Cell Reports
Report

Cell PRESS
Open ACCESS

Macrohistone Variants Preserve Cell Identity by Preventing the Gain of H3K4me2 during Reprogramming to Pluripotency

María J. Barrero,[1] Borja Sese,[1] Bernd Kuebler,[1] Josipa Bilic,[1] Stephanie Boue,[1] Mercè Martí,[1] and Juan Carlos Izpisua Belmonte[1,2,*]

[1]Center for Regenerative Medicine in Barcelona, Barcelona 08003, Spain

[2]Gene Expression Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037, USA

*Correspondence: belmonte@salk.edu

http://dx.doi.org/10.1016/j.celrep.2013.02.029

SUMMARY

Transcription-factor-induced reprogramming of somatic cells to pluripotency is a very inefficient process, probably due to the existence of important epigenetic barriers that are imposed during differentiation and that contribute to preserving cell identity. In an effort to decipher the molecular nature of these barriers, we followed a genome-wide approach, in which we identified macrohistone variants (macroH2A) as highly expressed in human somatic cells but downregulated after reprogramming to pluripotency, as well as strongly induced during differentiation. Knockdown of macrohistone variants in human keratinocytes increased the efficiency of reprogramming to pluripotency, whereas overexpression had opposite effects. Genome-wide occupancy profiles show that in human keratinocytes, macroH2A.1 preferentially occupies genes that are expressed at low levels and are marked with H3K27me3, including pluripotency-related genes and bivalent developmental regulators. The presence of macroH2A.1 at these genes prevents the regain of H3K4me2 during reprogramming, imposing an additional layer of repression that preserves cell identity.

INTRODUCTION

Chromatin structure plays fundamental roles in the regulation of gene expression during development and contributes to defining and preserving cell identity. The nucleosome, the basic subunit of eukaryotic chromatin, is composed of two molecules of each of the core histone proteins—H2A, H2B, H3, and H4—around which DNA is wrapped. Histones are among the most highly conserved proteins in terms of both structure and sequence, but in higher organisms, each histone subtype is represented by a family of genes encoding multiple variants. Although most variants appear highly similar and likely redundant (for example, all H4 family members encode the same protein), others show unique features and play roles in critical aspects of chromatin structure regulation and are, therefore, known as replacement variants. Such is the case of the H3 variants H3.3 (Ahmad and Henikoff, 2002) and CENPA (Sullivan et al., 1994), and the H2A variants H2A.Z (Faast et al., 2001), H2A.X (Rogakou et al., 1998), and macroH2A (Pehrson and Fried, 1992).

The macroH2A variant contains an amino-terminal histone-like region that is similar to histone H2A and a carboxy-terminal globular domain called macrodomain (Pehrson and Fried, 1992). Although generally referred to as macroH2A, two different genes, *H2AFY* and *H2AFY2*, encode for two macroH2A isoforms called macroH2A.1 and macroH2A.2, respectively (Chadwick and Willard, 2001). In addition, the mRNA of macroH2A.1 can be alternatively spliced, giving rise to the isoforms macroH2A.1.1 and 1.2 (Rasmussen et al., 1999). The macrodomain protrudes out of the nucleosome and in the case of macroH2A.1.1 variant, can act as a ligand binding domain for metabolites of NAD+ (Karras et al., 2005). MacroH2A has been suggested to play roles in transcriptional repression by physically impeding the access of transcription factors to their binding sites and blocking the action of ATP-dependent chromatin remodelers such as the SWI/SNF complex (Angelov et al., 2003). Moreover, the macrodomain can interact with histone deacetylases (HDACs) (Chakravarthy et al., 2005) and certain subunits of the Polycomb complex (Buschbeck et al., 2009) and, thus, might participate in the recruitment or stabilization of repressor complexes at defined genomic regions. Accordingly, macroH2A has been suggested to be involved in X inactivation (Costanzi and Pehrson, 1998), although more recent reports have located macroH2A in several autosomal genes (Buschbeck et al., 2009; Gamble et al., 2010). Despite the fact that macroH2A occupies mostly transcriptionally inactive genes, its knockdown has been reported to block the induction of genes by serum starvation (Gamble et al., 2010). Therefore, the potential involvement of macrohistone variants in transcriptional activation remains controversial.

Several reports have described the induction of macrohistone variants during the in vitro differentiation of mouse embryonic stem cells (ESCs) (Creppe et al., 2012; Dai and Rasmussen, 2007; Pehrson et al., 1997), and recently, they have been suggested to confer resistance to nuclear reprogramming of mouse cells (Pasque et al., 2011, 2012). However, the regulation of the expression of these variants and its mechanistic consequences remain practically unaddressed in human cells.

The epigenetic mechanisms that contribute to establish cell identity during differentiation and to preserve it in somatic cells are mostly unknown. These mechanisms, which presumably protect cells against transformation, constitute an important barrier to reprogramming (Barrero et al., 2010). Our data show that macroH2A.1 expression protects human cells against reprogramming to pluripotency, probably through the repression of pluripotency genes and bivalent germ layer-specific genes, where it prevents the gain of H3K4me2 that takes place during this process.

