VETBIONET

Veterinary Biocontained facility Network for excellence in animal infectiology research and experimentation

Deliverable D8.4

Development of high throughput class I MHC genotyping methods for sheep and cattle

Due date of deliverable: M60

Actual submission date: M72

Start date of the project: March 1st, 2017 **Duration:** 72 months

Organisation name of lead contractor: MRI

Revision: V1

Dissemination level	
Public	**X**
Confidential, only for members of the consortium (including Commission Services)	
Classified, as referred to in Commission Decision 2001/844/EC	

Table of contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°731014

1. Summary

Objectives: The protein products of the major histocompatibility complex or MHC class I genes are key regulators of both adaptive and innate arms of the immune response by presenting pathogen or vaccine antigens for recognition by T lymphocytes or by acting as ligands for natural killer cell receptors. The MHC class I genes are amongst the most variable genes in almost all vertebrate populations with thousands of variants or alleles identified at multiple genes in the highly studied human MHC. We have only begun to unravel the complexity of these important genes in livestock populations. Deliverable D8.4 therefore seeks to develop next generation sequencing technologies that targets genetic variation in the MHC class I genes of two important ruminant livestock species in European agriculture, sheep, and cattle. Defining the complexity of MHC class I diversity in sheep and cattle will improve our ability to identify protective pathogen antigens, design better vaccines, understand the role of MHC diversity in variations in response to infection or vaccination, identify and conserve unique MHC allelic diversity that may be present in the range of European livestock breeds.

Rationale: The power of 2nd and 3rd generation nucleotide sequencing technologies allows vast amounts of sequence data to be generated in a relatively short time. We sought to develop methods that will allow the sequencing of the highly variable region of functional MHC class I genes in large numbers of sheep and cattle. This will identify the combinations of genes or haplotypes inherited together from each parent, provide an indication of the level of expression of each gene within each haplotype and assign each sequence to either classical, non-classical or unknown function. To achieve this, we chose to use the Illumina MiSeq platform with paired end multiplex sequencing of barcoded amplicons and cDNA template for the PCR reactions. This allowed library sequencing of hundreds of animals simultaneously with subsequent bioinformatic pipelines developed or refined to deconvolute the data based on unique barcode combinations. Haplotype were defined based on identification of the same combinations of sequences in different animals and read frequency data for each sequence provided an indication of transcript expression. A range of peer reviewed publications have been produced, are under review or in preparation.

Teams involved: MRI, RI, TPI, INRAE.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°731014

2. Introduction

The major histocompatibility complex or MHC is a genetic region present in all jawed vertebrates. Originally identified as a region of the genome that determined the success or failure of tissue transplantation or histocompatibility, it was later discovered to include families of MHC class I and class II genes that are central to immune surveillance for pathogen infection or malignant transformation. A feature of the MHC is the high level of allelic diversity at many of the MHC class I and class II genes. This diversity is mostly associated with the region of the gene that codes for the groove where the antigenic peptide interacts with the MHC molecule. Allelic diversity is associated with changes in the range of peptide that are presented to T cells hence variation in immune responses. For this reason, conserved regions of the second and third exons of MHC class I genes are targeted for primer design. Up to three sets off overlapping barcoded primers were designed to target the MHC class I genes. This reduces the impact of primer bias and allowed more confident estimations of transcript frequency. To reduce the complexity of sequence analysis, we also chose to target only transcribed class I genes. RNA was prepared from cells isolated from peripheral blood samples from large numbers of cattle and sheep. This was used to prepare cDNA to be used at templated for PCR reactions. Sequence libraries were constructed by pooling samples from each animal. Amplification conditions, library construction and paired-end sequencing using Illumina MiSeq v3 chemistry and the bioinformatics pipeline are as described in Vasoya et al., 2021.

3. Results

3.1 Developing a high throughput MHC class I genotyping method for cattle

Initially, two MHC genotyping planning and progress meetings took place between WP partners from the Roslin Institute (RI) and Moredun Research Institutes (MRI). Training in practical aspects of high-throughput MHC class I genotyping technologies took place at the Roslin Institute while an MHC class I sequence analysis bioinformatics pipeline developed was developed at RI.

The first cattle MHC class I sequence and haplotype analysis using two sets of barcoded primers was completed at RI using the bioinformatics pipeline. The details of this method, subsequent refinements, and expansion of the deliverable into the highly variable MHC class II gene DRB3 to provide full class I and class II haplotypes are detailed in the publication listed

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°731014

below. This encompassed Milestone 36 (Prototype class I MHC genotyping methods validated in homozygous animals) and Deliverable 8.4 for cattle.

- Vasoya D, Oliveira PS, Muriel LA, Tzelos T, Vrettou C, Morrison WI, de Miranda Santos IKF, Connelley T. High throughput analysis of MHC-I and MHC-DR diversity of Brazilian cattle populations. HLA. 2021 Aug;98(2):93-113. doi: 10.1111/tan.14339. Epub 2021 Jun 17. PMID: 34102036.

The MHC analysis methods using the MiSeq platform has been shared with EU partners at the University of Perugia, Italy, for analysis of Italian cattle breeds and was presented at the 2nd annual meeting in Madrid, May 2019.

