MCB
Journals.ASM.org

Genomic and Proteomic Analyses of Prdm5 Reveal Interactions with Insulator Binding Proteins in Embryonic Stem Cells

Giorgio Giacomo Galli,[a]* Matteo Carrara,[b] Chiara Francavilla,[c] Kristian Honnens de Lichtenberg,[a] Jesper Velgaard Olsen,[c] Raffaele Adolfo Calogero,[b] Anders Henrik Lund[a]

Biotech Research and Innovation Centre and Centre for Epigenetics, University of Copenhagen, Copenhagen, Denmark[a]; Molecular Biotechnology Center, Department of Clinical and Biological Sciences, University of Turin, Turin, Italy[b]; Department of Proteomics, NNF Center for Protein Research, Faculty of Health Science, University of Copenhagen, Copenhagen, Denmark[c]

PRDM proteins belong to the SET domain protein family, which is involved in the regulation of gene expression. Although few PRDM members possess histone methyltransferase activity, the molecular mechanisms by which the other members exert transcriptional regulation remain to be delineated. In this study, we find that Prdm5 is highly expressed in mouse embryonic stem (mES) cells and exploit this cellular system to characterize molecular functions of Prdm5. By combining proteomics and next-generation sequencing technologies, we identify Prdm5 interaction partners and genomic occupancy. We demonstrate that although Prdm5 is dispensable for mES cell maintenance, it directly targets genomic regions involved in early embryonic development and affects the expression of a subset of developmental regulators during cell differentiation. Importantly, Prdm5 interacts with Ctcf, cohesin, and TFIIIC and cooccupies genomic loci. In summary, our data indicate how Prdm5 modulates transcription by interacting with factors involved in genome organization in mouse embryonic stem cells.

PRDM proteins constitute a family of transcriptional regulators characterized by the presence of variable numbers of zinc finger domains and an N-terminal PR domain which shares similarity to the SET domain of histone methyltransferases (1, 2). Although few members of the family have been found to possess intrinsic histone methyltransferase activity (3–5), the PR domain, which has been demonstrated to serve as a protein interaction surface (6), and the variable number of zinc fingers provide plasticity to PRDM proteins in terms of molecular functions. PRDM proteins typically display tissue-specific patterns of expression (7, 8) and are often involved in the differentiation of specific cell lineages (1). In accordance with the important role of PRDM proteins in development, several members have been reported to be expressed in multi/pluripotent stem cells populations and regulate their differentiation. Indeed, PRDM14 was shown to be essential for the maintenance of human embryonic stem cells (9) and for the differentiation of mouse embryonic stem cells into extra-embryonic endoderm (10). PRDM16 is the master regulator of the skeletal muscle/brown fat switch (11), and it regulates oxidative stress genes in neural stem cells (12). Finally, Prdm3 is involved in hematopoietic stem cell maintenance and differentiation (13).

PRDM5 is a recently cloned member of the PRDM family (14), and most studies have focused on its promoter hypermethylation in cancer, suggesting a role in tumor suppression (15–17). The role of Prdm5 in development has been addressed in zebrafish, where Prdm5 performs an essential function during embryonic convergent extension movements through regulation of Wnt signaling (18). We recently characterized a gene trap knockout mouse allele of *Prdm5* and demonstrated that Prdm5 mutant mice are viable and fertile but display ossification defects due to Prdm5-dependent regulation of collagenous extracellular matrix genes (19). These data are in agreement with the identification of *PRDM5* mutations in patients suffering from brittle cornea syndrome, a connective tissue disease characterized by impaired extracellular matrix (20).

Mechanistically, PRDM5 appears not to be able to methylate histones (21); however, in different cellular contexts, it has been shown to act as a transcriptional repressor by recruiting the G9a histone methyltransferase and histone deacetylases (21) or as an activator by promoting elongating RNA polymerase II (Pol II) occupancy within transcriptional units or by binding enhancer-like elements (19).

Gene expression involves multiple levels of regulation in the nucleus. Aside from the recruitment of RNA polymerases and the general transcription machinery by sequence specific transcription factors, research in recent decades has proven the importance of epigenetic modifications and chromatin remodelling complexes to achieve proper gene regulation. Moreover, in the recent years, a strong body of evidence has demonstrated the pivotal role of genomic organization in terms of intra- and interchromosomal interactions and nuclear compartmentalization in transcriptional regulation (reviewed in references 22 and 23). A number of molecules have been shown to be involved in chromatin organization, such as laminin, which interacts with the genome in large transcriptionally repressed domains (24), and the zinc finger protein CTCF, which is considered to be the main insulator protein in mammals (25). Indeed, CTCF appears to have multiple functions, such as mediating inter- and intrachromosomal interactions to-

Received 6 May 2013 **Returned for modification** 6 June 2013
Accepted 6 September 2013

Published ahead of print 16 September 2013

Address correspondence to Anders Henrik Lund, anders.lund@bric.ku.dk.

* Present address: Giorgio Giacomo Galli, Stem Cell Program and Department of Hematology/Oncology, Children's Hospital, Boston, and Department of Stem Cell and Regenerative Biology, Harvard University, and Harvard Stem Cell Institute, Cambridge, Massachusetts, USA.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/MCB.00545-13.

doi:10.1128/MCB.00545-13

gether with cohesins (26), to juxtapose enhancer-promoter regions to mediate transcriptional regulation (27), or to act as an enhancer barrier to prevent spreading of epigenetic modification domains (28).

Recently, insulator function also has been assigned to the TFIIIC complex. This is a multiprotein complex that predominantly recognizes internal sequences in tRNA genes (called A and B boxes), and it is able to recruit the RNA polymerase III transcriptional machinery for tRNA gene transcription. In addition, TFIIIC also appears to bind in regions of the genome regardless of the presence of B boxes and RNA Pol III in general (called extra TFIIIC [ETC] sites) (29–32), and such regions appear to be cooccupied by CTCF and affect neighboring RNA polymerase II-driven transcription (29, 32). Importantly, tRNA genes have been shown to act as chromatin insulators, demonstrating the important role of TFIIIC binding to tRNA genes and of ETC sites in chromatin organization (33).

In this study, we demonstrate that Prdm5 is highly expressed in mouse embryonic stem cells and use this cellular system to gain insights into the molecular functions of Prdm5. Mass spectrometry and chromatin immunoprecipitation (ChIP)-seq-based analyses reveal that Prdm5 associates with complexes involved in chromatin organization, such as Ctcf, Smc1 (cohesin), and TFIIIC. Indeed, Prdm5 acts as a neighboring factor for TFIIIC at both ETC sites and RNA polymerase III-driven tRNA genes, while the binding sites of Prdm5, Ctcf, and cohesin overlap. Interestingly, common regions presenting the Prdm5-specific consensus sequence are often devoid of Ctcf consensus sites, suggesting that Prdm5 affects Ctcf and cohesin occupancy at common sites.

Collectively, our data suggest that Prdm5 is involved in chromatin organization in mouse embryonic stem cells via interactions with CTCF and TFIIIC complexes.

MATERIALS AND METHODS

Feeders and mouse embryonic stem cell (mES) generation. All animal experimentation was performed with approval from Danish authorities (protocol number 2006/562-43; Dyreforsøgstilsynet). Mouse embryo fibroblasts (MEFs) were isolated by mechanical disruption of embryonic day 13.5 (E13.5) embryos followed by trypsin digestion. Feeder cells were MEFs mitotically inactivated by treatment with mitomycin C (Sigma) for 2 h at 10 μg/ml.

Wild-type and *Prdm5*$^{LacZ/LacZ}$ mouse embryonic stem cells were isolated essentially as previously described (34). Briefly, *Prdm5*$^{+/LacZ}$ females were superovulated by injection of pregnant mouse serum gonadotrophin (PMSG) and chorionic gonadotrophin (hCG) and mated to *Prdm5*$^{+/LacZ}$ males. Females were sacrificed 3.5 days after the observation of the vaginal plug, their uteruses were isolated, and blastocysts were flushed out to be cultivated in dishes containing a layer of feeder cells growing in knockout Dulbecco's modified Eagle medium (DMEM) supplemented with 20% knockout serum replacement, 1% penicillin-streptomycin, 2 mM L-glutamine, 1× MEM-nonessential amino acids, 100 mM β-mercaptoethanol, and recombinant mouse leukemia inhibitory factor (LIF) (1,000 U/ml of ESGRO; Chemicon).

After zona pellucida hatching, inner cell masses were isolated and disrupted by trypsinization with 10× trypsin solution (Sigma) and subsequently cultivated in the same medium supplemented with 20% fetal bovine serum (FBS) instead of knockout serum replacement on a layer of mouse embryonic fibroblasts as feeders.

The E14 mES cell line was maintained in Glasgow minimal essential medium (MEM) (Sigma) supplemented with 10% FBS, 1% penicillin-streptomycin, 2 mM L-glutamine, 1× MEM-nonessential amino acids, 1× sodium pyruvate, 100 mM β-mercaptoethanol, and recombinant mouse LIF.

mES cell characterization (genotyping, karyotyping, and proliferation). DNA for genotyping was isolated by lysing cells at 55°C in buffer containing 20 mM Tris-Cl, pH 8.0, 5 mM EDTA, 400 mM NaCl, 1% SDS, and 400 μg/ml proteinase K and subsequent phenol-chloroform isolation and isopropanol precipitation according to standard procedures. Primers used for genotyping were PRDM5_geno_rev (5′-AGTTTGAACAGGGTACCATCCAT-3′), PRDM5_geno_wt_fwd (5′-5-TATGGCCCTTTGATGTCTTCACTA-3′), PRDM5_geno_ko_fwd (5′-5-CGCCTTCTTGACGAGTTCTTCTG-3′), Xist_Fwd (5′-5-GCTTTGTTCACTTTCTCTGG-3′), Xist_Rev (5′-5-ATTCGGACCATGGGA-3′), Sry_Fwd (5′-5-GCATATGGTGTGGTCC-3′), and Sry_Rev (5′-5-CCAGTCTTGCCTGTATGTGA-3′).

Karyotyping was performed by treatment of mES cells with Colcemid (Sigma) for 2 h, followed by hypotonic treatment with 0.075 M KCl for 25 min at 37°C. Chromosome preparations were fixed with Carnoid solution and spread on glass slides for staining with Giemsa stain. Images from at least 20 spreads were acquired, and chromosome numbers were counted manually.

