Legend to supplementary figures

Fig.S1 CpG density at FOXA1 binding sites. CpG density was analyzed at the indicated FOXA1 binding sites by first identifying the center of these regions and expanding them 1 kilobase (kb) on each side. These regions were then split into 40 windows where the number of CpGs was counted. CpG density refers to the average CpG number per window across all analyzed binding sites. As a control, randomly selected regions within the non-repetitive sequences of chromosomes covered by the DNA chips were used. . This analysis indicate that the observed DNA hypomethylation at FOXA1 binding sites from MCF7 and LNCaP cells is not associated with a marked lower CpG density as compared to neighboring regions or control sites. On the other hand, FOXA1 binding sites from stably transfected MDA-MB231 cells show a lower CpG density.

Fig.S2 DNA methylation levels at cell type-specific FOXA1 recruitment sites inferred from MeDIP-seq data. Average tag counts issued from high-throughput sequencing of input or MeDIP DNA from MCF7 cells (Ruike et al. 2010). Analyses were performed as in Fig.1. Note that differences in DNA methylation levels between the center of analyzed sites and the flanking regions is more pronounced for sites bound by FOXA1 in MCF7 cells (MCF7-specific and shared) than for sites bound specifically in LNCaP cells.

Fig.S3 Cell type-specific CTCF recruitment to chromatin in MCF7 and LNCaP cells is linked to differential DNA hypomethylation. **A)** ChIP-chip experiments were performed to identify CTCF binding sites in MCF7 and LNCaP cells. Enrichment peaks were called using MAT at FDR 1%. Numbers of overlapping and non-overlapping binding sites are indicated on the Venn diagram. Consistent with previous studies, even though many CTCF recruitment sites are conserved within the genome of different cell types, a non-negligible fraction is

different (Kim et al. 2007; Bushey et al. 2009; Cuddapah et al. 2009; Phillips and Corces 2009). **B)** CTCF and DNA methylation levels at MCF7-specific, LNCaP-specific or shared CTCF binding sites were determined from tiling arrays data obtained in MCF7 and LNCaP cells. A random set of regions from chromosomes 8, 11 and 12 was also analyzed. Overall, cell type-specific CTCF binding to chromatin is linked to differential DNA hypomethylation between MCF7 and LNCaP cells. These results are in line with previous reports showing on a few sites that CTCF binding to chromatin could be modulated by DNA methylation (Bell and Felsenfeld 2000; Hark et al. 2000; Parelho et al. 2008). **C)** CpG density was analyzed within CTCF binding sites as in Fig.S1.

Fig.S4 FOXA1-bound but transcriptionally inactive enhancers are DNA hypomethylated. Average MAT scores from tiling array data obtained in MCF7 cells for FOXA1 (**A**), FAIRE (**B**), H3K4me2 (**C**), H2AFZ (**D**) and DNA methylation (**E**) levels. MCF7-specific FOXA1 binding sites previously determined as active or inactive enhancers were analyzed (Eeckhoute et al. 2009). As evidenced in **B**, active but not inactive FOXA1-bound enhancers are enriched by Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE) (Nagy et al. 2003; Hogan et al. 2006; Giresi et al. 2007; Eeckhoute et al. 2009; Gaulton et al. 2010). Nevertheless, inactive FOXA1-bound enhancers show the presence of H3K4me2 (**C**) and H2AFZ (**D**) even though levels of these epigenetic marks at these sites are lower than at active FOXA1-bound *cis*-regulatory elements. Remarkably, both active and inactive FOXA1-bound enhancers show DNA hypomethylation (**E**). Note that the overall higher MeDIP enrichments observed for active sites might be explained at least in part by their higher CpG content (**F**). For instance, when we compared MeDIP signals from MCF7 and LNCaP cells (where FOXA1 does not bind to the analyzed enhancers), we observed a similar drop in MeDIP levels in MCF7 cells at active and inactive FOXA1-bound enhancers

(**G**). Hence, FOXA1 recruitment rather than transcriptional activities of enhancers *per se* is linked to DNA hypomethylation suggesting that this epigenetic feature could be required but not sufficient for enhancer functions.

