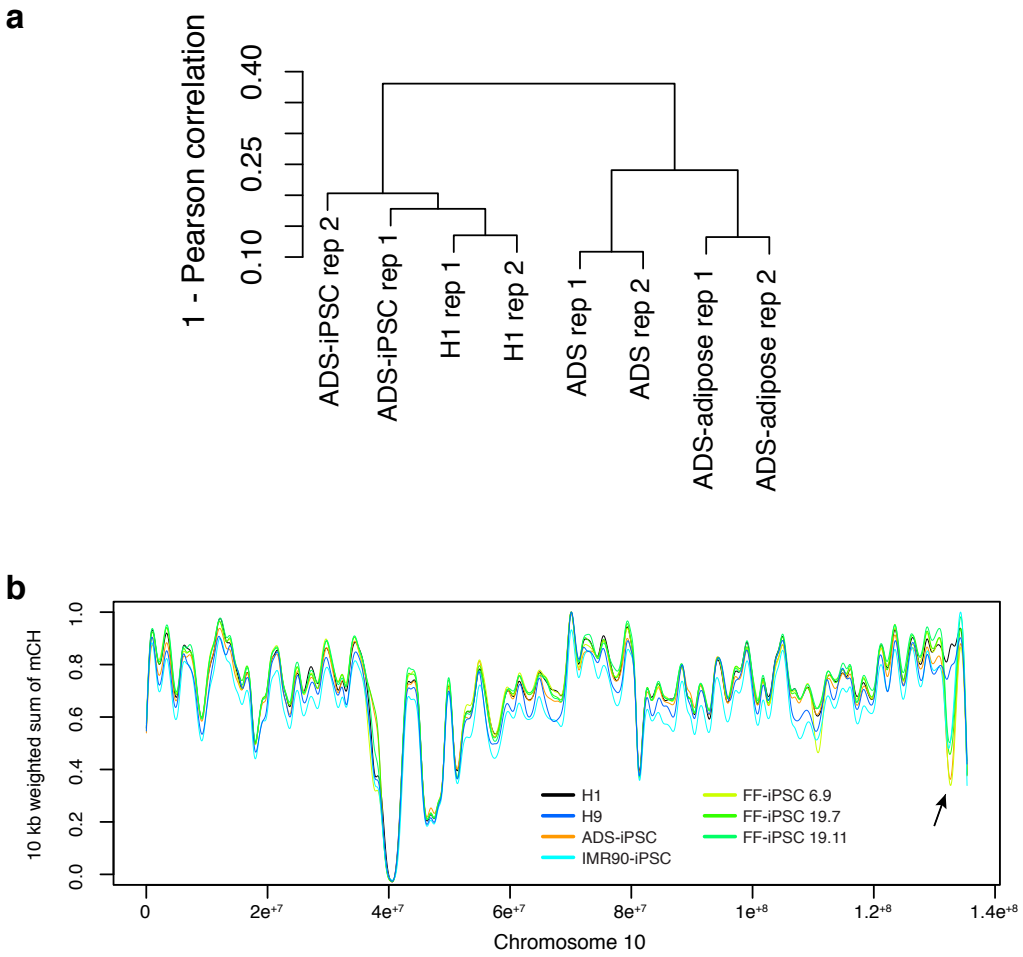
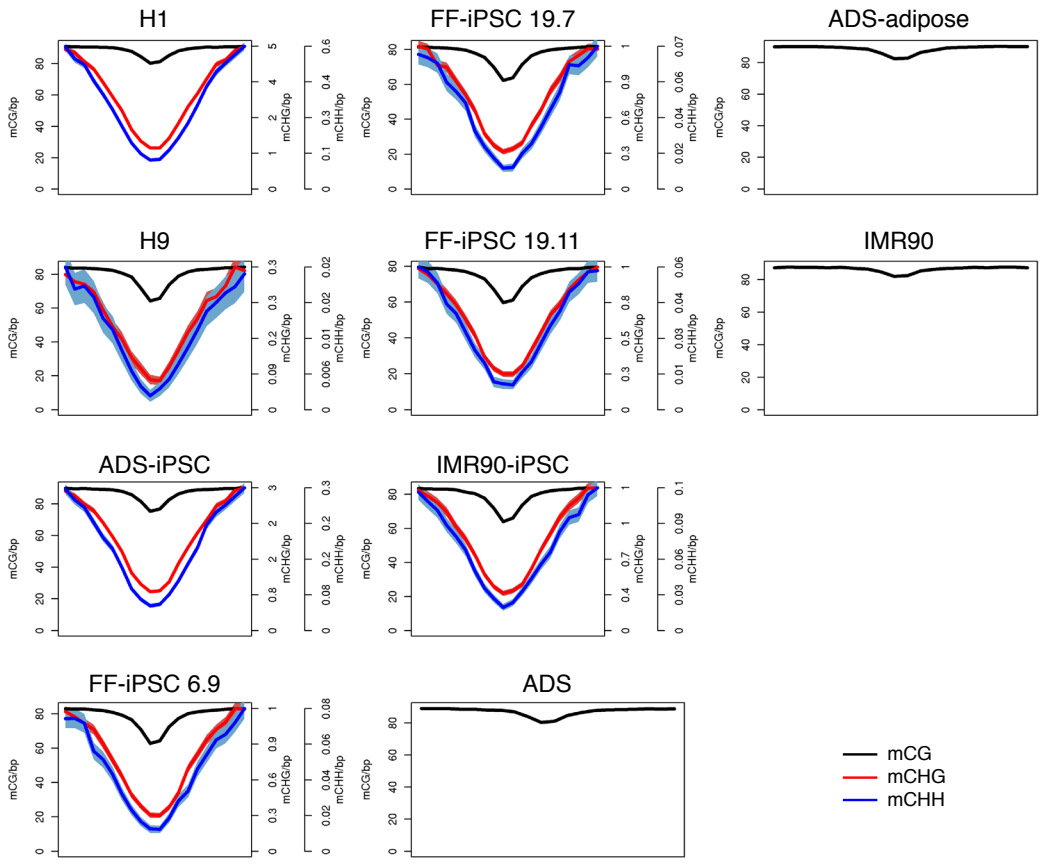
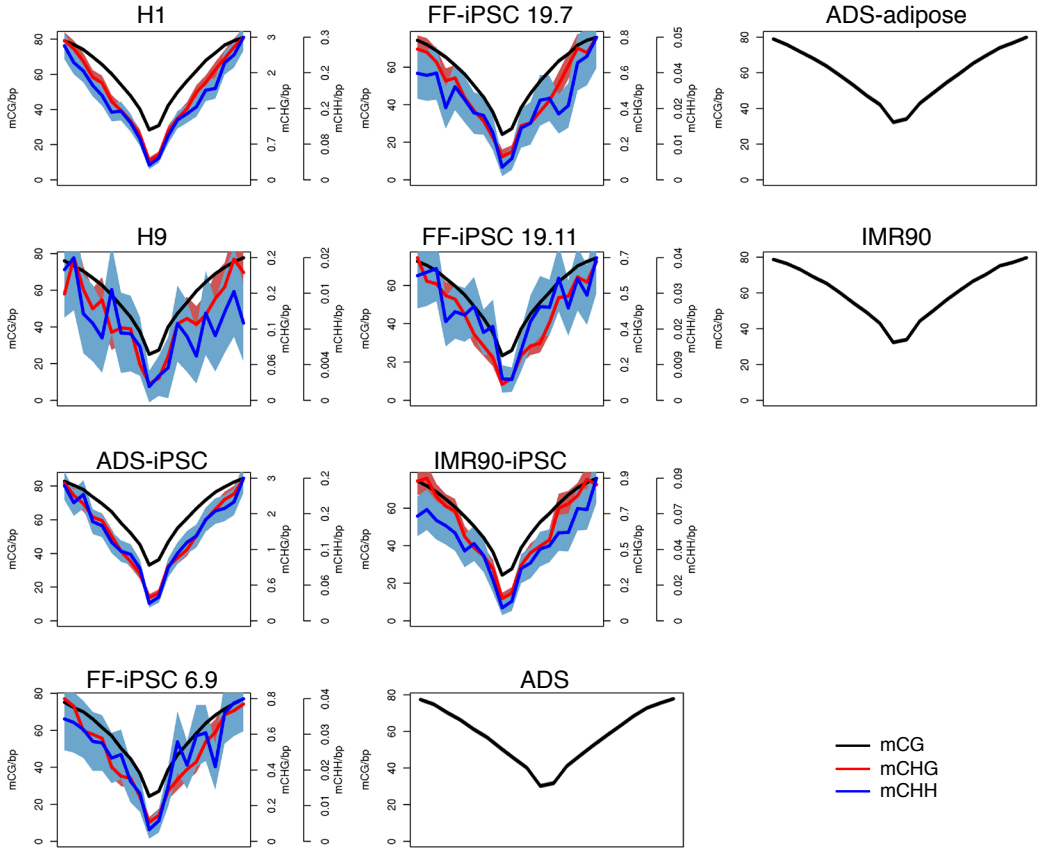
doi:10.1038/nature09798



