The counterpart congenital overgrowth syndromes Beckwith-Wiedemann Syndrome in human and large offspring syndrome in bovine involve alternations in DNA methylation, transcription, and chromatin configuration.

Yahan Li [1], Ping Xiao [2,a], Frimpong Boadu [3,a], Anna K. Goldkamp [2,a], Snehal Nirgude [4], Jianlin Cheng [3], Darren E. Hagen [2,*], Jennifer M. Kalish [4,5,*], Rocío Melissa Rivera [1,*]

[a] All contributed equally to the work

[1]Division of Animal Sciences, University of Missouri, Columbia, MO, 65211, USA

[2]Department of Animal and Food Sciences, Oklahoma State University, Stillwater, OK 74078, USA

[3]Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO 65211, USA

[4]Division of Human Genetics and Center for Childhood Cancer Research, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

[5]Departments of Pediatrics and Genetics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA

* Correspondence: riverarm@missouri.edu, kalishj@chop.edu, darren.hagen@okstate.edu

Short Title: Molecular analyses of the counterpart overgrowth syndromes BWS and LOS

Abstract

Beckwith-Wiedemann Syndrome (BWS, OMIM #130650) is a congenital epigenetic disorder in humans which affects approximately 1 in 10,340 children. The incidence is likely an underestimation as the condition is usually recognized based on observable phenotypes at birth. BWS children have up to a 28% risk of developing tumors and currently, only 80% of patients can be corroborated molecularly (epimutations/variants). It is unknown how the subtypes of this condition are molecularly similar/dissimilar globally, therefore there is a need to deeply characterize the syndrome at the molecular level. Here we characterize the methylome, transcriptome and chromatin configuration of 18 BWS individuals together with the animal model of the condition, the bovine large offspring syndrome (LOS). Sex specific comparisons are performed for a subset of the BWS patients and LOS. Given that this epigenetic overgrowth syndrome has been characterized as a loss-of-imprinting condition, parental allele-specific comparisons were performed using the bovine animal model. In general, the differentially methylated regions (DMRs) detected in BWS and LOS showed significant enrichment for CTCF binding sites. Altered chromosome compartments in BWS and LOS were positively correlated with gene expression changes, and the promoters of differentially expressed genes showed significant enrichment for DMRs, differential topologically associating domains, and differential A/B compartments in some comparisons of BWS subtypes and LOS. We show shared regions of dysregulation between BWS and LOS, including several HOX gene clusters, and also demonstrate that altered DNA methylation differs between the clinically epigenetically identified BWS patients and those identified as having DNA variants (i.e. *CDKN1C*

microdeletion). Lastly, we highlight additional genes and genomic regions that have the potential to serve as targets for biomarker development to improve current molecular methodologies. In summary, our results suggest that genome-wide alternation of chromosome architecture, which is partially caused by DNA methylation changes, also contribute to the development of BWS and LOS.

Keywords: Beckwith-Wiedemann Syndrome, large offspring syndrome, DNA methylation, chromosome architecture, Hi-C sequencing

Introduction

Beckwith-Wiedemann Syndrome (BWS, OMIM #130650) in humans and its bovine counterpart large/abnormal offspring syndrome (LOS/AOS) are well-known congenital epigenetic disorders [1–3]. The incidence of spontaneous BWS has been reported to be approximately 1 in 10,340 live births but the incidence of spontaneous LOS is currently unknown [4–6]. BWS and LOS share a broad spectrum of clinical features including macrosomia, macroglossia, abdominal wall defects (umbilical hernia/omphalocele), lateralized overgrowth, ear malformation, organomegaly, and placentomegaly [2,7–19]. In addition, BWS is also associated with increased embryonal tumor incidence, hyperinsulinism, and facial nevus simplex [2], and LOS often results in dystocia and demise of offspring and dam due to offspring oversize [20]. BWS and LOS share epigenetic defects in the form of loss-of-imprinting at *IGF2*, *KCNQ1*, *IGF2R*, *PLAGL1*, *PEG3*, and *DLK1* [2,6–8,21–30]. In addition to primary epimutations, subtypes of BWS involve genetic defects which lead to loss of gene function or secondary epimutations that result in loss-of-imprinting [31,32].

The development of the mammalian embryo is a finely regulated process and adverse external insults during this window can have permanent influence on the health of the individual long after the exposure, a phenomenon known as Developmental Origins of Health and Disease [33,34]. Both BWS and LOS have been reported to occur more frequently in offspring conceived using assisted reproductive technologies (ART). Children conceived using ART have over a 10-fold increased risk (1 in 1,126 live births) for developing BWS [35], while in the bovine genetic improvement industry, the incidence of overgrowth was recently reported to be 10% for males and 24% for

82 females [36]. The causal relationship between ART and the increased incidence of
83 these syndromes is considered to be achieved by inducing errors in the epigenome of
84 oocytes and/or embryos since ART manipulations overlap with critical developmental
85 windows for epigenetic reprogramming [37–41].

86 We and others have shown that alterations in epigenetic mechanisms (*i.e.*
87 genomic imprinting, DNA methylation, and locus-specific chromosome architecture) and
88 the resultant altered expression of small/long non-coding RNAs are detectable during
89 gestation and after birth in these syndromes [42–44]. However, whether genome wide
90 alteration of chromosome architecture exist in these conditions is not known, nor is how
91 altered genome organization is involved in the observed global alterations in methylome
92 and transcriptome and vice versa. Our hypothesis is that genome-wide alteration of
93 chromosome architecture occurs in the counterpart overgrowth syndromes BWS and
94 LOS that lead to regional and global gene expression changes and that some of these
95 topological aberrations are shared between the two species. In this study, we
96 systematically analyzed the alterations of DNA methylation, chromosome architecture,
97 and gene expression in fibroblast primary cells from human BWS patients and bovine
98 LOS fetuses and explored the association among these changes.

99

100 **Results and Discussion**

101 For human, five controls (CON_H) and 18 BWS individuals were used for
102 analyses. The BWS group is divided into three subgroups according to molecular
103 subtypes, namely imprinting center 1 gain of methylation (BWS_IC1_GOM, n=3),
104 imprinting center 2 loss of methylation (BWS_IC2_LOM, n=12), and *CDKN1C* variant

105 (BWS_CDKN1C_Mut, n=3). For bovine, the experimental groups were control (CON_B,
106 n=5) and LOS (LOS, n=5). It should be noted that one of the LOS samples, LOS_#1, is
107 considered an outlier according to all the sequencing results, therefore it is excluded
108 from the results presented in the main document. Data including LOS_#1 is available in
109 the supplement (LOS_All group, n=6). For each statistical test, overall (non-sex-specific)
110 and sex-specific comparisons between groups were conducted for human and bovine.
111 The sex of human patient BWS_16_Mut shows conflict between all the sequencing
112 results (female) and hospital record (male), so it was excluded from the sex-specific
113 comparisons. Given that BWS and LOS have been characterized as loss-of-imprinting
114 conditions [2,22], parental allele-specific comparisons were performed using the bovine
115 animal model as this type of analysis is not possible with the human samples. A
116 summary of performed comparisons and numbers of statistically significant differences
117 for all the analyses can be found in Table 1 (index 1-36).

118 **DNA methylation profile in BWS and LOS**

119 Whole genome bisulfite sequencing (WGBS) was conducted for human and
120 bovine fibroblast samples to determine CpG methylation status. Read alignment
121 information can be found in Table S1. Differentially methylated regions (DMR) were
122 determined for all the comparisons in Table 1 and the DMR details can be found in Table
123 S1.

124 Principal component analysis (PCA) clusters the majority of BWS samples while
125 the control samples are more dispersed (Figure S1A), perhaps due to limitations in
126 obtaining age matched healthy human samples. Similarly, hierarchical cluster analysis

127 groups most of the BWS samples irrespective of subtype (Figure 1A). 5,708 DMRs were
128 detected when comparing all the BWS samples together with the control group, and
129 ~66% of these DMRs were hypermethylated (Figure 1A). A greater number of DMRs
130 were detected between BWS and controls when each subtype was analyzed
131 independently, and hypermethylation was observed in 67% of DMRs in IC2 LOM group,
132 30% in IC1 GOM group and 96% in the *CDKN1C* variant group (Figure S2A, B, C). We
133 next compared the methylomes of the three BWS subtypes and identified clear and
134 specific differences between them (Figure S2D, E, F). For bovine samples, 10,860
135 DMRs were identified in LOS with ~75% of those being hypermethylated (Figure 1B), as
136 per the overall BWS analysis.
137
138
139 **Figure 1. CpG methylation heatmap and genomic context enrichment for**
140 **differentially methylated regions (DMR). (A)** and **(B)** DMR for human BWS group
141 compared with control group **(A)**, and bovine LOS group compared with control group
142 **(B)**. The left panel is the heatmap with hierarchical clustering and sex annotation which
143 include male (M), female (F), and conflict (C). The right panels are the distribution of
144 genomic contexts at the DMR (centered) and surrounding regions. Genomic contexts
145 include 1kb promoter, exon, intron, CpG island, predicted CTCF binding sites, public
146 CTCF ChIP-seq peaks, and repetitive sequences. The line chart on top of each panel is
147 the average enrichment score for all the DMRs. **(C)** Number of unique and overlapped
148 DMRs between BWS IC2 LOM subtypes and LOS.
149
150
151 The predominant epimutation in BWS is loss of methylation (LOM) on maternal
152 allele at the imprinting control region (ICR) of *KCNQ1OT1*, namely imprinting center 2

153 (IC2; i.e. KvDMR1), which accounts for ~50% of documented cases [2]. Normal levels of
154 DNA methylation at this ICR should be ~50%, since only the maternal allele is normally
155 methylated. Average DNA methylation at this locus (i.e. chr11:2719755-2722554) was
156 17.59% in fibroblasts from IC2 patients and 45.18% in control samples, indicating LOM
157 in this BWS subtype. This LOM at IC2 is not complete (0%), because the epimutations
158 in BWS are mosaic which lead to a mixture of affected and not affected cells [2].
159 Hierarchical clustering based on these DMRs separated most BWS with IC2 LOM
160 samples away from controls (8/12; Figure S2B). No clustering based on sex was
161 observed. To further investigate whether the other four IC2 LOM samples that
162 intermixed with controls could be considered a novel IC2 subgroup, we subdivided the
163 IC2 LOM group into subgroup 1 (BWS_IC2_LOM_S1, n=8) and 2 (BWS_IC2_LOM_S2,
164 n=4) and compared them with the control group and each other. Subgroup-specific
165 comparisons against controls identified a larger number of DMRs (23,582) in subgroup
166 1 than in subgroup 2 (1,915). 672 DMRs were identified as shared between the two IC2
167 LOM subgroups (Figure 1C). We then compared DMRs between LOS and the two IC2
168 LOM subgroups and found that 2.89% (683) DMRs of subgroup 1 and 1.93% (37)
169 DMRs of subgroup 2 are shared with LOS (Figure 1C and Table S1). In total, 17 DMRs
170 were shared between the two IC2 LOM subgroups and LOS, including at the promoters
171 of *KCNQ1OT1* (i.e., KvDMR1) and *HOXA-AS2* and gene bodies of the epigenetic
172 modifiers *HDAC4* and *PRDM8* (Figure 1C and Table S1). Further, one of the shared
173 DMRs is in the gene body of *NR2F1-AS1*, a gene associated with tumors reported in
174 BWS, namely rhabdomyosarcoma and hepatocellular carcinoma [45].
175 The second most common epimutation in BWS (~5% of cases [2]) is gain of

176 methylation (GOM) at the *H19/IGF2* imprinted domain (i.e. IC1). Compared to controls,
177 fibroblasts from BWS IC1 samples have 19.09% more DNA methylation at
178 chr11:2015075-2022914, which encompass the IC1 and the entire *H19* gene body. This
179 synchronized DNA methylation change in IC1 and *H19* gene body has also been
180 reported in recurrent spontaneous abortion patients [46]. Cluster analysis of the DMRs
181 grouped the three BWS samples with IC1 epimutation away from controls (Figure S2A).
182 No clustering based on sex was observed. Heritable genetic mutations (including large-
183 scale chromosomal rearrangement but not uniparental disomy) caused loss-of-
184 imprinting accounts for up to 9% of BWS cases [2]. Specifically, genetic mutations at the
185 *CDKN1C* gene occur in ~5% of BWS cases [2]. Figure S2C shows the cluster analysis
186 of DMRs between BWS with *CDKN1C* variant and controls. Of note, no IC1 or IC2
187 epimutations were detected in this subgroup.

188 Sex-specific comparisons of the IC2 LOM subtype identified a larger number of
189 DMRs in males than females when compared to controls (53,488 vs. 26,459,
190 respectively). As per BWS, LOS males have twice the number of methylome
191 epimutations than females (75,370 and 34,270 DMRs, respectively). Interestingly,
192 DMRs in male LOS are mainly hypermethylated while DMRs in female LOS are mainly
193 hypomethylated (Figure S3B and S3C).

194 Allele-specific methylation analysis in the bovine control group clearly separates
195 the maternal and paternal alleles in the PCA plot (Figure S1C) and identified 10,143
196 DMRs between the parental alleles, with the majority (>99%) being hypomethylated on
197 the paternal allele (Figure S3D). Some of the identified DMRs overlap with known
198 imprinted genes and exhibit the expected allelic patterns. For example, lower paternal

methylation at *IGF2R* ICR (89.37%; chr9:96222606-96223965), lower paternal methylation at IC2 (89.55%; chr29:48908847-48909926), and higher paternal methylation at IC1 (68.16% and 73.02% at chr29:49504687-49505606 and chr29:49507607-49510366, respectively). Interestingly, there seems to be a spread of DNA methylation on the paternal allele from the IC1 to *H19*, creating allelic specific methylation over the gene body (chr29:49511527-49512886) with 9.70% methylation on maternal allele and 55.24% on paternal allele. Allele-specific comparisons between LOS and controls, identified that both alleles are mainly hypermethylated in LOS, which is similar to the finding of the overall comparison (Figure S3E and S3F). However, the LOS-associated allele-specific DMRs are, for the most part (>80%), not shared between the parental alleles. As expected, the normally methylated maternal allele of the *IGF2R* ICR (chr9:96222606-96223885) is hypomethylated in LOS when compared to controls, similar to what has been previously published by us and others [21,24]. We have previously reported maternal LOM at IC2 in some LOS fetuses [7]. In the current study, one LOS fetus (LOS_#3, the heaviest female) shows a 54.36% decrease in DNA methylation at maternal IC2 (chr29:48908887-48909846) than the control group mean (90.98%). However, the group mean difference between LOS and control is 14.7% at maternal IC2 which does not pass the 15% filtering threshold and is not included in the results, although it indeed shows some level of hypomethylation. Similar to our previous findings, maternal hypermethylation at IC1, was not observed in LOS, however, we detected hypermethylation at IC1 (chr29:49504927-49505566, 17.98%) and the *H19* gene body (chr29:49512727-49512926, 21.02%) on the paternal allele, which is a novel finding of unknown significance.

222 Overall, hypomethylated DMRs in BWS and LOS (including overall, sex-
223 specific, and allelic-specific comparisons) are significantly enriched for gene body, exon,
224 intron, CpG island, CpG shore, CpG shelf, predicted CTCF binding sites, and public
225 CTCF ChIP-seq peaks, and depletion for repetitive sequence according to permutation
226 tests (Figure 1, S2, S3, and Table S1). Hypomethylated DMRs also show significant
227 enrichment for gene promoter (1kb) in BWS related comparisons and LOS allele-
228 specific comparisons. For hypermethylated DMRs in most of the LOS and BWS
229 comparisons, there is also enrichment for CpG island, CpG shore, CpG shelf, and
230 CTCF binding sites. However, the pattern is varied for gene promoter, exon, intron, and
231 repetitive sequence based on sex and BWS subtype. For example, BWS with IC2 LOM
232 and BWS with IC1 GOM show enrichment at these elements but BWS with *CDKN1C*
233 variant shows depletion.

234 Interestingly, among the top 50 more severely dysregulated regions (both
235 hyper- and hypomethylated) in the BWS with IC2 LOM group, 11 hypomethylated DMRs
236 overlap the gene promoter and/or body of several *HOX* genes, including *HOX* A3, A4,
237 A5, A6, A7, B7, B9, C4, C10, and D4 from four *HOX* clusters located on chromosome 7
238 (A), 17 (B), 12 (C), and 2 (D) (Figure 2 and Table S1). Most of these genes also had
239 hypomethylated DMRs in the BWS with IC1 GOM group. For example, the gene body of
240 *HOXC4* (chr12:54446778-54448977) had a 41.17% decrease in the BWS IC2 LOM
241 group and 58.58% decrease in the BWS IC1 GOM group. Further, the promoter of
242 *HOXD4* (chr2:177015615-177017974) also had reduced DNA methylation in these
243 groups (30.26% and 36.05%, respectively). In LOS, *HOX* A3, A5, A6, B3, C4, C10, D4,
244 and D11, also have modest alterations (15%-20%) in DNA methylation when compared

245 to controls. *HOX* genes play roles as both oncogenes and tumor suppressors in
246 regulating cell proliferation and differentiation [47]. Both hyper- and hypomethylated
247 *HOX* genes are frequently observed in cancer, including Wilms tumor, which is the most
248 commonly diagnosed tumor in BWS [47,48]. *HOX* genes are important transcription
249 factors during embryonic development and exhibit spatial and temporal-specific
250 expression patterns [49]. In human, the number in the name of *HOX* genes are based
251 on their order of spatial expression pattern with *HOX* 1 being expressed in the most
252 anterior side (i.e. the head) and *HOX* 13 is the most posterior side (i.e. the trunk) [47].
253 *HOX* 1 to 4, expressed in the head region, are more frequently associated with DMRs in
254 BWS, especially *HOX* 4. This observation may be partially the result of the source of our
255 BWS fibroblast cells as some samples were collected from the skin behind the ear;
256 further studies with various sources of fibroblasts will be needed to validate this finding.
257 Importantly, hypomethylated DMRs were identified in the top DMR lists in the IC2 LOM
258 and IC1 GOM when compared to BWS with *CDKN1C* variant. Identified DMRs in BWS
259 reflect altered *HOX* expression during embryonic development and have the potential to
260 serve as molecular markers to diagnose and distinguish different subtypes.
261

262 **Figure 2. DNA methylation alterations at *HOX* gene clusters in BWS and**
263 **LOS. (A)** *HOXA* cluster. **(B)** *HOXB*. **(C)** *HOXC*. **(D)** *HOXD*. Top panel is human genome
264 and bottom panel is bovine genome for each plot. Additional gene annotation from NCBI
265 RefSeq is added since some *HOX* genes are not annotated in the R package for UCSC
266 annotation. All the *HOX* genes from the same cluster are located on the same strand.
267 The order of *HOX* genes may be opposite for human and bovine.
268

269

270

271 In summary, the abovementioned findings demonstrate methylome distinctions
272 between the epigenetically- and also the genetically-identified syndrome beyond the
273 signature DMRs or deletions. BWS children have up to a 28% risk of developing tumors
274 and currently, only 80% of patients can be corroborated molecularly, future studies will
275 need to be performed to determine their usefulness in diagnosing these children better
276 and more specifically.

277 **Gene expression profile and alternative splicing events in BWS and LOS**

278 Ribosomal RNA depleted total RNA sequencing was conducted to determine
279 the global transcript profile in human and bovine fibroblast samples. In accordance with
280 the cell-type used here, extracellular matrix genes show high expression level in both
281 species, including collagen (*COL1A1/2*, *COL3A1*, *COL5A1/2, COL6A1/2/3*, and
282 *COL12A1*), thrombospondin (*THBS1*), vimentin (*VIM*), filamin (*FLNA*), fibronectin (*FN1*),
283 and tenascin (*TNC*). Differentially expressed genes (DEGs) were determined for all the
284 comparisons in Table 1 and the DEG details can be found in Table S2.

