scientific reports

OPEN

Comparative analysis of histone H3K4me3 modifications between blastocysts and somatic tissues in cattle

Mao Ishibashi, Shuntaro Ikeda [ID][✉] & Naojiro Minami

Epigenetic changes induced in the early developmental stages by the surrounding environment can have not only short-term but also long-term consequences throughout life. This concept constitutes the "Developmental Origins of Health and Disease" (DOHaD) hypothesis and encompasses the possibility of controlling livestock health and diseases by epigenetic regulation during early development. As a preliminary step for examining changes of epigenetic modifications in early embryos and their long-lasting effects in fully differentiated somatic tissues, we aimed to obtain high-throughput genome-wide histone H3 lysine 4 trimethylation (H3K4me3) profiles of bovine blastocysts and to compare these data with those from adult somatic tissues in order to extract common and typical features between these tissues in terms of H3K4me3 modifications. Bovine blastocysts were produced in vitro and subjected to chromatin immunoprecipitation-sequencing analysis of H3K4me3. Comparative analysis of the blastocyst-derived H3K4me3 profile with publicly available data from adult liver and muscle tissues revealed that the blastocyst profile could be used as a "sieve" to extract somatic tissue-specific modifications in genes closely related to tissue-specific functions. Furthermore, principal component analysis of the level of common modifications between blastocysts and somatic tissues in meat production-related and imprinted genes well characterized inter- and intra-tissue differences. The results of this study produced a referential genome-wide H3K4me3 profile of bovine blastocysts within the limits of their in vitro source and revealed its common and typical features in relation to the profiles of adult tissues.

The periconceptional period of mammalian embryonic development is a critical window during which diverse environmental factors surrounding the embryo have not only short-term consequences such as effects on immediate embryonic development, but also long-term consequences including lasting influences on metabolic, developmental, and etiologic processes throughout gestation and even during postnatal and adult life[1]. During preimplantation development, dynamic epigenetic rearrangements occur, including substantial changes in DNA methylation and histone modifications, which regulate specific and heritable patterns of gene expression[2–4]. As the epigenome is dynamically formed during the preimplantation period, environmental intervention-induced changes in epigenome formation during this period have been considered as one of the possible causes of the long-lasting influences induced by the periconceptional environment[1]. This concept constitutes the "Developmental Origins of Health and Disease" (DOHaD) hypothesis and encompasses the possibility of controlling health and diseases in later life by epigenetic regulation during early development.

Some phenotypic changes in the field of livestock production can be discussed in the context of DOHaD given that early life events during prenatal and early postnatal development often affect traits, including those of economic importance[5–7]. For example, in vitro handling of ruminant early embryos in assisted reproductive technology increases the risk of fetal overgrowth syndrome[8,9], whereas maternal nutrition, stress, or illness during pregnancy can affect productive traits such as postnatal growth, milk yield, carcass composition, and fertility[5–7].

Although the epigenetic modifications that occur during the early developmental period, which persist in differentiated tissues, are considered a major mechanism of DOHaD, there is few comparative studies of histone modifications between early embryos and fully differentiated somatic tissues[10,11]. As a preliminary step for examining the changes of epigenetic modifications in early embryos and their long-lasting effects in later life, the

Laboratory of Reproductive Biology, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan. [✉]email: ikeda.syuntaro.6u@kyoto-u.ac.jp

elucidation of the common and typical features of epigenetic modifications between these two developmental stages is worth studying. In particular, accumulating evidence suggests that histone modifications at developmentally important genes, compared with DNA methylation, are more susceptible to the surrounding environment during the preimplantation period[12,13]. In the present study, we aimed to obtain genome-wide profiles of histone H3 lysine 4 trimethylation (H3K4me3), a representative marker of active chromatin, in bovine blastocysts produced in vitro and to compare these profiles with those of adult liver and muscle tissues, which have been deposited in public databases. We examined whether we can extract the common and typical H3K4me3 features of early embryos and adult somatic tissues.

Results

H3K4me3 profile in bovine blastocysts.
We performed Chromatin Immunoprecipitation (ChIP)-seq analysis of H3K4me3 using three groups of bovine blastocysts (n = ~11 per replicate) derived from two independent IVF procedures. Pairwise comparisons of the ChIP signals in the triplicates showed the high reproducibility of our method (Supplementary Fig. S1). We detected about 20,000 significant peaks throughout the genome (Supplementary Table S1 and Fig. 1a). Approximately 20% of the peaks were located on gene promotor regions (Fig. 1b). Figure 1c shows a snapshot of the H3K4me3 landscape in a 50-kb region (chromosome 5) that encompasses the transcription start site (TSS) of *GAPDH*, which is a representative positive region for H3K4me3 modifications[14]. Figure 1c demonstrates the clear enrichment of H3K4me3 at this region. Average profile plotting of the H3K4me3 signal around the genome-wide TSS regions showed similar profiles among the replicates and exhibited asymmetric bimodal peaks with a valley at TSSs that are considered to be nucleosome-free regions[15] (Fig. 1d). As expected, average profile plotting around gene bodies categorized by gene groups with different expression levels revealed that highly expressed gene groups had more extensive H3K4me3 modifications (Fig. 1e).

Characterization of embryonic- and somatic tissue-specific H3K4me3 modifications.
We explored the tissue-specific H3K4me3 modifications between preimplantation embryos and somatic tissues. First, 14,018 overlapping peaks identified from two liver ChIP datasets in a public database (bulls 4 and 5 in E-MTAB-2633[16]) were designated as the liver peaks. On the other hand, the 20,298 overlapping peaks identified from our ChIP data from blastocysts (blastocysts 1 and 3, which exhibited the two highest peak numbers) were designated as the blastocyst peaks. Then, we merged these peak groups and extracted 1899 and 7901 liver- and blastocyst-specific peaks, respectively (Fig. 2a). From the genes that harbored these tissue-specific peaks within ± 3000 bp of the TSS, those with the top 100 peak occupancy rates were subjected to gene ontology (GO) analysis using the web-based DAVID tool[17,18]. As a result, the genes with liver-specific peaks enriched the GO terms closely related to liver function such as "organic acid metabolic processes", whereas the genes with blastocyst-specific peaks significantly enriched embryonic development-related GO terms such as "embryo development" and "cell fate commitment" (Fig. 2b). Such tissue function-related GO enrichment was not obtained from the liver and blastocyst peaks, without subtraction of the common peaks; in other words, analysis using these peak-associated genes generated only common and broad GO terms (Table S2). Figure 2c shows the representative liver- and blastocyst-specific peaks around the TSSs of *ARG1* and *GATA2*, respectively.