Proliferation indexes were obtained by measuring bromodeoxyuridine (BrdU) incorporation. Briefly, cells were treated with 33 μM BrdU solution for 30 min, trypsinized, and fixed in 70% ethanol (EtOH). DNA was denatured by treatment with 2 M HCl, and cells were incubated with anti-BrdU antibody (Beckton Dickinson). Detection was achieved by incubation with secondary Alexa Fluor 488-conjugated antibody (Invitrogen). Stained cells were analyzed on a FACSCalibur machine (Beckton Dickinson).

mES cell differentiation assays. Embryoid bodies from E14 cells were generated by growing in hanging drops according to standard procedures and harvesting at the indicated times. Cell clones from the wild-type or *Prdm5*$^{LacZ/LacZ}$ genotype were allowed to grow feeder free on gelatin-coated dishes and then seeded in 6-cm plates for differentiation assays 24 h previous to treatment. Cells treated for the so-called −LIF condition were grown in complete ES cell medium (described above) without LIF, while cells for the so-called −LIF/+ATRA condition were also treated with 1 μM all-*trans* retinoic acid (ATRA) with medium changed every 12 h. After 60 h of treatment, cells were harvested for gene expression analyses.

Expression analyses. Cells were harvested and lysed in radioimmunoprecipitation assay (RIPA) buffer and subjected to standard SDS-PAGE. Antibodies used are the following: polyclonal PRDM5 (19), monoclonal PRDM5 (described below), tubulin (ab11304; Abcam), Oct4 (ab19857; Abcam), Smc1, TFIIIC63, and TFIIIC220 (A300-055A, A301-242A, and A301-293A, respectively; Bethyl Laboratories), Ctcf (07-729; Millipore), vinculin (V9131; Sigma), and hemagglutinin (HA; PRB-101C; Biosite).

Total RNA was extracted from cell pellets using the RNeasy plus mini-kit (Qiagen) according to the manufacturer's instructions. cDNA was synthetized using TaqMan reverse transcription reagents (Applied Biosystems), and quantitative reverse transcription-PCR (qRT-PCR) was performed with the One Step plus sequence detection system (Applied Biosystems) using Fast SYBR green master mix reagent (Applied Biosystems).

Generation of Prdm5 mouse monoclonal antibody. The PRDM5 antigen fragment (corresponding to amino acids [aa] 1 to 142 of human PRDM5) was amplified by PCR and cloned into pGEX-6P-1 (GE Healthcare). It was expressed in BL21 cells (Invitrogen) by induction with 0.5 mmol/liter isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 4 h and subsequently purified with glutathione-Sepharose beads. The purified antigen was injected subcutaneously into BALB/c mice, and after 6 months hybridomas were made according to standard procedures. Positive clones were selected by screening the hybridoma libraries by enzyme-linked immunosorbent assay (ELISA) with plates coated with purified glutathione *S*-transferase (GST) or GST-PRDM5(1-142). The specificity of clone 57-20 was verified by immunoblotting samples from wild-type and *Prdm5*$^{LacZ/LacZ}$ MEFs (data not shown).

Protein interaction analyses. PRDM5 complexes were purified for mass spectrometry analysis by single-step affinity purification on HA beads. Briefly, control cells and HA-PRDM5 cells were harvested in NP-40 buffer (0.1% Igepal, 10 mM Tris-Cl, pH 7.5, 150 mM NaCl, 3 mM $MgCl_2$, 1 mM $CaCl_2$, and 250 mM sucrose) containing protease and phosphatase inhibitor cocktails. DNA was digested by incubation with 2.5 U/mg of micrococcal nuclease, and lysates were cleared by ultracentrifugation. Equal amounts of triplicate lysates were incubated at 4°C with HA-agarose beads (Sigma) overnight. After washes, immunocomplexes were eluted from beads by incubation with HA peptide (Sigma). Samples were treated with dithiothreitol (DTT; Sigma) and cloroacetamide (CAA; Sigma) to alkylate cysteines before being run on SDS-PAGE. Separated proteins were fixed in the gel and visualized with colloidal Coomassie staining (Invitrogen).

For coimmunoprecipitation (co-IP) experiments, cells were lysed in ELB buffer (150 mM NaCl, 50 mM HEPES, 0.1% Igepal, and 100 μg of ethidium bromide, where indicated) and incubated with HA agarose beads or beads coated with specific antibodies overnight. After washes, proteins were recovered by boiling beads in SDS Laemmli buffer. Gel filtration experiments were performed with a Superose 6 column (GE Healthcare) on an Akta fast protein liquid chromatography (FPLC) system according to the manufacturer's recommendations.

MS analysis. Digestion of proteins from the gel was performed as described previously (35). Briefly, each gel lane was separated into four slices that were destained with 50% EtOH in acetonitrile and dehydrated with EtOH. Proteins were digested with sequence-grade trypsin (Sigma) overnight. Trypsin activity was quenched by acidification with trifluoroacetic acid (pH 2.7), and peptides were extracted from the gel plugs with increasing concentrations of acetonitrile. Organic solvent was evaporated in a vacuum centrifuge, and peptides were purified on STAGE tips with two C_{18} filters. Eluted peptides were analyzed by on-line nanoflow liquid chromatography-tandem MS (LC-MS/MS). Peptide separation was performed by reverse-phase C_{18} high-performance liquid chromatography (HPLC) on an Easy nLC system (Thermo Fisher Scientific) loading sample with a constant flow onto 15-cm-long in-house-packed columns and eluting peptides using a 120-min gradient of increasing (3 to 70%) buffer B (80% acetonitrile, 0.5% acetic acid) at a constant flow. The effluent was directly electrosprayed into a Q-Exactive mass spectrometer (Thermo Fisher Scientific) through a nanospray ion source. The spray voltage was set to 2.1 kV, sheet and auxiliary gas to 0, and the temperature of the heated capillary to 275°C. The peptide mixture was analyzed by full-scan MS spectra (*m/z* from 200 to 1750; resolution at *m/z* 200, 70,000) in an Orbitrap analyzer after accumulation of ions at a 1e6 target value based on predictive automatic gain control (AGC) from the previous scan. Q-Exactive was operated in data-dependent mode to automatically switch between MS and MS/MS acquisition. For every full scan, the eight most intense ions were isolated and fragmented (collision energy, 25%) by high-energy collisional dissociation (HCD) for a fixed full time at 250 ms and 70,000 resolution. Finally, the dynamic exclusion list was restricted to a maximum period of 30 s (36).

The acquired data were processed with MaxQuant software (version 1.2.6.13; Max Plank Institute of Biochemistry, Department of Proteomics and Signal Transduction, Munich, Germany) (37), in which peptides and proteins are identified by the Andromeda search algorithm via matching all MS and MS/MS spectra against a target/decoy version of the mouse Uniprot complete proteome database supplemented with reversed copies of all sequences and frequently observed contaminants, like human keratin and albumin. The HA tag version of human PRDM5 sequence was added to the database. The false discovery rate (FDR) was set to 0.01 for both peptides and proteins. Carbamidomethylated cysteines were set as fixed and oxidation of methionine and N-terminal acetylation as variable modifications. The minimal peptide length was 6 amino acids.

Label-free quantification was performed with MaxQuant as described previously (38, 39). Briefly, protein intensities were normalized and proteins were quantified between control and case experiments by the MaxQuant label-free algorithm, resulting in LFQ (label-free quantitation) protein intensities. The downstream analysis was performed with Excel (Microsoft) and Perseus (Max Planck Institute of Biochemistry, Department of Proteomics and Signal Transduction, Munich, Germany) software. The triplicates of each bait IP were analyzed against the three control IPs. Protein identifications were filtered for contaminants and reverse hits. A minimum of three peptide identifications, with at least two being uniquely assigned to the particular protein, were required. Significant interactors were determined by a volcano plot-based strategy, combining *t* test *P* value and ratio information, as described previously (39). *t* test-based comparisons of bait IPs versus control IPs were performed with a *P* value threshold set at 0.05 and the bend-of-the-curve value set at 1. LFQ protein intensity ratios of bait relative to the control was plotted against the negative logarithmic *P* value of the *t* test, as was a stipulated line representing the permutation-based false discovery rate separating specific from nonspecific binders. Significant interactors of the bait proteins were color coded.

ChIP and high-throughput sequencing. For chromatin immunoprecipitation assays (ChIP), cells were fixed in 1% formaldehyde for 10 min and cross-linking was quenched by addition of 0.125 M glycine. Cells were then harvested in ChIP buffer and sonicated with a Bioruptor UCD-300 (Diagenode) to obtain fragments around 250 bp in size. Immunoprecipitations were performed with the following antibodies: polyclonal PRDM5 (19), Ctcf (07-729; Millipore), Smc1 (A300-055A; Bethyl Laboratories), TFIIIC63 (A301-242A; Bethyl Laboratories), and normal rabbit IgG (Sigma). Precipitated DNA was recovered by purification with a QIAquick PCR purification kit (Qiagen). For ChIP-quantitative PCR (qPCR), real-time PCRs were performed with the One Step plus sequence detection system (Applied Biosystems) using Fast SYBR green master mix reagent (Applied Biosystems). Libraries for sequencing were obtained using a ChIP-seq DNA sample prep kit (Illumina) according to the manufacturer's recommendations, and samples were sequenced on a Genome Analyzer II sequencer (Illumina).

For RNA-seq, double-stranded cDNA was transcribed by using a SuperScript double-stranded cDNA synthesis kit (Invitrogen), and libraries were obtained using the TruSeq RNA sample prep kit v2 (Illumina). All of the experiments involving high-throughput sequencing were run on Hiseq2000 sequences at the Danish national high-throughput sequencing center.

Bioinformatic analyses. (i) RNA sequencing. The quality of the RNA-seq data was evaluated with fastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/) software. RNA-seq data were analyzed using a pipeline based on the use of TopHat (40) and CuffLinks (41) essentially as previously described (42). The reference genome used for alignment was mm9 and uses UCSC annotation (genome.ucsc.edu). Differentially expressed genes were detected using the Bioconductor package (43) cummerBund.

(ii) ChIP-seq. ChIP-seq data were mapped over the assembly mm9, and peak finding and filtering were performed as previously described (19). A total of 25.7×10^6 reads were mapped for Prdm5-Ab1 in wild-type cells, 25.1×10^6 reads for Prdm5-Ab2 in wild-type cells, 27.8×10^6 reads for Prdm5-Ab1 in $Prdm5^{LacZ/LacZ}$ cells, and 21.5×10^6 reads for Prdm5-Ab2 in $Prdm5^{LacZ/LacZ}$ cells. For each antibody, the data from $Prdm5^{LacZ/LacZ}$ cells was used as the background signal to define peaks, obtaining 4,161 peaks detected by Prdm5-Ab1 and 2,019 detected by Prdm5-Ab2. A total of 1,490 peaks shared overlaps between the two antibodies and consequently were retained for further analysis (see Table S3 in the supplemental material). The Prdm5 consensus matrix was calculated as previously described (19).

