

# Epigenetic Studies in Bovine Mastitis Cell Models

Dissertation  
zur Erlangung des akademischen Grades  
doctor rerum naturalium (Dr. rer. nat.)  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Universität Rostock

vorgelegt von  
Anne Berthold, geb. am 16.12.1987 in Rostock  
aus Halle (Saale)  
Rostock, 10.01.2024



Dieses Werk ist lizenziert unter einer  
Creative Commons Namensnennung 4.0 International Lizenz.

Datum der Einreichung: 10.01.2024

Datum der Promotionsverteidigung: 07.06.2024

Gutachter:

Prof. Dr. Reinhard Schröder

Universität Rostock, Mathematisch-Naturwissenschaftliche Fakultät, Institut  
für Biowissenschaften, Abt. Genetik

Prof. Dr. Christa Kühn

Universität Rostock, Agrar- und Umweltwissenschaftliche Fakultät, Abt.  
Genetik der Krankheitsresistenz / Präsidentin Friedrich-Löffler-Institut  
Bundesforschungsinstitut für Tiergesundheit

Prof. Dr. Jens Tetens

Georg-August-Universität Göttingen, Department für Nutztierwissenschaften,  
Abt. Functional Breeding

## Table of Content

<b>List of Figures .....</b>	<b>IV</b>
<b>List of Tables.....</b>	<b>VI</b>
<b>List of Abbreviations.....</b>	<b>VIII</b>
<b>1 Introduction .....</b>	<b>1</b>
1.1 Relevance of Dairy Cattle in Livestock Farming.....	1
1.2 Mastitis .....	1
1.2.1 <i>E. coli</i> Mastitis .....	2
1.2.2 <i>S. aureus</i> Mastitis.....	2
1.3 Cellular Composition and Specification of the Mammary Gland.....	3
1.4 Mammary Epithelial Cell Models .....	4
1.5 Players of the Immune Response in Mammary Epithelial Cells .....	5
1.5.1 Toll-like Receptor Signaling .....	5
1.5.2 Cytokines .....	6
1.5.3 Chemokines.....	6
1.5.4 Defensins.....	7
1.6 Transcriptomics.....	7
1.7 Epigenomics .....	8
1.7.1 Epigenetics.....	8
1.7.2 Histone Modifications.....	9
1.7.3 Chromatin Immunoprecipitation.....	10
1.8 Epigenetics for Genome Regulation in Livestock Species.....	10
1.9 Epigenetics and Infection.....	11
1.9.1 Host-pathogen Interaction.....	11
1.9.2 Epigenetics in Immunity .....	11
1.9.3 Epigenetics in Immunity of Livestock.....	12
1.10 Study Objectives .....	14
<b>2 Material and Methods.....</b>	<b>15</b>
2.1 Experimental Design .....	15
2.1.1 Detection of Histone Modifications and Immune Gene Expression in MACT .....	16
2.1.2 Genome-wide Effect of Mastitis Pathogens on Epigenome and Transcriptome in MACT and pbMEC .....	16
2.2 Chemicals, Oligonucleotides, Materials, and Equipment .....	17
2.3 Mammary Epithelial Cell Models .....	18
2.4 Mastitis Pathogen Preparation.....	18
2.5 Cell Biological Techniques.....	19

2.5.1 Primary Cell Extraction.....	19
2.5.2 Cell Culture .....	20
2.5.3 Cell Thawing and Freezing .....	21
2.5.4 Selective Trypsinization.....	21
2.5.5 Cell Exposure Experiments.....	22
<b>2.6 Molecular Biological Technique.....</b>	<b>22</b>
2.6.1 Nucleic Acid Isolation from Cells.....	22
2.6.2 Determination of the Concentration and Purity of Nucleic Acids .....	23
2.6.3 Polymerase Chain Reaction .....	24
2.6.4 Quantitative PCR .....	27
2.6.5 Agarose Gel Electrophoresis.....	31
2.6.6 Purification of DNA from Agarose Gel .....	32
2.6.7 Competent Cells.....	32
2.6.8 Cloning.....	32
2.6.9 Restriction Enzyme Digestion of Plasmid DNA.....	34
2.6.10 DNA Sanger Sequencing .....	34
<b>2.7 N-ChIP.....</b>	<b>34</b>
2.7.1 Cell Harvest and Enzymatic Shearing.....	34
2.7.2 Chromatin Immunoprecipitation .....	35
<b>2.8 X-ChIP.....</b>	<b>36</b>
2.8.1 Fixation and Cell Harvest.....	36
2.8.2 Cell Lysis and Chromatin Shearing .....	37
2.8.3 Magnetic Immunoprecipitation .....	37
2.8.4 Elution and Decross-linking.....	38
2.8.5 DNA Purification .....	38
<b>2.9 Library Preparation and Next-Generation Sequencing.....</b>	<b>39</b>
2.9.1 Library Preparation for RNA-seq.....	39
2.9.2 Library Preparation for ChIP-seq.....	39
<b>2.10 Computational Methods.....</b>	<b>39</b>
2.10.1 RNA-seq Data Analysis Pipeline .....	39
2.10.2 ChIP-seq Data Analysis Pipeline .....	40
2.10.3 R-based Analysis and Visualization.....	41
<b>3 Results.....</b>	<b>42</b>
<b>3.1 Histone Modification in MAC-T in Response to Mastitis Pathogens.....</b>	<b>42</b>
3.1.1 Expression of Immune Response Genes in Response to <i>E. coli</i> .....	42
3.1.2 Divergent Expression of Immune Response Genes in Response to <i>E. coli</i> and <i>S. aureus</i> .....	43
3.1.3 Optimization of N-ChIP Procedure in MAC-T.....	44
3.1.4 N-ChIP-qPCR in Promoter Regions of Immune Response Genes after Exposure to Mastitis Pathogens .....	45
<b>3.2 Genome-wide Transcriptome Sequencing in Bovine Mastitis Cell Models.....</b>	<b>48</b>
3.2.1 Comparison of MAC-T and pbMEC.....	48
3.2.2 Effect of Pathogen on Transcriptome of MAC-T .....	56
3.2.3 Effect of Pathogen on Transcriptome of pbMEC.....	58
<b>3.3 Genome-wide Sequencing in Bovine Mastitis Cell Models Targeting the Epigenome .....</b>	<b>62</b>
3.3.1 X-ChIP Establishment and Verification.....	63
3.3.2 X-ChIP-seq Exploration and Quality Check.....	67
3.3.3 Differential Binding Analysis and Annotation of X-ChIP-seq Peaks.....	70



---

<b>3.4 Data Integration of ChIP-seq and RNA-seq.....</b>	<b>75</b>
3.4.1 MAC-T versus pbMEC control.....	76
3.4.2 pbMEC <i>E. coli</i> versus control .....	80
<b>4 Discussion .....</b>	<b>85</b>
<b>4.1 Methodological Aspect of ChIP .....</b>	<b>85</b>
4.1.1 Fixation, Douncing, and Shearing.....	85
4.1.2 Choice of Antibody .....	86
4.1.3 X-ChIP-qPCR .....	86
<b>4.2 Effect of Pathogen Challenge on Immune Response Genes.....</b>	<b>87</b>
<b>4.3 Measuring Histone Modifications in MAC-T .....</b>	<b>89</b>
<b>4.4 RNA-seq Identified a Strong Discrepancy in Commonly Used Cell Models.....</b>	<b>89</b>
<b>4.5 Response to Pathogens Assessed by RNA-seq.....</b>	<b>91</b>
<b>4.6 ChIP-seq Data Analysis.....</b>	<b>92</b>
4.6.1 Sequencing Depth and Read Mapping .....	92
4.6.2 Reproducibility of ChIP-seq .....	93
4.6.3 ChIP-seq Peak Calling and Differential Binding Analysis .....	93
<b>4.7 Genome-wide Sequencing Maps Epigenetic Profile of MAC-T and pbMEC .....</b>	<b>94</b>
<b>4.8 Usability of the Cell Models.....</b>	<b>96</b>
<b>4.9 Future work.....</b>	<b>96</b>
<b>Summary .....</b>	<b>99</b>
<b>Zusammenfassung .....</b>	<b>100</b>
<b>References .....</b>	<b>101</b>
<b>Appendix .....</b>	<b>A-1</b>
<b>A. Supplementary Materials and Methods .....</b>	<b>A-1</b>
<b>B. Supplementary Results .....</b>	<b>A-12</b>
<b>Acknowledgment .....</b>	<b>A-32</b>
<b>Eidesstattliche Erklärung .....</b>	<b>A-34</b>

## List of Figures

Figure 1-1: The bovine udder's physical barriers and representative cellular components. ....	3
Figure 1-2: Main epigenetic mechanisms.....	9
Figure 2-1: Experimental design for detecting transcriptomic and epigenomic changes in bovine mammary epithelial cell models.....	15
Figure 2-2: Origin of primary mammary epithelial cells for in vitro studies .....	19
Figure 2-3: DNA purification from agarose gel .....	32
Figure 2-4: Cloning procedure. ....	33
Figure 2-5: Data analysis pipeline for RNA-seq. ....	40
Figure 2-6: Data analysis pipeline for ChIP-seq. ....	41
Figure 3-1: Expression level of immune genes after challenge with <i>E. coli</i> <sub>1303</sub> in MAC-T.....	42
Figure 3-2: Expression level of immune genes after challenge with <i>E. coli</i> <sub>1303</sub> or <i>S. aureus</i> <sub>1027</sub> in MAC-T. ....	43
Figure 3-3: Optimization steps for N-ChIP in MAC-T. Lysis and shearing required testing to establish N-ChIP. ....	44
Figure 3-4: N-ChIP-qPCR in MAC-T after <i>E. coli</i> <sub>1303</sub> contact.....	46
Figure 3-5: N-ChIP-qPCR in MAC-T after pathogen challenge. ....	47
Figure 3-6: Principal component analysis for RNA-seq data in mammary epithelial cells. ....	49
Figure 3-7: Principal component analysis for RNA-seq data in individual mammary epithelial cells. ....	49
Figure 3-8: Visualization of differential gene expression.. ....	50
Figure 3-9: Expression of top regulated genes across sample groups.....	51
Figure 3-10: FPKM values for selected immune response genes from RNA-seq.....	52
Figure 3-11: Top 50 GO terms for DE genes of comparison of MAC-T and pbMEC control groups. ....	54
Figure 3-12: Relationship between top 50 GO terms from functional analysis of DE genes of the comparison of MAC-T and pbMEC control groups.....	55
Figure 3-13: Category netplot for functional analysis of DE genes of the comparison of MAC-T and pbMEC control groups. ....	56
Figure 3-14: Volcano plots for MAC-T comparing pathogen versus control-challenged cells. ....	57
Figure 3-15: Gene Ontology (GO) Analysis for biological processes for DEGs between MAC-T <i>E. coli</i> <sub>1303</sub> versus Control. ....	58
Figure 3-16: Volcano plots for pbMEC comparing pathogen versus control-challenged cells.....	59
Figure 3-17: Functional Analysis of DEGs between <i>E. coli</i> <sub>1303</sub> versus control in pbMEC. ....	61
Figure 3-18: Functional Analysis of DEGs between <i>S. aureus</i> <sub>1027</sub> and control in pbMEC. ....	62
Figure 3-19: Shearing optimization experiment in bovine model cells.....	63

Figure 3-20: X-ChIP control-PCR to determine control regions and antibody amount for MAC-T and pbMEC. ....	65
Figure 3-21: X-ChIP control qPCR in bovine model cells.....	66
Figure 3-22: X-ChIP control qPCR in MAC-T.....	67
Figure 3-23: Plotfingerprints generated by deepTools for ChIP assays in two bovine cell models....	69
Figure 3-24: Plot correlation analysis for ChIP-seq in bovine model cells based on genomic reads..	69
Figure 3-25: PCA of ChIP experiments for control-challenged samples.....	70
Figure 3-26: Differential Binding Analysis for H3K4me3 peaks in MAC-T versus pbMEC in control samples.....	72
Figure 3-27: Differential Binding Analysis for H3K27me3 peaks in MAC-T versus pbMEC in control samples.....	73
Figure 3-28: Differential Binding Analysis for H3K4me3 peaks in pbMEC comparing <i>E. coli</i> <sub>1303</sub> and control samples.....	74
Figure 3-29: ChIP-seq peak annotation of H3K4me3 and H3K27me3 in MAC-T and pbMEC. ....	75
Figure 3-30: RNA-seq and ChIP-seq data integration for MAC-T versus pbMEC control samples..	77
Figure 3-31: Example of histone modifications detected near OLR1 gene. ....	77
Figure 3-32: Functional analysis for 118 genes under epigenetic regulation.....	78
Figure 3-33: Functional analysis for 3999 genes under epigenetic regulation.....	79
Figure 3-34: Functional analysis for 22 genes under epigenetic regulation.....	80
Figure 3-35: Functional analysis for 422 genes under epigenetic regulation.....	80
Figure 3-36: Data integration of RNA-seq and ChIP-seq for comparing <i>E. coli</i> <sub>1303</sub> versus control challenged pbMEC.....	82
Figure 3-37: Functional analysis for 15 genes under epigenetic regulation.....	83
Figure 3-38: Functional analysis for 394 genes from RNA-seq.....	84
Figure B-1: Gene regulation of MAC-T after <i>E. coli</i> <sub>1303</sub> challenge.....	A-12
Figure B-2: Gene regulation of MAC-T after <i>E. coli</i> <sub>1303</sub> and <i>S. aureus</i> <sub>1027</sub> challenge.....	A-12
Figure B-3: Gene Ontology (GO) Analysis for Cellular Components for differentially expressed genes (DEGs) between MAC-T and pbMEC.....	A-13
Figure B-4: Relationship between the GO terms for Cellular Components for differentially expressed genes (DEGs) between MAC-T and pbMEC.....	A-14
Figure B-5: Top five GO terms for Cellular Components with associated up and down-regulated differentially expressed genes (DEGs) between MAC-T and pbMEC. ....	A-15
Figure B-6: Gene Ontology (GO) Analysis for Molecular Function for differentially expressed genes (DEGs) between MAC-T and pbMEC.....	A-16
Figure B-7: Relationship between the GO terms for Molecular Function for differentially expressed genes (DEGs) between MAC-T and pbMEC.....	A-17

Figure B-8: Top five GO terms for Molecular Function with associated up and down-regulated differentially expressed genes (DEGs) between MAC-T and pbMEC..	A-18
Figure B-9: Gene Ontology (GO) Analysis for Cellular Component for DEGs between MAC-T <i>E. coli</i> <sub>1303</sub> versus Control.	A-19
Figure B-10: Gene Ontology (GO) Analysis for Molecular Functions for DEGs between MAC-T <i>E. coli</i> <sub>1303</sub> versus Control.	A-20
Figure B-11: Gene Ontology (GO) Analysis for Cellular Components for DEGs between pbMEC <i>E. coli</i> <sub>1303</sub> versus Control.	A-21
Figure B-12: Gene Ontology (GO) Analysis for Molecular Function for DEGs between pbMEC <i>E. coli</i> <sub>1303</sub> versus Control.	A-22
Figure B-13: Gene Ontology (GO) Analysis for Cellular Components for DEGs between pbMEC <i>S. aureus</i> <sub>1027</sub> versus Control.	A-23
Figure B-14: Gene Ontology (GO) Analysis for Molecular Functions for DEGs between pbMEC <i>S. aureus</i> <sub>1027</sub> versus Control.	A-24
Figure B-15: Correlation heatmap.	A-29
Figure B-16: PCA of ChIP-seq for challenged cells.	A-30
Figure B-17: Detailed ChIP-seq peak annotation of H3K4me3 and H3K27me3 in MAC-T and pbMEC.	A-31

## List of Tables

Table 2-1: RPMI 1640 complete media composition.	17
Table 2-2: DMEM complete media composition.	17
Table 2-3: Cows information for tissue sampling.	18
Table 2-4: PCR reaction mixture with Taq polymerase.	24
Table 2-5: Cyclor Program used for Taq PCR.	24
Table 2-6: Primers for measuring RNA purity.	25
Table 2-7: PCR reaction mix for PFU polymerase.	25
Table 2-8: Cyclor Program used for semi-nested PFU PCR.	25
Table 2-9: Cyclor Program used for Taq semi-nested PCR.	26
Table 2-10: Cyclor Program used for Colony PCR.	26
Table 2-11: List of cDNA primers used for the reverse transcription.	27
Table 2-12: Enzyme master mix of the reverse transcription reaction.	27
Table 2-13: Reference list for counting plasmids for gene expression measurement by qPCR.	27
Table 2-14: Reaction Master Mixture for LightCycler® 2 capillary Real-Time PCR system (Roche)	29
Table 2-15: Reaction Master Mixture for LightCycler® 480 (Roche).	29
Table 2-16: qPCR program for LightCycler® 2 capillary Real-Time PCR system (Roche)	29

Table 2-17: qPCR program for LightCycler® 480 (Roche Life Science) .....	29
Table 2-18: Reaction Master Mixture for LightCycler® 96 (Roche Life Science) .....	30
Table 2-19: X-ChIP control-qPCR primer. ....	31
Table 2-20: qPCR program for LightCycler® 96 (Roche Life Science) .....	31
Table 2-21: Antibodies used for N-ChIP experiments. ....	35
Table 2-22: Antibodies used for X-ChIP experiment.....	38
Table 2-23: Conditions used for X-ChIP experiment. ....	38
Table 2-24: Parameters for Peak calling. ....	41
Table 3-1: Summary of the conditions for the ChIP experiment used for the bovine cell models. ....	64
Table 3-2: Data integration of ChIP-seq and RNA-seq for comparing MAC-T and pbMEC.....	76
Table 3-3: Data integration of ChIP-seq and RNA-seq for comparing <i>E. coli</i> <sub>1303</sub> versus control-challenged pbMEC.....	81
Table 3-4: Gene IDs for 15 common genes under epigenetic regulation.....	82
Table 3-5: Gene IDs for 9 genes under epigenetic regulation.....	83
Table A-1: First round of semi-nested PCR for amplifying DNA insert spanning immune response gene fragment for cloning. ....	A-1
Table A-2: Second round of semi-nested PCR for amplifying DNA insert spanning immune response gene fragment for cloning. ....	A-1
Table A-3: N-ChIP-qPCR Primers.....	A-1
Table A-4: Material list cell culture. ....	A-4
Table A-5: Consumables for molecular biological work. ....	A-4
Table A-6: Kit and reagents used for molecular biological work. ....	A-4
Table A-7: Material List for microbiological work.....	A-5
Table A-8: Kits and competent cells for microbiological work. ....	A-5
Table A-9: Substances and liquid components for cell culture work.....	A-5
Table A-10: Substances for microbiological work.....	A-6
Table A-11: Size standards for nucleic acids. ....	A-7
Table A-12: Restriction enzymes. ....	A-7
Table A-13: Devices and accessories in general use in the laboratory. ....	A-7
Table A-14: Cell culture equipment.....	A-7
Table A-15: Molecular biological equipment. ....	A-8
Table B-1: Establishment of FAANG core assays for bovine cell models.. ....	A-25
Table B-2: RNA-seq statistics.). ....	A-26
Table B-3: ChIP-seq statistics. ....	A-26
Table B-4: Average of ChIP peaks. ....	A-28

## List of Abbreviations

%RSD	percentage relative standard deviation
AMP	antimicrobial peptide
BHI	brain-heart-infusion
BNBD5	$\beta$ -defensin five
bp	base pair
BP	biological processes
BR	broad range
BSA	bovine serum albumin
CAVIN1	caveolae associated protein one
CBP	CREB-binding protein
CC	cellular compartments
CCL20	chemokine ligand 20
CD14	cluster of differentiation 14
cDNA	complementary DNA
CDU	collagenase degrading units
ChIP	chromatin immunoprecipitation
CL	cell line
CLIC1	chloride intracellular channel one
CO <sub>2</sub>	carbon dioxide
Ct	threshold cycle
CTCF	CCCTC-binding factor
CXCL3	C-X-C motif chemokine ligand three
CXCL8 / IL8	interleukin eight
CYP1A1	cytochrome P450 family one subfamily A member one
CYP1B1	cytochrome P450 family one subfamily B member one
DEGs	differentially expressed genes
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
DNA	desoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
DOTL1	DOT1-like, histone H3 methyltransferase
dsDNA	double-stranded DNA
<i>E. coli</i>	<i>Escherichia coli</i>
EBV	Epstein–Barr virus
ECM	extracellular matrix
EcoRI	<i>E. coli</i> RY13 restriction enzyme I

---

EDTA	ethylenediamine tetra acetic acid
EGFR	epidermal growth factor receptor
F2RL1	F2R-like trypsin receptor one
FA buffer	formaldehyde buffer for RNA gels
FAANG	Functional Annotation of Animal Genomes
FBN	Research Institute for Farm Animal Biology
FBS	fetal bovine serum
FC	fold change
FCS	fetal calf serum
FN1	glycoprotein fibronectin one
FOXP3	forkhead box P3
FPKM	mean values of the normalized expression counts during RNA-seq analysis
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
gDNA	genomic DNA
GO	gene ontology
H <sub>2</sub> O	water
H3K14Ac	histone three lysine 14 acetylation
H3K18Ac	histone three lysine 18 acetylation
H3K27me3	histone three lysine 27 trimethylation
H3K4	histone three lysine four
H3K4me3	histone three lysine four trimethylation
H3K9me3	histone three lysine nine trimethylation
HBSS	Hanks' buffered salt solution
HDAC	histone deacetylase
HS	high sensitivity
IGFBP2	insulin-like growth factor binding protein two
IgG	immunoglobulin G
IL6	interleukin 6
IP	immunoprecipitation
IPTG	isopropyl-b-D-thiogalactopyranoside
IRF	interferon regulatory factors
JARID2	histone methyltransferase jumonji, AT rich interactive domain 2
JMJD3	Jumonji domain-containing protein-3
KLF4	Krüppel-like factor four
KRT17	cytokeratin 17
LAP	lingual antimicrobial peptide
lncRNA	long non-coding RNA

---

LPS	lipopolysaccharide
LSD1	lysine-specific demethylase one
mAB	monoclonal antibody
MAC-T	mammary gland alveolar cells stably transfected with an SV40 large T antigen
MEC	mammary epithelial cells
MF	molecular functions
miRNA	micro RNA
MNase	micrococcal nuclease
mRNA	messenger RNA
MyD88	myeloid differentiation factor 88
N-ChIP	native chromatin immunoprecipitation
ncRNA	non-coding RNAs
NF- $\kappa$ B	Nuclear factor kappa-light-chain-enhancer of activated B cells
NO	nitric oxide
NOS2A / iNOS	inducible nitric oxide synthase
Nrf2	nuclear factor erythroid 2-related factor two
OD	optical density
P/S	penicillin / streptomycin
pAB	polyclonal antibody
p <sub>adj</sub>	adjusted p-value according to the Benjamini & Hochberg method
PAMP	pathogen-associated molecular
pbMEC	primary bovine mammary epithelial cells
PC	primary cells
PC1	first principal component
PC2	second principal component
PCA	principal component analysis
PCR	polymerase chain reaction
PFU polymerase	<i>Pyrococcus furiosus</i> polymerase
PIC	protease inhibitor cocktail
PMSF	phenylmethylsulfonyl fluoride
PRSS1	protease, serine one
qPCR	quantitative polymerase chain reaction
RIN	RNA integrity number
RNA	ribonucleic acid
RNA-seq	ribonucleic acid sequencing
RPMI	culture medium developed by Moore et al. at Roswell Park Memorial Institute
rRNA	ribosomal RNA



---

RSAD2	radical S-adenosyl methionine domain-containing protein two
RT	room temperature
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
S100A9	calgranulin B
ScaI	<i>Streptomyces caespitosus</i> restriction enzyme I
SCC	somatic cell count
snRNA	small nuclear RNA
SOB	super optimal broth
SOCS3	suppressor of cytokine signaling three
STAT	signal transducer and activator of transcription
STAT6	signal transducer and activator of transcription six
STK10	serine/threonine-protein kinase 10
STSM	short-term scientific mission
TAE	tris-acetate-EDTA
TFIID	transcription factor II D
TIRAP	TIR-associated protein
TLR	toll-like receptor
TNF- $\alpha$	tumor necrosis factor-alpha
TRAF6	tumor necrosis factor receptor-associated factor 6
tRNA	transfer RNA
TSH2B	testis-specific histone H2B variant
TSS	transcription start site
TTS	transcription termination site
TY agar	tryptone yeast
VIM	type III intermediate filament protein vimentin
X-ChIP	cross-linked chromatin immunoprecipitation
X-Gal	5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside
XhoI	<i>Xanthomonas holcicola</i> restriction enzyme I



# 1 Introduction

The introductory chapter reviews the broader context and significance of mastitis in cattle and cell models used to study the inflammatory disease of the udder and host immune response. It introduces the terminology of transcriptomics and epigenomics used throughout the dissertation. The role of epigenetics in infectious disease, particularly in livestock, is detailed before outlining the research objectives of this study, which focus on targeted and genome-wide gene expression thought to be regulated by epigenetic mechanisms such as histone modifications.

## 1.1 Relevance of Dairy Cattle in Livestock Farming

Throughout history, humans have lived in close relationships with animals, such as companion, wild, and livestock animals. Cattle are among the variety of animals domesticated from humans 11,000 years ago<sup>1</sup> in some parts of the world. Cattle served humans as a valuable resource for nutrition, leather production, and as transportation animal. Nowadays, meat and dairy products are in high consumer demand and are considered an essential part of a healthy and balanced diet by many of the world's population. Globally, traditional and modern cattle production provides livestock products to meet rising consumer demand. Dairy cattle are very efficient at feed conversion. The Holstein dairy breed is valued for its high milk yield, while other dairy breeds distinguish by higher fat, protein, and solids content in milk<sup>2</sup>. However, due to the high milk yield, fat, protein, and solids are also above average for Holstein cattle<sup>2</sup>. Dairy farming plays a significant societal role, providing animal products and employment. Farming practices impact animals, humans, and ecosystems, including their health, which are highly interdependent<sup>3</sup>. Therefore, the shift of cattle production toward sustainability, including management strategies and animal welfare, is becoming increasingly important in dairy farming<sup>4</sup>. Dairy cattle can only perform at their total capacity in good health. The cow's performance can be affected by some common production diseases<sup>5</sup> caused by bacterial and viral infections. For economical and societal reasons, actions are needed to ensure animal products' safety and health, which is closely related to animal welfare. Undetected and untreated disease outbreaks are a significant concern for farmers and the health of the animals and consumers.

## 1.2 Mastitis

One of the most critical diseases is mammary gland inflammation, known as mastitis. Cows suffering from mastitis produce less milk and carry the risk of deterioration of the general health of the animal and the herd, which can mean economic losses for farmers<sup>6</sup>. The milk cannot be sold due to bacterial or antibiotic contamination, and animals are sorted out due to performance reductions associated with lactation problems<sup>7</sup>. Costs also arise from treatment, veterinary services, and livestock losses<sup>8</sup>. This economic impact of the livestock disease is a global problem in dairy farming<sup>9,10</sup>.

Measures are in place so the cow's milk content is controlled very carefully and densely, e.g., somatic milk cell count (SCC) or microbial culture methods provide diagnostic indicators for mastitis<sup>11</sup>. In addition, adapted management strategies improve herd hygiene and udder disease<sup>10</sup>. Furthermore, molecular and genomic mastitis research has led to increased knowledge in the last decades through *in*

*vivo* and *in vitro* models. This research, among others, can help improve and manage mastitis in dairy cows.

The anatomy of the mammary gland differs between mammalian species<sup>12</sup>. Cattle possess an udder with four separate glands, each with one ductal system and one teat<sup>13</sup>. The cow's mastitis represents an inflammation of the mammary gland in its entire milk-forming, storing, and draining parts, referred to as the ductal system<sup>12</sup>. Each udder quarter can be affected by mastitis independently of the others. The trigger for the disease can be exposure of the udder to various stimuli, such as pathogen infection, trauma, or, less often, injury<sup>14</sup>. In addition, several host and environmental factors determine the diversity of a course of mastitis<sup>14</sup>. The course of mastitis can be acute or chronic, clinical, or subclinical. Among various microbial species<sup>15</sup> that cause mastitis in livestock animals, such as cattle<sup>16,17</sup>, sheep<sup>18</sup>, or pig<sup>19,20</sup>, the highly relevant bacteria are *Escherichia coli* (*E. coli*) and *Staphylococcus aureus* (*S. aureus*). These two bacterial species show an opposite clinical manifestation in the bovine host.

### 1.2.1 *E. coli* Mastitis

*E. coli* is a pathogen that considerably causes mastitis; some strains are more likely to cause the disease<sup>21</sup>. *E. coli* is a gram-negative, rod-shaped bacterium typically found in the intestines of cattle. As an environmental source of mastitis, the coliform bacteria primarily occur in bedding material, excretions, and soil<sup>14,22,23</sup>. The pathogen *E. coli* triggers local inflammation when entering the udder via the teat. This infection usually leads to acute up to highly acute mastitis with clinical symptoms. In the case of excessive cytokine secretion due to *E. coli* infection, a severe systemic inflammatory response syndrome can occur in the cow<sup>24</sup>, which may also result in the death of the animal<sup>25</sup>. Certain traits in the udder and milk<sup>26</sup> accompany clinical mastitis. The inflamed udder or udder quarters appear swollen and heated, causing pain to the animal<sup>23,27</sup>. There are deviations in milk quality and quantity. Due to the local immune response, more immune cells migrate from tissue to the alveolar lumen and are found in the milk, representing a large proportion of the detectable increased SCC<sup>16,27–29</sup>. Neutrophils are most instrumental for coliform bacteria clearance in the udder. Besides spontaneous recovery cured by the adequate immune defense of the host, broad-spectrum antibiotic treatments eliminate *E. coli* bacteria in the udder<sup>23</sup>. However, antibiotic resistance in cattle may occur, albeit at low frequency; antibiotics are used in dairy cattle for both treatment and prophylaxis<sup>30</sup>.

### 1.2.2 *S. aureus* Mastitis

*S. aureus* is a facultative anaerobic gram-positive coccal bacterium found among the commensal microbiota on the healthy skin and in the nose and throat of mammals<sup>31</sup>. At the same time, if the host defense collapses, *S. aureus* can cause mild to severe pathogenesis through cell invasion<sup>32,33</sup>. The infected bovine mammary gland constitutes an important *S. aureus* reservoir in the dairy production chain<sup>33–35</sup>. Transmission of this cow-associated pathogen occurs primarily by milking<sup>36,37</sup>. A mild *S. aureus* mastitis is usually accompanied by delayed inflammatory reaction with low cytokine release<sup>38,39</sup>. Typically, *S. aureus* causes subclinical chronic infections in which the pathogens survive encapsulated in the host cells<sup>40–42</sup>. Elevated SCC, but no changes in milk composition are detectable for

*S. aureus* mastitis<sup>16</sup>. Treatment of mastitis caused by *S. aureus* is difficult. First, *S. aureus* developed resistance to several antibiotics<sup>22,43</sup>. Second, the coccal bacterium can form a biofilm that interferes with the effectiveness of the host immune system and antimicrobial agents, thus contributing to the persistence of the pathogen in the mammary gland<sup>38,44,45</sup>. Antibiotics currently only cure 10 to 30% of *S. aureus* cases<sup>22</sup>. A present study on *S. aureus* in dairy cattle worldwide suggests genetic differences of the bacterial species due to geographical location<sup>42</sup>. It simultaneously reveals virulence factors associated with host invasion or impairment of immune defense<sup>42</sup>.

### 1.3 Cellular Composition and Specification of the Mammary Gland

The fully developed bovine mammary gland is a multi-cellular organ with pronounced cellular properties. The first physical cell structures that form barriers in the udder are the teat duct, teat cistern, and gland cistern, as illustrated in Figure 1-1. The unique organ consists of luminal epithelial and myoepithelial cells and stromal components, including fibroblasts, vascular endothelial cells, adipocytes, stromal matrix, and immune cells<sup>12,46</sup>. These different cell types take over the physiological component of, e.g., the initial immune defense in the udder (Figure 1-1).

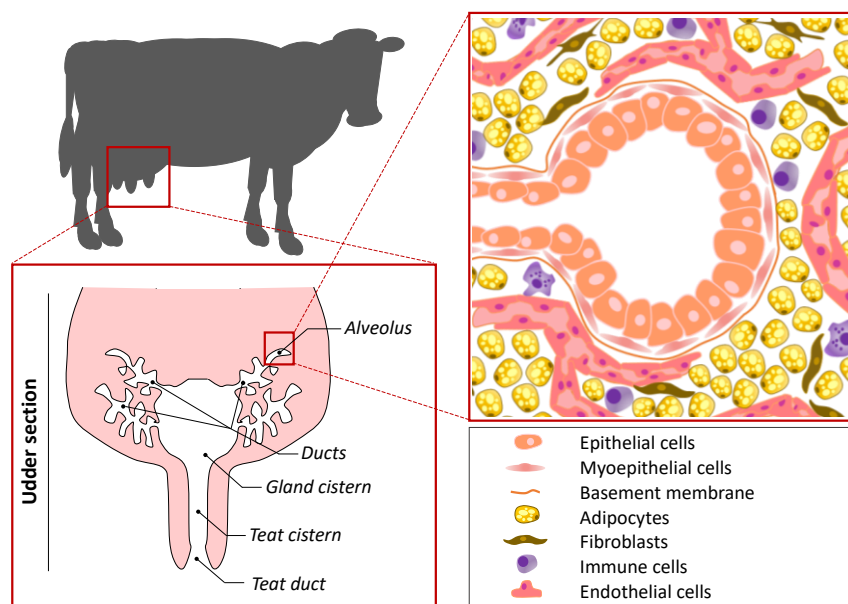


Figure 1-1: The bovine udder's physical barriers and representative cellular components. Image adapted from <sup>47-49</sup>. The cow's udder is composed of 4 teats. Milk is produced in the alveoli, transported down the ducts, and collected in the gland cistern, directly connected with the teat cistern, as shown in the udder section. The milk-producing alveoli are composed of or surrounded by various cells, as shown in the illustration.

The mammary glands of humans and ruminants consist of adipocytes together with fibroblasts, which provide the structure of the stroma<sup>47,48</sup>. As summarized in Inman *et al.* 2015<sup>46</sup>, these cells also possess additional characteristics of cellular communication and contribute to the structure and development of the organ. Likewise, the vascular endothelium is essential for the growth and development of the mammary gland and for the initiation and maintenance of milk production and host defense against bacteria<sup>50</sup>. Myoepithelial and stromal cells synthesize extracellular matrix (ECM) components, which support the architecture and function of epithelial cells<sup>46</sup>. Luminal epithelial cells represent the most common cell type in the lactating mammary gland, which line the milk ducts and alveoli, and constitute

a barrier separating exterior from interior. They secrete milk and form a functional unit contributing to the immune response network<sup>51,52</sup>. Mammary epithelial cells (MEC) are the first barrier to pathogens that enter the bovine mammary gland. Due to the widespread distribution of MEC, they are very likely to trigger the first effective immune response, which subsequently leads to the recruitment of cellular factors of immune defense in the udder. Immune cells, such as macrophages, eosinophils, and mast cells, also exist in healthy tissue. Together with hormones, they are essential for morphogenesis<sup>53,54</sup>. The immune cells are present in increased numbers in inflamed tissue by reacting to immunological mediators and contributing to the immune defense mechanism<sup>55</sup>.

#### 1.4 Mammary Epithelial Cell Models

As introduced, MECs are highly relevant sentinel and effector cells of udder immunity<sup>56–58</sup>. Deciphering the pathogen-specific regulatory mechanisms of mastitis pathogenesis at the molecular level is increasingly being researched. Global transcriptome profiling and targeted gene analysis have demonstrated that *E. coli* upregulates more genes in udder and model cells faster and more robustly than *S. aureus*<sup>39,58–61</sup>. An *in vitro* study showed that *S. aureus* triggers a delayed and much weaker immune response than *E. coli* in bovine MEC<sup>61</sup>. In addition to the genes relevant to the immune and inflammatory response, *CYP11A1* was a strongly regulated gene after stimulation with heat-inactivated *S. aureus* and *E. coli*<sup>61</sup>.

In cattle mastitis research, bovine MECs are preferred model systems for *in vitro* and functional studies<sup>40,58,62–66</sup>. This approach is also promising for ethical reasons, thus reducing animal testing to a minimum. In addition, the cell culture approach is preferable over the *in vivo* models as it is less difficult to standardize. Notably, the collection and alteration of identity over time limits the use of primary cells. Therefore, they could only be studied for a short test duration, whereas immortalized clonal cell cultures can theoretically be cultivated and used indefinitely for experiments.

Few immortalized mammary epithelial cell lines of bovine exist. Among those are the spontaneously immortalized mammary epithelial cell lines BMGE+H (bovine mammary epithelial cells of the hormone-adapted)<sup>67</sup> and HH2A<sup>68</sup>. Also, a buffalo mammary epithelial cell (BuMEC) line<sup>69</sup> exists. Bovine mammary epithelial cell lines ET-C (epithelial and myoepithelial-like characteristics)<sup>70</sup>, BME-UV<sup>71</sup>, and MAC-T (mammary alveolar cells)<sup>65</sup> were immortalized by stable integrating the simian virus large T antigen (SV40LTA) gene. The typical morphological characteristics of epithelial cells and sensitivity to lactogenic hormones is a decisive criterion of a representative cell line of the bovine mammary gland. However, only a few cell lines express lactation-specific proteins. MAC-T cells exhibit diverse lactation proteins, including  $\alpha$ - and  $\beta$ -caseins<sup>65</sup>. Contrasting reports on these properties and the potential use of MACT were postulated more recently<sup>72</sup>. Beyond milk protein and lipid synthesis, immortalized mammary epithelial cell lines are used to study epithelial function, such as cell communication and immune response. Primary bovine mammary epithelial cells (pbMEC) and the established clonal cell line (MAC-T) are frequently implemented as they adequately reproduce critical

aspects of the pathogen-specific response during udder infections<sup>57,73–75</sup>. However, there is a growing indication that fundamental differences exist between MAC-T and pbMEC.

## 1.5 Players of the Immune Response in Mammary Epithelial Cells

The molecular background of the course of mastitis was explored in numerous studies of cow udders and MEC, including critical factors contributing to the immune response during mastitis locally and systemically<sup>38,39,56,57,60,61,76–78</sup>. The infection in the mammary gland by pathogens leads to the immune system's immediate activation of several precise recognition and reaction mechanisms to protect the host. Mammals exhibit the innate (natural) and the adaptive (acquired) immune response<sup>79</sup>. Innate mechanisms are switched on after recognizing pathogenic structures and initiate clearance mechanisms of microorganisms to limit the pathogen's spread. The innate immune response stimulates the adaptive mechanisms, represented mainly by the function of B- and T-cells. Selectively stimulated by antigens, the adaptive and innate immune interact and efficiently clear infections. The following section details some immune response mechanisms.

### 1.5.1 Toll-like Receptor Signaling

As part of the innate immune response, toll-like receptors (TLRs), known as pattern recognition receptors (PRRs), are integral components of the epithelial cell membrane<sup>52</sup>. These receptors recognize and bind characteristic patterns on the microbial cell surface, such as conserved lipoproteins and other microbial cell wall components, and initiate downstream cellular response signaling pathways<sup>80</sup>. It was shown that increased messenger RNA (mRNA) expression levels of TLR2 and TLR4 occur in bovine udders infected by mastitis-causing bacteria<sup>81</sup>. TLR4 functions in a complex with associated molecules like the cluster of differentiation 14 (CD14); they recognize lipopolysaccharide (LPS), a basic pathogen-associated molecular pattern (PAMP) of *E. coli*<sup>52</sup>. TLR2 recognizes staphylococci peptidoglycan and lipoteichoic acid (LTA)<sup>82–84</sup>. TLR2 also depends on additional molecules, like CD14<sup>82,85</sup>, or forms clusters with other receptors, such as TLR1 or TLR6<sup>86,87</sup>, subsequently mediating the innate immune response and activating the adaptive immune system. Intracellular adapter proteins such as adapter myeloid differentiation factor 88 (MyD88), TIR-associated protein (TIRAP), TIR-domain-containing adaptor protein including interferon- $\beta$  (TRIF), and tumor necrosis factor receptor-associated factor 6 (TRAF6) lead to NF- $\kappa$ B activation, cytokine secretion, and pro-inflammatory cytokine secretion<sup>80</sup>. The synthesis and release of immune mediators in the mammary gland trigger the immune response primarily through cell recruitment and the synthesis of antimicrobial peptides<sup>45,88</sup>.

So far, this chapter has focused on recognizing pathogens and downstream signaling. The following sections will cover the immune mediators. Various immune mediators respond to mastitis pathogen infection time-dependently, as Günther et al. (2011) exemplified. They describe the kinetics of mastitis immune genes in a microarray study in MEC<sup>61</sup>. The same study classified genes according to when they reached their maximum altered mRNA expression after *E. coli* challenge. These classes of inflammatory response genes are immediate early (1 h post-challenge), delayed early (3 h post-challenge), and late (6 to 24 h post-challenge)<sup>61</sup>. To be more precise, the immediate early response includes modulation of

cytokines, such as tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), and chemokines, such as CXCL8 and CCL20. The cytokine interleukin-6 (IL6) is essential in the late early response.

Moreover, secondary response genes with antimicrobial functions are relevant to eliminate or contain bacterial propagation in the early phase of infection. One of these genes encodes inducible nitric oxide synthase (NOS2A), predicted to respond within 3 h post-challenge. Second, lingual antimicrobial peptide (LAP) shows a late response, as suggested by Liu *et al.* (2011)<sup>89</sup>. The coordinated expression of immune mediators is essential for an adequate inflammatory response.

### 1.5.2 Cytokines

Cytokines are instrumental in cell signaling of the immune system by acting either on the cell producing them (autocrine) or on the surrounding cells (paracrine)<sup>90</sup>. They divide into different classes according to their functionality. Besides its contribution to the immune response, cytokines are associated with the induction of cell proliferation, survival, apoptosis, or differentiation<sup>91</sup>. Pro-inflammatory cytokines are among the first secreted signal molecules after microbial contact in the mammary gland. Anti-inflammatory cytokines inhibit the activity of pro-inflammatory cytokines and thus intervene in the inflammatory process<sup>91,92</sup>.

Immune cells, such as macrophages, but also non-immune cells produce TNF- $\alpha$ <sup>92</sup>. The distinctive feature of TNF- $\alpha$  is its involvement in the early stages of infections and stimulation of a cascade of downstream messengers. Depending on the context, TNF- $\alpha$  stimulates, e.g., the activation of acute phase reaction, the promotion of endothelial adhesion molecules, chemotaxis, and phagocytosis<sup>38</sup>. Especially in coliform mastitis, TNF- $\alpha$  has been established as one of the main inflammatory factors during pathogenesis<sup>93–95</sup>.

IL6 plays a central role as a molecular messenger for acute phase reactions<sup>96</sup>. Moreover, IL6 is essential for B-cell differentiation and the interplay between the innate and adaptive immune system<sup>97,98</sup>. The cytokine IL6 is associated with pro- and anti-inflammatory effects<sup>98</sup>. The presence of bacteria, viruses, and other cytokines such as TNF- $\alpha$  and IL1 $\beta$  increases IL6<sup>99,100</sup>, leading to increased neutrophil recruitment and thus promoting inflammation. In contrast, the inhibition of IL1 $\beta$  and TNF- $\alpha$  and the induction of IL1-receptor antagonists and soluble TNF receptors are considered anti-inflammatory properties of IL6<sup>38</sup>. Various cell types, including mast cells, macrophages, neutrophils, fibroblasts, endothelial, and epithelial cells, produce the multifunctional cytokine IL6<sup>38,91,101</sup>. Studies demonstrated significant upregulation of IL6 in MEC and udder quarters 3 h after *E. coli* challenge, whereas IL6 was weakly upregulated after *S. aureus* challenge<sup>61,102</sup>.

### 1.5.3 Chemokines

Chemokines are a group of small proteins (8–12 kDa) characterized by three to four conserved cysteine (C) residues<sup>92</sup>. Depending on the position of the N-terminal cysteine residues, four subgroups of chemokines have been defined: C, CC, CXC, and CX3C<sup>91,92</sup>. Cytokines, such as IL1 $\beta$  and TNF- $\alpha$ , and exogenous stimuli at the site of infection induce chemokines<sup>103</sup>. Subsequently, chemokines control the



recruitment of effector cells by inducing chemotaxis and cell migration of numerous cell types, e.g., of the immune system, such as neutrophils, monocytes, lymphocytes, and eosinophils<sup>91</sup>.

*CXCL8* and *IL8* are interchangeable and refer to a pro-inflammatory cytokine secreted by various cell types, such as epithelial, endothelial, and neutrophil cells<sup>103</sup>. *IL8* possesses a strong chemotactic effect on all known types of immune cells<sup>103</sup>. In MEC, *IL8* expression has been documented following stimulation with LPS or the pathogens *E. coli* or *S. aureus*; it is one of the first upregulated genes after pathogen contact<sup>56,77</sup>. *IL8* was also strongly upregulated by *E. coli in vivo*<sup>102</sup>.

*CCL20* secretion of epithelial cells upon pathogen infection triggers lymphocyte chemotaxis. *In vitro* studies in pbMEC found *CCL20* as one of the highest upregulated genes following challenge with heat-inactivated *E. coli*<sup>56</sup>. Strong upregulation of *CCL20* was also found in *E. coli*-infected bovine udder quarters<sup>104</sup>.

### 1.5.4 Defensins

Defensins are components that belong to the group of antimicrobial peptides (AMPs) and are part of the fast and complex mechanisms of the innate immune response, e.g., in the mammary gland<sup>105</sup>. AMPs eliminate different pathogens, like microorganisms, fungi, and viruses<sup>106</sup>. The chemical conformation and charge of defensins allow them to disrupt the microbial membrane and induce lysis. The classification into  $\alpha$ ,  $\beta$ , and  $\theta$ -defensins is based on their chemical properties (size, disulfide bonding, and amino acid conformation), while all share a cysteine motif<sup>107</sup>.

Lingual antimicrobial peptide (LAP) belongs to the bovine bactericidal  $\beta$ -defensins and represents an essential innate immune factor in the host during mastitis<sup>108</sup>. Interestingly, the *LAP* signal in pbMEC drastically reduces during the first three passages of the cells, which suggests only short-time cultures of pbMEC to study pathogen-dependent gene regulation<sup>56</sup>. Additionally,  $\beta$ -defensin 5 (*BNBD5*) is upregulated in the udder tissue of cattle with mastitis<sup>81</sup>, a finding confirmed *in vitro* in MEC<sup>57</sup>.

Nitric oxide (NO) also serves as an inorganic disinfectant<sup>79</sup>. The enzyme inducible nitric oxygen synthase (*NOS2A* or *iNOS*) synthesizes nitric oxygen and thus is involved in host defense against infectious agents. Expression of *NOS2A* was upregulated by the challenge with *E. coli* or *S. aureus*, as demonstrated in a microarray study in pbMEC<sup>76</sup>.

## 1.6 Transcriptomics

The basic principle of molecular biology outlines the flow of sequence information stored in the form of genes in DNA, which is transcribed into RNA and translated into the final functional unit of proteins<sup>109</sup>. A gene consists of coding regions (exons), non-coding regions (introns), regulatory sequence regions, and sequences that code for the start and stop of transcription. Gene transcription requires an accessible promoter for transcription factors that recruit the transcription machinery and thus initiate the process. Transcription starts at the transcription start site (TSS) and ends at a specific transcription termination site (TTS). The reading direction is from 3' to 5' on the template strand of the DNA. Regulatory elements of a gene are enhancers and silencers, which are located close to or far from the gene. The term transcriptome refers to the complex variety of different coding RNAs, such as messenger

RNA (mRNA), and non-coding RNAs (ncRNAs), such as transfer RNA (tRNA), ribosomal RNA (rRNA), micro RNA (miRNA), small nuclear RNA (snRNA), long non-coding RNA (lncRNAs), at a given time point in a cell<sup>110</sup>. Diverse approaches exist to study the transcriptome. Quantitative polymerase chain reaction (qPCR) can measure single transcripts. Genome-wide quantification of the transcriptome can be achieved using hybridization-based microarray or sequence-based technologies. The recent development of the high throughput technique analyzing the transcriptome by next-generation sequencing is called RNA sequencing (RNA-seq). RNA-seq captures genome-wide coding and non-coding transcripts within a cell or cell population.

The principle of the latter technique is that RNA is converted into a library of complementary DNA (cDNA) fragments with adaptor(s) attached<sup>111</sup>. RNA molecules, with or without amplification, are sequenced in a high-throughput manner to obtain short sequences from one end (single-end sequencing) or both ends (pair-end sequencing)<sup>111</sup>. Sequence reads can subsequently be aligned to the genome assembly of the target species to enable a genome-wide identification or quantification of transcripts<sup>111</sup>. The present study will use qPCR and RNA-seq for transcript quantification.

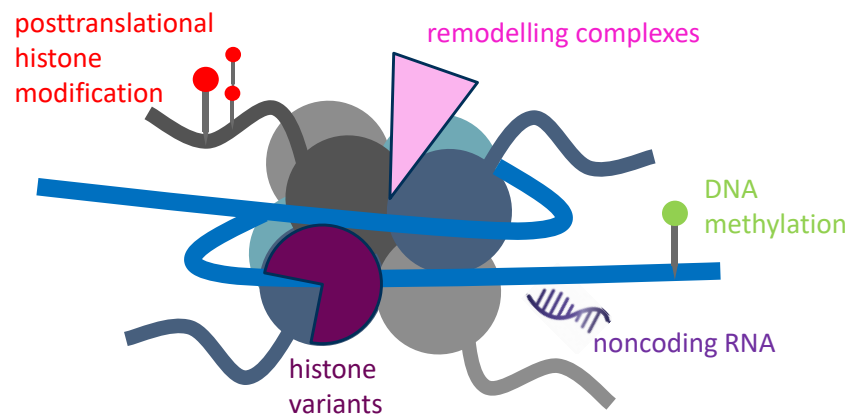
## **1.7 Epigenomics**

The term epigenomics refers to the study of the totality of epigenetic changes in the genetic material of a cell or a whole organism. It depends on age and environmental stimuli dictating various epigenetic mechanisms' density, presence, and absence. These mechanisms regulate how specific genes are transcriptionally active or inactive without affecting the underlying DNA sequence<sup>112</sup>. In general, epigenomics improves understanding of diseases' complex molecular regulatory mechanisms. In the following, epigenetic mechanisms will be introduced, focusing on histone modifications and their study through chromatin immunoprecipitation (ChIP).

### **1.7.1 Epigenetics**

Organisms, both plants and animals, interact extensively with the environment. Various internal and external stimuli can translate into gene activity and function changes. Epigenetic mechanisms mediate in which way genes are expressed. These mechanisms can have a temporary, permanent, or hereditary effect across generations<sup>113</sup>. Waddington originally defined epigenetics as bridging the gap between phenotype and genotype<sup>114</sup> and includes all causal processes that affect gene activity without altering the DNA code<sup>112</sup>. Figure 1-2 shows an overview of the major epigenetic mechanisms, including chemical modification of DNA bases and packaging proteins (histones), insertion of histone variants and remodeling complexes, and effects of non-coding RNAs<sup>115</sup>. Methods used to study epigenetic mechanisms include DNA methylation profiling, chromatin immunoprecipitation, bisulfite sequencing, and RNA sequencing, targeting, e.g., microRNAs. One of the most studied epigenetic mechanisms is the methylation of CpG dinucleotides of DNA, which is closely linked to the repression of gene expression in mammals<sup>116</sup>. The biological function of DNA methylation is well described in the development of the embryo<sup>117,118</sup> and differentiation of stem cells<sup>119</sup>. Comprehension of epigenetic

transcriptional regulation and dysregulation is critical for distinguishing between healthy and pathological states.



*Figure 1-2: Main epigenetic mechanisms. DNA is wrapped around histone proteins to condense genetic material in the nucleus. Epigenetic mechanisms mainly determine gene activity and chromatin conformation.*

### 1.7.2 Histone Modifications

Both histone and non-histone proteins bind and coil DNA, concentrating it into chromatin in the nucleus. The nucleosome is the basic module of chromatin. It consists of 147 bp of DNA<sup>120</sup> and an octamer of two copies of the four core histone proteins (H2A, H2B, H3, and H4)<sup>121</sup>. Each chromosome contains hundreds of thousands of nucleosomes. Various histone-modifying enzymes exist, which are responsible for specific post-translational modifications of the amino-terminal tail of histones. The common histone modifications are methylation and acetylation of a lysine (K), histidine (H), arginine (R), or phosphorylation at serine (S). Adding or removing chemical groups to DNA or histone proteins can either facilitate or inhibit the binding of transcription factors to DNA and subsequent gene transcription. Inactive and densely packed heterochromatin is not accessible, while the structural changes to active euchromatin allow access to the DNA<sup>122</sup>. Different histone modifications in specific gene regions, such as the promoter region and the transcription start site (TSS), are well documented for mammals<sup>123,124</sup>. The multitude of regulatory effects of histone modifications is vital, as they can be involved in diverse biological processes.

The study of histone modification is well established, particularly the methylation and acetylation of lysine (K) residues on the amino-terminal tail of histone H3. Histone methyltransferases and demethylases control the methylation status of histones. Mono-, di- or tri-methylation at histone three lysine four (H3K4) is often associated with an active gene expression<sup>125</sup>. Histone acetyltransferases and deacetylases determine the acetylation status of histones. Histone three lysine 14 acetylation (H3K14Ac) is often associated with an open chromatin structure, allowing active gene expression<sup>126</sup>. H3K14Ac was related to the activation of immune genes and is critical for recruiting co-factors such as transcription factor II D (TFIID) at the interferon-gamma locus, thereby directing an appropriate immune response. Conversely, removing specific acetyl residues leads to a closed chromatin structure and the repression of transcription<sup>127</sup>. Tri-methylation of lysine 27 and mono-, di-, or tri-methylation of lysine 9 of histone three are considered modifications associated with transcriptional repression<sup>124,128</sup>.

### 1.7.3 Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) combined with high-throughput sequencing (seq) is a state-of-the-art technique to study the epigenetic profile of, for instance, a model cell system on a genome-wide level. Two approaches to the ChIP method have evolved. These are native and cross-linked ChIP<sup>129</sup>. Protein-DNA interactions, as such, can be tight interactions<sup>130</sup>. As introduced previously, an octamer of histones binds ~150 bp DNA in the nucleosome. Native ChIP (N-ChIP) involves the use of the enzyme micrococcal nuclease (MNase) after the lysis of starting material (e.g., cells). MNase digests the linker DNA between nucleosomes. Thus, N-ChIP produces single mono-nucleosomal DNA fragments of ~150 bp length<sup>131</sup>. In the cross-linked ChIP (X-ChIP) approach, protein-DNA interactions are chemically stabilized, and chromatin is sheared into fragments of 100 to 800 bp<sup>132</sup> by sonication. Both ChIP approaches rely on antibodies binding specific histones such as histone modifications. To release the DNA and dissociate the protein, the N-ChIP and X-ChIP material is reverse-cross-linked and destabilized by heat. Then, the DNA is purified and used for PCR, qPCR, microarray, or sequencing analyses.

## 1.8 Epigenetics for Genome Regulation in Livestock Species

The International Consortium for Functional Annotation of Animal Genomes (FAANG) aims to standardize assays and bioinformatic pipelines and generate transcriptomic and epigenomic data in different animal species and tissues<sup>133</sup>. The mission of FAANG is to identify and describe all functional elements of animal genomes. These elements include protein-coding genes, non-coding RNAs, regulatory regions, and epigenetic modifications. Findings on epigenetic control in livestock have the potential of new therapeutic strategies for the benefit of the animal, as they can ultimately improve animal health and increase animal well-being.

Epigenetics complements the principle of activation and inactivation of the genome and is thus involved in dynamically fine-tuning all biological processes across species. On the one hand, genetically determined changes account for the phenotype; on the other hand, additional factors such as nutrition, stress, and exposure to pathogens or chemicals can lead to plastic epigenetic changes. These changes affect individuals and their offspring *in utero* or via transgenerational inheritance<sup>134</sup>. Consequently, size, development rates, reproduction, or other animal-related traits are compromised. In-depth knowledge of epigenetic transcriptional regulation in livestock can help facilitate an understanding of the normal operation and pathophysiological changes, e.g., attributable to environmental stimuli.

Histone modifications and DNA methylation have been studied in bovine embryogenesis. The timing and occurrence of genetic and epigenetic mechanisms contribute profoundly to chromatin conformation and gene activity, determining the proper development of cells, tissue, and, eventually, the embryo<sup>135</sup>. In a 2011 study, epigenetics was addressed by profiling histone marks in cloned and *in vitro* fertilized bovine embryos<sup>136</sup>. In the early stage of bovine development, the status of histone modifications (e.g., acetylation of lysine 18 and 9 at histone H3) varied among cloned embryos and *in vitro* fertilized control; however, in the developmental stage from morula to blastula, an equal distribution of the histone

modifications is established<sup>136</sup>. The disruption of the physiological epigenetic profile of the gametes and embryos could cause *in vitro* embryos to suffer health complications, for example, by making them more susceptible to disease or sudden death<sup>137</sup>. In addition, DNA methylation contributes to the control of imprinted genes that are exclusively expressed from maternal or paternal alleles. In livestock species such as cattle, pigs, and sheep, imprinted genes have been found to regulate fetal growth and development<sup>138</sup>. Still, they are also associated with muscle development, meat quality, and milk traits<sup>138</sup>.

## 1.9 Epigenetics and Infection

The following section will first outline the concept of host-pathogen interaction and its role in gene regulation through epigenetics. Selected studies will then illustrate the role of epigenetics during infection before focusing on epigenetic insights in dairy cattle and during mastitis.

### 1.9.1 *Host-pathogen Interaction*

Exposure of a host to a pathogen causes specific genes to be turned on and off with underlying mechanisms including epigenetics<sup>139</sup>. Bacteria can modify the epigenome of the host they target<sup>139,140</sup> with a long-lasting effect on the organism and offspring<sup>141</sup>. These changes can be either beneficial or detrimental to the host organism<sup>139</sup>. Pathogens penetrate the hosts' protective barriers to gain access to the host and cause infection. Epigenetic mechanisms on the part of the host are also critical to maintain a proper host immune response to prevent or inhibit pathogen invasion and growth.

Consequently, a cellular network initiates signal- and gene-specific functions, including the activation of specific genes for antimicrobial defense, immune response, tissue repair, and remodeling<sup>142</sup>. At the same time, pathogens have developed strategies to make the best possible use of the host and ensure bacterial survival, e.g., by circumventing the hosts' defense mechanisms. An infection-induced reprogramming of host genes can change the host physiology, morphology, and behavior. It has been suggested that pathogen-induced host transcriptional dysregulation is an ordered, sequential process with diverse genes altered throughout the infection process<sup>140</sup>. Such regulatory mechanisms include, e.g., the secretion of pathogen-derived molecules that may use host cell trafficking to access specific subcellular compartments and influence transcription by targeting chromatin, DNA, or proteins<sup>140</sup>. Any collapse of the host defense mechanisms makes a compromised host more vulnerable to infection. Long-term epigenetic changes based on host-pathogen interaction can be reprogramming or transforming host cells and induction of oncogenesis leading to chronic diseases<sup>140</sup>.

### 1.9.2 *Epigenetics in Immunity*

It is established that processes to protect the body against stress, injury, or infection are critically regulated by transcription factors of the Nuclear factor kappa-light-chain-enhancer of activated B cells (*NF-κB*), Forkhead box P3 (*FOXP3*), interferon regulatory factor (*IRF*), and signal transducer and activator of transcription (*STAT*) families combined with co-regulators and epigenetic transcriptional regulation<sup>142,143</sup>. In many diseases, such as cancer<sup>144–146</sup> or asthma<sup>147,148</sup>, epigenetics have been well documented in human and related animal models, while some knowledge gaps remain in livestock.

Histone modifications play a key role in inflammation and disease and affect e.g. immune and epithelial cells. In airway epithelial cells, H3K18ac and H3K9me3 modifications have been described to be associated with gene expression, especially for asthma-related genes such as EGFR and STAT6<sup>149</sup>. Another example is Epstein-Barr virus infection in epithelial cells of the nasopharynx: H3K4me3 and H3K27me3 modifications lead to downregulation of genes of the DNA repair pathway, highlighting the profound impact of histone modifications on disease development and immune response.<sup>150</sup>

Jumonji domain-containing protein-3 (JMJD3), the enzyme removing the methylation of histones, plays a vital role in macrophages by contributing to cell type differentiation, activation, and polarization of the cell type<sup>122</sup>. This mechanism represents an essential link between inflammatory processes and the alteration of the epigenome.

The CREB-binding protein *CBP* and its close homolog p300<sup>151</sup>, acting as histone acetyltransferases, contribute to chromatin condensation. Histone deacetylases are counterplayers of acetyltransferases. In a study on human monocytes, the mechanistic background by which *IL10* suppressed the LPS-induced expression of *CXCL8* and *TNF- $\alpha$*  was associated with histone deacetylase 2<sup>152</sup>. Chromatin activity change prevented transcription. In monocytes of chronic obstructive pulmonary disease (COPD) patients, HDAC2 levels were also reduced, affecting the function of *IL10* as an inflammatory prevention<sup>152</sup>.

These studies have shown that microorganisms interfere with host gene activity in various ways via influencing epigenetics or epigenetic machinery.

### **1.9.3 Epigenetics in Immunity of Livestock**

Immunity is a complex process. For example, it depends on the diet, which may alter the epigenetic landscape and the immunity of dairy cattle<sup>153,154</sup>. However, feeding dairy cows a highly concentrated diet to meet their energy needs poses a health risk in the form of a cow's systemic inflammation. In the liver of cows, demethylation and chromatin decompaction at the promoter region of the immune candidate genes were associated with an upregulation of these immune response genes (*Hp*, *SAA3*, *LBP*, *TLR4*) in response to the diet<sup>153</sup>. A second study<sup>154</sup> showed that a high-concentrate diet decreases histone H3 acetylation levels in the mammary gland, which negatively correlates with increased LPS concentration in the mammary blood of dairy cows<sup>154</sup>. Moreover, methylation of genes related to fat synthesis increased in response to a high-concentrate diet, while it decreased for genes related to protein synthesis<sup>154</sup>. These examples demonstrate the interconnectedness and complexity of epigenetic transcriptional regulation by internal and external stimuli with local and systemic effects.

Some details of epigenetic mechanisms associated with immune functions in the udder or mastitis have been uncovered in livestock. However, the epigenetic role of histone modifications during mastitis in cattle must be better understood and explored.

The density level of chromatin was studied *in vivo* after *E. coli* challenge of udder cells. It was found that lingual antimicrobial peptide (LAP) must open the densely packed chromatin of the quiescent promoter to recruit NF- $\kappa$ B p65 transcription factor for expression activation<sup>89</sup>. In addition, chromatin

compaction and activity of lactation genes and immune genes seem to be highly dependent on hyper- and hypomethylation of the genome during mastitis pathogen challenge<sup>155,156</sup>.

DNA methylation is involved in the pathogenesis of bovine and porcine mastitis. It has been found that DNA methylation affects gene and microRNA expression in neutrophils, one of the first effector cells during bovine *E. coli* mastitis<sup>157</sup>. Moreover, DNA methylation loci were characterized in porcine mammary epithelial cells during *E. coli* mastitis, providing mechanistic insights into the course of infection<sup>158</sup>.

In a study on bovine blood lymphocytes from unchallenged cows, the genome-wide epigenetic landscape for the repressive histone mark H3K27me3 was associated with genomic regulatory regions<sup>159</sup>. Exploration of the epigenome of lymphocytes from *S. aureus* infected cattle showed evidence for increased H3K27me3 levels in infected cows, compared to healthy controls<sup>160</sup>. IL10 and PTX3 were postulated within the same study as putative epigenetic biomarkers in bovine *S. aureus* mastitis<sup>160</sup>.

Previous data from a global transcriptome study of pbMEC challenged with mastitis pathogens provided reasonable evidence that *E. coli*-challenge regulates various transcription factors consistent with changes in histone-modifying enzymes<sup>61</sup>. Thus, after exposure to *E. coli*, the expression of histone methyltransferases jumonji, AT rich interactive domain 2 (*JARID2*), and DOT1-like, histone H3 methyltransferase (*DOT1L*) was increased, whereas that of the histone deacetylase HDAC7 was decreased<sup>61</sup>. The altered expression of these epigenetic regulators may indicate involvement in the immune response. In the same study, *S. aureus* caused a significantly lower activation of specific transcription factors, while no enzymes responsible for histone modifications were stimulated compared to *E. coli*<sup>61</sup>. *S. aureus* exposure upregulated gene expression for the NR4A subfamily (*NR4A1*, -2, -3)<sup>61</sup>. One model suggests that these “orphan” receptors, which function as transcription factors, are activated by ligand binding in response to, for example, environmental or pathophysiological changes<sup>161</sup>. NR4As act differently depending on the tissue, while they may be involved in apoptosis, proliferation, and immune responses<sup>162</sup>. In immune cells, it was reported that NR4As interact with essential transcription factors such as IRF4 and NF- $\kappa$ B<sup>162</sup>. These factors, in turn, lead to the association of other proteins, which activate, e.g., histone-modifying enzymes. The transcriptional corepressor complex CoREST that binds NF- $\kappa$ B, recruits histone deacetylases (HDAC1, -2, -3)<sup>163,164</sup> and the lysine-specific demethylase LSD1<sup>165</sup>. These enzymes’ actions affect the level of the epigenetic mark of chromatin or promoter regions of genes, causing changes in gene expression. In addition to the molecular background of the interaction of “orphan” receptors with regulatory gene elements, this suggests that epigenetic events occur at the chromatin level leading to divergent gene expression.

Considerable evidence suggests that epigenetic mechanisms play an essential role in bovine mastitis. Epigenetic profiles of bovine mammary epithelial cells need to be explored to demonstrate how and where histone modifications act as epigenetic transcriptional regulators during mastitis.

### 1.10 Study Objectives

Prior transcriptomic studies identified essential immune response genes specifically upregulated upon pathogen exposure in the bovine mammary gland. The respective immune response genes, *CCL20*, *TNF- $\alpha$* , *IL6*, *IL8*, *iNOS*, *LAP*, and *JMJD3* encoding an epigenetic regulator, were selected for analysis in this study. A large and growing body of literature addresses epigenetic mechanisms in physiological and pathophysiological processes, with many unresolved questions remaining. In particular, the transcription of immune genes during mastitis is thought to be influenced by epigenetic mechanisms, but this question is yet to be resolved. The epigenetic landscape of udder cells, representing the first cellular barrier during infection, is poorly documented. However, there is consensus about the role of bacteria contributing to epigenetic changes, such as alterations of histone modifications. I hypothesize that contact with mastitis pathogens induces transcription and coregulatory factors in the nucleus, which in turn trigger the function of histone-modifying enzymes in mammary epithelial cells. The action of these enzymes significantly alters the accessibility of chromatin and, thus, the transcription of the DNA. Various histone modifications in promoter regions have been documented; a role in promoter regions of immune genes in mastitis cell models is still undetermined. This study seeks to obtain data to help address the identified research gaps. The central questions of this study are:

**1) Do epigenetic changes in the form of histone modifications contribute to the transcription of immune genes in response to mastitis pathogens in mammary epithelial cells?**

Above all, it remains open to what extent the observations in an immortalized cell line can be extrapolated to the natural environment. Therefore, the immortalized cell line was compared with primary epithelial cells from mammary gland ducts. Furthermore, targeting specific genes gives limited insight into epigenetic changes in these cell models. Consequently, I propose to examine genome-wide gene expression and distribution of histone modifications in control and bacteria-exposed cells. I anticipate that this approach will provide a better understanding of the epigenetic transcriptional regulatory network in mammary epithelial cells during mastitis. The resulting research questions to be investigated are:

**2) Do mastitis pathogens have a measurable impact on histone modifications in representative bovine epithelial cell models beyond the candidate regions?**

**3) Do transcriptomic and epigenomic data correlate in mammary epithelial cells?**

Although this study was relatively exploratory, the findings should contribute to multi-omics research in livestock. However, a complete picture of epigenetic mechanisms lies beyond the scope of this study.



## 2 Material and Methods

This chapter of the dissertation describes the methods used in the study. The first section outlines the experimental design. The following sections contain information about the materials and describe in more detail the methods used to address the research questions. The methods include cell biological, molecular and computational analytical approaches.

### 2.1 Experimental Design

This dissertation used mammary epithelial cell models to explore transcriptomic and epigenomic responses to mastitis pathogens *in vitro*. Figure 2-1 visualizes the layout of the experiments. They are described below.

