



Research article

Establishment and transcriptomic analyses of a cattle rumen epithelial primary cells (REPC) culture by bulk and single-cell RNA sequencing to elucidate interactions of butyrate and rumen development

Shudai Lin ^{a,b}, Lingzhao Fang ^{a,c}, Xiaolong Kang ^{a,d}, Shuli Liu ^{a,e}, Mei Liu ^{a,f}, Erin E. Connor ^a, Ransom L. Baldwin VI ^a, George Liu ^{a,**}, Cong-Jun Li ^{a,*}^a Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD, USA^b Guangdong Provincial Key Lab of Agro-Animal Genomics and Molecular Breeding and Key Lab of Chicken Genetics, Breeding and Reproduction, Ministry of Agriculture, College of Animal Science of South China Agricultural University, Guangzhou, 510642, China^c Medical Research Council Human Genetics Unit at the Medical Research Council Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, United Kingdom^d College of Agriculture, Ningxia University, Yinchuan, 750021, China^e College of Animal Science and Technology, China Agricultural University, Beijing, 100193, China^f College of Animal Science and Technology, Shaanxi Key Laboratory of Agricultural Molecular Biology, Northwest A&F University, Yangling, 712100, Shaanxi, China

ARTICLE INFO

Keywords:

Agriculture
Bioinformatics
Cell biology
Developmental biology
Butyrate
Epigenetic regulation
Epithelium
Rumen
Transcriptome

ABSTRACT

As a critical and high-value tool to study the development of rumen, we established a stable rumen epithelial primary cell (REPC) culture from a two-week-old Holstein bull calf rumen epithelial tissue. The transcriptomic profiling of the REPC and the direct effects of butyrate on gene expression were assessed. Correlated gene networks elucidated the putative roles and mechanisms of butyrate action in rumen epithelial development. The top networks perturbed by butyrate were associated with epithelial tissue development. Additionally, two critical upstream regulators, E2F1 and TGFB1, were identified to play critical roles in the differentiation, development, and growth of epithelial cells. Significant expression changes of upstream regulators and transcription factors provided further evidence in support that butyrate plays a specific and central role in regulating genomic and epigenomic activities influencing rumen development. This work is the essential component to obtain a complete global landscape of regulatory elements in cattle and to explore the dynamics of chromatin states in rumen epithelial cells induced by butyrate at early developmental stages.

1. Introduction

The rumen is incompletely developed both physically and metabolically at birth and requires the establishment of a viable ruminal fermentation before the maturation process proceeds. Neonatal rumen exhibits poorly developed papillae without a high degree of keratinization characteristic of the mature organ. Metabolically, the rumen of neonates is essentially nonfunctional relative to its mature ketogenic capacity [1]. In production settings, one of the most critical factors affecting calf health during rearing and the key to successful weaning is ensuring healthy rumen development. The rumen is central to feed efficiency, methane emission, and productive performance in ruminants. Rumen microbes digest simple and complex carbohydrates (fiber) and

convert them into short-chain fatty acids (SCFAs, mainly acetic, propionic, and butyric acids), and in fact, SCFAs can provide 50 to 70 percent of cattle energy requirements. Prior to weaning the rumen must differentiate the epithelial layer to facilitate absorption and metabolism of SCFAs. Butyrate plays a specific and central role in regulating genomic and epigenomic activities influencing rumen development, nutrient utilization, and rumen function [2, 3].

An increase in butyric acid concentration in the rumen is highly correlated with the enlargement of the ruminal epithelium absorptive surface area [4] and the associated metabolic changes resulting in tissue capacity for the oxidation of SCFA in the ketogenesis pathway [5, 6]. The microbiome-driven generation of butyrate mediated the regulation of growth-related genes in the ruminal epithelium [7]. Butyrate stimulates

* Corresponding author.

** Corresponding author.

E-mail addresses: George.liu@usda.gov (G. Liu), Congjun.li@ars.usda.gov (C.-J. Li).<https://doi.org/10.1016/j.heliyon.2020.e04112>

Received 18 February 2020; Received in revised form 28 April 2020; Accepted 28 May 2020

2405-8440/Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the process resulting in functional (physical and metabolic) development of ruminal epithelium [2, 8].

Recently, we demonstrated that butyrate plays a prominent role in regulating genetic and epigenetic activities influencing cellular functions [9, 10, 11] in cultured MDBK (The Madin-Darby bovine kidney) cells. Butyrate is also reported to promote the growth of the colonic epithelium, but it predominately elicits inhibitory effects on colorectal cancer growth in humans [12, 13]. The impact of directly fed butyrate (in milk or top-dressed in feeds) on calf performance and functional capacity of rumen tissues is actively being investigated as a putative on-farm additive to enhance rumen development [3]. Although rumen development has attracted a lot of attention recently, a basic understanding of rumen morphology and physiology, as well as the genetic activities for the necessity of those biological activities, remain elusive. The hurdles of investigation of rumen function and development are not only the cost and ethical limitations impacting *in vivo* study but also the lack of adequate *in vitro* model such as a well-defined and characterized rumen epithelial cell line.

The cell is the essential unit of living organisms and possesses specific information of the whole organism. Primary cell culture has been extensively used in basic biological research as model systems. Those primary cell cultures model systems hold great promise for prosperous information on many aspects of biology. The use of primary cells also allows the researchers to avoid complications in using animal models, such as availability, cost, and ethics. Currently, there is no such primary cell or cell-line of rumen epithelial origin available worldwide.

Primary cell cultures provide excellent model systems for studying the normal cell functions [14]. Complete characterization of the established primary cell culture to meet the crucial genotyping information is inevitably required for an *in vitro* cell experimental system. The transcriptome is the fundamental determinant of the phenotype of a cell [15]. An objective of this research was to establish an *in vitro* model for evaluation of rumen epithelial function and developmental regulation, as well as to characterize the functional roles of butyrate in gene regulation and rumen development. In this report, we established a stable primary cell culture from rumen epithelium (REPC) and utilized both bulk-cell and single-cell RNA sequencing technologies to profile the transcriptome of rumen epithelial cells. Additionally, we treated the primary rumen epithelial cells with media containing butyrate to identify differentially expressed genes (DEGs) affected by butyrate and to characterize the genes and gene networks putatively related to rumen growth and development and to better elucidate butyrate related mechanisms regulating rumen developmental processes. This work is the essential component to obtain a complete global landscape of regulatory elements in cattle and to explore the dynamics of chromatin states in rumen epithelial cells induced by butyrate at early developmental stages [16].

2. Materials and methods

2.1. Rumen epithelial cell isolation

Animal care and tissue isolation work were approved by the Beltsville Area Animal Care and Use Committee Protocol Number 07–025. The methods for epithelial cell isolation and culture have been described previously [17]. Briefly, Rumen epithelial tissue was collected from a two-week-old Holstein bull calf fed with milk replacer only. The epithelial layer of the rumen tissue was manually separated from the muscular layer and rinsed in tap water to remove residual feed particles. Samples were further rinsed in ice-cold saline. The tissue was then added to a 50 ml digestion solution (2% trypsin and 1.15 mmol CaCl₂ in phosphate-buffered saline) and then was incubated in 37 °C incubator for 15 min. Rumen epithelial fragments generally underwent 5–6 cycles of digestion with fresh trypsin solution. The first two rounds of digestion solution were discarded, and the 3rd, fourth and fifth rounds of digestion were collected and combined. Cells then subjected to three wash cycles with sterile PBS with antibiotic-antimycotic (100 units/ml of Penicillin G

sodium, streptomycin sulfate, 0.25µg amphotericin B as Fungizone, MP Biomedical, Inc). Cells were counted using a hemocytometer, and cell viabilities were estimated by trypan blue dye exclusion.

2.2. Establishment of rumen primary epithelial cell (REPC) culture

Isolated rumen epithelial cells were plated in a 25 cm plate at a density of 1 million cells/dish in DMEM with antibiotic-antimycotic and 5% fetal bovine serum (DMEM-FBS). After 24h in culture, the cell media were removed and replaced with fresh DMEM-FBS. Subsequently, cell media was changed every 48h until the cells achieved confluence (4–7 days). Cells were removed from the dish by trypsinization, quantified and reseeded for treatment, or frozen in liquid nitrogen for future culture. Photographs in Figure 1 depict the cross-section of the rumen tissue, the inner surface of the rumen and the cultured cells isolated from the rumen tissue.

