

**INVOLVEMENT OF ABERRANT CHROMOSOME ARCHITECTURE
AND LOCUS-SPECIFIC VULNERABILITY TO DNA METHYLATION
EPIMUTATIONS IN BOVINE LARGE OFFSPRING SYNDROME**

A Dissertation

Presented to the Faculty of the Graduate School

At the University of Missouri

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

YAHAN LI

Dr. Rocío Melissa Rivera, Dissertation Supervisor

May 2022

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

**INVOLVEMENT OF ABERRANT CHROMOSOME ARCHITECTURE
AND LOCUS-SPECIFIC VULNERABILITY TO DNA METHYLATION
EPIMUTATIONS IN BOVINE LARGE OFFSPRING SYNDROME**

presented by Yahan Li,

a candidate for the degree of Doctor of Philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Rocío Melissa Rivera

Dr. Jianlin Cheng

Dr. Christine Elsik

Dr. Kiho Lee

Dr. Darren Hagen

ACKNOWLEDGEMENTS

There are so many people I would like to thank for their generous help and continuous support during the four years of my PhD studies. First, I am extremely grateful to my mentor Dr. Rocío Rivera for having me in her laboratory at Mizzou. The training I received from Dr. Rivera not only improved my experimental skills but also helped me to become an independent thinker and researcher. Her diligence at work, passion for science, trust, patience, and sense of responsibility to her trainees made her a great mentor. Second, I would like to thank my committee members Dr. Jianlin Cheng, Dr. Christine Elisk, Dr. Kiho Lee, and Dr. Darren Hagen, for their guidance on my research projects and professional development. I would especially like to thank Dr. Cheng and Dr. Hagen who provided valuable advice for the bioinformatic analyses of this work. I also thank Dr. Tieming Ji who participated in my committee prior to embarking on another career opportunity.

Next, I would like to thank current and previous Rivera laboratory members for their participation and hard work in the bovine projects, including Chris Kim, Edgar Joel Soto-Moreno, Astrid Roshealy Brau Rodríguez, Bhaumik Patel, Ali Patten, Monique Ferrell, Amanda Moreno, and Anna Goldkamp (Dr. Hagen's laboratory). I also want to recognize the positive environment for life and

research created by my laboratory members, including above mentioned and Godwin Iroanya, Mohamed Aboul EZZ, Emma Goodwin, Ruthanne Doebler, Camryn Habben, Anna Hagedorn, Marinel Ocasio-Rivera, María Velez-Colón, Arelis Acevedo-Santiago, Faith Korpus, Zhoulin Wu, and Alondra Figueroa.

I would like to thank faculty, staff, and other graduate students from our Division of Animal Sciences for aiding my research through their equipment sharing, feedback, and advice. Especially, I want to acknowledge Dr. Elsik and Dr. Zhiyuan Chen (previous laboratory member) for teaching me Perl coding language during my MS program, and Dr. M. Sofía Ortega for teaching me how to do cell culture. Their help was fundamental to the success of my research. I also appreciate the advice given by Dr. Robert Schnabel for the analyses of genomic sequencing data.

I would like to thank our collaborators for their active involvement in my research projects, including Dr. Pilar Coy Fuster, Dr. Jordana Sena Lopes, Dr. Callum Donnelly, Dr. Fred Williams III, Dr. Jennifer Kalish, Frimpong Boadu (Dr. Cheng's laboratory), and Dr. Max R. Highsmith (Dr. Cheng's laboratory). I also want to thank staff from the University of Missouri Genomics Technology Core for their help with deep sequencing and troubleshooting, including Nathan Bivens, Ming-Yi Zhou, and Karen Bromert.

Special thanks to Mrs. Carol Morrison and her family, the donors of Dr. Roger L. Morrison Scholarship that supplemented my assistantship.

Last but not least, I must thank my parents, my relatives, and friends for their continuous support and who encouraged me to pursue the PhD degree, and who supported me emotionally during the COVID-19 pandemic.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
LIST OF FIGURES.....	vii
LIST OF TABLES.....	x
NOMENCLATURE.....	xi
ABSTRACT.....	xx
Chapter 1: Literature review.....	1
1.1 General introduction.....	1
1.2 Epigenetic regulation of gene expression.....	3
1.2.1 DNA methylation and histone post-translational modifications.....	3
1.2.1.1 DNA methylation.....	3
1.2.1.2 Histone post-translational modifications.....	5
1.2.1.3 Interplay between DNA methylation and histone post-translational modifications.....	7
1.2.2 Chromosome architecture.....	10
1.2.2.1 Technologies used to study chromosomal architecture.....	12
1.2.2.2 Topologically associating domains.....	14
1.2.2.3 Chromosome compartments.....	19
1.2.2.4 Chromosome territories.....	22
1.2.3 Genomic imprinting.....	25
1.2.4 Epigenetic reprogramming.....	28
1.3 Beckwith-Wiedemann Syndrome.....	32
1.3.1 Molecular aberrations in BWS.....	34
1.3.2 Clinical features of BWS and corresponding treatments.....	36
1.4 Large/abnormal offspring syndrome.....	42
1.4.1 Molecular aberrations in LOS.....	43
1.4.2 Clinical features of LOS.....	44
1.5 Concluding remarks.....	45
Rationale for Dissertation.....	47

Chapter 2: Allele-specific aberration of imprinted domain chromosome architecture associates with large offspring syndrome	49
2.1 Abstract.....	49
2.2 Introduction	50
2.3 Results.....	52
2.4 Discussion.....	64
2.5 Limitations of Study	69
2.6 Materials and methods.....	70
2.7 Acknowledgements.....	99
Chapter 3: Spontaneous and ART-induced large offspring syndrome: similarities and differences in DNA methylome	134
3.1 Abstract.....	134
3.2 Introduction	135
3.3 Results.....	138
3.4 Discussion.....	146
3.5 Materials and Methods.....	154
3.6 Acknowledgements.....	162
Chapter 4: General discussion	186
4.1 How does ART induce LOS and BWS?.....	187
4.2 How do DNA methylation changes occur in LOS and BWS?	192
4.3 Conclusion.....	195
Appendix 1: Ongoing Hi-C sequencing project.....	196
Appendix 2: Overgrowth Syndrome.....	198
Appendix 3: Collaborative publications.....	216
Bibliography.....	226
Vita	268

LIST OF FIGURES

Chapter 2: Allele-specific aberration of imprinted domain chromosome architecture associates with large offspring syndrome

Figure 1. Differentially methylated regions identified within *IGF2R* ICR and KvDMR1 in LOS when compared with controls. 101

Figure 2. Validation of CTCF binding by chromatin immunoprecipitation (ChIP).
..... 102

Figure 3. 4C identified allele-specific cis and trans contacts with *IGF2R* ICR. . 104

Figure 4. 4C identified allele-specific cis and trans contacts with KvDMR1. 106

Figure 5. Distribution of altered *IGF2R* ICR far-cis and trans contact across various genomic contexts. 108

Figure 6. Location based clustering tendency of differentially expressed genes indicates global alteration of chromosome architecture in LOS. 110

Figure S1. Maternal allele CpG DNA methylation status at KvDMR1, related to Figure 1. 111

Figure S2. 4C identified overall (not allele specific) cis and trans contacts, related to Figure 3 and Figure 4. 113

Figure S3. 4Cker statistical results for overall (not allele specific) contacts, related to Figure 3 and Figure 4. 114

Figure S4. Significant contacts of the bait (*IGF2R* ICR) detected by fourSig, related to Figure 3. 115

Figure S5. 4C identified overall (not allele specific) cis contacts for individual samples, related to Figure 3 and Figure 4. 117

Figure S6. 4C sequencing reads allelic alignment, related to Figure 3 and Figure 4. 118

Figure S7. 4C identified allele-specific cis and trans contacts with *IGF2R* ICR, related to Figure 3. 120

Figure S8. 4C identified paternal allele cis contacts for individual samples, related to Figure 3 and Figure 4. 122

Figure S9. 4C identified maternal allele cis contacts for individual samples, related to Figure 3 and Figure 4. 124

Figure S10. 4Cker statistical results for allele-specific cis and far-cis contacts with *IGF2R* ICR, related to Figure 3. 125

Figure S11. 4Cker statistical results for allele-specific cis and far-cis contacts with <i>IGF2R</i> ICR, related to Figure 3.....	126
Figure S12. 4Cker statistical results for allele-specific cis and far-cis contacts with KvDMR1, related to Figure 4.	127
Figure S13. Distribution of altered <i>IGF2R</i> ICR trans contact across various genomic contexts, related to Figure 5.....	129
Figure S14. Distribution of altered <i>IGF2R</i> ICR trans contact across various genomic contexts, related to Figure 5.....	130
Figure S15. Distribution of altered <i>IGF2R</i> ICR trans contact across various genomic contexts, related to Figure 5.....	132
Figure S16. Distribution of permutation test for clustering tendency of differentially expressed genes (DEGs), related to Figure 6.	133

Chapter 3: Spontaneous and ART-induced large offspring syndrome: similarities and differences in DNA methylome

Figure 1. Example of phenotypic abnormalities of SLOS and ART-LOS calves.	165
Figure 2. Distribution of LOS associated differentially methylated regions (DMRs) across various genomic contexts.....	167
Figure 3. Example of LOS-associated vulnerable loci.	168
Figure 4. LOS-vulnerable loci around promoter regions.....	169
Figure 5. LOS-vulnerable loci overlapping CpG islands in body of imprinted genes.....	170
Figure 6. DNA methylation at the 25 highly vulnerable loci of LOS (≥ 3 experiments) in US_SLOS_#6 calf, its sire, dam, and full-sibling.....	172
Figure 7. Distribution of ART associated differentially methylated regions (DMRs) across various genomic contents.	174
Figure 8. Conservation of DNA methylation at LOS-vulnerable loci between muscle and blood in ES_RF group.	175
Figure S1. Distribution of LOS associated differentially methylated regions (DMRs) across various genomic contexts.	176
Figure S2. Distribution of LOS associated differentially methylated regions (DMRs) across various genomic contexts.	178
Figure S3. LOS- vulnerable loci overlapping CpG islands in gene bodies.....	179

Figure S4. LOS- vulnerable loci in gene bodies.....	180
Figure S5. LOS- vulnerable loci overlapping CpG islands in intergenic regions.	181
Figure S6. DNA methylation at all LOS-vulnerable loci in US_SLOS_#6 calf and its sire, dam, and full-sibling.	182
Figure S7. Distribution of tissue specific differentially methylated regions (DMRs) across various genomic contents.	184
Figure S8. Distribution of ART associated differentially methylated regions (DMRs) across various genomic contents.	185

Appendix 2: Overgrowth Syndrome

Fig. 1. ART-produced LOS.	213
Fig. 2. Spontaneous LOS.	215

LIST OF TABLES

Chapter 3: Spontaneous and ART-induced large offspring syndrome: similarities and differences in DNA methylome	
Table 1. Information of calves in this study.....	164

NOMENCLATURE

3C	Chromosome conformation capture
4C	Circular chromosome conformation capture/chromosome conformation capture on chip
5C	Chromosome conformation capture carbon copy
ADD	ATRX-DNMT3-DNMT3L
AI	Artificial insemination
<i>AIRN</i>	Antisense of IGF2R non-protein coding RNA
AOS	Abnormal offspring syndrome
ART	Assisted reproductive technologies
<i>ATP10D</i>	ATPase phospholipid transporting 10D
<i>B. t. indicus</i>	<i>Bos taurus indicus</i>
<i>B. t. taurus</i>	<i>Bos taurus taurus</i>
BAF	BRM-associated factors
<i>BLCAP</i>	BLCAP apoptosis inducing factor
bp	Base pairs
BWS	Beckwith-Wiedemann syndrome

CBX5/HP1A	Chromobox 5 protein
<i>CDKN1C</i>	Cyclin dependent kinase inhibitor 1C
ChIA-PET	Chromatin interaction analysis with paired-end tag
ChIP	Chromatin immunoprecipitation
<i>CHMP6</i>	Charged multivesicular body protein 6
CP190	Centrosomal protein 190kD
cpm	Counts per million
CTCF	CCCTC-binding factor
CXXC	Cys-X-X-Cys
DEG	Differentially expressed gene
DMR	Differentially methylated region
<i>DMRT2</i>	Doublesex and mab-3 related transcription factor 2
<i>DNMT</i>	DNA methyltransferase
<i>DNMT1</i>	DNA methyltransferase 1
<i>DNMT3A/B</i>	DNA methyltransferase 3 alpha/beta
<i>DNMT3L</i>	DNA methyltransferase 3 like
<i>Dppa3/Pgc7/Stella</i>	Developmental pluripotency associated 3
EAF	ELL-associated factor

ECM	Extracellular matrix
EMD	Emerin
<i>ENTPD1</i>	Ectonucleoside triphosphate diphosphohydrolase 1
<i>EPHA5</i>	EPH receptor A5
ET	Embryo transfer
FDR	False discovery rate
FISH	Fluorescence in situ hybridization
<i>FOS</i>	Fos proto-oncogene, AP-1 transcription factor subunit
GEO	NCBI Gene Expression Omnibus
<i>GNAS</i>	GNAS complex locus
<i>GNAS-AS1/NESPAS</i>	GNAS antisense RNA 1
<i>GRB10</i>	Growth factor receptor bound protein 10
<i>H19</i>	H19 imprinted maternally expressed transcript
H2AW/HIST3H2A	H2A.W histone
H2BU1/HIST3H2BB	H2B.U histone 1
H3-4/LOC518318	H3.4 histone
H3K27	Histone H3 lysine 27

H3K36	Histone 3 on lysine 36
H3K4	Histone H3 lysine 4
H3K9	Histone H3 lysine 9
HIV	Human immunodeficiency virus
hnRNPs	Heterogeneous ribonucleoproteins
IC	Imprinting center
IC1	Imprinting center 1
IC1-GOM	Gain of methylation at IC1
IC2	Imprinting center 2
IC2-LOM	Loss of methylation at IC2
ICR	Imprinting control region
<i>IGF1R</i>	Insulin like growth factor 1 receptor
<i>IGF2</i>	Insulin like growth factor 2
<i>IGF2R</i>	Insulin like growth factor 2 receptor
IVF	<i>In vitro</i> fertilization
kb	Kilobases
<i>KCNQ1</i>	Potassium voltage-gated channel subfamily Q member 1
<i>KCNQ1OT1/LIT1/KvLQT1-AS</i>	<i>KCNQ1</i> opposite strand/antisense transcript 1

<i>KDM1B</i>	Lysine demethylase 1B
KRAB	Krüppel-associated box
LAD	Lamina-associated domain
LINC	Linker of nucleoskeleton and cytoskeleton
LLPS	Liquid-liquid phase separation
lncRNA	Long non-coding RNA
LOS	Large offspring syndrome
LSPS	Liquid–solid phase separation
<i>MAPK1/ERK2</i>	Mitogen-activated protein kinase 1
<i>MAPK3/ERK1</i>	Mitogen-activated protein kinase 3
<i>MAPK8/9</i>	Mitogen-activated protein kinase 8/9
MAPQ	Mapping quality
mb	Megabases
<i>MEST/PEG1</i>	Mesoderm specific transcript
miRNAs	microRNAs
<i>MMP2</i>	Matrix metalloproteinase 2
mRNA	Messenger RNA
MTase	Methyltransferase

MUGTC	Genomics Technology Core, University of Missouri
<i>MYH9/NMIIA</i>	Myosin heavy chain 9
<i>NBL1/DAN</i>	Neuroblastoma suppressor of tumorigenicity 1-like
<i>NNAT</i>	Neuronatin
OGS	Overgrowth syndrome
PCR	Polymerase chain reactions
<i>PEG10</i>	Paternally expressed 10
PGC	Primordial germ cell
PHD	Plant homeodomain
<i>PLAGL1/ZAC</i>	PLAG1 like zinc finger 1
PPPS	Polymer-polymer phase separation
<i>PRDM8</i>	PR/SET domain 8
PWWP	Pro-Trp-Trp-Pro
<i>QKI</i>	QKI, KH domain containing RNA binding
qRT-PCR	Quantitative reverse transcription PCR
<i>RAD21/SCC1</i>	RAD21 cohesin complex component
RE	Restriction enzyme
RFTS/RFD	Replication foci targeting sequence

<i>RGS17</i>	Regulator of G protein signaling 17
RING	Really interesting new gene
SBWS	Spontaneous BWS
SCNT	Somatic cell nuclear transfer
<i>SETD2/SET2/HYPB</i>	SET domain containing 2, histone lysine methyltransferase
SHI	SETD2-hnRNP interaction
SINE	Short interspersed nuclear element
SLOS	Spontaneous LOS
SMCHD1	Structural maintenance of chromosomes flexible hinge domain containing 1
SNP	Single nucleotide polymorphism
<i>SNRPN</i>	Small nuclear ribonucleoprotein polypeptide N
SOF	Synthetic oviductal fluid
SRI	SET2–RPB1 interacting
STAG2/SA2	Stromal antigen 2
Su(Hw)	Suppressor of Hairy wing
SUN	Sad1 and UNC84 domain containing 1
SYNE/nesprin	Spectrin repeat containing nuclear envelope protein

TAD	Topologically associating domain
<i>TBX18</i>	T-box transcription factor 18
TES	Transcription end site
<i>TET</i>	Ten-eleven translocation
<i>TET1/2/3</i>	Tet methylcytosine dioxygenase 1/2/3
<i>TRIM28/KAP1</i>	Tripartite motif containing 28
tRNAs	Transfer RNAs
TSS	Transcription start sites
TTD	Tandem Tudor domain
UBL	Ubiquitin-like
<i>Uhrf1/Np95/Icbp90</i>	Ubiquitin like with PHD and ring finger domains 1
UPD	Uniparental disomy
WGBS	Whole genome bisulfite sequencing
<i>Zfp57</i>	Zinc finger protein 57
<i>Znf445/202</i>	Zinc finger protein 445/202

Rules of gene nomenclature:

Human, ovine, and bovine gene and transcript name – capitalization and italicization of all the letters, like *IGF2R*.

Mouse and rat gene and transcript name – capitalization of the first letter and italicization of all the letters, like *Igf2r*.

Human, ovine, bovine, mouse, and rat protein name – capitalization of all the letters, like IGF2R.

ABSTRACT

Large/abnormal offspring syndrome (LOS/AOS) and Beckwith-Wiedemann syndrome (BWS) are similar congenital overgrowth syndromes which occur naturally in ruminants and humans, respectively. The incidence of these syndromes increases when offspring are conceived with the use of assisted reproductive technologies (ART; i.e. *in vitro* oocyte maturation, *in vitro* fertilization, and embryo culture). Molecular defects reported in both syndromes include global gene misregulation, DNA methylome epimutations, and disruption of genomic imprinting (parental-allele-specific gene expression). Although we have reported that bovine LOS occurs spontaneously (SLOS) based on phenotypic similarities to ART-LOS, to date no study has been conducted to determine if SLOS has the same methylome epimutations as ART-LOS. One goal of my dissertation research is to characterize DNA methylation profiles in bovine SLOS and ART-LOS to determine whether there are conserved genomic loci with DNA methylation defects between these overgrowth conditions.

In addition, while it is known that LOS is characterized by global alterations in DNA methylation, it is largely unknown how altered DNA methylation drives the development of LOS, as the methylation errors (i.e., differentially methylated regions; DMRs) observed in the syndrome only explain <4% of the gene misregulation in short range (the flanking 20,000 DNA bases

from the DMR). Therefore, another goal of my dissertation research is to determine whether long-range regulatory mechanisms of gene expression, such as chromosome architecture, is altered in LOS as a result of aberrant DNA methylation.

In this dissertation, Chapter 1 is the literature review and will introduce epigenetic regulation of gene expression including chromosome architecture and clinical features and molecular findings of LOS and BWS. Chapter 2 and 3 are the research chapters. In Chapter 2, I characterize allele-specific chromosome architecture of *IGF2R* imprinted domain in fibroblast cells derived from control bovine fetuses and identified disrupted chromosome architecture in LOS. I also observed genomic location-based clustering tendency of misregulated genes in LOS. This study has been published in the Journal *iScience* (<https://doi.org/10.1016/j.isci.2022.104269>) (Li et al., 2022). In Chapter 3, I determined that bovine SLOS has DNA methylation defects with some similarities and differences when compared to ART-LOS. I also identified vulnerable genomic loci for DNA methylation defects in LOS, which could serve as molecular markers for the diagnosis of the syndrome during early pregnancy. This study has been published in the journal *Epigenetics* (<https://doi.org/10.1080/15592294.2022.2067938>). Chapter 4 is the general discussion in which my research findings are incorporated into the general knowledge of the field and implications and directions of future studies are

discussed.

In Appendix 1, I briefly introduce our ongoing Hi-C (global chromosome architecture), methylome and transcriptome project in which samples from LOS and BWS will be analyzed together to further shed light into the etiology of these syndromes, knowledge that will equally help Agriculture and Biomedicine. I anticipate submitting this manuscript for peer review and publication in July of 2022, thus becoming the third primary literature manuscript from my dissertation research. Appendix 2 is a review paper in which I am main contributor author published in *Veterinary Clinics of North America: Food Animal Practice* in 2019 (PMID: 31103180, <https://doi.org/10.1016/j.cvfa.2019.02.007>). This review summarized clinical and molecular findings in LOS and for the first time reported the existence of SLOS. Lastly, Appendix 3 summarizes my contributions of five other publications in which I collaborated with groups in the Division, at Mizzou, and other Academic institutions in the USA during my tenure as a PhD student.

Chapter 1: Literature review

1.1 General introduction

The process of embryo development requires fine regulation of gene expression, and epigenetic mechanisms play very important roles in it (Chason et al., 2011; Niakan et al., 2012). Epigenetic regulation is a broad concept including DNA methylation, histone variants and post-translational modifications, genomic imprinting, small/long non-coding RNAs, and chromosome architectures (Inbar-Feigenberg et al., 2013). The disruption of epigenetic regulation during embryo development can lead to fetal mortality or severe disorders in human infants, such as Beckwith-Wiedemann Syndrome (BWS), Silver–Russell syndrome, Angelman syndrome, and Prader–Willi syndrome, and in cattle offspring, such as large/abnormal offspring syndrome (LOS/AOS) (Elhamamsy, 2017; Osborne-Majnik et al., 2013; Rivera et al., 2021).

LOS and BWS are similar congenital generalized overgrowth syndromes, conditions with broad spectrum of symptoms including a frequent observation of excessive growth (i.e., 2-3 standard deviations increase in overall growth parameters including body weight, height, and head circumference) (Elliott et al., 1994; Lapunzina, 2005; Opitz et al., 1998). Genome wide alterations of DNA methylation and expression of imprinted and non-imprinted genes have been

reported in LOS and BWS, however there is a lack of knowledge on the global correlation between these two changes (Chang et al., 2021; Chen et al., 2016a, 2015, 2017; Krzyzewska et al., 2019; Li et al., 2019a). Recent BWS studies have shown clues of regional changes of chromosome architectures at the megabase scale, but it is still unknown whether genome wide alteration of chromosome architecture is associated with LOS and BWS (Naveh et al., 2021; Rovina et al., 2020).

Assisted reproductive technologies (ART) refer to a series of fertility treatments used to produce offspring and include oocyte collection and *in vitro* maturation, *in vitro* fertilization, intracytoplasmic sperm injection, embryo culture, and embryo transfer, but do not include procedures in which only sperm is handled, such as artificial insemination (AI) (Joao Viana, 2021; Sunderam et al., 2020). ART is known to induce errors in the epigenome of offspring in humans and ruminants (Fauser et al., 2014; Urrego et al., 2014).

In humans, BWS occurs spontaneously, and the use of ART is known to increase its incidence (Mussa et al., 2017; Vermeiden and Bernardus, 2013). Bovine LOS has long been known to be induced by ART and we recently documented the occurrence of spontaneous LOS in cattle (SLOS; conceived by natural mating or AI) (Behboodi et al., 1995; Farin et al., 2001; Hasler et al., 1995; Lazzari et al., 2002; Rivera et al., 2021; van Wagtendonk-de Leeuw et al., 2000).

1.2 Epigenetic regulation of gene expression

The epigenetic regulation of gene expression lays the foundation for cell lineage commitment during embryo development and is the basis for studying LOS and BWS. (Hemberger et al., 2009; Phillips-Cremins et al., 2013). This section will introduce different aspects of epigenetic regulation - with focus on DNA methylation, chromosome architecture, and genomic imprinting, since they are the focus of my dissertation work (Inbar-Feigenberg et al., 2013).

1.2.1 DNA methylation and histone post-translational modifications

1.2.1.1 DNA methylation

DNA methylation is the addition of a methyl group to DNA (in most cases at the fifth carbon of cytosine in CpG dinucleotides) and can regulate DNA binding accessibility and gene expression in mammals (Tate and Bird, 1993). The global DNA methylation level is about 70-80% at CpG context and is less than 1% at other contexts (i.e., CHH and CHG, H = A or C or T) in mammals (Chen et al., 2017; Li and Zhang, 2014). Thus, the phrase “DNA methylation” in this dissertation refers to CpG context unless otherwise specified. Relative to the global average level, the pattern of DNA methylation around actively transcribed genes shows a huge reduction at promoter and transcription start sites (TSS),

equal or higher level at gene bodies, and slightly lower around transcription end sites (TES) (Cao et al., 2020; Chen et al., 2017). DNA methylation in non-CpG contexts shows similar enrichment in gene bodies and depletion at promoters and enhancers (Lister et al., 2009). DNA methylation level of gene bodies is correlated with the frequency of transcription in a parabolic pattern such that the most highly and lowly expressed genes have low level of methylation and genes with intermediate level of expression have high methylation level (Jjingo et al., 2012).

The *de novo* establishment and maintenance of DNA methylation relies on the DNA methyltransferases (*Dnmts*) (Bourc'his et al., 2001; Okano et al., 1999; Song et al., 2011). *Dnmt3a* and *Dnmt3b* have methyltransferase activity through their catalytic MTase domain and are responsible for *de novo* methylation (Okano et al., 1999). The genome-wide distributions of *Dnmt3a* and *Dnmt3b* have similar patterns, but *Dnmt3b* is more enriched at bodies of actively transcribed genes (Baubec et al., 2015). *Dnmt3L* has no methyltransferase activity but can bind to *Dnmt3a* and *Dnmt3b* to increase their activity during gametogenesis and plays critical roles in the establishment of both maternal and paternal genomic imprints (Bourc'his et al., 2001; Webster et al., 2005). *Dnmt1* has methyltransferase activity and is responsible for maintaining DNA methylation in the newly synthesized strand during DNA replication (Song et al., 2011). Some studies have also shown that *Dnmt1* can introduce *de novo*

methylation, especially at transposable elements (Fatemi et al., 2002; Haggerty et al., 2021).

The removal of DNA methylation can be either through an active or passive process (Wu and Zhang, 2017). Passive DNA demethylation is caused by a lack of maintenance of DNA methylation on the newly synthesized strand during DNA replication. Active DNA demethylation is accomplished by enzymes from the ten-eleven translocation (TET) family (Tahiliani et al., 2009). TET proteins can iteratively oxidize DNA methylation (i.e., 5-methylcytosine; 5mC) through 5-hydroxymethylcytosine (5hmC) and 5-formylcytosine (5fC) to 5-carboxylcytosine (5caC) which can be converted back to cytosine by other enzymes (He et al., 2011; Ito et al., 2011; Tahiliani et al., 2009). Three members of the TET family have been found, which are tet methylcytosine dioxygenase 1 (*TET1*), *TET2*, and *TET3* (Ito et al., 2011). Although they all have the catalytic activity to oxidize 5mC, TETs have different affinities for different forms of cytosines as their substrates (Hu et al., 2015; Ito et al., 2011). The oxidized forms of 5mC can also be passively lost during DNA replication, as the affinity of Dnmt1 for this modification is lower (Hashimoto et al., 2012; Otani et al., 2013).

1.2.1.2 Histone post-translational modifications

DNA interacts with histone proteins to form nucleosome and

chromatosome (Maeshima et al., 2010). Each nucleosome consists of a linker histone H1, a nucleosome core particle which is an octamer of core histone H2A, H2B, H3, and H4 (two for each), and about 200 base pairs (bp) of DNA (145-147 bp in the nucleosome core particle) (Maeshima et al., 2010). Post-translational modifications of histone regulate chromatin compaction and accessibility and include acetylation, phosphorylation, methylation, ubiquitylation, sumoylation, deamination, propionylation, and butyrylation (Goudarzi et al., 2016; Kebede et al., 2017; Shilatifard, 2012). The majority of modifications occur on the N-terminal regions of histone H3. Acetylation and phosphorylation generally mark transcriptionally active genomic regions (Lawrence et al., 2016).

Methylations of histone H3 lysine 4 (H3K4) mark active promoters and enhancers (Shilatifard, 2012). Particularly, mono-methylation of H3K4 (H3K4me1) is enriched at active and poised enhancers (Heintzman et al., 2009). H3K4me2 and H3K4me3 are enriched at promoters of active genes (Santos-Rosa et al., 2002). Methylations of histone H3 lysine 9 (H3K9) mark transcriptionally repressed regions (Becker et al., 2016). For example, H3K9me2 marks facultative heterochromatin which is developmental stage/cell type specific heterochromatin and H3K9me3 marks constitutive heterochromatin which forms at gene poor regions and is enriched for repetitive sequences. Mono-methylation of histone H3 lysine 27 (H3K27me1) is accumulated in actively transcribed genes (Ferrari et al., 2014). H3K27me2 is the major form of H3K27 modifications and is

widely distributed in the genome to protect H3K27 from unspecific acetylation (Ferrari et al., 2014). H3K27me3 marks facultative heterochromatin and poised enhancers (Jamieson et al., 2016; Zentner et al., 2011).

During active transcription, the SET domain containing 2, histone lysine methyltransferase (*SETD2/SET2/HYPB*) binds to the hyperphosphorylated C-terminal of the large subunit of RNA polymerase II through its SET2-RPB1 interacting (SRI) domain and tri-methylate histone 3 on lysine 36 (H3K36me3) (Rebehmed et al., 2014; Sun et al., 2005). This process can be enhanced by the RNA splicing event through binding of the SETD2-hnRNP interaction (SHI) domain of SETD2 to the heterogeneous ribonucleoproteins (hnRNPs) which are RNA binding proteins during RNA splicing (Bhattacharya et al., 2021; De Almeida et al., 2011).

1.2.1.3 Interplay between DNA methylation and histone post-translational modifications

The cysteine enriched ATRX-DNMT3-DNMT3L (ADD) domain of DNMT3 plays regulatory roles on its activity and specificity through three-dimensional structural changes of the protein (Guo et al., 2015; Ooi et al., 2007; Otani et al., 2009; Zhang et al., 2010). When there is no histone 3 present, ADD domain binds to the catalytic domain (i.e., MTase) and inhibit its activity (Guo et al.,

2015). When H3K4 is unmethylated (me₀), the ADD domain can bind to H3K4me₀ and release the activity of the MTase domain, with higher affinity of DNMT3A than DNMT3L (Ooi et al., 2007; Otani et al., 2009; Zhang et al., 2010). The binding of the ADD domain to H3 N-terminus is mutually exclusive to the binding of HP1 to H3K9me₃, which indicates the process of *de novo* DNA methylation requires loose chromatin structure (Otani et al., 2009). The binding affinity of ADD domain is reduced by methylation of H3K4 and is negatively correlated with the number of methyl groups (Ooi et al., 2007; Otani et al., 2009; Zhang et al., 2010).

The Pro-Trp-Trp-Pro (PWWP) domain of DNMT3A/B can also regulate their functions through interacting with H3K36me₃ (Baubec et al., 2015; Dhayalan et al., 2010; Gong et al., 2020). Binding of PWWP domain to H3K36me₃ increases the activity of DNMT3A/B and maintains their subnuclear localization as enriched spots instead of homogeneous distribution (Dhayalan et al., 2010). Global loss of H3K36me₃ by mutation of *Setd2* leads to loss of the original enrichment at H3K36me₃ (Baubec et al., 2015). In addition, H3K36me₃ of both histone H3 protein in a nucleosome affects the *de novo* methylation efficiency in a cumulative manner in yeast (Gong et al., 2020).

Like DNMT3, the activity of DNMT1 is also self-regulated by its replication foci targeting sequence (RFTS/RFD) and Cys-X-X-Cys (CXXC) domains (Song et al., 2011; Syeda et al., 2011; Takeshita et al., 2011). When

DNMT1 is by itself, the RFTS domain is deeply inserted into the DNA binding pocket of its catalytic domain and inhibits its MTase activity (Syeda et al., 2011; Takeshita et al., 2011). Ubiquitin like with PHD and ring finger domains 1 (*Uhrf1/Np95/Icbp90*), which is a E3 ubiquitin ligase, binds specifically to hemimethylated CpG sites through its SET and RING associated (SRA) domain and recruits DNMT1 to the loci of DNA replication to maintain DNA methylation in the newly synthesized strand (Bostick et al., 2007; Sharif et al., 2007). This recruitment is accomplished between the ubiquitin-like (UBL) domain of UHRF1 and RFTS domain of DNMT1 (Li et al., 2018a). The plant homeodomain (PHD) of UHRF1 binds to unmodified N-terminus of histone H3, including H3R2me0 and H3K4me0, and the tandem Tudor domain (TTD) of UHRF1 can recognize H3K9me2 and H3K9me3, which leads to a enriched nuclear localization of UHRF1 and DNMT1 at heterochromatin (Arita et al., 2012; Nady et al., 2011; Rothbart et al., 2012, 2013). The really interesting new gene (RING) domain of UHRF1 has the activity to ubiquitylate H3K18 and H3K23, which then can be bound by the RFTS domain of DNMT1, resulting in the release of catalytic domain from self-inhibition and increased MTase activity of Dnmt1 (Ishiyama et al., 2017; Nishiyama et al., 2013; Qin et al., 2015). In addition, when both strands of DNA are not methylated, the CXXC domain of DNMT1 specifically binds to unmethylated CpG and insert the linker after CXXC domain into the DNA binding pocket of its catalytic domain to prevent *de novo* DNA methylation (Song et al.,

2011).

Active DNA demethylation is associated with histone modifications through the gene developmental pluripotency associated 3 (*Dppa3/Pgc7/Stella*) (Bian and Yu, 2014; Nakamura et al., 2012). DPPA3 interacts with TET2 and TET3 and can suppress their enzymatic activities (Bian and Yu, 2014). DPPA3 binds specifically to H3K9me2, which allows its regulation to be locus specific (Nakamura et al., 2012). In addition, DPPA3 can interact with UHRF1 to direct its subcellular localization from the nucleus to the cytoplasm and indirectly regulates the maintenance and *de novo* DNA methylation mediated by DNMT1, thus is also involved in passive DNA demethylation (Du et al., 2019; Funaki et al., 2014; Han et al., 2019; Li et al., 2018b; Mulholland et al., 2020).

1.2.2 Chromosome architecture

The genomes of human, cattle, mouse, and rat have similar length which are about 2.6-2.8 billion nucleotides (O'Leary et al., 2016). The linear DNA in the genome is heavily folded and fits in the cell nucleus of ~6 um diameter (Alberts, 2002). This three-dimensional organization process includes wrapping DNA around an octamer of histone proteins to form nucleosomes and chromatosomes, folding into a 30 nm fiber, and ultimately looping and compressing into chromatin (Maeshima et al., 2010). During the interphase of

cell cycles, different chromatin loci interact to form functional structures which are referred to as topologically associating domains (TAD) (Pope et al., 2014). Based on the level of condensation, chromatin can be divided into two types, namely euchromatin and heterochromatin, which are transcriptionally active and silent, respectively. Euchromatin is characterized by low condensation, high accessibility, and the presence of active genes and histone modifications including acetylation, H3K4me3, and H3K36me3 (Lawrence et al., 2016). On the contrary, heterochromatin is characterized by high condensation, low accessibility, and the presence of silent genes and histone modifications including H3K9me3 and H3K27me3 (Lawrence et al., 2016). Euchromatin and heterochromatin are often spatially separated to form different chromosome compartments (Lieberman-Aiden et al., 2009). Further, the non-random positioning of chromosomes within the nucleus, namely chromosome territories, permit specific functional interactions to occur between chromosomes (Cremer et al., 1993).

This section will review higher-order chromatin structures, in other words, chromosome architecture, including TAD, chromosome compartments, and chromosome territories. Chromosome architecture of large scale (i.e., megabases) is generally constant across different cell types within the same species, and many conformational features have been conserved during evolution in mammals (Dixon et al., 2012; Lieberman-Aiden et al., 2009; Rudan

et al., 2015; Tanabe et al., 2002). However, the local chromosome architecture can be dynamic across different cell types or even between different individual cells of the same type, as revealed by single cell studies (Nagano et al., 2013; Stevens et al., 2017).

1.2.2.1 Technologies used to study chromosomal architecture

Technologies used to study chromosomal architecture include fluorescence in situ hybridization (FISH), chromosome conformation capture (3C), circular 3C or 3C on chip (4C), 3C carbon copy (5C), Hi-C, chromatin interaction analysis with paired-end tag (ChIA-PET), and some other variants of these technologies (Xie et al., 2016).

For FISH studies, DNA or RNA probes with fluorescence attached are used to specifically bind to and show target sites, and followed by detection of the fluorescence signal by microscopy or flow cytometry to view the localization of targets (Levsky and Singer, 2003). FISH can be used to detect physical interaction with high specificity, but has limitations on the number of target sites to study in one experiment (Levsky and Singer, 2003).

3C is used to study interactions between two regions with known genomic locations (Dekker et al., 2002). For 3C studies, remote chromosomal interactions through proteins are crosslinked using formaldehyde before cell

lysis, then the genomic DNA is digested with a restriction enzyme to release crosslinked DNA-protein complexes (Dekker et al., 2002). Next, the two fragments of DNA within one DNA-protein-DNA complex are ligated together (referred to as 3C library). After decrosslinking, quantitative polymerase chain reaction (qPCR) is performed with a pair of primers such that each primer binds to one target region to verify hypothesized interactions between the two regions. 3C is the basis for 4C, 5C, ChIA-PET, and Hi-C technologies.

4C is used to study interactions between a region with known genomic location and unknown genomic regions (Simonis et al., 2006). For 4C, the 3C library is digested again with a different restriction enzyme to shorten the length of ligated DNA and increase resolution (Simonis et al., 2006). After the second DNA ligation, circular DNAs are used as templates for PCR with both primers binding to the known region to amplify the unknown regions (4C library). The primers usually contain adaptor sequences and the 4C library can be sequenced by microarray or high-throughput sequencing to determine the unknown regions that interact with the known region (i.e., the bait) (Simonis et al., 2006; Splinter et al., 2012).

5C is used to study interactions between many but not all unknown genomic regions (Dostie et al., 2006). For 5C, the 3C library is incubated with a mixture of thousands of different 5C primers, each containing the upstream or downstream restriction enzyme sequence followed by random sequences and

ended with an adaptor sequence (Dostie et al., 2006). During ligation, only the pair of 5C primers annealing to the 3C library at restriction enzyme cutting sites can be ligated. The ligated 5C primer pairs will be PCR amplified with universal primers that bind to the adaptor sequences (5C library). Then the 5C library can be analyzed by microarray or high-throughput sequencing.

Hi-C is used to study all chromatin interactions (Lieberman-Aiden et al., 2009). For Hi-C, during the 3C library preparation, prior to ligation, biotin labelled DNA oligos are added to fill in the sticky ends generated by restriction enzyme digestion (Lieberman-Aiden et al., 2009). After ligation, biotin from not ligated ends will be removed. Biotin serves to pull down interactive DNA with streptavidin beads which have high affinity for biotin. Then the library is sheared to increase resolution, ligated with universal adaptors, and sequenced using high throughput methodologies. ChIA-PET is similar to Hi-C with an additional step of chromatin immunoprecipitation to enrich chromatin interactions mediated by proteins of interest (Zhang et al., 2012).

1.2.2.2 Topologically associating domains

TADs refer to self-interacting genomic regions within a chromosome and constitutes the primary units of interphase chromosome folding (Dixon et al., 2012; Pope et al., 2014). TADs were initially known as topological domain of

kilobases (kb) scale in late 20th century (Kramer et al., 1999; Sinden and Ussery, 1992). Over 90% of the mammalian genome is folded by TADs and the size of TADs can range from tens of kb to 2 megabases (mb) (Dekker and Heard, 2015; Dixon et al., 2012). TADs exist in multiple layers in which a larger TAD (referred to as metaTAD) may contain many smaller TADs (referred to as sub-TADs) within it, and this structure is referred to as a metaTAD tree (Fraser et al., 2015; Phillips-Cremins et al., 2013).

Architectural proteins, such as CCCTC-binding factor (CTCF) and cohesin protein complex, are the building blocks for TAD formation, and deletion or mutation of these proteins will lead to major loss of TADs (Li et al., 2020b; Nora et al., 2017; Phillips-Cremins et al., 2013; Rao et al., 2017; Zuin et al., 2014). The process of TAD formation by CTCF and cohesin can be illustrated with the extrusion model (Sanborn et al., 2015). In this model, the looping of chromatin is initiated by two cohesin complexes at a genomic locus, and the loop will extend along with the two cohesin complexes moving towards opposite directions of the chromatin. A pair of CTCF proteins bind at their binding sites to define the boundaries of the TAD (Sanborn et al., 2015). The cohesin complexes can be trapped by CTCF through physical binding between the N-terminus of CTCF and two subunits of cohesin complexes, namely RAD21 cohesin complex component (RAD21/SCC1) and stromal antigen 2 (STAG2/SA2) (Li et al., 2020b). In this case, an opposite orientation of the two CTCF binding sites with 3'

side towards the loop of TAD is required (Rao et al., 2014; Sanborn et al., 2015). The maintenance of TADs is a dynamic process as studies have shown that disrupted TADs caused by temporary CTCF degradation can be rapidly restored after CTCF recovery (Nora et al., 2017). Loss of CTCF may lead to gain of DNA methylation at originally unmethylated regions (Dávalos-Salas et al., 2011; Fedorow et al., 2004) and DNA methylation can inhibit CTCF binding to the DNA (Lai et al., 2010; Rodriguez et al., 2010; Wiehle et al., 2019).

In addition to CTCF and cohesin, other architectural proteins have been identified in mammals, including structural maintenance of chromosomes flexible hinge domain containing 1 (SMCHD1) (Wang et al., 2018). SMCHD1 functions specifically during the X chromosome inactivation process, which occurs in female mammals to achieve equal levels of gene expression between males (XY) and females (XX) (Gdula et al., 2019; Wang et al., 2018). SMCHD1 suppresses CTCF and cohesin binding on the inactive X (Xi) chromosome and facilitates the merging of chromosome compartments for complete silencing of gene expression (Gdula et al., 2019; Wang et al., 2018). Other architectural proteins have also been identified in *Drosophila*, including ELL-associated factor (EAF), suppressor of Hairy wing (Su(Hw)), and centrosomal protein 190kD (CP190) (Hou et al., 2012). TADs in *Drosophila* are divided into active and repressive TADs based on their enrichment of epigenetic marks and architectural proteins around the boundaries. The repressive TADs are further subdivided into

Polycomb-associated, HP1/Centromere-associated, and null-associated TADs (Sexton et al., 2012).

Although CTCFs are enriched in majority of the TAD boundaries, CTCF-free boundaries exist in mice and humans (Dixon et al., 2012). The enrichment of other factors such as H3K4me3, H3K36me3, transcription start sites, housekeeping genes, tRNA genes, and repetitive elements such as short interspersed nuclear element (SINE), and exclusion of H3K9me3 have also been identified at the boundaries of TADs (Dixon et al., 2012).

The formation of TADs can either facilitate or block spatial interactions between different chromosomal loci (Doyle et al., 2014). The facilitating effect is achieved by looping-induced reduction of spatial distance between two loci inside or outside of a TAD (Doyle et al., 2014). Indeed, the chromosomal interactions within a TAD were found much more frequent than between two TADs by Hi-C studies (Dixon et al., 2012). High resolution DNA FISH studies also validated these chromosomal interactions detected by 5C and Hi-C studies (Giorgetti et al., 2014; Williamson et al., 2014). The blocking effect of TADs on chromosomal interactions is achieved by physical insulation of the TAD loops from the rest of the genome (Doyle et al., 2014). By manipulating the genome to alter the relative location of a locus to a TAD, many studies have shown this insulation effects of TADs (Andrey et al., 2013; Lupiáñez et al., 2015; Tsujimura et al., 2015).

The abilities of TADs to regulate spatial interactions further lead to their

functions in gene expression regulation. Enhancers and silencers are DNA sequences that can be bound by transcription factors to facilitate or repress expression of other genes, respectively (Maston et al., 2006). The capability of enhancers and silencers to regulate their target genes' expression is decided by their physical availability to the target genes (Della Rosa and Spivakov, 2020; Ong and Corces, 2011). Many studies have shown that TADs are responsible for restricting physical interactions between enhancers/silencers and promoters within the TAD (Anderson et al., 2014; Lupiáñez et al., 2015; Pang and Snyder, 2020; Symmons et al., 2014). Disruption of TADs may lead to aberrant expression of genes by gaining or losing interactions with active enhancers, and can cause severe diseases in human (Lupiáñez et al., 2015).

A genome-wide conservation of TADs has been found among some mammalian species (Rao et al., 2014; Rudan et al., 2015). This conservation of TADs is paralleled with the conservation of CTCF binding sites (Rudan et al., 2015). However, many sub-TADs have been found to be different between species, suggesting evolutionary adaption to changes in the genome (Rudan et al., 2015). In addition, TADs have some level of flexibility, as seen in changes of TADs during cell differentiation and lineage commitment (Fraser et al., 2015; Phillips-Cremins et al., 2013).

1.2.2.3 Chromosome compartments

When considering chromosomes in a linear manner, euchromatin and heterochromatin organize in alternating positions throughout the chromosome (Boyle et al., 2008; Buenrostro et al., 2013). However, Hi-C and DNA FISH studies demonstrate that the spatial organization of each chromosome allows the separation and grouping of euchromatic or heterochromatic regions into chromosome compartments (Boyle et al., 2011; Lieberman-Aiden et al., 2009). As defined by Lieberman-Aiden and others (the first report on this topic), compartment A refers to the chromosome compartment that is enriched in euchromatin, and compartment B refers to the one enriched in heterochromatin (Lieberman-Aiden et al., 2009). Similar to the localization of heterochromatin and euchromatin in the nucleus, compartment B mainly localizes in the peripheral regions and the regions surrounding the nucleoli, while compartment A mainly localizes in the interior of nucleus between two compartment B regions, as revealed by single cell high resolution Hi-C data (Stevens et al., 2017). Corresponding to the chromatin status, unique genetic and epigenetic features are associated with each chromosome compartment (Lieberman-Aiden et al., 2009; Rao et al., 2014). Compartment A is enriched in highly accessible chromatin, highly expressed genes, and activating histone modifications such as H3K36me3, H3K79me2, H3K27ac, and H3K4me1 (Lieberman-Aiden et al., 2009;

Rao et al., 2014). In addition, chromosome compartments are correlated with replication timing of DNA during S phase, in that compartment A is replicated earlier than B (Ryba et al., 2010).

The existence of subcompartments within compartment A and B have been reported in human and mice (Rao et al., 2014; Robson et al., 2017; Yaffe and Tanay, 2011). Subcompartments A1 and A2 belong to compartment A, and their differences are exhibited by replication timing, in which A1 finishes at the beginning of the S phase while A2 finishes replicating at mid S phase (Rao et al., 2014). In addition, A2 contains more repressive histone modification H3K9me3, lower GC content, and longer genes than A1 (Rao et al., 2014). Silent genes being released from B compartments are more likely to become part of the A2 subcompartments rather than A1, as demonstrated to occur during lymphocyte activation (Robson et al., 2017). Subcompartments B1, B2, B3, and B4 belong to compartment B (Rao et al., 2014), the heterochromatin containing compartment. B1 is enriched for H3K27me3 histone modification and replicates during the middle of S phase (Rao et al., 2014). B2, B3 and B4 replicate at the end of S phase. B2 is enriched for pericentromeric heterochromatin, localized in both peripheral and near nucleoli regions, B3 is localized only in the peripheral region, and B4 is enriched for KRAB-ZNF superfamily genes, and contain both activating and repressive histone modifications H3K36me3, H3K9me3, and H4K20me3 (Rao et al., 2014).

The formation of chromosome compartments and subcompartments has been proposed to be the result of separation based on physicochemical properties (i.e. phase separation mechanisms) (Erdel and Rippe, 2018; Hildebrand and Dekker, 2020). These mechanisms include polymer-polymer phase separation (PPPS), liquid-liquid phase separation (LLPS), and liquid–solid phase separation (LSPS) (Erdel and Rippe, 2018; Hildebrand and Dekker, 2020). PPPS can be induced by bridging factors, mainly proteins, binding to different chromosome segments with similar properties and forming a phase that separates from other phases (Erdel and Rippe, 2018; Hildebrand and Dekker, 2020; Michieletto et al., 2016). LLPS is achieved by enrichment of factors with multivalent interactions at chromosome segments with similar properties and forming a liquid-like phase, which will be separated from other phases based on their physical properties, like oil and water (Erdel and Rippe, 2018; Hildebrand and Dekker, 2020; Larson et al., 2017; Strom et al., 2017). The factors function in LLPS are mainly proteins with intrinsically disordered regions, such as the chromobox 5 protein (CBX5/HP1A) (Erdel and Rippe, 2018; Hildebrand and Dekker, 2020; Larson et al., 2017; Strom et al., 2017). LSPS is the spatial separation of a liquid-like phase, as in LLPS, and a solid-like phase, as in PPPS (Hildebrand and Dekker, 2020).

Different genes at discrete chromosomal sites with similar transcription requirements can be recruited to the same nuclear region including the nuclear

matrix (Karki et al., 2018). These stabilized loci, namely transcription factories, contain multiple RNA polymerases and high concentration of transcription factors (Papantonis and Cook, 2013). Transcription factories are associated with concurrent transcription of multiple loci although the existence of transcription factories can be independent of transcriptional activity (Karki et al., 2018; Larkin et al., 2013; Mitchell and Fraser, 2008; Osborne et al., 2004, 2007; Papantonis et al., 2012). Organization of chromosome compartments has been associated with transcription factories (Karbassi et al., 2019; Papantonis and Cook, 2013). It has been proposed that the compartment A is more associated with transcription factories than B, and when genes translocated from compartment B to A, they may gain accessibility to transcription factories (Karbassi et al., 2019; Robson et al., 2017).

1.2.2.4 Chromosome territories

Chromosome territories refer to a phenomenon in which different regions of the nucleus are preferentially occupied by particular chromosomes during interphase (Cremer et al., 1993). This concept was first introduced by Carl Rabl (1885) and Theodor Boveri (1909) based on cytological examination of the cell nucleus of salamander and roundworm, respectively (Cremer and Cremer, 2006a). The idea of chromosome territories was abandoned during 1950s to

1970s due to the failure of identification of chromosome territories by early electron microscope (Cremer and Cremer, 2006b). However, more recently, FISH and 3C based technologies have shown chromosome territories do exist (Bolzer et al., 2005; Boyle et al., 2001; Kalhor et al., 2012; Lieberman-Aiden et al., 2009; Tanabe et al., 2002). The specificities of these territories for particular chromosomes vary between datasets depending on whether single cells or millions of cells were used for the analyses. Even though studies in which millions of cells have been used propose static positioning of chromosomes (average position based on probabilities), single cell studies show that chromosome territories are dynamic in nature and can vary between similar cells (Meaburn and Misteli, 2007; Nagano et al., 2013; Ramani et al., 2017; Stevens et al., 2017).

Chromosome territories aid or restrict interactions between specific chromosomes based on their spatial proximity (Handoko et al., 2011; Kalhor et al., 2012; Lieberman-Aiden et al., 2009). Largely, interchromosomal interactions occur among regions with similar transcription activity, and are more frequently observed for active rather than repressive regions (Belyaeva et al., 2017; Kalhor et al., 2012). Nevertheless, specific interchromosomal interactions with regulatory roles in gene expression have been reported (Lomvardas et al., 2006; Spilianakis et al., 2005). When compared to interactions within a chromosome, interchromosomal interactions were found to be generally much weaker (less

frequently observed) (Handoko et al., 2011). In addition, these interactions are restricted to specific loci on each chromosome, such as CTCF binding sites, and most of these interactions have been shown as random events with frequencies that are low yet above experimental noise (Handoko et al., 2011; Kalhor et al., 2012). When there is no interaction between adjacent chromosomes, a channel-like chromatin-free space can be formed, which is known as interchromatin compartment (Albiez et al., 2006). Interchromatin compartments have been found connected to the nuclear pores, enriched for nuclear speckles, and functioning in RNA processing and transportation (Markaki et al., 2010).

During interphase, chromatin is mostly immobile that its diffusion is usually constrained within a radius of 0.4 μm , which reflects the boundaries of chromosome territories (Abney et al., 1997; Chubb et al., 2002; Gasser, 2002). These relatively stable chromosome territories are anchored by nuclear substructures such as nucleoli or nuclear periphery (Abney et al., 1997; Chubb et al., 2002; Gasser, 2002). However, long range movements (1-5 μm) of genomic loci from nuclear periphery to interior regions during gene activation have been reported, which reflects changes in chromosome compartments and/or territories (Chuang et al., 2006; Dundr et al., 2007).

1.2.3 Genomic imprinting

Genomic imprinting is an epigenetic mechanism that controls the parent allele-specific expression of approximately 150 genes in mammals (Blake et al., 2009; Chen et al., 2016b; Morison et al., 2001; Tian, 2014). Imprinted genes are often clustered in the genome and these clusters are regulated by allele-specific DNA methylation at their imprinting control regions (ICRs) (Reik and Walter, 2001). Two mechanisms have been demonstrated for ICR's regulation, namely non-coding RNA model and insulator model (Lewis and Reik, 2006). In the non-coding RNA model, the ICR, as a promoter, directly controls the expression of a non-coding RNA whose transcription event inhibits the expression of imprinted genes in that domain. In the insulator model, the ICR harbors the binding sites of CTCF proteins which can form TADs to insulate or promote the physical interaction between promoters and enhancers of imprinted genes. Recent studies in mice have revealed several cases of the involvement of TADs in imprinting regulation of domains following the non-coding RNA model, such as *Kcnq1ot1/Cdkn1c* and *Dlk1-Dio3* domain (Battistelli et al., 2014; Llères et al., 2019). In addition, allele-specific deposition of histone post-translational modifications has also been found for ICRs, such as H3K4me3 on the active allele (Xu et al., 2019).

Imprinting center (IC) is used to indicate the ICR of two specific imprinted

domains in the epigenetic (loss-of-imprinting) condition BWS. IC1 regulates the *IGF2/H19* imprinted domain through the insulator model and is DNA methylated on the paternal allele and unmethylated on the maternal allele (Steenman et al., 1994; Weksberg et al., 1993). The CTCF binds on the maternal allele of IC1 and blocks the interaction between insulin like growth factor 2 (*IGF2*) promoter and *H19* imprinted maternally expressed transcript (*H19*) downstream enhancer, which results in paternal expression of protein-coding gene *IGF2* and maternal expression of non-coding gene *H19* (Steenman et al., 1994; Weksberg et al., 1993). Human *IGF2* promotes growth and development and has two developmental stage-specific promoters for prenatal and postnatal expression (de Pagter-Holthuizen et al., 1987). Rodent *Igf2* is highly expressed during embryo development and decreases after birth (DeChiara et al., 1991; Stylianopoulou et al., 1988a, 1988b). Null mutation of paternal *Igf2* gene results in ~40% decreased birth weight without other morphological changes in mice when compared to wild-type littermates (DeChiara et al., 1990). Overexpression of *Igf2* resulted in macrosomia, disproportionate organ overgrowth, and skeletal abnormalities (Sun et al., 1997). However, either *Igf2* deficiency or overexpression do not affect postnatal growth rate in mice (DeChiara et al., 1991; Leighton et al., 1995). In addition, deletion of *H19* with the adjacent 10 kb sequence on the maternal allele, where the IC1 locates (Thorvaldsen et al., 1998), resulted in ~2.2 fold increased *Igf2* expression and ~27% increased fetal

weight without other defects in mice when compared with wild-type littermates (Leighton et al., 1995).

Insulin like growth factor 2 receptor (*IGF2R*) is a maternally-expressed imprinted gene and its protein can bind to IGF2 with high affinity to mediate the degradation of excess IGF2 (Denley et al., 2005). Maternally mutant *Igf2r* resulted in increase of IGF2 and IGF binding proteins, ~25%-30% increase of body weight, proportionately increased organ size, and perinatal death (Lau et al., 1994). Double mutation of *Igf2r* and *H19* with the adjacent 10 kb sequence results in biallelic expression of *Igf2* with much higher transcript level in embryo (~7.6 fold), tissues (~48.6, ~68.4, and ~38.1 fold in kidney, liver, and heart, respectively), and serum (~11.7 fold), and various abnormal phenotypes including macrosomia, omphalocele, visceromegaly, placentomegaly, cleft palate, and skeletal, cardiac, and adrenal defects (Eggenschwiler et al., 1997).

IC2 regulates the *KCNQ1/CDKN1C* imprinted domain through the non-coding RNA model and is DNA methylated on the maternal allele and unmethylated on paternal allele (Horike et al., 2000; Lee et al., 1999; Smilinich et al., 1999). IC2 contains the promoter of paternally expressed non-coding gene *KCNQ1* opposite strand/antisense transcript 1 (*KCNQ1OT1/LIT1/KvLQT1-AS*) and results in maternal expression of protein-coding genes including potassium voltage-gated channel subfamily Q member 1 (*KCNQ1*) and cyclin dependent kinase inhibitor 1C (*CDKN1C*) (Horike et al., 2000; Lee et al., 1999; Smilinich et

al., 1999). *Cdkn1c* is required for normal embryo development as shown in mouse studies that its deletion on the maternal allele leads to various abnormalities including umbilical hernia/omphalocele, cleft palate, short limbs, bone defects, renal dysplasia, adrenal hyperplasia, and gastrointestinal abnormalities (Yan et al., 1997; Zhang et al., 1997). Double mutation of *Cdkn1c* and *H19* with the adjacent 10 kb sequence results in omphalocele, macroglossia, placentomegaly, renal dysplasia, cleft palate, and polydactyly (Caspary et al., 1999). In addition, reduced expression of *Cdkn1c* is shown in either mouse primary embryo fibroblasts treated with IGF2 protein or mice with high serum IGF2 level, which indicates interplay between these two genes (Grandjean et al., 2000).

1.2.4 Epigenetic reprogramming

Epigenetic reprogramming refers to a genome-wide removal and reestablishment of epigenetic marks (i.e., DNA methylation and histone modifications) during development of germ cells and embryos (Messerschmidt et al., 2014; Ross and Sampaio, 2018). For germ cell reprogramming, genome-wide removal of epigenetic marks, including for imprinted domains, occurs in primordial germ cells (PGCs) after their migration into the genital ridge (Hajkova et al., 2002). Rapid removal of DNA methylation in PGCs is mediated by TET (Ni

et al., 2016; Yu et al., 2013). For male germ cells, the acquisition of DNA methylation and establishment of paternal imprinting occur prior to meiosis in the spermatocytes before birth (Oakes et al., 2007), but for female germ cells, these occur during meiosis after birth (Smallwood et al., 2011). For reprogramming in preimplantation embryos, the paternal pronucleus undergoes rapid active removal of DNA methylation by TET in zygotes and the maternal pronucleus loses DNA methylation in a passive manner (Santos et al., 2002). The acquisition of DNA methylation in embryos largely completes before implantation (Borgel et al., 2010). During this wave of reprogramming, imprinted epigenetic marks inherited from the germ cells are maintained (Hirasawa et al., 2008).

Histone modifications have been found to regulate the reprogramming of DNA methylation. Genome-wide distribution of H3K36me3 is different in oocytes and sperm cells and is highly dynamic during pre-implantation embryo development, which is associated with the transcription activity of genes (Xu et al., 2019). Proper H3K36me3 deposition is required for the establishment of normal DNA methylation patterns during oogenesis, as *Setd2* mutation leads to dramatic loss of H3K36me3 and DNA methylation accompanied with disruption of H3K4me3, H3K27me3, and H3K27ac (Xu et al., 2019). DPPA3 is highly and specifically expressed in germ cells and preimplantation embryos (Saitou et al., 2002; Sato et al., 2002). In zygotes, allele specific deposition of H3K9me2, mainly on the maternal allele, recruits Dppa3 and suppresses TET activity which

leads to the passive demethylation of maternal pronucleus (Nakamura et al., 2012). Mitogen-activated protein kinase 3 (*MAPK3/ERK1*) and 1 (*MAPK1/ERK2*), which are highly expressed during zygotic genome activation, are involved in H3K9me2 and TET3 regulation (Chen et al., 2022).

Members from the KRAB-ZFP family, such as zinc finger protein 57 (*Zfp57*), zinc finger protein 445 (*Znf445*), and zinc finger protein 202 (*Znf202*), have been found to be involved in the establishment and maintenance of genomic imprinting during reprogramming (Monteagudo-Sánchez et al., 2020; Quenneville et al., 2011; Takahashi et al., 2019). The Krüppel-associated box (KRAB) domain of these proteins can specifically recognize methylated TGCCGC motif which is enriched in ICRs (Quenneville et al., 2011). These proteins can recruit tripartite motif containing 28 (TRIM28/KAP1) and further recruit UHRF1, DNMT1, and DNMT3A/B to maintain DNA methylation at the ICRs (Monteagudo-Sánchez et al., 2020; Quenneville et al., 2011; Takahashi et al., 2019). In addition, the binding of KRAB-ZFP family proteins to ICRs is locus specific, which indicates regulation specificities of different proteins (Monteagudo-Sánchez et al., 2020). *Uhrf1*, which can recruit *Dnmt1*, is needed to maintain imprinted DNA methylation at ICR in newly synthesized strand (Sharif et al., 2007). The mutation of lysine demethylase 1B (KDM1B), which is a histone demethylase, results in increased H3K4me2 and inhibited acquisition of DNA methylation at some imprinted domains in oocytes (Ciccione et al., 2009).

CTCF exists universally in most of cell types including germ cells and is required to maintain germ cell specific chromosome architectures and DNA methylation patterns (Fedoriw et al., 2004; Jung et al., 2019). CTCF is also necessary during early embryo development and is highly expressed after zygotic genome activation (Chen et al., 2019; Moore et al., 2012; Wan et al., 2008). Accordingly, TADs start to form after two cell stage in human embryos (Chen et al., 2019). Chromosome architecture has been reported to participate in the reprogramming events (Du et al., 2017; Ke et al., 2017). Allele specific TADs are found in preimplantation embryos (Du et al., 2017). The genome wide changes of DNA methylation profiles are associated with chromosome compartments (Ke et al., 2017).

Germ cells and embryos are sensitive to changes of environmental factors, including chemicals, oxygen level, and temperature, which could affect their reprogramming procedure (Desmet et al., 2016; Rahman et al., 2014; Skiles et al., 2018; Zhang et al., 2016). Given that ART procedures overlap with the critical periods of epigenetic reprogramming of oocytes and preimplantation embryos, suboptimal environments during ART could contribute to epigenetic errors in the embryos by disturbing reprogramming. Historically, serum supplementation during bovine ART has been used to promote blastocyst formation and is considered as the cause of epigenetic errors (Carolan et al., 1995; Edwards, 1965; Edwards et al., 1970; Thompson et al., 1995, 1998).