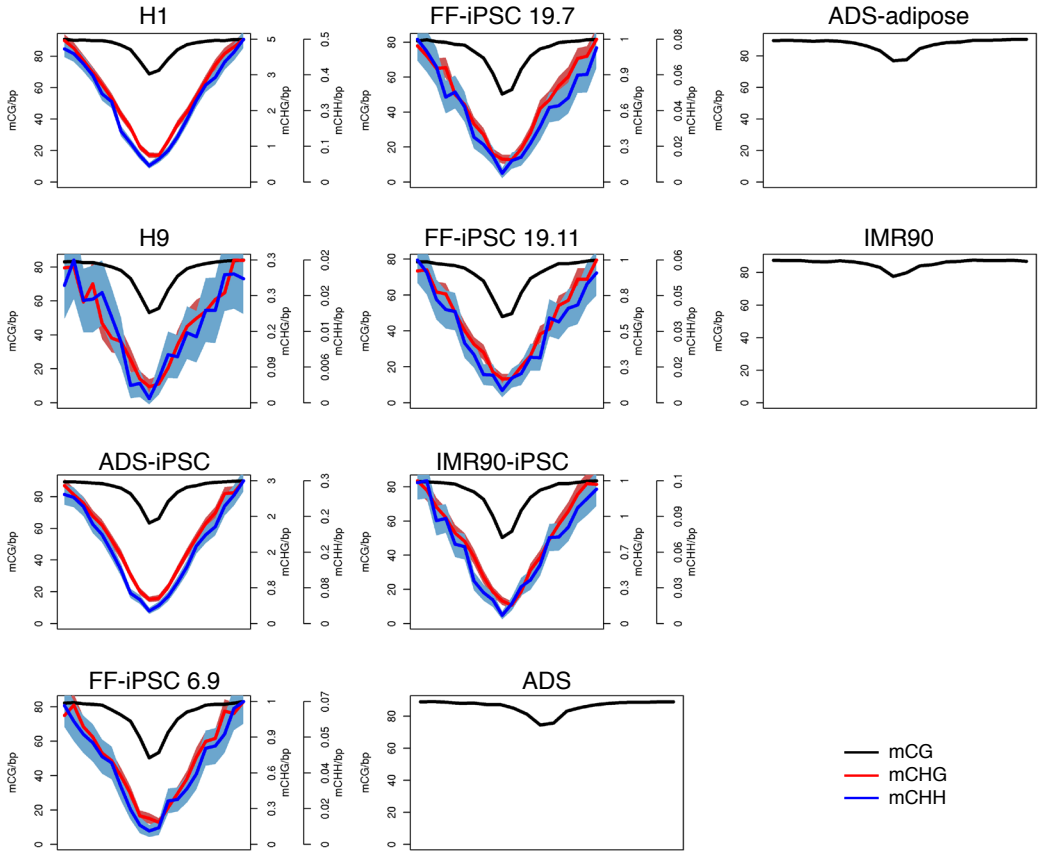
Supplementary Figure 1 | Global similarity in transcriptional and DNA methylation patterns between ESCs and iPSCs. a, Complete linkage hierarchical clustering of both biological replicates of ADS-adipose somatic cells, ADS cells, ADS-iPSCs and H1 ESCs, based on RPKM values from strand-specific RNA-seq. **b**, Plot of the density of mCH identified in chromosome 10. Lines represent smoothing of mCH density in 10 kb windows. Black arrow indicates a large region of dissimilarity between ESCs and iPSCs. Abbreviations: mCG/mCH, methylated cytosine in the CG/CH context.



