# University of Groningen

**University of Groningen**

## Undifferentiated Embryonic Cell Transcription Factor 1 Regulates ESC Chromatin Organization and Gene Expression

Kooistra, Susanne M.; van den Boom, Vincent; Thummer, Rajkumar P.; Johannes, Frank; Wardenaar, Rene; Tesson, Bruno M.; Veenhoff, Liesbeth M.; Fusetti, Fabrizia; O'Neill, Laura P.; Turner, Bryan M.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

*Citation for published version (APA):*
Kooistra, S. M., van den Boom, V., Thummer, R. P., Johannes, F., Wardenaar, R., Tesson, B. M., Veenhoff, L. M., Fusetti, F., O'Neill, L. P., Turner, B. M., de Haan, G., Eggen, B. J. L., & O'Neill, L. P. (2010). Undifferentiated Embryonic Cell Transcription Factor 1 Regulates ESC Chromatin Organization and Gene Expression. *STEM CELLS, 28*(10), 1703-1714. https://doi.org/10.1002/stem.497

**Inventory of Supplemental Information**

**Figure S1, related to Figure 1.** UTF1 interacts with core histones. Coomassie stained 4-20% SDS-PAGE gels of UTF1 co-immunoprecipitation experiments are shown and the proteins identified using mass-spectrometry are indicated.

**Figure S2, related to Figure 1.** Identification of UTF1-enriched tiles using a mixture model approach.

**Figure S3, related to Figure 1.** Distribution of UTF1 target genes over the chromosomes.

**Figure S4, related to Figure 1.** Peak-like and wide-spread occupation patterns of UTF1 on target regions.

**Figure S5, related to Figure 1.** Chromosal view of UTF1 targets on chromosome Y-random and a fragment of chromosome 13.

**Table S1, related to Figure 1.** List of UTF1 target genes. Chromosomal location, gene name and NCBI gene ID are given.

**Figure S6, related to Figure 2.** Western analysis of wild type and UTF1 KD IB10 ES cells used for expression array.

**Table S2, related to Figure 2.** List of genes with significantly upregulated expression levels in UTF1 KD ES cells. Chromosomal location, gene name and NCBI gene ID are listed. The log2 expression levels of these transcripts in biological replicates (A

and B) in wild type (wt) and two UTF1 KD ES lines (#1 and #2) with their rank products are given.

**Table S3, related to Figure 2.** List of genes with significantly downregulated expression levels in UTF1 KD ES cells. Chromosomal location, gene name and NCBI gene ID are listed. The log2 expression levels of these transcripts in biological replicates (A and B) in wild type (wt) and two UTF1 KD ES lines (#1 and #2) with their rank products are given.

**Figure S7, related to Figure 3.** Expression levels of all and differentially expressed UTF1 enriched genes in wild type and UTF1-knock down ES cells.

**Figure S8, related to Figure 3.** Bootstrap null distribution showing the significance (P=0.0001) of the overlap between genes differentially expressed in wild type and UTF1-knock down ES cells and UTF1-bound genes.

**Figure S9, related to Figure 4.** Western analysis of various histone modifications in wt and UTF1 P19CL6 cells. Overall histone modification levels were not altered in UTF1 KD P19CL6 cells.

**Figure S10, related to Figure 4.** Subnuclear fractionation P19CL6 cells differentiated with 1% DMSO for 0, 2 and 4 days, indicating that UTF1 remains tightly chromatin associated (fractionates to the ammonium sulphate fraction) during differentiation.

**Figure S11, related to Figure 4.** UTF1-associated chromatin is enriched for H3K27me3. The subnuclear distributions of UTF1 in relation to H3K27me3, H3K4me3, H3K9me3, H4K20me3 and H2B-GFP was determined.

**Figure S12, related to Figure 4.** DMSO-induced differentiation of P19CL6 cells was confirmed using western analysis of UTF1, Oct4, Troma1 and Actin levels at day 0, 3, 5 and 8 days of DMSO treatment.

**Figure S13, related to Figure 4.** Immunofluorescent analysis of chromocenters using H3K9me3, H4K20me3, HP1$\gamma$ and H3K4me3 antibodies in wild type and UTF1 knock down EC cells.

