

Fig. S1. Stepwise epigenetic modification of cytosine. (a) DNA methyltransferase (DNMT) methylates cytosine to create (b) 5-methylcytosine (5mC), which ten-eleven translocation (TET) enzymes oxidize to create (c) 5-hydroxymethylcytosine (5hmC), which TET again oxidizes to create (d) 5-formylmethylcytosine (5fC). Finally, TET can further oxidize 5fC to (e) 5-carboxymethylcytosine (5caC), which can then return to an unmodified cytosine through decarboxylation or thymine DNA glycosylase (TDG) mediated excision, followed by base excision repair. R indicates deoxyribose and the rest of a DNA molecule. Purple rectangles indicate functional groups changed in the reaction.

Fig. S2 (preceding page). Relationship between unmodified versus modified motif statistical significance of central enrichment (from CentriMo⁵⁵) and modified base calling thresholds across different whole genome bisulfite sequencing (WGBS) and oxidativeWGBS (oxWGBS) specimens, in mice.⁵¹ We compared each unmodified motif, at each threshold, to its top three most significant modifications for c-Myc, C/EBP β , and replicate 2 of the C57BL/6 ZFP57 samples, but only the single most significant modification for all other ZFP57 samples. The displayed motif pairs changes at individual thresholds, depending on which motif pairs stay in the top three. We called all ZFP57 peaks, excepting Quenneville et al.²², using a stringent MACS⁶² threshold of $q = 0.00001$. Sign of value indicates preference for the unmodified (negative) motif or the modified (positive) motif. Rows: single chromatin immunoprecipitation-sequencing (ChIP-seq) replicates for a particular transcription factor target, consisting of: 3 c-Myc (Krepelova et al.⁵⁹, ENCF001YHU, and ENCF001YJE), 6 ZFP57 (from Quenneville et al.²², and BC8 and CB9 cells, from Strogantsev et al.²³), and 3 C/EBP β (ENCF001XUR, ENCF001XUS, and ENCF001XUT). Columns: replicates of WGBS and oxWGBS. We additionally depict C/EBP β alone in Fig. S4.

