

ARTICLE

Received 11 Mar 2014 | Accepted 28 Apr 2014 | Published 29 May 2014

DOI: 10.1038/ncomms4995

OPEN

# PADI4 acts as a coactivator of Tal1 by counteracting repressive histone arginine methylation

Stephan Kolodziej<sup>1</sup>, Olga N. Kuvardina<sup>1</sup>, Thomas Oellerich<sup>2</sup>, Julia Herglotz<sup>1,†</sup>, Ingo Backert<sup>1,†</sup>, Nicole Kohrs<sup>1</sup>, Estel.la Buscató<sup>3</sup>, Sandra K. Wittmann<sup>3</sup>, Gabriela Salinas-Riester<sup>4</sup>, Halvard Bonig<sup>5</sup>, Michael Karas<sup>3</sup>, Hubert Serve<sup>2,6</sup>, Ewgenij Proschak<sup>3,6</sup> & Jörn Lausen<sup>1</sup>

The transcription factor Tal1 is a critical activator or repressor of gene expression in hematopoiesis and leukaemia. The mechanism by which Tal1 differentially influences transcription of distinct genes is not fully understood. Here we show that Tal1 interacts with the peptidylarginine deiminase IV (PADI4). We demonstrate that PADI4 can act as an epigenetic coactivator through influencing H3R2me2a. At the Tal1/PADI4 target gene *IL6ST* the repressive H3R2me2a mark triggered by PRMT6 is counteracted by PADI4, which augments the active H3K4me3 mark and thus increases *IL6ST* expression. In contrast, at the *CTCF* promoter PADI4 acts as a repressor. We propose that the influence of PADI4 on *IL6ST* transcription plays a role in the control of *IL6ST* expression during lineage differentiation of hematopoietic stem/progenitor cells. These results open the possibility to pharmacologically influence Tal1 in leukaemia.

<sup>1</sup>Georg-Speyer-Haus, Institute for Tumor Biology and Experimental Therapy, Paul-Ehrlich-Strasse 42-44, D-60596 Frankfurt am Main, Germany.

<sup>2</sup>Department of Medicine, Hematology/Oncology, Johann-Wolfgang-Goethe University, Theodor-Stern-Kai 7, D-60590 Frankfurt am Main, Germany.

<sup>3</sup>Institute of Pharmaceutical Chemistry, Johann-Wolfgang-Goethe University, Max-von-Laue-Strasse 9, D-60438 Frankfurt am Main, Germany. <sup>4</sup>Medical-University Goettingen, Transcriptome Analysis Laboratory, Justus-von-Liebig-Weg 11, D-37077 Goettingen, Germany. <sup>5</sup>German Red Cross Blood Service and Institute for Transfusion Medicine and Immunohematology, Johann-Wolfgang-Goethe University, Sandhofstrasse 1, D-60528 Frankfurt am Main, Germany.

<sup>6</sup>German Cancer Consortium (DKTK), Heidelberg, Germany. † Present addresses: Heinrich-Pette-Institute, Leibniz Institute for Experimental Virology, Martinistrasse 52, D-20251 Hamburg, Germany (J.H.); University Hospital Erlangen, Kussmaul Campus for Medical Research, Hartmannstrasse 14, D-91052 Erlangen, Germany (I.B.). Correspondence and requests for materials should be addressed to J.L. (email: lausen@em.uni-frankfurt.de).

The translocated in leukaemia 1 (Tal1; or stem cell leukaemia, Scl) transcription factor is essential for the generation of embryonic hematopoietic stem cells (HSCs). Tal1-deficient mice are embryonically lethal owing to the absence of hematopoiesis and to vascular defects<sup>1–4</sup>. In adult hematopoiesis, Tal1 plays a role in gene regulation and differentiation of monocytes and osteoclasts<sup>5,6</sup>, and at the megakaryocyte/erythrocyte branching point<sup>7–10</sup>. Tal1 also participates in the control of the balance between proliferation and differentiation during erythroid differentiation and influences adult HSC quiescence<sup>11,12</sup>. Furthermore, ectopic expression of Tal1 caused by chromosomal translocation is linked to T-cell acute lymphoblastic leukaemia<sup>13–15</sup>. Mouse models suggest that additional alterations in cofactors of Tal1, such as LMO2, facilitate manifestation of the leukemogenic potential of Tal1<sup>16–19</sup>. Tal1 is a bHLH-class transcription factor, which binds to the DNA-recognition E-box motif CANNTG as a heterodimer with E-proteins such as HEB or E47. Tal1 acts in gene and cell-type-specific gene regulatory complexes containing the E2A proteins, GATA1, RUNX1, FLI1, LMO2, Ldb1, ETS1 and ETO2 (MTG16)<sup>11,20–24</sup>. Furthermore, Tal1 can interact with the histone modifying coactivator proteins CBP/P300, P/CAF<sup>25</sup> and the corepressor proteins Sin3a and LSD1<sup>26–28</sup>. These interactions determine whether Tal1 acts as a repressor or an activator of gene expression. For this reason, knowledge of the distinctive association of Tal1 with epigenetic cofactors is crucial to the understanding of Tal1 function. Moreover, because histone-modifying enzymes are accessible to small-molecule inhibitors, identification of Tal1 interacting enzymes could uncover targets for a molecular therapy of Tal1-dependent leukaemias.

The peptidylarginine deiminase IV (PADI4 or PAD4) converts arginine residues at histone tails to citrulline<sup>29</sup>. Citrullination can take place at unmodified or methylated arginines and counteracts arginine methylation by protein arginine methyltransferases (PRMTs)<sup>30–34</sup>. *In vitro* PADI4 can deiminate arginine residues at positions H3R2, H3R8, H3R17 and H3R26, and PADI4 has been demonstrated to prevent arginine methylation at position H3R17 by PRMT4 (or coactivator-associated arginine methyltransferase 1)<sup>30</sup> in cells. H3R17 methylation by PRMT4 activates gene expression, and the repressive function of PADI4 on target genes has been linked to the negative influence on H3R17 methylation and its association with HDAC2<sup>35</sup>. However, citrullination of H3R8 by PADI4 can result in activation due to inhibition of HP1 binding<sup>36</sup>. Recently, citrullination of H4R3 by PADI4 has been linked to the p53 pathway<sup>37</sup>. A genome-wide analysis of PADI4 activity revealed that PADI4 can associate with active genes and act as an activator of gene expression by citrullination of transcription factors<sup>38</sup>. These results argue for a context-specific repressive or activating transcriptional cofactor activity of PADI4. PADI4 expression has been detected in human CD34+ stem cell/progenitor cells<sup>39</sup>, and PADI4 plays a role in regulation of hematopoietic progenitor proliferation<sup>40</sup>. Recently, it was shown that PADI4 is part of a transcriptional network, which regulates pluripotency<sup>41</sup>. In cancer, high PADI4 expression was connected to tumor growth and the PADI4 inhibitor Cl-amidine reduces growth of a subset of cell lines<sup>42</sup>.

In this study we identified PADI4 as an interaction partner of Tal1 by affinity purification and subsequent stable isotope labelling by amino acids in cell culture (SILAC)-based mass spectrometric identification of co-purified proteins<sup>43</sup>. For this approach, we used the high-affinity biotin-tag birA-ligase system<sup>11,23,44</sup>. By knockdown of PADI4 and Tal1 in the erythroleukemia cell line HEL, we identified co-regulated genes and direct Tal1/PADI4 target genes. We found that Tal1 and PADI4 are bound to the promoters of *IL6ST* (interleukin 6 signal transducer, gp130), which mediates IL6 signalling, and *CTCF*

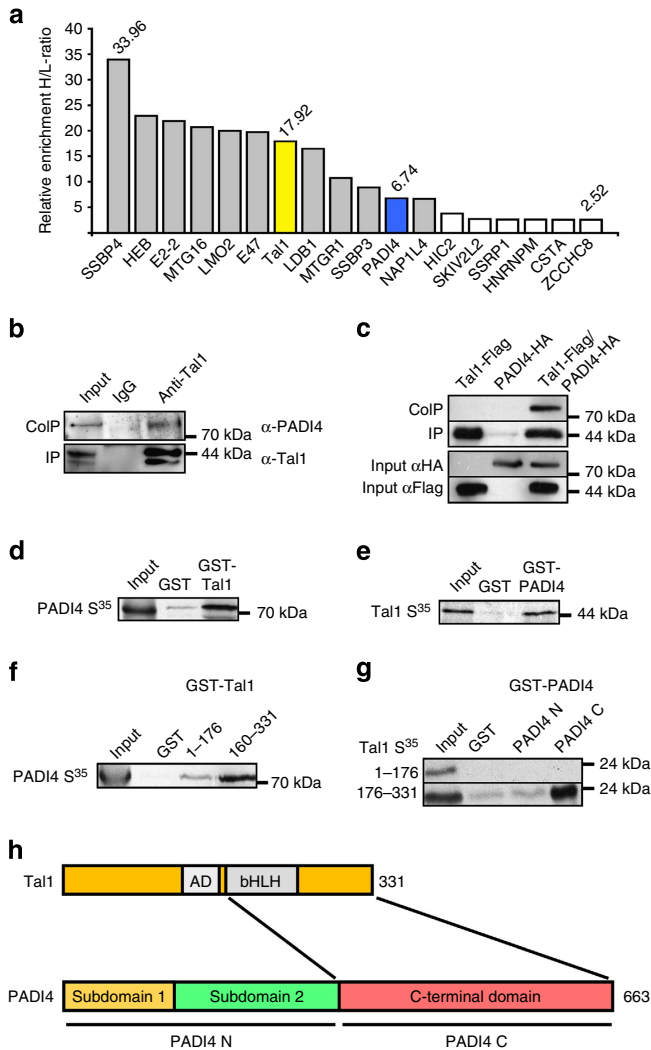
(CCCTC-binding factor), a central transcriptional regulator<sup>45–48</sup>. In particular, we demonstrate that PADI4 acts as a coactivator of Tal1 by counteracting the repressive H3R2me2a mark at the *IL6ST* promoter. This novel function of PADI4 sheds light on the mechanisms of gene regulation by the central hematopoietic transcription factor Tal1.

## Results

**PADI4 is a novel Tal1 interaction partner.** For identification of Tal1 interaction partners, we combined affinity purification using the birA-biotin tag system with quantitative mass spectrometric analysis using SILAC<sup>43</sup>. Erythroleukemia cells harbouring the birA-ligase together with a bir-tagged Tal1 (bir-Tal1) were grown in heavy SILAC medium (H)<sup>10</sup>. *In vivo* biotinylated bir-Tal1 protein was affinity purified using streptavidine beads and mixed in equimolar amounts with a control sample from birA-ligase cells, which were grown in light SILAC medium (L). Purified Tal1 and copurified proteins were separated by one-dimensional-polyacrylamide gel electrophoresis (PAGE) and proteins were identified by mass spectrometry. The relative abundance of proteins in the bir-Tal1 sample compared with the birA-ligase only cells was determined by calculating the ratio between peak intensities of identified peptides from the heavy (H, bir-Tal1) versus the light (L, birA-ligase) sample. An H/L ratio of > 5 was considered as a highly significant enrichment in the bir-Tal1 sample. A detailed scheme of the experimental procedure is given (Supplementary Fig. 1). Besides Tal1 itself, we identified known members of Tal1 complexes such as MTG1, MTG16 (ETO2), LDB1, LMO2 and the basic helix-loop-helix transcription factors HEB, E2-2 and E47 (E2A) (Fig. 1a). Furthermore, the single-stranded DNA-binding proteins 3 and 4 (SSBP3 and SSBP4) and the nucleosome assembly protein 1-like-4 were identified. Figure 1a shows significantly enriched proteins (ratio H/L > 5) and proteins with an H/L ratio between 2.5 and 5. An extended list of identified and quantified proteins is given in (Supplementary Table 1) as well as a complete list is provided in the (Supplementary Data 1).