285 The transcriptomes of the human fibroblasts are not distinct between them and
286 as a result the samples are intermixed in the PCA (Figure S1D). Twelve DEGs were
287 detected when all BWS samples were compared to controls and were all downregulated
288 in BWS. Those genes were *LINGO2*, *MDGA2*, *NGFR*, *MIR92A1*, *FLG*, *BMP2*,
289 *SIGLEC15*, *SLAMF7*, *LIPH*, and three novel genes. Downregulation of *MDGA2*, *NGFR*
290 and *BMP2* have been reported in tumorigenesis [50–52]. DEG based hierarchical
291 clustering separates each of the BWS subtypes from the control samples (Figure 3B);

292 however, no obvious clustering based on sex was observed in the IC2 LOM group.
293 Subtype-specific analyses revealed that 23 genes were different from controls in IC1
294 GOM, 8 in IC2 LOM and 10 in *CDKN1C* variant (Figure 3A, 3B, and 3C). When the
295 subgroups of IC2 LOM were taken into consideration, 742 transcripts were identified as
296 different from controls in subgroup 1 and two in subgroup 2. Nine DEGs were shared
297 between IC2 LOM subgroup 1 and IC1 GOM (Table S2), making them potential
298 subtype-specific biomarkers. Those were the upregulated *GRID1*, *KCNA4*, *SESN3*,
299 *GOLGA8CP*, *NTRK3*, *NKD1*, and the downregulated *PTPRE*, *GPRC5A*, and
300 *SERPINB2*. The other DEGs were unique to each BWS subtype. Some of the DEGs are
301 associated with tumorigenesis, including downregulation of *LOXL4* in BWS with IC1
302 GOM, upregulation of *PECAM1* in BWS with IC2 LOM, and downregulation of *TAP2* in
303 BWS with *CDKN1C* variant [53–55]. In addition, *SESN3* is involved in insulin sensitivity
304 and in tumorigenesis [56,57]. There was no significant group difference, however, in
305 transcript abundance for the signature imprinted genes of BWS (e.g. *H19*, *IGF2*,
306 *CDKN1C*, and *KCNQ1OT1*) in these comparisons. High variations in transcript level
307 were observed for *H19* and *IGF2* in the control group (more than 70-fold difference
308 between the highest and lowest individual). *IGF2* transcript abundance was very high in
309 BWS_#11 (IC2 LOM) and #16 (*CDKN1C* variant), more than four times higher than the
310 highest control sample, and *KCNQ1OT1* level was three-fold higher than controls in
311 BWS_#7 and #11 (IC2 LOM).
312
313 **Figure 3. Expression heatmap and genomic context enrichment for**
314 **differentially expressed gene (DEG).** DEG for human BWS subtypes compared with
315 control group **(A-C)**, and bovine LOS group compared with control group **(D)**. The left

panel is the heatmap with hierarchical clustering and sex annotation which include male (M), female (F), and conflict (C). Gene expression is shown as log2 transformed difference between sample and row average. The right panels are the distribution of genomic contexts at the DEG (centered at 1kb promoter or gene body) and surrounding regions. Genomic contexts include differentially methylated regions (DMR), differential interacting regions (DIR), differential topologically associating domains (diff_TAD), and differential compartments identified in corresponding group comparisons. The line chart on top of each panel is the average enrichment score for all the DEGs.

For bovine, PCA separates the samples based on treatment (Figure S1E) although not by parental allele (Figure S1F) indicating similar expression profiles from both alleles. Transcriptome comparison between LOS and controls identified 112 DEGs and hierarchical clustering clearly separated the two groups (Figure 3D). As expected, the signature imprinted gene of LOS (*IGF2R*) was found downregulated two-fold in the LOS group. *IGF2R* is normally highly expressed from the maternal allele and its protein can bind to IGF2 protein with high affinity to mediate the degradation of excess IGF2 [58]. Studies in mice have shown that maternal inheritance of *Igf2r* null variant causes increase of IGF2 and IGF binding proteins, ~25%-30% increase of body weight, proportionately increased organ size, and perinatal death [59]. The overgrowth phenotype in LOS could be largely attributed to changes in the IGF2 signaling pathway. In addition, several DEGs in LOS have also been reported in tumor studies in human, such as downregulation of *AOX1*, *CYGB*, and *TMOD2*, and upregulation of *BCL11A*, *AGRN*, and *LEF1* [60–65]. Four DEGs were shared between LOS and BWS IC2 LOM

341 subgroup 1, namely the upregulated *EMID1*, and the downregulated *SERPINE2*,
342 *LRRC8D*, *GPX1*. When analyzing the transcriptome in a sex-specific manner, male LOS
343 has 109 DEGs when compared to controls while the number of DEGs is doubled in
344 female LOS (213). Of those DEGs, eight were detected in both male and female LOS,
345 namely *FBLN7* (down/down in male/female), *CYGB* (down/down), *LOC112446407*
346 (up/up), *ABLIM3* (down/down), *IGF2R* (down/down), and three novel genes
347 MSTRG.7897 (up/up), MSTRG.13383 (up/down), and MSTRG.22799 (down/up).
348 Downregulation of *ABLIM3* has also been reported in hepatocellular carcinoma [66].
349 Allele-specific comparison in the bovine control group identified 122 genes with
350 significant biased expression from one allele, including *IGF2R* which has about three-
351 fold higher expression from the maternal allele than the paternal allele. In addition,
352 *GSTK1*, *MAN1B1*, *GSTZ1*, *ARSG*, *SLC25A29*, *TLDC1*, *PDIA4*, *RWDD2A*, and *NARS*
353 have been reported as imprinted genes or to have biased allelic expression in various
354 mammalian species by other researchers[67–75]. The allelic expression pattern of
355 *Gstz1* is associated with expression quantitative trait loci (eQTL) in mice [67].
356 *SLC25A29*, together with our previously reported *BEGAIN* [76], show biased paternal
357 expression and are located close to the *DLK1-DIO3* imprinted domain, which suggest
358 broader regulatory impacts of this domain in bovine. *TLDC1* is associated with gamete-
359 specific promoter methylation patterns in human [77]. Although not possible to
360 determine the imprinting status of the remaining genes in this study, as we only have F1
361 samples with one parental cross (*i.e. B.t. indicus X B.t. taurus*), other genes in the list
362 should be further characterized to determine if they are imprinted or rather they behave
363 as eQTL as we have previously shown [78].

364 Allele-specific DEGs were identified in LOS. *IGF2R* shows a significant
365 decrease (~2.4 fold) from the maternal allele making it similar in expression level to the
366 paternal allele. The expression of this gene from the paternal allele in LOS was similar
367 to control. In addition, among the 122 genes with biased allelic expression in controls,
368 *ECHDC2* has higher paternal expression in controls and is downregulated on the
369 paternal allele in the LOS group, and *IGF1R* has lower maternal expression in controls
370 and is upregulated on the maternal allele in LOS group. *ECHDC2* is involved in
371 branched chain amino acid metabolism and its lower expression increases cell viability,
372 and *IGF1R* encodes a membrane receptor for the IGF signaling and is positively
373 correlated with growth [79,80].
374 Regarding the association between changes in gene expression and DNA
375 methylation, upregulated DEGs in female BWS with IC2 LOM group showed enrichment
376 for hypomethylated DMRs at promoter and gene body, and downregulated DEGs in this
377 group showed enrichment for both hypo- and hypermethylated DMRs at promoter and
378 gene body according to permutation tests (Table S2). In addition, upregulated DEGs in
379 LOS group showed enrichment for both hypo- and hypermethylated DMRs at promoter
380 and hypomethylated DMRs at gene body, and downregulated DEGs in this group
381 showed enrichment for hypermethylated DMRs at promoter and gene body. Given that
382 DMRs overlapping exons can influence alternative splicing, we analyzed the total RNA
383 sequencing data to identify alternative splicing (AS) events in BWS and LOS [81]. For
384 humans, 1,183, 1,292, and 1,151 AS events were identified in BWS with IC1 GOM, IC2
385 LOM, and *CDKN1C* variant groups when compared to the control group (Figure 4A,
386 S4A, S4B, and Table S2). For bovine, 746 AS events were detected in the LOS group

387 when compared to the control group (Figure 4B and Table S2). Regardless of species or
388 subtypes, most AS events were skipped exons with an average of 43.5% across
389 pairwise comparisons. AS events appeared to be evenly distributed across
390 chromosomes in BWS and LOS. Few genes shared alternative splicing events between
391 BWS and LOS, however, we did find instances of shared splicing events in pathways
392 relevant to overgrowth. For example, genes involved in the WNT signaling pathway
393 (*TCF7L2*, *HDAC1*, *PRKC1*) had skipped exons in both BWS with *CDKN1C* variant and
394 LOS groups. WNT signaling pathway regulates cell proliferation and differentiation, and
395 is associated with Hippo pathway to control organ size [82]. YAP1, which is a
396 transcription co-activator in Hippo pathway, has been reported with increased non-
397 phosphorylated (active) protein level in skeletal muscle of LOS fetuses by us [26].
398 Although none of the exons of the mentioned genes were differentially methylated,
399 these data suggest alternative splicing as a contributor or outcome of the overgrowth
400 phenotype.
401
402
403 **Figure 4. Identification of alternative splicing events in Beckwith-**
404 **Wiedemann Syndrome (BWS) and Large Offspring Syndrome (LOS) and**
405 **correlation network in BWS. (A)** and **(B)** The genomic distribution, numbers, and
406 types of alternative splicing events identified in human BWS IC2 LOM (A) and bovine
407 LOS (B) groups when compared to their control groups. Manhattan plots show
408 alternatively spliced genes aligned along the x-axis by their position on the chromosome
409 and on the y-axis by their log10(false discovery rate). The horizontal dashed black line
410 represents a significance threshold of 0.05. **(C)** Correlation between gene expression
411 profile in 15 modules and BWS diagnosis and scores determined by weighted gene co-
412 expression network analysis (WGCNA). Numbers in the heatmap represent correlation

413 coefficient. Significance is shown as * (<0.05), ** (<0.01), and ns (≥0.05) in parentheses.
414 (D) Gene interaction network within purple module and hub genes were highlighted.
415
416
417 To understand the correlation between transcriptional regulation and phenotypic
418 characteristics of human BWS features, weighted correlation network analysis
419 (WGCNA) was performed to determine gene clusters with high correlation in BWS. As a
420 result, genes were grouped into 15 modules based on similarity of expression. The grey
421 and purple modules showed a positive correlation (p-value < 0.05) with macroglossia
422 and omphalocele (Figure S4C) suggesting that genes with high expression in these
423 modules may serve as biomarkers of these BWS pathologies. Other suggestive
424 features such as a birthweight greater than two standard deviations (Birthweight_2SDs),
425 facial nevus simplex, and ear creases or pits also showed higher correlation with purple
426 modules (Figure S4D). Based on the standard of BWS scoring from Wang's study [83],
427 a comprehensive BWS score was generated for each individual and expressional trends
428 of purple module showed a significantly positive correlation with BWS score compared
429 to other modules (Figure 4C). The 94 genes in this module showed enriched functions
430 in RNA processing and regulation of immune system process (Table S2). Among the
431 genes within the purple module, 21 hub genes with high BWS score correlation (BWS
432 correlation ≥ 0.5) and module membership score (MM_purple ≥ 0.8) were identified
433 (Figure 4D and Table S2), and one of these genes, Myb-Like, SWIRM And MPN
434 Domain-Containing Protein 1 (*MYSM1*) is a transcription coactivator and participates in
435 chromatin remodeling and repression of innate immunity and autoimmunity [84].
436 *KCNQ1OT1*, a long-noncoding RNA implicated in BWS, was also enriched [83]. In

437 addition, *MIR145* which encodes has-miR-145-5p was enriched in this module, which
438 has been reported to be differentially expressed in BWS patients with macroglossia by
439 us [26].

440 **Characterization of chromosome architecture in human and bovine**

441 Hi-C sequencing was performed to determine genome-wide chromosome
442 architecture and its relationship with the methylome and transcriptome in human and
443 bovine fibroblasts. Read alignment information can be found in Table S3.
444 We first characterized the two fundamental elements of chromosome
445 architecture, namely chromosome compartments and topologically associating domains
446 (TAD), in the human and bovine control groups (Table S3). For chromosome
447 compartment, the A (positive eigenvalue, enriched in euchromatin) and B (negative
448 eigenvalue, heterochromatin) compartments each roughly occupies half of the
449 sequenceable autosomal genome in human (49.4% for A and 50.6% for B) and bovine
450 (47.3% and 52.7%). Sex-specific and allele-specific analyses show a similar ratio (Table
451 S3). 13,912 and 12,699 TADs (including sub-TADs) were found in the autosomes of
452 human and bovine control groups, respectively, with similar TAD length distributions
453 (Figure 5A, 6A, and Table S3). BWS, LOS, and sex-specific groups have similar TAD
454 length distribution as their corresponding control group. For bovine allele-specific
455 groups, the TAD length is decreased likely due to largely reduced read coverage after
456 filtering for the parental alleles. The TAD boundaries in both species and all groups
457 show enrichment for predicted CTCF binding sites, public CTCF ChIP-seq peaks,
458 transcription start sites, and CpG islands, and depletion for repetitive sequences (Figure

459 5B and 6B).

460

461 **Figure 5.** Characterization of human chromosome architecture and alterations in
462 Beckwith-Wiedemann Syndrome (BWS). (A) Length distribution of topologically
463 associating domains (TAD) detected in control group. Level 1 TAD refers to the outmost
464 TADs. The red line marks the median length. (B) Genomic context enrichment at control
465 group TAD and subTAD boundaries (centered) and flanking regions. Genomic contexts
466 include predicted CTCF binding sites, CTCF public ChIP-seq peaks, transcription start
467 sites (TSS), CpG island, and repetitive sequences. The line chart on top of each panel
468 is the average enrichment score for all the boundaries. (C) Genome-wide distribution of
469 differentially methylated regions (DMR), differential compartments (diff_compartment),
470 differential TAD (diff_TAD), differential interacting regions (DIR) frequency, and
471 differentially expressed genes (DEG) identified when comparing BWS group with control
472 group. Red dots indicate an increase in CpG methylation, eigenvalue (towards A
473 compartment), and expression level, and black dots indicate decrease. The vertical
474 location of dots indicates the level of changes. Mb = megabases. (D) Correlation
475 between chromosome compartment changes (measured by eigenvalue) at gene
476 promoters and log2 transformed fold change (logFC) of protein-coding genes and
477 lncRNA. The left panel shows genome-wide correlations, and the right panel shows
478 correlation at regions of differential compartment. (E) Examples of alterations of
479 chromosome architecture associated with gene expression changes in BWS patients.
480 The top matrix plot shows the Hi-C contacts at 10kb resolution. A deeper blue color
481 indicates higher physical contact frequency. Red lines mark the TADs. All the statistical
482 analyses are conducted by comparing the bottom right group with the top left group.
483 Tracks below the matrix show the location of different elements along this genomic
484 region. The compartment track shows the A/B compartment represented by eigenvalue.
485 The blue and red lines indicate the top left and bottom right group, respectively. The
486 grey area marks significantly differential compartments between groups detected at
487 different resolutions which may overlap and show deeper color. The diff_TAD tracks

488 show the differential TADs in the bottom right group. The DIR tracks show the DIRs that
489 gain or lose contact. The DIRs are also marked in the matrix with purple or black
490 squares. The DMR tracks show DMRs that gain or lose CpG methylation. The DEG
491 tracks show up- or down-regulated DEGs. The DMR_CTCF_pre track shows predicted
492 CTCF binding sites that overlap with DMRs. The CTCF_pub_ChIP track shows the
493 public CTCF ChIP-seq peaks. The UCSC annotation tracks show the UCSC gene
494 annotation. (F) Contact frequency changes between DEG *SESN3* and super-enhancers
495 on Chromosome 11 (position denoted on the X axis) in BWS IC1 GOM group. Two
496 genes (not DEG) surrounding *SESN3* are also plotted for comparisons.

497

498 **Chromosome architecture disruption in BWS and LOS**

499 Differential chromosome compartments were detected between groups at 1 mb,
500 500 kb, 250 kb, 100 kb, 50 kb, and 25 kb resolutions for all the comparisons in Table 1
501 (Table S4). Differential TADs were detected between groups at 25kb resolution for all
502 the comparisons in Table 1 (Table S5). The group mean Hi-C contact matrix was used
503 for the differential TAD analysis instead of all individual samples in each group, due to
504 software limitations. Types of differential TAD between groups are categorized as loss,
505 strength change, merge, split, zoom, and complex (for description refer to Methods). In
506 addition, differentially interacting regions (DIR) were detected for binned genomic
507 regions between groups at 1 mb, 500 kb, 250 kb, 100 kb, and 50 kb resolutions for all
508 the comparisons in Table 1 (Table S6), and changes are shown as gain and loss in
509 contact frequency. Changes in chromosome architecture were associated with DEGs
510 according to permutation tests. For example, upregulated DEGs in BWS with IC1 GOM
511 group showed enrichment for differential compartment with increased eigenvalue

512 (change towards A compartment) at promoters, downregulated DEGs in female BWS
513 with IC1 GOM group showed enrichment for differential TAD at promoters, and
514 downregulated DEGs in LOS group showed enrichment for differential compartment
515 with decreased eigenvalue at promoters (Table S2). Together with the previously
516 mentioned association between DEGs and DMRs, these observations indicate that the
517 causes of gene misregulation in BWS and LOS are diverse and there is not a single
518 dominant mechanism.

519 For human BWS subtypes, BWS with IC1 GOM group has higher number of
520 unique differential compartments with longer length (50 and 500 kb for 25th and 75th
521 percentile, same format below) than BWS with IC2 LOM (50 and 300 kb) and BWS with
522 CDKN1C variant (50 and 400 kb) groups when compared to the control group (Table 1
523 and S4). Increased number and decreased length of unique differential compartments
524 were found in sex-specific comparisons, including male BWS with IC2 LOM (25 and 200
525 kb), female BWS with IC1 GOM (25 and 250 kb), female BWS with IC2 LOM (25 and
526 125 kb), and female BWS with CDKN1C variant (25 and 250 kb). For bovine, the
527 lengths of unique differential compartments in the LOS group (250 kb and 1 mb) are
528 longer than any of the human BWS subtypes. Unique differential compartments in male
529 LOS (100 kb and 1 mb) are lower in number but longer in length than female LOS (25
530 and 250 kb), and in LOS maternal allele (25 and 500 kb) are also lower in number but
531 longer in length than LOS paternal allele (25 and 75 kb).