(iii) Ctcf colocalization. Publicly available ChIP-seq peak data of Ctcf sequencing in mouse stem cells from the ENCODE/LICR project were downloaded (44). The annotation was lifted from the mouse reference genome assembly mm8 to the assembly mm9 using a UCSC online tool (http://genome.ucsc.edu/cgi-bin/hgLiftOver) and intersected with Smc1 and Smc3 tag coverage signal (45) using SeqMiner software (46). A total of

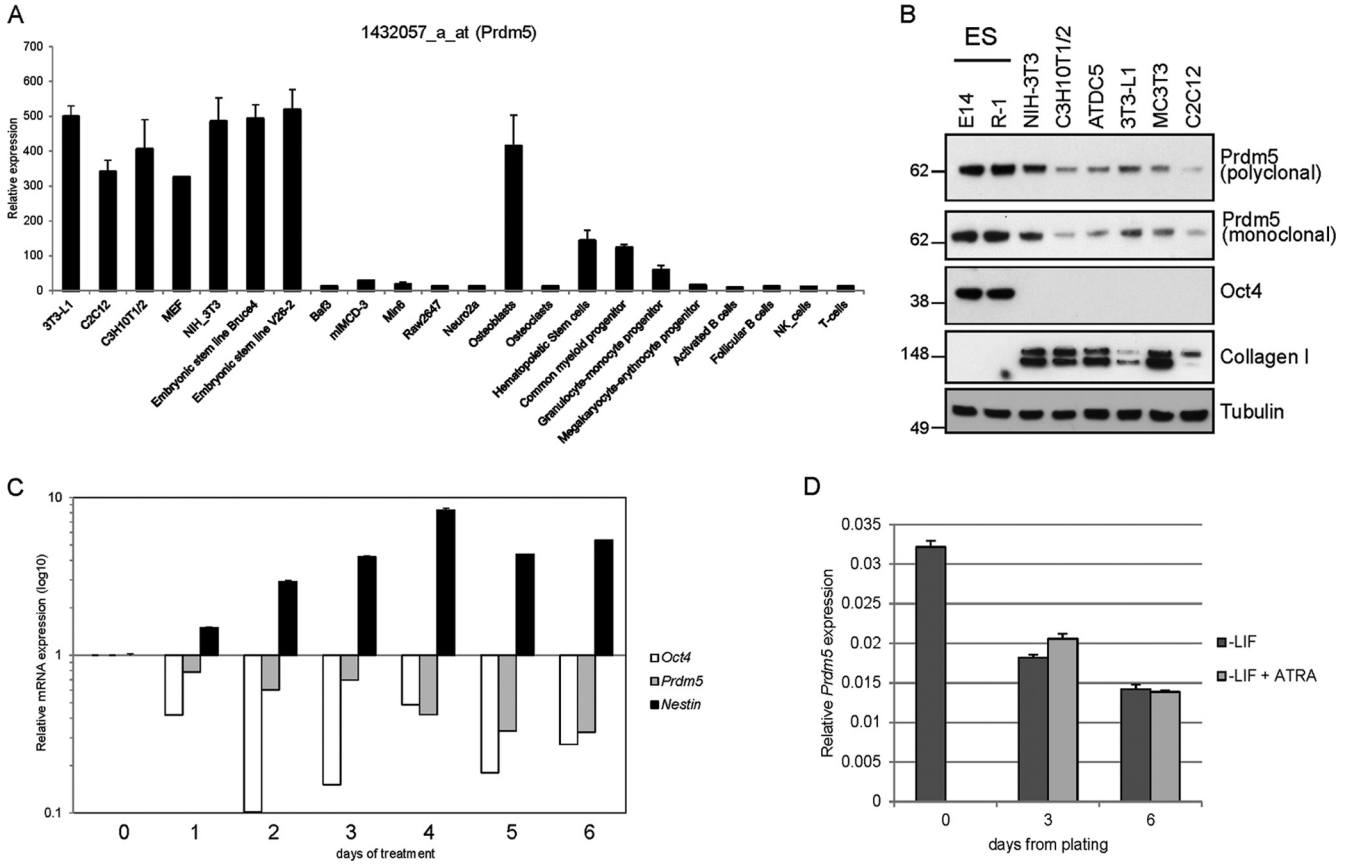

FIG 1 Prdm5 is highly expressed in mES cells and is downregulated upon differentiation. (A) Analysis of Prdm5 expression from a Biogps.org microarray data set on a panel of mouse cell types. (B) Western blot analysis with the indicated antibodies on a panel of cell lines, including two mES cell lines and embryo fibroblastic cells. Oct4 is used as a stemness marker and type I collagen as a fibroblast marker. Tubulin is used for equal protein loading. (C) qPCR analysis for *Prdm5*, *Oct4*, and *Nestin* in E14 mES cells treated with ATRA. Error bars represent standard deviations from duplicate experiments. Data are represented on a $\log_{10}$ (fold change) scale. (D) Representative qPCR analysis for Prdm5 in E14 cells grown as embryoid bodies for the indicated number of days with the addition of ATRA. Error bars represents standard deviations from triplicate technical replicates.

47,523 peaks were available afterwards. Ctcf position weight matrix (47) was used alongside the Prdm5 matrix to search for putative binding sites inside the available peak data using the Biostrings package from Bioconductor and a similarity cutoff set to 80%.

Functional annotation of genomic regions according to developmental stage of expression was performed using the GREAT database (48), while pathway annotation of Prdm5 target genes and deregulated genes was performed with Ingenuity pathway analysis (IPA). For the classification of TFIIIC sites we used the previously published annotation (31). Heat maps for signal comparison around Prdm5 or TFIIIC peaks were generated by Seqminer (46).

Microarray data accession number. The microarray data have been deposited in the GEO database, and the accession number is available on request.

RESULTS

Prdm5 is expressed in mouse embryonic stem cells. In order to investigate the molecular functions of Prdm5 in a relevant cellular system, we analyzed its expression pattern in a variety of mouse cell lines. We first interrogated publicly available microarray data sets and observed an interesting pattern of expression, with *Prdm5* being highly expressed in mouse embryonic stem (mES) cells, fibroblastic cell lines from mouse embryos, and osteoblasts (Fig. 1A). By Western blot analysis we detected substantial Prdm5 expression in a panel of mouse embryonic stem cells (characterized by Oct4 expression) and fibroblast cell lines (expressing type I collagen), with the highest expression in two mES cell lines (Fig. 1B), suggesting a role for Prdm5 in mES cells. To further corroborate this hypothesis, we evaluated the expression of *Prdm5* in a retinoic acid (ATRA)-induced differentiation model. As expected, ATRA treatment of the E14 mES cell line decreased the expression of the Oct4 stemness marker and induced the differentiation marker Nestin (Fig. 1C). Interestingly, during a 6-day experiment, we observed a progressive decrease in *Prdm5* transcript levels in cells treated with ATRA relative to mock-treated cells (Fig. 1C). Moreover, differentiation of mES cells by embryoid body formation confirmed that the downregulation of *Prdm5* upon differentiation is not specific to retinoic acid treatment (Fig. 1D). These findings are supported by publicly available data sets, which confirmed *Prdm5* to be downregulated in mES cells upon differentiation induced by embryoid body formation or Oct4 knockdown by short interfering RNA (siRNA) (available as GDS2905, GDS2666, and GDS1824 NCBI GEO data sets).

In summary, we demonstrate that Prdm5 is expressed in mouse embryonic stem cells, and it is downregulated upon cellular differentiation.

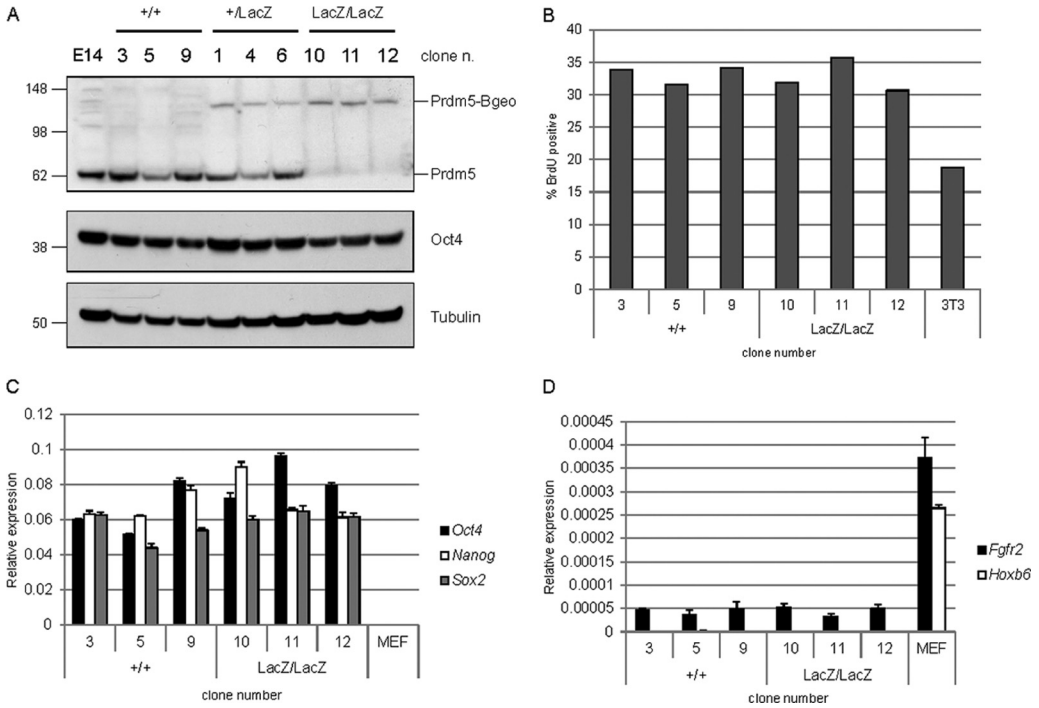

FIG 2 Generation of *Prdm5*$^{LacZ/LacZ}$ mouse embryonic stem cells. (A) Western blot analysis of primary mouse embryonic stem cell lines generated for this study. Oct4 is used as a stemness marker and tubulin for equal protein loading. (B) Percentage of cells incorporating BrdU as measured by flow cytometry analysis of mES cell lines used in this study. Representative experiment showing individual clones. NIH 3T3 cells are included as a reference. (C) qPCR analysis of mES cell lines for stemness markers *Oct4*, *Nanog*, and *Sox2*. MEFs are used as negative controls representing differentiated cells. (D) qPCR analysis of mES cell lines for differentiation markers *Fgfr2* and *HoxB6*. MEFs are used as negative controls representing differentiated cells.