Interestingly, PADI4 was identified as a potential Tal1-associated protein (Fig. 1a; Supplementary Fig. 2). PADI4 is an epigenetic enzyme, which converts arginine residues at histone tails to citrulline<sup>29</sup>. Because we are especially interested in epigenetic mechanisms connected to Tal1, we further analysed the Tal1/PADI4 association. We could show the interaction of Tal1 with PADI4 at the endogenous level from the erythroleukemia cell line HEL (Fig. 1b). We verified the interaction of Tal1 with PADI4 by coimmunoprecipitation (CoIP) in transiently transfected HEK293 cells (Fig. 1c). Interaction between the proteins was also shown in a glutathione *S*-transferase (GST)-pull-down with *in vitro*-translated PADI4 and in *Escherichia coli* produced GST-Tal1, and *vice versa* (Fig. 1d,e). The main interaction domain of Tal1 with PADI4 is located within the C terminus of the protein. The C-terminal GST-Tal1 construct of amino acids 160–331 interacted stronger with full-length PADI4 than the N-terminal GST-Tal1 of amino acids 1–176 (Fig. 1f). The interaction domain on the PADI4 side involves the C-terminus of PADI4<sup>49</sup>, which displayed a robust interaction with the C terminus of Tal1 (Fig. 1g). Taken together, we identified PADI4 as a novel interaction partner of Tal1 and the C-terminal parts of the proteins mediate this interaction (Fig. 1h).

**PADI4 and Tal1 coinfluence gene expression.** To functionally analyse Tal1 interaction with PADI4, we knocked down PADI4 and Tal1 by short hairpin RNA (shRNA) in HEL cells. Knockdown was confirmed by a reduction of Tal1 and PADI4 at the



**Figure 1 | PADI4 is a novel Tal1 interaction partner.** (a) Identification of members of the Tal1 interactome by affinity purification and SILAC-based MS analysis. Proteins that showed an enrichment of heavy (H)- to light (L)-labelled peptides above 5 (ratio H/L > 5) were identified as components of the Tal1 interactome in the SILAC-based MS analysis and are marked in grey, yellow (Tal1) and blue (PADI4). Further not significant enriched proteins with an H/L ratio of fivefold to 2.5-fold are shown. Detailed information including total number of all identified and quantified proteins is provided in Supplementary Data 1. (b) Endogenous Tal1 interacts with endogenous PADI4 in HEL cells. Immunoprecipitation (IP) was performed using an anti-Tal1 antibody, coimmunoprecipitated (CoIP) PADI4 was detected with an anti-PADI4 antibody. (c) Tal1 interacts with PADI4 in a CoIP performed with lysates of transfected HEK293 cells. IP was performed with anti-Flag M2 beads, which pulls out Flag-Tal1. Coimmunoprecipitated HA-tagged PADI4 was detected with an anti-HA antibody. (d,e) *In vitro* translated PADI4 interacts with GST-Tal1 and *in vitro* translated Tal1 interacts with GST-PADI4 in a GST-pull-down. (f) Interaction of the C terminus of Tal1 with PADI4 in GST-pull-down. (g) Interaction of the C terminus of PADI4 with the C terminus of Tal1 in GST-pull-down. (h) Schematic representation of the interaction between Tal1 and PADI4. The approximate position of proteins is given.

messenger RNA (mRNA) and protein levels (Fig. 2a,b; Supplementary Fig. 3). Subsequently, we performed microarray gene expression analysis. Knockdown of Tal1 altered expression of 1,203 genes and knockdown of PADI4 changed expression of 702 genes (Fig. 2c). When we merged the gene lists of Tal1- and

PADI4-influenced genes, we found that the expression of 414 genes was differentially regulated in both data sets (Fig. 2d).

Gene ontology analysis of the 414 common genes using DAVID<sup>50</sup> revealed that the largest number of these is involved in regulation of transcription (Fig. 2e). Within the regulated genes with the most strongly enriched gene ontology (GO)-term categories were genes involved in leukocyte differentiation, chromosome organization and hemopoiesis (Table 1). Of 414 common genes, 162 genes were upregulated and 248 were downregulated upon Tal1 and PADI4 knockdown. The large number of co-regulated genes is consistent with an intricate relationship of Tal1 and PADI4 in gene regulation. When we compared the list of 414 Tal1- and PADI4-influenced genes from HEL erythroleukemia cells with recently identified direct Tal1 target genes in the murine hematopoietic cell line HPC-7<sup>21</sup>, we found that 162 genes were already described as direct Tal1 target genes in HPC-7 cells.

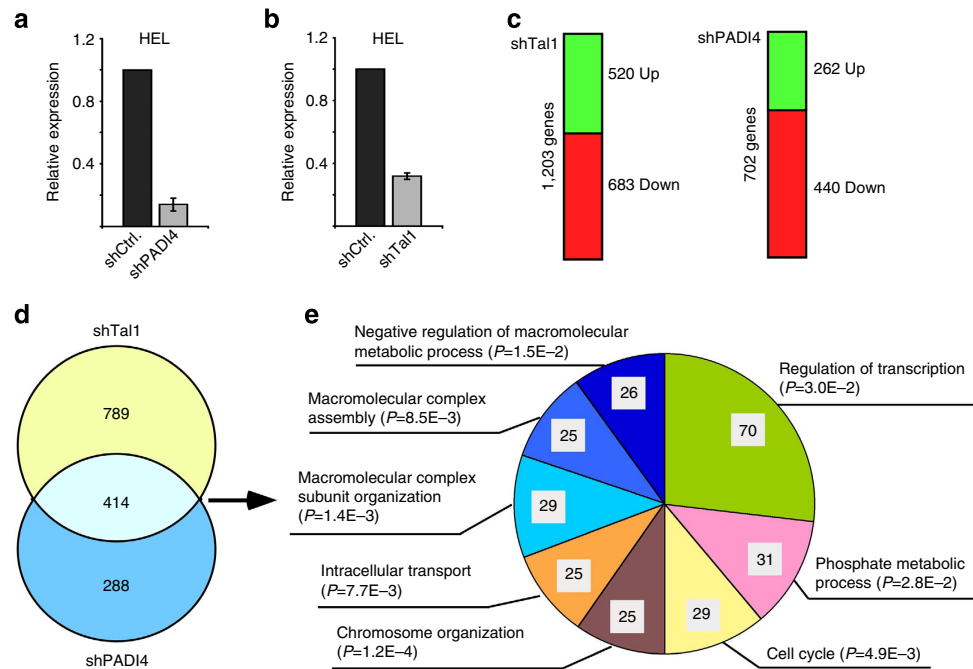
We reanalysed Tal1/PADI4-regulated genes, which are involved in leukocyte differentiation and hematopoiesis (compare Table 1) by quantitative reverse transcriptase-PCR (Supplementary Fig. 4). *CTCF*, *FGFR1OP2* and *TNF* were upregulated upon Tal1 or PADI4 knockdown (Fig. 3a-c); *IL6ST*, *CEBPε* and *VEGFA* were downregulated upon Tal1 or PADI4 knockdown (Fig. 3d-f).

Subsequently, we performed chromatin immunoprecipitation (ChIP) using specific antibodies against Tal1 or PADI4 to determine whether identified genes are direct target genes of Tal1 or PADI4. Guided by published ChIP-seq data, we designed primer pairs covering E-boxes at the promoters of identified genes and detected Tal1 and PADI4 binding to the promoter regions of *IL6ST* and *CTCF* by ChIP (Fig. 4; Supplementary Fig. 5). Figure 4a indicates the genomic position of the analysed promoter region of *IL6ST* and the location of primers for subsequent ChIP analysis. Using primer pairs at different locations at the *IL6ST* promoter for ChIP-qPCR analysis, we mapped Tal1 binding to the *IL6ST* promoter to the proximity of the first non-coding exon of *IL6ST* (Fig. 4b). Similarly, PADI4 binding was localized to the same region of the *IL6ST* promoter (Fig. 4c). To determine whether Tal1 and PADI4 concomitantly bind to the *IL6ST* promoter, we performed ChIP-ReChIP analysis and could detect Tal1 and PADI4 together close to the transcriptional start site but not at a downstream region, which served as a negative control (Fig. 4d-f). A similar analysis was performed for *CTCF* (Supplementary Fig. 5).

To evaluate the contribution of Tal1 on transcription, we performed ChIP experiments upon knockdown of Tal1. Knockdown of Tal1 resulted in diminished Tal1 occupancy on the *IL6ST* promoter (Fig. 4g). PADI4 binding is decreased upon knockdown of Tal1, which indicates that Tal1 takes part in PADI4 recruitment (Fig. 4h). Furthermore, knockdown resulted in reduced presence of the active histone marks H3K9ac and H3K4me3 (Fig. 4i), which is in line with the downregulation of *IL6ST* expression upon loss of Tal1 (compare Fig. 3). The repressive H3R2me2a mark is unchanged upon Tal1 knockdown and the active H3R17me2a mark is increased (Fig. 4j). Moreover, we observed that PRMT6 and PRMT4 binding is slightly reduced upon Tal1 knockdown (Fig. 4k).

Taken together, we confirmed *IL6ST* and *CTCF* as direct Tal1/PADI4 target genes. *CTCF* is a central regulator of gene expression and plays a role at the megakaryocytic/erythrocytic branching point<sup>45-48</sup>. *IL6ST* (gp130) is an important transducer of cytokine signalling, which plays a central role in hematopoiesis and is a target molecule for cancer therapy<sup>51-54</sup>.

**PADI4 influences histone arginine methylation.** PADI4 can modify histone arginines by citrullination and thus counteract



**Figure 2 | Gene expression analysis upon Tal1 and PADI4 knockdown.** (a) HEL cells were transfected with shRNAs against PADI4 (shPADI4) and the knockdown evaluated by quantitative real-time PCR. Error bars represent the s.d. of at least four determinations from two independent transductions. (b) HEL cells were transfected with shRNAs against Tal1 (shTal1) and the knockdown evaluated by quantitative real-time PCR. Error bars represent the s.d. of at least four determinations from two independent transductions. (c) Gene expression array analysis was performed with shTal1 and shPADI4 cells. Knockdown of Tal1 altered expression of 1,203 genes and knockdown of PADI4 altered expression of 702 genes. (d) Comparison of Tal1- and PADI4-influenced genes. Genes (414) are changed in both data sets. (e) Gene ontology analysis of the 414 genes changed upon Tal1 and PADI4 knockdown. The pie chart gives the enriched GO-terms (GO-term-BP) with the highest numbers of included genes. Analysis was performed using DAVID using standard settings. *P*-values are given.