**Supplemental Figures**
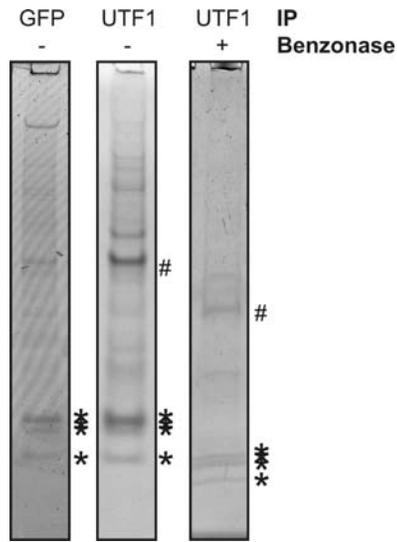


**Figure S1.** Coomassie stained 4-20% SDS-PAGE gels of co-immunoprecipitation experiments. P19CL6 or GFP-UTF1 expressing P19CL6 cell lysates were prepared in the presence (+) or absence (-) of benzonase as indicated and used in immunoprecipitations (IP) with anti-GFP or anti-UTF1 antibodies as indicated. Co-precipitating proteins were analyzed using mass-spectrometry, resulting in the identification of the core histones H2A, H2B, H3 and H4 (indicated by asterisks) and UTF1 (indicated by #).
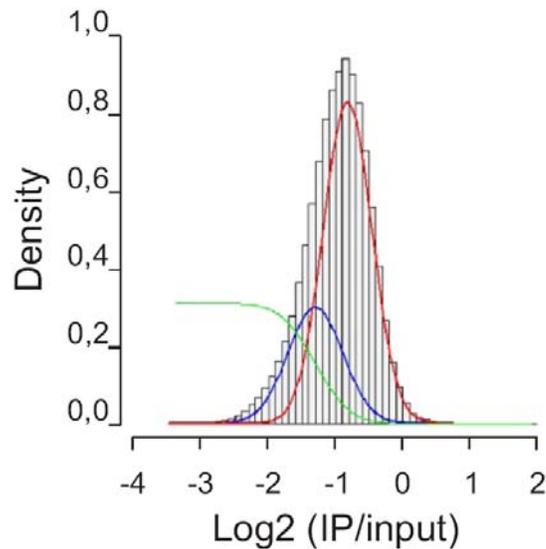


**Figure S2.** Mixture model of log2 (IP/input) values of all tiles. Populations of UTF1 enriched (red line) and non-enriched (blue line) tiles are indicated. The green line represents the genome-wide false positive cutoff.
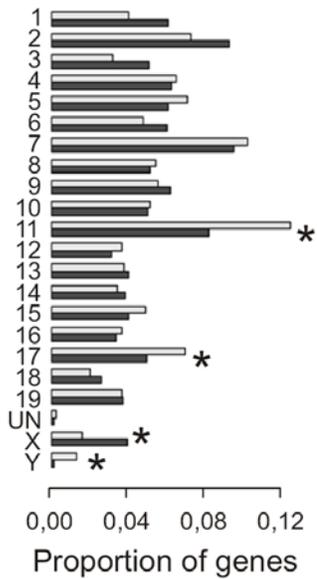
**Figure S3.** Distribution of UTF1 target genes over the chromosomes. The relative proportion of genes within the UTF-bound group (light grey bars) and the proportion based on chance (dark grey bars) are indicated. Asterisks indicate chromosomes that are significantly (P<0.05, Bonferoni adjusted for multiple testing) enriched (chr. 11, 17 and Y) or depleted (chr. X) for UTF1 targets.
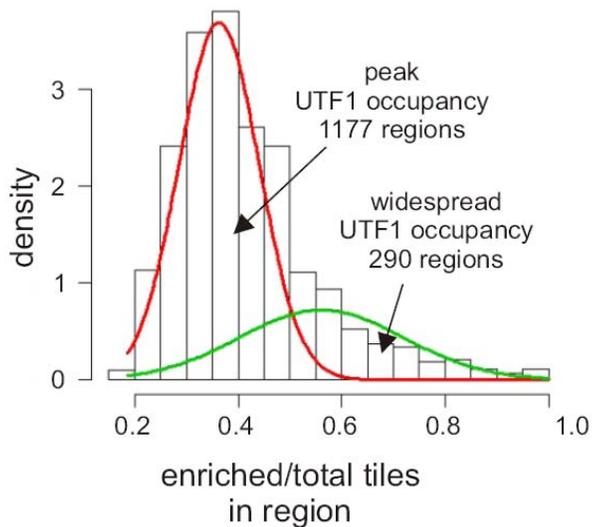


