Gene Regulation and Systems Biology

ORIGINAL RESEARCH

Butyrate Induced IGF2 Activation Correlated with Distinct Chromatin Signatures Due to Histone Modification

Joo Heon Shin[1], Robert W. Li[3], Yuan Gao[1], Derek M. Bickhart[3], George E. Liu[3], Weizhong Li[2], Sitao Wu[2] and Cong-jun Li[3]

[1]Lieber Institute for Brain Development, Johns Hopkins University, Baltimore, Maryland, United States of America. [2]Center for Research in Biological Systems, University of California, San Diego, California, United States of America. [3]United States Department of Agriculture, Agricultural Research Service, Bovine Functional Genomics Laboratory, Beltsville, Maryland, United States of America. Corresponding author email: congjun.li@ars.usda.gov

Abstract: Histone modification has emerged as a very important mechanism regulating the transcriptional status of the genome. Insulin-like growth factor 2 (IGF2) is a peptide hormone controlling various cellular processes, including proliferation and apoptosis. H19 gene is closely linked to IGF2 gene, and IGF2 and H19 are reciprocally regulated imprinted genes. The epigenetic signature of H19 promoter (hypermethylation) on the paternal allele plays a vital role in allowing the expression of the paternal allele of IGF2.[46] Our previous studies demonstrate that butyrate regulates the expression of IGF2 as well as genes encoding IGF Binding proteins. To obtain further understanding of histone modification and its regulatory potentials in controlling IGF2/H19 gene expression, we investigated the histone modification status of some key histones associated with the expression of IGF2/H19 genes in bovine cells using RNA-seq in combination with Chip-seq technology. A high-resolution map of the major chromatin modification at the IGF2/H19 locus induced by butyrate was constructed to illustrate the fundamental association of the chromatin modification landscape that may play a role in the activation of the IGF2 gene. High-definition epigenomic landscape mapping revealed that IGF2 and H19 have distinct chromatin modification patterns at their coding and promoter regions, such as TSSs and TTSs. Moreover, the correlation between the differentially methylated regions (DMRs) of IGF2/H19 locus and histone modification (acetylation and methylation) indicated that epigenetic signatures/markers of DNA methylation, histone methylation and histone acetylation were differentially distributed on the expressed IGF2 and silenced H19 genes. Our evidence also suggests that butyrate-induced regional changes of histone acetylation statusin the upstream regulation domain of H19 may be related to the reduced expression of H19 and strong activation of IGF2. Our results provided insights into the mechanism of butyrate-induced loss of imprinting (LOI) of IGF2 and regulation of gene expression by histone modification.

Keywords: bovine, butyrate, ChIP-seq, chromatin, histone modification, IGF2.

Gene Regulation and Systems Biology 2013:7 57–70

doi: 10.4137/GRSB.S11243

This article is available from http://www.la-press.com.

Introduction

The insulin-like growth factor (IGF) system plays an essential role in cell growth, proliferation, differentiation, transformation and apoptosis.[1,2] Insulin-like growth factor 2 (IGF2) is an imprinted gene with the parental allele expressed and the maternal allele silenced. IGF2 activity is further controlled through differential expression of receptors and IGF-binding proteins (IGFBPs) that determine protein availability. This complex and multifaceted regulation emphasizes the importance of accurate IGF2 expression and activity.[3] IGFBPs act as carrier/transport proteins in biologic fluids to control IGF efflux. In addition to their roles in prolonging IGF half-lives, IGFBPs themselves possess biological functions independent of their ability to modulate IGF activities.[4] For example, IGFBP3 has been shown to act as a ligand for nuclear receptor RXRα, which is involved in apoptosis.[5]

Butyrate serves as an inhibitor of histone deacetylases (HDACs), which are critical epigenetic regulators.[6–8] Butyrate can reactivate epigenetically-silenced genes by increasing global histone acetylation;[9] consequently butyrate induced histone acetylation is an excellent model to study epigenetic effects of histone modifications. Additionally, as a natural microbial fermentation product in the gastrointestinal tract, butyrate is an important nutrient in ruminants. Butyrate is produced during microbial fermentation of dietary fiber in the gastrointestinal tract, and is directly absorbed at the site of production and oxidized for cell energy production and use.[10] In humans, the colonic microbiota converts dietary fiber to produce prodigious amounts of butyrate that benefits the host through metabolic, trophic, and chemo-preventative effects. The human IGF2 locus includes a genomic interval of approximately 150 kb on human chromosome 11, containing two imprinted genes, IGF2 and H19, shared enhancers, and *cis*-acting regulatory elements such as the imprinting control region (ICR). DNA methylation and histone post- translational modifications, such as acetylation, methylation and phosphorylation, play a key role in IGF2/H19 gene expression regulation.[11] Differential histone acetylation may be one of the potential mechanisms by which the imprinting of IGF2/H19 is regulated.

We have previously examined potential mechanisms of butyrate-induced gene regulation in bovine epithelial cells (Madin-Darby bovine kidney epithelial cells (MDBK)). As a result, there is a significant amount of information on how butyrate regulates the transcriptome and various biological pathways. In this report, we investigated how butyrate regulated expression of IGFs, IGF receptors, and IGFBP; specifically, we assessed the major histone modification status associated with the expression of IGF2/H19 genes using RNAseq and Chip- seq technologies. To illustrate the fundamental association of the chromatin modification landscape that may play a role in the activation of the IGF2 gene, a high-resolution map of the major chromatin modification at the IGF2/H19 locus was constructed by mining our RNA-seq and ChIP-seq data. ChIP-seq data reveals that IGF2 may be regulated by dispersed epigenomic domains, which show a strong correlation between the structure of the IGF2 gene and the clustered imprinting gene H19.

Material and Methods

Cell culture and butyrate treatments

The Madin-Darby bovine kidney epithelial cells (MDBK, American Type Culture Collection, Manassas, VA, Catalog No. CCL-22) were cultured in Eagle's minimal essential medium and supplemented with 5% fetal bovine serum (Invitrogen, Carlsbad, CA) in 25 cm2 flasks, as described in our previous report.[6] At approximately 50% confluence, during the exponential phase, the cells were treated with 10 mM of sodium butyrate for 24 hours (Calbiochem, San Diego, CA). A butyrate concentration of 10 mM was selected as it represents a physiologically relevant dose and has previously been successfully used to evoke desired changes in cell cycle dynamics.[6] Four replicate flasks of cells for both treatment and control groups—a total of 8 samples—were used for the RNA-seq experiments.

ChIP

A ChIP-IT™ Express Enzymatic Kit (Active Motif North America, Carlsbad, CA) was used for this ChIP experiment. Antibodies against H3 and acetyl-H3K18, ChIP Formulated, were obtained from Cell Signaling Technology; acetyl-H3K9, acetyl-H3K27 and H4 (ChIP Grade) antibodies were obtained from Abcam. H4 tetra-acetyl, acetyl-H4K12, acetyl-H4K5, acetyl K16, and H3K9 trimethylation were obtained

from Active Motif. All antibodies were characterized and validated for cross-reactivity before the ChIP experiments with Western blots and immunoprecipitation (IP). The preparation of sheared chromatin (cell fixation, enzymatic shearing) was performed per the instructions of the product manufacturer. Briefly, cells were fixed with 1% formaldehyde in flasks at room temperature for 10 minutes and subsequently washed with ice-cold PBS before a 0.1 M glycine solution was added to stop the fixation. The cells were scraped from the flask and homogenized with an ice-cold Dounce homogenizer to release the nuclei. The collected nuclei were again suspended in a digestion buffer and an enzymatic shearing cocktail was added and incubated for 20 minutes at 37 °C. The shearing reaction was stopped by adding 10 µL ice-cold 0.5 M EDTA prior to centrifugation at 15,000 rpm in a 4 °C microcentrifuge. The supernatant containing the sheared chromatin was carefully collected and prepared for ChIP sequencing. Next, 10 µg of primary antibody were added to 100 µL of chromatin and incubated on an end-to-end rotator overnight at 4 °C. Subsequently, 50 µL of protein G magnetic beads were added, and incubation continued for another 4 hours. The magnetic beads were washed with ChIP buffer 1 once and ChIP buffer 2 twice, before the bound chromatin was eluted and cross-links were reversed and treated with protease K. Traditionally, chromatin shearing has been performed by different pulses of sonication. Although sonication can be an effective method for shearing chromatin, it is very difficult to optimize due to emulsification and overheat, as well as the quality of sonicator. We used enzymatic shearing because it is solely dependent on time and temperature. The results of ChIP are very satisfying in our experiments.[12]

