**iScience, Volume *25***

## Supplemental information

# Allele-specific aberration of imprinted domain chromosome architecture associates with large offspring syndrome

**Yahan Li, Frimpong Boadu, Max R. Highsmith, Darren E. Hagen, Jianlin Cheng, and Rocío Melissa Rivera**

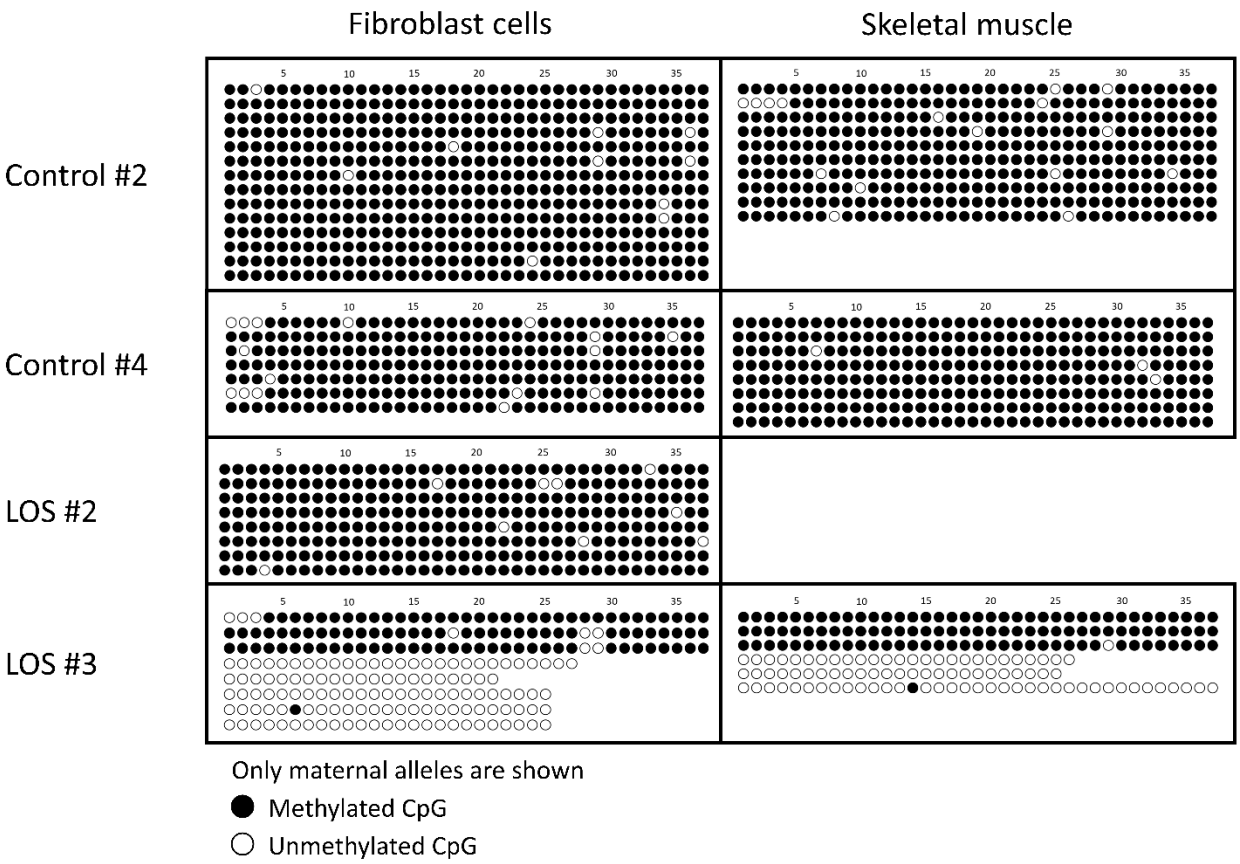


**Figure S1. Maternal allele CpG DNA methylation status at KvDMR1, related to Figure 1.** Methylation status was determined by bisulfite PCR, molecular cloning, and Sanger sequencing. Open circles represent unmethylated CpG and filled circles indicate methylated CpG. Only maternal alleles were shown, and paternal alleles had no methylated CpG found. For sample LOS#2, we did not perform bisulfite PCR and cloning for the muscle sample since the locus is methylated over this region in fibroblast.

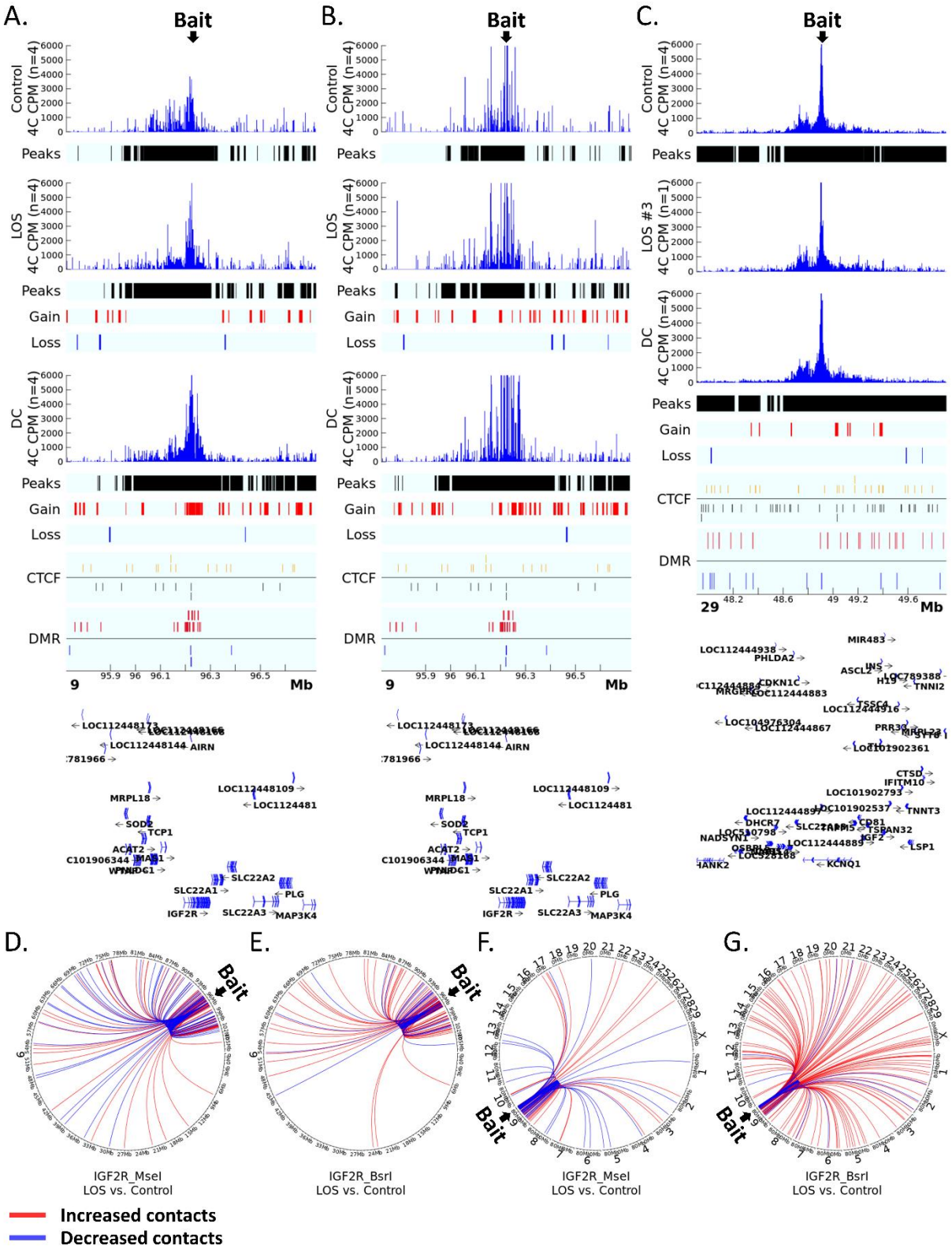


**Figure S2. 4C identified overall (not allele specific) cis and trans contacts, related**

**to Figure 3 and Figure 4.**

(A-C) Comparison of cis contacts between control, LOS, and DC. (A) IGF2R_MseI assay, (B) IGF2R_BsrI assay, and (C) KvDMR1 assays. Shown are the comparison of LOS and DC groups vs controls. Track ‘4C CPM’ shows the mean normalized count of reads aligned to the genome indicating physical contacts with the bait. Track ‘Peaks’ show regions with statistically significant contacts with the bait identified by fourSig software within a group. Track ‘Gain’ (red line) and ‘Loss’ (blue line) indicate regions with statistically significant difference in contacts with the bait regions identified by DESeq2 between groups. Track ‘CTCF’ shows predicted CTCF binding sites on the sense (gold line) or antisense (black line) strand. Track ‘DMR’ shows non-allelic differentially methylated regions identified between the LOS and the control group with the red line indicating increased and blue line indicating decreased methylation levels. The gene annotation is at the bottom of the figure. Mb = megabases. CPM = counts per million reads.