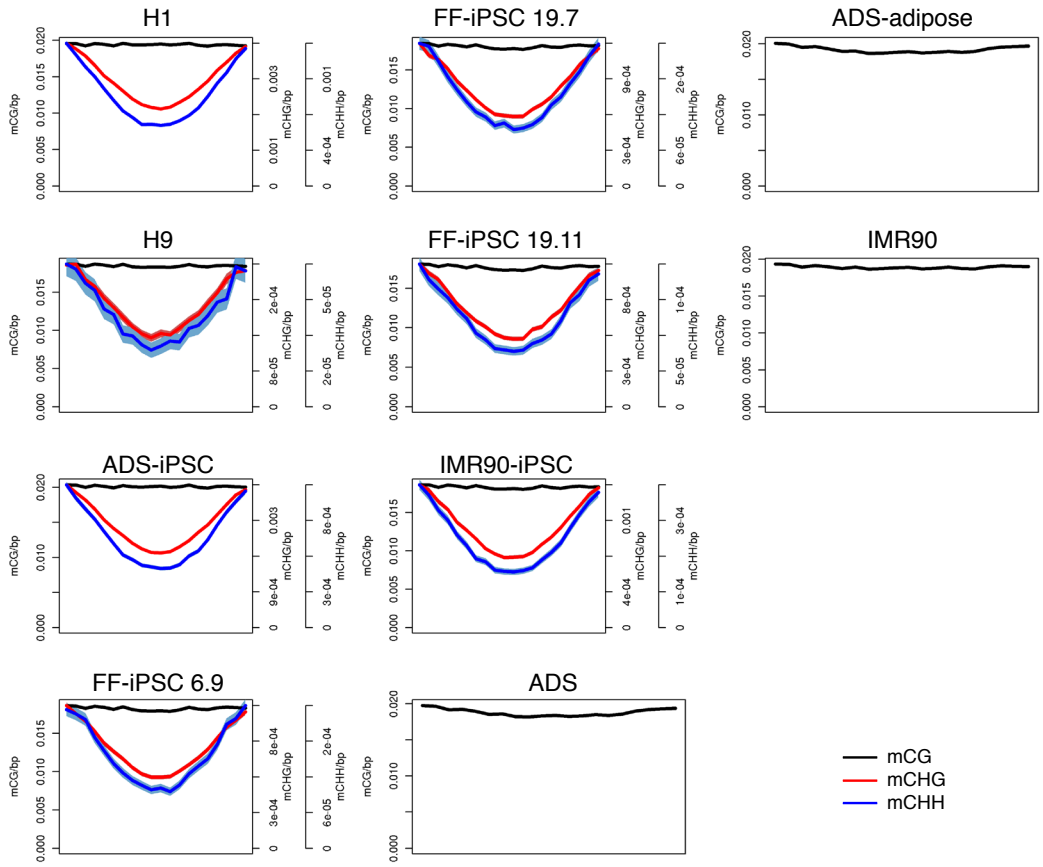
Supplementary Figure 2 | Density of DNA methylation at NANOG binding sites. The average relative DNA methylation densities in each sequence context are shown from 1.5 kb upstream to 1.5 kb downstream of the predicted sites of DNA-protein interaction of NANOG identified by ChIP-seq in H1 embryonic stem cells. Shaded areas indicate the 95th confidence interval for the mean.



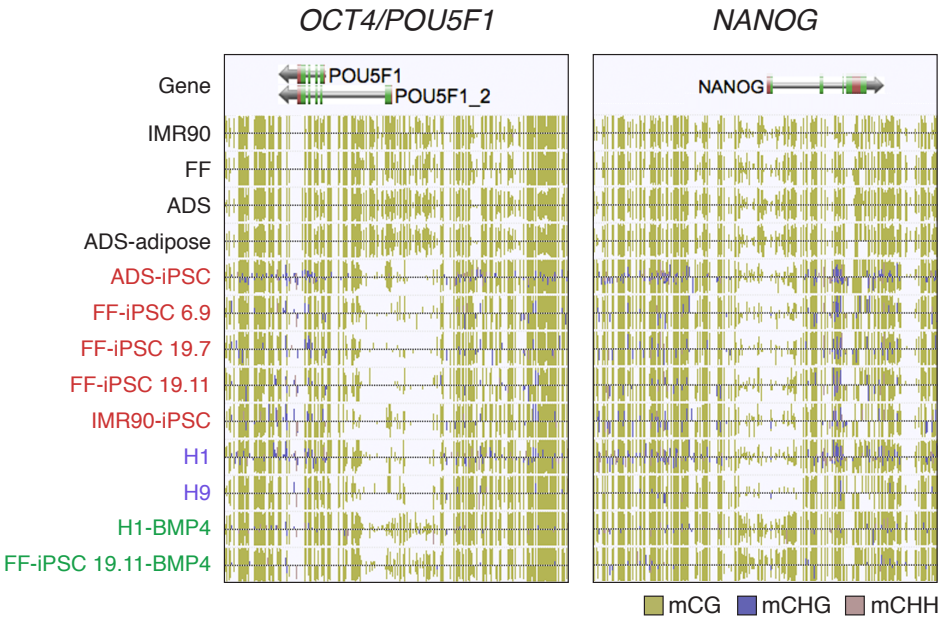
Supplementary Figure 3 | Density of DNA methylation at OCT4 binding sites. The average relative DNA methylation densities in each sequence context are shown from 1.5 kb upstream to 1.5 kb downstream of the predicted sites of DNA-protein interaction of OCT4 identified by ChIP-seq in H1 embryonic stem cells. Shaded areas indicate the 95th confidence interval for the mean.



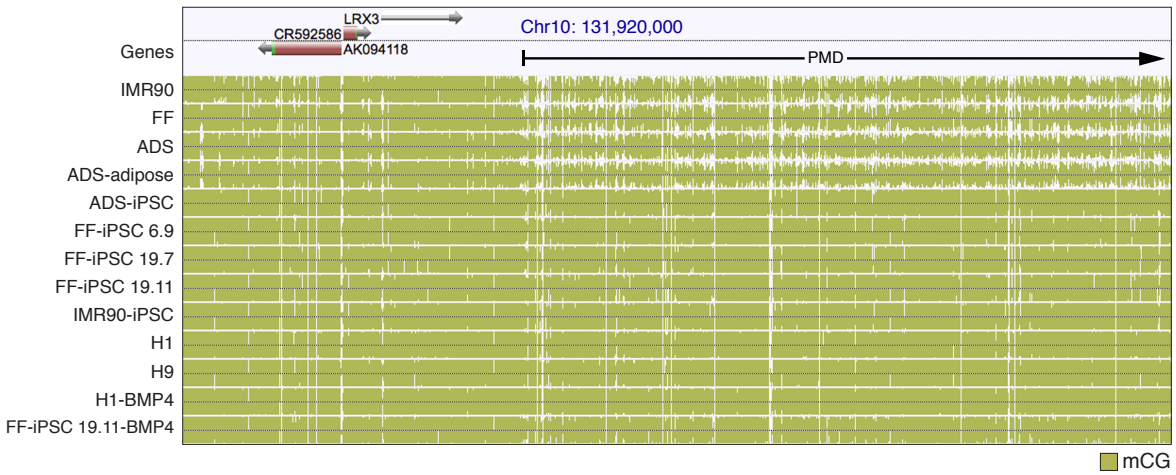
Supplementary Figure 4 | Density of DNA methylation at SOX2 binding sites. The average relative DNA methylation densities in each sequence context are shown from 1.5 kb upstream to 1.5 kb downstream of the predicted sites of DNA-protein interaction of SOX2 identified by ChIP-seq in H1 embryonic stem cells. Shaded areas indicate the 95th confidence interval for the mean.



