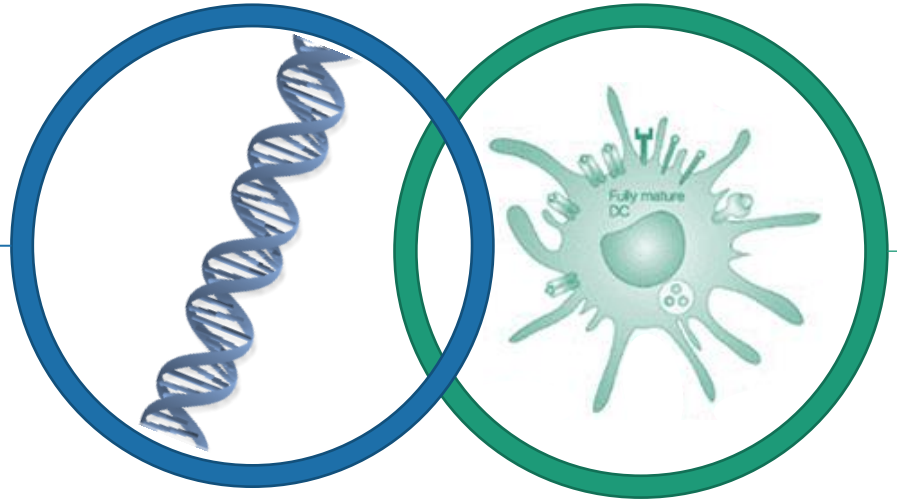


Unraveling the genetics of infectious disease susceptibility in livestock using HT technology.

