Published in final edited form as:
Nature. 2014 March 6; 507(7490): 104–108. doi:10.1038/nature12942.

Citrullination regulates pluripotency and histone H1 binding to chromatin

Maria A. Christophorou[1,*], **Gonçalo Castelo-Branco**[1,3,*], **Richard P. Halley-Stott**[1,4], **Clara Slade Oliveira**[1,5], **Remco Loos**[6], **Aliaksandra Radzisheuskaya**[7,8], **Kerri A. Mowen**[9], **Paul Bertone**[6,7,10], **José Silva**[7,8], **Magdalena Zernicka-Goetz**[1], **Michael L. Nielsen**[11], **John Gurdon**[1,4], and **Tony Kouzarides**[1,2]

[1]The Gurdon Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, United Kingdom

[2]Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, United Kingdom

[3]Laboratory of Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-17177 Stockholm, Sweden

[4]Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK

[5]EMBRAPA Dairy Cattle Research Center, Juiz de Fora, Brazil

[6]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SD, United Kingdom

[7]Wellcome Trust - Medical Research Council Stem Cell Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, United Kingdom

[8]Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, United Kingdom

[9]Department of Chemical Physiology, The Scripps Research Institute, La Jolla, CA, USA

[10]Genome Biology and Developmental Biology Units, European Molecular Biology Laboratory, Meyerhofstraβe 1, 69117 Heidelberg, Germany

[11]Department of proteomics, The Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Faculty of Health Sciences, Blegdamsvej 3b, DK-2200 Copenhagen, Denmark

Abstract

[*]These authors contributed equally to this work.

Author contributions:

M.A.C., G.C-B. and T.K. conceived the idea for this project, designed experiments and wrote the manuscript with the help of all the authors. G.C-B. and M.A.C performed ES cell transductions, established transgenic pre-iPS and ES cell lines and performed gene expression analyses. M.A.C carried out mutagenesis, protein isolation, biochemical and chromatin immunoprecipitation experiments, and performed citrullination analyses with the help of K.A.M.; G.C-B performed reprogramming experiments, with the help of J.S. and A.R. M.L.N. and M.A.C. performed Mass Spectrometric analyses. R.P.H-S and M.A.C. performed PADI4 treatments of permeabilized cells and subsequent chromatin compaction analyses, with the help of J.B.G. C.S.O and M.Z-G designed and performed mouse embryo experiments. R.L. and P.B. performed bioinformatic analyses of microarray data. T.K. supervised the study.

Europe PMC Funders Author Manuscripts

Citrullination is the post-translational conversion of an arginine residue within a protein to the non-coded amino acid citrulline[1]. This modification leads to the loss of a positive charge and reduction in hydrogen bonding ability. It is carried out by a small family of tissue-specific vertebrate enzymes called peptidylarginine deiminases (PADIs)[2] and is associated with the development of diverse pathological states such as autoimmunity, cancer, neurodegenerative disorders, prion diseases and thrombosis[2,3]. Nonetheless, the physiological functions of citrullination remain ill-defined, though citrullination of core histones has been linked to transcriptional regulation and the DNA damage response[4–8]. PADI4 (or PAD4/PADV), the only PADI with a nuclear localization signal[9], was previously shown to act in myeloid cells where it mediates profound chromatin decondensation during the innate immune response to infection[10]. Here we show that the expression and enzymatic activity of PADI4 are also induced under conditions of ground state pluripotency and during reprogramming. PADI4 is part of the pluripotency transcriptional network, binding to regulatory elements of key stem cell genes and activating their expression. Its inhibition lowers the percentage of pluripotent cells in the early mouse embryo and significantly reduces reprogramming efficiency. Using an unbiased proteomic approach we identify linker histone H1 variants, which are involved in the generation of compact chromatin[11], as novel PADI4 substrates. Citrullination of a single arginine residue within the DNA binding site of H1 results in its displacement from chromatin and global chromatin decondensation. Together, these results uncover a role for citrullination in the regulation of pluripotency and provide new mechanistic insights into how citrullination regulates chromatin compaction.

Pluripotent cells have the capacity to self-renew and differentiate into all somatic and germ cell lineages and, hence, possess immense therapeutic potential for a multitude of medical conditions. Their generation by reprogramming of differentiated somatic cells has been achieved by nuclear transfer, cell fusion and transduction of transcription factors, such as Oct4, Sox2, Klf4 and c-Myc[12]. Pluripotent cells have a distinctly open chromatin structure that is essential for unrestricted developmental potential[13,14] and reprogramming involves an almost complete epigenetic resetting of somatic cells[13]. The ability of PADI4-mediated histone citrullination to induce chromatin decondensation in neutrophils[10] prompted us to ask whether it can play a role in pluripotency, where chromatin decondensation is also necessary. To investigate this we first assessed the expression of *Padi4* in the embryonic stem cell line ES Oct4-GIP (ES), the neural stem cell line NSO4G (NS) and in induced pluripotent stem (iPS) cells derived NSO4G (see Methods). *Padi4* is expressed in pluripotent ES and iPS but not multipotent NS cells (Fig.1a). Culture of ES cells in 2i/LIF medium establishes a ground state of pluripotency[15]. This leads to the down-regulation of lineage specific markers and the up-regulation of pluripotency factors, as well as rapid induction of *Padi4* (Fig. 1b). The pattern of *Padi4* expression follows closely that of *Nanog,* an essential transcription factor for the transition to ground state pluripotency[16](Fig 1a,b). While other PADIs are expressed in pluripotent cells, PADI4 is the only one whose expression clearly associates with naïve pluripotency (Extended Data Fig. 1a,b). Citrullination of histone H3 (H3Cit), a modification shown previously to be carried out specifically by PADI4[17], is detectable in ES and iPS cells (Extended Data Fig. 1c), indicating that PADI4 is also enzymatically active. H3 and global citrullination are undetectable in NS cells (Extended Data Fig. 1c,d).

To determine the kinetics of PADI4 activation during the establishment of pluripotency, we examined RNA and protein samples collected daily during the course of reprogramming of NSO4G into iPS cells[16]. *Padi4* is rapidly induced in NS cells after transduction of reprogramming factors but only becomes active to citrullinate H3 after introduction of 2i/LIF, closely following the onset of *Nanog* expression (Fig. 1c). These observations strongly suggested that PADI4 activity is associated with ground state pluripotency and prompted us to examine whether *Padi4* is part of the pluripotency transcriptional network.

First, we asked whether the reprogramming factors regulate *Padi4* expression, utilizing the ZHBTc4.1 and 2TS22C cell lines where Oct4 and Sox2, respectively, can be deleted acutely in response to doxycycline treatment (see Methods). Deletion of Oct4, but not Sox2, led to a decrease in *Padi4* mRNA levels (Extended Data Fig. 1e). Furthermore, while Oct4 and Klf4 occupy the *Padi4* promoter in ES, but not NS cells, Sox2 is bound in both cell types (Extended Data Fig. 1f). To understand the effects of PADI4 on transcriptional regulation in pluripotent cells, we analyzed the transcriptome of ES cells upon PADI4 over-expression and inhibition. Several key pluripotency genes are up-regulated in response to PADI4 over-expression (Fig. 1d, Extended Data Figure 2a and Supplementary Table 2), including *Klf2, Tcl1, Tcfap2c*, and *Kit*. *Tcl1* was previously identified as the only regulator of self-renewal up-regulated in ground state pluripotency[15] and over-expression of *Tcl1* or *Tcfap2c* positively influence this process[18]. Gene Ontology (GO) analysis of this dataset indicates an enrichment of genes involved in stem cell development and maintenance (Fig. 1e). In addition, knockdown of *Padi4* in mES cells leads to decreased expression of *Tcl1* and *Nanog,* which is rescued by exogenous expression of RNAi-resistant human PADI4 (Fig. 1f, Extended Data Fig. 2b,c). These genes are under the control of PADI4 enzymatic activity since treatment with the chemical inhibitor Cl-amidine, which disrupts citrullination by PADI4[19], down-regulates their expression (Fig. 1g). Chromatin immunoprecipitation (ChIP) analysis indicated that H3Cit is present on regulatory regions of *Tcl1* and *Nanog* in ES and iPS cells, but not NS cells (Fig. 1h and Extended Data Fig. 3a). Accordingly, exogenously expressed human PADI4 localizes to and is enzymatically active on these regions, as well as regulatory regions of *Klf2* and *Kit* in ES cells (Extended Data Fig. 3b,c). In contrast to PADI4 over-expression, treatment of ES cells with Cl-amidine led to up-regulation of differentiation markers such as *Prickle1, EphA1* and *Wnt8a* and down-regulation of pluripotency markers such as *Klf5* (Extended Data Fig. 4a,b, and Supplementary Table 3), in addition to *Nanog* and *Tcl1* (Fig. 1g). GO analysis of this dataset indicated enrichment in genes involved in cell differentiation (Extended Data Fig. 4c). *Pou5f1 (Oct4)*, *Klf4*, *Sox2* and *c-Myc*, were not affected by PADI4 modulation (Extended Data Fig. 2a and Supplementary Tables 2 and 3). Cumulatively, the above results place PADI4 within the pluripotency transcriptional network, suggesting that it acts downstream of some of the cardinal reprogramming factors to regulate a specific subset of pluripotency genes.

Prompted by the above observations, we investigated whether PADI4 is necessary for pluripotency, as assessed during reprogramming (Extended Data Fig. 5a) and in the early stages of embryo development. NSO4G cells express a GFP reporter under the control of the Oct4 regulatory sequences, which is activated upon acquisition of pluripotency[16], allowing us to trace reprogrammed cells. Knock-down of *Padi4* in NSO4G-derived pre-iPS cells impaired the ability of the cells to establish H3Cit upon switch to 2i/LIF medium and led to

a clear reduction in reprogramming (Fig. 2a,b, Extended Data Fig. 5b-e and Extended Data time-lapse video). Consistent with this, levels of *Tcl1* and *Nanog* were not elevated upon reprogramming to the same extent as in control cells (Fig. 2b and Extended Data Fig. 5f). Cl-amidine treatment led to a dramatic reduction of reprogramming efficiency and H3Cit (Fig. 2c and Extended Data Fig. 5g-i), suggesting that the catalytic activity of PADI4 is important for the induction of pluripotency.

Padi4 expression and H3Cit are detected in the early embryo[20,21] and *Padi4-null* mice are born in lower numbers than would be expected by Mendelian inheritance[22], suggesting that PADI4 loss affects embryonic development. To assess the role of PADI4 in early development, we cultured mouse embryos in Cl-amidine-containing medium from the 2-cell stage and throughout pre-implantation development (see Methods and Extended Data Fig. 6a-c). Using 200μM Cl-amidine resulted in a complete developmental arrest of the embryos at the 8-cell stage (Extended Data Fig. 6a). We therefore used the maximum dose of Cl-amidine that reduced H3Cit (Extended Data Fig. S6b,c) but did not induce arrest (10μM). This led to a reduced percentage of pluripotent Nanog-positive epiblast cells and an increased percentage of differentiated trophectoderm cells at the blastocyst stage (Fig. 2d,e and Extended Data Fig. 6d,e). Time-course analyses of the cleavage patterns and cell fate decisions in early embryos showed that Cl-amidine increased the number of symmetric cell divisions at the expense of asymmetric divisions at the 8 to 16- and 16 to 32-cell transitions (Fig. 2f,g). This resulted in 16-cell stage embryos with fewer inner cells (destined for pluripotency) and greater numbers of outer cells (destined for differentiation into trophectoderm, reviewed in [23]) (Extended Data Fig. 6f). Treatment with another PADI4 inhibitor, TDFA[24], but not the HDAC inhibitor TSA, had similar effects (Extended Data Fig. 7, 8). These results indicate that PADI4 activity also promotes the maintenance of pluripotent cells in the early mouse embryo.

To elucidate the molecular mechanisms by which PADI4 regulates pluripotency, we aimed to identify PADI4 substrates in the chromatin fraction of mES cells using Stable Isotope Labeling of Amino acids in Cell culture (SILAC) (Fig. 3a and Extended Data Fig. 9a,b). Among the identified PADI4 substrates were Atrx, Dnmt3b, Trim28 and variants of linker histone H1 (Fig. 3b-e, Extended Data Fig. S9c-f, S10 and Supplementary Table 4), all of which can impact pluripotency. Histone H1 stabilizes the nucleosome and facilitates chromatin condensation, a state that is less permissive to processes that require access to the DNA, such as transcription[11]. The identified citrullinated H1 peptides correspond to, and are common between, variants H1.2, H1.3 and H1.4 (Fig. 3d,e), while an additional peptide corresponds to the same residue in H1.5 (Extended Data Fig. 10a,b). In ES cells, H1.2, H1.3 and H1.4 are required for chromatin compaction[25], while their depletion leads to increased expression of pluripotency genes such as *Nanog* and stalls them in a self-renewal state with impaired differentiation capability[26]. Notably, H1 is more loosely bound to chromatin in ES cells than in differentiated cells[14] and its genomic localization in cancer cells was shown to anti-correlate with that of PADI4[7].