RESULTS

Macrohistone Variants Are Mainly Expressed in Somatic Cells and Downregulated after Reprogramming

We previously (Boué et al., 2010) compiled genome-wide expression data from several independent reprogramming experiments and extracted a list of candidates that showed differential expression between human induced pluripotent stem cells (iPSCs) and somatic cells and identified the macroH2A isoform macroH2A.1 as one of the top 1,000 differentially expressed genes. The ratio of expression levels of several histone variants in somatic cells to that in ESCs or to their corresponding iPSCs shows that the gene encoding for macroH2A.1 (*H2AFY*) is consistently expressed at higher levels in somatic cells (Figure 1A). qPCR confirmed that macroH2A.1 was upregulated in keratinocytes compared to iPSCs derived from keratinocytes (KiPSCs) (Figure 1B). However, macroH2A.2 (*H2AFY2*) was expressed at lower levels, and the differences between pluripotent and somatic cells were milder (Figures 1A and 1B). Western blot (Figure 1C) and immunocytochemistry (Figures S1A and S1B) showed that both macroH2A.1 and macroH2A.2 were almost undetectable in pluripotent cells (ESCs and iPSCs) but clearly expressed in somatic cell types (keratinocytes and fibroblasts), whereas the levels of the canonical histone H2A remained unchanged (Figure 1C). Interestingly, the presence of macroH2A.1 in human ESCs (hESCs) grown under self-renewal conditions could be readily detected in a few cells that underwent spontaneous differentiation characterized by the loss of expression of the pluripotency marker OCT4 and expression of the differentiation marker FOXA2 (Figure S1C).

Macrohistone Variants Are Strongly Induced during the Differentiation of Human Pluripotent Cell Lines

In agreement with the prevalent expression of macroH2A in somatic cells, macroH2A.1 and macroH2A.2 were the most conspicuously induced core histone variants during the differentiation of hESCs (Figure 1A). These results were confirmed by qPCR during the differentiation of two pluripotent cell lines (Figure 1D), western blot (Figure 1C), and immunocytochemistry (Figures S1A and S1B). At day 20 of embryoid body (EB) differentiation, the expression of macroH2A.1 and macroH2A.2 could be detected in most cells' nuclei and colocalized with markers of all three germ layers (Figures S2A and S2B), suggesting that its induction takes place in cells of all three embryonic layers.

We also tested the possibility that macrohistone variants could be induced during the differentiation of adult stem cells. Neuronal precursors within neural rosettes marked with SOX9 expressed very low levels of macroH2A.2, but its expression could be clearly detected at peripheral mature neurons positive for TUJ1 and MAP2 (Figure S2C), indicating that macroH2A.2 expression is induced during the differentiation of neural precursors.

The Knockdown of Macrohistone Variants Increases the Reprogramming Efficiency

Because macrohistone variants are prevalent in somatic cells and downregulated during their reprogramming to pluripotency, we asked whether the knockdown of these proteins could affect the efficiency of reprogramming. For that, human keratinocytes were infected four independent times with GFP and previously described shRNA (Buschbeck et al., 2009) encoding lentiviruses that were able to efficiently knock down the expression of macroH2A.1 (shM1) or macroH2A.2 (shM2) (Figures 2A and 2B), and the efficiency of reprogramming to keratinocytes infected with a nontarget shRNA (shRD) was compared (Figure 2C). The knockdown of macroH2A.1 increased the number of alkaline phosphatase (AP)-positive colonies in all four reprogramming experiments. However, the knockdown of macroH2A.2 had milder effects on the efficiency of reprogramming. These results are in agreement with the predominant expression of macroH2A.1 compared to macroH2A.2 in human keratinocytes (Figures 1B, S3A, and S3B).

About 3 weeks after viral transduction, reprogrammed colonies were picked and expanded. Three lines of KiPSCs (induced pluripotent cells from keratinocytes) were established for each condition, which remained AP positive after several passages (Figure S4A) and showed expression of the endogenous pluripotency markers (Figures S4B and S4C). The downregulation of the expression of the transgenes (Figure S4D) further confirmed that cells had reached a fully reprogrammed pluripotent state. Karyotype was found to be normal (Figure S4E). Moreover, all the tested lines were able to produce teratomas positive for markers of the three germ layers in immunosuppressed mice (Figure S5A). Importantly, the expression of macrohistone variants was downregulated in all KiPSC clones compared to the original keratinocytes (Figures S5B and S5C).

To confirm that the effects of knocking down the macrohistone variants in the reprogramming of human keratinocytes are specific, we tested the effects of macroH2A.1 overexpression (Figure 2D). Figure 2E shows that the number of AP-positive colonies obtained after the reprogramming of keratinocytes that overexpress a HA-tagged version of macroH2A.1 is significantly reduced compared to keratinocytes transduced with an empty vector.

MacroH2A.1 Occupies Pluripotency and Bivalent Genes in Somatic Cells

To gain insight into the process by which macrohistone variant expression affects the efficiency of reprogramming, we analyzed the genome-wide occupancy of macroH2A.1 by chromatin immunoprecipitation sequencing (ChIP-seq) in human keratinocytes. Because the available specific antibodies against macroH2A.1 were not displaying good enough signal to background ratios to carry out genome-wide studies, we took advantage of the established lines of keratinocytes that overexpress the HA-tagged version of macroH2A.1 and used antibodies

