sup>1</sup>

<sup>1</sup> AgResearch, Invermay Agricultural Centre, Mosgiel, Otago, New Zealand

<sup>2</sup> AgResearch, Ruakura Research Centre, Hamilton, New Zealand

<sup>3</sup> 859 Russell Road, RD4, Hikurangi 0184, New Zealand

<sup>4</sup> Cameron Road, Meredith, Victoria 3333, Australia

<sup>5</sup> Gavins, 319 Ballard Rd, Gordonton, New Zealand

<sup>6</sup> NSG, Box 104, 1431 Ås, Norway

<sup>7</sup> Dairy Goat Co-operative, 18 Gallagher Drive, Hamilton 3240, New Zealand

In dairy goats, industry uptake of genomic technologies has been slow due to the small size of the industries coupled with the limited market for SNP array based technologies. However, the potential benefits of genomic selection in dairy goats are large, because key traits are sex limited, recorded post selection and pedigree recording in large dairy goat herds is problematic. This has led our laboratory to develop low-cost genotyping strategies based on sequencing. Two technologies, both of which can be multiplexed are used. The first is restriction enzyme-reduced representational sequencing (RE-RRS), commonly called genotyping by sequencing (Elshire et al., 2011). The second is an amplicon based method for specific putatively functional SNPs called GT-seq (Campbell et al., 2015). Historically, these approaches have suffered, in part because they are subject to missing or probabilistic genotyping calls. This made them difficult to integrate with existing genetic evaluation software. One alternative option, using imputation, also demanded extensive computation and development of higher density training sets. Over the past 7 years, statistical software has been developed to estimate genetic relationships, parentage, breed proportion, sex, linkage disequilibrium, Fst, linkage mapping, genome wide association, plus mate allocation while handling the probabilistic nature of DNA sequencing data directly (e.g. Dodds et al., 2019) and the developed software is available from <https://github.com/AgResearch>. The sequencing based technologies described above are currently used for 5 separate genetic evaluations from Australia, New Zealand and Norway, with more than 58,000 samples genotyped to date.

## 20. Using sequence variants to better characterize a QTL region and improve the accuracy of genomic evaluation in Saanen goats

Estelle Talouarn\*, Isabelle Palhière\*, Marc Teissier\*, Philippe Bardou†, Hélène Larroque\*, Virginie Clément‡, Gwenola Tosser-Klopp\*, Christèle Robert-Granié and **Rachel Rupp\***

\*GenPhySE, Université de Toulouse, INRAE, ENVT, F-31326 Castanet-Tolosan, France

† Sigenae, INRAE, F-31326 Castanet-Tolosan, France

‡ Institut de l'Élevage, F-31326 Castanet-Tolosan, France

The objective of this study was to assess the usefulness of sequence data in Saanen goats (N = 33) to better capture a QTL on chromosome 19 (CHI19) associated with production, udder type and SCS in order to improve the accuracy of predictions and fine map the QTL

A total of 1,207 50k genotypes were imputed to the sequence level. Four scenarios, each using a subset of CHI19 imputed variants, were tested. Sequence derived information included all CHI19 variants (529,576), all variants in the QTL region (22,269), 178 variants selected in the QTL region and added to the updated 50K chip, or 178 randomly selected variants on CHI19. Best overall results were obtained using ssGBLUP on 50k genotypes completed with all variants in the QTL region with +6.2% average increase in accuracy for nine traits (and up to +17.9% for fat yield). Using all CHI19 variants led to an overall decrease of 4.8% in prediction accuracy. The updated chip improved genomic evaluations by 3.1% to 6.4%.

No obvious candidate mutation was identified for the QTL. Phasing imputed sequence variants the region of the QTL highlighted a haplotype however that was present in Saanen (50 %) but almost absent in Alpine. This haplotype was associated with pleiotropic phenotypes : large animals with good production and less favorable udder type. The Alpine like haplotype showed opposite profiles: smaller animals with good udder type and less favorable milk production.

Altogether, using QTL-region variants imputed from sequence data offered a promising perspective for improving genomic evaluation and using QTL information in breeding schemes.

