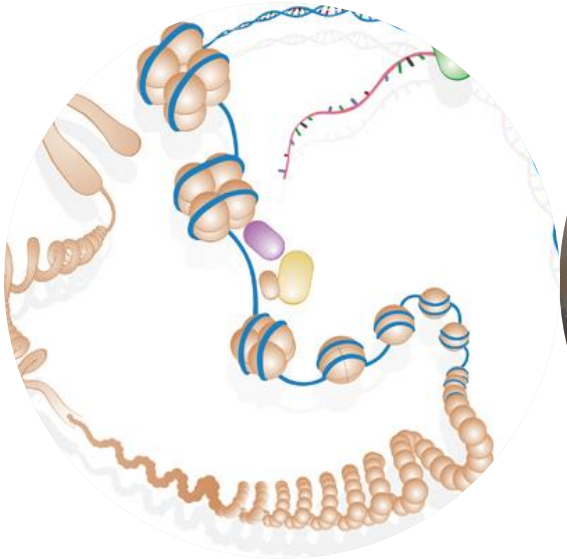
Functional annotation of farm animal genomes: ChIP-seq

Richard Crooijmans

Richard.Crooijmans@wur.nl

2018, PAGXXVI

FAANG
Functional Annotation of Animal Genomes



WAGENINGEN
UNIVERSITY & RESEARCH

Why FAANG is important

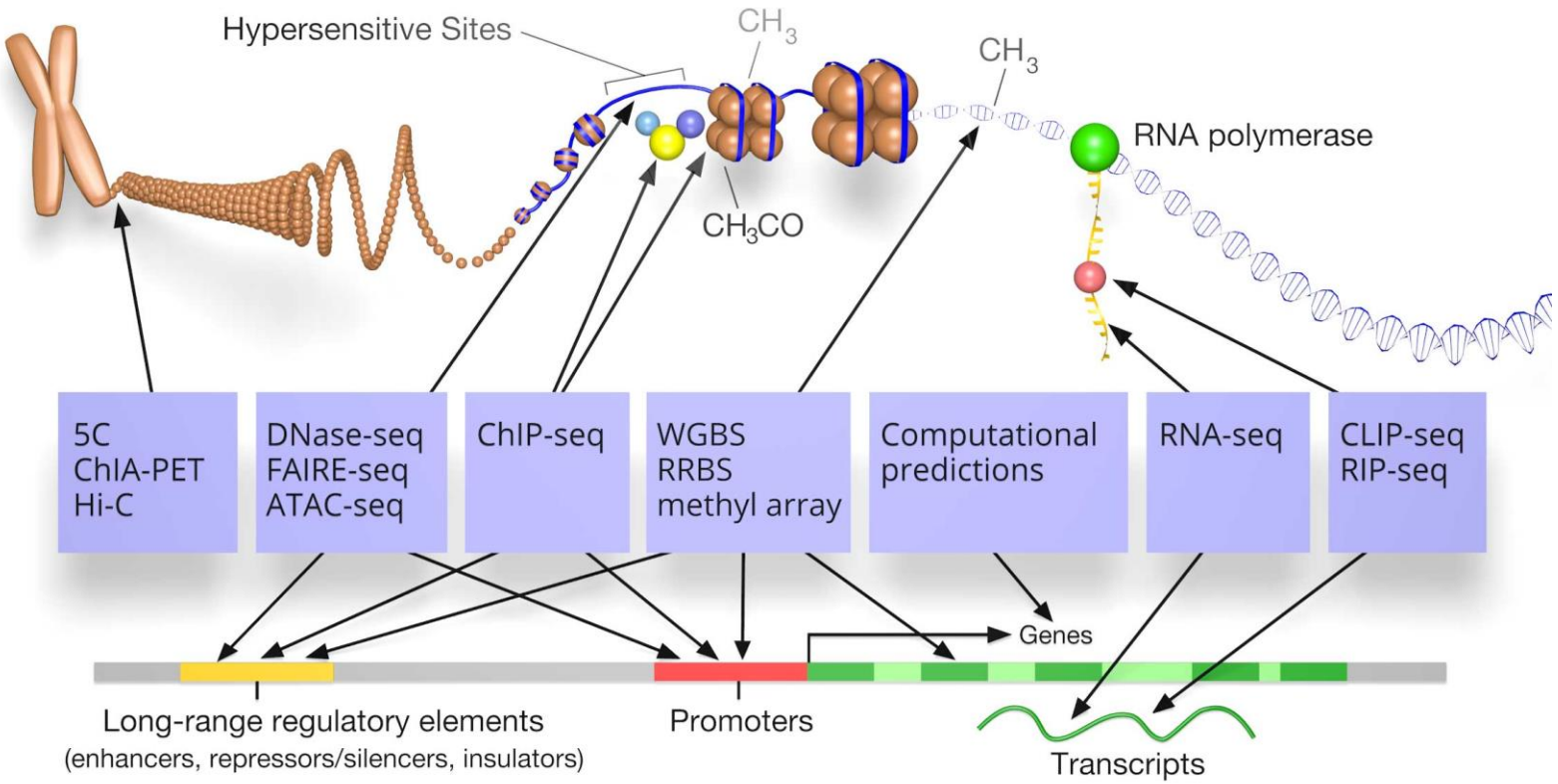
Understanding the genotype to phenotype link

- genomic selection
- improving fundamental understanding of biology

This needs:

- high quality reference genome sequence
- comprehensive annotation of the functional elements
- common infrastructure providing
  - biological resources
  - Bioinformatics tools
  - databases

What to study?