Russell Fraser, DVM
PhD Candidate
Trainee in Anatomic Pathology

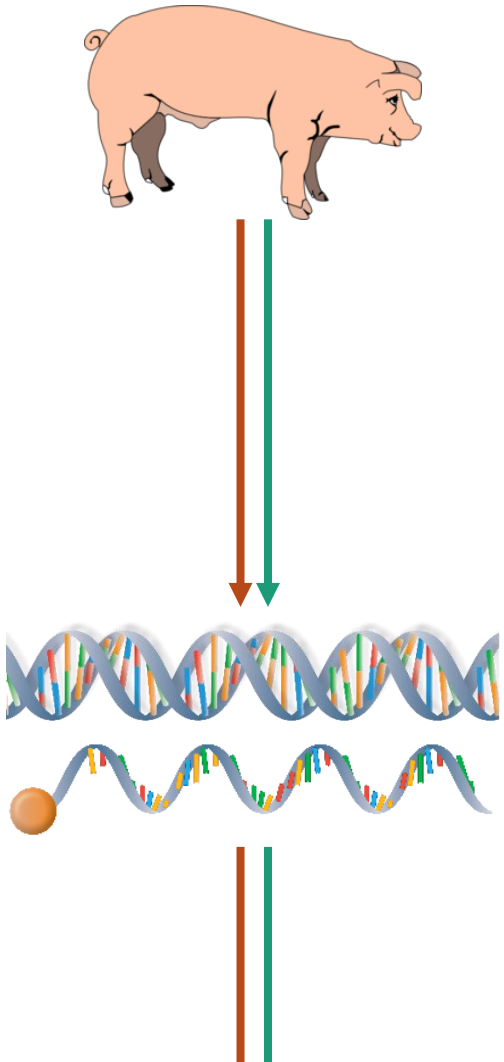
Introduction

Innate immune system



Genetics

Experimental Design

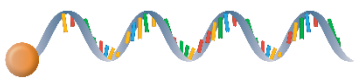


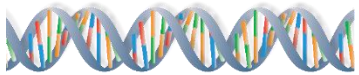
Population

- Healthy
- Market weight

Samples

- Liver
  - DNA
  - RNA



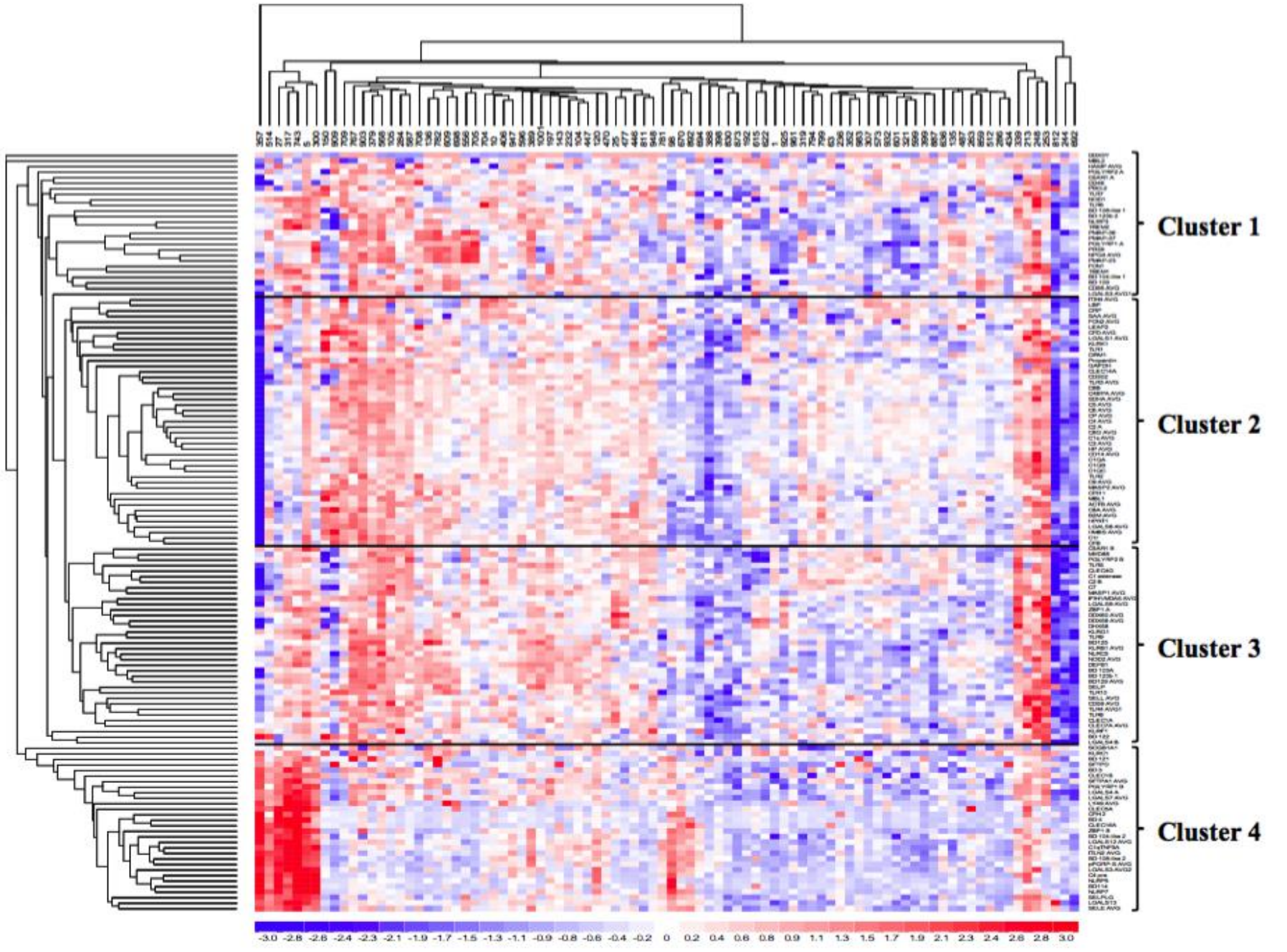


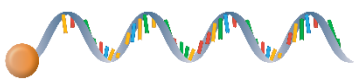
Experiment 1: EXPRESSION

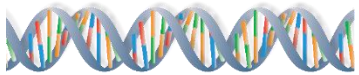
- Microarray on 89 pigs

Experiment 2: MUTATIONS

eQTL Analysis


Cluster 1
Cluster 2
Cluster 3
Cluster 4
-3.0 -2.8 -2.6 -2.4 -2.3 -2.1 -1.9 -1.7 -1.5 -1.3 -1.1 -0.9 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 0.9 1.1 1.3 1.5 1.7 1.9 2.1 2.3 2.4 2.6 2.8 3.0