**Table 1 | GO-terms with the most significant enrichment.**

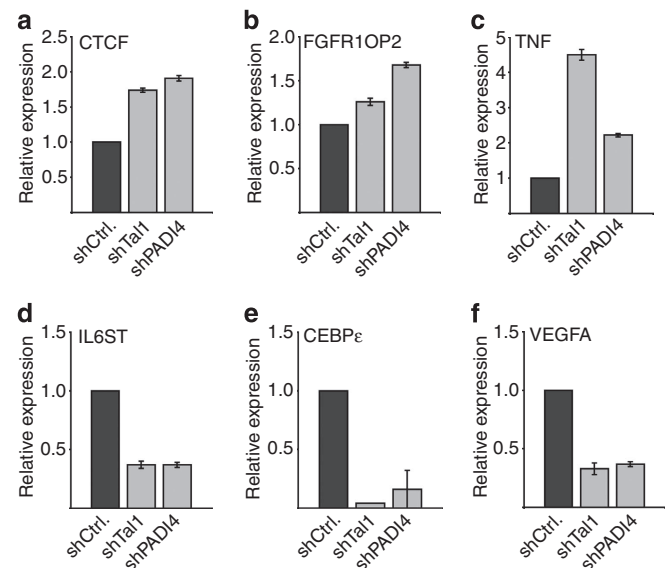
GO-term BP	Count	<i>P</i> -value
Leukocyte differentiation	13	$2.5E-5$
T-cell differentiation	9	$7.1E-5$
Chromatin assembly	12	$8.9E-5$
Chromosome assembly	25	$1.2E-4$
Nucleotide and nucleic acid transport	11	$1.6E-4$
Hemopoiesis	16	$1.8E-4$
RNA localization	10	$2.9E-4$
Lymphocyte differentiation	10	$3.6E-4$

GO, gene ontology.

arginine methylation by PRMTs. Because *CTCF* and *IL6ST* expression are differentially effected upon Tal1 and PADI4 knockdown, we performed a site-by-site comparison of the histone modification status of the *CTCF* and *IL6ST* promoters after PADI4 knockdown.

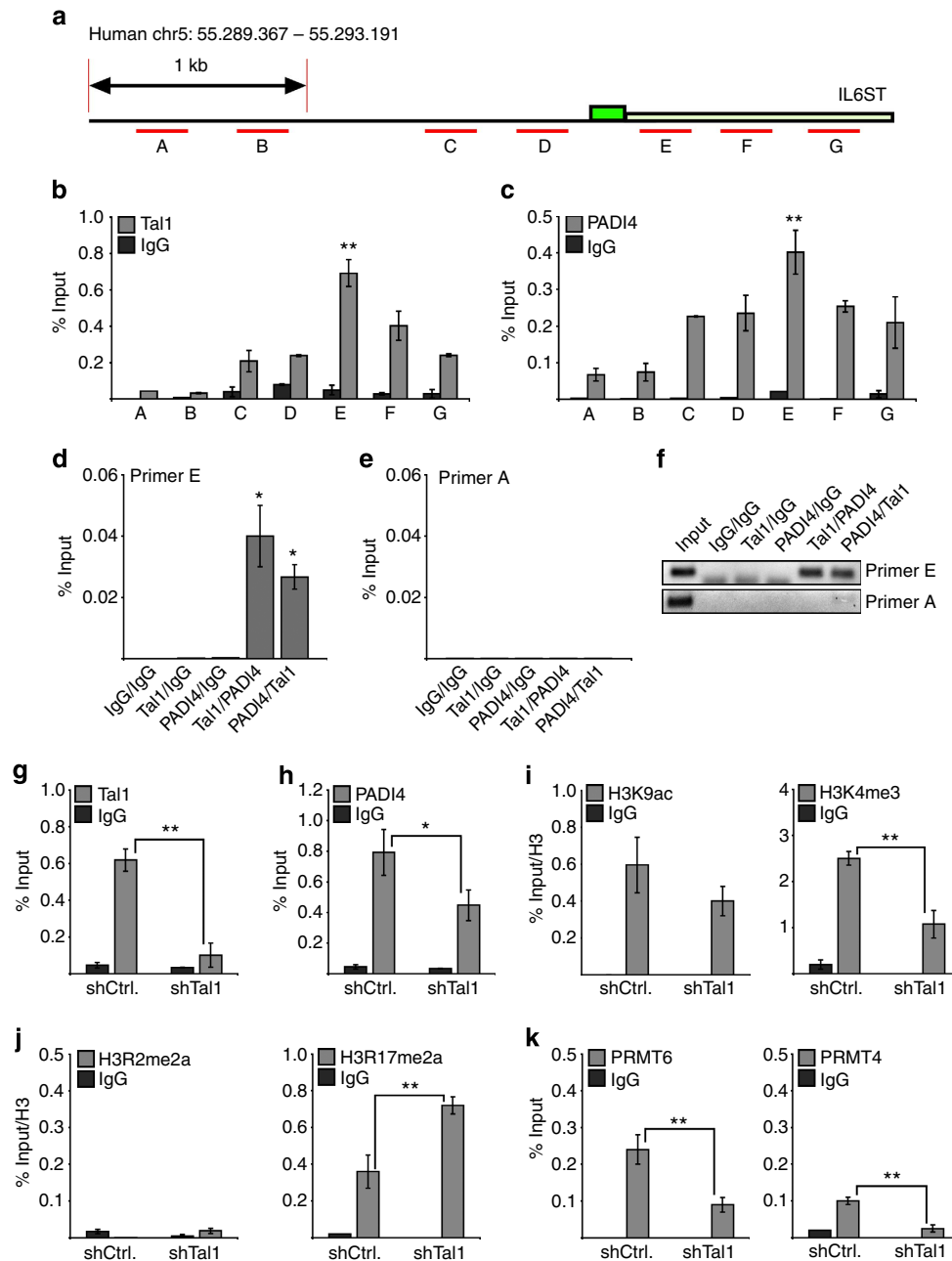
As expected, upon knockdown of PADI4 by shRNA the amount of PADI4 on the promoters of *CTCF* and *IL6ST* was strongly reduced (Fig. 5a), while Tal1 occupancy was not changed significantly (Fig. 5b). PADI4 is able to counteract the active H3R17me2a mark, which is mediated by PRMT4. H3R17me2a increased to a robust level on the *CTCF* promoter and on the *IL6ST* promoter already some H3R17me2a was detected, which was further increased upon PADI4 knockdown (Fig. 5c). Concomitantly, we detected an increase of PRMT4 occupancy on the *CTCF* and *IL6ST* promoters (Fig. 5d).

Interestingly, upon PADI4 knockdown we detected the appearance of the repressive H3R2me2a mark at the *IL6ST* promoter, the opposite of what we observed at the *CTCF*



**Figure 3 | Verification of gene expression changes upon Tal1 and PADI4 knockdown.** (a-f) A subset of genes from the array analysis was reanalysed by quantitative real-time PCR. Error bars represent the s.d. from at least four determinations and two independent knockdowns.

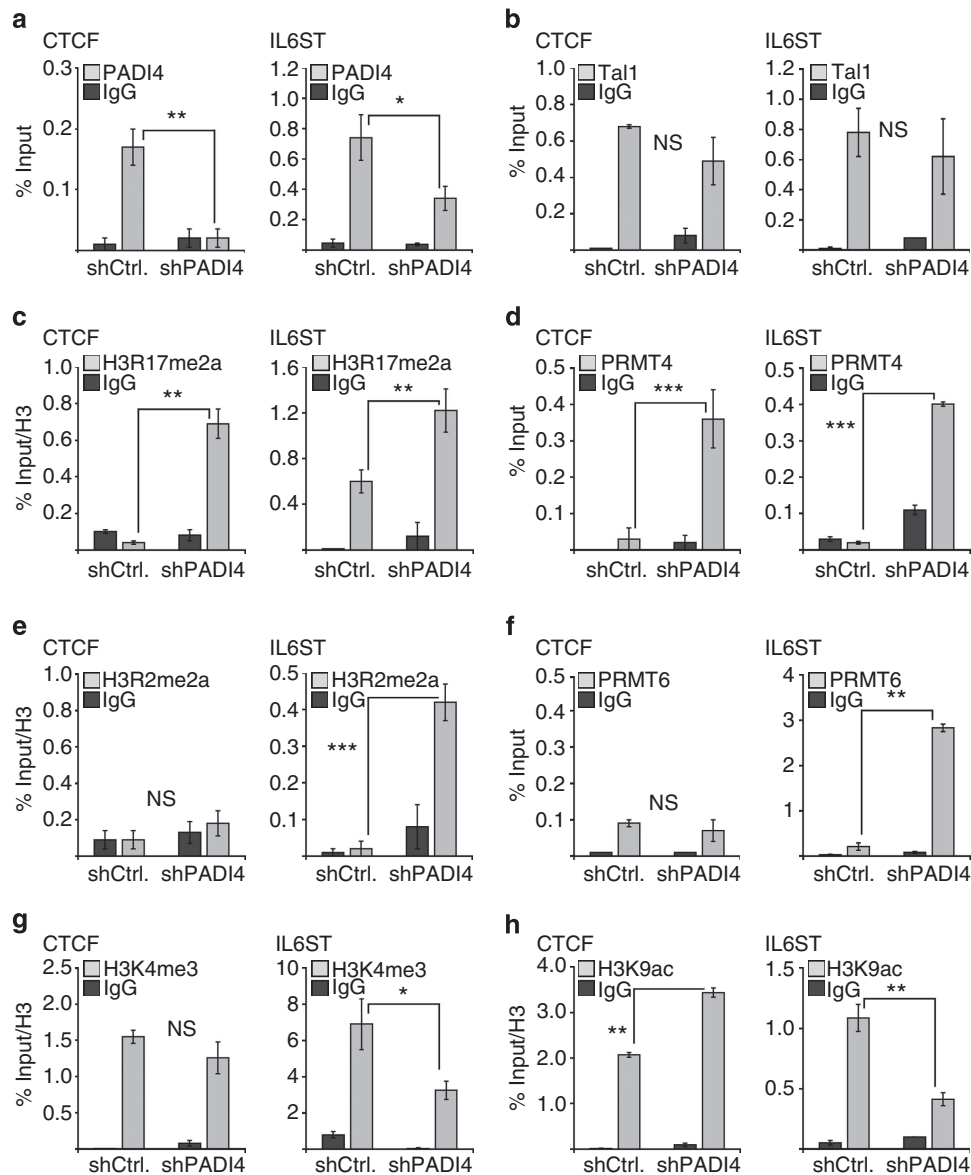
promoter (Fig. 5e). At the same time, we found an increased occupancy of PRMT6 at the *IL6ST* promoter, but not at the *CTCF* promoter (Fig. 5f). PRMT6-mediated H3R2me2a acts in an inhibitory fashion on the H3K4me3 mark. In agreement with this notion, we found the activating H3K4me3 mark unchanged in the case of *CTCF* and decreased in the case of *IL6ST* (Fig. 5g).