(D-G) Comparison of contacts in far-cis and trans between groups. (D and F) IGF2R_MseI assay and (E and G) IGF2R_BsrI. (D-E) far-cis contacts (chromosome 9) and (F-G) trans contacts (interchromosomal). Circos plots showing DESeq2-identified statistically different contacts with the bait in LOS vs. controls. Red line indicates increased contacts and blue line indicates decreased contacts.

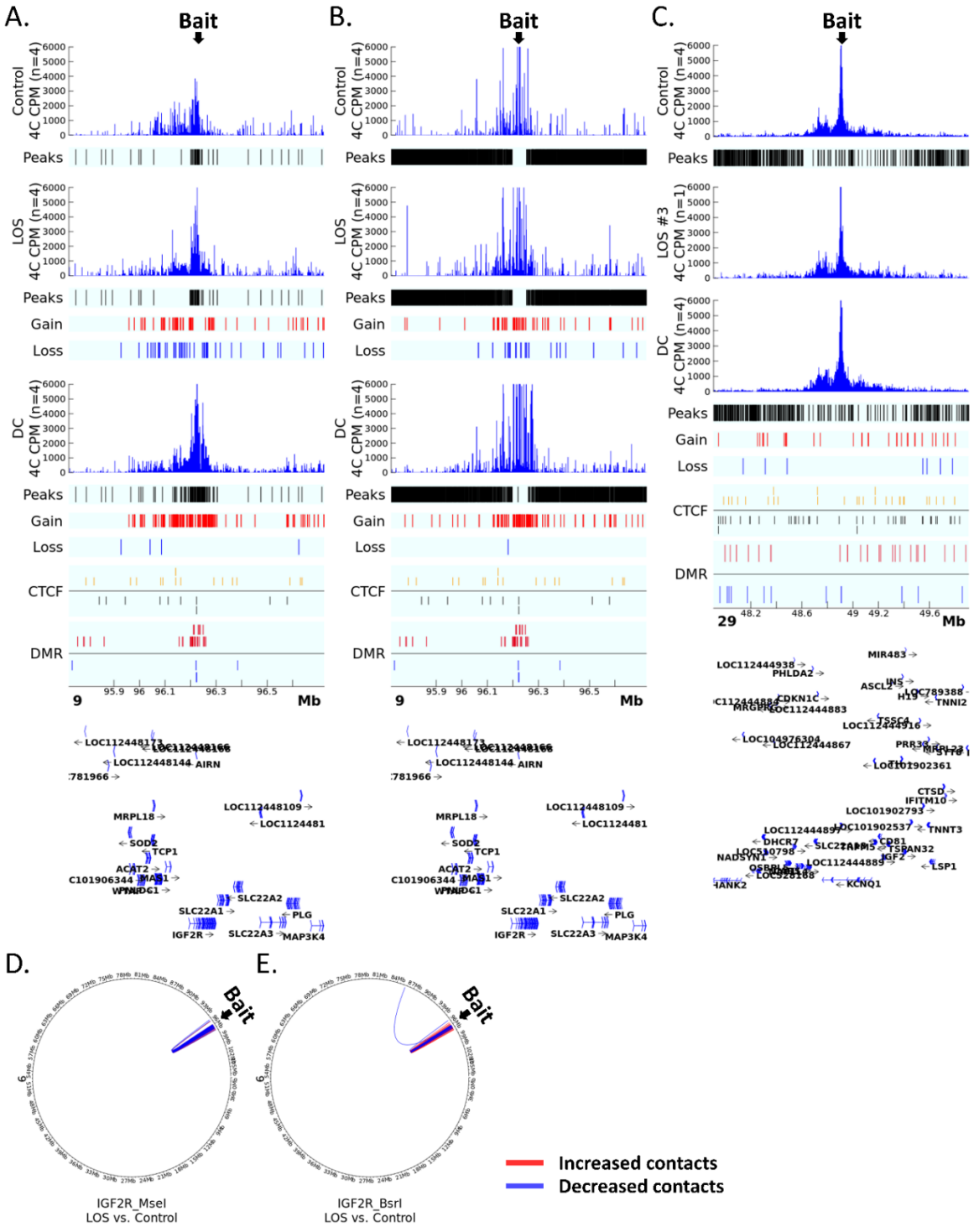


**Figure S3. 4Cker statistical results for overall (not allele specific) contacts, related to Figure 3 and Figure 4.**

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

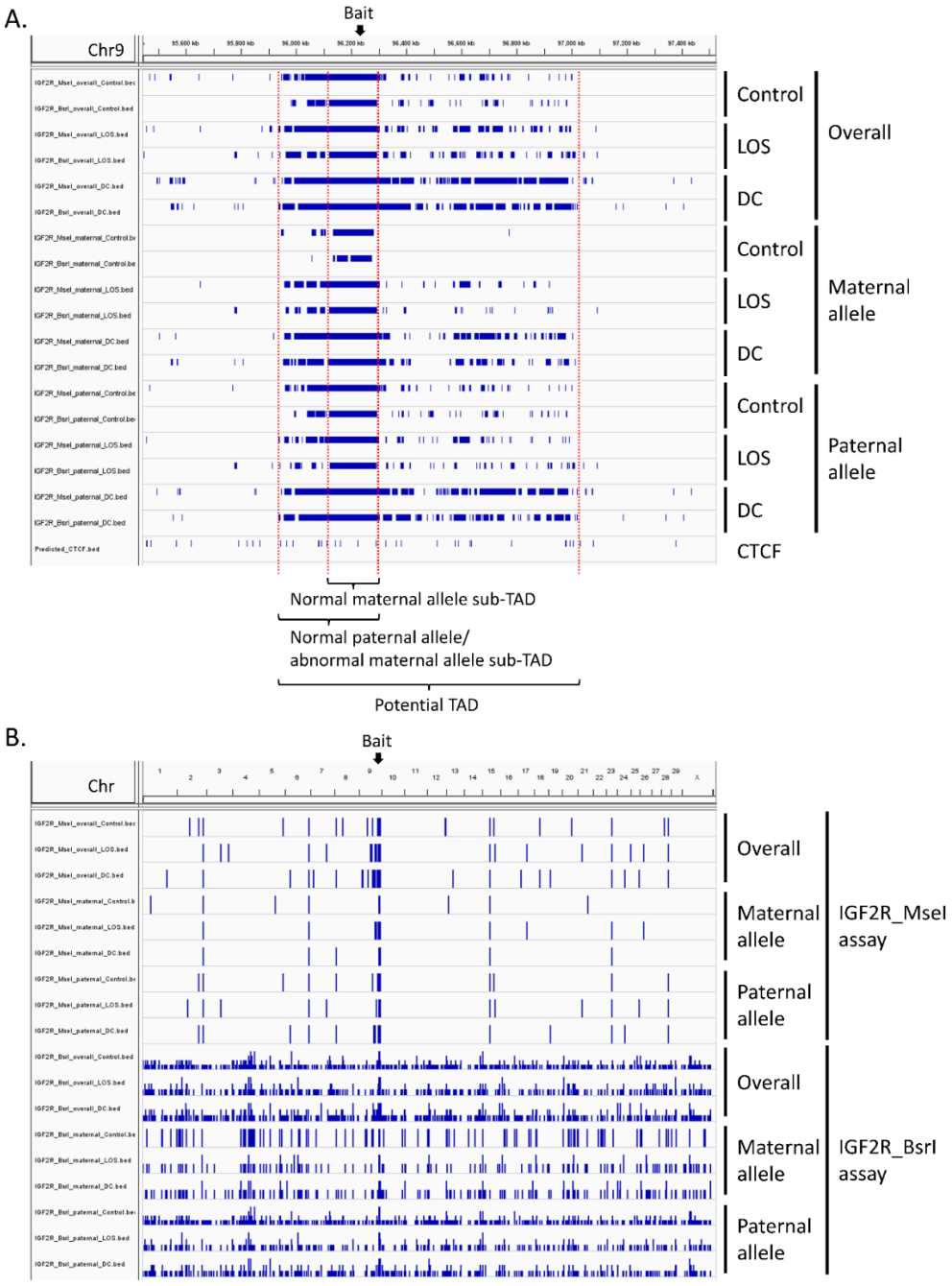


**Figure S4. Significant contacts of the bait (*IGF2R* ICR) detected by fourSig, related to Figure 3.**

(A) Potential TAD and allelic sub-TAD for *IGF2R* domain. Overall = not allelic.

(B) Genome-wide contacts of *IGF2R* ICR. Conserved contacts were detected between treatment groups within each assay.