Indeed, differences in embryo culture conditions such as culture media and serum supplementation results in various changes of DNA methylation and gene expression in the embryos (Blondin et al., 2000; Doherty et al., 2000; Khosla et al., 2001; Wrenzycki et al., 1999, 2001). However, serum supplementation is not commonly used in human ART and is still associated with higher incidence of epigenetic defects, which suggests that other unknown factors exist to disrupt the reprogramming events and current ART procedures need improvement (Mussa et al., 2017).

1.3 Beckwith-Wiedemann Syndrome

Beckwith-Wiedemann syndrome (BWS, OMIM #130650) was first reported in the 1960s by two pediatricians Hans-Rudolf Wiedemann and J. Bruce Beckwith (Beckwith et al., 1964; Wiedemann, 1964). This syndrome was initially called EMG syndrome according to its major symptoms (i.e., **e**xomphalos, **m**acroglossia, and **g**igantism) and was renamed later (Filippi and Mckusick, 1970; Thorburn et al., 1970). The spontaneous incidence of BWS was reported as ~1/13,700 natural births in 1970s and as ~1/11,000 natural births in 2010s from two different countries (Mussa et al., 2013; Thorburn et al., 1970). Owing to its syndromic nature, these incidences may still be an underestimate since newborns with only minor clinical features may not go through molecular tests

and will not be diagnosed (Brioude et al., 2018). Most BWS cases are sporadic in families, and only ~15% of BWS patients showed autosomal dominant inheritance (Elliott et al., 1994; Pettenati et al., 1986).

The use of ART has been reported to increase the incidence of BWS. A meta-analysis of eight studies showed a relative risk of 5.2 to develop BWS for ART-conceived children than naturally conceived children (Vermeiden and Bernardus, 2013). Another study in Italy showed a higher relative risk of 10.7 based on ten years' live birth data (Mussa et al., 2017). However, given that in most cases infertility is the reason for patients undergoing fertility treatments, there was a concern for the observed correlation between ART and increased BWS incidence, which could be due to infertility of the parents (Doornbos et al., 2007; Fauser et al., 2014). One BWS case report suggested the causal effects of ART on BWS (Kuentz et al., 2011). In this report, a couple went through ART procedures to prevent the transmission of human immunodeficiency virus (HIV) to the infant since the father was a HIV carrier. The couple had no infertility issue detected and had normal genetic and epigenetic status at BWS associated imprinted domains, but their child was born with typical BWS clinical features and epimutations.

BWS consists of a broad spectrum of molecular aberrations and clinical features and will be reviewed below.

1.3.1 Molecular aberrations in BWS

The cause of BWS is associated with a very broad spectrum of genetic and epigenetic defects, which mainly belong to three categories: DNA methylation defects (in 65%-73% BWS patients with positive molecular diagnosis), changes of chromosomal contents (16%-33%), and gene mutations (1%-8%) at chromosome 11p15 (Brioude et al., 2013; Cooper et al., 2005; Ibrahim et al., 2014; Mussa et al., 2012, 2016).

Genome wide alterations of DNA methylation has been reported in BWS and some of the differentially methylated regions (DMRs) play regulatory roles for imprinted and non-imprinted gene expression (Krzyzewska et al., 2019). BWS was coined as a loss-of-imprinting syndrome since several imprinted genes were frequently found to be misregulated (Reik et al., 1995; Weksberg et al., 2010). Imprinting center 2 (IC2; KvDMR1) of *KCNQ1/CDKN1C* imprinted domain and imprinting center 1 (IC1) of *H19/IGF2* imprinted domain are the two main genomic regions affected by DNA methylation defects in BWS (Brioude et al., 2018). Loss of methylation at IC2 (IC2-LOM) and gain of methylation at IC1 (IC1-GOM) on the maternal allele are the common epimutations found in BWS patients, and ~87% patients with DNA methylation defects show IC2-LOM when summarizing five studies (Brioude et al., 2013; Cooper et al., 2005; Ibrahim et al., 2014; Mussa et al., 2012, 2016). Other ICs being reported to have DNA

methylation defects in BWS include the imprinted genes mesoderm specific transcript (*MEST/PEG1*; chromosome 7q32.2), *PLAG1* like zinc finger 1 (*PLAGL1/ZAC*; 6q24), insulin like growth factor 2 receptor (*IGF2R*; 6q25.3), GNAS complex locus (*GNAS*; 20q13.32), *GNAS* antisense RNA 1 (*GNAS-AS1/NESPAS*; 20q13.32), small nuclear ribonucleoprotein polypeptide N (*SNRPN*; 15q11), and growth factor receptor bound protein 10 (*GRB10*; 7p21) (Arima et al., 2005; Bliiek et al., 2009; Rossignol et al., 2006). These ICs often show maternal allele hypomethylation in patients with IC2-LOM, which was referred to as multi-locus imprinting disturbance and was considered to be the result of abnormal imprinting establishment in the oocyte (Brioude et al., 2018). The number of affected ICs in multi-locus imprinting disturbance and the level of hypomethylation is variable in different patients (Bliiek et al., 2009). The ratio of IC2-LOM patients with multi-locus imprinting disturbance is about equal between naturally and ART conceived BWS (Rossignol et al., 2006). As mentioned before, *DNMT3L* plays important roles in the establishment of both maternal and paternal genomic imprints, however, no correlation could be found between multi-locus imprinting disturbance and *DNMT3L* mutations in BWS patients (Bliiek et al., 2009).

BWS originating from changes of chromosomal contents include uniparental disomy (UPD, (Henry et al., 1991)), duplications (Waziri et al., 1983), deletions (Beygo et al., 2016; Schmutz, 1986), trisomy (Okano et al., 1986;

Turleau et al., 1984), translocations (Kaltenbach et al., 2013), and inversions (Mannens et al., 1994). *CDKN1C* mutation is the main form of gene mutation in BWS (Hatada et al., 1996). BWS patients with these two types of molecular defects (changes of chromosomal contents and gene mutations) are inheritable (Brioude et al., 2018).

1.3.2 Clinical features of BWS and corresponding treatments

Owing to its various molecular aberrations, BWS patients show a broad spectrum and various combination of symptoms, and the criteria for clinical diagnosis advanced over time (Brioude et al., 2018; Elliott et al., 1994). Initially, the criteria for BWS diagnosis was that newborns with three primary features (macroglossia [large tongue], abdominal wall defects [omphalocele/hernia], and macrosomia [large body size, greater than 90th centile of control/population]) or with two primary plus three secondary features (ear malformations, facial nevus flammeus, organomegaly, and neonatal hypoglycemia) were considered to be BWS (Elliott et al., 1994).

Along with the development of technologies for molecular diagnosis, recently, an expert consensus defined cardinal (worth two points) and suggestive (one point) clinical features for BWS diagnosis and divided BWS into three subtypes based on phenotypic and (epi)genetic characteristics, namely classical

BWS, isolated lateralized overgrowth, and atypical BWS (Brioude et al., 2018).

Cardinal features of BWS include macroglossia, omphalocele, lateralized overgrowth, multifocal and/or bilateral Wilms tumor or hepatoblastoma, and hyperinsulinism (Brioude et al., 2018). Suggestive features of BWS include macrosomia (greater than two times of standard deviation above the mean of control/population), facial naevus simplex, ear malformations, transient hypoglycemia, umbilical hernia, diastasis recti, some other kinds of tumors, and some types of organomegaly (placentomegaly, nephromegaly, hepatomegaly) (Brioude et al., 2018). Infants with a total clinical feature score greater than two points will undergo molecular tests to evaluate DNA methylation levels at IC1 and IC2, copy number variation, *CDKN1C* mutations, and chromosomal abnormalities. If none of the molecular subtypes can be confirmed, infants with score greater than four points will still be diagnosed as BWS. Classical BWS shows a combination of some of the cardinal and suggestive features with or without molecular aberrations, isolated lateralized overgrowth shows asymmetrical hemihypertrophy or hemihyperplasia with molecular aberrations, and atypical BWS shows fewer cardinal and suggestive features than classical BWS but have molecular aberrations (Brioude et al., 2018; Kalish et al., 2017).

The frequencies of BWS clinical features vary in different studies and the summarized ranges are: macroglossia (72%-100% of patients), abdominal wall defects (>60%), lateralized overgrowth (~20%-38%), tumorigenesis (5%-25%),

hypoglycemia (32%-63%), macrosomia (43%-88%), and facial nevus simplex (30%-73%) (Bliek et al., 2004; Brioude et al., 2013; Cooper et al., 2005; DeBaun et al., 2002; Elliott et al., 1994; Engström et al., 1988; Gaston et al., 2001; Goldman et al., 2002; Ibrahim et al., 2014; Mussa et al., 2012, 2016; Pettenati et al., 1986; Thorburn et al., 1970). Their associations with specific molecular aberrations and treatments will be reviewed below.

Macroglossia showed a higher frequency (~10-20%) in UPD patients than in IC1-GOM patients in three studies, but was not consistent in IC2-LOM and *CDKN1C* mutation patients (Brioude et al., 2013; Ibrahim et al., 2014; Mussa et al., 2016). Severe macroglossia can cause difficulties in feeding/breathing, obstructive sleep apnea, malocclusion, and speech delay, and requires tongue reduction surgeries (a partial glossectomy) (Elliott et al., 1994; Style et al., 2018). Feature of macroglossia will still present in the adult without surgery but won't be as obvious as in neonates (Thorburn et al., 1970).

The frequency of abdominal wall defects, including omphalocele, umbilical hernia, and diastasis recti, is significantly higher in IC2-LOM and *CDKN1C* mutation patients than in IC1-GOM and UPD patients (Brioude et al., 2013; Cooper et al., 2005; Ibrahim et al., 2014; Mussa et al., 2016). An omphalocele is the outward protrusion of abdominal organs through the umbilical cord and these organs are covered by membranes (i.e., amnion, peritoneum, and Wharton's jelly) but not skin (Bair et al., 1986). An umbilical hernia is a bulge of

abdominal organs covered by skin at the umbilicus, which is caused by incomplete closure of umbilical ring (Jackson and Moglen, 1970). A diastasis recti is the separation of rectus abdominis muscle at the linea alba sometimes accompanied with bulging of abdominal organs (Boissonnault and Blaschak, 1988). Omphalocele and large umbilical hernia requires immediate corrective surgeries after birth, and for a severe omphalocele, staged surgeries to fit the organs into the abdomen need to take several months (Style et al., 2018). Minor umbilical hernia can heal spontaneously before six years of age.

Lateralized overgrowth (hemihyperplasia or hemihypertrophy) is the asymmetric growth of body parts and is significantly higher in UPD patients than others (Brioude et al., 2013; Cooper et al., 2005; DeBaun et al., 2002; Ibrahim et al., 2014; Mussa et al., 2016). Lateralized overgrowth has a lower frequency in IC2-LOM with multi-locus imprinting disturbance than in IC2-LOM only patients (Bliek et al., 2009). Lateralized overgrowth has also been shown closely related to increased risk of tumor development (DeBaun and Tucker, 1998). Surgical treatments for lateralized overgrowth include epiphysiodesis which is to temporarily fuse the epiphyseal plate of a bone to slow down growth of the side with overgrowth (Style et al., 2018).

BWS patients have an increased incidence (4-7.5%) of childhood tumorigenesis before eight years of age (DeBaun et al., 1998). The frequency of tumorigenesis is higher in IC1-GOM and UPD patients than in IC2-LOM and

CDKN1C mutation patients (Ibrahim et al., 2014; Mussa et al., 2016). Wilms tumor (50%-80% tumor cases), hepatoblastoma (10%-30%), and some low frequency malignant tumors have been reported in BWS (Bliek et al., 2004; Brioude et al., 2013; DeBaun et al., 2002; Gaston et al., 2001; Goldman et al., 2002; Mussa et al., 2016; Pettenati et al., 1986). BWS patients need to undertake a series of regular screenings to monitor risk of tumorigenesis, including abdominal ultrasound and urinalysis every three months until age eight, blood draws to check serum alpha-fetoprotein levels every two to three months until age four, and physical examination every six months until body growth is complete (Clericuzio, 1999; Shuman et al., 2016). Tumors will be treated with chemotherapy and organ removal or transplant (Style et al., 2018; Trobaugh-Lotrario et al., 2014).

Hypoglycemia caused by hyperinsulinism could last as long as three years in BWS patients (Moncrieff et al., 1977; Schiff et al., 1973). One study has shown a higher frequency of hypoglycemia in UPD patients than other molecular subtypes (Brioude et al., 2013), but this was not shown in several other studies (Cooper et al., 2005; Ibrahim et al., 2014; Mussa et al., 2016). Most (~80%) of BWS patients with hyperinsulinism and hypoglycemia show mild symptoms and require only dietary supplementation or short-term intravenous administration of glucose, and partial pancreatectomy could be applied for severe cases (Elliott et al., 1994; Style et al., 2018).

Macrosomia was considered as a primary BWS clinical feature until many patients without macrosomia were molecularly diagnosed (Brioude et al., 2013). Two studies have shown opposite associations between frequency of macrosomia and IC2-LOM with multi-locus imprinting disturbance (Bliek et al., 2009; Rossignol et al., 2006). Macrosomia is often coupled with accelerated childhood growth that will slow down before adolescence (Elliott and Maher, 1994).

Facial nevus simplex is caused by malformation of the capillary as a flat pink area at birth, will become darker and thicker due to superficial vessel ectasia, and tends to be lifelong if not treated (Patel et al., 2012; Raulin et al., 1999; Tan et al., 1989). Facial nevus simplex could cause psychological morbidities (e.g., impaired personality development) and easy bleeding after trauma (Tan et al., 1989). The frequency of facial nevus simplex is lower in IC1-GOM patients than other molecular subtypes (Ibrahim et al., 2014; Mussa et al., 2016). In addition, the frequency of facial nevus simplex is lower in IC2-LOM with multi-locus imprinting disturbance than IC2-LOM only patients (Bliek et al., 2009). Facial nevus simplex is also positively correlated with macroglossia and ear malformations (Cooper et al., 2005). Facial nevus simplex can be treated with pulsed dye laser to shrink the size of blood vessel and improve appearance (Jeon et al., 2019).

1.4 Large/abnormal offspring syndrome

LOS/AOS is a naturally occurring congenital overgrowth syndrome in ruminants and its incidence increases with the use of ART (Rivera et al., 2021). The incidences of spontaneous LOS (SLOS; conceived by natural mating or AI) and ART-induced LOS are currently unknown for bovine. ART procedures are widely used in the cattle industry, especially for dairy, to improve genetic merit (i.e., ability to produce superior offspring) of the offspring in a shorted time period when compared to natural reproduction (Kadarmideen et al., 2018). Although there is no official documentation of the incidence of LOS in ART produced calves, the estimated incidence is as high as 10% according to information provided to us by two embryo transfer companies in the US (Li et al., 2019b). Until recently, all the reports on LOS involved the use of ART (Behboodi et al., 1995; Chen et al., 2013; Farin and Farin, 1995; Farin et al., 2001; Gao et al., 2019; Hasler et al., 1995; Hori et al., 2010; Jacobsen et al., 2000; Kruij and Den Daas, 1997; Lazzari et al., 2002; Liu et al., 2013; Su et al., 2011b; van Wagtendonk-de Leeuw et al., 1998, 2000; Willadsen et al., 1991; Wilson et al., 1995; Young et al., 1996). In 2019, we were the first group to document its existence in naturally conceived individuals (Li et al., 2019b) (also please refer to Appendix 2 in this dissertation). Previous studies from our laboratory have shown similarities between LOS and BWS in both clinical features and molecular

aberrations, which also suggests LOS as a good animal model for BWS (Chen et al., 2013, 2015, 2017; Li et al., 2019a).

1.4.1 Molecular aberrations in LOS

Like BWS, LOS is also closely associated with epigenetic defects which could be results of incorrect reprogramming of epigenome during the formation of gametes and during early embryo development (Farin et al., 2006). Genome-wide misregulation of protein-coding gene, lncRNAs, and miRNA, loss of imprinting, and changes in DNA methylome have been reported in bovine LOS (Chen et al., 2013, 2015, 2017; Li et al., 2019a; O'Doherty et al., 2018; Su et al., 2014). Different patterns of DNA methylation changes have been reported to associate with the timing of embryo transfer during *in vitro* production of embryos (Salilew-Wondim et al., 2015). However, there is a lack of correlation between changed DNA methylation and gene expression in short distance (i.e., 20kb), which indicates the involvement of other mechanisms for long distance regulation (Chen et al., 2017). The misregulation of hundreds of miRNAs in bovine LOS has been reported in my MS studies (Li et al., 2019a). These misregulated miRNAs exhibit clustering tendency in their genomic locations and could affect the expression of thousands of gene at protein level given their roles in translation inhibition. Loss-of-imprinting at imprinted domains in bovine LOS have been

reported for *IGF2R*, *KCNQ1*, *IGF2*, *PLAGL1*, *PEG3*, and *DLK1* (Chen et al., 2013, 2015, 2017; Hori et al., 2010; Li et al., 2019a, 2022; Sangalli et al., 2014; Su et al., 2011a, 2011b). Aberrations of histone modifications are often associated with somatic cell nuclear transfer (SCNT) derived bovine fetuses (Arnold et al., 2017; Lin et al., 2008; Santos et al., 2003), which is partially due to original somatic histone modifications escaping reprogramming (Wee et al., 2006).

1.4.2 Clinical features of LOS

Microsomia is the most common clinical feature identified in LOS, including enlarged body and limb sizes coupled with increased skeletal length which can be two and five times larger than the average size at birth, respectively (Chen et al., 2013; Farin and Farin, 1995; Farin et al., 2001; Walker et al., 1996; Young et al., 1998). These increases can be detected as early as the fifth week of gestation in cattle (Hansen et al., 2016). Calves with microsomia at birth could be mediated during growth and reach normal mature body weight (Wilson et al., 1995). Although microsomia is no longer considered as a primary clinical feature for BWS, the identification of LOS in the bovine industry is still largely based on this clinical feature due to the lack of knowledge on biomarkers for molecular tests and neglect of other minor symptoms. Indeed, LOS is not always

characterized by overgrowth and other clinical features include macroglossia, abdominal wall defects (umbilical hernias and omphalocele), organomegaly of heart, liver, and kidney, placentomegaly, muscle and skeleton malformation, hydrallantois, abnormal placental vasculature, and even increased early embryo or fetus death rates (Farin and Farin, 1995; Farin et al., 2001, 2006; Hasler et al., 1995; McEvoy et al., 1998; van Wagtendonk-de Leeuw et al., 1998). In addition, LOS can also affect the welfare of dam through increased gestation length and dystocia rate (Kruip and Den Daas, 1997; Sinclair et al., 1995), and even lead to the death of dam, which will bring financial loss to producers with an estimation of \$29,000 per case (Rivera et al., 2021).

1.5 Concluding remarks

Given the important roles of chromosome architecture in gene expression regulation, it is necessary to investigate its genome-wide involvement in bovine LOS and human BWS. Identification and characterization of SLOS is needed to promote the understanding of etiology of this syndrome. In addition, determination of robust biomarkers for the early diagnosis of LOS is required urgently. The following studies present in this dissertation, regarding to chromosome architecture of bovine control and ART-LOS fibroblasts and DNA methylome of control, SLOS, and ART-LOS, added information to these

knowledge gaps.

Rationale for Dissertation

As mentioned before, there is low correlation between aberrant DNA methylation and altered gene expression in bovine LOS (Chen et al., 2017; Li et al., 2019a). Given the important roles of chromosome architecture in remote gene expression regulation, our overall hypothesis is that genome-wide DNA methylation defects alter chromosome architecture in bovine LOS resulting in aberrant gene expression from remote genomic loci. Our *in silico* prediction of CTCF binding sites within *IGF2R* ICR and KvDMR1 identified putative binding sites for this insulator protein thus suggesting that chromosome architecture may be involved in the normal regulation of these imprinted domains. In Chapter 2, we hypothesized that *IGF2R* ICR and KvDMR1 have allele-specific chromosome architecture in control bovine fetuses and disrupted chromosome architecture in LOS fetuses. To test this hypothesis, we conducted 4C sequencing for these domains to characterize their chromosome architecture as well as whole genome bisulfite sequencing (WGBS) and RNA sequencing to query DNA methylation status and gene expression levels using fibroblast cells derived from control and LOS fetuses. This work is presented in Chapter 2. In addition, to test the overall hypothesis that LOS and BWS are the result of genome-wide defects of chromosome topology, we initiated a Hi-C project (ongoing, Appendix 1) which will characterize the global chromosome architecture, DNA methylome and

transcriptome in human and bovine samples.

Although we have recently reported the presence of bovine SLOS based on clinical features in two reviews (Li et al., 2019b; Rivera et al., 2021), no data exists to demonstrate that SLOS shares epimutations with ART-LOS. It is important to characterize molecular aberrations in SLOS to better understand the etiology and the impact of method of conception (i.e., ART) on the development of LOS. In addition, it is still largely unknown whether the thousands of genomic loci with DNA methylation defects in LOS are equally important for LOS development or if there exist some key loci which are fundamental for LOS development (Chen et al., 2017). We hypothesized that bovine SLOS has similar DNA methylation defects as ART-LOS, and that there exist vulnerable genomic loci to DNA methylation epimutations in LOS. To test this hypothesis, we performed WGBS to identify DNA methylation defects in bovine SLOS and ART-LOS using samples from various breeds, developmental stages, and tissue types. This work is presented in Chapter 3.

Chapters 2 and 3 have already been published in *iScience* (<https://doi.org/10.1016/j.isci.2022.104269>) and *Epigenetics* (<https://doi.org/10.1080/15592294.2022.2067938>), respectively. For each chapter, the main figures, tables, and supplemental figures are included at the end of chapter, after the acknowledgements section, and the supplemental Excel tables are available from the online version of the manuscript.

Chapter 2: Allele-specific aberration of imprinted domain chromosome architecture associates with large offspring syndrome

2.1 Abstract

Large offspring syndrome (LOS) and Beckwith-Wiedemann syndrome are a similar epigenetic congenital overgrowth conditions in ruminants and humans, respectively. We have reported global loss-of-imprinting, methylome epimutations, and gene misregulation in LOS. However, less than 4% of gene misregulation can be explained with short range (<20kb) alterations in DNA methylation. Therefore, we hypothesized that methylome epimutations in LOS affect chromosome architecture which results in misregulation of genes located at distances >20kb in cis and in trans (other chromosomes). Our analyses focused on two imprinted domains that frequently show misregulation in these syndromes, namely KvDMR1 and *IGF2R*. Using bovine fetal fibroblasts, we identified CTCF binding at *IGF2R* imprinting control region but not KvDMR1, and allele-specific chromosome architecture of these domains in controls. In LOS, analyses identified erroneous long-range contacts and clustering tendency in the direction of expression of misregulated genes. In conclusion, altered chromosome architecture is associated with LOS.

2.2 Introduction

Large/abnormal offspring syndrome (LOS/AOS) is a naturally occurring congenital overgrowth syndrome in ruminants and its incidence increases with the use of assisted reproductive technologies (ART) (Rivera et al., 2021). Frequently observed abnormal phenotypes in LOS include macrosomia, omphalocele, and abnormal organ development (Rivera et al., 2021). Gestational problems associated with the dam include dystocia and in some cases fetal and/or maternal death (Rivera et al., 2021). In humans, a similar congenital overgrowth syndrome, Beckwith-Wiedemann syndrome (BWS, OMIM #130650), shares phenotypes and molecular aberrations with LOS (Brioude et al., 2018; Li et al., 2019a).

Previous studies from our laboratory have reported global alterations of DNA methylation and global loss-of-imprinting, including at the *KCNQ1* (KvDMR1) and *IGF2R* imprinted domains in LOS fetuses (Chen et al., 2013, 2015, 2017), signatures also observed in BWS (Rossignol et al., 2006). Genomic imprinting is an epigenetic phenomenon which controls parental-allele-specific expression of a subset of genes involved in fetal development in mammals by regulating allele-specific DNA methylation status at their regulatory regions (i.e., imprinting control region; ICR) (Verona et al., 2003). We also reported global misregulation of protein-coding genes and non-coding RNAs in these LOS

fetuses (Chen et al., 2017). In addition, we found global misregulation of microRNA genes in several tissues of LOS fetuses and human tongue of BWS patients (Li et al., 2019a). However, a limited number of misregulated genes (<4%) located within 20 kb (short range) of regions with aberrant DNA methylation showed associations in LOS (Chen et al., 2017; Li et al., 2019a).

It is well established that DNA methylation can affect gene expression at remote regulatory regions (i.e., kilobases/megabases away and also interchromosomal) by regulating chromosomal architecture (Ong and Corces, 2014). Chromosomal architecture defines spatial organization of the genome during interphase and includes topologically associating domains (TADs), chromosome compartments, and chromosome territories (Dixon et al., 2012). TADs are self-interacting genomic regions within a chromosome that range from kilobases to megabases in length. The formation and maintenance of TADs primarily relies on architectural proteins such as CCCTC-binding factor (CTCF) and cohesin protein complex (Nora et al., 2017; Rao et al., 2017), although CTCF-free TAD boundaries exist in mice and humans (Kagey et al., 2010). Binding of CTCF can be inhibited by DNA methylation (Rodriguez et al., 2010). TADs can either facilitate spatial interactions between chromosomal loci by reducing distance through looping or block spatial interactions by physical insulation of the TADs from the rest of the genome (Doyle et al., 2014; Lupiáñez et al., 2015). TADs also serve as regulators of gene expression, including

imprinted genes, by orchestrating physical interactions of regulatory regions with gene promoters (Llères et al., 2019), and their disruption can lead to aberrant expression of genes and cause severe diseases or malformations (Lupiáñez et al., 2015).

In this study, we hypothesized that *IGF2R* ICR and KvDMR1 have allele-specific chromosome architecture in control bovine fetuses and disrupted chromosome architecture in LOS fetuses. We used fibroblast cells derived from skin of day 105 *Bos taurus indicus* x *Bos taurus taurus* F1 hybrid control and LOS fetuses to perform circular chromosome conformation capture (4C) sequencing and detected allele-specific and condition-specific chromosome architectures which are associated with allele-specific DNA methylation status and global gene misregulation.

2.3 Results

We investigated the potential contribution of chromosome architecture in LOS by querying two of its commonly misregulated imprinted domains, namely *IGF2R* ICR and KvDMR1 using control (n=4) and LOS (n=4) fibroblasts.

Identification of paternal genomic variants

Genomic DNA of the bull's spermatozoa that sired all fetuses was

sequenced to determine parental origin of fetal DNA. Bovine genome assembly ARS-UCD1.2 (Rosen et al., 2020) was used for all sequencing alignments. In total, ~ 17.6 million raw short variants were identified from which ~ 17.5 million were retained after filtering and used as reference for whole genome bisulfite sequencing (WGBS) data analyses.

Global determination of DNA methylation status

WGBS was conducted for control and LOS fibroblasts to determine their global DNA methylation status. Read alignment information may be found in Table S1.A. Group comparison between control and LOS identified 9,634 significantly differentially methylated regions (DMRs) on chromosomes 1-29 (Table S1.B). Loss-of-methylation in LOS was observed for *IGF2R* ICR (Chr9:96220970-96221849 and 96222010-96223969, 26.8% and 50.6% decrease, respectively) and KvDMR1 (Chr29:48907678-48909917, 24.7% decrease) (Figure 1).

Allele-specific DNA methylation of a 385 bp region (Chr29:48908122-48908506) within the identified DMR of KvDMR1 was also determined through bisulfite PCR, cloning, and Sanger sequencing with primers previously used by us (Chen et al., 2013). A SNP at Chr29:48908415 (maternal=C, paternal=G, missing in LOS #4) was used to assign parental-allele identity. Fibroblast and

muscle samples from Control #2 and #4 showed >94% maternal allele methylation and 0% paternal allele methylation. For the LOS group, only LOS #3 showed <50% methylation on the maternal alleles for this region (Figure S1).

CTCF binding site prediction

In silico prediction identified 32,732 potential CTCF binding sites in the bovine genome with two localizing within the *IGF2R* ICR (Chr9:96223387-96223405 and 96223413-96223431) and two within KvDMR1 (Chr29:48908185-48908198 and 48908388-48908396). Comparisons between predicted CTCF binding sites and global DMRs showed that 125 CTCF sites overlap with 122 DMRs (Table S1.C).

Confirmation of CTCF binding to *IGF2R* ICR and KvDMR1

Binding of CTCF protein to the *in silico* predicted CTCF binding sites was investigated by chromatin immunoprecipitation (ChIP). For this, PCR primers were designed to include the putative CTCF binding sites and a SNP for *IGF2R* ICR and KvDMR1. The SNP for *IGF2R* ICR is located at Chr9:96223677 (maternal=C, paternal=T) and for KvDMR1 is located at Chr29:48908415 (maternal=C, paternal=G). In addition, PCR primers were designed for a region of *IGF2R*'s intron 3 which contained no predicted CTCF binding site to set

background levels, which could be caused by cohesin bound CTCF during chromatin extrusion activity (Davidson et al., 2019b; Sanborn et al., 2015). For *IGF2R* ICR, a PCR amplicon was visible for the CTCF ChIP, positive controls 2% input DNA and Histone H3 ChIP, but not for the negative control (rabbit IgG) in all the control and LOS fibroblasts (Figure 2.A). The ratio of CTCF ChIP to input was significantly higher ($p < 0.00001$) for *IGF2R* ICR (~143.39% on average) than for *IGF2R* intron 3 (~38.85% on average), indicating specific binding of CTCF at *IGF2R* ICR (Figure 2.B). There was more CTCF bound at the *IGF2R* ICR in LOS when compared to controls ($p < 0.01$). For KvDMR1, the ratio of CTCF to input was ~18.65% on average and was not higher than the background level identified using *IGF2R* intron 3, suggesting lack of CTCF binding at KvDMR1 (Figure 2.B). We also determined the ratio of the parental alleles in the CTCF ChIP PCR amplicons based on the Sanger sequencing fluorescence intensity of the SNP, and found a higher ratio of maternal alleles in LOS than controls (Figure 2.C-D).

4C assay design and sequencing

To determine whether alteration of chromosome architecture occurs in LOS or not, we designed and performed 4C sequencing, a technique used to identify contacts between a target region (referred to as bait) and other regions of

the genome (Krijger et al., 2020). 4C assays were designed for *IGF2R* ICR which included the CTCF binding sites within the bait region. It is known that in 4C sequencing the use of different restriction enzymes has impact on the results (Krijger et al., 2020; van de Werken et al., 2012). This is because the recognition sequence of each enzyme has a different prevalence in the genome, thus resulting in length differences of digested and ligated DNA fragments during sequencing library preparation. Very short DNA fragments are difficult to uniquely align back to the genome, and the sequencing quality and successfulness are decreased for long fragments (Tan et al., 2019). In this study we used two restriction enzymes for the second digestion to maximize the identification of contacts, and the assays were named IGF2R_MseI (4 base cutter) and IGF2R_BsrI (5 base cutter). We present the data generated from the MseI digestion in the main manuscript and for BsrI in the supplement.

For KvDMR1, even though no CTCF binding was detected, we were still interested in identifying potential loss-of-methylation related changes in genome architecture at this domain, this assay was named KvDMR1. For this assay, we only included one LOS sample (i.e., #3) since our bisulfite sequencing results only showed this individual as having maternal loss-of-methylation. For these assays, the previously mentioned SNPs were included in the sequencing reads to differentiate the parental alleles and only one enzyme for the second digestion was used (i.e., Tsp45I) since no other specificity was available.

In addition to the control and LOS groups, fibroblasts from the control group were treated with 0.5 uM Decitabine (5-aza-2'-deoxycytidine) for 96 hours to serve as a positive control for loss-of-methylation and named as DC group. Alignment information for all assays and samples are found in Table S2.A.

General TAD structure identified by 4C sequencing

We first compared the overall (not allelic) pattern of chromosomal interactions with each bait regions using a running window size of 50 NlaIII restriction fragments, since we were interested in identifying large-scale structural changes in LOS. For *IGF2R*, the control group of IGF2R_MseI and IGF2R_BsrI assays showed reads enriched regions of ~ 273 kb (chr9:96033799-96306553) and ~ 174kb (chr9:96121421-96295401) identified by software fourSig (Williams Jr et al., 2014), respectively (Figure S2.A-B). These interaction enriched regions around the bait are considered to be sub-TADs (Lupiáñez et al., 2015). We identified that both the LOS and DC groups have a large number of differentially interacting regions when compared with controls identified by software DESeq2 (Love et al., 2014) (Table S3.A). Of note, for the DC-induced global demethylation group, we only focused on the cis results to characterize local topology as far-cis and trans results could be affected by additional demethylation events. In addition, statistical analyses for read enrichment and

differences between groups were also conducted with a second software 4Cker as corroboration (Raviram et al., 2016) (Figure S3). However, due to 4Cker's lack of consistency among different assays in identifying far-cis contacts ($>\pm 500$ kb on the same chromosome) and inability to perform trans (interchromosomal) analysis, fourSig and DESeq2 results were reported as main results and used for downstream analyses.

For cis contacts (± 500 kb flanking the bait), the IGF2R_Msel assay identified gain of contact in 14 merged windows and loss of contact in 3 merged windows in LOS when compared to controls (Figure S2.A, Table S2, and Table S3.A) and the IGF2R_Bsrl assay identified gain of contact in 33 merged windows and loss of contact in 4 merged windows in LOS when compared to controls (Figure S2.B, Table S2, and Table S3.A). Both assays showed that the DC group mainly identified gain of cis contacts (Table S2 and Table S3.A). For visual comparison of sub-TADs between assays and treatments please refer to Figure S4. The read enrichment in each individual sample can be found in Figure S5. For far-cis contacts ($>\pm 500$ kb on the same chromosome), the IGF2R_Msel assay identified gain of contact in 91 merged windows and loss of contact in 64 merged windows in LOS when compared to controls (Figure S2.D, Table S2, and Table S3.A). While the IGF2R_Bsrl assay identified gain of contact in 81 merged windows and loss of contact in 29 merged windows in LOS when compared to controls (Figure S2.E, Table S2, and Table S3.A). For trans contacts

(interchromosomal), the IGF2R_MseI assay identified gain of contact in 19 merged windows and loss of contact in 19 merged windows in LOS when compared to controls (Figure S2.F, Table S2, and Table S3.A). While the IGF2R_BsrI assay identified gain of contact in 107 merged windows and loss of contact in 9 merged windows in LOS when compared to controls (Figure S2.G, Table S2, and Table S3.A). Of relevance, several change of contact regions identified by IGF2R_MseI and IGF2R_BsrI assays colocalized with predicted CTCF binding sites or DMRs (Table S2 and Table S3.A).

For the KvDMR1 assay, the control group showed a cis reads enriched region of ~ 635 kb (chr29:48610455-49245209) (Figure S2.C, Table S2 and Table S3.A). Statistical comparisons with LOS were not done since only one fetus was analyzed. DC group showed gain of contact in 14 merged windows in cis when compared to controls (Table S2 and Table S3.A).

Allele specific sub-TAD in control group

Parental allele specific analyses of 4C sequencing data were conducted based on the SNP retained in the sequencing reads, as previously mentioned. It should be noted that allele specificity only applies to the bait sequence but not the interacting regions. We first compared the paternal allele to the maternal allele in the control group for the three 4C assays. A higher percentage of 4C

reads were captured by the paternal allele bait in all three datasets (Figure S6), indicating that the unmethylated paternal alleles of KvDMR1 and *IGF2R* ICR have higher frequency of physical interactions with other chromosomal loci. For *IGF2R* domain, the paternal allele had a larger sub-TAD than the maternal allele, for example the IGF2R_Msel assay identified sub-TADs of ~ 149 kb (chr9:96134602-96283672) and ~ 265 kb (chr9:96041522-96306553), for the maternal and paternal alleles, respectively (Figure 3.A, Table S3.B, and Table S4.A). The paternal sub-TAD incorporates an additional four protein-coding genes, namely *TCP1*, *PNLDC1*, *MRPL18*, and *MAS1*, and two small nucleolar RNA genes, namely *LOC112448166* and *LOC112448168*. For far-cis and trans contacts, two contacts are different between the alleles (Figure 3.D-E). In addition, as expected, the enrichment of contacts is different between the IGF2R_Bsrl and IGF2R_Msel assays, nonetheless, they identified similar pattern of contacts around the bait (Figure S7.A, Figure S4, Table S3.B, and Table S4.B).

For KvDMR1, the maternal and paternal alleles in the control group showed reads enriched regions of ~ 575 kb (chr29:48634040-49209049) and ~ 640 kb (chr29:48605573-49245209), respectively (Figure 4.A, Table S3.B, and Table S4.C). About equal number of increased (6 merged windows) and decreased (7 merged windows) far-cis contacts were identified on chromosome 29 for paternal alleles when compared with maternal alleles (Figure 4.D). Trans contacts identified only two differences with chromosome 11 (increase in

paternal) and chromosome 14 (decrease in paternal; Figure 4.E). The read enrichment in each individual sample can be found in Figure S8-S9. 4Cker results show general gain of cis contacts for all three assays (Figure S10-S12.A).

Altered allele specificity of sub-TAD in LOS

Next, we compared the unmethylated paternal allele between LOS and controls. For the *IGF2R* domain, both IGF2R_Bsrl and IGF2R_Msel assays showed similar sub-TAD structure between treatments (Figure 3.B, Figure S7.B, Figure S4, Table S3.C, and Table S4). The two assays identified similar number of gain of far-cis contacts but IGF2R_Msel showed more loss of far-cis contacts in LOS compared to controls (Figure 5.A and Figure S7.E) and IGF2R_Bsrl assay captured more gain of trans contacts and less loss of trans contacts than IGF2R_Msel assay for LOS (Figure 5.D and Figure S7.H). The DC group showed extended sub-TAD and more gain of contacts at surrounding regions, which could be due to the impacts of global loss-of-methylation on chromosome architecture (Figure 3.B and Figure S7.B). For KvDMR1, the paternal allele behaved similarly between groups (Figure 4.B, Table S3.C, and Table S4). 4Cker results for the three assays can be found in Figure S10-S12.

Lastly, we determined the effects of loss-of-methylation on the maternal allele in LOS. For the *IGF2R* domain, we observed that the sub-TAD structure of

the maternal allele in LOS resembles the paternal allele of controls for both IGF2R_MseI (chr9:96087833-96303166) and IGF2R_BsrI (chr9:96121486-96295401) assays which is likely the result of loss-of-methylation, as the sub-TAD structure is the same in the DC and LOS maternal alleles (Figure 3.C, Figure S7.C, Figure S4, Table S3.D, and Table S4). Further, this allele gained more far-cis and trans contacts in LOS for both assays (Figure 5.G and I, Figure S7.F and I). For KvDMR1, the maternal allele behaved similarly between groups (Figure 4.C, Table S3.D, and Table S4). 4Cker results for the three assays can be found in Figure S10-S12.

Next, we analyzed the genomic contexts (i.e., predicted CTCF binding sites, repetitive sequences, promoters, gene bodies, exons, introns and CpG islands, shores, and shelves) of the *IGF2R* ICR's far-cis and trans contacts which are significantly different between LOS and control (Figure 5 and Figure S14). Further, in order to determine the density of genes in the regions overlapping the contacts, we calculated the number of genes/million bases (Figure S13 and Figure S15). The specificity of the enrichment of different genomic contexts over changed contacts was determined by a permutation test by shuffling each region ten thousand times. For this, the shuffle in cis included \pm 500 kb of the bait, the far-cis shuffle included the chromosome without the baits' cis region, and the trans shuffle was performed with all chromosomes except the chromosome containing the bait region (Figure S13 and Figure S15). For both IGF2R_MseI

and IGF2R_Bsrl assays, the increased far-cis contacts in LOS show enrichment for CpG islands, shores, and shelves, especially for shores, and depletion of repetitive sequences (Figure S13. A and E, and Figure 15. A and D). This enrichment of CpG shelves could be partially due to the uneven distribution of CpG islands on chromosome 9 since there is an enrichment around chr9:101000000-105000000. This is not the case for repetitive sequences as they do not show uneven distribution on chromosome 9.

Gene expression and its association with 4C identified interactions

Transcriptome analyses identified differences between LOS and controls (Table S5.A). As expected, genes associated with extracellular matrix, including collagen, vimentin, thrombospondin, tenascin, fibronectin, and filamin were highly expressed in fibroblasts of both groups. In total, there were 548 differentially expressed genes (DEG) between controls and LOS, including *IGF2R* with ~ 3.5 folds downregulation in LOS (Figure 6 and Table S5.A), similar to the results observed in muscle by quantitative RT-PCR (~ 3.3 folds downregulation in LOS). Enriched signaling pathways include lysosome, glycan degradation, and glycosaminoglycan degradation (Table S5.B).

Genome coordinate based analyses identified clustering tendency of DEG. In total, 149 (27.2%) DEGs were found within 200 kb of another DEG,

which is the average size of sub-TADs (Montefiori et al., 2018) (Figure 6 and Table S5.C). Most of the clustered DEGs show the same direction of misregulation. Given that the distribution of genes in the genome is not even, we conducted a permutation test to confirm this clustering tendency. The test was repeated ten thousand times, and for each time 548 genes (same number as DEGs) were randomly picked from the 15,042 expressed genes in our sample and the number of clustered genes in 200kb was calculated. The identified 149 clustered DEGs is significantly higher ($p = 0$) than the mean of permutation tests (91.46 with standard deviation 8.49) (Figure S16).

Further, DMRs were identified within the promoter region of 0.9% DEGs ($n=5$). These showed an inverse correlation between DNA methylation and transcript abundance (Table S5.D). When associating DEGs within a 100 kb flanking region from the gain/loss of contacts identified by 4C sequencing, eight (1.5%) DEGs were found for the IGF2R_Msel assay (*RAB8B*, *MMP2*, *LOC100848985*, *FAM126A*, *EPHA5*, *RGS17*, *QKI*, *RFNG*) and eight (1.5%) DEGs were found for the IGF2R_Bsrl assay (*DCLK1*, *FJX1*, *PYCR1*, *LOC112442278*, *MIR2887-1*, *ENTPD1*, *ASH1L*, *QKI*; Table S2 and Table S4).

2.4 Discussion

In this study, we used primary fibroblast cell lines derived from bovine

control and LOS fetuses to determine imprinted domain chromosome architecture. Our results show that chromosome architecture at *IGF2R* imprinted domain is different between the parental alleles, a conformation disrupted in LOS. The alteration of the maternal sub-TAD in LOS was very likely associated with loss-of-methylation at the *IGF2R* ICR since similar changes were observed in the DC group, which served as the positive control for loss-of-methylation. We observed biallelic binding of CTCF protein to the predicted binding sites within *IGF2R* ICR in control and LOS fibroblasts, with increased maternal allele preference in LOS. Biallelic binding of CTCF has been reported in mouse embryos (Marcho et al., 2015).

From the 4C sequencing data, we were not able to conclude whether the CTCF binding sites in the 4C bait were involved in the formation of the observed sub-TADs at the *IGF2R* domain. For instance, each of these two binding sites may serve as the boundary for one of two neighboring sub-TADs. Alternatively, the boundaries of the observed sub-TAD could be defined by one of the predicted CTCF binding sites around 95.96 and 96.3 Mb, thus the binding site within the bait could be involved in formation of far-cis or trans contacts, a known function of CTCF (Ling et al., 2006). Although most of the far-cis and trans interacting regions did not colocalize with predicted CTCF binding sites, they still served as indicators for spatial closeness as the maintenance of chromosome architecture is a highly dynamic process (Nora et al., 2017).

Loss-of-imprinting is defined as biallelic silencing/expression of imprinted genes, a phenomenon correlated with loss or gain of methylation at their ICR (Verona et al., 2003). Methylation on the maternal allele of the *IGF2R* ICR prevents the expression of the long non-coding RNA (lncRNA) *AIRN* (Sleutels et al., 2002), which when expressed, silences *IGF2R* by attracting Polycomb repressive complexes to the locus (Schertzer et al., 2019). Consistent with this, we observed a ~3.5 fold decrease of *IGF2R* transcript in the LOS group which is associated with hypomethylation on the maternal allele's ICR. The extended sub-TAD on the maternal allele in LOS harbors six non-imprinted genes. It is possible that their regulatory regions (i.e. enhancer/silencers), alter the expression of *IGF2R*, as we previously showed biallelic *AIRN* in the muscle of day 105 bovine fetuses (Chen et al., 2017). Future studies will determine the validity of this hypothesis.

Initially we only expected to see alteration of chromosome architecture on the maternal allele of *IGF2R* ICR in LOS since this is the allele that suffers loss-of-methylation. As expected, some gain and loss of interactions were detected in LOS. Unexpectedly, the normally unmethylated paternal allele showed a larger number of altered far-cis and trans contacts in LOS than in controls, indicating changes of the three-dimensional shape of chromosome 9 and spatial changes within the nucleus. We could not find a reference on our search of the literature on this phenomenon, but this is an observation we intend

to follow up on as this can potentially extend the definition of loss-of-imprinting.

For the KvDMR1 domain, none of the imprinted genes (*CDKN1C*, *KCNQ1*, *PHLDA2*) were differentially expressed in LOS. Even though allelic comparison in controls identified hundreds of differentially interacting regions in cis, few allelic differences were detected between the DC and control groups, suggesting limited regulatory effects of DNA methylation at this locus. In mice, CTCF-driven chromosome architecture at KvDMR1 regulates *Cdkn1c* expression during myoblast differentiation, a regulation affected by DNA methylation (Battistelli et al., 2014). However, we did not detect CTCF binding at this locus perhaps indicating species-specific regulation of KvDMR1.

For the five misregulated genes reversely correlated with DMRs, four were protein-coding genes and one was a lncRNA. The downregulated *LOC535280*, also known as neuroblastoma suppressor of tumorigenicity 1-like (*NBL1/DAN*), is a tumor suppressor (Cui et al., 2016). The downregulated gene charged multivesicular body protein 6 (*CHMP6*) plays roles in plasma membrane receptor downregulation and recycling (Yorikawa et al., 2005). Downregulated *LOC101907348* is a lncRNA of unknown function. The downregulated gene ectonucleoside triphosphate diphosphohydrolase 1 (*ENTPD1*) has immunosuppression functions via degradation of adenosine triphosphate (Feng et al., 2011). Lastly, paternally expressed 10 (*PEG10*) is an imprinted gene that was upregulated in the LOS group. *PEG10* has roles in promoting cell

proliferation and its upregulation has been reported in BWS, Wilms tumor, and hepatoblastoma, which are two frequent tumors in BWS (Berland et al., 2021; Cairo et al., 2008; Jiménez Martín et al., 2021).

When analyzing clustering tendency of misregulated genes, a 200 kb distance was chosen based on the average sub-TAD size found in human (~185-208 kb (Montefiori et al., 2018)), analyses identified 149 clustered misregulated genes, many of which had the same direction of misregulation. Of these, two clusters involve detected DMRs at CTCF binding sites in LOS. This, could however be an underestimation of the effect of DNA methylation on CTCF regulation since the statistical methods utilized identified a DMR only when 10 contiguous CpG sites in a running window had altered methylation.

When we queried the flanking 100 kb of the IGF2R-associated altered contacts, we found that the paternal allele has association with eight dysregulated genes while the maternal allele has only one association with QKI. Several of these genes have been reported to be involved in development and tumorigenesis. For example, the downregulated matrix metalloproteinase 2 (*MMP2*) functions in cleaving extracellular matrix components and signal molecules (Bauvois, 2012). Mutation of *MMP2* in human has been reported in several syndromes showing muscle and bone malformation (Rouzier et al., 2006). The downregulated EPH receptor A5 (*EPHA5*) is the membrane-bound receptor for Ephrin and its downregulation has also been reported in

tumorigenesis (Li et al., 2015). The upregulated gene regulator of G protein signaling 17 (*RGS17*) is an oncogene and shows increased expression in several tumor types (James et al., 2009; Sokolov et al., 2011).

In summary, our study characterized allele specific chromosome architecture at *IGF2R* and *KVDMR1* imprinted domains. In addition, we determined clustering tendency of LOS misregulated genes indicating genome-wide location-based cause of misregulation. Importantly, architectural changes at *IGF2R* occurs in both the maternal and paternal alleles in LOS. We conclude that altered chromosome architecture is associated with LOS.

2.5 Limitations of Study

First, in this study, we used skin fibroblast primary cells to characterize chromosome architecture of control and LOS fetuses. Given that chromosome architecture could be tissue/cell-type specific (Fraser et al., 2015), the patterns identified in fibroblasts may not in its entirety recapitulate what happened in other cell types. Second, we identified changes of allele-specific chromosome architecture at *IGF2R* imprinted domain in LOS, which was coupled with alterations in DNA methylation level, CTCF binding, and *IGF2R* expression. However, we could not determine whether the changed architecture in LOS was involved in genomic imprinting regulation and/or whether the altered

chromosome contacts were the cause or result of DNA methylation defects.

Future studies would need to be done to clarify this point using genome editing tools. Third, with 4C sequencing, we could not determine whether there was only one or multiple sub-TADs included in the read enriched regions around the bait. Ongoing Hi-C studies will address this question.

2.6 Materials and methods

STAR METHODS

KEY RESOURCES TABLE

REAGENT RESOURCE	or	SOURCE	IDENTIFIER
Antibodies			
CTCF		Cell Signaling Technology	Cat#3418; RRID: AB_2086791
Histone H3		Cell Signaling Technology	Cat#4620; RRID: AB_1904005
Normal Rabbit IgG		Cell Signaling Technology	Cat#2729; RRID: AB_1031062
Bacterial and virus strains			
DH10B Cells	Competent	Thermo Scientific	EC0113
Biological samples			
Bovine fetal tissues		This study	N/A
Chemicals, peptides, and recombinant proteins			
DMEM		Gibco	11885084
Fetal bovine serum		Atlanta Biologicals	S11150H
Antibiotic-antimycotic		Gibco	15240062
HEPES		Sigma-Aldrich	H4034

0.05% trypsin-EDTA	Gibco	25300054
DMSO	Sigma-Aldrich	D2650
Decitabine	Sigma-Aldrich	A3656
Phenol:Chloroform:Isoamyl Alcohol	Sigma-Aldrich	P3803
TRIzol™ Reagent	Invitrogen	15596026
RQ1 RNase-Free DNase	Promega	M6101
random hexamers	Promega	C1181
formaldehyde	Electron Microscopy Sciences	157-4
T4 DNA Ligase	New England Biolabs	M0202L
Proteinase K	Fisher	BP1700
RNase A	Roche	10109142001
NlaIII	New England Biolabs	R0125S
Tsp45I	New England Biolabs	R0583S
MseI	New England Biolabs	R0525S
BsrI	New England Biolabs	R0527S
Critical commercial assays		
EZ DNA Methylation-Direct™ Kit	ZYMO RESEARCH	D5021
SuperScript® IV Reverse Transcriptase	Invitrogen	18090010
GoTaq® Flexi DNA Polymerase	Promega	M8295
Wizard® SV Gel and PCR Clean-Up System	Promega	A9282
CloneJET PCR Cloning Kit	Thermo Scientific	K1231
Qubit dsDNA HS Assay Kit	Invitrogen	Q32851
QIAquick PCR Purification Kit	QIAGEN	28104
Platinum Taq DNA	Invitrogen	11304-011

Polymerase High Fidelity		
AxyPrep MAG PCR Clean-Up Kit	Axygen	MAG-PCR-CL-5
NEBNext Library Quant Kit for Illumina	New England Biolabs	E7630S
SimpleChIP Enzymatic Chromatin IP Kit	Cell Signaling Technology	9003S
Deposited data		
RNA-seq, WGBS, 4C-seq, and DNA-seq	This study	GEO: GSE197130
Experimental models: Cell lines		
Bovine fetal fibroblast primary cells	This study	N/A
Oligonucleotides		
GE_KvDMR1_F1	Integrated DNA Technologies	5'-AATCCGATCGCAAGGGT
GE_KvDMR1_R1	Integrated DNA Technologies	5'-GCTTCTCGGTGAGGAGAG
GE_IGF2R_ICR_F	Integrated DNA Technologies	5'-GGGGGAGGGTCTTTAAGGTTG
GE_IGF2R_ICR_R	Integrated DNA Technologies	5'-TGGCTTTCAGGCTCCATAGAA
BI_KvDMR1_F	Integrated DNA Technologies	5'-GTGAGGAGTATGGTATTGAGG
BI_KvDMR1_R	Integrated DNA Technologies	5'-CCCCTACAACTATCCAATCAACT
4C_KvDMR1_F	Integrated DNA Technologies	5'-TACACGACGCTCTTCCGATCT/CTCAGCGCCCAGCTTAC
4C_KvDMR1_R	Integrated DNA Technologies	5'-CAGACGTGTGCTCTTCCGATCT/TCACGACTTGGCTCTTCTC
4C_IGF2R_ICR_F	Integrated DNA Technologies	5'-TACACGACGCTCTTCCGATCT/TTTAGGCGCGGAAGAACGAT
4C_IGF2R_ICR_R	Integrated DNA Technologies	5'-CAGACGTGTGCTCTTCCGATCT/GTGCGCACA GCCGCCAGAA
GE_KvDMR1_F2	Integrated DNA	5'-GCACACCGCTTTCCACACC

	Technologies	
GE_KvDMR1_R2	Integrated DNA Technologies	5'-GCACTGAGGTGACTGCGG
GE_IGF2R_INT3_F	Integrated DNA Technologies	5'-CTCTGGAGGGTTTCAGCGTC
GE_IGF2R_INT3_R	Integrated DNA Technologies	5'-AGGGAATACGCTTTCCCACG
Software and algorithms		
4Cker	Open source	https://github.com/rr1859/R.4Cker
BBMap	Open source	https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbmap-guide/
bedtools	Open source	https://bedtools.readthedocs.io/en/latest/
bismark	Open source	https://www.bioinformatics.babraham.ac.uk/projects/bismark/
BisSNP	Open source	https://github.com/dnaase/Bis-tools
bowtie2	Open source	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
bwa-mem2	Open source	https://github.com/bwa-mem2/bwa-mem2
circular	Open source	https://cran.r-project.org/web/packages/circular/index.html
CTCFBSDB2.0	Open source	https://insulatordb.uthsc.edu/
DAVID Bioinformatics Resources	LHR1	https://david.ncicrf.gov/
DESeq2	Open source	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
edgeR	Open source	https://bioconductor.org/packages/release/bioc/html/edgeR.html
fourSig	Open source	https://sourceforge.net/projects/foursig/
GATK	Open source	https://gatk.broadinstitute.org/hc/en-us
ggplot2	Open source	https://cran.r-project.org/web/packages/ggplot2/index.html
HISAT2	Open source	http://daehwankimlab.github.io/hisat2/
HTSeq	Open source	https://htseq.readthedocs.io/en/master/
hummingbird	Open source	https://www.bioconductor.org/packages/release/bioc/html/hummingbird.html
ImageJ	NIH	https://imagej.nih.gov/ij/
Integrative Genomics Viewer	Open source	https://software.broadinstitute.org/software/igv/
JASPAR2020	Open source	https://bioconductor.org/packages/release/data/annotation/html/JASPAR2020.html

picard	Open source	https://broadinstitute.github.io/picard/
Samtools	Open source	http://www.htslib.org/
StringTie	Open source	https://ccb.jhu.edu/software/stringtie/
Sushi	Open source	https://bioconductor.org/packages/release/bioc/html/Sushi.html
TFBSTools	Open source	https://bioconductor.org/packages/release/bioc/html/TFBSTools.html
trimmomatic	Open source	http://www.usadellab.org/cms/?page=trimmomatic
Other		
Custom code	This Study	Zenodo: https://doi.org/10.5281/zenodo.6449167

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Rocío Melissa Rivera (riverarm@missouri.edu).

Materials availability

This study did not generate new unique reagents. Commercially available reagents are indicated in the key resources table.

Data and code availability

- All raw data of RNA-seq, WGBS, 4C-seq, and DNA-seq reported in this paper are publicly available in the GEO database with accession numbers

GSE197130 as of the date of publication. Original gel images reported in this paper are available from the lead contact upon request.

- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- All additional information required to reanalyze the data reported in this paper are available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal tissues

Day 105 *Bos taurus indicus* (*B. t. indicus*; Brahman breed) x *Bos taurus taurus* (*B. t. taurus*; Angus breed) F1 hybrid fetuses were generated by our laboratory in 2019 as reported before (Chen et al., 2013; Rivera and Hansen, 2001) and used as tissue donors. This breeding strategy aimed to introduce genetic variants, including single-nucleotide polymorphisms (SNPs), to differentiate parental alleles. The control group was generated using artificial insemination (AI) and the ART group was generated by *in vitro* production procedures. The LOS group was defined as individuals from the ART group with body weight greater than 97th centile of controls. On day 105 of pregnancy, conceptuses were collected by caesarean section to maintain nucleic acid integrity. The identifier, original ID, sex, and body weight of fetuses used in this

study were as follows; 1) control fetuses: Control#1 (original ID 533, female, 388g), Control#2 (647, female, 396g), Control#3 (640, male, 448g), and Control#4 (648, male, 466g); 2) LOS fetuses: LOS#1 (656, female, 704g), LOS#2 (602, male, 752g), LOS#3 (604B, female, 986g), and LOS#4 (664, male, 1080g).

All the animal procedures were approved by University of Missouri Animal Care and Use Committee under protocol 9455. Trained personnel and Veterinarians performed all animal handling and surgeries.

Establishment of skin fibroblast primary cell line, cell culture, and Decitabine treatment

Fibroblast cells were chosen since they originate from mesoderm and share lineage with skeletal muscle and kidney (Chan et al., 2016; Davidson et al., 2019a; Vodyanik et al., 2010). Muscle mass is a main contributor to the increased birth weight in LOS, and Wilms tumor of the kidney and rhabdomyosarcoma are tumors observed in BWS (Brioude et al., 2018; Li et al., 2019b). We reasoned that the molecular aberrations in skeletal muscle and kidney would be conserved in fibroblast if they occurred during early embryo development. We showed this to be the case for expression of *IGF2R* and DNA methylation at KvDMR1 in this study. Further, using fibroblast cells will allow comparison of findings with the human counterpart syndrome BWS, since skin

fibroblasts is one of the frequently obtained tissue samples from these patients and is being used to characterize the syndrome (Brioude et al., 2018; Naveh et al., 2021).

Fetal skins were collected to establish fibroblast primary cell line using protocol adapted from https://animal.ifas.ufl.edu/hansen/lab_protocol_docs/bovine_fetal_fibroblasts.pdf (Dobbs et al., 2013). Briefly, approximately one square centimeter piece of skin was collected from each fetus during fetal collection and incubated in 1 ml fresh bovine embryonic fibroblast medium (BEF; 89% (v/v) DMEM (Gibco, 11885084), 10% (v/v) fetal bovine serum (FBS; Atlanta Biologicals, S11150H), and 1% (v/v) Antibiotic-antimycotic (10000 units/ml penicillin, 10000 ug/ml streptomycin, and 25 ug/ml Gibco Amphotericin B; Gibco, 15240062)) containing 25mM HEPES (Sigma-Aldrich, H4034) at 38.5°C until further processing within 12 hours. The skin pieces were washed in homemade DPBS containing 1% (v/v) antibiotic-antimycotic three times, diced into smaller pieces, and transferred into a well of a 12 well plate containing 1 ml BEF medium. The skin pieces were cultured at 38.5°C (body temperature of cattle) with 5% CO₂ and 1 ml fresh BEF medium was added every two days. When outgrowing fibroblast cells reached confluency, the skin pieces were removed, and cells were trypsinized with 0.05% trypsin-EDTA (Gibco, 25300054) and transferred to T75 flasks (MIDSCI, TP90076) with 12 ml BEF medium. The medium was changed every two days until cells reached

80-90% confluency and were cryopreserved to keep them at low passage number (i.e., $n < 10$). For cryopreservation, cells were trypsinized, counted with a hemacytometer, centrifuged at 250 x g, resuspended in BEF medium containing 10% (v/v) DMSO (Sigma-Aldrich, D2650) at a concentration of ~1-2 million cells/ml, and transferred to a cryotube (MIDSCI, CM-4). The cryotubes were kept at -80°C overnight and then transferred into liquid nitrogen for long term storage. When recovering from cryopreservation, the cells were thawed at 38.5°C, centrifuged and washed with BEF medium to remove remaining DMSO, and cultured as described above.

To induce loss of DNA methylation, a subgroup of fibroblasts derived from control fetuses were treated with 0.5 uM Decitabine (5-aza-2'-deoxycytidine) for 96 hours. BEF medium with 0.5 uM Decitabine were changed every 24 hours to maintain the concentration of Decitabine. The identifier of Decitabine treated control samples are DC#1 to DC#4 corresponding to Control#1 to Control#4.

METHOD DETAILS

All the chromosomal coordinates in this manuscript refer to bovine genome assembly ARS-UCD1.2 unless otherwise specified (Rosen et al., 2020).

Genomic DNA extraction

Fibroblasts, semen of the sire of the fetuses (JDH MR. MANSO 7 860958 154BR599 11200 EBS/INC CSS 2), or fetal tissue samples were lysed in lysis buffer (0.05 M Tris-HCl (pH 8.0), 0.1 M EDTA, and 0.5% (w/v) SDS) with proteinase K (Fisher BioReagents, BP1700) at 55°C for four hours (cells and semen) or overnight (tissue). Genomic DNA was extracted with Phenol:Chloroform:Isoamyl Alcohol (SIGMA, P3803) following the manufacturer's instructions. The concentration of DNA was measured by using a NanoDrop® ND-1000 Spectrophotometer (Thermo Fisher Scientific) and DNA integrity was confirmed by electrophoresis on a 0.7% agarose gel. Genomic DNA samples were stored at -20°C.

Bisulfite conversion of genomic DNA

Genomic DNA was bisulfite converted with EZ DNA Methylation-Direct™ Kit (ZYMO RESEARCH, D5021) following the manufacturer's instructions. Bisulfite converted DNA samples were stored at -80°C.

RNA isolation

Total RNA from cultured fibroblast cells and tissue samples was isolated using TRIzol™ Reagent (Invitrogen, 15596026) following the manufacturer's

instructions. The concentration of RNA samples was measured by using the NanoDrop® ND-1000 Spectrophotometer. RNA samples were stored at -80°C.

Reverse transcription of mRNAs

Total RNA samples were treated with RQ1 Rnase-Free Dnase (Promega, M6101) following the manufacturer's instructions to remove genomic DNA contamination. cDNA was synthesized using SuperScript® IV Reverse Transcriptase (Invitrogen, 18090010) with random hexamers (Promega, C1181) following the manufacturer's instructions. cDNA samples were stored at -20°C.

Polymerase chain reaction (PCR), molecular cloning, and Sanger sequencing

GoTaq® Flexi DNA Polymerase (Promega, M8295) was used for end-point PCR following the manufacturer's instructions. Genomic primer GE_KvDMR1_F1 (5'-AATCCGATCGCAAGGGT, Chr29: 48907972-48907988) and GE_KvDMR1_R1 (5'-GCTTCTCGGTGAGGAGAG, Chr29: 48908541-48908558) were used to amplify KvDMR1 region to identify SNP, and the thermocycler conditions were: 95°C 2min; 95°C 30s, touchdown from 71.8°C to 61.8°C by 1°C per cycle 30s, 72°C 38s, 40 cycles; 72°C 5min. Genomic primer GE_IGF2R_ICR_F (5'-GGGGGAGGGTCTTTAAGGTTG, Chr9: 96223334-

96223354) and GE_IGF2R_ICR_R (5'-TGGCTTTCAGGCTCCATAGAA, Chr9: 96223732-96223752) were used to amplify *IGF2R* ICR to identify SNP, and the thermocycler conditions were: 95°C 2min; 95°C 30s, 64°C 30s, 72°C 30s, 35 cycles; 72°C 5min.