Prdm5 is dispensable for mES cell stemness. Given the high expression level of Prdm5 in mES cells, we decided to analyze its molecular functions in this cellular system. To circumvent the issues of off-target effects from RNA interference (RNAi) loss-of-function experiments, mouse embryonic stem cells from the *Prdm5*$^{LacZ/LacZ}$ mouse strain were generated. Three clones for each genotype were selected for further studies. The genotypes were confirmed by standard PCR (data not shown) and by Western blotting, where the expression of the mutant allele could be appreciated by the appearance of a 135-kDa band resulting from the fusion of the product of the first 2 exons from the *Prdm5* gene and the β-geo gene trap cassette (Fig. 2A). Cells from all genotypes retained characteristic ES cell morphology (see Fig. S1A in the supplemental material) and displayed normal karyotypes (see Fig. S1B). Moreover, Prdm5 loss in ES cells did not alter the cell cycle distribution (data not shown) or proliferation index, as measured by BrdU incorporation (Fig. 2B), compared to control cells. Importantly, Prdm5 mutant mES cells expressed stemness markers such as *Oct4*, *Nanog*, and *Sox2* at levels comparable to those of control cells (Fig. 2A and C), and cells of all genotypes expressed low levels of *Fgfr2* and did not express the differentiation marker *HoxB6* compared to mouse embryonic fibroblasts (Fig. 2D).

These results suggest that Prdm5 is dispensable for the stemness of mES cells.

Prdm5 mutant embryonic stem cells display deregulation of a set of developmental regulators following differentiation. A number of chromatin modifiers have been reported to be dispensable for mES cell maintenance but to play a role in gene regulation during cellular differentiation (49, 50). Along the same lines, we sought to characterize genes regulated by Prdm5 during the differentiation of mouse embryonic stem cells.

We performed this in two differentiation models. In the first one (−LIF), we deprived the cells of the cytokine LIF in order to detect transcriptional changes imposed by Prdm5 during early stages of spontaneous differentiation; in the second model (−LIF/+ATRA), other than withdrawing LIF, we treated the cells with ATRA to evaluate the results at a later differentiation stage following a defined differentiation cue. Cells were treated for 60 h, and RNA was isolated for expression analysis of stemness and differentiation markers. In agreement with the literature, in the two models we observed an increase in the expression of the differentiation markers *Nestin* and *T* (*Brachyury*) (see Fig. S2A in the supplemental material) and a decrease of the stemness marker *Oct4* (Fig. 3A) other than the Prdm5 protein level itself (Fig. 3A). Importantly, no differences in activation/repression of main lineage or stemness markers could be detected between wild-type and *Prdm5*$^{LacZ/LacZ}$ mES cells (see Fig. S2A; also data not shown), further indicating a nonessential role for Prdm5 in the early steps of ES cell differentiation. To broaden the analyses, we sought to identify genes regulated by Prdm5 during cellular differentiation by RNA sequencing. Data from the three cell clones of each genotype were treated as biological replicates in order to obtain deregulated genes strictly dependent upon Prdm5 loss. Differential expression analyses revealed that a relatively low number of deregulated genes reached statistical significance (FDR, <0.1; $n = 964$ for −LIF and $n = 305$ for −LIF/+ATRA). We subsequently applied a fold change cutoff of 1.3 in order to restrict our analyses to more robustly regulated genes ($n = 481$ for −LIF and $n = 136$

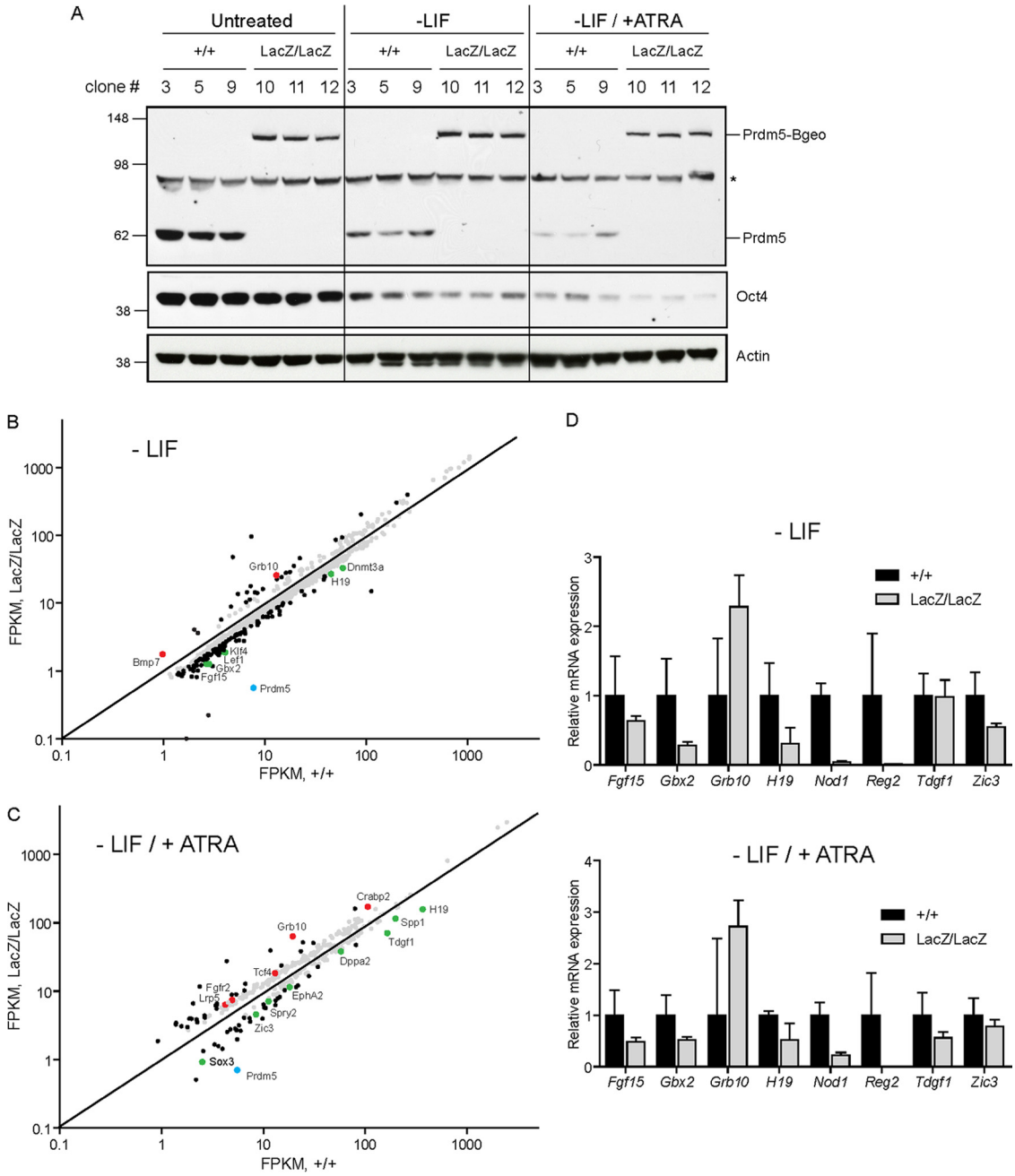

FIG 3 Developmental regulators are differentially expressed in *Prdm5*$^{LacZ/LacZ}$ cells upon differentiation. (A) Western blot analysis from ES differentiation experiments used for RNA-seq experiments. An unspecific band (asterisk) and actin were used for equal protein loading. (B and C) Scatterplot of expression data from RNA-seq experiments in cells deprived of LIF (B) and deprived of LIF but treated with ATRA (C). Each point corresponds to one RefSeq transcript that showed differential expression with fold changes of >1.3 (in gray) and >1.5 (in black). The Prdm5 point is in blue, while selected down- and upregulated genes involved in embryonic development are highlighted in green and red, respectively. (D) qPCR analyses of validation samples from the RNA-seq experiment in −LIF and −LIF/+ATRA conditions. Error bars represent standard deviations from assaying the three cell clones used in the study.

for −LIF/+ATRA) and performed gene ontology annotation on the lists of differentially expressed genes (see Fig. S2B). We observed enrichment for genes involved in early embryo development and, in particular, for the differentiation of specific tissues, such as those from the nervous and skeletal systems (see Fig. S2B), with the latter being in agreement with our previous results (19). Interestingly, upon LIF withdrawal we observed regulated genes that were mainly involved in embryonic development, while in cells treated with retinoic acid we observed a number of genes related to cell-to-cell signaling (see Fig. S2B).

Detailed analysis of differentially expressed genes in cells withdrawn from LIF revealed that only 14% of the genes that passed statistical significance were actually downregulated more than 1.5-fold (data not shown), indicating that Prdm5 loss predominantly induced small but consistent changes in the transcriptome. The large majority of deregulated genes were downregulated in

Prdm5$^{LacZ/LacZ}$ cells, and among the most regulated genes we observed genes coding for developmental regulators, such as *Klf4*, *Dnmt3a*, and *Lef1* (Fig. 3B). When we analyzed the results from cells treated with ATRA, although the number of genes differentially expressed was significantly lower, we observed stronger regulation, particularly for a number of developmental regulators involved in main signaling pathways, such as *Tdgf1* (also known as *Cripto*), *EphA2*, and *Fgfr2* (Fig. 3C and A). Complete lists of deregulated genes can be found in Table S1 and S2 in the supplemental material. To determine the reliability of our RNA-seq data, a random set of genes was validated by qPCR on an independent set of RNAs (Fig. 3D).

In summary, Prdm5 loss in mES cells induces the deregulation of a subset of developmental regulators and signaling genes during cellular differentiation.

Genome-wide analysis of Prdm5 occupancy reveals binding to genes encoding developmental regulators. In order to identify direct Prdm5 targets, we sought to map its genome-wide occupancy in mES cells by ChIP-seq. To identify Prdm5 binding sites in ES cells with high confidence, we included three conditions in the experimental design: (i) mES cells were grown on *Prdm5*$^{LacZ/LacZ}$ feeders to avoid contamination from Prdm5 expressed in supporting fibroblasts; (ii) *Prdm5*$^{LacZ/LacZ}$ mES cells were used in ChIP-seq analysis as negative background samples; and (iii) two antibodies displaying a different set of epitopes against Prdm5 were used for the analysis (see the scheme in Fig. S3A in the supplemental material).

We first tested the validity of our strategy by performing ChIP-qPCR using the previously identified *Pdgfra* transcription start site (TSS) region (19) as a positive control for a Prdm5-bound region. Indeed, when we performed ChIP in wild-type cells we could observe specific Prdm5 enrichment with both antibodies on *Pdgfra* TSS, while this signal was reduced to background levels in mutant cells (background was measured both by ChIP with unrelated IgG and by ChIP-qPCR on a negative region) (see Fig. S3B in the supplemental material).