532 Given the prevalence of differential compartments in BWS and LOS, Pearson
533 correlation tests were performed between eigenvalue changes at gene promoters,
534 which reflect A/B compartment changes, and expression changes of all protein-coding

535 genes and lncRNA (not just DEGs). Small RNA genes were not included in the test due
536 to the uncertainty of their promoter locations. For genome-wide test, a very weak
537 positive correlation was observed for both BWS (correlation coefficient equals to 0.078)
538 and LOS (0.13; Figure 5D and 6D). The positive correlation increased in BWS (0.32)
539 and LOS (0.29) when only considering the differential compartment regions, which
540 indicates that the compartment changes have impacts on gene expression regulation in
541 the differential compartment regions (Figure 5D and 6D). The same trend was detected
542 in most other BWS and LOS involved comparisons excluding BWS with IC2 LOM,
543 female LOS, and female LOS paternal allele (figure not shown). These exceptions have
544 similar weak correlation at differential compartment regions as their genome-wide
545 correlation, which indicates the compartment changes have less impact on gene
546 expression in these groups. Interestingly, the differential compartments in these
547 exceptions have a relatively shorter length as mentioned before. The Pearson
548 correlations between eigenvalue changes and methylation changes at DMRs are
549 generally very low (correlation coefficient absolute values are less than 0.05) in BWS
550 and LOS involved comparisons (figure not shown).
551 BWS with *CDKN1C* variant group had the highest number (286) of differential
552 TAD among BWS subtypes, while in LOS there were 369 differential TADs when
553 compared to the control group (Table 1 and S5). For sex-specific comparisons, both
554 female BWS with IC2 LOM (695) and female LOS (526) showed a higher differential
555 TAD number than male BWS with IC2 LOM (239) and male LOS (231) respectively,
556 when compared to their controls. This pattern could be related to the genome-wide sex-
557 specific DNA methylation changes. As previously mentioned in bovine, DMRs are

558 mainly hypermethylated in male LOS and occur at regions that have intermediate to
559 high DNA methylation level in the control group. Although these regions are enriched for
560 predicted CTCF binding sites, the normally high DNA methylation level is expected to
561 block the CTCF binding [85] which is seen as depletion of public CTCF ChIP-seq peaks
562 in the heatmap (Figure S3B). And further hypermethylation could have minor impact on
563 CTCF binding change in male LOS. In female LOS, DMRs also mainly occur at regions
564 that have high DNA methylation level in the control group, and have enrichment for
565 predicted CTCF binding sites and depletion for public CTCF ChIP-seq peaks (Figure
566 S3C). However, the global loss of methylation could lead to activation of these potential
567 CTCF binding sites and change the CTCF binding profile and TADs, which is reflected
568 by the higher number of differential TAD detected in female LOS than male LOS.
569 Similarly, male BWS with IC2 LOM also has a higher ratio of hypermethylated DMR than
570 female BWS with IC2 LOM. Another interesting observation is that the boundaries for
571 differential compartments are usually also the boundaries for TADs or sub-TADs, which
572 means TAD is the unit for compartment changes. So, the more frequently disrupted
573 TADs in female LOS may explain the observed higher number but shorter length of
574 differential compartments in female LOS than male LOS, and that the potentially
575 increased active CTCF binding sites in female LOS leads to smaller TADs and further
576 causes fragmentation or length decrease of nearby differential compartments. These
577 shortened differential compartments in female LOS have a much lower impact on gene
578 expression change (correlation coefficient = 0.027) than male LOS (0.27).

579 **Vulnerable chromosomes and chromosomal regions in BWS and LOS**

580 The comparative genomic distribution of DMRs, differential compartments,
581 DIRs, and DEGs are summarized in Figure 5C, 6C, S5, and S6. For humans, when
582 ordering chromosomes by the ratio of total differential compartment length to
583 chromosome length, chr2, 3, 6, 9, 10, 11, 12, 13, 17, 18, and 20 are often in the top ten
584 list in BWS (Figure 5C and S5). It is worth noting that the differential compartments
585 detected on chr9 and 18 are mainly around the unmappable region at the centromere.
586 In sex-specific comparisons, chrX and Y also exhibit frequent compartment changes
587 (figure not shown). For bovine, chr27, 21, 4, 10, 28, 8, 1, 14, 23, and 13 are the top ten
588 in LOS group when ordered by the ratio of total differential compartment length to
589 chromosome length (Figure 6C). Sex-specific comparisons show a somewhat different
590 top ten list for male LOS (chr27, X, 26, 21, 28, 8, 1, 13, 4, 14) and female LOS (chr27,
591 4, 7, 16, 10, 14, 18, 5, 8, 22; Figure S6B and S6C). When comparing paternal allele to
592 maternal allele in the control group, chr27, 9, 4, and 6 are the most different
593 chromosomes regarding compartment (Figure S6D). Chr27 is largely altered in LOS
594 paternal allele (Figure S6F) and male LOS paternal allele group (figure not shown).
595 Compartment changes on chr9 become frequently detected in LOS involved sex-
596 specific and allele-specific comparisons that female LOS paternal allele shows a similar
597 pattern of change as in Figure S6D, and male LOS maternal allele shows an opposite
598 pattern (figure not shown). Chr4 is largely altered in LOS paternal allele (Figure S6F)
599 and female LOS paternal allele group (figure not shown). Interestingly, the differential
600 compartment pattern on chr4 in LOS paternal allele is almost the opposite of the pattern
601 of allelic comparison in control group (Figure S6D), which suggests the paternal allele of
602 chr4 in LOS has a level of maternalization in chromosome compartment. Chr6 is largely

603 altered in LOS maternal allele (Figure S6E), male LOS maternal allele, female LOS
604 maternal allele, and female LOS paternal allele group (figure not shown).
605
606
607 **Figure 6. Characterization of bovine chromosome architecture and**
608 **alterations in large offspring syndrome (LOS). (A)** Length distribution of topologically
609 associating domains (TAD) detected in control group. **(B)** Genomic context enrichment
610 at control group TAD and subTAD boundaries (centered) and surrounding regions. **(C)**
611 Genome-wide distribution of differentially methylated regions (DMR), differential
612 compartments (diff_compartment), differential TAD (diff_TAD), differential interacting
613 regions (DIR) frequency, and differentially expressed genes (DEG) identified when
614 comparing LOS group with control group. **(D)** Correlation between chromosome
615 compartment changes (measured by eigenvalue) at gene promoters and log2
616 transformed fold change (logFC) of protein-coding genes and lncRNA. **(E)** Examples of
617 alterations of chromosome architecture associated with gene expression changes in
618 LOS fetuses. Detailed descriptions refer to Figure 5.
619
620
621 The maternal and paternal alleles show a similar large scale (chromosome-
622 wide) genomic distribution of allele-specific DMRs in LOS (Figure S6E and S6F). For
623 example, over 61% of DMRs on chr1 are found in the last 30 mb (19%) of this
624 chromosome in LOS maternal allele and LOS paternal allele. This enrichment is not
625 caused by a biased distribution of genomic variants since the methylation differences
626 between control paternal allele and control maternal allele are evenly distributed on this
627 chromosome (Figure S6D). However, as mentioned before, the precise genomic
628 coordinates of the DMRs are largely not shared (>80%) between maternal and paternal

629 alleles. This observation indicates that the two alleles of certain chromosomes or parts
630 of a chromosome are vulnerable for DNA methylation alterations in LOS, but the
631 alterations are largely not driven by the DNA sequence. The chromosome territory,
632 which refers to the special location of chromosomes in the nucleus, could be involved in
633 this vulnerability. According to human and mouse three-dimensional fluorescence in situ
634 hybridization (3D-FISH) images, many chromosomes have their two alleles separated
635 from each other [86,87]. Evidence have also shown that the two alleles of each
636 chromosome have a symmetrical relative location to other chromosomes in mouse [88].
637 Assuming bovine has the same features, if changes of microenvironment at symmetrical
638 spaces occur, they could have impacts on the same part of the two alleles. Other large
639 scale chromosome markers may also directly or indirectly contribute to this vulnerability,
640 such as histone modifications, which undergo rapid and dynamic reprogramming during
641 early embryo development, and are known to regulate the *de novo* establishment of
642 DNA methylation [89–91]. Future studies on chromosome territories in bovine will be
643 needed to understand this process.

644 Given the very low correlation between DMRs and differential compartments
645 and generally similar genomic distribution pattern of DMRs between maternal and
646 paternal alleles in LOS, the distinct distribution pattern of differential compartments
647 between maternal and paternal alleles suggests that other allele-specific alterations
648 exist in LOS and are involved in the regulation of chromosome compartment, such as
649 histone modifications and non-coding RNAs (ncRNAs). Chromosome compartment is
650 considered mainly regulated by phase separation mechanism that chromatin regions
651 with different physicochemical properties form distinct phases; phases with different

652 properties are spatially separated and phases with similar properties are clustered
653 [92,93]. Histone modifications and ncRNAs have been found to attract binding proteins
654 with intrinsically disordered domains that can lead to phase separation and
655 compartment change [94–96]. In addition, studies in mouse have shown that the
656 maternal and paternal alleles exhibit distinct histone modification profiles in early
657 embryo, which may account for possible allele-specific disruptions [89,91]. Future
658 studies on the histone modification profiles and ncRNA nuclear localization in BWS and
659 LOS will help build the connection and improve the understanding of the underlying
660 mechanisms.

661 **Association among DMRs, altered chromosome architectures, and DEGs in BWS**
662 **and LOS**

663 Several genomic regions are shown to illustrate the possible impact of DNA
664 methylation and chromosome architecture on gene expression changes in BWS and
665 LOS (Figure 5E, 6E, S7, and S8). For humans, gene sestrin 3 (*SESN3*) is upregulated
666 and located within a differential compartment region that changes toward A
667 compartment in BWS with IC1 GOM group, and no DMR is found within gene body or
668 promoter of *SESN3* (Figure 5E). Downregulated gene platelet derived growth factor
669 subunit A (*PDGFA*) resides in an A compartment region that gains contact with a B
670 compartment region in female BWS with IC2 LOM (Figure S7A). A small change toward
671 B compartment is observed although not significant, and two hypomethylated DMRs
672 overlap the gene body and promoter of *PDGFA*. Gene IQCJ-SCHIP1 readthrough
673 (*IQCJ-SCHIP1*) is downregulated but located within a differential compartment region

674 that changes toward A compartment in female BWS with CDKN1C variant group, and
675 the promoter and first exon of *IQCJ-SCHIP1* overlap with a hypermethylated DMR
676 (chr3:158785716-158788035, 12.33% increase) which was excluded in Table S1 due to
677 the 15% difference cutoff (Figure S7B). These changes are also shown in female
678 comparison with the other two BWS subtypes (figure not shown), which indicates the
679 alterations around *IQCJ-SCHIP1* are specific to this BWS subtype.

680 For bovine, gene KIT ligand (*KITLG*) is downregulated and located within a
681 differential compartment region that changes toward B compartment in LOS group, and
682 no DMR is found within gene body or promoter of *KITLG* (Figure 6E). The 5 mb
683 upstream region of the *DLK1-DIO3* imprinted domain shows different patterns of
684 chromosome compartment in male control and female control, with sex-specific
685 changes in male LOS and female LOS (Figure S8). A gigantic TAD (4 mb) is observed
686 mainly belonging to the B compartment in male control but not in female control. The
687 upregulated gene BCL11 transcription factor B (*BCL11B*) resides in this 4 mb TAD
688 which changes toward A compartment in male LOS group, and overlaps with multiple
689 DMRs (mainly hypomethylation) in the gene body (Figure S8A). In female LOS, the 4
690 mb TAD is not altered, but a small region that changes toward B compartment overlaps
691 with downregulated gene bradykinin receptor B2 (*BDKRB2*), and a hypermethylated
692 DMR is found in its gene body (Figure S8B). We have previously reported dysregulation
693 of multiple microRNAs in the *DLK1-DIO3* imprinted domain in both BWS and LOS [26],
694 and further investigation will be needed to validate whether this is a consequence of
695 compartment changes.

696 As mentioned before, DNA methylation at promoters and enhancers have

697 repressive impacts on gene expression through blocking binding sites for transcription
698 factors [97]. Chromosome compartments also have impacts on gene expression. The B
699 compartments are enriched for heterochromatin, mainly localized in the peripheral
700 regions in the nucleus or surrounding the nucleoli, and isolated from areas with
701 enrichment of transcription machineries (such as transcription factories) [98–102].
702 Oppositely, the A compartments are enriched for euchromatin, mainly localized in the
703 interior of nucleus between the two B compartment regions and associated with
704 transcription factories. The example loci above suggest a model that changes of DNA
705 methylation and chromosome compartment in BWS and LOS can simultaneously affect
706 gene expression in an independent manner. In the DEG examples of human *SESN3*
707 and bovine *KITLG*, *BCL11B*, and *BDKRB2*, no DMR was found at the promoters of
708 these DEGs, and their expression changes follow the alterations of their chromosome
709 compartments that changing towards A compartment leads to upregulation and
710 changing towards B compartment leads to downregulation. In the example of human
711 DEG *PDGFA*, the hypomethylated DMR at its promoter will further open the promoter
712 for transcription factor binding, however, its spatial change towards B compartment will
713 reduce the availability of transcription machineries in its surrounding microenvironment.
714 In the opposite example of human DEG *IQCJ-SCHIP1*, its spatial change towards A
715 compartment will provide more transcription machineries, but the hypermethylated DMR
716 at its promoter will prevent their binding. So, in these two cases, DNA methylation and
717 chromosome compartment form an antagonistic situation where the one with more
718 severe change or stronger effect might overcome the effect of the other on gene
719 expression regulation. This model could explain some of the conflicts between promoter

720 DNA methylation changes and gene expression changes observed in our BWS and
721 LOS datasets. In addition, alterations in TAD and chromosome compartment in BWS
722 and LOS could change enhancer availability and contacts with gene promoters, and
723 further affect gene expression. To investigate this mechanism, we embarked on the
724 identification of long-range contacts between super-enhancers from dbSUPER
725 database and promoter of the DEG *SESN3* in BWS IC1 GOM group as an example.
726 *SESN3* showed increased promoter contacts with the super-enhancer region
727 chr11:86430802-86457226, chr11:86168452-86236631, and chr11:85855833-85863446
728 from the closest upstream broad A compartment (Figure 5F). These super-enhancers
729 could also contribute to the upregulation of *SESN3* in BWS.

730 **Conservation between human and bovine**

731 Next, we investigated the conservation of TAD and chromosome compartment
732 between human and bovine genomes. For TAD in the control groups (with both sexes
733 included) a corresponding genomic region of the majority of human TADs can be found
734 in bovine (13,425 out of 13,912). Over these regions, 2,065 (~15.4%) human TADs
735 were found conserved in bovine with greater than 90% overlapping in both species
736 (Table S7). For sex specific analyses that include chrX, 1,960 (~13.7%) and 2,032
737 (~13.9%) human TADs were conserved in bovine for male and female control group,
738 respectively (Figure 7A and Table S7). For chromosome compartments in the control
739 groups (with both sexes included), continuous 25 kb bins with the same A or B
740 compartment assignment were merged that results in 4,086 regions in human, and a
741 corresponding genomic region can be found in bovine for most of them (3,713). Over

742 these regions, 1,813 (~48.8%) human compartments were found conserved in bovine
743 with greater than 90% overlapping in any species (Table S7). For sex specific analyses
744 that include chrX, 2,569 (~47.6%) and 1,989 (~46.8%) human compartments were
745 conserved in bovine for male and female control group, respectively (Figure 7C and
746 Table S7). For both TAD and chromosome compartment, the conservation pattern is
747 similar between male and female for autosomes but is different for chromosome X,
748 which is expected for inactivated X to show megadomain (lose TADs) and change
749 towards B compartment in female.
750
751 **Figure 7. Chromosome architecture conservations between human and**
752 **bovine. (A)** Genome-wide conservation of control group topologically associating
753 domains (TAD) between human and bovine. Each line indicates a conserved TAD.
754 Different colors refer to different human chromosomes. Mb = megabases. **(B)**
755 Overlapped differential TAD identified in some comparisons of Beckwith-Wiedemann
756 Syndrome (BWS) and large offspring syndrome (LOS). Information regarding
757 comparisons used refer to Methods. Each line indicates a shared differential TAD
758 between combined BWS and LOS results with greater than 90% overlapping in both
759 species. Green and black refer to shared differential TAD with same and different type
760 between BWS and LOS, respectively. **(C)** Genome-wide conservation of control group
761 A/B compartment between human and bovine. Each line indicates a conserved
762 compartment. Red and blue refer to A and B compartment, respectively. **(D)** Overlapped
763 differential compartments identified in BWS and LOS. Each line indicates an overlapped
764 differential compartment. Green and black refer to differential compartments with same
765 or opposite direction of change between BWS and LOS, respectively. **(E)** and **(F)** An
766 example region of conserved chromosome architecture in control group between human
767 and bovine, and shared compartment changes between BWS and LOS. Detailed
768 descriptions refer to Figure 5E.
769

770

771 Similar approaches were used to find shared differential compartments between
772 BWS and LOS. For humans, the lists of differential compartments from BWS involved
773 comparisons (index 1, 2, 3, 4, 8, 9, 10, 11, and 12) were combined, which resulted in
774 2,117 records (allowing overlap and identical records), and a corresponding genomic
775 region can be found in bovine for 1,689 records. The combined bovine list from LOS
776 involved comparison (index 16, 18, and 19) contains 1,118 records. In total, 772
777 (~45.7%) records of differential compartments from human overlap with bovine records
778 without a cutoff for overlap percentage and regardless of direction of change (Figure 7D
779 and Table S7). If the overlap cutoff in both species is set to 20%, 50% or 80%, 193
780 (~11.4%), 80 (~4.7%) or 28 (~1.7%) records will remain, respectively. Chromosome X,
781 6, 10, 9, 7, and 2 in human have the highest number of overlapped differential
782 compartments, and in bovine the most frequent chromosomes are X, 8, 4, 23, 13, and
783 28 (Figure 7D). Similarly for differential TADs, 2,739 and 1,126 records were combined
784 for BWS and LOS involved comparisons, respectively, and 2,635 human records have a
785 corresponding genomic region found in bovine. And the overlapped record number is
786 153 (5.8%) with a 90% overlap cutoff in both species. Chromosome 1, 5, 3, 2, and 14 in
787 human have the highest number of overlapped differential TADs, and in bovine the most
788 frequent chromosomes are 7, 16, 1, 10, and 2 (Figure 7B). A 6 mb region is shown as
789 an example to illustrate the TAD and compartment conservation and shared
790 compartment changes between human and bovine (Figure 7E and 7F). In control
791 groups, TADs and compartments are largely conserved in the right half of this region,
792 and a 500 kb differential compartment region that changes toward B compartment is

793 shared by BWS (chr2:231000000-231500000) and LOS (chr2:118000000-118500000)
794 groups. These shared changes of chromosome compartment and TAD between BWS
795 and LOS will be worthwhile for further detailed investigation and have the potential to
796 serve as molecular biomarkers for this syndrome.

797 The differential compartments identified in BWS and LOS show preference for
798 conservation on certain chromosomes, and this preference indicates these
799 chromosomes have some similar features that lead to vulnerabilities. As discussed
800 before, the territories of chromosome could be a source of vulnerability for molecular
801 disruptions. Among the human chromosomes that have the highest number of
802 conserved differential compartments in BWS, chr6, 10, 9, and 7 have relatively close
803 chromosome territories according to the human fibroblast 3D-FISH image [87].
804 Elucidating whether the corresponding chromosomes in bovine have such spatial
805 closeness requires future studies on bovine chromosome territories.

806 **Conclusion**

807 In summary, our results suggest that genome-wide alternation of chromosome
808 architecture, which is partially caused by DNA methylation changes, also contribute to
809 the development of BWS and LOS. We show shared regions of dysregulation between
810 BWS and LOS, including several *HOX* gene clusters, and also demonstrate that altered
811 DNA methylation differs between the clinically epigenetically identified BWS patients
812 and those identified as having DNA variants (i.e. *CDKN1C* microdeletion). Importantly,
813 we demonstrate, for the first time, that subpopulations of IC2 LOM patients exists and
814 can be further separated molecularly into two subpopulations. We also highlight

815 additional genes and genomic regions that have the potential to serve as targets for
816 biomarker development to improve current molecular methodologies. With the advent of
817 personalized medicine, the bovine model of overgrowth will be invaluable to be able to
818 perform animal and in vitro studies, such as CRISPR/Cas9-based epigenetic editing, of
819 regions identified as altered, in order to develop methodologies to ameliorate the clinical
820 features and treat BWS.