**Figure 4 | Analysis of Tal1 and PADI4 binding to the *IL6ST* 5'-region and influence of Tal1.** (a) Schematic representation of the analysed *IL6ST* genomic locus including the 5'-region. The genomic position is given and the position of primer pairs used for ChIP analysis is indicated. The first non-coding exon is marked in dark green and the first intron in light green. (b) Mapping of Tal1 binding to the *IL6ST* 5'-region. Upon ChIP with an anti-Tal1 antibody qPCR with primers at different positions of the *IL6ST* 5'-region was performed. (c) ChIP shows binding of PADI4 to the *IL6ST* 5'-region. Upon ChIP with an anti-PADI4 antibody qPCR with primers at different positions of the *IL6ST* 5'-region was performed. The *P*-value gives the statistical significant difference between the values gathered with primer E compared with primer A. (d-f) Analysis of concomitant Tal1 and PADI4 binding to the *IL6ST* 5'-region by ChIP-ReChIP. (d) Concomitant Tal1 and PADI4 binding is detected at a region close to the first exon of *IL6ST* (primer E). (e) No Tal1/PADI4 is detected with an upstream primer (primer A). (f) Analysis of the ChIP-ReChIP qPCR products by gel electrophoresis at the end point of the qPCR reaction confirms qPCR results. Antibody combinations and primer pairs are given. The *P*-value ( $*P < 0.05$ ) gives the statistical significant enrichment of the ChIP-ReChIP compared with the IgG controls according to Student's *t*-test. (g) ChIP demonstrates decreased Tal1 binding to the *IL6ST* promoter upon Tal1 knockdown. (h) ChIP shows decreased PADI4 binding to the *IL6ST* promoter upon Tal1 knockdown. (i) The histone modification marks H3K9ac and H3K4me3 are decreased upon Tal1 knockdown at the *IL6ST* promoter. (j) The histone modification H3R2me2a remains at a low level upon Tal1 knockdown and H3R17me2a is increased at the *IL6ST* promoter. (k) PRMT6 and PRMT4 binding are decreased upon Tal1 knockdown at the *IL6ST* promoter. Values are shown as percent enrichment compared with the input. Error bars represent the s.d. from at least four determinations. The *P*-values were calculated using the *t*-test,  $*P < 0.05$ ,  $**P < 0.01$ .

In addition, we analysed the H4R3me2a status, which was recently connected to PADI4 activity<sup>37</sup>. H4R3me2a was increased on the *CTCF* and *IL6ST* promoter upon PADI4 knockdown

(Supplementary Fig. 6). The activating histone mark H3K9ac increased at the *CTCF* promoter and decreased at the *IL6ST* promoter (Fig. 5h), which is in agreement with observed



**Figure 5 | Epigenetic changes upon PADI4 knockdown at the *CTCF* and *IL6ST* promoter.** (a) ChIP demonstrates decreased PADI4 occupancy of the *CTCF* and *IL6ST* promoters upon PADI4 knockdown. (b) ChIP shows no changes in Tal1 binding to the *CTCF* and *IL6ST* promoters upon PADI4 knockdown. (c) The PRMT4-mediated H3R17me2a mark is increased upon PADI4 knockdown at the *CTCF* and *IL6ST* promoters. (d) PRMT4 binding is increased at the *CTCF* and *IL6ST* promoters upon PADI4 knockdown. (e) H3R2me2a is increased upon PADI4 knockdown at the *IL6ST* promoter, but not at the *CTCF* promoter. (f) PRMT6 binding is increased at the *IL6ST* promoter upon PADI4 knockdown. (g) The H3K4me3 histone mark is significantly decreased at the *IL6ST* promoter. (h) The positive histone mark H3K9ac decreases upon PADI4 knockdown at the *IL6ST* promoter. Values are shown as percent enrichment compared with the input. Error bars represent the s.d. from at least four determinations. The *P*-values were calculated using the *t*-test, \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001. NS, not significant.

expression changes (compare Fig. 3; Supplementary Fig. 3 for changes of *IL6ST* surface expression).

Altogether, we found that PADI4 acts repressively on *CTCF* expression by keeping the activating H3R17me2a histone mark at a low level. In contrast, the main effect of PADI4 on the *IL6ST* promoter was suppression of the repressive H3R2me2a mark, which counteracts H3K4me3. By this mechanism PADI4 acts as an activator of *IL6ST* expression. These data show that PADI4 influences the epigenetic status of Tal1-target regulatory elements by changing arginine methylation patterns at distinct histone arginines.

**Pharmacological inhibition of PADI4.** Our data show that PADI4 negatively influences expression of Tal1 target genes such

as *CTCF* and promotes transcriptional activation of others, such as *IL6ST*. *IL6ST* (gp130) is an important transducer of cytokine signalling, for example, triggered by interleukin (IL)-6. Furthermore, *IL6ST* has a role in self-renewal of HSCs and hematopoietic differentiation<sup>51,53</sup>. In addition, *IL6ST* signalling contributes to growth control and is thus regarded as a promising target for cancer therapy<sup>54,55</sup>.

The small-molecule Cl-amidine inhibits PADI4 activity<sup>56</sup>. Treatment of HEL cells with Cl-amidine changed a subset of leukocyte or hematopoiesis-associated genes, which were altered upon Tal1 or PADI4 knockdown (Supplementary Fig. 4). Notably, pharmacological inhibition of PADI4 by Cl-amidine decreased *IL6ST* expression in HEL cells (Fig. 6a). Treatment with Cl-amidine did not immediately influence growth of HEL cells,

but on day 9 of treatment a significant decrease in cell numbers compared with the mock control was observed (Supplementary Fig. 7). Other cell lines also displayed sensitivity to Cl-amidine (Supplementary Fig. 7). Upon Cl-amidine treatment, the amount of Tal1 and PADI4 on the *IL6ST* promoter was slightly reduced (Fig. 6b,c). In agreement with our data gathered upon PADI4 knockdown, we detected a strong increase in H3R2me2a upon inhibitor treatment (Fig. 6d, compare Figure 5e). Consistent with the notion that H3R2me2a inhibits the H3K4me3 mark, we also detected a significant decrease in H3K4me3 on the *IL6ST* promoter (Fig. 6e). H3R17me2a did not change upon inhibitor treatment (Fig. 6f). In contrast, at the *CTCF* promoter H3R17me2a was increased and H3R2me2a remained unchanged upon Cl-amidine treatment (Fig. 6g,h).

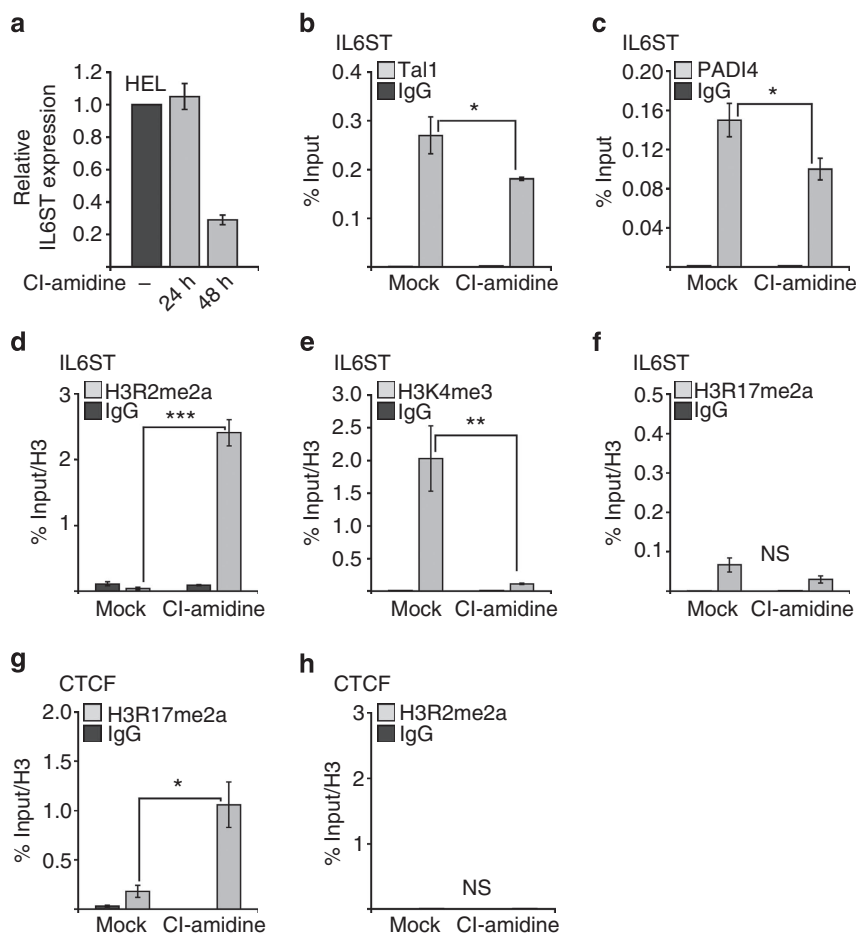
To evaluate the influence of PADI4 on *IL6ST* expression in primary cells, we examined *PADI4* and *IL6ST* expression during differentiation of human CD34<sup>+</sup> progenitor cells. We observed upregulation of *IL6ST* expression during erythroid and monocytic differentiation (Fig. 7a). *IL6ST* mRNA levels remained unchanged upon megakaryocytic differentiation compared with non-induced cells (Fig. 7a). Tal1 binding to the *IL6ST* promoter remained similar during erythroid or megakaryocytic differentiation and was slightly decreased upon monocytic differentiation (Fig. 7b). Concomitantly with upregulation of *IL6ST*, we detected increased

binding of PADI4 to the *IL6ST* promoter upon erythroid and monocytic differentiation (Fig. 7c). During megakaryocytic differentiation PADI4 binding was decreased (Fig. 7c) and H3R2me2a increased (Fig. 7d). During erythroid and monocytic differentiation H3R2me2a was slightly decreased. These data are in agreement with our results gathered in HEL cells, which implicate that PADI4 acts in an activating manner on *IL6ST* expression (compare Fig. 5). Subsequently, we examined whether PADI4 influences hematopoietic differentiation of hCD34<sup>+</sup> cells in a colony-formation assay. When we overexpressed PADI4, we observed fewer monocytic colonies, whereas an enzymatically inactive mutant did not show this effect (Fig. 7e). Inversely, knockdown of PADI4 increased the number of monocytic colonies (Fig. 7f).

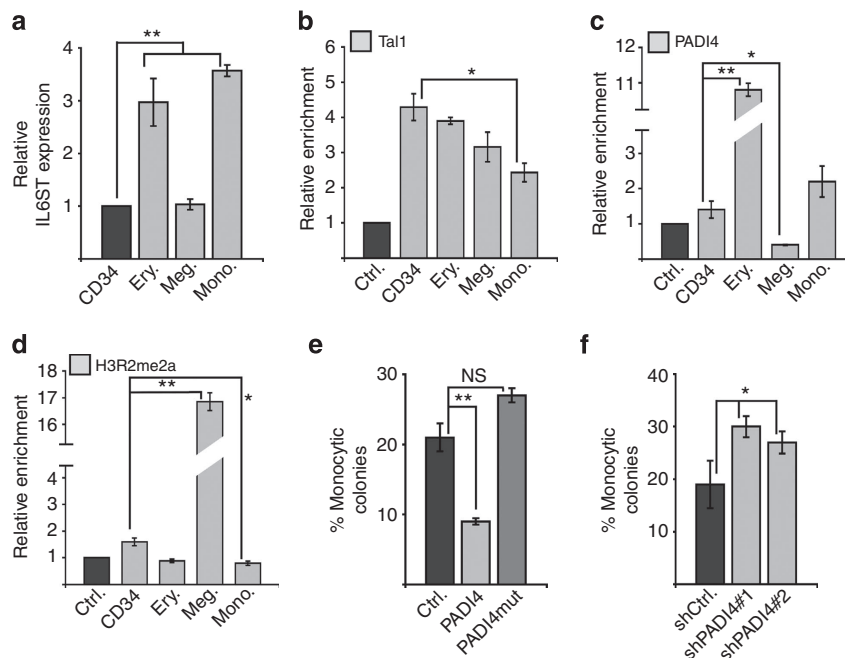
Taken together, PADI4 can act as an activator of Tal1-target gene expression, such as *IL6ST*, by counteracting H3R2me2a methylation triggered by PRMT6 (Fig. 8a). In contrast, at distinct target genes PADI4 is a repressor of gene expression (Fig. 8b).