Illumina sequencing

DNA recovered from a conventional ChIP procedure was quantified using a QuantiFluor fluorometer (Promega, Madison, WI). The DNA integrity was verified using an Agilent Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA). The DNA was then processed—including end repair, adaptor ligation, and size selection—using an Illumina sample prep kit following the manufacturer's instructions (Illumina, San Diego, CA, USA). Final DNA libraries were validated and sequenced at 75 bp per sequence read, using an Illumina HiSeq-2000 sequencer at a depth of approximately 30 million sequences per sample (mean ± SD = 28 ± 4 million per sample) with single end. For RNA-seq, total RNA was extracted using Trizol (Invitrogen, Carlsbad, CA, USA) followed by DNase digestion and Qiagen RNeasy column purification (Qiagen, Valencia, CA, USA) as previously described.[7] (Li and Gasbarre, 2009).[13] The RNA integrity was verified using an Agilent Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA). High-quality RNA (RNA Integrity number or RIN $>$ 9.0) was processed using an Illumina TruSeq RNA sample prep kit following the manufacturer's instruction (Illumina, San Diego, CA, USA). After quality control procedures, individual RNA-seq libraries were pooled based on their respective 6-bp adaptors and sequenced at 50 bp/sequence read using an Illumina HiSeq 2000 sequencer as described previously.[14] Approximately 67.5 million reads/sample (mean ± sd = 67,527,111 ± 8,330,388.6) were generated. Four replicate flasks of cells for both treatment and control groups—a total of 8 samples—were used for the RNA-seq experiments.

ChIP-seq analysis

Sequencing tags were aligned to the bosTau4 Oct 2007 release of the reference genome using Bowtie, an ultrafast memory-efficient short read aligner.[15] We considered those tags that aligned uniquely with less than 2 mismatches. For enriched-region (peak) identification (peak calling), we used the MACS (Model-based Analysis of ChIP-seq) algorithm.[16,17] A simple yet effective technique for the analysis of eukaryotes, MACS was designed to identify transcription factor binding sites and histone modification-enriched regions in ChIP-seq data sets, with or without control samples. The application of ChIP-seq facilitated the identification of binding sites associated with transcription factors and regions of histone modification. A PeakAnalyzer[18] was used for Peak annotation, Nearest Downstream Gene analysis, and Transcription start site (TSS) analysis. The overall processing procedure is summarized in our previous publications. A local implementation of the UCSC Genome Browser[19] was used to visualize the details of gene structure and chromatin modification landscape. The UCSC genome Browser also provides displays with other aligned annotation tracks such as known genes, predicted genes, ESTs, and CpG islands.

RNA extraction and sequencing using RNA-seq

Total RNA was extracted using Trizol (Invitrogen, Carlsbad, CA, USA) followed by DNase digestion and Qiagen RNeasy column purification (Qiagen, Valencia, CA, USA), as previously described.[7] The RNA integrity was verified using an Agilent Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA). High-quality RNA (RNA Integrity number or RIN $>$ 9.0) was processed using an Illumina TruSeq RNA sample prep kit following the manufacturer's instruction (Illumina, San Diego, CA, USA). After quality control procedures, individual RNA-seq libraries were then pooled based on their respective 6-bp adaptors, sequenced at 50 bp/sequence, and read using an Illumina HiSeq 2000 sequencer, as described previously.[20] Approximately 67.5 million reads per sample (mean ± sd = 67,527, 111 ± 8,330,388.6) were generated.

The molecular processes, molecular functions and genetic networks following butyrate treatment were further evaluated by analyzing differentially expressed genes using Ingenuity Pathways Analysis (IPA, Ingenuity® Systems, and http://www.ingenuity.com). IPA is a software application that enables biologists to identify the biological mechanisms, pathways and functions most relevant to their experimental datasets or genes of interest.[21–25] The data sets supporting the results of this article have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE38973.

Results

Butyrate-induced differential expression of IGFs, IGFRs and IGFBPs

Gene expression of IGFs, IGF receptors (IGFRs), and IGFBPs induced by butyrate was evaluated by RNA-seq and verified by real-time PCR (Table 1). IGF1 was barely detectable in the bovine epithelial cells, while butyrate significantly up-regulated IGF2 gene expression (479 fold, FDR $<$ 0.01). Conversely, the IGF2 receptor (IGF2R) was also significantly induced by butyrate. Butyrate induced a strong expression of insulin-like growth factor 2 mRNA binding proteins (IGF2BPs), especially IGF2BP1 and IGF2BP3. These proteins function by binding to the 5′ UTR of the insulin-like growth factor 2 (IGF2) mRNA and regulating IGF2 translation. Of IGFBPs, only IGFBP3 was strongly up-regulated by butyrate, while other IGFBPs, such as IGFBP4, IGFBP6 and IGFBP7, were significantly down-regulated by butyrate. The expression level of IGFBP2 and IGFBP5 remained unchanged by butyrate. These results were consistent with the earlier report on the effects of butyrate on the expression of IGFs and IGFBPs.[26]

Table 1. Gene expression induced by butyrate showing deregulation among IGFs and IGFBPs.

Gene	RNA-seq	Real-time RT-PCR*	Significant**
IGF2	479.356	143.4	Yes
IGF1R	–1.873	–1.25	Yes
IGF2BP2	1.174	N/A	Yes
IGF2BP3	85.933	N/A	Yes
IGF2R	2.605	4.7	Yes
IGFALS	1.578	N/A	Yes
IGFBP2	–1.033	–1.67	No
IGFBP3	21.937	21.2	Yes
IGFBP4	–2.800	–3.33	Yes
IGFBP5	–1.040	2.3	No
IGFBP6	–8.272	–1.42	Yes
IGFBP7	–3.108	N/A	Yes
IGF1	BD	BD	No
IGFLR1	1.397	N/A	Yes

Notes: *Data were extracted from (Li & Li 2007); **FDR $<$ 0.01 and $P < 0.01$.

Function and canonical pathways analysis

IPA pathway analysis revealed the biological function and canonical pathways that were related to IGFs and IGFBPs gene expression induced by butyrate. As shown in Figure 1, the significant functions related to IGFs and IGFBPs included cell cycle, cellular growth and proliferation, and DNA replication, recombination and repair, which were closely correlated to the butyrate-induced biological effects, such as cell cycle arrest and apoptosis.[6] There are eight pathways significantly affected by butyrate (Fig. 2). The IPA analysis also identified a major regulatory network and two minor networks (Fig. 3). The major network was directly related to cellular growth and proliferation, cellular development, and skeletal and muscular system development and function; the two minor networks were involved in cellular movement, immune cell trafficking, gene expression, and lipid metabolism.

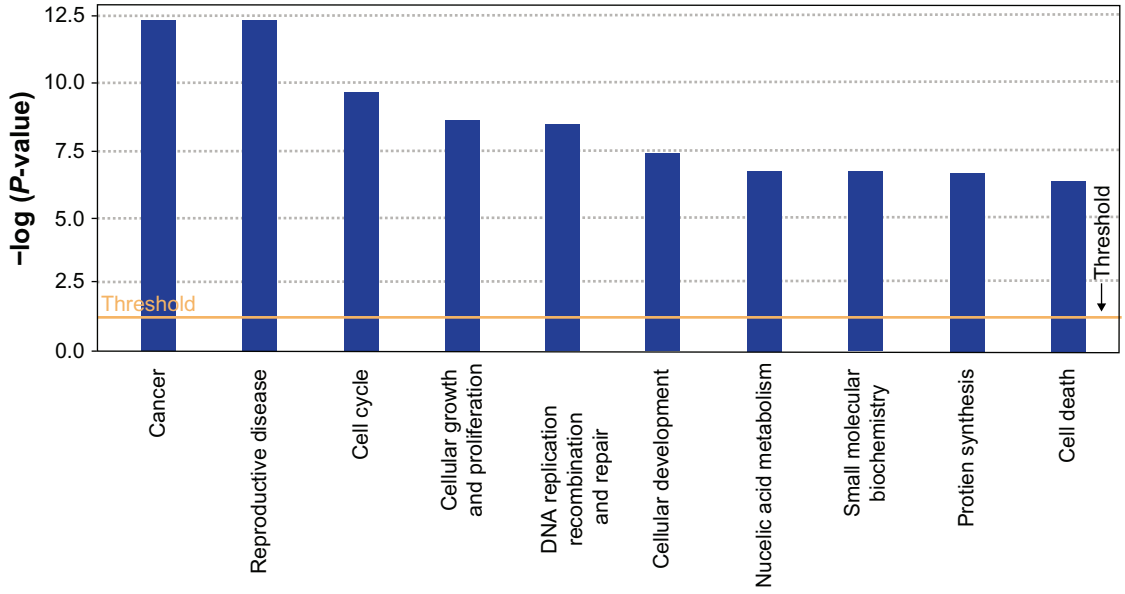

Figure 1. Global functional analysis.
Notes: Datasets were analyzed by the Ingenuity Pathways Analysis software (Ingenuity® Systems, http://www.ingenuity.com). The significance value associated with a function in Global Analysis is a measure for how likely it is that genes from the dataset file under investigation participate in that function. The significance is expressed as a *P*-value, which is calculated using the right-tailed Fisher's Exact Test.