**Figure S4.** Mixture model of UTF1-enriched/total tiles in UTF1-enriched regions. Populations of genes where >51% of the tiles was significantly enriched for UTF1 (widespread UTF1 occupancy, green line) and less then <51% of the tiles was enriched for UTF1 (peak UTF1 occupancy, red line) are indicated.
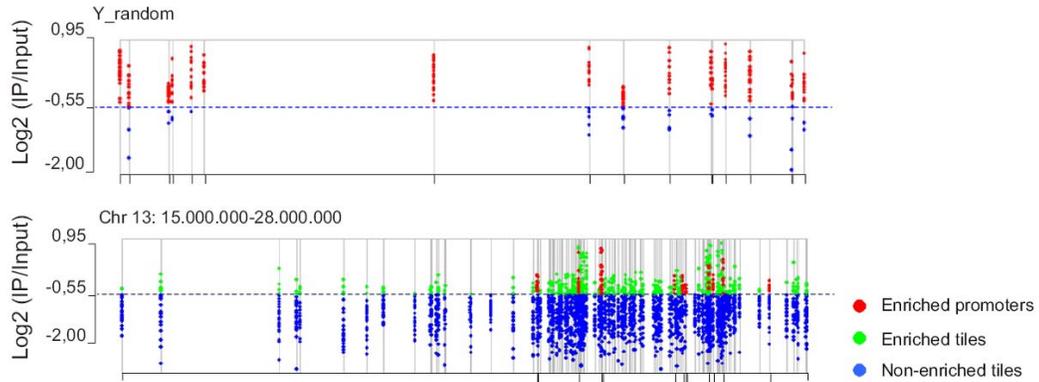
**Figure S5.** Chromosomal view of UTF1 targets. Log2 (IP/input) values of all tiles on chromosomes Y-random and 13 (positions 15.000.000-28.000.000, mouse genome build MM8). The location of each tiled region is indicated by a vertical grey line. Tiles with log2 values below false positive cutoff (indicated by blue dotted line) are indicated in blue, tiles above threshold are indicated in green. Tiles with log2 values above threshold that are within an enriched promoter (Log2 values of 5 out of 6 consecutive tiles are above threshold) are indicated in red.
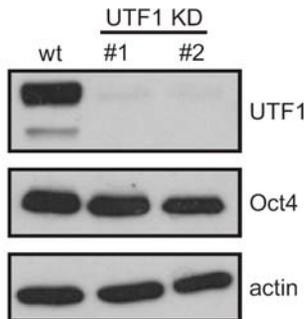