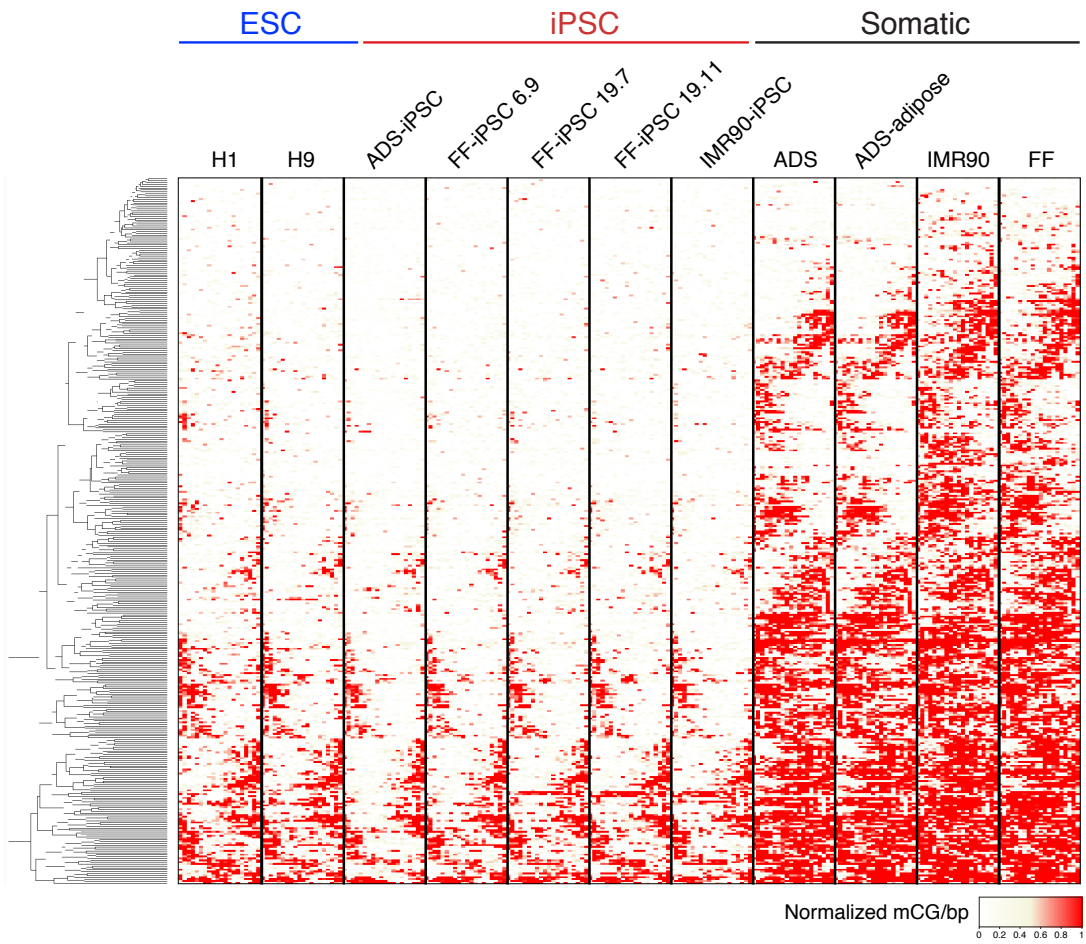
Supplementary Figure 5 | Density of DNA methylation at enhancer sites. The average relative DNA methylation densities in each sequence context in 100 bp windows are displayed throughout 5 kb upstream to 5 kb downstream of enhancers identified in H1 embryonic stem cells. Shaded areas indicate the 95th confidence interval for the mean.



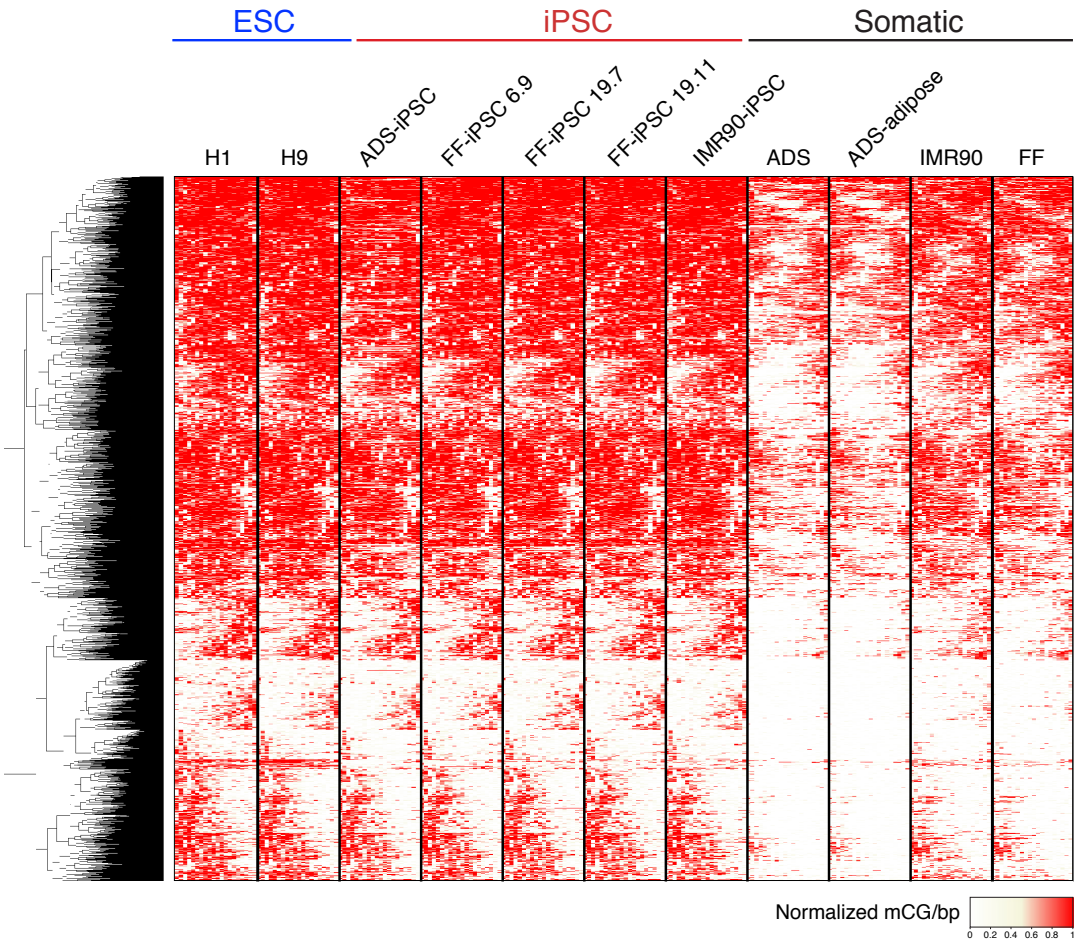
Supplementary Figure 6 | DNA methylation patterns at pluripotency-related genes. AnnoJ data browser representation of DNA methylation in all cell lines at pluripotency-related genes.