Bisulfite primer BI_KvDMR1_F (5'-GTGAGGAGTATGGTATTGAGG, Chr29: 48908486-48908506) and BI_KvDMR1_R (5'-CCCCTACAAACTATCCAATCAACT, Chr29: 48908205-48908229) were used to amplify KvDMR1 region to determine DNA methylation status, and the thermocycler conditions were: 95°C 2min; 95°C 30s, 58.7°C 30s, 72°C 30s, 35 cycles; 72°C 5min.

The PCR products were resolved on a 2% (w/v) agarose gel and visualized using ethidium bromide. Bands of expected sizes were cut and DNA was retrieved from the gel using Wizard® SV Gel and PCR Clean-Up System (Promega, A9282). Sanger sequencing for the retrieved DNA was performed at the University of Missouri Genomics Technology Core.

Molecular cloning was performed using CloneJET PCR Cloning Kit (Thermo Scientific, K1231) and DH10B Competent Cells (Thermo Scientific, EC0113) to determine allelic DNA methylation level or allelic CTCF ChIP enrichment following the manufacturer's instructions.

Quantitative Reverse-transcriptase PCR for *IGF2R*

Quantitative RT-PCR of *IGF2R* was done using TaqMan® probes (ThermoFisher Scientific, Hanover Park, IL) and a QuantStudio 3 Real-Time PCR System (Applied BioSystems, Waltham, Massachusetts). The mRNA level of each target transcript was normalized to the geometric mean of three endogenous normalizers, namely *NUCKS1*, *RBM39*, *SF3B1*. Amplifications were performed in at least duplicates. Each group's cycle threshold difference and $2^{-\Delta\Delta C_t}$ was calculated to determine the fold difference in transcript levels.

CTCF binding site prediction in bovine genome

Potential CTCF binding sites were predicted globally for bovine genome assembly ARS-UCD1.2 (Rosen et al., 2020) using TFBSTools 1.26.0 (Tan and Lenhard, 2016) with database JASPAR2020 (Fornes et al., 2020). CTCF motifs of 'vertebrates' were used for prediction and min.score was set to 90%. CTCFBSDB2.0 (Ziebarth et al., 2012) was also used for CTCF binding sites prediction in local regions.

Circular chromosome conformation capture (4C) library preparation and sequencing

The 4C library preparation procedure mainly followed the Krijger protocol

(Krijger et al., 2020) with some adaptations from other published protocols (Gheldof et al., 2012; Splinter et al., 2012; van de Werken et al., 2012). Briefly, for each sample, six million fibroblast cells were fixed with 2% (v/v) formaldehyde (Electron Microscopy Sciences, 157-4) at a concentration of two million cells per ml. Fixed cells were lysed, washed, and underwent first restriction enzyme (RE) digestion overnight as described in Krijger protocol (Krijger et al., 2020). Note: the specifics of the REs are explained below. On day two, after confirming good digestion efficiency by 0.7% agarose gel electrophoresis (downwards shift of the DNA smear), the first RE was inactivated according to manufacturer's instructions and samples were diluted and ligated with T4 DNA Ligase (New England Biolabs, M0202L) overnight. On day three, after confirming good ligation efficiency by 0.7% agarose gel electrophoresis (upwards shift of the DNA smear), the samples were treated with Proteinase K (Fisher, BP1700) overnight to reverse the formaldehyde cross-links between protein and DNA. On day four, samples were treated with Rnase A (Roche, 10109142001) and ligated DNA was extracted using Phenol-Chloroform. Next, ligated DNA samples were digested with the second RE overnight. On day five, after confirming good digestion efficiency by electrophoresis, the second RE was inactivated by heating or removed by Phenol-Chloroform extraction if heat insensitive. DNA concentration was measured by Qubit dsDNA HS Assay Kit (Invitrogen, Q32851) and these DNA samples were ligated with T4 DNA Ligase overnight at a concentration of 5

ng/ul. On day six, after confirming good ligation efficiency by electrophoresis, DNA was ethanol precipitated and purified with QIAquick PCR Purification Kit (QIAGEN, 28104). Concentration of the purified DNA, in other words 4C template, was measured by Qubit assay. The 4C templates were stored at -20°C.

Primers for the two PCR steps were designed as described in Krijger protocol (Krijger et al., 2020). Platinum Taq DNA Polymerase High Fidelity (Invitrogen, 11304-011) was used for 4C PCR following the manufacturer's instructions. Following each PCR step, the products were purified with AxyPrep MAG PCR Clean-Up Kit (Axygen, MAG-PCR-CL-5) to remove remaining primers, primer dimers, and self-ligation bands. 4C libraries were stored at -20°C. Concentration of the 4C libraries were measured by Qubit assay and the integrity of 4C libraries were confirmed by NEBNext Library Quant Kit for Illumina (New England Biolabs, E7630S). Average DNA fragment size of 4C libraries were measured by fragment analyses using Fragment Analyzer Systems (Agilent) at University of Missouri Genomics Technology Core. Molar concentration of 4C libraries were calculated based on Qubit concentration and average DNA fragment size, and equal molar amount of 4C libraries were pooled. A 15% spike-in of PhiX Control v3 (Illumina, FC-110-3001) was included in the final library to increase base diversity and improve color balance (proportions and distribution of dyes used to report different nucleotides) during sequencing for the bait (4C target region in the genome) sequence. The final library was sequenced on the

NovaSeq platform for 250bp paired-end reads at University of Missouri

Genomics Technology Core.

KvDMR1 and *IGF2R* ICR were selected for 4C sequencing. Fibroblast cells of Control#1 to Control#4 and LOS#3, and Decitabine treated fibroblasts DC#1 to DC#4 were used for KvDMR1. For *IGF2R* ICR, all the samples Control#1 to #4, LOS#1 to #4, and DC#1 to #4 were used.

In total, three 4C assays were designed, one for KvDMR1 and two for *IGF2R* ICR. For each region, the 4C bait contains predicted CTCF binding sites and a SNP. In order to include the SNP in the sequencing reads, one of the two restriction enzyme (RE) digestion sites has to be adjacent to the SNP which limited the choice of RE. For KvDMR1, NlaIII (New England Biolabs, R0125S) and Tsp45I (New England Biolabs, R0583S) were selected for the first and second RE digestion, respectively. For *IGF2R* ICR, two different RE were used as the second RE, which resulted in two 4C assays, namely IGF2R_MseI and IGF2R_BsrI. NlaIII (first RE) and MseI (New England Biolabs, R0525S) were used for IGF2R_MseI, and NlaIII (first RE) and BsrI (New England Biolabs, R0527S) were used for IGF2R_BsrI. For the first round of PCR, primer 4C_KvDMR1_F (5'- TACACGACGCTCTTCCGATCT/CTCAGCGCCCAGCTTAC, Chr29: 48907934-48907950, '/' indicates the split site where sequences on the left side are complementary to Illumina_i5 or i7 primers, and sequences on the right side are complementary to the genome) and 4C_KvDMR1_R (5'-

CAGACGTGTGCTCTTCCGATCT/TCACGACTTGGCTCTTCTC, Chr29:

48908379-48908397) were used for KvDMR1, and the thermal conditions were:

94°C 2min; 94°C 15s, 65.3°C 1min, 68°C 3min, 16 cycles; 68°C 5min. Primer

4C_IGF2R_ICR_F (5'-

TACACGACGCTCTTCCGATCT/TTTAGGCGCGGAAGAACGAT, Chr9:

96223648-96223667) and 4C_IGF2R_ICR_R (5'-

CAGACGTGTGCTCTTCCGATCT/GTGCGCACAGCCGCCAGAA, Chr9:

96223397-96223415) were used for IGF2R_MseI and IGF2R_BsrI, and the

thermal conditions were: 94°C 2min; 94°C 15s, 62.1°C 1min, 68°C 3min, 16

cycles; 68°C 5min. For the second round of PCR, 17 pairs of Illumina_i5 (5'-

AATGATACGGCGACCACCGAGATCTACAC-index-

ACACTCTTTCCCTACACGACGCTCTTCCGATCT) and Illumina_i7 (5'-

CAAGCAGAAGACGGCATACGAGAT-index-

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT) primers were designed

with different index sequences (UDI0001-UDI0017 in the official manual 'Illumina

Adapter Sequences v16') and assigned to different samples. For IGF2R_MseI

and IGF2R_BsrI assays, UDI0001-12 were used for control #2, #4, #1, #3, LOS

#2, #1, #4, #3, DC #2, #4, #1, and #3, respectively. For KvDMR1 assay,

UDI0001, 4, 7, 10, 13-17 were used for control #2, #4, LOS #3, DC #2, control

#1, #3, DC #1, #3, and #4, respectively. For some indexes, the three samples

from different 4C assays that shared the index can be separated based on bait

sequence in the reads, which served as a secondary barcode. The thermal conditions for Illumina_i5 and Illumina_i7 were: 94°C 2min; 94°C 15s, 65.5°C 1min, 68°C 3min, 20 cycles; 68°C 5min. These designs resulted in three 4C datasets after sequencing.

4C data analyses

Raw sequencing reads were first sorted into the three 4C datasets (KvDMR1, IGF2R_MseI, and IGF2R_BsrI) based on the bait sequences. This was accomplished by aligning the bait part of the reads to the bait sequences using bowtie2 (Langmead and Salzberg, 2012) and filtering for unique alignments. Next, the parental alleles were assigned to each read pairs based on the SNP in the bait sequences. Then the bait sequences were removed from the reads and reads were trimmed for adapter sequences and low quality bases using trimmomatic 0.39 (Bolger et al., 2014) with parameters 'ILLUMINACLIP:adapter_seq:2:30:10:1:true LEADING:20 TRAILING:20 AVGQUAL:20 MAXINFO:0:0.5'. The trimmed reads were split into multiple fragments by the recognition sites of corresponding first and second RE used. The first fragment on the 5' side of each read was kept and read pairs were combined and aligned to genome as single-end reads using bowtie2 with parameters '-N 1 -L 15 -no-unal'. Aligned reads with mapping quality less than 20

were filtered out using Samtools 1.13 (Li et al., 2009).

The genome was fragmented by NlaIII recognition site from hereunto referred as NlaIII restriction fragments. Read coverage was calculated for NlaIII restriction fragments using bedtools 2.30.0 with parameters 'coverage -sorted -counts' (Quinlan and Hall, 2010). Self-ligation reads aligned to the NlaIII restriction fragments covering the bait region and one up/downstream fragment were excluded for statistical analyses. 4C peaks were called using fourSig with parameters 'cis.only=FALSE, window.size=50, iterations=1000, fdr=0.01, fdr.prob=0.05, only.mappable=FALSE' (Williams Jr et al., 2014). fourSig reported 4C peaks for individual samples were converted to NlaIII restriction fragments, and the restriction fragments present in at least two samples in a group were reported as the 4C peaks for the group.

Statistical comparison between groups to identify change of 4C interactions were conducted using a running window approach which has been widely used for 4C analyses (Simonis et al., 2006; van de Werken et al., 2012). The sum read coverage of a running window of 50 NlaIII restriction fragments were calculated and used as input for statistical comparison by DESeq2 (Love et al., 2014). A prefilter for low coverage window using 'rowSums(counts>2) >= 2' were conducted to make the input list manageable. Default settings of DESeq2 were used for other steps. False discovery rate (FDR) was controlled at 0.05 for DESeq2 results (labelled as Padj) by the Benjamini-Hochberg method. The

coordinates of the 13th restriction fragment in each running window were used to indicate that window. Overlapped and continuous significant window with the same direction of changes were merged. In addition, statistical analyses for read enrichment and differences between groups were also conducted with a second software 4Cker as corroboration (Raviram et al., 2016). Parameters used for 4Cker functions are k=5 for 'nearBaitAnalysis', k=10 for 'cisAnalysis', and pval=0.05 for 'differentialAnalysis' as recommended by the manual. Trans analyses were not performed with 4Cker according to the software manual recommendation and also due to very low efficiency in processing multiple samples as a group.

Genomic content related analyses and permutation test

Information of gene annotation was obtained from NCBI (GCF_002263795.1_ARC-UCD1.2_genomic.gff) (O'Leary et al., 2016). Per million base gene density was calculated based on the annotation. Repeated and overlapped exons were merged for each gene, and introns were calculated based on merged exons. Promoters (1kb) were calculated based on transcription start sites annotation and only included protein coding genes and long non-coding RNAs. Annotation of CpG islands and repeated sequences were obtained from UCSC Genome Browser (Kent et al., 2002). Locations of CpG shores

(flanking 2kb from CpG islands) and shelves (flanking 2-4kb from the CpG island) were calculated based on CpG island annotation. Bedtools and custom Perl scripts were used for permutation test and identify overlapped genomic location and make tables (Quinlan and Hall, 2010). R package Sushi, circular, and ggplot2 were used for making figures (Lund et al., 2017; Phanstiel et al., 2014; Wickham, 2011). Integrative Genomics Viewer was also used for visualization (Robinson et al., 2011).

Bull semen genomic sequencing and data analyses

Genomic sequencing for semen DNA of the bull used to sire all the fetuses in this study was conducted by University of Missouri Genomics Technology Core. Information on library preparation and sequencing obtained from the Core is as follows: The library was constructed following the manufacturer's protocol with reagents supplied in Illumina's TruSeq DNA PCR-Free sample preparation kit (#FC-121-3001). Briefly, DNA was sheared using standard Covaris methods to generate average fragmented sizes of 350 bp. The resulting 3' and 5' overhangs were converted to blunt ends by an end repair reaction which uses a 3' to 5' exonuclease activity and polymerase activity. A single adenosine nucleotide was added to the 3' ends of the blunt fragment followed by the ligation of Illumina indexed paired-end adapters. The adaptor

ligated library was purified twice with AxyPrep Mag purification beads. The purified library was quantified using KAPA library quantification kit (KK4824) and library fragment size confirmed by Fragment Analyzer (Agilent Technologies, Inc.). Libraries were diluted and sequenced according to Illumina's standard sequencing protocol for the NovaSeq 6000.

For genomic sequencing data analyses, we followed the pipeline for 1000 bull genome project (Hayes and Daetwyler, 2019). Briefly, raw sequencing reads were trimmed for adapter sequences and low quality bases using trimmomatic 0.39 (Bolger et al., 2014) with parameters 'ILLUMINACLIP:adapter_seq:2:30:10:1:true LEADING:20 TRAILING:20 AVGQUAL:20 SLIDINGWINDOW:3:15'. Trimmed reads were aligned to the bovine genome using bwa-mem2 (Vasimuddin et al., 2019) with default parameters. Samtools 1.13 (Li et al., 2009) was used to convert, sort, filter, and index bam files. Aligned reads with mapping quality (MAPQ) less than 20 were excluded from downstream analyses. Read groups were added using AddOrReplaceReadGroups function of picard 2.25.5 (Broad Institute, 2021). The dataset of known variants in bovine was acquired from the 1000 bull genome project, namely ARS1.2PlusY_BQSR_v3.vcf.gz. GATK 4.2.1.0 (Van der Auwera and O'Connor, 2020) was used to recalibrate base quality and identify variants in the genomic sequencing data with the known variant dataset as reference.

Parameters used for BaseRecalibrator and HaplotypeCaller were '-bqsr-baq-

gap-open-penalty 45' and '-pcr-indel-model NONE', respectively. Raw variants were scored using 2D model of CNNScoreVariants function of GATK with parameter '-tensor-type read_tensor'. Scored variants were filtered using FilterVariantTranches function of GATK with parameter '-info-key CNN_2D -invalidate-previous-filters -snp-tranche 99.95 -indel-tranche 99.4'.

Whole genome bisulfite sequencing (WGBS) and data analyses

WGBS for cultured fibroblast cells was conducted by CD Genomics. Information on library preparation and sequencing obtained from the company is as follows: For WGBS library preparation, 1 ug of genomic DNA was fragmented by sonication to a mean size of approximately 200-400 bp. Fragmented DNA was end-repaired, 5'-phosphorylated, 3'-dA-tailed and then ligated to methylated adapters. The methylated adapter-ligated DNAs were purified using 0.8x Agencourt AMPure XP magnetic beads and subjected to bisulfite conversion by ZYMO EZ DNA Methylation-Gold Kit (zyzo). The converted DNAs were then amplified using 25 µl KAPA HiFi HotStart Uracil+ ReadyMix (2X) and 8-bp index primers with a final concentration of 1 µM each. The constructed WGBS libraries were then analyzed by Agilent 2100 Bioanalyzer and quantified by a Qubit fluorometer with Quant-iT dsDNA HS Assay Kit (Invitrogen), and finally sequenced on Illumina HiSeq X ten sequencer. 0.1-1% lambda DNA were added

during the library preparation to monitor bisulfite conversion rate.

For WGBS data analyses, duplicated reads generated during PCR and sequencing were removed from raw sequencing reads using the clumpify function of BMap 38.90 (Bushnell, 2021). The remaining raw reads were trimmed for adapter sequences and low quality bases using trimmomatic 0.39 (Bolger et al., 2014) with parameters 'ILLUMINACLIP:adapter_seq:2:30:10:1:true LEADING:20 TRAILING:20 AVGQUAL:20 MAXINFO:0:0.5'. Trimmed reads were aligned to the bovine genome using bismark 0.23.0 (Krueger and Andrews, 2011) with parameters '-X 900 -unmapped -ambiguous -non_bs_mm'. Trimmed reads were also aligned to lambda phage genome to determine bisulfite conversion rates. Samtools 1.13 (Li et al., 2009) was used to convert, sort, filter, and index bam files. MarkDuplicates function of picard 2.25.5 (Broad Institute, 2021) was used to further remove duplicated reads after alignment. Read groups were added for each samples using AddOrReplaceReadGroups function of picard. Variants identified in bull semen genomic sequencing data and the previously mentioned variants acquired from the 1000 bull genome project served as known variants to identify genomic variants in WGBS data. Indel realignment was performed using RealignerTargetCreator and IndelRealigner functions of BisSNP 1.0.1 (Liu et al., 2012). Base quality recalibration was carried out using BisulfiteCountCovariates and BisulfiteTableRecalibration functions of BisSNP 0.82.2 since these functions are missing in version 1.0.0 and 1.0.1. Parameters

used for BisulfiteCountCovariates were '-cov ReadGroupCovariate -cov QualityScoreCovariate -cov CycleCovariate -baqGOP 30'. Genomic variants were identified using BisSNP 1.0.1 with default setting expect that '-bsRate' was changed to bisulfite conversion rate observed from lambda phage genome alignment for each sample. BisSNP identified variants were filtered by its VCFpostprocess function with parameter '-windSizeForSNPfilter 0'. Additionally, genomic variants were identified using BS-SNPer 1.0 (Gao et al., 2015) with parameters '-minhetfreq 0.1 -minhomfreq 0.85 -minquali 15 -mincover 5 -maxcover 1000 -minread2 2 -errorate 0.02 -mapvalue 20'. M-bias plots were generated using bismark and the first 3 bases of R1 reads and the first 4 bases of R2 reads showed biased CpG methylation level, thus these bases were excluded from downstream analyses. CpG methylation information were extracted from the bam files using bismark_methylation_extractor function of bismark with parameters '-p -ignore 3 -ignore_r2 4 -comprehensive -no_header -gzip -bedGraph -buffer_size 50% --cytosine_report'. Statistical analyses were conducted using R package hummingbird (Ji, 2019) with parameter 'minCpGs = 10, minLength = 100, maxGap = 300' to identify differentially methylated regions (DMRs) between LOS and Control groups. DMRs with at least 15% difference in methylation level (both gain and loss of methylation) and at least 2 mean read coverage at CpG sites were reported. The sex chromosomes were not analyzed to circumvent confounding created by X chromosome inactivation associated

DNA methylation.

Chromatin immunoprecipitation (ChIP) for CTCF protein

ChIP for CTCF protein was conducted in fibroblasts derived from control and LOS fetuses to verify *in silico* predicted CTCF binding site within the region used as bait for the 4C assay. SimpleChIP Enzymatic Chromatin IP Kit (Cell Signaling Technology, 9003S) and CTCF (D31H2) XP Rabbit mAb (Cell Signaling Technology, 3418S) were used for this experiment following the manufacturer's instructions. Briefly, three ChIP assays with different antibodies (CTCF, Histone H3 [positive control (Cell Signaling Technology, 4620S)], and Normal Rabbit IgG [negative control (Cell Signaling Technology, 2729S)]) were conducted for each sample. For each sample, 12 million fibroblast cells, equivalent to 4 million per ChIP assay, were fixed with 1% (v/v) formaldehyde (Electron Microscopy Sciences, 157-4) at a concentration of 0.5 million cells per ml. Fixed cells were washed, lysed, digested with 0.5 ul of micrococcal nuclease, and sonicated to break nuclear membrane as described in the manual. Once good digestion efficiency (about 150-900 bp DNA fragments, equivalent to 1-5 nucleosomes) was confirmed by 0.7% agarose gel electrophoresis, ChIP buffer and antibodies of recommended amount (1:50 dilution for CTCF and Histone H3 and 1.5 ug for Rabbit IgG) were added into each sample and incubated overnight at 4°C. On

day two, ChIP-Grade Protein G Magnetic Beads were used to pull down antibodies and bound protein and DNA. After washing, elution, and reversing crosslink, the DNA was eventually purified with spin columns.

To confirm *in silico* predicted CTCF binding sites, genomic primers were designed to amplify a short region covering the binding sites. The size of amplicon used in this study was longer than recommended by the manual since we included the SNPs in the amplicon to differentiate parental alleles. For *IGF2R* ICR, primer GE_IGF2R_ICR_F and GE_IGF2R_ICR_R were used and the thermal conditions were: 95°C 2min; 95°C 30s, 64°C 30s (0.3°C/s ramp temperature), 72°C 30s, 35 cycles; 72°C 5min. As allelic bias during PCR is possible, we also report results using 1M betaine (to relax secondary structures; Sigma B2629). For KvDMR1, primer GE_KvDMR1_F2 (5'-GCACACCGCTTTCCACACC, Chr29: 48908151-48908169) and GE_KvDMR1_R2 (5'-GCACTGAGGTGACTGCGG, Chr29: 48908477-48908494) were used and the thermal conditions were: 95°C 2min; 95°C 30s, 67.3°C 30s, 72°C 30s, 35 cycles; 72°C 5min. In addition, primer GE_IGF2R_INT3_F (5'-CTCTGGAGGGTTTCAGCGTC, Chr9: 96229536-96229555) and GE_IGF2R_INT3_R (5'-AGGGAATACGCTTTCCCACG, Chr9: 96229935-96229954) were used to amplify a region of *IGF2R*'s intron 3 which contained no predicted CTCF binding site to set background levels for traveling CTCF, and the thermal conditions were: 95°C 2min; 95°C 30s, 64°C 30s, 72°C 30s, 35 cycles;

72°C 5min. PCR amplicons were visualized on 7% acrylamide gel and the intensity of bands were measured using ImageJ (Schneider et al., 2012). T-test was conducted using online T-Test Calculator (<https://www.socscistatistics.com/tests/studentttest/default2.aspx>).

RNA sequencing (RNA-seq) and data analyses

RNA-seq for cultured fibroblast cells was conducted by BGI. Information on library preparation and sequencing obtained from the company is as follows: mRNA molecules were purified from total RNA using oligo (dT)-attached magnetic beads. mRNA molecules were fragmented into small pieces using fragmentation reagent after reaction a certain period in proper temperature. First strand cDNA was generated using random hexamer primed reverse transcription, followed by a second strand cDNA synthesis. The synthesized cDNA was subjected to end repair and then was 3' adenylated. Adapters were ligated to the ends of these 3' adenylated cDNA fragments. PCR was used to amplify the cDNA fragments with adapters from previous step. PCR products were purified with Ampure XP Beads (AGENCOURT) and dissolved in EB solution. Library was validated on the Agilent Technologies 2100 bioanalyzer. The double stranded PCR products were heat denatured and circularized by the splint oligo sequence. The single strand circle DNA (ssCir DNA) were formatted as the final library. The

library was amplified with phi29 to make DNA nanoball (DNB) which had more than 300 copies of one molecular. The DNBs were load into the patterned nanoarray and pair end 100 bases reads were generated in the way of combinatorial Probe Anchor Synthesis (cPAS).

Reads were aligned to the *Bos taurus* reference genome, ARS-UCD1.2, using HISAT2 with the `-dta` flag to allow for downstream transcript assembly (Pertea et al., 2016). Reads aligned to the genome were assembled into transcripts using StringTie and all transcripts merged (Pertea et al., 2016). Transcript abundance was estimated using HTSeq with the following flags – `order=pos`, `--idattr=gene`, and `--stranded=no` (Anders et al., 2015). To note that the gene symbols were in accordance with “gene_id” column instead of “gene” column in GCF_002263795.1_ARC-UCD1.2_genomic.gtf file (from NCBI) since the former differentiates repeated genes from different genomic location by adding a “_X” tag. Statistical comparison between Control and LOS groups were conducted using DESeq2. A prefilter for low abundance genes were conducted using `'rowSums(cpm(counts)>0.2) >= 3'` which resulted in 15,042 identified as expressed in this study (Robinson et al., 2010). Default settings of DESeq2 were used for other steps. FDR was controlled at 0.05. DE genes enriched signaling pathways were identified using DAVID Bioinformatics Resources 6.8 (Huang et al., 2009). Per million base gene expression level for control and LOS groups were calculated as the sum of the group mean CPM of all genes detected.

QUANTIFICATION AND STATISTICAL ANALYSIS

A two-tailed t-test was used for ChIP related comparisons and p less than 0.05 was considered as significant. fourSig was used to identify significant 4C contacts with parameters 'cis.only=FALSE, window.size=50, iterations=1000, fdr=0.01, fdr.prob=0.05, only.mappable=FALSE' (Williams Jr et al., 2014). DESeq2, which performs a Wald test, was used to detect significant differences in 4C contacts and gene expression with false discovery rate controlled at 0.05 using the Benjamini-Hochberg method (Love et al., 2014). 4Cker was also used to identify significant 4C contacts and significant differences in 4C contacts with parameters k=5 for 'nearBaitAnalysis', k=10 for 'cisAnalysis', and pval=0.05 for 'differentialAnalysis' (Raviram et al., 2016). Hummingbird, which is based on a Bayesian Hidden Markov Model, was used to identify significantly differentially methylated regions with parameter 'minCpGs = 10, minLength = 100, maxGap = 300' (Ji, 2019). Signaling pathway analyses were conducted using DAVID Bioinformatics Resources 6.8 (Huang et al., 2009). For permutation tests (10,000 shuffles), the p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|) / 10000$.

2.7 Acknowledgements

This work was supported by Agriculture and Food Research Initiative

(grant AFRI - 2018-67015-27598); National Science Foundation (grant IOS1545780); Genomics Technology Core, University of Missouri (MUGTC, Tier1 Block Grant); College of Agriculture, Food & Natural Resources, University of Missouri (Dr. Roger L. Morrison Scholarship to Y.L.); and the William and Nancy Thompson professorship (J.C.). We thank Dr. Robert Schnabel for assistance with the DNA sequencing analyses, Astrid Roshealy Brau for *IGF2R* qRT-PCR, and MUGTC staff for 4C sequencing assistance.

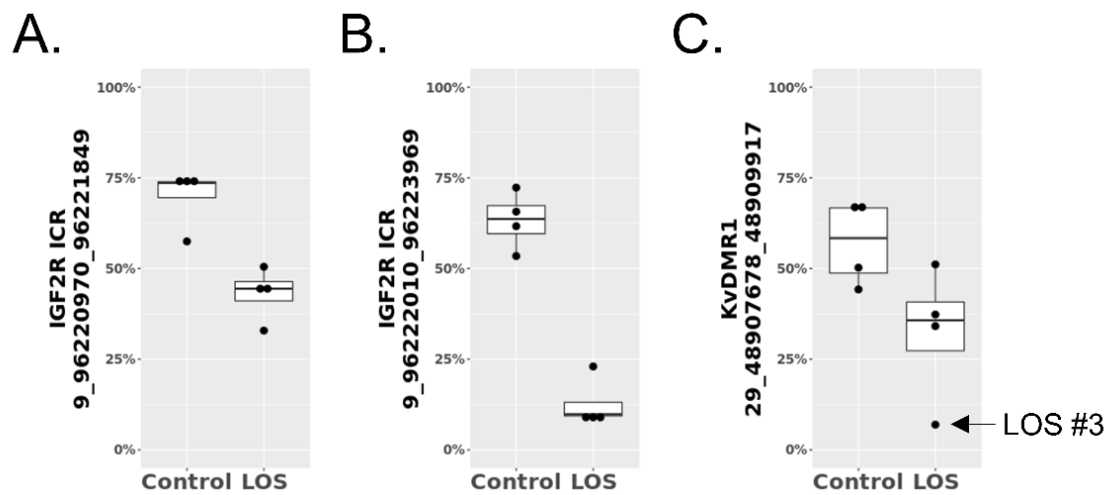


Figure 1. Differentially methylated regions identified within *IGF2R* ICR and *KvDMR1* in LOS when compared with controls.

(A-C) Data are represented as box plots with dots indicating individual samples. Y-axis shows average CpG DNA methylation level (not allelic).

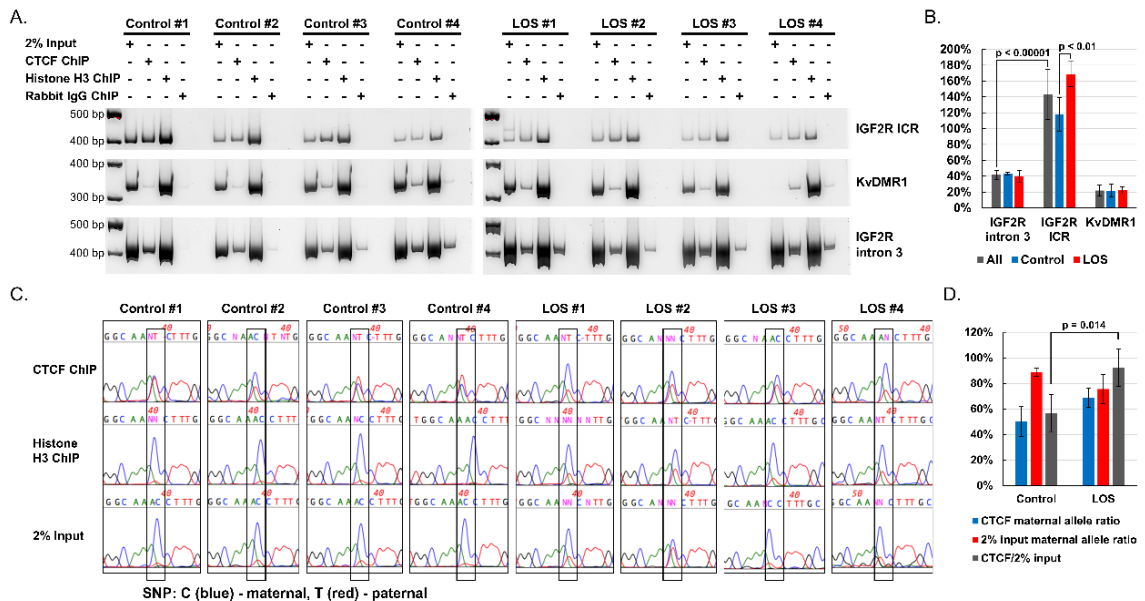


Figure 2. Validation of CTCF binding by chromatin immunoprecipitation (ChIP).

(A) PCR amplifications of ChIP products and input genomic DNA of predicted CTCF binding sites within the *IGF2R* ICR and KvDMR1 and within a region of *IGF2R* with no predicted CTCF binding site, namely intron 3. PCR amplicons were visualized on a 7% acrylamide gels. 2% input = input genomic DNA after micrococcal nuclease digestion without ChIP; Histone H3 ChIP = positive control; Rabbit IgG ChIP = negative control (for unspecific binding).

(B) Band intensity ratio between CTCF ChIP and 2% input DNA from (A) indicating increased presence of CTCF at *IGF2R* ICR in LOS and no binding at KvDMR1. Data are represented as mean \pm SD. P-values were from t-test.

(C) Allele-specific binding of CTCF at *IGF2R* ICR shown by Sanger sequencing. Peaks show the intensity of fluorescence signal for each nucleotide. The nucleotide enclosed in a box denotes a SNP between the maternal (C, blue) and paternal (T, red) alleles.

(D) Increased maternal allele binding of CTCF in LOS samples. Maternal allele ratio of CTCF ChIP and 2% input DNA, and corresponding CTCF/2% input ratio calculated from (C). The high ratio of maternal allele in the 2% input could be caused by micrococcal nuclease digestion during the ChIP procedure, as its efficiency is known to be affected by the status of chromatin compression (Grewal and Elgin, 2002). Data are represented as mean \pm SD. P-value was from t-test.

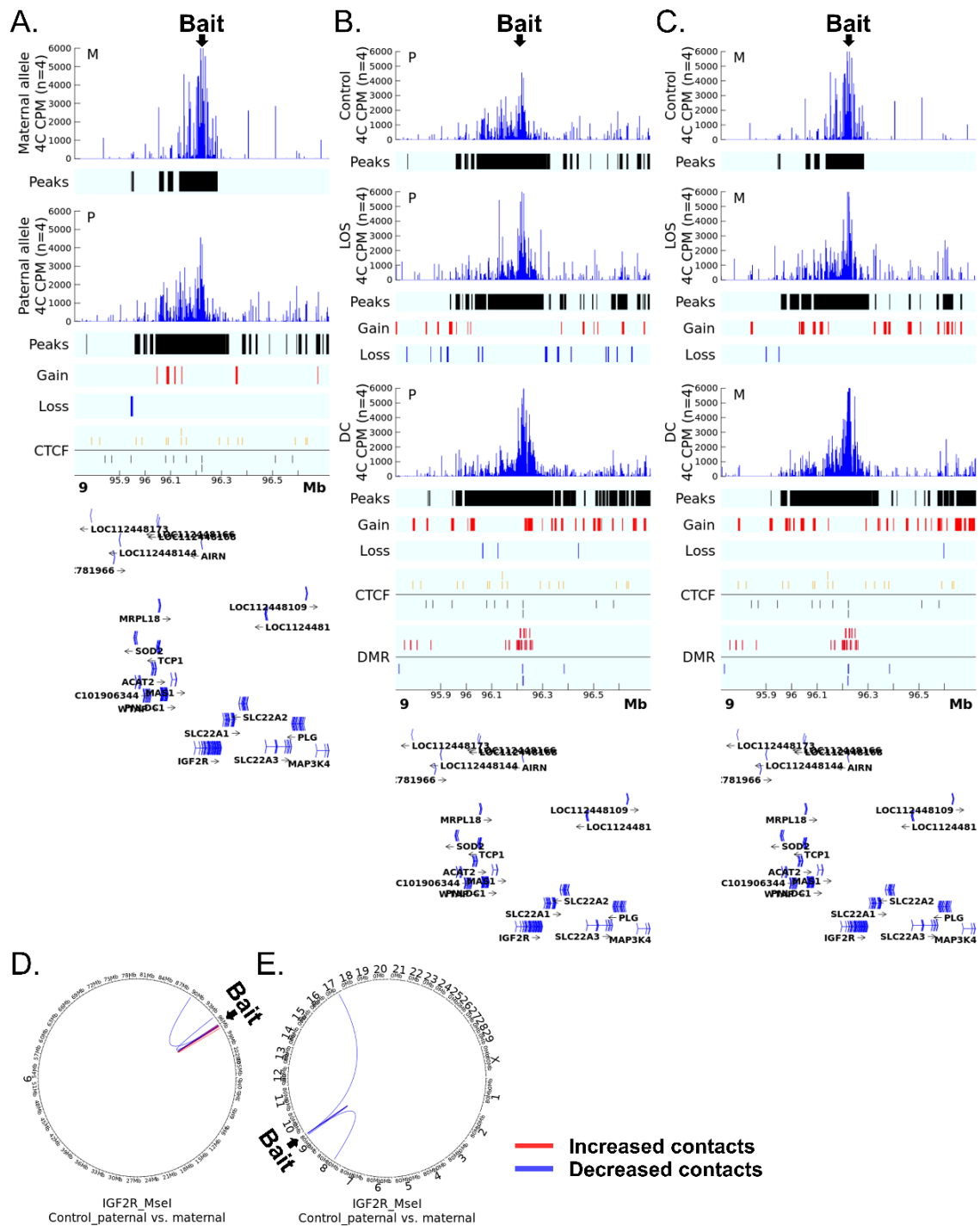


Figure 3.

Figure 3. 4C identified allele-specific cis and trans contacts with *IGF2R* ICR.

Shown are data for the *IGF2R*_Msel assay.

(A) Comparison of cis contacts between the paternal and maternal alleles in controls. Track '4C CPM' shows the mean normalized count of reads aligned to the genome indicating physical contacts with the bait. Track 'Peaks' show regions with statistically significant contacts with the bait identified by fourSig software within a group. Track 'Gain' (red line) and 'Loss' (blue line) indicate regions with statistically significant difference in contacts with the bait regions identified by DESeq2 between alleles. Track 'CTCF' shows predicted CTCF binding sites on the sense (gold line) or antisense (black line) strand. The gene annotation is at the bottom of the figure. Mb = megabases. CPM = counts per million reads. M = maternal allele. P = paternal allele.

(B and C) Comparison of allele-specific cis contacts between control, LOS, and DC. Shown are the comparison of LOS and DC groups vs controls. Track 'Gain' (red line) and 'Loss' (blue line) indicate regions with statistically significant difference in contacts with the bait regions identified by DESeq2 between groups. Track 'DMR' shows non-allelic differentially methylated regions identified between the LOS and the control group with the red line indicating increased and blue line indicating decreased methylation levels. All other track information as in (A).

(D and E) Comparison of contacts in far-cis and trans between parental alleles in controls. (D) far-cis contacts (chromosome 9) and (E) trans contacts (interchromosomal) in controls. Circos plots showing DESeq2-identified statistically different contacts with the bait in the paternal vs the maternal allele. Red line indicates increased contacts and blue line indicates decreased contacts.

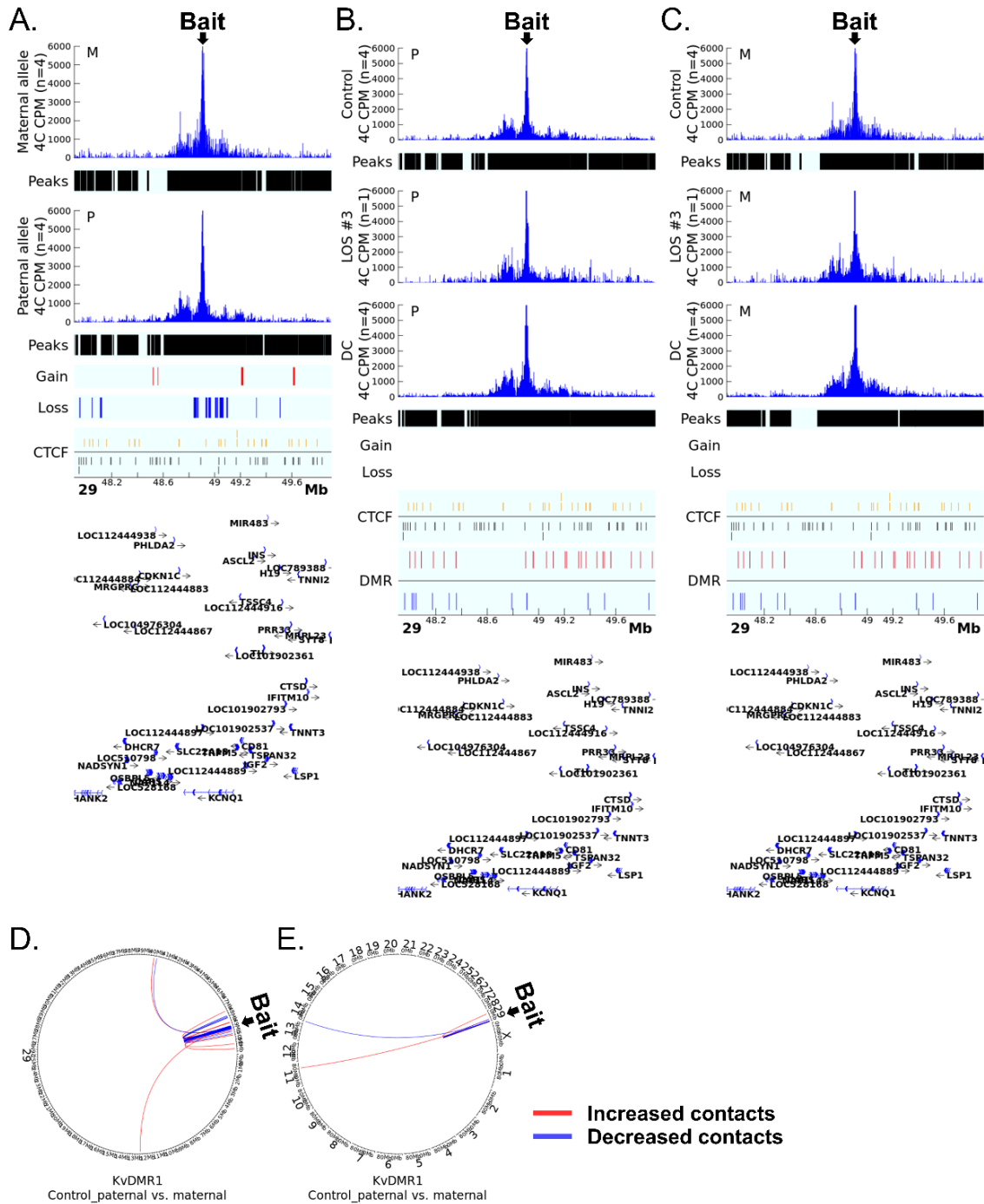


Figure 4.

Figure 4. 4C identified allele-specific cis and trans contacts with KvDMR1.

(A) Comparison of cis contacts between the paternal and maternal alleles in controls. Track '4C CPM' shows the mean normalized count of reads aligned to the genome indicating physical contacts with the bait. Track 'Peaks' show regions with statistically significant contacts with the bait identified by fourSig software within a group. Track 'Gain' (red line) and 'Loss' (blue line) indicate regions with statistically significant difference in contacts with the bait regions identified by DESeq2 between alleles. Track 'CTCF' shows predicted CTCF binding sites on the sense (gold line) or antisense (black line) strand. The gene annotation is at the bottom of the figure. Mb = megabases. CPM = counts per million reads. M = maternal allele. P = paternal allele.

(B and C) Comparison of allele-specific cis contacts between control, LOS, and DC. Shown are the comparison of LOS and DC groups vs controls. Track 'Gain' (red line) and 'Loss' (blue line) indicate regions with statistically significant difference in contacts with the bait regions identified by DESeq2 between groups. Track 'DMR' shows non-allelic differentially methylated regions identified between the LOS and the control group with the red line indicating increased and blue line indicating decreased methylation levels. All other track information as in (A).

(D and E) Comparison of contacts in far-cis and trans between parental alleles in controls. (D) far-cis contacts (chromosome 29) and (E) trans contacts (interchromosomal) in controls. Circos plots showing DESeq2-identified statistically different contacts with the bait in the paternal vs the maternal allele. Red line indicates increased contacts and blue line indicates decreased contacts.

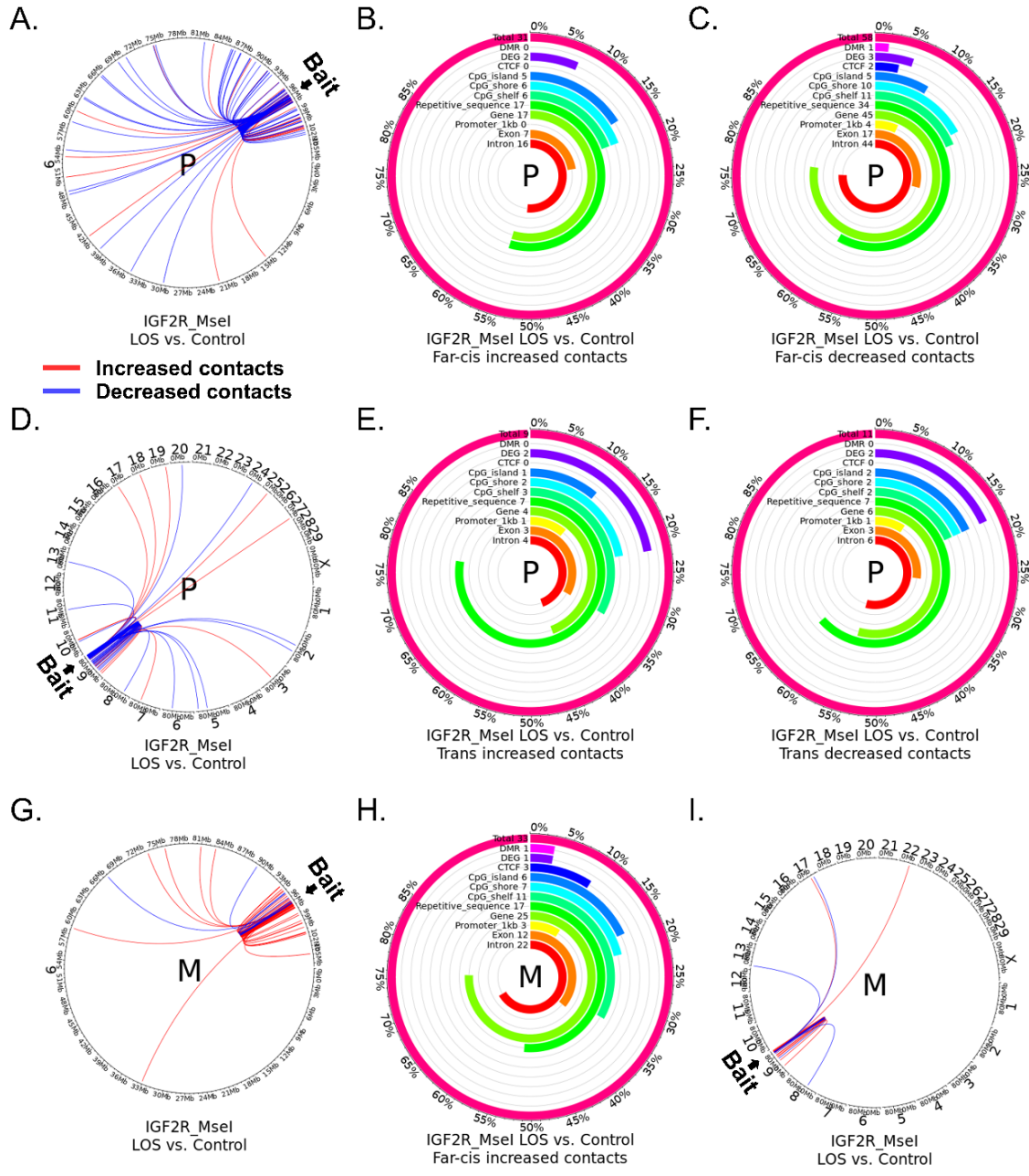


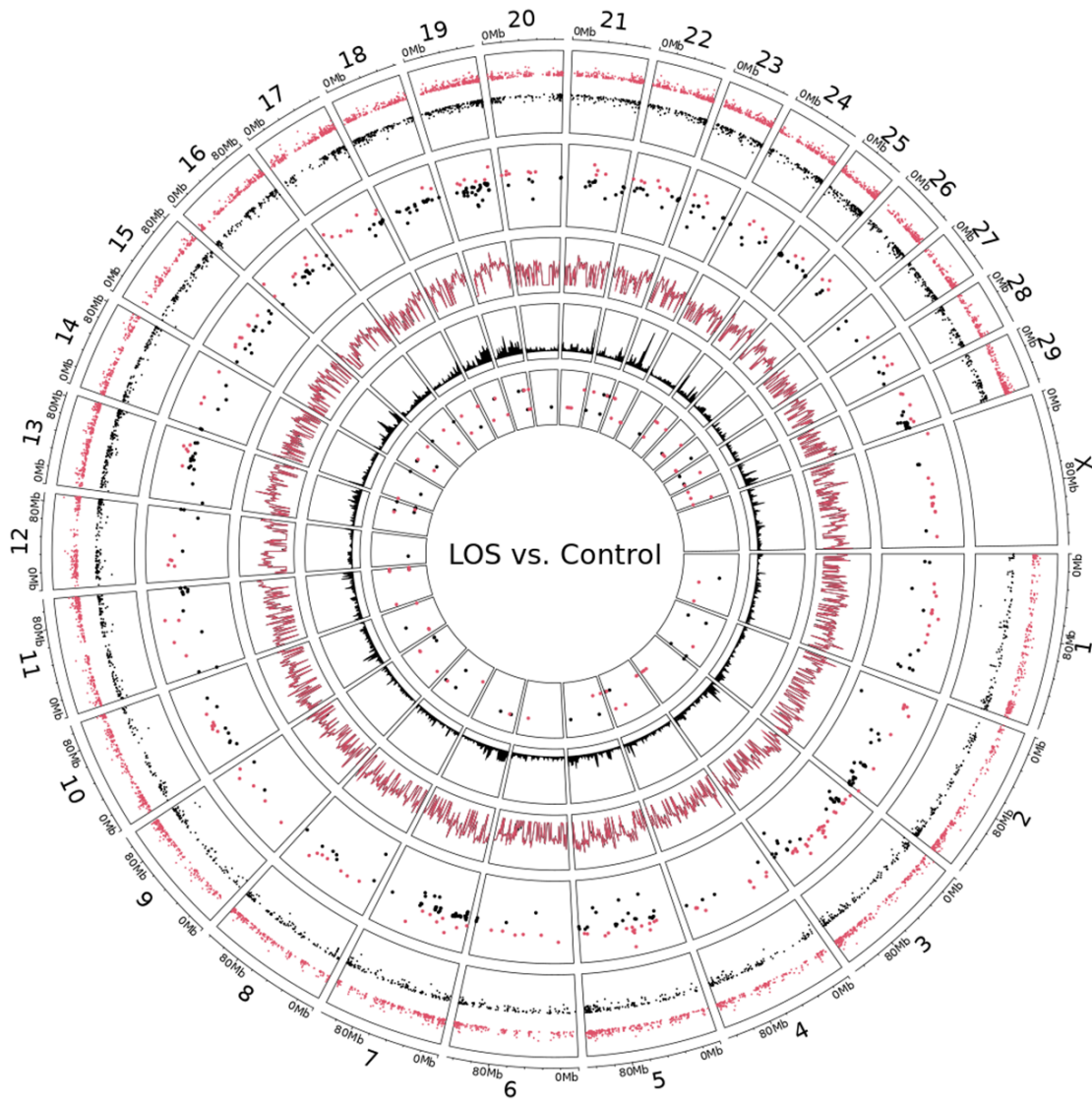
Figure 5.

Figure 5. Distribution of altered *IGF2R* ICR far-cis and trans contact across various genomic contexts.

Shown are data for the *IGF2R*_MseI assay.

(A, D, G, and I) Circos plots showing DESeq2-identified statistically different contacts with the bait in LOS vs. controls. Red line indicates increased contacts and blue line indicates decreased contacts. (A and D) Paternal and (G and I) maternal allele-specific comparisons. (A and G) far-cis contacts (chromosome 9) and (D and I) trans contacts (interchromosomal). P = paternal allele. M = maternal allele.

(B-C, E-F, and H) Figures show the total number of altered far-cis (B-C and H) or trans (E-F) contacts identified and the number and percent of increased (B, E, and H) and decreased (C and F) contacts over each genomic context. For example, ~55% of increased far-cis contacts in the paternal allele overlap with repetitive sequences (n=17) and genes (n=17). In addition, the figures include the number and percent of altered contacts that overlap differentially methylated regions (DMR) and within 100kb of differentially expressed genes (DEG) reported in this work. Analyses were only conducted for conditions with greater than five altered contacts.



- Layer (outer to inner):
1. DMR – Hyper (pink) or hypo (black) methylation
 2. DEG – Increased (pink) or decreased (black) expression
 3. Expression level – LOS (pink) or control (black) group mean
 4. Gene density
 5. DMR-CTCF – Hyper (pink) or hypo (black) DMR overlapped predicted CTCF binding sites

Figure 6.

Figure 6. Location based clustering tendency of differentially expressed genes indicates global alteration of chromosome architecture in LOS.

Shown are the genomic locations of differentially methylated regions (DMRs), differentially expressed genes (DEG) and predicted CTCF binding sites that overlap LOS DMRs. In addition, gene density and log₁₀ transformed gene expression level per million bases are shown. The vertical location indicates level of misregulation for DMR and DEG, and sense (external) or antisense (internal) strands of CTCF binding sites. Mb = megabases. Note: track three (expression level) shows that for the most part the level of expression of genes globally is similar between LOS and controls.

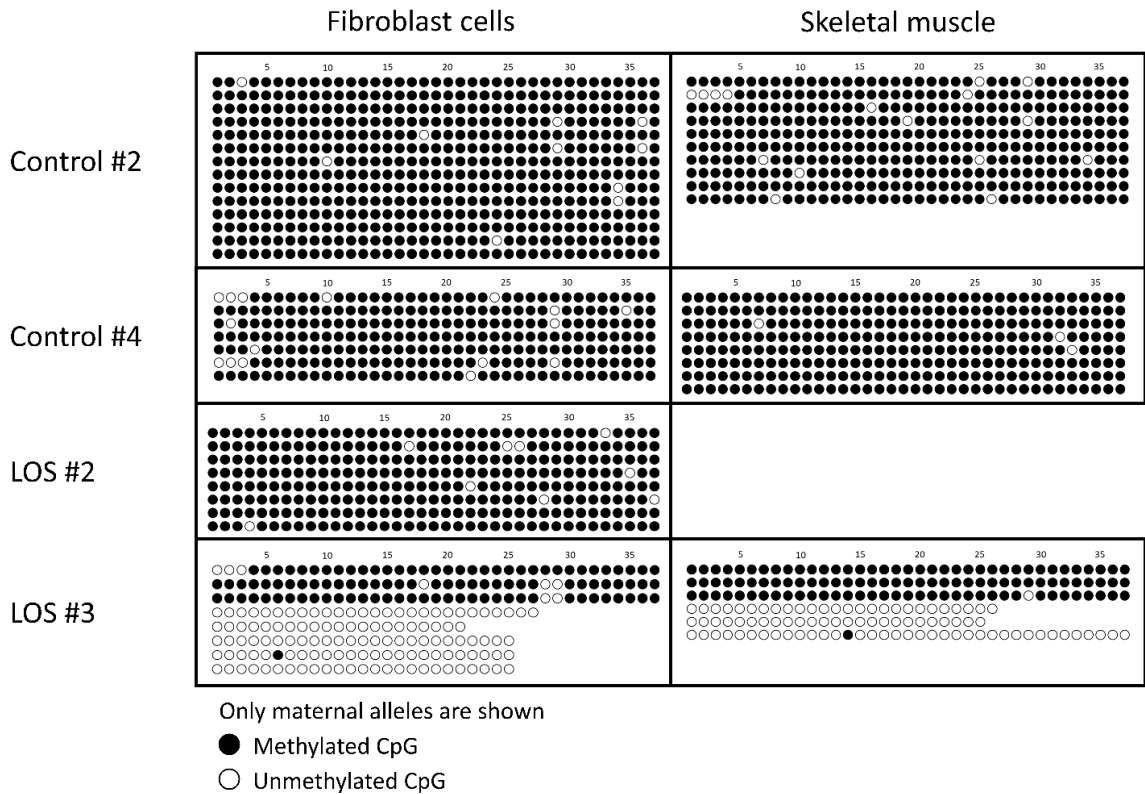


Figure S1. Maternal allele CpG DNA methylation status at KvDMR1, related to Figure 1.

Methylation status was determined by bisulfite PCR, molecular cloning, and Sanger sequencing. Open circles represent unmethylated CpG and filled circles indicate methylated CpG. Only maternal alleles were shown, and paternal alleles had no methylated CpG found. For sample LOS#2, we did not perform bisulfite PCR and cloning for the muscle sample since the locus is methylated over this region in fibroblast.

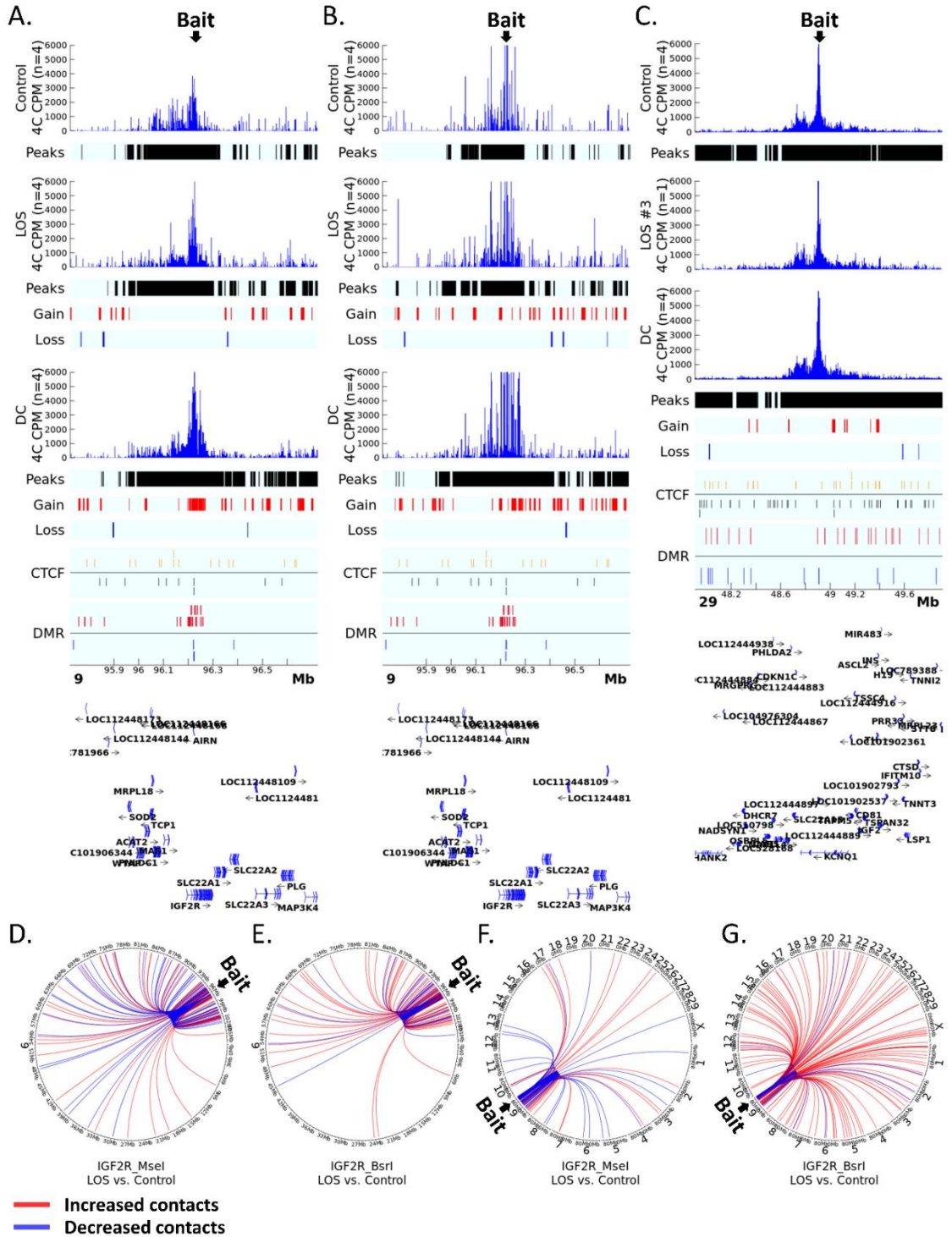


Figure S2.

Figure S2. 4C identified overall (not allele specific) cis and trans contacts, related to Figure 3 and Figure 4.

(A-C) Comparison of cis contacts between control, LOS, and DC. (A) IGF2R_Msel assay, (B) IGF2R_Bsrl assay, and (C) KvDMR1 assays. Shown are the comparison of LOS and DC groups vs controls. Track '4C CPM' shows the mean normalized count of reads aligned to the genome indicating physical contacts with the bait. Track 'Peaks' show regions with statistically significant contacts with the bait identified by fourSig software within a group. Track 'Gain' (red line) and 'Loss' (blue line) indicate regions with statistically significant difference in contacts with the bait regions identified by DESeq2 between groups. Track 'CTCF' shows predicted CTCF binding sites on the sense (gold line) or antisense (black line) strand. Track 'DMR' shows non-allelic differentially methylated regions identified between the LOS and the control group with the red line indicating increased and blue line indicating decreased methylation levels. The gene annotation is at the bottom of the figure. Mb = megabases. CPM = counts per million reads.

(D-G) Comparison of contacts in far-cis and trans between groups. (D and F) IGF2R_Msel assay and (E and G) IGF2R_Bsrl. (D-E) far-cis contacts (chromosome 9) and (F-G) trans contacts (interchromosomal). Circos plots showing DESeq2-identified statistically different contacts with the bait in LOS vs. controls. Red line indicates increased contacts and blue line indicates decreased contacts.