## Discussion

During hematopoietic differentiation chromatin is extensively reorganized and epigenetic marks are adjusted. This contributes to epigenetic refinement of gene expression programs during

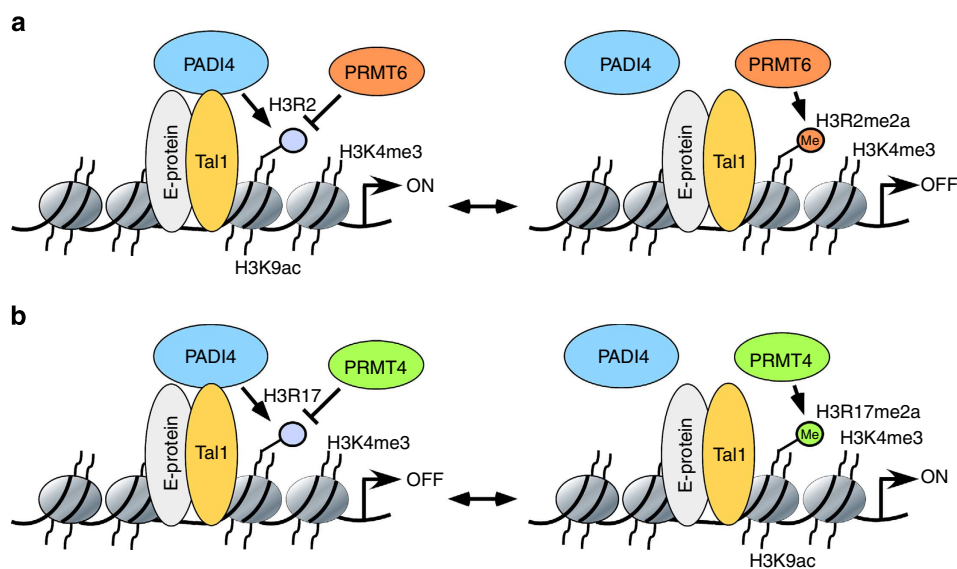


**Figure 6 | The PADI4 inhibitor Cl-amidine influences *IL6ST* gene expression and epigenetic marking.** (a) *IL6ST* expression is decreased upon treatment of HEL cells with Cl-amidine (200  $\mu$ M) for 48 h. (b) Tal1 occupancy of the *IL6ST* promoter upon Cl-amidine treatment of HEL cells measured by ChIP. (c) PADI4 occupancy of the *IL6ST* promoter upon Cl-amidine treatment of HEL cells measured by ChIP. (d) The negative histone mark H3R2me2a is increased at the *IL6ST* promoter upon Cl-amidine treatment of HEL cells measured by ChIP. (e) H3K4me3 is decreased at the *IL6ST* promoter upon Cl-amidine treatment in HEL cells. (f) H3R17me2a at the *IL6ST* promoter is unaffected by Cl-amidine treatment. (g,h) In contrast to the *IL6ST* promoter, at the *CTCF* promoter H3R17me2a was increased and H3R2me2a remained unchanged upon Cl-amidine treatment. Values are shown as percent enrichment compared with the input. Error bars represent the s.d. of at least four determinations. The *P*-values were calculated using the *t*-test, \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001. NS, not significant.



**Figure 7 | *IL6ST* expression changes during differentiation of hCD34<sup>+</sup> cells that correlate with PADI4 recruitment and epigenetic alterations.**

(a) *IL6ST* expression is increased upon erythroid and monocytic differentiation of primary human CD34<sup>+</sup> cells and remains unchanged upon megakaryocytic differentiation. (b) Tal1 binding to the *IL6ST* promoter is unchanged upon erythroid and megakaryocytic differentiation. (c) PADI4 binding to the *IL6ST* promoter increased upon erythroid and monocytic differentiation and decreased upon megakaryocytic differentiation. (d) The repressive H3R2me2a mark is strongly upregulated upon megakaryocytic differentiation and decreased upon erythroid and monocytic differentiation. ChIP assays were performed with at least two independent preparations and measured in duplicates. Values are shown normalized to the input and as relative enrichment over background. (e) Influence of PADI4 on hematopoietic colony formation. Overexpression of PADI4 reduces the number of monocytic colonies in a CFU-C assay. An enzymatic inactive PADI4 mutant does not inhibit monocytic differentiation. (f) Knock down of PADI4 increases the number of monocytic colonies in a colony-formation unit-C assay. hCD34<sup>+</sup> cells were expanded for 4 days, transduced and sorted according to their green fluorescent protein signal. Subsequently, the transduced cells were seeded in methocult. Colonies were evaluated 2 weeks after seeding. Error bars represent the s.d. of three determinations. The *P*-values were calculated using the *t*-test, \**P* < 0.05, \*\**P* < 0.01. NS, not significant.



**Figure 8 | Model of PADI4 activity with Tal1 on regulatory elements of genes, such as *IL6ST* and *CTCF*.** (a) PADI4 as an activator. Tal1 recruits PADI4, which inhibits H3R2me2a by PRMT6. As a consequence H3K4me3 remains at a high level and the gene is transcribed. Upon loss of PADI4 binding H3R2me2a is triggered by PRMT6. Subsequently, the H3K4me3 level is decreased and expression downregulated. (b) PADI4 as a repressor. Tal1 recruits PADI4, which inhibits H3R17me2a by PRMT4 and contributes to low expression of the gene. Upon loss of PADI4 binding H3R17me2a is triggered by PRMT4. Subsequently, the H3K4me3 level increases and expression is upregulated.



differentiation. Transcription factors play a decisive role in the assembly of gene and cell-specific transcriptional complexes, and their deregulation can result in uncontrolled cell growth and subsequently cancer. Consequently, small-molecule inhibitors of epigenetic cofactors such as histone deacetylase inhibitors are highly desired.

The transcription factor Tal1 is a critical regulator of hematopoietic gene expression and can act as an oncogene if aberrantly expressed. In this work, we have identified PADI4 as a novel interaction partner of Tal1 and we found that a large number of genes are co-regulated by Tal1 and PADI4. Strikingly, we identified genes that are concomitantly downregulated by Tal1 or PADI4 knockdown as well as concurrently upregulated genes. This observation strengthens the notion of a close functional connection between Tal1 and PADI4, and also demonstrates the existence of a repressive and an activating mode of cooperation between Tal1 and PADI4.

PADI4 is able to citrullinate monomethylated arginine residues at histones and thus counteracts the activity of PRMTs<sup>30,31</sup>. As was previously shown, H3R17 demethylation by PADI4 counteracts H3R17 methylation by PRMT4<sup>57</sup>. This block of H3R17 methylation has a repressive effect, as a consequence PADI4 acts as a corepressor of gene expression on these genes. This is in line with our observation that knockdown of PADI4 increases PRMT4 recruitment and H3R17 methylation on the *CTCF* promoter and augments *CTCF* gene expression. Thus on the *CTCF* promoter PADI4 exerts inhibitory effects on expression by counteracting H3R17me2a. Because some degree of *CTCF* expression was detectable in the presence of PADI4, a possible scenario is dynamic competition of PADI4 and PRMT4 for access to H3R17.

Unexpectedly, upon PADI4 knockdown we observed a strong increase of H3R2me2a at the *IL6ST* promoter. This increase of the repressive H3R2me2a histone mark<sup>58–60</sup> goes hand-in-hand with downregulation of *IL6ST* expression. H3R2me2a mediated by PRMT6 is able to inhibit the prominent H3K4me3 mark<sup>58–62</sup>. In agreement with this notion, we detected decreased H3K4me3 methylation concomitant with the upregulation of H3R2me2a upon PADI4 knockdown. This is the first report of an involvement of PADI4 in repression of the H3R2me2a mark and an indirect support of the H3K4me3 mark by PADI4. Interestingly, we could also detect an increase of the repressive H3R2me2a mark, a decrease of H3K4me3 and a concomitant downregulation of *IL6ST* expression, when we pharmacologically inhibited PADI4 with the small-molecule compound Cl-amidine. Notably, the slight increase of H3R17me2a upon PADI4 knockdown at the *IL6ST* promoter was not detectable upon Cl-amidine treatment. This could indicate that the presence of PADI4 protein also influences complex formation structurally.

Furthermore, we found that during differentiation of primary human CD34+ progenitor cells *IL6ST* expression is correlated with PADI4 occupancy of the *IL6ST* promoter. Low PADI4 occupancy during megakaryocytic differentiation is related to high levels of H3R2me2a. This result demonstrates that increased PADI4 binding can be associated with gene activation during differentiation. Our observation that PADI4 influences monocytic differentiation in a colony-formation assay shows that the function of PADI4 might be critical for hematopoietic differentiation. The observation that a number of genes important for leukocyte differentiation such as CEBP $\epsilon$  are changed upon Tal1 and PADI4 knockdown, could offer an explanation for the influence of PADI4 on differentiation. Interestingly, it was recently demonstrated that PADI4 regulates c-myc expression in conjunction with the transcription factor LEF1 and in this manner influences multipotent hematopoietic progenitor proliferation<sup>40</sup>. For the detailed analysis of PADI4 function in distinct

hematopoietic lineages, cell-type-specific PADI4 knockout in the mouse model will be required.

Taken together, PADI4 can target different histone arginines<sup>30,31</sup> and our data support the notion of a function of PADI4 as a general opponent of PRMTs. Thus, PADI4 might act as a repressor if activating arginine methylation is inhibited, or an activator if the balance of PADI4 activity is shifted towards suppression of repressive arginine methylation. In this regard, our data show that PADI4 contributes to epigenetic fine-tuning of Tal1 target gene expression in hematopoiesis. A further advance in this field of study would be the identification of an enzyme that reverses citrullination by PADI4 and would facilitate renewed access to arginines for methylation by PRMTs. Our study revealed a novel role of PADI4 as a promoter-dependent epigenetic cofactor of Tal1, which interplays with important epigenetic histone marks and is accessible to pharmacological manipulation.