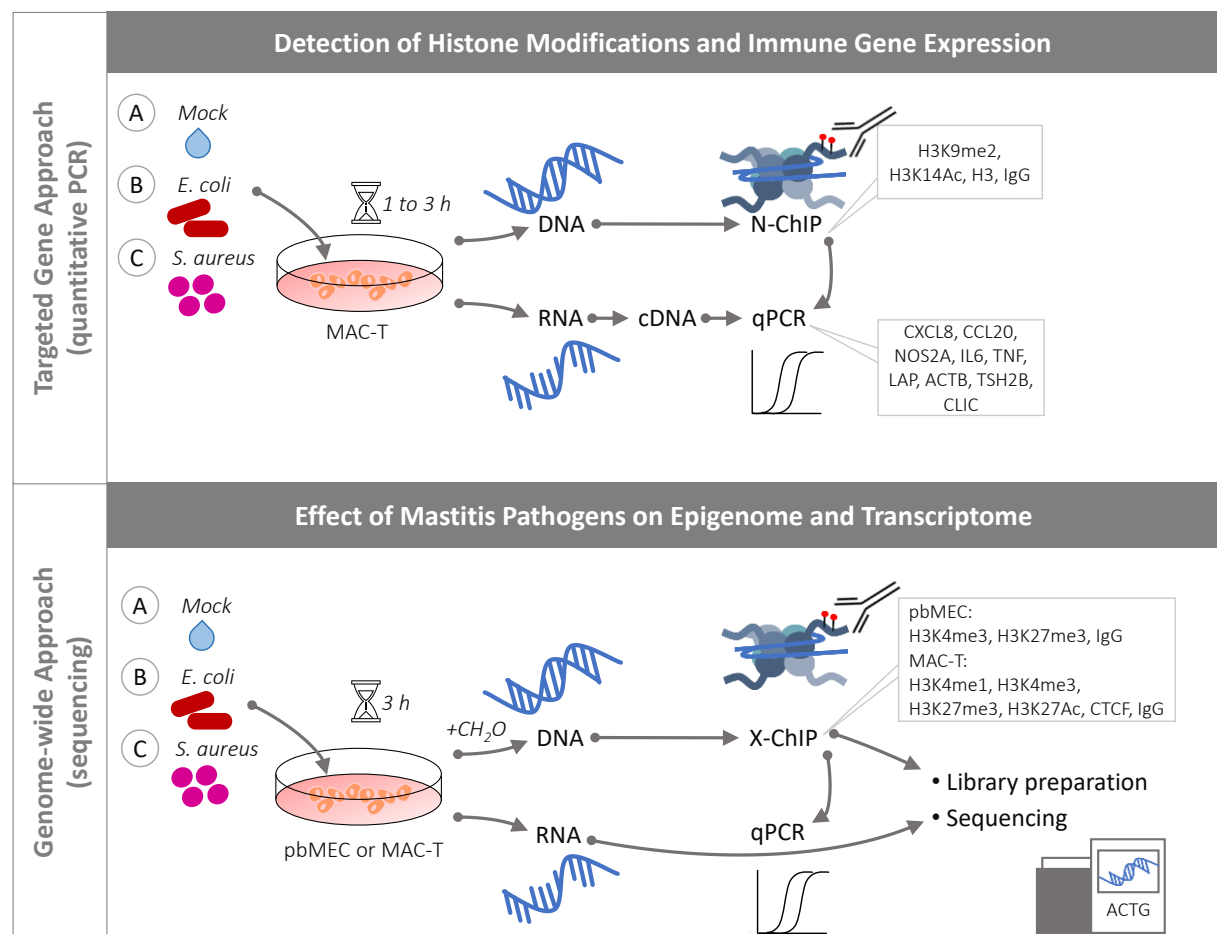


Figure 2-1: Experimental design for detecting transcriptomic and epigenomic changes in bovine mammary epithelial cell models. First, in the targeted gene approach, native chromatin immunoprecipitation (N-ChIP) followed by qPCR was established in the immortalized cell line MAC-T to detect histone modifications. Immunoprecipitated DNA was used for qPCR targeting promoter regions of immune genes. In parallel, RNA was isolated from MAC-T cells challenged with mock (A) or bacteria (B or C), followed by reverse transcription (RT)-qPCR targeting transcription of immune genes. Second, in the genome-wide approach, the immortalized cell line MAC-T and primary cells pbMEC were used to perform cross-linked ChIP (X-ChIP)-seq and RNA-seq. The aim was to gain insights into the genome-wide effects of mastitis pathogens on the whole genome. X-ChIP was verified by qPCR.

### 2.1.1 *Detection of Histone Modifications and Immune Gene Expression in MACT*

First, the immortalized cell line MAC-T was used to determine epigenetic mechanisms in the form of histone modifications. MAC-T were challenged with mastitis pathogen *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub> for 1 or 3 h or remained unchallenged as control (mock). As introduced, histone proteins and DNA naturally form close bonds. Therefore, a native chromatin immunoprecipitation (N-ChIP) protocol was pursued without cross-linking the protein-DNA complex. Antibodies against different histone modifications of histone H3, such as H3K9Me2 and H3K14Ac, were used to precipitate the protein antigen. Moreover, antibodies against histone H3 and non-specific Immunoglobulin G (IgG) served as control detecting the background signal. Immunoprecipitated DNA was used for quantitative polymerase chain reaction (qPCR) measurements in promoter regions of immune response genes.

In parallel, cells were treated similarly to extract RNA and measure the expression of immune response genes by reverse transcription (RT)-qPCR. The immune response genes are known to be activated *in vivo* and *in vitro* in response to mastitis pathogens in a time-dependent manner. I aimed to reproduce these findings for MAC-T. Evidence of the regulation of immune response genes was critical for the possible role of epigenetic mechanisms in supporting the response of the MEC model to the pathogens.

### 2.1.2 *Genome-wide Effect of Mastitis Pathogens on Epigenome and Transcriptome in MACT and pbMEC*

The cross-linked ChIP (X-ChIP) followed by sequencing was established to explore the genome-wide response to pathogen challenge in primary cells pbMEC and the immortalized cell line MAC-T. This work was part of a Short-Term Scientific Mission (STSM) of myself. This was part of a collaboration within the Functional Annotation of Animal Genomes -European network (FAANG-Europe) between the Leibniz Institute for Farm Animal Biology, Dummerstorf, Germany, and Diagenode s.a., Liège, Belgium. The two bovine cell models (pbMEC and MAC-T) were incubated with bacteria (*E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub>) for 3 h in three independent experiments in Dummerstorf. Adding formaldehyde to the culture vessels allowed cross-linking of proteins and DNA. Cells were harvested and deep frozen at -80°C before shipping to Diagenode and processing along the X-ChIP protocol on-site during my STSM. Diagenode performed the library preparation, sequencing, and first exploratory bioinformatic analysis. As the paid data analyses services were unsatisfactory, these analyses were adapted and revised accordingly by Dummerstorf experts. Eventually, data exploration, visualization, and functional analysis were performed independently from Diagenode using a variety of computer-based approaches (see 2.10.1).

In parallel to the X-ChIP-seq, RNA from both cell models incubated under identical conditions (3 h *E. coli*<sub>1303</sub>, *S. aureus*<sub>1027</sub>, or mock) underwent transcriptome analysis by RNA-seq at the Leibniz Institute for Farm Animal Biology, Dummerstorf, Germany. Bioinformatic analysis was used to determine statistically significant results and perform functional analysis (see 2.10.2 and 2.10.3). The overlap of the obtained data sets determined the correlation between histone modifications and gene expression to draw a possible conclusion about epigenetic transcriptional regulation.

## 2.2 Chemicals, Oligonucleotides, Materials, and Equipment

Chemicals used for the study were purchased from AppliChem GmbH, Carl Roth GmbH, Merck KGaA, and Sigma-Aldrich Chemie GmbH unless otherwise stated. All solutions, buffers, and media were prepared with autoclaved ultrapure water. The recipes for all used solutions are listed in the appendix. Oligonucleotides were required as primers in polymerase chain reactions (PCR) or for sequencing. Customer-specific oligonucleotides were purchased from Sigma-Aldrich and Integrated DNA Technologies, Inc. Primer pairs were designed by Primer3 Input software (version 4.1.0) with adjustment of the following parameters: product size, 19 to 25 bp primer length, 58 to 62°C melting temperature range, 40 to 60% GC content per primer, and low complementarity properties for primers and primer pairs. The primer stock solution was adjusted to 50 pmol/ $\mu\text{L}$   $\leftrightarrow$   $\mu\text{M}$ . The primer working solution was used at 10 or 20 pmol/ $\mu\text{L}$ . All oligonucleotides used and generated in the present study can be found in appendix A1.

All details of buffers, substances, reagent kits, equipment, and consumables used in this study can be found in tabular form in appendix A2 to A12. These lists include growth broths used for cloning and buffers required for plasmid DNA extraction. For the selective growth of bacteria, agar plates with different antibiotic supplements were poured.

HBSS and PBS buffer were regularly used for cell extraction and cell washing in cell culture experiments. Complete media were prepared for the growth of mammary epithelial cell models. Primary cells required special media and additives for culture, including hormones, which were contained in RPMI 1640 complete media (Table 2-1). The MAC-T cell line was cultured under standard culture conditions in DMEM complete media (Table 2-2).

Table 2-1: RPMI 1640 complete media composition. Adapted from Hensen *et al.*, 2000<sup>64</sup>

For 50 mL Media	Volumes	Stock concentration	Working concentration
RPMI 1640	43.425 mL		
FCS	5 mL		10%
L-glutamine	1 mL	200 mM	2%
L-methionine	25 $\mu\text{L}$	200 mM	0.1 mM
L-lysine	20 $\mu\text{L}$	1 M	0.4 mM
Sodium pyruvate	5 $\mu\text{L}$	100 mM	0.01 mM
Prolactin	10 $\mu\text{L}$	5 mg/mL	1 $\mu\text{g/mL}$
Hydrocortisone	5 $\mu\text{L}$	10 mg/mL	1 $\mu\text{g/mL}$
Insulin	10 $\mu\text{L}$	5 mg/mL	1 $\mu\text{g/mL}$
P/S	500 $\mu\text{L}$	100x	1x

Table 2-2: DMEM complete media composition

For 50 mL Media	Volumes	Stock concentration	Working concentration
DMEM	43.5 mL		
FBS	5 mL		10%
L-glutamine	1 mL	200 mM	2%
P/S	500 $\mu\text{L}$	100x	1x

### 2.3 Mammary Epithelial Cell Models

Frequently implemented *in vitro* or *in vivo* models advance the understanding of cellular and transcriptional events during milk synthesis and immune defense in the udder. In the present study, I focused on *in vitro* cell models. I implemented primary bovine mammary gland epithelial cells (pbMEC) and mammary gland alveolar cells stably transfected with a T antigen (MAC-T).

There are two ways to obtain pbMEC. One way is to use milk to extract pbMEC<sup>166,167</sup>; the other is enzymatic digestion from tissue<sup>64,168</sup> derived from an udder region with high alveolar density. The second way was methodologically established on-site. The extraction and culture of primary cells are described in section 2.5.1. Udder tissue samples for primary cell extraction of pbMEC were obtained from three healthy Holstein cows in their first pregnancy. Animal culling and sampling were performed at the local Dummerstorf slaughterhouse (EU-approved ES1635) in compliance with all relevant ethical and legal requirements regarding animal welfare and the protection of animals used for scientific purposes. The cows had been purchased and routinely killed without animal experimentation.

Table 2-3: Cows information for tissue sampling.

Internal Animal code	Birth Date	Specimen Collection Date	Animal Age at Collection (days)	Developmental Stage	Health Status at Collection
animal_AAA	01.03.2010	19.06.2013	1206	prime adult stage	normal
animal_AAD	20.01.2012	07.10.2014	991	prime adult stage	normal
animal_AAF	07.06.2012	18.11.2014	874	prime adult stage	normal

An aliquot of MAC-T was kindly provided from Prof. Dr. Ulrich Dobrindts' Laboratory in Würzburg to Dr. Juliane Günther and Prof. Dr. Hans-Martin Seyfert at the Leibniz Institute for Farm Animal Biology (FBN) in Dummerstorf. MAC-T cultures were routinely used, and subcultures were long-term stored in liquid nitrogen until used for experiments in this study.

### 2.4 Mastitis Pathogen Preparation

The mastitis pathogenic strains *E. coli*<sub>1303</sub> and *S. aureus*<sub>1027</sub>, previously extracted from udder secretions of cattle with clinical mastitis<sup>39</sup>, were established in the laboratory. A batch of each pathogen (*E. coli*<sub>1303</sub> and *S. aureus*<sub>1027</sub>) was prepared and used throughout the study to maintain constant challenge conditions. Brain-heart-infusion (BHI) broth for microbiology (Carl Roth) and BHI broth agar were prepared and autoclaved before being used to prepare heat-killed *E. coli*<sub>1303</sub> and *S. aureus*<sub>1027</sub>. BHI broth agar plates without additives were poured under sterile working conditions, inoculated with a glycerol stock of a bacterial culture, and incubated overnight at 37°C. A grown culture of the overnight plate was used to inoculate a preculture in 5 mL BHI broth, which was then subjected to 37 °C incubation in a shaking bath. One mL of preculture was transferred to 100 mL broth without further additives to inoculate the main culture. The culture was incubated for 2 h at 37°C in a shaking bath. *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub> were cultivated until the logarithmic growth phase. The growth of the bacteria was determined by optical density (OD) measurement at 600 nm, with OD<sub>600</sub>=0.5 corresponding to 5.0\*10<sup>7</sup> bacteria mL<sup>-1</sup>. Bacteria were heat-killed for 60 min at 80°C in a shaking water bath. The killing of bacteria was tested by

spreading 50  $\mu\text{L}$  of killed bacteria on BHI broth agar plates, observing growth or, preferably, lack of growth. Main cultures were centrifuged for 15 min at  $2100 \times g$  at  $15^\circ\text{C}$ . Bacterial cell pellets were washed twice with 5 mL of cell culture media RPMI 1640 (Biochrom) without additives using the same centrifugation conditions. The bacteria were resuspended in RPMI 1640 at a  $5.0 \times 10^8$  bacteria  $\text{mL}^{-1}$  concentration and used for protein determination in 96-well microtiter plates. To this end, a bovine serum albumin (BSA) standard dilution series (0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 30, 40, 50, 75, 100  $\mu\text{g mL}^{-1}$ ) and sample dilution of 1:20, 1:40, 1:100, and 1:200 were prepared. For the measurements, diluted BSA and samples were aliquoted in triplicates into a 96-well plate with a fill height of 360  $\mu\text{L}$  per well. The protein absorbance was measured at 595 nm on a microplate reader. Linear regression was calculated using the BSA standard curve, and the protein concentration of the bacteria was determined. The bacterial concentration was then adjusted to 1  $\text{mg mL}^{-1}$ , aliquoted into 1.5 mL reaction vials, and stored at  $-20^\circ\text{C}$  until further use for cell experiments with pathogens.

## 2.5 Cell Biological Techniques

### 2.5.1 Primary Cell Extraction

Fist-sized tissue samples from the udder of three individual healthy cows provided the starting material for primary cell isolation. A day before sampling, 5 L Hanks' Buffered Salt Solution (HBSS) buffer was prepared. Collagenase (Serva) was dissolved in HBSS

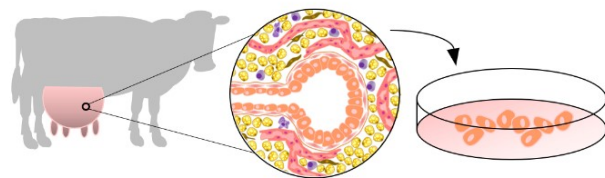


Figure 2-2: Origin of primary mammary epithelial cells for *in vitro* studies

buffer and diluted to reach 20,000 collagenase degrading units (CDU)  $\text{mL}^{-1}$ . Unused aliquots of the collagenase were stored at  $-20^\circ\text{C}$ . The collagenase was diluted 1:100 for usage.

At the slaughterhouse, the udder was placed in a bowl with teats facing up. The udder was washed with water and then extensively cleaned with 70% ethanol. The teats were cut open from two udder quarters and attached with surgical clamps to pull the tissue upward. A knife was used to cut out a fist-sized tissue sample that contained the large milk ducts. The tissue was placed in a beaker with warm HBSS buffer. The beaker in a Styrofoam box was transported immediately into the cell culture laboratory. The primary cell preparation was performed in a biological safety cabinet under conditions that minimized contamination. The tissue samples were rinsed thrice with HBSS buffer, cut into slices and strips with a knife, and smaller pieces with a scalpel. The cutting process optimized the access of collagenase to tissue digestion. The small tissue pieces were washed thrice with HBSS to rinse out milk and blood residues before being placed into a glass bottle containing collagenase in HBSS ( $200 \text{ CDU mL}^{-1}$ ). The bottle underwent a 45-min incubation in a  $37^\circ\text{C}$  shaking water bath to allow initial enzymatic digestion. The first digest was sieved through a household sieve. The tissue pieces held back in the sieve were placed back into the bottle. Fresh collagenase was added to continue digestion for another 45 min at  $37^\circ\text{C}$  under shaking conditions. The digestion was repeated two more times. The second, third, and fourth digest flow-through was collected and processed individually.

Sieving digested tissue through a household sieve into a glass dish with HBSS buffer helped to dilute the collagenase rapidly. After that, the digest was filtered through a sieve tower with different pore sizes from big to small size (500  $\mu\text{m}$ , 300  $\mu\text{m}$ , 150  $\mu\text{m}$ , and 90  $\mu\text{m}$ ). The filtrate was distributed into 14 50 mL reaction tubes and filled with HBSS. Again, adding a large volume of HBSS decreased the collagenase concentration in the digested tissue. The tubes were centrifuged for 10 min, at 280 x g, at 15°C without brake in a swinging-bucket rotor assembly in an Allegra X12R benchtop centrifuge (Beckman). The wash buffer was aspirated. Cells were resuspended in HBSS for washing before centrifuging. The pelleted cells from the different tubes in a digestion round were combined into one tube by washing at least five times with HBSS and repeated centrifugation. Before the last centrifugation, the cells were sieved through a falcon sieve of 100  $\mu\text{m}$  pore size. After the last centrifugation, the cell pellet was resuspended in FCS and stored at 2-8°C in a refrigerator. The third and fourth digestions were handled the same way until the resuspension in FCS. Finally, the cells were combined as 3 mL cell suspension in one reaction tube. The cells were prepared for long-term storage in liquid nitrogen. FCS containing 20% DMSO was prepared in one tube. The cells suspended in FCS and FCS with DMSO were combined and brought to a final DMSO concentration of 10% by carefully pipetting the mix up and down. The cells were transferred into cryotubes (1 mL each) and gradually cooled down in a Mr. Frosty™ freezing box to -80°C for one day. Then, the cryotubes were transferred into the storage boxes in liquid nitrogen tanks.

As a control of the primary cell extraction, a small aliquot of cells was resuspended in RPMI 1640 complete media (composition in Table 2-1) and seeded into collagen-coated cell culture vessels (Greiner Bio-One). Cultivation of pbMEC took approximately one week. During this time, the RPMI complete media was changed every second day. The cell morphology and density were surveyed using an inverse cell culture microscope (Nikon TMS-F). Proper activity and proliferation of primary cells were enabled by simulating *in vivo* conditions in an artificial culture environment by supplying amino acids, vitamins, growth factors, and appropriate pH. The culture was realized at stable environmental conditions of 37°C, 5% CO<sub>2</sub>, and humidified air. Before using primary cells for experiments, they were selectively trypsinized, as explained in 2.5.4.

### 2.5.2 Cell Culture

Cell culture is the process of growing cells in a well-designed artificial environment. Cell culture requires a change of media and division into subcultures. Nutrients, cell surface, and cell contacts limit cell growth. All cell culture procedures were performed under sterile working conditions in a biological safety cabinet. The two mammary epithelial cell models included in this study were cultured in humidified incubators (Binder) at 37°C and 5% CO<sub>2</sub> in the required growth medium with additives. MAC-T cells grew on tissue culture-treated vessels, while pbMEC required additional collagenase-treated growth surfaces (both Greiner Bio-One). It was part of the regular supply to provide the cells with fresh medium twice to thrice a week. At 80% confluence, the cells were passaged onto new cell culture dishes to enable continuous growth. Otherwise, the cells were used or seeded for experiments

into vessels of the appropriate size (for DNA extraction 10 cm dish, for RNA extraction 6- or 12-well plate). For passaging, cells grown in 10 cm culture dishes were washed with PBS and incubated with 1 mL trypsin/ ethylenediamine tetraacetic acid (EDTA) solution for 5 min (MAC-T) or 10 min (pbMEC) at 37°C. Cells detached from the culture vessel surface were resuspended in fresh medium and centrifuged for 5 min at 230 x g and 15°C, with a maximum centrifuge brake. The supernatant was aspirated. The cell pellet was resuspended in fresh medium. Cell number was determined in a Fuchs Rosenthal counting chamber. The trypan exclusion method was used to distinguish live from dead cells. Toward the end of the experiments, an automatic cell counter (Cellometer Auto 2000, Nexcelom Bioscience) complemented cell counting. A defined number of cells was seeded into culture vessels for experiments, kept in culture as a backup (~25,000 cells/cm<sup>2</sup>), or frozen for cell storage.

### **2.5.3 Cell Thawing and Freezing**

Cells stored in liquid nitrogen were briefly thawed in a 37°C water bath, resuspended, and washed in a pre-warmed medium aliquot. The freezing medium was removed by centrifugation (5 min at 230 x g, 15°C). The cell pellet was resuspended in fresh, complete growth medium and transferred to a cell culture dish. Cultivation was performed under constant environmental conditions. After 24 h, the growth of the cells was checked under an inverted cell culture microscope, and the media was replaced with fresh media.

In the initial step for preserving the mammary epithelial cells in liquid nitrogen, the medium was removed from cells grown in a culture vessel. Then, the cells were washed with PBS and detached from the culture vessel surface with trypsin/EDTA solution. The detached cells were resuspended in fresh medium and centrifuged for 5 min at 230 x g and 15°C. For MAC-T cell preservation, 1 x 10<sup>6</sup> cells were resuspended in 1 mL complete growth media containing 10% DMSO and transferred to cryotubes (Greiner). The preservation of pbMEC was performed in FCS complemented with 10% DMSO. The cells were gradually cooled to -80°C in a Mr. Frosty™ freezing box and stored in liquid nitrogen tanks.

### **2.5.4 Selective Trypsinization**

Co-culture of fibroblasts and primary epithelial cells was not desired in this study. Fibroblast cells were selectively detached from pbMEC in culture by a gentle method with trypsin/EDTA. The enzyme trypsin acts at 37°C and catalyzes the hydrolysis of peptide bonds between the carboxy group of arginine or lysine<sup>169</sup>. Trypsin detaches adherent cells from the growth surface<sup>170</sup>. Because of its high affinity for metal ions as a chelating agent, EDTA supports the trypsin action and weakens the cell-cell and cell-matrix interaction<sup>170</sup>. Fibroblasts, which are long, narrow cells, could be detached more rapidly than epithelial cells with their typical cobblestone morphology. For cells in a highly confluent monolayer, the culture medium was aspirated. They were rinsed twice with PBS to remove medium residues that could negatively affect trypsin action. Prewarmed trypsin (1 mL) was added to a 10 cm culture dish incubated at 37°C and 5% CO<sub>2</sub> for 5 to 10 min. The detachment of fibroblasts was monitored under an inverted microscope (Nikon). The trypsin action was stopped after a maximum of 10 min by adding media, and floating fibroblasts were removed. The timing was crucial in this step. Renewed addition of

trypsin and incubation at 37°C for 10 min detached pbMEC. The remaining pbMECs attached to the surface were mechanically detached with a cell scraper. Cells were resuspended and washed off the growing surface with growth media. Then, the cells were pelleted by centrifugation (5 min at 230 x g and 15°C) and either resuspended in complete growth media and seeded for experiments or resuspended in freezing media for cell storage.

### **2.5.5 Cell Exposure Experiments**

As the experimental design (see 2.1) outlined, mammary epithelial cells were exposed to mastitis pathogens or remained unexposed. First, cells were seeded in 10 cm culture plates for the ChIP experiment and 6-well plates for RNA extraction to proceed with qPCR or RNA-seq. At a density of  $1.1 \times 10^5$  cells  $\text{cm}^{-2}$ , which equals 80% confluency, the cells were challenged with pathogens *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub> for 1 or 3 h or remained unchallenged as a control (mock). For this purpose, cells were washed with PBS. An antibiotic-free medium was added. The bacterial exposure plates contained antibiotic-free medium with 30  $\mu\text{g}$  of heat-killed bacteria per mL of culture medium. At the end of the challenge, the cells were examined under the microscope, washed with PBS, and harvested according to the subsequent processing steps.

## **2.6 Molecular Biological Technique**

### **2.6.1 Nucleic Acid Isolation from Cells**

The Direct-zol™ RNA Miniprep kit from Zymo Research was applied to isolate the total RNA from cells instead of using chloroform, phase-separation, and precipitation steps (appendix A14). According to the manufacturer's instructions, all centrifugation and working steps were performed at room temperature (RT). The animal cells grown in 6 or 12-well plates (Cellstar Greiner Bio-one) were harvested in TRIzol Reagent (Ambion®, Life Technologies) with 100  $\mu\text{L}$   $\text{cm}^{-2}$ . The cells in TRIzol were stored at -20°C until RNA was prepared. On the preparation day, the cells lysed in TRIzol were thawed at RT.

Meanwhile, DNase and DNase reaction buffer were thawed on ice. At a ratio of 1:1 volume to sample homogenate, 95-100% ethanol was added to each sample and mixed with a vortex mixer. The mixture was transferred onto a Zymo-Spin™ IIC Column in a collection tube and centrifuged for 1 min at 12,000 x g. The column was transferred into a new collection tube and washed with 400  $\mu\text{L}$  RNA wash buffer, applying a centrifugation step at 12,000 x g for 30 s. DNase was diluted in reaction buffer and nuclease-free H<sub>2</sub>O. 80  $\mu\text{L}$  of the DNase reaction mix was loaded directly onto the column and incubated for 15 min at RT. The column was centrifuged at 12,000 x g for 30 s and washed twice with 400  $\mu\text{L}$  of Direct-zol™ RNA PreWash solution (centrifugation for 1 min at 12,000 x g at RT). A washing step with 700  $\mu\text{L}$  of RNA Wash Buffer followed using the same centrifugation conditions. The column was then centrifuged at maximum speed for 2 min before being transferred into an RNase-free reaction tube. Elution of RNA was performed in 20  $\mu\text{L}$  nuclease-free H<sub>2</sub>O (included in the kit) (centrifugation for 1 min at 12,000 x g). The RNA was stored at -20°C for subsequent use. A small aliquot of RNA was



used to determine the yield using a Nanodrop® spectrophotometer (Peqlab, Thermo Fisher Scientific) and to assess RNA quality on a 1.2% denaturing agarose gel.

The RNeasy Plus Mini Kit was used for RNA extraction from cells for the RNA-seq experiment, which required the elimination of genomic DNA (appendix A16 for full manual instructions). Briefly, cells were lysed and homogenized, and genomic DNA (gDNA) was removed in a column-based approach. The flow through containing total RNA was complemented with ethanol and then bound to a column, which was washed, and RNA was eluted. Downstream RNeasy Plus Mini Kit, RNeasy® MinElute® Cleanup (Qiagen) was applied to remove excess DNA before sequencing. RNA purity was validated by control PCR with PGC (Peroxisome Proliferator-Activated Receptor Gamma Coactivator 1-Alpha) primers. The quantity of RNA was determined with a Nanodrop® spectrophotometer and Qubit 1.0 Fluorometer (both Thermo Fisher Scientific). RNA quality was checked using the Agilent Bioanalyzer 2100 and found good with an RNA integrity number (RIN) equal to or greater than 8.

The gDNA was extracted from MAC-T cells with the Master Pure™ DNA Purification Kit (Lucigen) according to the manual instructions (appendix A18). The gDNA was RNase treated in an additional step, quantified, and used for PCR or qPCR in serial dilutions.

### **2.6.2 Determination of the Concentration and Purity of Nucleic Acids**

Nucleic acid concentration and purity were determined spectrophotometrically by use of a small volume (1-2  $\mu\text{L}$ ) on the NanoDrop®ND-1000 (Peqlab, until 2015) or ND-2000 (Thermo Fisher Scientific) UV/Vis spectrophotometer. Based on the Lambert-Beer equation, which assumes a linear relationship between absorbance and concentration, the quantity of nucleic acids was calculated and given in the instrument software. An absorbance value at 260 nm of 1.0 corresponds to 40 ng  $\mu\text{L}^{-1}$  RNA or 50 ng  $\mu\text{L}^{-1}$  double-stranded DNA. The quality of the nucleic acid was assessed by determining the absorbance quotient of the 260 nm and 280 nm wavelengths - A260/A280. For good quality, the DNA quotient's value was between 1.8 and 2, while for RNA, a value above 2 was acceptable.

The RNA for RNA-seq was quantified with a Qubit system and Qubit™ RNA HS Assay Kit (Thermo Fisher Scientific). In addition, these RNA samples were examined for their spectrometric properties with the 2100 Bioanalyzer instrument (Agilent) to determine the RNA integrity number (RIN) according to the manufacturer's instructions.

DNA samples used for ChIP experiments and subsequent sequencing were analyzed for quality and quantity with facilities at Diagenode, which were Qubit (Thermo Fisher Scientific), Fragment Analyzer (Advanced Analytical), and 2100 Bioanalyzer (Agilent). All assays were performed according to the manufacturer's instructions. For DNA quantification on the Qubit device, broad range (BR) and high sensitivity (HS) assay kits (Thermo Fisher Scientific) were applied. DNA Assay 1000 was used to quantify DNA and cDNA before sequencing.

### 2.6.3 Polymerase Chain Reaction

Primer-mediated amplification of specific DNA fragments was performed using the polymerase chain reaction (PCR) method<sup>171</sup>. PCR was used during this project in different variants (standard PCR, touch-down PCR, semi-nested PCR, qPCR) and employing different nucleic acid starting material (RNA transcribed into complementary DNA (cDNA), genomic DNA (gDNA), and plasmid DNA). Customer-specific oligonucleotides were purchased from Sigma-Aldrich and Integrated DNA Technologies, Inc. Primers are listed in appendix A1 unless otherwise indicated. Thermocyclers (Biometra) and different qPCR devices (Roche Life Science) were used throughout the present study. PCR reactions were performed in 0.2 mL reaction tubes (Sarstedt), while qPCR was performed with 32 capillaries per carousel (Roche Life Science) or 96-well plates (Roche Life Science). In the course of the experiments, the use of the number of cycles, temperature profiles, polymerases, and primers were explicitly adapted to the requirements. PCR products were loaded on agarose gels to visualize the amplicons.

#### *Standard PCR with Taq-Polymerase*

A standard PCR recipe for testing the melting temperature of each of the ChIP-relevant primers designed with small fragments in the promotor region of *CXCL8*, *CCL20*, *NOS2A*, *IL6*, *TNF*, *actin (ACTB)*, *testis-specific histone H2B variant (TSH2B)*, and *LAP* is listed in Table 2-4. PCR was performed with GoTaq® DNA Polymerase and Master Mix (Promega). The thermocycler program for amplifying these standard PCR reactions is provided in Table 2-5. After amplification, all PCR products were loaded onto an agarose gel for assessment. The same recipe was used for PGC control PCR for RNA samples testing for remaining DNA. The primers are derived from the Peroxisome Proliferator-Activated Receptor Gamma Coactivator 1-alpha gene (Table 2-6), which regulates mitochondrial biogenesis. These were established in the laboratory routine for verification of RNA before RNA-seq.

Table 2-4: PCR reaction mixture with Taq polymerase

Stock solutions	1x reaction ( $\mu$ L)
Primer forward (20 pmol/ $\mu$ L)	0.2
Primer reverse (20 pmol/ $\mu$ L)	0.2
dNTPs (10 mM)	0.8
5x GoTaq-Buffer (7.5 mM MgCl <sub>2</sub> )	2
GoTaq-Polymerase (5 U/ $\mu$ L)	0.05
Nuclease-free H <sub>2</sub> O	5.25
DNA or H <sub>2</sub> O control	1.5

Table 2-5: Cycler Program used for Taq PCR

Step name	Temperature (°C)	Time (sec)	Cycles
Initial denaturation	94	60	1x
Denaturation	94	30	34x
Annealing	TM (Variable)	60	
Extension	72	30	
Final elongation	72	420	1x
Gradual cooling	15	60	1x
Final hold	6		PAUSE

*Semi-nested PCR with PFU and Taq Polymerase for cloning*

A semi-nested PCR allowed the amplification of a target sequence by two consecutive PCR runs. For the second PCR run, the forward primer was used again, and the used reverse primer was located within the target sequence. In preparation for the cloning procedure, targeted gene fragments were amplified first with *Pyrococcus furiosus* (PFU) DNA polymerase (Thermo Fisher Scientific), generating blunt-ended DNA fragments. The reaction volume per sample was 10  $\mu$ L. The extension time was set to 2 min  $\text{kb}^{-1}$  according to the respective fragment size, which varied from ~1500 bp to ~4600 bp.

Table 2-6: Primers for measuring RNA purity

Primer Name	Sequence (5'-3')	Fragment Size (bp)
PGC_E4	AAGAAGCTCTTACTG	1178
U1	GCACC	
PGC_E5	ATGTTGTGTCTGCGAT	
L1	TGTG	

Table 2-7: PCR reaction mix for PFU polymerase

Stock Solutions	1x Reaktion ( $\mu$ L)
Primer forward (20 pmol/ $\mu$ L)	0.1
Primer reverse (20 pmol/ $\mu$ L)	0.1
dNTP Mix (2 mM)	0.2
10xPfu-Puffer with $\text{MgCl}_2$	1
PFU DNA Polymerase (2.5 U/ $\mu$ L)	0.15
Nuclease-free $\text{H}_2\text{O}$	6.95
DNA or $\text{H}_2\text{O}$ control	1.5

Table 2-8: Cyclor Program used for semi-nested PFU PCR. Extension time was adapted to 2 min/kb up to 2kb, and for larger products prolonged by 1 min/kb.

Step name	Temperature ( $^{\circ}\text{C}$ )	Time (sec)	Cycles
Initial denaturation	95	180	1x
Denaturation	95	30 s	34x
Annealing	TM (Variable)	60	
Extension	72	120/kb	
Final elongation	72	420	1x
Gradual cooling	15	60	1x
Final hold	6		PAUSE

In the event of challenges during the cloning procedure, a switch of polymerases was made so that large fragments were amplified for cloning with Taq polymerase according to the same recipe as for PCR with PFU polymerase. Buffer and polymerase for the semi-nested PCR were replaced accordingly and ran with the cyclor program given in Table 2-9.

Table 2-9: *Cycler Program used for Taq semi-nested PCR*

Step name	Temperature (°C)	Time (sec)	Cycles
Initial denaturation	95	180	1x
Denaturation	95	30	34x
Annealing	TM (Variable)	60	
Extension	72	45 - 120	
Final elongation	72	420	1x
Gradual cooling	15	60	1x
Final hold	6		PAUSE

*Colony PCR*

A colony PCR was performed according to the CloneJET PCR Cloning Kit (Thermo Scientific) instructions. The reaction volume was 10  $\mu\text{L}$  containing 1.5  $\mu\text{L}$  sample and 8.5  $\mu\text{L}$  enzyme master mix according to the standard reaction mix. The pJET forward and reverse primers were used in 10 pmol  $\mu\text{L}^{-1}$  stock concentration. The PCR cycler program is given in Table 2-10.

Table 2-10: *Cycler Program used for Colony PCR*

Step name	Temperature (°C)	Time (sec)	Cycles
Initial denaturation	95	180	1x
Denaturation	94	30	25x
Annealing	60	30	
Extension	72	120 (60 s/kb)	
Final elongation	72	420	1x
Gradual cooling	15	60	1x
Final hold	6		PAUSE

*Reverse Transcriptase PCR*

RNA was transcribed into cDNA with SuperScript<sup>TM</sup>II Reverse Transcriptase (Invitrogen) and gene-specific cDNA primers for immune response genes (IL6, JMJD3, iNOS, IL8, CCL20, TNF $\alpha$ ) and one housekeeping gene chloride intracellular channel one (CLIC1). A 20  $\mu\text{L}$  reaction volume contained 9  $\mu\text{L}$  primer master mix and 8  $\mu\text{L}$  enzyme master mix. The remaining volume was added up with 500 ng RNA diluted in nuclease-free  $\text{H}_2\text{O}$ . First, the primer master mix was pipetted to each RNA sample and incubated at 70°C for 10 min, then transferred to ice, and the enzyme master mix was added to each sample. The mix was incubated at 42°C for 1 h in a thermocycler. The cDNA was then purified in a column-based approach of the High Pure PCR Product Purification Kit (Roche Life Science). The cDNA was eluted in 100  $\mu\text{L}$  nuclease-free  $\text{H}_2\text{O}$  and was used as a template for amplification in qPCR. Samples were stored at -70°C until use in experiments.

Table 2-11: List of cDNA primers used for the reverse transcription.

Primer Name	Sequence
BT_IL6_r2	GGGAGCCCCAGCTACTTCAT
BT_JMJD3_r2	AACACCTCCACGTCGCACT
BT_iNOS2_r2	CCGGGGTCCTATGGTCAAA
BT_CCL20_r2	TTCCATCCCCAAAAGCATCC
BT_CLIC1_r2	GATCCCCTCATCCTCAGCAC
BT_NR4A2_r2	AATAGTTGGGGCGGTTCAAA
BT_IL8_r1	CATGGAACAATGTACATGCGAC
BT_TNFa_r1	CTGTGAGTAGATGAGGTAAAGC
Oligo(dT)20	TTTTTTTTTTTTTTTTTTTT

## 2.6.4 Quantitative PCR

In contrast to conventional PCR, which was used to amplify the sample and to provide a qualitative endpoint analysis, real-time or quantitative PCR (qPCR) can quantify how much of a particular target sequence is present in a sample. The quantitative measurement applied in this study is based on using DNA-binding fluorescent dyes and detecting the fluorescent signal as the cycles progress. An increase in the fluorescence signal correlates with the increased amount of the DNA amplicon. The recorded threshold value of the fluorescence signal is assigned to a threshold cycle (Ct) and is the starting point for underlying calculations. A low Ct represents a high, and a high Ct represents a low DNA concentration in the sample.

### Gene Expression Measurement by qPCR

Gene-specific plasmid DNA was used as a template in serial dilutions of copy numbers ( $10^6$ ,  $10^5$ ,  $10^4$ ,  $10^3$  and  $10^2$  copies/5  $\mu$ L) to establish a standard curve for the targeted gene sequence of immune response genes. Copy numbers for a standard counting plasmid (Table 2-13) were calculated as follows:

$$\text{Copy Number} = \frac{(\text{Amount of Template DNA in ng}) * \text{Avogadro's Constant}}{(\text{Template DNA Length in bp}) * (\text{Average Mass of 1 bp dsDNA}) * 1 * 10^9 \frac{\text{ng}}{\text{g}}}$$

$$\begin{aligned} \text{Avogadro's Constant} &= 6.0221 * 10^{23} \text{ mol}^{-1} \\ \text{Average mass of 1 bp dsDNA} &= 660 \text{ g/mol} \end{aligned}$$

Table 2-13: Reference list for counting plasmids for gene expression measurement by qPCR.

Plasmid for Gene	Plasmid ID #	Primer ID#: Primer Name	Primer Sequence
BT_iNOS2	1520	N110: BT_iNOS2_f1	ACAGGATGACCCCAAACGTC
		N109: BT_iNOS2_r1	TCTGGTGAAGCGTGTCTTGG
BT_CCL20new	1554	N112: BT_CCL20_f3	CAGCAAGTCAGAAGCAAGCAA
		N111: BT_CCL20_r1	CCCACTTCTTCTTTGGATCTGC
BT_JMJD3	2292	N108: BT_JMJD3_f1	ACTGTGCTGCTGACGTTGTG
		N107: BT_JMJD3_r1	CATCCCGTTTGCTTTCCAG
BT_IL8	1324	N133: BT_IL8_f2	CCTCTTGTTCAATATGACTTCCA
		N132: BT_IL8_r3	GGCCCACTCTCAATAACTCTC
BT_TNFa	1252	N136: BT_TNFa_f1	GAGTTGATGTGCGCTACAACG
		N135: BT_TNFa_r2	CTTCTGCCTGCTGCACTTCG
BT_LAP	1404	N118: LAP_sf	AGGCTCCATCACCTGCTCCTT
		N117: LAP_sr	CCTGCAGCATTTTACTTGGGCT
BT_CLIC1	1959	N114: BT_CLIC_f1	AGAACAACCGCAGGTGCAAT
		N113: BT_CLIC_r1	GTCTCAGTCCGCCTCTTGGT
BT_IL6	1519	N106: BT_IL6_f1	GGAGGAAAAGGACGGATGCT
		N105: BT_IL6_r1	GGTCAGTGTGTTGTGGCTGGA

Table 2-12: Enzyme master mix of the reverse transcription reaction

Stock Solutions	1x Reaction ( $\mu$ L)
DTT (0.1 M)	2
dNTP's (10 nM)	1
5X First-Strand Buffer	4
Reverse Transcriptase	1

The series of diluted plasmid DNA was always measured on one of the qPCR instruments (LightCycler® 2 capillary Real-Time PCR system, LightCycler® 480 & 96; Roche Life Science) before the actual experiment. A negative control of nuclease-free water and a positive control of gene-specific plasmid DNA ( $10^6$  copies/5  $\mu$ L) were included for each primer set alongside the measured samples in each experimental run. A qPCR reaction volume was 10  $\mu$ L (5  $\mu$ L sample + 5  $\mu$ L Master Mixture). Recipes were adapted to the used SYBR green reagent, as were the cycler programs to the primers melting and product melting temperatures on the respective instrumentation. The melting temperature of the primers was set to 60°C, if possible. An equivalent of 75 ng of cDNA in 5  $\mu$ L and 400 nM primers were used for real-time qPCR. For measurements with LightCycler® 2, capillaries were positioned in the carousel using tweezers, letting them slip loosely into the holes. Master mixture (5  $\mu$ L), prepared under a CleanPrep station, was added at the upper edge into all capillaries. Water samples were added into negative control capillaries. DNA samples were loaded into individual capillaries. Capillaries were capped and sealed with a Roche capping key. The sample carousel was centrifuged using a carousel centrifuge (Roche), followed by visually verifying volume and bubbles before the run. For LightCycler® 480, reactions were pipetted into individual wells of a 96-well plate, which was sealed with transparent sealing film and centrifuged before measurement. For each of the individual qPCR machines, specific machine software from Roche Life Science (LightCycler® 3 Data Analysis Module, LightCycler® 480 SW 1.5, and LightCycler® 96 SW 1.1) was used to analyze melting curves, the efficiency of the PCR, Ct values, and export of data for analysis. Based on the quality matrix of the standard curve, including Y-intercept, slope, performance coefficient ( $R^2$ ), and efficiency, qPCR experiments were analyzed. A powerful and reproducible experiment required a linear standard curve with a high-performance coefficient ( $R^2 > 0.98$ ). A good reaction showed an efficiency between 90% and 110%, corresponding to a slope between -3.58 and -3.10<sup>172</sup>. All qPCR products were also loaded onto an agarose gel to confirm the reaction quality.

Copy numbers were calculated based on the established standard curve for the targeted gene sequence as follows:

$$\text{Copy Number} = 10^{\frac{Ct - \text{intercept}}{\text{slope}}}$$

Transcriptional measurements included two cytokines of the "immediate early" response (*IL8*, *TNF- $\alpha$* ), the "late early" cytokine *IL6* and the cytokine *CCL20*, two factors with antimicrobial functions *NOS2A* and *LAP*, and *JMJD3*, which encodes a lysine-specific demethylase. The gene expression analysis data were normalized to an internal control gene (*CLIC1*). For this purpose, *CLIC1* was measured in a standard curve using a dilution series of plasmid DNA and measured for all experimental samples. Copy numbers were determined based on the Ct. The average of *CLIC1* copy numbers was multiplied by the sample copy for the specific gene, which was then divided by the determined *CLIC1* copy for each sample. Fold enrichment was calculated by the ratio of bacterial-challenged samples versus unchallenged control (= t0).

Table 2-14: Reaction Master Mixture for LightCycler® 2 capillary Real-Time PCR system (Roche)

Reaction components	1x reaction volume (μL)
H <sub>2</sub> O	2.6
Forward primer	0.2
Reverse primer	0.2
Fast Start DNA Plus SYBR Green 1	2
DNA or H <sub>2</sub> O control	1.5

Table 2-15: Reaction Master Mixture for LightCycler® 480 (Roche)

Reaction components	1x reaction volume (μL)
H <sub>2</sub> O	0.6
Forward primer	0.2
Reverse primer	0.2
LightCycler® 480 SYBR Green I Master (2x)	6
DNA or H <sub>2</sub> O control	5

Table 2-16: qPCR program for LightCycler® 2 capillary Real-Time PCR system (Roche)

Program Step name	Temperature Target (°C)	Hold Time (sec)	Slope (°C/sec)	Acquisition	Cycles
Preincubation	95	600	20	None	1
Amplification	95	15	20	None	40
	Primer T <sub>m</sub>	10	20	None	
	72	30	20	None	
	Product T <sub>m</sub>	5	20	Single	
Melting Curve	95	0	20	None	1
	65	15	20	None	
	95	0	0.1	Continuous	
Cooling	40	60	20	None	1

Table 2-17: qPCR program for LightCycler® 480 (Roche Life Science)

Program Step name	Temperature Target (°C)	Hold Time (sec)	Ramp Rate (°C/sec)	Acquisition	Cycles
Preincubation	95	600	4,40	None	1
Amplification	95	20	4,40	None	45
	Primer T <sub>m</sub>	15	2,20	None	
	72	15	4,40	None	
	Product T <sub>m</sub>	10	4,40	Single	
Melting Curve	95	0	4,40	None	1
	70	15	2,20	None	
	95	0	0,29	Continuous	
Cooling	40	30	1,50	None	1

### N-ChIP-qPCR

Primers in promoter regions of the immune response genes (*IL6*, *IL8*, *CCL20*, *TNF-α*, *NOS2A*, *LAP*) with putative binding sites for histone modifying enzymes were designed and used for N-ChIP-qPCR. Each primer pair spans a short fragment within the promoter region. ChIP primers for *actin* and *TSH2B* served as control regions for histone modifications and thus were included in measurements. The standard counting plasmids for N-ChIP-qPCR, spanning each of the entire gene sections, were established through PCR, cloning, and verification by sequencing. Standards consisted of a 10-fold dilution series of the respective plasmid DNA ranging from 10<sup>6</sup> to 10<sup>2</sup> copies/5 μL. Genes (*IL6*, *IL8*, *CCL20*, *TNF-α*, *actin*) were measured by qPCR after N-ChIP of cells with or without bacterial contact

targeting lysine modification on histone H3. Measurement with a 20 ng DNA template was performed with LightCycler® 2 or 480 (both Roche Life Science). Copy numbers were calculated for all ChIP (H3, H3K9me2, H3K14Ac, IgG) and input DNA with slope and intercept for each standard curve of the measured gene target regions based on the Ct values. This calculation was followed by normalizing the copy numbers to the internal control gene *actin* by multiplying the copy number of a target gene with the determined copy numbers of *actin*, divided by the average of all *actin* measurements. Next, a fold enrichment of the ChIP sample relative to the IgG sample was calculated.

#### *X-ChIP control qPCR*

Primer pairs for positive and negative genomic regions for each histone mark were developed for both bovine epithelial cell models to validate X-ChIP-seq. Sets of primers (Table 2-19) were tested on a serial dilution of bovine gDNA (isolated from MAC-T with EpiCentre MasterPure DNA Kit) using the LightCycler® 96 (Roche Life Science), establishing a standard curve for DNA amplification. To this end, duplicates of a 2-fold serial dilution series with concentrations ranging from 8 ng to 0.125 ng bovine gDNA were measured. Melting curves, primer efficiency, and amplification of a single, specific product were criteria for final primer selection. ChIP-qPCR measurement on the LightCycler® 96 System included the input DNA, a no amplification control (H<sub>2</sub>O), the immuno-precipitated histone sample, and a negative IgG control for each positive and negative control primer set. Per well of a 96-well-plate, 5 µL of the sample was added to 15 µL of qPCR mix (10 µL SyberGreen (Kapa Biosystems), 1 µL primer (10 µM total primer concentration, forward+reverse), 4 µL DNase/RNase-free water). The qPCR program is given in Table 2-20. The recovery or ‘% Input’ was calculated with the formula:

$$\% \text{ Input} = 2^{(\text{Ct Input} - \text{Ct ChIP sample})}$$

The % Input is the relative amount of immunoprecipitated DNA compared to input DNA. Fold enrichment was calculated, putting into relation recovery on the target DNA (positive control) over background (negative control) for the appropriate histone mark. A 5-fold enrichment of target DNA should be observed in the ChIP sample compared to a negative control <sup>173</sup>.

Table 2-18: Reaction Master Mixture for LightCycler® 96 (Roche Life Science)

Reaction components	1x reaction volume (µL)
H2O	4
Forward primer	0.5
Reverse primer	0.5
Kappa SYBR Fast qPCR Master Mix (2x)	10
DNA or H2O control	5



Table 2-19: X-ChIP control-qPCR primer. The list contains the sequence for Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), protease, serine one (PRSS1), and testis-specific histone H2B variant, including fragment size of the primers used for X-ChIP-qPCR; the melting temperature for all primers was 60 °C.

Primer name	Sequence 5' → 3'		Amplification product [bp]
	Forward	Reverse	
GAPDH_b	GGGATTCCATCCTCTTAGCC	TCACTCCGACCTTCACCATC	91
PRSS1	CACACTCGAGACAACCAAGG	TCACAGAACAGGGCTACAGG	70
TSH2B	TGACGGTAGGTTGTGCTTTGC	TGGATACCTGGGAGTGTTCATGT	107

Table 2-20: qPCR program for LightCycler® 96 (Roche Life Science)

Program Step name	Temperature Target (°C)	Hold Time (sec)	Ramp Rate (°C/sec)	Acquisition	Cycles
Preincubation	95	600	4,40	None	1
Amplification	95	10	4,40		45
	60	30	2,20		
	72	30	4,40		
Melting Curve	95	10	4,40	None	1
	55	60	2,20	None	
	97	1	0,20	Continuous	
Cooling	40	30	1,50	None	1

### 2.6.5 Agarose Gel Electrophoresis

Agarose gels were used to determine nucleic acid's size and purity. Those macromolecules migrate through an agarose matrix based on the size and charge. The molecule separation was achieved by adjusting the agarose concentration (w/v) from low (0.8%) for large DNA molecules (gDNA, plasmid DNA) to high (2%) for small DNA fragments (PCR amplicon). 1x Tris-acetate-EDTA (TAE) served as DNA electrophoresis running and gel preparation buffer. Ethidium bromide was used as a nucleic acid intercalating dye at a concentration of 0.5 µg mL<sup>-1</sup>, allowing visualization of DNA or RNA by a UV transilluminator. Before loading onto the gel, samples were mixed with a loading dye to allow the sample to sink into the well and follow the subsequent gel run with a color indicator. Separation was performed at a voltage of 5 – 10 V cm<sup>-1</sup>. The nucleic acid ladder pBR328 (Carl Roth) was loaded onto the gel as a comigrating nucleic acid size standard.

For visualization of RNA, a 1.2% RNA denaturing agarose gel was prepared. To this end, 0.6 g agarose was mixed with 5 mL 10x formaldehyde (FA) Gel Buffer and 45 mL RNase-free water; the mixture was heated in a microwave. Ethidium bromide and 0.9 mL 37% formaldehyde were added to the cooled mixture before the gel was poured, solidified for 30 min, and equilibrated in 1x FA Gel Running Buffer for another 30 min. Before loading, 2 µL 5x RNA loading dye was added to an equivalent of 500 ng RNA in 8 µL H<sub>2</sub>O, incubated for 5 min at 65°C, and transferred on ice. The gel ran at 80 V for 60 min and was visualized with a UV transilluminator.

Sharp bands on the gel represent either 28S and 18S rRNA of intact RNA or DNA, e.g., of a specific gene amplicon.

### 2.6.6 Purification of DNA from Agarose Gel

The Kit Nucleo Spin Extract II (Macherey-Nagel) and Cycle Pure Kit (Peqlab) were used to elute DNA fragments of specific sizes from agarose gels.

After electrophoresis, the corresponding band was cut out of the gel with a scalpel. Isolation was performed according to the manufacturer's

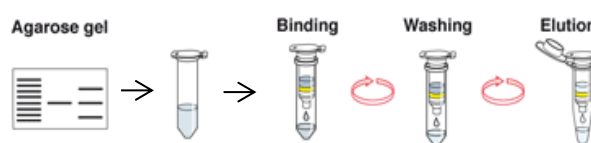


Figure 2-3: DNA purification from agarose gel  
Adapted from Macherey Nagel illustration.

instructions. All centrifugation steps were performed at RT at 10,000 x g. After weighing and dissolving the gel in the appropriate buffer, the DNA was bound to a silica gel membrane, washed, and eluted with 30  $\mu$ L ultrapure nuclease-free H<sub>2</sub>O. The quantity and quality of DNA were determined.

### 2.6.7 Competent Cells

The protocol for preparing competent cells was adapted to Hanahan (1983)<sup>174</sup>. All recipes for prepared agar and solutions can be found in appendix A4. TY agar plates supplemented with antibiotic tetracycline (15  $\mu$ g mL<sup>-1</sup>) were poured. A flamed inoculation loop was dipped into a glycerol stock of *E. coli* XL1-Blue (Stratagene) and plated on TY-agar plates with tetracycline. After overnight incubation at 37°C, individual colonies were picked and transferred into 5 mL Hanahan's Broth (Super Optimal Broth (SOB) Medium) in separate culture tubes, which were incubated overnight at 37°C in a shaker. With 1 mL of the preculture, 100 mL of the main culture (Psi-Medium including MgSO<sub>4</sub>) was inoculated. The main culture was grown in an Erlenmeyer flask until OD<sub>600</sub> between 0.45 and 0.55 was reached. Bacterial cells were pelleted by centrifugation (at 2100 x g, 4°C, for 15 min), then carefully resuspended in 33.3 mL RF1-solution. The resuspended cells were chilled on ice for 10 min before pelleting again by centrifugation (at 2100 x g, 4°C, for 15 min). The supernatant was discarded. The pellet was carefully resuspended in 8 mL RF2-solution and incubated on ice for 10 min. Bacterial cells were transferred in 40  $\mu$ L aliquots to 1.5 mL reaction tubes (Sarstedt), which were snap-frozen with liquid nitrogen and stored at -70°C until use.

A glycerol stock was prepared by removing the required volume from the main culture at the appropriate time and adding glycerol. The final glycerol concentration in 1 mL of bacterial stock was 17%. These bacterial stocks were also kept at -70°C until use.

As an alternative, chemically competent cells were produced by the use of TransformAid Bacterial Transformation Kit (Thermo Fisher Scientific) according to manufacturer's instructions using tryptone yeast (TY) agar plates supplemented with ampicillin, 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-Gal) and isopropyl-b-D-thiogalactopyranoside (IPTG) for blue/white selection of clones.

### 2.6.8 Cloning

Recombinant DNA technology enables the sequence determination of large DNA fragments and the large-scale amplification of standard counting plasmid DNAs for qPCR experiments. Figure 2-4 illustrates the cloning procedure.

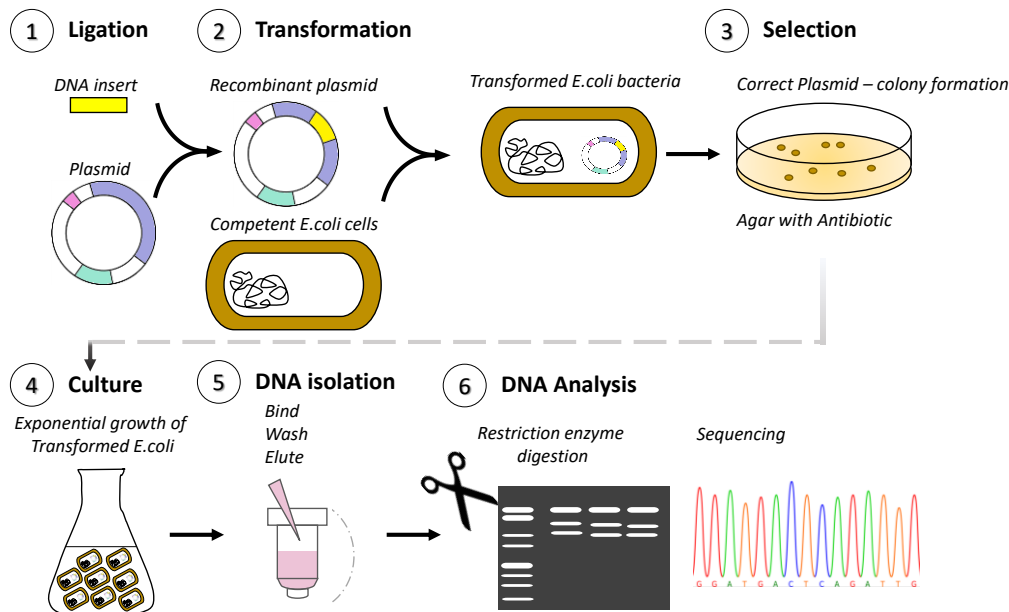


Figure 2-4: Cloning procedure. Steps of recombinant DNA technology include: (1.) ligation of plasmid and isolated gene of interest (DNA insert); (2.) transformation of the recombinant plasmid into competent bacteria cells; (3.) plating of transformed bacteria cells on agar with an antibiotic for selection of clones; (4.) large-scale culture of the picked colony of transformed bacteria; (5.) DNA isolation and (6.) analysis of the nucleic acid sequence. Figure adapted from Deans (2014)<sup>175</sup>.

For reproducible and reliable cloning, blunt-end and TA cloning techniques are common and applied in this work. The first technique uses the characteristic proofreading PFU polymerase to remove overhang DNA, thus producing blunt-end PCR products. The 5'-ends of the vector contain phosphoryl groups and can be ligated to blunt-end DNA fragments. The second technique uses the addition of a 3' adenine produced by Taq Polymerase at the DNA fragment and 3' thymidine at vector DNA. These are compatible and can be covalently enzymatically linked during ligation. The targeted gene fragments were amplified by PCR and purified with PCR Cleanup Gel Extraction Kit (Macherey-Nagel) and Cycle Pure Kit (Peqlab). During ligation, DNA samples were incorporated into vector DNA. With the CloneJET PCR Cloning Kit (Thermo Fisher Scientific), the blunt-end DNA fragments representing and spanning each of the targeted ChIP-Promoter regions (*IL6*, *IL8*, *CCL20*, *TNF- $\alpha$* , *NOS2A*, *LAP*, *TSH2B*, *actin*), were cloned into the cloning vector pJET1.2/blunt following instructions of Thermo Fisher Scientific. Gene fragments amplified with Taq Polymerase required an additional enzymatic step (DNA blunting) for cloning with vector pJET1.2. Counting plasmids for candidate immune response genes (Table 2-13) for transcriptomic studies were cloned into the pGEM®-T Easy vector (Promega) following the manufacturer's instructions. Recombinant plasmids were transformed into competent *E. coli* XL1 blue host cells, in which they were present as circular plasmid DNA. As desired, this recombinant plasmid DNA was replicated independently from the existing genetic material. Through antibiotic selection (ampicillin agar plates) of the transformed *E. coli*, only those bacterial cells containing the plasmid could multiply. Three individual colonies from those grown on agar plates were first picked and cultured in growth media on a small scale. Then, plasmid DNA was isolated, quantified, and verified by restriction enzyme digestion, colony PCR, and sequencing. Upon verification, the

plasmid DNA was finally processed on a larger scale with the NucleoSpin Plasmid kit (Macherey-Nagel) or maxi preparation protocol (appendix A17) with subsequent verification of DNA sequence.

### **2.6.9 Restriction Enzyme Digestion of Plasmid DNA**

Restriction enzymes are molecular scissors working in a sequence-specific manner. For this study, restriction enzymes EcoRI, ScaI, or XhoI (all Thermo Fisher Scientific) were applied. The starting material was 10 µg plasmid DNA. The mixture of 35 µL enzyme-specific reaction buffer, 5 µL restriction enzyme, and 10 µg plasmid DNA was topped up with nuclease-free H<sub>2</sub>O to the maximum volume of 350 µL. The digestion occurred at 37°C during overnight incubation. ScaI and XhoI linearized the plasmid DNA vector as desired. XhoI was applied for IL6 plasmid DNA. EcoRI cleaved out the desired insert from the pGEM®-T Easy vector. After digestion, an aliquot of the sample was loaded onto a low-percentage agarose gel. Passing this quality inspection, plasmid DNA was precipitated by adding 35 µL 3 M sodium acetate and 350 µL isopropanol. The mixture was incubated on ice for 10 min before centrifugation (15 min, maximum speed, at RT). The supernatant was decanted. The pellet was briefly dried and washed with ethanol (centrifugation 7 min, maximum speed) before eluting in 50 µL nuclease-free H<sub>2</sub>O. DNA concentration was measured. The plasmid DNA concentration was adjusted to 100 ng µL<sup>-1</sup>. The produced linearized plasmid DNAs served as standards for qPCR and allowed the detection of pipetting errors and the efficiency of qPCR.

### **2.6.10 DNA Sanger Sequencing**

The sequencing according to the Sanger principle<sup>176</sup> required for the present work was performed in the DNA sequencing laboratory of the genome biology department by Marlies Fuchs at the Leibniz Institute for Farm Animal Biology in Dummerstorf. For an initial verification of sequencing peaks and CTAG-code, the software BioEdit (version 7.2.5.0) was used. The sequences were compared with the nucleotide BLAST tool of the NCBI database (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) against existing sequences of the database<sup>177</sup>.

## **2.7 N-ChIP**

The protocol version F5 of the ChIP-IT® Express Enzymatic kit (Active Motif) was followed for the entire working procedure of N-ChIP<sup>178</sup>. The ChIP DNA Clean & Concentrator kit (Zymo Research) was used for subsequent column-based purification of ChIP DNA (instructions in appendix A19).

### **2.7.1 Cell Harvest and Enzymatic Shearing**

As starting material for the IP, cells were harvested in ice-cold 1x PBS and pelleted by centrifugation for 10 min at 1000 x g at 4°C. No cell fixation was performed for the native immunoprecipitation (IP) of the protein-DNA complex between DNA and histone, which are naturally linked. Before cell pellets were frozen at -80°C until further use, 1 µL 100 mM phenylmethylsulfonyl fluoride (PMSF) and 1 µL of protease inhibitor cocktail (PIC) were added. Preparation of sheared chromatin started by thawing the cell pellet on ice, then resuspending the pellet in 1 mL ice-cold Lysis Buffer supplemented with 5 µL PMSF + 5 µL PIC, remaining on ice for 30 min. According to the manufacturer's instructions, the enzymatic shearing cocktail was diluted with 50% glycerol to a working solution of 200 units mL<sup>-1</sup>.

Cells in the Lysis Buffer were mechanically sheared on ice in a cold Dounce homogenizer with a small clearance width by 10 to 15 strokes. This step involves pressing the cell solution through a narrow glass vessel while a pestle is slowly moved up and down<sup>179,180</sup> to achieve lysis and nuclear release of epithelial cells. The progress of lysis was monitored on a hemocytometer with a microscope, with nuclei looking like dots and surrounded by asymmetric debris. The nuclei were pelleted by centrifugation for 10 min at 2,400 x g at 4°C in a microcentrifuge, resuspended in 350 µL Digestion Buffer (supplemented with 1.75 µL PMSF and 1.75 µL PIC), and pre-incubated for 5 min at 37°C. Adding the enzymatic cocktail (17 µL) to lysed cells induces sequence-independent cleavage in open DNA regions between nucleosomes. The shearing time was optimized by testing a 5 to 15 min window. The reaction tubes were mixed every 2 min using a vortex mixer and incubated for 6 min 30 s at 37°C. The enzyme reaction was stopped by adding 7 µL ice-cold 0.5 M EDTA to each tube, then incubating on ice for 10 min. Sheared chromatin was centrifuged for 10 min at 16,000 x g at 4°C. The supernatant containing the sheared chromatin was transferred to a new 1.7 mL reaction tube, aliquoted, and stored at -80°C until further use. An aliquot was used to assess the quantity and quality of the sheared DNA. 1 µL RNase (included in the kit) was added to the sheared chromatin aliquot and incubated for 15 min at 37°C. Then, 10 µL of Proteinase K was added and incubated at 42°C for 1.5 h. DNA was phenol/chloroform extracted (according to A15) and precipitated to determine the DNA concentration of the sample and the shearing efficiency. The DNA sample mixed with 6x Loading Buffer was loaded onto a 1% TAE agarose gel run at 100 V for 45 - 60 min. The shearing process ideally generated single 150 bp to multiples of 150 bp fragments that showed a banding pattern on the gel.

### 2.7.2 Chromatin Immunoprecipitation

Antibodies were selected according to their validation for the ChIP method. Antibodies directed against active (H3K14Ac) and repressive (H3K9me2) modifications of histone H3, IgG as a negative control, and H3 to detect the endogenous concentration of the protein were included in this study, as listed in Table 2-21.

Table 2-21: Antibodies used for N-ChIP experiments. mAB= monoclonal antibody, pAB= polyclonal antibody

Antibody	Host	Reference#	Lot#	Company	ChIP application (µL)
H3K9me2 mAB	mouse	ab1220	GR138716-5	Abcam	2
H3K9me2 pAB (200 µL)	rabbit	39239	3208001	Active Motif	10
H3K14Ac pAB (200 µL)	rabbit	39599	11709001	Active Motif	10
normal rabbit IgG (200 µg/0.5 mL)	rabbit	SC-2027	L2414	Santa Cruz	2
H3 pAB	rabbit	2650S	N.A.	Cell Signaling	10

Sheared chromatin of all sample conditions (MAC-T challenged with mock, *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub>) was thawed on ice. 10 µL of input DNA per condition was set aside for later reprocessing. The reaction mixture for one IP consisted of 10 µg chromatin, the company's recommended µL of ChIP antibody, 25 µL G-protein-coupled magnetic beads, 20 µL ChIP Buffer I, 2 µL PIC and was filled up to 200 µL with nuclease-free H<sub>2</sub>O in siliconized 1.7 mL reaction tubes (Active Motif). The IP was incubated on a

rotating wheel overnight at 4°C. A brief spin collected liquid from the inside of the cap. The tubes were placed on a magnetic stand to pellet magnetic beads on the tube side. For handling reasons and to prevent the beads from drying out, only 4 to 8 samples were processed at a time. The supernatant was carefully removed, and the beads were washed with 700 µL ChIP Buffer 1.2, followed by two washing steps with 800 µL ChIP Buffer 2.3. After the final washing step, the remaining supernatant was removed without disturbing the beads. 50 µL of Elution Buffer AM2.2 was added to the beads. They were incubated for 15 min at RT on a rotating wheel. After a quick spin, 50 µL of the Reverse Cross-linking Buffer was added to the eluted chromatin and mixed by pipetting up and down. Tubes were placed back into the magnetic stand, allowing the beads to pellet to the sides of the tubes. The supernatant containing the chromatin was transferred to a new reaction tube. 88 µL ChIP buffer 2 and 2 µL 5 M NaCl were added to the 10 µL input DNA in an additional reaction tube. ChIP and Input DNA samples were incubated for 2.5 h at 65°C. After a quick spin, collecting all liquid from the side of the tube, 2 µL Proteinase K was added. Well-mixed samples were incubated at 37°C for 1 h. The Proteinase K Stop Solution was placed at RT to equilibrate during this incubation. After a quick spin of the tubes, 2 µL Proteinase K Stop Solution was added to each tube. Samples were kept at -80°C until the column-based purification of DNA (appendix A19). The immunoprecipitated DNA was subjected to qPCR analysis for several primer sets (*IL6*, *IL8*, *CCL20*, *TNF-α*, *actin*, *TSH2B*) (refer to appendix Table A-3). Because of the sample volume limitation, only one measurement was performed in N-ChIP-qPCR.

## 2.8 X-ChIP

The target histone modifications and research approach were reconsidered and revised in line with the FAANG community, maintaining the underlying scope of the research. Here, the focus was on the FAANG core assays, which are established in the scientific community, to contribute to the description of the functional elements of the bovine genome. The FAANG histone marks the mono - and trimethylation of lysine (K) 4, trimethylation and acetylation of lysine (K) 27 of histone H3, and the CCCTC-binding factor (CTCF) were included. The X-ChIP method followed by sequencing was established to investigate histone modifications in mastitis cell models genome-wide. FAANG core assays were established for two cell models with the cooperation partner Diagenode. The on-site equipment and assay optimization expertise was used. The protocol for Auto iDeal ChIP-seq kit for Transcription Factors (Diagenode) Version 1 (06.15) and the IP-Star® Compact automated system were used to perform the X-ChIP experiments. The pipetting robot-based approach reduced experimental inaccuracy compared to manual pipetting.

### 2.8.1 Fixation and Cell Harvest

The bovine cell models pbMEC and MAC-T, which were challenged with mock, *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub> for 3 h in three independent experiments, were the starting material for X-ChIP. DNA-protein cross-linking was performed by fixation of the cells directly in the culture plate with 10 v/v % formaldehyde in fixation buffer (1.1% final formaldehyde concentration) for 15 min at RT with gentle shaking. Fixation was stopped by adding 10 v/v % 1.25 M glycine to the cell culture medium, continuing

incubation for 5 min at RT with gentle shaking. Once the fixation was completed, the following steps were performed on ice. The supernatant was removed, and cells were washed once with 20 mL cold PBS, collected by scraping in ice-cold PBS, and transferred into a 15 mL reaction tube. Cells were pelleted by centrifugation at 1600 x g and 4 °C for 5 min, the supernatant was discarded, and samples were stored at -80 °C until further use.

### **2.8.2 Cell Lysis and Chromatin Shearing**

Chromatin of an estimated 1 million cells for the histone modifications and 3.6 million cells for CTCF was used for the ChIP protocol. Samples were thawed on ice, while all kit components for this step were also chilled on ice. The cooling system for the sonication apparatus was switched on to reach a constant 4°C. 1 mL of lysis buffer iL1b was added to 1 million cells, resuspended, and incubated at 4 °C for 20 min with gentle mixing on a rotating wheel. Cells were pelleted by centrifugation at 500 x g and 4°C for 5 min. The supernatant was discarded. Then, 600 µL of lysis buffer iL2 per 1 million cells was added to the cell pellet. The mix was resuspended and incubated at 4°C for 10 min with gentle mixing on a rotating wheel. Cells were pelleted by centrifugation again.

Meanwhile centrifugation, shearing buffer iS1b was completed by adding 1x PIC (200x). A cell concentration of 1.8 million cells per 100 µL iS1b was used for the MEC models. Complete shearing buffer was added to the cell pellet, chilled on ice for 10 min, and transferred into a 2 mL Dounce homogenizer. Douncing was performed with 15 strokes. The lysed cell suspension was aliquoted in 150 µL aliquots into Bioruptor® Microtubes (Diagenode). The chromatin was sheared using the Bioruptor sonication apparatus at 4 °C, initially testing 5 to 15 shearing cycles (30 s ON/30 s OFF). Eventually, ten sonication cycles were used for pbMEC and 12 cycles for MAC-T. After shearing, samples were kept on ice, shortly spun, and contents of Bioruptor® Microtubes were transferred to new 1.5 mL reaction tubes and centrifuged at 16000 x g and 4 °C for 10 min. Supernatants containing the sheared chromatin were pooled per condition. An aliquot of 50 µL was set aside for chromatin shearing assessment. The remaining chromatin was stored at -80 °C.