Based on an image by Darryl Leja (NHGRI), Ian Dunham (EBI), Michael Pazin (NHGRI)

International reference material

Pig IPEC-J2

- Established from normal intestinal epithelium cells isolated from the jejunum of a neonatal, unsuckled pig
- Available from DSMZ in Germany

Chicken SL-29

- Established from a 11 day old decapitated chicken embryo by standard trypsinization
- Layer (leghorn)
- Available from ATCC

Both primary untransformed cell lines

FAANG assays

- Whole genome re-sequence (WGS: 30x)
- RNA-seq, poly A and stranded
- Whole genome bisulfite sequencing (WGBS: 50x)
- Reduced representation bisulfite sequencing (RRBS)
- ChIP-seq (histone marks and CTCF)
  - H3K27ac active enhancers and promoters
  - H3K4me3 promoter active genes and transcription start sites
  - H3K27me3 silenced genes (active during developmental stages)
  - H3K4me1 active enhancers
  - CTCF insulator activity
- ATAC-seq (University of Leiden, The Netherlands; in progress using different protocols)

Preliminary analysis

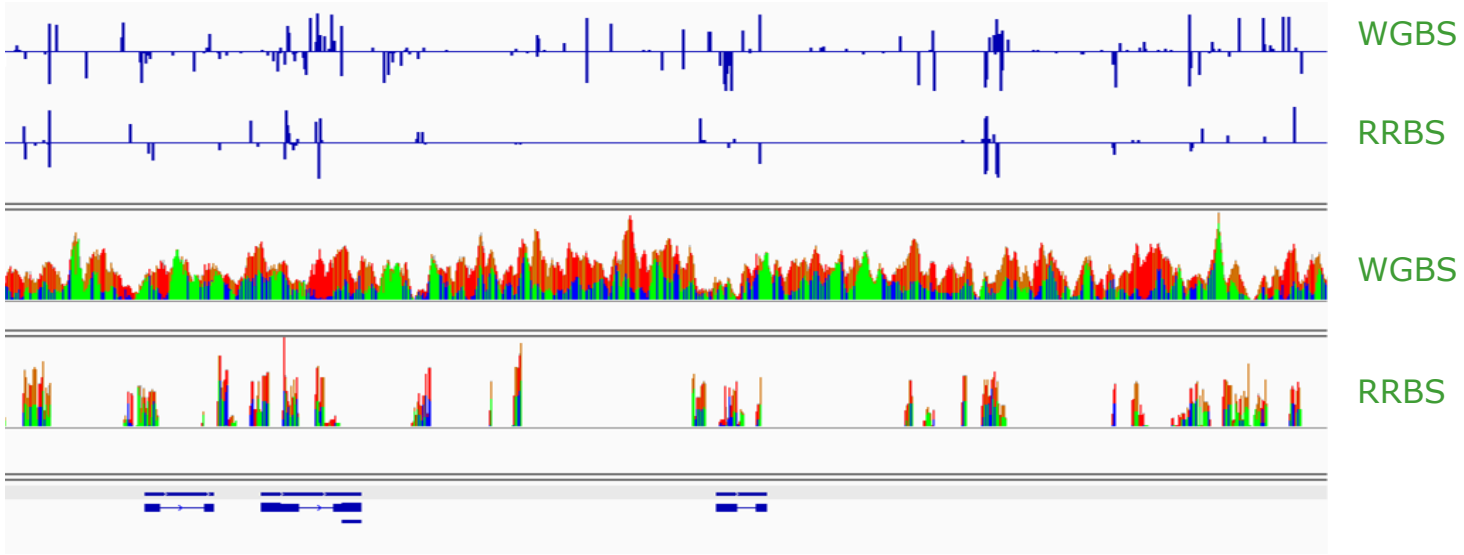
- Standard analysis
  - WGS: BWA mem
  - RRBS: BSeeker2
  - WGBS: BSseeker2
  - RNA-seq: Tophat and Cufflinks
  - ChIP-seq: Bowtie and MACS
- Future: FAANG analysis pipelines

Methylation

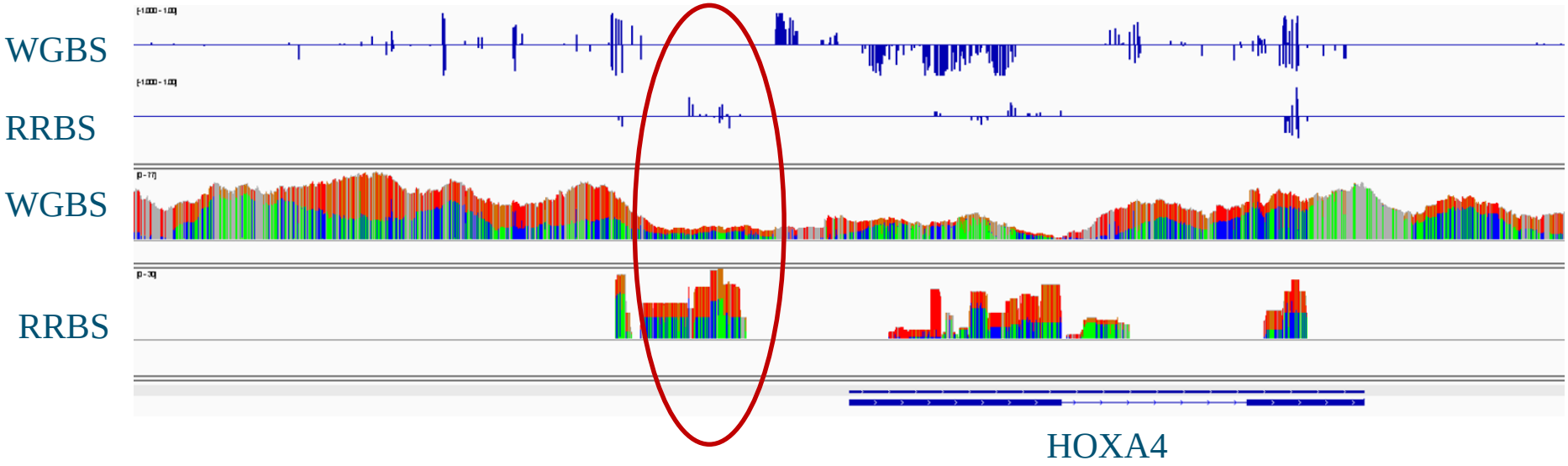
- Whole Genome Bisulphite Sequencing (WGBS)
  - High coverage needed
  - Quiet expensive
- Reduced Representation Bisulphite Sequencing (RRBS)
  - Much lower cost
  - Enrichment of the specific regions