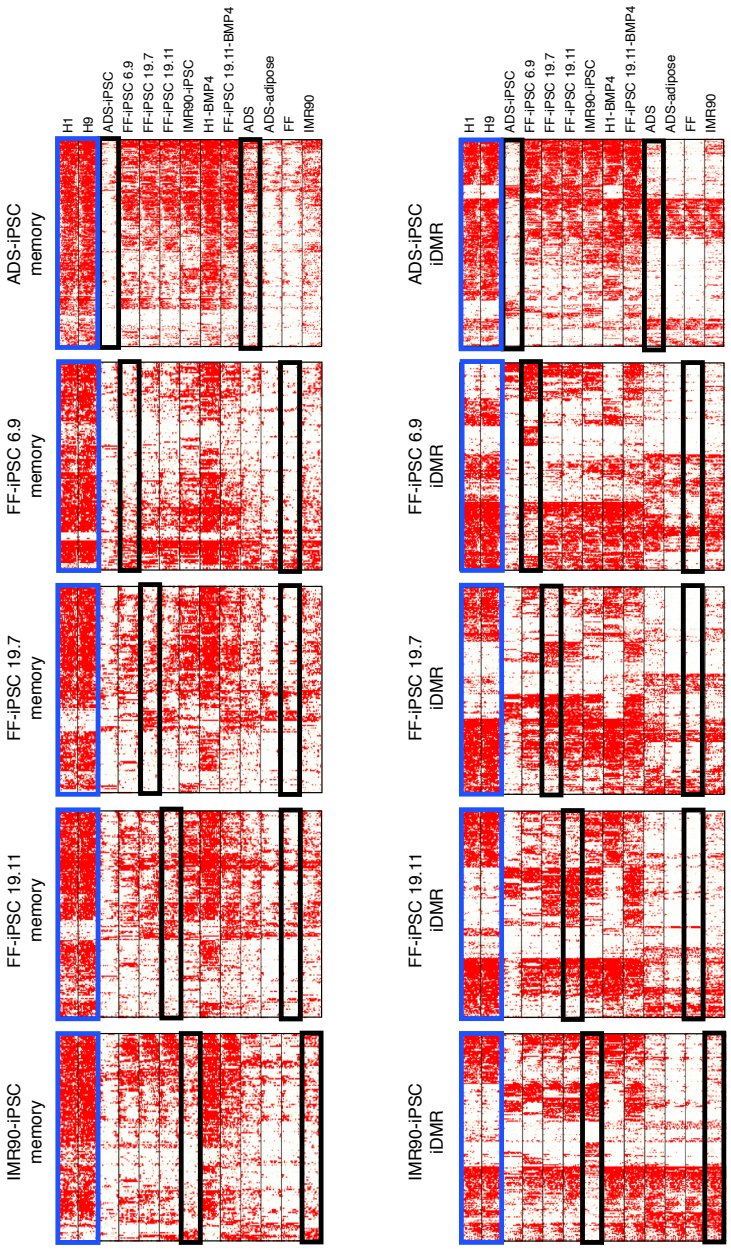
Supplementary Figure 7 | Restoration of fully-methylated state in iPSCs at somatic PMDs. AnnoJ browser representation of DNA methylation at a PMD genomic region that is partially methylated in somatic cell types, but fully methylated in both ESCs and iPSCs. For the DNA methylation tracks, vertical lines above and below the dotted central line indicate the presence of methylcytosines on the Watson and Crick strands, respectively. Only DNA methylation sites in the CG context are displayed, and the vertical height of the line indicates the methylation level of each methylcytosine. Abbreviations: mCG, methylcytosine (CG context); PMD, partially methylated domain.



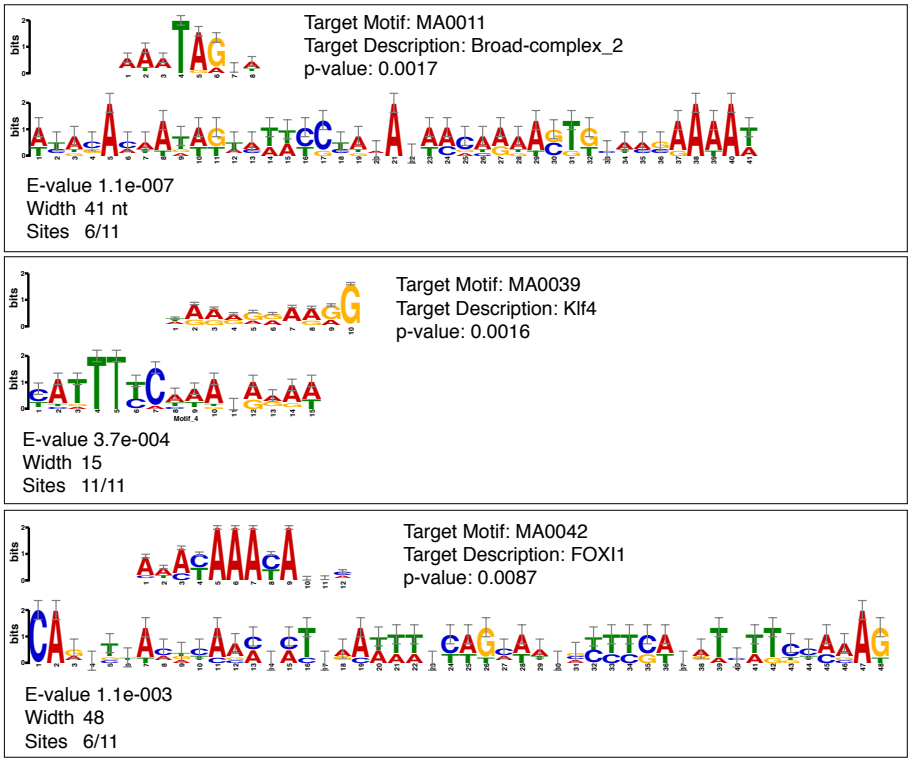
Supplementary Figure 8 | Clustering of CG-DMRs at CGIs hypomethylated in ESCs relative to somatic cells. Complete linkage hierarchical clustering of mCG density within CG-DMRs identified between all ESCs and somatic cells, hypomethylated in ESCs relative to somatic cells, and coincident with CGIs. Each CG-DMR was profiled over 20 equally sized bins.



