6608–6626 *Nucleic Acids Research, 2018, Vol. 46, No. 13* *Published online 29 May 2018*
doi: 10.1093/nar/gky461

PHD finger protein 1 (PHF1) is a novel reader for histone H4R3 symmetric dimethylation and coordinates with PRMT5–WDR77/CRL4B complex to promote tumorigenesis

Ruiqiong Liu[1,†], Jie Gao[1,†], Yang Yang[1], Rongfang Qiu[1], Yu Zheng[1], Wei Huang[1], Yi Zeng[1], Yongqiang Hou[1], Shuang Wang[1], Shuai Leng[1], Dandan Feng[1], Wenqian Yu[1], Gancheng Sun[1], Hang Shi[1], Xu Teng[2] and Yan Wang[1,2,*]

[1]2011 Collaborative Innovation Center of Tianjin for Medical Epigenetics, Tianjin Key Laboratory of Cellular and Molecular Immunology, Key Laboratory of Immune Microenvironment and Disease (Ministry of Education), Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China and [2]Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Capital Medical University, Beijing 100069, China

Received June 25, 2017; Revised May 02, 2018; Editorial Decision May 10, 2018; Accepted May 18, 2018

ABSTRACT

Histone post–translational modifications regulate chromatin structure and function largely through interactions with effector proteins that often contain multiple histone-binding domains. PHF1 [plant homeodomain (PHD) finger protein 1], which contains two kinds of histone reader modules, a Tudor domain and two PHD fingers, is an essential factor for epigenetic regulation and genome maintenance. While significant progress has been made in characterizing the function of the Tudor domain, the roles of the two PHD fingers are poorly defined. Here, we demonstrated that the N-terminal PHD finger of PHF1 recognizes symmetric dimethylation of H4R3 (H4R3me2s) catalyzed by PRMT5–WDR77. However, the C-terminal PHD finger of PHF1, instead of binding to modified histones, directly interacts with DDB1, the main component of the CUL4B-Ring E3 ligase complex (CRL4B), which is responsible for H2AK119 mono-ubiquitination (H2AK119ub1). We showed that PHF1, PRMT5–WDR77, and CRL4B reciprocally interact with one another and collaborate as a functional unit. Genome-wide analysis of PHF1/PRMT5/CUL4B targets identified a cohort of genes including E-cadherin and *FBXW7,* which are critically involved in cell growth and migration. We demonstrated that PHF1 promotes cell proliferation, invasion, and tumorigenesis *in vivo* and *in vitro* and found that its expression is markedly upregulated in a variety of human cancers. Our data identified a new reader for H4R3me2s and provided a molecular basis for the functional interplay between histone arginine methylation and ubiquitination. The results also indicated that PHF1 is a key factor in cancer progression, supporting the pursuit of PHF1 as a target for cancer therapy.

INTRODUCTION

Covalent chromatin modifications, which include DNA methylation and histone posttranslational modifications (PTMs), play an important role in epigenetic regulation. Histone methylation occurs on both lysine (K) and arginine (R) residues, and its dynamic regulation plays a critical role in chromatin biology. Several protein domain families have been identified that specifically recognize histone methylation, including the plant homeodomain (PHD), PWWP, chromodomains and Tudor domain (1,2). Such recognition mechanisms facilitate coordinated regulatory functions in a plethora of chromatin template-based biological processes including gene regulation, DNA replication, and recombination. While significant progress has been made in characterizing the effectors of methyllysine (3), less is known regarding how histone arginine residues are recognized, although arginine methylation plays equally important roles (4).

The Polycomblike (Pcl) protein was first identified in *Drosophila* (5). Early studies with Pcl showed that it was localized to polycomb group (PcG) sites on *Drosophila*

*To whom correspondence should be addressed. Tel: +86 22 83336946; Fax: +86 22 83336946; Email: yanwang@ccmu.edu.cn
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

polytene chromosomes (6). Biochemical analysis of the *Drosophila* Polycomblike protein (dPcl) showed that it is present in a 1 MDa complex in early embryos and that this complex is distinct from the 600-kDa E(z) complex PRC2 (7). The Pcl protein has a Tudor domain and two tandem PHD finger domains that are also conserved in its three human homologues, PHF1 (hPcl1), MTF2 (hPcl2), and PHF19 (hPcl3) (8,9). PHF1 is essential in epigenetic regulation and genome maintenance. It is reported that PHF1 interacts with PRC2 and plays an important role in H3K27 methylation and *HOX* gene silencing (10,11). PHF1 has also been implicated in DNA damage repair. It physically interacts with Ku70-Ku80, which binds to DSB ends, and may play a role in nonhomologous end joining (NHEJ) and maintaining genomic stability (12). The Tudor domain of PHF1 recognizes trimethylation of H3K36 (13–15). However, the precise function of the two PHD domains remains unclear.

PRMT5 is a type II arginine methyltransferase that has been linked to gene silencing through the establishment of repressive histone marks (16), including symmetrical dimethylation of H4R3, H2AR3, H3R2 and H3R8 (17–19). PRMT5 is essential for stem cell maintenance and developmental growth in mouse embryonic and induced pluripotent stem cells (20), and is required for initiation of differentiation in myogenesis (21). Genetic depletion of PRMT5 has previously been reported to impair cancer cell viability by promoting G1 cell cycle arrest and apoptosis (22). The transcription-repressive functions of PRMT5 are also crucial for the process of epithelial–mesenchymal transition (EMT), a hallmark of tumor progression (23). As a consequence of EMT, cancer cells are able to migrate and invade much more efficiently, and this process probably has an active role in metastasis. PRMT5 has oncogenic properties because of its ability to repress the expression of tumor suppressor genes, such as E-cadherin (24). It is believed PRMT5 is complexed with the WD repeat protein WDR77 (also known as MEP50 or androgen coactivator p44) in vertebrates as a tetramer of heterodimers. WDR77 is hypothesized to be required for the recruitment of protein substrates to the catalytic domain of PRMT5. Interestingly, shRNAs targeting either *PRMT5* or *WDR77* reduced the levels of both proteins, consistent with depletion of the methylosome complex using either shRNA (25).

Cullin 4B-Ring E3 ligases (CRL4B), assembled with CUL4B, DDB1 and ROC1 as core components, are involved in a broad variety of physiologically- and developmentally-controlled processes such as cell cycle progression, replication and DNA damage responses (26). Loss-of-function mutation in CUL4B results in X-linked mental retardation, absence of speech, short statue, restlessness and other phenotypes in humans (27), and *Cul4b* knockout mice are embryonic lethal (28). Our previous study revealed that CRL4B possesses an intrinsic transcription repressive activity by catalyzing histone H2AK119 mono-ubiquitination. CRL4B physically associates with PRC2 to transcriptionally repress several tumor suppressors and to promote tumorigenesis (29). In addition, CRL4B also coordinates with SUV39H1/HP1/DNMT3A to facilitate epigenetic silencing of tumor suppressors such as *IGFBP3* (30).

In this study, we identified PHF1 as a novel reader that specifically recognizes symmetrically methylated H4R3, a substrate of PRMT5–WDR77. PHF1 facilitates the PRMT5–WDR77/CRL4B complex to conform coordinated H4R3me2s/H2AK119ub1 modifications for transcriptional silencing. We demonstrated that PHF1 promotes cell proliferation, invasion, and tumorigenesis *in vitro* and *in vivo* and found that its expression is markedly upregulated in various human cancers. Our data indicated that PHF1 promotes tumorigenesis, supporting the pursuit of PHF1 as a target for cancer therapy.

MATERIALS AND METHODS

Antibodies and reagents

The sources of the antibodies were: anti-FLAG (F1408), anti-PHF1 (AV39528), anti-CUL4B (C9995), anti-CUL4A (C0371), anti-HDAC1 (H3284), anti-HDAC2 (H3159), anti-Fibronectin (F3648), anti-Vimentin (V6630), and anti-β-actin (A1978) (Sigma-Aldrich); anti-DDB1 (sc-25367), anti-MTA1 (sc-10813), anti-SIN3A (sc-994), anti-GRIP1 (sc-8996), and anti-SRC1 (sc-8995) (Santa Cruz Biotechnology); anti-PRMT5 (ab109451), anti-PHF1 (ab80042), anti-ROC1 (ab2977), anti-LSD1 (ab17721), anti-H4R3me2s (ab5823), anti-H4R3me2a (ab129231), anti-H3R8me2s (ab130740), anti-H3R8me2a (ab127163), anti-H3R2me2a (ab80075), anti-H3K36me1 (ab9048), and anti-H3K36me3 (ab9050) (Abcam); anti-H2AK119ub1 (05-678), anti-H3R2me2s (ABE460), anti-MTA3 (IM1012), and anti-EED (17-10034) (Millipore); anti-SUZ12 (3737s) (Cell Signaling Technology); anti-SAP180 (A302-233A), anti-SAP30 (A303-551A), and anti-WDR77 (A301-561A) (Bethyl); anti-E-cadherin (610181), anti-α-catenin (610193), anti-γ-catenin (610253), anti-N-cadherin (610920), anti-EZH2 (612667) and anti-AIB1 (611104) (BD Bioscience); anti-SETD2 (55377-1-AP) (Proteintech). The peptides of histone tails were from Abcam or purchased from Ontores (Zhejiang, China). Protein A/G Sepharose CL-4B beads were from Amersham Biosciences, and protease inhibitor mixture cocktail was from Roche Applied Science. siRNAs were purchased from Sigma-Aldrich. shRNAs were obtained from GenePharma Co Ltd (Shanghai, China).

Cell culture and transfection

The cell lines used were obtained from the American Type Culture Collection. HEK 293T cells and Hs 578T cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Cells were maintained in a humidified incubator equilibrated with 5% CO_2 at 37°C. MDA-MB-231 cells and MDA-MB-453 cells were cultured in L-15 medium supplemented with 10% FBS and without CO_2. Transfections were carried out using Lipofectamine 2000 or Lipofectamine® RNAiMAX Reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Each experiment was performed in triplicate and repeated at least three times. For RNAi experiments, at least four independent shRNA sequences were tested for each gene and the one with the highest efficiency was used.

Immunopurification and mass spectrometry

HEK 293T cells stably expressing FLAG-PHF1 were generated by transfection of the cells with FLAG-tagged PHF1 and selection in medium containing 1 mg/ml of G418. Anti-FLAG immunoaffinity columns were prepared using anti-FLAG M2 affinity gel (Sigma-Aldrich) following the manufacturer's suggestions. Cell lysate was obtained from approximately 5×10^8 cells and applied to an equilibrated FLAG column of 1ml bed volume to allow for adsorption of the protein complexes to the column resin. After binding, the column was washed with cold BC500 buffer containing 50 mM Tris, 2 mM EDTA, 500 mM KCl, 10% glycerol, and protease inhibitors. FLAG peptide (0.2 mg/ml, Sigma-Aldrich) was applied to the column to elute the FLAG protein complex as described by the vendor. Fractions of the bed volume were collected and resolved on SDS-polyacrylamide gels, silver stained, and subjected to LC-MS/MS sequencing and data analysis.

Fast protein liquid chromatography (FPLC)

MDA-MB-231 nuclear extracts were prepared and dialyzed against buffer D (20 mM HEPES, pH 8.0, 10% glycerol, 0.1 mM EDTA, 300 mM NaCl) (Applygen Technologies, Beijing, China). Approximately 6 mg of nuclear protein was concentrated to 1 ml using Millipore Ultrafree centrifugal filter apparatus (10 kDa nominal molecular mass limit), and then applied to an 850 × 20 mm Superose 6 size exclusion column (Amersham Biosciences, Salt Lake City, UT, USA) that had been equilibrated with buffer D containing 1 mM dithiothreitol and calibrated with protein standards (blue dextran, 2000 kDa; thyroglobulin, 667 kDa; Ferritin, 440 kDa; Aldolase, 158 kDa; Ovalbumin, 43 kDa; all from Amersham Biosciences). The column was eluted at a flow rate of 0.5 ml/min and fractions were collected.

Immunoprecipitation and western blotting

For immunoprecipitation assays, cells were washed with cold phosphate buffered saline (PBS) and lysed with cold cell lysis buffer for 30 min at 4°C. Then, 500 μg of cellular extract was incubated with appropriate specific antibodies or normal rabbit/mouse immunoglobin G (IgG) at 4°C overnight with constant rotation, followed by the addition of protein A/G Sepharose beads and incubation for 2 h at 4°C. Beads were then washed five times with cell lysis buffer (50 mM Tris–HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 0.5% NP-40, 0.25% sodium deoxycholate and protease inhibitor mixture). The immune complexes were subjected to SDS-PAGE followed by immunoblotting with secondary antibodies. Immunodetection was performed using enhanced chemiluminescence (ECL System, Thermo Scientific) according to the manufacturer's instructions.

Glutathione S-transferase (GST) pull-down experiments

GST fusion constructs were expressed in *Escherichia coli* BL21 cells, and crude bacterial lysates were prepared by sonication in cold PBS in the presence of a protease inhibitor mixture. The *in vitro* transcription and translation experiments were performed with rabbit reticulocyte lysate (TNT Systems; Promega). In GST pull-down assays, ∼10 μg of the appropriate GST fusion protein was mixed with 5–8 μl of the *in vitro* transcribed/translated products and incubated in binding buffer (0.8% BSA in PBS in the presence of the protease inhibitor mixture). The binding reaction was then added to 30 μl of glutathione-Sepharose beads and mixed at 4°C for 2 h. The beads were washed five times with binding buffer, resuspended in 30 μl of 2 × SDS-PAGE loading buffer, and resolved on 12% gels. Protein bands were detected with specific antibodies by western blotting.

Histone peptide array

A MODified Histone Peptide Array (Active Motif) was used for PHD1 binding detection. A MODified Protein Domain Kit and analysis software (Active Motif) were used in accordance with the manufacturer's instructions. These peptide identities are listed in Supplementary File 3.

Isothermal titration calorimetry (ITC)

A MicroCal ITC-200 system (Malvern Instruments Ltd.) was used for ITC experiments. Briefly, the synthesized peptides (Ontores Biotechnologies) and proteins were all subjected to extensive dialysis against PBS. Peptides at concentrations of ∼500 μM were loaded into the ITC syringe and proteins at concentrations of ∼50 μM were loaded into the ITC cell. Then, 19 injections of 2 μl peptide each were made into the cell at 25°C, automatically. The results of binding isotherms were analyzed using Origin 7.0 software package (Origin Lab).