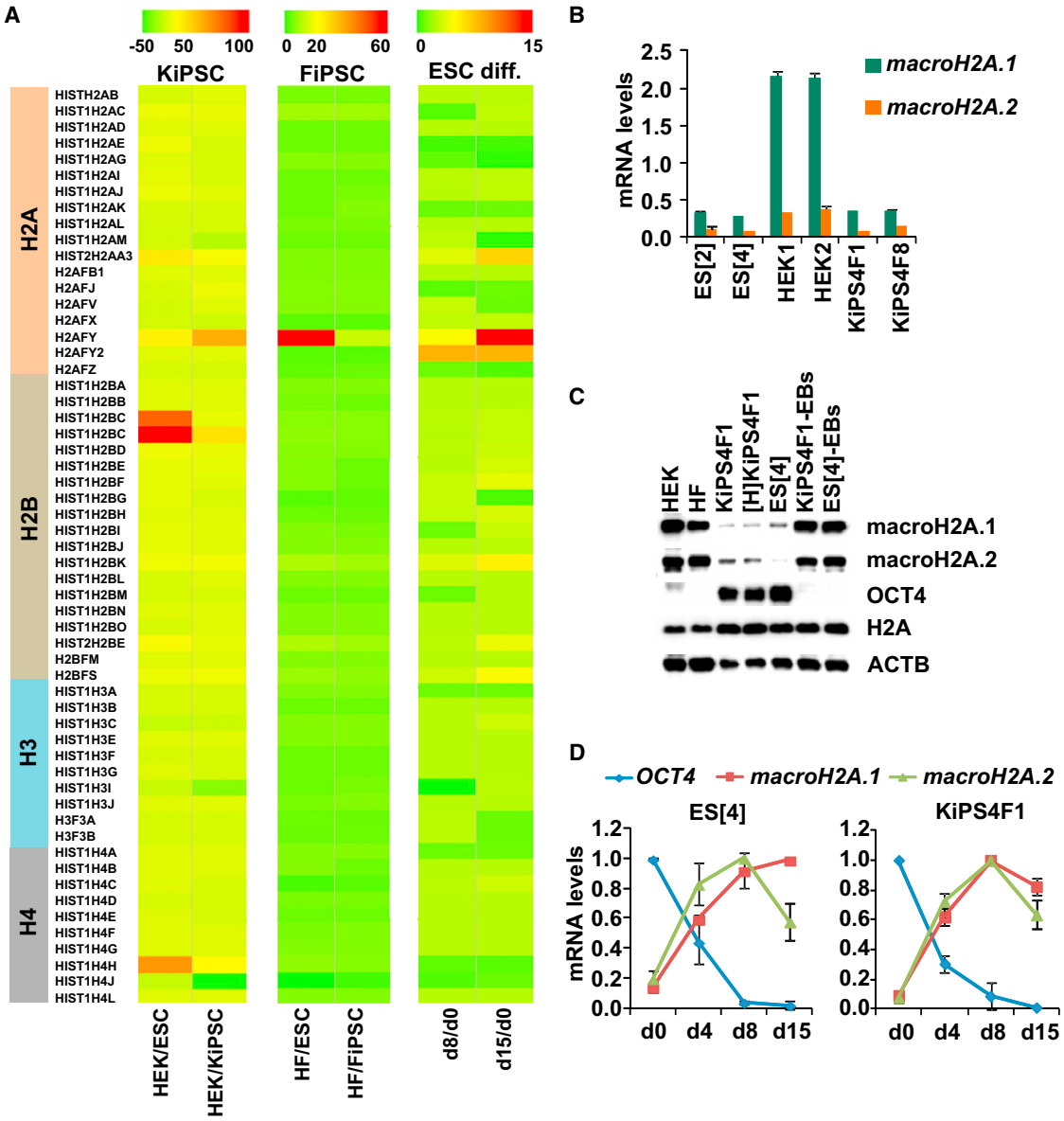

Figure 1. Differential Expression of Macrohistone Variants in Human Pluripotent and Differentiated Cells

(A) Heatmaps showing log fold changes of expression levels. Left panel shows fold change between keratinocytes (HEK) and hESCs, and keratinocytes and KiPSCs. Middle panel shows fold change between fibroblasts (HF) and hESCs, and fibroblast and induced pluripotent cells from fibroblast (FiPSCs). Right panel shows fold change between undifferentiated ESCs and cells at day 8 or 15 (d8 or d15) of differentiation.

(B) mRNA levels of macroH2A.1 and macroH2A.2 in two lines of keratinocytes (HEK1 and HEK2), two lines of hESCs (ES[4] and ES[2]), and two lines of KiPSCs (KiPS4F1 and KiPS4F8) determined by qPCR. Primers for macroH2A.1 amplify both splice variants. Levels were corrected according to the amplification efficiency of each pair of oligonucleotides and normalized to GAPDH, HPRT, and HMBS. The mean and SD (n = 3) are shown.

(C) Protein levels of macroH2A.1, macroH2A.2, OCT4, and canonical H2A determined by western blot in keratinocytes (HEK), fibroblasts (HF), KiPSCs (KiPS4F1 and [H]KiPS4F1), ESCs (ES[4]), and EBs from ES[4] and KiPS4F1 at day 15 of differentiation. Levels of ACTB were used as a loading control.

(D) mRNA levels of macroH2A.1, macroH2A.2, and OCT4 during the differentiation of ES[4] or KiPS4F1 determined by qPCR. Levels were normalized to GAPDH and represented relative to the maximum expression. The mean and SD of two independent experiments per cell line are shown.

See Figures S1 and S2 for immunostaining analysis. See also Table S2.

against the HA tag to perform the ChIP. To rule out potential unspecific binding of the tagged histone to chromatin, we confirmed its differential occupancy in two cell lines of different germ layer origin (Figure S6A) and that its expression levels resembled those of the endogenous protein (Figure 2D).