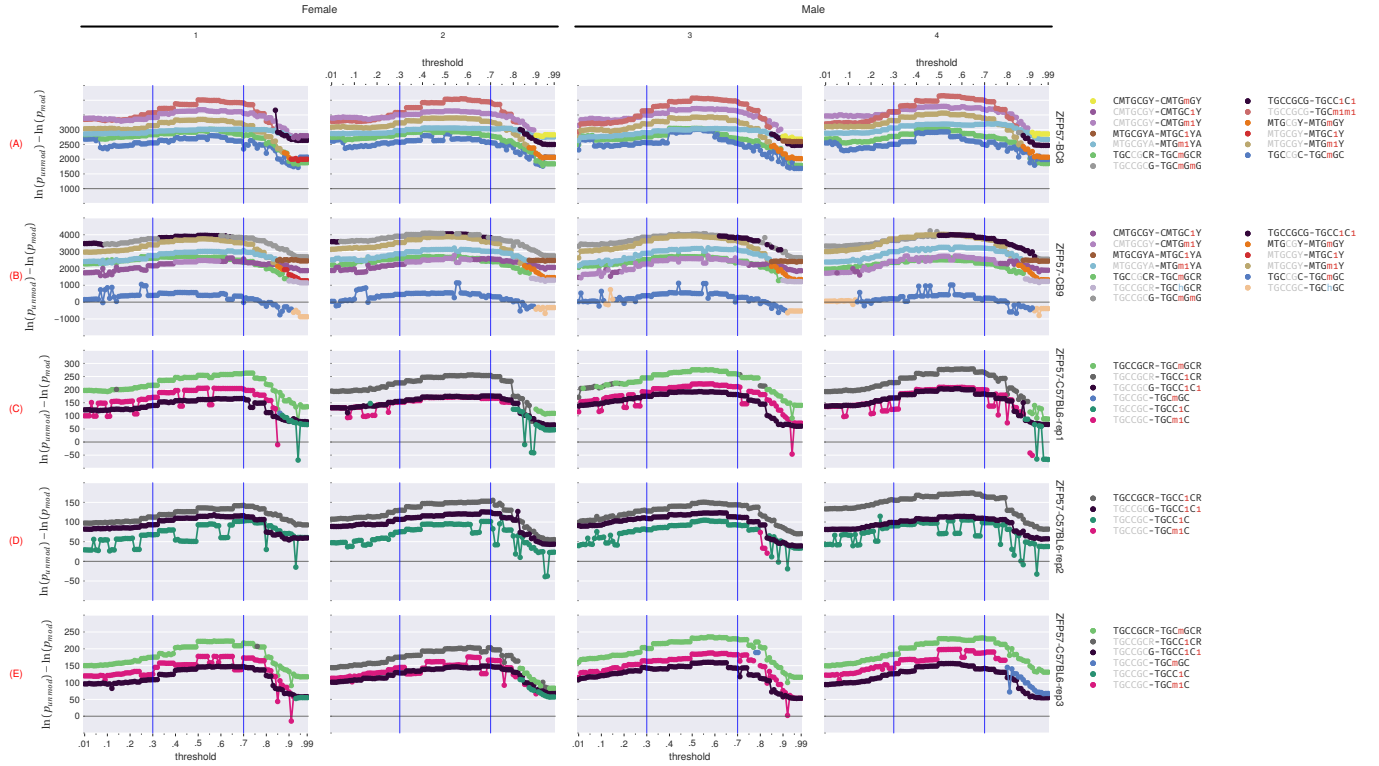


Fig. S3. Relationship between unmodified versus modified ZFP57 statistical significance of central enrichment (from CentriMo⁵⁵) and modified base calling thresholds across different WGBS and oxWGBS specimens, in mice.⁵¹ We compared each unmodified motif, at each threshold, to its top most significant modification. The displayed motif pairs changes at individual thresholds, depending on which motif pairs stay at the top. We called all ZFP57 peaks using the default MACS⁶² threshold of $q = 0.05$. Sign of value indicates preference for the unmodified (negative) motif or the modified (positive) motif. Rows: single ChIP-seq replicates for a particular transcription factor target, consisting of all ZFP57 replicates (from Quenneville et al.²², and BC8 and CB9 cells, from Strogantsev et al.²³). Columns: replicates of WGBS and oxWGBS.