2.3. Butyrate treatment of REPC

To test the response of the primary rumen epithelial cells, 5 mM of butyrate was added to the culture for 24 h before harvesting for RNA-Sequencing. Based on results from previous *in vitro* experiments [10], treatment of 5 mM butyrate of bovine cells can induce significant changes in transcription activities of cells without inducing significant apoptosis. Accordingly, REPC culture was treated with 5 mM butyrate when cells reached 50% confluence for 24 h during the exponential phase of growth. Three replicate flasks of cells for both treatment and control groups (a total of 6 samples) were prepared for final RNA extraction and RNA sequencing. The gene expression value was based on the average of replicates.

2.4. Library preparation and whole transcriptome sequencing

The RNA extraction procedure was reported previously [18]. After quality control (QC) procedures, individual RNA-Seq libraries were pooled after indexing with their respective sample-specific 6-bp (base pairs) adaptors and sequenced at 50bp/single sequence read using an Illumina HiSeq 2500 sequencer (Illumina, Inc. San Diego, CA). RNA library preparation and sequence were performed by RNA-sequencing service of Novogene Corporation Inc, UC Davis sequencing center.

Single-cell RNA-Seq: Single-cell RNA sequencing enables the high-resolution transcriptome profiling of a single cell and has broad utility for investigating developmental processes and gene regulatory networks, and ultimately, for revealing intricate gene expression patterns within cell cultures, tissues, and organs. In this study, single cells were randomly isolated using QIAscout device (QIAGEN) with a high-density microwell array that can be used to isolate and recover individual cells from a cell suspension. Single cells were randomly selected following the manufacturer's instruction. The SMARTer kit (Takara Bio, USA) was used for single-cell RNA amplification, which reduces amplification costs, improves amplification rates, and has been utilized in multiple publications [19, 20, 21].

2.5. RNA-seq data analysis

The computational pipeline for expression quantification is based on STAR aligner [22] and Cufflinks software tool [23, 24]. The pipeline is recommended in a recent review paper (see Figure 1, left panel) [25]. Reads from RNA-Seq were subjected to quality control using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>; version 0.11.4), quality trimmiOng using Trim_Galore (version 0.4.1) and aligned to cow reference genome (Bos taurus UMD3.1.1/bosTau8) using STAR (version 020201; options: `-outSAMAttrIHstart 0 -outSAMstrandField intronMotif -outFilterIntronMotifs RemoveNoncanonical -alignIntronMin 20 -alignIntronMax 1000000 -outFilterMultimapNmax 1`) [22]. Duplicated reads were discovered using Picard tools (version 1.119) and

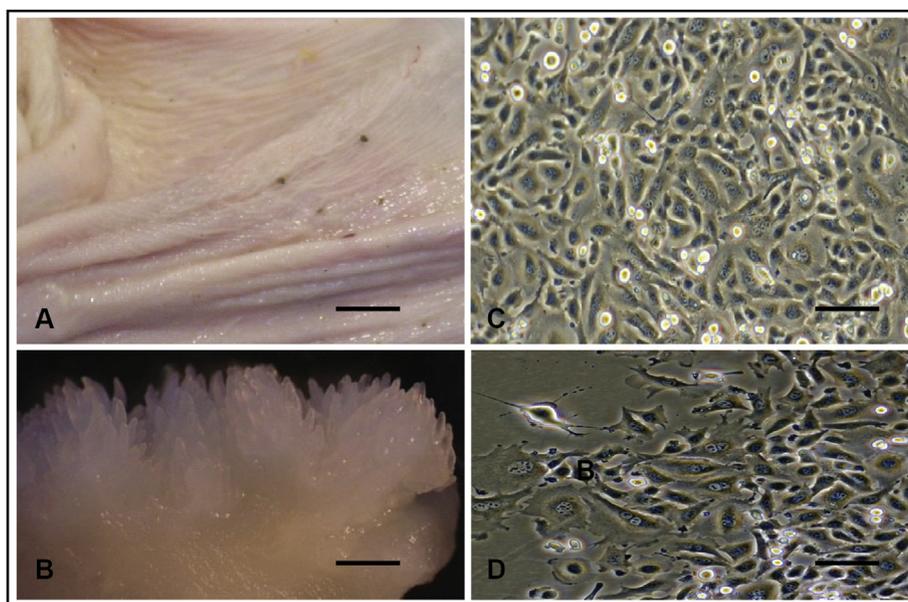


Figure 1. Rumen epithelial tissue and cells in primary culture from a two-week-old calf. A. The inner surface of the rumen of two-week-old calf (measuring bar = 5 cm), B. A magnification of the inner surface of the rumen (measuring bar = 1cm), C. Cells isolated from rumen epithelium were growing in a culture dish (confluence, measuring bar = 50 μ m). D. Cells isolated from rumen epithelium were growing in a culture dish (colony, measuring bar = 50 μ m).

removed. Gene annotations (gff file; version UMD_3.1.1) were obtained from NCBI. Cufflinks version 2.2.1 was used to estimate the expression level of each detected gene or Fragments Per Kilobase Million (FPKM) value [23].

In this study, the CLC Genomics Workbench (v12; Qiagen Bioinformatics) was used for further RNA-Seq data analysis. Trimmed reads were aligned to the bovine reference genome (BosTau UMD3.1). Gene expression levels of mapped reads were normalized as reads per kilobase of exon model per million mapped reads (RPKM) using the CLC transcriptomic analysis tool. To ensure the accuracy of estimated RPKM values and remove the auxiliary data, only genes with RPKM >1 in at least one sample was analyzed. Expression levels of each gene in all samples were log₂ converted in the following analysis. Principal component analysis (PCA), heatmap, DEGs, Venn diagram and gene ontology (GO) analysis of DEGs were all performed using CLC genomics workbench (Figure 2). The enrichment of specific GO terms was determined based on the Fisher exact test. DEGs were defined only if the corresponding P values were less than 0.05 and the false discovery rate

(FDR) was less than 0.05 with a fold change of log₂-converted absolute RPKM larger than 2. Pearson's correlation coefficient was calculated for all genes to each pattern. Thus, genes that contributed most to separate different cell groups were determined.

As described before [26], the analysis of canonical pathways identified the pathways from the IPA library of canonical pathways that were most significantly represented in the data set. The significance of the association between the data set and the canonical pathway was measured in 2 ways. One was to use a ratio of the number of genes from the data set that map to the pathway divided by the total number of genes that map to the canonical pathway. The other was to use the Fisher exact test to calculate a P-value to determine the probability of the association between the genes in the data set and the canonical pathway.

Pathway analysis and network generation: The set of gene identifiers and corresponding expression values was uploaded into the IPA application. Each gene identifier was mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base. These genes, called focus genes, were overlaid onto a global molecular network developed from

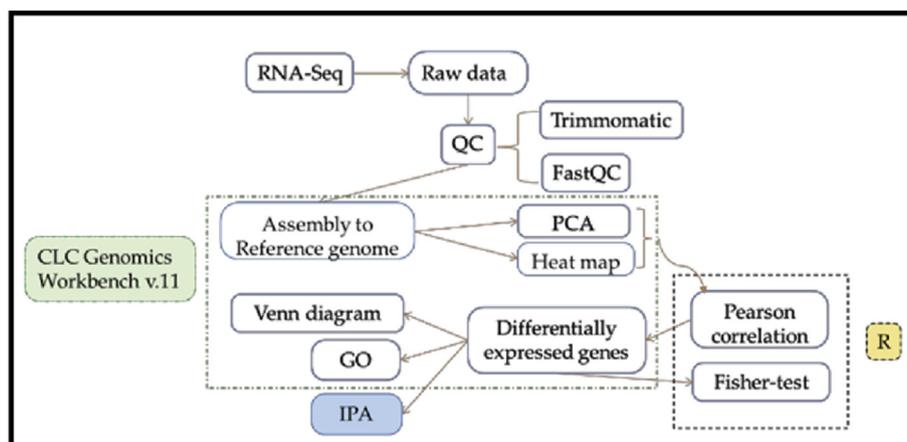


Figure 2. Bioinformatics flowchart of tools and methods used to process and analyze the RNA-Seq data and produce the transcriptome. QC: quality control; PCA: principal component analysis; GO: gene ontology; IPA: Ingenuity Pathway Analysis (Qiagen Bioinformatics).

information contained in the Ingenuity Pathways Knowledge Base. Networks of these focus genes were then algorithmically generated based on their connectivity in the IPA database.

3. Results

3.1. Transcriptomic profiling of REPC shows the high level of E-cadherin and low-level N-E-cadherin expression

From six replicates of REPC cultures (three standard cell culture and three culture treated with butyrate) and nine single cells isolated from each group (total 18 cells), RNA-Seq reads were mapped and quantified for each sample. The number of genes expressed per sample was 13111 for bulk cell samples and 12531 for the single cells.

