A multi-omics analysis of the postnatal development of intestinal epithelial cells

Von der Fakultät für Mathematik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Johannes Schöneich, M.Sc.

aus

Münster, Deutschland

Berichter:
Univ.-Prof. Dr. med. Mathias Hornef
Univ.-Prof. Dr. rer. nat. Raplph Panstruga
Univ.-Prof. Dr. rer. nat. Lars Küpfer

Tag der mündlichen Prüfung: 27.05.2026

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.

Table of contents

1. Abstract 3
2. Zusammenfassung 4
3. List of Figures 5
4. Abbreviation list 7
5. Table of contributors 12
6. Introduction 13
6.1. Structure and main functions of the small intestine 13
6.2. Intestinal epithelial cells (IECs) 14
6.2.1. Stem and transit amplifying cells 14
6.2.2. Absorptive enterocytes 15
6.2.3. Goblet cells 15
6.2.4. Paneth cells 16
6.2.5. Enteroendocrine cells (EEC) 17
6.2.6. Tuft cells 17
6.3. Regulators of spatial identity in the intestinal epithelium 18
6.4. Structural and functional differences between neonatal and adult intestines 20
6.4.1. Key developmental and structural features 20
6.4.2. Fetal-type enterocytes and the immature intestinal epithelium 22
6.4.3. BLIMP1 and MAFB as regulators of the neonate-to-adult transition 22
6.4.4. Breast milk as a source of nutrients 23
6.4.5. How early life microbiota colonization shapes the immune system 24
6.5. *Salmonella* as a major causative agent of gastrointestinal infections 25
6.6. Molecular profiling approaches and techniques 26
6.7. Aim 28
7. Material & Methods 29
7.1. Laboratory work 29
7.1.1. Animals 29
7.1.2. Sample preparation for bulk RNA-seq and proteomics 29
7.1.3. Crypt and villi sample preparation for 3’ RNA-seq 30
7.1.4. Sample preparation for single-cell RNA-seq 30
7.1.5. Mass spectrometry analysis of proteomic data 31
7.2. Bioinformatic analysis 32
7.2.1. Preprocessing 32
7.2.2. Sex- and gender-based analyses 32
7.2.3. Batch correction 32
7.2.4. Principal component and differential expression analysis 34
7.2.5. Temporal DE gene analysis 34

7.2.6. Analysis of proteomic LFQ values 34
7.2.7. Analysis of single-cell data 35
7.2.8. Transcription factor analysis 36
7.2.9. Gene Ontology and pathway analysis 36
7.2.10. RNA Velocity 37
7.3. Research data management 37
7.4. Code availability 37
8. Results 38
8.1. Bulk RNA-seq reveals developmental factors, instead of microbial factors, as main drivers of the early postnatal period 38
8.2. Proteomic transformations highlight developmental adaptations in small intestinal epithelium during weaning 43
8.3. Age-related shifts in gene expression along the crypt-villus axis 48
8.4. Segment-specific transcriptional profiles along the proximal-to-distal axis 52
8.5. Single-cell analysis reveals age- and cell-type-specific developmental trajectories 56
8.6. Global and cell-type-specific responses upon infection with *Salmonella* Typhimurium 63
9. Discussion 66
9.1. Summary of key findings 66
9.2. Wnt signaling changes during neonatal intestinal development 67
9.3. Age-dependent programs override microbiota influences 68
9.4. Complementary roles of transcriptomics and proteomics 69
9.5. Early manifestation of immune-related epithelial programs in the distal small intestine 69
9.6. Early spatial positional signals shape neonatal epithelial differentiation 70
9.7. Neonatal foundations of adult intestinal domain architecture 71
9.8. Cell type-specific neonatal epithelial maturation by transcriptional networks 73
9.9. Neonatal epithelial adaptation to *Salmonella* infection 74
9.10. Technical considerations and limitations of the study 75
10. Concluding remarks 78
11. Acknowledgements 80
12. Supplementary Figures 81
13. References 87
14. Eidesstattliche Erklärung 102

1. Abstract

The adult murine small intestinal epithelium is a highly dynamic and continuously renewing tissue composed of multiple specialized epithelial cell types arranged along defined anatomical and functional axes. In contrast, the neonatal intestine differs markedly in both structure and function. Despite the importance of early life as a period marked by the shift from placental nutrient supply to enteral feeding, the initial establishment of the intestinal microbiota, and rapid immune maturation, the temporal, regional, and cellular heterogeneity of the small intestinal epithelium during this postnatal window has not been comprehensively characterized.

This thesis combines complementary approaches, including bulk RNA sequencing, global proteomics, and single-cell RNA sequencing, to characterize the intestinal epithelium at multiple levels. By integrating whole-tissue, regional (proximal-to-distal), and crypt-villus-specific analyses, the epithelial landscape was examined in specific pathogen-free, germ-free, and *Salmonella*-infected mice throughout the postnatal period. Bulk RNA-seq data were processed with DESeq2 to identify differentially expressed genes, proteomic data were analyzed with limma to detect differentially abundant proteins, and single-cell RNA-seq data were analyzed with Seurat, including normalization, clustering, and identification of cell-type-specific expression patterns. These analyses were followed by Gene Ontology, pathway enrichment, and clustering analyses to interpret functional and regulatory signatures.

Global bulk transcriptomics and mass-spectrometry-based proteomics identified age-dependent shifts in metabolic and immune pathways, while complementary laser capture microdissection of crypt and villus areas showed early crypt-villus gradients despite the absence of sharply defined adult-like architecture. Regional analyses along the proximal-to-distal axis demonstrated that spatial programs of epithelial specialization are already detectable early after birth and become progressively refined with maturation, including the emergence of distal immune signatures prior to full structural development, while single-cell RNA sequencing resolved the cellular heterogeneity underlying these changes and highlighted dynamic transitions in epithelial lineage composition.

Taken together, the results presented in this thesis demonstrate coordinated epithelial differentiation, metabolic adaptation, and host-microbe interactions in the neonatal intestine, highlighting early distal immune signatures prior to full maturation. They extend current knowledge of neonatal intestinal biology and provide a framework for understanding vulnerabilities of the immature gut and potentially increase our understanding of intestinal health and immune response in children and preterm infants.

2. Zusammenfassung

Das Epithel des adulten murinen Dünndarms ist ein hochdynamisches, kontinuierlich erneuertes Gewebe, das aus spezialisierten Epithelzelltypen besteht, die entlang definierter anatomischer und funktioneller Achsen organisiert sind. Im Gegensatz dazu unterscheidet sich der Darm von neugeborenen Mäusen deutlich in Struktur und Funktion. Obwohl die frühe postnatale Phase durch den Übergang von der intrauterinen Versorgung zur Nahrungsaufnahme über den Gastrointestinaltrakt, die initiale Etablierung des Darmmikrobioms sowie die rasche Entwicklung der angeborenen Immunantwort geprägt ist, wurde die zeitliche, regionale und zelluläre Heterogenität des Dünndarmepithels in diesem Zeitraum bislang nicht umfassend charakterisiert.

Diese Arbeit kombiniert komplementäre Omics-Ansätze wie Bulk-RNA-Sequenzierung, globale Proteomik und Einzelzell-RNA-Sequenzierung, um das Darmepithel auf mehreren Ebenen zu analysieren. Durch die Integration von Ganzgewebe-, proximal-distalen sowie krypten- und zottenspezifischen Analysen wurde die epitheliale Landschaft bei spezifisch pathogenfreien, keimfreien und mit *Salmonella* Typhimurium infizierten Mäusen über die gesamte postnatale Phase hinweg untersucht. Differenzielle Expressions- und Abundanzanalysen (DESeq2, limma, Seurat) wurden durch Gene Ontology-, Signalweg- und Clusteringanalysen ergänzt.

Globale Bulk-Transkriptomik und Proteomik zeigten altersabhängige Verschiebungen metabolischer und immunologischer Signalwege. Lasermikrodissektion offenbarte frühe Krypten-Zotten-Gradienten, obwohl eine klar abgegrenzte, adultähnliche Architektur noch nicht etabliert war. Regionale Analysen entlang der proximal-distalen Achse zeigten, dass regionale Spezialisierungsprogramme bereits kurz nach der Geburt erkennbar sind und sich im Verlauf der postnatalen Entwicklung weiter ausprägen. Dazu gehört unter anderem das Auftreten distaler Immunsignaturen vor der vollständigen strukturellen Entwicklung des Darms. Die Einzelzell-RNA-Sequenzierung machte die zugrunde liegende zelluläre Heterogenität sowie dynamische Verschiebungen in der Zusammensetzung epithelialer Zelllinien sichtbar.

Zusammenfassend zeigen die Ergebnisse eine koordinierte epitheliale Differenzierung, metabolische Anpassung und Wirt-Mikroben-Interaktion im neonatalen Darm. Sie erweitern das Verständnis der frühen Darmbiologie und liefern einen Rahmen zur Einordnung der erhöhten Vulnerabilität des unreifen Darms, insbesondere im Kontext kindlicher und frühgeborener Entwicklung.

3. List of Figures

Figure 1: The interaction between BMP and Wnt signaling establishes the different epithelial turnover rates and villus length in the proximal and distal small intestine. 19
Figure 2: Schematic overview of the development of the murine small intestine. 28
Figure 3: Comparison of different batch correction methods for the global transcriptomic data. 33
Figure 4: Principal component analysis (PCA) and biplot of the global transcriptomic datasets. 38
Figure 5: Differential gene expression across time in the global transcriptomic data. 39
Figure 6: Gene Ontology analysis of differentially expressed genes in global transcriptomic data. 40
Figure 7: Pathway activity inference on the global transcriptomic data using PROGENy. 41
Figure 8: Temporally DE gene expression patterns and association with known major transcriptional regulators BLIMP1 and MAFB. 41
Figure 9: Gene Ontology Analysis of the Temporal DE genes of the global transcriptome analysis. 42
Figure 10: Principal component and differential gene analysis of SPF vs. GF mice in the global transcriptome of SPF and GF mice. 43
Figure 11: Comparative Analysis of RNA-seq and Proteomics. 43
Figure 12: Principal component analysis (PCA) and biplot of changes over time in the proteomic data. 44
Figure 13: Differential gene expression across time in the proteomic data. 45
Figure 14: Temporally DE protein expression patterns and corresponding Gene Ontology analysis in the proteomic data. 46
Figure 15: Expression of proteins that are absent at the transcriptomic level in the Tabula Muris. 47
Figure 16: Pathway activity inference on the proteomic data using PROGENy. 48
Figure 17: Spatial epithelial gene expression along the crypt-villus axis. 49
Figure 18: Differential gene expression intra-location and across time along the crypt-villus axis. 50
Figure 19: Gene Ontology analysis of differentially expressed downregulated over time in crypts and villi. 51
Figure 20: Gene Ontology analysis of differentially expressed upregulated over time in crypts and villi. 52
Figure 21: Principal component analysis (PCA) and biplot of samples from the proximal-to-distal axis. 52
Figure 22: Differential gene expression between consecutive time points across the proximal-to-distal small intestine. 54
Figure 23: Spatiotemporal clustering of genes regulated along the proximal-to-distal axis 55
Figure 24: Gene Ontology analysis of the identified spatiotemporal clusters. 56

Figure 25: Quality control and clustering of the single-cell data. 57
Figure 26: Occurrence heatmap of the top 15 transcription factors (TF) identified by DoRothEA. 58
Figure 27: RNA velocity analysis of the single-cell data. 59
Figure 28: Subclustering of absorptive enterocytes. 60
Figure 29: Subclustering of goblet/Paneth cells. 61
Figure 30: Gene Ontology analysis of the absorptive enterocytes subcluster. 61
Figure 31: Gene Ontology analysis of goblet and Paneth cell subcluster. 62
Figure 32: Subclustering of the EEC cells. 62
Figure 33: PCA plot and volcano plot of global transcriptomic data, including the Salmonella-infected data. 63
Figure 34: UMAP visualization of uninfected and infected cells and cell-type-specific differential gene expression upon infection. 64
Figure 35: Volcano plots showing differential gene expression between 5-day-old uninfected and *Salmonella*-infected animals for each epithelial cell type. 64
Figure 36: Gene Ontology analysis of the DE genes of different cell types corresponding to infection.. 65
Supplementary Figure 1: Heatmap of temporally differentially expressed genes based on the global transcriptomic data. 81
Supplementary Figure 2: Gene Ontology analysis of differentially expressed genes in the proteomics data. 81
Supplementary Figure 3: Differential expression between consecutive time points in either crypts or villi. 82
Supplementary Figure 4: Differential gene expression between intestinal segments per time point. .82
Supplementary Figure 5: Gene Ontology analysis for the genes downregulated with age in proximal, medial, and distal small intestine. 83
Supplementary Figure 6: Gene Ontology analysis for the genes upregulated with age in proximal, medial, and distal small intestine. 83
Supplementary Figure 7: Table for the number of cells per single-cell cluster 84
Supplementary Figure 8: DoRothEA Transcription Factor Heatmaps. 84
Supplementary Figure 9: Upset plots for the number of DoRothEA-predicted transcription factors. .84
Supplementary Figure 10: Tables for the number of cells per cluster in absorptive enterocytes and goblet/Paneth cells. (a-b) 84
Supplementary Figure 11: Barplot and table for the frequency and number of cells per EEC subcluster. 85
Supplementary Figure 12: Table of common infection response genes. 86

4. Abbreviation list

ACS	Apical canalicular system
ADAM	A disintegrin and metalloproteinase
Adamts9	A disintegrin and metalloproteinase with thrombospondin motifs 9
AGC	Automatic Gain Control
AMP	Antimicrobial peptide
ANOVA	Analysis of Variance
Apoa1	Apolipoprotein A1
Apoa4	Apolipoprotein A4
Apob	Apolipoprotein B
Aqp8	Aquaporin-8
Arg2	Arginase 2
Arid2	AT-Rich Interaction Domain 2
Ascl2	Achaete-scute complex homolog 2
ASS1	Argininosuccinate synthase 1
Atoh1	Atonal BHLH Transcription Factor 1
Avil	Advilin
BAM	Binary Alignment Map
BHLH	Basic Helix-Loop-Helix
BLIMP1	B lymphocyte-induced maturation protein-1
BMP	Bone morphogenetic protein
CCK	Cholecystokinin
CCL5	CC-chemokine ligand 5
CD45	Cluster of differentiation 45
Cd74	Cluster of differentiation 74
cDNA	Complementary DNA
Cdx1	Caudal Type Homeobox 1
CDX2	Caudal Type Homeobox 2
CEA	Carcinoembryonic Antigen-Related
Ceacam18	CEA cell adhesion molecule 18
Cebpe	CCAAT Enhancer Binding Protein Epsilon
Cel	Carboxyl ester lipase
Cela3a	Chymotrypsin-like elastase family member 3A
Cela3b	Chymotrypsin-like elastase family member 3B
CFU	Colony forming units
c-MAF	cellular musculoaponeurotic fibrosarcoma oncogene homolog
Cnbp2	Cellular nucleic acid-binding protein 2
Col6a1	Collagen type VI alpha 1 chain
Cpa1	Carboxypeptidase A1
Cpa2	Carboxypeptidase A2
Cpb1	Carboxypeptidase B1
CSN1S1	Casein alpha S1
CTSL	Cathepsin L1
CYP2B10	cytochrome P450, family 2, subfamily b, polypeptide 10
Da	Dalton
Dab2	Disabled homolog 2
DAPI	4’,6-Diamidin-2-phenylindol
DDN	downregulated-downregulated-non-regulated
Defa	Alpha-defensin
DEFA10	Alpha-defensin 10

Defa17	Alpha-defensin 17
DEFA23	Alpha-defensin 23
Defa26	Alpha-defensin 26
Defa29	Alpha-defensin 29
Defa30	Alpha-defensin 30
Defa38	Alpha-defensin 38
DEFA5	Alpha-defensin 5
DEFA6	Alpha-defensin 6
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
DNN	Downregulated-non-regulated-non-regulated
DNU	Downregulated-non-regulated-upregulated
DPEP1	Dipeptidase 1
DTT	Dithiothreitol
EC	Enterochromaffin
ECM	Extracellular matrix
Edn3	Endothelin-3
EDTA	Ethylenediaminetetraacetic acid
EEC	Enteroendocrine cells
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EpCAM	Epithelial cell adhesion molecule
EPEC	Enteropathogenic *Escherichia coli*
Ephb2	Ephrin type-B receptor 2
Esr2	Estrogen receptor beta
FCS	Fetal calf serum
FDR	False discovery rate
FELASA	Federation of European Laboratory Animal Science Associations
FITC	Fluorescein isothiocyanate
Folr1	Folate receptor 1
Ftx	Five prime to Xist
FWD	Forward
GALT	Gut-associated lymphoid tissue
GAP	Goblet cell-associated passage
GATA4	Transcription factor GATA-4
GATM	Glycine amidinotransferase
Gbp2	Guanylate binding protein 2
GEO	Gene Expression Omnibus
GHRL	Ghrelin
GIP	Glucose-dependent insulinotropic polypeptide
GIT	Gastrointestinal tract
GLP-1	Glucagon-like peptide 1
Gpx3	Glutathione peroxidase
GRCm38	Genome Reference Consortium Mouse Build 38
Grem1	Gremlin 1
H19	H19 imprinted maternally expressed transcript
HCD	Higher-energy collisional dissociation
HD5	Human alpha defensin 5
HD6	Human alpha defensin 6
Hes1	Hairy and enhancer of split-1
HF	High field

HGF	Hepatocyte growth factor
HMGA1	High mobility group AT-hook 1
Hmga2	High mobility group AT-hook 2
HMO	Human milk oligosaccharides
Hmox1	Heme oxygenase 1
Hnf1a	Hepatocyte nuclear factor 1 homeobox A
HNF4	Hepatocyte nuclear factor 4
HNF4A	Hepatocyte nuclear factor 4 alpha
HSD	Honestly significant difference
Ido1	Indoleamine 2,3-dioxygenase 1
IEC	Intestinal epithelial cell
IGF	Insulin-like growth factor
Igha	Immunoglobulin heavy constant alpha
Igkc	Immunoglobulin kappa constant
IRAK1	Interleukin 1 receptor associated kinase 1
Itln1	Intelectin 1
IZKF	Interdisziplinäres Zentrum für Klinische Forschung
JAK	Janus kinase
Jpx	Just proximal to Xist
Lama1	Laminin subunit alpha-1
Lamc1	Laminin subunit gamma-1
LAVE	Landesamt für Verbraucherschutz und Ernährung
LC	Liquid chromatography
LCM	Laser capture microdissection
Lct	Lactase
LFQ	Label-free quantification
Lgr5	Leucine-rich repeat-containing G-protein coupled receptor 5
LPS	Lipopolysaccharide
LRE	Lysosome-rich enterocytes
LRP2	Low density lipoprotein receptor-related protein 2 / Megalin
LRT	Likelihood ratio test
Ly6f	Lymphocyte antigen 6 complex
LYZ	Lysozyme
Lyz1	Lysozyme 1
MAFB	V-maf avian musculoaponeurotic fibrosarcoma oncogene homolog B
MAMDC4	MAM Domain Containing Protein 4/ Endotubin
MAPK	Mitogen-activated protein kinase
Mbl2	Mannose-binding lectin 2
Mchr1	Melanin-concentrating hormone receptor 1
Meltf	Melanotransferrin
MFGE8	Milk fat globule EGF factor 8 domain containing protein
MGAM	Maltase glucoamylase
MLN	Motilin
mRNA	Messenger RNA
MS	Mass spectrometry
Msx2	Msh homeobox 2
Muc2	Mucin 2
NAGA	Alpha-N-acetylgalactosaminidase
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
NEC	Necrotizing enterocolitis

NEUROD1	Neurogenic differentiation 1
Neurog3	Neurogenin 3
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NGS	Next-generation sequencing
NND	Non-regulated-non-regulated-downregulated
NNU	Non-regulated-non-regulated-upregulated
Notum	Notum, palmitoleoyl-protein carboxylesterase
NTS	Neurotensin
Olfm4	Olfactomedin 4
p53	Tumor protein p53
PBS	Phosphate-buffered saline
PC1	Principal component 1
PC2	Principal component 2
PCA	Principal component analysis
PEN	Polyethylene naphthalate
PFA	Polyformaldehyde
PI3K	Phosphoinositide 3-kinase
Pkp1	Plakophilin 1
Pla2g2a	Phospholipase A2
Pnliprp1	Pancreatic lipase-related protein 1
POU2F3	POU domain, class 2, transcription factor 3
Prdm1	PR domain zinc finger protein 1
Prss1	Serine protease 1
Prss2	Serine protease 2
PYY	Peptide YY
Reg	Regenerating islet-derived
Reg3a	Regenerating islet-derived 3 alpha
Reg3b	Regenerating islet-derived 3 beta
Reg3g	Regenerating islet-derived 3 gamma
RELM-β	Resistin-like molecule-beta
Retnlb	Resistin-like beta
RNA	Ribonucleic acid
RPMI	Roswell Park Memorial Institute (medium)
RWTH	Rheinisch-Westfälische Technische Hochschule
S100a8	S100 calcium-binding protein A8
S100a9	S100 calcium-binding protein A9
SAM	Sterile alpha motif
SCT	Secretin
SDS	Sodium dodecyl sulfate
seq	Sequencing
Sgk2	Serum/glucocorticoid regulated kinase 2
Sis	Sucrase-isomaltase
Smad3	SMAD family member 3
SMAD4	SMAD family member 4
Sox13	SRY-box transcription factor 13
Sox9	SRY-box transcription factor 9
SPDEF	SAM Pointed Domain ETS Factor
SPF	Specific pathogen-free
Spib	SPI-B transcription factor
Spink4	Serine peptidase inhibitor kazal type 4
SRY	Sex-determining Region Y

SST	Somatostatin
STAT	Signal Transducers and Activators of Transcription
Susd2	Sushi domain containing 2
Tcf4	Transcription factor 4
Tcf7l2	Transcription factor 7-like 2
TF	Transcription factor
Tff1	Trefoil factor 1
TGF-β	Transforming growth factor-β
TLR	Toll-like receptor
TLR3	Toll-like receptor 3
TLR4	Toll-like receptor 4
TLR5	Toll-like receptor 5
TNFα	Tumor necrosis factor (alpha)
TREH	Trehalase
Tsix	Antisense to Xist
UMAP	Uniform manifold approximation and projection
UMI	Unique molecular identifier
UND	Upregulated-non-regulated-downregulated
UNN	Upregulated-non-regulated-non-regulated
UNU	Upregulated-non-regulated-upregulated
Upb1	Ureidopropionase beta
UV	Ultraviolet
VEGF	Vascular endothelial growth factor
VIPER	Virtual Inference of Protein-activity by Enriched Regulon analysis
VST	Variance-stabilizing transformation
WAP	Whey acid protein
WNT/Wnt	Wingless/Int-1
Xist	X-inactive specific transcript
ZG16	Zymogen granule protein 16
ZO1	Zonula occludens-1

5. Table of contributors

Contribution / Task	Contributor(s)	Affiliation	Category
All in vivo experimental animal work (infection, tissue collection)	Aline Dupont	AG Hornef, RWTH Aachen	Internal collaboration
Intestinal epithelial cell isolation for bulk RNA-seq and proteomics	Aline Dupont	AG Hornef, RWTH Aachen	Internal collaboration
RNA isolation for bulk RNA-seq	Aline Dupont	AG Hornef, RWTH Aachen	Internal collaboration
cDNA library preparation for bulk RNA-seq	Jasmin Hübner, Anna Rudzinski	IZKF Genomics Facility, RWTH Aachen	Service / Core facility
Bulk RNA-seq sequencing (NextSeq 500 / NovaSeq 6000)	Genomics Facility, IZKF	RWTH Aachen University	Service / Core facility
Crypt and villus tissue LCM, RNA isolation, library preparation for 3' RNA-seq	Isabel Richter, Matthias Schmitz	AG Hornef, RWTH Aachen	Internal collaboration (MD thesis, bachelor's thesis)
Bulk RNA-seq sequencing of LCM samples, pre-processing of files	Fabian Imdahl, Christophe Toussaint	AG Saliba, Helmholtz-Institut für RNA-basierte Infektionsforschung (HIRI)	External collaboration
Adapter trimming and quality filtering of RNA-seq reads	Keshav Prasad Gubbi	AG Hornef, RWTH Aachen	Internal collaboration
Single-cell RNA-seq sample preparation and cell sorting	Stefan Schlößer, Aline Dupont	AG Hornef, RWTH Aachen	Internal collaboration
scRNA-seq library preparation	Stefan Schlößer	AG Hornef, RWTH Aachen	Internal collaboration
scRNA-seq sequencing, FASTQ demultiplexing and genome alignment (Cellranger)	Genomics Facility, IZKF	RWTH Aachen University	Service / Core facility
Mass spectrometry (LC-MS/MS) analysis	Sofia Jäverfelt	AG Pelaseyed, Gothenburg University	External collaboration
MaxQuant preprocessing of MS raw files	Thaher Pelaseyed	AG Pelaseyed, Gothenburg University	External collaboration

6. Introduction

6.1. Structure and main functions of the small intestine

The main function of the intestinal epithelium is to absorb nutrients and water while acting as a barrier against microbial intrusion. Anatomically, the intestinal epithelium consists of two major structures: villi, finger-like protrusions that maximize the surface area for nutrient absorption, and crypts, also called crypts of Lieberkühn, gland-like invaginations between the villi, where the pool of pluripotent stem cells lies. Each villus is functionally linked to at least six crypts (Barker, 2014).

The small intestine makes up 80% of the gastrointestinal tract (GIT) of mice and is regionally segmented into the relatively short duodenum, the longer jejunum, and the even shorter ileum. The regional segmentation of the small intestine, originating in ancient Greek anatomical classifications, reflects functional specialization in nutrient absorption. The duodenum, the first segment, receives partly digested food from the stomach and is the initial site of chemical digestion, using enzymes and bile acids from the pancreas, liver, and gallbladder. The jejunum, the second segment of the small intestine, is primarily responsible for the absorption of micronutrients such as zinc and folate (vitamin B9). Most macronutrient and water absorption occurs across the duodenum and ileum, where 95% of carbohydrates and proteins and about 90% of water are absorbed. The ileum, the final and longest segment, serves as the principal site for prolonged nutrient exposure and absorption before the contents are passed into the large intestine (Fish et al., 2024).

Although the classical segmentation of the small intestine implies distinct anatomical boundaries, these transitions are gradual and continuous. Along the proximal-to-distal axis, the villus length decreases and crypt depth increases, reflecting a morphological gradient rather than sharply defined regions (Beumer & Clevers, 2021). Recent studies further support this view: at least five distinct molecular domains have been found using single-cell sequencing along the mouse small intestine. This suggests overlapping transcriptional and functional gradients along the mouse small intestine and further highlights the idea of a continuous zonation rather than sharply defined segments (Zwick et al., 2024).

The epithelium is supported by a basal membrane that borders the *lamina propria*, a thin layer of loose connective tissue. The *lamina propria* provides the structural framework of the intestinal mucosa and is populated by mesenchymal cells such as fibroblasts, myofibroblasts, endothelial cells, and smooth muscle cells, as well as immune cells such as lymphocytes, plasma cells, and macrophages (Powell et al., 2011). In addition to its structural role, the *lamina propria* is a key component of the gut-associated lymphoid tissue (GALT) and serves as an important site for the mucosal immune response, containing scattered immune cells and isolated lymphoid follicles. It is also closely associated with Peyer's patches,

organized lymphoid aggregates located in the submucosa, which act as inductive sites for adaptive immune responses to luminal antigens (Torow et al., 2023). Luminal antigens originate from both dietary components and resident microbiota, providing a continuous and diverse source of stimuli to the intestinal epithelium and immune system. At the luminal interface, the epithelium is protected by a continuously renewed mucus layer, whose secretion is stimulated by neuroimmune signals such as acetylcholine and cytokines (Raya Tonetti et al., 2024). This layer forms a protective physicochemical barrier that separates the heavily colonized gut lumen from the sterile underlying mucosal tissue (Zhang et al., 2015).

The luminal microbial density increases throughout the gastrointestinal tract, ranging from 10^2 bacteria per gram of luminal material in the proximal small intestine to 10^8 in the distal ileum and 10^{12} in the colon (Santaolalla et al., 2011). Most bacteria in the small intestine are commensal or transient and are generally beneficial to the host. However, some foodborne pathogens like *Salmonella* can survive the gastric acid and establish infection when buffered by food (Smith, 2003). The small intestine must therefore maintain a balance between immune tolerance to commensal microbes and defense against foodborne pathogens. To support this dual function and its other physiological roles, a variety of specialized epithelial cell types exist.

6.2. Intestinal epithelial cells (IECs)

6.2.1. Stem and transit amplifying cells

Intestinal epithelial stem cells are characterized by the expression of leucine-rich repeat-containing G-protein coupled receptor 5 (*Lgr5*) and olfactomedin 4 (*Olfm4*). About four to six stem cells lie at the bottom of every crypt; these self-renewing stem cells cycle steadily to produce the rapidly proliferating transit-amplifying (TA) cells, which are in turn capable of differentiating towards all lineages (Barker et al., 2007; Cheng & Leblond, 1974). Differentiated cells that emerge from the TA zone travel upward towards the villus tip. After a turnover time of about 3 days in mice and 5 days in humans, they reach the villus tip, undergo anoikis, a form of programmed cell death, and are exfoliated into the gut lumen for disposal (Cheng & Leblond, 1974; de Santa Barbara et al., 2003).

The Notch and Wnt signaling pathways are the two major signaling pathways driving epithelial cell differentiation by regulating lineage-specific transcription factors. Both pathways work as a gradient in the crypt, most prominently at the base, and decrease along the crypt-villus axis.

Canonical Wnt signaling, mediated in part by β-catenin stabilization, is essential for maintaining intestinal stem cells and driving proliferation of transit-amplifying progenitors, thereby supporting differentiation into epithelial lineages. In contrast, Notch activation induces the transcription factor hairy and enhancer of split-1 (*Hes1*), which, in turn, represses atonal bHLH transcription factor 1 (*Atoh1*), thereby blocking secretory lineage commitment and promoting differentiation towards

absorptive enterocytes, which constitute approximately 80% of the intestinal epithelial cell population. When Notch signaling is low or absent, *Atoh1* expression is no longer actively inhibited, promoting differentiation into the secretory lineages, including goblet cells (4-16% of all intestinal epithelial cells, depending on location in the small intestine), Paneth cells (2-4%), enteroendocrine cells (EECs, 1%), and tuft cells (0.5-2%) (Barker, 2014; Clevers, 2013; Mah et al., 2016; Noah & Shroyer, 2013).

6.2.2. Absorptive enterocytes

Absorptive enterocytes make up most of the intestinal epithelial cells and are responsible for nutrient absorption. Their differentiation from transit-amplifying progenitors is driven by high Notch signaling activity, which promotes the absorptive lineage at the expense of secretory cells. Their function differs based on their location as they migrate along the crypt-villus axis. Absorptive enterocytes located at the base of the villi express a set of antibacterial genes from the Reg family, regulated by signals from the gut microbiome. This microbiome-dependent gene program enables absorptive enterocytes to contribute to the intestinal epithelial antimicrobial defense by responding to microbial cues and maintaining barrier integrity. Absorptive enterocytes located at the mid-villus section play an important role in nutrient absorption and uptake by expressing carbohydrate, amino acid, and peptide transporters. In contrast, absorptive enterocytes located at the villus tip express mainly apolipoproteins (*Apoa1*, *Apoa4*, *Apob*) and genes involved in tight junction reorganization like the epidermal growth factor receptor (*Egfr*) (Moor et al., 2018). Absorptive enterocytes are also sub-specialized according to their position along the proximal-to-distal axis. Thus, proximal absorptive enterocytes have been shown to be specialized in the absorption of iron and macronutrients, whereas distal enterocytes absorb bile acids and vitamin B12 (Haber et al., 2017).

6.2.3. Goblet cells

Goblet cells are distributed throughout the small intestine, increasing in abundance towards the distal gut and extending from the middle of the crypt to the tip of the villus (de Santa Barbara et al., 2003). Goblet cells mainly produce and secrete Mucin 2 (*Muc2*), a large gel-forming mucin that is the primary structural component of the intestinal mucus layer. Based on their mucus-producing properties, two main subsets of goblet cells can be distinguished: crypt-type and villi-type. In the upper crypt, goblet cells feature large mucus reservoirs, giving them a granular, secretory appearance. In contrast, the crypt-villus junction and villi tip lack mucus-filled goblet cells and instead feature a distinct subset characterized by higher mucus production and secretion rate (Nyström et al., 2021; Schneider et al., 2018). Beyond their role in mucin production, goblet cells also play a role in immune regulation. They release antimicrobial peptides (e.g., resistin-like molecule-beta (RELM-β) or trefoil factor 1 (*Tff1*)), chemokines, and cytokines that recruit and modulate immune cells and deliver luminal antigens to underlying dendritic cells via the goblet cell-associated passages (GAP) (Raya Tonetti et al., 2024).

6.2.4. Paneth cells

First noted by Schwalbe in 1872 and more fully described by Paneth in 1887, Paneth cells are unique among intestinal epithelial lineages in that they migrate downward towards the crypt base rather than upward towards the lumen (Schwalbe, 1872; Paneth, 1887). They ultimately reside at the base of the small intestinal crypts, where they play essential roles in mucosal immunity and epithelial homeostasis. In contrast to the other intestinal epithelial cell types, Paneth cells also possess a longer turnover period. They can stay for 20 days up to two months at the crypt base and only are cleared by infiltrating macrophages once they have reached the end of their lifespan (Beumer & Clevers, 2021; de Santa Barbara et al., 2003). Low Notch activity, high Wnt/β-catenin signaling, and reduced MAPK pathway activation direct secretory progenitors towards the Paneth cell lineage, while shifts in this balance can redirect cell fate towards goblet or other secretory cells (Heuberger et al., 2014).

Morphologically, Paneth cells are characterized by a pyramidal shape, a prominent Golgi apparatus, and apically localized secretory granules. They secrete a variety of antimicrobial peptides (AMPs), which are critical for shaping the composition of the intestinal microbiota and for protecting the host stem cell niches from microbial invasion (Wallaeys et al., 2023).

AMPs comprise a diverse group of small effector molecules, usually less than 40 amino acids, with broad antimicrobial activity against Gram-negative and Gram-positive bacteria, fungi, and protozoa, acting primarily through membrane disruption or interference with microbial physiology (Ayabe et al., 2004; Ouellette & Bevins, 2001). Within this group, defensins represent a structurally defined subfamily characterized by conserved cysteine residues and disulfide bond formation and are overall the largest and most extensively studied family of AMPs in mammals (Duarte-Mata & Salinas-Carmona, 2023). Among the different defensin classes, α-defensins, cryptdins in mice, are preferentially expressed by Paneth cells and contribute substantially to antimicrobial defense in the small intestinal crypts. β-defensins, in contrast, display a broader epithelial expression pattern and are generally present at lower levels in the small intestine, with comparatively higher expression in the colon (Huttner et al., 1997).

While murine Paneth cells express a broad set of over 20 α-defensins, human Paneth cells produce primarily HD5 (encoded by DEFA5) and HD6 (encoded by DEFA6). HD5 displays potent bactericidal activity similar to murine cryptdins, whereas HD6 exhibits limited direct antimicrobial activity and instead forms higher-order peptide assemblies that entrap bacteria and restrict their access to the epithelial surface (Zhao et al., 2015).

Paneth cell numbers increase from the proximal to the distal small intestine, ranging from 5 to 16 cells per crypt depending on strain and detection method. This distribution correlates with the elevated bacterial load in distal regions and signifies their role in antimicrobial defense (Nakamura et al., 2020).

Lysozyme (LYZ in humans; *Lyz1* in mice), the first AMP identified in Paneth cells, remains a widely used Paneth cell marker, particularly in the ileum (Haber et al., 2017).

6.2.5. Enteroendocrine cells (EEC)

Enteroendocrine cells (EECs) are a rare epithelial cell type that plays a central role in the integration of environmental signals and systemic physiology. EECs are positioned along the crypt-villus axis and throughout the gastrointestinal tract, and despite constituting less than 1% of the intestinal epithelium, they collectively form the body's largest endocrine system. They are responsible for the secretion of over 20 different hormones that regulate digestion, appetite, glucose homeostasis, and gut motility. The specific hormone repertoire varies by EEC subtype, and EECs have therefore traditionally been classified using single- or two-letter systems based on the hormones they produce (Sanchez et al., 2022).

Enterochromaffin (EC) cells, named after the staining process with chrome salts, produce more than 90% of the body's serotonin (5-HT) and were the first EEC subtype to be discovered (Heidenhain, 1870). Later, additional types of EECs were classified by the size of their vesicles. They were classified as small (S cells, producing secretin, SCT), intermediate (I cells, producing cholecystokinin, CCK), large (L cells, producing glucagon-like peptide 1, GLP-1, and peptide YY, PYY), and large-like (K cells, producing glucose-dependent insulinotropic polypeptide, GIP) (Barton et al., 2023; Fothergill & Furness, 2018). Other groups have been defined by their morphological resemblance to pancreatic cells (A cells, producing ghrelin, GHRL, and D cells, producing somatostatin, SST) or by their immunoreactivity to hormone antibodies, such as M cells (motilin, MLN), N cells (neurotensin, NTS), S cells (secretin, SCT), and X cells (also ghrelin) (Barton et al., 2023). EEC subtypes secreting serotonin and ghrelin are usually found more proximally, whereas distally located EECs mostly secrete GLP-1 (Haber et al., 2017). These classifications, grounded in early immunohistochemical and physiological studies, formed the foundation for understanding EEC diversity (Gribble & Reimann, 2016). However, recent advances, using single-cell RNA-seq, have challenged traditional classification. At least 12 transcriptionally distinct EEC clusters have been identified in the mouse small intestine, with many cells co-expressing multiple hormones, thus challenging the "one cell, one hormone" paradigm (Beumer et al., 2020; Haber et al., 2017; Habib et al., 2012).

6.2.6. Tuft cells

Tuft cells were first described in 1956 in the mouse stomach (Järvi & Keyriläinen, 1956) and rat trachea (Rhodin & Dalhamn, 1956) using electron microscopy. They were named after their apical "tuft" of long, actin-rich microvilli. In the small intestine, they arise from Lgr5+ intestinal stem cells through an Atoh1-independent but POU domain, class 2, transcription factor 3 (POU2F3)-dependent differentiation pathway. They are among the rarest intestinal epithelial cell types, sparsely distributed

along the crypt-villus axis, with slightly higher abundance in the proximal duodenum, where they can comprise more than 2% of the epithelial population (Akimori et al., 2011).

Tuft cells function as chemosensory sentinels, detecting luminal signals through taste receptor pathways and contributing to mucosal immunity by secreting interleukin-25 (IL-25), which in turn activates type 2 innate lymphoid cells. They also support the epithelial barrier by influencing mucus secretion and interacting with immune and microbial signals. (Bas et al., 2023). In addition to their intestinal roles, tuft cells are abundant in the gastric groove, where they protect the squamous epithelium of the forestomach and esophagus by secreting an alkaline solution (sodium bicarbonate, $NaHCO_3$) to neutralize stomach acid. In a similar way, the relatively high proportion of tuft cells in the proximal duodenum may help protect the mucosa from residual stomach acid entering from the stomach, as well as from organic acids produced by gut bacteria (Akimori et al., 2011).

The actin-binding protein villin was one of the first markers used to identify tuft cells by immunohistochemistry. However, because villin is also highly expressed by absorptive enterocytes, its utility as a tuft cell–specific genetic marker is limited. In contrast, the related advillin (*Avil*) has been reported to serve as a more specific marker for tuft cells in the murine GIT (Ruppert et al., 2020).

6.3. Regulators of spatial identity in the intestinal epithelium

A key feature of the intestinal stem cell niche is the maintenance of low bone morphogenetic protein (BMP) signaling, which supports stem cell proliferation and self-renewal. This gradient is maintained by local inhibitors such as Gremlin 1 (*Grem1*), which suppress BMP activity at the crypt base to support stem cell proliferation and self-renewal. BMP signaling increases along the crypt–villus axis, reaching higher levels toward the villus. Increased BMP activity restricts proliferation and promotes the differentiation of epithelial cells into the major functional subtypes as they migrate upwards (Beumer et al., 2022). However, BMP signals do not drive terminal absorptive enterocyte differentiation; instead, they stimulate a switch between two states of differentiation from villus bottom to villus top. High BMP levels at the villus tip support expression of nutrient transporters and lipid uptake in absorptive enterocytes, whereas bottom villus absorptive enterocytes show higher expression of antimicrobial genes from the *Reg* family and increased protein translation. Similarly, goblet cells also show a BMP-induced zonation, with high Intelectin-1 (*Itln1*) and Serine peptidase inhibitor kazal type 4 (*Spink4*) expression in the crypts and at the bottom of the villi and high expression levels of zymogen granule protein 16 (*Zg16*) at the villus tip, which contributes to mucosal defense by promoting the aggregation of Gram-positive bacteria (Bergström et al., 2016; Beumer et al., 2022). In addition, BMP signaling has emerged as a key regulator of hormonal plasticity in enteroendocrine cells (EECs), where it can shift the spatial expression of EEC hormones, extending crypt-restricted hormone profiles into the villus region while repressing villus-enriched hormones (Beumer et al., 2018, 2020). BMP signaling

exhibits a proximal-to-distal gradient along the small intestine, with higher activity in the ileum. Increased BMP signaling antagonizes Wnt-driven epithelial proliferation and promotes epithelial cell apoptosis, in part through suppression of integrin-mediated survival signals. Wnt signaling exerts feedback control over BMP signaling, whereas BMP signaling inhibits Wnt activity, together forming a reciprocal regulatory loop (Figure 1). Together, these effects reduce epithelial cell turnover and limit villus growth, contributing to the progressive shortening of villi toward the distal small intestine (Wang et al., 2025a).