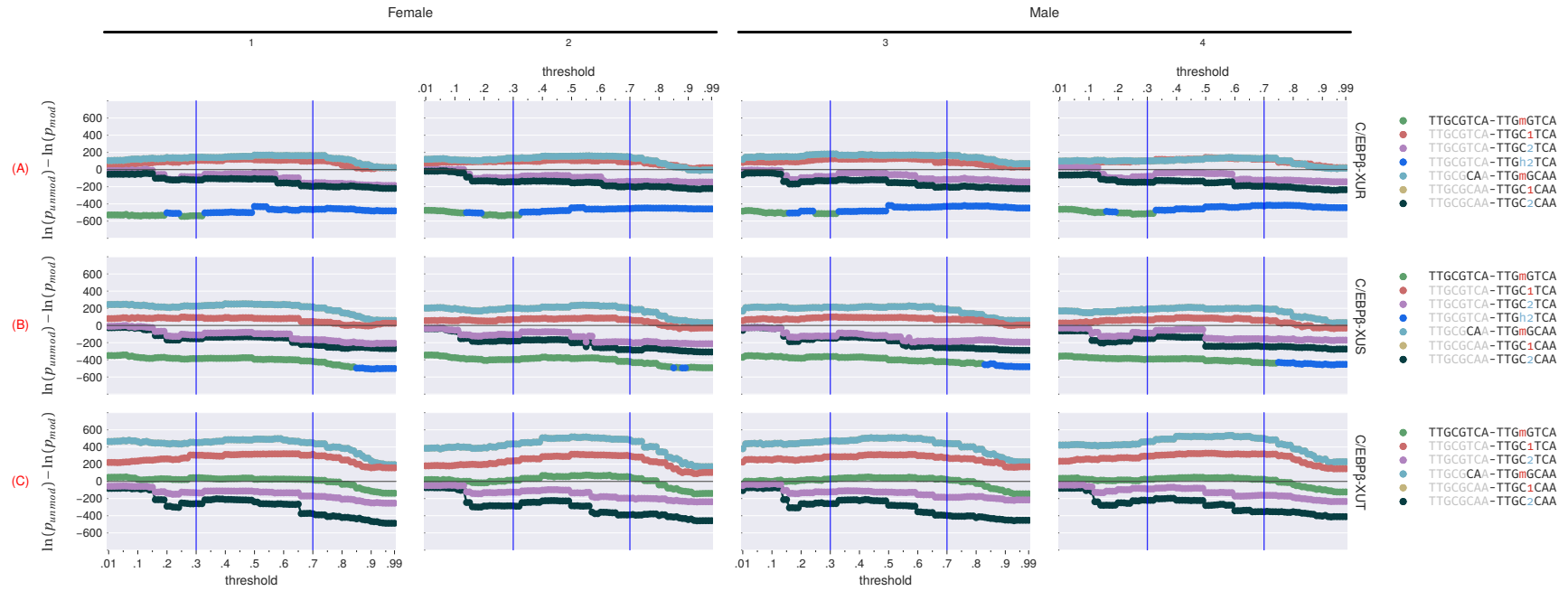


Fig. S4. Relationship between unmodified versus modified C/EBP β statistical significance of central enrichment (from CentriMo⁵⁵) and modified base calling thresholds across different WGBS and oxWGBS specimens, in mice.⁵¹ We compared each unmodified motif, at each threshold, to its top three most significant modifications. The displayed motif pairs changes at individual thresholds, depending on which motif pairs stay in the top three. Sign of value indicates preference for the unmodified (negative) motif or the modified (positive) motif. Rows: single ChIP-seq replicates for a particular transcription factor target, consisting of all C/EBP β replicates (ENCFF001XUR, ENCFF001XUS, and ENCFF001XUT). Columns: replicates of WGBS and oxWGBS. We additionally depict all of our other tested transcription factors in Fig. S2.

(A)