Approximately 34 million reads were obtained from each sequencing reaction with an average mapping percentage of 73% (data not shown), and peak finding using *Prdm5*$^{LacZ/LacZ}$ samples as the background sample for each antibody resulted in 4,161 and 2,018 peaks for the two Prdm5 antibodies (Fig. 4A). Overlay of the ChIP-seq data for the two antibodies resulted in 1,490 peaks, which we defined as high-confidence Prdm5 binding sites (Fig. 4A; also see Table S3 in the supplemental material). Finally, when we analyzed the signal distribution for Prdm5 in a 20-kb region around targeted TSSs, we observed enrichment for Prdm5 signal in wild-type cells that was lost in mutant cells (see Fig. S3C in the supplemental material), demonstrating the efficacy of our experimental design.

Analysis of sequences underlying peak centers with a *de novo* motif-finding algorithm identified a Prdm5 consensus sequence that shares similarity to the one previously identified in MC3T3 cells (19) (Fig. 4B), indicating that the Prdm5 binding motif is conserved across different cell types.

Annotation of Prdm5 peaks relative to the nearest transcriptional unit showed that Prdm5 binding was evenly distributed among the TSSs of genes, gene bodies, or intergenic regions (Fig. 4C). This suggests that Prdm5 exerts other functions than the canonical transcription factor acting on the proximal promoters of target genes. Interestingly, this binding pattern differs from the one previously identified in preosteoblastic cells, suggesting that Prdm5 has different functions in different cell types (19).

Functional annotation of genes assigned to Prdm5-bound regions displayed enrichment for categories such as embryonic development and organ morphology, further supporting a role for Prdm5 in differentiation processes (Fig. 4D). Indeed, when we performed validation of peaks identified by ChIP-seq, we confirmed Prdm5 binding to regions around major developmental regulators such as *Gata2* and *Cdx2* (Fig. 4E). Moreover, functional annotation of Prdm5 target genes in known pathways displayed enrichment for Prdm5 targets to be part of the WNT pathway (see Fig. S3D and E in the supplemental material), further confirming the previously published role for Prdm5 in regulating Wnt signaling (15, 18).

Proteomic analysis of Prdm5 binding partners reveals its association with complexes involved in chromatin organization. In order to understand the molecular mechanisms involving Prdm5 in transcriptional regulation, we sought to identify Prdm5 interaction partners from the E14 mES cell line. We employed a label-free interaction proteomic approach by comparing triplicate purifications from control cells or cells expressing HA-tagged (N-terminal tag) PRDM5 at levels comparable to the endogenous one (Fig. 5A).

We first adopted a previously published bioinformatics approach (38, 39) to discriminate between specific interactors and contaminants and identified 11 proteins as putative Prdm5 interaction partners (Fig. 5B; also see Table S4 in the supplemental material). Interestingly, two members of the TFIIIC complex (Gtf3c2 and Gtf3c4) were found to reach statistical significance as Prdm5 interactors, and other two members of the same complex were just below significance but still enriched in HA-PRDM5 purifications compared to control samples (Fig. 5B). Independent coimmunoprecipitation experiments validated that overexpressed Prdm5 associates with endogenous TFIIIC complex components such as TFIIIC-220 and TFIIIC-63 (Fig. 5D).

Given the recently suggested role for TFIIIC in chromosome organization (51) and the reported binding of Prdm5 to a distal enhancer-like site in osteoblastic cells (19), we hypothesized that Prdm5 can associate with other proteins involved in chromatin interactions. Indeed, by manual analysis of the list of candidate interaction partners for Prdm5, we noticed that cohesin complex members Smc1a and Smc3 were retrieved with increased MS/MS counts in IPs from HA-PRDM5 cells compared to those of control cells (see Fig. S4A in the supplemental material). The validity of our analysis was confirmed by the finding of RNA polymerase II subunit as a Prdm5 interaction partner (see Fig. S4), which we validated in a previous study (19).

Importantly, overexpressed Prdm5 readily bound endogenous SMC1 and CTCF (a known cohesin interaction partner [26]) (Fig. 5D). By size-exclusion chromatography experiments, we showed that Prdm5 elutes in fractions of size corresponding to high-molecular-weight complexes in E14 mES cells, and in the same fractions, we identified TFIIIC components as well as Ctcf and Smc1 as coeluting proteins (Fig. 5C). We demonstrate that Prdm5 can associate at the endogenous level with both CTCF and TFIIIC complex (see Fig. S4B in the supplemental material). Moreover, Prdm5-TFIIIC interaction does not depend on the presence of residual DNA, as it also occurs in the presence of ethidium bromide (see Fig. S4C).

In summary, analyses of Prdm5 binding partners suggest that a

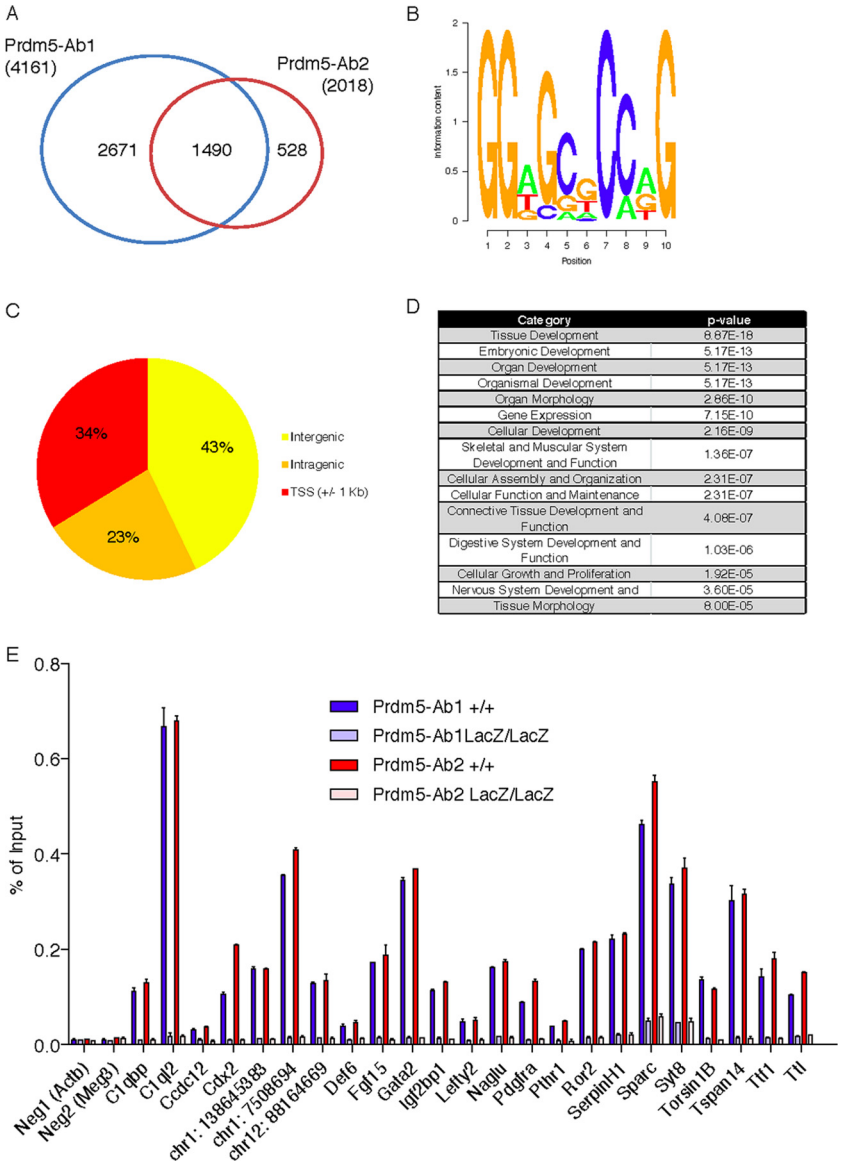

FIG 4 Prdm5 ChIP-seq reveals binding to developmental regulators. (A) Venn diagram representing the overlap of peaks identified by ChIP-seq with two polyclonal Prdm5 antibodies. (B) Slogo representation of Prdm5 consensus sequence obtained by ChIP-seq. (C) Distribution of the position of Prdm5 peaks divided in categories of chromosomal regions. (D) Functional annotation according to IPA database for Prdm5 target genes. (E) ChIP-qPCR validation of Prdm5 target regions with both Prdm5-Ab1 and Prdm5-Ab2 in wild-type and *Prdm5*$^{LacZ/LacZ}$ mES cells.

pool of Prdm5 resides in complexes involved in chromosome organization.

Prdm5 cooccupies genomic regions with TFIIIC and Ctcf complexes. The finding of Prdm5 associating with the TFIIIC complex, CTCF, and cohesins prompted us to investigate possible cooccupancy between Prdm5 and its binding partners. ChIP-seq experiments for TFIIIC (31), CTCF (27), and cohesins (45) in mES cells have been reported, so we analyzed these publicly available data sets.

We first addressed a possible interplay between Prdm5 and the TFIIIC complex by analyzing the presence of ChIP-seq signal for TFIIIC components around Prdm5 peaks, as shown in Fig. 6A. Only a subset (approximately 10%) of Prdm5 peaks displays enrichment for TFIIIC subunits, but strikingly, the binding of TFIIIC subunits occurred next to Prdm5 peaks in the 2-kb area inspected (Fig. 6A). When we visually inspected ChIP-seq tracks, we noted that Prdm5 binding indeed occurred predominantly next to TFIIIC subunit binding (Fig. 6B).

Importantly, gene annotation of regions bound by Prdm5 and surrounded by TFIIIC signal revealed enrichment for genes expressed during early embryonic stages. On the contrary, regions devoid of TFIIIC signal were assigned to genes expressed during limb/skeleton development (Fig. 6A).