Genomic intervals occupied by macroH2A.1 showed larger overlap with those marked with H3K27me3 and lower with those marked with H3K36me3, H3K4me3, or occupied by RNA Polymerase II (Pol II) (Figure S6B), suggesting that macroH2A.1 is associated with transcriptional repression. Analysis of binding

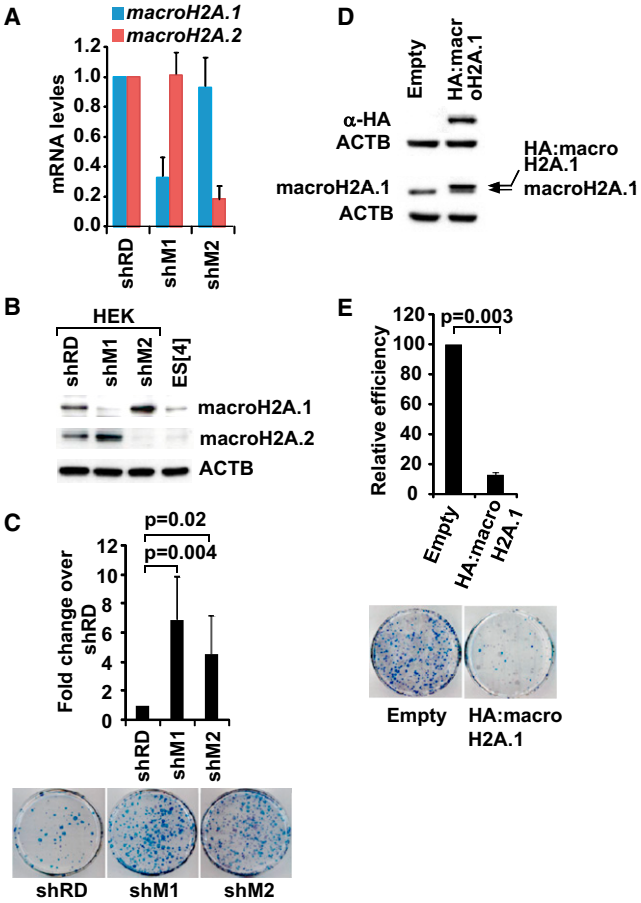

Figure 2. MacroH2A Knockdown Facilitates Reprogramming

(A) Levels of mRNA for each variant quantified by qPCR, normalized to the levels of GAPDH, and plotted relative to the control levels (shRD). Means and SDs (n = 3) are shown.

(B) Protein levels for each variant determined by western blot. ACTB was used as loading control. Relative protein levels of macrohistone variants in human keratinocytes can be found in Figure S3.

(C) Upper panel shows the mean and SD of the number of AP-positive colonies obtained after four independent reprogramming experiments of keratinocytes knocked down for macroH2A.1 (shM1) or macroH2A.2 (shM2) or control (shRD). Each experiment was carried out in duplicate. Lower panel shows a representative AP staining of colonies after the reprogramming of human keratinocytes knocked down for the macrohistone variants. See Figures S4 and S5 for characterization of the established KiPSC lines.

(D) Western blot with antibodies against HA or against macroH2A.1 in keratinocytes transduced with empty or HA:macroH2A.1-overexpressing vectors.

(E) Relative efficiency of reprogramming in keratinocytes transduced with empty or HA:macroH2A.1-overexpressing vectors. AP-positive colonies were counted and plotted as percent relative to the empty vector. The mean and SD (n = 3) are shown. Lower panel shows a representative AP staining of colonies after the reprogramming of human keratinocytes overexpressing HA:macroH2A.1.

See also Table S2.

enrichment along the genome (Figure S6C) shows that macroH2A.1 is mainly located at promoter regions and depleted from distal intergenic regions. Therefore, we examined the presence of macroH2A.1 at ±5 kb from the TSS of known UCSC genes (Table S1). MacroH2A.1 occupies about one-fourth of total genes, which are depleted of RNA Pol II ($p = 1.57 \times 10^{-206}$) and H3K36me3 ($p = 0$) (Figure 3A) and are enriched in H3K27me3 ($p = 0$) (Figures 3B and S6D). In accordance, the presence of macroH2A.1 correlated with lower levels of gene expression (Figure S6E).

To clarify the role of macrohistone variants in reprogramming, we compared the profile of histone modifications around the TSS of macroH2A.1 target genes in keratinocytes and hESCs (Figure 3C). According to this, macroH2A.1 target genes could be distributed in five major clusters. Importantly, clusters 1, 3, and 4 correspond to genes that show increased levels of H3K4me2 in pluripotent cells compared to keratinocytes, a mark that has been suggested to change early during reprogramming (Koche et al., 2011). Both clusters 1 and 3 are enriched in genes marked with bivalent domains in both keratinocytes and ESCs ($p = 0$) that are involved in developmental processes ($p = 4.80 \times 10^{-64}$ and $p = 9.10 \times 10^{-22}$ for clusters 1 and 3, respectively). Interestingly, cluster 1 was found to be enriched in transcription factors ($p = 1.80 \times 10^{-83}$) that are tightly repressed in both keratinocytes and ESCs (Figure 3D), whereas cluster 3 was enriched in transmembrane activities ($p = 3.30 \times 10^{-20}$) and showed a more permissive expression in ESCs (Figure 3D). Cluster 4 corresponds to genes that are highly expressed in ESCs compared to keratinocytes (Figure 3D), including important pluripotency transcription factors such as *NANOG* and *SALL2*. Genes included in cluster 5 are silent and show DNA hypermethylation in both cell types ($p = 0$) with a low content in CpG islands ($p = 0$) and likely correspond to tissue-specific genes (Barrero et al., 2012; Fouse et al., 2008). Cluster 2 contains a large number of macroH2A.1 target genes that are marked with H3K4me3 alone in human keratinocytes and are highly expressed in both ESCs and keratinocytes; however, the levels of macroH2A.1 at these genes are low compared to genes marked with H3K27me3 (Figure 3E).

