



HAL
open science

Samples collection and assays-by-sequence provided to the scientific community

Hervé Acloque

► **To cite this version:**

Hervé Acloque. Samples collection and assays-by-sequence provided to the scientific community. Final conference of the GENE-SWitCH project, EAAP; INRAE; INRAE-T, Nov 2023, Brussels, Belgium. hal-04498288

HAL Id: hal-04498288

<https://hal.inrae.fr/hal-04498288>

Submitted on 11 Mar 2024

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

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



GENE-SWitCH

The regulatory GENomE of SWine and CHicken:
functional annotation during development

*Samples collection and assays-by-sequence
provided to the scientific community*

Hervé Acloque - INRAE

November 6, 2023

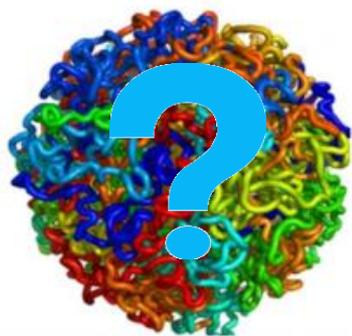
GENE-SWitCH Final Conference



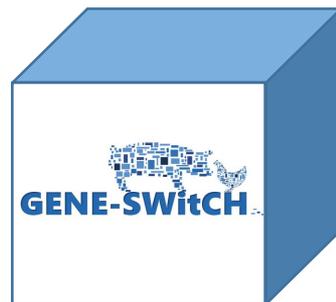
This project has received funding from the European Union's **Horizon 2020**
Research and Innovation Programme under the grant agreement n° 817998

@GeneSwitch - www.gene-switch.eu

The regulatory GENomE of SWine and CHicken: functional annotation during development

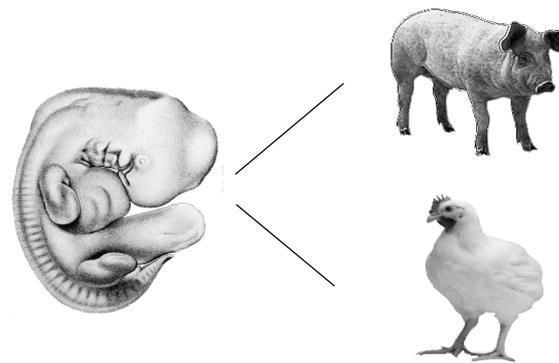


Genome and epigenome from chicken and pig

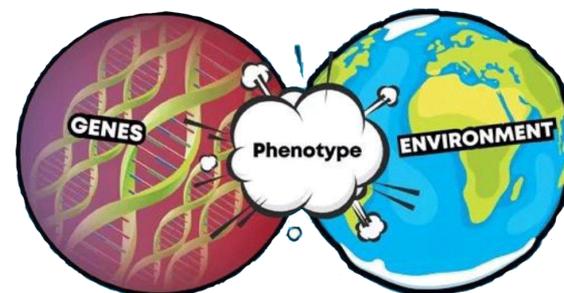


Functional Annotation

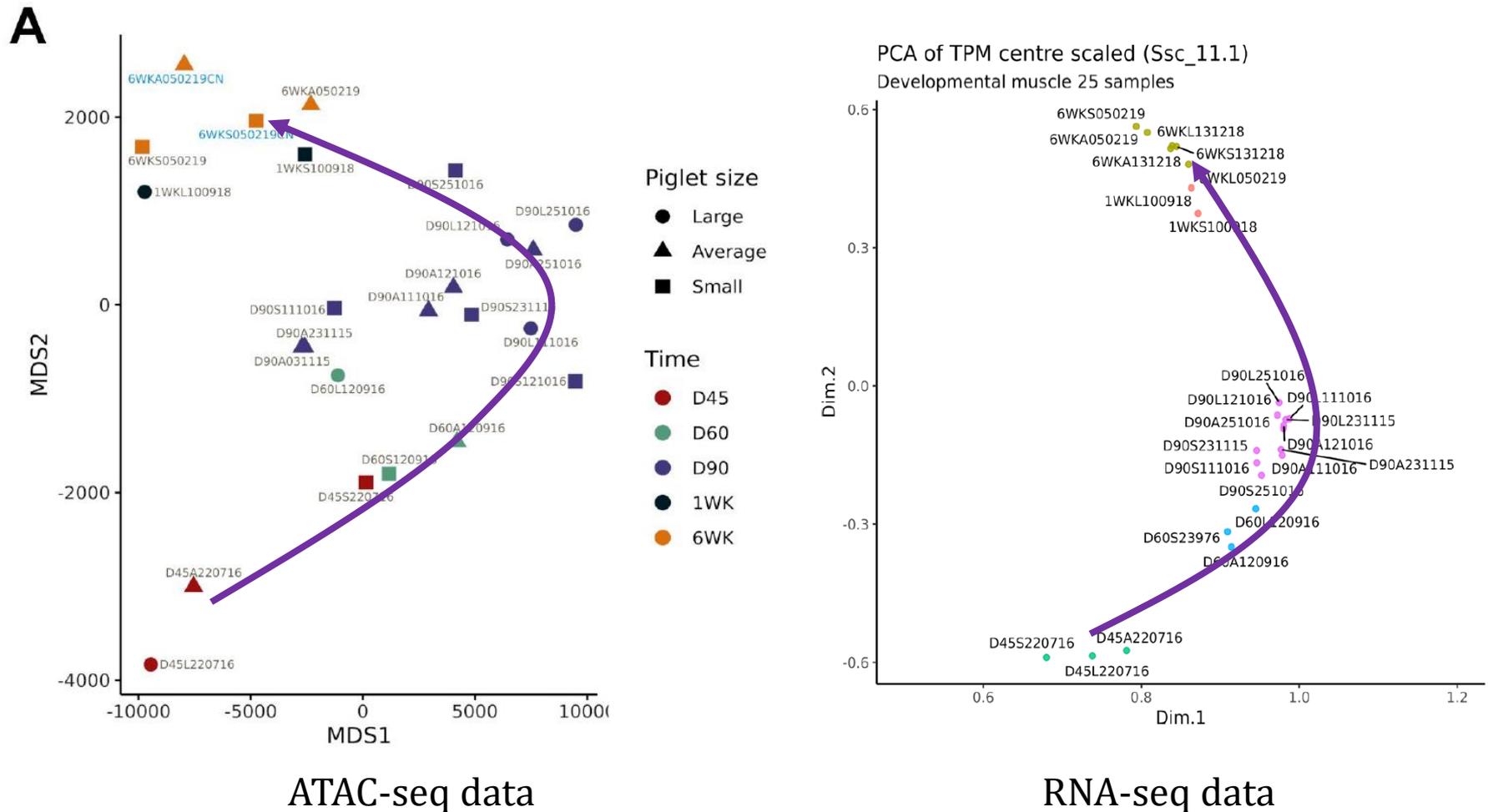
Key features to build organisms



Key features to build phenotypes



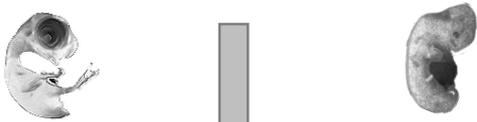
A dynamic of regulatory elements exist in the skeletal muscle during pig fetal and adult development



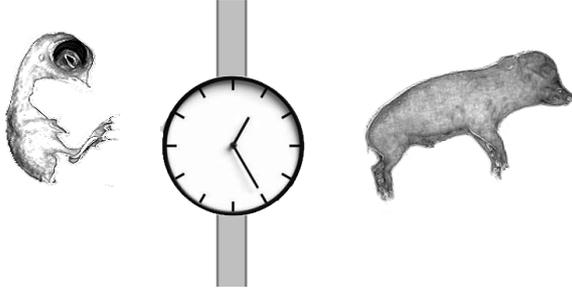
Functional annotation of 7 tissues during development

➤ 3 developmental stages:

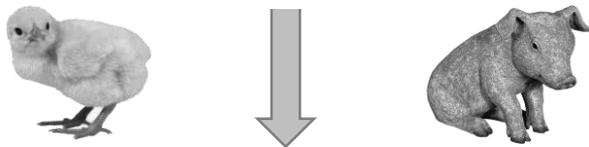
Early organogenesis



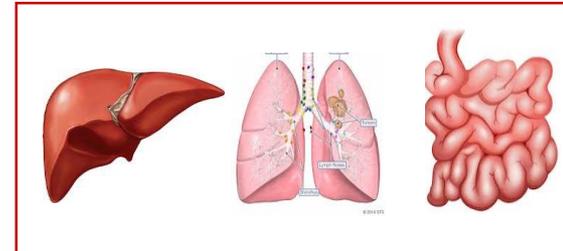
Late organogenesis



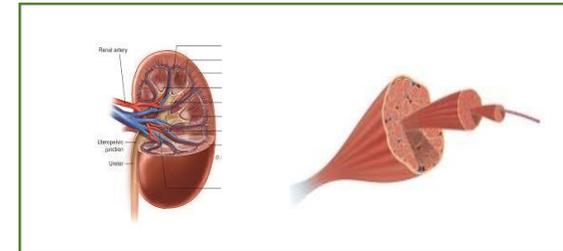
Newborns



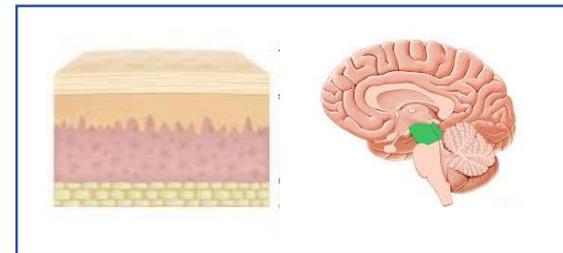
➤ 7 Tissues: liver, skeletal muscle, small intestine, cerebellum, dorsal epidermis, lung and kidney



Endoderm



Mesoderm



Ectoderm

Sampling and biobanking of tissues





	Fetus 30D	Fetus 70D	Piglet	Embryo 8D	Embryo 15D	Chick
	Large White			Advanced Intercross Line (AIL)		
	2 pools of males and 2 pools of females	2 males and 2 females	2 males and 2 females	2 pools of males and 2 pools of females	2 males and 2 females	2 males and 2 females

Biobanking of 6 additional tissues: **gonads, cerebral cortex, large intestine, heart, stomach, spleen**

Metadata on individuals, tissues and pool of tissues are available on the FAANG Data Portal and the BioSamples database.

All the protocols for tissue sampling and DNA/RNA extraction are also publicly available on the FAANG Data Portal and the BioSamples database.

Sampling and biobanking of tissues

[SSC_INRAE_GS_WP1_FT_30dpf_1](#)

SAMEA7628703

Updated on: 22-10-2020 08:53

birth date 2019-11-21 (YYYY-MM-DD) birth location Nouzilly birth location latitude 47.32432 (decimal degree)

birth location longitude 0.46596 (decimal degree) birth weight 1.89 (grams) breed Large white

father id FR22XPX201805117 health status normal insemination date 2019-10-22 (YYYY-MM-DD)

material organism mother id FR17MAG201711684 mother parity 5.0 mother weight 252.0 (kilograms)

organism *Sus scrofa* pregnancy length 30.0 (day) project FAANG

sample description Pig_large_white_fetus_30dpf sample name SSC_INRAE_GS_WP1_FT_30dpf_1

secondary project GENE-SWitCH sex Male

submission description Samples used for GENE-SWitCH WorkPackage 1. It includes fetus of 30d, 70d and one day old piglets.

submission title INRAE-FAANG-ACLOQUE-Pig-200915 derived from (reverse) SAMEA7628751

derived from (reverse) SAMEA7628787 derived from (reverse) SAMEA7628823

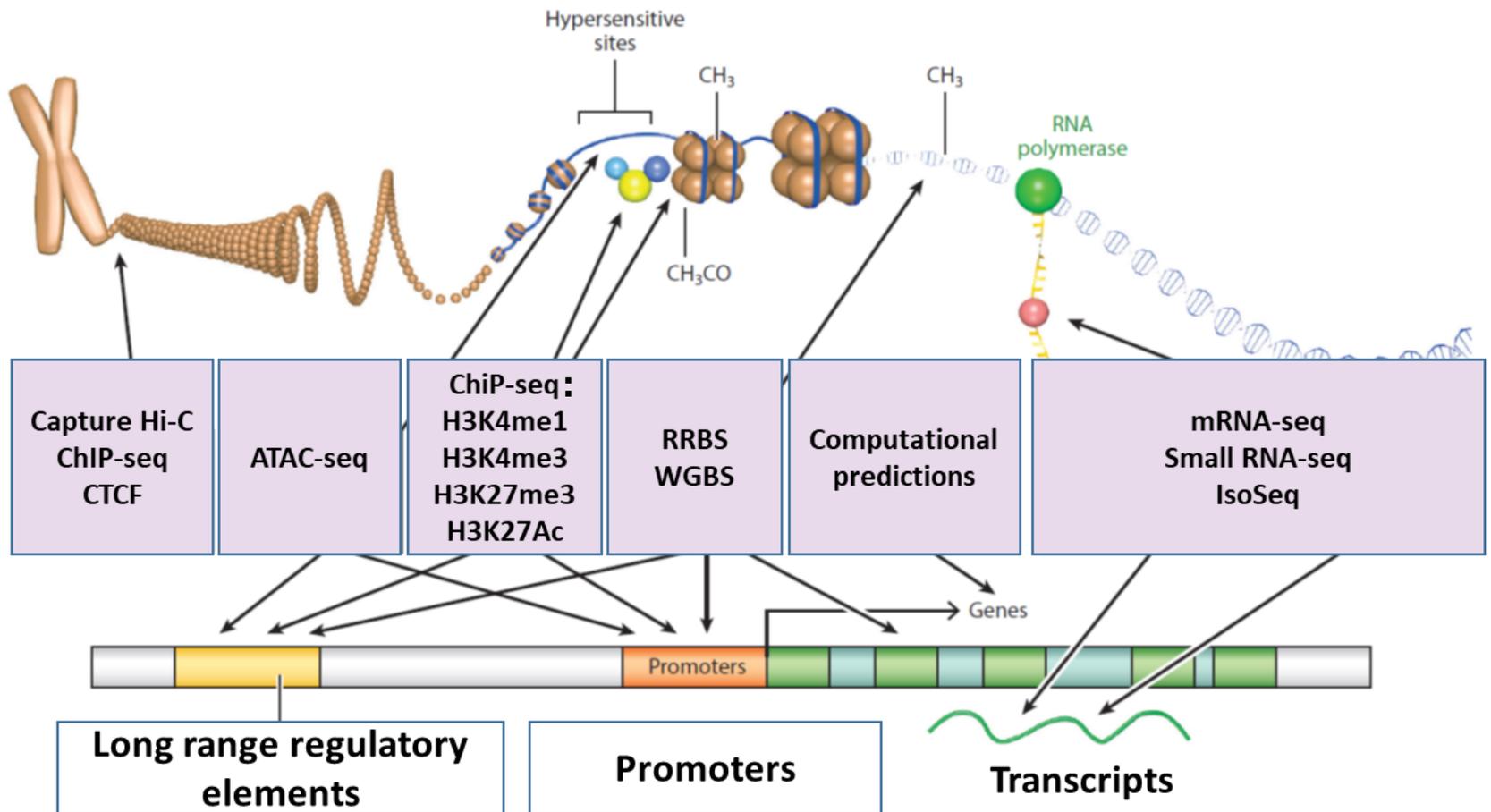
derived from (reverse) SAMEA7628859 derived from (reverse) SAMEA7628895

derived from (reverse) SAMEA7628931 derived from (reverse) SAMEA7628967

derived from (reverse) SAMEA7629059 derived from (reverse) SAMEA7629095

derived from (reverse) SAMEA7629131 derived from (reverse) SAMEA7629167

A set of complementary molecular assays to unravel regulatory developmental switches



Assay		Aim	Accession	Libraries
ATAC-seq		Chromatin accessibility	PRJEB44468	84
			PRJEB45945	84
ChIP-seq	H3K27Ac	Chromatin conformation Map regulatory elements (promoters, enhancers and repressors) Structural annotations	On-going	
	H3K4me1			
	H3K4me3			
	H3K27me3			
	CTCF			
RNA-seq	totRNA-seq (polyA+)	TSS and TTS identification Transcripts structures and categories Transcripts numbers	PRJEB41970	84
			PRJEB42025	84
	smallRNA-seq		PRJEB42001	84
			PRJEB42041	84
	Iso-seq		PRJEB50963	21
		PRJEB48060	21	
DNA methylation	RRBS	Functional Elements Active and inactive promoters Active and inactive chromatin	PRJEB41822	63
			PRJEB41829	63
	WGBS		PRJEB42772	21
			PRJEB42775	21
Promoter Capture Hi-C (liver and muscle only)		Interaction maps Enhancer-promoter interactions	PRJEB44486	12
			PRJEB53986	12