Experiment 1: EXPRESSION

- Microarray on 89 pigs

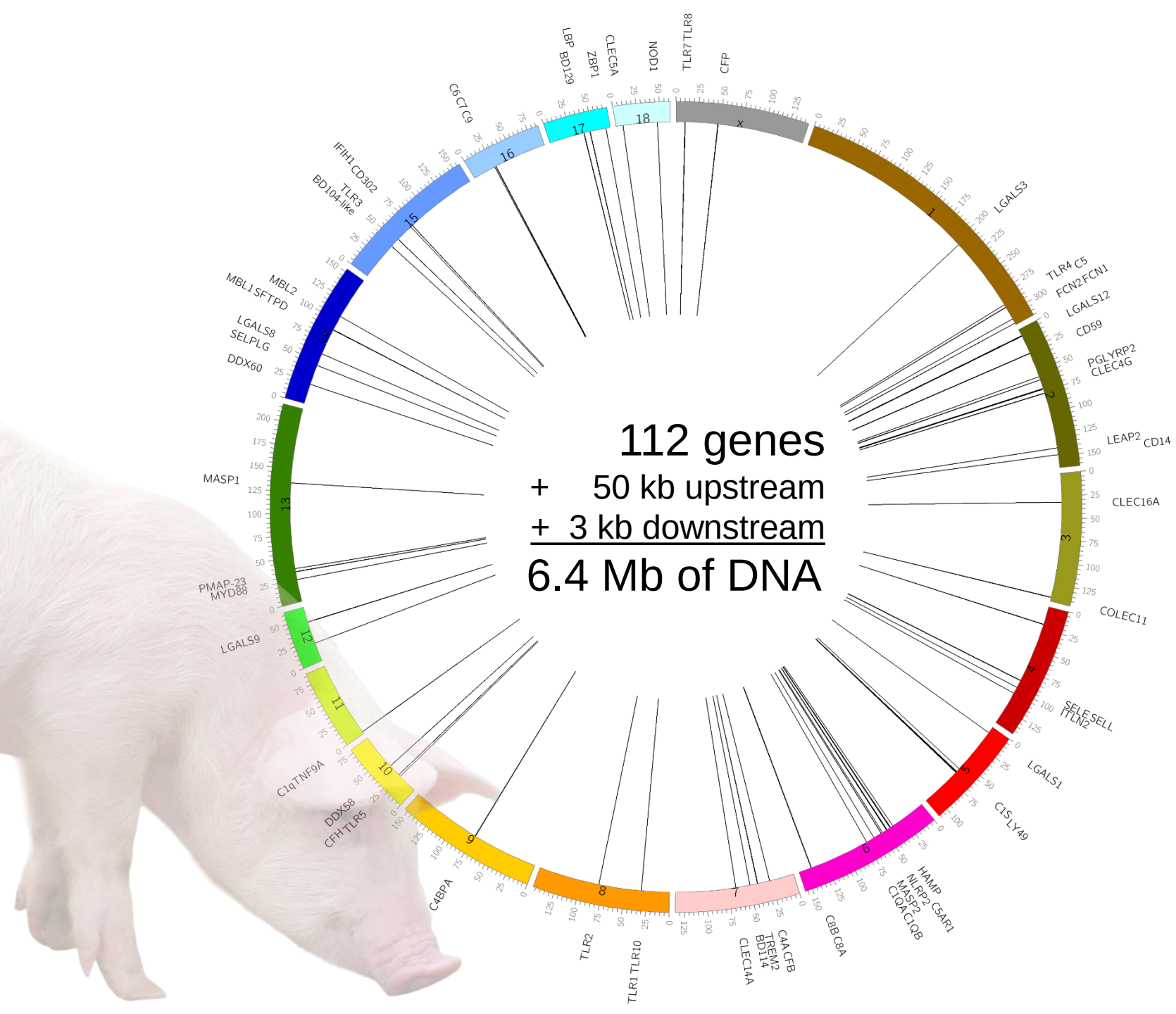
Results

- 112 innate immune genes differentially expressed

Experiment 2: MUTATIONS

- Identify genetic variation
- Top 72 most variable pigs
- Targeted, high throughput re-sequencing

eQTL Analysis


112 genes
+ 50 kb upstream
+ 3 kb downstream
6.4 Mb of DNA
LBP BD129
CLEC5A
ZBP1
NOD1
TLR7 TLR8
CFP
C6 C7 C9
IFIH1 CD302
TLR3
BD104-like
MBL2
MBL1 SFTPD
LGALS8
SELPLG
DDX60
MASP1
PMAP-23
MYD88
LGALS9
C1qTNF9A
DDX58
CFH TLR5
C4BPA
TLR2
TLR1 TLR10
CLEC14A
TREM2
BD114
C4A CFB
C8B C8A
HAMP
NLRP2 C5AR1
MASP2
C1QA C1QB
C1S LY49
LGALS1
SELE SELL
ITLN2
COLEC11
CLEC16A
LEAP2 CD14
PGLYRP2
CLEC4G
CD59
LGALS12
TLR4 C5
FCN2 FCN1
LGALS3