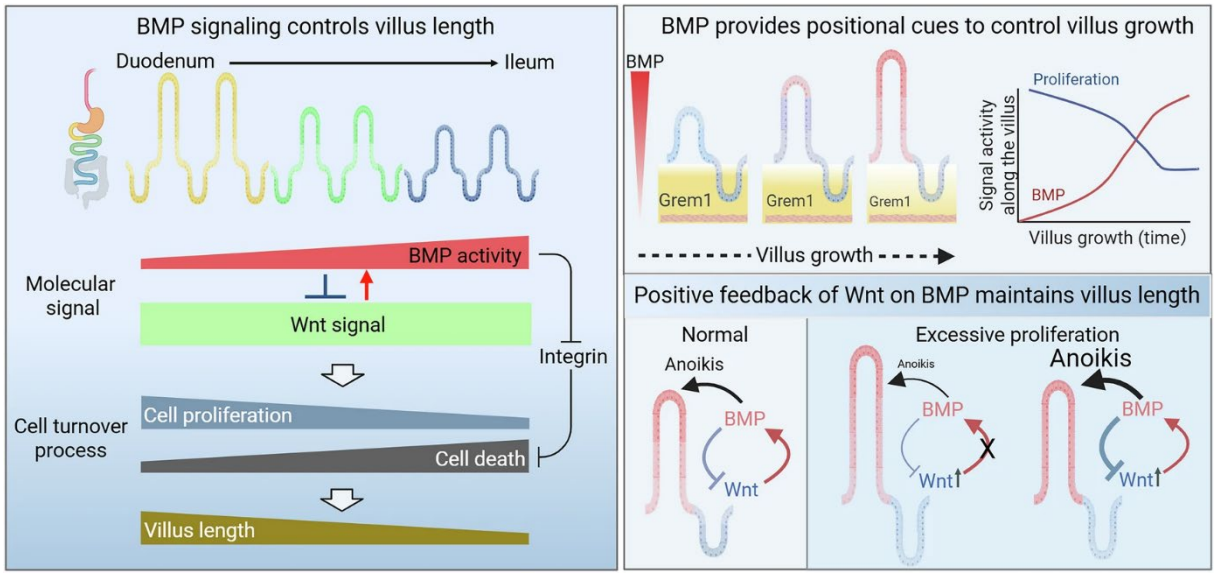


Figure 1: The interaction between BMP and Wnt signaling establishes the different epithelial turnover rates and villus length in the proximal and distal small intestine. Schematic representation of the increase of BMP signaling activity from proximal to distal small intestine, while Wnt signaling activity remains constant. For cell turnover BMP inhibits cell proliferation and promotes cell death (apoptosis) by inhibiting integrin expression. The BMP gradient corresponds with the steady decline in villus length from the proximal to distal small intestine. Along the crypt-villus axis, the increase in BMP signaling antagonizes Wnt activity, leading to a progressive reduction in epithelial proliferation. Reproduced unchanged from Wang et al., 2025, Nature Communications, under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

In parallel to BMP signaling pathways, transcription factors play a leading role in establishing and maintaining spatial identity within the intestinal epithelium. Among these, caudal-related homeobox protein 2 (CDX2) is a known regulator of intestinal identity. Functionally, CDX2 is essential for establishing a midgut identity during development, suppressing foregut programs, and driving differentiation of adult epithelial lineages. CDX2 functions downstream of Wnt and can repress Wnt target genes in differentiated cells, thereby reinforcing epithelial maturation (Guo et al., 2004). CDX2 activity has also been associated with the promotion of epithelial cell maturation and crypt morphogenesis, including the induction of premature crypt formation. Disruptions in crypt organization, such as changes in Paneth cell differentiation and the migration of TA cells to the crypt base, have been linked to altered CDX2 function. These effects reflect CDX2-dependent modulation of Wnt/β-catenin signaling and interactions with other transcriptional regulators. Through these mechanisms, CDX2 contributes to the spatial organization of the crypt-villus axis and plays a critical role in coordinating epithelial identity, crypt architecture, and signaling dynamics (Crissey et al., 2011).

GATA4 is another key transcription factor that acts along the proximal-to-distal axis of the small intestine. It is expressed primarily in proximal and medial segments, where it represses distal ("ileal") gene expression through direct DNA binding and chromatin modulation and thereby supports regional differentiation. Specifically, GATA4 suppresses gene programs by limiting histone acetylation at distal gene loci, thereby restricting their activation in jejunal regions and supporting region-specific differentiation (Aronson et al., 2014; Bosse et al., 2006).

Aside from spatial identity, various transcription factors exist that promote cell-type-specific differentiation. Notable examples include the Wnt target genes Achaete-scute complex homolog 2 (*Ascl2*), which is a master regulator of intestinal stem cell identity (Schuijers et al., 2015), as well as SRY-box transcription factor 9 (*Sox9*) and transcription factor 4 (*Tcf4,* also known as *Tcf7l2* in mice), both of which are necessary for Paneth cell development (Bastide et al., 2007; Van Es et al., 2012). The SAM Pointed Domain ETS Factor (SPDEF) promotes goblet cell terminal differentiation by driving mucin and goblet gene programs and converting crypt progenitors into goblet cells (Noah et al., 2010). Neurogenin 3 (*Neurog3*) expression reflects the first step of enteroendocrine development, and without it, enteroendocrine cells do not develop, whereas the subsequent expression of the transcription factor neuronal differentiation 1 (NEUROD1) appears to signify a later stage of differentiation for mature enteroendocrine cells (Li et al., 2011). Hepatocyte nuclear factor 4 alpha (HNF4A) stabilizes absorptive enterocyte identity and metabolic specialization by acting within feed-forward transcriptional circuits, in which HNF4A and signaling-responsive factors such as SMAD4 co-regulate overlapping target genes, reinforcing enterocyte gene expression and functional maturation (Chen et al., 2019).

6.4. Structural and functional differences between neonatal and adult intestines

6.4.1. Key developmental and structural features

The neonatal small intestine differs greatly from the adult gut in terms of form and function, indicating its role in the perinatal transition as it adjusts to microbial colonization and enteral feeding.

In humans, its intestinal architecture and cellular composition are structurally mature starting with week 19 of gestation. Thus, at birth, the human small intestinal epithelium already has a well-established crypt-villus architecture, with clearly defined crypts containing stem and TA cells, as well as completely differentiated epithelial lines. Nonetheless, the intestinal immune system is immature at birth because antigenic stimulation of the colonizing microbiota is still required for its complete maturation (Tourneur & Chassin, 2013).

In mice, however, due to the short gestational period of approximately 20 days, the neonatal small intestine is structurally immature at birth. The epithelium displays a relatively smooth surface with shallow intervillus regions rather than well-defined crypt invaginations. At this stage, only absorptive enterocytes, goblet cells, enteroendocrine cells (EECs), and transit-amplifying (TA) cells are present, whereas Paneth cells are absent. Paneth cells emerge only postnatally during crypt formation and maturation of the small intestinal epithelium (de Santa Barbara et al., 2003).

Following birth, epithelial cell proliferation in the murine small intestine remains low during the suckling period, which is adapted to a milk-based diet. Proliferative activity increases around days 10-12 after birth, coinciding with crypt development and the adaptation of the gut epithelium to more complex nutritional components (Dutta & Sengupta, 2016; Ménard et al., 2008). In fact, intestinal crypts only begin to form around day 7 and become morphologically distinct by day 14. Before crypts form, epithelial proliferation takes place in the intervillus region rather than in a designated stem cell zone. As crypts mature, proliferative activity becomes spatially restricted to the crypt base, corresponding with the appearance of Paneth cells and the formation of a permanent stem cell niche. Before Paneth cells fully develop in neonatal mice, goblet cells can produce α-defensins and partially compensate for the antimicrobial function of Paneth cells (Bry et al., 1994). However, full crypt maturation and the establishment of a complete crypt-villus axis, including stem cell localization and Paneth cell differentiation, are typically achieved around the weaning period at day 21 (Dehmer et al., 2011).

In the neonatal intestine, barrier function is still immature, resulting in increased epithelial permeability during the early postnatal period prior to what has been called “gut closure”. This may reflect incomplete tight junction assembly, which normally limits the paracellular passage of ions, microbial components, and dietary antigens (Weström et al., 2020). At the same time, the neonatal epithelium exhibits heightened expression of Toll-like receptors (TLRs) TLR4 and TLR5. TLR4, which detects bacterial lipopolysaccharide (LPS), is broadly expressed across neonatal enterocytes, but its signaling is rapidly reduced after birth due to loss of the signaling molecule interleukin 1 receptor associated kinase 1 (IRAK1), reducing epithelial LPS sensitivity under homeostatic conditions (Chassin et al., 2010; Lotz et al., 2006). TLR5, the flagellin receptor, follows a similar trajectory, being initially widespread but later restricted to Paneth cells in the adult intestine, where it supports the detection of microbial patterns in the crypt environment (Price et al., 2018). In contrast, TLR3, which recognizes viral double-stranded RNA, is expressed at low levels in neonates but increases postnatally in the villus epithelium, correlating with the development of antimicrobial defenses (Pott et al., 2012).

Given the heightened microbial sensing, the immature epithelial barrier, and the postnatal exposure to microbial ligands released by the establishing microbiota, nuclear factor-κB (NF-κB) signaling is physiologically stimulated by postnatal microbial exposure in the neonatal gut to protect epithelial

cells from apoptosis induced by inflammatory cytokines, reactive oxygen species, and microbial stimuli. This controlled NF-κB stimulation supports epithelial cell survival and barrier integrity and prevents excessive inflammation during the critical early-life window in which the mucosal immune system is still adapting to the colonizing microbiota (Chassin et al., 2010).

6.4.2. Fetal-type enterocytes and the immature intestinal epithelium

The immature intestinal epithelium contains vacuolated fetal-type enterocytes, which possess an apical canalicular system (ACS) that forms large supranuclear vacuoles (Skrzypek et al., 2007). These cells are conserved across species, including other mammals and zebrafish, and are in some studies also referred to as lysosome-rich enterocytes, reflecting their prominent endo-lysosomal compartment (Block et al., 2025; Moxey & Trier, 1979; Park et al., 2019; Skrzypek et al., 2007). These vacuoles are present along the entire small intestine but are most abundant in the distal region, reflecting regional specialization for macromolecule uptake. This specialized system enables pinocytosis and transcytosis of maternal immunoglobulins and other milk macromolecules, allowing their uptake and transport across the epithelium while preserving biological activity. This system of fetal-type enterocytes and the ACS contributes to the increased intestinal permeability during the neonatal period and gradually decreases as cells containing an ACS are replaced with adult enterocytes. This proximal-to-distal maturation wave begins in the proximal small intestine, then gradually declines, but persists in the distal small intestine until weaning, coinciding with structural and functional remodeling of the epithelium (Muncan et al., 2011; Skrzypek et al., 2007).

Markers of the immature epithelium include the neonatal Fc (fragment, crystallizable) receptor (*FcRn*), which facilitates immunoglobulin G (IgG) uptake from milk, and argininosuccinate synthase 1 (*Ass1*), which is essential for arginine biosynthesis (Hurwitz & Kretchmer, 1986; Israel et al., 1997). During the suckling-to-weaning transition, the fetal-type vacuolated enterocytes disappear, the epithelium tightens, and intestinal permeability becomes more selective (Muncan et al., 2011).

These structural and molecular changes set the stage for the transcriptional programs, including B-lymphocyte-induced maturation protein 1 (BLIMP1) and the transcription factor musculoaponeurotic fibrosarcoma oncogene homolog B (MAFB), that orchestrate the neonate-to-adult intestinal transition.

6.4.3. BLIMP1 and MAFB as regulators of the neonate-to-adult transition

In the mouse, postnatal epithelial maturation is controlled by the transcriptional repressor BLIMP1 (encoded by the gene *Prdm1*). BLIMP1 is widely expressed throughout the embryonic intestinal epithelium. However, following birth, BLIMP1 expression is progressively lost from newly generated enterocytes. As enterocytes migrate upwards along the villi, BLIMP1 expression first becomes excluded from the intervillus domain and emerging crypts, and by the second and third week of life, is only detected at the villus tips (Chin et al., 2017; Muncan et al., 2011). In the intestinal epithelium, BLIMP1

suppresses the premature expression of genes involved in adult metabolic processes, such as those related to lipid and bile acid metabolism, as well as enzymes adapted for solid food digestion. Intestinal epithelial-specific deletion of BLIMP1 in mice (using Blimp1flox/flox VillinCre+ animals) leads to the premature differentiation of crypts and Paneth cells, an enhanced goblet cell development, and a pronounced metabolic transition towards the adult epithelial phenotype in neonatal mice at day 7 after birth. This occurs at the cost of enzymes typically expressed and crucial during the early pre-weaning period, such as lactase (*Lct*) (Harper et al., 2011; Muncan et al., 2011).

The transcription factors MAFB and c-MAF are essential for maintaining neonatal-specific enterocyte function. They are expressed early during enterocyte differentiation and regulate a gene program involved in endo-lysosomal processing. This pathway enables the uptake and intracellular digestion of intact milk-derived macromolecules, a nutrient acquisition mode important for suckling mice. This functional defect resembles aspects of the BLIMP1-deficient phenotype, suggesting that MAF proteins and BLIMP1 both support neonatal epithelial identity and nutrient uptake. However, while BLIMP1 knockout mice show a global, premature neonatal-to-adult transition, MAFB/c-MAF double knockout mice display a selective impairment in neonatal enterocyte function without a complete adult shift. The MAF factors act more in a lineage-restricted and functional capacity rather than as broad repressors of maturation like BLIMP1 (Bara et al., 2022).

6.4.4. Breast milk as a source of nutrients

Whereas adult mice obtain nutrients from a solid food diet with mainly complex carbohydrates, the growth of infant mice, particularly in the early postnatal period, depends on access to maternal breast milk. Newborn pups begin nursing almost immediately after birth, and within hours, a characteristic white “milk spot” becomes visible through their translucent abdominal skin. Five days after birth, hearing begins to develop, the eyes open around day 12, and pups gradually start exploring beyond the nest. Around day 16, they begin to nibble on solid food if it is accessible near the dam, and approximately by day 19, they consume mainly solid food and become independent of the dam (Williams & Scott, 1954). Around this time, neonatal markers decrease in expression, while the brush-border enzymes sucrase-isomaltase (*Sis*) and arginase 2 (*Arg2*) start to be expressed to help digest solid food (De Jonge et al., 1998; Krasinski et al., 1994).

During early postnatal development, breast milk serves not only as a source of nutrition but also as a key regulator of epithelial development, innate immune response, and microbiota establishment. Breast milk is a rich source of secretory IgA (sIgA), which is also released on the mucosal surfaces and coats a portion of the gut microbiota, preventing pathogen adhesion and excessive immune activation. This passive immunity supports mucosal homeostasis during a period when the neonatal immune system is immature (Rogier et al., 2014). Similarly, breast milk in the first week after birth is particularly

rich in IgG, and the application of small amounts of IgG to newborn mice during the first week has been shown to prevent over-activation of gut immune cells after weaning (Shenoy et al., 2025).

In addition, breast milk contains various bioactive molecules, such as growth factors, cytokines, and early AMPs, that shape the early immune response. These growth factors, such as epidermal growth factor (EGF), hepatocyte growth factor (HGF), insulin-like growth factor (IGF), and transforming growth factor-β (TGF-β), regulate cell proliferation and stimulate cell differentiation in the crypts. Their expression in the small intestine increases until the end of the second week of life and then drastically decreases in the third week of life before reaching a plateau around weaning (Zhang et al., 2016).

Breast milk also contains a high concentration of exosomes, extracellular nanovesicles that mediate cell-cell communication via the transfer of messenger RNA (mRNA), beneficial immune-related microRNAs (miRNAs), and proteins. These molecules are protected from degradation in the digestive tract and are taken up by endocytosis, contributing to immune regulation in the neonatal gut (Lee et al., 2012). Exosomes have also been shown to reduce inflammation and injury of intestinal epithelial cells (IECs) by protecting the tight-junction protein-based epithelial barrier with claudins, occludins, and zonula occludens-1 (ZO1) from inflammatory damage (He et al., 2021).

Similarly, milk-associated commensal bacteria, such as *Bifidobacteria*, prevalent in humans due to their ability to metabolize human milk oligosaccharides (HMOs), support barrier maturation and exert strong immunomodulatory effects. A similar role is fulfilled by *Lactobacilli* in mice (Van Best et al., 2020). These bacteria induce host mechanisms and pathways that strengthen tight junctions, reduce epithelial permeability, and suppress pro-inflammatory cytokines, consistent with their established probiotic roles in immune tolerance and epithelial development (Le Doare et al., 2018; Lyons et al., 2020).

6.4.5. How early life microbiota colonization shapes the immune system

The development of the intestinal microbiota begins at birth, when the previously sterile gut is exposed to maternal and environmental microbes. In fact, current evidence does not support the presence of a consistent prenatal microbiota in murine fetal or placental tissues, making birth the primary microbial colonization event (De Goffau et al., 2021; Theis et al., 2020). During the early postnatal period, microbial diversity remains low, and community composition is shaped by maternal transmission through fecal contact, skin-to-skin licking, and breastfeeding. Neonates have increased epithelial permeability, which allows better macromolecule movement throughout the gastrointestinal tract. This heightened permeability is a consequence of both the immature barrier architecture and the increased endocytic capacity of fetal-type enterocytes. Through fluid-phase or receptor-mediated endocytosis, these fetal-type enterocytes import luminal material containing

macromolecules for either intracellular digestion in digestive vacuoles or vesicular transport through the cell and release on the basolateral side (transcytosis) (Weström et al., 2020).

This early-life period represents a critical neonatal "window of opportunity" in which host–microbe interactions shape long-term immune function, tolerance, and overall gut health (Torow et al., 2023). Disruption of microbial colonization during this neonatal window has been shown to induce durable alterations in immune programming that influence susceptibility to inflammatory and infectious diseases later in life (Borbet et al., 2023).

Previous studies have shown that conventionally raised, specific pathogen-free (SPF) adult mice have greater numbers of Paneth cells compared to germ-free mice, despite their overall epithelial composition being mostly the same. This increase has been interpreted as an adaptation to microbial colonization, as the host must rapidly establish epithelial defense mechanisms like antimicrobial peptide production, followed by the formation of a biochemical barrier that limits microbial access to the crypts and protects stem cells (Schoenborn et al., 2019; Tsang et al., 2022). Nonetheless, this interpretation remains inconclusive, as alternative functional studies have demonstrated that germ-free and colonized mice produce comparable numbers of enteric defensin products, suggesting that Paneth cell number and antimicrobial output may be uncoupled from microbial colonization (Pütsep et al., 2000).

While microbial colonization in early life is essential for shaping immune development and maintaining intestinal homeostasis, the unique physiological features of the neonatal gut also influence its interaction with potential pathogens. Epithelial immaturity with limited antimicrobial defense and a developing immune system can affect susceptibility and response to gastrointestinal infections induced by enteric pathogens. In addition, the increased epithelial permeability in neonates also raises the possibility of potentially harmful bacteria crossing the gut barrier.

6.5. *Salmonella* as a major causative agent of gastrointestinal infections

Infectious diseases are the leading cause of mortality in children under the age of five in developing countries worldwide. Infections in the gastrointestinal tract account for over 10% of all child deaths, and the neonatal period (first 28 days of life in humans) accounts for nearly half of these infection-caused deaths, with sepsis, respiratory, and diarrheal diseases being the main contributors (Torow et al., 2023). Infectious diarrheal diseases are commonly caused by enteropathogenic viruses like rotavirus, parasites like cryptosporidia, and bacteria like *Salmonella* or enteropathogenic *Escherichia coli* (EPEC).

In humans, infection with enteric (i.e., non-typhoid) *Salmonella* results in salmonellosis, a gastrointestinal disease characterized by diarrhea, abdominal cramps, fever, and vomiting (Keestra-

Gounder et al., 2015). Severe cases of salmonellosis, especially in infants, elderly, and immunocompromised individuals, may lead to bacteremia (Arshad et al., 2008). Infants and young children are disproportionately affected due to the immaturity of the gastrointestinal tract, including hypochlorhydria and overall lower gastric acidity, as well as an overall accelerated gastric emptying (Bula-Rudas et al., 2015). A common model organism to study gastrointestinal infection in mice is *Salmonella enterica* subspecies *enterica* serovar Typhimurium (*S.* Typhimurium), because its progression mimics the progression of *Salmonella enterica* subspecies *enterica* serovar Typhi in humans (Walker et al., 2023). *Salmonella*, a gram-negative bacterium belonging to the Enterobacteriaceae family, is typically transmitted through the ingestion of contaminated food via the fecal-oral route. After colonization of the gut lumen, *Salmonella* crosses the epithelial barrier either by uptake via M cells overlaying the Peyer's patches in the distal small intestine or by direct enterocyte invasion (Gal-Mor, 2018; Jones et al., 1994; Zhang et al., 2014).

Neonatal mice display a distinct susceptibility compared with adults: they allow intestinal colonization by *Salmonella* and the formation of prominent intraepithelial microcolonies without prior antibiotic treatment (Zhang et al., 2014). This contrasts with adult mice, which generally require antibiotic pretreatment to permit colonization, and without antibiotic pretreatment, only a fraction of exposed mice will get infected (Barthel et al., 2003). Despite anatomical and physiological differences between murine and human neonates, this model provides a powerful system to study infection-induced changes in early life under conditions of an underdeveloped immune system.

Importantly, the neonatal period coincides with the milk-dependent phase of intestinal development. Milk not only serves as the primary nutrient source but also drives epithelial maturation and gut homeostasis. Consistent with this, neonatal *Salmonella* infection markedly reduces intestinal *Lct* expression and activity, impairing lactose, milk's main carbohydrate, digestion and modeling transient lactose intolerance during enteric infection. Such metabolic disturbances can increase disease severity and prolong symptoms in neonates compared to adults (Ren et al., 2018; Schlößer et al., 2025).

6.6. Molecular profiling approaches and techniques

Understanding how the intestinal epithelium responds to developmental and microbial signals requires methods that capture molecular changes across different levels of regulation. This thesis will feature three commonly used approaches: bulk RNA sequencing (RNA-seq), proteomics, and single-cell RNA sequencing (scRNA-seq).

Massively parallel sequencing was made possible by earlier sequencing technologies like pyrosequencing (also called 454 sequencing), which is based on the enzymatic detection of pyrophosphate release during nucleotide incorporation (Margulies et al., 2005; Ronaghi et al., 1998). However, pyrosequencing has been mostly replaced by Illumina-based sequencing-by-synthesis

technologies for bulk RNA-seq applications due to its limited throughput, higher costs, and shorter read output. Bulk RNA-seq using Next Generation Sequencing (NGS) has become the preferred method for transcriptome analysis because, compared to older sequencing techniques, NGS allows millions of RNA fragments to be simultaneously sequenced, does not rely on predefined probes, and can therefore capture all RNA molecules, including low-abundant or unknown transcripts (Satam et al., 2023). During the process of bulk RNA-seq, the fragile RNA is extracted and, once isolated, reverse-transcribed into the corresponding complementary DNA (cDNA), which is more stable and therefore better suited for sequencing. The cDNA is fragmented into smaller pieces; afterwards, adapters are added, creating sequencing libraries, which can then be loaded onto a high-throughput sequencing platform, which in the end generates the raw data in the form of millions of short sequence reads (Wang et al., 2009).

Proteomics analysis shares similarities with bulk RNA sequencing, as both aim to generate quantitative abundance profiles across biological samples. Whereas bulk RNA-seq relies on high-throughput sequencing to measure transcript levels, proteomics uses mass spectrometry to detect and quantify peptides. A major advantage of proteomics is its ability to capture post-transcriptional and post-translational modifications, which are not detectable using transcriptomic methods. As such, proteomics acts as a complementary method to RNA-seq for understanding gene regulation and cellular states (Low et al., 2019).

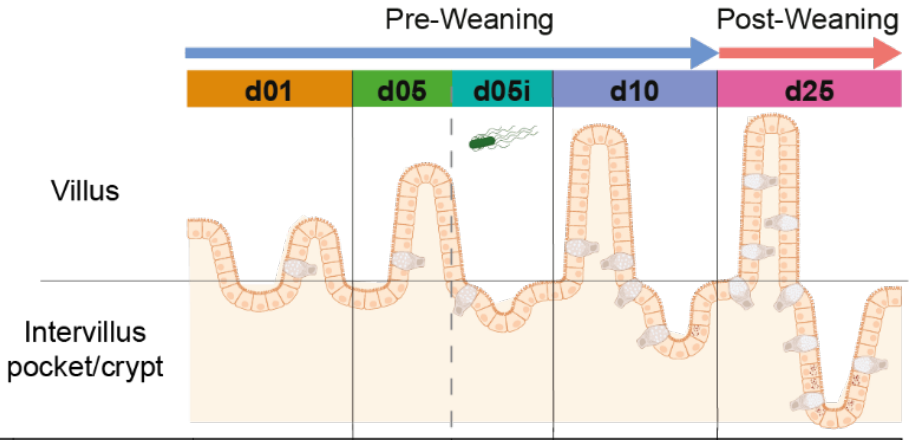
While bulk RNA-seq and proteomics calculate the average expression of multiple cells in a sample, scRNA-seq allows a zoom-in on the single-cell level, revealing the transcriptional profile in rare cell populations as well as the overall transcriptional cellular heterogeneity. To perform scRNA-seq, individual cells are first isolated using flow cytometric sorting or microfluidic-based techniques. Among these, droplet-based microfluidics is now the most widely adopted. In this method, each droplet contains a single cell and a barcoded gel bead. During reverse transcription, cell-specific barcodes and unique molecular identifiers (UMIs) are added to the cDNA, allowing each transcript to be traced back to its cell of origin (Eberwine et al., 2014; Li & Wang, 2021).

6.7. Aim

The aim of this thesis was to characterize the transcriptional and translational landscape of the small intestinal epithelium at both the tissue-wide and single-cell level across early postnatal development. Particular emphasis was placed on spatial organization along the crypt-villus axis and the proximal-to-distal axis of the small intestine. In addition, this thesis aimed to assess how enteric bacterial colonization and infection shape the global and single-cell transcriptional profile of the neonatal intestinal epithelium (Figure 2).

To address these questions, bulk RNA sequencing, mass spectrometry-based proteomics, and single-cell RNA sequencing were employed. These approaches were complemented by bulk RNA-seq data generated from crypts/intervillus regions and villi using laser capture microdissection, from proximal, medial, and distal segments of the small intestine, and from germ-free mice. All datasets were analyzed using established computational frameworks.

Based on these datasets, epithelial maturation, cellular heterogeneity, and regulatory signaling pathways active during early postnatal intestinal development were investigated using functional enrichment (Gene Ontology) and pathway activity-based analytical frameworks. This detailed characterization of postnatal epithelial maturation is expected to provide new insights into the transition of cell types from pre- to post-weaning and the establishment of host-microbial and immune-related homeostasis in the small intestine.



RNA-Seq	SPF (full)	4	4	3	3	3
	GF (full)	4	4	-	4	-
	SPF (P/M/D)	3	3	-	3	3
3'RNA-Seq	SPF (C/V)	3	3	-	3	3
scRNA-Seq	SPF (full)	1	1	1	1	1
Proteomics	SPF (full)	3	3	-	3	3

Figure 2: Schematic overview of the development of the murine small intestine. Schematic representation of the murine small intestine development during the early postnatal period (top) and a table summarizing the different samples (number of biological replicates indicated) analyzed in this study (bottom). P: proximal, M: medial, D: distal, C: crypt, V: villus. Figure created using BioRender.com.

7. Material & Methods

7.1. Laboratory work

7.1.1. Animals

All animal work was performed according to the regulations specified by the Federation for Laboratory Animal Science Associations (FELASA) and the German Society of Laboratory Animal Science (GV-SOLAS, http://www.gv-solas.de). C57BL/6J mice were bred and maintained under specific pathogen-free conditions in ventilated cages, with a 12-hour light/dark cycle and ad libitum access to food and water at the Institute of Laboratory Animal Science at the RWTH University Hospital Aachen. GF mice were provided by Tom Clavel and Susan A. Jennings (RWTH University Hospital, Aachen). Experimental animal work was approved by the North Rhine-Westphalia Landesamt für Verbraucherschutz und Ernährung (LAVE, approvals 81-02.04.2017.A397). 1-day-old C57BL/6J SPF newborn mice were orally infected with 1 µl PBS containing 100-300 log-phase *S.* Typhimurium (ATCC14028). 1-, 5-, 10-, and 25-day-old untreated and 5-day-old infected (4 days post-infection) animals were decapitated, and their small intestines were removed and processed as described below. All animal work was performed by Dr. Aline Dupont.

7.1.2. Sample preparation for bulk RNA-seq and proteomics

Intestinal epithelial cells were isolated as described (Pott et al., 2012). Briefly, full small intestines or parts of the small intestines (the proximal part corresponding to 5 to 20%, the medial part to 37.5 to 62.5%, and the distal part to 80 to 95% of the full small intestine) were cut into small pieces, incubated at 37°C in 30 mM EDTA for 10 minutes, and shaken for 20 seconds. Detached epithelial cells were then filtered through a 100 µm cell strainer, washed with PBS, and pelleted by centrifugation. Pellets were then either resuspended in TRIzol reagent (for RNA-seq) or snap-frozen in liquid nitrogen (for proteomics) and stored at -80°C. IEC isolation was done by Aline Dupont.

For bulk RNA-seq, RNA was subsequently isolated by Aline Dupont following the manufacturer's instructions, and cDNA libraries were prepared by Jasmin Hübner and Anna Rudzinski using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina with the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs) and sequenced in paired-end mode on a NextSeq 500 or a NovaSeq 6000 sequencer (Illumina) at the Genomics Facility, a core facility of the Interdisciplinary Centre for Clinical Research (IZKF) within the Faculty of Medicine at RWTH Aachen University. Each sample corresponded to one given animal, except for day 1 samples, for which the small intestines of two 1-day-old animals were pooled.

For the sample preparation for proteomics, snap-frozen cells were lysed by the addition of 150 µL lysis buffer (4% SDS, 100 mM Tris-HCl pH 7.5, 100 mM DTT) and heated for 5 min at 95°C. Cell lysates were

sonicated for 10 s, centrifuged at 16,000 x g in a tabletop centrifuge at RT, and supernatant added onto 10 kDa cutoff filters (#OD010C33, PALL). Protein concentrations were measured at 280 nm using a NanoDrop (Thermo Fisher Scientific). 250 µg of protein of each sample was alkylated and digested using filter-aided sample preparation (Wiśniewski et al., 2009). Every sample was incubated with 0.45 µg trypsin at 37°C overnight, apart from the pooled sample of three 1-day-old animals, where 0.65 µg trypsin was used. Peptide concentration after elution was measured at 280 nm using a NanoDrop, and peptides were cleaned with StageTip C18 columns prior to mass spectrometry (MS) analysis (Rappsilber et al., 2007). For proteomics, each sample corresponded to one given animal, except for day 1 samples, for which the small intestines of three 1-day-old animals were pooled.

7.1.3. Crypt and villi sample preparation for 3' RNA-seq

Small intestines were rolled and snap-frozen in liquid nitrogen for 30 s to 1 min before being transferred to a -80°C freezer until further use. 12 µm-thick cryosections were then cut using a Leica cryotome and mounted onto UV-treated Zeiss PEN membrane slides. Next, slides were stained by being transferred to successive baths of 95% ethanol, 75% ethanol, 50% ethanol, 1% cresyl violet solution in diethylpyrocarbonate (DEPC)-treated water, 50% ethanol, 75% ethanol, 95% ethanol, and 100% ethanol (x2) for 30 seconds each. The slides were finally left for 5 minutes in 100% ethanol before being allowed to dry for 5 minutes at room temperature. Crypts (1 million µm²) and villi (2 million µm²) were then microdissected using a Zeiss PALM MicroBeam laser microdissection system (10x objective, cut energy 57 (1-100), cut focus 58-60 (1-100)) and catapulted into opaque UV-treated Zeiss adhesive caps.

RNA was immediately isolated from the microdissected tissue pieces using the microRNeasy kit (Qiagen) following the manufacturer's instructions. Finally, cDNA libraries were prepared using the QuantSeq 3' mRNA-seq library prep kit FWD with the i5 6 nt dual indexing add-on module (Lexogen) and sequenced in single-end mode on a NextSeq 500 sequencer. Each sample corresponded to the small intestine of two pooled animals. In addition, 4µm-thick PFA-fixed paraffin-embedded small intestinal tissue sections were stained with hematoxylin and eosin to illustrate postnatal crypt development. Sample preparation of crypt and villus samples was done by Isabel Richter and Matthias Schmitz.

7.1.4. Sample preparation for single-cell RNA-seq

Single intestinal epithelial cells were isolated using a protocol adapted from Haber et al. Longitudinally cut-open small intestines were placed in a 50 mL tube filled with PBS, shortly vortexed to remove the luminal content, and further cut into 2- to 5-mm-long pieces. The intestinal pieces were then transferred into a 50 mL tube prefilled with 30 mL room-temperature RPMI supplemented with 5% fetal calf serum (FCS) and 2 mM EDTA. Next, the tube was placed on a shaker set at 100 rpm at room

temperature. After 15 minutes of incubation, single epithelial cells were filtered through a 100 µm cell strainer and pelleted by centrifugation at 300xg at 4°C. The pellet was resuspended in 1 mL PBS containing 5% FCS and 0.01 mM EDTA and kept on ice. The remaining tissue pieces were transferred into a new 50 mL tube containing 30 mL room-temperature RPMI supplemented with 5% FCS and 2 mM EDTA, and the cell isolation process was repeated three more times. The first fraction collected was discarded, while the epithelial cells contained in the last three fractions were pooled and stained using a rat anti-mouse PE-Cy7-labelled anti-EpCAM antibody (118215, BioLegend), a rat anti-mouse FITC-labelled anti-CD45 antibody (103108, BioLegend), and DAPI (Carl Roth). Single, alive EpCAM+ CD45- cells were sorted using the SH800 Cell Sorter (Sony). Isolation of barcoded RNA and subsequent cDNA library preparation were performed using the Chromium single-cell 3' reagent kits v2 and the Chromium Controller (10x Genomics), following the manufacturer's instructions. cDNA libraries were sequenced in paired-end mode on a NextSeq 500 sequencer (Illumina) at the Genomics Facility of the IZKF. Each sample corresponded to three to six pooled animals. Two day 25 samples were pooled for the final analysis to increase the overall number of cells. Sample preparation was done by Stefan Schlößer and Aline Dupont.

7.1.5. Mass spectrometry analysis of proteomic data

Nano liquid chromatography (LC)-MS/MS was performed on a Q-Exactive High Field (HF) mass spectrometer (Thermo Fischer Scientific), connected with an EASY-nLC 1000 system (Thermo Fischer Scientific) through a nanoelectrospray ion source. Peptides were loaded on a reverse-phase column (150 mm^3 0.075 mm inner diameter, New Objective, Woburn, MA) packed in-house with Reprosil-Pur C18-AQ 3 mm particles (Dr. Maisch, Ammerbuch, Germany). Peptides were separated with a 230-minute gradient: from 3 to 25% B in 175 min, 25 to 45% B in 30 min, and 45 to 100% B in 5 min, followed by a 20-min wash with 100% B (A: 0.1% formic acid, B: 0.1% formic acid/80% acetonitrile) using a flow rate of 250 nl/min. The Q-Exactive HF was operated at a 250°C capillary temperature and 2.0 kV spray voltage. Full mass spectra were acquired in the Orbitrap mass analyzer over a mass range from m/z 350 to 1600 with a resolution of 60,000 (m/z 200) after accumulation of ions to a 3e6 target value based on predictive automatic gain control (AGC) from the previous full scan. Twelve of the most intense peaks with a charge state ≥ 2 were fragmented in the higher-energy collisional dissociation (HCD) collision cell with normalized collision energy of 27%, and a tandem mass spectrum was acquired in the Orbitrap mass analyzer with a resolution of 15,000 after accumulation of ions to a 1e5 target value. Dynamic exclusion was set to 30 seconds. The maximum allowed ion accumulation times were 20 ms for full MS scans and 50 ms for tandem mass spectrum. MS analysis was done in the lab of Dr. Thaher Pelaseyed (Gothenburg University) by Sofia Jäverfelt.

7.2. Bioinformatic analysis

All bioinformatics analyses were performed using R (version 4.4, R Core Team, 2024). Volcano plots were generated using EnhancedVolcano (version 1.24, Blighe, 2018), other figures were generated using the ggplot2 package (version 3.5.2, Wickham, 2016), unless otherwise specified.

7.2.1. Preprocessing

MS raw files were processed by Dr. Thaher Pelasyed with MaxQuant software version 1.5.7.4 (Cox & Mann, 2008), and peak lists were identified by searching against the UniProt database for mouse proteins (downloaded 2018.07.11). Searches were performed using trypsin as an enzyme, a maximum of 2 missed cleavages, and a precursor tolerance of 20 ppm in the first search used for recalibration, followed by 7 ppm for the main search and 0.5 Da for fragment ions. Carbamidomethylation of cysteine was set as a fixed modification, and methionine oxidation and protein N-terminal acetylation were set as variable modifications. The required false discovery rate (FDR) was set to 1% for both peptide and protein levels, and the minimum required peptide length was set to seven amino acids. Label-free quantification (LFQ) based on two peptides was used.

RNA-seq raw sequence reads were trimmed for adapter/low-quality reads using cutadapt (version 4.4), aligned to the reference genome (mm10, GRCm38) using STAR (version 2.7.10a), and afterwards a read count matrix was generated using the featureCounts function from the Rsubread package (version 2.14.2, Liao et al., 2014). Trimming was performed by Keshav Prasad Gubbi, a member of the Hornef lab. For the preprocessing of the scRNA-seq data, demultiplexing of FASTQ files and alignment to the mm10 mouse genome were performed using the Cellranger toolkit (version 3.1.0) by 10x Genomics and version 2.5.1b of STAR included in Cellranger. This was performed at the Genomics Facility of the IZKF at the Uniklinik Aachen. The output BAM (Binary Alignment Map) files were stored for RNA velocity.

7.2.2. Sex- and gender-based analyses

Since the focus of this study was on neonatal and infant samples collected prior to the animals' puberty, sex-specific characteristics were not analyzed. Both male and female neonatal and infant animals were included in the study; sex-specific genes from the X-chromosome inactivation center (*Xist*, *Jpx*, *Ftx*, *Tsix,* and *Cnbp2*) were removed from the datasets for subsequent analyses (Chow et al., 2005; Chureau et al., 2002).

7.2.3. Batch correction

Batch correction refers to the computational process of identifying and removing technical variations (batch effects) from high-throughput biological data, such as bulk RNA-seq or proteomics. These batch effects result from non-biological differences in sample processing, such as differences in time,

equipment, or reagents, and they can obscure true biological signals. Correcting for these effects is essential to ensure that downstream analyses reflect genuine biological variation rather than technical artifacts (Leek & Storey, 2007). Batch correction was performed using ComBat-Seq (version 3.44, Zhang et al., 2020) for both the global transcriptomic dataset (including infected and uninfected samples) and the crypt-villus dataset. For the global dataset, the fourth replicates from day 1 (d01) and day 5 (d05), as well as all three replicates from *Salmonella* Typhimurium-infected mice at day 4 post-infection (d05i), were sequenced on a different cartridge than the rest of the samples. The availability of additional replicates for d01 and d05 allowed batch effects associated with the d05i samples to be distinguished from the already strong transcriptional changes induced by infection. In the crypt-villus dataset, batch effects arose from both differences in sample generation (by two different experimenters) and sequencing across different runs. Specifically, samples were prepared either by Isabel Richter (n = 1 per time point and intestinal region) or by Matthias Schmitz (n = 2 per time point and region). To account for this effect, experimenter identity was included as a batch variable during correction.

The effectiveness of different batch correction methods was first evaluated using the global bulk RNA-seq datasets. In the uncorrected principal component analysis (PCA), the effects introduced by the second batch of samples are visible (Figure 3a). Three different techniques were evaluated against each other. ComBat-Seq (Figure 3b), which utilized a negative binomial regression model (Zhang et al., 2020), was evaluated against two alternative approaches: modeling batch effects as a covariate in the design formula of DESeq2 (Figure 3c) and applying the removeBatchEffect function from the limma R package (Figure 3d, version 3.54.2, Ritchie et al., 2015).