Methylation: Comparison WGBS and RRBS

- Comparison of overlapping sites (Coverage >10 reads)
- Similar methylation distribution
- Very good correlation between 2 RRBS samples (0.96)
- Good correlation between WGBS and RRBS (0.93)



Methylation: RRBS and WGBS provide complementary information



Region with very low coverage in WGBS but sufficient coverage in RRBS data

Chr 18:50,074,454-50,078,736

ChIP-seq

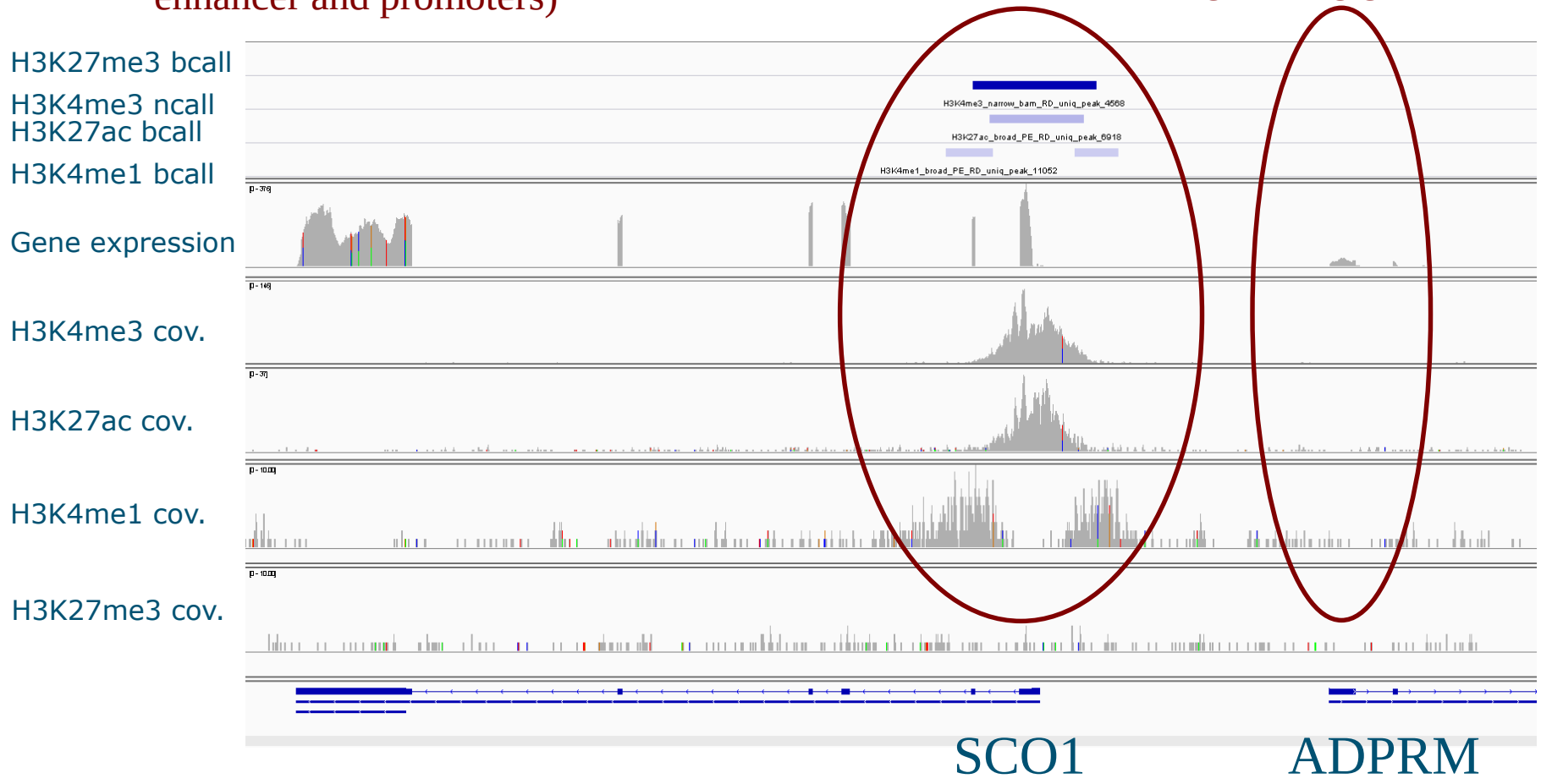
- Active elements that interfere in gene expression (enhancers, silencers)

- Different protocols
  - Extracting nuclei
  - Fixation
  - Amount of input for ChIP experiment
  - Antibodies
- Different controls in the procedure (positive and negative)

Example expressed gene: SCO1

Signal for H3K4me3, H3K27ac and H3K4me1 (active enhancer and promoters)