## **21. Accounting for genotype uncertainty of genotyping-by-sequencing data for genomic analyses in New Zealand dairy goats**

**A. Hess<sup>1</sup>, J. Rich<sup>1</sup>, B. Foote<sup>2</sup>, J. Foote<sup>2</sup>, R. Brauning<sup>1</sup>, A. Caulton<sup>1</sup>, M. Hess<sup>1</sup>, J. McEwan<sup>1</sup> and S. Clarke<sup>1</sup>**

1. AgResearch, New Zealand 2. Foote Goat Farm, New Zealand

Low-depth genotyping-by-sequencing (GBS) is commonplace in New Zealand dairy goat herds that utilise genomic information because it is a high-throughput and cost-effective alternative to SNP genotyping arrays. Low-depth sequencing results in uncertainty in genotyping, particularly for heterozygous animals which may be incorrectly called homozygous. Therefore, methods to account for genotype uncertainty are needed. A GRM that considers sequencing depth has previously been developed; however, other models may be more suitable for trait analysis, either to identify major QTL associated with a trait or to better model the genomic architecture of a trait. How this uncertainty is accounted for when utilising these models may impact the results of the analysis. We employed different methods to account for genotype uncertainty in a New Zealand dairy goat herd with milk records and low-depth GBS data. Different methods for accounting for genotype uncertainty resulted in differences in genetic parameter estimates, with methods that utilized more information resulting in parameter estimates closest to those previously reported. Furthermore, while all methods found similar peaks, the more informed methods resulted in larger and more precise GWAS peaks. Using genotypes obtained from targeted amplicon sequencing (i.e. GT-seq) at the putative QTL in DGAT1 for fat content suggested the observed differences were due to more informed models more accurately capturing the LD between the marker and the QTL. Interestingly, however, genotype calling method made no practical difference on genomic prediction accuracy. Ultimately, the best strategy may be to capture SNPs more confidently in regions containing large QTL and rely on low-depth sequencing to capture the relationships between individuals through a GRM.

## 22. Selection signature analysis of South African Merino-derived breeds identifies sweeps important in production traits

E.F. Dzomba<sup>1</sup>; M.A. van Der Nest<sup>2</sup>; J.N. Mthembu<sup>1</sup>; P. Soma<sup>3</sup>; M. Synmann<sup>4</sup>; M. Chimonyo<sup>5</sup> and F.C. Muchadeyi<sup>2</sup>

<sup>1</sup>Discipline of Genetics, School of Life Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, South Africa

<sup>2</sup>Agricultural Research Council, Biotechnology Platform, Private Bag X5, Onderstepoort, 0110, South Africa

<sup>3</sup>Agricultural Research Council, Animal Production and Improvement, Irene, South Africa

<sup>4</sup>Grootfontein Agricultural Development Institute, Private Bag X529, Middelburg, EC, 5900, South Africa

<sup>5</sup>Discipline of Animal and Poultry Science, School of Agricultural, Earth & Environmental Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, South Africa

Merino sheep are a breed of choice across the world popularly kept for their wool and mutton value. They are often reared as a pure breed or employed in crossbreeding and are a common component in synthetic breed development. This study evaluated genetic diversity, population structure and breed divergence in 279 animals of Merino and Merino-based sheep breeds of South Africa using the Illumina Ovine50K BeadChip. The sheep breeds analysed included the three Merino-derived breeds and five presumed ancestral populations of Merinos and non-Merino founding breeds of Damara, Ronderib Afrikaner and Nguni.

Highest genetic diversity values were observed in Dohne Merino (DM) with  $H_o = 0.39 \pm 0.01$  followed by Meatmaster and South African Merino (SAM) with  $H_o = 0.37 \pm 0.03$ . The level of inbreeding ranged from  $0.0 \pm 0.02$  (DM) to  $0.27 \pm 0.05$  (Nguni). The first Principal Component (PC1) separated the Merino, South African Mutton Merino (SAMM), DM and Afrino (AFR) from the Meatmaster, Damara, Nguni and Ronderib Afrikaner (RAfr). PC2 aligned each Merino derived breed with its presumed ancestors and separated the SAMM from the Merino and SAM. The  $|iHS|$  analysis yielded selection sweeps across the AFR (12 sweeps), Meatmaster (4 sweeps) and DM (29 sweeps). Genes associated with hair/wool traits such as *FGF12*, metabolic genes of *ICA1*, *NXPH1* and *GPR171* and immune response genes of *IL22*, *IL26*, *IFNAR1* and *IL10RB* were reported.

