



Developing the Functional Annotation of the Sheep Genome

Brenda Murdoch¹, Stephen White^{2,3}, Michelle Mousel², Alisha Massa³, Kim Worley⁴, Alan L. Archibald⁵, Emily Clark⁵, Brian Dalrymple⁶, James Kijas⁷, Shannon Clarke⁸, Rudiger Brauning⁸, Timothy P.L. Smith⁹, Tracy Hadfield¹⁰, Noelle Cockett¹⁰

¹University of Idaho, ²USDA, ARS U.S. Animal Disease Research Unit, ³Washington State University, ⁴Baylor College of Medicine-Human Genome Sequencing Center, ⁵The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, ⁶University of Western Australia, ⁷CSIRO Agricultural Flagship, ⁸AgResearch, ⁹USDA, ARS U.S. Meat Animal Research Center, ¹⁰Utah State University



Baylor College of Medicine®

GIVING LIFE TO POSSIBLE



THE UNIVERSITY OF
WESTERN
AUSTRALIA



United States
Department of
Agriculture

National Institute
of Food and
Agriculture

Award number USDA-NIFA-2017-67016-26301



New Reference Genome (PacBio assembly)

- 7 year-old U.S. Rambouillet ewe
- FAANG - Annotations from the same animal as the new reference genome
- 100 tissues collected in April, 2016



Rambouillet ewes Benz 2616 (right) and her full-sib sister (left).

FAANG Sample Coordination & Collection: Noelle Cockett (USU), Stephen White (ARS, WSU), Kim Worley (BCM), James Kijas (CSIRO), Brian Dalrymple (UWA), Brian Sayre (VSU), Tim Smith (MARC), Mike Heaton (MARC), Michelle Mousel (ARS, ADRU), Alisha Massa (WSU), Brenda Murdoch (U Idaho), Tracy Hadfield (USU), Tom Baldwin (USU), Rusty Stott (USU), Arnaud Van Wettere (USU), Gordon Hullinger (USU), Holly Mason (USU), Jaqualine LaRose (USU), Dave Forrester (USU), Codie Durfee (ARS, ADRU), Corey Wareham (USU), Maria Herndon (WSU), Caylee Birge (ARS, ADRU), Sarah Behunin (USU), Kara Thornton (USU), Rachael Christianson (USU), Nicole Ineck (USU), Angie Robinson (USU), Dallin Wengert (USU), Kerry Rood (USU), Erica Moscoso (USU), Rickie Warr (USU), Dustin Kinney (USU), Abbey Benninghoff (USU), Sumira Phatak (USU), Kevin Contreras (USU), Braden Abercrombie (USU), Misha Regouski (USU).