821 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

822 **Human patients**

823 Patient skin fibroblast samples and clinical information were collected through
824 the BWS Registry, under the oversight of the Children's Hospital of Philadelphia (CHOP)
825 Institutional Review Board protocol (IRB 13-010658) and in accordance with the
826 Declaration of Helsinki. In brief, consent was obtained from all patients and/or legal
827 guardians to collect longitudinal clinical information, in addition to samples that became
828 available through clinical care and publish the findings. Patients were identified from the
829 registry cohort and selected based on the availability of skin fibroblasts and their genetic
830 or epigenetic type of BWS. Non-BWS samples (control) were collected for consideration
831 of other growth difference (i.e. have an omphalocele but do not have BWS).
832 Patient skin fibroblasts were cultured as previously described [43]. Briefly, skin
833 samples collected from regions either behind the ear or abdomen were split to process
834 one chemically and other mechanically. Chemical disruption was performed using
835 collagenase and the other was mechanically minced using a scalpel blade. Both
836 explants were seeded into a T25 flask and supplemented with RPMI skin media (RPMI

837 with fetal bovine serum, penicillin–streptomycin antibiotic, and a final concentration of 2
838 mM L-glutamine). All flasks were incubated at 37°C for up to one month, with periodic
839 media changes. Successful explant cultures were passaged for sustained growth and
840 trypsinized to freeze and stored in liquid nitrogen. Clinical testing for BWS was
841 performed at the University of Pennsylvania Genetic Diagnostic Laboratory as
842 previously described [103].

843 The identifier and sex of patient included in this study were as follows; 1)
844 control: CON_H_#1 to #3 (female), CON_H_#4 and #5 (male); 2) BWS with IC2 loss of
845 methylation: BWS_#1_IC2 to BWS_#6_IC2 (female), BWS_#7_IC2 to BWS_#12_IC2
846 (male); 3) BWS with IC1 gain of methylation: BWS_#13_IC1 to BWS_#15_IC1 (female);
847 4) BWS with *CDKN1C* variant: BWS_#16_Mut (male according to record but female
848 according to sequencing results), BWS_#17_Mut and BWS_#18_Mut (female). The age
849 of fibroblast sample collection (in days with gestational age correction) for each group
850 are control (average 139, range 1-505), BWS with IC2 loss of methylation (327, 1-
851 1283), BWS with IC1 gain of methylation (1274, 116-3459), BWS with *CDKN1C* variant
852 (206, 6-488).

853 **Bovine animals**

854 Day 105 *Bos taurus indicus* (*B. t. indicus*; Brahman breed) x *Bos taurus taurus*
855 (*B. t. taurus*; Angus breed) F1 hybrid fetuses were generated by our laboratory in 2019
856 and used as tissue donors [19,44]. This breeding strategy introduced genetic variants
857 between the maternal and paternal alleles in the for allele-specific analyses. The control
858 group was generated using artificial insemination (AI) and the ART group was generated

859 by *in vitro* production procedures. The LOS group was defined as individuals from the
860 ART group with body weight greater than 97th centile of controls (548g). Conceptuses
861 were collected by caesarean section to maintain nucleic acid integrity. Fetal skin
862 fibroblast primary cell lines were established and maintained as described in our
863 previous study [44]. The identifier, original ID, sex, and body weight of fetuses used in
864 this study were as follows; 1) control fetuses: CON_B_#1 (original ID 533, female,
865 388g), CON_B_#2 (647, female, 396g), CON_B_#3 (640, male, 448g), CON_B_#4
866 (648, male, 466g), and CON_B_#5 (527, male, 514g); 2) LOS fetuses: LOS_#1 (616,
867 female, 638g), LOS_#2 (656, female, 704g), LOS_#3 (604B, female, 986g), LOS_#4
868 (512, male, 648g), LOS_#5 (602, male, 752g), and LOS_#6 (664, male, 1080g).
869 All the animal procedures were approved by the University of Missouri Animal
870 Care and Use Committee under protocol 9455. Trained personnel and Veterinarians
871 performed all animal handling and surgeries.
872
873 **METHOD DETAILS**
874 In this manuscript, all the chromosomal coordinates for bovine genome refers to
875 assembly ARS-UCD1.2, and for human genome refers to GRCh37 (hg19) [104].

876 **Genomic DNA extraction**

877 Genomic DNA from human skin fibroblasts was isolated using the AllPrep
878 DNA/RNA Micro Kit (QIAGEN) per the manufacturer's instructions, as previously
879 described [43]. Briefly, the cells were lysed in the Buffer RLT Plus. The lysate was then
880 homogenized into a QIAshredder spin column. The homogenized lysate was then

881 transferred to an AllPrep DNA spin column, washed with washing buffer and then DNA
882 was eluted using elution buffer. The concentration of DNA was measured by using a
883 NanoDrop® ND-1000 Spectrophotometer (Thermo Fisher Scientific). Genomic DNA
884 samples were stored at -20°C.

885 Bovine fibroblast cells were lysed in lysis buffer (0.05 M Tris-HCl (pH 8.0), 0.1 M
886 EDTA, and 0.5% (w/v) SDS) with proteinase K (Fisher BioReagents, BP1700) at 55°C
887 for four hours. Genomic DNA was extracted with Phenol:Chloroform:Isoamyl Alcohol
888 (SIGMA, P3803) following the manufacturer's instructions. The concentration of DNA
889 was measured by using a NanoDrop® ND-1000 Spectrophotometer and DNA integrity
890 was confirmed by electrophoresis on a 0.7% agarose gel. Genomic DNA samples were
891 stored at -20°C.

892 **RNA isolation**

893 Total RNA from human skin fibroblasts was isolated using the AllPrep DNA/RNA
894 Micro Kit (QIAGEN) per the manufacturer's instructions, as previously described [43].
895 Briefly, the cells were lysed in the Buffer RLT Plus. The lysate was then homogenized
896 into a QIAshredder spin column. The homogenized lysate was then transferred to an
897 AllPrep DNA spin column, and the flowthrough was used for RNA extraction using 70%
898 ethanol. The concentration of RNA was measured by using a NanoDrop® ND-1000
899 Spectrophotometer and RNA samples were stored at -80°C.

900 Total RNA was isolated from bovine fibroblast cells using TRIzol™ Reagent
901 (Invitrogen, 15596026) following the manufacturer's instructions. The concentration of
902 RNA was measured by using a NanoDrop® ND-1000 Spectrophotometer. RNA samples

903 were stored at -80°C.

904 **CTCF binding site prediction**

905 Potential CTCF binding sites were predicted globally for human and bovine
906 genome using TFBSTools 1.26.0 [105] with database JASPAR2020 [106]. CTCF motifs
907 of 'vertebrates' were used for prediction and min.score was set to 81%.

908 **Hi-C sequencing and data analyses**

909 Hi-C sequencing for fibroblast cells was conducted by CD Genomics.
910 Information on library preparation and sequencing obtained from the company is as
911 follows: Cells were cross-linked with 1% formaldehyde at room temperature and
912 quenched with glycine. The cross-linked cells were subsequently lysed. Endogenous
913 nucleases were inactivated with 0.3% SDS, then chromatin DNA were digested by MboI
914 (NEB), marked with biotin-14-dCTP (Invitrogen), and ligated by T4 DNA ligase (NEB).
915 After reversing cross-links, the ligated DNA was extracted through QIAamp DNA Mini Kit
916 (Qiagen) according to manufacturers' instructions. Purified DNA was sheared to 300-
917 500 bp fragments and were further blunt-end repaired, A-tailed and adaptor-added,
918 followed by purification and PCR amplification. Finally, equimolar pooling of libraries
919 was performed based on QC values and sequenced on Novaseq6000 sequencer
920 (Illumina) with a read length configuration of 150 PE for 1,000M PE reads per sample
921 (1,000M in each direction).
922 The processing of Hi-C raw sequencing reads was based on the 'Hi-C
923 Processing Pipeline' from 4D Nucleome Data Portal

924 (https://data.4dnucleome.org/resources/data-analysis/hi_c-processing-pipeline). Briefly,
925 reads were aligned to the corresponding reference genome using bwa 0.7.17 with
926 parameters 'mem -SP5M' [107]. Reads with mapping quality (MAPQ) equal or greater
927 than 30 were kept for analyses. Then, pairtools 0.3.0 was used to parse, sort, and
928 deduplicate aligned reads, and pair types 'UU', 'UR', and 'RU' were selected [108].
929 Cooler and Juicer were used to convert contact matrix to .cool and .hic files [109,110].
930 Hi-C data was also used to identify genomic variants following the pipeline for
931 1000 bull genome project [111]. Briefly, Read groups were added using
932 AddOrReplaceReadGroups function of picard 2.25.5 [112] for aligned Hi-C sequencing
933 reads. Duplicated reads were marked using MarkDuplicates function of picard with
934 parameter '--OPTICAL_DUPLICATE_PIXEL_DISTANCE 2500'. For human, the dataset
935 of known variants was acquired from GATK resource bundle
936 (https://gatk.broadinstitute.org/hc/en-us/articles/360035890811-Resource-bundle). For
937 bovine, the dataset of known variants was acquired from the 1000 bull genome project,
938 namely ARS1.2PlusY_BQSR_v3.vcf.gz. GATK 4.2.1.0 [113] was used to recalibrate
939 base quality and identify genomic variants in the Hi-C data with the known variant
940 dataset as reference. Parameters used for BaseRecalibrator were '–bqsr-baq-gap-
941 open-penalty 45 --read-filter MappingQualityReadFilter --minimum-mapping-quality 20'
942 for bovine. For human, the same settings were used except with default value for --bqsr-
943 baq-gap-open-penalty. In addition, for bovine, our previously generated genomic
944 sequencing data for semen DNA of the bull used to sire all the fetuses in this study was
945 included for joint genotyping in order to assign alleles for the genomic variants (GEO:
946 GSE197130) [44]. HaplotypeCaller was used with parameter '-ERC GVCF', and

947 GenomicsDBImport and GenotypeGVCFs were used with default settings. For bovine,
948 raw SNP and INDEL were separately scored using VariantRecalibrator with parameter '-
949 -resource:1000G,known=false,training=true,truth=true,prior=10.0 known_variant -an QD
950 -an MQ -an MQRankSum -an ReadPosRankSum -an FS -an SOR' was used. For
951 human, parameter '--resource:hapmap,known=false,training=true,truth=true,prior=15.0
952 known_variant1 --resource:omni,known=false,training=true,truth=true,prior=12.0
953 known_variant2 --resource:1000GSNP,known=false,training=true,truth=true,prior=10.0
954 known_variant3 --
955 resource:1000GINDEL,known=false,training=true,truth=true,prior=10.0 known_variant4
956 --resource:dbsnp,known=true,training=false,truth=false,prior=2.0 known_variant5 -an
957 QD -an MQ -an MQRankSum -an ReadPosRankSum -an FS -an SOR' was used.
958 Scored variants were filtered using ApplyVQSR with parameter '--truth-sensitivity-filter-
959 level 99.0'.
960 For bovine, variants used for allele assignment met the following criteria: 1)
961 marked as PASS by ApplyVQSR, 2) GQ>=40 in the bull and the fetus, 3) DP>=10 in the
962 bull, 4) no read supports not assigned alleles in the bull and the fetus, 5) only one allele
963 of the bull is shown in the fetus. In addition, variants on chrX identified in male fetuses
964 were excluded from analyses, since which is likely due to the lack of chrY in the
965 reference genome. Alleles were assigned for bovine Hi-C reads based on qualified
966 variants using custom Perl scripts. For SNP, the nucleotide in the reads at the SNP
967 location was directly used for allele assignment. For INDEL, local maternal and paternal
968 allele pseudo genomes were generated, and the reads were realigned to both pseudo
969 genomes. The reads were assigned to the alleles with more aligned bases, or less

970 mismatches if aligned base numbers were the same. Next, variants with biased allelic
971 ratio of Hi-C reads (one allele is less than 15%) were excluded from analyses. Lastly,
972 reads that cover multiple variants and have conflicting allele assignments were
973 excluded from the analyses.

974 Statistical analyses for detecting differentially interacting regions (DIR),
975 differential TAD, and differential chromosome compartments were performed. The sex
976 chromosomes were only analyzed in sex-specific comparisons to circumvent
977 confounding created by X chromosome inactivation. R package multiHiCcompare
978 1.12.0 was used to identify DIR at 1mb, 500kb, 250kb, 100kb, and 50kb resolutions
979 between groups [114]. The make_hicexp function was used with default parameters
980 except 'remove.regions = NULL', and cyclic_loess and hic_exactTest functions were
981 used with default settings. False discovery rate (FDR) was controlled at 0.05 for
982 significance.

983 For visualization and TAD detection for each group, the size of contact matrix
984 was normalized to the first sample in the Control group for both human (CON_H_#1)
985 and bovine (CON_B_#1), and group mean contact matrix was calculated based on
986 normalized matrix of each sample in the group. R package SpectralTAD 1.10.0 was
987 used with parameter 'qual_filter = TRUE, z_clust = FALSE, levels = level, resolution =
988 25000, window_size = window_size, gap_thresh = 0.2, min_size = 6' to identify TAD
989 structure ranging from 150kb to 4mb at 25kb resolution [115]. Level and window_size
990 was initially set to 10 and 120, respectively, and reduced by 1 and 30 each time if error
991 reported. In addition, diffDomain was used with function and parameter 'dvsd multiple --
992 min_nbin 1 --hicnorm NONE' and 'adjustment fdr_bh' to detect differential TAD

993 boundaries at 25kb resolution between groups based on TAD identified by SpectralTAD
994 [116]. Since SpectralTAD and diffDomain only support single sample comparison
995 instead of group comparison, to reduce the impacts of sample variations within each
996 group and obtain more consistent differences between groups, filtering for inconsistent
997 TADs in the control groups was performed for detected differential TADs. Inconsistent
998 TADs were obtained by applying diffDomain between each sample and its group mean
999 contact matrix and merging all the differential results from each sample. It is worth
1000 noting that TAD comparisons between BWS subtypes were not filtered and resulted in
1001 much larger number of differential TAD since there was no control group involved.
1002 Types of differential TAD include loss, strength change, merge, split, zoom, and
1003 complex. Strength change refers to the situation that the same TAD is detected in the
1004 two groups for comparing but the chromosomal contact frequency is different within the
1005 TAD. TAD with strength change could reflect forming or losing small sub-TADs and
1006 loops within it that requires higher resolution and cannot be properly annotated with the
1007 sequencing depth in this study. TAD with strength change could also be a result of
1008 chromosome compartment change of the sub-TADs within it. Zoom refers to the
1009 extension or shrink of a TAD in another group. Complex refers to the situation that
1010 cannot be simply categorized into any other type.
1011 dcHiC was used to assign chromosome compartments and identify differential
1012 compartments at 1mb, 500kb, 250kb, 100kb, 50kb, and 25kb resolutions between
1013 groups [117]. Default settings for cis, select, analyze, and subcomp functions of dcHiC
1014 were used, and FDR was controlled at 0.05 for significance during the analyze step.
1015 The quantile normalized principal component (eigenvalue) calculated at 25kb resolution

1016 was used for assigning compartments and plotting.

1017 **Whole genome bisulfite sequencing (WGBS) and data analyses**

1018 WGBS for fibroblast cells was conducted by CD Genomics. Information on
1019 library preparation and sequencing obtained from the company is as follows: Genomic
1020 DNA was fragmented by sonication to a mean size of approximately 200-400 bp.
1021 Fragmented DNA was end-repaired, 5'-phosphorylated, 3'-dA-tailed and then ligated to
1022 methylated adapters. The methylated adapter-ligated DNAs were purified using 0.8×
1023 Agencourt AMPure XP magnetic beads and subjected to bisulfite conversion by ZYMO
1024 EZ DNA Methylation-Gold Kit (Zymo Research). The converted DNAs then have the
1025 addition of 8-bp index primers to create the final cDNA library. Equimolar pooling of
1026 libraries was performed based on QC values and sequenced on Novaseq6000
1027 sequencer (Illumina) with a read length configuration of 150 PE for 300M PE reads per
1028 sample (300M in each direction). 0.1-1% lambda DNA were added during the library
1029 preparation to monitor bisulfite conversion rate.

1030 For WGBS data, raw sequencing reads were deduplicated using the clumpify
1031 function of BBMap 38.90 [118] and trimmed for adapter sequences and low quality
1032 bases using trimmomatic 0.39 [119] with parameters
1033 'ILLUMINACLIP:adapter_seq:2:30:10:1:true LEADING:20 TRAILING:20 AVGQUAL:20
1034 MAXINFO:0:0.5'. Trimmed reads were aligned to the corresponding reference genome
1035 using BSBolt 1.5.0 [120] with default setting. Trimmed reads were also aligned to
1036 lambda phage genome to determine bisulfite conversion rates. Samtools 1.13 [121] was
1037 used to convert, sort, filter, and index bam files. MarkDuplicates function of picard

1038 2.25.5 [112] was used with parameter '--OPTICAL_DUPLICATE_PIXEL_DISTANCE
1039 2500' to further remove duplicated reads after alignment. CpG methylation information
1040 was extracted from the bam files using the function CallMethylation of BSBolt with
1041 parameter '-CG -min 1 -BQ 20'. For both human and bovine samples, CpG methylation
1042 at C/T SNPs detected in their corresponding Hi-C data were excluded for sense strand
1043 reads, and CpG methylation at G/A SNPs were excluded for antisense strand reads. For
1044 bovine samples, SNPs identified from Hi-C dataset were also used to assign alleles for
1045 WGBS reads using custom Perl scripts. C/T SNPs (based on sense strand sequence)
1046 were only applied for antisense strand reads, and G/A SNPs were only applied for
1047 sense strand reads. Other types of SNP were applied for both strands. Reads that cover
1048 multiple SNPs and have conflicting allele assignments were excluded. For both overall
1049 and allele-specific comparisons, statistical analyses were conducted using R package
1050 hummingbird [122] with parameter 'minCpGs = 10, minLength = 100, maxGap = 300' to
1051 identify differentially methylated regions (DMRs) between groups. DMRs with at least
1052 15% difference in methylation level (both gain and loss of methylation) and at least 4
1053 mean read coverage at CpG sites in all the samples were reported. For allele-specific
1054 analyses, the read coverage filtering was not applied since only about 20% of the
1055 usable SNPs are shared by all the sample. The sex chromosomes were only analyzed
1056 in sex-specific comparisons to circumvent confounding created by X chromosome
1057 inactivation.