We applied the same strategy to characterize muscle tissue in terms of the H3K4me3 profile. We used publicly available H3K4me3 data from the longissimus dorsi muscle of beef cattle[19] and compared them with our blastocyst data. GO analysis using the muscle peaks generated broad terms again; however, the top ten significant GO terms generated by the muscle-specific peaks extracted by comparison with the blastocyst peaks all contained the word "muscle" (Fig. 3a,b). The genes participating in the enriched GO terms identified muscle function-related genes with H3K4me3 modifications (Fig. 3c). In addition, when blastocyst-specific peaks relative to liver (Fig. 2a) were further narrowed down by subtracting muscle peaks and subjected to the GO analysis, the result indicated more embryo development-oriented GO terms than the result using only liver peaks did (Table S3).

Characterization of common H3K4me3 modifications in blastocysts and liver.
We next characterized the genes that harbored common H3K4me3 modifications in blastocyst and liver. Among the publicly available bovine H3K4me3 data from somatic tissues (liver and muscle), the liver data were relatively comparable with our blastocyst data in terms of available read numbers, average profiles of the gene groups of interest (Figs. 4a and 5a), and the enrichment profiles of representative positive regions (Fig. S2a). However, the muscle data did not meet these criteria (Fig. S2a,b) and we did not analyze these data further. At first, we focused on meat production-related genes given that meat production is one of the most important economic traits in beef cattle. We used the list of meat production-related genes produced by Williams et al. on the basis of biological roles, which might influence muscle development, structure, metabolism, or meat maturation[20]. The 504 listed genes were curated with RefSeq mRNA accession numbers and official gene symbols, which narrowed down the gene number to 438. Of these, 216 harbored H3K4me3 peaks within ± 3000 bp of TSSs both in blastocysts and liver, and 203 genes were recognized by the ngsplot program[21] to calculate peak area around the TSSs. The processed gene list is shown in Table S4. Figure 4a shows the average profile plot of the 203 genes in the samples (three blastocyst and two liver samples). Given that the overall profiles were similar, we considered that