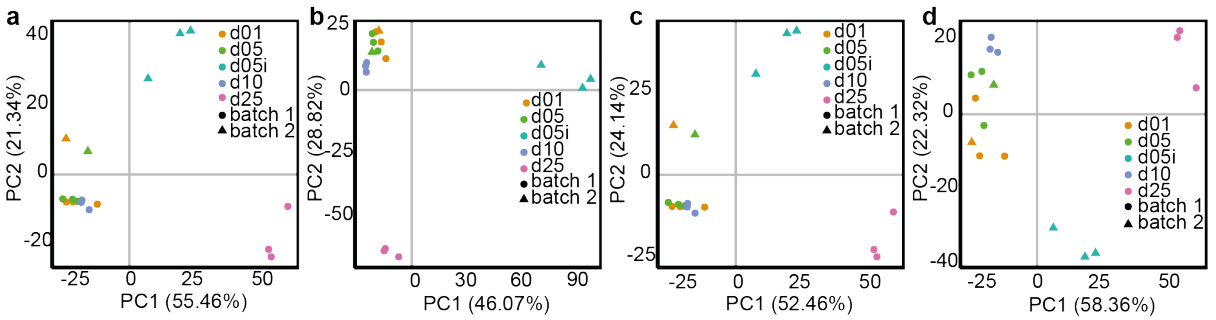


Figure 3: Comparison of different batch correction methods for the global transcriptomic data. (a-d) PCA plots representing the epithelial transcriptome of the small intestine of 1-day-old (orange, n = 4), 5-day-old (green, n = 4), 10-day-old (blue, n = 3), and 25-day-old (pink, n = 3) uninfected mice, as well as 5-day-old mice infected at birth with S. Typhimurium (turquoise, n = 3). **(a)** PCA plot without any batch correction, **(b)** PCA plot after adjustment of batch effects using ComBat-Seq, **(c)** PCA plot after adjustment of batch effects by incorporating the batch into DESeq2's design formula, **(d)** PCA plot after adjustment of batch effects using the removeBatchEffect function from the limma package.

Whereas the incorporation of the batch to the formula in DESeq2 failed to reduce the batch effect, limma and ComBat-Seq both eliminated the batch-associated variation and clustered all replicates of d01, d05, and d10 closely together without any obvious batch effect and therefore preserved biological signal rather than technical artifacts. ComBat-Seq clustered these samples even closer together than

any of the other methods, which was empirically chosen to be closest to the biological expectation and therefore chosen as the batch correction method of choice.

7.2.4. Principal component and differential expression analysis

A PCA on the first two principal components was conducted after batch correction, using the top 500 driver genes of the first and second principal components. Ellipses for optical differentiation between pre- and postweaning samples were estimated using the Khachiyan algorithm implemented inside the ggforce package (Pedersen, 2025).

Subsequent pre-analysis and differential expression analysis of the RNA-seq data were carried out using the DESeq2 package (version 1.38.3, Love et al., 2014). Raw read counts were normalized utilizing DESeq2's VST method, and low-count genes (overall expressions below 60 counts in the full dataset) were filtered out, and zeroes were imputed by adding one.

To identify differentially expressed (DE) genes, a generalized linear model was employed to fit dispersion estimates, with the likelihood ratios test (LRT) used to establish statistical significance. Genes that showed an adjusted p-value (adjustment method: false discovery rate (Benjamini & Hochberg, 1995)) of less than 0.05 and $\log_2$ fold change (FC) higher/lower than one were identified as differentially expressed (DE). Additionally, the ashr method was applied to shrink log2 fold changes (Stephens, 2016). To allow overall comparability, the fold changes and adjusted p-values in each analysis were capped at maximum and minimum thresholds corresponding to the 0.01 and 99.9 percent quantiles (negative and positive fold changes) for fold changes and the 0.01 percentage quantile for adjusted p-values in the volcano plots.

7.2.5. Temporal DE gene analysis

Temporally DE genes were identified by comparing DE genes between any two consecutive time points (d05 vs. d01; d10 vs. d05; d25 vs. d10). Hierarchical clustering was performed on gene expression values; the ward.D2 method was chosen empirically for gene clustering. The expression of the temporal DE genes in each pattern was compared to the overall expression of genes either up- or down-regulated in the presence of the transcription factors BLIMP1 and MAFB (Bara et al., 2022; Harper et al., 2011; Muncan et al., 2011). A Fisher's exact test was performed to determine significance, and results were plotted in a heatmap below the patterns.

7.2.6. Analysis of proteomic LFQ values

The proteomic raw data were further analyzed using the limma R package (Ritchie et al., 2015, version 3.54.2). Protein intensities were log-transformed, normalized, and fitted to a linear model, with an

empirical Bayes model applied to stabilize variance estimates. Differentially abundant proteins were identified using the same thresholds as previously described (chapter 7.2.4).

To identify the origin of the proteins that were present in the proteomic dataset but absent in the bulk RNA-seq count matrix, the processed Seurat R-object files for mammary gland tissue published inside the *Tabula Muris* dataset were obtained from the published repository (The Tabula Muris Consortium, 2018). The R object already contained annotations and clustered cell types. This annotation was consequently used to map the uncorrelated proteins to the specific cell types of the mammary gland tissue.

7.2.7. Analysis of single-cell data

Quality control, normalization, dimensionality reduction, unsupervised clustering, and identification of differentially expressed genes were performed using the Seurat R package version 4 (Stuart et al., 2019). During the quality control, cells with less than 600 genes (nFeature_RNA) and more than 25% mitochondrial counts were filtered out. Additionally, cells with fewer than 1,000 or more than 40,000 total RNA counts (nCount_RNA) were excluded to remove low-quality cells (e.g., empty droplets) and potential doublets. A subset of the 2,000 highest variable genes was identified using Seurat's FindVariableFeatures function. Seurat's default integration algorithm was used to combine datasets for the main analysis. Seurat processes scRNA-seq data by normalizing, identifying highly variable genes, and reducing dimensionality to cluster cells based on transcriptional similarity.

Clusters were annotated based on canonical marker genes from literature. Smaller subclusters were merged to generate five clusters to accommodate the five major small intestinal cell types (absorptive enterocytes, EECs, goblet and Paneth cells, stem cells, and tuft cells). Three of the five identified subclusters of cells (all but stem and tuft cells) were further divided into subsets and reanalyzed to explore gene expression differences within each cell population. For the EEC subset analyses, the Harmony algorithm was used to integrate the data (Korsunsky et al., 2019).

This reanalysis was followed by a new marker gene analysis to identify smaller subpopulations. To visualize the gradient of absorptive enterocytes from crypts to villi, the expression of the top 50 DE genes (based on log2 fold change) of pre-weaning crypts against pre-weaning villi from the 3' bulk RNA-seq of crypts and villi was aggregated and overlaid onto the single-cell data to identify areas of crypts and villi. The gradient from goblet and Paneth cells was identified by the expression of *Zg16* (goblet) and *Lyz1* (Paneth). For the subset of EECs, known marker genes were chosen, and the clusters were annotated according to literature (Haber et al., 2017).

Differentially expressed marker genes (DE genes) for each cluster were identified using a Wilcoxon signed-rank test with false-discovery-rate correction (Seurat's FindMarkers method). Overall, for the DE analysis, the same thresholds as the ones described in chapter 7.2.4 were chosen. To reduce overall

complexity, only consecutive time points (day 5 vs. day 1, day 10 vs. day 5, and day 25 vs. day 10) were subsequently chosen for further analysis.

7.2.8. Transcription factor analysis

For transcription factor analysis, the DoRothEA R package (version 1.18.0) was used (Badia-i-Mompel et al., 2022; Garcia-Alonso et al., 2019) to estimate transcription factor (TF) activities from gene expression data. DoRothEA employs a curated collection of TF-target interactions weighted by confidence levels (A to E). TF activities were computed using the VIPER algorithm, and results were scaled using z-scores for downstream analysis. TF activity was calculated per cluster (absorptive enterocytes, goblet/Paneth, EEC, and stem; tuft cells were excluded due to low cell count) for each time point (days 1, 5, 10, and 25). To assess TF activity consistency across time points, significant TF activity results for the clusters of absorptive enterocytes and goblet/Paneth cells were extracted, and the top 15 most active TFs per time point based on z-score ranking were identified. Presence-absence matrices were generated, with presence coded as 1 and absence as 0. Occurrence heatmaps were then created using the ComplexHeatmap package (version 2.22.0), and hierarchical clustering was applied to visualize TF presence patterns across time points (Gu, 2022; Gu et al., 2016).

7.2.9. Gene Ontology and pathway analysis

DE gene-based Gene Ontology (GO) analysis was calculated in the same manner for all datasets. A comparative cluster analysis was done using clusterProfiler version 4.12.6 (Xu et al., 2024). Comparisons were only performed against the ontology of biological processes (bp). Significant GO terms were only reported with an adjusted p-value of 0.05 (adjustment method: false discovery rate (Benjamini & Hochberg, 1995)). Fold enrichment was calculated as the ratio of the observed proportion of genes associated with a given GO term in the input gene list to the expected proportion based on the background gene set, reflecting how much more frequently the GO term appears than expected by chance. The significant GO terms were afterwards clustered using the complete clustering method. For the global transcriptomic and proteomics data, pathway activities were inferred using the PROGENy (version 3.19) and decoupleR (version 2.10) R packages (Badia-i-Mompel et al., 2022; Schubert et al., 2018). PROGENy computes scores for 14 signaling pathways based on gene expression data. Default settings were used, and pathway scores were scaled for comparisons. Boxplots were generated on these inferred pathway activities per time point, followed by an analysis of variance (ANOVA) and a Tukey-HSD (honestly significant difference) post-hoc test.

7.2.10. RNA Velocity

RNA velocity analysis was performed using the scVelo Python package (version 0.2.5, Bergen et al., 2020), which models transcriptional dynamics by leveraging the ratio of unspliced and spliced mRNA transcripts. The count matrices were combined with the BAM files, which contained the raw alignments needed to quantify spliced and unspliced transcript counts. The data were afterwards normalized and log-transformed, and neighborhood relationships were computed in a low-dimensional embedding space.

RNA velocities were estimated using the dynamical model, which models transcriptional dynamics to capture transient cellular states and estimate latent time. Latent time was inferred from the dynamical model to approximate the internal progression of cells along continuous processes, such as differentiation. Moments of velocity were computed across the nearest-neighbor graph to estimate the RNA velocity vector field. The inferred velocities and latent time were each projected onto the pre-calculated uniform manifold approximation and projection (UMAP) embedding to visualize the direction and magnitude of transcriptional changes across cell states.

7.3. Research data management

The bulk RNA-seq data for the transcriptomic data on organ site and anatomical location dependent as well as age, infection, and microbiota-driven transcriptional changes, are available in Gene Expression Omnibus (GEO) from the National Center for Biotechnology Information (NCBI) under the following accession numbers: GSE283142 (Samples GSM8656671-GSM8677457), GSE283143 (GSM8656695-GSM8656730), and GSE283495 (GSM8663964-GSM8663992). The single-cell data have been made available in GEO under the accession number GSE284074 (GSM8677452-GSM7688487).

Mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE 75 partner repository with the dataset identifier PXD059639.

7.4. Code availability

The complete code used for this study is available on GitHub under the following link: https://github.com/jschoeneich/epithelial_cell_diversity. The repository includes all necessary scripts for reproducing the analysis.

8. Results

8.1. Bulk RNA-seq reveals developmental factors, instead of microbial factors, as main drivers of the early postnatal period

After preprocessing of files and the removal of gender-specific genes, global bulk RNA-seq was performed on small intestinal epithelial cells isolated from the full small intestine of 1-, 5-, 10-, and 25-day-old C57BL/6J SPF mice. A principal component analysis (PCA) captures the greatest sources of variance across samples, and the associated biplot identifies the key genes contributing to each principal component. In this analysis, the PCA revealed that weaning exerts a major impact on the transcriptional profile of intestinal epithelial cells (IECs). In fact, the pre-weaning samples (d01, d05, and d10, marked on the PCA with a blue ellipse) and the post-weaning (d25, pink ellipse) samples clustered on each side of the PC1 axis (explaining 72.69% of the observed variance), while the pre-weaning samples clustered on the PC2 axis (explaining only 12.33% of the observed variance) according to their age (Figure 4a).

Antimicrobial molecules such as defensins (*Defa*), lysozyme 1 (*Lyz1*), phospholipase A2 (*Pla2g2a*), and regenerating islet-derived 3 gamma (*Reg3g*) were identified as driver genes of the post-weaning samples, correlating with the emergence of Paneth cells around weaning, while endotubin (*Mamdc4*), melanotransferrin (*Meltf*), folate receptor (*Folr1*), and glutathione peroxidase (*Gpx3*) were identified as driver genes of the pre-weaning samples, correlating with the presence of fetal-type enterocytes. *H19*, a gene known to be abundantly expressed during embryonic development and downregulated after birth (Poirier et al., 1991), was the main driver gene associated with day 1 samples (Figure 4b).

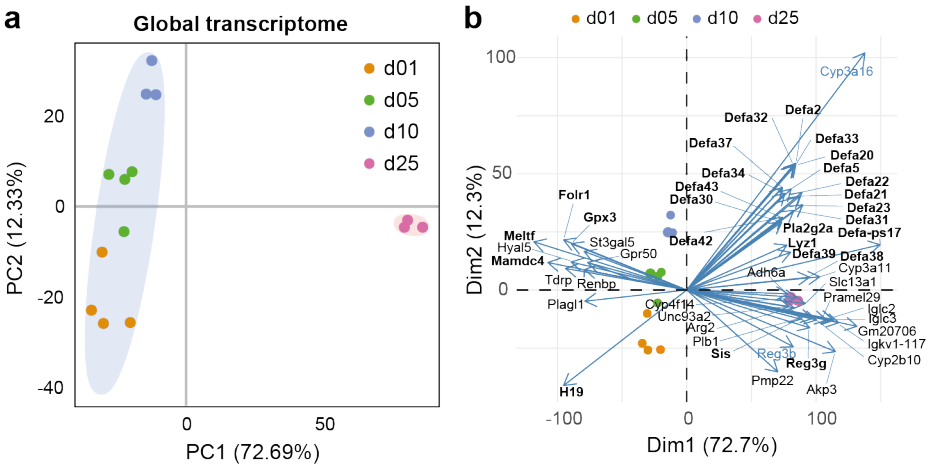


Figure 4: Principal component analysis (PCA) and biplot of the global transcriptomic datasets. (a) PCA plot representing the full small intestinal epithelial transcriptome of 1-day-old (orange, n = 4), 5-day-old (green, n = 4), 10-day-old (blue, n = 3), and 25-day-old (pink, n = 3) mice. Ellipses were estimated using the Khachiyan algorithm and show the distribution of the pre-weaning (light blue) and post-weaning (light pink) samples. **(b)** Biplot displaying the top 50 driver genes for the 1st and 2nd dimensions of the PCA plot shown in (a). Black/blue names represent genes that were identified as DE/non-DE between pre- and post-weaning samples (see Figure 5). 1-day-old (orange dots), 5-day-old (green dots), 10-day-old (blue dots), and 25-day-old (pink dots).

Next, a more nuanced approach was used by identifying differentially expressed genes (DE genes) between consecutive time points (Figure 5). DE genes were considered significant with an adjusted p-value < 0.05 and |log2FC| > 1. The difference in expression between known fetal genes (major histocompatibility complex class Ib gene (*H2-Q10*), *H19,* and coagulation factor IX (*F9*)) and the emergence of early immune genes (defensins) was noted while examining the individual variations between day 1 and day 5. Defensins accounted for the majority of genes upregulated with age between day 1 and day 5 and day 5 and day 10. Stronger differences were observed when comparing day 10 to day 25. In fact, more genes (3253) were identified as DE genes between day 10 and day 25 than between day 1 and day 5 (980) and between day 5 and day 10 (178) combined. Here a significant increase was observed in the expression of the regenerating islet-derived (*Reg*) family genes *Reg3a*, *Reg3b*, and *Reg3g* at day 25, which, similarly to the defensins, suggests an activation of pathways associated with epithelial regeneration and antimicrobial defense.

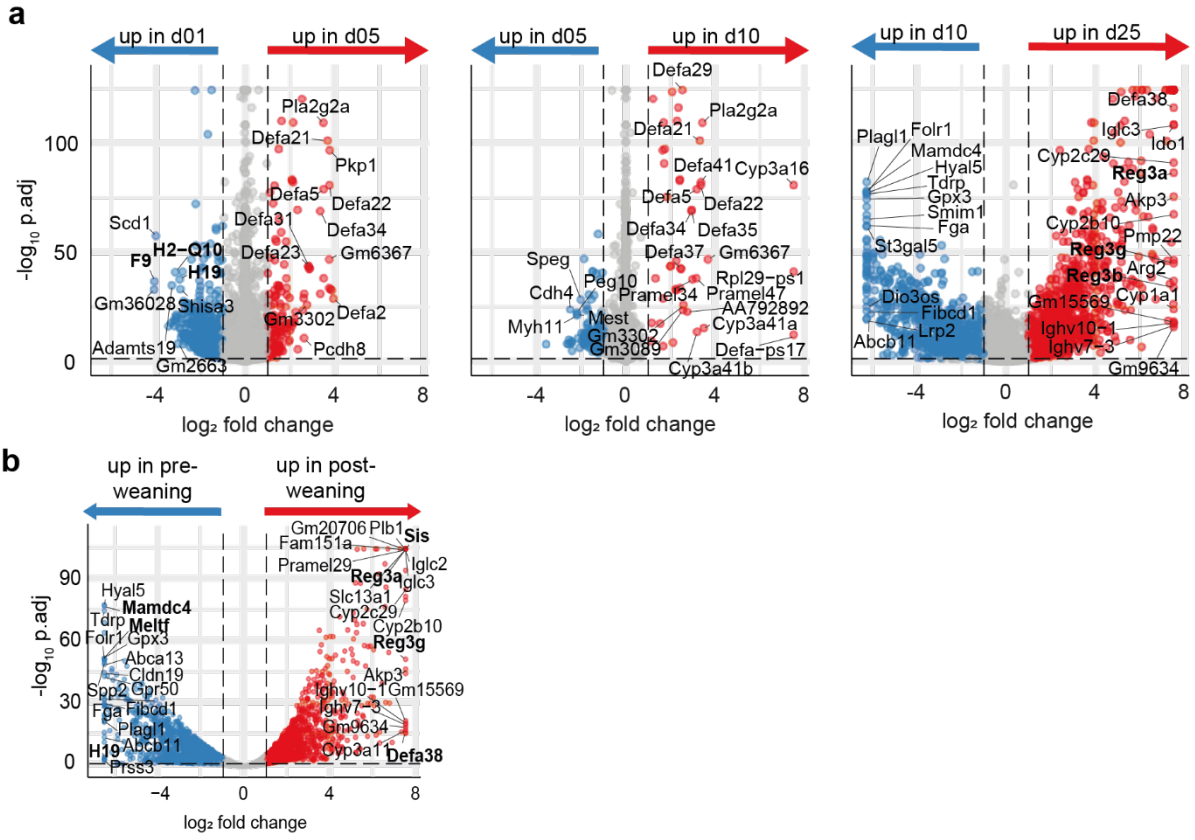


Figure 5: Differential gene expression across time in the global transcriptomic data. (a) Volcano plots showing differential gene expression between 1- and 5-day-old (left), 5- and 10-day-old (middle), and 10- and 25-day-old (right) samples. Blue and red dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in the younger and older animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes. **(b)** Volcano plot showing differential gene expression between pre-weaning (1-, 5-, and 10-day-old) and post-weaning (25-day-old) samples. Blue and red dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in the younger and older animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes.

The genes previously found to drive the biplot were identified when examining the total alterations between pre-weaning and post-weaning samples (Figure 5b). Defensins (*Defa38*, with additional defensins not in the top 10 highlighted genes but substantially expressed in the post-weaning samples),

together with *Reg3g* and *Reg3a*, were significantly higher in the post-weaning samples, while *H19*, *Meltf*, and *Mamdc4* were significantly upregulated in the pre-weaning samples. The digestive enzyme sucrase isomaltase (*Sis*) was also among the genes upregulated during the post-weaning period, emphasizing the transition from breast milk to solid food absorption.

Next, significant DE genes (adjusted p-value < 0.05 and |log2FC| > 1) were evaluated using GO enrichment analysis. In the d01 vs. d05 comparison, the most strongly enriched downregulated GO terms (Figure 6a) were "external encapsulating structure", "extracellular structure organization", and "extracellular matrix organization". These terms suggested that some tissue remodeling occurred shortly after birth. GO terms corresponding to the upregulated DE genes, in contrast, were, as expected, from immune-related categories (Figure 6b). Highly enriched GO terms at all days were "innate immune response in mucosa", "mucosal immune response", and "defense response to Gram-negative bacterium", pointing to the activation of immune defense pathways as animals transition to the post-weaning phase. Whereas all terms that were upregulated between d01 and d05 were also upregulated between d05 and d10, the terms "fatty acid metabolic process", "icosanoid metabolic process", "unsaturated fatty acid metabolic process", and "carboxylic acid biosynthetic process" showed up only between d10 and d25. All these terms show the importance of lipid metabolism and biosynthesis at weaning.

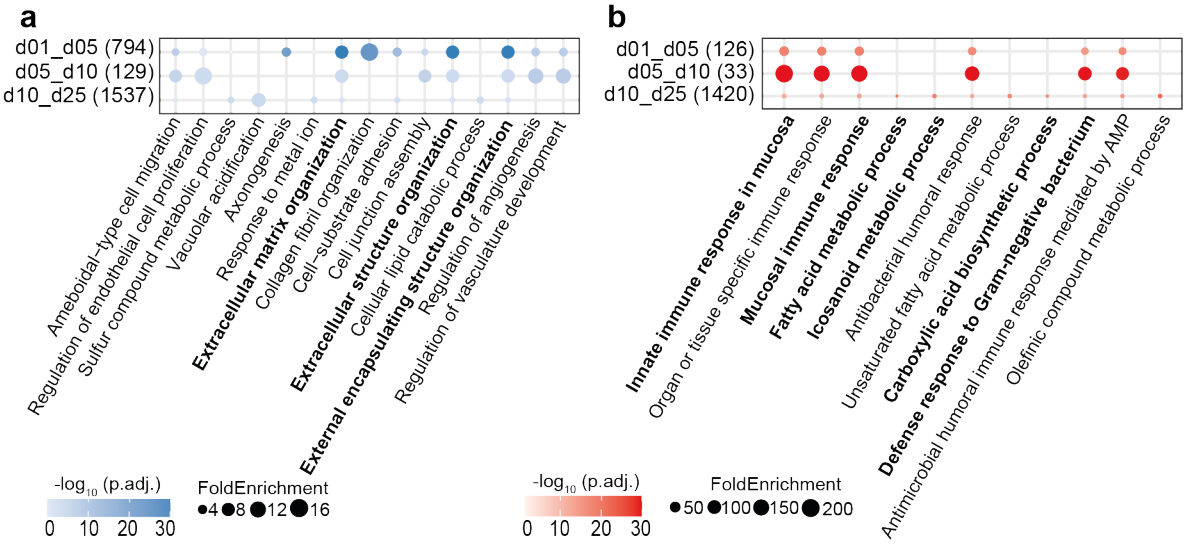


Figure 6: Gene Ontology analysis of differentially expressed genes in global transcriptomic data. (a) Dot plot showing the top 5 enriched GO terms (by adjusted p-value) for the differentially expressed genes downregulated with age (blue DE genes from Fig. 4). Brackets indicate the number of DE genes associated with a GO term for each comparison. **(b)** Dot plot showing the top 5 enriched GO terms (by adjusted p-value) for the differentially expressed genes upregulated with age (red DE genes from Fig. 4). The number of DE genes associated with a GO term for each comparison is indicated in brackets.

The analysis of pathway activity using PROGENy in the neonatal small intestine of mice revealed strong age-dependent variations across several signaling pathways critical for intestinal development. A significant increase was observed in the MAPK, EGFR, and PI3K pathways between d10 and d25, suggesting an increase of cellular signaling associated with growth and maturation during this period. Conversly, JAK-STAT, TGFβ, Trail, and WNT pathways showed a decline over time, potentially a shift

away from their roles in early development as maturation progresses. Hypoxia and PI3K activity decreased between days 1 and 5 but increased again between days 10 and 25, suggesting dynamic regulatory mechanisms at play during intestinal development. The NF-κB, TNFα, p53, and VEGF signaling pathways remained unaltered throughout the postnatal period.

Additionally, both androgen and estrogen levels stayed consistent, suggesting they do not significantly influence developmental processes in the neonatal small intestine during this specific period.

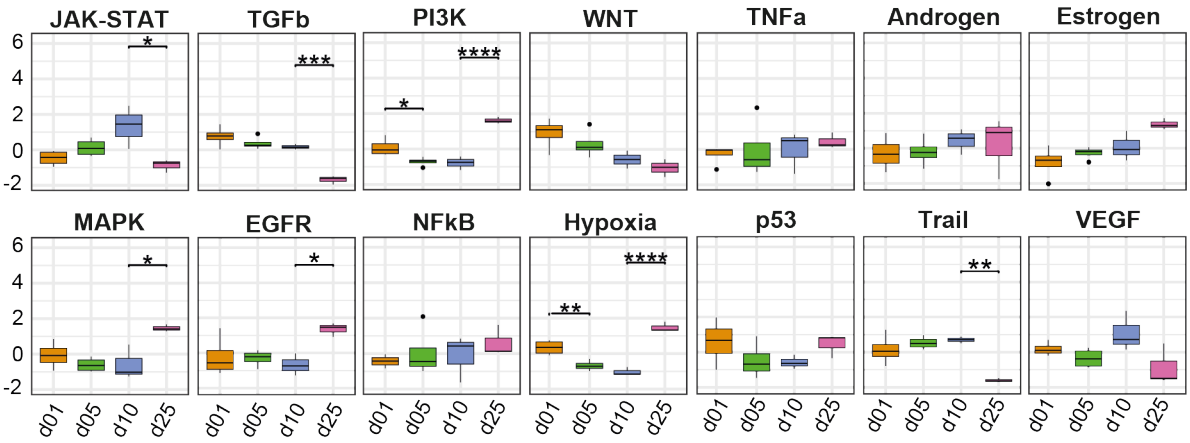


Figure 7: Pathway activity inference on the global transcriptomic data using PROGENy. Boxplots showing PROGENy pathway analysis at different ages. Color code as in Figure 4. Significance was tested using an ANOVA followed by a post hoc Tukey HSD test. The median value is indicated via a black bar. Significance levels are indicated as follows: * p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.0001. Only comparisons between consecutive time points are shown.

As transcriptional changes were observed in specific gene groups at defined postnatal time points, a temporal pattern analysis of gene expression between consecutive time points (1-5, 5-10, and 10-25 days after birth) was conducted (Supplementary Figure 1). DE genes were then further classified as either upregulated (U), downregulated (D), or non-regulated (N) at each transition and further categorized into patterns (Figure 8). Most genes could be categorized into the pattern NND (1568 genes) and NNU (1415), strongly highlighting the importance of the weaning process occurring between the two latest time points of d10 and d25.

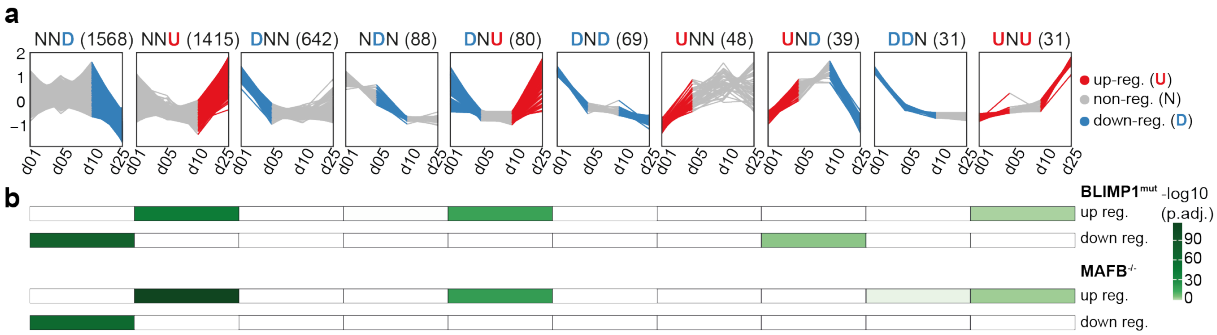


Figure 8: Temporally DE gene expression patterns and association with known major transcriptional regulators BLIMP1 and MAFB. The ten most abundant temporal postnatal gene expression patterns as identified by temporal DE analysis (U, upregulated, red; N, non-regulated, grey; D, downregulated, blue) and their association with reported BLIMP/MAFB gene regulation. The number of DE genes for each pattern is indicated in brackets.

The third largest cluster, DNN (642 genes), consists of genes being downregulated between d01 and d05 and represents fetal genes that still play a role in the early developmental stages of the newborn

mice and then subsequently get less important after birth. Other patterns were found, but all of them were represented by less than 100 genes.

Genes from these temporal patterns were then correlated against other genes known to be regulated by the two known transcription factors BLIMP1 (encoded by *Prdm1*) and MAFB. These two factors are known to play a major role in the transition from neonate to adult (Bara et al., 2022; Harper et al., 2011; Muncan et al., 2011). It could be observed that the genes significantly co-expressed with BLIMP1 and MAFB belonged to the patterns NND, NNU, DNU, and UNU. Genes within DDN (31 genes) were only co-expressed with genes upregulated with MAFB, and genes within UND (39 genes) were only co-expressed with genes that were downregulated with BLIMP1.

GO enrichment analysis was performed on the three most abundant patterns (NND, NNU, and DNN) as well as the UNN pattern, which is the opposite of DNN. In the pattern of NND, downregulation of genes related to “cellular lipid catabolic process”, “sulfur compound metabolic process”, “lysosome organization”, and “vacuolar acidification” was observed, reflecting a developmental shift as the small intestine matures. Interestingly, genes associated with GO terms “alcohol metabolic process”, “fatty acid metabolic process”, “olefinic metabolic process” and “isoprenoid metabolic process“ were upregulated in the NNU pattern during this timeframe, indicating that these pathways could play a crucial role in adapting to changing nutritional demands at this developmental stage. For the pattern DNN, terms corresponding to tissue remodeling after birth (e.g., “extracellular structure organization”, “external encapsulating structure organization”, “collagen fibril organization”) had a high fold enrichment. UNN and DNN both shared two terms corresponding to WNT signaling, indicating the idea that the WNT pathway is not globally up- or downregulated after birth but selectively modulated, with different components playing distinct roles at specific times.

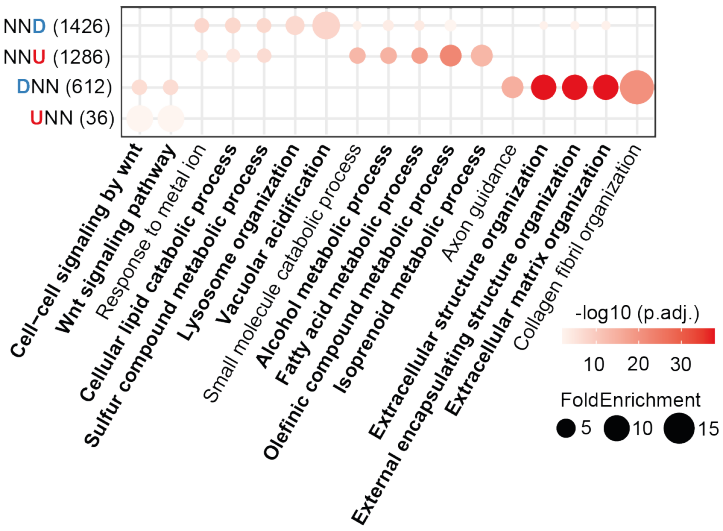


Figure 9: Gene Ontology Analysis of the Temporal DE genes of the global transcriptome analysis. Dot plot showing the top 5 enriched GO terms (by adjusted p-value) for the differentially expressed genes identified as belonging to the NND, NNU, DNN, and UNN expression patterns. The number of DE genes associated with a GO term for each pattern is indicated in brackets.

Further analysis followed on, investigating the differences between pre-weaned conventional mice (SPF) and germ-free mice (GF). Only minor differences and few differentially expressed genes were found when comparing GF and SPF mice (Figure 10). This was in contrast to the findings in adult individuals, where microbiota had been previously identified as affecting immune defense, glycosylation, and metabolism (Tsang et al., 2022).

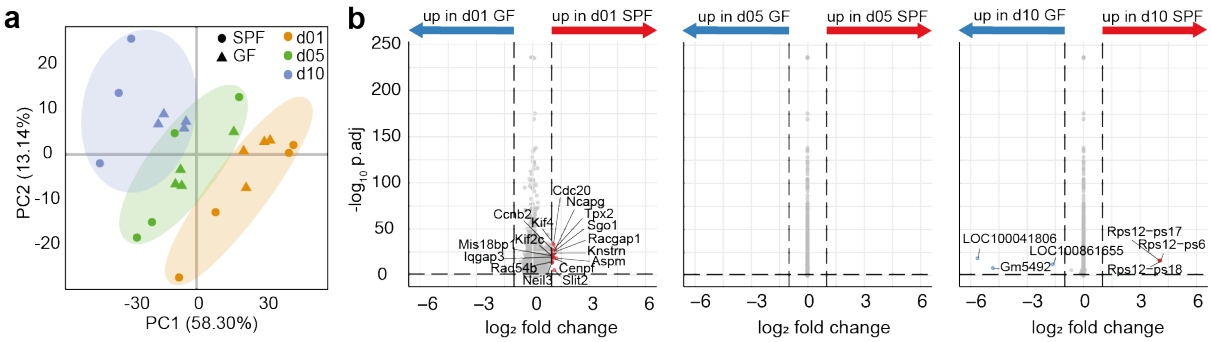


Figure 10: Principal component and differential gene analysis of SPF vs. GF mice in the global transcriptome of SPF and GF mice. (a) PCA plot representing the full small intestinal epithelial transcriptome of 1-day-old (orange, n = 4/4 SPF/GF), 5-day-old (green, n = 4/4), and 10-day-old (blue, n = 3/4) specific pathogen-free (SPF, circles) and germ-free (GF, triangles) mice. Ellipses were estimated using the Khachiyan algorithm and show the distribution of samples for each age group. **(b)** Volcano plots show differential gene expression between GF and SPF mice at 1 (left), 5 (middle), and 10 (right) days of age. Blue and red dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in GF and SPF animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes.

Together, these findings characterize the overall time course and relative importance of epithelial transcriptional alterations that occur during the formation of mucosal host-microbial homeostasis.

8.2. Proteomic transformations highlight developmental adaptations in small intestinal epithelium during weaning

The proteomics analysis of total isolated small intestinal epithelium detected 4562 proteins, of which 4160 were also detected at the transcriptional level by bulk RNA-seq (Figure 11a).

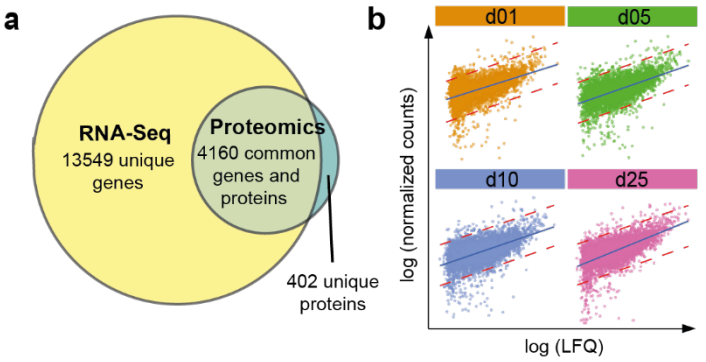


Figure 11: Comparative Analysis of RNA-seq and Proteomics. (a) Venn diagram representing the overlap between the RNA-seq and proteomic datasets. **(b)** Correlation of the expression level of the 4160 common genes and proteins detected by RNA-seq and proteomics in 1-day-old (orange), 5-day-old (green), 10-day-old (blue), and 25-day-old (pink) animals. On the x-axis are log-normalized LFQ intensities, and on the y-axis are log-normalized bulk RNA-seq counts. The blue line indicates the regression line, and the red dashed line the 95% confidence interval.

As expected, comparison of the log-normalized RNA-seq counts and the log-normalized label-free quantification (LFQ) values revealed a strong positive correlation between protein and mRNA expression levels (Figure 11b). Notably, at days 5 and 10, most proteins correlated well with their RNA-seq levels, but several proteins corresponding to low-abundance transcripts exhibited relatively higher LFQ values than the counts of their corresponding transcripts.

The PCA analysis of the proteomic dataset showed a strong difference between the pre-weaning and post-weaning samples (PC1 47.46%), as previously observed for the global transcriptomic RNA-seq dataset analysis (Figure 4a, Figure 12a). However, in contrast to the RNA-seq analysis, the samples of day 1 clustered away from those of day 5 and day 10 (PC2 32.01%). Furthermore, the samples of day 5 and day 10 clustered more tightly together than observed in the transcriptomic data.

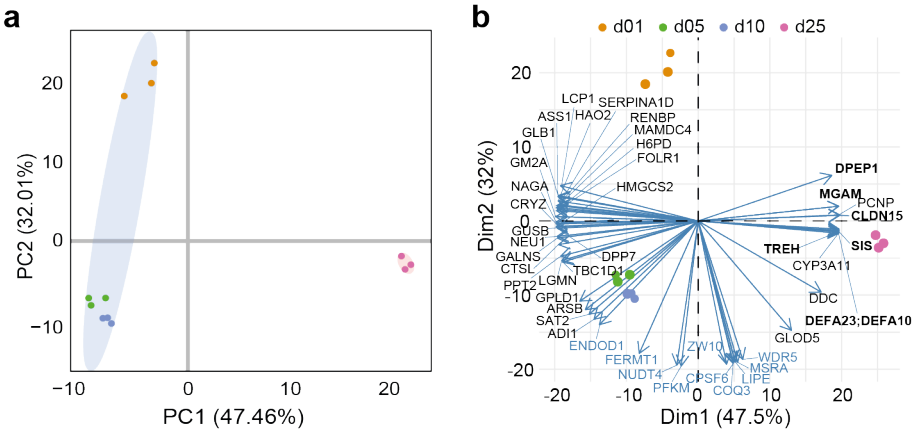


Figure 12: Principal component analysis (PCA) and biplot of changes over time in the proteomic data. (a) PCA plot representing the full small intestinal epithelial proteome of 1-day-old (orange, n = 3), 5-day-old (green, n = 3), 10-day-old (blue, n = 3), and 25-day-old (pink, n = 3) mice. Ellipses were estimated using the Khachiyan algorithm and show the pre-weaning (light blue) and post-weaning (light pink) samples. **(b)** Biplot displaying the top 50 driver proteins for the 1st and 2nd dimensions of the PCA plot shown in (a). Black/blue names represent proteins that were identified as DE/non-DE between pre- and post-weaning samples. Multiple proteins separated by a semicolon indicate a double annotation, where two proteins fit the LFQ intensity. 1-day-old (orange dots), 5-day-old (green dots), 10-day-old (blue dots), and 25-day-old (pink dots).

Analysis of the biplot (Figure 12b) and more distinctly of the volcano plots from the DE analysis (Figure 13a-b) revealed that postnatal maturation and an increased innate immune activity are also reflected at the proteomics level, notably through the emergence of Paneth cells (indicated by defensins DEFA23 and DEFA10). It could also be noticed that metabolic proteins were playing a role during weaning (e.g., SIS, trehalase (TREH), dipeptidase 1 (DPEP1), and maltase glucoamylase (MGAM)). Proteins whose expression decreased at weaning were consistent with previously observed changes in the transcriptomic data. Several of them belonged to fetal-type enterocytes (e.g., MAMDC4, cathepsin L (CTSL), alpha-N-acetylgalactosaminidase (NAGA)), or were metabolic enzymes (e.g., argininosuccinate synthase 1 (ASS1) and glycine amidinotransferase (GATM)).