Overall, the study demonstrated some genetic similarities between the Merino and Merino-derived breeds emanating from common founding populations as well as some divergence driven by breed-specific selection goals.

*Key words: Merino-type sheep, population genetic structure, breed divergence, SNP genotypes, selection sweeps, EHH signatures*

## 23. Y-chromosomal haplotypes in domestic and wild goats reveal ancient paternal bottlenecks and recent introgressions

Vargoats Consortium<sup>1</sup>, Isaïc J. Nijman<sup>2,3</sup>, Benjamin D. Rosen<sup>4</sup>, Thomas Faraut<sup>5</sup>, Philippe Bardou<sup>6</sup>, Tristan Cumer<sup>5</sup>, Kevin G. Daly<sup>7</sup>, Zhuqing Zheng<sup>8</sup>, Jiang Yu<sup>8</sup>, Chuzhao Lei<sup>8</sup>, Yudong Cai<sup>8</sup>, Valentin A.

Bâlțeanu<sup>9</sup>, Diana Barfield<sup>10</sup>, Hailey Beard<sup>11</sup>, Beate Berger<sup>12</sup>, Thor Blichfeldt<sup>13</sup>, Geert Boink<sup>14</sup>, Sean Carolan<sup>15</sup>, Licia Colli<sup>16</sup>, Vlatka Cubric-Curik<sup>17</sup>, Louise Drew<sup>18</sup>, Cord Droegemueller<sup>19</sup>, Juha Kantanen<sup>20</sup>, Negar Khayatzadeh<sup>12</sup>, Namshin Kim<sup>21</sup>, Amparo Martínez<sup>22</sup>, Raffaele Mazza<sup>23</sup>, John McEwan<sup>11</sup>, Nadjat-Amina Ouchene-Kheliffi<sup>24</sup>, Filipe Pereira<sup>25</sup>, Anne da Silva<sup>26</sup>, Mojca Simčič<sup>27</sup>, Johann Sölkner<sup>12</sup>, Alison Sutherland<sup>28</sup>, Johannes Tigchelaar<sup>14</sup>, Econogene Consortium<sup>29</sup>, Paolo Ajmone-Marsan<sup>16</sup>, Daniel G. Bradley<sup>7</sup>, François Pompanon<sup>5</sup>, Gwenola Tosser-Klop<sup>6</sup> and **Johannes A. Lenstra**<sup>2\*</sup>

By its paternal transmission, Y-chromosomal haplotypes are sensitive markers of population history and male-mediated introgression. Previous studies have identified in domestic goats four major Y-chromosomal haplotypes Y1A, Y1B, Y2A and Y2B with a marked geographic differentiation and several regional variants. In this study we combine whole-genome sequences (WGSs) of 392 male goats from 64 modern breeds and 7 wild goat species generated by the Vargoa goats genome project, 136 entries in the Short-Read Archive and genotypes of diagnostic SNPs of several modern and ancient samples. We identified single-copy male-specific SNPs in four scaffolds, containing *SRY*, *ZFY*, *DBY* with *SSX3Y* and *UTY*, and *USP9Y* with *UMN2001*, respectively. Phylogenetic analyses indicated haplogroups corresponding to Y1B, Y2A and Y2B, respectively, but Y1A is split into Y1AA and Y1AB. All these haplogroups were detected in ancient DNA samples from southeast Europe. In different branches of the tree, the haplotypes from domestic goats, Iranian bezoars and Anatolian bezoars, respectively, have separate positions. Notably, the Anatolian haplotypes are near the roots of the tree or of subtrees, suggesting that the Iranian bezoars are closer to the wild ancestors of domestic sheep. Haplogroup distributions for 150 domestic breeds indicates on all three continents ancient paternal population bottlenecks, especially during the migration into northern Europe. In addition, sharing of haplogroups reveals male-mediated introgression of Boer goat into Uganda, Kenya, Tanzania, Malawi and Zimbabwe, and of European goats into the native Korean goat population. This study illustrates the power of the Y-chromosomal variation for the reconstructing the history of mammalian species with a wide geographic range.