## Methods

**Biotin purification and mass spectrometry.** Cells expressing the birA-Tal1 protein were grown in heavy (H) SILAC medium and control cells expressing the birA-ligase only, were grown in light SILAC medium (L) (RPMI SILAC medium, Thermo Scientific, Rockford, USA). For biotin purification of bir-Tal1, we used nuclear extracts of  $1 \times 10^8$  K562-BirA control and K562-BirA-Tal1 cells<sup>10</sup>. Extracts were essentially prepared as described<sup>63</sup>. Biotinylated Tal1 was pulled out from cell extracts using streptavidin beads (Dynabeads M-280, Life Technologies, Oslo, Norway) and the beads were washed five times. As a control a pull-down from cells expressing only the birA-ligase was performed. BirA-Tal1 and the birA-ligase control samples were combined. Proteins were separated by one-dimensional SDS-PAGE (4–12% NuPAGE Bis-Tris Gel, Invitrogen) and the entire lane of the Coomassie blue-stained gel was cut into 23 slices. All slices were reduced with 10 mM dithiothreitol (DTT) for 55 min at 56 °C, alkylated with 55 mM iodoacetamide (IAA) for 20 min at 26 °C and digested with trypsin (Serva) overnight at 37 °C. Tryptic peptides were injected into a C18 precolumn (2.5 cm, 360  $\mu$ m o.d., 100  $\mu$ m i.d., Reprosil-Pur 120 Å, 5  $\mu$ m, C18-AQ, Dr Maisch GmbH) at a flow rate of 10  $\mu$ l min<sup>-1</sup>. Bound peptides were eluted and separated on a C18 capillary column (20 cm, 360  $\mu$ m o.d., 75  $\mu$ m i.d., Reprosil-Pur 120 Å, 3  $\mu$ m, C18-AQ, Dr Maisch GmbH) at a flow rate of 300 nl min<sup>-1</sup>, with a gradient from 5 to 35% acetonitrile (ACN) in 0.1% formic acid for 90 min using an proxeon-nano-LC coupled to an Q-exactive mass spectrometer (Thermo Electron). Mass spectrometry (MS) conditions were as follows: spray voltage, 1.8 kV; heated capillary temperature, 270 °C; normalized collision energy 28. An underfill ratio of 1.2% and intensity threshold of 4.0 e4 were used. The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS/MS acquisition. Survey MS spectra were acquired in the Orbitrap (*m/z* 350–1,600) with the resolution set to 70,000 at *m/z* 200 and automatic gain control target at  $2 \times 10^5$ . The 15 most intense ions were sequentially isolated for higher-energy collisional dissociation MS/MS fragmentation and detection. Ions with single and unrecognized charge states were excluded. Raw data were analysed with 'MaxQuant' software (version 1.3.0.5). Uniprot human (version 27.08.2012 with 86,725 entries) was used as a sequence database. Up to two missed cleavages of trypsin were allowed. Oxidized methionine was searched as variable modifications, whereas cysteine carbamido-methylation as a fixed modification. The modifications corresponding to arginine and lysine labelled with heavy stable isotopes were handled as fixed modifications in the Maxquant search, if applicable, after identification of SILAC pairs by MaxQuant. The false-positive rate was set to 1% at the peptide level, the false discovery rate was set to 1% at the protein level and the minimum required peptide length was set to six amino acids. Proteins with an H/L ratio of > 5 were defined as high-confidence members of the Tal1 interactome. The full list of identified proteins is given in Supplementary Data 1.

**Interaction assays.** For endogenous CoIP  $5 \times 10^6$  HEL cells were lysed in 500  $\mu$ l IP-buffer (50 mM Tris pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.5% sodium deoxycholate, 5% glycerol and proteinase inhibitor cocktail) for 30 min at 4 °C on a rotating wheel. Lysates were centrifuged at 15,000g for 30 min and the supernatants were divided into 100- $\mu$ l aliquots. For CoIP, 200  $\mu$ l of IP-buffer was added to each 100- $\mu$ l aliquot. Three microgram of Tal1 antibody (Abcam, ab75738) or immunoglobulin (Ig)-G isotype control antibody (Santa-Cruz, sc-2027) and magnetic protein-G beads (Dynabeads protein-G beads, Life Technologies) were added. Upon incubation overnight at 4 °C on a rotating wheel, the beads were washed three times with IP-buffer and subjected to SDS-PAGE. Detection of immunoprecipitated Tal1 and coimmunoprecipitated PADI4 was performed by western blot using a PADI4 antibody (ab128086, Abcam 4H5, 1:2,000 dilution) and a Tal1 antibody (04–123, Millipore, BTL73, 1:500 dilution), respectively. For CoIP from HEK293 cells we used  $3 \times 10^4$  HEK293 cells transfected with metafectene (Biontex). CoIP was performed as detailed<sup>60</sup>. Briefly, transiently transfected

HEK293 cells were lysed in lysis buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 0.1% Triton X-100, 1 mM DTT, 1 mM EDTA, 10% glycerol and protease inhibitors). After centrifugation the supernatant was collected and incubated with 10  $\mu$ l of anti-FLAG agarose beads (M2, Sigma) at 4 °C for 4 h with rotation. The beads were washed four times with lysis buffer and the beads were resuspended in SDS-loading dye. The CoIP samples were analysed by western blot using an anti-HA antibody for detection of HA-PAD14 (sc-805, Santa-Cruz, 1:1,000) and anti-Tal1 antibody for the detection of Flag-Tal1 (sc-12984, Santa Cruz, 1:1,000 dilution). GST pull-down assays were performed as described<sup>64</sup>. Briefly, GST fusion proteins were expressed in *E. coli* (strain BL21). The GST protein was purified from lysates prepared in pull-down buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 1 mM EDTA, 10% glycerol and 0.01% Nonidet P-40) using glutathione-agarose beads (Sigma-Aldrich). <sup>35</sup>S-labelled proteins were produced by coupled *in vitro* transcription/translation in the presence of [<sup>35</sup>S]Met using the reticulocyte TNT-kit (Promega). The GST fusion protein was immobilized on glutathione-agarose beads and incubated with 10  $\mu$ l of TNT reaction mixture in 300  $\mu$ l of pull-down buffer at 4 °C for 1 h. The beads were washed four times using pull-down buffer. Upon elution of the proteins by boiling in SDS-loading-dye the samples were subjected to SDS-PAGE. Equal loading was assured by coomassie staining. Subsequently, the gel was dried, exposed to an X-ray film and analysed after 48 h.

**ChIP.** ChIP assays were performed in at least two independent experiments according to the Abcam-X-ChIP protocol with modifications<sup>60</sup>. For immunoprecipitation, the following antibodies were used: H3R2me2a (#07-585, Millipore, 4  $\mu$ g) and H3K4me3 (ab1012, Abcam, 4  $\mu$ g). H3K9Ac (ab10812, Abcam, 2  $\mu$ g), H4R3me2a (39705, Active Motif, 4  $\mu$ g), H3 (ab1791, Abcam, 3  $\mu$ g), H3R17me2a (ab8284, Abcam, 4  $\mu$ g), PRMT4/CARM1 (09-818, Millipore, 4  $\mu$ g), Tal1 (ab75739, Abcam, 5  $\mu$ g) and PAD14 (ab96758 in Figs 6 and 7 and ab128086 in Figs 4 and 5, Abcam, 5  $\mu$ g). The following IgGs were used as controls: IgG goat (Santa-Cruz, sc-2028), IgG mouse (Santa-Cruz, sc-2025) and IgG rabbit (Santa-Cruz, sc-2027). Purification of ChIP-DNA was performed using DNA purification columns (ChIP DNA Clean and Concentrator, Zymo Research, USA). ChIP-DNA was eluted with 40  $\mu$ l of buffer and analysed by quantitative PCR using 2  $\mu$ l of chromatin. Primer pairs are given in (Supplementary Table 2). DNA recovery calculated as percent of the input, error bars represent the s.d. from at least four determinations. All ChIP data were confirmed with at least two independent chromatin preparations. Histone modification ChIP values were corrected for nucleosome density using values gathered with a Histone 3 ChIP. For ChIP-ReChIP, we diluted 70  $\mu$ l of ChIP lysate with 210  $\mu$ l RIPA buffer without SDS (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM EDTA pH 8, 1% NP-40 and 0.5% sodium deoxycholate). In the first round of ChIP, two samples per antibody were prepared. Four micrograms was used for each antibody and 15  $\mu$ l of magnetic protein-G beads (Life Technologies). Incubation was performed overnight at 4 °C on a rotating wheel. Subsequently, the beads were washed four times with wash buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM EDTA pH 8 and 1% Triton X-100) and eluted in 50  $\mu$ l elution buffer-1 (50 mM Tris-HCl pH 7.5, 1% SDS and 15 mM DTT). At this stage the two samples were pooled and 700  $\mu$ l of incubation buffer was added (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP-40 and 0.25% sodium deoxycholate). For the second ChIP, 3  $\mu$ g of antibody was added and incubated overnight at 4 °C. Upon elution of the DNA with elution buffer-2 (1% SDS and 100 mM NaCO<sub>3</sub>) the sample was incubated with proteinase K, decross-linked at 65 °C and column purified as described for standard ChIP. *P*-values were calculated by an unpaired *t*-test (\**P*-value < 0.05, \*\**P*-value < 0.01 and \*\*\**P*-value < 0.001).

**Cell culture.** K562 and HEL cells were obtained from ATCC and maintained in RPMI medium supplemented with 10% fetal calf serum, 1% penicillin/streptomycin and 1% glutamine. Lentiviral packaging was performed using HEK293T cells. HEL cells were at a cell density of  $1.4 \times 10^6$  cells per well in a six-well plate. Transduction was performed with a multiplicity of infection of 100 and 8  $\mu$ g ml<sup>-1</sup> protamin sulphate. Human primary CD34<sup>+</sup> cells from granulocyte colony-stimulating factor (G-CSF)-mobilized apheresis samples from healthy volunteer donors, with written informed consent and approval by the Ethics Committee (permit #329-10), were immunomagnetically enriched according to the manufacturer's instructions (Miltenyi Biotec, Bergisch-Gladbach, Germany). Experiments were performed with CD34<sup>+</sup> cells from at least two independent donors. hCD34<sup>+</sup> cells were expanded for 3 days, transduced and sorted according to their green fluorescent protein signal. Subsequently, the transduced cells were seeded in MethoCult (Stem Cell Technologies, Grenoble, France). Colonies were evaluated 2 weeks after seeding, Supplementary Fig. 8. The CD34<sup>+</sup> cell isolation was controlled by flow cytometry using an anti-CD34 antibody (anti-CD34 pacific blue, Biolegend, Germany). Positive-selected cells were directly used or cryopreserved using CryoStor CS10 (Stem Cell Technologies) until further usage. Cells were expanded for 4 days at a cell density of 400,000–500,000 cells per millilitre in serum-free expansion medium (SFEM, Stem Cell Technologies) with cytokines (100 ng ml<sup>-1</sup> FLT3, 100 ng ml<sup>-1</sup> stem cell factor (SCF), 20 ng ml<sup>-1</sup> IL-3 and 20 ng ml<sup>-1</sup> IL-6). The cells were subjected to megakaryocytic (SFEM with 1 ng ml<sup>-1</sup> SCF, 30 ng ml<sup>-1</sup> thrombopoietin (TPO), 13.5 ng ml<sup>-1</sup> IL-9 and 7.5 ng ml<sup>-1</sup> IL-6) or erythrocytic differentiation (SFEM with 20 ng ml<sup>-1</sup> SCF, 5 ng ml<sup>-1</sup> IL-3, 2  $\mu$ M dexamethasone, 0.2  $\mu$ M estradiol and 1 U ml<sup>-1</sup>

erythropoietin (EPO)) for 6 days and used for mRNA analysis or ChIP. The efficiency of differentiation was controlled by flow cytometry using an CD41 or GYPA (GPA) antibody, respectively. Furthermore, quantitative reverse transcriptase-PCR with primers for beta-globin and CD41 were used to monitor erythrocytic and megakaryocytic differentiation. For monocytic differentiation, CD34<sup>+</sup> cells were grown in monocytic differentiation medium (iscove's modified dulbecco's medium (IMDM) with 20% fetal calf serum, 30 ng ml<sup>-1</sup> SCF, 30 ng ml<sup>-1</sup> FLT-3, 30 ng ml<sup>-1</sup> IL-3 and 30 ng ml<sup>-1</sup> macrophage colony-stimulating factor (M-CSF)). Monocytic cells were isolated by CD14 magnetic microbeads (Miltenyi Biotec) to a purity of 95%. The efficiency was controlled by flow cytometry using an anti-CD14 antibody (anti-CD14 APC, Miltenyi Biotec). See also Supplementary Fig. 9.