For chromatin assessment, the samples were treated with Ambion RNase (Thermo Fisher Scientific) for 1 h at 37°C before adding 50 µL of elution buffer iE1 and 4 µL of iE2 to the samples. The mix was incubated at 65 °C for 4 h (or overnight) for reverse cross-linking. DNA was purified with QIAquick PCR Purification Kit (Qiagen), quantified with Qubit dsDNA Broad Range Assay (Thermo Fisher Scientific), and quality of shearing was assessed on the Bioanalyzer (DNA 1000 Assay, Agilent) or Fragment Analyzer (DNF 373 High Sensitivity NGS Fragment Analysis Kit, Advanced Analytical Technologies, Inc.).

### **2.8.3 Magnetic Immunoprecipitation**

In this study, ChIP grade antibodies (Diagenode) against four different histone modifications (H3K4me3, H3K27me3, H3K27Ac, H3K4me1) and the transcription factor CTCF were included in this study (Table 2-22). IgG was used as a negative control. Two amounts of antibodies were tested before the actual experiments. On the IP-Star® Compact automated system, the direct method protocol

ChIP\_IPure\_16\_200\_D was chosen with the setup: 3 h Ab coating, 12 h IP reaction, and 5 min for washes. The instrument had accessories (container, tips, tube strips). ChIP buffer, Ab coating mix, and IP mix were prepared, and appropriate volumes were added to 200  $\mu$ L tube strips. ChIP buffer contained 60  $\mu$ L 5x ChIP buffer iC1b, 234  $\mu$ L nuclease-free H<sub>2</sub>O, and 6  $\mu$ L 5% BSA (DNA free) per IP. Ab coating mix contained an equivalent of the amount of antibody tested or used (Table 2-22, Table 2-23), 0.5  $\mu$ L PIC, and was filled up to 100  $\mu$ L with 1x ChIP buffer iC1b + BSA. IP mix contained the appropriate amount of chromatin (chromatin from 3.6 million cells for CTCF and 1 million for histone modifications), 1  $\mu$ L PIC, 4  $\mu$ L 5% BSA, and was filled to 200  $\mu$ L with shearing buffer iS1b. 1% input DNA for each condition and antibody used was set aside for later use. 10  $\mu$ L of homogenously mixed DiaMag protein A-coated magnetic beads were added into another 200  $\mu$ L tube strip and placed in the instrument (according to screen instruction). The reagent racks were filled with buffer (iW1,2,3,4 and iE1) according to the manufacturer's instructions. All settings and buffers were double-checked before the IP proceeded entirely automatically.

*Table 2-22: Antibodies used for X-ChIP experiment. All listed antibodies (Diagenode) against histone modifications (Diagenode) were polyclonal antibodies raised in rabbits. Two quantities of ChIP-grade antibodies were tested initially.*

Antibody	Reference#	Lot#	Tested Antibody Amount ( $\mu$ g)
CTCF	C15410210	A235900234P	1, 2
H3K27Ac	C15410196	A17230041D	0.5, 1
H3K4me1	C15410194	A1862D	0.5, 1
H3K27me3	C15410195	A1811-001P	0.5, 1
H3K4me3	C15410003	A1052D	0.5, 1
IgG	C15410206		1

*Table 2-23: Conditions used for X-ChIP experiment.*

Cell type:	Shearing cycles	ChIP Antibody (Ab)	Ab Amount ( $\mu$ g)
pbMEC	10 x	H3K4me3	0.5
MAC-T	12 x		1
pbMEC	10 x	H3K27me3	0.5
MAC-T	12 x		
MAC-T	12 x	H3K27Ac	0.5
MAC-T	12 x	H3K4me1	1

#### 2.8.4 Elution and Decross-linking

After the overnight IP run, the eluted samples (~100  $\mu$ L) were equilibrated for 10 min at RT and placed on the DiaMag 0.2 mL magnetic rack, waiting 1 min to let the beads settle. The supernatant was transferred into new labeled tube strips. Beads were discarded. Input samples were filled up to 96  $\mu$ L with iE1. 4  $\mu$ L of elution buffer iE2 was added to all samples and input before tubes were incubated at 65 °C for 4 h decross-linking.

#### 2.8.5 DNA Purification

After the reverse cross-link, DNA was purified using the IPure-seq protocol of the IP-Star® for elution in 25  $\mu$ L. All required consumables and buffers (wash buffer 1, 2, and buffer C) were added to the appropriate positions and containers in the instrument. 2  $\mu$ L of Carrier was added to all IP and input samples. 10  $\mu$ L of IPure magnetic beads for each sample and input were loaded on a 96-well plate. Then, the run proceeded. After that, samples in tube strips were placed in the DiaMag 0.2 mL magnetic rack. After 1 min waiting time, the supernatant containing the immunoprecipitated DNA was transferred into new tubes. DNA was quantified using dsDNA High Sensitivity Assay. All samples were diluted 1/20



with ChIP-Seq grade water in 200 µL strip tubes and used for qPCR analyses afterward. All remaining samples, diluted and undiluted, were frozen at -20 °C until further use. Based on qPCR measurement, the % input and fold enrichment were calculated. In mock-challenged cells, the positive and negative genomic regions were established.

## **2.9 Library Preparation and Next-Generation Sequencing**

### **2.9.1 Library Preparation for RNA-seq**

Simone Wöhl performed library preparation at the Leibniz Institute for Farm Animal Biology. Starting material for library preparation for RNA-seq was 1 µg RNA, which originated from pbMEC and MAC-T, incubated for 3 h with *E. coli*<sub>1303</sub>, *S. aureus*<sub>1027</sub>, or mock, in three independent experiments with subsequent RNA isolation with RNeasy Plus Mini Kit and RNeasy® MinElute® Cleanup (both Qiagen). TruSeq RNA sample preparation kit set A (Illumina) version Version 2015\_06\_23 was used to prepare the libraries for sequencing from total RNA samples. A polyA tail was used at the very beginning of the protocol to select processed non-ribosomal RNA. Briefly, the working steps for library preparation included purification and fragmentation of RNA, first and second cDNA synthesis, adapter ligation, and PCR amplification. The quality of the libraries was assessed by 2100 Bioanalyzer (DNA 1000 Kit, Agilent) before use for sequencing. The libraries were subjected to paired-end mRNA sequencing on the HiSeq 2500 Sequencing System (Illumina) present at the Institute for Genome Biology at the Leibniz Institute for Farm Animal Biology, Dummerstorf, Germany.

### **2.9.2 Library Preparation for ChIP-seq**

The service provider Diagenode prepared the libraries for the 48 ChIP-seq samples of MAC-T and pbMEC after immunoprecipitation of the FAANG core assays and input DNA according to the protocols of MicroPlex library preparation Kit v2 (Diagenode) with 1 ng of DNA. Amplification was performed with 10 - 15 cycles, followed by quality verification with Fragment Analyzer (Advanced Analytical Technologies, Inc). Agencourt® AMPure beads (Beckman Coulter) were used for the 300 bp size selection of DNA fragments. The libraries' concentrations were quantified using the Qubit ds DNA HS kit and analyzed on the Fragment Analyzer to assess their insert sizes. The molar concentration of each library was calculated using the quantification values and the size measurement. Then, the different libraries were diluted to reach a concentration of 5 nM each. After the dilution, the libraries were separated into eight different pools. ChIP-seq libraries were sequenced, generating 50 bp single reads on an Illumina HiSeq 4000.

## **2.10 Computational Methods**

### **2.10.1 RNA-seq Data Analysis Pipeline**

Analyses were performed by a series of commands in a Linux-based and R environment (pipeline shown in Figure 2-5) with the help of Prof. Kühn. Raw data were quality controlled<sup>181,182</sup>, trimmed with Cutadapt (version 2.9 with Python 3.6.9)<sup>183</sup> and QualityTrim (version 1.6.0)<sup>184</sup>, and aligned to bovine genome (ARS\_UCD 1.2, Ensembl 95) with HISAT2 (version 2.2.0)<sup>185</sup>. Bam files were indexed with samtools (version 1.10)<sup>186</sup>, and reads were counted across samples with featureCounts (version 2.0.0)<sup>187</sup>.

Data were subjected to DeSeq2 (version 1.26.0)<sup>188</sup> for differential expression analyses. Deseq2 was used to normalize the count value by the parameter *FPKM*: fragments per kilobase per million mapped fragments. Thereafter, I performed the analysis. The principal component analysis (PCA) was performed with the samples' normalized count matrix to test the data's dimensions of differences and similarities. EDAs<sup>189</sup> (version 2.22.0), an exploratory data analysis tool, combined with ggscatter from ggpubr<sup>190</sup> (version 0.4.0) package in R, was used to visualize better a classical multidimensional scaling of log2 transformed data. Statistically significant differentially expressed genes between the two groups were determined using negative binomial modeling in DeSeq2. The data files were subsequently filtered for statistical significance by an adjusted p-value  $\leq 0.05$ . Optionally a log2 fold change ( $\log_2FC$ )  $< |1|$  or more stringent  $\log_2FC < |2|$  was applied. Data visualization and functional analysis, including visualization, were performed independently, testing different tools David<sup>191,192</sup>, Panther<sup>193</sup>, g:Profiler<sup>194</sup>, and packages ClusterProfiler<sup>195</sup>, pathview<sup>196</sup>, and enrichplot<sup>197</sup> in R<sup>198</sup>. The volcano plot showing significant differentially expressed genes was produced by plotting  $\log_2\text{FoldChange}$  versus negative  $\log_{10}(\text{padj})$ . Heatmaps were produced using the pretty heatmaps package pheatmap<sup>199</sup> in R. The functional analysis employed the Benjamini-Hochberg method using the corrected p-values ( $p_{\text{adjust}} = \text{FDR}$ )  $< 0.05$  for significance filtering. Functional analysis ClusterProfiler<sup>195</sup> allowed visualization of overrepresented terms in dot plots, significantly enriched terms in enrichment maps, and cnet plots showing involved genes linked to enriched terms. In each plot, the number of shown categories was chosen between 5, 10, 20, and 50 terms.

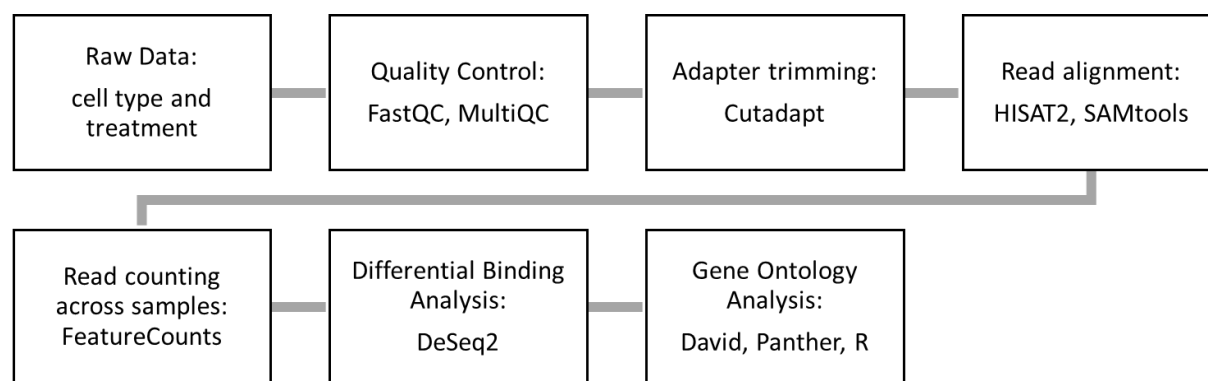


Figure 2-5: Data analysis pipeline for RNA-seq.

### 2.10.2 ChIP-seq Data Analysis Pipeline

The data analyses were initially performed by the service provider Diagenode on a payment basis. With the release of the new bovine reference genome assembly, the data analysis was reproduced and adapted according to the pipeline outlined in Figure 2-6 with support of Frieder Hadlich and Prof. Kühn. The sequencing quality of the raw data was controlled with FASTQC<sup>181</sup>. Reads were trimmed (trimmomatic, version 0.39)<sup>200</sup> and controlled for quality. Reads were aligned with Burrows-Wheeler Aligner (BWA)<sup>201,202</sup> (version 0.7.17-r1198-dirty) to the reference genome ARS-UCD1.2 (Ensembl release 95) obtained from the Ensembl genome browser<sup>203</sup>. Only uniquely mapping reads with no PCR duplicates were included (Genome Analysis Toolkit (GATK) 3.8-1-0-gf15c1c3ef<sup>204</sup>, PICARD 2.18.2<sup>205</sup>). I also

applied the quality analyses deepTools (version 3.1.0)<sup>206</sup> using the bin option of multiBamsummary for the input data in correlation, PCA, and heatmap analyses. Peak calling was performed for histone modifications with EPIC2 (at least version 0.033)<sup>207</sup>. The settings for peak calling for the histone modifications are shown in Table 2-24. Processing of ChIP-seq data was performed with the R/Bioconductor package: DiffBind<sup>208,209</sup>. Differential binding sites were identified with the DESeq2<sup>188</sup> method in DiffBind<sup>208,209</sup>. The annotation of differential binding sites was realized with HOMER (version 4.10)<sup>210</sup> with the script annotatePeaks.pl using the Ensembl genome file ARS-UCD1.2 (Ensembl release 95 Bos\_taurus.ARS-UCD1.2.95.dna.toplevel.fa) and the gtf annotation file (Bos\_taurus.ARS-UCD1.2.95.gtf). Data integration of RNA-seq and ChIP-seq was performed with the online tool to draw custom Venn diagrams<sup>211</sup>. Functional annotation and visualization were performed with g:Profiler (version e106\_eg53\_p16\_5d9a51b)<sup>194</sup>.

Table 2-24: Parameters for Peak calling. EPIC2 was used for peak calling of the histone modifications.

SICER	Redundancy threshold	Window size (bp)	Fragment size	Effective genome fraction	Gap size (bp)	FDR
H3K4me3	1	200	300	0.8	600	0.05
H3K27me3	1	5000	300	0.8	50000	0.1
H3K4me1	1	200	300	0.8	600	0.05
H3K27ac	1	200	300	0.8	600	0.05

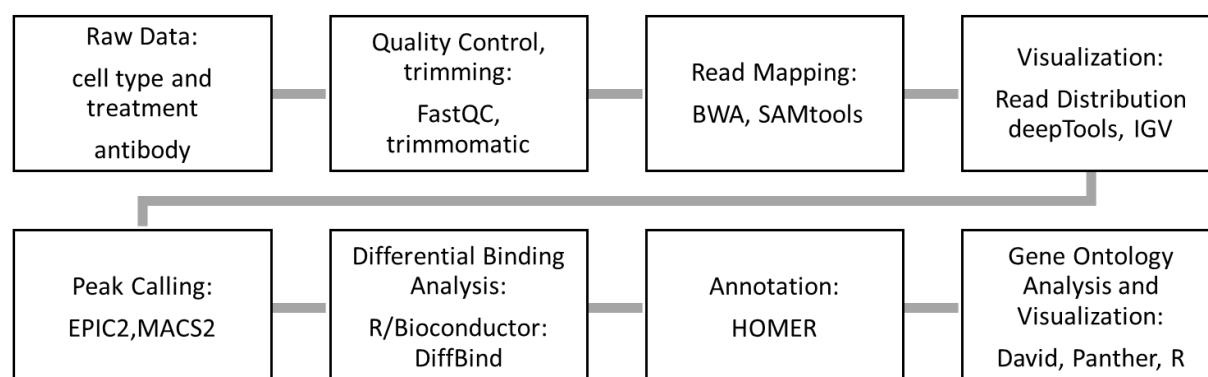


Figure 2-6: Data analysis pipeline for ChIP-seq.

### 2.10.3 R-based Analysis and Visualization

Based on the measured Ct from qPCR, absolute and relative quantification of qPCR data was performed in Microsoft Office Excel. Statistical analysis during data screening included pairwise t-test, ANOVA, multiple comparisons with the Post Hoc Tukey test, Shapiro-Wilk test for normality of the data, Bartlett's test, and Levene's test for homogeneity of variance across groups. The statistical analyses were computed with R<sup>198</sup> in the R Studio environment<sup>212</sup>. Data visualization was performed in R<sup>198</sup> using ggpubr<sup>190</sup> and ggplot2<sup>213</sup>.

NGS data exploration and visualization were performed with R packages dplyr<sup>214</sup>, tidyverse<sup>215</sup>, tidyr<sup>216</sup>, data.table<sup>217</sup>, eulerr<sup>218</sup>, ggplot2<sup>213</sup>, patchwork<sup>219</sup> and cowplot<sup>220</sup>. Graphical visualization of functional overlap between histone modifications was also performed with a free online tool to calculate and draw custom Venn diagrams<sup>211</sup>.

### 3 Results

#### 3.1 Histone Modification in MAC-T in Response to Mastitis Pathogens

##### 3.1.1 Expression of Immune Response Genes in Response to *E. coli*

The initial experiments focused on using the immortalized cell line MAC-T. Expression of immune response genes after 0, 1, and 3 h of pathogen contact was measured in biological replicates ( $n=3$ ) to confirm that the cell model recognized pathogens. The immune response of MAC-T challenged with the bacterium *E. coli*<sub>1303</sub> was determined by isolating RNA at each time point, reverse transcribing RNA into cDNA, and measuring gene expression by qPCR for *CCL20*, *TNF- $\alpha$* , *IL6*, *IL8*, *iNOS*, *JMJD3*, and *LAP*. Standard counting plasmids for these listed genes were successfully cloned and linearized by enzymatic digestion for use in qPCR experiments. Serial dilutions of these standard counting plasmids were a reference for calculating mRNA copy numbers of unknown templates from test samples.

The experimental data revealed that pathogens upregulated the immune response genes in MAC-T. *CCL20*, *IL6*, *IL8*, *iNOS*, and *TNF- $\alpha$*  showed significant differences at 1 or 3 h *E. coli*<sub>1303</sub>-challenge relative to control or within the course of pathogen-challenge (Figure 3-1, t-test,  $p \leq 0.05$ ).

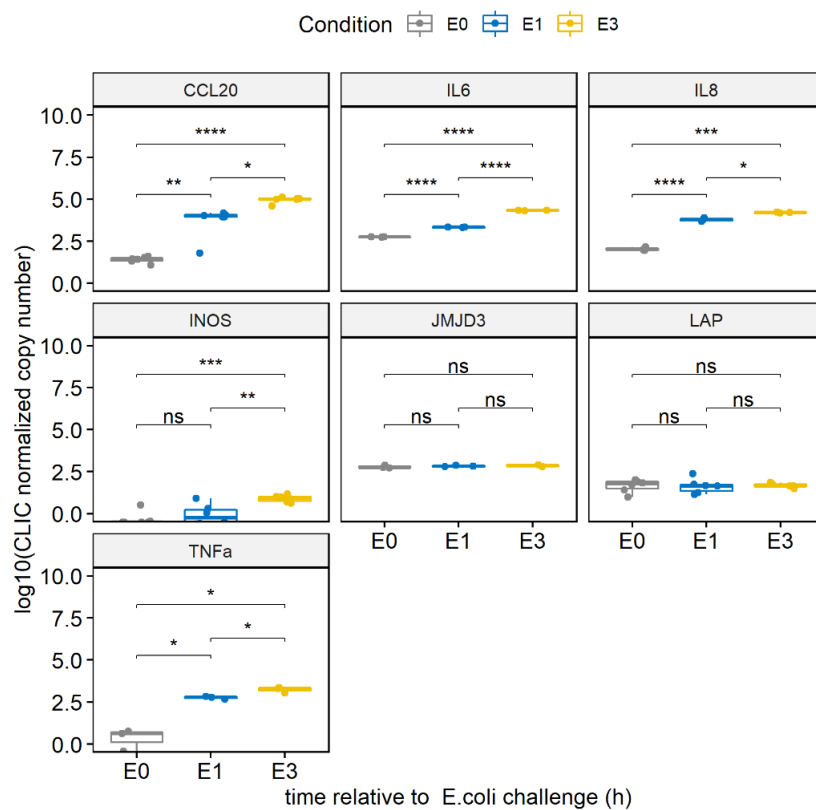


Figure 3-1: Expression level of immune genes after challenge with *E. coli*<sub>1303</sub> in MAC-T. Cells were exposed to the pathogen for 0, 1, and 3 h, shown as E0, E1, and E3 on the x-axis. Two cytokines of the "immediate early" response (*IL8*, *TNF- $\alpha$* ), the "late early" cytokine *IL6*, and three factors with antimicrobial functions (*iNOS*, *LAP*, and *CCL20*), and *JMJD3*, which encodes a lysine-specific demethylase were measured with RT-qPCR thrice in duplicates ( $n=6$ ). The copy number was estimated based on external standards counting plasmids of the same gene target. Samples were additionally normalized to a control gene (*CLIC1*). Statistical analysis was performed in R using pairwise comparisons in t-test and indicating significance with ns  $> 0.05$ , \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ , and \*\*\*\*  $p \leq 0.0001$ . The y-scale of the plots is log<sub>10</sub> transformed. Plots were generated with the R package ggpubr.

In addition to the pairwise comparisons in the t-test (Figure 3-1), the Shapiro-Wilk test for normality was performed for one variable (E0, E1, or E3) per target gene. *IL8*, *IL6*, *iNOS* (E3), *JMJD3*, *CCL20*, *TNF- $\alpha$* , and *LAP* passed it. The test for normality was rejected only for *iNOS* (E0 and E1). Using Bartlett's test with one independent variable, the sample groups were tested for homogeneity of variance across groups. The sample groups showed a p-value smaller than 0.05 for *IL8*, *IL6*, *CCL20*, and *TNF- $\alpha$* , while p-values larger than 0.05 were determined for *JMJD3*, *LAP*, and *iNOS*. P-values above the threshold in Bartlett's test indicate outliers in the samples, a limitation by the number of replicates, or a measurement error, resulting in a more substantial inconsistency with normality. Multiple comparisons of means for significance with the Post Hoc Tukey test revealed significance for all target genes in either one, two, or all comparisons except for *JMJD3* and *LAP* (data not shown). These findings match pairwise comparisons with t-tests. The measurements for *iNOS* showed the lowest abundance in copy numbers, while *JMJD3* and *LAP* revealed no measurable effect of the pathogen challenge. Another approach to show and interpret the results is to plot the fold enrichment of time points 1 or 3 h versus 0 h (appendix Figure B-1). This plot revealed the same trends of immune regulation of *CCL20*, *IL6*, *IL8*, *iNOS*, and *TNF- $\alpha$* .

### 3.1.2 Divergent Expression of Immune Response Genes in Response to *E. coli* and *S. aureus*

An additional experiment focused on the expression of immune response genes when comparing challenges of MAC-T with either of the mastitis pathogens *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub>. *INOS*, which had previously shown below ten copies, was excluded from the measurement program. Data were collected in two replications of the experiment and measured at least in duplicates with qPCR (n=5 to 6). Results for *CCL20*, *IL6*, and *TNF- $\alpha$*  for *E. coli*<sub>1303</sub>-challenged cells were reproduced (see Figure 3-2), although partly less significant when compared with the results above (Figure 3-1).

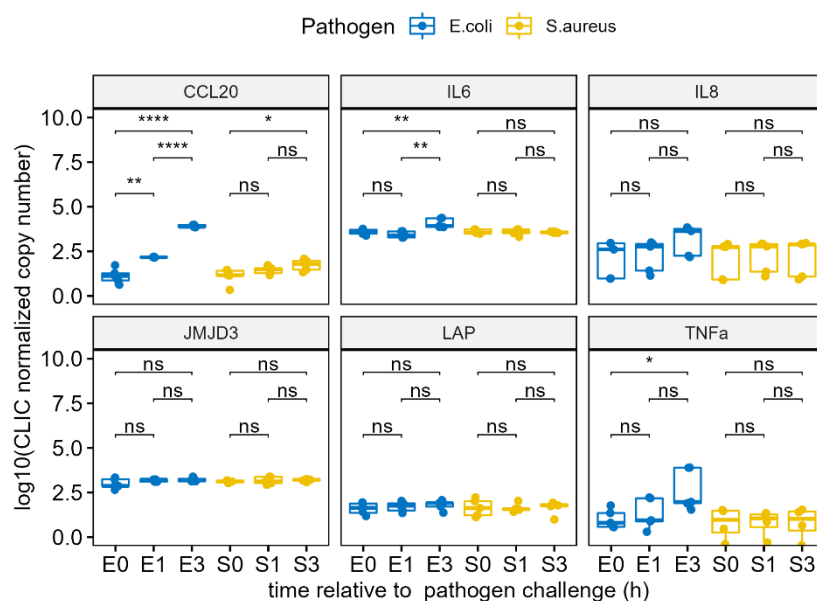


Figure 3-2: Expression level of immune genes after challenge with *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub> in MAC-T. Cells were exposed to the pathogen for 0, 1, and 3 h; RNA was used for RT-qPCR measuring two cytokines of the "immediate early" response (*IL8*, *TNF- $\alpha$* ), the "late early" cytokine *IL6* and factors with antimicrobial functions (*LAP*, and

*CCL20*), and *JMJD3*, which encodes a lysine-specific demethylase ( $n=5$  to  $6$ ). The copy number was estimated based on standard counting plasmids of the same gene target. Samples were additionally normalized to a control gene (*CLIC1*). Statistical analysis was performed in R using pairwise comparisons in t-test and indicating significance with  $ns > 0.05$ ,  $*p \leq 0.05$ ,  $**p \leq 0.01$ ,  $***p \leq 0.001$ , and  $****p \leq 0.0001$ . The y-scale of the plots is  $\log_{10}$  transformed. Plots were generated with the R package ggpubr.

This experiment observed no significant effect for *IL8* after challenge with *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub>. Moreover, the results provided evidence that *E. coli* upregulated *IL6* and *IL8* to a lesser extent when the copy numbers of the genes were compared (Figure 3-2 and Figure 3-1). This finding suggests limited replication of the experimental measurements.

When comparing the effect of both pathogens on MAC-T, *S. aureus*<sub>1027</sub>-exposed cells did show a lower expression for all measured genes than *E. coli*<sub>1303</sub>-exposed cells (Figure 3-2). *CCL20* was the only gene significantly upregulated comparing the 3 h- *S. aureus*<sub>1027</sub>-challenge (S3) to the control (S0). A pairwise comparison with a t-test confirmed these findings ( $p \leq 0.05$ ). The replicates were largely heterogeneous, as suggested by Bartlett's test, and visible in the boxplots, for example, for *TNF- $\alpha$*  and *IL8* (Figure 3-2). Analog to the above, the fold enrichment was plotted (see appendix Figure B-2).

Evidence of immune gene regulation in response to mastitis pathogen *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub> was a prerequisite to determining a possible epigenetic effect, as determined by ChIP.

### 3.1.3 Optimization of N-ChIP Procedure in MAC-T

N-ChIP was used to test the applicability of the epigenetic profiling method for specific histone marks in MAC-T. The lysis of cells harvested in PBS without cross-linking was monitored microscopically. Mechanical membrane disruption was achieved after 15 strokes in a Dounce homogenizer, as exemplified in Figure 3-3A. A time series of the effects of micrococcal nuclease on DNA fragmentation was tested and verified by gel electrophoresis. Figure 3-3B shows a preserved DNA ladder after 5 min. The fragmentation is strongest after 10- and 15-min enzyme action. The small DNA fragments indicated the degradation of histone-bound DNA. Correctly fragmented DNA after short enzyme action was used for the IP to target histone modifications.

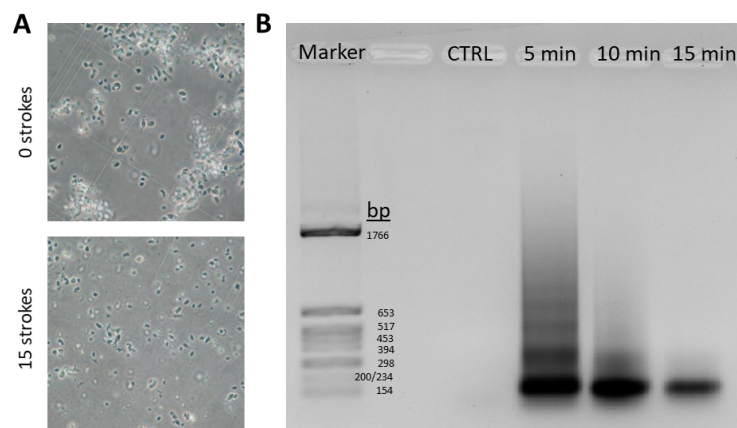


Figure 3-3: Optimization steps for N-ChIP in MAC-T. Lysis and shearing required testing to establish N-ChIP. The release of nuclei and disruption of cell sheets was checked microscopically (A). The enzyme was used for 5 to 15 min, and the digested DNA was verified on a 1% agarose gel (B).

### 3.1.4 *N-ChIP-qPCR in Promoter Regions of Immune Response Genes after Exposure to Mastitis Pathogens*

A commercially available kit from Active Motif and antibodies validated for the method according to the manufacturer's instructions were used for N-ChIP. The N-ChIP antibody selection included an active histone mark (H3K14Ac), a repressive histone mark (H3K9me2), and IgG and histone H3 controls. Downstream of the IP, a DNA clean-up and qPCR followed. The optimal primer melting temperature of all primer sets designed for the promoter region of immune response genes was determined by gradient PCRs. Standard counting plasmids were prepared, verified by sequencing, and used for qPCR measurements. They allowed the estimation of copy numbers per sample and measurement. NOS2A and LAP standard counting plasmids for N-ChIP and the designed primer sets caused many difficulties and were therefore excluded from the measurement program for N-ChIP. The immunoprecipitated DNA from N-ChIP in MAC-T was used in precise 20 ng/5  $\mu$ L for qPCR. Given the variable yield of cleaned N-ChIP DNA, partially fewer primer sets were tested.

One experiment used immunoprecipitated DNA from MAC-T incubated with *E. coli*<sub>1303</sub> for 0, 1, or 3 h for qPCR measurements in various promoter regions of selected immune genes (Figure 3-4). The measurements revealed that the IgG negative control representing the background noise was as strong as the determined signal for histone H3 positive control and input DNA (Figure 3-4), suggesting antibody or measurement issues. These findings affected the analysis. Thus, the conventional calculations of fold enrichment, which divides the histone signal by the background, were of limited applicability here. Instead, the results are shown as actin normalized copy numbers.

Another experiment used immunoprecipitated DNA from MAC-T challenged with *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub> for 1 h (Figure 3-5). Here, the IgG antibody supplier was changed, while the remaining antibodies were the same as in the previous N-ChIP experiment. When visually inspected, the signals for histone modifications differ from the measured background signal (Figure 3-5). *TNF- $\alpha$*  regions showed the highest measured copy numbers in all measurements compared with regions of other immune target genes. Notably, a decrease in copy number for all genes studied after 1 h-pathogen-exposure for the repressive ChIP mark H3K9me2 matches the increase in expression of the immune genes studied (Figure 3-4, Figure 3-5).

The RT-qPCR measurements for the expression of immune genes in MAC-T (sections 3.1.1 and 3.1.2) were performed multiple times and allowed statistical analyses. In contrast, measurements for N-ChIP-qPCR were performed once and were limited by the total sample size, resulting in some gaps in the measurements of the 23 primer sets.

Despite its exploratory nature, this study offers insight into the N-ChIP results after bacterial challenge. The measured signal was different from the control-challenged cells. The method was established, and different starting materials and antibodies were tested. It was found that epigenetics may play a role in selected target regions of immune genes. The indication that histone modifications can be determined in MAC-T made it worthwhile to turn to whole-genome analysis. Genome-wide gene expression and

histone patterns were investigated to determine epigenetic transcriptional regulation in targeted immune genes and beyond.

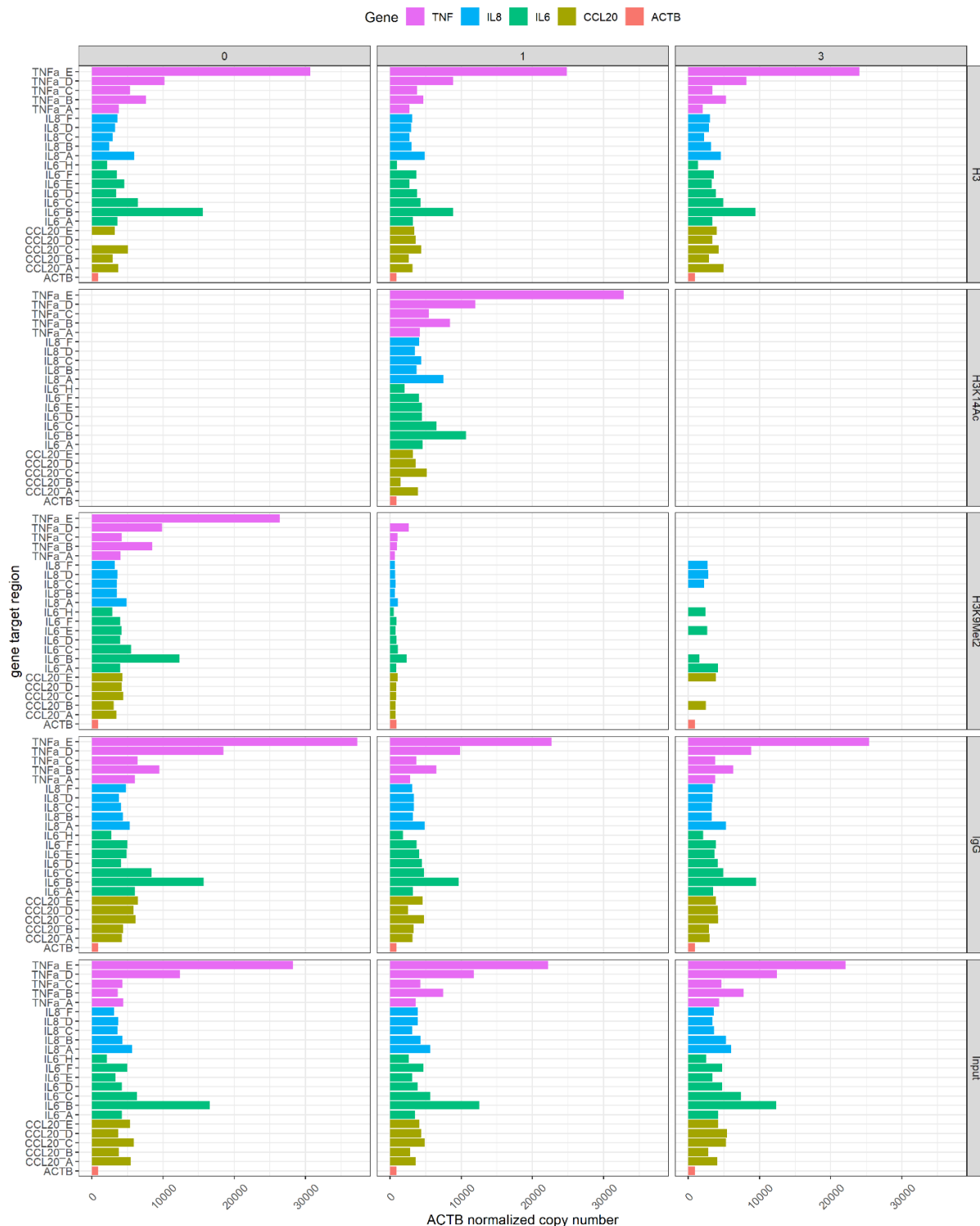


Figure 3-4: N-ChIP-qPCR in MAC-T after *E. coli*<sub>1303</sub> contact. Chromatin isolated from MAC-T incubated with *E. coli*<sub>1303</sub> for 0, 1, or 3 h was immunoprecipitated with H3, H3K14Ac, H3K9Me2, and IgG antibodies. DNA from IPs and input DNA was purified, and the enrichment of histone marks in promoter regions of immune response genes (TNF, IL8, IL6, CCL20) or control gene actin (ACTB) was checked with qPCR. Besides the active and repressive histone marks, H3K14Ac and H3K9Me2, the measurement included global enrichment of H3 and negative control antibody IgG.



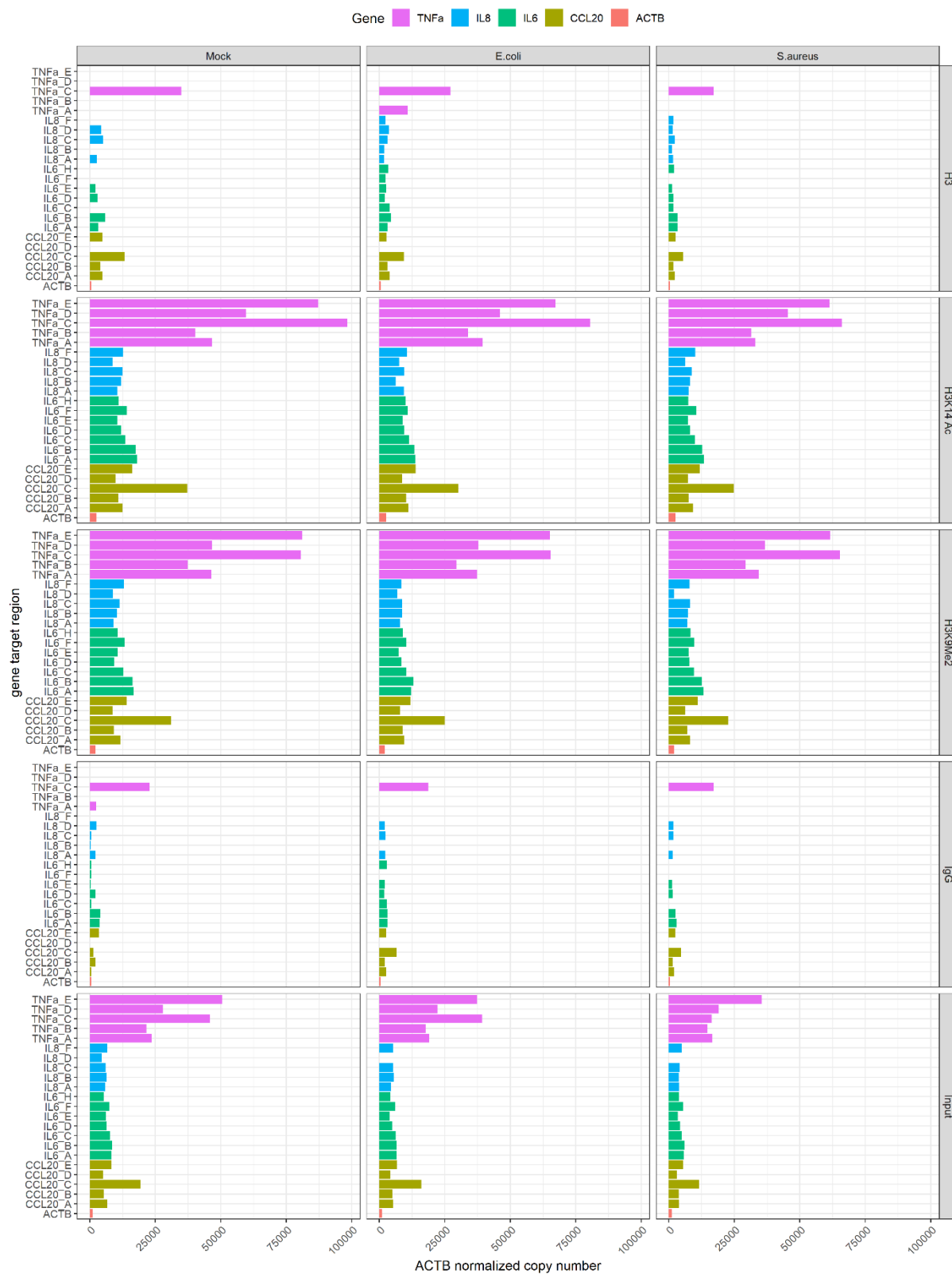


Figure 3-5: N-ChIP-qPCR in MAC-T after pathogen challenge. Chromatin isolated from MAC-T incubated with *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub> for 1 h was immunoprecipitated with H3, H3K14Ac, H3K9Me2, and IgG antibodies. DNA from IPs and input DNA was purified, and the enrichment of histone marks in promoter regions of immune response genes (TNF, IL8, IL6, CCL20) or control gene actin (ACTB) was checked by qPCR.

### 3.2 Genome-wide Transcriptome Sequencing in Bovine Mastitis Cell Models

While the qPCR experiments targeted specific immune genes, genome-wide gene expression of the two cell models, MAC-T and pbMEC, was analyzed by RNA-seq after a 3 h challenge with control, *E. coli*<sub>1303</sub>, or *S. aureus*<sub>1027</sub>. The following subsections will present comparisons of the cell types and the effect of pathogen challenge in each cell type.

#### 3.2.1 Comparison of MAC-T and pbMEC

The goal was to investigate, compare, and describe the transcriptome of the immortalized cell line MAC-T and the primary cells pbMEC. RNA-seq generated, on average, 76.8 million reads per sample. All files were subjected to quality control and trimming. It can be seen from the data in appendix Table B-2 that, on average, 89.6 % of the reads passed quality control. Mapping, alignment, and normalization were performed analogously to previous bovine studies. On average, 97.4 % of the reads were mapped to the bovine reference genome (ARS-UCD1.2, Ensembl 95).

The initial RNA-seq analysis included a principal component analysis (PCA), which provided a dimensionality reduction of the comprehensive data. PCA was used to test how similar or different the RNA-seq data sets were. Figure 3-6 illustrates the relationship between the samples and replicates. Data are separated by cell type (MACT versus pbMEC), with the most considerable apparent variance being the first principal component (PC1), which explains ~83% of the total inter-group variability along the x-axis. The biological or technical replicates were critical for cluster resolution for the second principal component (PC2) along the y-axis, which explains ~8% of the data variance. The challenge *per se* did not affect the data pattern in the PCA. All conditions per replicate were close to each other. Primary cells exhibited greater intra-group variability than the immortalized cell line MAC-T, as evidenced by the larger scatter of the different cell samples.

Cell type was the most prominent variance direction in the previous PCA (Figure 3-6). Thus, cell types were analyzed by performing individual PCAs (Figure 3-7). For both cell types, biological replicates explained the main variance of the data along PC1 on the x-axis with ~61% and ~60% for MAC-T and pbMEC, respectively. PC2 explained ~21% and ~26% of the data variance for MAC-T and pbMEC, respectively. The bacterial challenge did not cluster in either cell type.

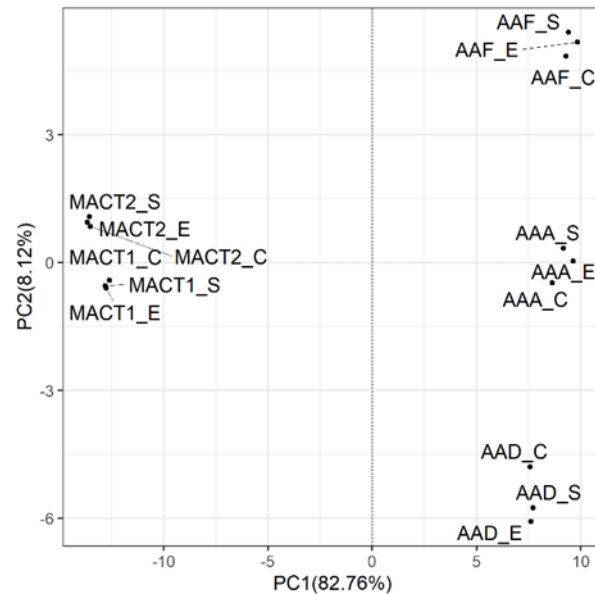


Figure 3-6: Principal component analysis for RNA-seq data in mammary epithelial cells. Two cell models, primary bovine mammary epithelial cells, pbMEC, and a cell line MAC-T, were exposed to *E. coli*<sub>1303</sub> (E), *S. aureus*<sub>1027</sub> (S) or media as a control (C) for 3 h. RNA was extracted from MAC-T in duplicates (MACT1, MACT2) and primary cells in triplicates (AAA, AAD, AAF) and used for RNA-seq. Data were aligned to the bovine reference genome, and differential expression analysis was performed with DESeq2.

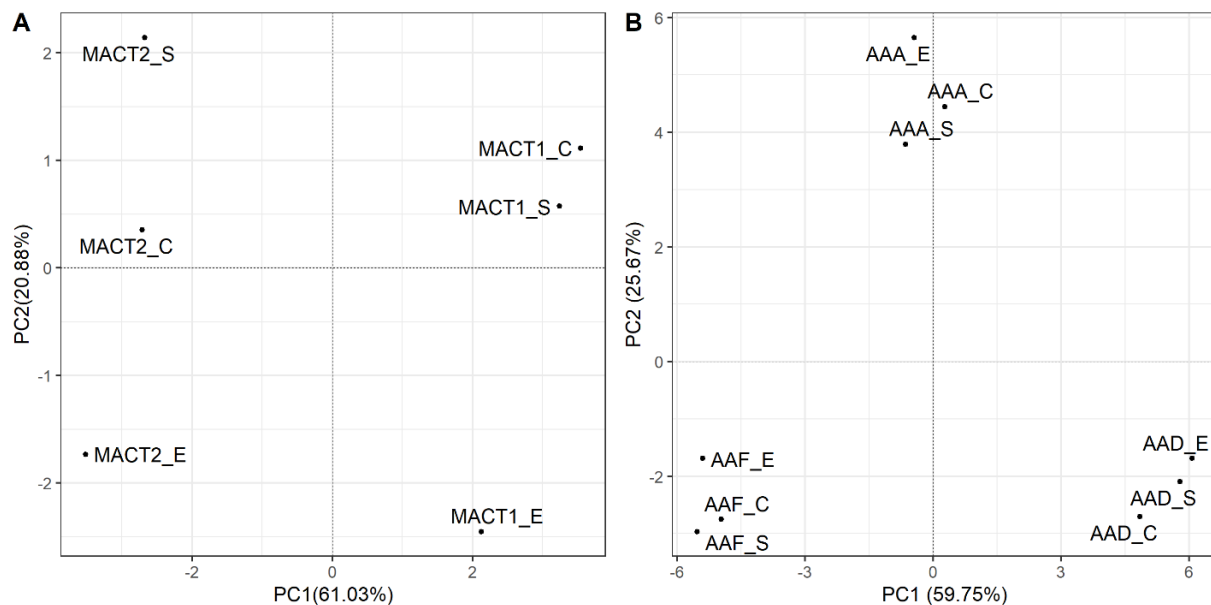


Figure 3-7: Principal component analysis for RNA-seq data in individual mammary epithelial cells. Data for the two cell models, cell line MAC-T (Diagram A) and primary bovine mammary epithelial cells pbMEC (Diagram B) challenged for 3 h with *E. coli*<sub>1303</sub> (E), *S. aureus*<sub>1027</sub> (S), or media as control (C), were plotted individually. MAC-T data were collected in duplicates (MACT1, MACT2), and data from primary cells in triplicates (AAA, AAD, AAF)—principal components (PC) on the x- and y-axis account for the main variance.

12781 gene loci were included in the analyses comparing the two cell models, filtered for significance with a  $p_{\text{adj}} < 0.05$  and a  $\log_2\text{FC} \geq |2|$ . A total of 1829 significant differentially expressed genes (DEGs) were identified in the comparison between control-challenged MAC-T and pbMEC. A volcano plot (Figure 3-8A) was used to illustrate the significant expression of all data points in the data set. All DEGs with  $p_{\text{adj}} < 0.05$  and a  $\log_2\text{FC} \geq 2$ , indicating that expression was at least fourfold higher in MAC-T

compared with pbMEC, were significantly upregulated in MAC-T and colored in red. In contrast, all DEGs with  $p_{adj} < 0.05$  and a  $\log_2FC \leq -2$ , indicating that expression was at least fourfold lower in MAC-T compared with pbMEC, were downregulated in MAC-T and colored blue (Figure 3-8A). The grey data points (in total 10952) on the volcano plot indicated no significance according to the above definition of  $p_{adj} < 0.05$  and a  $\log_2FC \geq 2$ . The results of significant gene expression levels between the cell types MAC-T and pbMEC in the control condition revealed more DEGs downregulated (in total 1464) than upregulated (in total 365), as depicted in the volcano plot (Figure 3-8A). Thus, the transcriptomic analysis implies a massive difference between the cell models.

A heatmap (Figure 3-8B) was used to visualize the hierarchical clustering of the data. The heatmap shows  $\log_2$ -transformed normalized expression values per cell type and challenge (column-wise) for all the significant genes (row-wise). Similar data are consequently clustered together. The column-wise dendrogram was color-coded in grey for controls (C), blue for *E. coli*<sub>1303</sub>(E), and yellow for *S. aureus*<sub>1027</sub> (S). As a result, the clusters indicate the grouping of technical replicates per cell type, not for the challenge, as depicted in the colored boxes. The color coding of the heatmap itself is consistent with the volcano plot with the upregulated genes in red and downregulated genes in blue. The pbMEC samples (AAA, AAD, and AAF) show noticeable variance.

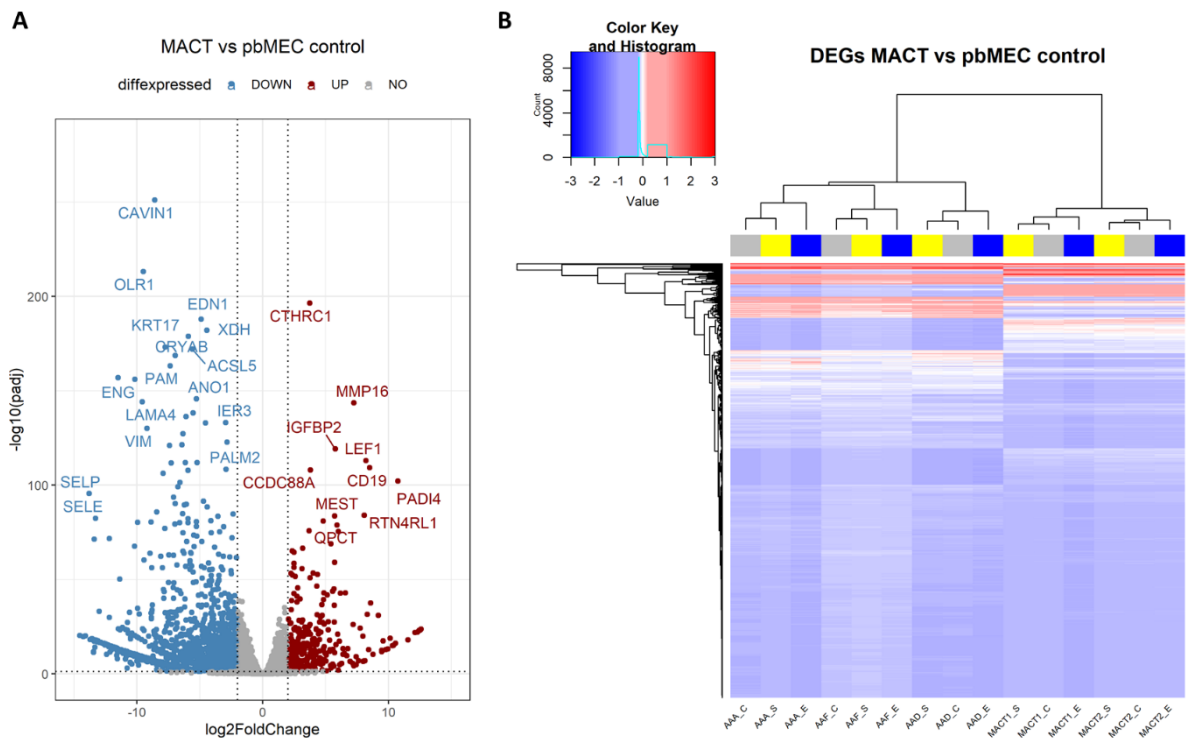


Figure 3-8: Visualization of differential gene expression. Expression data from MAC-T and pbMEC control-challenged were compared and visualized in a volcano plot (A) and heatmap (B). In A, the  $\log_2FC$  on the X-axis, representing the level of change between the two exposure conditions, was plotted against statistical significance in terms of  $p_{adj}$  on the Y-axis. The horizontal line marks  $p_{adj} = 0.05$ , whereas the vertical lines represent the cut-off of  $\log_2FC$  at -2 and 2. The blue dots represent significantly downregulated genes (in total 1464), while the red dots represent the significantly upregulated genes (in total 365) in the control-challenged MAC-T cells. The genes highlighted in gray are outside the cut-offs and considered not significant. The volcano plot was drawn with ggplot2 applying ggrepel in R. The heatmap (B) supports the volcano diagram based on single replicates for expressed genes in cattle. The column-wise dendrogram supports clustering according to the challenge, as seen by color coding grey for control (C), blue for *E. coli*<sub>1303</sub> (E), and yellow for *S. aureus*<sub>1027</sub>.

From the identified set of genes that were significantly expressed between MAC-T and pbMEC (1829 DEGs), the top 20 upregulated genes ( $\log_2FC \geq 2$ ) and the top 20 downregulated genes ( $\log_2FC \leq -2$ ) were selected based on sorting by  $p_{adj}$  to assess the effect of challenge on these genes. Mean values of the normalized expression counts (FPKM) of MAC-T and pbMEC replicates for the top 20 upregulated genes are shown in Figure 3-9A, complemented by the top 20 downregulated genes in Figure 3-9B. The dot plots show differing expression levels between cell types and how little the challenge with *E. coli*<sub>1303</sub> (blue dots) or *S. aureus*<sub>1027</sub> (yellow dots) affected these genes in comparison to the control (grey dots). Please note that the normalized counts on the y-axis are logged ( $\log_{10}$ ) to ensure that any large differences in expression are plotted without compromising the quality of the visualization. Considering the genes previously measured by qPCR, their FPKM values were extracted from the list of DEGs from RNA-seq and plotted in Figure 3-10. The genes CCL20, CXCL8, IL6, and NOS2 were visualized across samples and showed very weak upregulation compared to significant genes in Figure 3-9. These plots confirm a trend of pbMEC responding dissimilar from MAC-T.

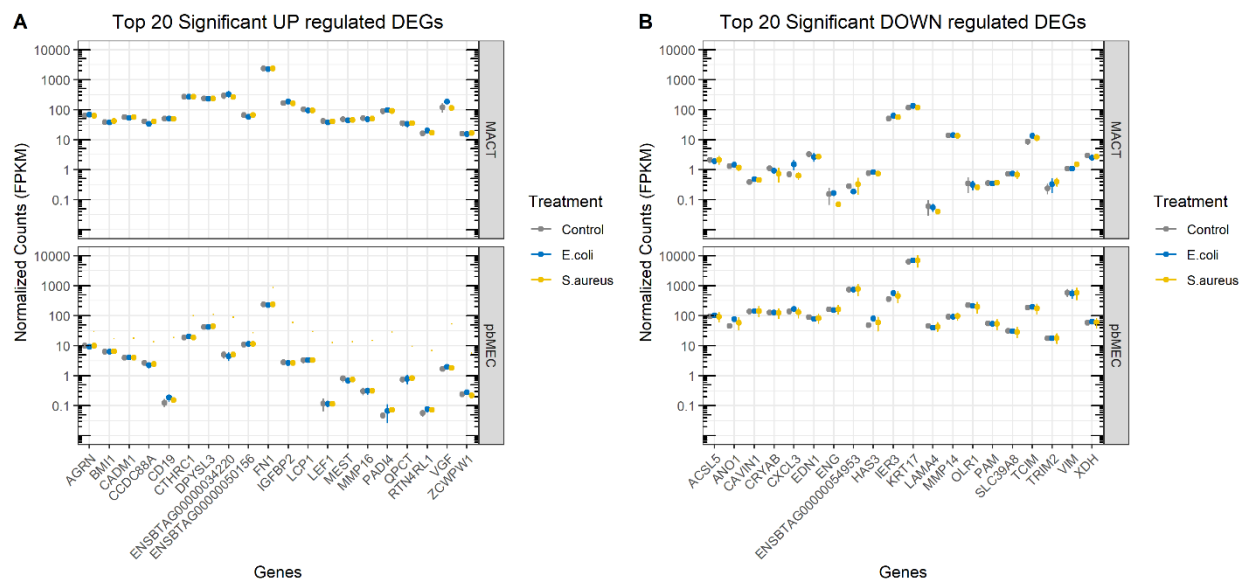


Figure 3-9: Expression of top regulated genes across sample groups. The top 20 up-and down-regulated genes were selected by their  $\log_2FC \geq 2$  or  $\leq -2$  and  $p_{adj} < 0.05$ . The dots represent the mean value of the replicates per cell type (upper panel MAC-T, bottom panel: pbMEC) and condition, color-coded for the control in grey, *E. coli*<sub>1303</sub> in blue, and *S. aureus*<sub>1027</sub> in yellow.

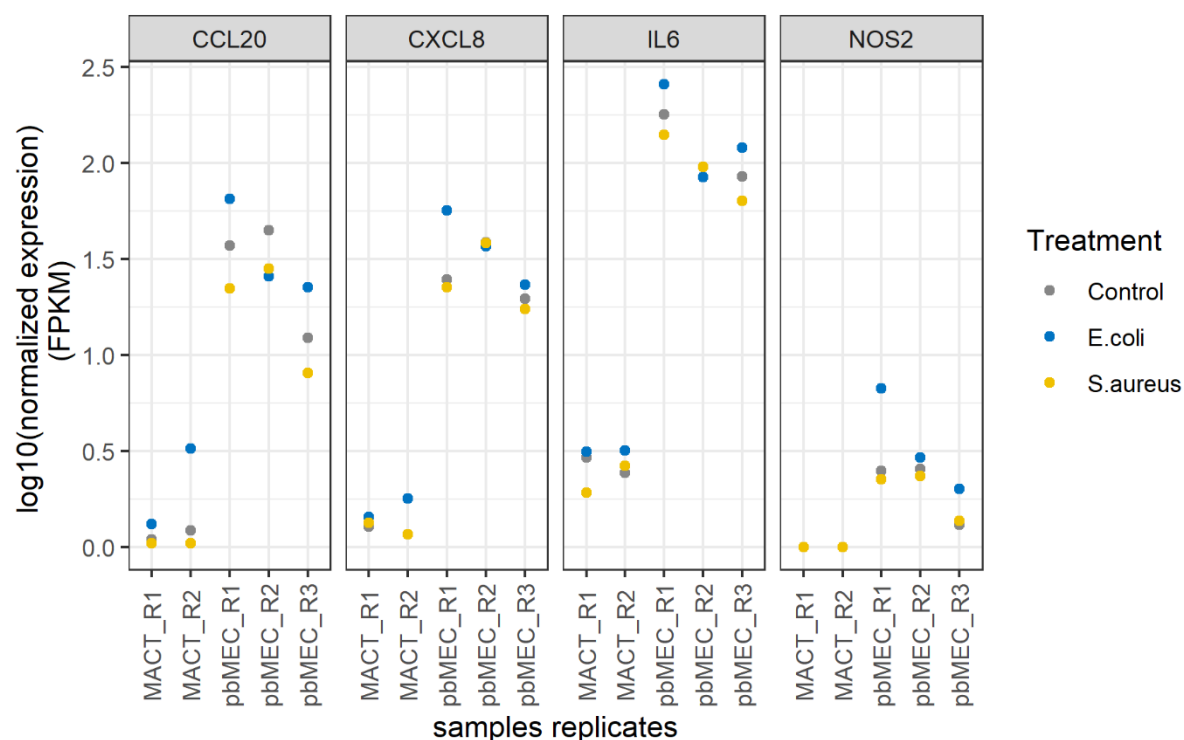


Figure 3-10: FPKM values for selected immune response genes from RNA-seq. Analogous to qPCR, results were presented for the two cell models for those genes found in the data.

The identified DEGs from RNA-seq comparing control-challenged MAC-T with pbMEC were subjected to functional annotation analysis using *enrichGO* method, a Gene Ontology (GO) Enrichment Analysis, in the *clusterProfiler* package in R. The analyses were used to determine the biological value of the gene sets by identifying significant overrepresented and enriched gene ontologies for biological processes (BP), molecular functions (MF), and cellular compartments (CC), and then visualizing the results graphically.

The dot plot (Figure 3-11) shows multiple features simultaneously. The determined top 50 GO terms of the BP are sorted in descending order by their gene ratio, which is the number of genes per term relative to the number of significant genes. The size of the dots represents the number of genes associated with the top 50 GO terms of the BP. The color scale of the dots represents the statistical significance in the form of  $p_{adj}$  for these terms. The top terms with the most genes associated with them included, among others, cellular developmental processes, cell differentiation, regulation of the multicellular organismal process, regulation of cell communication, and animal organ development.

The enrichment GO map (Figure 3-12) illustrates the relationship between the 50 most enriched GO terms ( $p_{adj}$ ) by clustering related terms. The linkage of terms is based on the number of genes present in each GO term. The color in the GO plot indicates the graduation of p-values relative to the other terms shown, with a lighter red dot having greater significance over the purple dots. The size of the dots associated with the terms represents the number of genes that are significant from the input list of DEGs coming from the comparison of gene expression of MAC-T and pbMEC in the control group. The relation of the terms condenses in four main node points, which can be described with superordinate

terms: biomineralization, differentiation and developmental processes, adhesion, and cellular response. Individual BP terms, such as G-Protein receptor coupled signaling pathway and collagen metabolic processes are in the periphery of the plot.

The category netplot (Figure 3-13) shows the linkages between genes associated with the five significant GO terms of BP: cell adhesion, biological adhesion, developmental process regulation, multicellular organismic process regulation, and system processes. Genes associated with these terms are color-coded based on the fold changes of the significant genes associated with them. Green dots are significantly down-regulated genes between control-challenged MAC-T and pbMEC. Red dots are significantly up-regulated genes between control-challenged MAC-T and pbMEC, of which there are fewer overall in the plot (Figure 3-13). In this display format, a fold change greater than 4 or less than -4 has been scaled to 4 or -4, respectively, to allow the readability of the results within the wide fold change range. Here, the size of the GO terms reflects the p-values of the terms, with the more significant terms being larger in size. Some DEGs are involved in several processes. *TGFB1*, *FNI*, and *NR1D1* are examples of upregulated DEGs, whereas *IL6*, *IL7*, *CXCL8*, and *NOD2* are examples of downregulated DEGs among several genes with multiple links to most GO terms of the BP. The latter genes, part of a list of genes significantly higher expressed in pbMEC, are predominantly related to immunity.

Moreover, most differentially expressed genes were associated with GO terms for cellular components such as membrane, junctions, the extracellular matrix, and more sparsely with filaments and cytoskeleton (appendix Figure B-3, Figure B-4, and Figure B-5). In addition, the identified differentially expressed genes are known to be involved in molecular functions like signaling from outside or the membrane into the cell, as depicted in appendix Figure B-8, Figure B-9, and Figure B-10.

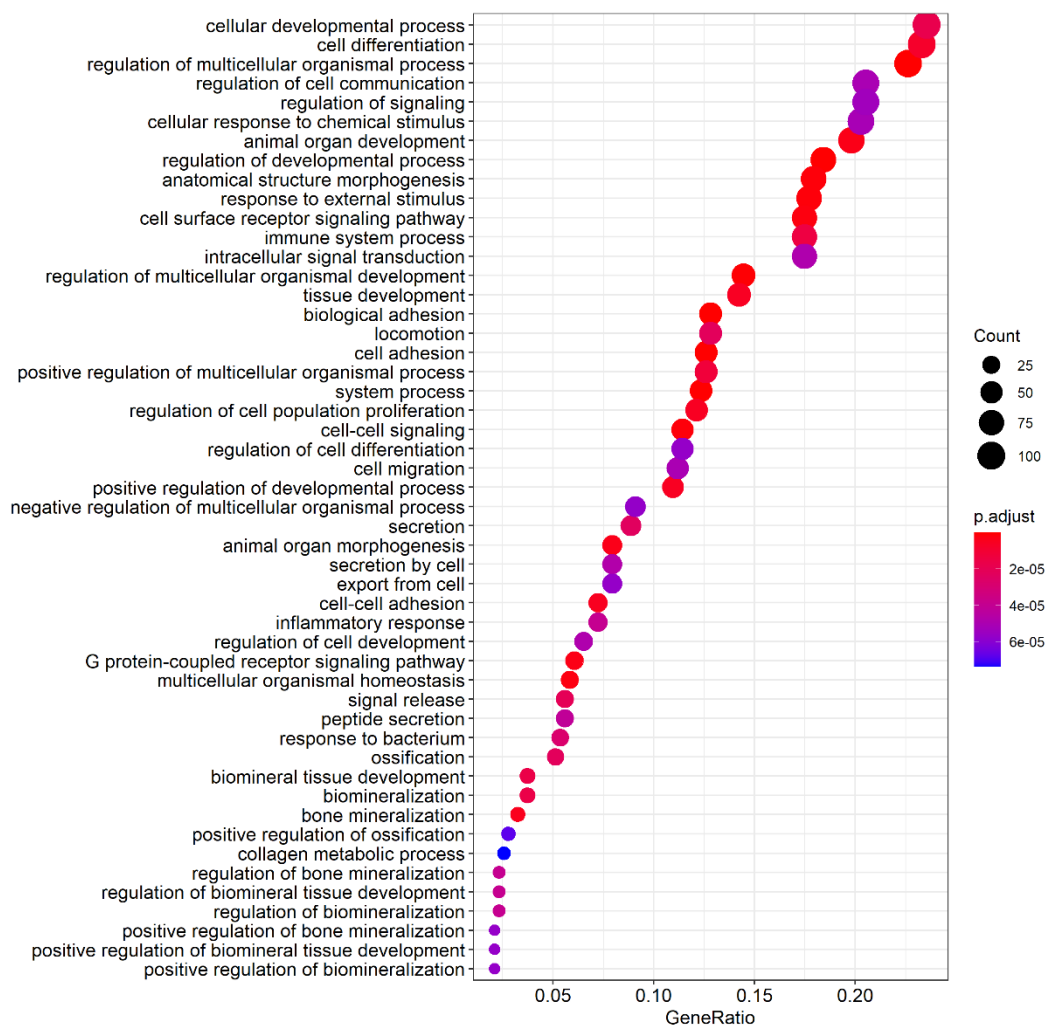


Figure 3-11: Top 50 GO terms for DE genes of comparison of MAC-T and pbMEC control groups. The Bioconductor package clusterProfiler was used for functional overrepresentation analysis of DEGs with the significance threshold of  $p_{adj} < 0.05$  and  $\log_2 FC \geq |2|$ . The term list of biological processes (BP) is represented in dots, with size representing the number of genes and color representing the  $p_{adj}$ . Terms are hierarchically listed according to their gene ratio, the number of genes related to GO term / total number of significant genes. The Bonferroni Hochberg correction was applied, and a  $q$ -value cutoff ( $FDR$ )  $\leq 0.05$  was set.



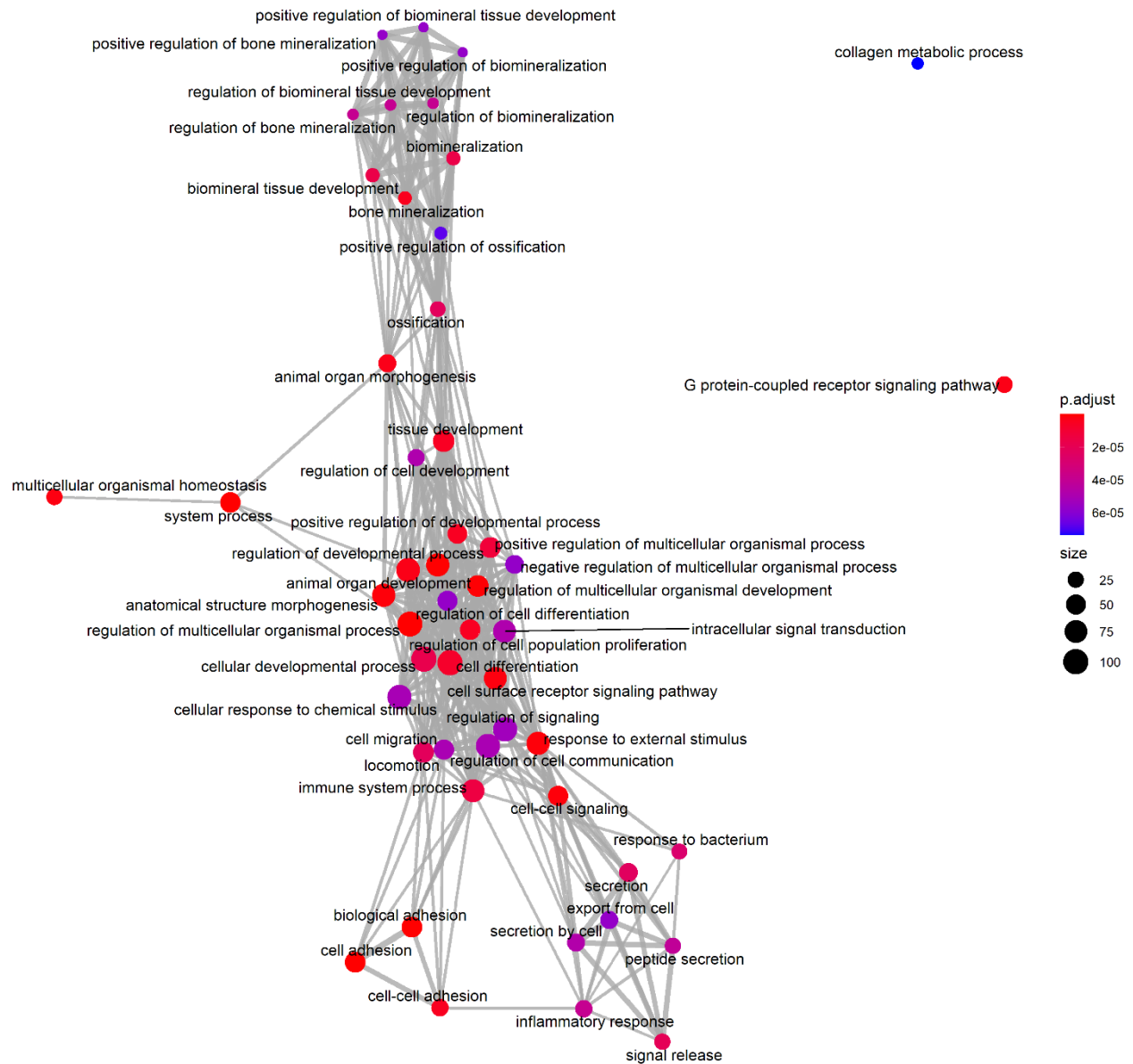


Figure 3-12: Relationship between top 50 GO terms from functional analysis of DE genes of the comparison of MAC-T and pbMEC control groups. These are the GO terms for biological processes (BP) in which the 1829 DEGs between the cell types were enriched. The Bonferroni Hochberg correction was applied, and a q-value cutoff ( $FDR \leq 0.05$ ) was set.