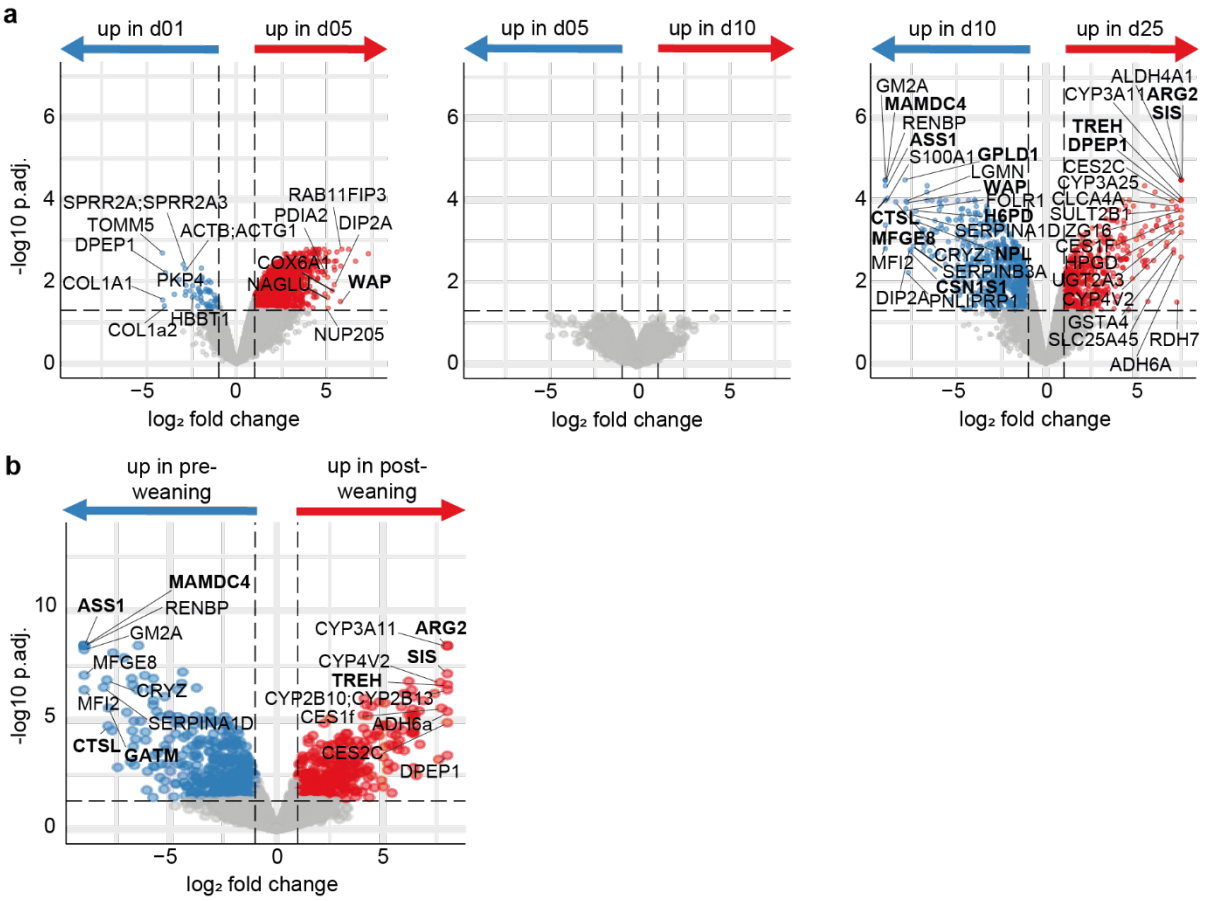


Figure 13: Differential gene expression across time in the proteomic data. (a) Volcano plots showing the differential expression of proteins between 1- and 5-day-old (left), 5- and 10-day-old (middle), and 10- and 25-day-old animals (right). Blue and red dots indicate statistically upregulated proteins (p.adj. < 0.05 & |log2FC| > 1) in the younger and older animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated proteins. **(b)** Volcano plots showing differential protein expression between pre-weaning (1-, 5-, and 10-day-old) and post-weaning (25-day-old) samples. Blue and red dots indicate statistically upregulated proteins (p.adj. < 0.05 & |log2FC| > 1) in the younger and older animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated proteins.

The distinct clustering of day 1 samples away from the other pre-weaning time points could be attributed to the high number of proteins significantly upregulated between day 1 and day 5 (1098), compared to only 58 downregulated. The high number of DE proteins at day 5 compared to day 1 contrasted with the transcriptomic dataset, where fewer genes (126) had been upregulated between day 1 and day 5 after birth. Whereas at least a small amount of DE genes was found when comparing day 5 and day 10 in the bulk RNA-seq dataset, no proteins were significantly differentially expressed between these two time points in the proteomic dataset (Figure 5).

When investigating the temporal changes of the differentially expressed proteins (Figure 14), a notable pattern emerged: UNN, a group with very low abundance in the transcriptomic data (46 genes), transformed into the one with the highest number of entries in the proteomics dataset (878). It was followed by the patterns NND and NNU, which continued to highlight the weaning transition (Figure 14a+b). GO analysis of the four most abundant patterns (UNN, NND, NNU, and UND) showed strong enrichment of metabolic processes in NNU, including steroid, alcohol, hormone, and terpenoid metabolism (Figure 14c). Although some of these metabolic GO terms were also detected in NND, NNU

could be distinguished by the postnatal upregulation of genes corresponding to these terms between days 10 and 25.

Simultaneously, changes occurred in enzymes and transporters, along with remodeling of lysosomes and vesicle trafficking, as evidenced by GO terms like “vacuolar acidification”, “synaptic vesicle maturation”, and “synaptic vesicle lumen acidification” related to pattern NND. These changes indicate that lysosomal and vesicular trafficking pathways are maintained during early neonatal stages and subsequently reduced after day 10.

Additionally, the early upregulation of “Golgi organization” and “Golgi vesicle transport” in patterns UND and UNN not only highlights the secretory needs of the small intestine at these time points but also underscores the importance of the autophagy machinery, which depends on Golgi-derived membranes for autophagosome formation and maturation (Wei et al., 2018).

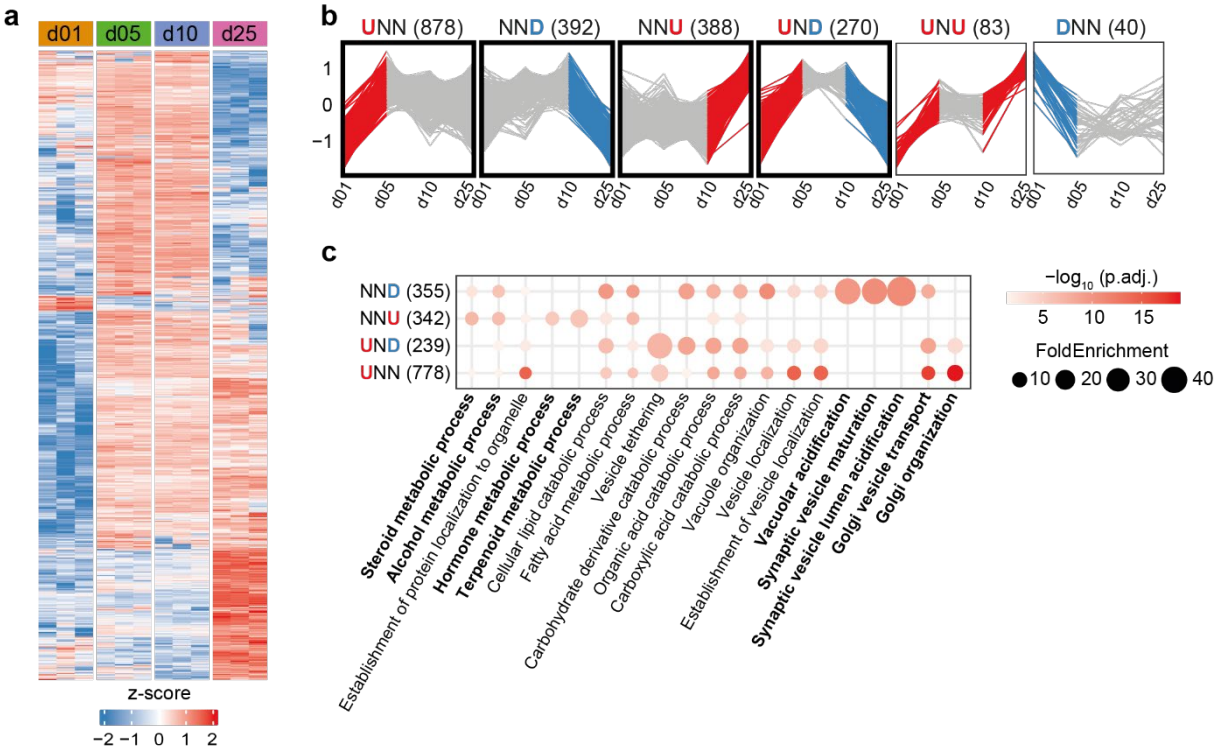


Figure 14: Temporally DE protein expression patterns and corresponding Gene Ontology analysis in the proteomic data. (a) Heatmap depicting the expression level of all proteins differentially regulated between either 1- and 5-day-old, 5- and 10-day-old, and/or 10- and 25-day-old animals. **(b)** Six most abundant temporal postnatal protein expression patterns as identified by temporal DE analysis. The number of DE proteins for each pattern is indicated in brackets. **(c)** Dot plot showing the top 5 enriched GO terms (by adjusted p-value) for the differentially expressed genes identified as belonging to the NND, NNU, UND, and UNN expression patterns. Brackets indicate the number of DE proteins associated with a GO term for each pattern.

One possible explanation for the increase of DE proteins at day 5 compared to day 1 was the presence of milk proteins, such as whey acid protein (WAP), in 5-day-old mice but not in 1-day-old mice (Figure 13). In fact, some milk proteins, such as caseins and whey proteins, are not synthesized by intestinal epithelial cells and therefore lack corresponding transcriptional signals. Instead, these proteins are maternally derived and ingested through breastfeeding (Schoenenberger et al., 1990). Once in the neonatal gut, they may persist transiently before being degraded by proteases. However, due to the

relatively immature digestive capacity of the neonatal intestine at early time points, some of these proteins can remain partially intact and be therefore detectable at the protein level. This observation led to further investigation into the 402 molecules found at protein but not mRNA levels. By utilizing the well-maintained *Tabula Muris* dataset (The Tabula Muris Consortium et al., 2018), which contains over 100,000 single-cell transcriptomes from various mouse organs, the aggregated expression of these proteins with no transcriptional counterpart revealed that 135 could be matched to genes expressed in luminal progenitor and mature luminal cells found within mammary gland tissue, suggesting a possible breast milk origin (Figure 15). Further supporting this hypothesis, additional milk proteins, including WAP, milk fat globule EGF factor 8 domain containing protein (MFGE8), and casein alpha S1 (CSN1S1), were detected in day 10 samples but absent from those collected at day 25 (Figure 13).

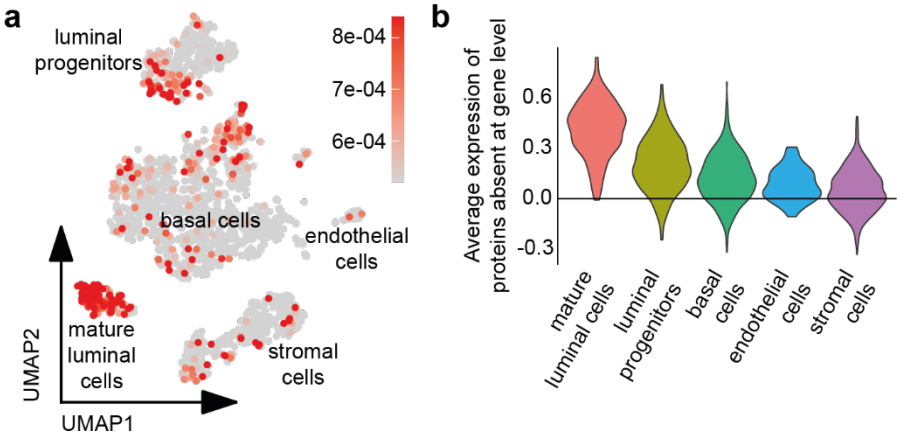


Figure 15: Expression of proteins that are absent at the transcriptomic level in the *Tabula Muris*. (a) Uniform Manifold Approximation and Projection (UMAP) plot displaying the different types of cells in breast tissue identified by the previously published *Tabula Muris* mammary gland dataset and overlapped with the expression level of the intestinal epithelial proteins that were not detected at the transcriptional level by RNA-seq. **(b)** Violin plot showing the quantification data of (a).

Other terms found in the Gene Ontology (Supplementary Figure 2), together with the PROGENy pathway analysis (Figure 16) of the proteomic dataset, confirmed many of the changes in pathway activity that were already observed in the transcriptomic dataset. Similar patterns in JAK-STAT, MAPK, and PI3K pathway activity could be observed, highlighting overall similarities in proteomic and transcriptomic data.

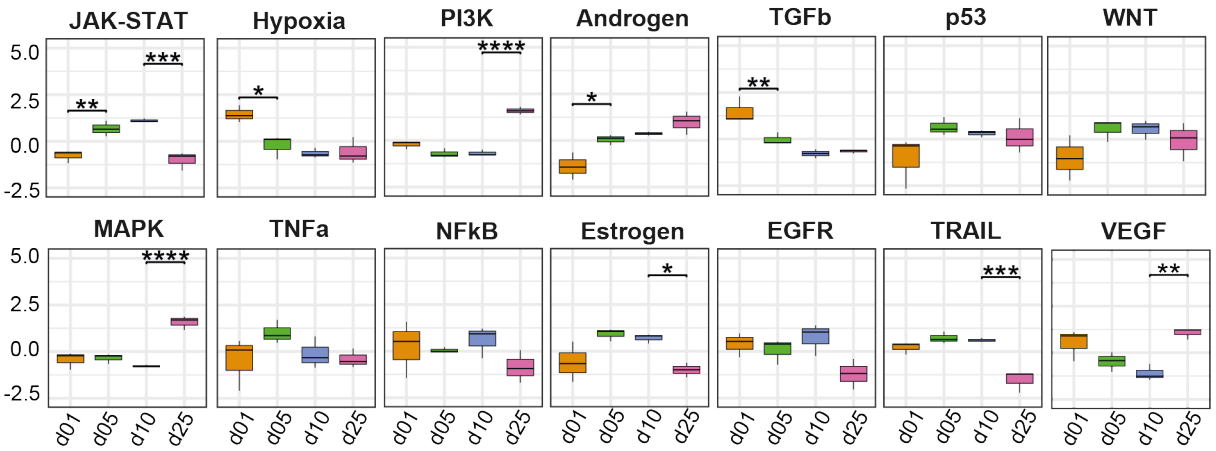


Figure 16: Pathway activity inference on the proteomic data using PROGENy. Boxplots showing PROGENy pathway analysis for the proteomic data at different ages. Significance was tested using an ANOVA followed by a post hoc Tukey HSD test. Significance levels are indicated as * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. Only comparisons between consecutive time points are shown.

Overall, proteomics confirmed many of the major transcriptional trends associated with postnatal maturation and weaning yet provided distinct insights into protein-level dynamics during early life. In particular, the detection of milk-derived proteins at pre-weaning time points underscores the contribution of maternal factors not visible in the transcriptome to the neonatal intestine and explains, at least in part, the pronounced divergence between day 1 and day 5 samples. Thus, proteomics complements transcriptomic analyses by capturing both the endogenous epithelial programs and the exogenous proteins present in the intestinal lumen.

8.3. Age-related shifts in gene expression along the crypt-villus axis

The global analysis of the intestinal epithelium does not allow for the interpretation of the anatomical location and functional specification of different areas of the small intestine. To account for differences along the crypt-villus axis, laser capture microdissection (LCM) was performed to cut and capture areas of crypts/intervillus areas and villi from neonatal mice at different ages and separately analyze them (Figure 17a). After batch correction was applied, PCA revealed a distinct separation between the transcriptomes of crypts/intervillus areas and villus areas at all ages (Figure 17b).

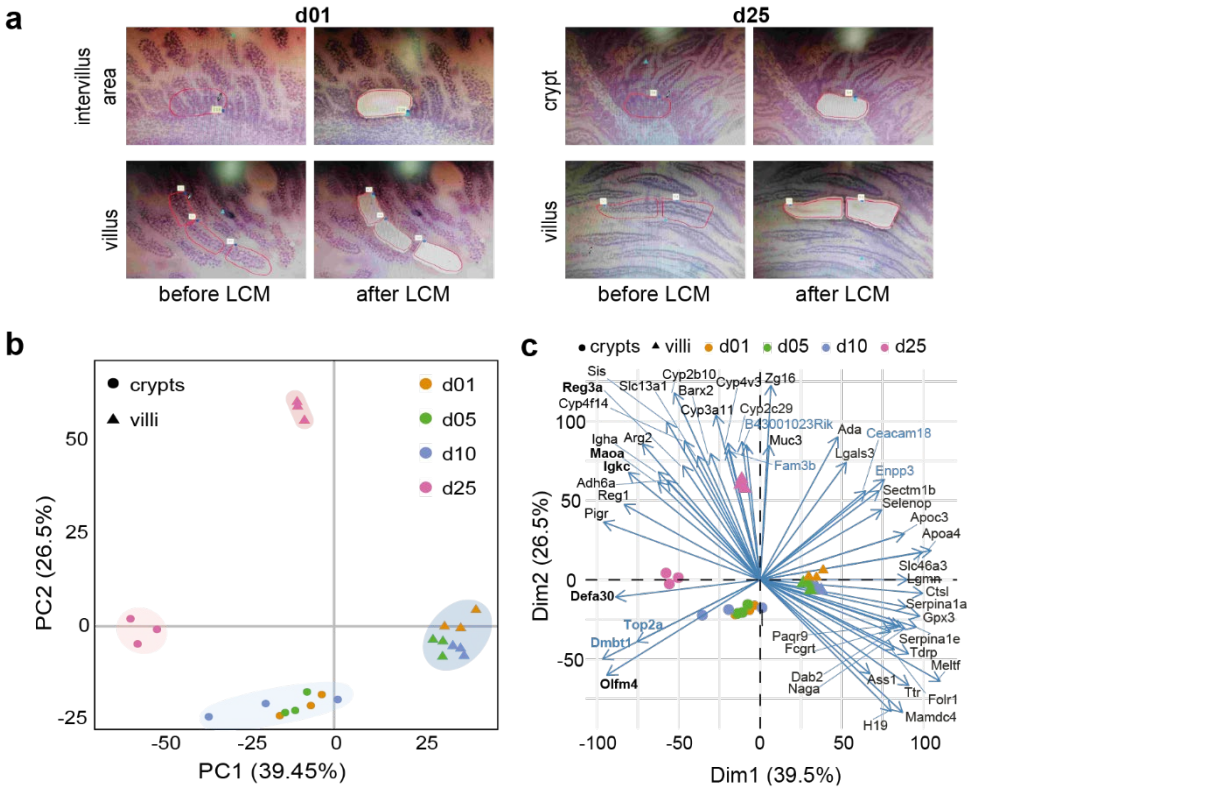


Figure 17: Spatial epithelial gene expression along the crypt-villus axis. (a) Pictures of small intestinal tissue sections, obtained from 1-day-old (left) and 25-day-old (right) animals and stained using cresyl violet. Pictures were taken before (left) and after (right) laser capture microdissection (LCM) of the intervillus/crypt (top) and villus (bottom) areas. **(b)** PCA plot, representing the small intestinal epithelial transcriptome of either the crypt/intervillus area (circles) or villi (triangles) of 1-day-old (orange, n = 3), 5-day-old (green, n = 3), 10-day-old (blue, n = 3), and 25-day-old (pink, n = 3) mice. Ellipses were estimated using the Khachiyan algorithm and show the pre-weaning (crypts: light blue; villi: dark blue) and post-weaning (crypts: light pink; villi: dark pink) samples. **(c)** Biplot displaying the top 50 driver genes for the 1st and 2nd dimensions of the PCA plot shown in (a). Black/blue names represent genes that were identified as DE/non-DE between pre- and post-weaning crypt or villi samples. 1-day-old (orange dots), 5-day-old (green dots), 10-day-old (blue dots), and 25-day-old (pink dots).

In pre-weaning samples, both crypt and villus regions were similar to the results in the global transcriptome and proteomics, characterized by high expression of genes corresponding to fetal-type enterocytes (i.e., *Meltf*, *H19*, and *Mamdc4*). In contrast, post-weaning samples showed increased expression of *Igha,* which encodes a segment of the immunoglobulin A (IgA) heavy chain, and the xenobiotic-metabolizing enzyme *Cyp2b10* in both crypts and villi, indicating a shift towards immune-associated and metabolic gene expression after weaning. As laser capture microdissection does not yield purely epithelial material but also epithelial-adjacent cells, the elevated *Igha* signal likely reflects IgA-producing immune cells adjacent to the epithelium rather than epithelial IgA production. Antimicrobial defensins (*Defa26*, *Defa29*, and *Defa38*) were predominantly enriched in post-weaning crypt samples but were absent from post-weaning villus samples (Figure 18a).

The biplot (Figure 17c) and the following analysis of DE genes showed, independent of age, the stem cell marker *Olfm4* (olfactomedin-4) as a crypt marker and the CEA cell adhesion molecule 18 (*Ceacam18*) as a villus marker (Figure 18b). In the pre-weaning crypts, genes indicative of active crypt proliferation and stem cell maintenance, such as *Ascl2*, were found. *Ascl2* has been shown to act as a Wnt-responsive regulator of intestinal stem cell identity in columnar cells at the crypt base and to drive

stemness in vivo (Schuijers et al., 2015). Additionally, the expression of high mobility group AT-hook 2 (*Hmga2*) could be observed. Hmga family members are known to be enriched in intestinal stem/progenitor compartments and promote proliferation and Wnt pathway activity, consistent with an undifferentiated, proliferative crypt profile (Xian et al., 2017).

In contrast, post-weaned crypts, in comparison to post-weaned villi, expressed defensin alpha genes (*Defa26, Defa38, Defa17, Defa30*), SPI-B transcription factor (*Spib*), and plakophilin 1 (*Pkp1*), reflecting Paneth cell maturation and increased antimicrobial defense (Figure 18b).

In pre-weaned villi, genes such as zymogen granule protein 16 (*Zg16*), mannose-binding lectin 2 (*Mbl2*), and serum/glucocorticoid regulated kinase 2 (*Sgk2*) were enriched compared to crypts. These genes highlight early epithelial defense functions; for example, *Zg16*, which is also a marker of goblet cells (Pelaseyed et al., 2014), contributes to mucus-mediated antimicrobial defense, while *Mbl2* encodes a lectin involved in pathogen recognition (Figure 18b) (Fraser et al., 1998). In post-weaned villi, CC-chemokine ligand 5 (*Ccl5*), sushi domain containing 2 (*Susd2*), and ureidopropionase beta (*Upb1*) were upregulated relative to post-weaning crypts. These genes reflect a shift towards enhanced immune signaling and metabolic processing, with *Ccl5* mediating immune cell recruitment and *Upb1* contributing to pyrimidine metabolism (Figure 18b) (Tyner et al., 2005).

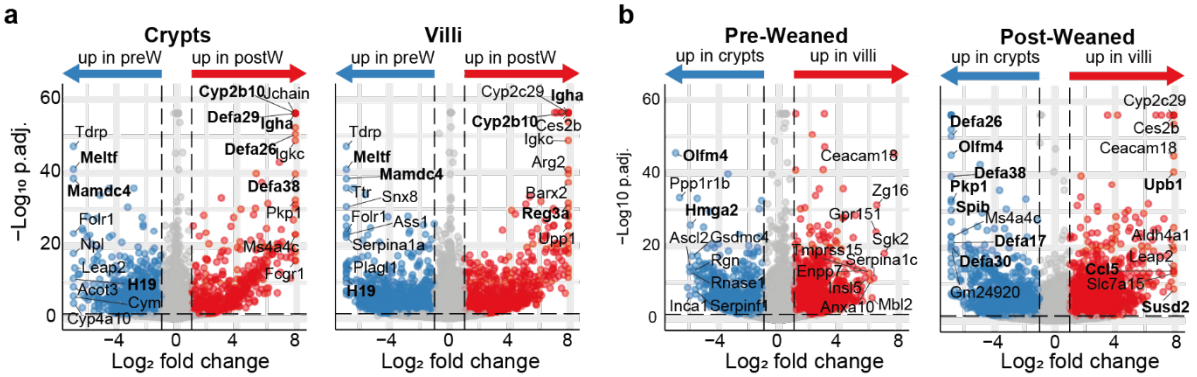


Figure 18: Differential gene expression intra-location and across time along the crypt-villus axis. (a) Volcano plots showing differential gene expression between pre-weaning (1-, 5-, and 10-day-old) and post-weaning (d25) crypt samples (left) and pre-weaning (1-, 5-, and 10-day-old) and post-weaning (d25) villi samples (right). Blue and red dots indicate genes statistically upregulated (p.adj. < 0.05 & |log2FC| > 1) in the younger and older animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes. **(b)** Volcano plots showing differential gene expression between pre-weaning (1-, 5-, and 10-day-old, left) and post-weaning (d25, right) crypt and villus samples, respectively. Blue and red dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in the crypts and villi, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes.

Genes associated with the GO terms “extracellular structure organization”, “extracellular matrix (ECM) organization”, and “cell-substrate adhesion” were downregulated in the crypt area at day 25 compared to day 10, indicating higher expression of ECM-related genes at earlier postnatal stages. This suggests an increased extracellular matrix remodeling activity in crypts during the early postnatal period, which is reduced as crypt architecture becomes established by day 25 (Figure 19). Significantly

downregulated GO terms were not observed in either crypt/intervillus area or villus samples during the interval between day 1 and day 5 and only between day 5 and day 10 for the villi.

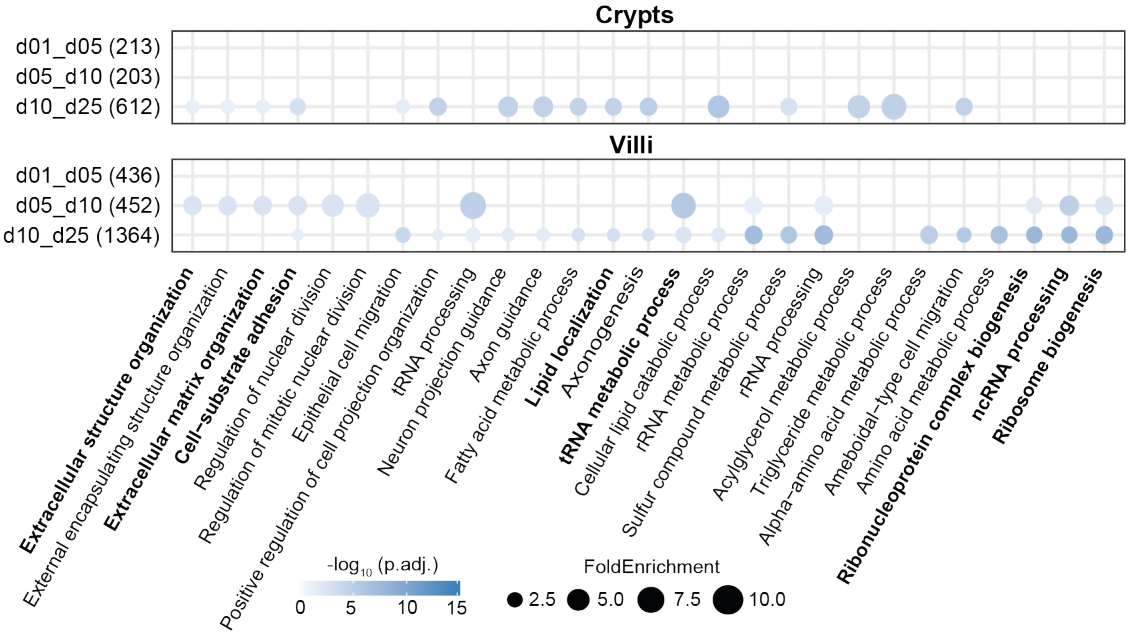


Figure 19: Gene Ontology analysis of differentially expressed downregulated over time in crypts and villi. Dot plot showing the top ten enriched GO terms (by adjusted p-value) for the differentially expressed genes downregulated with age in crypts (top) and in villi (bottom). The number of DE genes associated with a GO term for each comparison is indicated in brackets.

GO terms for DE genes upregulated in the villi at day 5 compared to day 1 (Figure 20) included terms "Response to BMP" and "Cellular response to BMP stimulus" which likely reflect the establishment of early BMP signaling in the intervillus regions, marking the beginning of crypt development. The GO terms seen at day 25 (i.e., terms containing "metabolic process") indicate increased energy requirements and adaptation to novel physiological demands like nutrient absorption and growth (Figure 20).

In the post-weaned crypts, upregulated genes like *Reg3a* and defensins, as well as GO terms corresponding to immune response (Figure 20), highlight the imminent immune response and the full maturation of Paneth cells in the crypts. Similar patterns were observed when investigating genes at consecutive time points (d01-d05, d05-d10, d10-d25) in either crypts or villi (Supplementary Figure 3). Interestingly, translational activity, as indicated by the GO terms "ribosome biogenesis", "tRNA metabolic process", "ncRNA processing", and "ribonucleoprotein complex biogenesis", was upregulated in the villi between day 1 and day 5 and afterwards subsequently downregulated after weaning. This pattern may reflect increased translational demand during early postnatal stages and a reduced requirement for translational activity at the villus tip following the establishment of continuous crypt-villus cell migration after weaning. Taken together, the LCM-derived transcriptomic data demonstrate that both anatomical location along the crypt-villus axis and developmental stage before and after weaning contribute to the dynamic regulation of epithelial gene expression in the neonatal small intestine.

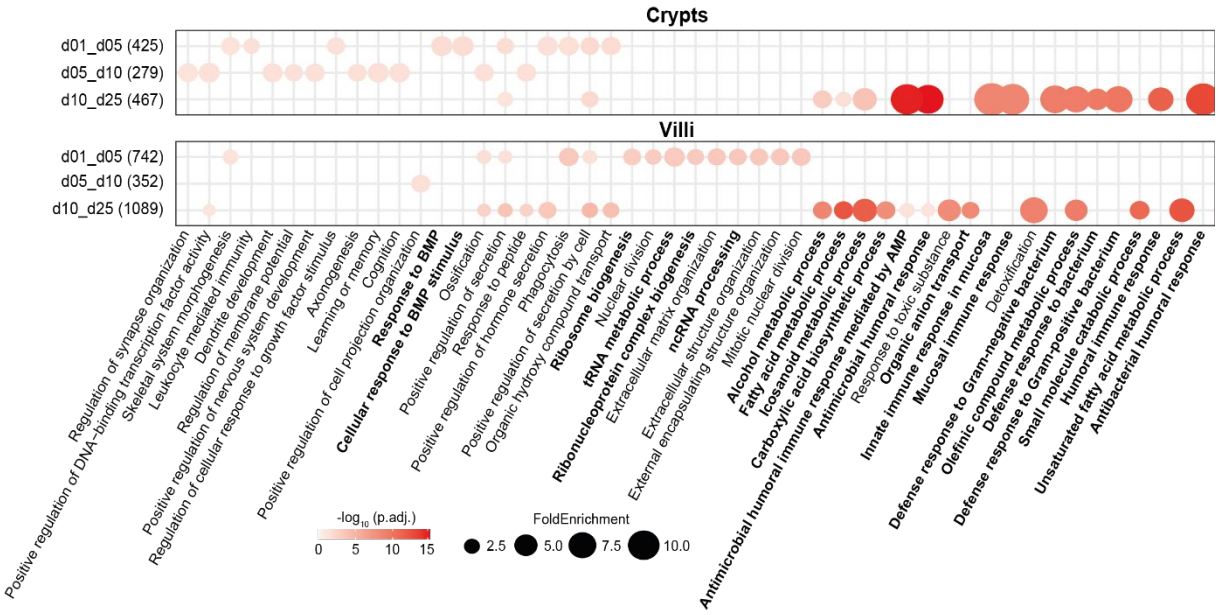


Figure 20: Gene Ontology analysis of differentially expressed upregulated over time in crypts and villi. Dot plot showing the top 10 enriched GO terms (by adjusted p-value) for the differentially expressed genes upregulated with age in crypts (top) and in villi (bottom). The number of DE genes associated with a GO term for each comparison is indicated in brackets.

8.4. Segment-specific transcriptional profiles along the proximal-to-distal axis

In contrast to the crypt-villus axis, which reflects vertical organization within intestinal units, the proximal-to-distal axis describes regional differences along the length of the small intestine. Segments of each part of the small intestine (proximal, medial, distal) were collected, and epithelial cells were isolated and sequenced (Figure 21a).

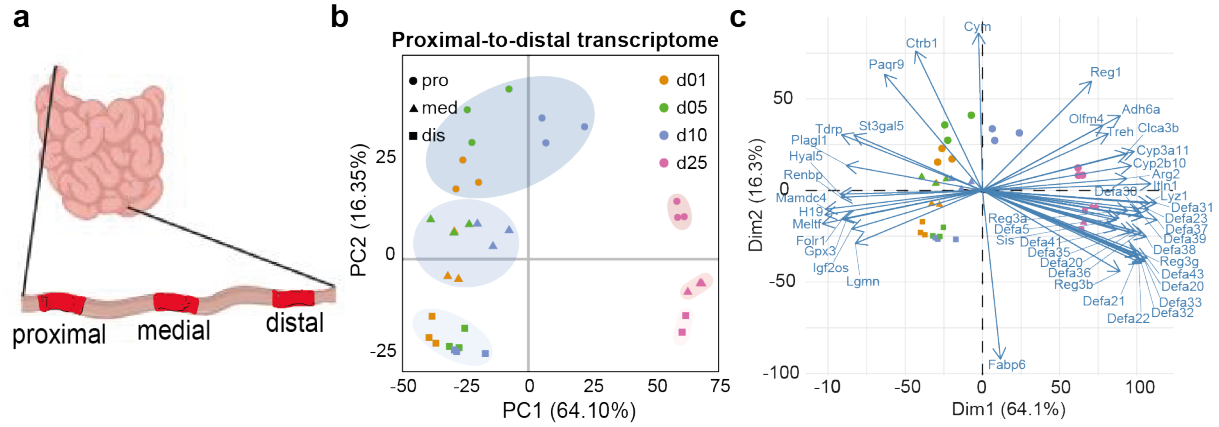


Figure 21: Principal component analysis (PCA) and biplot of samples from the proximal-to-distal axis. (a) Schematic representation of the small intestine showing the different samples collected. Figure created using Biorender.com. **(b)** PCA plot representing the epithelial transcriptome of either the proximal (circles), medial (triangles), or distal (squares) part of the small intestine of 1-day-old (orange, n = 3), 5-day-old (green, n = 3), 10-day-old (blue, n = 3), and 25-day-old (pink, n = 3) mice. Ellipses were estimated using the Khachiyan algorithm and show the distribution of the pre-weaning (proximal: light blue; medial: blue; distal: grey blue) and post-weaning (proximal: light pink; medial: pink; distal: dark pink) samples. **(c)** Biplot displaying the top 50 driver genes for the 1st and 2nd dimensions of the PCA plot shown in (b). 1-day-old (orange dots), 5-day-old (green dots), 10-day-old (blue dots), and 25-day-old (pink dots). Shapes correspond to the proximal (circles), medial (triangles), or distal (squares) part of the small intestine.

Interestingly, each anatomical segment separated from proximal to distal along the second principal component of the PCA (PC2, 16.4%); however, these segmental differences remained smaller than the pre- to post-weaning transition (PC1, 64.1%) (Figure 21b-c).

In the proximal region, genes involved in digestive enzyme production were upregulated between day 1 and day 5 (Figure 22, Supplementary Figure 4). These include pancreatic lipase-related protein 1 (*Pnliprp1*), serine proteases (*Prss1*, *Prss2*), carboxypeptidases (*Cpa1, Cpa2, Cpb1*), elastases (*Cela3a, Cela3b*), and carboxyl ester lipase (*Cel*). This indicates an early increase in transcripts associated with digestive functions in the proximal intestine. In addition, GO terms (Supplementary Figure 5), previously described to be associated with the early neonatal period (Figure 6) as “extracellular matrix organization” and “extracellular structure organization“, were also found, consistent with early remodeling and structural development of the tissue over the whole length of the small intestine.

In the medial region, notable genes upregulated at day 5 included aquaporin 8 (*Aqp8*), known to be involved in fluid and water transport in the intestinal epithelium; *Notum*, a regulator of Wnt signaling that modulates epithelial renewal (Kakugawa et al., 2015), and ephrin type-B receptor 2 (*Ephb2*), known for its role in maintaining crypt-villus organization (Figure 22) (Batlle et al., 2002).

Defensins and intelectin 1 (*Itln1*) were detected in the distal intestine as early as day 5. *Itln1*, which has previously been shown to be expressed in Paneth cells (Komiya et al., 1998), may act alongside defensins to support antimicrobial defense (Figure 22). Corresponding GO terms (Supplementary Figure 6), particularly those terms containing “defense response“, were driven by the defensin expression, painting a similar picture of early immune response. Across all regions, *H2-Q10*, heme oxygenase 1 (*Hmox1*), lymphocyte antigen 6 complex (*Ly6f*) and melanin-concentrating hormone receptor 1 (*Mchr1*) showed higher expression at day 1 compared to day 5 (Figure 22). These genes reflect epithelial and stress-response and immune pathways present in the early neonatal intestine and were also found to be highly regulated at day 1 compared to day 5 (*H2-Q10* and *Ly6f*) or in the pre-weaning (*Hmox1* and *Mchr1*) compared to the post-weaning samples of the global transcriptome (Figure 5) (Flanagan et al., 2008; Goodall et al., 2019; Kokkotou et al., 2008; Sebastián et al., 2018). Following the timeline, more defensins were found to be significantly expressed at day 10 compared to day 5, hinting at the development of Paneth cells in the medial part of the small intestine between both time points. Interestingly, no defensins were found to be differentially expressed at day 10 in the distal region and were only expressed with relatively small log2-fold changes ranging between values of 1 and 2 in the proximal and medial regions. On day 25, compared to day 10, defensins were found to be upregulated in all segments, showing now the complete development of the small intestine (Figure 22).

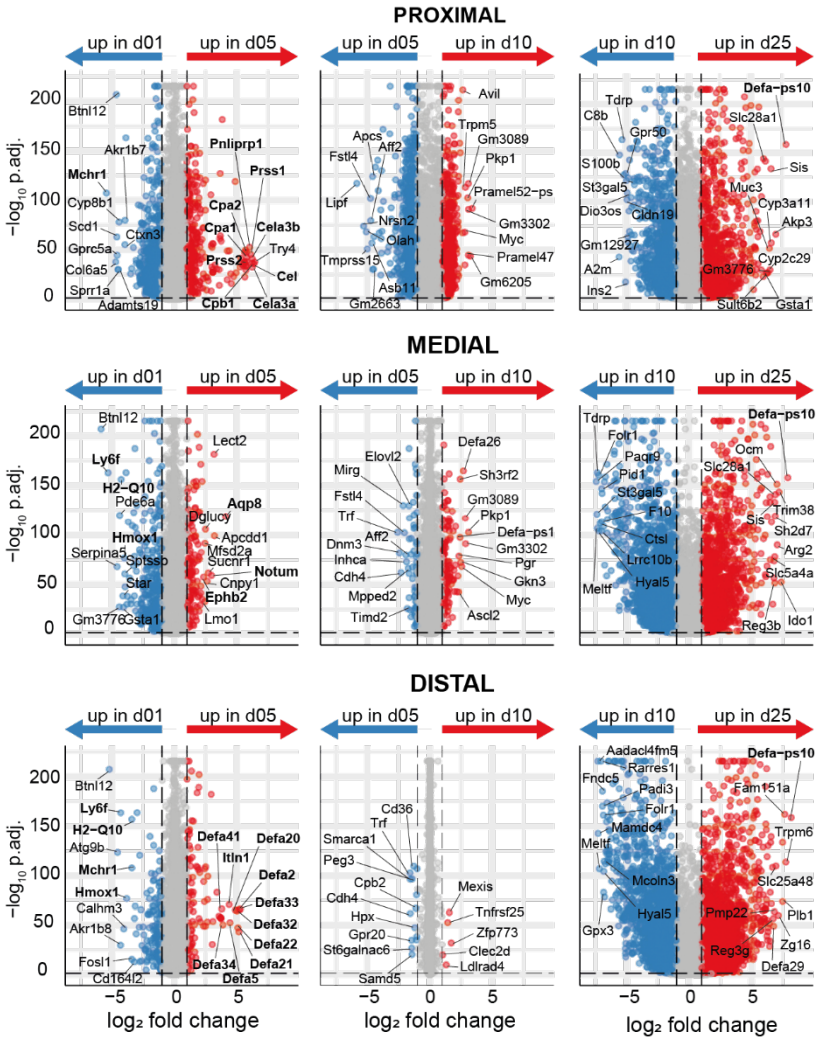


Figure 22: Differential gene expression between consecutive time points across the proximal-to-distal small intestine. Volcano plots showing differential gene expression between 1- and 5-day-old (left), 5- and 10-day-old (middle), and 10- and 25-day-old (right) proximal, medial, and distal samples (top to bottom). Blue and red dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in the younger and older animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes.

Spatiotemporal clustering of genes along the proximal-to-distal axis was performed to reduce data complexity and reveal coherent patterns of expression. This approach highlights region-specific programs and their changes during postnatal intestinal development. This hierarchical clustering of DE genes regulated either between the proximal and medial part or the medial and distal part at any age identified 5 major spatiotemporal expression patterns (Figure 23).