Pipelines used for the different assays

Assay		Pipelines
ATAC-seq		nf-core ATACseq : https://nf-co.re/atacseq
ChIP-seq		nf-core ChIPseq : https://nf-co.re/chipseq
RNA-seq	totRNAseq	dedicated pipeline TAGADA (Kurylo et al. 2023) : https://github.com/FAANG/analysis-TAGADA
	smallRNAseq	nf-core smRNAseq : https://nf-co.re/smrnaseq
	Iso-seq	Dedicated pipeline (Guizard et al. 2023): https://github.com/FAANG/analysis-iseq nf-core/iseq
DNA methylation	RRBS	Dedicated pipeline adapted from nf-core/methylseq https://github.com/FAANG/GSM-pipeline
	WGBS	
Promoter Capture Hi-C		under development

Achievements: ATAC-seq

Assay	Aim	Accession	Libraries
ATAC-seq	Chromatin accessibility	PRJEB44468	84
		PRJEB45945	84

	Raw reads 150PE (median by assay)	Uniquely mapped PE reads (median by assay)	% MT reads (mean by assay)	Number of peaks (median per tissue)
Pig	84,710,247	57,104,004	2,33%	68,490
Chicken	92,284,762	59,681,384	4,88%	80,306

Protocol:

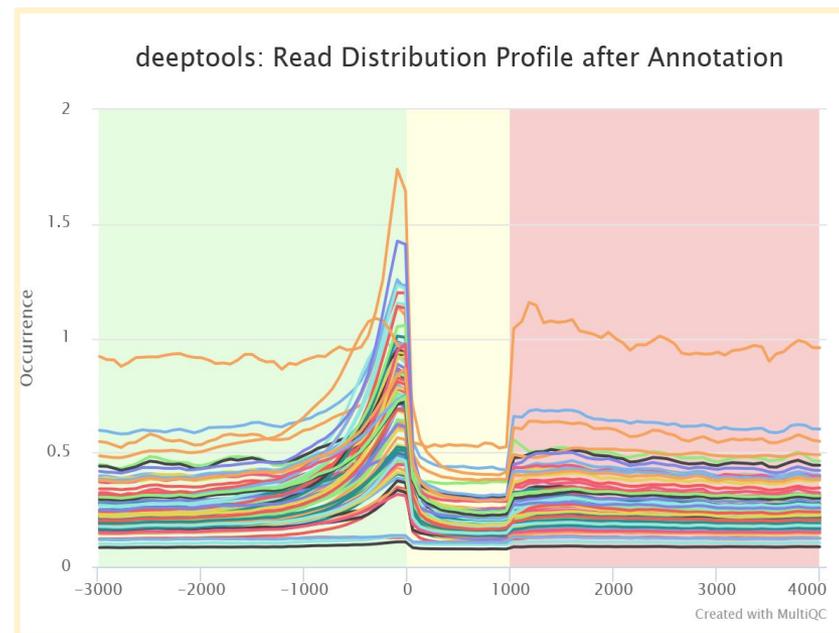
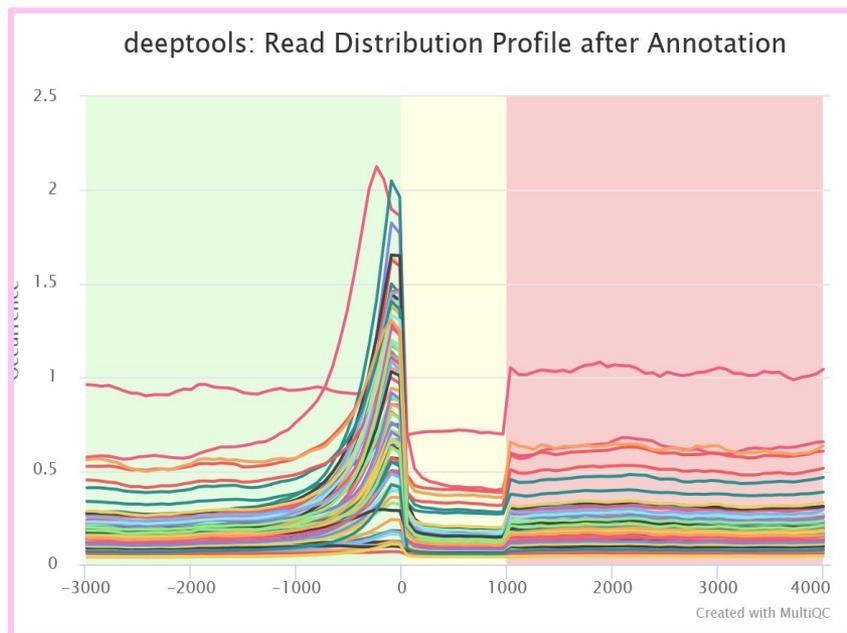
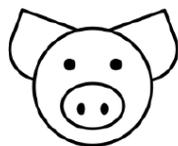
https://data.faang.org/api/fire_api/experiments/DIAGENODE_SOP_GENESWITCH_ATAC-SEQ_20201118.pdf

Commercial kit (DIAGENODE):

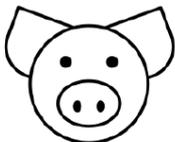
ATAC-seq package for tissue (C01080006)

Achievements: ATAC-seq

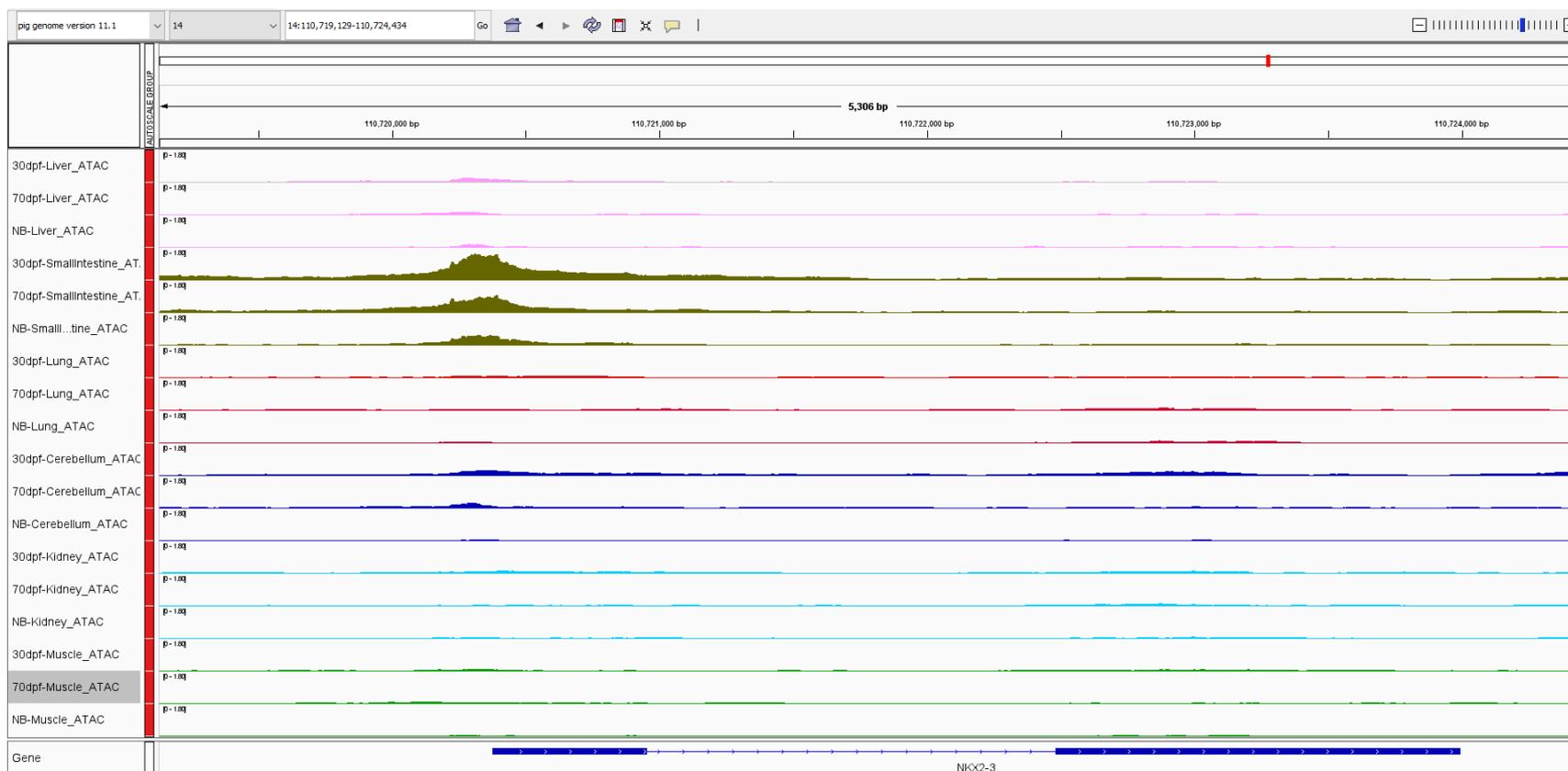
Quality control: reads distribution at TSS for pig and chicken libraries