the study-dependent bias in the peak area was negligible and the peak area could be compared among the samples. Therefore, the peak areas at these genes were subjected to Principal component analysis (PCA). As a result, PC1 well characterized the inter-tissue differences of H3K4me3 modifications, the levels of which are largely different between blastocysts and liver, even though they are common modifications (Fig. 4b,c). For example, *CROT* had lower H3K4me3 levels in blastocysts (autoscaled area [mean ± standard deviation], − 0.68 ± 0.26) compared with liver (1.02 ± 0.62), whereas *ALDH5A1* exhibited the opposite pattern (0.71 ± 0.23 for blastocysts and − 1.07 ± 0.14
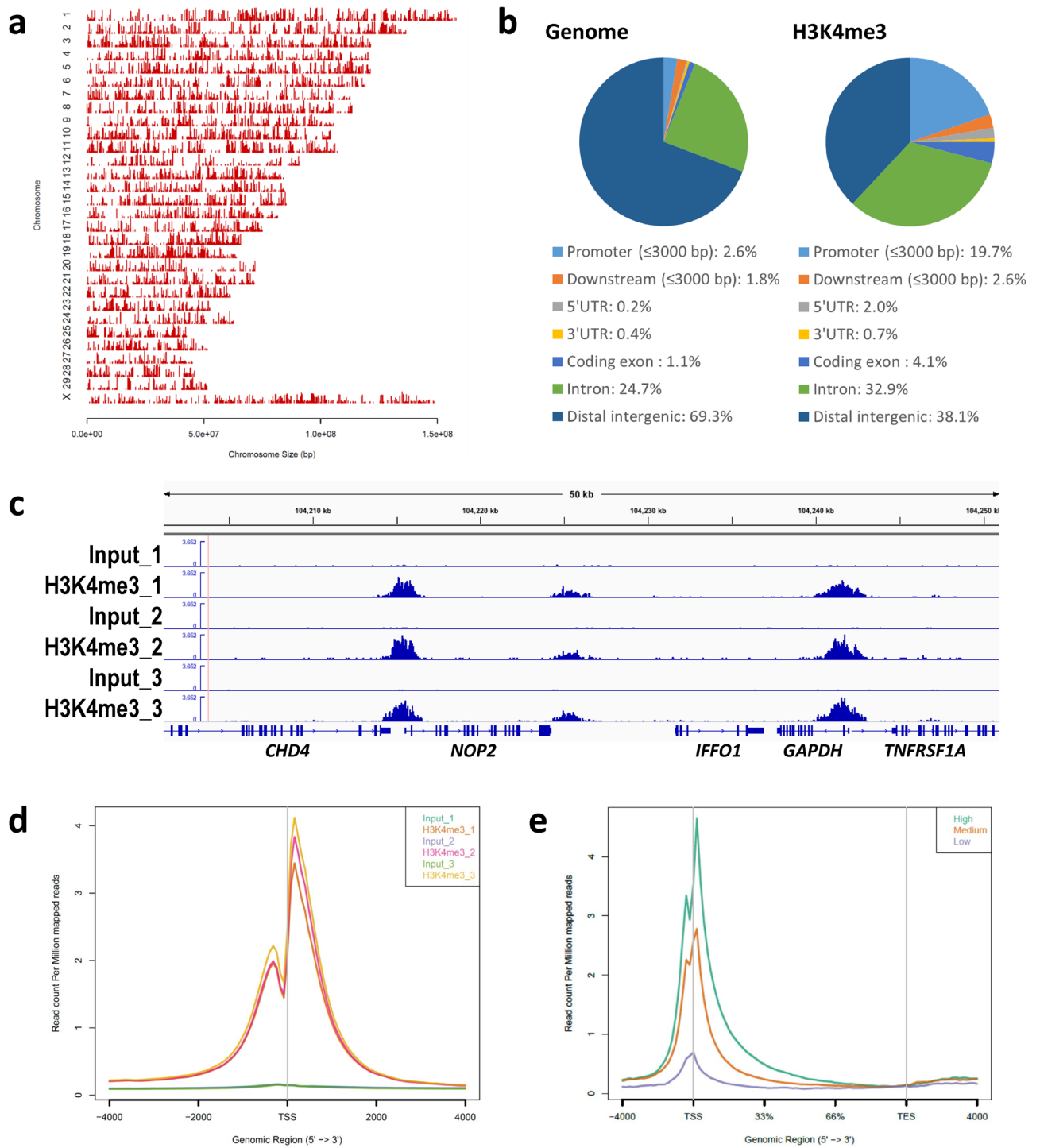
Figure 1. Overview of H3K4me3 ChIP-seq results in bovine blastocysts. (**a, b**) Distribution of H3K4me3 peaks in each chromosome (**a**) and in corresponding genic and intergenic regions (**b**). These figures were generated by the CEAS[35] tool using the H3K4me3_1 (Blastocyst_1) sample. (**c**) Snapshot of the H3K4me3 landscape in a 50-kb region (chromosome 5) that encompasses the *GAPDH* TSS. The ChIP peaks in the triplicates were visualized using the Integrative Genomics Viewer[36]. (**d**) Average profile plot of the H3K4me3 signal around the genome-wide TSSs. The signals from the triplicates are shown. (**e**) Average profile plot around gene bodies categorized by gene groups with different expression levels based on GSE52415[37]. The H3K4me3_1 sample was used to generate the figure. The average profile plots were generated by ngs.plot[21].
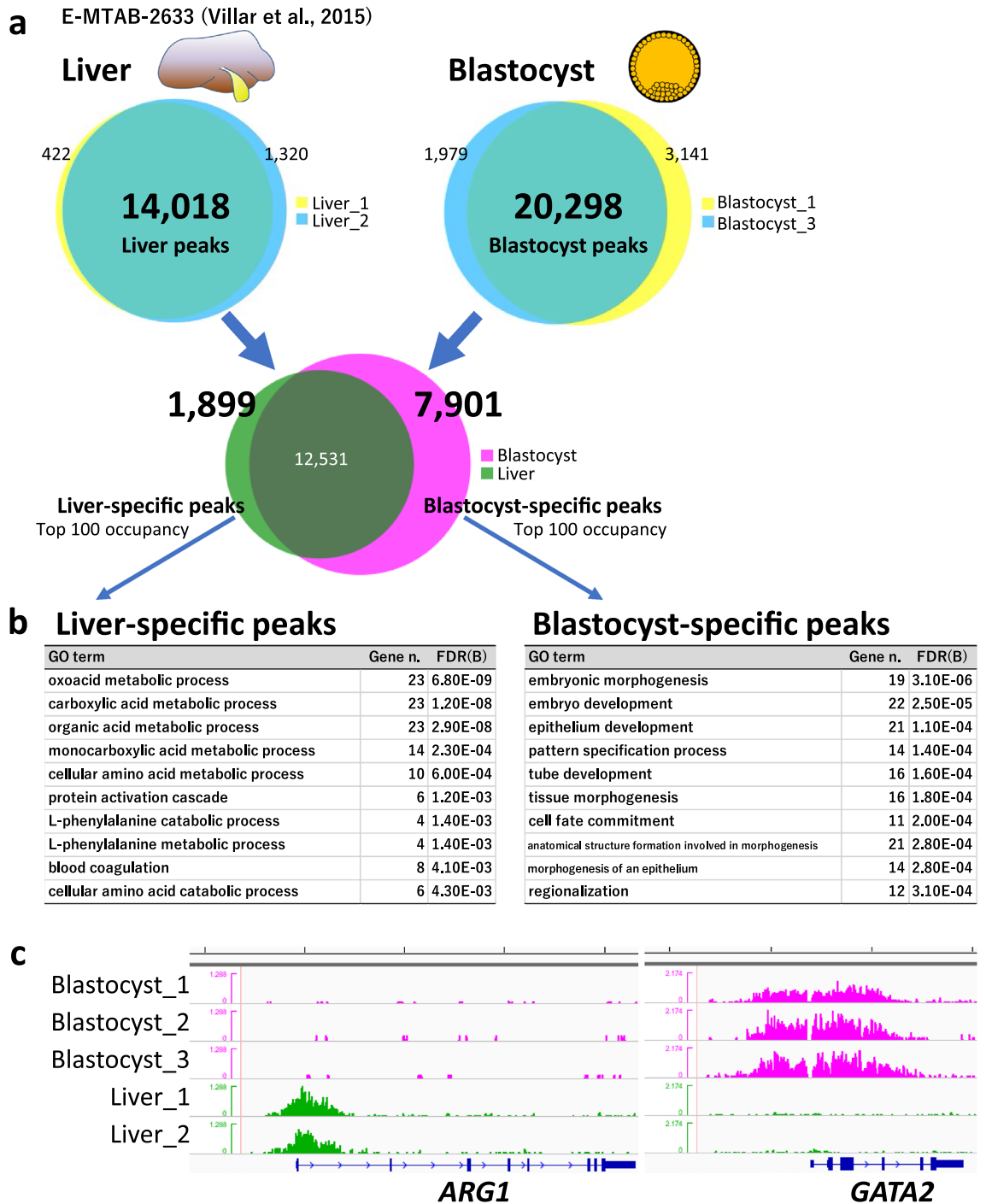
Figure 2. Characterization of blastocyst- and liver-specific H3K4me3 modifications. (**a**) Processing of ChIP-peaks in liver and blastocyst samples. The numbers show those of the peaks. The sum of common (intersect) and specific peak numbers is not equal to the original peak number in each sample because some peaks were separated into multiple peaks to represent intersects. (**b**) Top 10 significant GO terms for biological process enriched by the genes with the top 100 highest peak occupancy rates. Gene n. represents the numbers of related genes. FDR(B) indicates the Benjamin false discovery rate. (**c**) Examples of the H3K4me3 landscape of liver- (*ARG1*) and blastocyst-specific (*GATA2*) peaks around their TSSs. Five-kb graduations are shown in the top scale.