The Knockdown of MacroH2A.1 Facilitates the Regain of H3K4me2 at Pluripotency Genes during Reprogramming

Our data suggest potential antagonistic roles for macroH2A.1 occupancy and presence of H3K4me2 because (1) pluripotency genes are active and marked with H3K4me2 in ESCs, while occupied by macroH2A.1 and silent in keratinocytes; and (2) the levels of macroH2A.1 at bivalent genes drop at the TSS where the levels of H3K4 methylation peak (Figure 3E). Therefore, we asked if the knockdown of macroH2A.1 could be facilitating the gain of H3K4me2 at pluripotency-related and developmental genes during reprogramming. We performed ChIP assays to monitor the gain of H3K4me2 at macroH2A.1 target genes during the early stages of reprogramming. Figure 4 shows that cells knocked down for macroH2A.1 show higher levels of H3K4me2 at four macroH2A.1 target genes compared to control cells at days 4 and 8 of reprogramming.

DISCUSSION

MacroH2A.1 was selected as a differentially expressed candidate based on the analysis of genome-wide expression data from several reprogramming experiments (Boué et al., 2010).

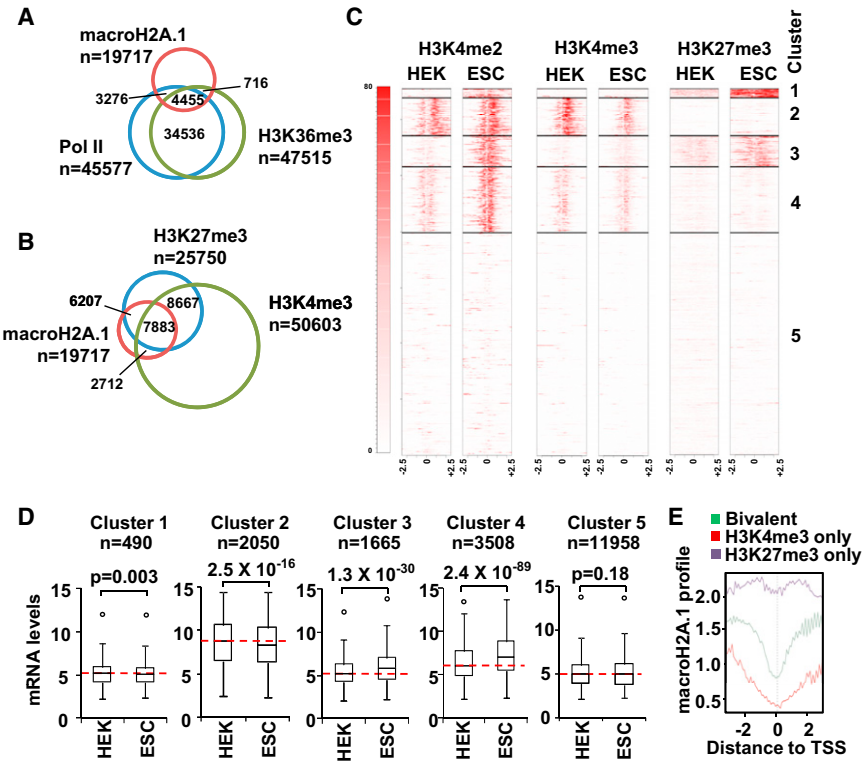

Figure 3. Overlap of MacroH2A.1 Occupancy and Histone Modifications

(A) Intersect of USCS known genes (n = 80,922) occupied by macroH2A.1 at ±5 kb of the TSS and presence of RNA Pol II and H3K36me3.

(B) Intersect of USCS known genes (n = 80,922) occupied by macroH2A.1 at ±5 kb of the TSS and presence of H3K4me3 and H3K27me3.

(C) Profile of histone modifications in human keratinocytes (HEK) and in hESCs at ±2.5 kb around the TSS of macroH2A.1 target genes.

(D) Box blot of the mRNA levels of genes belonging to the indicated clusters in keratinocytes (HEK) and ESCs. Number of genes in each category and p values are shown.

(E) Frequency of macroH2A.1 signal at 3 kb around the TSS of target genes marked with bivalent domains, H3K27me3 only, or H3K4me3 only. See also Figure S6. List of genes occupied by macroH2A.1 can be found in Table S1. See also Table S2.

Comparison of the levels of expression of histone variants between reprogramming experiments (Figure 1A) revealed histone variants, such as HIST1H2BC, with differential expression depending on the cell of origin. However, macroH2A.1 was the most consistent differentially expressed variant among all experiments. Although macroH2A.2 follows a similar pattern of differential expression between somatic and pluripotent cells at the protein level (Figure 1C), it did not show up in our list of the top 1,000 differentially regulated genes, likely due to less dramatic differences in expression of this gene at the mRNA level. Both variants have been reported likely to be recruited to the same set of target genes (Gamble et al., 2010), suggesting that their functions might be redundant. However, the prevalent expression of macroH2A.1 over macroH2A.2 may explain why the macroH2A.1 knockdown had more consistent functional consequences for reprogramming, whereas the knockdown of macroH2A.2 showed milder effects, most likely dependent on whether the efficiency of its knockdown could contribute to significantly reduce the total macrohistone pool.