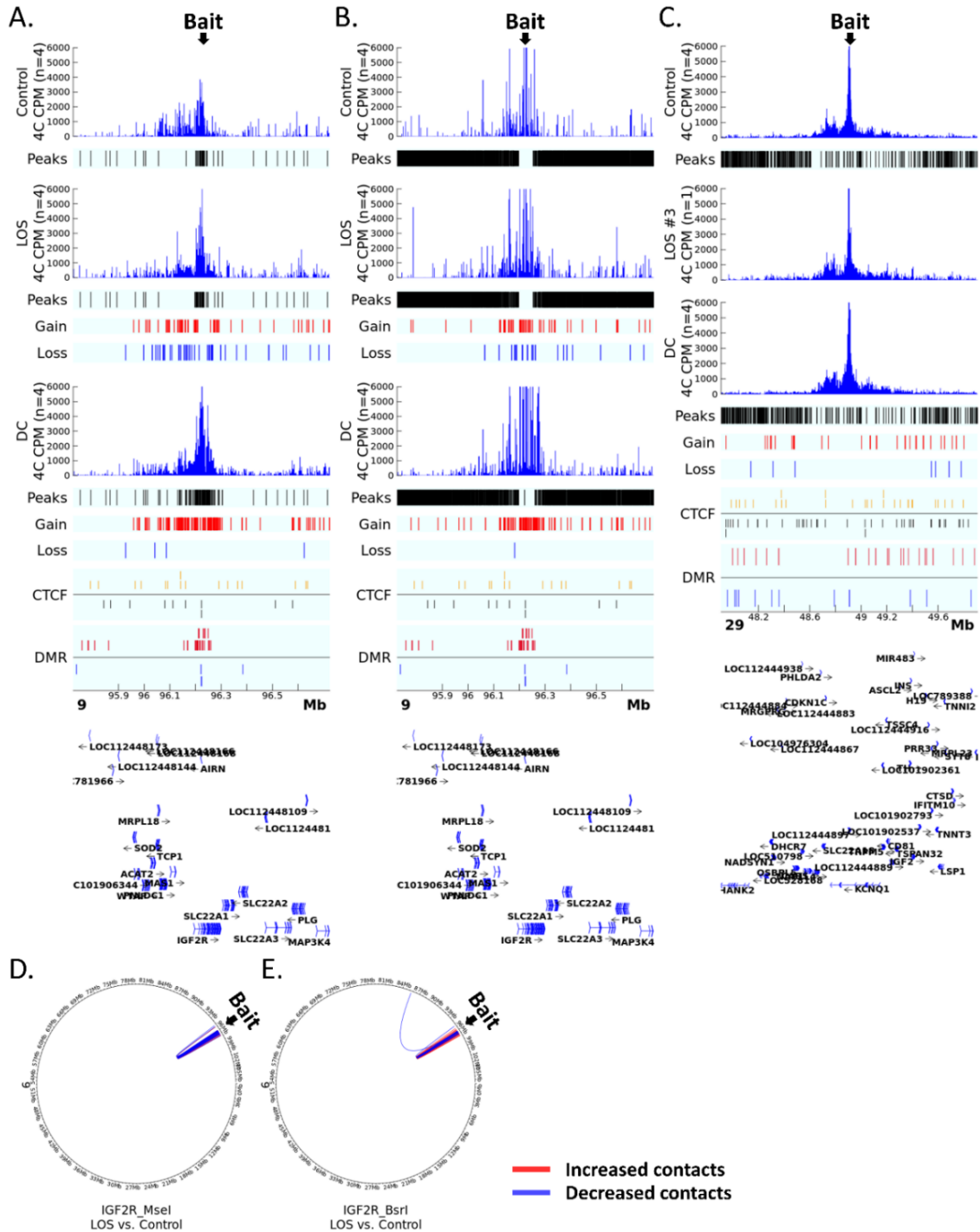
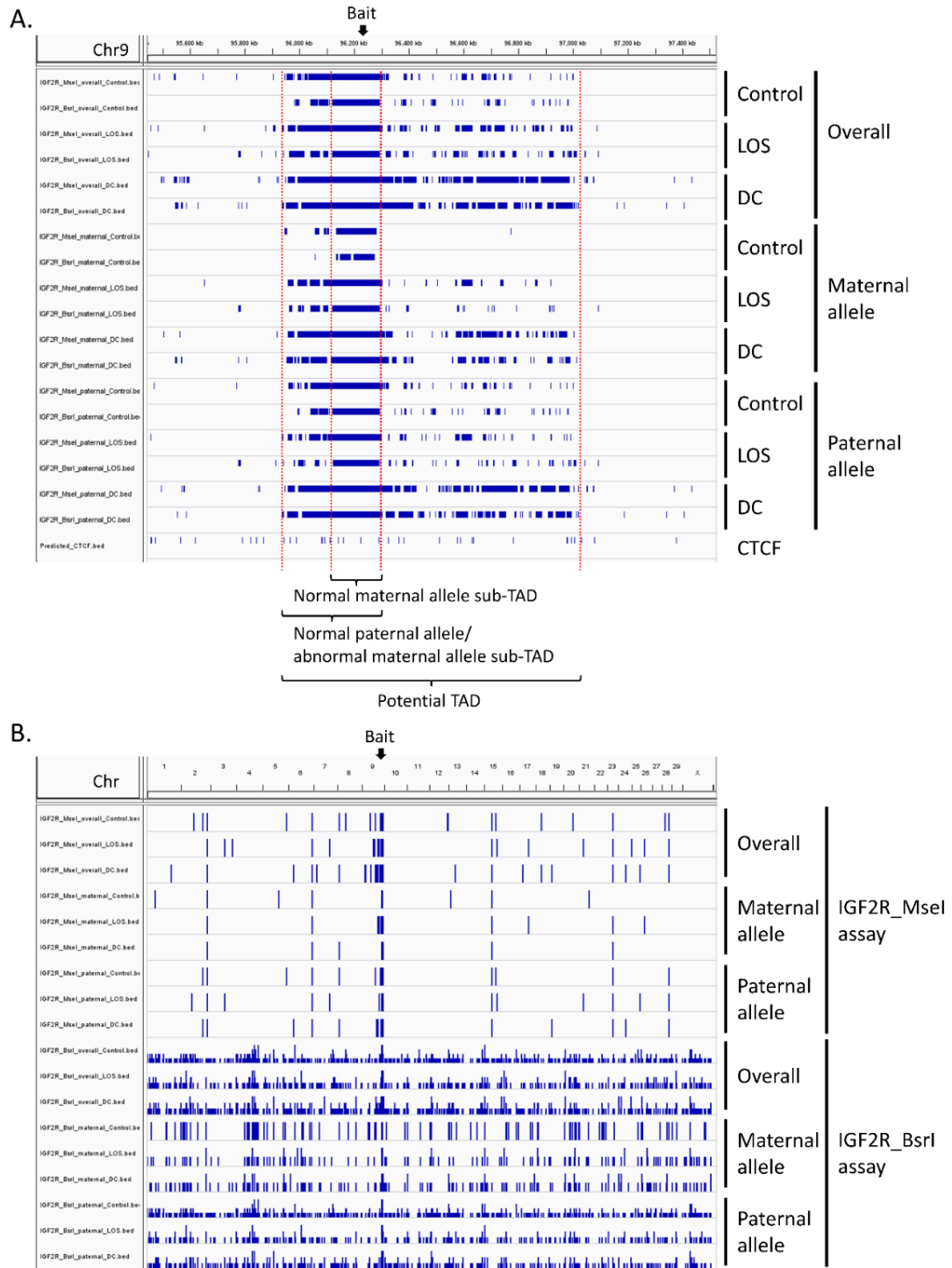


Figure S3. 4Cker statistical results for overall (not allele specific) contacts, related to Figure 3 and Figure 4.

(A-E) Same data as in Figure S2 A-E analyzed with 4Cker.



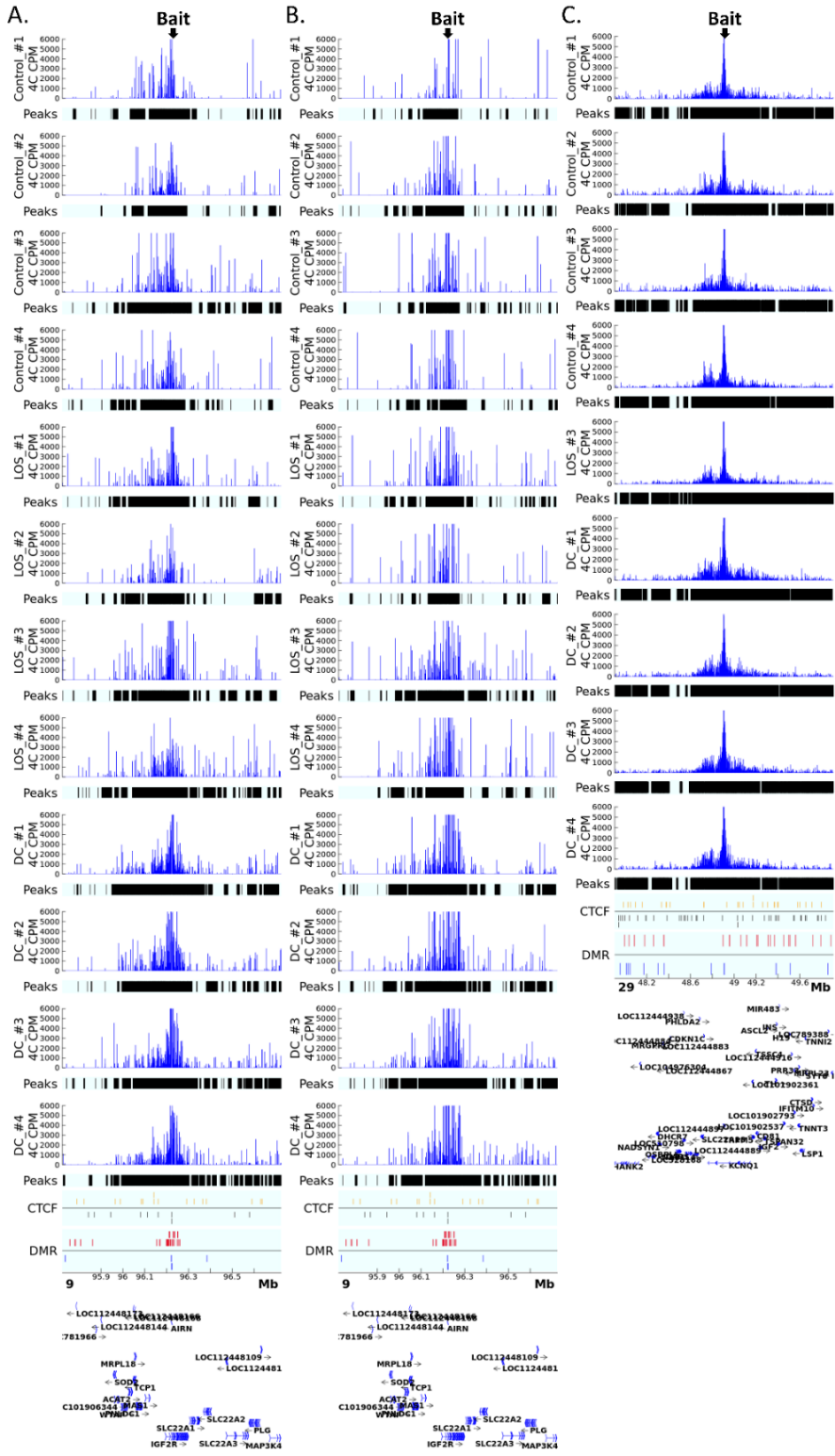


Figure S5.

Figure S5. 4C identified overall (not allele specific) cis contacts for individual samples, related to Figure 3 and Figure 4.

(A-C) Cis contacts of control, LOS, and DC samples for (A) IGF2R_MseI assay, (B) IGF2R_BsrI assay, and (C) KvDMR1 assays. Track '4C CPM' shows the normalized count of reads aligned to the genome indicating physical contacts with the bait. Track 'Peaks' show regions with statistically significant contacts with the bait identified by fourSig software for individual sample. Track 'CTCF' shows predicted CTCF binding sites on the sense (gold line) or antisense (black line) strand. Track 'DMR' shows non-allelic differentially methylated regions identified between the LOS and the control group with the red line indicating increased and blue line indicating decreased methylation levels. The gene annotation is at the bottom of the figure. Mb = megabases. CPM = counts per million reads.

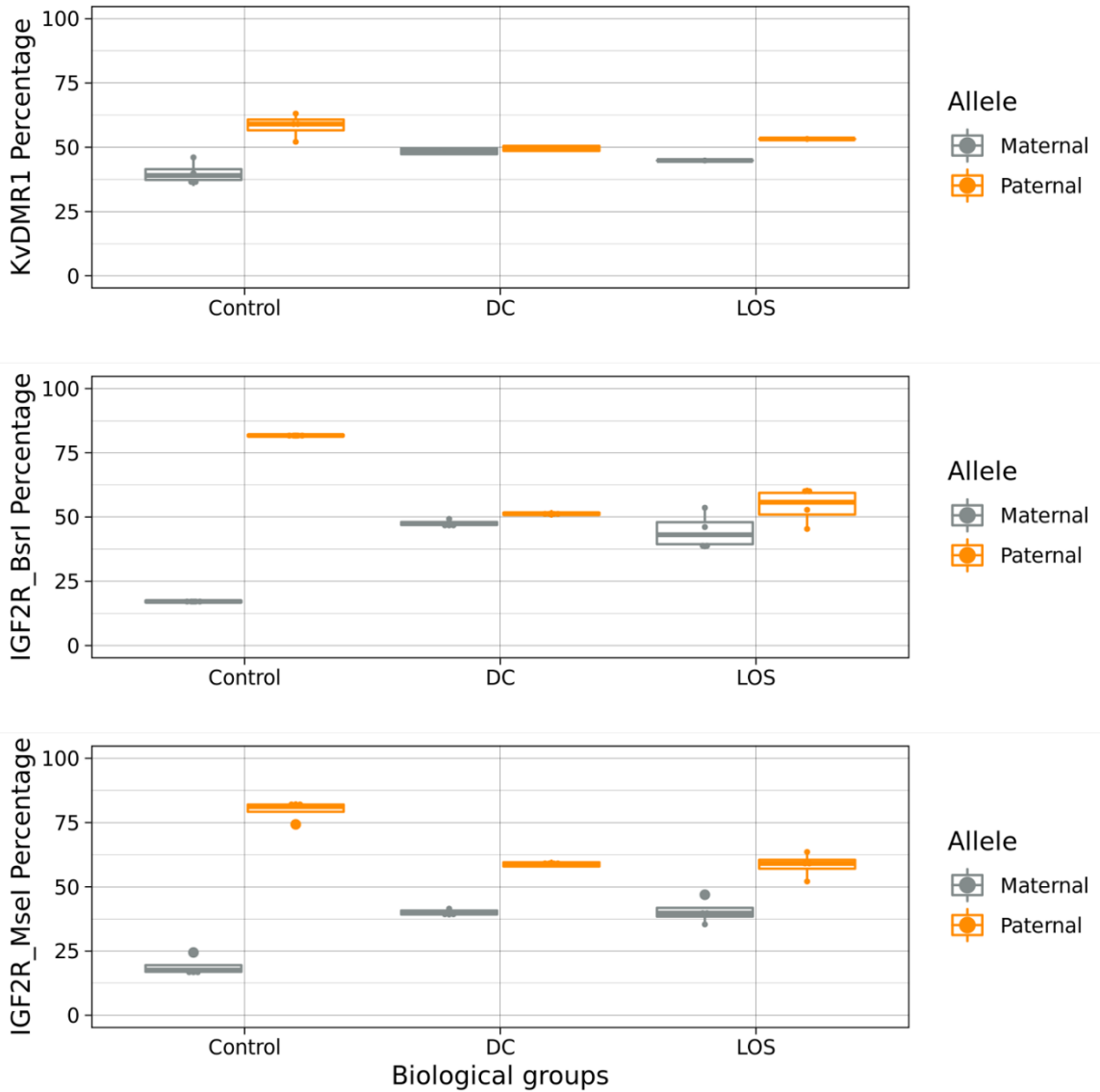


Figure S6. 4C sequencing reads allelic alignment, related to Figure 3 and Figure 4.

Percentage of aligned reads belonging to maternal allele (gray, left) or paternal allele (orange, right). Data are represented as box plots with dots indicating individual samples.

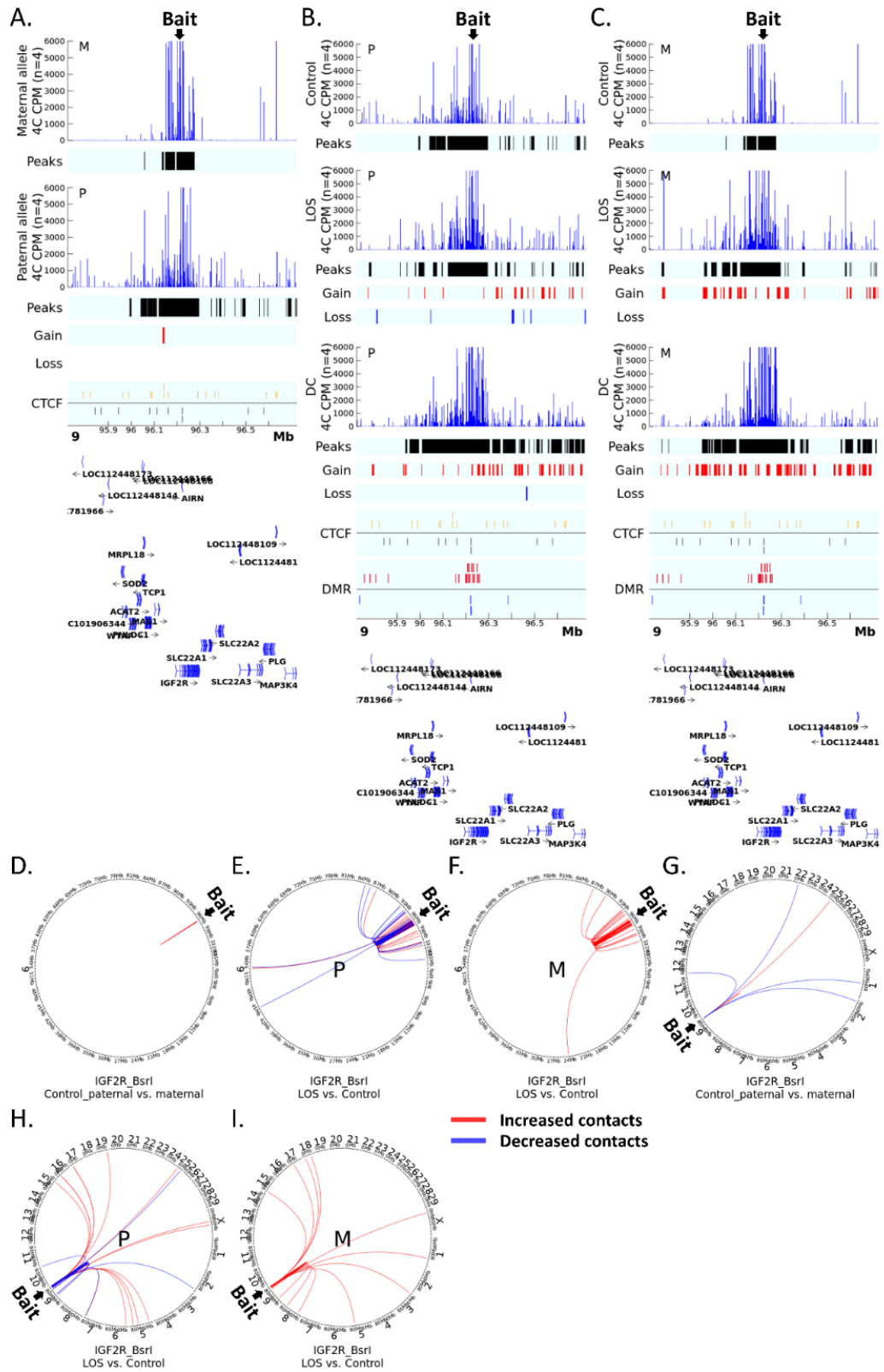


Figure S7.

Figure S7. 4C identified allele-specific cis and trans contacts with *IGF2R* ICR, related to Figure 3.

Shown are data for the *IGF2R*_Bsrl assay.

(A) Comparison of cis contacts between the paternal and maternal alleles in controls. Track '4C CPM' shows the mean normalized count of reads aligned to the genome indicating physical contacts with the bait. Track 'Peaks' show regions with statistically significant contacts with the bait identified by fourSig software within a group. Track 'Gain' (red line) and 'Loss' (blue line) indicate regions with statistically significant difference in contacts with the bait regions identified by DESeq2 between alleles. Track 'CTCF' shows predicted CTCF binding sites on the sense (gold line) or antisense (black line) strand. The gene annotation is at the bottom of the figure. Mb = megabases. CPM = counts per million reads. M = maternal allele. P = paternal allele.

(B and C) Comparison of allele-specific cis contacts between control, LOS, and DC. Shown are the comparison of LOS and DC groups vs controls. Track 'Gain' (red line) and 'Loss' (blue line) indicate regions with statistically significant difference in contacts with the bait regions identified by DESeq2 between groups. Track 'DMR' shows non-allelic differentially methylated regions identified between the LOS and the control group with the red line indicating increased and blue line indicating decreased methylation levels. All other track information as in (A).

(D and G) Comparison of contacts in far-cis and trans between parental alleles in controls. (D) far-cis contacts (chromosome 9) and (G) trans contacts (interchromosomal) in controls. Circos plots showing DESeq2-identified statistically different contacts with the bait in the paternal vs the maternal allele. Red line indicates increased contacts and blue line indicates decreased contacts.

(E-F and H-I) Comparison of contacts in far-cis and trans between control, LOS, and DC. (E-F) far-cis contacts (chromosome 9) and (H-I) trans contacts (interchromosomal). (E and H) Paternal allele and (F and I) maternal allele. Circos plots showing DESeq2-identified statistically different contacts with the bait in LOS vs. controls. Red line indicates increased contacts and blue line indicates decreased contacts.

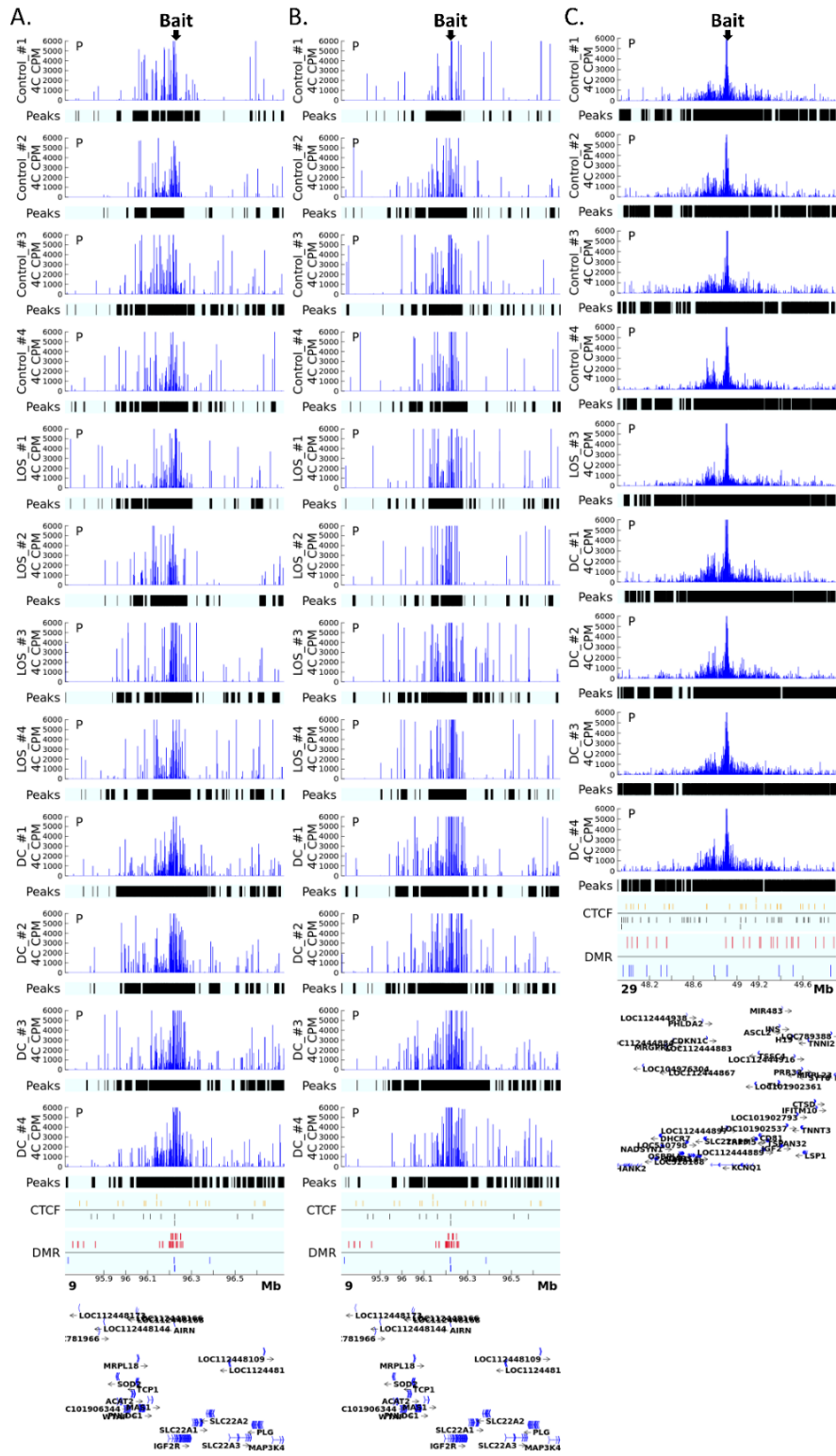


Figure S8.

Figure S8. 4C identified paternal allele cis contacts for individual samples, related to Figure 3 and Figure 4.

(A-C) Cis contacts of control, LOS, and DC samples for (A) IGF2R_MseI assay, (B) IGF2R_BsrI assay, and (C) KvDMR1 assays. Track '4C CPM' shows the normalized count of reads aligned to the genome indicating physical contacts with the bait. Track 'Peaks' show regions with statistically significant contacts with the bait identified by fourSig software for individual sample. Track 'CTCF' shows predicted CTCF binding sites on the sense (gold line) or antisense (black line) strand. Track 'DMR' shows non-allelic differentially methylated regions identified between the LOS and the control group with the red line indicating increased and blue line indicating decreased methylation levels. The gene annotation is at the bottom of the figure. Mb = megabases. CPM = counts per million reads.

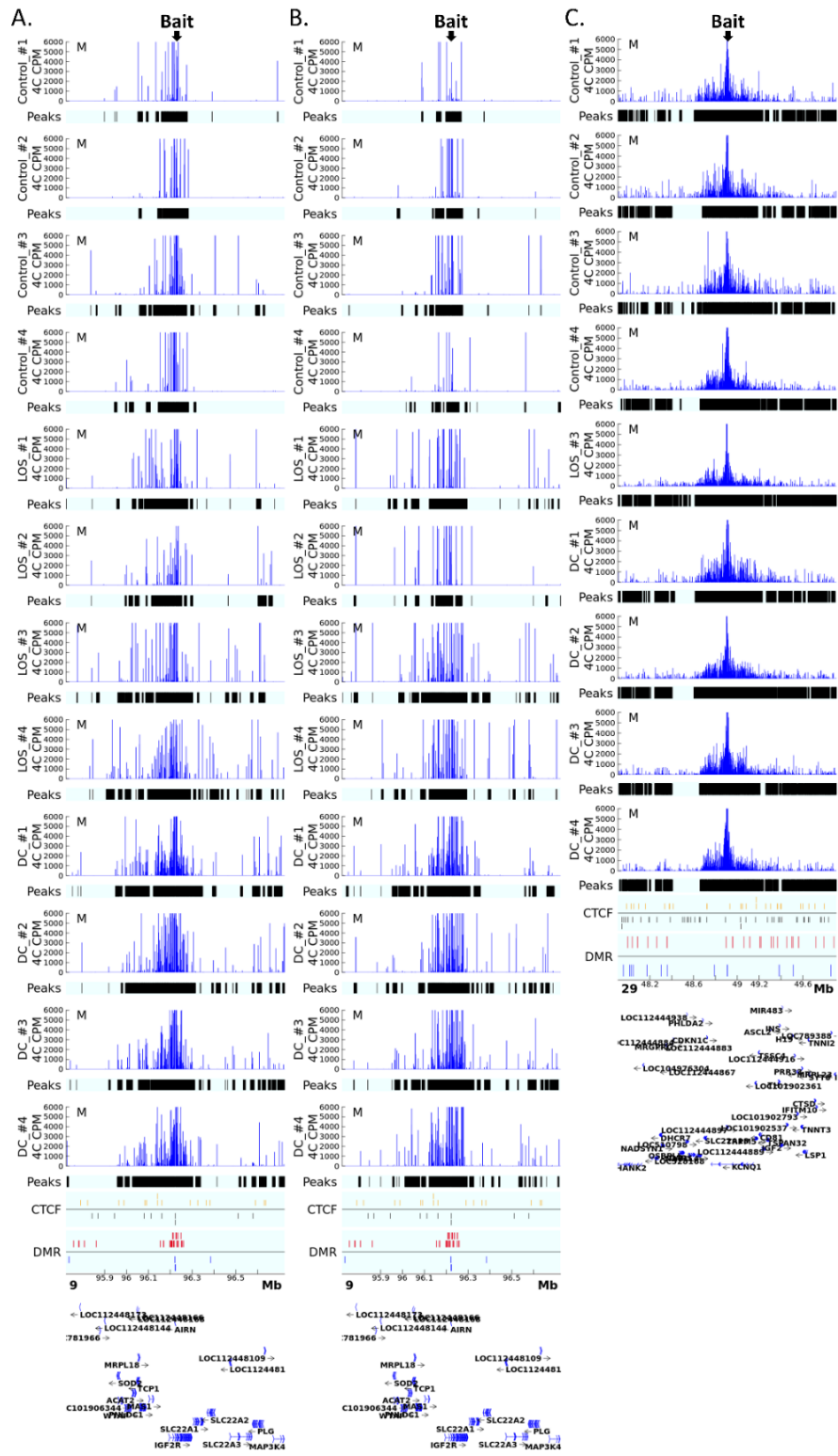


Figure S9.

Figure S9. 4C identified maternal allele cis contacts for individual samples, related to Figure 3 and Figure 4.

(A-C) Cis contacts of control, LOS, and DC samples for (A) IGF2R_MseI assay, (B) IGF2R_BsrI assay, and (C) KvDMR1 assays. Track '4C CPM' shows the normalized count of reads aligned to the genome indicating physical contacts with the bait. Track 'Peaks' show regions with statistically significant contacts with the bait identified by fourSig software for individual sample. Track 'CTCF' shows predicted CTCF binding sites on the sense (gold line) or antisense (black line) strand. Track 'DMR' shows non-allelic differentially methylated regions identified between the LOS and the control group with the red line indicating increased and blue line indicating decreased methylation levels. The gene annotation is at the bottom of the figure. Mb = megabases. CPM = counts per million reads.

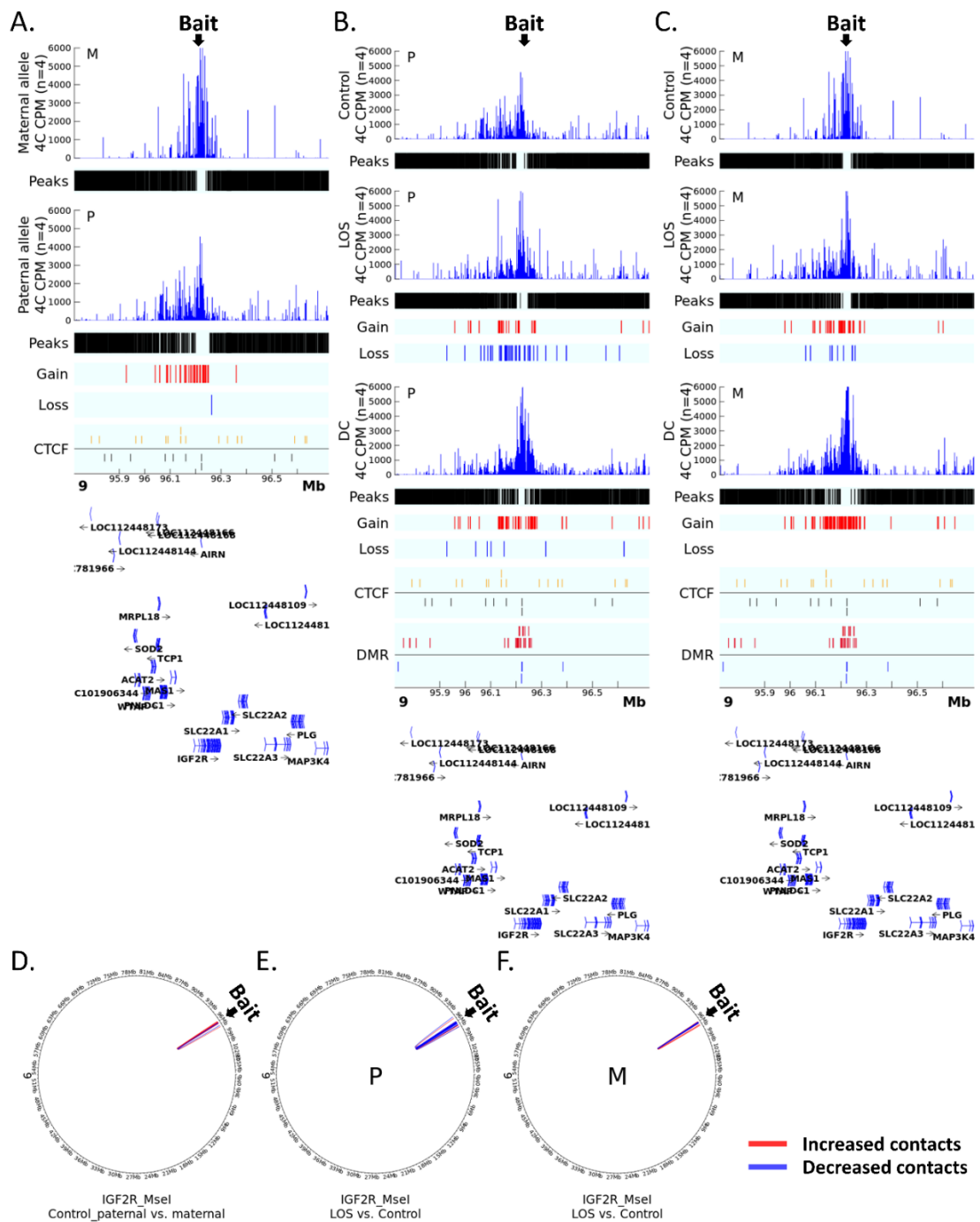


Figure S10. 4Cker statistical results for allele-specific cis and far-cis contacts with *IGF2R* ICR, related to Figure 3.

(A-F) Same *IGF2R*_MseI data as in Figure 3 A-D and Figure 5 A and G analyzed with 4Cker.

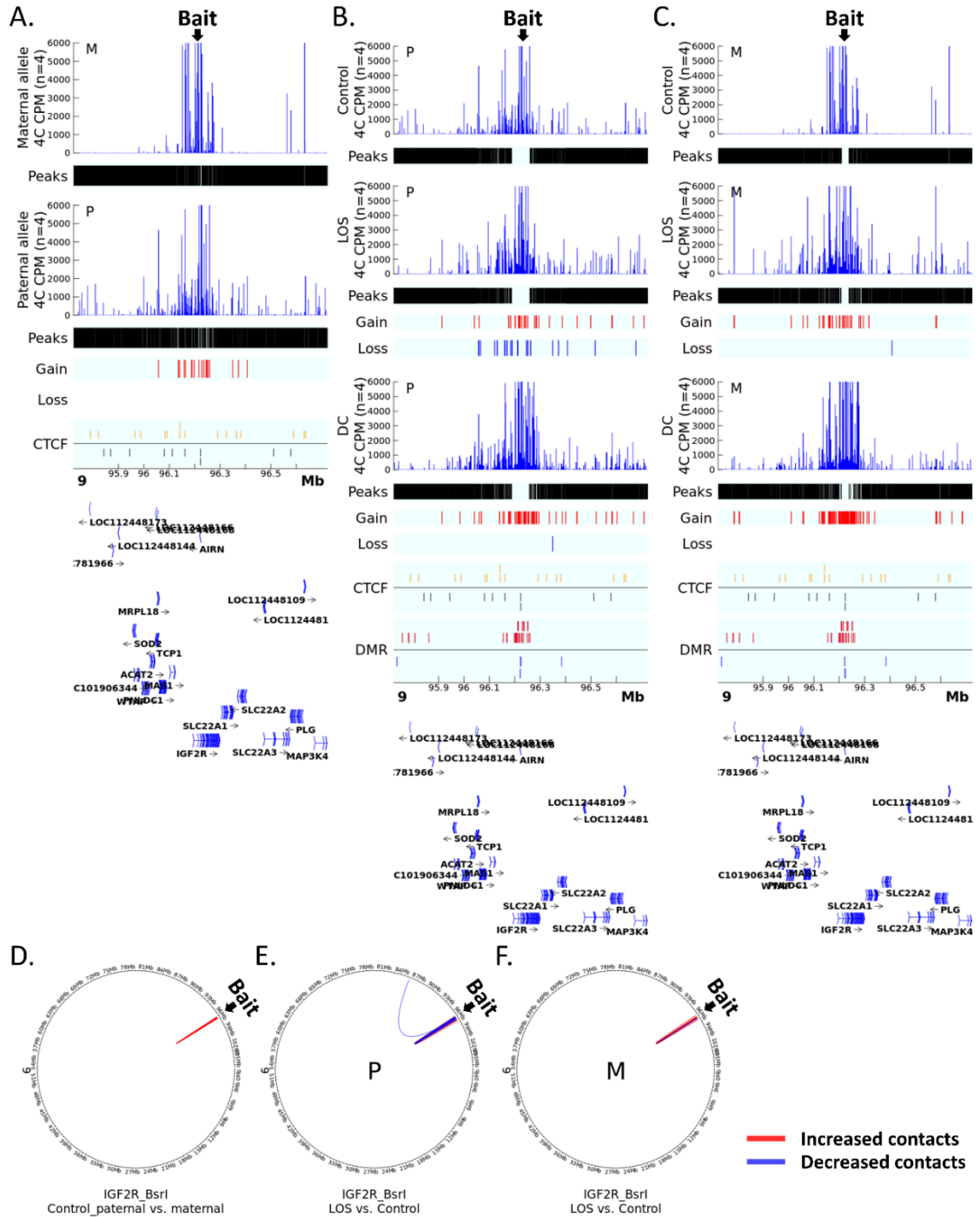


Figure S11. 4Cker statistical results for allele-specific cis and far-cis contacts with *IGF2R* ICR, related to Figure 3.

(A-F) Same *IGF2R_BsrI* data as in Figure S7 A-F analyzed with 4Cker.

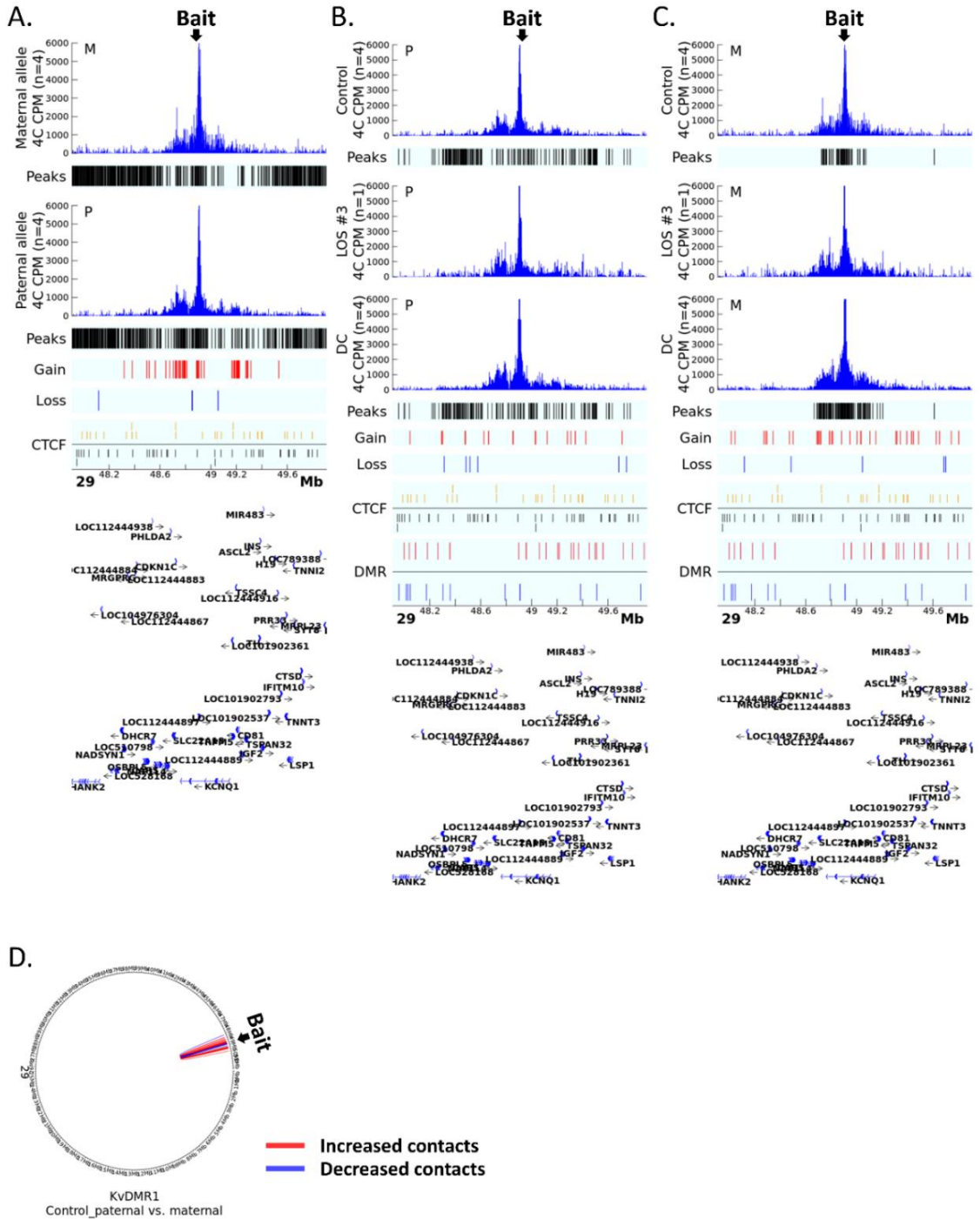


Figure S12. 4Cker statistical results for allele-specific cis and far-cis contacts with KvDMR1, related to Figure 4.

(A-D) Same KvDMR1 data as in Figure 4 A-D analyzed with 4Cker.

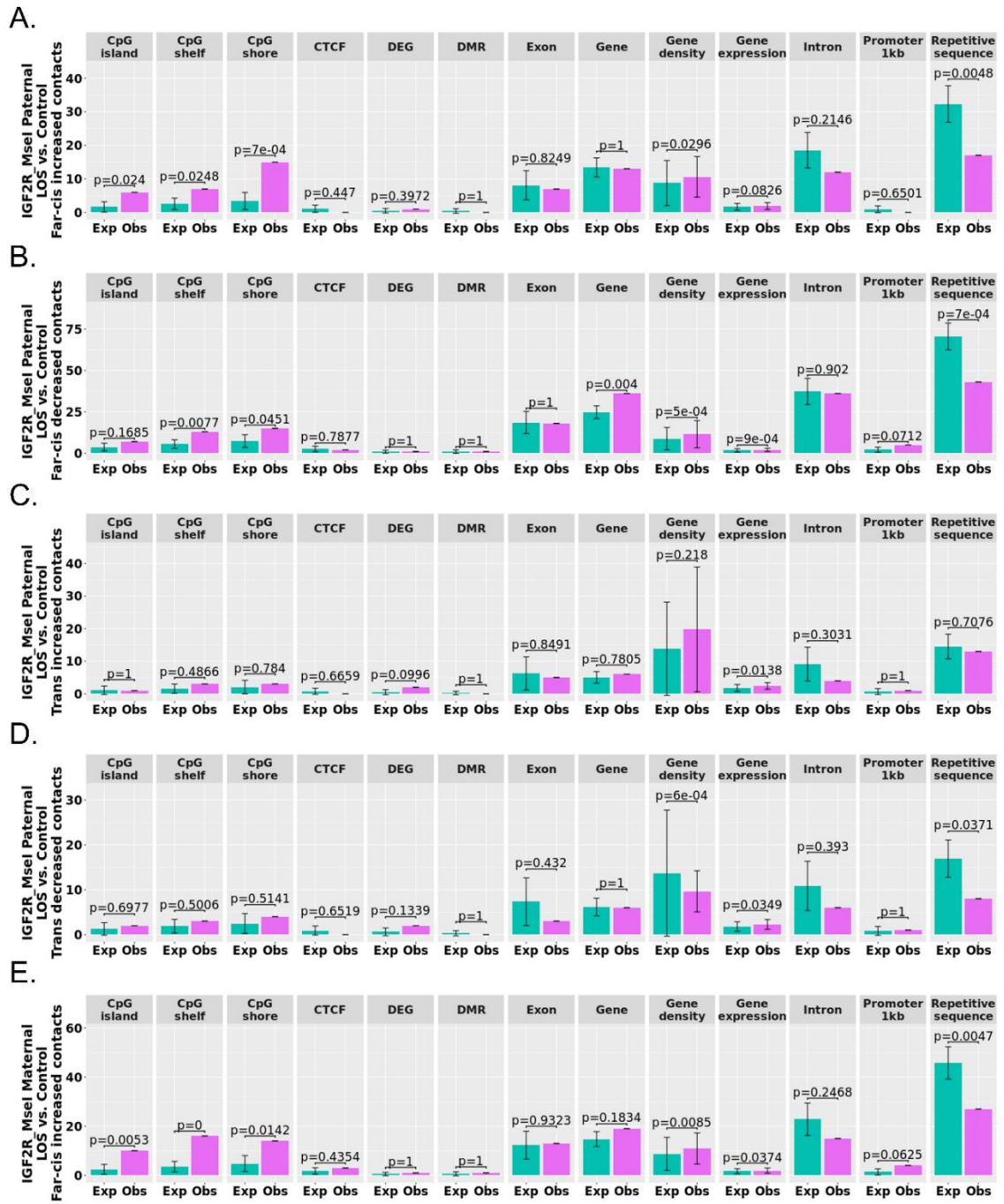


Figure S13.

Figure S13. Distribution of altered *IGF2R* ICR trans contact across various genomic contexts, related to Figure 5.

Results shown are for IGF2R_Msel.

(A-E) Observed and expected number of unique genomic context that overlaps with altered far-cis (A-B and E) or trans (C-D) contacts. Differentially methylated regions (DMR) that overlap with altered contacts and differentially expressed genes (DEG) within 100kb of altered contacts were also examined. In addition, gene expression level of LOS (log10 transformed) and gene density per million bases were calculated for altered contacts and permuted results. Analyses were only conducted for conditions with greater than five altered contacts. Obs = observed number. Exp = expected number (obtained from shuffling altered contacts across corresponding genomic region 10,000 times). Data are represented as mean \pm SD. For gene expression level and gene density, the p values between Obs and Exp were obtained from t-test. For other genomic contents, the p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|) / 10000$.

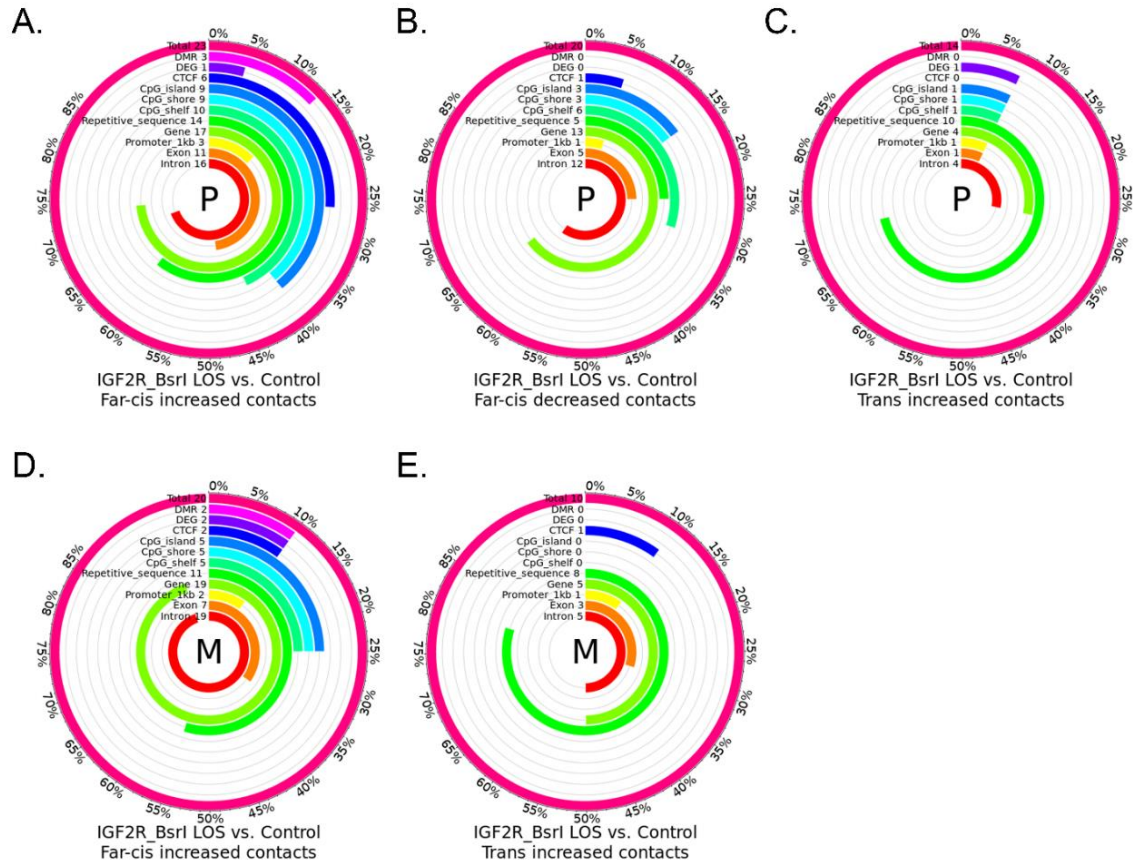


Figure S14. Distribution of altered *IGF2R* ICR trans contact across various genomic contexts, related to Figure 5.

Shown results are for IGF2R_Bsrl.

(A-E) Figure show the total number of altered far-cis (A-B and D) and trans (C and E) contacts identified and the number and percent of increased (A, and C-E) and decreased (B) contacts over each genomic context. In addition, the figures include the number and percent of altered contacts that overlap differentially methylated regions (DMR) and within 100kb of differentially expressed genes (DEG) reported in this work. Analyses were only conducted for conditions with greater than five altered contacts. P = paternal allele. M = maternal allele.

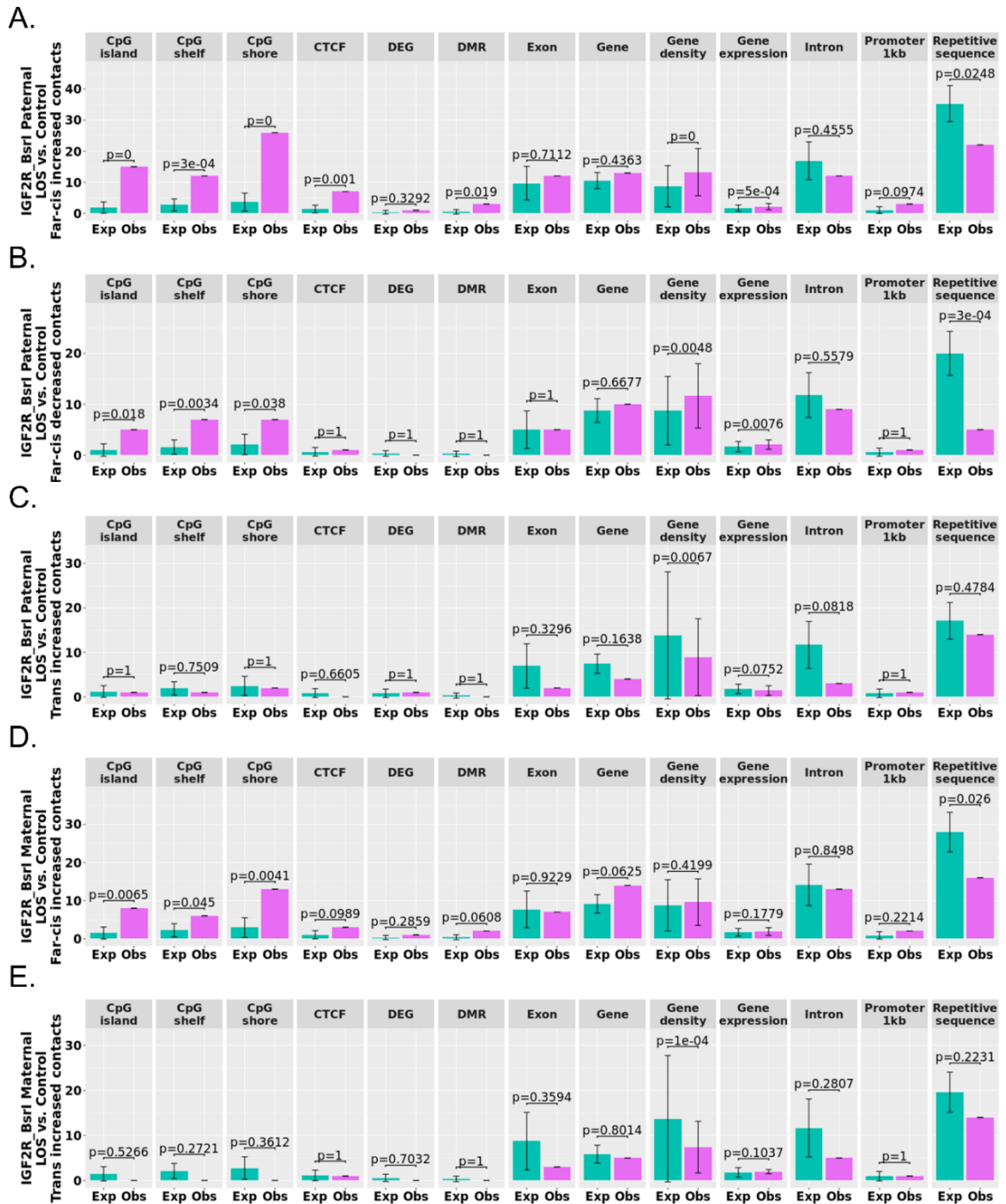


Figure S15.

Figure S15. Distribution of altered *IGF2R* ICR trans contact across various genomic contexts, related to Figure 5.

Shown results are for *IGF2R*_Bsrl.

(A-E) Observed and expected number of unique genomic context that overlaps with altered trans contacts. Differentially methylated regions (DMR) that overlap with altered contacts and differentially expressed genes (DEG) within 100kb of altered contacts were also examined. In addition, gene expression level of LOS (log10 transformed) and gene density per million bases were calculated for altered contacts and permuted results. Analyses were only conducted for conditions with greater than five altered contacts. Obs = observed number. Exp = expected number (obtained from shuffling altered contacts across corresponding genomic region, trans in this case, 10,000 times). Data are represented as mean \pm SD. For gene expression level and gene density, the p values between Obs and Exp were obtained from t-test. For other genomic contents, the p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|) / 10000$.

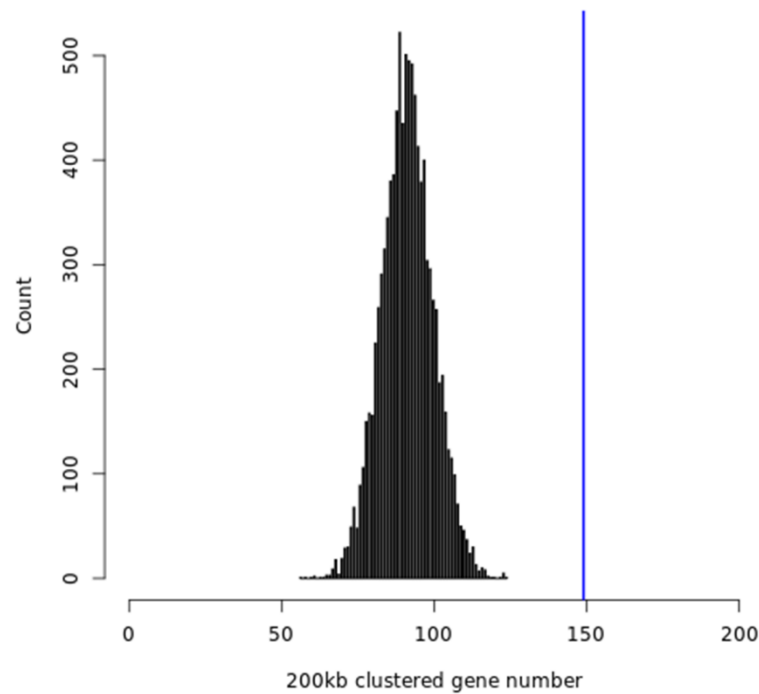


Figure S16. Distribution of permutation test for clustering tendency of differentially expressed genes (DEGs), related to Figure 6.

Black lines show results of 10,000 times of permutating of 548 genes (same number as DEGs) against the 15,042 expressed genes in this study, and the 200kb clustered gene number (Exp) was calculated each time. Blue line indicates the number (Obs) of the 149 clustered DEGs identified in this work, which is significantly higher ($p = 0$) than the mean of permutation tests (91.46 with standard deviation 8.49). The p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|) / 10000$.

Chapter 3: Spontaneous and ART-induced large offspring syndrome: similarities and differences in DNA methylome

3.1 Abstract

Large/abnormal offspring syndrome (LOS/AOS) is a congenital overgrowth syndrome reported in ruminants produced by assisted reproduction (ART-LOS) which exhibit global disruption of the epigenome and transcriptome. LOS/AOS shares phenotypes and epigenotypes with the human congenital overgrowth condition Beckwith-Wiedemann syndrome. We have reported that LOS occurs spontaneously (SLOS); however, to date, no study has been conducted to determine if SLOS has the same methylome epimutations as ART-LOS. In this study, we performed whole genome bisulfite sequencing to examine global DNA methylation in bovine SLOS and ART-LOS tissues. We observed unique patterns of global distribution of differentially methylated regions (DMRs) over different genomic contexts, such as promoters, CpG islands, shores and shelves, as well as at repetitive sequences. In addition, we included data from two previous LOS studies to identify shared vulnerable genomic loci in LOS. Overall, we identified 320 genomic loci in LOS that have alterations in DNA methylation when compared to controls. Specifically, there are 25 highly

vulnerable loci that could potentially serve as molecular markers for the diagnosis of LOS, including at the promoters of *DMRT2* and *TBX18*, at the imprinted gene bodies of *IGF2R*, *PRDM8*, and *BLCAP/NNAT*, and at multiple CpG islands. We also observed tissue-specific DNA methylation patterns between muscle and blood, and conservation of ART-induced DNA methylation changes between muscle and blood. We conclude that as ART-LOS, SLOS is an epigenetic condition. In addition, SLOS and ART-LOS share similarities in methylome epimutations.

3.2 Introduction

Large/abnormal offspring syndrome (LOS/AOS) is a congenital overgrowth syndrome that has been reported in ruminants (Farin et al., 2006; Rivera et al., 2021). Frequently observed features include macrosomia, macroglossia, umbilical hernia, organomegaly, placentomegaly, hydrallantois, increased gestation length, and increased dystocia rate (Behboodi et al., 1995; Chen et al., 2013; Farin and Farin, 1995; Farin et al., 2001; Hasler et al., 1995; Hori et al., 2010; Kruij and Den Daas, 1997; Lazzari et al., 2002; McEvoy et al., 1998; Sinclair et al., 1995; van Wagtenonk-de Leeuw et al., 1998). All of the bovine LOS reported in the literature have involved the use of *in vitro* production procedures (i.e. *in vitro* maturation, fertilization and culture) or nuclear transfer

(hereunto referred to as ART for assisted reproductive technologies) (Behboodi et al., 1995; Chen et al., 2013; Farin and Farin, 1995; Farin et al., 2001; Hasler et al., 1995; Hori et al., 2010; Kruij and Den Daas, 1997; Lazzari et al., 2002; McEvoy et al., 1998; Sinclair et al., 1995; van Wagtenonk-de Leeuw et al., 1998). ART is known to induce errors in the epigenome including DNA methylation and genomic imprinting to offspring in humans and ruminants (Fauser et al., 2014; Urrego et al., 2014). We recently reported that LOS can occur spontaneously (Li et al., 2019b; Rivera et al., 2021), a phenomenon that in some cases may have been incorrectly ascribed to the sire's genetics (Coleman et al., 2021; Thompson et al., 1998). Currently, there is a lack of documented incidence for both spontaneous LOS (SLOS) and ART associated LOS (ART-LOS) from the industry, although those experiencing them in their farm or practice incur steep financial losses (Rivera et al., 2021).

We and others have reported that ART-LOS is an epigenetic disorder (Chen et al., 2015, 2017) with global alterations of transcriptome and methylome, changes in chromosomal architecture, and loss-of-imprinting at multiple imprinted domains including *IGF2R*, *KCNQ1*, *IGF2*, *PLAGL1*, *PEG3*, and *DLK1* (Chen et al., 2013, 2015, 2017; Hori et al., 2010; Li et al., 2019a, 2022; Sangalli et al., 2014; Su et al., 2011a, 2011b). Although we recently documented that LOS occurs spontaneously, at least based on phenotypes (Li et al., 2019b; Rivera et al., 2021), no data exist to demonstrate that the spontaneous overgrowth

syndrome shares epigenotype with the ART-induced LOS.

Beckwith-Wiedemann syndrome (BWS, OMIM #130650) is the most common congenital overgrowth syndrome in humans. The incidence of BWS is approximately 1 in 10,340 live births and children conceived with the use of ART have a 10.7 relative risk of suffering from BWS (Mussa et al., 2013, 2017).

Clinical features frequently observed in BWS include macrosomia, macroglossia, abdominal wall defects (umbilical hernia/exomphalos), lateralized overgrowth, increased tumor incidence, hyperinsulinism, facial naevus simplex, ear malformation, organomegaly, and placentomegaly (Brioude et al., 2018).

Molecular defects found in BWS include global alteration of transcriptome and methylome, changes of chromosomal architectures, loss-of-imprinting at imprinted domains including *IGF2*, *KCNQ1*, *IGF2R*, *PLAGL1*, *PEG3*, *PEG10*, *GRB10*, *MEST*, *DLK1*, *IGF1R*, and *GNAS* (Brioude et al., 2018; Krzyzewska et al., 2019; Li et al., 2019a; Naveh et al., 2021; Rossignol et al., 2006; Rovina et al., 2020; Tee et al., 2013; Tenorio et al., 2016). In addition, a subset of BWS are the result of secondary epimutations (genetic defects) which result in loss-of-imprinting (Cooper et al., 2005; Mussa et al., 2016). We have shown that ART-LOS shares phenotypes and molecular aberrations with BWS (Brioude et al., 2018; Li et al., 2019a).

Given the phenotypic similarities between the spontaneous and the ART-induced syndromes we hypothesized that SLOS has similar methylome

epimutations as ART-LOS. In this study, we performed whole genome bisulfite sequencing in bovine tissues of control, SLOS and ART-LOS to identify conserved signatures of this syndrome. In addition, we included data from two previous LOS studies to identify shared vulnerable genomic loci in LOS. Overall, we identified 320 genomic loci in LOS that have alterations in DNA methylation when compared to controls. Specifically, there are 25 highly vulnerable loci in LOS, including *DMRT2*, *TBX18*, *IGF2R*, *PRDM8*, and *BLCAP/NNAT*. We also observed tissue-specific DNA methylation patterns between muscle and blood, and conservation of ART-induced DNA methylation changes between muscle and blood. We conclude that as ART-LOS, SLOS is an epigenetic condition. In addition, SLOS and ART-LOS share similarities in methylome epimutations.

3.3 Results

Animal information and phenotypes

In total, 26 animals were included in this study and were assigned to different groups (Table 1). The US_Control group contains three AI conceived Holstein breed neonate calves of average weight and with no clinical abnormalities and serves as control for other animals from the United States. The US_SLOS group contains eight SLOS calves found in the United States and the observable phenotypic abnormalities include macrosomia, macroglossia, and

abdominal wall defects (Table 1 and Figure 1). The dam, sire, and sibling of US_SLOS_#6 showed no clinical abnormalities and are included for analyses to determine whether there exist inheritable methylation-specific causal effects of LOS or not.

The ES_Control group contains three AI conceived calves with no clinical abnormalities identified and serves as control for other animals from Spain. The ES_ART group contains four ART conceived calves with no clinical abnormalities except one had macroglossia (Figure 1). The ES_RF group contains three calves conceived by ART supplemented reproductive fluids with one having some clinical abnormalities. Last, the ES_RF_necropsy group contains two dead calves from the ES_RF group with typical LOS clinical features.

Genomic context of differentially methylated regions in SLOS calves

Whole genome bisulfite sequencing (WGBS) identified 2,839 differentially methylated regions (DMRs) in US_SLOS muscle samples when compared with US_Control muscle samples, namely US_SLOS_muscle_DMR, and ~ 66% of them were hypomethylated (Figure 2. A-C and Table S1. A). Hypomethylated DMRs showed significant enrichment for promoters, CpG islands, CpG shores, and predicted CTCF binding sites than random distribution within the genome, but were depleted from gene bodies (Figure 2. C).