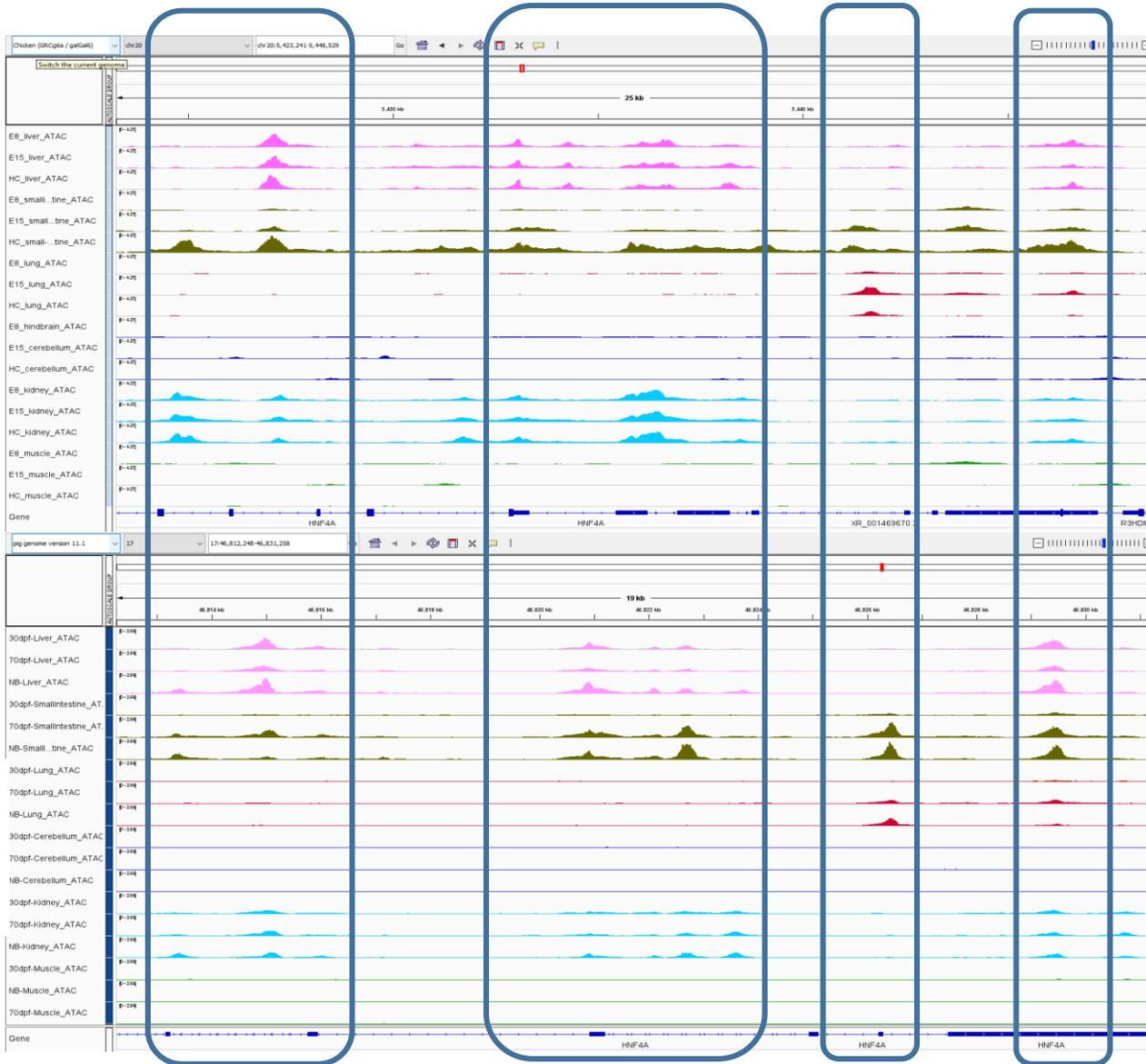
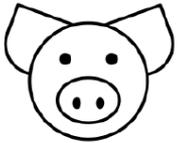
Achievements: ATAC-seq



Small intestine marker: NKX2.3



Achievements: ATAC-seq



HNF4A

Achievements: RNA-seq

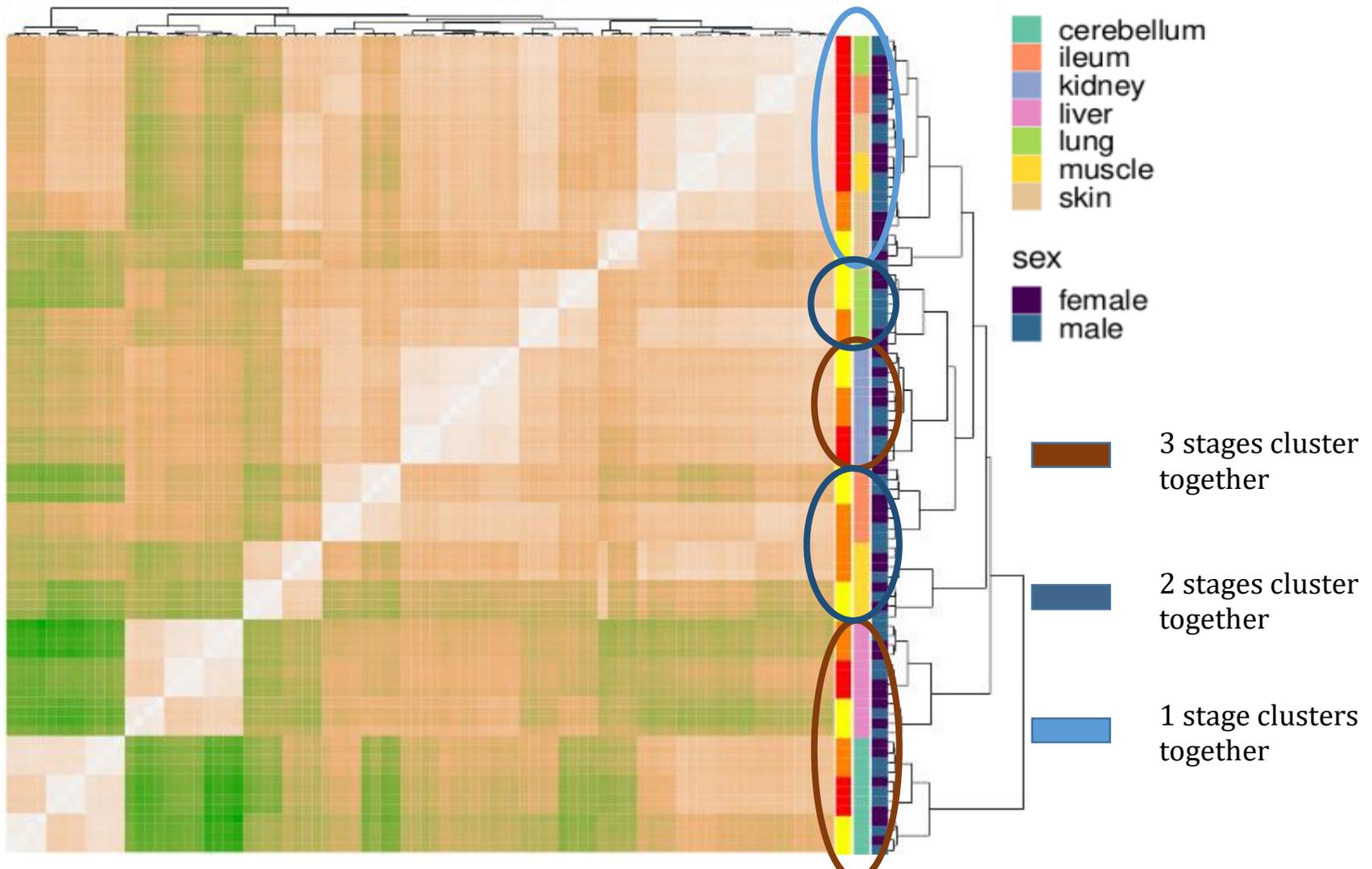


Assays	Species	Date of released	Accession number	Number of libraries
totRNAseq (polyA+)	Pig	2020-12	PRJEB41970	84
	Chicken	2020-12	PRJEB42025	84
smallRNA-seq	Pig	2020-12	PRJEB42001	84
	Chicken	2020-12	PRJEB42041	84