**Figure S6** Western analysis of UTF1, Oct4 and actin levels in wild type (wt) and UTF1 knock down (UTF1 KD) IB10 ES cells analyzed in the expression arrays.
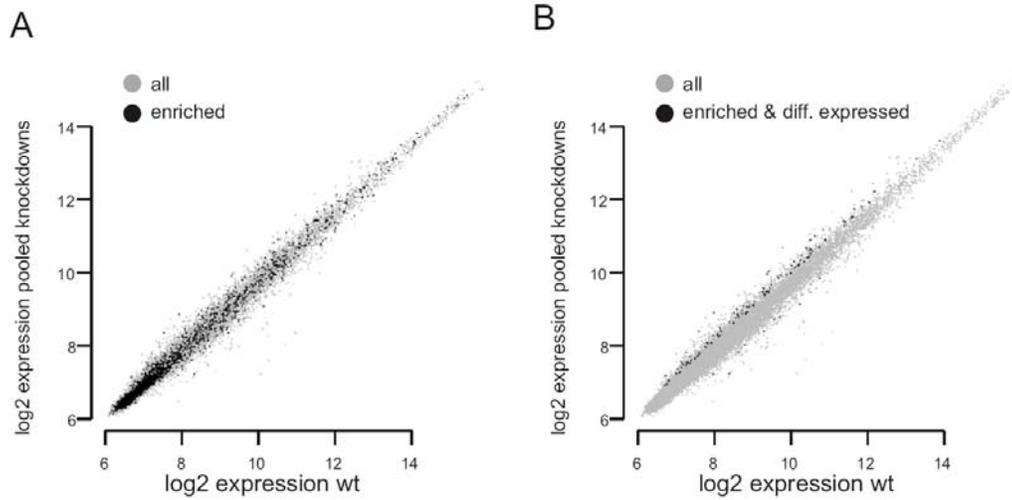
**Figure S7**. Expression of UTF1 target genes **(A)** Log2 expression level of all probes in wild type (wt) versus UTF1 KD ES cells. All probes are shown (grey dots) and the probes of UTF1 target genes are indicated by black dots. **(B)** Log2 expression level of all probes in wild type versus UTF1 KD ES cells. All probes are represented by grey dots and the probes of UTF1 target genes that are differentially expressed between wild type and UTF1 KD ES cells are indicated with black dots.
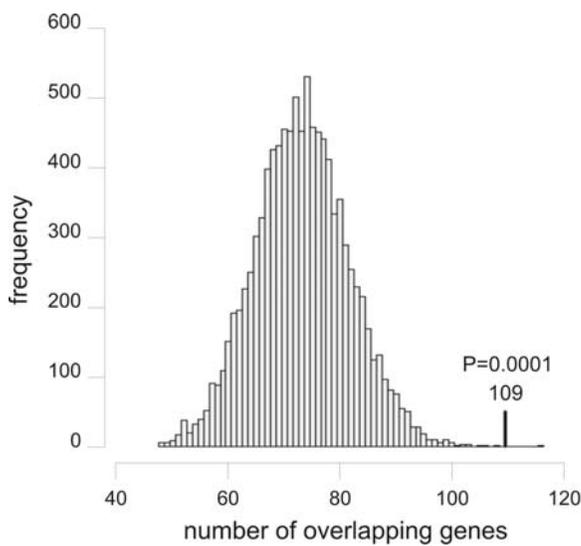


**Figure S8.** Bootstrap null distribution showing the significance (P=0.0001) of the overlap between differentially expressed and UTF1-bound genes (109 genes).

UTF1

wt  R  #1  #2

H3K27me1

H3K27me2

H3K27me3

H3K4me2

H3K4me3

H3K9me3

H4K20me3

H3K9ac

UTF1

Actin
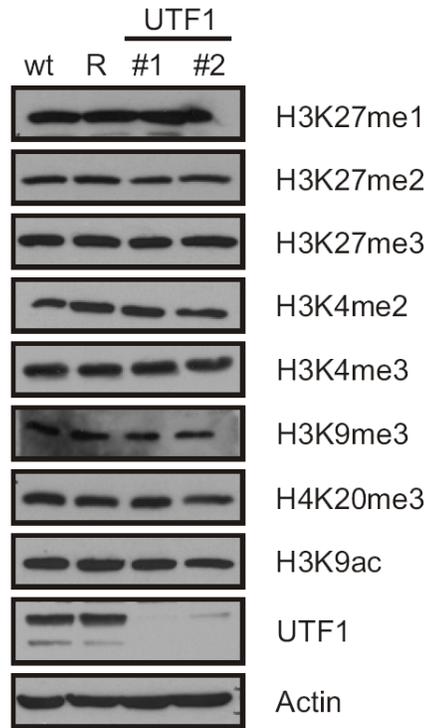
**Figure S9** Cell lysates of wild type (wt), renilla luciferase knock down (R) and UTF1 knock down (UTF1 #1 and #2) IB10 ES cells were analysed by western blotting and the antibodies indicated.