To confirm that cells in culture are of epithelial origin, the expression of E-cadherins, the epithelial cell biomarker, was assessed. Highly expressed epithelial cell's biomarker E-cadherin (*CDH1*) and low-level of N-cadherin (*CDH2*) confirmed the epithelial origin of the established REPC cells. We observed our primary cell cultures of rumen exhibited a high level of expression of *CDH1* concomitant with low expression of *CDH2* (Table 1). Notably, butyrate treatment did not induce changes the pattern in the expression of *CDH1* and *CDH2*, yet it did significantly change the expression of some members of the cadherin family (Table 2).

3.2. The treatment of butyrate induces significant changes in transcriptomic profiles of REPC

Based on the mapped-count matrix of the Ensembl annotation, we systematically investigated differentially expressed genes induced by butyrate-treatment in REPCs. Using high stringency, (false discovery rate (FDR) ≤ 0.05 and \log_2 fold change of RPKM absolute value >2), 1977 genes were identified as differentially expressed between untreated REPC and butyrate-treated REPC (Fisher's test P value of $6.44e-126$) in bulk cell samples. The DEGs were displayed in a volcano plot (Figure 3A). The hierarchical clustering of selected genes significantly regulated by butyrate in the REPC was generated and shown in Figure 3B. The analyses also demonstrated that group-derived transcriptomes are reproducible, as the biological replicates cluster together and the transcriptomes are different between the groups.

3.3. Essential gene network and pathways associated with the development of cattle epithelial tissue

To further investigate the potential functions of the DEGs induced by butyrate infusion, we performed a Gene Ontology (GO) enrichment analysis. The changes in GO terms from enrichment analysis reflects the impact of butyrate treatment on rumen epithelial transcriptome. The GO terms in the biological processes enriched in DEGs from the bulk cell study are presented in Table 3. Mitotic cell cycle process, cell cycle process, and nucleic acid metabolic process were the most highly enriched GO terms in rumen epithelial bulk cells (24 h) after butyrate treatment. All the genes represented in each GO terms are presented in Supplementary Table 1.

By performing the function and pathway analysis of DEGs using Ingenuity Pathways Analysis (IPA), we found that significantly affected gene networks were associated with cell cycle, cell morphology, and embryonic development (Table 4, Supplementary Table 2).

Specifically, 33 DEGs were identified, which are known to involve in embryonic development through the *WNT* pathway and 26 DEGs involved in lipid metabolism by P38 mitogen-activated protein kinase (*MAPK*) pathway (Table 4, Supplementary Table 1). Transcription regulators *E2F1*, troponin I1/3 (*TNNI1/3*) and cyclin B were identified which affect cell cycle; Janus kinase and microtubule interacting protein 2 (*JAKMIP2*), c-Jun N-terminal kinase (*Jnk*) and nuclear factor-kappaB (*NFkB*) are related to cell-to-cell signaling and interaction (Figure 4 A, B, and C), Supplementary Table 2). Interestingly, there was also a network identified with all down-regulated genes which are associated with the regulation of cell cycle and DNA replication (Figure 4 D).

3.4. Essential genes associated with cattle epithelial tissue development

Two key upstream regulators appear crucial for epithelial tissue development. Namely, *E2F1* which plays regulatory role in development (P-value = $2.79E-09$), differentiation (P-value = $4.67E-11$) and growth (P-value = $3.54E-11$) of epithelial tissue; and *TGFB1* which likewise plays known functions in development (P-value = $1.70E-09$), differentiation (P-value = $1.04E-13$) and growth (P-value = $3.54E-11$) of epithelial tissue (Figure 5). In these two key pathways, although *CCNE1* and *HDAC1* genes are not remarkably changed, there were 10 DEG

Table 1. Epithelial cell biomarkers CDH1 and CDH2 by FPKM from the bulk-cell transcriptomic profiling with and without butyrate treatment.

Genes	Bulk-C1	Bulk-C2	Bulk-C3	Bulk-BT1	Bulk-BT2	Bulk-BT3
CDH1	52.85	52.72	54.38	58.98	55.52	54.13
CDH2	1.02	1.92	2.31	1.7	1.8	1.58
CDH4	–	–	–	0.52	0.33	0.09
CDH6	–	–	0.04	0.04	–	–

“–” represents the FPKM of the target gene was not detectable; Bulk-C and Bulk-BT represent bulk REPC and butyrate-treated REPC, respectively, *CDH*: cadherin.

Table 2. CDH (cadherin) superfamily gene expression changes from the bulk cell cultures with 24 h butyrate treatment.

Gene name	Log ₂ Fold change	P value	FDR p-value
CDH11	0.51	1.42E-06	5.59E-06
CDH13	0.63	0	0
CDH15	8.23	0	0
CDH17	2.74	0	0
CDH23	1.34	0.01	0.03
CDH24	0.49	4.55E-05	1.52E-04
CDH3	1.38	7.95E-07	3.20E-06
CDH4	6.84	1.10E-06	4.38E-06

“Log₂ fold change” represents the DEGs value between Bulk-BT and Bulk-C group.

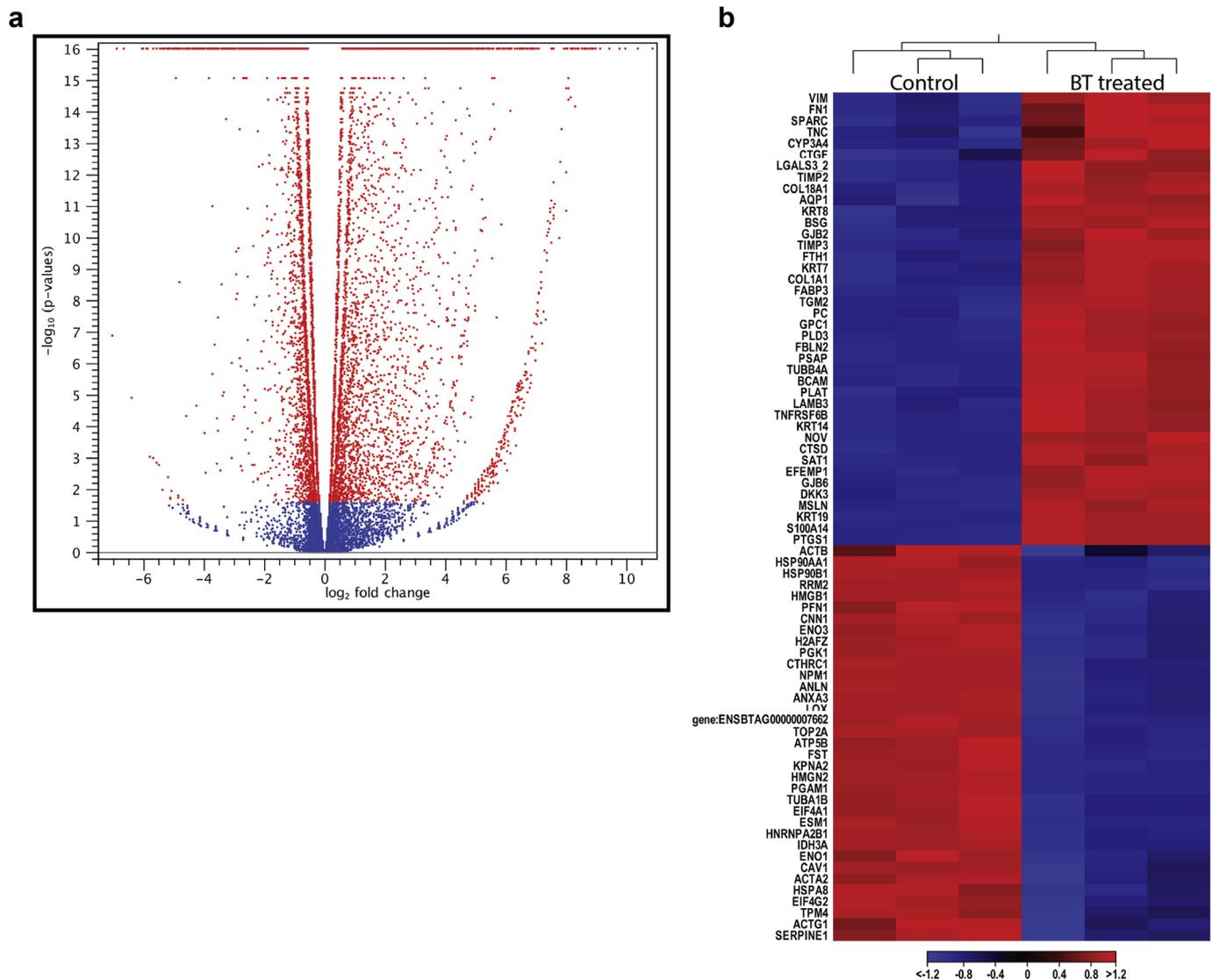


Figure 3. A Volcano plot displaying the DEGs between butyrate-treated cells (Bulk-BT) and control bulk cells (Bulk-C). Red dots represent DEGs with log10 FDR value >1.3. B. The hierarchical heatmap (Euclidean distance) of selected DEGs in bulk cells RNA sequencing following butyrate treatment of REPC. The heatmap shows the most differentially expressed genes, as well as the consistency within the control and butyrate-treated groups.

increased (FDR <0.05) which are known to be regulated by *E2F1* and *TGFB1* (Table 5). Moreover, two transcription factors, grainy head-like protein, two homologs (*GRHL2*) and ETS-related transcription factor 3 (*ELF3*), were up-regulated by butyrate treatment, with the log2 fold change of RPKM 29.42 (FDR = 0.02) and 38.53 (FDR = 0.00), respectively.