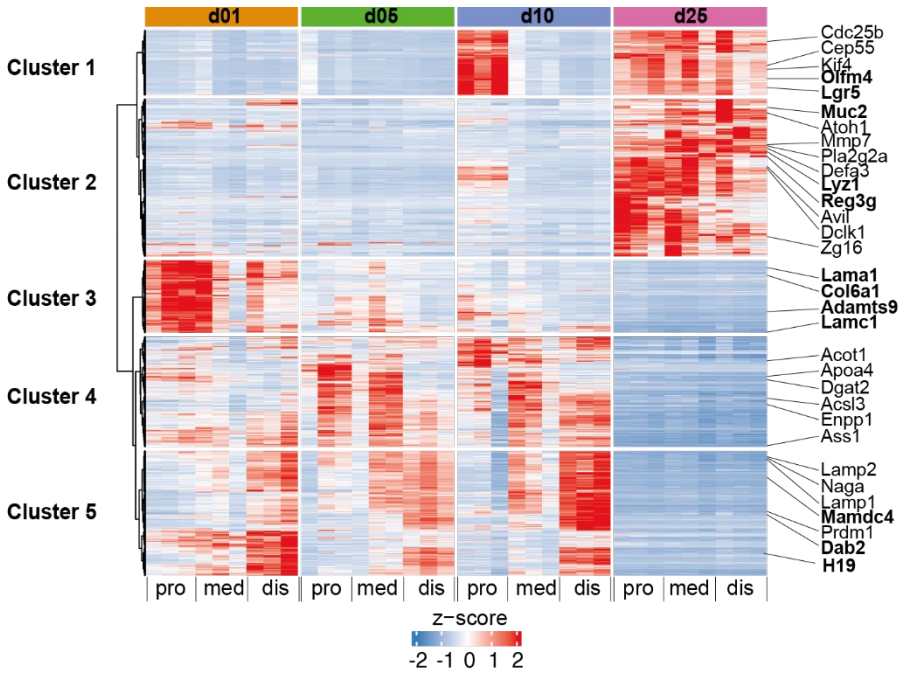


Figure 23: Spatiotemporal clustering of genes regulated along the proximal-to-distal axis. Clustered heatmap depicting the expression level of all genes differentially regulated between the proximal and medial part or the medial and distal part of 1-, 5-, 10-, and/or 25-day-old animals.

The first expression pattern represented by Cluster 1 was characterized by genes that were upregulated at day 25 throughout the entire length of the intestine, as well as in the proximal region at day 10. Genes included in this cluster such as *Olfm4* and *Lgr5* and GO terms related to cell division processes, including "sister chromatid segregation", "nuclear division", and "chromosome segregation", suggest a strong emphasis on cellular proliferation and division during this developmental phase (Figure 24).

Cluster 2 consisted of genes upregulated at day 25 throughout the entire length of the intestine and featured the goblet cell marker *Muc2*, as well as *Lyz1* and *Reg3g*, which are linked to Paneth cell function (Figure 23). The presence of GO terms reflecting immune response activities emphasizes the importance of maintaining intestinal homeostasis and defending against pathogens occurring after weaning (Figure 24).

Cluster 3 was characterized by genes that were upregulated at day 1 mainly in the proximal region, but to some extent also in the medial and distal regions. It contained genes such as laminin subunit alpha 1 (*Lama1*), collagen type VI alpha 1 chain (*Col6a1*), ADAM metallopeptidase with thrombospondin type 1 motif 9 (*Adamts9*), and laminin subunit gamma 1 (*Lamc1*) were significantly expressed (Figure 23). Those genes were involved in "collagen fibril organization" and various developmental processes. Other relevant GO terms included those related to "digestive system development", "extracellular matrix organization" and "external encapsulating structure organization", indicating that genes in this cluster are vital for structural integrity and proper functioning of the intestinal architecture especially in the proximal small intestine at day 1 (Figure 24).

Cluster 4 consisted of genes that were upregulated across all intestinal segments during the pre-weaning phase, but mainly at days 5 and 10. These genes were predominantly associated with metabolic processes and homeostasis, indicating their potential role in energy balance and nutrient absorption during a critical growth period.

Lastly, Cluster 5 contained expressions of distal genes prior to weaning. Interestingly, no GO terms related to immune response were identified within this cluster. This could suggest a focus on other physiological functions or adaptations specific to distal regions of the intestine during early development. The expression of the gene Disabled homolog 2 (*Dab2*) and enriched GO terms in cluster 5 such as “vacuole organization”, “lytic vacuole organization” or “lysosome organization” most likely represent lysosome-rich enterocytes (LREs), a functionally specialized subset of fetal-type enterocytes characterized by enhanced endocytic and lysosomal capacity (Block et al., 2025; Park et al., 2019) (Figure 23, Figure 24).

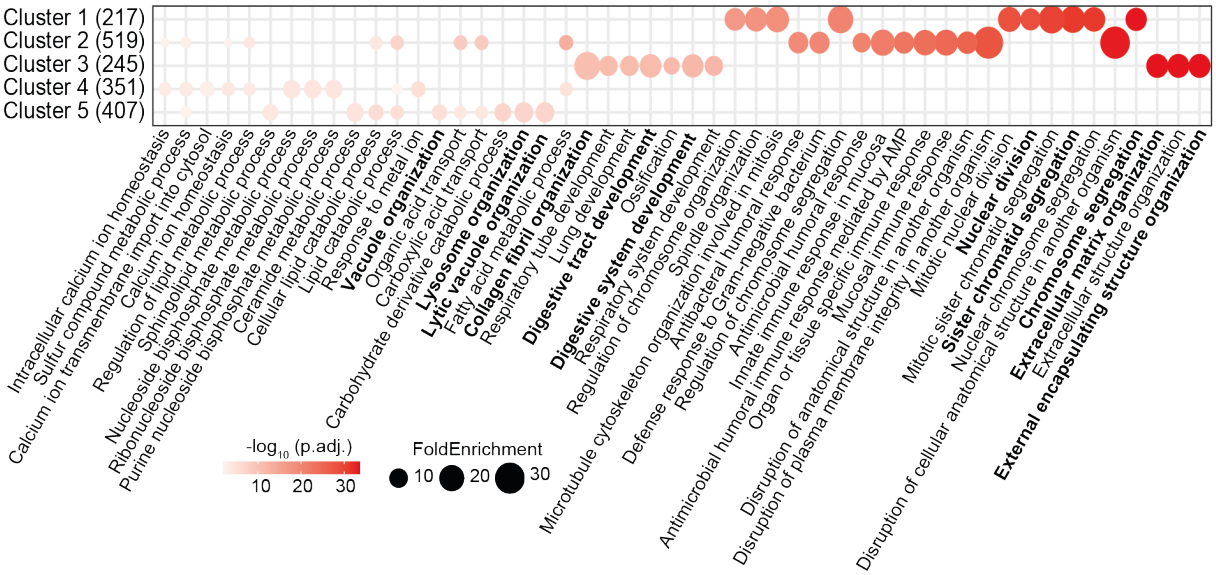


Figure 24: Gene Ontology analysis of the identified spatiotemporal clusters. Dot plot showing the top 10 enriched GO terms (by adjusted p-value) for the differentially expressed genes associated with each of the five clusters identified in Figure 23. The number of DE genes associated with a GO term for each comparison is indicated in brackets.

The analysis of differences along the proximal-to-distal axis revealed that while some developmental and metabolic programs are conserved across the small intestine, others are deployed in a region-specific manner, reflecting the functional specialization of proximal, medial, and distal segments. These insights offer a foundation for exploring the cellular basis of intestinal development.

8.5. Single-cell analysis reveals age- and cell-type-specific developmental trajectories

To analyze cell-type-specific genes, single-cell RNA sequencing (scRNA-seq) was performed on alive EpCAM+, CD45- epithelial cells isolated from 1-, 5-, 10-, and 25-day-old mice, as well as from *Salmonella*-infected animals (4 days post-infection/5-day-old). After quality control and nFeature and

nCount filtering and the filtering of mitochondrial genes (see chapter 7.2.1), a total of 9244 high-quality cells were further analyzed (Figure 25a+b). All major intestinal epithelial cell types, such as absorptive enterocytes, goblet and Paneth cells, enteroendocrine cells, stem cells, and tuft cells, could be identified using established marker genes (Figure 25c-d; Haber et al., 2017; Ruppert et al., 2020). Epithelial cells from day 25 mice exhibited a reduced overall viability, as only 661 cells were left after quality control on day 25. This loss in viability was a result of the challenging process to isolate and sort viable intestinal epithelial cells from crypts and villi, in the end leading to a severe underrepresentation of absorptive enterocytes in the day 25 samples and stem cells at time points of day 10 and day 25 (Supplementary Figure 7).

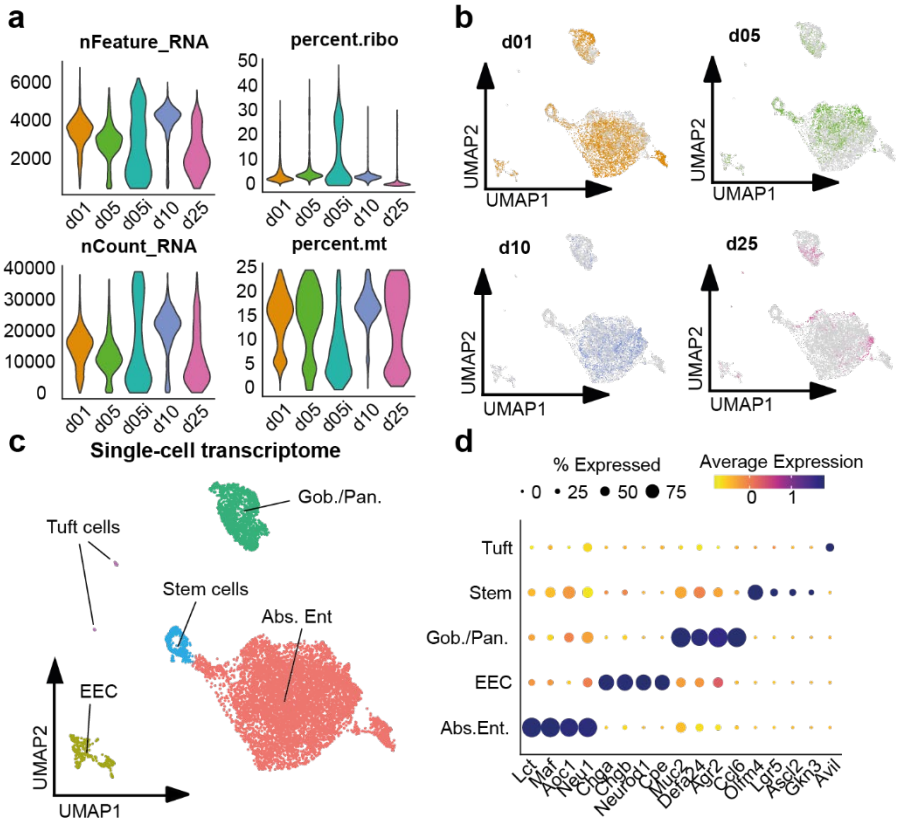


Figure 25: Quality control and clustering of the single-cell data. (a) Violin plots, showing the feature counts, gene counts, and percentage of ribosomal and mitochondrial genes after filtering of the scRNA-seq dataset. Uninfected 1-day-old, orange; uninfected 5-day-old, green; 5-day-old infected at birth with S. Typhimurium, turquoise; uninfected 10-day-old, blue; uninfected 25-day-old, pink. **(b)** UMAP plots displaying the cells recovered at each time point. Same color code as in (a). **(c)** UMAP displaying the different types of epithelial cells recovered from uninfected pre- and post-weaning murine small intestines: absorptive enterocytes (red), goblet/Paneth (green), stem (blue), enteroendocrine (ochre), and tuft (purple). **(d)** Dot plot displaying the marker genes for each cluster identified in (c).

The application of scRNA-seq allowed the investigation of cell-type-specific gene expression and transcription factor (TF) profiles, offering insight into regulatory mechanisms of cell differentiation during neonatal and early life stages that are not yet fully understood. DoRothEA, a curated collection of TF gene interactions, was employed to identify cell type and age-specific TF activity by the expression of their target genes (Badia-i-Mompel et al., 2022; Garcia-Alonso et al., 2019) (Supplementary Figure

8). This was followed by an analysis of the top 15 TFs with the strongest activity in each age group for absorptive enterocytes and goblet/Paneth cells (Figure 26a-b).

The analysis revealed three main distinctive patterns: first, a group of TFs that act independently of age (around 50% of the top TFs). Examples of this group were known and expected TFs *Hnf1a*, *Hnf4a*, and *Cebpe* for absorptive enterocytes or *Spdef*, *Tcf4*, and *Sox9* for goblet/Paneth cells. The second largest group encompassed TFs that acted only at the pre-weaning stage. Thus, TFs that have not yet been associated with cell differentiation in absorptive enterocytes, such as *Msx2*, *Esr2*, *Arid2,* and *Sox13,* were specifically associated at all ages with pre-weaning but not post-weaning absorptive enterocytes. The TF *Gata4* had previously been associated with absorptive enterocyte differentiation in adults but was here associated with goblet/Paneth cell development during the neonatal period (Earley et al., 2023), whereas *Smad3* had previously been described as an early goblet cell TF in other epithelial tissues (McCauley et al., 2014)

The last major group was a set of TFs specifically acting during the post-weaning stage. Examples of this group were *Cdx1* and *Spib,* which were associated with absorptive enterocytes and goblet/Paneth cells, respectively. When looking at the overall distribution of TFs in these subsets, several TFs specific to each age group were also identified (Supplementary Figure 9).

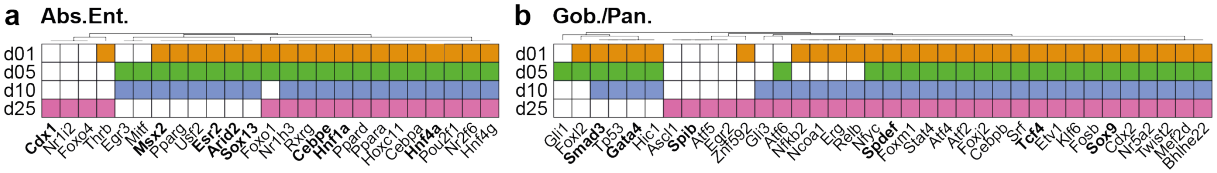


Figure 26: Occurrence heatmap of the top 15 transcription factors (TF) identified by DoRothEA. (a-b) Graphical representation of the top 15 TFs identified by DoRothEA for 1- (orange), 5- (green), 10- (blue), and 25-day-old (pink) absorptive enterocytes **(a)** and goblet/Paneth cells **(b)**. Colored cells depict transcription factors predicted by DoRothEA for each age group. White cells represent TFs that were not detected by DoRothEA.

To determine cell fate and track development over time, RNA velocity was performed using scVelo (Bergen et al., 2020). RNA velocity leverages the relative amounts of spliced and unspliced mRNA in each cell and, by comparing these differences across all cells, infers a latent transcriptional time. Based on this latent time, the transcriptional momentum between cells can be estimated. For each gene, a steady-state ratio between premature (unspliced) and mature (spliced) mRNA counts is fitted, representing a constant transcriptional state. RNA velocities are then calculated as residuals from this ratio. By combining velocities across genes, the future transcriptional state of an individual cell can be inferred. These velocities are subsequently projected onto a UMAP embedding and visualized as vectors (arrows) indicating the inferred developmental direction. This was performed during the ongoing preparation of the experiments shown in this thesis, therefore, a previous iteration of the UMAP depicted in Figure 25 was used to display latent time and velocity.

Analysis of the latent time (Figure 27a) revealed a low latent time, and therefore a starting point of velocity, at the upper right end of the UMAP (purple-blue area). When looking at the cell clusters and

age distributions (Figure 27b-c), scVelo predicted a start of the velocity for the absorptive enterocytes at day 25, developing from the middle of the UMAP towards a single cluster of day 1 absorptive enterocytes. A secondary trajectory within the absorptive enterocyte cluster could be observed, starting in the left middle part of the UMAP, just below the stem cell cluster. Here a more predominantly pre-weaning directional flow of arrows was observed starting at this point and going towards the right part of the UMAP.

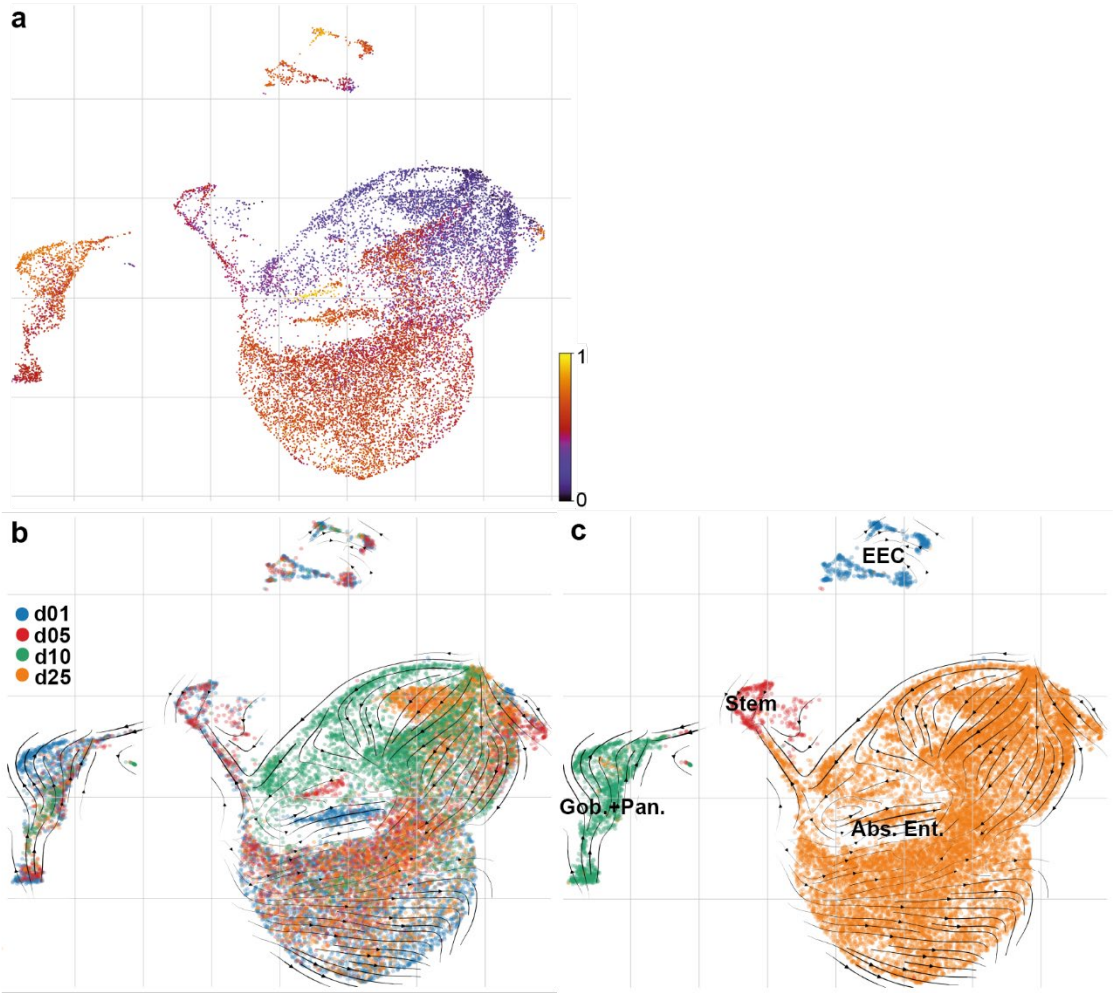


Figure 27: RNA velocity analysis of the single-cell data. (a) RNA velocity-inferred latent time visualized on a UMAP embedding for the single-cell data from all uninfected animals (1 to 25 days old). A previous iteration of the UMAP using different filtering thresholds is shown. Each point represents a single cell, colored by latent time from early (purple) to late (yellow) states. Latent time reflects the cell's progression along a dynamic trajectory, inferred from RNA velocity analysis. **(b)** RNA velocity analysis on the same UMAP plot as (a). Vector fields (arrows) represent the local average velocity calculated on a regular grid, illustrating the predicted future positions of the cells. 1-day-old (blue dots), 5-day-old (red dots), 10-day-old (green dots), and 25-day-old (orange dots). **(c)** Same UMAP as in (b), now colored according to cluster identity. Absorptive enterocyte cells are orange, EEC cells are blue, goblet/Paneth cells are green, and stem cells are red.

In the goblet/Paneth cluster, two directional flows of arrows were found. On one hand, two arrows come from the cluster of stem cells and go towards the goblet/Paneth cell cluster; on the other hand, a large group of the arrows start at the lower end of the goblet/Paneth cell cluster and go towards the top of the cluster. Noticeably, no clear directional flow of arrows could be observed in EEC or stem cells. For tuft cells, the number of cells was too low to give a meaningful interpretation of the vector

field generated here. Overall, and using prior biological knowledge, the RNA velocity analysis revealed surprising and incoherent results, and the analysis was therefore not pursued any further.

In the absence of interpretable RNA velocity results, cell fate dynamics of absorptive enterocytes, goblet/Paneth cells, and EEC were explored by manually analyzing subcluster relationships. Each of the three cell types was individually subclustered and reanalyzed. Furthermore, a spatial dimension was added to the new absorptive enterocyte cluster by overlapping gene expression signatures of genes previously associated with crypts and villi using bulk RNA-seq (Figure 17, Figure 28a).

In the cluster heatmap for absorptive enterocytes, the genes were ordered in a staircase pattern according to their clustering order (Figure 28a). However, many of the top 8 DE genes expressed most in one cluster compared against the other clusters were not specific to the cluster per se but to an age, as can be seen for d05 and d10 (Figure 28b).

Each age also contributed differently to the absorptive enterocyte subclusters. Thus, clusters 0, 3, 4, and 9 were mainly composed of cells isolated from day 1 animals; clusters 1 and 6 corresponded to cells isolated from day 5 animals, while cluster 5 was enriched for cells isolated from day 10 animals. Cluster 2 included a mix of cells from day 5, day 10, and day 25 (Supplementary Figure 10a).

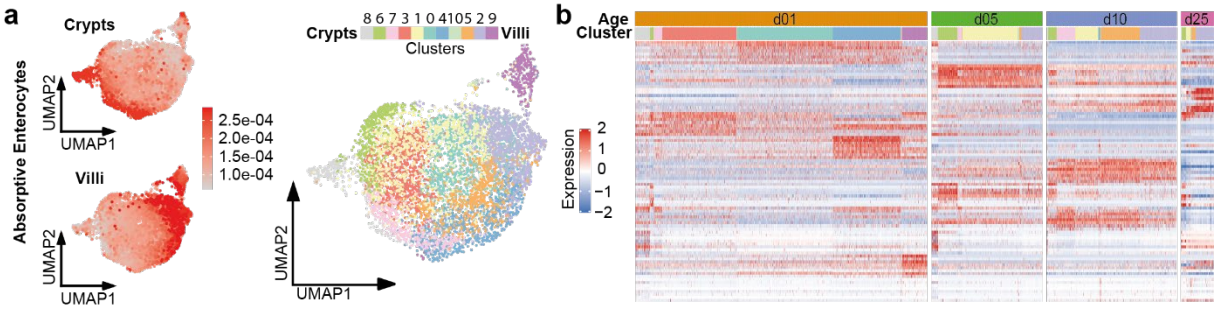


Figure 28: Subclustering of absorptive enterocytes. (a) UMAP plots showing the expression level of genes previously found to be associated with pre-weaned crypts (top left) and pre-weaned villi (bottom left) and the different subclusters recovered from the absorptive enterocyte cluster (right). The red color intensity of the left plots reflects the overall gene expression level. **(b)** Heatmap representing the top 8 DE genes per absorptive enterocyte subcluster. Age/cluster color code as in (a).

Similarly, the subcluster of goblet and Paneth cells was investigated, as well as the transitional shift between goblet cells, which were identified by the higher expression of *Zg16*, and Paneth cells, identified by the higher expression of *Lyz1* (Haber et al., 2017, Figure 29a). Here clusters 1, 3, and 4 were primarily associated with day 1, whereas cluster 2 was composed mainly of cells from day 25 (Supplementary Figure 10b, Figure 29b).

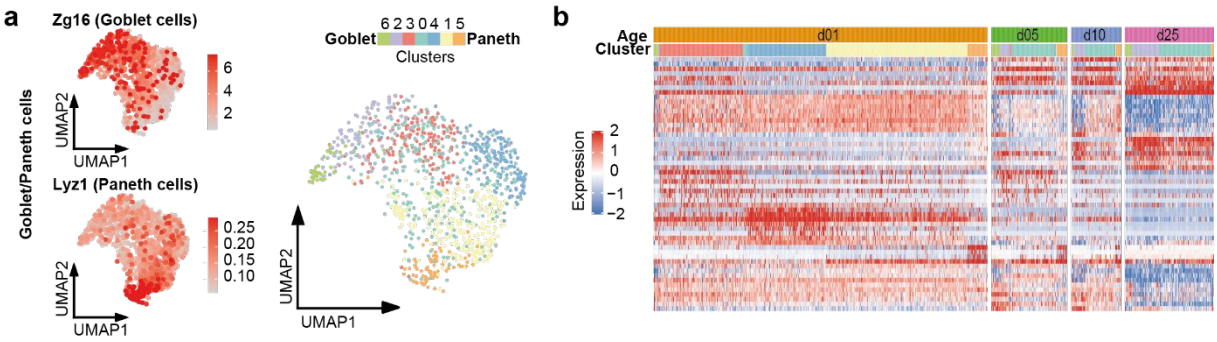


Figure 29: Subclustering of goblet/Paneth cells. (a) UMAP plots showing the expression level of known goblet (*Zg16*, top left) and Paneth cell (*Lyz1*, bottom left) markers and the different subclusters recovered from the goblet/Paneth cluster (right). The red color intensity of the left plots reflects the gene expression level. **(b)** Heatmap representing the top 8 DE genes per goblet/Paneth subcluster. Age/subcluster color code as in (a).

Upon investigating the GO analysis of the differentially expressed genes within the absorptive enterocytes (Figure 30) and goblet/Paneth cells (Figure 31) clusters, the previously observed trends, such as those of developmental changes associated with fatty acid and alcohol metabolic processes, could be observed once more. Furthermore, an increase in TGFβ activity was observed specifically in absorptive enterocytes between day 5 and day 10. This was evidenced by terms related to "response to TGF beta" and "TGF beta receptor signaling pathway" (Figure 30b).

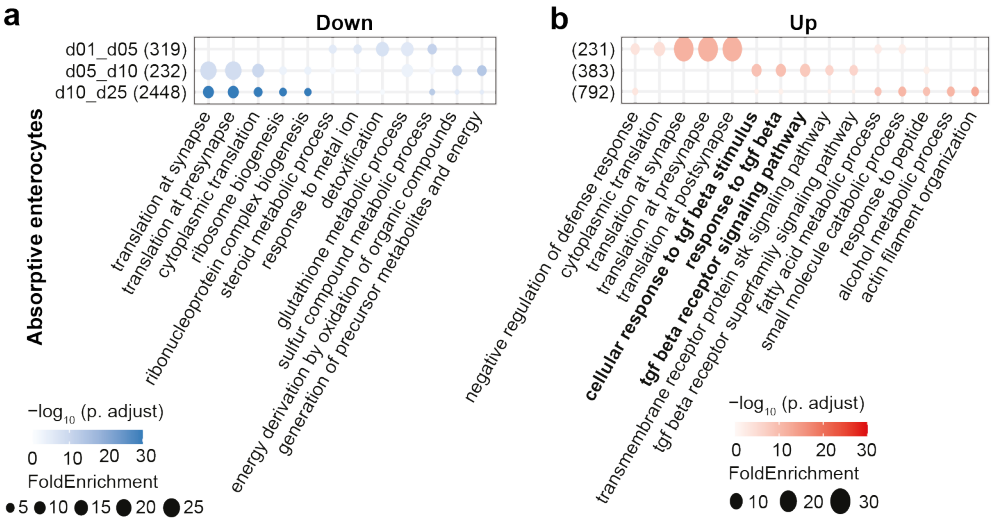


Figure 30: Gene Ontology analysis of the absorptive enterocytes subcluster. (a-b) Dot plot showing the top 5 enriched GO terms (by p.adj.) for the genes differentially downregulated **(a)** and upregulated **(b)** with age in absorptive enterocytes. The number of DE genes associated with a GO term for each comparison is indicated in brackets.

This increase was not detected via the pathway analysis based on the global transcriptomic or proteomic profiles at this timeframe, further emphasizing and highlighting the importance of performing cell type-specific analyses to capture these nuanced biological processes accurately (Figure 30, Figure 31, Figure 7, Figure 16).

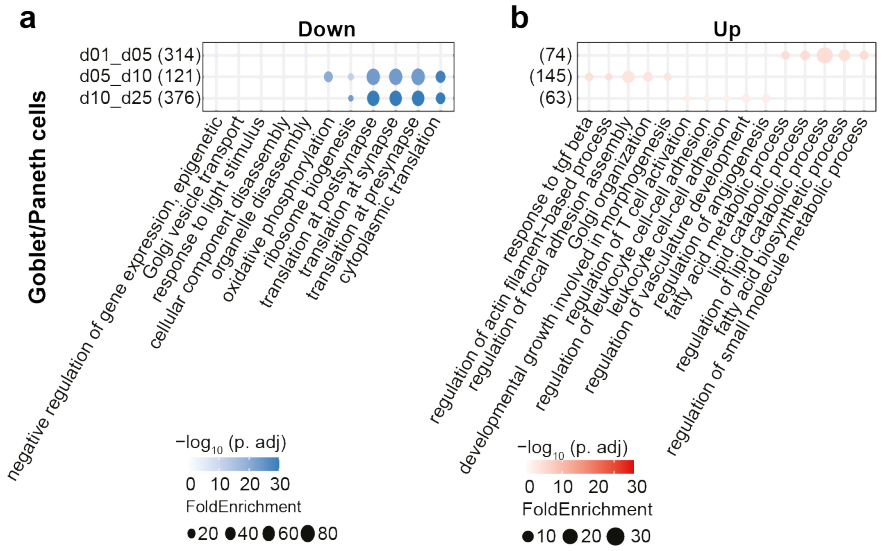


Figure 31: Gene Ontology analysis of goblet and Paneth cell subcluster. (a-b) Dot plot showing the top 5 enriched GO terms (by p.adj.) for the genes differentially downregulated **(a)** and upregulated **(b)** with age in goblet and Paneth cells. The number of DE genes associated with a GO term for each comparison is indicated in brackets.

Unlike the absorptive enterocytes and goblet/Paneth cells, all seven identified EEC subclusters (Figure 32a) matched previously described EEC subtypes based on known marker genes (Haber et al., 2017) (Figure 32b). These subsets were present across all ages with similar abundance (Supplementary Figure 11). The heatmap of the top 8 DE genes per subcluster and age showed distinctive patterns of markers for each age group (i.e., the “staircase pattern”), which suggests that EECs, in contrast to absorptive enterocytes and goblet/Paneth cells, develop similarly in neonatal and adult animals (Figure 32c).

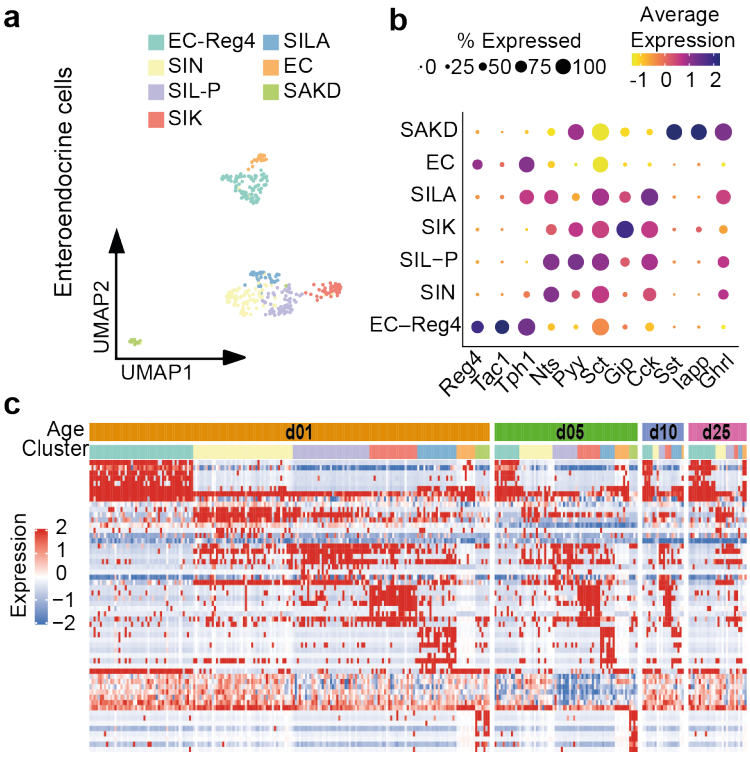


Figure 32: Subclustering of the EEC cells. (a) UMAP plots showing the different subclusters recovered from the EEC cluster. **(b)** Dot plot displaying the marker genes for each cluster identified in (a). **(c)** Heatmap representing the top 8 DE genes per EEC subcluster per age. Subcluster color code as in (a).

The usage of scRNA-seq revealed dynamic, cell type-specific changes during neonatal intestinal development, particularly in absorptive enterocytes and goblet/Paneth cells, while enteroendocrine cells remained relatively stable. The subclustering revealed age-dependent developmental trajectories and gene expression patterns that the bulk analyses failed to capture. These findings demonstrate the power of single-cell resolution to uncover the processes and developmental processes of the cells of the neonatal intestinal epithelium. To further explore how neonatal intestinal epithelial programs are altered under challenge, the impact of *Salmonella* infection on both bulk and single-cell transcriptional profiles was investigated.

8.6. Global and cell-type-specific responses upon infection with *Salmonella* Typhimurium

Epithelial cells collected at four days post infection from mice infected at day 1 after birth with 100-300 colony units (CFU) of *S.* Typhimurium (i.e., day 5 after birth) were also analyzed by RNA-seq and scRNA-seq (Zhang et al., 2014). First, the RNA-seq data obtained for the infected samples were mapped together with those previously analyzed from the global transcriptomic bulk RNA-seq (Figure 4) and batch correction was applied to adjust for technical variations between the infected and uninfected samples. PCA revealed that infection induced a strong transcriptional change (PC1 43.5%), that outweighed the age-induced change (PC2 34.6%) (Figure 33a). A subsequent DE gene analysis comparing infected day 5 animals against healthy age-matched controls (Figure 33b) revealed the expected upregulation of antimicrobial host response genes (e.g., *Reg3b*, *Reg3g*, *S100a8*, *S100a9*). In addition, infected epithelial cells upregulated immune regulatory and interferon-responsive genes such as indoleamine 2,3-dioxygenase 1 (*Ido1*), guanylate binding protein 2 (*Gbp2*), and *Cd74*, indicating a coordinated epithelial response of barrier defense and immune signaling.

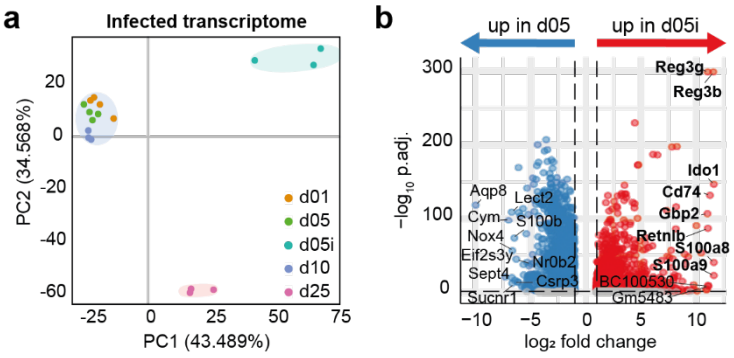


Figure 33: PCA plot and volcano plot of global transcriptomic data, including the *Salmonella*-infected data. (a) PCA plot representing the epithelial transcriptome of the small intestine of 1-day-old (orange, n = 4), 5-day-old (green, n = 4), 10-day-old (blue, n = 3), and 25-day-old (pink, n = 3) uninfected mice, as well as 5-day-old mice infected at birth with *S.* Typhimurium (turquoise, n = 3). Ellipses were estimated using the Khachiyan algorithm and show the distribution of the pre-weaning (light blue) and post-weaning (light pink) uninfected and 5-day-old infected (light turquoise) samples. **(b)** Volcano plots showing differential gene expression between 5-day-old *S.* Typhimurium-infected and uninfected samples. Blue and red dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in the uninfected and infected animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes.

At the single-cell level, UMAPs visualizing the cell distribution of the uninfected (d05) and infected datasets (d05i) showed a strong depletion of absorptive enterocytes in the infected sample (Figure 34a). This finding deviated from previous reports of intact tissue structure in infected animals (Zhang et al., 2014, 2018), suggesting that epithelial cells may have experienced increased stress during isolation and sorting, leading to reduced data quality and subsequent exclusion from the analysis.

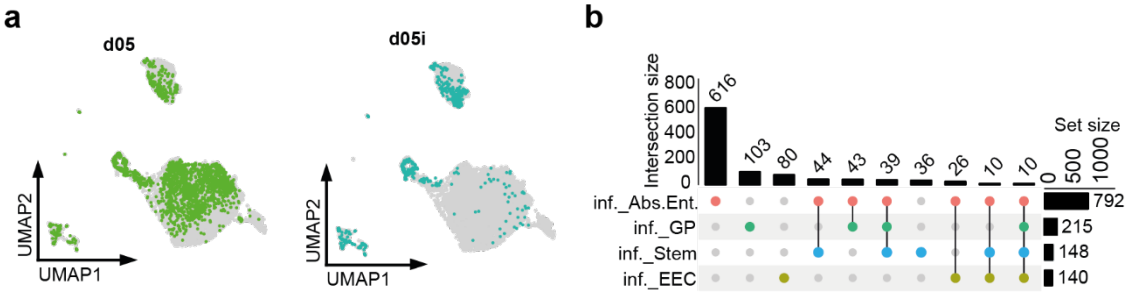


Figure 34: UMAP visualization of uninfected and infected cells and cell-type-specific differential gene expression upon infection. (a) UMAP plots displaying epithelial cells isolated from uninfected (left; green; same dataset as in Fig. 24b) and *Salmonella*-infected (right; turquoise) 5-day-old newborn mice. **(b)** Upset plot showing the number of genes differently regulated by one or more epithelial cell types upon infection with *S.* Typhimurium. Absorptive enterocytes (Abs. Ent.; red), goblet/Paneth cells (GP; green), stem cells (blue), and enteroendocrine cells (EEC; ochre). Only interactions with 10 or more DE genes are shown.

However, no such reduction in abundance was observed for the other cell type clusters, suggesting a specific lower resistance to apoptosis or stress-induced loss during sample processing of absorptive enterocytes (Supplementary Figure 7, Figure 34a). Cell-type-specific DE analysis of the epithelial response to infection revealed a common gene signature response by all cell types represented by genes such as *Reg3b* and *Reg3g*, as well as a unique gene response specific for each cell type (marked in purple) (Figure 34b, Figure 35). Notably, 42 genes made up the common response to infection across all four cell types, with 10 genes being upregulated and 32 genes being downregulated across all cell types in response to infection (Supplementary Figure 12). The uniform downregulation of genes such as *Fabp1*, *Fabp2*, and *Lct* indicates broad, infection-associated metabolic perturbations across the epithelium.

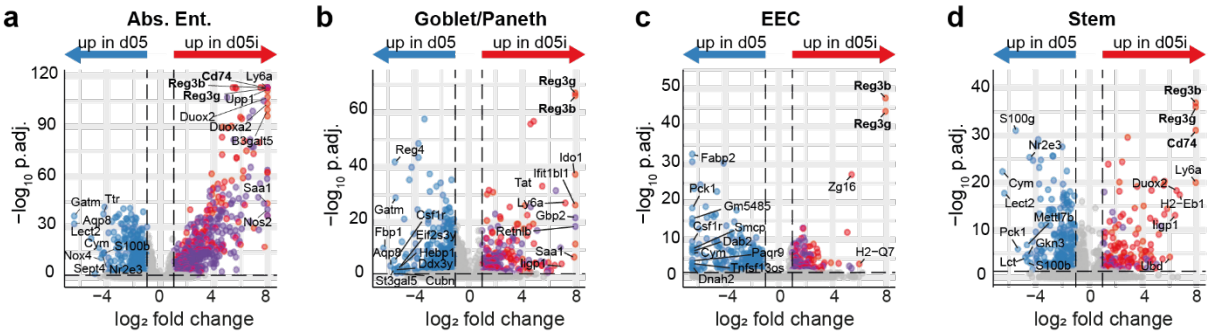


Figure 35: Volcano plots showing differential gene expression between 5-day-old uninfected and *Salmonella*-infected animals for each epithelial cell type. (a-d) Volcano plots showing differential gene expression between 5-day-old uninfected (d05) and *Salmonella*-infected (d05i) **(a)** absorptive enterocytes, **(b)** goblet/Paneth, **(c)** EECs, and **(d)** stem cells. Blue and red/purple dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in the uninfected and infected samples, respectively. Purple dots indicate genes uniquely differentially regulated upon infection by one cell type, whereas red dots indicate genes differentially regulated by multiple cell types. Grey dots indicate non-significant and/or non-differentially regulated genes.