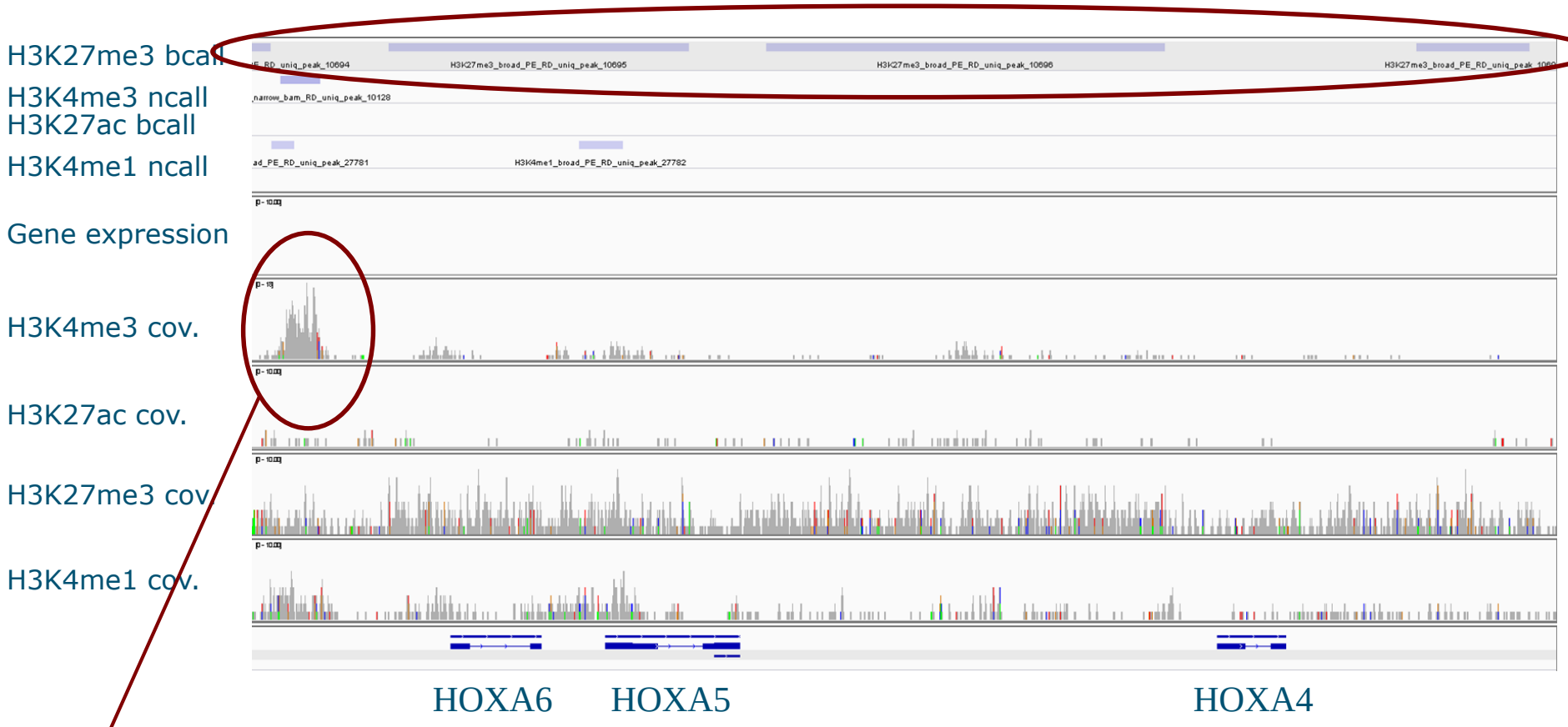
No signal for silent neighboring gene



chr 12:58,037,835-58,069,194

Example silenced genes (HOX genes chr18)