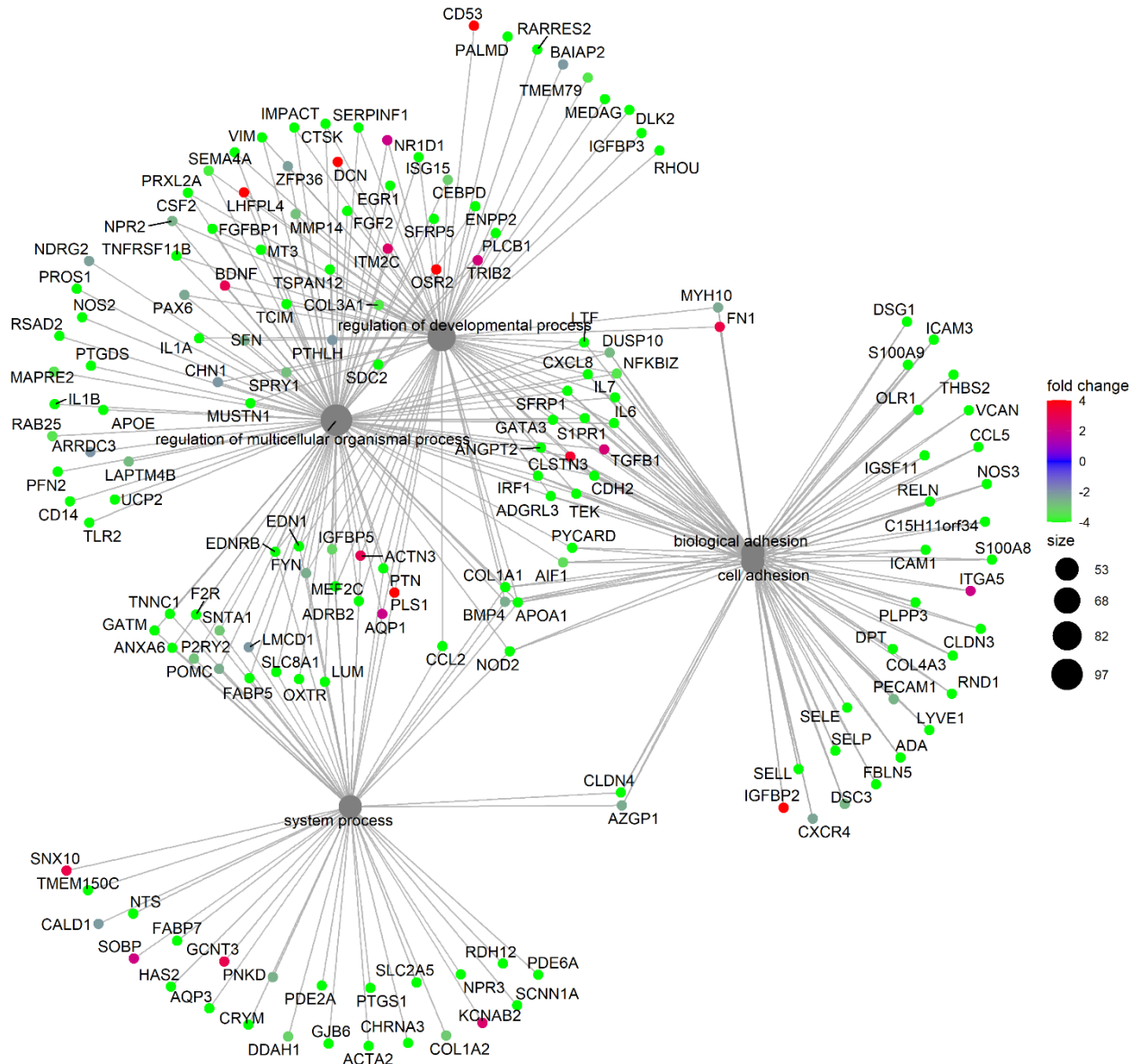


Figure 3-13: Category netplot for functional analysis of DE genes of the comparison of MAC-T and pbMEC control groups. The plot shows the relationships between the genes associated with the top five most significant GO terms for biological processes (BP) and the fold changes of the significant genes related to these terms (color). Since some of the high fold changes are drowned out due to an extensive range, a maximum fold change value of 4 and -4 was set.

### 3.2.2 Effect of Pathogen on Transcriptome of MAC-T

The genome-wide transcriptome analysis of MAC-T pathogen and control-challenged samples revealed a total of 12547 expressed genes. Data filtering identified differentially expressed genes at a statistically significant level of  $p_{adj} < 0.05$ . During data exploration and performing the analysis analog to the above (3.2.1), a stringent  $\log_2FC \geq |2|$  was applied and revealed four genes (*CYP1A1*, *CYP1B1*, *NR4A3*, *CCL2*) for the MAC-T comparison of *E. coli*<sub>1303</sub> versus control and two genes (*CYP1A1*, *CYP1B1*) for MAC-T comparison of *S. aureus*<sub>1027</sub> versus control. If  $\log_2FC$  cut-off was disregarded and only  $p_{adj} < 0.05$  was used to identify significant differentially expressed genes, 102 DEGs significantly differed between *E. coli*<sub>1303</sub> and control in MAC-T. Of these, 25 DEGs were downregulated, and 77 DEGs were

upregulated in response to *E. coli*<sub>1303</sub>, as shown in Figure 3-14A. Interestingly, two of the upregulated DEGs were bta-mir-2887-1 localized on chromosomes 2 and 27.

In contrast to the *E. coli* results, only three DEGs (two genes coding for enzymes of the cytochrome P450 superfamily *CYP1A1*, *CYP1B1*, and receptor protein-tyrosine kinase *EPHA4*) were found to be significantly upregulated between *S. aureus*<sub>1027</sub> and control in MAC-T (Figure 3-14B).

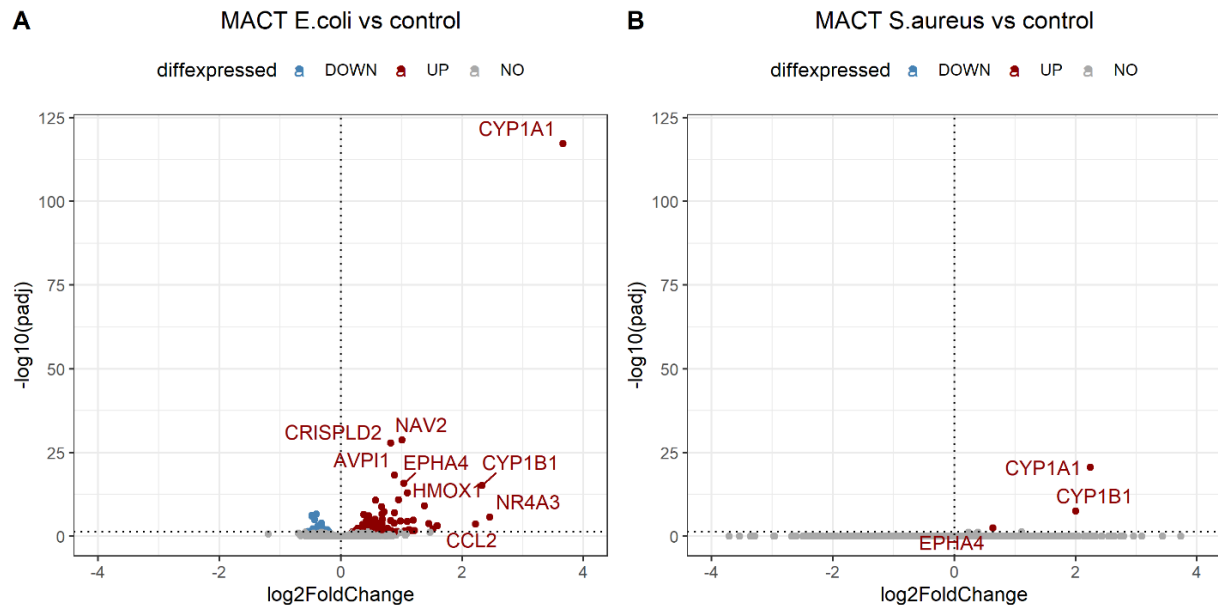


Figure 3-14: Volcano plots for MAC-T comparing pathogen versus control-challenged cells. The horizontal line marks  $p_{adj} = 0.05$ , whereas the vertical line separates positive and negative numerical values for the  $\log_2\text{FoldChange}$ . 102 DEGs were found to be significantly different between *E. coli* and control in MAC-T (A). 25 of the DEGs were downregulated, and 77 DEGs were upregulated

The 110 DEGs between *E. coli* and control in MAC-T were subjected to enrichGO analysis in R. The overrepresentation and enrichment of DEGs only worked if the statistical power had been removed in the functional analysis in this specific case. Among the top 10 overrepresented GO terms of the BP were intracellular signal transduction, response to the bacterium, and leukocyte chemotaxis, as shown in Figure 3-15A. The top 50 enriched GO terms showed a tight relationship network (Figure 3-15B). Immune genes *CCL2* and *CXCL5*, CCAAT/enhancer-binding protein beta (CEBPB), Ras GTPase-activating protein 1 (*RASAI*), and Gap junction alpha-1 protein (*GJAI*) had the most connections within critical BPs in MAC-T exposed to *E. coli*<sub>1303</sub> versus control.

In contrast, the small data input for functional analysis revealed no significant GO results for BP, CC, and MF for MACT exposed to *S. aureus*<sub>1027</sub> versus control.

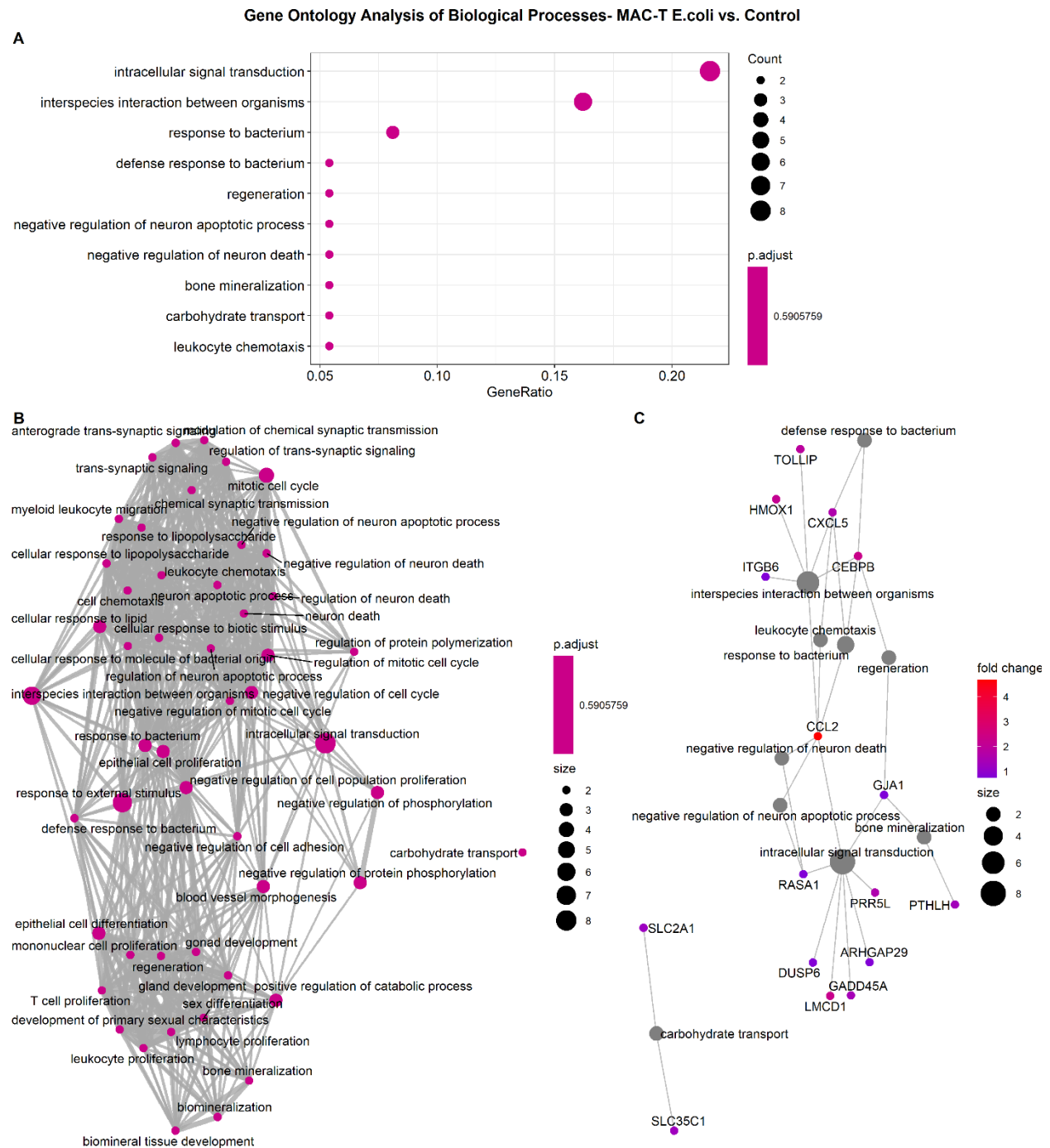


Figure 3-15: Gene Ontology (GO) Analysis for biological processes for DEGs between MAC-T *E. coli*<sub>1303</sub> versus Control. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) but without applying the Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The dot plot (A) shows the top 10 overrepresented terms for DEGs. The size of the dots correlates with the number of genes (count) per term. Gene set enrichment map (B) depicts the relationship of the top 50 categories in which the DEGs are enriched. The size of the dots in the network correlates with the number of genes per term. The Cnetplot (C) shows the top 10 terms with genes associated with them.

### 3.2.3 Effect of Pathogen on Transcriptome of pbMEC

The analysis of pbMEC control and pathogen-challenged samples revealed a total of 12547 expressed genes. Data filtering for  $p_{adj} < 0.05$  again identified differentially expressed genes at a statistically significant level. Applying a stringent  $\log_2FC \geq |2|$  revealed seven genes (*CYP11A1*, *CYP11B1*, *SLC5A8*,

*MXII*, *CA4*, *ENSBTAG00000026527*, *ARTN*) to be different between *E. coli*<sub>1303</sub> and control in pbMEC, and two genes (*CYP1A1*, *CYP1B1*) comparing *S. aureus*<sub>1027</sub> with control in pbMEC.

Analysis disregarding the  $\log_2FC \geq |2|$  cut-off revealed significantly expressed genes up- or downregulated as depicted in the volcano plots (Figure 3-16). 409 DEGs were identified comparing *E. coli*<sub>1303</sub> and control in pbMEC, of which 49 DEGs were downregulated, while 360 DEGs were upregulated (Figure 3-16A). 88 DEGs were identified comparing *S. aureus*<sub>1027</sub> and control in pbMEC, of which 15 DEGs were downregulated and 73 DEGs were upregulated (Figure 3-16B).

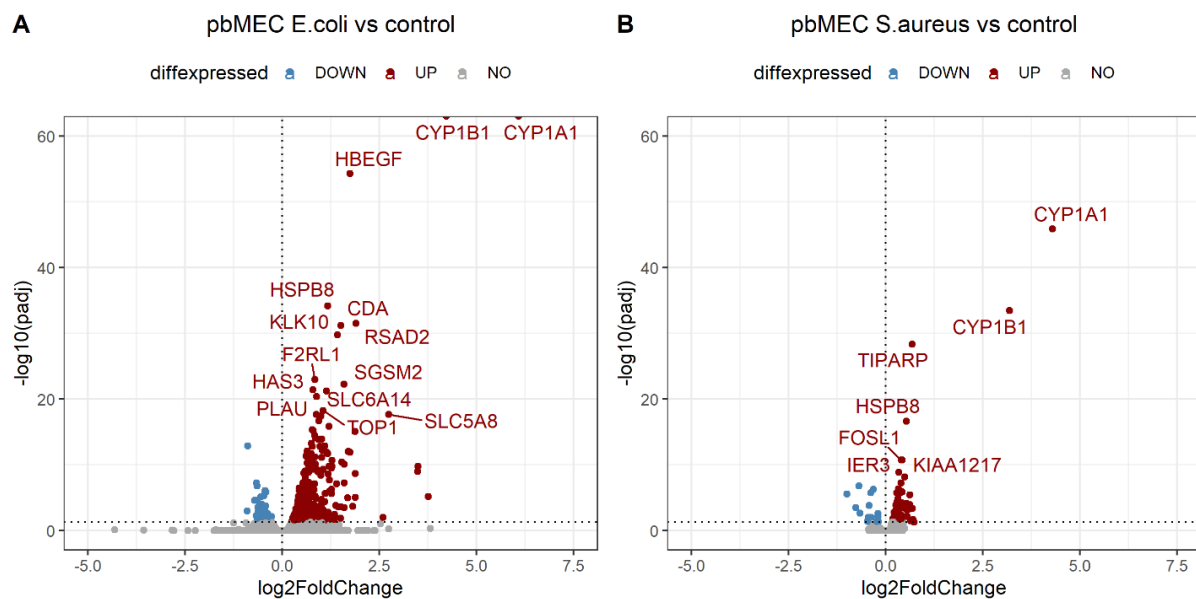


Figure 3-16: Volcano plots for pbMEC comparing pathogen versus control-challenged cells. 409 DEGs were identified in comparing *E. coli*<sub>1303</sub> and control in pbMEC (A). 49 of the DEGs are downregulated, while 360 DEGs are upregulated. 88 DEGs were identified in comparing *S. aureus*<sub>1027</sub> and control in pbMEC (B). 15 DEGs were downregulated, and 73 DEGs were upregulated, as shown in B. The horizontal line marks  $p_{adj} = 0.05$ , whereas the vertical line separates positive and negative numerical values for the  $\log_2\text{FoldChange}$ .

The identified 409 DEGs between *E. coli* and control-challenged pbMEC were subjected to functional analysis analog to previous analyses. The genes in pbMECs correspond to similar responses to an *E. coli* challenge as in MAC-T. The GO analysis provided insights into the potential roles of genes in response to external stimulus, immune system process, granulocyte chemotaxis, and apoptotic processes. In the tightly interwoven network, the most significant BP stands out by being colored in red, representing the smallest  $p_{adj}$  (Figure 3-17B). When comparing the results of pbMEC with MAC-T, more DEGs between *E. coli* and control in pbMEC were associated with GO terms of BP. This finding is evident in the size of the dots in the dot plots and in the cnetplots with several genes associated with enriched terms (Figure 3-17).

Functional annotation using GO terms of CC indicated that the differentially expressed genes between pbMEC exposed to *E. coli*<sub>1303</sub> and control were associated with a widespread of cellular components (appendix Figure B-11). Among them are the early endosome, fibrillar center, external side of the plasma membrane, extracellular region, and bicellular tight junctions. The analysis using GO terms of MF

highlighted the importance of signaling receptor activity, molecular transducer activity, and integrin binding, as depicted in appendix Figure B-12.

The significance level of the GO analysis differs when comparing the two pathogen challenges in pbMEC, as evidenced by smaller values in *E. coli* versus control than in *S. aureus* versus control indicated in the legends of Figures 3-17A and 3-18A. Also, fewer genes are linked to the identified functional terms of *S. aureus* than those for *E. coli*, comparing Figures 3-17C and 3-18C. Thus, the data demonstrate that more genes were differentially expressed in response to *E. coli*<sub>1303</sub>, and the cells responded stronger than to *S. aureus*<sub>1027</sub>.

The GO analysis showed that the most significantly enriched BP categories with DEGs from comparing *S. aureus*<sub>1027</sub> and control in pbMEC were immune system processes, cell migration, and muscle development processes (Figure 3-18). Of the 11 genes associated with immune system processes, *CYP26B1* showed the highest fold change in this context. The GO term analysis of MF revealed that the DEGs for the comparison of *S. aureus*<sub>1027</sub> versus control in pbMEC were involved in an array of binding processes (appendix Figure B-14). These were mainly associated with membranes. Notably, these findings were not significant for the GO analysis of MF (appendix Figure B-13).





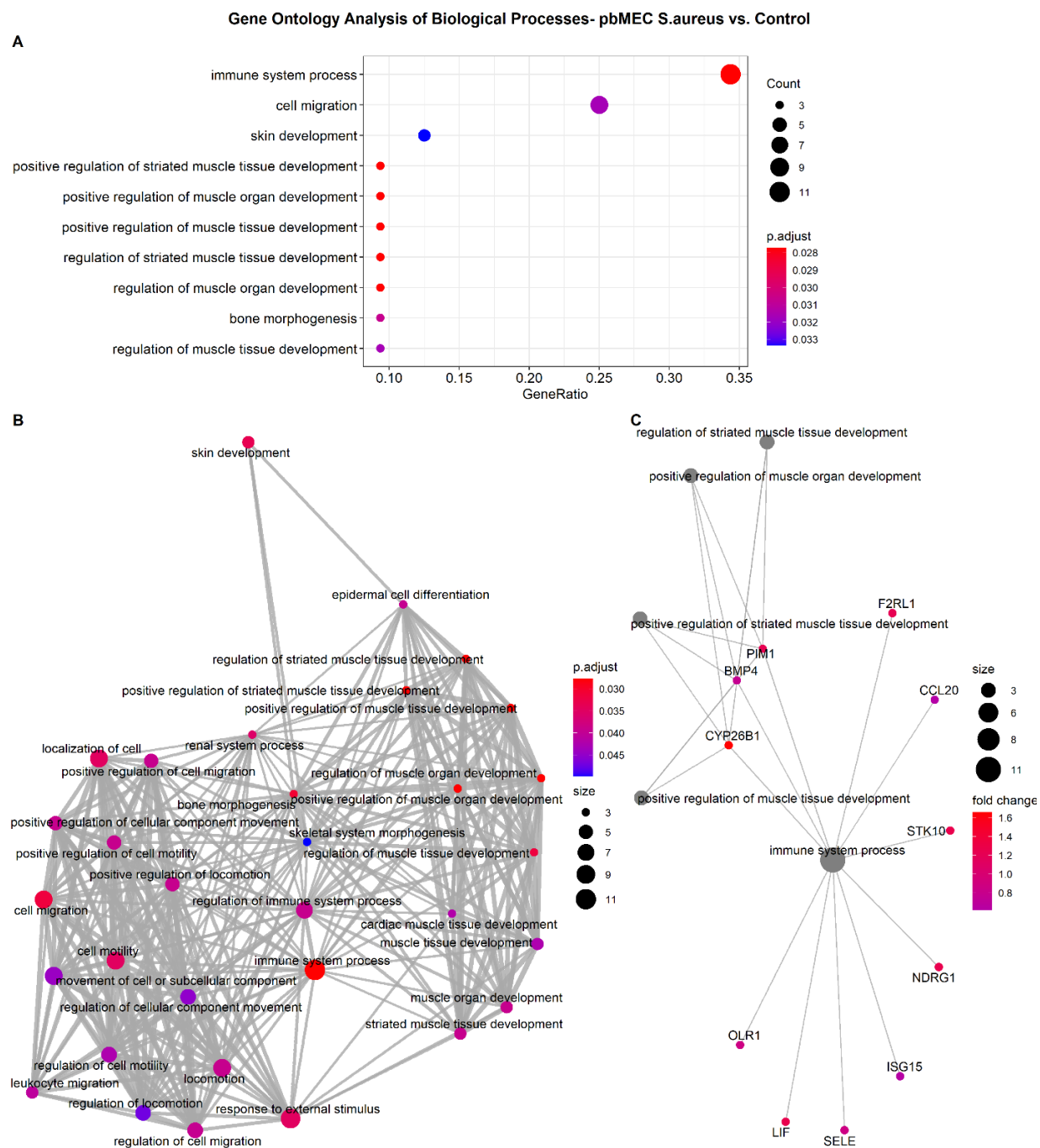


Figure 3-18: Functional Analysis of DEGs between *S. aureus*<sub>1027</sub> and control in pbMEC. Dot plot (A) shows the top 10 overrepresented GO terms of the biological processes (BP). The enrichment map (B) indicates the relationship between the top 50 GO terms of the BP. The cnetplot (C) visualizes the top 5 enrichment terms and associated genes.

### 3.3 Genome-wide Sequencing in Bovine Mastitis Cell Models Targeting the Epigenome

While genome-wide gene expression was examined using RNA-seq, genome-wide sequencing for epigenetic profiling of histone modifications was performed with ChIP-seq. Several verification steps were performed for experimentation and data analysis to ensure the quality of the data. These steps are explained in further detail below.



### 3.3.1 X-ChIP Establishment and Verification

ChIP-seq was used to identify histone modifications associated with DNA throughout the genome. The primary objective was to describe and compare the cell line MAC-T and the primary cells pbMEC in their epigenetic profile for both an active and a repressive histone mark (H3K4me3 versus H3K27me3). Henceforth, epigenetic research materials and kits were obtained from Diagenode, with part of the work performed in the wet lab on site in Liege, Belgium (the company's headquarters), and in accordance with FAANG requirements.

The extraction of primary cells from udder tissue was based on protocols available at the Leibniz Institute for Farm Animal Biology (FBN) that could be used and adapted for the study's purposes and in collaboration with another project. Cryo-preserved pbMEC and MAC-T were used as starting material for the experiments. Both cell models, MAC-T and pbMEC, were exposed to the bacteria *E. coli*<sub>1303</sub> and *S. aureus*<sub>1027</sub> for 3 h or remained unchallenged as a reference condition (mock). Subsequently, cells were fixed with formaldehyde, harvested, and pelleted for a subsequent experiment using the iDeal ChIP-seq kit (Diagenode). Decisive steps of the ChIP procedure required optimization for the bovine epithelial cell models. Lysis of the cell membrane was performed with buffers according to the manufacturer's instruction and included 15x douncing strokes and the consistent use of the same starting material of the IP. Next, shearing of the chromatin material at a steady 4°C was performed, testing the number of shear cycles for each cell model at a constant volume. Assessment of the sheared DNA with the Bioanalyzer revealed profiles that showed a pronounced bimodal distribution of fragments, with a lower and a higher bp fraction (Figure 3-19). The smaller fragments (100 to 500 bp) representing DNA from single or multiple nucleosomes were particularly interesting to ChIP. In the following, ten sonication cycles for the pbMEC and 12 sonication cycles for the MAC-T cell line were used.

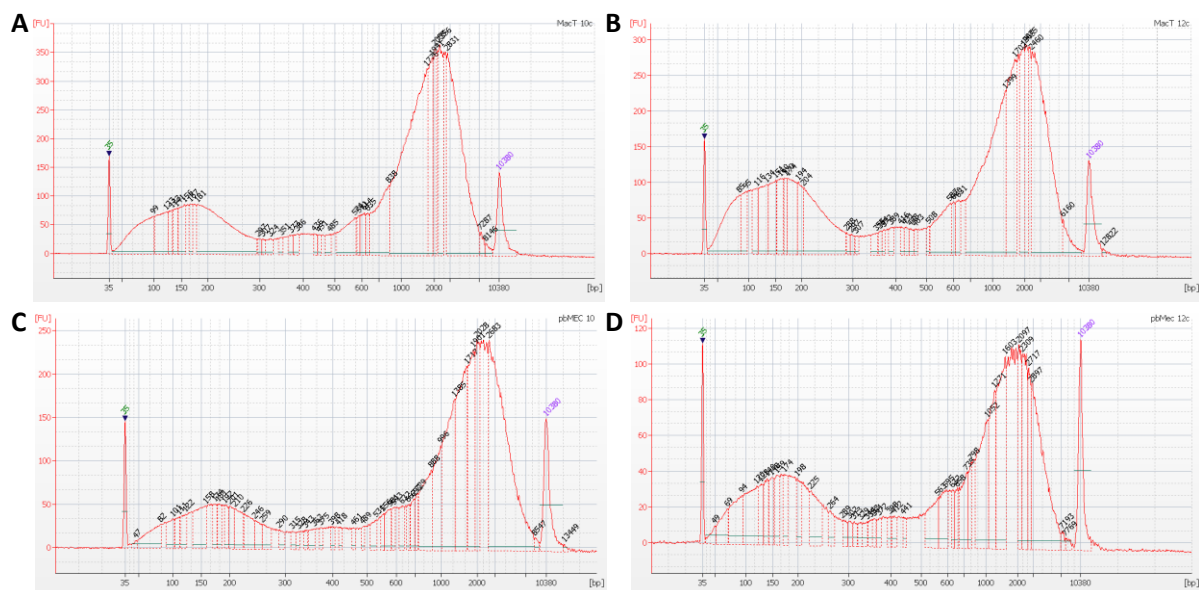


Figure 3-19: Shearing optimization experiment in bovine model cells. Fixed MAC-T and pbMEC cells that remained unchallenged were lysed, and nuclei were extracted using the Lysis Buffers iL1b and iL2 of the Diagenode iDeal ChIP-seq kit and a Dounce homogenizer. Samples corresponding to 3.6 million cells were suspended in 200  $\mu$ L of shearing buffer iS1b and sheared separately during ten cycles and 12 cycles (30 sec

ON/30 sec OFF) with the Bioruptor® Pico combined with the Bioruptor® Water cooler. 50 µl of DNA was used for the shearing analysis. Upper panels: chromatin extracted from MAC-T cells sheared for 10 (A) or 12 cycles (B). Bottom panels: chromatin extracted from pbMEC cells sheared for 10 (C) or 12 (D) cycles. Bioanalyzer profiles for both cell models show base pairs (bp) on the x-axis and fluorescence unit (FU) on the y-axis. Note that the y-scaling varies between the graphs. Lower and upper markers are represented by peaks at 35 and 10380 bp.

The optimization experiments for X-ChIP also included testing two amounts of antibodies for each of the cell types. Testing antibody amounts has yet to be considered further in N-ChIP. The efficiency of the automated X-ChIP method performed with the IP Star automated system (Diagenode) was determined by qPCR measurements. All primers were first tested on a standard curve on gDNA. Primers for positively and negatively enriched genomic regions were established for each antibody against the epigenetic marks to be tested and were used in X-ChIP control qPCR. The determined Ct values were used to determine the %Input per genomic region and the respective antibody amount (Figure 3-20, appendix Table B-1). IgG served again as a negative control. Positive genomic regions showed strong enrichment for specific histone marks, while the target protein was unbound in negative genomic regions. Primer sets and antibody amounts were chosen based on the fold enrichment of positive over the negative enriched genomic region (referred to as recovery) (Table 3-1). Some control regions differ between the two cell models, as does the amount of antibody used. Positive controls for active histone marks were considered negative controls for repressive histone marks, as they were available for H3K4me3 and H3K27me3. H3K4me3 control regions differed markedly in higher recovery compared to other epigenetic marks in this study.

Table 3-1: Summary of the conditions for the ChIP experiment used for the bovine cell models. The settings listed here proved suitable for the experiments during the optimization tests. They differ for shearing cycles and partly for antibody amount and control regions for X-ChIP control qPCR between the cell models, pbMEC and MAC-T.

Cell type:	shearing cycles [30 sec ON, 30 sec OFF]	ChIP antibody (Ab)	Ab amount	Positive genomic region	Negative genomic region
pbMEC	10 x	H3K4me3	0.5 µg	GAPDH_TSS_3	PRSS1 or TSH2B
		H3K27me3	0.5 µg	TSH2B	GAPDH_TSS_3
MAC-T	12 x	H3K4me3	1 µg	GAPDH_TSS_3	PRSS1 or TSH2B
		H3K27me3	0.5 µg	PRSS1	GAPDH_TSS_3
		H3K27Ac	0.5 µg	GAPDH_TSS_3	TSH2B
		H3K4me1	1 µg	ActB1	TSH2B
		CTCF	2 µg	ActB3	TSH2Bthe

Before the library preparation for sequencing with MicroPLEX v2 protocol, the X-ChIP control-qPCR measurements were reproduced for all five FAANG marks for all sample conditions, considering cell models and challenges (Figure 3-21, Figure 3-22). The measurements for the five marks yielded a similar recovery as in the establishment of the assay. The results are shown as mean values of % Input of replicate measurements, and the error bars indicate the standard error of the mean (Figure 3-21, Figure 3-22). Across conditions (Mock, *E. coli*<sub>1303</sub>, *S. aureus*<sub>1027</sub>), the chosen genomic regions served as positive or negative genomic regions for all marks. For H3K4me3, the same genomic regions could be applied for both cell models (Figure 3-21). For H3K27me3, protease, serine one (*PRSS1*) was a stronger positive region than testis-specific histone H2B variant (*TSH2B*) in MAC-T, while it was vice versa in pbMEC

(Figure 3-21). Control regions were also established and verified for the three epigenetic markers, H3K4me1, H3K27Ac, and CTCF, as shown in Figure 3-22.

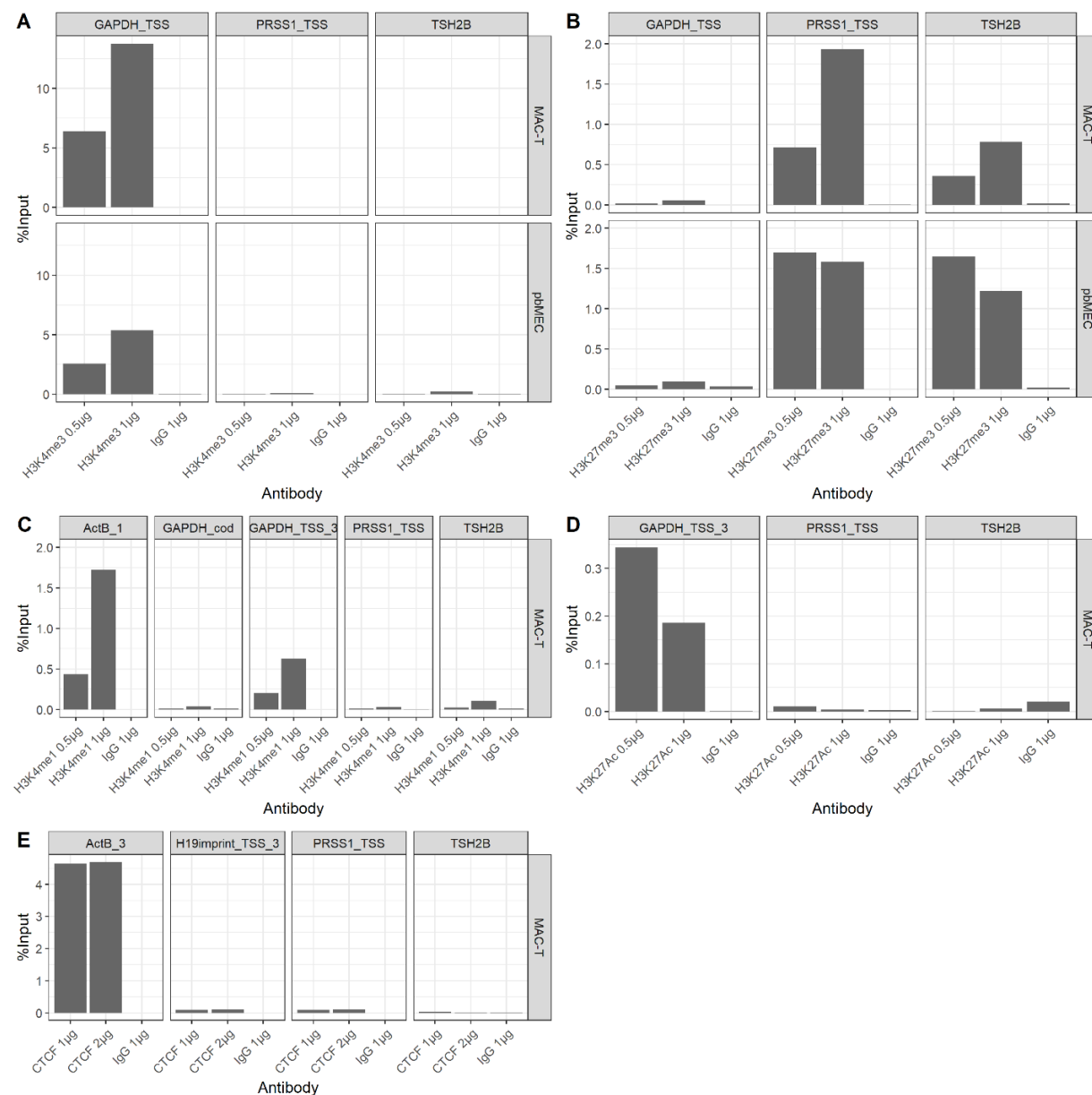


Figure 3-20: X-ChIP control-PCR to determine control regions and antibody amount for MAC-T and pbMEC. Plots show measured enriched signal (%Input) for the respective ChIP positive and negative controls developed in unchallenged cell samples for the antibodies H3K4me3 (A), H3K27me3 (B), H3K4me1 (C), H3K27Ac (D), and CTCF (E). IgG served as a negative control IP with a low background signal. Two amounts of antibody were tested as given in µg on the x-scale for each FAANG mark.

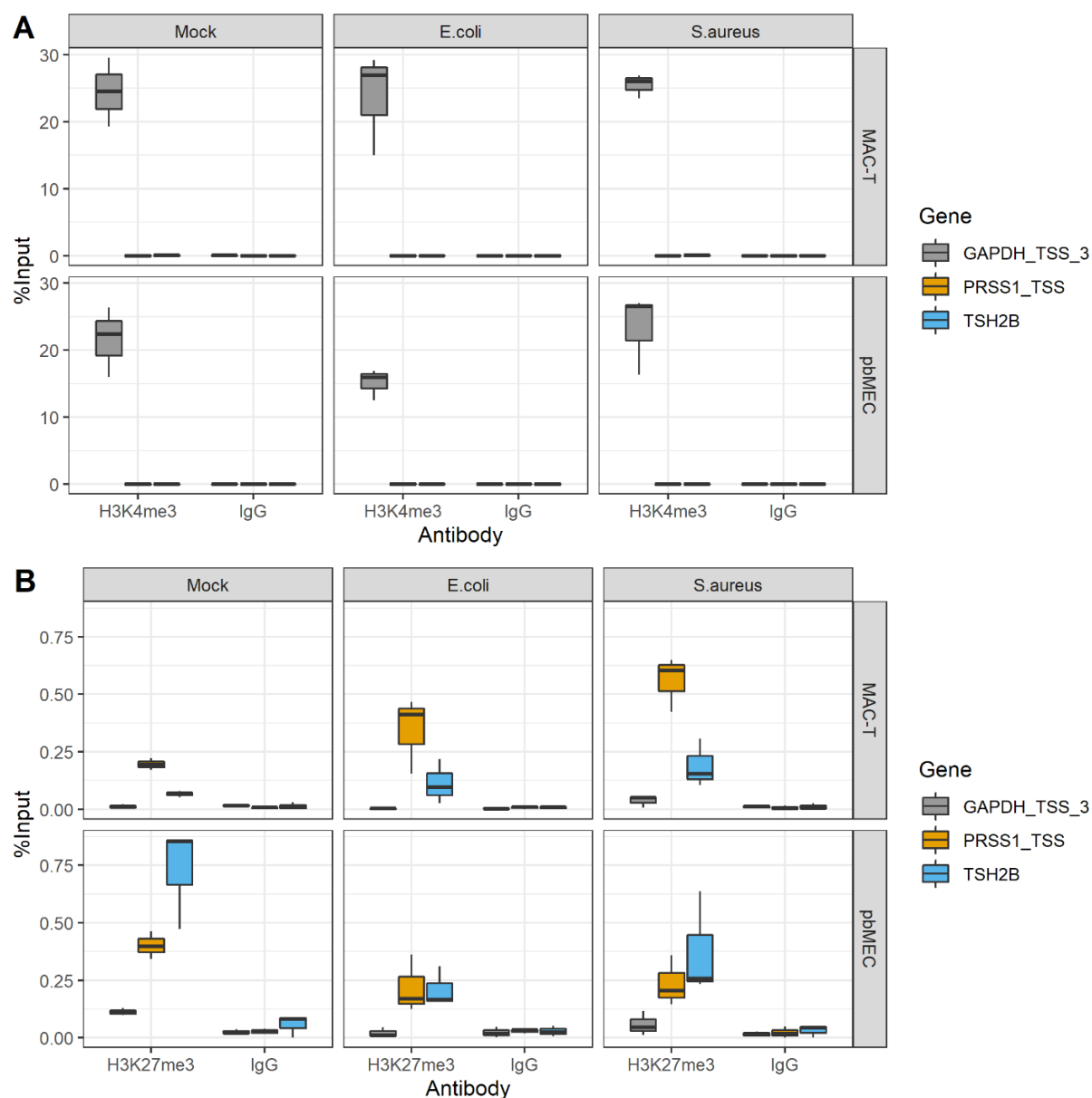


Figure 3-21: X-ChIP control qPCR in bovine model cells. Chromatin from MAC-T and pbMEC challenged for 3 h with pathogens *E. coli*<sub>1303</sub> or *S. aureus*<sub>1027</sub> or unchallenged (Mock) were processed with the iDeal ChIP-seq Kit (Diagenode) targeting FAANG core marks H3K4me3 and H3K27me3. Purified DNA was used for X-ChIP-qPCR analyses to verify genomic control regions (GAPDH, PRSS1, TSH2B) for active histone mark H3K4me3 (A) and repressive histone mark H3K27me3 (B). Shown are means of biological triplicates (H3K4me3, H3K27me3) as %Input with error bars indicating the standard error of the mean (SEM) of duplicate measurements in qPCR. IgG was a negative control antibody measuring background signal in each genomic region. Plots were generated with ggplot2 and cowplot package in R.

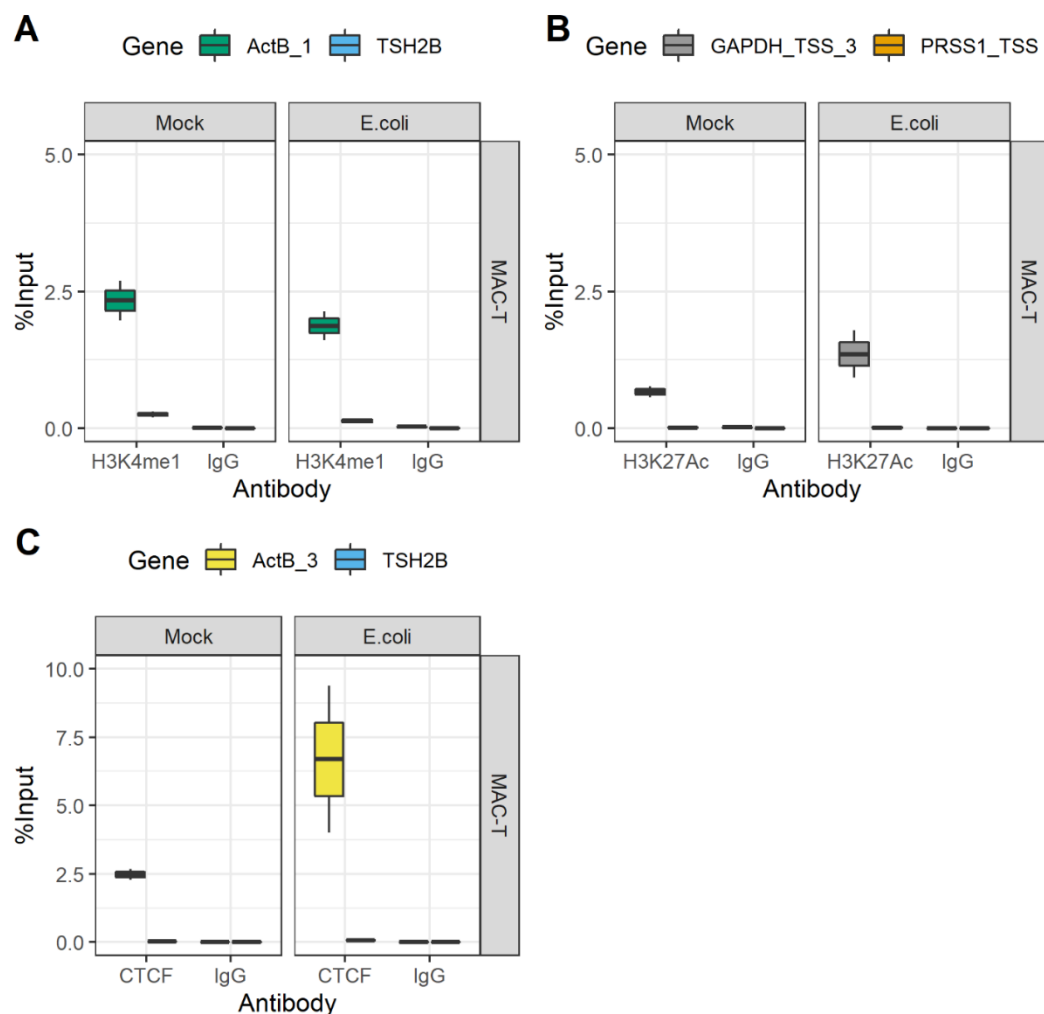


Figure 3-22: X-ChIP control qPCR in MAC-T. Chromatin from MAC-T challenged for 3 h with pathogens *E. coli*<sub>1303</sub> or unchallenged (Mock) was processed with the iDeal ChIP-seq Kit (Diagenode) targeting FAANG core marks H3K4me1, H3K27Ac, and CTCF. Purified DNA was used for X-ChIP-qPCR analyses to verify genomic control regions (ACTB, GAPDH, PRSS1, TSH2B) for histone mark H3K4me1 (A) and H3K27meAc (B), and transcription factor CTCF (C). Shown are means of biological duplicates (CTCF, H3K27ac, H3K4me1) as %Input with error bars indicating the standard error of the mean (SEM) of duplicate measurements in qPCR. IgG was a negative control antibody measuring background signal in each genomic region. Plots were generated with ggplot2 and cowplot package in R.

### 3.3.2 X-ChIP-seq Exploration and Quality Check

The ChIP-Seq libraries were sequenced, and 50 bp single reads were generated on an Illumina HiSeq 4000 instrument. Initial computational analysis of the data was outsourced to Diagenode. With the release of the new genome assembly (ARS-UCD1.2), the data analysis pipeline was rerun to compare the results to the obtained RNA-seq data. Visualization and exploration of the data were performed independently. The high-throughput ChIP-seq raw data passed the first stage of quality assessment with the FASTQC tool, followed by trimming, read mapping, filtering out redundant reads, peak calling, differential binding analysis, and peak annotation. H3K4me3 was saturated on average at a lower sequencing depth (44.2 million reads) than H3K27me3 (60.7 million reads), according to the data. The ChIP-seq statistics (appendix Table B-3) revealed a high read mapping efficiency with an average of 85.7% (range: 74.6 to 87.6%) with four outliers below 80%. More peaks were identified for the active

histone mark H3K4me3 (average 24,227, range: 20,218 to 30,197) than for H3K27me3 (average 3,225, range: 2,722 to 4,076). Peak numbers between the two cell types differed slightly. Between replicates per IP per cell type and challenge, the percentage relative standard deviation (%RSD) was less than 10% for most samples with a few outliers in the range of less than 20%RSD (appendix Table B-4).

Further assessment of the reproducibility of the samples was done by plotting the profiles of the cumulative read coverages for each sample using the plotFingerprint option of deepTools (Figure 3-23). The difference of the ChIP signals over the background in control conditions was thus determined. Close overlap of the curves of each condition was taken as an indication of a high replicate match. Typical fingerprints were detected for each histone label, with H3K27me3 closer to the input than H3K4me3. Moreover, the respective fingerprint for H3K4me3 and H3K27me3 was comparable between MAC-T and pbMEC (Figure 3-23); a Pearson correlation heatmap can be found in the appendix (Figure B-15) showing the clustering of samples for the control challenge. The fingerprint H3K27me3 was identified closer to the inputs. Inputs represent control DNA that did not undergo the ChIP procedure but have been sequenced to determine the background signal. A Pearson correlation heatmap in Figure 3-24 depicts the hierarchical clustering of the samples into two main clusters for the histone marks H3K4me3 and H3K27me3 and, within them, subclusters for the different cell models. However, no clusters were found for the different challenges. The results are supported by correlation coefficients close to zero, indicating no linear relationship, and a coefficient close to 1, indicating a correlation of the samples, simultaneously shown in intense color.

In addition to clustering in the heatmap, a PCA was used to identify patterns in the data based on the read distribution across the genome and express the data in such a way as to highlight their similarities and differences (Figure 3-25). Distinct clusters per cell replicates and histone marks were detected that overlapped closely. Similarly, the fingerprint H3K27me3 was identified closer to the inputs. Similarly, the effect of the pathogen challenge was not identifiable in the PCA when challenge samples for *E. coli* and *S. aureus* were included alongside control samples and plotted for H3K4me3 and H3K27me3 (see appendix Figure B-16). The scree plot supports the emphasis on the two principal components included as the driving factors for the analysis.



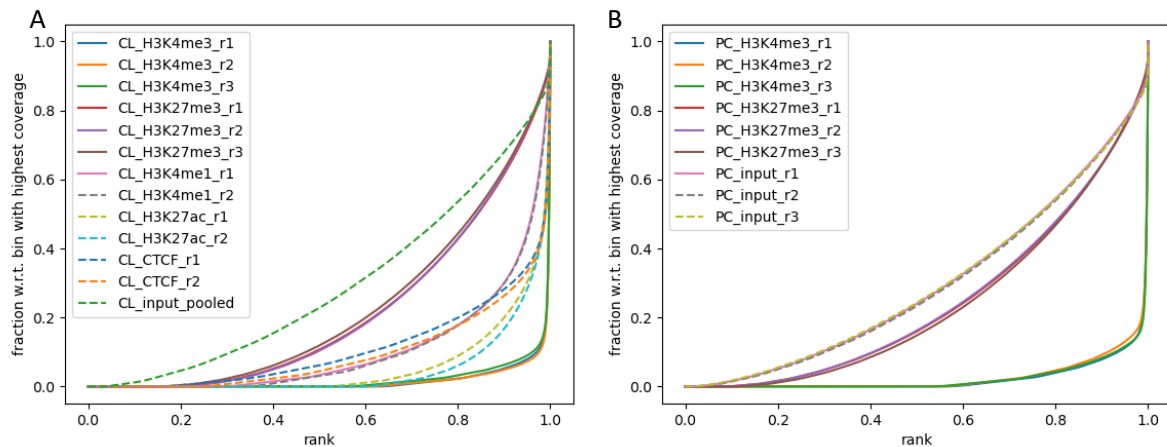


Figure 3-23: Plotfingerprints generated by deepTools for ChIP assays in two bovine cell models. The application assesses the enrichment strength of a ChIP experiment. The plots show the distribution of the total number of reads (y-axis) along the percentages of the genome (x-axis) for cell line (CL) MAC-T (A) and primary cells (PC) pbMEC (B). The diagonal line represents the input, which shows a differentiable profile from the performed ChIP assays. The similar shape of the curves between the same ChIP assay and different cell types is identifiable.

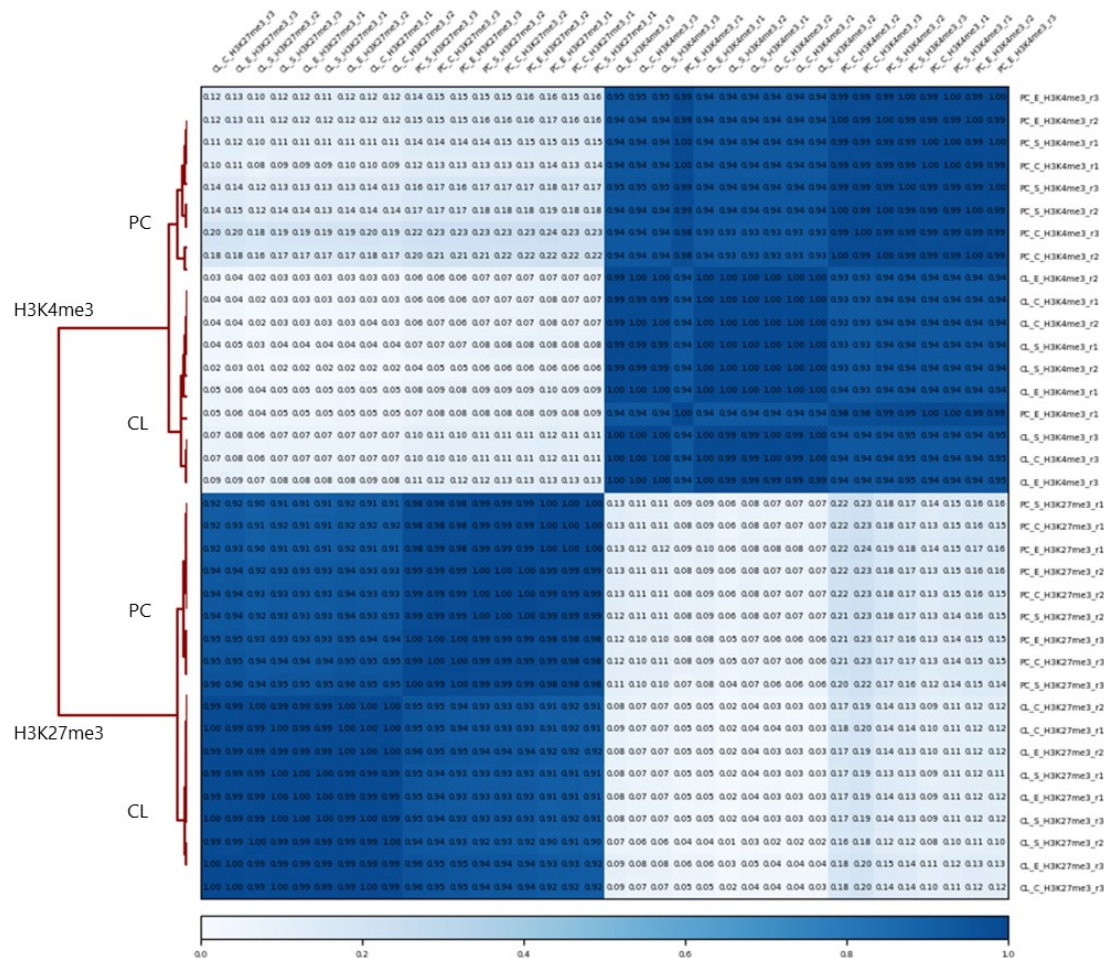


Figure 3-24: Plot correlation analysis for ChIP-seq in bovine model cells based on genomic reads. Pearson correlation between ChIP replicates of pbMEC (primary cells = PC) and MAC-T (cell line = CL) samples is shown in color shading and ranges from 0 to 1, as shown in numbers in a correlation matrix. Sample replicates (r1 to r3) are hierarchically clustered by the condition in the heatmap supported by the dendrogram on the left. The plot was generated with the deepTools plotCorrelation tool.

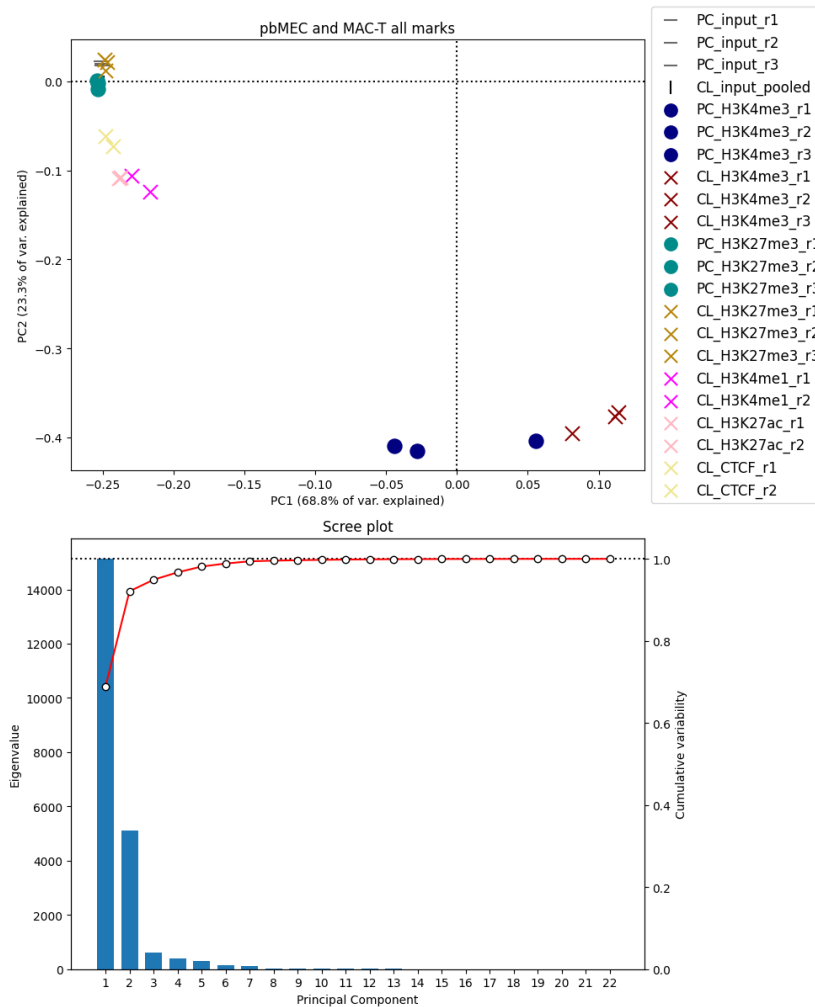


Figure 3-25: PCA of ChIP experiments for control-challenged samples. Five ChIP assays (H3K4me3, H3K27me3, H3K4me1, H3K27Ac, and CTCF) and control DNA (input) are shown in the PCA. The scatter of sample replicates ( $r$ ) for the primary cells (PC) pbMEC, shown as dots, and the cell line (CL) MAC-T, shown as crosses, is based on the similarity or difference of the data along the principal components along the axes.

### 3.3.3 Differential Binding Analysis and Annotation of X-ChIP-seq Peaks

Peak sets are associated with genomic coordinates that can be assigned to specific genes and genomic regions. Based on these histone-DNA binding sites, functional characteristics for specific histone marks can be inferred. Differential binding analysis of ChIP-seq peaks for each histone mark H3K4me3 and H3K27me3 first identified statistically significant differential binding sites between sample groups from MAC-T and pbMEC control samples. For H3K4me3, 24449 differential binding sites were identified, whereas, for H3K27me3, 2480 sites were identified. Secondly, bacteria and control challenges within the cell types were analyzed. The analysis revealed significant binding sites for the active histone mark H3K4me3 when comparing *E. coli*<sub>1303</sub> and control-challenged pbMECs. All other comparisons of pathogen challenge with control cells in MAC-T and pbMEC revealed non-significant results.

Versatile plotting allowed the visualization of the differential binding analysis (Figure 3-26 to Figure 3-28). A correlation heatmap (panel A in each of Figures 3-26 to 3-28) provided initial clustering on the sample's basis using the cross-correlations of each row of the binding matrix. An MA plot of MAC-T versus pbMEC shows identified sites as significantly differentially bound colored as red dots in the plot



(panel D in Figure 3-26, Figure 3-27). A PCA shows the clustering of sample groups based on all binding sites (panel B); a second PCA is based on all differential binding sites (panel C). A final heatmap (panel E) shows the clustering of differential binding sites across sample groups.

The statistically relevant genomic binding sites were annotated with HOMER from the identified peak sets per histone mark. The identified differential binding sites between MAC-T and pbMEC for H4K4me3 and H3K37me3 in the control condition and *E. coli*<sub>1303</sub>-challenged versus control results in pbMEC H3K4me3 are illustrated in Figure 3-29. While under control-challenge conditions, most differential binding sites fell into intergenic and intron regions for both histone marks; there is a relative lower occurrence in intergenic regions, but a gain in the promoter-TSS region after the *E. coli*<sub>1303</sub> challenge in pbMEC for H3K4me3 (panel E and F in Figure 3-29). The total number of differential binding sites identifiable under the same peak width conditions for H3K4me3 varied in a 10<sup>3</sup> scale across samples. Detailed annotation (appendix Figure B-17) also revealed that differential binding sites between MAC-T and pbMEC fell into genomic regions coding for lncRNAs.

## H3K4me3, summit=1,000, MACT vs pbMEC control

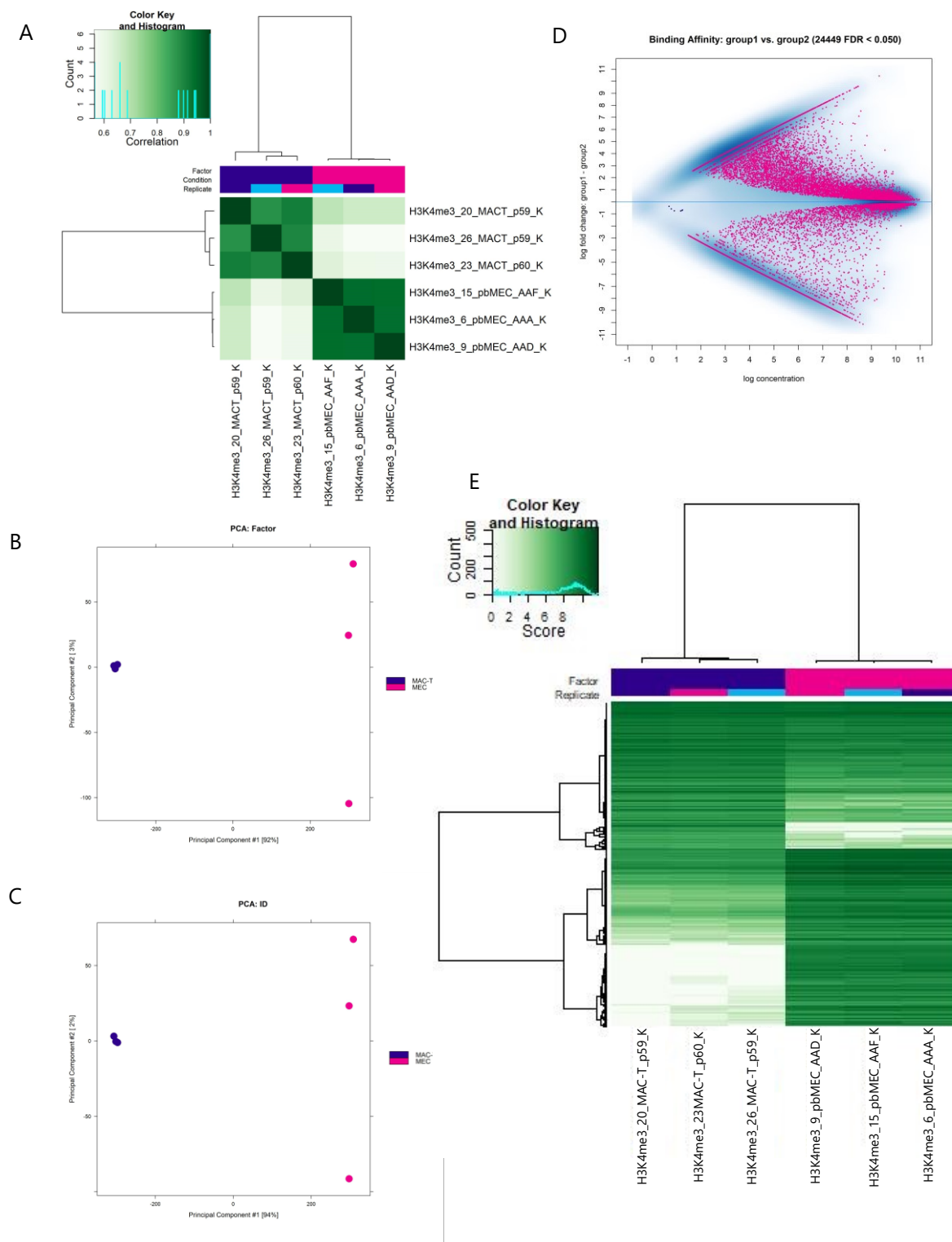


Figure 3-26: Differential Binding Analysis for H3K4me3 peaks in MAC-T versus pbMEC in control samples. Peak summits were set at 1000 bp for counting reads. Differential binding analysis was performed with the DESeq2 method in Diffbind. Plots show a correlation heatmap (A), using occupancy (peak caller score) data, PCAs based on all binding sites (B) and on differentially bound sites (C), an MA plot (D) showing significantly differentially bound sites in red, and a heatmap (E) showing cluster based on differentially bound sites per cell type.

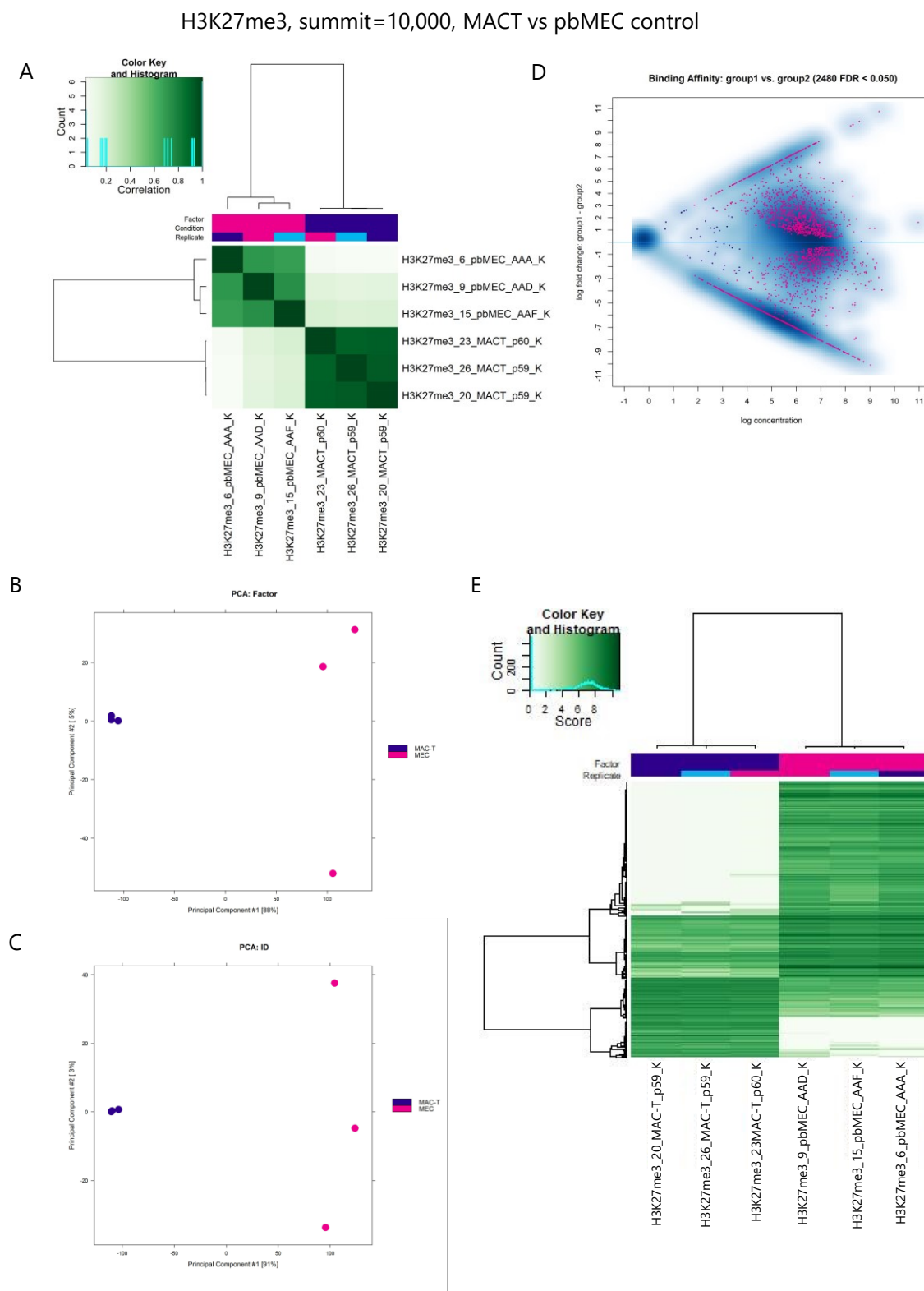


Figure 3-27: Differential Binding Analysis for H3K27me3 peaks in MAC-T versus pbMEC in control samples. Peak summits were set at 1000 bp for counting reads. Differential binding analysis was performed with the DESeq2 method in Diffbind. Plots show a correlation heatmap (A), using occupancy (peak caller score) data, PCAs based on all binding sites (B) and differentially bound sites (C), an MA plot (D) showing significantly differentially bound sites in red, and a heatmap (E) showing cluster based on differentially bound sites per cell type.