Butyrate induces expression of IGF2 mRNA isoforms

Mixture-of-Isoforms algorithm (MISO), a probabilistic framework[27] that enables quantification of the expression level of alternatively spliced genes and identification of differentially regulated isoforms from RNA-Seq data, was used to analyze the alternative splicing events induced by butyrate. A significant number of alternative splicing events was discovered in the RNA-seq dataset. Ten selected alternative splicing events, including 2 exon-skipping event, 2 gene fusion events, and 6 transcript isoforms, were tested

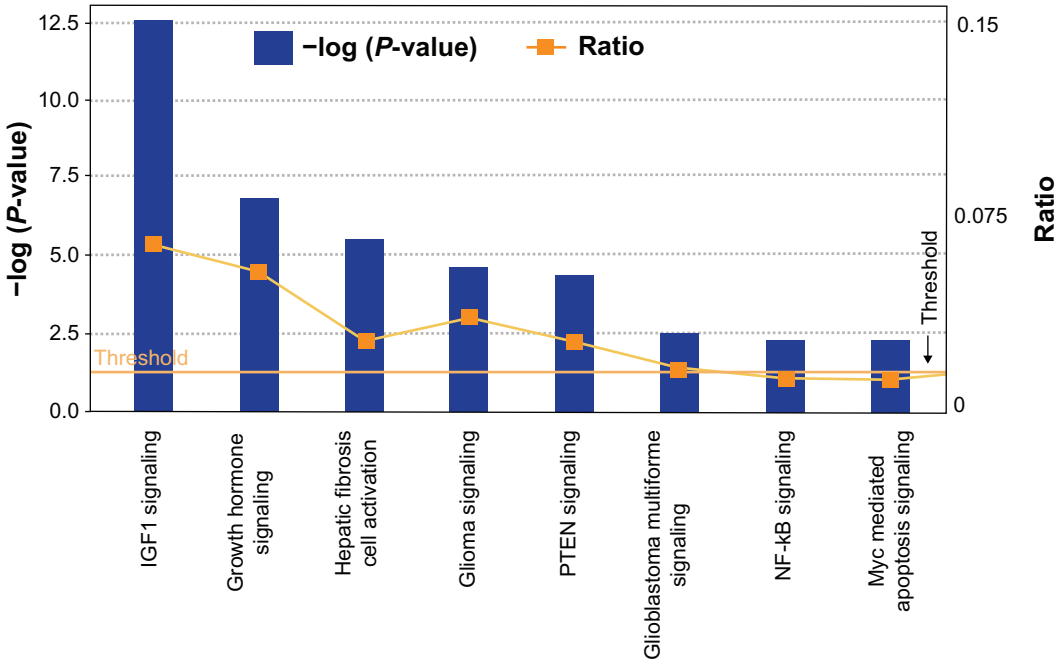

Figure 2. Global canonical pathway analysis.
Notes: Datasets were analyzed by the Ingenuity Pathways Analysis software (Ingenuity® Systems, http://www.ingenuity.com). The significance is expressed as a *P*-value, which is calculated using the right-tailed Fisher's Exact Test.

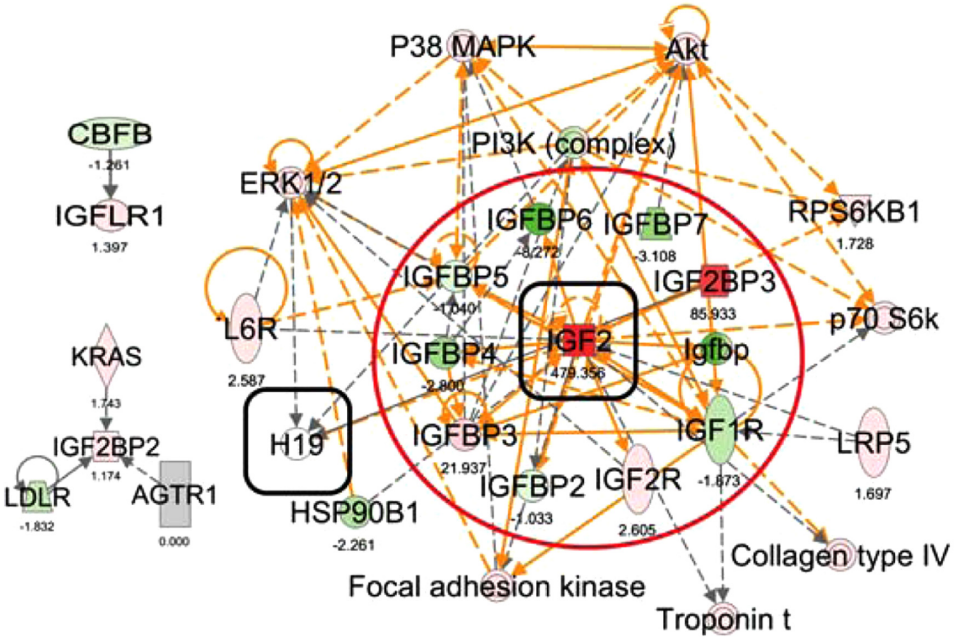

Figure 3. The biologically relevant network: IGF2 and associated genes.
Notes: The dataset was analyzed by the Ingenuity Pathways Analysis software (Ingenuity® Systems, http://www.ingenuity.com). The color indicates the expression level of the genes (red indicating up- regulated genes and green indicating down-regulated genes). Two round-cornered squares highlight the IGF2 and H19 genes.

using real-time RT-PCR and confirmed the RNA-seq findings.[28] Very interestingly, data analyses also detect differentially expressed IGF2 isoforms induced by butyrate in bovine epithelial cells. Two major isoforms of IGF2 mRNA were detected using this method (Table 2). Exon skipping is one of the major forms of alternative splicing, which generates multiple isoforms with different combination of exon sequences.[29] The possible exon skipping events of IGF2 gene were detected in butyrate-treated cells (Fig. 4). The number of junction reads—sequence reads spanning exon-exon junctions—supporting this event was significantly higher (456 fold) in butyrate treated cells than in the control. Filtering parameters, such as the sum of inclusion and exclusion reads greater than 10 (≥1 inclusion read and ≥1 exclusion read), ΔΨ greater than 0.20, and the Bayes factor ≥ 10, were used to detect differentially expressed isoforms. Figure 4 illustrates the isoforms of IGF2 mRNA due to the changes of splicing. This result demonstrates that epigenetic features such as histone modifications and DNA methylation are associated at different levels of transcription regulation such as alternative splicing.

Epigenomic landscapes of butyrate-induced histone modifications across IGF2 gene and associated loci

To demonstrate the fundamental association of the chromatin modification landscape that may play a role in the activation of the IGF2 gene, we mapped epigenomic landscapes of normal histone H3 and H4; we also mapped their major butyrate-induced post-translational modifications, including histone H3 acetylation at the positions lysine 9 (H3K9ac), lysine 18 (H3K18ac), lysine 27 (H3K27ac), H4 acetylation at positions lysine 5 (H4K5ac), lysine 12 (H4K12ac), lysine 16 (H4K16ac), and H4 tetra-acetylation (H4ac4) using Chip-seq. Distribution of histone-associated

Table 2. Butyrate-induced IGF2 mRNA isoforms by alternative splicing.