## 24. Evolutionary history of Pachyceriform sheep

**Sarah H. D. Santos**<sup>1</sup>, Rhiannon M. Peery<sup>1</sup>, Joshua M. Miller<sup>1</sup>, Anh Dao<sup>1</sup>, Feng-Hua Lyu<sup>2</sup>, Xin Li<sup>3,4</sup>, Meng-Hua Li<sup>3</sup>, David W. Coltman<sup>1</sup>

<sup>1</sup> Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada

<sup>2</sup> College of Animal Science and Technology, China Agricultural University, Beijing 100193, China

<sup>3</sup> CAS Key Laboratory of Animal Ecology and Conservation Biology, Chinese Academy of Sciences (CAS), Beijing 100101, China

<sup>4</sup> University of Chinese Academy of Sciences (UCAS), Beijing, China

Species with rapid and recent radiation present interesting patterns of ancient hybridization. These events can leave detectable genomic footprints, such as introgression blocks. The advancement of whole-genome sequencing has improved our detection and understanding of these genealogical discordances. The evolutionary history of North American mountain sheep – including both subspecies of thinhorn (Dall - *Ovis dalli dalli*; Stone - *O. dalli stonei*) and bighorn (*O. canadensis*) - presents an interesting case study of a genealogically discordant evolutionary history. These species had a recent and rapid radiation, which could have enabled secondary contact between them, and subsequent exchange and incorporation of genomic blocks. Here we employed multiple whole-genome sequences of these species, together with snow (*O. nivicola*) and domestic sheep (*O. aries*) as outgroups, to study their phylogenomic history, potential introgression patterns and adaptive consequences. We observed a consistent monophyletic tree when considering all genomic fragments together, yet alternative topologies across many genome fragments. The conventional species tree was the most frequently observed, followed by phylogenies placing Stone as sister to bighorn, and Dall as sister to bighorn.

Genetic divergence between Stone and bighorn was less divergent than that between Dall and bighorn, and we observed more introgression signal between the former. We often observed more coat color genes within genomic fragments with introgression signal between Stone and bighorn. These results suggest that Stone might have originated from a complex series of events, and that it had posterior ancient secondary contact with bighorn.

## **25. Runs of homozygosity to unravel genetic changes associated with domestication and modern selection in goats**

Heidi Signer-Hasler<sup>1\*</sup>, Jan Henkel<sup>2</sup>, Erika Bangerter<sup>3</sup>, Zafer Bulut<sup>4</sup>, The VarGoats Consortium, Cord Drögemüller<sup>2</sup>, Tosso Leeb<sup>2+</sup> and **Christine Flury<sup>1+</sup>**

We analyzed WGS data from 217 modern goats and 9 wild Bezoar goats (*Capra aegagrus*). After quality control, 27,728,288 biallelic SNVs were used for the identification of runs of homozygosity (ROHs) and the detection of ROH-islands. Across all 226 caprine genomes of 11 populations, we found 389,629 SNVs in 344 ROH-islands containing 1220 annotated genes. ROH-islands of the modern breeds were contrasted with those in Bezoar goats. We identified two missense variants in *STC1*:p.Lys139Arg and *TSHR*:p.Ala239Thr, which might represent causative functional variants for domestication signatures. We hypothesize that the *TSHR*:p.Ala239Thr variant played a role in changing the seasonality of reproduction in modern domesticated goats. The exact functional significance of the *STC1*:p.Lys139Arg variant remains unclear and requires further investigation. Nonetheless, *STC1* might represent a new domestication gene affecting relevant traits such as body size and/or milk yield in goats. The information from ROH-islands using WGS data is suitable for the analysis of signatures of selection and allowed the detection of protein coding variants that may have conferred beneficial phenotypes during goat domestication.