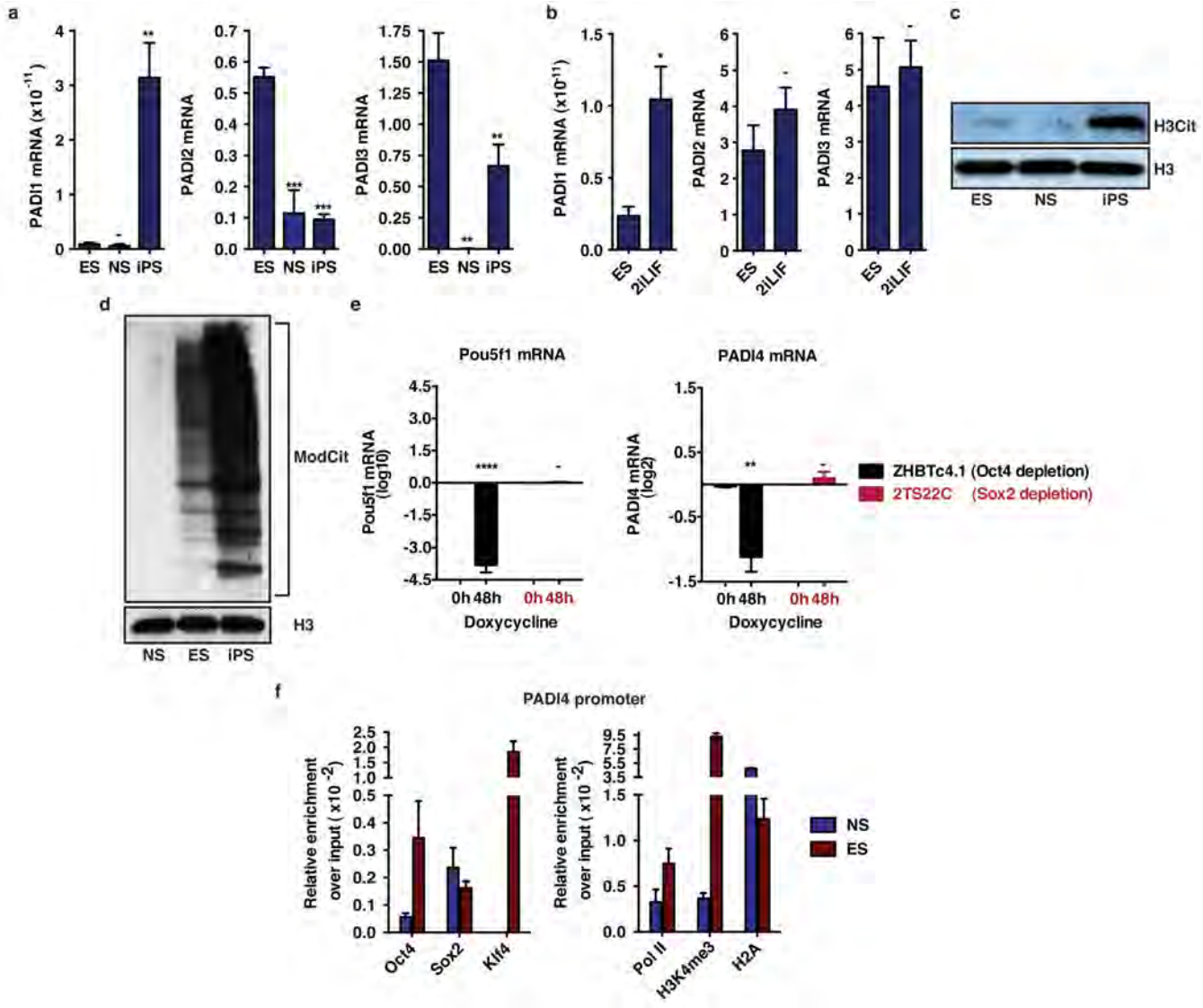
Mass spectrometric analysis accounted for all arginine residues within H1 but indicated that Arg54 (H1R54) is the only site citrullinated by PADI4 (Fig. 3e and Extended Data Fig. 9c). Indeed, we found that while H1.2 is citrullinated in ES cells (already by endogenous PADI4,

and significantly increased upon PADI4 over-expression), it is refractory to modification when Arg54 is mutated (Fig. 3f). Similar results were obtained in *in vitro* citrullination assays (Extended Data Fig. 11a). H1R54 lies within the globular domain of H1 (Extended Data Fig. 11b), which is highly conserved among the linker histone family and is necessary for interaction with nucleosomal DNA[27,28]. To test whether H1R54 is necessary for binding of H1 to nucleosomes, we mutated and assessed it in nucleosome-binding assays. Figure 3g shows that an R54A mutant, which mimics the charge change that accompanies citrullination, is impaired for nucleosome binding. An R54K mutant, which retains the positive charge, is impaired to a lesser extend (Fig. 3g) suggesting that H1R54 is important for electrostatic interactions between H1.2 and the nucleosome.

The above results open up the possibility that PADI4 may affect chromatin compaction in pluripotent cells. To test this hypothesis, we first assessed whether citrullination by ectopic PADI4 can lead to decondensation of differentiated cell chromatin. Recombinant PADI4 was added to permeabilized and stabilized differentiated C2C12 mouse myoblast nuclei (Fig. 4a). This protocol ensures stabilization of the nuclear component while allowing the free diffusion of non-chromatin bound nuclear proteins into the extra-nuclear fraction, and their collection by washing. Incubation with active PADI4 (Extended Data Fig. 12a,b), leads to the eviction of H1 from the chromatin and its diffusion out of the permeabilized nucleus (Fig. 4b). The evicted H1 is citrullinated on R54, as determined by mass spectrometry (Extended Data Figure 12c,d). Consistent with this, PADI4-treated cells showed evidence of decondensed chromatin, as determined by nuclear swelling, diffuse DAPI staining and increased sensitivity to micrococcal nuclease (Fig. 4c,d and Extended Data Figure 12e). Similar results were observed when *PADI4* was over-expressed in C2C12 cells (Fig. 4e) or NS cells (data not shown). To monitor if PADI4 can affect H1 binding on pluripotent cell chromatin, we performed ChIP-qPCR analyses of H1.2 on the regulatory regions of *Tcl1* and *Nanog* and found that it is stabilized upon *Padi4* knockdown (Fig. 4f). The ability of PADI4 to disrupt the binding of H1 to nucleosomal DNA provides a novel mechanistic example of how citrullination regulates protein function and chromatin condensation.

The work described above identifies citrullination of chromatin components by PADI4 as a feature of pluripotency (Fig. 4g), in addition to its previously described role in the myeloid lineage. One of the reasons for the restricted expression pattern of PADI4 may be the requirement for an open chromatin state in these cell types. The selective expression characteristics and the inducible nature of the catalytic activity of PADI4, suggest that it is under tight spatial and temporal regulation, giving it a unique status among chromatin modifying enzymes. As such, inappropriate PADI4 activity may have deleterious consequences, which may explain its activation in cancers of varying origin during progression to malignancy[29]. Indeed, citrullination is a common feature of several unrelated diseases, suggesting that strict regulation is likely to be a requirement for the physiological function of all PADIs. During review of this manuscript, Coonrod and colleagues suggested that PADI2, thought to be mainly cytoplasmic, can also citrullinate histones and lead to transcriptional activation[30]. This opens the intriguing possibility that other PADIs may mediate nuclear events in specific contexts, including in pluripotent cells. Further research into the function and targets of PADIs is likely to shed light into the etiology of several pathologies.

Extended Data



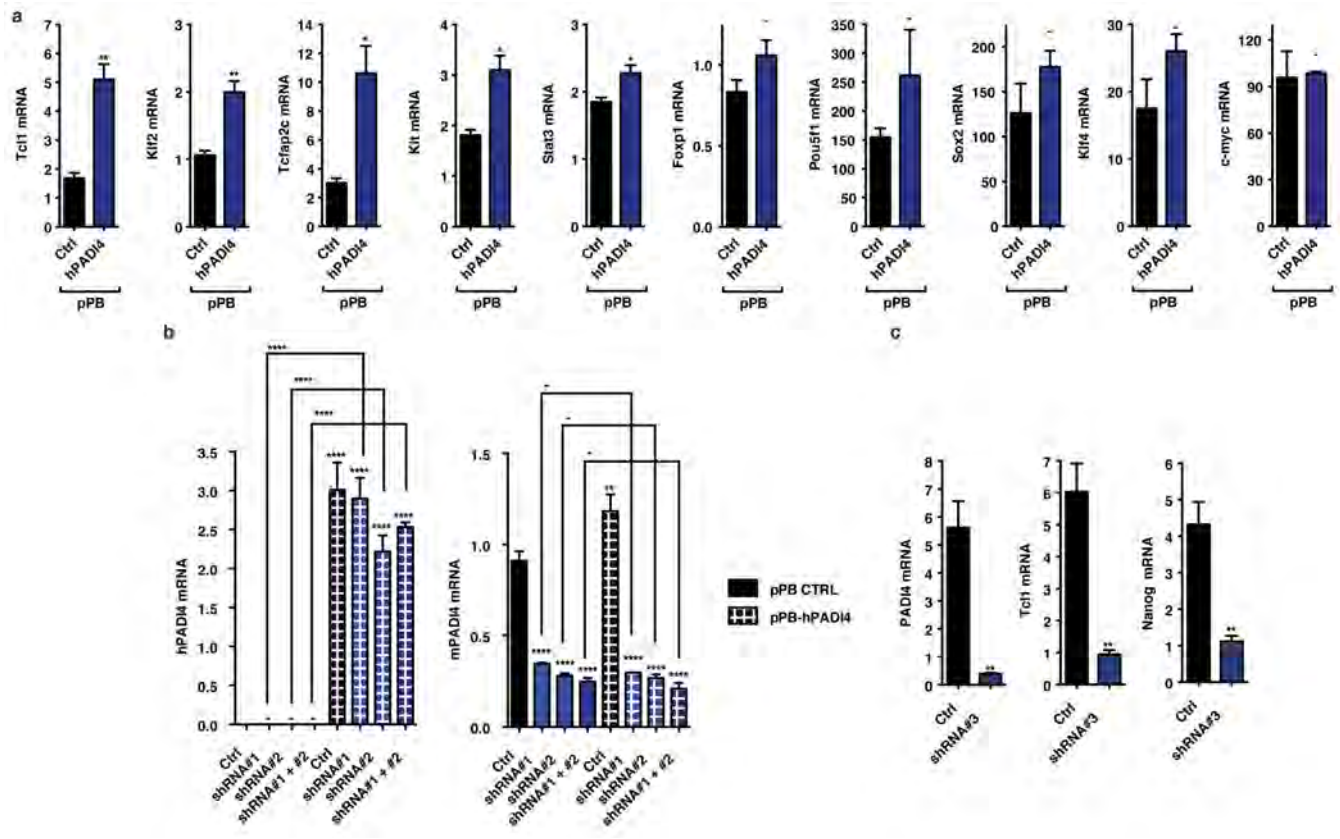
Extended Data Figure 1:

(a) Transcript levels for *Padi1, Padi2 and Padi3* in ES, NS and iPS cells, as assessed by qRT-PCR. *Padi6* was undetectable in all three cell types. Expression normalized to endogenous levels of *Ubiquitin (UbC)*. Error bars represent the standard error of the mean of three biological replicates.

(b) Transcript levels of *Padi1, Padi2 and Padi3* in ES cells upon switch to 2i containing medium for one passage, as assessed by qRT-PCR. *Padi6* was undetectable in both conditions. Expression normalized to *UbC*. Error bars represent the standard error of the mean of three biological replicates.

(c) Immunoblot analysis of H3Cit levels in ES, NS and iPS cells. Total H3 is presented as loading control.