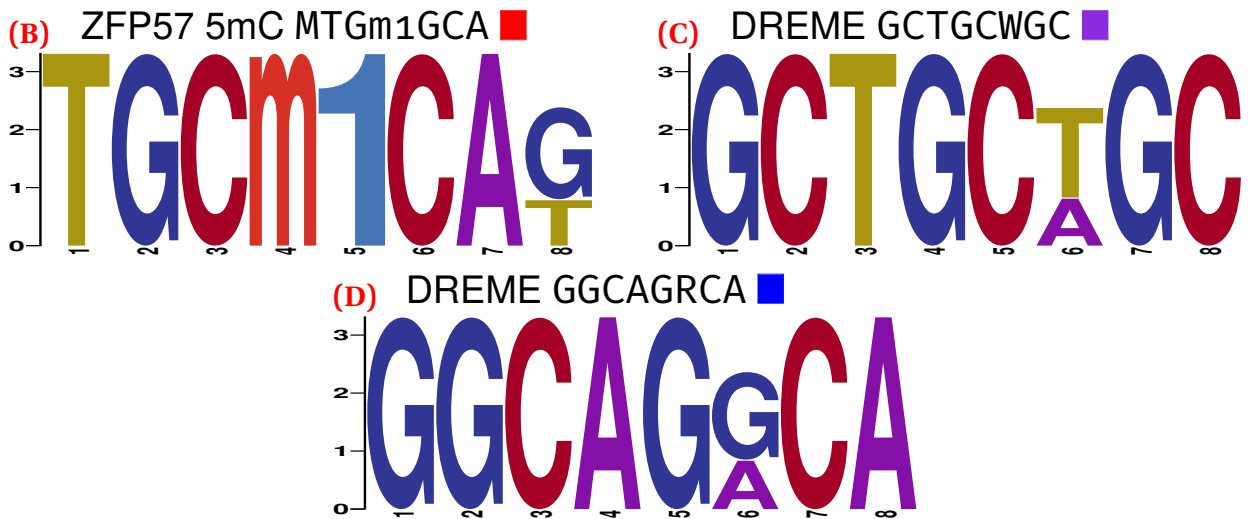
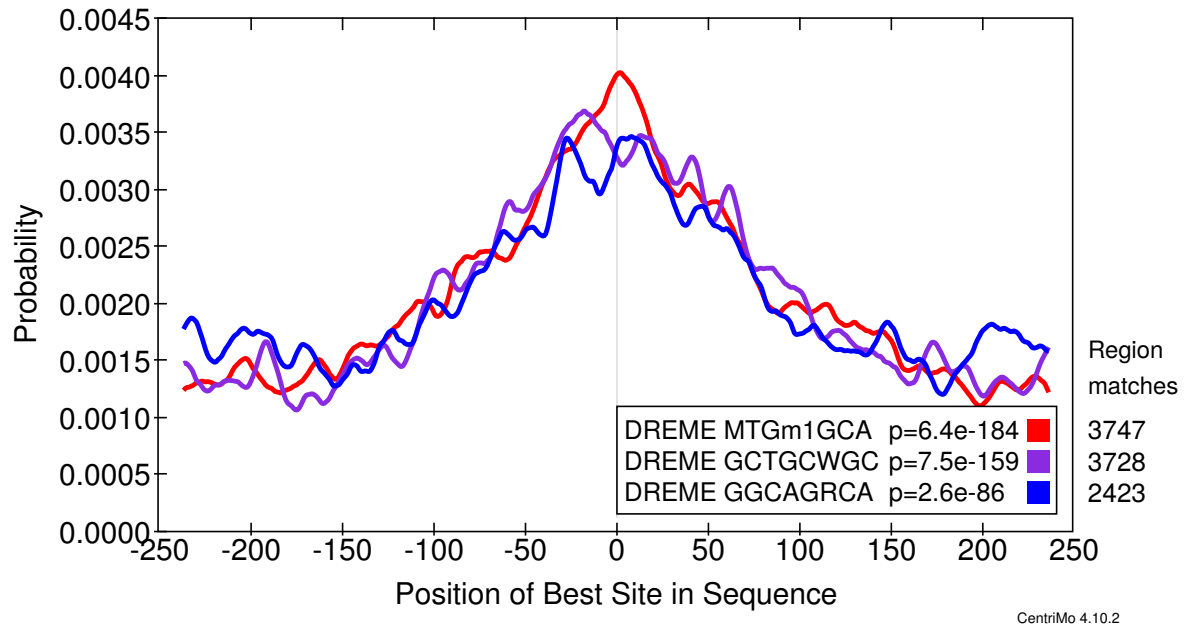