	Raw reads (median)	Uniquely mapped PE reads (median)
Pig dataset (totRNAseq)	150,082,484	143,342,894
Pig dataset (smallRNAseq)	67,614,012	16,898,621
Chicken dataset (totRNAseq)	112,414,695	105,321,362
Chicken dataset (smallRNAseq)	66,104,048	20,486,160

Achievements: mRNA-seq

Hierarchical clustering

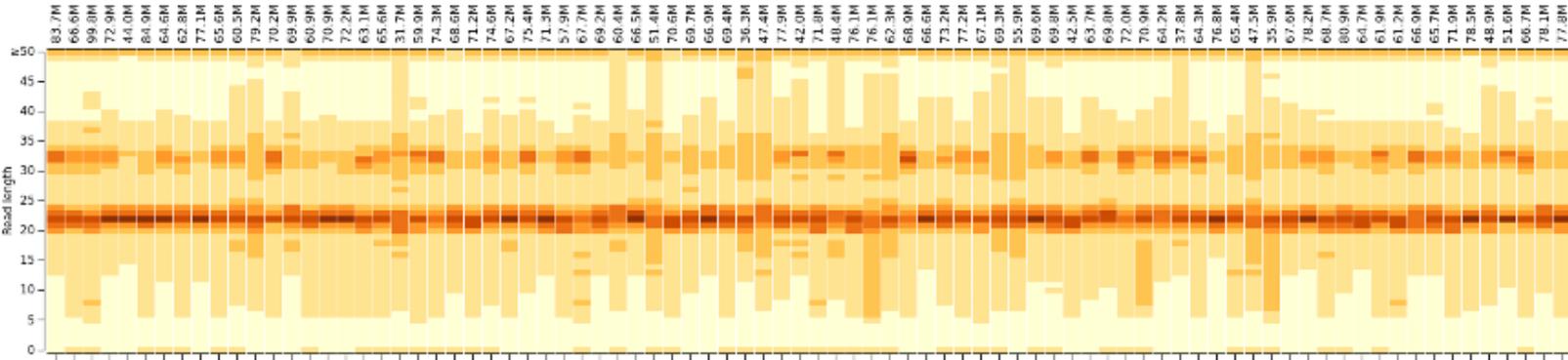


Achievements: smallRNA-seq



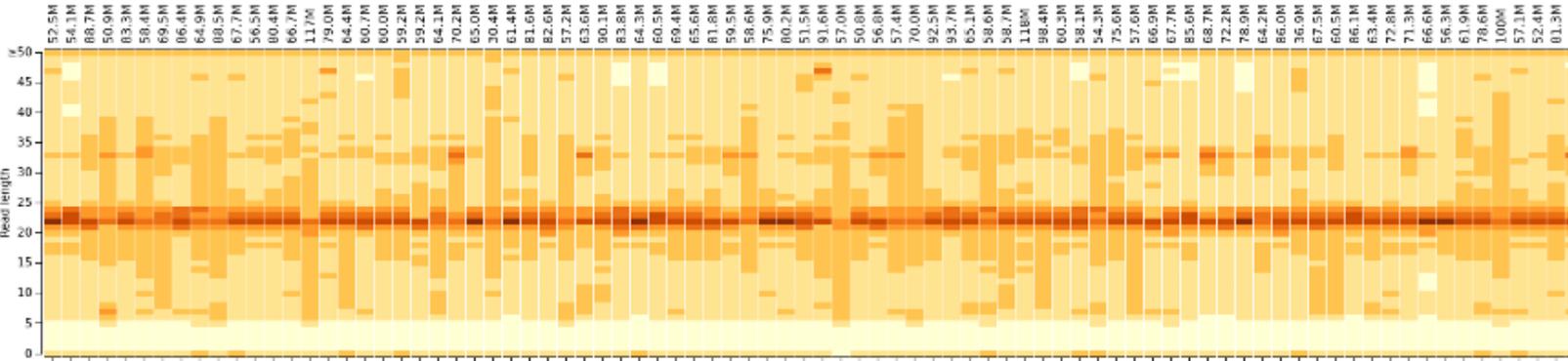
Read Length Distribution

Percentage of reads of each length.

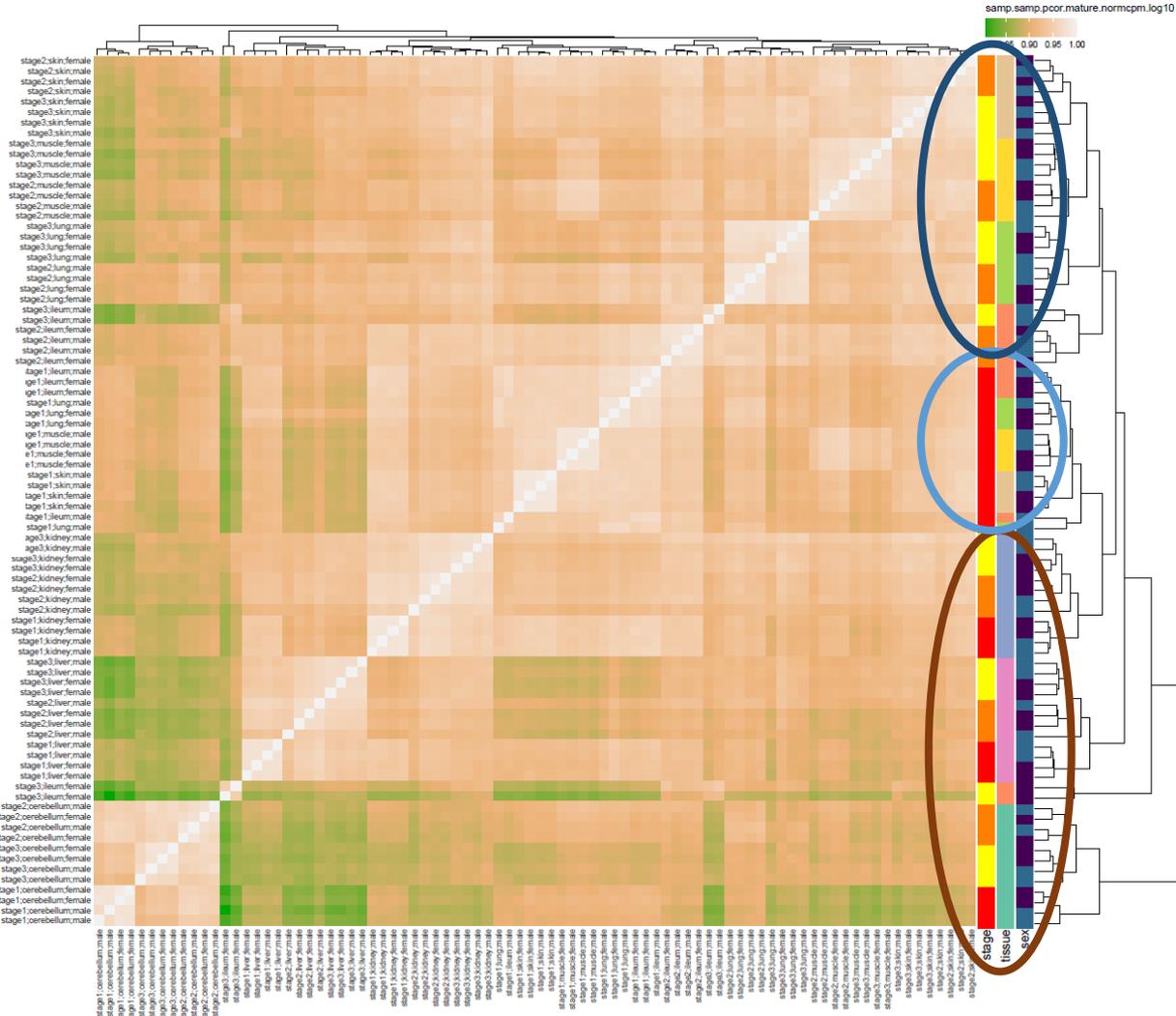


Read Length Distribution

Percentage of reads of each length.