1058 **Total RNA sequencing (total RNA-seq) and data analyses**

1059 Total RNA-seq for fibroblast cells was conducted by CD Genomics. Information

1060 on library preparation and sequencing obtained from the company is as follows: The
1061 first step involves the removal of ribosomal RNA using RiboZero kit (Illumina).
1062 Subsequently, the RNA is fragmented into small pieces using divalent cations under
1063 elevated temperatures. The cleaved RNA fragments are copied into first strand cDNA
1064 using reverse transcriptase and random primers, followed by second strand cDNA
1065 synthesis. These cDNA fragments then have the addition of a single 'A' base and
1066 subsequent ligation of the adapter. The products are purified and enriched with PCR to
1067 create the final cDNA library. This results in libraries with inserts ranging in size from
1068 120–200 bp with a median size of 150 bp. Equimolar pooling of libraries was performed
1069 based on QC values and sequenced on Novaseq6000 sequencer (Illumina) with a read
1070 length configuration of 150 PE for 33M PE reads per sample (33M in each direction).
1071 All read data was assessed for quality with FastQC 0.11.7 and low-quality
1072 bases (Phred scores ≤ 20) were trimmed using the dynamictrim function of SolexaQA++
1073 3.1.7.1. Reads less than 60 bases in length were removed using SolexaQA++
1074 LengthSort. Known splice sites were extracted from the gtf file of each reference
1075 genome using the Python script hisat2_extract_splice_sites.py. Hisat2 2.1.0 was used
1076 to establish a genome index with the hisat2-build and also to align the quality trimmed
1077 reads to the reference genome with the following adjusted parameters: --mp 6,6; –
1078 score-min L,0,-0.2; known-splicesite-infile. Sam files were converted to bam files and
1079 coordinate-sorted using Samtools 1.8. Then, Stringtie 2.1.4 was used with the sorted
1080 bam files to assemble transcriptomes for each sample. Stringtie merge was used to
1081 combine the assembled transcripts from all samples into a reference assembly.
1082 Assembled transcripts were extracted using the gffread module in cufflinks 2.2.1. An

1083 index was built for the merged transcriptome and reads were quantified using Salmon
1084 0.11.1. IsoformSwitchAnalyeR 1.17.04 was used to rescue stringtie annotations and
1085 generate a gene count matrix.

1086 For alternative splicing (AS) event analyses, the merged Stringtie annotation file
1087 and coordinate-sorted bam files were used, and AS events were called using rMATs
1088 4.0.1 with parameters '-t paired –readlength 150 -cstat 0.0001 -libType fr-firststrand'.
1089 FDR was controlled at 0.05 for significance. Manhattan plots of were generated using R
1090 package Sushi [123].

1091 To assign allele specificity for total RNA-seq reads in bovine, variants were
1092 called from the aligned data according to GATK's Best Practices for Calling Variants in
1093 RNA-seq data. Duplicate reads were removed and read groups were assigned using
1094 Picard 2.22.4. A dictionary was created for the reference genome using GATK's
1095 CreateSequenceDictionary. GATK SplitNCigarReads was used to hard clip sequences
1096 extending into intronic regions and GATK BaseRecalibrator to recalibrate base quality
1097 scores. Then, HaplotypeCaller was run on calibrated data and called variants were
1098 filtered with adjusted parameters (AB < 0.2 & MQ0 > 50). Variants that were shared
1099 between genomic and RNA-seq data for each sample were kept. Variants that were
1100 heterozygous in the fetus and homozygous in the bull were used to assign alleles. A
1101 variant annotation file was created using the annotate function of bcftools 1.8 and gene
1102 annotations were added using custom scripts. Custom R scripts were used to retrieve
1103 maternal and paternal allele depth read counts from VCF files in order to quantify allele-
1104 specific expressions.

1105 Statistical analyses were conducted using R package edgeR 3.36.0 [124] to

1106 identify differentially expressed genes (DEGs) between groups. Parameters used
1107 include ‘method = "TMM"’ for calcNormFactors function and ‘robust = TRUE’ for glmFit
1108 function. FDR was controlled at 0.05 for significance.

1109 **Functional correlation between gene expression and BWS clinical diagnosis**

1110 To determine the co-expressional network that gathered by genes showing
1111 similar trends across samples, WGCNA was performed in R [125]. Briefly, gene
1112 abundance matrix (normalized expression matrix) from IsoformSwitchAnalyeR was log2
1113 transformed and only genes with high expressional variability values (top 60%, 15304
1114 genes retained) were remained for identification of co-expression network and a soft
1115 threshold β value was set as 10 according to the criterion of approximate scale-free
1116 topology. BWS specific clinical diagnosis was generated through the evaluated features
1117 from Wang et al reported study [83]. Clinical BWS score consists of the combination of
1118 cardinal and suggestive features, with one more point (plus two points) was given for
1119 each diagnosed cardinal feature, while one point was given for each diagnosed
1120 suggestive feature. The correlation between eigengene (a representative of the gene
1121 expression profiles in a module) in each module and clinical diagnoses was calculated
1122 using Pearson’s correlation method, and only relationships with p-value < 0.05 were
1123 considered as significantly positive/negative correlation. Hub genes within BWS-related
1124 module were identified with both absolute gene-BWS score significance ≥ 0.5 and
1125 absolute module membership ≥ 0.8.

1126 **Correlation analyses between gene expression and chromosome compartment**

1127 Pearson correlation analyses between gene expression changes and
1128 eigenvalue changes at gene promoter were performed using the R base function 'cor'.

1129 **Genomic content overlap and enrichment analyses**

1130 Gene annotation information was obtained from NCBI
1131 (GCF_002263795.1_ARS-UCD1.2_genomic.gtf and
1132 GCF_000001405.25_GRCh37.p13_genomic.gtf) [126]. Repeated and overlapped
1133 exons were merged for each gene, and introns were calculated based on merged
1134 exons. Promoters (1kb) were calculated based on transcription start sites annotation
1135 and only included protein coding genes and long non-coding RNAs. Annotation of CpG
1136 islands and repeated sequences were obtained from UCSC Genome Browser [127].
1137 Locations of CpG shores (flanking 2kb from CpG islands) and shelves (flanking 2-4kb
1138 from the CpG island) were calculated based on CpG island annotation. Human public
1139 CTCF ChIP-seq data of fibroblast primary cells from different tissues were obtained
1140 from the ENCODE portal with the following identifiers: ENCFF148VQH (lung),
1141 ENCFF337WIE (foreskin), ENCFF344UBE (cardiac), ENCFF438XHB (mammary
1142 gland), ENCFF570FLB (lung), ENCFF638OUB (dermis), ENCFF652MLT (foreskin),
1143 ENCFF734DZF (lung), ENCFF738CXX (villous mesenchyme), ENCFF859PRV (aortic
1144 adventitia), and ENCFF933VBD (pulmonary artery) [128]. Peaks from these files were
1145 merged to use as common CTCF sites. Bovine public CTCF ChIP-seq data of different
1146 tissues (no fibroblast primary cell data available) were obtained from the Gene
1147 Expression Omnibus (GEO) database with the following identifiers: GSE158430
1148 (replicate for adipose, cerebellum, cortex, hypothalamus, liver, lung, muscle, and

1149 spleen) and GSE129423 (control rumen epithelial primary cells) [129]. GSE129423 was
1150 based on genome build UMD3.1.1 and was converted to ARS-UCD1.2 using liftOver
1151 from UCSC Genome Browser with parameter '-minMatch=0.1'. Due to lower quality (i.e.
1152 occasionally very broad peaks) of these data than the human data, peak regions found
1153 in at least four files were kept as common CTCF sites. Bedtools and custom Perl scripts
1154 were used to identify overlapped genomic contexts for regions with altered DNA
1155 methylation, gene expression, and chromosome architecture, and calculate observed
1156 frequencies (Obs) [130]. Permutation tests were performed to determine expected
1157 frequencies (Exp) of overlapping by randomly shuffling genomic contexts across the
1158 genome for 10,000 times. For each test, the p value was calculated as p = n(|Exp -
1159 mean(Exp)| ≥ |Obs - mean(Exp)|)/10,000. EnrichedHeatmap was also used for
1160 enrichment visualization in the heatmaps and different flanking distance was chosen
1161 based on the length of genomic content examined to best show the background level
1162 and enrichment pattern [131].
1163 The genomic coordinates of human super-enhancers were obtained from
1164 database dbSUPER and used for finding enhancer-promoter contact for gene *SESN3* at
1165 10kb resolution with Python library bioframe [132,133].

1166 **Conservation analyses**

1167 For TAD/compartment conservation analyses between human and bovine
1168 control groups, the genomic coordinates of human TAD/compartment were converted to
1169 bovine coordinates using liftOver from UCSC Genome Browser with parameter '-
1170 minMatch=0.1'. Then, custom Perl scripts were used to identify bovine

TAD/compartment overlapping with human TAD/compartment that the overlapped region is greater than 90% in both species for TAD and in any species for compartment.

To identify shared differential compartments between BWS and LOS, for each comparison, bedtools was first used to merge significant results of different resolutions into non-overlapping regions [130]. Merged results of comparisons that compare BWS/LOS to control group were combined for each species. For human, BWS vs Control (index 1 in Table 1), BWS_IC1_GOM vs Control (2), BWS_IC2_LOM vs Control (3), BWS_CDKN1C_Mut vs Control (4), BWS_IC2_LOM_Male vs Control_Male (8), BWS_Female vs Control_Female (9), BWS_IC1_GOM_Female vs Control_Female (10), BWS_IC2_LOM_Female vs Control_Female (11), BWS_CDKN1C_Mut_Female vs Control_Female (12) were used for analyses. For bovine, LOS vs Control (16), LOS_Male vs Control_Male (18), and LOS_Female vs Control_Female (19) were used for analyses. Then human coordinates were converted to bovine using liftOver and regions overlapped with bovine differential compartments were identified using custom Perl scripts.

Plot generation

R package ComplexHeatmap, EnrichedHeatmap, plotgardener, circlize, methylKit, ggpubr, ggfortify, ggplot2, and Python library matplotlib were used for making plots [131,134–141].

Data availability

All the raw sequencing data for human samples, including whole genome

1193 bisulfite sequencing, total RNA sequencing, and Hi-C sequencing, are available in the
1194 database of Genotypes and Phenotypes (dbGaP) with accession number
1195 (phs001794.v1.p1). The subject ID of each sample used in dbGaP can be found in
1196 Table S8. Some of the patients in this study were also included in our previous study
1197 and deposited under the same dbGaP subject ID [26]. All the raw sequencing data for
1198 bovine samples, including whole genome bisulfite sequencing, total RNA sequencing,
1199 and Hi-C sequencing, are available in the Gene Expression Omnibus (GEO) database
1200 with accession number (pending).
1201

1202 **Acknowledgements**

1203 We acknowledge the ENCODE Consortium and the John Stamatoyannopoulos
1204 (UW) and Bradley Bernstein (Broad) laboratories that generated the human CTCF
1205 ChIP-seq datasets. This research was partially funded by the United States Department
1206 of Agriculture-Agriculture and Food Research Initiative (USDA-AFRI), grant 2018-
1207 67015-27598. This work was partially supported by National Cancer Institute [K08
1208 CA193915], Alex's Lemonade Stand Foundation, a Damon Runyon Clinical Investigator
1209 Award provided by the Damon Runyon Cancer Research Foundation (105–19), the
1210 Lorenzo "Turtle" Sartini Jr. Endowed Chair in Beckwith-Wiedemann Syndrome
1211 Research, and the Victoria Fertitta Fund through the Lorenzo "Turtle" Sartini Jr.
1212 Endowed Chair in Beckwith-Wiedemann Syndrome Research (J.M.K.). The funders had
1213 no role in study design, data collection and analysis, decision to publish, or preparation
1214 of the manuscript.
1215

Author contributions

Y.L., J.M.K, D.E.H., J.C., and R.M.R. designed research; Y.L. and S.N. performed research; Y.L., P.X., F.B., A.K.G., analyzed data; Y.L., P.X., F.B., A.K.G., S.N., and R.M.R. wrote the paper; all authors edited the paper.

Declaration of interests

The authors declare no competing interests.

Authors' ORCID

Yahan Li, yl5c7@mail.missouri.edu, ORCID 0000-0003-3062-4626
Ping Xiao, ping.xiao@okstate.edu, ORCID 0000-0001-5678-501X
Frimpong Boadu, fbqc9@missouri.edu, ORCID 0000-0002-4464-6191
Anna Goldkamp, anna.goldkamp@okstate.edu, ORCID 0000-0002-1510-465X
Snehal Nirgude, nirgudes@chop.edu, ORCID 0000-0001-8471-1384
Jianlin Cheng, chengji@missouri.edu, ORCID 0000-0003-0305-2853
Darren Hagen, darren.hagen@okstate.edu, ORCID 0000-0001-8295-020X
Jennifer Kalish, kalishj@chop.edu, ORCID 0000-0003-1500-9713
Rocío Melissa Rivera, riverarm@missouri.edu, ORCID 0000-0002-9832-9618

References

1. Rivera RM, Donnelly CG, Patel BN, Li Y, Soto-Moreno EJ. Abnormal Offspring Syndrome. Bovine Reproduction. 2021; 876–895.

2. Brioude F, Kalish JM, Mussa A, Foster AC, Bliek J, Ferrero GB, et al. Clinical and molecular diagnosis, screening and management of Beckwith–Wiedemann syndrome: an international consensus statement. Nature Reviews Endocrinology. 2018;14: 229–249.

3. Farin PW, Piedrahita JA, Farin CE. Errors in development of fetuses and placentas from in vitro-produced bovine embryos. Theriogenology. 2006;65: 178–91. doi:10.1016/j.theriogenology.2005.09.022

4. Li Y, Donnelly CG, Rivera RM. Overgrowth syndrome. Veterinary Clinics: Food Animal Practice. 2019;35: 265–276.

5. Mussa A, Russo S, De Crescenzo A, Chiesa N, Molinatto C, Selicorni A, et al. Prevalence of Beckwith–Wiedemann syndrome in north west of Italy. American journal of medical genetics Part A. 2013;161: 2481–2486.

6. Li Y, Sena Lopes J, Fuster PC, Rivera RM. Spontaneous and ART-induced large offspring syndrome: similarities and differences in DNA methylome. Epigenetics. 2022; 1–20.

7. Chen Z, Robbins KM, Wells KD, Rivera RM. Large offspring syndrome: a bovine model for the human loss-of-imprinting overgrowth syndrome Beckwith-Wiedemann. Epigenetics. 2013;8: 591–601. doi:10.4161/epi.24655

8. Hori N, Nagai M, Hirayama M, Hirai T, Matsuda K, Hayashi M, et al. 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. 2010;122: 303–312.

9. Lazzari G, Wrenzycki C, Herrmann D, Duchi R, Kruip T, Niemann H, et al. Cellular and molecular deviations in bovine in vitro-produced embryos are related to the large offspring syndrome. Biology of reproduction. 2002;67: 767–75.

10. Behboodi E, Anderson GB, BonDurant RH, Cargill SL, Kreuscher BR, Medrano JF, et al. Birth of large calves that developed from in vitro-derived bovine embryos. Theriogenology. 1995;44: 227–32.

11. Farin PW, Farin CE. Transfer of bovine embryos produced in vivo or in vitro: survival and fetal development. Biology of reproduction. 1995;52: 676–682.

12. McEvoy TG, Sinclair KD, Broadbent PJ, Goodhand KL, Robinson JJ. Post-natal growth and development of Simmental calves derived from in vivo or in vitro embryos. Reproduction, Fertility and Development. 1998;10: 459–464.

13. Farin PW, Crosier AE, Farin CE. Influence of in vitro systems on embryo survival and fetal development in cattle. Theriogenology. 2001;55: 151–170.

14. Hasler JF, Henderson WB, Hurtgen PJ, Jin ZQ, McCauley AD, Mower SA, et al. Production, freezing and transfer of bovine IVF embryos and subsequent calving results. Theriogenology. 1995;43: 141–152.

15. van Wagtendonk-de Leeuw AM, Aerts BJ, den Daas JH. Abnormal offspring following in vitro production of bovine preimplantation embryos: a field study. Theriogenology. 1998;49: 883–94.

16. Sinclair KD, Broadbent PJ, Dolman DF. In vitro produced embryos as a means of achieving pregnancy and improving productivity in beef cows. Animal Science. 1995;60: 55–64.

17. Kruip TA, Den Daas JHG. In vitro produced and cloned embryos: effects on pregnancy, parturition and offspring. Theriogenology. 1997;47: 43–52.

18. van Wagtendonk-de Leeuw AM, Mullaart E, de Roos AP, Merton JS, den Daas JH, Kemp B, et al. Effects of different reproduction techniques: AI MOET or IVP, on health and welfare of bovine offspring. Theriogenology. 2000;53: 575–97.

19. Rivera RM, Goldkamp AK, Patel BN, Hagen DE, Soto-Moreno EJ, Li Y, et al. Identification of large offspring syndrome during pregnancy through ultrasonography and maternal blood transcriptome analyses. Scientific Reports. 2022;12: 10540. doi:10.1038/s41598-022-14597-w

20. Nava-Trujillo H, Rivera RM. Large offspring syndrome in ruminants: current status and prediction during pregnancy. animal. 2023;17: 100740.

21. Chen Z, Hagen DE, Ji T, Elsik CG, Rivera RM. Global misregulation of genes largely uncoupled to DNA methylome epimutations characterizes a congenital overgrowth syndrome. Scientific reports. 2017;7: 12667.

22. Chen Z, Hagen DE, Elsik CG, Ji T, Morris CJ, Moon LE, et al. Characterization of global loss of imprinting in fetal overgrowth syndrome induced by assisted reproduction. Proc Natl Acad Sci U S A. 2015;112: 4618–23. doi:10.1073/pnas.1422088112

23. Sangalli JR, Chiaratti MR, De Bem THC, de Araujo RR, Bressan FF, Sampaio RV, et al. Development to term of cloned cattle derived from donor cells treated with valproic acid. PLoS One. 2014;9: e101022.

24. Su J, Wang Y, Liu Q, Yang B, Wu Y, Luo Y, et al. Aberrant mRNA expression and DNA methylation levels of imprinted genes in cloned transgenic calves that died of large offspring syndrome. Livestock Science. 2011;141: 24–35.

25. Su H, Li D, Hou X, Tan B, Hu J, Zhang C, et al. Molecular structure of bovine Gtl2 gene and DNA methylation status of Dlk1-Gtl2 imprinted domain in cloned bovines. Animal reproduction science. 2011;127: 23–30.

26. Li Y, Hagen DE, Ji T, Bakhtiarizadeh MR, Frederic WM, Traxler EM, et al. Altered microRNA expression profiles in large offspring syndrome and Beckwith-Wiedemann syndrome. Epigenetics. 2019; 1–27.

27. Krzyzewska IM, Alders M, Maas SM, Bliek J, Venema A, Henneman P, et al. Genome-wide methylation profiling of Beckwith-Wiedemann syndrome patients without molecular confirmation after routine diagnostics. Clinical epigenetics. 2019;11: 53.

28. Tee L, Lim DH, Dias RP, Baudement M-O, Slater AA, Kirby G, et al. Epimutation profiling in Beckwith-Wiedemann syndrome: relationship with assisted reproductive technology. Clinical epigenetics. 2013;5: 1–10.

29. Tenorio J, Romanelli V, Martin-Trujillo A, Fernández G-M, Segovia M, Perandones C, et al. Clinical and molecular analyses of Beckwith–Wiedemann syndrome: Comparison between spontaneous conception and assisted reproduction techniques. American journal of medical genetics Part A. 2016;170: 2740–2749.

30. Rossignol S, Steunou V, Chalas C, Kerjean A, Rigolet M, Viegas-Pequignot E, et al. 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. 2006;43: 902–907.

31. Cooper WN, Luharia A, Evans GA, Raza H, Haire AC, Grundy R, et al. Molecular subtypes and phenotypic expression of Beckwith–Wiedemann syndrome. European journal of human genetics. 2005;13: 1025.

32. Mussa A, Russo S, De Crescenzo A, Freschi A, Calzari L, Maitz S, et al. (Epi) genotype–phenotype correlations in Beckwith–Wiedemann syndrome. European journal of human genetics. 2016;24: 183.