FASTQ

Trimmomatic v0.32 — Trimming

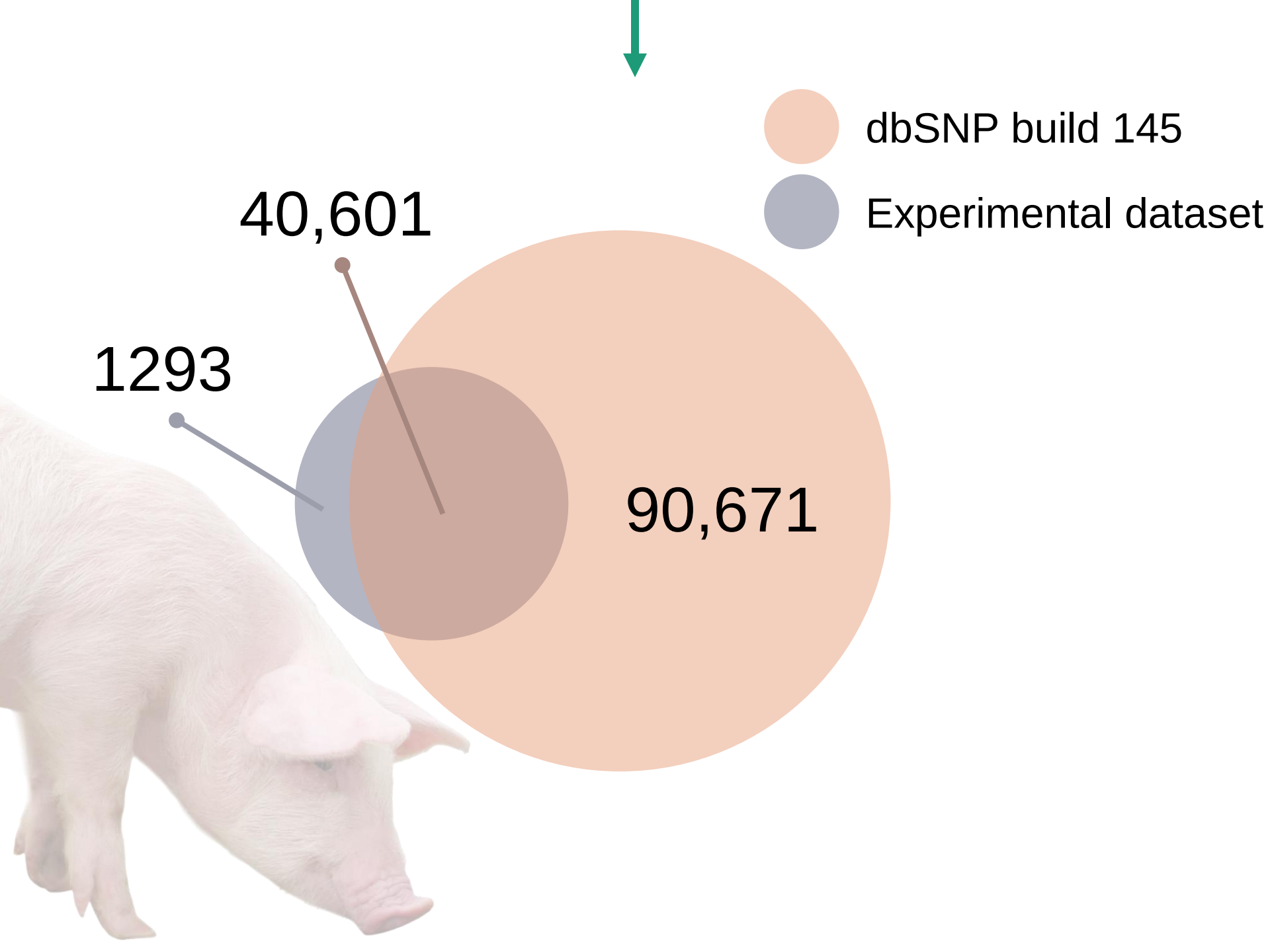
BWA v0.75 — Map reads

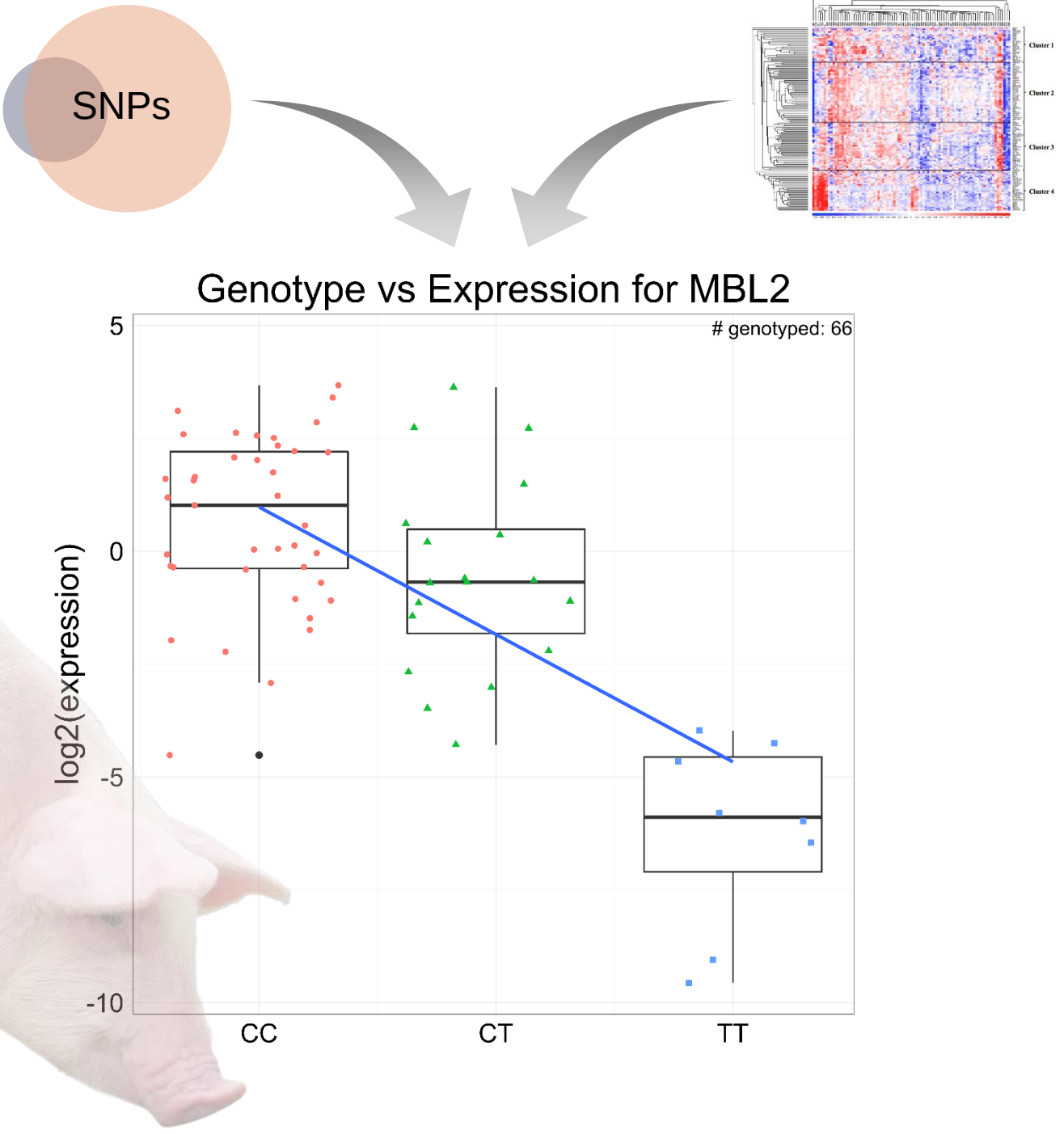
Samtools v1.2 — Remove duplicates

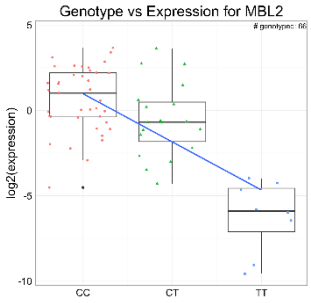