F  D  AS  HS  M    1% DMSO (days)

0

2

4
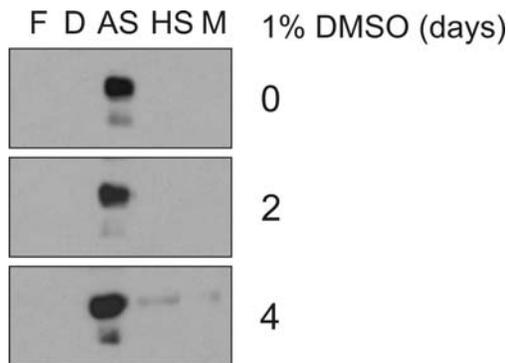
**Figure S10.** Subnuclear fractionation of UTF1 in P19CL6 cells differentiated with 1% DMSO for 0, 2 and 4 days. F: free diffusing/cytoplasmic fraction; D: DNaseI fraction ; AS: ammonium sulphate fraction; HS: high salt fraction; M: nuclear matrix fraction. Since UTF1 is downregulated during differentiation, more protein was loaded for the 4 day DMSO-treated sample. At all time points investigated, UTF1 was detected in the AS fraction containing strongly DNA-associated proteins.
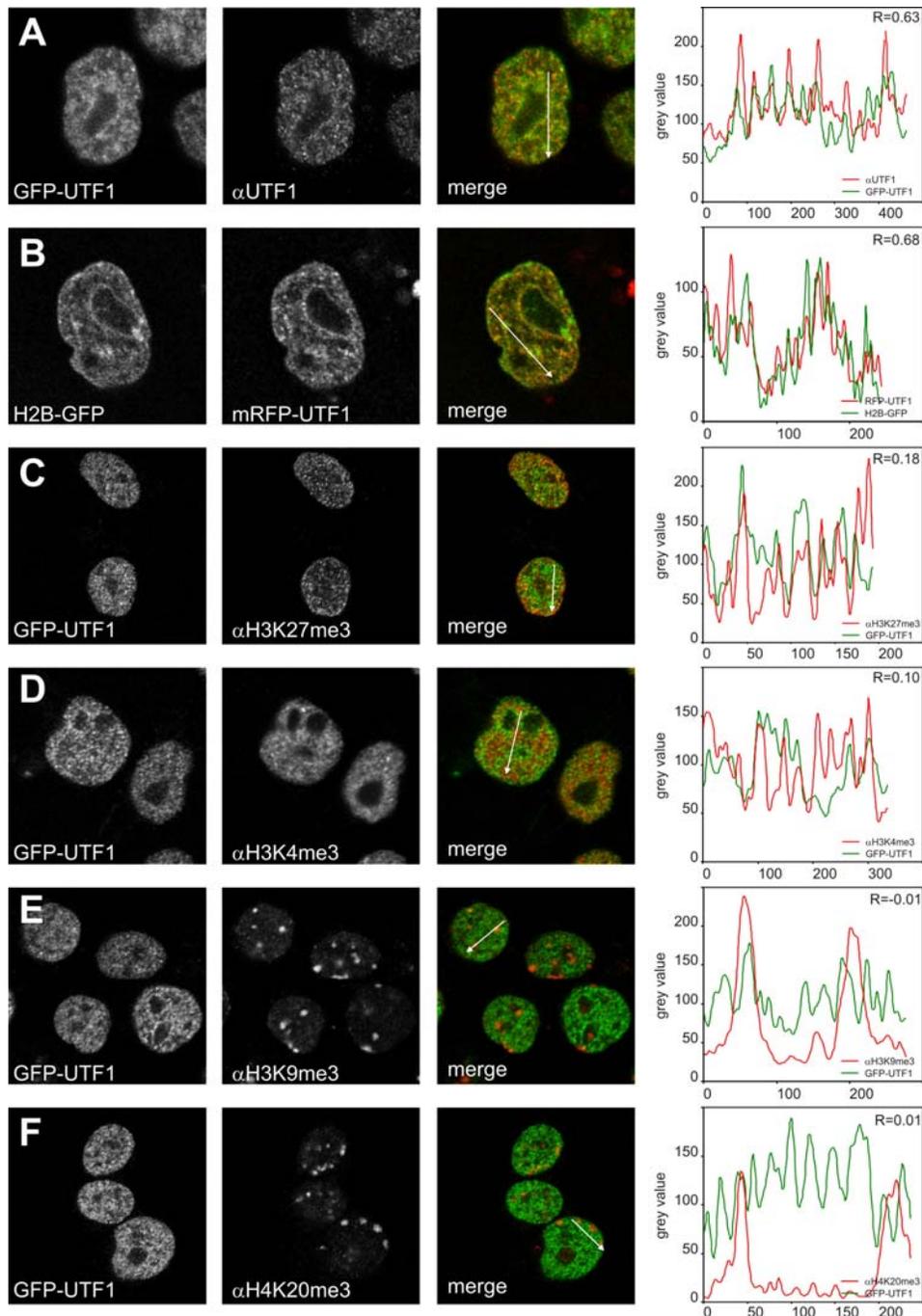