Cell-type-specific GO analysis of unique infection-induced gene sets revealed distinct functional signatures across all epithelial lineages (Figure 36).

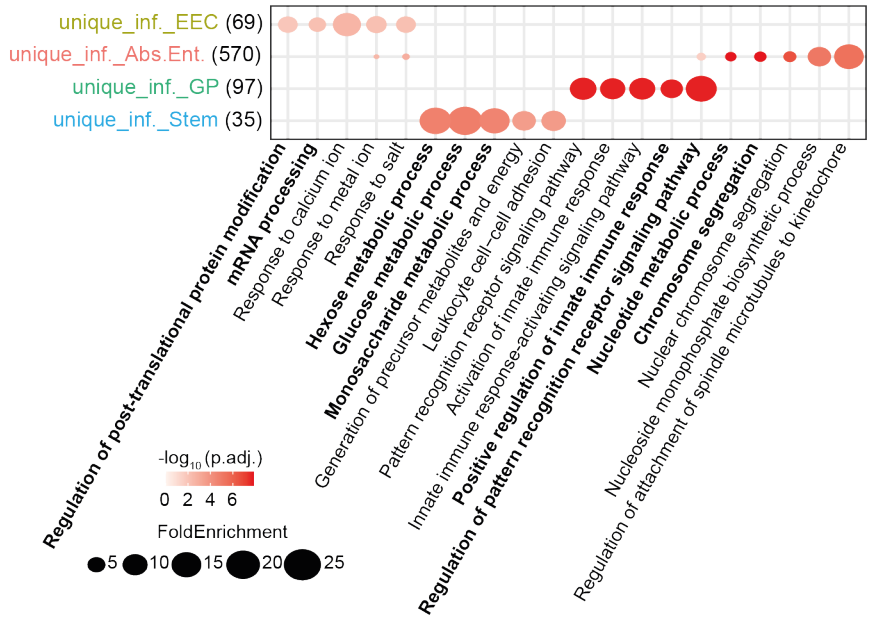


Figure 36: Gene Ontology analysis of the DE genes of different cell types corresponding to infection. Dot plot depicting the top 5 enriched GO terms (by adjusted p-value) for the infection-induced DE genes that were unique to each intestinal epithelial cell type. The number of DE genes associated with a GO term for each comparison is indicated in brackets.

In goblet and Paneth cells, upregulated GO terms included “Regulation of pattern recognition receptor signaling” and “positive regulation of innate immune responses”, reflecting their role in microbial sensing and early epithelial defense. Absorptive enterocytes showed enrichment for terms related to DNA replication and mitosis, such as “chromosome segregation” and “nucleotide metabolic process”, suggesting activation of proliferative and cell cycle programs in response to infection. In EECs, enriched pathways were associated with “mRNA processing”, “post-translational modification” and terms that can be grouped together as “ion-responsive signaling”, consistent with their secretory role and sensitivity to luminal cues. Finally, stem cells exhibited upregulation of metabolic pathways involved in glucose, hexose, and monosaccharide utilization (with glucose being a hexose and a monosaccharide), indicating a shift towards enhanced energy production and biosynthetic activity to support regeneration.

The analysis of *Salmonella* infection revealed a notable epithelial response distinct from maturation processes, with cell-type-specific transcriptional changes uncovering unique pathways and regulatory programs not apparent in the bulk analysis and therefore providing novel and unique insight into how individual intestinal epithelial lineages react to microbial challenges.

9. Discussion

9.1. Summary of key findings

The results of my thesis highlight the complex, heterogeneous, and dynamic nature of the murine small intestinal epithelium during the early postnatal period. This work is unique in its way, as it integrates multiple omics approaches. In fact, the bioinformatic analysis of bulk global transcriptomic and proteomic datasets, together with spatially restricted datasets (crypt-villus differentiation and regional differences along the proximal-to-distal axis) and single-cell transcriptomic datasets at multiple timepoints from birth to weaning, had never been performed with this intensity before. Previously published studies focused on either one or two of these techniques and usually focused on either the pre- or post-weaning period (Bara et al., 2022; Burclaff et al., 2022; Haber et al., 2017; Moor et al., 2018; Tsang et al., 2022; Zwick et al., 2024). The major conclusion from my analysis was that age, particularly weaning-associated processes, induced more pronounced transcriptional changes than regional differences (crypt/villus, proximal-to-distal axis). Many of the changes observed at weaning are consistent with the known roles of BLIMP1 (*Prdm1*) and MAFB, two key transcriptional regulators of the pre- to post-weaning transition. BLIMP1 is known to coordinate epithelial maturation and changes in nutrient absorption and barrier function, while MAFB contributes to differentiation programs in specific epithelial lineages (Bara et al., 2022; Harper et al., 2011; Muncan et al., 2011). The temporal patterns detected in this study align with these established functions.

Analysis of the global transcriptome and of the global proteome mostly converge and highlight the expected activation of pathways related to epithelial regeneration and antimicrobial defense around postnatal day 10, with a particularly pronounced signature by day 25, coinciding with crypt formation and the emergence of crypt-based Paneth cells, as reflected by the upregulation of canonical Paneth cell genes such as α-defensins and phospholipase A2 (Bry et al., 1994; Ménard et al., 2008; Sumigray et al., 2018). The previously described “weaning reaction”, which marks immune maturation in the mucosa and involves increased expression of pro-inflammatory cytokines, may contribute to the observed enhancement of epithelial signaling through JAK-STAT, MAPK, and PI3K pathways (Al Nabhani et al., 2019). Interestingly, divergences between the transcriptomic and proteomic analysis were detectable, with the major one occurring between day 1 and 5 after birth, with 1098 proteins but only 126 genes that were significantly upregulated during this time period. In part, this could be explained by the presence of milk proteins like WAP, which is ingested through breastfeeding and not synthesized by epithelial cells (see chapter 9.4.).

Another notable finding was the apparent acceleration of postnatal differentiation in the proximal small intestine compared to the medial and distal segments. A large proportion of genes associated with early postnatal tissue remodeling were already strongly expressed in cluster 3 at day 1 in the

proximal intestine, indicating an advanced maturation state shortly after birth. This spatial gradient was further supported by the PCA, where the proximal samples of day 10 clustered closer to their day 25 counterparts than the medial or distal samples. Together, these observations are consistent with a proximal-to-distal progression of intestinal maturation, highlighting that structural differentiation and functional specialization follow a spatially ordered developmental trajectory (Chin et al., 2017; Kolev & Kaestner, 2023).

9.2. Wnt signaling changes during neonatal intestinal development

The observation that both early upregulated and downregulated gene sets (UNN and DNN) were enriched for Wnt-related GO terms suggests that postnatal intestinal development is not characterized by uniform activation or repression of the Wnt pathway. Rather, these data support a model in which Wnt signaling is selectively modulated, with distinct pathway components and transcriptional outputs engaged at different developmental stages. In the adult intestine, canonical Wnt signaling regulates epithelial cell proliferation, secretory cell differentiation, and crypt-specific functions (Fevr et al., 2007; Mah et al., 2016; Noah et al., 2011; Pinto et al., 2003; Van Es et al., 2012). In contrast, in the neonatal intestine, Wnt signaling appears to be uncoupled from these adult crypt-associated outputs, reflecting a developmental shift in pathway function.

This is consistent with earlier research on neonates, which showed a perinatal window during which canonical Wnt signaling is mostly missing from intervillus regions but active in newly formed villi. Fetal intestinal villi express Wnt signaling ligands, whereas adult intestinal villi lack Wnt (Clevers, 2013; Van Es et al., 2012). This fetal Wnt activity localized in the villi is closely linked to epithelial differentiation and the expression of mature epithelial markers rather than to proliferative expansion. Consistent with this, late fetal villus cells exhibiting high Wnt activity do not express classical proliferative Wnt effectors such as *Tcf4*, *Cd44*, or Cyclin D1 (BCL1) (Kim et al., 2007). Instead, they preferentially express the related transcription factor 3 (*Tcf3*) and the alternative Wnt target Myc proto-oncogene protein (MYC), indicating that fetal villus Wnt signaling is coupled to differentiation-oriented transcriptional outputs rather than canonical stem-cell-associated programs. After birth, Wnt activity is only detectable in a small subset of villus structures, after which signaling rapidly redistributes from the villi to the intervillus regions within the first two postnatal days (Kim et al., 2007). This spatial shift precedes and accompanies crypt formation and subsequently supports the establishment of adult-like stem cell niches. In the neonatal intestine, where epithelial architecture and functional boundaries are still emerging, this pattern reflects a developmental relocation of Wnt activity from villus-dominant early programs toward crypt-centered regulatory control characteristic of the mature intestine (Kim et al., 2007). This context-dependent behavior of Wnt signaling supports the idea that the observed transcriptional changes reflect developmental rewiring of pathway outputs after birth.

9.3. Age-dependent programs override microbiota influences

Despite earlier research demonstrating an impact of intestinal microbiota on the intestinal epithelial expression of host defense and glycosylation-associated genes in adult animals, only minimal differences were observed between SPF and GF mice during the pre-weaning period (Tsang et al., 2022). While the gut microbiota clearly impacts postnatal development, developmental transitions are also largely driven by endogenous programs, including epigenetic mechanisms such as DNA methylation, that operate independently of microbial cues (Pan et al., 2018). These programs can mask microbiota-dependent transcriptional changes that are prominent in adults. As a result, age-related differences can appear much stronger than microbiota-related differences at this stage. In addition, maternal influences during gestation and lactation profoundly shape neonatal physiology and immune development. In fact, in addition to nutritional factors, neonates receive abundant immune factors, which dampen host immune activation, such as secretory IgA and IgG, through maternal milk. These antibodies and accompanying immunomodulatory molecules (e.g., cytokines and growth factors) promote mucosal homeostasis and structural maturation of the intestine and thereby effectively reduce the need for microbiota-driven epithelial defense programs, like the ones prominent in adults (Rogier et al., 2014; Shenoy et al., 2025).

Dietary protein absorption in neonates depends on a specialized epithelial population located predominantly in the ileal region of the small intestine, historically described as vacuolated or fetal-type enterocytes and more recently referred to as lysosome-rich enterocytes (LREs) (Park et al., 2019; Skrzypek et al., 2007). These cells are characterized by pronounced endocytic activity and are adapted for the uptake of luminal macromolecules, which are subsequently degraded within enlarged lysosomal vacuoles (Fujita et al., 2007; Gonnella et al., 1987; Park et al., 2019). In mammals, LREs are largely restricted to the suckling period and are progressively lost until weaning (Harper et al., 2011). In line with these reports, GO terms associated with this fetal-type enterocyte program (e.g. “vacuole organization”, “lytic vacuole organization” or “lysosome organization”) were most prominently enriched in the distal small intestine across all pre-weaning animals, supporting a regionally delayed maturation pattern (Gonnella et al., 1987; Park et al., 2019).

The enrichment of Disabled homolog 2 (*Dab2*) in this region also supports a functional interpretation, as DAB2 is involved in the endocytosis of milk-derived macromolecules and co-localizes with megalin (LRP2), cubilin (*Cubn*), and amnionless (*Amn*) at the microvilli and within the apical endocytic apparatus of ileal enterocytes (Park et al., 2019; Vázquez-Carretero et al., 2014). Conditional deletion of *Dab2* affects protein but not antibody transcytosis, indicating the existence of separate uptake mechanisms for nutritional and immunological cargos (Block et al., 2025).

9.4. Complementary roles of transcriptomics and proteomics

Bulk RNA-seq quantifies transcript levels, indicating transcriptional activity, whereas mass-spectrometry-based proteomics directly measures protein abundance. Even though RNA-seq provides high sensitivity and a broad dynamic range, mRNA levels do not necessarily predict protein abundance due to post-transcriptional regulation, translation efficiency, and protein stability. On the other hand, label-free proteomics captures the regulation beyond transcription but is biased toward abundant and stable proteins and has a lower coverage of low-abundance regulators. Therefore, integrating both datasets enables transcriptional changes to be interpreted considering their corresponding proteomic changes and reveals regulatory layers that would be missed by either approach alone.

Of the 4,562 proteins detected by proteomics, most of them correlated well with the transcriptome, and only 402 lacked corresponding transcripts. Of these, 135 could be mapped to genes expressed in luminal progenitor and mature luminal cells of the mammary gland, as reported in the well-curated *Tabula Muris* single-cell dataset, highlighting their origin from breast milk. The neonatal gastrointestinal tract is specifically adapted to preserve bioactive milk components while meeting the high nutritional demands in the early days of life. In fact, because the stomach is not yet fully acidified, proteolysis is limited (Heinrichs et al., unpublished), and milk proteins are able to reach the intestinal lumen largely intact, potentially contributing to early proteomic signatures (Chatterton et al., 2004; Neal-Kluever et al., 2019). Among the 267 remaining proteins, 147 corresponded to the same LFQ intensity as at least one other protein, preventing unambiguous matching to a single transcript for technical reasons. Another 35 proteins had detectable transcripts in the raw RNA-seq data, but their transcriptomic counterparts fell below the initial filtering threshold due to low expression. The remaining 82 proteins may reflect enhanced autophagic activity in the neonatal epithelium, during which cells recycle intracellular components to generate energy and nutrients critical for growth and function. Together with the uptake of milk-derived proteins, which supply external nutrients and passive immune factors, autophagy likely supports the neonate in overcoming the energy deficit immediately after birth, prior to the establishment of regular feeding (Kuma et al., 2004; Li et al., 2022).

9.5. Early manifestation of immune-related epithelial programs in the distal small intestine

Although morphologically distinct Paneth cells are generally reported to develop by day 7 after birth (Bry et al., 1994), defensin transcripts, usually Paneth cell-associated, were already detectable in the distal transcriptome at day 5 and were likewise visible in the global transcriptome at day 5. This apparent inversion suggests that innate immune response and epithelial differentiation are at least partially uncoupled processes, despite an overall proximal-to-distal structural maturation and an

earlier antimicrobial transcriptional activity distally. Notably, prior observations that goblet cells are also capable of producing antimicrobial peptides (Bry et al., 1994) suggest that these early distal defensin signals may not solely reflect Paneth cells but rather reflect a broader, stage-specific, epithelial contribution to innate defense during early postnatal development. Because both goblet and Paneth cells derive from the same secretory lineage, these patterns may also reflect intermediate, or hybrid, epithelial cell states present during early life (Bry et al., 1994; Lueschow & McElroy, 2020; Shroyer et al., 2007)

A potential explanation for this regional bias may lie in differences in microbial composition and functional exposure along the intestinal axis rather than absolute microbial richness. While overall bacterial diversity in the small intestine is low immediately after birth and increases only from approximately postnatal day 7 onwards, with significantly higher bacterial richness becoming evident in the colon primarily after weaning (Van Best et al., 2020), the distal small intestine is nevertheless positioned closer to the colon and may therefore encounter microbial signals or metabolites earlier or more consistently than proximal segments. Under such conditions, antimicrobial transcriptional priming could occur in immature secretory lineages prior to full morphological differentiation, thereby establishing a functional defensive gradient toward the colon, where microbial density and diversity subsequently increase further (Santaolalla et al., 2011).

9.6. Early spatial positional signals shape neonatal epithelial differentiation

The utilization of LCM also bridged bulk and single-cell approaches by capturing spatially defined epithelial regions and successfully allowed the overlay of location-specific transcriptional identity onto single-cell data. In the adult small intestine, Moor et al. used a similar strategy to map transcriptional zonation of enterocytes along the crypt-villus axis, revealing well-defined spatial programs: cells at the villus base expressed antibacterial and metabolic genes, mid-villus enterocytes specialized in nutrient absorption, and villus-tip cells exhibited signatures linked to immune modulation and adhesion (Moor et al., 2018). In contrast, the current data capture full crypts and full villi during neonatal development, a period when these spatial programs are detectable but not yet fully established. At early postnatal ages, crypt-villus architecture is still immature, and spatial functional boundaries are still forming. Although LCM was aimed at specifically microdissecting the superficial epithelial layer, transcripts from adjacent stromal (e.g., endothelin *Edn3*) and immune cells (e.g., *Igha*, *Igkc*) were also detected, reflecting the close proximity and molecular interplay within the epithelial microenvironment. Nevertheless, LCM performed robustly overall and revealed the expected transcriptional differences between spatial compartments, supporting its suitability for region-resolved epithelial profiling.

Although they lacked the sharp, adult-like structure of mature intestines, crypt-villus transcriptional gradients were already detectable even at the earliest postnatal stages, as illustrated by the early enrichment of BMP-associated signaling in villus fractions and extracellular organization programs in crypt regions. This indicates that molecular programs underlying villus zonation are initiated very early, before morphological maturation is complete. Importantly, the early gradients observed here are more strongly associated with developmental signaling pathways, such as BMP (e.g. GO terms “Response to BMP” and “Cellular response to BMP stimulus” up in day 5 villi) and with age-dependent transcriptional programs (e.g. GO terms “extracellular structure organization”, “extracellular matrix organization” and “cell-substrate adhesion” in the crypts of day 10 compared to those of day 25) than with the terminal metabolic specialization described in adults. Together, these findings extend the concept of enterocyte zonation into a developmental context, indicating that spatial patterning is initiated in parallel with epithelial maturation.

9.7. Neonatal foundations of adult intestinal domain architecture

The comprehensive analysis of epithelial transcriptomes along the proximal-to-distal axis revealed that spatiotemporal programs of epithelial regionalization emerge early in murine development and become progressively refined with age. Notably, while in the PCA the pre-to-post-weaning transition (PC1) accounted for the majority of transcriptional variance, segmentation along the proximal-to-distal axis (PC2) was evident even in the pre-weaning period.

Functional studies in neonatal rodents further support the presence of region-specific epithelial roles early in life. Jejunal enterocytes have been described as being specialized for transcytosis of intact proteins, whereas ileal enterocytes preferentially engage lysosomal degradation pathways. These functional differences are associated with distinct endolysosomal gene-expression programs and indicate regional specialization for nutrient handling well before full maturation is reached (Block et al., 2025; Gonnella et al., 1987). Consistently, the spatiotemporal clustering analysis revealed remnants of LREs, indicated by *Dab2* expression and GO terms related to “lysosome organization” and similar processes, underscoring early regional differentiation within the epithelium.

These results contrast with studies on human fetuses (first and second trimesters) and newborn children, where samples from different regions strongly clustered by age rather than by region, indicating that the cells shared similar properties across location and only later in adults showed signs of regional differences (Gu et al., 2024). Human datasets, however, often rely on limited prenatal or neonatal material, which can potentially mask subtle spatial gradients. Preterm and fetal human cohorts also typically include substantial biological and clinical heterogeneity in gestational week differences, type of medical interventions, nutrition, hypoxia, and inflammation (Eidem et al., 2015). These factors strongly influence epithelial transcription and can easily outweigh spatial differences. In

contrast, laboratory mice are genetically homogeneous and environmentally standardized, which enhances detectability of regional effects. Furthermore, precise protocols and a higher sample rate may enhance the detectability of regional differences. Thus, part of the discrepancy between human fetuses and neonatal mice may be technical rather than biological. Nevertheless, the findings of this thesis are in line with observations of epithelial zonation in the adult murine and human small intestine, where five discrete metabolic and functional domains have been identified along the proximal-to-distal axis (Zwick et al., 2024), suggesting that elements of this domain architecture may already be transcriptionally prefigured during early postnatal development.

In adult mice, the domains A and B were most strongly associated with the fatty acid metabolism, with domain A also being associated with the uptake of iron. Domain C was defined by the lack of expression of genes found in other domains but also associated with carbohydrate metabolism, domain D with chylomicron and lipoprotein metabolism as well as amino acid transport, and domain E with cholesterol and steroid metabolism (Zwick et al., 2024). While the published work primarily addresses the mature intestine and adult zonation signatures, the developmental data presented in my study indicate that transcriptional hallmarks of regional specialization are already detectable prior to the establishment of fully defined adult domains. For instance, the early proximal enrichment of digestive programs (GO term “digestive system development”) corresponds to the lipid-absorptive and carbohydrate-metabolic domains A, B and C described by Zwick et al., whereas distal immune and defense-associated signatures relate to adult regions D and partially E, which are defined by genes enriched for bile salt recycling and microbe-associated functions. Importantly, unlike the sharply defined and metabolically specialized domains of the adult intestine, neonatal zonation appears more gradual and closely intertwined with overall maturation. This is also evident in my dataset: PCA of the proximal-to-distal data shows that age accounts for far more transcriptional variation than anatomical position, and spatiotemporal clustering identifies gene modules that are co-regulated along the entire intestinal length rather than confined to individual segments.

Ultimately, achieving a comprehensive understanding of how the mammalian small intestine transitions from the neonatal state into a finely zonated and segmented organ specialized for diverse absorptive, metabolic, and immunological functions remains challenging, as multiple overlapping processes contribute to this transformation (Block et al., 2025; Kim et al., 2007; Moor et al., 2018; Zwick et al., 2024). Nevertheless, the present work provides a first insight into this critical developmental window.

9.8. Cell type-specific neonatal epithelial maturation by transcriptional networks

Bulk transcriptomic approaches provide robust global trends but inherently average signals across diverse epithelial populations and therefore cannot resolve spatial gradients or cell-type-specific programs along the crypt-villus axis, which are particularly relevant in a rapidly adapting tissue such as the neonatal intestine. Single-cell transcriptomic analyses partially overcome this limitation by enabling cell-type-resolved insights and the investigation of differentiation trajectories.

While RNA velocity was applied to infer lineage relationships, overlapping transcriptional programs and rapid developmental transitions limited the robustness of trajectory predictions in this dataset. Manual cell fate inference was based on established marker expression and therefore provided a complementary strategy, as the availability of LCM data from crypt and villus areas and the expression of known Goblet/Paneth markers *Zg16* and *Lyz1* offered an independent reference that allowed for a spatial orientation of the UMAP and supported the interpretation of differentiation patterns where the RNA velocity estimates remained inconclusive.

The observed patterns of transcription factor activity in absorptive and secretory epithelial cells indicate that neonatal intestinal maturation is controlled by a finely tuned regulatory hierarchy, in which age-specific TFs direct the timing of differentiation and functional specialization. Pre-weaning TFs may coordinate adaptation to a milk-based diet and the establishment of early barrier functions, whereas post-weaning TFs contribute to the emergence of adult-like metabolic and secretory programs. The presence of conserved, age-independent TFs across lineages highlights core regulatory circuits that maintain lineage identity throughout development, ensuring coordinated maturation despite changing environmental and dietary conditions.

The expression of *Gata4* and *Smad3* in pre-weaning goblet and Paneth cells likely reflects the ongoing coordination of regional and lineage-specific maturation in the neonatal intestine. *Gata4*, while classically associated with proximal absorptive enterocytes and repression of distal programs (Aronson et al., 2014; Bosse et al., 2006), may act more broadly in early secretory cells to prime gene programs and chromatin states, ensuring that these cells acquire region-appropriate features while maintaining developmental plasticity. Rather than enforcing strict distal repression, its early activity may establish a gradient of influence, allowing spatial patterning to emerge alongside functional differentiation. Similarly, *Smad3*, a key mediator of TGF-β signaling, could support the balance between differentiation and functional specialization, priming goblet and Paneth cells for early barrier function and innate defense (Piek et al., 2001). Interestingly, observations from the ocular epithelium demonstrate a *Smad3*-dependent regulation of SPDEF, a master regulator of goblet cell differentiation, suggesting

that a comparable regulatory logic may guide related epithelial cell-fate decisions across multiple tissues (McCauley et al., 2014).

By contrast, the relative stability of transcriptional programs in EEC cells across the different ages examined suggests that endocrine hormone specification is largely established early and maintained throughout life, reflecting the essential role of gut hormone signaling in nutrient sensing and digestive physiology.

9.9. Neonatal epithelial adaptation to *Salmonella* infection

While investigating the infected samples, it became evident that infection exerted a stronger influence on transcriptional activities than age, as PC1 of the PCA was primarily driven by the infected samples. This observation is in line with previous findings from my group that suggested an accelerated intestinal epithelial maturation following neonatal enteric infection (Schlößer et al., 2025). In this work to which I contributed, neonatal infection with *Salmonella* Typhimurium was shown to increase epithelial proliferation and induce features of premature epithelial maturation as early as 4 days post-infection (day 5 of life). Gene set enrichment analysis revealed an overrepresentation of *Salmonella*-induced genes within a gene set characteristic of the adult small intestinal epithelium, a finding validated by PCR for selected metabolic, antimicrobial, and developmental markers (e.g., *Sis*, *Lyz1* and *Spdef*). Notably, this infection-associated maturation occurred without the canonical downregulation of the negative master regulator Blimp1, which normally governs the suckling-to-weaning transition under homeostatic conditions. This interpretation is consistent with the PCA observations, in which PC1 captures transcriptional changes driven by infection and PC2 reflects the underlying maturation processes associated with post-weaning and aging. Whereas this previous analysis was limited to a small set of marker genes, the present study provides a more comprehensive view of the epithelial response, allowing us to resolve both global and cell-type-specific adaptations.

Beyond a global defense signature induced by *Salmonella* infection, distinct cell type-specific responses became apparent. Goblet and Paneth cells showed increased activity, accompanied by upregulation of resistin-like beta (*Retnlb*), consistent with enhanced mucus secretion and antimicrobial peptide production and potentially reflective of a strong response to bacterial infection (Bergstrom et al., 2015). In contrast, enterocytes adopted a more proliferative phenotype, likely supporting epithelial restitution following infection-induced damage, which comes at the expense of nutrient absorption. As nutrient intake during the neonatal period relies almost exclusively on maternal milk, *Salmonella*-induced disturbances in digestive capacity at this stage may have critical consequences for growth and disease progression and overall outcome. Metabolic complications, including transient lactose intolerance, have been reported in young children following non-typhoidal *Salmonella* infection (Ren et al. 2018), and reduced lactase activity has also previously been observed in infected murine

neonates (Schlößer et al., 2025). Consistently, my work shows significant downregulation of *Lct* across epithelial cell types, indicating impaired lactose digestion during infection.

Infection amplifies pre-existing gradients of regional differentiation and immune readiness, prioritizing barrier reinforcement and antimicrobial defense over homeostatic processes such as nutrient absorption. At this early stage, when the epithelium and immune system are still maturing, microbial and pathogen-derived signals can strongly shape gene expression and cellular behavior. In the context of *Salmonella* infection, this heightened responsiveness drives premature epithelial maturation, including increased proliferation, enhanced antimicrobial peptide production, and metabolic adaptation, reflecting the unique plasticity of the neonatal intestinal epithelium (Schlößer et al., 2025). Together, these findings highlight how early-life infection can reorganize developmental programs to rapidly bolster defense, even at the cost of digestive capacity and nutrient uptake.

9.10. Technical considerations and limitations of the study

Several challenges exist when isolating epithelial cells. First, detachment from the basal membrane, which is required for epithelial cell isolation, leads to a rapid reduction in cell viability. While this effect is largely negligible when isolating epithelial sheets for bulk RNA sequencing and proteomic analyses, it represents a major limitation during single-cell isolation, where extended dissociation procedures substantially compromise cell survival. This loss of viability may be further aggravated by tissue inflammation, such as that occurring during infection; therefore, it is critical to reduce the time between cell isolation and processing. The protocols used in this study were established in 2019 and 2020, and the field of scRNA-seq has since rapidly evolved. Subsequent methodological advances and updated best practices have enabled the sequencing of substantially larger numbers of cells with increased sensitivity, improving the detection of rare or short-lived transient cell states and allowing deeper insights into differentiation processes (Heumos et al., 2023; Ortolano, 2024).

Second, the highly structured anatomical organization of the intestinal epithelium complicates the simultaneous isolation of representative and viable cell populations from both villus and crypt compartments across different developmental stages. Intestinal enterocytes have a high mortality rate (via anoikis) of post-mitotic differentiated cells, and the protocol used in this thesis was tailored to neonatal tissue, whereas many protocols developed for adult animals are designed to enrich for crypt-associated cells through selective crypt release (Haber et al., 2017). While such approaches are appropriate for studying stem and progenitor populations in the mature intestine, they are less suitable in the neonatal context, where crypt structures are still immature. In addition, crypt-enrichment strategies may bias cell recovery toward proliferative populations and underrepresent differentiated villus cells, thereby influencing cell viability and transcriptional readouts. Accordingly,

this analysis prioritized villus cell viability over comprehensive crypt coverage in adult samples, which likely explains the underrepresentation of crypt stem and Paneth cells.

The analysis of transcriptional dynamics introduced further complexity. RNA velocity uses the relative abundance of spliced and unspliced mRNA to infer gene expression trajectories and offers predictive insights into cell-state transitions (Bergen et al., 2020). However, these velocity models rely on simplifying assumptions, including constant transcriptional rates, which may not hold in tissues with rapid epithelial turnover such as the small intestine. Genes with multiple or complex expression dynamics can violate these assumptions. Another important phenomenon is transcriptional bursting, in which genes switch between inactive and active states and produce RNA in intermittent pulses rather than continuously. Such bursting can lead to spurious or inconsistent velocity estimates (Gorin et al., 2022; Suter et al., 2011; Weiler et al., 2023). In the context of the pronounced regional and age-dependent transcriptional heterogeneity in the small intestine, these limitations likely explain why RNA velocity projections did not produce biologically meaningful trajectories in this study. Trajectory analysis, which orders cells along a continuous trajectory based on transcriptional profiles (Trapnell et al., 2014), provides a more flexible alternative. Single cells are ordered along an inferred developmental path according to similarities in their gene-expression profiles, thereby reconstructing a virtual timeline of biological progression (Qiu et al., 2017; Trapnell et al., 2014). This timeline is referred to as “pseudotime” because it reflects an estimated developmental order rather than real chronological time. ArchR is a trajectory inference tool primarily designed for single-cell ATAC-seq to integrate regulatory information along pseudotime but also works on single-cell data (Granja et al., 2021). Here, it was applied to the generated single-cell RNA-seq data to reconstruct differentiation trajectories based on gene expression, without using its chromatin- or motif-based regulatory features. Different pseudotime analyses using first Monocle (versions 2 and 3) (Qiu et al., 2017) and subsequently ArchR proved difficult to interpret robustly given prior biological knowledge at an early stage of this analysis and were therefore not included in this thesis.

Another key consideration is quality control. Filtering criteria for scRNA-seq vary with tissue, cell type, and experimental conditions, and mitochondrial gene content remains a debated metric. Cells with high mitochondrial RNA are typically excluded due to associations with dissociation-induced stress and necrosis (Heumos et al., 2023; Ilicic et al., 2016; Osorio & Cai, 2021; Van Den Brink et al., 2017). ScRNA-seq studies of intestinal epithelial cells remain relatively sparse and are largely focused on adult tissues, and no established mitochondrial filtering criteria are applied. Based on prior experience from other datasets, a threshold as low as 5% mitochondrial genes was initially proposed to perform well in distinguishing healthy from low-quality cells in mouse tissues (Osorio & Cai, 2021). Subsequent studies adopted substantially more permissive strategies, either filtering cells with up to 50% mitochondrial content (Tsang et al., 2022) or strongly regressing out the effects of mitochondrial genes (Zwick et al.,

2024). At the time of the analysis, a 25% mitochondrial threshold represented a reasonable compromise between stringent quality control and retention of biological information, though more recent data-driven approaches highlight the limitations of fixed cutoffs, particularly given cell-type-specific metabolic activity (Yates et al., 2025).

Finally, pathway inference methods such as PROGENy and decoupleR, while widely used, have inherent limitations. These approaches focus on a restricted set of well-characterized, canonical signaling pathways, meaning that biological processes operating outside these curated routes may be overlooked. In particular, PROGENy infers pathway activity by evaluating the expression of experimentally validated target genes rather than simply assessing pathway component expression. This provides several advantages: it generates quantitative estimates of pathway activity, reduces noise by concentrating on the most responsive targets, and captures downstream signaling effects that may not be reflected in upstream gene expression alone (Badia-i-Mompel et al., 2022). Nevertheless, the reliance on predefined gene sets may bias analyses toward previously studied pathways, and rare or context-specific signaling events could remain undetected. Despite these limitations, the curated and weighted signatures used here make PROGENy and decoupleR suitable for focused assessment of canonical signaling pathways, allowing a robust, quantitative interpretation of pathway activity in the neonatal intestinal epithelium.

10. Concluding remarks

This study builds on existing knowledge of neonatal intestinal development and provides new insights into how epithelial differentiation, metabolism, and host-microbe interactions are coordinated during early life. Many of the transcriptional and proteomic patterns observed align with previous reports, reinforcing the reliability of the dataset, while other findings, such as early distal innate immune signatures, metabolic adaptations, and autophagic activity, highlight aspects of neonatal adaptation that have been less explored. By combining bulk and single-cell analyses with information about crypts and villi from LCM and regional localization of the proximal-to-distal axis, this study was able to resolve cell-type and region-specific programs that are invisible in global measurements alone.

Developmental trajectories in mice and humans are broadly comparable, though on different timelines. The murine intestine before day 10 after birth resembles the early fetal human intestine. At day 14, it aligns with approximately 22 to 23 weeks of gestation in humans, and at day 28, it approaches the intestine of a full-term human newborn (Stanford et al., 2020). Understanding neonatal intestinal development in mice is therefore highly relevant for preterm human infants, whose intestines are immature at birth.

The data presented in this study potentially suggests an early distal immune signature starting at day 5 in mice. Investigating whether a similar distal signature emerges in preterm or term human intestines could clarify how early microbial exposure shapes region-specific immune maturation and potentially also susceptibility to inflammatory disorders. Similar programs in humans may influence susceptibility to infection, responsiveness to oral vaccines, and vulnerability to conditions such as necrotizing enterocolitis (NEC), where immaturity, dysregulated microbial interactions, and excessive inflammation come together (Egozi et al., 2023; Jilling et al., 2006; Juhl et al., 2017).

Despite the LCM-based efforts, a major limitation of this thesis is the lack of comprehensive spatial information on the anatomical organization of intestinal epithelial cells and the restriction to $EpCAM^+$ epithelial cells, which prohibits insights into the surrounding stromal or immune populations. Spatially resolved transcriptomics (Ståhl et al., 2016), awarded “Method of the Year 2020” by Nature Methods, offers a powerful approach to overcome these limitations. By combining transcriptomic profiling with precise spatial barcoding, this method allows the mapping of gene expression directly onto tissue architecture, enabling the localization of both abundant and rare cell types within their native microenvironment (Marx, 2021). Since then, spatial transcriptomics at the single-cell level has become widely commercially available, further increasing its accessibility and utility for detailed tissue mapping (Cheng et al., 2025; Fu et al., 2025; Oliveira et al., 2025). Applying spatial transcriptomics to the neonatal small intestine would provide critical information on the anatomical positioning of epithelial subtypes, reveal local signaling interactions with stromal and immune cells that drive differentiation,

and offer a more complete picture of early postnatal intestinal development (Torow et al., 2015; Venkatraman et al., 2021). As spatial transcriptomics does not require tissue dissociation, the limitations of low viability enterocytes in the infected samples would be overcome; therefore, this approach could also be extended to study host responses to infection with enteropathogens such as *Salmonella* and enteropathogenic *E. coli* (EPEC), providing new insight into spatially coordinated host-microbe interactions (Dupont et al., 2016; Zhang et al., 2014).

Future studies could also leverage advanced trajectory-inference approaches such as InterVelo, which integrates pseudotime with RNA velocity in a joint modeling framework and aims to mitigate limitations inherent to either method alone (Wang et al., 2025b). In addition, different pseudotime methods, such as Slingshot, which are specifically designed to infer and visualize multiple lineages within a single embedding (Street et al., 2018), may enable the simultaneous reconstruction of proximal-to-distal and crypt-villus maturation programs in one UMAP which previously proved challenging. Applying these approaches could help overcome current limitations in resolving regional and temporal differentiation, providing a more comprehensive map of neonatal intestinal development.

Collectively, this work highlights the plasticity of the neonatal intestine during a critical developmental window, illustrating how structural, functional, and microbial factors converge to balance defense, growth, and nutrient absorption. These findings may also extend beyond infection: the timing and regional specialization of epithelial and immune maturation could shape responses to orally delivered antigens, including vaccines. A deeper understanding of the interplay between epithelial development, microbial exposure, and infection-driven transcriptional shifts could inform strategies to optimize oral vaccine delivery and immune priming in early life.

11. Acknowledgements

Here I would like to thank all the people who helped me over the last years with this thesis. It was a long ride with ups and downs along the road from the beginning of November 2019 until now.

First, I want to thank my direct supervisors at the AG Hornef: Aline Dupont and Mathias Hornef. Aline, I think we needed some time to find our common ground between the world of bioinformatics and the wet lab, but over the years I think we complemented both our strengths and weaknesses, and I know without you this thesis would not have been possible. You gave me input and feedback and encouraged me to keep on writing and pushing my boundaries. Thank you for all the time and effort you put into me and my special needs and thank you for all the rounds of corrections this thesis went through to become the final version.

Mathias, thank you for giving me the opportunity to work in your lab. I was the first bioinformatician in the group and kind of "guinea pig" for the way of bioinformatics in a wet lab-focused immunological group. Thank you for believing in me very early on and allowing me to work on my own terms.

Next, I need to thank Ivan Costa, who was a mentor in bioinformatics and who allowed me to work in his group and participate in group meetings, discussions, and group social events like I was one of his students. I also would like to thank Prof. Lars Küpfer, who agreed to be my second supervisor from the CRC, and Prof. Ralph Panstruga, who agreed to be the third supervisor from the department of biology.

Furthermore, I want to thank Christin Meier as coordinator of the CRC graduate school for helpful encouragement along the way and patient support even when I was having panic attacks, especially over the last weeks.

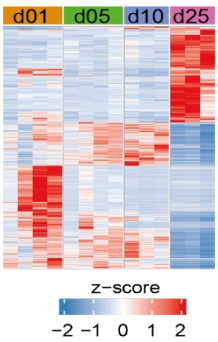
Special thanks also to all CRC members, lab members of AG Hornef and AG Costa, and colleagues of other groups whom I have met over the last years and who shaped this thesis over multiple discussions.

A very special thanks to Keshav; you were a great colleague and friend and helped me when I needed someone to talk about bioinformatics in AG Hornef.

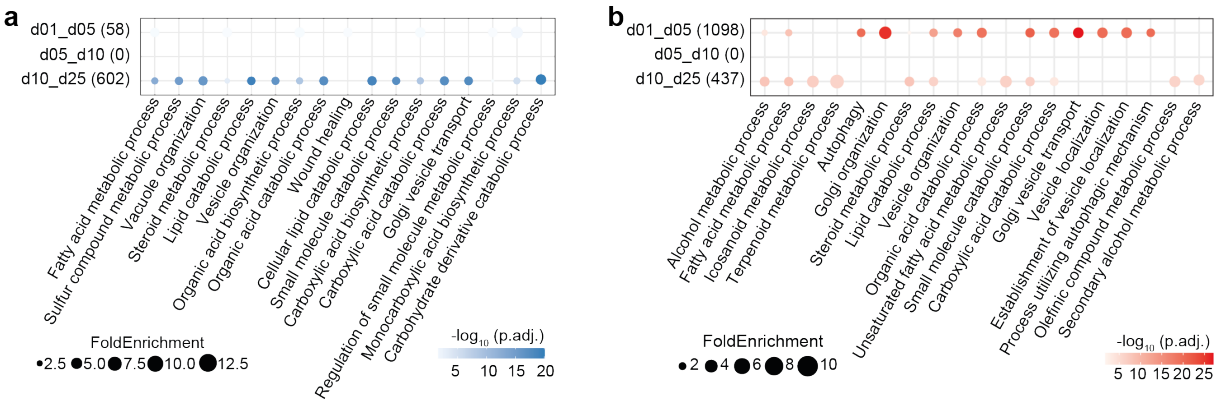
A special shoutout also to Tiago, Mingbo, and James from AG Costa, who were helpful discussion partners and good friends and sometimes teachers of complicated tools and topics. Another shoutout to Anna-Lena for patiently listening to my nerdy topics during commutes, and to Fabian for matching my craziness and bringing positivity throughout the years. I also want to thank Kaiyi, because occasionally all you need in life is a crazy Chinese.

Last, I want to thank my family: my sister, who gave me helpful advice on how to structure my discussion and what parts could be interesting; my father, who was always supportive and never stressed me although everything took a bit longer; and especially my mother, at whose place I could crash over the last weeks of writing to finish the nasty parts and who gave me the special support I needed.