(d) Immunoblot analysis of total citrullination the levels in ES, NS and iPS cells, using an antibody against Modified Citrulline (ModCit) which recognizes peptidylcitrulline irrespective of amino acid sequence. Total H3 is presented as loading control.
(e) ZHBTc4.1 and 2TS22C ES cell lines were treated with 1μg/ml doxycycline for 48 hours, resulting in depletion of Oct4 or Sox2 (data not shown). *Padi4* mRNA was significantly reduced upon Oct4, but not Sox2 knockdown, as assessed by qPCR. Error bars represent standard error of the mean of four biological replicates.
(f) ChIP-qPCR for Oct4, Sox2, Klf4, RNA polymerase II (PolII), H3K4me3 and H2A on the promoter of *Padi4* in mES and NS cells. For each cell condition, the signal is presented as fold enrichment over Input and after subtracting background signal from the beads. Error bars represent the standard deviation of three technical qPCR replicates.
Asterisks denote difference with ES cells (a) or media (b), and 0h time point (e); - not significant, * P≤0.05, ** P≤0.01, *** P≤0.001, **** P≤0.0001 by ANOVA (a) or t-test (b,e).

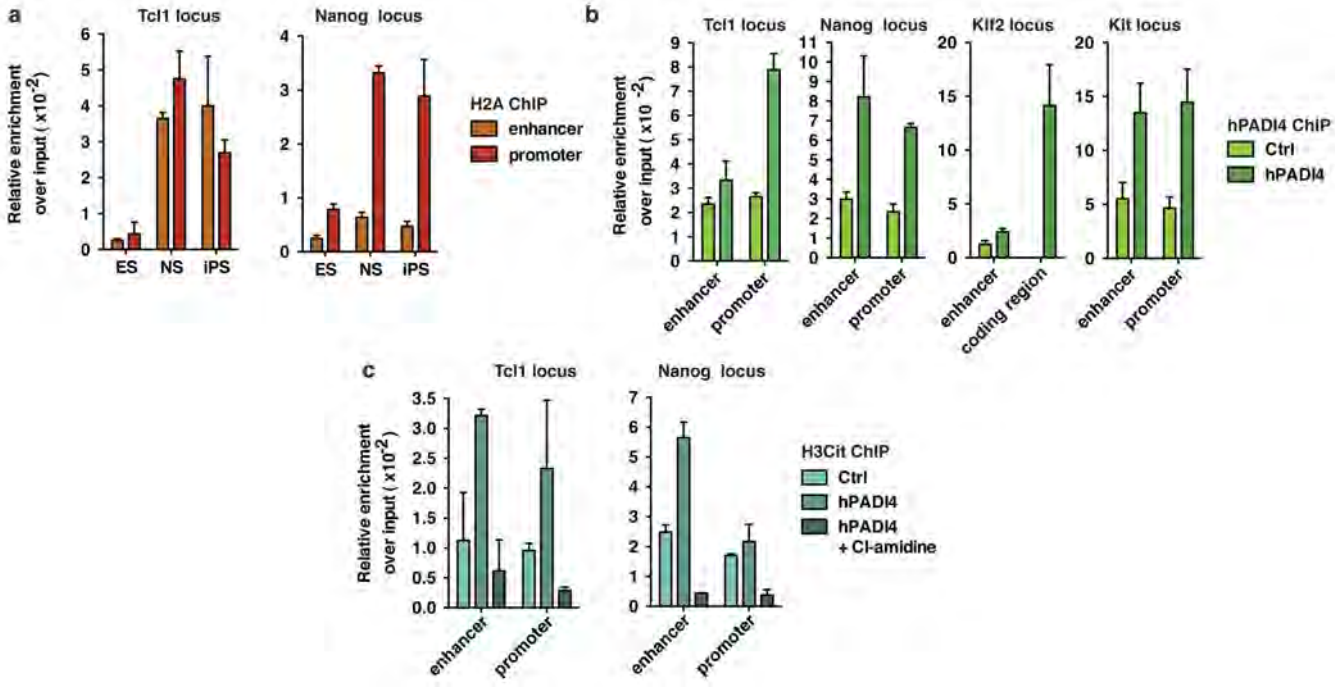


Extended Data Figure 2:
(a) Validation of selected targets from the PADI4 over-expression microarray dataset by qRT-PCR. Expression of *Pou5f1*, *Sox2*, *Klf4* and *c-Myc* is not affected by PADI4 over-expression. Expression normalised to *UbC*. Error bars presented as standard error of the mean of three biological replicates.
(b) Transcript levels of mouse *Padi4* and human *PADI4* in mES cells after transient knock-down with *Padi4* or control (Ctrl) shRNA, and over-expression of human *PADI4* or control

vector (pPB CTRL), as assessed by qRT-PCR. Expression normalized to *UbC*. Error bars represent the standard error of the mean of three biological replicates.

(c) Transcript levels of mouse *Padi4, Tcl1* and *Nanog* in mES cell clones stably expressing *Padi4* or control (Ctrl) shRNA, as assessed by qRT-PCR. Expression normalized to *UbC*. Error bars represent the standard error of the mean of three biological replicates.

Asterisks denote difference with Ctrl (a,b,c) and between samples (b); - not significant, * P≤0.05, ** P≤0.01, *** P≤0.001, **** P≤0.0001 by ANOVA (b) or t-test (a,c).



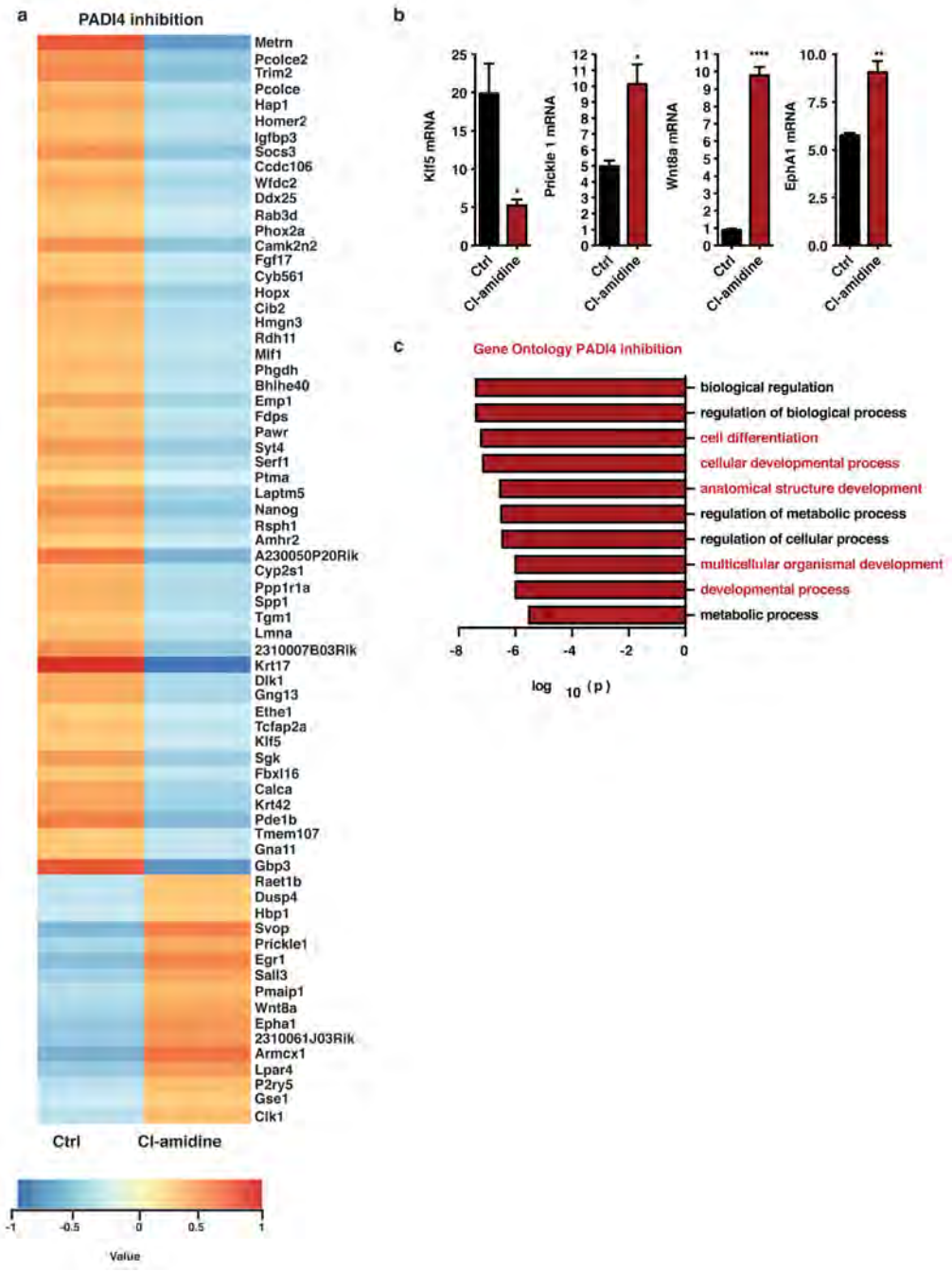
Extended Data Figure 3:

(a) Representative ChIP-qPCR for H2A on regulatory regions of *Tcl1* and *Nanog* in mES, NS and iPS cells (corresponding to Fig. 1h). For each cell condition, the signal is presented as fold enrichment over Input and after subtracting background signal from the beads. Error bars represent the standard deviation of three technical qPCR replicates.

(b) ChIP-qPCR for hPADI4 on regulatory regions of *Tcl1, Nanog, Klf2 and Kit*, which are up-regulated by hPADI4 over-expression, in mES cells stably expressing hPADI4. For each cell condition, the signal is presented as fold enrichment over Input and after subtracting background signal from the beads. Error bars represent the standard deviation of three technical qPCR replicates.

(c) Representative ChIP-qPCR for H3Cit on regulatory regions of *Tcl1* and *Nanog* in mES cells stably expressing hPADI4 and treated with 200μM Cl-amidine for 48h. For each cell condition, the signal is presented as fold enrichment over Input and after subtracting background signal from the beads. Error bars represent the standard deviation of three technical qPCR replicates.

Europe PMC Funders Author Manuscripts



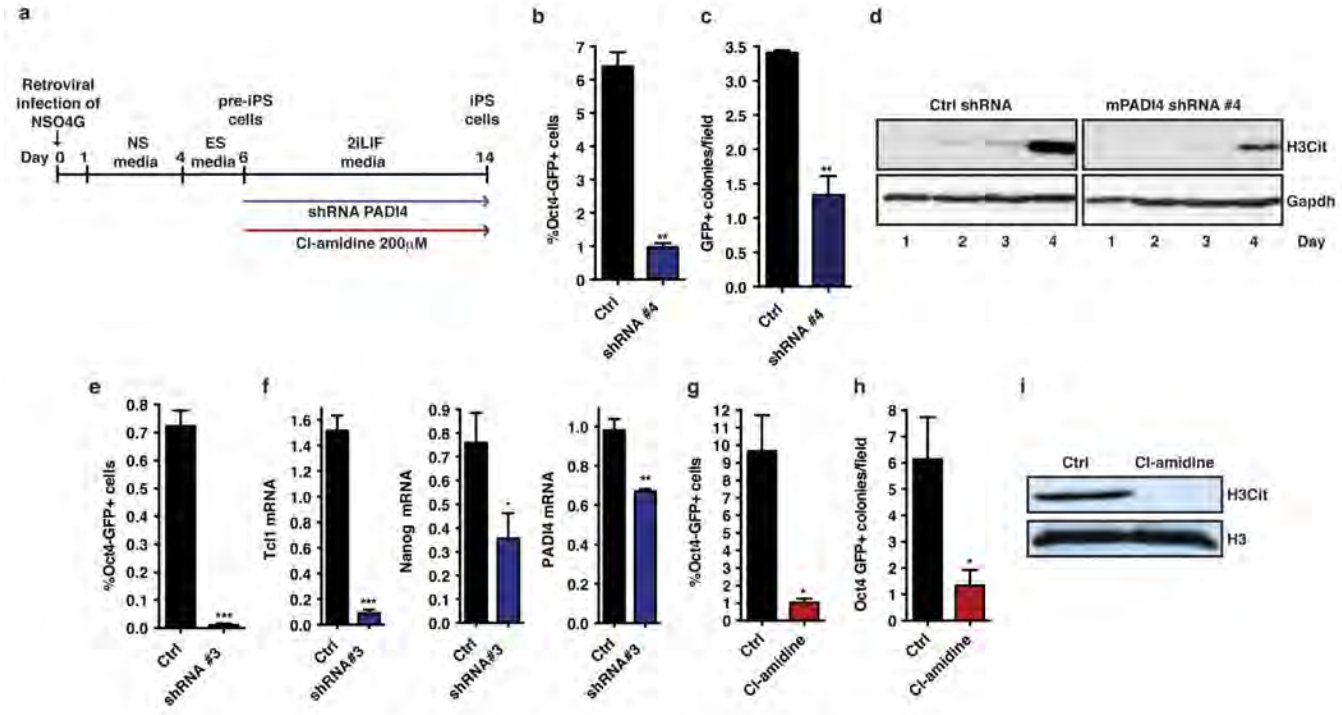
Extended Data Figure 4:

(a) Heat map of the top 70 genes that showed differential expression after PADI4 inhibition in ES cells by with 200μM Cl-amidine for 48h, as determined by microarray analysis. Displayed values are normalized log intensities, minus the mean expression of the gene across the two samples. Hierarchical clustering based on correlation.