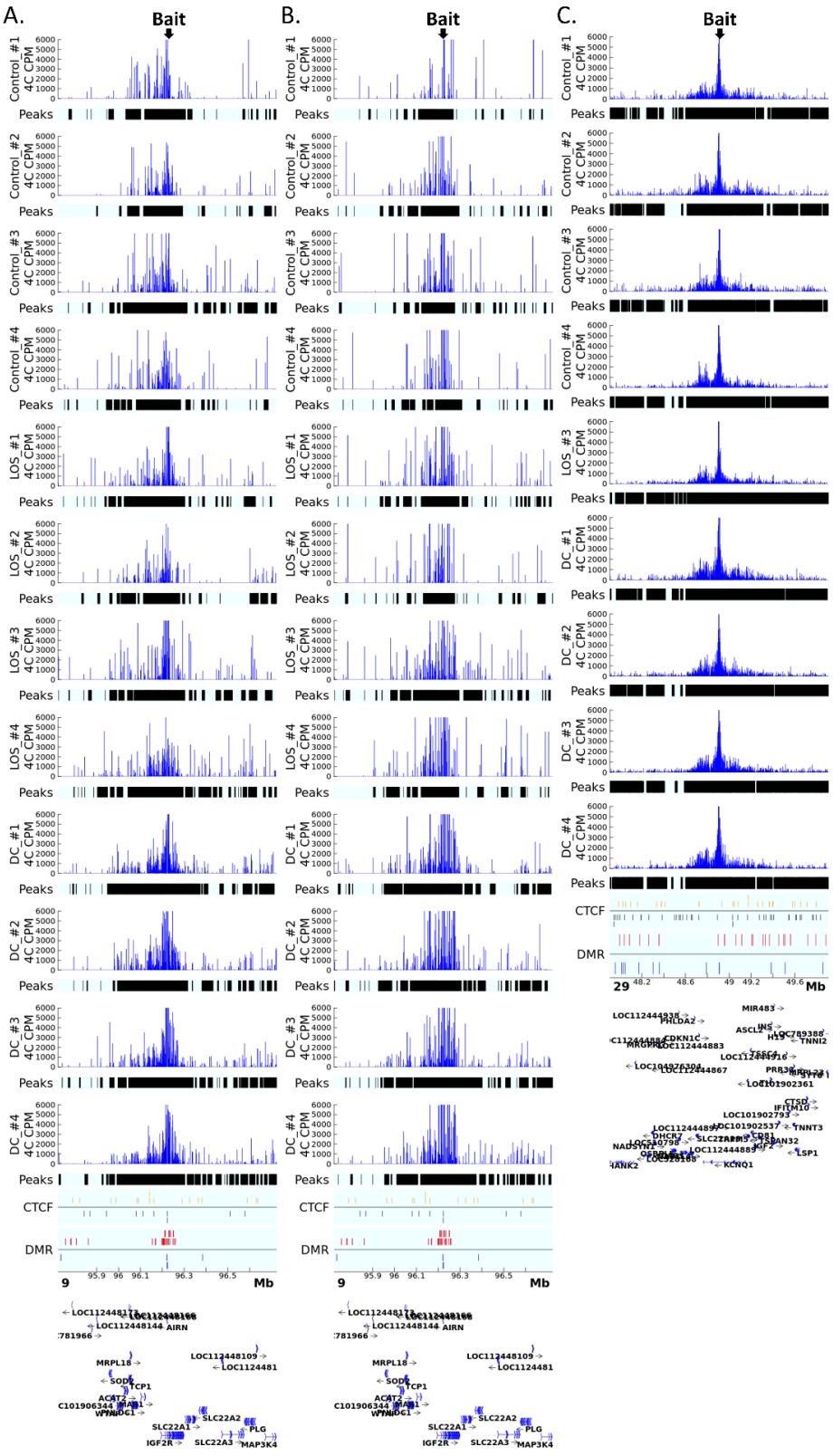


**Figure S5. 4C identified overall (not allele specific) cis contacts for individual**

**samples, related to Figure 3 and Figure 4.**

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

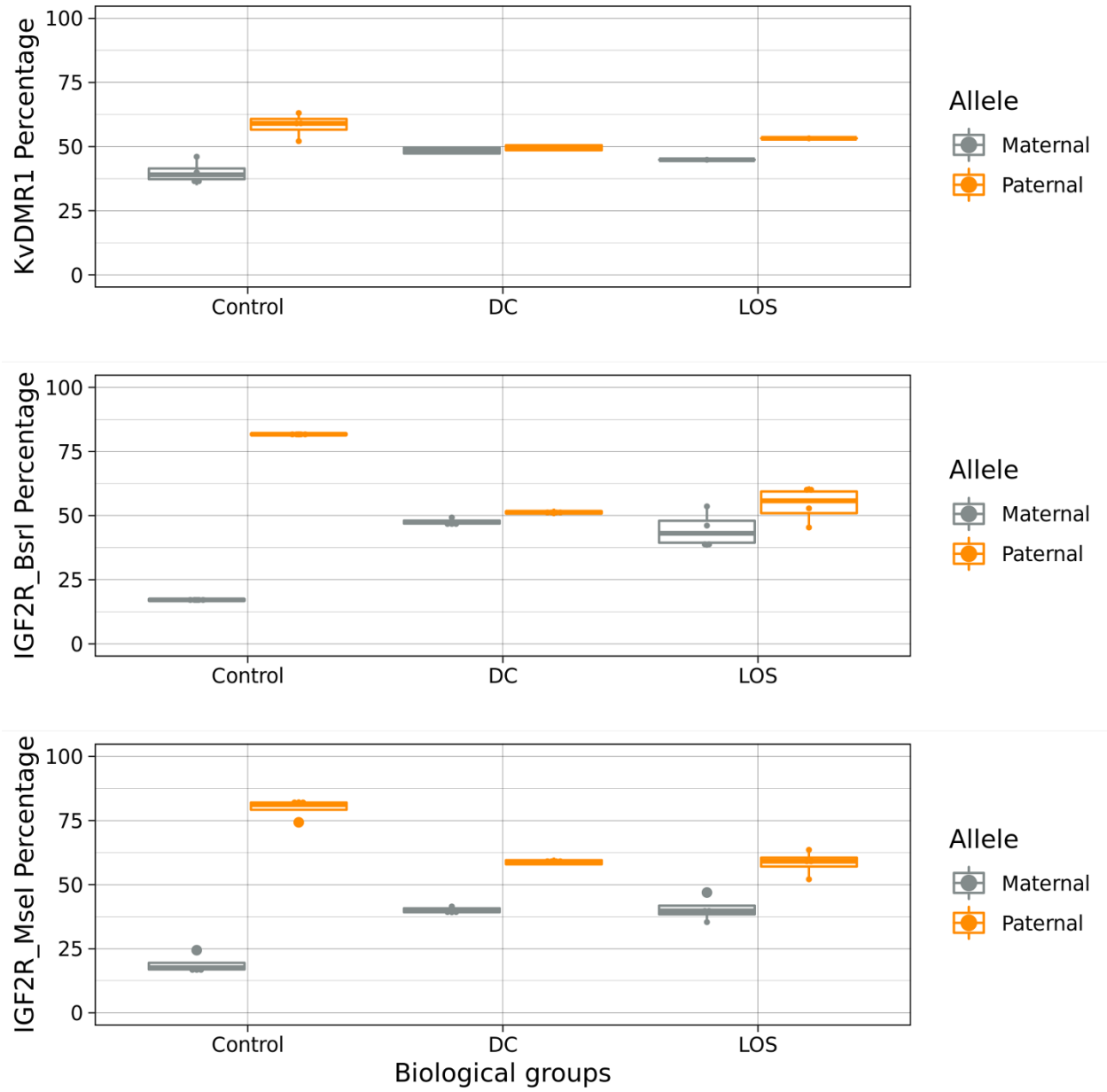


**Figure S6. 4C sequencing reads allelic alignment, related to Figure 3 and Figure 4.**

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

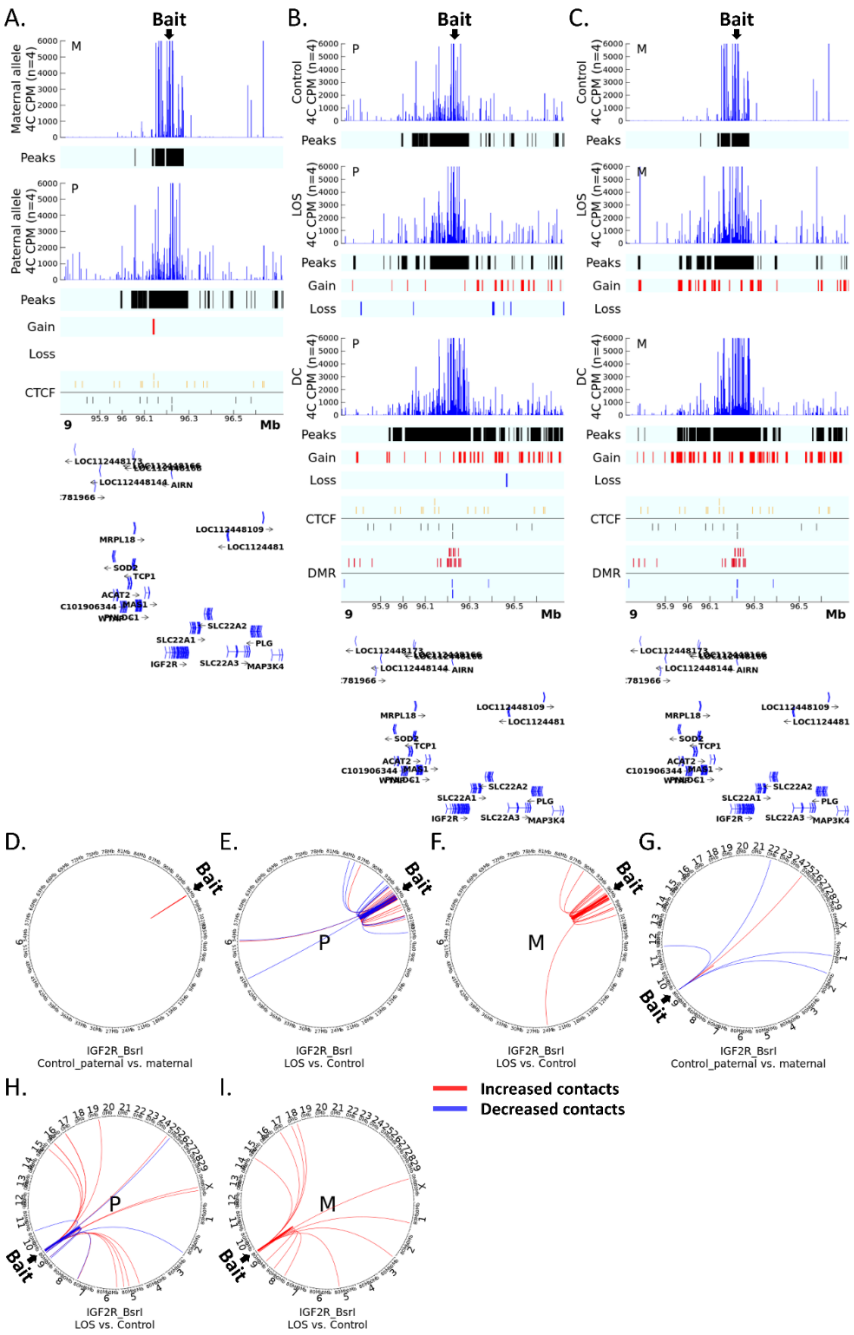


**Figure S7. 4C identified allele-specific cis and trans contacts with *IGF2R* ICR, related to Figure 3.** Shown are data for the IGF2R_BsrI assay.

(A) Comparison of cis contacts between the paternal and maternal alleles in controls. Track '4C CPM' shows the mean normalized count of reads aligned to the genome indicating physical contacts with the bait. Track 'Peaks' show regions with statistically

significant contacts with the bait identified by fourSig software within a group. Track ‘Gain’ (red line) and ‘Loss’ (blue line) indicate regions with statistically significant difference in contacts with the bait regions identified by DESeq2 between alleles. Track ‘CTCF’ shows predicted CTCF binding sites on the sense (gold line) or antisense (black line) strand. The gene annotation is at the bottom of the figure. Mb = megabases. CPM = counts per million reads. M = maternal allele. P = paternal allele.