Achievements: smallRNA-seq Hierarchical clustering



stage
■ stage1
■ stage2
■ stage3

tissue
■ cerebellum
■ ileum
■ kidney
■ liver
■ lung
■ muscle
■ skin

sex
■ female
■ male

■ 3 stages cluster together

■ 2 stages cluster together

■ 1 stage clusters together

Achievements: Iso-seq



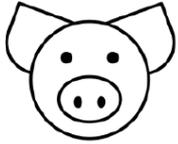
Assays	Species	Date of released	Accession number	Number of libraries
Iso-seq	Pig	2022-03	PRJEB50963	21
	Chicken	2022-03	PRJEB48060	21

Species	Mapped reads (median)	Total Genes (median per tissue)	Total Transcripts (median per tissue)
Pig	3,080,464	19,578	85,311
Chicken (GRCg7w)	2,416,260	12,224	68,802

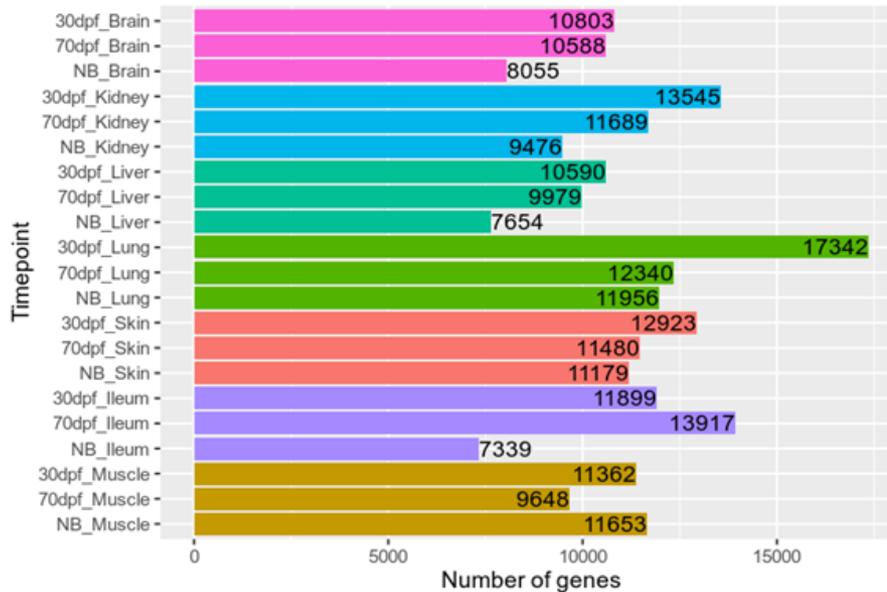
Protocol:

https://data.fang.org/api/fire_api/experiments/ROSLIN_SOP_GENESWITCH_WP1_Iso-seq_20210809.pdf

Achievements: Iso-seq

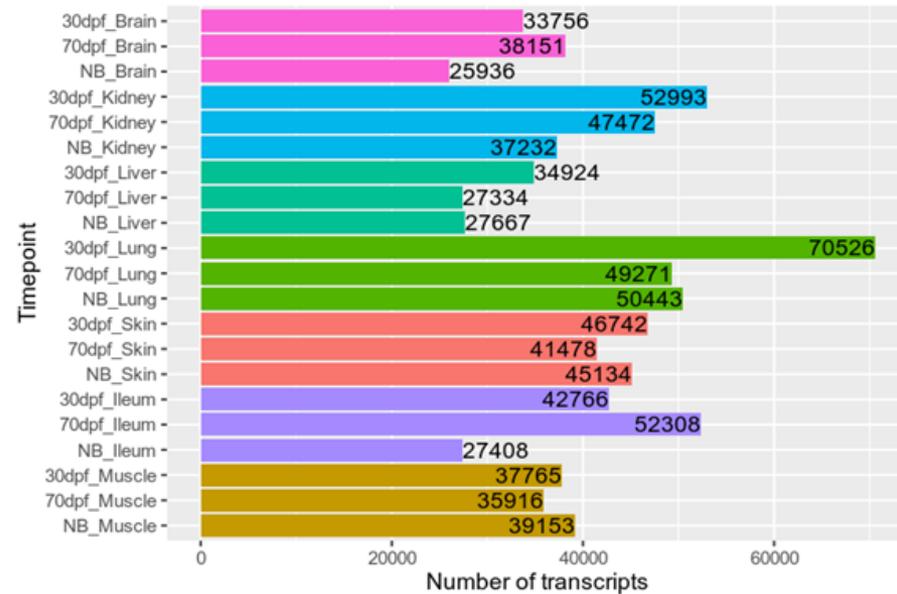


Sscrofa11.1 - Number of genes per Timepoint



Tissue Skin Muscle Lung Liver Kidney Ileum Brain

Sscrofa11.1 - Number of transcripts per Timepoint



Tissue Brain Ileum Kidney Liver Lung Muscle Skin

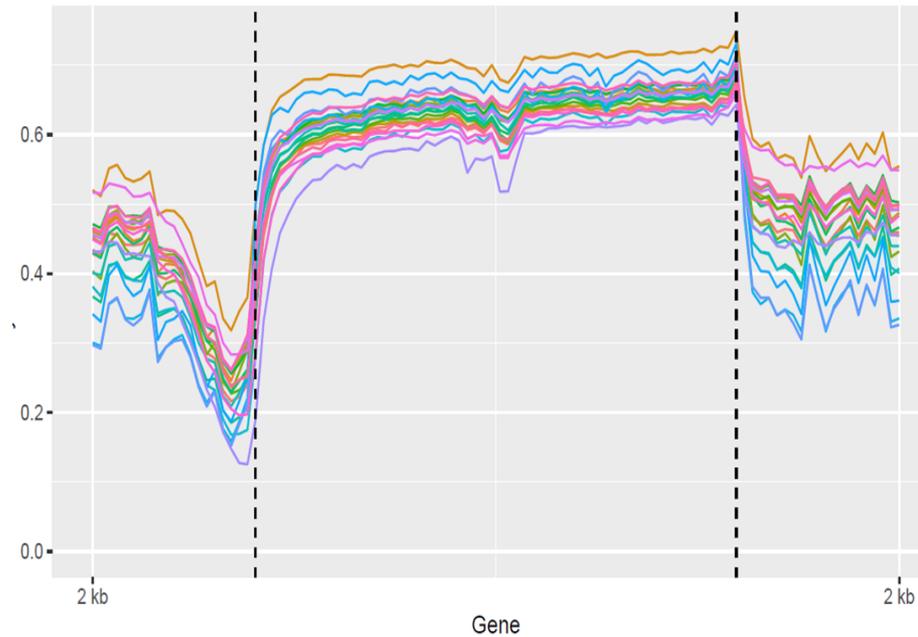
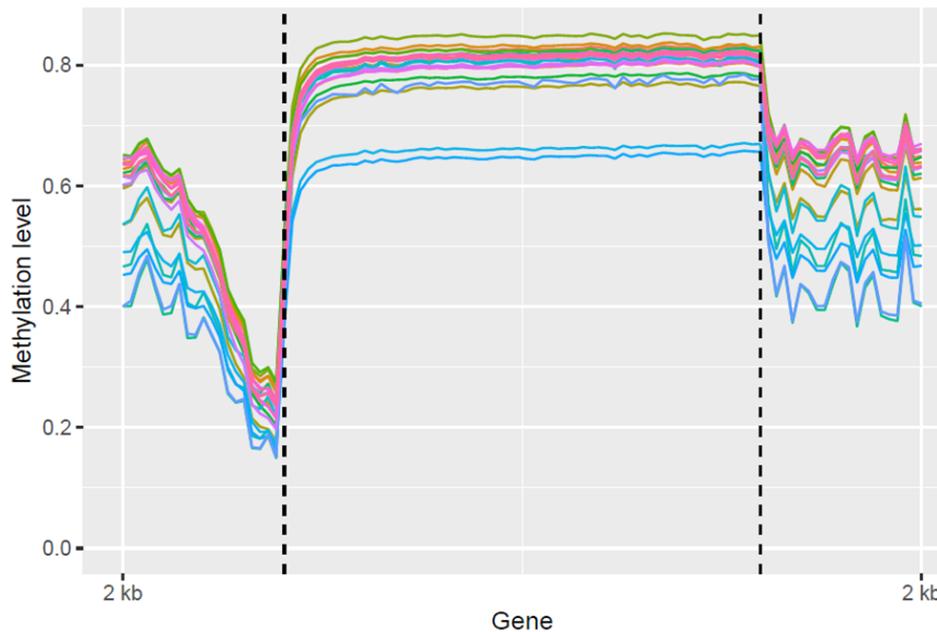
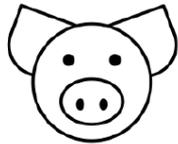
Achievements: DNA methylation