GATK v3.5 — In/Del realignment

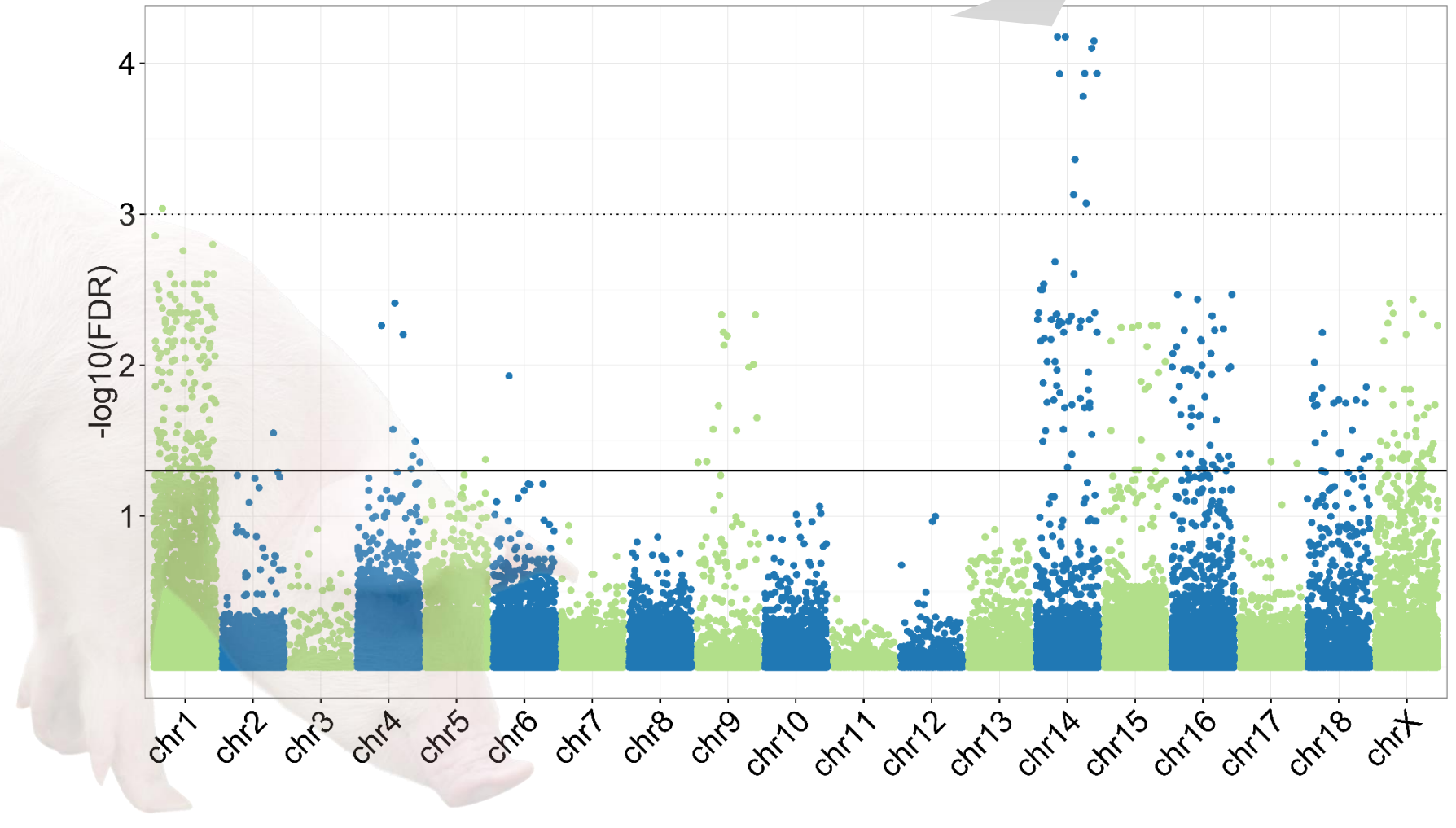
GATK v3.5 — Base score recalibration

GATK v3.5 — Variant calling

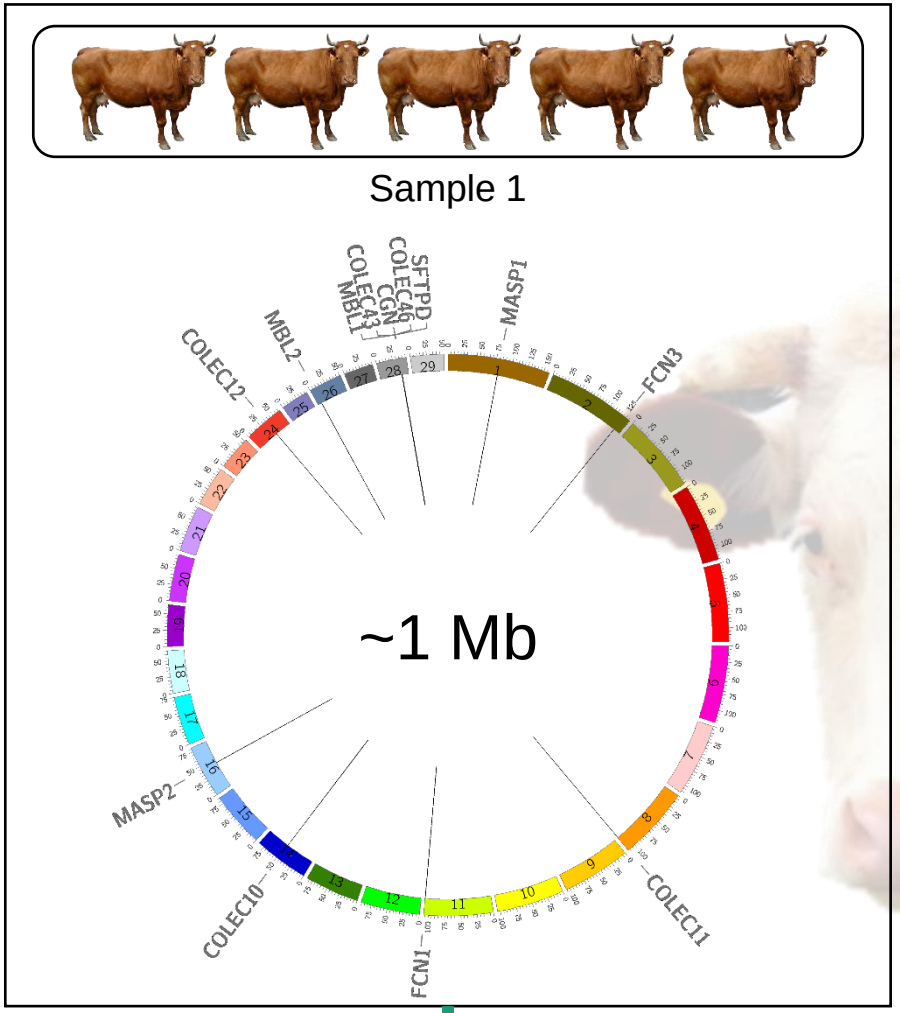

dbSNP build 145
Experimental dataset
40,601
1293
90,671


SNPs
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Genotype vs Expression for MBL2
# genotyped: 66
log2(expression)
5
0
-5
-10
CC
CT
TT