TFIIIC has been proposed to have multiple roles by concurrent binding (and recruitment) of RNA polymerase III subunits or by binding independently at so-called ETC sites (30). However, when we analyzed the presence of Prdm5 around coordinates corresponding to TFIIIC-ETC sites or common TFIIIC-Rpc4 sites, we

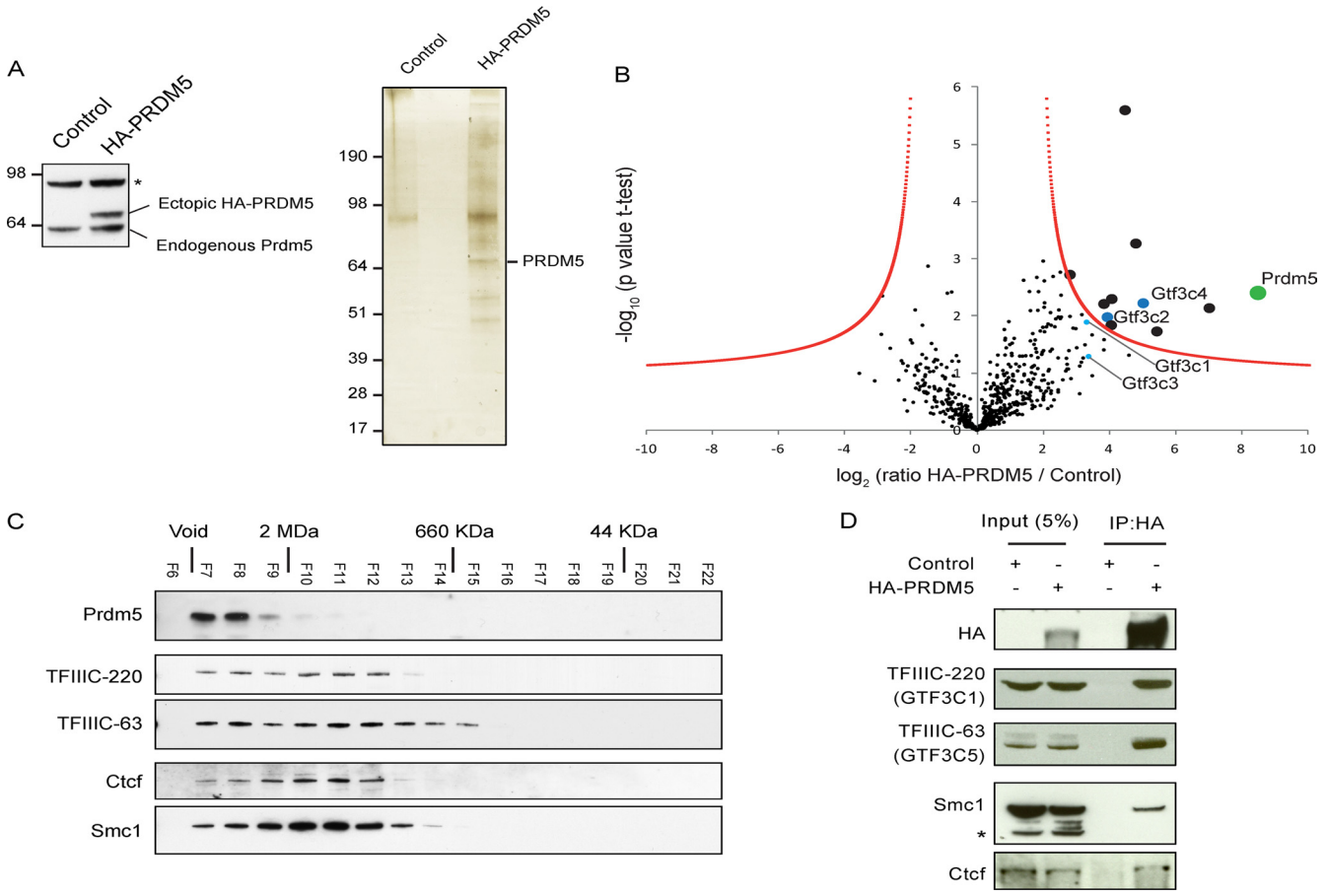

FIG 5 Prdm5 interacts with TFIIIC, Ctcf, and Smc1. (A, left) Western blot with anti-Prdm5 antibody from control cells and cells expressing Flag-HA-tagged PRDM5 (the asterisk indicates an unspecific band used as a loading control). (Right) Representative silver-staining analysis of samples from HA immunoprecipitation experiments. (B) Volcano plot representing results of the label-free pulldown of PRDM5. The logarithmic ratio of protein intensities in the PRDM5/control pulldowns were plotted against negative logarithmic *P* values of the *t* test performed on triplicate pulldown experiments. A hyperbolic red curve separates specific PRDM5 interactors (green dot, PRDM5; blue dots, TFIIIC components) from the background. (C) Western blot analysis for the indicated proteins on fractions resulting from size-exclusion chromatography from E14 mES cell nuclear extracts. Fractions corresponding to the molecular mass of standard markers are indicated. (D) Coimmunoprecipitation experiments between overexpressed HA-tagged PRDM5 and the indicated endogenous proteins. An asterisk indicates an unspecific band.

did not find differences, with Prdm5 occupancy always situated around, and not centered on, a subset of TFIIIC sites (Fig. 6C). On the contrary, when we classified common TFIIIC-Rpc4 sites according to the type of transcribed genes, we observed an increase in Prdm5 signal around tRNA genes versus regions that were previously validated as repetitive elements, such as SINEs or LINEs (Fig. 6D; also see Fig. S5A and B in the supplemental material), suggesting that Prdm5-TFIIIC common sites are enriched in previously described insulator elements (33).

Our proteomic analysis also identified an association between Prdm5, Ctcf, and Smc1. When analyzing the ChIP-seq signal for these proteins around Prdm5 peaks, we observed a large extent of cooccupancy between Prdm5 and each of the proteins (Fig. 6E). Different from the overlap with TFIIIC, the cooccupancy of Ctcf, Smc1, and Smc3 occurs centered on Prdm5 peaks (Fig. 6E). Indeed, by inspection of ChIP-seq tracks, we observed Prdm5 peaks, such as the one downstream of *Cdx2*, to be cooccupied by Ctcf, Smc1, and Smc3 (see Fig. S5C in the supplemental material).

Ctcf is known to be involved in chromatin organization, and recently it has been reported to associate with TFIIIC (29, 31, 32). Since both Prdm5 and Ctcf appear to bind DNA via specific consensus sequences, we asked which DNA sequence would drive the occupancy of Prdm5-Ctcf-Smc1 complex. To this end, we annotated common Prdm5-Ctcf-Smc1 peaks according to the presence of either Prdm5 or Ctcf consensus motifs and found that 64.3% of common Prdm5-Ctcf-Smc1 complexes exclusively contained a Prdm5 motif, and only 1.75% showed a Ctcf motif (Fig. 6F) ($P < 0.0001$ by chi-square test), suggesting that Prdm5 mediates Ctcf presence on common sites.

In summary, our data demonstrate that Prdm5 cooccupies genomic sites with TFIIIC and/or Ctcf-Smc1 complexes, with the latter being due to the overrepresentation of Prdm5 consensus sequences on common target sites. This Prdm5 interaction network indicates that Prdm5 is involved in chromatin organization in mouse embryonic stem cells in cooperation with insulator binding proteins.

DISCUSSION

This study provides a global characterization of Prdm5 based on the isolation of Prdm5 knockout mES cells derived from a previ-

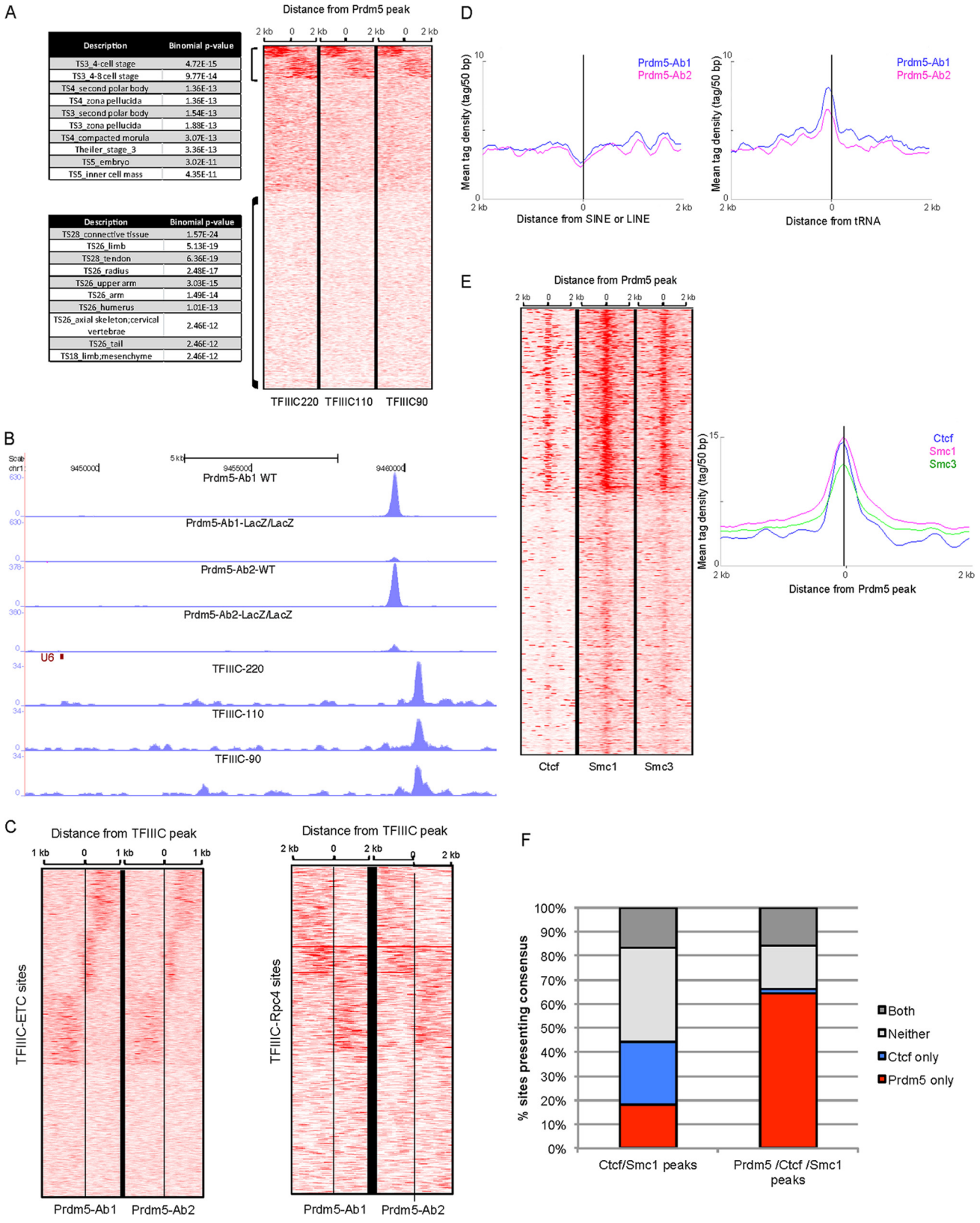

FIG 6 Prdm5 interacts with TFIIIC and Ctcf complexes. (A, left) Gene ontology annotation of common Prdm5-TFIIIC sites or Prdm5-specific sites. (Right) Heat map displaying the ChIP-seq signal for TFIIIC subunits in a ±2-kb window around Prdm5 target regions (n = 1,490). (B) Example of ChIP-seq tracks for a Prdm5 binding site upstream of U6 and the neighboring peaks for TFIIIC subunits. (C) Heat maps displaying Prdm5-Ab1 and Prdm5-Ab2 ChIP-seq signals in a ±2-kb window around TFIIIC ETC target regions (n = 2,233) (left) or TFIIIC-Rpc4 common sites (n = 419) (right) classified according to reference 31. (D) Signal distribution for Prdm5-Ab1 (blue) and Prdm5-Ab2 (pink) in a ±2-kb window centered on SINE and LINE elements (left) or tRNA genes (right). (E, left) Heat map displaying ChIP-seq signal for Ctcf, Smc1, and Smc3 in a ±2-kb window around Prdm5 target regions (n = 1490). (Right) Signal distribution for Ctcf, Smc1, and Smc3 in a ±2-kb window centered on Prdm5 target regions. (F) Histogram representing the percentage of sites containing Prdm5 and/or Ctcf consensus sequence in Ctcf/Smc1-only common peaks or Prdm5/Ctcf/Smc1 common peaks.

ously characterized gene trap model (19). We employed high-throughput technologies, such as label-free quantitative interaction proteomics, RNA-seq, and ChIP-seq, to identify Prdm5 interactions partners as well as the transcriptional program controlled by Prdm5. By integration of such robust data sets, we identified a novel role for Prdm5 in chromatin organization.