## **26. Molecular characterization of the Montecristo feral goats in the Mediterranean context**

**Elisa Somenzi**(1), Gabriele Senczuk(2), Roberta Ciampolini(3), Matteo Cortellari(4), Ettore Randi(5), Gwenola Tosser-Klopp(6), Fabio Pilla(2), Paolo Ajmone-Marsan(1,7), Paola Crepaldi(4), Licia Colli(1,8) and the Italian Goat Consortium

The Montecristo wild goat is an endangered feral population occurring on the homonymous island in the Tuscan Archipelago since a long time. The origins of Montecristo goats are still debated, with authors dating their introduction either back to the Neolithic times or between the 6<sup>th</sup> and 13<sup>th</sup> century of the Common Era. To investigate the evolutionary history and genomic relationships of this population we assembled a 50K SNP dataset including 55 Mediterranean breeds and two nuclei of Montecristo goats sampled on the island and from an *ex situ* conservation project on the Italian mainland, respectively. Genetic diversity, gene flow, population structure and relationships were assessed through multiple approaches. The insular population scored the lowest values of both observed and expected heterozygosity, highlighting reduced genetic variation, while the *ex situ* nucleus showed a less severe reduction. Multivariate statistics, Neighbour-network and population structure analyses clearly separated the insular nucleus from all other breeds, but also the two Montecristo populations from each other. Treemix software analysis pinpointed possible genetic inputs received by the two Montecristo goat nuclei from different sources, while Runs Of Homozygosity (ROHs) indicated an ancient bottleneck/founder effect in the insular population and recent inbreeding in the *ex situ* one. Overall, our results suggest that Montecristo goats experienced several demographic fluctuations combined with admixture events over time, and highlighted a noticeable differentiation between the two *nuclei*. This



evidence can serve as a starting point to implement marker-assisted conservation plans for the endangered Montecristo feral goat.

- (1) DIANA Dipartimento di Scienze Animali, della Nutrizione e degli Alimenti, Università Cattolica del S. Cuore, Piacenza, Italy
- (2) Dipartimento di Agricoltura, Ambiente e Alimenti, University of Molise, 86100, Campobasso, Italy
- (3) Dipartimento di Scienze Veterinarie, University of Pisa, 56100 Pisa, Italy
- (4) Dipartimento di Scienze Agrarie ed Ambientali - Produzione, Territorio, Agroenergia, University of Milano, 20133, Milan, Italy
- (5) University of Aalborg, Faculty of Engineering and Science, Department of Chemistry and Bioscience, Aalborg, Denmark.
- (6) GenPhySE, Université de Toulouse, INRAE, ENVT, F-31326 Castanet-Tolosan, France
- (7) Centro di Ricerca Nutrigenomica e Proteomica-PRONUTRIGEN, Università Cattolica del Sacro Cuore, Piacenza, Italia
- (8) Centro di Ricerca sulla Biodiversità e sul DNA Antico-BioDNA, Università Cattolica del Sacro Cuore, Piacenza, Italia

## **27. Goat genomes from earliest managed herds in the Zagros Mountains, c. 8,000 BC**

**Kevin G. Daly**

Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland

Early goat genomes from the Zagros Mountain at the dawn of herding, c. 8,000 BC

The Zagros Mountains provides some of the earliest archaeological evidence of goat herding and demographic manipulation by early farmers, dating to c. 8,000 BC. The genetic consequences of the earliest transition to livestock population management by humans is still unknown, as is their relationship to present day goats. To explore this we present a combination of genomic and archaeological data from two Aceramic Neolithic Zagros sites. These data confirm that goat were both managed and genetically distinct from wild bezoar by this time, and that these early managed goats fall at the basal to domestic goat variation. We show contrasting matrilineal and patrilineal diversity, some traces of recent inbreeding, and an absence of reported selection signals. This study demonstrates the power of cross-disciplinary research, combining ancient genomic and faunal analysis to reveal the consequences of livestock herding at the dawn of agriculture, and points to the Zagros region as being a key genetic source for the domestic goat gene pool. Additionally, in a separate analysis we report preliminary results of the presence of an unreported lineage of *Capra* in the Taurus Mountains, c. 11,000 BC.