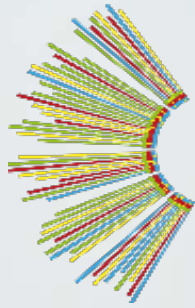
Tissues Collected

Tier 1 Tissues (n = 12)	Tier 2 Tissues (n = 47)		Tier 3 Tissues (n = 41)	
Cerebral cortex (frontal cortex)	Spleen	Gall bladder	CD14 positive cells (monocytes)	Skeletal muscle - gluteus medius
Cerebellum	Adipose - subcutaneous	Adipose - kidney capsule	CD4 positive cells (T cells)	Skeletal muscle – infraspinatus
Hypothalamus	Rumen – ventral	Kidney medulla	CD8 positive cells (T cells)	Skeletal muscle – intercostals
Abomasum	Medulla oblongata (brainstem)	Urinary bladder	WC1 positive cells ($\gamma\delta$ T cells)	Skeletal muscle - psoas major
Ileum	Hippocampus	Adrenal medulla	Gall bladder	Pons
Liver	Olfactory bulb	Uterus – caruncle	Peripheral nerve - sciatic	Pulmonary vein
Kidney cortex	Tongue	Mammary gland	Tonsil – palatine	Salivary gland - mandibular
Ovary	Lymph node – mandibular	Left ventricle	Adipose - mesenteric	Semilunar valve (aortic)
Blood leukocytes	Thyroid gland	Right atrium/auricle	Aorta – thoracic	Mesencephalon (midbrain)
Lung	Skin - non-haired	Right ventricle	Atrioventricular valve - left)	Oviduct
Skeletal muscle - biceps femoris	Rumen – atrium	Left atrium/auricle	Atrioventricular valve - right)	Parathyroid gland
Adrenal cortex	Rumen – dorsal	Trachea	Bronchi - left and right)	Pineal gland
	Reticulum	Skeletal muscle – diaphragm	Cartilage - epiglottis	Soft palate
	Omasum	Esophagus – thoracic	Cartilage - sternal	Tendon
	Pylorus – abomasum	Alveolar macrophages	Caudal vena cava	Thalamus
	Duodenum	Bone marrow	Cervix	Trigeminal ganglia
	Jejunum	Haired skin - dorsal lumbar	Corpus callosum	Ureter
	Cecum	Skeletal muscle - longissimus dorsi	Retina	Urethra
	Spiral colon	Skeletal muscle - semimembranosis	Hard palate	Uterus - intercaruncular
	Descending colon	Skeletal muscle - supraspinatus	Lymph node - mediastinal	Vagina
	Rectum	Spinal cord - cervical	Lymph node - prescapular	
	Pancreas	Peripheral nerve - brachial plexus		
	Lymph node – mesenteric	Adenohypophysis (pituitary gland)		
	Peyer's patch			



Goals

1. Deliver enhanced functional annotation of the ovine genome.
2. Facilitate the understanding of gene regulation in sheep by generating and distributing curated transcriptome data.
3. Deliver tools and resources for studying the ovine genome through online databases, including those supported by the international FAANG Consortium.



