Supplementary Information

Supplementary methods

Chromatin immunoprecipitation (ChIP) assays

Note H2A.Z (Abcam#ab4174) antibody recognizes a 100aa epitope in the C-terminal region of H2A.Z, and therefore detects both acetylated and unacetylated H2A.Z (Bruce et al. 2005), whereas H2A.Z acetyl K4+K7+K11 (Abcam#ab18262) only recognises acH2AZ. Ten µg of antibody was used for each immunoprecipitation. Non-immune rabbit IgG controls (Upstate#PP64B) were also included for each ChIP assay and no precipitation was observed by quantitative Real-Time PCR (qPCR) analysis. Duplicate input samples were processed in parallel.

Gene expression array analysis

Affymetrix GeneChip HuGene 1.0 ST expression array data was preprocessed using Robust Multichip Analysis (RMA) (Irizarry et al. 2003) using the implementation in the R aroma.affymetrix (http://aroma-project.org/) package. The summarized expression values for each gene and experimental condition were averaged over replicates. An RMA value of 4.5 or less (log2 scale) was considered to be below detection. For the assessment of differential expression between LNCaP, DU-145, PC-3 and PrEC cell lines, a linear model was fitted to the RMA-processed data, and moderated t-statistics (t-stats) were obtained using the R package limma (Smyth 2004).

Affymetrix GeneChip Human Promoter 1.0R Array Analysis

The Affymetrix GeneChip Human Promoter 1.0R array data was processed by MAT normalization (Johnson et al. 2006) as implemented in the R package

aroma.affymetrix (http://aroma-project.org/). Microarray probes were mapped to regions of interest (i.e. windows overlapping gene TSSs). To achieve a promoter-level summary of change in tiling array signal, the 'blocksStats' procedure from the Repitools R package (Statham et al., 2010) was used. Briefly, for each region, a score for each probe was calculated according to the difference of interest (e.g. LNCaP signal - PrEC signal for each probe). To test the significance of a change at the promoter level, the scores for all probes were used within a single sample two-sided t-test (t-stats). Enrichment of ChIP signals were visualised using UCSC Genome Browser software (http://genome.ucsc.edu/).

MBDCap analysis

The differences between LNCaP and PrEC were analysed by an exact test method for count data (Robinson and Oshlack 2010) between -2000 and +1500 bases of each genes' TSSs. A modification of the published method was implemented, where the count data was split into ten copy number levels. This was based on fold change ratios of tag counts between the input samples of LNCaP and PrEC, in windows 0f 20kb and along the entire genome. Each copy number level was separately tested with an effective library size of the total counts per lane within all of the TSS-surrounding windows, multiplied by the average copy number of the windows within the copy number level. For absolute statistics, an exact test was performed on the counts of reads again between -2000 to +1500 of TSSs, for the IP and input samples, per cell line.

Significance plots

Absolute calls. For the study of the correlation between DNA methylation and H2A.Z promoter regions having > 80 adjusted counts in the DNA treated with SssI (fully methylated control, used as an indicator of what genomic regions can be interrogated using MBDCap.seq) were used, limiting the information to 11,000 genes. Genes were deemed highly methylated when their adjusted counts were > 80% relative to the SssI fully methylated control. On the other hand, genes were deemed lowly methylated when their adjusted count was < 20% of the control.

The genes sets for high/low of H2A.Z, acH2A.Z and acH2A.Z/H2A.Z correspond to promoter level tiling array t-stats of +1.50/-1.50

Relative change calls. Hypermethylated genes in prostate cancer had 2 fold of gain in methylation in LNCaP than PrEC (3150 genes). On the other hand, hypomethylated genes in cancer correspond to those ones with 2 fold less methylation in LNCaP than PrEC (225 genes). Then from these two gene lists we generate a shorter list containing the hypermethylated genes, which had a loss in gene expression (≥ 1.5 t-stats) in cancer obtaining 903 genes and hypomethylated genes with gain in gene expression in LNCaP (75 genes). The genes sets for gain/loss of H2A.Z, acH2A.Z and acH2A.Z/H2A.Z correspond to promoter level tiling array t-stats of +1.50/-1.50

DNA methylation analysis using Sequenom analysis

Following bisulphite conversion, DNA was PCR amplified in triplicate using primers listed in **Supplementary Table 5**. The primers were designed using the EpiDesigner ^{BETA} software from Sequenom in such a way that the amplicons covered the same region of the CpG island and the regions where the ChIP primers were designed. Each reverse primer has a T7-promoter tag (5-cagtaatacgactcactatagggagaaggct-3) and each

forward primer has a 10-mer tag (5-aggaagagag-3). The results were analysed by the EpiTYPER software V 1.0 and methylation ratios were calculated as previously described (Nair et al. 2011).

Supplementary figure legends:

Supplementary Figure 1: Correlation between gene expression and H2A.Z, H3, H2A, acH2A.Z promoter occupancy in LNCaP cells and H3 promoter occupancy in PrEC cells. Heatmaps (left panel) showing levels of (**A**) H2A.Z, (**B**) H3, (**C**) H2A, (**D**) acH2A.Z and (**E**) acH2A.Z/H2A.Z across gene promoters according to gene expression in LNCaP cell line and, in the case of (**B**) H3, for PrEC as well. Lineplots (right panel) for (**A**) H2A.Z, (**B**) H3, (**C**) H2A, (**D)** acH2A.Z and (**E**) acH2A.Z/H2A.Z show enrichment in LNCaP cells according to different levels of expression and, in the case of (**B**) H3, also for PrEC. Abbreviations, colours and symbols are the same as **Figure 1**.