In addition, a PCR method to generate near-full length bovine MHC class I amplicons for sequencing with long read PacBio technology is under development. This will increase the amount of MHC class I sequence data available for each transcript, improve the assignment of sequences to individual gene loci and allow official nomenclature to be assigned to each sequence. The broad application of this MHC class I amplicon and MiSeq analysis methodology has also been tested for goat, pig and horse.

3.2 Developing a high throughput MHC class I genotyping methods for sheep

The MHC class I amplicon and MiSeq analysis methodology and bioinformatics pipeline developed for cattle was adapted for sheep. A third set of amplification primers was included for comparison with the initial two primer sets used in cattle and to provide more robust individual gene expression data. From 100 Scottish Blackface sheep, using high-throughput amplicon sequencing with the three sets of barcoded primers, we identified 132 MHC class I transcripts within 38 haplotypes. Haplotypes were identified with between two and five MHC class I genes. One or two highly transcribed transcripts dominate each haplotype indicative of two highly polymorphic, classical MHC class I genes. Additional medium, low, and very low expressed transcripts, indicative of non-classical genes were identified. This analysis included cDNA from MHC homozygous animals developed at MRI and INRA to validate the methodologies. This completed Milestone 36 for sheep (Prototype class I MHC genotyping ghthrouput gh methods validated in homozygous animals). A publication describing the methodology and data analysis is currently in preparation.

- Large scale transcriptional analysis of MHC class I haplotype diversity in sheep. Deepali Vesoya, Timothy Connelley, Thomas Tzelos, Helen Todd and, Keith T. Ballingall, *In preparation.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°731014

Other publications that acknowledge the contribution of VetBioNet funding

- Contemporary selection on MHC genes in a free-living ruminant population (2022). Huang, W., Dicks, K.L., Hadfield, J.D., Johnston, S.E., Ballingall, K.T. & Pemberton, J.M. Ecology Letters, 25, 828–838. https: //doi.org /10.1111/e l e.13957.

- Development of a potential yeast-based vaccine platform for *Theileria parva* infection in cattle (2021) Shan Goh, Jeannine Kolakowski, Angela Holder, Mark Pfuhl, Daniel Ngugi, Keith Ballingall, Kata Tombacz and Dirk Werling. Front. Immunol, 12:674484. doi: 10.3389/fimmu.2021.674484.

- MHC class IIa haplotypes derived by high throughput SNP screening in an isolated sheep population (2021) Kara L Dicks, Josephine M Pemberton, Keith T Ballingall, Susan E Johnston. G3, Genes|Genomes|Genetics, 11, jkab200, https://doi.org/10.1093/g3journal/jkab200.

- Characterisation of Major histocompatibility complex class IIa haplotypes in an island sheep population (2019). Kara L Dicks, Josephine M Pemberton and Keith T. Ballingall. *Immunogenetics*, 71, 383-393. https://doi.org/10.1007/s00251-019-01109-w.

- Allelic nomenclature for the duplicated MHC class II *DQ* genes in sheep (2019) Keith T. Ballingall, Kara Dicks, Panoraia Kyriazopoulou, Lynne Herrmann-Hoesing. *Immunogenetics*, 71, 347-351, DOI: 10.1007/s00251-018-1096-9.

- Immunological homeostasis at the ovine placenta may reflect the degree of maternal foetal interaction (2018) Sean R. Wattegedera, Laura Doull, Mariya Goncheva, Nick Wheelhouse, Donna Watson, Julian Pearce, Julio Benavides, Javier Palarea-Albaladejo, Colin McInnes, Keith Ballingall, Gary Entrican: *Frontiers in Immunology*, 9:3025.doi: 10.3389/fimmu.2018.03025.

- Comparative MHC Nomenclature: report from the ISAG/IUIS-VIC committee (2018) Keith T. Ballingall, Ronald E. Bontrop, Shirley A. Ellis, Unni Grimholt, John A. Hammond, Chak-Sum Ho, Jim Kaufman, Lorna J. Kennedy, Giuseppe Maccari, Donald Miller, James Robinson, Steven G.E. Marsh (2018) *Immunogenetics,* 70, 625-632. doi.org/10.1007/s00251-018-1073-3.

- IPD-MHC: nomenclature requirements for the non-human major histocompatibility complex in the next-generation sequencing era. (2018) Giuseppe Maccari, James Robinson, Ronald E. Bontrop, Nel Otting, Natasja G. de Groot, Chak-Sum Ho, Keith T.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°731014

Ballingall, Steven G.E. Marsh and John A. Hammond *Immunogenetics,* 70:619–623. doi.org/10.1007/s00251-018-1072-4.

Interactions with other WP: Planning discussions for MHC class genotyping of Spanish Colmenareña sheep took place in collaboration with INIA, WP7 at the 2nd annual meeting. Genomic DNA was subsequently extracted from 47 sheep of the Colmenareña breed, which is commonly used in infectious disease studies at INIA, to provide an indication of the level of MHC diversity in the flock. DNA was transferred to MRI and each animal genotyped at the highly polymorphic MHC class II *DRB1* locus using a sequence-based genotyping approach.

4. Conclusions

We have successfully developed high throughput MHC class I sequencing technologies for sheep and cattle. This has greatly improved our understanding of MHC haplotype structure and allelic diversity in both these key agricultural species. We have expanded the initial focus of the deliverable by including MHC class II loci and now have the tools to rapidly define complex MHC haplotypes for sheep and cattle involved in experimental infectious disease studies.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°731014