Many H3K27me3 peaks: Silenced genes

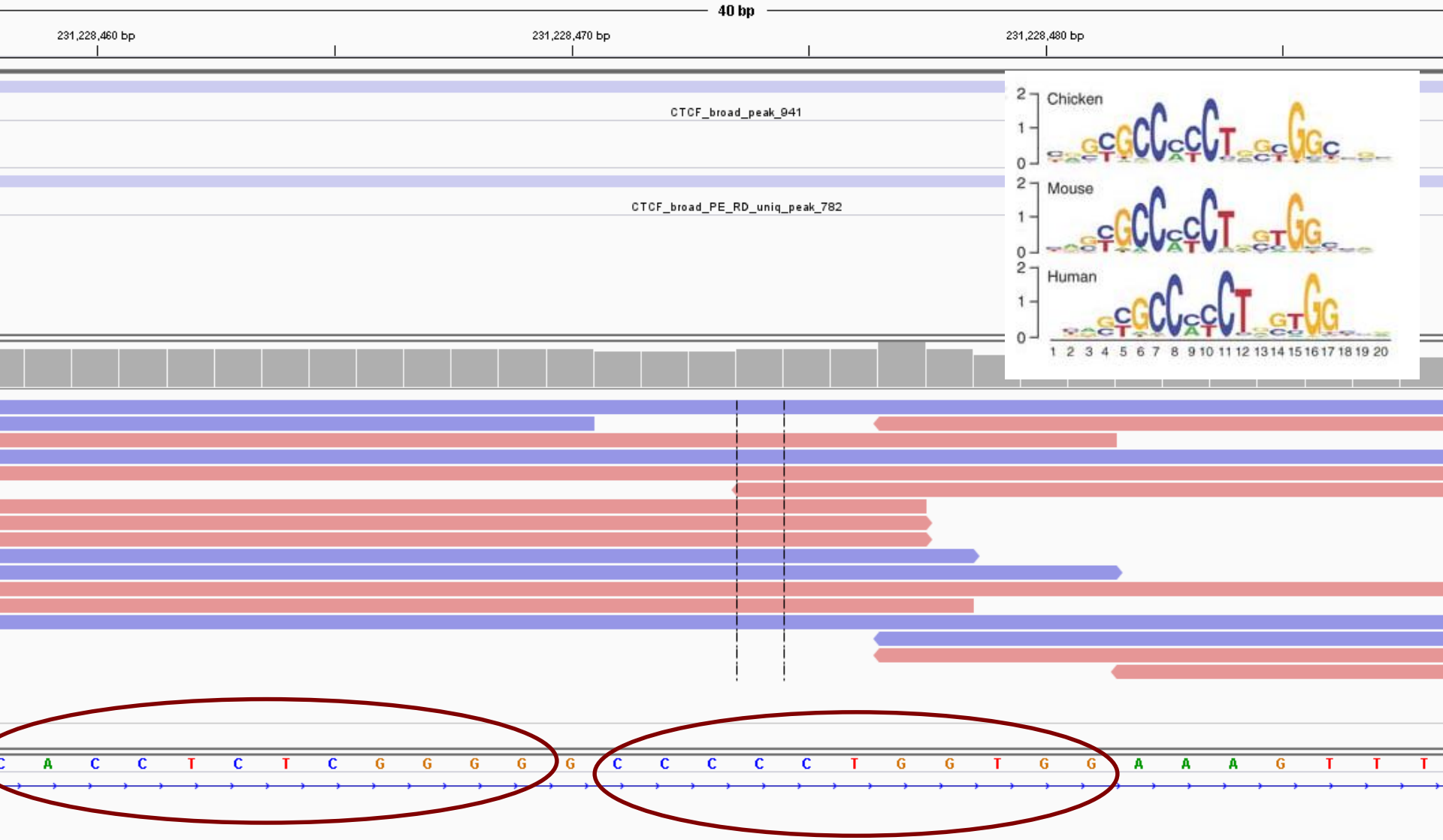


A promoter element but not active due to missing H3K27ac?