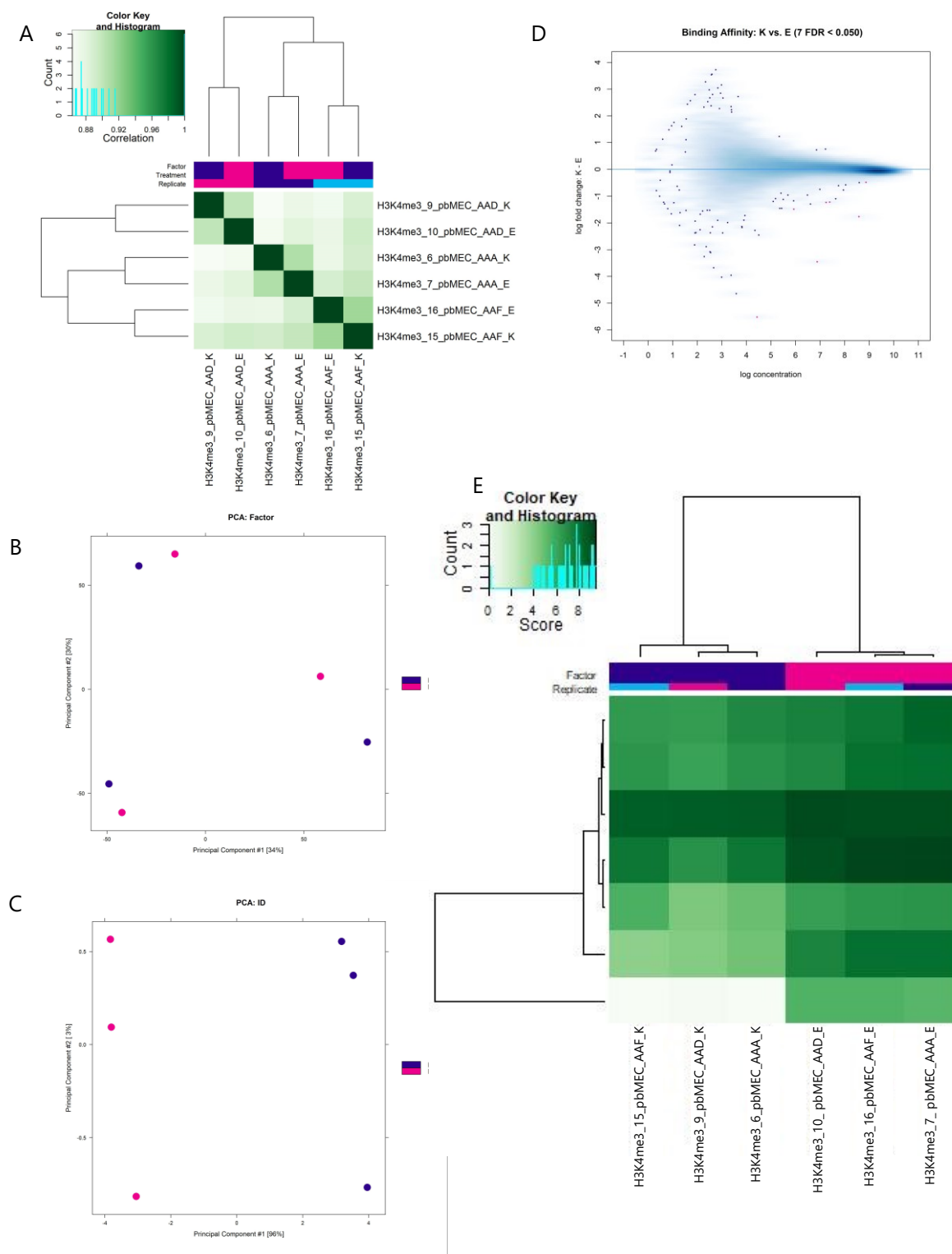
H3K4me3, summit=500, pbMEC *E.coli* vs control

Figure 3-28: Differential Binding Analysis for H3K4me3 peaks in pbMEC comparing *E. coli*<sub>1303</sub> and control samples. Here, peak summits were set at 500 bp for counting reads. Differential binding analysis was performed with the DESeq2 method in Diffbind. Plots show a correlation heatmap (A), using occupancy (peak caller score) data, PCAs based on all binding sites (B) and differentially bound sites (C), an MA plot (D) showing significantly differentially bound sites in red, and a heatmap (E) showing cluster based on differentially bound sites per cell type.

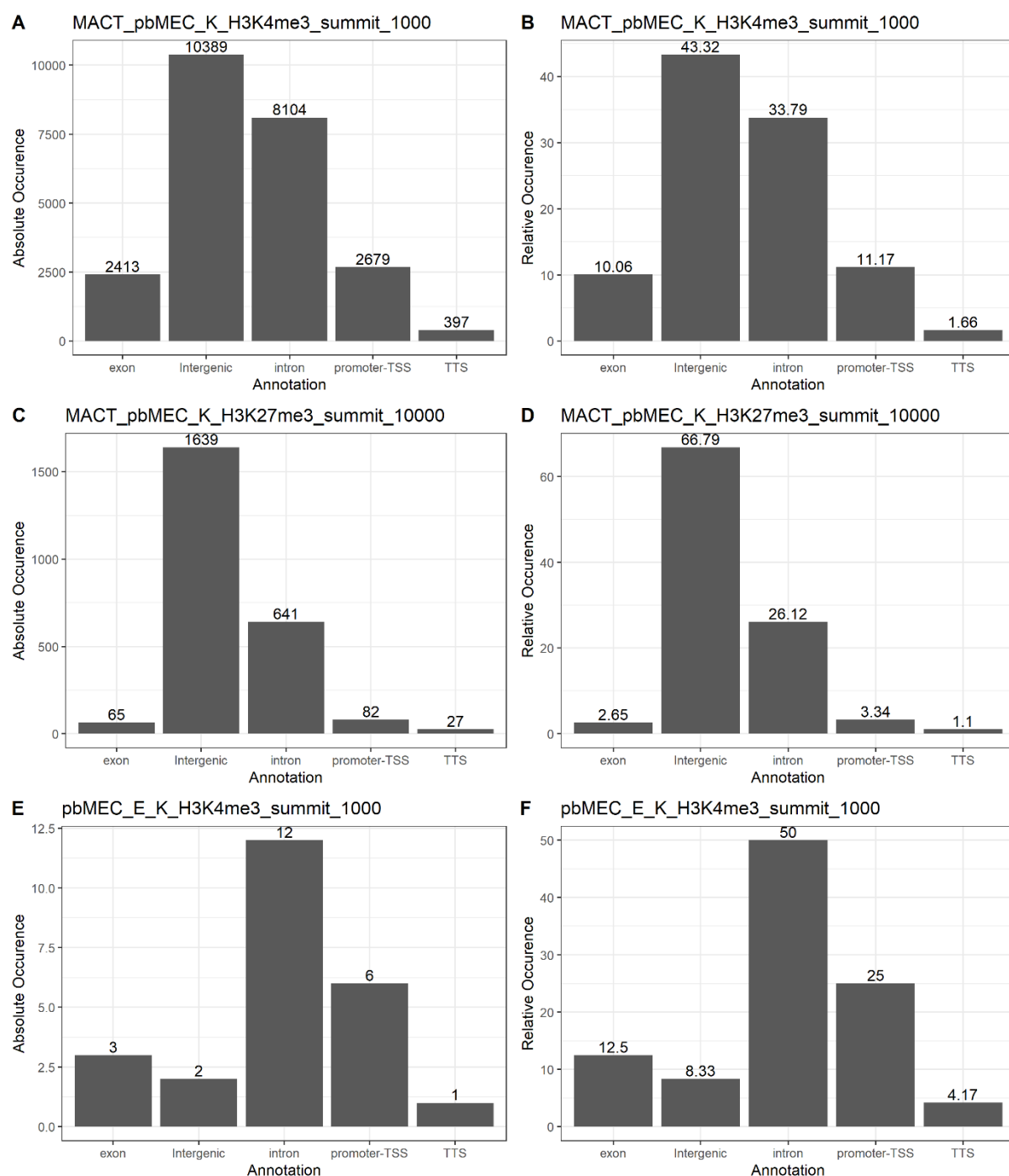


Figure 3-29: ChIP-seq peak annotation of H3K4me3 and H3K27me3 in MAC-T and pbMEC. Bar charts show absolute and relative occurrence in genomic regions of identified differential binding sites.

### 3.4 Data Integration of ChIP-seq and RNA-seq

Based on data integration, it was tested whether the identified gene expression patterns from RNA-seq overlapped with the binding sites from ChIP-seq of active or repressive histone marks, H3K4me3 and H3K27me3, respectively. The colocalization data set was then compared to reference databases to identify the biological functional annotation. First, the transcriptome and epigenome data of control-

challenged cells were integrated. Secondly, the data sets of *E. coli*<sub>1303</sub>-challenged pbMEC were analyzed. These findings will be presented in the following pages.

### 3.4.1 *MAC-T versus pbMEC control*

First, it was found that several different binding sites were associated with the same gene for both histone marks, H3K4me3 and H3K27me3, in the data set comparing the two MEC models. Consequently, these lists have been reduced to unique identifiers for integrating the data sets as listed by number in Table 3-2. The set of RNA-seq data with a gene count of 5794 was obtained by applying the significance filter  $\text{padj} < 0.05$  to the original data set, without taking the log2FC filter into account.

Table 3-2: Data integration of ChIP-seq and RNA-seq for comparing MAC-T and pbMEC.

List names	Number of elements	Number of unique elements
ChIP-seq H3K27me3	2465	752
ChIP-seq H3K4me3	23982	13694
RNA-seq	5794	5794
The overall number of unique elements		15561

As a result of the data integration, the Venn diagram in Figure 3-30A shows that the data sets included in the analysis have clear intersections. Of 13,694 H3K4me3 binding sites, 3999 genes overlapped with DEGs from RNA-seq, corresponding to ~30% consensus. The 118 genes associated with H3K4me3 binding sites overlapped with the repressive histone mark H3K27me3 and DEGs from RNA-seq. Thus, these differentially expressed genes were subject to dual histone regulation. Moreover, 22 genes associated with the repressive histone mark H3K27me3 overlapped with the identified DEGs. Notably, 422 genes associated with H3K4me3 binding sites overlapped with the repressive histone marks only, without overlapping with DEGs.

When comparing RNA-seq with ChIP-seq, it was shown that genes that show near a peak for H3K4m3, H3K27me3, or with both histone marks in control-challenged MEC, the fold changes of these genes are manifold, either positive or negative. This result is illustrated in the violin plot (Figure 3-30B), showing the fold enrichment distribution of the DEGs under single or dual histone regulation. A positive fold change indicated a higher expression of genes in MAC-T, while a negative fold change indicated a reduced expression of these genes in MAC-T.

Figure 3-31 shows an example of the expression profile and histone profiles in the different sample groups for the *OLRI* gene, which encodes oxidized low-density lipoprotein receptor 1 (lectin-like). These results illustrate the different responses of the cell models at the molecular level by the disparity of MAC-T and pbMEC profiles for each histone mark.

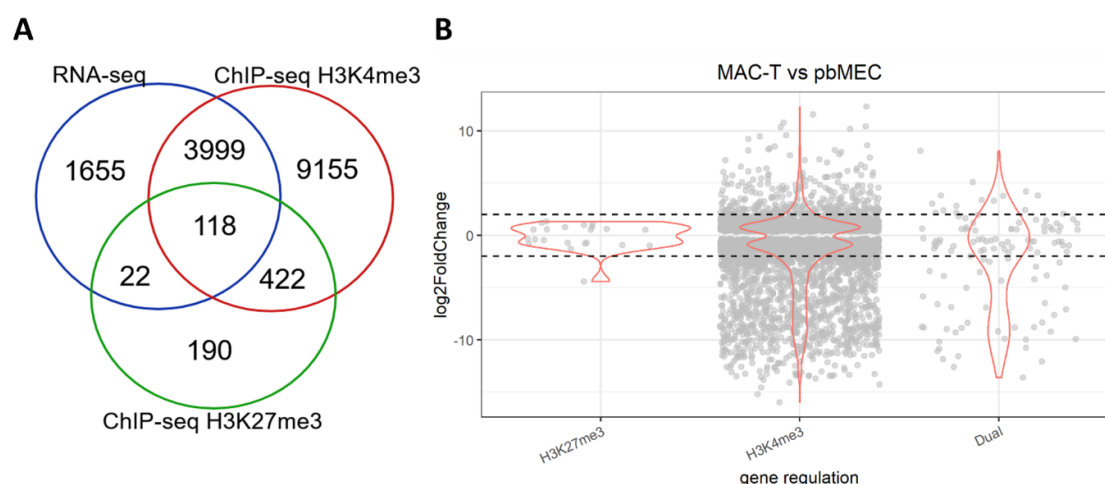


Figure 3-30: RNA-seq and ChIP-seq data integration for MAC-T versus pbMEC control samples. A: The overlap in the Venn is based on Ensembl Gene IDs (ENSBTA). B: The violin plot shows the subsets of genes regulated by histone marks and how their fold change was in RNA-seq.

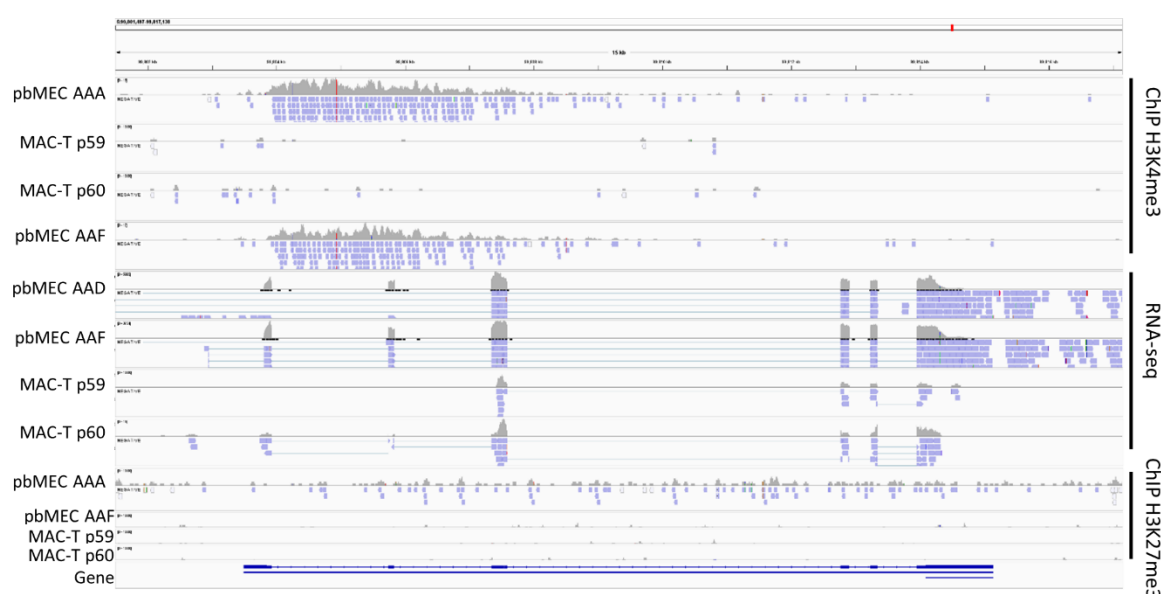


Figure 3-31: Example of histone modifications detected near *OLR1* gene. The Integrative Genomics Viewer (IGV) snapshot visualizes the genomic and epigenomic data from the experiment over a range of 15 kilobases (kb). The top four rows show H3K4me3 peaks for pbMEC AAA, MAC-T p59, MAC-T p60, and pbMEC AAF, all control challenged. The following four rows show read coverage determined with *hsat* during RNA-seq data analysis in order of pbMEC AAA, AAF, MAC-T 1(p59), and MAC-T 2(p60). The last four rows show the H3K27me3 peak distribution for pbMEC AAA, AAF, MAC-T p59, and MAC-T p60. The bottom row shows the gene structure.

The identified gene intersections were used for functional analysis. The GO analysis for BP revealed that the 118 genes associated with both histone marks might play an important role in cellular development, epithelial cell development, epithelial cell differentiation, and anatomical structure development (Figure 3-32). The genes associated with the activating histone mark are enriched in KEGG pathways such as multicellular development, focal adhesion, PI3K-Akt signaling, MAPK signaling, regulation of actin cytoskeleton, TNF signaling, and Ras signaling (Figure 3-33). Given the considerable input of 3999 genes, the output of functional elements is the highest of the overall results (Figure 3-33). Organelle transport along microtubules was mainly enriched with the 22 genes under the control of the

repressive histone mark (Figure 3-34). In contrast, the 422 genes associated with both histone marks without overlap with DEGs were associated with GO terms such as neurogenesis, synaptic signaling, and the KEGG pathway dopaminergic synapse (Figure 3-35).

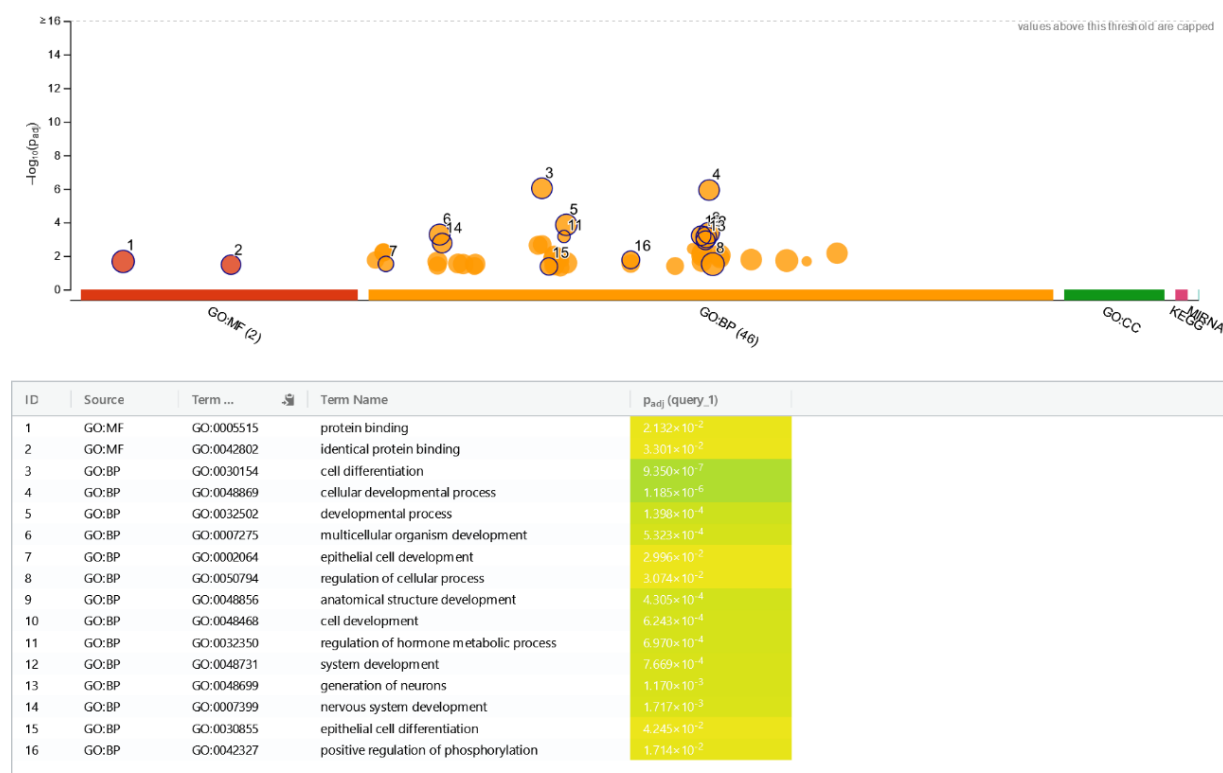


Figure 3-32: Functional analysis for 118 genes under epigenetic regulation. The overlapping genes identified with RNA-seq match with H3K4me3 and H3K27me3 binding sites. The gene set was used for functional analysis with g:Profiler under the following reference <https://biit.cs.ut.ee/gplink/l/71b2cfZaSx>



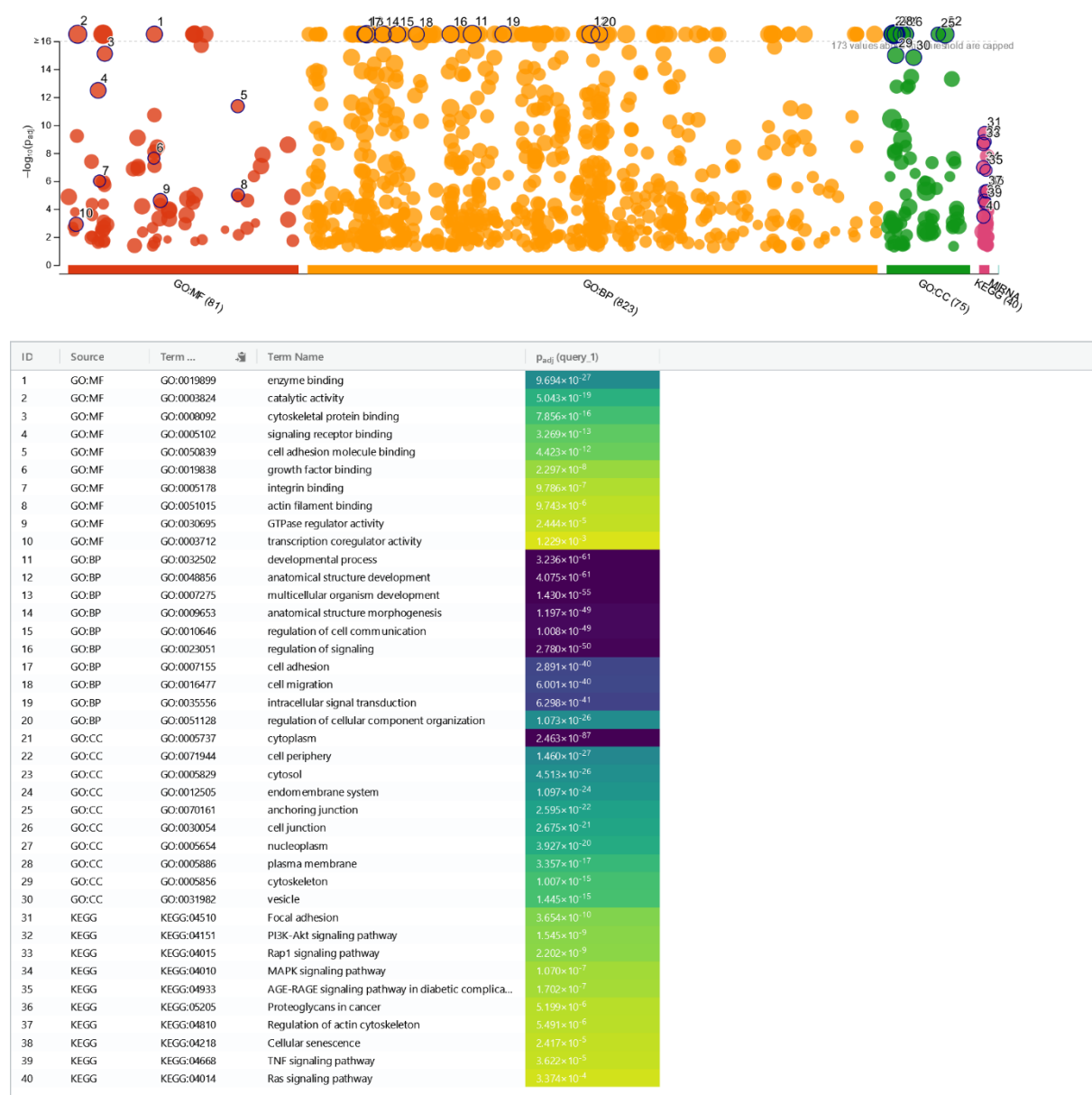
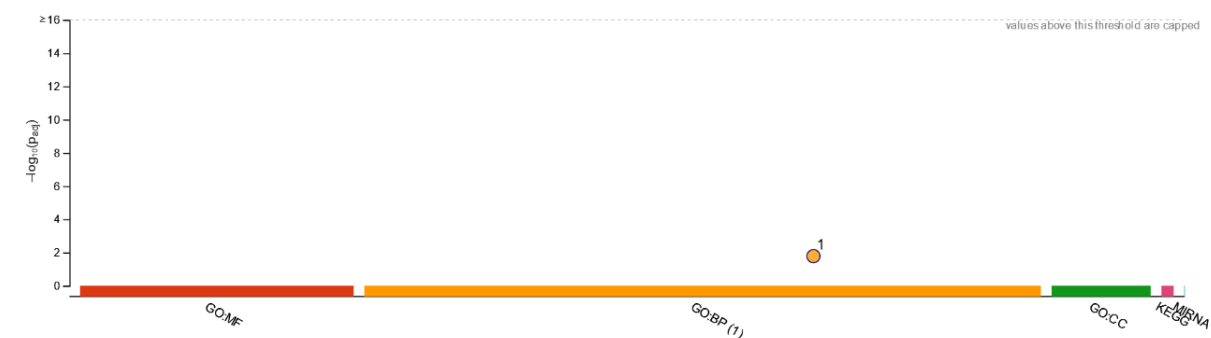
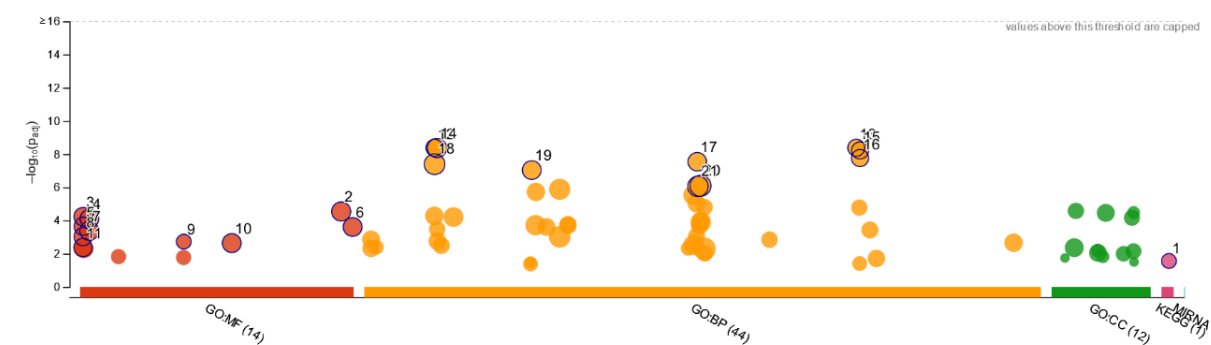


Figure 3-33: Functional analysis for 3999 genes under epigenetic regulation. The overlapping genes identified with RNA-seq match with H3K4me3 binding sites. The gene set was used for functional analysis with g:Profiler under the following reference <https://biit.cs.ut.ee/gplink//HOn4NGu1Q0>.



ID	Source	Term ...	Term Name	Padj (query_1)
1	GO:BP	GO:0072384	organelle transport along microtubule	1.650 × 10 <sup>-2</sup>

Figure 3-34: Functional analysis for 22 genes under epigenetic regulation. The overlapping genes identified with RNA-seq match with H3K27me3 binding sites. The gene set was used for functional analysis with g:Profiler under the following reference <https://biit.cs.ut.ee/gplink/l/K1-XUrtmTe>.



ID	Source	Term ...	Term Name	Padj (query_1)
1	KEGG	KEGG:04728	Dopaminergic synapse	2.740 × 10 <sup>-2</sup>
2	GO:MF	GO:0140110	transcription regulator activity	2.888 × 10 <sup>-5</sup>
3	GO:MF	GO:0000981	DNA-binding transcription factor activity, RNA pol...	5.864 × 10 <sup>-5</sup>
4	GO:MF	GO:0003700	DNA-binding transcription factor activity	7.840 × 10 <sup>-5</sup>
5	GO:MF	GO:0000987	cis-regulatory region sequence-specific DNA bindi...	2.228 × 10 <sup>-4</sup>
6	GO:MF	GO:1990837	sequence-specific double-stranded DNA binding	2.457 × 10 <sup>-4</sup>
7	GO:MF	GO:0003690	double-stranded DNA binding	4.241 × 10 <sup>-4</sup>
8	GO:MF	GO:0000978	RNA polymerase II cis-regulatory region sequence...	9.381 × 10 <sup>-4</sup>
9	GO:MF	GO:0022843	voltage-gated cation channel activity	1.878 × 10 <sup>-3</sup>
10	GO:MF	GO:0043565	sequence-specific DNA binding	2.298 × 10 <sup>-3</sup>
11	GO:MF	GO:0000977	RNA polymerase II transcription regulatory region ...	4.211 × 10 <sup>-3</sup>
12	GO:BP	GO:0007268	chemical synaptic transmission	4.165 × 10 <sup>-9</sup>
13	GO:BP	GO:0098916	anterograde trans-synaptic signalling	4.165 × 10 <sup>-9</sup>
14	GO:BP	GO:0007399	nervous system development	4.288 × 10 <sup>-9</sup>
15	GO:BP	GO:009537	trans-synaptic signaling	6.005 × 10 <sup>-9</sup>
16	GO:BP	GO:009536	synaptic signaling	1.731 × 10 <sup>-8</sup>
17	GO:BP	GO:0048699	generation of neurons	2.878 × 10 <sup>-8</sup>
18	GO:BP	GO:0007275	multicellular organism development	4.058 × 10 <sup>-8</sup>
19	GO:BP	GO:0022008	neurogenesis	9.159 × 10 <sup>-8</sup>
20	GO:BP	GO:0048856	anatomical structure development	8.091 × 10 <sup>-7</sup>
21	GO:BP	GO:0048731	system development	8.829 × 10 <sup>-7</sup>

Figure 3-35: Functional analysis for 422 genes under epigenetic regulation. The overlapping genes were identified as match from binding sites for both histone marks and are thus dual-regulated. The gene set was used for functional analysis with g:Profiler under the following reference <https://biit.cs.ut.ee/gplink/l/RVLbsd5yTf>.

### 3.4.2 pbMEC *E. coli* versus control

Given that cells are thought to use a variety of mechanisms to regulate gene expression, these may change under different cellular conditions and be environmentally influenced. Other than for the analysis

above (section 3.4.1), there were no duplicate IDs in the differential histone binding site data for pbMEC comparing *E. coli* and the control (Table 3-3).

Table 3-3: Data integration of ChIP-seq and RNA-seq for comparing *E. coli*<sub>1303</sub> versus control-challenged pbMEC.

List names	Number of elements	Number of unique elements
ChIP-seq H3K4me3	24	24
RNA-seq	409	409
The overall number of unique elements		418

The results for the intersection of RNA-seq and ChIP-seq data are presented in the form of a Venn diagram. Thus, the diagram (Figure 3-36A) illustrates the consensus of 15 genes of the 409 DEGs from RNA-seq with the 24 differential binding sites identified for the activating histone mark H3K4me3 when comparing *E. coli*<sub>1303</sub> to the control sample in pbMEC. Genes located near a peak for H3K4m3 in pbMEC challenged with *E. coli*<sub>1303</sub> were found upregulated in their gene expression detected by RNA-seq. The violin plot in Figure 3-36B depicts this in a positive fold change, indicating that *E. coli* upregulated gene expression. Table 3-4 lists all Ensembl IDs, gene abbreviations, and descriptions for the 15 genes regulated by the active histone mark, of which some are reoccurring candidates like *CYP1A1* and *CYP1B1*. Functional analysis, including the subset of the 15 genes, revealed involvement in pathways like tryptophan metabolism and receptor activation (Figure 3-37). From the data intersection, 9 genes emerged being differential binding sites for the active histone mark only. These 9 genes are listed by name in Table 3-5. No significant gene ontology annotation was determined for these nine binding sites, given the small input data.

As identified previously (section 3.2.3), the identified differentially expressed genes from RNA-seq are involved in multiple signaling pathways in the cell. Here, functional annotation analysis for the 394 genes revealed that these were associated with the TNF pathway, NF- $\kappa$  B signaling, MAPK-and PI3K-Akt signaling, microRNAs in cancer, and NOD-like receptor signaling (Figure 3-38).

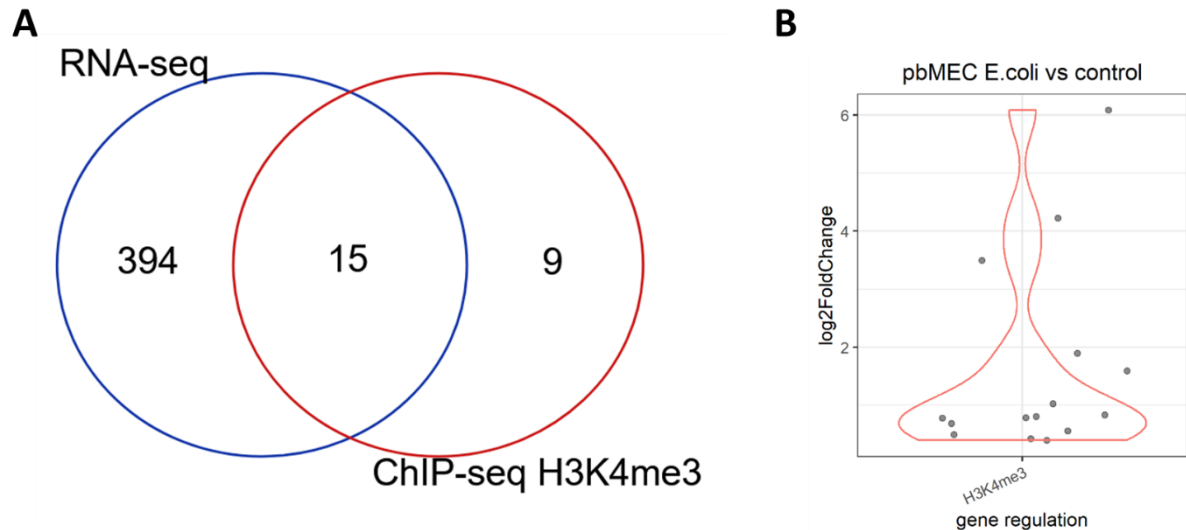


Figure 3-36: Data integration of RNA-seq and ChIP-seq for comparing *E. coli*<sub>1303</sub> versus control challenged pbMEC. A: The overlap in the Venn is based on Ensembl Gene IDs (ENSBTA). B: The violin plot shows the subsets of genes regulated by the active histone mark H3K4me3 and how the fold change was in RNA-seq.

Table 3-4: Gene IDs for 15 common genes under epigenetic regulation. The ENSBTA conversion was performed with g:Profiler under the following reference: <https://biit.cs.ut.ee/gplink/l/JYr3nxh4TT>.

initial_alias	Name	Description
ENSBTAG00000005431	LMCD1	LIM and cysteine rich domains 1 [Source:VGNC Symbol;Acc:VGNC:30926]
ENSBTAG00000010531	CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1 [Source:NCBI gene (formerly Entrezgene);Acc:511470]
ENSBTAG00000001294	PPP1R15A	protein phosphatase 1 regulatory subunit 15A [Source:VGNC Symbol;Acc:VGNC:33228]
ENSBTAG00000004052	EPGN	epithelial mitogen [Source:VGNC Symbol;Acc:VGNC:53711]
ENSBTAG00000020355	KLF4	Kruppel like factor 4 [Source:VGNC Symbol;Acc:VGNC:52790]
ENSBTAG00000017263	MXI1	MAX interactor 1, dimerization protein [Source:VGNC Symbol;Acc:VGNC:31773]
ENSBTAG00000013912	TXNRD1	thioredoxin reductase 1 [Source:VGNC Symbol;Acc:VGNC:36545]
ENSBTAG00000016519	MBOAT1	membrane-bound O-acyltransferase domain containing 1 [Source:VGNC Symbol;Acc:VGNC:31286]
ENSBTAG00000009513	TGFB1	transforming growth factor beta-induced [Source:VGNC Symbol;Acc:VGNC:52886]
ENSBTAG00000000245	NHSL1	NHS like 1 [Source:VGNC Symbol;Acc:VGNC:32072]
ENSBTAG00000045500	SGSM2	small G protein signaling modulator 2 [Source:VGNC Symbol;Acc:VGNC:34551]
ENSBTAG00000021717	BDKRB2	bradykinin receptor B2 [Source:VGNC Symbol;Acc:VGNC:26460]
ENSBTAG00000001021	CYP1A1	cytochrome P450 family 1 subfamily A member 1 [Source:HGNC Symbol;Acc:HGNC:2595]
ENSBTAG00000009997	ANO1	anoctamin 1 [Source:VGNC Symbol;Acc:VGNC:50077]
ENSBTAG00000021499	PSAP	prosaposin [Source:VGNC Symbol;Acc:VGNC:33431]

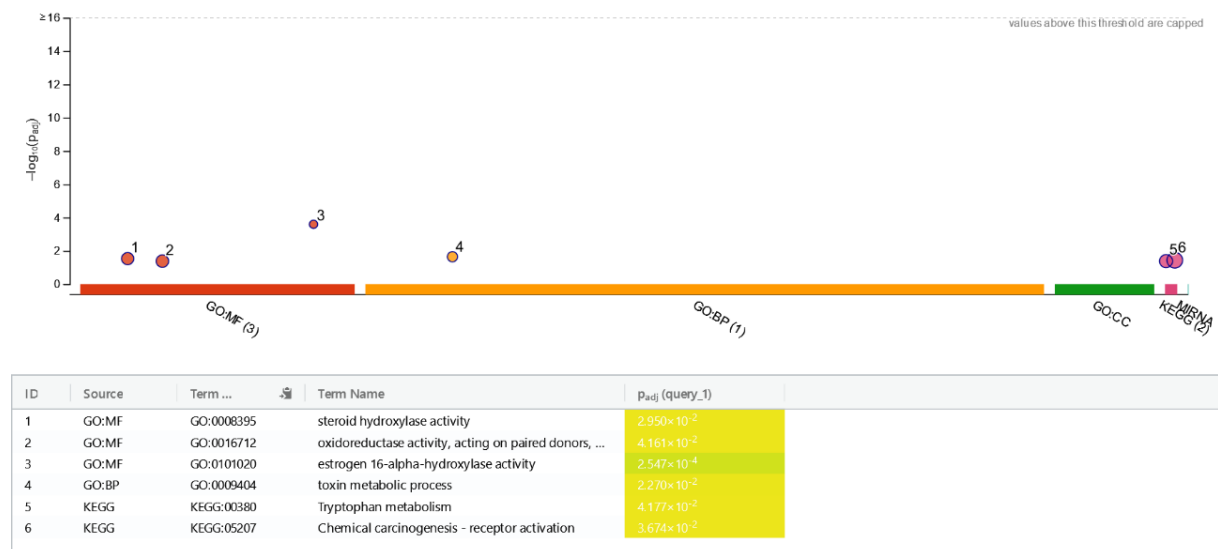


Figure 3-37: Functional analysis for 15 genes under epigenetic regulation. The overlapping genes were identified as match from binding sites for histone mark H3K4me3 and RNA-seq. The gene set was used for functional analysis with g:Profiler under the following reference: <https://biit.cs.ut.ee/gplink/I/JYr3nxh4TT>

Table 3-5: Gene IDs for 9 genes under epigenetic regulation.

initial_alias	Name	Description
ENSBTAG00000009226	SCUBE3	signal peptide, CUB domain and EGF-like domain containing 3 [Source:VGNC Symbol;Acc:VGNC:34373]
ENSBTAG00000050129	ENSBTAG00000050129	None
ENSBTAG00000019203	S100A4	S100 calcium-binding protein A4 [Source:VGNC Symbol;Acc:VGNC:34244]
ENSBTAG00000014614	ACTA2	actin alpha 2, smooth muscle [Source:VGNC Symbol;Acc:VGNC:106627]
ENSBTAG00000052375	AHRR	aryl-hydrocarbon receptor repressor [Source:NCBI gene (formerly Entrezgene);Acc:529500]
ENSBTAG00000052600	ENSBTAG00000052600	None
ENSBTAG00000008410	OVOL2	ovo like zinc finger 2 [Source:VGNC Symbol;Acc:VGNC:32506]
ENSBTAG00000021024	MACF1	microtubule actin cross-linking factor 1 [Source:VGNC Symbol;Acc:VGNC:55610]
ENSBTAG00000033680	CEP112	centrosomal protein 112 [Source:VGNC Symbol;Acc:VGNC:27190]

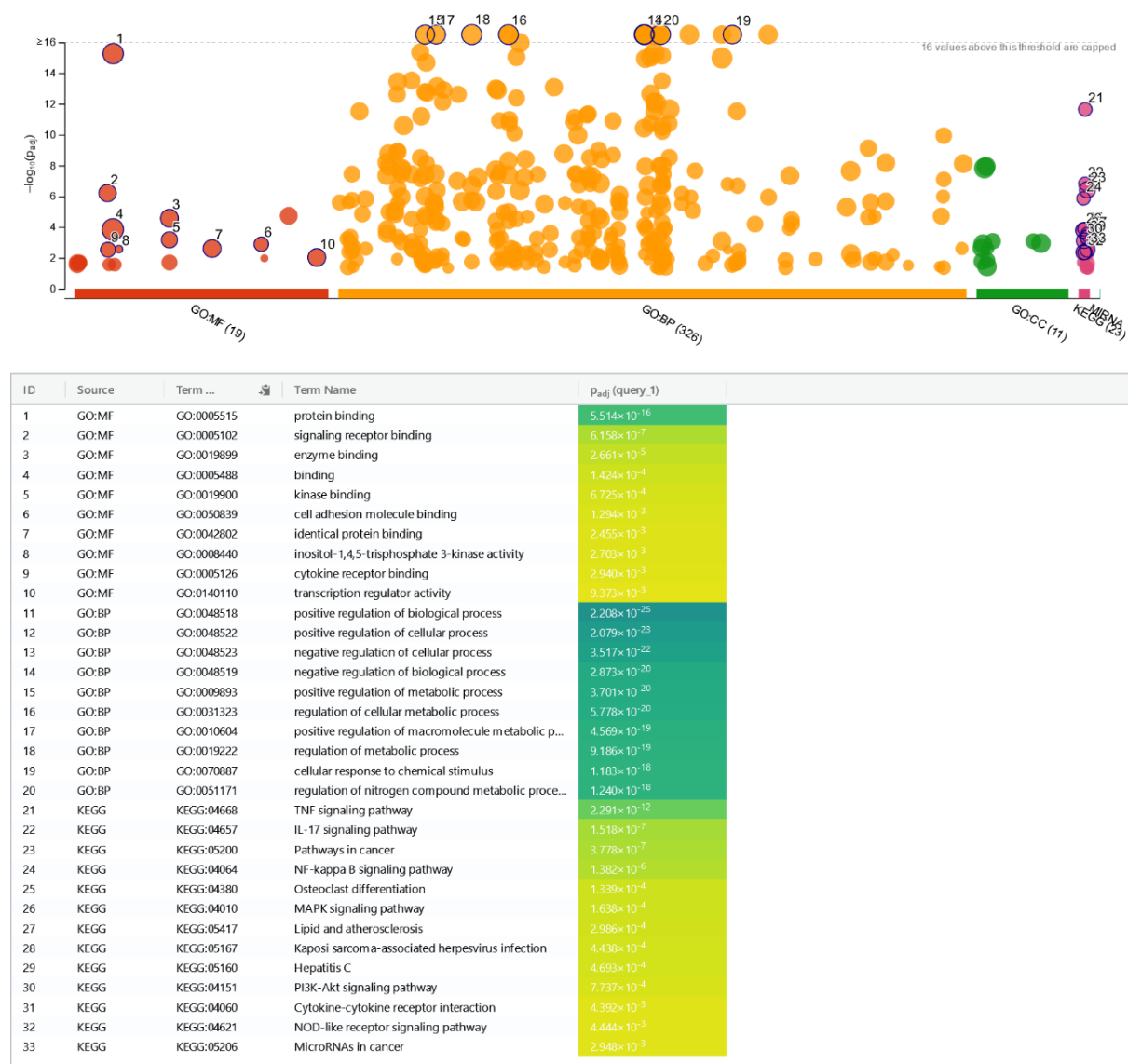


Figure 3-38: Functional analysis for 394 genes from RNA-seq. The genes were identified as differentially expressed in pbMEC, comparing *E. coli*<sub>1303</sub> versus control. The gene set was used for functional analysis with g:Profiler under the following reference: <https://biit.cs.ut.ee/gplink/l/94TNBmF4R6> <https://biit.cs.ut.ee/gplink/l/ndfrDp09TR>

## 4 Discussion

The complexity of host-pathogen interactions in bovine mastitis has yet to be fully elucidated. This *in vitro* study aimed to investigate the involvement of epigenetic mechanisms in the form of histone modifications and their role in changes in gene expression by mastitis pathogens (1.) in targeted immune response genes, (2.) at the genome-wide level, and (3.) their correlation and predicted function in bovine cell models. *In vitro* cell models were used to establish ChIP and allowed the identification of specific histone modifications that are thought to influence cellular gene selection, response, and signal transduction. This chapter addresses nuances in optimizing the methodological approaches used in this study in bovine mammary epithelial cells. Subsequently, the transcriptome and epigenome results of the analysis presented in the previous chapter are discussed.

### 4.1 Methodological Aspect of ChIP

Superficially, it seems trivial, but efforts to establish chromatin immunoprecipitation (ChIP) for relatively unexplored specific cell lines or tissues required rigorous optimization and adaptation experiments. In the first instance, the individual steps of the ChIP procedure for the bovine MEC models had to be optimized to study the histone modifications. N-ChIP in MAC-T was started before including pbMEC in X-ChIP.

#### 4.1.1 Fixation, Douncing, and Shearing

For the experimental setup, both ChIP approaches were tested. Although the histone examination is also possible in the native chromatin state, as stable interactions exist between DNA and histone protein<sup>221</sup>, studies point out the limitations of the N-ChIP procedure. Chances exist that rearrangements occur during processing, that excessive concentration of nucleases interferes with DNA-protein interaction, and that conservation of non-histone proteins may fail<sup>221–223</sup>. In contrast to N-ChIP, X-ChIP can use lower amounts of samples and antibodies, and chromatin rearrangements are unlikely<sup>221</sup>. For X-ChIP, the bovine mammary epithelial cells have been fixed with formaldehyde as the cross-linking reagent of the first choice<sup>224</sup>. The use of secondary membrane-permeable cross-linking reagents was excluded in the current study. To preserve the association of proteins to DNA, the timing of fixation is crucial<sup>225</sup>. One limitation of the approach used in the presented study is the lack of testing for several fixation times. Fixation for 15 min was chosen because an extended cross-linking generally influences subsequent shearing outcomes or might introduce cross-linking artefacts<sup>226</sup>. Previous studies reported that formaldehyde-fixed epithelial cells possess strong adhesive characteristics, leading to cell sheet formation and thus affecting the quality of DNA for ChIP<sup>227</sup>. Unlike previous studies<sup>227,228</sup>, trypsinization of cells was omitted in the protocol during the harvesting step. Instead, a douncing step before sonication was included, eliminating still-existing cell sheets, shearing the cell membrane, and liberating the nuclei<sup>179</sup>. This step is an addition to the chemical lysis step.

The formaldehyde cross-linking is entirely reversible at high temperatures in an aqueous solution<sup>224</sup>. Therefore, strict temperature control is required to obtain high-quality sheared chromatin with sonication<sup>173</sup>. There are two commonly used devices: the Bioruptor (Diagenode) and the Covaris system,

in which DNA is fragmented by physical shearing. Both approaches seem to work equally, while the first one has been used in this study. The determination of DNA fragment sizes followed the testing of sonication cycles. The applied trial-and-error approach focused on sonication setting adjustment with reaching a sufficient proportion of small fragments, though large fragments may still be present. All conditions set in this step also include volumes and cell concentration that are kept at a constant level to ensure the reproducibility of the results. Large chromatin fragments (above 800 bp) can interfere with chromatin quality<sup>132</sup>. Therefore, adjustment of the sonication conditions by adding more pulses is recommended<sup>229</sup>. Chromatin that is condensed is less efficiently sonicated than accessible chromatin regions<sup>229</sup>. At the same time, over-shearing may harm ChIP-seq as it causes protein denaturation<sup>173</sup>. The small fragment portion should have an average chromatin length of 200 to 500 bp<sup>229</sup>, whereas the total enrichment of DNA fragments from sonication can result in a range of 100 to 1000 bp<sup>173</sup>. O'Geen *et al.* (2011) pointed out that the smaller portion of sonicated chromatin (200 to 400 bp) contained more promoter fragments and gave the highest enrichment for histone marks of open chromatin<sup>229</sup>. In addition, the same study found that larger fragment selection for library preparation (400 to 600 bp) resulted in increased enrichment for histone marks of condensed chromatin<sup>229</sup>. This aspect could be worth considering in future studies.

#### 4.1.2 Choice of Antibody

Since histone modifications are likely to be evolutionarily conserved between species<sup>230</sup>, it was assumed that the antibodies could also be used in cattle, despite their development in rabbits. The results obtained in the presented study agree that antibody selection considerably influences the outcome of a ChIP experiment. The findings suggest that either not all antibodies worked for the N-ChIP experiments or that the concentrations were not well chosen when introducing the method. It is established that polyclonal antibodies with affinity for multiple epitopes of the target protein appear to perform better. For X-ChIP, high-quality polyclonal antibodies were used, which were extensively validated and confirmed for their specificity and high level of performance in the ChIP experiment in human HeLa cells<sup>231,232</sup>. Nowadays, most suppliers of epigenetic reagents provide species spectrum information to demonstrate the broad applicability of their antibodies. Since the development of pipetting robots and the affordability of NGS, the development and applicability of reagents for a broad species spectrum have advanced recently. This information was only partially available when this research study began.

#### 4.1.3 X-ChIP-qPCR

More quality checks have been included with the switch from N-ChIP to X-ChIP. X-ChIP DNA from MAC-T and pbMEC experiments was verified by qPCR to assess the quality of the method before sequencing. The principle is based on the fact that the positive genomic region is enriched for the histone mark of interest. At the same time, the negative genomic region should not be enriched and thus represents the background signal of the ChIP experiment. Specificity is revealed by high target DNA signal to low background signal (fold enrichment  $\geq 5$ )<sup>173</sup>. At least one positive and negative control primer set was developed for each histone mark (H3K4me3, H3K27me3, H3K4me1, H3K27Ac) and



the transcription factor CTCF, which are the central five ChIP assays in the FAANG community. *Actin* or *GAPDH*, which are usually constantly expressed within a cell, served as a positive control for histone marks associated with active chromatin and a negative control for inactive chromatin. Cell-specifically expressed genes *PRSSI* and *TSH2B*, expressed in pancreas cells and germ cells in testis<sup>233,234</sup>, respectively, were used as negative controls for active chromatin and positive control for inactive chromatin.

The calculation of the ChIP-qPCR results can vary and will be affected by differences in volumes and dilutions during the assay. One reliable ChIP-qPCR normalization method is the normalization to the input DNA signal, which is principally applied and corrects technical deviations, such as precipitation efficiency and variation in the recovery of DNA<sup>129</sup>. The constant volume input method was used instead of the recently postulated constant quantity input method<sup>235</sup> to normalize ChIP-qPCR. The % Input varies depending on the chosen histone mark and the location of the selected target primer sets<sup>229</sup>, which is also evident in the observations made in this study for X-ChIP-qPCR. Our results for H3K4me3 ranged above the % Input level of ChIP assays performed with HeLa cells<sup>231</sup>. The % Input for both MEC models assayed for H3K27me3 was below 1%, which differs from previously published data<sup>231</sup>. However, these measurements have not fallen below the limit for the relative enrichment of good quality ChIP DNA of  $\geq 5$ <sup>173</sup>. The targeted genomic regions are specific to the cell types. However, the MAC-T cell line differs from primary cells because it does not map everything one-to-one; some control regions differ. Based on the availability of ChIP-seq data produced within this study for both bovine epithelial cell models, other regions can be considered and confirmed with X-ChIP-qPCR for future studies.

## 4.2 Effect of Pathogen Challenge on Immune Response Genes

The interaction of bacteria and bovine epithelial cells and the resulting molecular mechanisms of innate immune response were measured with qPCR for a set of targeted immune genes, including cytokines, chemokines, and antimicrobial peptides. Those genes have been measured analogously to *in vivo* and *in vitro* studies, validating the response to mastitis pathogens. Time points considered for the qPCR experiment were 0, 1, and 3 h of pathogen exposure. Summarizing the qPCR data, it is evident that *E. coli*<sub>1303</sub> challenge of MAC-T upregulates gene expression for *CCL20*, *IL8*, *IL6*, and *TNF- $\alpha$*  at a higher level than *S. aureus*<sub>1027</sub> challenge, which ties in with previous results *in vivo* and *in vitro*<sup>56,76,236–238</sup>. Epithelial cells are involved in *E. coli*-specific secretion of chemokines and pro-inflammatory cytokines required for immune effector cell recruitment. *In vivo* results suggest *S. aureus* triggers short-term immunosuppression instead of inflammation and invasion of host epithelial cells<sup>102</sup>. Thus, *S. aureus* contributes to a microenvironment that enables its persistence instead of elimination.

The observations made in this study show that repeating the experiment with the *E. coli* challenge led to less significant results but the same trend of gene upregulation. This finding may have been introduced during data collection as the qPCR machines were switched from a carousel to a plate approach during the study. From unpublished data, it is apparent that MAC-T responsiveness to pathogens varies with

the passage the cells are cultured in (personal communication with Prof. Christa Kühn), which suggests a limitation of using MAC-T as a cell model.

The epigenetic regulator *JMJD3*, known as histone H3 lysine 27 (H3K27) demethylases, was not induced in response to mastitis pathogens in MAC-T. These results contradict *JMJD3* expression induced in macrophages<sup>239,240</sup> and vascular endothelial cells<sup>241</sup> 2 h after LPS exposure. *JMJD3* has been instrumental in the NF- $\kappa$ B pathway in inflammatory diseases<sup>239,241</sup> and is, therefore, suspected to be involved in molecular mechanisms during mastitis.

Another limitation of MAC-T as a cell model that needs to be stressed is the lack of expression of bactericidal genes *LAP* and *NOS2A*. Other udder-derived cell models express these genes<sup>56,74</sup>. However, *LAP* expression has been shown to decrease when pbMECs are sub-cultured, suggesting limited use for infection studies<sup>56</sup>. In contrast, induction of *LAP* and udder-like immunological gene regulation could also be mimicked in isolated perfused bovine udders<sup>242</sup>, providing another alternative to *in vivo* mastitis experiments.

The Holstein breed is one of the most widely used breeds in the world because of its high milk yield, so for the thesis study, primary cells were obtained from three Holstein cows. One aspect of biological variability of primary cells emerges from bovine cell preparations<sup>243</sup>. It was postulated that one way to overcome this factor is to use the established cell line MAC-T, thus excluding variability through the animal from which cells have been extracted for primary culture. This example highlights the justifiable controversy over the most appropriate model. Other factors leading to variability in primary cells are the isolation of cells from tissue or milk and notably from different cattle breeds.

Besides the targeted qPCR, whole-genome RNA-seq was performed to verify immune gene regulation and detect other patterns and trends in gene expression in both cell models, MAC-T and pbMEC. The RNA-seq results showed no or weak changes of expression of those immune genes targeted with qPCR after pathogen challenge. Possibly, no transient expression changes could be detected 3 h after pathogen exposure with the NGS approach. Testing several time points within the early infection could improve the holistic picture. Another possible explanation for the apparent mismatch in immune gene regulation could be that the cells did not respond to the challenge by the mastitis pathogen as expected.

*E. coli* strain 1303 and *S. aureus* strain 1027 used in this study were initially isolated from cows with clinical mastitis<sup>39</sup>. Genome sequencing of *E. coli* strain 1303 confirmed unique virulence factors to the strain<sup>244,245</sup>. At the same time, not all *E. coli* strains cause mastitis similarly. Specific *E. coli* are associated with specified clinical manifestations of bovine mammary gland infections, depending on their virulence factors<sup>246</sup> and host factors. Cell models could also respond with varying sensitivity or depending on variable pathogen strains. A data integration study revealed that global hosts react differently to live and heat-inactivated *E. coli*<sup>247</sup>. Other studies used *E. coli* LPS or *S. aureus* culture supernatant for testing the immune response<sup>76,248</sup>. Also, a more robust inflammatory response in mammal cells has been recorded for live versus killed bacteria such as *Staphylococcus epidermidis*<sup>249</sup> and *Staphylococcus aureus*<sup>250</sup>. Although killing the bacteria reduces experimental variability and

standardizes experimental conditions, it is essential to note that living bacteria can elicit other cellular responses and use different mechanisms during host infection<sup>249</sup>. These mechanisms may include mechanical movement and invasion of the host, active action of pathogenic factors, and bacterial growth and signal transduction. Addressing this variability in mastitis pathogens was outside the scope of this study. Still, it could be interesting for epigenetic regulation of gene expression and application in animals.

### 4.3 Measuring Histone Modifications in MAC-T

Inconclusive results in the promoter regions of immune response genes emerged from the targeted N-ChIP-qPCR. In this experiment, the focus was on a 1 h challenge with pathogens. Sample data were normalized to a housekeeping gene (*actin*) since the other results, e.g., % Input, did not give interpretable results or were biased. The background signal of IgG worked poorly, even after the antibody supplier had been changed. The use of the correct amounts of input DNA was not entirely implemented because the N-ChIP method still needed to be fully refined. This observation is reflected in the N-ChIP-qPCR results in the different input bars and explains the substantial deviations (Figure 3-4 and Figure 3-5). This fact also limited the standard ChIP-qPCR normalization. Likewise, no RNA treatment for N-ChIP was included, which might have misled the ChIP procedure's estimation of nucleic acid yield. Drawing from the established knowledge gain and quality checkpoints might be possible. Specific genes identified with ChIP-seq could then be compared with N-ChIP results.

The inconclusiveness of the results also came from the limitation of replicates of the measurement in qPCR given the limitation of sample volume and secondly due to the need for more repetition of the experiment per se. The primer design could have been more optimal. In addition, the goal was to cover many promoter regions. The disadvantage was that the primers were selected too extensively and limited multiple measurements.

The research questions of whether histone modifications show an altered pattern of presence or absence in mammary epithelial cells, in MAC-T, in response to mastitis pathogens in promoter regions of immune response genes cannot be proven or disproven with the N-ChIP-qPCR. Nonetheless, the results indicated the presence of histone modifications, so X-ChIP analysis followed by sequencing was performed.

### 4.4 RNA-seq Identified a Strong Discrepancy in Commonly Used Cell Models

To my knowledge, no previous comparative RNA-seq transcriptomic study of MAC-T and pbMEC exists. The present results extend previous findings from a microarray study by Hosseini *et al.* (2013)<sup>72</sup>, in which MAC-T was compared to biopsies of mammary tissue and identified as substantially different, focusing on lactation. The transcriptome comparison analysis of MAC-T and pbMEC revealed a substantial discrepancy between the two cell models. The key genes upregulated in MAC-T included genes encoding immunoglobulin CD19, glycoprotein fibronectin one (FN1), and insulin-like growth factor binding protein two (IGFBP2). Among the key genes upregulated in pbMEC were genes coding

for the type III intermediate filament protein vimentin (VIM), cytokeratin 17 (KRT17), the C-X-C motif chemokine ligand three (CXCL3), and the caveolae associated protein one (CAVIN1). No distinct change in these genes was detectable when bacterial challenge was included in the analysis.

*CD19* is established as a lymphocyte marker, and altered levels have been associated with various tumor cells in humans<sup>251</sup>. It appears to be the case that high expression of *FNI* promotes proliferation and is also related to numerous cancer cell types in humans<sup>252–255</sup>. *IGFBP2* allows growth and survival of bovine mammary epithelial cell line MAC-T<sup>256</sup>, while in pbMEC, *IGFBP2* suggested autocrine regulation<sup>257</sup>. In the human context, *IGFBP2* was proposed as a tumor promoter<sup>258</sup>. The list of genes specifically expressed in MAC-T justifies that MAC-T has undergone a drastic change, i.e., in proliferative properties that allow immortal culture and resemble cancer cells to some extent.

*VIM* is an epithelial marker that seems to get lost with a continuous culture. As determined by other studies that used immunofluorescence and quantitative PCR, *VIM* can be found in MAC-T depending on the FBS supplementation of the culture media<sup>259</sup>. In contrast, in pbMEC, *VIM* progressively diminished with passaging<sup>259</sup>. *KRT* genes also serve as epithelial cell markers<sup>260</sup>. In the present study, VIM and KRT17 were drastically downregulated in MAC-T compared to pbMEC, with a magnitude of  $\sim 10^3$  (Figure 3-9). In general, there was a decisive change in the cellular components of the membrane and extracellular matrix that affects cellular signaling in the two cell models.

High expression levels of *CXCL3* in pbMEC were consistent with the constitutive secretion by the lactating mammary epithelium confirmed by immunohistochemistry in mammary epithelial cells and the abundance of mRNA in non-inflamed mammary tissue<sup>261</sup>. *CXCL3* is reported among a set of inflammatory cytokines and chemokines that are upregulated in response to LPS in bovine mammary epithelial cells *in vitro*<sup>236</sup> and thus are prerequisites for combating pathogens. In the present study's findings, differentially expressed genes, such as the chemokines *IL6*, *IL7*, *CXCL8*, and pattern recognition receptor NOD2, were associated with several of the top five GO terms of BP (Figure 3-13). This finding could indicate a latent pathogen challenge in the cell donor cows.

Caveolins and cavins form specific lipid raft domains in the plasma membrane, contribute to the curvature of the membrane<sup>262,263</sup>, and are involved in cellular processes, including signal transduction<sup>264</sup>. The membrane shape and bending are consistent with the nature of the cells in the alveoli of the bovine udder. Insulin, a component of the media of pbMEC, regulates *CAVIN1* to translocate to focal adhesions in adipocytes<sup>265</sup>, which might also be relevant for epithelial cells *in vitro* or *in vivo*. During human lactation, prolactin downregulates caveolae-associated proteins such as caveolin-1, whereas these are reconstituted upon weaning<sup>266</sup>. In mice mammary epithelial cells, caveolin-1 downregulation led to milk production, which was confirmed to be mediated through STAT5a signaling<sup>264</sup>. While the microanatomical function of *CAVIN1* in bovine mammary epithelial cells can only be conjectured, it is evident that low expression exists in MAC-T. Therefore, *CAVIN1* may have no role in these cells. The results show that pbMECs were very similar to the *in vivo* system in that they had epithelial markers,

which gradually reduced or disappeared in the MAC-T. However, the RNA-seq results also suggest that pbMECs exhibited strong animal individualism as a driving factor.

In the investigation, most genes with strong statistical significance were associated with cell development, cell differentiation, cellular response to stimuli, secretion, and organ morphogenesis (Figure 3-11). These terms are closely related biological processes that underscore the role of epithelial cells during lactation and mammary gland development. The physiological significance of these genes and gene products is, in part, better understood in humans than in cattle. Since both are mammals, it is reasonable to assume that the genes have corresponding functions. Among other things, both humans and cattle suffer from mastitis. However, given the anatomical differences between cattle and humans in mastitis, these conclusions cannot be extrapolated per se from one species to the other.

#### 4.5 Response to Pathogens Assessed by RNA-seq

The transcriptome comparison analysis of MAC-T and pbMEC identified many DEGs. In this special case, it was reasonable to apply a cut-off of log2FC to obtain significant essential information and reduce the number of thousand DEGs to a suitable size for downstream GO analysis. However, some regulators with fewer expression changes might get filtered out this way. In contrast, a distinctly smaller number of genes has been identified as differentially expressed in comparing bacteria and control challenges in the two cell models. A mathematical approach to identify significant DEGs in RNA-seq analysis is to control for the proportion of false positives of less than 5 % using the BH-adjusted p-value  $< 0.05$ .

The effect of pathogens on the transcriptomic level of MAC-T revealed that more genes were altered in response to *E. coli* than *S. aureus*. However, the small data set identified no significant functionalities for MAC-T. Although insignificant, the response to bacteria and intracellular signal transduction are biological processes in which most genes are enriched in MAC-T exposed to *E. coli*<sub>1303</sub> compared with the control. Interestingly, one of the genes coding for the transcription factor *CEBPB*, which has several interconnections in the cnetpot (Figure 3-15), is associated with nuclear factor  $\kappa$ B (NF- $\kappa$ B), a downstream signaling molecule in the TLR pathway and leads to the induction of *IL6*<sup>267</sup>. Thus, the results found in response to *E. coli*<sub>1303</sub> in MAC-T are biologically relevant, although not statistically significant.

More genes were identified as DEGs in the effect of the pathogen on the transcriptome level of pbMEC than in MAC-T. Cytochrome P450 family members *CYP1A1* and *CYP1B1* show recurrent involvement in response to the pathogen stimulus (Figure 3-14 and Figure 3-16). While *CYP1A1* is proven to be involved in *E. coli*-induced inflammatory processes, e.g., in macrophages<sup>268</sup> and bovine MEC<sup>269</sup>, *CYP1B1* is understudied. In a 2020 study, multiple bovine tissues were examined for ground levels of Cytochrome P450 family members; in healthy cows, the mammary gland showed the lowest expression, while the liver had the highest expression<sup>270</sup>. In the same study, ground levels of expression of oxylipid-producing *CYP450* in cultured bovine cells were measured in which MAC-T expressed *CYP1A1* and *CYP2J2*<sup>270</sup> to some extent. The authors link the metabolites of cytochrome P450 family members, such

as oxylipids, to stress and disease by affecting the metabolism of fat-soluble vitamins and thus contributing to the animal's overall health.

The GO analysis for BP highlighted that most DEGs of pbMEC challenged with *E. coli*<sub>1303</sub> compared with control were associated with response to external stimulus and cell death. Among the top five BP, the genes with large fold change were suppressor of cytokine signaling three (*SOCS3*), calgranulin B (*SI00A9*), radical S-adenosyl methionine domain-containing protein two (*RSAD2*), and chemokines *CXCR4*, and *CCL5*. Other immune genes are similarly important, as shown in Figure 3-17C. The upregulation of, for example, *SOCS3* and other nuclear factors and transcription factors by *E. coli* are consistent with a microarray study in udder tissue<sup>102</sup>. *SOCS3* also acts as a crucial immune regulator in human bladder epithelial cells in response to uropathogenic *E. coli* to the survival advantage of the pathogen<sup>271</sup>. The upregulation of chemotactic protein *SI00A9* aligns with previous findings in pbMEC after exposure to *E. coli*<sup>56,59,75</sup>. Recently, *SI00A9* appeared to be involved in *IL1β* secretion in interferon- $\gamma$ -primed monocytes in a *TLR4/MyD88/NF- $\kappa$ B*-independent pathway and has been proposed as a damage-associated molecular pattern<sup>272</sup>. In virus-exposed MAC-T, *RSAD2* and *CCL5* were among the interferon-stimulated genes and chemokines<sup>273</sup>. These genes, therefore, highlight the responsiveness of the epithelial cell model to pathogens, which is equally applicable to both viruses and mastitis bacteria *E. coli*.

The response of pbMEC to *S. aureus* is slightly different from the response to *E. coli*. However, the GO term immune system process indicates an immune response to challenge with *S. aureus*<sub>1027</sub>. A pathogen-specific response of pbMEC to *E. coli* or *S. aureus* is generally established. A 2021 study has shown that *S. aureus* impairs the activation of nuclear factor erythroid 2-related factor two (*Nrf2*), a stress-responsive transcription factor<sup>274</sup>. In the present research, cytochrome P450 family member *CYP26B1*, F2R-like trypsin receptor one (*F2RL1*), and Serine/threonine-protein kinase 10 (*STK10*) showed the most significant fold change in response to *S. aureus*<sub>1027</sub> in pbMEC (Figure 3-18). While *CYP26B1* regulates developmental processes of epithelial cells and organs<sup>275</sup>, *F2RL1* and *STK10* are strongly involved in the cellular immune response mechanisms, such as recognizing proteolytic enzymes<sup>276</sup> and recruiting lymphocytes<sup>277</sup>, respectively. These findings align with *S. aureus* being known to express proteases as virulent factors<sup>278</sup>, and immune cells help to clear the infection.

## 4.6 ChIP-seq Data Analysis

### 4.6.1 Sequencing Depth and Read Mapping

The high-throughput ChIP-seq replicate raw data passed the first quality assessment stage with the FASTQC tool<sup>181</sup>. With the obtained ChIP-seq data, the critical sequencing depth of up to 60 million reads for ChIP-seq data analysis of mammals<sup>279</sup> was met. The sequencing depth can have a major effect on ChIP-seq results<sup>279</sup> for the histone marks studied. Notably and in line with previous studies, H3K4me3 was saturated on average at a lower sequencing depth than H3K27me3<sup>280</sup>. The use of one of the standard mapping tools, BWA<sup>201,202</sup> or Bowtie<sup>281</sup>, forms the basis of aligned read information. Several studies exclude multi mapping reads from their analysis, which is the common practice<sup>282</sup>.

Others keep multimapping reads to increase sequencing depth and ability to detect novel peaks<sup>283,284</sup>. We followed the common sequencing data processing and included only uniquely mapping reads.

#### **4.6.2 Reproducibility of ChIP-seq**

The scientific community requirement of at least two replicates per ChIP-seq experiment<sup>282,285</sup> was met with three independent biological replicate experiments for primary cells and three technical replicates for the MAC-T cell line used for ChIP-seq. Under certain assumptions, an alternative can include repeated ChIP-seq measurements of the exact experimental replicate.

In addition to targeting the specific histone marks H3K4me3 and H3K27me3, input DNA was sequenced as a control sample. The input DNA is a common control sample that filters out the background noise<sup>282</sup>. Contrary to other studies<sup>282,286</sup>, no mock ChIP was performed for the X-ChIP-seq targeting IgG or H3. However, differences between H3 and input DNA have been reported to have a minor impact on the quality of a standard analysis<sup>286</sup>.

We applied plotfingerprint, a CHANCE<sup>279,287</sup> integrated tool for quality control and validation of ChIP-seq data, and thus followed computational analysis workflow to assess the strength of the individual ChIP-seq files<sup>285</sup>. The fingerprint curve for the H3K4me3 ChIP assay remains near X-axes until 0.6, which could indicate insufficient coverage in these immunoprecipitations. Interpreting the remaining profile argues for a characteristic point-source enrichment<sup>280</sup>. The H3K27me3 ChIP profile and input are very similar. Nonetheless, this profile is typical for H3K27me3, which produces large enrichment domains<sup>280,287–290</sup>. Correlation analysis and PCA were used to visualize further the degree of overlap between ChIP assays and replicates of the cell types. Replicates of the two histone assays in MAC-T and pbMEC exhibit similar read distributions, forming clusters in PCA and heatmap. The Pearson correlation coefficient was close to 1, thus matching the consistency and replicate samples in high-quality experiments<sup>285,291</sup>.

#### **4.6.3 ChIP-seq Peak Calling and Differential Binding Analysis**

Peak callers, such as EPICS2 and SICER<sup>207,292</sup>, are specifically designed software to identify regions of the genome, which are significantly enriched for histone ChIP reads<sup>285,293</sup>. The peak calling algorithm integrates histone modification ChIP-seq with input reads and outputs putative genomic binding sites for the epigenetic mark. In the vast majority of peak calling programs, the binding pattern is accounted for by the shape of the peak<sup>229</sup>. A significant difficulty in identifying enriched regions is that there are three types: sharp, broad, and mixed<sup>294</sup>. Therefore, algorithms must be adapted when histone marks extend over larger areas of the genome<sup>229</sup>. Adjustments to the settings may introduce bias and cause discrepancies in peak detection. For comparative ChIP-seq analysis, the differential binding analysis was performed using DESeq2<sup>188</sup>, one reliable method that outperforms other tools through short runtime and low memory requirements<sup>295</sup>. Similar to a previous study in mouse cells<sup>296</sup>, we detected more peaks in H3K4me3 than in H3K27me3. Genomic annotation data is consistent with this prior study, showing that H3K4me3 is enriched at promoter-TSS stronger than the repressive mark H3K27me3. In addition, H3K27me3 is widely found to be more enriched in intergenic and intron regions<sup>293,296,297</sup>. As the

complete description of the bovine genome has yet to be established, there may be additional, hitherto unknown genes in these intergenic regions.