for liver). The genes largely contributing to PC2 exhibited relatively large intra-tissue deviations; for example, *PPP3CA* exhibited -0.21 ± 1.01 in blastocysts and 0.32 ± 1.25 in liver (Fig. 4c).

We then investigated H3K4me3 peaks that are common in blastocysts and liver in terms of imprinted genes given that these genes are profoundly involved in the etiology of fetal developmental disorders[8] and genomic imprinting is also closely associated with the meat production traits of livestock[22,23]. The TSSs of 105 known
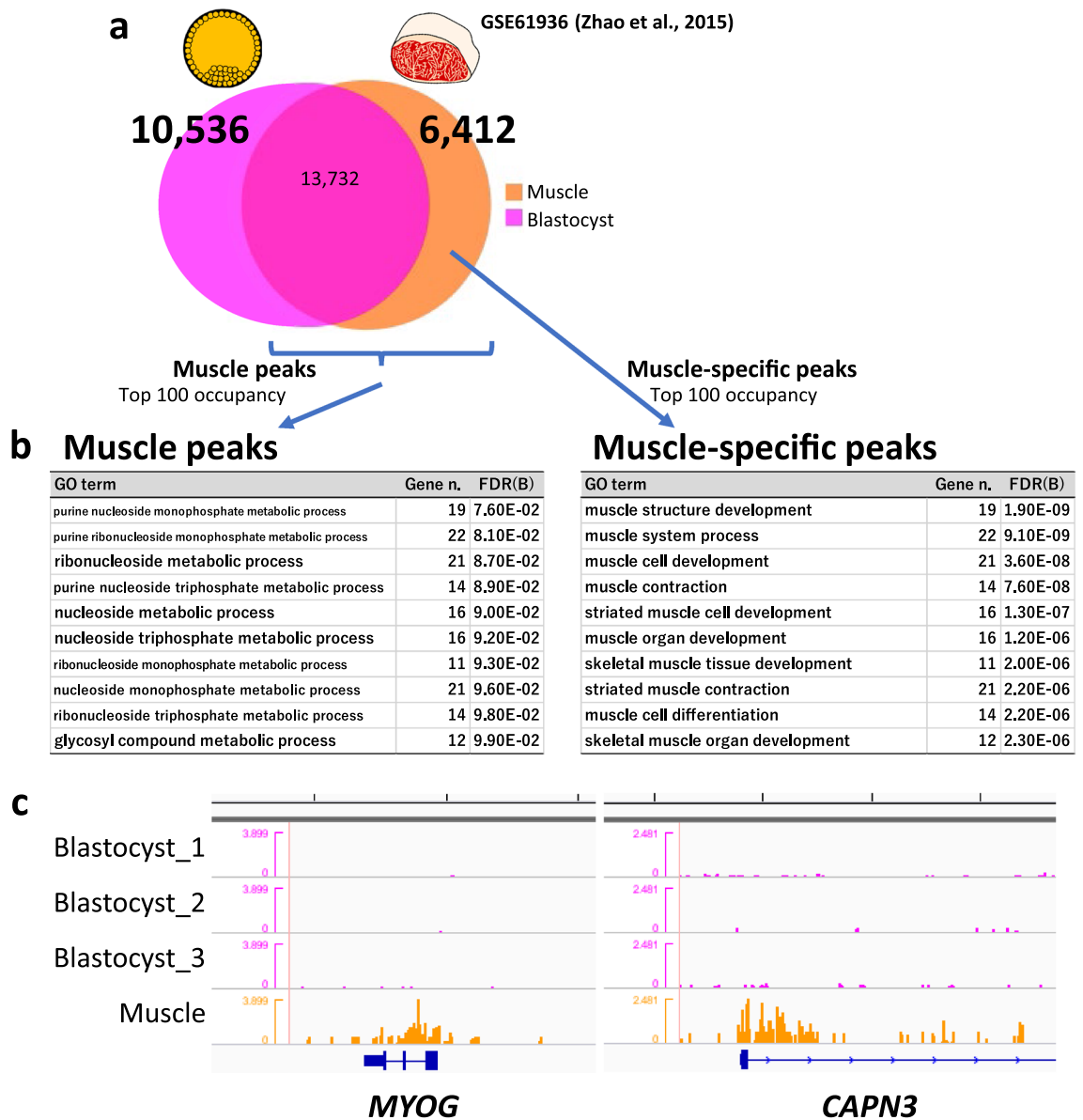
Figure 3. Characterization of blastocyst- and muscle-specific H3K4me3 modifications. (**a**) Processing of ChIP-peaks in muscle and blastocyst samples. The numbers show those of the peaks. The sum of common (intersect) and specific peak numbers is not identical to the original peak number in each sample because some peaks were separated into multiple peaks to represent intersects. (**b**) Top 10 significant GO terms for biological process enriched by the genes with the top 100 highest peak occupancy rates. Gene n. represents the numbers of related genes. FDR(B) indicates the Benjamin false discovery rate. (**c**) Examples of the H3K4me3 landscape for muscle-specific peaks (*MYOG* and *CAPN3*) around their TSSs. Five-kb graduations are shown in the top scale.

imprinted genes listed in a previous bovine study[8] and two differentially methylated regions (DMRs) related to bovine fetal overgrowth (KvDMR1 of *KCNQ1OT1* and *IGF2R/AIRN* DMR) were selected for investigation, and we found that 30 TSSs and the two DMRs had H3K4me3 peaks both in blastocysts and liver. The average profile plot of the 30 TSSs again exhibited overall similar profiles between the two different tissues (Fig. 5a). Meanwhile, we could also extract tissue-specific H3K4me3 modifications, as shown in Fig. S3. PCA of the common peaks in the two tissues extracted the inter- (PC1) and intra-tissue (PC2) differences represented by differential patterns of H3K4me3 modifications among the tissues and samples (Fig. 5b,c). The imprinted genes *IGF2R* (with *IGF2R/AIRN* DMR) and *KCNQ1OT1* (with KvDMR1), both of which are implicated in fetal overgrowth syndrome[8], exhibited a relatively high contribution to intra-tissue differences (PC2). In addition, we further investigated the correlation between the levels of H3K4me3 around TSSs and the expression levels of these imprinted genes calculated from publicly available RNA-seq data. Despite the independent experiments, the levels of H3K4me3 in both tissues exhibited significant correlations with the gene expression levels (Fig. 6a). Furthermore, the relative difference between blastocyst and liver in terms of H3K4me3 modifications around the TSSs of the imprinted genes well reflected the relative difference in gene expression levels between the two tissues (Fig. 6b).
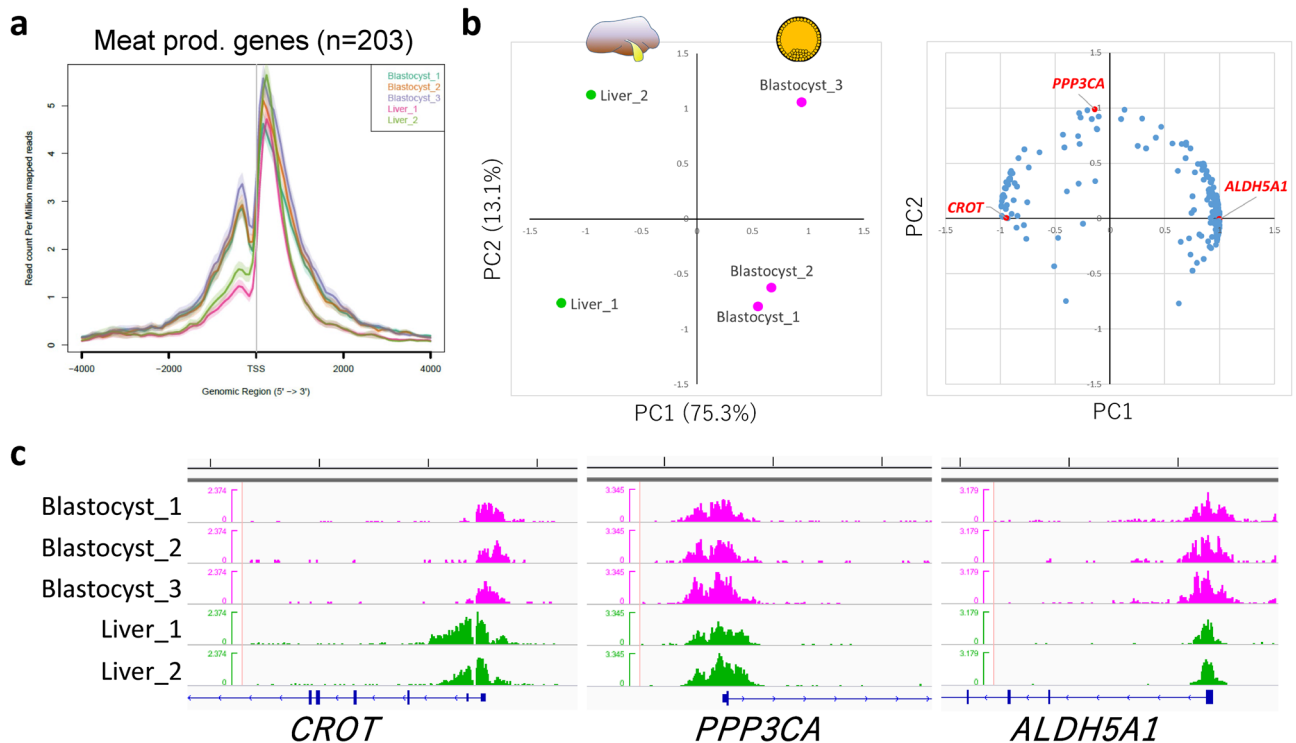
Figure 4. Characterization of H3K4me3 modifications common in blastocysts and liver in terms of meat production-related genes. (**a**) Average profile plot of the H3K4me3 signal around the TSSs of 203 meat production-related genes. (**b**) PCA of liver and blastocyst samples considering the peak areas around the TSSs of 203 meat production-related genes. Left and right panels show the principal component plot of all samples and loading plot of the 203 genes, respectively. The genes whose H3K4me3 landscapes are shown in (**c**) are highlighted in red. (**c**) H3K4me3 landscapes of *CROT*, *PPP3CA*, and *ALDH5A1* around their TSSs. Five-kb graduations are shown in the top scale.