3.5. Single-cell RNA-Sequencing highlights divergence in gene expression among the REPC population

The putative biomarkers of epithelial cells were not only detected in bulk cell samples but were also present in each single cell sample (Supplementary Table 1) we tested. The principal component analysis (PCA)

Table 3. Top GO* terms in the biological process significantly impacted by butyrate infusion.

GO term	Description	Ratio	P value	FDR P value
1903047	Mitotic cell cycle process	34/34	1.53E-08	5.25E-05
22402	Cell cycle process	44/42	3.36E-08	5.79E-05
90304	Nucleic acid metabolic process	50/44	6.96E-06	6.61E-03
6139	Nucleobase-containing compound metabolic process	68/57	7.67E-06	6.61E-03
33554	Cellular response to stress	37/34	1.00E-05	6.89E-03
44260	Cellular macromolecule metabolic process	96/76	1.43E-05	7.81E-03
6259	DNA metabolic process	36/33	1.59E-05	7.81E-03

*GO, gene ontology. Ratio = the number of all genes assigned to this GO term to the number of significantly regulated genes in the data set that can be assigned to this GO term.

Table 4. Top five gene networks of DEGs induced by butyrate detected in bulk cell RNA sequencing.

ID	Genes in Network	Focus Molecules	Top Functions
1	AMY2B,AURKA, CCNB1,Cdc2,CDC25B,CENPU,CEP55,CIT,E2F1,ECT2,ERCC6L, Gamma tubulin, GINS1,GM2A,KIF23,KIF2C,KIF4A,KIF4C,MTHFD1,NCAPD3,NCAPG, NCAPG2,NCAPH,NDC80,PLK1, POLA2,PRC1,RACGAP1,SGO2,SMC2,SMC4,TACC3,TRPV2,UHRF1,ZYG11A	33	Cell Cycle, Cellular Assembly and Organization, DNA Replication, Recombination, and Repair
2	ADGRL3,APBA2,BIRC5,CDCA8,CHGA, CHGB,CLMN,CPE,Creb,CYP17A1,EMX2,ESPL1,GDPD3,GPBAR1, GPR85,GSTM3,GZMM, Hdac,KCNH2,KLF9,NTRK1,ONECUT2,OTX1,OXT,PAX2,PCSK1,PSMC3IP,RET, SHANK1,SLC5A8,SSTR2,TRAIP,WNT1,WNT4,WNT9B	33	Embryonic Development, Organ Development, Organismal Development
3	ATAD5,AURKB, BARD1,BLM,BRCA1,BRCA2,BRIP1,CHEK2,CHTF18,DNA2,DSCC1,DTL,FANCB, FANCG, FEN1,FIGNL1,I kappa b kinase, Igm, MCM8,PALB2,PARBP, PCLAF,PCNA, POLH,RAD18,RAD51,RAD54B, RAD54L,RECQL4,RFC3,RFC4,RPA,SMC1A,TOBP1,XRCC2	32	Cell Cycle, DNA Replication, Recombination, and Repair, Cell Morphology
4	ADGRF5,ADORA2A,ATP2B3,C3,CENPE, CENPF,CFB,CFD,COCH, Collagen type II, CXCL2, CXCL3, DSN1,FBL,GABRB3,GN7,GNL3,GPLD1,HRH3,MAC,MST1R,NPHS1,NPM1,NUF2,OGG1,Pde4,PMF1/PMF1-BGLAP, SERPING1, SPC24,SPC25, ST14, SYT11,TNNI1,TNNI3,ZWINT	32	Cell Cycle, Cellular Assembly and Organization, DNA Replication, Recombination, and Repair
5	ABCG5,ACACB,C2,CBLC, CIDEB,CLCA1,CREB3L3,CYP11A1,CYP2F1,DHH,EBF1,ELOVL3,ENO2,FBP2,FLT4,FOXA3, GGT1,GGT5,GRIN2C,ITGA9, KCP,LCN2,MLXIPL,MMP15,Mucin,Nr1h,OSTN, PEPCK,PPARGC1A,RORA, RORC, SERPINF1,SLC29A4,SYT2,SYT7	32	Metabolic Disease, Lipid Metabolism, Small Molecule Biochemistry

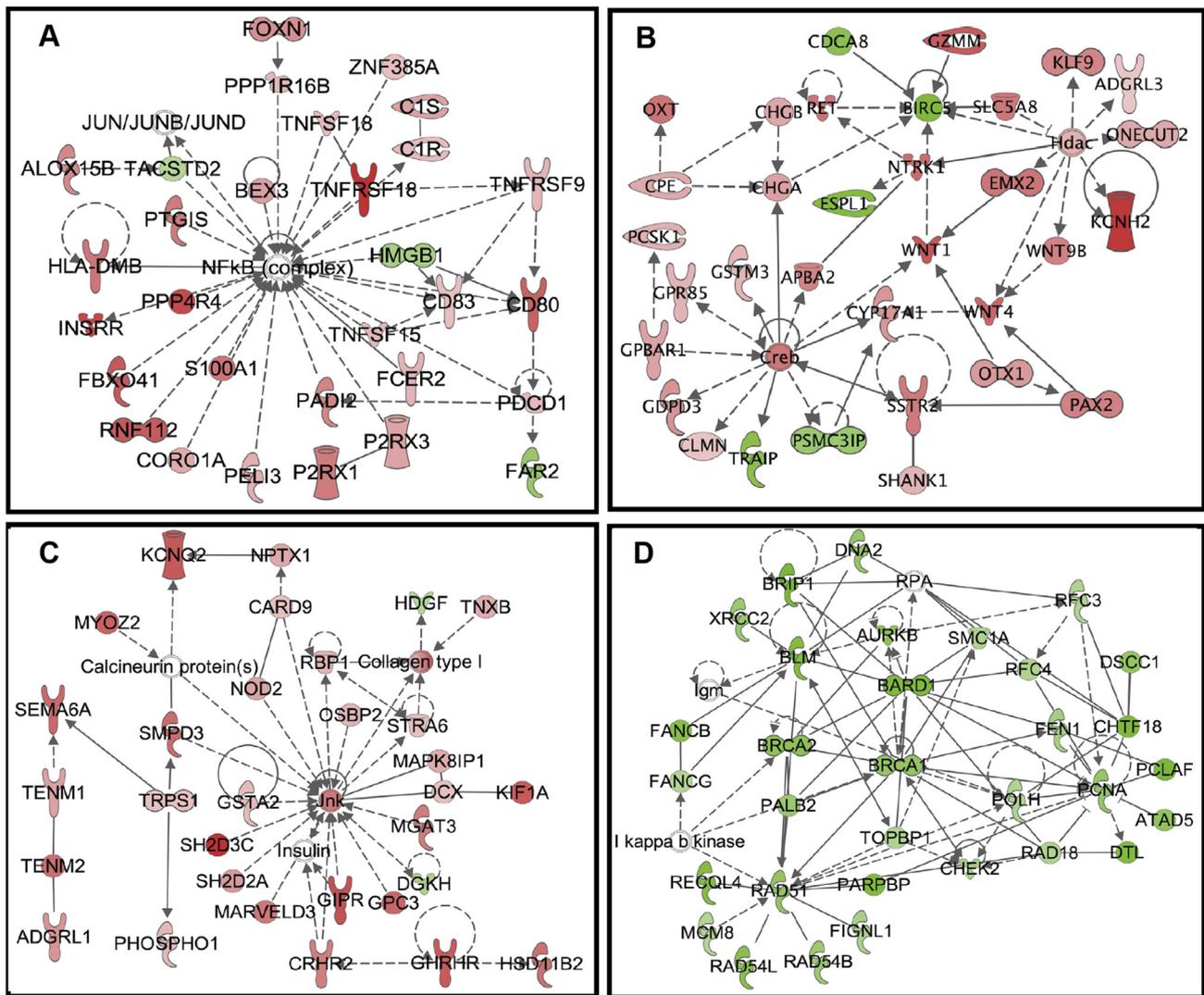


Figure 4. Four key networks involved in the development of cattle epithelial tissues. A. Cell-To-Cell Signaling and Interaction, Cellular Growth and Proliferation, Hematological System Development and Function. B. Embryonic Development, Organ Development, Organismal Development. C. Cell-To-Cell Signaling and Interaction, Cellular Growth and Proliferation, Nervous System Development and Function. D. Cell cycle, DNA replication, Recombination and Repair, and Cell Morphology. Red dots represent up-regulated genes and green dots represent down-regulated genes.