Surface plasmon resonance (SPR)

SPR experiments were carried out using a Biacore T200 (GE Healthcare). All SPR-based materials were purchased from GE Healthcare. Biotin-polypeptides (Ontores Biotechnologies) were diluted in HBS-EP and immobilized on an SA chip. Approximately 600 RU of the immobilized peptides were obtained. Interaction analyses were tested using HBS-EP as a running buffer. Increasing concentrations of His-PHF_{PHD1} (0.4438, 0.8875, 1.775, 3.55, 7.1, 14.2 μM) were injected using the 'Kinetics/Affinity' program. A flow cell without immobilized peptide served as a non-specific binding control. The SA chip surface was regenerated after each cycle by injecting 10 mM NaOH for 30 s. k_a, k_d and K_D were determined using the 'Kinetics' model in the Biacore T200 evaluation software version 2.0.

Chromatin immunoprecipitation (ChIP) and Re-ChIP

ChIPs and Re-ChIPs were performed in MDA-MB-231 cells as previously described (31). Briefly, 1×10^7 cells were cross-linked with 1% formaldehyde, sonicated, pre-cleared, and incubated with 5–10 μg of antibody per reaction. Complexes were washed with low and high salt buffers, and the DNA was extracted and precipitated. For Re-ChIP assays, immune complexes were eluted from the beads with 20 mM dithiothreitol. Eluents were then diluted 30-fold with ChIP dilution buffer and subjected to a second immunoprecipitation reaction. The final elution step was performed using

1% SDS solution in Tris-EDTA buffer, pH 8.0. DNA template enrichment was analyzed by conventional PCR using primers specific for each target gene promoter. The primers used are listed in Supplementary File 4.

Real-time quantitative RT-PCR

Total RNA was isolated from samples with Trizol reagents following the manufacturer's instructions (Invitrogen). Any potential DNA contamination was removed by RNase-free DNase treatment (Promega). cDNA was prepared using the MMLV Reverse Transcriptase (Fermentas). Relative quantitation was performed using the ABI PRISM 7500 sequence detection system (Applied Biosystems) which measures real-time SYBR Green fluorescence. Briefly, cDNA was mixed with 1 μl forward and reverse primers (5 μM of each), 8 μl RNase-free water, and 10 μl 2 × PCR SYBR Green Mix buffer in a 20 μl reaction. Then, 40 cycles of PCR were conducted at 95°C for 15 s, and 60°C for 1 min within each cycle. Quantitation was then performed by means of the comparative Ct method ($2^{-\Delta\Delta Ct}$) with the expression of *GAPDH* as an internal control. The primers used are listed in Supplementary File 4.

Lentiviral production and infection

Recombinant lentiviruses expressing shPHF1, shCUL4B, shPRMT5 were constructed by Shanghai GenePharma. Concentrated viruses were used to infect 5×10^5 cells in a 60 mm dish with 8 μg/ml polybrene. Infected cells were then subject to sorting target expression. The shRNA sequences are as follows and ∗ indicates the shRNA chosen for further study. For the shRNA-resistant PHF1 overexpression vector, the shPHF1#2 target site bases 253–273 were mutated (GGTGAAGAGCTGCTATGCTGC).

shSCR	TTCTCCGAACGTGTCACGT
shPHF1#1	GGTGTGTCTGGTCCAGTTTGA
shPHF1#2∗	GGAGAGGAACTCCTCTGTTGT
shPHF1#3	GCAACCGACAGCAGAGTTACT
shPHF1#4	GCCACAAGGACCGTTTCATTT
shPRMT5#1	GGGACTGGAATACGCTAATTG
shPRMT5#2	GCCATCACTCTTCCATGTTCT
shPRMT5#3∗	GGAATCTCAGACATATGAAGT
shPRMT5#4	GCTATTGCACCTTGGAATTTC
shCUL4B#1	GCCACGTACCGATACAGAAGA
shCUL4B#2∗	GGATTCATTGGATAGCGTTCT
shCUL4B#3	GGATAGAACTTACGTTCTTCA
shCUL4B#4	GCAGACAGACTTATTACTTAC

Cell invasion assay

Transwell chamber filters (Chemicon Incorporation) were coated with Matrigel. After infection with lentivirus, cells were resuspended in serum-free media and 2.5×10^4 cells in 0.5 ml serum-free media were placed in the upper chamber of the transwell. The chamber was then transferred to a well containing 500 μl of media containing 10% FBS. Cells were incubated for 18–48 h at 37°C. Cells in the upper well were removed by wiping the top of the membrane with cotton swabs. The membranes were then stained, and the remaining cells were counted. Three high-powered fields were counted for each membrane.

Mouse xenograft models

MDA-MB-231 cells that had been transfected to stably express firefly luciferase (Xenogen) were infected with lentiviruses carrying control shRNA or shPHF1. These cells were inoculated into the left abdominal mammary fat pad (5×10^6 cells) or injected into the lateral tail vein (2×10^6 cells) of 6-week-old female SCID mice. For bioluminescence imaging, mice were injected abdominally with 200 mg/kg of D-luciferin in PBS. At 15 min after injection, mice were anesthetized, and bioluminescence was imaged with a charge-coupled device camera (IVIS; Xenogen). Bioluminescence images were obtained with a 15 cm field of view, binning (resolution) factor of eight, 1/*f* stop, open filter, and an imaging time of 30 s to 2 min. Bioluminescence from relative optical intensity was defined manually. Photon flux was normalized to background, which was defined from a relative optical intensity drawn over a mouse not given an injection of luciferin. Animal handling and procedures were approved by Tianjin Medical University Institutional Animal Care.

Tissue specimens and immunohistochemistry

Samples were frozen in liquid nitrogen immediately after surgical removal and maintained at –80°C until analysis. Samples were fixed in 4% paraformaldehyde (Sigma-Aldrich) at 4°C overnight and then embedded in paraffin, sectioned at 8 μm onto Superfrost-Plus Slides, processed per standard protocols using DAB staining, and monitored microscopically.

Statistical analysis

Results are reported as mean ± SD unless otherwise noted. Comparisons were performed using two tailed unpaired *t* test. SPSS V.17.0 was used for statistical analysis. Tumor data sets were downloaded from http://www.ncbi.nlm.nih.gov/geo and GSE numbers are shown in the Figures. Data for Kaplan–Meier survival analysis were from http://kmplot.com/analysis.

RESULTS

PHF1 is physically associated with PRMT5–WDR77 and CRL4B complex

In an effort to better understand the mechanistic role of PHF1, we employed affinity purification and mass spectrometry to identify the proteins that are associated with PHF1. In these experiments, FLAG-tagged PHF1 (FLAG-PHF1) was stably expressed in HEK 293T cells. Cellular extracts were prepared and subjected to affinity purification using an anti-FLAG affinity gel. Mass spectrometric analysis showed that PHF1 was co-purified with various epigenetic factors including ZNF217, AIB1, SIN3A, DDB1, CUL4B, SUZ12, EZH2, MTA1/2/3, PRMT5, HDAC1/2, RbAp46/48, WDR77 and EED (Figure 1A). Among the listed proteins, the association of PHF1 with PRC2 including EZH2, SUZ12, and EED has been previously reported (10,11). The detailed results of mass spectrometric analysis are provided in Supplementary File 1. The presence

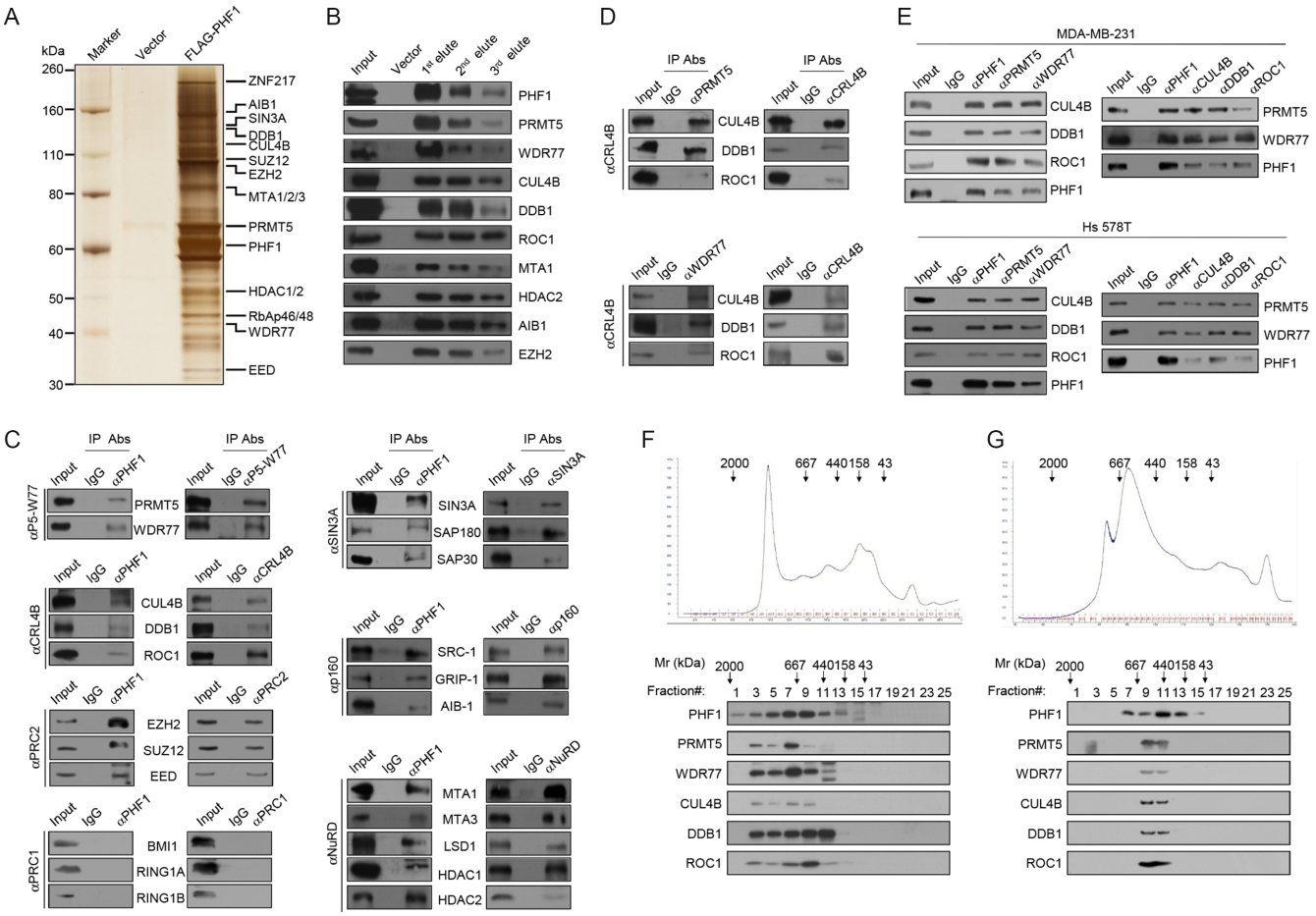

Figure 1. PHF1 Interacts with CRL4B and PRMT5–WDR77 Complex. (**A**) Immunoaffinity purification of PHF1-containing protein complexes. Cellular extracts from HEK 293T cells stably expressing FLAG (control) or FLAG-PHF1 were immunopurified with anti-FLAG affinity columns and eluents with FLAG peptide. The eluates were resolved by SDS-PAGE and silver-stained. The protein bands were retrieved and analyzed by mass spectrometry. Detailed results from the mass spectrometric analysis are provided in Supplementary File 1. (**B**) Western blot analysis of the purified fractions using antibodies against the indicated proteins. (**C**) Association of PHF1 with the P5-W77 (PRMT5, WDR77) complex, the CRL4B complex (CUL4B, DDB1, ROC1), the NuRD components (MTA1, MTA3, LSD1, HDAC1, HDAC2), the SIN3A complex (SIN3A, SAP180, SAP30) and the p160 SRC family (SRC-1, GRIP-1, AIB-1) in HEK 293T cells. Whole cell lysates from HEK 293T cells were prepared and co-immunoprecipitation (co-IP) assays with anti-PHF1 followed by immunoblotting (IB) with antibodies against the indicated proteins, or with antibodies against the indicated proteins followed by IB with anti-PHF1. PHF1 did not interact with the PRC1 complex (BMI1, RING1A, RING1B). IgG served as a negative control. This figure consists of seven sub-panels which showed the association of PHF1 with the seven indicated complexes. In the left part of each sub-panel, co-IP assays were performed with PHF1 antibody; immunocomplexes were then immunoblotted using antibodies against the intermediate proteins. In the right part of each sub-panel, co-IP assays were performed with antibodies against each of intermediate proteins respectively and then immunoblotted using PHF1's antibody. The slanting labels identify the IP antibodies, and the horizontal labels in the middle indicate the detailed IB antibodies for the left panels and detailed IP antibodies for the right panels. (**D**) Association of CRL4B complex with PRMT5–WDR77 complex in HEK 293T cells. Whole cell lysates from HEK 293T cells were prepared and co-IP was performed with antibodies against the indicated proteins. Immunocomplexes were then IB tested using antibodies against the indicated proteins. The figure consists of two sub-panels which showed the association of PRMT5–WDR77 with CRL4B complex. In the left part of each sub-panel, co-IP assays were performed with PRMT5 antibody in the upper sub-panel and WDR77 antibody in the lower sub-panel; immunocomplexes were then immunoblotted using antibodies against the intermediate proteins. In the right part of each sub-panel, co-IP assays were performed with antibodies against each of intermediate proteins respectively and then immunoblotted using PRMT5's antibody in the upper sub-panel and WDR77's antibody in the lower sub-panel. The slanting labels identify the IP antibodies, and the horizontal labels of CRL4B complex in the middle indicate the detailed IB antibodies for the left panels and detailed IP antibodies for the right panels. (**E**) Association of PHF1, PRMT5–WDR77 complex, and CRL4B complex in MDA-MB-231 and Hs 578T cells. Whole cell lysates were immunoprecipitated with antibodies against the indicated proteins. Immunocomplexes were then IB tested using antibodies against the indicated proteins. (**F**) Co-fractionation of PHF1, PRMT5–WDR77 complex, and CRL4B complex by fast protein liquid chromatography (FPLC). Nuclear extracts of MDA-MB-231 cells were fractionated on a DEAE sepharose column followed by a Superose 6 gel filtration column. The fractions were analyzed by western blotting. Molecular weight standards are shown at top (in kDa). (**G**) Western blot of the PHF1-containing complex fractionated by Superose 6 gel filtration.

of PRMT5–WDR77, CRL4B, MTA1, HDAC2, AIB1 and EZH2 in the PHF1-associated protein complex was confirmed by western blot analysis (Figure 1B).