Discussion

To our knowledge, only one study has assessed genome-wide histone modifications of bovine early embryos[24]. That report was a pioneering study of the issue; however, due to its emphasis on an unconventional ChIP methodology to reduce sample cell numbers, H3K4me3 landscape that was produced was compromised by a lack of valley-like patterns around TSSs that correspond to nucleosome-free regions (Fig. S4)[15]. The present ChIP-seq analysis of H3K4me3 successfully generated the typical landscape of epigenetic modifications around TSSs (Fig. 1d and Fig. S4). In addition, the clear enrichment of H3K4me3 modifications at TSSs and their correlation with the expression levels of the corresponding genes (Fig. 1) support the accuracy of our results. Therefore, the H3K4me3 profile produced in this study might be useful as a new reference for bovine blastocysts.

We compared this high-resolution H3K4me3 profile of bovine blastocysts with those of somatic tissues (liver and muscle) to elucidate common and typical epigenetic modifications between early embryos and differentiated somatic tissues. Interestingly, the H3K4me3 peaks of blastocysts could be used like a "sieve" to extract somatic tissue-specific peaks in genes closely related to tissue-specific functions. Although the genes with the top 100 highest peak occupancy rates in blastocysts, liver, and muscle did not represent the specific functions of each tissue, their counterparts after subtracting genes with common peaks (i.e., "sieving") clearly did (Figs. 2b and 3b and Table S2). These results suggest that high-throughput histone methylome data from early embryos are useful for sieving the methylome of other somatic tissues to characterize them in terms of the tissue-specific modifications that are related to their functions. In other words, the major peaks in a given tissue (not tissue-specific and redundant with other tissues in many cases) are supposed to be related to general biological functions e.g., housekeeping functions, and hence, these peaks are not likely to represent the tissue-specificity. In addition, as a note, one of the limitations of the present comparative analysis is that all three ChIP-seq experiments (our present and two published studies) used different antibodies for detecting H3K4me3. Further analyses with using identical antibody condition may enhance the accuracy of the inter-experiment comparisons.