(b) Validation of selected targets from the above microarray dataset by qRT-PCR. Expression normalised to *UbC*. Error bars presented as standard error of the mean of three

biological replicates. Asterisks denote difference with Ctrl; - not significant, * P≤0.05, ** P≤0.01, *** P≤0.001, **** P≤0.0001 by t-test.

(c) Gene Ontology for Biological Process (GOBP) analysis for the most regulated gene categories within the microarray dataset of Cl-amidine treatment in mES cells. *P*-value is corrected for multiple testing using Benjamini and Hochberg FDR.



Extended Data Figure 5:

(a) Scheme of reprogramming of neural stem cells to pluripotent state. NSO4G cells were retrovirally transduced with Oct4, Klf4 and c-Myc. After 6 days, partially reprogrammed pre-iPS cells arose. For shRNA experiments, pre-iPS cells were stably transfected with Ctrl or PADI4 shRNA and then full reprogramming was performed in the presence of 2iLIF media for 8 days. For PADI4 enzymatic inhibition, pre-iPS cells were immediately changed to 2iLIF media in the presence of the inhibitor Cl-amidine for 8 days.

(b) Quantification of flow cytometry analysis for the assessment of Oct4-GFP reporter expression in a reprogramming assay using pre-iPS cells stably expressing *Padi4* shRNA #4 and Ctrl shRNA. Error bars represent standard error of the mean of triplicate samples within a representative from four reprogramming experiments.

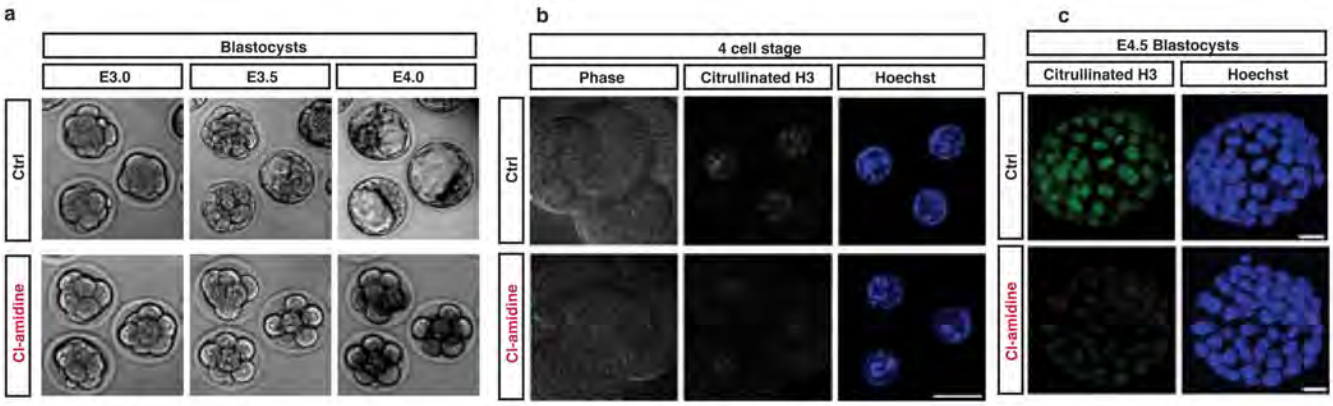
(c) Quantification of Oct4-GFP positive colonies in the reprogramming assay where pre-iPS cells were *Padi4* shRNA #4 versus control (*see* Fig. 2a), after time-lapse image acquisition with Biostation CT. Error bars represent standard error of the mean of triplicate samples within a representative reprogramming experiment. Time-lapse video in supplementary data online.

(d) Immunoblot analysis of H3Cit in pre-iPS cells treated with 2i, after *Padi4* knock-down (PADI4 shRNA #4) versus control cells (Ctrl shRNA). 2i-containing medium was added on day 2. Total histone H3 presented as loading control.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

(e) Quantification of flow cytometry analysis for the assessment of Oct4-GFP reporter expression in a reprogramming assay using pre-iPS cells stably expressing *Padi4* shRNA #3 and Ctrl shRNA. Error bars represent standard error of the mean of triplicate samples.
(f) qRT-PCR analysis for the expression of *Tcl, Nanog* and *Padi4* mRNAs at the end of the above reprogramming assay (e). Error bars represent standard error of the mean of triplicate samples.
(g) Quantification of flow cytometry analysis for the assessment of Oct4-GFP reporter expression in a reprogramming assay were treated with 200μM Cl-amidine. Error bars represent standard error of the mean of triplicate samples within a representative from three reprogramming experiment.
(h) Quantification of Oct4-GFP positive colonies in the reprogramming assay where pre-iPS cells were treated with 200μM Cl-amidine (*see* Fig. 2c) after time-lapse image acquisition with Biostation CT. Error bars represent standard error of the mean of triplicate samples within a representative reprogramming experiment.
(i) Immunoblot analysis for the presence of H3Cit at the end of the above reprogramming assay (g). Total histone H3 presented as loading control.
Asterisks denote difference with Control; - not significant, * P≤0.05, ** P≤0.01, *** P≤0.001, **** P≤0.0001 by t-test



d

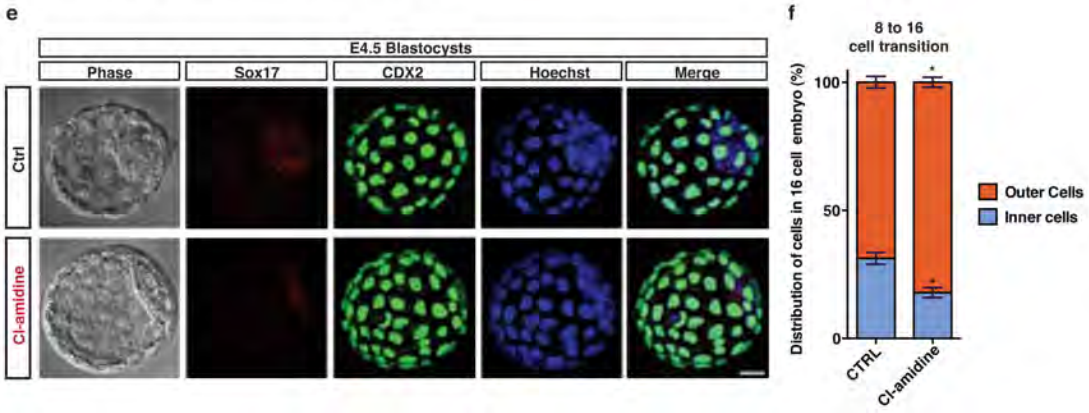
Lineage commitment in E4.5 blastocysts developed in the presence of Cl-amidine

	Number of EPI cells (mean±SE)	Percentage of EPI cells (mean±SE)	Number of TE cells (mean±SE)	Percentage of TE cells (mean±SE)	Number of PE cells (mean±SE)	Percentage of PE cells (mean±SE)	Total (mean±SE)
Control	10,00±0,73	12,37±0,91	60,15±3,23	77,10±1,43	6,61±0,76	7,60±0,78	77,84±3,94
Cl-amidine	10,26±0,63	9,48±0,40*	82,29±4,30*	80,79±0,87*	8±0,43	7,59±0,44	101,33±4,69*

Data from 3 replicates and 50 E4.5 blastocysts (n = 24-26 per group)

*Asterisks indicate difference within the same column (Unpaired T-test, P<0.05)

EPI: epiblast, TE: trophectoderm, PE: primitive endoderm



Extended Data Figure 6:

(a) Embryos at 2-cell stage were treated with 200μM Cl-amidine and snapshots were taken at E3.0, E3.5 and E4.0. 200μM Cl-amidine embryos arrested at 8-cell stage, while Control embryos continued development to form blastocysts. Phase contrast images are shown.

(b) Embryos at 2-cell stage were treated with 10μM Cl-amidine for 12 hours, fixed and stained for H3Cit at the 4-cell stage. Phase contrast, H3Cit (white) and HOECHST 33342 (blue) images are shown. Bar represents 20μm.

(c) Embryos at E3.5 were treated with 10μM Cl-amidine for 24 hours, fixed and stained for H3Cit at E4.5. H3Cit (green) and HOECHST 33342 (blue) images are shown. Bar represents 20μm.

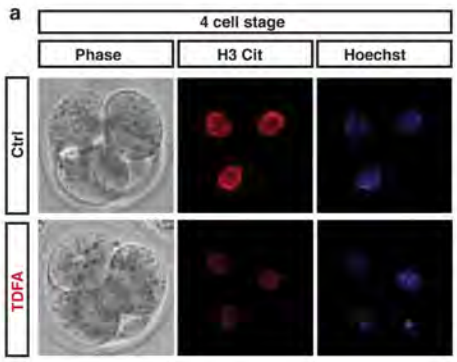
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

(d) Table with quantifications of lineage commitment in E4.5 blastocysts treated with 10μM Cl-amidine from the 2-cell stage. Asterisks denote difference with Control, unpaired t-test, * = $p<0.05$. n=3 (50 embryos).

(e) Embryos were cultured in medium supplemented with 10μM Cl-amidine from 2-cell stage and through preimplantation development. E4.5 blastocysts were fixed and stained for SOX17 (primitive endoderm marker, red), Cdx2 (trophectoderm marker, green) and HOECHST 33342 (blue). Bar represents 20μm.

(f) Time-lapse analysis of distribution of inner and outer cells at the 8 to 16-cell transition, upon culturing of embryos with medium containing 10μM Cl-amidine from 2-cell stage. Error bars represent standard error of the mean. Statistical significance was determined by unpaired t-test or Mann Whitney test upon non-normal distribution. Asterisks denote difference with Control; * $P\leq0.05$.



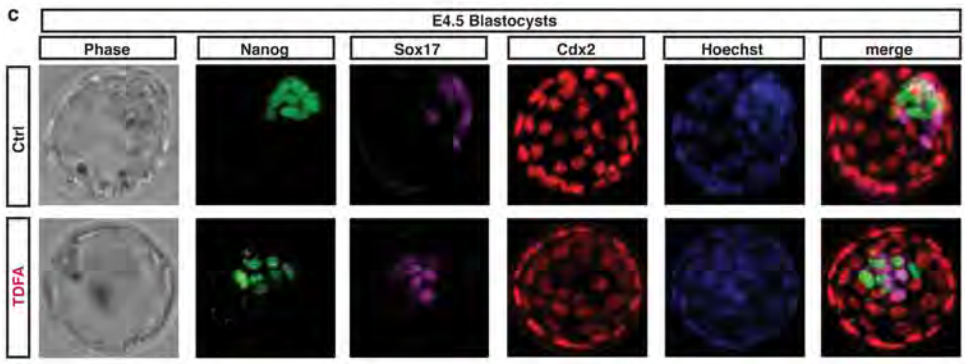
b Lineage commitment in E4.5 blastocysts developed in the presence of TDFA

	Percentage of EPI cells (mean±SE)	Percentage of TE cells (mean±SE)	Percentage of PE cells (mean±SE)
Control	12,60±0,79	77,02±0,94	10,38±0,60
TDFA	10,43±0,62*	78,70±0,68	10,87±0,62

Data from 3 replicates and 60 E4.5 embryos (n=28-32 per group)

* Asterisks indicate difference within the same column (Mann Whitney test, P<0.05)

EPI: epiblast, TE: trophoectoderm , PE: primitive endoderm

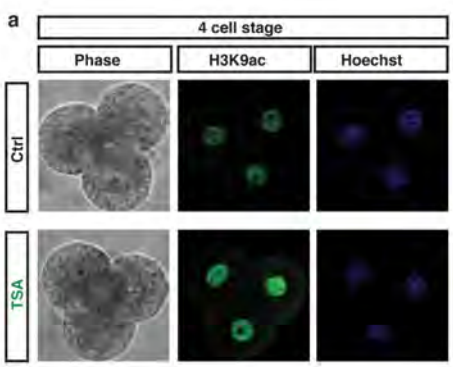


Extended Data Figure 7:

(a) Embryos at 2-cell stage were treated with 100μM TDFA for 12 hours and fixed and stained for H3Cit at 4-cell stage. H3Cit and HOECHST 33342 images are shown.