In agreement with our results, it has been reported that macroH2A is quickly degraded after nuclear transfer (Chang et al., 2010), and a recent report describes that the knockdown of macroH2A in mouse cells facilitates nuclear transfer and highlights a role for macroH2A in X chromosome inactivation (Pasque et al., 2011). However, our findings argue against a major role of macroH2A in X chromosome inactivation because macroH2A is barely expressed in the female human pluripotent cell line [H] KiPS4F1 in which the X chromosome is inactivated (Barrero et al., 2012; Tchieu et al., 2010), and we found functional implications for the reprogramming of male cells. Our results indicate that macroH2A.1 occupies a subset of pluripotency and bivalent developmental genes marked with H3K27me3 in somatic cells, where it likely provides an additional layer of repression that prevents the gain of H3K4me2 at those genes during reprogramming to pluripotency. Our findings are in agreement with a recent report describing that the gain of H3K4me2 methylation at pluripotency-related and developmental genes is a key event that takes place during the early stages of reprogramming (Koche et al., 2011). The fact that the macrodomain has been previously shown to block the access of transcription factors to their binding sites (Angelov et al., 2003) tempts us to speculate that the presence of macroH2A.1 might be compromising the accessibility of the reprogramming factors to their binding sites. However, further dissection of the reprogramming process will be needed in order to determine the critical responsive elements that might be influenced by the presence of macroH2A.1 during the reprogramming of human cells.

In humans, macroH2A.1 was found specifically downregulated in a significant fraction of breast and lung tumors, and its reduced expression correlated with the risk of cancer recurrence (Sporn et al., 2009). Similarly, macroH2A loss has also been reported to correlate with increasing melanoma malignancy (Kapoor et al., 2010). These data, together with our results, suggest that these histone variants act as critical players that preserve cell identity in humans.

EXPERIMENTAL PROCEDURES

Cell Culture

The hESC lines ES[4] and ES[2] (Raya et al., 2008) and KiPSCs lines (Aasen et al., 2008) were grown in Matrigel-coated plates, hESC media, and subcultured using trypsin. For in vitro differentiation, cells were trypsinized and seeded into 96-well v-bottom plates in hESC media. After 3 days, EBs were transferred to gelatin-coated dishes and cultured in differentiation media up to 20 days. Human keratinocytes were isolated from juvenile foreskin as

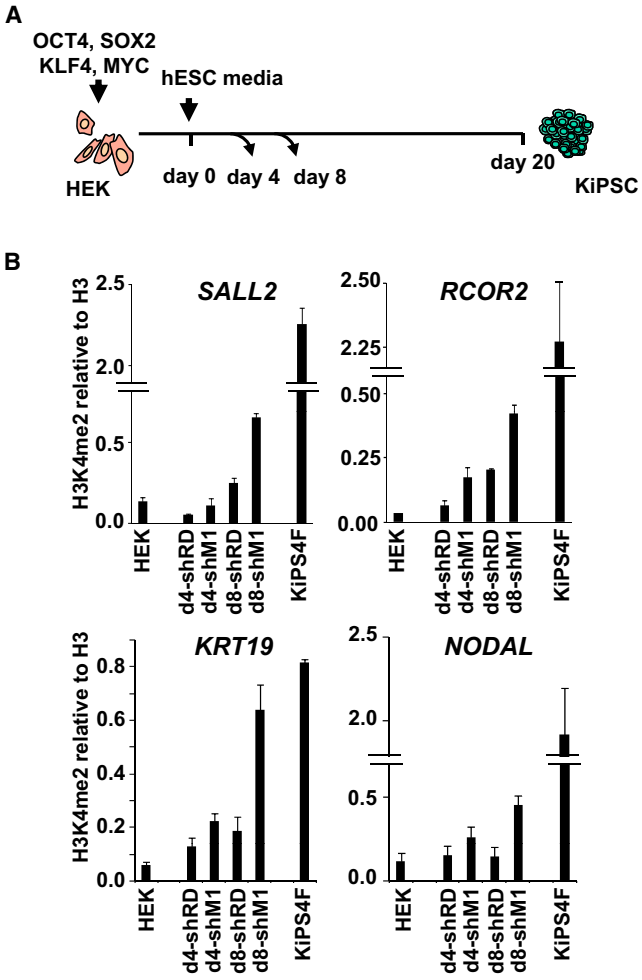

Figure 4. MacroH2A.1 Knockdown Facilitates the Gain of H3K4me2 during Reprogramming

(A) Timeline of the reprogramming experiment.

(B) The levels of H3K4me2 relative to total H3 content at the regulatory regions of the indicated genes were determined by ChIP in human keratinocytes (HEK), KiPSCs, and in two independent reprogramming experiments of keratinocytes control (shRD) and keratinocytes knocked down for macroH2A.1 (shM1) collected at days 4 or 8 after addition of hESC media. Means and SDs are shown.

See also Table S2.

described by Aasen et al. (2008) and cultured in EpiLife (Invitrogen). Human fibroblasts were isolated and expanded in DMEM supplemented with 10% fetal bovine serum and 2 mM L-glutamine.

This study was done in accordance with Spanish laws and regulations regarding the generation of human iPSCs and following protocols approved by the Spanish competent authorities (Comisión de Seguimiento y Control de la Donación de Células y Tejidos Humanos del Instituto de Salud Carlos III).