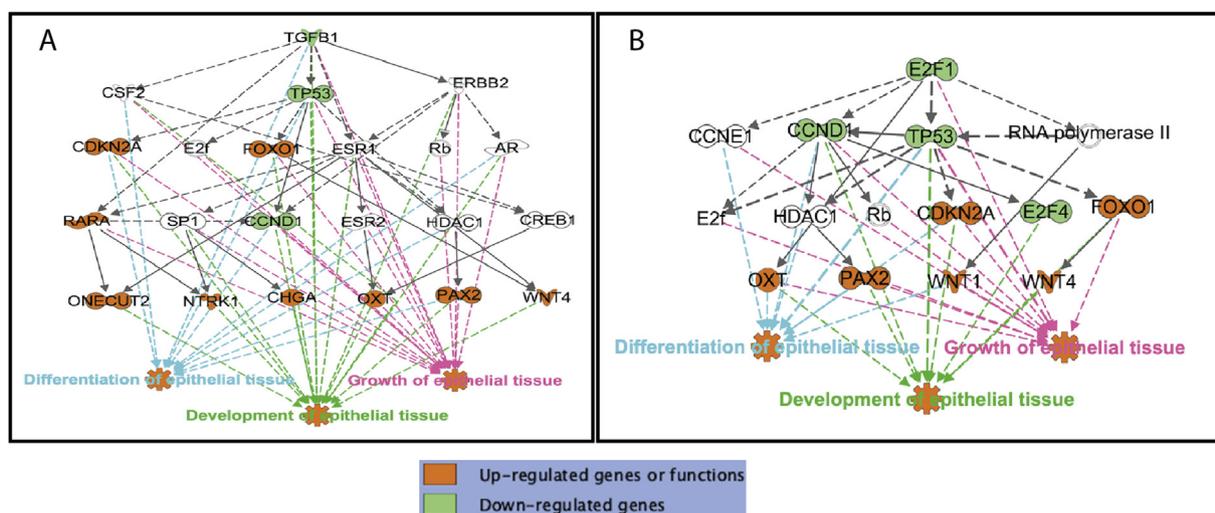


Figure 5. Upstream regulator effects of *TGFβ1* and *E2F1*. A. *TGFβ1*, and B. *E2F1*, two upstream negative regulators and their downstream genes, involving in regulating differentiation, development and growth of epithelial tissue.

Table 5. Regulator effects: genes significantly changed by two key upstream regulators *E2F1* and *TGFβ1*.

Gene name	Description	Log ₂ fold change	P value	FDR P value
CCND1	cyclin D1	-1.15	0	0
CDKN2A	cyclin-dependent kinase inhibitor 2A	0.74	0	0
CHGA	chromogranin A	3.41	0	0
E2F1	E2F transcription factor 1	-2.93	5.25E-11	2.94E-10
E2F4	E2F transcription factor 4	-0.75	6.72E-12	4.01E-11
FOXO1	forkhead box protein O1	1.02	3.13E-05	1.07E-04
NTRK1	neurotrophic receptor tyrosine kinase1	5.59	0	0
ONECUT2	one cut homeobox 2	3.12	7.17E-03	0.02
OXT	oxytocin/neurophysin I prepropeptide	5.31	0.000963	0.00272
PAX2	paired box 2	4.73	0.01	0.03
TGFβ1	transforming growth factor beta 1	-2.63	0	0
TP53	tumor protein p53	-1.62	0	0
RARA	retinoic acid receptor alpha	0.96	0	0
WNT1	Wnt family member 1	6.59	1.47E-06	5.77E-06
WNT4	Wnt family member 4	5.86	9.30E-05	3.30E-03

of the gene expression data of total of 18 single cells along with the clustering of the differentially expressed transcripts showed high divergence and variation in gene expression (Figure 6 A and B) among the REPC cells. Such high divergence and variation in gene expression may present an indication of considerable inter-individual diversity in the gene and genetic networks that were activated during the individual cell's physiological statuses or phenotypes. A total of 402 genes were in common as DEG from both bulk and single-cell subgroups (Figure 6 C).

We observed 1,204 DEGs in a single cell induced by butyrate treatment (Figure 7A). Given the high divergence and variation observed among the single-cell RNA expression, we subsequently estimated the Pearson correlation between the bulk and single-cell RNA-Sequencing. With Pearson's $R = 0.79$ (Figure 7B), indicating that bulk cell RNA sequencing data and single-cell RNA sequencing data have very similar expression profiles and thus both bulk cells preparations and single-cell samples represent their derivation from the epithelial cell population they originated from and similarly directed response to butyrate treatment. IPA was also used to clarify further the affected biological functions in response to butyrate treatment from the single cells analyzed as well as comparisons between bulk and single-cell RNA sequencing results. There are only 402 genes that overlapping the response to butyrate between the bulk and single-cell RNA-Seq analyses. The results of the

analyses of canonical pathways, upstream regulators and molecular functions highlight the commonality between the two sample preparation techniques (bulk cells and single-cell sequencing) (Figure 7C).

4. Discussion

In the current study, a culture of primary cells isolated from a single newborn calf (2 weeks of age) rumen epithelium was established to mimic the preweaning physiological status of cells *in vivo* specifically to elucidate the effects of butyrate and rumen development likely to occur at this stage of maturity. A stable rumen epithelial primary cell (REPC) culture is a critical and high-value tool to closely mimic the physiological state of immature rumen epithelial cells *in vivo* to elucidate the effects of butyrate on rumen development. Using bulk-cell and single-cell RNA sequencing, the transcriptomic profile of REPC was established, and the effects of butyrate on the transcriptome of the rumen epithelial cells was tested. Highly expressed epithelial cells biomarker E-Cadherin/Cadherin-1 (ECAD) confirms that cells in our established REPC culture, at least the majority of the cells, are of the epithelial origin. E-Cadherin/Cadherin-1 (ECAD) is a transmembrane cell adhesion protein. It forms homophilic clusters between epithelial cells and is a critical component of adherens junctions. These junctions are critical for the maintenance of the