#### 4.7 Genome-wide Sequencing Maps Epigenetic Profile of MAC-T and pbMEC

A common way to interpret several genome-wide data sets is to focus on significant overlapping genes and compare them to reference databases<sup>298</sup>. This study applied the outlined approach to answering the proposed research question of whether mastitis pathogens have an epigenetic regulatory effect on the transcriptome beyond immune genes in representative bovine epithelial cell models.

The results confirmed an association exclusively for the active histone mark H3K4me3 with differentially expressed genes in *E. coli*<sub>1303</sub>-challenged pbMEC. The binding sites overlap analysis revealed 15 differentially expressed genes in pbMEC challenged with *E. coli*<sub>1303</sub> that are associated with active histone marks (Figure 3-36). Among these genes are *CYP11A1*, *CYP11B1*, and Krüppel-like factor four (*KLF4*), which were associated with functional pathways such as tryptophan metabolism and chemical carcinogenesis-receptor activation. It is reported for intestinal cells that microbial tryptophan catabolites can affect host physiology, including immune response<sup>299</sup>, which might also be relevant in MEC. *KLF4* is a transcription factor associated with microRNAs and transcriptional regulation in mammary epithelial and other cancer cell types<sup>300</sup>, supporting the theory of crosstalk of epigenetic mechanisms during gene expression regulation. Bovine peripheral blood mir-25, which targets *KLF4*, was among the ten most highly expressed miRNAs in mastitis, potentially mediating inflammation<sup>301,302</sup>. *CYP11A1* and *CYP11B1* induce genotoxicity, which can translate into epigenetic changes and is critical for disease onset<sup>303</sup>. The exact and causal role in the complex cellular network during *E. coli*-challenge would need to be solved in pbMEC.

The subset of genes from gene expression analysis *E. coli* versus control challenge in pbMEC (394 genes from RNA-seq), which was without histone association (section 3.4.2), was associated with specific KEGG signaling pathways such as *IL-17* signaling pathway, *TNF* signaling pathway, *NF-κB* signaling pathway, *MAPK* signaling pathway, *PI3K-Akt* signaling pathway, and NOD-like receptor signaling pathway. These results are consistent with the typical pathways of pathogen recognition and subsequent upregulation of antimicrobial peptides, cytokines, and chemokines in MEC.

Given the findings in RNA-seq, in which many genes were differentially expressed between control-challenged MAC-T and pbMEC, another research question arose. It was whether histone modifications link to gene expression in these cell models. The results from the comparison of control-challenged MAC-T and pbMEC suggest strong evidence of epigenetic transcription regulation, as many genes were associated with one or both histone marks (Figure 3-30). With approximately 4,000 genes accounting for nearly two-thirds of the DEGs associated with the active histone mark H3K4me3, it can be strongly presumed that active histone marks participate in the molecular regulatory network in MEC. Bivalent histone-regulated genes have both active and repressive histone modifications that help rapidly and selectively turn on or off the expression of these genes. In cattle, it is established that these bivalent



marked genes can be essential for regulating various processes, such as development and cell differentiation. An important novel finding of this study was the identification of 118 DEGs overlapping with both active and repressive histone modifications in bovine MEC. These bivalent regulated genes exhibit functions for epithelial development and epithelial differentiation functions, among others, highlighting the contribution to an epithelial phenotype (Figure 3-32).

In contrast, the 22 DEGs overlapping with binding sites for the repressive histone mark H3K27me3 were associated with the biological process of organelle transport along microtubules (Figure 3-34). Microtubules function as tracks for intracellular transport connecting cells or top (apical) and bottom (basolateral) surfaces<sup>304,305</sup>. Suppression of organelle and cargo transport in epithelial cells indicates impairment of critical cellular functions of the epithelium.

The simultaneous presence of H3K4me3 and H3K27me3 associated with 422 genes that did not overlap with DEGs may putatively present a set of silenced or paused genes (Figure 3-35). The functional analysis for this gene subset revealed GO terms such as synaptic signaling, neurogenesis, and transcription regulatory activity. These bivalent genes need further investigation to clarify the exact function and regulation in MEC.

Besides histone modification and DNA methylation, the effect of non-coding RNAs contributes to the function and activity of a gene. When comparing the two cell models, intron and promoter TSS comprised most of the relative occurrence of peaks (Figure 3-29). Interestingly, some of the peaks comparing the two cell models fell into the categories of lncRNA exon or intron or promoter TSS (Figure B-17), assuming these epigenetic mechanisms are involved in the cellular network and controlled through H3K4me3 and therefore suggest interrelation.

The example of bta-mir-2887-1 from the *E. coli* effect in MAC-T suggests a contributing regulatory role of microRNAs in the cellular network. These results tie in with the postulated role of bta-miR-2887 as one of two microRNAs in the immune response of Johne's disease in cattle caused by mycobacterium species<sup>306</sup>. Further evidence for the role in the immune response is the identification of bta-miR-2887 among 77 differentially expressed miRNA in *S. aureus*-induced mastitis in Chinese Holstein<sup>307</sup>. Others have shown that bta-miR-2887 was among the most significantly differentially expressed miRNAs in lactation stages<sup>308</sup>. For the DEGs of *E. coli* *l*<sub>1303</sub> versus the control challenge in pbMEC, the genes were, among others, enriched in the KEGG pathway microRNAs in cancer. At this stage of understanding, it can be considered that microRNAs, among other epigenetic factors, play a role during mammary gland infection with *E. coli*.

In principle, the relationship between the binding sites of epigenetic factors and target genes is complex. The approach used in this study could be further refined regarding the fine resolution of the results in terms of location in the gene body. The proximity or overlap of bivalent markers in the exact location of the gene body and their function need to be further investigated. Moreover, proximity is not necessarily evidence of causal interaction. Epigenetic factors can bind multiple sites throughout the genome. In addition, one factor may bind to various target genes, whereas numerous epigenetic factors

may regulate a single gene.<sup>309</sup> Thus, the molecular function and operation of the identified epigenetic mechanisms for bovine epithelial cells require further research.

#### 4.8 Usability of the Cell Models

A striking finding from the present study was that the commonly used cell models, MAC-T and pbMEC, are dissimilar rather than similar as identified in RNA-seq and amplified with ChIP-seq. This finding was unexpected. There is considerable debate among researchers about appropriate cell models because they reflect some aspects of, e.g., inflammatory responses, but there are also reproducibility issues. Indeed, the body of literature varies regarding the supplementation of MAC-T culture media, in which some researchers add insulin alone<sup>310,311</sup>. In contrast, others add insulin, hydrocortisone, and prolactin<sup>65,312,313</sup>, similar to the used supplementation of pbMEC media in this dissertation. However, there are also inconsistencies in the medium composition for pbMEC in the field. It was found that pbMEC respond to *E. coli* depends on supplementing media with FBS<sup>314</sup>. This result suggests that besides animal bias, an experimental bias is possible. An established MAC-T cell line from one laboratory may differ from that of another laboratory due to the culture microenvironment. Self-critically, the design with different media compositions could have been more optimal, as it can affect the experimental outcome.

Moreover, MAC-T may have diverged in laboratories worldwide due to the passaging. The passages for MAC-T in the current study ranged between early 50 and 60. We suspect passaging increases the diversity of MAC-T and causes the loss of essential receptors for pathogen recognition. Similar observations have been made that cell lines change their phenotype from the original tissue<sup>315</sup>. Thus, the genetic and epigenetic plasticity of the MECs depends on the anatomical origin, how the cell line was immortalized, or the environmental conditions under which the cells were grown. This discrepancy raises doubts about whether the two models are truly interchangeable in use as postulated in the livestock research community.

#### 4.9 Future work

Observations from this dissertation should be borne in mind. The arguments in the discussion suggest focusing on a cross-validation study of established MAC-T cell lines from various laboratories worldwide. Until then, the results indicate that the future focus should be on primary cells. In addition, testing the effect of live bacteria on cells over multiple time points is an essential avenue for future research on epigenetic transcriptional regulation.

Several layers of epigenetic regulation still need to be addressed in a future study. A more detailed understanding of epigenetic transcriptional regulation is required to provide a complete picture of crucial mechanisms during mastitis. A follow-up study may focus on studying the remaining FAANG core marks (H3K4me1, H3K27Ac, and transcription factor CTCF) and integrate those with present findings in pbMEC. In addition, ncRNAs and DNA methylation could be determined, resolving their contribution to the regulatory network of MECs. Further studies could also include examining the spatial organization of chromatin in a cell with Hi-C<sup>316</sup> or measuring chromatin accessibility and mapping nucleosomes and

non-histone proteins with ATAC-seq<sup>316</sup>. These methodological innovations align with key areas the FAANG community plans to address in the next decade<sup>317</sup>. In primary cell culture or in the context of taking mastitis tissue biopsies, a single-cell sequencing approach could resolve cellular interaction's complexity and contribution to the immune response during mastitis. Once their functional role during mastitis is established, the epigenetic marker can be used as a diagnostic biomarker or for a drug target development<sup>316</sup>.

With the advances in epigenetic profiling techniques, various data analysis tools and web servers were developed to facilitate data integration, visualization, and interpretation. A possible future direction with a more bioinformatics focus is that the data underlying this study could be used to perform target prediction analyses more detailed and efficiently with software packages such as BETA<sup>309</sup>, LOLA<sup>298</sup>, or MotifStack<sup>318</sup>. In general, it seems essential to turn to the collection and integration of omics data such as transcriptomics, genomics, phenomics, and epigenomics<sup>316</sup> to understand the functional implications of epigenetics. Combining and analyzing these diverse data poses a significant computational challenge and demands investment in computational infrastructure.

Identified functionalities must be verified in *in vitro* and *in vivo* experiments in mice, humans, and cattle to ascertain mammalian similarities or species specificities.

Future research should also target epigenetic mechanisms and collect supportive data for functional validation. These studies include immunohistochemistry and western blot experiments to prove the presence of the gene product. Versatile knockdown or overexpression experiments of enzymes responsible for histone modification or other epigenetic mechanisms *in vitro* and *in vivo* may be complementary. The causal relationship between the presence of the epigenetic mark and the pathophysiological state can be determined.

Based on the established patterns of genome-wide epigenetic patterns in MEC from the dissertation, applying modulators of epigenetic marks can advance the understanding of complex cellular mechanisms. Epigenetic modulators are drugs or compounds that can selectively alter the activity of enzymes responsible for adding or removing epigenetic marks on DNA or histone proteins. Two inhibitor examples are S2101 and C646. S2101 inhibits lysine-specific demethylase 1A, while C646 selectively inhibits histone acetyltransferases (p300 and CBP). Both inhibitory compounds affect transcriptional machinery and resulting cellular processes. It is desirable to co-treat pbMEC with *E. coli* and the inhibitors and determine the cellular response, either gene-targeted or genome-wide.

The targeted dysregulation of histone modifications allows new insights into various biological processes, including pathophysiology, and could be critical for developing new therapies. These alternative treatment approaches are urgent since extensive use of antibiotics and drugs in livestock farming causes a significant risk to humans and the environment and gives rise to antibiotic resistance. The potential applications of epigenetic modulators have been described in the literature for numerous diseases. Some epigenetic modulators have even passed the clinical trial stage and have now been approved for the treatment of cancer in human<sup>319</sup>. Probiotics, natural products, vitamins, and designed

chemicals as modulators or inhibitors of epigenetic mechanisms are promising tools to support the host immune system and prevent inflammation. Targeted modulators of epigenetic marks have the potential to enable more effective and individualized treatments. Animal- or herd-specific therapies and treatment plans could be developed based on the identification of mastitis-related epigenetic changes.

The financial aspect of these new approaches and treatment specificity are restraining factors. The great potential in light of the reversibility and transience of histone modifications makes them desirable therapeutic targets. Specialized delivery methods need to be generated to achieve specificity and precision<sup>319</sup>. Validation and approval of epigenetic modulators can be a lengthy process. It is unlikely that there will be a universal solution overnight, but development and expanded knowledge in this area will be to the benefit of animals, humans, and the environment.

Despite significant progress in epigenetic research, there are still technical bottlenecks, such as the difficulty in obtaining high-quality samples from specific tissues. Overcoming these technical problems is crucial to improving the accuracy and reliability of epigenetic measurements. The results of the *in vitro* culture of isolated cells from udder tissue or immortalized cells should be applied with caution to the *in vivo* context. This discussion goes to the heart of the community about which approach and model are the most appropriate. *In vivo*, bacteria are confronted with a community of cells when they enter the host. However, homogenous *in vitro* models outside the organism are a reliable strategy to clarify each contribution to the cellular response in the udder. As established in humans, it might be worthwhile to establish a 3D cell model for mammary tissue in the same 3D microenvironment, including stromal, parenchymal, epithelial, and fibroblast cells<sup>320</sup> to study mastitis. An alternative approach can be using isolated perfused bovine udders<sup>242</sup>.

Overall, the data of this study suggests a multifaceted role of epigenetics in bovine mastitis that needs further investigation. Future studies should aim to replicate the results on a larger scale and focus mainly on X-ChIP to investigate histone modifications or specific transcription factors.

## Summary

Mastitis is a severe animal disease of considerable concern to farmers, the general public, and animals. Despite the extensive literature on the molecular mechanisms of pathogen-specific determinants in mastitis pathogenesis, the involvement and significance of epigenetic regulatory mechanisms, such as histone modifications, for udder immune defense in cattle remained underexamined. Therefore, this study focused on determining epigenetically regulated host gene expression after pathogen contact. Bovine mastitis epithelial cell culture models were exposed to heat-inactivated *Escherichia coli* or *Staphylococcus aureus* for 0 to 3 h, and downstream extracted nucleic acids served for subsequent methods and analysis.

Although inconsistently, the cellular immune gene response could be mimicked, with various underlying reasons such as culturing microenvironment and passage of MAC-T. Using N-ChIP-qPCR, I assessed histone modification, modulation, and the effects on gene expression of selected immune response genes in MAC-T. Problems in N-ChIP potentially arose from antibodies, sample size, and unreliable data normalization, so the results were inconclusive. Overall, genome-wide analysis of H3K4me3 and H3K27me3 and transcriptomics demonstrated no pronounced effect of pathogens. However, the active histone mark H3K4me3 showed overlap with differentially expressed genes in pbMEC challenged with *E. coli*. This discovery is the first report of gene expression regulation by histone modifications in MEC challenged with *E. coli*. A complex network of epigenetic mechanisms is likely involved in mastitis.

Another important finding was the considerable discrepancy between the cell models, MAC-T and pbMEC, on the transcriptomic level. Overlap analysis of differentially expressed genes with differential binding sites identified an association with either or both tested histone marks. Functional analysis suggests histone modifications controlling, in part, the epithelial phenotype.

These experiments add to a growing corpus of research showing the involvement of epigenetics in cattle, in particular in a disease context. The broader implication of this research is a better understanding of the network of regulatory mechanisms in cells contributing to the immune response in the bovine udder. Interfering during mastitis onset or the course of the infection would be a valuable therapeutic target. In summary, the results of this study provide a blueprint and first insight into histone mark profiles of these two bovine epithelial cell models, a previously unexplored aspect. Overall, evidence of epigenetic regulators in mastitis warrants follow-up studies.

## Zusammenfassung

Mastitis ist eine schwere Tierkrankheit, die Landwirten, Menschen und Tieren erhebliche Sorgen bereitet. Trotz der umfangreichen Literatur über die molekularen Mechanismen pathogen-spezifischer Determinanten in der Entstehung von Mastitis sind die Beteiligung und die Bedeutung epigenetischer Regulationsmechanismen, wie z.B. Histon-Modifikationen, für die Immunabwehr im Euter von Rindern noch nicht ausreichend erforscht. Daher konzentrierte sich diese Studie auf die Bestimmung der epigenetisch regulierten Wirtsgenexpression nach Pathogenkontakt. Mastitisepithelzellmodelle (MEC) des Rindes wurden 0 bis 3 h lang hitzeinaktivierten *Escherichia coli* oder *Staphylococcus aureus* ausgesetzt, und die anschließend extrahierten Nukleinsäuren dienten für anschließende Methoden, Ergebniserhebung und darauf aufbauenden Analysen.

Die zelluläre Immunreaktion konnte, wenn auch uneinheitlich, nachgeahmt werden, wobei verschiedene Gründe, wie z. B. die Kultivierungsumgebung und die Passage von MAC-T, eine Rolle spielten. Mithilfe von N-ChIP-qPCR untersuchten wir die Histonmodifikation, die Modulation und die Auswirkungen auf die Genexpression ausgewählter Gene der Immunantwort in MAC-T. Probleme bei N-ChIP ergaben sich möglicherweise aus den Antikörpern, der Probengröße und einer unzulänglichen Datennormalisierung, sodass die Ergebnisse nicht schlüssig waren. Insgesamt zeigten die genomweite Analyse von H3K4me3 und H3K27me3 sowie die Transkriptomik keinen ausgeprägten Effekt der Krankheitserreger. Allerdings zeigte die aktive Histonmarkierung H3K4me3 Überschneidungen mit differentiell exprimierten Genen in pbMEC, die mit *E. coli* behandelt waren. Diese Entdeckung ist der erste Bericht über die Regulierung der Genexpression durch Histonmodifikationen in MEC, die mit *E. coli* infiziert sind. Es ist wahrscheinlich, dass ein komplexes Netzwerk epigenetischer Mechanismen an der Mastitis beteiligt ist. Ein weiteres wichtiges Ergebnis war die erhebliche Diskrepanz zwischen den Zellmodellen MAC-T und pbMEC auf transkriptomischer Ebene. Eine Überschneidungsanalyse von differentiell exprimierten Genen mit differentiellen Bindungsstellen ergab einen Zusammenhang mit einer oder beiden getesteten Histonmarkierungen. Die funktionelle Analyse deutet darauf hin, dass Histonmodifikationen teilweise den epithelialen Phänotyp kontrollieren.

Diese Studien fügen sich in eine wachsende Zahl von Forschungsergebnissen ein, die zeigen, dass die Epigenetik bei Rindern eine Rolle spielt, insbesondere im Zusammenhang mit einer Krankheit. Dieser Forschungsansatz führt zu einem besseren Verständnis des Netzwerks von Regulationsmechanismen in Zellen, die zur Immunantwort im Rindereuter beitragen. Ein Eingriff während des Ausbruchs der Mastitis oder des Verlaufs der Infektion wäre ein wertvolles therapeutisches Ziel. Zusammenfassend, erlauben die Ergebnisse dieser Arbeit einen ersten Einblick in die Histonmarkierungsprofile der untersuchten beiden Rinderepithelzellmodelle, was bisher ein unerforschter Aspekt war. Insgesamt geben diese gewonnenen Ergebnisse Anlass für weiterführende Studien auf epigenetische Regulatoren bei Mastitis.

## References

1. Zeder, M. A. The domestication of animals. *J. Anthropol. Res.* **68**, 161–190 (2012).
2. Buchanan, D. S. *Major bos taurus breeds. Encyclopedia of Dairy Sciences: Third edition* vol. 1 (Elsevier, 2021).
3. Adisasmito, W. B. *et al.* One Health: A new definition for a sustainable and healthy future. *PLoS Pathog.* **18**, (2022).
4. Engle, T., Klingborg, D. J. & Rollin, B. E. *The welfare of cattle. The Welfare of Cattle* (2018). doi:10.1201/b21911.
5. FAO. *Impact of mastitis in small scale dairy production systems. Production and Health Working Paper.* vol. No. 13 <http://www.fao.org/3/a-i3377e.pdf> (2014).
6. van Soest, F. J. S., Santman-Berends, I. M. G. A., Lam, T. J. G. M. & Hogeveen, H. Failure and preventive costs of mastitis on Dutch dairy farms. *J. Dairy Sci.* **99**, 8365–8374 (2016).
7. Kvist, L. J. Diagnostic methods for mastitis in cows are not appropriate for use in humans: Commentary. *Int. Breastfeed. J.* (2016) doi:10.1186/s13006-016-0061-1.
8. DeGraves, F. J. & Fetrow, J. Economics of mastitis and mastitis control. *Vet. Clin. North Am. Food Anim. Pract.* **9**, 421–434 (1993).
9. Tiwari, J. G. *et al.* Trends in therapeutic and prevention strategies for management of Bovine Mastitis: An overview. *J. Vaccines Vaccin.* **4**, (2013).
10. Ruegg, P. L. A 100-Year Review: Mastitis detection, management, and prevention. *J. Dairy Sci.* **100**, 10381–10397 (2017).
11. Gonçalves, J. L. *et al.* Using milk leukocyte differentials for diagnosis of subclinical bovine mastitis. *J. Dairy Res.* **84**, 309–317 (2017).
12. Hughes, K. & Watson, C. J. The Mammary Microenvironment in Mastitis in Humans, Dairy Ruminants, Rabbits and Rodents: A One Health Focus. *J. Mammary Gland Biol. Neoplasia* **23**, 27–41 (2018).
13. Rowson, A. R., Daniels, K. M., Ellis, S. E. & Hovey, R. C. Growth and development of the mammary glands of livestock: A veritable barnyard of opportunities. *Semin. Cell Dev. Biol.* **23**, 557–566 (2012).
14. Jones. Understanding the basics of Mastitis. *VCE Publ.* (2009).
15. Watts, J. L. Etiological agents of bovine mastitis. *Vet. Microbiol.* **16**, 41–66 (1988).
16. Harmon, R. J. Physiology of Mastitis and Factors Affecting Somatic Cell Counts. *J. Dairy Sci.* **77**, 2103–2112 (1994).
17. Zadoks, R. N., Middleton, J. R., McDougall, S., Katholm, J. & Schukken, Y. H. Molecular epidemiology of mastitis pathogens of dairy cattle and comparative relevance to humans. *J. Mammary Gland Biol. Neoplasia* **16**, 357–372 (2011).
18. Abdalhamed, A. M., Zeedan, G. S. G. & Zeina, H. A. A. Isolation and identification of bacteria causing mastitis in small ruminants and their susceptibility to antibiotics, honey, essential oils, and plant extracts. *Vet. World* **11**, 355–362 (2018).
19. Gerjets, I. & Kemper, N. Coliform mastitis in sows: A review. *J. Swine Heal. Prod.* **17**, 97–105 (2009).
20. Jenny, B., Vidondo, B., Pendl, W., Kümmerlen, D. & Sidler, X. Evaluation of risk factors for mastitis-metritis-agalactia in pig farms in Switzerland. *Schweiz. Arch. Tierheilkd.* **157**, 689–696 (2015).
21. Goldstone, R. J., Harris, S. & Smith, D. G. E. Genomic content typifying a prevalent clade of bovine mastitis-associated *Escherichia coli*. *Sci. Rep.* **6**, 1–15 (2016).
22. Gomes, F. & Henriques, M. Control of Bovine Mastitis: Old and Recent Therapeutic Approaches. *Curr. Microbiol.* **72**, 377–382 (2016).
23. Smith, K. L., Todhunter, D. A. & Schoenberger, P. S. Environmental Mastitis: Cause, Prevalence, Prevention. *J. Dairy Sci.* **68**, 1531–1553 (1985).
24. Burvenich, C., Van Merris, V., Mehrzad, J., Diez-Fraile, A. & Duchateau, L. Severity of *E. coli* mastitis is mainly determined by cow factors. *Vet. Res.* **34**, 521–564 (2003).
25. Smith, K. L. & Hogan, J. S. Environmental mastitis. *Vet. Clin. North Am. Food Anim. Pract.* **9**, 489–498 (1993).
26. Hillerton, J. E. & Berry, E. A. Treating mastitis in the cow - A tradition or an archaism. *J. Appl. Microbiol.* **98**, 1250–1255 (2005).

27. Hogan, J. S. *et al.* Field Survey of Clinical Mastitis in Low Somatic Cell Count Herds. *J. Dairy Sci.* **72**, 1547–1556 (1989).
28. Sharma, N., Singh, N. K. & Bhadwal, M. S. Relationship of somatic cell count and mastitis: An overview. *Asian-Australasian J. Anim. Sci.* **24**, 429–438 (2011).
29. Leitner, G., Shoshani, E., Krifucks, O., Chaffer, M. & Saran, A. Milk leucocyte population patterns in bovine udder infection of different aetiology. *J. Vet. Med. Ser. B* **47**, 581–589 (2000).
30. Botrel, M. A. *et al.* Distribution and antimicrobial resistance of clinical and subclinical mastitis pathogens in Dairy Cows in Rhône-Alpes, France. *Foodborne Pathog. Dis.* **7**, 479–487 (2010).
31. Lozano, C., Gharsa, H., Ben Slama, K., Zarazaga, M. & Torres, C. Staphylococcus aureus in Animals and Food: Methicillin Resistance, Prevalence and Population Structure. A Review in the African Continent. *Microorganisms* **4**, 12 (2016).
32. Lowy, F. D. Lowy 1998. *N. Engl. J. Med.* **339**, 520–532 (1998).
33. Haveri, M., Hovinen, M., Roslöf, A. & Pyörälä, S. Molecular types and genetic profiles of Staphylococcus aureus strains isolated from bovine intramammary infections and extramammary sites. *J. Clin. Microbiol.* **46**, 3728–3735 (2008).
34. Kümmel, J. *et al.* Staphylococcus aureus entrance into the Dairy Chain: Tracking S. aureus from dairy cow to cheese. *Front. Microbiol.* **7**, 1–11 (2016).
35. Roberson, J. R., Fox, L. K., Hancock, D. D., Gay, J. M. & Besser, T. E. Ecology of Staphylococcus aureus Isolated from Various Sites on Dairy Farms. *J. Dairy Sci.* **77**, 3354–3364 (1994).
36. Philpot, W. N. Control of Mastitis by Hygiene and Therapy. *J. Dairy Sci.* **62**, 168–176 (1979).
37. Abebe, R., Hatiya, H., Abera, M., Megersa, B. & Asmare, K. Bovine mastitis: Prevalence, risk factors and isolation of Staphylococcus aureus in dairy herds at Hawassa milk shed, South Ethiopia. *BMC Vet. Res.* **12**, (2016).
38. Bannerman, D. D. Pathogen-dependent induction of cytokines and other soluble inflammatory mediators during intramammary infection of dairy cows. *J. Anim. Sci.* **87**, 10–25 (2009).
39. Petzl, W. *et al.* Escherichia coli, but not Staphylococcus aureus triggers an early increased expression of factors contributing to the innate immune defense in the udder of the cow. *Vet. Res.* **39**, (2008).
40. Almeida, P. A., Matthews, K. R., Cifrian, E., Guidry, A. J. & Oliver, S. P. Staphylococcus aureus Invasion of Bovine Mammary Epithelial Cells. *J. Dairy Sci.* **79**, 1021–1026 (1996).
41. Garzoni, C. & Kelley, W. L. Staphylococcus aureus: new evidence for intracellular persistence. *Trends Microbiol.* **17**, 59–65 (2009).
42. Monistero, V. *et al.* Staphylococcus aureus isolates from bovine mastitis in eight countries: Genotypes, detection of genes encoding different toxins and other virulence genes. *Toxins (Basel)*. **10**, (2018).
43. Bjorland, J. *et al.* Widespread distribution of disinfectant resistance genes among staphylococci of bovine and caprine origin in Norway. *J. Clin. Microbiol.* **43**, 4363–4368 (2005).
44. Melchior, M. B. *et al.* Biofilm formation and genotyping of Staphylococcus aureus bovine mastitis isolates: Evidence for lack of penicillin-resistance in Agr-type II strains. *Vet. Microbiol.* **137**, 83–89 (2009).
45. Schukken, Y. H. *et al.* Host-response patterns of intramammary infections in dairy cows. *Vet. Immunol. Immunopathol.* **144**, 270–289 (2011).
46. Inman, J. L., Robertson, C., Mott, J. D. & Bissell, M. J. Mammary gland development: Cell fate specification, stem cells and the microenvironment. *Dev.* **142**, 1028–1042 (2015).
47. Cristea, S. & Polyak, K. Dissecting the mammary gland one cell at a time. *Nat. Commun.* **9**, 9–11 (2018).
48. Dziegielewska, Ż. & Gajewska, M. Stromal-Epithelial Interactions during Mammary Gland Development. *Stromal Cells - Struct. Funct. Ther. Implic.* **i**, 13 (2019).
49. Nickerson, S. C. & Akers, R. M. Mammary Gland: Anatomy. in *Encyclopedia of Dairy Sciences: Second Edition* 328–337 (2011). doi:10.1016/B978-0-12-374407-4.00290-9.
50. Ryman, V. E., Packiriswamy, N. & Sordillo, L. M. Role of endothelial cells in bovine mammary gland health and disease. *Anim. Heal. Res. Rev.* **16**, 135–149 (2015).
51. Shaykhiev, R. & Bals, R. Interactions between epithelial cells and leukocytes in immunity and tissue homeostasis. *J. Leukoc. Biol.* **82**, 1–15 (2007).
52. Günther, J. & Seyfert, H. M. The first line of defence: insights into mechanisms and relevance



- of phagocytosis in epithelial cells. *Semin. Immunopathol.* 555–565 (2018) doi:10.1007/s00281-018-0701-1.
53. Chua, A. C. L., Hodson, L. J., Moldenhauer, L. M., Robertson, S. A. & Ingman, W. V. Dual roles for macrophages in ovarian cycle-associated development and remodelling of the mammary gland epithelium. *Development* **137**, 4229–4238 (2010).
  54. Aupperlee, M. D. *et al.* Epidermal growth factor receptor (EGFR) signaling is a key mediator of hormone-induced leukocyte infiltration in the pubertal female mammary gland. *Endocrinology* **155**, 2301–2313 (2014).
  55. Need, E. F., Atashgaran, V., Ingman, W. V. & Dasari, P. Hormonal regulation of the immune microenvironment in the mammary gland. *J. Mammary Gland Biol. Neoplasia* **19**, 229–239 (2014).
  56. Günther, J. *et al.* Assessment of the immune capacity of mammary epithelial cells: Comparison with mammary tissue after challenge with *Escherichia coli*. *Vet. Res.* **40**, (2009).
  57. Strandberg, Y. *et al.* Lipopolysaccharide and lipoteichoic acid induce different innate immune responses in bovine mammary epithelial cells. *Cytokine* **31**, 72–86 (2005).
  58. Lahouassa, H., Moussay, E., Rainard, P. & Riollot, C. Differential cytokine and chemokine responses of bovine mammary epithelial cells to *Staphylococcus aureus* and *Escherichia coli*. *Cytokine* **38**, 12–21 (2007).
  59. Mitterhuemer, S. *et al.* *Escherichia coli* infection induces distinct local and systemic transcriptome responses in the mammary gland. *BMC Genomics* **11**, (2010).
  60. Griesbeck-Zilch, B., Meyer, H. H. D., Kühn, C., Schwerin, M. & Wellnitz, O. *Staphylococcus aureus* and *Escherichia coli* cause deviating expression profiles of cytokines and lactoferrin messenger ribonucleic acid in mammary epithelial cells. *J. Dairy Sci.* **91**, 2215–2224 (2008).
  61. Günther, J. *et al.* Comparative kinetics of *Escherichia coli*- and *Staphylococcus aureus*-specific activation of key immune pathways in mammary epithelial cells demonstrates that *S. aureus* elicits a delayed response dominated by interleukin-6 (IL-6) but not by IL-1A or tumor n. *Infect. Immun.* **79**, 695–707 (2011).
  62. Fu, M., Chen, Y., Xiong, X., Lan, D. & Li, J. Establishment of mammary gland model in vitro: Culture and evaluation of a yak mammary epithelial cell line. *PLoS One* **9**, (2014).
  63. Jedrzejczak, M. & Szatkowska, I. Bovine mammary epithelial cell cultures for the study of mammary gland functions. *Vitr. Cell. Dev. Biol. - Anim.* **50**, 389–398 (2014).
  64. Hensen, S. M., Pavičić, M. J. A. M. P., Lohuis, J. A. C. M. & Poutrel, B. Use of bovine primary mammary epithelial cells for the comparison of adherence and invasion ability of *Staphylococcus aureus* strains. *J. Dairy Sci.* **83**, 418–429 (2000).
  65. Huynh, H. T., Robitaille, G. & Turner, J. D. Establishment of bovine mammary epithelial cells (MAC-T): An in vitro model for bovine lactation. *Exp. Cell Res.* **197**, 191–199 (1991).
  66. Zhao, K., Liu, H., Zhou, M. & Liu, J. Establishment and characterization of a lactating bovine mammary epithelial cell model for the study of milk synthesis. *Cell Biol. Int.* **34**, 717–721 (2010).
  67. Schmid, E., Schiller, D. L., Grund, C., Stadler, J. & Franke, W. W. Tissue type specific expression of intermediate filament proteins in a cultured epithelial cell line from bovine mammary gland. *J. Cell Biol.* **96**, 37–50 (1983).
  68. Huynh, H. & Pollak, M. HH2A, an immortalized bovine mammary epithelial cell line, expresses the gene encoding mammary derived growth inhibitor (MDGI). *Vitr. Cell. Dev. Biol. - Anim.* **31**, 25–29 (1995).
  69. Anand, V. *et al.* Establishment and characterization of a buffalo (*bubalus bubalis*) mammary epithelial cell line. *PLoS One* **7**, (2012).
  70. Zavizion, B., Van Duffelen, M., Schaeffer, W. & Politis, I. Use of microinjection to generate an immortalized bovine mammary cell line with both epithelial and myoepithelial characteristics. *Methods Cell Sci.* **17**, 271–282 (1995).
  71. Zavizion, B. Establishment and characterization of a bovine mammary myoepithelial cell line. *Vitr. Cell. Dev. Biol. - Anim.* **32**, 149–158 (1996).
  72. Hosseini, A., Sharma, R., Bionaz, M. & Loor, J. J. Transcriptomics Comparisons of Mac-T cells Versus Mammary Tissue during Late Pregnancy and Peak Lactation. *Adv. Dairy Res.* **01**, 68–70 (2013).
  73. Gray, C., Strandberg, Y., Donaldson, L. & Tellam, R. L. Bovine mammary epithelial cells, initiators of innate immune responses to mastitis. *Aust. J. Exp. Agric.* **45**, 757–761 (2005).

74. Günther, J., Koy, M., Berthold, A., Schuberth, H. J. & Seyfert, H. M. Comparison of the pathogen species-specific immune response in udder derived cell types and their models. *Vet. Res.* **47**, 1–11 (2016).
75. Brand, B. *et al.* Comparative expression profiling of *E. coli* and *S. aureus* inoculated primary mammary gland cells sampled from cows with different genetic predispositions for somatic cell score. *Genet. Sel. Evol.* **43**, 24 (2011).
76. Gilbert, F. B. *et al.* Differential response of bovine mammary epithelial cells to *Staphylococcus aureus* or *Escherichia coli* agonists of the innate immune system. *Vet. Res.* **44**, (2013).
77. Bannerman, D. D. *et al.* *Escherichia coli* and *Staphylococcus aureus* elicit differential innate immune responses following intramammary infection. *Clin. Diagn. Lab. Immunol.* **11**, 463–472 (2004).
78. Petzl, W. *et al.* Pathogen-specific responses in the bovine udder. Models and immunoprophylactic concepts. *Res. Vet. Sci.* **116**, 55–61 (2018).
79. Yang, D., Biragyn, A., Kwak, L. W. & Oppenheim, J. J. Mammalian defensins in immunity: More than just microbicidal. *Trends Immunol.* **23**, 291–296 (2002).
80. Askarian, F., Wagner, T., Johannessen, M. & Nizet, V. *Staphylococcus aureus* modulation of innate immune responses through Toll-like (TLR), (NOD)-like (NLR) and C-type lectin (CLR) receptors. *FEMS Microbiol. Rev.* **42**, 656–671 (2018).
81. Goldammer, T. *et al.* Mastitis Increases Mammary mRNA Abundance of  $\alpha$ -Defensin 5, but Not TLR9 in Cattle. *Society* **11**, 174–185 (2004).
82. Yoshimura, A. *et al.* Cutting edge: recognition of Gram-positive bacterial cell wall components by the innate immune system occurs via Toll-like receptor 2. *J. Immunol.* **163**, 1–5 (1999).
83. Schwandner, R., Dziarski, R., Wesche, H., Rothe, M. & Kirschning, C. J. Peptidoglycan- and lipoteichoic acid-induced cell activation is mediated by Toll-like receptor 2. *J. Biol. Chem.* **274**, 17406–17409 (1999).
84. Fournier, B. & Philpott, D. J. Recognition of *Staphylococcus aureus* by the innate immune system. *Clin. Microbiol. Rev.* **18**, 521–540 (2005).
85. Schröder, N. W. J. *et al.* Lipoteichoic acid (LTA) of *Streptococcus pneumoniae* and *Staphylococcus aureus* activates immune cells via Toll-like receptor (TLR)-2, lipopolysaccharide-binding protein (LBP), and CD14, whereas TLR-4 and MD-2 are not involved. *J. Biol. Chem.* **278**, 15587–15594 (2003).
86. Ozinsky, A. *et al.* The repertoire for pattern recognition of pathogens by the innate immune system is defined by cooperation between Toll-like receptors. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13766–13771 (2000).
87. Morr, M., Takeuchi, O., Akira, S., Simon, M. M. & Mühlradt, P. F. Differential recognition of structural details of bacterial lipopeptides by toll-like receptors. *Eur. J. Immunol.* **32**, 3337–3347 (2002).
88. Rainard, P. & Riollot, C. Innate immunity of the bovine mammary gland. *Vet. Res.* **37**, 369–400 (2006).
89. Liu, S., Shi, X., Bauer, I., Günther, J. & Seyfert, H. M. Lingual antimicrobial peptide and IL-8 expression are oppositely regulated by the antagonistic effects of NF- $\kappa$ B p65 and C/EBP $\beta$  in mammary epithelial cells. *Mol. Immunol.* **48**, 895–908 (2011).
90. Burska, A., Boissinot, M. & Ponchel, F. Cytokines as biomarkers in rheumatoid arthritis. *Mediators Inflamm.* **2014**, (2014).
91. Commins, S. P., Borish, L. & Steinke, J. W. Immunologic messenger molecules: Cytokines, interferons, and chemokines. *J. Allergy Clin. Immunol.* **125**, S53–S72 (2010).
92. Turner, M. D., Nedjai, B., Hurst, T. & Pennington, D. J. Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 2563–2582 (2014).
93. Nakajima, Y. *et al.* Elevated levels of tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ) and interleukin-6 (IL-6) activities in the sera and milk of cows with naturally occurring coliform mastitis. *Res. Vet. Sci.* **62**, 297–298 (1997).
94. Blum, J. W. *et al.* Tumor necrosis factor- $\alpha$  and nitrite/nitrate responses during acute mastitis induced by *Escherichia coli* infection and endotoxin in dairy cows. *Domest. Anim. Endocrinol.* **19**, 223–235 (2000).
95. Hoebe, D. *et al.* Role of endotoxin and TNF- $\alpha$  in the pathogenesis of experimentally induced

- coliform mastitis in periparturient cows. *J. Dairy Res.* **67**, 503–514 (2000).
96. Hirano, Toshio, Shizuo Akira, T. T. and T. K. Biological and clinical aspects of IL-6. *Immunol. Today* **1**, 1–3 (1990).
97. Yoshizaki, K. *et al.* Isolation and characterization of B cell differentiation factor (BCDF) secreted from a human B lymphoblastoid cell line. *J. Immunol.* **132**, 2948–54 (1984).
98. Scheller, J., Chalaris, A., Schmidt-Arras, D. & Rose-John, S. The pro- and anti-inflammatory properties of the cytokine interleukin-6. *Biochim. Biophys. Acta - Mol. Cell Res.* **1813**, 878–888 (2011).
99. van der Poll, T. & van Deventer, S. J. The role of interleukin 6 in endotoxin-induced inflammatory responses. *Prog. Clin. Biol. Res.* **397**, 365–377 (1998).
100. Biffl, W. L., Moore, E. E., Moore, F. A. & Peterson, V. M. Interleukin-6 in the injured patient: Marker of injury or mediator of inflammation? *Ann. Surg.* **224**, 647–664 (1996).
101. Krueger, J., Ray, A., Tamm, I. & Sehgal, P. B. Expression and function of interleukin-6 in epithelial cells. *J. Cell. Biochem.* **45**, 327–334 (1991).
102. Günther, J. *et al.* Differentiating Staphylococcus aureus from Escherichia coli mastitis: S. aureus triggers unbalanced immune-dampening and host cell invasion immediately after udder infection. *Sci. Rep.* **7**, 1–14 (2017).
103. Qazi, B. S., Tang, K. & Qazi, A. Recent Advances in Underlying Pathologies Provide Insight into Interleukin-8 Expression-Mediated Inflammation and Angiogenesis. *Int. J. Inflamm.* **2011**, 1–13 (2011).
104. Rinaldi, M. *et al.* A sentinel function for teat tissues in dairy cows: Dominant innate immune response elements define early response to E. coli mastitis. *Funct. Integr. Genomics* **10**, 21–38 (2010).
105. Gurao, A., Kashyap, S. K. & Singh, R.  $\beta$ -defensins: An innate defense for bovine mastitis. *Vet. World* **10**, 990–998 (2017).
106. A. Brogden, K. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat. Rev. Microbiol.* 238–250 (2005).
107. Lehrer, R. I., Ganz, T. & Selsted, M. E. Defensins: Endogenous antibiotic peptides of animal cells. *Cell* **64**, 229–230 (1991).
108. Swanson, K. *et al.* Expression of a  $\beta$ -defensin mRNA, lingual antimicrobial peptide, in bovine mammary epithelial tissue is induced by mastitis. *Infect. Immun.* **72**, 7311–7314 (2004).
109. Crick, F. Central dogma of molecular biology. *Nature* **227**, 561–563 (1970).
110. Kukurba, K. R. & Montgomery, S. B. RNA sequencing and analysis. *Cold Spring Harb. Protoc.* **2015**, 951–969 (2015).
111. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomic. *Nat. Rev. Genet.* **10**, 57–63 (2009).
112. Goldberg, A. D., Allis, C. D. & Bernstein, E. Epigenetics: A Landscape Takes Shape. *Cell* **128**, 635–638 (2007).
113. Kanherkar, R. R., Bhatia-Dey, N. & Csoka, A. B. Epigenetics across the human lifespan. *Front. Cell Dev. Biol.* **2**, (2014).
114. Waddington, C. H. The epigenotype. *Endeavour* **1**, 18–20 (1942).
115. Perrone, L., Matrone, C. & Singh, L. P. Epigenetic Modifications and Potential New Treatment Targets in Diabetic Retinopathy. *J. Ophthalmol.* **2014**, (2014).
116. Moore, L. D., Le, T. & Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology* **38**, 23–38 (2013).
117. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
118. Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* (80-. ). **293**, 1089–1093 (2001).
119. Bröske, A. M. *et al.* DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat. Genet.* **41**, 1207–1215 (2009).
120. Handy, D. E., Castro, R. & Loscalzo, J. Epigenetic modifications: Basic mechanisms and role in cardiovascular disease. *Circulation* **123**, 2145–2156 (2011).
121. Clapier, C. R. & Cairns, B. R. The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304 (2009).
122. Van Den Bossche, J., Neele, A. E., Hoeksema, M. A. & De Winther, M. P. J. Macrophage

- polarization: the epigenetic point of view. *Curr. Opin. Lipidol.* **25**, 367–373 (2014).
123. Kim, M. & Lin, S. Characterization of histone modification patterns and prediction of novel promoters using functional principal component analysis. *PLoS One* **15**, (2020).
124. Barski, A. *et al.* High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **129**, 823–837 (2007).
125. Liang, G. *et al.* Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 7357–7362 (2004).
126. Agalioti, T., Chen, G. & Thanos, D. Deciphering the transcriptional histone acetylation code for a human gene. *Cell* **111**, 381–392 (2002).
127. Shahbazian, M. D. & Grunstein, M. Functions of Site-Specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* **76**, 75–100 (2007).
128. Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* **40**, 897–903 (2008).
129. Haring, M. *et al.* Chromatin immunoprecipitation: Optimization, quantitative analysis and data normalization. *Plant Methods* **3**, 1–16 (2007).
130. Keren, L. & Segal, E. Fixated on fixation - using ChIP to interrogate the dynamics of chromatin interactions. *Genome Biol.* **14**, 2–4 (2013).
131. Cuatrecasas, P., Edelhoch, H. & Anfinsen, C. B. Fluorescence studies of the interaction of nucleotides with the active site of the nuclease of *Staphylococcus aureus*. in *Proceedings of the National Academy of Sciences of the United States of America* vol. 58 2043–2050 (1967).
132. Arrigoni, L. *et al.* Standardizing chromatin research: A simple and universal method for ChIP-seq. *Nucleic Acids Res.* **44**, (2015).
133. Giuffra, E. & Tuggle, C. K. Functional Annotation of Animal Genomes (FAANG): Current Achievements and Roadmap. *Annu. Rev. Anim. Biosci.* **7**, 65–88 (2019).
134. González-Recio, O., Toro, M. A. & Bach, A. Past, present, and future of epigenetics applied to livestock breeding. *Frontiers in Genetics* (2015) doi:10.3389/fgene.2015.00305.
135. Sinclair, K. D. *et al.* Epigenetics and developmental programming of welfare and production traits in farm animals. *Reproduction, Fertility and Development* (2016) doi:10.1071/RD16102.
136. Wu, X. *et al.* Multiple histone site epigenetic modifications in nuclear transfer and in vitro fertilized bovine embryos. *Zygote* **19**, 31–45 (2011).
137. Urrego, R., Rodriguez-Orsorio, N. & Niemann, H. Epigenetic disorders and altered gene expression after use of assisted reproductive technologies in domestic cattle. *Epigenetics* **9**, 803–815 (2014).
138. O'Doherty, A. M., MacHugh, D. E., Spillane, C. & Magee, D. A. Genomic imprinting effects on complex traits in domesticated animal species. *Front. Genet.* **6**, (2015).
139. Bierne, H., Hamon, M. & Cossart, P. Epigenetics and bacterial infections. *Cold Spring Harb. Perspect. Med.* **2**, (2012).
140. Silmon De Monerri, N. C. & Kim, K. Pathogens hijack the epigenome: A new twist on host-pathogen interactions. *Am. J. Pathol.* **184**, 897–911 (2014).
141. Handel, A. E., Ebers, G. C. & Ramagopalan, S. V. Epigenetics: molecular mechanisms and implications for disease. *Trends Mol. Med.* **16**, 7–16 (2010).
142. Bayarsaihan, D. Epigenetic mechanisms in inflammation. *J. Dent. Res.* **90**, 9–17 (2011).
143. Medzhitov, R. & Horng, T. Transcriptional control of the inflammatory response. *Nat. Rev. Immunol.* **9**, 692–703 (2009).
144. Fraga, M. F. *et al.* Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* **37**, 391–400 (2005).
145. Esteller, M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev. Genet.* **8**, 286–298 (2007).
146. Hattori, N. & Ushijima, T. Epigenetic impact of infection on carcinogenesis: Mechanisms and applications. *Genome Med.* **8**, 1–13 (2016).
147. Adcock, I. M., Ito, K. & Barnes, P. J. Histone deacetylation: An important mechanism in inflammatory lung diseases. *COPD J. Chronic Obstr. Pulm. Dis.* **2**, 445–455 (2005).
148. Moffatt, M. F. & Cookson, W. O. The genetics of asthma. Maternal effects in atopic disease. *Clin. Exp. Allergy* **28 Suppl 1**, 56–61; discussion 65–6 (1998).
149. Stefanowicz, D. *et al.* Elevated H3K18 acetylation in airway epithelial cells of asthmatic

- subjects. *Respir. Res.* **16**, 1–13 (2015).
150. Leong, M. M. L. *et al.* EBV infection is associated with histone bivalent switch modifications in squamous epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 14144–14153 (2019).
151. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953–959 (1996).
152. Castellucci, M. *et al.* IL-10 disrupts the Brd4-docking sites to inhibit LPS-induced CXCL8 and TNF- $\alpha$  expression in monocytes: Implications for chronic obstructive pulmonary disease. *J. Allergy Clin. Immunol.* **136**, 781–791.e9 (2015).
153. Chang, G. *et al.* Epigenetic mechanisms contribute to the expression of immune related genes in the livers of dairy cows fed a high concentrate diet. *PLoS One* **10**, 1–19 (2015).
154. Dong, G. *et al.* Feeding a high-concentrate corn straw diet induced epigenetic alterations in the mammary tissue of dairy cows. *PLoS One* **9**, 1–7 (2014).
155. Vanselow, J. *et al.* DNA-remethylation around a STAT5-binding enhancer in the  $\alpha$ S1-casein promoter is associated with abrupt shutdown of a  $\alpha$ S1-casein synthesis during acute mastitis. *J. Mol. Endocrinol.* **37**, 463–477 (2006).
156. Chen, J. *et al.* Bacterial lipopolysaccharide induced alterations of genome-wide DNA methylation and promoter methylation of lactation-related genes in bovine mammary epithelial cells. *Toxins (Basel)*. **11**, (2019).
157. Ju, Z. *et al.* Genome-wide methylation and transcriptome of blood neutrophils reveal the roles of DNA methylation in affecting transcription of protein-coding genes and miRNAs in E. coli-infected mastitis cows. *BMC Genomics* **21**, 1–14 (2020).
158. Sajjanar, B., Trakooljul, N., Wimmers, K. & Ponsuksili, S. DNA methylation analysis of porcine mammary epithelial cells reveals differentially methylated loci associated with immune response against Escherichia coli challenge. *BMC Genomics* **20**, 1–15 (2019).
159. He, Y. *et al.* Genome-wide bovine H3K27me3 modifications and the regulatory effects on genes expressions in peripheral blood lymphocytes. *PLoS One* **7**, (2012).
160. He, Y. *et al.* Whole-genome regulation analysis of histone H3 lysin 27 trimethylation in subclinical mastitis cows infected by Staphylococcus aureus. *BMC Genomics* **17**, 1–12 (2016).
161. Sladek, F. M. What are nuclear receptor ligands? *Mol. Cell. Endocrinol.* **334**, 3–13 (2011).
162. Herring, J. A., Elison, W. S. & Tessem, J. S. Function of Nr4a Orphan Nuclear Receptors in Proliferation, Apoptosis and Fuel Utilization Across Tissues. *Cells* **8**, (2019).
163. Kwapis, J. L. *et al.* HDAC3-mediated repression of the Nr4a family contributes to age-related impairments in long-term memory. *J. Neurosci.* **39**, 4999–5009 (2019).
164. Saijo, K. *et al.* A Nurr1/CoREST Pathway in Microglia and Astrocytes Protects Dopaminergic Neurons from Inflammation-Induced Death. *Cell* **137**, 47–59 (2009).
165. Kooistra, S. M. & Helin, K. Post-translational modifications: Molecular mechanisms and potential functions of histone demethylases. *Nat. Rev. Mol. Cell Biol.* **13**, 297–311 (2012).
166. Boutinaud, M., Ben Chedly, M. H., Delamaire, E. & Guinard-Flament, J. Milking and feed restriction regulate transcripts of mammary epithelial cells purified from milk. *J. Dairy Sci.* **91**, 988–998 (2008).
167. Boutinaud, M., Herve, L. & Lollivier, V. Mammary epithelial cells isolated from milk are a valuable, non-invasive source of mammary transcripts. *Front. Genet.* **6**, (2015).
168. Cifrian, E., Guidry, A. J., O'Brien, C. N., Keys, J. E. & Marquardt, W. W. Bovine mammary teat and ductal epithelial cell cultures. *Am. J. Vet. Res.* **55**, 239–246 (1994).
169. Simpson, R. J. Fragmentation of Protein Using Trypsin. *Cold Spring Harb. Protoc.* **2006**, pdb.prot4550-pdb.prot4550 (2006).
170. Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and P. W. Cell-Cell Adhesion. in *Molecular Biology of the Cell*. (Garland Science, 2002).
171. Saiki, R. K. *et al.* Enzymatic amplification of  $\beta$ -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science (80-. )*. **230**, 1350–1354 (1985).
172. Thermo Fisher Scientific Inc. Real-time PCR: understanding Ct. *Application Note Real-time PCR* <https://www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/PG1503-PJ9169-CO019879-Re-brand-Real-Time-PCR-Understanding-Ct-Value-Americas-FHR.pdf> (2016).
173. Raha, D., Hong, M. & Snyder, M. ChIP-Seq: A method for global identification of regulatory elements in the genome. *Curr. Protoc. Mol. Biol.* (2010) doi:10.1002/0471142727.mb2119s91.
174. Hanahan, D. Studies on transformation of Escherichia coli with plasmids. *J. Mol. Biol.* **166**, 557–

- 580 (1983).
175. Deans, T. L. Parallel networks: Synthetic biology and artificial intelligence. *ACM J. Emerg. Technol. Comput. Syst.* **11**, (2014).
176. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467 (1977).
177. Madden, T. The BLAST sequence analysis tool. in *The BLAST Sequence Analysis Tool* 1–17 (2013).
178. Active Motif Inc. ChIP-IT® ExpressMagnetic ChromatinImmunoprecipitation Kit & Sonication Shearing Kit (version F5). <https://www.activemotif.com/documents/1574.pdf> (2014).
179. DOUNCE, A. L., WITTER, R. F., MONTY, K. J., PATE, S. & COTTONE, M. A. A method for isolating intact mitochondria and nuclei from the same homogenate, and the influence of mitochondrial destruction on the properties of cell nuclei. *J. Biophys. Biochem. Cytol.* **1**, 139–153 (1955).
180. Folco, E. G., Lei, H., Hsu, J. L. & Reed, R. Small-scale nuclear extracts for functional assays of gene-expression machineries. *J. Vis. Exp.* 1–6 (2012) doi:10.3791/4140.
181. Andrews, S. FASTQC A Quality Control tool for High Throughput Sequence Data. *Babraham Institute* 2010 (2015).
182. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* vol. 32 3047–3048 (2016).
183. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* vol. 17 10 (2011).
184. Robinson, A. Quality Trim version 1.6.0. (2015).
185. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: A fast spliced aligner with low memory requirements. *Nature Methods* vol. 12 357–360 (2015).
186. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* vol. 25 2078–2079 (2009).
187. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* vol. 30 923–930 (2014).
188. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* vol. 15 (2014).
189. Risso, D., Schwartz, K., Sherlock, G. & Dudoit, S. GC-Content Normalization for RNA-Seq Data. (2011).
190. Kassambara, A. ggpubr: ‘ggplot2’ Based Publication Ready Plots. R package version 0.2. <https://CRAN.R-project.org/package=ggpubr> 153 (2020).
191. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* vol. 4 44–57 (2009).
192. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research* vol. 37 1–13 (2009).
193. Thomas, P. D. *et al.* PANTHER: A library of protein families and subfamilies indexed by function. *Genome Research* vol. 13 2129–2141 (2003).
194. Raudvere, U. *et al.* G:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research* vol. 47 W191–W198 (2019).
195. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS A Journal of Integrative Biology* vol. 16 284–287 (2012).
196. Luo, W. & Brouwer, C. Pathview: An R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* vol. 29 1830–1831 (2013).
197. Yu, G. enrichplot: Visualization of Functional Enrichment Result. R package version 1.8.1. (2020).
198. R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing: Vienna, Austria* online (2020).
199. Kolde, R. pheatmap: Pretty Heatmaps. R package. (2019).
200. Bolger, a. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible read trimming tool for Illumina NGS data. *Bioinformatics* vol. 30 2114–2120 (2014).
201. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler Transform.

- Bioinformatics* vol. 25 1754–1760 (2009).
202. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* vol. 26 589–595 (2010).
  203. Cunningham, F. *et al.* Ensembl 2022. *Nucleic Acids Res.* **50**, D988–D995 (2022).
  204. Van der Auwera, G., O'Connor, B. & Safari, an O. M. C. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. Genomics in the Cloud* (2020).
  205. Broad Institute. Picard toolkit. *Broad Institute, GitHub repository* (2019).
  206. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research* vol. 44 W160–W165 (2016).
  207. Stovner, E. B. & Sætrom, P. Epic2 efficiently finds diffuse domains in ChIP-seq data. *Bioinformatics* vol. 35 4392–4393 (2019).
  208. Stark, R. & Brown, G. DiffBind: differential binding analysis of ChIP-Seq peak data. *Bioconductor* 1–27 (2011).
  209. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* vol. 481 389–393 (2012).
  210. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell* vol. 38 576–589 (2010).
  211. Van de Peer, Y. Draw Venn Diagram: Calculate and Draw Custom Venn Diagrams. 1276–1277 <http://bioinformatics.psb.ugent.be/webtools/%0AVenn/> (2015).
  212. RStudio Team. RStudio: Integrated Development Environment for R. (2021).
  213. Wickham, H. ggplot2. *ggplot2* (2016) doi:10.1007/978-3-319-24277-4.
  214. Wickman, H., François, R., Henry, L. & Muller, K. dplyr: A Grammar of Data Manipulation. (2021).
  215. Wickham, H. *et al.* Welcome to the Tidyverse. *Journal of Open Source Software* vol. 4 1686 (2019).
  216. Wickham, H. tidy: Tidy Messy Data. R package version 1.1.3. (2021).
  217. Dowle, M. & Srinivasan, A. data.table: Extension of 'data.frame'. R package version 1.13.2. <https://CRAN.R-project.org/package=data.table>. (2021).
  218. Larsson, J. *et al.* eulerr: area-proportional euler and venn diagrams with ellipses. R package version 6.1.0. 1–18 (2020).
  219. Pedersen, T. L. patchwork: The Composer of Plots The Composer of Plots. R package version 1.1.1. *Cran* (2020).
  220. Wilke, C. O. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 0.7.0. URL: <https://CRAN.R-project.org/package=cowplot>. Accessed 18 October 2018 (2016).
  221. Turner, B. ChIP with Native Chromatin: Advantages and Problems Relative to Methods Using Cross-Linked Material. in *Mapping Protein/DNA Interactions by Cross-Linking* (2001).
  222. Das, P. M., Ramachandran, K., VanWert, J. & Singal, R. Chromatin immunoprecipitation assay. *Biotechniques* **37**, 1–9 (2004).
  223. O'Neill, L. P. & Turner, B. M. Immunoprecipitation of native chromatin: NChIP. *Methods* **31**, 76–82 (2003).
  224. Qiu, H. & Wang, Y. Exploring DNA-binding proteins with in vivo chemical cross-linking and mass spectrometry. *J. Proteome Res.* **8**, 1983–1991 (2009).
  225. Thavarajah, R., Mudimbaimannar, V. K., Elizabeth, J., Rao, U. K. & Ranganathan, K. Chemical and physical basics of routine formaldehyde fixation. *J. Oral Maxillofac. Pathol.* **16**, 400–405 (2012).
  226. Baranello, L., Kouzine, F., Sanford, S. & Levens, D. ChIP bias as a function of cross-linking time. *Chromosom. Res.* **24**, 175–181 (2016).
  227. Browne, J. A., Harris, A. & Leir, S. H. An optimized protocol for isolating primary epithelial cell chromatin for ChIP. *PLoS One* **9**, (2014).
  228. Weinmann, A. S. & Farnham, P. J. Identification of unknown target genes of human transcription factors using chromatin immunoprecipitation. *Methods* **26**, 37–47 (2002).
  229. O'Geen, H., Echipare, L. & Farnham, P. J. Using ChIP-seq technology to generate high-resolution profiles of histone modifications. *Methods Mol. Biol.* **791**, 265–286 (2011).
  230. Woo, Y. H. & Li, W. H. Evolutionary conservation of histone modifications in mammals. *Mol.*

- Biol. Evol.* **29**, 1757–1767 (2012).
231. Berguet, G. *et al.* Automating ChIP-seq experiments to generate epigenetic profiles on 10,000 HeLa Cells. *J. Vis. Exp.* (2014) doi:10.3791/52150.
232. Diagenode. H3K4me3 polyclonal antibody - Premium. [www.diagenode.com/en/p/h3k4me3-polyclonal-antibody-premium-50-ug-50-ul](http://www.diagenode.com/en/p/h3k4me3-polyclonal-antibody-premium-50-ug-50-ul) (2018).
233. Teich, N., Rosendahl, J., Tóth, M., Mössner, J. & Sahin-Tóth, M. Mutations of human cationic trypsinogen (PRSS1) and chronic pancreatitis. *Hum. Mutat.* **27**, 721–730 (2006).
234. Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457–466 (2007).
235. Solomon, E. R., Caldwell, K. K. & Allan, A. M. A novel method for the normalization of ChIP-qPCR data. *MethodsX* **8**, 101504 (2021).
236. Islam, M. A. *et al.* Transcriptome analysis of the inflammatory responses of bovine mammary epithelial cells: Exploring immunomodulatory target genes for bovine mastitis. *Pathogens* **9**, (2020).
237. Fu, Y. *et al.* Staphylococcus aureus and Escherichia coli elicit different innate immune responses from bovine mammary epithelial cells. *Vet. Immunol. Immunopathol.* **155**, 245–252 (2013).
238. Petzl, W. *et al.* Early transcriptional events in the udder and teat after intra-mammary Escherichia coli and Staphylococcus aureus challenge. *Innate Immun.* **22**, 294–304 (2016).
239. De Santa, F. *et al.* The Histone H3 Lysine-27 Demethylase Jmjd3 Links Inflammation to Inhibition of Polycomb-Mediated Gene Silencing. *Cell* **130**, 1083–1094 (2007).
240. De Santa, F. *et al.* Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. *EMBO J.* **28**, 3341–3352 (2009).
241. Yu, S. *et al.* The regulation of Jmjd3 upon the expression of NF- $\kappa$ B downstream inflammatory genes in LPS activated vascular endothelial cells. *Biochem. Biophys. Res. Commun.* **485**, 62–68 (2017).
242. Brand, K. S., Filor, V. & Bäumer, W. Early inflammatory events of mastitis—a pilot study with the isolated perfused bovine udder. *BMC Vet. Res.* **17**, 1–10 (2021).
243. Rose, M. T. *et al.* In vitro differentiation of a cloned bovine mammary epithelial cell. *J. Dairy Res.* **69**, 345–355 (2002).
244. Leimbach, A. *et al.* Complete genome sequences of Escherichia coli strains 1303 and ECC-1470 isolated from bovine mastitis. *Genome Announc.* **3**, (2016).
245. Leimbach, A. *et al.* No evidence for a bovine mastitis Escherichia coli pathotype. *BMC Genomics* **18**, 1–22 (2017).
246. Zaatout, N. An overview on mastitis-associated Escherichia coli: Pathogenicity, host immunity and the use of alternative therapies. *Microbiol. Res.* **256**, 126960 (2022).
247. Younis, S., Javed, Q. & Blumenberg, M. Meta-analysis of transcriptional responses to mastitis-causing Escherichia coli. *PLoS One* **11**, 1–18 (2016).
248. Védrine, M. *et al.* Sensing of Escherichia coli and LPS by mammary epithelial cells is modulated by O-antigen chain and CD14. *PLoS One* **13**, (2018).
249. Strunk, T. *et al.* Method of bacterial killing differentially affects the human innate immune response to Staphylococcus epidermidis. *Innate Immun.* **17**, 508–516 (2011).
250. Tartaglia, N. R. *et al.* Staphylococcus aureus extracellular vesicles elicit an immunostimulatory response in vivo on the murine mammary gland. *Front. Cell. Infect. Microbiol.* **8**, (2018).
251. Wang, K., Wei, G. & Liu, D. CD19: a biomarker for B cell development, lymphoma diagnosis and therapy. *Exp. Hematol. Oncol.* **1**, 1–7 (2012).
252. Cai, X. *et al.* Down-regulation of FN1 inhibits colorectal carcinogenesis by suppressing proliferation, migration, and invasion. *J. Cell. Biochem.* **119**, 4717–4728 (2018).
253. Sponziello, M. *et al.* Fibronectin-1 expression is increased in aggressive thyroid cancer and favors the migration and invasion of cancer cells. *Mol. Cell. Endocrinol.* **431**, 123–132 (2016).
254. Lou, X. *et al.* SOX2 targets fibronectin 1 to promote cell migration and invasion in ovarian cancer: New molecular leads for therapeutic intervention. *Omi. A J. Integr. Biol.* **17**, 510–518 (2013).
255. Wu, J., Weening, E. H., Faske, J. B., Höök, M. & Skare, J. T. Invasion of eukaryotic cells by Borrelia burgdorferi requires  $\beta$ 1 integrins and Src kinase activity. *Infect. Immun.* **79**, 1338–1348 (2011).
256. Woodward, T. L., Turner, J. D., Hung, H. T. & Zhao, X. Inhibition of cellular proliferation and



- modulation of insulin-like growth factor binding proteins by retinoids in a bovine mammary epithelial cell line. *J. Cell. Physiol.* **167**, 488–499 (1996).
257. Gibson, C. A., Staley, M. D. & Baumrucker, C. R. Identification of IGF binding proteins in bovine milk and the demonstration of IGFBP-3 synthesis and release by bovine mammary epithelial cells. *J. Anim. Sci.* **77**, 1547–1557 (1999).
258. Beattie, J., Hawsawi, Y., Alkharobi, H. & El-Gendy, R. IGFBP-2 and -5: important regulators of normal and neoplastic mammary gland physiology. *J. Cell Commun. Signal.* **9**, 151–158 (2015).
259. Silvestrelli, G., Ulbrich, S. E. & Saenz-de-Juano, M. D. Assessing extracellular vesicles from bovine mammary gland epithelial cells cultured in FBS-free medium. *Extracell. Vesicles Circ. Nucleic Acids* 252–267 (2021) doi:10.20517/evcna.2021.18.
260. Becker, D., Weikard, R., Hadlich, F. & Kühn, C. Single-cell RNA sequencing of freshly isolated bovine milk cells and cultured primary mammary epithelial cells. *Sci. Data* **8**, 1–9 (2021).
261. Rainard, P. *et al.* The chemokine CXCL3 is responsible for the constitutive chemotactic activity of bovine milk for neutrophils. *Mol. Immunol.* **45**, 4020–4027 (2008).
262. Zhou, Y. *et al.* Caveolin-1 and cavin1 act synergistically to generate a unique lipid environment in caveolae. *J. Cell Biol.* **220**, (2021).
263. Tillu, V. A. *et al.* Cavin1 intrinsically disordered domains are essential for fuzzy electrostatic interactions and caveola formation. *Nat. Commun.* **12**, 1–18 (2021).
264. Sotgia, F., Schubert, W., Pestell, R. G. & Lisanti, M. P. Genetic ablation of caveolin-1 in mammary epithelial cells increases milk production and hyper-activates STAT5a signaling. *Cancer Biol. Ther.* **5**, 292–297 (2006).
265. Wang, H., Pilch, P. F. & Liu, L. Cavin-1/PTRF mediates insulin-dependent focal adhesion remodeling and ameliorates high-fat diet-induced inflammatory responses in mice. *J. Biol. Chem.* **294**, 10544–10552 (2019).
266. Park, D. S. *et al.* Prolactin Negatively Regulates Caveolin-1 Gene Expression in the Mammary Gland during Lactation, via a Ras-dependent Mechanism. *J. Biol. Chem.* **276**, 48389–48397 (2001).
267. Akira, S. *et al.* A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. *EMBO J.* **9**, 1897–1906 (1990).
268. Tian, L. X. *et al.* Cytochrome P450 1A1 enhances inflammatory responses and impedes phagocytosis of bacteria in macrophages during sepsis. *Cell Commun. Signal.* **18**, (2020).
269. Zhang, W. Y. *et al.* CYP1A1 relieves lipopolysaccharide-induced inflammatory responses in bovine mammary epithelial cells. *Mediators Inflamm.* **2018**, (2018).
270. Kuhn, M. J., Putman, A. K. & Sordillo, L. M. Widespread basal cytochrome P450 expression in extrahepatic bovine tissues and isolated cells. *J. Dairy Sci.* **103**, 625–637 (2020).
271. Demirel, I., Savaş, S., Kruse, R. & Persson, K. Expression of suppressor of cytokine signalling 3 (SOCS3) in human bladder epithelial cells infected with uropathogenic Escherichia coli. *Apmis* **121**, 158–167 (2013).
272. Koy, M. *et al.* Recombinant bovine S100A8 and A9 enhance IL-1 $\beta$  secretion of interferon-gamma primed monocytes. *Vet. Immunol. Immunopathol.* **155**, 162–170 (2013).
273. Cuesta, L. M., Liron, J. P., Farias, M. V. N., Dolcini, G. L. & Ceriani, M. C. Effect of bovine leukemia virus (BLV) infection on bovine mammary epithelial cells RNA-seq transcriptome profile. *PLoS One* **15**, (2020).
274. Ying, Y. T. *et al.* Escherichia coli and staphylococcus aureus differentially regulate Nrf2 pathway in bovine mammary epithelial cells: Relation to distinct innate immune response. *Cells* **10**, (2021).
275. Daniel, E. *et al.* Cyp26b1 is an essential regulator of distal airway epithelial differentiation during lung development. *Dev.* **147**, (2020).
276. Chandrabalan, A. & Ramachandran, R. Molecular mechanisms regulating Proteinase-Activated Receptors (PARs). *FEBS J.* **288**, 2697–2726 (2021).
277. Kuramochi, S. *et al.* LOK is a novel mouse STE20-like protein kinase that is expressed predominantly in lymphocytes. *J. Biol. Chem.* **272**, 22679–22684 (1997).
278. Kolar, S. L. *et al.* Extracellular proteases are key mediators of Staphylococcus aureus virulence via the global modulation of virulence-determinant stability. *Microbiologyopen* **2**, 18–34 (2013).
279. Chen, Y. *et al.* Systematic evaluation of factors influencing ChIP-seq fidelity. *Nat. Methods* **9**,

- 609–614 (2012).
280. Jung, Y. L. *et al.* Impact of sequencing depth in ChIP-seq experiments. *Nucleic Acids Res.* **42**, 1–10 (2014).
281. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, (2009).
282. Landt, S. G. *et al.* ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* **22**, 1813–1831 (2012).
283. Zhang, Q. & Keleş, S. CNV-guided multi-read allocation for ChIP-seq. *Bioinformatics* **30**, 2860–2867 (2014).
284. Chung, D. *et al.* Discovering transcription factor binding sites in highly repetitive regions of genomes with multi-read analysis of ChIP-seq data. *PLoS Comput. Biol.* **7**, (2011).
285. Bailey, T. *et al.* Practical Guidelines for the Comprehensive Analysis of ChIP-seq Data. *PLoS Comput. Biol.* **9**, 5–12 (2013).
286. Flensburg, C., Kinkel, S. A., Keniry, A., Blewitt, M. E. & Oshlack, A. A comparison of control samples for ChIP-seq of histone modifications. *Front. Genet.* **5**, (2014).
287. Diaz, A., Nellore, A. & Song, J. S. CHANCE: comprehensive software for quality control and validation of ChIP-seq data. *Genome Biol.* **13**, R98 (2012).
288. Andersson, L. *et al.* Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. *Genome Biol.* **16**, 4–9 (2015).
289. Nakato, R. & Shirahige, K. Sensitive and robust assessment of ChIP-seq read distribution using a strand-shift profile. *Bioinformatics* **34**, 2356–2363 (2018).
290. Fang, L. *et al.* Functional annotation of the cattle genome through systematic discovery and characterization of chromatin states and butyrate-induced variations. *BMC Biol.* **17**, 1–16 (2019).
291. Nakato, R. & Sakata, T. Methods for ChIP-seq analysis: A practical workflow and advanced applications. *Methods* **187**, 44–53 (2021).
292. Zang, C. *et al.* A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952–1958 (2009).
293. Young, M. D. *et al.* ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. *Nucleic Acids Res.* **39**, 7415–7427 (2011).
294. Park, P. J. ChIP-seq: Advantages and challenges of a maturing technology. *Nat. Rev. Genet.* **10**, 669–680 (2009).
295. Eder, T. & Grebien, F. Comprehensive assessment of differential ChIP-seq tools guides optimal algorithm selection. *Genome Biol.* **23**, 1–27 (2022).
296. Yang, Y. & Wilson, M. J. Genome-wide analysis of H3K4me3 and H3K27me3 modifications throughout the mouse urogenital ridge at E11.5. *Gene Reports* **16**, (2019).
297. Baker, K. *et al.* Chromatin state analysis of the barley epigenome reveals a higher-order structure defined by H3K27me1 and H3K27me3 abundance. *Plant J.* **84**, 111–124 (2015).
298. Sheffield, N. C. & Bock, C. LOLA: Enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinformatics* **32**, 587–589 (2016).
299. Roager, H. M. & Licht, T. R. Microbial tryptophan catabolites in health and disease. *Nat. Commun.* **9**, 1–10 (2018).
300. Taracha-Wisniewska, A., Kotarba, G., Dworkin, S. & Wilanowski, T. Recent discoveries on the involvement of Krüppel-like factor 4 in the most common cancer types. *Int. J. Mol. Sci.* **21**, 1–29 (2020).
301. Li, Z. *et al.* Identification and characterization of novel and differentially expressed microRNAs in peripheral blood from healthy and mastitis Holstein cattle by deep sequencing. *Anim. Genet.* **45**, 20–27 (2014).
302. Kuhn, A. R. *et al.* MicroRNA expression in human airway smooth muscle cells: Role of miR-25 in regulation of airway smooth muscle phenotype. *Am. J. Respir. Cell Mol. Biol.* **42**, 506–513 (2010).
303. Ren, N., Atyah, M., Chen, W. Y. & Zhou, C. H. The various aspects of genetic and epigenetic toxicology: Testing methods and clinical applications. *J. Transl. Med.* **15**, 1–13 (2017).
304. Garcin, C. & Straube, A. Microtubules in cell migration. *Essays Biochem.* **63**, 509–520 (2019).
305. Vasileva, E. & Citi, S. The role of microtubules in the regulation of epithelial junctions. *Tissue Barriers* **6**, 1–20 (2018).
306. Choi, S. W. *et al.* MicroRNA profiling in bovine serum according to the stage of Mycobacterium

- avium subsp. paratuberculosis infection. *PLoS One* **16**, 1–16 (2021).
307. Li, R. *et al.* Transcriptome microRNA profiling of bovine mammary glands infected with *Staphylococcus aureus*. *Int. J. Mol. Sci.* **16**, 4997–5013 (2015).
308. Do, D. N., Li, R., Dudemaine, P. L. & Ibeagha-Awemu, E. M. MicroRNA roles in signalling during lactation: An insight from differential expression, time course and pathway analyses of deep sequence data. *Sci. Rep.* **7**, (2017).
309. Wang, S. *et al.* Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nat. Protoc.* **8**, 2502–2515 (2013).
310. Ranga Niroshan Appuhamy, J. A. D., Knoebel, N. A., Deepthi Nayananjali, W. A., Escobar, J. & Hanigan, M. D. Isoleucine and leucine independently regulate mTOR signaling and protein synthesis in MAC-T cells and bovine mammary tissue slices. *J. Nutr.* **142**, 484–491 (2012).
311. Cabezas-Cruz, A., Estrada-Peña, A., Rego, R. O. M. & De la Fuente, J. Tick-pathogen Ensembles: Do molecular interactions lead ecological innovation? *Front. Cell. Infect. Microbiol.* **7**, 1–5 (2017).
312. Silva, L. G., Ferguson, B. S. & Faciola, A. P. Rapid Communication: Prolactin and hydrocortisone impact TNF $\alpha$ -mediated mitogen-activated protein kinase signaling and inflammation of bovine mammary epithelial (MAC-T) cells. *J. Anim. Sci.* **95**, 5524–5531 (2017).
313. German, T. & Barash, I. Characterization of an epithelial cell line from bovine mammary gland. *Vitr. Cell. Dev. Biol. - Anim.* **38**, 282–292 (2002).
314. Bauer, I., Günther, J., Wheeler, T. T., Engelmann, S. & Seyfert, H. M. Extracellular milieu grossly alters pathogen-specific immune response of mammary epithelial cells. *BMC Vet. Res.* **11**, 1–14 (2015).
315. Pan, C., Kumar, C., Bohl, S., Klingmueller, U. & Mann, M. Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. *Mol. Cell. Proteomics* **8**, 443–450 (2009).
316. Cazaly, E. *et al.* Making sense of the epigenome using data integration approaches. *Front. Pharmacol.* **10**, 1–15 (2019).
317. Clark, E. L. *et al.* From FAANG to fork: application of highly annotated genomes to improve farmed animal production. *Genome Biol.* **21**, 1–9 (2020).
318. Ou, J., Wolfe, S. A., Brodsky, M. H. & Zhu, L. J. MotifStack for the analysis of transcription factor binding site evolution. *Nat. Methods* **15**, 8–9 (2018).
319. Feehley, T., O'Donnell, C. W., Mendlein, J., Karande, M. & McCauley, T. Drugging the epigenome in the age of precision medicine. *Clin. Epigenetics* **15**, 1–13 (2023).
320. Barros da Silva, P., Coelho, M., Bidarra, S. J., Neves, S. C. & Barrias, C. C. Reshaping in vitro Models of Breast Tissue: Integration of Stromal and Parenchymal Compartments in 3D Printed Hydrogels. *Front. Bioeng. Biotechnol.* **8**, 1–15 (2020).