18:50,054,955-50,084,206

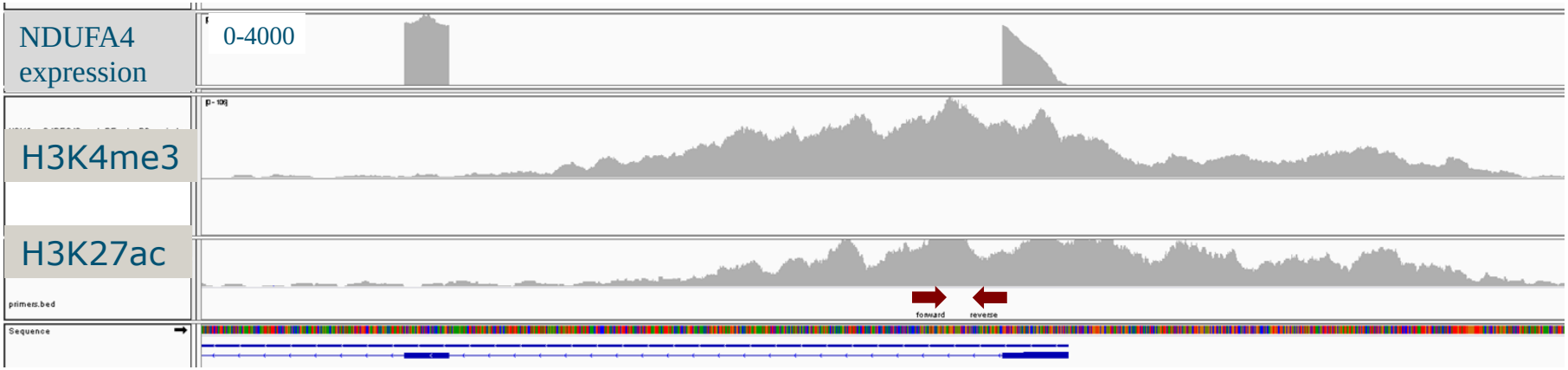
WAGENINGEN
UNIVERSITY & RESEARCH

Example CTCF consensus at CTCF peak

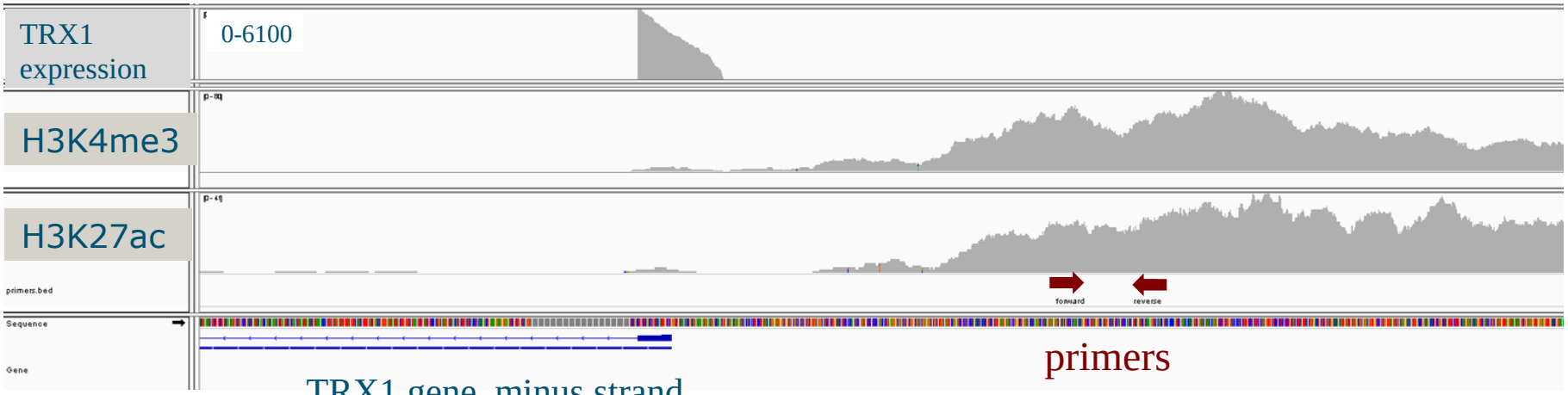


WAGENINGEN
UNIVERSITY & RESEARCH

qPCR: Positive controls for H3K4me3 and H3K27ac



NDUFA4 gene, minus strand primers

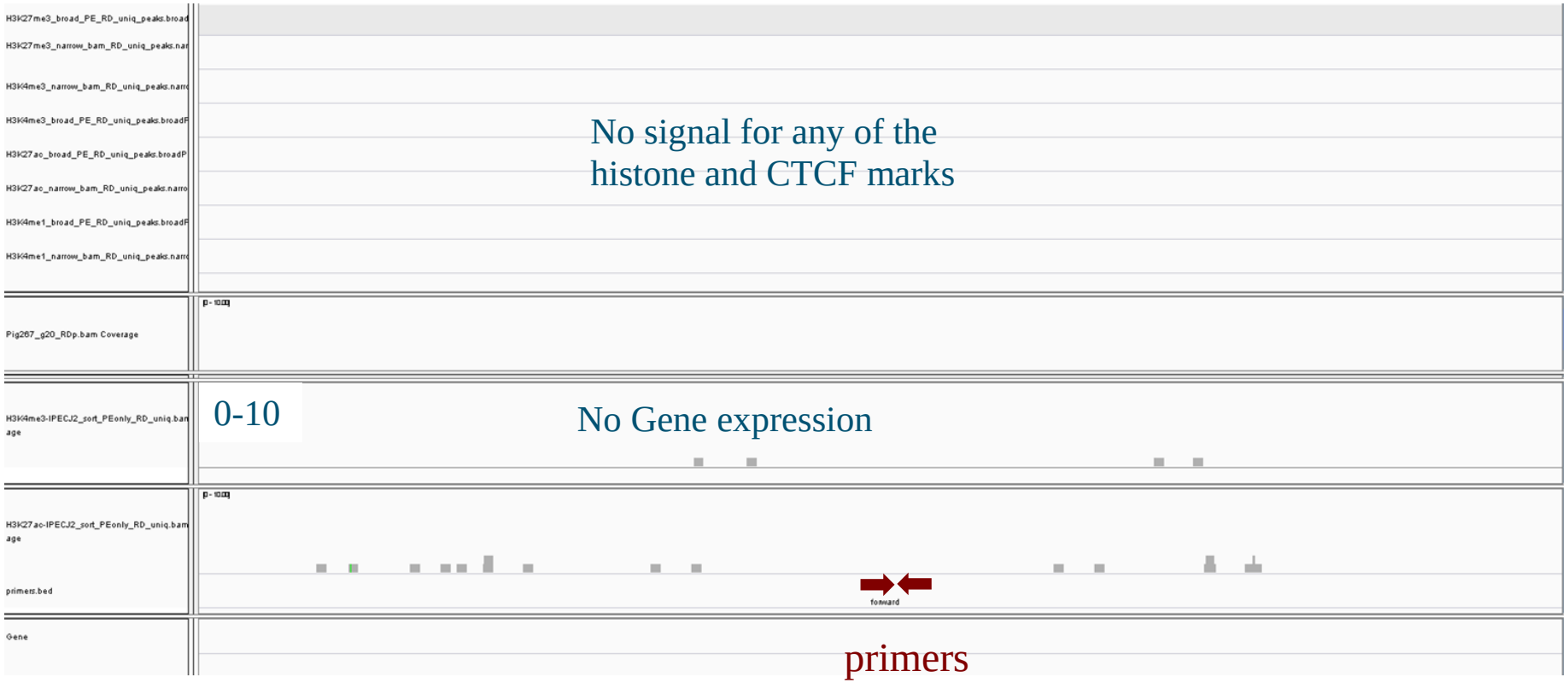


primers

TRX1 gene, minus strand

WAGENINGEN
UNIVERSITY & RESEARCH

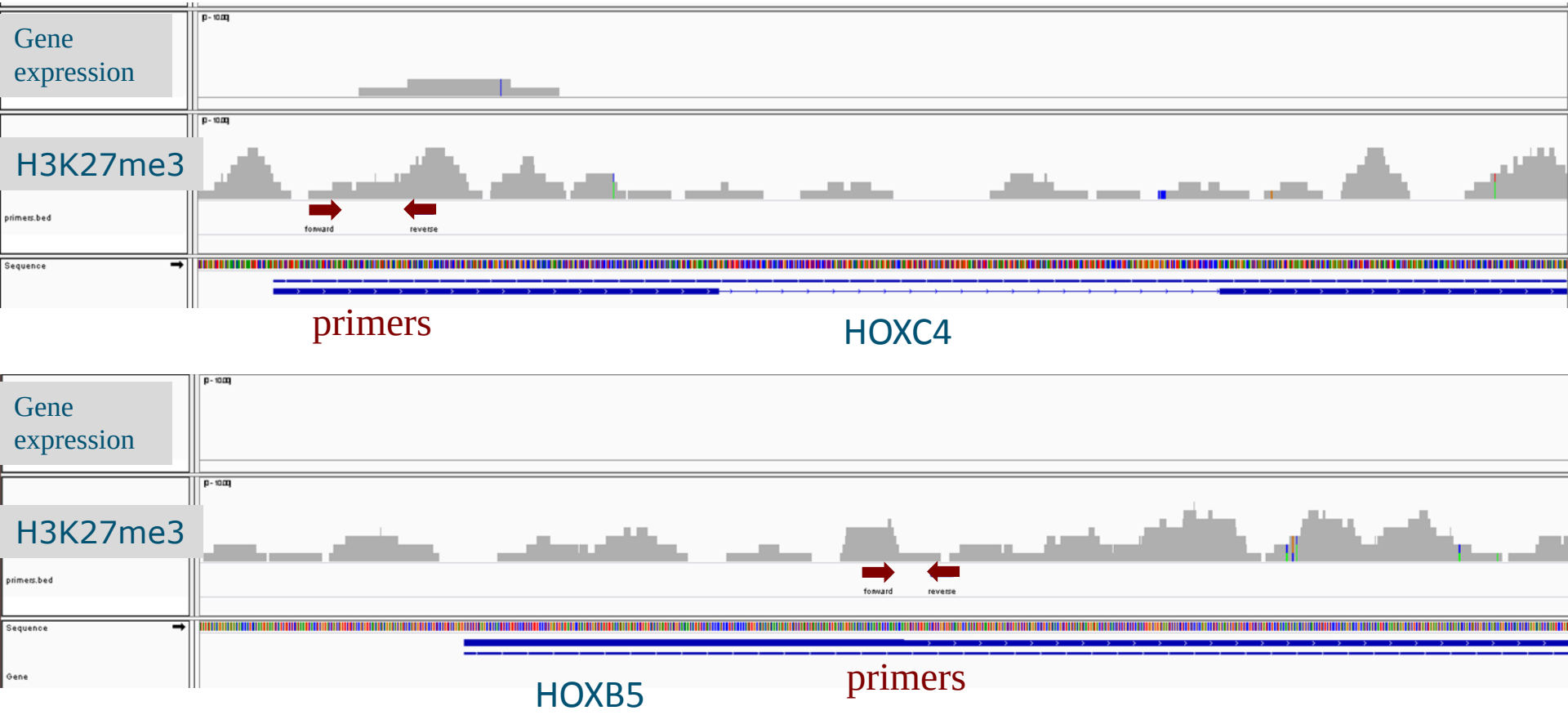
qPCR: Negative control primers (gene desert)