Genotype vs Expression for MBL2
log2(expression)
5
0
-5
-10
CC
CT
TT


-log10(FDR)
4
3
2
1
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chrX

Experimental setup



FASTQ

Trimmomatic v0.32 — Trimming

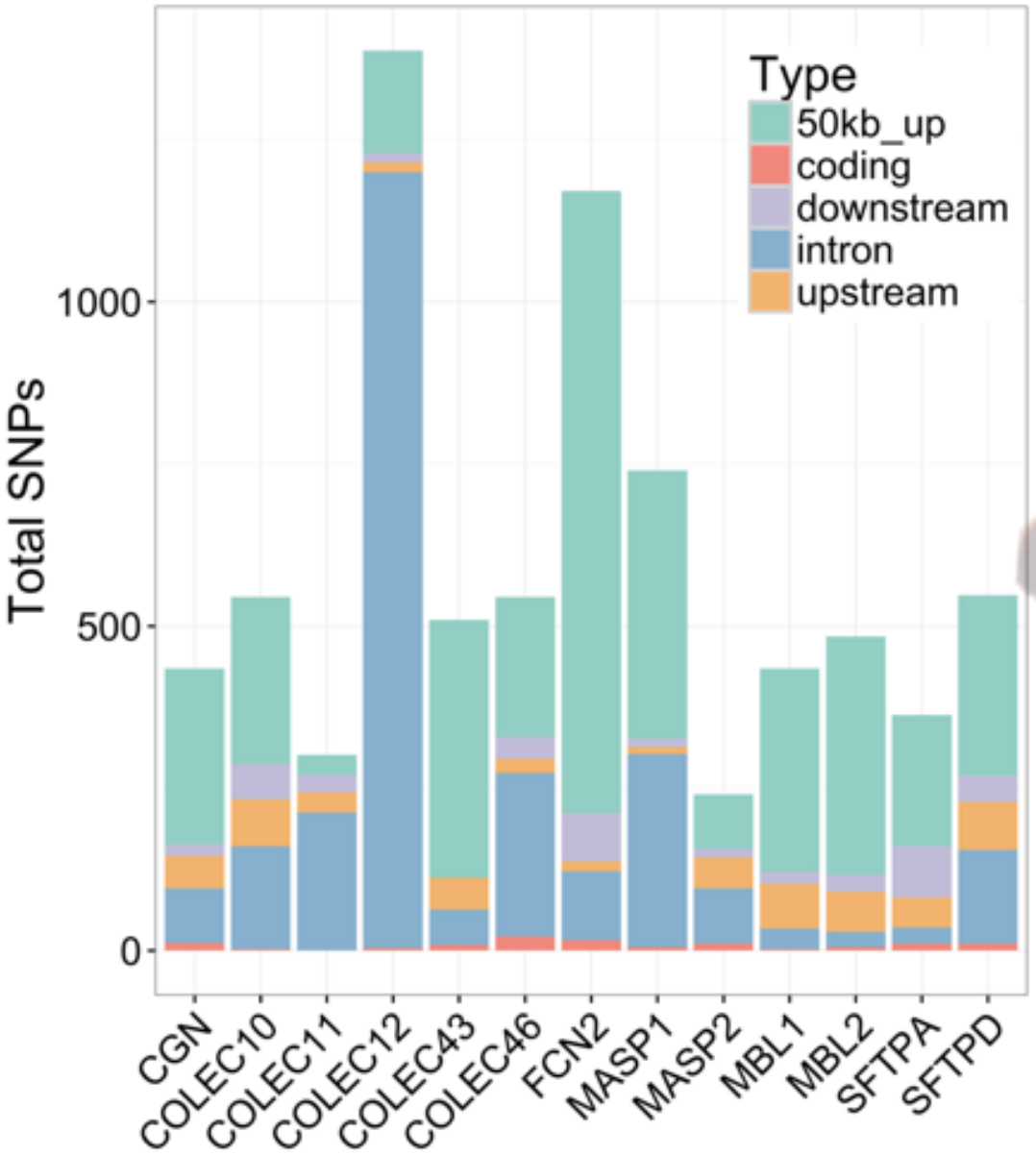
BWA v0.75 — Map reads

Samtools v1.2 — Remove duplicates

GATK v3.5 — In/Del realignment

Base score recalibration

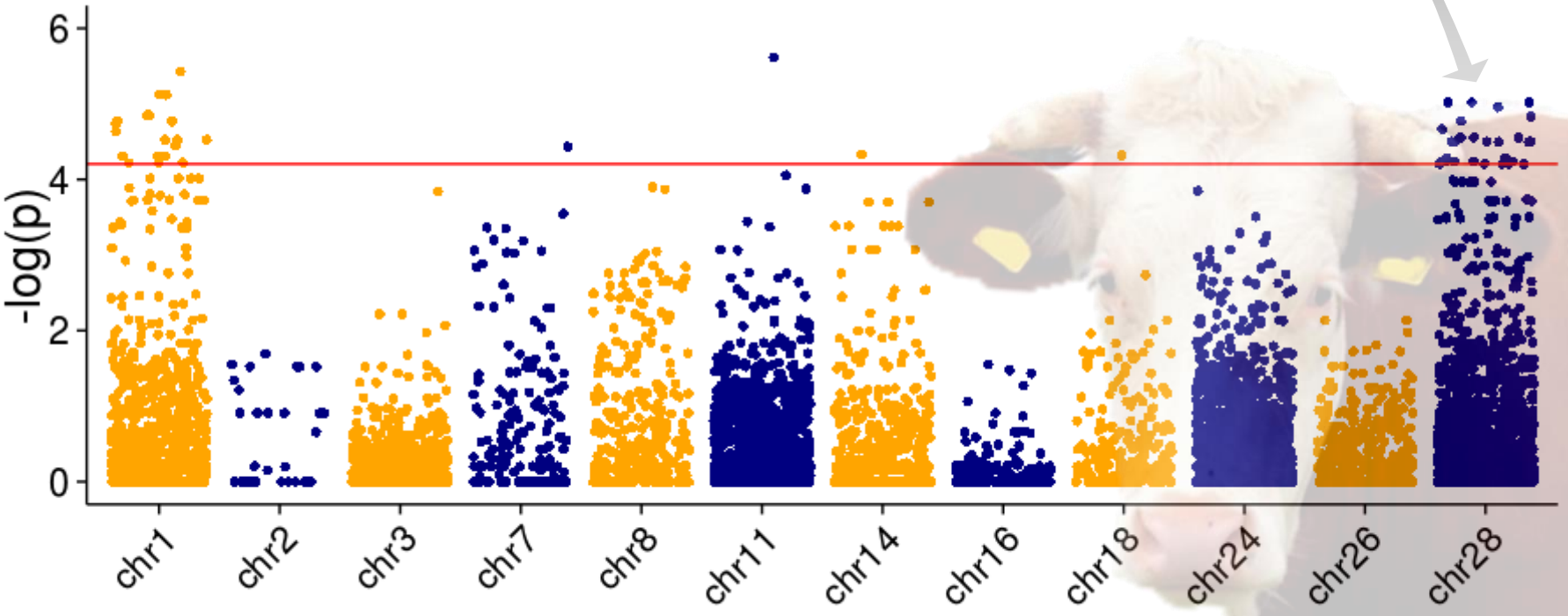
Variant calling


Type
50kb_up
coding
downstream
intron
upstream
Total SNPs
1000
500
0
CGN
COLEC10
COLEC11
COLEC12
COLEC43
COLEC46
FCN2
MASP1
MASP2
MBL1
MBL2
SFTPA
SFTPD

GENE	CHROM	POS	ID	REF	ALT	AA	POLYPHEN
MBL2	chr26	6344919	rs210611099	C	A	Pro42Gln	**1**
CL43	chr28	35721661	.	G	T	Gly229Trp	**1**
CL43	chr28	35721662	.	G	T	Gly229Val	**1**
SFTPA1	chr28	35853355	rs378044423	G	C	Pro97Arg	**1**
CL46	chr28	35700022	rs433242968	T	A	Val348Glu	**0.998**
CL43	chr28	35721644	.	T	A	Leu223His	**0.995**
CL46	chr28	35699941	rs468921789	C	A	Pro321His	**0.993**
FCN2	chr11	106828710	rs385211468	C	T	Thr193Met	**0.991**