Besides PRC2, the newly identified PHF1-associated proteins indicated that PHF1 interacts with various protein complexes and families including the PRMT5–WDR77 complex, the CRL4B complex, the NuRD complex, the SIN3A/HDAC complex, and the p160 nuclear receptor coactivator family. To further support the physical association of PHF1, total proteins from HEK 293T cells were extracted and co-immunoprecipitation (co-IP) experiments were performed with antibodies detecting the endogenous proteins. Moreover, immunoprecipitation (IP) with antibodies against PHF1 followed by immunoblotting (IB) with antibodies against PRMT5, WDR77, CUL4B, DDB1, ROC1, MTA1, MTA3, LSD1, HDAC1, HDAC2, SIN3A, SAP30, SAP180, SRC1, GRIP1 and AIB1 demonstrated that PHF1 co-immunoprecipitated with PRMT5–WDR77 complex, CRL4B complex, NuRD components, SIN3A complex and the p160 nuclear receptor coactivator family. Reciprocally, IP with antibodies against the proteins mentioned above and IB with antibodies against PHF1 also confirmed these interactions (Figure 1C). However, co-IP assays with antibodies against PRC1 components BMI1, RING1A or RING1B demonstrated that PHF1 did not interact with PRC1 complex (Figure 1C).

Interestingly, we noticed that DDB1 functions as a linker molecule to recruit receptor WD40 proteins to CUL4–ROC1 ubiquitin ligases and WDR77 is just a WD40 repeat-containing protein (32). We found that the components of PRMT5–WDR77 and CRL4B could be easily immunoprecipitated by each other (Figure 1D). Based on these interactions, it is reasonable to speculate that the histone modification effector PHF1 and the two histone modification enzymes PRMT5–WDR77 and CRL4B might form a functional complex. To further validate the interaction between PHF1, PRMT5–WDR77 and CRL4B in breast cancer cells, cellular extracts from MDA-MB-231 cells and Hs 578T cells were immunoprecipitated with antibodies against PHF1, PRMT5 and WDR77 followed by IB with antibodies against CRL4B subunits and PHF1. Reciprocal IPs with antibodies against PHF1, CUL4B, DDB1 and ROC1 followed by IB with antibodies against PRMT5, WDR77, and PHF1 further confirmed the associations (Figure 1E).

To ascertain the existence of a complex composed of PHF1/PRMT5–WDR77/CRL4B, protein fractionation experiments were carried out with nuclear proteins by FPLC. Nuclear extracts derived from MDA-MB-231 cells were fractionated by DEAE sepharose, followed by Superose 6 gel filtration chromatography. Western blotting revealed a major peak at approximately 669–1000 kDa for PHF1, PRMT5–WDR77, and the CRL4B components CUL4B, DDB1, and ROC1 (Figure 1F). Significantly, the chromatographic profiles of PHF1, PRMT5–WDR77 and CRL4B were largely overlapping, supporting the argument that PHF1, PRMT5–WDR77 and CRL4B complex may cooperate functionally *in vivo*. Furthermore, western blotting of FLAG-PHF1 affinity eluent from FPLC after Superose 6 gel filtration revealed that PHF1 co-immunoprecipitated with PRMT5–WDR77 and CRL4B (Figure 1G). The interaction between PHF1, PRMT5–WDR77 and CRL4B suggests that these proteins may function in a concerted manner to regulate expression of certain target genes.

Molecular interactions between PHF1, PRMT5–WDR77 and CRL4B complex

In order to further explore the molecular basis for the interaction between PHF1, PRMT5–WDR77, and CRL4B, a GST pull-down assay was performed using GST-fused PHF1 and *in vitro* transcribed/translated PRMT5, WDR77, CUL4B, DDB1 and ROC1. These experiments revealed that PHF1 could interact directly with PRMT5 and DDB1. Reciprocally, GST pull-down assays using GST-fused PRMT5, WDR77, CUL4B, DDB1, ROC1, and *in vitro* transcribed/translated PHF1 also verified these results (Figure 2A). As we expected, DDB1 could interact with WDR77 directly through the DWD motif (DDB1-binding WD40 protein) of WDR77 (Figure 2B).

It is considered that PHF1 contains one Tudor domain, two adjacent PHD fingers immediately followed by a domain of extended homology (EH) with *Drosophila* PCL, and the C-terminal end harbors a 'chromo-like domain' weakly homologous to the chromodomain (8,33). To determine which domain mediates the interactions of PHF1, a series of truncation vectors of PHF1 (N1–N5, C1–C5) were used to generate GST fusion proteins (Figure 2C). The results showed that the C-terminus of PHF1 is essential for PRMT5 binding and PHF1's second PHD domain (PHD2) is necessary for DDB1 binding (Figure 2D). It has been reported that DDB1 consists of three propeller (BP) domains: BPA, BPB, and BPC (Figure 2E), of which the BPB domain is responsible for CUL4-DDB1 interaction (34,35). In order to map the interaction interface of DDB1 with PHF1 and WDR77, GST pull-down assays were performed with GST-fused BPA, BPB, BPC, and BPA+BPC domain of DDB1. The results indicated that the BPA_{DDB1} and BPC_{DDB1} domains were involved in directly interacting with PHF1 or WDR77 (Figure 2F). Taken together, these results provide strong support for a physical association between PRMT5–WDR77 and CRL4B, which is bridged by PHF1 and WDR77 molecules (Figure 2G).

The N terminal PHD finger of PHF1 binds specifically to histone H4R3me2s

The Tudor domain of PHF1 recognizes trimethylation of H3K36 (13–15). However, the precise function of the PHD domains remains unclear. Considering the comprehensive interaction model described in Figure 2G, we speculated that PHF1 might be involved in the recognition of histone arginine methylation that may be generated or affected by PRMT5. In order to test this hypothesis, we next determined the histone specificity towards nucleosomal substrates by GST pull-down assays using GST-PHF1 and histone octamers purified from calf thymus. Surprisingly, besides the H3K36me3, the GST pull-down results revealed that PHF1 could also bind octamers containing H4R3me2s, which is precisely the product of PRMT5 as previously reported. However, octamers marked with the other catalysates of PRMT5, H3R8me2s and H3R2me2s,

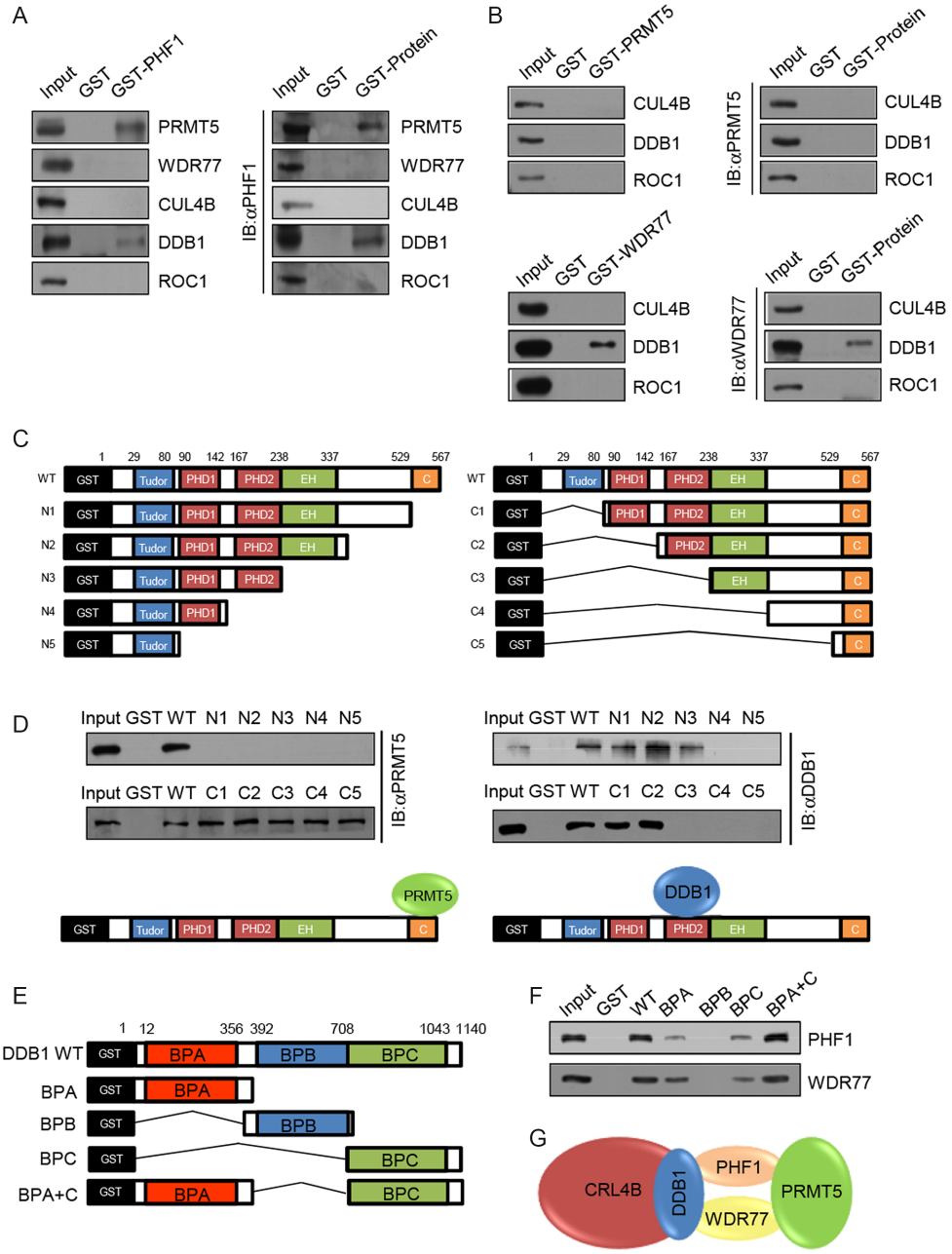

Figure 2. Molecular Interaction between PHF1, PRMT5–WDR77, and CRL4B Complex subunits. (**A**, **B**) GST pull-down experiments with bacterially expressed GST-fused proteins and the indicated *in vitro* transcribed/translated proteins. The slanting labels show the GST-fused proteins, the vertical labels indicate the IB antibodies, and the horizontal labels in the right panels show the detailed GST-fused proteins. (**C**, **D**) Identification of the essential domains required for the interaction with PRMT5 or DDB1 of PHF1. GST pull-down experiments with a bacterially expressed series of truncation constructs of PHF1 (N1–N5, C1–C5) to generate GST fusion proteins and the *in vitro* transcribed/translated indicated proteins. (**E**, **F**) Identification of the essential domains required for interactions with PHF1 or WDR77 of DDB1. GST pull-down experiments with bacterially expressed GST-fused truncated DDB1 proteins and the indicated *in vitro* transcribed/translated proteins. (**G**) Schematic diagram depicting the molecular interactions between PHF1, PRMT5–WDR77 and CRL4B.

did not bind to PHF1. Asymmetric dimethylation of H3R8, H3R2 and H4R3 also did not result in binding to PHF1 (Figure 3A). Furthermore, total proteins from HEK 293T were extracted and co-IP experiments were performed with antibodies against PHF1. Immunoblotting with antibodies against dimethylated histone marks showed that PHF1 co-immunoprecipitated with endogenous histones that contained H4R3me2s rather than H3R8me2s, H3R2me2s or asymmetric dimethylation (Figure 3B).

To confirm the interaction between PHF1 and H4R3me2s, we performed GST pull-down assays using GST-fused PHF1 and peptides of histones with residues H3R8, H4R3 and H3R2 symmetrically or asymmetrically dimethylated. Immunoblotting with antibodies against the histone marks showed that PHF1 specifically recognized

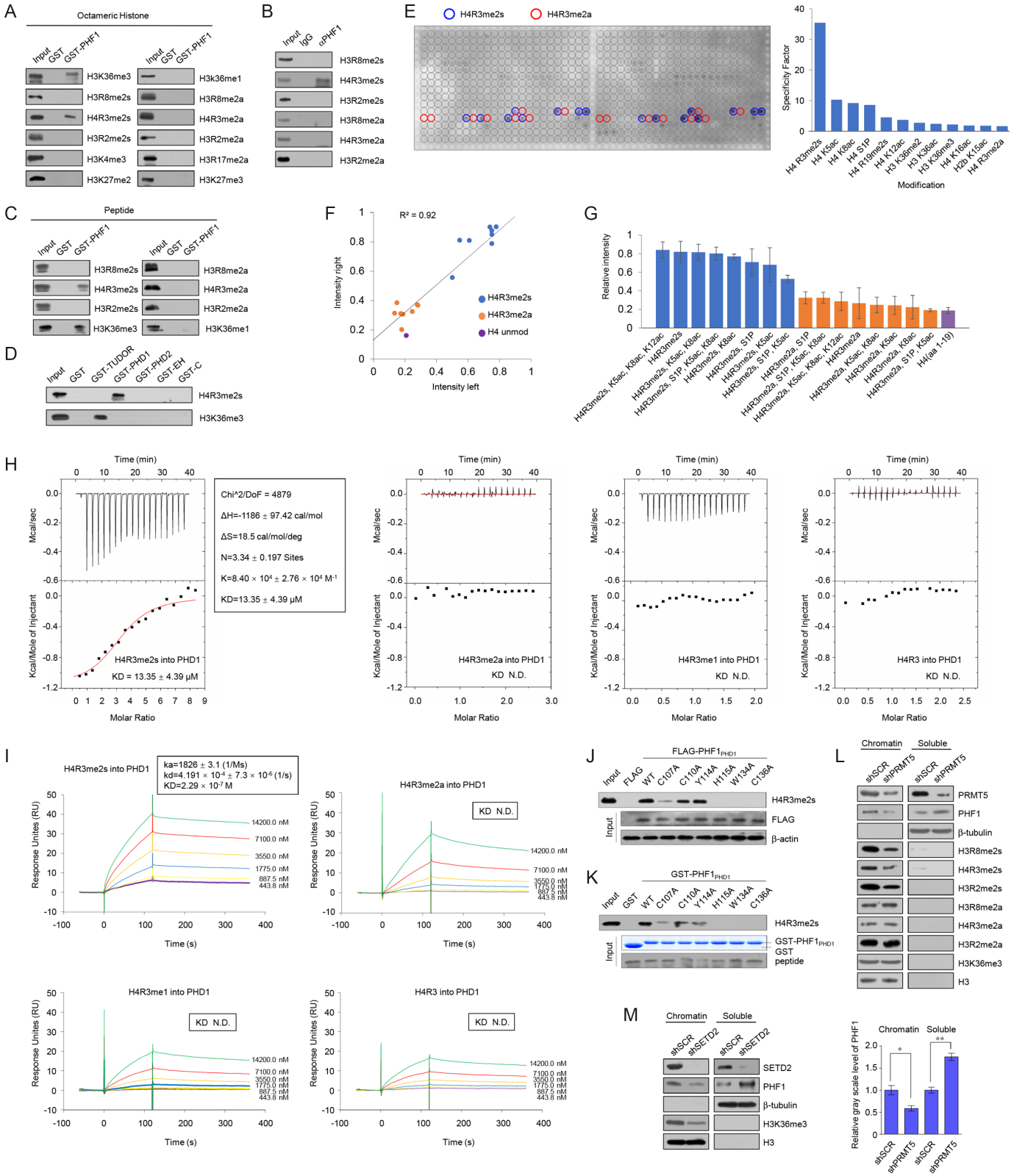

Figure 3. The PHD domain of PHF1 recognizes H4R3me2s. (**A**) GST pull-down experiments with bacterially expressed GST-PHF1 and histone octamers. Immunoblot using antibodies against various histone methylation modifications. (**B**) Whole cell lysates from HEK 293T cells were immunoprecipitated with antibodies against PHF1. Immunocomplexes were then IB tested using antibodies against the indicated proteins. (**C**) GST pull-down experiments with bacterially expressed GST-PHF1 and peptides with indicated residues. (**D**) GST pull-down experiments with bacterially expressed GST-fused truncated PHF1 proteins and peptides with H4R3me2s or H3K36me3. (**E**) Anti-His immunoblot (left panel) and graphical analysis of the highest binding events detected (right panel) showing the binding specificity of the His-PHF1$_{PHD1}$ domain measured on Histone Peptide Array. Peptides were spotted in duplicate