(b) Table representing the percentage of cells committed to each embryonic lineage in E4.5 blastocysts upon treatment of embryos at 2-cell stage with 100μM TDFA. Bars represent mean percentage (±SEM). Asterisks denote difference with Control, Mann-Whitney test, * = $p<0.05$. n=3 (60 embryos).

(c) Embryos at 2-cell stage were treated with 100μM TDFA and fixed at embryonic day E4.5. Phase contrast, Nanog (green), Sox17 (purple), Cdx2 (red) and HOECHST 33342 (blue) images are shown.



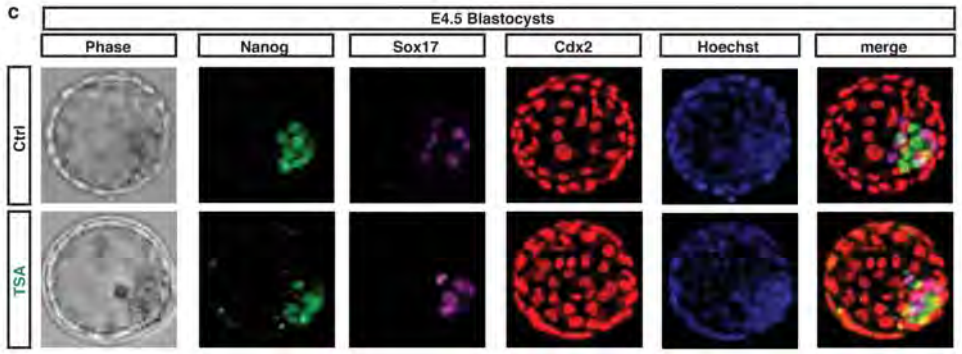
b Lineage commitment in E4.5 blastocysts developed in the presence of TSA

	Percentage of EPI cells (mean±SE)	Percentage of TE cells (mean±SE)	Percentage of PE cells (mean±SE)
Control	10,06±0,67	77,29±0,86	12,64±0,61
TSA	10,71±0,80	79,34±0,91	9,95±0,60*

Data from 2 replicates and 32 E4.5 embryos (n=15-17 per group)

* Asterisks indicate difference within the same column (unpaired T-test, P<0.05)

EPI: epiblast, TE: trophoectoderm , PE: primitive endoderm



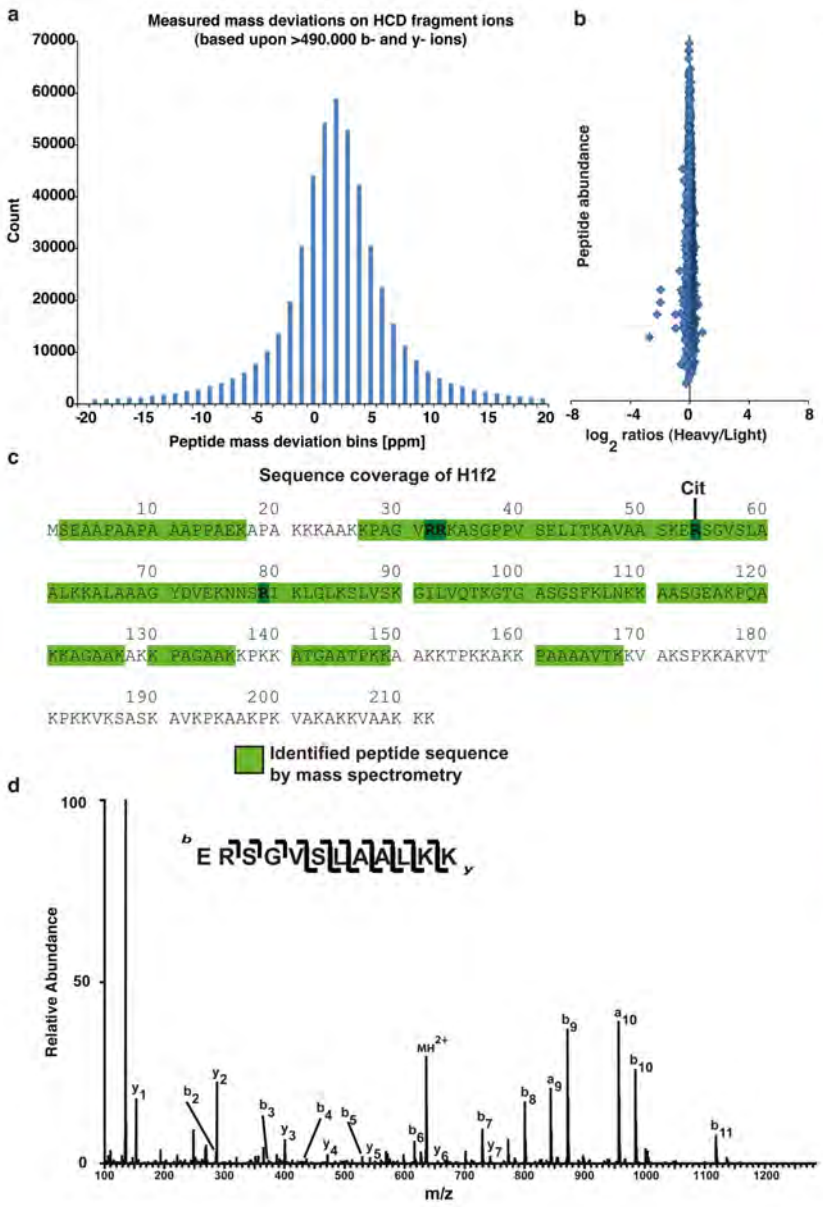
Extended Data Figure 8:

(a) Embryos at 2-cell stage were treated with 10nM TSA for 12 hours, and fixed and stained for H3K9ac at 4-cell stage. H3K9ac and HOECHST 33342 images are shown.

(b) Table representing the percentage of cells committed to each embryonic lineage in E4.5 blastocysts upon treatment of embryos at 2-cell stage with 10nM TSA. Bars represent mean percentage (±SEM). Asterisks denote difference with Control, unpaired t-test, * = $p<0.05$. n=2 (32 embryos).

(c) Embryos at 2-cell stage were treated with 10nM TSA and fixed at embryonic day E4.5. Phase contrast, Nanog (green), Sox17 (purple), Cdx2 (red) and HOECHST 33342 (blue) images are shown.

Europe PMC Funders Author Manuscripts



e

b-ion	Theo.Mass	Mass	PPM
b2	287.1355	287.1343	4.18
b3	374.1676	374.1649	7.22
b4	431.1890	431.1849	9.51
y-ion	**Theo.Mass**	**Mass**	**PPM**
y1	153.1335	153.1325	6.53
y2	287.2486	287.2474	4.18
y3	400.3326	400.3315	2.75
y4	471.3698	471.3683	3.18
y5	542.4069	542.4055	2.58
y6	655.4909	655.4880	4.42
y7	742.5230	742.5204	3.50
y8	841.5914	841.5886	3.33
y9	898.6128	898.6099	3.23
y10	985.6449	985.6422	2.74

f

b-ion	Theo.Mass	Mass	PPM
b2	286.1515	286.1502	4.54
b3	373.1836	373.1808	7.50
b4	430.2050	430.2068	4.18
b5	529.2734	529.2703	5.86
b6	616.3055	616.3045	1.62
b7	729.3895	729.3867	3.84
b8	800.4266	800.4236	3.75
b9	871.4638	871.4616	2.52
b10	984.5478	984.5458	2.03
b11	1118.6629	1118.6654	2.23
y-ion	**Theo.Mass**	**Mass**	
y1	153.1335	153.1326	5.88
y2	287.2486	287.2474	4.18
y3	400.3326	400.3318	2.00
y4	471.3698	471.3679	4.03
y5	542.4069	542.4033	6.64
y6	655.4909	655.4884	3.81
y7	742.5230	742.5139	12.26

Extended Data Figure 9:

(a) Histogram demonstrating the mass accuracies of all fragment ion masses used for identifying citrullinated peptides in our HCD MS/MS spectra. >490.000 y- and b-ion masses are depicted. The average absolute mass accuracy for all of these fragment ions is 3.97 ppm. **(b)** Scatter plot representing SILAC ratios in ES cells cultured in $^{13}C_6$ L-Lysine (HEAVY) and LIGHT medium separately, to assess extend and quality of SILAC labeling. No significant outliers are observed, indicating equal labeling.

(c) Peptide coverage of histone H1 by LC-MS analysis. Detected peptides are highlighted in light green and cover >60% of H1. While all arginine residues of Histone H1 (highlighted in dark green) were accounted for by the analysis, Arg54 was the only one found citrullinated.

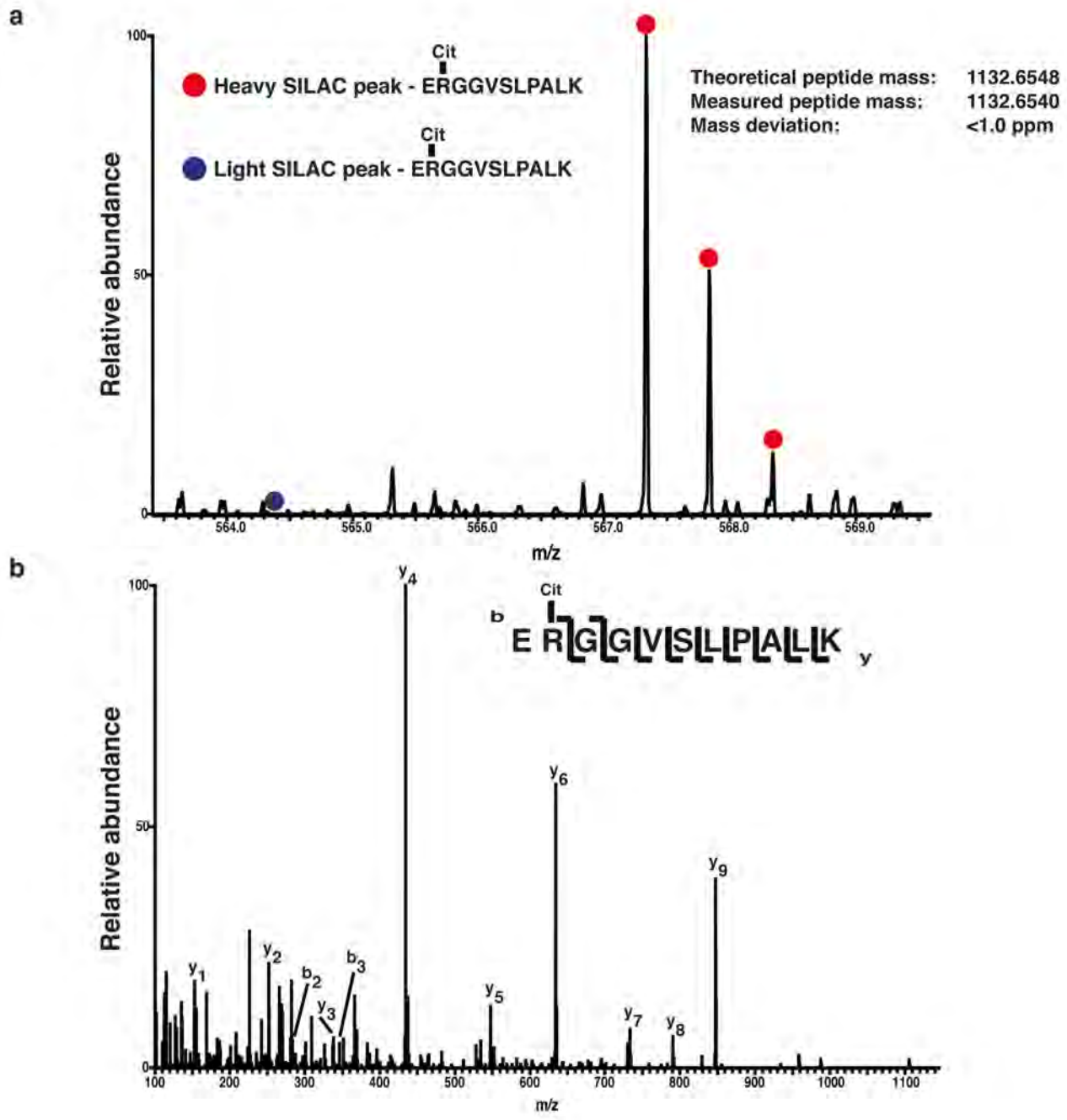
Europe PMC Funders Author Manuscripts

(d) Fragmentation spectra of the unmodified LysC peptide ERSGVSLAALKK surrounding Arginine 54 of H1.2 (unmodified counterpart of citrullinated peptide depicted in Fig. 3d). The y and b series indicate fragments at amide bonds of the peptide.