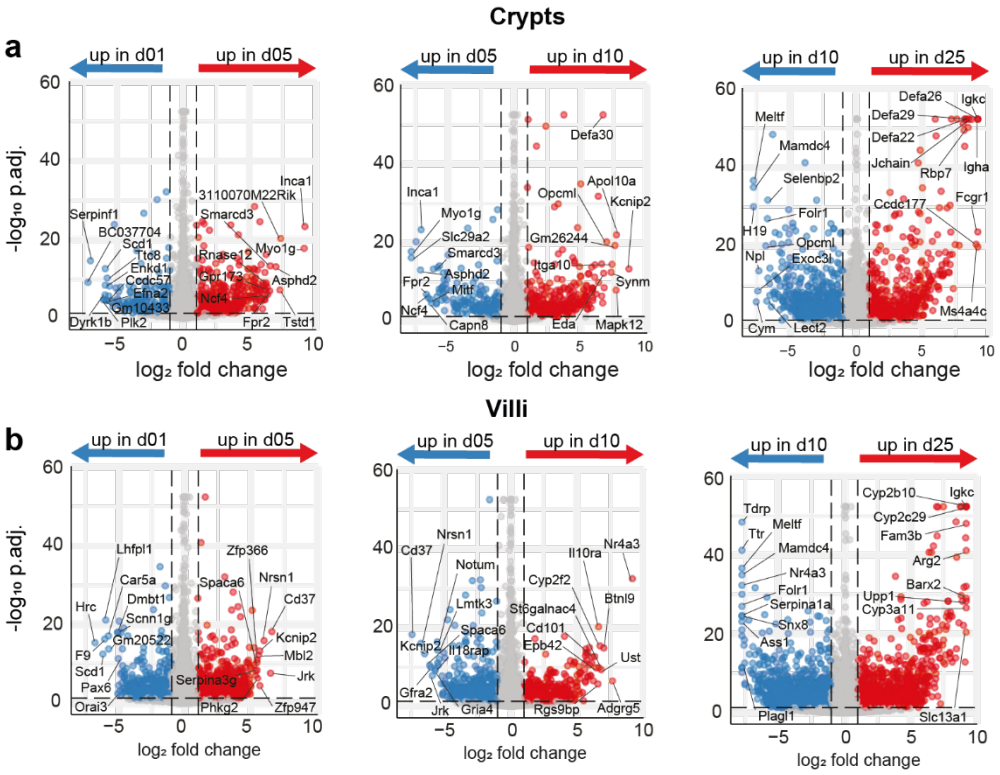
12. Supplementary Figures



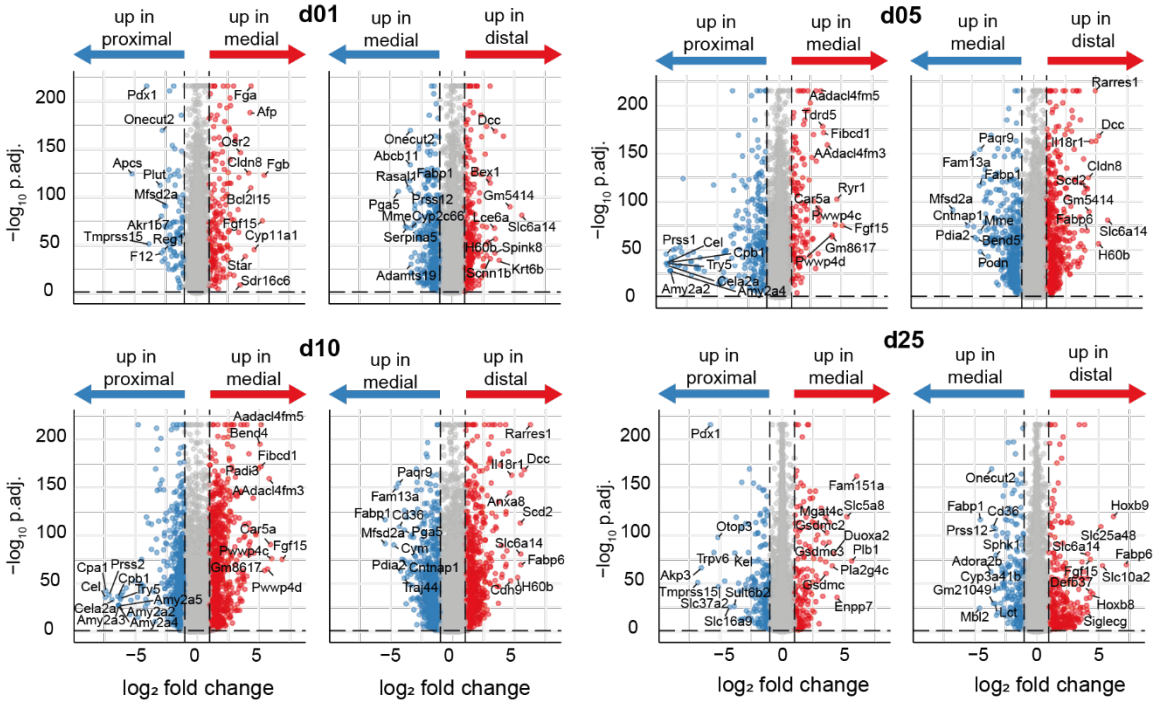
Supplementary Figure 1: Heatmap of temporally differentially expressed genes based on the global transcriptomic data. Heatmap depicting the expression level of all genes differentially regulated between either 1- and 5-day-old, 5- and 10-day-old, and/or 10- and 25-day-old animals.



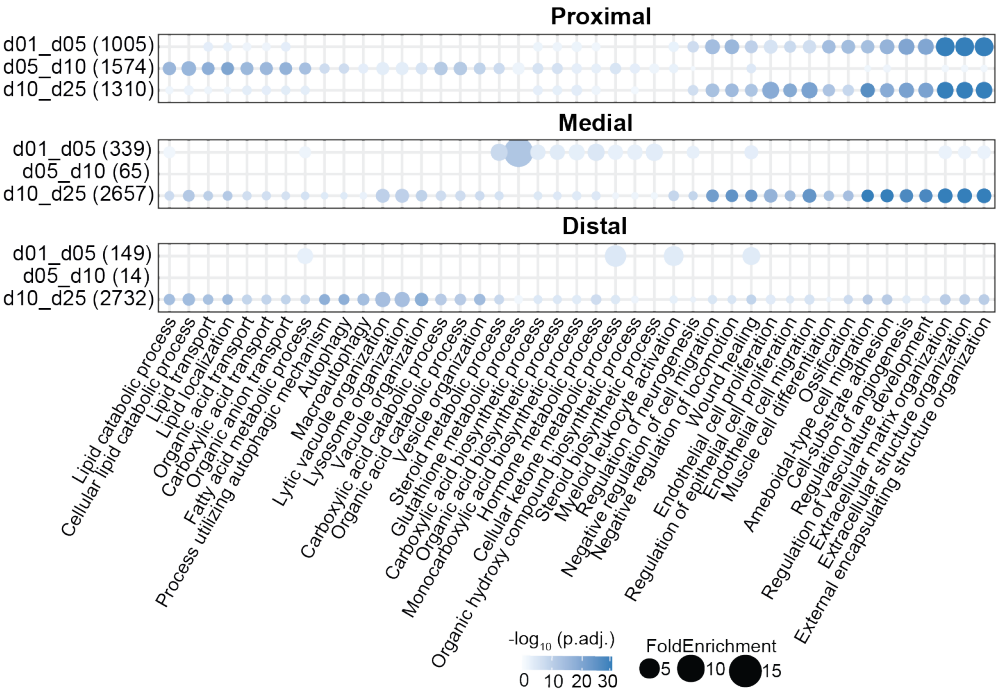
Supplementary Figure 2: Gene Ontology analysis of differentially expressed genes in the proteomics data. (a) Dot plot showing the top 10 enriched GO terms (by adjusted p-value) for the differentially expressed genes downregulated with age. Brackets indicate the number of DE genes associated with a GO term for each comparison. **(b)** Dot plot showing the top 10 enriched GO terms (by adjusted p-value) for the differentially expressed genes upregulated with age. The number of DE genes associated with a GO term for each comparison is indicated in brackets.



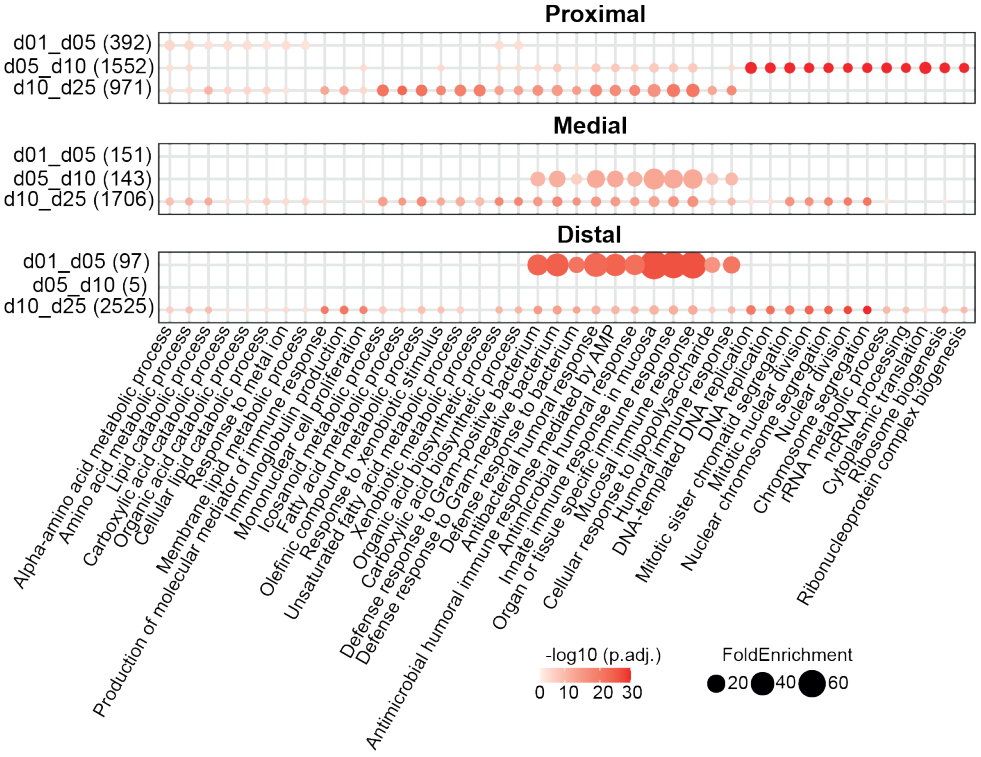
Supplementary Figure 3: Differential expression between consecutive time points in either crypts or villi. (a) Volcano plots showing differential gene expression between 1- and 5-day-old (left), 5- and 10-day-old (middle), and 10- and 25-day-old (right) crypt samples. Blue and red dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in the younger and older animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes. **(b)** Volcano plots showing differential gene expression between 1- and 5-day-old (left), 5- and 10-day-old (middle), and 10- and 25-day-old (right) villi samples. Blue and red dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in the younger and older animals, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes.



Supplementary Figure 4: Differential gene expression between intestinal segments per time point. Volcano plots showing differential gene expression between proximal and medial (left) and medial and distal (right) of 1-, 5-, 10-, and 25-day-old animals (top left to bottom right). Blue and red dots indicate statistically upregulated genes (p.adj. < 0.05 & |log2FC| > 1) in the most proximal and most distal parts, respectively. Grey dots indicate non-significant and/or non-differentially regulated genes.



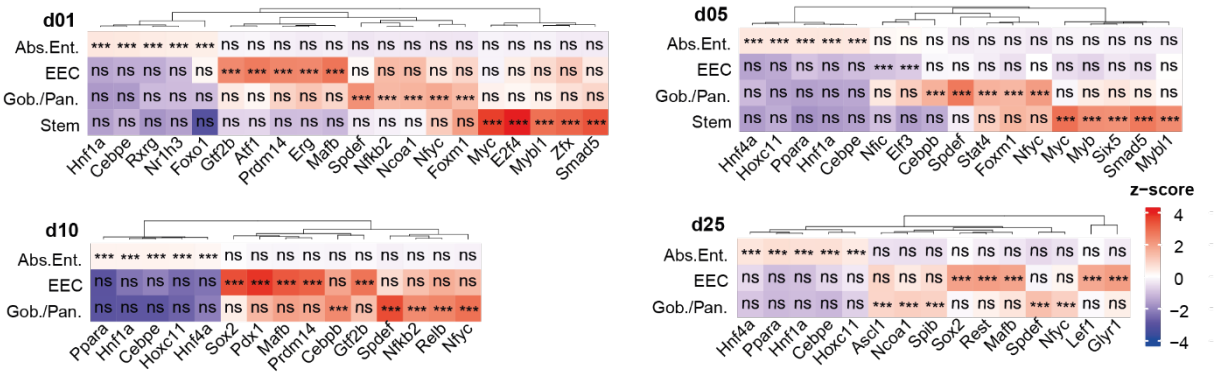
Supplementary Figure 5: Gene Ontology analysis for the genes downregulated with age in proximal, medial, and distal small intestine. Dot plot showing the top 10 enriched GO terms (by adjusted p-value) for the differentially expressed genes downregulated with age in the proximal, medial, and distal parts (top to bottom). The number of DE genes associated with a GO term for each comparison is indicated in brackets.



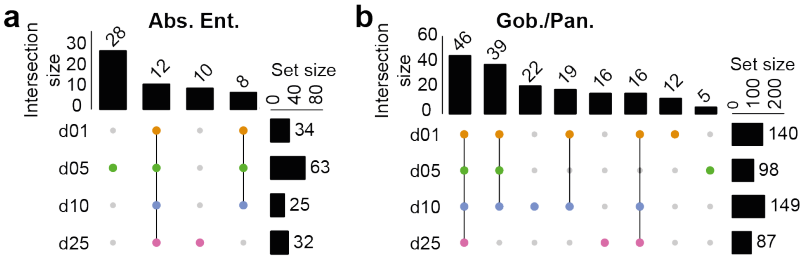
Supplementary Figure 6: Gene Ontology analysis for the genes upregulated with age in proximal, medial, and distal small intestine. Dot plot showing the top 10 enriched GO terms (by adjusted p-value) for the differentially expressed genes upregulated with age in the proximal, medial, and distal parts (top to bottom). The number of DE genes associated with a GO term for each comparison is indicated in brackets.

	Abs.Ent.	EEC	Gob./Pan.	Stem	Tuft	Rowsums
d01	3530	195	704	152	2	4583
d05	1339	153	183	127	3	1805
d05i	81	100	207	89	7	484
d10	1573	21	108	4	5	1711
d25	396	31	198	1	35	661
colsums	6919	500	1400	373	52	9244

Supplementary Figure 7: Table for the number of cells per single-cell cluster. Number of cells recovered per cell type and per time point after filtering.



Supplementary Figure 8: DoRothEA Transcription Factor Heatmaps. Heatmaps depicting the top 5 transcription factors (TF) per cell type predicted by DoRothEA in 1-, 5-, 10-, and 25-day-old animals.



Supplementary Figure 9: Upset plots for the number of DoRothEA-predicted transcription factors. (a-b) Upset plots showing the number of DoRothEA-predicted transcription factors specific to one or more age groups for the absorptive enterocyte (a) and goblet/Paneth clusters (b). Interactions with less than 5 TFs are not shown.

a

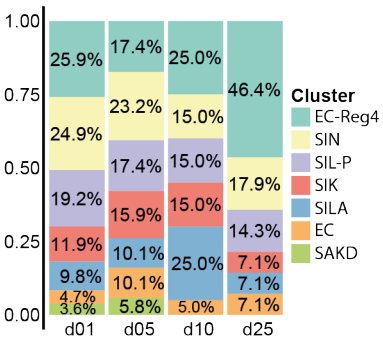
Abs.Ent.	0	1	2	3	4	5	6	7	8	9	10	rowsums
d01	1159	3	14	898	818	3	43	104	185	298	5	3530
d05	1	675	249	3	4	31	235	52	80	0	9	1339
d10	4	278	446	2	18	470	103	221	21	0	10	1573
d25	0	53	216	0	2	54	49	4	11	0	7	396
colsums	1164	1009	925	903	842	558	430	381	297	298	31	6838

b

Gob./Pan.	0	1	2	3	4	5	6	rowsums
d01	10	283	2	167	158	40	10	670
d05	86	2	21	3	3	21	16	152
d10	57	2	24	0	2	10	5	100
d25	105	1	55	0	0	4	13	178
colsums	258	288	102	170	163	75	44	1100

Supplementary Figure 10: Tables for the number of cells per cluster in absorptive enterocytes and goblet/Paneth cells. (a-b) Number of cells per absorptive enterocyte (a) and goblet/Paneth (b) subcluster and per time point.

a



b

EEC	EC-Reg4	SIN	SIL-P	SIK	SILA	EC	SAKD	rowsums
d01	50	48	37	23	19	9	7	193
d05	12	16	12	11	7	7	4	69
d10	5	3	3	3	5	1	0	20
d25	13	5	4	2	2	2	0	28
colsums	80	72	56	39	33	19	11	310

Supplementary Figure 11: Barplot and table for the frequency and number of cells per EEC subcluster. (a) Frequency of the 7 subtypes of EECs per time point. Subcluster color code as in Figure 32a. **(b)** Number of cells per EEC subcluster and per time point.

Genename	Description	Log2FC
Reg3b	regenerating islet-derived 3 beta	7.966557
Reg3g	regenerating islet-derived 3 gamma	7.966557
Gpx2	glutathione peroxidase 2	7.745384
Dmbt1	deleted in malignant brain tumors 1	6.256353
Zg16	zymogen granule protein 16	4.991891
Cd9	CD9 antigen	4.591181
Ceacam1	CEA cell adhesion molecule 1	3.459187
H2-Q7	histocompatibility 2, Q region locus 7	3.421775
Nfkbia	nuclear factor of kappa light polypeptide gene enhancer in B cells inhibitor, alpha	3.173203
Mxd1	MAX dimerization protein 1	2.718314
Hadh	hydroxyacyl-Coenzyme A dehydrogenase	-1.032798
Apoa4	apolipoprotein A-IV	-1.038476
Mogat2	monoacylglycerol O-acyltransferase 2	-1.058274
Cbr1	carbonyl reductase 1	-1.091934
Calml4	calmodulin-like 4	-1.098461
Acaa2	acetyl-CoA acyltransferase 2	-1.198362
Fabp1	fatty acid binding protein 1, liver	-1.213633
Dhrs1	dehydrogenase/reductase 1	-1.252412
Fabp2	fatty acid binding protein 2, intestinal	-1.353597
Meltf	melanotransferrin	-1.421864
Plin2	perilipin 2	-1.545676
Cotl1	coactosin like F-actin binding protein 1	-1.573325
Cat	catalase	-1.614216
Mettl7b	methyltransferase-like protein 7B	-1.624850
Ech1	enoyl coenzyme A hydratase 1, peroxisomal	-1.633428
Cidec	cell death-inducing DFFA-like effector c	-1.652115
Apoc3	apolipoprotein C-III	-1.698664
Oat	ornithine aminotransferase	-1.800532
Dgat2	diacylglycerol O-acyltransferase 2	-1.880357
Galk1	galactokinase 1	-1.915045
Leap2	liver-expressed antimicrobial peptide 2	-2.100387
Hmgcs2	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2	-2.183070
Fcgrt	Fc fragment of IgG receptor and transporter	-2.263679
Tdrp	testis development related protein	-2.438317
Pck1	phosphoenolpyruvate carboxykinase 1, cytosolic	-2.524579
Acat1	acetyl-Coenzyme A acetyltransferase 1	-2.545930
Smim1	small integral membrane protein 1	-2.684966
Lct	lactase	-2.754195
Lbh	limb-bud and heart	-2.824772
Fbp1	fructose bisphosphatase 1	-4.160972
Ttr	transthyretin	-4.248549
Cym	chymosin	-4.570286

Supplementary Figure 12: Table of common infection response genes. Table of the significant common response DE genes across all 4 major cell types between 5-day-old uninfected (d05) and *Salmonella*-infected (d05i) animals. Negative log2 fold changes (Log2FC) signify genes downregulated; positive log2 fold changes signify genes upregulated with infection.

13. References

Akimori, T., Hanazaki, K., Okabayashi, T., Okamoto, K., Kobayashi, M., & Ogata, T. (2011). Quantitative distribution of brush cells in the rat gastrointestinal tract: Brush cell population coincides with NaHCO3 secretion. *Medical Molecular Morphology*, *44*(1), 7–14. https://doi.org/10.1007/s00795-009-0488-1

Al Nabhani, Z., Dulauroy, S., Marques, R., Cousu, C., Al Bounny, S., Déjardin, F., Sparwasser, T., Bérard, M., Cerf-Bensussan, N., & Eberl, G. (2019). A Weaning Reaction to Microbiota Is Required for Resistance to Immunopathologies in the Adult. *Immunity*, *50*(5), 1276-1288.e5. https://doi.org/10.1016/j.immuni.2019.02.014

Aronson, B. E., Rabello Aronson, S., Berkhout, R. P., Chavoushi, S. F., He, A., Pu, W. T., Verzi, M. P., & Krasinski, S. D. (2014). GATA4 represses an ileal program of gene expression in the proximal small intestine by inhibiting the acetylation of histone H3, lysine 27. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1839*(11), 1273–1282. https://doi.org/10.1016/j.bbagrm.2014.05.018

Arshad, M. M., Wilkins, M. J., Downes, F. P., Rahbar, M. H., Erskine, R. J., Boulton, M. L., Younus, M., & Saeed, A. M. (2008). Epidemiologic attributes of invasive non-typhoidal Salmonella infections in Michigan, 1995–2001. *International Journal of Infectious Diseases*, *12*(2), 176–182. https://doi.org/10.1016/j.ijid.2007.06.006

Ayabe, T., Ashida, T., Kohgo, Y., & Kono, T. (2004). The role of Paneth cells and their antimicrobial peptides in innate host defense. *Trends in Microbiology*, *12*(8), 394–398. https://doi.org/10.1016/j.tim.2004.06.007

Badia-i-Mompel, P., Vélez Santiago, J., Braunger, J., Geiss, C., Dimitrov, D., Müller-Dott, S., Taus, P., Dugourd, A., Holland, C. H., Ramirez Flores, R. O., & Saez-Rodriguez, J. (2022). decoupleR: Ensemble of computational methods to infer biological activities from omics data. *Bioinformatics Advances*, *2*(1), vbac016. https://doi.org/10.1093/bioadv/vbac016

Bara, A. M., Chen, L., Ma, C., Underwood, J., Moreci, R. S., Sumigray, K., Sun, T., Diao, Y., Verzi, M., & Lechler, T. (2022). Maf family transcription factors are required for nutrient uptake in the mouse neonatal gut. *Development*, *149*(23), dev201251. https://doi.org/10.1242/dev.201251

Barker, N. (2014). Adult intestinal stem cells: Critical drivers of epithelial homeostasis and regeneration. *Nature Reviews Molecular Cell Biology*, *15*(1), 19–33. https://doi.org/10.1038/nrm3721

Barker, N., Van Es, J. H., Kuipers, J., Kujala, P., Van Den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P. J., & Clevers, H. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature*, *449*(7165), 1003–1007. https://doi.org/10.1038/nature06196

Barthel, M., Hapfelmeier, S., Quintanilla-Martínez, L., Kremer, M., Rohde, M., Hogardt, M., Pfeffer, K., Rüssmann, H., & Hardt, W.-D. (2003). Pretreatment of Mice with Streptomycin Provides a *Salmonella enterica* Serovar Typhimurium Colitis Model That Allows Analysis of Both Pathogen and Host. *Infection and Immunity*, *71*(5), 2839–2858. https://doi.org/10.1128/IAI.71.5.2839-2858.2003

Barton, J. R., Londregan, A. K., Alexander, T. D., Entezari, A. A., Covarrubias, M., & Waldman, S. A. (2023). Enteroendocrine cell regulation of the gut-brain axis. *Frontiers in Neuroscience*, *17*, 1272955. https://doi.org/10.3389/fnins.2023.1272955

Bas, J., Jay, P., & Gerbe, F. (2023). Intestinal tuft cells: Sentinels, what else? *Seminars in Cell & Developmental Biology*, *150–151*, 35–42. https://doi.org/10.1016/j.semcdb.2023.02.012

Bastide, P., Darido, C., Pannequin, J., Kist, R., Robine, S., Marty-Double, C., Bibeau, F., Scherer, G., Joubert, D., Hollande, F., Blache, P., & Jay, P. (2007). Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium. *The Journal of Cell Biology*, *178*(4), 635–648. https://doi.org/10.1083/jcb.200704152

Batlle, E., Henderson, J. T., Beghtel, H., Van Den Born, M. M. W., Sancho, E., Huls, G., Meeldijk, J., Robertson, J., Van De Wetering, M., Pawson, T., & Clevers, H. (2002). β-Catenin and TCF Mediate Cell Positioning in the Intestinal Epithelium by Controlling the Expression of EphB/EphrinB. *Cell*, *111*(2), 251–263. https://doi.org/10.1016/S0092-8674(02)01015-2

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Bergen, V., Lange, M., Peidli, S., Wolf, F. A., & Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology*, *38*(12), 1408–1414. https://doi.org/10.1038/s41587-020-0591-3

Bergström, J. H., Birchenough, G. M. H., Katona, G., Schroeder, B. O., Schütte, A., Ermund, A., Johansson, M. E. V., & Hansson, G. C. (2016). Gram-positive bacteria are held at a distance in the colon mucus by the lectin-like protein ZG16. *Proceedings of the National Academy of Sciences*, *113*(48), 13833–13838. https://doi.org/10.1073/pnas.1611400113

Bergstrom, K. S. B., Morampudi, V., Chan, J. M., Bhinder, G., Lau, J., Yang, H., Ma, C., Huang, T., Ryz, N., Sham, H. P., Zarepour, M., Zaph, C., Artis, D., Nair, M., & Vallance, B. A. (2015). Goblet Cell Derived RELM-β Recruits CD4+ T Cells during Infectious Colitis to Promote Protective Intestinal Epithelial Cell Proliferation. *PLOS Pathogens*, *11*(8), e1005108. https://doi.org/10.1371/journal.ppat.1005108

Beumer, J., Artegiani, B., Post, Y., Reimann, F., Gribble, F., Nguyen, T. N., Zeng, H., Van Den Born, M., Van Es, J. H., & Clevers, H. (2018). Enteroendocrine cells switch hormone expression along the crypt-to-villus BMP signalling gradient. *Nature Cell Biology*, *20*(8), 909–916. https://doi.org/10.1038/s41556-018-0143-y

Beumer, J., & Clevers, H. (2021). Cell fate specification and differentiation in the adult mammalian intestine. *Nature Reviews Molecular Cell Biology*, *22*(1), 39–53. https://doi.org/10.1038/s41580-020-0278-0

Beumer, J., Gehart, H., & Clevers, H. (2020). Enteroendocrine Dynamics – New Tools Reveal Hormonal Plasticity in the Gut. *Endocrine Reviews*, *41*(5), bnaa018. https://doi.org/10.1210/endrev/bnaa018

Beumer, J., Puschhof, J., Yengej, F. Y., Zhao, L., Martinez-Silgado, A., Blotenburg, M., Begthel, H., Boot, C., Van Oudenaarden, A., Chen, Y.-G., & Clevers, H. (2022). BMP gradient along the intestinal villus axis controls zonated enterocyte and goblet cell states. *Cell Reports*, *38*(9), 110438. https://doi.org/10.1016/j.celrep.2022.110438

Blighe, K. (2018). *EnhancedVolcano*. Bioconductor. https://doi.org/10.18129/B9.BIOC.ENHANCEDVOLCANO

Block, C. L., Childers, L., Cortez, A. L., Sakers, K., Lewis, T. R., Levic, D. S., Frazer, L. C., Jania, C. M., Arshavsky, V. Y., Good, M., Eroglu, C., & Bagnat, M. (2025). The mouse neonatal small intestine is regionally specialized for protein absorption and transepithelial transport. *Development*, *152*(23), dev205127. https://doi.org/10.1242/dev.205127

Borbet, T. C., Pawline, M. B., Li, J., Ho, M. L., Yin, Y. S., Zhang, X., Novikova, E., Jackson, K., Mullins, B. J., Ruiz, V. E., Hines, M. J., Zhang, X.-S., Müller, A., Koralov, S. B., & Blaser, M. J. (2023). Disruption of the early-life microbiota alters Peyer's patch development and germinal center formation in gastrointestinal-associated lymphoid tissue. *iScience*, *26*(6), 106810. https://doi.org/10.1016/j.isci.2023.106810

Bosse, T., Piaseckyj, C. M., Burghard, E., Fialkovich, J. J., Rajagopal, S., Pu, W. T., & Krasinski, S. D. (2006). Gata4 Is Essential for the Maintenance of Jejunal-Ileal Identities in the Adult Mouse Small Intestine. *Molecular and Cellular Biology*, *26*(23), 9060–9070. https://doi.org/10.1128/MCB.00124-06

Bry, L., Falk, P., Huttner, K., Ouellette, A., Midtvedt, T., & Gordon, J. I. (1994). Paneth cell differentiation in the developing intestine of normal and transgenic mice. *Proceedings of the National Academy of Sciences*, *91*(22), 10335–10339. https://doi.org/10.1073/pnas.91.22.10335

Bula-Rudas, F. J., Rathore, M. H., & Maraqa, N. F. (2015). Salmonella Infections in Childhood. *Advances in Pediatrics*, *62*(1), 29–58. https://doi.org/10.1016/j.yapd.2015.04.005

Burclaff, J., Bliton, R. J., Breau, K. A., Ok, M. T., Gomez-Martinez, I., Ranek, J. S., Bhatt, A. P., Purvis, J. E., Woosley, J. T., & Magness, S. T. (2022). A Proximal-to-Distal Survey of Healthy Adult Human Small Intestine and Colon Epithelium by Single-Cell Transcriptomics. *Cellular and Molecular Gastroenterology and Hepatology*, *13*(5), 1554–1589. https://doi.org/10.1016/j.jcmgh.2022.02.007

Chassin, C., Kocur, M., Pott, J., Duerr, C. U., Gütle, D., Lotz, M., & Hornef, M. W. (2010). miR-146a Mediates Protective Innate Immune Tolerance in the Neonate Intestine. *Cell Host & Microbe*, *8*(4), 358–368. https://doi.org/10.1016/j.chom.2010.09.005

Chatterton, D. E. W., Rasmussen, J. T., Heegaard, C. W., Sørensen, E. S., & Petersen, T. E. (2004). In vitro digestion of novel milk protein ingredients for use in infant formulas: Research on biological functions. *Trends in Food Science & Technology*, *15*(7–8), 373–383. https://doi.org/10.1016/j.tifs.2003.12.004

Chen, L., Toke, N. H., Luo, S., Vasoya, R. P., Fullem, R. L., Parthasarathy, A., Perekatt, A. O., & Verzi, M. P. (2019). A reinforcing HNF4–SMAD4 feed-forward module stabilizes enterocyte identity. *Nature Genetics*, *51*(5), 777–785. https://doi.org/10.1038/s41588-019-0384-0

Cheng, H., & Leblond, C. P. (1974). Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine V. Unitarian theory of the origin of the four epithelial cell types. *American Journal of Anatomy*, *141*(4), 537–561. https://doi.org/10.1002/aja.1001410407

Cheng, Y., Dang, S., Zhang, Y., Chen, Y., Yu, R., Liu, M., Jin, S., Han, A., Katz, S., & Wang, S. (2025). Sequencing-free whole-genome spatial transcriptomics at single-molecule resolution. *Cell*, *188*(24), 6953-6970.e12. https://doi.org/10.1016/j.cell.2025.09.006

Chin, A. M., Hill, D. R., Aurora, M., & Spence, J. R. (2017). Morphogenesis and maturation of the embryonic and postnatal intestine. *Seminars in Cell & Developmental Biology*, *66*, 81–93. https://doi.org/10.1016/j.semcdb.2017.01.011

Chow, J. C., Yen, Z., Ziesche, S. M., & Brown, C. J. (2005). Silencing of the Mammalian X Chromosome. *Annual Review of Genomics and Human Genetics*, *6*(1), 69–92. https://doi.org/10.1146/annurev.genom.6.080604.162350

Chureau, C., Prissette, M., Bourdet, A., Barbe, V., Cattolico, L., Jones, L., Eggen, A., Avner, P., & Duret, L. (2002). Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. *Genome Research*, *12*(6), 894–908. https://doi.org/10.1101/gr.152902

Clevers, H. (2013). The Intestinal Crypt, A Prototype Stem Cell Compartment. *Cell*, *154*(2), 274–284. https://doi.org/10.1016/j.cell.2013.07.004

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372. https://doi.org/10.1038/nbt.1511

Crissey, M. A. S., Guo, R., Funakoshi, S., Kong, J., Liu, J., & Lynch, J. P. (2011). Cdx2 Levels Modulate Intestinal Epithelium Maturity and Paneth Cell Development. *Gastroenterology*, *140*(2), 517-528.e8. https://doi.org/10.1053/j.gastro.2010.11.033

De Goffau, M. C., Charnock-Jones, D. S., Smith, G. C. S., & Parkhill, J. (2021). Batch effects account for the main findings of an in utero human intestinal bacterial colonization study. *Microbiome*, *9*(1), 6. https://doi.org/10.1186/s40168-020-00949-z

De Jonge, W. J., Dingemanse, M. A., De Boer, P. A. J., Lamers, W. H., & Moorman, A. F. M. (1998). Arginine-Metabolizing Enzymes in the Developing Rat Small Intestine. *Pediatric Research*, *43*(4), 442–451. https://doi.org/10.1203/00006450-199804000-00002

de Santa Barbara, P., van den Brink, G. R., & Roberts, D. J. (2003). Development and differentiation of the intestinal epithelium. *Cellular and Molecular Life Sciences (CMLS)*, *60*(7), 1322–1332. https://doi.org/10.1007/s00018-003-2289-3

Dehmer, J. J., Garrison, A. P., Speck, K. E., Dekaney, C. M., Van Landeghem, L., Sun, X., Henning, S. J., & Helmrath, M. A. (2011). Expansion of Intestinal Epithelial Stem Cells during Murine Development. *PLoS ONE*, *6*(11), e27070. https://doi.org/10.1371/journal.pone.0027070

Duarte-Mata, D. I., & Salinas-Carmona, M. C. (2023). Antimicrobial peptides´ immune modulation role in intracellular bacterial infection. *Frontiers in Immunology*, *14*, 1119574. https://doi.org/10.3389/fimmu.2023.1119574

Dupont, A., Sommer, F., Zhang, K., Repnik, U., Basic, M., Bleich, A., Kühnel, M., Bäckhed, F., Litvak, Y., Fulde, M., Rosenshine, I., & Hornef, M. W. (2016). Age-Dependent Susceptibility to Enteropathogenic Escherichia coli (EPEC) Infection in Mice. *PLOS Pathogens*, *12*(5), e1005616. https://doi.org/10.1371/journal.ppat.1005616

Dutta, S., & Sengupta, P. (2016). Men and mice: Relating their ages. *Life Sciences*, *152*, 244–248. https://doi.org/10.1016/j.lfs.2015.10.025

Earley, Z. M., Lisicka, W., Sifakis, J. J., Aguirre-Gamboa, R., Kowalczyk, A., Barlow, J. T., Shaw, D. G., Discepolo, V., Tan, I. L., Gona, S., Ernest, J. D., Matzinger, P., Barreiro, L. B., Morgun, A., Bendelac, A., Ismagilov, R. F., Shulzhenko, N., Riesenfeld, S. J., & Jabri, B. (2023). GATA4 controls regionalization of tissue immunity and commensal-driven immunopathology. *Immunity*, *56*(1), 43-57.e10. https://doi.org/10.1016/j.immuni.2022.12.009

Eberwine, J., Sul, J.-Y., Bartfai, T., & Kim, J. (2014). The promise of single-cell sequencing. *Nature Methods*, *11*(1), 25–27. https://doi.org/10.1038/nmeth.2769

Egozi, A., Olaloye, O., Werner, L., Silva, T., McCourt, B., Pierce, R. W., An, X., Wang, F., Chen, K., Pober, J. S., Shouval, D., Itzkovitz, S., & Konnikova, L. (2023). Single-cell atlas of the human neonatal small intestine affected by necrotizing enterocolitis. *PLOS Biology*, *21*(5), e3002124. https://doi.org/10.1371/journal.pbio.3002124

Eidem, H. R., Ackerman, W. E., McGary, K. L., Abbot, P., & Rokas, A. (2015). Gestational tissue transcriptomics in term and preterm human pregnancies: A systematic review and meta-analysis. *BMC Medical Genomics*, *8*(1), 27. https://doi.org/10.1186/s12920-015-0099-8

Fevr, T., Robine, S., Louvard, D., & Huelsken, J. (2007). Wnt/β-Catenin Is Essential for Intestinal Homeostasis and Maintenance of Intestinal Stem Cells. *Molecular and Cellular Biology*, *27*(21), 7551–7559. https://doi.org/10.1128/MCB.01034-07

Fish, E. M., Shumway, K. R., & Burns, B. (2024). Physiology, Small Bowel. In *StatPearls*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK532263/

Flanagan, K., Modrusan, Z., Cornelius, J., Chavali, A., Kasman, I., Komuves, L., Mo, L., & Diehl, L. (2008). Intestinal Epithelial Cell Up-Regulation of LY6 Molecules during Colitis Results in Enhanced Chemokine Secretion. *The Journal of Immunology*, *180*(6), 3874–3881. https://doi.org/10.4049/jimmunol.180.6.3874

Fothergill, L. J., & Furness, J. B. (2018). Diversity of enteroendocrine cells investigated at cellular and subcellular levels: The need for a new classification scheme. *Histochemistry and Cell Biology*, *150*(6), 693–702. https://doi.org/10.1007/s00418-018-1746-x

Fraser, I. P., Koziel, H., & Ezekowitz, R. A. B. (1998). The serum mannose-binding protein and the macrophage mannose receptor are pattern recognition molecules that link innate and adaptive immunity. *Seminars in Immunology*, *10*(5), 363–372. https://doi.org/10.1006/smim.1998.0141

Fu, X., Cao, Y., Bian, B., Wang, C., Graham, D., Pathmanathan, N., Patrick, E., Kim, J., & Yang, J. Y. H. (2025). Spatial gene expression at single-cell resolution from histology using deep learning with GHIST. *Nature Methods*, *22*(9), 1900–1910. https://doi.org/10.1038/s41592-025-02795-z

Fujita, M., Baba, R., Shimamoto, M., Sakuma, Y., & Fujimoto, S. (2007). Molecular morphology of the digestive tract; macromolecules and food allergens are transferred intact across the intestinal absorptive cells during the neonatal-suckling period. *Medical Molecular Morphology*, *40*(1), 1–7. https://doi.org/10.1007/s00795-006-0346-3

Gal-Mor, O. (2018). Persistent Infection and Long-Term Carriage of Typhoidal and Nontyphoidal Salmonellae. *Clinical Microbiology Reviews*, *32*(1), e00088-18. https://doi.org/10.1128/CMR.00088-18

Garcia-Alonso, L., Holland, C. H., Ibrahim, M. M., Turei, D., & Saez-Rodriguez, J. (2019). Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Research*, *29*(8), 1363–1375. https://doi.org/10.1101/gr.240663.118

Gonnella, P. A., Siminoski, K., Murphy, R. A., & Neutra, M. R. (1987). Transepithelial transport of epidermal growth factor by absorptive cells of suckling rat ileum. *Journal of Clinical Investigation*, *80*(1), 22–32. https://doi.org/10.1172/JCI113051

Goodall, K. J., Nguyen, A., Matsumoto, A., McMullen, J. R., Eckle, S. B., Bertolino, P., Sullivan, L. C., & Andrews, D. M. (2019). Multiple receptors converge on H2-Q10 to regulate NK and γδT-cell development. *Immunology & Cell Biology*, *97*(3), 326–339. https://doi.org/10.1111/imcb.12222

Gorin, G., Fang, M., Chari, T., & Pachter, L. (2022). RNA velocity unraveled. *PLOS Computational Biology*, *18*(9), e1010492. https://doi.org/10.1371/journal.pcbi.1010492

Granja, J. M., Corces, M. R., Pierce, S. E., Bagdatli, S. T., Choudhry, H., Chang, H. Y., & Greenleaf, W. J. (2021). ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nature Genetics*, *53*(3), 403–411. https://doi.org/10.1038/s41588-021-00790-6

Gribble, F. M., & Reimann, F. (2016). Enteroendocrine Cells: Chemosensors in the Intestinal Epithelium. *Annual Review of Physiology*, *78*(1), 277–299. https://doi.org/10.1146/annurev-physiol-021115-105439

Gu, W., Eke, C., Gonzalez Santiago, E., Olaloye, O., & Konnikova, L. (2024). Single-cell atlas of the small intestine throughout the human lifespan demonstrates unique features of fetal immune cells. *Mucosal Immunology*, *17*(4), 599–617. https://doi.org/10.1016/j.mucimm.2024.03.011

Gu, Z. (2022). Complex heatmap visualization. *iMeta*, *1*(3), e43. https://doi.org/10.1002/imt2.43

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, *32*(18), 2847–2849. https://doi.org/10.1093/bioinformatics/btw313

Guo, R.-J., Suh, E. R., & Lynch, J. P. (2004). The role of Cdx proteins in intestinal development and cancer. *Cancer Biology & Therapy*, *3*(7), 593–601. https://doi.org/10.4161/cbt.3.7.913

Haber, A. L., Biton, M., Rogel, N., Herbst, R. H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T. M., Howitt, M. R., Katz, Y., Tirosh, I., Beyaz, S., Dionne, D., Zhang, M., Raychowdhury, R., Garrett, W. S., Rozenblatt-Rosen, O., Shi, H. N., Yilmaz, O., … Regev, A. (2017). A single-cell survey of the small intestinal epithelium. *Nature*, *551*(7680), 333–339. https://doi.org/10.1038/nature24489

Habib, A. M., Richards, P., Cairns, L. S., Rogers, G. J., Bannon, C. A. M., Parker, H. E., Morley, T. C. E., Yeo, G. S. H., Reimann, F., & Gribble, F. M. (2012). Overlap of Endocrine Hormone Expression in the Mouse Intestine Revealed by Transcriptional Profiling and Flow Cytometry. *Endocrinology*, *153*(7), 3054–3065. https://doi.org/10.1210/en.2011-2170

Harper, J., Mould, A., Andrews, R. M., Bikoff, E. K., & Robertson, E. J. (2011). The transcriptional repressor Blimp1/Prdm1 regulates postnatal reprogramming of intestinal enterocytes. *Proceedings of the National Academy of Sciences*, *108*(26), 10585–10590. https://doi.org/10.1073/pnas.1105852108

He, S., Liu, G., & Zhu, X. (2021). Human breast milk-derived exosomes may help maintain intestinal epithelial barrier integrity. *Pediatric Research*, *90*(2), 366–372. https://doi.org/10.1038/s41390-021-01449-y

Heidenhain, R. (1870). Untersuchungen über den Bau der Labdrüsen. *Archiv für Mikroskopische Anatomie*, *6*(1), 368–406. https://doi.org/10.1007/BF02955986

Heinrichs, A., Brück, J., Stange, E., Nguyen, T.-H., van Best, N., Pes, J., Buhl, E. M., Schöneich, J., Brandt, E., Simons, B., Brandenburg, K., Freudenberg, M., Pabst, O., Berres, M.-L., Dupont, A., Zhang, K., Riba, A., Torow, N., & Hornef, M. W. (2025). *Early life hepatic innate immune cell priming prior to gut closure*.