Transcript ID	Gene	Isoform	Exon*	Control**	Butyrate**	*P* value
ENSBTAT00000044139	IGF2	1	Exon 1–4	4.95	2167.27	0
ENSBTAT00000017372	IGF2	2	Exon 1–5	38.3	19277.44	0

Notes: *Exons involved; **the number denotes normalized counts of junction reads that support the events.

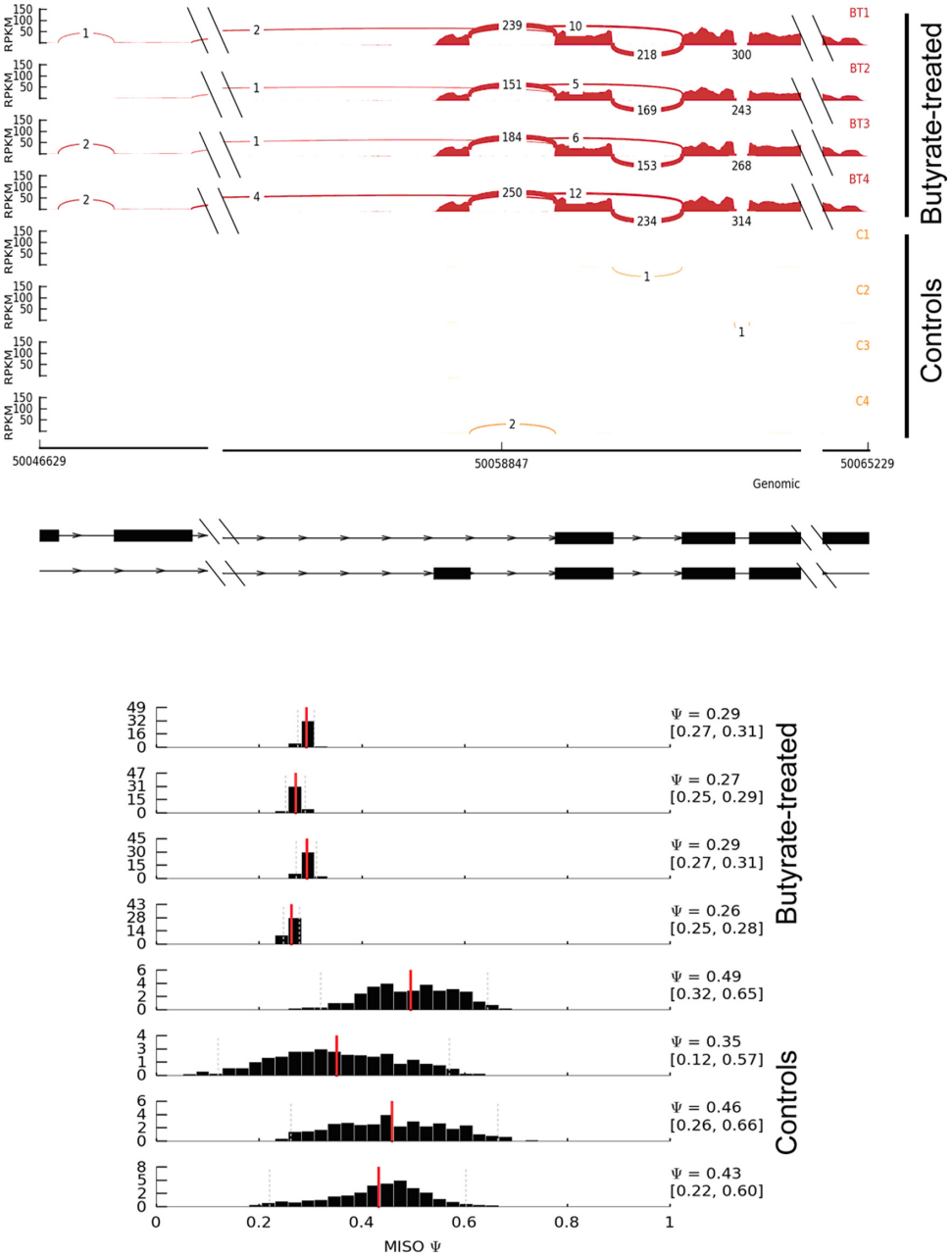

Figure 4. Relative abundance of transcript isoforms of IGF2 gene induced by butyrate in bovine epithelial cells.
Notes: Reads per kilobase of exon model per million mapped reads (RPKM) were displayed in the Y-axis for each sample tested. Mixture-of-Isoforms algorithm (MISO), a probabilistic framework that enables quantification of the expression level of alternatively spliced genes and the identification of differentially regulated isoforms from RNA-Seq data, was used. Arcs with numbers represent junction reads. Bold arcs show the junction supported by >10 junction reads. Diagrams below show the gene structures of two transcript isoforms. Lower panel: the posterior distribution with Ψ value. BT1 to BT4: Four butyrate-treated replicates; C1 to C4: Four Control replicates. Differentially expressed isoforms were detected using the following filtering parameters: the sum of inclusion and exclusion reads is greater than 10 (≥1 inclusion read and ≥1 exclusion read), ΔΨ greater than 0.20, and the Bayes factor ≥ 10.

sequence tags across the IGF2 gene and associated loci, as well as their chromosomes of origin, are mapped and visualized with a Circos visualization tool.[30] Figure 5 shows the integrated global visual representation of expression of IGF2 and its associated genes, the epigenomic landscape of histones H3 and H4, and their major post-translational modifications induced by butyrate. Figure 5 illustrates the major epigenomic landscape alterations induced by butyrate with histone-associated sequence tags within the coding sequence of the each gene (colored according to the figure legend). The plot also exemplifies the deregulated IGF2 and its associated genes and their locations in relation to their genomic position.