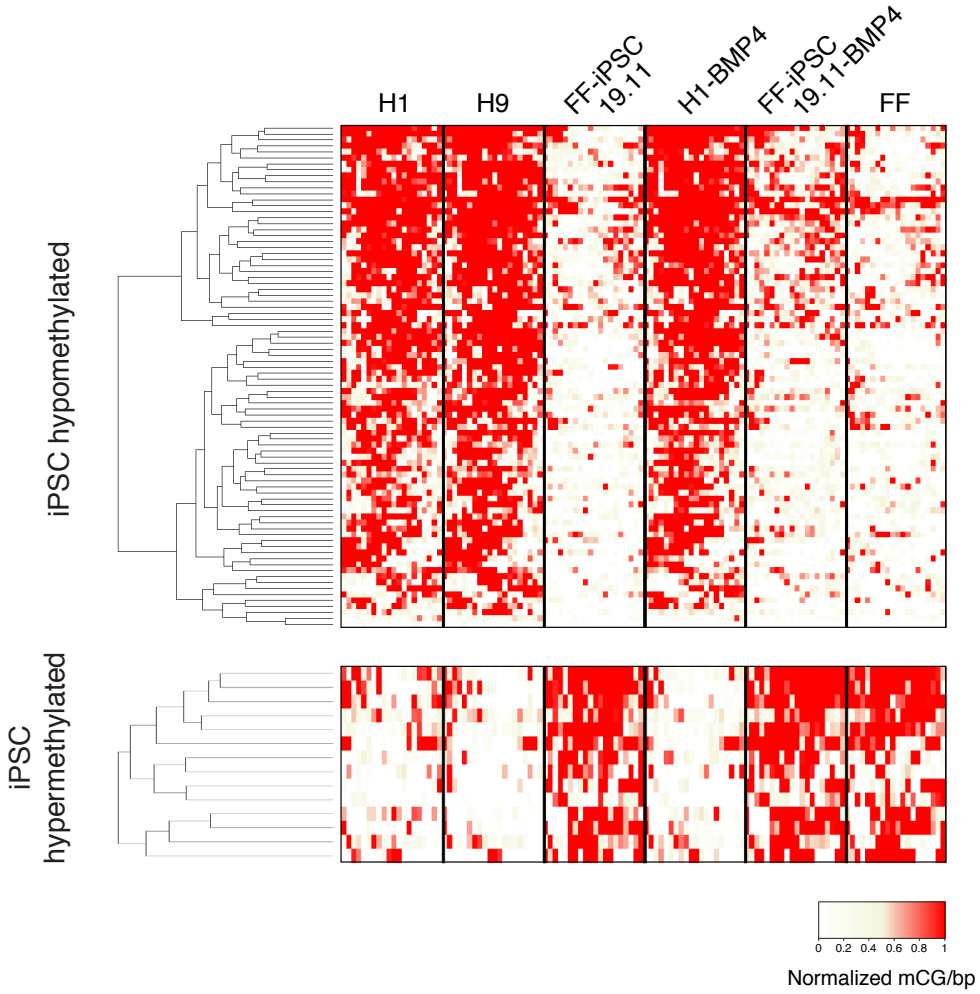
Supplementary Figure 9 | Clustering of CG-DMRs at CGIs hypomethylated in somatic cells relative to ESCs. Complete linkage hierarchical clustering of mCG density within CG-DMRs identified between all ESCs and somatic cells, hypomethylated in somatic cells relative to ESCs, and coincident with CGIs. Each CG-DMR was profiled over 20 equally sized bins.



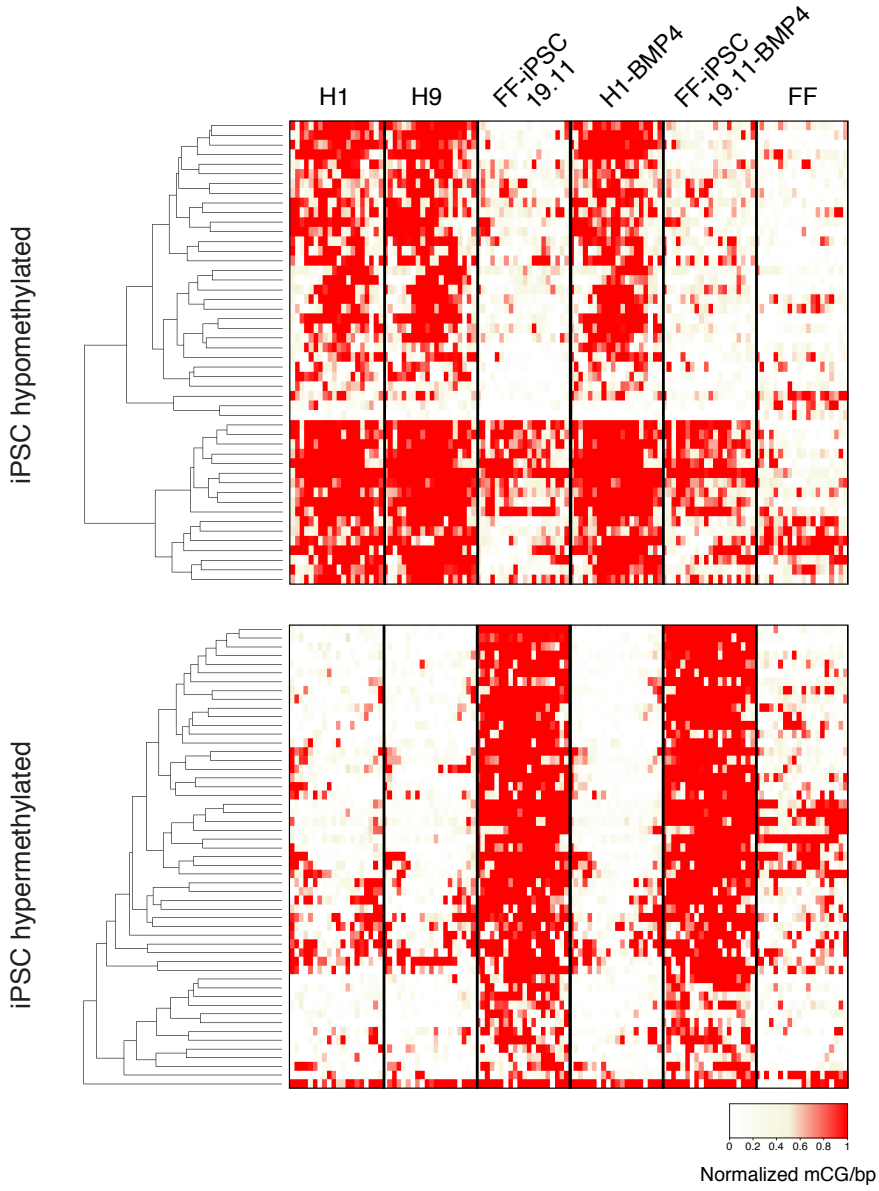
Supplementary Figure 10 | Heatmaps of mCG density for the CG-DMRs in each iPSC line that are similar or dissimilar to their progenitor somatic lines. mCG density is profiled in 20 bins over the CG-DMRs aberrant in each iPSC line (left black box in each heatmap) in respect to both H1 and H9 hESC (blue box). The CG-DMRs are divided according to the similarity to the methylation status of the progenitor somatic lines (right black box in each heatmap): CG-DMRs similar to progenitor somatic line (memory) in the left column, and CG-DMRs dissimilar to parental line on the right column (iDMR).