FCN2	chr11	106828043	rs380421551	C	T	Arg167Trp	**0.984**
CL46	chr28	35674970	rs463601056	C	A	Pro171Thr	**0.958**
CL46	chr28	35675371	rs383278255	C	T	Pro185Leu	**0.958**
SFTPD	chr28	35820078	rs380240341	C	T	Pro132Ser	**0.958**
SFTPD	chr28	35819626	rs134154426	C	T	Ala75Val	**0.904**
FCN2	chr11	106827968	rs382216843	C	T	Arg142Cys	**0.896**
MASP2	chr16	43463621	rs207667073	G	A	Gly102Ser	**0.85**
CL43	chr28	35718034	rs42967143	A	G	Thr117Ala	**0.659**
CL43	chr28	35718807	rs211678602	G	T	Gln185His	**0.608**
CL46	chr28	35678387	rs378728838	C	G	Leu247Val	**0.508**
CL46	chr28	35675453	rs43119196	G	T	Glu212Asp	**0.455**

	Ref	Alt
Non-diseased	0.8	0.2
Diseased	0.3	0.7

Fisher's Exact Test



Summary

- High-throughput technology identified SNVs that:
  - Impact gene expression
  - Modify protein structure and function
  - Are present at a higher frequency in diseased vs normal animals.
- These represent attractive candidates for further study.

Acknowledgements

- Brandon Lillie
- Hein Snyman
- Jutta Hammermueller
- Luis Arroyo
- Ann Meyer
- John Lumsden
- Tony Hayes