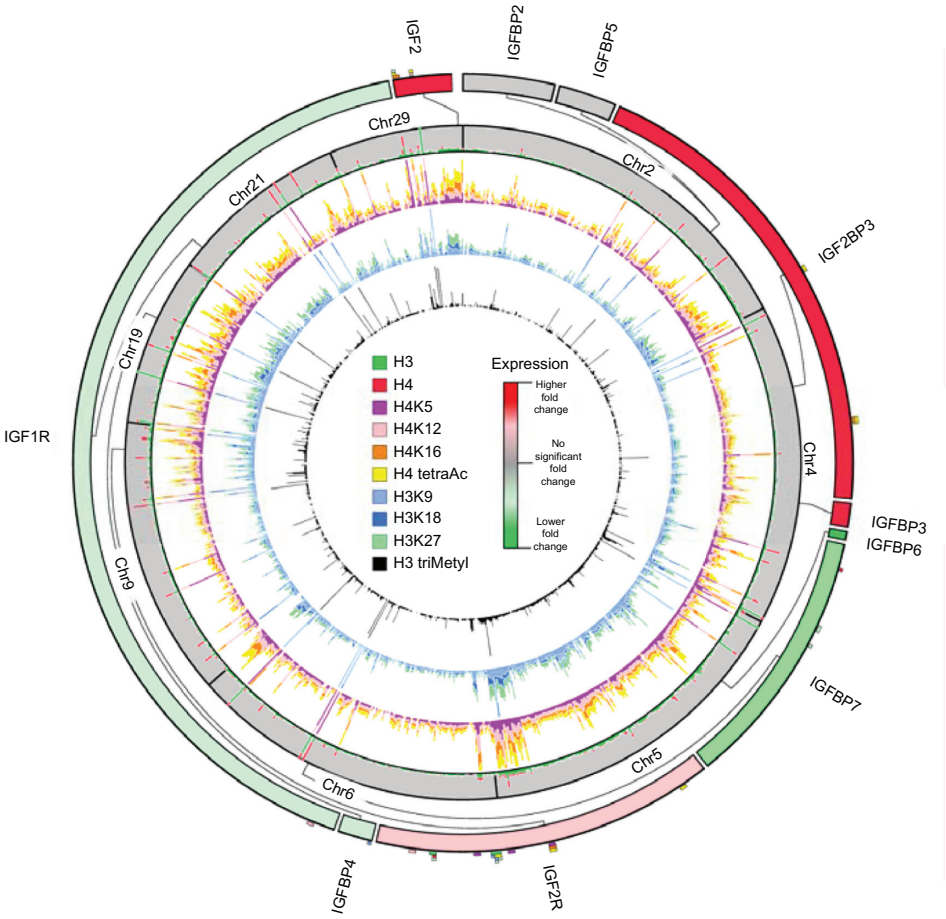

Figure 5. Distribution of histone-associated sequence tags across IGF2 gene associated loci and their chromosomes of origin.
Notes: The outermost track depicts the IGF2 associated genes linked to their chromosome coordinates (grey bars with chromosome number labels). IGF2 associated genes are colored according to their fold change of expression relative to the addition of butyrate with dark green indicating a higher fold change and dark red indicating a fold reduction of expression. Small boxes on the exterior of the genes indicate the presence of histone-associated sequence tags within the coding sequence of the gene (colored according to the figure legend, center). The histogram within the center of the chromosome track represents the presence of H3 and H4 associated sequence tags (green and red; respectively). Histograms on the interior of the chromosome track represent the following (outermost to innermost): (1) H4K5, H4K12, H4K16 and H4 tetra-acetylated associated sequence tags (purple, light red, orange and yellow; respectively); (2) H3K9, H3K18 and H3K27 associated sequence tags (light blue, blue and light green; respectively); (3) H3 tri-methylated associated sequence tags (black).

Epigenomic domains correlated to IGF2/H19 gene structures

To reveal the possible correlation between butyrate-induced histone acetylation and IGF2 activation, the local scale of epigenomic landscape at IGF2/H19 locus was specifically explored by mining our ChIP-seq data. IGF2 and H19, two clustered imprinting but antagonistically co-regulated genes, are reciprocally regulated imprinted genes co-located in the same locus. Our data show that while IGF2 was induced by butyrate-treatment significantly, and as expected, H19 remains unexpressed. We constructed a high-resolution map of the major chromatin modification at the IGF2/H19 locus induced by butyrate, and exemplified the fundamental association of chromatin modification landscape and integrated IGF2/H19 expression as presented in Figure 6 (using the Circos visualization tool). The plot further illustrates the diversity of expression of IGF2 and H19 genes, as well as butyrate induced histone H3 acetylation at the positions, lysine 9 (H3K9ac), lysine 18 (H3K18ac), lysine 27 (H3K27ac), butyrate-induced H4 acetylation at positions lysine 5 (H4K5ac), lysine 12 (H4K12ac), lysine 16 (H4K16ac), and H4 tetra-acetylation (H4ac4) using ChIP-seq. We also mapped histone H3 tri-methylation at position lysine 9 (H3K9me3), which is considered the binding target of Heterochromatin Protein 1a (HP1a) and is generally perceived as an epigenetic silencing mark.[31,32] The distribution diversity of post-translational modified histone at

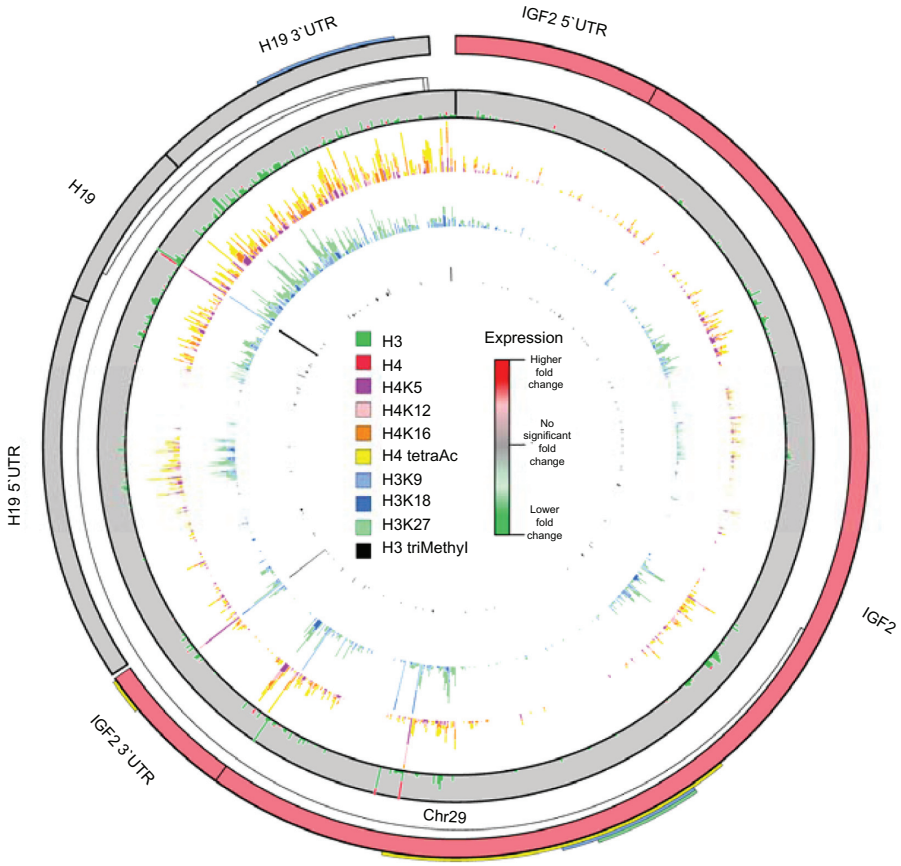

Figure 6. Distribution of histone-associated sequence tags across the IGF2/H19 gene locus.
Notes: The outermost track depicts the IGF2 associated genes linked to their chromosome coordinates (gray bars with chromosome number labels). IGF2 associated genes are colored according to their fold change of expression relative to the addition of butyrate with dark green indicating a higher fold change and dark red indicating a fold reduction of expression. Small boxes on the exterior of the genes indicate the presence of histone-associated sequence tags within the coding sequence of the gene (colored according to the figure legend). The histogram within the center of the chromosome track represents the presence of H3 and H4 associated sequence tags (green and red, respectively). Histograms on the interior of the chromosome track represent the following (outermost to innermost): (1) H4K5, H4K12, H4K16 and H4 tetra-acetylated associated sequence tags (purple, light red, orange and yellow, respectively); (2) H3K9, H3K18 and H3K27 associated sequence tags (light blue, blue and light green, respectively); (3) H3 tri-methylated associated sequence tags (black).

H19 and IGF2 genes was very revealing, with distinct chromatin modification landscape signatures/marks within their gene bodies. While distribution of histone at IGF2 gene was presented with distinctive dispersed domains, distribution of histone present with very high density and lacking the clear boundaries at H19 gene (Fig. 6). The correlation of the histone distribution pattern and gene expression pattern of the IGF2/H19 gene strongly suggest that chromatin modification landscape may play a role in the regulation of expression of IGF2/H19 genes.

To look into the details of gene structure and chromatin modification landscape, we applied all 10 ChIP-seq mapping datasets as histone modification marks in the IGF2/H19 locus and the ChIP-seq chromatin data were visualized in details using a local implementation of the UCSC Genome Browser (Figs. 7 and 8).[19]

While dynamic changes in genome-wide histone distribution pattern were evident due to the acetylation response to a 10 mM butyrate-treatment, there was a strong association between histone modification and the specific types of genomic sequences such as the transcription start site (TSS) and transcription termination site (TTS) of IGF2 (Fig. 7). ChIP-seq mapping and analysis demonstrated that butyrate induced acetylation developed as dispersed epigenomic domains at TSS and immediate up-stream of the region, the coding region, and TTS of IGF2 gene. These domains displayed strikingly distinct distribution patterns of chromatin modification. In domain 1, which covers the TSS region of the ING2 gene, there are several tightly concentrated binding peaks of histone modifications, whereas in domains 2 and 4 all histone modifications were presented as diffused patterns.