Histone Variant Expression Profiling

Gene expression profiles for all core histone variants were extracted from GSE12583, GSE9865, GSE9832, GSE13828, GSE15148, GSE14711, GSE12390, and GSE16093. Data were imported in R, normalized by experiments using GCRMA, and a log fold change was calculated between the states of interest. Heatmaps showing log fold changes were generated using the heatmap2 function from the gplots package in R.

Reprogramming of Human Keratinocytes

At passage 2, keratinocytes were transduced with lentiviral vectors to knock down the expression of macroH2A.1 or macroH2A.2 or to overexpress HA:macroH2A.1. Infections were done at a high moi (five to ten) to ensure at least 90% of infection efficiency. At passage 3, 50,000–100,000 cells were infected with retroviruses encoding for Oct4, Sox2, Klf4, and c-Myc by spinfection as previously described by Aasen et al. (2008), transferred to Matrigel-coated dishes, and cultured in hESC media. AP staining or expansion of clones was performed after 3 weeks.

ChIP-Seq

About 4 × 10^6 cells were used for immunoprecipitation using anti-HA antibodies (Ab9110 from Abcam) as described in Extended Experimental Procedures. Two independent immunoprecipitations were performed. Libraries for input and immunoprecipitated material were constructed selecting fragments from 200 to 500 bp and processed in the Solexa GAIIx Sample Sequencing (single reads of 36 nt). Sequence analysis was carried out using Galaxy (https://main.g2.bx.psu.edu/) and Galaxy Cistrome (http://cistrome.org/). Biological replicates were mapped against the human genome using Bowtie (Langmead et al., 2009) and pooled, resulting in more than 40 × 10^6 uniquely mapped reads. Genomic intervals bound by macroH2A.1 were called using SICER (Zang et al., 2009) (window size, 200; fragment size, 260; gap size, 600; FDR, 0.01). The relative enrichment level of ChIP regions in each gene feature with respect to the whole genome and profiles around the TSS was analyzed using *cis*-regulatory element annotation system (Shin et al., 2009). Data corresponding to histone modifications and Pol II occupancy in human keratinocytes and ESCs were generated in the Bradley E. Bernstein lab and downloaded from the ENCODE Project at UCSC (http://genome.ucsc.edu/). DNA methylation data in hESCs and keratinocytes are described in Barrero et al. (2012).

Statistical Methods

p values regarding the enrichment of macroH2A.1 occupancy at given genomic locations were calculated using the chi-square test. The significance of the differences in gene expression between specific categories was analyzed using the Student's t test. p values for gene ontology functional annotation were obtained using the DAVID analysis tool (http://david.abcc.ncifcrf.gov/).

ACCESSION NUMBERS

Chip-seq data have been deposited into the NCBI Gene Expression Omnibus database under accession number GSE44400.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, two tables, and Extended Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.02.029.

LICENSING INFORMATION

ACKNOWLEDGMENTS

We would like to thank J. Castaño, M. Carrió, C. Gómez, and L. Casano and the platforms at the Center for Regenerative Medicine in Barcelona for technical assistance and members of the J.C.I.B. laboratory for feedback and discussions. We also thank S. Malik for discussion on ChIP-seq data and A. Jordan for advice on the immunoprecipitation of HA-tagged proteins. Antibodies against macroH2A.2 and short hairpin vectors were provided by M. Bushbeck. This work was supported by grants from MICINN (RYC-2007-01510 and SAF2009-08588 to M.J.B.). M.J.B. is partially supported by the

Ramón y Cajal program. B.S. is a recipient of a FPU predoctoral fellowship from MICINN. J.B. is a recipient of a Juan de la Cierva fellowship from the MICINN. Work in the laboratory of J.C.I.B. was funded by the G. Harold and Leila Y. Mathers Charitable Foundation, The Leona M. and Harry B. Helmsley Charitable Trust, TERCEL-ISCIII-MINECO, CIBER-BBN, and Fundacion Cellex.

Received: November 16, 2012
Revised: January 22, 2013
Accepted: February 28, 2013
Published: March 28, 2013

REFERENCES

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilić, J., Pekarik, V., Tiscornia, G., et al. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat. Biotechnol. *26*, 1276–1284.

Ahmad, K., and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol. Cell *9*, 1191–1200.

Angelov, D., Molla, A., Perche, P.Y., Hans, F., Côté, J., Khochbin, S., Bouvet, P., and Dimitrov, S. (2003). The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. Mol. Cell *11*, 1033–1041.

Barrero, M.J., Boué, S., and Izpisúa Belmonte, J.C. (2010). Epigenetic mechanisms that regulate cell identity. Cell Stem Cell *7*, 565–570.

Barrero, M.J., Berdasco, M., Paramonov, I., Bilic, J., Vitaloni, M., Esteller, M., and Izpisua Belmonte, J.C. (2012). DNA hypermethylation in somatic cells correlates with higher reprogramming efficiency. Stem Cells *30*, 1696–1702.

Boué, S., Paramonov, I., Barrero, M.J., and Izpisúa Belmonte, J.C. (2010). Analysis of human and mouse reprogramming of somatic cells to induced pluripotent stem cells. What is in the plate? PLoS One *5*, e12664.