H4R3me2s (Figure 3C). In order to map the regions of PHF1 required for H4R3me2s binding, similar peptide pull-down assays were performed with a GST-fused PHF1 Tudor domain (1–80 aa), the two PHD finger domains (90–142 aa, 167–238 aa), the EH fragment (238–337 aa), the chromo-like domain (529–567 aa), with H4R3me2s and H3K36me3 peptides (Figure 3D). The results showed that the Tudor domain was responsible for H3K36me3 binding while the N-terminal PHD finger domain (PHD1) specifically recognized H4R3me2s. We next sought to determine the histone binding specificity of $PHF1_{PHD1}$ in an unbiased manner by using peptide array assays. Interrogation of the His-tagged $PHF1_{PHD1}$ using peptide arrays (Figure 3E, left panel) revealed a strong preference for various H4R3me2s-containing peptides (Figure 3E, right panel). In addition, $PHF1_{PHD1}$ displayed little interaction with H4R3me2a (Figure 3F and G), indicating that PHD1 can accurately distinguish H4R3me2s from H4R3me2a. Using isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR) assays, we next quantified these observed interactions. Consistent with the results of peptide array assays and peptide pull-down assays, ITC measurements demonstrated a specific interaction of $PHF1_{PHD1}$ with H4R3me2s., with a dissociation constant (KD) of 13.35 $\pm$ 4.39 μM. Compared to H4R3me2s, interactions with H4R3me2a, H4R3me1, and H4R3me0 were not detected (Figure 3H). SPR showed similar results (Figure 3I). To examine which residues in PHD1 are important for the interaction between PHD1 and H4R3me2s, we introduced individual alanine substitutions to each of the PHD1 evolutionarily conserved residues. We found that single point mutations including C107A, H115A, W134A, and C136A, abrogated the binding of $PHF1_{PHD1}$ to H4R3me2s peptides in Co-IPs *in vivo* (Figure 3J) and peptide pull-downs *in vitro* (Figure 3K). In conclusion, these results indicated that the N-terminal PHD finger domain of PHF1 is responsible for the specific recognition of H4R3me2s.

It has been reported that full-length PHF1 localizes to the nucleus and its association with chromatin is dependent on H3K36me3 (13). We next determined whether H4R3me2s plays a similar role to PHF1. To do this, we diminished global levels of H4R3me2s by interfering with the expression of a H4R3 symmetric methyltransferase, PRMT5, in HEK 293T cells. We separated the cells into chromatin-bound and soluble cytoplasm/nucleoplasm fractions. As previously reported, we found a reduction in symmetrically methylated H3R8, H4R3 and H3R2, while no change was detected involving the asymmetrically methylated histone residues and trimethylated H3K36. Remarkably, the results showed a reduction of PHF1 in the chromatin-bound fraction, concurrent with an increase in the soluble fraction (Figure 3L), which indicated that PHF1 association with chromatin is dependent on H4R3me2s. We also examined the level of H3K36me3 and found that decreased levels of PRMT5 did not affect H3K36me3 levels (Figure 3L). In addition, when SETD2 (an H3K36 methyltransferase) was depleted (36), the amount of PHF1 in chromatin-bound components was decreased (Figure 3M). It is possible that both marks are essential for PHF1 chromatin binding, or that they are required for PHF1 chromatin binding in different contexts. This will require further investigation. Taken together, our data demonstrate that PHF1 specifically binds to H4R3me2s and that this association plays a role in targeting PHF1 to chromatin.

Genome-wide identification of transcription targets for the PHF1/PRMT5–WDR77/CRL4B complex

To further investigate the functional associations between PHF1, PRMT5 and CRL4B and to explore the biological significance of these interactions, we next analyzed the genome-wide transcriptional targets of the PHF1/PRMT5/CRL4B complex using the ChIP-on-chip approach. In these experiments, ChIP was conducted using MDA-MB-231 cells with antibodies against PHF1, PRMT5, or CUL4B. Following ChIP, PHF1-, PRMT5- and CUL4B-associated DNAs were amplified using non-biased conditions, labeled, and hybridized to an oligonucleotide array covering 2.7 kb (–2.2 kb to +0.5 kb with respect to transcription start sites) of 18 028 annotated transcripts in the NCBI data base with a false discovery rate less than 0.05 (GEO accession number: GSE84071). The detailed results of the ChIP-on-chip experiments are summarized in Supplementary File 2. A total of 703 unique promoters were found to be targeted by PHF1 and they were then classified into various cellular signaling pathways using KEGG pathway software (http://www.kegg.jp/kegg/pathway.html) (Figure 4A). These signaling pathways include MAPK, Jak-STAT, Wnt, adheren junction, TGF-beta, cell cycle and pathways in cancer that are critically involved in cell growth, survival, migration and invasion. The data from PHF1 antibodies (703 genes), PRMT5 antibodies (1498 genes), and CUL4B antibodies (1358 genes) were then analyzed with the others for the identification of common targets, or co-targets (Figure 4B).

Quantitative ChIP analysis in MDA-MB-231 cells using specific antibodies against PHF1, PRMT5 and CUL4B on selected genes showed a strong enrichment

←

as shown in two boxes on the same array. The positions of H4R3me2s-containing peptides are highlighted with blue circles and H4R3me2a with red circles. (**F**) Scatter plot of duplicate peptide arrays probed with the $PHF1_{PHD1}$ domain. Peptides are colored according to the legend. The correlation coefficient was calculated to measure linear association. (**G**) Graphical representation of H4R3me2-containing peptides showing the binding specificity of GST-$PHF1_{PHD1}$ measured on Histone Peptide Array. (**H**) Experimental ITC titration curves of $PHF1_{PHD1}$ binding to differentially methylated H4R3 peptides, as indicated. Concentrations of protein and peptide were approximately 50 and 500 μM, respectively. The parameters are labeled beside the figures. N.D., not determined. (**I**) SPR analysis of the interaction of $PHF1_{PHD1}$ with differentially methylated H4R3 peptides, as indicated. The parameters are labeled in the figures. N.D., not determined. (**J**) Whole cell lysates from HEK 293T cells transfected with indicated FLAG-PHD1 mutations were immunoprecipitated with antibodies against FLAG. Immunocomplexes were then IB tested using antibodies against H4R3me2s. (**K**) GST pull-down experiments with the indicated bacterially expressed GST-fused PHD1 mutations and H4R3me2s peptides. (**L**) Western blot analysis of chromatin-bound and soluble cell fractions in HEK 293T cells after transfection with control shRNA or shRNA targeting PRMT5. Gray scanning analysis of the PHF1 blot is shown in the bottom panel. (**M**) Western blot analysis of chromatin-bound and soluble cell fractions in HEK 293T cells after transfection with control shRNA or shRNA targeting SETD2.

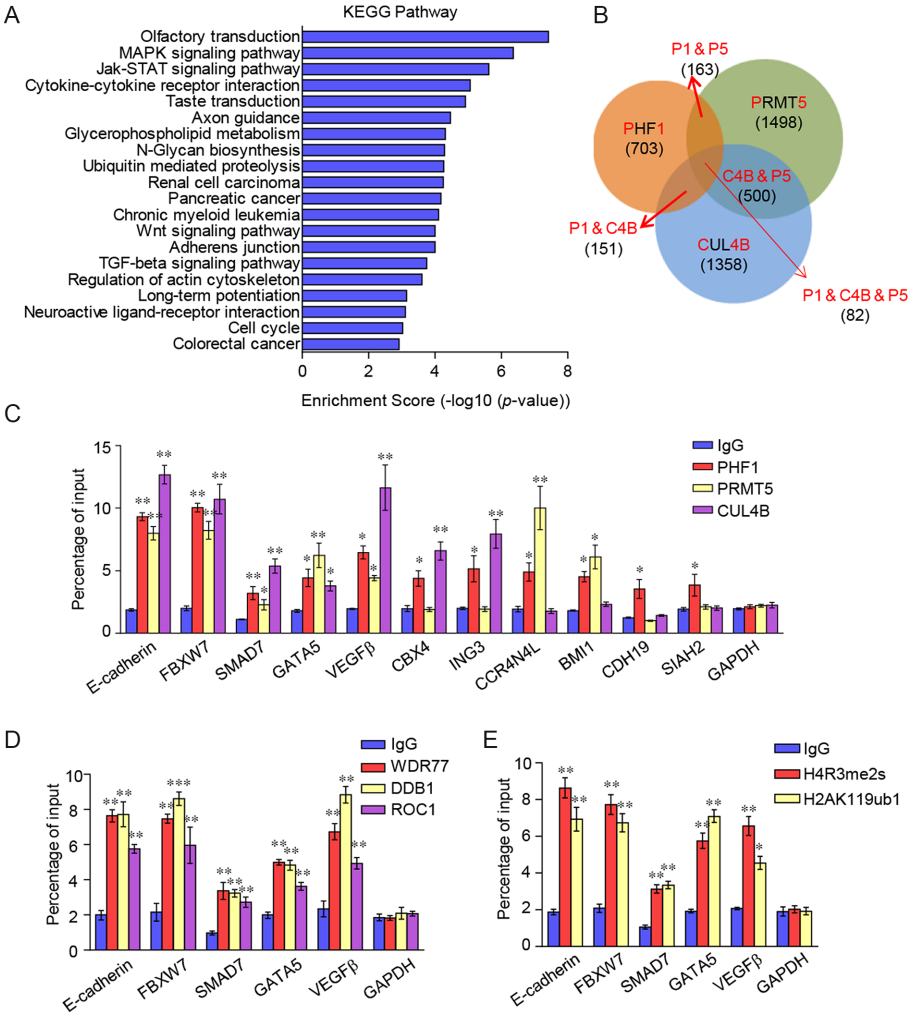

Figure 4. Genome-wide Transcriptional Target Analysis for the PHF1/PRMT5/CRL4B Complex. (**A**) Pathway analysis of the 703 target genes of PHF1 arranged into functional groups. (**B**) Venn diagram of overlapping promoters bound by PHF1, PRMT5 and CUL4B in MDA-MB-231 cells. The numbers represent the number of the promoters that were targeted by the indicated proteins. The detailed results of the ChIP-on-chip experiments are provided in Supplementary File 2. (**C**–**E**) Verification of the ChIP-on-chip results by qChIP analysis of the indicated genes in MDA-MB-231 cells. Results are presented as percentage of input with glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) as a negative control. Error bars represent mean ± SD for three independent experiments ($^{*}P < 0.05$, $^{**}P < 0.01$, and two-tailed unpaired *t* test).

on the promoters of these genes, including *CDH1* (E-cadherin), *FBXW7*, *SMAD7*, *GATA5*, *VEGF*, *CBX4*, *ING3*, *CCR4N4L*, *BMI1*, *CDH19* and *SIAH2*, which represent each of the classified pathways (Figure 4C). Moreover, similar assays using specific antibodies against WDR77, DDB1, ROC1, H4R3me2s and H2AK119ub1 on selected co-target genes showed strong correlations between PHF1 and PRMT5/H4R3me2s or CUL4B/H2AK119ub1 enrichment, validating the ChIP-on-chip results (Figure 4D and E).

Regulation of E-cadherin and *FBXW7* by the PHF1/PRMT5–WDR77/CRL4B complex

For further investigation, we tested lentivirus-delivered short hairpin RNA (shRNA) packages targeted to *PHF1*, *PRMT5*, and *CUL4B* mRNA (Figure 5A), and chose the most efficient ones (marked in red) for the following experiments. Among the common target genes identified earlier, E-cadherin is a well-established tumor suppressor gene, which is involved in cell adhesion (37) and FBXW7 plays an important tumor suppressor role in cancer development and progression (38). We therefore further investigated the transcriptional regulation of E-cadherin and *FBXW7* by the PHF1/PRMT5/CRL4B complex. MDA-MB-231, Hs 578T, and MDA-MB-453 cell clones with PHF1, PRMT5 or CUL4B stably depleted were generated by the chosen lentivirus-delivered shRNA along with scrambled shRNA control (Figure 5B). Each of these shRNAs led to a significant reduction in the expression of their target genes without causing detectable changes to non-targeted genes. Knockdown of PHF1, PRMT5 or CUL4B in MDA-MB-231, Hs 578T and MDA-MB-453 cells led to increased expression of E-cadherin and *FBXW7* at both the transcriptional (Figure 5B) and protein levels (Figure 5C), as expected.