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

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

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

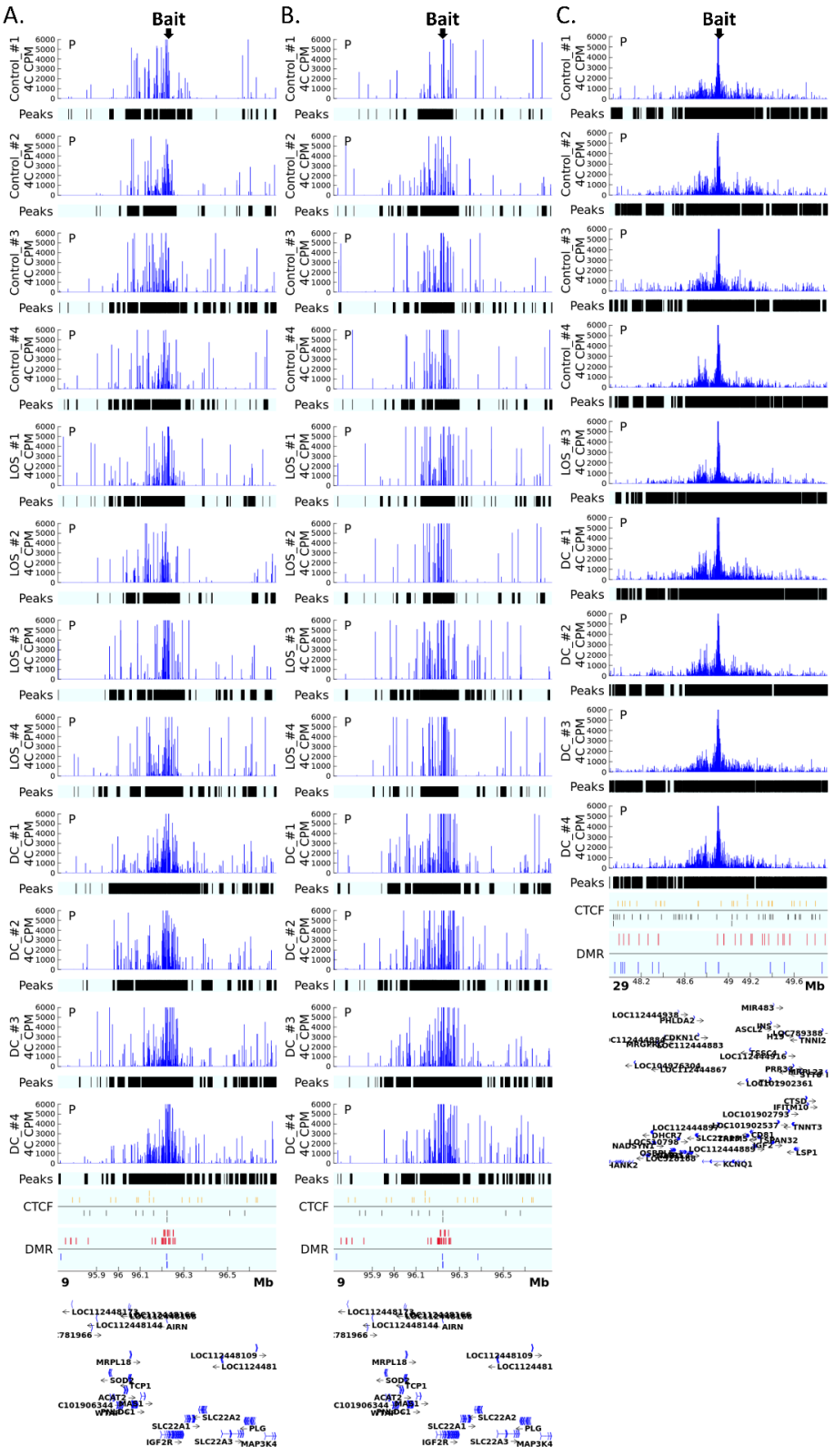


**Figure S8. 4C identified paternal allele cis contacts for individual samples, related**

**to Figure 3 and Figure 4.**

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

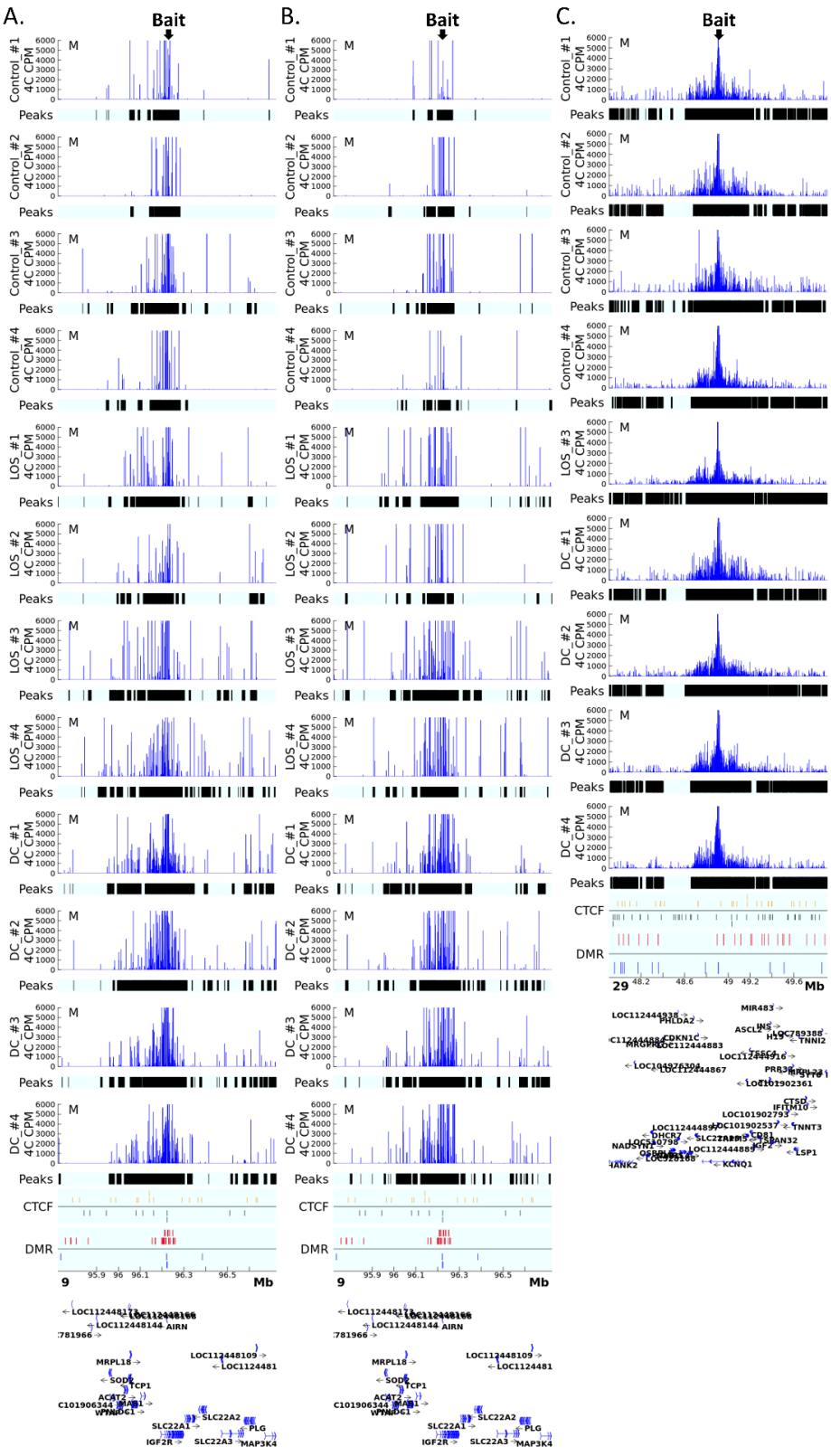


**Figure S9. 4C identified maternal allele cis contacts for individual samples,**

**related to Figure 3 and Figure 4.