Buschbeck, M., Uribesalgo, I., Wibowo, I., Rué, P., Martin, D., Gutierrez, A., Morey, L., Guigó, R., López-Schier, H., and Di Croce, L. (2009). The histone variant macroH2A is an epigenetic regulator of key developmental genes. Nat. Struct. Mol. Biol. *16*, 1074–1079.

Chadwick, B.P., and Willard, H.F. (2001). Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. Hum. Mol. Genet. *10*, 1101–1113.

Chakravarthy, S., Gundimella, S.K., Caron, C., Perche, P.Y., Pehrson, J.R., Khochbin, S., and Luger, K. (2005). Structural characterization of the histone variant macroH2A. Mol. Cell. Biol. *25*, 7616–7624.

Chang, C.C., Gao, S., Sung, L.Y., Corry, G.N., Ma, Y., Nagy, Z.P., Tian, X.C., and Rasmussen, T.P. (2010). Rapid elimination of the histone variant MacroH2A from somatic cell heterochromatin after nuclear transfer. Cell. Reprogram. *12*, 43–53.

Costanzi, C., and Pehrson, J.R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature *393*, 599–601.

Creppe, C., Janich, P., Cantariño, N., Noguera, M., Valero, V., Musulén, E., Douet, J., Posavec, M., Martín-Caballero, J., Sumoy, L., et al. (2012). MacroH2A1 regulates the balance between self-renewal and differentiation commitment in embryonic and adult stem cells. Mol. Cell. Biol. *32*, 1442–1452.

Dai, B., and Rasmussen, T.P. (2007). Global epiproteomic signatures distinguish embryonic stem cells from differentiated cells. Stem Cells *25*, 2567–2574.

Faast, R., Thonglairoam, V., Schulz, T.C., Beall, J., Wells, J.R., Taylor, H., Matthaei, K., Rathjen, P.D., Tremethick, D.J., and Lyons, I. (2001). Histone variant H2A.Z is required for early mammalian development. Curr. Biol. *11*, 1183–1187.

Fouse, S.D., Shen, Y., Pellegrini, M., Cole, S., Meissner, A., Van Neste, L., Jaenisch, R., and Fan, G. (2008). Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell Stem Cell *2*, 160–169.

Gamble, M.J., Frizzell, K.M., Yang, C., Krishnakumar, R., and Kraus, W.L. (2010). The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. Genes Dev. *24*, 21–32.

Kapoor, A., Goldberg, M.S., Cumberland, L.K., Ratnakumar, K., Segura, M.F., Emanuel, P.O., Menendez, S., Vardabasso, C., Leroy, G., Vidal, C.I., et al. (2010). The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. Nature *468*, 1105–1109.

Karras, G.I., Kustatscher, G., Buhecha, H.R., Allen, M.D., Pugieux, C., Sait, F., Bycroft, M., and Ladurner, A.G. (2005). The macro domain is an ADP-ribose binding module. EMBO J. *24*, 1911–1920.

Koche, R.P., Smith, Z.D., Adli, M., Gu, H., Ku, M., Gnirke, A., Bernstein, B.E., and Meissner, A. (2011). Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell *8*, 96–105.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Pasque, V., Gillich, A., Garrett, N., and Gurdon, J.B. (2011). Histone variant macroH2A confers resistance to nuclear reprogramming. EMBO J. *30*, 2373–2387.

Pasque, V., Radzisheuskaya, A., Gillich, A., Halley-Stott, R.P., Panamarova, M., Zernicka-Goetz, M., Surani, M.A., and Silva, J.C. (2012). Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. J. Cell Sci. *125*, 6094–6104.

Pehrson, J.R., and Fried, V.A. (1992). MacroH2A, a core histone containing a large nonhistone region. Science *257*, 1398–1400.

Pehrson, J.R., Costanzi, C., and Dharia, C. (1997). Developmental and tissue expression patterns of histone macroH2A1 subtypes. J. Cell. Biochem. *65*, 107–113.

Rasmussen, T.P., Huang, T., Mastrangelo, M.A., Loring, J., Panning, B., and Jaenisch, R. (1999). Messenger RNAs encoding mouse histone macroH2A1 isoforms are expressed at similar levels in male and female cells and result from alternative splicing. Nucleic Acids Res. *27*, 3685–3689.

Raya, A., Rodríguez-Pizà, I., Arán, B., Consiglio, A., Barri, P.N., Veiga, A., and Izpisúa Belmonte, J.C. (2008). Generation of cardiomyocytes from new human embryonic stem cell lines derived from poor-quality blastocysts. Cold Spring Harb. Symp. Quant. Biol. *73*, 127–135.

Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S., and Bonner, W.M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J. Biol. Chem. *273*, 5858–5868.

Shin, H., Liu, T., Manrai, A.K., and Liu, X.S. (2009). CEAS: cis-regulatory element annotation system. Bioinformatics *25*, 2605–2606.

Sporn, J.C., Kustatscher, G., Hothorn, T., Collado, M., Serrano, M., Muley, T., Schnabel, P., and Ladurner, A.G. (2009). Histone macroH2A isoforms predict the risk of lung cancer recurrence. Oncogene *28*, 3423–3428.

Sullivan, K.F., Hechenberger, M., and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J. Cell Biol. *127*, 581–592.

Tchieu, J., Kuoy, E., Chin, M.H., Trinh, H., Patterson, M., Sherman, S.P., Aimiuwu, O., Lindgren, A., Hakimian, S., Zack, J.A., et al. (2010). Female human iPSCs retain an inactive X chromosome. Cell Stem Cell *7*, 329–342.

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics *25*, 1952–1958.