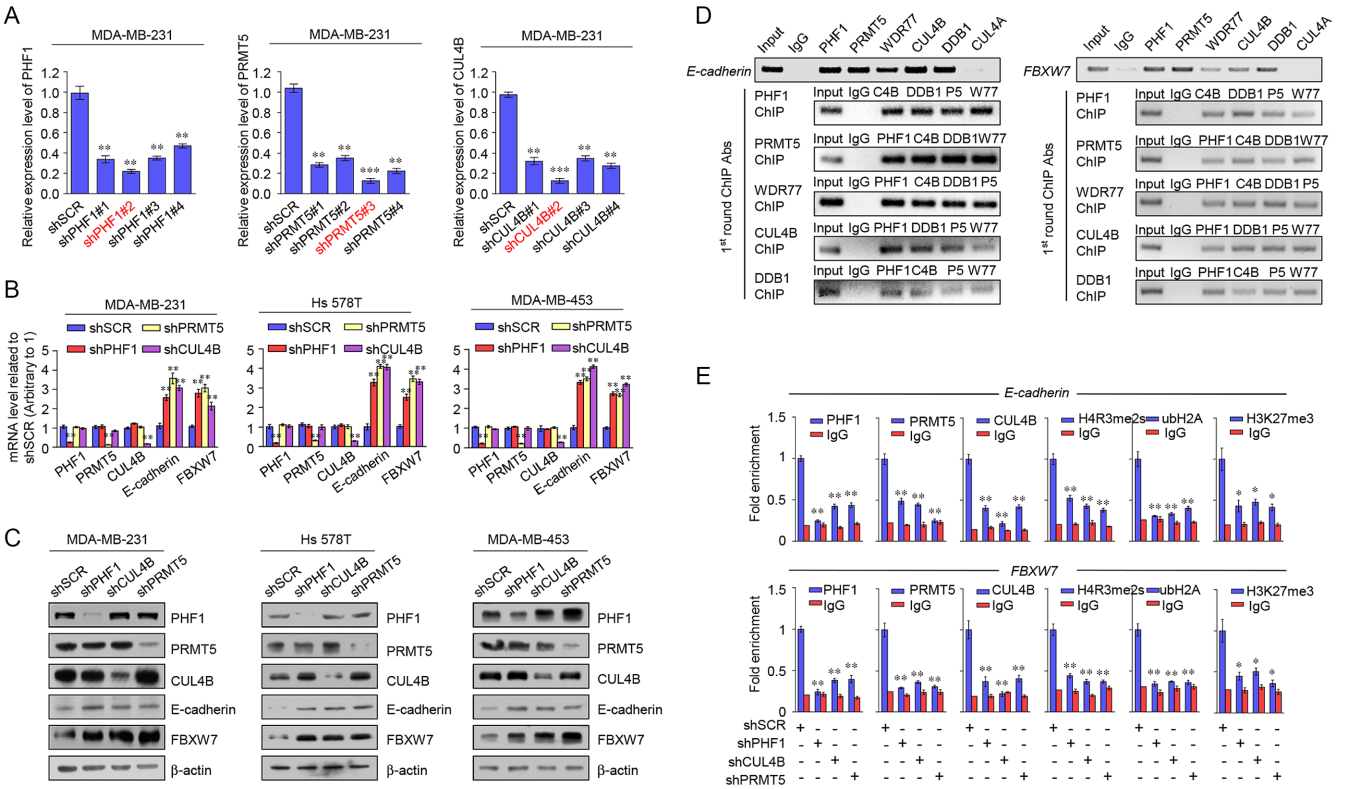

Figure 5. Tumor suppressor genes E-cadherin and FBXW7 are cotargeted by PHF1/PRMT5/CRL4B complex. (**A**) The efficiency of four shRNAs targeting *PHF1*, *PRMT5* or *CUL4B*, respectively. MDA-MB-231 cells were infected with lentiviruses carrying control shRNA (shSCR) or four different shRNAs targeting *PHF1*, *PRMT5* or *CUL4B*. The knockdown efficiencies of *PHF1*, *PRMT5*, and *CUL4B* were verified by RT-qPCR. We chose shPHF1#2, shPRMT5#3, and shCUL4B#2 (marked in red) for further study. Error bars represent mean ± SD for three independent experiments. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ (two-tailed unpaired *t* test). (**B** and **C**) Clones in which *PHF1*, *PRMT5*, or *CUL4B* were stably knocked down were compared with the parental cell lines with respect to the levels of mRNA and protein of E-cadherin and *FBXW7* in MDA-MB-231, Hs 578T and MDA-MB-453 cells. The mRNA levels were normalized to those of *GAPDH* and β-actin served as a loading control for western blotting. Error bars represent mean ± SD of three independent experiments. ($^{*}P < 0.05$, $^{**}P < 0.01$, and two-tailed unpaired *t* test). (**D**) PHF1, PRMT5–WDR77, and CRL4B complexes exist in the same protein complex on the E-cadherin and *FBXW7* promoters. CUL4A served as a negative control. ChIP and Re-ChIP experiments were performed in MDA-MB-231 cells with the indicated antibodies. (**E**) qChIP analysis of the recruitment of the indicated proteins on E-cadherin and *FBXW7* promoters in MDA-MB-231 cells after transfection with control shRNA or shRNAs targeting PHF1, PRMT5 or CUL4B. Purified rabbit IgG was used as a negative control. Error bars represent mean ± SD of three independent experiments ($^{*}P < 0.05$, $^{**}P < 0.01$, and two-tailed unpaired *t* test).

We next investigated the regulation of E-cadherin and *FBXW7* by the PHF1/PRMT5/CRL4B complex. We demonstrated that PHF1, PRMT5–WDR77, and CRL4B co-occupied the promoters of E-cadherin and *FBXW7* through ChIP assays using MDA-MB-231 cells and antibodies against PHF1, PRMT5, WDR77, CUL4B, DDB1 or control immunoglobulin G (IgG). To further support our proposition that PHF1, PRMT5–WDR77,and CRL4B function in the same protein complex at the E-cadherin and *FBXW7* promoters, we performed sequential ChIP or ChIP/Re-ChIP assays (31). In these experiments, soluble chromatin was first immunoprecipitated with antibodies against PHF1, PRMT5, WDR77, CUL4B or DDB1, while CUL4A served as a negative control. The immunoprecipitates were subsequently re-immunoprecipitated with appropriate antibodies. The results showed that the E-cadherin and *FBXW7* promoters that were immunoprecipitated with antibodies against PHF1 could be re-immunoprecipitated with antibodies against PRMT5, WDR77, CUL4B or DDB1. Similar results were obtained when an initial ChIP was performed with antibodies against PRMT5, WDR77, CUL4B or DDB1 (Figure 5D). These results support the idea that PHF1, PRMT5–WDR77 and CRL4B occupy the E-cadherin and *FBXW7* promoters in one concerted complex.

In order to identify a functional association between PHF1, PRMT5, and CRL4B on E-cadherin and *FBXW7* promoters, MDA-MB-231 cell clones with PHF1, PRMT5 or CUL4B stably depleted were used. qChIP analyses showed that knockdown of the expression of PHF1, PRMT5 or CUL4B led to a significant reduction in the binding of corresponding proteins to the promoters of E-cadherin and *FBXW7* (Figure 5E), implying that PHF1, PRMT5, and CUL4B act as a whole and each component is essential for the complex to bind to chromatin. Notably, except for their own products, knockdown of PRMT5 resulted in a marked reduction in the level of H2AK119ub1, while knockdown of CUL4B led to a reduction of H4R3me2s at the promoters of E-cadherin and *FBXW7*, suggesting that PRMT5 and CUL4B affect each other`s enzymatic activity. Otherwise, the levels of H4R3me2s and H2AK119ub1 were markedly decreased at the promoters of E-cadherin and

FBXW7 upon knockdown of PHF1 (Figure 5E), suggesting that PHF1 acts coordinately with PRMT5-mediated H4R3me2s and CRL4B-catalyzed H2AK119ub1. Furthermore, knockdown of PHF1, PRMT5, or CUL4B led to a significant reduction in H3K27me3 binding to the promoters (Figure 5E). It has been reported that CRL4B and PHF1 are essential for EZH2 recruitment (29,39). In light of our results, we believe EZH2 recruitment is a down-stream event following PHF1–PRMT5–CRL4B interaction.

PHF1 promotes the proliferation and invasion of breast cancer cells

As previously stated, E-cadherin and FBXW7 have been proven to play important roles in proliferation and invasion of tumor cells and are considered to be tumor suppressors. Based on the roles of CRL4B and PRMT5 in tumorigenesis and cancer progression (29,40,41) and our observation that PHF1/PRMT5/CRL4B functions as a complex, we next investigated what role PHF1 plays in tumorigenesis and metastasis in breast cancer. For gain-of-function, breast cancer cells were infected with lentivirus delivering a PHF1 CDS; and for-loss-of-function, PHF1 stably depleted breast cancer cell lines were generated using two different lentivirus-delivered shRNAs. The shRNAs led to a significant reduction in the expression of their target genes as previously observed. We also designed a shPHF1#2-resistant PHF1 overexpressed lentivirus (PHF1 res) to perform rescue experiments and to exclude off-target effects. We first performed growth assays in cells with PHF1 gain-of-function and loss-of-function. The resulting growth curves showed that PHF1 overexpression was associated with increased cell proliferation rate, whereas PHF1 knockdown was associated with a significant decrease in cell numbers, which could be rescued by the re-expression of shRNA-resistant PHF1. We also found that both of the H4R3me2s-binding mutants, H115A and C136A, could partially rescue the cell proliferation in PHF1-depleted cells than the wild type did (Figure 6A). Moreover, colony formation assays showed that PHF1 overexpression was associated with a marked increase in colony numbers, whereas PHF1 knockdown was associated with a significant decrease in colony numbers, which could be rescued by the re-expression of shRNA-resistant PHF1 (Figure 6B). In addition, shRNA-resistant PHF1 rescued the expression of E-cadherin and *FBXW7* in PHF1-depleted cells, while both the H4R3me2s-binding mutants H115A and C136A could partially rescue the target genes in PHF1-depleted cells than the wild type did (Figure 6C). Next, MDA-MB-231 and Hs 578T cells were treated with EdU for 2 hours. The results of this treatment revealed that the PHF1 knockdown cells showed a much lower percentage of EdU-labeled cells (Figure 6D). Together, data from these experiments support the hypothesis that PHF1 promotes the proliferation of breast cancer.

As reported earlier, E-cadherin has been proven to participate in EMT (42). Therefore, we next investigated whether PHF1 plays a role in EMT and tumor metastasis. For this purpose, we detected the expression levels of some invasion markers of EMT under the influence of loss-of-function of PHF1. qPCR and western blot analysis showed that the expression of some epithelial markers (α-catenin, γ-catenin, and E-cadherin) increased, while certain mesenchymal markers (N-cadherin, Fibronectin, and Vimentin) decreased (Figure 6E and F). Moreover, the results from transwell invasion assays in three breast cancer cell lines (MCF-7, MDA-MB-231 and Hs 578T) showed that the overexpression of PHF1 resulted in a greater than two-fold increase in cell invasion, while knockdown of PHF1 in the MDA-MB-231 and Hs 578T cells by two separate shRNAs resulted in apparent decreases in cell invasion potential, which could be partially rescued by co-knockdown of E-cadherin, indicating that PHF1 could promote breast cancer metastasis through transcriptional repression of E-cadherin (Figure 6G). Collectively, these results indicate that PHF1 promotes the invasive potential of breast cancer cells, possibly by repressing the expression of E-cadherin.

In order to investigate the role of PHF1 in tumor development and progression *in vivo*, MDA-MB-231 cells engineered to stably express firefly luciferase (MDA-MB-231-Luc-D3H2LN, Xenogen Corporation) were co-infected with lentiviruses carrying PHF1 shRNA. These cells were then orthotopically implanted or intravenously injected into immunocompromised SCID mice to measure orthotopic tumorigenesis or seeding lung metastasis, respectively. Specifically, MDA-MB-231-Luc-D3H2LN cells were either implanted into the abdominal mammary fat pad ($n = 6$) or injected into the lateral tail vein ($n = 6$) of 6-week-old female SCID mice. The growth/dissemination of tumors was monitored weekly by bioluminescence imaging with an IVIS imaging system (Xenogen Corporation). Orthotopic tumors and lung metastases were measured using quantitative bioluminescence imaging after 8 weeks. The results showed that, in orthotopically implanted groups, knockdown of PHF1 resulted in a significant reduction in MDA-MB-231-Luc-D3H2LN tumor growth and metastatic potential (Figure 6H and I). In addition, in the intravenous groups, PHF1 knockdown led to a dramatic decrease in lung metastasis of the MDA-MB-231-Luc-D3H2LN tumors (Figure 6J and K). The metastases in lung tissue were verified by bioluminescence imaging. In conclusion, these experiments demonstrated that PHF1 promotes tumorigenesis and metastasis of breast cancer cells *in vitro* and *in vivo*.

Expression of PHF1 is upregulated in breast cancer and negatively correlated with expression of E-cadherin and FBXW7

In order to further investigate the role of PHF1 in breast cancer progression, we collected 81 breast cancer samples and adjacent normal tissues from breast cancer patients and performed tissue microarrays by means of immunohistochemical staining. PHF1 was found to be significantly upregulated in tumors and its expression appeared to be positively correlated with histological grades (Figure 7A and B). Next, we analyzed published data downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/). The results support the idea that PHF1 is more highly expressed in breast cancer cells than in adjacent normal tissues (Figure 7C). In addition, its expression gradually increases across subtypes of breast cancer including Luminal A, Luminal B, Her2-enriched and Basal-like (Figure 7D).