**

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

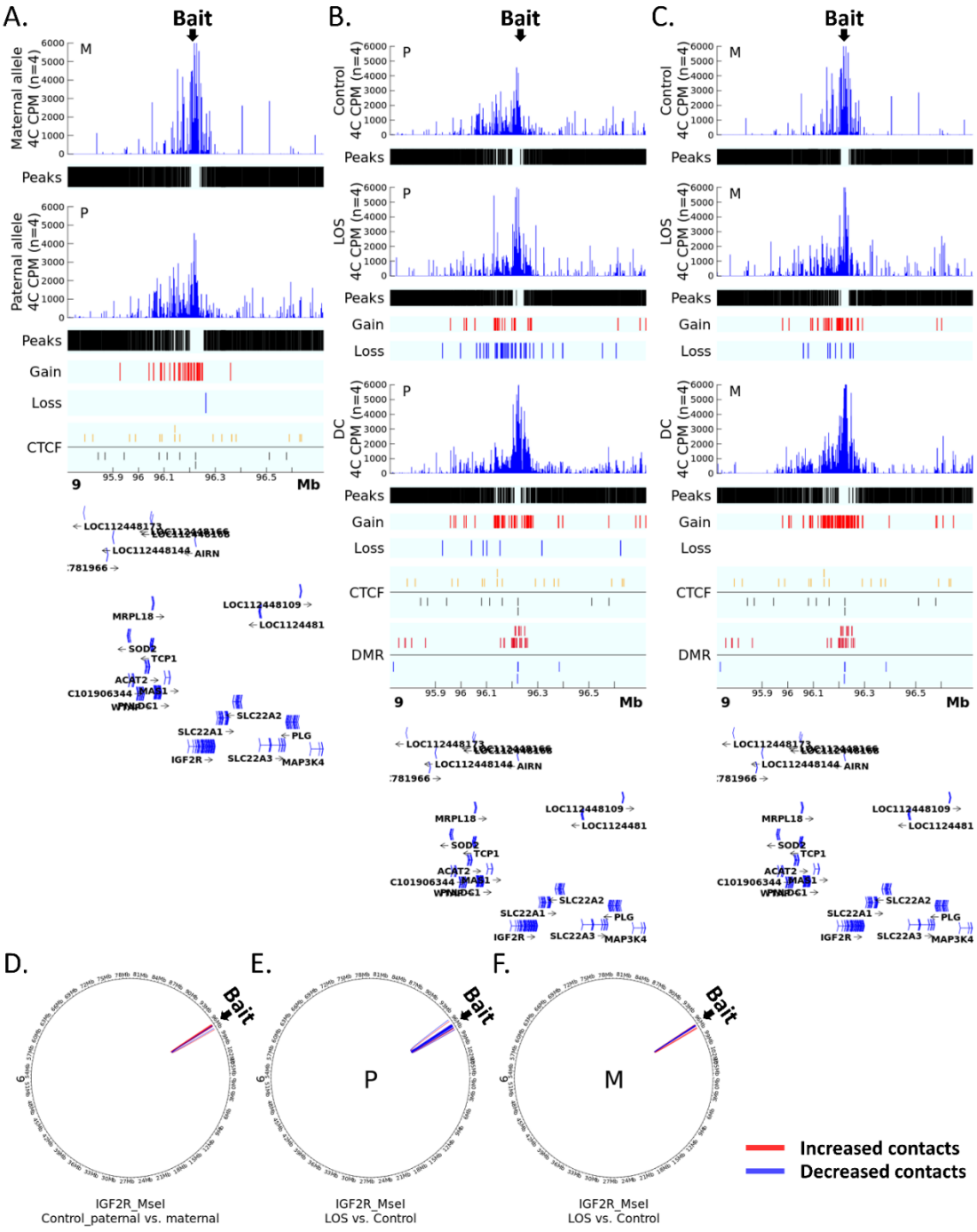


**Figure S10. 4Cker statistical results for allele-specific cis and far-cis contacts with *IGF2R* ICR, related to Figure 3.**

(A-F) Same IGF2R_MseI data as in Figure 3 A-D and Figure 5 A and G analyzed with 4Cker.

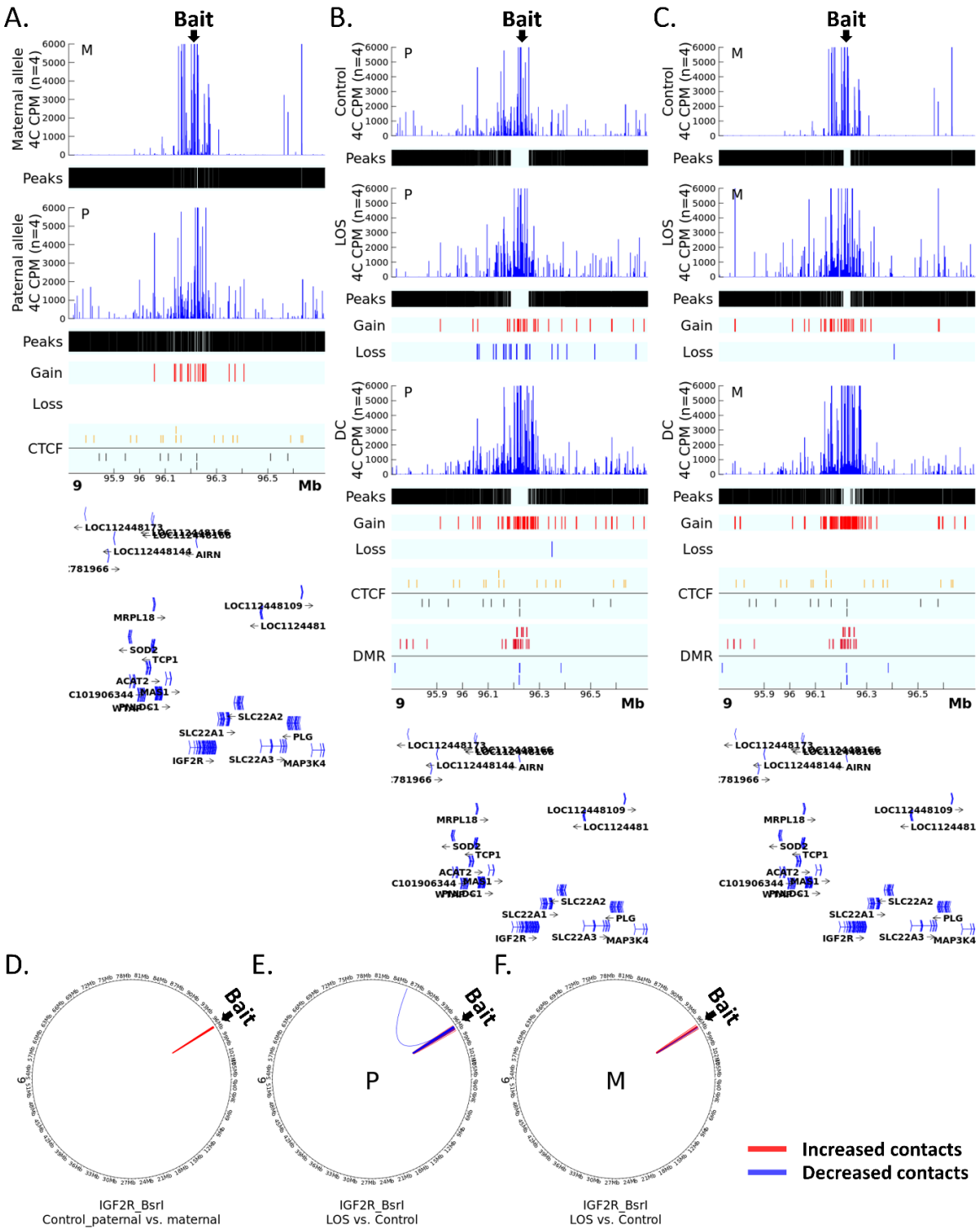


**Figure S11. 4Cker statistical results for allele-specific cis and far-cis contacts with *IGF2R* ICR, related to Figure 3.**