## **28. Genomic analyses reveal the influence of geographic origin and signatures of selection in Boer goats from different countries**

Gcabashe, Sithembisile<sup>1</sup>, Dzomba, Edgar F<sup>1</sup>, Muchadeyi, Farai C<sup>2</sup> and Hadebe, Khanyisile <sup>2\*</sup>

<sup>1</sup> Discipline of Genetics, School of Life Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, South Africa

<sup>2</sup> Agricultural Research Council, Biotechnology Platform, Private Bag X5, Onderstepoort, 0110, South Africa

Boer goats have become one of the leading meat-type breeds with worldwide distribution and utilization. The availability of Illumina SNP60K BeadChip and the collection of data through the AdaptMap project has allowed for comprehensive evaluation of goats worldwide. These studies did not include the South African Boer population despite the breed originating from the country. In this study, Boer goat genotypes of 366 animals from Australia (BOEAU; n = 61), New Zealand (BOENZ; n = 14), Tanzania (BOETZ; n = 4), South Africa (BOESA; n = 33), Switzerland (BOECH; n = 190), Uganda (BOEUG; n = 5), United States of America (BOEUS; n = 34) and Zimbabwe (BOEZW; n = 25) were used to investigate the genetic relationships among these populations. We implemented two complementary approaches of the iHS (Integrated haplotype score) and XP-EHH (Cross Population Extended Haplotype Homozygosity) to detect selection signatures within and across populations. BOEZW formed a cluster at PC2 (1.94%) supported by the optimal  $K$  at 15 (cv error = 0.582) and a higher  $F_{ST}$  (range 0.05 to 0.08). BOESA and BOEUS clustered together, showed similar genetic backgrounds and the lowest  $F_{ST}$  value ( $F_{ST} = 0.01$ ). BOECH had different genetic backgrounds. A total of 51 and 111 genomic regions revealed 136 and 161 candidate genes identified by iHS and XP-EHH, respectively. These genes are involved in pathways for metabolism, adaptation and immune response. The study findings highlight the effect of divergent selection on the Boer breed genome which will ultimately result in distinct Boer populations.

Keywords: *Boer, with-in breed relationship, signatures of selection*

## 29. Genomic estimation of inbreeding and effective population size in three Czech goat populations

<sup>1</sup>Vostry L., <sup>1,2</sup>Vostra-Vydrova H., <sup>3</sup>Cubric-Curik V., <sup>3</sup>Drzaic I., <sup>2</sup>Rychtarova J., <sup>3</sup>Curik, I.

<sup>1</sup>Czech University of Life Sciences, Kamycka 129, 16500 Prague, Czech Republic, <sup>2</sup>Institute of Animal Science, Pratelstvi 815, 10400 Prague, Czech Republic, <sup>3</sup>University of Zagreb, Faculty of Agriculture, Svetošimunska 25, Zagreb, Croatia.

Indigenous Brown Shorthair (BSG) and White Shorthair (WSG), together with the international Alpine goat (ALG), are the most important goat breeds in Czech Republic. In this study, we genotyped 91 individuals (32 BSG, 33 WSG and 26 Alpine goats) using the Illumina goat SNP50 BeadChip. To determine key conservation risk parameters, we estimated inbreeding ( $F_{ROH}$ ) and newly developed current and historical effective population size ( $Ne_{GONE}$ ) based on 38108 autosomal SNPs remaining after quality control (GC score 0.7, GT score 0.4, mind 0.9, geno and no null positions). Highest remote ( $F_{ROH>4Mb} = 0.076$ ) and close ( $F_{ROH>8Mb} = 0.052$ ) inbreeding was observed in BSG, while intermediate remote ( $F_{ROH>4Mb} = 0.046$ ) and close ( $F_{ROH>8Mb} = 0.028$ ) inbreeding was observed in ALG. On the other hand, very low remote ( $F_{ROH>4Mb} = 0.019$ ) and close ( $F_{ROH>8Mb} = 0.014$ ) inbreeding was observed in WSG. Thus, the estimated current  $Ne_{GONE-0}$  was 130, but even higher than 12 generations ago when it was 108 ( $Ne_{GONE-12}$ ). ALG had the smallest current  $Ne_{GONE-0}$  (107), which was much higher 12 generations ago ( $Ne_{GONE-12} = 237$ ). Current effective population size was relatively large in WSG ( $Ne_{GONE-0} = 244$ ), but with a warning rapid decline in the last 12 generations ( $Ne_{GONE-12} = 698$ ). The observed estimates will be used for conservation and breeding management of these three breeds. The study was supported by projects QK1910156 and in part (man-hours) ANAGRAMS -IP-2018-01-8708.