**Knockdown constructs.** ShRNA oligos against human PAD14 and human Tal1 were designed using the InvivoGen siRNA wizard program (<http://www.sirnawizard.com/design.php>) and cloned into a lentiviral backbone<sup>60</sup>. A non-specific shRNA was used as a control. Sequences targeted by shRNAs were: control-sh 5'-GACAAGAACTAAGCAAGAAT-3'; shPAD14 5'-GCGAAGACCTGCAG GACAT-3'; 5'-GCCAACCCAGAGCTGTGAAAGAT-3'; shTal1 5'-GACAAGAAG CTCAGCAAGAAT-3'; 5'-AAAGTTGTGCGCGGTATCTTC-3'.

**Cl-amidine.** The PAD14 inhibitor Cl-amidine was synthesized according to the slightly modified procedure described by Knuckley *et al.*<sup>65</sup>, or obtained commercially (Merck, Darmstadt, Germany). Step 1: (S)-2-benzamido-5-((*tert*-butoxycarbonyl)amino)pentanoic acid: (S)-2-amino-5-((*tert*-butoxycarbonyl)amino)pentanoic acid (0.93 g, 4.0 mmol) was dissolved in a biphasic mixture of aqueous 3.5 M NaOH (1.2 ml) and Et<sub>2</sub>O (4 ml). The mixture was cooled to 0 °C and alternating portions of benzoyl chloride (0.46 ml, 4.0 mmol) and aqueous 8.5 M NaOH (4.0 mmol, 0.46 ml) were added every 5 min for 30 min. The reaction mixture was allowed to warm to room temperature and stirred overnight. The reaction was acidified (pH 1) with concentrated HCl and extracted with DCM (3  $\times$  20 ml). The organics were combined, washed with brine, dried over MgSO<sub>4</sub> and concentrated to afford the product as a white powder (89% yield). Step 2: (S)-*tert*-butyl (5-amino-4-benzamido-5-oxopentyl)carbamate: (S)-2-benzamido-5-((*tert*-butoxycarbonyl)amino)pentanoic acid (2) (0.50 g, 1.5 mmol), HOTT (S-(1-oxido-2-pyridinyl)-1,1,3,3-tetra-methylthiuronium hexafluorophosphate) (0.82 g, 2.23 mmol) and DIPEA (*N,N*-diisopropylethylamine) (0.52 ml, 2.97 mmol) in DMF (*N,N*-dimethylformamide) (5 ml). After stirring for 10 min, ammonium chloride (0.159 g, 2.97 mmol) was added and stirring continued for 45 min. The reaction was partitioned between EtOAc (ethylacetate) (25 ml) and brine (60 ml). The organics were combined and washed with 2 N HCl (2  $\times$  12 ml), water (2  $\times$  12 ml), saturated NaHCO<sub>3</sub> (2  $\times$  12 ml), water (3  $\times$  12 ml) and brine 12 ml. The organics were then dried over MgSO<sub>4</sub> and concentrated to afford the product as a white powder (92% yield). Step 3: (S)-*N*-(1,5-diamino-1-oxopentan-2-yl)benzamide: (S)-*tert*-butyl (5-amino-4-benzamido-5-oxopentyl)carbamate (3) (0.2 g, 0.60 mmol) was dissolved in cold trifluoroacetic acid (TFA) (6 ml) and stirred for 0 °C for 45 min. After warming to room temperature, TFA was removed under reduced pressure and the residue dissolved in water (5 ml) and washed with Et<sub>2</sub>O (5 ml) before being lyophilized (96% yield). Step 4: Ethyl 2-chloroacetimidate hydrochloride: in a 100-ml three-neck round-bottomed flask was 2-chloroacetonitrile (4.2 ml, 66.1 mmol) dissolved in EtOH (4.25 ml, 72.8 mmol) and Et<sub>2</sub>O (50 ml) to give a colourless solution. The mixture was cooled in an acetone/ice bath to -10 °C and HCl (g) was bubbled through the reaction. The temperature rose to 0 °C and HCl (g) was continuously bubbled (1 h). A white precipitate is formed and collected by filtration and rinsed with cold ether. A white solid is obtained (77% yield). Step 5: (S)-*N*-(1-amino-5-(2-chloroacetimidamido)-1-oxopentan-2-yl)benzamide (Cl-amidine): (S)-*N*-(1,5-diamino-1-oxopentan-2-yl)benzamide-TFA salt (0.194 g, 0.630 mmol) and triethylamine (0.26 ml, 1.89 mmol) were dissolved in MeOH (2 ml). Ethyl 2-chloroacetimidate hydrochloride (0.2 g, 1.26 mmol) was added and the mixture stirred overnight. The reaction mixture was quenched with water (5 ml) and acidified with TFA. The product was isolated by preparative HPLC-MS using an isocratic method (18% MeOH, 0.1% formic acid and water) as a white powder (5% yield). See also Supplementary Fig. 10.

**Gene expression analysis.** Total RNA was isolated using the RNeasy mini kit (Qiagen, Hilden, Germany). Complementary DNA (cDNA) was synthesized using the Omniscript reverse transcriptase (Qiagen). Quantitative PCR was performed on a LightCycler 480 (Roche, Mannheim, Germany) using SYBR-Green PCR MasterMix (Eurogentec, Liege Belgium). Values were normalized against glyceraldehyde 3-phosphate dehydrogenase expression. Error bars represent the s.d. of at least four independent evaluations. Primer sequences are given in Supplementary Table 3.

**Microarray analysis and data analysis.** Gene expression array analysis of shTal1 and shPAD14 HEL cells was done with a whole-genome 4  $\times$  44K gene array (GPL13497, Agilent Technologies, Santa Clara, CA, USA). Array analysis was performed in triplicate. Microarrays were done using the 'Low RNA Input linear Amplification Kit Plus, One Color' protocol (Agilent Technologies) and the Agilent RNA Spike-In Kit for One color (Agilent Technologies) following the

manufacturer's standard protocol. Total RNA (200 ng) were used as a starting material to prepare cDNA. cDNA synthesis and *in vitro* transcription were performed according to the manufacturer's recommendation. Quantity and efficiency of the labelled amplified cDNA were determined using the NanoDrop ND-1000 UV-vis Spectrophotometer version 3.2.1. The hybridizations were performed for 17 h at 10 r.p.m. and 65 °C in the Hybridization Oven (Agilent Technologies). Washing and staining of the arrays were done according to the manufacturer's recommendation. Cy3 intensities were detected by one-colour scanning using an Agilent DNA microarray scanner. Scanned image files were visually inspected for artifacts and then analysed<sup>66</sup>.

Intensity data were extracted using Agilent's Feature Extraction software (version 11.5.1.1) including a quality control based on internal controls using Agilent's protocol GE1\_107\_Sep09. All chips passed the quality control and were analyzed using the Limma package<sup>67</sup> of Bioconductor<sup>68</sup>. Candidate genes were filtered to a minimum of twofold change and false discovery rate-corrected *P*-value < 0.05. For functional characterization of the differentially expressed transcripts, GO enrichment analysis was conducted via gseq accounting for gene length bias and correcting for multiple testing. Further functional association of candidate genes was performed with the webtool DAVID.

**Mutagenesis.** Mutagenesis of PADI4 was performed using site-directed mutagenesis. The inactive PADI4 double-mutant R374A/C645A was constructed according to published data<sup>49,69</sup> (Supplementary Fig. 2). Sequencing validated the integrity of wild-type and mutant PADI4. Expression and activity of mutant PADI4 was tested by enzymatic assay and western blot with transfected PADI4 and PADI4 mutant (Supplementary Fig. 11).

**Immunoblotting and FACS.** Immunoblot analysis was performed using anti-Tal1 (sc-12984, Santa-Cruz, 1:1,000 dilution and 04–123, Millipore, 1:500 dilution), anti-PADI4 (ab96758, 1:500 dilution and ab128086, 1:2,000 dilution, Abcam), anti-HA (sc-805, Santa-Cruz, 1:1,000 dilution), anti-Flag (F3165, 1:4,000 dilution, Sigma-Aldrich), anti-LaminB1 (ab16048, Abcam, 1:1,000 dilution) and anti-CTCF (ab70303, Abcam, 1:1,000 dilution). Western blots were analysed using X-ray film or an imaging system (Fusion FX7, PEQLAB, Erlangen, Germany). Uncropped images are provided in Supplementary Figs 12–15. FACS analysis of gp130 surface staining was performed using an APC-labelled gp130 antibody (R&D Systems, FAB228A).