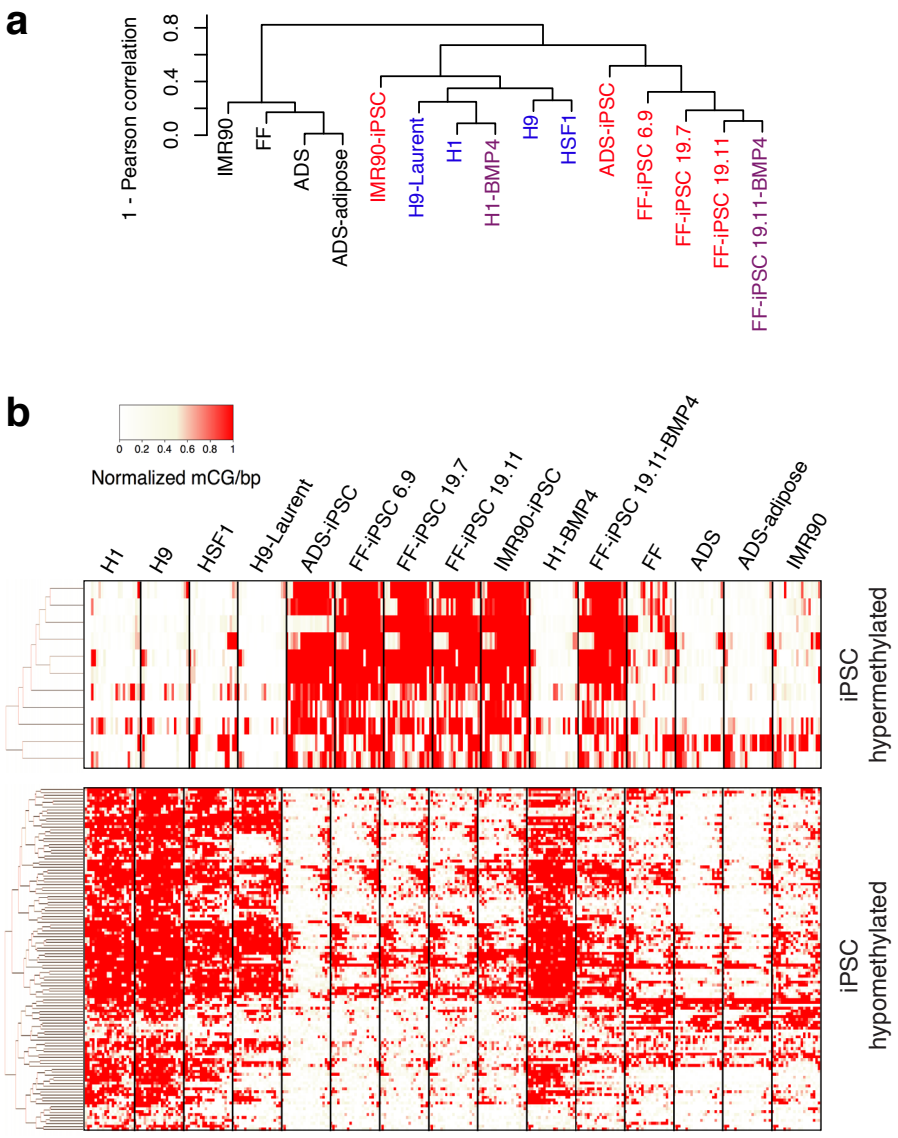
Supplementary Figure 11 | Analysis of over-represented motifs in the CG-DMRs conserved among all iPSC lines. 11 hypermethylated and 119 hypomethylated CG-DMRs shared by all five iPSC lines were considered, and sequences 1 kb upstream and downstream of the center of each DMR were retrieved (Watson strand). MEME was used to identify sequence patterns that occur repeatedly in the two group of DNA sequences. As a control, since the CG-DMRs are highly enriched for CG islands, a set of 2 kb sequences were retrieved from 100 random CG islands. The top 5 sequences within each group were considered (hypermethyltaed and hypomethylated CG-DMRs and random CG islands; E-value < 0.01). These sequences were analyzed with TOMTOM searching for transcription factor binding sites (TFBS, based on the JASPAR Core DB; q-value<0.5). Three out of the five sequences identified within the set of hypermethylated CG-DMRs were found to contain binding sites that were not found in the reference CG island set. The logo plots for these three sequences are displayed, with a length of 41, 15 and 48 nt, together with the logo plot of the matching TFBS. All analysis were performed with the MEME suite (http://meme.nbcr.net/meme4_5_0/intro.html).



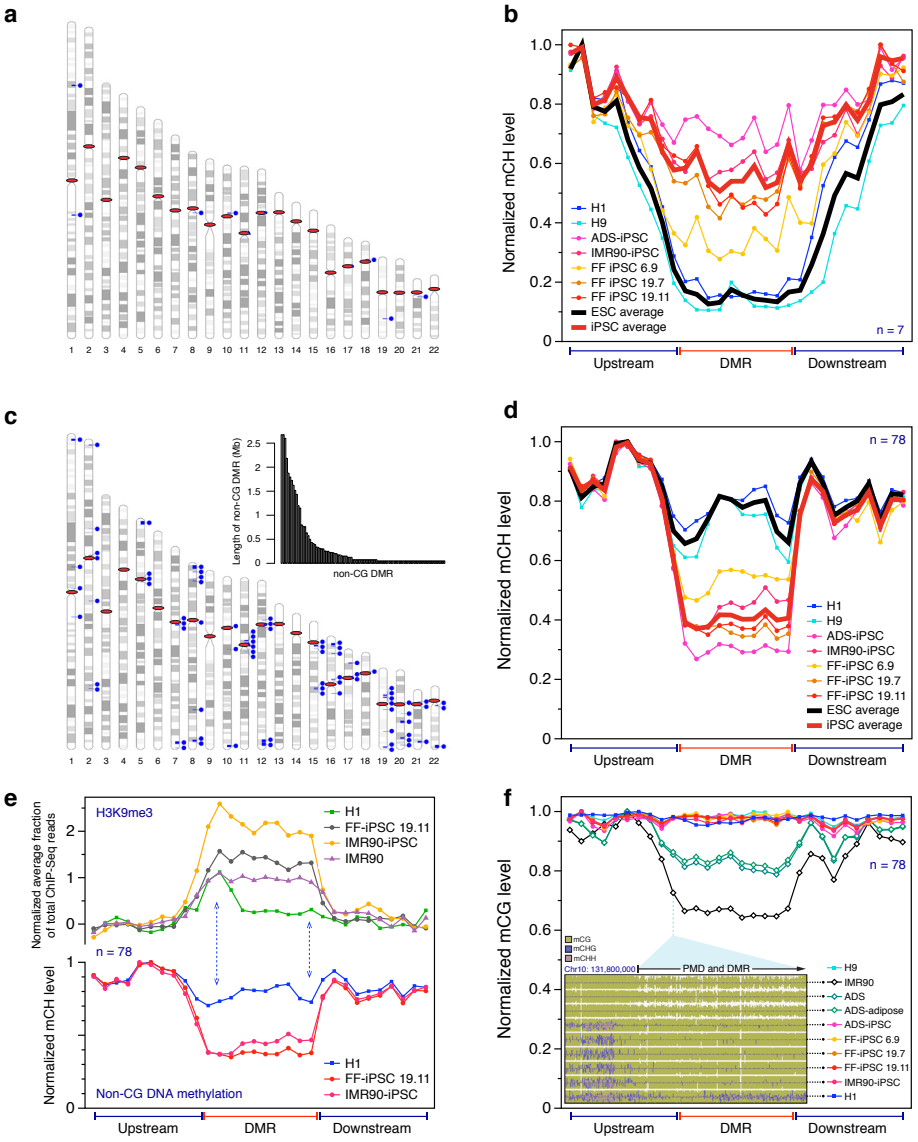
Supplementary Figure 12 | Clustering of FF-iPSC 19.11 CG-DMRs relative to ESCs, where somatic DNA methylation memory in FF-iPSC 19.11 is transmitted through cellular differentiation. Complete linkage hierarchical clustering of mCG density within CG-DMRs identified between FF-iPSC 19.11 and both ESCs, where FF-iPSC 19.11, FF progenitor, and FF-iPSC 19.11 differentiated with BMP4 to trophoblast lineage display the same methylation state. Separate heatmaps are shown for iPSC hypermethylation and hypomethylation relative to the the ESCs. Each CG-DMR was profiled over 20 equally sized bins.