Chip_NegC2_2 position 2:44,492,350-44,498,036

qPCR: Positive control for H3K27me3 at HOX genes

- Primer combinations not very good: design new combinations



Next step: from cultured cells to tissue

* Optimize protocols extracting nuclei
    - from different tissue
    - from different cells
    - from limited number of cells/tissue (biopsies, organoids)
* Optimize controls qPCR
    - positive and negative controls per tissue

* Reports and protocols on the FAANG website
* Samples deposited in Biosamples (EBI)

Overview ChIP–seq assays

Species	Tissue/cell line	name	Mark	Ab provider	reference	Ab quantity (µg)	starting amount/IP	chromatin preparation
	Frozen Tissue		H3K4me3	Diagenode	C15410003	0,5 and 1 ug	2mg of tissue	DIAGENODE_SOP_ChIP-seq_for_Histone_Marks_20170630
	Frozen Tissue		H3K27me3	Diagenode	C15410194	0,5 and 1 ug	2mg of tissue	DIAGENODE_SOP_ChIP-seq_for_Histone_Marks_20170630
	Frozen Tissue		H3K27ac	Diagenode	C15410195	0,5 and 1 ug	2mg of tissue	DIAGENODE_SOP_ChIP-seq_for_Histone_Marks_20170630
	Frozen Tissue		H3K4me1	Diagenode	C15410196	0,5 and 1 ug	2mg of tissue	DIAGENODE_SOP_ChIP-seq_for_Histone_Marks_20170630
	Frozen Tissue		H3K4me3	Diagenode	C15410003	0,5 and 1 ug	50mg of tissue	DIAGENODE_SOP_ChIP-seq_for_low_input_20170630
	Frozen Tissue		H3K27me3	Diagenode	C15410194	0,5 and 1 ug	50mg of tissue	DIAGENODE_SOP_ChIP-seq_for_low_input_20170630
	Frozen Tissue		H3K27ac	Diagenode	C15410195	0,5 and 1 ug	50mg of tissue	DIAGENODE_SOP_ChIP-seq_for_low_input_20170630
	Frozen Tissue		H3K4me1	Diagenode	C15410196	0,5 and 1 ug	50mg of tissue	DIAGENODE_SOP_ChIP-seq_for_low_input_20170630
Chicken	Cell		H3K4me3	Diagenode	C15410003	5.7 µg	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105
Chicken	Cell		H3K27me3	Diagenode	C15410194	5.2 µg	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105
Chicken	Cell		H3K27ac	Abcam	ab4729	3.0 µg	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105
Chicken	Cell		H3K4me1	Abcam	ab8895	2.7 µg	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105
Chicken	Cell		CTCF	Merck	07-729	3.0 µl	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105
Pig	Cell		H3K4me3	Diagenode	C15410003	5.7 ug	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105
Pig	Cell		H3K27me3	Diagenode	C15410194	5.2 ug	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105
Pig	Cell		H3K27ac	Abcam	ab4729	3.0 µg	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105
Pig	Cell		H3K4me1	Abcam	ab8895	2.7 µg	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105
Pig	Cell		CTCF	Merck	07-729	3.0 µl	5x 10^6 cells	WUR_SOP_ChIP-seq_20170105