## 30. Whole-genome sequencing reveals selection signals among Chinese, Pakistani and Nepalese goats

Yefang Li<sup>1,2†</sup>, Shen Song<sup>1,2,3†</sup>, Xuexue Liu<sup>1,2</sup>, Yanli Zhang<sup>1,2</sup>, Dandan Wang<sup>1,2</sup>, Xiaohong He<sup>1,2</sup>,

Qianjun Zhao<sup>1,2</sup>, Yabin Pu<sup>1,2</sup>, Weijun Guan<sup>1,2</sup>, Guangxin E<sup>4</sup>, Chander Kumer<sup>5</sup>, Yuehui Ma<sup>1,2\*</sup>, Lin Jiang<sup>1,2\*</sup>

1 Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, P. R. China

2 CAAS-ILRI Joint Laboratory on Livestock and Forage Genetic Resources, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing, 100193, P. R. China

3 State Key Laboratory of Cardiovascular Disease Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100037, P. R. China

4 College of Animal Science and Technology, Chongqing Key Laboratory of Forage & Herbivore, Chongqing Engineering Research Centre for Herbivores Resource Protection and Utilization, Southwest University, Chongqing, P. R. China

5 Department of Animal Breeding and Genetics, Faculty of Animal Husbandry and Veterinary Sciences, Sindh Agriculture University, Tando Jam, Pakistan

Domestic goats have an excellent adaptability to harsh environments and various traits of economical importance. Whole-genome sequencing is a powerful platform to rapidly gain novel insights into the identification of genetic basis underlying specific traits in goats. Here we collected whole-genome sequences of 115 domestic goats representing 15 breeds from China, Nepal and Pakistan. Comparative genomic analyses among goats varying in phenotypic traits, detected a set of candidate genes that might be associated with heat adaptation, milk production and cashmere growth. We also identified that a 505-bp indel variant at the fibroblast growth factor 5 (FGF5) gene locus appeared to be strongly associated with cashmere growth. Functional validation showed that the insertion variant may serve as an enhancer for transcription factor binding, resulting in increased transcription of the upstream FGF5 gene in non-cashmere goats. Our study provides useful information of the genetic diversity and traits exploration for Chinese, Pakistani and Nepalese goats, and reveals that an indel mutation in the FGF5 gene could potentially serve as a molecular marker for cashmere growth in cashmere goat breeding.

Key words: whole-genome sequencing, heat adaptation, milk production, cashmere growth, FGF5, indel

### **31. Run of Homozygosity in the main historical lines of Spanish Merino Sheep**

**Gabriel Anaya**(1) - Antonio Granero - (2) - Sebastián Demyda-Peyrás (3) - Nora Laseca (1) - María Jesús Alcalde (4) - Molina Alcalá (1)

Departamento de Genética, University of Córdoba, Córdoba, Spain (1)

Secretario Técnico del Libro Genealógico, Asociación Nacional de Criadores de Ganado Merino, Madrid, Spain (2)

Departamento de Producción Animal, National University of La Plata, Argentina; CONICET CCT La Plata(3)

Ciencias Agroforestales (ETSIAM), University of Sevilla, Sevilla, Spain (4)

The Spanish Merino is one of the most important sheep breeds over the world due to its economic and historical impact. It was originally formed by several Spanish ancient lines, but nowadays it is clear its profound influence in numerous breeds. Most of these historical lines have disappeared in Spain but several of them are still bred in closed herds during the last 200 years. Among them we can mention the Maesso, Egea, Granda, Lopez-Montenegro, Hidalgo and Donoso lines. The aim of this study was

to analyze Runs of Homozygosity (ROH) patterns on 6 historical lines of the Spanish White Merino sheeps to determine genomic differences among lines. To this end, 344 animals were genotyped with the Axiom™Ovine Genotyping Array 50K chip. ROH were determined using a hidden Markov model-based procedure including 7 different classes, with a minimum length of 500kb only in the autosomes. Thereafter, a ROH-incidence p-value was determined per marker and line. We identified clear differences between lines across the genome, suggesting the existence of different genetic footprints and selective processes in each historical group of animals. These results could help to achieve a better understanding of the origin and evolution of the historical lines of Spanish Merino, but also as start-point to determine unique genomic characteristics that can explain their distinctive productive aptitudes.