Assays	Species	Date of released	Accession number	Number of libraries
RRBS	Pig	2020-12	PRJEB41822	63
	Chicken	2020-12	PRJEB41829	63
WGBS	Pig	2021-01	PRJEB42772	21
	Chicken	2021-01	PRJEB42775	21

	Raw reads (RRBS,mean) or coverage (WGBS)	Uniquely mapped reads (mean)
Pig dataset (RRBS)	59,265,612	41,485,928
Pig dataset (WGBS)	36X	
Chicken dataset (RRBS)	54,671,484	37,176,609
Chicken dataset (WGBS)	31X	

Achievements: DNA methylation

Methylation levels at promoters



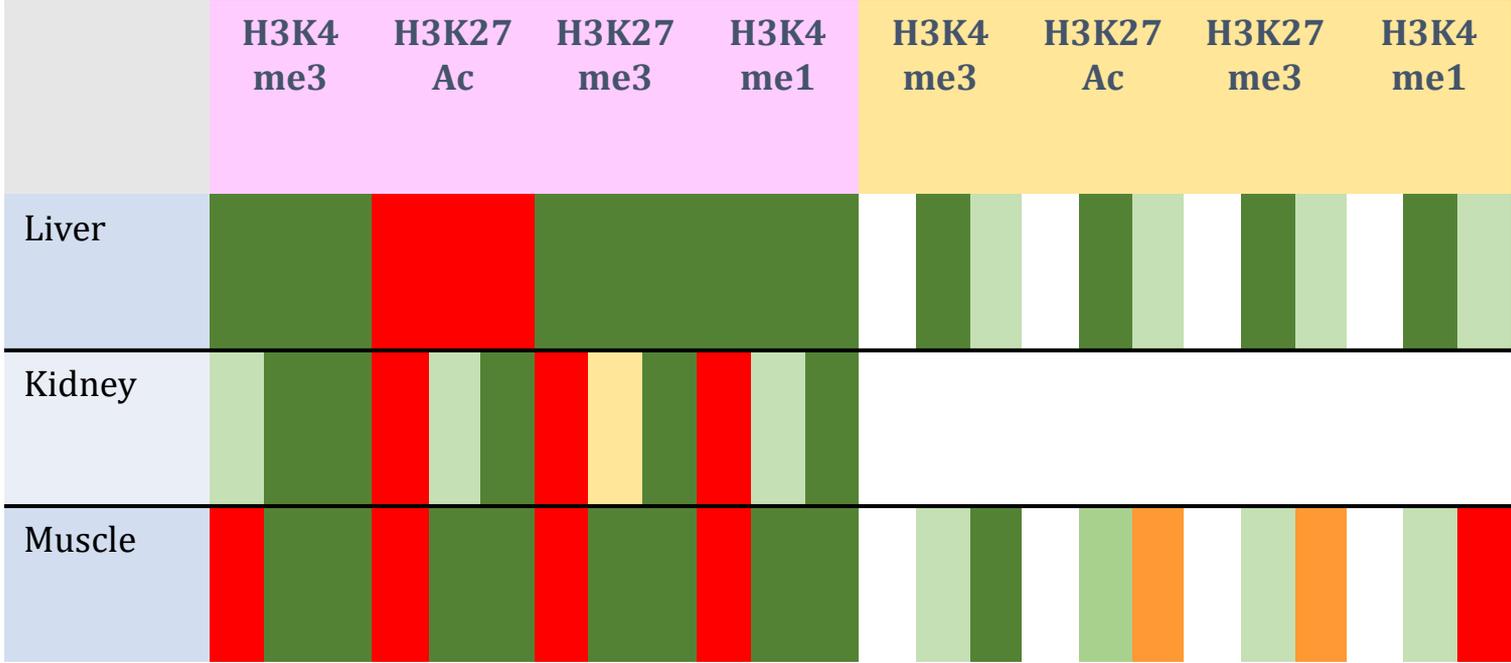
On-going ChIP-seq

The logo for INRAE, consisting of the letters 'INRAE' in a teal, sans-serif font.The logo for diagenode, featuring the word 'diagenode' in white lowercase letters on a red rectangular background.

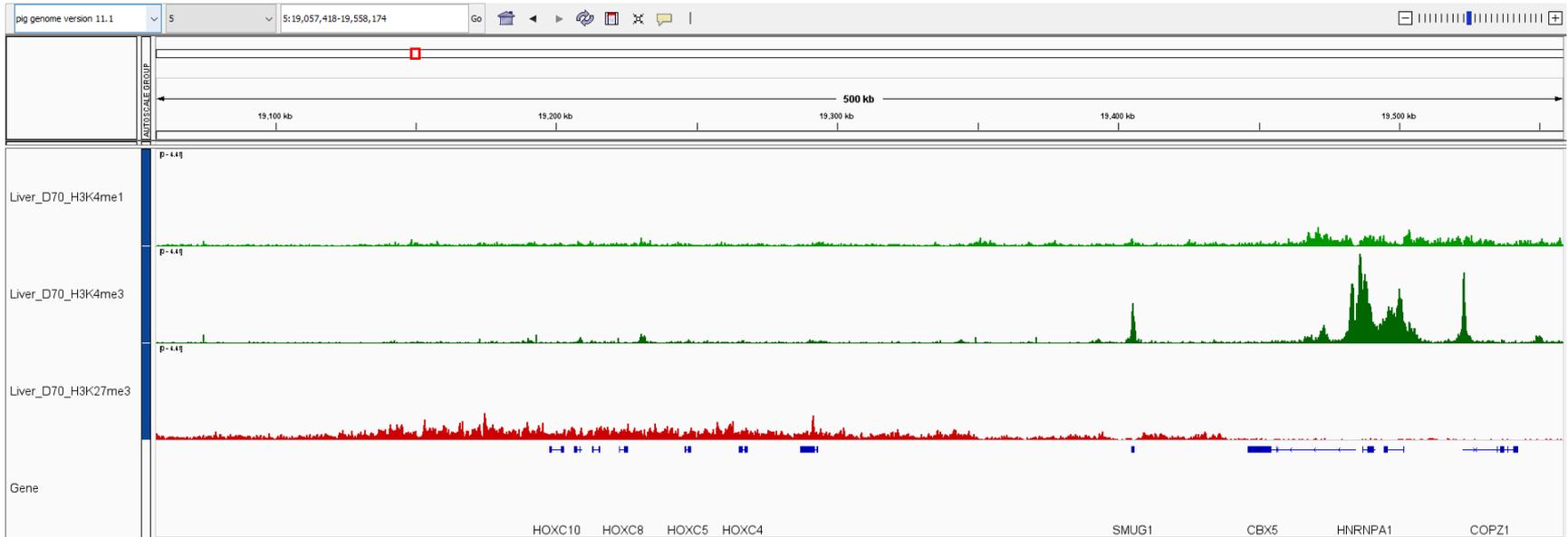
WAGENINGEN
UNIVERSITY & RESEARCH

- **Protocols have been optimized for pig and chicken (ChIPmentation)**
- **Pig and chicken libraries have been produced and sequenced**
- **QC and dataset validation are on-going**

Achievements: ChIP-seq (on-going)



Achievements: ChIP-seq (on-going)