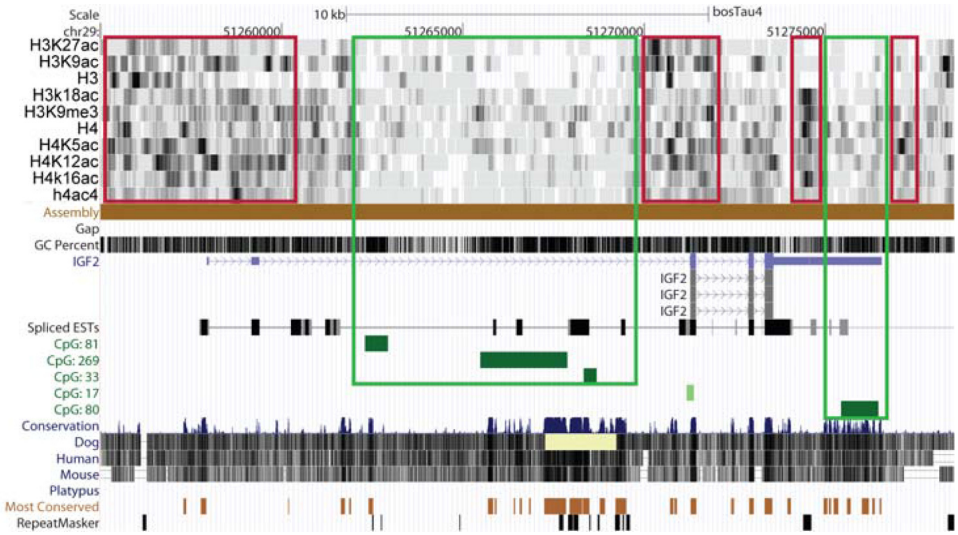

Figure 7. UCSC browser view of epigenomic data of IGF2 gene locus: Butyrate-induced epigenetic landscapes modifications are distributed as dispersed epigenomic domains.
Notes: Distribution of histone H3 acetylation at the positions of lysine 9 (H3K9ac), lysine 18 (H3K18ac), lysine 27 (H3K27ac); butyrate-induced H4 acetylation at positions lysine 5 (H4K5ac), lysine 12 (H4K12ac), lysine 16 (H4K16ac), H4 tetra-acetylation (H4ac4); as well as histone H3 tri-methylation at the position lysine 9 (H3K9me3) enriched region in the proximal promoter at the 5′ untranscribed region (UTR) from the transcription start site (TSS), exon, intron, and intergenic region. The grayscale represents the enriched region for all ten ChIP-Seq data sets and at least 6 domains with either high density binding sites or diffused distribution domains are presented at IGF2 gene. Red marks the high density chromatin distribution regions and green marks the differentially methylated regions (DMRs) with defused background and lacking significant peaks of chromatin distribution.

Interestingly, these diffused regions co-located with differentially methylated regions (DMRs). On the other hand, while the peak distribution of the ChIP-seq reads (presented as grayscale) of normal histone 3 and histone 4 (H3 and H4) exhibited mostly as a typical priority background distribution pattern, there were few high density binding sites of H3 and H4 at upstream of IGF2 gene. Specifically, there were some very interesting features among the different H3 or H4 modifications (acetylation or methylation at the different lysine residues of histone H3 and H4 tails). Among these changes, acetylation at H3K9, H4K5, H4K12 and H4K16 showed dramatically high intensity of binding at TSS and the promoter region of IGF2 (Fig. 7). However acetylation at H3K27, H3K18, and especially trimethylation at

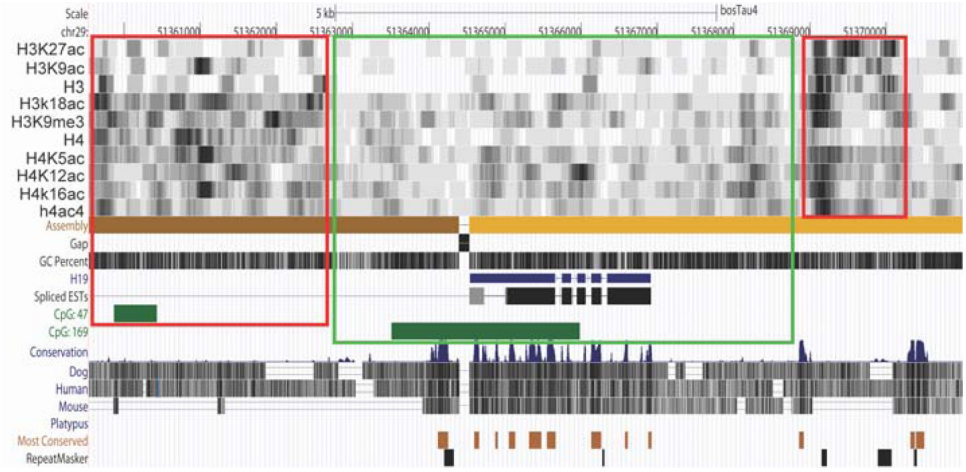

Figure 8. UCSC browser view of epigenomic data of H19 gene locus.
Notes: Distribution of histone H3 acetylation at the positions of lysine 9 (H3K9ac), lysine 18 (H3K18ac), lysine 27 (H3K27ac); butyrate-induced H4 acetylation at positions lysine 4 (H4K5ac), lysine 12 (H4K12ac), lysine 16 (H4K16ac), H4 tetra-acetylation (H4ac4); and histone H3 tri-methylation at the position lysine 9 (H3K9me3) enriched region in the proximal promoter at the 5 untranscribed region (UTR) from the transcription start site (TSS), exon, intron, and intergenic region. Red marks the high density chromatin distribution regions and the green regions mark the differentially methylated regions (DMRs) regions and defused background and lacking significant peaks of chromatin distribution regions.

H3K9 (H3K9me3), occurred at a much lower intensity or a narrower range of nucleotides near TSS of IGF2. While the histone modifications detected at the TSS of IGF2 were intense and varied among the acetylation or methylation of histone H3 or H4, in the regions containing CpG islands (CpG:81, CpG:269 and CpG:33) the coding and/or intergenic regions of IGF2 gene did not display any detectable peaks of histone modifications.

The detailed chromatin landscape of another imprinting gene H19, which is antagonistically co-regulated and co-located in the same cluster with IGF2, also displayed a distinctive pattern (Fig. 8). In the H19 gene, CpG islands (CpG: 169) overlap the TSS of the gene. In contrast to the chromatin landscape of IGF2 genes, the histone modifications were diffuse and no significant histone acetylation peaks were evident in the H19 gene, even at the regions near TSS and TTS. However, at the up-stream intergenic region of H19 gene, which contains a putative imprinting control region (ICR),[33] moderate to high levels of modified histone acetylation were detected in a region of about 3–4 kb, including epigenetic silencing marks of H3K9me3 (Fig. 3B). The epigenetic landscape of H3K9ac, which is associated with transcriptional activity,[34,35] was also distinguishable at the H19 and IGF2 genes. No enrichment can be seen at TSS, TTS or the intergenic region of the H19 gene for this histone modification. Two regions of high density enrichment of modified histone H3 and histone H4 were located at about 2–3 kb upstream or downstream of H19 gene, respectively. Interestingly, an ICR for IGF2/H19 cluster is located 2 kb upstream of the TSS of the H19 gene. This ICR has been suggested to act as an enhancer-blocking element controlling the imprinted expression of the IGF2 gene.[32] Additionally, a modified H3 binding domain with a relatively high density was evident at the downstream region of H19 gene.[36]

Discussion

The protein product of IGF2 gene is insulin-like growth factor 2. This protein has an essential function in growth and development before birth. IGF2 promotes the growth and proliferation of cells in many different tissues. Although the IGF2 gene is highly active during fetal development, it is much less active in the adult body due to the phenomenon called genomic imprinting. IGF2 is a part of gene cluster with H19 gene which is also involved in growth and development. Abnormal expression of IGF2 has been linked with several human diseases such as Beckwith-Wiedemann syndrome and several types of cancer. In this study, we report the first comprehensive approach for the epigenetic changes correlated with butyrate-induced activation of IGF2 gene in bovine cells. We profiled histone H3 acetylation at the positions lysine 9 (H3K9ac), lysine 18 (H3K18ac), lysine 27 (H3K27ac); H4 acetylation at positions lysine 5 (H4K5ac), lysine 12 (H4K12ac), lysine 16 (H4K16ac); and H4 tetra-acetylation (H4ac4) using ChIP-seq. We also mapped histone H3 tri-methylation at the position lysine 9 (H3K9me3), which is considered the binding target of Heterochromatin Protein 1a (HP1a) and is generally perceived as an epigenetic silencing mark.[31,32]

Using microarray[3] and RT-PCR,[33] as well as RNA-seq techniques in this report, we have confirmed indisputably that butyrate-induced histone acetylation can activate IGF2 expression. A similar finding was reported in the human fibroblast cell line (HSK09) and proved the activation of IGF2 is due to the loss of imprinting (LOI).[5] Our results also indicated that IGF1 was below detection and IGF1R was down regulated by the butyrate treatment. Both IGF2BP3 and IGFBP3 were also significantly up-regulated by butyrate treatment. Expression of IGF2 and IGF2R was significantly induced by butyrate, suggesting that IGF2, not IGF1, may play an essential role in butyrate-induced cell cycle progression and programmed cell death by apoptosis; over-expression of IGF2R induced by butyrate may be involved in pro-apoptotic processes in MDBK cells. While the functions of IGFBPs seemed diverse, IGFBP3 and IGF2BP3 were shown to be the primary regulators of butyrate-induced cell growth arrest and apoptosis.