FAANG

Functional Annotation of Animal Genomes

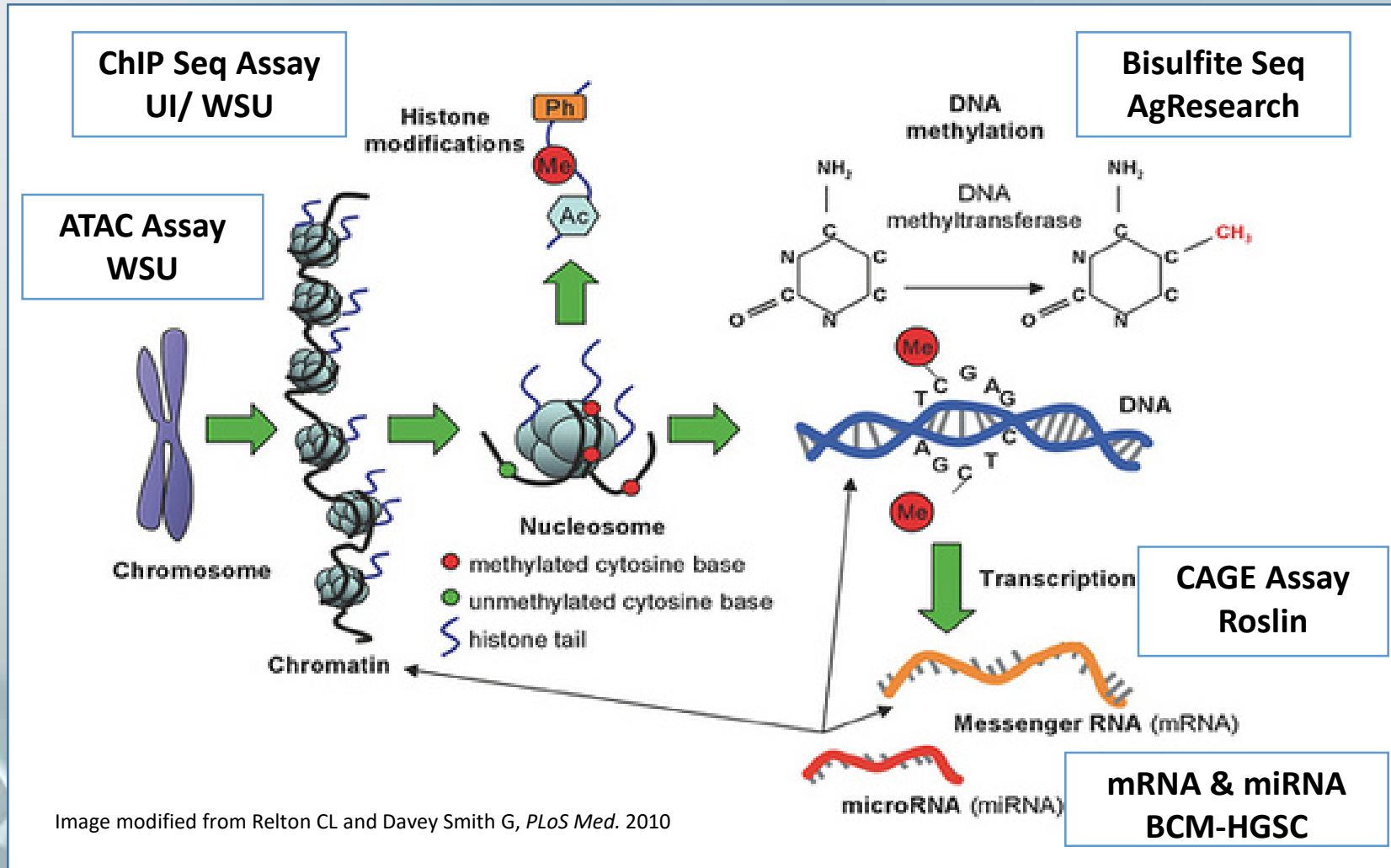
Requirement

- High quality reference sequence
- Support of the research community
- Critical mass of key investigators
- Selection of specific animal
- Core set of tissue samples
- Storing of tissues
- Sharing of tissues
- Core set of assays
- Data management & quality control
- Data publicly available & archived
- Annotation of the genome assembly

Ovine FAANG Project

- https://www.ncbi.nlm.nih.gov/assembly/GCA_002742125.1
- International Sheep Genomics Consortium
- 1 PD, 6 Co-PD, 3 Collaborators
- Benz 2616
- 100 tissues
- USU and USDA, ARS U.S. MARC
- AnGen posting in May, 2016
- FAANG core
- EMBL-EBI DCC and DAC
- EMBL-EBI, NCBI
- Completed within this project

Image of Transcriptional Assay



Project assays for each PD

Assay	Number of Tissues	PD	Institute
Transcriptome			
○ mRNA Sequence	Tier 1, 2 & 3 (94) ¹	K. Worley	BCM-HGSC
○ microRNA Sequence	Tier 1, 2 & 3 (94) ¹		
○ PacBio Iso-Seq	Tier 1 (12) ²		
Transcriptome		A. Archibald	Roslin
○ CAGE	All (100)	E. Clark	Institute
Histone Modification ChIP Seq			
○ H3K4me3	Tier 1 & 2 (59) ³	B. Murdoch	U I
○ H3K27me3	Tier 1 & 2 (59) ³	A. Massa	WSU
○ H3K4me1	Tier 1 & 2 (59) ³	S. White	WSU
○ H3K27ac	Tier 1 & 2 (59) ³		
Chromatin Accessibility		S. White	WSU
○ ATAC-seq	All (100)	M. Mousel	USDA, ARS
DNA methylation			
○ WG Bisulfite sequencing	FAANG core (22) & GI tract (6)	S. Clarke	AgResearch
○ epiGBS (RR Bisulfite sequencing)	All (100) ⁴		

¹ Sequencing of 59 samples is funded through the 2013 NIFA Sheep Genomes Database project.

² Sequencing funded through the 2013 NIFA Sheep Genomes Database project.

³ Partial fund for assay from UI and sequencing by USDA ARS.

⁴ Sequencing funded through matching NZ funds.



Project timeline

Year 1: April 1, 2017

- Perform transcriptome sequencing
- Perform transcript regulatory assays
- Data to DCC

Year 2: 2018

- Perform transcriptome sequencing
- Perform transcript regulatory assays
- DNA methylation
- Data to DCC
- Publication

Year 3, 2019

- Complete assays
- Data DCC
- Publication