HoxC complex

Achievement: Promoter Capture Hi-C

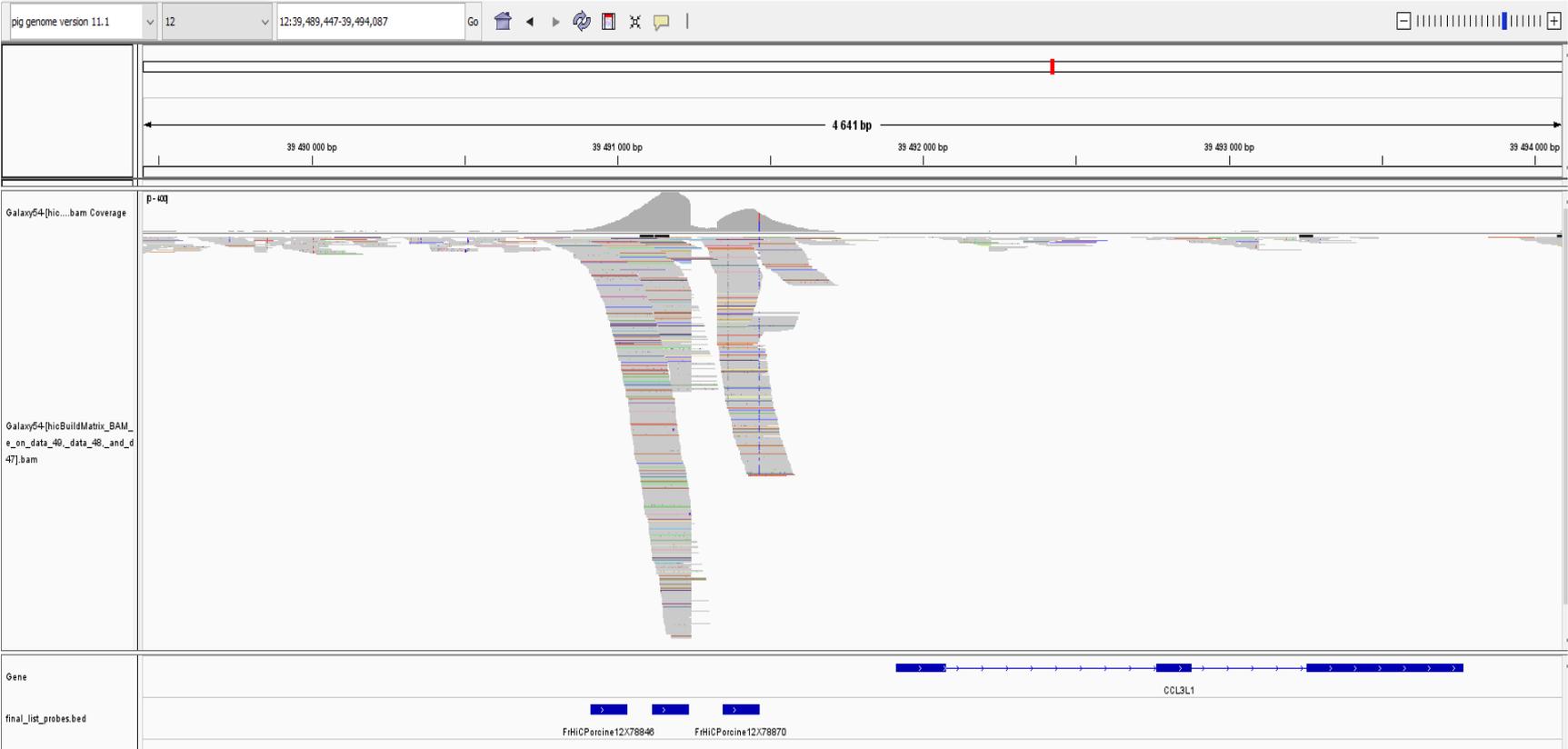
Assays	Species	Date of released	Accession number	Number of libraries
Capture Hi-C	Pig	2021-04	PRJEB44486	12
	Chicken	2022-06	PRJEB53986	12

- **Protocols have been optimized for pig and chicken (Arima HiC + Agilent SureSelect)**

https://data.fang.org/api/fire_api/experiments/INRAE_SOP_GENESWITCH_WP1_PROMOTER_CAPTURE_HI-C_20210325.pdf

- **~15,000 promoters being captured for each species**

Achievement: Promoter Capture Hi-C



• **FAIR principles applied to GENE-SWitCH**

All metadata regarding samples, experiments and analysis are publicly available on the FAANG Data Portal. It includes protocols, sample descriptors, experimental parameters and analysis parameters.

All raw sequences have been deposited in the European Nucleotide Archive (ENA) , pipelines are publicly available either on github or nf-core.

All the information and material related to GENE-SWitCH are easily Findable, Accessible, Interoperable and Reusable from the FAANG Data Portal.

• Accessible through the FAANG data portal

<https://data.faang.org/projects/GENE-SWitCH>


Home Data ▾ Projects Summary Submit ▾ Ontology Improver ^{Beta} Search Help ▾



<https://www.gene-switch.eu/>

GENE-SWitCH aims to deliver new underpinning knowledge on the functional genomes of two main monogastric farm species (pig and chicken) and to enable immediate translation to the pig and poultry sectors.



Horizon2020
European Union Funding
for Research & Innovation

Tweets by @GeneSwitch ①

 GENE-SWitCH Retweeted

 **EuroFAANG**
@EuroFaang

EuroFAANG is a coordinated effort to unravel the connection between the genetic make-up of an animal and the observable physical and physiological traits. EuroFAANG will address key challenges in farmed animal production in collaboration with the global [@faangomics](#) initiative.

Related Datasets Column selection ▾

Study name	Species	Archive	Assay type
PRJEB40576	Sus scrofa	ENA	Hi-C
PRJEB42041	Gallus gallus	ENA	microRNA profiling by high throughput sequencing
PRJEB41822	Sus scrofa	ENA	methylation profiling by high throughput sequencing
PRJEB45945	Gallus gallus	ENA	ATAC-seq
PRJEB42775	Gallus gallus	ENA	methylation profiling by high throughput sequencing
PRJEB48060	Gallus gallus	ENA	ISO-Seq
PRJEB42025	Gallus gallus	ENA	RNA-seq of total RNA
PRJEB42001	Sus scrofa	ENA	microRNA profiling by high throughput sequencing
PRJEB42772	Sus scrofa	ENA	methylation profiling by high throughput sequencing
PRJEB44468	Sus scrofa	ENA	ATAC-seq

« Previous 1 2 Next »

Acknowledgments



Megan Davey, Jonathan Smith, Jacqueline Smith, Kasia Miedzinska, Sebastien Guizard, Mick Watson



Mayrone Mongellaz, Sylvain Foissac, Cyril Kurilo, Cervin Guyomard, Marine Beinat, Jérôme Lluch, Tabatha Bulach, Jules Sabban, Hervé Acloque, Elisabetta Giuffra



Celine Sabatel, Wassim Lakhal, Laurence Servais



Sarah Djebali



Jani de Vos, Ole Madsen, Martijn Derks



Alexey Sokolov, Peter Harrison



*This project has received funding from the European Union's **Horizon 2020** Research and Innovation Programme under the grant agreement n° 817998*

@GeneSwitch - www.gene-switch.eu

Thank you for your attention

INRAE

Name: Elisabeth Guillaud, Organization: INRAE

Name: Anne Helou, Organization: INRAE

Name: Sylvain Fossas, Organization: INRAE

Name: Fabrice Audebert, Organization: INRAE

Name: Andréa Roux, Organization: INRAE

Name: Julien Caron, Organization: INRAE

Name: Pascal Astoux, Organization: INRAE

INRAE > transfert

Inserm

Name: Sarah Dyeball, Organization: INSERM

Name: Carole Bénard, Organization: IT

EMBL
European Molecular Biology Laboratory

Name: Fergal Martin, Organization: EMBL EDI

Name: Peter Harrison, Organization: EMBL EDI

Name: Alchutha Dwyer, Organization: EMBL EDI

EAAP
European Federation of Animal Science

Name: Marlene Scaretta, Organization: EAAP

Name: Bruno Perez, Organization: EAAP

Name: Federica Matera, Organization: EAAP

Name: Federico Ligotti, Organization: EAAP

IRTA

Name: Joan Pardo Sanchez, Organization: IRTA

Name: Maria Balcells, Organization: IRTA

Name: Antoni Galobardes, Organization: IRTA

EFFAB
European Forum of Farm Animal Breeders

Name: Olga Kovač, Organization: EFFAB

Name: Ana Grubisic, Organization: EFFAB

UPPSALA UNIVERSITET

Name: Ogor Carlberg, Organization: Uppsala University

Name: Ying Guo, Organization: Uppsala University

HENDRIX GENETICS

Name: Bruno Perez, Organization: Hendrix Genetics

Name: Marco Bink, Organization: Hendrix Genetics

SAB member

Name: Graham Phillips, Organization: UKBARTA

Name: Geert Bruggeman, Organization: SAB Member

Name: Marie-José Mercat, Organization: IFPI Alliance R&D

Name: Filippo Miglior, Organization: University of Georgia

WAGENINGEN UNIVERSITY & RESEARCH

Name: Marco Mengo, Organization: WUR

Name: Ole, Organization: WUR

Name: Jani, Organization: WUR

Name: Linda Looen, Organization: WUR

THE UNIVERSITY OF EDINBURGH

Name: Sébastien Guibert, Organization: Roslin

Name: Jacqueline Smith, Organization: UEDIN

Name: Alan Archibald, Organization: The Roslin Institute

Name: BioMick, Organization: Roslin

diagenoë

Name: Wassim Lakkaj, Organization: DIAGEN