Heuberger, J., Kosel, F., Qi, J., Grossmann, K. S., Rajewsky, K., & Birchmeier, W. (2014). Shp2/MAPK signaling controls goblet/paneth cell fate decisions in the intestine. *Proceedings of the National Academy of Sciences*, *111*(9), 3472–3477. https://doi.org/10.1073/pnas.1309342111

Heumos, L., Schaar, A. C., Lance, C., Litinetskaya, A., Drost, F., Zappia, L., Lücken, M. D., Strobl, D. C., Henao, J., Curion, F., Single-cell Best Practices Consortium, Aliee, H., Ansari, M., Badia-i-Mompel, P., Büttner, M., Dann, E., Dimitrov, D., Dony, L., Frishberg, A., … Theis, F. J. (2023). Best practices for single-cell analysis across modalities. *Nature Reviews Genetics*, *24*(8), 550–572. https://doi.org/10.1038/s41576-023-00586-w

Hurwitz, R., & Kretchmer, N. (1986). Development of arginine-synthesizing enzymes in mouse intestine. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *251*(1), G103–G110. https://doi.org/10.1152/ajpgi.1986.251.1.G103

Huttner, K. M., Kozak, C. A., & Bevins, C. L. (1997). The mouse genome encodes a single homolog of the antimicrobial peptide human β-defensin 1. *FEBS Letters*, *413*(1), 45–49. https://doi.org/10.1016/S0014-5793(97)00875-2

Ilicic, T., Kim, J. K., Kolodziejczyk, A. A., Bagger, F. O., McCarthy, D. J., Marioni, J. C., & Teichmann, S. A. (2016). Classification of low quality cells from single-cell RNA-seq data. *Genome Biology*, *17*(1), 29. https://doi.org/10.1186/s13059-016-0888-1

Israel, E. J., Taylor, S., Wu, Z., Mizoguchi, E., Blumberg, R. S., Bhan, A., & Simister, N. E. (1997). Expression of the neonatal Fc receptor, FcRn, on human intestinal epithelial cells. *Immunology*, *92*(1), 69–74. https://doi.org/10.1046/j.1365-2567.1997.00326.x

Järvi, O., & Keyriläinen, O. (1956). On the Cellular Structures of the Epithelial Invasions in the Glandular Stomach of Mice Caused by Intramural Application of 20-Methylcholantren. *Acta Pathologica Microbiologica Scandinavica*, *38*(S111), 72–73. https://doi.org/10.1111/j.1600-0463.1956.tb06739.x

Jilling, T., Simon, D., Lu, J., Meng, F. J., Li, D., Schy, R., Thomson, R. B., Soliman, A., Arditi, M., & Caplan, M. S. (2006). The Roles of Bacteria and TLR4 in Rat and Murine Models of Necrotizing Enterocolitis. *The Journal of Immunology*, *177*(5), 3273–3282. https://doi.org/10.4049/jimmunol.177.5.3273

Jones, B. D., Ghori, N., & Falkow, S. (1994). Salmonella typhimurium initiates murine infection by penetrating and destroying the specialized epithelial M cells of the Peyer's patches. *The Journal of Experimental Medicine*, *180*(1), 15–23. https://doi.org/10.1084/jem.180.1.15

Juhl, S., Hansen, M., Fonnest, G., Gormsen, M., Lambæk, I., & Greisen, G. (2017). Poor validity of the routine diagnosis of necrotising enterocolitis in preterm infants at discharge. *Acta Paediatrica*, *106*(3), 394–398. https://doi.org/10.1111/apa.13541

Kakugawa, S., Langton, P. F., Zebisch, M., Howell, S. A., Chang, T.-H., Liu, Y., Feizi, T., Bineva, G., O'Reilly, N., Snijders, A. P., Jones, E. Y., & Vincent, J.-P. (2015). Notum deacylates Wnt proteins to suppress signalling activity. *Nature*, *519*(7542), 187–192. https://doi.org/10.1038/nature14259

Keestra-Gounder, A. M., Tsolis, R. M., & Bäumler, A. J. (2015). Now you see me, now you don't: The interaction of Salmonella with innate immune receptors. *Nature Reviews Microbiology*, *13*(4), 206–216. https://doi.org/10.1038/nrmicro3428

Kim, B., Mao, J., Taketo, M. M., & Shivdasani, R. A. (2007). Phases of Canonical Wnt Signaling During the Development of Mouse Intestinal Epithelium. *Gastroenterology*, *133*(2), 529–538. https://doi.org/10.1053/j.gastro.2007.04.072

Kokkotou, E., Moss, A. C., Torres, D., Karagiannides, I., Cheifetz, A., Liu, S., O'Brien, M., Maratos-Flier, E., & Pothoulakis, C. (2008). Melanin-concentrating hormone as a mediator of intestinal inflammation. *Proceedings of the National Academy of Sciences*, *105*(30), 10613–10618. https://doi.org/10.1073/pnas.0804536105

Kolev, H. M., & Kaestner, K. H. (2023). Mammalian Intestinal Development and Differentiation—The State of the Art. *Cellular and Molecular Gastroenterology and Hepatology*, *16*(5), 809–821. https://doi.org/10.1016/j.jcmgh.2023.07.011

Komiya, T., Tanigawa, Y., & Hirohashi, S. (1998). Cloning of the Novel Gene Intelectin, Which Is Expressed in Intestinal Paneth Cells in Mice. *Biochemical and Biophysical Research Communications*, *251*(3), 759–762. https://doi.org/10.1006/bbrc.1998.9513

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., & Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods*, *16*(12), 1289–1296. https://doi.org/10.1038/s41592-019-0619-0

Krasinski, S. D., Estrada, G., Yeh, K. Y., Yeh, M., Traber, P. G., Rings, E. H., Buller, H. A., Verhave, M., Montgomery, R. K., & Grand, R. J. (1994). Transcriptional regulation of intestinal hydrolase biosynthesis during postnatal development in rats. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *267*(4), G584–G594. https://doi.org/10.1152/ajpgi.1994.267.4.G584

Kuma, A., Hatano, M., Matsui, M., Yamamoto, A., Nakaya, H., Yoshimori, T., Ohsumi, Y., Tokuhisa, T., & Mizushima, N. (2004). The role of autophagy during the early neonatal starvation period. *Nature*, *432*(7020), 1032–1036. https://doi.org/10.1038/nature03029

Le Doare, K., Holder, B., Bassett, A., & Pannaraj, P. S. (2018). Mother's Milk: A Purposeful Contribution to the Development of the Infant Microbiota and Immunity. *Frontiers in Immunology*, *9*, 361. https://doi.org/10.3389/fimmu.2018.00361

Lee, Y., El Andaloussi, S., & Wood, M. J. A. (2012). Exosomes and microvesicles: Extracellular vesicles for genetic information transfer and gene therapy. *Human Molecular Genetics*, *21*(R1), R125–R134. https://doi.org/10.1093/hmg/dds317

Leek, J. T., & Storey, J. D. (2007). Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. *PLoS Genetics*, *3*(9), e161. https://doi.org/10.1371/journal.pgen.0030161

Li, H. J., Ray, S. K., Singh, N. K., Johnston, B., & Leiter, A. B. (2011). Basic helix-loop-helix transcription factors and enteroendocrine cell differentiation. *Diabetes, Obesity and Metabolism*, *13*(s1), 5–12. https://doi.org/10.1111/j.1463-1326.2011.01438.x

Li, H., Li, C., Zhai, W., Zhang, X., Li, L., Wu, B., Yu, B., Zhang, P., Li, J., Cui, C.-P., & Zhang, L. (2022). Destabilization of TP53 by USP10 is essential for neonatal autophagy and survival. *Cell Reports*, *41*(1), 111435. https://doi.org/10.1016/j.celrep.2022.111435

Li, X., & Wang, C.-Y. (2021). From bulk, single-cell to spatial RNA sequencing. *International Journal of Oral Science*, *13*(1), 36. https://doi.org/10.1038/s41368-021-00146-0

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930. https://doi.org/10.1093/bioinformatics/btt656

Lotz, M., Gütle, D., Walther, S., Ménard, S., Bogdan, C., & Hornef, M. W. (2006). Postnatal acquisition of endotoxin tolerance in intestinal epithelial cells. *The Journal of Experimental Medicine*, *203*(4), 973–984. https://doi.org/10.1084/jem.20050625

Low, T. Y., Mohtar, M. A., Ang, M. Y., & Jamal, R. (2019). Connecting Proteomics to Next-Generation Sequencing: Proteogenomics and Its Current Applications in Biology. *PROTEOMICS*, *19*(10), 1800235. https://doi.org/10.1002/pmic.201800235

Lueschow, S. R., & McElroy, S. J. (2020). The Paneth Cell: The Curator and Defender of the Immature Small Intestine. *Frontiers in Immunology*, *11*, 587. https://doi.org/10.3389/fimmu.2020.00587

Lyons, K. E., Ryan, C. A., Dempsey, E. M., Ross, R. P., & Stanton, C. (2020). Breast Milk, a Source of Beneficial Microbes and Associated Benefits for Infant Health. *Nutrients*, *12*(4), 1039. https://doi.org/10.3390/nu12041039

Mah, A. T., Yan, K. S., & Kuo, C. J. (2016). Wnt pathway regulation of intestinal stem cells. *The Journal of Physiology*, *594*(17), 4837–4847. https://doi.org/10.1113/JP271754

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., … Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, *437*(7057), 376–380. https://doi.org/10.1038/nature03959

Marx, V. (2021). Method of the Year: Spatially resolved transcriptomics. *Nature Methods*, *18*(1), 9–14. https://doi.org/10.1038/s41592-020-01033-y

McCauley, H. A., Liu, C.-Y., Attia, A. C., Wikenheiser-Brokamp, K. A., Zhang, Y., Whitsett, J. A., & Guasch, G. (2014). TGFβ signaling inhibits goblet cell differentiation via SPDEF in conjunctival epithelium. *Development*, *141*(23), 4628–4639. https://doi.org/10.1242/dev.117804

Ménard, S., Förster, V., Lotz, M., Gütle, D., Duerr, C. U., Gallo, R. L., Henriques-Normark, B., Pütsep, K., Andersson, M., Glocker, E. O., & Hornef, M. W. (2008). Developmental switch of intestinal antimicrobial peptide expression. *The Journal of Experimental Medicine*, *205*(1), 183–193. https://doi.org/10.1084/jem.20071022

Moor, A. E., Harnik, Y., Ben-Moshe, S., Massasa, E. E., Rozenberg, M., Eilam, R., Bahar Halpern, K., & Itzkovitz, S. (2018). Spatial Reconstruction of Single Enterocytes Uncovers Broad Zonation along the Intestinal Villus Axis. *Cell*, *175*(4), 1156-1167.e15. https://doi.org/10.1016/j.cell.2018.08.063

Moxey, P. C., & Trier, J. S. (1979). Development of villus absorptive cells in the human fetal small intestine: A morphological and morphometric study. *The Anatomical Record*, *195*(3), 463–482. https://doi.org/10.1002/ar.1091950307

Muncan, V., Heijmans, J., Krasinski, S. D., Büller, N. V., Wildenberg, M. E., Meisner, S., Radonjic, M., Stapleton, K. A., Lamers, W. H., Biemond, I., Van Den Bergh Weerman, M. A., O'Carroll, D., Hardwick, J. C., Hommes, D. W., & Van Den Brink, G. R. (2011). Blimp1 regulates the transition of neonatal to adult intestinal epithelium. *Nature Communications*, *2*(1), 452. https://doi.org/10.1038/ncomms1463

Nakamura, K., Yokoi, Y., Fukaya, R., Ohira, S., Shinozaki, R., Nishida, T., Kikuchi, M., & Ayabe, T. (2020). Expression and Localization of Paneth Cells and Their α-Defensins in the Small Intestine of Adult Mouse. *Frontiers in Immunology*, *11*, 570296. https://doi.org/10.3389/fimmu.2020.570296

Neal-Kluever, A., Fisher, J., Grylack, L., Kakiuchi-Kiyota, S., & Halpern, W. (2019). Physiology of the Neonatal Gastrointestinal System Relevant to the Disposition of Orally Administered Medications. *Drug Metabolism and Disposition*, *47*(3), 296–313. https://doi.org/10.1124/dmd.118.084418

Noah, T. K., Donahue, B., & Shroyer, N. F. (2011). Intestinal development and differentiation. *Experimental Cell Research*, *317*(19), 2702–2710. https://doi.org/10.1016/j.yexcr.2011.09.006

Noah, T. K., Kazanjian, A., Whitsett, J., & Shroyer, N. F. (2010). SAM pointed domain ETS factor (SPDEF) regulates terminal differentiation and maturation of intestinal goblet cells. *Experimental Cell Research*, *316*(3), 452–465. https://doi.org/10.1016/j.yexcr.2009.09.020

Noah, T. K., & Shroyer, N. F. (2013). Notch in the Intestine: Regulation of Homeostasis and Pathogenesis. *Annual Review of Physiology*, *75*(1), 263–288. https://doi.org/10.1146/annurev-physiol-030212-183741

Nyström, E. E. L., Martinez-Abad, B., Arike, L., Birchenough, G. M. H., Nonnecke, E. B., Castillo, P. A., Svensson, F., Bevins, C. L., Hansson, G. C., & Johansson, M. E. V. (2021). An intercrypt subpopulation of goblet cells is essential for colonic mucus barrier function. *Science*, *372*(6539), eabb1590. https://doi.org/10.1126/science.abb1590

Oliveira, M. F. D., Romero, J. P., Chung, M., Williams, S. R., Gottscho, A. D., Gupta, A., Pilipauskas, S. E., Mohabbat, S., Raman, N., Sukovich, D. J., Patterson, D. M., Visium HD Development Team, De Oliveira, M. F., Romero, J. P., Williams, S. R., Patterson, D. M., Taylor, S. E. B., & Taylor, S. E. B. (2025). High-definition spatial transcriptomic profiling of immune cell populations in colorectal cancer. *Nature Genetics*, *57*(6), 1512–1523. https://doi.org/10.1038/s41588-025-02193-3

Ortolano, N. (2024). The neXt generation of single cell RNA-seq: An introduction to GEM-X technology. *The neXt Generation of Single Cell RNA-Seq: An Introduction to GEM-X Technology*. https://www.10xgenomics.com/blog/the-next-generation-of-single-cell-rna-seq-an-introduction-to-gem-x-technology

Osorio, D., & Cai, J. J. (2021). Systematic determination of the mitochondrial proportion in human and mice tissues for single-cell RNA-sequencing data quality control. *Bioinformatics*, *37*(7), 963–967. https://doi.org/10.1093/bioinformatics/btaa751

Ouellette, A. J., & Bevins, C. L. (2001). Paneth Cell Defensins and Innate Immunity of the Small Bowel: *Inflammatory Bowel Disease*, *7*(1), 43–50. https://doi.org/10.1097/00054725-200102000-00007

Pan, W.-H., Sommer, F., Falk-Paulsen, M., Ulas, T., Best, L., Fazio, A., Kachroo, P., Luzius, A., Jentzsch, M., Rehman, A., Müller, F., Lengauer, T., Walter, J., Künzel, S., Baines, J. F., Schreiber, S., Franke, A., Schultze, J. L., Bäckhed, F., & Rosenstiel, P. (2018). Exposure to the gut microbiota drives distinct methylome and transcriptome changes in intestinal epithelial cells during postnatal development. *Genome Medicine*, *10*(1), 27. https://doi.org/10.1186/s13073-018-0534-5

Paneth, J. (1887). Ueber die secernirenden zellen des dünndarm-epithels. *Archiv Für Mikroskopische Anatomie*, *31*, 113–191.

Park, J., Levic, D. S., Sumigray, K. D., Bagwell, J., Eroglu, O., Block, C. L., Eroglu, C., Barry, R., Lickwar, C. R., Rawls, J. F., Watts, S. A., Lechler, T., & Bagnat, M. (2019). Lysosome-Rich Enterocytes Mediate Protein Absorption in the Vertebrate Gut. *Developmental Cell*, *51*(1), 7-20.e6. https://doi.org/10.1016/j.devcel.2019.08.001

Pedersen, T. L. (2025). *ggforce: Accelerating "ggplot2."* https://github.com/thomasp85/ggforce

Pelaseyed, T., Bergström, J. H., Gustafsson, J. K., Ermund, A., Birchenough, G. M. H., Schütte, A., Van Der Post, S., Svensson, F., Rodríguez-Piñeiro, A. M., Nyström, E. E. L., Wising, C., Johansson, M. E. V., & Hansson, G. C. (2014). The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunological Reviews*, *260*(1), 8–20. https://doi.org/10.1111/imr.12182

Piek, E., Ju, W. J., Heyer, J., Escalante-Alcalde, D., Stewart, C. L., Weinstein, M., Deng, C., Kucherlapati, R., Böttinger, E. P., & Roberts, A. B. (2001). Functional Characterization of Transforming Growth Factor β Signaling in Smad2- and Smad3-deficient Fibroblasts. *Journal of Biological Chemistry*, *276*(23), 19945–19953. https://doi.org/10.1074/jbc.M102382200

Pinto, D., Gregorieff, A., Begthel, H., & Clevers, H. (2003). Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes & Development*, *17*(14), 1709–1713. https://doi.org/10.1101/gad.267103

Poirier, F., Chan, C.-T. J., Timmons, P. M., Robertson, E. J., Evans, M. J., & Rigby, P. W. J. (1991). The murine *H19* gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo. *Development*, *113*(4), 1105–1114. https://doi.org/10.1242/dev.113.4.1105

Pott, J., Stockinger, S., Torow, N., Smoczek, A., Lindner, C., McInerney, G., Bäckhed, F., Baumann, U., Pabst, O., Bleich, A., & Hornef, M. W. (2012). Age-Dependent TLR3 Expression of the Intestinal Epithelium Contributes to Rotavirus Susceptibility. *PLoS Pathogens*, *8*(5), e1002670. https://doi.org/10.1371/journal.ppat.1002670

Powell, D. W., Pinchuk, I. V., Saada, J. I., Chen, X., & Mifflin, R. C. (2011). Mesenchymal Cells of the Intestinal Lamina Propria. *Annual Review of Physiology*, *73*(1), 213–237. https://doi.org/10.1146/annurev.physiol.70.113006.100646

Price, A. E., Shamardani, K., Lugo, K. A., Deguine, J., Roberts, A. W., Lee, B. L., & Barton, G. M. (2018). A Map of Toll-like Receptor Expression in the Intestinal Epithelium Reveals Distinct Spatial, Cell Type-Specific, and Temporal Patterns. *Immunity*, *49*(3), 560-575.e6. https://doi.org/10.1016/j.immuni.2018.07.016

Pütsep, K., Axelsson, L.-G., Boman, A., Midtvedt, T., Normark, S., Boman, H. G., & Andersson, M. (2000). Germ-free and Colonized Mice Generate the Same Products from Enteric Prodefensins. *Journal of Biological Chemistry*, *275*(51), 40478–40482. https://doi.org/10.1074/jbc.M007816200

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., & Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods*, *14*(10), 979–982. https://doi.org/10.1038/nmeth.4402

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rappsilber, J., Mann, M., & Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols*, *2*(8), 1896–1906. https://doi.org/10.1038/nprot.2007.261

Raya Tonetti, F., Eguileor, A., & Llorente, C. (2024). Goblet cells: Guardians of gut immunity and their role in gastrointestinal diseases. *eGastroenterology*, *2*(3), e100098. https://doi.org/10.1136/egastro-2024-100098

Ren, L., Yang, M., Geng, L., Chen, P., Chen, H., Gong, S., & Li, D.-Y. (2018). Nontyphoidal *Salmonella* Gastroenteritis in a Tertiary Children's Hospital in Southern China: Characteristics and Dietary Considerations. *Gastroenterology Research and Practice*, *2018*, 1–4. https://doi.org/10.1155/2018/3097468

Rhodin, J., & Dalhamn, T. (1956). Electron microscopy of the tracheal ciliated mucosa in rat. *Zeitschrift Für Zellforschung Und Mikroskopische Anatomie*, *44*(4), 345–412. https://doi.org/10.1007/BF00345847

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47–e47. https://doi.org/10.1093/nar/gkv007

Rogier, E. W., Frantz, A. L., Bruno, M. E. C., Wedlund, L., Cohen, D. A., Stromberg, A. J., & Kaetzel, C. S. (2014). Secretory antibodies in breast milk promote long-term intestinal homeostasis by regulating the gut microbiota and host gene expression. *Proceedings of the National Academy of Sciences*, *111*(8), 3074–3079. https://doi.org/10.1073/pnas.1315792111

Ronaghi, M., Uhlén, M., & Nyrén, P. (1998). A Sequencing Method Based on Real-Time Pyrophosphate. *Science*, *281*(5375), 363–365. https://doi.org/10.1126/science.281.5375.363

Ruppert, A.-L., Keshavarz, M., Winterberg, S., Oberwinkler, J., Kummer, W., & Schütz, B. (2020). Advillin is a tuft cell marker in the mouse alimentary tract. *Journal of Molecular Histology*, *51*(4), 421–435. https://doi.org/10.1007/s10735-020-09893-6

Sanchez, J. G., Enriquez, J. R., & Wells, J. M. (2022). Enteroendocrine cell differentiation and function in the intestine. *Current Opinion in Endocrinology, Diabetes & Obesity*, *29*(2), 169–176. https://doi.org/10.1097/MED.0000000000000709

Santaolalla, R., Fukata, M., & Abreu, M. T. (2011). Innate immunity in the small intestine: *Current Opinion in Gastroenterology*, *27*(2), 125–131. https://doi.org/10.1097/MOG.0b013e3283438dea

Satam, H., Joshi, K., Mangrolia, U., Waghoo, S., Zaidi, G., Rawool, S., Thakare, R. P., Banday, S., Mishra, A. K., Das, G., & Malonia, S. K. (2023). Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology*, *12*(7), 997. https://doi.org/10.3390/biology12070997

Schlößer, S., Ullrich, A.-L., Modares, N. F., Schmitz, M. A., Schöneich, J., Zhang, K., Richter, I., Robrahn, L., Schraven, S., Nagai, J. S., Haange, S.-B., Jennings, S. A. V., Clavel, T., Rolle-Kampczyk, U., Kiessling, F., Costa, I. G., Muncan, V., Repnik, U., Von Bergen, M., … Hornef, M. W. (2025).

Salmonella infection accelerates postnatal maturation of the intestinal epithelium. *Proceedings of the National Academy of Sciences*, *122*(1), e2403344122. https://doi.org/10.1073/pnas.2403344122

Schneider, H., Pelaseyed, T., Svensson, F., & Johansson, M. E. V. (2018). Study of mucin turnover in the small intestine by in vivo labeling. *Scientific Reports*, *8*(1), 5760. https://doi.org/10.1038/s41598-018-24148-x

Schoenborn, A. A., Von Furstenberg, R. J., Valsaraj, S., Hussain, F. S., Stein, M., Shanahan, M. T., Henning, S. J., & Gulati, A. S. (2019). The enteric microbiota regulates jejunal Paneth cell number and function without impacting intestinal stem cells. *Gut Microbes*, *10*(1), 45–58. https://doi.org/10.1080/19490976.2018.1474321

Schoenenberger, C.-A., Zuk, A., Groner, B., Jones, W., & Andres, A.-C. (1990). Induction of the endogenous whey acidic protein (Wap) gene and a Wap-myc hybrid gene in primary murine mammary organoids. *Developmental Biology*, *139*(2), 327–337. https://doi.org/10.1016/0012-1606(90)90302-Y

Schubert, M., Klinger, B., Klünemann, M., Sieber, A., Uhlitz, F., Sauer, S., Garnett, M. J., Blüthgen, N., & Saez-Rodriguez, J. (2018). Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nature Communications*, *9*(1), 20. https://doi.org/10.1038/s41467-017-02391-6

Schuijers, J., Junker, J. P., Mokry, M., Hatzis, P., Koo, B.-K., Sasselli, V., van der Flier, L. G., Cuppen, E., van Oudenaarden, A., & Clevers, H. (2015). Ascl2 Acts as an R-spondin/Wnt-Responsive Switch to Control Stemness in Intestinal Crypts. *Cell Stem Cell*, *16*(2), 158–170. https://doi.org/10.1016/j.stem.2014.12.006

Schwalbe, G. (1872). Beiträge zur kenntniss der drüsen in den darmwandungen, in's besondere der brunner'schen drüsen. *Archiv Für Mikroskopische Anatomie*, *8*(1), 92–140.

Sebastián, V. P., Salazar, G. A., Coronado-Arrázola, I., Schultz, B. M., Vallejos, O. P., Berkowitz, L., Álvarez-Lobos, M. M., Riedel, C. A., Kalergis, A. M., & Bueno, S. M. (2018). Heme Oxygenase-1 as a Modulator of Intestinal Inflammation Development and Progression. *Frontiers in Immunology*, *9*, 1956. https://doi.org/10.3389/fimmu.2018.01956

Shenoy, M. K., Rico, D. M., Lorant, A. K., Touré, H., Gordon, S., Milburn, L. J., Schwensen, J. S., Cabán, M. E., & Koch, M. A. (2025). Breast milk IgG engages the mouse neonatal immune system to instruct responses to gut antigens. *Science*, *389*(6761), eado5294. https://doi.org/10.1126/science.ado5294

Shroyer, N. F., Helmrath, M. A., Wang, V. Y. –C., Antalffy, B., Henning, S. J., & Zoghbi, H. Y. (2007). Intestine-Specific Ablation of Mouse atonal homolog 1 (Math1) Reveals a Role in Cellular Homeostasis. *Gastroenterology*, *132*(7), 2478–2488. https://doi.org/10.1053/j.gastro.2007.03.047

Skrzypek, T., Valverde Piedra, J. L., Skrzypek, H., Kazimierczak, W., Biernat, M., & Zabielski, R. (2007). Gradual disappearance of vacuolated enterocytes in the small intestine of neonatal piglets. *Journal of Physiology and Pharmacology : An Official Journal of the Polish Physiological Society*, *58 Suppl 3*, 87–95.

Smith, J. L. (2003). The Role of Gastric Acid in Preventing Foodborne Disease and How Bacteria Overcome Acid Conditions. *Journal of Food Protection*, *66*(7), 1292–1303. https://doi.org/10.4315/0362-028X-66.7.1292

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J. O., Huss, M., Mollbrink, A., Linnarsson, S., Codeluppi, S., Borg, Å., Pontén, F., Costea, P. I., Sahlén, P., Mulder, J., Bergmann, O., … Frisén, J. (2016). Visualization and analysis

of gene expression in tissue sections by spatial transcriptomics. *Science*, *353*(6294), 78–82. https://doi.org/10.1126/science.aaf2403

Stanford, A. H., Gong, H., Noonan, M., Lewis, A. N., Gong, Q., Lanik, W. E., Hsieh, J. J., Lueschow, S. R., Frey, M. R., Good, M., & McElroy, S. J. (2020). A direct comparison of mouse and human intestinal development using epithelial gene expression patterns. *Pediatric Research*, *88*(1), 66–76. https://doi.org/10.1038/s41390-019-0472-y

Stephens, M. (2016). False discovery rates: A new deal. *Biostatistics*, kxw041. https://doi.org/10.1093/biostatistics/kxw041

Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., Purdom, E., & Dudoit, S. (2018). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*, *19*(1), 477. https://doi.org/10.1186/s12864-018-4772-0

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., & Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell*, *177*(7), 1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031

Sumigray, K. D., Terwilliger, M., & Lechler, T. (2018). Morphogenesis and Compartmentalization of the Intestinal Crypt. *Developmental Cell*, *45*(2), 183-197.e5. https://doi.org/10.1016/j.devcel.2018.03.024

Suter, D. M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., & Naef, F. (2011). Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science*, *332*(6028), 472–474. https://doi.org/10.1126/science.1198817

The Tabula Muris Consortium, Overall coordination, Logistical coordination, Organ collection and processing, Library preparation and sequencing, Computational data analysis, Cell type annotation, Writing group, Supplemental text writing group, & Principal investigators. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*, *562*(7727), 367–372. https://doi.org/10.1038/s41586-018-0590-4

Theis, K. R., Romero, R., Greenberg, J. M., Winters, A. D., Garcia-Flores, V., Motomura, K., Ahmad, M. M., Galaz, J., Arenas-Hernandez, M., & Gomez-Lopez, N. (2020). No Consistent Evidence for Microbiota in Murine Placental and Fetal Tissues. *mSphere*, *5*(1), e00933-19. https://doi.org/10.1128/mSphere.00933-19

Torow, N., Hand, T. W., & Hornef, M. W. (2023). Programmed and environmental determinants driving neonatal mucosal immune development. *Immunity*, *56*(3), 485–499. https://doi.org/10.1016/j.immuni.2023.02.013

Torow, N., Yu, K., Hassani, K., Freitag, J., Schulz, O., Basic, M., Brennecke, A., Sparwasser, T., Wagner, N., Bleich, A., Lochner, M., Weiss, S., Förster, R., Pabst, O., & Hornef, M. W. (2015). Active suppression of intestinal CD4+TCRαβ+ T-lymphocyte maturation during the postnatal period. *Nature Communications*, *6*(1), 7725. https://doi.org/10.1038/ncomms8725

Tourneur, E., & Chassin, C. (2013). Neonatal Immune Adaptation of the Gut and Its Role during Infections. *Clinical and Developmental Immunology*, *2013*, 1–17. https://doi.org/10.1155/2013/270301

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S., & Rinn, J. L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology*, *32*(4), 381–386. https://doi.org/10.1038/nbt.2859

Tsang, D. K. L., Wang, R. J., De Sa, O., Ayyaz, A., Foerster, E. G., Bayer, G., Goyal, S., Trcka, D., Ghoshal, B., Wrana, J. L., Girardin, S. E., & Philpott, D. J. (2022). A single cell survey of the microbial

impacts on the mouse small intestinal epithelium. *Gut Microbes*, *14*(1), 2108281. https://doi.org/10.1080/19490976.2022.2108281

Tyner, J. W., Uchida, O., Kajiwara, N., Kim, E. Y., Patel, A. C., O'Sullivan, M. P., Walter, M. J., Schwendener, R. A., Cook, D. N., Danoff, T. M., & Holtzman, M. J. (2005). CCL5-CCR5 interaction provides antiapoptotic signals for macrophage survival during viral infection. *Nature Medicine*, *11*(11), 1180–1187. https://doi.org/10.1038/nm1303

Van Best, N., Rolle-Kampczyk, U., Schaap, F. G., Basic, M., Olde Damink, S. W. M., Bleich, A., Savelkoul, P. H. M., Von Bergen, M., Penders, J., & Hornef, M. W. (2020). Bile acids drive the newborn's gut microbiota maturation. *Nature Communications*, *11*(1), 3692. https://doi.org/10.1038/s41467-020-17183-8

Van Den Brink, S. C., Sage, F., Vértesy, Á., Spanjaard, B., Peterson-Maduro, J., Baron, C. S., Robin, C., & Van Oudenaarden, A. (2017). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nature Methods*, *14*(10), 935–936. https://doi.org/10.1038/nmeth.4437

Van Es, J. H., Haegebarth, A., Kujala, P., Itzkovitz, S., Koo, B.-K., Boj, S. F., Korving, J., Van Den Born, M., Van Oudenaarden, A., Robine, S., & Clevers, H. (2012). A Critical Role for the Wnt Effector Tcf4 in Adult Intestinal Homeostatic Self-Renewal. *Molecular and Cellular Biology*, *32*(10), 1918–1927. https://doi.org/10.1128/MCB.06288-11

Vázquez-Carretero, M. D., Palomo, M., García-Miranda, P., Sánchez-Aguayo, I., Peral, M. J., Calonge, M. L., & Ilundain, A. A. (2014). Dab2, Megalin, Cubilin and Amnionless Receptor Complex Might Mediate Intestinal Endocytosis in the Suckling Rat. *Journal of Cellular Biochemistry*, *115*(3), 510–522. https://doi.org/10.1002/jcb.24685

Venkatraman, A., Yu, W., Nitkin, C., & Sampath, V. (2021). Intestinal Stem Cell Development in the Neonatal Gut: Pathways Regulating Development and Relevance to Necrotizing Enterocolitis. *Cells*, *10*(2), 312. https://doi.org/10.3390/cells10020312

Walker, G. T., Gerner, R. R., Nuccio, S., & Raffatellu, M. (2023). Murine Models of *Salmonella* Infection. *Current Protocols*, *3*(7), e824. https://doi.org/10.1002/cpz1.824

Wallaeys, C., Garcia-Gonzalez, N., & Libert, C. (2023). Paneth cells as the cornerstones of intestinal and organismal health: A primer. *EMBO Molecular Medicine*, *15*(2), e16427. https://doi.org/10.15252/emmm.202216427

Wang, X., Li, S., Liu, Y., Kuang, X., Chen, J., Yu, P., Zhao, L., Zhang, Z., Huang, M., Liu, L., Zhang, M., Wang, Y., Ji, W., Li, J., Zhang, L., & Chen, Y.-G. (2025a). Positional BMP signaling orchestrates villus length in the small intestine. *Nature Communications*, *16*(1), 5461. https://doi.org/10.1038/s41467-025-60643-2

Wang, Y., Lin, Z., & Wang, T. (2025b). InterVelo: A mutually enhancing model for estimating pseudotime and RNA velocity in multi-omic single-cell data. *Bioinformatics*, *41*(10), btaf500. https://doi.org/10.1093/bioinformatics/btaf500

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57–63. https://doi.org/10.1038/nrg2484

Wei, Y., Liu, M., Li, X., Liu, J., & Li, H. (2018). Origin of the Autophagosome Membrane in Mammals. *BioMed Research International*, *2018*, 1012789. https://doi.org/10.1155/2018/1012789

Weiler, P., Van Den Berge, K., Street, K., & Tiberi, S. (2023). A Guide to Trajectory Inference and RNA Velocity. In R. A. Calogero & V. Benes (Eds.), *Single Cell Transcriptomics* (Vol. 2584, pp. 269–292). Springer US. https://doi.org/10.1007/978-1-0716-2756-3_14

Weström, B., Arévalo Sureda, E., Pierzynowska, K., Pierzynowski, S. G., & Pérez-Cano, F.-J. (2020). The Immature Gut Barrier and Its Importance in Establishing Immunity in Newborn Mammals. *Frontiers in Immunology*, *11*, 1153. https://doi.org/10.3389/fimmu.2020.01153

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Williams, E., & Scott, J. P. (1954). The Development of Social Behavior Patterns in the Mouse, in Relation To Natural Periods 1). *Behaviour*, *6*(1), 35–64. https://doi.org/10.1163/156853954X00031

Wiśniewski, J. R., Zougman, A., Nagaraj, N., & Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nature Methods*, *6*(5), 359–362. https://doi.org/10.1038/nmeth.1322

Xian, L., Georgess, D., Huso, T., Cope, L., Belton, A., Chang, Y.-T., Kuang, W., Gu, Q., Zhang, X., Senger, S., Fasano, A., Huso, D. L., Ewald, A. J., & Resar, L. M. S. (2017). HMGA1 amplifies Wnt signalling and expands the intestinal stem cell compartment and Paneth cell niche. *Nature Communications*, *8*(1), 15008. https://doi.org/10.1038/ncomms15008

Xu, S., Hu, E., Cai, Y., Xie, Z., Luo, X., Zhan, L., Tang, W., Wang, Q., Liu, B., Wang, R., Xie, W., Wu, T., Xie, L., & Yu, G. (2024). Using clusterProfiler to characterize multiomics data. *Nature Protocols*, *19*(11), 3292–3320. https://doi.org/10.1038/s41596-024-01020-z

Yates, J., Kraft, A., & Boeva, V. (2025). Filtering cells with high mitochondrial content depletes viable metabolically altered malignant cell populations in cancer single-cell studies. *Genome Biology*, *26*(1), 91. https://doi.org/10.1186/s13059-025-03559-w

Zhang, K., Dupont, A., Torow, N., Gohde, F., Leschner, S., Lienenklaus, S., Weiss, S., Brinkmann, M. M., Kühnel, M., Hensel, M., Fulde, M., & Hornef, M. W. (2014). Age-Dependent Enterocyte Invasion and Microcolony Formation by Salmonella. *PLoS Pathogens*, *10*(9), e1004385. https://doi.org/10.1371/journal.ppat.1004385

Zhang, K., Hornef, M. W., & Dupont, A. (2015). The intestinal epithelium as guardian of gut barrier integrity: The epithelium as a barrier to infection. *Cellular Microbiology*, *17*(11), 1561–1569. https://doi.org/10.1111/cmi.12501

Zhang, K., Riba, A., Nietschke, M., Torow, N., Repnik, U., Pütz, A., Fulde, M., Dupont, A., Hensel, M., & Hornef, M. (2018). Minimal SPI1-T3SS effector requirement for Salmonella enterocyte invasion and intracellular proliferation in vivo. *PLOS Pathogens*, *14*(3), e1006925. https://doi.org/10.1371/journal.ppat.1006925

Zhang, M., Liao, Y., & Lönnerdal, B. (2016). Milk growth factors and expression of small intestinal growth factor receptors during the perinatal period in mice. *Pediatric Research*, *80*(5), 759–765. https://doi.org/10.1038/pr.2016.150

Zhang, Y., Parmigiani, G., & Johnson, W. E. (2020). ComBat-seq: Batch effect adjustment for RNA-seq count data. *NAR Genomics and Bioinformatics*, *2*(3), lqaa078. https://doi.org/10.1093/nargab/lqaa078

Zhao, A., Lu, W., & De Leeuw, E. (2015). Functional synergism of Human Defensin 5 and Human Defensin 6. *Biochemical and Biophysical Research Communications*, *467*(4), 967–972. https://doi.org/10.1016/j.bbrc.2015.10.035

Zwick, R. K., Kasparek, P., Palikuqi, B., Viragova, S., Weichselbaum, L., McGinnis, C. S., McKinley, K. L., Rathnayake, A., Vaka, D., Nguyen, V., Trentesaux, C., Reyes, E., Gupta, A. R., Gartner, Z. J., Locksley, R. M., Gardner, J. M., Itzkovitz, S., Boffelli, D., & Klein, O. D. (2024). Epithelial zonation along the mouse and human small intestine defines five discrete metabolic domains. *Nature Cell Biology*, *26*(2), 250–262. https://doi.org/10.1038/s41556-023-01337-z

14. Eidesstattliche Erklärung

Johannes Schöneich erklärt hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Promovierender dieser Fakultät und Universität angefertigt;

2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;

3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;

4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;

5. Alle wesentlichen Quellen von Unterstützung wurden benannt;

6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;

7. Ein Teil dieser Arbeit wurde vor deren Einreichung veröffentlicht in: Schöneich, J., Dupont, A., Schlößer, S., Schmitz, M. A., Richter, I., Gubbi, N. M. K. P., Cheng, M., Maié, T., Jäverfelt, S., Imdahl, F., Toussaint, C., Saliba, A.-E., Kuppe, C., Pelaseyed, T., Costa, I. G., & Hornef, M. W. (2025). *Developmental factors drive the compartmentalised and discontinuous maturation of the small intestinal epithelium during the early postnatal period*. bioRxiv. https://doi.org/10.1101/2025.05.15.654215

Datum und Unterschrift