Hypermethylated DMRs only showed significant depletion from CpG islands. Of note, in this study we only included the promoters of protein coding genes and long non-coding RNAs (lncRNAs) since the location of promoters for small ncRNA are not well characterized in bovine (Li et al., 2010). We also compared the DMRs identified here with those previously published for ART-LOS skeletal muscle and skin fibroblast cells (Chen et al., 2017; Li et al., 2022) and identified an overlap of 22 and 134 DMRs, respectively (Figure 2. A-B). Due to the lack of proper control samples for tongue and ear tissues, a separate comparison was conducted by combining US_SLOS muscle, ear, and tongue samples and compared them to US_Control muscle samples. Similar results were observed as the muscle comparison (Figure S1. A-C and Table S1.B).

Genomic context of differentially methylated regions in ART-LOS calves

In total, 1,552 DMRs were identified in ES_RF_necropsy muscle samples when compared with ES_Control muscle, namely ES_RF_necropsy_muscle_DMR, and like US_SLOS_muscle_DMR, ~ 66% of DMRs were hypomethylated (Figure 2. D-F and Table S1. C). These hypomethylated DMRs showed a significant enrichment for gene bodies which was different from US_SLOS results, and also enriched for CpG islands, shores, and shelves (Figure 2. F). Hypermethylated DMRs were significantly enriched for

promoters, repetitive sequences, CpG shores, and shelves, and depleted from exons.

DNA methylation changes among different tissue, breeds, and developmental stages

In order to identify molecular markers for LOS, we searched for vulnerable loci regardless of direction of DNA methylation changes (i.e. hypomethylated or hypermethylated) when compared to controls. For this analysis we included four independent experiments, the above mentioned US_SLOS muscle and ES_RF_necropsy muscle, and our two previously published datasets, namely Li_LOS_fibroblast and Chen_LOS_muscle (Chen et al., 2017; Li et al., 2022). It should be noted that the raw data for the Chen_LOS_muscle experiment were reanalyzed with methods described here. The enrichment of genomic context for DMRs identified in these two previously published datasets can be found in Figure S2. In total, four loci were found to be vulnerable in all four experiments, 21 loci were found in three of the four experiments, and 295 loci were found in two of the four experiments (Figure 3 and Table S2). Overall, the vulnerable loci found in ≥ 3 experiments were enriched for CpG islands and CpG shores (Figure 3). The DNA methylation level and coverage for several of the LOS-associated vulnerable loci are illustrated in

Figures 4 and 5 and Figures S3-S5.

The disruption of *IGF2R* imprinted domain has been frequently reported in LOS and BWS and we identified nine vulnerable loci within this domain (Figure 3 and Table S2). However, although there is ~20% and ~40% reduction in DNA methylation levels in the imprinting control region (ICR) of *IGF2R* in the US_SLOS muscle and ES_RF_necropsy muscle, respectively (Figure 5.B), hypomethylation is not reported for these samples because of the read coverage being lower than the cutoff used in this study. In addition to the imprinted genes shown in Figure 5, several other imprinted genes known in bovine or other species overlap with vulnerable loci, including *SGCE* (4_12059801_12060360), *PRKN* (9_97743481_97744740), *GNAS* (13_57485061_57485680 and 13_57520861_57521040), *KBTBD3* (15_1648961_1650300), *TSHZ3* (18_41905361_41906000), *HOXB3* (19_37915781_37915960 and 19_37922621_37923120), *INPP5F* (26_39936921_39938980), and *KCNQ1* (29_48957301_48958880; Table S2). These loci mainly located within introns of imprinted genes, except that *GNAS* and *HOXB3* also included their promoters.

A case study for DNA methylation at LOS-associated vulnerable loci in a SLOS calf, its sire, dam, and full-sibling

To determine each parent's impact on LOS development, we conducted

a case study for DNA methylation at LOS-associated vulnerable loci in US_SLOS_#6 and its sire, dam, and full-sibling. When comparing the level of DNA methylation in blood of the 25 highly vulnerable loci between the sire, dam, and full-sibling against the average of US_control blood samples, the dam showed a trend ($p=0.085$; mean \pm SD = -11.59 ± 19.23) of overall lower DNA methylation level than the sire (-3.09 ± 13.77 ; Figure 6.A). In addition, for all 320 LOS-vulnerable loci, this trend still exists between dam (-1.79 ± 17.17) and sire (0.47 ± 15.1 ; Figure S6). Due to the lack of comparable tissue samples and limited sample number, we only draw plots for visual examination of the trend of DNA methylation changes without statistical tests for individual vulnerable loci (Figure 6. B-P). Nine loci showed differences ($>10\%$) of DNA methylation in parental blood samples when compared to the mean of the control group, including one in sire only (Figure 6. G), two in both the sire and dam (Figure 6. H and I), and six in dam only (Figure 6. J to O). The higher number of LOS-vulnerable loci with altered DNA methylation level in the dam than in the sire suggests a higher proportion of maternal contribution to the SLOS development in the US_SLOS_#6.

Reproductive fluid supplementation partially improved methylome outcomes of ART

In total, 857 DMRs were found between ES_ART muscle and ES_Control muscle, namely ES_ART_muscle_DMR, with a bias (~84%) towards hypermethylation (Figure 7. A-C and Table S3. A). The hypermethylated DMRs showed significant enrichment for promoters, gene bodies, CpG islands, shores, and shelves, and hypomethylated DMRs were enriched for CpG shores and shelves (Figure 7. C).

The supplementation of reproductive fluids during ART resulted in 419 DMRs identified between ES_RF muscle and ES_Control muscle, namely ES_RF_muscle_DMR, with an equal ratio for hyper and hypomethylation (Figure 7. D-F and Table S3. B). Hypermethylated DMRs showed similar patterns of genomic context enrichment as ART groups. However, hypomethylated DMRs showed enrichment for CpG islands, shores, shelves and gene bodies (Figure 7. F). When comparing the DMRs between ES_ART_muscle_DMR and ES_RF_muscle_DMR, 62 were shared and had similar hypo or hypermethylation (Table S3. C).

Tissue specific DNA methylation pattern between muscle and leukocytes

Overall, the genome of muscle is hypomethylated when compared to

leukocytes. There were a total of 25,466 and 9,961 DMRs between blood and muscle for US_Control and ES_Control, respectively, and ~90% of them were hypomethylated in muscle (Figure S7. A-F and Table S4. A-B). In addition, 5,169 DMRs were shared by these two different breeds of cattle and had similar direction of change in muscle (Table S4. C).

Conservation of ART induced DNA methylation between muscle and leukocytes

In total, 591 DMRs were identified when comparing ES_RF blood samples to ES_Control blood, namely ES_RF_blood_DMR, and ~88% of these DMRs show hypermethylation (Figure S8. A-C and Table S5. A). This pattern was not similar to ES_RF_muscle_DMR but resembled ES_ART_muscle_DMR (Figure 7). When comparing the DMRs identified in ES_RF blood and muscle, 38 were found to be shared with 37 having the same direction in methylation change (Table S5. B). In addition, when comparing between ES_RF_blood_DMR and ES_ART_muscle_DMR, 16 DMRs were found to be shared and all had the same direction of DNA methylation (Table S5. C). Several LOS-vulnerable loci that show conserved DNA methylation changes between tissues in ES_RF group are illustrated in Figure 8. Additionally, similar DNA methylation changes were observed for these DMRs in the blood sample of ES_ART_#2 which has

macroglossia (Figure 8).

3.4 Discussion

In this study, we observed typical LOS/AOS/BWS clinical abnormalities from SLOS calves, including macrosomia, macroglossia, and abdominal wall defects, and some atypical features. Spontaneous BWS (SBWS) shows no differences on the frequency of symptoms including macroglossia, hemihyperplasia, abdominal wall defects, hypoglycemia, but a significantly higher frequency of ear malformation than ART-induced BWS (ART-BWS) (Tenorio et al., 2016). In addition, SBWS patients have significantly longer gestational age and heavier birth weight than ART-BWS (Tenorio et al., 2016). Studies with larger sample size are needed to draw conclusions for these types of frequencies for SLOS.

From the analyses of DMR distribution over various genomic contexts, we observed different preferences between LOS-associated hyper- and hypomethylated DMRs, and similarities and differences between SLOS and ART-LOS. As a regulatory element, CpG islands and shores are enriched for enhancers in human, and the activity of enhancers is regulated by DNA methylation (Moran et al., 2016; Plank and Dean, 2014). For both US_SLOS_muscle_DMR and ES_RF_necropsy_muscle_DMR, the observed

frequencies at CpG islands and shores are higher than expected, and with increased hypomethylation preference. This enrichment of hypomethylated DMRs over CpG shore resembles the observation of cancer-specific DMRs in human which are associated with cell proliferation and growth (Irizarry et al., 2009). It is well known that gene expression is negatively correlated with promoter DNA methylation level (Vincent et al., 2011). Both US_SLOS_muscle_DMR and ES_RF_necropsy_muscle_DMR showed higher frequencies overlapping promoter regions than expected, but the former had increased hypomethylation. DNA methylation level of gene bodies reflects expression level (Jjingo et al., 2012). Interestingly, US_SLOS_muscle_DMR and ES_RF_necropsy_muscle_DMR showed opposite trend of hypomethylated DMR enrichment in gene body, which suggests differences in global gene expression level.

When comparing ES_ART_muscle_DMR and ES_RF_muscle_DMR, we found the supplementation of the culture medium with reproductive fluids largely reduced the number of hypermethylated DMRs caused by ART. However, only ~15% of the DMRs of ES_RF_muscle_DMR were shared with ES_ART_muscle_DMR, indicating the supplementation of reproductive fluids also induced new changes in the methylome, which we have previously reported (Canovas et al., 2017). Next, when adding ES_RF_blood_DMR into this comparison, we found that although the distribution pattern of

ES_RF_blood_DMR resembled ES_ART_muscle_DMR, ES_RF_blood_DMR still shared more DMRs with ES_RF_muscle_DMR instead of ES_ART_muscle_DMR. This indicates that the progenitors of muscle and blood have cell type-specific response to the supplementation of reproductive fluids. Nevertheless, the shared DMRs were very consistent in the direction of changes, which suggests that a proportion of identified DMRs could be used as diagnostic biomarkers in blood for muscle.

Several of the LOS-vulnerable loci were found close to the promoter of genes, including 7_2579941_2581100, 8_43514181_43518080, and 9_64659781_64660620. The frequently hypomethylated loci 7_2579941_2581100 resides in a region enriched for histone proteins including H2A.W histone (*H2AW/HIST3H2A*), H2B.U histone 1 (*H2BU1/HIST3H2BB*), and H3.4 histone (*H3-4/LOC518318*). However, the transcript level of these histone genes were barely detected from our previous RNA-seq results of fibroblast cells and muscle of ART-LOS fetuses (Chen et al., 2015; Li et al., 2022), although that may be indicative of different developmental stages (d105 fetuses vs. newborn calves). 8_43514181_43518080 covers the promoter of protein coding gene doublesex and mab-3 related transcription factor 2 (*DMRT2*) and was detected both hypo- and hypermethylation in different experiments. *DMRT2* is a polycomb associated transcription factor and is known to be regulated by promoter DNA methylation (Daino et al., 2018; Renner et al., 2013). Interestingly, significant

downregulation of *DMRT2* transcript was reported in muscle of ART-LOS fetuses although no DNA methylation differences detected in the corresponding samples (Chen et al., 2015). 9_64659781_64660620 is one of the four most frequent LOS-vulnerable loci mainly showing hypermethylation and located close (1.3 kb) to the transcription start site of gene T-box transcription factor 18 (*TBX18*). As a critical transcription factor during embryo development in various tissues, *TBX18* can also be regulated by promoter DNA methylation (Christoffels et al., 2009; Haraguchi et al., 2015). The downregulation of *TBX18* transcript has been reported in muscle of ART-LOS fetuses (Chen et al., 2015).

We have shown that ART-LOS, like BWS is a global loss-of-imprinting disorder. Several LOS-vulnerable loci were found within bodies of imprinted genes and overlapped with CpG islands. 6_94882141_94883160 overlapped a CpG island within the third intron of PR/SET domain 8 (*PRDM8*). *PRDM8* is a histone methyltransferase and can inhibit cell proliferation through PI3K/AKT/mTOR signaling pathway thus functioning as a tumor suppressor (Chen et al., 2018). Although not completely confirmed, *PRDM8* is considered as a candidate of imprinted genes and the predicted ICR in human located in its last exon (Bina, 2020; Daelemans et al., 2010). The gene structure of *PRDM8* is highly conserved between human and bovine, thus this locus' overlapped CpG island is not likely to be the ICR (Bina, 2020). Accordingly, *PRDM8* was not found to be misregulated in fibroblast cells nor in muscle of ART-LOS fetuses (Chen et

al., 2015; Li et al., 2022). The hypomethylated DMR 13_66465461_66466640 localizes within the second intron of imprinted gene *BLCAP* apoptosis inducing factor (*BLCAP*) and covers most of another imprinted gene neuronatin (*NNAT*) (Schulz et al., 2008). *BLCAP* is known as a tumor suppressor through inducing cell cycle arrest and apoptosis (Yao et al., 2007). *NNAT* is a proteolipid that regulates calcium channels (Nass et al., 2017). Increased expression of *NNAT* is often found in tumor development, including the Wilms tumor of kidney, and associated with poor outcomes of patients (Hubertus et al., 2013; Nass et al., 2017). Similarly, *NNAT* showed significant upregulation in both fibroblast cells and muscle of ART-LOS fetuses (Chen et al., 2015; Li et al., 2022). Interestingly, the proposed model of imprinting regulation in human at this locus relies on CTCF binding within the second intron of *BLCAP* (Hubertus et al., 2013). However, there is no putative CTCF binding sites predicted in bovine based on vertebrate CTCF motifs, which suggests either there is undiscovered unique motif in bovine, or the mechanism of regulation is not conserved in bovine.

As previously mentioned, *IGF2R* imprinted domain contains the highest number of LOS-vulnerable loci, which is nine. These hypermethylated loci in LOS are located within the first four introns of *IGF2R* surrounding (i.e. not including) the ICR. This ICR is the promoter of lncRNA *AIRN* and normally its methylated state on the maternal allele prevents *AIRN*'s expression and allows *IGF2R* expression (Sleutels et al., 2002). On the contrary, an unmethylated ICR on the

paternal allele allows the expression of *AIRN* which silences *IGF2R* by attracting Polycomb repressive complexes to the locus (Schertzer et al., 2019).

Hypomethylation of *IGF2R* ICR occurs frequently in LOS, but the low read coverage prevented us to include it in the list of vulnerable loci although we observed similar results regardless of coverage (Chen et al., 2017; Smith et al., 2015; Young et al., 2001). Compared to Chen_LOS_muscle, the decreased read coverage at *IGF2R* ICR in the other three LOS experiments is likely caused by differences in the process of sequencing library preparation. Further studies are needed to determine the reasons of this inconsistency of sequencing results of this region. DNA methylation level of gene bodies is associated with transcription frequency in a parabolic pattern that the most highly and lowly expressed genes have low level of methylation but genes with intermediate level of expression have high methylation level (Jjingo et al., 2012). This pattern matches our observation for the hypermethylated DMRs in *IGF2R* gene body and *IGF2R* transcripts were downregulated in both fibroblast cells and muscle of ART-LOS fetuses (Chen et al., 2015; Li et al., 2022). For example, the *IGF2R* expression ranked 127 (top 0.7%) in the control group (~880 counts per million reads (cpm)) of fibroblast cells and decreased by ~3.5 folds to ~260 cpm in LOS group which ranked 581 (top 3%) (Li et al., 2022).

Among the four most frequently LOS-vulnerable loci,

6_66245821_66247640 is the only one located within a gene body that does not

overlap with a CpG island. This locus covers the 20th exon and surrounding 19th and 20th introns of ATPase phospholipid transporting 10D (*ATP10D*) gene and is hypomethylated in all four LOS experiments. *ATP10D* functions in the modulation of high density lipoprotein and is associated with susceptibility of obesity under high fat diet in mice studies (Sigrüener et al., 2017). Our previous RNA-seq results did not show misregulation of *ATP10D* transcript in either fibroblast cells or muscle of LOS fetuses (Chen et al., 2015; Li et al., 2022). For the other two LOS-vulnerable loci found in four LOS experiments, namely 4_102068961_102070360 and 9_54438201_54439220, they always show hypomethylation at intergenic regions covering CpG islands. Further studies are needed to determine whether these loci serve as remote regulatory elements for gene expression.

In human, although the hierarchical cluster analyses based on DNA methylation status at imprinted/non-imprinted genes cannot completely separate ART-BWS and SBWS groups, the two group still show different preferences of DNA methylation changes for imprinted domains including *PEG10*, *MEST*, *GNAS*, *PLAGL1*, and *IGF2R* (Tee et al., 2013; Tenorio et al., 2016). Similarly, we also observed inconsistency of DNA methylation disruption between SLOS and ART-LOS at some imprinted domains, including *NNAT* and *IGF2R*. Additionally, SBWS is associated more with genetic defects including changes of chromosomal contents and gene mutations when compared to ART-BWS

(Tenorio et al., 2016). Further studies on DNA sequencing of LOS are needed to investigate if there is a genetic contribution to the susceptibility of LOS/AOS development.

Finally, we did a comparison of the methylome of a SLOS, namely US_SLOS_#6, with its relatives (dam, sire, and full-sibling) to determine if the epimutations were inherited or occurred de novo in the offspring. We identified that some of the DMRs may have been inherited through the maternal or paternal genomes, although the dam seems to contribute more to the abnormal offspring's methylome. While some of the differences detected may be breed specific, it appears that the abnormalities in the SLOS may be partly due to the higher number of epimutations inherited from the parents as its full-sibling was born healthy and of normal size, even though it shares some inherited epimutations.

In summary, unique patterns of distribution over different genomic contexts were observed for DMRs as a result of ART, reproductive fluid supplementation of culture media, ART-LOS, and SLOS. Hundreds of LOS-vulnerable loci determined in this study could serve as molecular markers for the diagnosis of LOS. Further studies are needed to determine the level of conservation of these DMRs in other tissue types of LOS fetuses that could be used for early diagnosis, such as amniotic fluid. In conclusion, alterations of epigenome are involved in the etiology of SLOS with certain levels of similarities

to ART induced LOS.

3.5 Materials and Methods

All the chromosomal coordinates in this manuscript refer to bovine genome assembly ARS-UCD1.2 (Rosen et al., 2020).

Animal tissues

Blood and tissue samples of animals from the United States (US) and Spain (ES) were used in this study (Table 1). Control animals from the US (US_Control) were conceived by artificial insemination (AI) at the University of Missouri Foremost Dairy Research Center and sacrificed immediately upon birth by a trained Veterinarian for blood and tissue collection. The three Holstein breed neonates were male, of average birth weight and without any abnormal phenotypes. Blood samples were collected from the jugular vein using K3EDTA vacutainers (BD) and processed as described by Ortega et al. (Ortega et al., 2018). Tissues were dissected, diced, sealed in aluminum foil pockets, snap frozen in liquid nitrogen, and stored at -80 °C.

SLOS animals were from various parts of the US and were stillborn or died within several hours/days after birth (US_SLOS). Tissue samples of SLOS animals were collected from carcass by their owners, veterinarians, or our

collaborators and shipped to University of Missouri and we do not have any information other than body weight, breed, sex (for most), and clinical features (for some). US_SLOS_#6 was processed by us and donated by a farmer from a nearby town so in this case we know that the stillborn calf had been conceived by natural service. We also collected blood samples of the dam, sire, and full-sibling of US_SLOS_#6 for methylome comparisons (i.e., case study).

Animals from Spain were generated as described previously (Lopes et al., 2020). Briefly, the control animals (ES_Control) were conceived by AI using frozen-thawed semen from one bull (Asturian Valley breed) among synchronized cows (Holstein breed) on the day of presumptive estrus. *In vitro* produced animals were generated using slaughterhouse oocytes (crossbred Limousin and Charolais) and semen from the same bull as controls. Following fertilization, embryos were separated in two different groups: one culture group (ES_ART) composed of synthetic oviductal fluid (SOF) media supplemented with bovine serum albumin during the 7-8 days of culture and another group (ES_RF) composed of SOF media supplemented with bovine oviductal fluid (NaturARTs-BOF-EL, EmbryoCloud, Spain) for the first 4 days and bovine uterine fluid (NaturARTs-BUF-ML, EmbryoCloud) for the following days. Embryos (blastocysts and expanded blastocysts) were vitrified on day 7 or 8 of culture and stored until use. Recipients (Holstein cows) were synchronized and on day 6 to 8 after presumptive estrus, each cow received one thawed embryo. After parturition,

calves were immediately assessed for general health parameters and continued to be monitored throughout their lives. Calves that did not survive parturition or died were collected for necropsy. Simultaneous blood and muscle samples were collected in two different days and calves' age ranged between 71 and 292 days (mean age 167 days and median 138). Necropsy muscle samples do not have corresponding blood sample and the age of the calves varied between 0 (at birth) and 13 days (mean age 5 days and median 2). Blood samples were collected from the jugular or coccygeal vein (according to the size of the animal) using EDTA tube (BD vacutainer, BD, Spain) and stored at 4°C. Samples (less than 2h after collection) were then aliquoted in 300 µL and 900 µL of Tris-HCl solution was added. The content was mixed and centrifuged at 14500 x g for one minute and the supernatant discarded. The procedure was repeated twice more and the final pellet was submerged in liquid nitrogen and stored at -80°C. Muscle biopsies were performed using a semiautomatic needle (ML18160, RI.MOS., Italy). Surgical preparation prior biopsy included minor restraint of the animal, shaving of the area, cleaning, and application of local anesthesia (lidocaine). The incision on the gluteus medius was ~1cm long, enough for the biopsy needle to pass. Samples were immediately collected and placed on ice. The incision was closed, and calves were monitored for any sign of infection. Samples were transported to the laboratory (less than 2h after collection), frozen in liquid nitrogen and stored at -80 °C. Blood and muscle samples were shipped on dry

ice to the University of Missouri.

Ethics approval

US_Control animals were purchased from the University of Missouri Foremost Dairy Research Center and euthanized by veterinarians. All the animal procedures were approved by University of Missouri Animal Care and Use Committee under protocol 9455.

Animals from Spain were handled by veterinarians following the Spanish Policy for Animal Protection RD 53/2013, which meets European Union Directive 2010/63/UE on animal protection. The Ethics Committee of Animal Experimentation of the University of Murcia and the Animal Production Service of the Agriculture Department of the Region of Murcia (Spain) (ref. no. A132141002) approved the procedures performed for these animals.

Genomic DNA extraction

Blood and tissue samples were lysed in lysis buffer (0.05 M Tris-HCl (pH 8.0), 0.1 M EDTA, and 0.5% (w/v) SDS) with proteinase K (Fisher BioReagents, BP1700) at 55°C for four hours (blood) or overnight (tissue). Genomic DNA was extracted with Phenol:Chloroform:Isoamyl Alcohol (SIGMA, P3803) following the manufacturer's instructions. The concentration of DNA was measured by using a

NanoDrop® ND-1000 Spectrophotometer (Thermo Fisher Scientific) and DNA integrity was confirmed by electrophoresis on a 0.7% agarose gel. Genomic DNA samples were stored at -20°C.

Whole genome bisulfite sequencing (WGBS) and data analyses

WGBS was conducted by CD Genomics (New York, United States). Information on library preparation and sequencing obtained from the company is as follows: For WGBS library preparation, 1 ug of genomic DNA was fragmented by sonication to a mean size of approximately 200-400 bp. Fragmented DNA was end-repaired, 5'-phosphorylated, 3'-dA-tailed and then ligated to methylated adapters. The methylated adapter-ligated DNAs were purified using 0.8× Agencourt AMPure XP magnetic beads and subjected to bisulfite conversion by ZYMO EZ DNA Methylation-Gold Kit (zymo). The converted DNAs were then amplified using 25 µl KAPA HiFi HotStart Uracil+ ReadyMix (2X) and 8-bp index primers with a final concentration of 1 µM each. The constructed WGBS libraries were then analyzed by Agilent 2100 Bioanalyzer and quantified by a Qubit fluorometer with Quant-iT dsDNA HS Assay Kit (Invitrogen), and finally sequenced on Illumina HiSeq X ten sequencer. 0.1-1% lambda DNA were added during the library preparation to monitor bisulfite conversion rate.

For WGBS data analyses, duplicated reads generated during PCR and

sequencing were removed from raw sequencing reads using the clumpify function of BMap 38.90 (Bushnell, 2021). The remaining raw reads were trimmed for adapter sequences and low quality bases using trimmomatic 0.39 (Bolger et al., 2014) with parameters 'ILLUMINACLIP:adapter_seq:2:30:10:1:true LEADING:20 TRAILING:20 AVGQUAL:20 MAXINFO:0:0.5'. Trimmed reads were aligned to the bovine genome using bismark 0.23.0 (Krueger and Andrews, 2011) with parameters '-X 900 --unmapped --ambiguous --non_bs_mm'. Trimmed reads were also aligned to lambda phage genome to determine bisulfite conversion rates. Samtools 1.13 (Li et al., 2009) was used to convert, sort, filter, and index bam files. MarkDuplicates function of picard 2.25.5 (Broad Institute, 2021) was used to further remove duplicated reads after alignment. Read groups were added for each samples using AddOrReplaceReadGroups function of picard. The dataset of known variants in bovine, namely ARS1.2PlusY_BQSR_v3.vcf.gz, was acquired from the 1000 bull genome project (Hayes and Daetwyler, 2019) and served as reference to identify genomic variants in WGBS data. Indel realignment was performed using RealignerTargetCreator and IndelRealigner functions of BisSNP 1.0.1 (Liu et al., 2012). Base quality recalibration was carried out using BisulfiteCountCovariates and BisulfiteTableRecalibration functions of BisSNP 0.82.2 since these functions are missing in version 1.0.0 and 1.0.1. Parameters used for BisulfiteCountCovariates were '-cov ReadGroupCovariate -cov QualityScoreCovariate -cov CycleCovariate -baqGOP 30'. Genomic variants

were identified using BisSNP 1.0.1 with default setting expect that '-bsRate' was changed to bisulfite conversion rate observed from lambda phage genome alignment for each sample. BisSNP identified variants were filtered by its VCFpostprocess function with parameter '-windSizeForSNPfilter 0'. Additionally, genomic variants were identified using BS-SNPer 1.0 (Gao et al., 2015) with parameters '-minhetfreq 0.1 --minhomfreq 0.85 --minquali 15 --mincover 5 --maxcover 1000 --minread2 2 --errorate 0.02 --mapvalue 20'. M-bias plots were generated using bismark and the first 3 bases of R1 reads and the first 4 bases of R2 reads showed biased CpG methylation level, thus these bases were excluded from downstream analyses. CpG methylation information were extracted from the bam files using bismark_methylation_extractor function of bismark with parameters '-p --ignore 3 --ignore_r2 4 --comprehensive --no_header -gzip --bedGraph --buffer_size 50% --cytosine_report'. Statistical analyses were conducted using R package hummingbird (Ji, 2019) with parameter 'minCpGs = 10, minLength = 100, maxGap = 300' to identify differentially methylated regions (DMRs) in various comparisons. DMRs with at least 15% difference in methylation level (both gain and loss of methylation) and at least 4 mean read coverage at CpG sites were reported. The sex chromosomes were not analyzed to circumvent confounding created by X chromosome inactivation associated DNA methylation.

In addition, WGBS data from two of our previous LOS experiments were

analyzed with the same methods mentioned above (Chen et al., 2017; Li et al., 2022). The GEO accession numbers for these data are GSE93775 and GSE197130.

Analyses of overlapping between DMRs and genomic contents

Information of gene annotation was obtained from NCBI (GCF_002263795.1_ARS-UCD1.2_genomic.gff) (O’Leary et al., 2016). Repeated and overlapped exons were merged for each gene, and introns were calculated based on merged exons. Promoters (1kb) were calculated based on transcription start sites annotation and only included protein coding genes and long non-coding RNAs. Annotation of CpG islands and repeated sequences were obtained from UCSC Genome Browser (Kent et al., 2002). Locations of CpG shores (flanking 2kb from CpG islands) and shelves (flanking 2-4kb from the CpG island) were calculated based on CpG island annotation. Potential CTCF binding sites were predicted as previously reported (Li et al., 2022). Bedtools and custom Perl scripts were used for these analyses to identify overlapped genomic location and make tables (Quinlan and Hall, 2010). R package Sushi, circular, and ggplot2 were used for making figures (Lund et al., 2017; Phanstiel et al., 2014; Wickham, 2011).

Sequencing data availability statement

The raw sequencing reads of WGBS used in this study are available in the GEO database with accession numbers (GSE199084).

3.6 Acknowledgements

The work in the US was supported by Agriculture and Food Research Initiative (grant AFRI - 2018-67015-27598) and Dr. Roger L. Morrison Scholarship from the College of Agriculture, Food & Natural Resources (to Yahan Li). We thank Astrid Roshealy Brau for technical assistance and Dr. Darren Hagen, Dr. Callum Donnelly and Mr. Stephenson for providing pictures and samples. In addition, we would like to thank Dr. Dietrich Volkmann, Professor in Theriogenology, College of Veterinary Medicine, University of Missouri for recognizing the US_SLOS_#6 and arranging for us to collect the tissues and Dr. Fred Williams III for helping with the tissue collection and identification of malformations.

The work in Spain was funded by European Union, Horizon 2020 Marie Skłodowska-Curie Action (Ref. REPBIOTECH 675526), by the Spanish Ministry of Science and Innovation (MCIN/AEI/10.13039/501100011033/ and FEDER, ref. I+D+I PID2020-113366RB-I00) and Fundación Séneca, Murcia, Spain (ref. 20040/GERM/16). We thank Dr. Rafael Latorre for aiding in biopsy collection and

the Physiology of Reproduction group for support with the animals.

Table 1. Information of calves in this study.

Sex with * = predicted sex based on WGBS reads alignment to chromosome Y (lack of sex information from original owners/sample providers). BW = birth weight. Used ID = animal ID used in previous publications.

Origin	Group	Animal	Breed	Sex	BW (kg)	Tissue	Clinical diagnosis	Used ID
United States	US_Control	US_Control_#1	Holstein	Male	-	Blood, muscle	None	-
		US_Control_#2	Holstein	Male	-	Blood, muscle	None	-
		US_Control_#3	Holstein	Male	-	Blood, muscle	None	-
	US_SLOS	US_SLOS_#1	Angus	Male	58	Ear	Macrosomia	-
		US_SLOS_#2	-	Male*	-	Muscle	N/A	-
		US_SLOS_#3	Simmental	Male*	54	Tongue	Macrosomia	-
		US_SLOS_#4	Hereford	Female*	61	Muscle	Macrosomia	-
		US_SLOS_#5	Angus	Female	-	Muscle	Ventral abdominal wall: cloacal exstrophy, omphalocele, diastasis recti; heart: patent ductus arteriosus, patent foramen ovale	-
		US_SLOS_#6	Charolais	Male	77	Muscle, tongue	Macroglossia, macrosomia	-
		US_SLOS_#7	Shorthorn crossbred	Female	44	Muscle, tongue	Dexamethasone and prostaglandin induced delivery 16 days preterm	-
		US_SLOS_#8	Shorthorn crossbred	Male	58	Muscle, tongue	Dexamethasone and prostaglandin induced delivery 16 days preterm, macroglossia, macrosomia, microtia, hardened umbilicus	-
US_SLOS_dam	US_SLOS_#6_dam	Charolais	Female	37	Blood	None	-	
US_SLOS_sibling	US_SLOS_#6_sibling	Charolais	Male	41	Blood	None	-	
US_SLOS_sire	US_SLOS_#6_sire	Charolais	Male	-	Blood	None	-	
Spain	ES_Control	ES_Control_#1	Asturian Valley x Holstein (AxH)	Male	32	Blood, muscle	None	AI 1
		ES_Control_#2	AxH	Male	30	Blood, muscle	None	AI 4
		ES_Control_#3	AxH	Male	45	Blood, muscle	None	AI 5
	ES_ART	ES_ART_#1	Asturian Valley x Limousin crossbred (AxL)	Male	35	Muscle	None	BSA 1
		ES_ART_#2	AxL	Female	46	Blood	Macroglossia, hypersensitivity to needles	BSA 3
		ES_ART_#3	AxL	Male	33	Muscle	None	BSA 4
		ES_ART_#4	AxL	Male	33	Muscle	None	BSA 8
	ES_RF	ES_RF_#1	AxL	Male	29	Blood, muscle	Heart murmur, trouble standing for the first three days	RF 5
		ES_RF_#2	AxL	Male	49	Blood, muscle	None	RF 7
		ES_RF_#3	AxL	Male	33	Blood, muscle	None	RF 8
ES_RF_necropsy	ES_RF_necropsy_#1	AxL	Male	56	Muscle	Stillbirth, dystocia, macrosomia, enlarged head, limb hyperflexion, cardiomegaly	RF 2	
	ES_RF_necropsy_#2	AxL	Female	56	Muscle	Died after birth due to spine fracture, dystocia, macrosomia, limb hyperflexion, cardiomegaly	RF 6	



Figure 1. Example of phenotypic abnormalities of SLOS and ART-LOS calves.

(A) Abdominal wall defect of US_SLOS_#5 (Angus breed). This spontaneous LOS calf was born alive and had to be euthanized due to the body wall malformation. (B and C) Macrosomia and macroglossia of US_SLOS_#6 (Charolais breed). This stillborn calf was ~77 Kg at birth. The average weight for calves of this breed is ~ 36 Kg. (D and E) Macroglossia of ES_ART_#2 and of stillborn ES_RF_necropsy_#1 (Asturian Valley x Limousin crossbred), respectively.

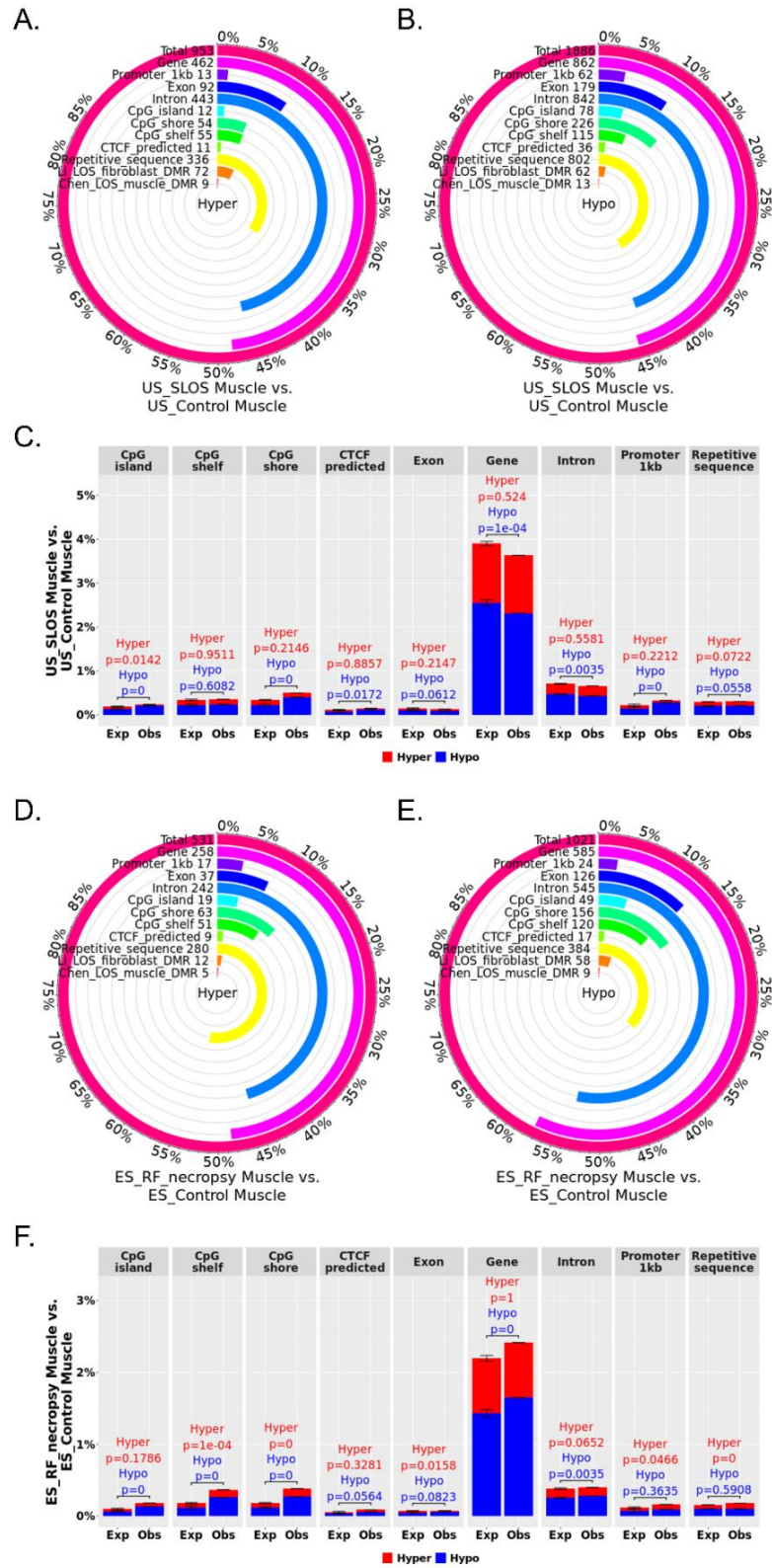


Figure 2.

Figure 2. Distribution of LOS associated differentially methylated regions (DMRs) across various genomic contexts.

(A-C) Muscle US_SLOS vs. US_Control DMRs. (D-F) Muscle ES_RF_necropsy vs. ES_Control DMRs. (A-B and D-E) Each figure shows the total number of DMRs in the comparison and the number and percent of the hypermethylated (hyper; A and D) and hypomethylated (hypo; B and E) DMRs over each genomic context. In addition, the figures include the number and percent of DMRs that overlap with two previous studies (Li (Li et al., 2022) and Chen (Chen et al., 2017)) for comparison purposes. (C and F) Percent of the genomic context that overlaps with DMRs. Obs = observed frequencies. Exp = expected frequencies (mean \pm standard deviation; obtained from randomly shuffling DMRs across genome 10,000 times). The p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|) / 10000$.

■ Hypo ■ Hyper ■ Overlap

DMR	LOS experiments					Count	ES_RF_bloo d_DMR	US_SLOS_tis sue_DMR	ES_ART_mu scl_DMR	ES_RF_musc le_DMR	Gene	Promoter 1kb	Exon	Intron	CpG island	CpG shore	CpG shelf	Repetitive sequence
	US_SLOS_mu scl_DMR	ES_RF_necro psy_muscle_ DMR	Li_LOS_fibrobl last_DMR	Chen_LOS_m uscle_DMR														
4_102069961_102070360					4					-								
6_66245821_66247640					4					ATP10D								
9_54438201_54439220					4					-								
9_64659791_64660620					4					-								
3_54800101_54801160					3					LOC511531								
4_8983421_8984640					3					-								
5_50897361_50898240					3					PPM1H								
5_57756761_57757480					3					-								
6_20896461_20896880					3					-								
6_94882141_94883160					3					PRDM8								
7_2579941_2581100					3					-								
7_12556261_12556900					3					NFIX								
7_84578941_84579760					3					-								
8_43514191_43518090					3					DMRT2								
9_96225361_96226140					3					IGF2R								
13_66465461_66466640					3					BLCAP_NNAT								
15_814141_814600					3					-								
15_1648961_1650300					3					KBTBD3								
16_42867461_42868700					3					CASZ1								
19_45140541_45141200					3					-								
19_64171061_64171520					3					LOC789803								
19_47821141_47821560					3					-								
20_58685361_58687420					3					TRIO								
23_33666491_33667620					3					-								
28_37047681_37048780					3					NRG3								

Figure 3. Example of LOS-associated vulnerable loci.

Hypo = hypomethylation. Hyper = hypermethylation. Differentially methylated regions (DMR) identifiers are the positions in the bovine genome assembly ARS-UCD1.2. For complete information please refer to Table S2.

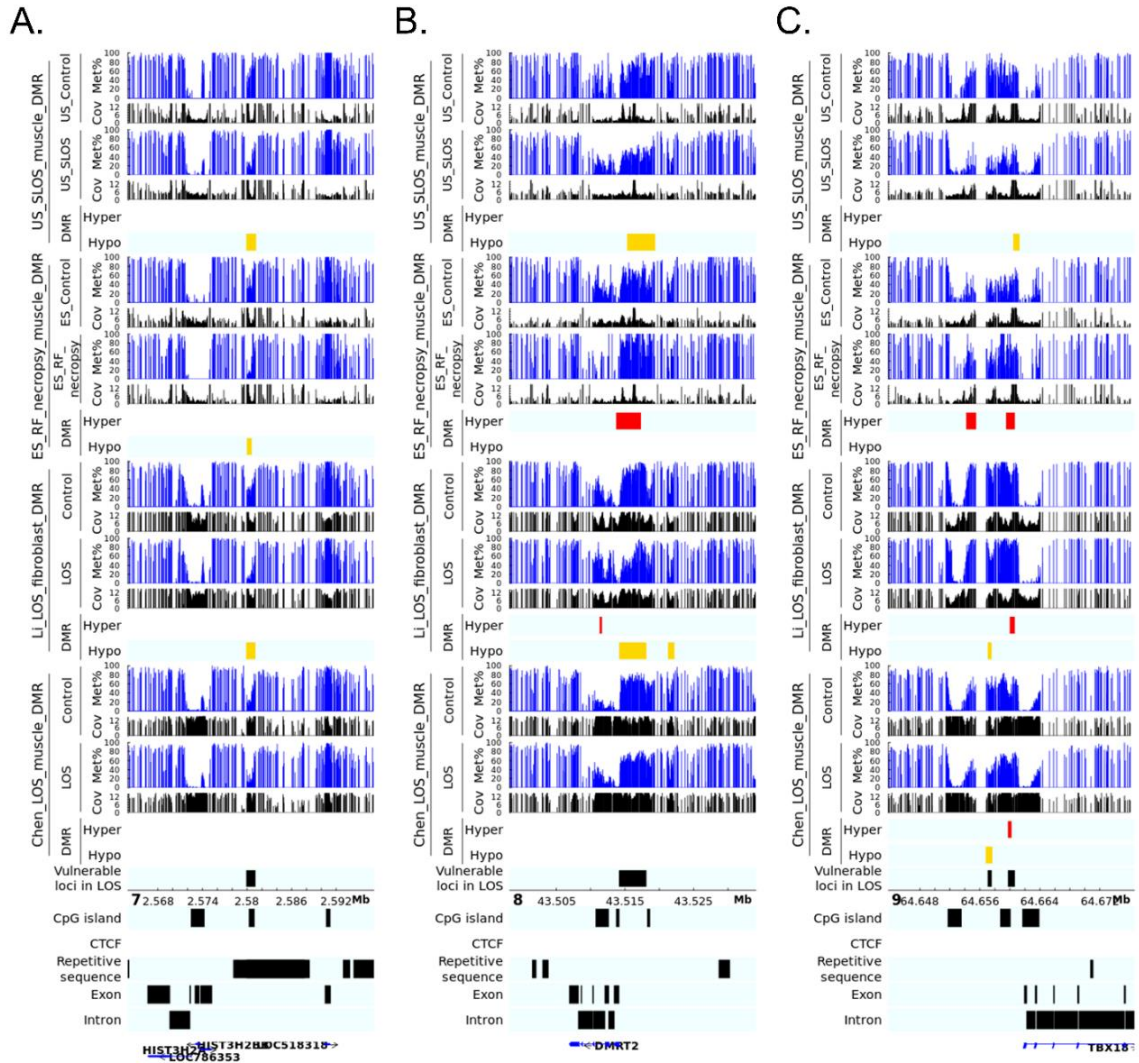


Figure 4. LOS-vulnerable loci around promoter regions.

This figure shows DNA methylation level of vulnerable loci 7_2579941_2581100 (A), 8_43514181_43518080 (B), and 9_64659781_64660620 (C) in four LOS experiments. The aforementioned numbers refer to the chromosomes and genomic position in bovine genome assembly ARS-UCD1.2. Met% = group mean CpG methylation level in percent. Cov = group mean CpG read coverage. DMR = differentially methylated regions. Hyper = hypermethylation (red). Hypo = hypomethylation (yellow).

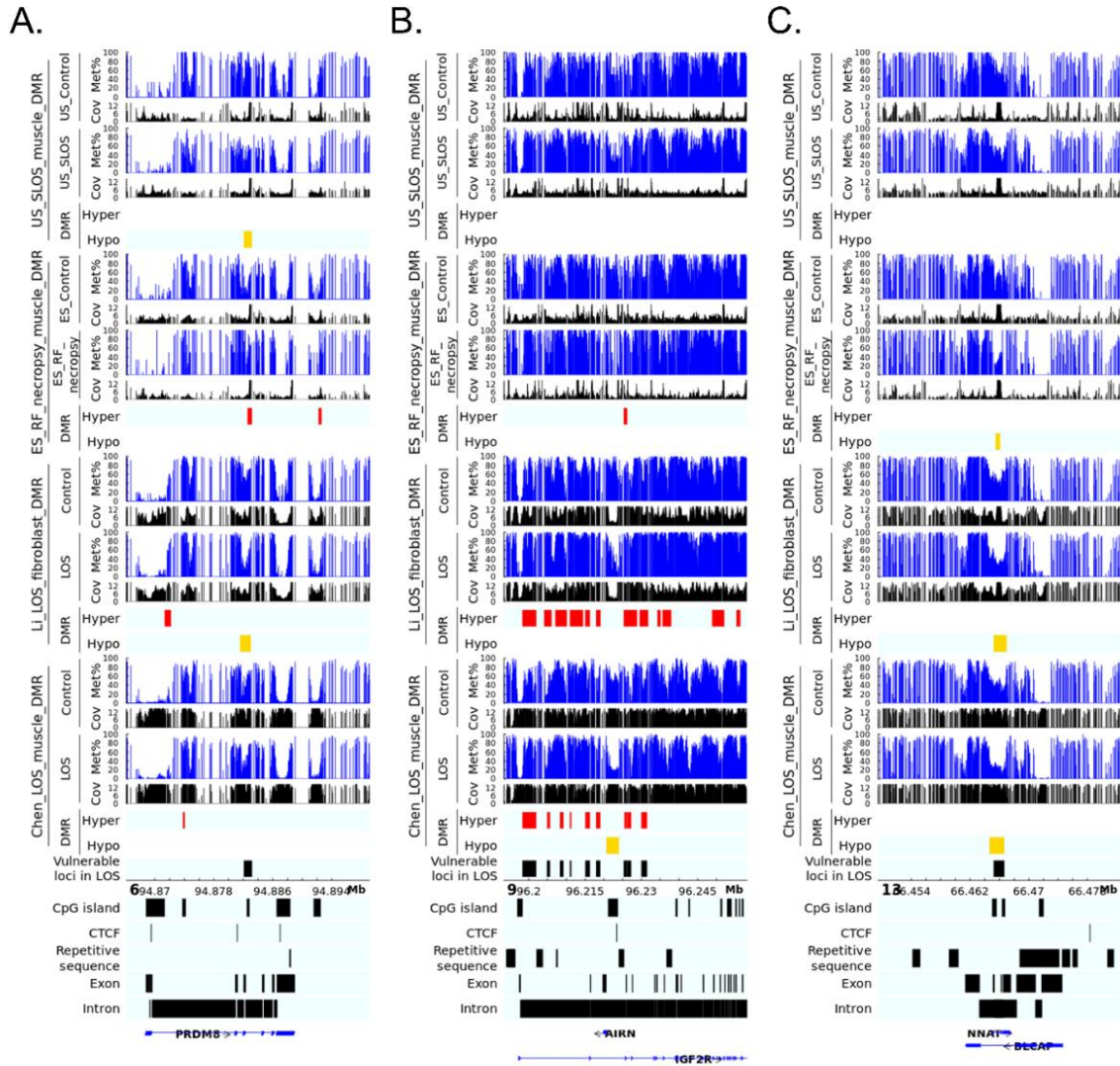


Figure 5. LOS-vulnerable loci overlapping CpG islands in body of imprinted genes.

This figure shows DNA methylation level of vulnerable loci 6_94882141_94883160 (A), 9_96225361_96226140 (B), and 13_66465461_66466640 (C) in four LOS experiments. The aforementioned numbers refer to the chromosomes and genomic position in bovine genome assembly ARS-UCD1.2. Met% = group mean CpG methylation level in percent. Cov = group mean CpG read coverage. DMR = differentially methylated regions. Hyper = hypermethylation (red). Hypo = hypomethylation (yellow).

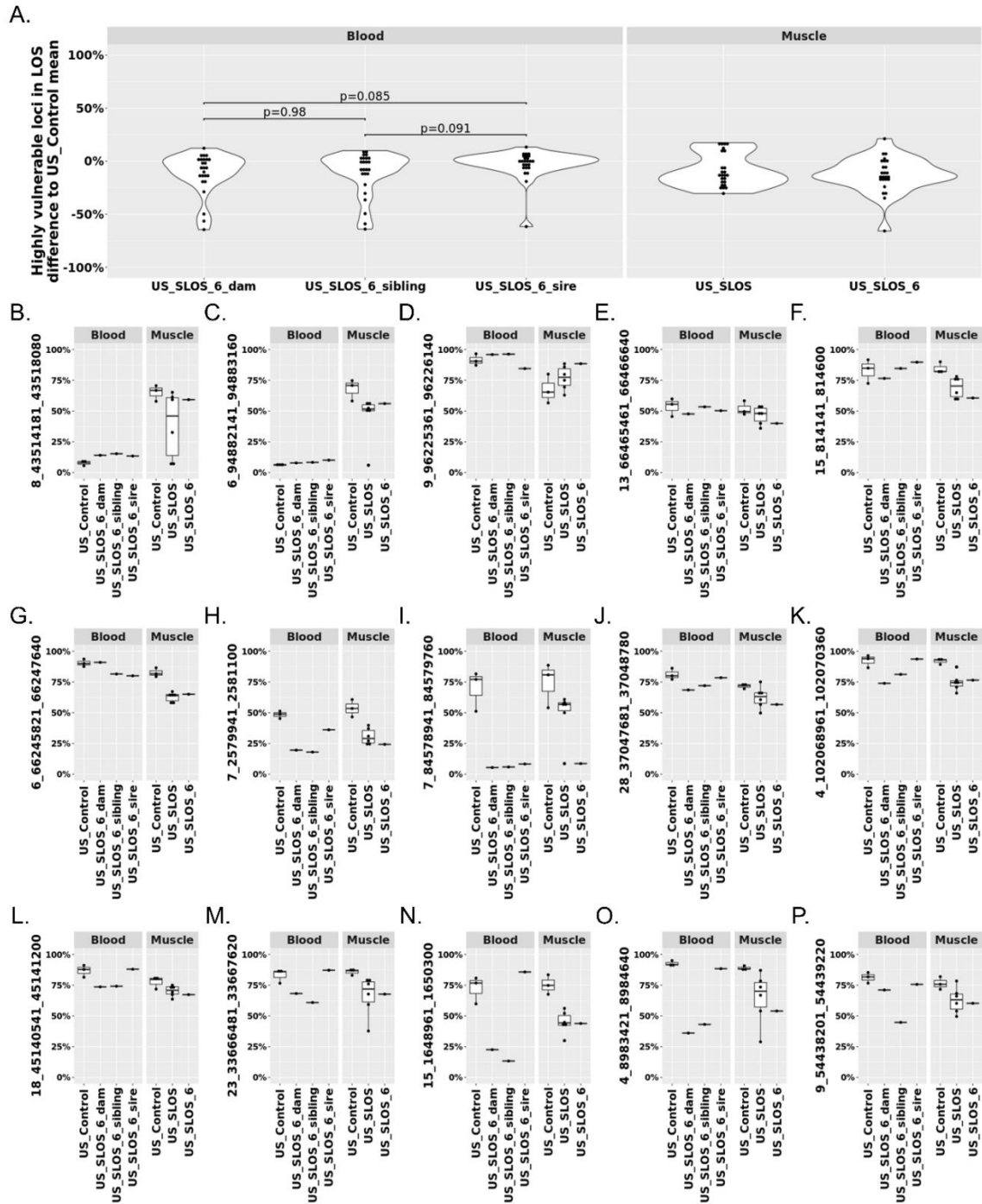


Figure 6.

Figure 6. DNA methylation at the 25 highly vulnerable loci of LOS (≥ 3 experiments) in US_SLOS_#6 calf, its sire, dam, and full-sibling.

(A) Violin plots with dots showing the difference of DNA methylation between examined individual/group (mean) and mean of US_Control for highly vulnerable loci in LOS (found in ≥ 3 LOS experiments). Each dot in the violin plot represents a vulnerable locus. P values were from t-test. The baseline for blood samples is US_Control blood and the baseline for muscle samples is US_Control muscle.

(B-P) Box plots with dots show DNA methylation level (y-axis) at highly vulnerable loci without obvious differences ($>10\%$) in parental blood samples (B-F), with obvious differences in sire only (G), both sire and dam (H-I), dam only (J-O), and sibling only (P) when compared to the mean of control group. Each dot in the box plot represents a sample.

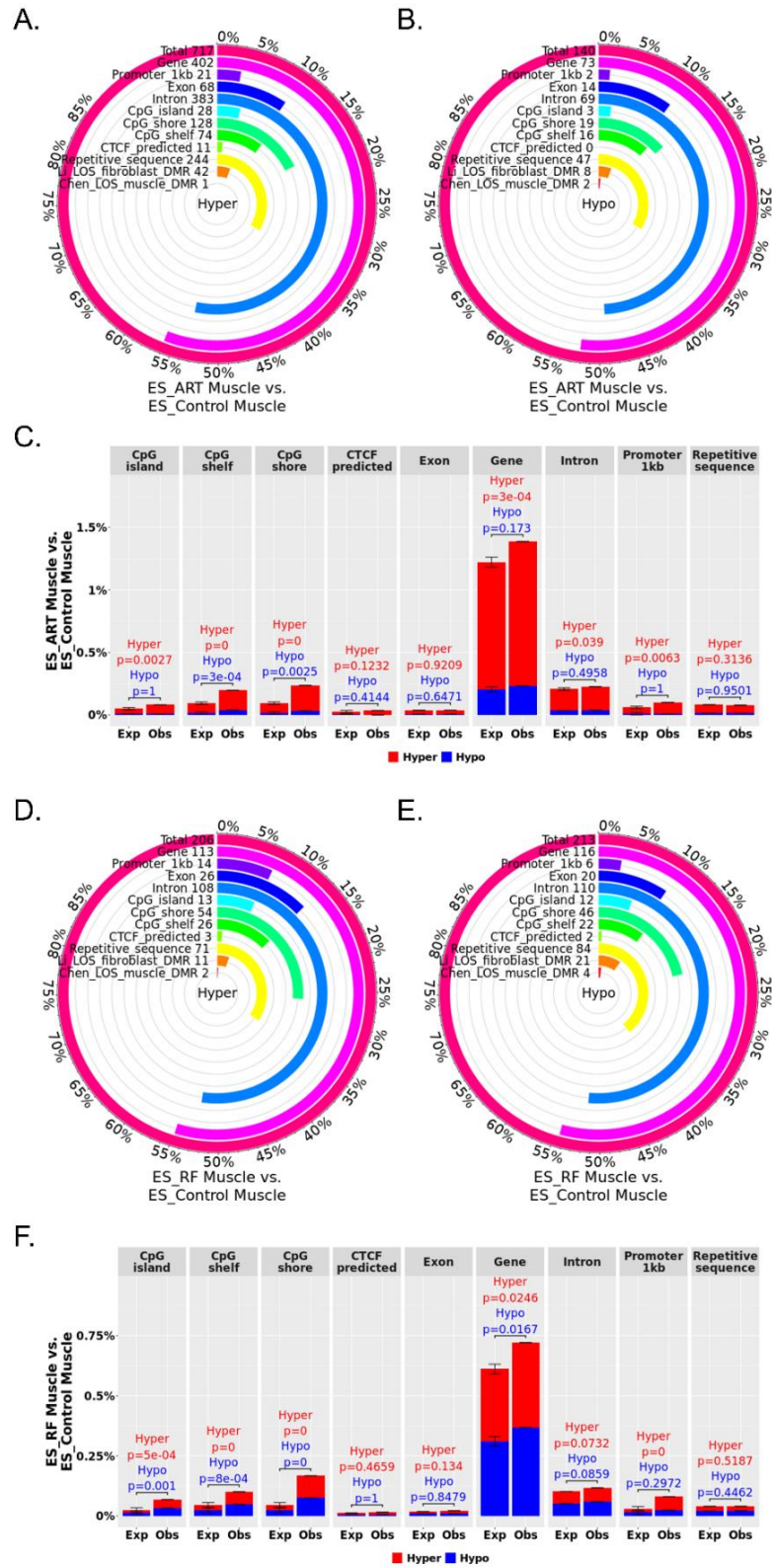


Figure 7.

Figure 7. Distribution of ART associated differentially methylated regions (DMRs) across various genomic contents.

(A-C) Muscle ES_ART vs. ES_Control DMRs. (D-F) Muscle ES_RF vs. ES_Control DMRs. (A-B and D-E) Each figure shows the total number of DMRs in the comparison and the number and percent of the hypermethylated (hyper; A and D) and hypomethylated (hypo; B and E) DMRs over each genomic context. In addition, the figures include the number and percent of DMRs that overlap with two previous studies (Li (Li et al., 2022) and Chen (Chen et al., 2017)) for comparison purposes. (C and F) Percent of the genomic context that overlaps with DMRs. Obs = observed frequencies. Exp = expected frequencies (mean \pm standard deviation; obtained from randomly shuffling DMRs across genome 10,000 times). The p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|) / 10000$.

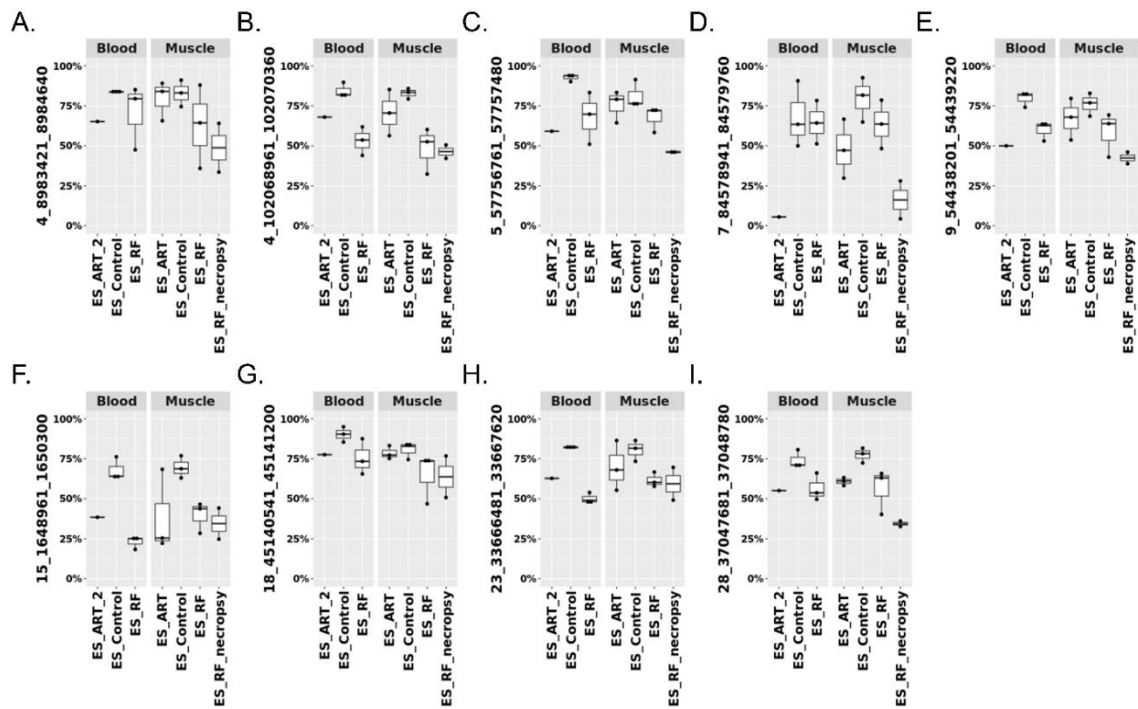


Figure 8. Conservation of DNA methylation at LOS-vulnerable loci between muscle and blood in ES_RF group.

Y-axis = DNA methylation level. Note: ES_ART_2 is the same animal as Figure 1(D).

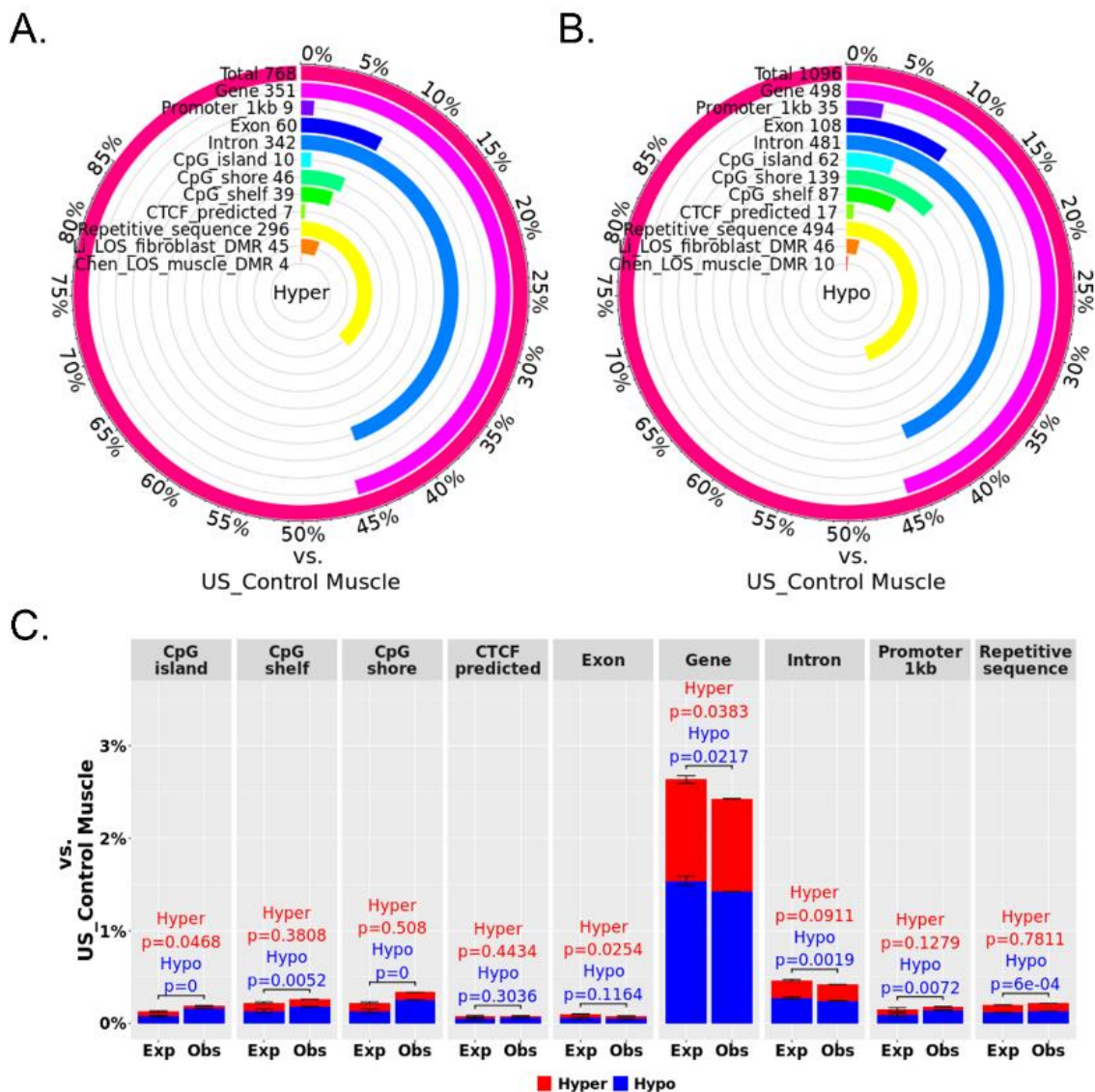


Figure S1. Distribution of LOS associated differentially methylated regions (DMRs) across various genomic contexts.