33. Wadhwa PD, Buss C, Entringer S, Swanson JM. Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. Seminars in reproductive medicine. © Thieme Medical Publishers; 2009. pp. 358–368.

34. Niakan KK, Han J, Pedersen RA, Simon C, Pera RAR. Human pre-implantation embryo development. Development. 2012;139: 829–841.

35. Mussa A, Molinatto C, Cerrato F, Palumbo O, Carella M, Baldassarre G, et al. Assisted Reproductive Techniques and Risk of Beckwith-Wiedemann Syndrome. Pediatrics. 2017;140. doi:10.1542/peds.2016-4311

36. Erik Mullaart, Geart Ludema, Helga Flapper, John Hepburn, Björn Oback. Birthweight data of IVP calves. 39th Annual Meeting of the Association of Embryo Technology in Europe Animal Reproduction. 2023;2. Available: https://animal-reproduction.org/article/64dfa00da953954a9b16f7a5

37. Fauser BC, Devroey P, Diedrich K, Balaban B, Bonduelle M, Delemarre-van de Waal HA, et al. Health outcomes of children born after IVF/ICSI: a review of current expert opinion and literature. Reprod Biomed Online. 2014;28: 162–82. doi:10.1016/j.rbmo.2013.10.013

38. Urrego R, Rodriguez-Osorio N, Niemann H. Epigenetic disorders and altered gene expression after use of assisted reproductive technologies in domestic cattle. Epigenetics. 2014;9: 803–815.

39. Sunderam S, Kissin DM, Zhang Y, Jewett A, Boulet SL, Warner L, et al. Assisted reproductive technology surveillance—United States, 2017. MMWR Surveillance Summaries. 2020;69: 1.

40. Joao Viana. 2020 Statistics of embryo production and transfer in domestic farm animals. Embryo Technology Newsletter. 2021;39.

41. Rivera RM, Stein P, Weaver JR, Mager J, Schultz RM, Bartolomei MS. Manipulations of mouse embryos prior to implantation result in aberrant expression of imprinted genes on day 9.5 of development. Human molecular genetics. 2008;17: 1–14.

42. Rovina D, La Vecchia M, Cortesi A, Fontana L, Pesant M, Maitz S, et al. Profound alterations of the chromatin architecture at chromosome 11p15. 5 in cells from Beckwith-Wiedemann and Silver-Russell syndromes patients. Scientific reports. 2020;10: 1–19.

43. Naveh NSS, Deegan DF, Huhn J, Traxler E, Lan Y, Weksberg R, et al. The role of CTCF in the organization of the centromeric 11p15 imprinted domain interactome. Nucleic Acids Research. 2021;49: 6315–6330.

44. Li Y, Boadu F, Highsmith MR, Hagen DE, Cheng J, Rivera RM. Allele-specific aberration of imprinted domain chromosome architecture associates with large offspring syndrome. iScience. 2022;25: 104269. doi:10.1016/j.isci.2022.104269

45. Li X, Li Y, Bai S, Zhang J, Liu Z, Yang J. NR2F1-AS1/miR-140/HK2 axis regulates hypoxia-induced glycolysis and migration in hepatocellular carcinoma. Cancer Management and Research. 2021; 427–437.

46. Wen X, Zhang Q, Zhou L, Li Z, Wei X, Yang W, et al. Intrachromosomal Looping and Histone K27 Methylation Coordinately Regulates the lncRNA H19-Fetal Mitogen IGF2 Imprinting Cluster in the Decidual Microenvironment of Early Pregnancy. Cells. 2022;11: 3130.

47. Paço A, de Bessa Garcia SA, Freitas R. Methylation in HOX clusters and its applications in cancer therapy. Cells. 2020;9: 1613.

48. Brzezinski J, Choufani S, Romao R, Shuman C, Chen H, Cunanan J, et al. Clinically and biologically relevant subgroups of Wilms tumour defined by genomic and epigenomic analyses. British Journal of Cancer. 2021;124: 437–446.

49. Mallo M, Alonso CR. The regulation of Hox gene expression during animal development. Development. 2013;140: 3951–3963.

50. Wang K, Liang Q, Li X, Tsoi H, Zhang J, Wang H, et al. MDGA2 is a novel tumour suppressor cooperating with DMAP1 in gastric cancer and is associated with disease outcome. Gut. 2016;65: 1619–1631.

51. Yang Z, Chen H, Huo L, Yang Z, Bai Y, Fan X, et al. Epigenetic inactivation and tumor-suppressor behavior of NGFR in human colorectal cancer. Molecular Cancer Research. 2015;13: 107–119.

52. Vishnubalaji R, Yue S, Alfayez M, Kassem M, Liu F-F, Aldahmash A, et al. Bone morphogenetic protein 2 (BMP2) induces growth suppression and enhances chemosensitivity of human colon cancer cells. Cancer cell international. 2016;16: 1–12.

53. Choi SK, Kim HS, Jin T, Moon WK. LOXL4 knockdown enhances tumor growth and lung metastasis through collagen-dependent extracellular matrix changes in triple-negative breast cancer. Oncotarget. 2017;8: 11977.

54. Abraham V, Cao G, Parambath A, Lawal F, Handumrongkul C, Debs R, et al. Involvement of TIMP-1 in PECAM-1-mediated tumor dissemination. International Journal of Oncology. 2018;53: 488–502.

55. Henle AM, Nassar A, Puglisi-Knutson D, Youssef B, Knutson KL. Downregulation of TAP1 and TAP2 in early stage breast cancer. PloS one. 2017;12: e0187323.

56. Tao R, Xiong X, Liangpunsakul S, Dong XC. Sestrin 3 protein enhances hepatic insulin sensitivity by direct activation of the mTORC2-Akt signaling. Diabetes. 2015;64: 1211–1223.

57. Chen Y, Huang T, Yu Z, Yu Q, Wang Y, Hu J, et al. The functions and roles of sestrins in regulating human diseases. Cellular & Molecular Biology Letters. 2022;27: 1–24.

58. Denley A, Cosgrove LJ, Booker GW, Wallace JC, Forbes BE. Molecular interactions of the IGF system. Cytokine & growth factor reviews. 2005;16: 421–439.

59. Lau MM, Stewart CE, Liu Z, Bhatt H, Rotwein P, Stewart CL. Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. Genes & development. 1994;8: 2953–2963.

60. Vantaku V, Putluri V, Bader DA, Maity S, Ma J, Arnold JM, et al. Epigenetic loss of AOX1 expression via EZH2 leads to metabolic deregulations and promotes bladder cancer progression. Oncogene. 2020;39: 6265–6285.

61. Khaled WT, Choon Lee S, Stingl J, Chen X, Raza Ali H, Rueda OM, et al. BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. Nature communications. 2015;6: 5987.

62. Shivapurkar N, Stastny V, Okumura N, Girard L, Xie Y, Prinsen C, et al. Cytoglobin, the newest member of the globin family, functions as a tumor suppressor gene. Cancer research. 2008;68: 7448–7456.

63. Wu X, Li C, Wang Z, Zhang Y, Liu S, Chen S, et al. A bioinformatic analysis study of m7G regulator-mediated methylation modification patterns and tumor microenvironment infiltration in glioblastoma. BMC cancer. 2022;22: 1–17.

64. Zhang H, Liang J, Lu T, Li M, Shan G, Bi G, et al. AGRN promotes lung adenocarcinoma progression by activating Notch signaling pathway and acts as a therapeutic target. Pharmacological Research. 2023; 106819.

65. Yang X, Liu H, Ye T, Ye Z. High expression of LEF1 correlates with poor prognosis in solid tumors, but not blood tumors: a meta-analysis. Bioscience Reports. 2020;40: BSR20202520.

66. Wang T, Dang N, Tang G, Li Z, Li X, Shi B, et al. Integrating bulk and single-cell RNA sequencing reveals cellular heterogeneity and immune infiltration in hepatocellular carcinoma. Molecular Oncology. 2022;16: 2195–2213.

67. Lu L, Pandey AK, Houseal MT, Mulligan MK. The genetic architecture of murine glutathione transferases. PLoS One. 2016;11: e0148230.

68. Lim K-S, Chang S-S, Choi B-H, Lee S-H, Lee K-T, Chai H-H, et al. Genome-Wide analysis of allele-specific expression patterns in seventeen tissues of Korean Cattle (Hanwoo). Animals. 2019;9: 727.

69. Li SM, Valo Z, Wang J, Gao H, Bowers CW, Singer-Sam J. Transcriptome-wide survey of mouse CNS-derived cells reveals monoallelic expression within novel gene families. PloS one. 2012;7: e31751.

70. de Souza MM, Zerlotini A, Rocha MIP, Bruscadin JJ, Diniz WJ da S, Cardoso TF, et al. Allele-specific expression is widespread in Bos indicus muscle and affects meat quality candidate genes. Scientific Reports. 2020;10: 10204.

71. Edwards CA, Watkinson WM, Telerman SB, Hulsmann LC, Hamilton RS, Ferguson-Smith AC. Reassessment of weak parent-of-origin expression bias shows it rarely exists outside of known imprinted regions. Elife. 2023;12: e83364.

72. Huo H-N, Zhang C, Wang K, Chen W-N, Zhang Y-J, Yu W-L, et al. A novel imprinted locus on bovine chromosome 18 homologous with human Chromosome 16q24. 1. 2023.

73. Lim K-S, Kim H-C, Choi B-H, Son J-W, Lee K-T, Choi T-J, et al. Identification of Monoallelically Expressed Genes Associated with Economic Traits in Hanwoo (Korean Native Cattle). Animals. 2021;12: 84.

74. Cao W, Douglas KC, Samollow PB, VandeBerg JL, Wang X, Clark AG. Origin and evolution of marsupial-specific imprinting clusters through lineage-specific gene duplications and acquisition of promoter differential methylation. Molecular Biology and Evolution. 2023;40: msad022.

75. Huang W-C, Ferris E, Cheng T, Hörndli CS, Gleason K, Tamminga C, et al. Diverse non-genetic, allele-specific expression effects shape genetic architecture at the cellular level in the mammalian brain. Neuron. 2017;93: 1094-1109. e7.

76. Xie J, Ji T, Ferreira MA, Li Y, Patel BN, Rivera RM. Modeling allele-specific expression at the gene and SNP levels simultaneously by a Bayesian logistic mixed regression model. BMC bioinformatics. 2019;20: 530.

77. Schulze KV, Szafranski P, Lesmana H, Hopkin RJ, Hamvas A, Wambach JA, et al. Novel parent-of-origin-specific differentially methylated loci on chromosome 16. Clinical Epigenetics. 2019;11: 1–10.

78. Chen Z, Hagen DE, Wang J, Elsik CG, Ji T, Siqueira LG, et al. Global assessment of imprinted gene expression in the bovine conceptus by next generation sequencing. Epigenetics. 2016;11: 501–16. doi:10.1080/15592294.2016.1184805

79. Du J, Li Z, Li Q-Z, Guan T, Yang Q, Xu H, et al. Enoyl coenzyme a hydratase domain–containing 2, a potential novel regulator of myocardial ischemia injury. Journal of the American Heart Association. 2013;2: e000233.

80. Riedemann J, Macaulay VM. IGF1R signalling and its inhibition. Endocrine-related cancer. 2006;13: S33–S43.

81. Maunakea AK, Chepelev I, Cui K, Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. Cell research. 2013;23: 1256–1269.

82. Heallen T, Zhang M, Wang J, Bonilla-Claudio M, Klysik E, Johnson RL, et al. Hippo pathway inhibits Wnt signaling to restrain cardiomyocyte proliferation and heart size. Science. 2011;332: 458–461.

83. Wang KH, Kupa J, Duffy KA, Kalish JM. Diagnosis and management of Beckwith-Wiedemann syndrome. Frontiers in pediatrics. 2020;7: 562.

84. Tian M, Liu W, Zhang Q, Huang Y, Li W, Wang W, et al. MYSM1 represses innate immunity and autoimmunity through suppressing the cGAS-STING pathway. Cell Reports. 2020;33.

85. Wiehle L, Thorn GJ, Raddatz G, Clarkson CT, Rippe K, Lyko F, et al. DNA (de) methylation in embryonic stem cells controls CTCF-dependent chromatin boundaries. Genome research. 2019;29: 750–761.

86. Markaki Y, Smeets D, Fiedler S, Schmid VJ, Schermelleh L, Cremer T, et al. The potential of 3D-FISH and super-resolution structured illumination microscopy for studies of 3D nuclear architecture: 3D structured illumination microscopy of defined chromosomal structures visualized by 3D (immuno)-FISH opens new perspectives for studies of nuclear architecture. Bioessays. 2012;34: 412–426.

87. Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, et al. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. 2005;3: e157.

88. Han Z, Cui K, Placek K, Hong N, Lin C, Chen W, et al. Diploid genome architecture revealed by multi-omic data of hybrid mice. Genome research. 2020;30: 1097–1106.

89. Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature. 2016;537: 553–557.

90. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nature Reviews Genetics. 2009;10: 295–304.

91. Wang C, Liu X, Gao Y, Yang L, Li C, Liu W, et al. Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. Nature cell biology. 2018;20: 620–631.

92. Ryu J-K, Hwang D-E, Choi J-M. Current understanding of molecular phase separation in chromosomes. International Journal of Molecular Sciences. 2021;22: 10736.

93. Hildebrand EM, Dekker J. Mechanisms and Functions of Chromosome Compartmentalization. Trends in Biochemical Sciences. 2020.

94. Quinodoz SA, Jachowicz JW, Bhat P, Ollikainen N, Banerjee AK, Goronzy IN, et al. RNA promotes the formation of spatial compartments in the nucleus. Cell. 2021.

95. Wang L, Gao Y, Zheng X, Liu C, Dong S, Li R, et al. Histone modifications regulate chromatin compartmentalization by contributing to a phase separation mechanism. Molecular cell. 2019;76: 646-659. e6.

96. Bian Q, Anderson EC, Yang Q, Meyer BJ. Histone H3K9 methylation promotes formation of genome compartments in Caenorhabditis elegans via chromosome compaction and perinuclear anchoring. Proceedings of the National Academy of Sciences. 2020;117: 11459–11470.

97. Li E, Zhang Y. DNA methylation in mammals. Cold Spring Harbor perspectives in biology. 2014;6: a019133.

98. Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science (New York, NY). 2009;326: 289–293.

99. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159: 1665–1680.

100. Karbassi E, Rosa-Garrido M, Chapski DJ, Wu Y, Ren S, Wang Y, et al. Direct visualization of cardiac transcription factories reveals regulatory principles of nuclear architecture during pathological remodeling. Journal of molecular and cellular cardiology. 2019;128: 198–211.

101. Robson MI, Jose I, Czapiewski R, Sivakumar A, Kerr AR, Schirmer EC. Constrained release of lamina-associated enhancers and genes from the nuclear envelope during T-cell activation facilitates their association in chromosome compartments. Genome research. 2017;27: 1126–1138.

102. Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature. 2017;544: 59–64.

103. Baker SW, Duffy KA, Richards-Yutz J, Deardorff MA, Kalish JM, Ganguly A. Improved molecular detection of mosaicism in Beckwith-Wiedemann Syndrome. Journal of medical genetics. 2021;58: 178–184.

104. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience. 2020;9: giaa021.

105. Tan G, Lenhard B. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. Bioinformatics (Oxford, England). 2016;32: 1555–1556.

106. Fornes O, Castro-Mondragon JA, Khan A, Van der Lee R, Zhang X, Richmond PA, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. Nucleic acids research. 2020;48: D87–D92.

107. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. bioinformatics. 2009;25: 1754–1760.

108. Open2C, Abdennur N, Fudenberg G, Flyamer IM, Galitsyna AA, Goloborodko A, et al. Pairtools: from sequencing data to chromosome contacts. bioRxiv. 2023; 2023.02. 13.528389.

109. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell systems. 2016;3: 95–98.

110. Abdennur N, Mirny LA. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics. 2020;36: 311–316.

111. Hayes BJ, Daetwyler HD. 1000 bull genomes project to map simple and complex genetic traits in cattle: applications and outcomes. Annual review of animal biosciences. 2019;7: 89–102.

112. Broad Institute. Picard toolkit. Broad Institute, GitHub repository. Broad Institute, GitHub Repository; 2021. Available: http://broadinstitute.github.io/picard/

113. Van der Auwera GA, O'Connor BD. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly Media; 2020.

114. Stansfield JC, Cresswell KG, Dozmorov MG. multiHiCcompare: joint normalization and comparative analysis of complex Hi-C experiments. Bioinformatics. 2019;35: 2916–2923.

115. Cresswell KG, Stansfield JC, Dozmorov MG. SpectralTAD: an R package for defining a hierarchy of topologically associated domains using spectral clustering. BMC bioinformatics. 2020;21: 1–19.

116. Hua D, Gu M, Du Y, Qi L, Du X, Bai Z, et al. DiffDomain enables identification of structurally reorganized topologically associating domains. Research in Computational Molecular Biology: 26th Annual International Conference, RECOMB 2022, San Diego, CA, USA, May 22–25, 2022, Proceedings. Springer; 2022. pp. 302–303.

117. Chakraborty A, Wang JG, Ay F. dcHiC detects differential compartments across multiple Hi-C datasets. Nature Communications. 2022;13: 6827.

118. Bushnell B. BBMap. 2021. Available: sourceforge.net/projects/bbmap/

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

120. Farrell C, Thompson M, Tosevska A, Oyetunde A, Pellegrini M. BiSulfite Bolt: A bisulfite sequencing analysis platform. GigaScience. 2021;10: giab033.

121. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25: 2078–2079.

122. Ji T. A Bayesian hidden Markov model for detecting differentially methylated regions. Biometrics. 2019;75: 663–673.

123. Phanstiel DH, Boyle AP, Araya CL, Snyder MP. Sushi. R: flexible, quantitative and integrative genomic visualizations for publication-quality multi-panel figures. Bioinformatics. 2014;30: 2808–2810.

124. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England). 2010;26: 139–140.

125. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics. 2008;9: 1–13.

126. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic acids research. 2016;44: D733-45. doi:10.1093/nar/gkv1189

127. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome research. 2002;12: 996–1006.

128. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489: 57.

129. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic acids research. 2002;30: 207–210.

130. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26: 841–842.

131. Gu Z, Eils R, Schlesner M, Ishaque N. EnrichedHeatmap: an R/Bioconductor package for comprehensive visualization of genomic signal associations. BMC genomics. 2018;19: 1–7.

132. Khan A, Zhang X. dbSUPER: a database of super-enhancers in mouse and human genome. Nucleic acids research. 2016;44: D164–D171.

133. Open2C, Abdennur N, Fudenberg G, Flyamer I, Galitsyna AA, Goloborodko A, et al. Bioframe: operations on genomic intervals in pandas dataframes. bioRxiv. 2022; 2022.02. 16.480748.

134. Wickham H. ggplot2. Wiley Interdisciplinary Reviews: Computational Statistics. 2011;3: 180–185.

135. Gu Z. Complex heatmap visualization. iMeta. 2022;1: e43.

136. Kramer NE, Davis ES, Wenger CD, Deoudes EM, Parker SM, Love MI, et al. Plotgardener: Cultivating precise multi-panel figures in R. Bioinformatics. 2022;38: 2042–2045.

137. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome biology. 2012;13: 1–9.

138. Gu Z, Gu L, Eils R, Schlesner M, Brors B. Circlize implements and enhances circular visualization in R. Bioinformatics. 2014;30: 2811–2812.

139. Kassambara A. ggpubr:“ggplot2” based publication ready plots. R package version 04 0. 2020;438.

140. Tang Y, Horikoshi M, Li W. ggfortify: unified interface to visualize statistical results of popular R packages. R J. 2016;8: 474.