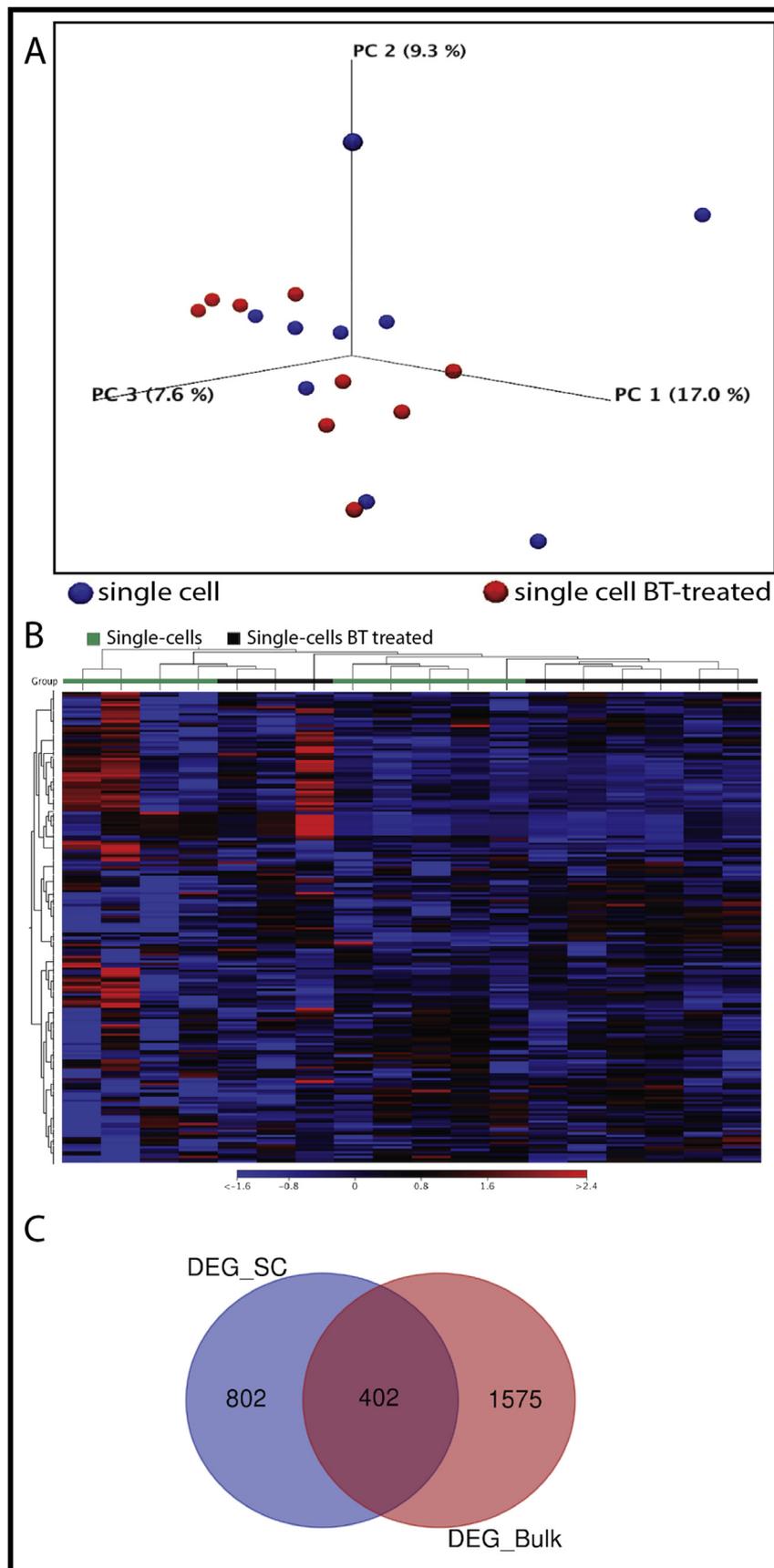


Figure 6. Single-cell RNA sequencing highlights heterogeneity in primary rumen epithelial cells. A. Principal component analysis of 18 single-cell transcriptomes. B. A heatmap showing the divergence and variability of the transcriptome of single-REPC, The color code represent normalized expression value (log2 fold change). C. a Venn diagram of DEG induced by butyrate in bulk cells vs. single-cell transcriptomic analyses.

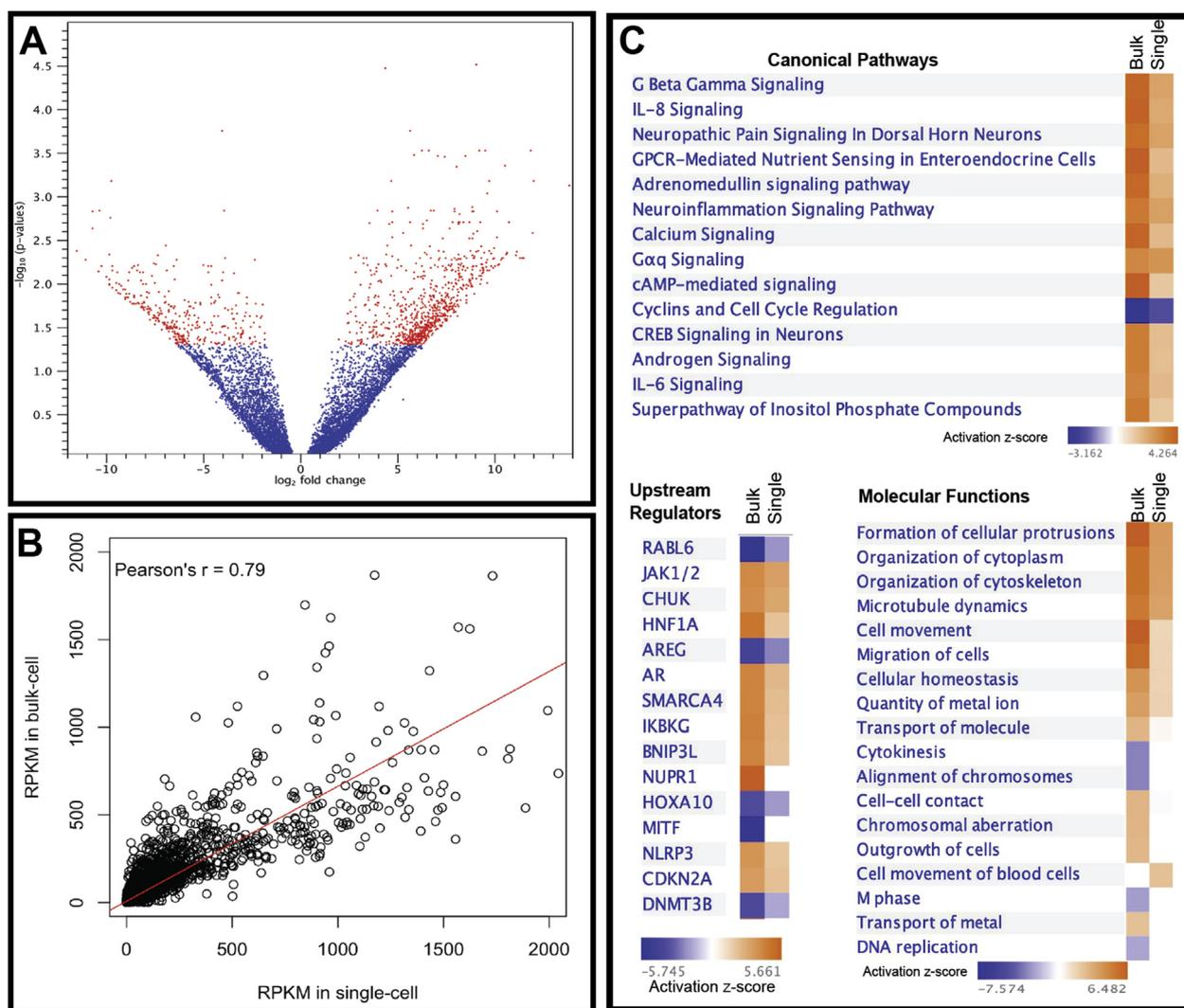


Figure 7. A. volcano plot showing the DEGs induced by butyrate in the single-cell transcriptomic analysis. B. Pearson's correlation between the expression values (RPKM, Reads Per Kilobase of transcript) of normal bulking samples and normal single-cell samples. C. Comparison of DEG in butyrate-induced transcriptomes in bulk and single-cell RNA-Seq showed similar trends in canonical pathways, upstream regulators and molecular functions.

epithelial barrier and epithelial tissue homeostasis. It is unavoidable that bulk isolation and culture of primary cells may contain different cell types. Therefore, we are in progress performing single-cell sequencing in large quantities (~10,000 cells) to examine the cell population in REPC. Currently, we are diligently analyzing the large scale single-cell sequencing data and will report our findings in the near future.

Analysis of the transcriptome indicated that butyrate could induce differentiation, development, and growth of cattle epithelial tissue through modification of expression of a number of genes. The role of butyrate in the development of ruminal epithelium and effect on isolated epithelial cells is receiving a great deal of research attention recently, and it has been established that butyrate is important as a substrate of cell metabolism and as a regulator of gene activities [10, 27, 28]. The ruminal concentration of butyrate in neonatal calves is low, ~0.002 mM/l [29, 30], while the calves were fed with only milk or commercial milk replacer before weaning. Calves start eating solid feeds during weaning transition and rumen is the site of fermentation. A massive community of microorganisms, bacteria and protozoa, ferments the plant material to short-chain fatty acids, methane, and carbon dioxide. The artificial elevation of the ruminal concentration of butyrate within the normal physiological range has demonstrated accelerated maturation processes in the ruminal epithelial cells in calves [31, 32].

Butyrate maintains the structural integrity of ruminal epithelial cells and intestinal epithelial cells and maintains intestinal homeostasis. Interestingly, butyrate is a prominent epigenetic regulator due to its strong inhibitory effects on histone deacetylase (HDAC) [26, 33], which results in specific changes in gene expression. Using butyrate treatment of REPC, we identified many signaling pathways, such as *WNT*, *MAPK*, *Jnk*, *NFκB*, *ERK*, and *PI3K* pathways, as putative factors affecting epithelial tissue development. Moreover, these identified pathways are known to play regulatory roles in biological functions relating to cell cycle, cellular assembly and organization, DNA replication, recombination and repair, cell morphology, and lipid metabolism. The *WNT* pathways, in particular, are known to direct the specific activation of sets of genes that regulate several cellular responses, including cell growth, differentiation, movement, migration, polarity, cell survival, and immune response [34]. Thus our findings are consistent with the accepted theory that the *WNT* signaling pathway is associated with cell differentiation, polarization, and migration during development [35].