US_SLOS tissue vs. US_Control muscle DMRs. (A-B) Each figure shows the total number of DMRs in the comparison and the number and percent of the hypermethylated (hyper; A) and hypomethylated (hypo; B) DMRs over each genomic context. In addition, the figures include the number and percent of DMRs that overlap with two previous studies (Li (Li et al., 2022) and Chen (Chen et al., 2017)) for comparison purposes. (C) Percent of the genomic context that overlaps with DMRs. Obs = observed frequencies. Exp = expected frequencies (mean \pm standard deviation; obtained from randomly shuffling DMRs across genome 10,000 times). The p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|) / 10000$.

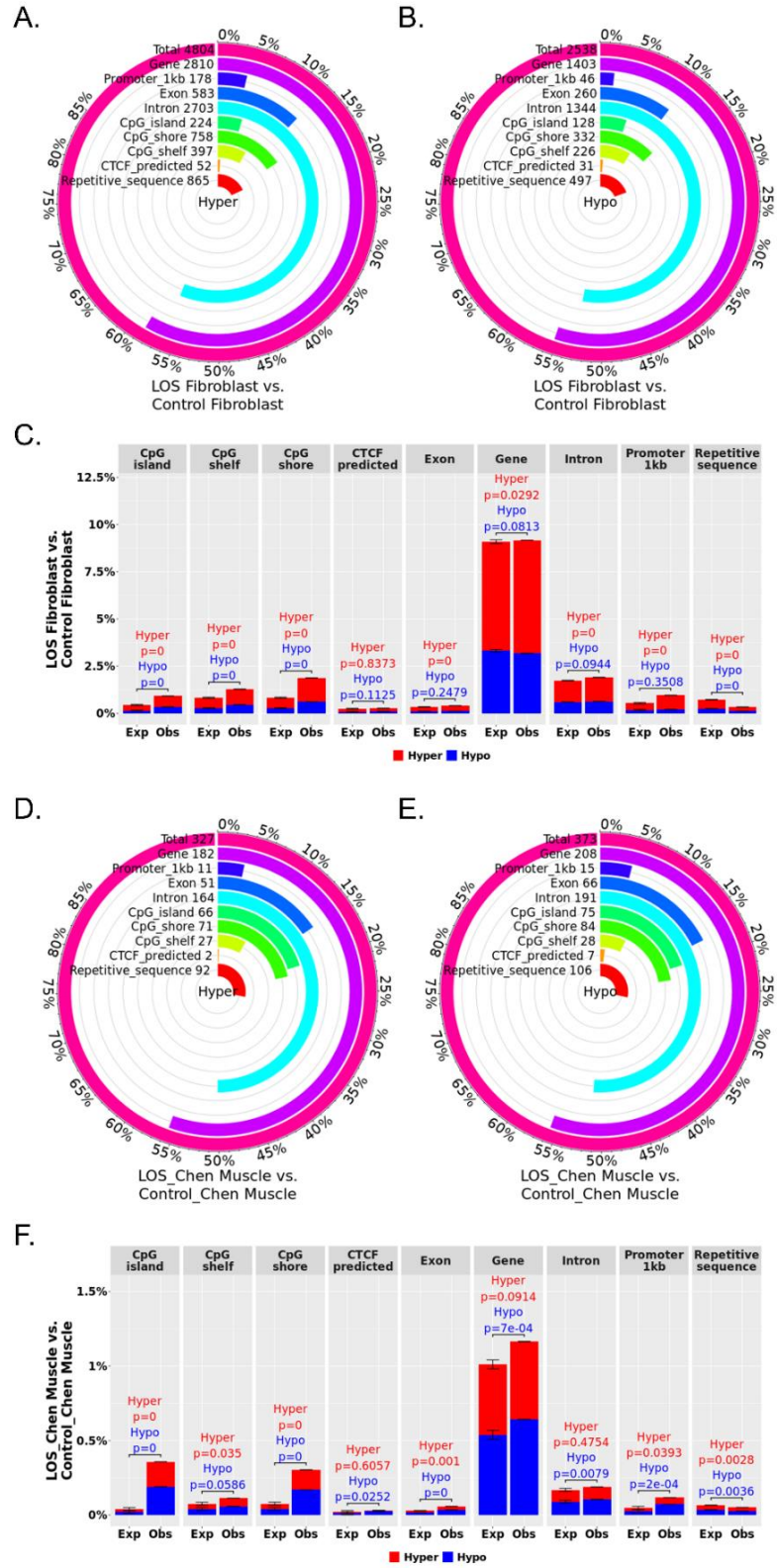


Figure S2.

Figure S2. Distribution of LOS associated differentially methylated regions (DMRs) across various genomic contexts.

(A-C) Fibroblast LOS vs. Control DMRs identified from our previous study (Li et al., 2022). (D-F) Fetal muscle LOS vs. Control DMRs identified from our previous study (Chen et al., 2017). (A-B and D-E) Each figure shows the total number of DMRs in the comparison and the number and percent of the hypermethylated (hyper; A and D) and hypomethylated (hypo; B and E) DMRs over each genomic context. (C and F) Percent of the genomic context that overlaps with DMRs. Obs = observed frequencies. Exp = expected frequencies (mean \pm standard deviation; obtained from randomly shuffling DMRs across genome for 10,000 times). The p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|) / 10000$.

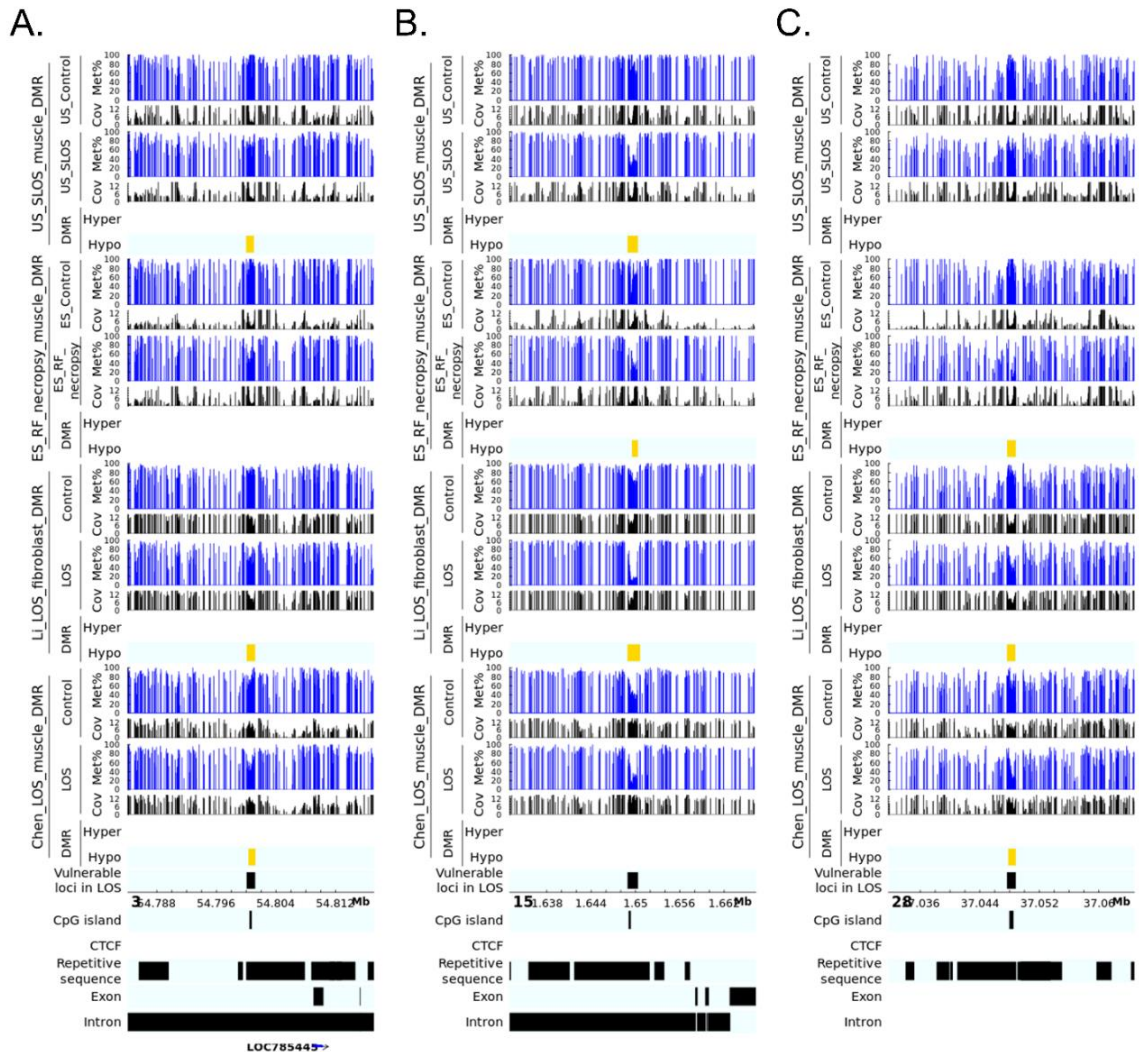


Figure S3. LOS- vulnerable loci overlapping CpG islands in gene bodies.

This figure shows DNA methylation level of vulnerable loci 3_54800101_54801160 (A), 15_1648961_1650300 (B), and 28_37047681_37048780 (C) in four LOS experiments. The aforementioned numbers refer to the chromosomes and genomic position in bovine genome assembly ARS-UCD1.2. Met% = group mean CpG methylation level in percent. Cov = group mean CpG read coverage. DMR = differentially methylated regions. Hyper = hypermethylation (red). Hypo = hypomethylation (yellow).

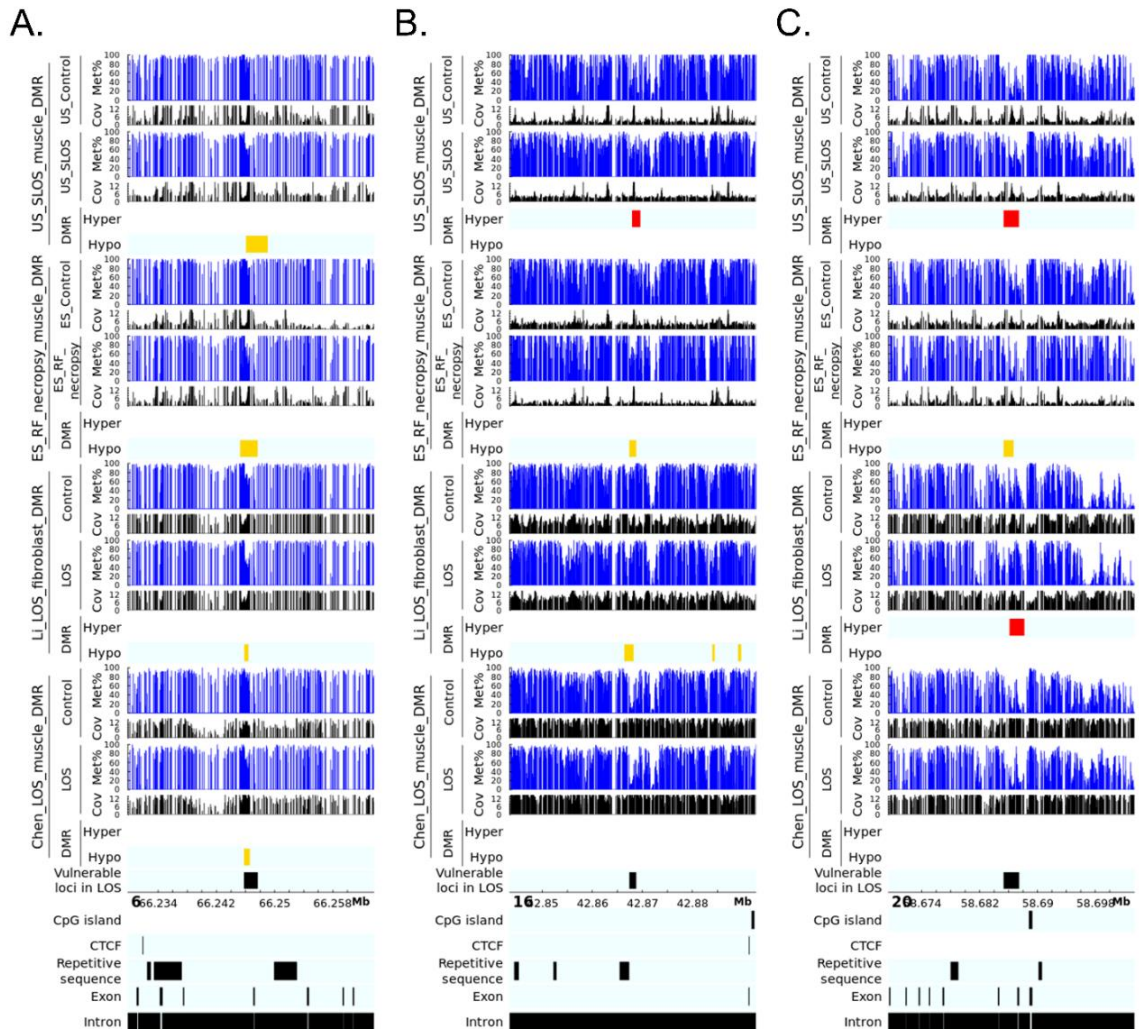


Figure S4. LOS- vulnerable loci in gene bodies.

This figure shows DNA methylation level of vulnerable loci 6_66245821_66247640 (A), 16_42867461_42868700 (B), and 20_58685361_58687420 (C) in four LOS experiments. The aforementioned numbers refer to the chromosomes and genomic position in bovine genome assembly ARS-UCD1.2. Met% = group mean CpG methylation level in percent. Cov = group mean CpG read coverage. DMR = differentially methylated regions. Hyper = hypermethylation (red). Hypo = hypomethylation (yellow).

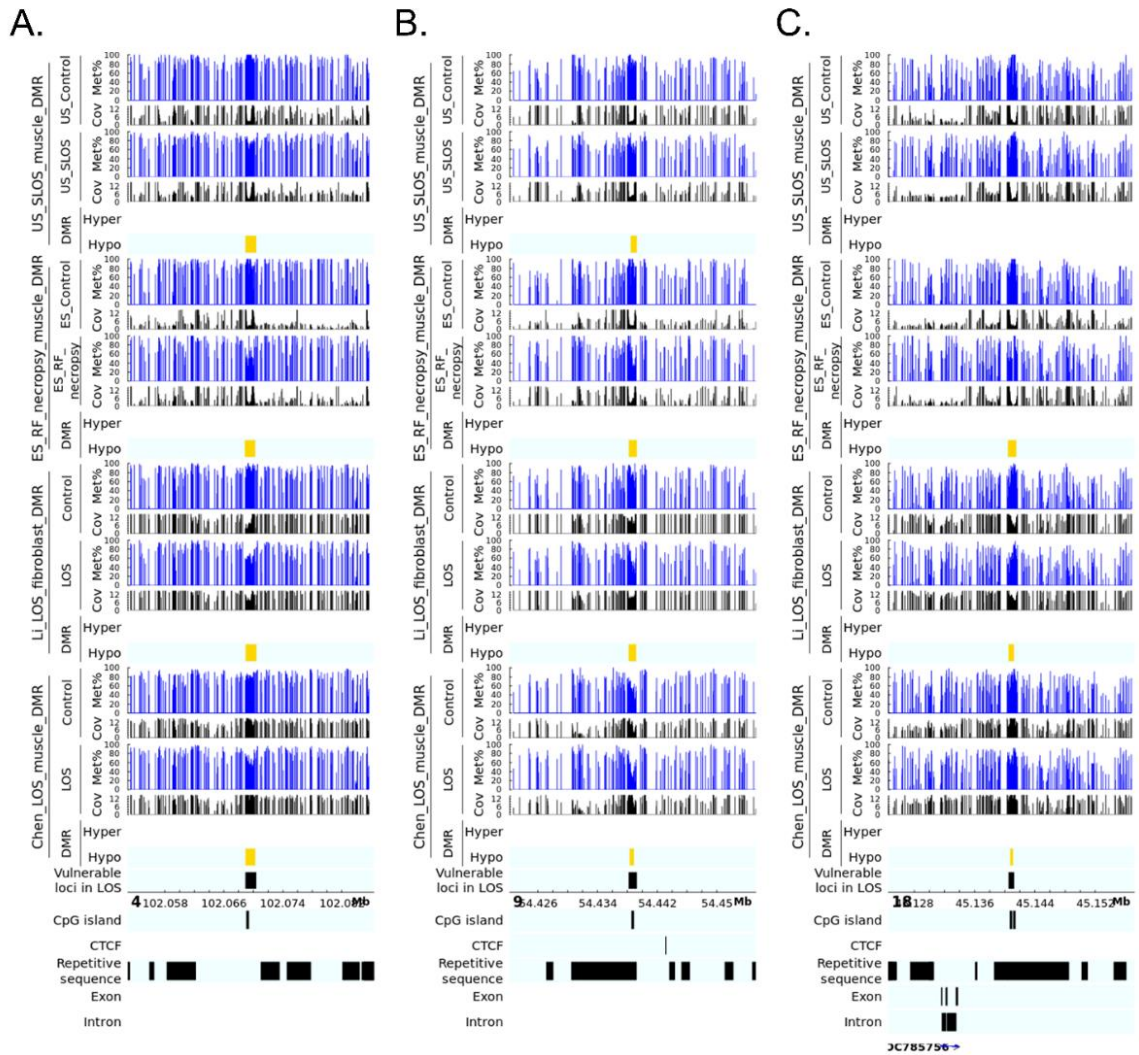


Figure S5. LOS- vulnerable loci overlapping CpG islands in intergenic regions.

This figure shows DNA methylation level of vulnerable loci 4_102068961_102070360 (A), 9_54438201_54439220 (B), and 18_45140541_45141200 (C) in four LOS experiments. The aforementioned numbers refer to the chromosomes and genomic position in bovine genome assembly ARS-UCD1.2. Met% = group mean CpG methylation level in percent. Cov = group mean CpG read coverage. DMR = differentially methylated regions. Hyper = hypermethylation (red). Hypo = hypomethylation (yellow).

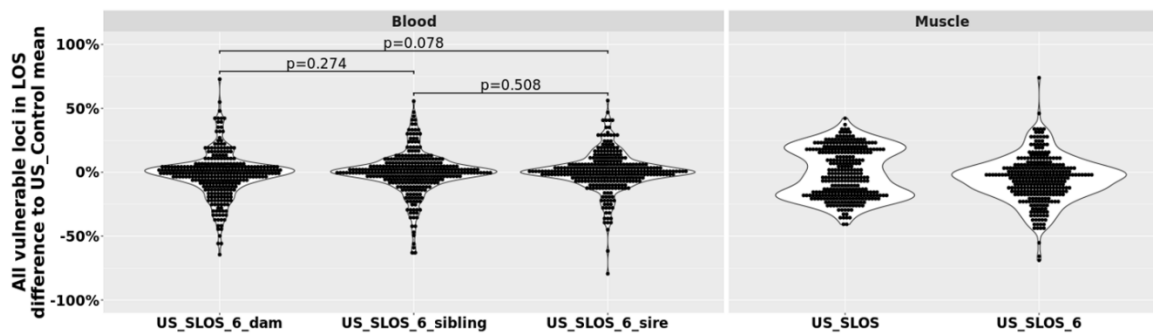


Figure S6. DNA methylation at all LOS-vulnerable loci in US_SLOS_#6 calf and its sire, dam, and full-sibling.

Violin plots with dots showing the difference of DNA methylation between examined individual/group (mean) and mean of US_Control for all 320 vulnerable loci in LOS. Each dot represents a vulnerable locus. P values were from t-test.

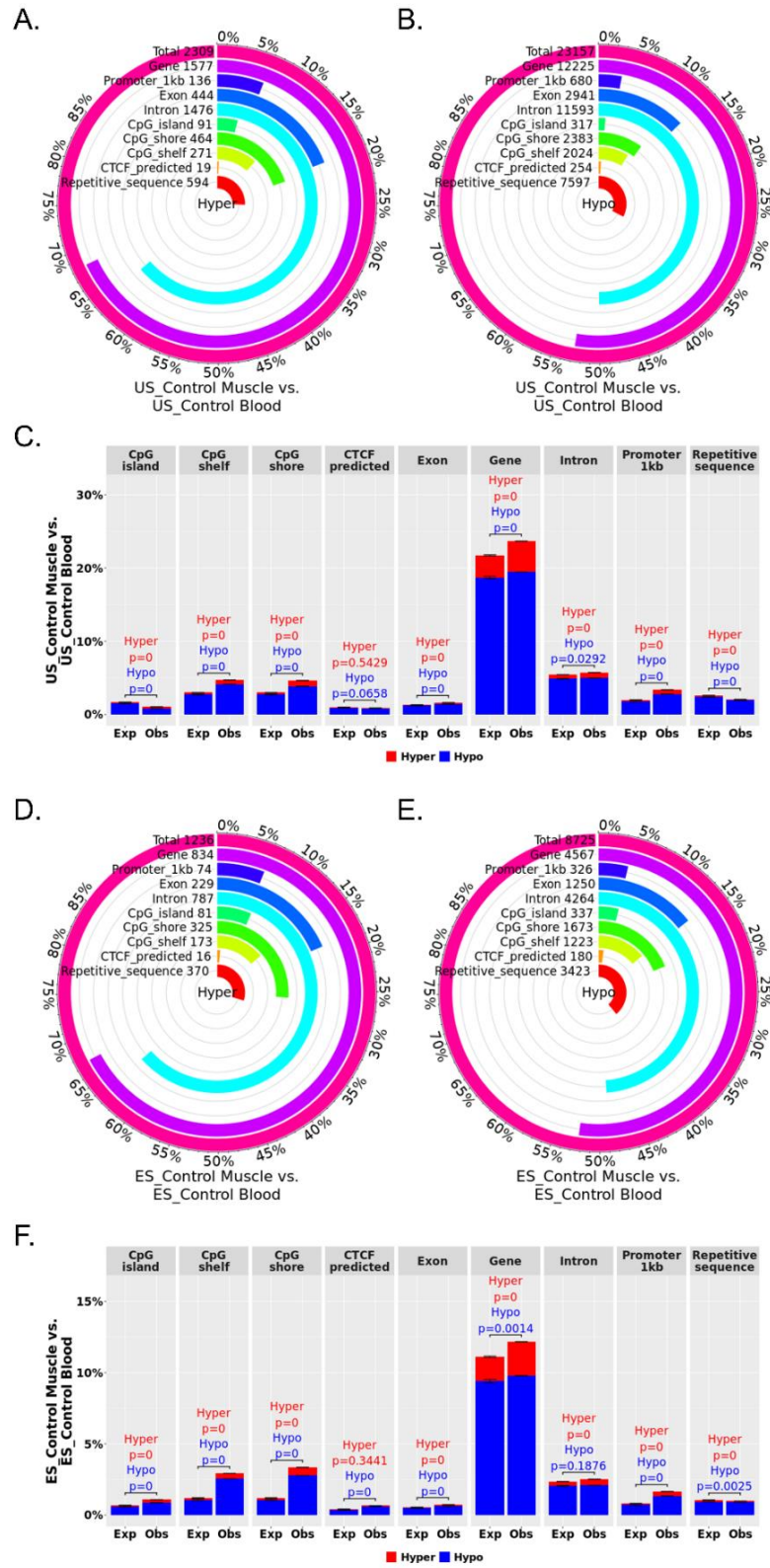


Figure S7.

Figure S7. Distribution of tissue specific differentially methylated regions (DMRs) across various genomic contents.

(A-C) US_Control muscle vs. blood DMRs. (D-F) ES_Control muscle vs. blood DMRs. (A-B and D-E) Each figure shows the total number of DMRs in the comparison and the number and percent of the hypermethylated (hyper; A and D) and hypomethylated (hypo; B and E) DMRs over each genomic context. (C and F) Percent of the genomic context that overlaps with DMRs. Obs = observed frequencies. Exp = expected frequencies (mean \pm standard deviation; obtained from randomly shuffling DMRs across genome 10,000 times). The p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|)/10000$.

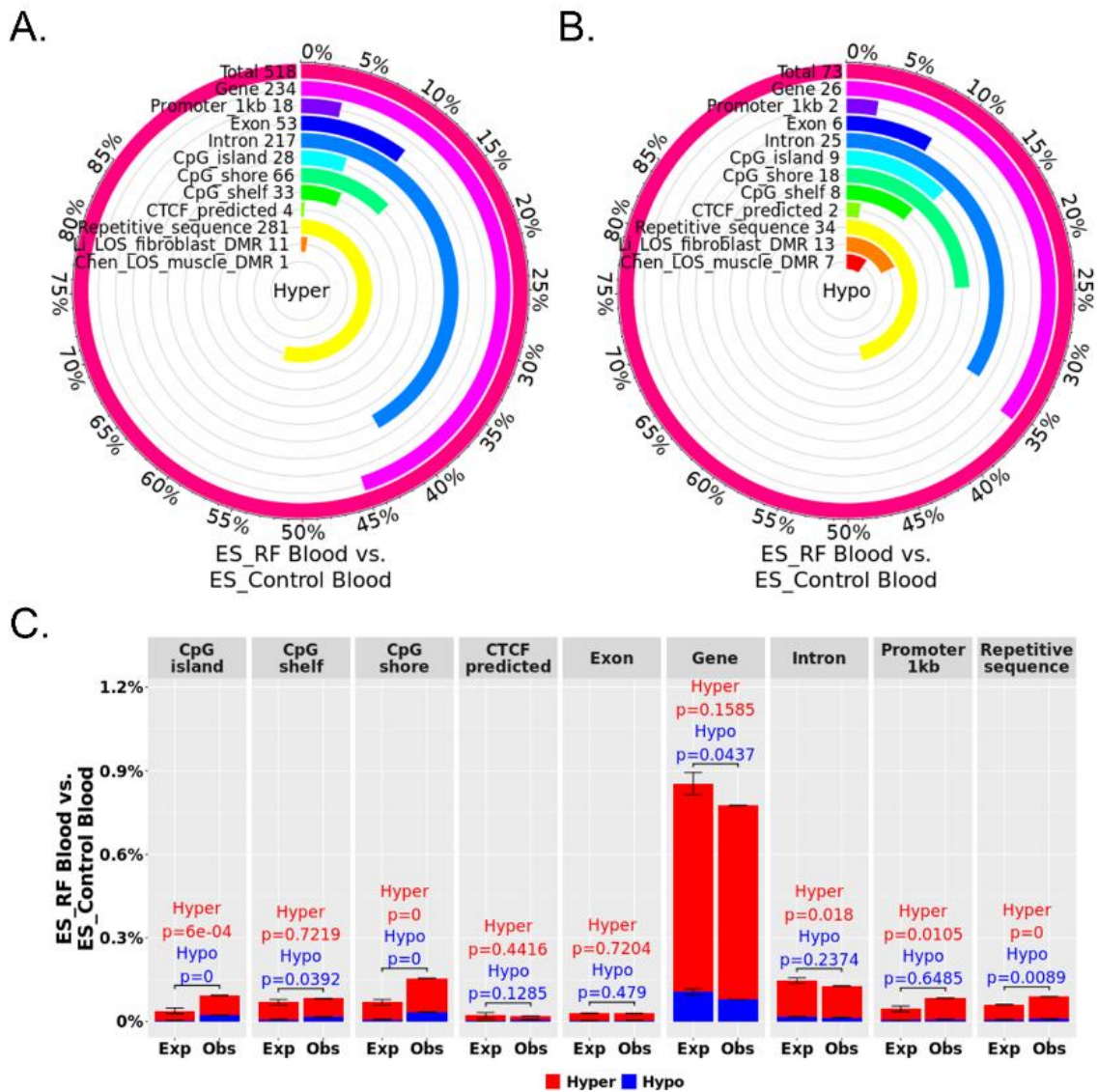


Figure S8. Distribution of ART associated differentially methylated regions (DMRs) across various genomic contents.

Blood ES_RF vs. ES_Control DMRs. (A-B) Each figure shows the total number of DMRs in the comparison and the number and percent of the hypermethylated (hyper; A) and hypomethylated (hypo; B) DMRs over each genomic context. In addition, the figures include the number and percent of DMRs that overlap with two previous studies (Li (Li et al., 2022) and Chen (Chen et al., 2017)) for comparison purposes. (C) Percent of the genomic context that overlaps with DMRs. Obs = observed frequencies. Exp = expected frequencies (mean \pm standard deviation; obtained from randomly shuffling DMRs across genome 10,000 times). The p values were calculated as $p = n(|\text{Exp} - \text{mean}(\text{Exp})| \geq |\text{Obs} - \text{mean}(\text{Exp})|) / 10000$.

Chapter 4: General discussion

In this dissertation, two studies have been presented. For one study, we used primary fibroblast cell lines derived from bovine control and LOS fetuses to determine chromosome architecture of imprinted domains. We found allele-specific chromosome architecture at the *IGF2R* imprinted domain in the control group which is disrupted in LOS group. From RNA-seq data, we also found genomic location-based clustering tendency of misregulated genes, which suggests genome-wide alterations of chromosome architecture is involved in LOS. For the other study, we examined the DNA methylome of four independent LOS studies, including one for SLOS. We identified hundreds of genomic loci that are vulnerable for DNA methylation defects in LOS, and 25 loci that are highly vulnerable among them. These 25 loci have the potential to serve as molecular markers for the diagnosis of LOS during pregnancy.

In addition, we have an on-going Hi-C study of LOS and BWS. This study will include WGBS, RNA-seq, and Hi-C sequencing for 11 bovine fibroblast samples (control and LOS group) and 23 human fibroblast samples (control and BWS group). This study aims to systemically and comparatively analyze the crosstalk among DNA methylation, chromosome architecture, and gene expression in LOS and BWS.

With the findings in the two present studies, several questions are

reasoned below.

4.1 How does ART induce LOS and BWS?

As mentioned in Chapter 1, the use of ART increases the incidence of BWS even without the supplementation of serum during embryo culture, and how ART leads to this remains a question. One possible answer is that the recipes for some medium used during ART are still not optimal and require further improvement or even personalization for different individuals to fit their genetics. The present studies in this dissertation and the previous miRNA study have shown chromosome architecture changes in the *IGF2R* imprinted domain and an indication of global changes (Li et al., 2019a). Based on these studies, another possible answer to the question is that the differences in mechanical force (e.g., shear stress) and physical status (e.g., cell shape and movement) between *in vivo* and *in vitro* treatments have impacts on oocyte maturation and embryo development through affecting cytoskeleton and chromosome architectures.

ART procedures often involve multiple steps of pipetting of oocytes and/or embryos and cannot faithfully recapitulate the physical environment of *in vivo*, such as organ motions (Mizrachi et al., 2018). Shear stress, caused by rapid movement of fluid over cell membrane, is known as a factor to induce embryonic stem cell differentiation towards endothelial cells (Yamamoto et al.,

2005). Alterations of gene expression and histone modification and reorganization of cytoskeletons have been reported during this process in endothelial cells (Illi et al., 2005; Tzima et al., 2002). In mouse preimplantation embryos, continuous shear stress (1.2 dynes/cm²) for 12 hours induces the activation (i.e., phosphorylation) of mitogen-activated protein kinase 8 (MAPK8) and MAPK9 which mediate stress responses, upregulation of Fos proto-oncogene, AP-1 transcription factor subunit (FOS) which is a marker of shear stress, and eventually leads to cell apoptosis and embryonic lethality (Xie et al., 2006). The impacts of shear stress are more obvious in denuded embryos than embryos with zona pellucida (Xie et al., 2006). In addition, pipetting of embryos leads to increase of MAPK8/9 phosphorylation and FOS expression in a dosage dependent manner (Xie et al., 2007).

Cytoskeleton refers to the skeleton structures in cytoplasm and consists mainly microtubule, microfilament, and intermediate filament (Hohmann and Dehghani, 2019). Different types of cytoskeletons often have crosstalk through linker proteins such as plakin (Mohammed et al., 2020). In addition, cytoskeletons are closely associated with extracellular matrix (ECM) for the transmission of mechanical forces (Peyton et al., 2007). Disruption of cytoskeletons have been shown to affect normal nuclear shape and position, chromatin condensation, and gene expression (Alam et al., 2016; Keeling et al., 2017). Importantly, cytoskeletons have been shown to regulate chromosome

architecture (Dundr et al., 2007; Pradhan et al., 2018).

Tubulin forms microtubules in cytoplasm, which play roles as tracks or supporting molecules in regulation of cell shape, cell migration, transportation of molecules, and spindle positioning during mitosis (Bhogaraju et al., 2013; Ganguly et al., 2011; Jung et al., 2013; Lee et al., 2007).

Actin mainly forms microfilaments in cytoplasm and nucleus, which play roles in regulation of cell shape, cell migration, vesicle transportation, endocytosis and exocytosis, and cytokinesis during cell division (Berven et al., 2004; Frémont et al., 2017; Fujimoto et al., 2000; Jiménez et al., 2000; McNiven et al., 2000; Murthy and Wadsworth, 2005; Schuh, 2011). In the nucleus, actin can also function in monomer, known as globular actin (G-actin), and is needed as a part of the transcription machinery during transcription initiation and elongation (Dopie et al., 2012; Hofmann et al., 2004; Le et al., 2016; Obrdlik and Percipalle, 2011; Philimonenko et al., 2004; Xu et al., 2010). Long range intranuclear translocation during gene activation has been reported as mediated by actin, reflecting its roles in the regulation of chromosome architecture (Dundr et al., 2007). Actin is a component of the chromatin remodeler BRM-associated factors (BAF) complex, which regulate epigenetic marks on chromatin during cell differentiation, and disruption of actin leads to abnormal deposition of H3K9me3 and H3K17me3 and increased heterochromatins (Varga et al., 2021; Xie et al., 2018). Nuclear actin is also involved in DNA replication and damage repair

processes (Belin et al., 2015; Caridi et al., 2018; Parisis et al., 2017; Schrank et al., 2018).

Lamin is a type (type V) of the proteins that form intermediate filaments found particularly in nucleus, known as lamina, to regulate spatial localization of heterochromatin (Solovei et al., 2013). Genomic regions that contact with nuclear lamina are known as lamina-associated domains (LADs) (Van Steensel and Belmont, 2017). The expression and phosphorylation of lamin is regulated by the stiffness of ECM (Buxboim et al., 2014; Swift et al., 2013). Other types (type I to IV) of intermediate filament-forming proteins exist in cytoplasm to mainly form a dense meshwork at perinuclear regions and also reaches cell membrane (Craggs et al., 2001; Kreis et al., 2005). This meshwork serves as a mechanical buffer and can organize the locations of organelles including nucleus, mitochondria, and the Golgi apparatus (Chang and Goldman, 2004; Dupin et al., 2011; Matveeva et al., 2015). Importantly, the nuclear lamina and cytoskeletons are physically connected by the linker of nucleoskeleton and cytoskeleton (LINC) complex consists members of Sad1 and UNC84 domain containing 1 (SUN) and spectrin repeat containing nuclear envelope protein (SYNE/nesprin) families (Ostlund et al., 2009; Wilhelmsen et al., 2005). LINC enables the force transmission between nuclear lamina and cytoskeletons, and the disruption of LINC can lead to genome-wide gene misregulation (Alam et al., 2016; Lombardi et al., 2011).

Emerin (EMD) is a member of the nuclear lamina-associated protein family located on nuclear membrane and plays roles in transmission of mechanical force signals into nucleus (Pradhan et al., 2018). The stiffness of cell culture substrate has been found to affect chromosome territories, positioning of nuclear lamina, and gene expression through EMD in human tumor cell lines (Pradhan et al., 2018).

Actin and lamin are closely associated with each other in the regulation of nuclear morphology. A study in human epidermal stem cells has revealed the mechanisms by which mechanical force (e.g., strain) regulates lineage commitment (Le et al., 2016). Strain induces the activation of myosin heavy chain 9 (MYH9/NMIIA) accompanied with enrichments of MYH9 and EMD at outer nuclear membrane, which promotes actin polymerization and reduces nuclear actin level. The lack of nuclear G-actin causes decreased transcription activity, and eventually leads to accumulation of H3K27me3 and increased heterochromatin anchored by nuclear lamina. Another study in mouse embryonic fibroblast cells has shown that nuclear lamin could rapidly recruit a perinuclear actin cap to maintain nuclear morphology upon mechanical stress (Kim et al., 2017).

Taken together, these studies suggest an axis of mechanical stress, cytoskeleton, chromosome architecture, DNA methylation and histone modifications, and gene expression for considering the involvement of ART in

LOS and BWS development. Whether genome-wide alteration of chromosome architecture is involved in LOS and BWS will be soon revealed by our Hi-C studies. Future studies are needed to determine whether mechanical stress is a causal effect for LOS and BWS and the involvement of cytoskeletons.

4.2 How do DNA methylation changes occur in LOS and BWS?

Chapter 1 mentioned many genes with important roles in the establishment and alterations of DNA methylation and histone modifications. Although the mutation or misregulation of some of these genes, including *DNMTs* (Dagar et al., 2018; Sano et al., 2016) and *ZFP57* (Boonen et al., 2012; Sano et al., 2016), have been reported in BWS patients, their frequencies remain very low. In addition, misregulation of most of these genes, including *DNMTs*, *TET2*, *TET3*, *SETD2*, *UHRF1*, *DPPA3*, *ZFP57*, *ZNF445*, *ZNF202*, *KDM1B*, and *TRIM28*, were not found in the RNA-seq results of LOS fibroblast cells, except for *TET1* which had a ~42% upregulation of transcript coupled with intronic hypermethylation (28_24908960_24909439) in the LOS group. Intronic hypermethylation (23_39430166_39430845) of *KDM1B* was also found in the LOS group but is not associated with gene misregulation. However, it is still largely unknown whether these genes have normal expression and/or epigenetic marks in the preimplantation embryos of the BWS patients or LOS animals and in

the germ cells that contribute to them. For preimplantation embryos, these examinations could be done with embryo biopsy (Gianaroli, 2000). For germ cells, this is a technically challenging question to answer since current technologies cannot complete the tests without damaging or consuming the oocytes and sperms. For oocytes, polar body biopsy could provide some insights but will still be difficult to faithfully reflect the mature ovum (Gianaroli, 2000). Alternatively, additional somatic cells or germ cells from the same donor could be used for the examinations and may partially reflect the germ cells that contribute to BWS or LOS. Future studies are needed to answer this question.

Interestingly, some other members of the ZNF family were found misregulated in LOS group and located across the genome, including downregulation of *ZNF385A*, *ZNF408*, *ZNF423* (coupled with gene body hyper- and hypomethylations), *ZNF503*, *ZNF668*, *ZNF703*, *ZNF771*, and upregulation of *ZNF207*, *ZNF644*. There is no report showing whether these ZNF genes are involved in genomic imprinting regulation and future studies are needed to determine this.

As reviewed in Chapter 1, DNA methylation over the CTCF binding sites affects their binding to DNA, and loss of CTCF binding could result in gain of DNA methylation at originally unmethylated regions. In addition, different chromosome compartments and subcompartments have their specific microenvironment features, which means when a genomic region is incorrectly

placed in a different compartment, it will have a higher chance to encounter molecules that normally cannot be recruited by it. The nuclear localizations of many enzymes with important roles in the establishment and alterations of DNA methylation and histone modifications, including DNMTs, TETs, UHRF1, DPPA3, KDM1B, and SETD2, have been found enriched as particles instead of even distribution (Ciccone et al., 2009; Du et al., 2019; Li et al., 2020a, 2018a, 2018b; Nady et al., 2011; Nakamura et al., 2012; Sharif et al., 2007; Song et al., 2018). These suggest an explanation for the genome-wide DNA methylation defects in LOS if global alteration of chromosome architecture exists in LOS. During epigenetic reprogramming in oocytes and/or embryos, temporary disruption of chromosome architecture caused by mechanical stress places some genomic regions in incorrect chromosome compartments and results in abnormal DNA methylation status and/or histone modifications applied to these regions. In return, these epigenetic marks affect the interactions of architectural proteins with these regions, and eventually lead to stabilized disruption of chromosome architecture with expanded impacts. This hypothesis needs experimental validations in the future, particularly with genome editing studies to delete genomic regions and disrupt normal chromosome architecture during early embryo development to determine the impacts on epigenetic marks.

4.3 Conclusion

In summary, this dissertation work highlighted the involvement of chromosome architecture alterations in bovine LOS and identified genomic loci with DNA methylation defect vulnerability in this overgrowth syndrome.

Appendix 1: Ongoing Hi-C sequencing project

The results of the 4C project in Chapter 2 suggested genome-wide alteration of chromosome architecture in bovine LOS. Recent studies in BWS also showed that chromosome architecture changes are not limited to *IGF2* imprinted domain (Naveh et al., 2021; Rovina et al., 2020). This inspired us to collaborate with Dr. Jennifer Kalish from the Children's Hospital of Philadelphia to conduct a Hi-C sequencing project in which LOS and BWS samples will be analyzed simultaneously. The project will determine genome-wide chromosome architecture, DNA methylome, and transcriptome for a comprehensive analysis of the correlation between DNA methylation defects and gene misregulation through chromosome architecture changes in LOS and BWS. In addition, comparative analyses between bovine and human will be conducted to determine whether conserved genomic regions between the two species also exhibit similar chromosome architectures in control group and similar changes in LOS and BWS. This will help us to better understand the roles and potential causes of the vulnerable genomic loci identified in Chapter 3.

For this project, we used bovine control and LOS fibroblast cells (in total 11 samples, including the samples used in Chapter 2) and human control and BWS fibroblast cells (in total 23 samples) for Hi-C sequencing, WGBS, and total RNA sequencing. Total RNA sequencing with ribosomal reduction, instead of

poly-A enriched RNA sequencing, will allow us to better characterize the expression of lncRNAs without a poly-A tail, such as *AIRN*. All the sequencings are currently underway by CD Genomics (New York, United States). I will analyze the sequencing data and prepare the manuscript for this project during the summer of 2022.

Appendix 2: Overgrowth Syndrome

INTRODUCTION

Overgrowth syndromes (OGSs) refer to a heterogeneous group of conditions found in many species (Lapunzina, 2005; Sinclair et al., 2000; Young et al., 1998), which show a common feature of excessive growth. According to the definitions used in humans, OGSs can be divided into 2 categories based on phenotypes: generalized OGSs and localized/partial OGSs (Lapunzina, 2005). Generalized OGSs often are characterized by a 2-3 standard deviations increase in overall growth parameters, including body weight, height, and head circumference (Elliott et al., 1994; Opitz et al., 1998). On the other hand, localized/partial OGSs result in overgrowth in 1 or few organs or regions of the body (Wiedemann et al., 1983). OGS also can be characterized as congenital and/or postnatal according to the age when phenotypes present (Biesecker, 2006; Cole and Hughes, 1994; DeBaun et al., 2003; Opitz et al., 1998). A greater risk of tumorigenesis is a shared feature of many OGSs found in humans (Lapunzina, 2005).

A generalized congenital OGS in bovine is known as LOS (Fig. 1) (Young et al., 1998). LOS refers to a group of abnormal phenotypes occurring in bovine and ovine fetuses, placentas, and newborns produced by assisted reproductive

technologies (ARTs). There are many ART-induced LOS calf reports from experimental studies (Behboodi et al., 1995; Farin and Farin, 1995; Hasler et al., 1995; Lazzari et al., 2002; van Wagtendonk-de Leeuw et al., 2000). Features of LOS include overgrowth, enlarged tongues, umbilical hernias, muscle and skeleton malformations, abnormal organ growth, allantois development defects, abnormal placental vasculature, and increased early embryo or fetus death rates (Farin et al., 2001, 2006; van Wagtendonk-de Leeuw et al., 1998; Walker et al., 1996). LOS can affect the dam and cause death of the afflicted animal, bringing financial loss to producers. Although there is a lack of published reports providing the incidence of LOS in ART-produced offspring, this incidence has been reported as high as 10% (Rocio M. Rivera, personal communications, 2017).

CLINICAL FEATURES OF LARGE OFFSPRING SYNDROME

Macrosomia refers to increased body and limb size, the most commonly identified feature of LOS (Farin et al., 2001; Walker et al., 1996; Young et al., 1998). This increased size of body and limbs can be 2 times and 5 times greater than the average size at birth, respectively (Walker et al., 1996), and the increased body size can be detected as early as the fifth week of gestation in cattle (Hansen et al., 2016). Increased skeletal lengths have been reported to be coupled with macrosomia (Farin and Farin, 1995). Calves with macrosomia at

birth, however, reach similar mature body weight as control animals (Wilson et al., 1995). Using the criteria for humans, birthweight greater than 2 times SD above the mean is defined as macrosomia (Brioude et al., 2018). Because LOS indicates overgrowth, it is easily taken for granted that macrosomia is a necessary feature. LOS, however, is not always characterized by overgrowth and thus sometimes referred to as abnormal offspring syndrome (Farin et al., 2006).

Macroglossia (enlarged tongue) is a feature of LOS (Chen et al., 2013). Severe macroglossia causes feeding and breathing difficulties. Abdominal wall defects, including omphalocele and umbilical hernia, also have been observed in LOS fetuses (Chen et al., 2013). An omphalocele is the outward protrusion of abdominal organs through the umbilical cord, with the organs not being covered by skin but by membranes (ie, amnion, peritoneum, and Wharton jelly (Bair et al., 1986)). Omphalocele is a severe defect present at birth and requires immediate corrective surgeries. An umbilical hernia is a bulge of abdominal organs at the umbilicus, which is caused by incomplete closure of umbilical ring and is covered by skin (Jackson and Moglen, 1970).

Organomegaly, the abnormal enlargement of organs, has been observed in heart, liver, and kidney of LOS calves (Farin and Farin, 1995; McEvoy et al., 1998). In addition, placentomegaly, an abnormally enlarged placenta, has been found in cows carrying *in vitro* fertilization (IVF)–conceived fetuses (Farin et al., 2001).

Other features, including increased incidence of hydrallantois (Hasler et al., 1995; van Wagtendonk-de Leeuw et al., 1998), increased gestation length (Sinclair et al., 1995), increased dystocia rate (Kruip and Den Daas, 1997), ataxia/paresis (Reichenbach et al., 1992; Schmidt et al., 1996), and abnormal limbs combined with abnormal spine (van Wagtendonk-de Leeuw et al., 1998), also have been observed in LOS calves.

ASSISTED REPRODUCTIVE TECHNOLOGIES AND LARGE OFFSPRING SYNDROME

In the late 1980s and 1990s, clinical epidemiologist Barker (Barker, 1995) suggested that the gestating maternal environment could have adverse consequences to the well-being of the offspring after birth. The phenomenon, which explains this permanent programming of the fetus, was named *fetal origins of adult disease* or the *Barker hypothesis*. The *developmental origins of health and disease* hypothesis, as the phenomenon is now known, propositions that the inherent developmental (genetic) program of an individual can be influenced by its environment, especially during critical periods of development, which can have significant long-term consequences for the well-being of the offspring during life. One artificial environment that has received much scientific attention for its potential to cause incorrect developmental programming to the resulting offspring

in humans and livestock animals is ARTs.

ARTs refer to a series of laboratory techniques and procedures used to conceive offspring. ART procedures include oocyte retrieval from ovaries, *in vitro* oocyte maturation, IVF, embryo culture, and embryo transfer (ET). ARTs are used in cattle to improve genetic merit of the offspring in a shorter length of time compared with natural reproduction. Genetic merit is defined as the rank of an animal for its ability to produce superior offspring relative to other selection candidates (Purdue Extension — <https://www.extension.purdue.edu/extmedia/nsif/nsif-8.pdf>). In addition, ART can be used to produce genetically manipulated animals with improved production traits (National Research Council, 2004).

Supplement of serum during *in vitro* embryo culture has been historically used to stimulate blastocyst formation (Edwards, 1965). Two experimental accounts suggest that serum can induce LOS in approximately 25% of ovine and bovine fetuses (Chen et al., 2015; Young et al., 2001). Adding fetal calf serum and bovine serum albumin during bovine embryo culture accelerates embryo development and improves blastocyst yield by day 6 but decreases embryo survival rate (Carolan et al., 1995; Thompson et al., 1998). When comparing ovine embryos cultured with or without human serum supplements, bovine serum albumin, and amino acid supplements, increased body weight and gestation length in the human serum group were observed (Thompson et al., 1995).

Coculture of embryos with various types of cells also has been used to increase blastocyst yield (Gandolfi and Moor, 1987). Similar to what has been found with serum supplementation in sheep, overgrowth (Holm et al., 1996; Maxfield et al., 1997) and increased gestation length (Holm et al., 1996) have been reported for ovine embryos that were cocultured with granulosa or oviduct epithelial cells. In addition, the size of the primary muscle fibers (which form during the first wave of myogenesis) and the ratio of secondary to primary fibers of the cocultured fetuses also were greater than in the controls, which indicate that hypertrophy of the primary fibers and hyperplasia of the secondary fibers are associated with the increased body weight observed in these fetuses (Maxfield et al., 1997).

A SIMILAR OVERGROWTH SYNDROME OCCURS IN HUMANS

In humans, Beckwith-Wiedemann syndrome (BWS) (OMIM #130650), a human OGS, has phenotypical and molecular similarities to LOS. The most current report indicates an incidence of BWS in approximately 1/11,000 natural births (Mussa et al., 2013). BWS is a heterogeneous condition for which various phenotypic and (epi)genetic defects have been reported. Clinical features of BWS include macroglossia, abdominal wall defects (omphalocele/hernia/diastasis recti), lateralized overgrowth, childhood tumors (Wilms tumor and hepatoblastoma), neonatal hypoglycemia, macrosomia (large

body size), ear malformations (creases/pits), facial nevus simplex (nevus flammeus or port-wine stain), and organomegaly (Brioude et al., 2018). The use of ARTs has been reported to increase the incidence of BWS by up to 10.7 times (Mussa et al., 2017; Vermeiden and Bernardus, 2013).

MOLECULAR FINDINGS OF LARGE OFFSPRING SYNDROME AND BECKWITH-WIEDEMANN SYNDROME

The main molecular defects of BWS occur on human chromosome 11p15 (bovine = chromosome 29) and include defects in DNA methylation, incorrect expression of imprinted genes, changes of chromosomal contents, and gene mutations (Beygo et al., 2016; DeBaun et al., 2003; Gicquel et al., 2003; Hatada et al., 1996; Henry et al., 1991; Okano et al., 1986; Reik et al., 1995; Schmutz, 1986; Turleau et al., 1984; Waziri et al., 1983). Among them, loss of imprinting caused by DNA methylation defects is the most frequently observed. DNA methylation (the addition of a methyl group [CH₃] to DNA) is an epigenetic mark involved in the control of gene expression. Genomic imprinting is an epigenetic phenomenon, which regulates parent-specific (ie, chromosome specific) gene expression of approximately 150 genes (ie, imprinted genes) in mammals (Blake et al., 2009; Chen et al., 2016b; Morison et al., 2001; Tian, 2014). These genes control growth and development of the fetus and the

placenta and their expression is tightly regulated by a discrete region of differential DNA methylation known as the imprinting center (IC) (Reik and Walter, 2001). One of these ICs, namely, IC2 (also known as KvDMR1) is the most common genomic region affected by DNA methylation defects in BWS and LOS (Chen et al., 2015). In a normal situation, the IC2 is methylated maternally inherited chromosome. This methylation state allows for the expression of the gene *KCNQ1OT1* from the paternal allele, which by attracting epigenetic modifiers, silences various flanking imprinted genes, including the cell cycle regulator *CDKN1C* (Horike et al., 2000; Lee et al., 1999; Smilnich et al., 1999). The methylated state of the maternal chromosome orchestrates the expression of several genes involved in fetal and placental growth (Horike et al., 2000; Lee et al., 1999; Smilnich et al., 1999). In LOS and BWS, imprinted gene expression regulated by the IC2 is lost as a result of loss of methylation of the maternal KvDMR1 (Brioude et al., 2013; Chen et al., 2015; Ibrahim et al., 2014; Mussa et al., 2016).

ALTERATIONS IN GENE EXPRESSION

Alterations in imprinted and nonimprinted gene expression as a result of *in vitro* embryo production have been reported in numerous studies in bovine (Chen et al., 2016b, 2017; Corcoran et al., 2006; Driver et al., 2012; Fair et al.,

2007; Rizos et al., 2002, 2003, 2004; Smith et al., 2009; Wrenzycki et al., 1999, 2001). Different culture media and supplementation with serum cause transcript abundance changes of several developmentally important genes involved in cell-cell junctions, transport, RNA processing, and stress in bovine embryos (Wrenzycki et al., 1999, 2001). The up-regulation of several developmentally important genes, including the imprinted gene *IGF1R*, have been suggested as early markers of LOS for bovine (Lazzari et al., 2002). A 2-fold increase in expression of the imprinted fetal growth factor *IGF2* transcript can be detected in liver of day 70 bovine fetuses cultured in medium containing estrus cow serum when compared with the serum-restricted group (Blondin et al., 2000).

SPONTANEOUS LARGE OFFSPRING SYNDROME

Although LOS cases in bovine have only been reported associated with ARTs, LOS can occur spontaneously. Spontaneous maternal-fetal disproportion is the predominant cause of dystocia in beef cattle (Zaborski et al., 2009).

Several environmental and genetic factors cause this disparity and it is associated most commonly with first-calf heifers (Holland and Odde, 1992; Zaborski et al., 2009). Although calves in this scenario may be relatively large to the dam, they may not be oversized in absolute terms of population normals. In the human literature, neonates that are large for gestational age typically are

above the 97th percentile for birthweight at delivery, although the definition varies slightly by condition (Kamien et al., 2018). Mechanistically, syndromes of overgrowth may be due to increased numbers of cells, hypertrophy, increases in the interstitium (such as fluid accumulation) or a combination of these conditions (Kamien et al., 2018). In humans, there are 2 broad categories of fetal OGSs. The first is those that are driven by the maternal environment, such as gestational diabetes, occurring in approximately 5% of all pregnancies (Kampmann et al., 2015). Conditions, such as these, predominantly result in symmetric hypertrophy of fetal tissue, particularly adipose tissue. The second category is neonates that are affected by either spontaneous or inherited genetic mutations, such as BWS, Sotos syndrome, and Proteus syndrome, to name a few. Although these conditions are rare, they are becoming increasingly recognized, due to increased utilization of IVF techniques (Bianci et al., 2010). Spontaneous fetal oversize syndromes have not been well recognized in food animal species outside of neonates generated by ART (Young et al., 1998). For the purposes of this review, LOS can be categorized by conditions associated with prolonged gestation or those of normal gestational length.

Gestational length is relatively constant in cattle within breeds and environmental conditions, ranging from 280 days to 290 days (Foote, 1981). Increasing gestational length within the normal range has been associated with larger birthweight (Holland and Odde, 1992). Gestational length is moderately

heritable and displays a gender bias, with male calves generally having longer gestational lengths than female calves (Holland and Odde, 1992). A definition for post-term for cattle has not been established but is generally in excess of 300 days. Prolonged gestation has long been associated with a poor outcome for the resultant neonate (Shibata and Ishihara, 1949). Several breeds, including Ayrshire, Holstein Friesian, Guernsey, Jersey, Swedish Red and White, and Belgian Blue cattle, have been documented with pathologically prolonged gestation (Buczinski et al., 2007; Cornillie, 2007; Graves et al., 1991). Prolonged gestation in these cases largely is secondary to a dysfunctional hypothalamic-pituitary-adrenal axis, with the calf failing to initiate parturition. These conditions have included adenohipophyseal hypoplasia/aplasia, cerebellar hypoplasia, and adrenal hypoplasia. A genetic mutation has been suspected in many of these cases, with the mode of inheritance established but putative genetic mutations not. Infectious causes, such as Akabane virus, bluetongue virus, and bovine viral diarrhea virus, and toxic causes, such as ingestion of *Veratrum californicum*, also may result in prolonged gestation due to dysfunction of the hypothalamic-pituitary axis (Constable et al., 2017). Prolonged gestation alone, however, does not always result in fetal oversize, with this syndrome only identified in Holstein Friesian, Swedish Red and White, and Ayrshire cattle (Holland and Odde, 1992; Young et al., 1998). Calves affected by overgrowth, commonly referred to as fetal giants, characteristically have long teeth, hair coats, and toes and otherwise

appear normal. These calves have been reported to weigh between 59 kg and 98 kg. Typically, these calves are delivered after induction of parturition and almost invariably necessitate caesarean section. The traits that produce prolonged gestation and fetal oversize largely are incompatible with postnatal life and these calves rarely survive for more than 24 hours.

Although rare, fetal giants also have been recognized in calves that are born at normal gestation lengths, with the predominant abnormality absolute oversize. No specific descriptions of these calves have been made, other than Roberts (Roberts, 1971) regarding any calf over 59 kg at birth to be a fetal giant. Despite reported by clinicians and producers, there are, to the authors' knowledge, no published descriptions of fetal giant LOS outside that of calves produced by IVF.

SPONTANEOUS LARGE OFFSPRING SYNDROME: CASE STUDIES

Three cases of S-LOS are illustrated in Fig. 2. The first calf was a purebred male Holstein Friesian calf that was delivered via caesarean section at 293 days of gestation. The calf weighed 83 kg at birth with the combined weight of the fetal membranes 12.7 kg. The calf had an appropriate hair coat, erupted teeth, and normal eponychium. The calf was proportionate with no obvious musculoskeletal defects. The calf did have an enlarged tongue (macroglossia)

and large umbilical hernia (omphalocele). The calf was unable to nurse due to the enlarged tongue and was humanely euthanized. Imaging of the brain by MRI and detailed necropsy definitively ruled out a structural alteration of the hypothalamus or pituitary or adrenal glands. The second calf was a female purebred Brown Swiss calf that was delivered at 283 days of gestation. The dam was initially presented due to concern for a hydrops condition with perceived overdistension of the abdomen at approximately 278 days of gestation. Based on palpation and transabdominal ultrasound, a hydrops condition was considered unlikely; however, the calf appeared very large. The calf was delivered by elective caesarean section after induction of parturition with cloprostenol and dexamethasone 40 hours prior to surgery. The viable calf weighed 63 kg at birth, with the fetal membranes unable to be weighed due to retention. The calf had small umbilical hernia and marked carpal contracture but was otherwise vigorous and healthy. The calf's legs were treated with splints and it was discharged at 10 days of age.

Although the definitive reason for fetal oversize has yet to be identified in these calves, it is strongly suspected that these calves represent S-LOS. After the identification of these 2 calves, other large-for-gestational-age calves not generated by IVF have been investigated and have been shown to possess the same epimutation as IVF-generated large offspring calves. This strongly emulates the analogous human condition BWS (Bianci et al., 2010).

Epigenetic conditions should be considered in cases of fetal gigantism where aberrations of gestational length have been excluded. At this time, risk factors have not been established and the occurrence of such calves is sporadic.

CONCLUDING REMARKS

The first ART calf was reported by Brackett and colleagues in 1982 (Brackett et al., 1982), whereas the fact that ARTs can induce the birth of abnormally large calves was first documented in the 1990s (Behboodi et al., 1995). Even though many reports have been published and almost 30 years have passed since the first LOS report, it is not yet possible to predict which embryos are molecularly programmed to suffer LOS, because the etiology of the syndrome is not known. Further, as Dr John F. Hasler states in a recent review of the ET industry, “there is a lack of peer-reviewed published data describing the current status of LOS problems in the commercial ET-IVP [embryo transfer - *in vitro* production] industry” (Hasler, 2014). Therefore, it is difficult to calculate what the current incidence of LOS is and how much producers are affected. As discussed previously, off-the-record conversations have stated that 10% LOS in some practices is not unusual. For obvious reasons, the authors do not believe that such numbers will ever be published but hope to be able to provide information that may be used by ET companies and others using ART embryos in

their practices to identify embryos molecularly programmed to suffer LOS prior to transfer.

Perinatal mortality—death occurring prior to, during, or within 48 hours of calving—is a recognized problem in the cattle industry (Mee, 2013). In developed countries, 25% to 46% of perinatal deaths in cattle are the result of dystocia, and, in the United States, 32% of perinatal deaths are due to unknown causes. One of the main causes of dystocia is fetal-maternal size mismatch. Dystocia has a direct negative impact on calves (Mee, 2013; Mee et al., 2014), dam survival and reproduction performance (Bicalho et al., 2007), and milk production (Bicalho et al., 2008). The authors' current research hopes to aims if similar genetic and epigenetic misregulation is the culprit of S-LOS, a previously uncharacterized syndrome in cattle, and to shed light on the contribution of S-LOS to the 25% to 46% rate of perinatal death in cattle resulting from unknown causes and dystocia.



Fig. 1. ART-produced LOS.

Large bull calf produced by *in vitro* procedures by Rocio Melissa Rivera while at the University of Florida. The picture was taken when the calf was 2 days of age. The calf weighed 98 kg at birth and died at 1 week of age as a result of complications relating to overgrowth, which included inability to stand up to suckle.



Fig. 2.

Fig. 2. Spontaneous LOS.

Case 1 (*top left and right*)—a post-term 86-kg (normal birth at birth 5 40–50 kg) Holstein bull calf was delivered by emergency cesarean section due to dystocia. The calf was macrosomic and had a marked omphalocele and macroglossia (*top right*). The calf was mentally inappropriate and was euthanized at 1 day of age. Immunohistochemistry of the anterior pituitary and hypothalamus could not demonstrate an aberration that could explain the macrosomia. Case 2 (*middle left and right*)—a preterm 63-kg Brown Swiss heifer (normal weight at birth approximately 45 kg) calf was delivered by planned cesarean section. The cow was referred for her large size approximately 2 weeks prior to being term. The calf was mentally appropriate at delivery and had an omphalocele and bilateral flexural deformities of the front metacarpophalangeal joint. The omphalocele was corrected surgically (*middle right*) and the flexural deformities by splints and physical therapy. The calf was discharged in good health and apparently is still performing well. Case 3 (*bottom left and right*)—1-day-old Holstein Friesian calf showing typical signs associated with LOS, including absolute macrosomia, omphalocele, and asymmetry of the pinna. The calf presented in respiratory distress and was later euthanized.

Appendix 3: Collaborative publications

In addition to the work presented in Chapters 2 and 3 and Appendix 1 and 2, I collaborated with other research groups within the Division of Animal Sciences, within Mizzou and with colleagues at other institutions in the USA during my PhD program. I also participated in one book chapter as coauthor. These publications will be briefly introduced below.

1. Modeling allele-specific expression at the gene and SNP levels simultaneously by a Bayesian logistic mixed regression model


Xie et al. *BMC Bioinformatics* (2019) 20:530
<https://doi.org/10.1186/s12859-019-3141-6>

BMC Bioinformatics

METHODOLOGY ARTICLE Open Access

Check for updates

Modeling allele-specific expression at the gene and SNP levels simultaneously by a Bayesian logistic mixed regression model

Jing Xie¹, Tieming Ji^{1*} , Marco A. R. Ferreira², Yahan Li³, Bhaumik N. Patel³ and Rocio M. Rivera³

Abstract

Background: High-throughput sequencing experiments, which can determine allele origins, have been used to assess genome-wide allele-specific expression. Despite the amount of data generated from high-throughput experiments, statistical methods are often too simplistic to understand the complexity of gene expression. Specifically, existing methods do not test allele-specific expression (ASE) of a gene as a whole and variation in ASE within a gene across exons separately and simultaneously.