(e) Fragment ion table (expected and observed masses for detected y and b ions) for the identified H1R54 citrullination of peptide ERSGVSLAALKK on histone H1.2 (as shown in Fig. 3d). All measured fragment ions were detected with mass accuracies <10ppm, unambiguously identifying that the detected peptide sequence harbors a citrullination at position R54.

(f) Theoretical and measured b- and y-ion fragment masses for the corresponding unmodified and heavy SILAC labeled H1.2 peptide, as presented in (d) above.

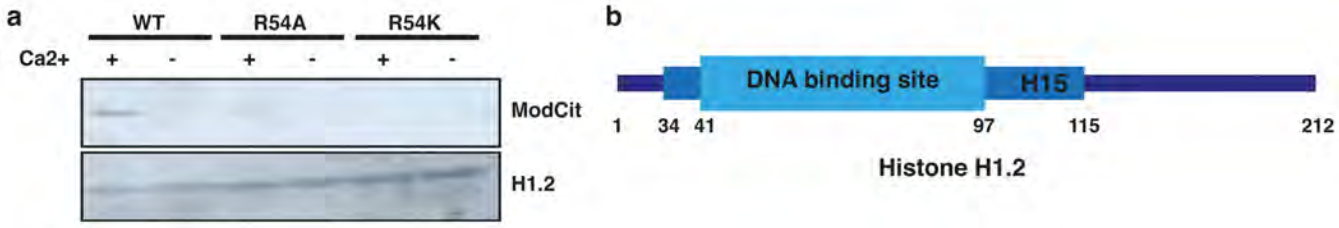
Europe PMC Funders Author Manuscripts



Extended Data Figure 10:

(a) MS spectrum of Histone H1.5 in a SILAC proteomic screen for identification of PADI4 substrates. Linker histone H1.5 is deiminated by PADI4, as identified by a highly increased SILAC ratio of the heavy labeled identified peptide (marked by a red dot).

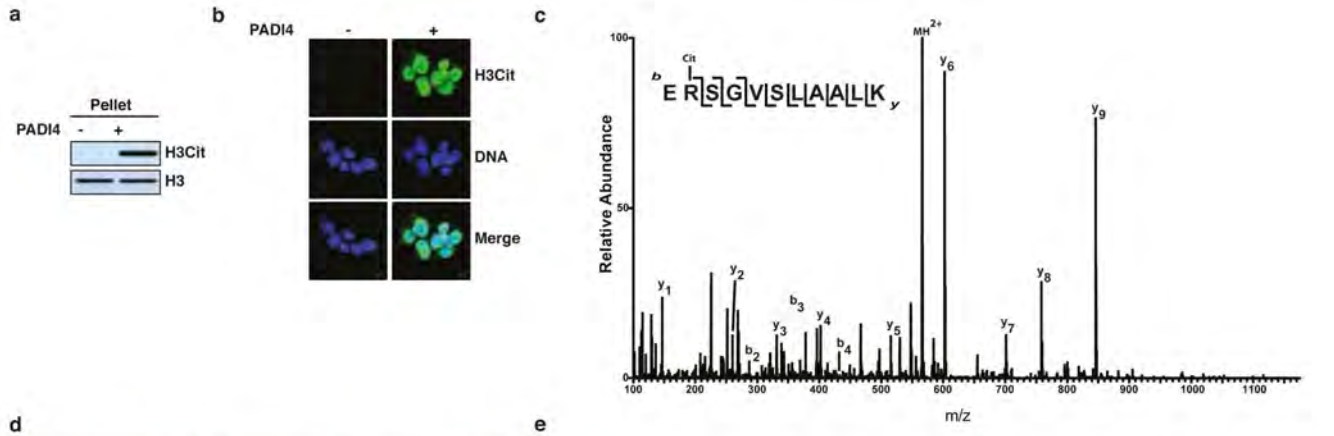
(b) Fragmentation spectra of the doubly charged LysC peptide ERGGVSLPALK surrounding Arginine 54 of H1.5. The y and b series indicate fragments at amide bonds of the peptide, unambiguously verifying the citrullinated peptide.



Extended Data Figure 11:

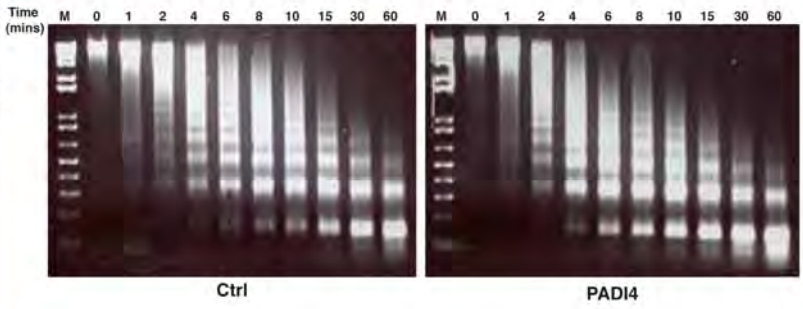
(a) Mutation of R54 renders histone H1.2 refractory to deimination. Immunoblot analysis of recombinant histone H1.2 using an antibody that detects all deimination events (ModCit). Wild type and R54-mutant H1.2 were treated with recombinant PADI4, in the presence of activating calcium. Only wild-type H1.2 can be deiminated, indicating that R54 is the only substrate of PADI4 in H1.2. No-calcium reactions presented as negative controls. Total H1.2 presented as loading control.

(b) Schematic representation of the position of R54 within the globular domain linker histone H1.2.



d

b-ion	Theo.Mass	Mass	PPM
b2	287.1355	287.1351	1.39
b3	374.1676	374.1667	2.41
b4	431.1890	431.1902	2.78
y-ion	**Theo.Mass**	**Mass**	**PPM**
y1	147.1134	147.1128	4.08
y2	260.1974	260.1969	1.92
y3	331.2345	331.2339	1.81
y4	402.2716	402.2711	1.24
y5	515.3557	515.3554	0.58
y6	602.3877	602.3871	1.00
y7	701.4562	701.4558	0.57
y8	758.4776	758.4770	0.79
y9	845.5096	845.5089	0.83



Extended Data Figure 12:

(a) Immunoblot analysis of the “Pellet” fraction of C2C12 permeabilised cells treated with recombinant PADI4. Presence of H3Cit species indicates PADI4 activity. Total H3 is presented as a control for equal use of starting material in the two experimental conditions.

Europe PMC Funders Author Manuscripts

(b) Immunofluorescence analysis of C2C12 nuclei after treatment with recombinant PADI4. Presence of H3Cit species indicates PADI4 activity. DNA is visualised by staining with DAPI.

(c) Fragmentation spectra of the citrullination site R54 on the evicted H1.2 peptide ERSGVSLAALK (corresponding to Fig. 4b). The evicted Histone H1 is citrulinated at R54.

(d) Theoretical and measured b- and y-ion fragment masses for the citrullinated H1.2 peptide (peptide sequence ERSGVSLAALK) evicted after treatment of C2C12 cells with recombinant hPADI4 (corresponding to Fig. 4b).

(e) Micrococcal nuclease digestion of C2C12 nuclei after treatment with recombinant PADI4, as described in Fig. 4a. M= size marker.

Supplemental Information

Refer to Web version on PubMed Central for supplementary material.

Acknlowledgements

This work was funded by programme grants from Cancer Research UK (T.K.) and EMBL (P.B., R.L.). R.P.H-S and JBG are supported by the Medical Research Council [G1001690] and the Wellcome Trust. G.C.-B. was funded by EMBO (Long-Term Post-Doctoral Fellowship), European Union (FP7 Marie Curie Intra-European Fellowship for Career Development) and Swedish Research Council. M.A.C was funded by an EMBO Long-Term Post-Doctoral Fellowship and a Human Frontier Science Programme Long-Term Post-Doctoral Fellowship. C.S.O was supported by FAPESP (Foundation for Research Support of the State of São Paulo) and mouse embryo work was supported by the Wellcome Trust programme grant to M.Z.G. M.L.N was partly supported by the Novo Nordisk Foundation Center for Protein Research, the Lundbeck Foundation, and by and the European Commission's 7th Framework Programme HEALTH-F7-2010-242129/SYBOSS. We would like to thank Sri Lestari, Alistair Cook and Cynthia Hill for technical assistance; Paul Thompson for kindly providing the TDFA compound; GSK Epinova for Cl-amidine; Till Bartke for the kind gift of histone octamers and help with nucleosome pull-down assays; Andrew Finch for help with FPLC chromatography; Agnieszka Jedrusik for help with embryo work; Rachael Walker at the Flow Cytometry Core Facility at Wellcome Trust Centre for Stem Cell Research, University of Cambridge and Thor Theunissen for help with the flow cytometry; and members of the Kouzarides laboratory for critical discussions of the work. 2TS22C cells were kindly provided by Dr. Hitoshi Niwa, at the RIKEN Center for Developmental Biology, Kobe, Japan. The ChIP grade H1.2 antibody was a generous gift from Prof. Arthur Skoultchi, Albert Einstein School of Medicine, New York, USA.

References

1. Vossenaar ER, Zendman AJ, van Venrooij WJ, Pruijn GJ. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. BioEssays: news and reviews in molecular, cellular and developmental biology. 2003; 25:1106–1118. DOI: 10.1002/bies.10357
2. Wang S, Wang Y. Peptidylarginine deiminases in citrullination, gene regulation, health and pathogenesis. Biochim Biophys Acta. 2013; 1829:1126–1135. DOI: 10.1016/j.bbagrm.2013.07.003 [PubMed: 23860259]
3. Martinod K, et al. Neutrophil histone modification by peptidylarginine deiminase 4 is critical for deep vein thrombosis in mice. Proc Natl Acad Sci U S A. 2013; 110:8674–8679. DOI: 10.1073/pnas.1301059110 [PubMed: 23650392]
4. Hagiwara T, Nakashima K, Hirano H, Senshu T, Yamada M. Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes. Biochemical and biophysical research communications. 2002; 290:979–983. DOI: 10.1006/bbrc.2001.6303 [PubMed: 11798170]
5. Cuthbert GL, et al. Histone deimination antagonizes arginine methylation. Cell. 2004; 118:545–553. DOI: 10.1016/j.cell.2004.08.020 [PubMed: 15339660]
6. Wang Y, et al. Human PAD4 regulates histone arginine methylation levels via demethylimination. Science. 2004; 306:279–283. DOI: 10.1126/science.1101400 [PubMed: 15345777]