We constructed a high-resolution map of the major chromatin modification at the IGF2/H19 locus induced by butyrate and to illustrate the fundamental association of chromatin modification landscape that may play a role in the activation of IGF2 gene. Due to the specific characteristics of IGF2 as an imprinting gene, we concentrated our investigation on the specific epigenetic landscape of the IGF2/H19 cluster induced by butyrate. ChIP-seq data reveals that IGF2 may be regulated by dispersed epigenomic

domains, which show strong correlation to the structure of IGF2 gene, as well as the clustered imprinting gene H19. Our results provide insights into the mechanism of butyrate induced loss of imprinting (LOI) of IGF2. Out data also extended the observation as it was reported in mouse experiments that there are allele specific epigenetic states such as DNA methylation and histone modification that may act in concert in setting up and maintaining reciprocal parental allelic histone acetylation at DMRs.[37]

In eukaryotes, the 'core promoter' is often located in the immediate vicinity of the TSS, which is assumed to dock the pre-initiation complex for RNA transcription. The main functional classes of metazoan promoters are configured with promoter signals, TSS position, nucleosome positions, and their epigenetic marks.[38] The general assessment of the epigenomic landscape at the IGF2/H19 cluster in this study shows the overall patterns of histone acetylation at these two genes. Our data revealed that IGF2 and H19, two clustered imprinting but antagonistically co-regulated genes, have distinct chromatin modification landscape signatures/marks within their gene bodies and functional structures such as TSSs and TTSs. While a high degree of histone acetylation was detected at the TSS and TTS of IGF2, the mid-body of the IGF2 gene had a reduced background of histone modifications with few significant peaks. A similarly reduced background pattern of histone modification covered the whole H19 gene body, as well as a 2–4 kb extension upstream and downstream of the H19 gene. There is some high to moderate intensity of modified histone binding domains at the intergenic region of the up-stream H19 gene, which has been characterized as a putative imprinting control region (ICR). Those histone modifications include epigenetic silencing marks of H3K9me3. These patterns suggest that many different histone modifications at different lysines must act in close correlation in regulating and maintaining reciprocal activation of imprinted genes in a cluster such as IGF2 and H19. The correlation between the DMRs of IGF2/H19 and chromatin modifications such as histone acetylation indicates that DNA methylation, histone methylation, and histone acetylation as epigenetic marks associated with imprinted genes at IGF2/H19 are differentially distributed on the expressed and silenced alleles. Evidence also indicates the correlation of butyrate induced regional changes of acetylation within the upstream and downstream regulation domain of H19 with the reduced expression of H19 and activation of expression of IGF2.

The mechanism of genomic imprinting is very complicated because some imprinted genes, such as IGF2 and H19, may be controlled by other genes which are also imprinted. Both IGF2 and H19 genes are co-located at one cluster about 88 kb apart on bovine chromosome 29, and are expressed from the maternal and paternal alleles, respectively.[39] The cluster of two gene clusters is very conservative and two genes are reciprocally imprinted and regulated. Although it has been known for a long time that IGF2 and H19 are imprinted in bovine, underlying mechanisms of imprinting regulation have not been fully understood.[40,41] Clustered imprinted genes are regulated by differentially methylated (DMRs) imprinting control regions (ICRs) that affect gene activity, and repression in *cis* over a large region. Although a primary imprint signal for each of these clusters is DNA methylation, different mechanisms are used to establish and maintain these marks.[33] The majority of ICRs are methylated in the maternal germline; the methylated ICRs are usually functioning as promoters for antisense transcripts whose elongation is associated with imprinting control in the domain. By contrast, ICRs methylated in the paternal germline do not appear to act as promoters and are located between genes. At least one ICR at the Igf2/H19 locus is known to function as an insulator. Analysis of ICRs suggests that maternal and paternal methylation imprints function in distinct ways.[33]

Histone modification has been implied to play a very important role in maintaining gene imprinting. It was determined that when human fibroblast cells were treated with HDACs such as butyrate or trichostatin, the IGF2 gene expressed biallelically (loss of imprinting, LOI), indicating that in addition to DNA methylation, differential histone acetylation of two parental alleles may be another potential mechanism by which the imprinting of *IGF*2 is regulated; this is done probably through changes in the local chromatin structure of the imprinted locus.[41] Our results showed

that butyrate significantly up-regulated the mRNA expression of IGF2 gene in the bovine epithelial cell (FDR $<$ 0.0001). Our results also suggest that butyrate differentially regulates the expression of the IGF2 gene at the gene and isoform levels. Alternative splicing contributes to protein diversities and results in mRNA forming coding for proteins with different chemical and biological activity. No known isoforms in IGF2 genes have been annotated in cattle so far. The regulation of exon skipping by butyrate remains unknown. It is possible that histone modification induced by butyrate may leave certain marks on newly synthesized mRNAs that were attributed to the disruption of exonic splicing enhancers.[42,43] Our data, however, are consistent with the recent report[44] that epigenetic features such as histone modifications and DNA methylation are significantly associated with alternative splicing. Our results provided a snapshot into complex transcriptome dynamics regulated by butyrate, which will facilitate our understanding of biological effect of butyrate and other HDAC inhibitors.

Conclusion

Chromatin modification has clearly emerged as a very important mechanism in regulating the transcriptional status of the genome. Our previous studies revealed that VFAs, especially butyrate, participate in metabolism both as nutrients and as regulators of histone modification, thereby regulating the 'epigenomic code'.[7,12,28,45] In this study, we confirmed that butyrate can activate the IGF2 gene through its HDAC inhibitory activity. ChIP-seq high-definition mapping of the epigenomic landscape modification induced by butyrate reveals that IGF2 may be regulated by dispersed epigenomic domains, which show strong correlation to the structure of IGF2 gene, as well as the clustered imprinting gene H19. Our results provide insights into the mechanism of butyrate induced loss of imprinting (LOI) of IGF2 and broadened our knowledge of regulation of gene expression by histone modification.

Acknowledgements

Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U S Department of Agriculture. The USDA is an equal opportunity provider and employer.

Author Contributions

Conceived and designed the experiments: CJL, RWL. Analysed the data: YG, JHS, DMB, GEL WL and SW. Wrote the first draft of the manuscript: CJL. Contributed to the writing of the manuscript: RWL. Agree with manuscript results and conclusions: JHS, RWL, YG, DMB, GEL, WL, SW and CJL. Jointly developed the structure and arguments for the paper: CJL and RWL. Made critical revisions and approved final version: CJL. All authors reviewed and approved of the final manuscript.

Funding

SW and WL were supported by Award R01HG005978 from the National Human Genome Research Institute (NHGRI) and CAMERA project funded by Gordon and Betty Moore Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NHGRI or the National Institutes of Health.

Competing Interests

Authors disclose no potential conflicts of interest.

Disclosures and Ethics

As a requirement of publication the authors have provided signed confirmation of their compliance with ethical and legal obligations including but not limited to compliance with ICMJE authorship and competing interests guidelines, that the article is neither under consideration for publication nor published elsewhere, of their compliance with legal and ethical guidelines concerning human and animal research participants (if applicable), and that permission has been obtained for reproduction of any copyrighted material. This article was subject to blind, independent, expert peer review. The reviewers reported no competing interests.

References

1. Stewart CE, James PL, Fant ME, Rotwein P. Overexpression of insulin-like growth factor-II induces accelerated myoblast differentiation. *J Cell Physiol*. 1996;169(1):23–32.
2. Stewart CE, Rotwein P. Growth, differentiation, and survival: multiple physiological functions for insulin-like growth factors. *Physiol Rev*. 1996;76(4): 1005–26.