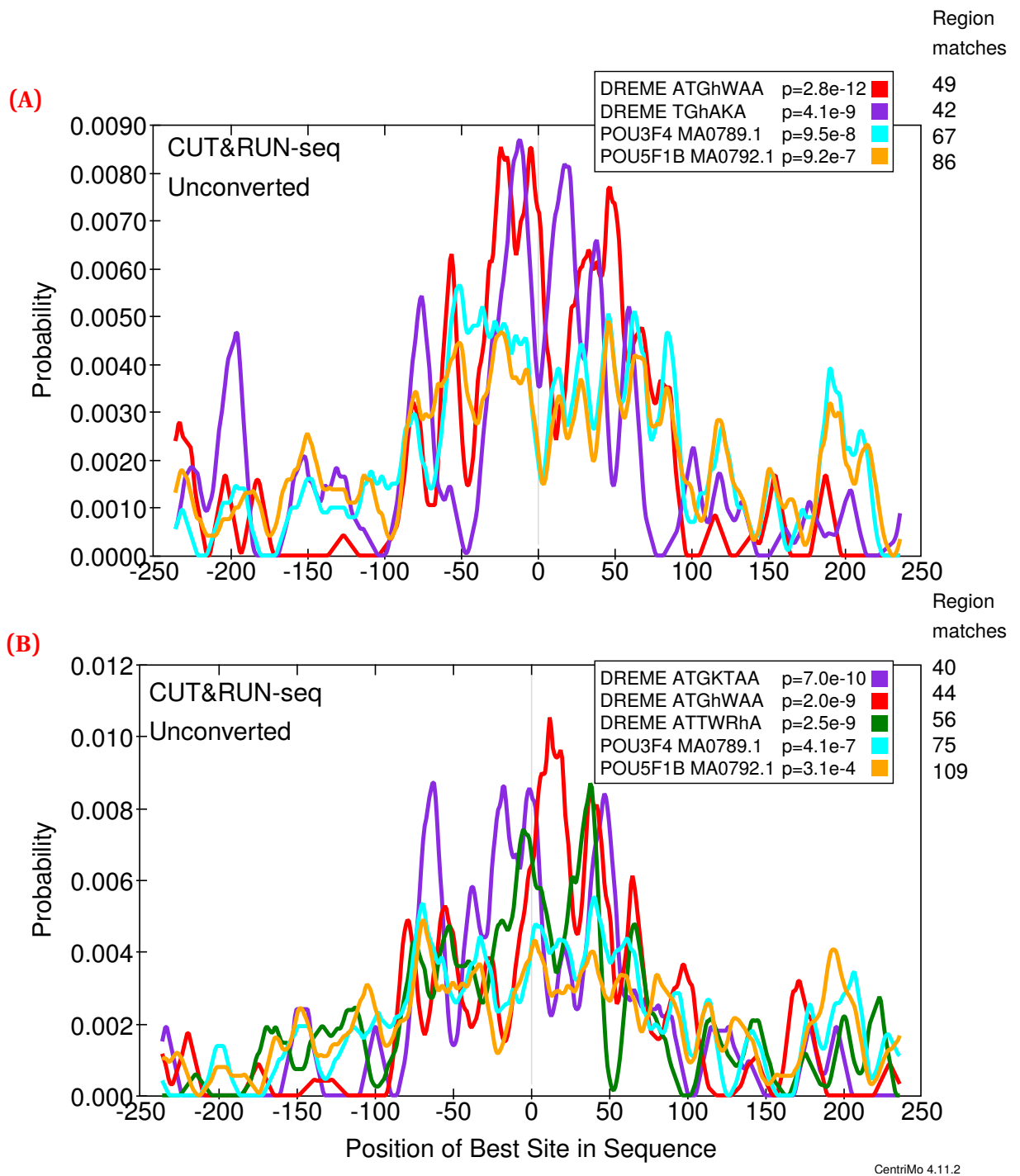