The results from our experiment with REPC reflect the impact of butyrate treatment on rumen epithelial transcriptome and essential gene network and pathways associated with the development of cattle epithelial tissue. Our findings not only consistent with our previous experiments with bovine tissue but also with other experiment systems. For

instance, existing literature reports that osteoblast differentiation is controlled via a complex mechanism involving extracellular signaling molecules such as *WNT*, *TGF β* , bone morphogenetic proteins (BMP), fibroblast growth factor, parathyroid hormone, Indian hedgehog, Sonic hedgehog, Notch, and insulin-like growth factors. Moreover, *WNT* and *TGF β* play critical roles at multiple steps during osteoblast differentiation [36, 37]. Likewise, *MAPKs*, which are activated by various environmental stimuli, have been shown to regulate the transcriptional activity of many genes involved in maintaining cellular homeostasis. There are three major groups of *MAPKs* in mammalian cells that are essential regulatory proteins that transduce various extracellular signals into intracellular events: the *ERK*, Jnk/stress-activated protein kinase, and p38 subfamilies [38]. The dynamics of *ERK* signaling underlies its versatile functions in cell differentiation, cell proliferation, and cell motility [39]. *NF κ B* is a critical transcription factor present in the nucleus, which not only participates in the process of controlling immune responses such as inflammation and tissue damage repair but also in the embryo development [40].

In the present analysis, a total of 1977 genes were found differentially expressed between Bulk-BT treated and Bulk-control cells. Among them, up to 7 highly upregulated genes (*CDKN2A*, *NTRK1*, *ONECUT2*, *OXT*, *PAX2*, *WNT1*, *WNT4*), which are known to be regulated by two upstream regulators *E2F1* and *TGF β 1*. The *E2F* family of transcription factors has previously been tied to the upregulation of transcript expression of products, which are crucial for cell cycle progression to commence. *E2F1* promotes gene expression by increasing RNA pol II phosphorylation, which promotes the recruitment of the methyl cap synthetic enzymes. The increase of RNA pol II phosphorylation is necessary for *E2F1*-dependent methyl cap formation, and the methyl cap is required for mRNA maturation, expression and stability [41].

E2F1 is a critical transcription factor modulating the expression of specific chromatin components in oligodendrocyte progenitor cells during the transition from proliferation to differentiation. Specifically, *E2F1* targeted cell cycle genes and chromatin components, including those modulating DNA methylation [42]. In our study, the *E2F1* gene expression was down-regulated by butyrate treatment. Thus, it is a putative upstream regulator for both proliferation and development in the ruminal epithelium. Likewise, *TGF β 1* is also a negative regulator of proliferation and is a known inducer of apoptosis beyond its fibrogenic effects, leading to trans differentiation of hepatic stellate cells into myofibroblasts [43]. In our preparation, *TGF β 1* gene expression was also found significantly down-regulated in primary rumen epithelial cells treated with butyrate and thus, may participate in the development and growth of epithelial tissue. This observation is consistent with previous work that showed *TGF β 1* mediates calf rumen epithelial tissue development [31 1473]. Therefore, this observation and the results discussed above signify that using REPC, we are able to reproduce results from rumen in vivo experiments and REPC is a suitable model for early ruminal development research. The polycomb repressor complex 2 molecule *EZH2* is now known to play a role in essential cellular processes, namely, cell fate decisions, cell cycle regulation, senescence, cell differentiation, and cancer development/progression. It was demonstrated that the depletion and inhibition of *EZH2* could successfully induce *NTRK1* transcripts and functional proteins, which indicated that *EZH2* plays crucial roles in preventing the differentiation of neuroblastoma cells and that *EZH2*-related *NTRK1* transcriptional regulation may be the critical pathway for neuroblastoma cell differentiation [44].

An oncogenic mechanism known to be under epigenetic regulation and also affected by *HDAC* inhibitor is the epithelial-mesenchymal transition (EMT). The EMT process is acquired by polarized epithelial cells from mesenchymal phenotypes upon a loss of epithelial cell-to-cell junction. Cells undergoing EMT display decreased expression of epithelial markers *CDH1*, increased expression of *CDH2*, as well as *GRHL2* and *ELF3/5* were down-regulated during EMT [45, 46]. Butyrate as an *HDAC* inhibitor changes gene transcription in ovarian cancer cells and affects pathways critical for EMT [47]. In REPC, as it is a normal cell population,

butyrate did not change the expression pattern of epithelial cell markers of *CDH1* and *CDH2*. However, two transcription factors *GRHL2* and *ELF3*, both belong to central families of proteins in differentiation, epithelialization and the negative EMT regulators [48, 49, 50], were found to markedly increase after treating with butyrate. This implies that butyrate may inhibit EMT activity in REPC, thereby promote and/or maintain the epithelial cell's differentiation and functionality. This observation may also have a role in the action of butyrate when used for chemoprevention and therapies to treat cancers.

Single-cell transcriptome complexity and single-cell transcriptome variation, especially within-cell-type variability has been reported in different cell types [20, 21, 51]. With the small size of sampling of a total of 18 cells, single-cell transcriptome sequencing of REPC demonstrated a high transcriptome variability, even though they all consistently expressed epithelial gene markers. While we cannot for certain eliminate technical replicate variability, it is conceivable that individual cell phenotypes such as the size, ultrastructure, stage of cell cycle may directly impact cell-to-cell transcriptome variability. Single-cell phenotypic variations may also be significant biologically for understanding cell-cycle dynamics caused by this apparent transcriptome variability [51]. Therefore, large-sampling of single-cell transcriptome sequencing may be required to characterize intracellular heterogeneity, identification of cell type from mixed cultures and cell cycle state, evaluation of dynamic cellular transitions with complex cell compositions such as stratified squamous epithelia, and describe impacts of cell-to-cell interactions among hundreds to tens of thousands of cells.

Declarations

Author contribution statement

Shudai Lin: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Lingzhao Fang, Xiaolong Kang, Shuli Liu, Mei Liu: Analyzed and interpreted the data.

Erin E. Connor: Conceived and designed the experiments; Performed the experiments.

Ransom L. Baldwin: Performed the experiments.

George Liu: Conceived and designed the experiments.

Cong-Jun Li: Conceived and designed the experiments; Wrote the paper.

Funding statement

G.E. Liu was supported by appropriated project 8042-31000-001-00-D, "Enhancing Genetic Merit of Ruminants Through Improved Genome Assembly, Annotation, and Selection" of the Agricultural Research Service of the United States Department of Agriculture. E.E. Connor, R.L. Baldwin, and C.-J. Li were supported by appropriated project 8042-31310-078-00-D, "Improving Feed Efficiency and Environmental Sustainability of Dairy Cattle through Genomics and Novel Technologies".

Competing interest statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at <https://doi.org/10.1016/j.heliyon.2020.e04112>.

Acknowledgements

We thank Reuben Anderson, Mary Bowman, Donald Carbaugh, Christina Clover, Cecelia Niland, and Sara McQueeney for technical assistance and sample collection. Mention of trade names or commercial

products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. The USDA is an equal opportunity provider and employer.