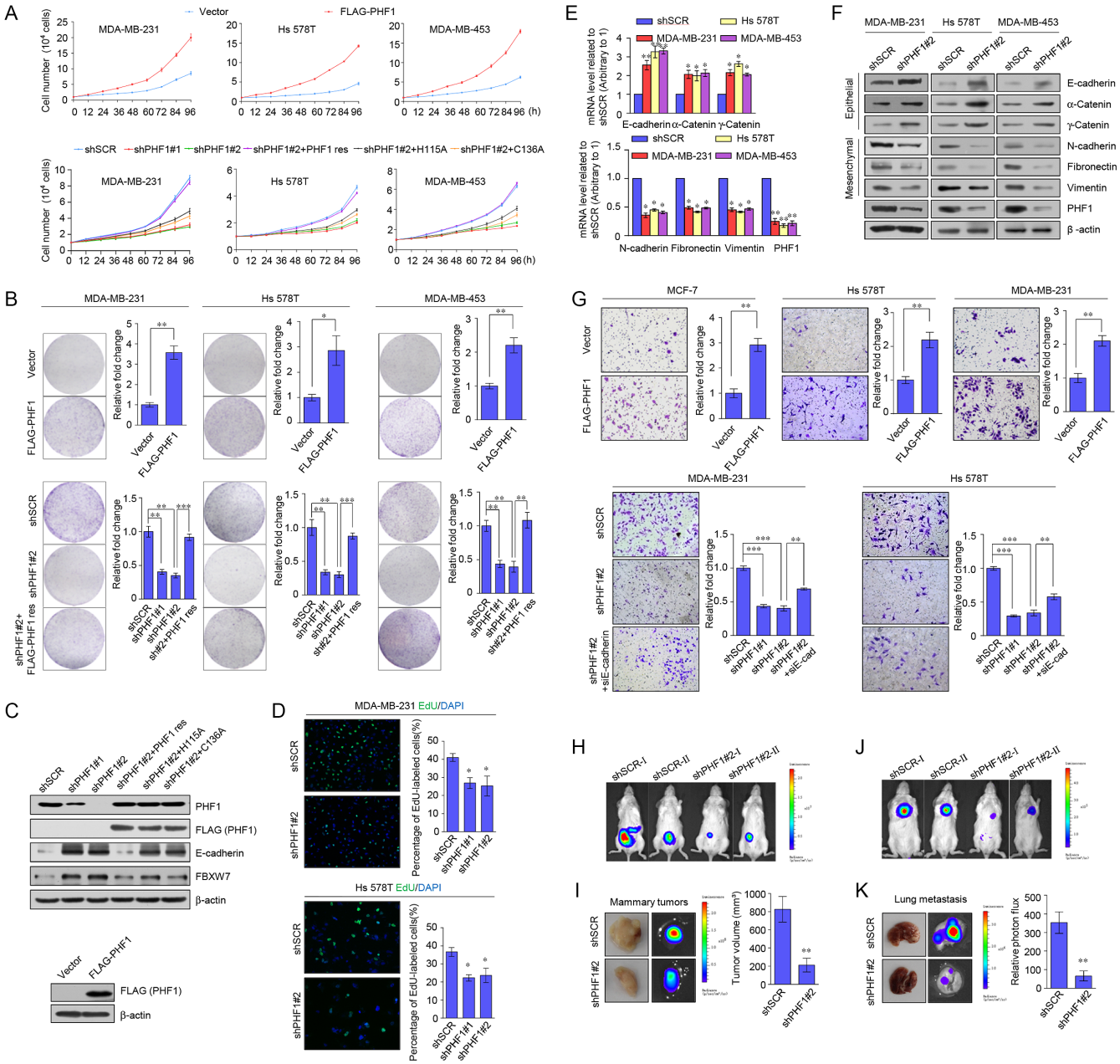

Figure 6. PHF1 promotes carcinogenesis and metastasis of breast cancer *in vitro* and *in vivo*. (**A**) PHF1 promotes cellular proliferation. Growth curve analysis was performed on MDA-MB-231, Hs 578T, and MDA-MB-453 cells transfected with vector or FLAG-PHF1, or transfected with shSCR, two different PHF1 shRNAs or together with shRNA-resistant FLAG-PHF1 wild type (FLAG-PHF1 res) or H115A, C136A mutants. (**B**) PHF1 enhances the colony-forming efficiency of breast cancer cells. MDA-MB-231, Hs 578T and MDA-MB-453 cells transfected with vector or FLAG-PHF1 were maintained for 4 days while cells transfected with shSCR, two different shRNAs of PHF1 or together with shRNA-resistant FLAG-PHF1 (FLAG-PHF1 res) were maintained in culture media for 7 days prior to crystal violet staining. Representative photos and statistical analyses are shown. (**C**) Western blot analysis of PHF1 expression in MDA-MB-231 cells transfected with different lentiviruses. β-actin served as a loading control. (**D**) MDA-MB-231 cells and Hs 578T cells were incubated with EdU for 2 h. A fluorescence microscope was used for EdU detection. (**E** and **F**) Expression of the indicated epithelial or mesenchymal markers was measured by RT-qPCR or western blotting in MDA-MB-231, Hs 578T and MDA-MB-453 cells with PHF1 depleted. (**G**) MCF-7, MDA-MB-231, and Hs 578T cells were transfected with vector or FLAG-PHF1 for cell invasion assays using Matrigel transwell filters. MDA-MB-231 cells and Hs 578T cells transfected with shSCR, two different shRNAs of *PHF1* or/and siE-cadherin were used for the same assays. The invading cells were stained and counted. The images represent one field under microscopy in each group. (**H** and **I**) MDA-MB-231-Luc-D3H2LN cells infected with lentiviruses carrying shSCR or shRNA against *PHF1* were inoculated orthotopically into the abdominal mammary fat pad of 6-week-old female SCID mice ($n = 6$). Primary tumors were quantified using bioluminescence imaging 6 weeks after initial implantation. Representative *in vivo* bioluminescent images are shown (H), and tumor specimens were examined by *in vitro* bioluminescent measurement (I). (**J** and **K**) The above-described MDA-MB-231-Luc-D3H2LN cells were injected intravenously through the tail vein of 6-week-old female SCID mice ($n = 6$). Lung metastases were quantified using bioluminescence imaging after 6 weeks. Representative *in vivo* bioluminescent images are shown (J). Lung cancer specimens were examined by *in vitro* bioluminescent measurement (K). (A–K) Two-tailed unpaired *t* test ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

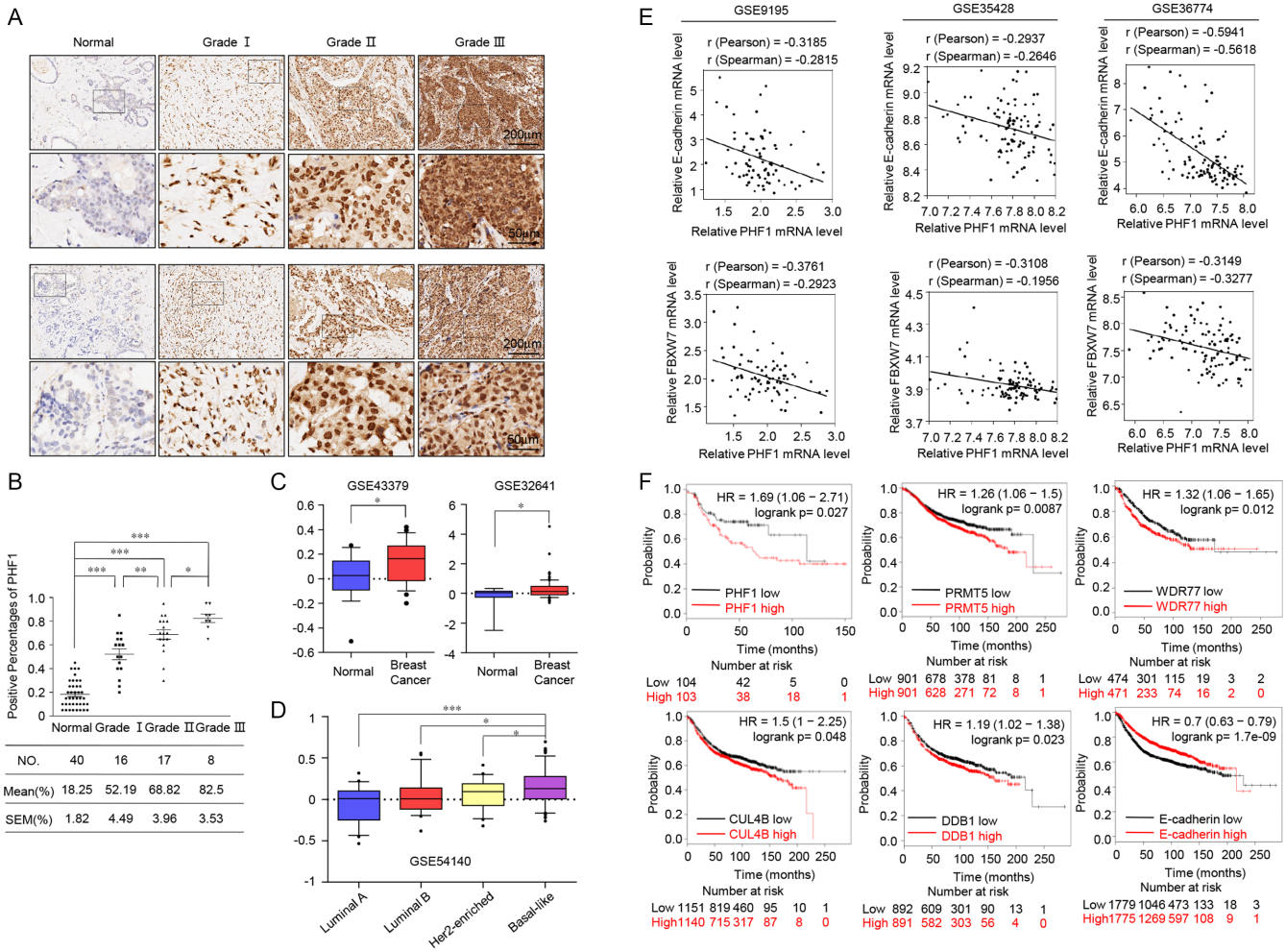

Figure 7. Expression of PHF1 is upregulated in breast cancer and negatively correlated with expression of E-cadherin and FBXW7. (**A**) Immunohistochemical staining of PHF1 in normal breast tissues and breast carcinomas (histological grades I, II, and III). (**B**) Positively stained nuclei (in percentages) in grouped samples were analyzed by two-tailed unpaired *t* test ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). (**C** and **D**) Analysis of public datasets (GSE43379, GSE32641 and GSE54140) for the expression of PHF1 by two-tailed unpaired *t* test ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). (**E**) Analysis of public datasets (GSE9195, GSE35428, and GSE36774) for the expression of PHF1 and E-cadherin/FBXW7 in breast cancer. The relative levels of E-cadherin or FBXW7 were plotted against that of PHF1. (**F**) Kaplan–Meier survival analysis for the relationship between survival time and PHF1, PRMT5, WDR77, CUL4B, DDB1 and E-cadherin signature in breast cancer using the online tool (http://kmplot.com/analysis/).

Analysis of three published clinical datasets (GSE9195, GSE35428 and GSE36774) revealed statistically significant negative correlation between the expression of E-cadherin or FBXW7 and the expression of PHF1, which supports our findings that E-cadherin and *FBXW7* were transcriptionally regulated by PHF1 (Figure 7E). To further extend our observations to a clinical and pathologically relevant context, we analyzed the expression of PHF1 and its correlation with clinical behaviors of breast cancer patients. Kaplan-Meier survival analysis of the relationship between survival time and the expression of PHF1, PRMT5, WDR77, CUL4B, DDB1, and E-cadherin was performed with the online tool (http://kmplot.com/analysis/). This showed that lower expression of PHF1 ($P = 0.027$), PRMT5 ($P = 0.0087$), WDR77 ($P = 0.012$), CUL4B ($P = 0.048$), and DDB1 ($P = 0.023$) with higher expression of E-cadherin ($P = 1.7e{-}09$) were associated with improved survival in breast cancer patients, when the influence of systemic treatment, endocrine therapy, and chemotherapy were excluded (Figure 7F). These data are consistent with PHF1 playing a role in promoting breast cancer progression and support the observation that E-cadherin and *FBXW7* are downstream targets of PHF1.

Expression of PHF1 is upregulated in multiple carcinomas and is a potential cancer biomarker

To investigate whether the effect of PHF1 could be extended to a broader scope of cancers, we collected a series of carcinoma samples from breast, lung, colon, stomach, ovarian, prostate, esophageal, liver, pancreatic, renal, rectal and uterine cancer patients, with each type of carcinoma having at least twelve samples paired with adjacent normal tissues. Tissue microarray analysis by immunohistochemical staining showed a statistically significant upregulation of PHF1 expression in carcinomas from multiple tissues compared to the adjacent normal tissues (Figure 8A and B). Next, we de-