**Figure S11.** Immunofluorescent analysis of P19CL6 cells using antibodies directed against the indicated histone modifications and UTF1. Line scans (white arrows) were performed on deconvoluted confocal images and Pearson's correlation coefficients between red and green channels were calculated using ImageJ.
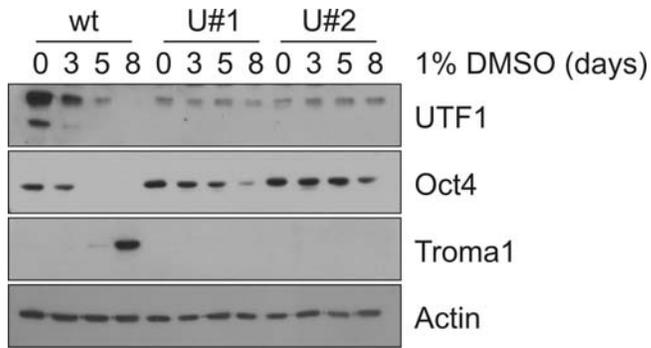
**Figure S12** Western analysis of protein markers during P19CL6 differentiation. UTF1, Oct4, Troma1 and Actin levels were determined at day 0, 3, 5 and 8 days of DMSO treatment.
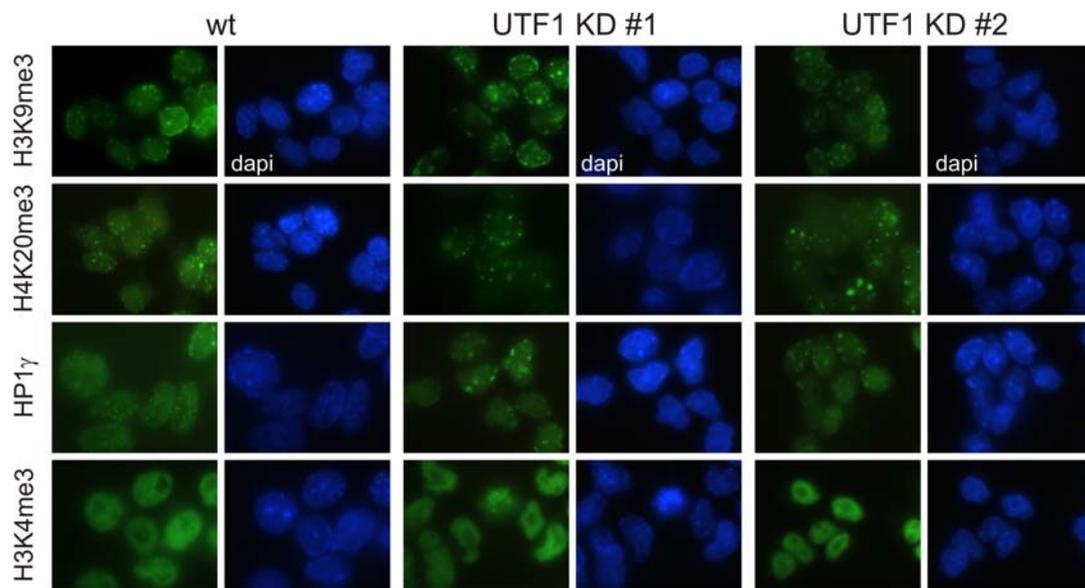