ChIP-seq: qPCR controls

Species	animal id (BioSamples)	sex	development stage/ age	tissue/cells	ChIP-seq antibody	positive controls Gene id	genome build	genome location	primer_F	primer_F annotation	primer_R	primer _R annotation	Negative control Gene id	primer_F	primer_F annotation	primer_R	primer _R annotation
Gallus gallus		male		cells	all		Ggallus5.0								2:44496220		2:44,496,315
Sus scrofa		male		cells	H3k27ac	NDUFA4	Sscrofa11.1	13: 30,448,569-30,449,050		13:30,448,759-13:30,448,779		13: 30,448,838-13:448,848					
Sus scrofa		male		cells	H3K4me	NDUFA4	Sscrofa11.1	13: 30,448,569-30,449,050		13:30,448,759-13:30,448,779		13: 30,448,838-13:448,848					
Sus scrofa		male		cells	H3k27ac	TRX1	Sscrofa11.1	1:251,264,178-251,277,550									
Sus scrofa		male		cells	H3K4me	TRX1	Sscrofa11.1	1:251,264,178-251,277,550									

	positive controls						
ChIP-seq antibody	Gene id	genome build	genome location	primer_F	primer_F annotation	primer_R	primer _R annotation
all		Ggallus5.0					
H3k27ac	NDUFA4	Sscrofa11.1	13: 30,448,569-30,449,050	ACGTGGATCGCCTGAGTC	13:30,448,759-13:30,448,779	CCTGACGTCTGACCGTGA	13: 30,448,838-13:448,848
H3K4me	NDUFA4	Sscrofa11.1	13: 30,448,569-30,449,050	ACGTGGATCGCCTGAGTC	13:30,448,759-13:30,448,779	CCTGACGTCTGACCGTGA	13: 30,448,838-13:448,848
H3k27ac	TRX1	Sscrofa11.1	1:251,264,178-251,277,550				
H3K4me	TRX1	Sscrofa11.1	1:251,264,178-251,277,550				

WAGENINGEN
UNIVERSITY & RESEARCH

Questions to the FAANG community

1. ChIP-seq protocols (fresh/frozen; cells/tissue/biopsy/single cell)
    a. tissue disaggregation and DNA-protein cross-linking
    b. Including tips and tricks /tissue/stage
    c. qPCR positive and negative controls/tissue/species

Mail to: richard.crooijmans@wur.nl

Summer school Wageningen, June 25-29, 2018 (pre-announcement)

Content:

- International speakers on functional annotation.

- Wet-lab experiments (hands-on)

- Data analysis (data made available)

FAANG
Functional Annotation of Animal Genomes

WAGENINGEN
UNIVERSITY & RESEARCH

diagenode

cost
EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

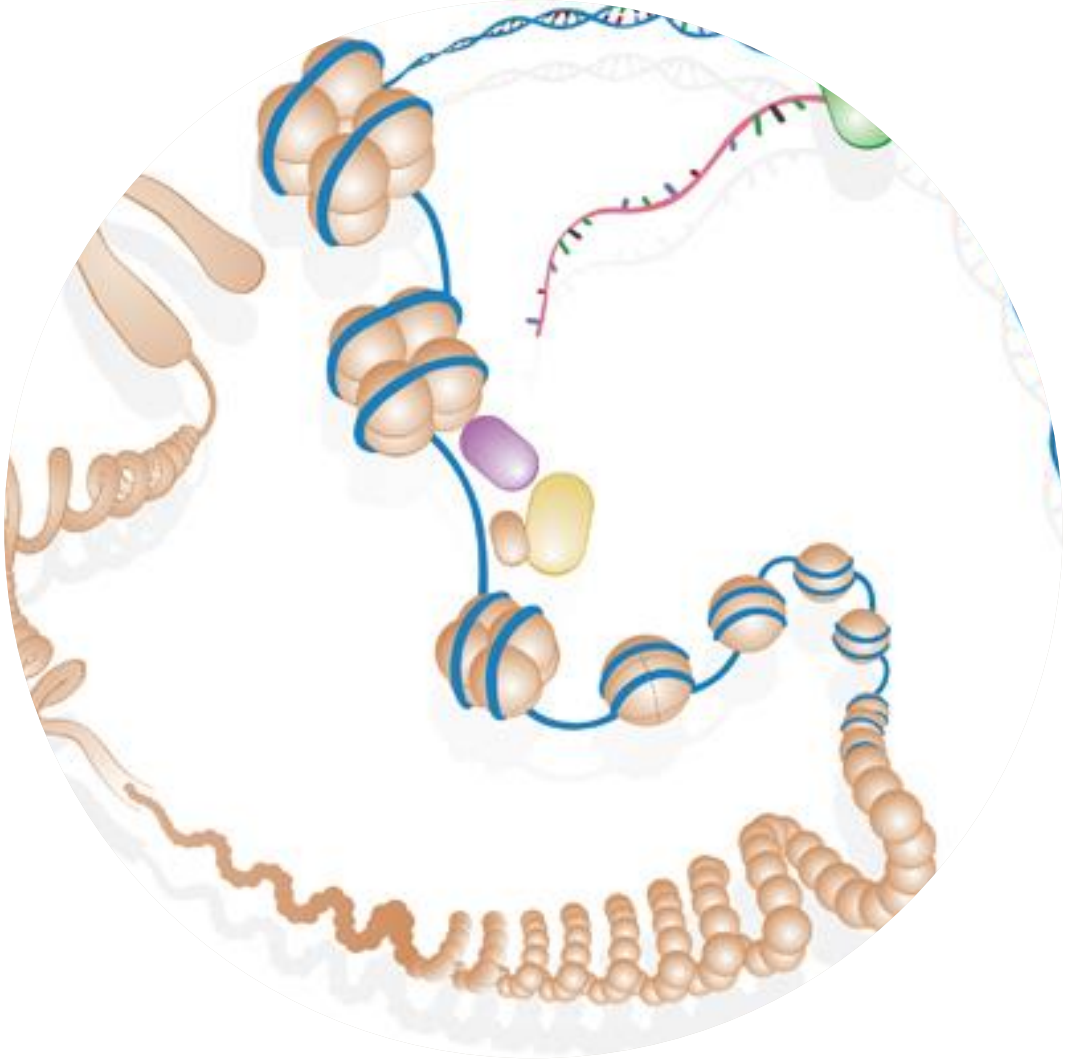
Thanks

Martien Groenen

Ole Madsen

Geoffrey Berguet

diagenode



FAANG

Functional Annotation of Animal Genomes

http://www.faang.com

WAGENINGEN

UNIVERSITY & RESEARCH