7. Zhang X, et al. Genome-wide analysis reveals PADI4 cooperates with Elk-1 to activate c-Fos expression in breast cancer cells. PLoS genetics. 2011; 7:e1002112.doi: 10.1371/journal.pgen.1002112 [PubMed: 21655091]
8. Tanikawa C, et al. Regulation of histone modification and chromatin structure by the p53-PADI4 pathway. Nature communications. 2012; 3:676.doi: 10.1038/ncomms1676
9. Asaga H, Nakashima K, Senshu T, Ishigami A, Yamada M. Immunocytochemical localization of peptidylarginine deiminase in human eosinophils and neutrophils. Journal of leukocyte biology. 2001; 70:46–51. [PubMed: 11435484]
10. Neeli I, Khan SN, Radic M. Histone deimination as a response to inflammatory stimuli in neutrophils. J Immunol. 2008; 180:1895–1902. [PubMed: 18209087]
11. Buttinelli M, Panetta G, Rhodes D, Travers A. The role of histone H1 in chromatin condensation and transcriptional repression. Genetica. 1999; 106:117–124. [PubMed: 10710717]
12. Yamanaka S, Blau HM. Nuclear reprogramming to a pluripotent state by three approaches. Nature. 2010; 465:704–712. DOI: 10.1038/nature09229 [PubMed: 20535199]
13. Gaspar-Maia A, Alajem A, Meshorer E, Ramalho-Santos M. Open chromatin in pluripotency and reprogramming. Nature reviews. Molecular cell biology. 2011; 12:36–47. DOI: 10.1038/nrm3036 [PubMed: 21179060]
14. Meshorer E, et al. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev Cell. 2006; 10:105–116. DOI: 10.1016/j.devcel.2005.10.017 [PubMed: 16399082]
15. Marks H, et al. The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. Cell. 2012; 149:590–604. DOI: 10.1016/j.cell.2012.03.026 [PubMed: 22541430]
16. Theunissen TW, et al. Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. Curr Biol. 2011; 21:65–71. DOI: 10.1016/j.cub.2010.11.074 [PubMed: 21194951]
17. Darrah E, Rosen A, Giles JT, Andrade F. Peptidylarginine deiminase 2, 3 and 4 have distinct specificities against cellular substrates: novel insights into autoantigen selection in rheumatoid arthritis. Annals of the rheumatic diseases. 2012; 71:92–98. DOI: 10.1136/ard.2011.151712 [PubMed: 21859690]
18. Polo JM, et al. A molecular roadmap of reprogramming somatic cells into iPS cells. Cell. 2012; 151:1617–1632. DOI: 10.1016/j.cell.2012.11.039 [PubMed: 23260147]
19. Luo Y, et al. Inhibitors and inactivators of protein arginine deiminase 4: functional and structural characterization. Biochemistry. 2006; 45:11727–11736. DOI: 10.1021/bi061180d [PubMed: 17002273]
20. Brahmajosyula M, Miyake M. Localization and expression of peptidylarginine deiminase 4 (PAD4) in mammalian oocytes and preimplantation embryos. Zygote. 2011; :1–11. DOI: 10.1017/S0967199411000633
21. Kan R, et al. Potential role for PADI-mediated histone citrullination in preimplantation development. BMC developmental biology. 2012; 12:19.doi: 10.1186/1471-213X-12-19 [PubMed: 22712504]
22. Li P, et al. PAD4 is essential for antibacterial innate immunity mediated by neutrophil extracellular traps. J Exp Med. 2010; 207:1853–1862. DOI: 10.1084/jem.20100239 [PubMed: 20733033]
23. Zernicka-Goetz M, Morris SA, Bruce AW. Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. Nat Rev Genet. 2009; 10:467–477. DOI: 10.1038/nrg2564 [PubMed: 19536196]
24. Jones JE, et al. Synthesis and screening of a haloacetamidine containing library to identify PAD4 selective inhibitors. ACS chemical biology. 2012; 7:160–165. DOI: 10.1021/cb200258q [PubMed: 22004374]
25. Fan Y, et al. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell. 2005; 123:1199–1212. DOI: 10.1016/j.cell.2005.10.028 [PubMed: 16377562]
26. Zhang Y, et al. Histone h1 depletion impairs embryonic stem cell differentiation. PLoS genetics. 2012; 8:e1002691.doi: 10.1371/journal.pgen.1002691 [PubMed: 22589736]
27. Goytisolo FA, et al. Identification of two DNA-binding sites on the globular domain of histone H5. EMBO J. 1996; 15:3421–3429. [PubMed: 8670844]

28. Brown DT, Izard T, Misteli T. Mapping the interaction surface of linker histone H1(0) with the nucleosome of native chromatin in vivo. Nat Struct Mol Biol. 2006; 13:250–255. DOI: 10.1038/nsmb1050 [PubMed: 16462749]

29. Chang X, et al. Increased PADI4 expression in blood and tissues of patients with malignant tumors. BMC cancer. 2009; 9:40.doi: 10.1186/1471-2407-9-40 [PubMed: 19183436]

30. Zhang X, et al. Peptidylarginine deiminase 2-catalyzed histone H3 arginine 26 citrullination facilitates estrogen receptor alpha target gene activation. Proc Natl Acad Sci U S A. 2012; 109:13331–13336. DOI: 10.1073/pnas.1203280109 [PubMed: 22853951]

Europe PMC Funders Author Manuscripts

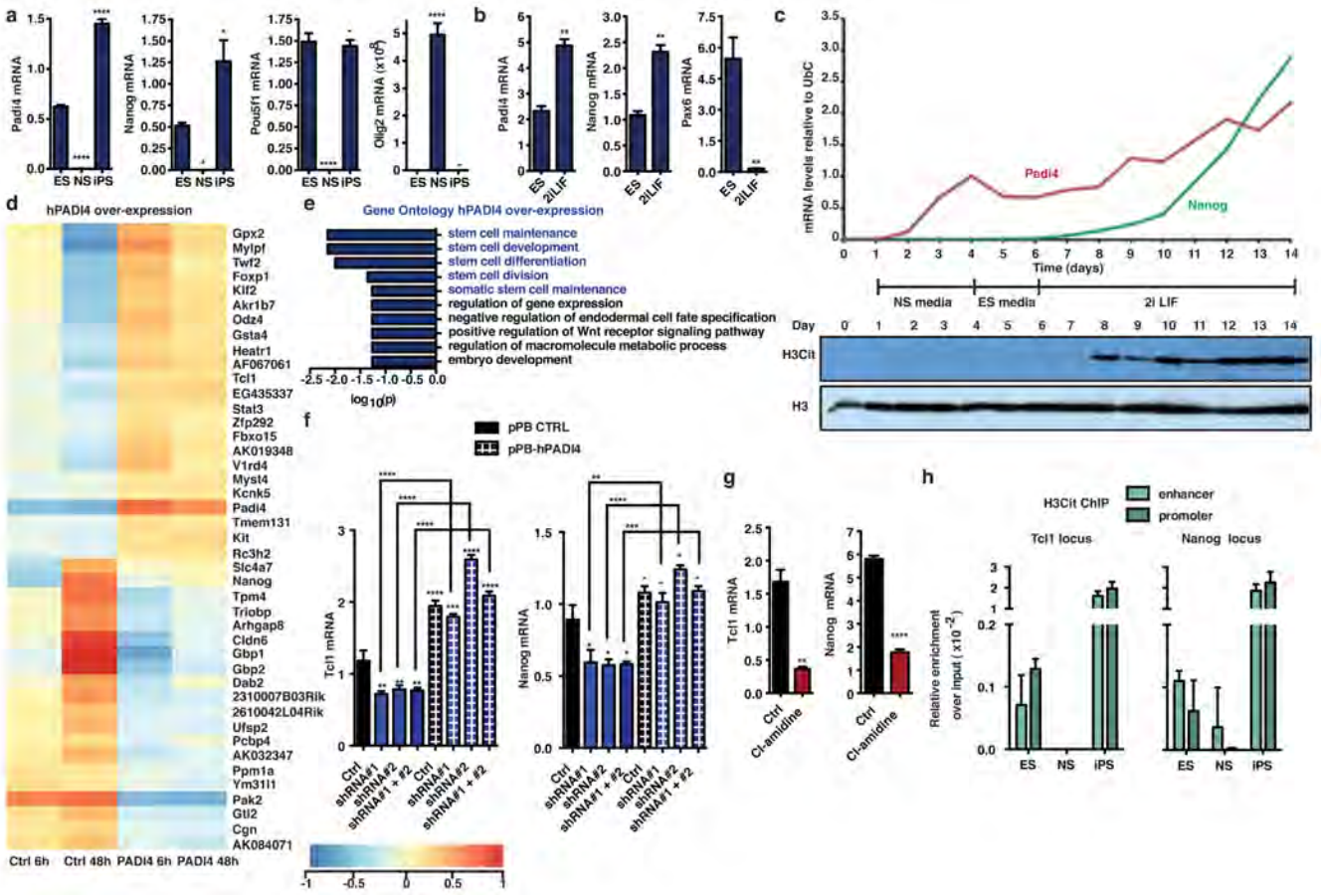


Figure 1: PADI4 expression and activity are features of pluripotent cells

(a,b) qRT-PCR for *Padi4* and *Nanog* expression in ES, NS and iPS cells **(a)**, and in ES cells upon culture in 2i/LIF for one passage **(b)**. *Pou5f1, Olig2* and *Pax6* are presented as controls. Expression normalized to *Ubiquitin (UbC)*. Error bars: standard error of the mean of three biological replicates.

(c) qRT-PCR for *Padi4* and *Nanog* expression and H3Cit immunoblot during the course of reprogramming (see also Extended Data Fig. 5a). Loading control: total histone H3. Representative of four experiments.

(d) Heat map of the genes regulated upon *hPADI4* over-expression in mES cells, as determined by microarray analysis. Displayed values are normalized log intensities, minus the mean expression of the gene across the four samples. Hierarchical clustering based on correlation.

(e) Gene Ontology for Biological Process (GOBP) analysis of the above microarray dataset. *P*- value is corrected for multiple testing using Benjamini and Hochberg False Discovery Rate (FDR).

(f,g) qRT-PCR for *Tcl1* and *Nanog* expression in mES cells after transient knock-down with *Padi4* or control (Ctrl) shRNAs, and over-expression of human *PADI4* or control vector (pPB CTRL) **(f)**, and after treatment with 200μM Cl-amidine **(g)**. Expression normalized to *UbC*. Error bars: standard error of the mean of three biological replicates.

Europe PMC Funders Author Manuscripts

(h) ChIP-qPCR for H3Cit on regulatory regions of *Tcl1* and *Nanog* in mES, NS and iPS cells. Error bars: standard deviation of three technical qPCR replicates. Representative of three experiments.

Asterisks denote difference with ES cells (a) or media (b), Control (f, g) and between samples (f); - not significant, * P≤0.05, ** P≤0.01, *** P≤0.001, **** P≤0.0001, by ANOVA (a,f) or t-test (b,g).

Europe PMC Funders Author Manuscripts

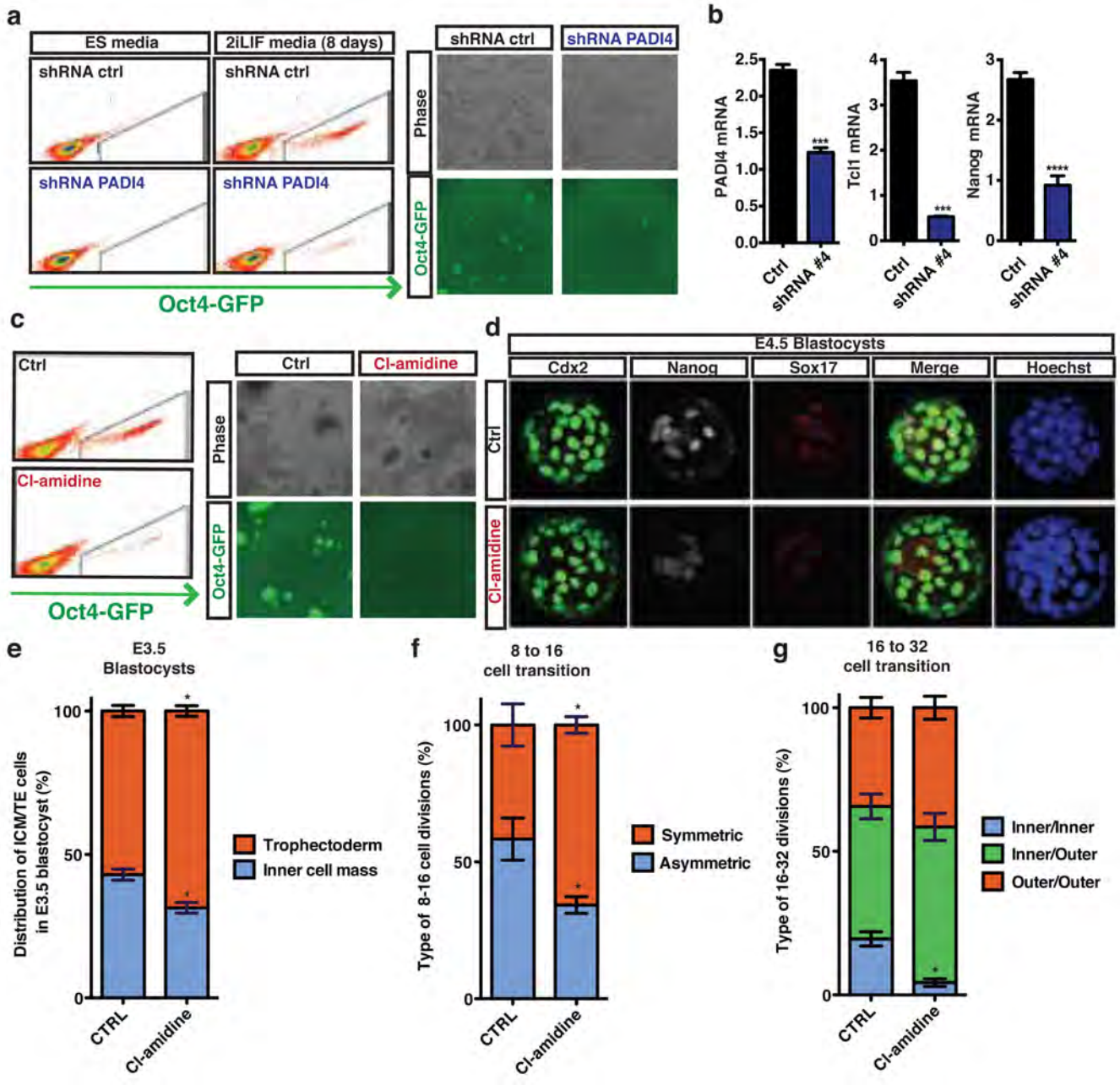


Figure 2: Citrullination and PADI4 regulate pluripotency during reprogramming and early embryo development