## References

- Shivdasani, R. A., Mayer, E. L. & Orkin, S. H. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* **373**, 432–434 (1995).
- Robb, L. *et al.* Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proc. Natl Acad. Sci. USA* **92**, 7075–7079 (1995).
- Porcher, C., Liao, E. C., Fujiwara, Y., Zon, L. I. & Orkin, S. H. Specification of hematopoietic and vascular development by the bHLH transcription factor SCL without direct DNA binding. *Development* **126**, 4603–4615 (1999).
- Visvader, J. E., Fujiwara, Y. & Orkin, S. H. Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development. *Genes Dev.* **12**, 473–479 (1998).
- Courtial, N. *et al.* Tal1 regulates osteoclast differentiation through suppression of the master regulator of cell fusion DC-STAMP. *FASEB J.* **26**, 523–532 (2011).
- Dey, S., Curtis, D. J., Jane, S. M. & Brandt, S. J. The TAL1/SCL transcription factor regulates cell cycle progression and proliferation in differentiating murine bone marrow monocyte precursors. *Mol. Cell. Biol.* **30**, 2181–2192 (2010).
- Hall, M. A. *et al.* The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12. *Proc. Natl Acad. Sci. USA* **100**, 992–997 (2003).
- Hall, M. A. *et al.* Functional but abnormal adult erythropoiesis in the absence of the stem cell leukemia gene. *Mol. Cell. Biol.* **25**, 6355–6362 (2005).
- Elwood, N. J., Zogos, H., Pereira, D. S., Dick, J. E. & Begley, C. G. Enhanced megakaryocyte and erythroid development from normal human CD34(+) cells: consequence of enforced expression of SCL. *Blood* **91**, 3756–3765 (1998).
- Lausen, J. *et al.* Targets of the Tal1 transcription factor in erythrocytes: E2 ubiquitin conjugase regulation by Tal1. *J. Biol. Chem.* **285**, 5338–5346 (2010).
- Goardon, N. *et al.* ETO2 coordinates cellular proliferation and differentiation during erythropoiesis. *EMBO J.* **25**, 357–366 (2006).
- Lacombe, J. *et al.* Scl regulates the quiescence and the long-term competence of hematopoietic stem cells. *Blood* **115**, 792–803 (2010).
- Chen, Q. *et al.* The tal gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix-loop-helix protein. *EMBO J.* **9**, 415–424 (1990).
- Finger, L. R. *et al.* Involvement of the TCL5 gene on human chromosome 1 in T-cell leukemia and melanoma. *Proc. Natl Acad. Sci. USA* **86**, 5039–5043 (1989).
- Palomero, T. *et al.* Transcriptional regulatory networks downstream of TAL1/SCL in T-cell acute lymphoblastic leukemia. *Blood* **108**, 986–992 (2006).
- Larson, R. C. *et al.* Protein dimerization between Lmo2 (Rbtl2) and Tal1 alters thymocyte development and potentiates T cell tumorigenesis in transgenic mice. *EMBO J.* **15**, 1021–1027 (1996).
- Draheim, K. M. *et al.* A DNA-binding mutant of TAL1 cooperates with LMO2 to cause T cell leukemia in mice. *Oncogene* **30**, 1252–1260 (2011).
- Condorelli, G. L. *et al.* T-cell-directed TAL-1 expression induces T-cell malignancies in transgenic mice. *Cancer Res.* **56**, 5113–5119 (1996).
- Kelliher, M. A., Seldin, D. C. & Leder, P. Tal-1 induces T cell acute lymphoblastic leukemia accelerated by casein kinase IIalpha. *EMBO J.* **15**, 5160–5166 (1996).
- Tijssen, M. R. *et al.* Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Dev. Cell* **20**, 597–609 (2011).
- Wilson, N. K. *et al.* The transcriptional program controlled by the stem cell leukemia gene Scl/Tal1 during early embryonic hematopoietic development. *Blood* **113**, 5456–5465 (2009).
- Wadman, I. A. *et al.* The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J.* **16**, 3145–3157 (1997).
- Schuh, A. H. *et al.* ETO-2 associates with SCL in erythroid cells and megakaryocytes and provides repressor functions in erythropoiesis. *Mol. Cell. Biol.* **25**, 10235–10250 (2005).
- Palii, C. G. *et al.* Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages. *EMBO J.* **30**, 494–509 (2011).
- Huang, S., Qiu, Y., Stein, R. W. & Brandt, S. J. p300 functions as a transcriptional coactivator for the TAL1/SCL oncoprotein. *Oncogene* **18**, 4958–4967 (1999).
- Huang, S. & Brandt, S. J. mSin3A regulates murine erythroleukemia cell differentiation through association with the TAL1 (or SCL) transcription factor. *Mol. Cell. Biol.* **20**, 2248–2259 (2000).
- Huang, S., Qiu, Y., Shi, Y., Xu, Z. & Brandt, S. J. P/CAF-mediated acetylation regulates the function of the basic helix-loop-helix transcription factor TAL1/SCL. *EMBO J.* **19**, 6792–6803 (2000).
- Hu, X. *et al.* LSD1-mediated epigenetic modification is required for TAL1 function and hematopoiesis. *Proc. Natl Acad. Sci. USA* **106**, 10141–10146 (2009).
- Hagiwara, T., Nakashima, K., Hirano, H., Senshu, T. & Yamada, M. Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes. *Biochem. Biophys. Res. Commun.* **290**, 979–983 (2002).
- Cuthbert, G. L. *et al.* Histone deimination antagonizes arginine methylation. *Cell* **118**, 545–553 (2004).
- Wang, Y. *et al.* Human PAD4 regulates histone arginine methylation levels via demethylination. *Science* **306**, 279–283 (2004).
- Denis, H. *et al.* Functional connection between deimination and deacetylation of histones. *Mol. Cell. Biol.* **29**, 4982–4993 (2009).
- Li, P. *et al.* Regulation of p53 target gene expression by peptidylarginine deiminase 4. *Mol. Cell. Biol.* **28**, 4745–4758 (2008).
- Yao, H. *et al.* Histone Arg modifications and p53 regulate the expression of OKL38, a mediator of apoptosis. *J. Biol. Chem.* **283**, 20060–20068 (2008).
- Li, P. *et al.* Coordination of PAD4 and HDAC2 in the regulation of p53-target gene expression. *Oncogene* **29**, 3153–3162 (2010).
- Sharma, P. *et al.* Citrullination of histone H3 interferes with HP1-mediated transcriptional repression. *PLoS Genet.* **8**, e1002934 (2012).
- Tanikawa, C. *et al.* Regulation of histone modification and chromatin structure by the p53-PADI4 pathway. *Nat. Commun.* **3**, 676 (2012).
- Zhang, X. *et al.* Genome-wide analysis reveals PADI4 cooperates with Elk-1 to activate c-Fos expression in breast cancer cells. *PLoS Genet.* **7**, e1002112 (2011).
- Chang, X. & Han, J. Expression of peptidylarginine deiminase type 4 (PAD4) in various tumors. *Mol. Carcinog.* **45**, 183–196 (2006).
- Nakashima, K. *et al.* PAD4 regulates proliferation of multipotent haematopoietic cells by controlling c-myc expression. *Nat. Commun.* **4**, 1836 (2013).
- Christophorou, M. A. *et al.* Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature* **507**, 104–108 (2014).
- Slack, J. L., Causey, C. P. & Thompson, P. R. Protein arginine deiminase 4: a target for an epigenetic cancer therapy. *Cell. Mol. Life Sci.* **68**, 709–720 (2011).
- Vermeulen, M., Hubner, N. C. & Mann, M. High confidence determination of specific protein-protein interactions using quantitative mass spectrometry. *Curr. Opin. Biotechnol.* **19**, 331–337 (2008).
- de Boer, E. *et al.* Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc. Natl Acad. Sci. USA* **100**, 7480–7485 (2003).
- Torrano, V. *et al.* CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *J. Biol. Chem.* **280**, 28152–28161 (2005).

46. Tanimoto, K. *et al.* Human beta-globin locus control region HS5 contains CTCF- and developmental stage-dependent enhancer-blocking activity in erythroid cells. *Mol. Cell. Biol.* **23**, 8946–8952 (2003).
47. Farrell, C. M., West, A. G. & Felsenfeld, G. Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol. Cell. Biol.* **22**, 3820–3831 (2002).
48. Mahajan, M. C., Karmakar, S., Newburger, P. E., Krause, D. S. & Weissman, S. M. Dynamics of alpha-globin locus chromatin structure and gene expression during erythroid differentiation of human CD34(+) cells in culture. *Exp. Hematol.* **37**, 1143–1156 e3 (2009).
49. Arita, K. *et al.* Structural basis for Ca(2+) -induced activation of human PAD4. *Nat. Struct. Mol. Biol.* **11**, 777–783 (2004).
50. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
51. Audet, J., Miller, C. L., Rose-John, S., Piret, J. M. & Eaves, C. J. Distinct role of gp130 activation in promoting self-renewal divisions by mitogenically stimulated murine hematopoietic stem cells. *Proc. Natl Acad. Sci. USA* **98**, 1757–1762 (2001).
52. Jenkins, B. J., Roberts, A. W., Najdovska, M., Grail, D. & Ernst, M. The threshold of gp130-dependent STAT3 signaling is critical for normal regulation of hematopoiesis. *Blood* **105**, 3512–3520 (2005).
53. Sui, X. *et al.* Soluble interleukin-6 (IL-6) receptor with IL-6 stimulates megakaryopoiesis from human CD34(+) cells through glycoprotein (gp)130 signaling. *Blood* **93**, 2525–2532 (1999).
54. Jones, S. A., Scheller, J. & Rose-John, S. Therapeutic strategies for the clinical blockade of IL-6/gp130 signaling. *J. Clin. Invest.* **121**, 3375–3383 (2011).
55. Hong, D. S., Angelo, L. S. & Kurzrock, R. Interleukin-6 and its receptor in cancer: implications for translational therapeutics. *Cancer* **110**, 1911–1928 (2007).
56. Luo, Y. *et al.* Inhibitors and inactivators of protein arginine deiminase 4: functional and structural characterization. *Biochemistry* **45**, 11727–11736 (2006).
57. Bauer, U. M., Daujat, S., Nielsen, S. J., Nightingale, K. & Kouzarides, T. Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Rep.* **3**, 39–44 (2002).
58. Guccione, E. *et al.* Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature* **449**, 933–937 (2007).
59. Hyllus, D. *et al.* PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes Dev.* **21**, 3369–3380 (2007).
60. Herglotz, J. *et al.* Histone arginine methylation keeps RUNX1 target genes in an intermediate state. *Oncogene* **32**, 2565–2575 (2013).
61. Kirmizis, A. *et al.* Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. *Nature* **449**, 928–932 (2007).
62. Iberg, A. N. *et al.* Arginine methylation of the histone H3 tail impedes effector binding. *J. Biol. Chem.* **283**, 3006–3010 (2008).
63. Schreiber, E., Matthias, P., Muller, M. M. & Schaffner, W. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. *Nucleic Acids Res.* **17**, 6419 (1989).
64. Lausen, J., Cho, S., Liu, S. & Werner, M. H. The nuclear receptor co-repressor (N-CoR) utilizes repression domains I and III for interaction and co-repression with ETO. *J. Biol. Chem.* **279**, 49281–49288 (2004).
65. Knuckley, B., Causey, C. P., Pellechia, P. J., Cook, P. F. & Thompson, P. R. Haloacetamide-based inactivators of protein arginine deiminase 4 (PAD4): evidence that general acid catalysis promotes efficient inactivation. *Chembiochem* **11**, 161–165 (2010).
66. Opitz, L. *et al.* Impact of RNA degradation on gene expression profiling. *BMC Med. Genomics* **3**, 36 (2010).
67. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3** Article 3 (2004).
68. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
69. Arita, K. *et al.* Structural basis for histone N-terminal recognition by human peptidylarginine deiminase 4. *Proc. Natl Acad. Sci. USA* **103**, 5291–5296 (2006).

## Acknowledgements

We thank Yanming Wang (Pennsylvania State University) for providing a PADI4 expression vector. We are grateful to Ole Pless and Achim Leutz (MDC-Berlin) for sharing the BirA-Tal1 expression cell line. We thank Helge Hussong for excellent technical assistance. J.L. is supported by the Deutsche Forschungsgemeinschaft (SPP-1463, LA 1389/5-1) and institutional funds of the Georg-Speyer-Haus. The Georg-Speyer-Haus is funded jointly by the German Federal Ministry of Health (BMG) and the Ministry of Higher Education, Research and the Arts of the state of Hessen (HMWK). E.P. was supported by the Deutsche Forschungsgemeinschaft (DFG, Sachbeihilfe PR 1405/2-1, SFB 1039 A07) and Oncogenic Signaling Frankfurt (OSF). This work was supported by the Alfred und Angelika Gutermuth-Stiftung, the LOEWE Center for Cell and Gene Therapy Frankfurt (CGT, funded by the HMWK: III L 4-518/17.004, 2013) and the Thyssen Foundation.

## Author contributions

S.K. performed most of the experiments and analysed the data. O.N.K., J.H., I.B., N.K. and S.K.W. performed experiments; E.B. and E.P. synthesized small molecules and performed experiments. H.B. provided essential material, G.S.-R. performed array analysis, T.O., H.S. and M.K. performed mass spectrometry, S.K., O.N.K., J.H. and H.B. helped writing the paper. J.L. designed and performed experiments, analysed data, supervised the work and wrote the paper.

## Additional information

**Accession Codes:** Microarray data have been deposited in NCBI's Gene Expression Omnibus under accession number GSE54075.

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://www.nature.com/reprintsandpermissions/>

**How to cite this article:** Kolodziej, S. *et al.* PADI4 acts as a coactivator of Tal1 by counteracting repressive histone arginine methylation. *Nat. Commun.* 5:3995 doi: 10.1038/ncomms4995 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>