## Appendix

### A Supplementary Materials and Methods

#### A1. Primer Lists

Table A-1: First round of semi-nested PCR for amplifying DNA insert spanning immune response gene fragment for cloning.

Internal ID#	Primer Name	Sequence	Amplicon Size (bp)	Primer T <sub>m</sub> (°C)
N051	BT_CCL20_ChIP_A_f	TGGGGCTGATCTTTGTACCACT	2600	61
N111	BT_CCL20_r1	CCCACTTCTTCTTTGGATCTGC		
N039	BT_CXCL8_ChIP_A_f	TCCTGGTTCTTATATTGCTTTCTTC	2566	57
N132	BT_IL8_r3	GGCCCACTCTCAATAACTCTC		
N140	BT_IL6_pm_f	TCAGGGTCTTCACGGGGAAC	4559	61
N141	BT_IL6_pm_r	CTCACAGCAGCGGACACAC		
N093	BT_TNF_ChIP_A_f	TGAGGGGCAGAGGGGAAATG	1546	63
N031	BT_TNF_neu_r1	TGGGCTACCGCTTGTTACT		
N119	BT_LAP_ChIP_A_f	GCAGATTCAAAGCCCTAAGCG	2046	63
N117	LAP_sr	CCTGCAGCATTTTACTTGGGCT		
N138	BT_NOS2_ChIP_I_f	TGGTACCCTCTCCCCTGTCA	1715	63
N139	BT_NOS2_ChIP_I_r	AAGGGTGGCTGAGAGGTTTCA		

Table A-2: Second round of semi-nested PCR for amplifying DNA insert spanning immune response gene fragment for cloning.

Internal ID#	Primer Name	Sequence	Amplicon Size (bp)	Primer T <sub>m</sub> (°C)
N051	BT_CCL20_ChIP_A_f	TGGGGCTGATCTTTGTACCACT	419	63
N060	BT_CCL20_ChIP_E_r	TGCTGCACATCATTTTCAACTCG		
N039	BT_CXCL8_ChIP_A_f	TCCTGGTTCTTATATTGCTTTCTTC	858	59
N050	BT_CXCL8_ChIP_F_r	AGGAAAGCTGCCAAGAGAGC		
N077	BT_IL6_ChIP_A_f	TGAGGAGGAAGCTGGCAGAT	733	61
N092	BT_IL6_ChIP_H_r	CCCGAGGGCAGACTGAGTC		
N093	BT_TNF_ChIP_A_f	TGAGGGGCAGAGGGGAAATG	415	61
N100	BT_TNF_ChIP_D_r	GGGTGTGTTGTCTGGGAGGT		
N119	BT_LAP_ChIP_A_f	GCAGATTCAAAGCCCTAAGCG	395	63
N124	BT_LAP_ChIP_C_r	AAGGAGCAGGTGATGGAGCC		
N061	BT_NOS2_ChIP_A_f	ACCCTCTCCCCTGTCAACCTT	1685	61
N076	BT_NOS2_ChIP_H_r	GTGACACTGAGTTCTCTAGGGCT		

Table A-3: N-ChIP-qPCR Primers. The list includes primer information about primer name, sequence, amplicon size, primer melting temperature (Primer T<sub>m</sub>), product T<sub>m</sub> targeting promoter regions of immune response genes chemokine (C-X-C motif) ligand 8 (IL8/CXCL8), chemokine (C-C motif) ligand 20 (CCL20), interleukin 6 (IL6), tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), Actin beta (ACTB) and histone cluster 1 H2ba (TSH2B). Primers were designed with Primer 3 software. Primer melting temperatures were tested with PCR before use in qPCR. Product T<sub>m</sub> was developed in silico (<http://mbcf.dfci.harvard.edu/docs/oligocalc.html>).

Internal ID#	Primer Name	Sequence	Amplicon Size (bp)	Primer T <sub>m</sub> (°C)	Product T <sub>m</sub> (°C)
N039	BT_CXCL8_ChIP_A_f	TCCTGGTTCTTATATTGCTTTCTTC	95	56	73
N040	BT_CXCL8_ChIP_A_r	CTGAGTAATGTGGGAATCTGGAGAC			
N041	BT_CXCL8_ChIP_B_f	TGAGAGTGATAATTGTTCTCTGGCA	69	58	71
N042	BT_CXCL8_ChIP_B_r	GGGACACAATTTACAAGAAGTCCA			
N043	BT_CXCL8_ChIP_C_f	CAAAGCTTGGGTCACACAGAA	91	57	76
N044	BT_CXCL8_ChIP_C_r	GCAACCATCCGCCCTTTGAG			
N045	BT_CXCL8_ChIP_D_f	CTCAAAGGGCGGATGGTTGC	119	60	78
N046	BT_CXCL8_ChIP_D_r	TGTTGGGCAGAGAACTCCTGT			

N049	BT_CXCL8_ChIP_F_f	GGACAGCAGAGCTCACAAGC	111	60	78
N050	BT_CXCL8_ChIP_F_r	AGGAAAGCTGCCAAGAGAGC			
N051	BT_CCL20_ChIP_A_f	TGGGGCTGATCTTTGTACCACT	101	58	79
N052	BT_CCL20_ChIP_A_r	TAGCTTTGCCCTGTGTCCCA			
N053	BT_CCL20_ChIP_B_f	GGGACACAGGGCAAAGCTAC	103	59	79
N054	BT_CCL20_ChIP_B_r	AGTCAGGAAGGTGTGAAGGTC			
N055	BT_CCL20_ChIP_C_f	AGGGAAAGGCAGGAAGTTTTC	90	57	77
N056	BT_CCL20_ChIP_C_r	TGTTCTCACGTGGGGTTTTC			
N057	BT_CCL20_ChIP_D_f	GACAACATACTGGGGCGAGG	97	57	77
N058	BT_CCL20_ChIP_D_r	GGATGGCCCTACTTATAGCAAA			
N059	BT_CCL20_ChIP_E_f	CCCAGGCTGCTATCAAACATCAC	92	59	79
N060	BT_CCL20_ChIP_E_r	TGCTGCACATCATTTTCAACTCG			
N077	BT_IL6_ChIP_A_f	TGAGGAGGAAGCTGGCAGAT	84	59	80
N078	BT_IL6_ChIP_A_r	TGTGGACGTTCTCTAGGAGTCT			
N079	BT_IL6_ChIP_B_f	TTCTGCAACAGCTGGCCAC	91	59	79
N080	BT_IL6_ChIP_B_r	TCTCCATCCCTCCACTCCCT			
N081	BT_IL6_ChIP_C_f	TTATTTGTAATCCCAGGCCAGCA	70	57	76
N082	BT_IL6_ChIP_C_r	AAACCTAAGATGTTCTGACCTCA			
N083	BT_IL6_ChIP_D_f	CACCCGCTACCCCTCTAAAA	82	56	73
N084	BT_IL6_ChIP_D_r	TTTAGTGACTCAGCACATTAGC			
N085	BT_IL6_ChIP_E_f	CATGCTAATGTGCTGAGTCACT	125	57	76
N086	BT_IL6_ChIP_E_r	TCACAGGAATCAAACCAGGGC			
N087	BT_IL6_ChIP_F_f	TCCTGTGATGCTAAAGGACG	60	58	71
N088	BT_IL6_ChIP_F_r	TGGGGCTGATTGGAACCTTA			
N091	BT_IL6_ChIP_H_f	CCCCACCCTCCAACAAAGATT	118	59	79
N092	BT_IL6_ChIP_H_r	CCCGAGGGCAGACTGAGTC			
N093	BT_TNF_ChIP_A_f	TGAGGGGCAGAGGGGAAATG	107	59	84
N094	BT_TNF_ChIP_A_r	TACTTCCTACACTCCGCCCG			
N095	BT_TNF_ChIP_B_f	CGGGCGGAGTGTAGGAAGTA	82	59	84
N096	BT_TNF_ChIP_B_r	CGGTTTCTTCTCCATCGCGG			
N097	BT_TNF_ChIP_C_f	GAATGAGAGCCTTTCCCCGC	86	59	79
N098	BT_TNF_ChIP_C_r	CTCTGGGAGCTTCTGTTGGC			
N099	BT_TNF_ChIP_D_f	GCCAACAGAAGCTCCCAGAG	100	60	84
N100	BT_TNF_ChIP_D_r	GGGTGTGTTGTCTGGGAGGT			
N142	BT_TNF_ChIP_E_f	CCGCGATGGAGAAGAAACCG	120	61	82
N143	BT_TNF_ChIP_E_r	GCGGGGAAAGGCTCTCATTC			
N101	BT_ACTB_ChIP_f	GCCGTTTCGAAAGTTGCCTT	63	59	80
N102	BT_ACTB_ChIP_r	CGCTGGGTTATATAGGGCGC			
N103	BT_TSH2B_ChIP_f	TGACGGTAGGTTGTGCTTTGC	107	60	78
N104	BT_TSH2B_ChIP_r	TGGATACCTGGGAGTGTCATGT			

## A2. Cell Culture Buffer

### HBSS Buffer

900 mL ultra-pure water  
100 mL 10x HBSS  
17 mL HEPES

### PBS

9.55 g PBS  
Fill to 1 L with *Aqua bidest.*

### Fixation Buffer (X-ChIP)

5 mL 1 M NaCl  
0.5 mL 100 mM EDTA  
0.5 mL 50 mM EGTA  
2.5 mL 1 M HEPES  
41.5 mL autoclaved *Aqua bidest.*  
Adjust pH to 7.6.  
Sterile filter before use

## A3. Molecular Biological Buffer

### 1x FA Gel Running Buffer

100 mL 10x FA Gel Buffer  
20 mL 37% Formaldehyde  
Fill to 1 L with DEPC-H<sub>2</sub>O

### 10x FA Gel Buffer

200 mL 1 M MOPS  
50 mL 1 M Sodium acetate  
20 mL 0.5 mL EDTA  
Fill to 1 L with DEPC-H<sub>2</sub>O

### DEPC-H<sub>2</sub>O

1 mL DEPC  
Fill to 1 L with *Aqua bidest.*

RNA Stopp Buffer(RNA Loading dye)

1.6 µL Bromphenolblue  
 8 µL 500 mM EDTA (pH 8.0)  
 72 µL 37% Formaldehyde  
 200 µL Glycerol  
 308.4 µL Formamide  
 400 µL 10x FA Gel Buffer  
 10 µL DEPC-H<sub>2</sub>O

Stopp Buffer(DNA loading dye)

50 mL Glycerol  
 1 mL Bromphenolblau  
 24 mL 0.5 mM EDTA  
 1 mL 1 M Tris  
 Fill to 100 mL with *Aqua bidest.*

TAE-Puffer

242.48 g 2 M Tris  
 37.2 g 100 mM EDTA  
 Adjust pH to 8.0  
 Fill to 1 L with *Aqua bidest.*

**A4. Microbiological Working Solutions**

For microbiological work, the following growth broths were prepared (list below). Trypton-Yeast (TY) agar was weighed, dissolved, and autoclaved. After cooling followed the addition of the indicated components from sterile stock solutions. Agar plates with different antibiotic supplements were prepared under aseptic working conditions in a biosafety cabinet and poured to grow bacteria selectively. TY agar plates were supplemented with tetracycline (final concentration 15 µg/mL) for preparation of competent cells or with X-gal (4%), IPTG (0.04 mM), and ampicillin (final concentration 100 µg/mL) for cloning with TransformAID Bacterial Transformation Kit (Thermo Fisher Scientific). Other TY-agar plates with only ampicillin (final concentration 100 µg/mL) were used for CloneJET PCR Cloning Kit (Thermo Fisher Scientific).

TY-Agar

16 g Trypton  
 10 g yeast extract  
 5 g NaCl  
 20 g Agar agar  
 Fill to 1 L with *Aqua bidest.*

BHI broth

18.5 g BHI broth  
 Fill to 500 mL with *Aqua bidest.*

BHI broth agar

18.5 g BHI broth  
 10 g Agar agar  
 Fill to 500 mL with *Aqua bidest.*

TY-Medium

16 g Trypton  
 10 g yeast extract  
 5 g NaCl  
 Fill to 1 L with *Aqua bidest.*

SOC-Medium

2 g Trypton  
 0.5 g yeast extract  
 Fill to 100 mL with *Aqua bidest.*  
 Autoclave, add  
 0.05 g NaCl  
 0.019 g KCl  
 0.5 mL sterile 2M MgCl<sub>2</sub>  
 0.5 mL sterile 2M MgSO<sub>4</sub>  
 1 g Glucose

**A5. Buffers for Preparation of Competent Cells**RF1 solution\*

1.21 g RbCl  
 0.809 g MnCl<sub>2</sub>  
 0.6 mL 5M KAc (pH 7,5)  
 0.15 g CaCl<sub>2</sub>  
 15% Glycerol  
 Adjust pH to 5.8  
 Fill to 100 mL with *Aqua bidest.*  
 Sterile filter solution (0.22 µm pore size)

RF2 solution\*

1 mL 1M MOPS  
 0.12 g RbCl  
 1.125 g CaCl<sub>2</sub>  
 15% Glycerol  
 Adjust pH 6.8  
 Fill to 100 mL with *Aqua bidest.*  
 sterile filter

\*Sterile filtered RF1 and RF2 kept at -20°C

SOB-Broth

2 g Trypton  
 0.5 g Yeast extract  
 0.05 g NaCl  
 0.019 g KCl  
 Fill to 99 mL with *Aqua bidest.*  
 Autoclave solution, add  
 0.5 mL sterile 2M MgCl<sub>2</sub>  
 0.5 mL sterile 2M MgSO<sub>4</sub>

**Psi Broth**

2 g Trypton

0.5 g Yeast extract

Fill to 100 mL

Autoclave

Add 1 mL sterile 2M MgSO<sub>4</sub>

0.058 g NaCl

0.038 g KCl

Adjust pH to 7.6

**A6. Cell Culture Materials***Table A-4: Material list cell culture.*

Item	Company
Beakers	Simax, Sazava, CZ
Cell culture plates (collagen-coated, 10 cm)	Greiner Bio-One, Kremsmünster, AT
Cell scraper	VWR, Darmstadt, DE
Cellstar® 6Well Plate	Greiner Bio-One, Kremsmünster, AT
Cellstar® cell culture vessels	Greiner Bio-One, Kremsmünster, AT
Cryotube (2 mL)	A. Hartenstein, Würzburg, DE
Cutting boards made of polyethylene	Behr, Düsseldorf, DE
Falcon sieves (100 µm)	Falcon, Corning, New York, USA
Forceps (normal and surgical)	N/A
Funnel	N/A
Knife	Feather, Osaka, JP
Mr. Frosty™ Freezing Container	ThermoFisher Scientific, Waltham, USA
Pasteur glass pipettes	Carl Roth, Karlsruhe, DE
Pipette tips (10 µL, 50 µL, 200 µL, 1 mL, 5 mL)	Sarstedt, Nümbrecht, DE
Pipettes	Starlab, Hamburg, DE
Polypropylen tubes (15 mL, 50 mL)	Greiner Bio-One, Kremsmünster, AT
Serological pipette (5 mL, 10 mL, 25 mL)	Sarstedt, Nümbrecht, DE
Sieve tower (500 µm, 300 µm, 150 µm and 90 µm)	Carl Roth, Karlsruhe, DE
Surgical disposable scalpels	B.Braun, Tuttlingen, DE

**A7. Molecular Biological Materials***Table A-5: Consumables for molecular biological work.*

Item	Company
Reaction tubes (0.5 mL, 1.5 mL, 2 mL)	Sarstedt, Nümbrecht, DE
Light Cycler Capillaries (20 µL)	Roche Life Science, Basel, CH
LightCycler®480 Multiwell Plate 96, clear	Roche Life Science, Basel, CH
LightCycler®480 Sealing Foil	Roche Life Science, Basel, CH
Pipette tips (10 µL to 5 mL)	Sarstedt, Nümbrecht, DE
Reaction tubes	Carl Roth, Karlsruhe, DE

*Table A-6: Kit and reagents used for molecular biological work.*

Kit and Reagent Name	Company
<b><u>Nucleic acid extraction kit:</u></b>	
Master Pure™ DNA Purification Kit	Lucigen, Middleton, USA
NucleoSpin® Extract II	Macherey Nagel, Düren, DE
NucleoSpin Gel and PCR clean-up	Macherey Nagel, Düren, DE
RNeasy® Plus Mini Kit	Qiagen, Hilden, DE



RNeasy® MinElute® Cleanup (Qiagen)	Qiagen, Hilden, DE
High Pure PCR Product Purification Kit	Roche Life Science, Basel, CH
Direct-zol™ RNA MiniPrep	Zymo Research, Irvine, USA
ChIP DNA Clean and Concentrator	Zymo Research, Irvine, USA
QIAquick PCR Purification Kit	Qiagen, Hilden, DE
<b><u>Molecular biological kits and reagents:</u></b>	
Ambion Rnase I	Thermo Fisher Scientific, Waltham, USA
ChIP-IT® Express Enzymatic kit	Active Motif, Carlsbad, USA
DNA 1000 Kit	Agilent, Santa Clara, USA
dNTP Mix (10 mM)	Thermo Fisher Scientific, Waltham, USA
GoTaq® DNA Polymerase	Promega, Madison, USA
GoTaq® DNA Polymerase and Master Mix	Promega, Madison, USA
Kapa SYBR® Fast qPCR Mastr Mix	Kapa Biosystems, Wilmington, USA
Light Cyclers Fast Start DNA Master Plus SYBR Green I	Roche Life Science, Basel, CH
LightCycler® 480 SYBR Green I Master	Roche Life Science, Basel, CH
Pfu DNA Polymerase, recombinant	Thermo Fisher Scientific, Waltham, USA
Qubit™ dsDNA Broad Range Assay	Thermo Fisher Scientific, Waltham, USA
Qubit™ HS DNA Assay	Thermo Fisher Scientific, Waltham, USA
Qubit™ RNA HS Assay Kit	Thermo Fisher Scientific, Waltham, USA
Rnase T1	Thermo Fisher Scientific, Waltham, USA
SuperScript™II Reverse Transcriptase	Thermo Fisher Scientific, Waltham, USA
TruSeq RNA sample preparation kit set A	Illumina, San Diego, USA
UltraPure™ DNase/RNase-Free Distilled Water	Thermo Fisher Scientific, Waltham, USA

## A8. Microbiological Materials

Table A-7: Material List for microbiological work

Item	Company
Glass culture tubes	N/A
Glassware Erlenmeyer flask (2 L, 500 mL)	Simax, Sazava,CZ
inoculation loop	N/A
Nunc™ Petri Dishes	Thermo Fisher Scientific, Waltham, USA
Sterile filter 0.22 µm	Carl Roth, Karlsruhe, DE
Omnifix® Solo 50 mL	B. Braun, Melsungen, DE

Table A-8: Kits and competent cells for microbiological work.

Item	Company
Clone Jet PCR cloning Kit	Thermo Fisher Scientific, Waltham, USA
Transform Aid Kit	Thermo Fisher Scientific, Waltham, USA
XL1-Blue MRF' Supercompetent Cells	Stratagene, La Jolla, USA

## A9. Cell Culture Substances

Table A-9: Substances and liquid components for cell culture work.

Item	Company
<b><u>Cell culture media:</u></b>	
DMEM with 4.5 g/L glucose without L-glutamine	Lonza, Basel, CH
RPMI 1640	Biochrom, Berlin, DE
<b><u>Cell culture media supplements, antibiotics, and reagents:</u></b>	

10x Hank's Balanced Salt Solution (HBSS)	PAN-Biotech, Aidenbach, DE
Amphotericin B (250 µg/mL)	Biochrom, Berlin, DE
Collagenase NB 4G Proved Grade	Serva, Heidelberg, DE
Dimethyl sulfoxide (DMSO)	Carl Roth, Karlsruhe, DE
Fetal Bovine Serum (FBS)	Biochrom, Berlin, DE
Fetal Calf Serum (FCS)	PAN-Biotech, Aidenbach, DE
Hydroxyethylpiperazine ethane sulfonic acid (HEPES)	Biochrom, Berlin, DE
Isopropanol	Carl Roth, Karlsruhe, DE
L-glutamine	Biochrom, Berlin, DE
L-lysine	Sigma Aldrich, Steinheim, DE
L-methionine	Sigma Aldrich, Steinheim, DE
Penicillin/Streptomycin (P/S)	Biochrom, Berlin, DE
Phosphate Buffered Saline (PBS)	Biochrom, Berlin, DE
Sodium pyruvate	Sigma Aldrich, Steinheim, DE
Trypan Blue Solution	Thermo Fisher Scientific, Waltham, USA
Trypsin/Ethylenediaminetetraacetate (EDTA)	Biochrom, Berlin, DE
Ultra Pure Water	Biochrom, Berlin, DE
<u>Hormones for primary cell culture:</u>	
Hydrocortisone	Sigma Aldrich, Steinheim, DE
Insulin	Sigma Aldrich, Steinheim, DE
Prolactin	Sigma Aldrich, Steinheim, DE

## A10. Molecular Biological and Microbiological Substances

Table A-10: Substances for microbiological work.

Item	Company
10x PBS (from ChIP kit)	Active Motif, Carlsbad, USA
37% Formaldehyde	Carl Roth, Karlsruhe, DE
95-100% Ethanol	Carl Roth, Karlsruhe, DE
Agar agar Kobe I	Carl Roth, Karlsruhe, DE
Agarose	Biodeal, Markkleeberg, DE
Ampicillin	AppliGene, Illkirch-Graffenstaden, DE
Bovine Serum Albumin	Sigma Aldrich, Steinheim, DE
Brain-Heart-Infusion (BHI) Broth	Carl Roth, Karlsruhe, DE
Bromophenol blue	Serva, Heidelberg, DE
C646	Sigma Aldrich, Steinheim, DE
Calcium chloride	Carl Roth, Karlsruhe, DE
Chloroform Carl	Carl Roth, Karlsruhe, DE
Diethyl pyrocarbonate (DEPC)	Applichem, Darmstadt, DE
EDTA	Carl Roth, Karlsruhe, DE
EGTA	Sigma Aldrich, Steinheim, DE
Ethidium bromide	Applichem, Darmstadt, DE
Formamide	Applichem, Darmstadt, DE
Glucose	Carl Roth, Karlsruhe, DE
Glycerol	Carl Roth, Karlsruhe, DE
Isopropyl-b-D-thiogalactopyranoside (IPTG)	Carl Roth, Karlsruhe, DE
Magnesium chloride (MgCl <sub>2</sub> )	Merck, Darmstadt, DE
Magnesium sulfate (MgSO <sub>4</sub> )	Merck, Darmstadt, DE
Manganese (II) chloride (MnCl <sub>2</sub> )	Merck, Darmstadt, DE
Morpholinopropane sulfonic acid (MOPS)	Carl Roth, Karlsruhe, DE

Potassium Acetate (KAc)	Merck, Darmstadt, DE
Potassium chloride (KCl)	Carl Roth, Karlsruhe, DE
RNase free water	Macherey Nagel, Düren, DE
Rubidium chloride (RbCl)	Applichem, Darmstadt, DE
S2101	Calbiochem, Merck Millipore, Darmstadt, DE
Sodium acetate	Merck, Darmstadt, DE
Sodium butyrate	Sigma Aldrich, Steinheim, DE
Sodium chloride (NaCl)	Carl Roth, Karlsruhe, DE
β-Mercaptoethanol	Serva, Heidelberg, DE
Tetracycline	Applichem, Darmstadt, DE
Tris	Applichem, Darmstadt, DE
TRIzol Reagent	Thermo Fisher Scientific, Waltham, USA
Trypton/Pepton aus Casein	Carl Roth, Karlsruhe, DE
X-Gal	Carl Roth, Karlsruhe, DE
Yeast extract	Carl Roth, Karlsruhe, DE

Table A-11: Size standards for nucleic acids.

Item	Company
Size marker pBR328	Carl Roth, Karlsruhe, DE
GeneRuler 1 kb Range DNA Ladder	ThermoScientific, Waltham, USA
FastRuler Low Range DNA Ladder	ThermoScientific, Waltham, USA

Table A-12: Restriction enzymes.

Item	Company
ScaI	ThermoScientific, Waltham, USA
XhoI, FastDigest™	ThermoScientific, Waltham, USA
EcoRI	ThermoScientific, Waltham, USA

## A11. General Devices and Accessories

Table A-13: Devices and accessories in general use in the laboratory.

Equipment	Company
Benchtop centrifuge Allegra X-12R	Beckman Coulter, Krefeld, DE
Bunsen burner	N/A
Centrifuge Universal 32	Hettich, Tuttlingen, DE
Digital scale PRJ	Kern, Balingen, DE
Eppendorf® 5415 D Micro Centrifuge	Eppendorf, Hamburg, DE
Eppendorf® 5417 R Centrifuge	Eppendorf, Hamburg, DE
Eppendorf® ThermoMixerC	Eppendorf, Hamburg, DE
Magnetic stirrer	N/A
Microwave	Panasonic, Osaka, JP
Vortex Mixer Vortex Genie 2	Scientific Industries, New York, USA

## A12. Cell Culture Equipment

Table A-14: Cell culture equipment.

Equipment	Company
accu-jet® pro Pipette Controller	BrandTech Scientific, Inc., Wertheim, DE
Benchtop centrifuge Allegra X-12R	Beckman Coulter, Krefeld, DE
Camera Nikon Coolpix	Nikon Corp., Tokio, JP
Cellometer Auto 2000	Nexcelom Bioscience, Lawrence, USA
CO <sub>2</sub> incubator CB 150	Binder, München, DE
Fuchs Rosenthal counting chamber	N/A

GyroTwister™ 3D shaker	Labnet International Inc, Edison, USA
Inverse microscope Nikon TMS-F	Nikon Corp., Tokio, JP
sterile safety cabinet Herasafe	Heraeus, Hanau, DE
Vacuum pump	N/A
Waterbath Medingen W6	Labortechnik Medingen, Arnsdorf, DE

### A13. Molecular Biological Equipment

Table A-15: Molecular biological equipment.

Equipment	Company
2100 Bioanalyzer	Agilent, Santa Clara, USA
ChemiDoc™ Imaging System	Bio-Rad, Hercules, USA
Consort Electrophoresis Power Supply, 800 Series, E844	Peqlab, Erlangen, DE
Dounce tissue grinder set (D8938)	Sigma Aldrich, Steinheim, DE
Electrophoresis chambers (DNA)	Carl Roth GmbH, Karlsruhe, DE
Electrophoresis chambers (RNA)	Thermo Fisher Scientific, Waltham, USA
High-Performance UV-Transilluminator	UVP, Cambridge, UK
Kisker Test tube rotator	Kisker Biotech, Steinfurt, DE
LightCycler Carousel Centrifuge	Roche Life Science, Basel, CH
LightCycler Sample Carousel	Roche Life Science, Basel, CH
LightCycler® 2 capillary Real-Time PCR system	Roche Life Science, Basel, CH
LightCycler® 480 Instrument	Roche Life Science, Basel, CH
LightCycler® 96 System	Roche Life Science, Basel, CH
Next-Generation-Sequencer HiSeq 2500	Illumina, San Diego, USA
T1 Thermocycler 96	Biometra, Göttingen, DE
Thermocycler T3000	Biometra, Göttingen, DE
UV/Vis-spectrophotometer Nano-Drop ND-2000	Thermo Fisher Scientific, Waltham, USA
UV/Vis-spectrophotometer NanoDrop®ND-1000	Peqlab, Erlangen, DE
Workstation DNA/RNA UV-Cleaner box	Kisker-Biotech, Steinfurt, DE

### A14. Phenol/Chloroform Extraction of RNA

Adherent cells were homogenized in the culture vessel with TRIzol at 100  $\mu\text{L}/\text{cm}^2$ , then stored at  $-20^\circ\text{C}$ . Samples were thawed for approx. 10 min before starting work. For phase separation, 200  $\mu\text{L}$  of chloroform was added to each sample, and tubes were mixed by shaking tubes in a reaction tube stand. Tubes were let stand for 3 min, before centrifugation for 15 min, at 10,000 x g at  $4^\circ\text{C}$ . Holding the reaction tube at a  $45^\circ$  angle, the upper colorless aqueous phase containing RNA was transferred into a new reaction tube. The interphase and lower phase (red, containing DNA and protein) remain in the reaction tube. Approximately 500  $\mu\text{L}$  of isopropanol (1:1 ratio of volumes to the upper phase) were added to RNA containing phase, shaken, and incubated for 10 min on ice, improving the precipitation step. The tubes were centrifuged for 15 min, at 10,000 x g at  $4^\circ\text{C}$ . The supernatant was carefully removed. The pellet was washed twice in 70% ethanol (centrifugation 5 min, at 10,000 x g at  $4^\circ\text{C}$ ). The pellet was dried briefly for 3 min at RT and resuspended in 30  $\mu\text{L}$  of nuclease-free  $\text{H}_2\text{O}$ . RNA was used or stored at  $-80^\circ\text{C}$ .

### A15. Phenol/Chloroform Extraction of DNA

DNA extraction with Phenol/Chloroform was performed during N-ChIP to verify shearing efficiency and determine the DNA concentration of the samples. Under a fume hood, 200  $\mu\text{L}$  of 1:1 phenol/chloroform TE saturated pH 8 was added to the sample, and mixed with a vortex, before centrifugation at maximum speed for 5 min. The aqueous phase of each sample was transferred to a new reaction tube, then 20  $\mu\text{L}$  of 3 M sodium acetate pH 5.2 and 500  $\mu\text{L}$  of 100% ethanol was added. The samples were mixed entirely using a vortex mixer and subsequently left at  $-20^\circ\text{C}$  overnight. Following a centrifugation step at maximum speed for 10 min at  $4^\circ\text{C}$ , the supernatant was carefully removed without disturbing the pellet. For washing the pellet, 500  $\mu\text{L}$  70% ice-cold ethanol was added, the tube was centrifuged at maximum speed for 5 min at  $4^\circ\text{C}$ , and all supernatant was removed.

The pellet was air-dried and subsequently resuspended in 30  $\mu$ L nuclease-free H<sub>2</sub>O. DNA concentration was determined on a Nanodrop® ND-1000 spectrophotometer.

### A16. RNA Isolation for RNA-seq

The adherent epithelial cells grown in a monolayer were directly lysed in the cell culture vessel. The cells were disrupted by adding 350  $\mu$ L Buffer RLT Plus from RNeasy Plus Mini Kit (Qiagen) and harvested with a cell scraper. Cells were homogenized by resuspension and transferred into a new 1.5 mL reaction tube, mixed with a vortex mixer, and kept at -80°C until RNA isolation. The thawed homogenized lysate was transferred to a gDNA Eliminator spin column in a 2 mL collection tube. During centrifugation for 30 s at 8000 x g, gDNA was held back while RNA flowed through the silicate membrane. Ethanol (70%) was added in a volume ratio of 1:1 to the sample flow-through and mixed well by pipetting. The 700  $\mu$ L mix of ethanol and sample was transferred to an RNeasy spin column placed in a 2 mL collection tube. RNA bound to the column membrane in the following centrifugation step (for 30 s at 8000 x g). With the addition of 700  $\mu$ L Buffer RW1 to the RNeasy spin column, the membrane was washed when centrifuged for 30 s at 8000 x g. The membrane was washed twice with 500  $\mu$ L buffer RPE and centrifuged at 8000 x g for 30 s in the first round and for 2 min in the second round. The optional centrifugation step to dry the membrane at full speed for 1 minute was included. The RNeasy spin column was placed in a new 1.5 mL reaction tube, and for elution of RNA, 30  $\mu$ L RNase-free H<sub>2</sub>O was added directly to the spin column membrane. The column in the tube was centrifuged for 1 min at 8000 x g to elute the RNA.

This procedure was followed by an extra gDNA elimination step with RNeasy® MinElute® Cleanup (Qiagen). To the remaining 21  $\mu$ L of RNA were 66.5  $\mu$ L nuclease-free H<sub>2</sub>O, 10  $\mu$ L Buffer RDD and 2.5  $\mu$ L DNase I stock solution were added. The mix was incubated on the benchtop at RT for 10 min before 350  $\mu$ L RLT buffer, and subsequently, 250  $\mu$ L of 96-100% ethanol was added. The solution was mixed by pipetting. The total volume was sequentially loaded onto the RNeasy MinElute column and centrifuged for 30 s at 8000 x g to bind the RNA to the membrane. Collection tubes were renewed at every step. The column was washed with the addition of 500  $\mu$ L RPE buffer, followed by centrifugation for 30 s at 8000 x g. The column was then washed with 500  $\mu$ L 80% ethanol and centrifuged for 2 min at 8000 x g. The column was placed in a new 2 mL collection tube and dried through centrifugation at full speed for 5 min. The column was placed into a new 1.5 mL reaction tube, and 10  $\mu$ L of RNase-free H<sub>2</sub>O were pipetted onto the membrane of the column, incubated at RT for 5 min, and centrifuged for 1 min at 8000 x g. The elution step was repeated in another 10  $\mu$ L of RNase-free H<sub>2</sub>O. The quantity of RNA was determined with a Nanodrop® spectrophotometer and Qubit 1.0 Fluorometer (both Thermo Fisher). The RNA quality was controlled with the Agilent Bioanalyzer 2100. 1  $\mu$ L of RNA sample was used for PGC control PCR, and the remaining RNA was stored at -80°C until library preparation.

### A17. Plasmid Preparation (Maxiprep)

1. Freshly prepared 100 mL Trypton-Yeast (TY)-medium + ampicillin was required. A new single colony in from agar plate + amp. was inoculated to 1.5 mL TY medium + amp. and allowed to grow for approx. 5 – 8 h at 37°C (in a test glass tube, shaking at 180 rpm)
2. 1 mL from the preculture was transferred into 100 mL TY medium + amp. (in 500 mL Erlenmeyer flask) and incubated overnight at 37°C, 180 rpm.
3. In the morning, divide the culture into two 50 mL falcons and centrifuge at 3750 rpm, 4°C for 10 min.
4. Decant supernatant and wash the pellet with 5 mL cold TE (bacterial pellet could be stored at -20°C)
5. Centrifugation for 10 min at 3750 rpm, 4°C.
6. Decant the pellet and add solution 1 (3.2 mL) to resuspend.
7. Add solution 2 (6.4 mL): invert and mix gently.
8. Add solution 3 (4.8 mL), and vortex vigorously for 10 s.
9. Centrifuge samples for 15 min at 3750 rpm, 4°C.
10. Filter the supernatant through sterile gauze and combine both falcons. NOTE: Place gauze on top of open 50 mL falcon, and pour slowly and carefully without letting the gauze fall in.
11. Precipitate with one volume of isopropanol, and let stand for 10 min.
12. Preheat water bath to 37°C - Centrifuge samples for 15 min at 3750 rpm, 4°C.
13. Decant samples and wash with 5 mL 70% ethanol.
14. Centrifugation for 10 min at 3750 rpm, 4°C.
15. Decant, dry the pellet and resuspend the sample in 3.2 mL cold TE buffer.
16. Add 8  $\mu$ L RNase A (20 mg/mL), and incubate samples for 5 min at 37°C.
17. Add 0.5  $\mu$ L RNase T1 (100,000 U/mL, Roche Life Science) for 25 min at RT.
18. Add 5  $\mu$ L proteinase K (20 mg/mL) for 30 min at 37°C.

19. Precipitate DNA with 1/10 volume of 3 M Na-acetate (320  $\mu$ L) and addition of 1 volume of isopropanol (3.5 mL).
20. Centrifugation for 10 min at 3750 rpm, 4°C.
21. Wash the pellet with 5 mL of 70% alcohol.
22. Centrifugation for 10 min at 3750 rpm, 4°C.
23. Dry the pellet and resuspend it in 1 mL nuclease-free H<sub>2</sub>O.
24. Aliquot the sample into 2 x 500  $\mu$ L microcentrifuge tubes, and add 500  $\mu$ L PCI (3:1) to each tube, shake, and centrifuge in a benchtop centrifuge at max. speed 13,000 rpm for 10 min at RT. Transfer supernatant (upper phase) into a new microcentrifuge tube.
25. Add 500  $\mu$ L CI (24:1) to each tube, shake, and centrifuge 13,000 rpm for 10 min at RT. Transfer supernatant (upper phase) into a new microcentrifuge tube.
26. Repeat the previous step by adding 500  $\mu$ L CI (24:1) to each tube, shaking, and centrifuging at 13,000 rpm for 10 min at RT. Transfer supernatant (upper phase) into a new microcentrifuge tube.
27. Precipitate DNA by adding 50  $\mu$ L 3 M Na-acetate and 500  $\mu$ L isopropanol, shake after addition, let stand briefly (1-2 min), then centrifuge at 13,000 rpm, 10 min.
28. Decant supernatant, wash pellet with 500  $\mu$ L 70% ethanol, centrifuge at 13,000 rpm, 10 min.
29. Decant supernatant, centrifuge again briefly, and remove residual ethanol.
30. Dry pellet, add 100  $\mu$ L nuclease-free H<sub>2</sub>O to each tube, then combine samples. Leave samples overnight at 2 – 8°C in a fridge for remaining dissolution. Measure nucleic acid concentration on the following day.

Solutions for plasmid preparation:

Solution 1: 50 mM glucose 6.25 mL (stock 2 M).

10 mM EDTA 5 mL (stock 0.5 M)

25 mM Tris pH 8.0 6.25 mL (stock 1 M)

Fill to 250 mL with H<sub>2</sub>O MilliQ

Solution 2: 0.2 N NaOH 10 mL (stock 5 M)

1% SDS 12.5 mL (stock 20% SDS solution)

Fill to 250 mL with H<sub>2</sub>O MilliQ

Solution 3: 150 mL KAc (5 M, autoclaved)

28.75 mL 96% acetic acid

71.25 mL H<sub>2</sub>O (MilliQ)

TE buffer:

*Note: make 10x buffer for plasmid preparation; you can use 10x or 1x buffer. Storage in the refrigerator.*

Recipe 1 for TE buffer: 100 mM Tris = 1.211 g pH 8.0, 100 mL

10 mM EDTA = 0.3722 g

Recipe 2 for TE buffer: 10 mL 1M Tris pH 8 for 1 l, verify pH 8.0

2 mL 500 mM EDTA pH 8

KAc: 5 M stock solution, weighing 98.15 g to 200 mL H<sub>2</sub>O MilliQ, autoclave.

20% SDS solution: weigh 50 g to 250 mL MilliQ H<sub>2</sub>O, stir at 37°C

CI (24:1): 24 mL chloroform + 1 mL isoamyl alcohol storage in refrigerator

PCI (3:1): 30 mL phenol + 10 mL CI

## **A18. Master Pure™ Complete DNA Purification Kit, Lucigen**

*Note: Cat.no. MC85200*

Cell samples:

1. Dilute 1  $\mu$ L of Proteinase K into 300  $\mu$ L of Tissue and Cell Lysis Solution for each sample
2. Pellet cells by centrifugation ( $1 \times 10^6$  cells) and discard the supernatant, leaving approximately 25  $\mu$ L liquid
3. Vortex mix 10 s to resuspend the cell pellet
4. Add 300  $\mu$ L Tissue and Cell Lysis Solution containing Proteinase K; mix thoroughly
5. Incubate at 65°C for 15 min; vortex mix for 5 min

6. Cool the samples to 37°C and add 1 µa. of 5 µg/µL RNase A to the sample; mix thoroughly
7. Incubate at 37°C for 30 min
8. Place the samples on ice for 3-5 min, and then proceed with DNA precipitation

DNA precipitation:

9. Add 175 µL of MPC protein precipitation reagent to 300 µL of lysed samples and vortex mix vigorously for 10 s.
10. Pellet debris by centrifugation at 4°C for 10 min at 10,000 x g in a microcentrifuge.
11. Transfer the supernatant to a clean microcentrifuge tube and discard the pellet.
12. Add 500 µL of isopropanol to the recovered supernatant. Invert the tube several times (30-40 times).
13. Pellet the DNA by centrifugation at 4°C for 10 min in a microcentrifuge.
14. Carefully pour off the isopropanol without dislodging the DNA pellet.
15. Rinse twice with 70% ethanol being careful not to dislodge the pellet. Centrifuge briefly if the pellet is dislodged. Remove all of the residual ethanol with a pipet.
16. Resuspend the DNA in 35 µL of TE Buffer.

The additional RNA removal step

17. Add 1 µL of RNase A to the sample, and mix thoroughly.
18. Incubate at 37°C for 30 min.
19. Add 14 µL TE Buffer and 50 µL of 2x Tissue and Cell Lysis Solution to each sample.
20. Place the samples on ice for 3-5 min. Add 100 µL of MPC Protein Precipitation Reagent and mix by vortexing for 10 s.
21. Pellet debris by centrifugation at 4°C for 10 min at 10,000 x g in a microcentrifuge
22. Transfer the supernatant to a clean microcentrifuge tube and discard the pellet.
23. Add 500 µL of isopropanol to the recovered supernatant. Invert the tube several times (30-40 times).
24. Pellet the DNA by centrifugation at 10,000 x g for 10 min in a microcentrifuge.
25. Carefully pour off the isopropanol without dislodging the DNA pellet.
26. Rinse twice with 70% ethanol being careful not to dislodge the pellet. Centrifuge briefly if the pellet is dislodged. Remove all of the residual ethanol with a pipet.
27. Resuspend the DNA in 35 µL of TE Buffer.

**A19. ChIP DNA Clean & Concentrator™ Zymo Research**

*Note: Catalog#: D5205*

All necessary reagents were prepared before starting, including the preparation of the final DNA wash buffer solution. In a 1.5 mL microcentrifuge tube containing the ChIP DNA or input DNA, ChIP DNA binding buffer (500 µL) was added to the sample (100 µL) in a 5:1 volume ratio and mixed briefly. For each sample, the mixture was loaded onto a Zymo-Spin™ column in a Collection Tube. After centrifugation at 10,000 x g for 1 min, the flow-through was discarded, and 200 µL of wash buffer was loaded onto the column. After centrifugation, the wash step was repeated. DNA was eluted from the column matrix by adding 40 µL elution buffer followed by centrifugation and collected in a new 1.5 mL microcentrifuge tube. The highly purified DNA was usable for downstream applications such as qPCR.

B. Supplementary Results

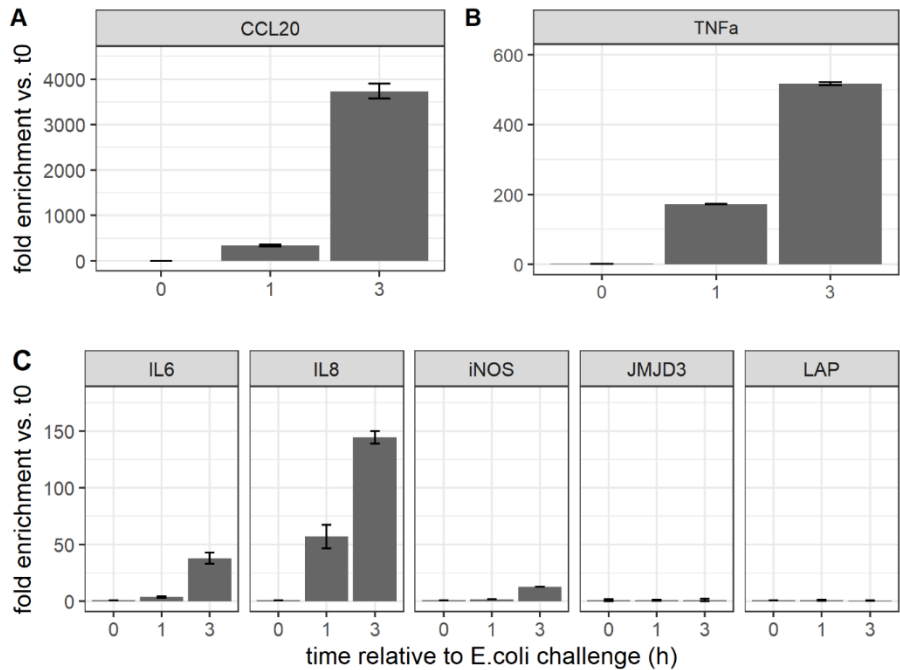


Figure B-1: Gene regulation of MAC-T after *E. coli*<sub>1303</sub> challenge. Time points 0, 1, and 3 h relative to infection are plotted against the fold enrichment of *CLIC1* normalized copy numbers relative to timepoint 0 h (t0). *CCL20* (A) and *TNF-α* (B) expression is higher than expression of genes *IL6*, *IL8*, *iNOS*, *JMJD3*, and *LAP* (C). Error bars are SEM. The figure was generated with R packages *ggplot2* and *cowplot*.

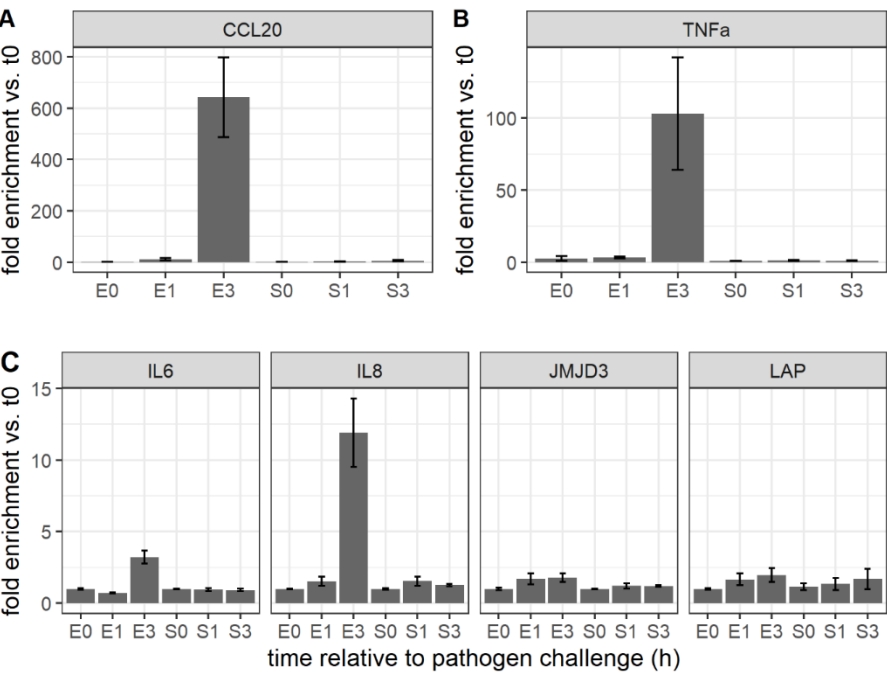


Figure B-2: Gene regulation of MAC-T after *E. coli*<sub>1303</sub> and *S. aureus*<sub>1027</sub> challenge. Time points 0, 1, and 3 h relative to infection with the pathogens *E. coli* (E) and *S. aureus* (S) are plotted against the fold enrichment of *CLIC1* normalized copy numbers relative to timepoint 0 h (t0). *CCL20* (A) and *TNF-α* (B) expression is higher than expression of genes *IL6*, *IL8*, *iNOS*, *JMJD3*, and *LAP* (C). Challenge with *S. aureus* caused a less pronounced immune response than *E. coli*. Error bars are SEM. The figure was generated with R packages *ggplot2* and *cowplot*.



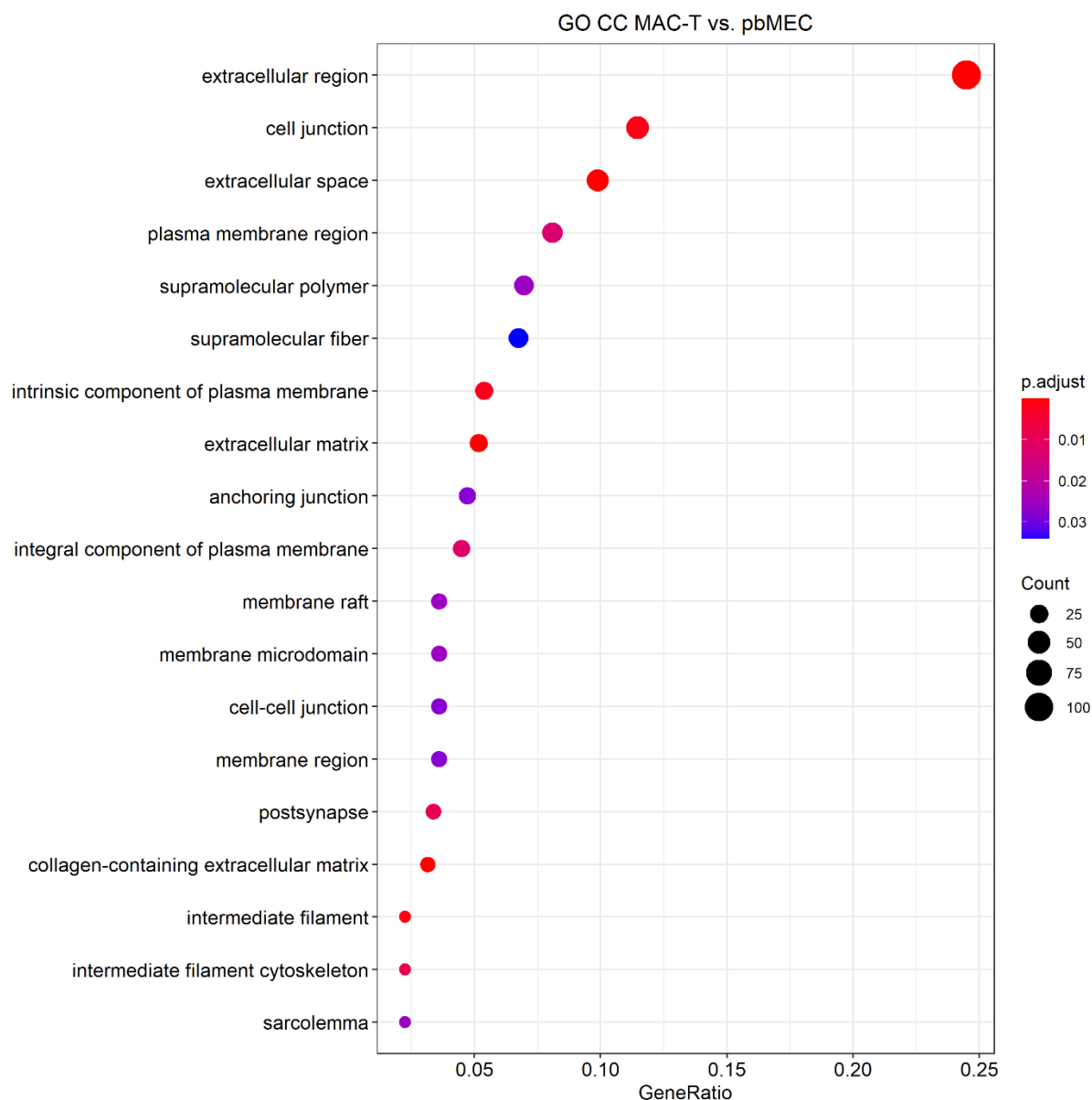


Figure B-3. Gene Ontology (GO) Analysis for Cellular Components for differentially expressed genes (DEGs) between MAC-T and pbMEC. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) and applying Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The dot plot shows the top overrepresented terms for DEGs. The size of the dots correlates with the number of genes (count) per term. The gene ratio plotted on the x-axis is calculated by the count of the genes per term divided by the set size of the DEGs.

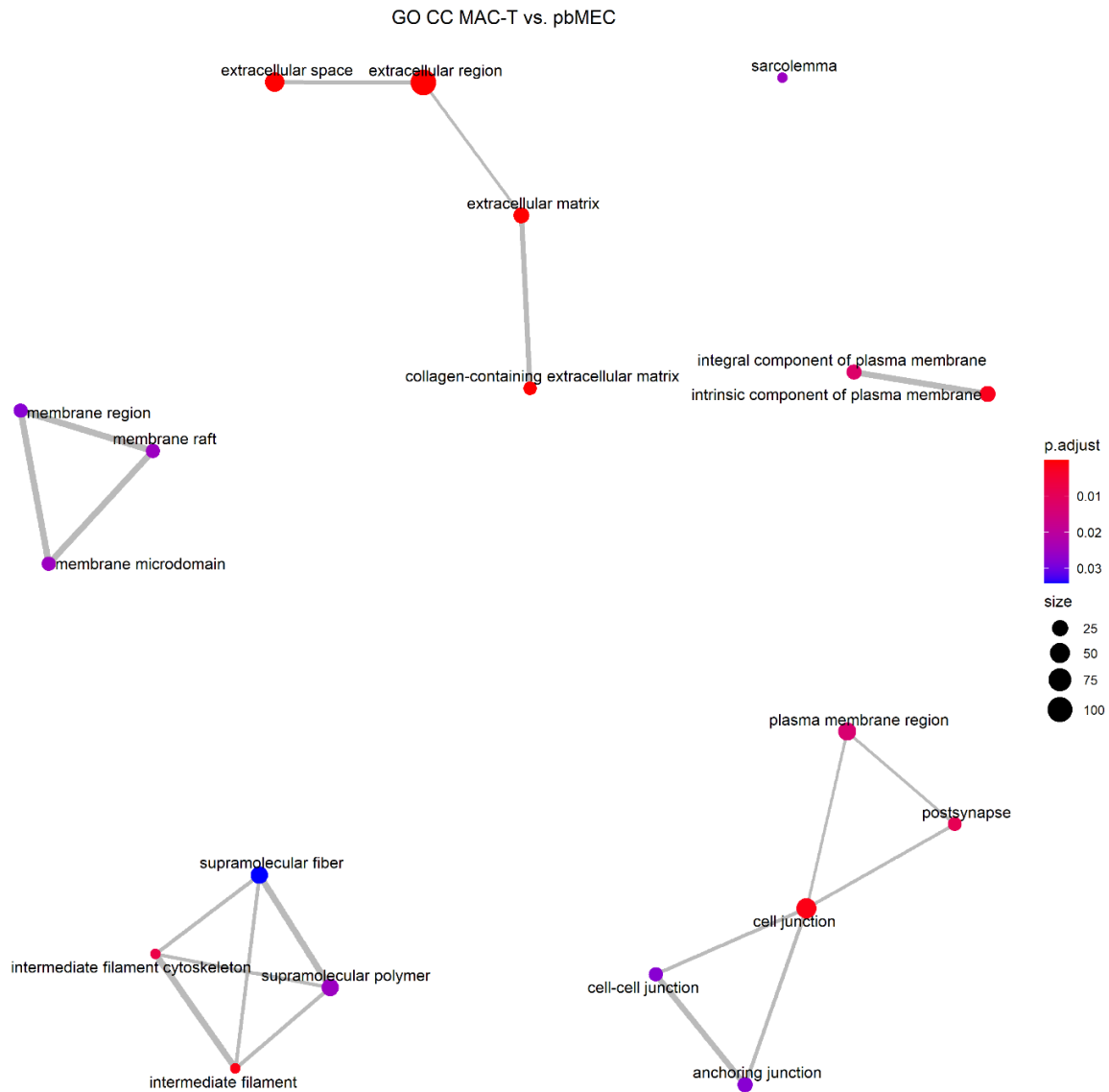
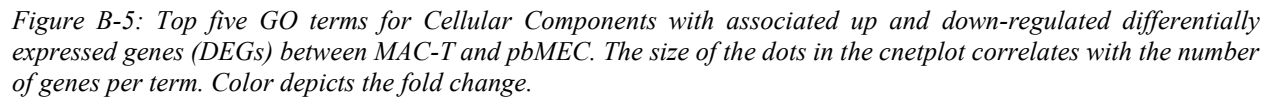


Figure B-4: Relationship between the GO terms for Cellular Components for differentially expressed genes (DEGs) between MAC-T and pbMEC. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) and applying Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The size of the dots in the network correlates with the number of genes per term. Color depicts the significance value.



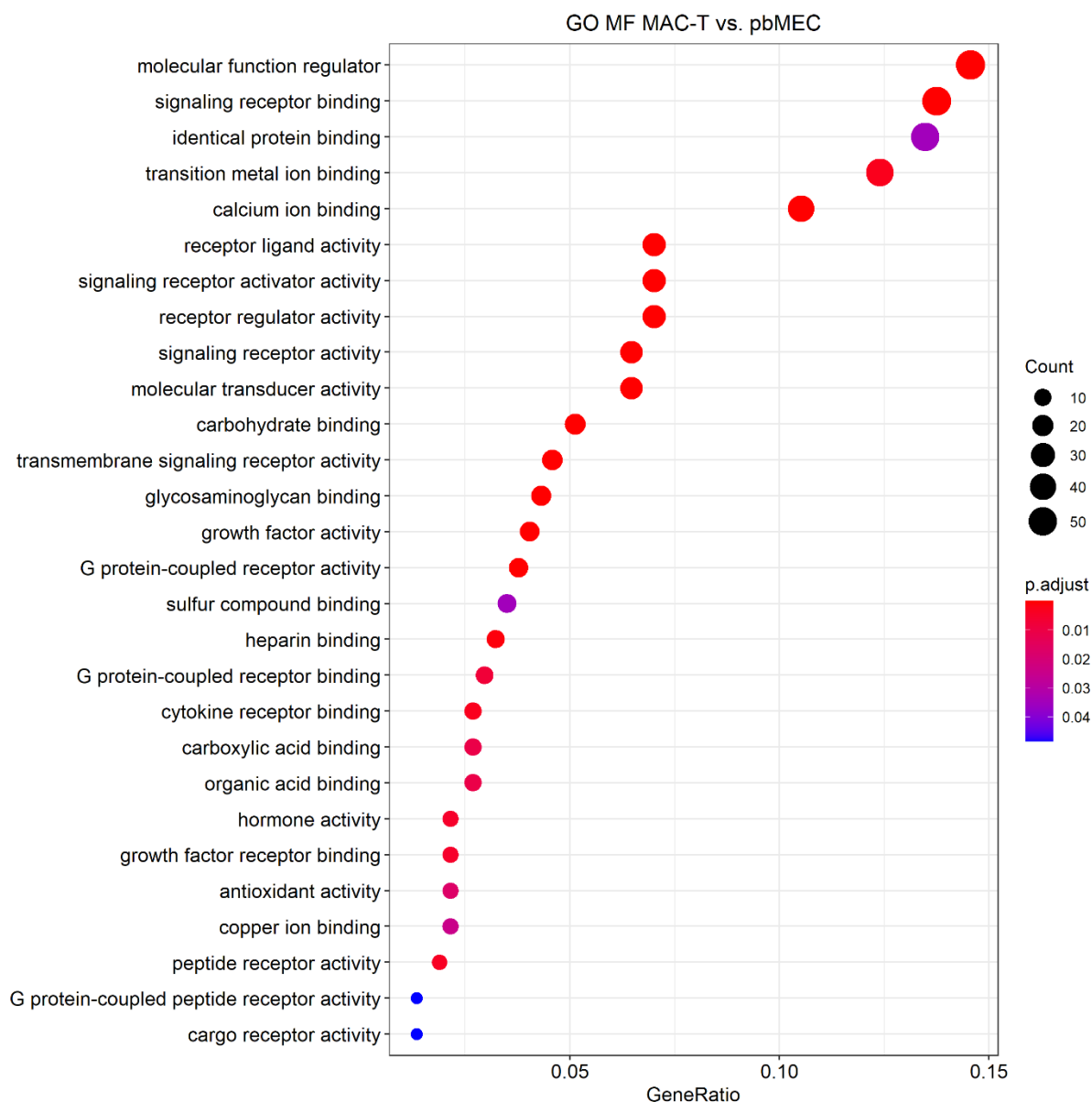


Figure B-6: Gene Ontology (GO) Analysis for Molecular Function for differentially expressed genes (DEGs) between MAC-T and pbMEC. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) and applying Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The dot plot shows the top overrepresented terms for DEGs. The size of the dots correlates with the number of genes (count) per term. The gene ratio plotted on the x-axis is calculated by the count of the genes per term divided by the set size of the DEGs.

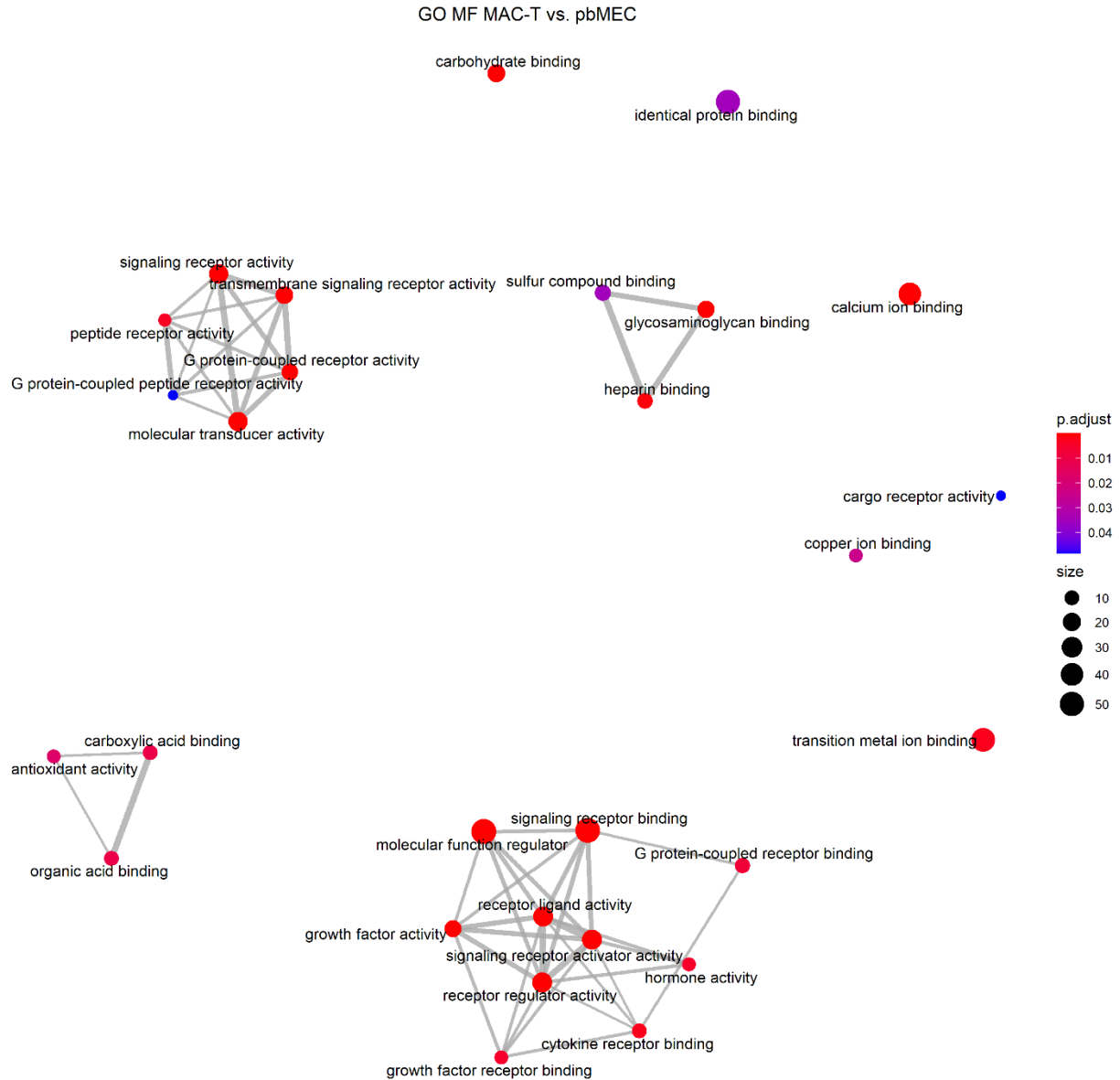


Figure B-7: Relationship between the GO terms for Molecular Function for differentially expressed genes (DEGs) between MAC-T and pbMEC. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) and applying Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The size of the dots in the network correlates with the number of genes per term. Color depicts the significance value.