**Figure S13.** Nuclear distribution of H3K9me3, H4K20me3, HP1$\gamma$ and H3K4me3 in wild type and UTF1 knock down EC cells.

Immunofluorescent analysis of wild type (wt) and UTF1 knock down (UTF1 KD#1 and #2) cells using antibodies against H3K9me3, H4K20me3, heterochromatin protein 1$\gamma$ (HP1$\gamma$) and H3K4me3. Cells were counterstained with DAPI.

**Table S1.** List of UTF1 target genes. Chromosomal location, gene name and NCBI gene ID are given.

**Table S2.** List of genes with significantly upregulated expression levels in UTF1 KD ES cells. Chromosomal location, gene name and NCBI gene ID are listed. The log2 expression levels of these transcripts in biological replicates (A and B) in wild type (wt) and two UTF1 KD ES lines (#1 and #2) with their rank products are given.

**Table S3.** List of genes with significantly downregulated expression levels in UTF1 KD ES cells. Chromosomal location, gene name and NCBI gene ID are listed. The log2 expression levels of these transcripts in biological replicates (A and B) in wild type (wt) and two UTF1 KD ES lines (#1 and #2) with their rank products are given.

**Supplemental Materials and Methods**

*Liquid chromatography-mass spectrometry*

Gel fragments were completely destained in 50 mM $NH_4HCO_3$ and 50% acetonitrile (ACN). Subsequently, the proteins in the gel slices were reduced and alkylated by incubation in 10 mM DTT in 50 mM $NH_4HCO_3$ for 45 min at 55°C followed by treatment with 55 mM iodoacetamide in 50 mM $NH_4HCO_3$ for 30 min at room temperature. The gel slices were dehydrated in 100% ACN and rehydrated in 10 µl trypsin solution (Promega V5111; 10 ng/µl in 40 mM $NH_4HCO_3$, 10% ACN), followed by overnight incubation at 37°C. The tryptic peptides were recovered from the gel slices by extraction with different concentrations of ACN and trifluoroacetic acid (TFA) The extracted peptides were pooled and concentrated under vacuum. Next, peptides were separated on C18 capillary column (C18 PepMap 300, 75 µm x 150 mm, 3 µm particle size, LC-Packing, Dionex, The Netherlands) mounted on an UltiMate 3000 nanoflow liquid chromatography system (Dionex, The Netherlands). 0.05% TFA and a 80% acetonitrile in 0.05% TFA solutions were used as mobile phase A and B respectively. A gradient from 4–55% B was performed in 30 min, at a flow rate of 300 nl/min. Column effluent was mixed 1:4 *v/v* with a matrix solution of 2.3 mg/ml α-cyano-4-hydroxycinnamic acid (LaserBio Labs, Sophia-Antipolis, France) containing 2 nM angiotensin II and 4 nM adrenocorticotrophin hormone 18–36 (Sigma, Zwijndrecht, The Netherlands) as internal standards for mass calibration. Fractions of 12-seconds were spotted on blank MALDI targets with a Probot system (LC Packings, Amsterdam, the Netherlands). MS analysis was performed using a MALDI-

TOF/TOF 4800 Proteomics Analyzer (Applied Biosystems, Foster City, CA, USA). MS data acquisition was performed in positive ion mode. The MSMS spectra were acquired using collision-induced dissociation at low air pressure and calibrated using default calibration. Protein identification was carried out using ProteinPilot (version 2.0.1, Applied biosystems/ MDS Sciex) using the Paragon algorithm and searching against the *Mus musculus* protein sequence database at NCBI combined with common keratin and trypsin contaminants. Protein identifications were accepted if they could be established at greater than 99.0% probability and were based on at least two unique peptides.

*PCR and qPCR*

UTF1 target regions and controls and differentially expressed were amplified by PCR using the following primersets:

Midline 1:     F: TGTCCTCTTCTGCAAAACGAATCTTACCG

              R: GCTCATTTCCGAAGCAAGCCAAGAAATC

Ssty1:        F: GAGTCTGGCACCATTTCTGACACTAAGAAC

              R: AGCCTTCAGTTTCTGATTCACCAACATCC

MGC107098:  F: GCAGAAACCGAAGACTTTCTCAGAACG

              R: TGCCAAAGAATTACACCCTTGCATGG

Sly:          F: CTTCCTTTCATTTCTTTCCACCAAACCAGC

              R: TGTGTATTTCATGTACCCAAGAGCATTCCG

Hist2h3c1:    F: AAGCTTTACTTTGAACCAGGCCATAGTAGC

              R: GCCTTGGGTTCCTGAAACTCCCATC

Hist3h2ba:    F: GAGCGACAGCCACATCTCTATGTGAG

              R: AGGCTGGCAAGGTGCCCTG

GAPDH:        F: CCAATGTGTCCGTCGTGGAT

              R: GCTGTTGAAGTCGCAGGAGA

Igf2:   F: TGACACGCTTCAGTTTGTCTGTTCGG

     R: TGACACGCTTCAGTTTGTCTGTTCGG

H19:   F: TCAGAACCACTACACTACCTGCCTCAG

     R: TGGGTGCTATGAGTCTGCTCTTTCAAAATG

P57:   F: GCCGGGTGATGAGCTGGGA

     R: ACCGCCTCCCCATTCTGCTC

Nanog:  F: AGGCAGCCCTGATTCTTCTACC

     R: GCATCTTCTGCTTCCTGGCAAG

Essrb:  F: TATGCCTCCCAACGATATCCCC

     R: CTTGGCCCAGTTGATGAGGAAC

Rhox2:  F: CCGTGGTCCACTGTTAATCCTG

     R: ACTGGAAGATGCTCTCCAGCTC

Klf4:   F: AGATGCAGTCACAAGTCCCCTC

     R: CACGACCTTCTTCCCCTCTTTG

Hist1h2bc: F: GTCGGAGGGCACCAAGGC

     R: TTGAGGTTACAGCATCCAGC


PCR products were analyzed on 2% agarose gels.

To determine the relative fold change, quantitative PCR was performed (MESA GREEN qPCR MasterMix Plus for SYBR Assay w/fluorescein; Eurogentec). Data were analyzed by the $2^{-\Delta\Delta Ct}$ method, using GAPDH as a reference (1).


*Immunofluerescence and microscopy*

For immunofluorescence, cells were cultured on poly-D-lysine coated glass coverslips. Cells were washed with PBS, and then fixed in freshly prepared 2% paraformaldehyde in PBS for 15 minutes at RT. After permeabilization with 0.1% Triton-X-100 in PBS, the cells were incubated with antibodies diluted in PBS with 1.5

mg/ml glycine and 5 mg/ml BSA. Slides were mounted using Vectashield and images were recorded within 24 hours. Fluorescent images were made using an Axiophot microscope (Carl Zeiss MicroImaging, Inc.) with a plan-NEOFLUAR 40×NA 0.70 lens.

*Antibodies*

Western blotting was performed using antibodies recognizing UTF1 (2), actin (C4; MP Biomedicals), Oct4 (H-134; Santa Cruz Biotechnology, Inc.), Troma1 (DSHB), HP1γ (Euromedex; (3)) H3K4me3 (ab8580, AbCam), H3K27me3 (07-449, Upstate), H3K4me2 (ab7766, AbCam), H3K27me2 (07-452, Upstate), H3K27me1 (07-448, Upstate), H4K20me3 (ab9053, Abcam), H3K9me3 (07-442, Upstate) and H3K9Ac (17-658, Upstate). Secondary antibodies for western blotting were donkey anti-rabbit IgG, HRP linked (GE healthcare) , donkey anti-goat IgG (Santa Cruz Biotechnology, Inc.) and goat anti-mouse IgG-HRP (Santa Cruz Biotechnology, Inc.).

1. Livak KJ, Schmittgen TD Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. **Methods** 2001, 25(4):402-428.

2. van den Boom, V, Kooistra SM, Boesjes M, Geverts B, Houtsmuller AB, Monzen K, Komuro I, Essers J, Drenth-Diephuis LJ, Eggen BJL UTF1 is a chromatin-associated protein involved in ES cell differentiation. **The Journal of Cell Biology** 2007 913-924.

3. Nielsen AL, Ortiz JA, You J, Oulad-Abdelghani M, Khechumian R, Gansmuller A, Chambon P, Losson R Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. **EMBO J.** 1999 18(22):6385-6395.