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

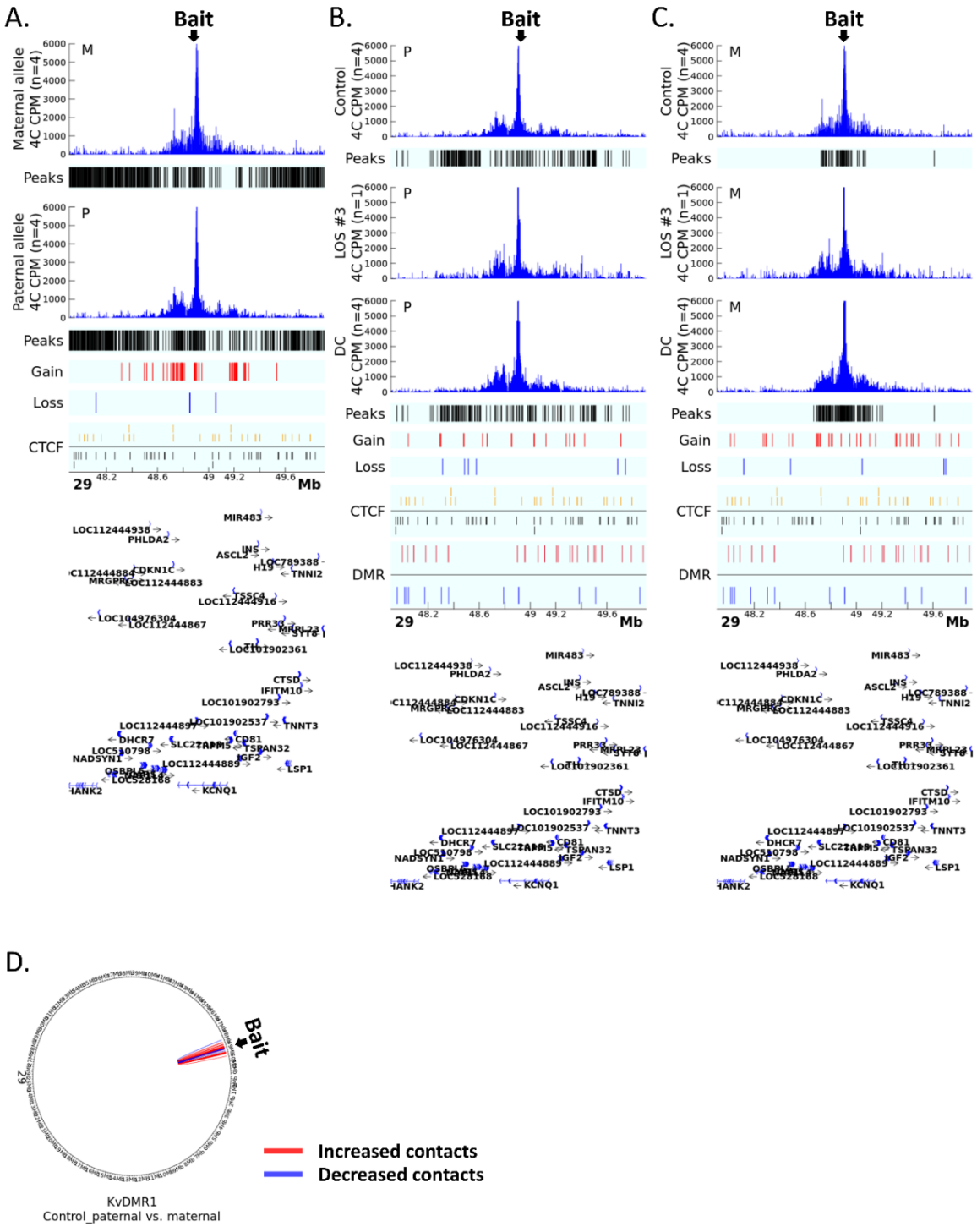


**Figure S12. 4Cker statistical results for allele-specific cis and far-cis contacts with KvDMR1, related to Figure 4.**

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

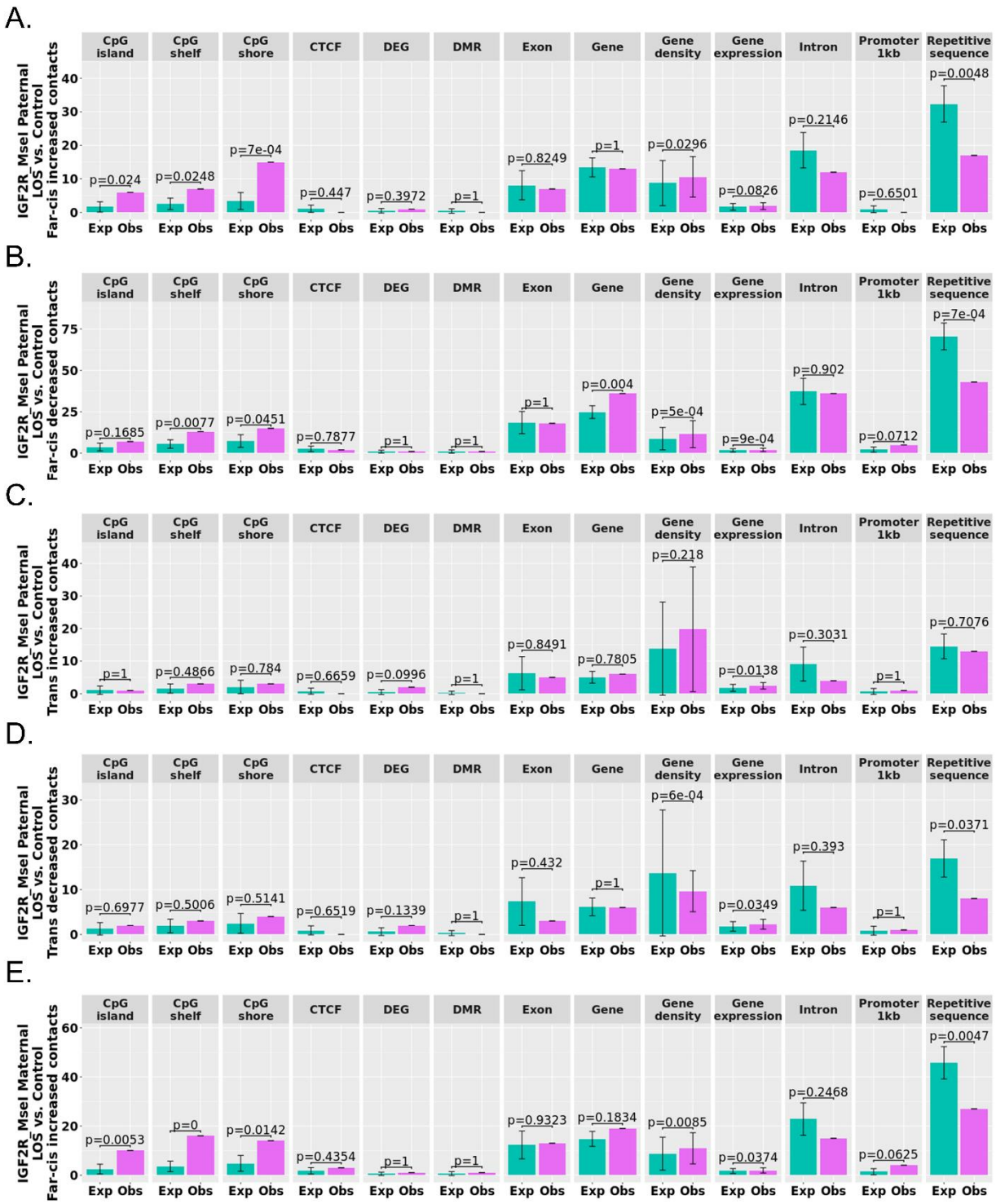


**Figure S13. Distribution of altered *IGF2R* ICR trans contact across various genomic contexts, related to Figure 5.** Results shown are for IGF2R_MseI.