The common epigenetic modifications between early embryos and somatic tissues are important from the viewpoint of DOHaD, because they represent candidate modifications responsible for the epigenetic persistence-derived long-term consequences of early life conditions. In this study, the common peaks between blastocysts and liver were investigated with a focus on meat production-related genes. PCA of peak areas at TSSs well characterized inter-tissue differences (Fig. 4b,c); in other words, the genes highly contributing to PC1 represented H3K4me3 modifications that are largely different between blastocysts and liver, even though they are "common" modifications (e.g., *ALDH5A1* and *CROT*). On the other hand, PC2 might mirror intra-tissue differences rather than inter-tissue ones, for example, such as *PPP3CA* (Fig. 4c).
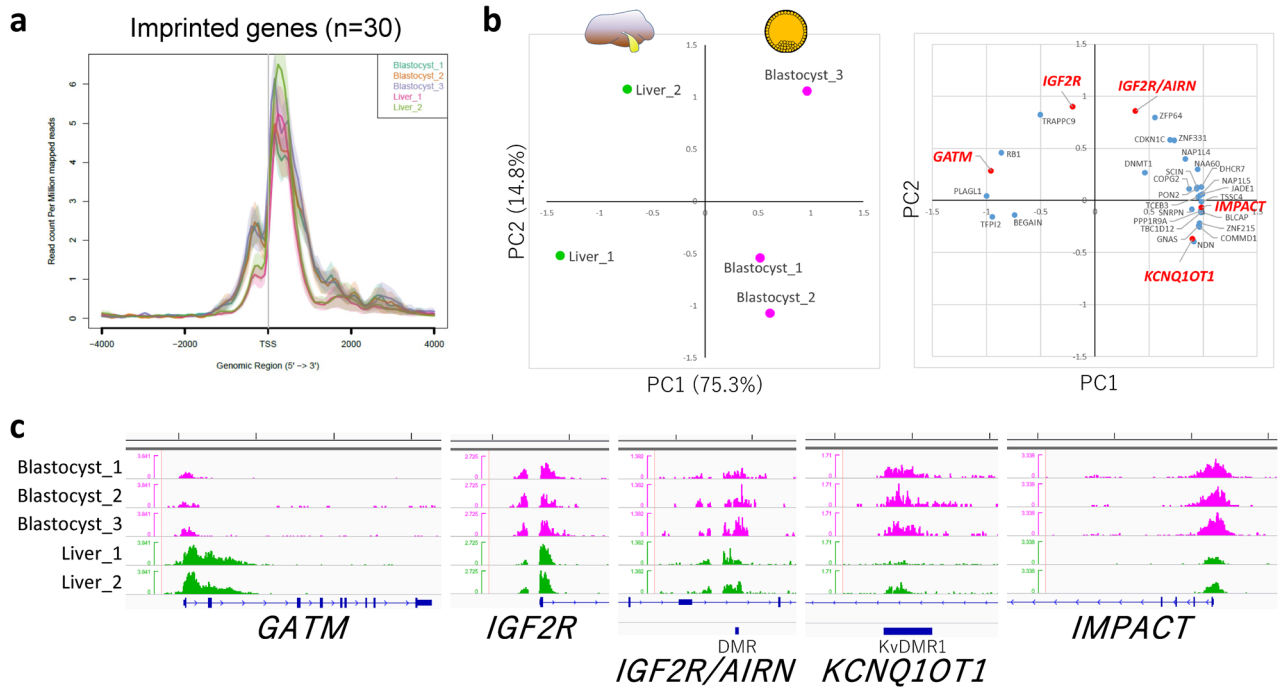
Figure 5. Characterization of H3K4me3 modifications common in blastocysts and liver in terms of imprinted genes. (**a**) Average profile plot of the H3K4me3 signal around the TSSs of 30 imprinted genes. (**b**) PCA of liver and blastocyst samples considering the peak areas around the 30 TSSs of the imprinted genes and two additional DMRs. Left and right panels show the principal component plot of all samples and loading plot of the loci examined, respectively. The genes whose H3K4me3 landscapes are shown in (**c**) are highlighted in red. (**c**) H3K4me3 landscapes of *GATM*, *IGF2R*, *IGF2R/AIRN* (DMR), *KCNQ1OT1* (DMR), and *IMPACT*. Five-kb graduations are shown in the top scale.
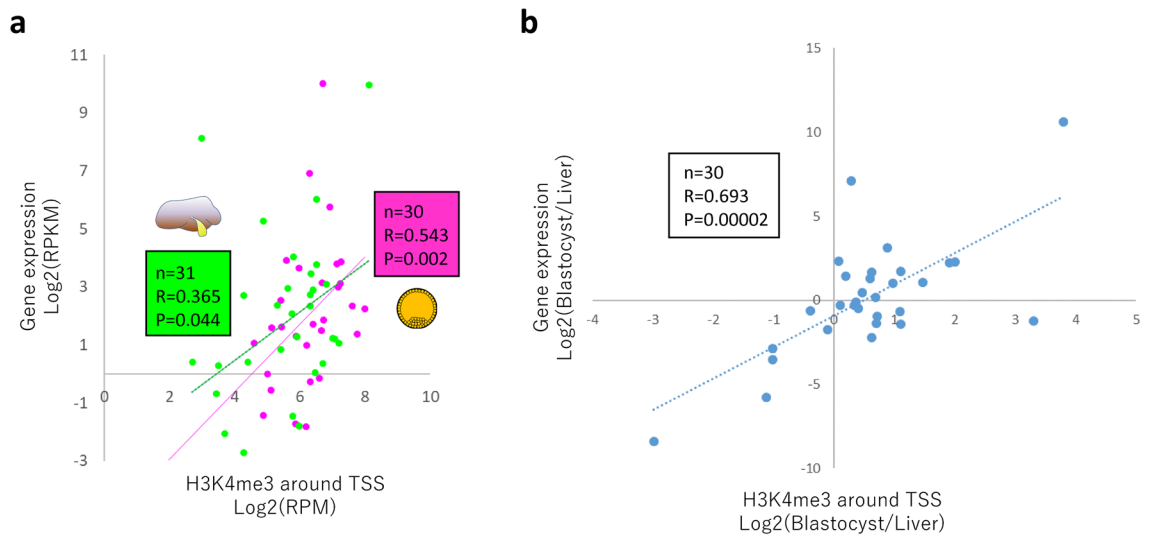

Figure 6. Pearson correlation values between the H3K4me3 signal around the TSSs of imprinted genes and their expression levels. (**a**) Scatter plots of the values for 31 (liver, green) and 30 (blastocyst, magenta) imprinted genes with the linear regression lines are shown. (**b**) Scatter plots of the values for 30 imprinted genes comparing the relative levels between blastocyst and liver. The gene expression levels in liver and blastocyst are based on GSE158430 and GSE52415[37], respectively. *BEGAIN* that showed RPKM = 0 for the expression was omitted from 31 blastocyst data due to the impossibility of log transformation.

In addition to meat production-related genes, we also focused on imprinted genes. Genomic imprinting is an epigenetic phenomenon that compels a subset of genes to be monoallelically expressed in a parent-of-origin-dependent manner in mammals[25]. Appropriate monoallelic expression of imprinted genes is crucial for normal fetal growth, and environmental perturbation during early embryonic development, including the use of assisted reproductive technology, can induce a loss of imprinting of these genes, leading to abnormal fetal overgrowth syndrome[8]. Furthermore, genomic imprinting effects have been widely documented in the economic traits of livestock animals. For example, in beef cattle, several reports have described large relative imprinting variance (i.e., the proportion of total genetic variance attributable to imprinted genes) such as for fat score (24.8% for German Simmental bulls[23]) and beef marbling score (35.2% for Japanese Black bulls[22]). Although DNA methylation has been characterized as a major code of genomic imprinting, recent studies reported DNA methylation-independent imprinting that is controlled by histone methylation[26]. In addition, the global dysregulation of imprinted genes in assisted reproductive technology-induced fetal overgrowth is often independent of DNA methylome epimutations[8,27]. These findings suggest the importance of histone methylation in the expression of imprinted genes. In the present study, we categorized H3K4me3 modifications on imprinted genes into tissue-specific (Fig. S3), common but relatively tissue-dependent (Fig. 5b, PC1 contributions), and common (Fig. 5b, PC2 contributions) modifications. Interestingly, the imprinted genes *IGF2R* and *KCNQ1OT1*, which are well documented in fetal overgrowth syndrome[8], exhibited a relatively high contribution to intra-tissue differences (PC2), suggesting that these H3K4me3 modifications are common in early embryos and hepatic tissues and diverse among individuals and/or given conditions. Interestingly, within the limits of comparing among independent (unrelated) experiments, we could observe the significant correlations between the H3K4me3 levels of these imprinted genes and their expression levels in both blastocysts and liver (Fig. 6). Therefore, the H3K4me3 modifications might be involved in the regulation of these genes alongside the regulation by canonical DNA methylation-dependent imprinting. The appropriate controlling of these histone modifications in developmentally, and hence, economically important genes in the early life period might result in phenotypic changes to improve the welfare and production traits of farm animals. One of the limitations of the present study is comparing unsexed cohort early embryos with adult tissues only from one sex (male) owing to unavailability of public female bovine data. For example, in relation to genomic imprinting, we could clearly recognize the difference in H3K4me3 signals on *XIST* TSS region on chromosome X between male somatic tissues and unsexed blastocysts (Fig. S5a). These differences may be due to the sex difference not due to the tissue specificity because the same differences can be observed between male and female human somatic tissues (Fig. S5b). The differential histone modifications between male and female embryos and bovine somatic tissues are left for future research.