Supplementary Figure 2: Gene examples of H2A and H2A.Z correlation at gene promoters. UCSC genome browser tracks for H2A.Z and H2A Chip on chip promoter arrays and expression arrays data in PrEC cells. **(A)** Highly expressed gene examples showed high levels of H2A.Z occupancy and no presence of H2A at their promoter. **(B)** Lowly expressed gene examples that exhibit intermediate levels of both H2A.Z and H2A.Z presence at their promoters levels. **(C)** Microarray hybridization signals for mRNA expression levels in PrEC of the three highly expressed representative genes (right side, light green) and the two poorly expressed representative genes (left side, dark green).

Supplementary Figure 3: H3K9 acetylation, H3K4 tri-methylation and H3K27 tri-methylation occupancy at gene promoters in PrEC and LNCaP cell lines.

(**A**) K9 acetylation in Histone3 3 (H3K9ac) and (**B**) K4 tri-methylation in Histone 3 (H3K4me3) profiles were used as active chromatin marks in PrEC (left panel) and

LNCaP (right panel) and (**C**) K27 tri-methylation in Histone 3 (H3K27me3) was used as an inactive chromatin mark.

Supplementary Figure 4: Correlation between DNA methylation and H2A in human prostate cancer. (A) Significance plots of H2A signal comparing DNA methylation along gene promoters in genes sorted by high methylation (red line) or low methylation (black line) in PrEC (upper panel) and LNCaP (bottom panel). These plots were made following the same criteria as **Figure 6A**. Significance plots to compare changes in H2A with the changes in DNA methylation in LNCaP-PrEC were made **(B and C)**. In B the red line represents DNA hypermethylated genes (≥ 2 fold) and the black line corresponds to DNA hypomethylated genes (≤ -2 fold) in LNCaP-PrEC. In C the red line correspond to DNA hypermethylated (≥ 2 fold) and down-regulated (≥ 1.5 t-stats) genes and the black line represents DNA hypomethylated (≤ -2 fold) and down-regulated genes (≤ -1.5 t-stats) in LNCaP-PrEC.

Supplementary Figure 5: Correlation between DNA methylation and H2A.Z, acH2A.Z or acH2A.Z/H2A.Z in human prostate cancer. Significance plots to compare changes in H2A.Z (left), acH2A.Z (middle) or acH2A.Z/H2A.Z (right) with changes in DNA methylation in LNCaP-PrEC. The red line represents genes DNA hypermethylated (≥ 2 fold) in LNCaP-PrEC. and the black line corresponds to DNA hypomethylated genes (≤ -2 fold) in LNCaP-PrEC

Supplementary Figure 6: Gene examples for changes in acH2A.Z/H2A.Z promoter enrichment and DNA methylation in prostate cancer. (A) Microarray hybridization signals for mRNA expression levels as explained in **Figure 3**, for two

hypomethylated gene examples: *MKRN3* and *LIMCH1*. **(B)** Gene expression validation by RT-qPCR reveals up-regulation in LNCaP relative to PrEC in *MKRN3* and *LIMCH1*. **(C)** UCSC genome browser tracks for DNA methylation, H3K27me3, H2A.Z, acH2A.Z and acH2A.Z/H2A.Z for these two genes. Abbreviations, colours and symbols are the same as **Figure 3 and 4. (D)** Quantitative DNA methylation Sequenom MALDI-TOF analysis of genomic DNA from PrEC and LNCaP cell lines. Average methylation ratios obtained from each interrogated CpG site of the promoter region of *RND3, MKRN3* and *LIMCH1* are shown in both LNCaP and PrEC.

Supplementary Table 1. Up-regulated genes in LNCaP-PrEC associated with a gain of acH2A.Z/H2A.Z (t-stats $\geq$ 1.5).

Supplementary Table 2. Down-regulated genes in LNCaP-PrEC associated with a loss of acH2A.Z/H2A.Z (t-stats $\leq$ -1.5).

Supplementary Table 3. DNA hypomethylated and up-regulated genes in LNCaP cell line.

Supplementary Table 4. DNA hypermethylated and down-regulated genes in LNCaP cell line.

Supplementary Table 5. Primers used for gene expression, ChIP and Sequenom analyses**.**

References

Bruce K, Myers FA, Mantouvalou E, Lefevre P, Greaves I, Bonifer C, Tremethick DJ, Thorne AW, Crane-Robinson C. 2005. The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. *Nucleic Acids Res* **33**(17): 5633-5639.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**(2): 249-264.

Johnson WE, Li W, Meyer CA, Gottardo R, Carroll JS, Brown M, Liu XS. 2006. Model-based analysis of tiling-arrays for ChIP-chip. *Proc Natl Acad Sci U S A* **103**(33): 12457-12462.

Nair SS, Coolen MW, Stirzaker C, Song JZ, Statham AL, Strbenac D, Robinson MD, Clark SJ. 2011. Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias. *Epigenetics* **6**(1).

Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**(3): R25.

Smyth GK. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**: Article3.