Results: We propose a generalized linear mixed model to close these gaps, incorporating variations due to genes, single nucleotide polymorphisms (SNPs), and biological replicates. To improve reliability of statistical inferences, we assign priors on each effect in the model so that information is shared across genes in the entire genome. We utilize Bayesian model selection to test the hypothesis of ASE for each gene and variations across SNPs within a gene. We apply our method to four tissue types in a bovine study to de novo detect ASE genes in the bovine genome, and uncover intriguing predictions of regulatory ASEs across gene exons and across tissue types. We compared our method to competing approaches through simulation studies that mimicked the real datasets. The R package, BLMRM, that implements our proposed algorithm, is publicly available for download at <https://github.com/JingXieMIZZOU/BLMRM>.

Conclusions: We will show that the proposed method exhibits improved control of the false discovery rate and improved power over existing methods when SNP variation and biological variation are present. Besides, our method also maintains low computational requirements that allows for whole genome analysis.


Keywords: Allelic imbalance, Hierarchical generalized linear mixed model, High-throughput sequencing experiments, Single nucleotide polymorphism

PMID: 31660858. <https://doi.org/10.1186/s12859-019-3141-6>

In this peer-reviewed article we demonstrated a statistical model to identify allele-specific expression of genes based on single nucleotide polymorphisms (SNP) detected by RNA sequencing (Xie et al., 2019). This model












can be applied to individual exons to study alternative splicing accompanied with allele-specific expression. I participated in this work for bioinformatic analyses of RNA sequencing and SNP data, demonstration of biological relevance, and manuscript writing and revising.

2. Using online tools at the Bovine Genome Database to manually annotate genes in the new reference genome

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics 

doi: 10.1111/age.12962

Using online tools at the Bovine Genome Database to manually annotate genes in the new reference genome

D. A. Triant*, J. J. Le Tourneau*, C. M. Diesh[†], D. R. Unni[‡], M. Shamimuzzaman*, A. T. Walsh*, J. Gardiner*, A. K. Goldkamp[§], Y. Li* , H. N. Nguyen* , C. Roberts*, Z. Zhao** , L. J. Alexander^{††}, J. E. Decker* , R. D. Schnabel* , S. G. Schroeder^{‡‡}, T. S. Sonstegard^{§§}, J. F. Taylor* , R. M. Rivera* , D. E. Hagen[§]  and C. G. Elsik*   

*Division of Animal Sciences, University of Missouri, Columbia, MO 65211, USA. [†]Department of Bioengineering, University of California, Berkeley, Berkeley, CA 94720, USA. [‡]Division of Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, CA 94608, USA. [§]Department of Animal and Food Sciences, Oklahoma State University, Stillwater, OK 74078, USA. [¶]MU Institute for Data Science and Informatics, University of Missouri, Columbia, MO 65211, USA. ^{**}Division of Plant Sciences, University of Missouri, Columbia, MO 65211, USA. ^{††}USDA-ARS-PA-Livestock & Range Research Lab, Miles City, MT 59301, USA. ^{‡‡}USDA-ARS Animal Genomics and Improvement Lab, Beltsville, MD 20705, USA. ^{§§}Acceligen, Eagan, MN 55121, USA.

Summary With the availability of a new highly contiguous *Bos taurus* reference genome assembly (ARS-UCD1.2), it is the opportune time to upgrade the bovine gene set by seeking input from researchers. Furthermore, advances in graphical genome annotation tools now make it possible for researchers to leverage sequence data generated with the latest technologies to collaboratively curate genes. For many years the Bovine Genome Database (BGD) has provided tools such as the APOLLO genome annotation editor to support manual bovine gene curation. The goal of this paper is to explain the reasoning behind the decisions made in the manual gene curation process while providing examples using the existing BGD tools. We will describe the sources of gene annotation evidence provided at the BGD, including RNA-seq and Iso-Seq data. We will also explain how to interpret various data visualizations when curating gene models, and will demonstrate the value of manual gene annotation. The process described here can be applied to manual gene curation for other species with similar tools. With a better understanding of manual gene annotation, researchers will be encouraged to edit gene models and contribute to the enhancement of livestock gene sets.

Keywords *Bos taurus*, gene prediction, genome annotation, genome annotation tools, RNA-seq

PMID: 32537769. <https://doi.org/10.1111/age.12962>

This peer-reviewed article demonstrated online tools for bovine genome annotation developed by Dr. Elsik's laboratory (Triant et al., 2020). These tools allow users to modify the annotation of bovine genes based on their own

sequencing data and also serve as a database. I participated in this work for providing advice for the functions of the tools as a user and manuscript revising.

3. Characterization of tRNA expression profiles in large offspring syndrome

Goldkamp et al. *BMC Genomics* (2022) 23:273
<https://doi.org/10.1186/s12864-022-08496-7>

BMC Genomics

RESEARCH Open Access

Check for updates

Characterization of tRNA expression profiles in large offspring syndrome

Anna K. Goldkamp¹, Yahan Li², Rocio M. Rivera² and Darren E. Hagen^{1*}

Abstract

Background: Assisted Reproductive Technologies (ART) use can increase the risk of congenital overgrowth syndromes, such as large offspring syndrome (LOS) in ruminants. Epigenetic variations are known to influence gene expression and differentially methylated regions (DMRs) were previously determined to be associated with LOS in cattle. We observed DMRs overlapping tRNA clusters which could affect tRNA abundance and be associated with tissue specificity or overgrowth. Variations in tRNA expression have been identified in several disease pathways suggesting an important role in the regulation of biological processes. Understanding the role of tRNA expression in cattle offers an opportunity to reveal mechanisms of regulation at the translational level. We analyzed tRNA expression in the skeletal muscle and liver tissues of day 105 artificial insemination-conceived, ART-conceived with a normal body weight, and ART-conceived bovine fetuses with a body weight above the 97th percentile compared to Control-AI.

Results: Despite the centrality of tRNAs to translation, in silico predictions have revealed dramatic differences in the number of tRNA genes between humans and cattle (597 vs 1,659). Consistent with reports in human, only a fraction of predicted tRNA genes are expressed. We detected the expression of 474 and 487 bovine tRNA genes in the muscle and liver with the remainder being unexpressed. 193 and 198 unique tRNA sequences were expressed in all treatment groups within muscle and liver respectively. In addition, an average of 193 tRNA sequences were expressed within the same treatment group in different tissues. Some tRNA isodecoders were differentially expressed between treatment groups. In the skeletal muscle and liver, we categorized 11 tRNA isoacceptors with undetected expression as well as an isodecoder that was unexpressed in the liver (Ser^{GGA}). Our results identified variation in the proportion of tRNA gene copies expressed between tissues and differences in the highest contributing tRNA anticodon within an amino acid family due to treatment and tissue type. Out of all amino acid families, roughly half of the most highly expressed tRNA isoacceptors correlated to their most frequent codon in the bovine genome.

Conclusion: Although the number of bovine tRNA genes is nearly triple of that of the tRNA genes in human, there is a shared occurrence of transcriptionally inactive tRNA genes in both species. We detected differential expression of tRNA genes as well as tissue- and treatment- specific tRNA transcripts with unique sequence variations that could modulate translation during protein homeostasis or cellular stress, and give rise to regulatory products targeting genes related to overgrowth in the skeletal muscle and/or tumor development in the liver of LOS individuals. While the absence of certain isodecoders may be relieved by wobble base pairing, missing tRNA species could increase the likelihood of mistranslation or mRNA degradation.

Keywords: tRNA, Bovine, Protein translation, Large Offspring Syndrome

PMID: 35392796. <https://doi.org/10.1186/s12864-022-08496-7>

In this peer-reviewed article we conducted the first in-depth analyses of transfer RNA (tRNA) expression in bovine and changes in LOS (Goldkamp et al.,

2022). We found tissue-specific expression of over 30 tRNA genes between muscle and liver in controls, as well as tissue-specific usage of 11 tRNA anticodons. The upregulation of tRNA genes His^{GUG} was found in LOS group in muscle when compared to control group. Misregulation of several other tRNA genes were found in both ART-normal and LOS groups which reflects the impacts of ART. I participated in this work for data analyses and interpretation and manuscript revising.

4. Identification of large offspring syndrome during pregnancy through ultrasonography and maternal blood transcriptome analyses

Article

Identification of large offspring syndrome during pregnancy through ultrasonography and maternal blood transcriptome analyses

Rocio Melissa Rivera, Anna Katherine Goldkamp, Bhaumik Narendrabhai Patel, and 16 more

This is a preprint; it has not been peer reviewed by a journal.

<https://doi.org/10.21203/rs.3.rs-1511098/v1>
This work is licensed under a CC BY 4.0 License

Abstract

The use of assisted reproductive technologies (ART) in cattle can result in large/abnormal offspring syndrome (LOS/AOS) which is characterized by macrosomia. LOS can cause dystocia and lead to the death of dam and calf. Currently, no test exists to identify LOS pregnancies. We hypothesized that fetal ultrasonography and/or maternal blood markers are useful to identify LOS. Bovine fetuses were generated by artificial insemination (control) or ART. Fetal ultrasonographies were taken on gestation day 55 (D55) and fetal collections performed on D56 or D105 (gestation in cattle ≈280 days). ART fetuses weighing ≥97 percentile of the control weight were considered LOS. Ultrasonography results show that the product of six D55 measurements can be used to identify extreme cases of LOS. To determine whether maternal blood can be used to identify LOS, leukocyte mRNA from 23 females was sequenced. Unsupervised hierarchical clustering grouped the transcriptomes of the two females carrying the two largest LOS fetuses. Comparison of the leukocyte transcriptomes of these two females to the transcriptome of all other females identified several misregulated transcripts on gestation D55 and D105 with *LOC783838* and *PCDH1* being misregulated at both time-points. Together our data suggest that LOS is identifiable during pregnancy in cattle.

Large offspring syndrome abnormal offspring syndrome assisted reproductive technologies fetal morphometry maternal blood biomarker Beckwith-Wiedemann Syndrome

Status: **Under Review**

scientific reports

Version 1
posted 08 Apr, 2022

- Reviews received at journal 26 Apr, 2022
- Reviewers agreed at journal 06 Apr, 2022
- Reviewers invited by journal 06 Apr, 2022
- Editor assigned by journal 06 Apr, 2022
- Editor invited by journal 06 Apr, 2022
- Submission checks completed at journal 06 Apr, 2022
- First submitted to journal 31 Mar, 2022

You are reading this latest preprint version

Badges 0

Citations See more

Engagement 31 views

Tweets 1

Comments 0

<https://doi.org/10.21203/rs.3.rs-1511098/v1>

This manuscript is currently under review by Scientific Reports. In this study, we investigated the correlation between day 55 ultrasonography measurements and fetal measurements at different developmental stages (day 55 or 105) for control and LOS fetuses. Some of the measurements, including abdominal diameter, abdominal height, crown-rump length, head length, thoracic diameter, and thoracic height show slight or moderate positive correlations in a sex and/or stage specific manner. This suggests that early-stage ultrasonography

could be used for LOS diagnosis. In addition, day 55 and 105 maternal blood transcriptome were analyzed and over 40 differentially expressed genes were identified between LOS and control groups, which have the potential to serve as biomarkers for early diagnosis of LOS. I participated in this work for animal experiments, data analyses and interpretation, and manuscript revising.

5. Abnormal Offspring Syndrome

Chapter 71

Abnormal Offspring Syndrome

Rocío Melissa Rivera, Callum George Donnelly, Bhaumik Narendrabhai Patel, Yahan Li, Edgar Joel Soto-Moreno

Book Editor(s): Richard M. Hopper DVM, Diplomat ACT

First published: 03 May 2021 | <https://doi.org/10.1002/9781119602484.ch71>

[OpenURL](#)

[PDF](#) [TOOLS](#) [SHARE](#)

 [Bovine Reproduction, Second Edition](#)

[Related](#) [Information](#)

Details

© 2021 John Wiley & Sons, Inc.

Keywords

abnormal offspring syndrome
assisted reproductive technologies
Beckwith–Wiedemann syndrome
DNA methylation
epigenetic defects
large offspring syndrome
loss-of-imprinting

Publication History

Published Online:
03 May 2021
Published Print:
20 August 2021

Summary

Large offspring syndrome (LOS) refers to an overgrowth phenotype sometimes observed in cattle fetuses and newborns produced by the use of assisted reproductive technologies (ART). This chapter discusses the current knowledge on ART-induced abnormal offspring syndrome (AOS). It also discusses findings from sheep as AOS has also been documented in this species, and from Beckwith–Wiedemann syndrome (BWS), a human congenital overgrowth condition which phenocopies AOS. In addition, the chapter presents evidence of spontaneous AOS in cattle and provides examples of financial loss incurred by the cattle industry as a result of this syndrome. As the causes of BWS at the molecular level have been largely attributed to epigenetic defects, which include alterations of DNA methylation and loss-of-imprinting, it discusses these epigenetic modifications/mechanisms. Early identification of AOS would allow early elective termination of the pregnancy in order to prevent the monetary losses and animal welfare concerns associated with AOS.

<https://doi.org/10.1002/9781119602484.ch71>

In this book chapter we reviewed clinical and molecular findings of LOS/AOS, discussed potential causes for molecular changes in LOS, and highlighted the presence of spontaneous LOS and economic loss due to LOS. I contributed to this book chapter for the epigenetic sections.

Bibliography

- Abney, J.R., Cutler, B., Fillbach, M.L., Axelrod, D., and Scalettar, B.A. (1997). Chromatin dynamics in interphase nuclei and its implications for nuclear structure. *The Journal of Cell Biology* *137*, 1459–1468. .
- Alam, S.G., Zhang, Q., Prasad, N., Li, Y., Chamala, S., Kuchibhotla, R., Kc, B., Aggarwal, V., Shrestha, S., and Jones, A.L. (2016). The mammalian LINC complex regulates genome transcriptional responses to substrate rigidity. *Scientific Reports* *6*, 1–11. .
- Alberts, B. (2002). *Molecular Biology of the Cell: Hauptbd* (Garland).
- Albiez, H., Cremer, M., Tiberi, C., Vecchio, L., Schermelleh, L., Dittrich, S., Küpper, K., Joffe, B., Thormeyer, T., and von Hase, J. (2006). Chromatin domains and the interchromatin compartment form structurally defined and functionally interacting nuclear networks. *Chromosome Research* *14*, 707–733. .
- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* *31*, 166–169. .
- Anderson, E., Devenney, P.S., Hill, R.E., and Lettice, L.A. (2014). Mapping the Shh long-range regulatory domain. *Development* *141*, 3934–3943. .
- Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., and Duboule, D. (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science (New York, N.Y.)* *340*, 1234167. .
- Arima, T., Kamikihara, T., Hayashida, T., Kato, K., Inoue, T., Shirayoshi, Y., Oshimura, M., Soejima, H., Mukai, T., and Wake, N. (2005). ZAC, LIT1 (KCNQ1OT1) and p57 KIP2 (CDKN1C) are in an imprinted gene network that may play a role in Beckwith–Wiedemann syndrome. *Nucleic Acids Research* *33*, 2650–2660. .
- Arita, K., Isogai, S., Oda, T., Unoki, M., Sugita, K., Sekiyama, N., Kuwata, K., Hamamoto, R., Tochio, H., and Sato, M. (2012). Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. *Proceedings of the National Academy of Sciences* *109*, 12950–12955. .
- Arnold, D.R., Gaspar, R.C., da Rocha, C.V., Sangalli, J.R., de Bem, T.H., Corrêa, C.A., Penteadó, J.C., Meirelles, F.V., and Lopes, F.L. (2017). Nuclear transfer alters placental gene expression and associated histone modifications of the placental-specific imprinted gene pleckstrin homology-like domain, family A, member 2 (PHLDA2) in cattle. *Reproduction, Fertility and Development* *29*, 458–467. .
- Bair, J.H., Russ, P.D., Pretorius, D.H., Manchester, D., and Manco-Johnson, M.L. (1986). Fetal

omphalocele and gastroschisis: a review of 24 cases. *American Journal of Roentgenology* 147, 1047–1051. .

Barker, D.J. (1995). The fetal and infant origins of disease. *European Journal of Clinical Investigation* 25, 457–463. .

Battistelli, C., Busanello, A., and Maione, R. (2014). Functional interplay between MyoD and CTCF in regulating long-range chromatin interactions during differentiation. *Journal of Cell Science* 127, 3757–3767. .

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* 520, 243–247. .

Bauvois, B. (2012). New facets of matrix metalloproteinases MMP-2 and MMP-9 as cell surface transducers: outside-in signaling and relationship to tumor progression. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* 1825, 29–36. .

Becker, J.S., Nicetto, D., and Zaret, K.S. (2016). H3K9me3-dependent heterochromatin: barrier to cell fate changes. *Trends in Genetics* 32, 29–41. .

Beckwith, J.B., Wang, C.I., Donnell, G.N., and Gwinn, J.L. (1964). Hyperplastic fetal visceromegaly with macroglossia, omphalocele, cytomegaly of adrenal fetal cortex, postnatal somatic gigantism, and other abnormalities: Newly recognized syndrome. *Proceedings of the American Pediatric Society, Seattle* 16–18. .

Behboodi, E., Anderson, G.B., BonDurant, R.H., Cargill, S.L., Kreuzer, B.R., Medrano, J.F., and Murray, J.D. (1995). Birth of large calves that developed from in vitro-derived bovine embryos. *Theriogenology* 44, 227–232. .

Belin, B.J., Lee, T., and Mullins, R.D. (2015). DNA damage induces nuclear actin filament assembly by Formin-2 and Spire-1/2 that promotes efficient DNA repair. *Elife* 4, e07735. .

Belyaeva, A., Venkatachalapathy, S., Nagarajan, M., Shivashankar, G.V., and Uhler, C. (2017). Network analysis identifies chromosome intermingling regions as regulatory hotspots for transcription. *Proceedings of the National Academy of Sciences* 114, 13714–13719. .

Berland, S., Rustad, C.F., Bentsen, M.H.L., Wollen, E.J., Turowski, G., Johansson, S., Houge, G., and Haukanes, B.I. (2021). Double paternal uniparental isodisomy 7 and 15 presenting with Beckwith-Wiedemann spectrum features. *Molecular Case Studies* mcs. a006113. .

Berven, L.A., Willard, F.S., and Crouch, M.F. (2004). Role of the p70S6K pathway in regulating the actin cytoskeleton and cell migration. *Experimental Cell Research* 296, 183–195. .

Beygo, J., Joksic, I., Strom, T.M., Lüdecke, H.-J., Kolarova, J., Siebert, R., Mikovic, Z., Horsthemke, B., and Buiting, K. (2016). A maternal deletion upstream of the imprint control region 2 in 11p15 causes loss of methylation and familial Beckwith–Wiedemann syndrome. *European Journal of Human Genetics* *24*, 1280. .

Bhattacharya, S., Levy, M.J., Zhang, N., Li, H., Florens, L., Washburn, M.P., and Workman, J.L. (2021). The methyltransferase SETD2 couples transcription and splicing by engaging mRNA processing factors through its SHI domain. *Nature Communications* *12*, 1–16. .

Bhogaraju, S., Cajanek, L., Fort, C., Blisnick, T., Weber, K., Taschner, M., Mizuno, N., Lamla, S., Bastin, P., and Nigg, E.A. (2013). Molecular basis of tubulin transport within the cilium by IFT74 and IFT81. *Science* *341*, 1009–1012. .

Bian, C., and Yu, X. (2014). PGC7 suppresses TET3 for protecting DNA methylation. *Nucleic Acids Research* *42*, 2893–2905. .

Bianci, D.W., Crombleholme, T.M., Dalton, M.E., and Malone, F.D. (2010). *Fetology: diagnosis and management of the fetal patient* 2nd edition (McGraw-Hill Professional).

Bicalho, R.C., Galvão, K.N., Cheong, S.H., Gilbert, R.O., Warnick, L.D., and Guard, C.L. (2007). Effect of stillbirths on dam survival and reproduction performance in Holstein dairy cows. *Journal of Dairy Science* *90*, 2797–2803. .

Bicalho, R.C., Galvão, K.N., Warnick, L.D., and Guard, C.L. (2008). Stillbirth parturition reduces milk production in Holstein cows. *Preventive Veterinary Medicine* *84*, 112–120. .

Biesecker, L. (2006). The challenges of Proteus syndrome: diagnosis and management. *European Journal of Human Genetics* *14*, 1151. .

Bina, M. (2020). Discovering candidate imprinted genes and imprinting control regions in the human genome. *BMC Genomics* *21*, 1–16. .

Blake, A., Pickford, K., Greenaway, S., Thomas, S., Pickard, A., Williamson, C.M., Adams, N.C., Walling, A., Beck, T., and Fray, M. (2009). MouseBook: an integrated portal of mouse resources. *Nucleic Acids Research* *38*, D593–D599. .

Bliek, J., Gicquel, C., Maas, S., Gaston, V., Le Bouc, Y., and Mannens, M. (2004). Epigenotyping as a tool for the prediction of tumor risk and tumor type in patients with Beckwith-Wiedemann syndrome (BWS). *The Journal of Pediatrics* *145*, 796–799. .

Bliek, J., Verde, G., Callaway, J., Maas, S.M., De Crescenzo, A., Sparago, A., Cerrato, F., Russo, S., Ferraiuolo, S., and Rinaldi, M.M. (2009). Hypomethylation at multiple maternally methylated imprinted regions including PLAGL1 and GNAS loci in Beckwith–Wiedemann syndrome. *European Journal of Human Genetics* *17*, 611. .

Blondin, P., Farin, P.W., Crosier, A.E., Alexander, J.E., and Farin, C.E. (2000). In vitro production of embryos alters levels of insulin-like growth factor-II messenger ribonucleic acid in bovine fetuses 63 days after transfer. *Biology of Reproduction* 62, 384–389. .

Boissonnault, J.S., and Blaschak, M.J. (1988). Incidence of diastasis recti abdominis during the childbearing year. *Physical Therapy* 68, 1082–1086. .

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 30, 2114–2120. .

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Müller, S., Eils, R., Cremer, C., and Speicher, M.R. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol* 3, e157. .

Boonen, S.E., Hahnemann, J., Mackay, D., Tommerup, N., Brøndum-Nielsen, K., Tümer, Z., and Grønskov, K. (2012). No evidence for pathogenic variants or maternal effect of ZFP57 as the cause of Beckwith–Wiedemann Syndrome. *European Journal of Human Genetics* 20, 119–121. .

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. *Nature Genetics* 42, 1093–1100. .

Bostick, M., Kim, J.K., Estève, P.-O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317, 1760–1764. .

Bourc'his, D., Xu, G.-L., Lin, C.-S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science (New York, N.Y.)* 294, 2536–2539. .

Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S., and Crawford, G.E. (2008). High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132, 311–322. .

Boyle, S., Gilchrist, S., Bridger, J.M., Mahy, N.L., Ellis, J.A., and Bickmore, W.A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Human Molecular Genetics* 10, 211–220. .

Boyle, S., Rodesch, M.J., Halvensleben, H.A., Jeddloh, J.A., and Bickmore, W.A. (2011). Fluorescence in situ hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis. *Chromosome Research* 19, 901–909. .

Brackett, B.G., Bousquet, D., Boice, M.L., Donawick, W.J., Evans, J.F., and Dressel, M.A. (1982). Normal development following in vitro fertilization in the cow. *Biology of Reproduction* 27, 147–158. .

Brioude, F., Lacoste, A., Netchine, I., Vazquez, M.-P., Auber, F., Audry, G., Gauthier-Villars, M., Brugieres,

L., Gicquel, C., and Le Bouc, Y. (2013). Beckwith-Wiedemann syndrome: growth pattern and tumor risk according to molecular mechanism, and guidelines for tumor surveillance. *Hormone Research in Paediatrics* *80*, 457–465. .

Brioude, F., Kalish, J.M., Mussa, A., Foster, A.C., Bliiek, J., Ferrero, G.B., Boonen, S.E., Cole, T., Baker, R., and Bertoletti, M. (2018). Clinical and molecular diagnosis, screening and management of Beckwith–Wiedemann syndrome: an international consensus statement. *Nature Reviews Endocrinology* *14*, 229–249. .

Broad Institute (2021). Picard toolkit (Broad Institute, GitHub Repository).

Buczinski, S., Bélanger, A.-M., Fecteau, G., and Roy, J.-P. (2007). Prolonged gestation in two Holstein cows: Transabdominal ultrasonographic findings in late pregnancy and pathologic findings in the fetuses. *Journal of Veterinary Medicine Series A* *54*, 624–626. .

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods* *10*, 1213. .

Bushnell, B. (2021). BBMap.

Buxboim, A., Swift, J., Irianto, J., Spinler, K.R., Dingal, P.D.P., Athirasala, A., Kao, Y.-R.C., Cho, S., Harada, T., and Shin, J.-W. (2014). Matrix elasticity regulates lamin-A, C phosphorylation and turnover with feedback to actomyosin. *Current Biology* *24*, 1909–1917. .

Cairo, S., Armengol, C., De Reyniès, A., Wei, Y., Thomas, E., Renard, C.-A., Goga, A., Balakrishnan, A., Semeraro, M., and Gresh, L. (2008). Hepatic stem-like phenotype and interplay of Wnt/ β -catenin and Myc signaling in aggressive childhood liver cancer. *Cancer Cell* *14*, 471–484. .

Canovas, S., Ivanova, E., Romar, R., García-Martínez, S., Soriano-Ubeda, C., García-Vázquez, F.A., Saadeh, H., Andrews, S., Kelsey, G., and Coy, P. (2017). DNA methylation and gene expression changes derived from assisted reproductive technologies can be decreased by reproductive fluids. *Elife* *6*, e23670. .

Cao, P., Li, H., Zuo, Y., and Nashun, B. (2020). Characterization of DNA methylation patterns and mining of epigenetic markers during genomic reprogramming in SCNT embryos. *Frontiers in Cell and Developmental Biology* *8*, 877. .

Caridi, C.P., D’Agostino, C., Ryu, T., Zapotoczny, G., Delabaere, L., Li, X., Khodaverdian, V.Y., Amaral, N., Lin, E., and Rau, A.R. (2018). Nuclear F-actin and myosins drive relocalization of heterochromatic breaks. *Nature* *559*, 54–60. .

Carolan, C., Lonergan, P., Van Langendonck, A., and Mermillod, P. (1995). Factors affecting bovine embryo development in synthetic oviduct fluid following oocyte maturation and fertilization in vitro. *Theriogenology* *43*, 1115–1128. .

- Caspary, T., Cleary, M.A., Perlman, E.J., Zhang, P., Elledge, S.J., and Tilghman, S.M. (1999). Oppositely imprinted genes p57Kip2 and Igf2 interact in a mouse model for Beckwith–Wiedemann syndrome. *Genes & Development* *13*, 3115–3124. .
- Chan, S.S.-K., Hagen, H.R., Swanson, S.A., Stewart, R., Boll, K.A., Aho, J., Thomson, J.A., and Kyba, M. (2016). Development of bipotent cardiac/skeletal myogenic progenitors from MESP1+ mesoderm. *Stem Cell Reports* *6*, 26–34. .
- Chang, L., and Goldman, R.D. (2004). Intermediate filaments mediate cytoskeletal crosstalk. *Nature Reviews Molecular Cell Biology* *5*, 601–613. .
- Chang, S., Hur, S.K., Naveh, N.S.S., Thorvaldsen, J.L., French, D.L., Gagne, A.L., Jobaliya, C.D., Anguera, M.C., Bartolomei, M.S., and Kalish, J.M. (2021). Derivation and investigation of the first human cell-based model of Beckwith-Wiedemann syndrome. *Epigenetics* *16*, 1295–1305. .
- Chason, R.J., Csokmay, J., Segars, J.H., DeCherney, A.H., and Armant, D.R. (2011). Environmental and epigenetic effects upon preimplantation embryo metabolism and development. *Trends in Endocrinology & Metabolism* *22*, 412–420. .
- Chen, B., Deng, M., Pan, M.-H., Sun, S.-C., and Liu, H. (2022). Regulation of paternal 5mC oxidation and H3K9me2 asymmetry by ERK1/2 in mouse zygotes. *Cell & Bioscience* *12*, 1–15. .
- Chen, J., Yao, Z.-X., Chen, J.-S., Gi, Y.J., Muñoz, N.M., Kundra, S., Herlong, H.F., Jeong, Y.S., Goltsov, A., and Ohshiro, K. (2016a). TGF- β / β 2-spectrin/CTCF-regulated tumor suppression in human stem cell disorder Beckwith-Wiedemann syndrome. *The Journal of Clinical Investigation* *126*, 527–542. .
- Chen, X., Ke, Y., Wu, K., Zhao, H., Sun, Y., Gao, L., Liu, Z., Zhang, J., Tao, W., and Hou, Z. (2019). Key role for CTCF in establishing chromatin structure in human embryos. *Nature* *576*, 306–310. .
- Chen, Z., Robbins, K.M., Wells, K.D., and Rivera, R.M. (2013). Large offspring syndrome: a bovine model for the human loss-of-imprinting overgrowth syndrome Beckwith-Wiedemann. *Epigenetics* *8*, 591–601. <https://doi.org/10.4161/epi.24655>.
- Chen, Z., Hagen, D.E., Elsik, C.G., Ji, T., Morris, C.J., Moon, L.E., and Rivera, R.M. (2015). Characterization of global loss of imprinting in fetal overgrowth syndrome induced by assisted reproduction. *Proc Natl Acad Sci U S A* *112*, 4618–4623. <https://doi.org/10.1073/pnas.1422088112>.
- Chen, Z., Hagen, D.E., Wang, J., Elsik, C.G., Ji, T., Siqueira, L.G., Hansen, P.J., and Rivera, R.M. (2016b). Global assessment of imprinted gene expression in the bovine conceptus by next generation sequencing. *Epigenetics* *11*, 501–516. <https://doi.org/10.1080/15592294.2016.1184805>.
- Chen, Z., Hagen, D.E., Ji, T., Elsik, C.G., and Rivera, R.M. (2017). Global misregulation of genes largely uncoupled to DNA methylome epimutations characterizes a congenital overgrowth syndrome. *Scientific Reports* *7*, 12667. .

Chen, Z., Gao, W., Pu, L., Zhang, L., Han, G., Zuo, X., Zhang, Y., Li, X., Shen, H., and Wu, J. (2018). PRDM8 exhibits antitumor activities toward hepatocellular carcinoma by targeting NAP1L1. *Hepatology* 68, 994–1009. .

Christoffels, V.M., Grieskamp, T., Norden, J., Mommersteeg, M.T., Rudat, C., and Kispert, A. (2009). Tbx18 and the fate of epicardial progenitors. *Nature* 458, E8–E9. .

Chuang, C.-H., Carpenter, A.E., Fuchsova, B., Johnson, T., de Lanerolle, P., and Belmont, A.S. (2006). Long-range directional movement of an interphase chromosome site. *Current Biology* 16, 825–831. .

Chubb, J.R., Boyle, S., Perry, P., and Bickmore, W.A. (2002). Chromatin motion is constrained by association with nuclear compartments in human cells. *Current Biology* 12, 439–445. .

Ciccone, D.N., Su, H., Hevi, S., Gay, F., Lei, H., Bajko, J., Xu, G., Li, E., and Chen, T. (2009). KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 461, 415–418. .

Clericuzio, C.L. (1999). Recognition and management of childhood cancer syndromes: a systems approach. *American Journal of Medical Genetics Part A* 89, 81–90. .

Cole, T.R., and Hughes, H.E. (1994). Sotos syndrome: a study of the diagnostic criteria and natural history. *Journal of Medical Genetics* 31, 20–32. .

Coleman, L., Back, P., Blair, H., López-Villalobos, N., and Hickson, R. (2021). Sire effects on birth weight, gestation length, and pre-weaning growth of beef-cross-dairy calves: a case study in New Zealand. *Dairy* 2, 385–395. .

Constable, P.D., Hinchcliff, K.W., Done, S.H., and Grünberg, W. (2017). *Veterinary medicine : A textbook of the diseases of cattle, horses, sheep, pigs, and goats (Edition 11)* (Elsevier Health Sciences).

Cooper, W.N., Luharia, A., Evans, G.A., Raza, H., Haire, A.C., Grundy, R., Bowdin, S.C., Riccio, A., Sebastio, G., and Blik, J. (2005). Molecular subtypes and phenotypic expression of Beckwith–Wiedemann syndrome. *European Journal of Human Genetics* 13, 1025. .

Corcoran, D., Fair, T., Park, S., Rizos, D., Patel, O.V., Smith, G.W., Coussens, P.M., Ireland, J.J., Boland, M.P., and Evans, A.C.O. (2006). Suppressed expression of genes involved in transcription and translation in in vitro compared with in vivo cultured bovine embryos. *Reproduction (Cambridge, England)* 131, 651–660. .

Cornillie, P.V.D.B. (2007). Prolonged gestation in two Belgian blue cows due to inherited adenyphophyseal hypoplasia in the fetuses. *The Veterinary Record* 161, 388–391. .

Craggs, G., Finan, P.M., Lawson, D., Wingfield, J., Perera, T., Gadher, S., Totty, N.F., and Kellie, S. (2001). A nuclear SH3 domain-binding protein that colocalizes with mRNA splicing factors and intermediate filament-containing perinuclear networks. *Journal of Biological Chemistry* 276, 30552–30560. .

Cremer, T., and Cremer, C. (2006a). Rise, fall and resurrection of chromosome territories: a historical perspective Part I. The rise of chromosome territories. *European Journal of Histochemistry* 161–176. .

Cremer, T., and Cremer, C. (2006b). Rise, fall and resurrection of chromosome territories: a historical perspective Part II. Fall and resurrection of chromosome territories during the 1950s to 1980s. Part III. Chromosome territories and the functional nuclear architecture: experiments and m. *European Journal of Histochemistry* 223–272. .

Cremer, T., Kurz, A., Zirbel, R., Dietzel, S., Rinke, B., Schröck, E., Speicher, M.R., Mathieu, U., Jauch, A., and Emmerich, P. (1993). Role of chromosome territories in the functional compartmentalization of the cell nucleus. (Cold Spring Harbor Laboratory Press), pp. 777–792.

Cui, C., Zhang, H., Guo, L.-N., Zhang, X., Meng, L., Pan, X., and Wei, Y. (2016). Inhibitory effect of NBL1 on PDGF-BB-induced human PSMC proliferation through blockade of PDGF β -p38MAPK pathway. *Bioscience Reports* 36, e00374. .

Daelemans, C., Ritchie, M.E., Smits, G., Abu-Amero, S., Sudbery, I.M., Forrest, M.S., Campino, S., Clark, T.G., Stanier, P., and Kwiatkowski, D. (2010). High-throughput analysis of candidate imprinted genes and allele-specific gene expression in the human term placenta. *BMC Genetics* 11, 25. .

Dagar, V., Hutchison, W., Muscat, A., Krishnan, A., Hoke, D., Buckle, A., Siswara, P., Amor, D.J., Mann, J., and Pinner, J. (2018). Genetic variation affecting DNA methylation and the human imprinting disorder, Beckwith-Wiedemann syndrome. *Clinical Epigenetics* 10, 1–13. .

Daino, K., Nishimura, M., Imaoka, T., Takabatake, M., Morioka, T., Nishimura, Y., Shimada, Y., and Kakinuma, S. (2018). Epigenetic dysregulation of key developmental genes in radiation-induced rat mammary carcinomas. *International Journal of Cancer* 143, 343–354. .

Dávalos-Salas, M., Furlan-Magaril, M., González-Buendía, E., Valdes-Quezada, C., Ayala-Ortega, E., and Recillas-Targa, F. (2011). Gain of DNA methylation is enhanced in the absence of CTCF at the human retinoblastoma gene promoter. *BMC Cancer* 11, 1–11. .

Davidson, A.J., Lewis, P., Przepiorski, A., and Sander, V. (2019a). Turning mesoderm into kidney. In *Seminars in Cell & Developmental Biology*, (Elsevier), pp. 86–93.

Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G., and Peters, J.-M. (2019b). DNA loop extrusion by human cohesin. *Science* 366, 1338–1345. .

De Almeida, S.F., Grosso, A.R., Koch, F., Fenouil, R., Carvalho, S., Andrade, J., Levezinho, H., Gut, M., Eick, D., and Gut, I. (2011). Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nature Structural & Molecular Biology* 18, 977–983. .

DeBaun, M.R., and Tucker, M.A. (1998). Risk of cancer during the first four years of life in children from The Beckwith-Wiedemann Syndrome Registry. *The Journal of Pediatrics* 132, 398–400. .

DeBaun, M.R., Siegel, M.J., and Choyke, P.L. (1998). Nephromegaly in infancy and early childhood: a risk factor for Wilms tumor in Beckwith-Wiedemann syndrome. *The Journal of Pediatrics* *132*, 401–404. .

DeBaun, M.R., Niemitz, E.L., McNeil, D.E., Brandenburg, S.A., Lee, M.P., and Feinberg, A.P. (2002). Epigenetic alterations of H19 and LIT1 distinguish patients with Beckwith-Wiedemann syndrome with cancer and birth defects. *The American Journal of Human Genetics* *70*, 604–611. .

DeBaun, M.R., Niemitz, E.L., and Feinberg, A.P. (2003). Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. *American Journal of Human Genetics* *72*, 156–160. <https://doi.org/10.1086/346031>.

DeChiara, T.M., Efstratiadis, A., and Robertsen, E.J. (1990). A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature* *345*, 78. .

DeChiara, T.M., Robertson, E.J., and Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* *64*, 849–859. .

Dekker, J., and Heard, E. (2015). Structural and functional diversity of topologically associating domains. *FEBS Letters* *589*, 2877–2884. .

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science (New York, N.Y.)* *295*, 1306–1311. .

Della Rosa, M., and Spivakov, M. (2020). Silencers in the spotlight. *Nature Genetics* *52*, 244–245. .

Denley, A., Cosgrove, L.J., Booker, G.W., Wallace, J.C., and Forbes, B.E. (2005). Molecular interactions of the IGF system. *Cytokine & Growth Factor Reviews* *16*, 421–439. .

Desmet, K.J., Van Hoeck, V., Gagné, D., Fournier, E., Thakur, A., O’doherly, A.M., Walsh, C.P., Sirard, M.A., Bols, P.E.J., and Leroy, J. (2016). Exposure of bovine oocytes and embryos to elevated non-esterified fatty acid concentrations: integration of epigenetic and transcriptomic signatures in resultant blastocysts. *BMC Genomics* *17*, 1004. .

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *Journal of Biological Chemistry* *285*, 26114–26120. .

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* *485*, 376–380. .

Dobbs, K.B., Khan, F.A., Sakatani, M., Moss, J.I., Ozawa, M., Ealy, A.D., and Hansen, P.J. (2013).

Regulation of pluripotency of inner cell mass and growth and differentiation of trophectoderm of the bovine embryo by colony stimulating factor 2. *Biology of Reproduction* *89*, 141, 1–10. .

Doherty, A.S., Mann, M.R., Tremblay, K.D., Bartolomei, M.S., and Schultz, R.M. (2000). Differential effects of culture on imprinted H19 expression in the preimplantation mouse embryo. *Biology of Reproduction* *62*, 1526–1535. .

Doornbos, M.E., Maas, S.M., McDonnell, J., Vermeiden, J.P., and Hennekam, R.C. (2007). Infertility, assisted reproduction technologies and imprinting disturbances: a Dutch study. *Hum Reprod* *22*, 2476–2480. <https://doi.org/10.1093/humrep/dem172>.

Dopie, J., Skarp, K.-P., Rajakylä, E.K., Tanhuanpää, K., and Vartiainen, M.K. (2012). Active maintenance of nuclear actin by importin 9 supports transcription. *Proceedings of the National Academy of Sciences* *109*, E544–E552. .

Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., and Nusbaum, C. (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Research* *16*, 1299–1309. .

Doyle, B., Fudenberg, G., Imakaev, M., and Mirny, L.A. (2014). Chromatin loops as allosteric modulators of enhancer-promoter interactions. *PLoS Computational Biology* *10*. .

Driver, A.M., Peñagaricano, F., Huang, W., Ahmad, K.R., Hackbart, K.S., Wiltbank, M.C., and Khatib, H. (2012). RNA-Seq analysis uncovers transcriptomic variations between morphologically similar in vivo- and in vitro-derived bovine blastocysts. *BMC Genomics* *13*, 118. .

Du, W., Dong, Q., Zhang, Z., Liu, B., Zhou, T., Xu, R., Wang, H., Zhu, B., and Li, Y. (2019). Stella protein facilitates DNA demethylation by disrupting the chromatin association of the RING finger-type E3 ubiquitin ligase UHRF1. *Journal of Biological Chemistry* *294*, 8907–8917. .

Du, Z., Zheng, H., Huang, B., Ma, R., Wu, J., Zhang, X., He, J., Xiang, Y., Wang, Q., and Li, Y. (2017). Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* *547*, 232–235. .

Dundr, M., Ospina, J.K., Sung, M.-H., John, S., Upender, M., Ried, T., Hager, G.L., and Matera, A.G. (2007). Actin-dependent intranuclear repositioning of an active gene locus in vivo. *The Journal of Cell Biology* *179*, 1095–1103. .

Dupin, I., Sakamoto, Y., and Etienne-Manneville, S. (2011). Cytoplasmic intermediate filaments mediate actin-driven positioning of the nucleus. *Journal of Cell Science* *124*, 865–872. .

Edwards, R.G. (1965). Maturation in vitro of mouse, sheep, cow, pig, rhesus monkey and human ovarian oocytes. *Nature* *208*, 349. .

Edwards, R.G., Steptoe, P.C., and Purdy, J.M. (1970). Fertilization and cleavage in vitro of preovulator human oocytes. *Nature* 227, 1307. .

Eggenchwiler, J., Ludwig, T., Fisher, P., Leighton, P.A., Tilghman, S.M., and Efstratiadis, A. (1997). Mouse mutant embryos overexpressing IGF-II exhibit phenotypic features of the Beckwith–Wiedemann and Simpson–Golabi–Behmel syndromes. *Genes & Development* 11, 3128–3142. .

Elhamamsy, A.R. (2017). Role of DNA methylation in imprinting disorders: an updated review. *Journal of Assisted Reproduction and Genetics* 34, 549–562. .

Elliott, M., and Maher, E.R. (1994). Beckwith-Wiedemann syndrome. *Journal of Medical Genetics* 31, 560. .

Elliott, M., Bayly, R., Cole, T., Temple, I.K., and Maher, E.R. (1994). Clinical features and natural history of Beckwith-Wiedemann syndrome: presentation of 74 new cases. *Clinical Genetics* 46, 168–174. .

Engström, W., Lindham, S., and Schofield, P. (1988). Wiedemann-Beckwith syndrome. *European Journal of Pediatrics* 147, 450–457. .

Erdel, F., and Rippe, K. (2018). Formation of chromatin subcompartments by phase separation. *Biophysical Journal* 114, 2262–2270. .

Fair, T., Carter, F., Park, S., Evans, A.C.O., and Lonergan, P. (2007). Global gene expression analysis during bovine oocyte in vitro maturation. *Theriogenology* 68, S91–S97. .

Farin, P.W., and Farin, C.E. (1995). Transfer of bovine embryos produced in vivo or in vitro: survival and fetal development. *Biology of Reproduction* 52, 676–682. .

Farin, P.W., Crosier, A.E., and Farin, C.E. (2001). Influence of in vitro systems on embryo survival and fetal development in cattle. *Theriogenology* 55, 151–170. .

Farin, P.W., Piedrahita, J.A., and Farin, C.E. (2006). Errors in development of fetuses and placentas from in vitro-produced bovine embryos. *Theriogenology* 65, 178–191. <https://doi.org/10.1016/j.theriogenology.2005.09.022>.

Fatemi, M., Hermann, A., Gowher, H., and Jeltsch, A. (2002). Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *European Journal of Biochemistry* 269, 4981–4984. .

Fauser, B.C., Devroey, P., Diedrich, K., Balaban, B., Bonduelle, M., Delemarre-van de Waal, H.A., Estella, C., Ezcurra, D., Geraedts, J.P., Howles, C.M., et al. (2014). Health outcomes of children born after IVF/ICSI: a review of current expert opinion and literature. *Reprod Biomed Online* 28, 162–182. <https://doi.org/10.1016/j.rbmo.2013.10.013>.

Fedoriw, A.M., Stein, P., Svoboda, P., Schultz, R.M., and Bartolomei, M.S. (2004). Transgenic RNAi

reveals essential function for CTCF in H19 gene imprinting. *Science* *303*, 238–240. .

Feng, L., Sun, X., Csizmadia, E., Han, L., Bian, S., Murakami, T., Wang, X., Robson, S.C., and Wu, Y. (2011). Vascular CD39/ENTPD1 directly promotes tumor cell growth by scavenging extracellular adenosine triphosphate. *Neoplasia* *13*, 206-IN2. .

Ferrari, K.J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stützer, A., Fischle, W., Bonaldi, T., and Pasini, D. (2014). Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Molecular Cell* *53*, 49–62. .

Filippi, G., and Mckusick, V.A. (1970). THE BECKWITH-WIEDEMANN SYNDROME: THE EXOMPHALOS-MACROGLOSSIA-GIGANTISM SYNDROME: REPORT OF TWO CASES AND REVIEW OF THE LITERATURE. *Medicine* *49*, 279–298. .

Foote, R.H. (1981). Factors affecting gestation length in dairy cattle. *Theriogenology* *15*, 553–559. .

Fornes, O., Castro-Mondragon, J.A., Khan, A., Van der Lee, R., Zhang, X., Richmond, P.A., Modi, B.P., Correard, S., Gheorghe, M., and Baranašić, D. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research* *48*, D87–D92. .

Fraser, J., Ferrai, C., Chiariello, A.M., Schueler, M., Rito, T., Laudanno, G., Barbieri, M., Moore, B.L., Kraemer, D.C., and Aitken, S. (2015). Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Molecular Systems Biology* *11*. .

Frémont, S., Hammich, H., Bai, J., Wioland, H., Klinkert, K., Rocancourt, M., Kikuti, C., Stroebel, D., Romet-Lemonne, G., and Pylypenko, O. (2017). Oxidation of F-actin controls the terminal steps of cytokinesis. *Nature Communications* *8*, 1–16. .

Fujimoto, L.M., Roth, R., Heuser, J.E., and Schmid, S.L. (2000). Actin assembly plays a variable, but not obligatory role in receptor-mediated endocytosis. *Traffic* *1*, 161–171. .

Funaki, S., Nakamura, T., Nakatani, T., Umehara, H., Nakashima, H., and Nakano, T. (2014). Inhibition of maintenance DNA methylation by Stella. *Biochemical and Biophysical Research Communications* *453*, 455–460. .

Gandolfi, F., and Moor, R.M. (1987). Stimulation of early embryonic development in the sheep by co-culture with oviduct epithelial cells. *Journal of Reproduction and Fertility* *81*, 23–28. .

Ganguly, A., Yang, H., and Cabral, F. (2011). Class III β -tubulin counteracts the ability of paclitaxel to inhibit cell migration. *Oncotarget* *2*, 368. .

Gao, G., Wang, S., Zhang, J., Su, G., Zheng, Z., Bai, C., Yang, L., Wei, Z., Wang, X., Liu, X., et al. (2019). Transcriptome-wide analysis of the SCNT bovine abnormal placenta during mid-to late gestation. *Scientific Reports* *9*, 1–10. .

Gao, S., Zou, D., Mao, L., Liu, H., Song, P., Chen, Y., Zhao, S., Gao, C., Li, X., and Gao, Z. (2015). BS-SNPer: SNP calling in bisulfite-seq data. *Bioinformatics* 31, 4006–4008. .

Gasser, S.M. (2002). Visualizing chromatin dynamics in interphase nuclei. *Science* 296, 1412–1416. .

Gaston, V., Le Bouc, Y., Soupre, V., Burglen, L., Donadieu, J., Oro, H., Audry, G., Vazquez, M.-P., and Gicquel, C. (2001). Analysis of the methylation status of the KCNQ10T and H19 genes in leukocyte DNA for the diagnosis and prognosis of Beckwith–Wiedemann syndrome. *European Journal of Human Genetics* 9, 409. .

Gdula, M.R., Nesterova, T.B., Pintacuda, G., Godwin, J., Zhan, Y., Ozadam, H., McClellan, M., Moralli, D., Krueger, F., and Green, C.M. (2019). The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome. *Nature Communications* 10, 1–14. .

Gheldof, N., Leleu, M., Noordermeer, D., Rougemont, J., and Reymond, A. (2012). Detecting long-range chromatin interactions using the chromosome conformation capture sequencing (4C-seq) method. In *Gene Regulatory Networks*, (Springer), pp. 211–225.

Gianaroli, L. (2000). Preimplantation genetic diagnosis: polar body and embryo biopsy. *Hum. Reprod* 15. .

Gicquel, C., Gaston, V., Mandelbaum, J., Siffroi, J.P., Flahault, A., and Le Bouc, Y. (2003). In vitro fertilization may increase the risk of Beckwith-Wiedemann syndrome related to the abnormal imprinting of the KCN10T gene. *American Journal of Human Genetics* 72, 1338–1341. .

Giorgetti, L., Galupa, R., Nora, E.P., Piolot, T., Lam, F., Dekker, J., Tiana, G., and Heard, E. (2014). Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* 157, 950–963. .

Goldkamp, A.K., Li, Y., Rivera, R.M., and Hagen, D.E. (2022). Characterization of tRNA expression profiles in large offspring syndrome. *BMC Genomics* 23, 1–16. .

Goldman, M., Smith, A., Shuman, C., Caluseriu, O., Wei, C., Steele, L., Ray, P., Sadowski, P., Squire, J., and Weksberg, R. (2002). Renal abnormalities in beckwith-wiedemann syndrome are associated with 11p15. 5 uniparental disomy. *Journal of the American Society of Nephrology* 13, 2077–2084. .

Gong, T., Gu, X., Liu, Y.-T., Zhou, Z., Zhang, L.-L., Wen, Y., Zhong, W.-L., Xu, G.-L., and Zhou, J.-Q. (2020). Both combinatorial K4me0-K36me3 marks on sister histone H3s of a nucleosome are required for Dnmt3a-Dnmt3L mediated de novo DNA methylation. *Journal of Genetics and Genomics* 47, 105–114. .

Goudarzi, A., Zhang, D., Huang, H., Barral, S., Kwon, O.K., Qi, S., Tang, Z., Buchou, T., Vitte, A.-L., and He, T. (2016). Dynamic competing histone H4 K5K8 acetylation and butyrylation are hallmarks of highly active gene promoters. *Molecular Cell* 62, 169–180. .

Grandjean, V., Smith, J., Schofield, P.N., and Ferguson-Smith, A.C. (2000). Increased IGF-II protein affects p57kip2 expression in vivo and in vitro: implications for Beckwith–Wiedemann syndrome. *Proceedings of the National Academy of Sciences* *97*, 5279–5284. .

Graves, T.K., Hansel, W., and Krook, L. (1991). Prolonged gestation in a Holstein cow: adenohipophyseal aplasia and skeletal pathology in the offspring. *The Cornell Veterinarian* *81*, 277–294. .

Grewal, S.I., and Elgin, S.C. (2002). Heterochromatin: new possibilities for the inheritance of structure. *Current Opinion in Genetics & Development* *12*, 178–187. .

Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., and Li, G. (2015). Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* *517*, 640–644. .

Haggerty, C., Kretzmer, H., Riemenschneider, C., Kumar, A.S., Mattei, A.L., Bailly, N., Gottfreund, J., Giesselmann, P., Weigert, R., and Brändl, B. (2021). Dnmt1 has de novo activity targeted to transposable elements. *Nature Structural & Molecular Biology* *28*, 594–603. .

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of Development* *117*, 15–23. .

Han, L., Ren, C., Zhang, J., Shu, W., and Wang, Q. (2019). Differential roles of Stella in the modulation of DNA methylation during oocyte and zygotic development. *Cell Discovery* *5*, 1–4. .

Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W.H., Ye, C., Ping, J.L.H., and Mulawadi, F. (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. *Nature Genetics* *43*, 630. .

Hansen, P.J., Dobbs, K.B., Denicol, A.C., and Siqueira, L.G. (2016). Sex and the preimplantation embryo: implications of sexual dimorphism in the preimplantation period for maternal programming of embryonic development. *Cell and Tissue Research* *363*, 237–247. .

Haraguchi, R., Kitazawa, R., and Kitazawa, S. (2015). Epigenetic regulation of Tbx18 gene expression during endochondral bone formation. *Cell and Tissue Research* *359*, 503–512. .

Hashimoto, H., Liu, Y., Upadhyay, A.K., Chang, Y., Howerton, S.B., Vertino, P.M., Zhang, X., and Cheng, X. (2012). Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Research* *40*, 4841–4849. .

Hasler, J.F. (2014). Forty years of embryo transfer in cattle: A review focusing on the journal *Theriogenology*, the growth of the industry in North America, and personal reminiscences. *Theriogenology* *81*, 152–169. .

Hasler, J.F., Henderson, W.B., Hurtgen, P.J., Jin, Z.Q., McCauley, A.D., Mower, S.A., Neely, B., Shuey, L.S., Stokes, J.E., and Trimmer, S.A. (1995). Production, freezing and transfer of bovine IVF embryos and

subsequent calving results. *Theriogenology* *43*, 141–152. .

Hatada, I., Ohashi, H., Fukushima, Y., Kaneko, Y., Inoue, M., Komoto, Y., Okada, A., Ohishi, S., Nabetani, A., and Morisaki, H. (1996). An imprinted gene p57KIP2 is mutated in Beckwith–Wiedemann syndrome. *Nature Genetics* *14*, 171. .

Hayes, B.J., and Daetwyler, H.D. (2019). 1000 bull genomes project to map simple and complex genetic traits in cattle: applications and outcomes. *Annual Review of Animal Biosciences* *7*, 89–102. .

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., and Li, L. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* *333*, 1303–1307. .

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., and Ching, C.W. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* *459*, 108–112. .

Hemberger, M., Dean, W., and Reik, W. (2009). Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington’s canal. *Nature Reviews Molecular Cell Biology* *10*, 526–537. .

Henry, I., Bonaiti-Pellie, C., Chehensse, V., Beldjord, C., Schwartz, C., Utermann, G., and Junien, C. (1991). Uniparental paternal disomy in a genetic cancer-predisposing syndrome. *Nature* *351*, 665. .

Hildebrand, E.M., and Dekker, J. (2020). Mechanisms and Functions of Chromosome Compartmentalization. *Trends in Biochemical Sciences*.

Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R., and Sasaki, H. (2008). Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes & Development* *22*, 1607–1616. .

Hofmann, W.A., Stojiljkovic, L., Fuchsova, B., Vargas, G.M., Mavrommatis, E., Philimonenko, V., Kysela, K., Goodrich, J.A., Lessard, J.L., and Hope, T.J. (2004). Actin is part of pre-initiation complexes and is necessary for transcription by RNA polymerase II. *Nature Cell Biology* *6*, 1094–1101. .

Hohmann, T., and Dehghani, F. (2019). The cytoskeleton—a complex interacting meshwork. *Cells* *8*, 362. .

Holland, M.D., and Odde, K.G. (1992). Factors affecting calf birth weight: a review. *Theriogenology* *38*, 769–798. .

Holm, P., Walker, S.K., and Seamark, R.F. (1996). Embryo viability, duration of gestation and birth weight in sheep after transfer of in vitro matured and in vitro fertilized zygotes cultured in vitro or in vivo. *Journal of Reproduction and Fertility* *107*, 175–181. .

Hori, N., Nagai, M., Hirayama, M., Hirai, T., Matsuda, K., Hayashi, M., Tanaka, T., Ozawa, T., and Horike,

S. (2010). Aberrant CpG methylation of the imprinting control region KvDMR1 detected in assisted reproductive technology-produced calves and pathogenesis of large offspring syndrome. *Animal Reproduction Science* *122*, 303–312. .

Horike, S., Mitsuya, K., Meguro, M., Kotobuki, N., Kashiwagi, A., Notsu, T., Schulz, T.C., Shirayoshi, Y., and Oshimura, M. (2000). Targeted disruption of the human LIT1 locus defines a putative imprinting control element playing an essential role in Beckwith–Wiedemann syndrome. *Human Molecular Genetics* *9*, 2075–2083. .

Hou, C., Li, L., Qin, Z.S., and Corces, V.G. (2012). Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Molecular Cell* *48*, 471–484. .

Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., and Gong, W. (2015). Structural insight into substrate preference for TET-mediated oxidation. *Nature* *527*, 118–122. .

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* *4*, 44–57. .

Hubertus, J., Zitzmann, F., Trippel, F., Müller-Höcker, J., Stehr, M., von Schweinitz, D., and Kappler, R. (2013). Selective methylation of CpGs at regulatory binding sites controls NNAT expression in Wilms tumors. *PLoS One* *8*, e67605. .

Ibrahim, A., Kirby, G., Hardy, C., Dias, R.P., Tee, L., Lim, D., Berg, J., MacDonald, F., Nightingale, P., and Maher, E.R. (2014). Methylation analysis and diagnostics of Beckwith-Wiedemann syndrome in 1,000 subjects. *Clinical Epigenetics* *6*, 11. .

Illi, B., Scopece, A., Nanni, S., Farsetti, A., Morgante, L., Biglioli, P., Capogrossi, M.C., and Gaetano, C. (2005). Epigenetic histone modification and cardiovascular lineage programming in mouse embryonic stem cells exposed to laminar shear stress. *Circulation Research* *96*, 501–508. .

Inbar-Feigenberg, M., Choufani, S., Butcher, D.T., Roifman, M., and Weksberg, R. (2013). Basic concepts of epigenetics. *Fertility and Sterility* *99*, 607–615. <https://doi.org/10.1016/j.fertnstert.2013.01.117>.

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., and Webster, M. (2009). Genome-wide methylation analysis of human colon cancer reveals similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics* *41*, 178. .

Ishiyama, S., Nishiyama, A., Saeki, Y., Moritsugu, K., Morimoto, D., Yamaguchi, L., Arai, N., Matsumura, R., Kawakami, T., and Mishima, Y. (2017). Structure of the Dnmt1 reader module complexed with a unique two-mono-ubiquitin mark on histone H3 reveals the basis for DNA methylation maintenance. *Molecular Cell* *68*, 350–360. e7. .

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* *333*, 1300–1303. .

- Jackson, O.J., and Moglen, L.H. (1970). Umbilical hernia—a retrospective study. *California Medicine* 113, 8. .
- Jacobsen, H., Schmidt, M., Holm, P., Sangild, P.T., Vajta, G., Greve, T., and Callesen, H. (2000). Body dimensions and birth and organ weights of calves derived from in vitro produced embryos cultured with or without serum and oviduct epithelium cells. *Theriogenology* 53, 1761–1769. .
- James, M.A., Lu, Y., Liu, Y., Vikis, H.G., and You, M. (2009). RGS17, an overexpressed gene in human lung and prostate cancer, induces tumor cell proliferation through the cyclic AMP-PKA-CREB pathway. *Cancer Research* 69, 2108–2116. .
- Jamieson, K., Wiles, E.T., McNaught, K.J., Sidoli, S., Leggett, N., Shao, Y., Garcia, B.A., and Selker, E.U. (2016). Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. *Genome Research* 26, 97–107. .
- Jeon, H., Bernstein, L.J., Belkin, D.A., Ghalili, S., and Geronemus, R.G. (2019). Pulsed dye laser treatment of port-wine stains in infancy without the need for general anesthesia. *JAMA Dermatology* 155, 435–441. .
- Ji, T. (2019). A Bayesian hidden Markov model for detecting differentially methylated regions. *Biometrics* 75, 663–673. .
- Jiménez, C., Portela, R.A., Mellado, M., Rodríguez-Frade, J.M., Collard, J., Serrano, A., Martínez-a, C., Avila, J., and Carrera, A.C. (2000). Role of the PI3K regulatory subunit in the control of actin organization and cell migration. *The Journal of Cell Biology* 151, 249–262. .
- Jiménez Martín, O., Schlosser, A., Furtwängler, R., Wegert, J., and Gessler, M. (2021). MYCN and MAX alterations in Wilms tumor and identification of novel N-MYC interaction partners as biomarker candidates. *Cancer Cell International* 21, 1–15. .
- Jjingo, D., Conley, A.B., Soojin, V.Y., Lunyak, V.V., and Jordan, I.K. (2012). On the presence and role of human gene-body DNA methylation. *Oncotarget* 3, 462. .
- Joao Viana (2021). 2020 Statistics of embryo production and transfer in domestic farm animals. *Embryo Technology Newsletter* 39. .
- Jung, H.-Y., Jung, J.S., Whang, Y.M., and Kim, Y.H. (2013). RASSF1A suppresses cell migration through inactivation of HDAC6 and increase of acetylated α -tubulin. *Cancer Research and Treatment: Official Journal of Korean Cancer Association* 45, 134. .
- Jung, Y.H., Kremisky, I., Gold, H.B., Rowley, M.J., Punyawai, K., Buonanotte, A., Lyu, X., Bixler, B.J., Chan, A.W., and Corces, V.G. (2019). Maintenance of CTCF-and transcription factor-mediated interactions from the gametes to the early mouse embryo. *Molecular Cell* 75, 154-171. e5. .

Kadarmideen, H., Mazzoni, G., Watanabe, Y., Strøbech, L., Baruselli, P., Meirelles, F., Callesen, H., Hyttel, P., Ferraz, J., and Nogueira, M. (2018). Genomic selection of in vitro produced and somatic cell nuclear transfer embryos for rapid genetic improvement in cattle production. *Animal Reproduction (AR)* *12*, 389–396. .

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., and Levine, S.S. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* *467*, 430–435. .

Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., and Chen, L. (2012). Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nature Biotechnology* *30*, 90. .

Kalish, J.M., Biesecker, L.G., Brioude, F., Deardorff, M.A., Cesare-Merlone, D., Druley, T., Ferrero, G.B., Lapunzina, P., Larizza, L., and Maas, S. (2017). Nomenclature and definition in asymmetric regional body overgrowth. *American Journal of Medical Genetics Part A* *173*, 1735–1738. .

Kaltenbach, S., Capri, Y., Rossignol, S., Denjoy, I., Soudée, S., Aboura, A., Baumann, C., and Verloes, A. (2013). Beckwith–Wiedemann syndrome and long QT syndrome due to familial-balanced translocation t (11; 17)(p15. 5; q21. 3) involving the KCNQ1 gene. *Clinical Genetics* *84*, 78–81. .

Kamien, B., Ronan, A., Poke, G., Sinnerbrink, I., Baynam, G., Ward, M., Gibson, W.T., Dudding-Byth, T., and Scott, R.J. (2018). A clinical review of generalized overgrowth syndromes in the era of massively parallel sequencing. *Molecular Syndromology* *9*, 70–82. .

Kampmann, U., Madsen, L.R., Skajaa, G.O., Iversen, D.S., Moeller, N., and Ovesen, P. (2015). Gestational diabetes: A clinical update. *World Journal of Diabetes* *6*, 1065. .

Karbassi, E., Rosa-Garrido, M., Chapski, D.J., Wu, Y., Ren, S., Wang, Y., Stefani, E., and Vondriska, T.M. (2019). Direct visualization of cardiac transcription factories reveals regulatory principles of nuclear architecture during pathological remodeling. *Journal of Molecular and Cellular Cardiology* *128*, 198–211. .