As is the case of imprinted genes, in vitro manipulation of early embryos can cause epigenetic anomalies[13,28]. The embryos used in the present study were in vitro produced and it is unknown which reproductive method was used to obtain the somatic tissues studied here. It is worth investigating the alterations in histone modifications caused by the conditions in which embryos are produced and tracking their persistency or alterations in adult somatic tissues of the same origin. Notably, preimplantation development harbors dynamic epigenetic reprogramming including histone methylations[3,11,29]. Although it is elusive which histone alteration under the establishment of embryonic epigenome can be persistent and affect phenotypes in later life, many reports have addressed the environment-induced histone modifications and their possible long term effects[13,30,31]. Collectively, the effects of environmental conditions and the development-dependent changes on epigenetic modifications and the feasibility of controlling them remain targets for future research.

In conclusion, the present study produced a referential H3K4me3 landscape of bovine blastocysts and revealed its common and typical features compared with adult tissues.

Methods

In vitro production of bovine embryos.   This study was approved by the Animal Research Committee of Kyoto University (Permit Number 31-10) and was carried out in accordance with the Regulation on Animal Experimentation at Kyoto University. The bovine ovaries used in the study were purchased from a commercial abattoir as by-products of meat processing, and the frozen bull semen used for in vitro fertilization (IVF) was also commercially available. In vitro production of bovine embryos by IVF was performed as previously described[32] except for that 50-μL drops of culture medium were used in in vitro culture after IVF. Blastocyst-stage embryos at 192 h post IVF were collected as approximately 11 embryos per replicate for ChIP. The blastocyst samples consisted of the mixture of early, expanded, hatching, and hatched blastocysts.

ChIP.   ChIP for small cell numbers was performed with a Low Cell ChIP-Seq Kit (Active Motif) according to the manufacturer's manual (version A3) with some modifications. The blastocysts were freed from the zona pellucida by using pronase before crosslinking with formaldehyde. After crosslink-quenching, the sample was sonicated to shear chromatin using a Bioruptor UCD-250 (Cosmo Bio) for 30×30 s with 30-s pauses in ice-water. The sample was centrifuged for 2 min at $18,000 \times g$ and the supernatant (200 μL) was transferred to a new tube. The 200-μL sample of sheared chromatin was divided into a 10-μL aliquot as "input" and the rest (190 μL). The latter aliquot was processed for ChIP using 3 μg anti-H3K4me3 antibody (pAb-003-050, Diagenode) as indicated in the user manual. The input and ChIPed DNA was decrosslinked, purified, and resuspended in 40-μL low-EDTA TE buffer. The DNA samples were processed for library preparation for next-generation sequencing by using a Next Gen DNA Library Kit (Active Motif) following the manufacturer's manual. The specific enrichment of H3K4me3 in the ChIP-seq libraries was validated by quantitative PCR for positive (1st exon–intron boundary of *GAPDH*) and negative (2nd exon of *MB*) regions.

Sequencing and data processing. Sequencing was performed on a HiSeq2500 (Illumina) as sigle-end 51-base reads. The sequencing reads were quality checked and aligned to the bovine genome (Bos_taurus_ UMD_3.1.1/bosTau8, June 2014) except for scaffolds using Bowtie[33]. The mapping duplicates were removed by Picard (http://broadinstitute.github.io/picard/). The peaks were called in the ChIP samples relative to the respective input samples using MACS[34]. The annotation of called peaks to genomic regions was generated using CEAS[35] and the peak occupancy rates were in its output. Average H3K4me3 enrichment profiles and heat maps were generated by ngs.plot[21], and peak areas were calculated from its output. Gene ontology analysis was performed using the DAVID tool[17,18]. ChIP-peaks were visualized using the Integrative Genomics Viewer (IGV)[36]. The publicly available raw data for bovine liver and muscle were processed as described above except for the lack of input sample in the muscle sample. The common and specific peaks between samples were identified using bedtools (https://bedtools.readthedocs.io/en/latest/) with the default and −v option, respectively. PCA with autoscaling of H3K4me3 peak areas and Pearson correlation analysis with average H3K4me3 peak areas and merged RNA-seq data were performed using SPSS (SPSS Inc.).

Publicly available data. The following raw data from publicly available databases were used: ChIP-seq of bovine liver, Bull4 and Bull5 of E-MTAB-2633[16]; ChIP-seq of bovine muscle (longissimus dorsi), GSM1517452 of GSE61936[19]; RNA-seq of bovine blastocysts, GSM1265773, GSM1265774, and GSM1265775 of GSE52415[37]; and RNA-seq of bovine liver, GSM4799566 and GSM4799567 of GSE158430 (one of the split files of pair-end reads were used). For RNA-seq data, the three (blastocyst) and two (liver) datasets were merged and expression levels in RPKM values were calculated as previously described[38]. The genes were evenly divided into three categories as high, medium, and low expression levels according to the calculated RPKM values.

Data availability
The ChIP-seq datasets for bovine blastocysts have been deposited in the Gene Expression Omnibus of NCBI with accession number GSE161221.