Previous studies have indicated versatile and cell type-specific functions for Prdm5 in gene regulation, acting in both transcriptional repression and activation, likely depending on the presence of specific cofactors (19, 21).

Functionally, PRDM5 has been coined a tumor suppressor, as the promoter is frequently found methylated in different cancer types (14–17); however, genetic evidence causally establishing tumor-suppressive functions of PRDM5 has been published only recently (60), and the Prdm5 gene trap mouse strain employed in this study does not develop spontaneous tumors (19). However, while Prdm5 gene trap animals are viable and fertile (19), morpholino experiments in zebrafish embryos revealed an essential role for Prdm5 (18). In zebrafish, the Prdm5 loss-of-function phenotype was assigned to the role of Prdm5 as a negative regulator of Wnt signaling. Although Prdm5 in the mouse does not appear to have an essential function (19), ChIP-seq analysis in mES cells revealed that Prdm5 might affect WNT signaling by directly targeting genomic regions carrying genes involved in the pathway (see Fig. S3E in the supplemental material). This is in agreement with studies of human cancer cell lines linking PRDM5 to deregulated levels of β-catenin (15).

In a previous study, we identified a role for Prdm5 in ossification processes during development and demonstrated that Prdm5 associates with all collagen loci in MC3T3 preosteoblastic cells (19). Interestingly, Prdm5 binding to collagen genes occurs predominantly within the body of the genes, and binding intensity correlates with the expression level of the collagen loci. Mechanistically, we found that Prdm5 associates with RNA polymerase II and is important for maintaining RNA polymerase II levels at the *Col1a1* locus (19), presumably via affecting local chromatin structures to promote transcriptional elongation or by direct modification of elongating RNA polymerase II.

In the present study, we exploited the high expression of Prdm5 in mES cells to gain a global perspective of the functions of Prdm5 in a dynamic model system. Interestingly, a previous analysis identified *PRDM5* together with *PRDM14* and *PRDM4* as the three PRDMs expressed at significant levels in hES cells (9). Although PRDM14 is essential for stemness maintenance of human ES cells by regulating OCT4 levels (9), in mouse ES cells loss-of-function experiments for Prdm14 revealed only a minor effect on stemness maintenance (10); indeed, *Prdm14* knockout mice are viable, displaying only fertility defects, indicating species-specific functions (52). In line with these data, we find that Prdm5 loss does not affect mES cell maintenance, while, similarly to PRDM14, it might have different functions in human cells or act redundantly with Prdm14 in mouse embryonic stem cells. The latter hypothesis could be valid, considering that Prdm5 and Prdm14 share approximately 20% of genes bound according to ChIP-seq data (this study and reference 10), encoding important signaling molecules such as bone morphogenetic proteins (BMPs), EphA2, and Meis2 (data not shown).

Gene expression regulation occurs at multiple levels, starting from the genetic code and more complex epigenetic regulations, such as histone modifications, to the recently described genomic interactions and higher-order chromatin organization.

Indeed, the rapid development of genome-wide-scale technologies has revealed the importance of spatial organization in the nucleus and the assembly of complex inter/intrachromosomal networks to establish cell-specific transcriptomes (53). Such phenomena have essential roles during development, where cellular differentiation of progenitor cells into mature cell types necessitates a plastic genome to be rearranged to express specific transcriptional programs.

A well-studied process of global chromosome organization occurs during X inactivation when female embryonic stem cells start to differentiate. Indeed, this process includes the spatial organization and interaction of the two X chromosomes (54) and the chromosome condensation driven by the long noncoding RNA molecule Xist (55). An example of local chromatin organization is represented by the differentially methylated region (DMR) around the H19-IGF2 locus that determines imprinting. Upstream of the H19 gene, DNA methylation of an insulator sequence determined the binding of CTCF, which acts through an enhancer-blocking function to determine expression of H19 and IGF2 genes (56). Ctcf binding induces local chromatin looping via association with cohesin complex both at the DMR (57) and on a more global scale (26). Indeed, cohesins have been shown to mediate chromatin looping in other complexes as well, such as the bridging between enhancers and proximal promoters involving the Mediator complex (45).

By mass spectrometry analysis of affinity-purified PRDM5 complexes, we identified TFIIIC components as Prdm5 interaction partners. TFIIIC is a multimeric protein complex, the main role of which is to recognize RNA polymerase III target genes containing internal promoters with A and B boxes, typically tRNA genes. TFIIIC recognition leads to the recruitment of the TFIIIB machinery containing the Brf1 subunit for RNA polymerase III-driven transcription (reviewed in reference 58). Interestingly, TFIIIC has been discovered to bind genomic sites independent of the RNA polymerase III machinery, called ETC (30). In recent years, TFIIIC was shown to be involved in a number of additional processes, such as chromatin insulation (reviewed in reference 51). Interestingly, tRNA genes themselves recently have been shown to act as chromatin insulators, indicating an additional role for these loci in chromatin biology other than tRNA transcription (33). Comparison of Prdm5 and TFIIIC ChIP-seq data sets revealed that TFIIIC binding occurs near Prdm5 binding peaks in a range of 2 kb, and these common sites are assigned to genes expressed in early embryogenesis according to GREAT functional annotation. When we dissected TFIIIC sites according to their genomic features, we identified ETC sites and tRNAs as TFIIIC targets also bound by Prdm5, indicating a role for Prdm5 in mediating chromatin insulation.

The other protein complex members we identified to associate with Prdm5 are Ctcf and cohesin. Ctcf is a multi-zinc finger transcriptional regulator that is known as the main mediator of insulation functions in vertebrates (reviewed in reference 25). Instead, cohesins are multimeric protein complexes with a main function in sister chromatid cohesion (59), although studies have revealed a mitosis-independent role for cohesins in mediating gene expression (45) and Ctcf-dependent insulator functions (26). Interestingly, we identified Ctcf, Smc1, and Smc3 as proteins exhibiting large cooccupancy with Prdm5 target sites, with their binding cen-

tered on Prdm5 peaks. Given that Ctcf was shown to direct cohesin enrichment at specific sites via its consensus sequence (26), we looked into the presence of Prdm5 and Ctcf binding motifs at the common binding sites. Strikingly, while the Prdm5 consensus binding motifs were found at the majority of common Prdm5-Ctcf binding sites, the Ctcf binding motif was underrepresented, indicating a role for Prdm5 in associating Ctcf with chromatin at these sites.

In summary, we provide evidence for a new function for Prdm5 in interacting with chromatin organizers. Further studies will be needed to determine if Prdm5 mainly affects local structures, such as the establishment of the transcriptional boundary and regional looping, or if Prdm5 also is important for higher-order chromatin structures, such as interchromosomal interactions.

ACKNOWLEDGMENTS

We are grateful to Anita Friismose and Cord Brakebusch for assistance in the transgenic core facility. We acknowledge Bettina Mentz and Maddalena Arigoni for excellent technical assistance and Klaus Hansen for suggestions and help with biochemical experiments.

Work in our laboratories was supported by the Danish National Research Foundation, the Danish National Advanced Technology Foundation, the Novo Nordisk Foundation, European Commission (EC) FP7 programs (ONCOMIRS, grant agreement number 201102, and APOSYS), the Lundbeck Foundation, and the Danish Cancer Society. G.G.G. is supported by an American-Italian Cancer Foundation postdoctoral research fellowship. C.F. is supported by a Marie Curie IEF and EMBO long-term fellowship.

This publication reflects only the authors' views. The EC is not liable for any use that may be made of the information herein.