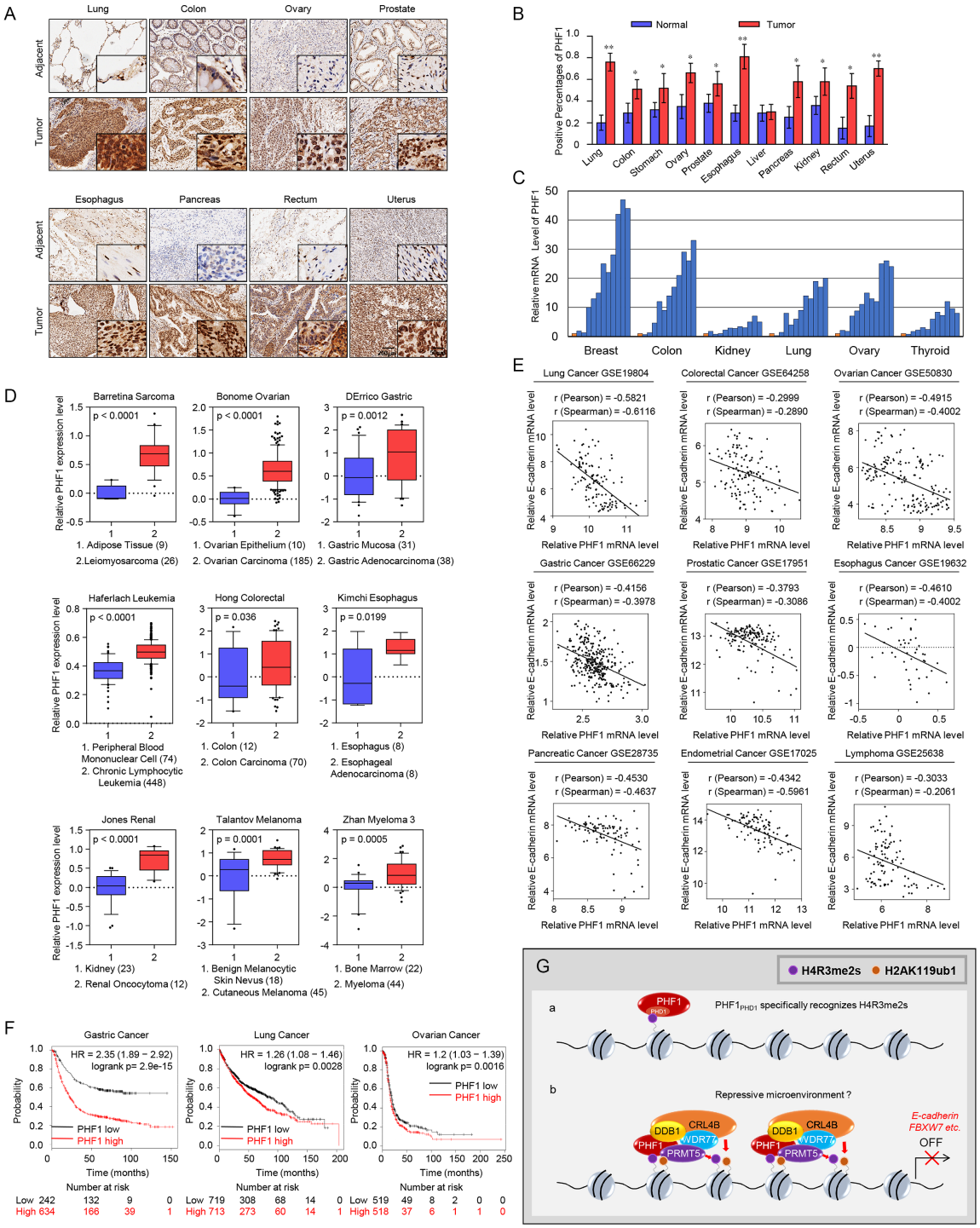

Figure 8. Expression of PHF1 is upregulated in multiple carcinomas and is a potential cancer biomarker. (**A** and **B**) PHF1 is upregulated in multiple carcinomas. Immunohistochemical staining of PHF1 in paired samples of breast, lung, colon, stomach, ovary, prostate, esophageal, liver, pancreas, kidney, rectal and uterine carcinoma versus adjacent normal tissues. Representative images of 200-fold magnifications of each type of paired tumor section are presented. Each bar represents the mean ± SD for triplicate experiments ($^{*}P < 0.05$ and $^{**}P < 0.01$). (**C**) Transcript levels of *PHF1* are significantly higher in multiple malignant tumor tissues. Relative levels of *PHF1* transcripts were normalized to that of *GAPDH* and calibrated to the mean mRNA level (arbitrary value of 1) in normal tissue (yellow bars). The fold increase in gene expression relative to the mean value for each disease sample is indicated. (**D**) PHF1 expression in multiple cancer microarray datasets available from Oncomine (https://www.oncomine.com/). (**E**) Analysis of public datasets (GSE19804, GSE64258, GSE50830, GSE66229, GSE17951, GSE19632, GSE28735, GSE17025 and GSE25638) for the expression of *PHF1* and E-cadherin in multiple carcinomas. The relative level of E-cadherin was plotted against that of *PHF1*. (**F**) Kaplan-Meier survival analysis for the relationship between survival time and *PHF1* signature in gastric cancer, lung cancer and ovarian cancer using the online tool. (**G**) Graphic model as discussed in the text. DNA (black line); nucleosomes with single N-terminus of H4 and C-terminus of H2A (blue ball). (a) PHF1 recognized the H4R3me2s modification. (b) PRMT5/WDR77 and CRL4B were recruited by PHF1, and the PHF1/PRMT5–WDR77/CRL4B complex was formed on the target promoters, which was stable and further catalyzed H4R3me2s/H2AK119ub1 formation to establish, as well as maintain repression at these sites.

termined the expression profile of PHF1 using the human cancer survey RT-qPCR gene expression panel (Origene) in tumors originating from breast, colon, kidney, lung, ovary, and thyroid tissue (Figure 8C). The results showed that PHF1 is overexpressed in all of these cancer groups compared to normal tissue controls. The oncomine database (https://www.oncomine.com/) also showed significant correlation of PHF1 expression with leiomyosarcoma, ovarian cancer, gastric cancer, leukemia, colon cancer, esophageal cancer, renal cancer, melanoma, and myeloma (Figure 8D). Further, analysis of nine published clinical datasets revealed statistically significant negative correlation between the expression of E-cadherin and PHF1 expression in multiple carcinomas (Figure 8E). Kaplan–Meier survival analysis showed that lower expression of PHF1 was associated with an improved overall survival in gastric, lung and ovarian cancer patients (Figure 8F). In summary, our analysis showed that PHF1 is upregulated in multiple carcinomas and is a potential cancer biomarker.

DISCUSSION

PHF1 was identified several years ago as a component of the PRC2 complex and is required to generate high levels of H3K27 trimethylation at the promoters of Polycomb target genes (43). Besides PRC2, our affinity purification revealed several PHF1-interacting epigenetic factors and complexes including PRMT5–WDR77 complex, CRL4B complex, SIN3A/HDAC complex, NuRD complex, and AIB1, indicating that it may participate in various epigenetic transcriptional regulation pathways. Specifically, PRMT5–WDR77 complex targets H4R3 for symmetric dimethylation (16), CRL4B catalyzes H2AK119ub1 (29), and the SIN3A/HDAC complex and NuRD complex possess histone deacetylation activity (44,45). All of these have primary function in transcriptional repression programs by virtue of their enzymatic activities and through their chromatin remodelling capabilities, suggesting that PHF1 may participate in a broad variety of epigenetic repression processes. Interestingly, in contrast to those repressive complexes, AIB1 (amplified in breast cancer-1, also called SRC-3 and NCoA-3), is a member of the p160 nuclear receptor co-activator family and is considered an important oncogene in breast cancer through transcriptional activation (46). Increased AIB1 levels in human breast cancer have been correlated with poor clinical prognosis. Although mass spectrometry only identified AIB1/NCOA3, co-IP assays confirmed that all three member of the p160 family, SRC-1, GRIP-1 and AIB1 have physical interactions with PHF1, indicating that PHF1 may respond to estrogen or androgen stimulation or be involved in relevant nuclear receptor physiological or pathological pathways (47). Considering the fact that WDR77 is an androgen-associated protein and PRMT5–WDR77 complex functions as a coactivator of androgen receptor (48), the precise function of the interaction of PHF1 and p160 proteins remains to be elucidated. It is not expected that PHF1 has relationships with all types of epigenetic regulators, but it will not be surprising if future investigations uncover additional epigenetic regulators that are associated with PHF1.

PHF1 was previously reported to be an epigenetic effector, which recognizes H3K36me3 through its Tudor domain to synergize PRC2's activity on H3K27 tri-methylation (13,14). Our study focused on the function of the PHD fingers of PHF1. PHD motifs of different proteins could bind to several methylated histone marks, including H3K4me2/3 (49–53), H3K36me3 (13,14), H3K9me3 (54), H3R2me2s (55) and H4R3me2s (56,57), indicating that PHD finger-containing proteins have emerging roles as important epigenetic effectors. These effectors (or 'readers') induce downstream functional consequences, including (i) presenting substrate histone residues to enzymes (e.g., the SANT domain could present *in vivo* histone substrates to Gcn5p-containing HAT complexes) (58); (ii) introduction of further post-translational modifications through interactions with 'writers' (e.g., HP1 could recognize H3K9me2 and recruit DNMT1 for further transcriptional silencing) (59). We demonstrated that the N-terminal PHD finger of PHF1 recognizes the symmetric dimethylated H4R3 catalyzed by PRMT5–WDR77 (Figure 8G-a), and the C-terminus PHD finger of PHF1, instead of binding modified histones, recruits the 'writer' CRL4B for H2AK119 mono-ubiquitination. We showed that PHF1, PRMT5–WDR77, and CRL4B reciprocally interact with one another and collaborate as a functional unit. We revealed that the enrichment of PHF1, PRMT5, or CUL4B at the promoters of a series of target genes such as E-cadherin and *FBXW7* was greatly decreased after the knockdown of any of these, supporting the hypothesis that PHF1, PRMT5–WDR77 and CRL4B act as a functional complex. Based on these findings, we speculate that in the absence of PHF1, PRMT5–WDR77 complex at promoters would not be stably tethered, and this also affects CRL4B recruitment. However, when PHF1 is present, it recognizes and binds H4R3me2s catalyzed by PRMT5–WDR77 and recruits CRL4B. The PHF1/PRMT5–WDR77/CRL4B complexes were formed on the targeted promoters, which creates an H4R3me2s/H2AK119ub1 coordinated repressive 'histone code' on gene promoters to transcriptionally repress their target genes (Figure 8G-b). The collaboration of newly identified H4R3me2s and H2AK119ub1 shed new light on the functional interactions of distinct enzymatic activities and epigenetic transcriptional regulation mechanisms. Future investigations are needed to explore the scope and the exact molecular mechanism of the formation of the PHF1/PRMT5–WDR77/CRL4B complex, and to determine whether this functionality involves additional elements.

It has been reported by several groups that *PHF1* was fused with several genes such as *JAZF1*, *EPC1* and *MEAF6* in endometrial stromal sarcomas (60,61). The chimeric fusion proteins commonly contain entire PHF1 protein sequences in the C-termini. Thus, the coding region of *PHF1* was brought under control of a different promoter; that is, it was regulated via the translocation partner's promoter. Therefore, altered *PHF1* regulation may possibly contribute to the pathogenetic effects of these fusion genes. However, the involvement of PHF1 in tumorigenesis has not been clarified. In our study, ChIP-Sequencing identified genes targeted by PHF1, which are involved in various oncogenic signaling pathways including MAPK, Jak-STAT, Wnt, ad-

heren junction, TGF-beta, cell cycle, and pathways in cancer and many of them are known tumor suppressors. Both PRMT5 and CRL4B are reported to promote carcinogenesis through transcriptional regulation and correlate with malignancy grade of various kinds of tumors (29,30,62–64). Among the overlapping targets of PHF1/PRMT5/CRL4B, we chose E-cadherin and *FBXW7* as representative target genes for further study. E-cadherin has been reported to be a single-pass transmembrane protein that mediates homophilic cell–cell interactions and its loss is often associated with tumor progression (65). E-cadherin plays an important role in the EMT and is regarded as a 'caretaker' of the epithelial phenotype (66). FBXW7 is the substrate recognition subunit of SCF E3 ubiquitin ligase which contains F-box and WD40 repeat domains (67). It is now considered to be a tumor suppressor and mutation or deficiency of FBXW7 occurs in diverse human tumor types (68). It has been reported that many oncogenes and oncogenic transcription factors including Cyclin E, c-Myc, c-Jun, Notch, KLF5, mTOR and NF1 are the degradation substrates of FBXW7 and its complex (38,69). Our findings have revealed that PHF1 promotes the proliferation and invasion of breast cancer cells and it does so, at least in part, through repression of the powerful tumor suppressors, E-cadherin and FBXW7. We have also shown that PHF1 expression was significantly higher in various kinds of tumor samples compared to adjacent normal tissues and that the level of PHF1 expression was negatively correlated with E-cadherin and FBXW7 expression. Finally, yet importantly, we found that PHF1 is negatively associated with patient survival in breast, gastric, lung and ovarian cancers, suggesting that PHF1 may be used as a potential cancer biomarker.

In summary, our study has revealed that PHF1 is a novel reader of H4R3me2s. We showed that PHF1, PRMT5–WDR77 and CRL4B reciprocally interact with one another and collaborate as a functional unit, providing the 'H4R3me2s-H2AK119ub1' readout of a defined combinatorial transcriptionally repressive epigenetic signature, thus providing a new transcriptional regulation model and a new molecular basis for the interplay between histone ubiquitination and arginine methylation in chromatin remodeling. Our data demonstrate that PHF1 contributes to epigenetic silencing of tumor suppressors, and increased PHF1 levels in human cancer are correlated with poorer clinical prognosis. Our data indicate that PHF1 is a potential oncogene as well as a promising biomarker, supporting the pursuit of PHF1 as a target for cancer therapy.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

AKNOWLEDGEMENTS

We thank the Core Facilities Center, Capital Medical University for the assistance of SPR analysis in this study.

FUNDING

Ministry of Science and Technology of China [2016YFA0102400 to Y.W.]; National Natural Science Foundation of China [81773017, 81472733 to Y.W., 81402334 to Y.Y., 81502446 to R.Q.]. Funding for open access charge: National Natural Science Foundation of China.
Conflict of interest statement. None declared.

REFERENCES

1. Lu,R. and Wang,G.G. (2013) Tudor: a versatile family of histone methylation 'readers'. *Trends Biochem. Sci.*, **38**, 546–555.
2. Taverna,S.D., Li,H., Ruthenburg,A.J., Allis,C.D. and Patel,D.J. (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.*, **14**, 1025–1040.
3. Yun,M., Wu,J., Workman,J.L. and Li,B. (2011) Readers of histone modifications. *Cell Res.*, **21**, 564–578.
4. Bedford,M.T. and Clarke,S.G. (2009) Protein arginine methylation in mammals: who, what, and why. *Mol. Cell*, **33**, 1–13.
5. Duncan,I.M. (1982) Polycomblike: a gene that appears to be required for the normal expression of the bithorax and antennapedia gene complexes of Drosophila melanogaster. *Genetics*, **102**, 49–70.
6. Lonie,A., D'Andrea,R., Paro,R. and Saint,R. (1994) Molecular characterisation of the Polycomblike gene of Drosophila melanogaster, a trans-acting negative regulator of homeotic gene expression. *Development*, **120**, 2629–2636.
7. Tie,F., Prasad-Sinha,J., Birve,A., Rasmuson-Lestander,A. and Harte,P.J. (2003) A 1-megadalton ESC/E(Z) complex from Drosophila that contains polycomblike and RPD3. *Mol. Cell. Biol.*, **23**, 3352–3362.
8. Boulay,G., Rosnoblet,C., Guerardel,C., Angrand,P.O. and Leprince,D. (2011) Functional characterization of human Polycomb-like 3 isoforms identifies them as components of distinct EZH2 protein complexes. *Biochem J.*, **434**, 333–342.
9. Coulson,M., Robert,S., Eyre,H.J. and Saint,R. (1998) The identification and localization of a human gene with sequence similarity to Polycomblike of Drosophila melanogaster. *Genomics*, **48**, 381–383.
10. Cao,R., Wang,H., He,J., Erdjument-Bromage,H., Tempst,P. and Zhang,Y. (2008) Role of hPHF1 in H3K27 methylation and Hox gene silencing. *Mol. Cell. Biol.*, **28**, 1862–1872.
11. Sarma,K., Margueron,R., Ivanov,A., Pirrotta,V. and Reinberg,D. (2008) Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. *Mol. Cell. Biol.*, **28**, 2718–2731.
12. Hong,Z., Jiang,J., Lan,L., Nakajima,S., Kanno,S., Koseki,H. and Yasui,A. (2008) A polycomb group protein, PHF1, is involved in the response to DNA double-strand breaks in human cell. *Nucleic Acids Res.*, **36**, 2939–2947.
13. Cai,L., Rothbart,S.B., Lu,R., Xu,B., Chen,W.Y., Tripathy,A., Rockowitz,S., Zheng,D., Patel,D.J., Allis,C.D. *et al.* (2013) An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. *Mol. Cell*, **49**, 571–582.
14. Musselman,C.A., Avvakumov,N., Watanabe,R., Abraham,C.G., Lalonde,M.E., Hong,Z., Allen,C., Roy,S., Nunez,J.K., Nickoloff,J. *et al.* (2012) Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nat. Struct. Mol. Biol.*, **19**, 1266–1272.
15. Qin,S., Guo,Y., Xu,C., Bian,C., Fu,M., Gong,S. and Min,J. (2013) Tudor domains of the PRC2 components PHF1 and PHF19 selectively bind to histone H3K36me3. *Biochem. Biophys. Res. Commun.*, **430**, 547–553.
16. Branscombe,T.L., Frankel,A., Lee,J.H., Cook,J.R., Yang,Z., Pestka,S. and Clarke,S. (2001) PRMT5 (Janus kinase-binding protein 1) catalyzes the formation of symmetric dimethylarginine residues in proteins. *J. Biol. Chem.*, **276**, 32971–32976.
17. Migliori,V., Muller,J., Phalke,S., Low,D., Bezzi,M., Mok,W.C., Sahu,S.K., Gunaratne,J., Capasso,P., Bassi,C. *et al.* (2012) Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat. Struct. Mol. Biol.*, **19**, 136–144.
18. Pal,S., Baiocchi,R.A., Byrd,J.C., Grever,M.R., Jacob,S.T. and Sif,S. (2007) Low levels of miR-92b/96 induce PRMT5 translation and H3R8/H4R3 methylation in mantle cell lymphoma. *EMBO J.*, **26**, 3558–3569.