Fig.S5 mRNA expression levels of the *FOXA* family members in MCF7, LNCaP, MDA-control and MDA-FOXA1 cells. RT-qPCR experiments were performed in MCF7, LNCaP, MDA-control and MDA-FOXA1 cells to monitor *FOXA1*, *2* and *3* mRNA expression levels. HepG2 cells were used as a reference since they express all three FOXA family members. Expression of analyzed genes was normalized using *RSP28*. Results shown are means and S.D. from two independent experiments.

Fig.S6 Changes in mRNA expression levels of genes closest to validated FOXA1 binding sites in MDA-FOXA1 compared to MDA-control cells. RT-qPCR experiments were performed in MDA-control and MDA-FOXA1 cells. Expression of analyzed genes was normalized using *RSP28* and is shown as fold change relative to expression in MDA-control, which was set to 1. Results are from two independent experiments performed in duplicates. ND, not detectable.

Fig.S7 mRNA expression levels of the *Foxa* family members in P19 cells. RT-qPCR experiments were performed in P19 cells before (0h) and after stimulation with RA (48 and 120h) to monitor *Foxa1*, *2* and *3* mRNA expression levels. Expression of analyzed genes was normalized using *Rsp28*. Results are means and S.D. from three independent experiments.

Fig.S8 CpGs are absent from the FOXA1 recognition motif but are found in its vicinity within enhancers. A) Matrix for FOXA1 DNA recognition motif from (Lupien et al. 2008).

Sequences of the FOXA1-bound enhancers analyzed by bisulfite pyrosequencing in Fig.2B (**B**), Fig.2C (**C**) and Fig.4D and.Fig.5 (**D**) are provided. The central 200 bp of the enhancers were identified by inspection of the FOXA1 ChIP-chip signals using Integrated Genome Browser (Nicol et al. 2009) and are indicated as red colored letters. Potential FOXA1 recognition sites within this central region of the enhancers appear underlined and in blue. CpGs that were analyzed by bisulfite pyrosequencing are highlighted in yellow and numbered according to the respective main figures.

Fig.S9 DNA demethylation kinetic in P19 cells analyzed using sensitivity to digestion by *HpyCH4IV*. Genomic DNA from P19 cells treated with RA for the indicated times was subjected to *HpyCH4IV* digestion. The percentage of non-digested DNA was determined by qPCR using primers allowing for amplification of a DNA fragment encompassing CpG 2 and 1 of FOXA1 binding sites 1 and 6, respectively (contained within ACGT motifs as shown in Fig.S8D). Results are means and S.D. from at least two independent experiments.

Fig.S10 Sequences of enhancers used in luciferase reporter assays. Sequences of enhancers used in Fig.6 are given. CpGs are highlighted in red.

Fig.S11 Effect of 5-aza-2'-deoxycytidine treatment on genes whose expression is stimulated during RA-induced neural differentiation of P19 cells. **A)** Bisulfite-pyrosequencing experiments were performed in DMSO vehicle or 5-aza-2'-deoxycytidine (5-azaC) treated P19 cells on FOXA1 binding sites 5 and 6. CpG numbering is consistent with Fig.5. Results are means and S.D. from duplicates. **B-C)** RT-qPCR experiments were performed in P19 cells treated with 5-azaC or vehicle only for 48h. Expression of analyzed genes was normalized using *Rsp28* and is shown as fold change relative to expression in

vehicle treated cells, which was set to 1. Results are from two independent experiments performed in duplicates. ND, not detectable.