Fig. S5. ZFP57 (Strogantsev et al.²³ CB9; 56 142 ChIP-seq peaks) CentriMo analysis of *de novo* and JASPAR motifs (Methods). Listed p-values computed by CentriMo.⁵⁵ Depicts female replicate 2 of the combined WGBS and oxWGBS data⁵¹ at a 0.7 modification threshold. (A) the CentriMo result with an expected ZFP57 methylated motif (red), top DREME unmodified motif (purple), and second DREME unmodified motif (blue). (B) Sequence logo of the reverse complement of the CentriMo result with an expected ZFP57 methylated motif. (C) Sequence logo of the top DREME unmodified motif. (D) Sequence logo of the second DREME unmodified motif.



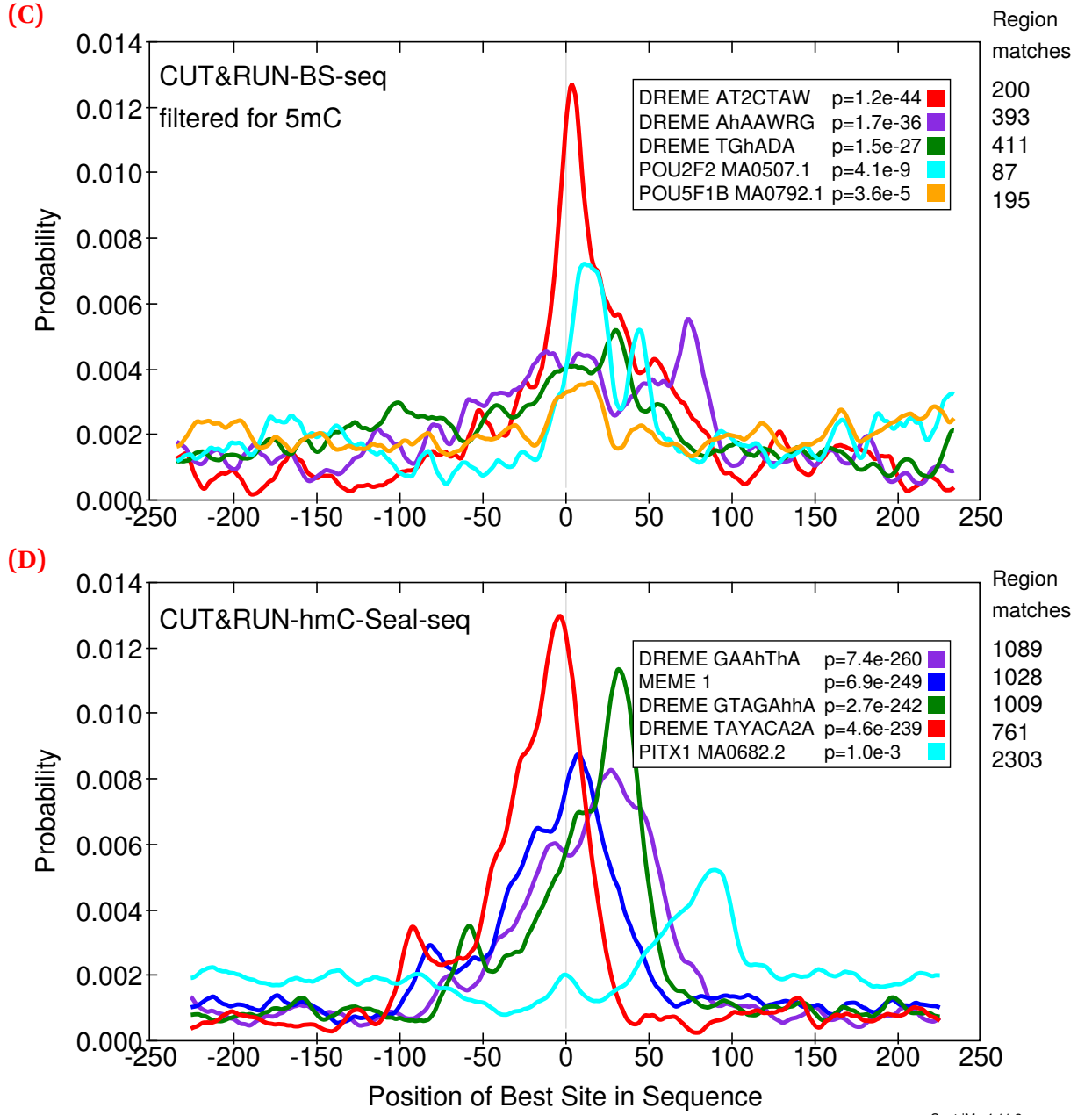


Fig. S6. CentriMo⁵⁵ results for OCT4 cleavage under targets and release using nuclease (CUT&RUN) in mouse embryonic stem cells (mESCs). Motifs include the top three DREME motifs with colour indicating rank: first (red); second (purple), third, where applicable (dark green). Motifs also include the top non-POU5F1 JASPAR motif (cyan), and the JASPAR POU5F1B motif (orange). Both of these motifs come from the JASPAR 2020⁶¹ core vertebrate set. We generated these results using 500 bp regions centred upon the summits of MACS 2⁶² peaks generated from those CUT&RUN fragments ≤ 120 bp. We called peaks using IgG controls and without any spike-in calibration (Methods). Listed p-values computed by CentriMo.⁵⁵ For consistency, we depict the JASPAR sequence logo using MEME's relative entropy calculation and colouring. Also depicts the top MEME⁶³ *de novo* motif (blue). We depict the first replicates of 5mC and 5hmC data in Fig. 5. (A) replicate 1 unconverted (227 CUT&RUN peaks) sequences. (B) replicate 2 unconverted (265 CUT&RUN peaks) sequences. (C) replicate 2 bisulfite-converted (methylated; 2077 CUT&RUN peaks) sequences. (D) replicate 2 hmC-Seal (11 615 CUT&RUN peaks) sequences.