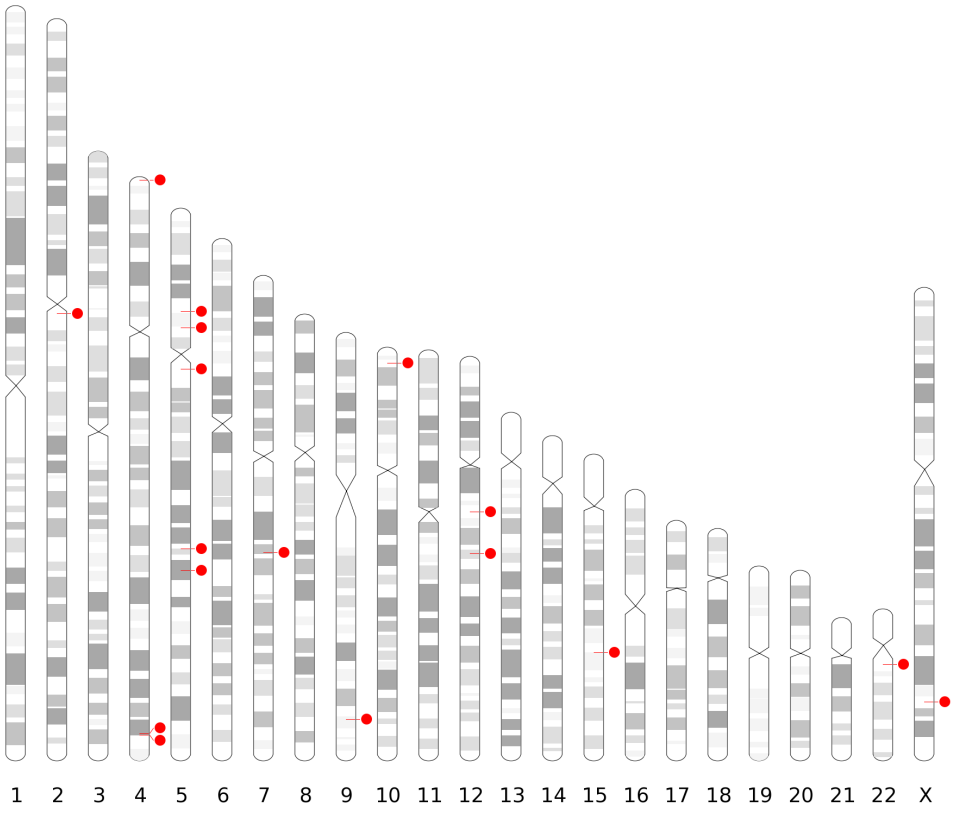
Supplementary Figure 13 | Clustering of FF-iPSC 19.11 CG-DMRs relative to ESCs, where FF-iPSC 19.11 methylation patterns unlike both the FF somatic progenitor and ESCs are transmitted through cellular differentiation. Complete linkage hierarchical clustering of mCG density within CG-DMRs identified between FF-iPSC 19.11 and both ESCs, where FF-iPSC 19.11 shows a methylation state not found in either ESCs or FF somatic progenitor, and FF-iPSC 19.11 differentiated with BMP4 to trophoblast displays the same methylation state as FF-iPSC 19.11. Separate heatmaps are shown for iPSC hypermethylation and hypomethylation relative to the the ESCs. Each CG-DMR was profiled over 20 equally sized bins.



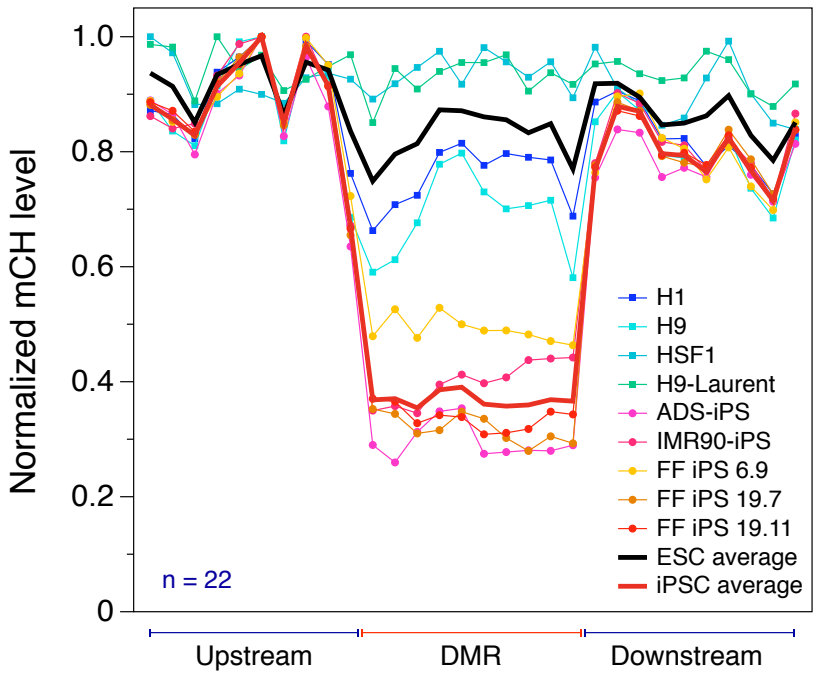
Supplementary Figure 14 | Validation of the CG-DMRs in two additional ESC lines. a, Complete linkage hierarchical clustering of mCG density within the 1175 CG-DMRs identified between iPSCs and H1 and H9 ESCs (Fig. 3a), which were also profiled in the HSF1 and H9-Laurent ESCs. **b,** as in (**a**) for the hypermethylated or hypomethylated CG-DMRs in all iPSCs. In both (**a**) and (**b**) the two additional ESCs show high similarity to H1 and H9 ESCs.



Supplementary Figure 15 | non-CG mega-DMRs. a, Chromosome ideograms of the 7 large non-CG DMRs where H1 is hypomethylated relative to ADS-iPSC. **b,** Normalized mCH levels over the 7 large non-CG DMRs and flanking genomic regions where H1 is hypomethylated relative to ADS-iPSC. **c,** Chromosome ideograms of all 78 non-CG DMRs where ADS-iPSC is hypomethylated relative to H1. **d,** Normalized mCH levels over all 78 non-CG DMRs where ADS-iPSC is hypomethylated relative to H1 and flanking genomic regions. **e,** Lower y-axis as in **(f)** for the cell lines indicated. Upper y-axis shows profiles of normalized H3K9me3 ChIP-Seq read density throughout the 78 non-CG DMRs. **g,** Plot displays normalized mCG levels over the 78 non-CG DMRs and flanking genomic regions.



Supplementary Figure 16 | Retroviral insertion sites in ADS-iPSC. Ideogram of the retroviral insertion sites in the genome of ADS-iPSC, indicated by red dots.



Supplementary Figure 17 | Validation of the non-CG mega-DMRs in two additional ESC lines. Normalized mCH levels over all non-CG mega-DMRs and flanking genomic regions, including the previously published HSF1 and H9-Laurent ESCs.