REFERENCES

1. **Fog CK, Galli GG, Lund AH.** 2012. PRDM proteins: important players in differentiation and disease. Bioessays **34:**50–60.
2. **Huang S.** 2002. Histone methyltransferases, diet nutrients and tumour suppressors. Nat. Rev. Cancer **2:**469–476.
3. **Hayashi K, Yoshida K, Matsui Y.** 2005. A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature **438:**374–378.
4. **Eom GH, Kim K, Kim SM, Kee HJ, Kim JY, Jin HM, Kim JR, Kim JH, Choe N, Kim KB, Lee J, Kook H, Kim N, Seo SB.** 2009. Histone methyltransferase PRDM8 regulates mouse testis steroidogenesis. Biochem. Biophys. Res. Commun. **388:**131–136.
5. **Kim KC, Geng L, Huang S.** 2003. Inactivation of a histone methyltransferase by mutations in human cancers. Cancer Res. **63:**7619–7623.
6. **Huang S, Shao G, Liu L.** 1998. The PR domain of the Rb-binding zinc finger protein RIZ1 is a protein binding interface and is related to the SET domain functioning in chromatin-mediated gene expression. J. Biol. Chem. **273:**15933–15939.
7. **Fumasoni I, Meani N, Rambaldi D, Scafetta G, Alcalay M, Ciccarelli FD.** 2007. Family expansion and gene rearrangements contributed to the functional specialization of PRDM genes in vertebrates. BMC Evol. Biol. **7:**187. doi:10.1186/1471-2148-7-187.
8. **Kinameri E, Inoue T, Aruga J, Imayoshi I, Kageyama R, Shimogori T, Moore AW.** 2008. Prdm proto-oncogene transcription factor family expression and interaction with the Notch-Hes pathway in mouse neurogenesis. PLoS One **3:**e3859. doi:10.1371/journal.pone.0003859.
9. **Chia NY, Chan YS, Feng B, Lu X, Orlov YL, Moreau D, Kumar P, Yang L, Jiang J, Lau MS, Huss M, Soh BS, Kraus P, Li P, Lufkin T, Lim B, Clarke ND, Bard F, Ng HH.** 2010. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature **468:**316–320.
10. **Ma Z, Swigut T, Valouev A, Rada-Iglesias A, Wysocka J.** 2011. Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates. Nat. Struct. Mol. Biol. **18:**120–127.
11. **Seale P, Bjork B, Yang W, Kajimura S, Chin S, Kuang S, Scime A, Devarakonda S, Conroe HM, Erdjument-Bromage H, Tempst P, Rudnicki MA, Beier DR, Spiegelman BM.** 2008. PRDM16 controls a brown fat/skeletal muscle switch. Nature **454:**961–967.
12. **Chuikov S, Levi BP, Smith ML, Morrison SJ.** 2010. Prdm16 promotes stem cell maintenance in multiple tissues, partly by regulating oxidative stress. Nat. Cell Biol. **12:**999–1006.
13. **Zhang Y, Stehling-Sun S, Lezon-Geyda K, Juneja SC, Coillard L, Chatterjee G, Wuertzer CA, Camargo F, Perkins AS.** 2011. PR-domain-containing Mds1-Evi1 is critical for long-term hematopoietic stem cell function. Blood **118:**3853–3861.
14. **Deng Q, Huang S.** 2004. PRDM5 is silenced in human cancers and has growth suppressive activities. Oncogene **23:**4903–4910.
15. **Shu XS, Geng H, Li L, Ying J, Ma C, Wang Y, Poon FF, Wang X, Ying Y, Yeo W, Srivastava G, Tsao SW, Yu J, Sung JJ, Huang S, Chan AT, Tao Q.** 2011. The epigenetic modifier PRDM5 functions as a tumor suppressor through modulating WNT/beta-catenin signaling and is frequently silenced in multiple tumors. PLoS One **6:**e27346. doi:10.1371/journal.pone.0027346.
16. **Watanabe Y, Kim HS, Castoro RJ, Chung W, Estecio MR, Kondo K, Guo Y, Ahmed SS, Toyota M, Itoh F, Suk KT, Cho MY, Shen L, Jelinek J, Issa JP.** 2009. Sensitive and specific detection of early gastric cancer with DNA methylation analysis of gastric washes. Gastroenterology **136:**2149–2158.
17. **Watanabe Y, Toyota M, Kondo Y, Suzuki H, Imai T, Ohe-Toyota M, Maruyama R, Nojima M, Sasaki Y, Sekido Y, Hiratsuka H, Shinomura Y, Imai K, Itoh F, Tokino T.** 2007. PRDM5 identified as a target of epigenetic silencing in colorectal and gastric cancer. Clin. Cancer Res. **13:**4786–4794.
18. **Meani N, Pezzimenti F, Deflorian G, Mione M, Alcalay M.** 2009. The tumor suppressor PRDM5 regulates Wnt signaling at early stages of zebrafish development. PLoS One **4:**e4273. doi:10.1371/journal.pone.0004273.
19. **Galli GG, de Lichtenberg KH, Carrara M, Hans W, Wuelling M, Mentz B, Multhaupt HA, Fog CK, Jensen KT, Rappsilber J, Vortkamp A, Coulton L, Fuchs H, Gailus-Durner V, Hrabe de Angelis M, Calogero RA, Couchman JR, Lund AH.** 2012. Prdm5 regulates collagen gene transcription by association with RNA polymerase II in developing bone. PLoS Genet. **8:**e1002711. doi:10.1371/journal.pgen.1002711.
20. **Burkitt Wright EM, Spencer HL, Daly SB, Manson FD, Zeef LA, Urquhart J, Zoppi N, Bonshek R, Tosounidis I, Mohan M, Madden C, Dodds A, Chandler KE, Banka S, Au L, Clayton-Smith J, Khan N, Biesecker LG, Wilson M, Rohrbach M, Colombi M, Giunta C, Black GC.** 2011. Mutations in PRDM5 in brittle cornea syndrome identify a pathway regulating extracellular matrix development and maintenance. Am. J. Hum. Genet. **88:**767–777.
21. **Duan Z, Person RE, Lee HH, Huang S, Donadieu J, Badolato R, Grimes HL, Papayannopoulou T, Horwitz MS.** 2007. Epigenetic regulation of protein-coding and microRNA genes by the Gfi1-interacting tumor suppressor PRDM5. Mol. Cell. Biol. **27:**6889–6902.
22. **Gondor A, Ohlsson R.** 2009. Chromosome crosstalk in three dimensions. Nature **461:**212–217.
23. **Dillon N.** 2006. Gene regulation and large-scale chromatin organization in the nucleus. Chromosome Res. **14:**117–126.
24. **Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, van Steensel B.** 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature **453:**948–951.
25. **Phillips JE, Corces VG.** 2009. CTCF: master weaver of the genome. Cell **137:**1194–1211.
26. **Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters JM.** 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature **451:**796–801.
27. **Handoko L, Xu H, Li G, Ngan CY, Chew E, Schnapp M, Lee CW, Ye C, Ping JL, Mulawadi F, Wong E, Sheng J, Zhang Y, Poh T, Chan CS, Kunarso G, Shahab A, Bourque G, Cacheux-Rataboul V, Sung WK, Ruan Y, Wei CL.** 2011. CTCF-mediated functional chromatin interactome in pluripotent cells. Nat. Genet. **43:**630–638.
28. **Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K.** 2009. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res. **19:**24–32.

29. **Moqtaderi Z, Wang J, Raha D, White RJ, Snyder M, Weng Z, Struhl K.** 2010. Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. Nat. Struct. Mol. Biol. **17:**635–640.
30. **Moqtaderi Z, Struhl K.** 2004. Genome-wide occupancy profile of the RNA polymerase III machinery in Saccharomyces cerevisiae reveals loci with incomplete transcription complexes. Mol. Cell. Biol. **24:**4118–4127.
31. **Carriere L, Graziani S, Alibert O, Ghavi-Helm Y, Boussouar F, Humbertclaude H, Jounier S, Aude JC, Keime C, Murvai J, Foglio M, Gut M, Gut I, Lathrop M, Soutourina J, Gerard M, Werner M.** 2012. Genomic binding of Pol III transcription machinery and relationship with TFIIS transcription factor distribution in mouse embryonic stem cells. Nucleic Acids Res. **40:**270–283.
32. **Oler AJ, Alla RK, Roberts DN, Wong A, Hollenhorst PC, Chandler KJ, Cassiday PA, Nelson CA, Hagedorn CH, Graves BJ, Cairns BR.** 2010. Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. Nat. Struct. Mol. Biol. **17:**620–628.
33. **Raab JR, Chiu J, Zhu J, Katzman S, Kurukuti S, Wade PA, Haussler D, Kamakaka RT.** 2012. Human tRNA genes function as chromatin insulators. EMBO J. **31:**330–350.
34. **Bryja V, Bonilla S, Arenas E.** 2006. Derivation of mouse embryonic stem cells. Nat. Protoc. **1:**2082–2087.
35. **Lundby A, Olsen JV.** 2011. GeLCMS for in-depth protein characterization and advanced analysis of proteomes. Methods Mol. Biol. **753:**143–155.
36. **Kelstrup CD, Young C, Lavallee R, Nielsen ML, Olsen JV.** 2012. Optimized fast and sensitive acquisition methods for shotgun proteomics on a quadrupole orbitrap mass spectrometer. J. Proteome Res. [Epub ahead of print.] doi:10.1021/pr3000249.
37. **Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M.** 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. **10:**1794–1805.
38. **Hubner NC, Mann M.** 2011. Extracting gene function from protein-protein interactions using quantitative BAC interactomics (QUBIC). Methods **53:**453–459.
39. **Hubner NC, Bird AW, Cox J, Splettstoesser B, Bandilla P, Poser I, Hyman A, Mann M.** 2010. Quantitative proteomics combined with BAC transgeneomics reveals in vivo protein interactions. J. Cell Biol. **189:**739–754.
40. **Kim D, Salzberg SL.** 2011. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol. **12:**R72.
41. **Roberts A, Pimentel H, Trapnell C, Pachter L.** 2011. Identification of novel transcripts in annotated genomes using RNA-Seq. Bioinformatics **27:**2325–2329.
42. **Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L.** 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. **7:**562–578.
43. **Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J.** 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. **5:**R80.
44. **ENCODE Project Consortium.** 2011. A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. **9:**e1001046. doi:10.1371/journal.pbio.1001046.
45. **Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA.** 2010. Mediator and cohesin connect gene expression and chromatin architecture. Nature **467:**430–435.
46. **Ye T, Krebs AR, Choukrallah MA, Keime C, Plewniak F, Davidson I, Tora L.** 2011. seqMINER: an integrated ChIP-seq data interpretation platform. Nucleic Acids Res. **39:**e35.
47. **Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, Zhang MQ, Lobanenkov VV, Ren B.** 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell **128:**1231–1245.
48. **McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G.** 2010. GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. **28:**495–501.
49. **Pasini D, Bracken AP, Hansen JB, Capillo M, Helin K.** 2007. The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol. Cell. Biol. **27:**3769–3779.
50. **Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, Nakayama J, Okano M.** 2006. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells **11:**805–814.
51. **Donze D.** 2012. Extra-transcriptional functions of RNA polymerase III complexes: TFIIIC as a potential global chromatin bookmark. Gene **493:**169–175.
52. **Yamaji M, Seki Y, Kurimoto K, Yabuta Y, Yuasa M, Shigeta M, Yamanaka K, Ohinata Y, Saitou M.** 2008. Critical function of Prdm14 for the establishment of the germ cell lineage in mice. Nat. Genet. **40:**1016–1022.
53. **Fraser P, Bickmore W.** 2007. Nuclear organization of the genome and the potential for gene regulation. Nature **447:**413–417.
54. **Xu N, Tsai CL, Lee JT.** 2006. Transient homologous chromosome pairing marks the onset of X inactivation. Science **311:**1149–1152.
55. **Chaumeil J, Le Baccon P, Wutz A, Heard E.** 2006. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev. **20:**2223–2237.
56. **Bell AC, Felsenfeld G.** 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature **405:**482–485.
57. **Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters JM, Murrell A.** 2009. Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. PLoS Genet **5:**e1000739. doi:10.1371/journal.pgen.1000739.
58. **White RJ.** 2011. Transcription by RNA polymerase III: more complex than we thought. Nat. Rev. Genet. **12:**459–463.
59. **Losada A, Hirano T.** 2005. Dynamic molecular linkers of the genome: the first decade of SMC proteins. Genes Dev. **19:**1269–1287.
60. **Galli GG, Multhaupt HA, Carrara M, de Lichtenberg KH, Christensen IB, Linnemann D, Santoni-Rugiu E, Calogero RA, Lund AH.** 22 July 2013. Prdm5 suppresses ApcMin-driven intestinal adenomas and regulates monoacylglycerol lipase expression. Oncogene. doi:10.1038/onc.2013.283.