References

- [1] R.L.t. Baldwin, E.E. Connor, Rumen function and development, *Vet. Clin. North Am. Food Anim. Pract.* 33 (3) (2017) 427–439.
- [2] B. Niwinska, et al., Review: exogenous butyrate: implications for the functional development of ruminal epithelium and calf performance, *Animal* 11 (9) (2017) 1522–1530.
- [3] P. Gorka, et al., Invited review: use of butyrate to promote gastrointestinal tract development in calves, *J. Dairy Sci.* 101 (6) (2018) 4785–4800.
- [4] M. Malhi, et al., Increased papillae growth and enhanced short-chain fatty acid absorption in the rumen of goats are associated with transient increases in cyclin D1 expression after ruminal butyrate infusion, *J. Dairy Sci.* 96 (12) (2013) 7603–7616.
- [5] E.E. Connor, et al., Gene expression in bovine rumen epithelium during weaning identifies molecular regulators of rumen development and growth, *Funct. Integr. Genom.* 13 (1) (2013) 133–142.
- [6] Q. Yan, et al., Proteomic profiling of the acid tolerance response (ATR) during the enhanced biomethanation process from Taihu Blue Algae with butyrate stress on anaerobic sludge, *J. Hazard Mater.* 235–236 (2012) 286–290.
- [7] L. Lin, et al., Ruminal microbiome-host crosstalk stimulates the development of the ruminal epithelium in a lamb model, *Microbiome* 7 (1) (2019) 83.
- [8] L. Liu, et al., Infusion of sodium butyrate promotes rumen papillae growth and enhances expression of genes related to rumen epithelial VFA uptake and metabolism in neonatal twin lambs, *J. Anim. Sci.* 97 (2) (2019) 909–921.
- [9] C.J. Li, T.H. Elsasser, Butyrate-induced apoptosis and cell cycle arrest in bovine kidney epithelial cells: involvement of caspase and proteasome pathways, *J. Anim. Sci.* 83 (1) (2005) 89–97.
- [10] C.J. Li, et al., Pathway analysis identifies perturbation of genetic networks induced by butyrate in a bovine kidney epithelial cell line, *Funct. Integr. Genom.* 7 (3) (2007) 193–205.
- [11] J.H. Shin, et al., Genome-wide ChIP-seq mapping and analysis reveal butyrate-induced acetylation of H3K9 and H3K27 correlated with transcription activity in bovine cells, *Funct. Integr. Genom.* 12 (1) (2012) 119–130.
- [12] E.N. Bergman, Energy contributions of volatile fatty acids from the gastrointestinal tract in various species, *Physiol. Rev.* 70 (2) (1990) 567–590.
- [13] S.M. McNabney, T.M. Henagan, Short chain fatty acids in the colon and peripheral tissues: a focus on butyrate, colon cancer, obesity and insulin resistance, *Nutrients* 9 (12) (2017).
- [14] S.A. Hudu, et al., Cell culture, technology: enhancing the culture of diagnosing human diseases, *J. Clin. Diagn. Res. : JCDR* 10 (3) (2016) DE01–DE5.
- [15] J. Kim, J. Eberwine, RNA: state memory and mediator of cellular phenotype, *Trends Cell Biol.* 20 (6) (2010) 311–318.
- [16] L. Fang, et al., Functional annotation of the cattle genome through systematic discovery and characterization of chromatin states and butyrate-induced variations, *BMC Biol.* 17 (1) (2019) 68.
- [17] R.L. Baldwin, The proliferative actions of insulin, insulin-like growth factor-I, epidermal growth factor, butyrate and propionate on ruminal epithelial cells in vitro, *Small Rumin. Res.* 32 (1999) 261–268.
- [18] S. Wu, et al., Transcriptome characterization by RNA-seq unravels the mechanisms of butyrate-induced epigenomic regulation in bovine cells, *PLoS One* 7 (5) (2012), e36940.
- [19] X. Fan, et al., Spatial transcriptomic survey of human embryonic cerebral cortex by single-cell RNA-seq analysis, *Cell Res.* 28 (7) (2018) 730–745.
- [20] L. Li, et al., Single-cell RNA-seq analysis maps development of human germline cells and gonadal niche interactions, *Cell Stem Cell* 20 (6) (2017) 858–873 e4.
- [21] F. Zhou, et al., Tracing haematopoietic stem cell formation at single-cell resolution, *Nature* 533 (7604) (2016) 487–492.
- [22] A. Dobin, et al., STAR: ultrafast universal RNA-seq aligner, *Bioinformatics* 29 (1) (2013) 15–21.
- [23] C. Trapnell, et al., Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks, *Nat. Protoc.* 7 (3) (2012) 562–578.
- [24] C. Trapnell, et al., Differential analysis of gene regulation at transcript resolution with RNA-seq, *Nat. Biotechnol.* 31 (1) (2013) 46–53.
- [25] A. Conesa, et al., A survey of best practices for RNA-seq data analysis, *Genome Biol.* 17 (2016) 13.
- [26] R.L.t. Baldwin, et al., Transcriptomic impacts of rumen epithelium induced by butyrate infusion in dairy cattle in dry period, *Gene Regul. Syst. Biol.* 12 (2018), 1177625018774798.
- [27] C.J. Li, R.W. Li, Butyrate induced cell cycle arrest in bovine cells through targeting gene expression relevant to DNA replication apparatus, *Gene Regul. Syst. Biol.* 2 (2008) 113–123.
- [28] C.J. Li, R.W. Li, Bioinformatic dissecting of TP53 regulation pathway underlying butyrate-induced histone modification in epigenetic regulation, *Genet. Epigenet.* 6 (2014) 1–7.
- [29] K.E. Lesmeister, A.J. Heinrichs, Effects of corn processing on growth characteristics, rumen development, and rumen parameters in neonatal dairy calves, *J. Dairy Sci.* 87 (10) (2004) 3439–3450.
- [30] K.E. Lesmeister, A.J. Heinrichs, M.T. Gabler, Effects of supplemental yeast (*Saccharomyces cerevisiae*) culture on rumen development, growth characteristics, and blood parameters in neonatal dairy calves, *J. Dairy Sci.* 87 (6) (2004) 1832–1839.
- [31] E.E. Connor, et al., Transcriptional regulators transforming growth factor-beta1 and estrogen-related receptor-alpha identified as putative mediators of calf rumen epithelial tissue development and function during weaning, *J. Dairy Sci.* 97 (7) (2014) 4193–4207.
- [32] A.H. Laarman, et al., Effects of feeding a calf starter on molecular adaptations in the ruminal epithelium and liver of Holstein dairy calves, *J. Dairy Sci.* 95 (5) (2012) 2585–2594.
- [33] D.R. Donohoe, et al., The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation, *Mol. Cell.* 48 (4) (2012) 612–626.
- [34] B. Garcia-Castro, et al., Restoration of WNT4 inhibits cell growth in leukemia-derived cell lines, *BMC Cancer* 13 (2013) 557.
- [35] B. Taciak, et al., Wnt signaling pathway in development and cancer, *J. Physiol. Pharmacol.* 69 (2) (2018).
- [36] R. Baron, M. Kneissel, WNT signaling in bone homeostasis and disease: from human mutations to treatments, *Nat. Med.* 19 (2) (2013) 179–192.
- [37] M. Wu, G. Chen, Y.P. Li, TGF-beta and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease, *Bone Res.* 4 (2016) 16009.
- [38] L. Chang, M. Karin, Mammalian MAP kinase signalling cascades, *Nature* 410 (6824) (2001) 37–40.
- [39] Y. Ogura, et al., A switch-like activation relay of EGFR-ERK signaling regulates a wave of cellular contractility for epithelial invagination, *Dev. Cell* 46 (2) (2018) 162–172 e5.
- [40] R.G. Jayasooriya, et al., Verrucarin A enhances TRAIL-induced apoptosis via NF-kappaB-mediated Fas overexpression, *Food Chem. Toxicol.* 55 (2013) 1–7.
- [41] M. Aregger, V.H. Cowling, E2F1-dependent methyl cap formation requires RNA pol II phosphorylation, *Cell Cycle* 11 (11) (2012) 2146–2148.
- [42] L. Magri, et al., E2F1 coregulates cell cycle genes and chromatin components during the transition of oligodendrocyte progenitors from proliferation to differentiation, *J. Neurosci.* 34 (4) (2014) 1481–1493.
- [43] A.M. Gressner, et al., Roles of TGF-beta in hepatic fibrosis, *Front. Biosci.* 7 (2002) d793–807.
- [44] Z. Li, et al., EZH2 regulates neuroblastoma cell differentiation via NTRK1 promoter epigenetic modifications, *Oncogene* 37 (20) (2018) 2714–2727.
- [45] B. De Craene, G. Bercx, Regulatory networks defining EMT during cancer initiation and progression, *Nat. Rev. Cancer* 13 (2) (2013) 97–110.
- [46] R. Chakrabarti, et al., E2F1 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2, *Nat. Cell Biol.* 14 (11) (2012) 1212–1222.
- [47] A. Mrkvicova, et al., The effect of sodium butyrate and cisplatin on expression of EMT markers, *PLoS One* 14 (1) (2019), e0210889.
- [48] F.J. Boivin, K.M. Schmidt-Ott, Transcriptional mechanisms coordinating tight junction assembly during epithelial differentiation, *Ann. N. Y. Acad. Sci.* 1397 (1) (2017) 80–99.
- [49] T.L. Yeung, et al., ELF3 is a negative regulator of epithelial-mesenchymal transition in ovarian cancer cells, *Oncotarget* 8 (10) (2017) 16951–16963.
- [50] H. Nishino, et al., Grainyhead-like 2 (GRHL2) regulates epithelial plasticity in pancreatic cancer progression, *Cancer Med.* 6 (11) (2017) 2686–2696.
- [51] H. Dueck, et al., Deep sequencing reveals cell-type-specific patterns of single-cell transcriptome variation, *Genome Biol.* 16 (2015) 122.