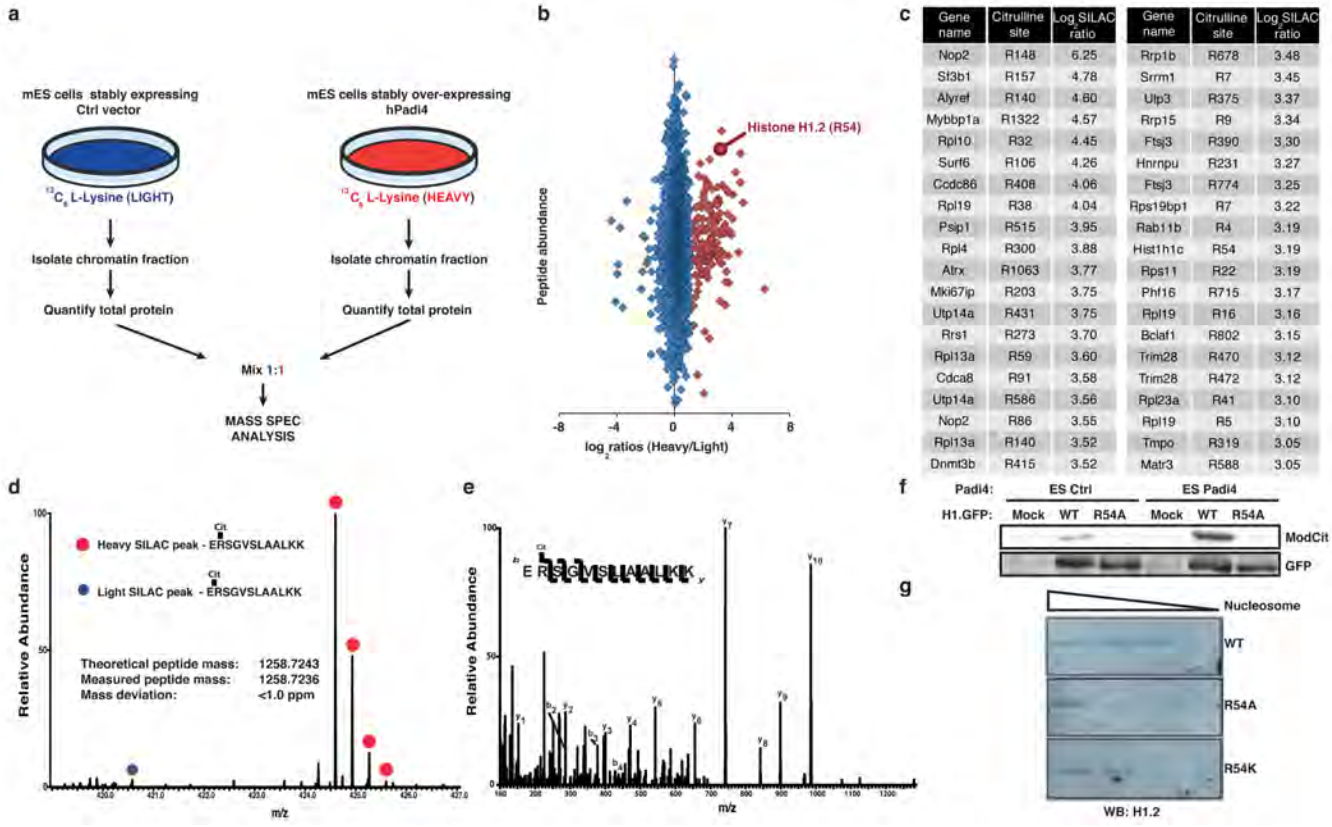
(a) Flow cytometry analysis and phase contrast/fluorescence images for the assessment of Oct4-GFP reporter expression after reprogramming of pre-iPS cells stably expressing *Padi4* and Ctrl shRNAs. Representative of four independent experiments. Time-lapse video in supplementary data online.

(b) qRT-PCR for expression of *Tcl1*, *Nanog* and *Padi4* at the end of the above reprogramming assay. Error bars: standard error of the mean of triplicate samples.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

(c) Flow cytometry analysis and phase contrast/fluorescence images for the assessment of Oct4-GFP reporter expression after reprogramming assay of pre-iPS cells treated with 200μM Cl-amidine. Representative of three independent experiments.
(d) E4.5 blastocysts from 2-cell stage embryos treated with 10μM Cl-amidine. SOX17 (primitive endoderm marker, red), Cdx2 (trophectoderm marker, green), Nanog (epiblast marker, white) and HOECHST 33342 (blue).
(e) Distribution of inner cell mass versus trophectoderm cells in E3.5 blastocyst treated as above.
(f,g) Time-lapse analysis of embryos in 10μM Cl-amidine from 2-cell stage. Number of symmetric versus asymmetric divisions at the 8 to 16-cell transition **(f)** and type of divisions at the 16 to 32-cell transition **(g)**. Error bars: standard error of the mean.
Statistical significance was determined by unpaired t-test (b), or Mann Whitney test upon non-normal distribution (e-g). Asterisks denote difference with Control; - not significant, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$.



Gene name	Citrulline site	Log2 SILAC ratio	Gene name	Citrulline site	Log2 SILAC ratio
Nop2	R148	6.25	Rrp1b	R678	3.48
Sf3b1	R157	4.78	Srrm1	R7	3.45
Alyref	R140	4.60	Utp3	R375	3.37
Mybbp1a	R1322	4.57	Rrp15	R9	3.34
Rpl10	R32	4.45	Ftsj3	R390	3.30
Surf6	R106	4.26	Hnrnpu	R231	3.27
Ccdc86	R408	4.06	Ftsj3	R774	3.25
Rpl19	R38	4.04	Rps19bp1	R7	3.22
Psip1	R515	3.95	Rab11b	R4	3.19
Rpl4	R300	3.88	Hist1h1c	R54	3.19
Atrx	R1063	3.77	Rps11	R22	3.19
Mki67ip	R203	3.75	Phf16	R715	3.17
Utp14a	R431	3.75	Rpl19	R16	3.16
Rrs1	R273	3.70	Bclaf1	R802	3.15
Rpl13a	R59	3.60	Trim28	R470	3.12
Cdca8	R91	3.58	Trim28	R472	3.12
Utp14a	R586	3.56	Rpl23a	R41	3.10
Nop2	R86	3.55	Rpl19	R5	3.10
Rpl13a	R140	3.52	Tmpo	R319	3.05
Dnmt3b	R415	3.52	Matr3	R588	3.05

Figure 3: PADI4 citrullinates Arg54 on linker histone H1 and affects its binding to nucleosomal DNA

(a) Experimental strategy for screening for PADI4 citrullination substrates in the chromatin fraction of ES cells.

(b) Scatter plot representing the overall fold change for all identified citrullination sites. Red diamonds: PADI4-regulated citrullinations.

(c) Table representing the 40 most highly regulated PADI4 substrates, their individual citrullination sites and the $\log_2$ SILAC ratio. Complete dataset in Supplementary Table 4.

(d) Quantification of citrullination site R54 on H1.2 through differential regulation of the triply charged peptide ERSGVSLAALKK.

(e) Fragmentation spectra of the triply charged and heavy SILAC labeled LysC peptide ERSGVSLAALKK surrounding Arginine 54 of H1.2. The y and b series indicate fragments at amide bonds of the peptide.

(f) Citrullination immunoblot of wild-type and R54A mutant GFP-tagged H1.2 expressed and pulled-down from ES cells expressing *PADI4* or control vector (Mock). Control for the efficiency of the pull-down: GFP.

(g) Nucleosome pull-down assay using wild-type and R54-mutant H1.2. WB: Western Blot.

Europe PMC Funders Author Manuscripts

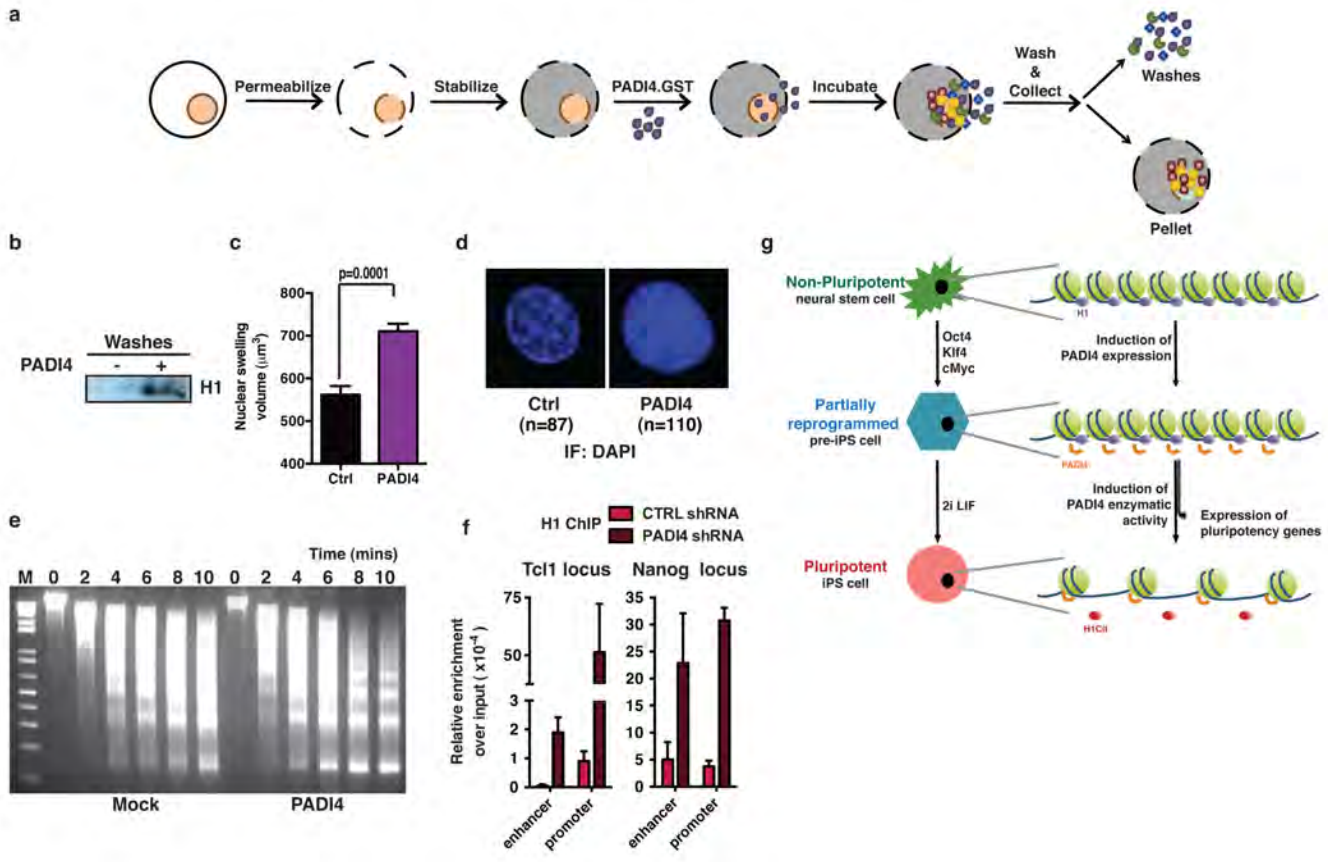


Figure 4: PADI4 evicts histone H1 from chromatin and affects chromatin condensation

(a) Schematic representation of treatment of C2C12 myoblast nuclei with recombinant PADI4.

(b) Immunoblot analysis of the wash fraction after the above treatment, for histone H1.2.

(c,d) Quantification of nuclear volume **(c)** and representative DAPI fluorescence **(d)** upon treatment of permeabilized C2C12 nuclei with recombinant PADI4. Error bars: standard error of the mean. Statistical significance determined by unpaired student t-test.

(e) Micrococcal nuclease digestion of C2C12 cells overexpressing an empty vector (Mock) or hPADI4.

(f) ChIP-qPCR for H1.2 on the regulatory regions of *Tcl1* and *Nanog* in mES cells stably expressing *Padi4* or Ctrl shRNA. Error bars: standard error of the mean of three technical qPCR replicates. Representative of two experiments.

(g) Proposed model for the role of PADI4 in the regulation of pluripotency.

Europe PMC Funders Author Manuscripts