19. Pal,S., Vishwanath,S.N., Erdjument-Bromage,H., Tempst,P. and Sif,S. (2004) Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes. *Mol. Cell. Biol.*, **24**, 9630–9645.
20. Tee,W.W., Pardo,M., Theunissen,T.W., Yu,L., Choudhary,J.S., Hajkova,P. and Surani,M.A. (2010) Prmt5 is essential for early mouse development and acts in the cytoplasm to maintain ES cell pluripotency. *Genes Dev.*, **24**, 2772–2777.
21. Dacwag,C.S., Bedford,M.T., Sif,S. and Imbalzano,A.N. (2009) Distinct protein arginine methyltransferases promote ATP-dependent chromatin remodeling function at different stages of skeletal muscle differentiation. *Mol. Cell. Biol.*, **29**, 1909–1921.
22. Scoumanne,A., Zhang,J. and Chen,X. (2009) PRMT5 is required for cell-cycle progression and p53 tumor suppressor function. *Nucleic Acids Res.*, **37**, 4965–4976.
23. Hanahan,D. and Weinberg,R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646–674.
24. Hou,Z., Peng,H., Ayyanathan,K., Yan,K.P., Langer,E.M., Longmore,G.D. and Rauscher,F.J. 3rd (2008) The LIM protein AJUBA recruits protein arginine methyltransferase 5 to mediate SNAIL-dependent transcriptional repression. *Mol. Cell. Biol.*, **28**, 3198–3207.
25. Kryukov,G.V., Wilson,F.H., Ruth,J.R., Paulk,J., Tsherniak,A., Marlow,S.E., Vazquez,F., Weir,B.A., Fitzgerald,M.E., Tanaka,M. *et al.* (2016) MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells. *Science*, **351**, 1214–1218.
26. Jackson,S. and Xiong,Y. (2009) CRL4s: the CUL4-RING E3 ubiquitin ligases. *Trends Biochem. Sci.*, **34**, 562–570.
27. Zou,Y., Liu,Q., Chen,B., Zhang,X., Guo,C., Zhou,H., Li,J., Gao,G., Guo,Y., Yan,C. *et al.* (2007) Mutation in CUL4B, which encodes a member of cullin-RING ubiquitin ligase complex, causes X-linked mental retardation. *Am. J. Hum. Genet.*, **80**, 561–566.
28. Jiang,B., Zhao,W., Yuan,J., Qian,Y., Sun,W., Zou,Y., Guo,C., Chen,B., Shao,C. and Gong,Y. (2012) Lack of Cul4b, an E3 ubiquitin ligase component, leads to embryonic lethality and abnormal placental development. *PLoS One*, **7**, e37070.
29. Hu,H., Yang,Y., Ji,Q., Zhao,W., Jiang,B., Liu,R., Yuan,J., Liu,Q., Li,X., Zou,Y. *et al.* (2012) CRL4B catalyzes H2AK119 monoubiquitination and coordinates with PRC2 to promote tumorigenesis. *Cancer Cell*, **22**, 781–795.
30. Yang,Y., Liu,R., Qiu,R., Zheng,Y., Huang,W., Hu,H., Ji,Q., He,H., Shang,Y., Gong,Y. *et al.* (2015) CRL4B promotes tumorigenesis by coordinating with SUV39H1/HP1/DNMT3A in DNA methylation-based epigenetic silencing. *Oncogene*, **34**, 104–118.
31. Wang,Y., Zhang,H., Chen,Y., Sun,Y., Yang,F., Yu,W., Liang,J., Sun,L., Yang,X., Shi,L. *et al.* (2009) LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. *Cell*, **138**, 660–672.
32. He,Y.J., McCall,C.M., Hu,J., Zeng,Y. and Xiong,Y. (2006) DDB1 functions as a linker to recruit receptor WD40 proteins to CUL4-ROC1 ubiquitin ligases. *Genes Dev.*, **20**, 2949–2954.
33. Wang,S., Robertson,G.P. and Zhu,J. (2004) A novel human homologue of Drosophila polycomblike gene is up-regulated in multiple cancers. *Gene*, **343**, 69–78.
34. Angers,S., Li,T., Yi,X., MacCoss,M.J., Moon,R.T. and Zheng,N. (2006) Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature*, **443**, 590–593.
35. Li,T., Chen,X., Garbutt,K.C., Zhou,P. and Zheng,N. (2006) Structure of DDB1 in complex with a paramyxovirus V protein: viral hijack of a propeller cluster in ubiquitin ligase. *Cell*, **124**, 105–117.
36. Edmunds,J.W., Mahadevan,L.C. and Clayton,A.L. (2008) Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J.*, **27**, 406–420.
37. Christofori,G. and Semb,H. (1999) The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene. *Trends Biochem. Sci.*, **24**, 73–76.
38. Wang,Z., Inuzuka,H., Zhong,J., Wan,L., Fukushima,H., Sarkar,F.H. and Wei,W. (2012) Tumor suppressor functions of FBW7 in cancer development and progression. *FEBS Lett.*, **586**, 1409–1418.
39. Li,H., Liefke,R., Jiang,J., Kurland,J.V., Tian,W., Deng,P., Zhang,W., He,Q., Patel,D.J., Bulyk,M.L. *et al.* (2017) Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature*, **549**, 287–291.
40. Ibrahim,R., Matsubara,D., Osman,W., Morikawa,T., Goto,A., Morita,S., Ishikawa,S., Aburatani,H., Takai,D., Nakajima,J. *et al.* (2014) Expression of PRMT5 in lung adenocarcinoma and its significance in epithelial-mesenchymal transition. *Hum. Pathol.*, **45**, 1397–1405.
41. Tabata,T., Kokura,K., Ten Dijke,P. and Ishii,S. (2009) Ski co-repressor complexes maintain the basal repressed state of the TGF-beta target gene, SMAD7, via HDAC3 and PRMT5. *Genes Cells*, **14**, 17–28.
42. Schmalhofer,O., Brabletz,S. and Brabletz,T. (2009) E-cadherin, beta-catenin, and ZEB1 in malignant progression of cancer. *Cancer Metast. Rev.*, **28**, 151–166.
43. Nekrasov,M., Klymenko,T., Fraterman,S., Papp,B., Oktaba,K., Kocher,T., Cohen,A., Stunnenberg,H.G., Wilm,M. and Muller,J. (2007) Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *EMBO J.*, **26**, 4078–4088.
44. Dannenberg,J.H., David,G., Zhong,S., van der Torre,J., Wong,W.H. and Depinho,R.A. (2005) mSin3A corepressor regulates diverse transcriptional networks governing normal and neoplastic growth and survival. *Genes Dev.*, **19**, 1581–1595.
45. Xue,Y., Wong,J., Moreno,G.T., Young,M.K., Cote,J. and Wang,W. (1998) NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol. Cell*, **2**, 851–861.
46. Lahusen,T., Henke,R.T., Kagan,B.L., Wellstein,A. and Riegel,A.T. (2009) The role and regulation of the nuclear receptor co-activator AIB1 in breast cancer. *Breast Cancer Res. Treat.*, **116**, 225–237.
47. Xu,J., Wu,R.C. and O'Malley,B.W. (2009) Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family. *Nat. Rev. Cancer*, **9**, 615–630.
48. Hosohata,K., Li,P., Hosohata,Y., Qin,J., Roeder,R.G. and Wang,Z. (2003) Purification and identification of a novel complex which is involved in androgen receptor-dependent transcription. *Mol. Cell. Biol.*, **23**, 7019–7029.
49. Li,H., Ilin,S., Wang,W., Duncan,E.M., Wysocka,J., Allis,C.D. and Patel,D.J. (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature*, **442**, 91–95.
50. Pena,P.V., Davrazou,F., Shi,X., Walter,K.L., Verkhusha,V.V., Gozani,O., Zhao,R. and Kutateladze,T.G. (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature*, **442**, 100–103.
51. Shi,X., Hong,T., Walter,K.L., Ewalt,M., Michishita,E., Hung,T., Carney,D., Pena,P., Lan,F., Kaadige,M.R. *et al.* (2006) ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature*, **442**, 96–99.
52. Vermeulen,M., Mulder,K.W., Denissov,S., Pijnappel,W.W., van Schaik,F.M., Varier,R.A., Baltissen,M.P., Stunnenberg,H.G., Mann,M. and Timmers,H.T. (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell*, **131**, 58–69.
53. Wysocka,J., Swigut,T., Xiao,H., Milne,T.A., Kwon,S.Y., Landry,J., Kauer,M., Tackett,A.J., Chait,B.T., Badenhorst,P. *et al.* (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature*, **442**, 86–90.
54. Iwase,S., Lan,F., Bayliss,P., de la Torre-Ubieta,L., Huarte,M., Qi,H.H., Whetstine,J.R., Bonni,A., Roberts,T.M. and Shi,Y. (2007) The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell*, **128**, 1077–1088.
55. Ramon-Maiques,S., Kuo,A.J., Carney,D., Matthews,A.G., Oettinger,M.A., Gozani,O. and Yang,W. (2007) The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 18993–18998.
56. Zhao,Q., Rank,G., Tan,Y.T., Li,H., Moritz,R.L., Simpson,R.J., Cerruti,L., Curtis,D.J., Patel,D.J., Allis,C.D. *et al.* (2009) PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nat. Struct. Mol. Biol.*, **16**, 304–311.
57. Dhar,S.S., Lee,S.H., Kan,P.Y., Voigt,P., Ma,L., Shi,X., Reinberg,D. and Lee,M.G. (2012) Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev.*, **26**, 2749–2762.
58. Boyer,L.A., Langer,M.R., Crowley,K.A., Tan,S., Denu,J.M. and Peterson,C.L. (2002) Essential role for the SANT domain in the

functioning of multiple chromatin remodeling enzymes. *Mol. Cell*, **10**, 935–942.
59. Smallwood,A., Esteve,P.O., Pradhan,S. and Carey,M. (2007) Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes Dev.*, **21**, 1169–1178.
60. Micci,F., Panagopoulos,I., Bjerkehagen,B. and Heim,S. (2006) Consistent rearrangement of chromosomal band 6p21 with generation of fusion genes JAZF1/PHF1 and EPC1/PHF1 in endometrial stromal sarcoma. *Cancer Res.*, **66**, 107–112.
61. Panagopoulos,I., Micci,F., Thorsen,J., Gorunova,L., Eibak,A.M., Bjerkehagen,B., Davidson,B. and Heim,S. (2012) Novel fusion of MYST/Esa1-associated factor 6 and PHF1 in endometrial stromal sarcoma. *PLoS One*, **7**, e39354.
62. Bao,X., Zhao,S., Liu,T., Liu,Y. and Yang,X. (2013) Overexpression of PRMT5 promotes tumor cell growth and is associated with poor disease prognosis in epithelial ovarian cancer. *J. Histochem. Cytochem.*, **61**, 206–217.
63. Gu,Z., Gao,S., Zhang,F., Wang,Z., Ma,W., Davis,R.E. and Wang,Z. (2012) Protein arginine methyltransferase 5 is essential for growth of lung cancer cells. *Biochem. J.*, **446**, 235–241.
64. Han,X., Li,R., Zhang,W., Yang,X., Wheeler,C.G., Friedman,G.K., Province,P., Ding,Q., You,Z., Fathallah-Shaykh,H.M. *et al.* (2014) Expression of PRMT5 correlates with malignant grade in gliomas and plays a pivotal role in tumor growth in vitro. *J. Neurooncol.*, **118**, 61–72.
65. Canel,M., Serrels,A., Frame,M.C. and Brunton,V.G. (2013) E-cadherin-integrin crosstalk in cancer invasion and metastasis. *J. Cell Sci.*, **126**, 393–401.
66. Thiery,J.P. (2002) Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer*, **2**, 442–454.
67. Nakayama,K.I. and Nakayama,K. (2006) Ubiquitin ligases: cell-cycle control and cancer. *Nat. Rev. Cancer*, **6**, 369–381.
68. Tan,Y.M., Sangfelt,O. and Spruck,C. (2008) The Fbxw7/hCdc4 tumor suppressor in human cancer. *Cancer Lett.*, **271**, 1–12.
69. Davis,R.J., Welcker,M. and Clurman,B.E. (2014) Tumor suppression by the Fbw7 ubiquitin ligase: mechanisms and opportunities. *Cancer Cell*, **26**, 455–464.