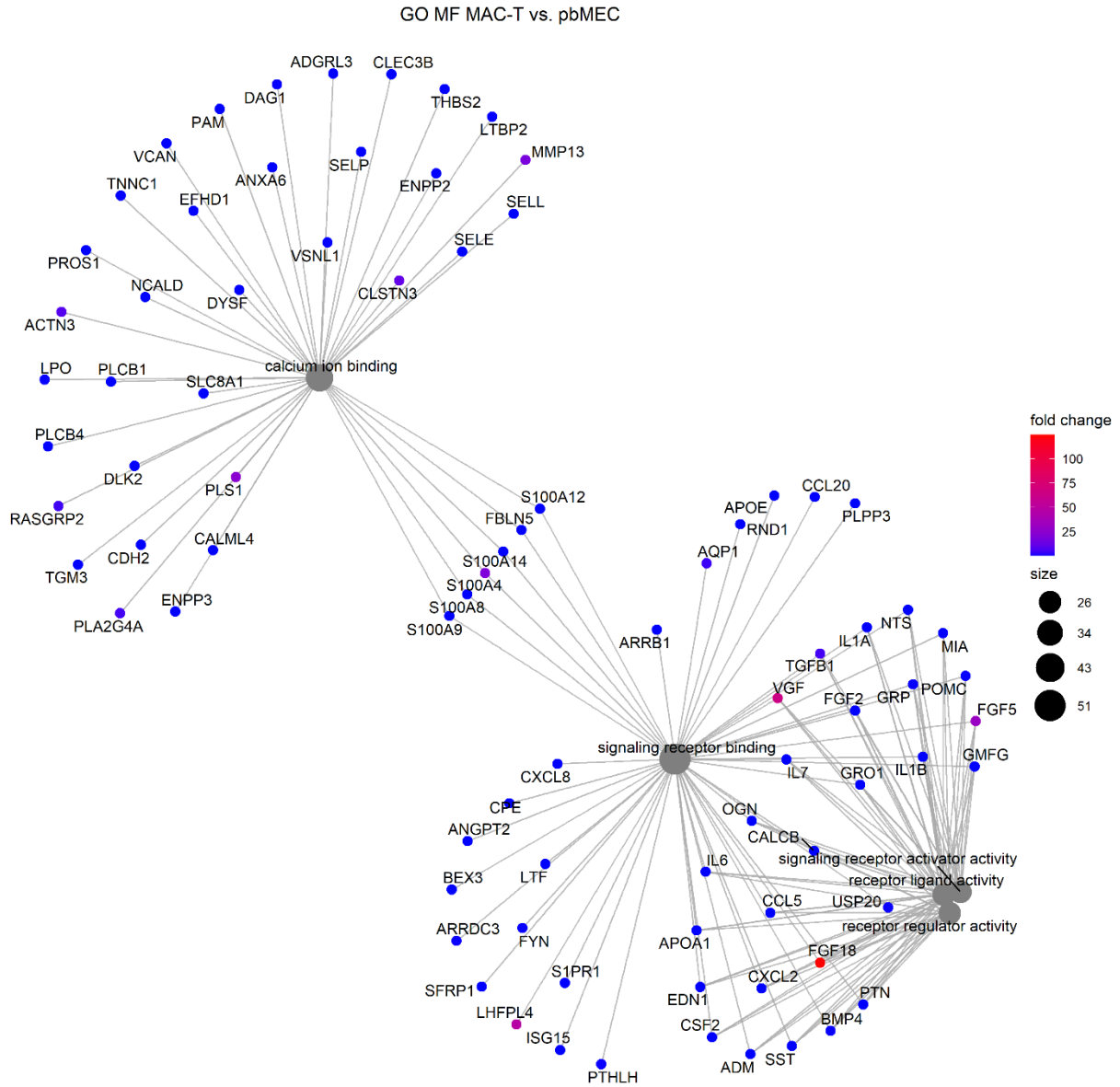


Figure B-8: Top five GO terms for Molecular Function with associated up and down-regulated differentially expressed genes (DEGs) between MAC-T and pbMEC. The size of the dots in the cnetplot correlates with the number of genes per term. Color depicts the fold change.

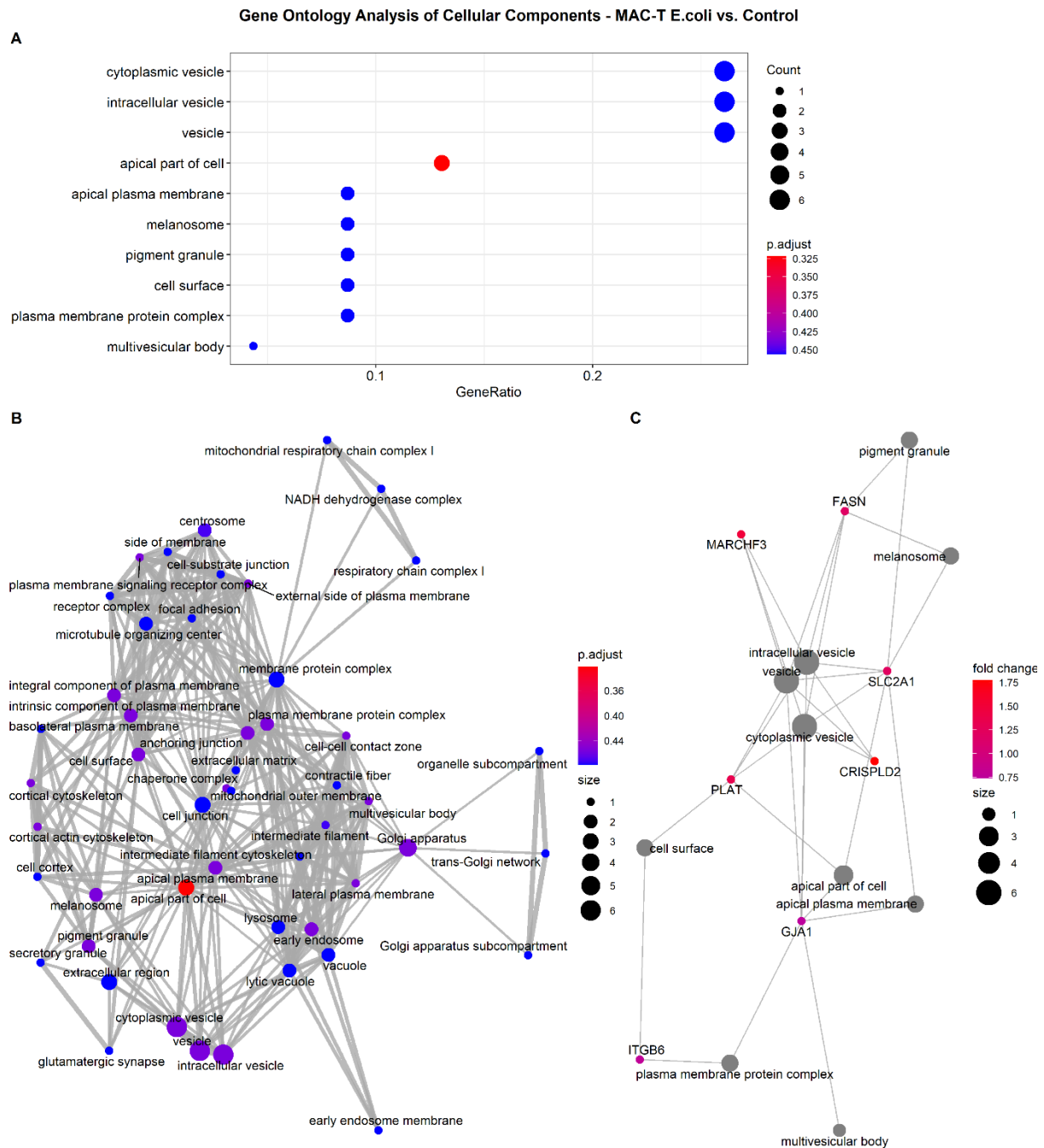


Figure B-9: Gene Ontology (GO) Analysis for Cellular Component for DEGs between MAC-T *E. coli*<sub>1303</sub> versus Control. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) but without applying the Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The presented results are, therefore, not significantly enriched terms. The dot plot (A) shows the top 10 overrepresented terms for DEGs. The size of the dots correlates with the number of genes (count) per term. Gene set enrichment map (B) depicts the relationship of the top 50 categories in which the DEGs are enriched. The size of the dots in the network correlates with the number of genes per term. The Cnetplot (C) shows the top 10 terms with genes associated with them.

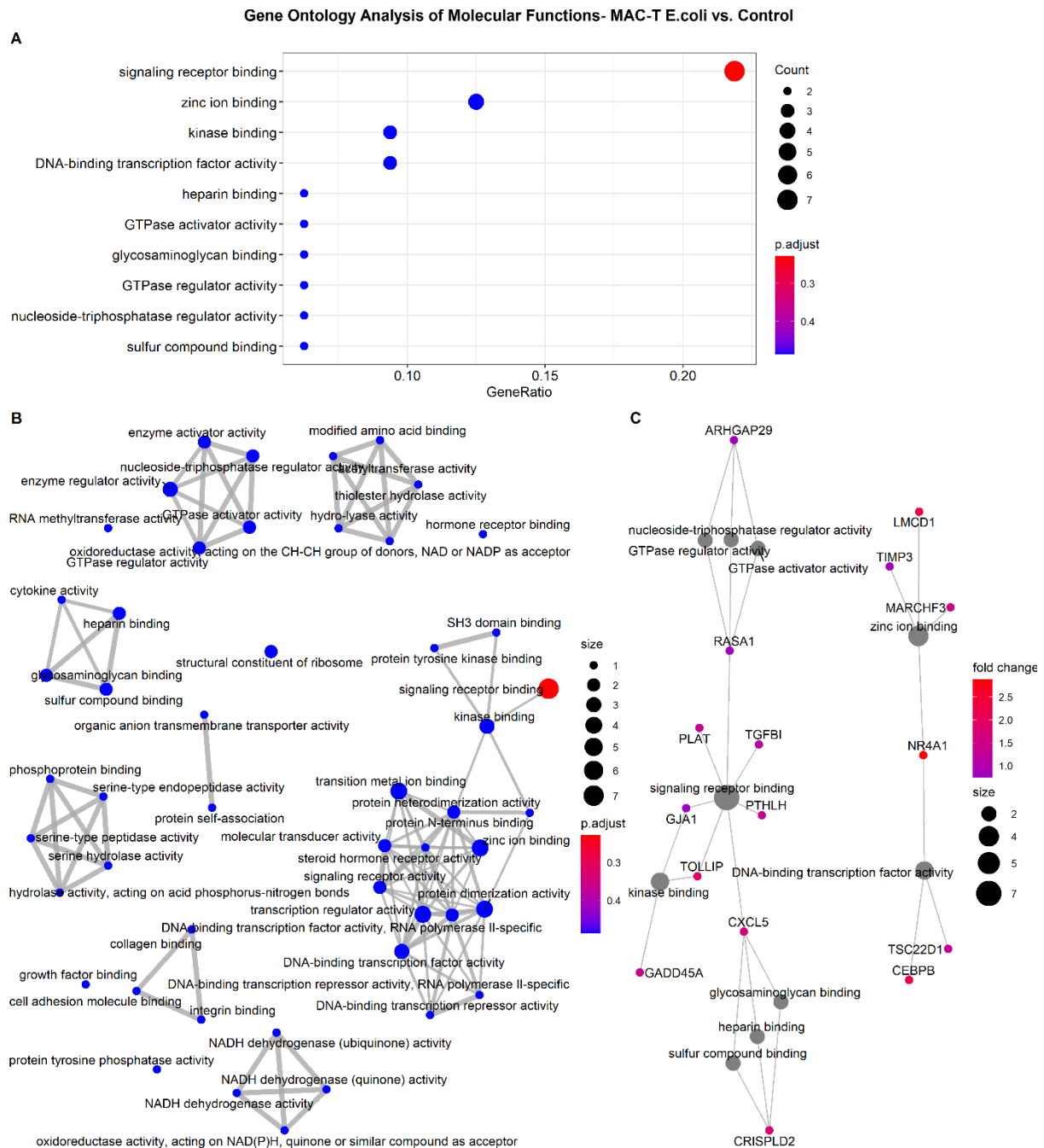


Figure B-10: Gene Ontology (GO) Analysis for Molecular Functions for DEGs between MAC-T *E. coli*<sub>1303</sub> versus Control. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) but without applying the Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The presented results are, therefore, not significantly enriched terms. The dot plot (A) shows the top 10 overrepresented terms for DEGs. The size of the dots correlates with the number of genes (count) per term. Gene set enrichment map (B) depicts the relationship of the top 50 categories in which the DEGs are enriched. The size of the dots in the network correlates with the number of genes per term. The Cnetplot (C) shows the top 10 terms with genes associated with them.



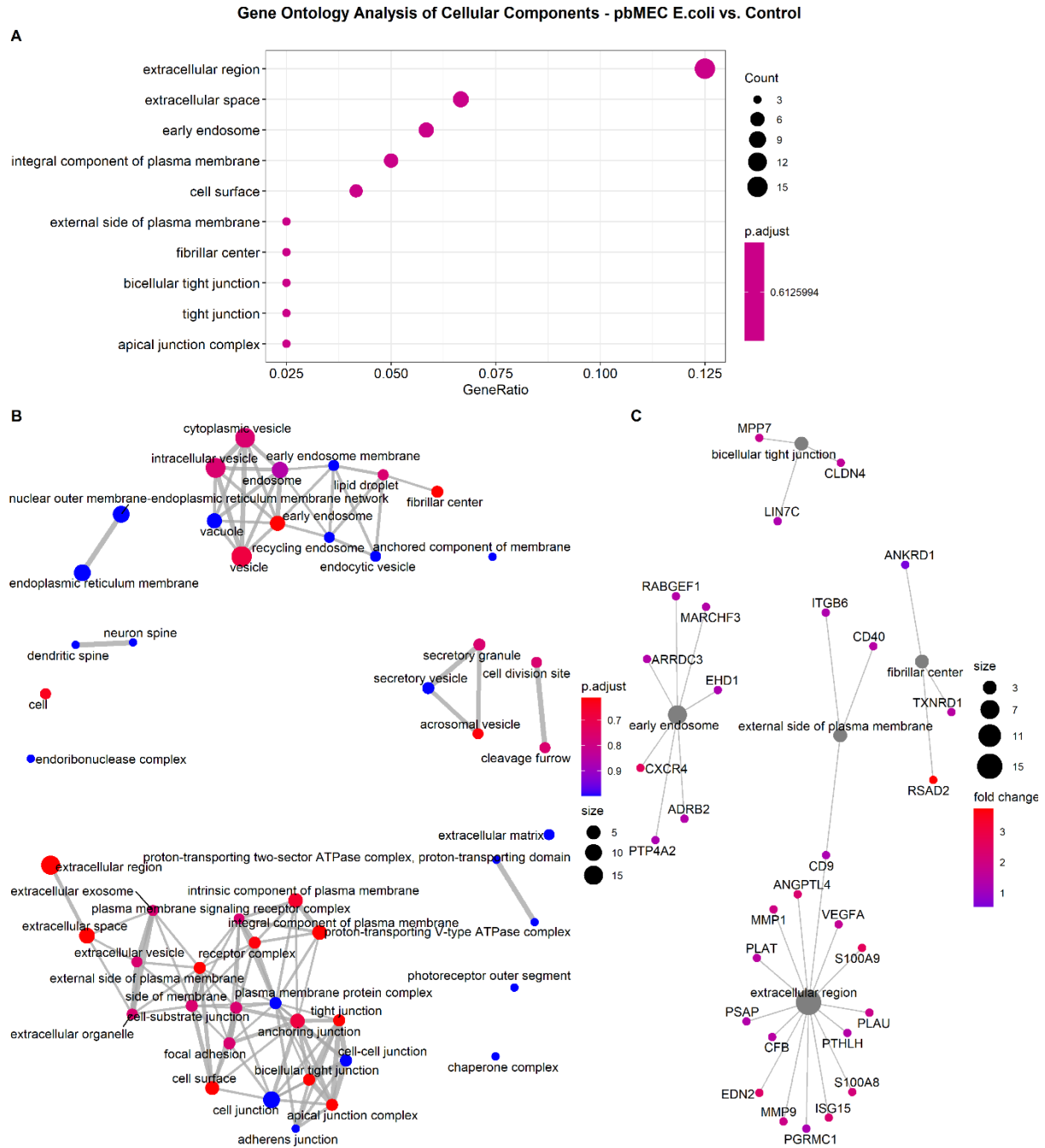


Figure B-11: Gene Ontology (GO) Analysis for Cellular Components for DEGs between pbMEC *E. coli*<sub>1303</sub> versus Control. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) but without applying the Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The presented results are, therefore, not significantly enriched terms. The dot plot (A) shows the top 10 overrepresented terms for DEGs. The size of the dots correlates with the number of genes (count) per term. Gene set enrichment map (B) depicts the relationship of the top 50 categories in which the DEGs are enriched. The size of the dots in the network correlates with the number of genes per term. The Cnetplot (C) shows the top five terms with genes associated with them.

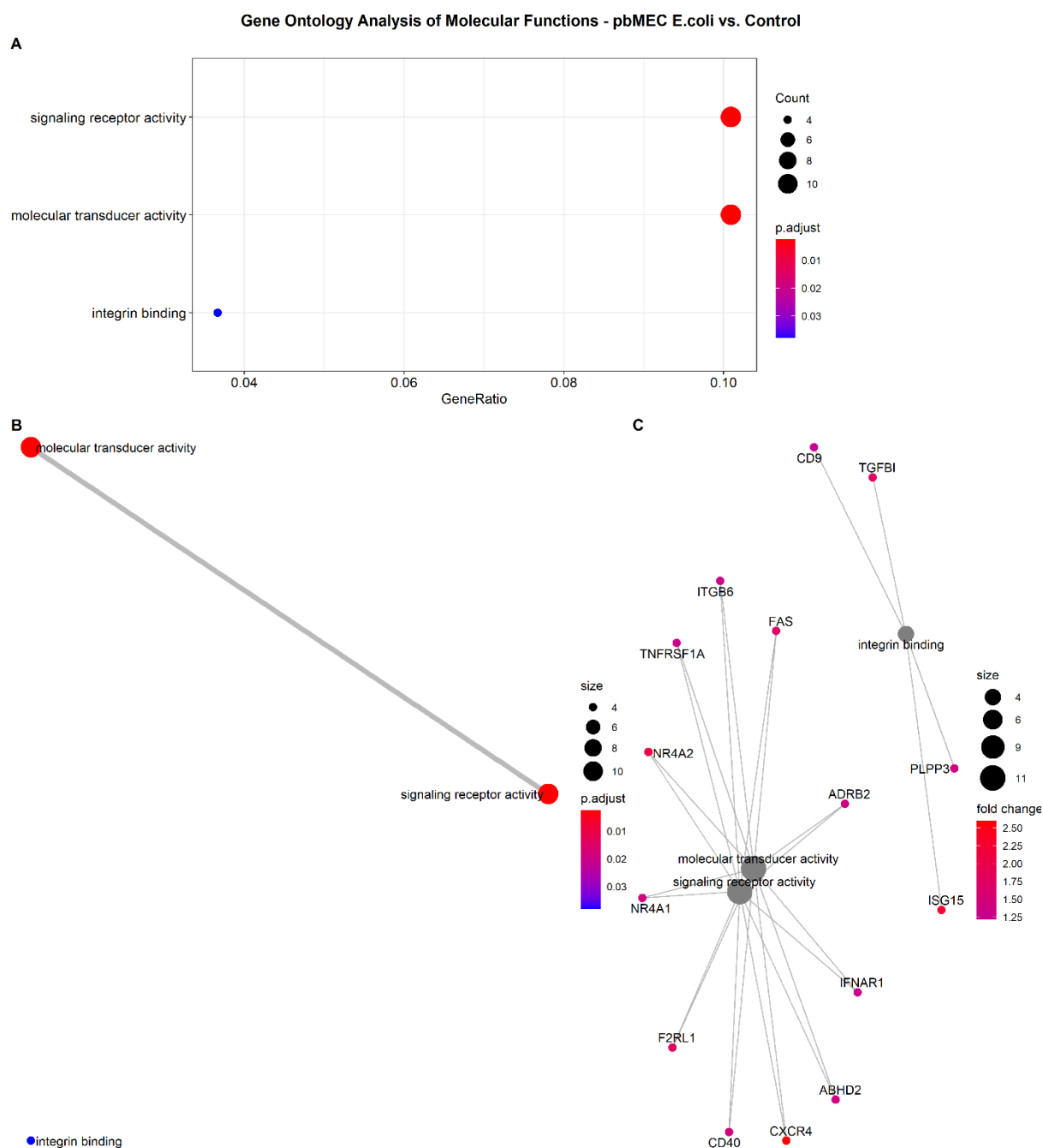


Figure B-12: Gene Ontology (GO) Analysis for Molecular Function for DEGs between pbMEC *E. coli*<sub>1303</sub> versus Control. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ), applying a Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The presented results are, therefore, significantly enriched terms. The dot plot (A) shows the top 10 overrepresented terms for DEGs. The size of the dots correlates with the number of genes (count) per term. Gene set enrichment map (B) depicts the relationship of the top 50 categories in which the DEGs are enriched. The size of the dots in the network correlates with the number of genes per term. The Cnetplot (C) shows the top five terms with genes associated with them.

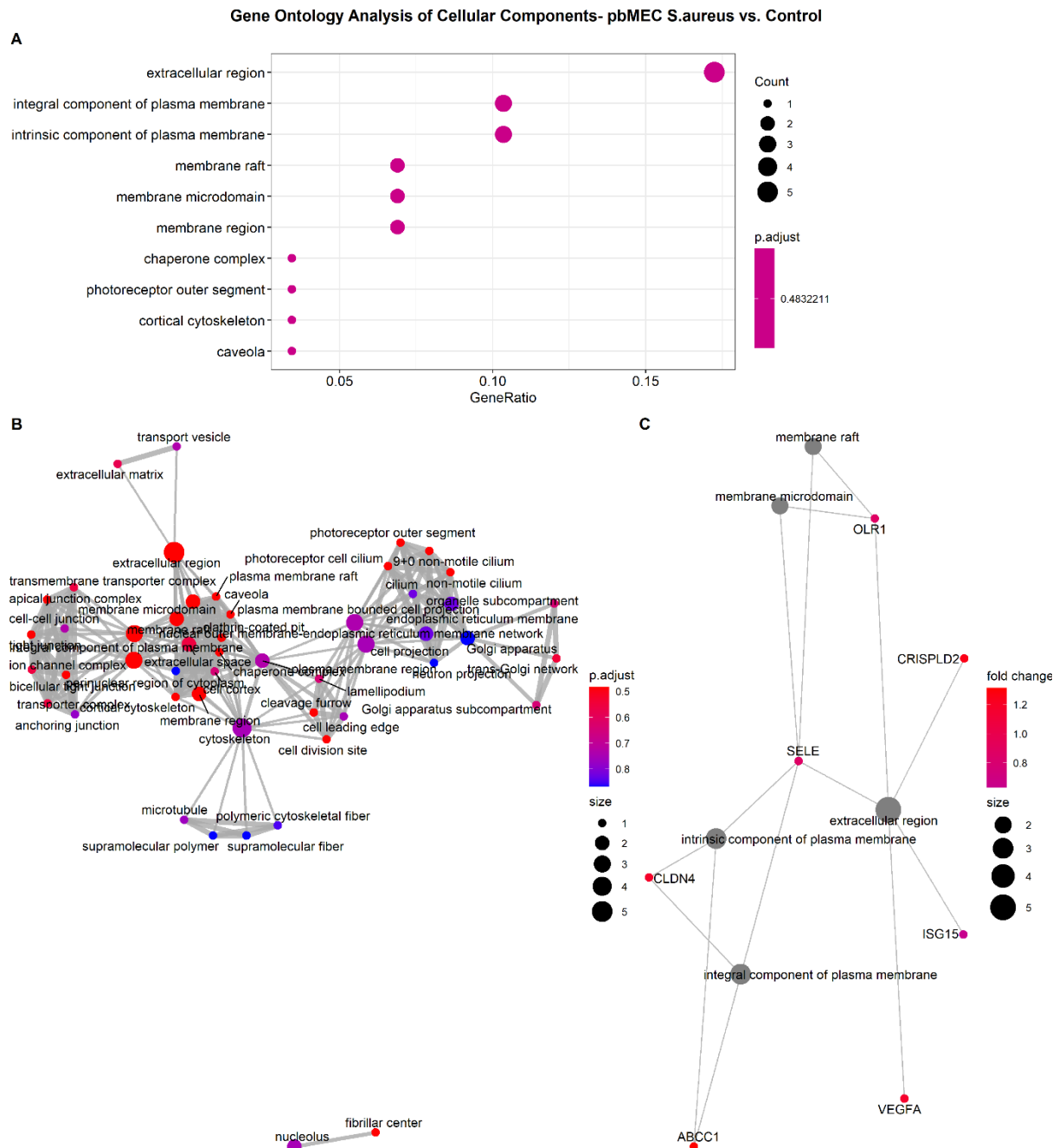


Figure B-13: Gene Ontology (GO) Analysis for Cellular Components for DEGs between pbMEC *S. aureus*<sub>1027</sub> versus Control. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) but without applying the Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The presented results are, therefore, not significantly enriched terms. The dot plot (A) shows the top 10 overrepresented terms for DEGs. The size of the dots correlates with the number of genes (count) per term. Gene set enrichment map (B) depicts the relationship of the top 50 categories in which the DEGs are enriched. The size of the dots in the network correlates with the number of genes per term. The Cnetplot (C) shows the top five terms with genes associated with them.

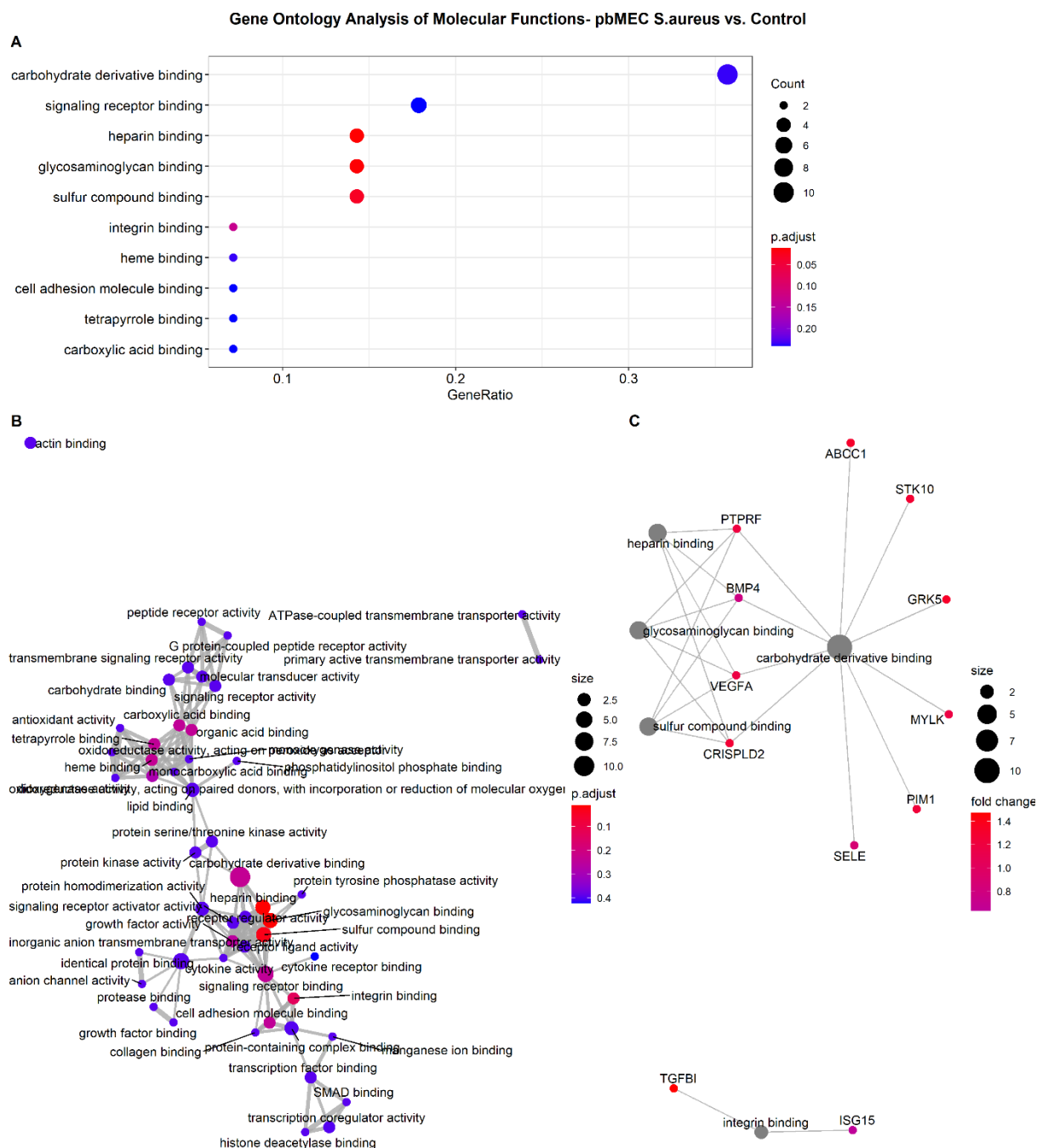


Figure B-14: Gene Ontology (GO) Analysis for Molecular Functions for DEGs between pbMEC *S. aureus*<sub>1027</sub> versus Control. GO analysis was performed with significant DEGs ( $p_{adj} < 0.05$ ) but without applying the Benjamini-Hochberg (BH) cutoff of  $p < 0.05$ . The presented results are, therefore, not all significantly enriched terms as can be taken by the color coding in the legend. The dot plot (A) shows the top 10 overrepresented terms for DEGs. The size of the dots correlates with the number of genes (count) per term. Gene set enrichment map (B) depicts the relationship of the top 50 categories in which the DEGs are enriched. The size of the dots in the network correlates with the number of genes per term. The Cnetplot (C) shows the top five terms with genes associated with them.

*Table B-1: Establishment of FAANG core assays for bovine cell models. The ChIP Optimization experiment included testing different antibody amounts and measuring several potential genomic regions to be used as positive or negative enriched regions for X-ChIP control qPCR. Unchallenged cells were processed with an ideal ChIP-seq Kit (Diagenode), and purified DNA was used for qPCR. Fold enrichment (recovery) > 5 was assumed to be accurate for these controls.*

Cell type	Antibody (Ab)	Ab amount	qPCR measurements for genomic regions (%Input)				fold enrichment= positive/negative genomic region				
			GAPDH_TSS_3	PRSS1_TSS	TSH2B		GAPDH_TSS_3/PRSS1		GAPDH_TSS_3/TSH2B		
MAC-T	IgG	1 µg	0,009	0,000	0,005		99,388		1,959		
	H3K4me3	0.5 µg	6,386	0,000	0,006		30573,625		1093,699		
	H3K4me3	1 µg	13,737	0,007	0,009		2055,110*		1509,652*		
pbMEC	IgG	1 µg	0,033	0,001	0,022		31,341		1,490		
	H3K4me3	0.5 µg	2,567	0,035	0,065		73,262*		39,671*		
	H3K4me3	1 µg	5,352	0,084	0,237		63,338		22,549		
			GAPDH_TSS_3	PRSS1_TSS	TSH2B		TSH2B/GAPDH_TSS_3		PRSS1/GAPDH_TSS_3		
MAC-T	IgG	1 µg	0,000	0,002	0,021		65,345		7,438		
	H3K27me3	0.5 µg	0,017	0,714	0,360		20,894*		41,499		
	H3K27me3	1 µg	0,055	1,932	0,785		14,320		35,261		
pbMEC	IgG	1 µg	0,033	0,001	0,022		0,671		0,032		
	H3K27me3	0.5 µg	0,051	1,699	1,647		32,223*		33,243		
	H3K27me3	1 µg	0,094	1,586	1,218		12,906		16,795		
			GAPDH_cod	GAPDH_TSS_3	PRRS1_TSS	ActB_1	TSH2B	GAPDH_cod/GAPDH_TSS	GAPDH_TSS/PRSS1_TSS	ActB_1/TSH2B	
MAC-T	IgG	1 µg	0,014	0,000	0,002	0,000	0,012	45,098	0,134		
	H3K4me1	0.5 µg	0,013	0,204	0,013	0,435	0,025	0,061	15,780	17,753	
	H3K4me1	1 µg	0,037	0,631	0,031	1,723	0,105	0,059	20,678	16,450*	
			GAPDH_TSS_3	PRRS1_TSS	TSH2B		GAPDH_TSS/TSH2B		GAPDH_TSS/PRSS1		
MAC-T	IgG	1 µg	0,000	0,002	0,021		0,015		0,134		
	H3K27Ac	0.5 µg	0,344	0,010	0,001		382,681*		33,591		
	H3K27Ac	1 µg	0,186	0,004	0,006		28,740		50,388		
			ActB_3	H19imprint_TSS_3	PRRS1_TSS	TSH2B		H19/TSH2B		H19/PRSS1	ActB_3/TSH2B
MAC-T	IgG	1 µg	0,000	0,000	0,002	0,021		0,002		0,014	0,014
	CTCF	1 µg	4,643	0,099	0,093	0,039		2,514		1,057	118,193
	CTCF	2 µg	4,691	0,112	0,107	0,019		5,796		1,042	243,032*

Table B-2: RNA-seq statistics. The sample ID indicates the cell type (pbMEC or MAC-T), the challenge (control, *E. coli*<sub>1303</sub>, or *S. aureus*<sub>1027</sub>), and the replicate (R1 to R3).

Sample ID	Total raw reads (fastqc)	Total reads (qualtrim)	Singleton	% Reads passed (trimmed)	QC passed reads	Mapped reads	% Mapped reads (versus QC passed)
pbMEC Control R1	71750082	56663934	2286119	82.16	58950053	54958350	96.99
pbMEC E.coli R1	96513130	84265694	3516107	90.95	87781801	81779856	97.05
pbMEC S.aureus R1	111839704	99447662	3427378	91.98	102875040	48242061	97.02
pbMEC Control R2	81487454	71521890	2789921	91.19	74311811	34777519	97.25
pbMEC E.coli R2	69068604	59184516	2328459	89.06	61512975	28577244	96.57
pbMEC S.aureus R2	77320500	68034834	2515806	91.24	70550640	33024108	97.08
pbMEC Control R3	76634202	67404030	2587801	91.33	69991831	32940349	97.74
pbMEC E.coli R3	81835410	71951474	2725659	91.25	74677133	35119514	97.62
pbMEC S.aureus R3	86266198	76417956	2776057	91.80	79194013	37410410	97.91
MAC-T Control R1	64209520	55361352	2062352	89.43	57423704	26955442	97.38
MAC-T E.coli R1	72611748	61013234	2789741	87.87	63802975	29676837	97.28
MAC-T S.aureus R1	67991646	57605974	2414706	88.28	60020680	28054109	97.4
MAC-T Control R2	64278712	54314240	2209987	87.94	56524227	26649282	98.13
MAC-T E.coli R2	68981872	58698760	2398211	88.57	61096971	28747718	97.95
MAC-T S.aureus R2	61366714	53490758	2201332	90.75	55692090	26298731	98.33

Table B-3: ChIP-seq statistics. The **sample ID** indicates the histone marks (*H3K4me3*, *H3K27me3*, *H3K27Ac*, *H3K4me1*) or the transcription factor (*CTCF*) targeted with ChIP-seq. The ID also includes decoding the three replicates of primary cells pbMEC (*AAA*, *AAD*, and *AAF*) or MAC-T used in passages p.59 and 60. The challenge conditions of the cells are abbreviated as *K* (control), *E* (*E. coli*), and *S* (*S. aureus*). The main sequencing and peak calling statistics include the total number of **raw reads** sequenced, the quality **trimmed** and the **mapped reads** to the reference genome ARS-UCD1.2, including both uniquely and multi-mapped reads, the percentage of **mapping efficiency** calculated by mapped reads relative to raw reads, number of **non-redundant filtered reads** that align only to one location in the genome one time, a **unique ratio** that represents the percentage of non-redundant filtered reads divided by mapped reads, and the number of identified **peaks** per sample.

Sample ID	Raw Reads (FastQC)	Trimmed Raw Reads (Trimmatic)	Mapped Reads (BWA)	Mapping efficiency (%) (BWA/FastQC*100)	Non-redundant Filtered Reads (DEDUP)	Unique Ratio (DEDUP/BWA*100)	Peaks
H3K4me3_6_pbMEC_AAA_K	41587186	36358307	36207544	87.06	16492187	45.55	22350
H3K4me3_7_pbMEC_AAA_E	38498037	32907015	31972708	83.05	15492809	48.46	21292
H3K4me3_8_pbMEC_AAA_S	48979415	42796287	42628004	87.03	16992075	39.86	21410
H3K4me3_9_pbMEC_AAD_K	41654963	36391216	36218719	86.95	14570829	40.23	21986
H3K4me3_10_pbMEC_AAD_E	41504912	36033245	35614758	85.81	14841489	41.67	21344
H3K4me3_11_pbMEC_AAD_S	41664160	36369597	36146738	86.76	16034115	44.36	21209
H3K4me3_15_pbMEC_AAF_K	49263939	43017492	42770899	86.82	16103936	37.65	20218
H3K4me3_16_pbMEC_AAF_E	40776858	35517915	35227220	86.39	14495374	41.15	21408
H3K4me3_17_pbMEC_AAF_S	43123411	37478654	37095948	86.02	14305570	38.56	21261
H3K4me3_20_MACT_p59_K	42889054	37444176	37265567	86.89	17837112	47.86	24304
H3K4me3_21_MACT_p59_E	41414929	36115401	35875347	86.62	18089732	50.42	25496
H3K4me3_22_MACT_p59_S	50157584	43789132	43561135	86.85	19498936	44.76	30197
H3K4me3_23_MACT_p60_K	52834989	46036079	45724433	86.54	21089327	46.12	28464
H3K4me3_24_MACT_p60_E	42421864	36960902	36691365	86.49	18429727	50.23	28137
H3K4me3_25_MACT_p60_S	37658340	32804683	32563162	86.47	17219544	52.88	23886
H3K4me3_26_MACT_p59_K	49894142	43499712	43222897	86.63	20547867	47.54	29247

H3K4me3_27_MACT_p59_E	39575054	34477938	34229290	86.49	18085629	52.84	25058
H3K4me3_28_MACT_p59_S	52022466	45357031	45057595	86.61	21031044	46.68	28810
H3K27me3_6_pbMEC_AAA_K	55527416	48628471	48279782	86.95	28741874	59.53	4011
H3K27me3_7_pbMEC_AAA_E	67138558	56058333	52799015	78.64	26341110	49.89	3657
H3K27me3_8_pbMEC_AAA_S	65746549	57646741	57289982	87.14	30438252	53.13	4076
H3K27me3_9_pbMEC_AAD_K	59399853	52035402	51715404	87.06	27949717	54.05	3654
H3K27me3_10_pbMEC_AAD_E	54031011	47250085	46843193	86.70	26704926	57.01	3589
H3K27me3_11_pbMEC_AAD_S	58618879	51253746	50788028	86.64	26884627	52.93	3494
H3K27me3_15_pbMEC_AAF_K	62657086	50889144	46720375	74.57	24015241	51.40	3512
H3K27me3_16_pbMEC_AAF_E	61120945	53495217	52993568	86.70	23332170	44.03	3439
H3K27me3_17_pbMEC_AAF_S	76894751	66688641	65533830	85.23	25710552	39.23	3322
H3K27me3_20_MACT_p59_K	58536483	50870774	50233714	85.82	23791284	47.36	2776
H3K27me3_21_MACT_p59_E	59213161	51630547	51078570	86.26	29137767	57.04	2813
H3K27me3_22_MACT_p59_S	65311887	57230133	56867132	87.07	29790874	52.39	2798
H3K27me3_23_MACT_p60_K	45189388	39487584	39133015	86.60	20963843	53.57	2722
H3K27me3_24_MACT_p60_E	53542333	46467341	45701792	85.36	22080464	48.31	2817
H3K27me3_25_MACT_p60_S	62557234	51772250	48268104	77.16	27473772	56.92	2947
H3K27me3_26_MACT_p59_K	60441519	52943966	52537475	86.92	24426657	46.49	2834
H3K27me3_27_MACT_p59_E	65147869	56833398	56248469	86.34	24434551	43.44	2822
H3K27me3_28_MACT_p59_S	60737348	53130902	52724788	86.81	27899547	52.92	2775
CTCF_MACT_P59_K	49554330	43233435	42894861	86.56	24544829	57.22	105680
CTCF_MACT_p59_E	46468505	40443345	40017547	86.12	22436152	56.07	116308
CTCF_MACT_P60_K	39803224	34624436	34268529	86.09	19726986	57.57	111482
CTCF_MACT_P60_E	42010177	36565302	36208536	86.19	21078671	58.21	113294
H3K27Ac_MACT_P59_K	40635778	35189303	34588292	85.12	8602066	24.87	32764
H3K27Ac_MACT_p59_E	39130990	33820814	33194533	84.83	12944448	39.00	38904
H3K27Ac_MACT_P60_K	31620421	26820401	25827956	81.68	6110795	23.66	30909
H3K27Ac_MACT_P60_E	37666377	32026265	30849055	81.90	8345307	27.05	29488
H3K4me1_MACT_P59_K	40893229	35824166	35611835	87.08	22608599	63.49	76255
H3K4me1_MACT_p59_E	39138199	34241813	33970934	86.80	22359254	65.82	75940
H3K4me1_MACT_P60_K	29118664	25488306	25320885	86.96	18499846	73.06	63220
H3K4me1_MACT_P60_E	33295124	29124387	28868350	86.70	18435291	63.86	61573
INPUT_pool_MACT_E	34106822	28619748	27045631	79.30	23353604	86.35	
INPUT_pool_MACT_K	33909876	29804219	29650897	87.44	24132220	81.39	
INPUT_pool_MACT_S	53773316	47287103	47057259	87.51	35899001	76.29	
6_pbMEC_AAA_K_INPUT	46067697	40488493	40330989	87.55	28816032	71.45	
7_pbMEC_AAA_E_INPUT	44920335	38764386	37818543	84.19	26909538	71.15	
8_pbMEC_AAA_S_INPUT	48132905	42254723	42068885	87.40	29742413	70.70	
9_pbMEC_AAD_K_INPUT	43806073	38533401	38395366	87.65	27068010	70.50	
10_pbMEC_AAD_E_INPUT	42355594	36516369	35587947	84.02	26305732	73.92	
11_pbMEC_AAD_S_INPUT	47063970	41384276	41201116	87.54	28807332	69.92	
15_pbMEC_AAF_K_INPUT	53376791	46844745	46652260	87.40	31954885	68.50	
16_pbMEC_AAF_E_INPUT	46362219	39744489	38553560	83.16	27688266	71.82	
17_pbMEC_AAF_S_INPUT	43271249	37985469	37806391	87.37	29458568	77.92	

*Table B-4: Average of ChIP peaks. The sample replicates were used to calculate the mean number of identified peaks, their standard deviation, and relative standard deviation in percent.*

<b>Sample ID</b>	<b>Mean of Identified Peaks</b>	<b>Standard Deviation</b>	<b>Relative Standard Deviation (%) (SD/Mean*100)</b>
H3K4me3_pbMEC_K	21518	1140	5.30
H3K4me3_pbMEC_E	21348	58	0.27
H3K4me3_pbMEC_S	21293	104,33	0.49
H3K4me3_MAC-T_K	27338	2657	9.72
H3K4me3_MAC-T_E	26230	1666	6.35
H3K4me3_MAC-T_S	27631	3316,58	12.00
H3K27me3_pbMEC_K	3726	257	6.90
H3K27me3_pbMEC_E	3562	112	3.13
H3K27me3_pbMEC_S	3631	395,14	10.88
H3K27me3_MAC-T_K	2777	56	2.02
H3K27me3_MAC-T_E	2817	5	0.16
H3K27me3_MAC-T_S	2840	93,38	3.29
CTCF_K	108581	4103	3.78
CTCF_E	114801	3412,50	2.97
H3K27Ac_K	31837	1312	4.12
H3K27Ac_E	34196	6658,12	19.47
H3K4me1_K	69738	9217	13.22
H3K4me1_E	68757	10159	14.78



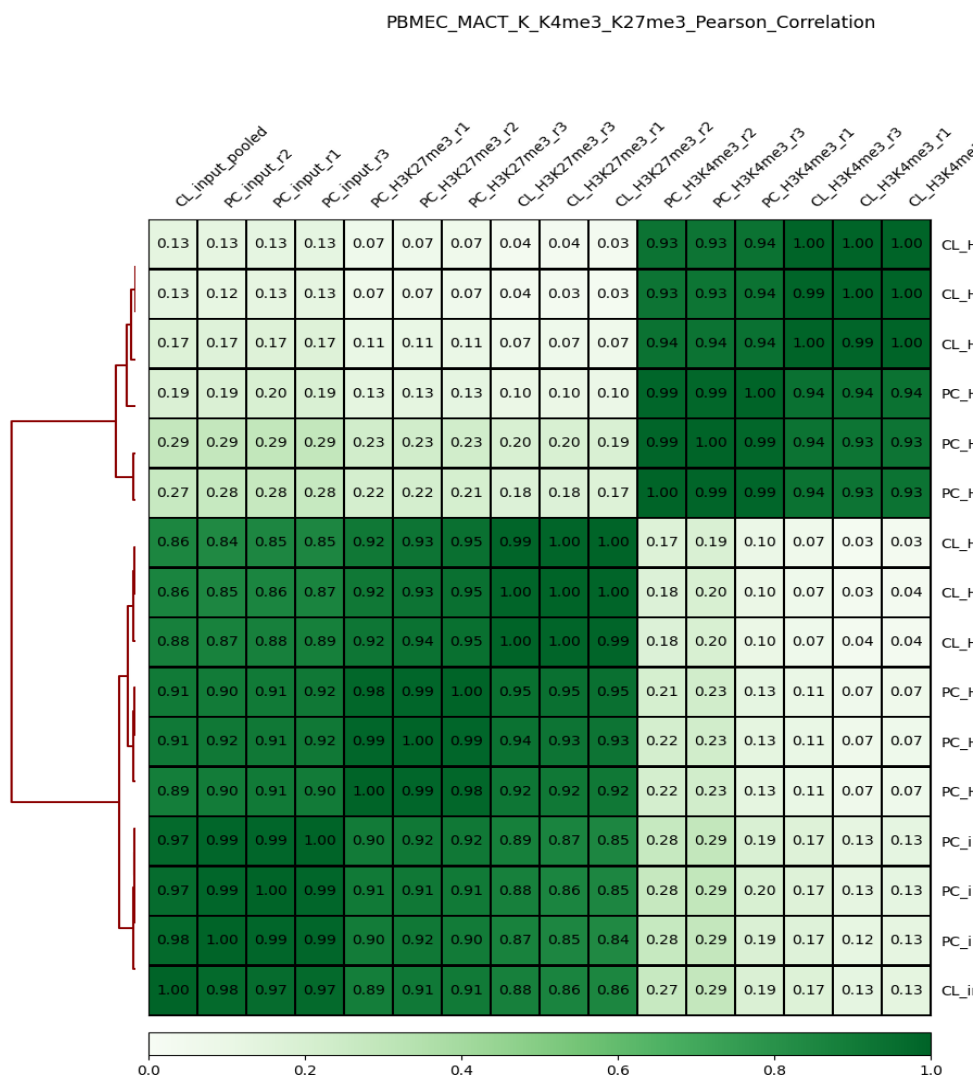


Figure B-15: Correlation heatmap. The diagram represents a correlation matrix of ChIP-seq samples from different cell types (primary cells = PC or cell line = CL) tested for two different histone marks clustered hierarchically. The intense color indicates a strong positive correlation, while the brighter color indicates no correlation. Inputs represent control DNA that did not undergo the ChIP procedure but have been sequenced to determine the background signal.

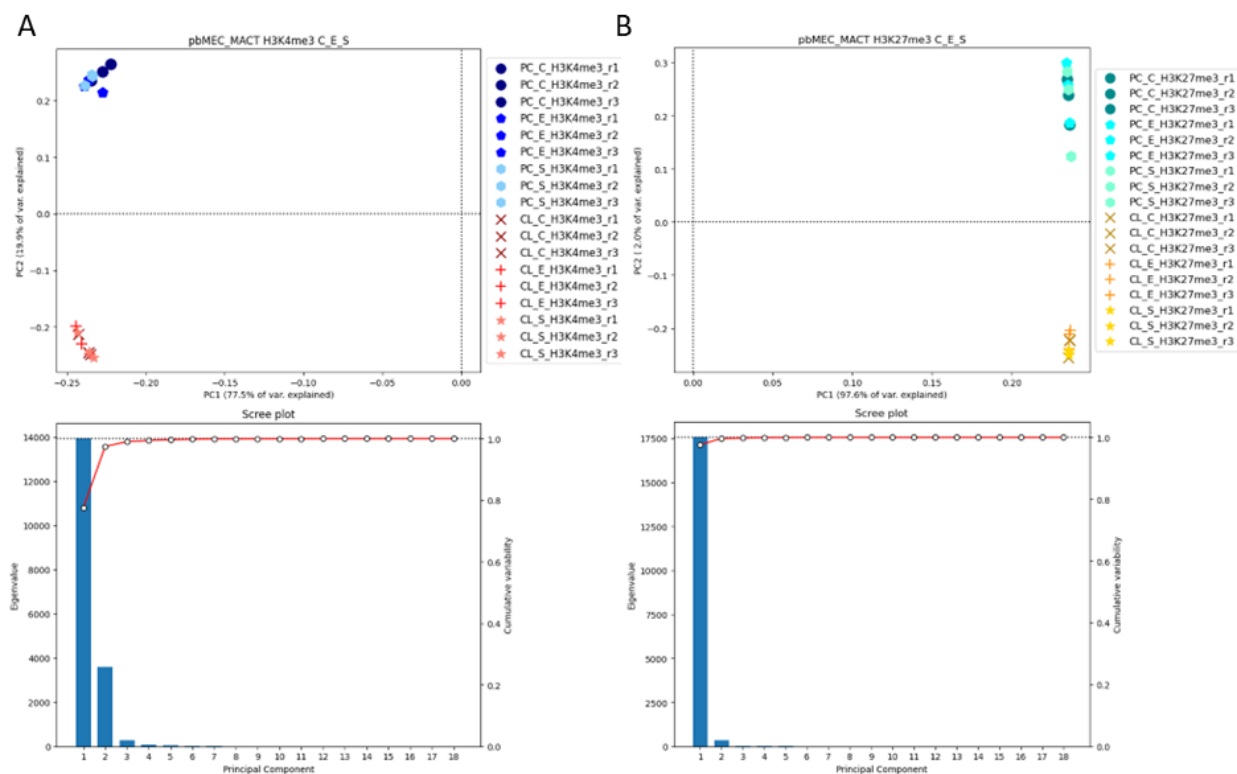


Figure B-16: PCA of ChIP-seq for challenged cells. The two ChIP assays, H3K4me3 (A) and H3K27me3 (B), are shown in the PCA. The scatter of sample replicates (r) for the primary cells (PC) pbMEC, shown as dots, and the cell line (CL) MAC-T, shown as crosses and stars, is based on the similarity or difference of the data along the principal components along the axes. The scree plot supports the emphasis on the first two principal components as the driving factor for the analysis.

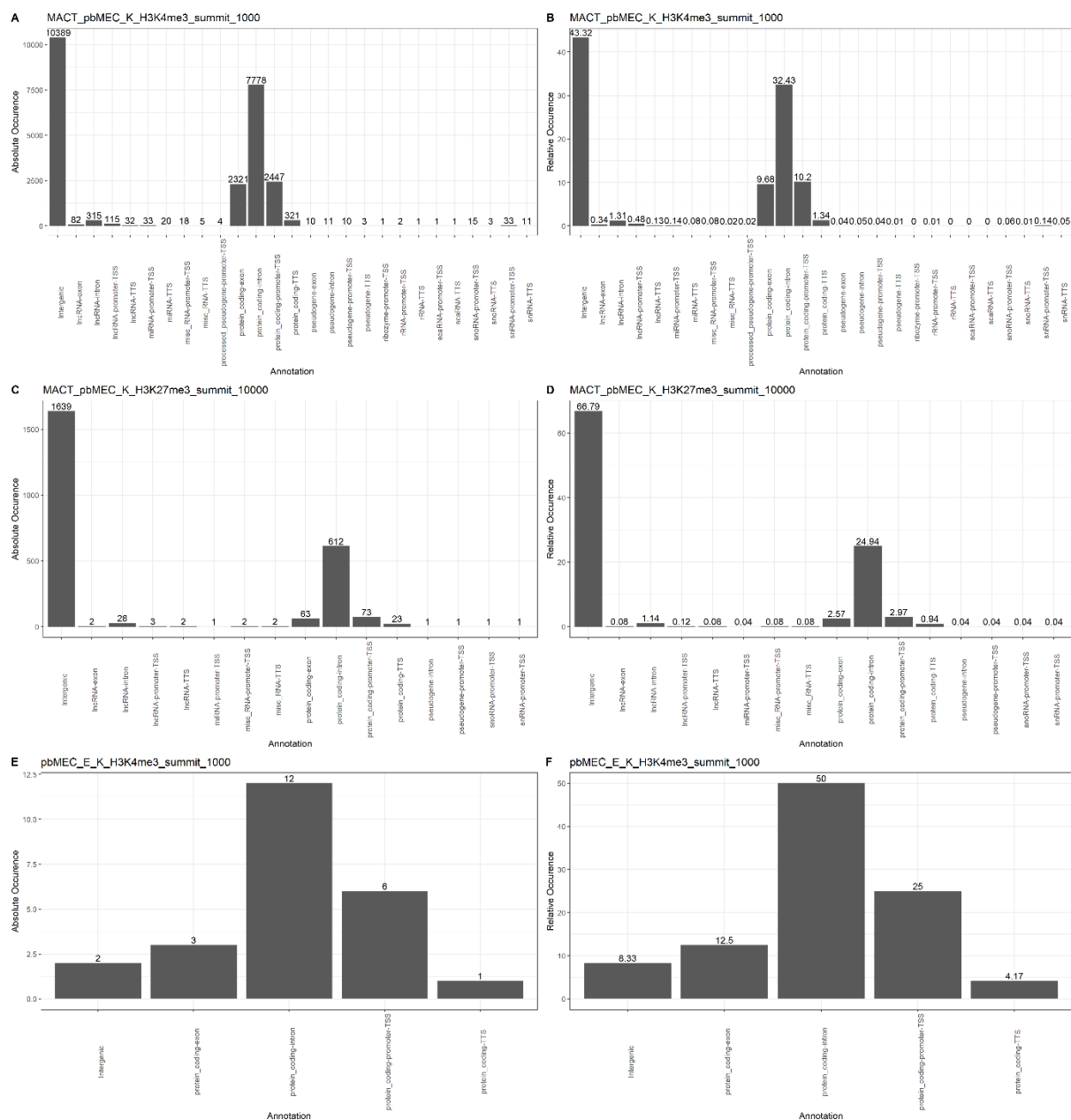


Figure B-17: Detailed ChIP-seq peak annotation of H3K4me3 and H3K27me3 in MAC-T and pbMEC. Bar charts show absolute and relative occurrence in genomic regions.

## Acknowledgment

I am grateful to all those who have contributed to the success of this work through their professional and personal support.

I would like to thank, Dr. Juliane Günther, for the opportunity to dedicate myself to this research. Her suggestions and guidance on experimental design, analysis, and her guiding role in the DFG project on epigenetics in cattle have contributed to the success of this study.

A heartfelt thanks goes to my academic mentor and supervisor, Prof. Dr. Christa Kühn. Her guidance and infinite patience were unsurpassed throughout my doctoral studies at the Leibniz Institute for Farm Animal Biology (FBN) in Dummerstorf and beyond. Furthermore, I am grateful for the constructive advice, the help with data analysis and the review of my written thesis.

I am equally grateful for the provided supervision by Prof. Dr. Reinhard Schröder at the Biological Institute of the University of Rostock. His persistent encouragement and advisory commitment have been a continuous source of motivation.

I also enjoyed working with past and present Genome Physiology Unit lab members and colleagues at the Leibniz Institute of Farm Animal Biology (FBN) in Dummerstorf. The pleasant working atmosphere and teamwork made the years memorable. I am incredibly grateful for the excellent technical support of Simone Wöhl, Bärbel Pletz, Emma Schröder, and Angelika Deike during the years of my Ph.D. at the FBN. I was also fortunate to be surrounded by a knowledgeable and outstanding group of colleagues with whom I enjoyed good coffee, cookies, and inspiring conversations. Especially, Dr. Wietje Nolte, Frieder Hadlich, Franziska Just, Dr. Alexander Rebl, Dr. Wiebke Demasius, Dr. Mareen Nipkow, and Kaja Heide provided me with moral support—many thanks to Frieder for the computational support with the sequencing data.

A special thank you to Geoffrey Berguet and Raphael Werdning for their support with the scientific exchange at Diagenode in Belgium. I would also like to thank Dr. Catherine Creppe, Dr. Sol Schvartman, Dr. Miklós Laczik, Dr. Matteo Tosolini and Dr. Paola Genevini for their generous hospitality and help in Liège.

I would like to thank the German Research Foundation (DFG) for the financial support of the research project GU 1487/1-1, the COST Action CA15112- Functional Annotation of Animal Genomes - European network (FAANG-Europe) for enabling my research exchange and the grants of the Graduate Academy of the University of Rostock and the Fazit Foundation for enabling my participation in external qualification courses, which greatly enriched my learning experience.

I am deeply grateful to Dr. Maximilian Berthold for his help while revising my work. His advice and impulses were greatly appreciated.

A special thank you goes to Dr. Fanny Romoth, Dr. Ekaterina Sergeeva, Dr. Katrin Gärtner, Dr. Merel Hofmeijer, and all other participants, as well as the coordination team of the University of Rostock

mentoring program for female PhDs (2016-2017). I also extend my thanks to Prof. Dr. Vett Lloyd and her team for their professional and individual support since 2019.

A big thank you to all my friends who supported me in word and deed. These people include Dr. Janin Knop, Dr. Henry Göhlich, Dr. Miriam Paulisch, Dr. Anne Haack, Dr. Julia Leonardy, Dr. Celina Genschel, Samantha Bishop, D&G, the Tomlik family, the Prest family, and Maria Thistle. Furthermore, I would like to express my deep gratitude to my family for their boundless patience, support, time, motivation and unwavering love during the difficult stages of this journey. All of you have played an important role in shaping my career and my identity, and I owe who I am today to you all. Finally, a special thanks goes to C<sup>3</sup> and M<sup>2</sup>!

## **Eidesstattliche Erklärung**

I, Anne Berthold, hereby declare in lieu of an oath that I have written this thesis independently and have not used any sources or aids other than those indicated. All statements taken verbatim or in spirit from other writings are identified. The work has not been presented in the same or similar form or in excerpts as part of any other examination.

Ich, Anne Berthold, erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit eigenständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle wörtlich oder sinngemäß aus anderen Schriften übernommenen Aussagen sind gekennzeichnet. Die Arbeit wurde nicht in gleicher oder ähnlicher Form oder in Auszügen im Rahmen einer anderen Prüfung vorgelegt.

Date/ Datum

Anne Berthold

10.06.2024

## **Lebenslauf – Anne Berthold, geb. Stielow**

### **Ausbildung**

**Promotionsstudium in Genetik** seit 07/2014 bis 06/2024

Universität Rostock, Deutschland

**Master of Science in Biochemie** 10/2011 bis 02/2014

Medizinische Hochschule Hannover (MHH), Deutschland

**Bachelor of Science in Biowissenschaften** 10/2007 bis 09/2010

Universität Rostock, Deutschland

**Abitur** 07/2007

Gymnasium Gadebusch, Deutschland

### **Akademische Leistungen**

**Masterarbeit** 07/2013 bis 02/2014

Thema: *Rolle der microRNAs 455-3p und 545 im Zusammenhang mit Mitochondrien-vermittelte Apoptose in verschiedenen Krebszelllinien*

Betreuer:

Prof. Dr. Brigitte M. Pützer und Dr. Vijay Alla, Institut für Experimentelle Gentherapie und Krebsforschung, Universitätsklinikum Rostock, Deutschland

Prof. Dr. Jürgen Alves, Institut für Biophysikalische Chemie, Medizinische Hochschule Hannover, Deutschland

**Erasmus Praktikum** 03/2013 bis 06/2013

Thema: *Reptin und Pontin im hepatozellulären Karzinom; Etablierung eines Degron-Systems für das Protein Flag-Reptin*

Betreuer: Dr. Jean Rosenbaum, INSERM U1053, Segalen Universität, Bordeaux, Frankreich

**Praktikum** 08/2012 bis 10/2012

Thema: *Untersuchung von E2F1 im Zusammenhang der Tumorgenese*

Betreuung: Dr. David Engelmann, Dr. Deborah M. Nüssle, Dr. Ottmar Herchenröder, Institut für Experimentelle Gentherapie und Tumorforschung, Universitätsmedizin Rostock

**Bachelorarbeit** 04/2010 bis 06/2010

Thema: *Untersuchung der Zytotoxizität und der Mikrotubuli-depolymerisierenden Aktivität neuer Podophyllotoxin-Derivate*

Betreuer: PD Sergei. A. Kuznetsov und Prof. Dr. Dieter. G. Weiss, Institut für Biowissenschaften, Abteilung für Tierphysiologie, Universität Rostock, Deutschland

### **Berufserfahrung**

**Wissenschaftliche Mitarbeiterin** seit 04/2024

Kompetenzzentrum für Wasserwirtschaft, Halle (Saale), Deutschland

**Wissenschaftliche Mitarbeiterin** 07/2019 bis 07/2023

Lloyd tick lab, Fachbereich Biologie, Mount Allison Universität, Sackville, NB, Kanada

**Labor-Managerin** 09/2022 bis 05/2023

Geneticks, Atlantic Canada Branch, Sackville, NB, Kanada

**Wissenschaftliche Mitarbeiterin** 07/2014 bis 03/2019

Institut für Genombiologie, Abteilung Genomphysiologie, Leibniz-Institut für Nutztierbiologie, FBN, Dummerstorf, Deutschland

**Gastwissenschaftlerin** 01/2018 bis 03/2018

Diagenode, Service Team, Liège, Belgien

**Technische Assistentin** 05/2014 bis 06/2014

Arbeitsgruppe Dr. Katrin Manda, Radiobiologisches Labor, Universitätsklinikum Rostock, Rostock, Deutschland

**Wissenschaftliche Hilfskraft** 10/2010 bis 06/2011

Arbeitsgruppe von Dr. Ralf Bastrop Abteilung Tierphysiologie, Universität Rostock  
*Untersuchung zur molekularen Systematik mariner Wirbeltiere und Wirbelloser.*

**Weiterbildung****Projektmanagement** 12/2023 bis 01/2024

Zertifizierte Weiterbildung in Projektmanagement, Alfatraining Bildungszentrum GmbH  
Basiszertifikat Projektmanagement, Deutsche Gesellschaft für Projektmanagement e.V. (GPM)

**Life Science Manager** 09/2023 bis 12/2023

Zertifizierte Weiterbildung in Good clinical practice (GCP), Good manufacturing practice (GMP), Qualitätsmanagement (QM) für Medizinprodukte und Projektmanagement  
Alfatraining Bildungszentrum GmbH

**Micro Certificate in Teaching and Learning** 09/2022 bis 06/2023

Vertiefung innovativer pädagogischer Praktiken, Erstellung eines Lehrportfolios.  
Maple League of Universities, Kanada

**Communicating with Government and Industry** 06/2022

Clarity Studios and New Brunswick Innovation Foundation

**Mental Health First Aider** 09/2020

Mental Health First Aid Canada

**Programmieren mit R** 05/2016 & 06 bis 07/2018

Thünen Institut, Braunschweig & Leibniz-Institut für Nutztierbiologie, FBN, Dummerstorf, Deutschland

**Mentoring für Doktorandinnen** 03/2016 bis 06/2017

Peer Mentoring, Mentee-Mentor-Beziehung, Soft-Skill-Training  
KarriereWegeMentoring, Universität Rostock, Deutschland

**Publikationen:****Zeitschriftenartikel:**

A Berthold\*, ML Faucillion\*, I Nilsson, M Golovchenko, V Lloyd, S Bergström, N Rudenko. *Cultivation Methods of Spirochetes from Borrelia burgdorferi Ssensu Lato Complex and Relapsing Fever Borrelia.* J Vis Exp. 2022 Nov 25;(189). doi: 10.3791/64431. \* Geteilte Erstautorenschaft

J Günther, M Koy, A Berthold, HJ Schuberth, HM Seyfert. *Comparison of the pathogen species-specific immune response in udder derived cell types and their models.* Vet Res. 2016 Feb 1;47:22. doi: 10.1186/s13567-016-0307-3.

MN Semenova, AS Kiselyov, DV Tsyganov, LD Konyushkin, SI Firgang, RV Semenov, OR Malyshev, MM Raihstat, F Fuchs, A Stielow, M Lantow, AA Philchenkov, MP Zavelevich, NS Zefirov, SA



Kuznetsov, VV Semenov. *Polyalkoxybenzenes from plants. 5. Parsley seed extract in synthesis of azapodophyllotoxins featuring strong tubulin destabilizing activity in the sea urchin embryo and cell culture assays*. J Med Chem. 2011 Oct 27;54(20):7138-49. doi: 10.1021/jm200737s.

**Konferenzbeiträge:** ° =Vortragende

**Vorträge:**

A Berthold°, V Lloyd. *The tick-borne pathogen Borrelia differentially affects gene transcription but not DNA methylation of mammalian cells*. 61st Annual Meeting of the Canadian Society of Zoologists, Moncton, New Brunswick, Canada, 9. -13. 05.2022 (online)

A Berthold°, V Lloyd. *In vitro expression changes after infection of endothelial cells with tick-borne pathogen Borrelia*. Meeting of the Atlantic Canada Association of Parasitologists 2021, Pictou, Nova Scotia, Canada, 23.10. 2021 (online)

A Nuyens°, A Berthold, V Lloyd. *Investigation of methylation status in relation to the presence of Borrelia burgdorferi in black-legged Ticks (Ixodes scapularis)*. Meeting of the Atlantic Canada Association of Parasitologists 2021, Pictou, Nova Scotia, Canada, 23.10.2021

A Berthold°, V Lloyd. *Epigenetic modifications - an additional player in tick borne diseases*. Meeting of the Atlantic Canada Association of Parasitologists 2020, Pictou, Nova Scotia, Canada, 24.10.2020 (online)

**Poster:**

A Berthold°, V Lloyd. *Epigenetics as a potential factor in Lyme disease*. EMBL Conference: Chromatin and epigenetics, Heidelberg, Deutschland, 15.- 18.05.2023,

A Berthold°, V Lloyd. *RNA sequencing of human cells provides insight into the role of non-coding RNAs and genes affected during infection with the Lyme disease pathogen Borrelia burgdorferi*. The International Human Epigenome Consortium (IHEC) Annual Meeting and the 8th Canadian Conference on Epigenetics, Esterel, Quebec, Canada, 4. - 7.10.2022

A Berthold°, V Lloyd. *Borrelia burgdorferi induces gene and ncRNA alterations but not DNA methylation in human cells*. 71st Annual Conference of the Canadian Society of Microbiologists, Guelph, Ontario Canada, 26.-30.06.2022

A Berthold°, V Lloyd. *High-throughput Expression and DNA methylation Screening in Human Cells after Borrelia burgdorferi Infection*. CEEHRC 7th Annual Canadian Conference on Epigenetics. 3. - 5.11.2021 (online)

A Nuyens, J Vickery, A Berthold°, V Lloyd. *Investigation of Methylation Status in Relation to the Presence of Borrelia burgdorferi in Ixodes scapularis (Black-legged Ticks)*. CEEHRC 7th Annual Canadian Conference on Epigenetics. 3. - 5.11.2021 (online)

A Berthold°, C Creppe, S Schwartzman, Ch Kühn, J Günther. *The path to ChIP-Seq: optimization of ChIP assays for bovine udder cell models*. Second Workshop Next-Generation-Sequencing, Dummerstorf, Deutschland, 8.10.2018

J Hehl°, M Koy, A Berthold, HJ Schuberth, M Weinert, S Engelmann, Ch Kühn, HM Seyfert, J Günther. *Mammary epithelial cells, rather than professional immune cells dictate the pathogen species-specific immune reaction of the udder*. 36th International Society for Animal Genetics Conference, (ISAG), Dublin, Irland, 16. - 21.07.2017

J Günther°, M Koy, A Berthold, HJ Schuberth, HM Seyfert. *Mammary epithelial cells, rather than professional immune cells dictate the pathogen-species specific immune reaction of the udder during mastitis*. World Buiatrics Congress 2016 (WBC 2016), Dublin, Irland. 3.- 8.07.2016

A Stielow<sup>o</sup>, HM Seyfert, W Petzl, Ch Kühn, J Günther. *Role of epigenetic regulatory mechanisms in pathogen specific mastitis of the cow*. 18th International Summer School on Immunology, FEBS Advanced Lecture Course, Rabac, Kroatien, 12.-19.09.2015

Halle (Saale), den 10.06.2024