Karki, S., Kennedy, D.E., Mclean, K., Grzybowski, A.T., Maienschein-Cline, M., Banerjee, S., Xu, H., Davis, E., Mandal, M., and Labno, C. (2018). Regulated capture of V κ gene topologically associating domains by transcription factories. *Cell Reports* *24*, 2443–2456. .

Ke, Y., Xu, Y., Chen, X., Feng, S., Liu, Z., Sun, Y., Yao, X., Li, F., Zhu, W., and Gao, L. (2017). 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* *170*, 367–381. e20. .

Kebede, A.F., Nieborak, A., Shahidian, L.Z., Le Gras, S., Richter, F., Gomez, D.A., Baltissen, M.P., Meszaros, G., de Fatima Magliarelli, H., and Taudt, A. (2017). Histone propionylation is a mark of active chromatin.

Nature Structural & Molecular Biology 24, 1048–1056. .

Keeling, M.C., Flores, L.R., Dodhy, A.H., Murray, E.R., and Gavara, N. (2017). Actomyosin and vimentin cytoskeletal networks regulate nuclear shape, mechanics and chromatin organization. *Scientific Reports* 7, 1–14. .

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. *Genome Research* 12, 996–1006. .

Khosla, S., Dean, W., Brown, D., Reik, W., and Feil, R. (2001). Culture of preimplantation mouse embryos affects fetal development and the expression of imprinted genes. *Biology of Reproduction* 64, 918–926. .

Kim, J.-K., Louhghalam, A., Lee, G., Schafer, B.W., Wirtz, D., and Kim, D.-H. (2017). Nuclear lamin A/C harnesses the perinuclear apical actin cables to protect nuclear morphology. *Nature Communications* 8, 1–13. .

Kramer, P.R., Fragoso, G., Pennie, W., Htun, H., Hager, G.L., and Sinden, R.R. (1999). Transcriptional state of the mouse mammary tumor virus promoter can affect topological domain size in vivo. *Journal of Biological Chemistry* 274, 28590–28597. .

Kreis, S., Schönfeld, H.-J., Melchior, C., Steiner, B., and Kieffer, N. (2005). The intermediate filament protein vimentin binds specifically to a recombinant integrin $\alpha2/\beta1$ cytoplasmic tail complex and co-localizes with native $\alpha2/\beta1$ in endothelial cell focal adhesions. *Experimental Cell Research* 305, 110–121. .

Krijger, P.H., Geeven, G., Bianchi, V., Hilvering, C.R., and de Laat, W. (2020). 4C-seq from beginning to end: A detailed protocol for sample preparation and data analysis. *Methods* 170, 17–32. .

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27, 1571–1572. .

Kruip, T.A., and Den Daas, J.H.G. (1997). In vitro produced and cloned embryos: effects on pregnancy, parturition and offspring. *Theriogenology* 47, 43–52. .

Krzyzewska, I.M., Alders, M., Maas, S.M., Bliet, J., Venema, A., Henneman, P., Rezwan, F.I., vd Lip, K., Mul, A.N., and Mackay, D.J.G. (2019). Genome-wide methylation profiling of Beckwith-Wiedemann syndrome patients without molecular confirmation after routine diagnostics. *Clinical Epigenetics* 11, 53. .

Kuentz, P., Bailly, A., Faure, A.C., Blagosklonov, O., Amiot, C., Bresson, J.L., and Roux, C. (2011). Child with Beckwith-Wiedemann syndrome born after assisted reproductive techniques to an human immunodeficiency virus serodiscordant couple. *Fertility and Sterility* 96, e35-8. <https://doi.org/10.1016/j.fertnstert.2011.04.030>.

Lai, A.Y., Fatemi, M., Dhasarathy, A., Malone, C., Sobol, S.E., Geigerman, C., Jaye, D.L., Mav, D., Shah, R., and Li, L. (2010). DNA methylation prevents CTCF-mediated silencing of the oncogene BCL6 in B cell lymphomas. *Journal of Experimental Medicine* 207, 1939–1950. .

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9, 357–359. .

Lapunzina, P. (2005). Risk of tumorigenesis in overgrowth syndromes: a comprehensive review. *Am J Med Genet C Semin Med Genet* 137C, 53–71. <https://doi.org/10.1002/ajmg.c.30064>.

Larkin, J.D., Papantonis, A., Cook, P.R., and Marenduzzo, D. (2013). Space exploration by the promoter of a long human gene during one transcription cycle. *Nucleic Acids Research* 41, 2216–2227. .

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* 547, 236–240. .

Lau, M.M., Stewart, C.E., Liu, Z., Bhatt, H., Rotwein, P., and Stewart, C.L. (1994). Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes & Development* 8, 2953–2963. .

Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral thinking: how histone modifications regulate gene expression. *TRENDS in Genetics* 32, 42–56. .

Lazzari, G., Wrenzycki, C., Herrmann, D., Duchi, R., Kruip, T., Niemann, H., and Galli, C. (2002). Cellular and molecular deviations in bovine in vitro-produced embryos are related to the large offspring syndrome. *Biology of Reproduction* 67, 767–775. .

Le, H.Q., Ghatak, S., Yeung, C.-Y.C., Tellkamp, F., Günschmann, C., Dieterich, C., Yeroslaviz, A., Habermann, B., Pombo, A., and Niessen, C.M. (2016). Mechanical regulation of transcription controls Polycomb-mediated gene silencing during lineage commitment. *Nature Cell Biology* 18, 864–875. .

Lee, C., Scherr, H.M., and Wallingford, J.B. (2007). Shroom family proteins regulate γ -tubulin distribution and microtubule architecture during epithelial cell shape change.

Lee, M.P., DeBaun, M.R., Mitsuya, K., Galonek, H.L., Brandenburg, S., Oshimura, M., and Feinberg, A.P. (1999). Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith–Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proceedings of the National Academy of Sciences* 96, 5203–5208. .

Leighton, P.A., Ingram, R.S., Eggenschwiler, J., Efstratiadis, A., and Tilghman, S.M. (1995). Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature* 375, 34. .

Levsky, J.M., and Singer, R.H. (2003). Fluorescence in situ hybridization: past, present and future.

Journal of Cell Science 116, 2833–2838. .

Lewis, A., and Reik, W. (2006). How imprinting centres work. *Cytogenetic and Genome Research* 113, 81–89. <https://doi.org/10.1159/000090818>.

Li, E., and Zhang, Y. (2014). DNA methylation in mammals. *Cold Spring Harbor Perspectives in Biology* 6, a019133. .

Li, C., Huang, Z., and Gu, L. (2020a). SETD2 reduction adversely affects the development of mouse early embryos. *Journal of Cellular Biochemistry* 121, 797–803. .

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. .

Li, S., Zhu, Y., Ma, C., Qiu, Z., Zhang, X., Kang, Z., Wu, Z., Wang, H., Xu, X., and Zhang, H. (2015). Downregulation of EphA5 by promoter methylation in human prostate cancer. *BMC Cancer* 15, 1–10. .

Li, T., Zhou, X., Wang, X., Zhu, D., and Zhang, Y. (2010). Identification and characterization of human snoRNA core promoters. *Genomics* 96, 50–56. .

Li, T., Wang, L., Du, Y., Xie, S., Yang, X., Lian, F., Zhou, Z., and Qian, C. (2018a). Structural and mechanistic insights into UHRF1-mediated DNMT1 activation in the maintenance DNA methylation. *Nucleic Acids Research* 46, 3218–3231. .

Li, Y., Zhang, Z., Chen, J., Liu, W., Lai, W., Liu, B., Li, X., Liu, L., Xu, S., and Dong, Q. (2018b). Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. *Nature* 564, 136–140. .

Li, Y., Hagen, D.E., Ji, T., Bakhtiarizadeh, M.R., Frederic, W.M., Traxler, E.M., Kalish, J.M., and Rivera, R.M. (2019a). Altered microRNA expression profiles in large offspring syndrome and Beckwith-Wiedemann syndrome. *Epigenetics* 1–27. .

Li, Y., Donnelly, C.G., and Rivera, R.M. (2019b). Overgrowth syndrome. *Veterinary Clinics: Food Animal Practice* 35, 265–276. .

Li, Y., Haarhuis, J.H., Cacciato, Á.S., Oldenkamp, R., van Ruiten, M.S., Willems, L., Teunissen, H., Muir, K.W., de Wit, E., and Rowland, B.D. (2020b). The structural basis for cohesin–CTCF-anchored loops. *Nature* 578, 472–476. .

Li, Y., Boadu, F., Highsmith, M.R., Hagen, D.E., Cheng, J., and Rivera, R.M. (2022). Allele-specific aberration of imprinted domain chromosome architecture associates with large offspring syndrome. *IScience* 25, 104269. <https://doi.org/10.1016/j.isci.2022.104269>.

Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie,

- B.R., Sabo, P.J., and Dorschner, M.O. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (New York, N.Y.)* *326*, 289–293. .
- Lin, L., Li, Q., Zhang, L., Zhao, D., Dai, Y., and Li, N. (2008). Aberrant epigenetic changes and gene expression in cloned cattle dying around birth. *BMC Developmental Biology* *8*, 14. .
- Ling, J.Q., Li, T., Hu, J.F., Vu, T.H., Chen, H.L., Qiu, X.W., Cherry, A.M., and Hoffman, A.R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* *312*, 269–272. .
- Lister, R., Pelizzola, M., Downen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., and Ngo, Q.-M. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* *462*, 315. .
- Liu, J., Wang, Y., Su, J., Luo, Y., Quan, F., and Zhang, Y. (2013). Nuclear donor cell lines considerably influence cloning efficiency and the incidence of large offspring syndrome in bovine somatic cell nuclear transfer. *Reproduction in Domestic Animals* *48*, 660–664. .
- Liu, Y., Siegmund, K.D., Laird, P.W., and Berman, B.P. (2012). Bis-SNP: combined DNA methylation and SNP calling for Bisulfite-seq data. *Genome Biology* *13*, R61. .
- Llères, D., Moindrot, B., Pathak, R., Piras, V., Matelot, M., Pignard, B., Marchand, A., Poncelet, M., Perrin, A., and Tellier, V. (2019). CTCF modulates allele-specific sub-TAD organization and imprinted gene activity at the mouse Dlk1-Dio3 and Igf2-H19 domains. *Genome Biology* *20*, 1–17. .
- Lombardi, M.L., Jaalouk, D.E., Shanahan, C.M., Burke, B., Roux, K.J., and Lammerding, J. (2011). The interaction between nesprins and sun proteins at the nuclear envelope is critical for force transmission between the nucleus and cytoskeleton. *Journal of Biological Chemistry* *286*, 26743–26753. .
- Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. *Cell* *126*, 403–413. .
- Lopes, J.S., Alcázar-Triviño, E., Soriano-Úbeda, C., Hamdi, M., Cánovas, S., Rizos, D., and Coy, P. (2020). Reproductive outcomes and endocrine profile in artificially inseminated versus embryo transferred cows. *Animals* *10*, 1359. .
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* *15*, 550. .
- Lund, U., Agostinelli, C., and Agostinelli, M.C. (2017). Package ‘circular.’ Repository CRAN.
- Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., and Laxova, R. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* *161*, 1012–1025. .

Maeshima, K., Hihara, S., and Eltsov, M. (2010). Chromatin structure: does the 30-nm fibre exist in vivo? *Current Opinion in Cell Biology* 22, 291–297. .

Mannens, M., Hoovers, J., Redeker, E., Verjaal, M., Feinberg, A., Little, P., Boavida, M., Coad, N., Steenman, M., Bliiek, J., et al. (1994). Parental imprinting of human chromosome region 11p15. 3-pter involved in the Beckwith-Wiedemann syndrome and various human neoplasia. *European Journal of Human Genetics* 2, 3–23. .

Marcho, C., Bevilacqua, A., Tremblay, K.D., and Mager, J. (2015). Tissue-specific regulation of Igf2r/Airn imprinting during gastrulation. *Epigenetics & Chromatin* 8, 10. .

Markaki, Y., Gunkel, M., Schermelleh, L., Beichmanis, S., Neumann, J., Heidemann, M., Leonhardt, H., Eick, D., Cremer, C., and Cremer, T. (2010). Functional nuclear organization of transcription and DNA replication a topographical marriage between chromatin domains and the interchromatin compartment. (Cold Spring Harbor Laboratory Press), pp. 475–492.

Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.* 7, 29–59. .

Matveeva, E.A., Venkova, L.S., Chernoiivanenko, I.S., and Minin, A.A. (2015). Vimentin is involved in regulation of mitochondrial motility and membrane potential by Rac1. *Biology Open* 4, 1290–1297. .

Maxfield, E.K., Sinclair, K.D., Dolman, D.F., Staines, M.E., and Maltin, C.A. (1997). In vitro culture of sheep embryos increases weight, primary fiber size and secondary to primary fiber ratio in fetal muscle at day 61 of gestation. *Theriogenology* 1, 376. .

McEvoy, T.G., Sinclair, K.D., Broadbent, P.J., Goodhand, K.L., and Robinson, J.J. (1998). Post-natal growth and development of Simmental calves derived from in vivo or in vitro embryos. *Reproduction, Fertility and Development* 10, 459–464. .

McNiven, M.A., Kim, L., Krueger, E.W., Orth, J.D., Cao, H., and Wong, T.W. (2000). Regulated interactions between dynamin and the actin-binding protein cortactin modulate cell shape. *The Journal of Cell Biology* 151, 187–198. .

Meaburn, K.J., and Misteli, T. (2007). Chromosome territories. *Nature* 445, 379–381. .

Mee, J.F. (2013). Why do so many calves die on modern dairy farms and what can we do about calf welfare in the future? *Animals* 3, 1036–1057. .

Mee, J.F., Sánchez-Miguel, C., and Doherty, M. (2014). Influence of modifiable risk factors on the incidence of stillbirth/perinatal mortality in dairy cattle. *The Veterinary Journal* 199, 19–23. .

Messerschmidt, D.M., Knowles, B.B., and Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes & Development* 28,

812–828. .

Michieletto, D., Orlandini, E., and Marenduzzo, D. (2016). Polymer model with epigenetic recoloring reveals a pathway for the de novo establishment and 3D organization of chromatin domains. *Physical Review X* 6, 041047. .

Mitchell, J.A., and Fraser, P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes & Development* 22, 20–25. .

Mizrachi, Y., Horowitz, E., Farhi, J., Levran, D., Raziel, A., and Weissman, A. (2018). Human chorionic gonadotropin serum levels following ovulation triggering and IVF cycle outcome. *Journal of Assisted Reproduction and Genetics* 35, 891–897. .

Mohammed, F., Trieber, C., Overduin, M., and Chidgey, M. (2020). Molecular mechanism of intermediate filament recognition by plakin proteins. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1867, 118801. .

Moncrieff, M.W., Lacey, K.A., and Malleson, P.N. (1977). Management of prolonged hypoglycaemia in Beckwith's syndrome. *Postgraduate Medical Journal* 53, 159–161. .

Monteagudo-Sánchez, A., Hernandez Mora, J.R., Simon, C., Burton, A., Tenorio, J., Lapunzina, P., Clark, S., Esteller, M., Kelsey, G., and López-Siguero, J.P. (2020). The role of ZFP57 and additional KRAB-zinc finger proteins in the maintenance of human imprinted methylation and multi-locus imprinting disturbances. *Nucleic Acids Research* 48, 11394–11407. .

Montefiori, L.E., Sobreira, D.R., Sakabe, N.J., Aneas, I., Joslin, A.C., Hansen, G.T., Bozek, G., Moskowitz, I.P., McNally, E.M., and Nóbrega, M.A. (2018). A promoter interaction map for cardiovascular disease genetics. *Elife* 7, e35788. .

Moore, J.M., Rabaia, N.A., Smith, L.E., Fagerlie, S., Gurley, K., Loukinov, D., Disteché, C.M., Collins, S.J., Kemp, C.J., and Lobanenkov, V.V. (2012). Loss of maternal CTCF is associated with peri-implantation lethality of Ctfc null embryos. *PLoS One* 7, e34915. .

Moran, S., Arribas, C., and Esteller, M. (2016). Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* 8, 389–399. .

Morison, I.M., Paton, C.J., and Cleverley, S.D. (2001). The imprinted gene and parent-of-origin effect database. *Nucleic Acids Research* 29, 275–276. .

Mulholland, C.B., Nishiyama, A., Ryan, J., Nakamura, R., Yiğit, M., Glück, I.M., Trummer, C., Qin, W., Bartoschek, M.D., and Traube, F.R. (2020). Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive DNA demethylation in mammals. *Nature Communications* 11, 1–24. .

Murthy, K., and Wadsworth, P. (2005). Myosin-II-dependent localization and dynamics of F-actin during

cytokinesis. *Current Biology* 15, 724–731. .

Mussa, A., Peruzzi, L., Chiesa, N., De Crescenzo, A., Russo, S., Melis, D., Tarani, L., Baldassarre, G., Larizza, L., and Riccio, A. (2012). Nephrological findings and genotype–phenotype correlation in Beckwith–Wiedemann syndrome. *Pediatric Nephrology* 27, 397–406. .

Mussa, A., Russo, S., De Crescenzo, A., Chiesa, N., Molinatto, C., Selicorni, A., Richiardi, L., Larizza, L., Silengo, M.C., and Riccio, A. (2013). Prevalence of Beckwith–Wiedemann syndrome in north west of Italy. *American Journal of Medical Genetics Part A* 161, 2481–2486. .

Mussa, A., Russo, S., De Crescenzo, A., Freschi, A., Calzari, L., Maitz, S., Macchiaiolo, M., Molinatto, C., Baldassarre, G., and Mariani, M. (2016). (Epi) genotype–phenotype correlations in Beckwith–Wiedemann syndrome. *European Journal of Human Genetics* 24, 183. .

Mussa, A., Molinatto, C., Cerrato, F., Palumbo, O., Carella, M., Baldassarre, G., Carli, D., Peris, C., Riccio, A., and Ferrero, G.B. (2017). Assisted Reproductive Techniques and Risk of Beckwith-Wiedemann Syndrome. *Pediatrics* 140. <https://doi.org/10.1542/peds.2016-4311>.

Nady, N., Lemak, A., Walker, J.R., Avvakumov, G.V., Kareta, M.S., Achour, M., Xue, S., Duan, S., Allali-Hassani, A., and Zuo, X. (2011). Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *Journal of Biological Chemistry* 286, 24300–24311. .

Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A., and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502, 59–64. .

Nakamura, T., Liu, Y.-J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S., Tachibana, M., Ogura, A., Shinkai, Y., and Nakano, T. (2012). PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* 486, 415–419. .

Nass, N., Walter, S., Jechorek, D., Weissenborn, C., Ignatov, A., Haybaeck, J., Sel, S., and Kalinski, T. (2017). High neuronatin (NNAT) expression is associated with poor outcome in breast cancer. *Virchows Archiv* 471, 23–30. .

National Research Council (2004). *Methods and Mechanisms for Genetic Manipulation of Plants, Animals, and Microorganisms*.

Naveh, N.S.S., Deegan, D.F., Huhn, J., Traxler, E., Lan, Y., Weksberg, R., Ganguly, A., Engel, N., and Kalish, J.M. (2021). The role of CTCF in the organization of the centromeric 11p15 imprinted domain interactome. *Nucleic Acids Research* 49, 6315–6330. .

Ni, K., Dansranjavin, T., Rogenhofer, N., Oeztuerk, N., Deuker, J., Bergmann, M., Schuppe, H.-C., Wagenlehner, F., Weidner, W., and Steger, K. (2016). TET enzymes are successively expressed during human spermatogenesis and their expression level is pivotal for male fertility. *Human Reproduction* 31, 1411–1424. .

Niakan, K.K., Han, J., Pedersen, R.A., Simon, C., and Pera, R.A.R. (2012). Human pre-implantation embryo development. *Development* 139, 829–841. .

Nishiyama, A., Yamaguchi, L., Sharif, J., Johmura, Y., Kawamura, T., Nakanishi, K., Shimamura, S., Arita, K., Kodama, T., and Ishikawa, F. (2013). Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* 502, 249–253. .

Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* 169, 930-944. e22. .

Oakes, C.C., La Salle, S., Smiraglia, D.J., Robaire, B., and Trasler, J.M. (2007). Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. *Developmental Biology* 307, 368–379. .

Obrdlik, A., and Percipalle, P. (2011). The F-actin severing protein cofilin-1 is required for RNA polymerase II transcription elongation. *Nucleus* 2, 72–79. .

O’Doherty, A.M., McGettigan, P., Irwin, R.E., Magee, D.A., Gagne, D., Fournier, E., Al-Naib, A., Sirard, M.-A., Walsh, C.P., and Robert, C. (2018). Intragenic sequences in the trophectoderm harbour the greatest proportion of methylation errors in day 17 bovine conceptuses generated using assisted reproductive technologies. *BMC Genomics* 19, 438. .

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247–257. .

Okano, Y., Osasa, Y., Yamamoto, H., Hase, Y., Tsuruhara, T., and Fujita, H. (1986). An infant with Beckwith-Wiedemann syndrome and chromosomal duplication 11p13→ pter.: Correlation of symptoms between 11p trisomy and Beckwith-Wiedemann syndrome. *Journal of Human Genetics* 31, 365. .

O’Leary, N.A., Wright, M.W., Brister, J.R., Ciufu, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research* 44, D733-45. <https://doi.org/10.1093/nar/gkv1189>.

Ong, C.-T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nature Reviews Genetics* 12, 283–293. .

Ong, C.-T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. *Nature Reviews Genetics* 15, 234–246. .

Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., and Allis, C.D. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of

DNA. *Nature* 448, 714–717. .

Opitz, J.M., Weaver, D.W., and Reynolds, J.F. (1998). The syndromes of Sotos and Weaver: reports and review. *American Journal of Medical Genetics Part A* 79, 294–304. .

Ortega, M.S., Moraes, J.G., Patterson, D.J., Smith, M.F., Behura, S.K., Poock, S., and Spencer, T.E. (2018). Influences of sire conception rate on pregnancy establishment in dairy cattle. *Biology of Reproduction* 99, 1244–1254. .

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., and Reik, W. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. *Nature Genetics* 36, 1065–1071. .

Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., and Fraser, P. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biology* 5. .

Osborne-Majnik, A., Fu, Q., and Lane, R.H. (2013). Epigenetic mechanisms in fetal origins of health and disease. *Clinical Obstetrics and Gynecology* 56, 622. .

Ostlund, C., Folker, E.S., Choi, J.C., Gomes, E.R., Gundersen, G.G., and Worman, H.J. (2009). Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton (LINC) complex proteins. *Journal of Cell Science* 122, 4099–4108. .

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. *EMBO Reports* 10, 1235–1241. .

Otani, J., Kimura, H., Sharif, J., Endo, T.A., Mishima, Y., Kawakami, T., Koseki, H., Shirakawa, M., Suetake, I., and Tajima, S. (2013). Cell cycle-dependent turnover of 5-hydroxymethyl cytosine in mouse embryonic stem cells. *PLoS One* 8, e82961. .

de Pagter-Holthuizen, P., Jansen, M., Van Schaik, F.M.A., Van der Kammen, R., Oosterwijk, C., Van den Brande, J.L., and Sussenbach, J.S. (1987). The human insulin-like growth factor II gene contains two development-specific promoters. *FEBS Letters* 214, 259–264. .

Pang, B., and Snyder, M.P. (2020). Systematic identification of silencers in human cells. *Nature Genetics* 52, 254–263. .

Papantonis, A., and Cook, P.R. (2013). Transcription factories: genome organization and gene regulation. *Chemical Reviews* 113, 8683–8705. .

Papantonis, A., Kohro, T., Baboo, S., Larkin, J.D., Deng, B., Short, P., Tsutsumi, S., Taylor, S., Kanki, Y., and Kobayashi, M. (2012). TNF α signals through specialized factories where responsive coding and miRNA

genes are transcribed. *The EMBO Journal* *31*, 4404–4414. .

Parisis, N., Krasinska, L., Harker, B., Urbach, S., Rossignol, M., Camasses, A., Dewar, J., Morin, N., and Fisher, D. (2017). Initiation of DNA replication requires actin dynamics and formin activity. *The EMBO Journal* *36*, 3212–3231. .

Patel, A.M., Chou, E.L., Findeiss, L., and Kelly, K.M. (2012). The horizon for treating cutaneous vascular lesions. (*Frontline Medical Communications*), pp. 98–104.

Perteau, M., Kim, D., Perteau, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols* *11*, 1650–1667. .

Pettenati, M.J., Haines, J.L., Higgins, R.R., Wappner, R.S., Palmer, C.G., and Weaver, Dd. (1986). Wiedemann-Beckwith syndrome: presentation of clinical and cytogenetic data on 22 new cases and review of the literature. *Human Genetics* *74*, 143–154. .

Peyton, S.R., Ghajar, C.M., Khatiwala, C.B., and Putnam, A.J. (2007). The emergence of ECM mechanics and cytoskeletal tension as important regulators of cell function. *Cell Biochemistry and Biophysics* *47*, 300–320. .

Phanstiel, D.H., Boyle, A.P., Araya, C.L., and Snyder, M.P. (2014). Sushi. R: flexible, quantitative and integrative genomic visualizations for publication-quality multi-panel figures. *Bioinformatics* *30*, 2808–2810. .

Philimonenko, V.V., Zhao, J., Iben, S., Dingová, H., Kyselá, K., Kahle, M., Zentgraf, H., Hofmann, W.A., de Lanerolle, P., and Hozák, P. (2004). Nuclear actin and myosin I are required for RNA polymerase I transcription. *Nature Cell Biology* *6*, 1165–1172. .

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.-T., Hookway, T.A., Guo, C., and Sun, Y. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* *153*, 1281–1295. .

Plank, J.L., and Dean, A. (2014). Enhancer function: mechanistic and genome-wide insights come together. *Molecular Cell* *55*, 5–14. .

Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., and Canfield, T.K. (2014). Topologically associating domains are stable units of replication-timing regulation. *Nature* *515*, 402–405. .

Pradhan, R., Ranade, D., and Sengupta, K. (2018). Emerin modulates spatial organization of chromosome territories in cells on softer matrices. *Nucleic Acids Research* *46*, 5561–5586. .

Qin, W., Wolf, P., Liu, N., Link, S., Smets, M., Mastra, F.L., Forné, I., Pichler, G., Hörl, D., and Fellingner, K. (2015). DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone

ubiquitination. *Cell Research* 25, 911–929. .

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P.V., Grimaldi, G., and Riccio, A. (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Molecular Cell* 44, 361–372. .

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. .

Rahman, M.B., Kamal, M.M., Rijsselaere, T., Vandaele, L., Shamsuddin, M., and Van Soom, A. (2014). Altered chromatin condensation of heat-stressed spermatozoa perturbs the dynamics of DNA methylation reprogramming in the paternal genome after in vitro fertilisation in cattle. *Reproduction, Fertility and Development* 26, 1107–1116. .

Ramani, V., Deng, X., Qiu, R., Gunderson, K.L., Steemers, F.J., Disteche, C.M., Noble, W.S., Duan, Z., and Shendure, J. (2017). Massively multiplex single-cell Hi-C. *Nature Methods* 14, 263–266. .

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., and Lander, E.S. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680. .

Rao, S.S., Huang, S.-C., St Hilaire, B.G., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.-R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., and Bochkov, I.D. (2017). Cohesin loss eliminates all loop domains. *Cell* 171, 305–320. e24. .

Raulin, C., Schroeter, C.A., Weiss, R.A., Keiner, M., and Werner, S. (1999). Treatment of port-wine stains with a noncoherent pulsed light source: a retrospective study. *Archives of Dermatology* 135, 679–683. .

Raviram, R., Rocha, P.P., Müller, C.L., Miraldi, E.R., Badri, S., Fu, Y., Swanzey, E., Proudhon, C., Snetkova, V., and Bonneau, R. (2016). 4C-ker: a method to reproducibly identify genome-wide interactions captured by 4C-Seq experiments. *PLoS Computational Biology* 12, e1004780. .

Rebehmed, J., Revy, P., Faure, G., de Villartay, J.-P., and Callebaut, I. (2014). Expanding the SRI domain family: a common scaffold for binding the phosphorylated C-terminal domain of RNA polymerase II. *FEBS Letters* 588, 4431–4437. .

Reichenbach, H.D., Liebrich, J., Berg, U., and Brem, G. (1992). Pregnancy rates and births after unilateral or bilateral transfer of bovine embryos produced in vitro. *Journal of Reproduction and Fertility* 95, 363–370. .

Reik, W., and Walter, J. (2001). Genomic imprinting: parental influence on the genome. *Nature Reviews. Genetics* 2, 21–32. <https://doi.org/10.1038/35047554>.

Reik, W., Brown, K.W., Schneid, H., Le Bouc, Y., Bickmore, W., and Maher, E.R. (1995). Imprinting mutations in the Beckwith—Wiedemann syndrome suggested by an altered imprinting pattern in the IGF2—H19 domain. *Human Molecular Genetics* 4, 2379–2385. .

Renner, M., Wolf, T., Meyer, H., Hartmann, W., Penzel, R., Ulrich, A., Lehner, B., Hovestadt, V., Czwan, E., and Egerer, G. (2013). Integrative DNA methylation and gene expression analysis in high-grade soft tissue sarcomas. *Genome Biology* 14, 1–26. .

Rivera, R.M., and Hansen, P.J. (2001). Development of cultured bovine embryos after exposure to high temperatures in the physiological range. *Reproduction (Cambridge, England)* 121, 107–115. .

Rivera, R.M., Donnelly, C.G., Patel, B.N., Li, Y., and Soto-Moreno, E.J. (2021). Abnormal Offspring Syndrome. *Bovine Reproduction* 876–895. .

Rizos, D., Lonergan, P., Boland, M.P., Arroyo-Garcia, R., Pintado, B., Fuente, J. de la, and Gutierrez-Adan, A. (2002). Analysis of differential messenger RNA expression between bovine blastocysts produced in different culture systems: implications for blastocyst quality. *Biology of Reproduction* 66, 589–595. .

Rizos, D., Gutierrez-Adan, A., Perez-Garnelo, S., De La Fuente, J., Boland, M.P., and Lonergan, P. (2003). Bovine embryo culture in the presence or absence of serum: implications for blastocyst development, cryotolerance, and messenger RNA expression. *Biology of Reproduction* 68, 236–243. .

Rizos, D., Fair, T., Moreira, P.N., Pintado, B., Boland, M.P., and Lonergan, P. (2004). Effect of speed of development on mRNA expression pattern in early bovine embryos cultured in vivo or in vitro. *Molecular Reproduction and Development* 68, 441–448. .

Roberts, S.J. (1971). *Veterinary obstetrics and genital diseases (theriogenology)* (Edition 2). Published by the Author and Distributed by Edwards Brothers Inc., Ann Arbor, Michigan.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nature Biotechnology* 29, 24–26. .

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* 26, 139–140. .

Robson, M.I., Jose, I., Czapiewski, R., Sivakumar, A., Kerr, A.R., and Schirmer, E.C. (2017). Constrained release of lamina-associated enhancers and genes from the nuclear envelope during T-cell activation facilitates their association in chromosome compartments. *Genome Research* 27, 1126–1138. .

Rodriguez, C., Borgel, J., Court, F., Cathala, G., Forné, T., and Piette, J. (2010). CTCF is a DNA methylation-sensitive positive regulator of the INK/ARF locus. *Biochemical and Biophysical Research Communications* 392, 129–134. .

Rosen, B.D., Bickhart, D.M., Schnabel, R.D., Koren, S., Elsik, C.G., Tseng, E., Rowan, T.N., Low, W.Y., Zimin,

- A., and Couldrey, C. (2020). De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* *9*, giaa021. .
- Ross, P.J., and Sampaio, R.V. (2018). Epigenetic remodeling in preimplantation embryos: cows are not big mice. *Animal Reproduction (AR)* *15*, 204–214. .
- Rossignol, S., Steunou, V., Chalas, C., Kerjean, A., Rigolet, M., Viegas-Pequignot, E., Jouannet, P., Le Bouc, Y., and Gicquel, C. (2006). The epigenetic imprinting defect of patients with Beckwith—Wiedemann syndrome born after assisted reproductive technology is not restricted to the 11p15 region. *Journal of Medical Genetics* *43*, 902–907. .
- Rothbart, S.B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A.I., Barsyte-Lovejoy, D., Martinez, J.Y., Bedford, M.T., and Fuchs, S.M. (2012). Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nature Structural & Molecular Biology* *19*, 1155–1160. .
- Rothbart, S.B., Dickson, B.M., Ong, M.S., Krajewski, K., Houliston, S., Kireev, D.B., Arrowsmith, C.H., and Strahl, B.D. (2013). Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes & Development* *27*, 1288–1298. .
- Rouzier, C., Vanatka, R., Bannwarth, S., Philip, N., Coussement, A., Paquis-Flucklinger, V., and Lambert, J.-C. (2006). A novel homozygous MMP2 mutation in a family with Winchester syndrome. *Clinical Genetics* *69*, 271–276. .
- Rovina, D., La Vecchia, M., Cortesi, A., Fontana, L., Pesant, M., Maitz, S., Tabano, S., Bodega, B., Miozzo, M., and Sirchia, S.M. (2020). Profound alterations of the chromatin architecture at chromosome 11p15.5 in cells from Beckwith-Wiedemann and Silver-Russell syndromes patients. *Scientific Reports* *10*, 1–19. .
- Rudan, M.V., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., and Hadjur, S. (2015). Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Reports* *10*, 1297–1309. .
- Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T.C., Robins, A.J., Dalton, S., and Gilbert, D.M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Research* *20*, 761–770. .
- Saitou, M., Barton, S.C., and Surani, M.A. (2002). A molecular programme for the specification of germ cell fate in mice. *Nature* *418*, 293–300. .
- Salilew-Wondim, D., Fournier, E., Hoelker, M., Saeed-Zidane, M., Tholen, E., Looft, C., Neuhoﬀ, C., Besenfelder, U., Havlicek, V., and Rings, F. (2015). Genome-wide DNA methylation patterns of bovine blastocysts developed in vivo from embryos completed different stages of development in vitro. *PLoS*

One 10, e0140467. .

Sanborn, A.L., Rao, S.S., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., and Li, J. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences* 112, E6456–E6465. .

Sangalli, J.R., Chiaratti, M.R., De Bem, T.H.C., de Araujo, R.R., Bressan, F.F., Sampaio, R.V., Perecin, F., Smith, L.C., King, W.A., and Meirelles, F.V. (2014). Development to term of cloned cattle derived from donor cells treated with valproic acid. *PLoS One* 9, e101022. .

Sano, S., Matsubara, K., Nagasaki, K., Kikuchi, T., Nakabayashi, K., Hata, K., Fukami, M., Kagami, M., and Ogata, T. (2016). Beckwith–Wiedemann syndrome and pseudohypoparathyroidism type Ib in a patient with multilocus imprinting disturbance: a female-dominant phenomenon? *Journal of Human Genetics* 61, 765–769. .

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental Biology* 241, 172–182. .

Santos, F., Zakhartchenko, V., Stojkovic, M., Peters, A., Jenuwein, T., Wolf, E., Reik, W., and Dean, W. (2003). Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. *Current Biology* 13, 1116–1121. .

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–411. .

Sato, M., Kimura, T., Kurokawa, K., Fujita, Y., Abe, K., Masuhara, M., Yasunaga, T., Ryo, A., Yamamoto, M., and Nakano, T. (2002). Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. *Mechanisms of Development* 113, 91–94. .

Schertzer, M.D., Bracerros, K.C., Starmer, J., Cherney, R.E., Lee, D.M., Salazar, G., Justice, M., Bischoff, S.R., Cowley, D.O., and Ariel, P. (2019). lncRNA-induced spread of polycomb controlled by genome architecture, RNA abundance, and CpG island DNA. *Molecular Cell* 75, 523-537. e10. .

Schiff, D., Colle, E., Wells, D., and Stern, L. (1973). Metabolic aspects of the Beckwith-Wiedemann syndrome. *The Journal of Pediatrics* 82, 258–262. .

Schmidt, M., Greve, T., Avery, B., Beckers, J.-F., Sulon, J., and Hansen, H.B. (1996). Pregnancies, calves and calf viability after transfer of in vitro produced bovine embryos. *Theriogenology* 46, 527–539. .

Schmutz, S.M. (1986). Deletion of chromosome 11 (p11p13) in a patient with Beckwith-Wiedemann syndrome. *Clinical Genetics* 30, 154–156. .

- Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9, 671–675. .
- Schrank, B.R., Aparicio, T., Li, Y., Chang, W., Chait, B.T., Gundersen, G.G., Gottesman, M.E., and Gautier, J. (2018). Nuclear ARP2/3 drives DNA break clustering for homology-directed repair. *Nature* 559, 61–66. .
- Schuh, M. (2011). An actin-dependent mechanism for long-range vesicle transport. *Nature Cell Biology* 13, 1431–1436. .
- Schulz, R., McCole, R.B., Woodfine, K., Wood, A.J., Chahal, M., Monk, D., Moore, G.E., and Oakey, R.J. (2008). Transcript- and tissue-specific imprinting of a tumour suppressor gene. *Human Molecular Genetics* 18, 118–127. .
- Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148, 458–472. .
- Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., and Okamura, K. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450, 908–912. .
- Shibata, S., and Ishihara, M. (1949). Studies on hereditary defects in Japanese native cattle. *Japan. J. Zootech. Sci.* 19, 63. .
- Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annual Review of Biochemistry* 81, 65–95. .
- Shuman, C., Beckwith, J.B., and Weksberg, R. (2016). Beckwith-Wiedemann syndrome.
- Sigruener, A., Wolfrum, C., Boettcher, A., Kopf, T., Liebisch, G., Orsó, E., and Schmitz, G. (2017). Lipidomic and metabolic changes in the P4-type ATPase ATP10D deficient C57BL/6J wild type mice upon rescue of ATP10D function. *PLoS One* 12, e0178368. .
- Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., De Wit, E., Van Steensel, B., and De Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nature Genetics* 38, 1348–1354. .
- Sinclair, K.D., Broadbent, P.J., and Dolman, D.F. (1995). In vitro produced embryos as a means of achieving pregnancy and improving productivity in beef cows. *Animal Science* 60, 55–64. .
- Sinclair, K.D., Young, L.E., Wilmut, I., and McEvoy, T.G. (2000). In-utero overgrowth in ruminants following embryo culture: lessons from mice and a warning to men. *Human Reproduction* 15, 68–86. .

Sinden, R.R., and Ussery, D.W. (1992). [18] Analysis of DNA structure in vivo using psoralen photobinding: Measurement of supercoiling, topological domains, and DNA-protein interactions. In *Methods in Enzymology*, (Elsevier), pp. 319–335.

Skiles, W.M., Kester, A., Pryor, J.H., Westhusin, M.E., Golding, M.C., and Long, C.R. (2018). Oxygen-induced alterations in the expression of chromatin modifying enzymes and the transcriptional regulation of imprinted genes. *Gene Expression Patterns* 28, 1–11. .

Sleutels, F., Zwart, R., and Barlow, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 415, 810–813. .

Smallwood, S.A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S.R., and Kelsey, G. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature Genetics* 43, 811–814. .

Smilinich, N.J., Day, C.D., Fitzpatrick, G.V., Caldwell, G.M., Lossie, A.C., Cooper, P.R., Smallwood, A.C., Joyce, J.A., Schofield, P.N., and Reik, W. (1999). A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith–Wiedemann syndrome. *Proceedings of the National Academy of Sciences* 96, 8064–8069. .

Smith, L.C., Therrien, J., Filion, F., Bressan, F., and Meirelles, F.V. (2015). Epigenetic consequences of artificial reproductive technologies to the bovine imprinted genes SNRPN, H19/IGF2, and IGF2R. *Frontiers in Genetics* 6, 58. .

Smith, S.L., Everts, R.E., Sung, L.-Y., Du, F., Page, R.L., Henderson, B., Rodriguez-Zas, S.L., Nedambale, T.L., Renard, J.-P., and Lewin, H.A. (2009). Gene expression profiling of single bovine embryos uncovers significant effects of in vitro maturation, fertilization and culture. *Molecular Reproduction and Development* 76, 38–47. .

Sokolov, E., Iannitti, D.A., Schrum, L.W., and McKillop, I.H. (2011). Altered expression and function of regulator of G-protein signaling-17 (RGS17) in hepatocellular carcinoma. *Cellular Signalling* 23, 1603–1610. .

Solovei, I., Wang, A.S., Thanisch, K., Schmidt, C.S., Krebs, S., Zwerger, M., Cohen, T.V., Devys, D., Foisner, R., and Peichl, L. (2013). LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* 152, 584–598. .

Song, C., Wang, L., Wu, X., Wang, K., Xie, D., Xiao, Q., Li, S., Jiang, K., Liao, L., and Yates, J.R. (2018). PML recruits TET2 to regulate DNA modification and cell proliferation in response to chemotherapeutic agent. *Cancer Research* 78, 2475–2489. .

Song, J., Rechkoblit, O., Bestor, T.H., and Patel, D.J. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science* 331, 1036–1040. .

Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. (2005). Interchromosomal associations between alternatively expressed loci. *Nature* 435, 637–645. .

Splinter, E., de Wit, E., van de Werken, H.J., Klous, P., and de Laat, W. (2012). Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods* 58, 221–230. .

Steenman, M.J., Rainier, S., Dobry, C.J., Grundy, P., Horon, I.L., and Feinberg, A.P. (1994). Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. *Nature Genetics* 7, 433. .

Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F., Leeb, M., Wohlfahrt, K.J., Boucher, W., and O'Shaughnessy-Kirwan, A. (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* 544, 59–64. .

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. *Nature* 547, 241–245. .

Style, C.C., Cruz, S.M., Lau, P.E., Lee, T.C., Wesson, D.E., and Olutoye, O.O. (2018). Surgical Outcomes of Patients with Beckwith-Wiedemann Syndrome. *Journal of Pediatric Surgery*.

Stylianopoulou, F., Herbert, J., Soares, M.B., and Efstratiadis, A. (1988a). Expression of the insulin-like growth factor II gene in the choroid plexus and the leptomeninges of the adult rat central nervous system. *Proceedings of the National Academy of Sciences* 85, 141–145. .

Stylianopoulou, F., Efstratiadis, A., Herbert, J., and Pintar, J. (1988b). Pattern of the insulin-like growth factor II gene expression during rat embryogenesis. *Development* 103, 497–506. .

Su, H., Li, D., Hou, X., Tan, B., Hu, J., Zhang, C., Dai, Y., Li, N., and Li, S. (2011a). Molecular structure of bovine Gtl2 gene and DNA methylation status of Dlk1-Gtl2 imprinted domain in cloned bovines. *Animal Reproduction Science* 127, 23–30. .

Su, J., Wang, Y., Liu, Q., Yang, B., Wu, Y., Luo, Y., Hu, G., and Zhang, Y. (2011b). Aberrant mRNA expression and DNA methylation levels of imprinted genes in cloned transgenic calves that died of large offspring syndrome. *Livestock Science* 141, 24–35. .

Su, J., Wang, Y., Xing, X., Liu, J., and Zhang, Y. (2014). Genome-wide analysis of DNA methylation in bovine placentas. *BMC Genomics* 15, 12. .

Sun, F.-L., Dean, W.L., Kelsey, G., Allen, N.D., and Reik, W. (1997). Transactivation of Igf2 in a mouse model of Beckwith–Wiedemann syndrome. *Nature* 389, 809. .

Sun, X.-J., Wei, J., Wu, X.-Y., Hu, M., Wang, L., Wang, H.-H., Zhang, Q.-H., Chen, S.-J., Huang, Q.-H., and Chen, Z. (2005). Identification and characterization of a novel human histone H3 lysine 36-specific

methyltransferase. *Journal of Biological Chemistry* 280, 35261–35271. .

Sunderam, S., Kissin, D.M., Zhang, Y., Jewett, A., Boulet, S.L., Warner, L., Kroelinger, C.D., and Barfield, W.D. (2020). Assisted reproductive technology surveillance—United States, 2017. *MMWR Surveillance Summaries* 69, 1. .

Swift, J., Ivanovska, I.L., Buxboim, A., Harada, T., Dingal, P.D.P., Pinter, J., Pajeroski, J.D., Spinler, K.R., Shin, J.-W., and Tewari, M. (2013). Nuclear lamin-A scales with tissue stiffness and enhances matrix-directed differentiation. *Science* 341, 1240104. .

Syeda, F., Fagan, R.L., Wean, M., Avvakumov, G.V., Walker, J.R., Xue, S., Dhe-Paganon, S., and Brenner, C. (2011). The replication focus targeting sequence (RFTS) domain is a DNA-competitive inhibitor of Dnmt1. *Journal of Biological Chemistry* 286, 15344–15351. .

Symmons, O., Uslu, V.V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., and Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Research* 24, 390–400. .

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., and Aravind, L. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935. .

Takahashi, N., Coluccio, A., Thorball, C.W., Planet, E., Shi, H., Offner, S., Turelli, P., Imbeault, M., Ferguson-Smith, A.C., and Trono, D. (2019). ZNF445 is a primary regulator of genomic imprinting. *Genes & Development* 33, 49–54. .

Takeshita, K., Suetake, I., Yamashita, E., Suga, M., Narita, H., Nakagawa, A., and Tajima, S. (2011). Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). *Proceedings of the National Academy of Sciences* 108, 9055–9059. .

Tan, G., and Lenhard, B. (2016). TFBSTools: an R/bioconductor package for transcription factor binding site analysis. *Bioinformatics (Oxford, England)* 32, 1555–1556. .

Tan, G., Opitz, L., Schlapbach, R., and Rehrauer, H. (2019). Long fragments achieve lower base quality in Illumina paired-end sequencing. *Scientific Reports* 9, 1–7. .

Tan, O.T., Sherwood, K., and Gilchrest, B.A. (1989). Treatment of children with port-wine stains using the flashlamp-pulsed tunable dye laser. *New England Journal of Medicine* 320, 416–421. .

Tanabe, H., Habermann, F.A., Solovej, I., Cremer, M., and Cremer, T. (2002). Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 504, 37–45. .

- Tate, P.H., and Bird, A.P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. *Current Opinion in Genetics & Development* 3, 226–231. .
- Tee, L., Lim, D.H., Dias, R.P., Baudement, M.-O., Slater, A.A., Kirby, G., Hancocks, T., Stewart, H., Hardy, C., and Macdonald, F. (2013). Epimutation profiling in Beckwith-Wiedemann syndrome: relationship with assisted reproductive technology. *Clinical Epigenetics* 5, 1–10. .
- Tenorio, J., Romanelli, V., Martin-Trujillo, A., Fernández, G.-M., Segovia, M., Perandones, C., Pérez Jurado, L.A., Esteller, M., Fraga, M., and Arias, P. (2016). Clinical and molecular analyses of Beckwith-Wiedemann syndrome: Comparison between spontaneous conception and assisted reproduction techniques. *American Journal of Medical Genetics Part A* 170, 2740–2749. .
- Thompson, J.G., Gardner, D.K., Anne Pugh, P., McMillan, W.H., and Robin Tervit, H. (1995). Lamb birth weight is affected by culture system utilized during in vitro pre-elongation development of ovine embryos. *Biology of Reproduction* 53, 1385–1391. .
- Thompson, J.G., Allen, N.W., McGowan, L.T., Bell, A.C.S., Lambert, M.G., and Tervit, H.R. (1998). Effect of delayed supplementation of fetal calf serum to culture medium on bovine embryo development in vitro and following transfer. *Theriogenology* 49, 1239–1249. .
- Thorburn, M.J., Wright, E.S., Miller, C.G., and Smith-Read, E.H.M. (1970). Exomphalos-macroglossia-gigantism syndrome in Jamaican infants. *American Journal of Diseases of Children* 119, 316–321. .
- Thorvaldsen, J.L., Duran, K.L., and Bartolomei, M.S. (1998). Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes & Development* 12, 3693–3702. .
- Tian, X. (2014). Genomic imprinting in farm animals. *Annu. Rev. Anim. Biosci.* 2, 23–40. .
- Triant, D.A., Le Tourneau, J.J., Diesh, C.M., Unni, D.R., Shamimuzzaman, M., Walsh, A.T., Gardiner, J., Goldkamp, A.K., Li, Y., and Nguyen, H.N. (2020). Using online tools at the Bovine Genome Database to manually annotate genes in the new reference genome. *Animal Genetics* 51, 675–682. .
- Trobaugh-Lotrario, A.D., Venkatramani, R., and Feusner, J.H. (2014). Hepatoblastoma in children with Beckwith-Wiedemann syndrome: does it warrant different treatment? *Journal of Pediatric Hematology/Oncology* 36, 369–373. .
- Tsujimura, T., Klein, F.A., Langenfeld, K., Glaser, J., Huber, W., and Spitz, F. (2015). A discrete transition zone organizes the topological and regulatory autonomy of the adjacent *tfap2c* and *bmp7* genes. *PLoS Genetics* 11. .
- Turleau, C., de Grouchy, J., Chavin-Colin, F., Martelli, H., Voyer, M., and Charlas, R. (1984). Trisomy 11p15 and Beckwith-Wiedemann syndrome. A report of two cases. *Human Genetics* 67, 219–221. .

- Tzima, E., Del Pozo, M.A., Kiosses, W.B., Mohamed, S.A., Li, S., Chien, S., and Schwartz, M.A. (2002). Activation of Rac1 by shear stress in endothelial cells mediates both cytoskeletal reorganization and effects on gene expression. *The EMBO Journal* 21, 6791–6800. .
- Urrego, R., Rodriguez-Osorio, N., and Niemann, H. (2014). Epigenetic disorders and altered gene expression after use of assisted reproductive technologies in domestic cattle. *Epigenetics* 9, 803–815. .
- Van der Auwera, G.A., and O'Connor, B.D. (2020). *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra* (O'Reilly Media).
- Van Steensel, B., and Belmont, A.S. (2017). Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. *Cell* 169, 780–791. .
- Varga, J., Kube, M., Luck, K., and Schick, S. (2021). The BAF chromatin remodeling complexes: structure, function, and synthetic lethality. *Biochemical Society Transactions* 49, 1489–1503. .
- Vasimuddin, M., Misra, S., Li, H., and Aluru, S. (2019). Efficient architecture-aware acceleration of BWA-MEM for multicore systems. In *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, (IEEE), pp. 314–324.
- Vermeiden, J.P., and Bernardus, R.E. (2013). Are imprinting disorders more prevalent after human in vitro fertilization or intracytoplasmic sperm injection? *Fertility and Sterility* 99, 642–651. <https://doi.org/10.1016/j.fertnstert.2013.01.125>.
- Verona, R.I., Mann, M.R., and Bartolomei, M.S. (2003). Genomic imprinting: intricacies of epigenetic regulation in clusters. *Annual Review of Cell and Developmental Biology* 19, 237–259. .
- Vincent, A., Omura, N., Hong, S.-M., Jaffe, A., Eshleman, J., and Goggins, M. (2011). Genome-wide analysis of promoter methylation associated with gene expression profile in pancreatic adenocarcinoma. *Clinical Cancer Research* 17, 4341–4354. .
- Vodyanik, M.A., Yu, J., Zhang, X., Tian, S., Stewart, R., Thomson, J.A., and Slukvin, I.I. (2010). A mesoderm-derived precursor for mesenchymal stem and endothelial cells. *Cell Stem Cell* 7, 718–729. .
- van Wagtenonk-de Leeuw, A.M., Aerts, B.J., and den Daas, J.H. (1998). Abnormal offspring following in vitro production of bovine preimplantation embryos: a field study. *Theriogenology* 49, 883–894. .
- van Wagtenonk-de Leeuw, A.M., Mullaart, E., de Roos, A.P., Merton, J.S., den Daas, J.H., Kemp, B., and de Ruigh, L. (2000). Effects of different reproduction techniques: AI MOET or IVP, on health and welfare of bovine offspring. *Theriogenology* 53, 575–597. .
- Walker, S.K., Hartwich, K.M., and Seamark, R.F. (1996). The production of unusually large offspring following embryo manipulation: concepts and challenges. *Theriogenology* 45, 111–120. .

- Wan, L.-B., Pan, H., Hannenhalli, S., Cheng, Y., Ma, J., Fedoriw, A., Lobanenkov, V., Latham, K.E., Schultz, R.M., and Bartolomei, M.S. (2008). Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. *Development* *135*, 2729–2738. .
- Wang, C.-Y., Jégu, T., Chu, H.-P., Oh, H.J., and Lee, J.T. (2018). SMCHD1 merges chromosome compartments and assists formation of super-structures on the inactive X. *Cell* *174*, 406-421. e25. .
- Waziri, M., Patil, S.R., Hanson, J.W., and Bartley, J.A. (1983). Abnormality of chromosome 11 in patients with features of Beckwith-Wiedemann syndrome. *The Journal of Pediatrics* *102*, 873–876. .
- Webster, K.E., O’Byrne, M.K., Fletcher, S., Crewther, P.E., Aapola, U., Craig, J., Harrison, D.K., Aung, H., Phutikanit, N., and Lyle, R. (2005). Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. *Proceedings of the National Academy of Sciences* *102*, 4068–4073. .
- Wee, G., Koo, D.-B., Song, B.-S., Kim, J.-S., Kang, M.-J., Moon, S.-J., Kang, Y.-K., Lee, K.-K., and Han, Y.-M. (2006). Inheritable histone H4 acetylation of somatic chromatin in cloned embryos. *Journal of Biological Chemistry* *281*, 6048–6057. .
- Weksberg, R., Shen, D.R., Fei, Y.L., Song, Q.L., and Squire, J. (1993). Disruption of insulin-like growth factor 2 imprinting in Beckwith–Wiedemann syndrome. *Nature Genetics* *5*, 143. .
- Weksberg, R., Shuman, C., and Beckwith, J.B. (2010). Beckwith-Wiedemann syndrome. *European Journal of Human Genetics : EJHG* *18*, 8–14. <https://doi.org/10.1038/ejhg.2009.106>.
- van de Werken, H.J., de Vree, P.J., Splinter, E., Holwerda, S.J., Klous, P., de Wit, E., and de Laat, W. (2012). 4C technology: protocols and data analysis. In *Methods in Enzymology*, (Elsevier), pp. 89–112.
- Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics* *3*, 180–185. .
- Wiedemann, H.R. (1964). Familial Malformation Complex with Umbilical Hernia and Macroglossia--a "New Syndrome"? *Journal de Génétique Humaine* *13*, 223. .
- Wiedemann, H.-R., Burgio, G.R., Aldenhoff, P., Kunze, J., Kaufmann, H.J., and Schirg, E. (1983). The proteus syndrome. *European Journal of Pediatrics* *140*, 5–12. .
- Wiehle, L., Thorn, G.J., Raddatz, G., Clarkson, C.T., Rippe, K., Lyko, F., Breiling, A., and Teif, V.B. (2019). DNA (de) methylation in embryonic stem cells controls CTCF-dependent chromatin boundaries. *Genome Research* *29*, 750–761. .
- Wilhelmsen, K., Litjens, S.H., Kuikman, I., Tshimbalanga, N., Janssen, H., van den Bout, I., Raymond, K., and Sonnenberg, A. (2005). Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin. *The Journal of Cell Biology* *171*, 799–810. .
- Willadsen, S.M., Janzen, R.E., McAlister, R.J., Shea, B.F., Hamilton, G., and McDermand, D. (1991). The

viability of late morulae and blastocysts produced by nuclear transplantation in cattle. *Theriogenology* 35, 161–170. .

Williams Jr, R.L., Starmer, J., Mugford, J.W., Calabrese, J.M., Mieczkowski, P., Yee, D., and Magnuson, T. (2014). fourSig: a method for determining chromosomal interactions in 4C-Seq data. *Nucleic Acids Research* 42, e68–e68. .

Williamson, I., Berlivet, S., Eskeland, R., Boyle, S., Illingworth, R.S., Paquette, D., Dostie, J., and Bickmore, W.A. (2014). Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. *Genes & Development* 28, 2778–2791. .

Wilson, J.M., Williams, J.D., Bondioli, K.R., Looney, C.R., Westhusin, M.E., and McCalla, D.F. (1995). Comparison of birth weight and growth characteristics of bovine calves produced by nuclear transfer (cloning), embryo transfer and natural mating. *Animal Reproduction Science* 38, 73–83. .

Wrenzycki, C., Herrmann, D., Carnwath, J.W., and Niemann, H. (1999). Alterations in the relative abundance of gene transcripts in preimplantation bovine embryos cultured in medium supplemented with either serum or PVA. *Molecular Reproduction and Development* 53, 8–18. .

Wrenzycki, C., Herrmann, D., Keskinetepe, L., Martins Jr, A., Sirisathien, S., Brackett, B., and Niemann, H. (2001). Effects of culture system and protein supplementation on mRNA expression in pre-implantation bovine embryos. *Human Reproduction* 16, 893–901. .

Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. *Nature Reviews Genetics* 18, 517–534. .

Xie, J., Ji, T., Ferreira, M.A., Li, Y., Patel, B.N., and Rivera, R.M. (2019). Modeling allele-specific expression at the gene and SNP levels simultaneously by a Bayesian logistic mixed regression model. *BMC Bioinformatics* 20, 530. .

Xie, X., Ma, W., Songyang, Z., Luo, Z., Huang, J., Dai, Z., and Xiong, Y. (2016). CCSI: a database providing chromatin–chromatin spatial interaction information. *Database* 2016. .

Xie, X., Jankauskas, R., Mazari, A.M., Drou, N., and Percipalle, P. (2018). β -actin regulates a heterochromatin landscape essential for optimal induction of neuronal programs during direct reprogramming. *PLoS Genetics* 14, e1007846. .

Xie, Y., Wang, F., Zhong, W., Puscheck, E., Shen, H., and Rappolee, D.A. (2006). Shear stress induces preimplantation embryo death that is delayed by the zona pellucida and associated with stress-activated protein kinase-mediated apoptosis. *Biology of Reproduction* 75, 45–55. .

Xie, Y., Wang, F., Puscheck, E.E., and Rappolee, D.A. (2007). Pipetting causes shear stress and elevation of phosphorylated stress-activated protein kinase/jun kinase in preimplantation embryos. *Molecular Reproduction and Development* 74, 1287–1294. .

Xu, Q., Xiang, Y., Wang, Q., Wang, L., Brind'Amour, J., Bogutz, A.B., Zhang, Y., Zhang, B., Yu, G., and Xia, W. (2019). SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. *Nature Genetics* 51, 844–856. .

Xu, Y.Z., Thuraisingam, T., de Lima Morais, D.A., Rola-Pleszczynski, M., and Radzioch, D. (2010). Nuclear translocation of β -actin is involved in transcriptional regulation during macrophage differentiation of HL-60 cells. *Molecular Biology of the Cell* 21, 811–820. .

Yaffe, E., and Tanay, A. (2011). Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nature Genetics* 43, 1059. .

Yamamoto, K., Sokabe, T., Watabe, T., Miyazono, K., Yamashita, J.K., Obi, S., Ohura, N., Matsushita, A., Kamiya, A., and Ando, J. (2005). Fluid shear stress induces differentiation of Flk-1-positive embryonic stem cells into vascular endothelial cells in vitro. *American Journal of Physiology-Heart and Circulatory Physiology* 288, H1915–H1924. .

Yan, Y., Frisen, J., Lee, M.-H., Massague, J., and Barbacid, M. (1997). Ablation of the CDK inhibitor p57Kip2 results in increased apoptosis and delayed differentiation during mouse development. *Genes & Development* 11, 973–983. .

Yao, J., Duan, L., Fan, M., Yuan, J., and Wu, X. (2007). Overexpression of BLCAP induces S phase arrest and apoptosis independent of p53 and NF- κ B in human tongue carcinoma. *Molecular and Cellular Biochemistry* 297, 81–92. .

Yorikawa, C., Shibata, H., Waguri, S., Hatta, K., Horii, M., Katoh, K., Kobayashi, T., Uchiyama, Y., and Maki, M. (2005). Human CHMP6, a myristoylated ESCRT-III protein, interacts directly with an ESCRT-II component EAP20 and regulates endosomal cargo sorting. *Biochemical Journal* 387, 17–26. .

Young, L.E., Butterwith, S.C., and Wilmut, I. (1996). Increased ovine foetal weight following transient asynchronous embryo transfer is not associated with increased placental weight at day 21 of gestation. *Theriogenology* 1, 231. .

Young, L.E., Sinclair, K.D., and Wilmut, I. (1998). Large offspring syndrome in cattle and sheep. *Rev Reprod* 3, 155–163. .

Young, L.E., Fernandes, K., McEvoy, T.G., Butterwith, S.C., Gutierrez, C.G., Carolan, C., Broadbent, P.J., Robinson, J.J., Wilmut, I., and Sinclair, K.D. (2001). Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. *Nature Genetics* 27, 153. .

Yu, C., Zhang, Y.-L., Pan, W.-W., Li, X.-M., Wang, Z.-W., Ge, Z.-J., Zhou, J.-J., Cang, Y., Tong, C., and Sun, Q.-Y. (2013). CRL4 complex regulates mammalian oocyte survival and reprogramming by activation of TET proteins. *Science* 342, 1518–1521. .

Zaborski, D., Grzesiak, W., Szatkowska, I., Dybus, A., Muszynska, M., and Jedrzejczak, M. (2009). *Factors*

affecting dystocia in cattle. *Reproduction in Domestic Animals* *44*, 540–551. .

Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Research* *21*, 1273–1283. .

Zhang, J., Poh, H.M., Peh, S.Q., Sia, Y.Y., Li, G., Mulawadi, F.H., Goh, Y., Fullwood, M.J., Sung, W.-K., and Ruan, X. (2012). ChIA-PET analysis of transcriptional chromatin interactions. *Methods* *58*, 289–299. .

Zhang, P., Liegeois, N.J., Wong, C., Finegold, M., Hou, H., Thompson, J.C., Silverman, A., Harper, J.W., DePinho, R.A., and Elledge, S.J. (1997). Altered cell differentiation and proliferation in mice lacking p57KIP2 indicates a role in Beckwith–Wiedemann syndrome. *Nature* *387*, 151–158. .

Zhang, S., Chen, X., Wang, F., An, X., Tang, B., Zhang, X., Sun, L., and Li, Z. (2016). Aberrant DNA methylation reprogramming in bovine SCNT preimplantation embryos. *Scientific Reports* *6*, 30345. .

Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., and Fischle, W. (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Research* *38*, 4246–4253. .

Ziebarth, J.D., Bhattacharya, A., and Cui, Y. (2012). CTCFBSDB 2.0: a database for CTCF-binding sites and genome organization. *Nucleic Acids Research* *41*, D188–D194. .

Zuin, J., Dixon, J.R., van der Reijden, M.I., Ye, Z., Kolovos, P., Brouwer, R.W., van de Corput, M.P., van de Werken, H.J., Knoch, T.A., and van IJcken, W.F. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences* *111*, 996–1001. .

Vita

Yahan Li was born on March 26, 1991, in Jinan, Shandong Province, China. He grew up with his father Xueyuan Li and mother Xiuli Guo in the same city. Yahan attended the Shandong Agricultural University in Shandong Province during 2010-2014, and obtained his Bachelor of Agriculture degree in Animal Sciences in June 2014. Yahan joined Dr. Rocío Rivera's laboratory at the University of Missouri in January 2016 to pursue a Master of Science degree with focus on developmental epigenetics. He graduated in July 2018 and continued his Doctor of Philosophy program with Dr. Rivera. Yahan successfully defended his PhD on April 26, 2022. He will remain in the Rivera Laboratory until July of 2022 to complete the third research project of his PhD work. He is currently in conversation with a potential postdoctoral mentor at an academic institution in the field of Epigenetics with the goal of joining their laboratory in August of 2022.