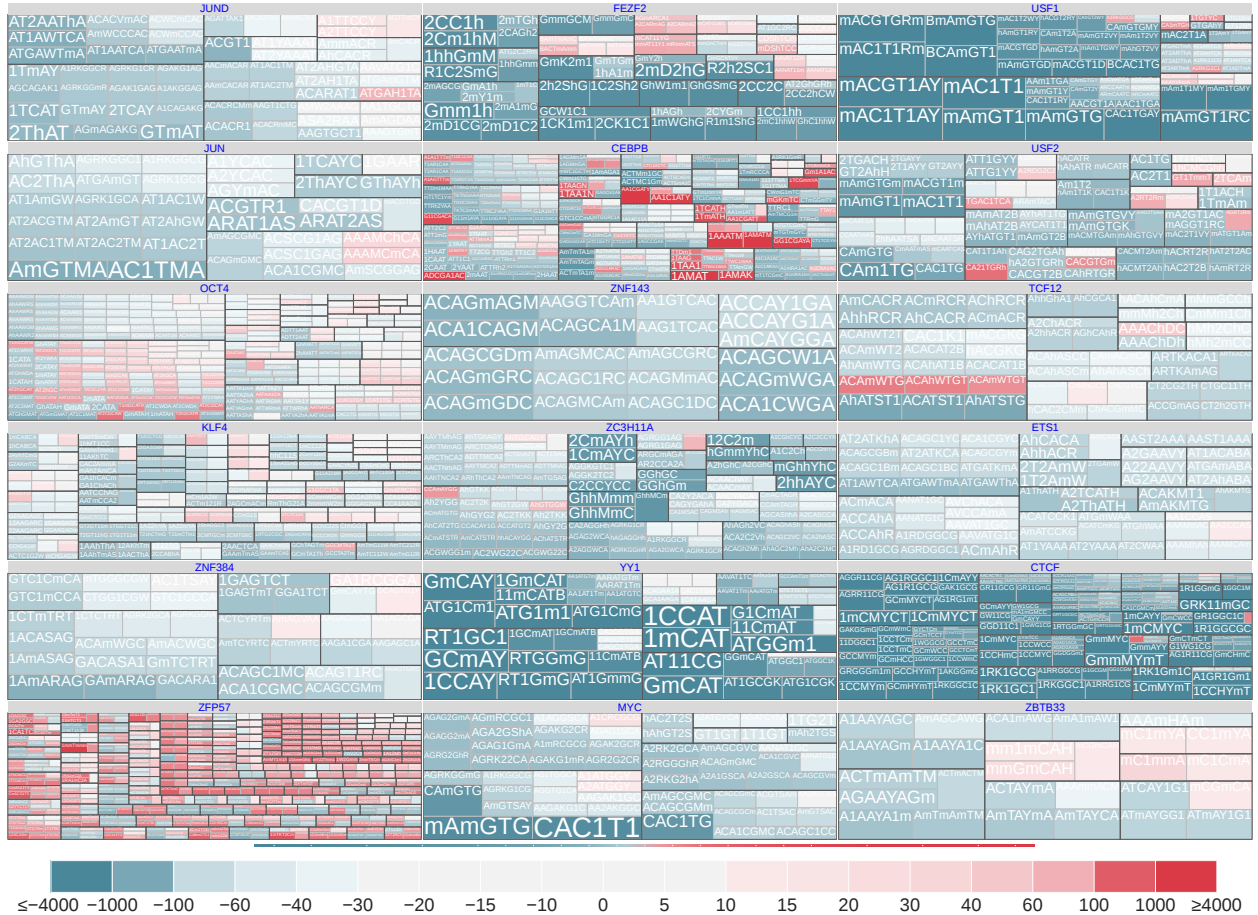


Fig. S7. Modified versus unmodified motifs, combining score and cluster information, for a wide array of transcription factors. These plots come from non-spike-in calibrated data, for the 500 bp regions surrounding peak summits. We clipped scores above or below ± 4000 and plotted them at a threshold that maximizes dynamic range where most scores occur. Some combinations of the displayed hypothesis pairs have multiple data points. For example, these points can correspond to multiple identical hypothesis pairs, but for different data sub-types or stringencies. We aggregated these multiple data points by plotting the maximum score. An asymmetric, diverging, colour scale further highlights modification-preferring motifs.