3. Chao W, D'Amore PA. IGF2: epigenetic regulation and role in development and disease. *Cytokine Growth Factor Rev*. 2008;19(2):111–20.
4. Jones JI, Clemmons DR. Insulin-like growth factors and their binding proteins: biological actions. *Endocr Rev*. 1995;16(1):3–34.
5. Liu B, Lee HY, Weinzimer SA, et al. Direct functional interactions between insulin-like growth factor-binding protein-3 and retinoid X receptor-alpha regulate transcriptional signaling and apoptosis. *J Biol Chem*. 2000;275(43): 33607–13.
6. Li CJ, Elsasser TH. Butyrate-induced apoptosis and cell cycle arrest in bovine kidney epithelial cells: involvement of caspase and proteasome pathways. *J Anim Sci*. 2005;83(1):89–97.
7. Li RW, Li C. Butyrate induces profound changes in gene expression related to multiple signal pathways in bovine kidney epithelial cells. *BMC Genomics*. 2006;7:234.
8. Goldberg AD, Allis CD, Bernstein E. Epigenetics: a landscape takes shape. *Cell*. 2007;128(4):635–8.
9. Berger SL. The complex language of chromatin regulation during transcription. *Nature*. 2007;447(7143):407–12.
10. Bugaut M. Occurrence, absorption and metabolism of short chain fatty acids in the digestive tract of mammals. *Comp Biochem Physiol B*. 1987;86(3): 439–72.
11. Grandjean V, O'Neill L, Sado T, Turner B, Ferguson-Smith A. Relationship between DNA methylation, histone H4 acetylation and gene expression in the mouse imprinted Igf2-H19 domain. *FEBS Lett*. 2001;488(3):165–9.
12. Shin JH, Li RW, Gao Y, Baldwin R 6th, Li CJ. Genome-wide ChIP-seq mapping and analysis reveal butyrate-induced acetylation of H3K9 and H3K27 correlated with transcription activity in bovine cells. *Funct Integr Genomics*. 2012;12(1):119–30.
13. Li RW, Gasbarre LC. A temporal shift in regulatory networks and pathways in the bovine small intestine during Cooperia oncophora infection. *Int J Parasitol*. Jun 2009;39(7):813–24.
14. Li RW, Schroeder SG. Cytoskeleton remodeling and alterations in smooth muscle contractility in the bovine jejunum during nematode infection. *Funct Integr Genomics*. 2012;12(1):35–44. doi: 10.1007/s10142-011-0259-7.
15. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25. doi: 10.1186/gb-2009-10-3-r25.
16. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137. doi: 10.1186/gb-2008-9-9-r137.
17. Feng J, Liu T, Zhang Y. Using MACS to identify peaks from ChIP-Seq data. *Curr Protoc Bioinformatics*. Jun 2011;Chapter 2:Unit 2.14. doi: 10.1002/0471250953.bi0214s34.
18. Salmon-Divon M, Dvinge H, Tammoja K, Bertone P. PeakAnalyzer: genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics*. 2010;11:415.
19. Karolchik D, Hinrichs AS, et al. The UCSC Genome Browser. Current protocols in bioinformatics/editoral board, Andreas D. Baxevanis ... [et al.] Chapter 1: Unit 14.
20. Li RW, Rinaldi M, Capuco AV. Characterization of the abomasal transcriptome for mechanisms of resistance to gastrointestinal nematodes in cattle. *Vet Res*. 2011;42(1):114. doi: 10.1186/1297-9716-42-114.
21. Calvano SE, Xiao W, Richards DR, et al. A network-based analysis of systemic inflammation in humans. *Nature*. 2005;437(7061):1032–7.
22. Mayburd AL, Martlínez A, Sackett D, et al. Ingenuity network-assisted transcription profiling: Identification of a new pharmacologic mechanism for MK886. *Clin Cancer Res*. 2006;12(6):1820–7.
23. Pospisil P, Iyer LK, Adelstein SJ, Kassis AI. A combined approach to data mining of textual and structured data to identify cancer-related targets. *BMC Bioinformatics*. 2006;7:354.
24. Abdel-Aziz HO, Takasaki I, Tabuchi Y, et al. High-density oligonucleotide microarrays and functional network analysis reveal extended lung carcinogenesis pathway maps and multiple interacting genes in NNK [4-(methylnitrosamino)-1-(3-pyridyle)-1-butanone] induced CD1 mouse lung tumor. *J Cancer Res Clin Oncol*. 2007;133(2):107–15.
25. Su YQ, Sugiura K, Woo Y, et al. Selective degradation of transcripts during meiotic maturation of mouse oocytes. *Dev Biol*. 2007;302(1):104–17.
26. Li RW, Li CJ. Effects of butyrate on the expression of insulin-like growth factor binding proteins in bovine kindery epithelial cells. *The Open Veternary Science Journal*. 2007;1:14–9.
27. Katz Y, Wang ET, Airoldi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods*. 2010;7(12):1009–15.
28. Li RW, Wu S, Baldwin RL 6th, Li W, Li C. Perturbation dynamics of the rumen microbiota in response to exogenous butyrate. *PLoS One*. 2012;7(1):e29392. doi: 10.1371/journal.pone.0029392.
29. Miriami E, Margalit H, Sperling R. Conserved sequence elements associated with exon skipping. *Nucleic Acids Res*. 2003;31(7):1974–83.
30. Krzywinski M, Schein J, Birol I. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19(9):1639–45.
31. Bannister AJ, Zegerman P, Partridge JF, et al. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*. 410(6824):120–4.
32. Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*. 2001;410(6824):116–20.
33. Edwards CA, Ferguson-Smith AC. Mechanisms regulating imprinted genes in clusters. *Curr Opin Cell Biol*. 2007;19(3):281–9.
34. Roh TY, Cuddapah S, Zhao K. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev*. 2005;19(5):542–52.
35. Yin H, Sweeney S, Raha D, Snyder M, Lin H. A high-resolution whole-genome map of key chromatin modifications in the adult Drosophila melanogaster. *PLoS Genet*. 2011;7(12):e1002380. doi: 10.1371/journal.pgen.1002380.
36. Baxter D, McInnes IB, Kurowska-Stolarska M. Novel regulatory mechanisms in inflammatory arthritis: a role for microRNA. *Immunol Cell Biol*. 2012;90(3):288–92.
37. Singh P, Cho J, Tsai SY, Rivas GE, Larson GP, Szabó PE. Coordinated allele-specific histone acetylation at the differentially methylated regions of imprinted genes. *Nucleic Acids Res*. 2010;38(22):7974–90.
38. Lenhard B, Sandelin A, Carninci P. Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat Rev Genet*. 2012;13(4):233–45.
39. Curchoe CL, Zhang S, Yang L, Page R, Tian XC. Hypomethylation trends in the intergenic region of the imprinted IGF2 and H19 genes in cloned cattle. *Anim Reprod Sci*. 2009;116(3–4):213–25.
40. Curchoe C, Zhang S, Bin Y, et al. Promoter-specific expression of the imprinted IGF2 gene in cattle (Bos taurus). *Biol Reprod*. 2005;73(6):1275–81.
41. Hu JF, Oruganti H, Vu TH, Hoffman AR. The Role of Histone Acetylation in the Allelic Expression of the Imprinted Human Insulin-like Growth Factor II Gene. *Biochem Biophys Res Commun*. 1998;251(2):403–8.
42. Cartegni L, Chew SL, Krainer AR, et al. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet*. 2002;3(4):285–98.
43. Cartegni L, Krainer AR. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. *Nat Genet*. 2002;30(4):377–84.
44. Zhou Y, Lu Y, Tian W. Epigenetic features are significantly associated with alternative splicing. *BMC Genomics*. 2012;13:123. doi: 10.1186/1471-2164-13-123.
45. Li CJ, Li RW, Wang YH, Elsasser TH. Pathway analysis identifies perturbation of genetic networks induced by butyrate in a bovine kidney epithelial cell line. *Funct Integr Genomics*. 2007;7(3):193–205.
46. Gao ZH, Suppola S, Liu J, Heikkilä P, Jänne J, Voutilainen R. Association of H19 promoter methylation with the expression of H19 and IGF-II genes in adrenocortical tumors. *J Clin Endocrinol Metab*. 2002;87(3):1170–6.