Fig.S12 Pre-existing DNA hypomethylation facilitates binding of ectopically-expressed ESR1. MeDIP-qPCR were performed in both MCF7, MDA-MB231 control cells (MDA-control) and MDA-MB231 stably expressing ESR1 (MDA-ESR1) (Metivier et al. 2003). ESR1 recruitment sites specific to MCF7 or MDA-ESR1 cells were inferred from (Carroll et al. 2006) and Quintin et al. (manuscript in preparation). Indeed, as previously observed by Krum et al. (Krum et al. 2008), ESR1 binding to chromatin in cells lacking FOXA1 is different from that in MCF7 cells. For each analyzed ESR1 binding site, relative MeDIP levels between MCF7 and MDA-control or MDA-ESR1 were determined and presented as in Fig.1E. Fold differences were inferred from data obtained in three independent experiments. Note that when expressed in MDA-MB231 cells, ESR1 binds to regions that were already DNA hypomethylated in the control cells.

Supplementary references

Bell, A.C. and Felsenfeld, G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405**(6785): 482-485.

Bushey, A.M., Ramos, E., and Corces, V.G. 2009. Three subclasses of a Drosophila insulator show distinct and cell type-specific genomic distributions. *Genes Dev* **23**(11): 1338-1350.

Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., Wang, Q., Bekiranov, S., Sementchenko, V., Fox, E.A., Silver, P.A., Gingeras, T.R., Liu, X.S., and Brown, M. 2006. Genome-wide analysis of estrogen receptor binding sites. *Nat Genet* **38**(11): 1289-1297.

Cuddapah, S., Jothi, R., Schones, D.E., Roh, T.Y., Cui, K., and Zhao, K. 2009. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* **19**(1): 24-32.

Eeckhoute, J., Lupien, M., Meyer, C.A., Verzi, M.P., Shivdasani, R.A., Liu, X.S., and Brown, M. 2009. Cell-type selective chromatin remodeling defines the active subset of FOXA1-bound enhancers. *Genome Res* **19**(3): 372-380.

Gaulton, K.J., Nammo, T., Pasquali, L., Simon, J.M., Giresi, P.G., Fogarty, M.P., Panhuis, T.M., Mieczkowski, P., Secchi, A., Bosco, D., Berney, T., Montanya, E., Mohlke,

K.L., Lieb, J.D., and Ferrer, J. 2010. A map of open chromatin in human pancreatic islets. *Nat Genet* **42**(3): 255-259.

Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R., and Lieb, J.D. 2007. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res* **17**(6): 877-885.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**(6785): 486-489.

Hogan, G.J., Lee, C.K., and Lieb, J.D. 2006. Cell cycle-specified fluctuation of nucleosome occupancy at gene promoters. *PLoS Genet* **2**(9): e158.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* **128**(6): 1231-1245.

Krum, S.A., Miranda-Carboni, G.A., Lupien, M., Eeckhoute, J., Carroll, J.S., and Brown, M. 2008. Unique ER{alpha} cistromes control cell type-specific gene regulation. *Mol Endocrinol* **22**(11): 2393-2406.

Lupien, M., Eeckhoute, J., Meyer, C.A., Wang, Q., Zhang, Y., Li, W., Carroll, J.S., Liu, X.S., and Brown, M. 2008. FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**(6): 958-970.

Metivier, R., Penot, G., Hubner, M.R., Reid, G., Brand, H., Kos, M., and Gannon, F. 2003. Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115**(6): 751-763.

Nagy, P.L., Cleary, M.L., Brown, P.O., and Lieb, J.D. 2003. Genomewide demarcation of RNA polymerase II transcription units revealed by physical fractionation of chromatin. *Proc Natl Acad Sci U S A* **100**(11): 6364-6369.

Nicol, J.W., Helt, G.A., Blanchard, S.G., Jr., Raja, A., and Loraine, A.E. 2009. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**(20): 2730-2731.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B.S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A.G., and Merkenschlager, M. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**(3): 422-433.

Phillips, J.E. and Corces, V.G. 2009. CTCF: master weaver of the genome. *Cell* **137**(7): 1194-1211.

Ruike, Y., Imanaka, Y., Sato, F., Shimizu, K., and Tsujimoto, G. 2010. Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing. *BMC Genomics* **11**: 137.