141. Hunter JD. Matplotlib: A 2D graphics environment. Computing in science & engineering. 2007;9: 90–95.

Figure Click here to access/download;Figure;Figure and legend 20231212.docx

Table 1. Summary of statistical comparisons

Species	Type	Index	Comparison	DMR	DEG	DIR*	Diff_TAD	Diff_compartment Raw*	Diff_compartment Unique
Human	Overall	1	BWS vs Control	5708	12	1489	191	482	83
		2	BWS_IC1_GOM vs Control	21124	23	702	105	950	157
		3	BWS_IC2_LOM vs Control	7865	8	1399	144	743	125
		4	BWS_IC2_LOM_S1 vs Control	23582	742	54921	256	2050	328
		5	BWS_IC2_LOM_S2 vs Control	1915	2	116	55	1909	532
		6	BWS_CDKN1C_Mut vs Control	37122	10	833	286	551	105
		7	BWS_IC2_LOM vs BWS_IC1_GOM	6535	4	445	2197	471	148
		8	BWS_CDKN1C_Mut vs BWS_IC1_GOM	66575	53	2563	3706	1252	276
		9	BWS_CDKN1C_Mut vs BWS_IC2_LOM	7169	8	1149	1985	560	150
		10	BWS_IC2_LOM_S1 vs BWS_IC2_LOM_S2	20657	484	68396	2508	4592	750
	Sex specific	11	BWS_IC2_LOM_Male vs Control_Male	53488	12	289	239	2623	486
		12	BWS_Female vs Control_Female	12700	7	600	434	926	157
		13	BWS_IC1_GOM_Female vs Control_Female	26034	20	682	105	1529	314
		14	BWS_IC2_LOM_Female vs Control_Female	26459	43	7867	695	1743	448
		15	BWS_CDKN1C_Mut_Female vs Control_Female	12320	19	1719	540	996	242
		16	BWS_IC2_LOM_Female vs BWS_IC1_GOM_Female	13100	3	218	2783	952	354
		17	BWS_CDKN1C_Mut_Female vs BWS_IC1_GOM_Female	69325	31	437	3059	1441	289
		18	BWS_CDKN1C_Mut_Female vs BWS_IC2_LOM_Female	27140	56	598	1055	966	243
Bovine	Overall	19	LOS vs Control	10860	112	20200	369	2328	192
		20	LOS_All vs Control	4071	6	566	216	1010	129
	Sex specific	21	LOS_Male vs Control_Male	75370	109	10299	231	3941	294
		22	LOS_Female vs Control_Female	34270	213	4671	526	2737	632
		23	LOS_All_Female vs Control_Female	26189	23	495	339	1489	471
	Allele specific	24	LOS_Maternal vs Control_Maternal	17611	29	1000	681	716	231
		25	LOS_Paternal vs Control_Paternal	15708	16	1115	651	1550	624
		26	Control_Paternal vs Control_Maternal	10143	122	4940	105	2367	1038
		27	LOS_All_Maternal vs Control_Maternal	15405	2	275	709	307	147
		28	LOS_All_Paternal vs Control_Paternal	14061	0	291	689	1112	511
	Sex and allele specific	29	LOS_Male_Maternal vs Control_Male_Maternal	26690	28	899	339	2146	838
		30	LOS_Male_Paternal vs Control_Male_Paternal	21682	25	1177	261	1566	520
		31	LOS_Female_Maternal vs Control_Female_Maternal	9286	19	1980	96	2609	1491
		32	LOS_Female_Paternal vs Control_Female_Paternal	7364	9	1737	66	4681	2830
		33	Control_Male_Paternal vs Control_Male_Maternal	1322	54	2262	23	1063	542
		34	Control_Female_Paternal vs Control_Female_Maternal	1302	14	1104	16	3551	2171
		35	LOS_All_Female_Maternal vs Control_Female_Maternal	11118	4	1035	138	1674	1139
		36	LOS_All_Female_Paternal vs Control_Female_Paternal	8819	0	956	92	3799	2440

DMR = differentially methylated regions; DEG = differentially expressed genes; DIR = differential interacting regions; diff_TAD = differential topologically associating domains; Diff_compartment = differential compartments; Unique = after merging overlapped and continuous regions with the same direction of change; *Count includes results from all the bin sizes

Table 1

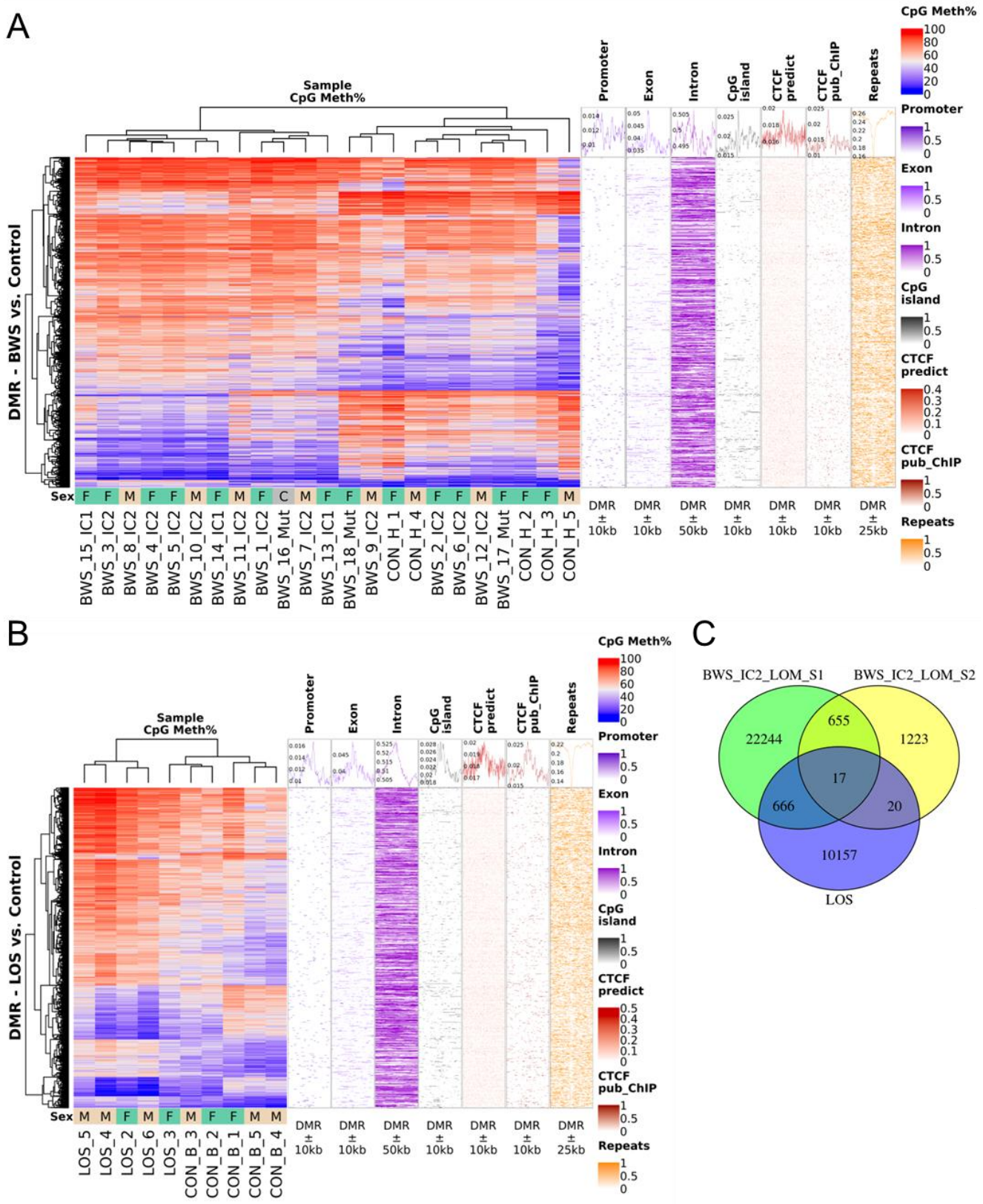
A
Sample
CpG Meth%
DMR - BWS vs. Control
Sex
BWS_15_IC1
BWS_3_IC2
BWS_8_IC2
BWS_4_IC2
BWS_5_IC2
BWS_10_IC2
BWS_14_IC1
BWS_11_IC2
BWS_1_IC2
BWS_16_Mut
BWS_7_IC2
BWS_13_IC1
BWS_18_Mut
BWS_9_IC2
CON_H_1
CON_H_4
BWS_2_IC2
BWS_6_IC2
BWS_12_IC2
BWS_17_Mut
CON_H_2
CON_H_3
CON_H_5
Promoter
Exon
Intron
CpG island
CTCF predict
CTCF pub_ChIP
Repeats
DMR ± 10kb
DMR ± 50kb
DMR ± 25kb
CpG Meth%
B
DMR - LOS vs. Control
LOS_5
LOS_4
LOS_2
LOS_6
LOS_3
CON_B_3
CON_B_2
CON_B_1
CON_B_5
CON_B_4
C
BWS_IC2_LOM_S1
BWS_IC2_LOM_S2
LOS
22244
655
1223
17
666
20
10157

Figure 1

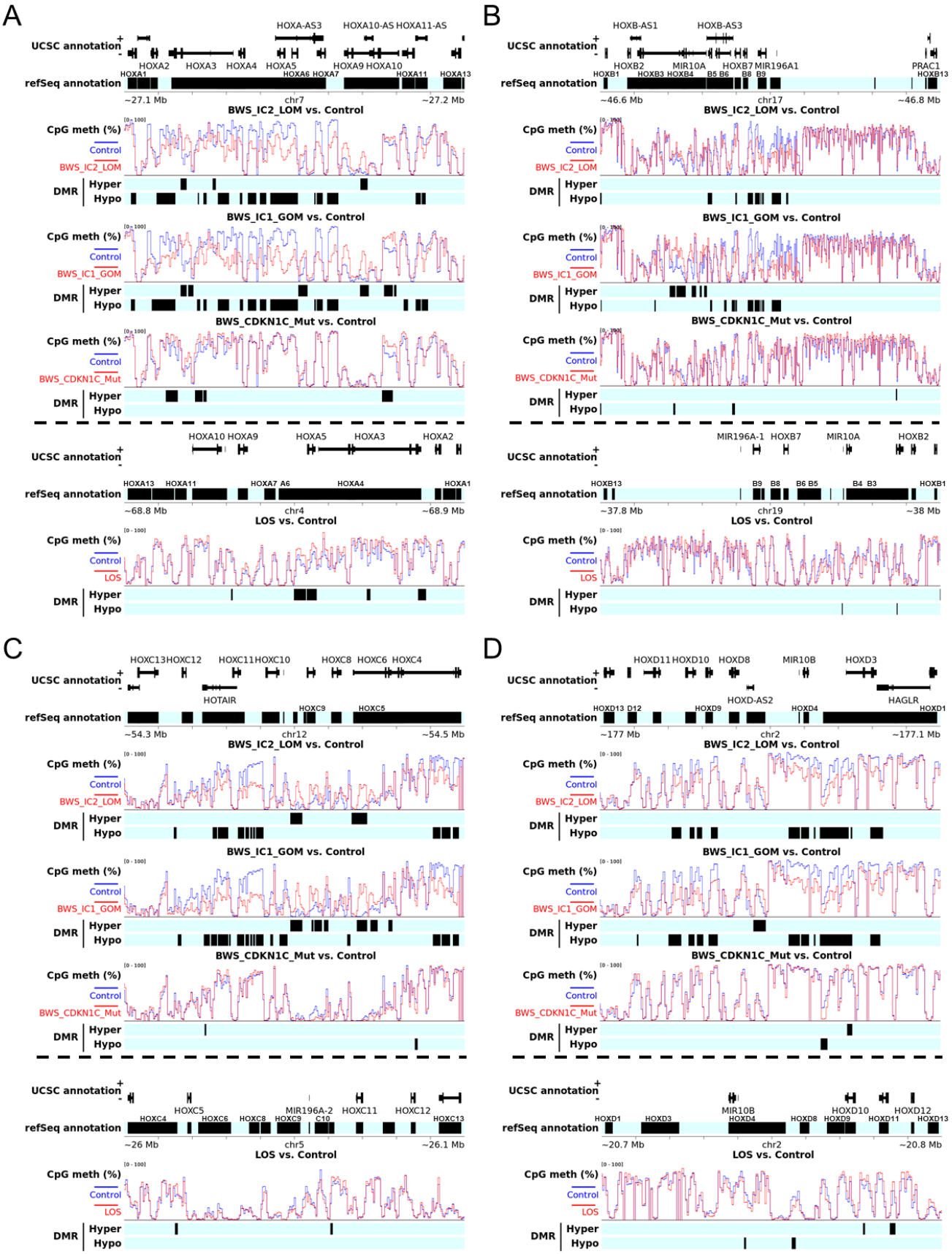
A
UCSC annotation
HOXA-AS3
HOXA10-AS HOXA11-AS
HOXA2 HOXA3 HOXA4 HOXA5 HOXA9 HOXA10
HOXA1 HOXA6 HOXA7 HOXA11 HOXA13
refSeq annotation
~27.1 Mb
chr7
~27.2 Mb
BWS_IC2_LOM vs. Control
CpG meth (%)
Control
BWS_IC2_LOM
DMR
Hyper
Hypo
BWS_IC1_GOM vs. Control
BWS_IC1_GOM
BWS_CDKN1C_Mut vs. Control
BWS_CDKN1C_Mut
HOXA10 HOXA9 HOXA5 HOXA3 HOXA2
HOXA13 HOXA11 HOXA7 A6 HOXA4 HOXA1
~68.8 Mb
chr4
~68.9 Mb
LOS vs. Control
LOS
B
HOXB-AS1
HOXB-AS3
HOXB2 MIR10A HOXB7 MIR196A1 PRAC1
HOXB1 HOXB3 HOXB4 B5 B6 B8 B9 HOXB13
~46.6 Mb
chr17
~46.8 Mb
MIR196A-1 HOXB7 MIR10A HOXB2
HOXB13 B9 B8 B6 B5 B4 B3 HOXB1
~37.8 Mb
chr19
~38 Mb
C
HOXC13 HOXC12 HOXC11 HOXC10 HOXC8 HOXC6 HOXC4
HOTAIR
HOXC9 HOXC5
~54.3 Mb
chr12
~54.5 Mb
HOXC5
MIR196A-2 HOXC11 HOXC12
HOXC4 HOXC6 HOXC8 HOXC9 C10 HOXC13
~26 Mb
chr5
~26.1 Mb
D
HOXD11 HOXD10 HOXD8 MIR10B HOXD3
HOXD-AS2
HAGLR
HOXD13 D12 HOXD9 HOXD4 HOXD1
~177 Mb
chr2
~177.1 Mb
MIR10B
HOXD10 HOXD12
HOXD1 HOXD3 HOXD4 HOXD8 HOXD9 HOXD11 HOXD13
~20.7 Mb
chr2
~20.8 Mb

Figure 2

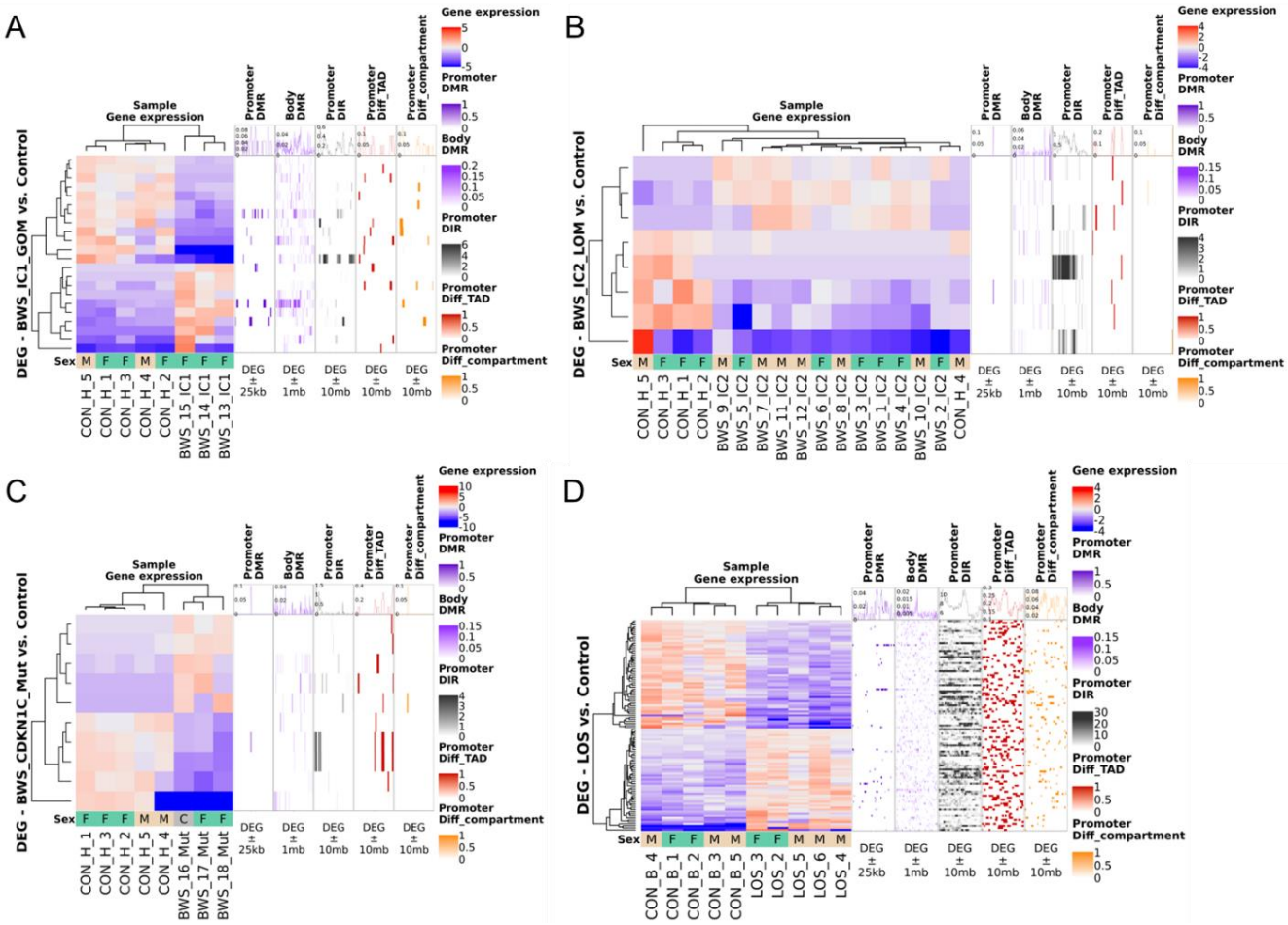
A
B
C
D
Sample
Gene expression
Gene expression
Promoter DMR
Body DMR
Promoter DIR
Promoter Diff_TAD
Promoter Diff_compartment
DEG - BWS_IC1_GOM vs. Control
DEG - BWS_IC2_LOM vs. Control
DEG - BWS_CDKN1C_Mut vs. Control
DEG - LOS vs. Control
Sex
DEG ± 25kb
DEG ± 1mb
DEG ± 10mb