(A-E) Observed and expected number of unique genomic context that overlaps with

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

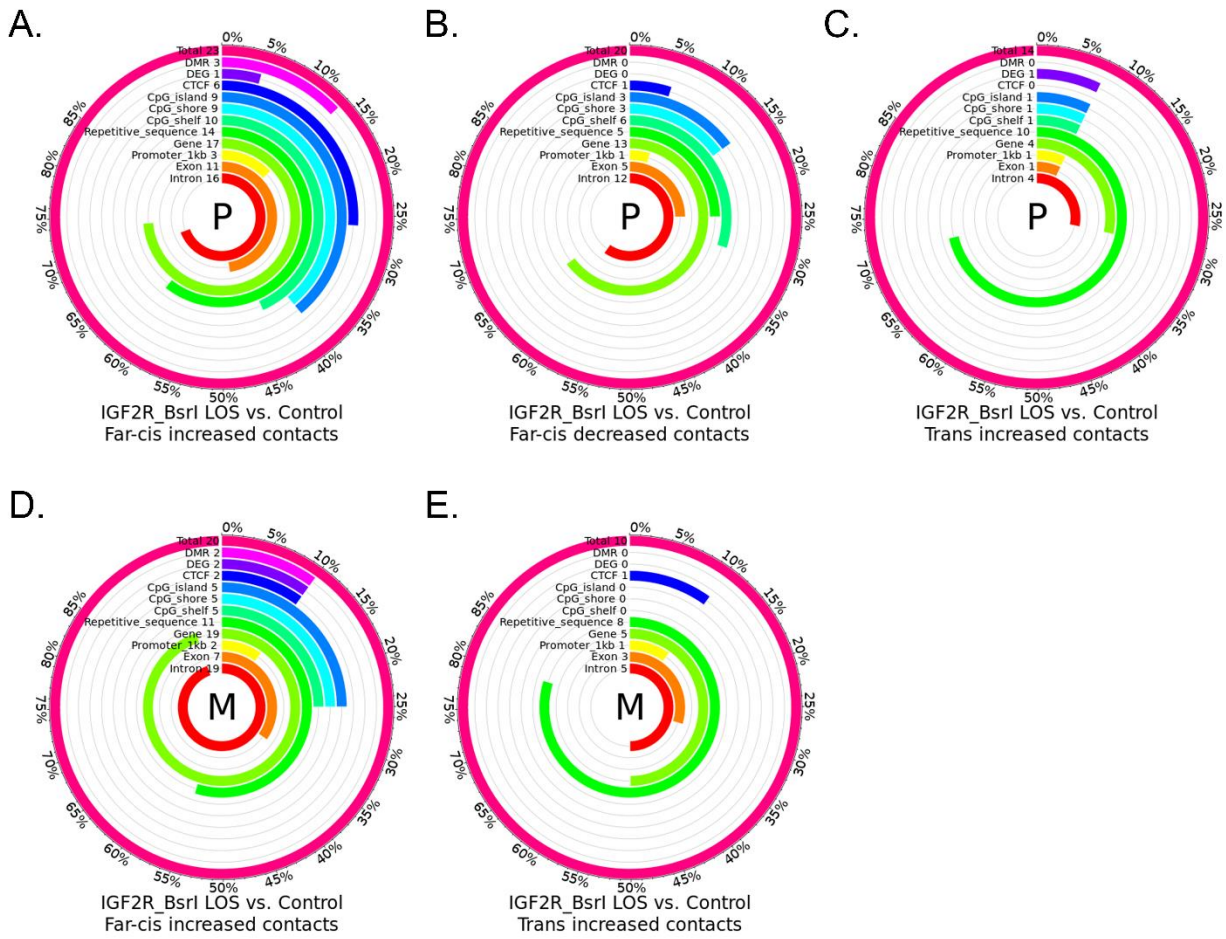


**Figure S14. Distribution of altered *IGF2R* ICR trans contact across various genomic contexts, related to Figure 5.** Shown results are for IGF2R_BsrI.

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

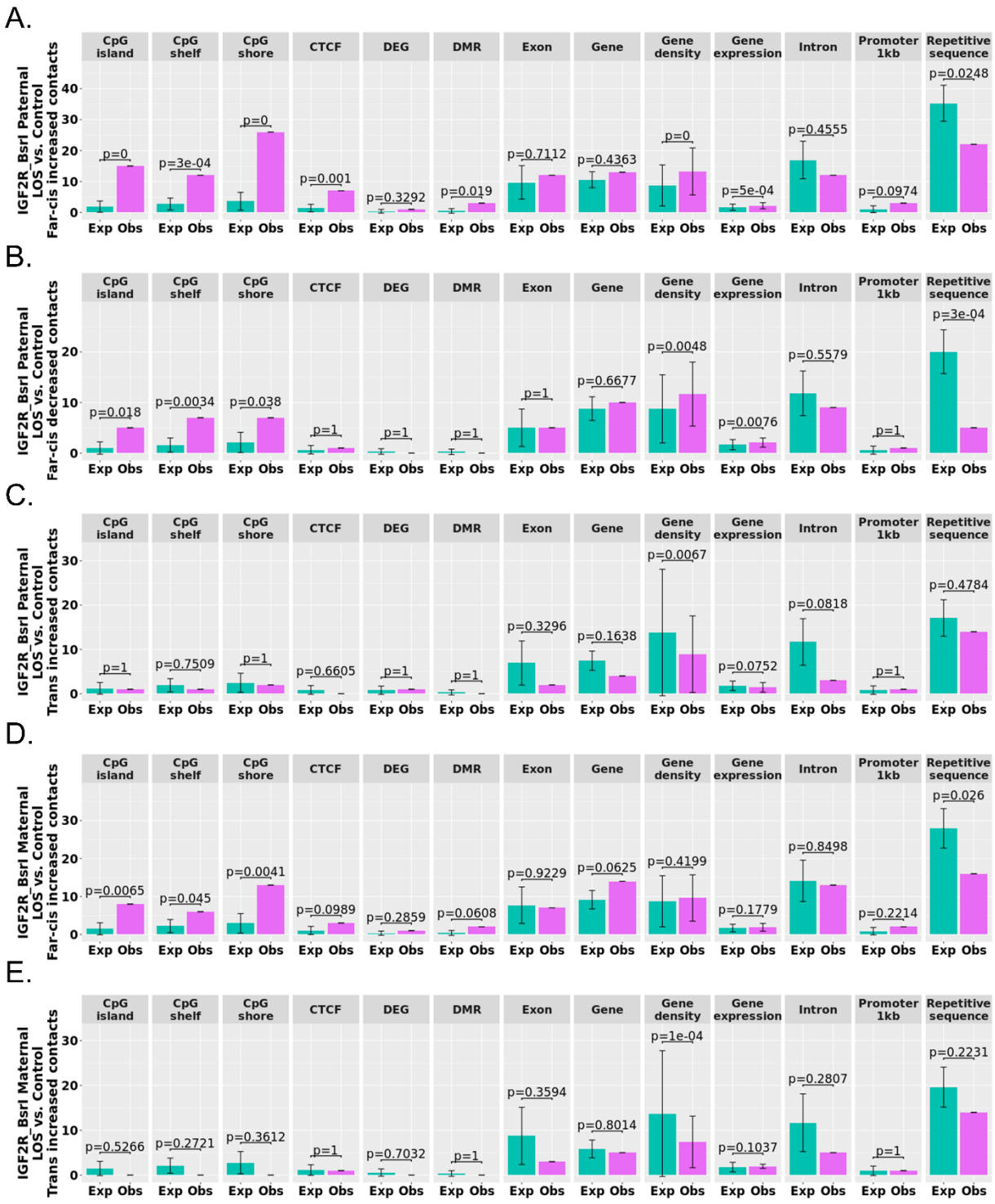


**Figure S15. Distribution of altered *IGF2R* ICR trans contact across various genomic contexts, related to Figure 5.** Shown results are for IGF2R_BsrI.

(A-E) Observed and expected number of unique genomic context that overlaps with

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

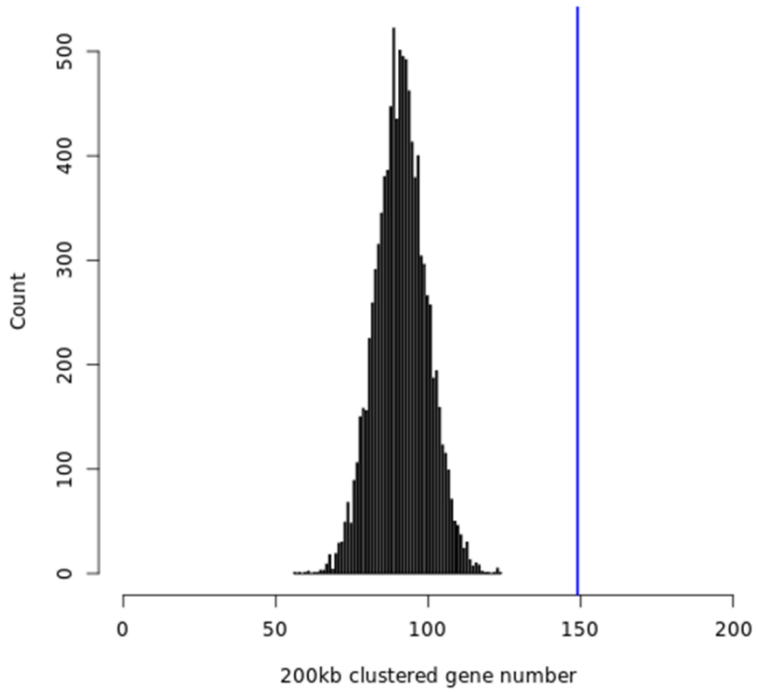


**Figure S16. Distribution of permutation test for clustering tendency of differentially expressed genes (DEGs), related to Figure 6.** Black lines show results of 10,000 times of permutating of 548 genes (same number as DEGs) against the 15,042 expressed genes in this study, and the 200kb clustered gene number (Exp) was calculated each time. Blue line indicates the number (Obs) of the 149 clustered DEGs identified in this work, which is significantly higher ($p = 0$) than the mean of permutation tests (91.46 with standard deviation 8.49). The p values were calculated as $p = n(|Exp - mean(Exp)| >= |Obs - mean(Exp)|)/10000$.