References
1. Sun, C., Velazquez, M. A. & Fleming, T. P. DOHaD and the periconceptional period, a critical window in time, in *The Epigenome and Developmental Origins of Health and Disease* (ed C. S. Rosenfeld) 33–47, https://doi.org/10.1016/C2013-0-23131-7 (Academic Press, 2015).
2. Wang, L. *et al.* Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**, 979–991. https://doi.org/10.1016/j.cell.2014.04.017 (2014).
3. Liu, X. *et al.* Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**, 558–562. https://doi.org/10.1038/nature19362 (2016).
4. Wang, C. *et al.* Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat. Cell Biol.* **20**, 620–631. https://doi.org/10.1038/s41556-018-0093-4 (2018).
5. Sinclair, K. D. *et al.* Epigenetics and developmental programming of welfare and production traits in farm animals. *Reprod. Fertil. Dev.* https://doi.org/10.1071/RD16102 (2016).
6. Gonzalez-Bulnes, A. *et al.* Developmental Origins of Health and Disease in swine: Implications for animal production and biomedical research. *Theriogenology* **86**, 110–119. https://doi.org/10.1016/j.theriogenology.2016.03.024 (2016).
7. Chavatte-Palmer, P. *et al.* The developmental origins of health and disease: Importance for animal production. *Anim. Reprod.* **12**, 505–520 (2015).
8. Chen, Z. *et al.* Characterization of global loss of imprinting in fetal overgrowth syndrome induced by assisted reproduction. *Proc. Natl. Acad. Sci. USA* **112**, 4618–4623. https://doi.org/10.1073/pnas.1422088112 (2015).
9. Young, L. E., Sinclair, K. D. & Wilmut, I. Large offspring syndrome in cattle and sheep. *Rev. Reprod.* **3**, 155–163. https://doi.org/10.1530/ror.0.0030155 (1998).
10. Huang, X. *et al.* Stable H3K4me3 is associated with transcription initiation during early embryo development. *Bioinformatics* **35**, 3931–3936. https://doi.org/10.1093/bioinformatics/btz173 (2019).
11. Xia, W. *et al.* Resetting histone modifications during human parental-to-zygotic transition. *Science* **365**, 353–360. https://doi.org/10.1126/science.aaw5118 (2019).
12. Kudo, M., Ikeda, S., Sugimoto, M. & Kume, S. Methionine-dependent histone methylation at developmentally important gene loci in mouse preimplantation embryos. *J. Nutr. Biochem.* **26**, 1664–1669. https://doi.org/10.1016/j.jnutbio.2015.08.009 (2015).
13. Feuer, S. K. *et al.* Use of a mouse in vitro fertilization model to understand the developmental origins of health and disease hypothesis. *Endocrinology* **155**, 1956–1969. https://doi.org/10.1210/en.2013-2081 (2014).
14. Herrmann, D., Dahl, J. A., Lucas-Hahn, A., Collas, P. & Niemann, H. Histone modifications and mRNA expression in the inner cell mass and trophectoderm of bovine blastocysts. *Epigenetics* **8**, 281–289. https://doi.org/10.4161/epi.23899 (2013).
15. Jiang, C. & Pugh, B. F. Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.* **10**, 161–172. https://doi.org/10.1038/nrg2522 (2009).
16. Villar, D. *et al.* Enhancer evolution across 20 mammalian species. *Cell* **160**, 554–566. https://doi.org/10.1016/j.cell.2015.01.006 (2015).
17. da Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57. https://doi.org/10.1038/nprot.2008.211 (2009).
18. da Huang, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13. https://doi.org/10.1093/nar/gkn923 (2009).
19. Zhao, C. *et al.* Genome-wide H3K4me3 analysis in angus cattle with divergent tenderness. *PLoS ONE* **10**, e0115358. https://doi.org/10.1371/journal.pone.0115358 (2015).
20. Williams, J. L. *et al.* Discovery, characterization and validation of single nucleotide polymorphisms within 206 bovine genes that may be considered as candidate genes for beef production and quality. *Anim. Genet.* **40**, 486–491. https://doi.org/10.1111/j.1365-2052.2009.01874.x (2009).
21. Shen, L., Shao, N., Liu, X. & Nestler, E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genom.* **15**, 284. https://doi.org/10.1186/1471-2164-15-284 (2014).

22. Okamoto, K. *et al.* Parent-of-origin effects on carcass traits in Japanese Black cattle. *J. Anim. Breed Genet.* **136**, 190–198. https://doi.org/10.1111/jbg.12382 (2019).
23. Neugebauer, N., Rader, I., Schild, H. J., Zimmer, D. & Reinsch, N. Evidence for parent-of-origin effects on genetic variability of beef traits. *J. Anim. Sci.* **88**, 523–532. https://doi.org/10.2527/jas.2009-2026 (2010).
24. Org, T. *et al.* Genome-wide histone modification profiling of inner cell mass and trophectoderm of bovine blastocysts by RAT-ChIP. *PLoS ONE* **14**, e0225801. https://doi.org/10.1371/journal.pone.0225801 (2019).
25. Bartolomei, M. S. & Ferguson-Smith, A. C. Mammalian genomic imprinting. *Cold Spring Harb. Perspect. Biol.* https://doi.org/10.1101/cshperspect.a002592 (2011).
26. Inoue, A., Jiang, L., Lu, F., Suzuki, T. & Zhang, Y. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* **547**, 419–424. https://doi.org/10.1038/nature23262 (2017).
27. Chen, Z., Hagen, D. E., Ji, T., Elsik, C. G. & Rivera, R. M. Global misregulation of genes largely uncoupled to DNA methylome epimutations characterizes a congenital overgrowth syndrome. *Sci. Rep.* **7**, 12667. https://doi.org/10.1038/s41598-017-13012-z (2017).
28. Canovas, S., Ross, P. J., Kelsey, G. & Coy, P. DNA methylation in embryo development: Epigenetic impact of ART (assisted reproductive technologies). *BioEssays* https://doi.org/10.1002/bies.201700106 (2017).
29. Zheng, H. *et al.* Resetting epigenetic memory by reprogramming of histone modifications in mammals. *Mol. Cell* **63**, 1066–1079. https://doi.org/10.1016/j.molcel.2016.08.032 (2016).
30. Dai, Z., Mentch, S. J., Gao, X., Nichenametla, S. N. & Locasale, J. W. Methionine metabolism influences genomic architecture and gene expression through H3K4me3 peak width. *Nat. Commun.* **9**, 1955. https://doi.org/10.1038/s41467-018-04426-y (2018).
31. Campos, E. I., Stafford, J. M. & Reinberg, D. Epigenetic inheritance: Histone bookmarks across generations. *Trends Cell Biol.* **24**, 664–674. https://doi.org/10.1016/j.tcb.2014.08.004 (2014).
32. Ikeda, S., Sugimoto, M. & Kume, S. The RPMI-1640 vitamin mixture promotes bovine blastocyst development in vitro and down-regulates gene expression of TXNIP with epigenetic modification of associated histones. *J. Dev. Orig. Health Dis.* **9**, 87–94. https://doi.org/10.1017/S2040174417000563 (2018).
33. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25. https://doi.org/10.1186/gb-2009-10-3-r25 (2009).
34. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137. https://doi.org/10.1186/gb-2008-9-9-r137 (2008).
35. Shin, H., Liu, T., Manrai, A. K. & Liu, X. S. CEAS: Cis-regulatory element annotation system. *Bioinformatics* **25**, 2605–2606. https://doi.org/10.1093/bioinformatics/btp479 (2009).
36. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26. https://doi.org/10.1038/nbt.1754 (2011).
37. Graf, A. *et al.* Fine mapping of genome activation in bovine embryos by RNA sequencing. *Proc. Natl. Acad. Sci. USA* **111**, 4139–4144. https://doi.org/10.1073/pnas.1321569111 (2014).
38. Ishitani, H. *et al.* Embryonic MTHFR contributes to blastocyst development. *J. Assist. Reprod. Genet.* **37**, 1807–1814. https://doi.org/10.1007/s10815-020-01898-0 (2020).

Acknowledgements

Author contributions

M.I., S.I., and N.M. conceived the experiments and drafted the manuscript. M.I. performed bovine IVF and ChIP-seq library preparation and sequencing. M.I. and S.I. analyzed the ChIP-seq results. N.M. supported the experiments and analyses. All authors discussed the results and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-87683-0.

Correspondence and requests for materials should be addressed to S.I.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.