Figure 3

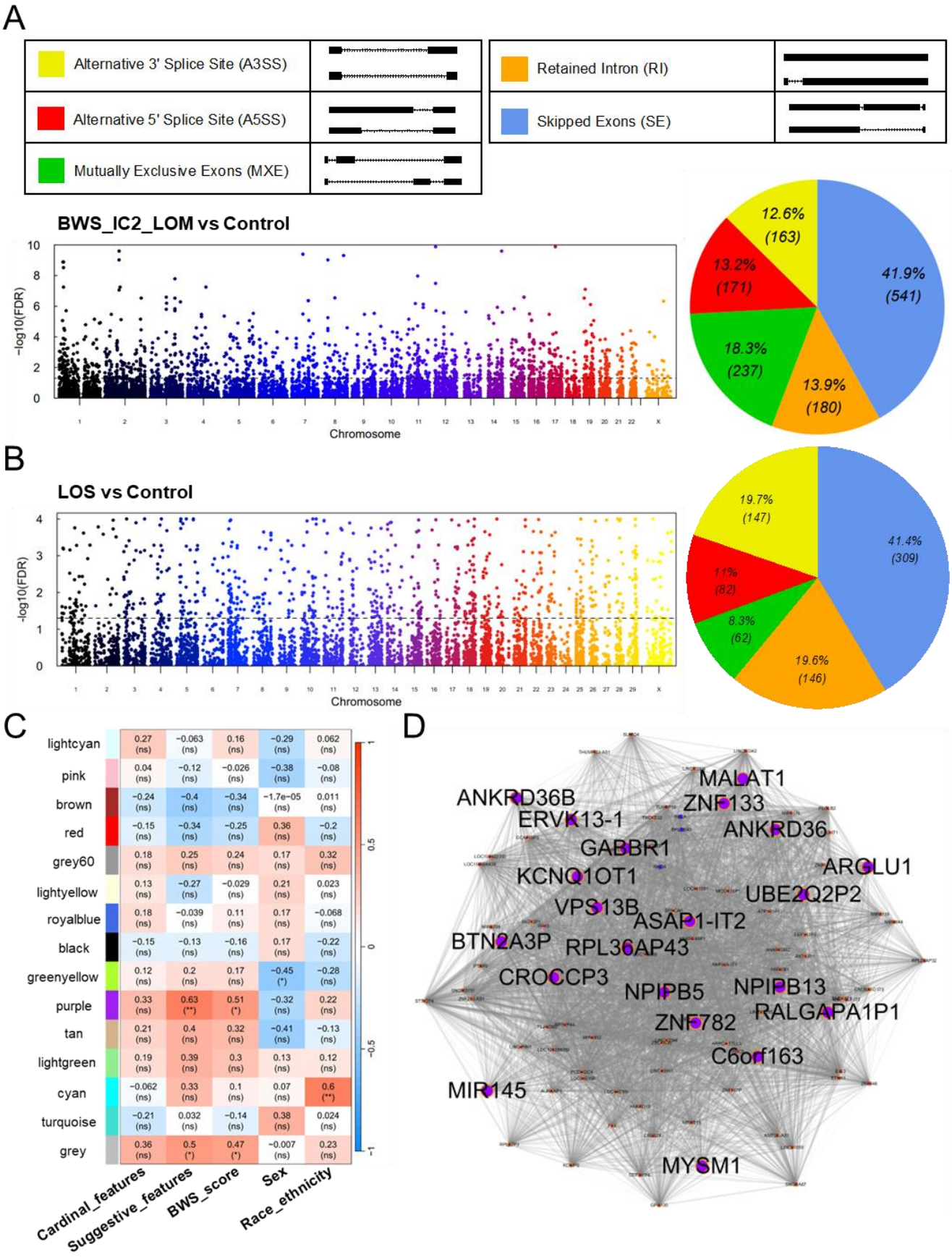
A
Alternative 3' Splice Site (A3SS)
Alternative 5' Splice Site (A5SS)
Mutually Exclusive Exons (MXE)
Retained Intron (RI)
Skipped Exons (SE)
BWS_IC2_LOM vs Control
-log10(FDR)
Chromosome
12.6% (163)
13.2% (171)
41.9% (541)
18.3% (237)
13.9% (180)
B
LOS vs Control
-log10(FDR)
Chromosome
19.7% (147)
41.4% (309)
11% (82)
8.3% (62)
19.6% (146)
C
lightcyan
pink
brown
red
grey60
lightyellow
royalblue
black
greenyellow
purple
tan
lightgreen
cyan
turquoise
grey
Cardinal_features
Suggestive_features
BWS_score
Sex
Race_ethnicity
D
MALAT1
ANKRD36B
ZNF133
ERVK13-1
ANKRD36
GABBR1
ARGLU1
KCNQ1OT1
UBE2Q2P2
VPS13B
ASAP1-IT2
BTN2A3P
RPL36AP43
CROCCP3
NPIPB5
NPIPB13
RALGAPA1P1
ZNF782
C6orf163
MIR145
MYSM1

Figure 4

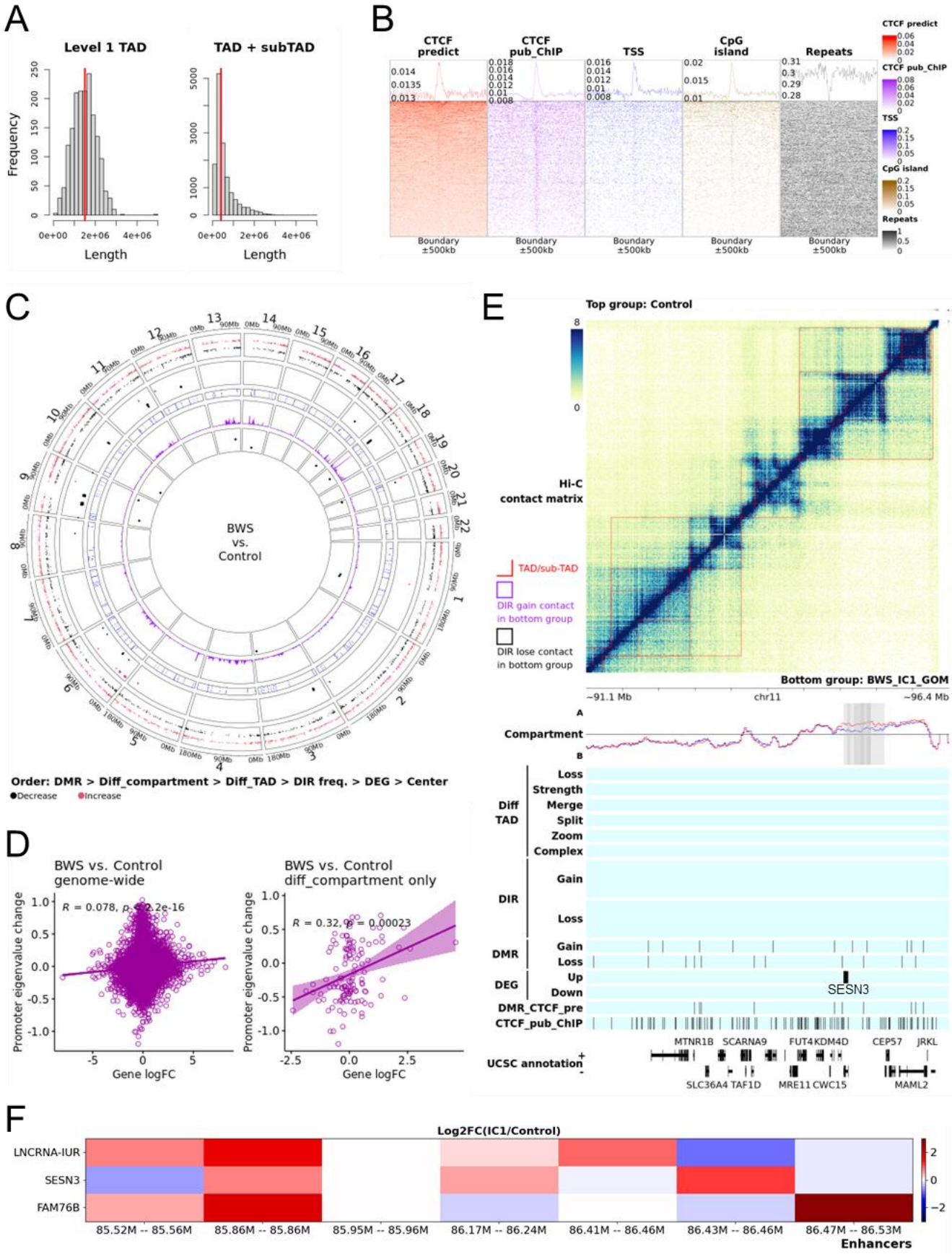
A
Level 1 TAD
TAD + subTAD
Frequency
Length
B
CTCF predict
CTCF pub_ChIP
TSS
CpG island
Repeats
Boundary ±500kb
C
BWS vs. Control
Order: DMR > Diff_compartment > Diff_TAD > DIR freq. > DEG > Center
Decrease
Increase
D
BWS vs. Control genome-wide
R = 0.078, p < 2.2e-16
BWS vs. Control diff_compartment only
R = 0.32, p = 0.00023
Promoter eigenvalue change
Gene logFC
E
Top group: Control
Hi-C contact matrix
TAD/sub-TAD
DIR gain contact in bottom group
DIR lose contact in bottom group
Bottom group: BWS_IC1_GOM
~91.1 Mb
chr11
~96.4 Mb
Compartment
Diff TAD
Loss
Strength
Merge
Split
Zoom
Complex
DIR
Gain
Loss
DMR
Gain
Loss
DEG
Up
Down
SESN3
DMR_CTCF_pre
CTCF_pub_ChIP
UCSC annotation
MTNR1B
SCARNA9
FUT4
KDM4D
CEP57
JRKL
SLC36A4
TAF1D
MRE11
CWC15
MAML2
F
Log2FC(IC1/Control)
LNCRNA-IUR
SESN3
FAM76B
85.52M -- 85.56M
85.86M -- 85.86M
85.95M -- 85.96M
86.17M -- 86.24M
86.41M -- 86.46M
86.43M -- 86.46M
86.47M -- 86.53M
Enhancers

Figure 5

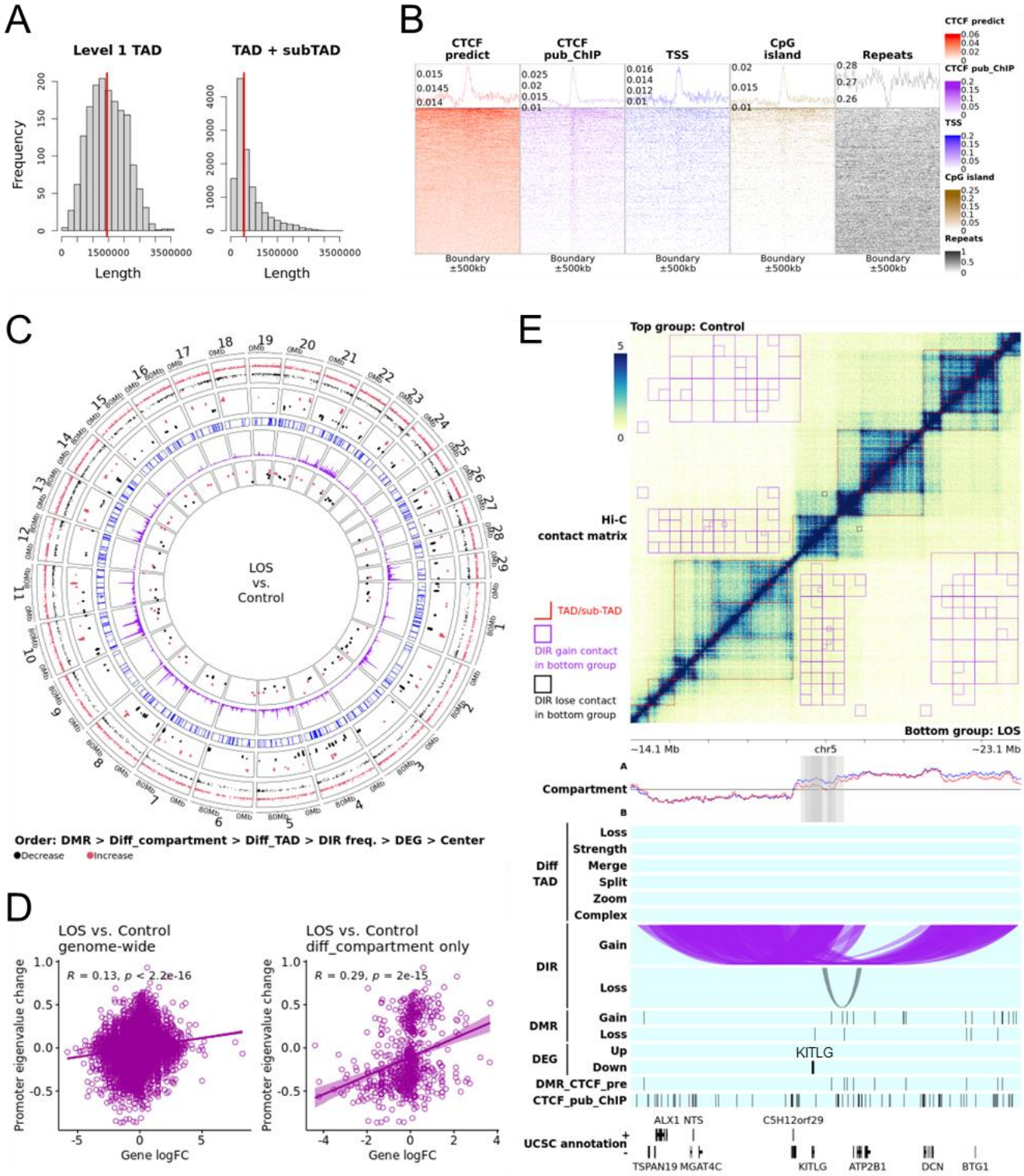
A
Level 1 TAD
TAD + subTAD
Frequency
Length
B
CTCF predict
CTCF pub_ChIP
TSS
CpG island
Repeats
Boundary ±500kb
C
LOS vs. Control
Order: DMR > Diff_compartment > Diff_TAD > DIR freq. > DEG > Center
Decrease
Increase
D
LOS vs. Control genome-wide
R = 0.13, p < 2.2e-16
LOS vs. Control diff_compartment only
R = 0.29, p = 2e-15
Promoter eigenvalue change
Gene logFC
E
Top group: Control
Hi-C contact matrix
TAD/sub-TAD
DIR gain contact in bottom group
DIR lose contact in bottom group
Bottom group: LOS
~14.1 Mb
chr5
~23.1 Mb
Compartment
Diff TAD
Loss
Strength
Merge
Split
Zoom
Complex
DIR
Gain
Loss
DMR
Gain
Loss
DEG
Up
Down
KITLG
DMR_CTCF_pre
CTCF_pub_ChIP
UCSC annotation
ALX1 NTS
C5H12orf29
TSPAN19 MGAT4C
KITLG
ATP2B1
DCN
BTG1

Figure 6

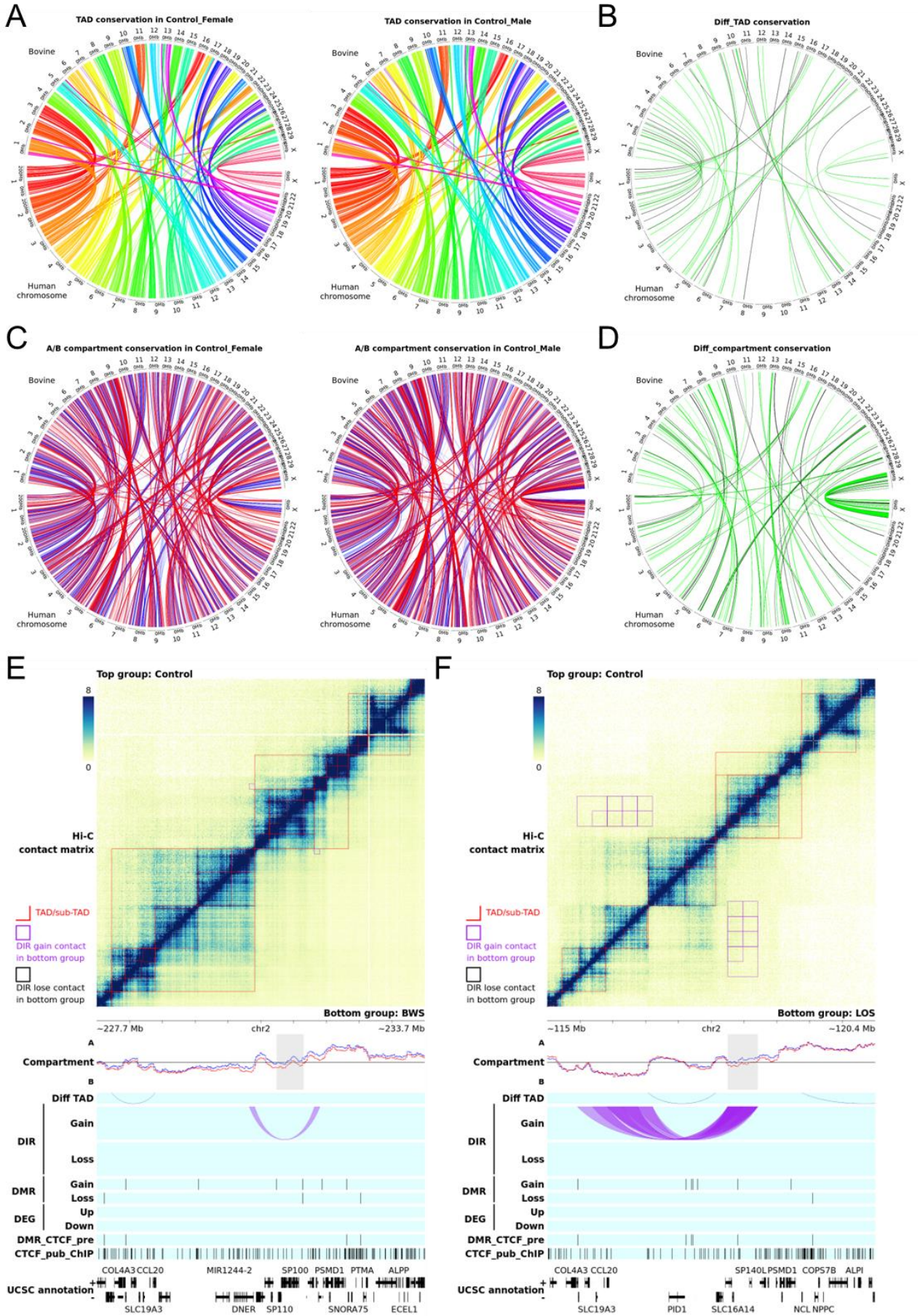
A
TAD conservation in Control_Female
TAD conservation in Control_Male
B
Diff_TAD conservation
C
A/B compartment conservation in Control_Female
A/B compartment conservation in Control_Male
D
Diff_compartment conservation
Bovine
Human chromosome
E
Top group: Control
Hi-C contact matrix
TAD/sub-TAD
DIR gain contact in bottom group
DIR lose contact in bottom group
Bottom group: BWS
~227.7 Mb
chr2
~233.7 Mb
Compartment
Diff TAD
DIR
Gain
Loss
DMR
DEG
Up
Down
DMR_CTCF_pre
CTCF_pub_ChIP
UCSC annotation
COL4A3 CCL20
MIR1244-2
SP100 PSMD1 PTMA ALPP
SLC19A3
DNER SP110
SNORA75
ECEL1
F
Top group: Control
Bottom group: LOS
~115 Mb
~120.4 Mb
SP140L PSMD1 COPS7B ALPI
PID1
SLC16A14
NCL NPPC

Figure 7